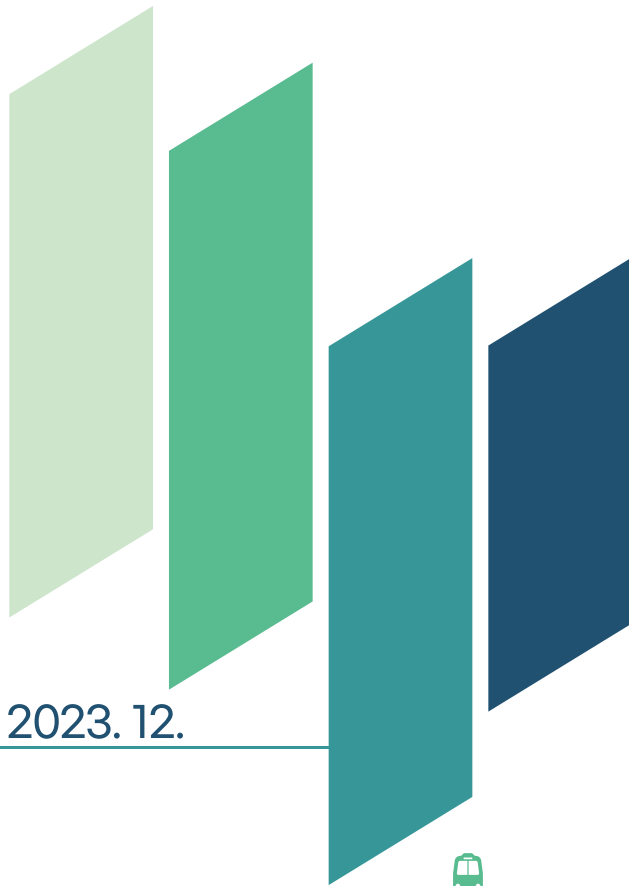


2023년 국가교통조사 및 분석

9

모바일 교통 빅데이터 구축 및 활용



국토교통부



한국교통연구원
KOREA TRANSPORT INSTITUTE

제 출 문

국토교통부 장관 귀하

본 보고서를 “2023년 국가교통조사 및 분석”의
최종보고서로 제출합니다.

2023년 12월

한국교통연구원
원장 오 재 학

본 『2023년 국가교통조사』는 다음 연구진에 의해
수행되었습니다.

참 여 연 구 진

<한국교통연구원>	
연구책임자	◦ 조종석 연구위원
연 구 진	◦ 조범철 연구위원 ◦ 김주영, 천승훈, 박용일 연구위원 ◦ 황순연, 장동익, 원민수, 이송봉, 이종우 부연구위원 ◦ 김동호, 신영권 책임전문원 ◦ 김규진 주임전문원 ◦ 가보연, 강국수, 곽명신, 김관용, 김수아, 김호용, 김 현, 박미란, 박성희, 박준호, 양태양, 오연선, 이동엽, 이새봄, 이선아, 이채영, 이해선, 홍성표 연구원 ◦ 홍연우 연구조원
<한국해양수산개발원>	
연 구 진	◦ 이호춘 부연구위원 ◦ 류희영 전문연구원 ◦ 박일란 선임사무원
<한국항공협회>	
연 구 진	◦ 성인영 책임연구원 ◦ 손병열, 최인영, 김지한, 김창욱, 김진성, 박다영 연구원

『2023년 국가교통조사』
보고서 구성 및 담당연구진

번 호	과 제 명	연 구 진
제 1권	요약보고서	조종석, 신영권, 가보연
제 2권	전국 여객 O/D 보완갱신	조종석, 강국수, 박미란, 이선아
제 3권	교통분석용 네트워크 구축	김동호, 이동엽, 이새봄
제 4권	항공여객 O/D 조사	한국항공협회
제 5권	전국화물 O/D 전수화	조범철, 황순연, 김호용, 오연선, 박준호, 김수아
제 6권	해상화물 O/D 전수화	한국해양수산개발원
제 7권	KTDB 모빌리티 기반지도 구축	이승봉, 양태양
제 8권	차량 GPS 빅데이터 구축	천승훈, 이채영
제9권	모바일통신 빅데이터 구축	원민수, 이종우, 이해선, 박성희
제10권	국가교통통계DB구축	박용일, 곽명신
제11권	특별교통대책기간 통행실태조사	김관용, 김 현
제12권	교통접근성지표 구축	장동익, 홍성표

『2023년도 국가교통조사』
과제별 공동참여·위탁용역 사업자

【공동사업 참여기관】
<ul style="list-style-type: none"> • 전국 여객O/D 현행화 공동사업 (수도권 부문) <ul style="list-style-type: none"> - 경기연구원, 인천연구원, 서울연구원 • 항공O/D 및 특성 조사 <ul style="list-style-type: none"> - (사)한국항공협회
【위탁용역 사업자】
<ul style="list-style-type: none"> • 전국여객 O/D 현행화(제주권) <ul style="list-style-type: none"> - 홍익대학교 산학협력단 • 전국여객 O/D 현행화(대구광역시권) <ul style="list-style-type: none"> - 홍익대학교 산학협력단 • 전국여객 O/D 현행화(대전세종충청권) <ul style="list-style-type: none"> - 신명이앤씨(주) • 개인통행실태 보완조사 <ul style="list-style-type: none"> - ㈜컨슈머인사이트 • 교통부문 네트워크 갱신을 위한 GIS기반 교통망 기초자료 구축 <ul style="list-style-type: none"> - 서울시립대학교 산학협력단 • T MAP·DTG 등 빅데이터를 활용한 화물자동차 운행특성 기초통계 구축 <ul style="list-style-type: none"> - ㈜노트스퀘어

【위탁용역 사업자】

- 모빌리티 빅데이터 DB구축 및 데이터 분석·활용체계 개선
 - ㈜큐빅웨어
- 모빌리티 빅데이터 DB 구축 및 데이터 분석·활용 체계 개선 감리
 - 악티보
- 모바일 통신 빅데이터 기반 기종점 통행량 검증
 - 인천대학교 산학협력단
- 객체 단위 모바일 통신 데이터 가공 및 통행 DB 구축
 - ㈜엔제로
- 특별교통통행실태조사
 - (주)컨슈머인사이트

최종보고서 목차

- 제 1권 요약보고서**
- 제 2권 전국여객 O/D 보완갱신**
- 제 3권 교통분석용 네트워크 구축**
- 제 4권 항공여객 O/D 조사**
- 제 5권 전국화물 O/D 전수화**
- 제 6권 해상화물 O/D 전수화**
- 제 7권 KTDB 모빌리티 기반지도 구축**
- 제 8권 차량 GPS 빅데이터 구축**
- 제 9권 모바일통신 빅데이터 구축**
- 제 10권 국가교통통계DB구축**
- 제 11권 특별교통대책기간 통행실태조사**
- 제 12권 교통접근성지표 구축**

● 목 차

요약

제1장 사업의 개요 3

제1절 사업의 배경 및 목적 3

1. 사업의 배경 3

2. 사업의 목적 3

제2절 사업의 범위 및 내용 4

1. 사업의 범위 4

2. 사업의 내용 5

제2장 모바일 통신데이터 기반 통행 DB 구축 및 결과 점검 9

제1절 모바일 통신데이터 기반 통행 DB 구축 개요 9

1. 모바일 통신데이터 활용목적을 고려한 통행 DB 구축 형태 검토 9

2. 통행 DB 구축 형태 검토 및 설계 11

제2절 모바일 통신데이터 기반 기종점 통행 DB 구축 12

1. 모바일 통신데이터 기반 기종점 통행 DB 형태 설계 12

2. 2022년 11월 일자별 기종점 통행 DB 구축 결과 점검 16

3. 2023년 월별 기종점 통행 DB 구축 결과 점검 32

제3절 객체단위 모바일 통신데이터 가공 및 통행 DB 구축 36

1. 개요 36

2. 객체단위 모바일 통신데이터 기초분석 38

3. 객체단위 모바일 통신데이터 오류 제거 및 전처리 48

4. 통행궤적 및 통행정보 데이터 구축 53

5. 객체단위 모바일 통신데이터 구축 및 활용 관련 검토사항 57

제3장 모바일 통신데이터 기반 통행 DB 검증 61

제1절 개요 61

1. 기본방향 61

2. 검증데이터 수집 및 가공 64

제2절 검증 결과 79

1. 데이터간 상호교차검증 결과 79

2. 실적자료와의 일치성 검증 결과 166

제3절 보정 방법 개발 169

1. 개요 169

2. 보정 가중치 산출 방법	169
3. 보정가중치 산출 결과	174
4. 통행량 보정 방법	189
5. 결론 및 향후과제	192

제4장 모바일 통신데이터 기반 교통지표 산출 및 제공 196

제1절 모바일 통신데이터 기반 교통지표 구축	196
1. 개요	196
2. 교통지표 구축을 위한 데이터 전처리	199
3. 교통지표 구축	203
제2절 온라인 서비스를 통한 교통지표 제공	225
1. 개요	225
2. 분석도구를 통한 교통지표 제공	226
3. 대시보드 및 월간리포트를 통한 교통지표 제공	243
4. 다운로드 기능을 통한 교통지표 제공	244

제5장 결론 및 차년도 수행계획 248

제1절 결론	248
제2절 차년도 수행계획	250

부록

● 표목차

〈표 2-1〉 모바일 통신 데이터 기반 개별 통행정보 DB 구축	9
〈표 2-1〉 모바일통신 빅데이터 DB 형태 (전국 일자별, 시간별 읍면동 단위 통행 정보) ...	13
〈표 2-2〉 모바일통신 빅데이터 DB 테이블 정의서 (전국 일자별, 시간별 읍면동 단위 통행 정보) 14	
〈표 2-3〉 모바일통신 빅데이터 DB형태 (전국 월별, 요일별 시군구 단위 통행 정보)	15
〈표 2-4〉 모바일통신 빅데이터 DB 테이블 정의서 (전국 월별, 요일별 시군구 단위 통행 정보) 15	
〈표 2-5〉 2022년 11월 평일 시도별 일평균 통행량	16
〈표 2-6〉 2022년 11월 평일 목적별 통행량	17
〈표 2-7〉 2022년 11월 평일 시도별 목적별 통행량	18
〈표 2-8〉 2022년 11월 평일 일자별 목적별 통행량	19
〈표 2-9〉 2022년 11월 주말 시도별 통행량	20
〈표 2-10〉 2022년 11월 주말 목적별 통행량	21
〈표 2-11〉 2022년 11월 주말 시도별 목적별 통행량	22
〈표 2-12〉 2022년 11월 주말 일자별 목적별 통행량	22
〈표 2-13〉 시간대별 통행량 패턴 분석	24
〈표 2-14〉 요일별 통행량 패턴 분석	26
〈표 2-15〉 지역별 통행 발생·도착 패턴 비교 결과	29
〈표 2-16〉 지역별 출퇴근 통행량 비교 결과	31
〈표 2-17〉 2023년 평일 시도별 목적별 통행시간	32
〈표 2-18〉 2023년 주말 시도별 목적별 통행시간	34
〈표 2-19〉 LG 유플러스 모바일 통신 원시데이터 테이블정의서	38
〈표 2-20〉 LG 유플러스 모바일 통신 격자&공간정보 데이터 테이블정의서	39
〈표 2-21〉 LG 유플러스 모바일 통신 원시 이벤트 건수	41
〈표 2-22〉 LG 유플러스 모바일 통신 원시 고유 트립ID 건수	41
〈표 2-23〉 LG 유플러스 모바일 통신 원시데이터 시도별 이벤트 건수	42
〈표 2-24〉 LG 유플러스 모바일 통신 원시데이터 시간대별 이벤트 건수	43
〈표 2-25〉 LG 유플러스 모바일 통신 원시데이터 기종점 이벤트 건수 (8일치)	44
〈표 2-26〉 LG 유플러스 모바일 통신 원시데이터 기종점 이벤트 비율 (내부통행 제외)	44
〈표 2-27〉 LG 유플러스 모바일 통신 전처리 데이터	49
〈표 2-28〉 LG 유플러스 모바일 통신 원시데이터 예시 (트립번호: 23669006)	50
〈표 2-29〉 LG 유플러스 모바일 통신 데이터 수단구분 예시 (트립번호: 23669006)	51
〈표 3-30〉 LG 유플러스 모바일 통신 원시데이터 도보, 도보 외 비율 분석	52
〈표 2-31〉 LG 유플러스 모바일 통신 원시데이터 예시 (트립번호: 23669006)	53
〈표 2-32〉 시퀀스 부여를 통한 개별 통행ID별 통행궤적 데이터 구축 (트립번호: 23669006) 53	
〈표 2-33〉 모바일 통신데이터 기반 개별 통행정보 DB 구축	54

〈표 3-1〉 KT 데이터 정의서	65
〈표 3-2〉 SKT 지라프 (GIRAF) 데이터 정의서	66
〈표 3-3〉 SKT 리트머스 (LITMUS) 데이터 정의서	67
〈표 3-4〉 기타 자료 수집 내용	70
〈표 3-5〉 자료별 공간적 정보 제공 방식	73
〈표 3-6〉 시간적 정보 통일 및 통행량 집계 1단계 예시	75
〈표 3-7〉 최종 가공 자료 (요약)	77
〈표 3-8〉 (통행발생량) 전체 - 데이터별 시도별 통행량	81
〈표 3-9〉 (통행발생량) 통근 - 데이터별 시도별 통행량 (조사자료)	83
〈표 3-10〉 (통행발생량) 통근 - 데이터별 시도별 통행량 (모바일 데이터)	84
〈표 3-11〉 (통행발생량) 통학 - 데이터별 시도별 통행량 (조사자료)	86
〈표 3-12〉 (통행발생량) 통학 - 데이터별 시도별 통행량 (모바일 데이터)	87
〈표 3-13〉 (통행발생량) 귀가 - 데이터별 시도별 통행량	89
〈표 3-14〉 (통행발생량) 기타 - 데이터별 시도별 통행량	91
〈표 3-15〉 (통행도착량) 전체 - 데이터별 시도별 통행량	93
〈표 3-16〉 (통행도착량) 통근 - 데이터별 시도별 통행량	95
〈표 3-17〉 (통행도착량) 통학 - 데이터별 시도별 통행량	97
〈표 3-18〉 (통행도착량) 귀가 - 데이터별 시도별 통행량	99
〈표 3-19〉 (통행도착량) 기타 - 데이터별 시도별 통행량	101
〈표 3-20〉 (통행분포량) 전체 - 데이터별 최다 통행지역 TOP 10 (시군구 내 통행)	158
〈표 3-21〉 (통행분포량) 전체 - 데이터별 최다 통행구간 TOP 10 (외부 통행)	159
〈표 3-22〉 읍면동별 통행량 분포 비교 결과 (2019년 가구통행실태조사 기준)	162
〈표 3-23〉 읍면동별 통행량 분포 비교 결과 (3가지 데이터 동시 비교)	162
〈표 3-24〉 읍면동별 통행량 분포 비교 결과 (모바일 통신데이터간 비교)	163
〈표 3-25〉 통근 통행에 대한 회귀모형 구축결과	164
〈표 3-26〉 통학 통행에 대한 회귀모형 구축결과	165
〈표 3-27〉 시군구별 KT·SKT·KTDB·교통카드 데이터의 통행발생량 통계검증 결과	166
〈표 3-28〉 모바일 통신데이터 기반 지역간 통행량 추정 결과와 수송실적간의 비교	168
〈표 3-29〉 통행목적별 설명변수	171
〈표 3-30〉 전체 통행 유형별 보정가중치보정기준	176
〈표 3-31〉 통근 통행 보정기준유형별 보정가중치	179
〈표 3-32〉 통학 통행 보정기준유형별 보정가중치	182
〈표 3-33〉 귀가 통행 유형별 보정가중치 보정기준	185
〈표 3-34〉 기타 통행 보정기준유형별 보정가중치	188

● 표목차

〈표 3-35〉 통행유형 분류를 위해 수집해야 하는 시군구 정보	189
〈표 3-36〉 통행유형별 변수 및 절단값	189
〈표 3-37〉 통행유형별 보정가중치 (평균, 표준편차 기준값)	191
〈표 3-38〉 통행량 보정 예시	192
〈표 4-1〉 모바일 통신데이터 기반 교통지표	197
〈표 4-2〉 격자와 매칭되지 않은 읍면동 리스트	201
〈표 4-3〉 격자 ID 변환 전, 후	201
〈표 4-4〉 모바일 데이터 전처리 표준 DB	202
〈표 4-5〉 읍면동-읍면동 단위 통행량 컬렉션의 구성	203
〈표 4-6〉 시군구-시군구 단위 통행량 컬렉션의 구성	204
〈표 4-7〉 시도-시도 단위 통행량 컬렉션의 구성	204
〈표 4-8〉 연휴기간의 통행량 컬렉션의 구성	205
〈표 4-9〉 계절별 통행량 컬렉션의 구성	206
〈표 4-10〉 계절별 기타통행량(도착량) 컬렉션의 구성	207
〈표 4-11〉 통행목적 구분 기준	208
〈표 4-12〉 통행목적별 통행량 컬렉션의 구성	208
〈표 4-13〉 고령자 통행량(발생량 기준) 컬렉션의 구성	210
〈표 4-14〉 고령자 통행량(OD통행량 기준) 컬렉션의 구성	210
〈표 4-15〉 출근 통행시간 컬렉션의 구성	211
〈표 4-16〉 퇴근 통행시간 컬렉션의 구성	212
〈표 4-17〉 내·외부 통행(시군구) 컬렉션의 구성	214
〈표 4-18〉 내·외부 통행(시도) 컬렉션의 구성	214
〈표 4-19〉 심야시간 출퇴근 통행 비율 컬렉션의 구성	216
〈표 4-20〉 월별 통근 통행량 컬렉션의 구성	217
〈표 4-21〉 도심공동화 심각도 컬렉션의 구성	218
〈표 4-22〉 경제활동의존도 컬렉션의 구성	219
〈표 4-23〉 경제활동 자체수용도 컬렉션의 구성	220
〈표 4-24〉 출·퇴근 통행연결성 컬렉션의 구성	222
〈표 4-25〉 초과통근량 컬렉션의 구성	224
〈부표 1〉 KT 모바일 통신데이터 기반의 통행 DB 구조	255
〈부표 2〉 지역별 통근량 통근거리 산출 예시	255
〈부표 3〉 지역별 통근거리 산출 예시	255
〈부표 4〉 초과통근량 기준 상위 10%, 하위 10% 지역	258
〈부표 5〉 시군구별 직주균형지표 결과 (전체)	258

〈그림 2-1〉 2022년 11월 평일 통행량 중 시도별 통행량 비율	16
〈그림 2-2〉 2022년 11월 평일 시도별 총 통행량 중 목적별 통행량 비율	17
〈그림 2-3〉 2022년 11월 주말 총 통행량 중 시도별 통행량 비율	20
〈그림 2-4〉 2022년 11월 주말 시도별 총 통행량 중 목적별 통행량 비율	21
〈그림 2-5〉 시간대별 지역별 통행 발생량 패턴	23
〈그림 2-6〉 시간대별 지역별 통행 도착량 패턴	23
〈그림 2-7〉 시간대별 요일별 발생량 패턴	25
〈그림 2-8〉 시간대별 요일별 도착량 패턴	25
〈그림 2-9〉 시간대별 출·퇴근 통근 통행량 패턴 (요일별)	26
〈그림 2-10〉 시간대별 출·퇴근 통근 통행량 패턴 (연령별)	27
〈그림 2-11〉 시간대별 출·퇴근 통근 통행량 패턴 (지역별)	27
〈그림 2-12〉 지역별 통행 발생량 비교 결과 (발생)	28
〈그림 2-13〉 지역별 통행 발생량 비교 결과 (도착)	28
〈그림 2-14〉 지역별 출근 통행량 비교 결과	30
〈그림 2-15〉 지역별 퇴근 통행량 비교 결과	30
〈그림 2-16〉 시도별 평일 출근 통행시간	33
〈그림 2-17〉 시도별 평일 등교 통행시간	33
〈그림 2-18〉 시도별 평일 귀가 통행시간	33
〈그림 2-19〉 시도별 평일 기타 통행시간	34
〈그림 2-20〉 시도별 주말 귀가 통행시간	35
〈그림 2-21〉 시도별 주말 기타 통행시간	35
〈그림 2-22〉 LG 통신 데이터 격자 전국 예시	40
〈그림 2-23〉 강남구 50m격자 단위 통행량 시각화	45
〈그림 2-24〉 제주시 50m격자 단위 통행량 시각화	45
〈그림 2-25〉 울릉군 50m격자 단위 통행량 시각화	46
〈그림 2-26〉 해운대구 50m격자 단위 통행량 시각화	46
〈그림 2-27〉 인천공항 50m격자 단위 통행량 시각화	47
〈그림 2-28〉 종로구 50m격자 단위 통행량 시각화	47
〈그림 2-29〉 트립 넘버 23669006 이동궤적 예시	51
〈그림 2-30〉 도보 이동속도 프로파일	55
〈그림 2-31〉 도보 이동거리 프로파일	55
〈그림 2-32〉 도보 외 이동속도 프로파일	56
〈그림 2-33〉 도보 외 이동거리 프로파일	56
〈그림 3-1〉 상호 검증 방식 예시 - 조사자료 및 통신사별 비교	62
〈그림 3-2〉 데이터 신뢰성 검증 및 보정 절차	63

● 그림목차

〈그림 3-3〉 (통행발생량) 전체 - 시도별 통행 비율	79
〈그림 3-4〉 (통행발생량) 통근 - 시도별 통행 비율	82
〈그림 3-5〉 (통행발생량) 통학 - 시도별 통행 비율	85
〈그림 3-6〉 (통행발생량) 귀가 - 시도별 통행 비율	88
〈그림 3-7〉 (통행발생량) 기타 - 시도별 통행 비율	90
〈그림 3-8〉 (통행도착량) 전체 - 시도별 통행 비율	92
〈그림 3-9〉 (통행도착량) 통근 - 시도별 통행 비율	94
〈그림 3-10〉 (통행도착량) 통학 - 시도별 통행 비율	96
〈그림 3-11〉 (통행도착량) 귀가 - 시도별 통행 비율	98
〈그림 3-12〉 (통행도착량) 기타 - 시도별 통행 비율	100
〈그림 3-13〉 (통행발생량) 전체 - 데이터별 통행량 밀도 비교 (1)	102
〈그림 3-14〉 (통행발생량) 전체 - 데이터별 통행량 밀도 비교 (2)	103
〈그림 3-15〉 (통행발생량) 전체 - 데이터간 상관관계 비교	103
〈그림 3-16〉 (통행발생량) 전체 통행 - 인구수에 따른 데이터별 통행량 비교	104
〈그림 3-17〉 (통행발생량) 전체 통행 - 인구수와 각 데이터별 상관관계	105
〈그림 3-18〉 (통행발생량) 전체 통행 - 인구밀도에 따른 데이터별 통행량 비교	106
〈그림 3-19〉 (통행발생량) 전체 통행 - 인구밀도와 각 데이터별 상관관계	107
〈그림 3-20〉 (통행발생량) 통근 통행 - 데이터별 통행량 빈도 비교 (1)	108
〈그림 3-21〉 (통행발생량) 통근 통행 - 데이터별 통행량 빈도 비교 (2)	108
〈그림 3-22〉 (통행발생량) 통근 통행 - 데이터간 상관관계 비교	109
〈그림 3-23〉 (통행발생량) 통근 통행 - 취업자수에 따른 데이터별 통행량 비교	110
〈그림 3-24〉 (통행발생량) 통근 통행 - 취업자수와 각 데이터별 상관관계	111
〈그림 3-25〉 (통행발생량) 통학 통행 - 데이터별 통행량 빈도 비교 (1)	112
〈그림 3-26〉 (통행발생량) 통학 통행 - 데이터별 통행량 빈도 비교 (2)	112
〈그림 3-27〉 (통행발생량) 통학 통행 - 데이터간 상관관계 비교	113
〈그림 3-28〉 (통행발생량) 통학 통행 - 학령인구수에 따른 데이터별 통행량 비교	114
〈그림 3-29〉 (통행발생량) 통학 통행 - 학령인구수와 각 데이터별 상관관계	115
〈그림 3-30〉 (통행발생량) 귀가 통행 - 데이터별 통행량 빈도 비교 (1)	116
〈그림 3-31〉 (통행발생량) 귀가 통행 - 데이터별 통행량 빈도 비교 (2)	116
〈그림 3-32〉 (통행발생량) 귀가 통행 - 데이터간 상관관계 비교	117
〈그림 3-33〉 (통행발생량) 귀가 통행 - 인구수에 따른 데이터별 통행량 비교	118
〈그림 3-34〉 (통행발생량) 귀가 통행 - 인구수와 각 데이터별 상관관계	119
〈그림 3-35〉 (통행발생량) 귀가 통행 - 인구밀도에 따른 데이터별 통행량 비교	120
〈그림 3-36〉 (통행발생량) 귀가 통행 - 인구밀도와 각 데이터별 상관관계	121
〈그림 3-37〉 (통행발생량) 기타 통행 - 데이터별 통행량 빈도 비교 (1)	122

〈그림 3-38〉 (통행발생량) 기타 통행 - 데이터별 통행량 빈도 비교 (2)	122
〈그림 3-39〉 (통행발생량) 기타 통행 - 데이터간 상관관계 비교	123
〈그림 3-40〉 (통행발생량) 기타 통행 - 인구수에 따른 데이터별 통행량 비교	124
〈그림 3-41〉 (통행발생량) 기타 통행 - 인구수와 각 데이터별 상관관계	125
〈그림 3-42〉 (통행발생량) 기타 통행 - 인구밀도에 따른 데이터별 통행량 비교	126
〈그림 3-43〉 (통행발생량) 기타 통행 - 인구밀도와 각 데이터별 상관관계	127
〈그림 3-44〉 (통행도착량) 전체 통행 - 데이터별 통행량 빈도 비교 (1)	128
〈그림 3-45〉 (통행도착량) 전체 통행 - 데이터별 통행량 빈도 비교 (2)	128
〈그림 3-46〉 (통행도착량) 전체 통행 - 데이터간 상관관계 비교	129
〈그림 3-47〉 (통행도착량) 전체 통행 - 인구수에 따른 데이터별 통행량 비교	130
〈그림 3-48〉 (통행도착량) 전체 통행 - 인구수와 각 데이터별 상관관계	131
〈그림 3-49〉 (통행도착량) 전체 통행 - 인구밀도에 따른 데이터별 통행량 비교	132
〈그림 3-50〉 (통행도착량) 전체 통행 - 인구밀도와 각 데이터별 상관관계	133
〈그림 3-51〉 (통행도착량) 통근 통행 - 데이터별 통행량 빈도 비교 (1)	134
〈그림 3-52〉 (통행도착량) 통근 통행 - 데이터별 통행량 빈도 비교 (2)	134
〈그림 3-53〉 (통행도착량) 통근 통행 - 데이터간 상관관계 비교	135
〈그림 3-54〉 (통행도착량) 통근 통행 - 종사자수에 따른 데이터별 통행량 비교	136
〈그림 3-55〉 (통행도착량) 통근 통행 - 종사자수와 각 데이터별 상관관계	137
〈그림 3-56〉 (통행도착량) 통학 통행 - 데이터별 통행량 빈도 비교 (1)	138
〈그림 3-57〉 (통행도착량) 통학 통행 - 데이터별 통행량 빈도 비교 (2)	138
〈그림 3-58〉 (통행도착량) 통학 통행 - 데이터간 상관관계 비교	139
〈그림 3-59〉 (통행도착량) 통학 통행 - 수용학생수에 따른 데이터별 통행량 비교	140
〈그림 3-60〉 (통행도착량) 통학 통행 - 수용학생수와 각 데이터별 상관관계	141
〈그림 3-61〉 (통행도착량) 통학 통행 - 학교수에 따른 데이터별 통행량 비교	142
〈그림 3-62〉 (통행도착량) 통학 통행 - 학교수와 각 데이터별 상관관계	143
〈그림 3-63〉 (통행도착량) 귀가 통행 - 데이터별 통행량 빈도 비교 (1)	144
〈그림 3-64〉 (통행도착량) 귀가 통행 - 데이터별 통행량 빈도 비교 (2)	144
〈그림 3-65〉 (통행도착량) 귀가 통행 - 데이터간 상관관계 비교	145
〈그림 3-66〉 (통행도착량) 귀가 통행 - 인구수에 따른 데이터별 통행량 비교	146
〈그림 3-67〉 (통행도착량) 귀가 통행 - 인구수와 각 데이터별 상관관계	147
〈그림 3-68〉 (통행도착량) 귀가 통행 - 인구밀도에 따른 데이터별 통행량 비교	148
〈그림 3-69〉 (통행도착량) 귀가 통행 - 인구밀도와 각 데이터별 상관관계	149
〈그림 3-70〉 (통행도착량) 기타 통행 - 데이터별 통행량 빈도 비교 (1)	150
〈그림 3-71〉 (통행도착량) 기타 통행 - 데이터별 통행량 빈도 비교 (2)	150
〈그림 3-72〉 (통행도착량) 기타 통행 - 데이터간 상관관계 비교	151

● 그림목차

〈그림 3-73〉 (통행도착량) 기타 통행 - 인구수에 따른 데이터별 통행량 비교	152
〈그림 3-74〉 (통행도착량) 기타 통행 - 인구수와 각 데이터별 상관관계	153
〈그림 3-75〉 (통행도착량) 기타 통행 - 인구밀도에 따른 데이터별 통행량 비교	154
〈그림 3-76〉 (통행도착량) 기타 통행 - 인구밀도와 각 데이터별 상관관계	155
〈그림 3-77〉 데이터별 최다 통행구간 TOP 10 (시군구 내부 통행)	156
〈그림 3-77〉 데이터별 최다 통행구간 TOP 10 (시군구 내부 통행)	156
〈그림 3-78〉 데이터별 최다 통행구간 TOP 10 (외부 통행)	157
〈그림 3-78〉 데이터별 최다 통행구간 TOP 10 (외부 통행)	157
〈그림 3-79〉 통행량이 가장 많은 구간 Top 1000	160
〈그림 3-80〉 시군구별 KT·SKT·KTDB·교통카드 데이터의 통행발생량 비교	166
〈그림 3-81〉 시간대별 KT·SKT 교통카드의 통행발생량 비교	167
〈그림 3-82〉 시간대별 KT·SKT 교통카드의 통행도착량 비교	167
〈그림 3-83〉 의사결정나무구조	170
〈그림 3-83〉 통행목적별 보정기준 산출 과정	171
〈그림 3-84〉 통행목적별 보정기준 산출 과정	173
〈그림 3-85〉 (전체 통행) 설명변수 선정 기준	174
〈그림 3-86〉 (전체 통행) 의사결정나무구조	175
〈그림 3-87〉 (통근 통행) 설명변수 선정 기준	177
〈그림 3-88〉 (통근 통행) 의사결정나무구조	178
〈그림 3-89〉 (통학 통행) 설명변수 선정 기준	180
〈그림 3-90〉 (통학 통행) 의사결정나무구조	181
〈그림 3-91〉 (귀가 통행) 설명변수 선정 기준	183
〈그림 3-92〉 (귀가 통행) 의사결정나무구조	184
〈그림 3-93〉 (기타 통행) 설명변수 선정 기준	186
〈그림 3-94〉 (기타 통행) 의사결정나무구조	187
〈그림 4-1〉 모바일 데이터 오류 유형 예시	199
〈그림 4-2〉 격자-행정구역 매칭 방법 (면적 비율 적용 예시)	200
〈그림 4-3〉 매칭 제외 대상 (해상에 위치하는 격자 예시)	200
〈그림 4-4〉 기종점 인구 통행량 분석도구 화면	226
〈그림 4-5〉 핫플레이스 분석도구 화면	227
〈그림 4-6〉 기종점 인구 통행구간 분석도구 화면	228
〈그림 4-7〉 연휴기간 통행 분석도구 화면	229
〈그림 4-8〉 계절별 핫플레이스 분석도구 화면	230
〈그림 4-9〉 관광특화지역 분석도구 화면	230
〈그림 4-10〉 출·퇴근 통행량 분석도구 화면	231

〈그림 4-11〉 고령자 주요통행구간 분석도구 화면	232
〈그림 4-12〉 통행시간·거리 분석도구 화면	233
〈그림 4-13〉 출·퇴근 통행시간 분석도구 화면	234
〈그림 4-14〉 내·외부 통행 비율 분석도구 화면	235
〈그림 4-15〉 고령자 통행 비중 분석도구 화면	236
〈그림 4-16〉 근무형태에 따른 통근통행 분석도구 화면	237
〈그림 4-17〉 고용인구 변화에 따른 통근통행 분석도구 화면	238
〈그림 4-18〉 도심공동화 심각도 분석도구 화면	239
〈그림 4-19〉 경제활동의존도 분석도구 화면	240
〈그림 4-20〉 경제활동 자체수용도 분석도구 화면	241
〈그림 4-21〉 출·퇴근 통행연결성 분석도구 화면	242
〈그림 4-22〉 직주 근접 분석도구 화면	242
〈그림 4-23〉 대시보드 화면	243
〈그림 4-24〉 모빌리티 월간리포트 화면	243
〈그림 4-25〉 다운로드 화면	244
〈부도 1〉 전국 시군구별 초과통근 정도	258



의약

요 약

1. 사업의 개요

가. 사업의 배경 및 목적

- 모바일통신 빅데이터는 교통 분야에서 활용 중인 기존 빅데이터보다 표본율이 높고, 개별 통행정보가 상세하여 활용 가치가 높음
 - 전 국민의 95% 이상이 모바일 기기를 사용하고 있고, 기기의 전원을 끄지 않는 이상 개인의 이동패적이 실시간으로 기록되어 데이터의 시·공간적 해상도가 높으며, 인적 정보가 담겨있어 사회 현상 분석에도 용이함
- 또한, 기존 인력식 조사 방식에 비해 경제적이고 효율적으로 통행정보를 취득할 수 있으며, 수집·분석 시간이 단축되어 시의성 있는 통행정보를 제공할 수 있음
- 모바일 기기를 통해 수집되는 모바일통신 빅데이터를 이용한 통행사슬(Trip Chain) DB를 구축하여 여객 및 화물OD 구축을 지원하고, 각종 교통관련 지표를 개발하여 교통정책 수립을 지원하고자 함

나. 사업의 범위 및 내용

1) 모바일 통신데이터 기반 통행 DB 구축 및 결과 점검

- 시간적 범위: SKT 리트머스 데이터 2022년 11월 (1개월), 2023년(1년)
- 공간적 범위: 전국

○ 내용적 범위

- 본 사업의 활용 목적, 통신사의 데이터 특성에 따라 통행 DB의 구조를 설계하고 정의된 로직에 따라 2023년 통행 DB를 생성함
- 통행정보(체류지 유형, 통행량, 통행시간, 통행속도, 통행거리 등) 검증 기준 정립하고 정립된 기준에 따라 통행DB 구축 결과를 점검함

2) 모바일 통신데이터 기반 통행 DB 검증

○ 시간적 범위

- 모바일 통신데이터: KT 데이터, SKT 리트머스 데이터, SKT 지라프 데이터 2022년 11월 (1개월)
- 조사자료: 2019년 가구통행실태조사, 2021년 개인통행실태조사, 2020년 인구총조사
- 기타자료(각종 사회경제지표 등): 모바일 통신데이터에 대한 해석이 용이하도록 2022년 자료로 수집함

○ 공간적 범위: 전국 (시도, 시군구, 읍면동 단위)

○ 내용적 범위

- (데이터간 상호 비교를 통한 검증) 데이터 6종(모바일 통신데이터 3종, 조사자료 3종)의 통행량, 통행분포를 상호 비교함
 - 시도, 시군구, 읍면동 단위로 전체, 통행목적별 발생량, 도착량을 각각 비교함
 - 시군구 단위로 최다 통행구간을 비교함
- (통행 발생·도착 요인을 기준으로 검증) 통행 발생·도착 요인(인구수, 수용학생수, 학령인구수 등)의 분포와 수집한 데이터 6종(모바일 통신데이터 3종, 조사자료 3종)의 통행량 분포가 유사한지 확인함
 - 시군구, 읍면동 단위로 전체, 통행목적별로 발생량, 도착량을 각각 비교함
- (실적자료를 기준으로 검증) 실적자료(대중교통카드자료, 수송실적 자료 등)의 통행량과 모바일 통신데이터의 통행량을 비교 분석함
- (모바일 통신데이터 보정 방법 개발) 검증 결과를 기반으로 모바일 통신데이터와

타 데이터의 통행량 차이를 보정하는 방법을 제시함

- 타 데이터와 통행량 차이가 발생하는 구간에 대한 특성 및 주요 요인을 분석함
- 모바일 통신데이터의 통행량 보정 대상(유형) 및 기준을 제시함
- 보정 통행량 산출 방법 및 적용 방법을 제시함

3) 모바일 통신데이터 기반 교통지표 산출 및 제공

- 시간적 범위: 2022년 1월~12월 (1년)
- 공간적 범위: 전국
- 내용적 범위
 - 문헌고찰을 통해 모바일 통신데이터에서 산출가능한 교통지표를 발굴함
 - 2022년 사업에서 구축한 KT 통신데이터 기반 통행 DB(2022년, 1년치)를 활용하여 교통지표 DB를 구축함
 - 교통지표를 확인할 수 있는 각종 분석도구, 대시보드, 분석 리포트를 개발함

2. 모바일 통신데이터 기반 통행 DB 구축 및 결과 점검

가. 개요

- 본 사업 내 모바일 통신데이터 구축 및 활용 목적과 현재 통신사에서 구매 및 확보 가능한 모바일 통신데이터의 공간적·시간적 범위 및 상세도를 고려하여 통행 DB 구축 형태를 설계함
 - 장기적으로는 개인속성정보 및 개인 통행궤적을 모두 포함하여 다양한 목적으로 활용 가능한 원천 통행 DB를 확보하여 구축할 필요가 있음
 - 그러나 현재 모바일통신 원천 통행 DB 확보는 현실적으로 어려운 상황으로, 본 사업에서는 활용 목적에 따라 개인 O/D 데이터와 개인별 통행궤적을 포함하는 데이터로 구분하여 확보 및 구축을 추진함
- 기종점 통행 DB는 여객 O/D 추정 및 지역간·지역내 통행분포 분석을 목적으로 구축함
 - 전국 읍면동 단위의 보다 상세한 시간적 단위(일자별·시간대별)의 통행정보 분석, 월별 통행 변화 분석 및 평일·주말 O/D 분석을 위한 DB를 구축함
 - 2022년 11월에 대한 일자별 읍면동 단위 기종점 통행 DB, 2023년에 대한 월별 시군구 단위 기종점 통행 DB를 구축 및 검토함
- 개인별 통행궤적 DB는 현재 확보 가능한 개별 객체단위 모바일 통신 빅데이터에 대한 전처리 및 가공, 통행경로 및 통행지표 DB 구축 등에 대한 기술적인 검토를 수행함
 - 객체 단위 모바일 통신데이터 오류 검토 및 필터링, 개별 통행경로 DB 구축, 통행량·통행속도 산출, 도보 통행 분석 및 통행지표 구축 관련 기술적 검토를 수행함
 - 2023년 10월 중 8일에 대한 개별 통행별 통행 궤적 데이터를 활용함

나. 기종점 통행 DB 구축

1) DB 설계

- 기종점 통행 DB는 여객 O/D 구축을 고려하여 기점 정보, 도착 정보, 통행자 정보를 중심으로 구축하며, 기종점간 통행량, 통행시간, 통행거리를 산출하여 구축함
- 「개인정보보호법」과 「위치정보의 보호 및 이용 등에 관한 법률」을 고려하여 개인이 식별되지 않도록 DB를 구축함
 - 출발지 및 도착지는 행정동 또는 시군구 단위로 기점 정보·도착 정보·통행자 정보를 기준으로 통행 정보가 같은 인구를 집계함
- 출발 시간 및 도착 시간 정보는 집계된 인구가 5명 이하가 될 가능성을 줄이면서, DB 사용자에게 가능한 섬세한 통행 정보를 제공할 수 있도록 한 시간 단위로 입력하도록 함
- 통행자에 대한 정보를 확인할 수 있도록 성, 연령 필드를 구성하고, 통행 정보가 지나치게 세분되지 않도록 가능한 통행 특성이 유사한 계층을 묶어 10세 단위로 연령을 구분하여 입력하도록 함
- 출발 및 도착 체류지 유형은 평소 통행자가 해당 위치에 얼마나 자주, 오래 머무르는지 추정할 수 있는 정보를 제공하기 위한 것으로, 「2019년 국가교통조사」에서 설정한 체류지 식별 기준에 따라 출발지(출발 기지국 좌표)와 도착지(도착 기지국 좌표)의 특성을 확인한 후 각각 체류지 유형을 구분하여 입력하도록 함
 - 체류지 유형을 집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 관광(5), 기타(4)로 구분하여 구축함
- 통행량 필드에는 전술한 기준에 의해 집계된 인구(단말기 수)를 입력함
- 통행시간 필드에는 집계된 통행 정보를 기준으로 산출한 평균 통행시간 값을 입력함

○ 전국에 대한 일자별 읍면동 단위 기종점 통행 DB는 아래와 같이 설계함

〈표 1〉 모바일통신 빅데이터 DB 형태 (전국 일자별, 시간별 읍면동 단위 통행 정보)

출발 정보							
일자	요일	시간대	읍면동	시군구	시도	체류지 유형*	
221101	2	10	11010530	11010	11	H	
도착 정보							
일자	요일	시간대	읍면동	시군구	시도	체류지 유형*	
221101	2	11	31091710	31091	3	0	
통행자 정보		통행 정보***					
성별	연령대**	총 통행량	통행시간			통행 거리 (평균)	통행 속도 (평균)
			평균	분산	중앙값		
M	30	6	4908	981.6	5040	44000	8.96

주: * 집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)

** 10세 간격

*** 시간은 sec 단위 / 거리는 m 단위 / 속도는 m/sec로 제공

〈표 2〉 모바일통신 빅데이터 DB 테이블 정의서 (전국 일자별, 시간별 읍면동 단위 통행 정보)

컬럼ID	컬럼명	비고
o_dt	출발 날짜	-
o_dow	출발 요일	1~7(월~일)
o_time	출발 시간대	0~23
o_adm	출발지 행정동	-
o_sigun	출발지 시군구	-
o_sido	출발지 시도	-
o_type	출발지 유형	집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)
d_dt	도착 날짜	-
d_time	도착 시간대	0~23
d_adm	도착지 행정동	-
d_sigun	도착지 시군구	-
d_sido	도착지 시도	-
d_type	도착지 유형	집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)
age	연령대	10세 간격
gender	성별	01:남자, 02:여자
tot_pop	총 통행량	-
move_time_avg	통행시간_평균	단위(sec)
move_time_var	통행시간_분산	-
move_time_mid	통행시간_중앙값	-
distance_avg	통행거리_평균	단위(m)
dspeed_avg	통행속도_평균	m/sec

○ 전국에 대한 월별 시군구 단위 기종점 통행 DB는 아래와 같이 설계함

〈표 3〉 모바일통신 빅데이터 DB 형태 (전국 월별, 요일별 시군구 단위 통행 정보)

출발 정보				
요일	시간대	시군구	시도	체류지 유형*
2	10	11010	11	H
도착 정보				
요일	시간대	시군구	시도	체류지 유형
2	11	31091	31	O
통행자 정보			통행 정보	
성별	연령대**		평균 통행량	
M	30		6	

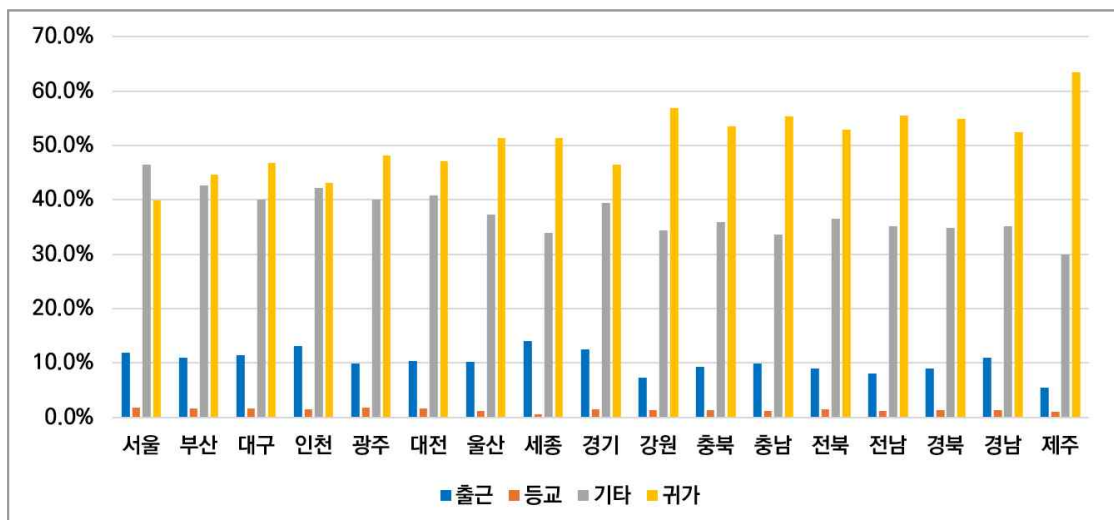
주: * 집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)
** 10세 간격

〈표 4〉 모바일통신 빅데이터 DB 테이블정의서 (전국 월별, 요일별 시군구 단위 통행 정보)

컬럼ID	컬럼명	비고
o_dow	출발요일	1~7(월~일)
o_time	출발시간대	0~23
o_sigun	출발지시군구	-
o_sido	출발지시도	-
o_type	출발지유형	집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)
d_time	도착시간대	0~23
d_sigun	도착지시군구	-
d_sido	도착지시도	-
d_type	도착지유형	집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)
age	연령대	10세 간격
gender	성별	01:남자, 02:여자
avg_pop	평균통행량	-

2) 2022년 11월 일자별 기종점 통행 DB 구축 결과 점검

○ 평일에 대한 시도별 목적별 통행량 분석 결과는 아래와 같음



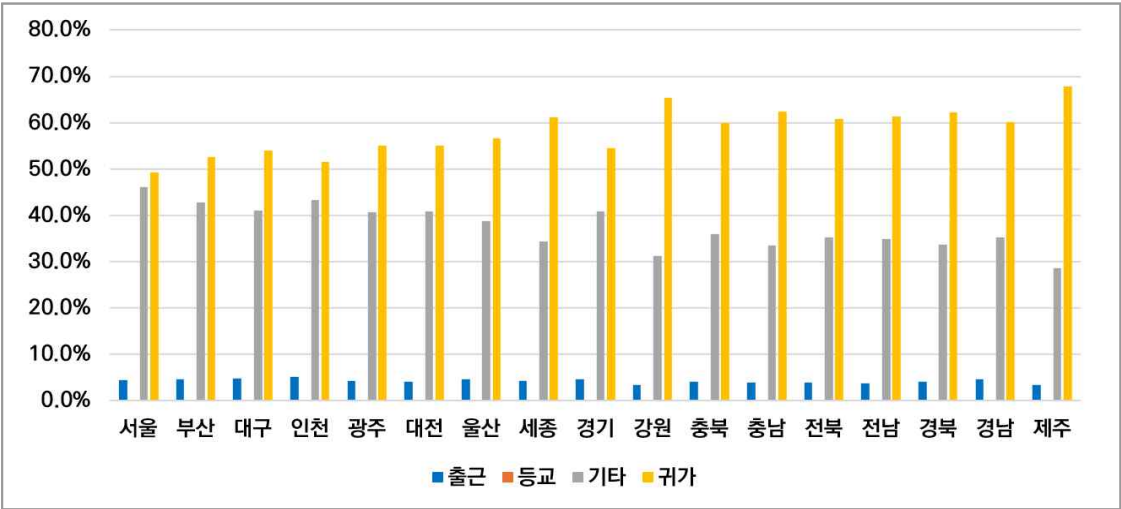
〈그림 1〉 2022년 11월 평일 시도별 총 통행량 중 목적별 통행량 비율

〈표 5〉 2022년 11월 평일 시도별 목적별 통행량

(단위: 통행/일)

시도	출근	등교	귀가	기타	합계
11 서울	2,065,946	320,182	8,074,332	6,927,078	17,387,538
26 부산	630,485	92,021	2,438,981	2,551,322	5,712,809
27 대구	463,065	69,185	1,630,808	1,907,320	4,070,377
28 인천	571,891	67,175	1,839,472	1,881,801	4,360,340
29 광주	261,095	47,465	1,057,907	1,271,545	2,638,012
30 대전	247,399	39,583	966,803	1,118,863	2,372,647
31 울산	193,562	21,794	700,823	965,669	1,881,847
36 세종	53,944	2,140	130,461	197,502	384,047
41 경기	2,317,792	272,741	7,292,170	8,590,679	18,473,382
42 강원	155,680	29,678	736,728	1,219,975	2,142,061
43 충북	208,043	29,272	796,429	1,188,058	2,221,802
44 충남	277,905	34,859	943,498	1,553,665	2,809,927
45 전북	246,810	43,236	1,010,352	1,461,630	2,762,028
46 전남	195,519	30,178	846,299	1,336,764	2,408,759
47 경북	330,612	49,231	1,289,206	2,027,550	3,696,598
48 경남	536,393	65,971	1,707,316	2,546,461	4,856,141
50 제주	70,676	14,302	387,444	821,616	1,294,038
합계	8,826,818	1,229,014	31,849,026	37,567,496	79,472,354

○ 주말에 대한 시도별 목적별 통행량 분석 결과는 아래와 같음



〈그림 2〉 2022년 11월 주말 시도별 총 통행량 중 목적별 통행량 비율

〈표 6〉 2022년 11월 주말 시도별 목적별 통행량

(단위: 통행/일)

시도		출근	등교	귀가	기타	총합계
11	서울	557,648	12,450	5,691,742	6,066,466	12,328,306
26	부산	206,824	3,000	1,942,323	2,385,812	4,537,958
27	대구	156,900	1,956	1,326,151	1,742,354	3,227,361
28	인천	172,010	1,239	1,459,854	1,739,410	3,372,513
29	광주	87,925	1,379	849,103	1,147,777	2,086,184
30	대전	78,480	1,361	770,713	1,038,508	1,889,061
31	울산	70,197	572	590,739	863,472	1,524,980
36	세종	12,681	38	100,329	178,193	291,242
41	경기	652,526	6,166	5,874,379	7,809,385	14,342,455
42	강원	69,156	650	636,588	1,331,845	2,038,239
43	충북	72,791	857	651,509	1,082,302	1,807,458
44	충남	93,297	705	784,455	1,460,906	2,339,363
45	전북	91,290	1,345	832,785	1,436,951	2,362,370
46	전남	79,261	472	743,602	1,309,529	2,132,863
47	경북	130,611	1,385	1,085,170	2,004,856	3,222,023
48	경남	183,694	1,614	1,419,791	2,423,910	4,029,008
50	제주	42,224	356	357,199	844,683	1,244,461
합계		2,757,515	35,543	25,116,430	34,866,357	62,775,845

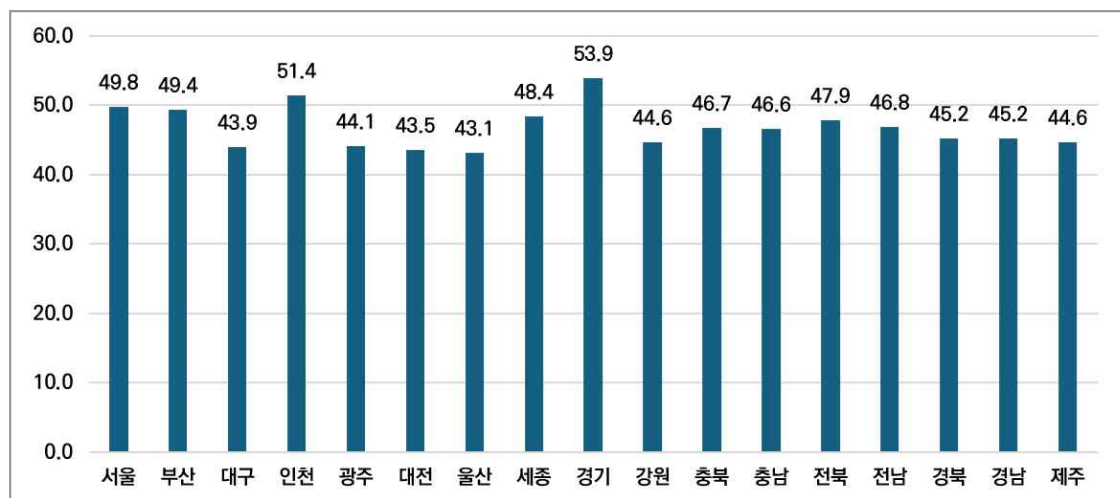
3) 2023년 월별 기종점 통행 DB 구축 결과 점검

- 평일 시도별·목적별 통행시간 분석 결과, 출근 목적 통행의 경우 수도권의 통행 시간이 타 시도 대비 길게 소요되는 것으로 나타난 반면, 기타 목적 통행의 경우 타 목적 통행에 비해 통행시간이 오래 소요되며 도 지역의 통행시간이 긴 것으로 분석됨

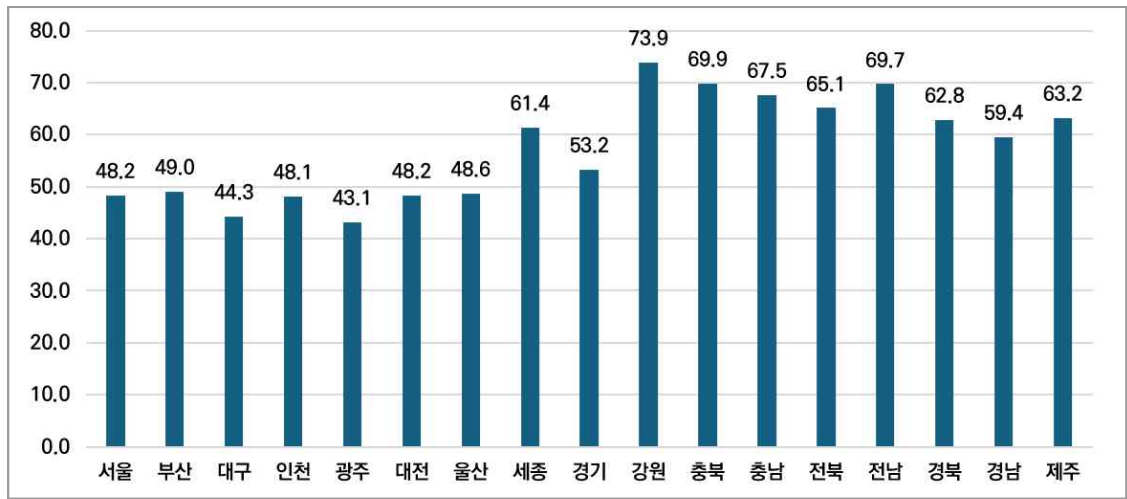
〈표 7〉 2023년 평일 시도별 목적별 통행시간

(단위: 분)

시도	출근	등교	귀가	기타
서울	49.8	45.9	48.2	69.6
부산	49.4	45.2	49.0	70.3
대구	43.9	43.3	44.3	68.6
인천	51.4	48.9	48.1	73.9
광주	44.1	43.6	43.1	69.5
대전	43.5	43.6	48.2	71.8
울산	43.1	45.4	48.6	71.6
세종	48.4	59.0	61.4	89.5
경기	53.9	51.3	53.2	78.4
강원	44.6	39.3	73.9	82.9
충북	46.7	43.6	69.9	84.3
충남	46.6	43.5	67.5	81.4
전북	47.9	42.0	65.1	80.4
전남	46.8	41.3	69.7	82.9
경북	45.2	42.2	62.8	78.9
경남	45.2	44.1	59.4	75.6
제주	44.6	45.7	63.2	70.4



〈그림 3〉 시도별 평일 출근 통행시간



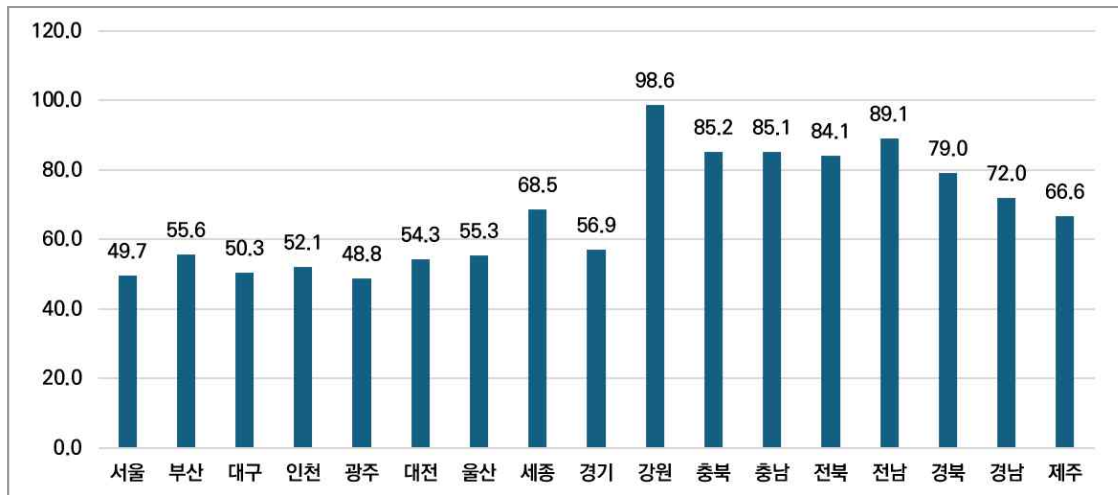
〈그림 4〉 시도별 평일 귀가 통행시간

- 주말 시도별·목적별 통행시간 분석 결과, 귀가 및 기타 목적 통행의 경우 서울 및 광역시 대비 도 지역의 통행시간이 긴 것으로 나타남

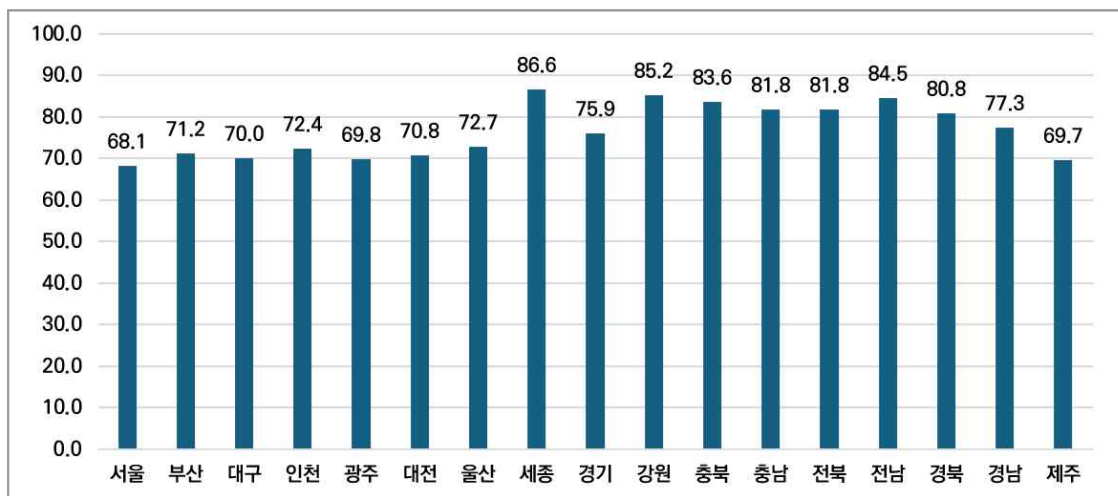
〈표 8〉 2023년 주말 시도별 목적별 통행시간

(단위: 분)

시도	출근	등교	귀가	기타
서울	50.6	46.2	49.7	68.1
부산	50.9	45.6	55.6	71.2
대구	45.6	43.9	50.3	70.0
인천	52.4	49.3	52.1	72.4
광주	45.8	44.5	48.8	69.8
대전	45.0	44.2	54.3	70.8
울산	44.8	46.2	55.3	72.7
세종	50.2	58.9	68.5	86.6
경기	54.9	51.6	56.9	75.9
강원	46.1	40.2	98.6	85.2
충북	48.2	44.3	85.2	83.6
충남	48.2	44.4	85.1	81.8
전북	49.5	43.1	84.1	81.8
전남	48.6	42.6	89.1	84.5
경북	46.9	43.2	79.0	80.8
경남	47.0	45.1	72.0	77.3
제주	45.3	47.8	66.6	69.7



〈그림 5〉 시도별 주말 귀가 통행시간



〈그림 6〉 시도별 주말 기타 통행시간

다. 객체단위 통행 DB 구축

1) 개요

- 모바일 통신데이터는 휴대전화 단말기를 소지한 개인별 속성 및 통행궤적에 대한 상세한 정보가 수집되어 교통 및 관련 분야에서 가치가 높음에도 불구하고 집계 수준에서 활용하여 개인의 통행에 관한 정보 활용에 한계가 존재함
- 이에, 모바일 원시데이터 기반 개별 객체의 이동에 대한 상세 정보를 포함하는 원천 통행DB 구축 확대를 통해 이용자가 필요로 하는 신규 데이터 재생산 체계 마련이 필요함
- 원천 통행DB 기반 개별 객체의 이동 행태 및 특성을 고려하여 기존 성능평가지표 구축 방법 고도화 및 신규 성능평가지표 발굴·구축이 필요하며, 궁극적으로 신규 모빌리티 서비스 분석과 모빌리티 영향 평가 및 의사결정 지원을 위한 데이터 기반 교통정책 지원 체계 구축이 필요함
- 이에 본 사업에서는 현재 확보 가능한 개별 객체단위 모바일 통신 빅데이터에 대한 전처리 및 가공, 통행경로 및 통행지표 DB 구축 등에 대한 기술적인 검토를 수행함

2) 객체단위 데이터 오류 제거 및 전처리

- 객체 단위 모바일 통신 데이터의 오류를 제거하고 대용량 데이터의 처리를 위해 다음과 같은 과정을 수행함
 - 파싱 오류, 기준일자·시간 오류, 참조 오류, 이동거리 및 이동시간 이상치를 제거함
 - 데이터의 효율적인 전처리를 위해 일자 분할, 시간 정보 변환, 행정동 정보 추가, 통행목적 컬럼 제외 등 전처리 과정을 수행함
- 최종 전처리된 데이터의 형태는 다음과 같음

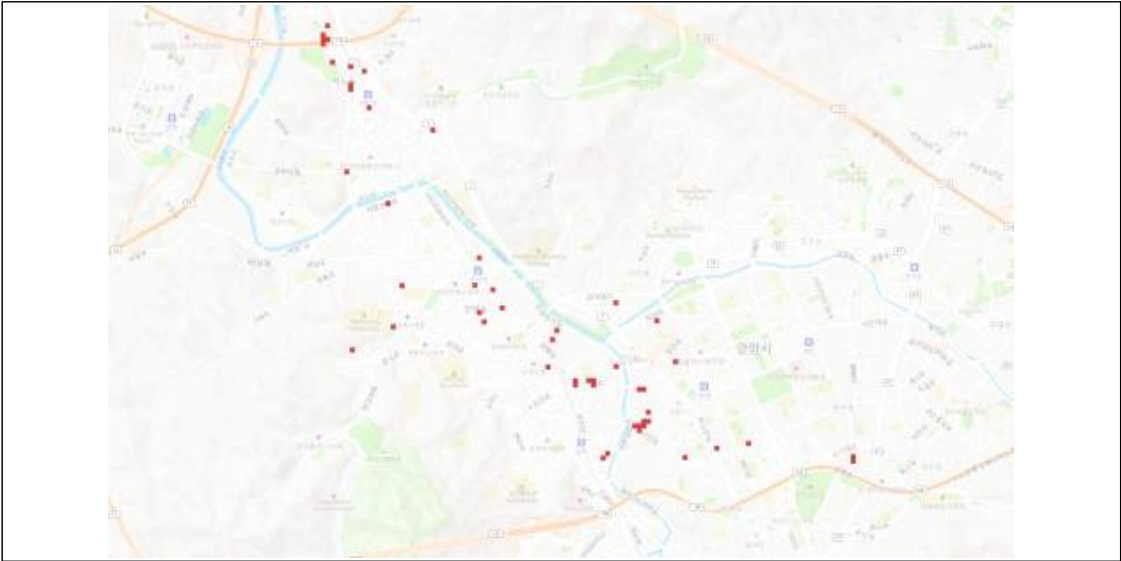
〈표 9〉 LG 유플러스 모바일 통신 전처리 데이터

순번	K-V	필드명	설명	자료형	비고
1	KEY	TRIP_ID	데이터 발생일자	VARCHAR	P_YYYYMMDD D + "_" + DAY_TRIP_NO
2	VALUE (반복)	O_CELL_ID	출발 격자 ID	NUMERIC	
3		D_CELL_ID	도착 격자 ID	NUMERIC	
4		O_ADMI_CD	출발 행정동 ID	NUMERIC	
5		D_ADMI_CD	도착 행정동 ID	NUMERIC	
6		DPR_MT10_UNIT_TM	출발 시간	NUMERIC	Unixtime, 출발시간으로 정렬함
7		ARV_MT10_UNIT_TM	도착 시간	NUMERIC	Unixtime
8		DYNA_MVMT_DISTANCE_KM	이동거리	NUMERIC	km
9		DYNA_MVMT_SPEED	속도	NUMERIC	km/h

- 개별 ID의 통행에는 다양한 수단이 혼재되어 있을 수 있으며, 이에 따라 이동 속도를 기준으로 도보 및 도보 외 통행을 구분함
 - 일반적인 성인의 걸음걸이 속도인 5km/h를 기준으로 구분함
 - 전체 통행이 아닌 통행의 한 단위마다 수단을 구분함

〈표 10〉 LG 유플러스 모바일 통신 데이터 수단구분 예시 (트립넘버: 23669006)

출발격자	도착격자	출발시간	도착시간	이동 거리 (km)	이동 속도 (km/h)	수단 구분
79111780	76167722	2023년 10월 17일 오전 3:00	2023년 10월 17일 오전 3:10	15.2	91.2	도보 외
79111780	77159622	2023년 10월 17일 오전 4:00	2023년 10월 17일 오전 4:10	15.27	91.63	도보 외
79111781	75735773	2023년 10월 17일 오전 7:00	2023년 10월 17일 오전 7:10	15.21	91.23	도보 외
79111781	76375699	2023년 10월 17일 오전 3:30	2023년 10월 17일 오전 3:40	15.27	91.61	도보 외
79111781	77079630	2023년 10월 17일 오전 1:50	2023년 10월 17일 오전 2:00	15.28	91.69	도보 외
79111781	79111781	2023년 10월 17일 오전 1:40	2023년 10월 17일 오전 1:50	0.35	2.12	도보
79111781	79111781	2023년 10월 17일 오전 3:20	2023년 10월 17일 오전 3:30	0.36	2.18	도보
⋮	⋮	⋮	⋮	⋮	⋮	



〈그림 7〉 트립 넘버 23669006 이동궤적 예시

3) 통행궤적 및 통행정보 데이터 구축

- 출발시간과 도착시간 정보를 활용하여 시간정렬을 통해 개별 통행 ID별 통행 궤적에 대한 시퀀스 정보를 포함하는 데이터를 구축하였음

〈표 11〉 시퀀스 부여를 통한 개별 통행ID별 통행궤적 데이터 구축 (트립넘버: 23669006)

순서	출발격자	도착격자	출발시간	도착시간
1	79111781	79111781	2023년 10월 17일 오전 1:40	2023년 10월 17일 오전 1:50
2	79111781	77079630	2023년 10월 17일 오전 1:50	2023년 10월 17일 오전 2:00
3	79111780	76167722	2023년 10월 17일 오전 3:00	2023년 10월 17일 오전 3:10
4	79111781	79111781	2023년 10월 17일 오전 3:20	2023년 10월 17일 오전 3:30
⋮	⋮	⋮	⋮	⋮

- 개별 ID별 다양한 통행정보 데이터를 포함하는 개별 통행정보 DB를 설계함

〈표 12〉 모바일 통신 데이터 기반 개별 통행정보 DB 구축

순번	필드명	설명	자료형	비고
1	TRIP_ID	통행 ID	VARCHAR	yyyy-MM-dd_{\$id}
2	O_CELL_ID	출발 격자 ID	NUMERIC	
3	D_CELL_ID	도착 격자 ID	NUMERIC	
4	O_ADMI_CD	시작 행정동 코드	NUMERIC	
5	D_ADMI_CD	도착 행정동 코드	NUMERIC	
6	LAYOVER_CNT	행정동 경유 수	NUMERIC	
7	WALK_RATE	도보 비율	NUMERIC	격자-격자 5km/h 미만 속도
8	WALK_CAL_SPD	도보 속도(시간, 속도로 재계산)	NUMERIC	km/h
9	WALK_CAL_DIST	도보 거리(시간, 속도로 재계산)	NUMERIC	km
10	WALK_SPD	도보 속도(평균)	NUMERIC	km/h
11	WALK_DIST	도보 거리(합)	NUMERIC	km
12	WALK_TIME	도보 시간(합)	NUMERIC	hour
13	NOT_WALK_RATE	도보외 비율	NUMERIC	격자-격자 5km/h 이상 속도
14	NOT_WALK_CAL_SPD	도보외 속도(시간, 속도로 재계산)	NUMERIC	km/h
15	NOT_WALK_CAL_DIST	도보외 거리(시간, 속도로 재계산)	NUMERIC	km
16	NOT_WALK_SPD	도보외 속도(평균)	NUMERIC	km/h
17	NOT_WALK_DIST	도보외 거리(합)	NUMERIC	km
18	NOT_WALK_TIME	도보외 시간(합)	NUMERIC	hour

4) 기술적 검토사항

- 도로 통행 상세 분석 및 통행지표 구축과 관련하여 기술적 검토를 아래와 같이 수행함

〈표 13〉 모바일 통신 데이터 기반 개별 통행정보DB 구축

구분	검토 항목	검토 내용
도로통행 상세분석	X,Y좌표 데이터	- 격자 단위 이동궤적 정보로는 실질적으로 도로를 구분해 내기에 한계가 있음 - X,Y좌표의 수집을 통해 도로를 구분해 낼 수 있는 기술 개발이 필요함
	시간 단위	- 10분 단위의 데이터로는 미시적인 분석 어려움 - 실질적으로 도로 통행 구분을 위해서는 초단위의 데이터가 필요함
	가속도 센서 데이터 분석	- 휴대폰에 내장된 가속도 센서는 사용자의 움직임의 방향과 강도를 측정하므로, 이를 이용하여 도보, 달리기, 자전거 타기 등 다양한 활동을 구분할 수 있음
	네트워크 맵매칭	- 다양한 수준의 교통망 네트워크와 통신 데이터의 결합을 통해 도로 통행을 명확히 구분할 수 있는 기술 개발이 필요함 - 해당 기술을 이용할 시 도로 통행 구분에 대한 정확도가 향상될 것으로 기대함
통행지표 구축	데이터 정확성	- 이동 경로 데이터의 정확성은 결과의 신뢰성에 큰 영향을 미치기 때문에 데이터 수집 과정에서 오류가 없었는지, 데이터가 충분히 정확한지에 대한 검증이 필요함
	데이터 처리 및 분석	- 모바일 통신 원시 데이터는 대용량 데이터로, 이를 처리하기 위해서는 최적화된 DB설계, 적절한 빅데이터 가공환경, 인프라 시스템이 중요함 - 구축된 데이터를 재가공하는 과정을 여러 차례 거쳐야 하는 경우 데이터를 빠르게 가공하기 위한 기술이 필요함
	통행지표 산출 알고리즘	- 최근 최적화, 머신러닝, 딥러닝 등 통행지표 생성을 위한 다양한 기술개발이 진행되고 있음 - 기술적으로 완성도 높은 알고리즘들을 활용하기 위해 고품질 데이터 셋이 필요하며 최신 알고리즘에 대한 연구 및 기술 개발이 필요함
	도메인 전문기술 필요	- 단순 알고리즘 뿐만 아니라 도메인 내에서 전문지식을 활용하여 산출해 낼 수 있는 통행지표 산출이 필요함 - 데이터에 대한 심도 있는 고찰과 분석을 통해 교통에서 유의미한 통행지표를 생산해 낼 수 있는 기술이 필요함
	데이터 속성정보 취득	- 성별, 연령, 통행 목적 등 통행지표 산출을 위한 다양한 속성정보 확보가 필요함

3. 모바일 통신데이터 기반 통행 DB 검증

가. 개요

- 모바일 통신데이터를 검증하기 위해서는 참값(ground truth)이 있어야 하나, 참값으로 볼 수 있는 기준 데이터가 없어, 통행정보를 담고 있는 데이터, 통행량과 연관성이 높은 각종 사회경제지표를 수집하여 상호교차검증을 수행함
- 검증 데이터로 모바일 통신데이터 3종(SKT 지라프 데이터, SKT 리트머스 데이터, KT 데이터), 조사자료 3종(통신사 데이터가 활용되지 않은 통행실태조사자료 중 가장 최신의 자료인 2019년 자료, 통신데이터가 활용된 통행실태조사자료 중 가장 최신의 자료인 2021년 자료, 2020년 인구총조사자료), 사회경제지표 자료 6종(인구, 학교, 산업, 취업률, 기지국, 행정경계), 실적자료 3종(수송실적 자료, 대중교통 카드 자료, TCS 자료)을 수집함
- 통신사, 가공알고리즘 차이에 따른 영향을 파악하기 위해 서로 다른 통신사, 각기 다른 가공 알고리즘이 적용된 모바일 통신데이터의 통행량, 통행분포를 비교하고, 기존 데이터 수집 방식의 차이를 파악하기 위해 조사를 통해 수집된 통행자료의 통행량, 통행 분포를 비교함
- 인구수, 학생수, 학령인구수 등 통행발생량·도착량과 연관이 있는 사회경제지표와 모바일 통신데이터, 조사자료의 통행량 분포를 비교하여 데이터에 대한 설명력을 확인함
- 일부 수단 통행에 한해서 참값으로 볼 수 있는 실적자료와 일치하는지 비교 분석함
- 또한 검증 결과를 토대로 모바일 통신데이터와 기준값을 설정하고, 데이터간 통행량 차이에 대해 분석하여 이를 보정할 수 있는 방법에 대해 제시함

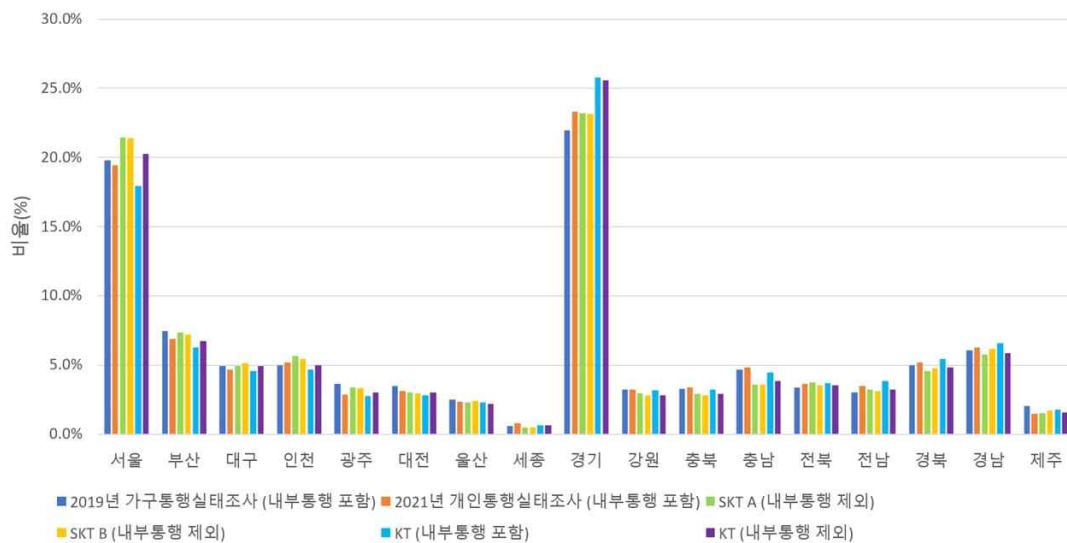
나. 검증 결과

1) 데이터간 상호교차검증 결과

① 시도

○ (전체 통행) 조사자료 2종과 모바일 통신데이터 4종의 시도별 통행 발생 비율, 도착 비율이 대체로 유사하게 나타났으며, 총량 측면에서는 2019년 가구통행실태조사(내부통행 포함)과 KT 데이터(내부통행 포함)가 가장 유사하게 나타남

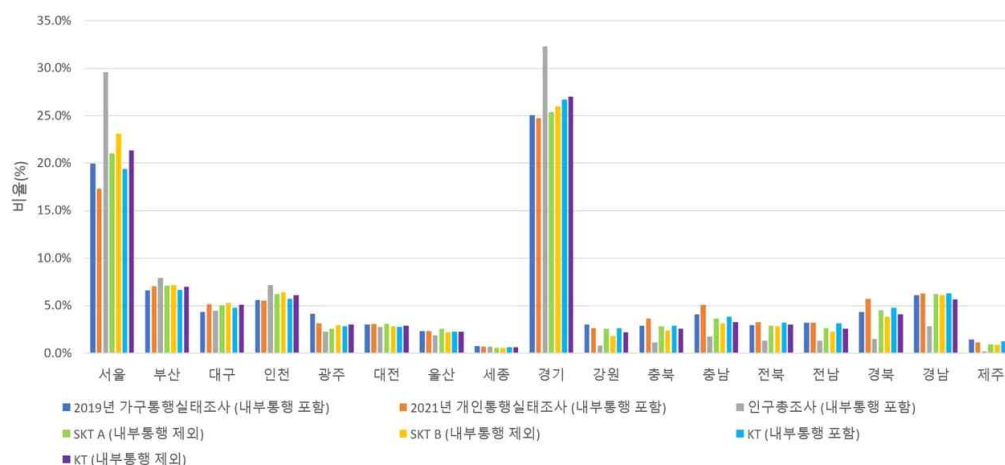
- 모바일 통신데이터는 대체로 조사자료보다 서울, 경기 지역에서 통행 비율이 다소 높게 나타났고, 충남, 경북 등 일부 지방지역에서는 통행 비율이 다소 낮게 나타남



〈그림 8〉 시도 단위 전체 통행에 대한 데이터간 상호교차검증 결과 (전체 - 발생량)

○ (통근 통행) 인구총조사 결과를 제외한 나머지 데이터의 시도별 분포가 유사하게 나타났으나, 통행량은 조사자료에 비해 모바일 통신데이터가 모든 지역에서 낮게 집계됨

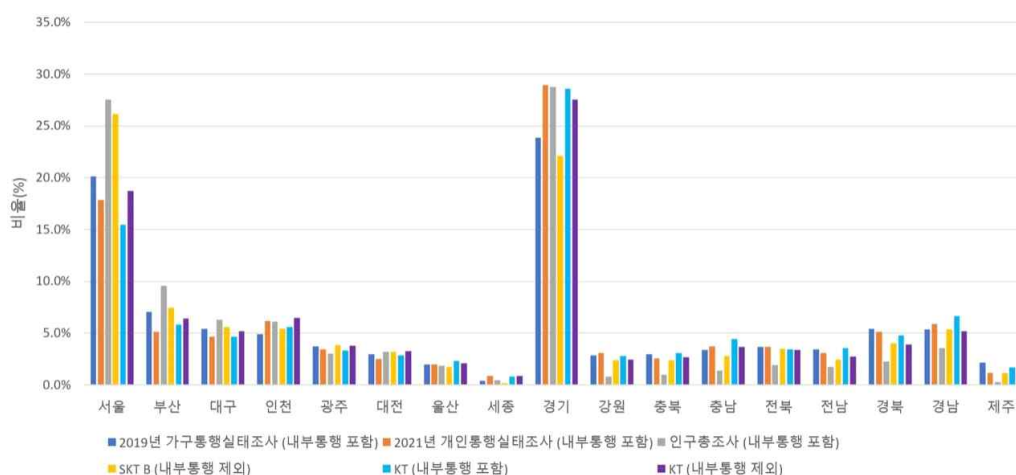
- 2019년 가구통행실태조사와 내부통행을 포함한 KT 데이터가 시도별 통행 비율이 가장 유사한 것으로 분석됨



〈그림 9〉 시도 단위 전체 통행에 대한 데이터간 상호교차검증 결과 (통근 - 발생량)

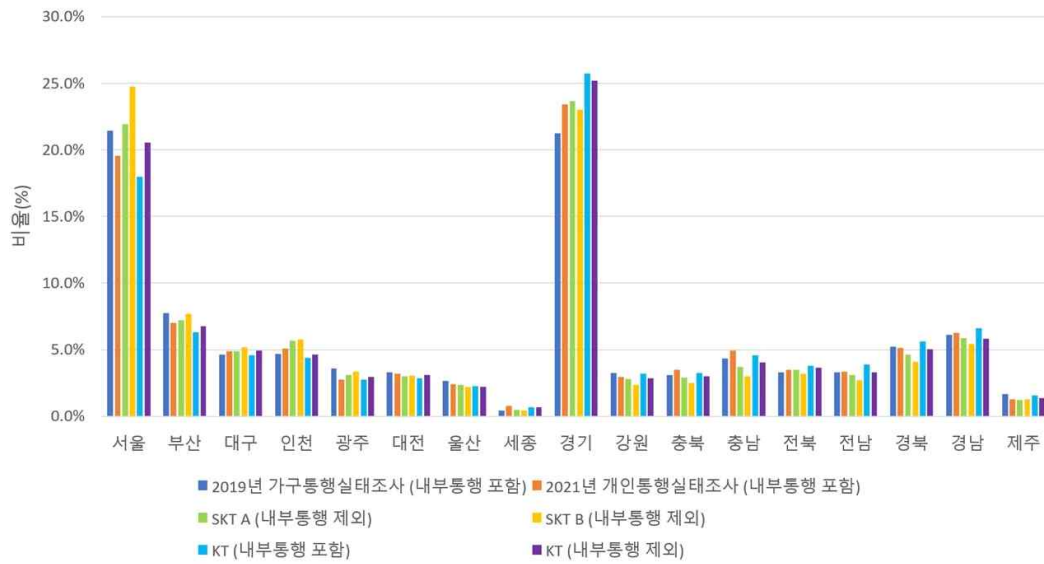
- (통학 통행) 서울, 경기 지역에서 데이터간 통행 비율 차이가 크게 나타나며, 통행량으로 보면, 지역별 차이뿐만 아니라 전체 총량에서도 데이터간 차이가 크게 나타남

- 2019년 가구통행실태조사와 내부통행을 제외한 KT 데이터가 시도별 통행 비율이 가장 유사한 것으로 분석됨



〈그림 10〉 시도 단위 전체 통행에 대한 데이터간 상호교차검증 결과 (통학 - 발생량)

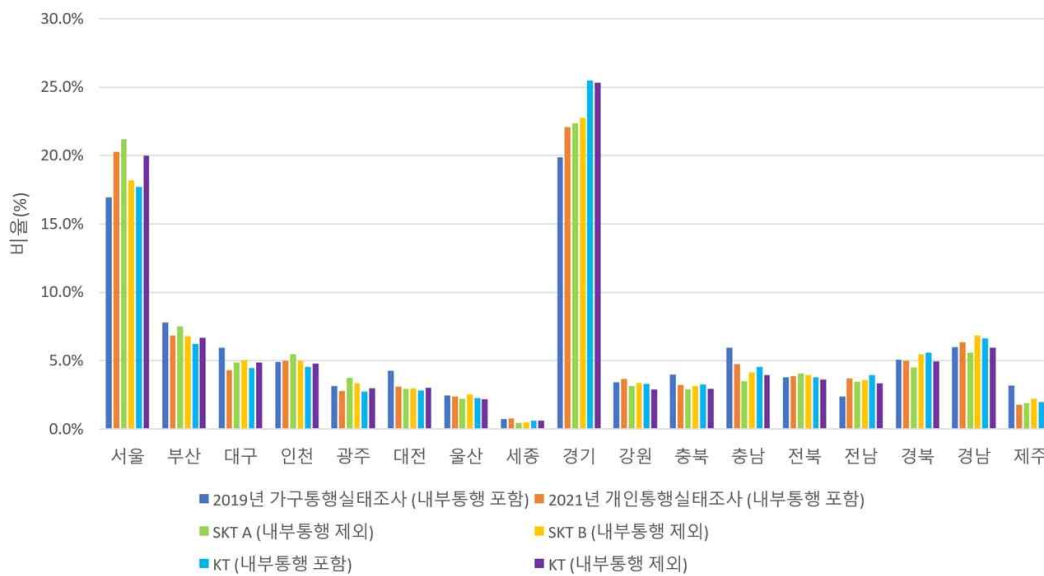
- (귀가 통행) 전체 통행과 마찬가지로 서울, 경기 지역을 제외하고는 조사자료와 모바일 데이터는 약 0.4% 정도의 근소한 차이를 보이며, 통행량은 조사자료에 비해 모바일 데이터의 결과가 매우 적은 편 (조사자료의 1/2~1/3 수준)



〈그림 11〉 시도 단위 전체 통행에 대한 데이터간 상호교차검증 결과 (귀가 - 발생량)

○ (기타 통행) 다른 목적통행에 비해 데이터간 통행량 차이가 다소 큰 편으로 서울, 경기, 충남 지역에서 데이터간 다소 큰 차이를 보임

- 통행량은 조사자료보다 모바일 통신데이터가 전반적으로 높게 나타남



〈그림 12〉 시도 단위 전체 통행에 대한 데이터간 상호교차검증 결과 (기타 - 발생량)

② 시군구

- (발생량 및 도착량) 전체, 통행목적별(통근, 통학, 귀가, 기타) 데이터간 발생량, 도착량에 대한 분포의 유사성과 통행발생·도착요인(인구수, 종사자수 등)을 기준으로 분포의 유사성을 검증한 결과는 다음 <표 14>와 같음

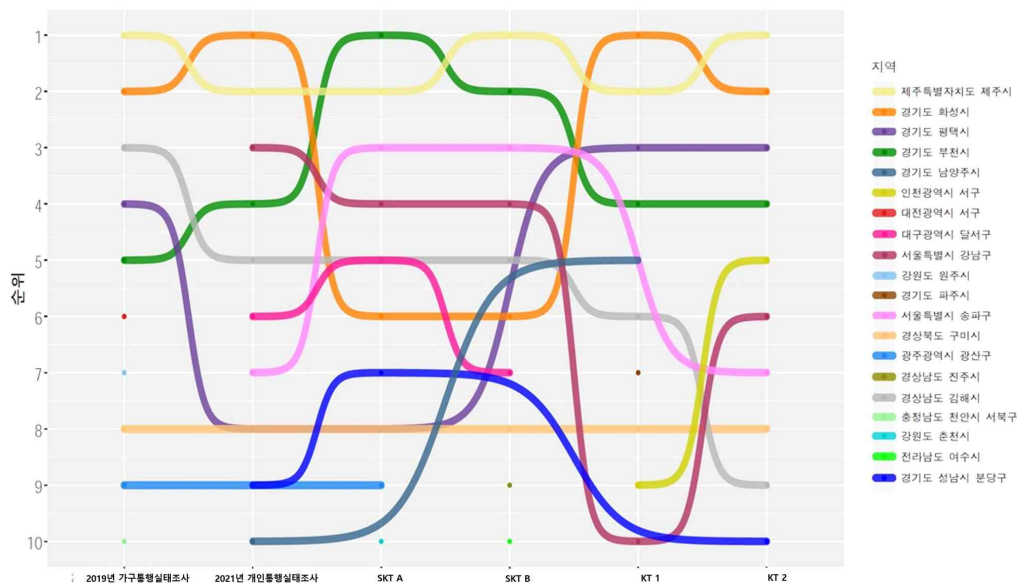
<표 14> 시군구 단위 발생량, 도착량에 대한 상호교차검증 결과 (요약)

통행 목적	비교 기준	발생량	도착량
전체	데이터간 유사성 비교	- 전반적으로 데이터간 상관관계가 높은 것으로 분석됨 (R: 0.94~0.99)	SKT A(내부통행 제외)와 KT(내부통행 포함)의 상관관계가 다소 낮은 것으로 분석됨(R: 0.87)
	통행발생·도착요인을 기준으로 설명력 비교	- 인구수를 기준으로 2019년 가구통행실태조사(내부통행 포함)의 설명력이 다소 낮은 것으로 분석됨 (2019년 가구통행실태조사 R-Squared: 0.79) - 인구밀도에 대한 설명력은 전반적으로 낮은 것으로 분석됨 (R-Squared: 0.2~0.38)	인구수에 대한 설명력은 전반적으로 다 높고(R-Squared: 0.80~0.92), 인구밀도에 대한 설명력은 전반적으로 다 낮은 것으로 분석됨 (R-Squared: 0.18~0.42)
통근	데이터간 유사성 비교	- 2019년 가구통행실태조사(내부통행 포함)와 인구총조사(내부통행 포함)의 상관관계만 다소 낮은 것으로 분석됨(R: 0.83)	전반적으로 데이터간 상관관계가 높은 것으로 분석됨(R: 0.95~0.99)
	통행발생·도착요인을 기준으로 설명력 비교	- 취업자수를 기준으로 인구총조사(내부통행 포함)를 제외한 나머지 데이터의 설명력이 높은 것으로 분석됨(인구총조사 R-Squared: 0.78)	종사자수에 대한 설명력은 전부 다 높게 나타남(R-Squared: 0.94~0.97)
통학	데이터간 유사성 비교	- 인구총조사(내부통행 포함)는 다른 데이터와 통행 분포 유사성이 다소 낮은 것으로 분석됨 (R: 0.80~0.89)	- 2019년 가구통행실태조사(내부통행 포함)는 다른 데이터와 상관관계가 다소 낮은 것으로 분석되었으며 (R: 0.63~0.81), 2021년 가구통행실태조사(내부통행 포함)와 SKT B(내부통행 제외)(R: 0.76), SKT B(내부통행 제외)와 KT(내부통행 포함)의 상관관계도 다소 낮은 것으로 분석됨(R: 0.73)
	통행발생·도착요인을 기준으로 설명력 비교	- 학령인구수를 기준으로 인구총조사(내부통행 포함)와 SKT B(내부통행 제외)는 설명력이 낮은 것으로 분석됨 (인구총조사 R-Squared: 0.69 / SKT B -	수용학생수에 대한 설명력은 전부 비슷한 것으로 분석됨(R-Squared: 0.69~0.77) - 학교수를 기준으로 2019년 가구통행실태조사(내부통행 포함)와 SKT

통행 목적	비교 기준	발생량	도착량
		R-Squared: 0.75)	B(내부통행 제외)의 설명력이 낮고(R-Squared: 0.41~0.44), 2021년 가구통행실태조사(내부통행 포함)와 KT(내부통행 포함)는 설명력이 높은 것으로 분석됨(R-Squared: 0.78~0.81)
귀가	데이터간 유사성 비교	- 전반적으로 데이터간 상관관계가 높은 것으로 분석됨(R: 0.89~0.98)	전반적으로 데이터간 상관관계가 높은 것으로 분석됨(R: 0.93~0.99)
	통행발생·도착요인을 기준으로 설명력 비교	- 인구수를 기준으로 대체로 모바일 데이터의 설명력이 높은 것으로 분석되었으며(모바일 데이터 R-Squared: 0.72~0.90), 2019년 가구통행실태조사(내부통행 포함)의 설명력이 가장 낮은 것으로 분석됨(R-Squared: 0.6) - 인구밀도에 대한 설명력은 전반적으로 낮은 것으로 분석됨(R-Squared: 0.18~0.40)	인구수에 대한 설명력은 전반적으로 다 높고(R-Squared: 0.86~0.97), 인구밀도에 대한 설명력은 전반적으로 다 낮은 것으로 분석됨(R-Squared: 0.2~0.49)
기타	데이터간 유사성 비교	- 전반적으로 데이터간 상관관계가 높은 것으로 분석됨(R: 0.90~0.99)	전반적으로 데이터간 상관관계가 높은 것으로 분석됨(R: 0.88~0.99)
	통행발생·도착요인을 기준으로 설명력 비교	- 인구수를 기준으로 2019년 가구통행실태조사(내부통행 포함)의 설명력이 다소 낮은 것으로 분석됨(R-Squared: 0.68) - 인구밀도에 대한 설명력은 전반적으로 낮은 것으로 분석됨(R-Squared: 0.15~0.36)	인구수에 대한 설명력은 전반적으로 높지 않으며, 2019년 가구통행실태조사(내부통행 포함)의 설명력이 가장 낮은 것으로 분석됨(R-Squared: 0.56~0.82) 인구밀도에 대한 설명력은 전반적으로 낮은 것으로 분석됨(R-Squared: 0.15~0.34)

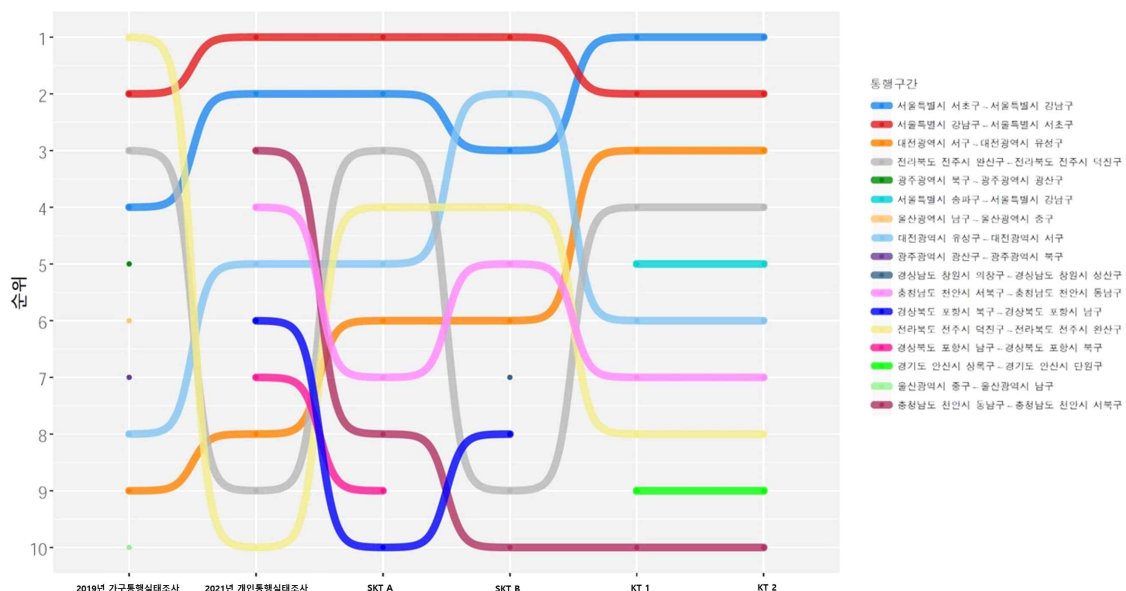
○ (OD 통행량) 데이터간 전체 통행에 대한 분포량(OD 통행량)을 비교한 결과, 1위부터 10위까지 통행량이 가장 많은 구간은 대체로 모든 데이터에서 유사하게 나타나며 순위 변동 폭이 크지 않은 것으로 분석됨

- 동일 시군구 내에서 이루어진 통행을 기준으로, 제주특별자치도 제주시, 경기도 화성시, 경상남도 김해시, 경기도 평택시, 경기도 부천시, 경상북도 구미시는 모든 데이터에서 최다 통행지역 10위 안에 포함되어 있음



〈그림 13〉 데이터별 최다 통행구간 TOP 10 (시군구 내부 통행)

- 외부 통행을 기준으로 서울특별시 강남구~서울특별시 서초구, 전라북도 전주시 덕진구~전라북도 전주시 완산구, 서울특별시 서초구~서울특별시 강남구, 전라북도 전주시 완산구~전라북도 전주시 덕진구, 대전광역시 유성구~대전광역시 서구, 대전광역시 서구~대전광역시 유성구 구간이 모든 데이터에서 10위 안에 포함되어 있음



〈그림 14〉 데이터별 최다 통행구간 TOP 10 (외부 통행)

③ 읍면동

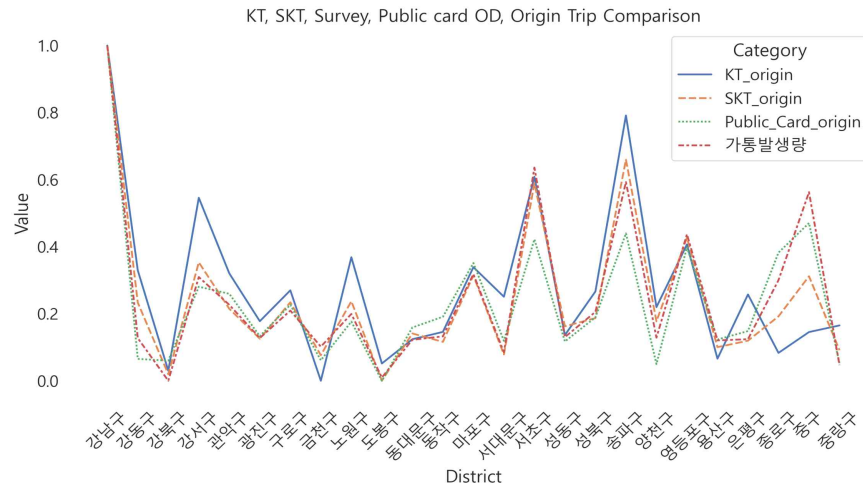
- 수도권 내부의 통근 통행, 통학 통행에 대하여, 2019년 가구통행실태조사(내부통행 포함), SKT B(내부통행 제외), KT(내부통행 포함) 세 가지 데이터를 비교 분석함
 - 2019년 가구통행실태조사를 기준으로 SKT B와 KT의 통행량에 대하여 상관관계, 분산의 유사도, 오차량을 비교하고, ANOVA 검정을 통해 세 가지 데이터에 대한 통행량 분포 차이를 확인하였으며, KS 검정을 통해 SKT B와 KT 두 가지 데이터의 통행량 분포 차이를 상호 비교함
 - 또한, 통행발생·도착요인과 각 데이터의 통행발생량·도착량의 관계를 분석함
- 분석 결과, 통근 통행은 발생량 도착량 모두 KT가 SKT B 2019년 가구통행실태조사와 유사하였고, 통학 통행은 SKT B가 KT보다 2019년 가구통행실태조사와 분산은 유사하나 오차량은 KT가 상대적으로 적은 것으로 나타났으며, ANOVA, KS 검정 결과 통근 통행은 대부분의 읍면동에서 분포 차이가 있으나, 통학 통행은 차이가 없는 것으로 분석됨
- 통행발생·도착요인에 대한 데이터의 설명력은 전반적으로 높은 편이었으나, 일부 지역에서 데이터별 설명력 차이가 존재하는 것으로 확인됨
 - 통근 통행의 경우 서울 지역에서 2019년 가구통행실태조사가 다른 데이터에 비해 취업자수에 대한 설명력이 낮게 나타남
 - 통학 통행의 경우 서울 지역에서는 KT가, 경기 지역에서는 2019년 가구통행실태조사가 다른 데이터에 비해 19세 미만의 학령인구수에 대한 설명력이 높게 나타남

2) 실적 자료와의 일치성 검증 결과

① 대중교통 카드자료 기준

- 서울시의 대중교통 실적자료를 기준으로 2019년 가구통행실태조사(내부통행 포함), SKT B(내부통행 제외), KT(내부통행 포함)을 검증하였으며, 분석 결과

KT·SKT·KTDB·교통카드 데이터는 모두 통계적으로 유사한 특성을 갖는 것으로 확인됨



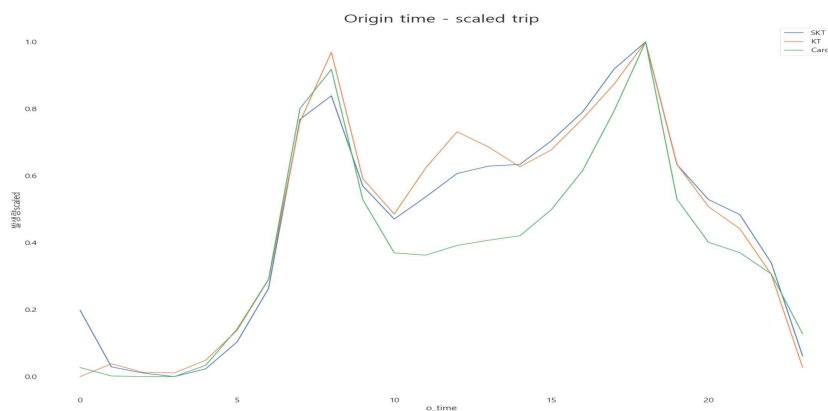
〈그림 15〉 시군구별 KT·SKT·KTDB·교통카드 데이터의 통행발생량 비교

〈표 15〉 시군구별 KT·SKT·KTDB·교통카드 데이터의 통행발생량 통계검증 결과

구분	ANOVA	Bartlett	Levene
P-value	0.88	0.93	0.96

주) ANOVA 검정은 집단의 평균을 비교하는 검정이며, Bartlett, Levene 검정은 ANOVA 검정 수행 전 대상 집단의 분산이 같은지 여부를 확인하는 테스트로, 유의값(p-value)이 0.05 이상인 경우에는 집단간 차이가 없다고 해석함

- 시간대별로 통행발생량·도착량을 비교 분석한 결과, 첨두시간대에는 모두 비슷한 패턴을 보이거나, 일부 시간대(10~15시)에서는 SKT, KT, 교통카드 모두 상이한 패턴을 보이는 것으로 확인됨



〈그림 16〉 시간대별 KT·SKT 교통카드의 통행발생량 비교

② 수송실적 기준

- 지역간 통행량, 권역간 통행량, 도서지역 통행량에 대하여 모바일 통신데이터와 실적자료의 결과가 일치하는지 확인하였으며, 비교한 결과는 <표 16>과 같음
 - 지역간 통행량: 전국 6X3 스크린라인을 통과하는 교통량을 비교한 결과, -58.7% ~ 15.5% 차이가 발생하는 것으로 나타남
 - 권역간 통행량: 수도권, 부산권, 강원권, 전라권 등 권역간 통행량을 비교한 결과, 25.8% ~ 80.5% 차이가 발생하는 것으로 나타남
 - 도서지역 통행량: 도서지역으로 유입·유출하는 통행량을 비교한 결과, -65.4% ~ 103.1% 차이가 발생하는 것으로 나타남

<표 16> 모바일 통신데이터 기반 지역간 통행량 추정 결과와 수송실적간의 비교

(단위 : 천통행/일)

구분		수송실적							통신	차이
		승용차	버스	트럭	철도	항공	해운	합계		
가로축 1	상→하	1,016	257	185	54	0	0	1,513	691	-54.3%
	하→상	1,160	294	213	52	0	0	1,719	711	-58.7%
가로축 2	상→하	512	184	164	93	10	0	963	736	-23.6%
	하→상	514	200	166	95	10	0	984	718	-27.0%
가로축 3	상→하	559	210	157	73	10	0	1,008	672	-33.4%
	하→상	575	213	159	70	10	0	1,028	684	-33.4%
가로축 4	상→하	201	76	83	61	10	0	431	464	7.7%
	하→상	248	86	104	58	10	0	505	461	-8.7%
가로축 5	상→하	226	98	104	43	10	0	480	467	-2.8%
	하→상	239	104	110	41	10	0	503	460	-8.5%
가로축 6	상→하	480	209	141	41	8	0	880	669	-23.9%
	하→상	487	215	155	40	8	0	904	680	-24.8%
세로축 1	좌→우	1,363	542	340	102	9	1	2,356	1,450	-38.5%
	하→상	1,491	642	387	105	9	1	2,636	1,400	-46.9%
세로축 2	좌→우	242	105	79	55	8	1	489	565	15.5%
	하→상	273	122	92	53	8	1	549	540	-1.8%
세로축 3	좌→우	315	102	88	33	8	0	545	603	10.6%
	하→상	496	142	156	31	8	0	833	608	-27.0%

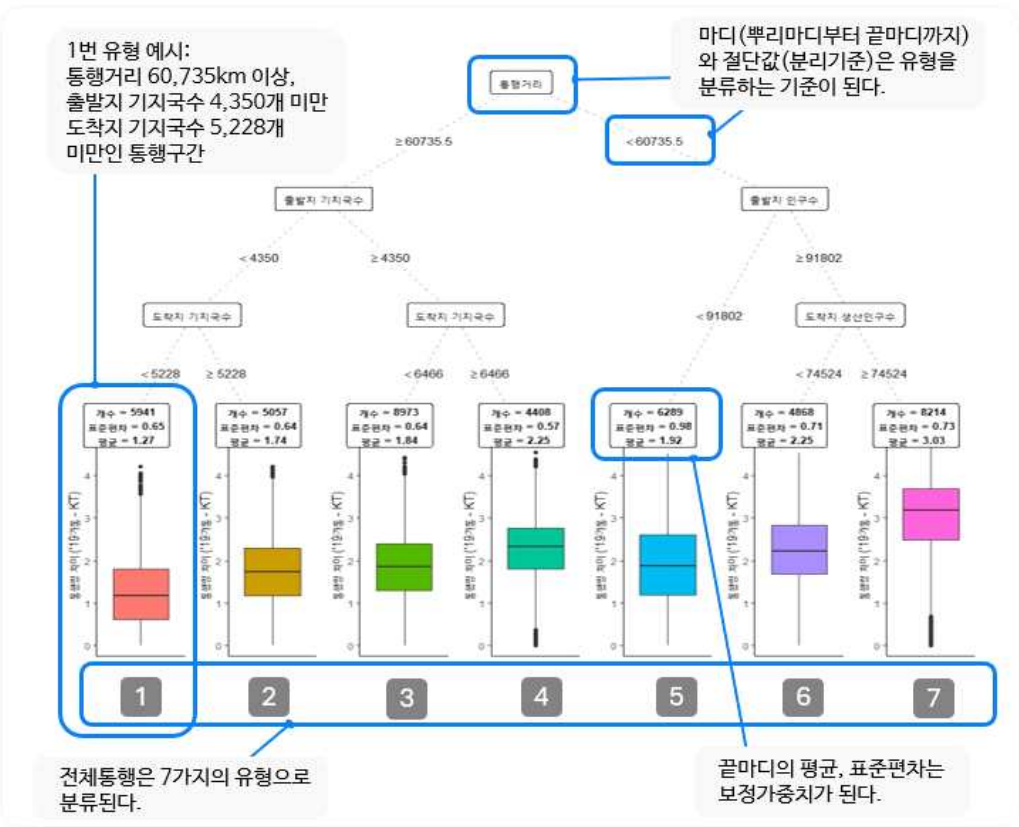
다. 보정 방법 개발

- 의사결정나무분석을 통해 데이터간 통행량 차이가 나타나는 통행구간의 특성을 파악하고, 각 유형별 보정 가중치를 산출함
 - 검증 결과를 토대로 한국교통연구원에서 구축한 KT 모바일 통신데이터를 보정 대상으로 2019년 가구통행실태조사를 기준값으로 설정함
 - 보정대상과 기준값의 차이가 두드러지게 나타나기 시작하는 시군구 단위의 OD 통행량(분포량)에 대하여 분석을 수행함
 - 전체 통행, 통행목적별(통근, 통학, 귀가, 기타)로 구분하여 기준값과 보정대상의 차이에 영향을 줄 것으로 예상되는 통행발생·도착요인을 설명변수로 선택하여 분석함(<표 17> 참조)
 - <표 17>에서 상관계수 $R \geq |0.2|$ 로, 종속변수와 영향이 적은 변수는 분석에서 제외함

<표 17> 통행목적별 설명변수

통행 목적	통행량과의 연관성		
	출발지 특성	도착지 특성	통행구간의 특성
전체	인구수, 학령인구수, 취업자수, 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	인구수, 수용학생수, 종사자수, 학교수, 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	통행거리
통근	인구수, 학령인구수, 생산인구수(비율), 고령인구수(비율), 취업자수, 기지국수(밀도), 면적	인구수, 생산인구수, 고령인구수, 종사자수, 기지국수(밀도), 면적	통행거리
통학	인구수, 학령인구수, 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	인구수, 수용학생수, 학교수(유·초·중·고등학교수, 유·초·중·고·대학교수), 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	통행거리
귀가	인구수, 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	인구수, 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	통행거리
기타	인구수, 생산인구수(비율), 고령인구수, 기지국수(밀도), 면적	인구수, 생산인구수(비율), 고령인구수, 기지국수(밀도), 면적	통행거리

- 분석 결과 중 데이터를 분류하는 데 사용한 분류기준(변수, 절단값)은 통행량 차이가 나타나는 통행구간을 유형화하는 기준으로 제시하고, 최종노드(끝마디)의 평균, 표준편차값을 보정가중치로 제시함(<그림 17> 참조)



<그림 17> 의사결정나무분석 수행결과 (전체 통행 예시)

- 분석 결과를 통해 보정유형과 보정가중치를 산출한 결과는 다음 <표 18>와 같음

<표 18> 통행 유형별 보정가중치

통행 유형	보정유형									보정가중치	
	통행 거리	인구수		기지국수 (밀도)		생산인구수 (비율)		수용 학생수	도착지	평균	표준 편차
		출발지	도착지	출발지	도착지	출발지	도착지				
전 체	①	≥ 60,736	-	-	< 4,350	< 5,228	-	-	-	1.27	0.65
	②	≥ 60,736	-	-	< 4,350	≥ 5,228	-	-	-	1.74	0.64
	③	≥ 60,736	-	-	≥ 4,350	< 6,466	-	-	-	1.84	0.64
	④	≥ 60,736	-	-	≥ 4,350	≥ 6,466	-	-	-	2.25	0.57
	⑤	< 60,736	< 91,802	-	-	-	-	-	-	1.92	0.98
	⑥	< 60,736	≥ 91,802	-	-	-	< 74,524	-	-	2.25	0.71
	⑦	< 60,736	≥ 91,802	-	-	-	≥ 74,524	-	-	3.03	0.73
통 근	①	≥ 60,315	-	-	< 4,350	-	-	-	-	0.13	0.06
	②	≥ 60,315	-	-	≥ 4,350	-	-	-	-	0.25	0.10

통행 유형		보정유형								보정가중치	
		통행 거리	인구수		기지국수 (밀도)		생산인구수 (비율)		수용 학생수	평균	표준 편차
			출발지	도착지	출발지	도착지	출발지	도착지	도착지		
	③	<60,315	-	-	-	-	<74,524	-	-	0.33	0.17
	④	<60,315	-	-	-	-	≥74,524	<74,524	-	0.37	0.18
	⑤	<60,315	-	-	-	-	≥74,524	≥74,524	-	0.87	0.18
통 학	①	≥50,473	-	-	-	-	<83,429	-	-	0.03	0.01
	②	≥50,473	-	-	-	-	≥83,429	-	<29,160	0.05	0.01
	③	≥50,473	-	-	-	-	≥83,429	-	≥29,160	0.14	0.03
	④	<50,473	-	-	(<18)	-	-	-	-	0.10	0.03
	⑤	<50,473	-	-	(≥18)	-	-	(<0.67)	-	0.11	0.03
	⑥	<50,473	-	-	(≥18)	-	-	(≥0.67)	-	0.38	0.05
	⑦	<50,473	-	-	(≥18)	-	-	(≥0.67)	-	0.38	0.05
귀 가	①	≥59,472	<108,763	<112,552	-	-	-	-	-	0.62	0.29
	②	≥59,472	≥108,763	<112,552	-	-	-	-	-	1.13	0.34
	③	≥59,472	-	≥112,552	<6,466	-	-	-	-	1.25	0.39
	④	≥59,472	-	≥112,552	≥6,466	-	-	-	-	1.58	0.34
	⑤	<59,472	-	-	-	-	-	<74,524	-	1.34	0.54
	⑥	<59,472	-	-	-	-	<74,524	≥74,524	-	1.55	0.47
	⑦	<59,472	-	-	-	-	≥74,524	≥74,524	-	2.26	0.38
기 타	①	≥58,582	-	-	<4,219	<5,247	-	-	-	1.02	0.55
	②	≥58,582	-	-	<4,219	≥5,247	-	-	-	1.42	0.58
	③	≥58,582	-	-	≥4,219	-	-	-	-	1.70	0.59
	④	<58,582	<91,802	-	-	-	-	-	-	1.61	0.89
	⑤	<58,582	≥91,802	-	-	-	-	<74,524	-	1.97	0.64
	⑥	<58,582	≥91,802	-	-	-	-	≥74,524	-	2.73	0.65

- 아래 산식에 보정가중치를 적용하여 보정 통행량을 산출하고, 보정할 대상에 보정 통행량을 더하거나 빼서 통행량을 보정할 수 있음(<표 19> 참조)

$$\text{보정 통행량} = \text{보정가중치의 평균 값} + \left\{ \frac{(\text{보정 전 통행량} - A)}{B} \right\} \times \text{보정가중치의 표준편차 값}$$

A: 해당통행 유형의 평균 통행량

B: 해당통행 유형의 통행량에 대한 표준편차

〈표 19〉 통행량 보정 예시

구분		통행구간 특성			통행 유형 분류	보정 전 통행량	→	보정 후 통행량*	보정 통행량
출발지	도착지	통행거리	출발지 기지국 수	도착지 기지국 수					
A	B	65,200	4,000	5,000	전체-①	10		11.59	1.59
C	D	68,000	4,900	4,300	전체-①	4		4.72	0.72
A	C	60,500	4,000	4,900	전체-①	8		9.3	1.30
E	F	70,000	2,000	1,000	전체-①	14		16.17	2.17
A	F	73,200	4,000	1,000	전체-①	3		3.58	0.58
평균 (A)						7.8	→		1.27
표준편차 (B)						4.5	→		0.65
(보정가중치 - 평균) 전체-① 유형의 평균						1.27			
(보정가중치 - 표준편차) 전체-① 유형의 표준편차						0.65			

*: 보정 전 통행량에서 보정 통행량을 더한 경우를 예시로 제시함

- 향후 이러한 보정 대상과 가중치를 이용하여 각 데이터간 산술평균 또는 가중
평균하여 구축된 기종점 통행량의 신뢰성을 개선할 수 있음

4. 모바일 통신데이터 기반 교통지표 산출 및 제공

가. 모바일 통신데이터 기반 교통지표

- 모바일 통신데이터는 기존 조사 기반 자료에 비해 높은 해상도의 시공간 정보(1년 365일 1시간 단위 500미터 격자단위)로 상세한 통행 및 활동 정보를 담고 있어, 다양한 교통지표 추출이 가능함
- 이에 2022년 사업에서 구축한 KT 모바일 통신데이터 기반 통행 DB를 기준으로 <표 20>과 같이 모바일 통신데이터로 산출 가능한 교통지표를 DB화 하여 구축함

<표 20> 모바일 통신데이터 기반 교통지표

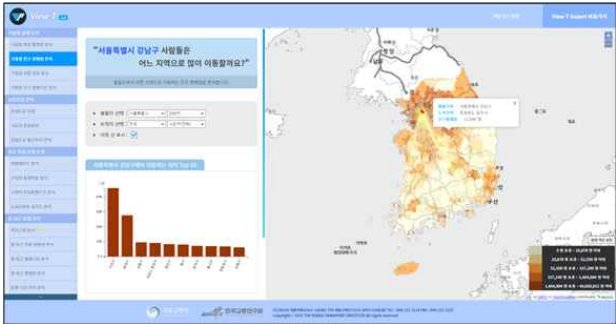
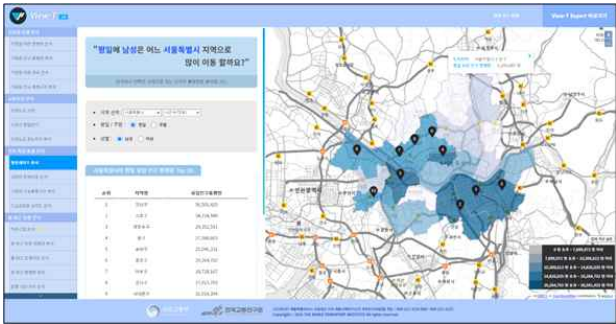
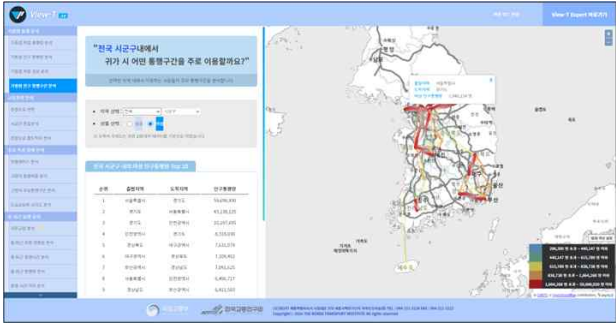
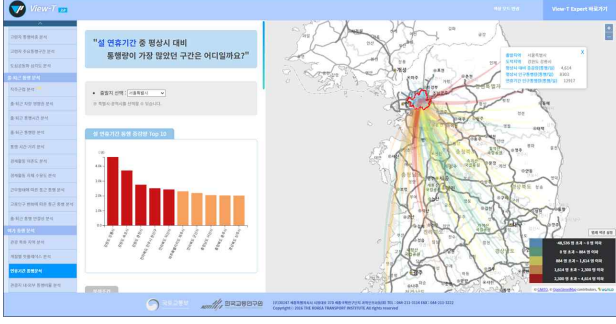
교통지표		설명	지표 추출 기준 (활용된 통행정보)
대분류	소분류		
통행량	발생량	출발지를 기준으로 집계한 통행량	출발지
	도착량	도착지를 기준으로 집계한 통행량	도착지
	분포량 (기중점통행량)	통행구간별로 집계한 통행량	출발지, 도착지
	기간별 통행량	연휴 기간 통행량	출발일자, 도착일자
		계절별 통행량	출발일자, 도착일자
	통행목적별 통행량	출근, 퇴근, 등교, 하교, 귀가, 여가, 관광 통행에 대해 집계한 통행량	출발지 체류유형, 도착지 체류유형
	통행자 특성별 통행량	연령대가 70대 이상의 고령인구에 대해 집계한 통행량	통행자의 연령
통행시간	평균통행시간	선택한 통행 대상의 통행시간을 평균한 값	통행시간
통행거리	평균통행거리	선택한 통행 대상의 통행거리를 평균한 값	출발지, 도착지
통행비율	내부 통행 비율	동일한 행정구역 내에서 이동하는 통행자가 전체 통행에서 차지하는 비율	출발지, 도착지
	외부 통행 비율	외부로 통행하는 통행자가 전체 통행에서 차지하는 비율	출발지, 도착지
	고령자 통행 비율	전체 통행자 중에서 고령 통행자가 차지하는 비율	통행자의 연령
	심야시간 출·퇴근 통행 비율	심야시간대 퇴근 통행량 대비 출근 통행량의 비율	출발시간, 도착시간, 출발지 체류유형, 도착지 체류유형
사회·경제 및	출·퇴근 증감량	전월 대비 기준월의 통근 목적 통행량의 증가량 또는 감소량	출발지 체류유형, 도착지 체류유형

교통지표		설명	지표 추출 기준 (활용된 통행정보)
대분류	소분류		
모빌리티 활동 지표	도심공동화 심각도	유동인구 대비 상주인구 비율	출발시간, 도착시간, 출발지 체류유형, 도착지 체류유형
	경제활동의존도	특별시·광역시 주변도시에서 발생한 출근 통행량 중 특별시·광역시로 출근한 통행량의 비율	출발지, 도착지, 출발지 체류유형, 도착지 체류유형
	경제활동 자체수용도	특별시·광역시에서 발생한 출근 통행량 중 지역 내 통행량의 비율	출발지, 도착지, 출발지 체류유형, 도착지 체류유형
	출·퇴근 통행연결성	출·퇴근 통행에 대한 기종점간의 모듈성	출발지, 도착지, 출발지 체류유형, 도착지 체류유형
	직주균형지표	거주지와 근무지간의 실제통근거리 대비 직주불일치로 발생하는 초과통근거리의 비율	출발지, 도착지, 출발지 체류유형, 도착지 체류유형

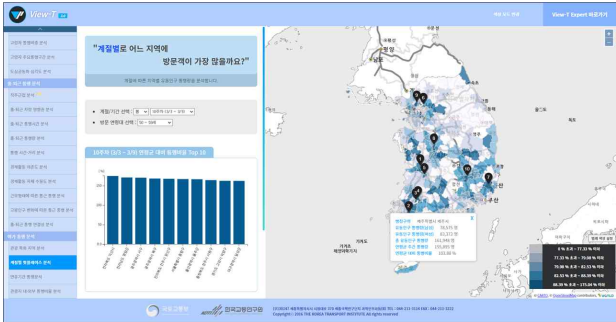
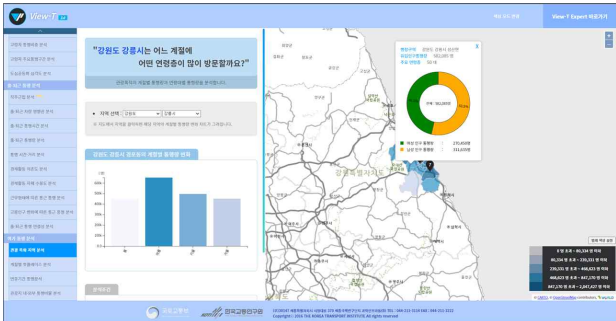
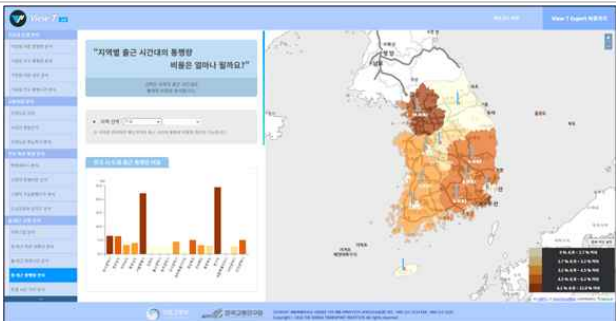

나. 온라인 서비스를 통한 교통지표 제공

- 모바일 통신데이터의 접근성·활용성을 높이기 위해 모바일 통신데이터로 구축한 교통지표를 온라인 서비스를 통해 제공함
 - 교통지표 DB를 온라인 서비스 내에 구축한 후, 교통지표를 확인할 수 있는 분석 도구를 개발하여 제공함
 - Light 버전의 분석도구: 일반 이용자를 대상으로 만든 분석도구로 기본적인 분석기능을 제공함
 - Expert 버전의 분석도구: 전문가를 대상으로 만든 분석도구로 Light 버전보다 상세 설정이 가능함
 - 온라인 서비스 내 구축한 교통지표(원본 데이터 포함)는 사용자가 다운받아 활용할 수 있도록 함

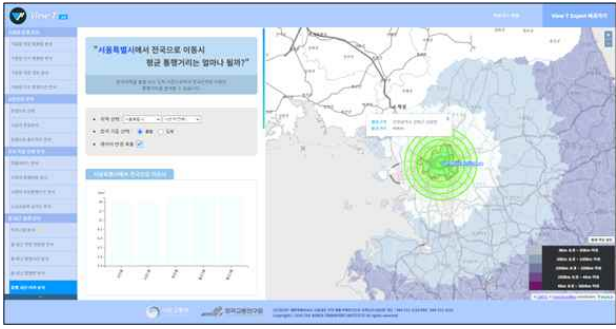
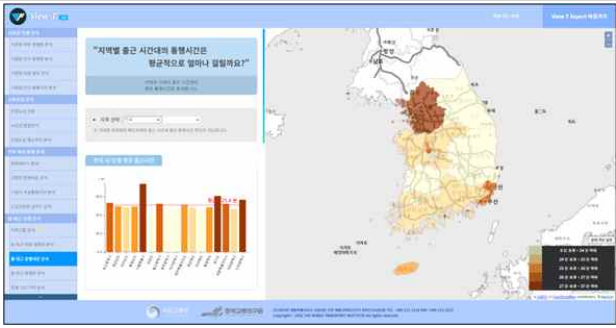
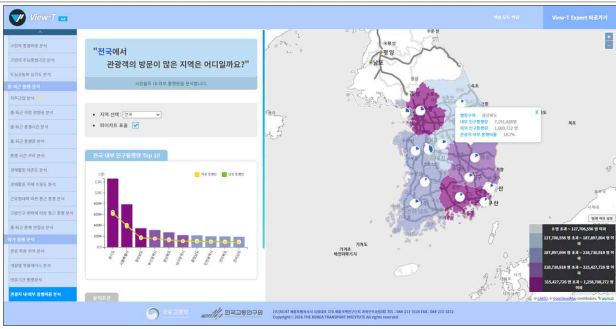
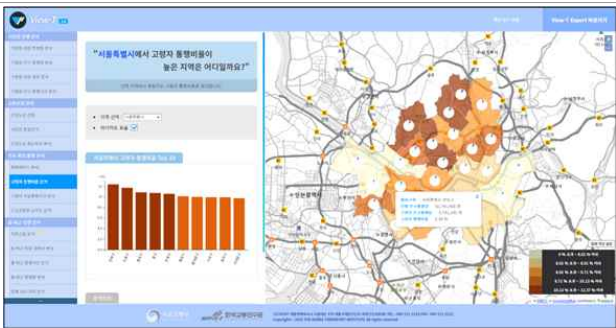
〈표 21〉 분석도구

교통 지표	분석도구		비고
	명칭	설명	
발생량 / 도착량	기종점 인구 통행량 분석	 <p>출발지에서 다른 지역으로 이동하는 인구 통행량을 분석</p>	Light, Expert 모두 제공
	핫 플레이스 분석	 <p>전국에서 선택한 지역으로 오는 인구의 통행량을 분석</p>	Light, Expert 모두 제공
기종점 통행량 (분포 량)	기종점 인구 통행 구간 분석	 <p>선택한 지역 내에서 이동하는 사람들의 주요 통행 구간을 분석</p>	Light, Expert 모두 제공
연휴 기간 통행량	연휴 기간 통행 분석	 <p>연휴기간 지역별 유동인구 통행량을 분석</p>	Light, Expert 모두 제공

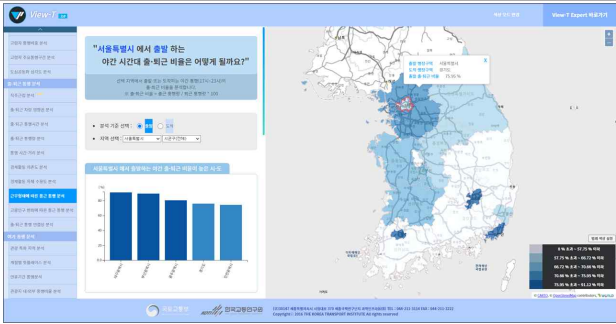
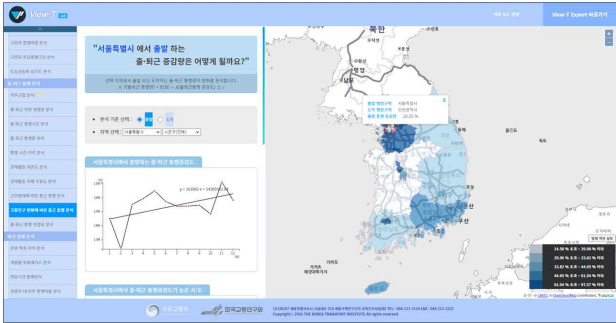

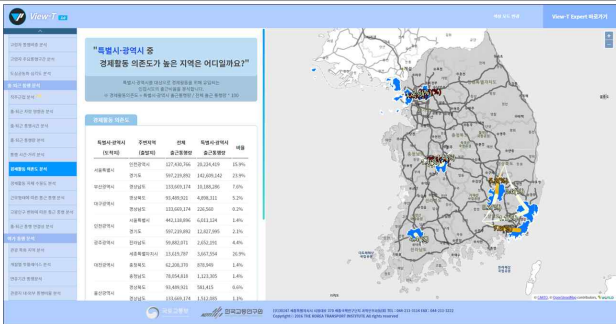
〈표 22〉 분석도구 (계속)

교통 지표	분석도구 설명		비고
	명칭		
계절별 통행량	계절별 핫 플레이스 분석	 <p>계절에 따른 지역별 유동인구 통행량을 분석</p>	Light 버전 한정
	관광 특화 지역 분석	 <p>관광목적의 계절별 통행량과 연령대별 통행량을 분석</p>	Light 버전 한정
출· 퇴근 통행량 , 통행 비율	출· 퇴근 통행량 분석	 <p>선택한 지역의 출근 시간대, 퇴근시간대의 통행량, 통행 비율을 분석</p>	Light 버전 한정
고령자 통행량	고령자 주요 통행 구간 분석	 <p>선택한 지역 내에서 고령자의 주요 통행구간을 분석</p>	Light, Expert 모두 제공

〈표 23〉 분석도구 (계속)

교통 지표	분석도구		비고
	명칭	설명	
통행 시간, 통행 거리	통행 시간·거리 분석	 <p>분석지역을 출발 또는 도착 기준으로 하여 전국단위로 이동한 통행거리를 분석</p>	Light 버전 한정
출·퇴근 평균 통행 시간	출·퇴근 통행 시간 분석	 <p>선택한 지역의 출근 시간대의 평균 통행시간을 분석</p>	Light, Expert 모두 제공
내·외부 통행 비율	관광지 내·외부 통행 비율 분석	 <p>사람들의 내·외부 통행량을 분석</p>	Light, Expert 모두 제공
고령자 통행 비율	고령자 통행 비중 분석	 <p>선택 지역에서 출발하는 고령자 통행비율을 분석</p>	Light, Expert 모두 제공

〈표 24〉 분석도구 (계속)

교통 지표	분석도구		비고
	명칭	설명	
심야 시간 출· 퇴근 통행 비율	근무 형태에 따른 통근 통행 분석	 <p>선택 지역에서 출발 또는 도착하는 야간 통행(17시~23시)의 출·퇴근 비율을 분석</p>	Light 버전 한정
출· 퇴근 증감량	고용 인구 변화에 따른 통근 통행 분석	 <p>선택 지역에서 출발 또는 도착하는 출·퇴근 통행량의 변화를 분석</p>	Light 버전 한정
도심 공동화 심각도	도심 공동화 심각도 분석	 <p>전체 유동인구 대비 상주인구를 분석하여 도심공동화가 심각한 지역을 분석</p>	Light, Expert 모두 제공
경제 활동 의존도	경제 활동 의존도 분석	 <p>특별시·광역시를 대상으로 경제활동을 위해 유입되는 인접시도의 출근비율을 분석</p>	Light, Expert 모두 제공

〈표 25〉 분석도구 (계속)

교통 지표	분석도구		비고
	명칭	설명	
경제 활동 자체 수용도	경제 활동 자체 수용도 분석	 <p>특별시·광역시에서 발생한 출근 통행량 중 지역 내 통행량이 차지하는 비율을 분석</p>	Light, Expert 모두 제공
출· 퇴근 통행 연결성	출· 퇴근 통행 연결성 분석	 <p>선택 지역에서 다른 지역으로 이동하는 출·퇴근 통행량을 분석</p>	Light 버전 한정
초과 통근량 ·통근 잠재력	직주 근접 분석 (신규)	 <p>통근비용을 최소화하는 이상적인 통행패턴과 현실의 통행패턴이 어느정도 차이 나는지를 분석</p>	Light 버전 한정

○ 연구, 정책 등에 자주 활용되는 통근(출근, 퇴근)시간 지표는 대시보드 및 월간 리포트 형태로 시각화하여 제공함

- ‘대시보드’를 통하여 교통지표를 종합적인 형태로 제공하고, 대시보드의 정보를 기반으로 매월 의미있는 정보를 추출하여 월간리포트로 발간함



〈그림 18〉 대시보드 및 월간리포트 (예시)

5. 결론 및 차년도 수행계획

가. 결론

- 본 사업에서는 모바일 통신데이터 구축 및 활용 목적을 고려하여 통행 DB 구축 형태를 설계하고 통행 DB를 구축하였음
- 국가교통조사사업의 주요 목적인 여객 O/D 추정 및 지역간·지역내 통행분포 분석을 목적으로 기점 정보, 도착 정보, 통행자 정보를 중심으로 기종점 통행 DB를 설계하고 구축하였음
 - 2022년 11월에 대한 일자별 기종점 통행 DB를 구축함
 - 2023년에 대한 월별 기종점 통행 DB를 구축함
- 이용자의 분석 및 활용 목적에 따른 데이터 가공 및 재생산이 가능한 체계를 마련하기 위해 개인별 통행궤적 DB를 구축하고 기술적 검토를 수행하였음
 - 2023년 10월 중 8일에 대한 개인별 통행궤적 DB를 구축함
 - 도보-비도보 구분 등 개인별 통행궤적 DB 관련 기술적 검토를 수행함
- 「개인정보보호법」, 「위치정보의 보호 및 이용 등에 관한 법률」에 저촉되지 않도록 통행 DB를 설계 및 구축하였음

- 또한, 모바일 통신데이터 기반의 통행 DB를 검증하고, 통행량 차이가 나타나는 지역, 구간에 대한 특성을 파악한 다음, 이를 보정할 수 있는 방법에 대하여 기술하였음
 - 모바일 통신데이터로 구축한 통행량이 실제 통행을 잘 반영하고 있는지 검증하기 위하여 조사자료, 사회경제지표, 실적자료와 상호 비교하여 검증하였음
 - 모바일 통신데이터는 통신사별 통신설비의 차이, 가공 알고리즘의 차이에 따라 구축되는 통행량의 차이가 크기 때문에, 3종의 모바일 통신데이터(KT, SKT 지라프, SKT 리트머스 데이터)를 수집하여 상호 비교함
 - 조사자료 중에서 통행 정보를 담고 있는 통행실태조사(2019년, 2021년)와 인구총조사를 수집하여 모바일 통신데이터와 비교 분석함
 - 전체 통행, 통행목적별(통근, 통학, 귀가, 기타) 통행에 대하여 통행 유발, 유인과 영향이 있을 것으로 예상되는 사회경제지표(인구수, 학령인구수 등)와 비교함
 - 또한, 조사자료도 추정치이기에 참값으로 볼 수 있는 실적자료를 취득하여 모바일 통신데이터에서 추출한 통행량과 비교함
 - 의사결정나무분석을 통해 통행량 차이가 나타나는 구간의 특성을 제시하고, 분석 결과를 토대로 보정기준과 보정 방법을 제시하였음
 - 전체 통행, 통행목적별(통근, 통학, 귀가, 기타)로 데이터간 통행량 차이가 나타나는 유형을 제시하고, 의사결정나무분석 결과에서 나온 절단값을 기준으로 통행량을 보정할 수 있는 방법을 제시하였음
 - 향후, 이러한 보정 대상과 가중치를 이용하여 각 데이터 간 산술평균 또는 가중평균하여 구축된 기종점 통행량의 신뢰성을 개선할 수 있을 것으로 판단함
- 아울러, 2022년 사업에서 구축한 KT 모바일 통신데이터를 활용하여 다양한 교통지표를 구축하고, 온라인 서비스의 분석도구, 대시보드, 월간 리포트 등을 통해 구축한 교통지표를 시각화하여 제공하였음
 - 기존 사업에서 개발한 교통지표와 새롭게 개발한 지표를 포함하여 다양한 교통지표를 구축함 (통행량, 통행시간, 통행거리, 공동화지수, 직주균형지표 등)
 - 일반인을 대상으로 개발한 Light 버전의 분석도구, 전문가를 대상으로 개발한 Expert 버전의 분석도구의 교통지표를 모두 2022년 기준으로 현행화함
 - 사용자, 관리자 편의를 고려하여 데이터 다운로드 기능을 업데이트하였음

나. 차년도 수행계획

- 시의성 있는 통행 정보를 제공하기 위하여 2024년 기준의 모바일 통신데이터를 구축할 예정이며, 구축 결과에 대한 점검을 통해 이상치 여부를 파악할 예정임
 - 본 사업의 활용 목적, 통신사의 데이터 특성을 고려하여 통행 DB 구조를 설계하고 이에 따라 통행 DB를 생성할 예정임
- 본 사업을 통해 구축한 2023년 기준의 모바일 통신데이터를 기준으로 본 사업에서 진행한 것과 같이 통행량 및 통행패턴에 대한 데이터간 상호 교차 검증을 수행할 예정임
 - 시공간별·목적별 개인통행실태조사, 인구총조사, 3개 통신사 자료, 차량 GPS, 대중교통카드, TCS, 사회경제지표 등의 자료를 대상으로 교차 검증할 예정임
- 모바일 통신데이터를 활용하여 통행지표를 산출 및 현행화할 예정임
 - 개별 객체 단위 모바일 통신데이터를 활용한 통행지표 산출 방안을 검토할 예정임
 - 2023년 모바일 통신데이터를 이용하여 교통지표(통행량, 통행시간 등)를 현행화할 예정임



제1장 사업의 개요

제1절 사업의 배경 및 목적

제2절 사업의 범위 및 내용

제1장 사업의 개요

제1절 사업의 배경 및 목적

1. 사업의 배경

- 모바일통신 빅데이터는 교통 분야에서 활용 중인 기존 빅데이터보다 표본율이 높고, 개별 통행정보가 상세하여 활용 가치가 높음
 - 전 국민의 95% 이상이 모바일 기기를 사용하고 있고, 기기의 전원을 끄지 않는 이상 개인의 이동궤적이 실시간으로 기록되어 데이터의 시·공간적 해상도가 높으며, 인적 정보가 담겨있어 사회 현상 분석에도 용이함
- 또한, 기존 인력식 조사 방식에 비해 경제적이고 효율적으로 통행정보를 취득할 수 있으며, 수집·분석 시간이 단축되어 시의성 있는 통행정보를 제공할 수 있음

2. 사업의 목적

- 모바일 기기를 통해 수집되는 모바일통신 빅데이터를 이용한 통행사슬(Trip Chain) DB를 구축하여 여객 및 화물OD 구축을 지원하고, 각종 교통관련 지표를 개발하여 교통정책 수립을 지원하고자 함

제2절 사업의 범위 및 내용

1. 사업의 범위

가. 모바일 통신데이터 기반 통행 DB 구축 및 결과 점검

- 시간적 범위: SKT 리트머스 데이터 2022년 11월 (1개월), 2023년 (1년)
- 공간적 범위: 전국

나. 모바일 통신데이터 기반 통행 DB 검증

- 시간적 범위
 - ① 모바일 통신데이터
 - KT 데이터: 2022년 11월 (1개월)
 - SKT 리트머스 데이터: 2022년 11월 (1개월)
 - SKT 지라프 데이터: 2022년 11월 (1개월)
 - ② 조사자료
 - 통행실태조사: 2019년 가구통행실태조사 (통신데이터가 활용되지 않은 조사 중 가장 최신의 자료), 2021년 개인통행실태조사 (통신데이터가 활용된 자료)
 - 인구총조사: 2020년
 - ③ 통행 유발 요인에 대한 자료
 - 인구수, 면적, 각종 사회경제지표, 기지국 자료를 모바일 통신데이터에 대한 해석이 용이하도록 2022년 자료로 수집함
 - 2022년에 대한 자료가 없는 경우 가장 최신의 자료를 활용함 (예: 종사자수와 같은 경우는 2020년 자료를 활용함)
- 공간적 범위: 전국

- 시도 단위, 시군구, 읍면동 단위로 구축함

다. 모바일 통신데이터 기반 교통지표 산출 및 제공

- 시간적 범위: 2022년 1월~12월 (1년)
 - 2022년 사업에서 구축한 KT 통신데이터 기반 통행 DB를 활용함
- 공간적 범위: 전국

2. 사업의 내용

가. 모바일 통신데이터 기반 통행 DB 구축 및 결과 점검

- 모바일 통신데이터 기반 통행 DB 구축
 - 본 사업의 활용 목적, 통신사의 데이터 특성에 따라 통행 DB의 구조를 설계하고 정의된 로직에 따라 2023년 통행 DB를 생성함
- 모바일 통신데이터 기반 통행 DB 결과 점검
 - 통행정보(채류지 유형, 통행량, 통행시간, 통행속도, 통행거리 등) 결과 점검 기준을 정립함
 - 정립된 기준에 따른 통행DB 구축 결과를 점검함

나. 모바일 통신데이터 기반 통행 DB 검증

- 모바일 통신데이터 검증
 - (데이터간 상호 비교를 통한 검증) 데이터 6종(모바일 통신데이터 3종, 조사자료 3종)의 통행량, 통행분포를 상호 비교함
 - 시도, 시군구, 읍면동 단위로 전체, 통행목적별 발생량, 도착량을 각각 비교함
 - 시군구 단위로 최다 통행구간을 비교함
 - (통행 발생·도착 요인을 기준으로 검증) 통행 발생·도착 요인(인구수, 수용학 생수, 학령인구수 등)의 분포와 수집한 데이터 6종(모바일 통신데이터 3종, 조

사자료 3종)의 통행량 분포가 유사한지 확인함

- 시군구, 읍면동 단위로 전체, 통행목적별로 발생량, 도착량을 각각 비교함
- (실적자료를 기준으로 검증) 실적자료(대중교통카드자료, 수송실적 자료 등)의 통행량과 모바일 통신데이터의 통행량을 비교 분석함

○ 모바일 통신데이터 보정 방법 개발

- 타 데이터와 통행량 차이가 발생하는 구간에 대하여 특성 및 주요 요인을 분석함
- 모바일 통신데이터의 통행량 보정 대상(유형) 및 기준을 제시함
- 보정 통행량 산출 방법 및 적용 방법을 제시함

다. 모바일 통신데이터 기반 교통지표 산출 및 제공

○ 교통지표 현행화

- 기존 사업에서 개발한 교통지표를 2022년 사업에서 구축한 KT 통신데이터 기반 통행 DB(2022년, 1년치)를 활용하여 업데이트함

○ 신규 교통지표 개발

- 문헌고찰을 통해 모바일 통신데이터에서 산출 가능한 교통지표를 발굴함

○ 데이터 개방 및 활용 환경 제공

- 모바일 통신데이터를 기반으로 교통지표별 DB를 구축함
- 모바일 통신데이터 및 교통지표 DB를 온라인 서비스 내 분석도구를 통하여 제공함
- 대시보드 및 분석 리포트를 통해 일부 교통지표의 정보를 제공함



제2장 모바일 통신데이터 기반 통행 DB 구축 및 결과 점검

제1절 모바일 통신데이터 기반 통행 DB 구축 개요

제2절 모바일 통신데이터 기반 기종점 통행 DB 구축

제3절 객체단위 모바일 통신데이터 가공 및 통행 DB 구축

제2장 모바일 통신데이터 기반 통행 DB 구축 및 결과 점검

제1절 모바일 통신데이터 기반 통행 DB 구축 개요

1. 모바일 통신데이터 활용목적을 고려한 통행 DB 구축 형태 검토

- 본 사업 내 모바일 통신데이터 구축 및 활용 목적을 고려하여 통행 DB 구축 형태를 검토함
 - (여객O/D 구축 보완자료) 총 여객 O/D 추정 및 지역간·지역내 통행분포 분석을 위한 보완자료로 활용
 - (전국단위 교통DB 구축) 모바일 통신데이터 기반 사람의 통행을 파악할 수 있는 기종점 인구 통행량 DB 구축
 - (전국단위 통행지표 구축) 통행목적, 지역 특성, 인적 특성에 따른 전국 전역에 대한 사람 통행 현황 분석을 위한 지표 구축
 - 출·퇴근 통행시간, 경제활동 의존도, 경제활동 자체 수용도, 도심공동화 심각도 등
 - (데이터 개방 및 활용 환경) 통행 분석도구 및 시의성 있는 정책지원 통행 정보 제공
 - (분석도구) 핫플레이스 분석, 고령자 통행 분석, 내·외부 통행 비율 분석 등
 - (정책지원 통행정보) 통근·통학 모빌리티 모니터링 대시보드, 분석 리포트
- 원천 통행DB에 대한 구축 및 활용 확대가 필요함
 - 개별 객체의 이동에 대한 상세 정보를 포함하는 DB 구축 확대를 통해 이용자가 필요로 하는 신규 데이터 재생산 체계 마련 필요

- 차량 GPS 데이터의 경우 개별 차량별 미시적 통행정보를 포함하는 원시데이터를 구매하여 개별차량 경로 DB 구축 및 고도화된 통행지표 구축
- 원천 통행DB 기반 개별 객체의 이동 행태 및 특성을 고려한 성능평가지표를 구축하여 생활밀착형 정책 지원 필요
- 이와 같이, 장기적으로는 개인속성정보 및 개인 통행궤적을 모두 포함하여 다양한 목적으로 활용 가능한 원천 통행DB를 확보하여 구축할 필요가 있음
- 그러나, 개인속성정보 및 개인 통행궤적을 모두 포함하는 모바일통신 원천 통행 DB 확보는 현재 현실적으로 어려운 상황임
- 이에 관련 기관 및 기업과의 지속적인 협의와 관련 법령·지침 개정을 통한 원천 통행 DB 확보 근거 마련이 필요함
- 본 사업에서는 활용 목적에 따라 개인속성정보를 포함하는 개인 O/D 데이터와 개인별 통행궤적을 포함하는 데이터로 구분하여 확보 및 구축을 추진함
- (개인속성정보 포함 O/D) 여객 O/D 추정, 지역간·지역내 통행분포 분석을 위해 성별·연령 등 인적 특성을 포함하는 데이터 확보 추진
- (개인별 통행궤적) 개별 객체 단위의 미시적 통행 특성 분석 및 기존 통행 DB 미구축 수단 관련 분석을 위한 개인별 통행궤적 데이터 확보 추진

2. 통행 DB 구축 형태 검토 및 설계

- 현재 통신사에서 구매 및 확보 가능한 모바일 통신데이터의 공간적·시간적 범위 및 상세도를 고려하여 통행 DB 구축 형태를 설계함
- 기종점 통행 DB는 여객 O/D 추정 및 지역간·지역내 통행분포 분석을 목적으로 구축함
 - 전국 읍면동 단위의 보다 상세한 시간적 단위(일자별·시간대별)의 통행정보 분석을 위한 DB 구축
 - 월별 통행 변화 분석 및 평일·주말 O/D 분석을 위한 DB 구축
- 개인별 통행궤적 DB는 객체 단위의 상세 정보를 포함한 통행 DB를 구축하여 이용자의 분석 및 활용 목적에 따른 데이터 가공 및 재생산이 가능한 체계를 마련하기 위한 검토를 수행함
 - 객체 단위의 출발지, 도착지, 통행궤적, 통행수단, 통행목적 등 상세 정보 포함
 - 대중교통, 보행, 자전거, PM 등 기존 통행DB 미구축 수단에 대한 분석을 위한 통행 DB 설계
 - 향후 신교통수단 및 서비스에 대한 분석을 위한 통행DB 구축 확대

1. 통행 DB 구축 및 검토

- 기종점 통행 DB 구축 및 검토
 - 2022년 11월에 대한 일자별 기종점 통행 DB 구축 및 검토
 - 2023년에 대한 월별 기종점 통행 DB 구축 및 검토
- 개인별 통행궤적 DB 구축 및 검토
 - 2023년 10월 중 8일에 대한 개인별 통행궤적 DB 구축
 - 도보·비도보 구분 등 기술적 검토 수행

제2절 모바일 통신데이터 기반 기종점 통행 DB 구축

1. 모바일 통신데이터 기반 기종점 통행 DB 형태 설계

가. 기종점 통행 DB 설계 방향

- 기종점 통행 DB는 국가교통조사사업의 주요 목적인 여객 O/D 구축을 고려하여 기점 정보, 도착 정보, 통행자 정보를 중심으로 구축하며, 기종점간 통행량, 통행시간, 통행거리를 산출하여 구축함
- 「개인정보보호법」과 「위치정보의 보호 및 이용 등에 관한 법률」을 고려하여 개인이 식별되지 않도록 DB를 구축함
 - 출발지와 도착지 위치는 「위치정보의 보호 및 이용 등에 관한 법률」에 저촉되지 않도록 기지국 좌표 그대로 노출하지 않고 행정동 또는 시군구 단위로 집계함
 - 개인의 통행 정보가 노출되지 않도록 기점 정보(출발 일자, 출발 시간대, 출발 위치, 출발 체류지 유형), 도착 정보(도착일자, 도착 시간대, 도착 위치, 도착 체류지 유형), 통행자 정보(성, 연령)를 기준으로 통행 정보가 같은 인구를 집계
- 출발 시간 및 도착 시간 정보는 집계된 인구가 5명 이하가 될 가능성을 줄이면서, DB 사용자에게 가능한 섬세한 통행 정보를 제공할 수 있도록 한 시간 단위로 입력하도록 함
- 통행자에 대한 정보를 확인할 수 있도록 성, 연령 필드를 구성하고, 통행 정보가 지나치게 세분되지 않도록 가능한 통행 특성이 유사한 계층을 묶어 10세 단위로 연령을 구분하여 입력하도록 함
- 출발 및 도착 체류지 유형은 평소 통행자가 해당 위치에 얼마나 자주, 오래 머무르는지 추정할 수 있는 정보를 제공하기 위한 것으로, 「2019년 국가교통조사」에서 설정한 체류지 식별 기준에 따라 출발지(출발 기지국 좌표)와 도

착지(도착 기지국 좌표)의 특성을 확인한 후 각각 체류지 유형을 구분하여 입력하도록 함

- 체류지 유형을 집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 관광(5), 기타(4)로 구분하여 구축함

○ 통행량 필드에는 전술한 기준에 의해 집계된 인구(단말기 수)를 입력함

○ 통행시간 필드에는 집계된 통행 정보를 기준으로 산출한 평균 통행시간 값을 입력함

나. 일자별·시간대별 기종점 통행 DB 설계

○ 전국에 대한 일자별 기종점 통행 DB 설계

- DB 구축 대상 기간의 개별 일자에 대한 시간대별 출발-도착 읍면동간 총 통행량 및 통행시간 분포 분석을 위한 DB를 설계함

〈표 2-1〉 모바일통신 빅데이터 DB 형태 (전국 일자별, 시간별 읍면동 단위 통행 정보)

출발 정보							
일자	요일	시간대	읍면동	시군구	시도	체류지 유형*	
221101	2	10	11010530	11010	11	H	
도착 정보							
일자	요일	시간대	읍면동	시군구	시도	체류지 유형*	
221101	2	11	31091710	31091	3	0	
통행자 정보		통행 정보***					
성별	연령대**	총 통행량	통행시간			통행 거리 (평균)	통행 속도 (평균)
			평균	분산	중앙값		
M	30	6	4908	981.6	5040	44000	8.96

주: * 집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)

** 10세 간격

*** 시간은 sec 단위 / 거리는 m 단위 / 속도는 m/sec로 제공

〈표 2-2〉 모바일통신 빅데이터 DB 테이블 정의서 (전국 일자별, 시간별 읍면동 단위 통행 정보)

컬럼ID	컬럼명	비고
o_dt	출발 날짜	-
o_dow	출발 요일	1~7(월~일)
o_time	출발 시간대	0~23
o_adm	출발지 행정동	-
o_sigun	출발지 시군구	-
o_sido	출발지 시도	-
o_type	출발지 유형	집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)
d_dt	도착 날짜	-
d_time	도착 시간대	0~23
d_adm	도착지 행정동	-
d_sigun	도착지 시군구	-
d_sido	도착지 시도	-
d_type	도착지 유형	집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)
age	연령대	10세 간격
gender	성별	01:남자, 02:여자
tot_pop	총 통행량	-
move_time_avg	통행시간_평균	단위(sec)
move_time_var	통행시간_분산	-
move_time_mid	통행시간_중앙값	-
distance_avg	통행거리_평균	단위(m)
dspeed_avg	통행속도_평균	m/sec

다. 월별 기종점 통행 DB 설계

○ 전국에 대한 월별 기종점 통행 DB 설계

- DB 구축 대상 기간의 각 월에 대해 요일별 시간대별 출발-도착 시군구간 일 평균 통행량 분석을 위한 DB를 설계함

〈표 2-3〉 모바일통신 빅데이터 DB형태 (전국 월별, 요일별 시군구 단위 통행 정보)

출발 정보				
요일	시간대	시군구	시도	체류지 유형*
2	10	11010	11	H
도착 정보				
요일	시간대	시군구	시도	체류지 유형
2	11	31091	31	O
통행자 정보			통행 정보	
성별	연령대**		평균 통행량	
M	30		6	

주: * 집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)

** 10세 간격

〈표 2-4〉 모바일통신 빅데이터 DB 테이블 정의서 (전국 월별, 요일별 시군구 단위 통행 정보)

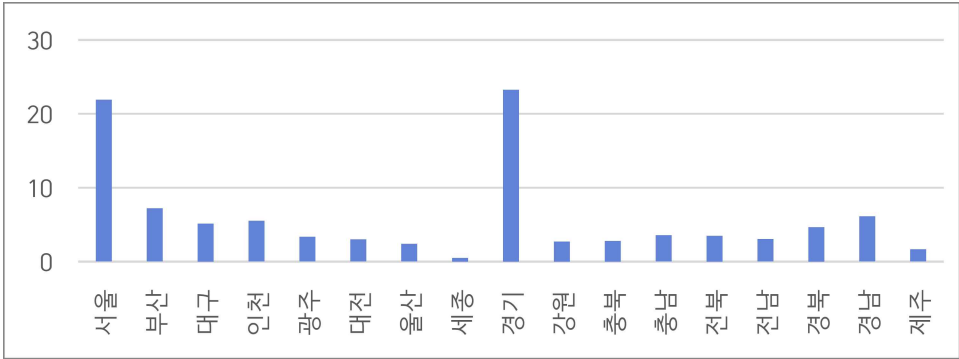
컬럼ID	컬럼명	비고
o_dow	출발요일	1~7(월~일)
o_time	출발시간대	0~23
o_sigun	출발지시군구	-
o_sido	출발지시도	-
o_type	출발지유형	집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)
d_time	도착시간대	0~23
d_sigun	도착지시군구	-
d_sido	도착지시도	-
d_type	도착지유형	집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)
age	연령대	10세 간격
gender	성별	01:남자, 02:여자
avg_pop	평균통행량	-

2. 2022년 11월 일자별 기종점 통행 DB 구축 결과 점검

가. 평일 일자별·시도별·목적별 통행량

1) 시도별 통행량

- 2022년 11월 평일 시도별 통행량은 경기, 서울, 부산, 경남 순으로 많이 발생하는 것으로 나타났으며, 세종의 통행량이 가장 적은 것으로 분석됨



<그림 2-1> 2022년 11월 평일 통행량 중 시도별 통행량 비율

<표 2-5> 2022년 11월 평일 시도별 일평균 통행량

(단위: 통행/일)

시도		통행량	비율
11	서울	17,387,538	21.9%
26	부산	5,712,809	7.2%
27	대구	4,070,377	5.1%
28	인천	4,360,340	5.5%
29	광주	2,638,012	3.3%
30	대전	2,372,647	3.0%
31	울산	1,881,847	2.4%
36	세종	384,047	0.5%
41	경기	18,473,382	23.2%
42	강원	2,142,061	2.7%
43	충북	2,221,802	2.8%
44	충남	2,809,927	3.5%
45	전북	2,762,028	3.5%
46	전남	2,408,760	3.0%
47	경북	3,696,598	4.7%
48	경남	4,856,141	6.1%
50	제주	1,294,038	1.6%
합계		79,472,354	100.0%

2) 목적별 통행량

- 출·도착지 유형(집, 회사, 학교, 기타 등)을 고려하여 4개 통행목적별(출근, 등교, 귀가, 기타) 통행량을 산출 및 분석함
 - 출근(집→회사), 등교(집→학교), 귀가(→집), 기타(→쇼핑(레저), 관광, 기타)
- 산출 결과, 기타, 귀가, 출근, 등교 순으로 통행량이 많은 것으로 나타남

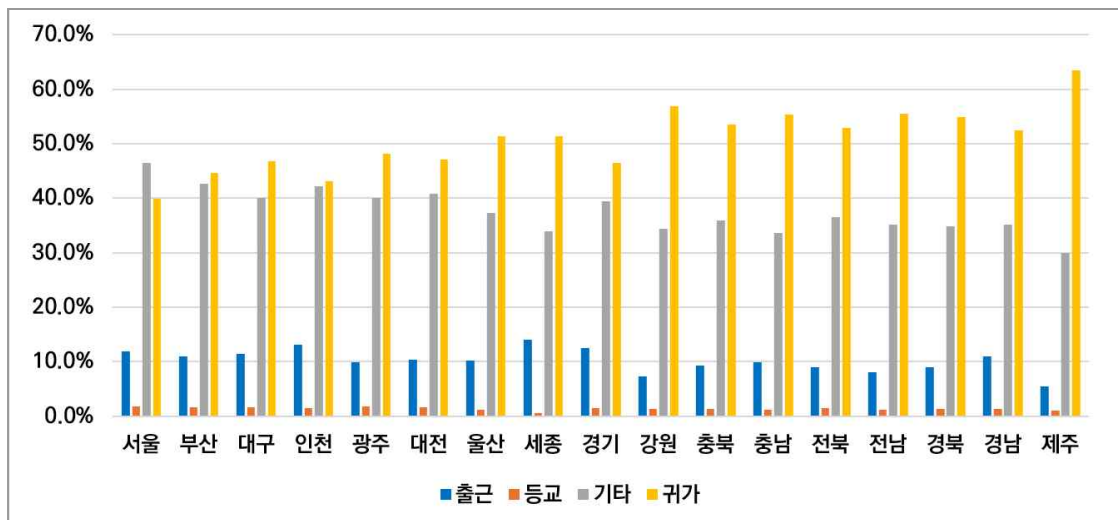
〈표 2-6〉 2022년 11월 평일 목적별 통행량

(단위: 통행/일)

목적	통행량	비율
출근	8,826,818	11.1%
등교	1,229,014	1.5%
귀가	31,849,026	40.1%
기타	37,567,496	47.3%
합계	79,472,354	100.0%

3) 시도별 목적별 통행량

- 전국에 대한 목적별 통행량과 같이 대부분의 시도에서 기타, 귀가, 출근, 등교 순으로 통행량이 많은 것으로 나타남
- 다만 서울에서는 기타 통행 대비 귀가 통행이 더 많이 발생한 것으로 분석됨



〈그림 2-2〉 2022년 11월 평일 시도별 총 통행량 중 목적별 통행량 비율

〈표 2-7〉 2022년 11월 평일 시도별 목적별 통행량

(단위: 통행/일)

시도	출근	등교	귀가	기타	합계
11 서울	2,065,946	320,182	8,074,332	6,927,078	17,387,538
26 부산	630,485	92,021	2,438,981	2,551,322	5,712,809
27 대구	463,065	69,185	1,630,808	1,907,320	4,070,377
28 인천	571,891	67,175	1,839,472	1,881,801	4,360,340
29 광주	261,095	47,465	1,057,907	1,271,545	2,638,012
30 대전	247,399	39,583	966,803	1,118,863	2,372,647
31 울산	193,562	21,794	700,823	965,669	1,881,847
36 세종	53,944	2,140	130,461	197,502	384,047
41 경기	2,317,792	272,741	7,292,170	8,590,679	18,473,382
42 강원	155,680	29,678	736,728	1,219,975	2,142,061
43 충북	208,043	29,272	796,429	1,188,058	2,221,802
44 충남	277,905	34,859	943,498	1,553,665	2,809,927
45 전북	246,810	43,236	1,010,352	1,461,630	2,762,028
46 전남	195,519	30,178	846,299	1,336,764	2,408,759
47 경북	330,612	49,231	1,289,206	2,027,550	3,696,598
48 경남	536,393	65,971	1,707,316	2,546,461	4,856,141
50 제주	70,676	14,302	387,444	821,616	1,294,038
합계	8,826,818	1,229,014	31,849,026	37,567,496	79,472,354

4) 일자별 목적별 통행량

- 출근의 경우 화요일에 가장 많은 통행이 발생하며, 수~금요일에는 통행량이 감소하는 패턴을 보임
- 등교의 경우 대체로 수요일의 통행량이 가장 많으며 금요일의 통행량이 가장 적은 패턴을 보임
 - 다만 대학수학능력시험이 시행된 11월 3주의 경우 타 주 대비 통행량이 상대적으로 적은 것으로 나타남
- 귀가의 경우 수요일까지는 통행량이 증가하다가 목요일, 금요일에는 통행량이 감소하는 패턴을 보임
- 기타 통행의 경우 월요일의 통행량이 가장 적고 금요일의 통행량이 가장 많은 것으로 나타남

〈표 2-8〉 2022년 11월 평일 일자별 목적별 통행량

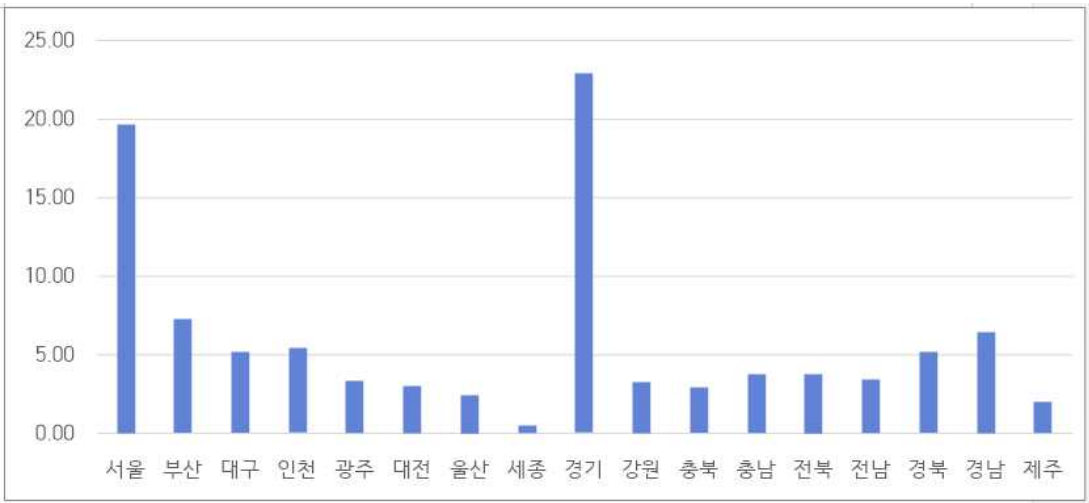
(단위: 통행/일)

일자(요일)	출근	등교	귀가	기타	합계
2022-11-01 (화)	9,353,326	1,105,515	32,171,477	37,213,928	79,844,245
2022-11-02 (수)	9,052,265	1,393,557	32,381,041	38,106,868	80,933,731
2022-11-03 (목)	8,975,001	1,358,573	32,001,315	37,591,630	79,926,520
2022-11-04 (금)	8,303,867	1,217,026	31,585,794	40,435,382	81,542,069
2022-11-07 (월)	8,832,370	1,377,555	31,451,927	36,026,645	77,688,497
2022-11-08 (화)	9,123,614	1,399,376	32,286,709	37,488,221	80,297,920
2022-11-09 (수)	9,100,299	1,410,362	32,475,236	37,806,218	80,792,114
2022-11-10 (목)	9,039,798	1,373,050	32,318,662	38,118,607	80,850,116
2022-11-11 (금)	8,293,486	1,239,095	32,013,120	41,399,347	82,945,048
2022-11-14 (월)	8,839,878	979,059	31,267,569	35,569,853	76,656,359
2022-11-15 (화)	9,020,009	992,160	31,715,556	36,608,998	78,336,723
2022-11-16 (수)	8,964,699	1,033,909	32,149,126	37,652,943	79,800,677
2022-11-17 (목)	8,719,132	756,059	31,923,789	38,312,717	79,711,697
2022-11-18 (금)	8,148,807	836,580	31,501,127	41,101,640	81,588,154
2022-11-21 (월)	8,840,674	1,334,918	31,369,224	35,753,092	77,297,909
2022-11-22 (화)	9,006,671	1,377,296	31,687,835	36,154,616	78,226,418
2022-11-23 (수)	8,994,316	1,358,402	32,066,126	37,091,497	79,510,341
2022-11-24 (목)	8,898,827	1,341,659	31,811,899	37,430,636	79,483,021
2022-11-25 (금)	8,183,472	1,191,735	31,991,899	40,508,917	81,876,023
2022-11-28 (월)	8,709,963	1,319,374	30,957,222	34,655,246	75,641,805
2022-11-29 (화)	8,844,431	1,310,677	31,908,691	35,618,498	77,682,298
2022-11-30 (수)	8,945,085	1,332,376	31,643,222	35,839,423	77,760,106

나. 주말 일자별·시도별·목적별 통행량

1) 시도별 통행량

- 2022년 11월 주말 시도별 통행량은 평일과 유사하게 경기, 서울, 부산, 경남 순으로 많이 발생하며 세종의 통행량이 가장 적은 것으로 나타남



<그림 2-3> 2022년 11월 주말 총 통행량 중 시도별 통행량 비율

<표 2-9> 2022년 11월 주말 시도별 통행량

(단위: 통행/일)

시도		통행량	비율
11	서울	12,328,306	19.6%
26	부산	4,537,959	7.2%
27	대구	3,227,361	5.1%
28	인천	3,372,513	5.4%
29	광주	2,086,184	3.3%
30	대전	1,889,061	3.0%
31	울산	1,524,980	2.4%
36	세종	291,242	0.5%
41	경기	14,342,455	22.8%
42	강원	2,038,239	3.2%
43	충북	1,807,458	2.9%
44	충남	2,339,363	3.7%
45	전북	2,362,371	3.8%
46	전남	2,132,863	3.4%
47	경북	3,222,023	5.1%
48	경남	4,029,008	6.4%
50	제주	1,244,461	2.0%
합계		62,775,845	100.0%

2) 목적별 통행량

- 평일과 동일한 방식으로 4개 통행목적별(출근, 등교, 귀가, 기타) 통행량을 산출 및 분석함
- 평일 대비 기타 통행의 비율이 상대적으로 높으며, 귀가 통행 비율은 평일과 유사한 것으로 나타남
- 다만, 통행량의 경우 평일 대비 전반적으로 적은 것으로 분석됨

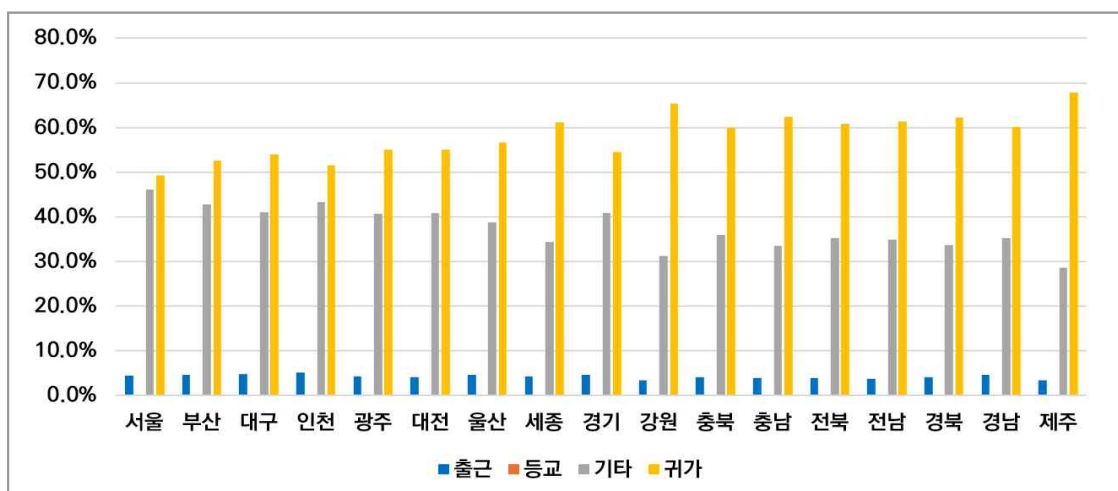
〈표 2-10〉 2022년 11월 주말 목적별 통행량

(단위: 통행/일)

목적	통행량	비율
출근	2,757,515	4.4%
등교	35,543	0.1%
귀가	25,116,430	40.0%
기타	34,866,357	55.5%
합계	62,775,846	100.0%

3) 시도별 목적별 통행량

- 모든 시도에서 기타, 귀가, 출근, 등교 순으로 통행량 비율이 높은 것으로 나타났으며, 평일 대비 기타 통행의 비율이 높은 것으로 분석됨
- 특히 제주, 강원, 기타 통행 비율이 높은 것으로 나타남



〈그림 2-4〉 2022년 11월 주말 시도별 총 통행량 중 목적별 통행량 비율

〈표 2-11〉 2022년 11월 주말 시도별 목적별 통행량

(단위: 통행/일)

시도	출근	등교	귀가	기타	총합계
11 서울	557,648	12,450	5,691,742	6,066,466	12,328,306
26 부산	206,824	3,000	1,942,323	2,385,812	4,537,958
27 대구	156,900	1,956	1,326,151	1,742,354	3,227,361
28 인천	172,010	1,239	1,459,854	1,739,410	3,372,513
29 광주	87,925	1,379	849,103	1,147,777	2,086,184
30 대전	78,480	1,361	770,713	1,038,508	1,889,061
31 울산	70,197	572	590,739	863,472	1,524,980
36 세종	12,681	38	100,329	178,193	291,242
41 경기	652,526	6,166	5,874,379	7,809,385	14,342,455
42 강원	69,156	650	636,588	1,331,845	2,038,239
43 충북	72,791	857	651,509	1,082,302	1,807,458
44 충남	93,297	705	784,455	1,460,906	2,339,363
45 전북	91,290	1,345	832,785	1,436,951	2,362,370
46 전남	79,261	472	743,602	1,309,529	2,132,863
47 경북	130,611	1,385	1,085,170	2,004,856	3,222,023
48 경남	183,694	1,614	1,419,791	2,423,910	4,029,008
50 제주	42,224	356	357,199	844,683	1,244,461
합계	2,757,515	35,543	25,116,430	34,866,357	62,775,845

4) 일자별 목적별 통행량

- 평일 대비 출근, 등교, 귀가 통행량은 현저히 낮은 수준으로 나타남
- 기타 통행의 경우, 토요일은 수요일, 목요일 수준으로 통행이 많은 반면 일요일의 기타 통행량은 일주일 중 가장 낮은 것으로 분석됨

〈표 2-12〉 2022년 11월 주말 일자별 목적별 통행량

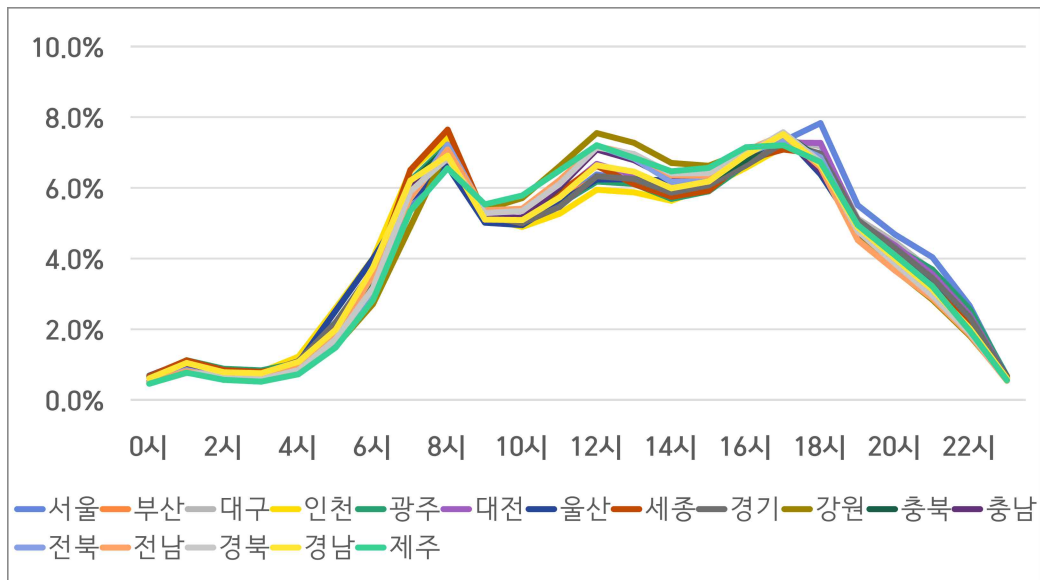
(단위: 통행/일)

일자(요일)	출근	등교	귀가	기타	합계
2022-11-05 (토)	3,433,451	39,848	26,099,578	38,922,354	68,495,231
2022-11-06 (일)	2,108,513	30,784	24,397,521	32,144,945	58,681,764
2022-11-12 (토)	3,411,686	40,180	26,026,690	37,993,957	67,472,513
2022-11-13 (일)	2,057,913	31,437	23,807,373	30,482,991	56,379,714
2022-11-19 (토)	3,437,673	40,082	26,219,052	39,139,451	68,836,259
2022-11-20 (일)	2,126,007	31,017	24,410,824	31,774,832	58,342,680
2022-11-26 (토)	3,410,514	40,096	25,957,150	37,971,269	67,379,028
2022-11-27 (일)	2,074,365	30,898	24,013,253	30,501,058	56,619,574

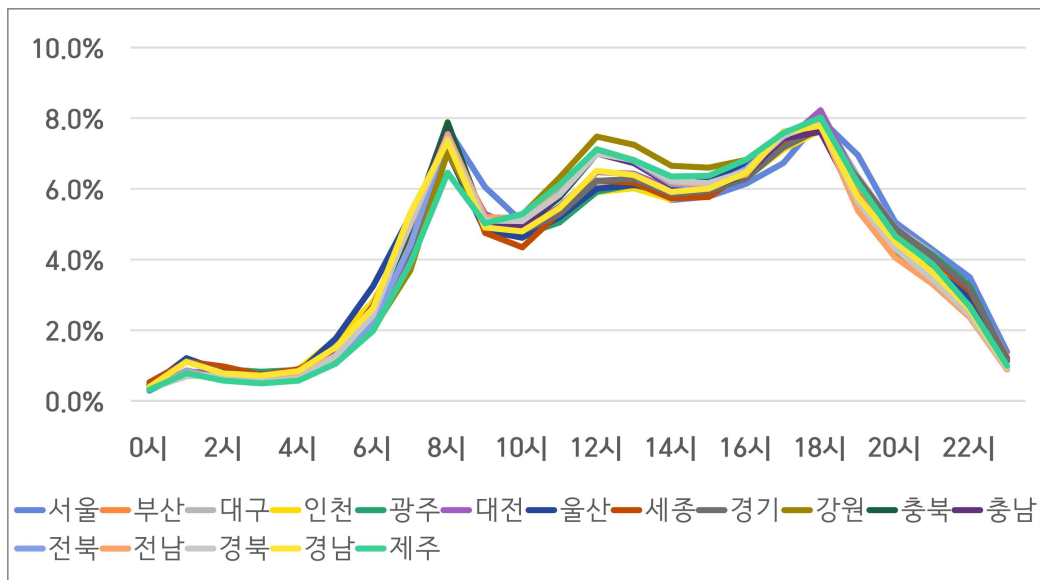
다. 시간대별 통행 발생량·도착량 패턴 분석

1) 시간대별 지역별 통행 발생량·도착량

- 일반적으로 오전 침두, 12시, 오후 침두에 타 시간대 대비 통행량이 많이 발생하는 패턴이 나타나며, 지역별로 시간대별 패턴에 큰 차이를 보이지 않음



〈그림 2-5〉 시간대별 지역별 통행 발생량 패턴



〈그림 2-6〉 시간대별 지역별 통행 도착량 패턴

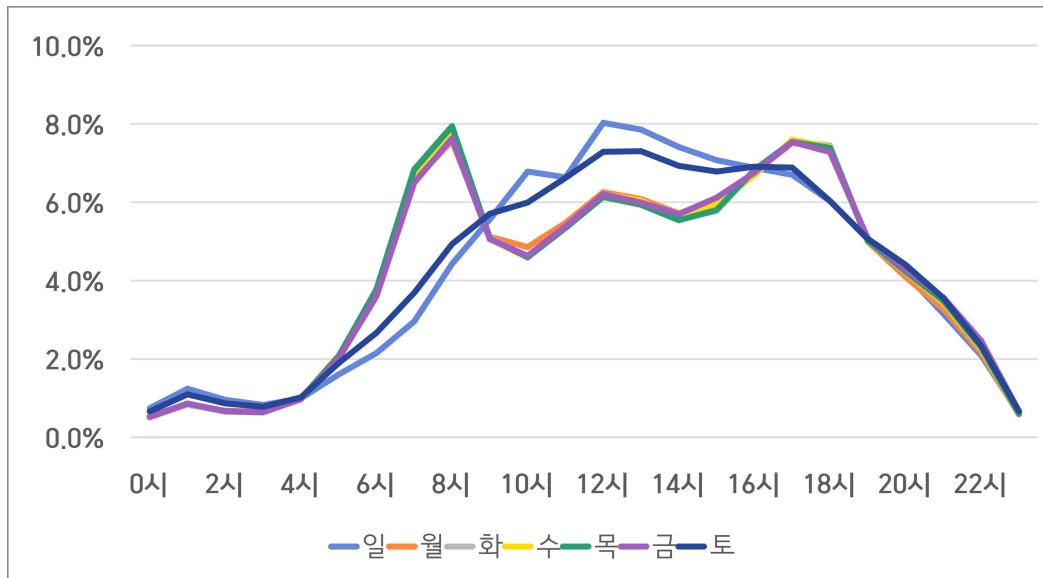
〈표 2-13〉 시간대별 통행량 패턴 분석

(단위: 통행/일)

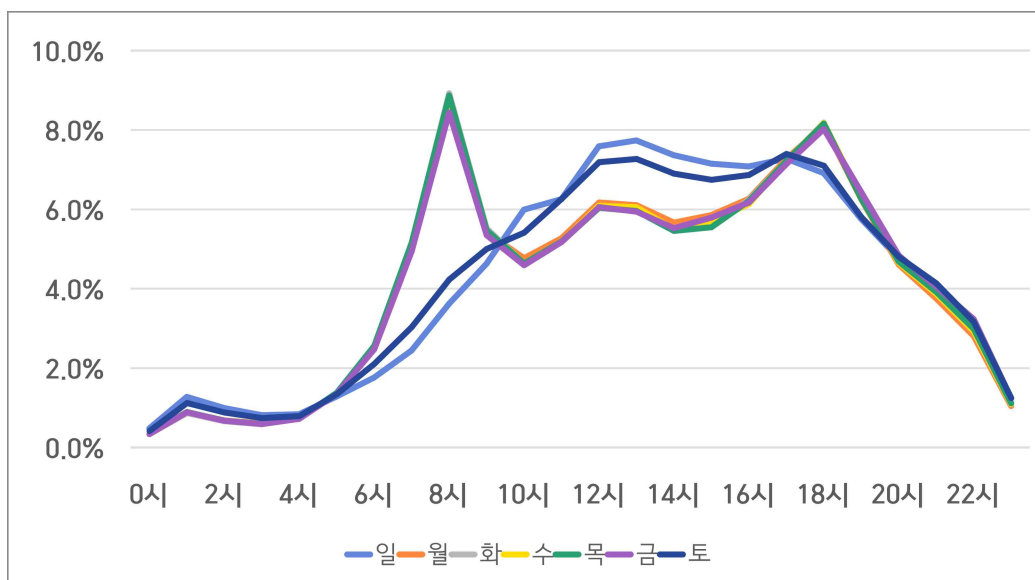
구분		새벽		오전		오후		저녁	
		0시~	3시~	6시~	9시~	12시~	15시~	18시~	21시~
발 생	서울	301,771	499,835	2,272,530	2,396,490	2,820,571	3,090,125	2,747,194	1,116,413
	부산	136,654	216,907	854,478	866,412	1,009,182	1,122,530	878,386	354,755
	대구	82,520	124,570	629,207	615,587	710,402	787,767	646,835	259,590
	인천	101,163	187,420	725,890	620,651	707,469	800,721	650,316	257,874
	광주	61,251	88,918	393,622	372,538	420,663	463,025	377,986	162,067
	대전	51,567	78,722	389,298	394,888	461,629	486,877	405,296	160,711
	울산	51,011	85,251	319,905	305,492	366,952	413,454	298,480	123,350
	세종	14,153	19,658	95,161	86,585	99,458	106,588	84,669	32,136
	경기	532,888	893,570	3,714,333	3,469,878	4,086,066	4,426,236	3,623,520	1,419,656
	강원	55,450	85,850	418,368	503,529	612,288	595,726	425,045	148,411
	충북	54,793	90,230	461,425	451,677	557,613	569,256	432,321	151,491
	충남	81,540	140,246	633,104	623,200	768,056	788,697	578,267	204,482
	전북	61,563	96,280	496,521	520,047	625,203	635,410	484,439	177,177
	전남	62,412	108,097	524,912	546,285	658,637	672,014	474,508	169,086
	경북	90,233	145,304	735,083	773,590	948,035	967,768	710,558	250,697
	경남	139,117	218,272	965,155	911,135	1,092,227	1,180,962	888,958	328,804
제주	27,916	42,424	229,113	277,098	319,387	325,580	245,392	89,101	
도 착	서울	271,913	359,270	2,181,125	2,619,617	2,838,740	2,959,450	3,170,028	1,454,182
	부산	128,110	166,090	746,892	836,588	997,680	1,088,418	1,003,782	445,974
	대구	77,510	94,276	512,599	573,968	693,406	765,181	729,993	321,413
	인천	95,233	132,972	599,329	575,121	685,446	754,013	722,317	327,202
	광주	57,205	70,544	321,416	338,514	407,846	456,720	433,952	195,469
	대전	48,256	59,736	327,331	374,790	453,027	476,934	459,588	198,502
	울산	47,955	64,868	305,006	287,341	354,673	404,006	348,249	153,179
	세종	13,427	15,774	78,143	74,512	93,667	100,506	97,602	43,132
	경기	503,623	654,212	3,073,200	3,289,131	3,965,112	4,187,428	4,080,865	1,826,270
	강원	51,814	69,842	371,092	483,454	621,254	608,785	505,122	194,723
	충북	50,867	70,216	424,915	438,894	553,556	555,105	505,528	193,082
	충남	76,113	108,540	594,611	611,795	768,631	771,273	680,491	263,665
	전북	57,476	76,324	440,789	501,310	622,130	624,975	564,028	220,896
	전남	57,386	84,593	496,710	540,000	660,560	660,811	566,276	214,927
	경북	84,019	116,198	704,311	758,081	947,756	954,382	842,698	322,151
	경남	130,794	176,687	877,438	870,620	1,080,710	1,152,550	1,037,666	414,654
제주	26,350	33,500	193,824	258,348	318,872	326,649	296,405	118,519	

2) 시간대별 요일별 통행 발생량·도착량

- 월요일~금요일의 경우, 오전 침두, 오후 침두, 12시 순으로 통행량이 많은 것으로 나타남
- 토요일, 일요일의 경우 12시 전후 시간대의 통행량이 가장 많은 것으로 나타남



〈그림 2-7〉 시간대별 요일별 발생량 패턴



〈그림 2-8〉 시간대별 요일별 도착량 패턴

〈표 2-14〉 요일별 통행량 패턴 분석

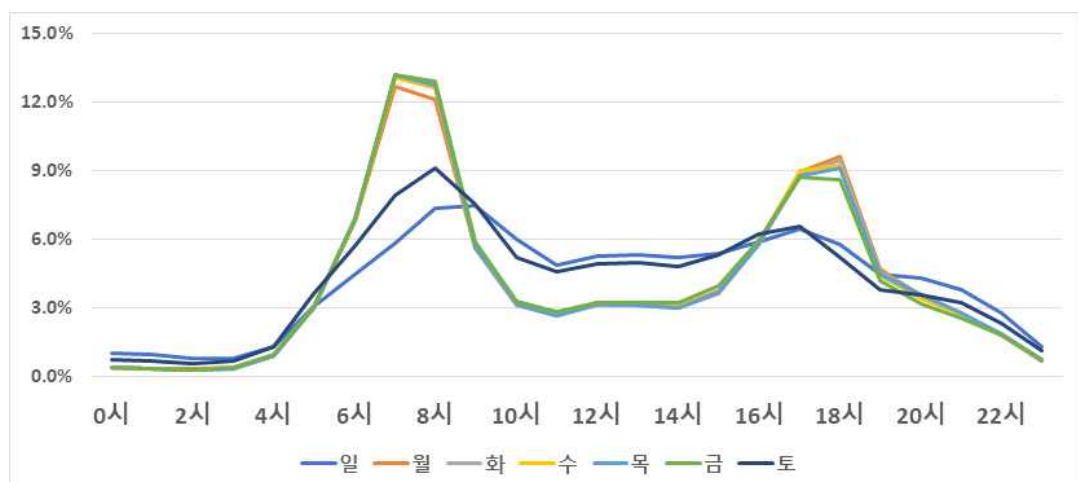
(단위: 통행/일)

구분		새벽		오전		오후		저녁	
		0시~	3시~	6시~	9시~	12시~	15시~	18시~	21시~
발생	월	249,478	456,709	2,178,518	1,886,072	2,204,086	2,479,939	2,011,199	736,277
	화	251,080	462,050	2,339,951	1,894,998	2,217,387	2,537,517	2,098,597	791,537
	수	258,542	464,225	2,320,644	1,899,830	2,254,970	2,564,674	2,105,807	798,318
	목	260,665	469,569	2,366,027	1,915,619	2,248,779	2,574,038	2,123,952	814,061
	금	265,010	468,421	2,302,139	1,957,608	2,324,788	2,651,475	2,160,812	868,243
	토	313,997	442,636	1,355,778	2,201,267	2,585,573	2,472,148	1,861,912	787,184
	일	307,230	357,946	995,048	1,979,688	2,428,255	2,152,945	1,589,891	610,179
도착	월	232,495	328,769	1,953,143	1,887,039	2,189,658	2,367,154	2,315,891	928,128
	화	234,927	338,379	2,094,728	1,930,211	2,198,736	2,393,144	2,410,773	992,218
	수	241,735	341,950	2,078,512	1,932,273	2,234,954	2,413,215	2,423,439	1,000,933
	목	243,603	344,580	2,117,570	1,952,277	2,228,725	2,422,641	2,439,516	1,023,798
	금	247,018	346,490	2,061,590	1,965,208	2,278,061	2,483,345	2,507,854	1,108,931
	토	291,073	346,448	1,127,110	2,005,524	2,567,771	2,525,533	2,128,935	1,028,102
	일	287,201	307,027	816,078	1,759,551	2,365,159	2,242,153	1,818,182	825,830

다. 시간대별 출·퇴근 통행량 패턴 분석

1) 시간대별 출·퇴근 통행량 (요일별)

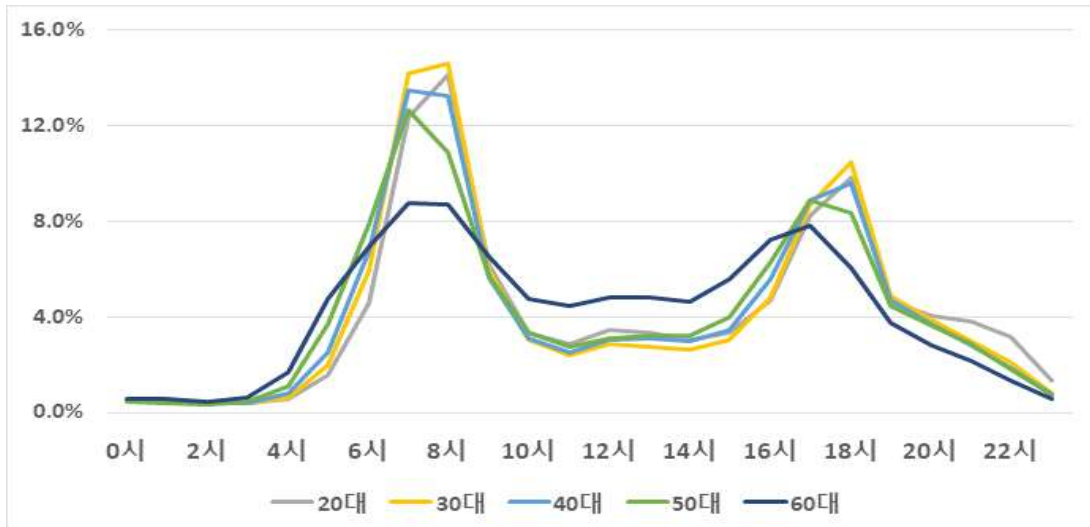
- 평일의 경우 오전 첨두와 오후 첨두에 타 시간대 대비 발생 비율이 높은 패턴이 나타나며, 오전 첨두의 통행 비율이 오후 첨두 대비 높은 것으로 분석됨
- 주말의 경우 오전 첨두·오후 첨두의 발생 비율이 높으나, 해당 시간대의 발생 비율은 타 평일 대비 낮은 것으로 분석됨



〈그림 2-9〉 시간대별 출·퇴근 통근 통행량 패턴 (요일별)

2) 시간대별 출·퇴근 통행량 (연령별)

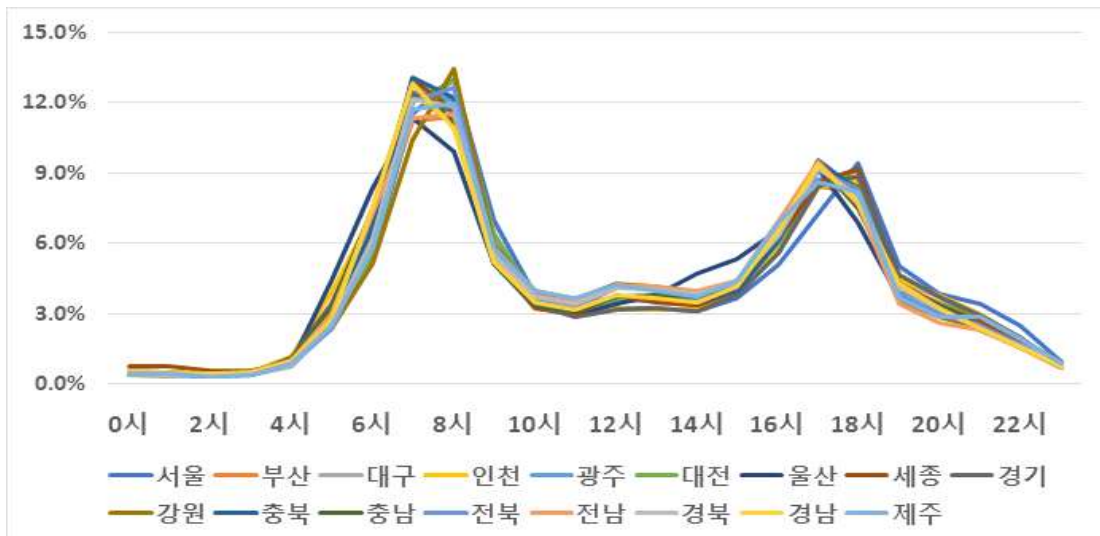
- 일반적으로 오전 첨두와 오후 첨두에 타 시간대 대비 발생 비율이 높은 패턴이 나타나며, 60대의 경우 타 연령대 대비 첨두시 발생 비율이 낮음



〈그림 2-10〉 시간대별 출·퇴근 통근 통행량 패턴 (연령별)

3) 시간대별 출·퇴근 통행량 (지역별)

- 일반적으로 오전 첨두와 오후 첨두에 타 시간대 대비 발생 비율이 높은 패턴이 나타나며, 지역별로 시간대별 패턴에 큰 차이를 보이지 않음

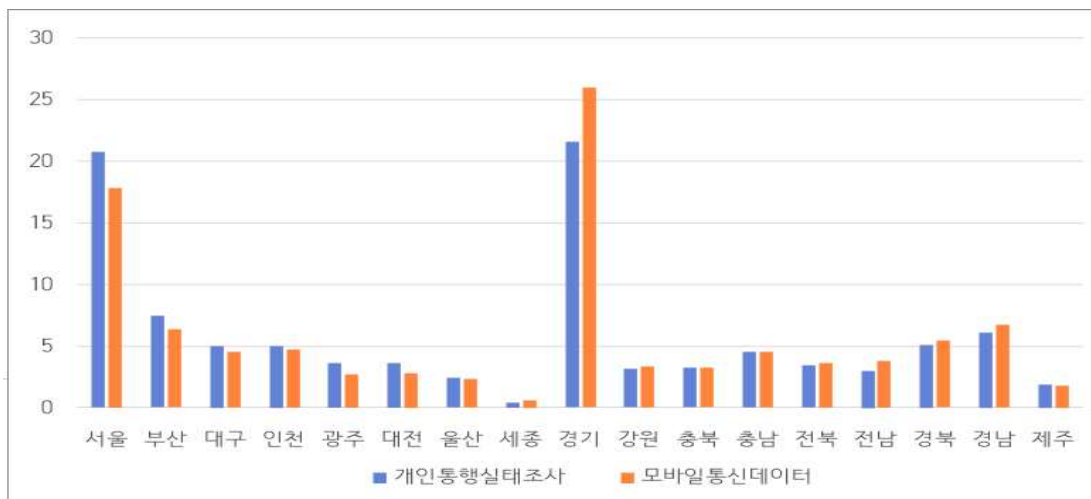


〈그림 2-11〉 시간대별 출·퇴근 통근 통행량 패턴 (지역별)

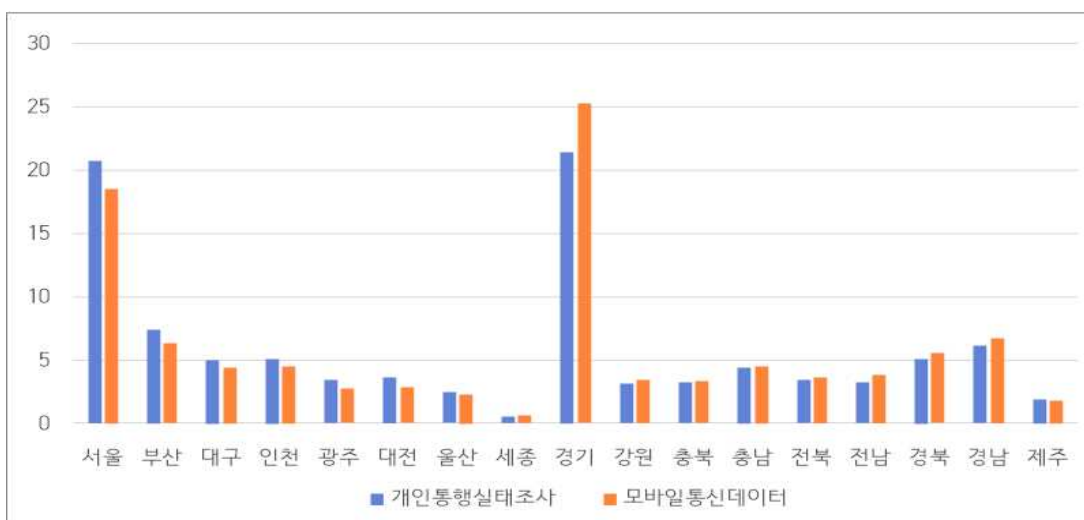
라. 통행발생·도착 패턴 비교·분석 결과

1) 지역별 통행 발생량 비교 결과

- 개인통행실태조사와 모바일 통신데이터 기반 지역별 통행 발생량 비교를 위해 각 자료별로 전체 통행 발생량 대비 각 시도별 발생량의 비율을 산출함
- 비교 결과, 서울 및 광역시, 세종, 제주, 충북, 충남의 경우 개인통행실태조사에 서의 발생 비율이 높으나, 타 지역의 경우 모바일 통신데이터의 비율이 높은 것으로 나타남



〈그림 2-12〉 지역별 통행 발생량 비교 결과 (발생)



〈그림 2-13〉 지역별 통행 발생량 비교 결과 (도착)

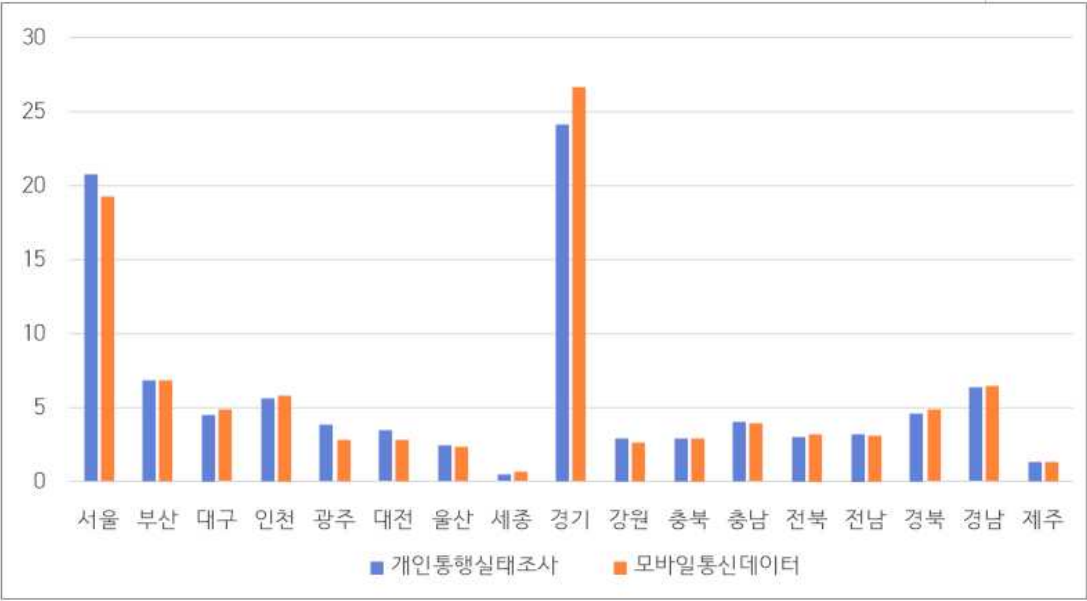
〈표 2-15〉 지역별 통행 발생·도착 패턴 비교 결과

(단위 : 통행/일, %)

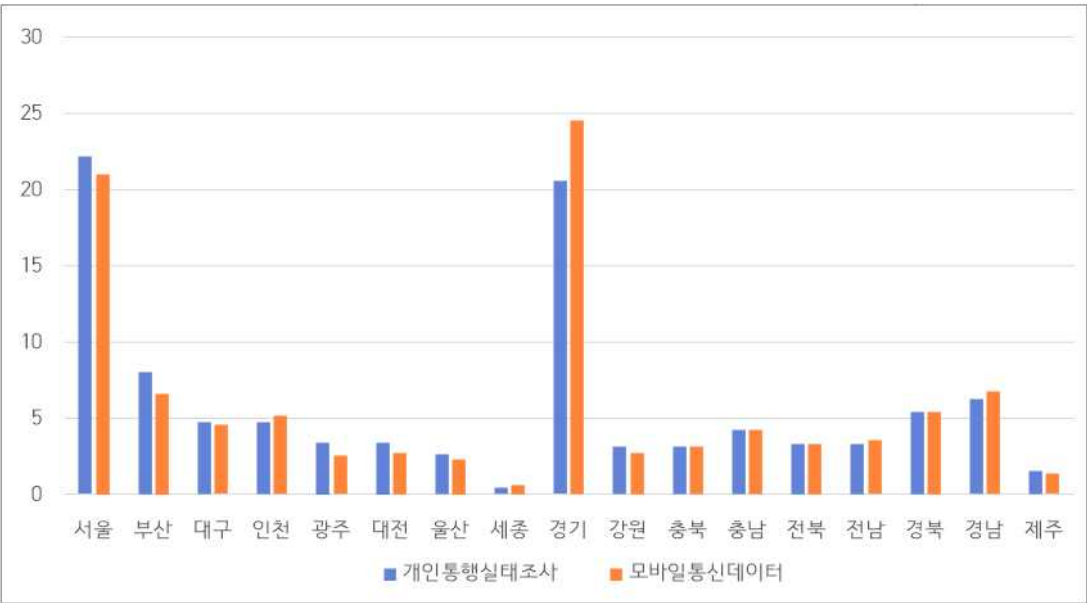
구분	개인통행실태조사 기준*		모바일통신 빅데이터 기준	
	발생 (비율)	도착 (비율)	발생 (비율)	도착 (비율)
서울	17,996,813(20.7)	17,971,878(20.7)	15,244,929(17.8)	15,854,325(18.5)
부산	6,457,538(7.4)	6,458,293(7.4)	5,439,304(6.3)	5,413,533(6.3)
대구	4,333,058(5.0)	4,330,603(5.0)	3,856,478(4.5)	3,768,346(4.4)
인천	4,370,318(5.0)	4,433,014(5.1)	4,051,505(4.7)	3,891,634(4.5)
광주	3,117,712(3.6)	2,942,575(3.4)	2,340,071(2.7)	2,281,667(2.7)
대전	3,091,862(3.6)	3,101,487(3.6)	2,428,989(2.8)	2,398,165(2.8)
울산	2,117,733(2.4)	2,109,519(2.4)	1,963,896(2.3)	1,965,278(2.3)
세종	389,864(0.4)	391,571(0.5)	538,406(0.6)	516,762(0.6)
경기	18,668,530(21.5)	18,628,440(21.4)	22,166,147(25.9)	21,579,840(25.2)
강원	2,714,015(3.1)	2,715,277(3.1)	2,844,666(3.3)	2,906,084(3.4)
충북	2,803,473(3.2)	2,793,476(3.2)	2,768,807(3.2)	2,792,163(3.3)
충남	3,871,861(4.5)	3,863,535(4.4)	3,817,592(4.5)	3,875,119(4.5)
전북	2,942,438(3.4)	2,943,452(3.4)	3,096,639(3.6)	3,107,928(3.6)
전남	2,606,282(3.0)	2,783,178(3.2)	3,215,951(3.8)	3,281,262(3.8)
경북	4,431,201(5.1)	4,442,218(5.1)	4,621,268(5.4)	4,729,595(5.5)
경남	5,297,159(6.1)	5,300,768(6.1)	5,724,630(6.7)	5,741,119(6.7)
제주	1,659,446(1.9)	1,660,019(1.9)	15,56,011(1.8)	1,572,467(1.8)
계	86,869,303(100)	86,869,303(100)	85,675,289(100)	85,675,287(100)

2) 지역별 출퇴근 통행량 비교 결과

- 개인통행실태조사와 모바일 통신데이터 기반 지역별 출퇴근 통행량에 대한 비교 결과는 지역별 전체 통행 발생량 비교 결과와 유사하게 나타남



<그림 2-14> 지역별 출근 통행량 비교 결과



<그림 2-15> 지역별 퇴근 통행량 비교 결과

〈표 2-16〉 지역별 출퇴근 통행량 비교 결과

(단위 : 통행/일, %)

구분	개인통행실태조사 기준		인구총조사 자료 기준	모바일통신 빅데이터 기준	
	출근 (비율)	귀가*	통근**	출근	퇴근
서울	4,516,466(20.7)	8,413,729(22.1)	4,597,322(19.6)	2,208,271(19.2)	2,183,315(21)
부산	1,492,300(6.8)	3,051,128(8.0)	1,500,101(6.4)	780,934(6.8)	687,264(6.6)
대구	974,423(4.5)	1,801,446(4.7)	1,064,227(4.5)	554,097(4.8)	462,558(4.5)
인천	1,213,843(5.6)	1,794,429(4.7)	1,344,311(5.7)	667,581(5.8)	530,206(5.1)
광주	838,601(3.8)	1,279,355(3.4)	639,412(2.7)	318,027(2.8)	261,230(2.5)
대전	736,843(3.4)	1,313,431(3.4)	672,856(2.9)	326,046(2.8)	280,836(2.7)
울산	514,493(2.4)	995,135(2.6)	538,654(2.3)	267,252(2.3)	241,670(2.3)
세종	89,203(0.4)	149,020(0.4)	92,403(0.4)	71,554(0.6)	63,317(0.6)
경기	5,272,517(24.1)	7,799,702(20.5)	5,745,634(24.4)	3,066,284(26.6)	2,538,255(24.5)
강원	644,578(2.9)	1,186,771(3.1)	703,266(3.0)	298,461(2.6)	277,706(2.7)
충북	637,895(2.9)	1,182,200(3.1)	742,399(3.2)	338,949(2.9)	324,180(3.1)
충남	875,177(4.0)	1,616,017(4.2)	1,032,913(4.4)	445,853(3.9)	435,901(4.2)
전북	665,597(3.0)	1,263,474(3.3)	811,899(3.5)	367,512(3.2)	340,853(3.3)
전남	706,911(3.2)	1,241,157(3.3)	864,498(3.7)	360,298(3.1)	361,776(3.5)
경북	1,007,670(4.6)	2,045,073(5.4)	1,280,135(5.4)	548,660(4.8)	558,259(5.4)
경남	1,372,291(6.3)	2,361,671(6.2)	1,560,582(6.6)	742,013(6.4)	692,468(6.7)
제주	291,636(1.3)	581,151(1.5)	309,133(1.3)	148,425(1.3)	135,968(1.3)
계	21,850,443(100)	38,074,889(100)	23,499,745(100)	11,510,217(100)	10,375,762(100)

주 : *개인통행실태조사자료는 2019년 기준 출근·귀가 통행에 대한 일평균시간이며, 귀가통행 통행목적에 상관없이 집으로 향하는 모든 통행을 의미함

**인구총조사(통계청)자료는 2020년도 기준이며, 출근과 퇴근이 구분되어 있지 않음

3. 2023년 월별 기종점 통행 DB 구축 결과 점검

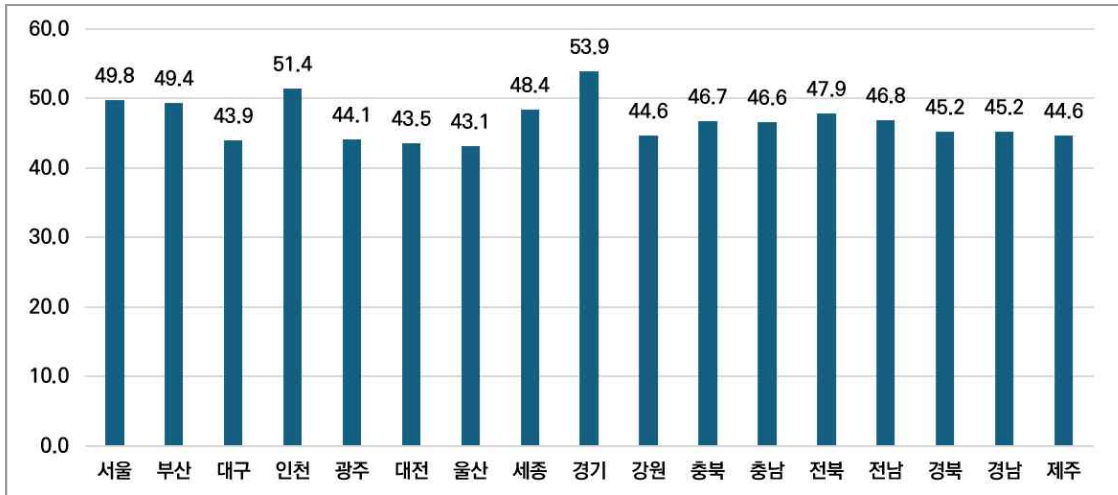
가. 평일 시도별·목적별 통행시간 분석

- 평일 출근 목적 통행의 경우 수도권, 특히 경기·인천의 통행시간이 타 시도 대비 길게 소요되는 것으로 나타남
- 평일 등교 목적 통행의 경우 세종, 경기, 인천 순으로 통행시간이 긴 것으로 나타남
- 평일 귀가 목적 통행의 경우 서울 및 광역시에 비해 도 지역의 통행시간이 긴 것으로 나타남
- 평일 기타 목적 통행의 경우 타 목적 통행에 비해 통행시간이 오래 소요되는 것으로 분석되었으며, 귀가 목적 통행과 유사하게 서울 및 광역시 대비 도 지역의 통행시간이 긴 것으로 나타남

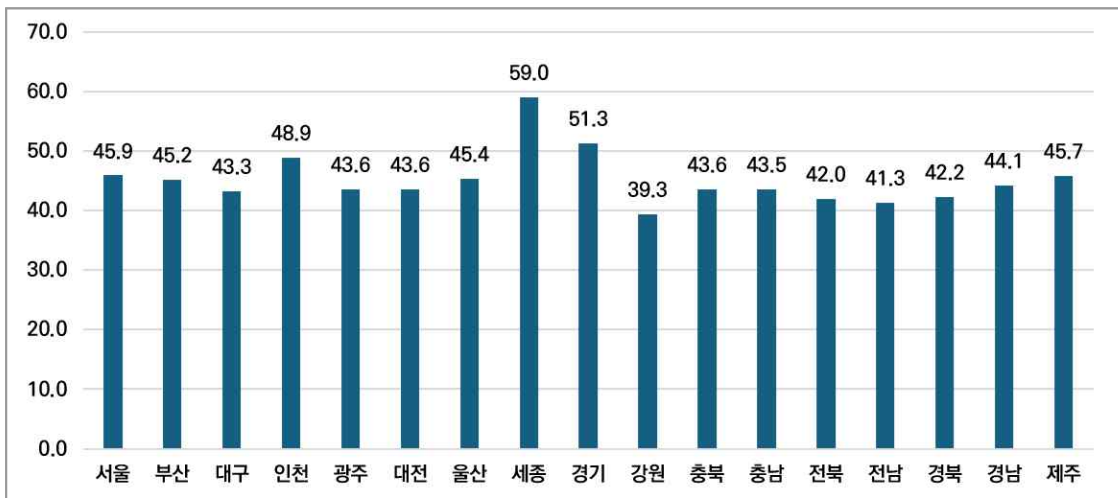
〈표 2-17〉 2023년 평일 시도별 목적별 통행시간

(단위: 분)

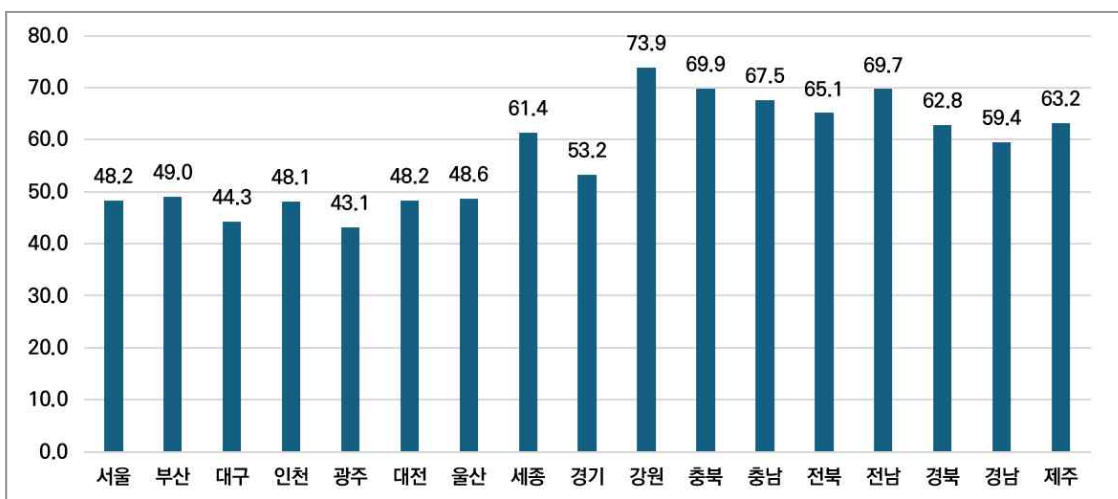
시도	출근	등교	귀가	기타
서울	49.8	45.9	48.2	69.6
부산	49.4	45.2	49.0	70.3
대구	43.9	43.3	44.3	68.6
인천	51.4	48.9	48.1	73.9
광주	44.1	43.6	43.1	69.5
대전	43.5	43.6	48.2	71.8
울산	43.1	45.4	48.6	71.6
세종	48.4	59.0	61.4	89.5
경기	53.9	51.3	53.2	78.4
강원	44.6	39.3	73.9	82.9
충북	46.7	43.6	69.9	84.3
충남	46.6	43.5	67.5	81.4
전북	47.9	42.0	65.1	80.4
전남	46.8	41.3	69.7	82.9
경북	45.2	42.2	62.8	78.9
경남	45.2	44.1	59.4	75.6
제주	44.6	45.7	63.2	70.4



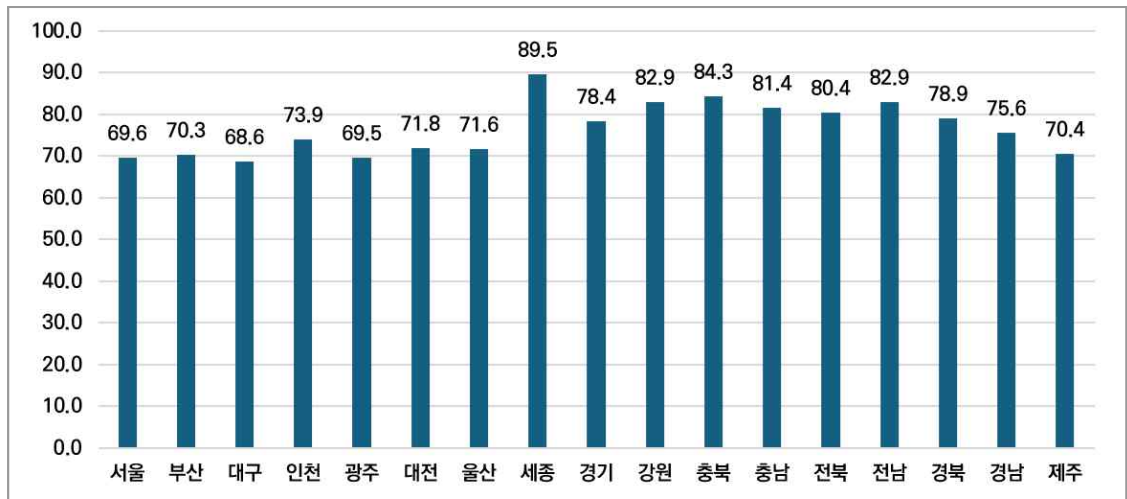
〈그림 2-16〉 시도별 평일 출근 통행시간



〈그림 2-17〉 시도별 평일 등교 통행시간



〈그림 2-18〉 시도별 평일 귀가 통행시간



〈그림 2-19〉 시도별 평일 기타 통행시간

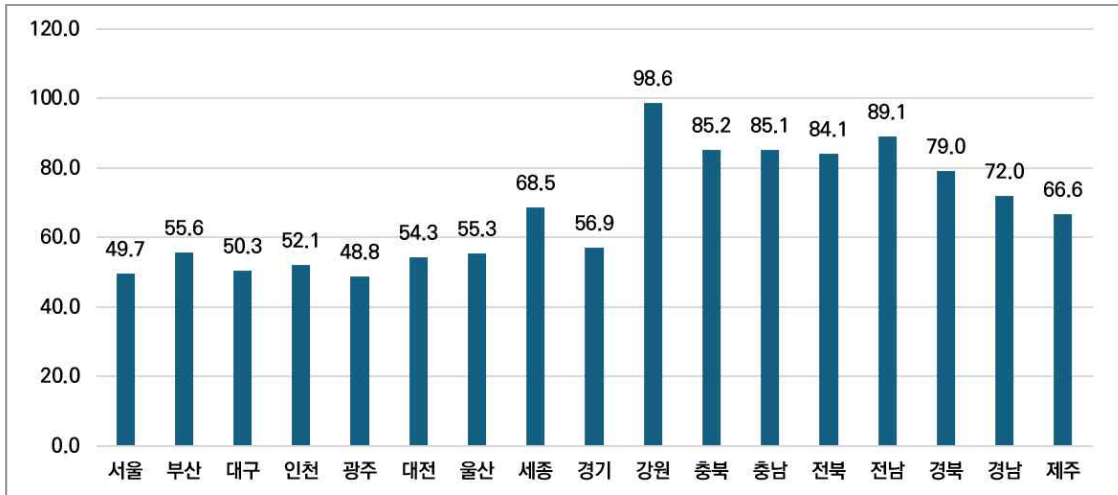
나. 주말 통행시간 분석

- 주말 귀가 및 기타 목적 통행의 경우 서울 및 광역시 대비 도 지역의 통행시간이 긴 것으로 나타남
- 서울 및 수도권, 광역시, 경남, 제주에서는 기타 목적 통행시간이 귀가 통행시간 대비 긴 것으로 분석되었으나, 강원, 충북, 충남, 전북, 전남 지역에서는 귀가 통행시간이 기타 목적 통행시간 대비 긴 것으로 나타남

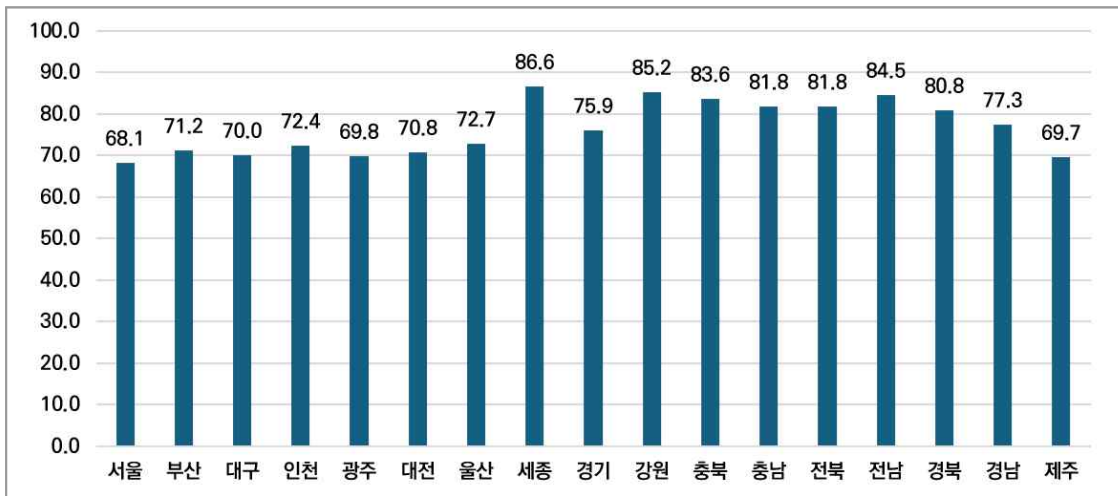
〈표 2-18〉 2023년 주말 시도별 목적별 통행시간

(단위: 분)

시도	출근	등교	귀가	기타
서울	50.6	46.2	49.7	68.1
부산	50.9	45.6	55.6	71.2
대구	45.6	43.9	50.3	70.0
인천	52.4	49.3	52.1	72.4
광주	45.8	44.5	48.8	69.8
대전	45.0	44.2	54.3	70.8
울산	44.8	46.2	55.3	72.7
세종	50.2	58.9	68.5	86.6
경기	54.9	51.6	56.9	75.9
강원	46.1	40.2	98.6	85.2
충북	48.2	44.3	85.2	83.6
충남	48.2	44.4	85.1	81.8
전북	49.5	43.1	84.1	81.8
전남	48.6	42.6	89.1	84.5
경북	46.9	43.2	79.0	80.8
경남	47.0	45.1	72.0	77.3
제주	45.3	47.8	66.6	69.7



〈그림 2-20〉 시도별 주말 귀가 통행시간



〈그림 2-21〉 시도별 주말 기타 통행시간

제3절 객체단위 모바일 통신데이터 가공 및 통행 DB 구축

1. 개요

가. 객체단위 DB 구축 목적

- 모바일 통신데이터는 휴대전화 단말기를 소지한 개인별 속성 및 통행궤적에 대한 상세한 정보가 수집되어 교통 및 관련 분야에서 가치가 높음
- 그러나, 현재까지 모바일 통신 데이터는 집계 수준에서 활용하여 교통분석에서 가치가 높은 개인의 통행에 관한 정보 활용에 한계가 존재함
- 이에, 모바일 원시데이터 기반 개별 객체의 이동에 대한 상세 정보를 포함하는 원천 통행DB 구축 확대를 통해 이용자가 필요로 하는 신규 데이터 재생산 체계 마련이 필요함
 - 객체 단위의 출발지, 도착지, 통행궤적, 통행수단, 통행목적 등 상세 통행정보를 포함한 통행DB 구축
 - 대중교통, 보행, 자전거, PM 등 기존 통행DB 미구축 수단과 신교통수단 및 서비스에 대한 통행DB 구축 확대
- 원천 통행DB 기반 개별 객체의 이동 행태 및 특성을 고려하여 기존 성능평가지표 구축 방법 고도화 및 신규 성능평가지표 발굴·구축이 필요함
 - 급격한 사회 변화 대응 및 국민이 체감하는 교통 서비스 개선 등 다양한 정책적 목표 지원 중심의 교통지표 구축 및 발굴
 - 자율주행, MaaS, 도심항공교통(UAM) 등 미래 기술 및 서비스 개발에 대비한 고도화된 지표 발굴 및 구축
- 궁극적으로, 신규 모빌리티 서비스 분석과 모빌리티 영향 평가 및 의사결정 지원을 위한 데이터 기반 교통정책 지원 체계 구축이 필요함

나. 객체단위 DB 구축 범위

- 본 사업에서는 현재 확보 가능한 개별 객체단위 모바일 통신 빅데이터에 대한 전처리 및 가공, 통행경로 및 통행지표 DB 구축 등 기술적 검토를 수행함
- DB 구축을 위한 객체단위 데이터는 본원에서 수행 중인 타 사업과 연계하여 확보한 개별 통행별 통행 궤적 데이터를 활용함
 - 공간적 범위: 전국
 - 시간적 범위: 2023. 10. 13. (금) ~ 2023. 10. 20. (금)
 - 통신사: LG유플러스
- 본 사업에서 수행한 기술적 검토 사항은 아래와 같음
 - 객체 단위 모바일 통신데이터 오류 검토 및 필터링
 - 모바일 통신데이터 기반 개별 통행경로 DB 구축
 - 통행량 및 통행속도 산출
 - 도로 통행 분석 관련 기술적 검토
 - 모바일 통신데이터 기반 통행지표 구축 관련 기술적 검토

2. 객체단위 모바일 통신데이터 기초분석

가. 객체단위 DB 개요 및 데이터 형태

- LG 유플러스 모바일 통신 원시 데이터는 총 10개의 컬럼으로 구성되어 있으며 데이터 발생일자, 트립ID, 출발격자ID, 출발시간, 도착격자ID, 도착시간, 이동거리, 속도, 출발체류지유형, 도착체류지유형으로 구성되어 있음
- LG 유플러스 모바일 통신 원시 데이터의 공간단위는 격자단위로 50m 단위로 구분되어 있으며, 시간단위는 10분 단위로 구분되어 있음
- 이동거리는 km 단위이며, 속도는 km/h 단위로 되어 있음

<표 2-19> LG 유플러스 모바일 통신 원시데이터 테이블정의서

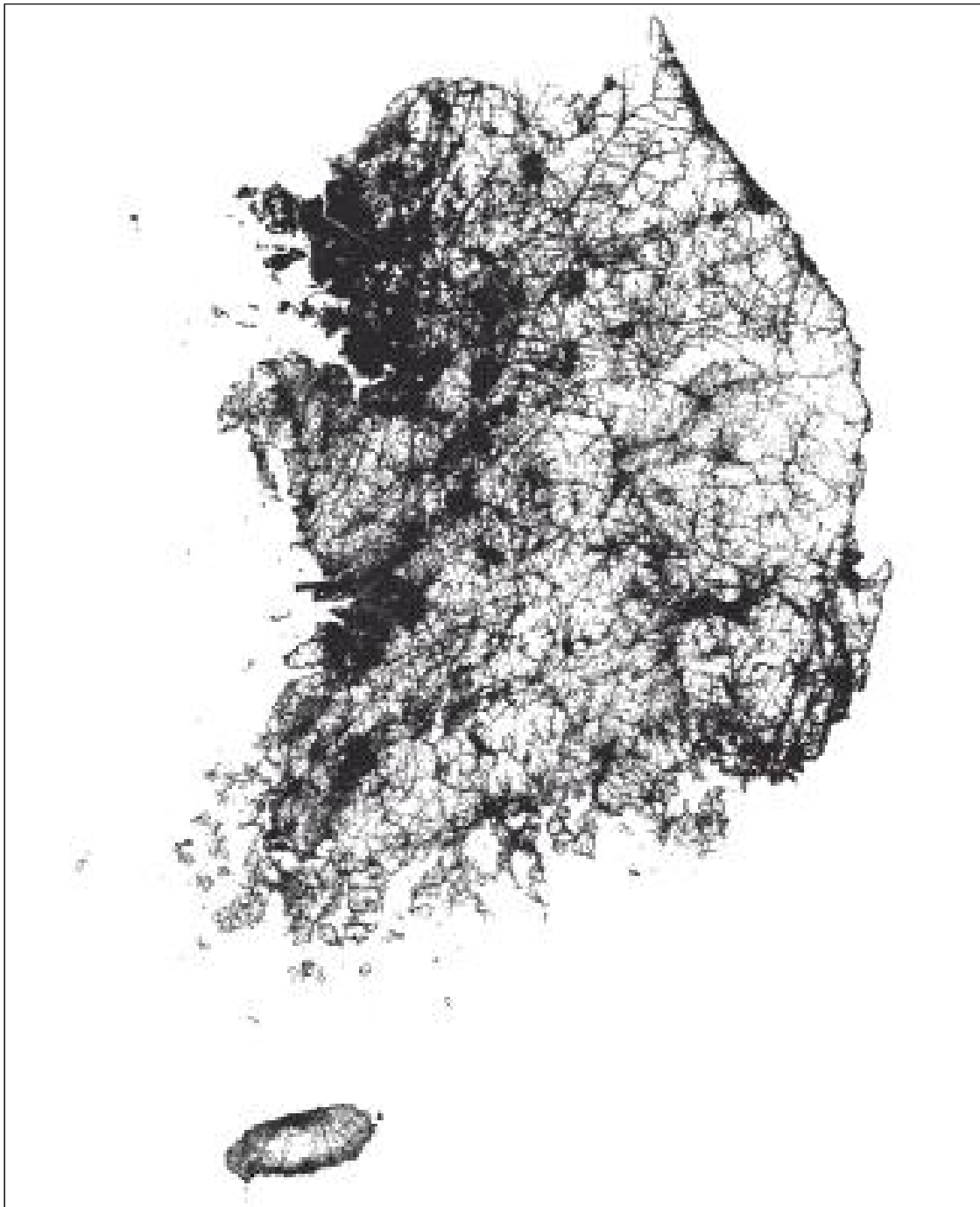
순번	필드명	설명	자료형	비고
1	P_YYYYMMDD	데이터 발생일자	VARCHAR	
2	DAY_TRIP_NO	트립ID	NUMERIC	
3	DPR_MT10_CELL_ID	출발격자ID	NUMERIC	50m 격자
4	DPR_MT10_UNIT_TM	출발시간	DATETIME	10분 단위
5	ARV_MT10_CELL_ID	도착격자ID	NUMERIC	50m 격자
6	ARV_MT10_UNIT_TM	도착시간	DATETIME	10분 단위
7	DYNA_MVMT_DIST_KM	이동거리	NUMERIC	km
8	DYNA_MVMT_SPED	속도	NUMERIC	km/h
9	DPR_CELL_STAY_PLC_DIVS_NM	출발체류지유형	VARCHAR	-
10	ARV_CELL_STAY_PLC_DIVS_NM	도착체류지유형	VARCHAR	-

- LG 유플러스 모바일 통신 원시 데이터에는 공간정보가 포함되어 있지 않아 격자와 공간정보 매칭테이블이 존재하는 격자&공간정보 데이터 테이블의 활용이 필요함
- 격자&공간정보 데이터 테이블에는 X, Y 좌표, 시도, 시군구, 행정동, 법정동에 대한 정보가 포함되어 있음

〈표 2-20〉 LG 유플러스 모바일 통신 격자&공간정보 데이터 테이블정의서

순번	필드명	설명	자료형	비고
1	CELL_ID	격자 ID	NUMERIC	-
2	SHAPE_WKT	형상정보	VARCHAR	-
3	X_AXIS	중심_X_UTMK	NUMERIC	-
4	Y_AXIS	중심_Y_UTMK	NUMERIC	-
5	X_AXIS_WGS	중심_X_WGS84	NUMERIC	-
6	Y_AXIS_WGS	중심_Y_WGS84	NUMERIC	-
7	MEGA_CD	시도코드	NUMERIC	-
8	MEGA_NM	시도명	VARCHAR	-
9	CTY_CD	시군구코드	NUMERIC	-
10	CTY_NM	시군구명	VARCHAR	-
11	ADMI_CD	행정동코드	NUMERIC	-
12	ADMI_NM	행정동명	VARCHAR	-
13	EMD_CD	법정동코드	NUMERIC	-
14	EMD_NM	법정동명	VARCHAR	-

- 공간정보 시각화 분석을 위해 별도로 격자 shp을 구축하였으며, 격자의 개수는 전국 총 6,419,645개임



〈그림 2-22〉 LG 통신 데이터 격자 전국 예시

나. 객체단위 데이터 건수 및 통행량

- LG 유플러스 모바일 통신 원시 이벤트 건수는 하루 약 1억4천건이 수집되며, 8일간 총 11.6억건이 수집되었음

〈표 2-21〉 LG 유플러스 모바일 통신 원시 이벤트 건수

일자	이벤트 건수 (건)
2023년 10월 13일 금요일	145,376,133
2023년 10월 14일 토요일	142,825,858
2023년 10월 15일 일요일	135,569,672
2023년 10월 16일 월요일	148,975,155
2023년 10월 17일 화요일	149,493,528
2023년 10월 18일 수요일	149,298,734
2023년 10월 19일 목요일	146,949,764
2023년 10월 20일 금요일	138,471,563
합계	1,156,960,407

- LG 유플러스 모바일 통신 원시 고유 트립ID 건수는 하루 약 2.6천만건이 수집되며, 8일간 총 2.08억건이 수집되었음

〈표 2-22〉 LG 유플러스 모바일 통신 원시 고유 트립ID 건수

일자	고유 트립ID 건수 (건)
2023년 10월 13일 금요일	25,824,839
2023년 10월 14일 토요일	25,561,484
2023년 10월 15일 일요일	24,507,357
2023년 10월 16일 월요일	26,771,125
2023년 10월 17일 화요일	26,811,113
2023년 10월 18일 수요일	26,918,088
2023년 10월 19일 목요일	26,689,415
2023년 10월 20일 금요일	25,251,567
합계	208,334,988

- LG 유플러스 모바일 통신 원시 시도별 이벤트 건수는 경기도에서 3.8억건으로 가장 많은 데이터가 수집되었으며 서울 1.7억건, 인천 0.91억건으로 수집되었음
- 인구가 가장 적은 세종시의 이벤트 건수가 6.7백만건으로 가장 적은 것으로 나타났음

〈표 2-23〉 LG 유플러스 모바일 통신 원시데이터 시도별 이벤트 건수

시도 구분	이벤트 건수 (건)	비율
서울특별시	166,803,589	14.4%
부산광역시	79,302,968	6.9%
대구광역시	34,881,836	3.0%
인천광역시	91,273,809	7.9%
광주광역시	36,597,976	3.2%
대전광역시	41,843,143	3.6%
울산광역시	25,862,262	2.2%
세종특별자치시	6,676,462	0.6%
경기도	377,553,028	32.6%
강원도	32,331,503	2.8%
충청북도	34,818,274	3.0%
충청남도	40,523,004	3.5%
전라북도	49,773,954	4.3%
전라남도	35,278,773	3.0%
경상북도	33,469,635	2.9%
경상남도	51,692,898	4.5%
제주특별자치도	18,277,293	1.6%
합계	1,156,960,407	100.0%

- LG 유플러스 모바일 통신 원시 시간대별 이벤트 건수는 17시-18시가 약 6.6천만 건으로 가장 많은 데이터가 수집되었으며, 03시-04시가 약 2.4천만건으로 가장 적은 데이터가 수집되었음
- 07시부터 19시까지는 약 6천만건의 이벤트가 유사한 수준으로 수집된 것을 확인할 수 있음

〈표 2-24〉 LG 유플러스 모바일 통신 원시데이터 시간대별 이벤트 건수

시간대	이벤트 건수 (건)	비율
00시-01시	32,571,166	2.8%
01시-02시	27,477,634	2.4%
02시-03시	24,854,225	2.1%
03시-04시	23,838,988	2.1%
04시-05시	25,226,424	2.2%
05시-06시	28,783,331	2.5%
06시-07시	40,449,274	3.5%
07시-08시	53,399,244	4.6%
08시-09시	61,962,709	5.4%
09시-10시	59,570,316	5.1%
10시-11시	59,443,967	5.1%
11시-12시	61,564,235	5.3%
12시-13시	63,157,570	5.5%
13시-14시	62,357,830	5.4%
14시-15시	61,952,379	5.4%
15시-16시	63,851,521	5.5%
16시-17시	64,979,262	5.6%
17시-18시	66,187,057	5.7%
18시-19시	62,441,131	5.4%
19시-20시	53,697,162	4.6%
20시-21시	48,166,734	4.2%
21시-22시	43,263,004	3.7%
22시-23시	37,640,476	3.3%
23시-24시	30,124,768	2.6%
합계	1,156,960,407	100.0%

○ 시도별 기종점 통행량의 유출입 비율이 유사하게 나타난 것을 알 수 있음

〈표 2-25〉 LG 유플러스 모바일 통신 원시데이터 기종점 이벤트 건수 (8일치)

(단위: 천건)

O\D	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주	합계
서울	162,139	1.8	0.5	127.4	0.8	1.4	0.5	0.4	4,515.8	2.8	2.0	3.4	1.8	1.1	1.1	1.6	1.9	166,804
부산	1.9	78,584	1.0	0.8	0.5	0.8	80.0	0.1	4.1	0.3	1.5	1.1	1.1	0.5	2.1	622.8	0.8	79,303
대구	0.5	0.9	34,407	0.3	0.2	0.5	1.7	0.1	2.4	0.1	1.1	1.0	0.5	0.3	439.1	25.9	0.4	34,882
인천	117.7	0.8	0.3	89,142	0.3	0.7	0.3	0.1	2,003.0	1.2	1.6	3.5	0.7	0.5	0.3	0.4	0.6	91,274
광주	0.8	0.5	0.2	0.3	34,682	7.8	0.1	1.1	3.7	2.1	1.7	4.0	383.5	1,409.2	0.3	0.5	100.0	36,598
대전	1.4	0.8	0.5	0.7	7.9	39,821	0.2	420.4	9.0	1.6	884.8	416.2	269.1	2.3	5.1	0.7	1.2	41,843
울산	0.5	81.7	1.9	0.3	0.1	0.2	25,581	0.0	1.4	0.1	0.4	0.5	0.4	0.2	90.1	103.2	0.1	25,862
세종	0.4	0.1	0.1	0.1	1.2	422.5	0.0	5,757	3.7	0.1	214.4	208.7	56.5	8.7	0.1	0.2	2.7	6,676
경기	4,527.7	4.1	2.4	1,992.7	3.8	9.4	1.5	4.3	370,227	137.7	105.6	509.1	5.1	13.2	3.3	3.4	2.7	377,553
강원	2.8	0.3	0.1	1.1	2.2	1.7	0.1	0.1	131.0	32,124	53.7	0.9	1.2	2.5	8.7	0.3	0.4	32,332
충북	2.0	1.5	0.8	1.6	1.8	879.3	0.3	215.2	106.5	53.5	33,303	116.8	36.6	4.2	88.8	0.9	5.3	34,818
충남	3.3	1.1	1.0	3.5	4.0	418.1	0.5	208.8	504.5	0.8	115.9	38,673	577.4	2.9	1.7	1.1	5.7	40,523
전북	1.8	1.1	0.5	0.7	385.9	267.3	0.5	55.9	5.0	1.2	36.3	576.7	48,103	293.3	1.0	25.9	18.2	49,774
전남	0.9	0.5	0.3	0.4	1,406.6	2.4	0.2	9.0	13.1	2.9	4.1	2.8	295.9	33,351	0.4	42.6	145.6	35,279
경북	1.0	2.1	440.7	0.3	0.3	4.9	90.5	0.1	3.3	8.8	86.3	1.7	1.0	0.4	32,777	50.9	0.1	33,470
경남	1.5	622.5	25.9	0.4	0.5	0.7	104.1	0.2	3.4	0.3	1.0	1.0	25.5	43.2	50.7	50,812	0.6	51,693
제주	2.0	0.8	0.2	0.3	99.0	1.2	0.1	2.7	2.8	0.5	5.4	5.7	18.2	143.8	0.2	0.9	17,994	18,277
합계	166,805	79,304	34,883	91,273	36,597	41,840	25,862	6,676	377,540	32,338	34,819	40,526	49,777	35,277	33,470	51,693	18,280	1,156,960

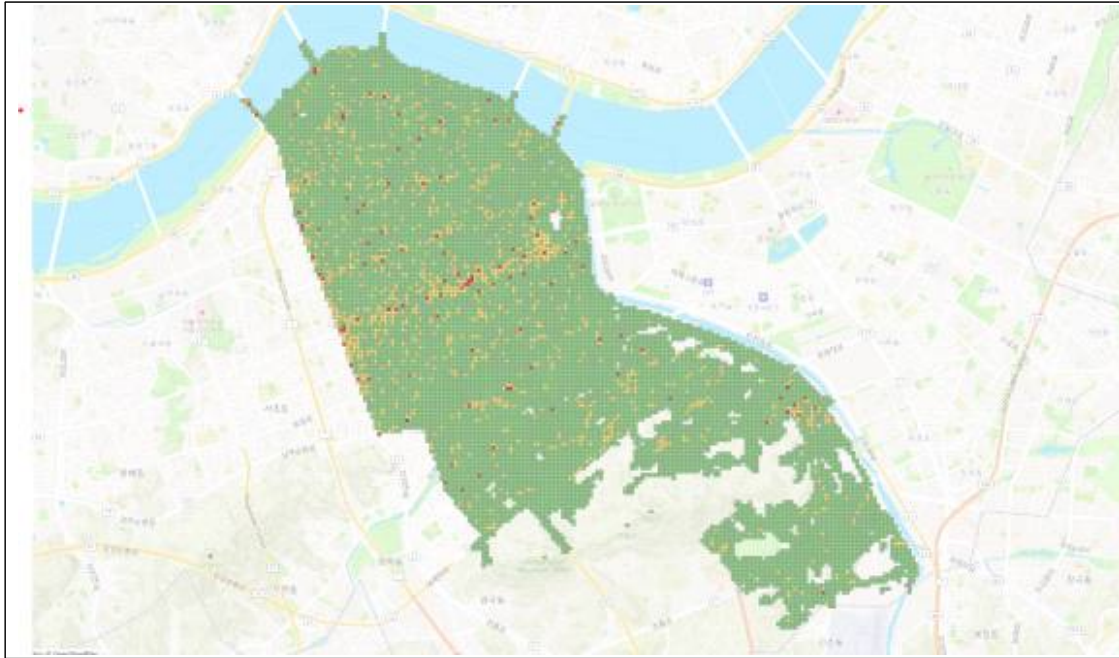
○ 내부 통행을 제외한 시도간 통행분포를 살펴보면 서울과 경기에서 약 40%의 통행이 발생하고 있는 것으로 나타남

〈표 2-26〉 LG 유플러스 모바일 통신 원시데이터 기종점 이벤트 비율 (내부통행 제외)

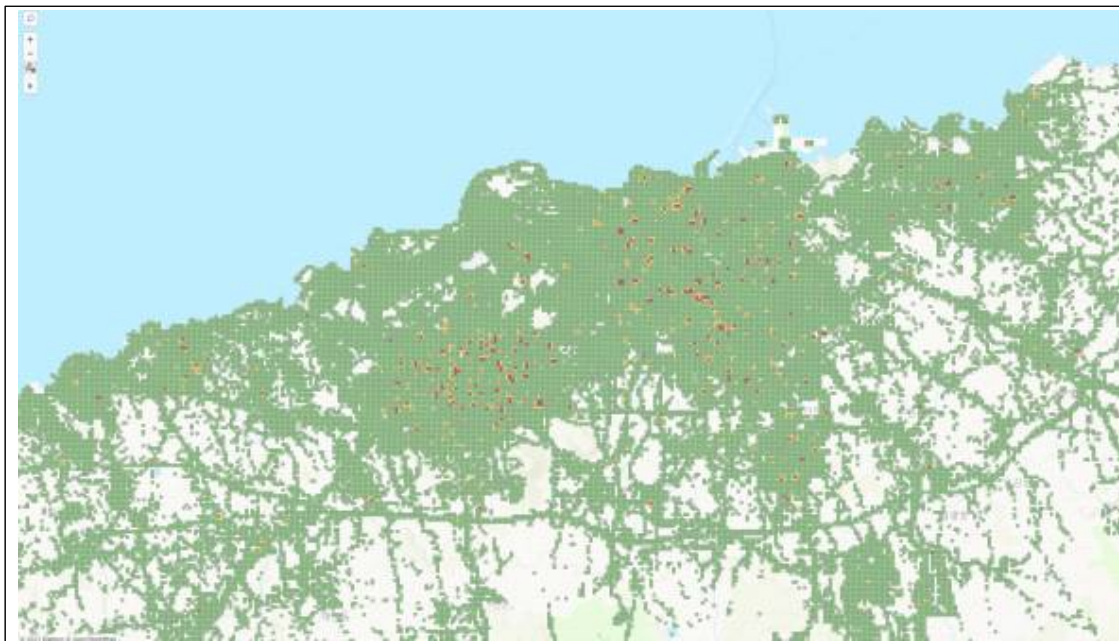
(단위: %)

O\D	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주	합계
서울	-	0.0	0.0	0.4	0.0	0.0	0.0	0.0	15.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.8
부산	0.0	-	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	2.4
대구	0.0	0.0	-	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.5	0.1	0.0	1.6
인천	0.4	0.0	0.0	-	0.0	0.0	0.0	0.0	6.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.2
광주	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	4.8	0.0	0.0	0.3	6.5
대전	0.0	0.0	0.0	0.0	0.0	-	0.0	1.4	0.0	0.0	3.0	1.4	0.9	0.0	0.0	0.0	0.0	6.9
울산	0.0	0.3	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.4	0.0	1.0
세종	0.0	0.0	0.0	0.0	0.0	1.4	0.0	-	0.0	0.0	0.7	0.7	0.2	0.0	0.0	0.0	0.0	3.1
경기	15.4	0.0	0.0	6.8	0.0	0.0	0.0	0.0	-	0.5	0.4	1.7	0.0	0.0	0.0	0.0	0.0	24.8
강원	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	-	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.7
충북	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.7	0.4	0.2	-	0.4	0.1	0.0	0.3	0.0	0.0	5.1
충남	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.7	1.7	0.0	0.4	-	2.0	0.0	0.0	0.0	0.0	6.3
전북	0.0	0.0	0.0	0.0	1.3	0.9	0.0	0.2	0.0	0.0	0.1	2.0	-	1.0	0.0	0.1	0.1	5.7
전남	0.0	0.0	0.0	0.0	4.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	-	0.0	0.1	0.5	6.5
경북	0.0	0.0	1.5	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.3	0.0	0.0	-	0.2	0.0	0.0	2.3
경남	0.0	2.1	0.1	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.2	-	0.0	3.0
제주	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.5	0.0	0.0	-	1.0
합계	15.8	2.4	1.6	7.2	6.5	6.8	1.0	3.1	24.8	0.7	5.1	6.3	5.7	6.5	2.4	3.0	1.0	100.0

○ 50m 격자 단위 통행량을 시각화한 결과는 아래와 같음



〈그림 2-23〉 강남구 50m격자 단위 통행량 시각화



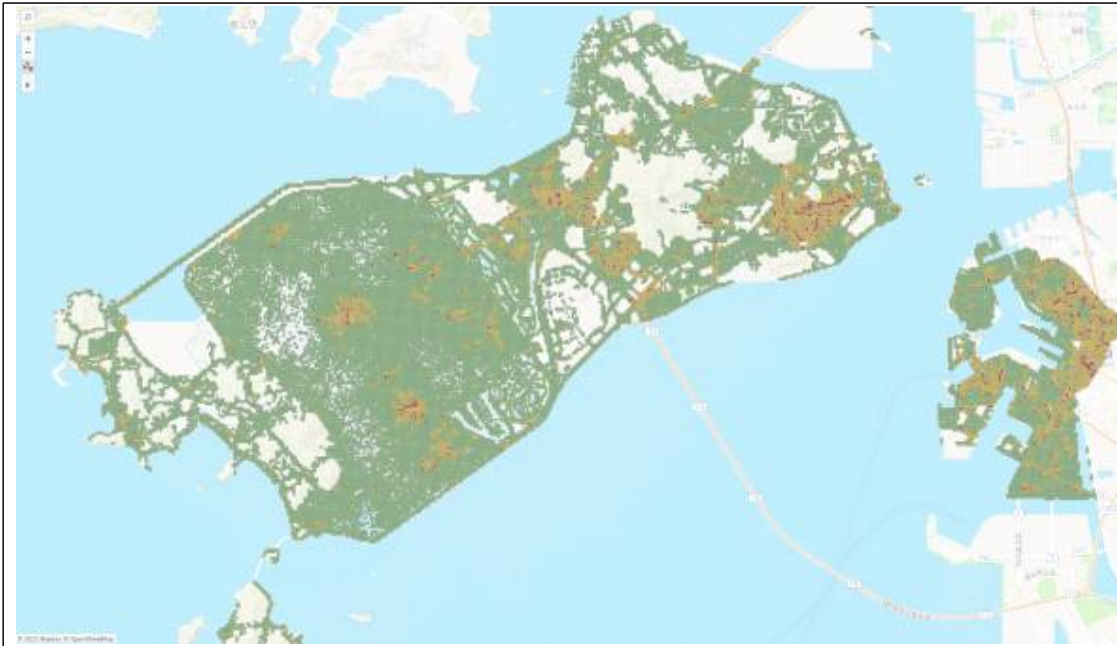
〈그림 2-24〉 제주시 50m격자 단위 통행량 시각화



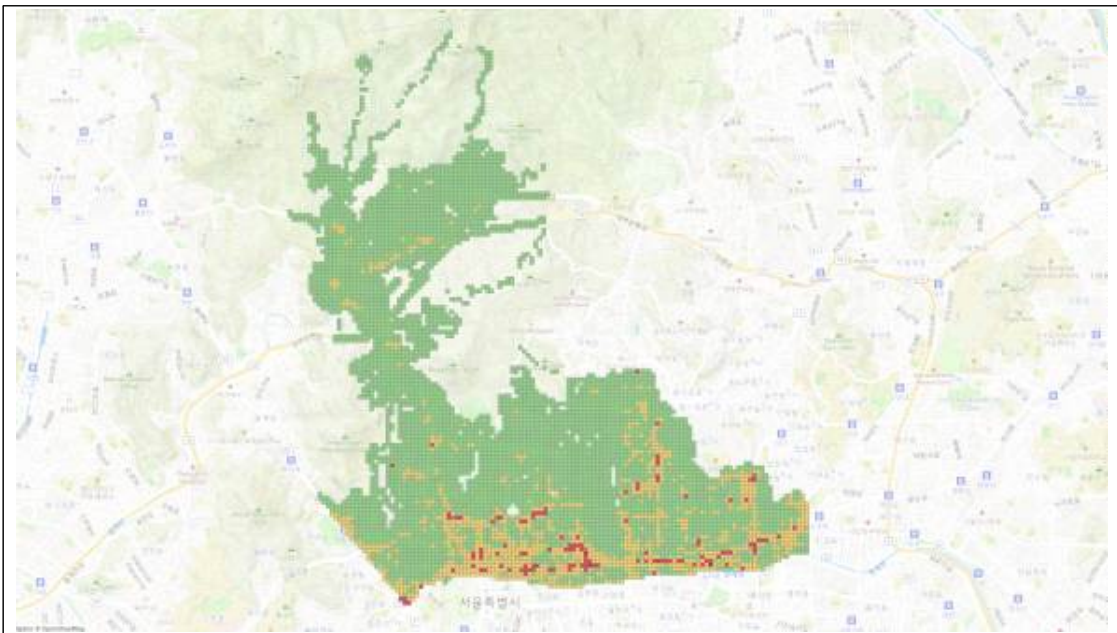
〈그림 2-25〉 울릉군 50m격자 단위 통행량 시각화



〈그림 2-26〉 해운대구 50m격자 단위 통행량 시각화



〈그림 2-27〉 인천공항 50m격자 단위 통행량 시각화



〈그림 2-28〉 종로구 50m격자 단위 통행량 시각화

3. 객체단위 모바일 통신데이터 오류 제거 및 전처리

가. 객체단위 데이터 오류 제거 및 전처리

- 객체 단위 모바일 통신 데이터의 오류를 제거하고 대용량 데이터의 처리를 위해 다음과 같은 과정을 수행함
- 오류 제거
 - 파싱 오류 제거: 한 라인의 필드 개수가 10개가 아닌 경우 제거
 - 날짜 오류 제거: 기준일자 필드가 비었거나 형식에 맞지 않은 경우 제거
 - 시간 오류 제거: 출발시간 또는 도착시간 필드가 비었거나 형식에 맞지 않은 경우 제거
 - 참조 오류 제거: 출발격자 또는 도착격자 필드가 격자-행정구역 매칭 테이블에 참조되지 않은 경우 제거
 - 이동거리 이상치 제거: 15km를 초과하는 이동거리 제거
 - 이동시간 이상치 제거: 300km/h를 초과하는 이동시간 제거
- 데이터의 효율적인 처리를 위해 다음과 같은 전처리 과정을 수행함
 - 일자 분할 : 일자별 trip_id로 파일 분할
 - 시간 정보 변환 : Datetime → Unixtime
 - 출발, 도착 행정동 정보 추가: 출발, 격자ID로 격자-행정구역 매칭 테이블의 행정동 정보 추가
 - 통행목적 제외: 무의미한 형태로 포함되어 있는 통행목적 컬럼 제외
- 최종 전처리된 데이터의 형태는 다음과 같음
- 데이터 발생일자와 트립 넘버를 조합한 Trip_ID로 키값을 설정하였으며, 출발 격자ID, 도착격자 ID, 출발 행정동 ID, 도착 행정동 ID, 출발시간, 도착시간, 이동거리, 속도로 컬럼을 구성하였음

〈표 2-27〉 LG 유플러스 모바일 통신 전처리 데이터

순번	K-V	필드명	설명	자료형	비고
1	KEY	TRIP_ID	데이터 발생일자	VARCHAR	P_YYYYMMD D + "_" + DAY_TRIP_NO
2	VALUE (반복)	O_CELL_ID	출발 격자 ID	NUMERIC	
3		D_CELL_ID	도착 격자 ID	NUMERIC	
4		O_ADMI_CD	출발 행정동 ID	NUMERIC	
5		D_ADMI_CD	도착 행정동 ID	NUMERIC	
6		DPR_MT10_UNIT _TM	출발 시간	NUMERIC	Unixtime, 출발시간으로 정렬함
7		ARV_MT10_UNIT _TM	도착 시간	NUMERIC	Unixtime
8		DYNA_MVMT_DI ST_KM	이동거리	NUMERIC	km
9		DYNA_MVMT_SP ED	속도	NUMERIC	km/h

나. 도보·비도보 수단 구분

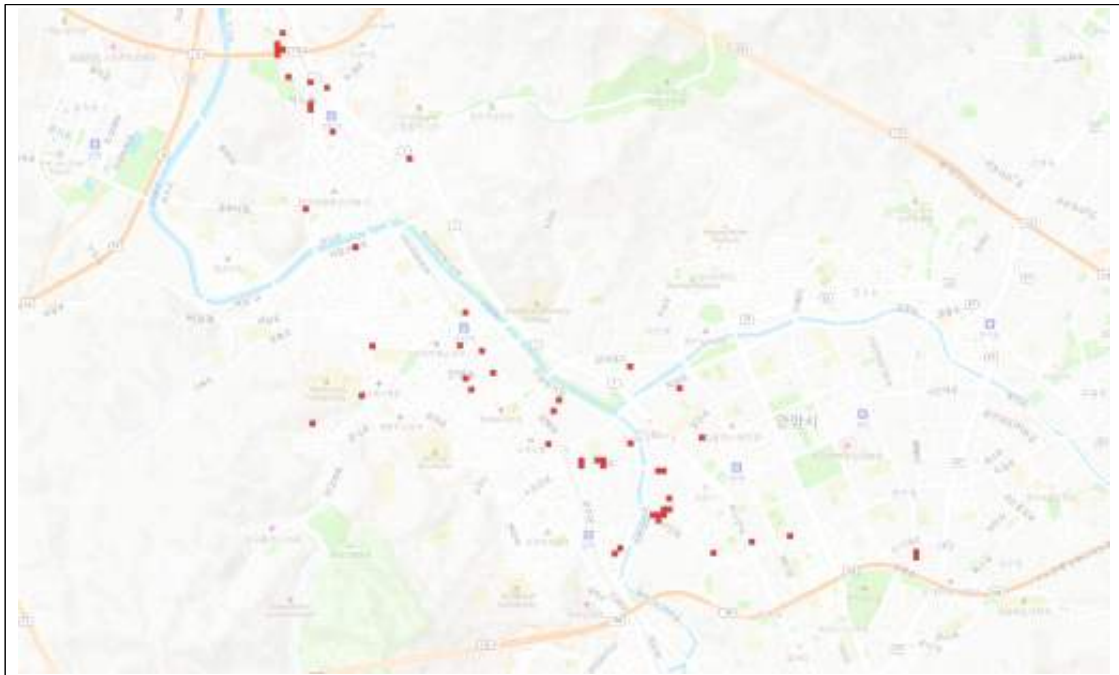
- 개별 ID의 트립에는 도보, 승용차, 버스, 기차, 비행기 등 다양한 수단이 혼재되어 있을 수 있음
- 이에 따라 이동속도를 기준으로 도보와 도보 외 통행을 구분하였음
- 도보를 구분하는 기준은 일반적인 성인의 걸음걸이 속도인 5km/h를 기준으로 5km/h 미만인 이동을 도보로 구분하였음
- 즉, 전체 모든 통행을 도보와 도보 외로 구분하는 것이 아닌 한 단위마다 도보와 도보 외를 구분하여 수단을 구분하였음
- 차량 혼잡, 수신 오류 등 다양한 케이스가 존재하여 100% 정확하게 구분하기 어렵지만 1차적으로 분류를 수행하고 향후 개선방안을 도출하고자 함

〈표 2-28〉 LG 유플러스 모바일 통신 원시데이터 예시 (트립번호: 23669006)

출발격자	도착격자	출발시간	도착시간	이동거리 (km)	이동속도 (km/h)
79111780	76167722	2023년 10월 17일 오전 3:00	2023년 10월 17일 오전 3:10	15.2	91.2
79111780	77159622	2023년 10월 17일 오전 4:00	2023년 10월 17일 오전 4:10	15.27	91.63
79111781	75735773	2023년 10월 17일 오전 7:00	2023년 10월 17일 오전 7:10	15.21	91.23
79111781	76375699	2023년 10월 17일 오전 3:30	2023년 10월 17일 오전 3:40	15.27	91.61
79111781	77079630	2023년 10월 17일 오전 1:50	2023년 10월 17일 오전 2:00	15.28	91.69
79111781	79111781	2023년 10월 17일 오전 1:40	2023년 10월 17일 오전 1:50	0.35	2.12
79111781	79111781	2023년 10월 17일 오전 3:20	2023년 10월 17일 오전 3:30	0.36	2.18
⋮	⋮	⋮	⋮	⋮	⋮

〈표 2-29〉 LG 유플러스 모바일 통신 데이터 수단구분 예시 (트립넘버: 23669006)

출발격자	도착격자	출발시간	도착시간	이동 거리 (km)	이동 속도 (km/h)	수단 구분
79111780	76167722	2023년 10월 17일 오전 3:00	2023년 10월 17일 오전 3:10	15.2	91.2	도보 외
79111780	77159622	2023년 10월 17일 오전 4:00	2023년 10월 17일 오전 4:10	15.27	91.63	도보 외
79111781	75735773	2023년 10월 17일 오전 7:00	2023년 10월 17일 오전 7:10	15.21	91.23	도보 외
79111781	76375699	2023년 10월 17일 오전 3:30	2023년 10월 17일 오전 3:40	15.27	91.61	도보 외
79111781	77079630	2023년 10월 17일 오전 1:50	2023년 10월 17일 오전 2:00	15.28	91.69	도보 외
79111781	79111781	2023년 10월 17일 오전 1:40	2023년 10월 17일 오전 1:50	0.35	2.12	도보
79111781	79111781	2023년 10월 17일 오전 3:20	2023년 10월 17일 오전 3:30	0.36	2.18	도보
⋮	⋮	⋮	⋮	⋮	⋮	



〈그림 2-29〉 트립 넘버 23669006 이동궤적 예시

- 도보 및 비도보 통행수단 구분 결과, 특별시와 광역시의 도보 통행 비율이 타 지역에 비해 상대적으로 높은 것으로 나타남
- 특히, 서울, 부산, 대구, 인천, 경기, 경남은 도보 통행의 비율이 70% 이상인 것으로 분석됨

〈표 3-30〉 LG 유플러스 모바일 통신 원시데이터 도보, 도보 외 비율 분석

시도 구분	도보	도보 외
서울특별시	78.8%	21.2%
부산광역시	77.0%	23.0%
대구광역시	75.0%	25.0%
인천광역시	79.0%	21.0%
광주광역시	59.7%	40.3%
대전광역시	59.6%	40.4%
울산광역시	72.1%	27.9%
세종특별자치시	65.0%	35.0%
경기도	73.8%	26.2%
강원도	67.5%	32.5%
충청북도	68.4%	31.6%
충청남도	61.9%	38.1%
전라북도	53.7%	46.3%
전라남도	51.3%	48.7%
경상북도	67.8%	32.2%
경상남도	72.4%	27.6%
제주특별자치도	49.7%	50.3%

4. 통행궤적 및 통행정보 데이터 구축

가. 개별 통행 ID별 통행궤적 및 통행정보 데이터 구축

- 앞서 살펴 보았듯이 원시 데이터에는 통행궤적에 대한 시퀀스 정보가 존재하지 않음
- 따라서 출발시간과 도착시간 정보를 활용하여 통행궤적에 대한 시퀀스 정보를 포함해야 온전한 개별 통행 ID별 통행궤적 데이터를 구축할 수 있음
- 원시 데이터는 아래와 같이 시퀀스 정보가 존재하지 않음

〈표 2-31〉 LG 유플러스 모바일 통신 원시데이터 예시 (트립넘버: 23669006)

순서	출발격자	도착격자	출발시간	도착시간
없음	79111780	76167722	2023년 10월 17일 오전 3:00	2023년 10월 17일 오전 3:10
	79111780	77159622	2023년 10월 17일 오전 4:00	2023년 10월 17일 오전 4:10
	79111781	75735773	2023년 10월 17일 오전 7:00	2023년 10월 17일 오전 7:10
	79111781	76375699	2023년 10월 17일 오전 3:30	2023년 10월 17일 오전 3:40
	⋮	⋮	⋮	

- 아래처럼 시간정렬을 통해 개별 통행 ID별 통행궤적 데이터를 구축하였음

〈표 2-32〉 시퀀스 부여를 통한 개별 통행ID별 통행궤적 데이터 구축 (트립넘버: 23669006)

순서	출발격자	도착격자	출발시간	도착시간
1	79111781	79111781	2023년 10월 17일 오전 1:40	2023년 10월 17일 오전 1:50
2	79111781	77079630	2023년 10월 17일 오전 1:50	2023년 10월 17일 오전 2:00
3	79111780	76167722	2023년 10월 17일 오전 3:00	2023년 10월 17일 오전 3:10
4	79111781	79111781	2023년 10월 17일 오전 3:20	2023년 10월 17일 오전 3:30
⋮	⋮	⋮	⋮	⋮

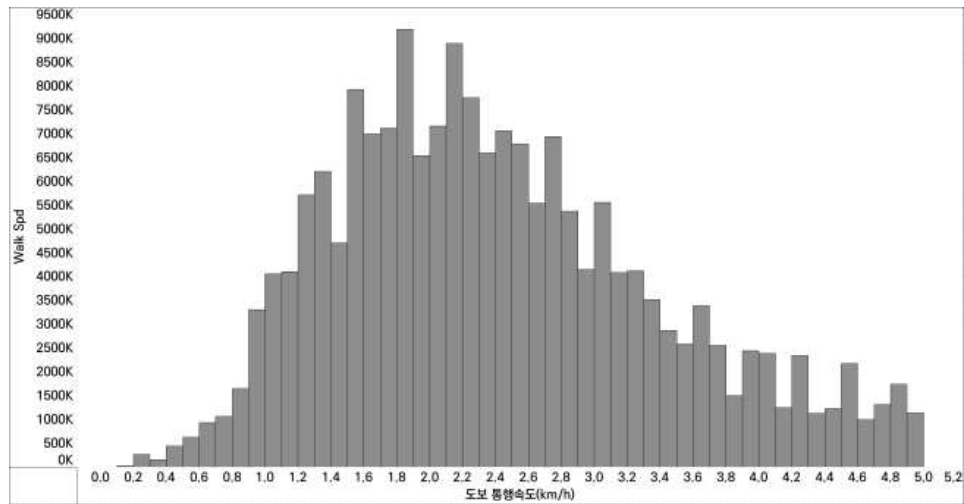
- 개별 ID의 트립에는 이동시간, 이동속도 등 다양한 통행정보가 포함되어 있음
- 본 사업에서는 개별 ID별 다양한 통행정보 데이터를 구축하기 위해 지역정보, 행정동 경유수, 도보 비율, 도보 이동시간, 도보 이동거리, 도보 이동시간, 도보 외 비율, 도보 외 이동시간, 도보 외 이동거리, 도보 외 이동시간에 대한 데이터를 구축하였음

〈표 2-33〉 모바일 통신데이터 기반 개별 통행정보 DB 구축

순번	필드명	설명	자료형	비고
1	TRIP_ID	통행 ID	VARCHAR	yyyy-MM-dd_\${id}
2	O_CELL_ID	출발 격자 ID	NUMERIC	
3	D_CELL_ID	도착 격자 ID	NUMERIC	
4	O_ADMI_CD	시작 행정동 코드	NUMERIC	
5	D_ADMI_CD	도착 행정동 코드	NUMERIC	
6	LAYOVER_CNT	행정동 경유 수	NUMERIC	
7	WALK_RATE	도보 비율	NUMERIC	격자-격자 5km/h 미만 속도
8	WALK_CAL_SPD	도보 속도(시간, 속도로 재계산)	NUMERIC	km/h
9	WALK_CAL_DIST	도보 거리(시간, 속도로 재계산)	NUMERIC	km
10	WALK_SPD	도보 속도(평균)	NUMERIC	km/h
11	WALK_DIST	도보 거리(합)	NUMERIC	km
12	WALK_TIME	도보 시간(합)	NUMERIC	hour
13	NOT_WALK_RATE	도보외 비율	NUMERIC	격자-격자 5km/h 이상 속도
14	NOT_WALK_CAL_SPD	도보외 속도(시간, 속도로 재계산)	NUMERIC	km/h
15	NOT_WALK_CAL_DIST	도보외 거리(시간, 속도로 재계산)	NUMERIC	km
16	NOT_WALK_SPD	도보외 속도(평균)	NUMERIC	km/h
17	NOT_WALK_DIST	도보외 거리(합)	NUMERIC	km
18	NOT_WALK_TIME	도보외 시간(합)	NUMERIC	hour

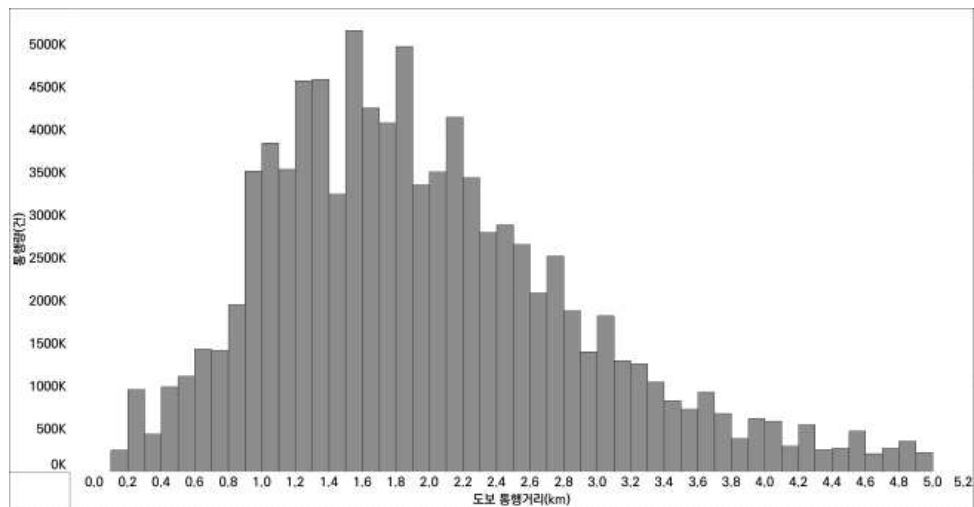
나. 개별 통행 ID별 통행속도 및 통행거리 프로파일 검토

- 도보에 대한 이동속도 프로파일은 다음과 같으며 이동속도는 1.8km/h가 가장 많은 빈도를 보임



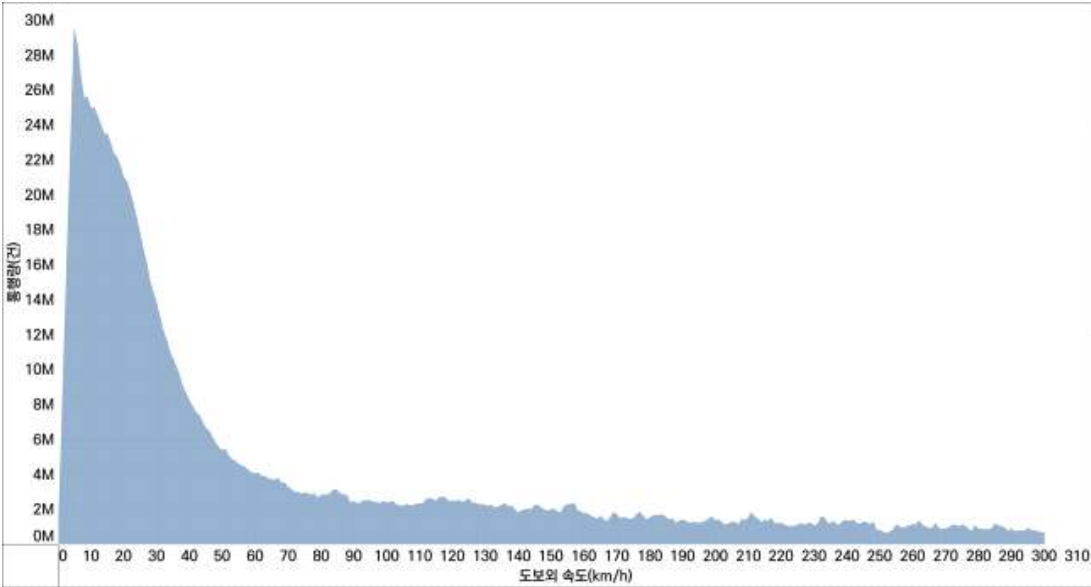
〈그림 2-30〉 도보 이동속도 프로파일

- 도보에 대한 이동거리 프로파일은 다음과 같으며 이동거리는 1.5km/h가 가장 많은 빈도를 보임



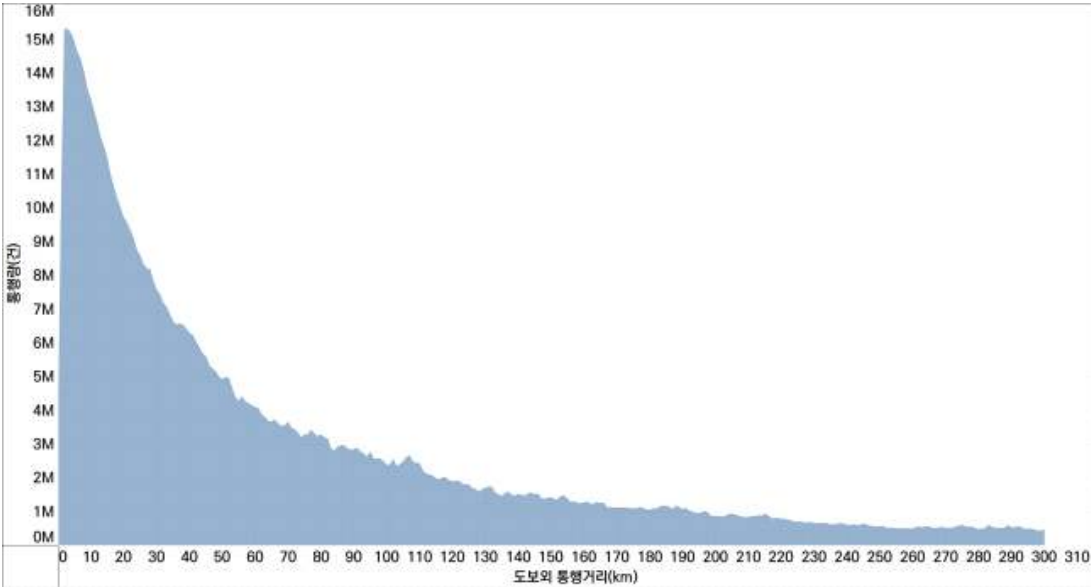
〈그림 2-31〉 도보 이동거리 프로파일

- 도보 외 통행에 대한 이동속도는 300km/h까지 속도를 보여주고 있으며, 5km/h가 가장 높은 빈도를 보임



〈그림 2-32〉 도보 외 이동속도 프로파일

- 도보 외 통행에 대한 이동거리는 3km가 가장 높은 빈도를 보이고 있으며, 3km 이후로 빈도가 점차 감소하는 것으로 분석됨



〈그림 2-33〉 도보 외 이동거리 프로파일

5. 객체단위 모바일 통신데이터 구축 및 활용 관련 검토사항

가. 도로 통행 상세 분석 관련 기술적 검토

- 도로 통행의 상세분석을 위해서는 다음과 같은 기술적 요구사항이 필요함
- GPS의 x,y좌표 데이터
 - 현재 GPS 좌표체계가 아닌 격자 체계로 이동궤적 정보를 받고 있기 때문에 실질적으로 도로를 구분해 내기에 한계가 있음
 - GPS 데이터에서 수집되는 x, y좌표의 수집을 통해 도로를 구분해 낼 수 있는 기술 개발이 필요함
- 시간 단위
 - 본 사업에서 수집한 데이터는 10분 간격의 데이터이기 때문에 매크로한 분석이 이루어 질 수밖에 없음
 - 초단위까지의 통신 데이터를 수집하여 실질적으로 도로 통행인지를 구분해 낼 수 있는 시간 데이터가 필요함
- 가속도 센서 데이터 분석
 - 휴대폰에 내장된 가속도 센서를 이용하여 사용자의 움직임을 측정하고, 이를 분석하여 도로를 추출할 수 있으며, 가속도 센서는 사용자의 움직임의 방향과 강도를 측정하므로, 이를 이용하여 도보, 달리기, 자전거 타기 등 다양한 활동을 구분할 수 있음
- 도로망, 대중교통망, 철도망, 도보망 등 네트워크 맵매칭
 - 다양한 교통네트워크와 통신 데이터의 결합을 통해 도로 통행을 명확히 구분할 수 있는 기술 개발이 필요함
 - 해당 기술을 이용할 시 도로 통행 구분에 대한 정확도가 향상될 것으로 기대함

나. 통행지표 구축 관련 기술적 검토

○ 데이터 정확성

- 이동 경로 데이터의 정확성은 결과의 신뢰성에 큰 영향을 미치기 때문에 데이터 수집 과정에서 오류가 없었는지, 데이터가 충분히 정확한지 검토하기 위한 기술 개발이 필요함

○ 데이터 처리 및 분석

- 모바일 통신 원시 데이터는 대용량 데이터로, 이를 처리하기 위해서는 최적화된 DB 설계, 적절한 빅데이터 가공환경, 인프라 시스템의 중요성이 매우 큼
- 3가지 요소 중 한가지만 적절하게 구축되지 않아도 데이터 가공을 위한 소요 시간이 크게 증가할 수 있음
- 구축된 데이터를 재가공하는 과정을 여러 차례 거쳐야 하는 경우 데이터를 빠르게 가공하기 위한 기술이 필요함

○ 통행지표 산출 알고리즘

- 최근 최적화 알고리즘, 머신러닝 알고리즘, 딥 러닝 알고리즘 등 통행지표 생성을 위한 다양한 기술개발들이 이루어지고 있음
- 기술적으로 완성도 높은 알고리즘들을 활용하기 위해 고품질 데이터 셋이 필요하며 최신 알고리즘에 대한 연구 및 기술 개발이 필요함

○ 도메인 전문 기술 필요

- 단순 알고리즘 뿐만 아니라 도메인 내에서 전문지식을 활용하여 산출해 낼 수 있는 통행지표 산출 기술 개발이 필요함
- 데이터에 대한 심도 있는 고찰과 분석을 통해 유의미한 직주균형 지표 등 통행지표를 생산해 낼 수 있는 기술이 필요함

○ 데이터 속성정보 취득

- 성별, 연령, 통행 목적 등 통행지표 산출을 위한 다양한 속성정보의 확보가 필요함



제3장 모바일 통신데이터 기반 통행 DB 검증

제1절 개요

제2절 검증 결과

제3절 보정 방법 개발

제3장 모바일 통신데이터 기반 통행 DB 검증

제1절 개요

1. 기본방향

- 모바일 통신데이터의 활용이 확대되면서, 가공된 정보의 신뢰성 검증을 위한 요구가 높아지고 있음
 - 21년 KTDB의 검증 결과, 모바일 통신데이터 기반 통행정보는 기초 사회경제적 지표(근로자수 및 학생수에 따른 통근통학량 등), 수송실적(대중교통카드자료, 철도 여객수송실적 등)과 상이한 결과를 보여주고 있다고 분석함
- 신뢰성을 검증하기 위해서는 기준이 되는 참값(ground truth)가 있어야 하는데, 이에 대한 논란이 존재함
 - 대표적인 통계인 한국교통연구원의 KTDB와 통계청 인구총조사의 자료는 검증하고자하는 모바일 통신데이터보다 샘플이 작고, 이 둘 조사자료 간에도 큰 차이가 존재하여 기준값을 설정하기 어려움
 - 수송실적(대중교통카드, 철도이용, TCS 등)을 이용하여 일부 수단통행량을 검증하는 것은 가능하나, 화물과 여객을 구분하고, 여객통행 안에서 목적통행을 검증하는 것은 사실상 어려움
 - KTDB에서 사회경제적지표를 활용한 통행 OD를 검증하고 있으나, 이는 공간 및 집단 간 동질성이 전제되어야 하나 실제로는 그렇지 않을 수 있음 (e.g., 지가, 상호 교차 검증역간 통행패턴의 이질성, 직업 간 재택근무 비율 상이 등)

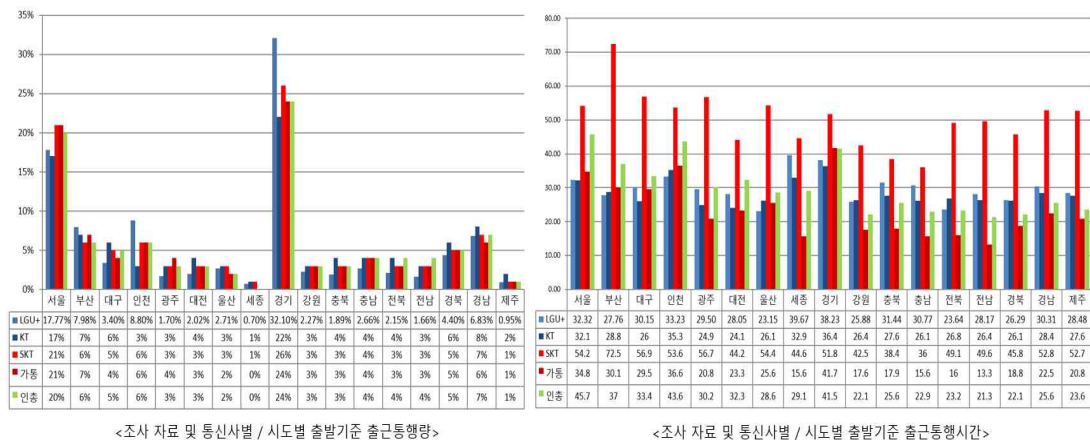
- 또한, 모바일 통신데이터와 같은 빅데이터에서는 기존 조사에서는 확인할 수 없었던 통행패턴을 확인할 수 있음에도 불구하고, 평균적 검증 접근방법이 적합하지 않을 수 있음

○ 이처럼 정확한 검증 기준값을 설정할 수 없는 상황에서는 각 데이터의 결과를 상호 교차 검증하여 각 데이터의 신뢰성을 간접적으로 검증하고자 함

- 데이터간 통행량 차이가 발생하는 구간에 대해 시공간적 특성, 사회적 특성 등을 분석함으로써 주요 오차 발생 요인을 파악할 수 있음. 이를 기준으로 각 데이터에 대한 오차율을 모델링하여 각 데이터별 보정방법을 제시할 수 있음

- ① 상호 교차 검증

- <그림 3-1>과 같이 KTDB, 인구총조사, 3개 통신사, 차량 GPS, 대중교통카드, 수송실적, 사회경제적지표를 두고, 각 정보에서 생성된 통행량 및 통행패턴을 통행수단·목적·시공간별로 교차검증하여 간접적으로 데이터의 신뢰성을 검증할 수 있음



<그림 3-1> 상호 검증 방식 예시 - 조사자료 및 통신사별 비교

- 또한, 학생수, 지역별 철도 승하차량, TCS 등 논란의 여지가 없는 전수에 가까운 명확한 수송실적의 값은 참값으로 수용하여 신뢰성 검증을 병행할 수 있음
- 이렇게 마련된 데이터 간 상호검증 결과는 데이터 전수화를 위한 가중치 값으로 활용할 수 있음

$$\text{통행량}_{ij} = \sum_{n=1}^m a_{ijn} \times T_{ijn}$$

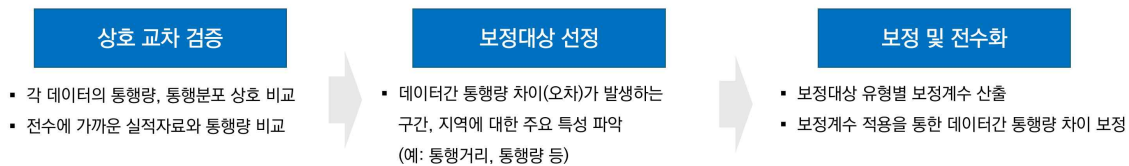
여기서, 데이터 종류 $n(1 \sim m)$ 에 대한 i 에서 j 까지 통행량 T_{ijn} 과 가중치 a_{ijn}

- ② 보정대상 선정

- 모바일 통신데이터는 기지국 또는 GNSS 정보(위치정보)를 이용하여 체류와 이동을 구분하고 통행정보를 추정하는 과정을 거쳐 통행 DB를 생성하므로, 이러한 데이터 전처리 및 가공 알고리즘 차이에 의해 데이터 간 통행량 차이가 발생할 수 있음
- * 예를 들어, 이동/체류 구분을 위한 체류시간 기준(30분 또는 60분)은 지역별 또는 장·단거리 통행의 과소 및 과대 추정의 원인이 되고 있음
- * 야간체류지를 기준으로 추정하는 집과 직장의 구분은 야간근무자를 잘못 추정할 수 있음
- 따라서 각 데이터별 통행량 차이가 발생하는 지역, 구간 등에 대한 특징(통행거리, 통행시간, 통행량, 시공간적 특성, POI 등)을 분류해봄으로써 주요 오차 발생 원인을 추정해볼 수 있으며, 주요 요인별 오차가 발생하기 시작하는 기점을 기준으로 하여 보정 대상으로 선정할 수 있음

- ③ 보정 및 전수화

- 모델링을 통해 오차 발생 요인이 미치는 영향(중요도)을 평가할 수 있으며, 모델링의 결과를 보정 및 전수화 기준으로 활용할 수 있음



〈그림 3-2〉 데이터 신뢰성 검증 및 보정 절차

- 따라서 본 사업에서는 상호 교차 검증 방식을 활용하여 모바일 통신데이터를 검증한 후, 사회경제지표를 기반으로 각 데이터별 통행량 차이가 발생하는 원인을 파악하고, 이를 기반으로 모델링하여 통행량 차이를 해석하고 보정할 수 있는 기준을 제시하고자 함

- ① 조사자료(통행실태조사, 인구총조사)와 모바일 데이터(SKT A, SKT B, KT)의 통행량을 상호 교차 검증하고자 함 ② 실적 자료(수송실적 자료, 교통카드 자료)와 모바일 데이터의 통행량 상호 교차 검증하고자 함

- 상호 교차 검증 대상은 본 사업을 통해 수집 가능한 데이터로 한정하고 공간 단위 변화에 따른 오차 변화를 확인하기 위해 분석 단위는 시도, 시군구, 읍면동 단위로 설정하고자 하며, 실적 자료의 상호 교차 검증은 상호 비교가 가능한 공간적 범위를 채택하여 검증하고자 함

- 각 자료의 통행량 차이에 대한 원인을 파악하기 위해, 사회경제지표를 기반으로 각 자료의 통행분포를 비교하고, 통행량 차이가 발생하는 지역, 구간에 대한 특성을 확인하고자 함
- 수집한 자료 중 한국교통연구원에서 수행한 자료를 활용하여 각 자료의 통행량 차이에 대해 모델링하고, 결과를 토대로 보정기준 및 방법을 제시하고자 함

2. 검증데이터 수집 및 가공

가. 데이터 수집

1) 모바일 통신데이터 기반 통행 DB

- 모바일 통신데이터는 통신사별로 통행 DB를 구축하는 알고리즘이 상이하여 최종 산출되는 통행 정보(통행량, 통행시간 등)도 통신사별로 다를 수 있기 때문에 통신사별 차이를 검증하기 위하여 서로 다른 통신사의 데이터로 구축된 통행 DB를 수집함
 - 본 사업을 통해 수집한 모바일 통신데이터 기반 통행 DB 총 3종을 수집함
 - ① KT 데이터, ② SKT 지라프 데이터, ③ SKT 리트머스 데이터를 기반으로 구축한 통행 DB를 수집하였으며, 상호 비교를 위하여 시간적·공간적·내용적 범위를 동일하게 설정함
 - 각 통신사별로 개인정보보호를 위한 특정 기준 이하의 통행량(3 또는 5 통행량)은 마스킹 처리함
 - 단, 마스킹 처리된 통행량의 총량은 별도의 정보(보정계수)로 확인할 수 있으며, 집계된 형태로 보정 전 총통행량을 확인할 수 있음 (예를 들어, 1~5 통행량을 각각 구분할 순 없지만, 마스킹 처리된 1~5 통행량의 평균인 2.5 통행량은 확인할 수 있음)
 - 단, KT는 전수화 전후의 통행량을 제공하며, 전수화 통행 정보의 경우 노이즈(noise)를 추가하는 익명화 처리(data anonymization)가 실시됨

① KT

- 시간적 범위: 2022년 11월 1일 ~ 2022년 11월 30일 (1개월)
- 공간적 범위: 전국 (500 × 500m 격자 단위)
- 내용적 범위: 읍면동 간 내부통행 포함 / 읍면동 간 내부통행 제외로 구분
- 출처: 2022년 국가교통조사사업을 통해 구축한 KT 통신데이터 기반의 2022년 기종점 통행량 데이터

〈표 3-1〉 KT 데이터 정의서

칼럼 ID	칼럼명	비고
o_date	출발 일자	yyyymmdd
o_dow	출발 요일	1~7(일~월)
o_time	출발 시간대	00~23 (1시간 단위)
o_polycode	출발 폴리곤 ID	국가지점번호 500m 격자 단위
o_sido	출발 시도	-
o_sgg	출발 시군구	-
o_emd	출발 읍면동	-
o_type	출발 체류지 유형	H: 집, C: 회사, S: 학교, R: 정기 잠재체류지, X/O:기타 잠재체류지
d_date	도착 일자	yyyymmdd
d_time	도착 시간대	00~23 (1시간 단위)
d_polycode	도착 폴리곤 ID	국가지점번호 500m 격자 단위
d_sido	도착 시도	-
d_sgg	도착 시군구	-
d_emd	도착 읍면동	-
d_type	도착 체류지 유형	H: 집, C: 회사, S: 학교, R: 정기 잠재체류지, X/O: 기타 잠재체류지
gender	성	M: 남성, F: 여성
age	연령	10세 단위 (0~110)
trans_gb	광역통행수단구분	F: 항공 / T: 고속철도 / B: 고속버스 / E: 기타
traffic_time_all_mean	통행시간_평균	1분 단위
traffic_distance_all_mean	통행거리_평균	1분 단위
traffic_speed_all_mean	통행속도_평균	km/h 단위
stay_traffic_volume	총 통행량 1	보정된 통행량 (3통행 이하 마스킹 처리)
tpop_weight	총 통행량 2	전수화된 통행량

② SKT 지라프 (GIRAF)

- 시간적 범위: 2022년 11월 1일 ~ 2022년 11월 30일 (1개월)
- 공간적 범위: 전국 (시군구 단위)
- 내용적 범위: 읍면동 간 내부통행 제외
- 출처: 2022년 국가교통조사사업을 통해 구축한 SKT 통신데이터 기반의 2022년 기종점 통행량 데이터

〈표 3-2〉 SKT 지라프 (GIRAF) 데이터 정의서

칼럼명	설명	비고
o_dt	출발 일자	yyyymmdd
o_dow	출발 요일	1~7(월~일)
o_adm	출발지 행정동	10자리 숫자코드
o_sigun	출발지 시군구	5자리 숫자코드
o_sido	출발지 시도	2자리 숫자코드
o_type	출발지 유형	집(H),W(회사),O(기타)
d_dt	도착 일자	yyyymmdd
d_adm	도착지 행정동	10자리 숫자코드
d_sigun	도착지 시군구	5자리 숫자코드
d_sido	도착지 시도	2자리 숫자코드
d_type	도착지 유형	집(H),W(회사),O(기타)
age	연령대	1세 간격
gender	성별	01:남자, 02:여자
tot_pop	총 통행량	보정된 통행량 (5통행 이하 마스킹 처리)
move_time_avg	통행시간_평균	단위(Sec)
distance_avg	통행거리_평균	단위(m)
speed_avg	통행속도_평균	m/sec

③ SKT 리트머스 (LITMUS)

- 시간적 범위: 2022년 11월 1일 ~ 2022년 11월 30일 (1개월)
- 공간적 범위: 전국 (읍면동 단위)
- 내용적 범위: 읍면동 간 내부통행 제외
- 출처: 2023년 국가교통조사사업을 통해 구축한 SKT 통신데이터 기반의 2023년 기종점 통행량 데이터

〈표 3-3〉 SKT 리트머스 (LITMUS) 데이터 정의서

칼럼명	설명	비고
o_dt	출발 일자	yyyymmdd
o_dow	출발 요일	1~7(월~일)
o_time	출발 시간대	0~23
o_adm	출발지 행정동	-
o_sigun	출발지 시군구	-
o_sido	출발지 시도	-
o_type	출발지 유형	집, 회사, 학교, 쇼핑여가, 관광, 기타
d_dt	도착 일자	yyyymmdd
d_time	도착 시간대	0~23
d_adm	도착지 행정동	-
d_sigun	도착지 시군구	-
d_sido	도착지 시도	-
d_type	도착지 유형	0:집, 1:회사, 2:학교, 3:쇼핑여가, 4:기타, 5:관광
age	연령대	10세 간격
gender	성별	01:남자, 02:여자
tot_pop	총 통행량	보정된 통행량 (5통행 이하 마스킹 처리)
move_time_avg	통행시간_평균	단위(Sec)
distance_avg	통행거리_평균	단위(m)
speed_avg	통행속도_평균	m/sec

2) 조사 자료

- 모바일 통신데이터와 상호교차 검증하기 위한 자료로 기존에 교통 수요 분석의 기초자료로 주로 활용되었던 조사자료를 수집함
 - 조사자료 중에서 통행량 정보를 제공하고 있는 한국교통연구원 개인통행실태조사, 통계청 인구총조사 총 2종의 자료를 수집함

① 통행실태조사

- 시간적 범위: 2019년, 2021년
 - 2019년 가구통행실태조사: 가구방문설문조사, 면접조사 등을 통해 수집한 통행자료
 - 2021년 개인통행실태조사: 기존 조사방식대로 수집한 통행자료를 통신사 데이터 (SKT 지라프 데이터)로 보정한 자료
- 공간적 범위: 전국 (시군구 단위)
 - 권역별 내부통행에 한해서 읍면동 단위로 구축되어 있음
- 내용적 범위: 내부통행 포함, 지역간 목적별 통행량
- 출처: 2020년, 2022년 국가교통조사사업을 통해 구축한 자료

② 인구총조사

- 시간적 범위: 2020년
- 공간적 범위: 전국 (시군구 단위)
- 내용적 범위: 내부통행 포함, 통근, 통학 목적 유출인구
- 출처: 국가통계포털(KOSIS)

3) 실적 자료

① 수송실적 자료

- 시간적 범위: 2022년
- 공간적 범위: 전국 (시군구 단위)
- 내용적 범위: 지역간 철도, 시외고속버스, 기타버스(통근·통학/일반전세)
- 출처: 2020년 국가교통조사사업을 통해 구축한 자료

② 대중교통 카드자료

- 시간적 범위: 2022년
- 공간적 범위: 대구, 광주, 부산, 대전 /전국
- 내용적 범위: 시내버스, 지하철
- 출처: 2020년 국가교통조사사업을 통해 구축한 자료

③ TCS 자료

- 시간적 범위: 2022년
- 공간적 범위: 전국
- 내용적 범위: 하이패스, 발권자료 등을 수집하여 영업소간 통행량으로 구축한 자료
- 출처: 2020년 국가교통조사사업을 통해 구축한 자료

4) 기타 및 사회경제지표

- 통행량, 통행분포와의 연관성을 확인하기 위하여 구축함
 - 전체 통행량, 목적별 통행량과 연관이 있어 보이는 자료 (인구, 학교, 산업, 취업률, 기지국, 행정경계 자료)를 수집하여 각 자료에서 필요한 정보를 추출함

〈표 3-4〉 기타 자료 수집 내용

구축 자료	추출 내용	통행량과의 연관성
인구	전체인구	전체, 통근·통학, 귀가, 기타 목적 통행발생량, 통행도착량과의 관계
	학령인구	통학목적 통행발생량과의 관계
	생산인구	전체, 통근·통학목적 통행발생량, 통행도착량과의 관계
	고령인구	전체, 통근·통학, 귀가, 기타 목적 통행발생량, 통행도착량과의 관계
학교	대학교 수	통학목적 통행도착량과의 관계
	유·초·중·고등학교 수	통학목적 통행도착량과의 관계
	대학생수	통학목적 통행도착량과의 관계
	유·초·중·고등학생 수	통학목적 통행도착량과의 관계
산업	종사자수	통근목적 통행도착량과의 관계
취업률	취업자수	통근목적 통행발생량과의 관계
기지국	기지국 수	전체 통행발생량, 통행도착량과의 관계 - 모바일 통신데이터는 기지국과 모바일 기기간의 송·수신 데이터이기 때문에 기지국의 설치현황에 따라 생성되는 모바일 통신데이터의 양과 질에 차이가 발생할 수 있고, 이를 기반으로 구축되는 통행 정보 또한 영향을 받을 수 있음
행정경계	면적	전체, 귀가, 기타 목적 통행발생량, 통행도착량과의 관계
	거리	전체, 귀가, 기타 목적 통행분포량과의 관계

① 인구 자료

- 시간적 범위: 2022년 12월 기준 (학령인구는 2021년 자료 활용)
- 공간적 범위: 전국 (시군구 단위)

- 내용적 범위: 0~99세까지 1세 단위, 성별 단위로 구축된 자료 (거주자·거주불명자·재외국민 포함, 외국인 제외)

- 출처: 국가통계포털(KOSIS)

② 학교 자료

- 시간적 범위: 2022년 10월 1일 기준

- 공간적 범위: 전국 (학교 단위로 구축)

- 내용적 범위: 유치원생, 초등학생, 중학생, 고등학생, 대학생 포함 (재학생·유학생·교원 포함, 휴학생·유예생 제외)

- 출처: 한국교육개발원 교육통계데이터베이스(KESS)

③ 산업 자료

- 시간적 범위: 2020년

- 공간적 범위: 전국 (시군구 단위, 7자리 코드)

- 내용적 범위: 종사자수

- 출처: 국가통계포털(KOSIS)

④ 취업률 자료

- 시간적 범위: 2020년

- 공간적 범위: 전국 (시도 단위)

- 내용적 범위: 건강보험 직장가입자, 교내취업자, 해외취업자, 농림어업종사자, 개인창작활동조사자, 1인창(사)업자, 프리랜서를 포함함

- 출처: 한국교육개발원 교육통계데이터베이스(KESS)

⑤ 기지국 자료

- 시간적 범위: 2022년 기준
- 공간적 범위: 전국 (시군구 단위)
- 내용적 범위: 기지국 수 정보를 수집함
- 출처: 전파누리 사이트 “무선국 현황”

⑥ 행정경계 자료

- 시간적 범위: 2022년 6월 기준
- 공간적 범위: 전국(시군구 단위)
- 내용적 범위: 행정경계
- 출처: 통계지리정보서비스(SGIS)

나. 데이터 가공

1) 데이터간 정보 통일

① 공간적 정보 통일

○ 데이터간 상호 비교를 위해서는 공간적 정보가 일치하여야 하나, 데이터별 공간 정보를 표기하는 형태가 상이하므로, 각 데이터를 연계하여 분석할 수 있는 기준 테이블을 구축하고 이를 기반으로 공간적 정보를 통일함

- KTDB에서 사용하는 시군구 250개의 존재계(대존, 소존)를 기준테이블로 설정하고 해당 코드를 기준으로 데이터를 매칭함

· 대존: 1~17 / 소존: 1~250

· 주소로 표기된 자료: 학교 정보는 좌표 정보로 변환하여 행정경계와 공간조인하여 행정동 정보를 매칭함

〈표 3-5〉 자료별 공간적 정보 제공 방식

자료	구분	정보 제공 방식	비고
모바일 통신데이터	KT	행정동 코드	5자리
	SKT 지라프	행정동 코드	"
	SKT 리트머스	행정동 코드	"
조사 자료	개인통행실태조사	존코드(TAZ)	1~250
	인구총조사	시군구 명칭	-
실적자료	수송실적	존코드(TAZ)	1~250
	대중교통카드	존코드(TAZ)	"
	TCS	존코드(TAZ)	"
	도로교통량	존코드(TAZ)	"
기타 자료	인구	시군구 명칭	-
	학교	주소	지번주소, 도로명주소 혼재
	산업	행정동코드, 시군구 명칭	5자리
	취업률	시도 명칭	-
	기지국	시군구 명칭	-
	행정경계	행정동코드, 시군구 명칭	5자리

② 통행목적 정보 통일

- 통행목적별 통행량 차이를 검증하기 위하여 통행목적 정보가 없는 실적자료를 제외하고, 모바일 통신데이터와 통행실태조사 자료를 대상으로 상이하게 구분되어있는 통행목적 분류를 “통근, 통학, 귀가, 기타” 네 가지 형태로 통일하여 재분류함

- ① KT

- 통근: 출발 체류지 유형이 ‘집(H)’이고, 도착 체류지 유형이 ‘회사(C)’인 통행
- 통학: 출발 체류지 유형이 ‘집(H)’이고, 도착 체류지 유형이 ‘학교(S)’인 통행
- 귀가: 출발 체류지 유형이 ‘집 이외의 모든 유형(C,S,R,X,O)’이고, 도착 체류지 유형이 ‘집(H)’인 통행
- 기타: 통근, 통학, 귀가를 제외한 나머지 통행

* 참고: KT 가공 알고리즘 특성상 연령을 기준(20세 기준)으로 체류지 유형을 회사(C)와 학교(S)로 구분하도록 되어 있어 통근 목적의 통행자는 전부 20대 이상으로 구성되어 있으며, 통학 목적의 통행자는 전부 20대 미만으로 구성되어 있음

- ② SKT 지라프

- 통근: 출발 체류지 유형이 ‘집(H)’이고, 도착 체류지 유형이 ‘회사(W)’인 통행
- 귀가: 출발 체류지 유형이 ‘집 이외의 모든 유형(W, O)’이고, 도착 체류지 유형이 ‘집(H)’인 통행
- 기타: 통근, 귀가를 제외한 나머지 통행

* 참고: SKT 지라프 데이터에는 통학을 추정할 수 있는 체류지 유형 정보가 없어 SKT 지라프 데이터는 통학을 제외하고 ‘통근, 귀가, 기타’ 세 가지의 목적으로 분류함

- ③ SKT 리트머스

- 통근: 출발 체류지 유형이 ‘집(0)’이고, 도착 체류지 유형이 ‘회사(1)’인 통행
- 통학: 출발 체류지 유형이 ‘집(0)’이고, 도착 체류지 유형이 ‘학교(2)’인 통행
- 귀가: 출발 체류지 유형이 ‘집 이외의 모든 유형(1~5)’이고, 도착 체류지 유형이 ‘집(0)’인 통행
- 기타: 통근, 통학, 귀가를 제외한 나머지 통행

- ④ 통행실태조사

- 7개의 목적(출근, 등교, 업무, 쇼핑, 귀가, 여가/오락/친지방문, 기타)에서 ‘출근’은 ‘통근’으로, ‘등교’는 ‘통학’으로 명칭을 변경하고, ‘출근’, ‘등교’, ‘귀가’를 제외한 나머지를 ‘기타’로 재집계함

③ 시간적 정보 통일 및 통행량 집계

- 모바일 통신데이터에는 일자별·시간대별 통행량 정보가 있으나, 조사자료에는 해당 정보가 없고 일평균 통행량 정보만 제공하므로, 모바일 통신데이터의 통행량을 조사자료에 맞추어 일평균 통행량으로 재산출함

- 1단계: 통행구간별·통행목적별·일자별 통행량 집계

- 모바일 통신데이터는 개인정보보호를 위해 일정 기준 이하의 통행량은 실제 통행량보다 더 큰 값으로 보정되어 있으므로, 보정계수를 적용하여 실제 통행량으로 변환함
- 지라프 데이터의 경우 출발일자와 도착일자가 다른 경우가 존재하여 해당 데이터는 분석 데이터에서 제외함

〈표 3-6〉 시간적 정보 통일 및 통행량 집계 1단계 예시

출발				도착				통행	통행	보정	비고
일자	지역	시간	유형	일자	지역	시간	유형	목적	량*	계수	
20221101	1	10	집	20221101	17	12	기타	기타	3	0.86	
20221101	2	13	기타	20221102	17	15	기타	기타	3	0.83	제외**
20221101	1	07	집	20221101	17	08	회사	통근	10	1.00	
20221101	1	08	집	20221101	5	09	학교	통학	7	1.00	
20221101	1	08	집	20221101	17	09	회사	통근	3	0.50	
20221102	2	17	기타	20221102	4	19	집	귀가	8	1.00	
20221102	2	16	기타	20221102	4	17	집	귀가	5	1.00	

↓

출발지	도착지	통행	기준일	통행량
		목적		***
1	17	기타	20221101	2.58
1	17	통근	20221101	11.50
1	5	통학	20221101	7.00
2	4	귀가	20221102	13.00

*: 개인정보보호를 위해 보정되어 있는 통행량

**: 출발일자와 도착일자가 다른 경우 제외

***: 보정계수를 적용한 실제 통행량

- 2단계: 한 달 총 통행량 집계

- 3단계: 2단계에서 집계한 통행량을 전체 분석일수(총 30일)로 나누어서 일평균 통행량 산출

$$\text{일평균 통행량} = \frac{\text{총 통행량}}{\text{분석일수}}$$

2) 자료 재집계 및 지표 산출

- 세분화되어 있는 자료를 각각 다른 집계 기준을 적용하여 유의미한 지표를 산출함
 - 학령인구수: 인구 자료에서 5~24세 인구를 집계함
 - 생산인구수: 인구 자료에서 15~64세 인구를 집계함
 - 고령인구수: 인구 자료에서 65~99세 인구를 집계함
 - 수용학생수: 학교 자료에서 학교별로 구분되어있는 인구수를 집계함
 - 수용학생수는 학교 카테고리를 기준으로 다시 유치원수, 중학생수, 고등학생수, 기타학생수, 대학생수로 구분하여 재집계함
 - 카테고리별 학교수: 학교 자료에서 학교별로 구분되어있는 학교수를 학교 카테고리별 기준으로 다시 유치원, 중학교, 고등학교, 기타학교, 대학교로 구분하여 재집계함
 - 취업자수: 시군구별 15세 이상 인구에 시도별 취업률을 곱하여 산출함
- 시군구별 스케일 문제로 인해 수치보다 비율로 환산하는 것이 의미 있을 것으로 판단되는 자료는 전체 인구 또는 면적으로 나누어서 추가로 자료를 구축함
 - 인구밀도: 해당 시군구의 인구수 / 해당 시군구의 면적
 - 생산인구 비율: 해당 시군구의 생산인구수 / 해당 시군구의 전체인구수
 - 고령인구 비율: 해당 시군구의 고령인구수 / 해당 시군구의 전체인구수
 - 기지국 밀도: 해당 시군구의 기지국 수 / 해당 시군구의 면적
- GIS 툴을 활용하여 유의미한 지표를 추출함
 - 면적: 행정경계 자료를 기반으로 필드계산기에 Area 함수를 입력하여 산출함
 - 거리: 행정경계 자료를 기반으로 시군구의 중심점을 추출한 후 필드계산기에 length 함수를 입력하여 중심점간 직선거리를 산출함

〈표 3-7〉 최종 가공 자료 (요약)

구분		기준년도	구축 기준	출처
인구	인구	2022	거주자·거주불명자·재외국민 포함, 외국인 제외	국가통계포털 (KOSIS)
	인구밀도		해당 시군구의 인구수를 해당 시군구의 면적으로 나눈 값	
학령인구수		2021	1세 단위로 구축된 인구 자료에서 5~24세 인구를 집계	국가통계포털 (KOSIS)
생산 인구	생산인구수	2022	1세 단위로 구축된 인구 자료에서 16~64세 인구를 집계	국가통계포털 (KOSIS)
	생산인구비율		해당 시군구의 생산인구수를 해당 시군구의 전체인구수로 나눈 값	
고령 인구	고령인구수	2022	1세 단위로 구축된 인구 자료에서 65~99세 인구를 집계	국가통계포털 (KOSIS)
	고령인구비율		해당 시군구의 고령인구수를 해당 시군구의 전체인구수로 나눈 값	
학교수	유·초·중·고등학 생수, 대학생수	2022	유치원, 초등학교, 중학교, 고등학교 수 집계	한국교육개발원 교육통계데이터 베이스(KESS)
	대학교 수		-	
수용학생수		2022	학교별(유,초,중,고,대학교)학생수 집계 - 대학생의 경우 재학생,유학생, 교원 포함 / 휴학생, 유예생 제외	한국교육개발원 교육통계데이터 베이스(KESS)
종사자수		2020	-	국가통계포털 (KOSIS)
취업자수		2021	해당 시군구의 15세 이상 인구에 해당 시도의 취업률을 곱한 값	한국교육개발원 교육통계데이터 베이스(KESS)
기지국	기지국수	2022	-	전파누리 사이트
	기지국 밀도		해당 시군구의 기지국 수를 해당 시군구의 면적으로 나눈 값	
면적		2022	-	통계지리정보 서비스(SGIS)
거리		2022	시군구의 중심점간의 직선거리	통계지리정보 서비스(SGIS)

3) 데이터간 조인 및 오차율 산출

- 공간 단위별 차이를 확인하기 위하여 시도, 시군구 단위로 데이터를 집계함
 - 통행발생량, 통행도착량, 통행분포량(OD pair) 세 가지 형태로 데이터를 구축함
- 전술한 바와 같이 재분류한 세 가지 통행목적(통근, 통학, 귀가, 기타)에 따라 통행량을 재집계함
- 상호 비교값은 각 자료에서 집계된 통행량(발생량, 도착량, 분포량) 자체의 값과 각 자료의 통행량을 빼거나 나누어서 산출한 절대오차(Absolute error)를 사용함

$$\text{절대오차} = |\text{기준값} - \text{비교값}|$$

제2절 검증 결과

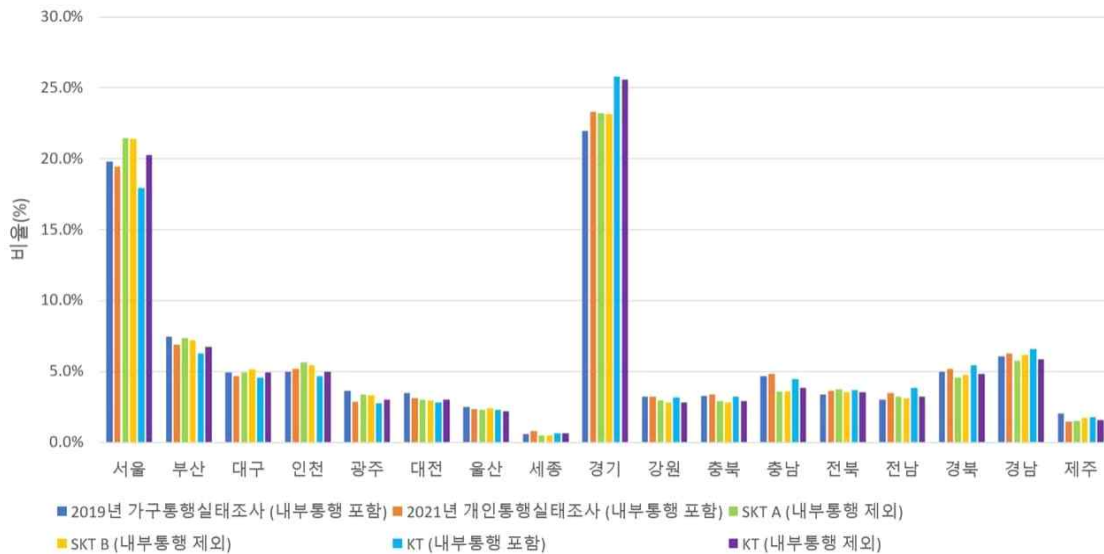
1. 데이터간 상호교차검증 결과

가. 시도 단위

1) 통행발생량 기준

① 전체

- 시도별 통행 발생비율은 조사자료 2종과 모바일 통신데이터 4종의 통행량 비율과 대체로 유사하게 나타남 (<그림 3-3> 참조)



〈그림 3-3〉 (통행발생량) 전체 - 시도별 통행 비율

- 2019년 가구통행실태조사 자료(내부통행 포함)를 기준으로 SKT A(내부통행 제외)가 평균 약 0.48%, SKT B(내부통행 제외)가 약 0.45%, KT(내부통행 포함)가 약 0.71%, KT(내부통행 제외)가 약 0.53% 정도 차이를 보이며, 2021년 개인통행실태조사 자료(내부통행 포함)를 기준으로 SKT A(내부통행 제외)

- 가 약 0.46%, SKT B(내부통행 제외)가 평균 약 0.44%, KT(내부통행 포함)가 0.45%, KT(내부통행 제외)가 평균 약 0.43% 차이를 보임
- 모바일 통신데이터는 대체로 조사자료보다 서울, 경기 지역에서 통행 비율이 다소 높게 나타났고, 충남, 경북 등 일부 지방지역에서는 통행 비율이 다소 낮게 나타남
 - SKT A 데이터와 SKT B 데이터는 2019년 가구통행실태조사에 비해 서울, 경기 지역의 통행 비율이 전반적으로 약 1% 정도 높게 나타났으며, 충남 지역은 약 1% 정도 통행 비율이 낮게 나타남
 - KT 데이터(읍면동 내부통행 포함, 제외)는 경기 지역이 차이가 다소 크게 나타났으며 (내부통행을 포함할 경우 +3.83%, 내부통행을 제외할 경우 +3.64%), SKT 데이터와는 반대로 서울 지역은 2019년 가구통행실태조사에 비해 낮게 나타났으며, 충남 지역은 큰 차이를 보이지 않고, 부산 지역에서 다소 낮게 나타남(내부통행을 포함할 경우 -1.16%, 내부통행을 제외할 경우 -0.742%)
 - 총량 측면에서는 2019년 가구통행실태조사(내부통행 포함)과 KT 데이터(내부통행 포함)가 가장 유사하게 나타남
 - 2021년 개인통행실태조사(내부통행 포함)은 2019년 가구통행실태조사보다 값이 약 1.6배 정도 높게 나타남

〈표 3-8〉 (통행발생량) 전체 - 데이터별 시도별 통행량

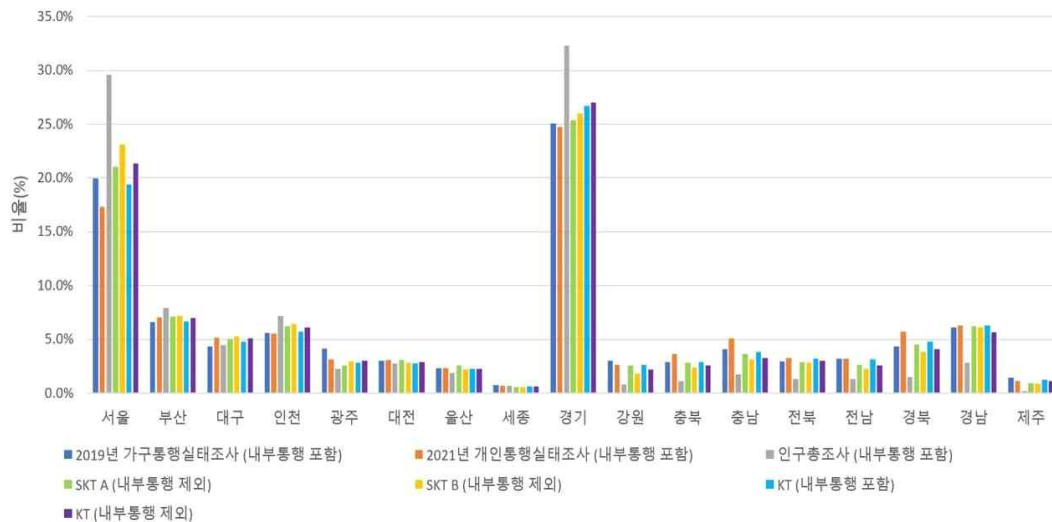
(단위: 통행, %)

시도	2019년 가구통행실태조사 (내부통행 포함)		2021년 개인통행실태조사 (내부통행 포함)		SKT A (내부통행 제외)		SKT B (내부통행 제외)		KT (내부통행 포함)		KT (내부통행 제외)	
	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율
서울	17,677,365	19.8	28,594,699	19.4	23,768,006	21.4	16,029,319	21.4	16,151,215	17.9	14,553,830	20.3
부산	6,648,435	7.4	10,101,969	6.9	8,145,629	7.4	5,396,309	7.2	5,663,298	6.3	4,829,688	6.7
대구	4,418,108	5.0	6,879,330	4.7	5,441,424	4.9	3,843,296	5.1	4,092,354	4.5	3,527,177	4.9
인천	4,465,002	5.0	7,595,075	5.2	6,256,402	5.6	4,094,423	5.5	4,223,380	4.7	3,572,832	5.0
광주	3,233,588	3.6	4,221,985	2.9	3,736,014	3.4	2,489,421	3.3	2,500,608	2.8	2,162,310	3.0
대전	3,117,532	3.5	4,574,007	3.1	3,323,029	3.0	2,242,325	3.0	2,547,187	2.8	2,169,253	3.0
울산	2,223,323	2.5	3,485,108	2.4	2,566,610	2.3	1,785,607	2.4	2,045,016	2.3	1,578,650	2.2
세종	539,194	0.6	1,145,857	0.8	534,417	0.5	359,078	0.5	576,280	0.6	458,165	0.6
경기	19,593,677	22.0	34,252,253	23.3	25,705,253	23.2	17,361,197	23.2	23,198,639	25.8	18,362,854	25.6
강원	2,867,736	3.2	4,728,882	3.2	3,271,863	3.0	2,112,921	2.8	2,866,676	3.2	2,004,868	2.8
충북	2,938,679	3.3	4,959,888	3.4	3,212,641	2.9	2,110,002	2.8	2,891,316	3.2	2,091,212	2.9
충남	4,173,250	4.7	7,086,783	4.8	3,984,155	3.6	2,682,747	3.6	4,003,281	4.4	2,772,133	3.9
전북	2,996,462	3.4	5,336,160	3.6	4,123,537	3.7	2,653,736	3.5	3,336,273	3.7	2,547,492	3.6
전남	2,713,338	3.0	5,091,184	3.5	3,593,049	3.2	2,333,675	3.1	3,441,587	3.8	2,305,701	3.2
경북	4,437,333	5.0	7,613,979	5.2	5,057,289	4.6	3,567,669	4.8	4,916,645	5.5	3,465,392	4.8
경남	5,399,909	6.1	9,259,633	6.3	6,390,199	5.8	4,632,740	6.2	5,922,746	6.6	4,211,863	5.9
제주	1,809,478	2.0	2,143,600	1.5	1,701,721	1.5	1,279,953	1.7	1,606,470	1.8	1,134,041	1.6
계	89,252,409	100.0	147,070,394	100.0	110,811,240	100.0	74,974,418	100.0	89,982,971	100.0	71,747,462	100.0

② 통근

○ 인구총조사 결과를 제외한 나머지 데이터의 분포가 유사하게 나타남

- 인구총조사의 경우 타 데이터와 서울, 경기 지역에서 다소 큰 차이를 보임
- 같은 조사자료인 가구통행실태조사와 비교하더라도 서울은 약 +11%, 경기도는 +7% 정도 높게 집계됨



〈그림 3-4〉 (통행발생량) 통근 - 시도별 통행 비율

○ 조사자료 중에서 모바일 데이터와 상대적으로 차이가 작게 나타난 것은 2019년 가구통행실태조사로 분석됨

- 2019년 가구통행실태조사(내부통행 포함)는 평균 약 0.53%, 2021년 개인통행실태조사(내부통행 포함)는 평균 약 0.72%, 인구총조사(내부통행 포함)는 평균 약 1.90% 차이를 보임

○ 2019년 가구통행실태조사와 상대적으로 유사한 모바일 데이터는 내부통행을 포함한 KT 데이터로 분석됨

- SKT A 데이터(내부통행 제외)는 평균 약 0.45% 차이를 보이며, SKT B 데이터(내부통행 제외)는 약 0.75%, KT 데이터(내부통행 포함)은 약 0.38%, KT 데이터(내부통행 제외)는 약 0.59% 차이를 보임

○ 전체 통행일 때보다 통근 통행일 때 2021년 개인통행실태조사 결과는 모바일

데이터와 차이가 더 벌어지는 것을 확인함

- 전체 통행일 때 모바일 데이터와의 차이가 0.4%였으나, 통근 통행일 때에는 그 차이가 0.6%로 다소 늘어남

○ 총량은 2019년 가구통행실태조사와 2021년 개인통행실태조사가 거의 비슷하고 모바일 데이터는 2019년 가구통행실태조사 결과보다 절반가량 적게 집계됨

- SKT 데이터는 내부통행을 포함할 경우 총량이 증가할 수 있으나 KT의 내부통행 비율을 고려했을 때 2019년 가구통행실태조사보다는 적을 것으로 예상됨

〈표 3-9〉 (통행발생량) 통근 - 데이터별 시도별 통행량 (조사자료)

(단위: 통행, %)

시도	2019년 가구통행실태조사 (내부통행 포함)		2021년 개인통행실태조사 (내부통행 포함)		2020년 인구총조사 (내부통행 포함)	
	통행량	비율	통행량	비율	통행량	비율
서울	4,581,716	20.0	3,951,487	17.3	2,424,563	29.6
부산	1,518,863	6.6	1,605,155	7.0	650,442	7.9
대구	1,003,973	4.4	1,186,541	5.2	368,341	4.5
인천	1,283,616	5.6	1,261,814	5.5	588,160	7.2
광주	954,846	4.2	715,588	3.1	184,828	2.3
대전	701,278	3.1	698,716	3.1	229,277	2.8
울산	528,631	2.3	534,102	2.3	154,338	1.9
세종	170,471	0.7	158,734	0.7	57,498	0.7
경기	5,756,461	25.1	5,656,541	24.8	2,646,044	32.3
강원	687,976	3.0	599,946	2.6	65,958	0.8
충북	662,113	2.9	833,756	3.7	93,720	1.1
충남	938,192	4.1	1,159,333	5.1	143,837	1.8
전북	677,960	3.0	751,413	3.3	106,949	1.3
전남	739,510	3.2	730,566	3.2	109,681	1.3
경북	1,002,824	4.4	1,303,867	5.7	124,319	1.5
경남	1,399,728	6.1	1,438,540	6.3	231,924	2.8
제주	330,106	1.4	255,477	1.1	15,835	0.2
계	22,938,262	100.0	22,841,575	100.0	8,195,714	100.0

〈표 3-10〉 (통행발생량) 통근 - 데이터별 시도별 통행량 (모바일 데이터)

(단위: 통행, %)

시도	SKT A (내부통행 제외)		SKT B (내부통행 제외)		KT (내부통행 포함)		KT (내부통행 제외)	
	통행량	비율	통행량	비율	통행량	비율	통행량	비율
서울	2,840,973	21.0	1,663,419	23.1	2,424,089	19.4	2,296,801	21.4
부산	962,995	7.1	517,406	7.2	831,867	6.6	752,099	7.0
대구	682,768	5.1	381,344	5.3	602,903	4.8	549,211	5.1
인천	842,045	6.2	465,165	6.5	719,556	5.7	656,142	6.1
광주	348,048	2.6	214,871	3.0	351,700	2.8	322,490	3.0
대전	420,493	3.1	202,308	2.8	350,313	2.8	314,626	2.9
울산	351,167	2.6	160,630	2.2	286,746	2.3	242,076	2.3
세종	73,546	0.5	42,933	0.6	78,020	0.6	65,942	0.6
경기	3,432,697	25.4	1,873,355	26.0	3,344,840	26.7	2,900,133	27.0
강원	350,805	2.6	132,567	1.8	327,810	2.6	236,507	2.2
충북	384,019	2.8	171,935	2.4	363,537	2.9	278,916	2.6
충남	492,118	3.6	228,623	3.2	479,289	3.8	351,637	3.3
전북	387,840	2.9	205,283	2.8	405,736	3.2	328,264	3.1
전남	362,010	2.7	164,473	2.3	397,429	3.2	276,333	2.6
경북	609,110	4.5	277,204	3.8	597,352	4.8	440,395	4.1
경남	841,093	6.2	442,249	6.1	792,218	6.3	611,723	5.7
제주	129,985	1.0	63,065	0.9	160,689	1.3	120,157	1.1
계	13,511,710	100.0	7,206,829	100.0	12,514,091	100.0	10,743,453	100.0

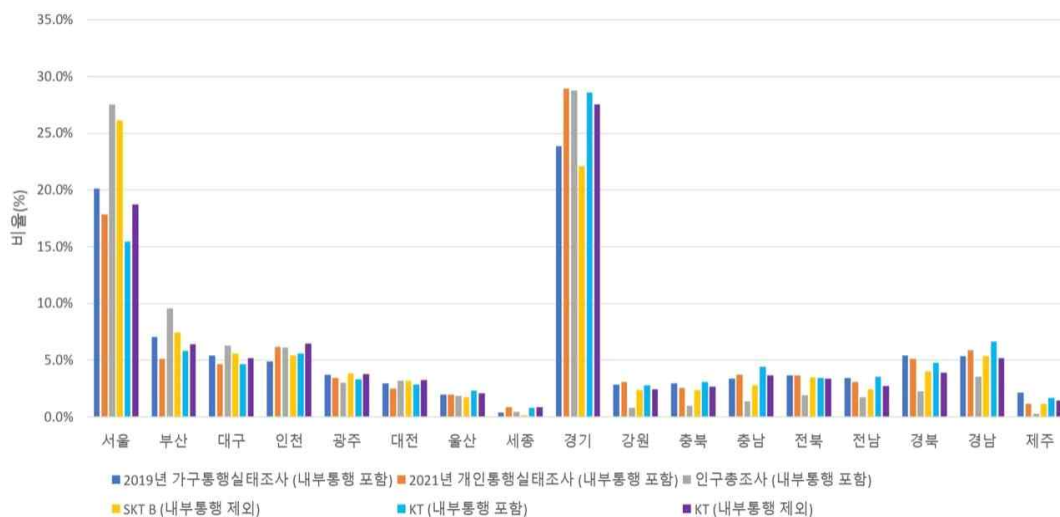
③ 통학

○ 가구통행실태조사 자료와 모바일 데이터의 통행 비율 차이가 통근 통행보다 더 크게 벌어지고, 인구총조사 자료와 모바일 데이터의 통행 비율 차이는 통근 통행보다 약간 줄어듦

- 통근 통행에서 가구통행실태조사와 모바일 데이터와의 차이가 0.6%였으나, 통학 통행에서는 0.8%인 것으로 분석됨
- 통근 통행에서 인구총조사와 모바일 데이터와의 차이는 1.9%, 통학 통행에서는 1.7%인 것으로 분석됨

○ 타 지역에 비해 서울, 경기 지역에서 데이터별 결과 차이가 큰 것으로 나타났으며, 통행 비율 측면에서 2019년 가구통행실태조사 결과와 가장 유사한 분포를 보이는 것은 KT 데이터(내부통행 제외)로 확인됨(<그림 3-5> 참조)

- 2019년 가구통행실태조사 자료와 비교하였을 때 SKT B 데이터는 서울 지역에서 큰 차이를 보이고, KT 데이터는 서울 지역보다는 경기 지역에서 큰 차이를 보임
 - SKT B 데이터는 서울(+5.98%), 경기(-1.75%), KT 데이터(내부통행 포함)는 서울(-4.71%), 경기(+4.71%), KT 데이터(내부통행 제외)는 서울(-1.41%), 경기(+3.67%) 차이를 보임



〈그림 3-5〉 (통행발생량) 통학 - 시도별 통행 비율

- 총량 측면에서는 동종의 자료, 데이터도 차이가 크게 나타남
 - 2019년 가구통행실태조사와 2021년 개인통행실태조사는 약 2배 가량 차이가 나타나며, 인구총조사는 2019년 가구통행실태조사 결과의 약 1/3 수준임
 - SKT B 데이터(내부통행 제외)와 KT 데이터(내부통행 제외)도 약 1.4배 차이를 보임

〈표 3-11〉 (통행발생량) 통학 - 데이터별 시도별 통행량 (조사자료)

(단위: 통행, %)

시도	2019년 가구통행실태조사 (내부통행 포함)		2021년 개인통행실태조사 (내부통행 포함)		2020년 인구총조사 (내부통행 포함)	
	통행량	비율	통행량	비율	통행량	비율
서울	651,732	20.2	1,132,595	17.9	244,226	27.6
부산	229,171	7.1	324,038	5.1	84,746	9.6
대구	175,128	5.4	297,244	4.7	56,118	6.3
인천	158,656	4.9	390,708	6.2	54,332	6.1
광주	121,461	3.8	216,982	3.4	27,077	3.1
대전	95,502	3.0	159,876	2.5	28,589	3.2
울산	63,679	2.0	127,411	2.0	16,755	1.9
세종	14,087	0.4	55,608	0.9	4,249	0.5
경기	772,131	23.9	1,833,258	28.9	254,968	28.8
강원	92,712	2.9	195,287	3.1	7,125	0.8
충북	96,276	3.0	162,632	2.6	8,580	1.0
충남	110,283	3.4	237,885	3.8	12,562	1.4
전북	119,172	3.7	233,590	3.7	17,208	1.9
전남	111,392	3.4	197,707	3.1	15,333	1.7
경북	176,476	5.5	326,871	5.2	20,162	2.3
경남	174,413	5.4	373,217	5.9	31,469	3.6
제주	69,700	2.2	72,930	1.2	2,719	0.3
계	3,231,971	100.0	6,337,838	100.0	886,218	100.0

〈표 3-12〉 (통행발생량) 통학 - 데이터별 시도별 통행량 (모바일 데이터)

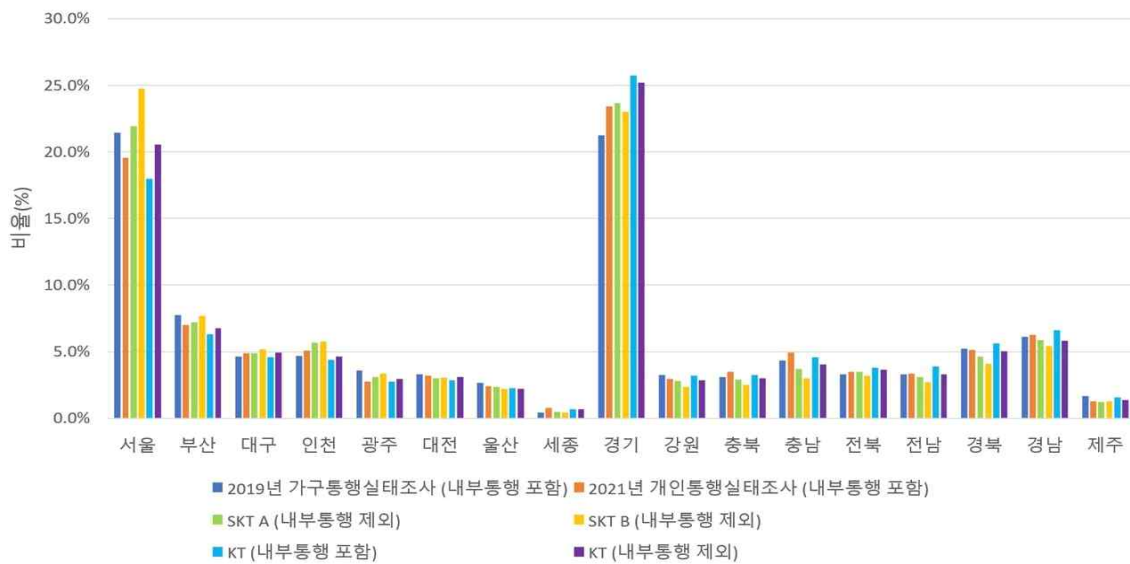
(단위: 통행, %)

시도	SKT B (내부통행 제외)		KT (내부통행 포함)		KT (내부통행 제외)	
	통행량	비율	통행량	비율	통행량	비율
서울	238,112	26.1	300,564	15.5	240,889	18.8
부산	68,279	7.5	113,438	5.8	82,712	6.4
대구	51,256	5.6	91,304	4.7	66,705	5.2
인천	49,591	5.4	109,447	5.6	83,172	6.5
광주	35,175	3.9	64,340	3.3	48,573	3.8
대전	29,390	3.2	55,802	2.9	41,915	3.3
울산	16,134	1.8	45,375	2.3	26,905	2.1
세종	1,580	0.2	15,930	0.8	11,396	0.9
경기	201,651	22.1	556,224	28.6	353,920	27.6
강원	21,937	2.4	54,917	2.8	31,484	2.5
충북	21,694	2.4	59,801	3.1	34,635	2.7
충남	25,751	2.8	85,923	4.4	47,023	3.7
전북	32,064	3.5	67,304	3.5	43,752	3.4
전남	22,256	2.4	69,167	3.6	35,402	2.8
경북	36,471	4.0	93,637	4.8	50,599	3.9
경남	48,808	5.4	129,391	6.7	66,675	5.2
제주	10,583	1.2	32,584	1.7	18,534	1.4
계	910,733	100.0	1,945,149	100.0	1,284,292	100.0

④ 귀가

○ 전체 통행과 마찬가지로 서울, 경기 지역을 제외하고는 조사자료와 모바일 데이터는 약 0.4% 정도의 근소한 차이를 보임

- 2019년 가구통행실태조사(내부통행 포함)를 기준으로, SKT A 데이터(내부통행 제외)와 SKT B 데이터(내부통행 제외)는 조사자료에 비해 서울(SK A +0.5%, SKT B +3.3%), 경기(SK A +2.4%, SKT B +1.7%) 모두 비율이 높게 나타났고, KT 데이터는 서울은 낮게(내부통행 포함 - 3.5%, 내부통행 제외 - 0.9%), 경기는 높게(내부통행 포함 +4.5%, 내부통행 제외 +4.0%) 나타남
- 2021년 개인통행실태조사를 기준으로, 서울 지역의 경우 SKT A 데이터(내부통행 제외)와 SKT B 데이터(내부통행 제외)는 조사자료에 비해 서울(SK A +2.4%, SKT B +5.2%)의 비율이 높게 나타났고, KT 데이터는 낮게(내부통행 포함 - 1.6%, 내부통행 제외 - 1.0%) 나타남



〈그림 3-6〉 (통행발생량) 귀가 - 시도별 통행 비율

○ 총량 측면에서 비교해보면 조사자료에 비해 모바일 데이터의 통행량은 매우 적은 편임

- 2019년 가구통행실태조사의 1/2, 2021년 개인통행실태조사의 1/3 수준임
- 조사자료에는 내부통행이 포함되어 있으므로, 내부통행이 포함되어 있는 KT 데이터와 비교함

〈표 3-13〉 (통행발생량) 귀가 - 데이터별 시도별 통행량

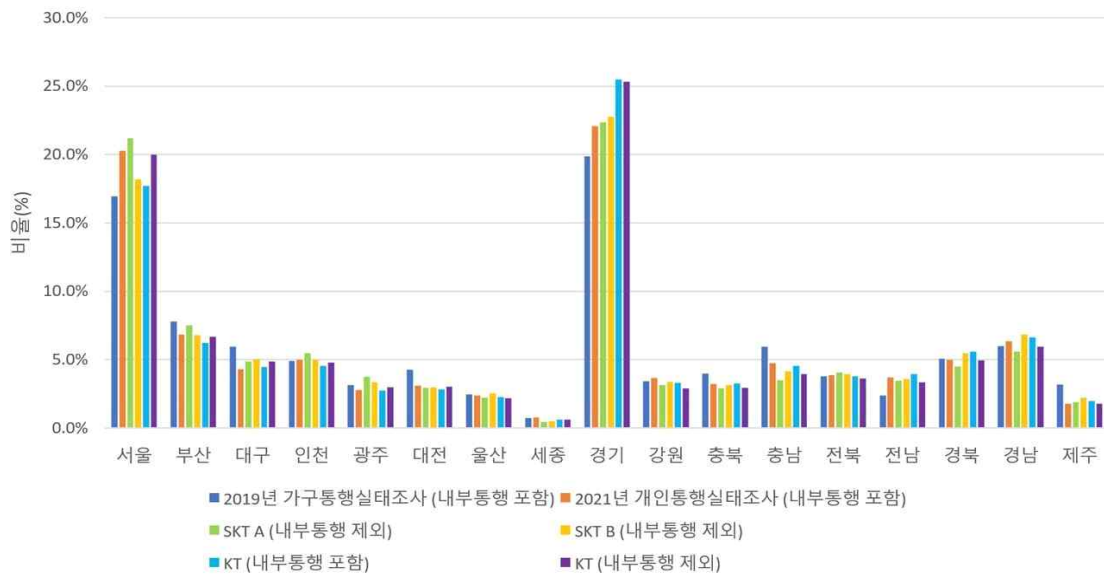
(단위: 통행, %)

시도	2019년 가구통행실태조사 (내부통행 포함)		2021년 개인통행실태조사 (내부통행 포함)		SKT A (내부통행 제외)		SKT B (내부통행 제외)		KT (내부통행 포함)		KT (내부통행 제외)	
	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율
서울	8,357,915	21.4	10,686,412	19.6	8,791,890	21.9	7,434,575	24.8	3,528,475	18.0	3,149,643	20.5
부산	3,014,738	7.7	3,831,550	7.0	2,886,779	7.2	2,305,079	7.7	1,239,910	6.3	1,034,654	6.7
대구	1,804,461	4.6	2,657,574	4.9	1,958,742	4.9	1,548,564	5.2	898,364	4.6	754,634	4.9
인천	1,834,852	4.7	2,773,216	5.1	2,271,545	5.7	1,737,082	5.8	857,522	4.4	707,523	4.6
광주	1,394,988	3.6	1,513,798	2.8	1,233,569	3.1	1,001,594	3.3	541,097	2.8	455,111	3.0
대전	1,286,784	3.3	1,734,804	3.2	1,208,001	3.0	913,921	3.0	564,163	2.9	473,461	3.1
울산	1,039,324	2.7	1,320,110	2.4	936,680	2.3	671,019	2.2	442,027	2.3	335,430	2.2
세종	175,063	0.4	433,206	0.8	197,966	0.5	122,341	0.4	128,768	0.7	100,908	0.7
경기	8,277,205	21.2	12,784,549	23.4	9,483,963	23.7	6,909,421	23.0	5,055,049	25.7	3,866,784	25.2
강원	1,258,059	3.2	1,610,172	2.9	1,127,754	2.8	709,528	2.4	627,371	3.2	436,920	2.8
충북	1,216,087	3.1	1,913,317	3.5	1,162,759	2.9	757,268	2.5	639,891	3.3	460,727	3.0
충남	1,692,786	4.3	2,687,664	4.9	1,480,544	3.7	900,465	3.0	898,796	4.6	621,488	4.1
전북	1,284,093	3.3	1,897,642	3.5	1,396,347	3.5	962,352	3.2	743,369	3.8	561,123	3.7
전남	1,282,806	3.3	1,819,101	3.3	1,240,619	3.1	818,338	2.7	764,625	3.9	509,251	3.3
경북	2,038,008	5.2	2,809,602	5.1	1,861,923	4.6	1,233,932	4.1	1,100,979	5.6	770,586	5.0
경남	2,383,541	6.1	3,427,440	6.3	2,352,823	5.9	1,629,565	5.4	1,297,726	6.6	890,419	5.8
제주	640,931	1.6	694,526	1.3	486,444	1.2	379,117	1.3	305,248	1.6	209,452	1.4
계	38,981,639	100.0	54,594,684	100.0	40,078,347	100.0	30,034,158	100.0	19,633,381	100.0	15,338,112	100.0

⑤ 기타

○ 2019년 가구통행실태조사와 모바일 통신데이터의 차이가 전체 통행, 통근 통행, 통학 통행, 귀가 통행과 비교했을 때 가장 크게 나타남

- 통행 비율 측면에서 전체 통행이 0.5%, 통근 통행이 0.5%, 통학 통행이 0.9%, 귀가 통행이 0.7%를 보일 때, 기타 통행은 1.0% 차이를 보임



〈그림 3-7〉 (통행발생량) 기타 - 시도별 통행 비율

- 다른 목적 통행에서는 대체로 2019년 가구통행실태조사자료보다 모바일 데이터의 통행량이 비해 적은 편이었으나, 기타 통행에서는 모바일 통행량이 2019년 가구통행실태조사자료보다 전반적으로 높게 나타남
- 서울, 경기 지역의 경우 SKT A(내부통행 제외)는 약 3배 높게, SKT B(내부통행 제외)는 약 1.6배 높게, KT(내부통행 포함, 제외)는 약 2배 높게 집계되었으며, 경기 지역의 경우 SKT A, KT 데이터 모두 약 2배 이상 높게 집계됨

〈표 3-14〉 (통행발생량) 기타 - 데이터별 시도별 통행량

(단위: 통행, %)

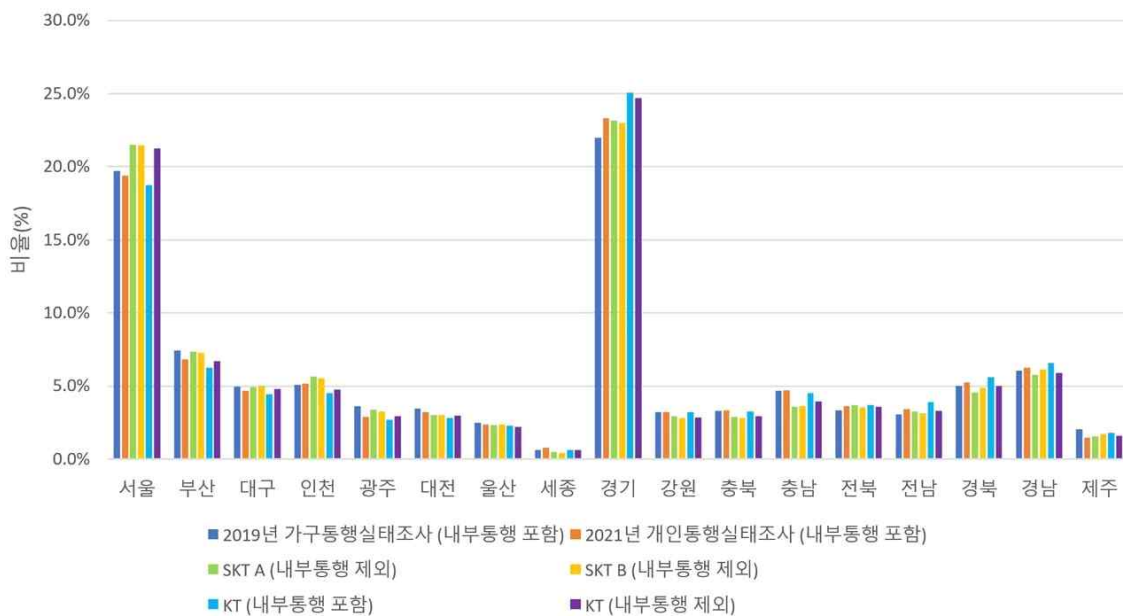
시도	2019년 가구통행실태조사 (내부통행 포함)		2021년 개인통행실태조사 (내부통행 포함)		SKT A (내부통행 제외)		SKT B (내부통행 제외)		KT (내부통행 포함)		KT (내부통행 제외)	
	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율
서울	4,086,002	17.0	12,824,205	20.3	12,135,142	21.2	6,693,213	18.2	9,898,088	17.7	8,866,497	20.0
부산	1,885,663	7.8	4,341,226	6.9	4,295,855	7.5	2,505,545	6.8	3,478,082	6.2	2,960,224	6.7
대구	1,434,546	6.0	2,737,971	4.3	2,799,915	4.9	1,862,132	5.1	2,499,783	4.5	2,156,626	4.9
인천	1,187,879	4.9	3,169,337	5.0	3,142,812	5.5	1,842,586	5.0	2,536,855	4.5	2,125,995	4.8
광주	762,293	3.2	1,775,618	2.8	2,154,397	3.8	1,237,782	3.4	1,543,472	2.8	1,336,136	3.0
대전	1,033,968	4.3	1,980,611	3.1	1,694,535	3.0	1,096,707	3.0	1,576,909	2.8	1,339,250	3.0
울산	591,689	2.5	1,503,485	2.4	1,278,763	2.2	937,824	2.5	1,270,868	2.3	974,239	2.2
세종	179,572	0.7	498,308	0.8	262,906	0.5	192,224	0.5	353,563	0.6	279,919	0.6
경기	4,787,881	19.9	13,977,905	22.1	12,788,594	22.3	8,376,771	22.7	14,242,527	25.5	11,242,018	25.3
강원	828,989	3.4	2,323,477	3.7	1,793,305	3.1	1,248,890	3.4	1,856,577	3.3	1,299,957	2.9
충북	964,203	4.0	2,050,183	3.2	1,665,864	2.9	1,159,104	3.1	1,828,086	3.3	1,316,935	3.0
충남	1,431,989	5.9	3,001,902	4.7	2,011,494	3.5	1,527,908	4.1	2,539,274	4.5	1,751,985	3.9
전북	915,237	3.8	2,453,515	3.9	2,339,350	4.1	1,454,037	3.9	2,119,865	3.8	1,614,353	3.6
전남	579,631	2.4	2,343,811	3.7	1,990,420	3.5	1,328,608	3.6	2,210,366	4.0	1,484,716	3.3
경북	1,220,026	5.1	3,173,640	5.0	2,586,256	4.5	2,020,062	5.5	3,124,677	5.6	2,203,811	5.0
경남	1,442,227	6.0	4,020,436	6.4	3,196,284	5.6	2,512,118	6.8	3,703,411	6.6	2,643,045	6.0
제주	768,741	3.2	1,120,668	1.8	1,085,293	1.9	827,189	2.2	1,107,948	2.0	785,897	1.8
계	24,100,537	100.0	63,296,296	100.0	57,221,184	100.0	36,822,699	100.0	55,890,350	100.0	44,381,605	100.0

2) 통행도착량 기준

① 전체

○ 통행발생량과 마찬가지로 데이터간 시도별 통행 비율은 약 0.5% 정도로 그 차이가 크지 않은 것으로 나타남(<그림 3-8> 참조)

- 2019년 가구통행실태조사 자료(내부통행 포함)를 기준으로 SKT A(내부통행 제외)가 평균 약 0.59%, SKT B(내부통행 제외)가 약 0.43%, KT(내부통행 포함)가 약 0.64%, KT(내부통행 제외)가 약 0.56% 정도 차이를 보이며, 2021년 개인통행실태조사 자료(내부통행 포함)를 기준으로 SKT A(내부통행 제외)가 약 0.58%, SKT B(내부통행 제외)가 평균 약 0.45%, KT(내부통행 포함)가 0.39%, KT(내부통행 제외)가 평균 약 0.41% 차이를 보임



<그림 3-8> (통행도착량) 전체 - 시도별 통행 비율

○ 내부통행이 포함되어 있는 조사자료(2019년 가구통행실태조사, 2021년 개인통행실태조사)와 모바일 데이터(KT)를 상호비교하였을 때 가장 큰 차이를 보이는 지역은 경기였으며, 2019년 가구통행실태조사보다는 SKT A 데이터로 보정된 2021년 개인통행실태조사와의 차이가 적은 편으로 나타남

- KT데이터의 경기 지역 통행 비율이 2019년 가구통행실태조사 자료보다 3.08%, 2021년 개인통행실태조사 자료보다 1.74% 높게 나타남

○ 총량은 통행발생량과 마찬가지로 2019년 가구통행실태조사(내부통행 포함)과 KT 데이터(내부통행 포함)가 가장 유사하게 나타남

〈표 3-15〉 (통행도착량) 전체 - 데이터별 시도별 통행량

(단위: 통행, %)

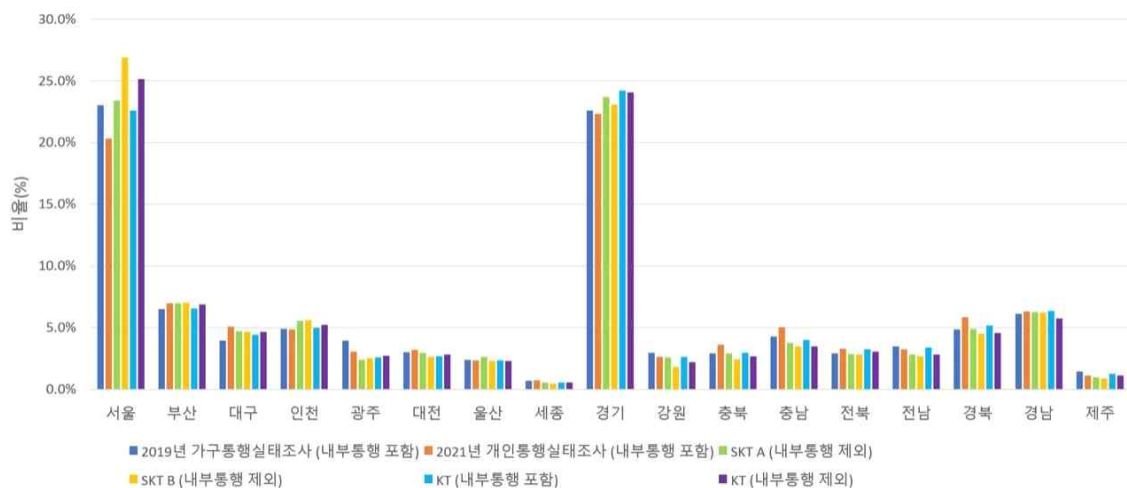
시도	2019년 가구통행실태조사 (내부통행 포함)		2021년 개인통행실태조사 (내부통행 포함)		SKT A (내부통행 제외)		SKT B (내부통행 제외)		KT (내부통행 포함)		KT (내부통행 제외)	
	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율
서울	17,591,441	19.7	28,537,163	19.4	23,808,122	21.5	16,086,083	21.5	16,850,981	18.7	15,253,596	21.3
부산	6,640,080	7.4	10,062,083	6.8	8,144,051	7.3	5,446,394	7.3	5,636,239	6.3	4,802,630	6.7
대구	4,418,369	5.0	6,859,090	4.7	5,441,876	4.9	3,764,306	5.0	3,995,512	4.4	3,430,335	4.8
인천	4,538,195	5.1	7,571,325	5.1	6,248,652	5.6	4,134,688	5.5	4,054,648	4.5	3,404,099	4.7
광주	3,228,518	3.6	4,269,697	2.9	3,735,572	3.4	2,453,483	3.3	2,437,272	2.7	2,098,974	2.9
대전	3,095,012	3.5	4,706,988	3.2	3,320,803	3.0	2,252,942	3.0	2,516,862	2.8	2,138,928	3.0
울산	2,206,629	2.5	3,507,550	2.4	2,566,489	2.3	1,774,530	2.4	2,046,491	2.3	1,580,125	2.2
세종	558,812	0.6	1,170,990	0.8	534,642	0.5	319,078	0.4	552,881	0.6	434,766	0.6
경기	19,619,380	22.0	34,306,547	23.3	25,670,167	23.2	17,238,586	23.0	22,555,613	25.1	17,719,827	24.7
강원	2,870,715	3.2	4,729,894	3.2	3,272,347	3.0	2,118,443	2.8	2,905,693	3.2	2,043,885	2.8
충북	2,929,392	3.3	4,939,237	3.4	3,214,112	2.9	2,114,524	2.8	2,917,603	3.2	2,117,500	3.0
충남	4,168,774	4.7	6,959,576	4.7	3,985,280	3.6	2,716,157	3.6	4,062,077	4.5	2,830,929	3.9
전북	2,995,094	3.4	5,336,633	3.6	4,122,532	3.7	2,660,421	3.5	3,350,303	3.7	2,561,522	3.6
전남	2,716,519	3.0	5,057,657	3.4	3,595,261	3.2	2,365,602	3.2	3,508,917	3.9	2,373,031	3.3
경북	4,448,150	5.0	7,689,728	5.2	5,058,175	4.6	3,647,629	4.9	5,034,471	5.6	3,583,218	5.0
경남	5,417,180	6.1	9,222,034	6.3	6,390,044	5.8	4,598,245	6.1	5,933,350	6.6	4,222,467	5.9
제주	1,810,150	2.0	2,144,202	1.5	1,703,115	1.5	1,283,307	1.7	1,624,059	1.8	1,151,630	1.6
계	89,252,409	100.0	147,070,394	100.0	110,811,240	100.0	74,974,418	100.0	89,982,971	100.0	71,747,462	100.0

② 통근

○ 시도별 통행 비율은 서울 지역을 제외하고는 대체로 데이터간 차이가 크지 않은 것으로 나타남

- 2019년 가구통행실태조사(내부통행 포함), 2021년 개인통행실태조사(내부통행 포함) 결과를 기준으로 모바일 데이터 4종과 비교하였을 때, 서울 지역은 최대 6.54%까지 차이를 보였으며, 데이터간 편차가 큰 편으로 나타남

- 2019년 가구통행실태조사(내부통행 포함)을 기준으로, SKT A(내부통행 제외)는 +0.35%, SKT B(내부통행 제외)는 +3.84%, KT(내부통행 포함)은 -0.42%, KT(내부통행 제외)는 -2.12% 차이를 보임
- 2021년 개인통행실태조사(내부통행 포함)을 기준으로, SKT A(내부통행 제외)는 +3.05%, SKT B(내부통행 제외)는 +6.54%, KT(내부통행 포함)은 +2.28%, KT(내부통행 제외)는 +4.82% 차이를 보임



〈그림 3-9〉 (통행도착량) 통근 - 시도별 통행 비율

○ 조사자료에 비해 모바일 통신데이터의 통행이 다소 낮게 집계된 편임

- 전체 통행, 귀가 통행, 기타 통행은 모바일 통신데이터가 조사자료에 비해 대체로 통행량이 높게 집계된 반면, 통근 통행은 SKT A 데이터, SKT B, KT 데이터 모두 모든 지역에서 조사자료에 비해 낮게 집계됨

〈표 3-16〉 (통행도착량) 통근 - 데이터별 시도별 통행량

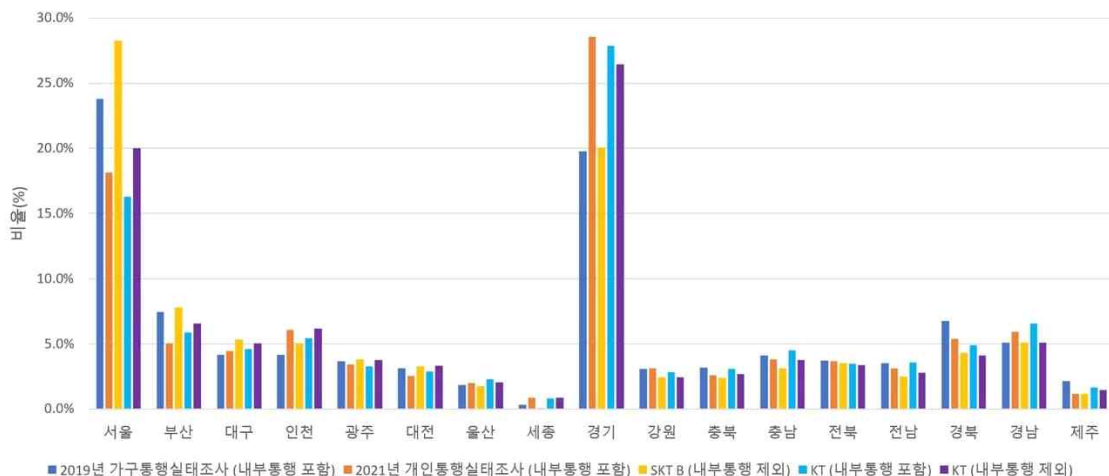
(단위: 통행, %)

시도	2019년 가구통행실태조사 (내부통행 포함)		2021년 개인통행실태조사 (내부통행 포함)		SKT A (내부통행 제외)		SKT B (내부통행 제외)		KT (내부통행 포함)		KT (내부통행 제외)	
	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율
서울	5,279,967	23.0	4,641,363	20.3	3,157,880	23.4	1,935,962	26.9	2,827,928	22.6	2,700,640	25.1
부산	1,489,636	6.5	1,593,109	7.0	944,315	7.0	505,693	7.0	817,929	6.5	738,161	6.9
대구	903,538	3.9	1,157,064	5.1	635,482	4.7	335,889	4.7	554,846	4.4	501,154	4.7
인천	1,122,354	4.9	1,103,824	4.8	750,542	5.6	402,755	5.6	626,109	5.0	562,695	5.2
광주	902,127	3.9	699,220	3.1	326,928	2.4	184,464	2.6	320,841	2.6	291,631	2.7
대전	693,745	3.0	732,092	3.2	399,312	3.0	190,910	2.6	336,330	2.7	300,644	2.8
울산	551,352	2.4	541,655	2.4	357,263	2.6	166,004	2.3	293,161	2.3	248,491	2.3
세종	161,422	0.7	165,052	0.7	75,588	0.6	33,079	0.5	70,725	0.6	58,648	0.5
경기	5,185,223	22.6	5,099,209	22.3	3,200,833	23.7	1,662,005	23.1	3,029,839	24.2	2,585,132	24.1
강원	681,103	3.0	603,957	2.6	351,249	2.6	131,236	1.8	326,683	2.6	235,380	2.2
충북	672,836	2.9	831,618	3.6	391,411	2.9	177,131	2.5	372,326	3.0	287,705	2.7
충남	985,833	4.3	1,145,650	5.0	509,658	3.8	249,552	3.5	499,315	4.0	371,664	3.5
전북	672,517	2.9	754,137	3.3	388,390	2.9	204,813	2.8	405,528	3.2	328,056	3.1
전남	797,662	3.5	745,683	3.3	381,767	2.8	192,417	2.7	426,162	3.4	305,066	2.8
경북	1,108,624	4.8	1,329,574	5.8	662,644	4.9	324,200	4.5	649,565	5.2	492,608	4.6
경남	1,400,182	6.1	1,441,097	6.3	848,518	6.3	447,654	6.2	796,091	6.4	615,596	5.7
제주	330,143	1.4	257,271	1.1	129,928	1.0	63,065	0.9	160,714	1.3	120,183	1.1
계	22,938,262	100.0	22,841,575	100.0	13,511,710	100.0	7,206,829	100.0	12,514,091	100.0	10,743,453	100.0

③ 통학

○ 서울, 경기 지역에서 데이터간 통행 비율 차이가 크게 나타남

- 서울 지역은 2019년 가구통행실태조사(내부통행 포함)를 기준으로 SKT B(내부통행 제외)와 +4.45%, KT(내부통행 포함)과 - 7.52%, KT(내부통행 제외)와 - 3.78% 차이를 보였으며, 2021년 개인통행실태조사(내부통행 포함)를 기준으로 SKT B(내부통행 제외)는 +10.10%, KT(내부통행 포함)과 - 1.86%, KT(내부통행 제외)와 +1.87% 차이를 보임
- 경기 지역은 2019년 가구통행실태조사(내부통행 포함)를 기준으로 SKT B(내부통행 제외)와 +0.27%, KT(내부통행 포함)과 +8.08%, KT(내부통행 제외)와 +6.67% 차이를 보였으며, 2021년 개인통행실태조사(내부통행 포함)를 기준으로 SKT B(내부통행 제외)는 -8.50%, KT(내부통행 포함)과 - 0.69%, KT(내부통행 제외)와 - 2.10% 차이를 보임



〈그림 3-10〉 (통행도착량) 통학 - 시도별 통행 비율

○ <표 3-17>과 같이 통행량으로 비교해보면, 지역별 차이뿐만 아니라 전체 총량에서도 데이터간 차이가 큰 것을 확인할 수 있음

- 동종의 조사자료인 2019년 가구통행실태조사와 2021년 개인통행실태조사자료도 2배 가까이 차이가 나는 것으로 나타났으며, 동일 조건(내부통행 포함)의 모바일 데이터(KT)와 비교했을 때에도 2019년 가구통행실태조사와 약 1.7배, 2021년 개인통행실태조사와 약 3.3배 차이를 보임

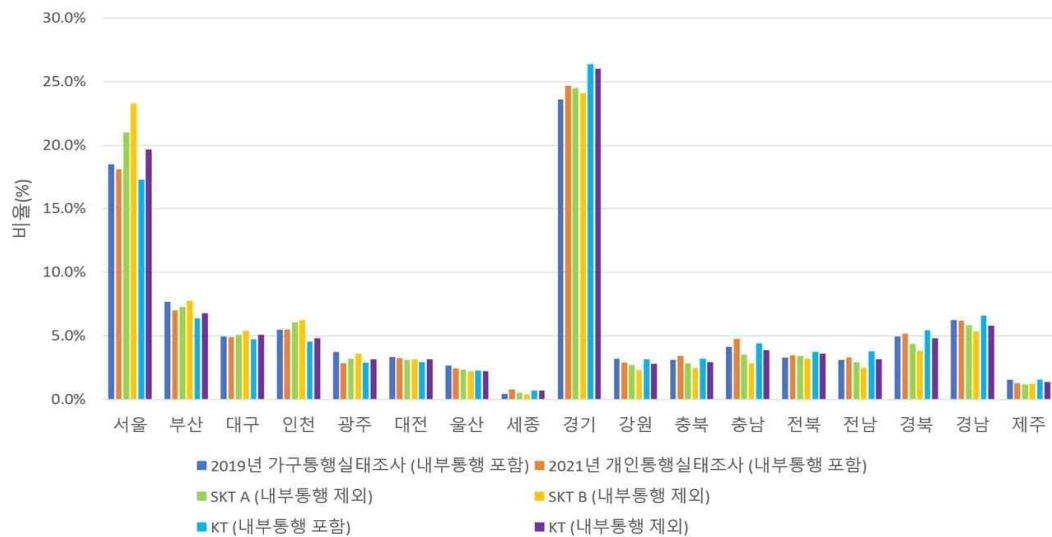
〈표 3-17〉 (통행도착량) 통학 - 데이터별 시도별 통행량

(단위: 통행, %)

시도	2019년 가구통행실태조사 (내부통행 포함)		2021년 개인통행실태조사 (내부통행 포함)		SKT B (내부통행 제외)		KT (내부통행 포함)		KT (내부통행 제외)	
	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율
서울	769,733	23.8	1,150,897	18.2	257,388	28.3	316,948	16.3	257,273	20.0
부산	240,717	7.4	321,090	5.1	70,788	7.8	114,873	5.9	84,147	6.6
대구	134,051	4.1	284,001	4.5	48,608	5.3	89,408	4.6	64,809	5.0
인천	134,843	4.2	385,221	6.1	46,180	5.1	105,395	5.4	79,120	6.2
광주	118,912	3.7	217,504	3.4	34,868	3.8	64,156	3.3	48,390	3.8
대전	100,675	3.1	161,442	2.5	30,066	3.3	56,366	2.9	42,479	3.3
울산	59,898	1.9	127,540	2.0	15,967	1.8	45,084	2.3	26,613	2.1
세종	11,168	0.3	55,655	0.9	915	0.1	15,599	0.8	11,064	0.9
경기	639,649	19.8	1,810,230	28.6	182,730	20.1	542,172	27.9	339,868	26.5
강원	99,490	3.1	197,899	3.1	22,317	2.5	55,112	2.8	31,679	2.5
충북	102,580	3.2	164,183	2.6	21,618	2.4	59,816	3.1	34,649	2.7
충남	132,531	4.1	241,145	3.8	28,519	3.1	87,281	4.5	48,381	3.8
전북	119,845	3.7	233,729	3.7	31,988	3.5	67,246	3.5	43,694	3.4
전남	114,288	3.5	197,802	3.1	22,595	2.5	69,312	3.6	35,548	2.8
경북	218,668	6.8	341,238	5.4	39,121	4.3	95,645	4.9	52,607	4.1
경남	165,222	5.1	375,023	5.9	46,481	5.1	128,151	6.6	65,435	5.1
제주	69,701	2.2	73,240	1.2	10,583	1.2	32,584	1.7	18,534	1.4
계	3,231,971	100.0	6,337,838	100.0	910,733	100.0	1,945,149	100.0	1,284,292	100.0

④ 귀가

- 타 목적 통행에 비해 데이터간 지역별 통행 비율 차이가 크지 않은 편임
 - 서울, 경기 지역을 제외하고는 약 0.4% 정도의 근소한 차이를 보임
- 데이터간 차이가 가장 크게 나타나는 서울을 기준으로 보면, 2019년 가구통행실태조사(내부통행 포함), 2021년 개인통행실태조사(내부통행 포함), KT(내부통행 포함, 제외) 데이터보다 SKT A 데이터(내부통행 제외), SKT B 데이터(내부통행 제외)가 다소 높은 비율을 보임



〈그림 3-11〉 (통행도착량) 귀가 - 시도별 통행 비율

- 도착량을 기준으로 보면 전반적으로 KT 기반 귀가 통행량이 상대적으로 작게 추정되는 것을 확인할 수 있음
 - 2019년 가구통행실태조사(내부통행 포함)는 2021년 개인통행실태조사(내부통행 포함)와 약 15백만통행 정도 차이를 보이며, KT(내부통행 포함)은 2021년 개인통행실태조사(내부통행 포함)과 약 34백만통행 정도 차이를 보임
 - KT기반 통행정보 가공 알고리즘의 경우, 더 정확한 통근통학 시간 산정을 위하여 회사나 학교에서 귀가 중 1개 이상의 다른 활동을 하는 경우 통행시간 산정에서 제외함. 이러한 알고리즘의 특성이 반영된 결과라고 판단됨¹⁾

1) 원민수, 최정윤, 이해선, & 김주영. (2021). 모바일기지국 데이터를 이용한 출퇴근 통행 분석 알고리즘 개발: 집, 직장 추

〈표 3-18〉 (통행도착량) 귀가 - 데이터별 시도별 통행량

(단위: 통행, %)

시도	2019년 가구통행실태조사 (내부통행 포함)		2021년 개인통행실태조사 (내부통행 포함)		SKT A (내부통행 제외)		SKT B (내부통행 제외)		KT (내부통행 포함)		KT (내부통행 제외)	
	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율
서울	7,213,878	18.5	9,880,158	18.1	8,429,388	21.0	6,995,571	23.3	3,398,914	17.3	3,020,081	19.7
부산	2,986,708	7.7	3,827,651	7.0	2,922,779	7.3	2,340,426	7.8	1,249,119	6.4	1,043,864	6.8
대구	1,921,133	4.9	2,681,383	4.9	2,042,641	5.1	1,620,160	5.4	924,597	4.7	780,867	5.1
인천	2,147,866	5.5	2,987,321	5.5	2,427,553	6.1	1,878,174	6.3	891,188	4.5	741,188	4.8
광주	1,455,655	3.7	1,548,232	2.8	1,290,109	3.2	1,077,713	3.6	569,172	2.9	483,186	3.2
대전	1,298,502	3.3	1,765,951	3.2	1,255,211	3.1	954,084	3.2	572,704	2.9	482,002	3.1
울산	1,033,635	2.7	1,328,542	2.4	944,264	2.4	669,970	2.2	448,119	2.3	341,522	2.2
세종	169,542	0.4	439,321	0.8	199,536	0.5	120,479	0.4	137,933	0.7	110,073	0.7
경기	9,201,447	23.6	13,480,235	24.7	9,820,226	24.5	7,236,218	24.1	5,184,410	26.4	3,996,145	26.1
강원	1,256,024	3.2	1,588,949	2.9	1,084,752	2.7	695,078	2.3	617,164	3.1	426,712	2.8
충북	1,207,192	3.1	1,881,122	3.4	1,132,462	2.8	745,262	2.5	626,125	3.2	446,961	2.9
충남	1,620,225	4.2	2,600,560	4.8	1,411,347	3.5	858,340	2.9	869,077	4.4	591,769	3.9
전북	1,280,802	3.3	1,891,840	3.5	1,376,905	3.4	962,352	3.2	735,540	3.7	553,294	3.6
전남	1,207,288	3.1	1,791,129	3.3	1,180,058	2.9	749,559	2.5	740,730	3.8	485,355	3.2
경북	1,931,585	5.0	2,815,001	5.2	1,755,032	4.4	1,153,921	3.8	1,067,406	5.4	737,013	4.8
경남	2,438,280	6.3	3,397,253	6.2	2,334,559	5.8	1,609,823	5.4	1,295,972	6.6	888,665	5.8
제주	611,875	1.6	690,036	1.3	471,523	1.2	367,027	1.2	305,211	1.6	209,415	1.4
계	38,981,639	100.0	54,594,684	100.0	40,078,347	100.0	30,034,158	100.0	19,633,381	100.0	15,338,112	100.0

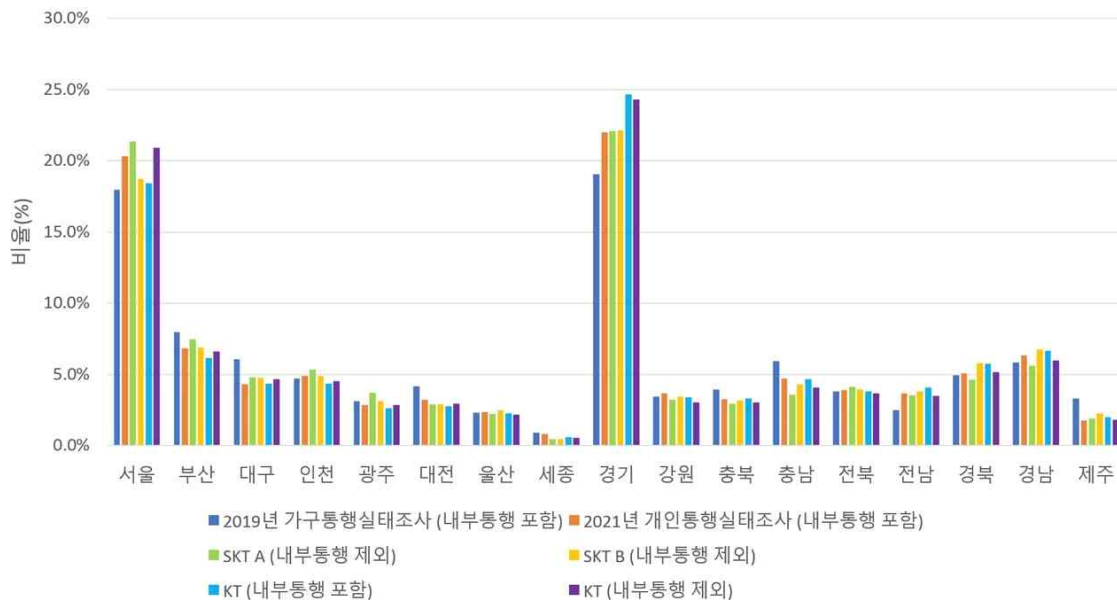
정을 중심으로. 대한교통학회지, 39(3), 383-398.

⑤ 기타

○ 시도별 통행 비율은 타 지역보다 서울, 경기 충남 지역에서 데이터간 다소 큰 차이를 보임

- 대체로 서울, 경기 지역은 조사자료가 모바일 데이터의 통행 비율보다 높고, 충남 지역은 조사자료보다 모바일 데이터의 통행 비율이 높은 것으로 나타났으며, SKT A 데이터로 보정된 2021년 개인통행실태조사보다 2019년 가구통행실태조사와 모바일 데이터간의 차이가 큰 것으로 나타남

· 2019년 가구통행실태조사(내부통행 포함)를 기준으로, SKT A 데이터(내부통행 제외)는 서울 지역에서 약 +3.40%, 경기 지역에서 약 +3.05% 충남 지역에서 약 -2.33% 차이를 보이고, SKT B 데이터(내부통행 제외)는 서울 지역에서 약 +0.77%, 경기 지역에서 +3.10%, 충남 지역에서 -1.64%, 차이를 보이며, KT 데이터(내부통행 포함)는 서울 지역에서 +0.48%, 경기 지역에서 +5.63%, 충남 지역에서 -1.27% 차이를 보임



〈그림 3-12〉 (통행도착량) 기타 - 시도별 통행 비율

○ 동일 조건(내부통행 포함)에서 총량이 상대적으로 유사한 것은 2021년 개인통행 실태조사와 KT 데이터로 확인됨

- 2019년 가구통행실태조사는 2021년 개인통행실태조사, KT 데이터의 약 1/3 수준으로 총량에서는 큰 차이를 보이는 것으로 나타남

〈표 3-19〉 (통행도착량) 기타 - 데이터별 시도별 통행량

(단위: 통행, %)

시도	2019년 가구통행실태조사 (내부통행 포함)		2021년 개인통행실태조사 (내부통행 포함)		SKT A (내부통행 제외)		SKT B (내부통행 제외)		KT (내부통행 포함)		KT (내부통행 제외)	
	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율	통행량	비율
서울	4,327,862	18.0	12,864,745	20.3	12,220,853	21.4	6,897,162	18.7	10,307,191	18.4	9,275,601	20.9
부산	1,923,020	8.0	4,320,233	6.8	4,276,957	7.5	2,529,487	6.9	3,454,317	6.2	2,936,458	6.6
대구	1,459,648	6.1	2,736,643	4.3	2,763,752	4.8	1,759,649	4.8	2,426,661	4.3	2,083,505	4.7
인천	1,133,132	4.7	3,094,958	4.9	3,070,556	5.4	1,807,578	4.9	2,431,956	4.4	2,021,097	4.6
광주	751,823	3.1	1,804,741	2.9	2,118,535	3.7	1,156,438	3.1	1,483,103	2.7	1,275,767	2.9
대전	1,002,089	4.2	2,047,503	3.2	1,666,280	2.9	1,077,882	2.9	1,551,462	2.8	1,313,803	3.0
울산	561,745	2.3	1,509,813	2.4	1,264,962	2.2	922,589	2.5	1,260,128	2.3	963,499	2.2
세종	216,680	0.9	510,963	0.8	259,518	0.5	164,605	0.4	328,624	0.6	254,980	0.6
경기	4,593,061	19.1	13,916,873	22.0	12,649,107	22.1	8,157,634	22.2	13,799,192	24.7	10,798,682	24.3
강원	834,098	3.5	2,339,089	3.7	1,836,346	3.2	1,269,811	3.4	1,906,733	3.4	1,350,113	3.0
충북	946,785	3.9	2,062,315	3.3	1,690,238	3.0	1,170,513	3.2	1,859,336	3.3	1,348,185	3.0
충남	1,430,185	5.9	2,972,221	4.7	2,064,275	3.6	1,579,745	4.3	2,606,403	4.7	1,819,115	4.1
전북	921,930	3.8	2,456,927	3.9	2,357,236	4.1	1,461,269	4.0	2,141,991	3.8	1,636,479	3.7
전남	597,280	2.5	2,323,044	3.7	2,033,436	3.6	1,401,032	3.8	2,272,713	4.1	1,547,062	3.5
경북	1,189,273	4.9	3,203,914	5.1	2,640,498	4.6	2,130,387	5.8	3,221,855	5.8	2,300,990	5.2
경남	1,413,495	5.9	4,008,661	6.3	3,206,967	5.6	2,494,287	6.8	3,713,136	6.6	2,652,771	6.0
제주	798,430	3.3	1,123,654	1.8	1,101,664	1.9	842,631	2.3	1,125,549	2.0	803,498	1.8
계	24,100,537	100.0	63,296,296	100.0	57,221,184	100.0	36,822,699	100.0	55,890,350	100.0	44,381,605	100.0

나. 시군구 단위

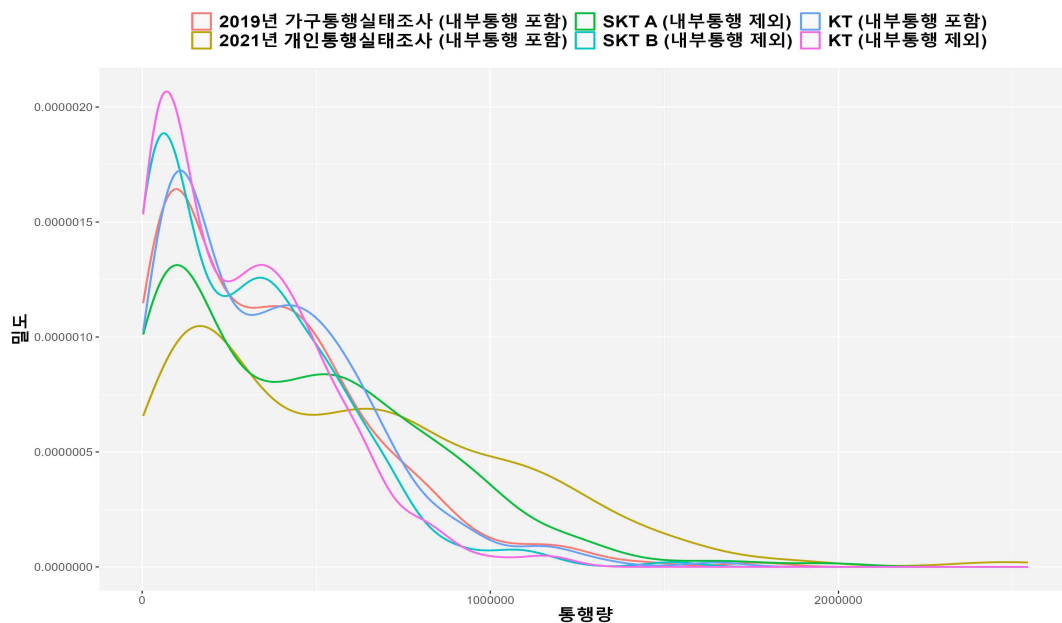
1) 통행발생량 기준

① 전체

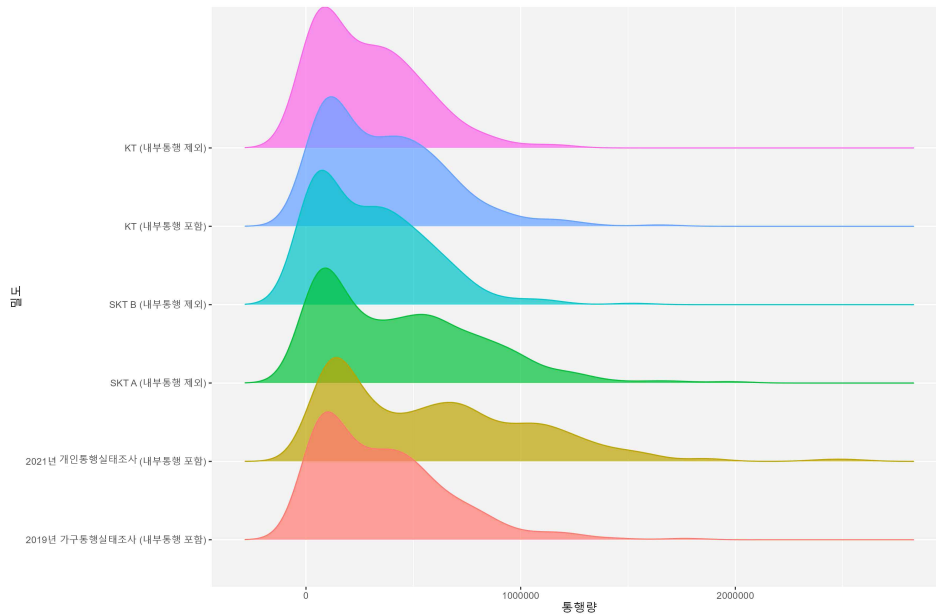
A) 데이터별 통행량 비교

○ 데이터별로 시군구별 통행량 분포를 비교한 결과는 다음 <그림 3-13>, <그림 3-14>와 같음

- 모든 데이터에서 1개의 변곡점이 나타나, 변곡점이 나타나는 기점은 각각 다르게 나타남
- 2021년 개인통행실태조사(내부통행 포함), SKT A 데이터(내부통행 제외)를 제외하고 나머지 데이터는 95% 이상의 통행량이 백만 통행 이하에서 나타남

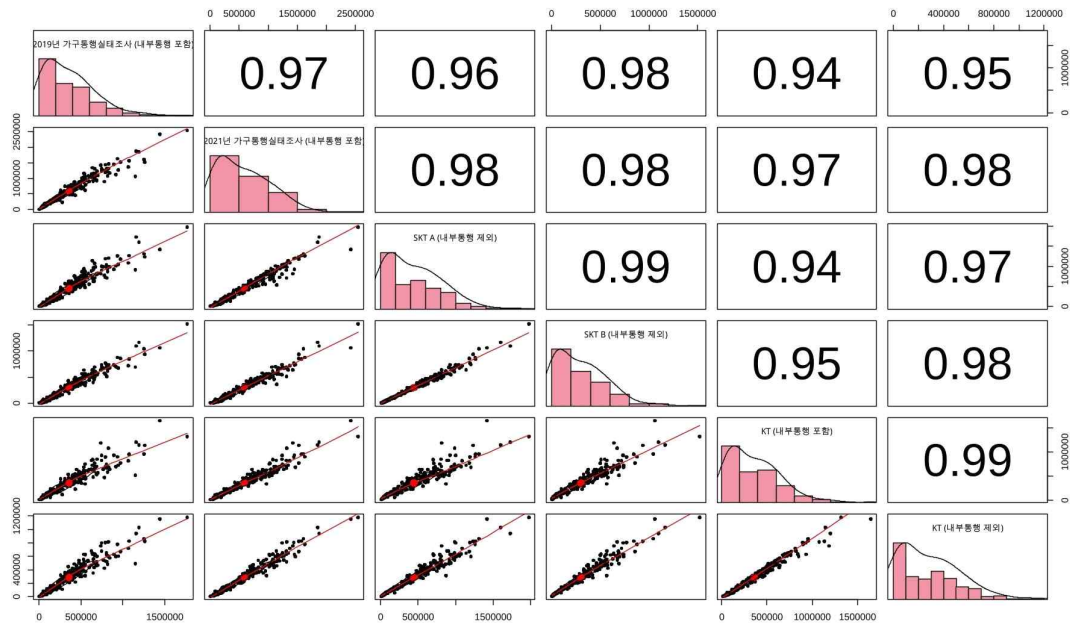


<그림 3-13> (통행발생량) 전체 - 데이터별 통행량 밀도 비교 (1)



〈그림 3-14〉 (통행발생량) 전체 -데이터별 통행량 밀도 비교 (2)

- 각 데이터간 유사성을 분석한 그래프는 다음 <그림 3-15>와 같음
- 통행실태조사(2019년, 2021년)와 모바일 통신데이터는 전체적으로 높은 상관 관계를 갖는 것으로 분석됨



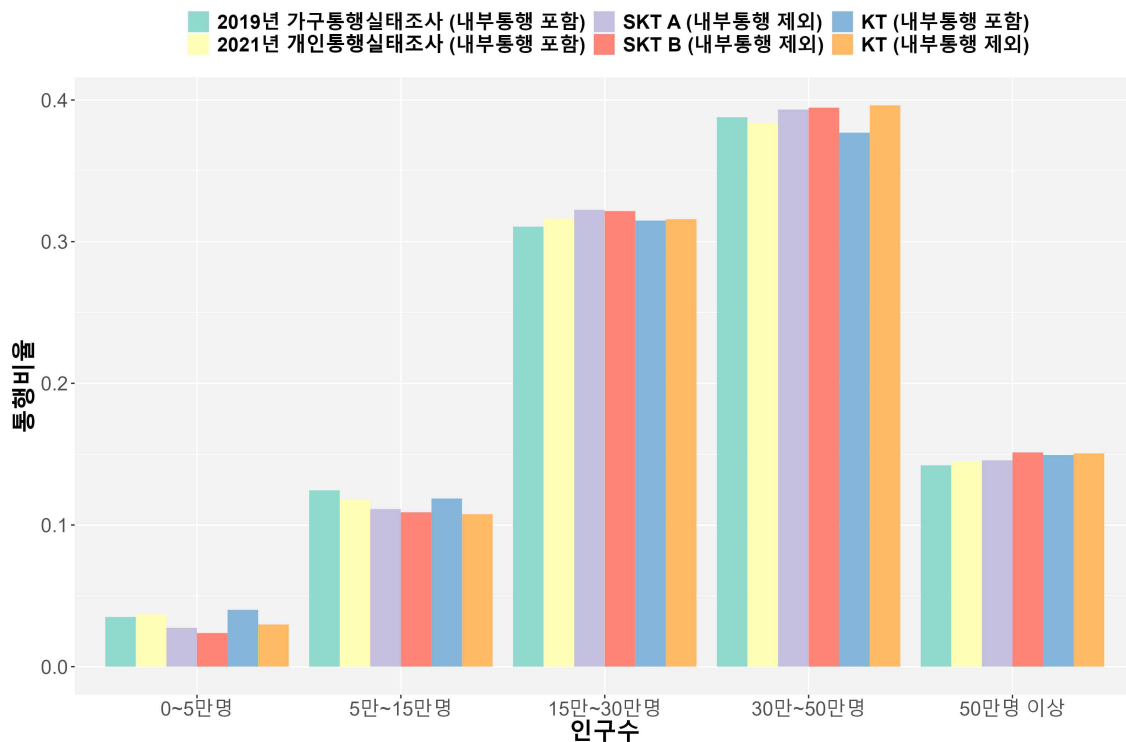
주: 그림에 표기된 수치는 R(상관계수)값을 의미함.

〈그림 3-15〉 (통행발생량) 전체 - 데이터간 상관관계 비교

B) 통행발생 요인별 통행량 비교

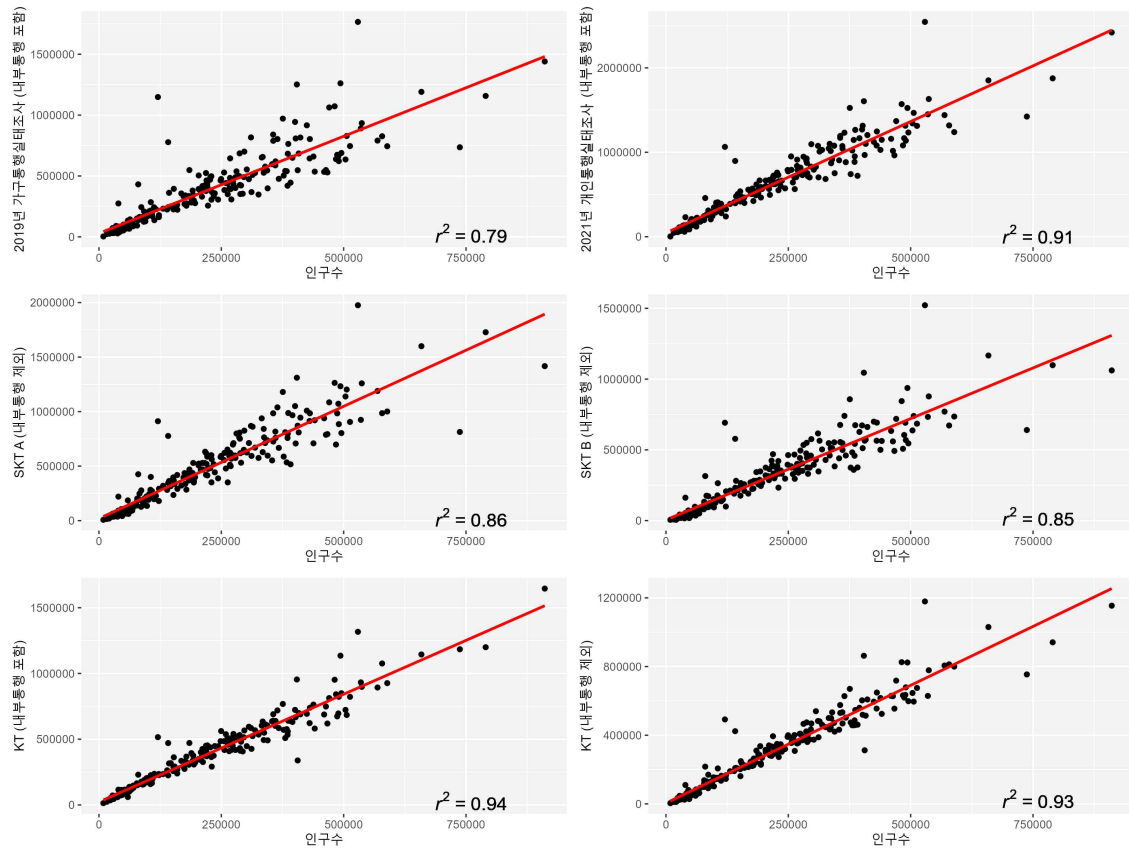
○ 인구수

- 인구 규모에 따른 통행비율 분포는 데이터간 큰 차이가 없는 것으로 나타남
- 15만 명 미만의 인구 규모에서는 2019년 가구통행실태조사(내부통행 포함), 2021년 개인통행실태조사(내부통행 포함), KT(내부통행 포함) 데이터의 통행 비율이 타 데이터 보다 약간 높게 집계되었으며, 15만 명 이상부터 50만 명 미만의 인구 규모에서는 2019년 가구통행실태조사(내부통행 포함), 2021년 개인통행실태조사(내부통행 포함), KT(내부통행 포함) 데이터의 통행 비율이 타 데이터보다 약간 낮게 집계됨



<그림 3-16> (통행발생량) 전체 통행 - 인구수에 따른 데이터별 통행량 비교

- 각 데이터의 시군구별 통행발생량과 인구수와의 관계를 비교한 결과는 다음 <그림 3-17>과 같음
- 데이터별 다소 차이는 있으나 전반적으로 인구수에 대한 설명력이 높은 것으로 나타남

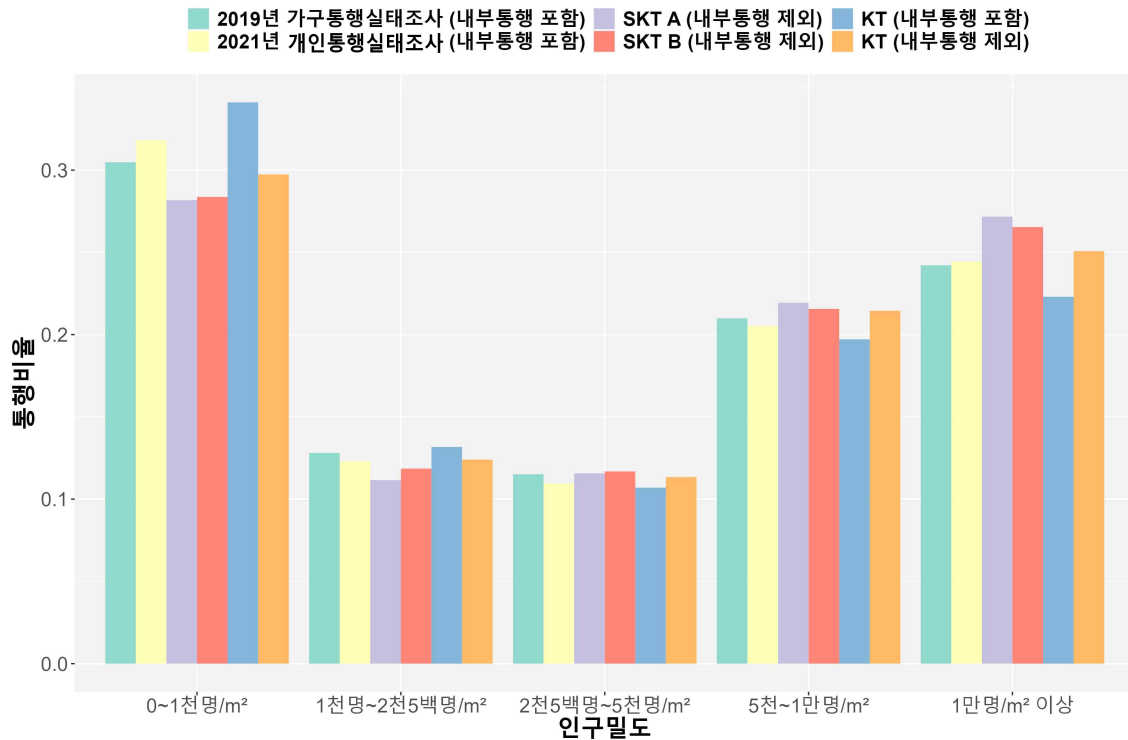


〈그림 3-17〉 (통행발생량) 전체 통행 - 인구수와 각 데이터별 상관관계

○ 인구밀도

- <그림 3-18>은 데이터별로 시군구별 인구밀도에 따른 통행량 결과를 나타낸 것임

· 인구밀도가 1천명/㎡ 미만, 1만명/㎡ 이상일 때 데이터간 통행 비율 차이가 큰 것으로 나타남

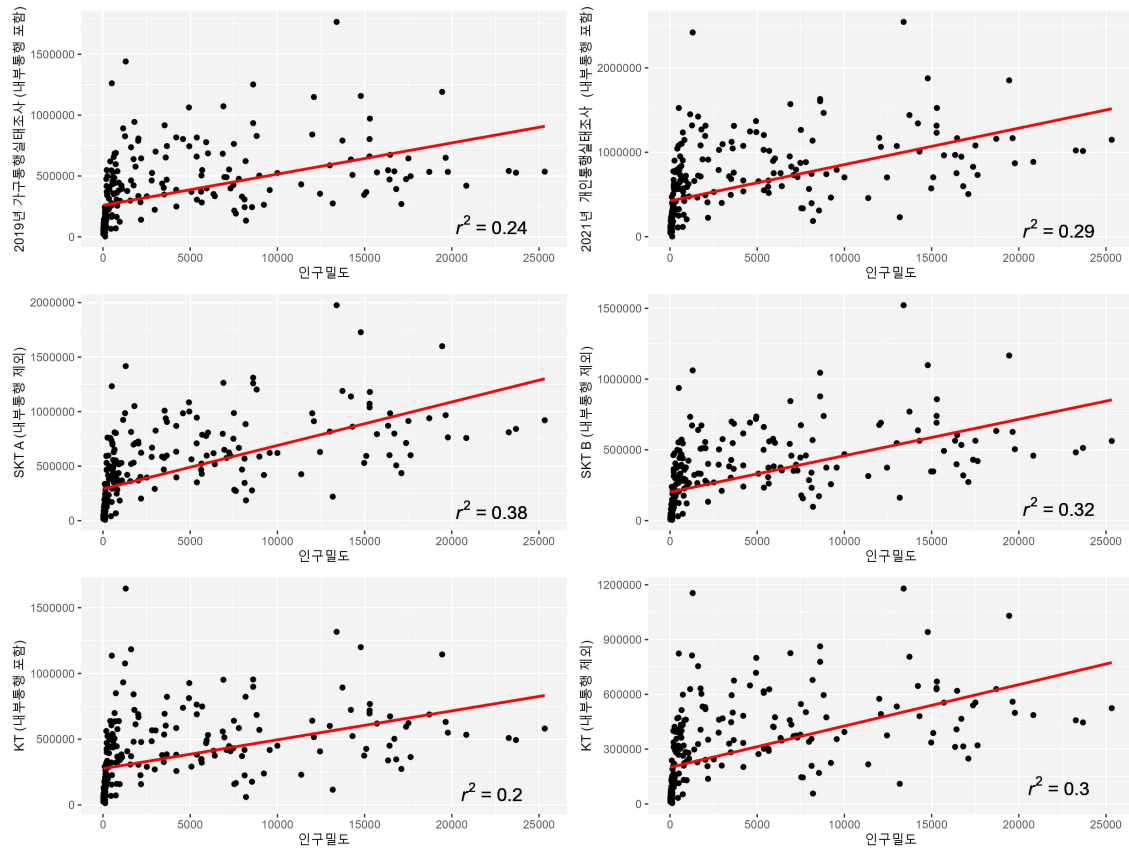


<그림 3-18> (통행발생량) 전체 통행 - 인구밀도에 따른 데이터별 통행량 비교

- 각 데이터의 시군구별 통행발생량과 인구밀도와의 관계를 비교한 결과는 다음

<그림 3-19>와 같음

· 전반적으로 인구밀도에 대한 설명력이 낮게 나타남

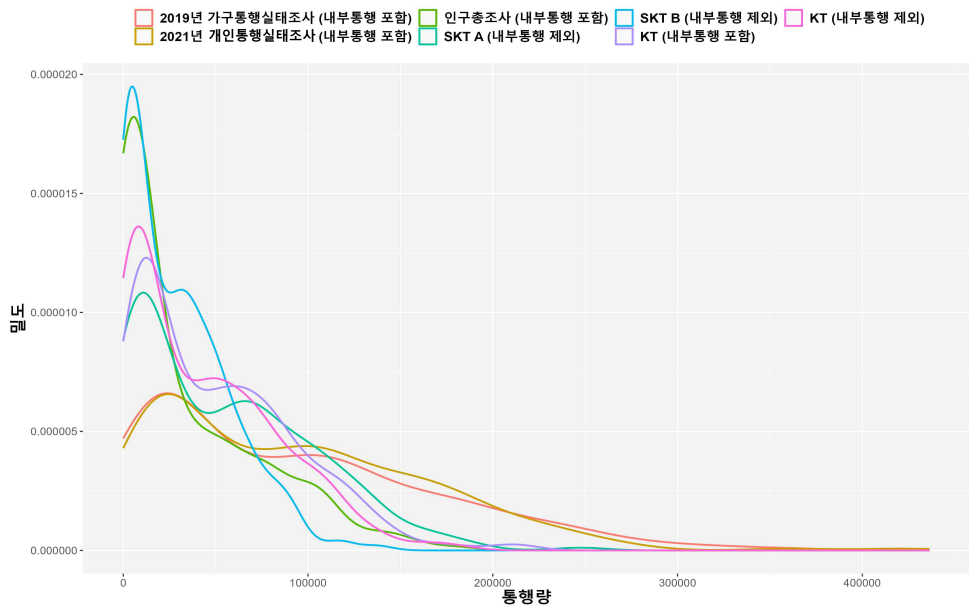


〈그림 3-19〉 (통행발생량) 전체 통행 - 인구밀도와 각 데이터별 상관관계

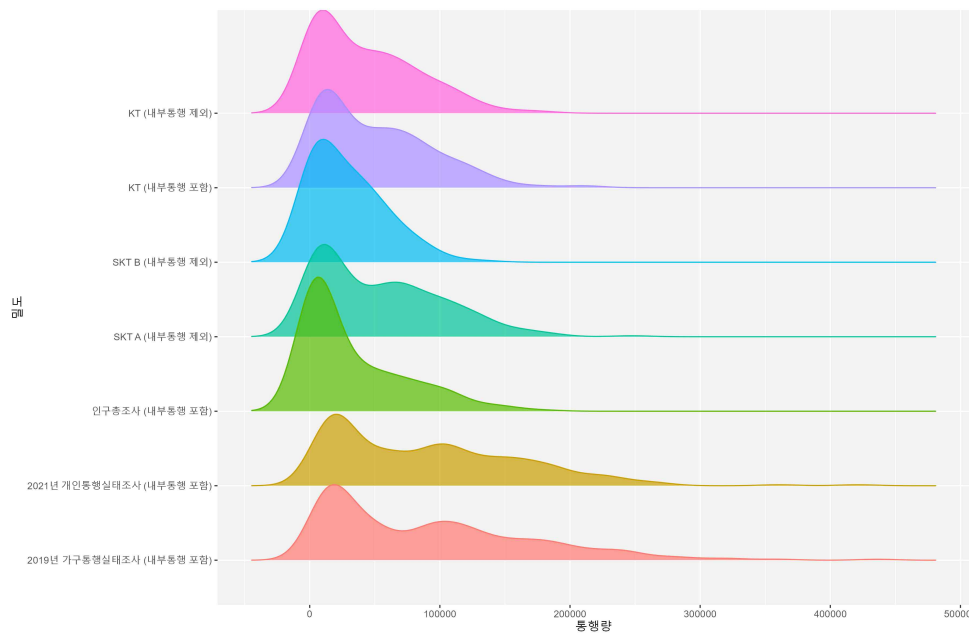
② 통근

A) 데이터별 통행량 비교

○ 다음 <그림 3-20>과 <그림 3-21>은 데이터별 통행량 빈도를 나타낸 것임



<그림 3-20> (통행발생량) 통근 통행 - 데이터별 통행량 빈도 비교 (1)

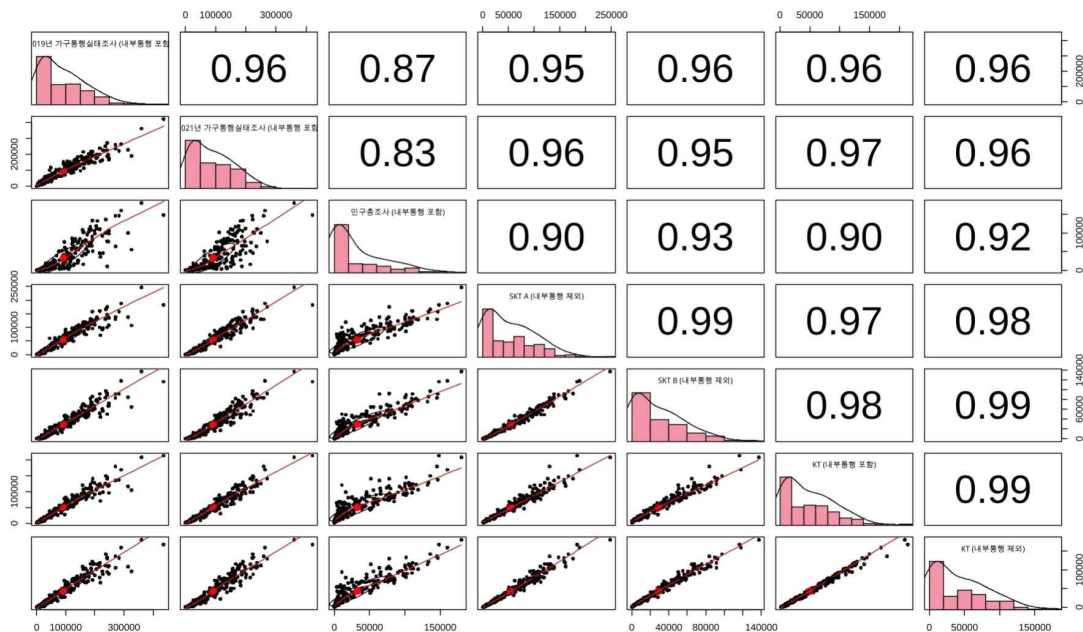


<그림 3-21> (통행발생량) 통근 통행 - 데이터별 통행량 빈도 비교 (2)

- 5만 통행 이하에서는 데이터간 분포 차이가 크게 나타남
 - SKT B(내부통행 제외)와 SKT A(내부통행 제외)는 낮은 통행량의 빈도가 집중적으로 높게 나타나는 반면 2019년 가구통행실태조사(내부통행 포함)와 2021년 개인통행실태조사(내부통행 포함)은 그보다 높은 통행량에서 빈도가 높게 나타나고 대체로 고루 분포하는 것을 알 수 있음

○ 다음 <그림 3-22>는 통근 통행에 대한 데이터간 상관관계를 비교한 것임

- 인구총조사와 가구통행실태조사의 상관관계는 타 데이터간의 상관관계보다 상대적으로 낮은 편이나, 전반적으로 데이터간 상관관계가 높게 나타남



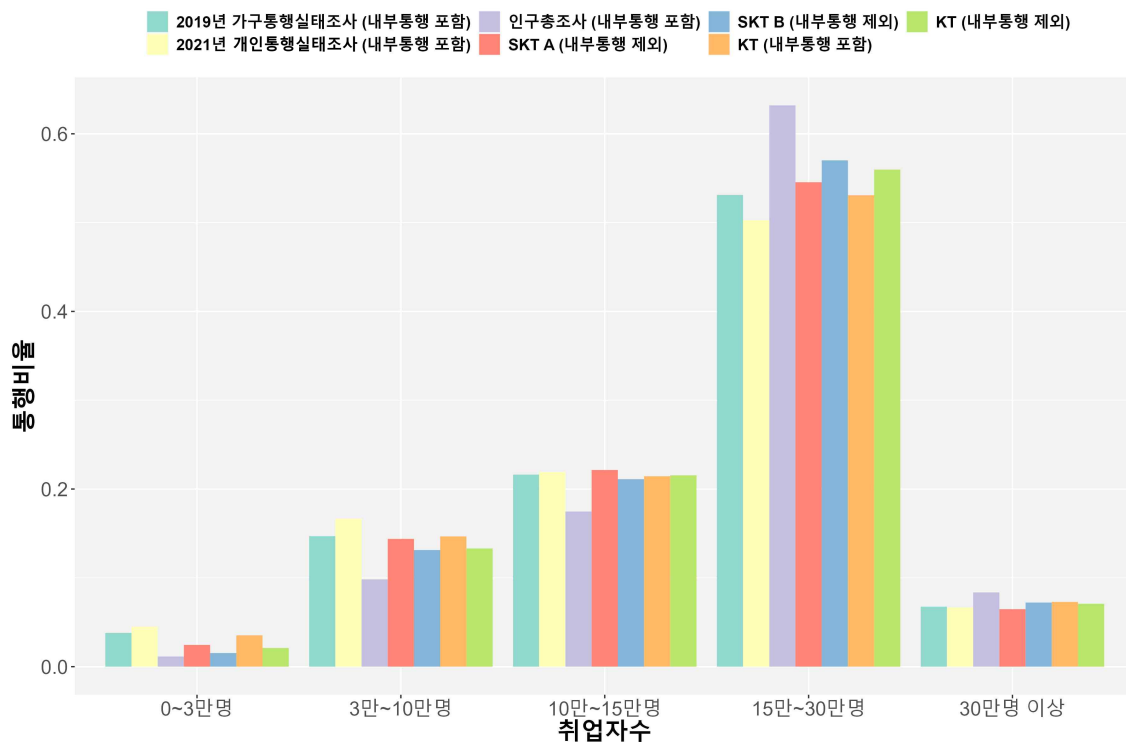
주: 그림에 표기된 수치는 R(상관계수)값을 의미함.

〈그림 3-22〉 (통행발생량) 통근 통행 - 데이터간 상관관계 비교

B) 통행발생 요인별 통행량 비교

○ 취업자수

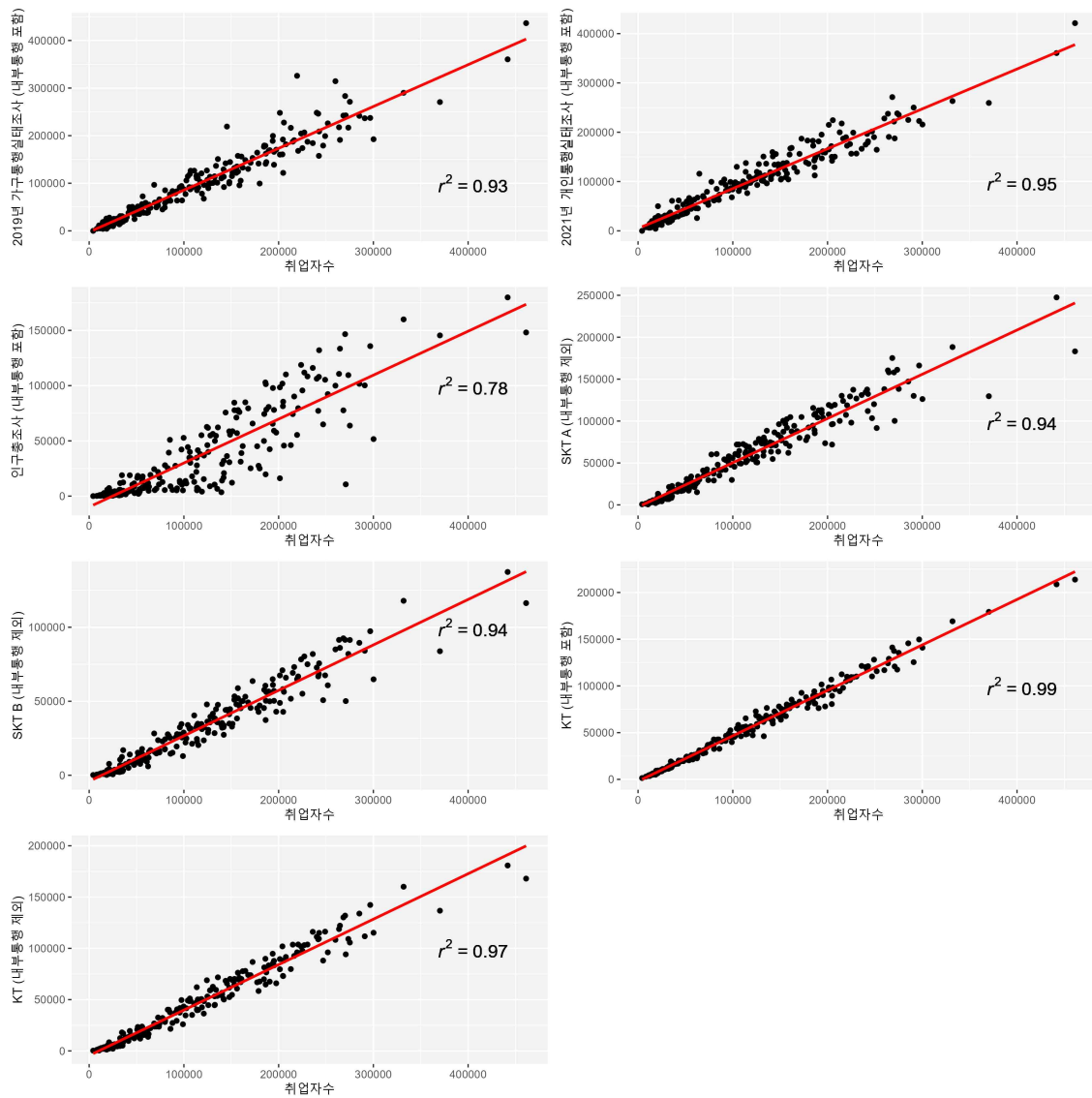
- <그림 3-23>은 취업자수에 따른 통근 통행 발생비율 변화를 데이터별로 비교한 것임
 - 인구총조사를 제외한 나머지 데이터는 분포가 대체로 유사한 편이며, 그 중에서도 2019년 가구통행실태조사(내부통행 포함)과 KT(내부통행 포함)의 통행 비율 분포가 가장 유사한 편임
 - 2021년 개인통행실태조사(내부통행 포함)는 15만 명 미만의 규모에서는 통행 비율이 타 데이터에 비해 높은 편이었으나, 15만 명 이상 30만 명 미만 구간에서 타 데이터에 비해 통행 비율이 다소 낮은 것으로 확인됨
 - SKT B(내부통행 제외), KT(내부통행 제외)는 이와 반대로 15만 명 미만의 규모에서 통행 비율이 낮은 편이었으나, 15만 명 이상 30만 명 미만 구간에서 통행 비율이 높은 것으로 분석됨



〈그림 3-23〉 (통행발생량) 통근 통행 - 취업자수에 따른 데이터별 통행량 비교

- <그림 3-24>는 각 데이터의 통근 통행발생량과 취업자수와의 관계를 비교한 것임
 - 인구총조사 자료를 제외한, 나머지 데이터는 설명력이 모두 높은 것으로 나타남

(R-Squared: 0.94~0.99)



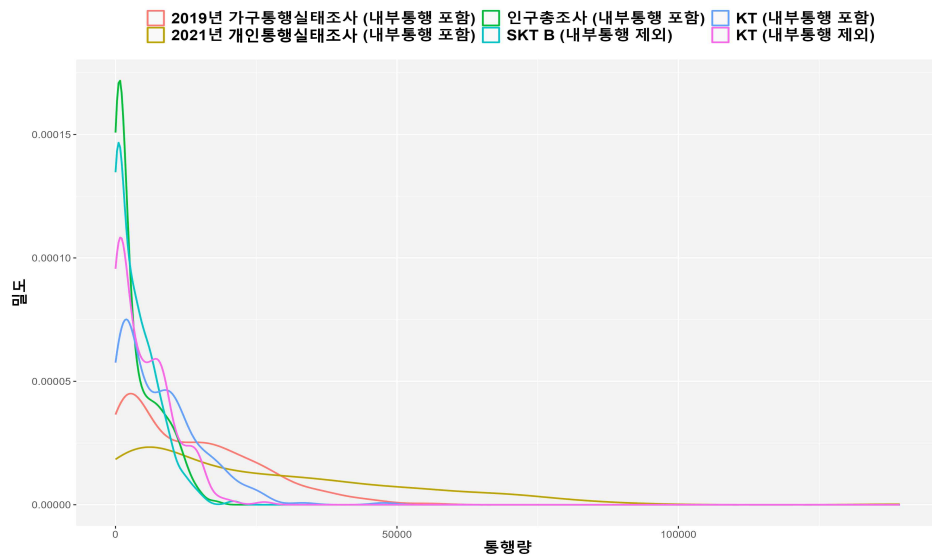
〈그림 3-24〉 (통행발생량) 통근 통행 - 취업자수와 각 데이터별 상관관계

③ 통학

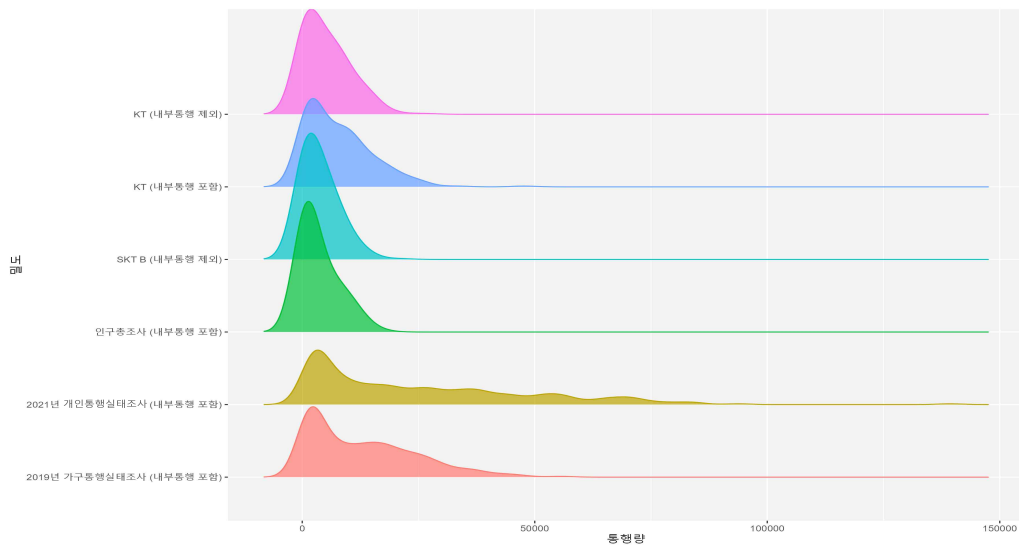
A) 데이터별 통행량 비교

○ 통학 통행은 대체로 2만 5천 명 미만의 낮은 통행량의 비율이 높게 나타나는 편임

- 조사자료보다는 모바일 통신데이터에서 그 특징이 더 두드러지는 편임
- 조사자료 중에서는 인구총조사만 침도가 높고, 2019년, 2021년 개인통행실태조사의 침도가 낮게 나타난 반면, 모바일 통신데이터는 모두 높은 침도를 보임



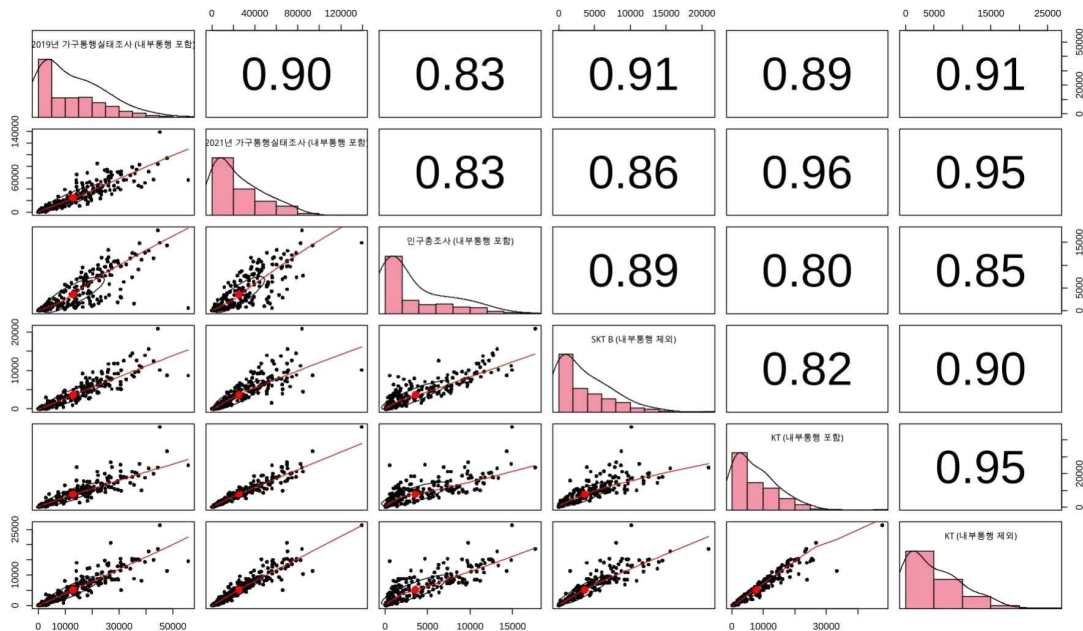
〈그림 3-25〉 (통행발생량) 통학 통행 - 데이터별 통행량 빈도 비교 (1)



〈그림 3-26〉 (통행발생량) 통학 통행 - 데이터별 통행량 빈도 비교 (2)

○ <그림 3-27>은 통학 목적의 통행발생량에 대하여 데이터간 상관관계를 분석한 결과임

- 인구총조사를 제외하고 SKT B(내부통행 제외)와 KT(내부통행 포함)도 타 데이터에 비해 분포 유사성이 낮은 편으로 분석됨 (R: 0.82)



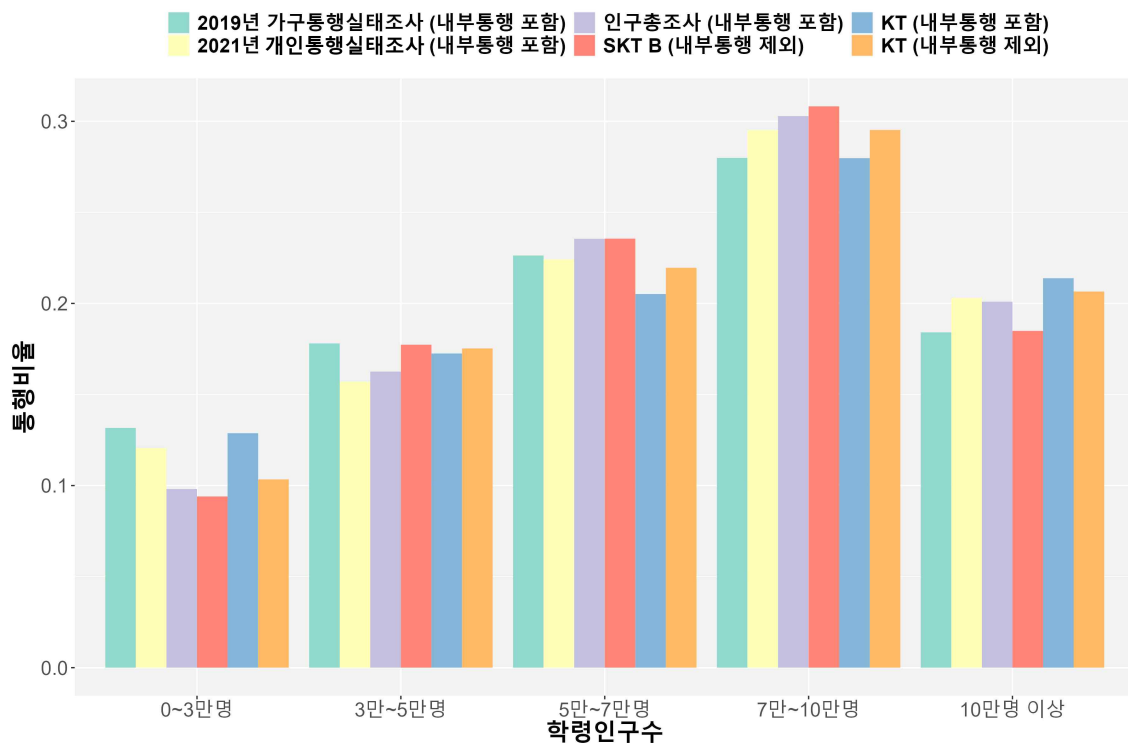
주: 그림에 표기된 수치는 R(상관계수)값을 의미함.

〈그림 3-27〉 (통행발생량) 통학 통행 - 데이터간 상관관계 비교

B) 통행발생요인별 통행량 비교

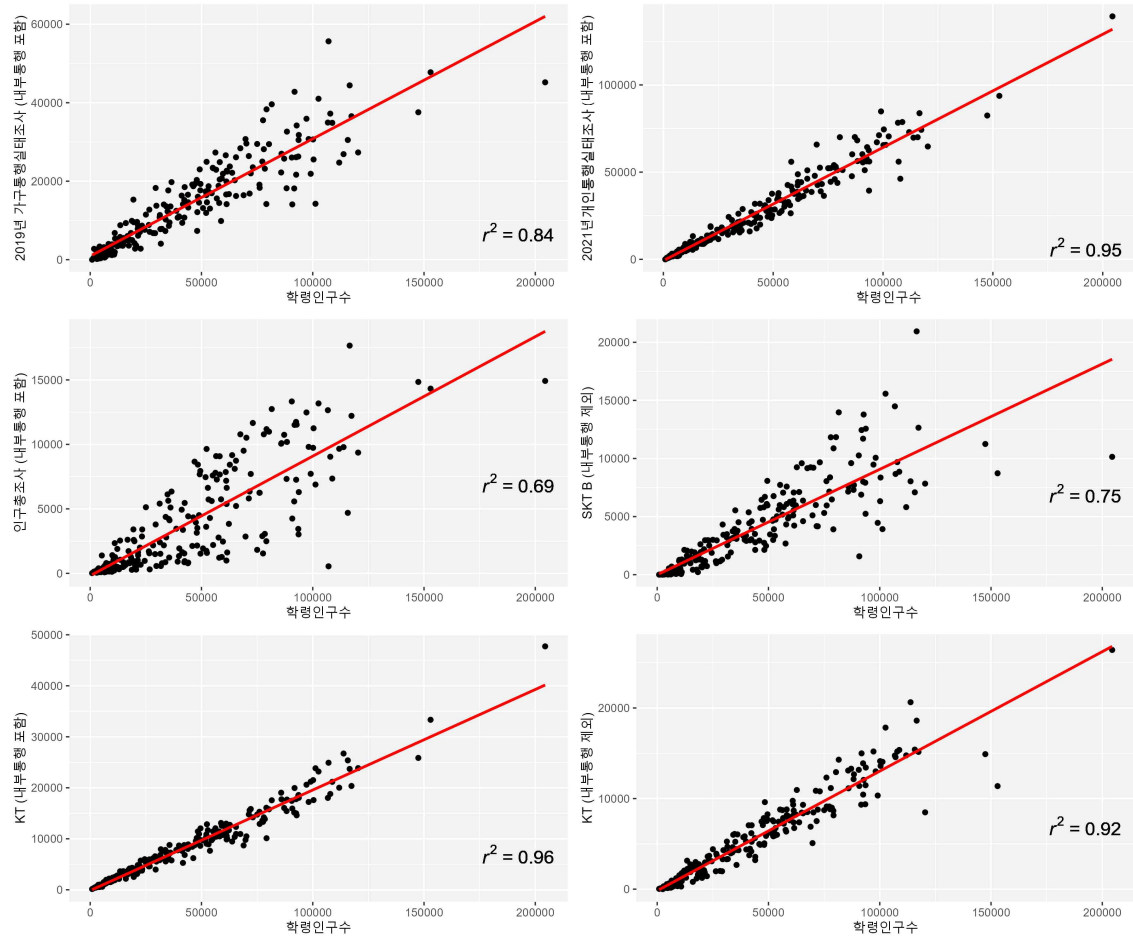
○ 학령인구수

- 학령인구수가 3만 명 미만인 시군구에서 데이터간 통행 비율 차이가 가장 크게 나타남
 - 해당 구간에서 조사자료와 모바일 데이터와의 차이가 가장 크게 나타남
- 2019년 가구통행실태조사(내부통행 포함)의 경우 타 데이터에 비해 학령인구수 규모가 작은 시군구의 통학 통행 비율이 높은 것으로 분석됨



〈그림 3-28〉 (통행발생량) 통학 통행 - 학령인구수에 따른 데이터별 통행량 비교

- 인구총조사와 SKT B(내부통행 포함)는 학령인구수에 대한 설명력이 부족한 것으로 나타남(인구총조사 - R-Squared: 0.69 / SKT B - R-Squared: 0.75)

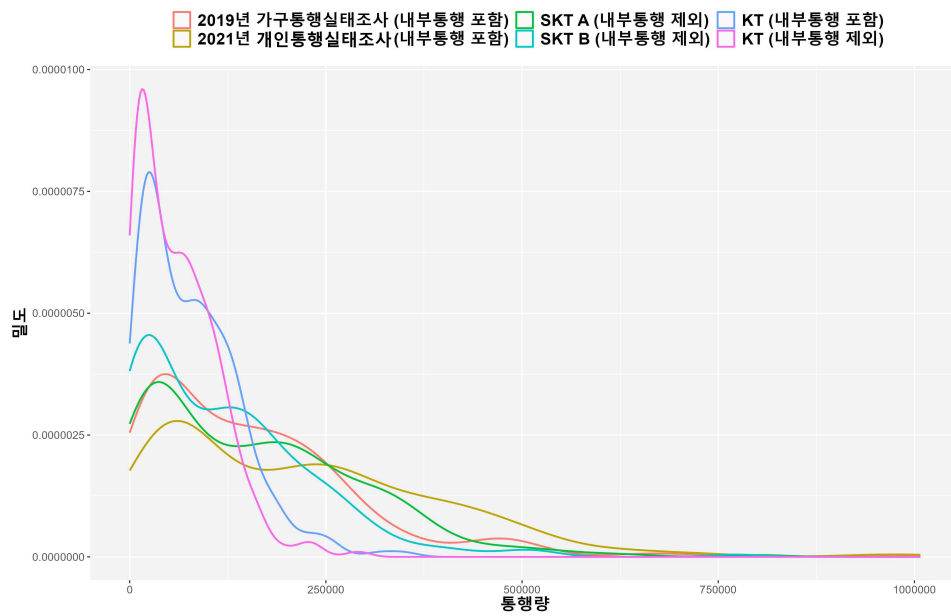


〈그림 3-29〉 (통행발생량) 통학 통행 - 학령인구수와 각 데이터별 상관관계

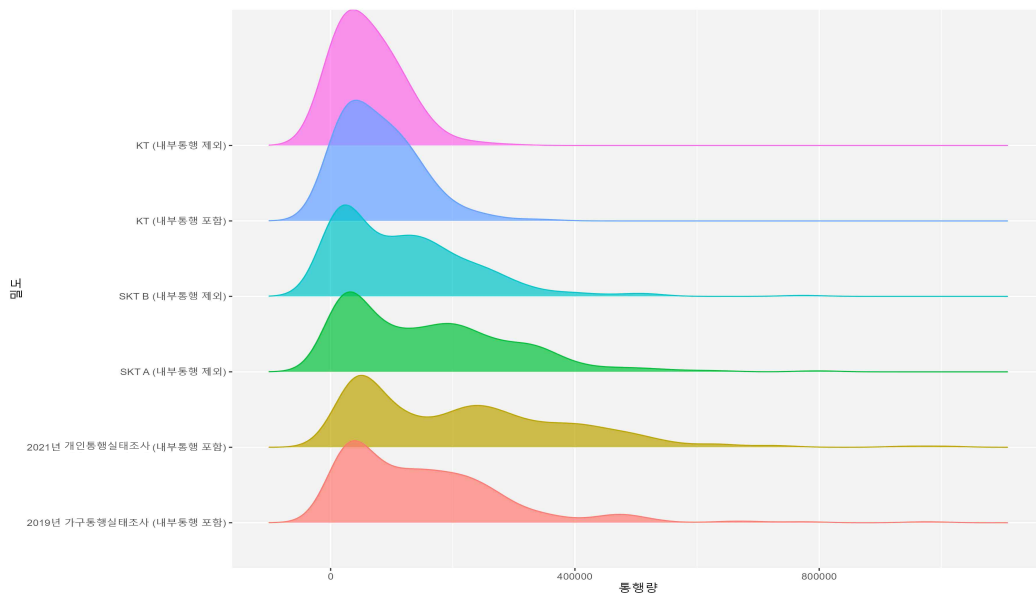
④ 귀가

A) 데이터별 통행량 비교

- <그림 3-30>, <그림 3-31>은 각 데이터별 귀가 통행에 대한 통행발생량별 빈도를 밀도 그래프로 표현한 것임



<그림 3-30> (통행발생량) 귀가 통행 - 데이터별 통행량 빈도 비교 (1)

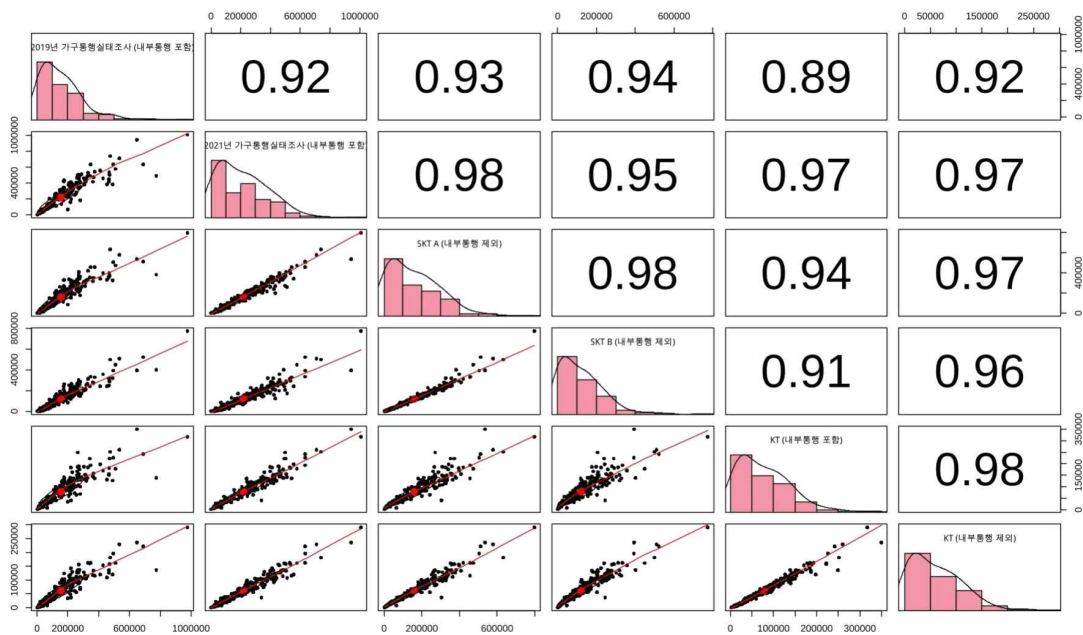


<그림 3-31> (통행발생량) 귀가 통행 - 데이터별 통행량 빈도 비교 (2)

- KT(내부통행 포함, 제외)는 통행량이 적은 데이터가 차지하는 비율이 높은 편이었으며, 그 외 데이터는 통행량이 적은 데이터가 차지하는 비율은 다소 낮은 편이었으나, 50만 통행 아래에 데이터가 분포하는 형태는 유사하게 나타남

○ <그림 3-32>는 데이터간 귀가 통행에 대한 상관관계를 비교한 결과임

- 전반적으로 귀가 통행에 대한 데이터별 상관관계는 유사한 것으로 나타남



주: 그림에 표기된 수치는 R(상관계수)값을 의미함.

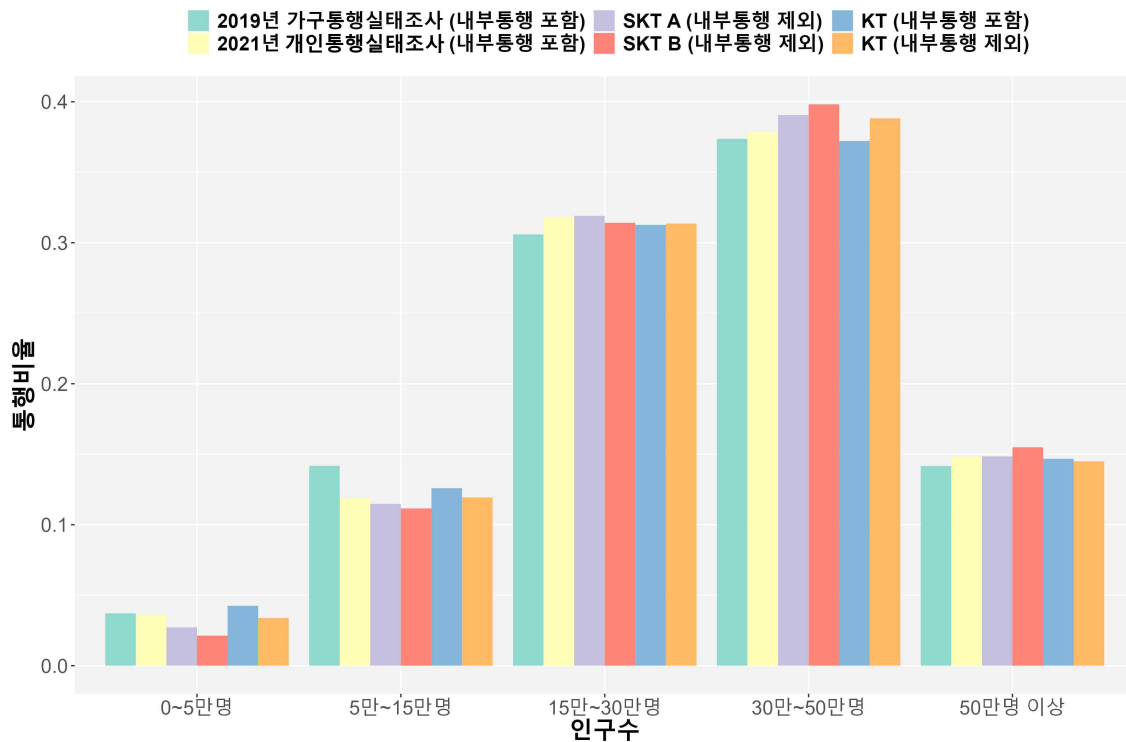
〈그림 3-32〉 (통행발생량) 귀가 통행 - 데이터간 상관관계 비교

B) 통행발생 요인별 통행량 비교

○ 인구수

- <그림 3-33>은 각 데이터의 귀가 통행발생량을 인구 규모에 따라 비교 분석한 결과임

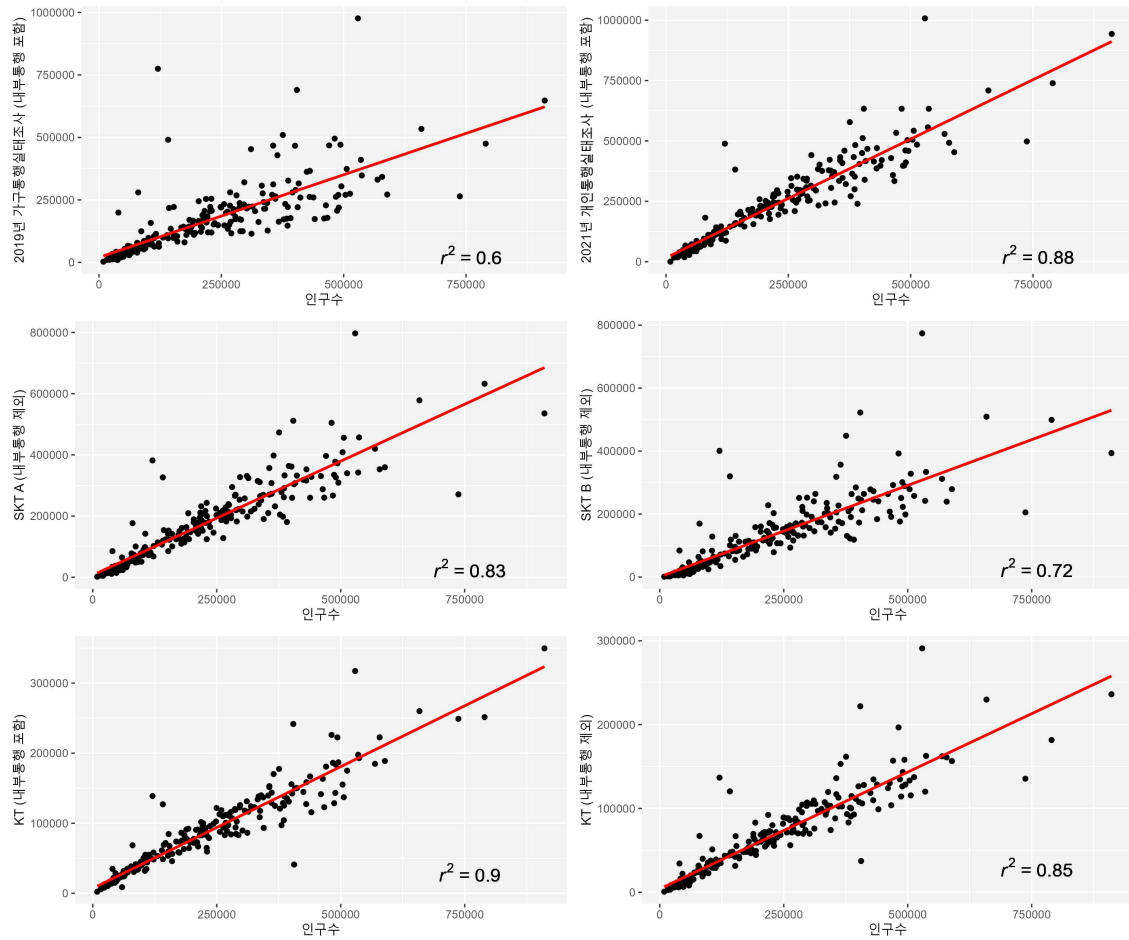
- 대체로 데이터간 차이가 크지 않으나, 인구수 5만 명 이상 15만 명 미만의 시군구에서 2019년 가구통행실태조사(내부통행 포함)와 나머지 데이터간 다소 차이가 나타남
- 30만 명 이상 50만 명 미만의 인구 규모를 가진 시군구에서 타 데이터에 비해 SKT B (내부통행 제외)의 통행 비율이 비교적 높게 나타남



<그림 3-33> (통행발생량) 귀가 통행 - 인구수에 따른 데이터별 통행량 비교

- <그림 3-34>는 각 데이터의 귀가 통행발생량과 인구수와의 관계를 비교한 결과임

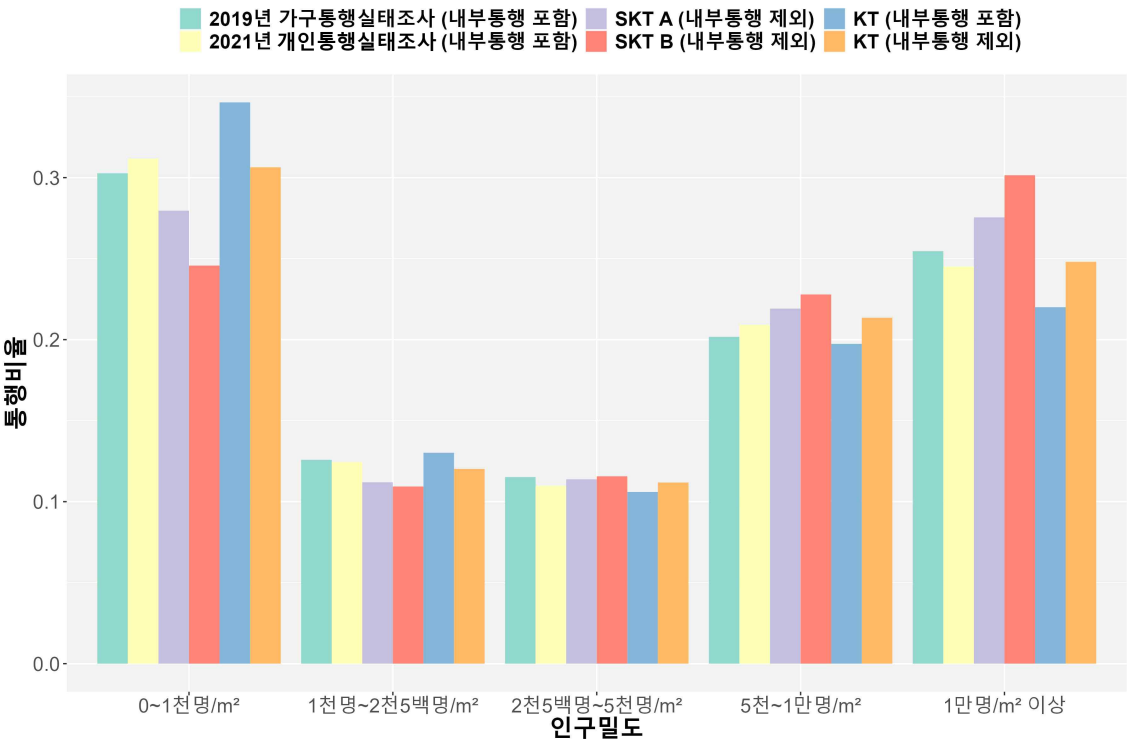
- 2019년 가구통행실태조사보다 모바일 데이터가 인구수 분포를 가장 잘 설명하는 것으로 나타났으며, 그 중에서도 내부통행을 포함한 KT 데이터의 설명력이 상대적으로 높은 것으로 분석됨(R-Squared: 0.90)



〈그림 3-34〉 (통행발생량) 귀가 통행 - 인구수와 각 데이터별 상관관계

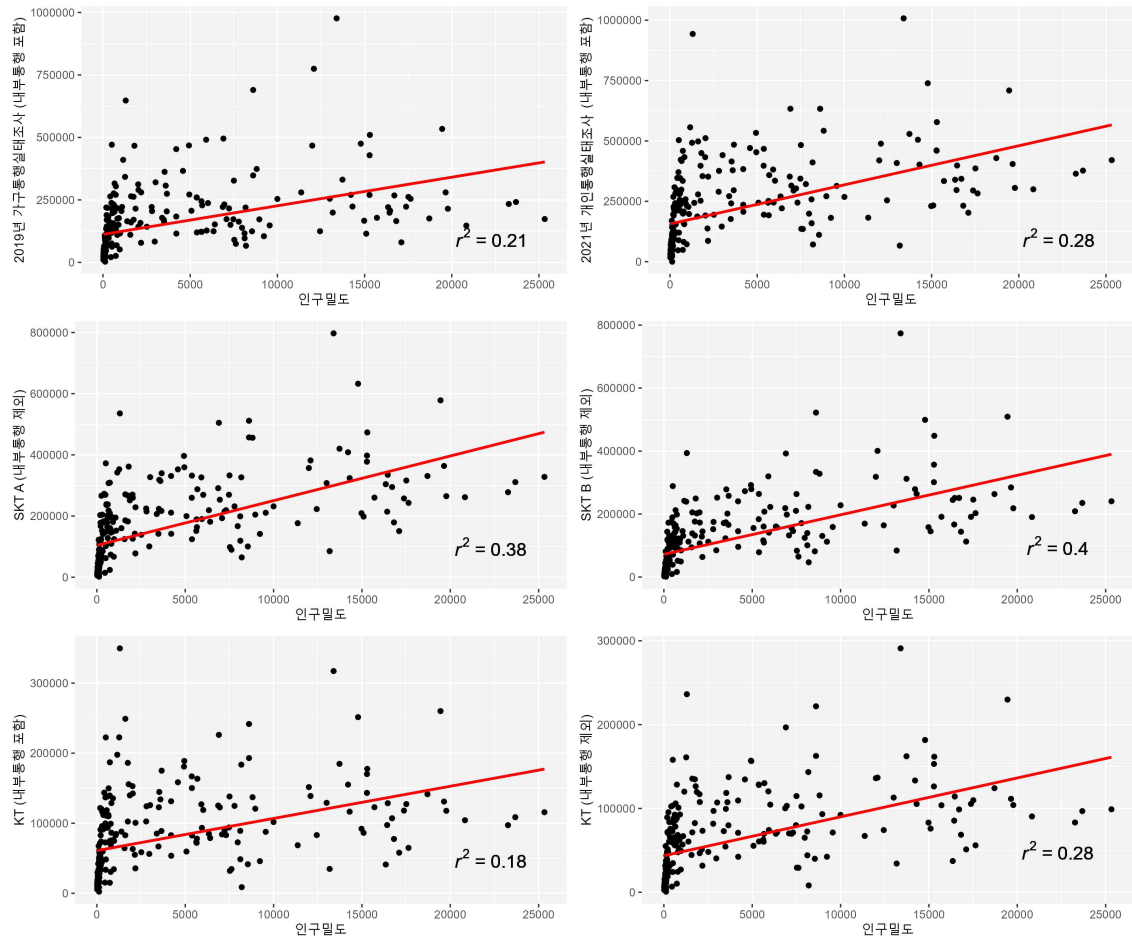
○ 인구밀도

- <그림 3-35>는 인구밀도를 기준으로 각 데이터의 귀가 통행발생량의 관계를 비교한 결과이며, 전반적으로 귀가 통행에 대한 인구밀도의 설명력이 낮은 것으로 분석됨



<그림 3-35> (통행발생량) 귀가 통행 - 인구밀도에 따른 데이터별 통행량 비교

- <그림 3-36>은 인구밀도를 기준으로 각 데이터의 귀가 통행발생량의 관계를 비교한 결과이며, 전반적으로 귀가 통행에 대한 인구밀도의 설명력이 낮은 것으로 분석됨

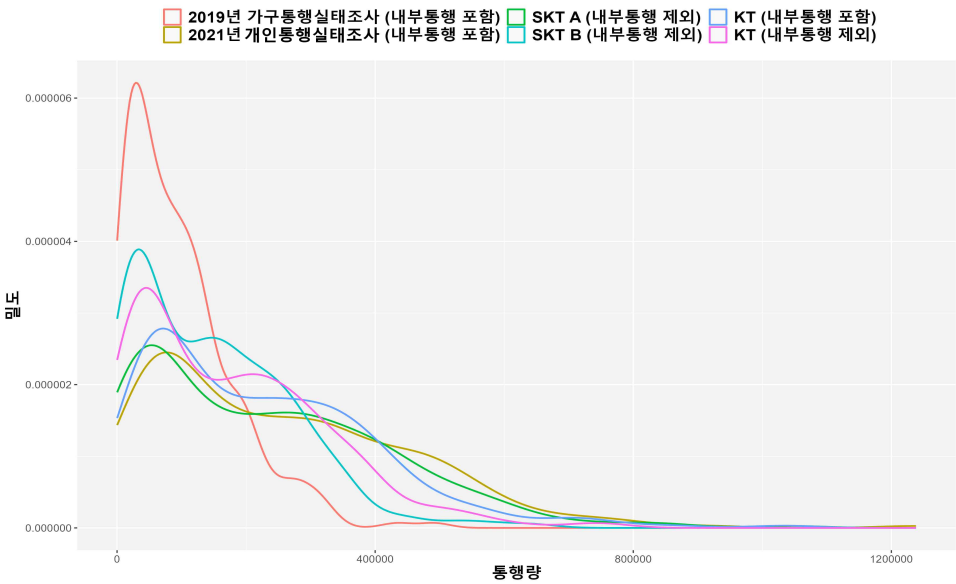


〈그림 3-36〉 (통행발생량) 귀가 통행 - 인구밀도와 각 데이터별 상관관계

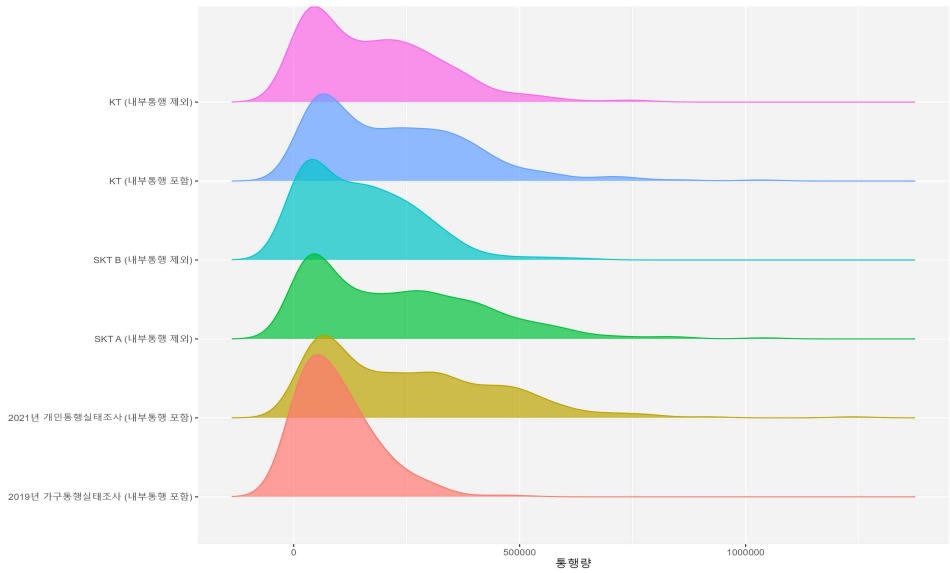
④ 기타

A) 데이터별 통행량 비교

- 데이터별 기타 통행에 대한 통행발생량 분포를 비교한 결과는 <그림 3-37>, <그림 3-38>과 같음



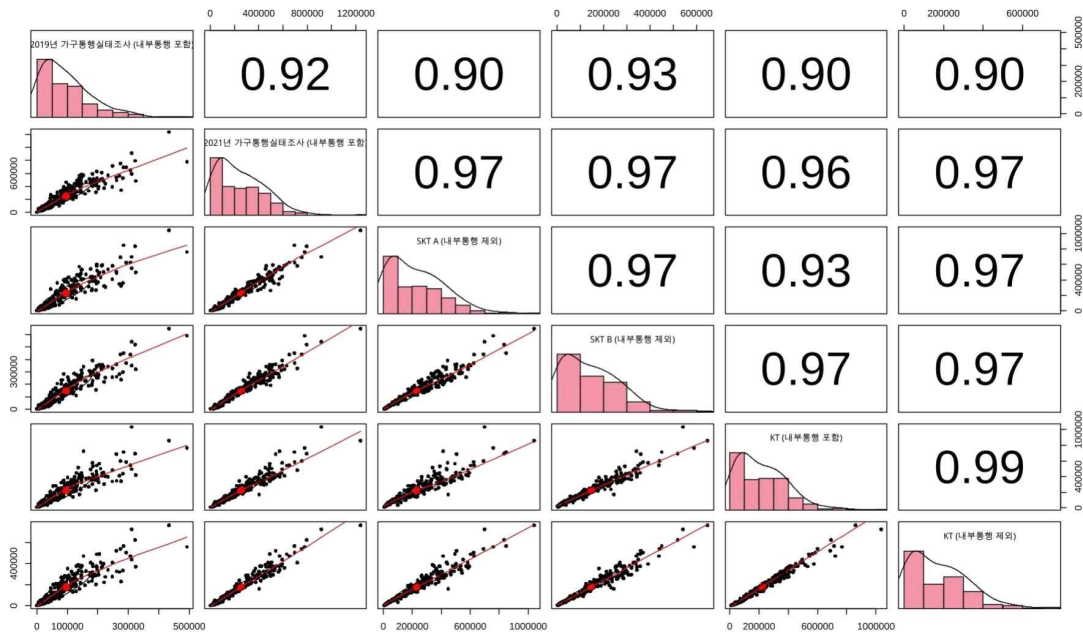
<그림 3-37> (통행발생량) 기타 통행 - 데이터별 통행량 빈도 비교 (1)



<그림 3-38> (통행발생량) 기타 통행 - 데이터별 통행량 빈도 비교 (2)

- 2019년 가구통행실태조사(내부통행 포함)는 다른 데이터에 비해 통행량이 적은 시군구가 차지하는 비율이 높은 편으로 나타남

○ 다음 <그림 3-39>와 같이 기타 통행에 대한 데이터별 상관관계를 비교한 결과, 전반적으로 데이터별 상관관계가 유사한 것으로 나타남



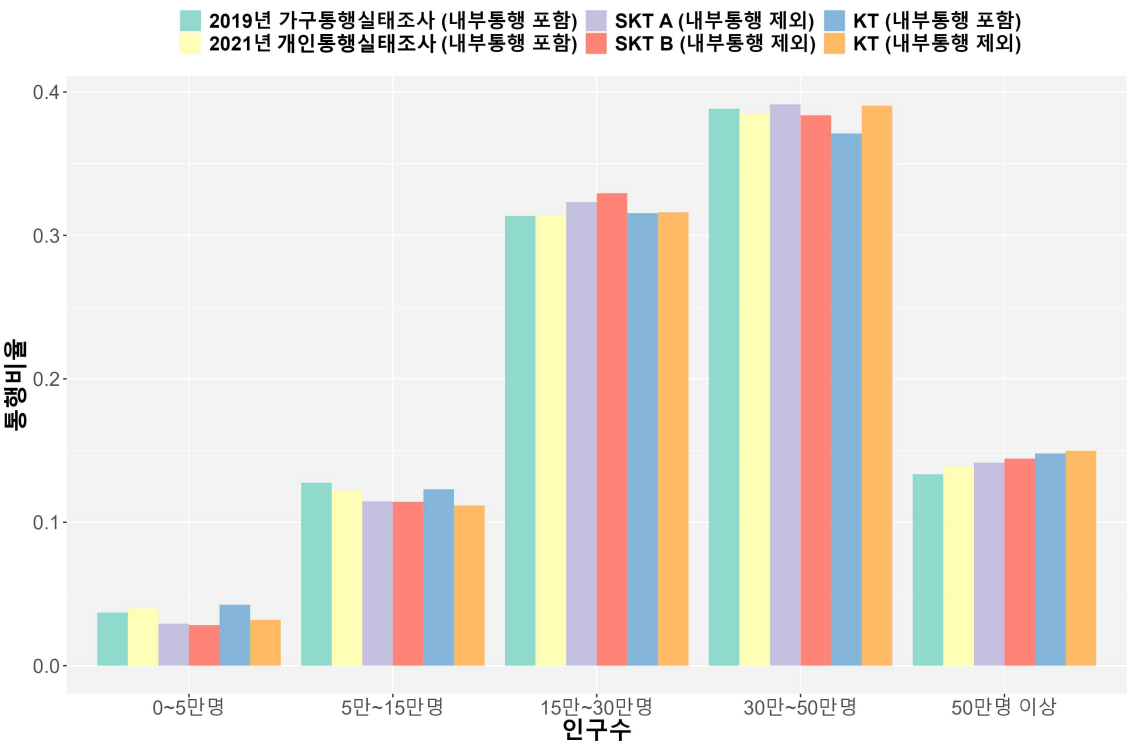
주: 그림에 표기된 수치는 R(상관계수)값을 의미함.

〈그림 3-39〉 (통행발생량) 기타 통행 - 데이터간 상관관계 비교

B) 통행발생 요인별 통행량 비교

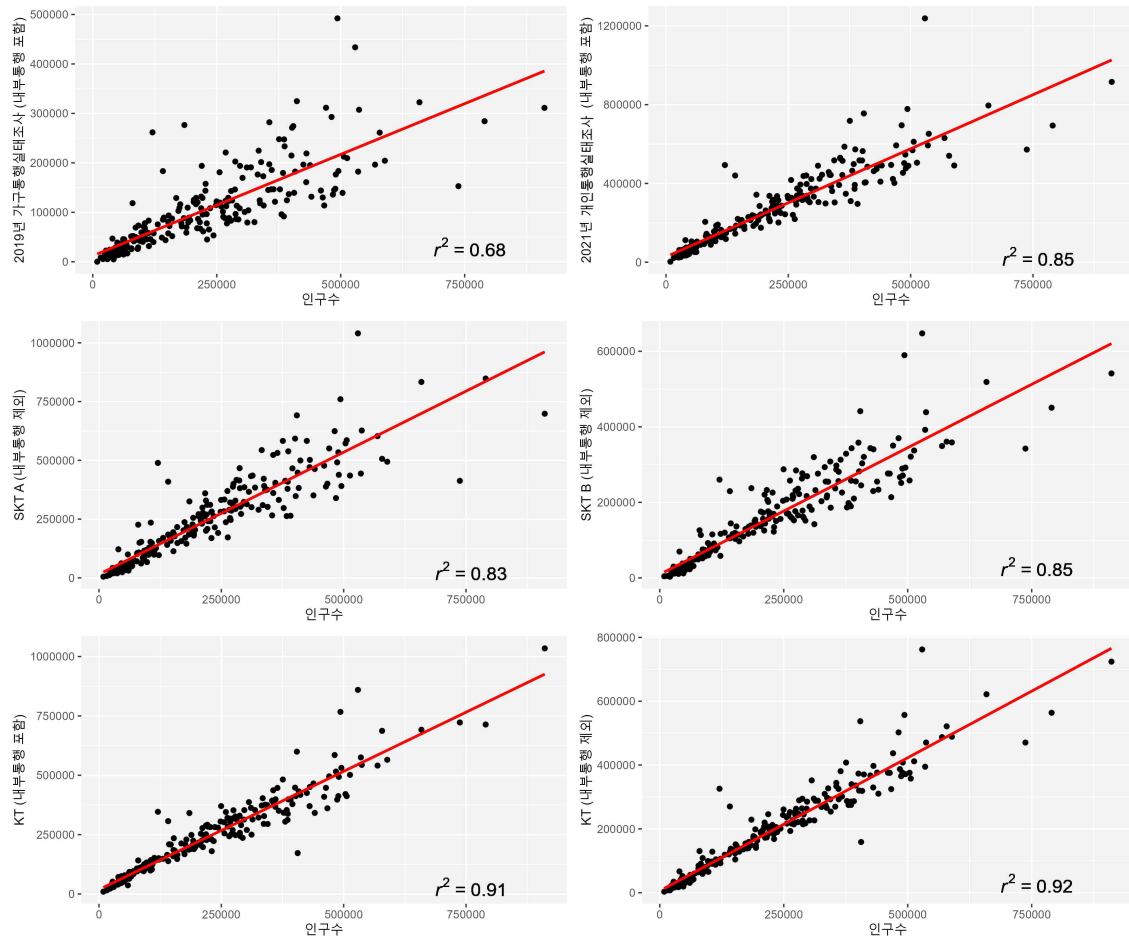
○ 인구수

- <그림 3-40>은 인구수에 따른 기타 통행에 대한 통행발생량의 관계를 데이터별로 분석한 결과로, 데이터 간 차이가 거의 없는 것을 확인할 수 있음



<그림 3-40> (통행발생량) 기타 통행 - 인구수에 따른 데이터별 통행량 비교

- <그림 3-41>은 각 데이터별 기타 통행에 대한 통행발생량과 인구수와의 관계를 비교한 결과임
- 2019년 가구통행실태조사(내부통행 포함)는 타 데이터에 비해 인구수에 대한 설명력이 낮은 편임 (R-Squared: 0.68)

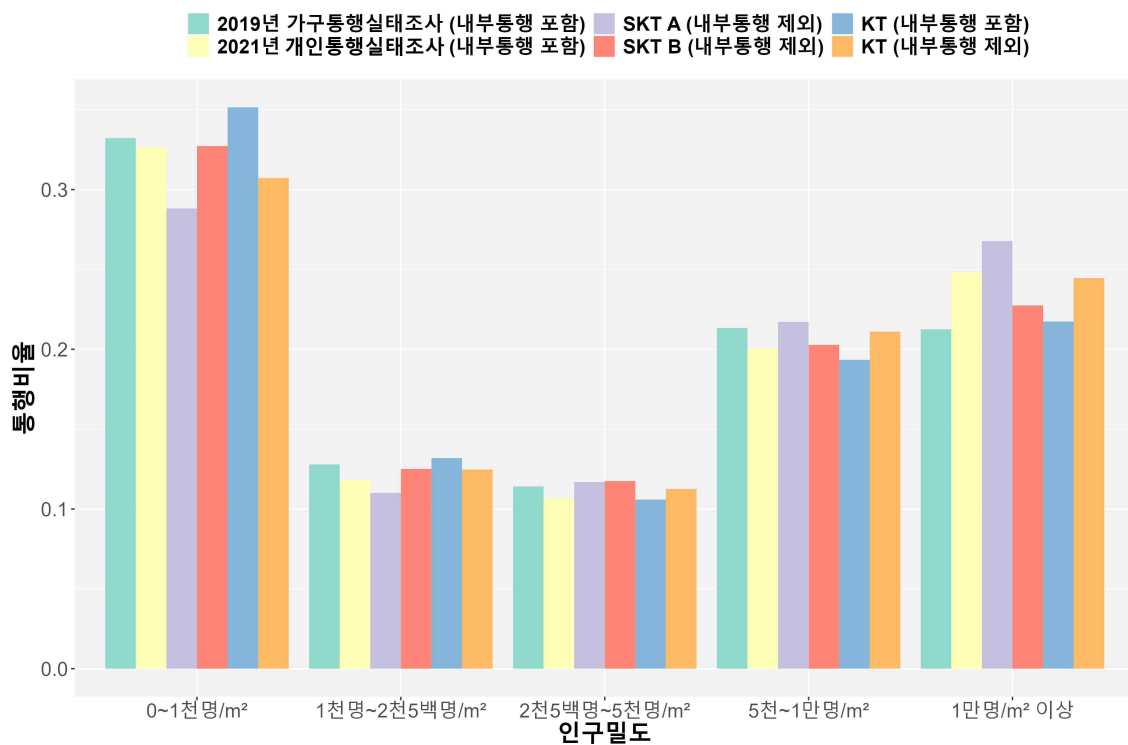


〈그림 3-41〉 (통행발생량) 기타 통행 - 인구수와 각 데이터별 상관관계

○ 인구밀도

- <그림 3-42>는 기타 통행에 대한 통행발생량의 분포를 인구밀도를 기준으로 카테고리화 하여 데이터별로 비교 분석한 결과임

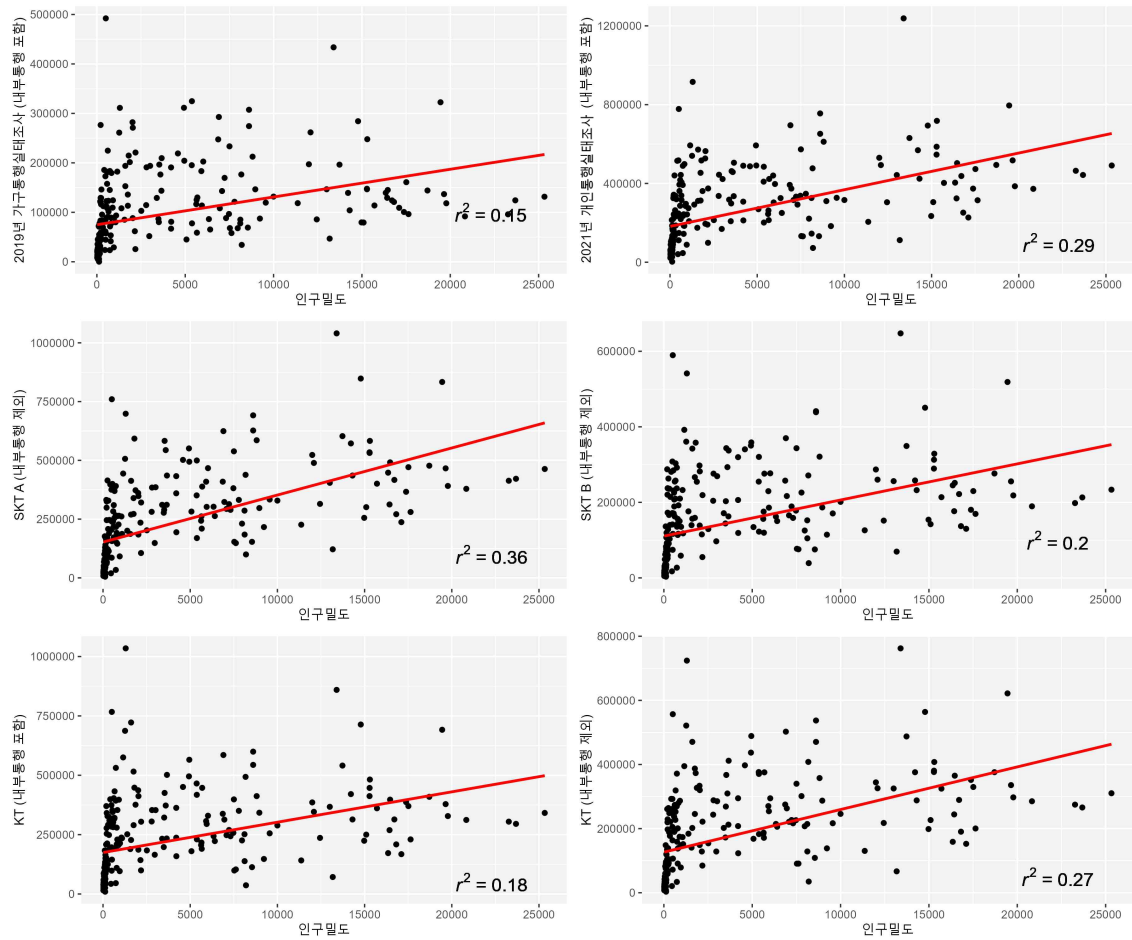
- 전반적으로 SKT A(내부통행 제외)가 타 데이터의 분석 결과와 다소 큰 차이를 보임
- KT(내부통행 포함)는 인구밀도 1천명/m² 미만의 구간에서 통행 비율이 다른 데이터에 비해 높게 나타나 2019년 가구통행실태조사(내부통행 포함), SKT B(내부통행 제외), KT(내부통행 제외)와 격차가 다소 크게 나타남



<그림 3-42> (통행발생량) 기타 통행 - 인구밀도에 따른 데이터별 통행량 비교

- <그림 3-43>은 인구밀도를 기준으로 각 데이터의 기타 통행에 대한 통행발생량과의 관계를 비교한 것임

- 모든 데이터가 인구밀도 분포에 대한 설명력이 떨어지는 편임



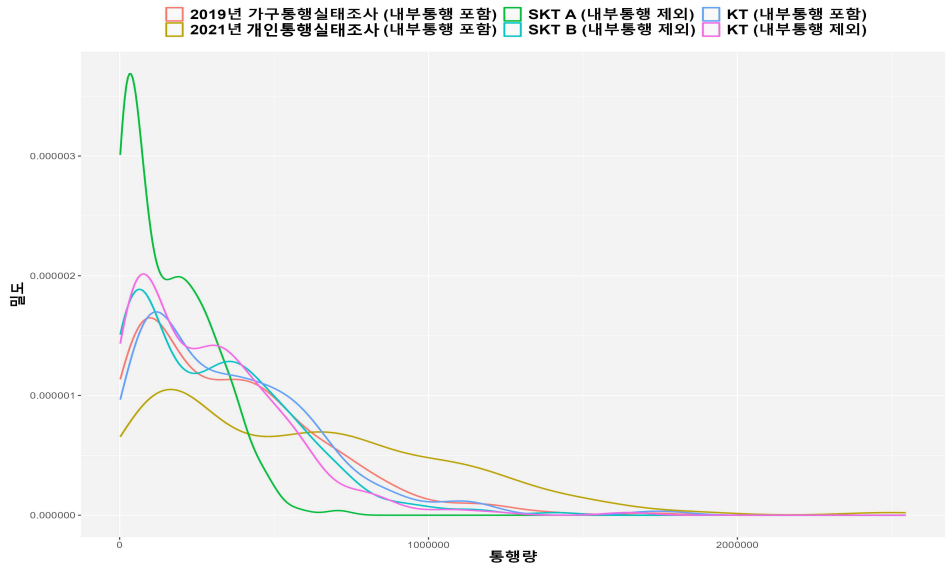
〈그림 3-43〉 (통행발생량) 기타 통행 - 인구밀도와 각 데이터별 상관관계

2) 통행도착량 기준

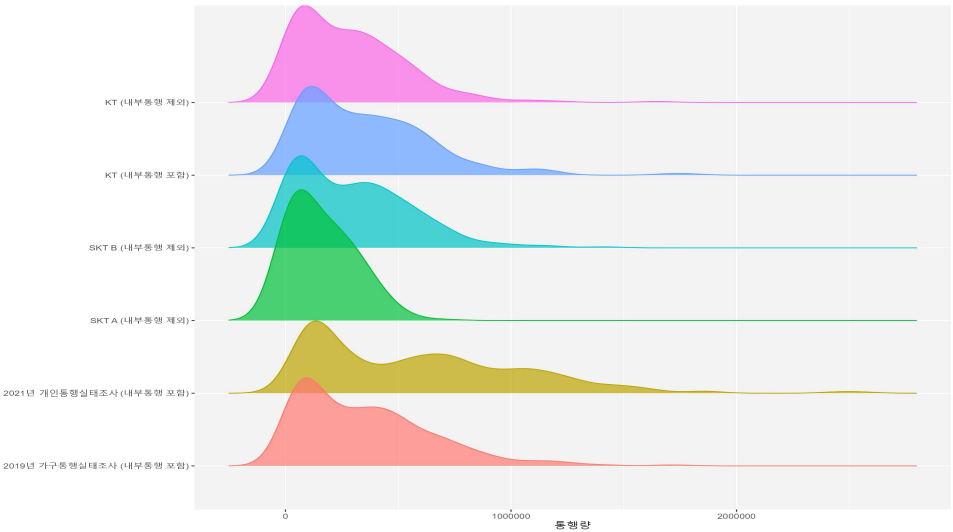
① 전체

A) 데이터별 통행량 비교

○ 시군구별 전체 통행도착량에 대한 분포를 데이터별로 비교한 결과는 다음 <그림 3-44>, <그림 3-45>와 같음



<그림 3-44> (통행도착량) 전체 통행 - 데이터별 통행량 빈도 비교 (1)



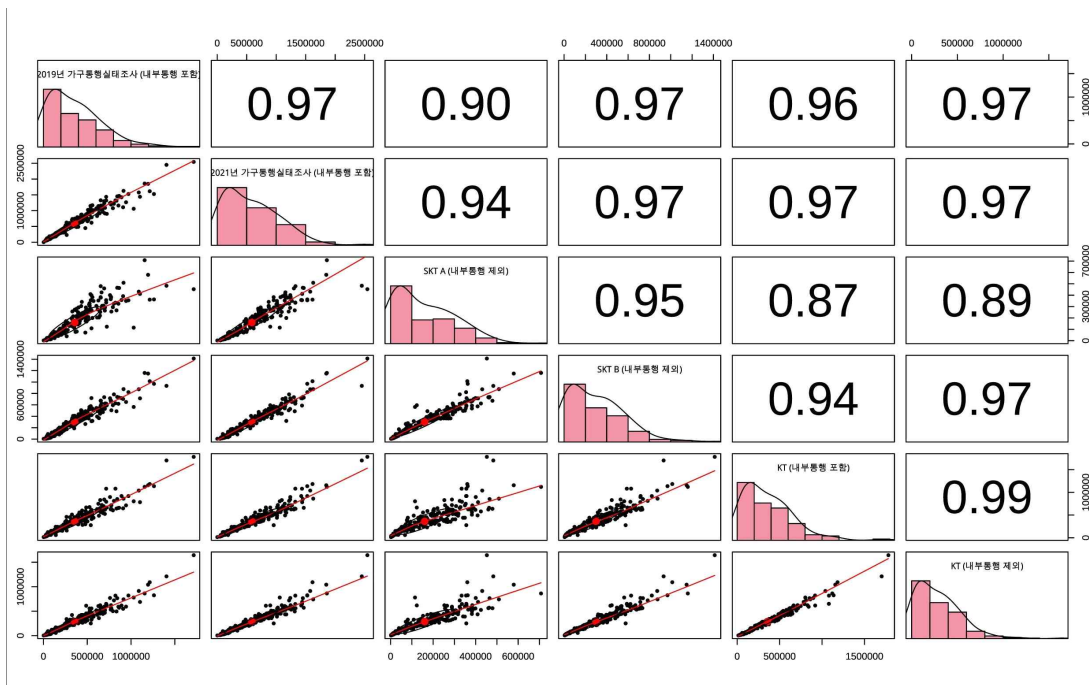
<그림 3-45> (통행도착량) 전체 통행 - 데이터별 통행량 빈도 비교 (2)

- 2021년 개인통행실태조사(내부통행 포함)와 SKT A(내부통행 제외)를 제외한 나머지 데이터의 분포가 유사한 것으로 나타남

- 2019년 가구통행실태조사(내부통행 포함), SKT B(내부통행 제외), KT(내부통행 포함, 제외)는 대체로 100만 통행량 미만에서 대부분의 통행이 이루어지고, 가장 높은 봉우리와 두 번째로 나타나는 봉우리의 높이 차가 크지 않은 데 비해, 2021년 개인통행실태조사(내부통행 포함)는 100만 통행 이상의 빈도도 비교적 높게 나타나고 여러 봉우리의 형태가 나타나며, SKT A(내부통행 제외)는 대부분의 통행이 50만 통행 미만에서 이루어지고 낮은 통행량의 비율이 높게 나타남

○ <그림 3-46>은 전체 통행에 대한 도착량을 기준으로 데이터간 상관관계를 분석한 것임

- 전반적으로 데이터간 상관관계가 높은 편이나, SKT A(내부통행 제외)와 KT (내부통행 포함, 제외)의 상관관계는 타 데이터간의 상관관계보다 상대적으로 낮은 것으로 분석됨



주: 그림에 표기된 수치는 R(상관계수)값을 의미함.

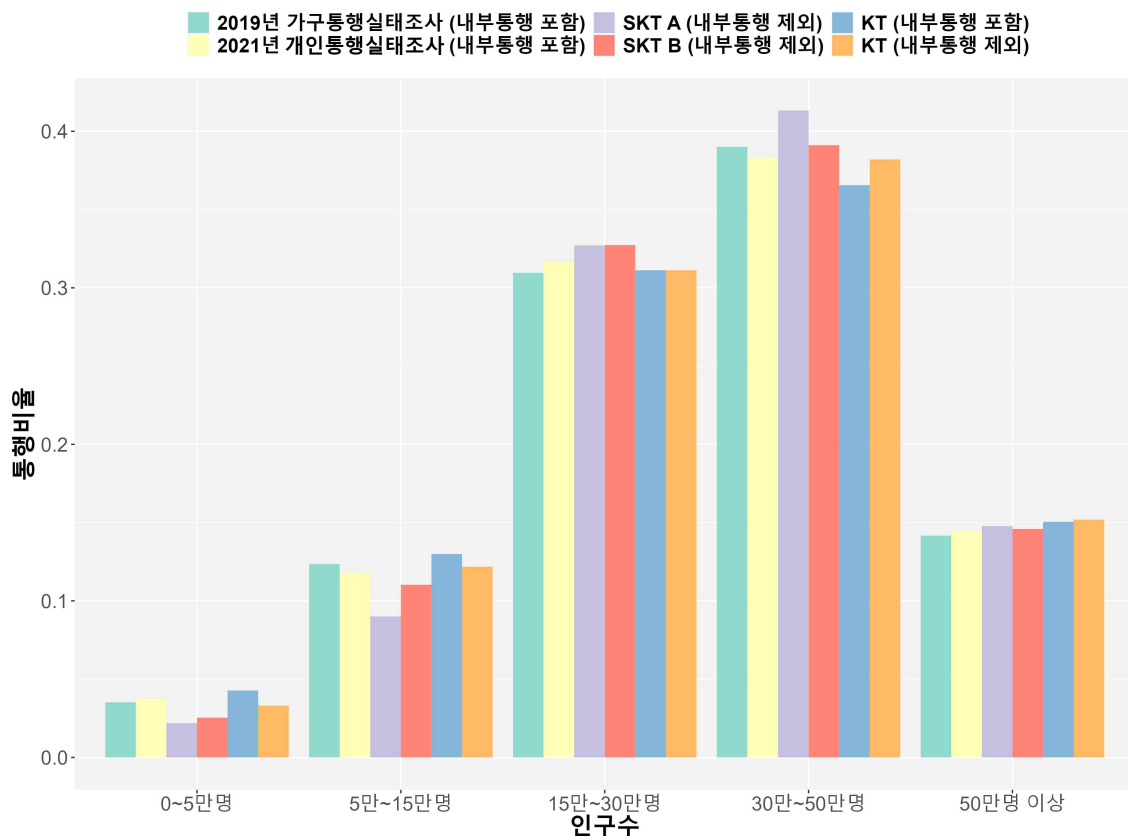
〈그림 3-46〉 (통행도착량) 전체 통행 - 데이터간 상관관계 비교

B) 통행도착 요인별 통행량 비교

○ 인구수

- <그림 3-47>은 인구 규모에 따라 전체 통행에 대한 도착량을 나누어서 데이터 별로 비교 분석한 그래프임

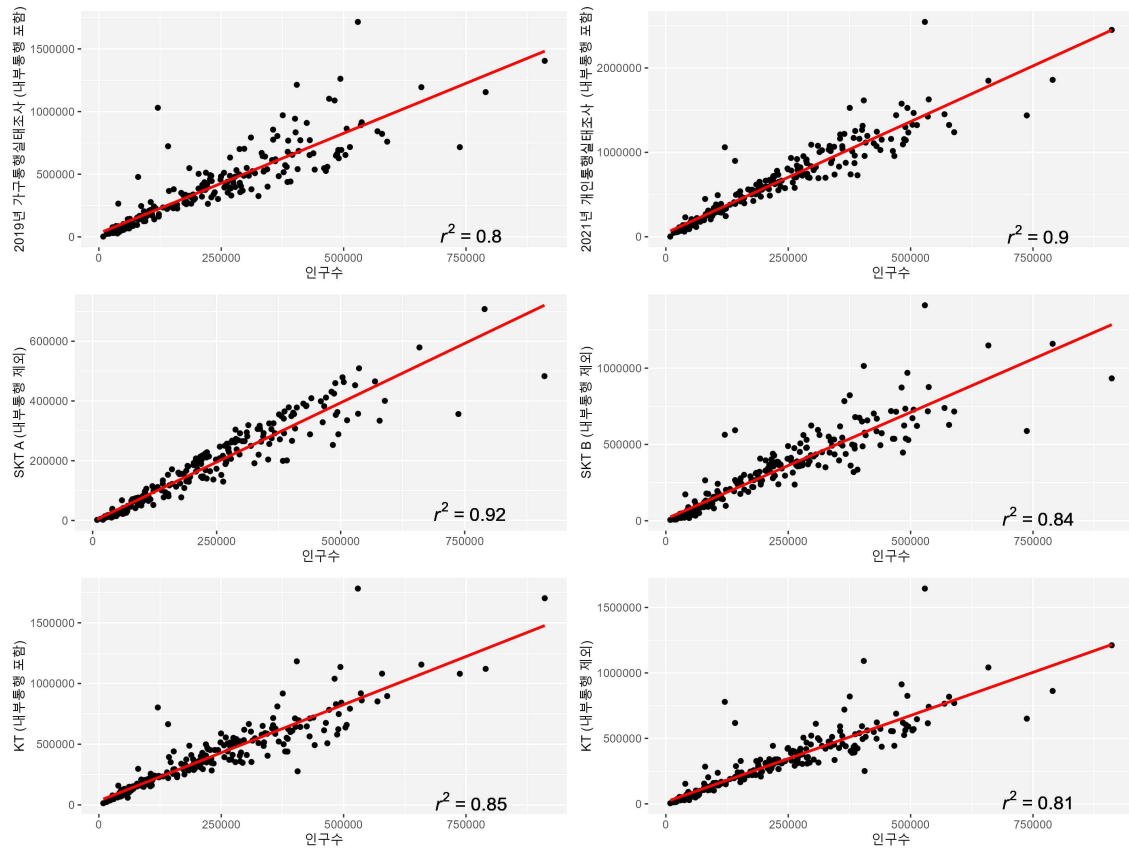
- 15만 명 이상 30만 명 미만, 50만 명 이상의 인구 규모에서는 데이터간 통행 비율이 거의 비슷한 편임
- 15만 명 미만의 인구 규모일 때에는 타 데이터에 비해 SKT A(내부통행 제외)의 통행 비율이 낮고, KT(내부통행 포함)의 통행 비율이 높았으나, 30만 명 이상 50만 명 미만의 인구 규모일 때에는 반대로 SKT A(내부통행 제외)의 통행 비율이 가장 높고, KT (내부통행 포함)의 통행 비율이 가장 낮게 나타남



<그림 3-47> (통행도착량) 전체 통행 - 인구수에 따른 데이터별 통행량 비교

- <그림 3-48>은 각 데이터의 시군구별 도착량과 인구수와의 관계를 비교한 것임

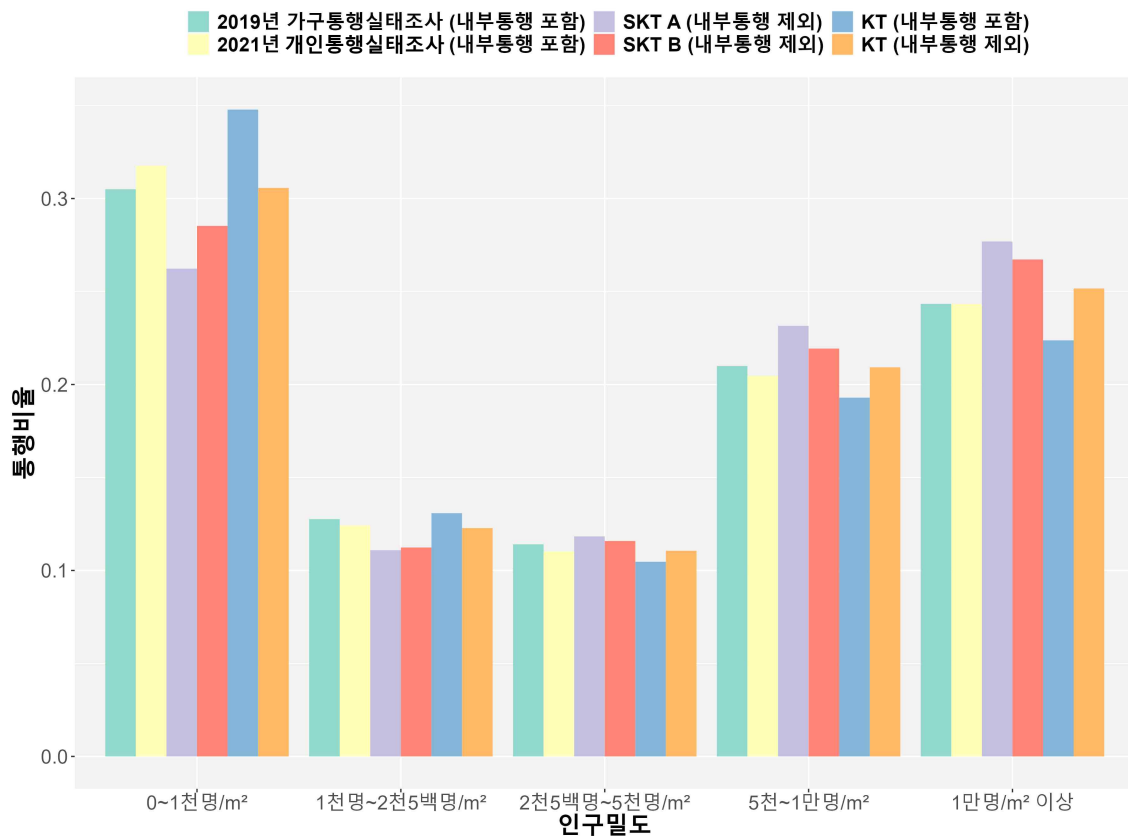
- 전반적으로 인구수에 대한 설명력이 높은 것으로 분석됨 (R-Squared: 0.80~0.92)



〈그림 3-48〉 (통행도착량) 전체 통행 - 인구수와 각 데이터별 상관관계

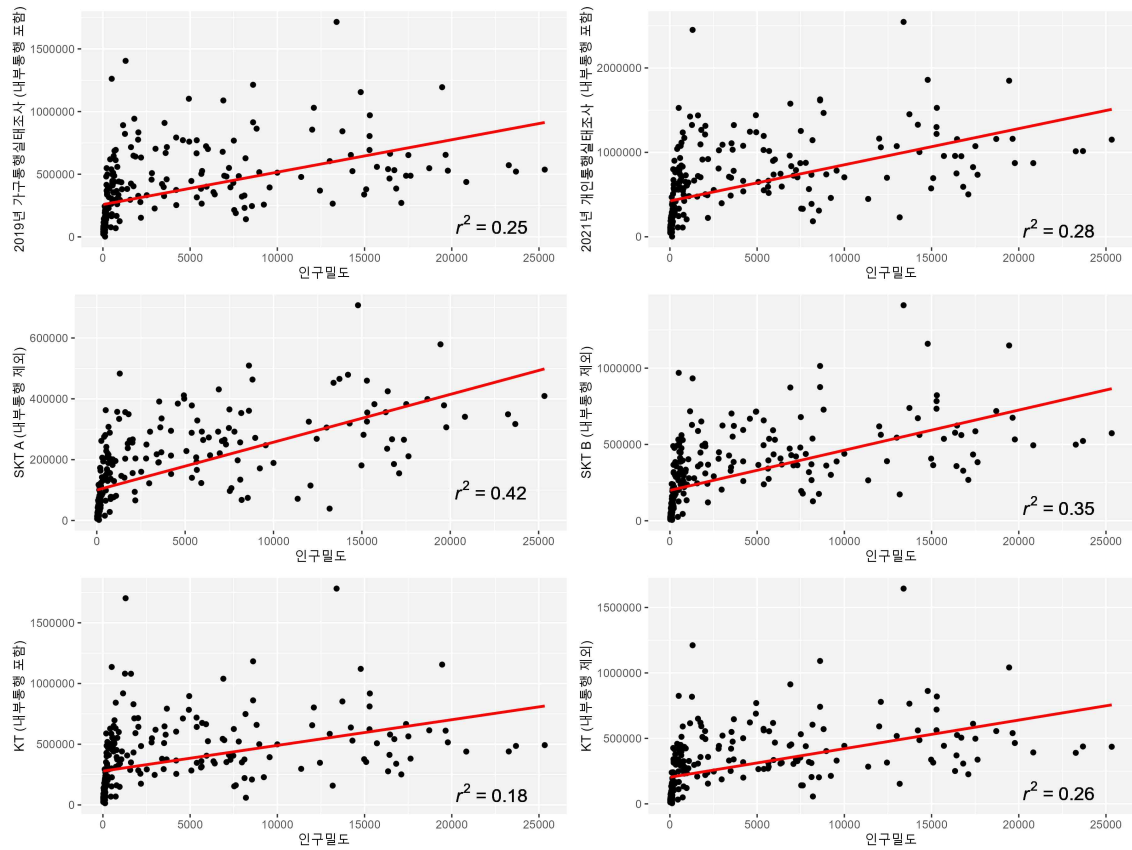
○ 인구밀도

- 전체 통행에 대한 도착량을 인구밀도에 따라 구분하여 데이터별 비교 분석한 결과는 다음 <그림 3-49>와 같음
- 인구밀도가 1천 명/㎡ 미만일 때 SKT A(내부통행 제외)의 통행 비율은 가장 낮게, KT(내부통행 포함)의 통행 비율은 가장 높게 나타나며, 데이터 간 차이가 가장 크게 나타남
- 인구밀도가 1만 명/㎡ 이상일 때에는 반대로 SKT A(내부통행 제외)의 통행 비율이 가장 높게, KT(내부통행 포함)의 통행 비율은 가장 낮게 나타남



<그림 3-49> (통행도착량) 전체 통행 - 인구밀도에 따른 데이터별 통행량 비교

- <그림 3-50>은 시군구별 인구밀도 분포와 각 데이터의 전체 통행에 대한 도착량의 관계를 비교 분석한 것임
- 전반적으로 인구밀도 분포를 잘 설명하지 못하는 것으로 확인됨



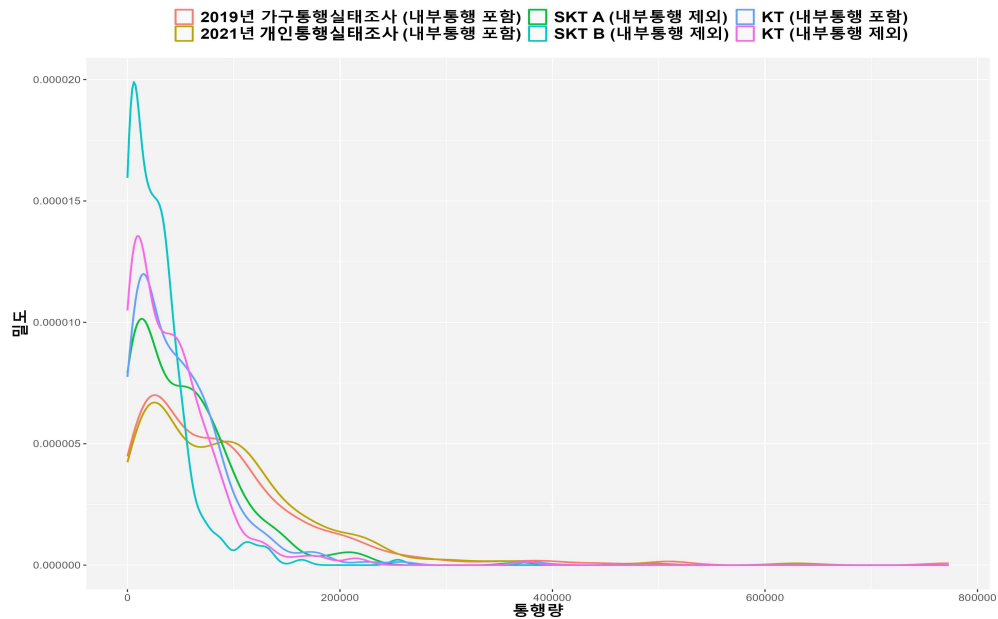
〈그림 3-50〉 (통행도착량) 전체 통행 - 인구밀도와 각 데이터별 상관관계

② 통근

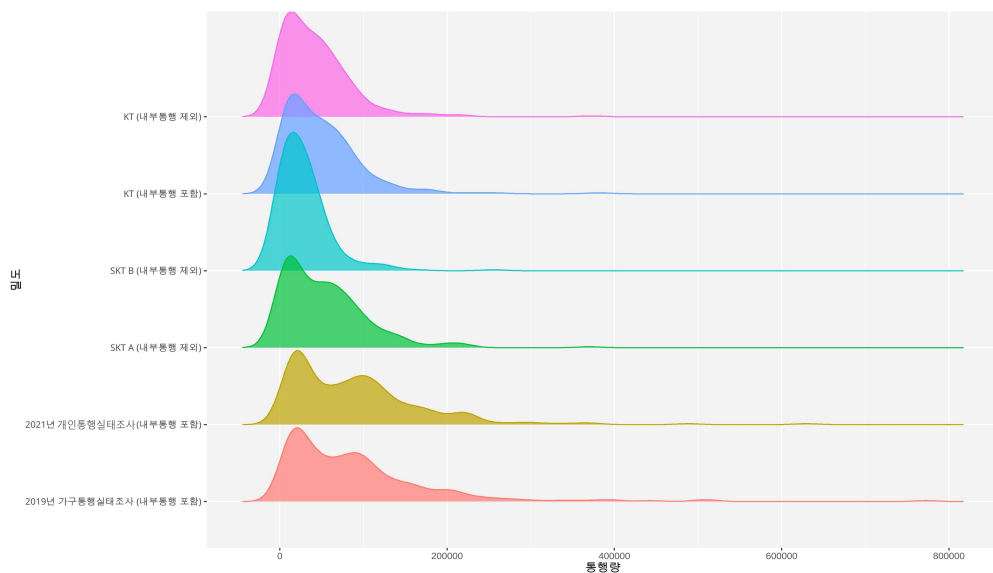
○ <그림 3-51>, <그림 3-52>는 데이터별 통근 목적 도착량 분포 그래프를 나타낸 것임

- SKT B(내부통행 제외)는 타 데이터에 비해 상대적으로 낮은 통행량의 비율이 높게 나타남

· 모바일 데이터는 조사자료에 비해 낮은 통행량의 비율이 높게 나타나는 편임



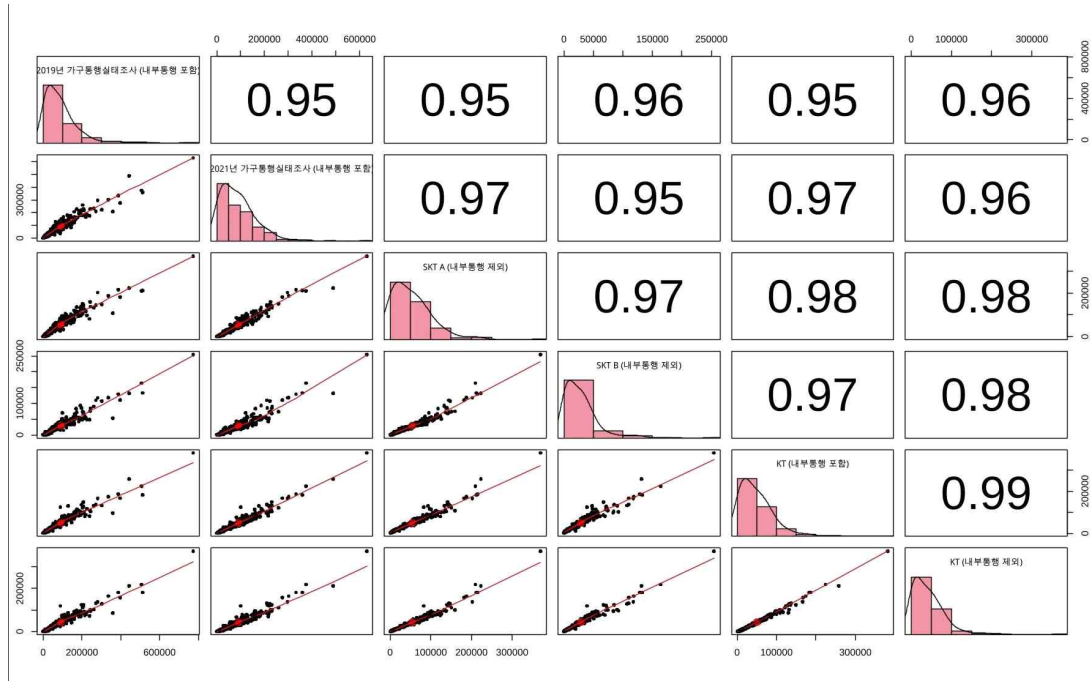
<그림 3-51> (통행도착량) 통근 통행 - 데이터별 통행량 빈도 비교 (1)



<그림 3-52> (통행도착량) 통근 통행 - 데이터별 통행량 빈도 비교 (2)

A) 데이터별 통행량 비교

- 통근 목적 통행의 도착량에 대하여 데이터 간 상관관계를 분석한 결과는 <그림 3-53>과 같음
 - 데이터 간 상관계수가 전부 0.95 이상으로 높은 상관관계를 갖는 것으로 분석됨



주: 그림에 표기된 수치는 R(상관계수)값을 의미함.

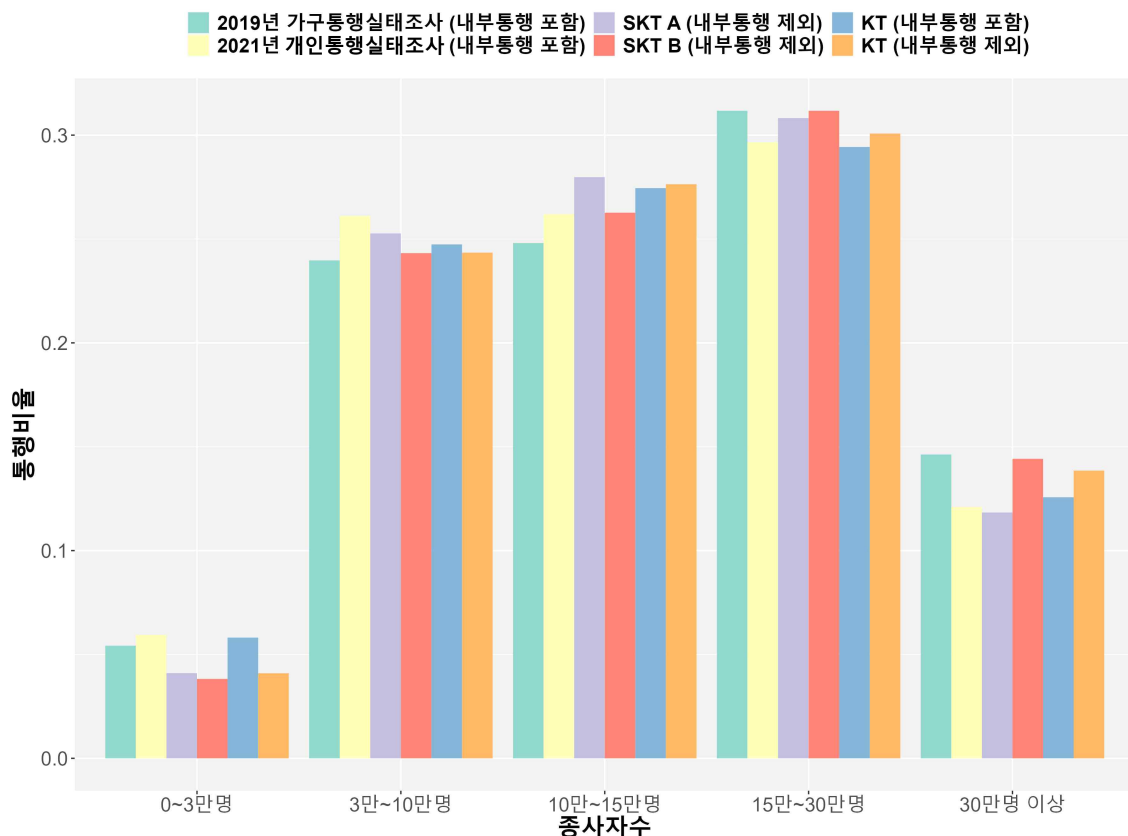
〈그림 3-53〉 (통행도착량) 통근 통행 - 데이터간 상관관계 비교

B) 통행도착 요인별 통행량 비교

○ 종사자수

- 종사자수 규모에 따라 통근 통행의 도착량을 데이터별로 비교한 결과는 다음 <그림 3-54>와 같음

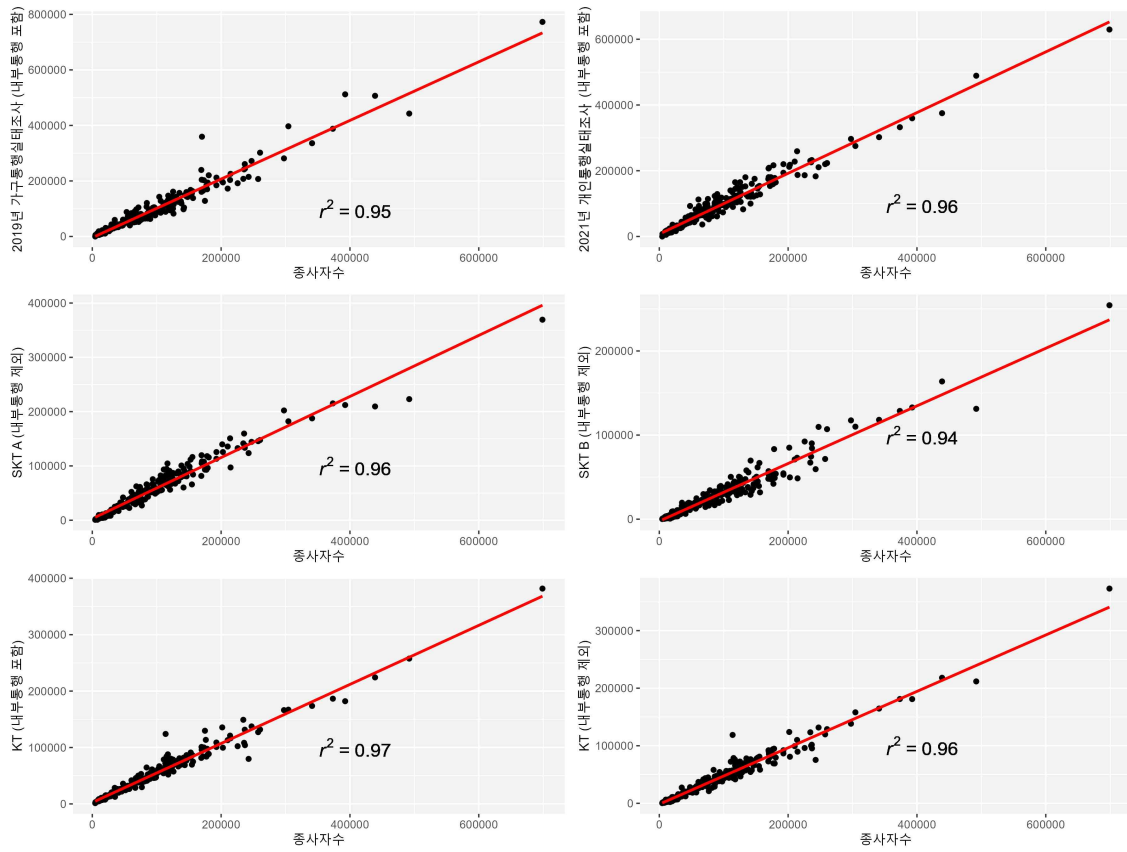
- 종사자수가 10만 명 이상 15만 명 미만일 때 2019년, 2021년 통행실태조사(내부통행 포함)의 통행 비율이 타 데이터에 비해 줄어들면서 데이터 간 격차가 다소 벌어지는 것을 확인함
- SKT A(내부통행 제외)는 종사자 수가 매우 적거나, 매우 큰 시군구에서는 (3만 명 미만이거나 30만 명 이상인 시군구) 타 데이터에 비해 통행 비율이 낮은 편임



<그림 3-54> (통행도착량) 통근 통행 - 종사자수에 따른 데이터별 통행량 비교

- 각 데이터의 통근 통행 도착량과 종사자수의 관계를 비교한 것은 다음 <그림 3-55>와 같음

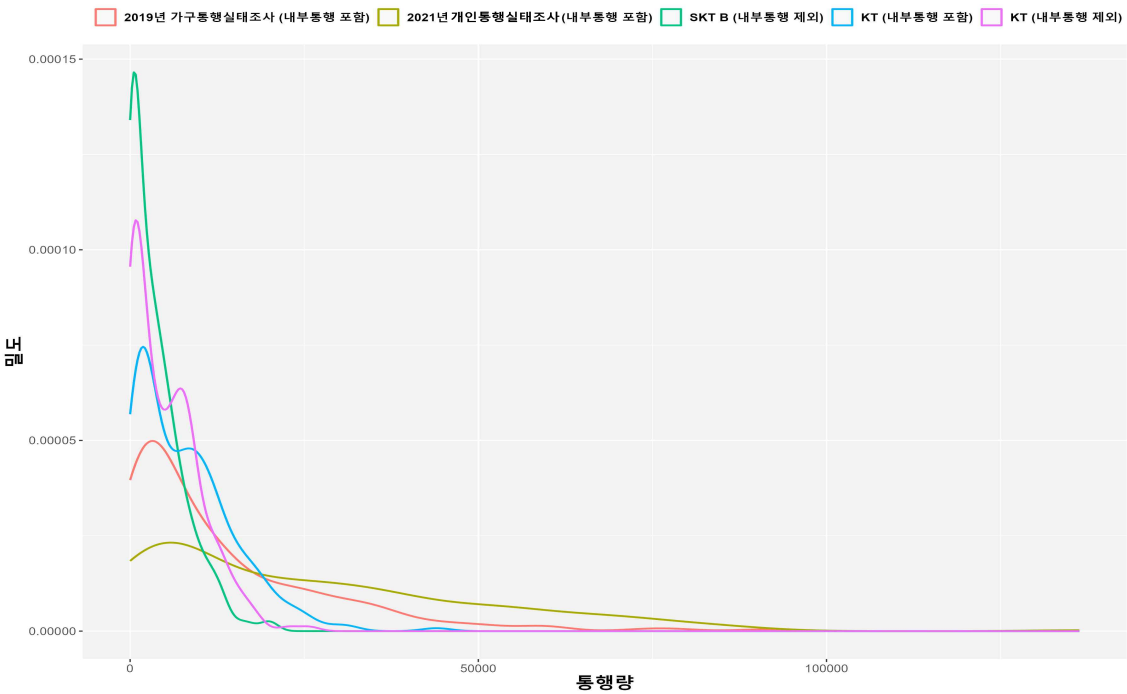
- 종사자수에 대한 설명력은 전부 0.94 이상으로 매우 높은 편으로 나타남



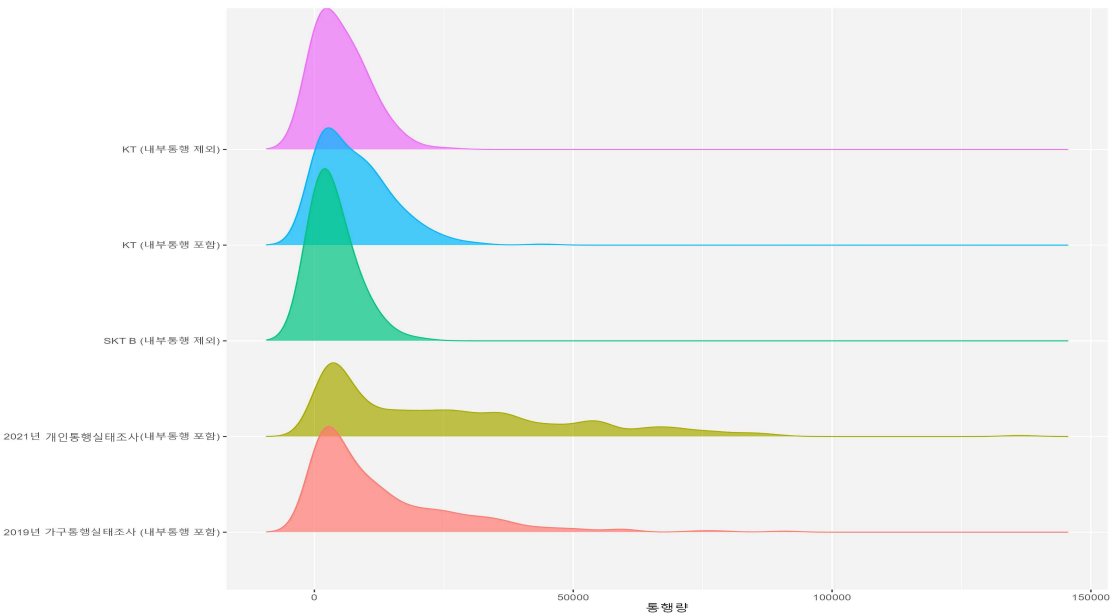
〈그림 3-55〉 (통행도착량) 통근 통행 - 종사자수와 각 데이터별 상관관계

③ 통학

- 통학 통행은 다른 목적 통행에 비해 데이터 간 도착량 분포 차이가 크게 나타남 (<그림 3-56>, <그림 3-57> 참조)



<그림 3-56> (통행도착량) 통학 통행 - 데이터별 통행량 빈도 비교 (1)



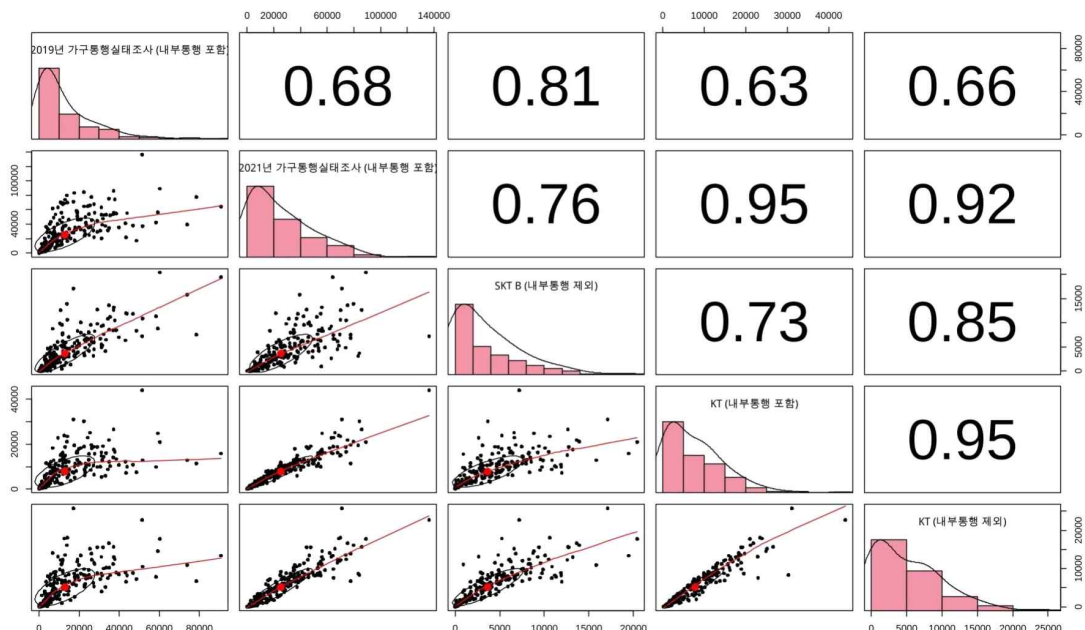
<그림 3-57> (통행도착량) 통학 통행 - 데이터별 통행량 빈도 비교 (2)

- 모바일 통신데이터는 상대적으로 조사자료에 비해 낮은 통행량의 빈도가 높은 편임
 - 모바일 통신데이터는 2만 5천 통행 미만 통행량 차지하는 비율이 95% 이상이며, 특히 SKT B(내부통행 제외)는 1만 통행 미만의 낮은 통행량의 비율이 매우 높은 편임
 - 조사자료는 2만 5천 통행 미만 통행량의 비율이 모바일 통신데이터에 비해 낮고 2만 5천 이상의 통행량도 상당 비율 차지하고 있음

A) 데이터별 통행량 비교

○ <그림 3-58>은 통학 목적 도착량에 대하여 데이터 간 상관관계를 분석한 결과임

- 2019년 가구통행실태조사(내부통행 포함)는 다른 데이터와의 분포 유사성이 다소 떨어지는 편임
 - 같은 조사자료인 2021년 개인통행실태조사(내부통행 포함)과도 차이를 보임(R: 0.68)
- 2021년 가구통행실태조사(내부통행 포함), KT(내부통행 포함), KT(내부통행 제외) 등의 상관관계가 높은 것으로 나타남



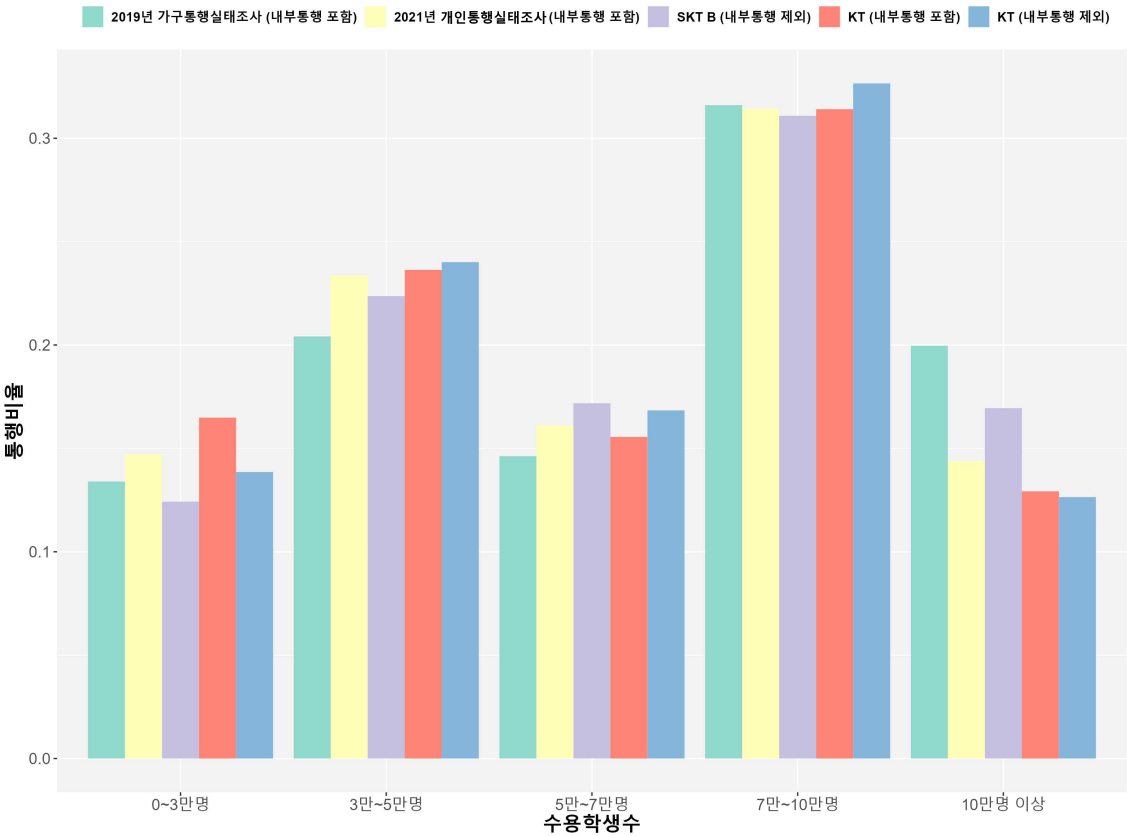
주: 그림에 표기된 수치는 R(상관계수)값을 의미함.

〈그림 3-58〉 (통행도착량) 통학 통행 - 데이터간 상관관계 비교

B) 통행도착 요인별 통행량 비교

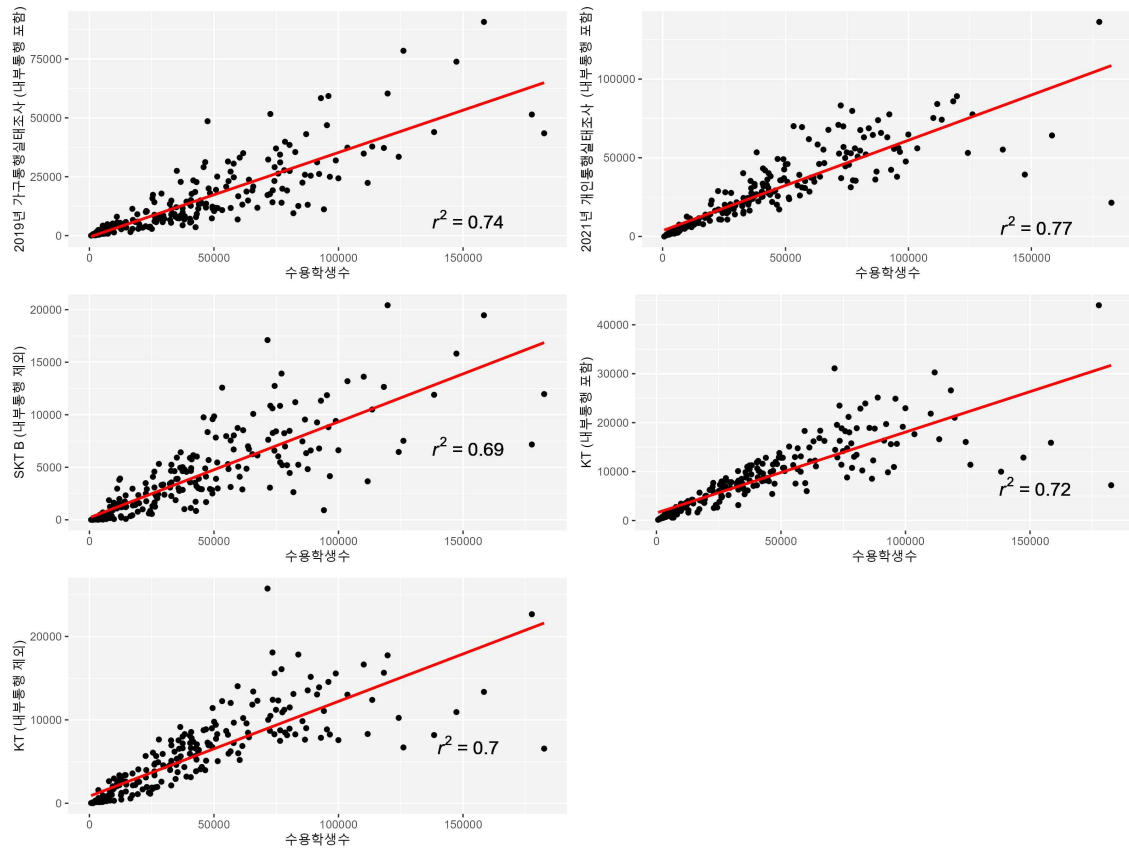
○ 수용학생수

- 수용학생수 규모에 따라 통학 통행의 도착량을 데이터별로 비교한 결과는 다음 <그림 3-59>와 같음
- 수용학생수가 10만 명 이상일 때 2019년 가구통행실태조사(내부통행 포함)와 SKT B (내부통행 제외)의 통행 비율이 타 데이터에 비해 높게 나타남



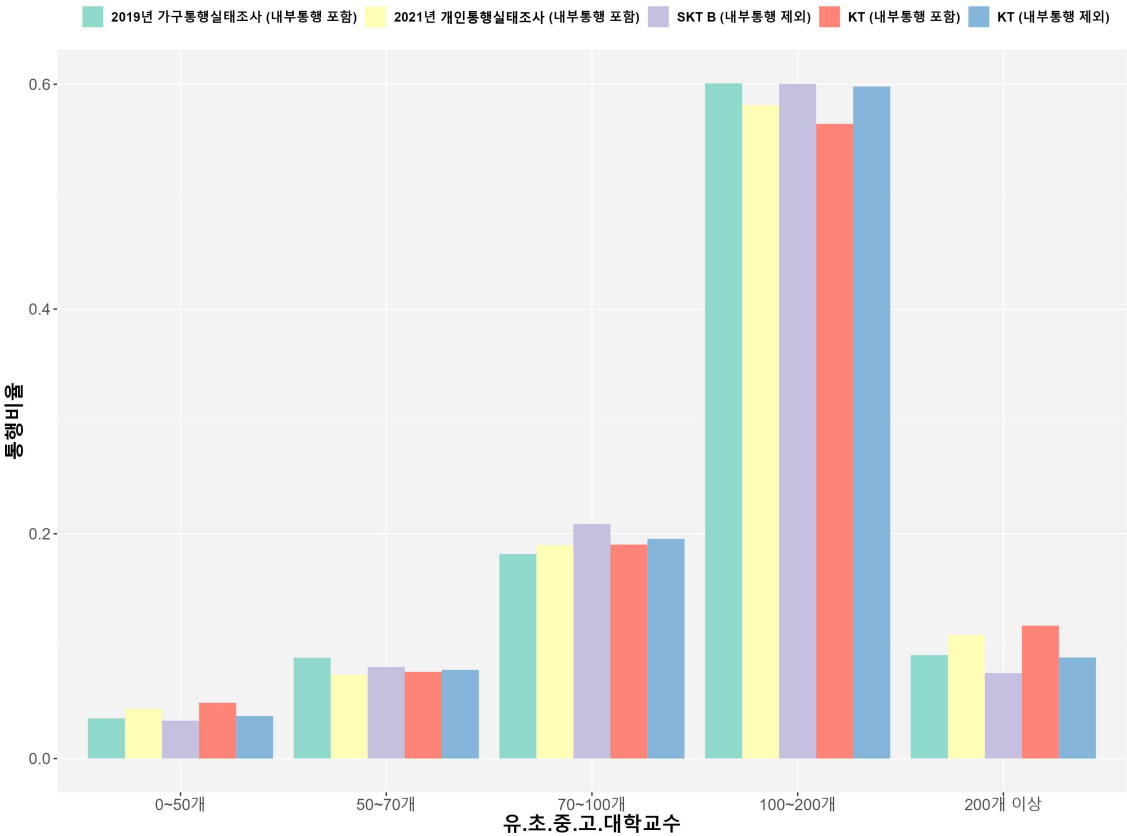
<그림 3-59> (통행도착량) 통학 통행 - 수용학생수에 따른 데이터별 통행량 비교

- <그림 3-60>은 데이터별 통학 목적 통행의 도착량과 수용학생수 간의 관계를 비교한 것임
- 수용학생수에 대한 설명력은 전반적으로 높지 않은 편 (R-Squared: 0.69~0.77)



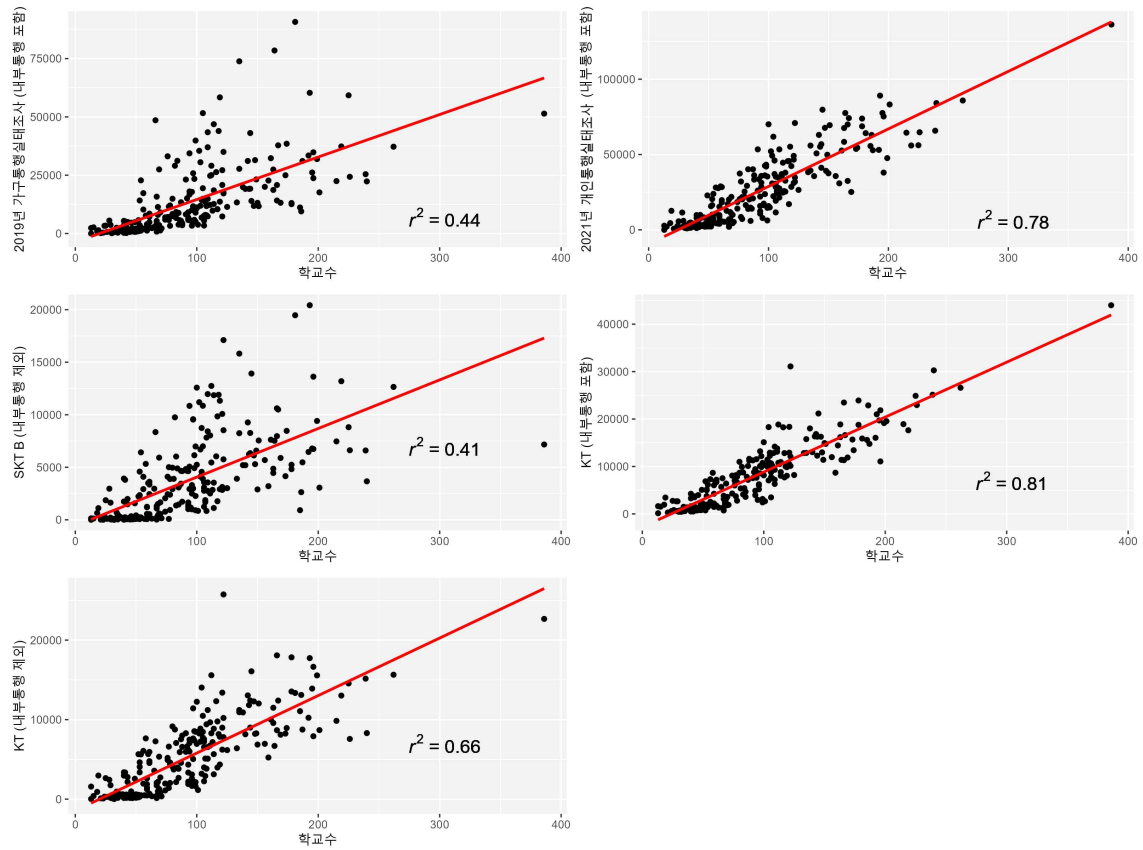
〈그림 3-60〉 (통행도착량) 통학 통행 - 수용학생수와 각 데이터별 상관관계

- 학교수 (유초중고등학교, 대학교 포함)
 - 학교수에 따라 통학 통행의 도착량을 데이터별로 비교한 결과는 다음 <그림 3-61>과 같음
 - 100개 이상 200개 미만인 시군구에서의 통행 비율이 가장 높으며, 데이터 간 통행량 비율도 유사한 것으로 나타남



<그림 3-61> (통행도착량) 통학 통행 - 학교수에 따른 데이터별 통행량 비교

- <그림 3-62>는 각 데이터별로 통학 목적 통행의 도착량과 학교수와의 관계를 비교 분석한 결과임
 - 2019년 가구통행실태조사(내부통행 포함)과 SKT B (내부통행 제외)는 R-Squared값이 0.5 미만으로 설명력이 다소 낮은 편이고, 2021년 개인통행실태조사(내부통행 포함)과 KT(내부통행 포함)은 R-Squared값이 0.7 이상으로 상대적으로 설명력이 높은 것으로 분석됨

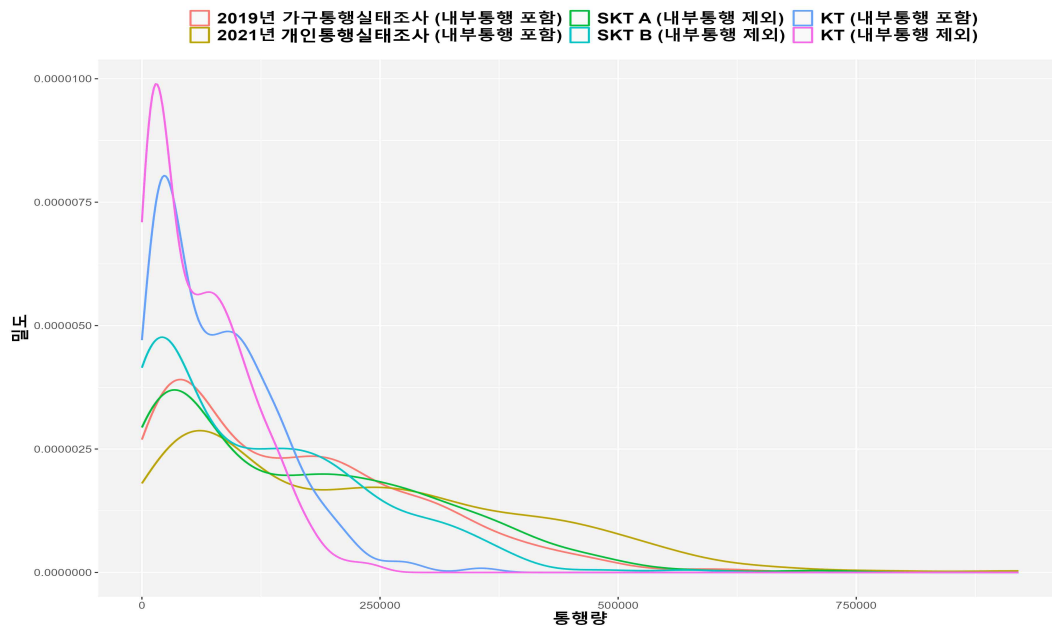


〈그림 3-62〉 (통행도착량) 통학 통행 - 학교수와 각 데이터별 상관관계

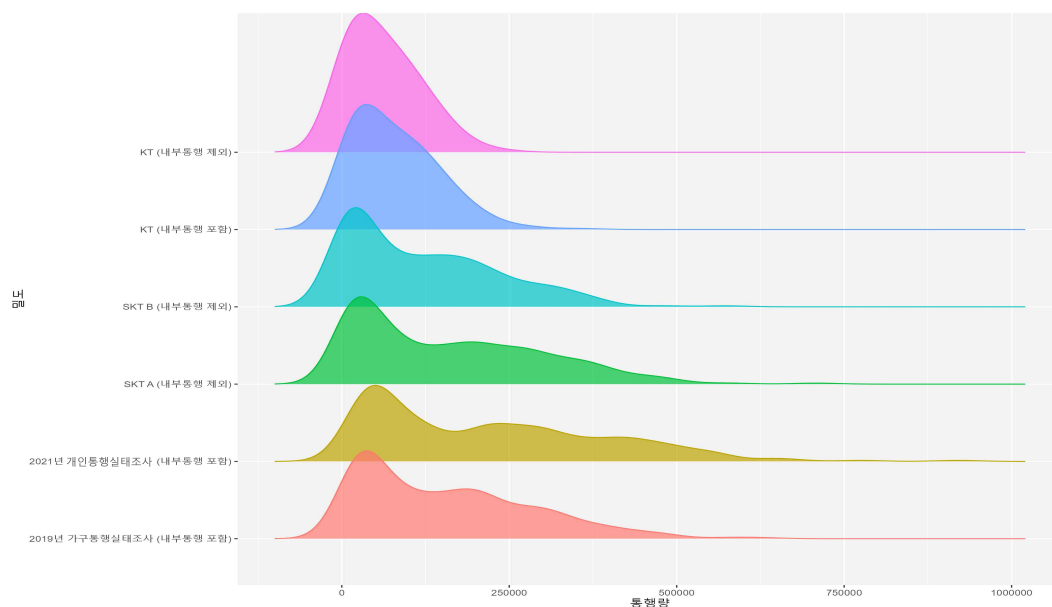
④ 귀가

A) 데이터별 통행량 비교

- 귀가 통행에 대한 도착량 분포를 데이터별로 비교한 결과는 다음 <그림 3-63>, <그림 3-64>와 같음



<그림 3-63> (통행도착량) 귀가 통행 - 데이터별 통행량 빈도 비교 (1)



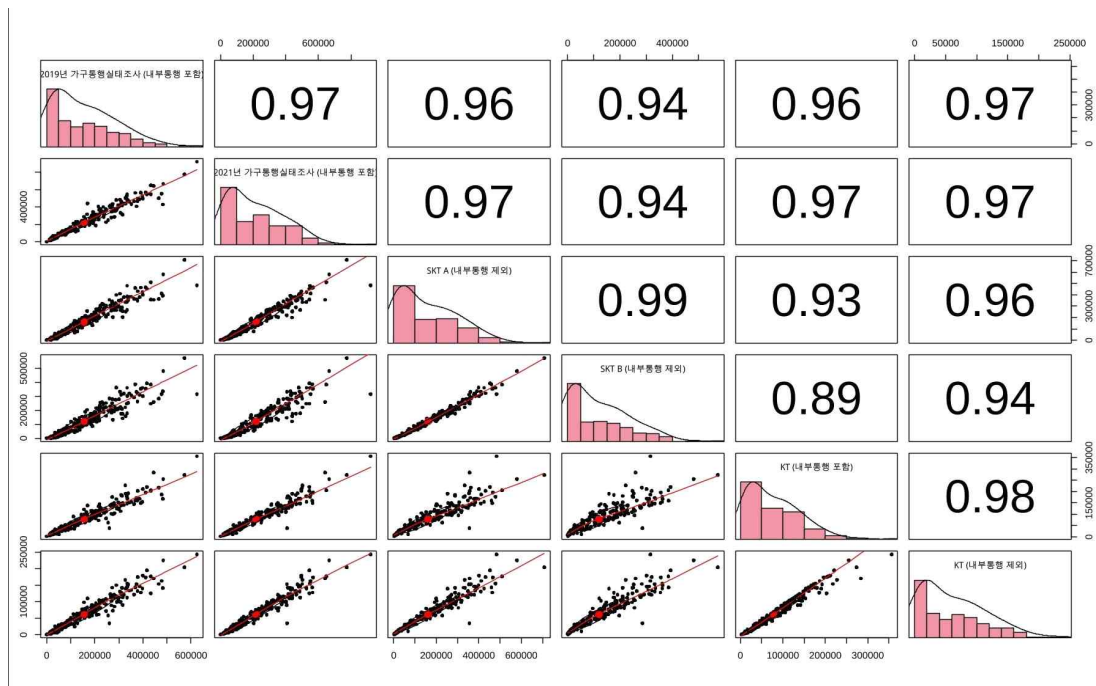
<그림 3-64> (통행도착량) 귀가 통행 - 데이터별 통행량 빈도 비교 (2)

- KT(내부통행 포함, 제외) 기준 귀가 통행량이 비교적 편향된 것으로 나타남

- KT는 25만 통행 미만에서 대부분의 통행이 이루어지는 반면, 타 데이터는 25만 통행 이상의 구간에서도 통행 빈도가 높게 나타남

○ <그림 3-65>은 귀가 목적 통행의 도착량에 대하여 데이터간 상관관계를 분석한 것임

- 모든 데이터간 상관관계가 높은 것으로 분석됨



주: 그림에 표기된 수치는 R(상관계수)값을 의미함.

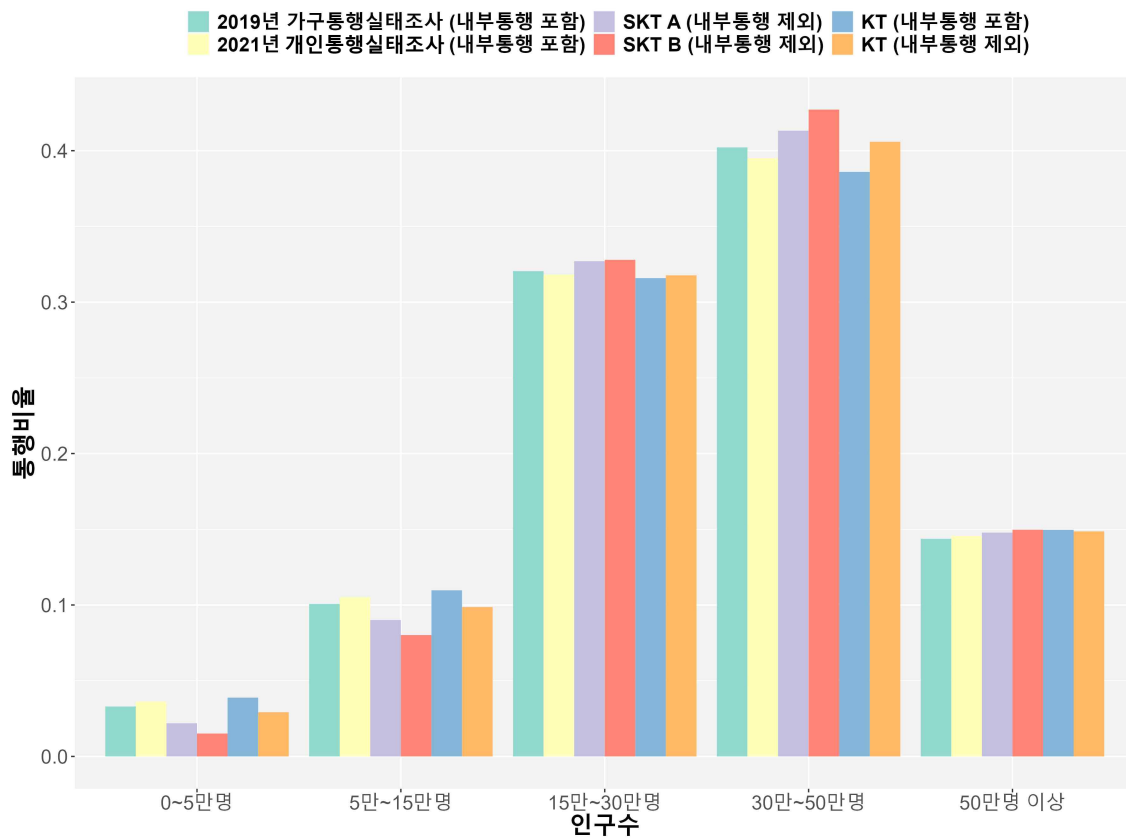
〈그림 3-65〉 (통행도착량) 귀가 통행 - 데이터간 상관관계 비교

B) 통행도착 요인별 통행량 비교

○ 인구수

- 인구수에 따라 데이터별 귀가 통행의 도착량을 비교한 결과는 다음 <그림 3-66>과 같음

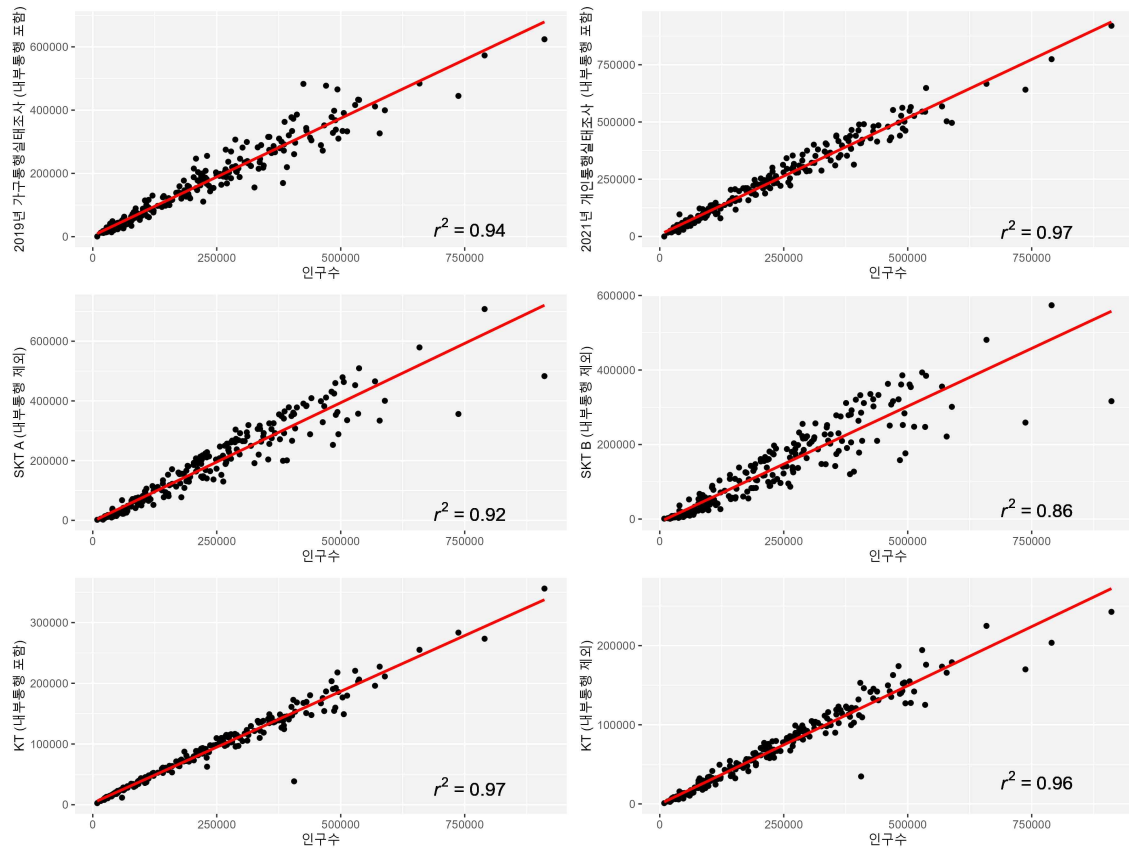
· 인구수 15만 명 미만의 시군구에서는 SKT A(내부통행 제외), SKT B(내부통행 제외)의 통행 비율이 타 데이터에 비해 다소 낮게 나타나며, 30만 명 이상 50만 명 미만의 시군구에서는 반대로 SKT A(내부통행 제외), SKT B(내부통행 제외)의 통행 비율이 타 데이터에 비해 다소 높게 나타남



<그림 3-66> (통행도착량) 귀가 통행 - 인구수에 따른 데이터별 통행량 비교

- <그림 3-67>은 각 데이터의 귀가 통행 도착량과 인구수와의 관계를 분석한 결과임

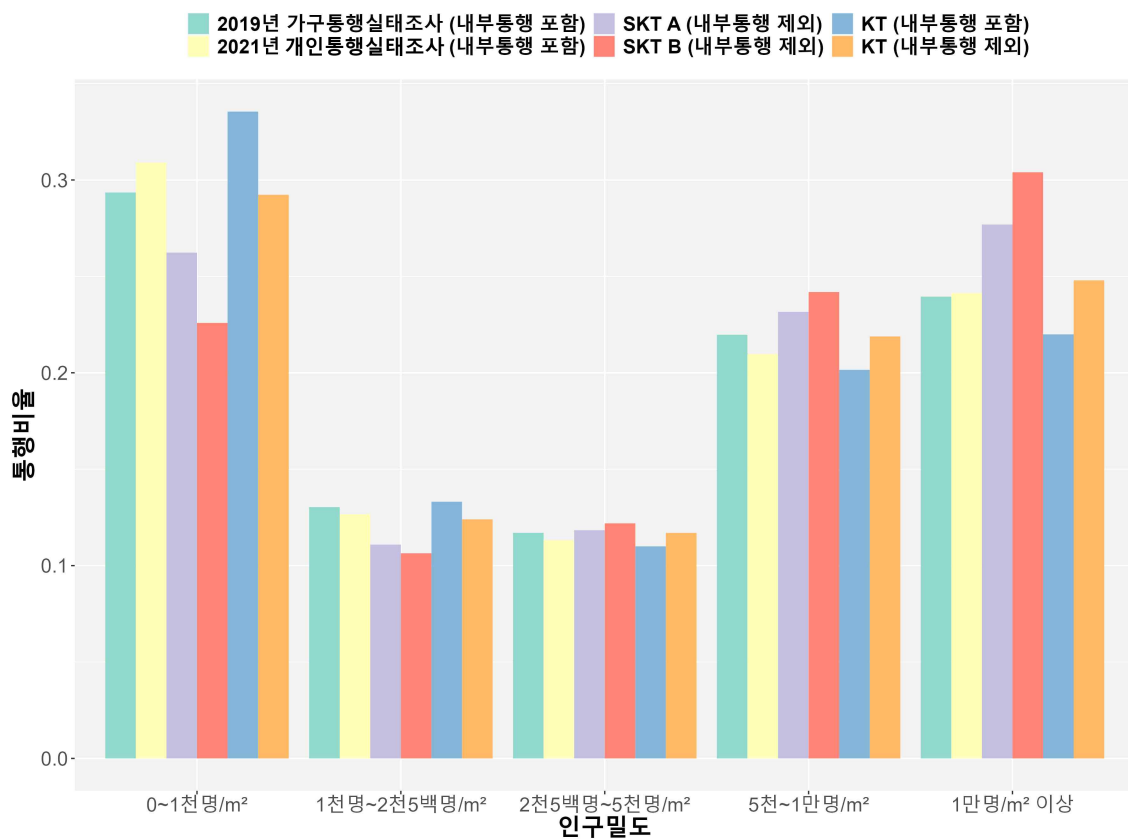
· 전반적으로 인구수에 대한 모든 데이터의 설명력이 높은 편임



〈그림 3-67〉 (통행도착량) 귀가 통행 - 인구수와 각 데이터별 상관관계

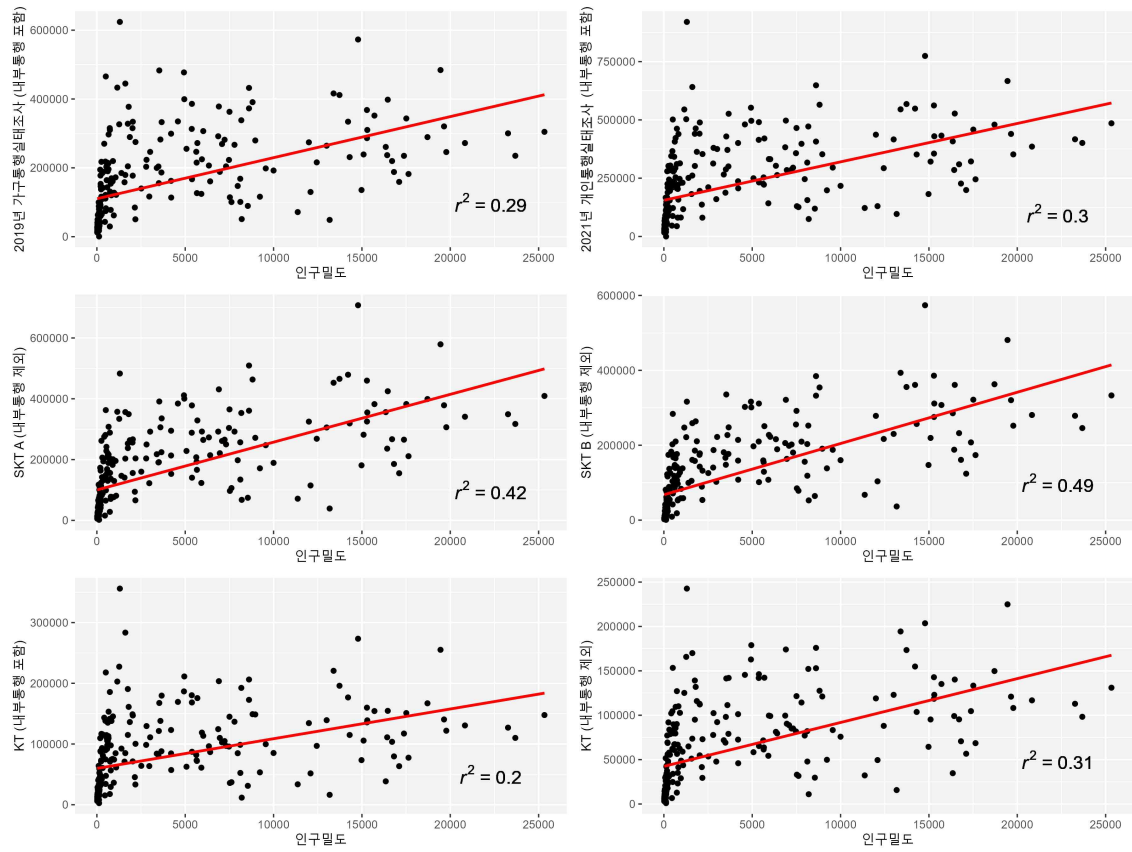
○ 인구밀도

- 인구밀도에 따라 데이터별 귀가 통행의 도착량을 비교한 결과는 다음 <그림 3-68>과 같음
- 인구밀도가 1천 명/㎡ 미만일 때 데이터 간 차이가 가장 크게 나타남
- KT(내부통행 포함)는 인구밀도가 낮은 구간에서 통행 비율이 높게 나타나며, SKT A(내부통행 제외), SKT B(내부통행 제외)는 인구밀도가 1만 명/㎡ 이상의 높은 구간에서 통행 비율이 높게 나타남



<그림 3-68> (통행도착량) 귀가 통행 - 인구밀도에 따른 데이터별 통행량 비교

- 각 데이터의 귀가 통행 도착량과 인구밀도와의 관계를 분석한 결과는 다음 <그림 3-69>와 같음
- 전반적으로 인구밀도를 잘 설명하지 못하는 것으로 나타남



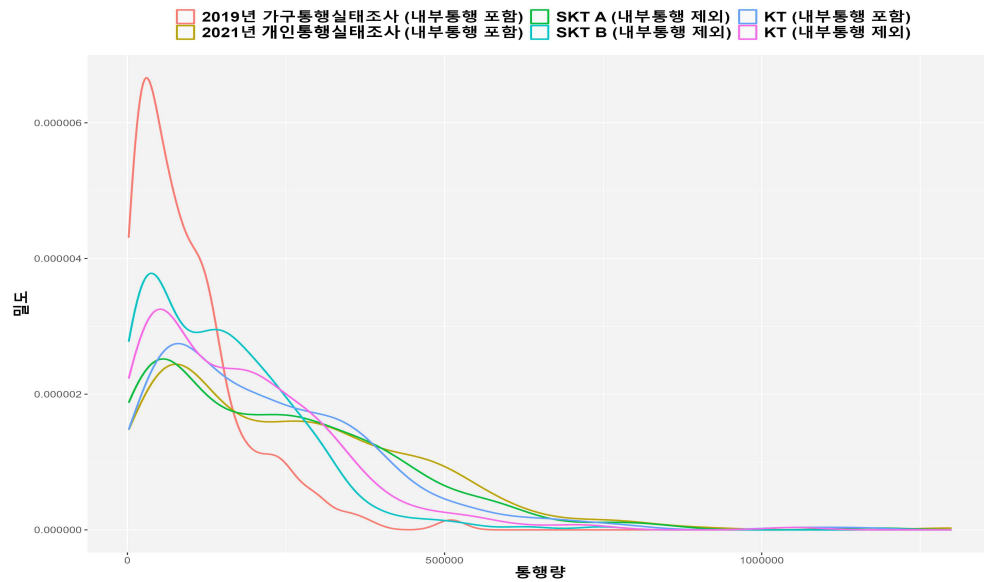
〈그림 3-69〉 (통행도착량) 귀가 통행 - 인구밀도와 각 데이터별 상관관계

⑤ 기타

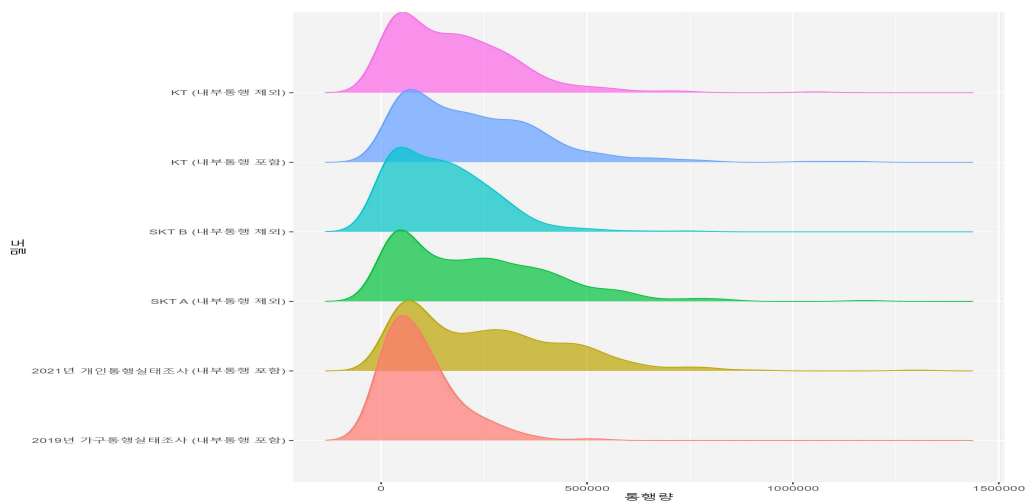
A) 데이터별 통행량 비교

○ <그림 3-70>, <그림 3-71>은 기타 통행에 대한 도착량 분포를 데이터별로 비교한 결과임

- 2019년 가구통행실태조사(내부통행 포함)는 타 데이터와 달리 도착량이 적은 시군구의 비율이 집중적으로 높은 것으로 나타남



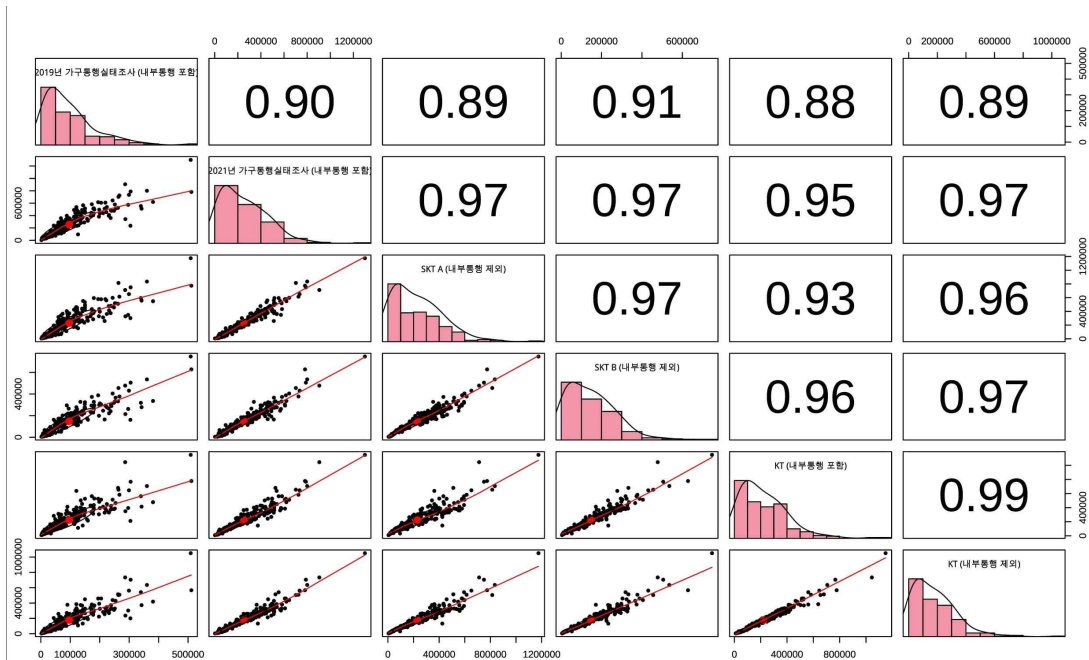
<그림 3-70> (통행도착량) 기타 통행 - 데이터별 통행량 빈도 비교 (1)



<그림 3-71> (통행도착량) 기타 통행 - 데이터별 통행량 빈도 비교 (2)

○ <그림 3-72>은 기타 목적 통행의 도착량에 대하여 데이터 간 유사성을 분석한 결과임

- 전반적으로 모든 데이터 간 상관관계가 높은 편으로 분석됨
- 2019년 가구통행실태조사(내부통행 포함)는 다른 데이터와의 상관계수가 0.88~0.91 정도로 타 데이터 간 상관관계가 상대적으로 낮은 것으로 나타남



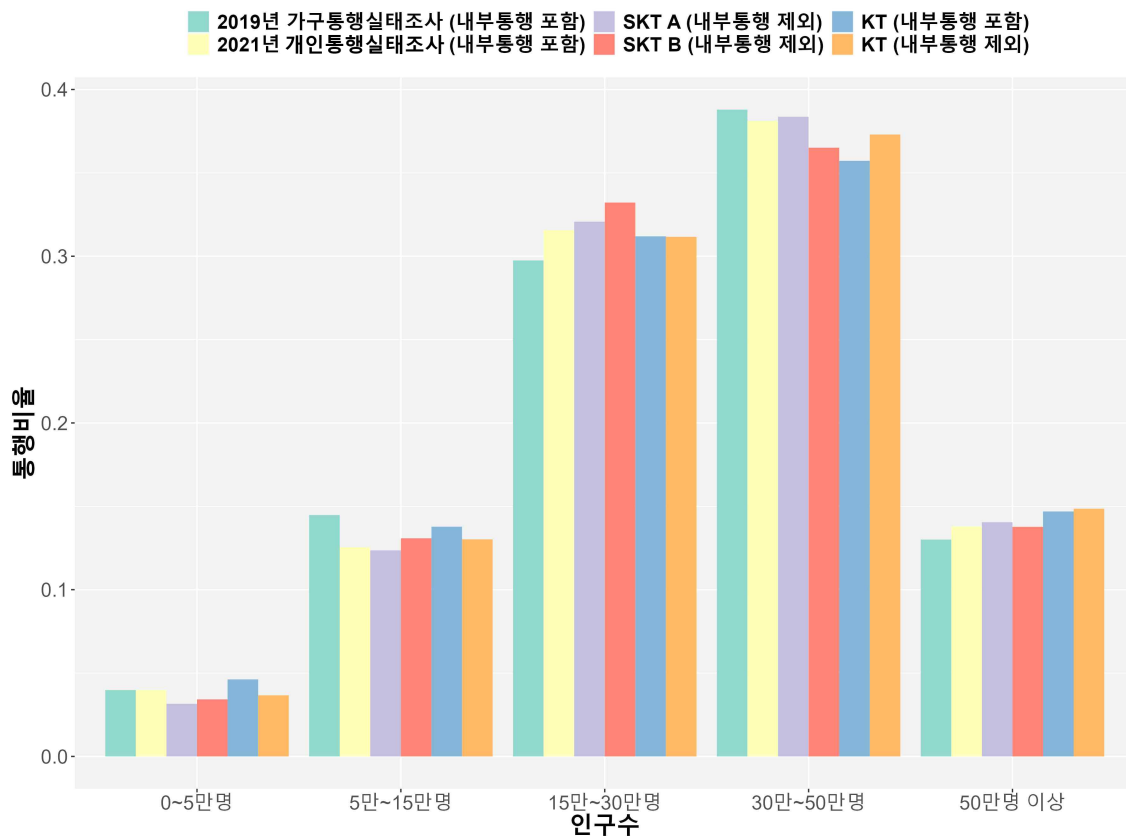
주: 그림에 표기된 수치는 R(상관계수)값을 의미함.

〈그림 3-72〉 (통행도착량) 기타 통행 - 데이터간 상관관계 비교

B) 통행도착 요인별 통행량 비교

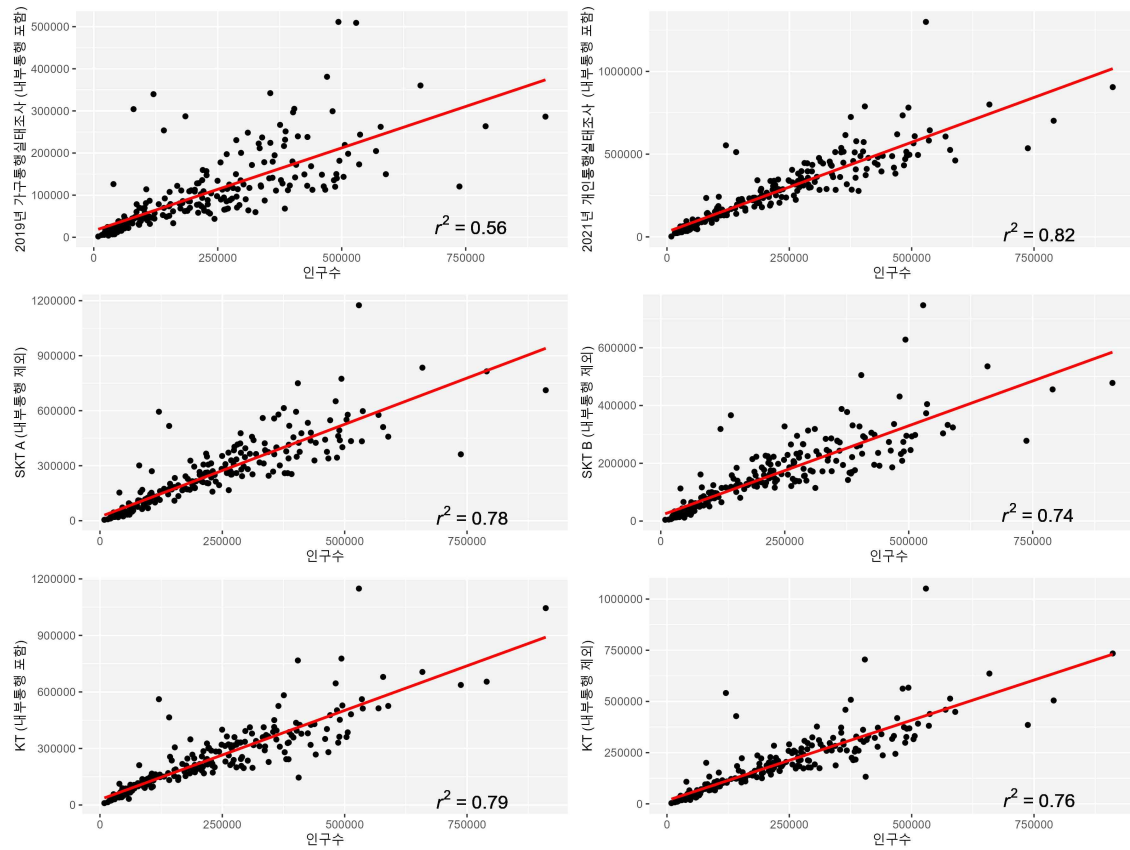
○ 인구수

- <그림 3-73>은 인구 규모에 따라 기타 통행에 대한 도착량을 각 데이터별로 산출하여 비교한 결과이며, 데이터 간 큰 차이를 보이지 않는 것으로 보여짐
- 2019년 가구통행실태조사(내부통행 포함)는 인구 5만 명 이상 15만 명 미만, 30만 명 이상 50만 명 미만일 때 타 데이터에 비해 통행 비율이 높은 것으로 분석되었으며, SKT B (내부통행 제외)는 인구 15만 명 이상 30만 명 미만일 때 타 데이터에 비해 통행 비율이 높은 것으로 분석됨



<그림 3-73> (통행도착량) 기타 통행 - 인구수에 따른 데이터별 통행량 비교

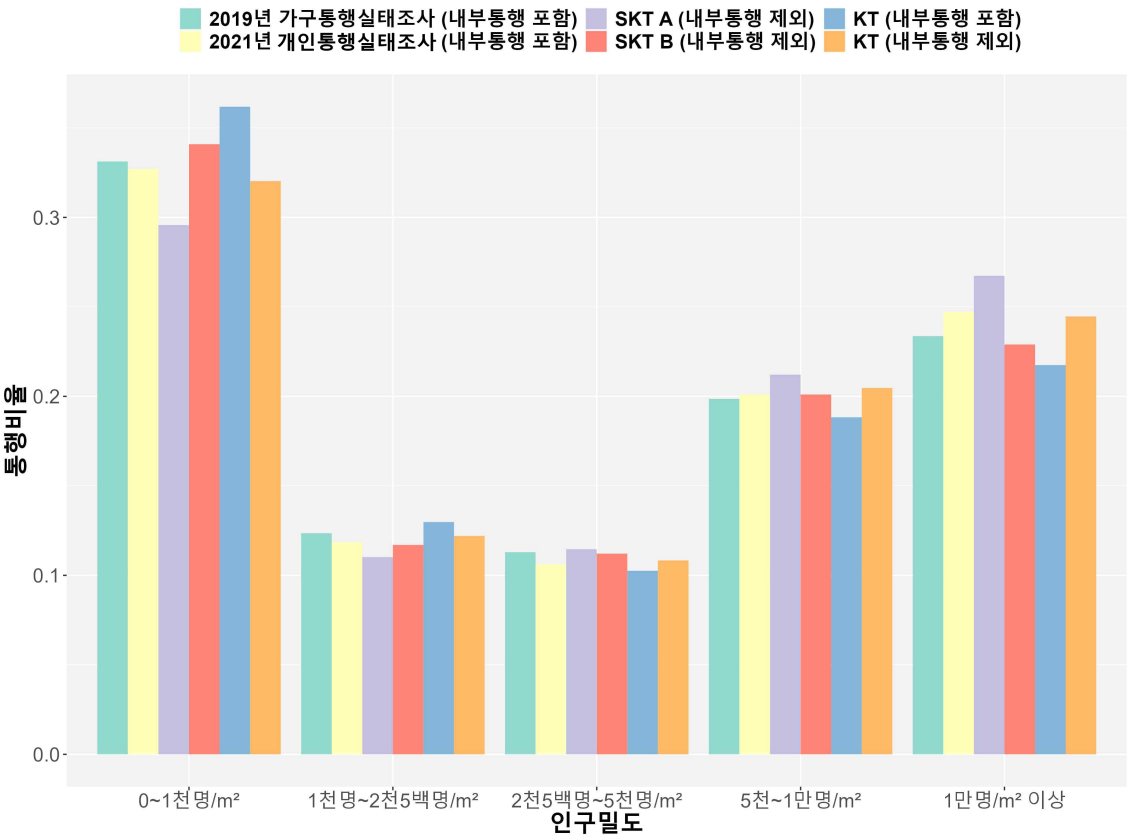
- <그림 3-74>는 인구수를 기준으로 각 데이터의 기타 통행 도착량의 관계를 비교한 것임
- 2019년 가구통행실태조사(내부통행 포함)를 제외한 나머지의 데이터는 인구수 분포에 대한 설명력이 유사한 편 (R-Squared: 0.74~0.82)



〈그림 3-74〉 (통행도착량) 기타 통행 - 인구수와 각 데이터별 상관관계

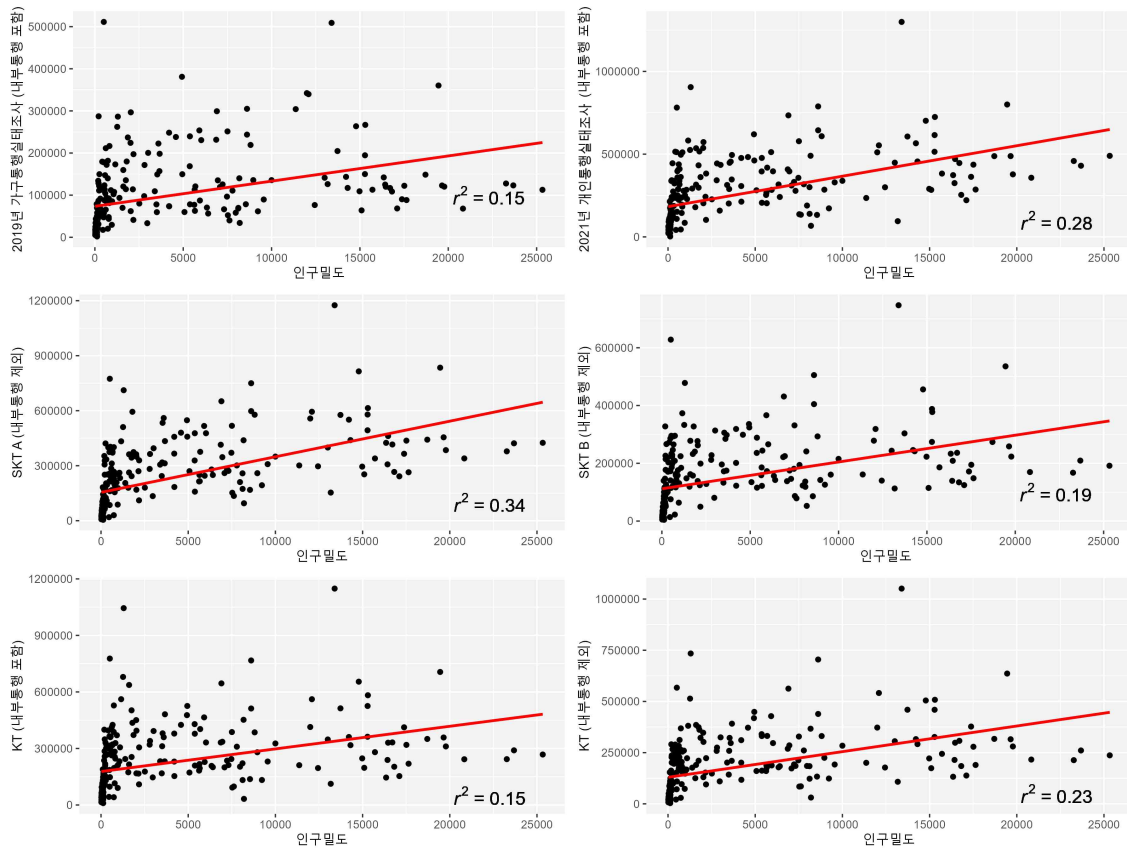
○ 인구밀도

- 인구밀도를 기준으로 각 데이터별 기타 통행에 대한 도착량을 비교하면 다음 <그림 3-75>와 같음
- 인구밀도가 1천 명/㎡ 미만, 1만 명/㎡ 이상일 때 데이터 간 차이가 다소 발생하는 것으로 나타남



<그림 3-75> (통행도착량) 기타 통행 - 인구밀도에 따른 데이터별 통행량 비교

- <그림 3-76>은 기타 통행에 대한 도착량과 인구밀도와의 관계를 데이터별로 분석한 것임
- 전반적으로 인구밀도에 대한 설명력이 낮은 것으로 분석됨



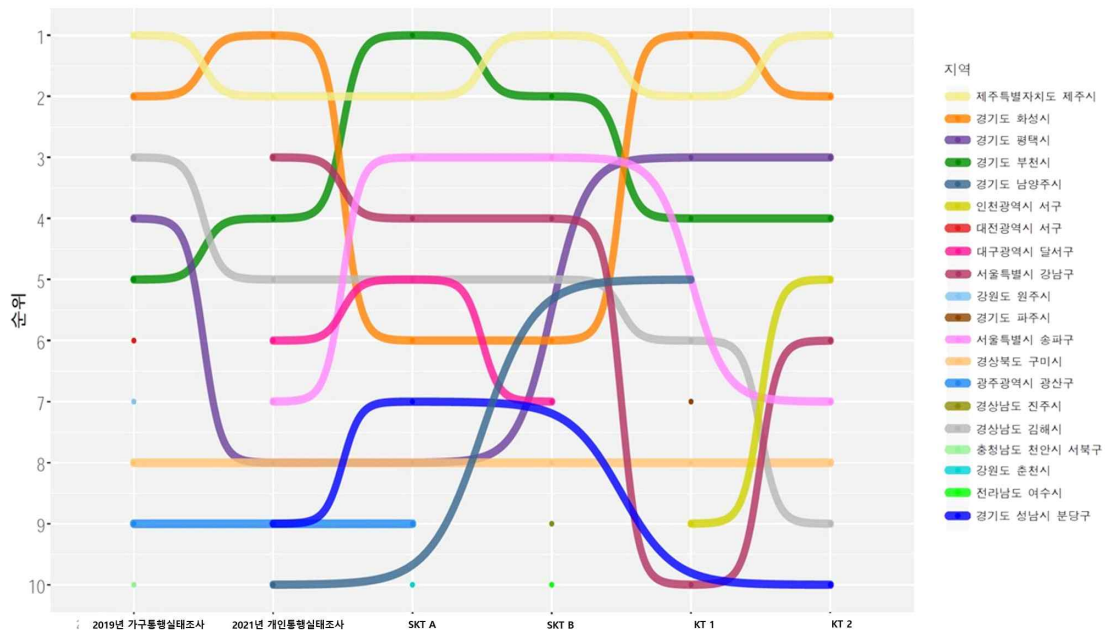
〈그림 3-76〉 (통행도착량) 기타 통행 - 인구밀도와 각 데이터별 상관관계

3) 통행분포량 기준

A) 데이터별 최다 통행구간 비교

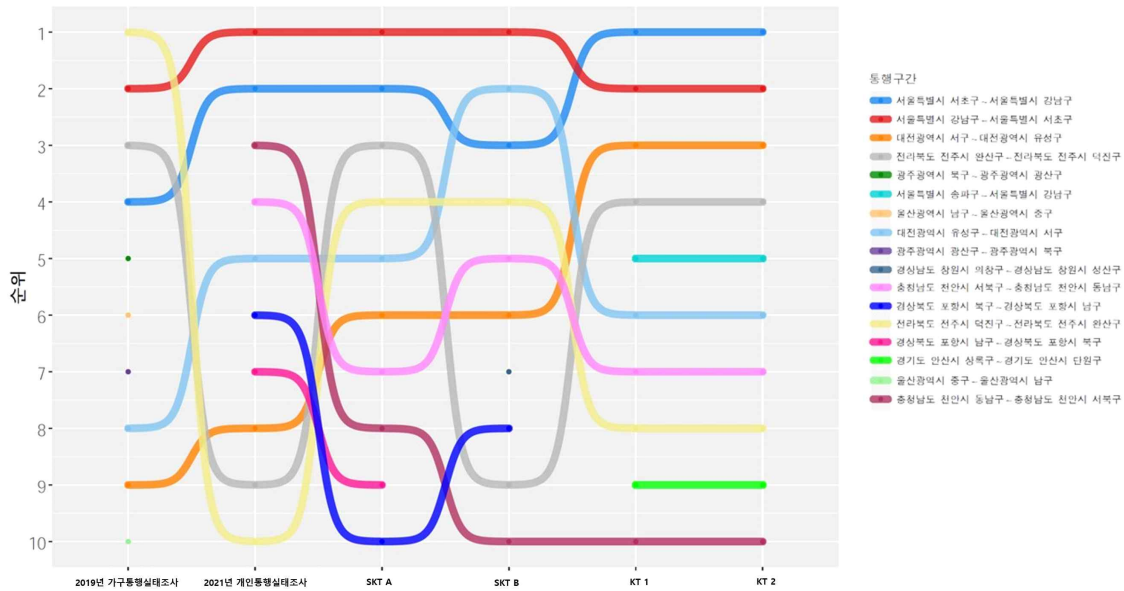
○ 1위부터 10위까지 통행량이 가장 많은 구간은 대체로 모든 데이터에서 유사하게 나타나며 순위 변동 폭이 크지 않음

- 동일 시군구 내에서 이루어진 통행을 기준으로, 제주특별자치도 제주시, 경기도 화성시, 경기도 평택시, 경기도 부천시, 경기도 남양주시, 인천광역시 서구, 대전광역시 서구, 대구광역시 달서구, 서울특별시 강남구, 강원도 원주시, 경기도 파주시, 서울특별시 송파구, 경상북도 구미시, 광주광역시 광산구, 경상남도 진주시, 경상남도 김해시, 충청남도 천안시 서북구, 강원도 춘천시, 전라남도 여수시, 경기도 성남시 분당구



〈그림 3-77〉 데이터별 최다 통행구간 TOP 10 (시군구 내부 통행)

- 외부 통행을 기준으로 서울특별시 강남구~서울특별시 서초구, 전라북도 전주시 덕진구~전라북도 전주시 완산구, 서울특별시 서초구~서울특별시 강남구, 전라북도 전주시 완산구~전라북도 전주시 덕진구, 대전광역시 유성구~대전광역시 서구, 대전광역시 서구~대전광역시 유성구 구간이 모든 데이터에서 10위 안에 포함되어 있음 (<그림 3-78> 참조)



〈그림 3-78〉 데이터별 최다 통행구간 TOP 10 (외부 통행)

- 최다 통행구간 순위권에 있는 지역들의 통행량의 절대값은 차이가 발생하는데, 상대적으로 외부 통행보다 내부 통행량 간 차이가 더 큼
 - 경기도 화성시, 서울특별시 강남구, 경기도 부천시, 대구광역시 달서구, 서울특별시 송파구, 경기도 남양주시, 서울특별시 노원구, 인천광역시 남동구, 경상남도 김해시 순으로 자료간 편차가 크게 나타남(<표 3-20> 참조)
- 조사자료(2019년, 2021년 개인통행실태조사)에서 통행량이 0인 구간(제로셀) 중 모바일 통신데이터 통행량이 100이상인 구간은 그 순위 차이가 크게 벌어짐
 - 가령 경기도 군포시~경기도 안성시 구간은 2019년 가구통행실태조사에서 61,236위, 2021년 개인통행실태조사에서 60,499위를 한 데 반해, SKT A 데이터에서는 6,533위, SKT B 데이터에서는 5,882위, KT데이터에서는 6,334위를 차지함
 - 즉, 샘플수가 상대적으로 작은 조사기반 자료에서는 제로셀이 다수 발생하며, 이러한 제로셀의 통행량은 샘플수가 많은 모바일 통신데이터에서 확인할 수 있음을 의미함

〈표 3-20〉 (통행분포량) 전체 - 데이터별 최다 통행지역 TOP 10 (시군구 내 통행)

순위	2019년 가구통행실태조사 (내부통행 포함)		SKT A (내부통행 제외)		KT (내부통행 포함)	
	지역	통행량	지역	통행량	지역	통행량
1위	제주특별자치도 제주시	1,134,595	경기도 부천시	1,156,470	경기도 화성시	1,072,724
2위	경기도 화성시	892,906	제주특별자치도 제주시	1,081,744	제주특별자치도 제주시	1,008,437
3위	경상남도 김해시	641,795	서울특별시 송파구	869,311	경기도 평택시	813,173
4위	경기도 평택시	620,184	서울특별시 강남구	849,093	경기도 부천시	704,951
5위	경기도 부천시	592,956	대구광역시 달서구	758,077	경기도 남양주시	699,494
6위	대전광역시 서구	571,555	경기도 화성시	743,900	경상남도 김해시	689,062
7위	강원도 원주시	534,895	경기도 성남시 분당구	732,884	경기도 파주시	585,532
8위	경상북도 구미시	528,522	경기도 평택시	699,157	경상북도 구미시	574,110
9위	광주광역시 광산구	514,208	광주광역시 광산구	687,875	인천광역시 서구	567,479
10위	충청남도 천안시 서북구	487,149	강원도 춘천시	679,564	서울특별시 강남구	562,149
순위	2021년 개인통행실태조사 (내부통행 포함)		SKT B (내부통행 제외)		KT (내부통행 제외)	
	지역	통행량	지역	통행량	지역	통행량
1위	경기도 화성시	1,668,091	제주특별자치도 제주시	837,319	제주특별자치도 제주시	696,972
2위	제주특별자치도 제주시	1,395,183	경기도 부천시	688,151	경기도 화성시	581,744
3위	서울특별시 강남구	1,345,444	서울특별시 송파구	525,363	경기도 평택시	550,031
4위	경기도 부천시	1,277,017	서울특별시 강남구	521,372	경기도 부천시	446,430
5위	경상남도 김해시	1,165,027	경상남도 김해시	483,079	인천광역시 서구	440,803
6위	대구광역시 달서구	1,158,316	경기도 화성시	476,302	서울특별시 강남구	424,556
7위	서울특별시 송파구	1,136,162	대구광역시 달서구	464,658	서울특별시 송파구	401,535
8위	경기도 평택시	1,046,293	경상북도 구미시	461,469	경상북도 구미시	389,050
9위	경기도 성남시 분당구	1,024,489	경상남도 진주시	459,694	경상남도 김해시	385,785
10위	경기도 남양주시	979,159	전라남도 여수시	458,302	경기도 성남시 분당구	383,908

〈표 3-21〉 (통행분포량) 전체 - 데이터별 최다 통행구간 TOP 10 (외부 통행)

순위	2019년 가구통행실태조사 (내부통행 포함)			SKT A (내부통행 제외)			KT (내부통행 포함)		
	출발지	~	도착지	통행량	출발지	~	도착지	통행량	통행량
1위	전라북도 전주시 덕진구	~	전라북도 전주시 완산구	183,789	서울특별시 강남구	~	서울특별시 서초구	195,635	141,213
2위	서울특별시 강남구	~	서울특별시 서초구	172,351	서울특별시 서초구	~	서울특별시 강남구	194,316	132,289
3위	전라북도 전주시 완산구	~	전라북도 전주시 덕진구	165,788	전라북도 전주시 완산구	~	전라북도 전주시 덕진구	185,664	128,317
4위	서울특별시 서초구	~	서울특별시 강남구	164,981	전라북도 전주시 덕진구	~	전라북도 전주시 완산구	184,732	115,925
5위	광주광역시 북구	~	광주광역시 광산구	162,448	대전광역시 유성구	~	대전광역시 서구	165,883	114,478
6위	울산광역시 남구	~	울산광역시 중구	154,605	대전광역시 서구	~	대전광역시 유성구	165,684	113,831
7위	광주광역시 광산구	~	광주광역시 북구	151,386	충청남도 천안시 서북구	~	충청남도 천안시 동남구	158,056	112,758
8위	대전광역시 유성구	~	대전광역시 서구	150,830	충청남도 천안시 동남구	~	충청남도 천안시 서북구	157,554	110,353
9위	대전광역시 서구	~	대전광역시 유성구	147,181	경상북도 포항시 남구	~	경상북도 포항시 북구	145,675	108,824
10위	울산광역시 중구	~	울산광역시 남구	144,772	경상북도 포항시 북구	~	경상북도 포항시 남구	145,335	106,064
순위	2021년 개인통행실태조사 (내부통행 포함)			SKT B (내부통행 제외)			KT (내부통행 제외)		
	출발지	~	도착지	통행량	출발지	~	도착지	통행량	통행량
1위	서울특별시 강남구	~	서울특별시 서초구	185,567	서울특별시 강남구	~	서울특별시 서초구	149,066	141,213
2위	서울특별시 서초구	~	서울특별시 강남구	183,775	대전광역시 유성구	~	대전광역시 서구	146,477	132,289
3위	충청남도 천안시 동남구	~	충청남도 천안시 서북구	172,999	서울특별시 서초구	~	서울특별시 강남구	144,764	128,317
4위	충청남도 천안시 서북구	~	충청남도 천안시 동남구	170,266	전라북도 전주시 덕진구	~	전라북도 전주시 완산구	135,679	115,925
5위	대전광역시 유성구	~	대전광역시 서구	166,849	충청남도 천안시 서북구	~	충청남도 천안시 동남구	131,481	114,478
6위	경상북도 포항시 북구	~	경상북도 포항시 남구	160,616	대전광역시 서구	~	대전광역시 유성구	126,601	113,831
7위	경상북도 포항시 남구	~	경상북도 포항시 북구	158,794	경상남도 창원시 의창구	~	경상남도 창원시 성산구	120,650	112,758
8위	대전광역시 서구	~	대전광역시 유성구	145,291	경상북도 포항시 북구	~	경상북도 포항시 남구	120,278	110,353
9위	전라북도 전주시 완산구	~	전라북도 전주시 덕진구	144,335	전라북도 전주시 완산구	~	전라북도 전주시 덕진구	117,626	108,824
10위	전라북도 전주시 덕진구	~	전라북도 전주시 완산구	141,303	충청남도 천안시 동남구	~	충청남도 천안시 서북구	116,190	106,064



〈그림 3-79〉 통행량이 가장 많은 구간 Top 1000

나. 읍면동 단위

- 수도권 내부의 통근 통행, 통학 통행에 대하여, 2019년 가구통행실태조사(내부통행 포함), SKT B(내부통행 제외), KT(내부통행 포함) 세 가지 데이터를 비교 분석함
 - 읍면동 단위로 데이터 수집·비교가 가능한 수도권 내부로 한정함
 - 2019년 가구통행실태조사에서 수도권 목적OD는 출퇴근, 등하교, 학원, 쇼핑, 기타 등 다양한 목적으로 분류되나, 통신데이터와 명확하게 비교할 수 있는 출퇴근, 등하교 통행만을 분석 대상으로 선정함
 - 2019년 가구통행실태조사의 목적 OD는 귀가 통행이 따로 분류되지 않고 출퇴근, 등하교가 하나의 범주로 묶여, 통신데이터의 OD 역시 귀가 통행을 구분하지 않았으며 출퇴근, 등하교로 구분함
 - 2019년 가구통행실태조사와 SKT B는 20세 이상의 통학 통행을 포함하고 있으나, KT는 20세 미만의 통학 통행만을 포함함
 - 읍면동별 통행량(발생량, 도착량) 분포의 유사성을 확인하고, 통행 발생·도착 요인을 기준으로 각 데이터의 통행량의 관계성을 살펴봄

1) 데이터별 통행량 비교

- 2019년 가구통행실태조사를 기준으로 SKT B와 KT 데이터의 통행발생량·도착량을 비교한 결과는 <표 3-22>와 같음
 - 통근 통행은 발생량, 도착량 모두 KT가 SKT B보다 2019년 가구통행실태조사와 유사한 것으로 나타남
 - 발생량의 경우 2019년 가구통행실태조사와 상관관계는 SKT B가 높게 나타나나, 그 외 데이터 분산의 유사도나 오차량으로 볼 때 KT가 SKT B에 비해 상대적으로 2019년 가구통행실태조사와 유사한 것을 알 수 있음
 - 통학 통행은 SKT B가 KT보다 2019년 가구통행실태조사와 분산이 유사하나, 오차량은 KT가 상대적으로 적은 것으로 나타남

〈표 3-22〉 읍면동별 통행량 분포 비교 결과 (2019년 가구통행실태조사 기준)

구분			R2 ¹⁾	Pearson R2 ²⁾	Explained Variance Score ³⁾	MAPE ⁴⁾
통행량	통행목적	데이터				
발생량	통근	SKT B	-0.337	0.841	0.327	0.669
		KT	0.261	0.326	0.617	0.593
	통학	SKT B	-0.17	0.542	0.242	2.497
		KT	-0.59	0.324	0.08	1.28
도착량	통근	SKT B	-0.345	0.769	0.343	0.695
		KT	0.242	0.787	0.619	0.596
	통학	SKT B	-0.133	0.557	0.223	1.72
		KT	-0.427	0.495	0.157	1.10

주 1) 예측값과 실측값의 선형관계를 평가하는 지표로서, 1에 가까울수록 선형성이 강함

2) 예측값과 실측값의 단순 상관관계를 평가함, -1~1 사이의 값을 가짐

3) 예측값과 실측값의 분산을 평가하는 지표로서, 1에 가까울수록 분산이 유사함

4) 오차량 절대값의 백분율량을 평가하는 지표로서, 0에 가까울수록 오차량이 적음

○ 2019년 가구통행실태조사, SKT B, KT 세 가지 데이터의 통행발생량과 도착량을 동시에 비교한 결과는 다음 <표 3-23>과 같음

- 통근 통행량은 거의 대부분의 읍면동에서 분포에 유의한 차이가 있는 것으로 분석되었으나, 통근 통행은 통근 통행보다 상대적으로 많은 읍면동에서 유의한 차이가 없는 것으로 분석됨

〈표 3-23〉 읍면동별 통행량 분포 비교 결과 (3가지 데이터 동시 비교)

P-value 기준 0.05		통근		통학	
		발생	도착	발생	도착
ANOVA 분석결과	분포가 유의한 차이가 있음	937	941	545	498
	분포가 유의한 차이가 없음	21	17	410	457

○ SKT B, KT 두 가지 데이터의 통행발생량과 도착량을 상호 비교한 결과는 <표 3-24>와 같음

- 통근 통행은 대부분의 읍면동에서 분포에 유의한 차이가 있는 것으로 확인되었으나, 통학 통행은 대부분 읍면동에서 분포에 유의한 차이가 없는 것으로 분석됨

〈표 3-24〉 읍면동별 통행량 분포 비교 결과 (모바일 통신데이터간 비교)

P-value 기준 0.05		통근		통학	
		발생	도착	발생	도착
KS 검정 (Kolmogorov-Smirnov) 결과	분포가 유의한 차이가 있음	861	729	320	413
	분포가 유의한 차이가 없음	169	301	707	614

2) 통행발생·도착요인별 통행량 비교

○ 통행발생·도착요인과 각 데이터의 통행발생량·도착량의 관계를 분석한 결과는 <표 3-25>, <표 3-26>와 같음

- 통근·통학 통행의 발생, 도착과 관련이 있는 사회경제지표를 기준으로 분석함
 - 통근 통행: 발생요인으로 취업자수를, 도착요인으로 종사자수를 설정함
 - 통학 통행: 발생요인으로 학령인구수를, 도착요인으로 수용학생수를 설정함
- 통근 통행 분석결과, 2019년 가구통행실태조사의 서울 지역에 대한 통근 통행 발생량을 제외한 나머지 통근 통행 발생량은 대체로 취업자수를 잘 설명하는 것으로 나타났으며, 통근 통행 도착량은 지역, 데이터에 상관없이 전반적으로 종사자수를 잘 설명하는 것으로 나타남
 - 경기 지역보다는 서울, 인천 지역에서 사회경제지표에 대한 데이터별 설명력 차이가 나타났으며, 취업자수의 경우 2019년 가구통행실태조사보다는 모바일 통신데이터의 설명력이 조금 더 높게 나타남
- 통학 통행 분석결과, KT는 20세 미만의 통학 통행량만을 대상으로 하고 있어 19세 이상의 학령인구수에 대한 설명력은 다소 낮게 나타나는 것으로 확인되었으며, 19세 미만의 학령인구수의 경우 서울 지역은 KT, 경기 지역은 2019년 가구통행실태조사의 설명력이 상대적으로 높고, 인천 지역은 데이터별 차이가 거의 없는 것으로 분석됨
- 또한 도착량을 기준으로 볼 때 수용학생수를 상대적으로 가장 잘 설명하는 것은 2019년 가구통행실태조사로 분석됨
 - SKT B는 경기 지역에서만 설명력이 높고, KT는 인천 지역에서만 설명력이 높았으나, 2019년 가구통행실태조사는 모든 지역에서 설명력이 높게 나타남

〈표 3-25〉 통근 통행에 대한 회귀모형 구축결과

구분			구분	계수값 ¹⁾	t값 ²⁾	p-value ³⁾	Adj-R2 ⁴⁾
통행량	변수	지역					
발생량	취업자수	서울	2019년 가구통행실태조사	2.362	21.826	0.000	0.529
			SKT B	0.763	34.558	0.000	0.740
			KT	1.109	68.520	0.000	0.918
		경기	2019년 가구통행실태조사	2.021	47.835	0.000	0.803
			SKT B	0.581	50.660	0.000	0.854
			KT	1.094	44.116	0.000	0.816
		인천	2019년 가구통행실태조사	2.044	22.937	0.000	0.779
			SKT B	0.643	28.943	0.000	0.896
			KT	1.063	28.300	0.000	0.892
도착량	종사자수	서울	2019년 가구통행실태조사	1.536	49.196	0.000	0.851
			SKT B	0.433	34.000	0.000	0.734
			KT	1.039	47.741	0.000	0.845
		경기	2019년 가구통행실태조사	1.936	47.569	0.000	0.802
			SKT B	0.502	46.255	0.000	0.829
			KT	0.969	46.330	0.000	0.830
		인천	2019년 가구통행실태조사	1.977	21.745	0.000	0.760
			SKT B	0.518	16.266	0.000	0.728
			KT	1.048	28.748	0.000	0.894

주 1) 회귀계수로서 독립변수 변화에 따른 종속변수의 변화를 설명함
2) 회귀계수가 유의한지 검정하기 위한 값으로, 독립변수와 종속변수간의 선형관계가 존재하는 정도를 나타냄
3) 독립변수의 유의확률을 나타내며, 0.05보다 작으면 유의미한 것으로 판단함
4) 조정된 결정계수로서, 독립변수의 개수와 표본 크기를 고려하여 결정계수를 조정한 값을 의미함

〈표 3-26〉 통학 통행에 대한 회귀모형 구축결과

통행량	구분		구분	계수값 ¹⁾	t값 ²⁾	p-value ³⁾	Adj-R2 ⁴⁾
	변수	지역					
발생량	학령인구수 (7~18세)	서울	2019년 가구통행실태조사	2.467	20.181	0.000	0.491
			SKT B	1.152	30.427	0.000	0.688
			KT	0.553	49.105	0.000	0.852
		경기	2019년 가구통행실태조사	1.918	54.586	0.000	0.842
			SKT B	0.573	32.398	0.000	0.704
			KT	0.428	29.571	0.000	0.666
		인천	2019년 가구통행실태조사	1.953	18.709	0.000	0.701
			SKT B	0.739	18.940	0.000	0.785
			KT	0.503	17.603	0.000	0.762
	학령인구수 (19~23세)	서울	2019년 가구통행실태조사	4.369	36.089	0.000	0.755
			SKT B	1.692	34.568	0.000	0.740
			KT	0.702	30.268	0.000	0.686
		경기	2019년 가구통행실태조사	3.725	52.172	0.000	0.829
			SKT B	1.104	35.418	0.000	0.740
			KT	0.788	27.750	0.000	0.637
		인천	2019년 가구통행실태조사	3.926	24.813	0.000	0.805
			SKT B	1.343	20.742	0.000	0.814
			KT	0.857	14.887	0.000	0.696
도착량	수용학생수	서울	2019년 가구통행실태조사	1.425	54.137	0.000	0.874
			SKT B	0.390	20.567	0.000	0.502
			KT	0.197	17.954	0.000	0.433
		경기	2019년 가구통행실태조사	1.543	61.494	0.000	0.871
			SKT B	0.521	15.077	0.000	0.701
			KT	0.297	24.924	0.000	0.586
		인천	2019년 가구통행실태조사	1.628	37.252	0.000	0.902
			SKT B	0.387	25.621	0.000	0.599
			KT	0.372	15.488	0.000	0.712

주 1) 회귀계수로써 독립변수 변화에 따른 종속변수의 변화를 설명함

2) 회귀계수가 유의한지 검정하기 위한 값으로, 독립변수와 종속변수간의 선형관계가 존재하는 정도를 나타냄

3) 독립변수의 유의확률을 나타내며, 0.05보다 작으면 유의미한 것으로 판단함

4) 조정된 결정계수로써, 독립변수의 개수와 표본 크기를 고려하여 결정계수를 조정한 값을 의미함

2. 실적자료와의 일치성 검증 결과

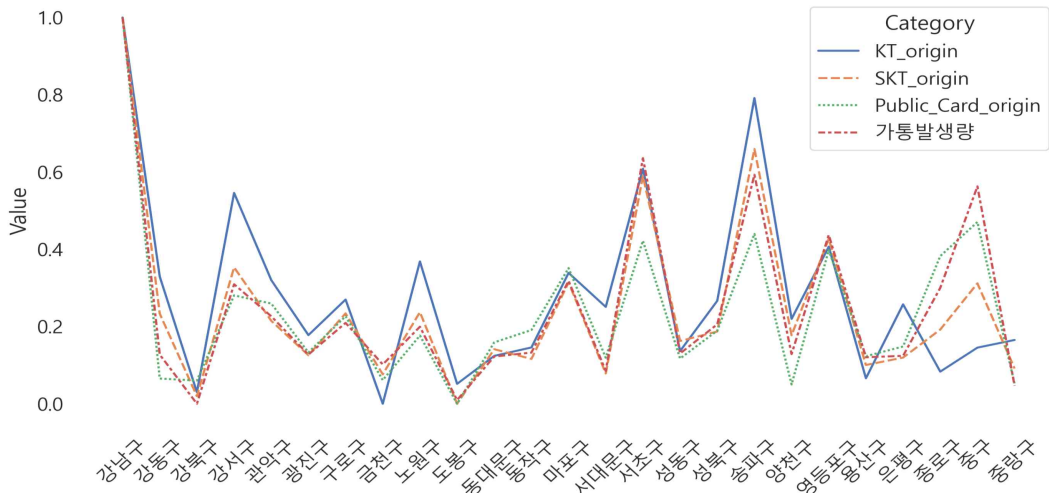
가. 대중교통 실적자료 기준

- 서울시의 대중교통 실적자료(교통카드 데이터)를 기준으로 2019년 가구통행실태조사(내부통행 포함), SKT B(내부통행 제외), KT(내부통행 포함)을 검증함
 - SKT·KT에는 대중교통 외 승용차의 통행도 포함되어 있으므로, 대중교통 실적자료와 스케일을 맞추기 위해 Min-Max scale을 한 후 ANOVA, Bartlett, Levene 검정을 수행함
- 분석결과, 2019년 가구통행실태조사, KT, SKT B, 교통카드 데이터는 대체로 통계적으로 유사한 특성을 갖는 것으로 확인됨
 - 종로구를 제외한 나머지 25개 시군구에서 통행발생량이 유사하게 나타남
 - 종로구에서 KT의 통행발생량이 타 데이터에 비해 다소 낮게 측정됨

<표 3-27> 시군구별 KT·SKT·KTDB·교통카드 데이터의 통행발생량 통계검증 결과

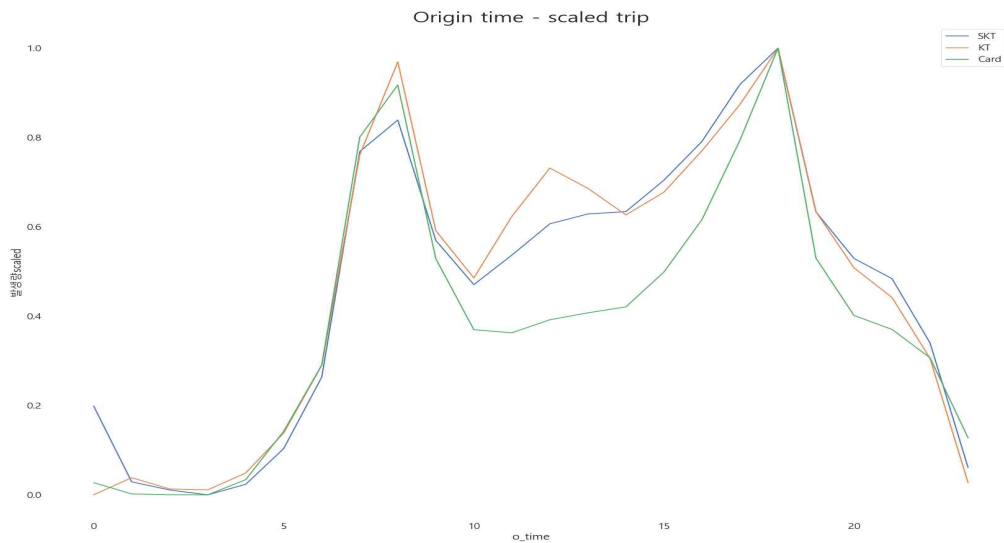
구분	ANOVA	Bartlett	Levene
P-value	0.88	0.93	0.96

주) ANOVA 검정은 집단의 평균을 비교하는 검정이며, Bartlett, Levene 검정은 ANOVA 검정 수행 전 대상 집단의 분산이 같은지 다른지 확인하는 테스트로, 유의값(p-value)이 0.05 이상인 경우에는 집단간 차이가 없다고 해석함

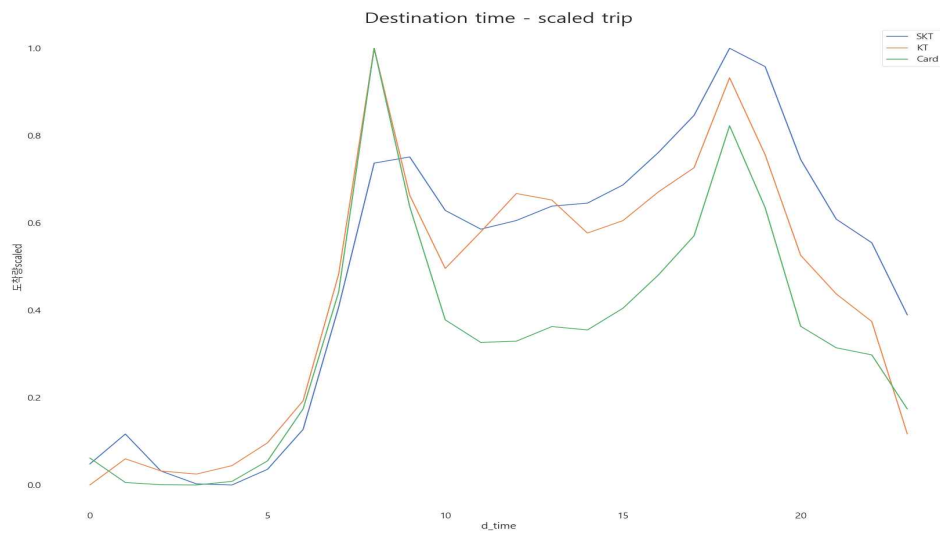


<그림 3-80> 시군구별 KT·SKT·KTDB·교통카드 데이터의 통행발생량 비교

- 시간대별로 통행발생량·도착량을 비교 분석한 결과, 첨두시간대에는 모두 비슷한 패턴을 보이거나, 일부 시간대(10~15시)에서는 SKT B, KT, 교통카드 모두 상이한 패턴을 보이는 것으로 확인됨
- 시간대별 통행량을 알 수 없는 2019년 가구통행실태조사는 분석에서 제외함
- 발생량은 오전, 오후 첨두시간대 모두 유사한 패턴을 보이거나(<그림 3-81> 참조), 도착량은 오전 첨두시간대에만 유사한 패턴을 보임(<그림 3-82> 참조)



〈그림 3-81〉 시간대별 KT·SKT 교통카드의 통행발생량 비교



〈그림 3-82〉 시간대별 KT·SKT 교통카드의 통행도착량 비교

나. 수송실적 자료 기준

- 지역간 통행량, 권역간 통행량, 도서지역 통행량에 대하여 모바일 통신데이터와 실적자료의 결과가 일치하는지 확인하였으며, 비교한 결과는 <표 3-28>과 같음
 - 지역간 통행량: 전국 6X3 스크린라인을 통과하는 교통량을 비교한 결과, -58.7% ~ 15.5% 차이가 발생하는 것으로 나타남
 - 권역간 통행량: 수도권, 부산권, 강원권, 전라권 등 권역간 통행량을 비교한 결과, 25.8% ~ 80.5% 차이가 발생하는 것으로 나타남
 - 도서지역 통행량: 도서지역으로 유입·유출하는 통행량을 비교한 결과, -65.4% ~ 103.1% 차이가 발생하는 것으로 나타남

<표 3-28> 모바일 통신데이터 기반 지역간 통행량 추정 결과와 수송실적간의 비교

(단위 : 천통행/일)

구분		수송실적							통신	차이
		승용차	버스	트럭	철도	항공	해운	합계		
가로축 1	상→하	1,016	257	185	54	0	0	1,513	691	-54.3%
	하→상	1,160	294	213	52	0	0	1,719	711	-58.7%
가로축 2	상→하	512	184	164	93	10	0	963	736	-23.6%
	하→상	514	200	166	95	10	0	984	718	-27.0%
가로축 3	상→하	559	210	157	73	10	0	1,008	672	-33.4%
	하→상	575	213	159	70	10	0	1,028	684	-33.4%
가로축 4	상→하	201	76	83	61	10	0	431	464	7.7%
	하→상	248	86	104	58	10	0	505	461	-8.7%
가로축 5	상→하	226	98	104	43	10	0	480	467	-2.8%
	하→상	239	104	110	41	10	0	503	460	-8.5%
가로축 6	상→하	480	209	141	41	8	0	880	669	-23.9%
	하→상	487	215	155	40	8	0	904	680	-24.8%
세로축 1	좌→우	1,363	542	340	102	9	1	2,356	1,450	-38.5%
	하→상	1,491	642	387	105	9	1	2,636	1,400	-46.9%
세로축 2	좌→우	242	105	79	55	8	1	489	565	15.5%
	하→상	273	122	92	53	8	1	549	540	-1.8%
세로축 3	좌→우	315	102	88	33	8	0	545	603	10.6%
	하→상	496	142	156	31	8	0	833	608	-27.0%

제3절 보정 방법 개발

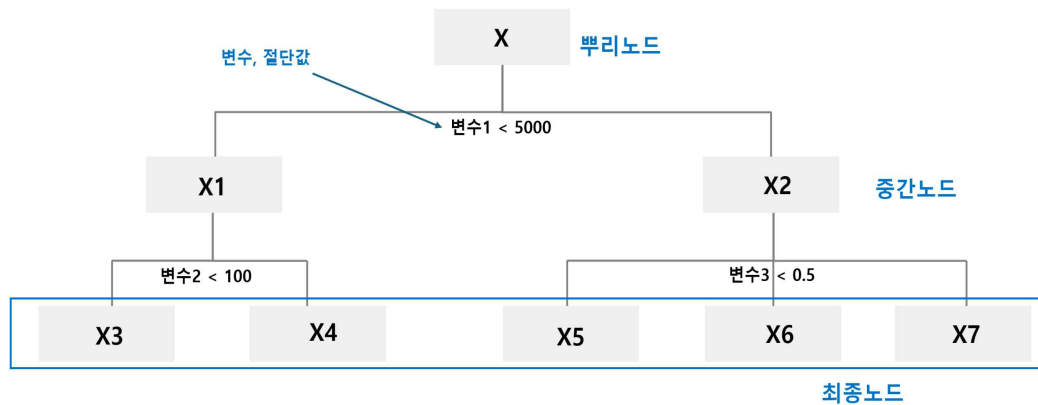
1. 개요

- 제2절에서 살펴본 것과 같이 통행량 및 통행분포는 시도 단위 또는 전체 통행과 같이 큰 카테고리 내에서는 데이터 간에 큰 차이가 없지만, 공간 단위, 통행 목적, 출발·도착 등 카테고리가 세분화될수록 그 차이가 점차 커짐
- 시공간적 해상도가 높다는 모바일 통신데이터의 장점을 유지하면서 실제 통행을 잘 반영하기 위해서는 기존에 통행자료로 활용되었던 데이터들과 차이가 크게 나타나는 구간을 실제 통행과 비교하여 보정할 필요가 있음
- 따라서 본 절에서는 제2절의 결과를 토대로 보정할 모바일 통신데이터와 보정 기준값을 선정하고 기준값과 모바일 통신데이터의 차이를 보정하는 방법에 대해 설명하고자 함

2. 보정가중치 산출 방법

가. 개요

- 기준값과 모바일 통신데이터의 통행량 차이가 나타나는 통행구간의 특성을 파악하고, 보정대상에 대한 가중치를 산출하기 위하여 의사결정나무분석기법을 활용함
 - 의사결정나무분석은 특정 기준값으로 집단을 분류하고 분류된 하위 집단을 다시 특정 기준을 찾아 분류하는 과정을 반복함으로써 종속변수와 설명변수 간의 관계를 확인하는 데 활용되는 분석 방법임
 - 분석 결과 중 데이터를 분류하는 데 사용한 분류기준(변수, 절단값)은 통행량 차이가 나타나는 통행구간을 유형화하는 기준으로 활용할 수 있으며, 최종노드(끝마디)의 평균, 표준편차값은 보정가중치로 활용할 수 있음



〈그림 3-83〉 의사결정나무구조

나. 대상 설정

- 분석 데이터: 모바일 통신데이터는 KT 모바일 통신데이터로 기준값은 2019년 가구통행실태조사로 선정함
 - 모바일 통신데이터는 가공 알고리즘에 따라 산출되는 결과값이 달라질 수 있음. 그러므로, 산출 결과를 해석하고 보정하기 위해서는 연구진이 알고리즘을 설계하고 직접 구축한 KT 모바일 통신데이터를 보정대상으로 선정함
 - 기준값은 조사자료 중 다른 통신사 데이터로 보정하지 않은 2019년 가구통행 실태조사 자료로 설정함
- 분석 대상: 시군구 단위의 OD 통행량(분포량)에 대하여 분석하고자 하며, 통행 목적별 통행량 차이가 나타나는 통행구간의 특성이 다를 수 있으므로, 전체 통행과 더불어 네 가지 통행목적(통근, 통학, 귀가, 기타)에 대하여 분석하고자 함
 - 제2절의 검증 결과에서 2019년 가구통행실태조사와 KT 모바일 통신데이터는 시군구 단위의 통행발생·도착요인을 기준으로 분석할 때 차이가 분명하게 나타나므로, 시군구 단위에서 통행발생·도착요인이 복합적으로 작용하는 OD 통행량을 기준으로 분석하고자 함

다. 산출 과정

- (1단계: 데이터 수집) 2019년 가구통행실태조사 통행 DB, KT 모바일 통신데이터 통행 DB, 그리고 이들의 관계를 설명할 설명변수를 수집함

- 설명변수는 통행목적별 통행량 차이에 영향을 줄 것으로 예상되는 통행발생·도착요인을 선택하여 <표 3-29>와 같이 설정함

〈표 3-29〉 통행목적별 설명변수

통행 목적	통행량과의 연관성		
	출발지 특성	도착지 특성	통행구간의 특성
전체	인구수, 학령인구수, 취업자수, 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	인구수, 수용학생수, 종사자수, 학교수, 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	통행거리
통근	인구수, 학령인구수, 생산인구수(비율), 고령인구수(비율), 취업자수, 기지국수(밀도), 면적	인구수, 생산인구수, 고령인구수, 종사자수, 기지국수(밀도), 면적	통행거리
통학	인구수, 학령인구수, 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	인구수, 수용학생수, 학교수(유·초·중·고등학교수, 유·초·중·고·대학교수), 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	통행거리
귀가	인구수, 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	인구수, 생산인구수(비율), 고령인구수(비율), 기지국수(밀도), 면적	통행거리
기타	인구수, 생산인구수(비율), 고령인구수, 기지국수(밀도), 면적	인구수, 생산인구수(비율), 고령인구수, 기지국수(밀도), 면적	통행거리

- (2단계: 데이터 연계 및 가공) 수집한 데이터를 통행구간별로 조인한 후 전체 통행에 대한 DB, 통근, 통학, 귀가, 기타 통행으로 구분한 통행목적별 DB를 각각 나누어 구축함
- (3단계: 데이터 분할) 의사결정나무구조를 작성하는 데 활용할 훈련데이터와 이를 검증할 테스트 데이터를 7:3의 비율로 분할함
- (4단계: 종속변수 및 설명변수 선택) 2019년 가구통행실태조사의 통행량과 KT 모바일 통신데이터의 통행량의 차이를 나타내는 값(절대오차값을 Box-cox로 변환한 값)을 종속변수로 설정하고, 수집한 설명변수간 상관관계를 분석하여 상관계수 $R \geq |0.2|$ 를 충족하는 변수를 설명변수로 선택함
- (5단계: 의사결정나무구조 작성) CART(Classification And Regression Trees)

알고리즘 적용하여, 분산의 감소량을 기준으로 전체, 통근, 통학, 귀가, 기타 통행에 대한 각 의사결정나무구조를 작성함

- (6단계: 가지치기 수행) 나무의 능력을 향상시키는 새로운 분할이 발견될 때까지 나무를 크게 만든 후 교차검증(cross validation)을 이용한 가지치기를 하여 과적합을 막고 최종 나무의 예측력을 높임

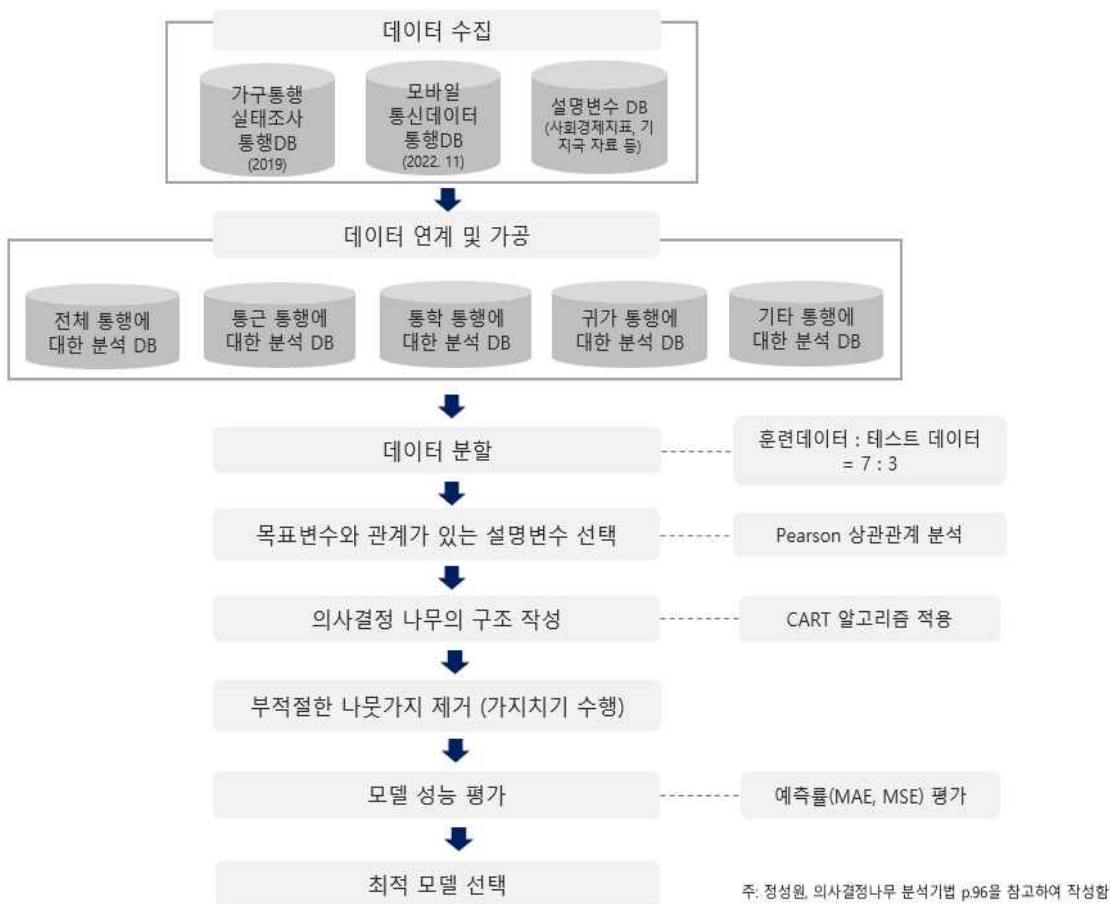
- (7단계: 모델 성능 평가) 테스트 데이터를 기준으로 의사결정나무 분석을 통해 예측한 종속변수 값과 실제값을 기준으로 평균제곱오차(Mean Squared Error, MSE), 평균절대오차(Mean Absolute Error, MAE)를 산출하여 모델의 성능을 평가함

- 평균제곱오차(MSE): 오차의 제곱에 대해 평균을 취한 값이며, 아래 산식에 의해 계산함

$$\frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2$$

- 평균절대오차(MAE): 절대오차에 대해 평균을 취한 값이며, 아래 산식에 의해 계산함

$$\frac{1}{n} \sum_{i=1}^n |\hat{Y}_i - Y_i|$$



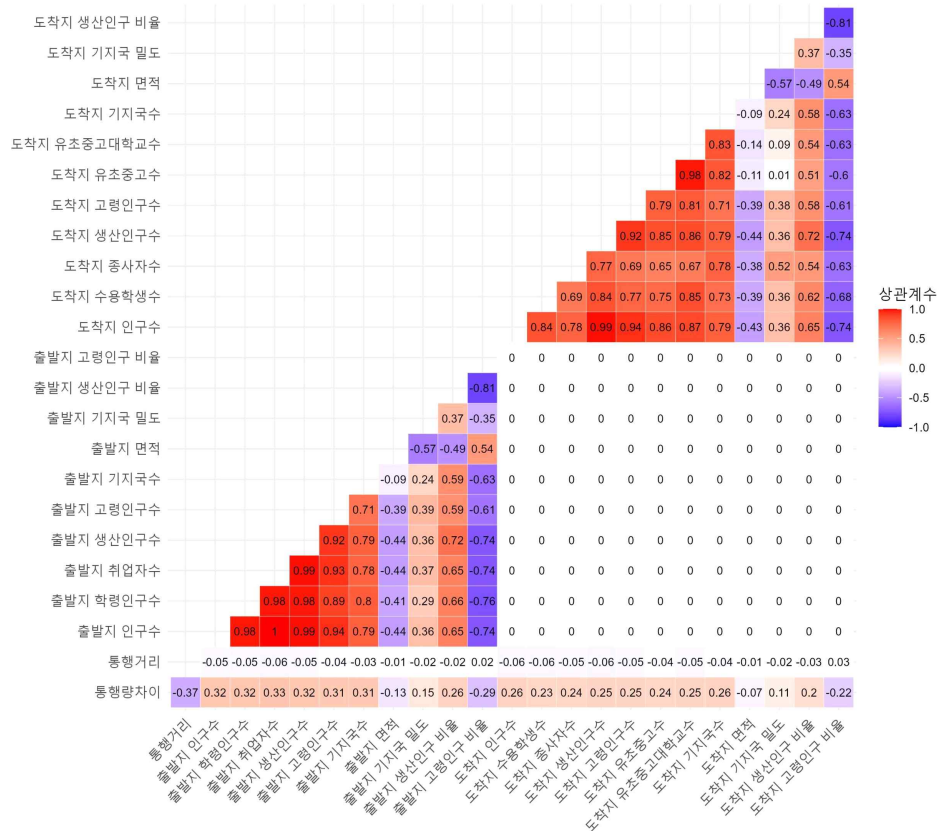
〈그림 3-84〉 통행목적별 보정기준 산출 과정

3. 보정가중치 산출 결과

가. 전체 통행

○ 분석조건

- 종속변수: 2019년 가구통행실태조사의 전체 통행량과 KT 모바일 통신데이터의 전체통행량의 절대오차를 Box-cox로 변환한 값
- 설명변수: 상관계수 분석 결과 종속변수와 영향이 적은 변수를 제외한 19개의 변수
 - 통행거리, 출발지 인구수, 출발지 학령인구수, 출발지 취업자수, 출발지 생산인구수, 출발지 고령인구수, 출발지 기지국수, 출발지 생산인구 비율, 출발지 고령인구 비율, 도착지 인구수, 도착지 수용학생수, 도착지 종사자수, 도착지 생산인구수, 도착지 고령인구수, 도착지 유·초·중·고등학교수, 도착지 유·초·중·고·대학교수, 도착지 기지국수, 도착지 생산인구 비율, 도착지 고령인구 비율

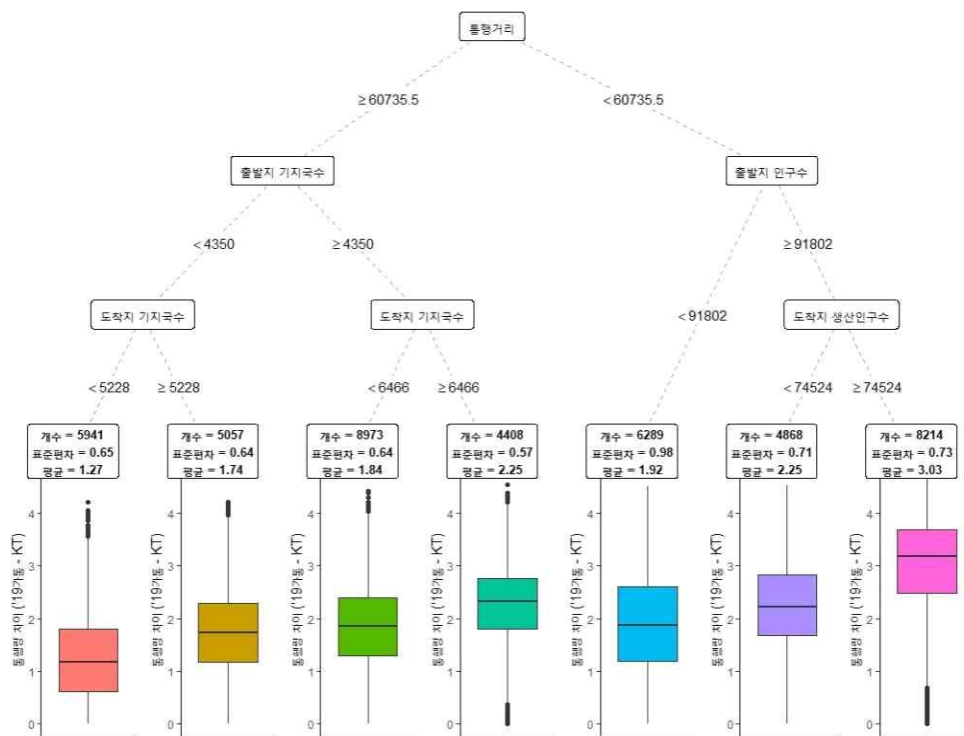


〈그림 3-85〉 (전체 통행) 설명변수 선정 기준

○ 분석결과

- 다음 7가지의 통행패턴에서 전체 통행량에 대한 차이가 발생하는 것으로 분석됨

- ① 통행거리가 60,736km 이상이고, 출발지 기지국수가 4,350개 미만이며, 도착지 기지국 수가 5,228개 미만인 경우
- ② 통행거리가 60,736km 이상이고, 출발지 기지국수가 4,350개 미만이며, 도착지 기지국 수가 5,228개 이상인 경우
- ③ 통행거리가 60,736km 이상이고, 출발지 기지국수가 4,350개 이상이며, 도착지 기지국 수가 6,466개 미만인 경우
- ④ 통행거리가 60,736km 이상이고, 출발지 기지국수가 4,350개 이상이며, 도착지 기지국 수가 6,466개 이상인 경우
- ⑤ 통행거리가 60,736km 미만이고, 출발지 인구수가 91,802명 미만인 경우
- ⑥ 통행거리가 60,736km 미만이고, 출발지 인구수가 91,802명 이상이며, 도착지 생산인구수가 74,524명 미만인 경우
- ⑦ 통행거리가 60,736km 미만이고, 출발지 인구수가 91,802명 이상이며, 도착지 생산인구수가 74,524명 이상인 경우



〈그림 3-86〉 (전체 통행) 의사결정나무구조

- 모델에 대한 성능은 평균제곱오차(MSE) 0.71, 절대평균오차(MAE) 0.68로 나타남
- 전체 통행에 대한 유형별 보정가중치 산출 결과는 <표 3-30>과 같음

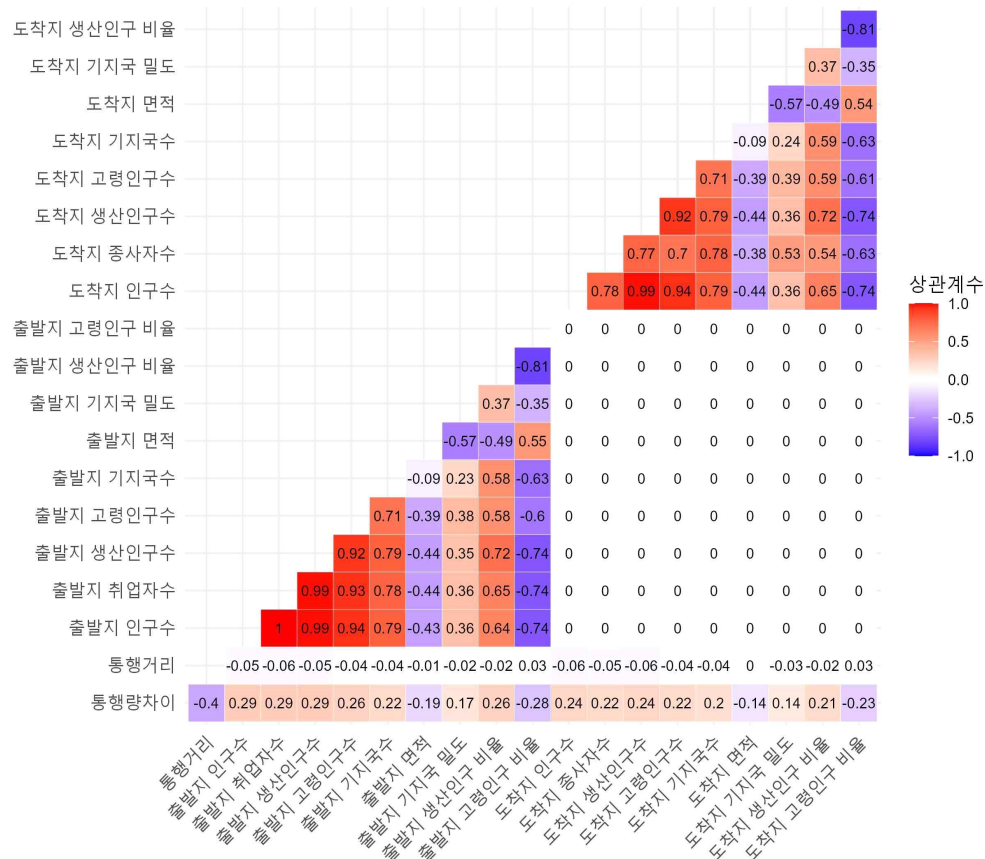
<표 3-30> 전체 통행 유형별 보정가중치

통행 유형	보정유형								보정가중치	
	통행 거리	인구수		기지국수 (밀도)		생산인구수 (비율)		수용 학생수	평균	표준 편차
		출발지	도착지	출발지	도착지	출발지	도착지	도착지		
①	≥ 60,736	-	-	<4,350	<5,228	-	-	-	1.27	0.65
②	≥ 60,736	-	-	<4,350	≥ 5,228	-	-	-	1.74	0.64
③	≥ 60,736	-	-	≥ 4,350	<6,466	-	-	-	1.84	0.64
④	≥ 60,736	-	-	≥ 4,350	≥ 6,466	-	-	-	2.25	0.57
⑤	<60,736	<91,802	-	-	-	-	-	-	1.92	0.98
⑥	<60,736	≥ 91,802	-	-	-	-	<74,524	-	2.25	0.71
⑦	<60,736	≥ 91,802	-	-	-	-	≥ 74,524	-	3.03	0.73

나. 통근 통행

○ 분석조건

- 종속변수: 2019년 가구통행실태조사의 통근 통행량과 KT 모바일 통신데이터의 통근 통행량의 절대오차를 Box-cox로 변환한 값
- 설명변수: 상관관계 분석 결과 종속변수와 영향이 적은 변수를 제외한 15개의 변수
 - 통행거리, 출발지 인구수, 출발지 취업자수, 출발지 생산인구수, 출발지 고령인구수, 출발지 기지국수, 출발지 생산인구 비율, 출발지 고령인구 비율, 도착지 인구수, 도착지 종사자수, 도착지 생산인구수, 도착지 고령인구수, 도착지 기지국수, 도착지 생산인구 비율, 도착지 고령인구 비율

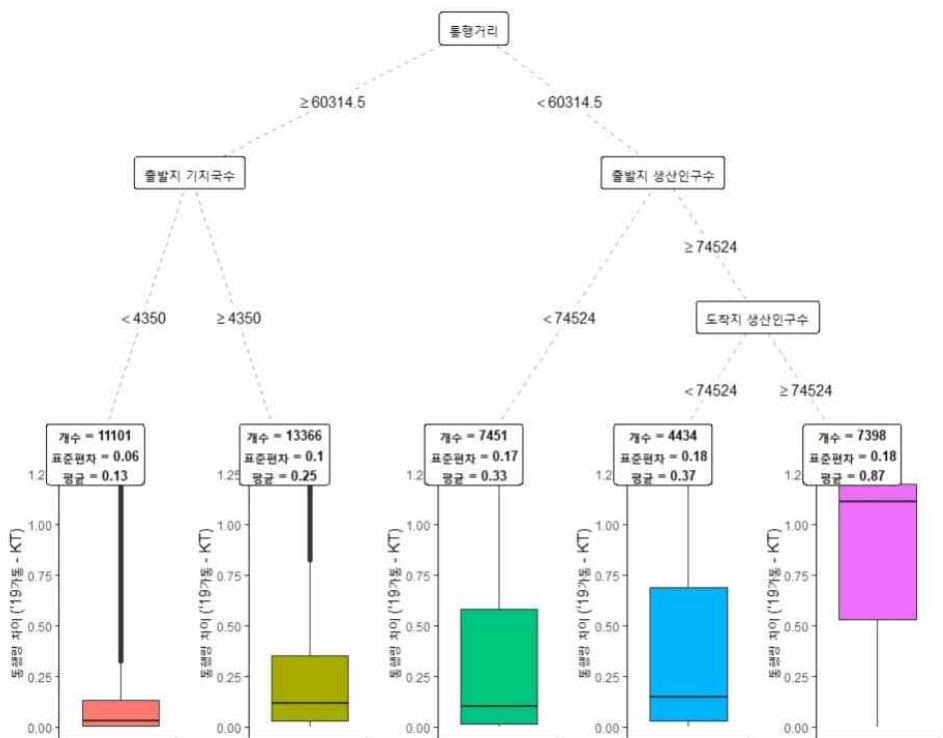


〈그림 3-87〉 (통근 통행) 설명변수 선정 기준

○ 분석결과

- 다음 5가지의 통행패턴에서 통근 통행량에 대한 차이가 발생하는 것으로 분석됨

- ① 통행거리가 60,315km 이상이고, 출발지 기지국수가 4,350개 미만인 경우
- ② 통행거리가 60,315km 이상이고, 출발지 기지국수가 4,350개 이상인 경우
- ③ 통행거리가 60,315km 미만이고, 출발지 생산인구수가 74,524명 미만인 경우
- ④ 통행거리가 60,315km 미만이고, 출발지 생산인구수가 74,524명 이상이며, 도착지 생산인구수가 74,524명 미만인 경우
- ⑤ 통행거리가 60,315km 미만이고, 출발지 생산인구수가 74,524명 이상이며, 도착지 생산인구수가 74,524명 이상인 경우



<그림 3-88> (통근 통행) 의사결정나무구조

- 모델에 대한 성능은 평균제곱오차(MSE) 0.12, 절대평균오차(MAE) 0.27로 나타남

○ 통근 통행에 대한 유형별 보정가중치 산출 결과는 <표 3-31>과 같음

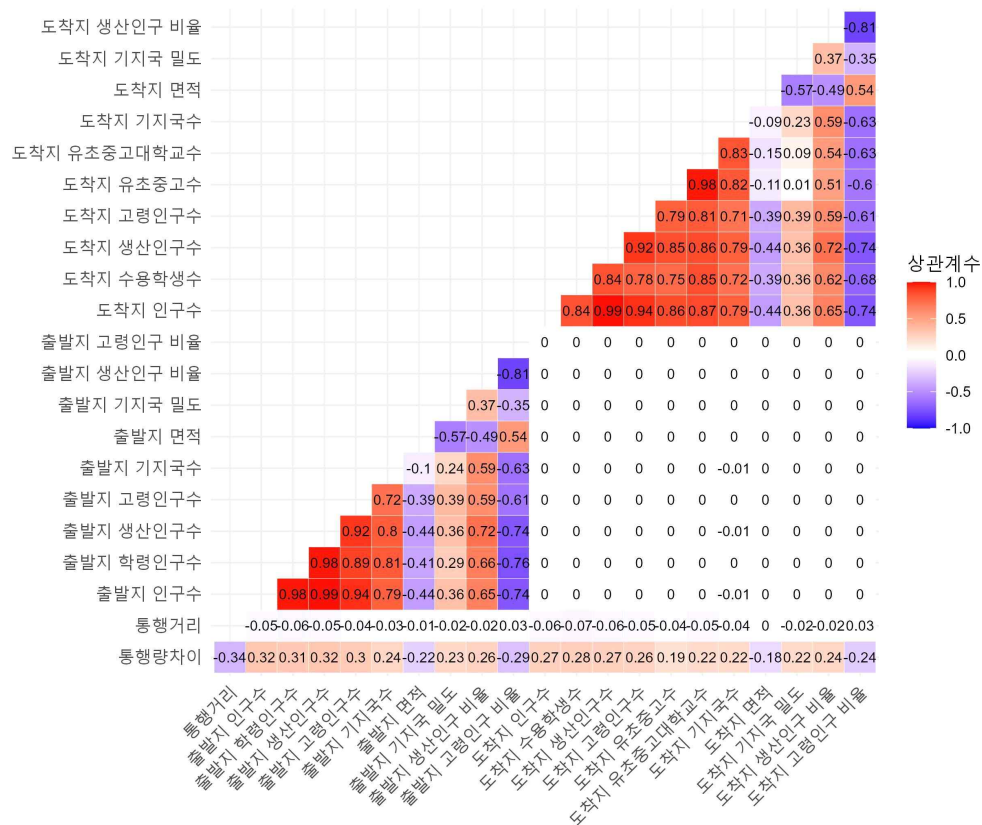
<표 3-31> 통근 통행 유형별 보정가중치

통행 유형	보정유형								보정가중치	
	통행 거리	인구수		기지국수 (밀도)		생산인구수 (비율)		수용 학생수	평균	표준 편차
		출발지	도착지	출발지	도착지	출발지	도착지	도착지		
①	≥ 60,315	-	-	<4,350	-	-	-	-	0.13	0.06
②	≥ 60,315	-	-	≥ 4,350	-	-	-	-	0.25	0.10
③	<60,315	-	-	-	-	<74,524	-	-	0.33	0.17
④	<60,315	-	-	-	-	≥ 74,524	<74,524	-	0.37	0.18
⑤	<60,315	-	-	-	-	≥ 74,524	≥ 74,524	-	0.87	0.18

다. 통학 통행

○ 분석조건

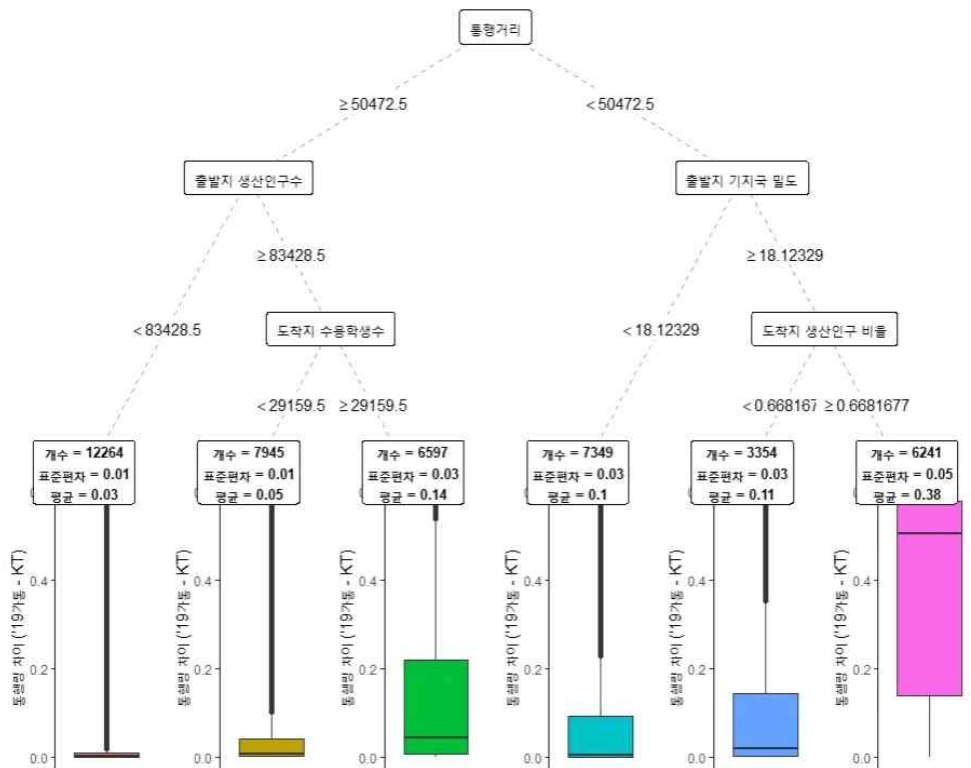
- 종속변수: 2019년 가구통행실태조사의 통학 통행량과 KT 모바일 통신데이터의 통학 통행량의 절대오차를 Box-cox로 변환한 값
- 설명변수: 상관관계 분석 결과 종속변수와 영향이 적은 변수를 제외한 19개의 변수
 - 통행거리, 출발지 인구수, 출발지 학령인구수, 출발지 생산인구수, 출발지 고령인구수, 출발지 기지국수, 출발지 면적, 출발지 기지국 밀도, 출발지 생산인구 비율, 출발지 고령인구 비율, 도착지 인구수, 도착지 수용학생수, 도착지 생산인구수, 도착지 고령인구수, 도착지 유·초·중·고·대학교수, 도착지 기지국수, 도착지 기지국 밀도, 도착지 생산인구 비율, 도착지 고령인구 비율



〈그림 3-89〉 (통학 통행) 설명변수 선정 기준

○ 분석결과

- 다음 6가지의 통행패턴에서 통학 통행량에 대한 차이가 발생하는 것으로 분석됨
 - ① 통행거리가 50,473km 이상이고, 출발지 생산인구수가 83,429명 미만인 경우
 - ② 통행거리가 50,473km 이상이고, 출발지 생산인구수가 83,429명 이상이며, 도착지 수용학생수가 29,160명 미만인 경우
 - ③ 통행거리가 50,473km 이상이고, 출발지 생산인구수가 83,429명 이상이며, 도착지 수용학생수가 29,160명 이상인 경우
 - ④ 통행거리가 50,473km 미만이고, 출발지 기지국 밀도가 18개/㎢ 미만인 경우
 - ⑤ 통행거리가 50,473km 미만이고, 출발지 기지국 밀도가 18개/㎢ 이상이며, 도착지 생산인구 비율이 0.67% 미만인 경우
 - ⑥ 통행거리가 50,473km 미만이고, 출발지 기지국 밀도가 18개/㎢ 이상이며, 도착지 생산인구 비율이 0.67% 이상인 경우



〈그림 3-90〉 (통학 통행) 의사결정나무구조

- 모델에 대한 성능은 평균제곱오차(MSE) 0.03, 절대평균오차(MAE) 0.11로 나타남

○ 통학 통행에 대한 유형별 보정가중치 산출 결과는 <표 3-32>와 같음

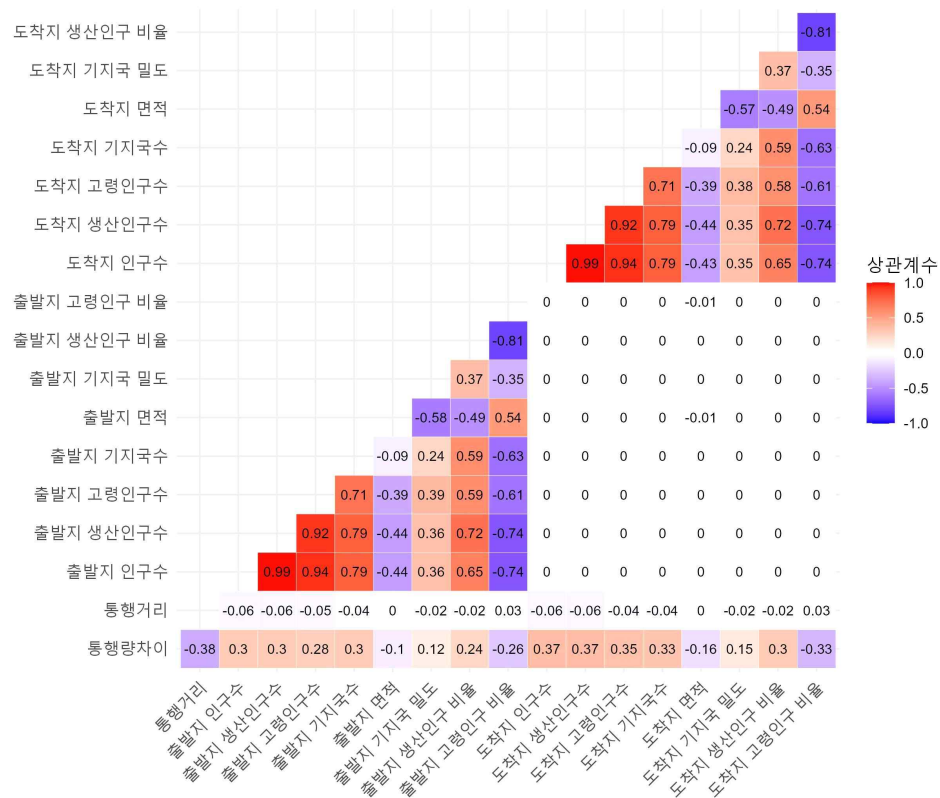
<표 3-32> 통학 통행 유형별 보정가중치

통행 유형	보정유형								보정가중치	
	통행 거리	인구수		기지국수 (밀도)		생산인구수 (비율)		수용 학생수	평균	표준 편차
		출발지	도착지	출발지	도착지	출발지	도착지	도착지		
①	≥ 50,473	-	-	-	-	<83,429	-	-	0.03	0.01
②	≥ 50,473	-	-	-	-	≥ 83,429	-	<29,160	0.05	0.01
③	≥ 50,473	-	-	-	-	≥ 83,429	-	≥ 29,160	0.14	0.03
④	<50,473	-	-	(<18)	-	-	-	-	0.10	0.03
⑤	<50,473	-	-	(≥18)	-	-	(<0.67)	-	0.11	0.03
⑥	<50,473	-	-	(≥18)	-	-	(≥0.67)	-	0.38	0.05

라. 귀가 통행

○ 분석조건

- 종속변수: 2019년 가구통행실태조사의 귀가 통행량과 KT 모바일 통신데이터의 귀가 통행량의 절대오차를 Box-cox로 변환한 값
- 설명변수: 상관관계 분석 결과 종속변수와 영향이 적은 변수를 제외한 13개의 변수
 - 통행거리, 출발지 인구수, 출발지 생산인구수, 출발지 고령인구수, 출발지 기지국수, 출발지 생산인구 비율, 출발지 고령인구 비율, 도착지 인구수, 도착지 생산인구수, 도착지 고령인구수, 도착지 기지국수, 도착지 생산인구 비율, 도착지 고령인구 비율

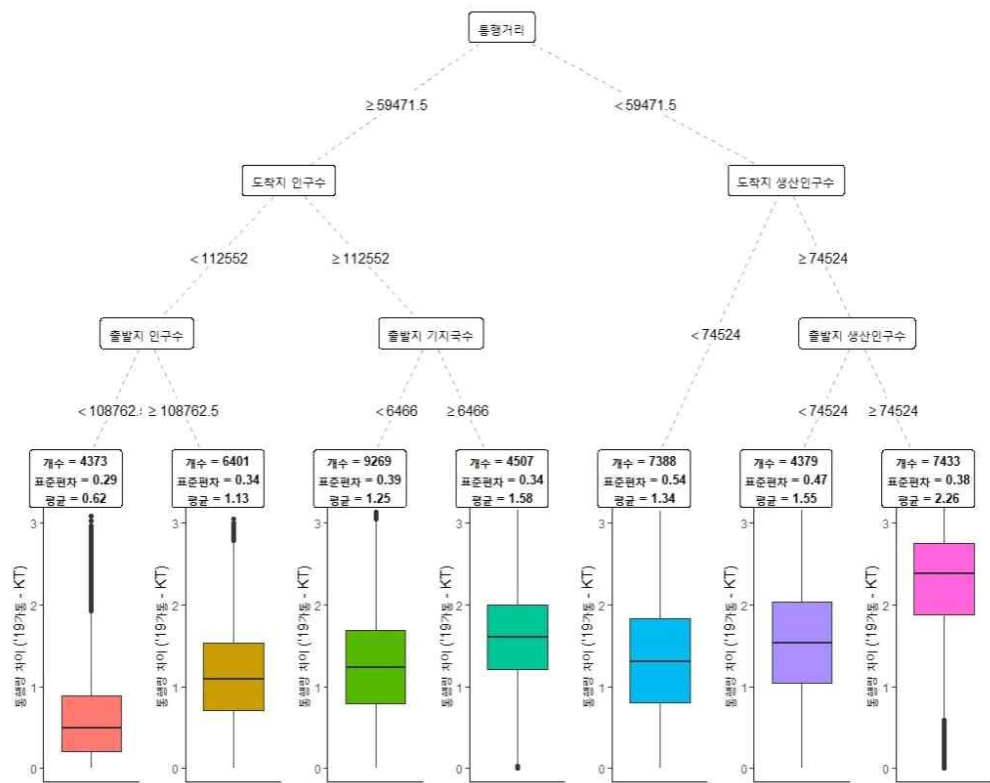


〈그림 3-91〉 (귀가 통행) 설명변수 선정 기준

○ 분석결과

- 다음 7가지의 통행패턴에서 귀가 통행량에 대한 차이가 발생하는 것으로 분석됨
 - ① 통행거리가 59,472km 이상이고, 도착지 인구수가 112,552명 미만이며, 출발지 인구가 108,763명 미만인 경우

- ② 통행거리가 59,472km 이상이고, 도착지 인구수가 112,552명 미만이며, 출발지 인구가 108,763명 이상인 경우
- ③ 통행거리가 59,472km 이상이고, 도착지 인구수가 112,552명 이상이며, 출발지 기지국 수가 6,466개 미만인 경우
- ④ 통행거리가 59,472km 이상이고, 도착지 인구수가 112,552명이상이며, 출발지 기지국 수가 6,466개 이상인 경우
- ⑤ 통행거리가 59,472km 미만이고, 도착지 생산인구수가 74,524명 미만인 경우
- ⑥ 통행거리가 59,472km 미만이고, 도착지 생산인구수가 74,524명 이상이며, 출발지 생산인구수가 74,524명 미만인 경우
- ⑦ 통행거리가 59,472km 미만이고, 도착지 생산인구수가 74,524명 이상이며, 출발지 생산인구수가 74,524명 이상인 경우



<그림 3-92> (귀가 통행) 의사결정나무구조

- 모델에 대한 성능은 평균제곱오차(MSE) 0.40, 절대평균오차(MAE) 0.51로 나타남

○ 귀가 통행에 대한 유형별 보정가중치 산출 결과는 <표 3-33>과 같음

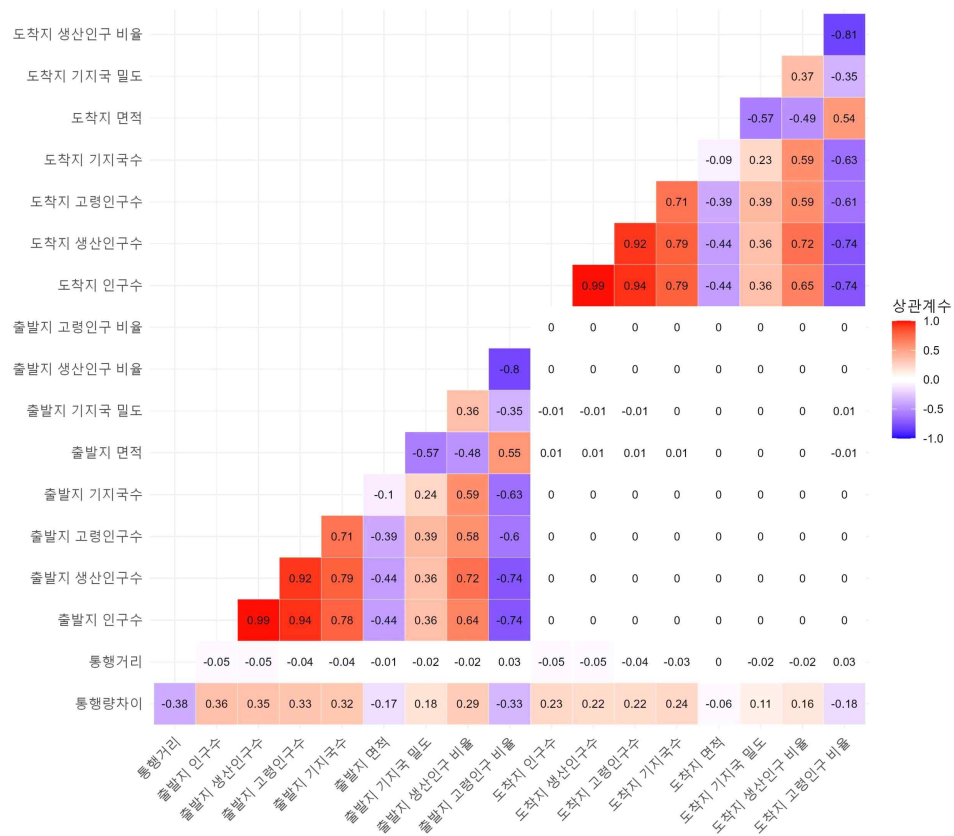
<표 3-33> 귀가 통행 유형별 보정가중치

통행 유형	보정유형								보정가중치	
	통행 거리	인구수		기지국수 (밀도)		생산인구수 (비율)		수용 학생수	평균	표준 편차
		출발지	도착지	출발지	도착지	출발지	도착지	도착지		
①	≥ 59,472	<108,763	<112,552	-	-	-	-	-	0.62	0.29
②	≥ 59,472	≥ 108,763	<112,552	-	-	-	-	-	1.13	0.34
③	≥ 59,472	-	≥ 112,552	<6,466	-	-	-	-	1.25	0.39
④	≥ 59,472	-	≥ 112,552	≥ 6,466	-	-	-	-	1.58	0.34
⑤	<59,472	-	-	-	-	-	<74,524	-	1.34	0.54
⑥	<59,472	-	-	-	-	<74,524	≥ 74,524	-	1.55	0.47
⑦	<59,472	-	-	-	-	≥ 74,524	≥ 74,524	-	2.26	0.38

마. 기타 통행

○ 분석조건

- 종속변수: 2019년 가구통행실태조사의 기타 통행량과 KT 모바일 통신데이터의 기타 통행량의 절대오차를 Box-cox로 변환한 값
- 설명변수: 상관관계 분석 결과 종속변수와 영향이 적은 변수를 제외한 11개의 변수
 - 통행거리, 출발지 인구수, 출발지 생산인구수, 출발지 고령인구수, 출발지 기지국수, 출발지 생산인구 비율, 출발지 고령인구 비율, 도착지 인구수, 도착지 생산인구수, 도착지 고령인구수, 도착지 기지국수

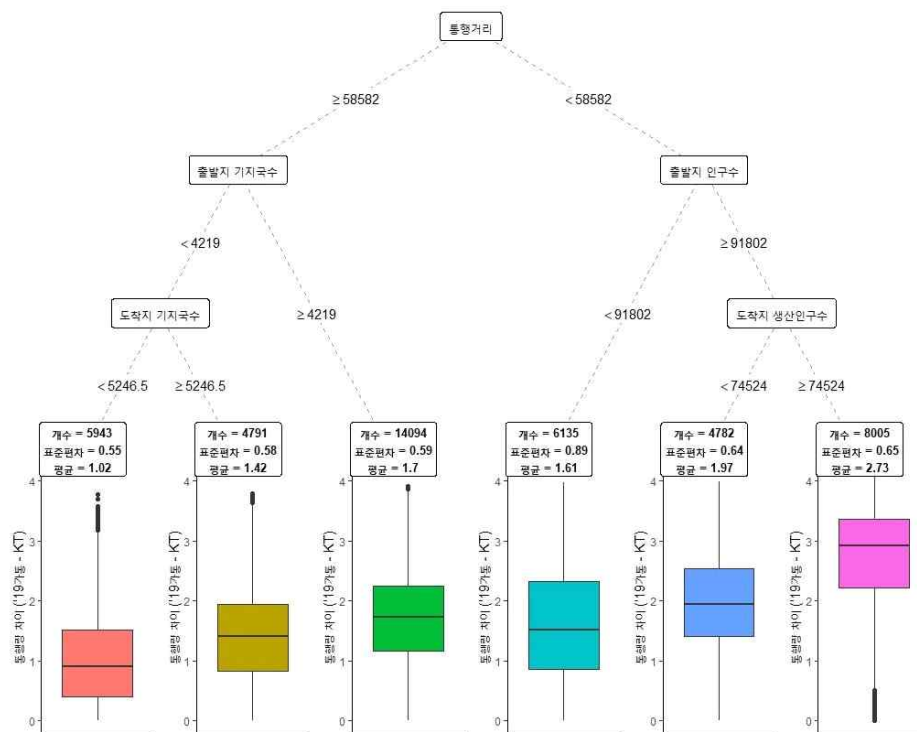


〈그림 3-93〉 (기타 통행) 설명변수 선정 기준

○ 분석결과

- 다음 6가지의 통행패턴에서 기타 통행량에 대한 차이가 발생하는 것으로 분석됨

- ① 통행거리가 58,582km 이상이고, 출발지 기지국수가 4,219개 미만이며, 도착지 기지국수가 5,247개 미만인 경우
- ② 통행거리가 58,582km 이상이고, 출발지 기지국수가 4,219개 미만이며, 도착지 기지국수가 5,247개 이상인 경우
- ③ 통행거리가 58,582km 이상이고, 출발지 기지국수가 4,219개 이상인 경우
- ④ 통행거리가 58,582km 미만이고, 출발지 인구수가 91,802명 미만인 경우
- ⑤ 통행거리가 58,582km 미만이고, 출발지 인구수가 91,802명 이상이며, 도착지 생산인구수가 74,524명 미만인 경우
- ⑥ 통행거리가 58,582km 미만이고, 출발지 인구수가 91,802명 이상이며, 도착지 생산인구수가 74,524명 이상인 경우



〈그림 3-94〉 (기타 통행) 의사결정나무구조

- 모델에 대한 성능은 평균제곱오차(MSE) 0.65, 절대평균오차(MAE) 0.66로 나타남

○ 기타 통행에 대한 유형별 보정가중치 산출 결과는 <표 3-34>과 같음

<표 3-34> 기타 통행 유형별 보정가중치

통행 유형	보정유형								보정가중치	
	통행 거리	인구수		기지국수 (밀도)		생산인구수 (비율)		수용 학생수	평균	표준 편차
		출발지	도착지	출발지	도착지	출발지	도착지	도착지		
①	≥ 58,582	-	-	<4,219	<5,247	-	-	-	1.02	0.55
②	≥ 58,582	-	-	<4,219	≥ 5,247	-	-	-	1.42	0.58
③	≥ 58,582	-	-	≥ 4,219	-	-	-	-	1.70	0.59
④	<58,582	<91,802	-	-	-	-	-	-	1.61	0.89
⑤	<58,582	≥ 91,802	-	-	-	-	<74,524	-	1.97	0.64
⑥	<58,582	≥ 91,802	-	-	-	-	≥ 74,524	-	2.73	0.65

4. 통행량 보정 방법

○ 전체 통행과 통행목적별 보정가중치를 실제 데이터에 적용하여 통행량을 보정하는 과정을 정리하면 다음과 같음

- 1단계: 보정하고자 하는 데이터를 보정대상 유형 분류가 가능한 형태로 준비함
 - 수집한 데이터를 전체, 통근, 통학, 귀가, 기타 5가지의 통행목적으로 분류하고 통행량을 집계한 후, <표 3-35>와 같이 통행 유형 분류에 필요한 정보를 연계함
 - 통행량은 1일 단위 시군구 단위로 집계하며, 집계 후 box-cox 변환하도록 함
 - 통행 유형 분류에 필요한 정보는 시군구 단위로 구축하고, 출발지, 도착지 코드를 기준으로 연계함

〈표 3-35〉 통행유형 분류를 위해 수집해야 하는 시군구 정보

구분		구축 목적
통행거리		전체, 통근, 통학, 귀가, 기타 통행 유형 분류시 기준으로 활용
인구수	출발지	전체, 귀가, 기타 통행 유형 분류시 기준으로 활용
	도착지	귀가 통행 유형 분류시 기준으로 활용
기지국수	출발지	전체, 통근, 통학, 귀가, 기타 통행 유형 분류시 기준으로 활용
	도착지	전체, 기타 통행 유형 분류시 기준으로 활용
면적	출발지	출발지 기지국 밀도 산출시 활용
	도착지	도착지 기지국 밀도 산출시 활용
생산인구수	출발지	통근, 통학, 귀가 통행 유형 분류시 기준으로 활용
	도착지	전체, 통근, 통학, 귀가, 기타 통행 유형 분류시 기준으로 활용
수용학생수	도착지	통학 통행 유형 분류시 기준으로 활용

- 2단계: 통행목적별 모델링의 변수, 절단값 기준을 이용하여 통행유형을 분류함
(<표 3-36> 참조)

〈표 3-36〉 통행유형별 변수 및 절단값

구분	통행 거리	인구수		기지국수 (밀도)		생산인구수 (비율)		수용 학생수	통행 유형
		출발지	도착지	출발지	도착지	출발지	도착지	도착지	
전체 통행	≥ 60,736	-	-	<4,350	<5,228	-	-	-	전체-①
	≥ 60,736	-	-	<4,350	≥ 5,228	-	-	-	전체-②
	≥ 60,736	-	-	≥ 4,350	<6,466	-	-	-	전체-③

구분	통행 거리	인구수		기지국수 (밀도)		생산인구수 (비율)		수용 학생수	통행 유형
		출발지	도착지	출발지	도착지	출발지	도착지	도착지	
	≥ 60,736	-	-	≥ 4,350	≥ 6,466	-	-	-	전체-④
	< 60,736	< 91,802	-	-	-	-	-	-	전체-⑤
	< 60,736	≥ 91,802	-	-	-	-	< 74,524	-	전체-⑥
	< 60,736	≥ 91,802	-	-	-	-	≥ 74,524	-	전체-⑦
통근 통행	≥ 60,315	-	-	< 4,350	-	-	-	-	통근-①
	≥ 60,315	-	-	≥ 4,350	-	-	-	-	통근-②
	< 60,315	-	-	-	-	< 74,524	-	-	통근-③
	< 60,315	-	-	-	-	≥ 74,524	< 74,524	-	통근-④
	< 60,315	-	-	-	-	≥ 74,524	≥ 74,524	-	통근-⑤
통학 통행	≥ 50,473	-	-	-	-	< 83,429	-	-	통학-①
	≥ 50,473	-	-	-	-	≥ 83,429	-	< 29,160	통학-②
	≥ 50,473	-	-	-	-	≥ 83,429	-	≥ 29,160	통학-③
	< 50,473	-	-	(< 18)	-	-	-	-	통학-④
	< 50,473	-	-	(≥ 18)	-	-	(< 0.67)	-	통학-⑤
	< 50,473	-	-	(≥ 18)	-	-	(≥ 0.67)	-	통학-⑥
귀가 통행	≥ 59,472	< 108,763	< 112,552	-	-	-	-	-	귀가-①
	≥ 59,472	≥ 108,763	< 112,552	-	-	-	-	-	귀가-②
	≥ 59,472	-	≥ 112,552	< 6,466	-	-	-	-	귀가-③
	≥ 59,472	-	≥ 112,552	≥ 6,466	-	-	-	-	귀가-④
	< 59,472	-	-	-	-	-	< 74,524	-	귀가-⑤
	< 59,472	-	-	-	-	< 74,524	≥ 74,524	-	귀가-⑥
	< 59,472	-	-	-	-	≥ 74,524	≥ 74,524	-	귀가-⑦
기타 통행	≥ 58,582	-	-	< 4,219	< 5,247	-	-	-	기타-①
	≥ 58,582	-	-	< 4,219	≥ 5,247	-	-	-	기타-②
	≥ 58,582	-	-	≥ 4,219	-	-	-	-	기타-③
	< 58,582	< 91,802	-	-	-	-	-	-	기타-④
	< 58,582	≥ 91,802	-	-	-	-	< 74,524	-	기타-⑤
	< 58,582	≥ 91,802	-	-	-	-	≥ 74,524	-	기타-⑥

- 3단계: 평균표준편차일치법을 활용하여 보정 통행량을 산출한 후, 보정할 대상에 보정 통행량을 더하거나 빼서 통행량을 보정함(<표 3-38> 참조)

- 보정 통행량은 모델링을 통해 산출된 통행유형별 평균, 표준편차(<표 3-37> 참조)와 보정 대상(모바일 데이터)의 평균, 표준편차를 아래 산식에 적용하여 산출함

$$\text{보정 통행량} = \text{보정가중치의 평균 값} + \left\{ \frac{(\text{보정 전 통행량} - A)}{B} \right\} \times \text{보정가중치의 표준편차 값}$$

A: 해당 통행 유형의 평균 통행량

B: 해당 통행 유형의 통행량에 대한 표준편차

〈표 3-37〉 통행유형별 보정가중치 (평균, 표준편차)

통행 유형		평균	표준편차
전체	①	1.27	0.65
	②	1.74	0.64
	③	1.84	0.64
	④	2.25	0.57
	⑤	1.92	0.98
	⑥	2.25	0.71
	⑦	3.03	0.73
통근	①	0.13	0.06
	②	0.25	0.10
	③	0.33	0.17
	④	0.37	0.18
	⑤	0.87	0.18
통학	①	0.03	0.01
	②	0.05	0.01
	③	0.14	0.03
	④	0.10	0.03
	⑤	0.11	0.03
	⑥	0.38	0.05
귀가	①	0.62	0.29
	②	1.13	0.34
	③	1.25	0.39
	④	1.58	0.34
	⑤	1.34	0.54
	⑥	1.55	0.47
	⑦	2.26	0.38
기타	①	1.02	0.55
	②	1.42	0.58
	③	1.70	0.59
	④	1.61	0.89
	⑤	1.97	0.64
	⑥	2.73	0.65

〈표 3-38〉 통행량 보정 예시

구분		통행구간 특성			통행 유형 분류	보정 전 통행량	→	보정 후 통행량*	보정 통행량
출발지	도착지	통행거리	출발지 기지국 수	도착지 기지국 수					
A	B	65,200	4,000	5,000	전체-①	10		11.59	1.59
C	D	68,000	4,900	4,300	전체-①	4		4.72	0.72
A	C	60,500	4,000	4,900	전체-①	8		9.3	1.30
E	F	70,000	2,000	1,000	전체-①	14		16.17	2.17
A	F	73,200	4,000	1,000	전체-①	3		3.58	0.58
평균 (A)						7.8	→		1.27
표준편차 (B)						4.5	→		0.65
(보정가중치 - 평균) 전체-① 유형의 평균						1.27			
(보정가중치 - 표준편차) 전체-① 유형의 표준편차						0.65			

*: 보정 전 통행량에서 보정 통행량을 더한 경우를 예시로 제시함

5. 결론 및 향후과제

- 본 과업에서는 조사 기반 자료와 빅데이터 기반 자료의 기종점 통행량 차이를 확인하고 보정 대상과 그 정도(가중치)를 파악함
 - 조사 기반 및 빅데이터 기반 기종점 통행량을 검증하기 위해서는 참이 되는 기준값 설정이 필요함
 - 그러나, 검증을 위한 참값(ground truth)을 설정하는 것은 여전히 논쟁의 여지가 있기 때문에, 본 연구에서는 조사 기반 및 빅데이터 기반 기종점 통행량을 상호 비교하여 그 차이의 패턴을 확인하고 보정 대상을 선정하여 보정의 정도(가중치)를 제시함
- 향후, 이러한 보정 대상과 가중치를 이용하여 각 데이터 간 산술평균 또는 가중평균하여 구축된 기종점 통행량의 신뢰성을 개선할 수 있을 것으로 판단함



제4장 모바일 통신데이터 기반 교통 지표 산출 및 제공

제1절 모바일 통신데이터 기반 교통지표 구축

제2절 온라인 서비스를 통한 교통지표 제공

제4장 모바일 통신데이터 기반 교통지표 산출 및 제공

제1절 모바일 통신데이터 기반 교통지표 구축

1. 개요

- 모바일 통신데이터는 기존 조사 기반 자료에 비해 높은 해상도의 시공간 정보(1년 365일 1시간 단위 500미터 격자단위)로 상세한 통행 및 활동 정보를 담고 있어, 다양한 교통지표 추출이 가능함
 - 기존 통행실태조사에서 최소 읍면동 단위로 제공되었던 통행자의 출발, 도착 위치 정보도 모바일 통신데이터에서는 50m 단위의 격자 단위로 제공하여 통행량(발생량, 도착량, 분포량)을 더욱 섬세한 단위로 집계할 수 있음
 - 또한 모바일 통신데이터에는 통행 정보가 일자별로 구축되어 있어, 일평균 통행량뿐만 아니라 특정 일자에 대한 통행량을 확인할 수 있음
 - 또한 기존에는 알 수 없었던 시간, 통행자의 인적 정보(성, 연령)도 포함되어 있어 특정 시간대나 특정 통행자에 대한 통행의 특성도 확인할 수 있음
 - <표 4-1>은 기존 사업과 본 사업에서 모바일 통신데이터에 포함되어 있는 출발 정보(지역, 시간대, 체류유형), 도착 정보(지역, 시간대, 체류유형), 통행자 정보(성, 연령), 구간 정보(통행시간)를 기준으로 개발한 교통지표임

〈표 4-1〉 모바일 통신데이터 기반 교통지표

교통지표		설명	지표 추출 기준 (활용된 통행정보)
대분류	소분류		
통행량	발생량	출발지를 기준으로 집계한 통행량	출발지
	도착량	도착지를 기준으로 집계한 통행량	도착지
	분포량 (기종점통행량)	통행구간별로 집계한 통행량	출발지, 도착지
	기간별 통행량	연휴 기간 통행량	출발일자, 도착일자
		계절별 통행량	출발일자, 도착일자
	통행목적별 통행량	출근, 퇴근, 등교, 하교, 귀가, 여가, 관광 통행에 대해 집계한 통행량	출발지 체류유형, 도착지 체류유형
	통행자 특성별 통행량	연령대가 70대 이상의 고령인구에 대해 집계한 통행량	통행자의 연령
통행시간	평균통행시간	선택한 통행 대상의 통행시간을 평균한 값	통행시간
통행거리	평균통행거리	선택한 통행 대상의 통행거리를 평균한 값	출발지, 도착지
통행비율	내부 통행 비율	동일한 행정구역 내에서 이동하는 통행자가 전체 통행에서 차지하는 비율	출발지, 도착지
	외부 통행 비율	외부로 통행하는 통행자가 전체 통행에서 차지하는 비율	출발지, 도착지
	고령자 통행 비율	전체 통행자 중에서 고령 통행자가 차지하는 비율	통행자의 연령
	심야시간 출·퇴근 통행 비율	심야시간대 퇴근 통행량 대비 출근 통행량의 비율	출발시간, 도착시간, 출발지 체류유형, 도착지 체류유형
사회·경제 및 모빌리티 활동 지표	출·퇴근 증감량	전월 대비 기준월의 통근 목적 통행량의 증가량 또는 감소량	출발지 체류유형, 도착지 체류유형
	도심공동화 심각도	유동인구 대비 상주인구 비율	출발시간, 도착시간, 출발지 체류유형, 도착지 체류유형
	경제활동의존도	특별시·광역시 주변도시에서 발생한 출근 통행량 중 특별시·광역시로 출근한 통행량의 비율	출발지, 도착지, 출발지 체류유형, 도착지 체류유형
	경제활동 자체수용도	특별시·광역시에서 발생한 출근 통행량 중 지역 내 통행량의 비율	출발지, 도착지, 출발지 체류유형, 도착지 체류유형
	출·퇴근 통행연결성	출·퇴근 통행에 대한 기종점간의 모듈성	출발지, 도착지, 출발지 체류유형, 도착지 체류유형
	직주균형지표	거주지와 근무지간의 실제통근거리 대비 직주불일치로 발생하는 초과통근거리의 비율	출발지, 도착지, 출발지 체류유형, 도착지 체류유형

- 모바일 통신데이터의 규모와 양은 조사 데이터에 비해 매우 방대하며, 그에 따른 적절한 형태로 가공되어야 함
 - 모바일 통신데이터는 1일 약 3천2백 건의 데이터로 구성되어 있어, 연평균 지표를 표출할 경우 약 1억 2천 건의 데이터를 조회해야 함
- 본 절에서는 이러한 대규모 빅데이터를 처리 및 가공하고, 정책 의사결정에 활용할 수 있는 다양한 사회·경제 및 모빌리티 지표를 만드는 과정을 설명함
 - 교통지표 산출 시간을 최소화할 수 있도록 기존 사업과 본 사업에서 개발한 모바일 통신데이터 기반의 교통지표 19종을 DB화함
 - 모바일 통신데이터 가공 과정과 최종 교통지표 구축과정을 설명함
 - 교통지표 DB 구축에는 2022년 사업에서 구축한 KT통신데이터 기반의 통행 DB를 활용함
 - 시간적 범위: 2022년 1월 1일~ 2022년 12월 31일(1년)
 - 공간적 범위: 전국

2. 교통지표 구축을 위한 데이터 전처리

가. 오류 데이터 제거

- 일부 정보가 누락되었거나, 통행으로 간주할 수 없는 데이터를 제거함
 - ① 출발지 또는 도착지 정보가 누락된 경우: 격자, 시도, 시군구, 읍면동 정보가 없는 경우
 - ② 통근, 통학 목적 적합하지 않은 연령대 정보가 포함된 경우: 통근 통행이면서 연령이 0, 10으로 표기된 경우, 통학 통행이면서 연령이 20대 이상인 경우
 - ③ 통행시간이 지나치게 짧거나 지나치게 긴 경우: 통근 통행이면서 통행시간이 1분 미만 4시간 이상인 경우

출발							연령	성별	시간(분)	비고	
출발일자	출발요일	출발시간	출발격자	출발시도	출발시군구	출발읍면동					출발체류지
20220101	7	7	다라09504900	36	36420	3642012	H	M	60	31.8857839	도착 정보 없음
20220101	7	10					H	F	50	7.308266651	출발 정보 없음
20220101	7	10					X	F	50	20.07457113	출발 정보 없음
20220101	7	20	마라51508750	21	21090	2109053	C	F	10	28.5137668	연령 오류(회사, 10대)
20220101	7	19	마라54009650	21	21310	2131011	S	F	50	21.12133597	연령 오류(학교, 50대)
20220101	7	6					H	M	60	65.0875429	출도착 정보 없음
20220101	7	12					H	F	0	25.81094333	출도착 정보 없음
20220101	7	7	다사22506200	31	31230	3123013	H	F	20	250	통행목적 출근 시간 오류
20220101	7	7	라라51008850	38	38030	3803068	H	M	30	1	통행목적 출근 시간 오류

〈그림 4-1〉 모바일 데이터 오류 유형 예시

나. 데이터 표준화

- 행정구역 정보를 기준시점의 정보로 통일함
 - KT 모바일 통신 빅데이터는 데이터 구축 시기별 다른 기준시점의 행정구역 정보를 사용하여 동일한 격자 ID에 서로 다른 행정구역 명칭이 매칭된 경우가 존재하므로, 2022년 6월 기준의 행정구역 정보로 전체 업데이트함
 - 행정구역 정보는 분기별로 업데이트되어 일부 지역에 대한 명칭이 변경되는데, 2022년 DB는 기존 구축된 2019~2020년 DB와 동일한 기준으로 구축되면서, 업데이트된 행정 구역 정보가 반영되지 않은 데이터가 존재하게 됨

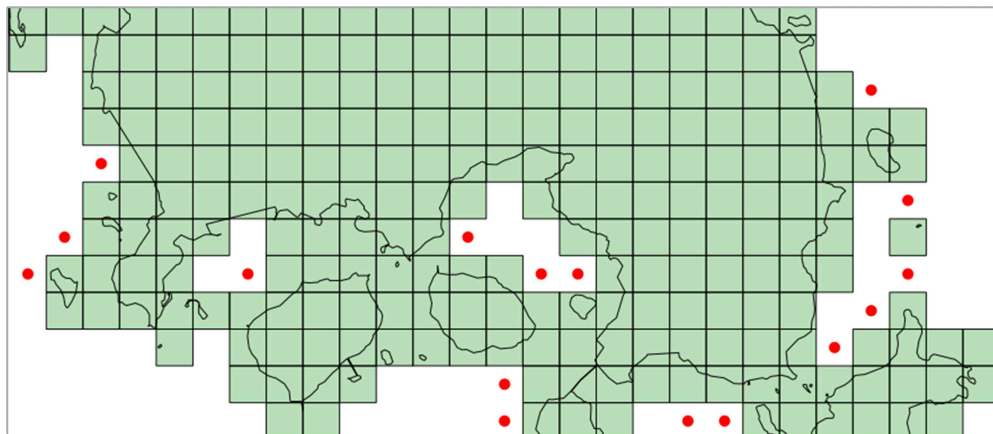
- 격자에 행정구역을 매칭시킬 때, 한 개의 격자에 복수개의 행정동이 매칭되는 경우 높은 면적 비율을 갖는 행정구역을 매칭시켜 대표성을 갖도록 함
- 기 입력된 행정동 정보는 행정동의 면적, 중심점 포함 여부와 관계없이 행정동 코드 숫자상 우선적으로 매칭되는 행정동만 남기는 방식으로 매칭된 것이라 대표성이 미흡하여 매칭 방식을 변경함



격자ID	연계 행정구역	면적비율	최종 매칭 행정구역
① 다사58b44a	역삼1동(11230640)	100%	역삼1동(11230640)
	역삼2동(11230650)	51%	
② 다사59a44a	역삼1동(11230640)	49%	역삼2동(11230650)
	역삼1동(11230640)	89%	
③ 다사58b43b	도곡1동(11230660)	6%	역삼1동(11230640)
	서초2동(11220520)	5%	
	역삼1동(11230640)	46%	
④ 다사59a43b	도곡1동(11230660)	44%	역삼1동(11230640)
	역삼2동(11230650)	10%	
	역삼2동(11230650)	10%	

〈그림 4-2〉 격자-행정구역 매칭 방법 (면적 비율 적용 예시)

- 해상에 위치하는 격자(전체의 약 0.004%)는 제외하고 육지에 포함되는 격자만을 대상으로 매칭함



〈그림 4-3〉 매칭 제외 대상 (해상에 위치하는 격자 예시)

- 매칭 이후 모든 격자ID, 행정구역 ID가 포함되었는지, ID 자릿수, 코드 등 오류는 없는지 검수하였고, 검수 결과 면적 비율 기준으로 매칭하는 과정에서 9개의 읍면동은 어떠한 격자에도 매칭되지 않은 것으로 확인됨

〈표 4-2〉 격자와 매칭되지 않은 읍면동 리스트

No	읍면동 코드	읍면동 명칭	전체 명칭
1	11010690	창신3동	서울특별시 종로구 창신3동
2	21010520	동광동	부산광역시 중구 동광동
3	21010560	부평동	부산광역시 중구 부평동
4	21010570	광복동	부산광역시 중구 광복동
5	21020510	동대신1동	부산광역시 서구 동대신1동
6	22030580	비산6동	대구광역시 서구 비산6동
7	23010570	율목동	인천광역시 중구 율목동
8	31060510	광명1동	경기도 광명시 광명1동
9	31060530	광명3동	경기도 광명시 광명3동

- 또한 행정동 정보를 매칭하기 위해서 국가지점번호 원형으로 표기된 격자 ID를 500m에 맞춰 변형한 국가지점번호로 변환함
- 격자에 행정구역 정보를 매칭시키기 위해서는 격자 맵 레이어와 행정구역 맵 레이어를 공간 조인(Spatial join)하여야 하는데, 국토지리정보원에서 제공 중인 격자 맵의 격자 ID는 국가지점번호를 변형하여 구성된 것으로서 기 입력된 격자ID와 체계가 달라서 격자 맵을 활용하려면 모바일 데이터에 입력된 격자 ID를 국토지리정보원 체계에 맞춰서 변환해야 함
- 모바일 데이터에 입력된 격자 ID의 3번째 4번째 숫자와 7번째 8번째 숫자가 '00'인 경우 'a'로, '50'인 경우 'b'로 변환하여 국토지리정보원 체계로 맞춤

〈표 4-3〉 격자 ID 변환 전, 후

구분	변환 전	변환 후
표기 방식	국가지점번호 원형 표기 방식으로, 한글 2+숫자 8로 구성됨	국가지점번호를 변형한 것으로, 한글 2+숫자2+영문1+숫자2+영문1로 구성
예시	나바98003400	나바98a34a

다. 추가 정보(거리) 생성

- 출발 격자와 도착 격자 간의 직선거리 정보를 생성함
- 전처리 결과 데이터를 <표 4-4>와 같이 표준 DB 형태로 변환 및 저장함

<표 4-4> 모바일 데이터 전처리 표준 DB

순번	필드명	필드설명	비고
1	o_date	출발지 날짜	
2	o_dow	출발지 요일	1~7 (일~토)
3	o_time	출발지 시간	
4	o_grid	출발지 격자	
5	o_sido	출발지 시도	
6	o_sgg	출발지 시군구	
7	o_emd	출발지 읍면동	
8	o_type	출발지 유형	H:집, C: 회사, S:학교, R:정기잠재체류지, X:기타잠재체류지, O: 기타
9	d_date	도착지 날짜	
10	d_time	도착지 시간	
11	d_grid	도착지 격자	
12	d_sido	도착지 시도	
13	d_sgg	도착지 시군구	
14	d_emd	도착지 읍면동	
15	d_type	도착지 유형	H:집, C: 회사, S:학교, R:정기잠재체류지, X:기타잠재체류지, O: 기타
16	gender	성별	M: 남성, F: 여성
17	age	연령대	0,10,20,30... (10코드는 : 10세~19세) 학교 나이코드: 0,10 회사나이코드:20,30,40,50,60(60이상은 모두60으로코드생성) 그외:0~110
18	stay_traffic_volume	총 통행량	
19	traffic_60	잠재체류 60분 이상 제외 통행비율	총 통행량 중의 비율
20	traffic_time_60_mean	통행시간_평균	분 단위
21	traffic_distance_60_mean	통행거리_평균	km 단위
22	traffic_speed_60_mean	통행속도_평균	km/h
23	final_volume	최종 통행량	

3. 교통지표 구축

1) 통행량

가. 발생량·도착량·기종점 통행량(분포량)

- 모바일 통신데이터에서 출발지를 기준으로 집계한 통행량을 발생량, 도착지를 기준으로 집계한 통행량은 도착량, 출발지와 도착지를 하나의 구간으로 보고 구간별로 집계한 통행량은 기종점 통행량으로 볼 수 있음
- 공간 위계별로 발생량, 도착량, 기종점 통행량을 확인할 수 있도록 격자 단위로 구축된 모바일 통신데이터를 읍면동, 시군구, 시도 단위로 집계하여 <표 4-5>, <표 4-6>,<표 4-7>과 같이 각각 구축함

<표 4-5> 읍면동-읍면동 단위 통행량 컬렉션의 구성

컬렉션명	E0_EMD2EMD			
컬렉션 설명	읍면동-읍면동 단위로 그룹화한 통행량 데이터			
데이터 수	약 2억 4천 6백만 건 (246,293,754)		데이터 크기	약 333GB
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { "A1" : 출발읍면동코드, "E1" : 도착읍면동코드, "J" : 성별, "I" : 연령대, "D" : 출발지 체류 유형, "H" : 도착지 체류 유형 }	A1, E1, D, H, J, I : integer 체류특성 코드 ·H:11, N:12, C:13, S:14, R:15, X: 16 성별 코드 ·F:1, M:2
출도착시간대 통행량	CG	Object 배열	CG : [{ "C" : 출발시간대, "G" : 도착시간대, "K" : [{ "SD":출발날짜, "ED":도착날짜, "SV":총 통행량, }] }]	C, G, : integer K : object 배열
출발시군구	A2	Integer	"A2" : 출발시군구코드	8자리로 변경
도착시군구	E2	Integer	"E2" : 도착시군구코드	8자리로 변경
출발시도	A3	Integer	"A3" : 출발시도코드	8자리로 변경
도착시도	E3	Integer	"E3" : 도착시도코드	8자리로 변경
거리	S	Double	"S" : 읍면동간 직선거리	-1: 읍면동 코드 네트워크상 없는 경우

〈표 4-6〉 시군구-시군구 단위 통행량 컬렉션의 구성

컬렉션명	E0_SGG2SGG			
컬렉션 설명	시군구-시군구 단위로 그룹화한 통행량 데이터			
데이터 수	약 8백 19만 건 (8,190,590)		데이터 크기	약 92.6GB
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { "A2" : 출발시군구코드, "E2" : 도착시군구코드, "J" : 성별, "I" : 연령대, "D" : 출발지 체류 유형, "H" : 도착지 체류 유형 }	A2, E2, D, H, J, I : integer 체류특성 코드 ·H:11, N:12, C:13, S:14, R:15, X: 16 성별 코드 ·F:1, M:2
출도착시간대 통행량	CG	Object 배열	CG : [{ "C" : 출발시간대, "G" : 도착시간대, "K" : [{ "SD":출발날짜, "ED":도착날짜, "SV":총 통행량, }] }]	C, G, : integer K : object 배열
출발시도	A3	Integer	"A3" : 출발시도코드	8자리로 변경
도착시도	E3	Integer	"E3" : 도착시도코드	8자리로 변경

〈표 4-7〉 시도-시도 단위 통행량 컬렉션의 구성

컬렉션명	E0_SIDO2SIDO			
컬렉션 설명	시도-시도 단위로 그룹화한 통행량 데이터			
데이터 수	약 5만 4천 건 (54,534)		데이터 크기	약 10.9GB
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { "A3" : 출발시도코드, "E3" : 도착시도코드, "J" : 성별, "I" : 연령대, "D" : 출발지 체류 유형, "H" : 도착지 체류 유형 }	A3, E3, D, H, J, I : integer 체류특성 코드 ·H:11, N:12, C:13, S:14, R:15, X: 16 성별 코드 ·F:1, M:2 시도 코드 ·8자리로 변경
출도착시간대 통행량	CG	Object 배열	CG : [{ "C" : 출발시간대, "G" : 도착시간대, "K" : [{ "SD":출발날짜, "ED":도착날짜, "SV":총 통행량, }] }]	C, G, : integer K : object 배열

나. 연휴기간 통행량

- 모바일 통신데이터의 출발일자와 도착일자 정보를 활용하여 연휴기간에 이루어진 통행을 추출한 후 집계한 통행량을 의미함
- 1년 중 대표적인 연휴기간인 ‘설, 가정의 달, 추석 연휴, 하계 휴가기간’에 대해 통행량을 집계하여 DB로 구축함(<표 4-8> 참조)
 - 연휴기간의 통행량은 집계한 기간으로 나누어 일평균 값으로 환산하여 구축하고, 평상시 통행량과 비교할 수 있도록 1년 평균 통행량 정보와 함께 연휴기간 평균 통행량과 연평균 통행량의 차이를 산출한 필드를 추가함
 - 2022년 DB에서 설, 가정의 달, 추석 연휴, 하계휴가 기간으로 설정한 기간은 다음과 같음
 - 설 기간 설정: 2022.01.29.(토)~2022.02.02.(수)
 - 가정의 달 기간 설정: 2022.05.05.(목)~2022.05.08.(일)
 - 추석 연휴 기간 설정: 2022.09.09.(금)~2022.09.12.(월)
 - 하계휴가 기간 설정: 2022.07.23.(토)~2022.08.15.(월)
 - 단, 집계시 출발 시도가 특별시, 광역시가 아닌 경우, 도착 시군구가 시 단위가 아닌 경우는 제거함

〈표 4-8〉 연휴기간의 통행량 컬렉션의 구성

컬렉션명	L0_HOLLIDAY_PASS			
컬렉션 설명	평상시와 연휴기간의 통행량 차이를 분석한 데이터			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
연휴 구분키	Key	Integer	“Key” : 연휴 구분 코드	출발날짜 기준 1:설연휴, 2:가정의달 3:하계휴가, 4:추석연휴
출발 시도	A3	Integer	“A3” : 출발시도코드	8자리로 변경 / 특별시, 광역시만
도착 시군구	E2	Integer	“E2” : 도착시군구코드	8자리로 변경 / 시단위
1년 평균 통행량	V2	Integer	“V2” : 1년 평균통행량	
연휴기간 평균 통행량	V	Integer	“V” : 연휴기간 평균 통행량	
통행량 차이	Diff	Integer	“Diff” : 통행량차이	연휴기간 평균(V) - 1년 평균(V2)

다. 계절별 통행량

- 계절별 통행량은 모바일 통신데이터의 출발 일자와 도착 일자 정보를 활용하여 통행이 이루어진 계절을 봄, 여름, 가을, 겨울로 구분하고, 각 계절별로 통행량을 집계한 통행량을 의미함
 - 도착일자 기준으로 3~5월은 봄, 6~8월은 여름, 9~11월은 가을 12~2월 겨울로 구분함
- 계절별 통행량은 다음 두 가지의 형태로 구축함
 - ① 일주일 단위로 집계한 통행량(<표 4-9> 참조)
 - 성별, 연령대별 통행 차이가 있을 것으로 예상되어 성별, 연령대별 통행량을 구분하여 집계한 필드를 추가함

<표 4-9> 계절별 통행량 컬렉션의 구성

컬렉션명	L0_SEASON_HOTPLACE_ALL			
컬렉션 설명	주차별로 남녀, 연령별 통행량을 분석한 데이터			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
계절 코드	Season	Integer	"Key" : 계절 구분 코드	도착일자 기준 1:봄,2:여름 3:가을,4:겨울 -1:그외
주차 코드	Week	Integer	"Week" : 주차 구분 코드	도착일자 기준 1~52:주차
도착 시도	E3	Integer	"E3" : 도착시도코드	8자리로 변경
도착 시군구	E2	Integer	"E2" : 도착시군구코드	8자리로 변경
연령대	I	Integer	"I" : 연령대코드	10,20,30... (10코드는 : 10세~19세)
남자 통행량	m	Integer	"m" : 남자통행량	
여자 통행량	f	Integer	"f" : 여자통행량	
전체 통행량	t	Integer	"t" : 총 통행량	남자 + 여자 통행량

② 계절 단위로 집계한 통행량(<표 4-10> 참조)

- 여가, 관광 통행의 특성이 부각될 수 있도록 출·퇴근, 등·하교 통행을 제외한 나머지 기타 통행에 대해서만 통행량을 집계함

〈표 4-10〉 계절별 기타통행량(도착량) 컬렉션의 구성

컬렉션명	L0_SIGHT_STATISTICS			
컬렉션 설명	읍면동별 연령별, 성별, 계절에 따른 통행량(도착량)을 분석한 데이터			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
도착 읍면동	EmdID	Integer	"EmdID" : 도착읍면동코드	
도착 시군구	SggID	Integer	"SggID" : 도착시군구코드	8자리로 변경
총 통행량	Total	Double	"Total" : 총 통행량	
연령대	MainAge	Integer	"MainAge" : 연령	
남자 통행량	Male	Double	"Male" : 남자 통행량	
여자 통행량	FeMale	Double	"FeMale" : 여자 통행량	
봄 통행량	Spring	Double	"Spring" : 봄 통행량	3~5월
여름 통행량	Summer	Double	"Summer" : 여름 통행량	6~8월
가을 통행량	Autumn	Double	"Autumn" : 가을 통행량	9~11월
겨울 통행량	Winter	Double	"Winter" : 겨울 통행량	12~2월

라. 통행목적별 통행량

- 모바일 통신데이터의 출발지 체류유형과 도착지 체류유형 정보를 활용하여 통행 목적을 출근, 퇴근, 하교, 귀가, 여가, 관광으로 구분하고, 각 통행목적별 집계한 통행량을 의미함
- 모바일 통신데이터에는 출발지에서의 체류유형과 도착지에서의 체류유형이 집(H), 회사(C), 학교(S), 잠재체류지(R, X), 기타(O)로 구분되어있어, <표 4-11>과 같이 출발지의 체류유형과 도착지의 체류유형 조합으로 통행목적을 구분할 수 있음
 - 단, 집(H)→기타 잠재(X)와 집(H)→기타(O) 통행 중 시군구간 이동만 “관광” 목적으로 구분함

<표 4-11> 통행목적 구분 기준

O \ D	H(집)	C(회사)	S(학교)	R(정기잠재)	X(기타잠재)	O(기타)
H(집)	귀가	출근	등교	여가	여가(관광)	여가(관광)
C(회사)	퇴근	제거	데이터없음	여가	여가	여가
S(학교)	하교	데이터없음	제거	여가	여가	여가
R(정기잠재)	귀가	제거	제거	여가	여가	여가
X(기타잠재)	귀가	제거	제거	여가	여가	여가
O(기타)	귀가	제거	제거	여가	여가	여가

- 통행목적이 구분된 식별자를 중심으로 평일, 주말 총 통행량을 집계하여 <표 4-12>와 같이 통행목적별 통행량을 구축함

<표 4-12> 통행목적별 통행량 컬렉션의 구성

컬렉션명	EL_PASSPURPOSE			
컬렉션 설명	통행목적별 통행량을 집계한 데이터			
데이터 수	약 6억 5천만 건 (651,671,120)		데이터 크기	약 50.1GB
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { “A1” : 출발읍면동코드, “E1” : 도착읍면동코드, “J” : 성별, “P” : 통행목적코드	A1, E1, J : integer 통행 목적 코드 코드 ·1: 출근(H→C), 2: 등교(H→S), 3: 퇴근(C→H), 4: 하교(S→H),

컬렉션명	EL_PASSPURPOSE			
컬렉션 설명	동행목적별 동행량을 집계한 데이터			
데이터 수	약 6억 5천만 건 (651,671,120)		데이터 크기	약 50.1GB
			}	5:귀가(R→H, X→H), 6: 여가(H→R, C→R, C→X, S→R, R→R, R→X, X→R, X→X), 8: 관광(H→X) 성별 코드 ·F:1, M:2
평일 총 통행량	WDFV	Integer	"WDFV" : 평일 총 통행량	
주말 총 통행량	WEFV	Integer	"WEFV" : 주말 총 통행량	
출발 시군구	A2	Integer	"A2" : 출발 시군구 코드	8자리로 변경
도착 시군구	E2	Integer	"E2" : 도착 시군구 코드	8자리로 변경
출발 시도	A3	Integer	"A3" : 출발 시도 코드	8자리로 변경
도착 시도	E3	Integer	"E3" : 도착 시도 코드	8자리로 변경
거리	S	Integer	"S" : 거리	직선거리, km 단위
평균시간	T	Integer	"T" : 평균시간	분단위

마. 고령자 통행량

- 모바일 통신데이터에 포함되어있는 통행자의 연령대 정보를 활용하여, 연령대가 70대 이상인 고령인구를 선별한 후 집계한 통행량을 의미함
- 고령자 통행량은 <표 4-13>과 같이 출발지를 기준으로 통행량을 집계한 발생량 DB와 <표 4-14>와 같이 OD기준으로 통행량을 집계한 기종점통행량 DB, 두 가지 형태로 구축함
 - 단, 출발지 체류유형이 기타(X, O), 도착지 체류유형이 정기잠재(R)인 경우 또는 도착지 체류유형이 기타(X, O), 출발지 체류유형이 집(H) 또는 정기잠재(R)인 경우만 집계함

<표 4-13> 고령자 통행량(발생량 기준) 컬렉션의 구성

컬렉션명	viewtlght_epp			
컬렉션 설명	출발 지역 기준의 전체 통행량, 고령자 통행량 데이터			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
출발 시도	s_sido	Integer	"s_sido" : 출발 시도 코드	8자리로 변경
출발 시군구	s_sgg	Integer	"s_sgg" : 출발 시군구 코드	8자리로 변경
출발 읍면동	s_umd	Integer	"s_umd" : 출발 읍면동 코드	
고령자 통행량	elderly_value	Double	"elderly_value" : 고령자통행량	
전체통행량	total_value	Double	"total_value" : 전체통행량	

<표 4-14> 고령자 통행량(OD통행량 기준) 컬렉션의 구성

컬렉션명	L0_ELDER_MAINPASS			
컬렉션 설명	70대 이상 고령자의 통행량을 지역 기준으로 그룹화한 데이터			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
출발 시도	A3	Integer	"A3" : 출발 시도 코드	8자리로 변경
출발 시군구	A2	Integer	"A2" : 출발 시군구 코드	8자리로 변경
출발 읍면동	A1	Integer	"A1" : 출발 읍면동 코드	
도착 시도	E3	Integer	"E3" : 도착 시도 코드	8자리로 변경
도착 시군구	E2	Integer	"E2" : 도착 시군구 코드	8자리로 변경
도착 읍면동	E1	Integer	"E1" : 도착 읍면동 코드	
통행량	sum	Double	"sum" : 고령자통행량	

2) 통행시간

가. 평균통행시간

- 평균통행시간은 통행자가 목적지까지 도달하는 데 걸린 통행시간을 평균한 값을 의미하며, 모바일 통신데이터에는 통행자별 통행시간 정보가 저장되어 있어, 해당 정보를 활용하여 평균 통행시간을 산출할 수 있음
- 평균통행시간 DB는 별도로 구축하지 않고, 통행목적별 통행량을 집계한 DB에 평균시간 필드를 추가하여 구축함 (<표 4-12> 참조)

나. 출·퇴근 통행시간

- 모바일 통신데이터의 출발지 체류유형과 도착지 체류유형 정보를 활용하여 출근, 퇴근으로 구분된 통행에 대하여 평균통행시간을 산출한 것을 의미함
- 출근 통행과 퇴근 통행에 대하여 각각 DB를 구축함 (출근 통행은 <표 4-15>, 퇴근 통행은 <표 4-16> 참조)
 - 연령대가 20대~60대이면서, 출발지 체류유형이 집(H)이고 도착지 체류유형이 회사(C)인 경우(“출근 통행”)와 출발지 체류유형이 회사(C)이고 도착지 체류유형이 집(H)인 경우(“퇴근 통행”)를 대상으로 선정함

〈표 4-15〉 출근 통행시간 컬렉션의 구성

컬렉션명	L0_HOME2COMPANY			
컬렉션 설명	출근 통행을 읍면동, 시간대 기준으로 그룹화한 통행량 데이터			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
출발 읍면동	A1	Integer	“A1” : 출발 읍면동 코드	
출발 시군구	A2	Integer	“A2” : 출발 시군구 코드	8자리로 변경
출발 시도	A3	Integer	“A3” : 출발 시도 코드	8자리로 변경
도착 읍면동	E1	Integer	“E1” : 도착 읍면동 코드	
도착 시군구	E2	Integer	“E2” : 도착 시군구 코드	8자리로 변경
도착 시도	E3	Integer	“E3” : 도착 시도 코드	8자리로 변경
출발 시간대	C	Integer	“C” : 출발시간대 코드	
도착 시간대	G	Integer	“G” : 도착시간대 코드	
평균통행시간	T	Integer	“T” : 평균통행시간	분 단위 (표준 DB의 “통행시간_평균”의 합/ 전체 출근

컬렉션명	L0_HOME2COMPANY			
컬렉션 설명	출근 통행을 읍면동, 시간대 기준으로 그룹화한 통행량 데이터			
				통행량)
통행량	V	Integer	"V" : 전체 통행량	출근 통행

〈표 4-16〉 퇴근 통행시간 컬렉션의 구성

컬렉션명	L0_COMPANY2HOME			
컬렉션 설명	퇴근 통행을 읍면동, 시간대 기준으로 그룹화한 통행량 데이터			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
출발 읍면동	A1	Integer	"A1" : 출발 읍면동 코드	
출발 시군구	A2	Integer	"A2" : 출발 시군구 코드	
출발 시도	A3	Integer	"A3" : 출발 시도 코드	
도착 읍면동	E1	Integer	"E1" : 도착 읍면동 코드	
도착 시군구	E2	Integer	"E2" : 도착 시군구 코드	
도착 시도	E3	Integer	"E3" : 도착 시도 코드	
출발 시간대	C	Integer	"C" : 출발시간대 코드	
도착 시간대	G	Integer	"G" : 도착시간대 코드	
평균통행시간	T	Integer	"T" : 평균통행시간	분 단위 (표준 DB의" 통행시간_평균"의 합/ 전체 출근 통행량)
통행량	V	Integer	"V" : 전체 통행량	출근 통행

3) 통행거리

가. 평균통행거리

- 평균통행거리는 모바일 통신데이터의 출발지, 도착지 정보를 활용하여, 통행자가 목적지까지 이동한 거리를 추정하여 평균한 것을 의미함
 - 원천데이터가 아닌 통행 DB의 형태로 가공된 모바일 통신데이터로는 통행자가 목적지까지 이동한 경로를 확인할 수 없어, 출발지, 도착지의 중심점을 기준으로 산출한 직선거리를 통행거리로 간주함
- 평균통행거리 DB는 별도로 구축하지 않고, 통행목적별 통행량을 집계한 DB에 거리 필드를 추가하여 구축함 (<표 4-12> 참조)

4) 통행비율

가. 내·외부 통행 비율

- 내부통행 비율은 전체 통행에서 내부 통행이 차지하는 비율을 의미하고, 외부통행 비율은 전체 통행에서 외부 통행이 차지하는 비율을 의미하며, 모바일 통신데이터의 출발지, 도착지 정보를 활용하여 산출 가능함
 - 내부통행량은 출발지역과 도착지역이 일치하는 통행량을 집계하고, 외부통행량은 출발지역과 도착지역이 일치하지 않는 통행량을 집계하여 산출할 수 있음

$$\text{내부통행비율} = \left(\frac{\text{내부 통행량}}{\text{전체 통행량}} \right) * 100$$

$$\text{외부통행비율} = \left(\frac{\text{외부 통행량}}{\text{전체 통행량}} \right) * 100$$

- 내부 통행 비율과 외부통행 비율 DB는 시군구, 시도 단위 2가지 형태로 구축함 (<표 4-17>, <표 4-18> 참조)
 - 내부 통행량과 외부 통행량을 집계한 필드를 생성하여 내부통행 비율, 외부통행비율 조회 시 위 산식에 의해 계산되도록 설계함

- 주말에 발생한 통행만 집계하고, 출근, 퇴근, 등교, 하교 통행에 해당하는 데이터는 제외함

〈표 4-17〉 내·외부 통행(시군구) 컬렉션의 구성

컬렉션명	L0_SIGHT_IO_SGG			
컬렉션 설명	시군구별 성별에 따른 집출입 통행량을 분석한 데이터			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
출발 시군구	A2	Integer	“A2” : 출발 시군구 코드	8자리로 변경
출발 시도	A3	Integer	“A3” : 출발 시도 코드	8자리로 변경
내부 통행량 (여성)	in_f	Integer	“in_f” : 여성 내부 통행량	
내부 통행량 (남성)	in_m	Integer	“in_m” : 남성 내부 통행량	
내부 총 통행량	in_t	Integer	“in_t” : 전체 내부 통행량	여성 내부 통행량 +남성 내부 통행량
외부 통행량 (여성)	out_f	Integer	“out_f” : 여성 외부 통행량	
외부 통행량 (남성)	out_m	Integer	“out_m” : 남성 외부 통행량	
외부 총 통행량	out_t	Integer	“out_t” : 전체 외부 통행량	여성 외부 통행량 +남성 외부 통행량

〈표 4-18〉 내·외부 통행(시도) 컬렉션의 구성

컬렉션명	L0_SIGHT_IO_SIDO			
컬렉션 설명	시도별 성별에 따른 집출입 통행량을 분석한 데이터			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
출발 시도	A3	Integer	“A3” : 출발 시군구 코드	8자리로 변경
내부 통행량 (여성)	in_f	Integer	“in_f” : 여성 내부 통행량	
내부 통행량 (남성)	in_m	Integer	“in_m” : 남성 내부 통행량	
내부 총 통행량	in_t	Integer	“in_t” : 전체 내부 통행량	여성 내부 통행량 +남성 내부 통행량
외부 통행량 (여성)	out_f	Integer	“out_f” : 여성 외부 통행량	
외부 통행량 (남성)	out_m	Integer	“out_m” : 남성 외부 통행량	
외부 총 통행량	out_t	Integer	“out_t” : 전체 외부 통행량	여성 외부 통행량 +남성 외부 통행량

나. 고령자 통행 비율

- 고령자 통행 비율은 전체 통행자 중에서 고령 통행자가 차지하는 비율을 의미함

$$\left(\frac{\text{고령인구 통행량}}{\text{전체 통행량}} \right) * 100$$

- 고령자 통행 비율은 별도로 DB를 구축하지 않고, 고령자 통행량 DB (<표 4-13>, <표 4-14> 참조)를 활용하여 위 산식에 의해 계산되도록 설계함

다. 심야시간 출·퇴근 통행 비율

- 전체 출·퇴근 통행자 중에서 심야시간대 출·퇴근하는 통행자가 차지하는 비율을 의미함

- 모바일 통신데이터의 출발지 체류유형, 도착지 체류유형 정보를 활용하여 출근 목적 통행과 퇴근 목적 통행을 구분하여 통행량을 집계한 후, 모바일 통신데이터의 출발시간, 도착시간 정보를 활용하여 심야시간대(17~23시) 이루어진 통행을 추출하여 전체 통행 중에 심야시간대 통행의 비율을 산출함

$$\left(\frac{\text{심야시간대 출·퇴근 목적 통행량}}{\text{전체 출·퇴근 목적 통행량}} \right) * 100$$

- 심야시간 출퇴근 통행 비율 DB는 출근 통행 비율과 퇴근 통행 비율을 구분하여 구축함 (<표 4-19> 참조)
- 단, DB 구축시 통행자의 연령대는 주요 경제활동인구인 20대~60대로 한정함

〈표 4-19〉 심야시간 출퇴근 통행 비율 컬렉션의 구성

컬렉션명	L0_WORKTYPE_COMMUTE			
컬렉션 설명	전체 통근통행 중 심야시간대 통근 통행량 분석			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
출발 시도	A3	Integer	“A3” : 출발 시도 코드	8자리로 변경
출발 시군구	A2	Integer	“A2” : 출발 시군구 코드	8자리로 변경
도착 시도	E3	Integer	“E3” : 도착 시도 코드	8자리로 변경
도착 시군구	E2	Integer	“E2” : 도착 시군구 코드	8자리로 변경
전체 통행량	V	Integer	“V” : 전체 통행량	출근 통행량 + 퇴근 통행량
출근 통행량	CV	Integer	“CV” : 출근 통행량	
퇴근 통행량	HV	Integer	“HV” : 퇴근 통행량	
심야시간 출근 통행량	NCV	Integer	“NCV” : 심야시간 출근 통행량	17~23시 출근
심야시간 출근 통행비율	NCR	Double	“NCR” : 심야시간 출근 통행 비율	(출근 심야 통행량(NCV) / 출근 통행량(CV)) * 100 (소수점 1자리)
심야시간 퇴근 통행량	NHV	Integer	“NHV” : 심야시간 퇴근 통행량	17~23시 퇴근
심야시간 퇴근 통행비율	NHR	Double	“NHR” : 심야시간 퇴근 통행 비율	(퇴근 심야 통행량(NHV) / 퇴근 통행량(HV)) * 100 (소수점 1자리)

5) 기타

가. 출·퇴근 증감량

- 출·퇴근 증감량은 특정 지역의 출·퇴근 통행량의 변화를 분석할 수 있는 지표로서, 전월 대비 기준월의 통근 목적 통행량의 증가량, 감소량을 측정한 것을 의미함
 - 모바일 통신데이터의 출발지 체류유형, 도착지 체류유형 정보를 활용하여 출·퇴근 통행을 추출하고, 월별로 통행량을 집계하여 변화량을 비교·분석하여 산출함
- 월별 출·퇴근 통행량을 <표 4-20>과 같이 집계하여 DB로 구축하고, 출·퇴근 증감량은 조회 시 전월 대비 기준월의 출·퇴근 통행량의 변화량을 산출하도록 설계함
 - 단, 연령대가 20대~60대인 통행자, 평일에 이루어진 통행으로 한정함

〈표 4-20〉 월별 통근 통행량 컬렉션의 구성

컬렉션명	EMPLOYMENTCHANGE_COMMUTE			
컬렉션 설명	월별 통근통행량 분석			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
출발 시도	A3	Integer	“A3” : 출발 시도 코드	8자리로 변경
출발 시군구	A2	Integer	“A2” : 출발 시군구 코드	8자리로 변경
도착 시도	E3	Integer	“E3” : 도착 시도 코드	8자리로 변경
도착 시군구	E2	Integer	“E2” : 도착 시군구 코드	8자리로 변경
월	M	Integer	“M” : 월	1~12, 출발 기준 월
통행량	R	Integer	“R” : 통행량	출근, 퇴근 통행량

나. 도심공동화 심각도

- 도심공동화 현상이 심각하게 나타나는 지역을 확인할 수 있는 지표로서, 유동인구 대비 상주인구의 비율로 측정함

$$\left(\frac{\text{상주인구수}}{\text{유동인구수}} \right) * 100$$

- 상주인구와 유동인구는 모바일 통신데이터의 도착시간, 도착지 체류유형 정보를

활용하여 추출함

- 상주인구와 유동인구에 대한 통행량을 집계하여 <표 4-21>과 같이 DB를 구축하고, 도심공동화 심각도는 조회 시 해당 DB에서 유동인구 대비 상주인구의 비율을 환산하도록 설계함
- 상주인구 집계 기준: 도착시간대가 19~6시에 속하고, 도착지 체류유형이 집(H)인 경우
- 유동인구 집계 기준: 도착시간대가 10~18시에 속하고, 도착 체류유형이 학교(S) 또는 회사(C)인 경우

〈표 4-21〉 도심공동화 심각도 컬렉션의 구성

컬렉션명		L0_DOUGHUNT_EFFECT		
컬렉션 설명		읍면동별 귀가 통행량을 분석한 데이터		
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
도착 시도	E3	Integer	"E3" : 도착 시도 코드	8자리로 변경
도착 시군구	E2	Integer	"E2" : 도착 시군구 코드	8자리로 변경
도착읍면동	E1	Integer	"E1" : 도착 읍면동 코드	
월	M	Integer	"M" : 월	1~12, 출발 기준 월
상주인구	R	Integer	"R" : 귀가 통행량(도착량)	19~6시 / 집
유동인구	A	Integer	"A" : 활동 통행량(도착량)	10~18시 / 회사, 학교

다. 경제활동의존도

- 경제활동의존도는 자족성과 반대되는 개념으로, 자기 지역이 아닌 타 지역에서 경제활동이 얼마나 이루어지고 있는지를 의미함
- 여기서 경제활동은 '회사에서의 근무'를 의미함
- 전체 출근 통행 중에서 자기 지역이 아닌 타 지역으로 이동한 출근 통행이 차지하는 비율을 산출하여 측정함
- 모바일 통신데이터의 출발지 체류유형과 도착지 체류유형 정보를 활용하여 출근 통행을 추출하고, 출발지와 도착지 정보를 활용하여 출근 통행 중에서 타 지역으로 통행한 출근 통행을 선별함

- 경제활동의존도 DB는 <표 4-22>와 같이 출발지, 도착지별 출근 통행량을 집계한 형태로 구축하고, 조회 시 아래 산식에 의해 경제활동의존도가 산출되도록 설계함

$$\left(\frac{\text{특별시·광역시 주변도시에서 특별시·광역시로 출근한 통행량}}{\text{특별시·광역시 주변도시에서 발생한 출근 통행량}} \right) * 100$$

- 대도시(특별시·광역시)에 대한 경제활동의존도를 파악할 수 있도록 특별시·광역시 주변도시에서 발생한 출근 통행 중에서 특별시·광역시로 출근한 통행의 비율이 산출될 수 있도록 설계함
- 단, 연령대가 20~60대인 통행자, 평일, 6~9시 사이에 출발하여 7~10시에 도착한 출근 목적 통행으로 한정함

〈표 4-22〉 경제활동의존도 컬렉션의 구성

컬렉션명	L0_ECONOMIC_DEPENDENCY			
컬렉션 설명	시군구별 출근 통행량을 분석한 데이터			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
도착 시도	E3	Integer	"E3" : 도착 시도 코드	8자리로 변경
도착 시군구	E2	Integer	"E2" : 도착 시군구 코드	8자리로 변경
출발 시도	A3	Integer	"A3" : 출발 시도 코드	8자리로 변경
출발 시군구	A2	Integer	"A2" : 출발 시군구 코드	8자리로 변경
통행량	CG	Integer	"CG" : 통행량	

라. 경제활동 자체수용도

- 경제활동 자체수용도는 경제활동의존도와 반대로, 자기 지역 내에서 경제활동이 얼마나 이루어지고 있는지를 의미함
- 전체 출근 통행 중에서 타 지역으로 이동하지 않고 자기 지역으로 출근한 통행이 차지하는 비율을 산출하여 측정함
 - 모바일 통신데이터의 출발지 체류유형과 도착지 체류유형 정보를 활용하여 출근 통행을 추출하고, 출발지와 도착지 정보를 활용하여 출근 통행 중에서 출발지와 도착지가 동일한 출근 통행을 선별함

- 경제활동 자체수용도 DB는 <표 4-23>과 같이 출발지 기준으로 경제활동 자체수용도를 구축하였으며, 경제활동 자체수용도 산출에 필요한 전체 출근 통행량, 지역 내 통행량, 지역 외 통행량도 추가로 구축함
- 단, 분석지역은 대도시(특별시·광역시)로 한정하여, 구축된 경제활동 자체수용도는 아래 산식에 의해 계산된 것임

$$\left(\frac{\text{특별시·광역시에서 특별시·광역시로 출근한 통행량}}{\text{특별시·광역시에서 발생한 출근 통행량}} \right) * 100$$

- 또한 경제활동의존도와 마찬가지로 분석 조건을 연령대가 20~60대인 통행자, 평일, 6~9시 사이에 출발하여 7~10시에 도착한 출근 목적 통행으로 한정함

<표 4-23> 경제활동 자체수용도 컬렉션의 구성

컬렉션명	L0_ECONOMIC_SELF_DEPENDENCY			
컬렉션 설명	지역 내 출근 통행량을 분석한 데이터			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
출발 시도	A3	Integer	“A3” : 출발 시도 코드	8자리로 변경
출발 시군구	A2	Integer	“A2” : 출발 시군구 코드	8자리로 변경
전체 출근 통행량	total	Integer	“total” : 전체 출근 통행량	
전체 출근 일평균 통행량	totalAVG	Integer	“totalAVG” : 전체 출근 일평균 통행량	평일(주말 제외) 수로 나눔
지역내 출근 통행량	in	Integer	“in” : 지역내 출근 통행량	
지역내 출근 일평균 통행량	inAVG	Integer	“inAVG” : 지역내 출근 일평균 통행량	평일(주말 제외) 수로 나눔
지역외 출근 통행량	out	Integer	“out” : 지역외 출근 통행량	전체 통행량 - 지역내 통행량
지역외 출근 일평균 통행량	outAVG	Integer	“outAVG” : 지역외 출근 일평균 통행량	전체 일평균 - 지역내 일평균
자체수용도	rate	Double	“rate” : 자체수용도	(지역내 통행량 / 전체 통행량) * 100 (소수점1자리)
시군구 Y좌표	Y	Double	“Y” : 시군구 Y좌표	WGS84, 경도
시군구 X좌표	X	Double	“X” : 시군구 X좌표	WGS84, 위도

마. 출·퇴근 통행연결성

- 출·퇴근 통행연결성은 출·퇴근 통행에 대한 기종점간의 모듈성을 의미하며, 아래 산식에 의해 산출 가능함

$$\frac{1}{2M} \sum_{i,j}^N (a_{ij} - \frac{k_i k_j}{2M}) \delta[C(i), C(j)]$$

M : 전체 링크 수

N : 전체 노드 수

a_{ij} : i, j 간 링크 (있을 경우 1, 없을 경우 0)

k_i : 노드 i 가 지나는 링크의 개수

$C(i)$: 노드 i 가 속하는 커뮤니티

$\delta[C(i), C(j)]$: $C(i)$ 와 $C(j)$ 가 같은 커뮤니티일 때 1, 다를 때 0

- 모바일 통신데이터에는 위 산식에 필요한 링크 수, 노드 수 등의 값이 포함되어 있지 않으므로, 분석하고자 하는 통행구간의 출·퇴근 통행량이 해당 지역의 전체 출·퇴근 통행량에서 얼마나 차지하는지로 분석을 대체함

- 모바일 통신데이터의 출발지 체류유형과 도착지 체류유형을 활용하여 출·퇴근 통행을 선별하고 통행구간별 출·퇴근 통행량을 집계한 후, 분석하고자 하는 통행구간의 출·퇴근 통행량이 분석지역(분석하고자 하는 통행구간이 속한 지역)의 출·퇴근 통행량에서 얼마나 차지하고 있는지 비율로 환산함

- 출·퇴근 통행연결성은 <표 4-24>와 같이 시군구 단위의 기종점간 출·퇴근 통행량의 형태로 DB를 구축한 후, 조회시 아래 산식에 의해 계산되어 표출되도록 설계함

$$\left(\frac{\text{분석구간의(시군구 단위) 출·퇴근 통행량}}{\text{분석지역의(시도 단위) 출·퇴근 통행량}} \right) * 100$$

- 단, 연령대가 20대~60대인 통행자, 평일에 이루어진 통행 중 출·퇴근 통행에 해당되는 데이터만 추출함

〈표 4-24〉 출·퇴근 통행연결성 컬렉션의 구성

컬렉션명	L0_PASSCONNECT_COMMUTE			
컬렉션 설명	시군구별 출퇴근 통행 연결성 분석			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
출발 시도	A3	Integer	“A3” : 출발 시도 코드	8자리로 변경
출발 시군구	A2	Integer	“A2” : 출발 시군구 코드	8자리로 변경
도착 시도	E3	Integer	“E3” : 도착 시도 코드	8자리로 변경
도착 시군구	E2	Integer	“E2” : 도착 시군구 코드	8자리로 변경
전체 통행량	V	Integer	“V” : 전체통행량	출근, 퇴근 통행

바. 직주균형(지표)

- 직주균형지표는 도시공간구조 측면에서 거주지와 근무지의 배치가 얼마나 효율적으로 이루어져있는지를 의미하며 직주균형지표를 토대로 통근 통행의 효율성을 확인할 수 있음
- 직주 불균형으로 인해 나타나는 ‘초과통근정도’, ‘통근잠재력’을 직주균형지표로 간주하고, 모바일 통신데이터를 통해 초과통근정도와 통근잠재력을 산출하여 DB로 구축함
 - ‘초과통근정도’는 실제통근거리 대비 실제통근거리와 최소통근거리간의 차이를 나타낸 값으로, 값이 작을수록 직주균형을 이루고 있고 통근이 효율적으로 이루어지고 있음을 의미함

$$\left(\frac{\text{실제통근거리} - \text{최소통근거리}}{\text{실제통근거리}}\right) * 100$$

- 최소통근거리: 거주지와 근무지간의 비용(통행거리, 시간)을 최소화하였을 때 나타나는 이상적인 통근거리로, 선형계획법을 통해 산출하며, 제약조건을 만족하는 가능해들 중에서 목적함수를 최소화하는 값을 의미함

$$\text{최소통근거리} = \min \frac{1}{T} * \left(\sum_h \sum_j t_{hj} * d_{hj}\right)$$

- 제약조건:

$$\sum_j t_{hj} = O_h, \quad \sum_h t_{hj} = D_j, \quad t_{hj} \geq 0$$

여기서,

O_h : 해당 지역에 거주하는 전체 통근자의 수

D_j : 해당 지역에 근무하는 통근자의 수

- 실제통근거리: 실제 통근에서 나타나는 통근거리로, 모바일 통신데이터에서 기종점간 출근통행량을 집계하여 산출함

$$ARC = \frac{\sum_h \sum_j t_{hj} * d_{hj}}{T}$$

여기서,

t_{hj} : 해당 지역의 통근자 수 (거주지에서 근무지로 이동하는 통행)

d_{hj} : 거주지와 근무지의 거리

T : 전체 통근자 수

- ‘통근잠재력’은 최대통근거리와 실제통근거리의 차이 대비 실제통근거리와 최소통근거리의 차이가 어느 정도인지를 측정한 값으로, 값이 작을수록 직주균형을 이루고 있고 통근이 효율적으로 이루어지고 있음을 의미함

$$\left(\frac{\text{실제통근거리} - \text{최소통근거리}}{\text{최대통근거리} - \text{실제통근거리}} \right) * 100$$

- 최대통근거리: 선형계획법을 통해 산출하며, 제약조건을 만족하는 가능해들 중에서 목적함수를 최대화하는 값을 의미함

$$\text{최대통근거리} = \max \frac{1}{T} * \left(\sum_h \sum_j t_{hj} * d_{hj} \right)$$

- 제약조건:

$$\sum_j t_{hj} = O_h, \quad \sum_h t_{hj} = D_j, \quad t_{hj} \geq 0$$

여기서,

O_h : 해당 지역에 거주하는 전체 통근자의 수

D_j : 해당 지역에 근무하는 통근자의 수

- 직주균형지표 DB는 <표 4-25>와 같은 형태로 시군구별 월별 최소통근거리, 실제통근거리, 최대통근거리, 초과통근정도, 통근잠재력을 산출하여 구축함

- 실제통근거리 산출시 거주지와 근무지간의 거리는 시군구별 중심점간의 직선 거리로 계산하였으며, 내부통행인 경우에는 해당 지역의 면적을 π 로 나눈 후 루트를 씌워서 계산함

〈표 4-25〉 초과통근량 컬렉션의 구성

컬렉션명	L0_PASSCONNECT_COMMUTE			
컬렉션 설명	지역 내 거주지와 근무지의 관계 분석			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
시군구 ID	sigungu_id	Integer	"sigungu_id" : 시군구 코드	8자리로 변경
월	month	Integer	"month" : 월	1~12
최소통근거리	mrc	Double	"mrc" : 최소통근거리	
실제통근거리	arc	Double	"arc" : 실제통근거리	
최대통근거리	maxrc	Double	"maxrc" : 실제통근거리	
초과통근정도	e	Double	"e" : 초과통근정도	0~100% (수치가 커질수록 비효율적인 통행으로 해석)
통근잠재력	c	Double	"c" : 통근잠재력	0~100% (수치가 커질수록 비효율적인 통행으로 해석)

제2절 온라인 서비스를 통한 교통지표 제공

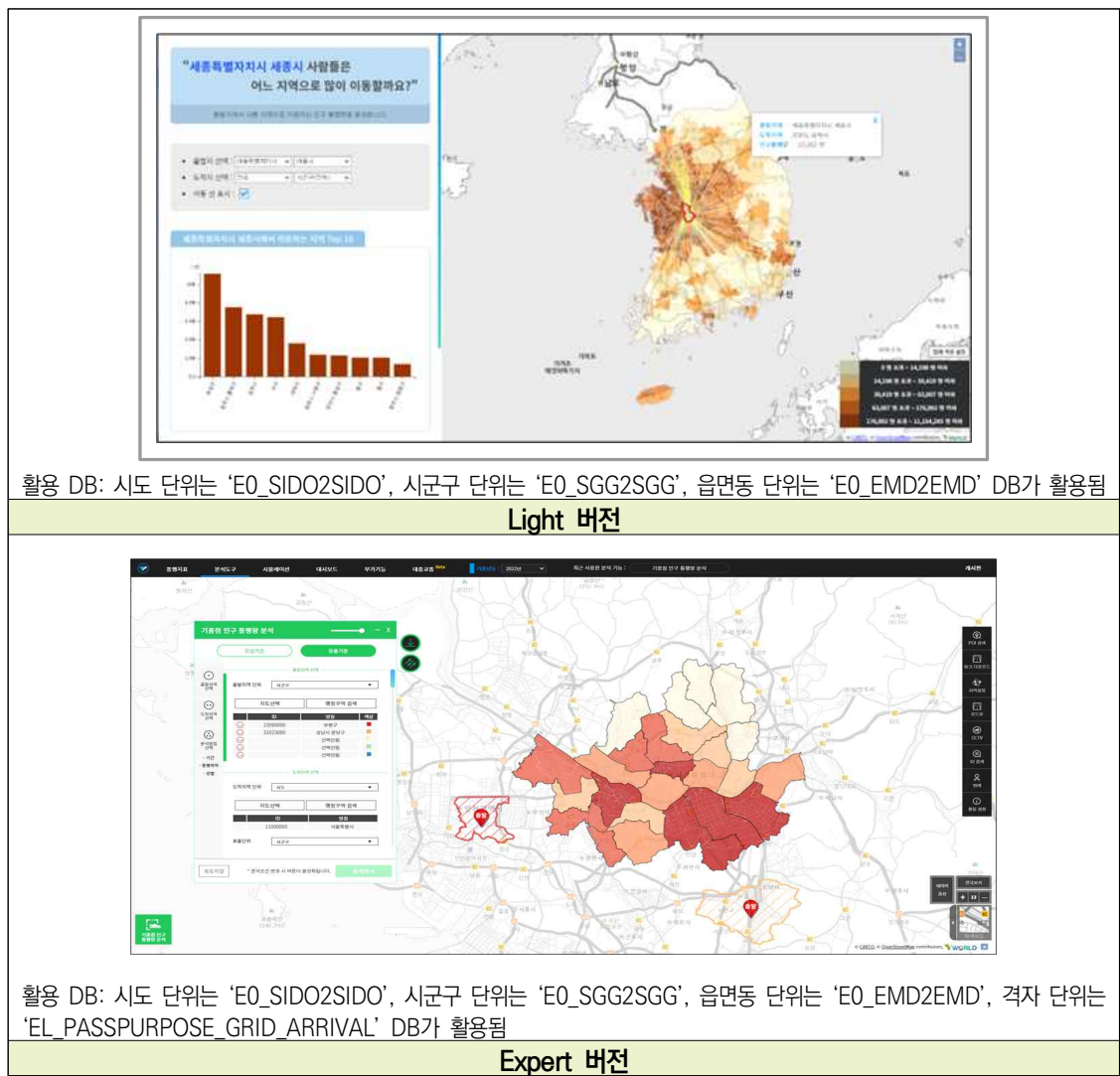
1. 개요

- 모바일 통신데이터의 접근성·활용성을 높이기 위해 모바일 통신데이터로 구축한 교통지표를 온라인 서비스를 통해 제공하고자 함
 - 모바일 통신데이터는 민간데이터로서, 통신사에서 분석 목적에 맞게 데이터를 판매하기 때문에 데이터 구입을 통해서만 접근할 수 있었음
 - 또한 모바일 통신데이터에 대한 이해부터, 분석 목적에 맞는 형태로 가공하고 결과를 확인하기까지의 많은 시간과 노력이 소요되어 일부 전문가에 한해서만 사용됨
 - 모바일 통신데이터로 구축한 교통지표 DB를 온라인 서비스 내에 구축하고, 교통지표를 확인할 수 있는 분석도구를 제공하고자 함
 - 데이터 분석프로그램(SAS, R, Python 등)을 다룰 수 없는 사용자도 쉽게 분석하고 이해할 수 있는 형태로 제공하고자 함
 - 기본적인 분석기능을 제공하는 Light 버전의 분석도구와 심층적인 분석기능을 제공하는 Expert 버전의 분석도구를 제공하고자 함
 - 연구, 정책 등에 자주 활용되는 교통지표는 다양한 정보를 한눈에 확인할 수 있도록 대시보드 형태로 제공하고, 시계열적 변화를 확인할 수 있도록 월간리포트로 제공하고자 함
 - 분석도구, 대시보드, 월간리포트로 제공되는 정보 외, 특정 분석 목적에 맞춘 정보가 필요한 경우 사용자가 직접 모바일 통신데이터를 가공할 수 있도록, 다운로드 기능을 통해 전처리 된 모바일 통신데이터 DB와 모바일 통신데이터로 구축한 모든 교통지표 DB를 제공하고자 함

2. 분석도구를 통한 교통지표 제공

가. 통행량 지표를 제공하는 분석도구

- 기종점 인구 통행량 분석
 - 출발지 기준으로 도착량이 가장 많은 지역, 도착지 기준으로 발생량이 가장 많은 지역을 분석하는 기능으로, 발생량·도착량 DB를 활용하여 분석되도록 설계함

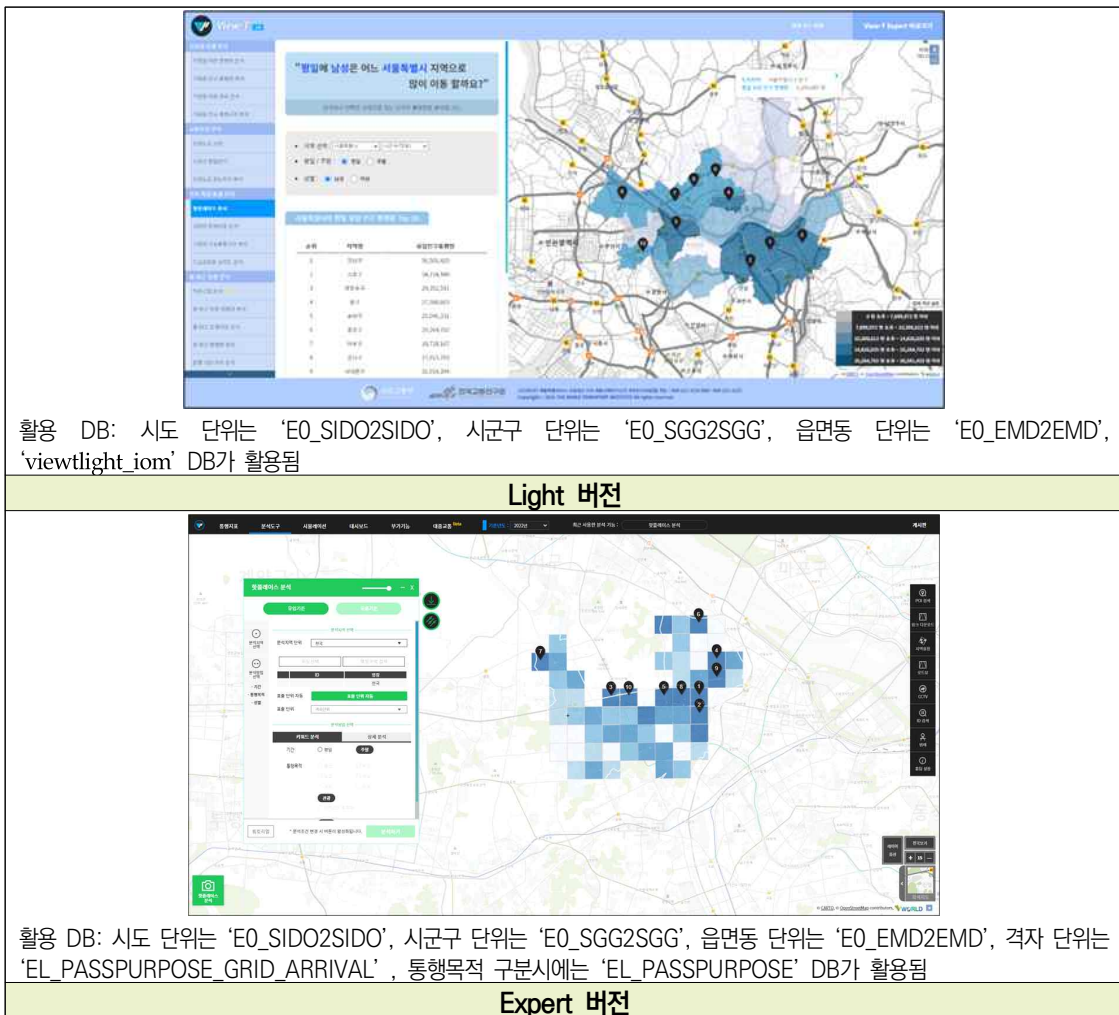


<그림 4-4> 기종점 인구 통행량 분석도구 화면

- Light 버전에서는 시도, 시군구, 읍면동 단위까지 분석 가능하고, Expert 버전에서는 시도, 시군구, 읍면동, 격자 단위까지 분석 가능함
- Expert 버전에서는 기간, 통행목적, 통행자 조건(성, 연령)을 상세 설정할 수 있음

○ 핫플레이스 분석

- 통행량(도착량)이 가장 많은 지역을 분석하는 기능으로, 도착량 DB를 활용하여 분석되도록 설계함
- Light 버전에서는 분석대상을 최소 시군구 단위로 선택할 수 있고, 분석결과는 읍면동 단위로 표출되며, Expert 버전에서는 분석대상을 최소 읍면동 단위로 선택할 수 있고, 분석결과는 격자 단위로 표출됨
- Expert 버전에서는 기간, 통행목적, 통행자 조건(성, 연령)을 상세 설정할 수 있음



<그림 4-5> 핫플레이스 분석도구 화면

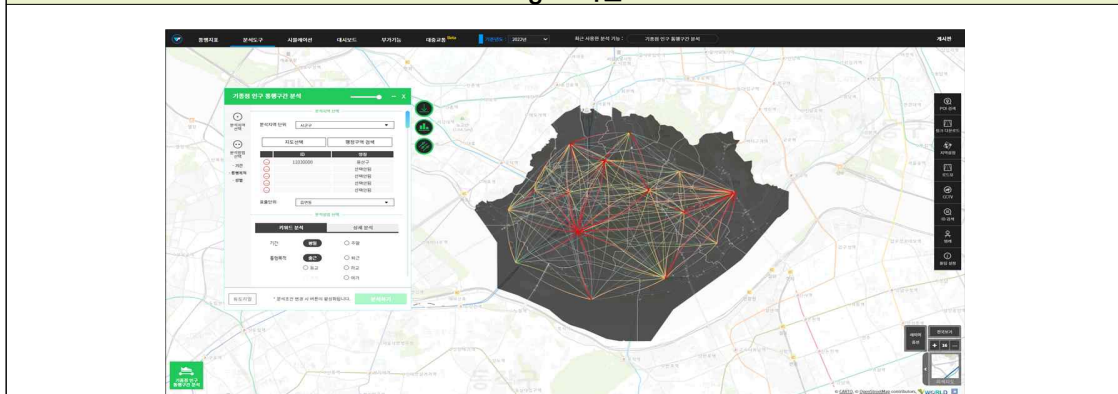
○ 기종점 인구 통행구간 분석

- 선택지역 내에서 가장 통행량이 많은 구간을 분석하는 기능으로, 기종점 통행량(분포량) DB를 활용하여 분석되도록 설계함
- Light 버전에서는 분석대상을 최소 시군구 단위로 선택할 수 있고 분석결과는 읍면동 단위로 표출되며, Expert 버전에서는 분석대상을 최소 읍면동 단위로 선택할 수 있고, 분석결과는 격자 단위로 표출됨
- Light 버전에서는 귀가 통행에 대해서만 분석 가능하며, Expert 버전에서는 기간, 통행목적, 통행자 조건(성, 연령)을 상세 설정할 수 있음



활용 DB: 시도 단위는 'E0_SIDO2SIDO', 시군구 단위는 'E0_SGG2SGG', 읍면동 단위는 'E0_EMD2EMD', 'viewlight_ion', 통행목적 구분시에는 'EL_PASSPURPOSE' DB가 활용됨

Light 버전



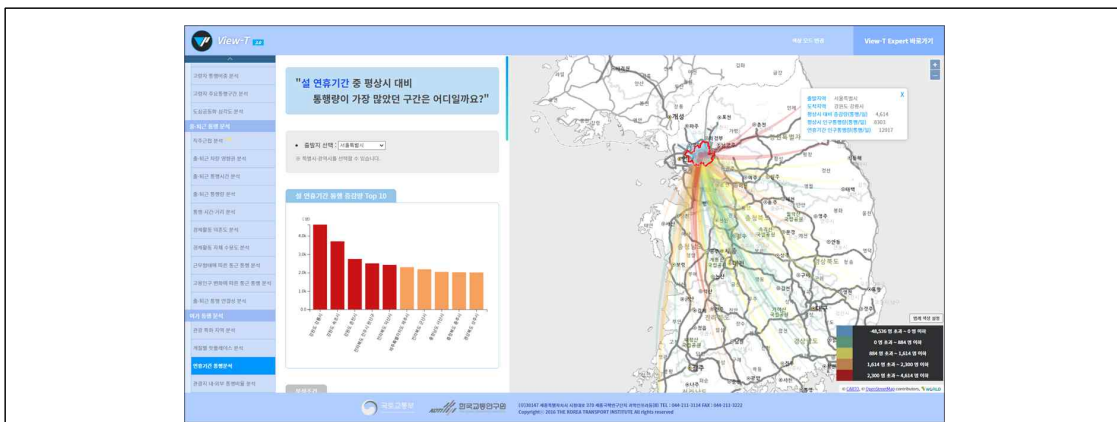
활용 DB: 시도 단위는 'E0_SIDO2SIDO', 시군구 단위는 'E0_SGG2SGG', 읍면동 단위는 'E0_EMD2EMD', 격자 단위는 'EL_PASSPURPOSE_GRID_ARRIVAL', 통행목적 구분시에는 'EL_PASSPURPOSE' DB가 활용됨

Expert 버전

<그림 4-6> 기종점 인구 통행구간 분석도구 화면

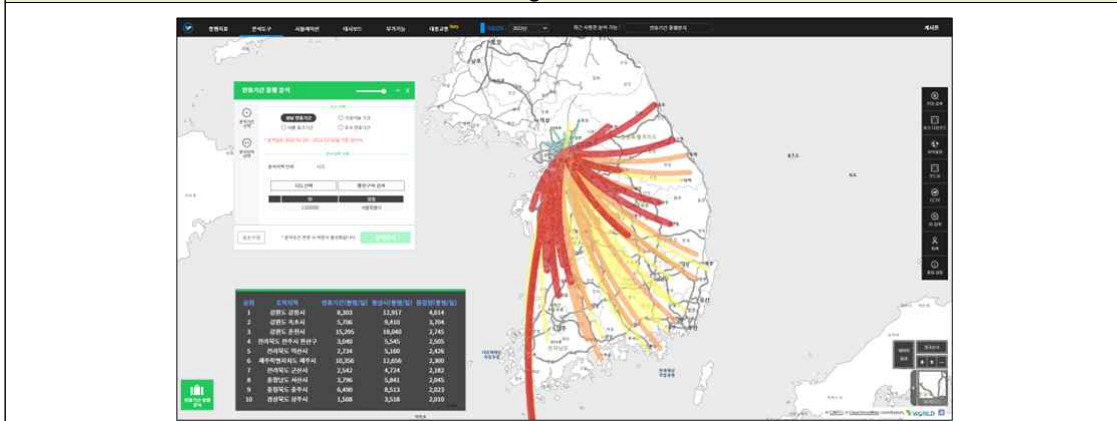
○ 연휴기간 통행분석

- 연휴기간(설, 가정의 달, 하계 연휴기간, 추석) 지역별 유동인구 통행량을 분석하는 기능으로, 연휴기간 통행량 DB를 활용하여 분석되도록 설계함
- 출발지(시도 단위)를 기준으로 네 가지 연휴기간에 대하여 최다 통행구간 10개가 표출됨
- 분석기능은 Light 버전, Expert 버전 모두 동일하나, Expert 버전은 Light 버전보다 표출 방식을 다양하게 설정할 수 있음 (다른 분석 결과와의 중첩, Base map 변경, 범례 조절 등)



활용 DB: 'LO_HOLLIDAY_PASS' DB가 활용됨

Light 버전



활용 DB: 'LO_HOLLIDAY_PASS' DB가 활용됨

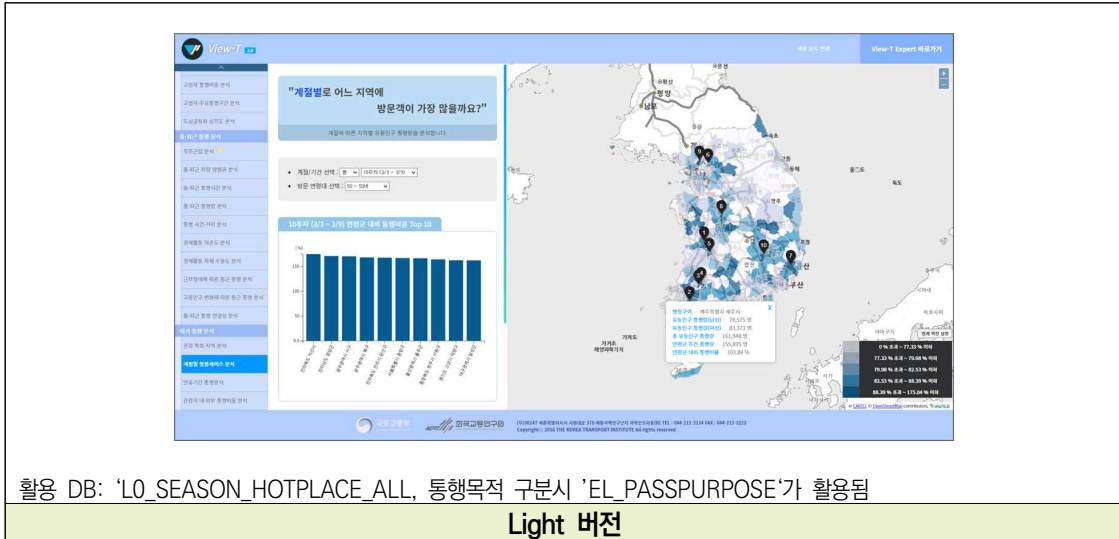
Expert 버전

〈그림 4-7〉 연휴기간 통행 분석도구 화면

○ 계절별 핫플레이스분석 (Light 버전 한정)

- 계절에 따른 지역별 유동인구 통행량을 분석하는 기능으로, 계절별 통행량 DB(일주일 단위로 집계한 통행량)를 활용하여 분석되도록 설계함

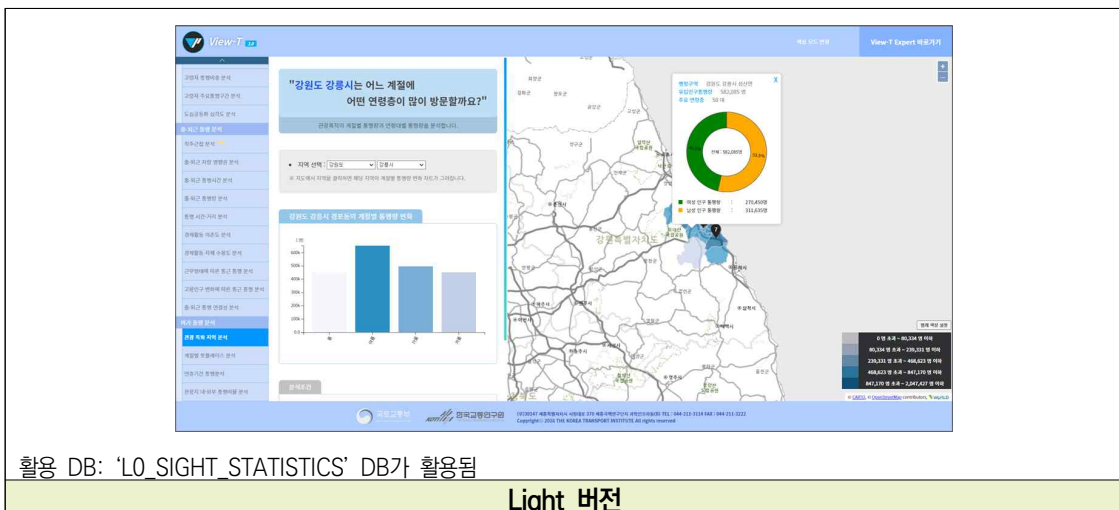
- 계절별 주차별(Week) 연령대별 통행량이 단계구분도(시군구 단위)로 표출되며, 최다 통행지역 10곳이 표기됨
- 분석지역별(시군구별) 연평균 주간 통행량과, 연평균 대비 주차별 통행량의 비율이 표출됨



<그림 4-8> 계절별 핫플레이스 분석도구 화면

○ 관광특화지역분석 (Light 버전 한정)

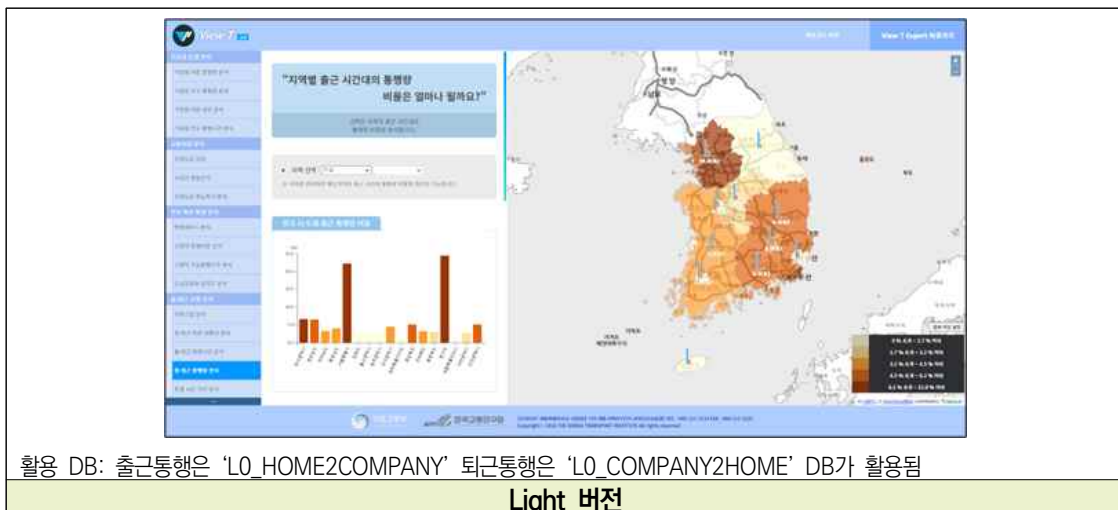
- 지역별(시군구 단위) 관광목적의 계절별 통행량을 분석하는 기능으로, 계절별 통행량 DB(계절 단위로 집계한 통행량)를 활용하여 분석되도록 설계함
- 분석지역에 대한 계절별 통행량이 그래프로 표출되며, 지역별 주요 통행자의 특성이 표출됨 (성비, 주요 연령층)



<그림 4-9> 관광특화지역 분석도구 화면

○ 출·퇴근 통행량 분석 (Light 버전 한정)

- 지역별(시도, 시군구 단위) 출·퇴근 통행량을 분석하는 기능으로, 출·퇴근 통행시간 DB의 통행량 필드를 활용하여 분석되도록 설계함
 - 분석지역의 출·퇴근 통행량이 전 지역(분석지역보다 상위 공간레벨)의 출·퇴근 통행량에서 차지하는 비율이 그래프와 단계구분도로 표출됨 (예: 전국 출·퇴근 통행량 중 서울의 출·퇴근 통행량이 차지하는 비율)
 - 출·퇴근목적 통행시간을 5가지(30분 미만, 30분 이상 60분 미만, 60분 이상 90분 미만, 90분 이상 120분 미만, 120분 이상)로 분류하고, 카테고리별 통행량이 차지하는 비율이 지도에 표기됨



〈그림 4-10〉 출·퇴근 통행량 분석도구 화면

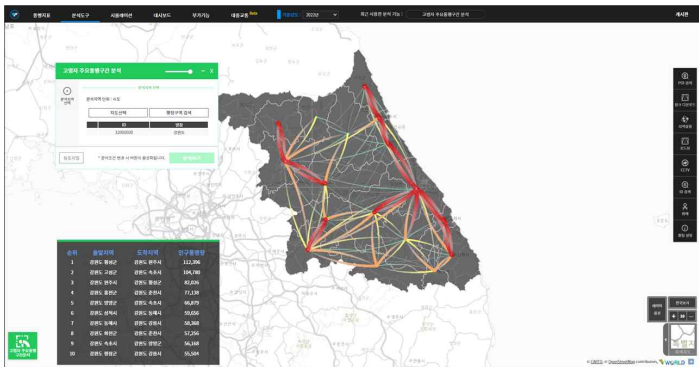
○ 고령자 주요통행구간 분석

- 지역별 고령자의 주요통행구간을 분석하는 기능으로, 고령자 통행량 DB를 활용하여 분석되도록 설계함
 - 선택지역(시군구 단위) 내에서 고령자의 통행이 많은 구간이 읍면동 단위로 지도에 표출되며, 최다 통행구간 10개가 표로 표출됨
 - 분석기능은 Light 버전, Expert 버전 모두 동일하나, Expert 버전은 Light 버전보다 표출 방식을 다양하게 설정할 수 있음 (다른 분석 결과와의 중첩, Base map 변경, 범례 조절 등)



활용 DB: 'viewlight_app' DB가 활용됨

Light 버전



활용 DB: 'LO_ELDER_MAINPASS' DB가 활용됨

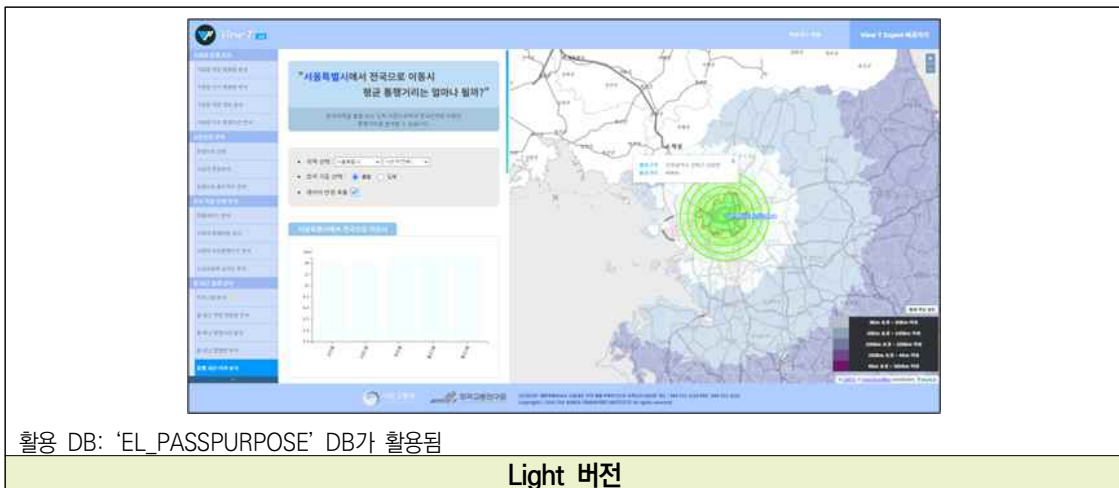
Expert 버전

<그림 4-11> 고령자 주요통행구간 분석도구 화면

나. 통행시간·통행거리 지표를 제공하는 분석도구

○ 통행시간·거리 분석(Light 버전 한정)

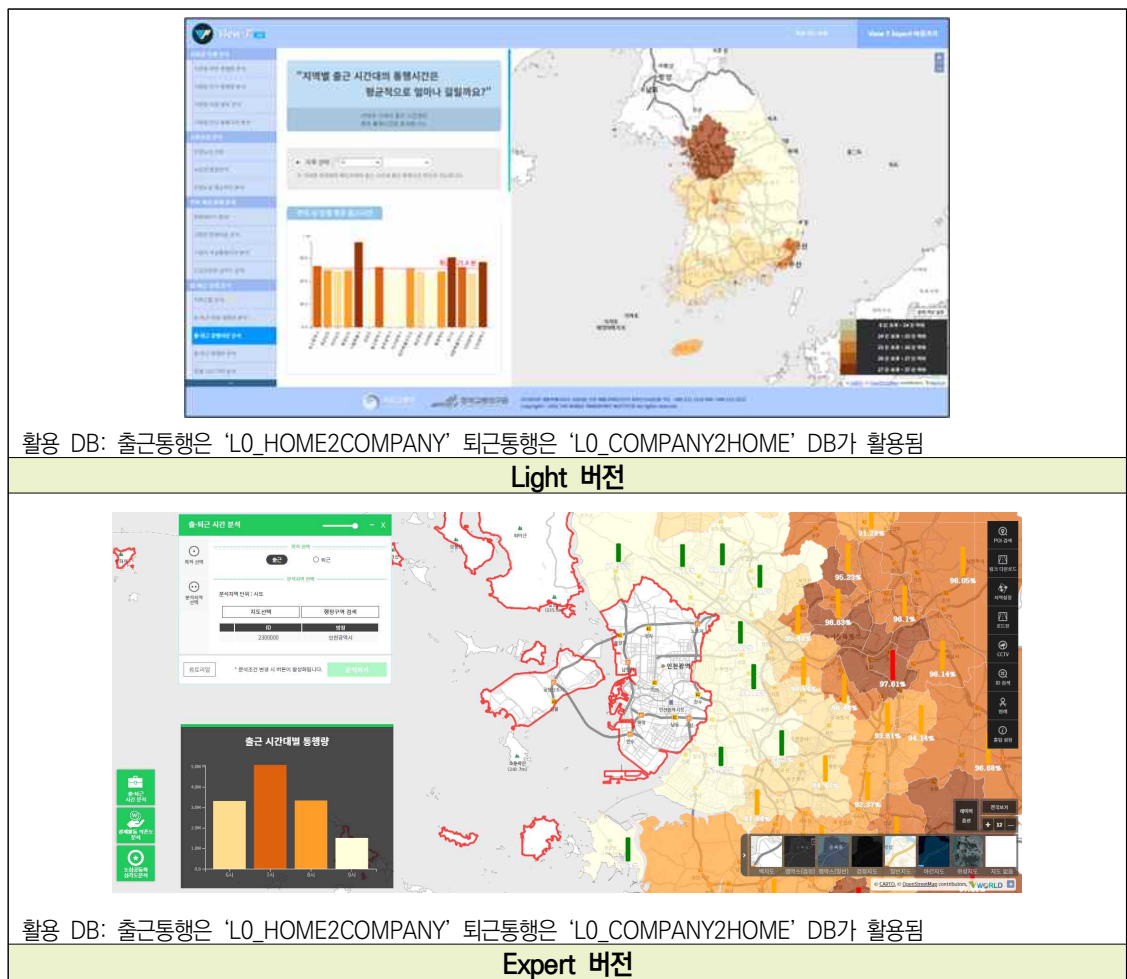
- 지역별 평균통행시간, 평균통행거리를 분석하는 기능으로, 통행목적별 DB에 포함된 평균시간, 거리 필드를 활용하여 분석되도록 설계함
- 최소 시군구 단위로 분석지역을 선택할 수 있으며, 분석지역을 출발지로 설정하였을 때의 평균통행시간·평균통행거리와 분석지역을 도착지로 설정하였을 때의 평균통행시간·평균통행거리가 계산되어 표출됨



〈그림 4-12〉 통행시간·거리 분석도구 화면

○ 출·퇴근 통행시간 분석

- 지역별 출근목적의 평균통행시간과 퇴근목적의 평균통행시간을 분석하는 기능으로, 출·퇴근 평균통행시간 DB를 활용하여 분석하도록 설계함
- Light 버전에서는 시도, 시군구별 평균통행시간이 그래프와 단계구분도로 표출되며, Expert 버전에서는 분석지역에 대한 출근시간대별 통행량, 평균통행시간, 1시간 이상 통행비율이 표출됨

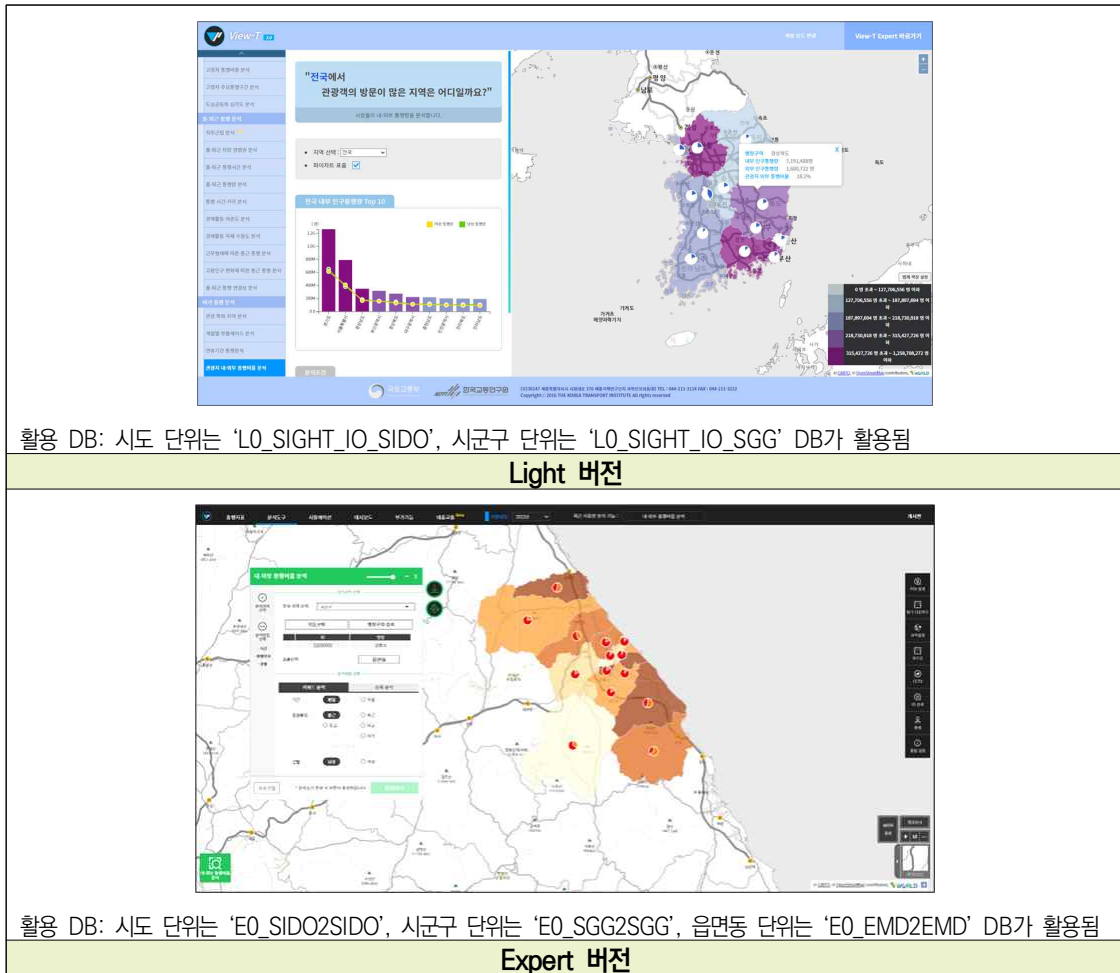


〈그림 4-13〉 출·퇴근 통행시간 분석도구 화면

다. 통행비율 지표를 제공하는 분석도구

○ 내·외부 통행 비율 분석

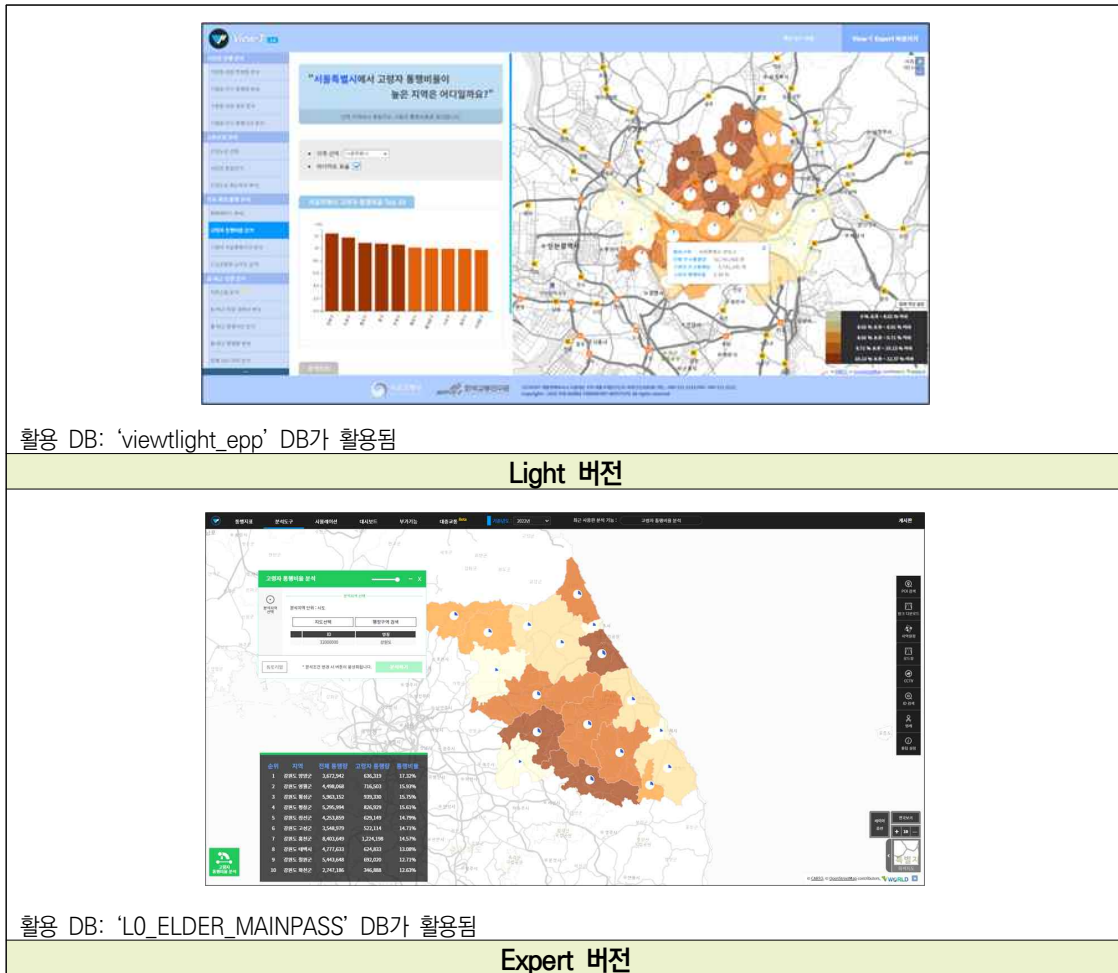
- 지역별 내부 통행 비율과 외부 통행 비율을 분석하는 기능으로, Light 버전의 분석도구는 내·외부 통행 비율 DB(시도, 시군구 2종)를 활용하고, Expert 버전의 분석도구는 발생량·도착량 DB(시도, 시군구, 읍면동 3종)를 활용하여 분석되도록 설계함
- Light 버전에서는 여가, 관광목적의 통행에 대해서만 분석되며, 시도별 내부통행 비율, 외부통행 비율이 산출되어 파이차트 형식으로 지도에 표기되고, 내부통행이 많은 지역은 그래프로 표출됨
- Expert 버전에서는 기간, 통행목적, 통행자 특성(성, 연령)을 선택하여 분석할 수 있으며, 최소 시군구 단위까지 분석할 수 있음



〈그림 4-14〉 내·외부 통행 비율 분석도구 화면

○ 고령자 통행 비중 분석

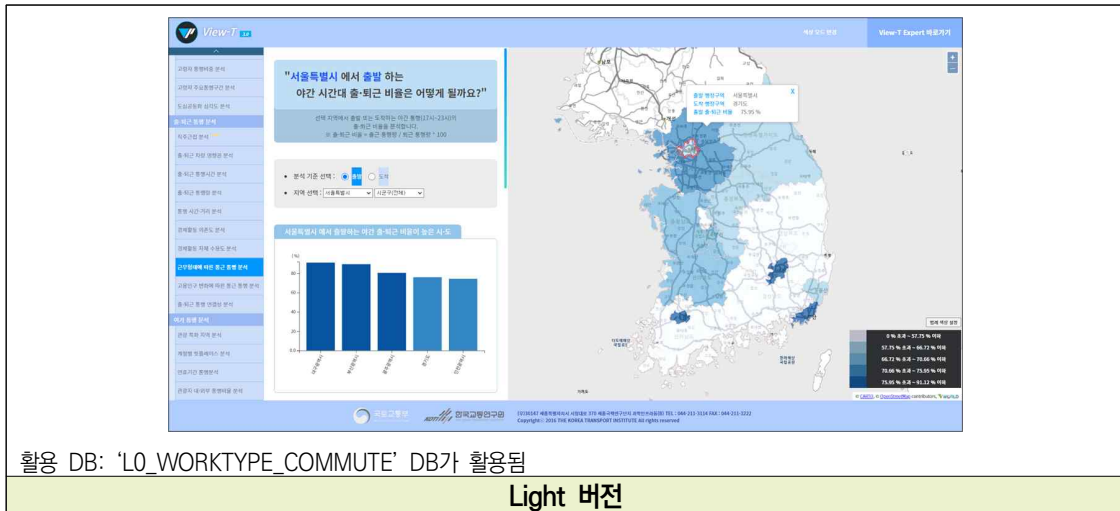
- 지역별 고령자 통행 비율을 분석하는 기능으로, 고령자 통행량 DB를 활용하여 분석되도록 설계함
- 시도 단위로 분석지역을 선택할 수 있으며, 분석지역에 대한 시군구별 고령자 통행 비율이 파이차트 형식으로 지도에 표기됨
- 분석기능은 Light 버전, Expert 버전 모두 동일하나, Expert 버전은 Light 버전보다 표출 방식을 다양하게 설정할 수 있음 (다른 분석 결과와의 중첩, Base map 변경, 범례 조절 등)



〈그림 4-15〉 고령자 통행 비중 분석도구 화면

○ 근무형태에 따른 통근통행 분석(Light 버전 한정)

- 지역별(시군구 단위) 심야시간 출·퇴근 통행 비율을 분석하는 기능으로, 심야 시간 출·퇴근 통행 비율 DB를 활용하여 분석되도록 설계함
- 분석기준을 출발지, 도착지 기준으로 변경할 수 있고, 분석기준에 대한 출·퇴근 통행 비율이 높은 지역이 그래프와 단계구분도로 표출됨 (예: 출발지를 분석기준으로 설정할 경우, 출·퇴근 통행 비율이 높은 도착지가 표출됨)
- 지가지수에 따른 심야시간대 출·퇴근 통행 비율 그래프도 좌측 하단에 표출됨

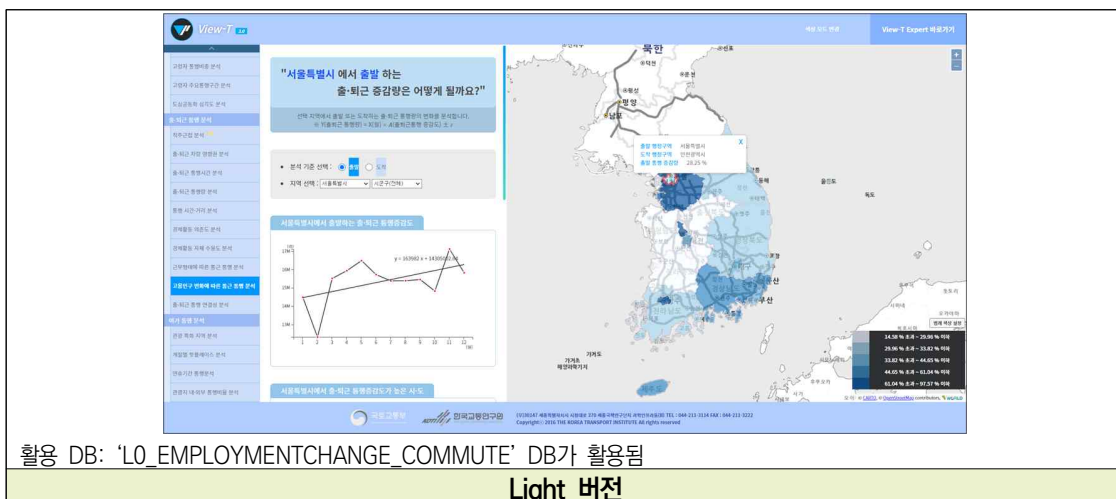


〈그림 4-16〉 근무형태에 따른 통근통행 분석도구 화면

라. 기타 지표를 제공하는 분석도구

○ 고용인구 변화에 따른 통근통행 분석(Light 버전 한정)

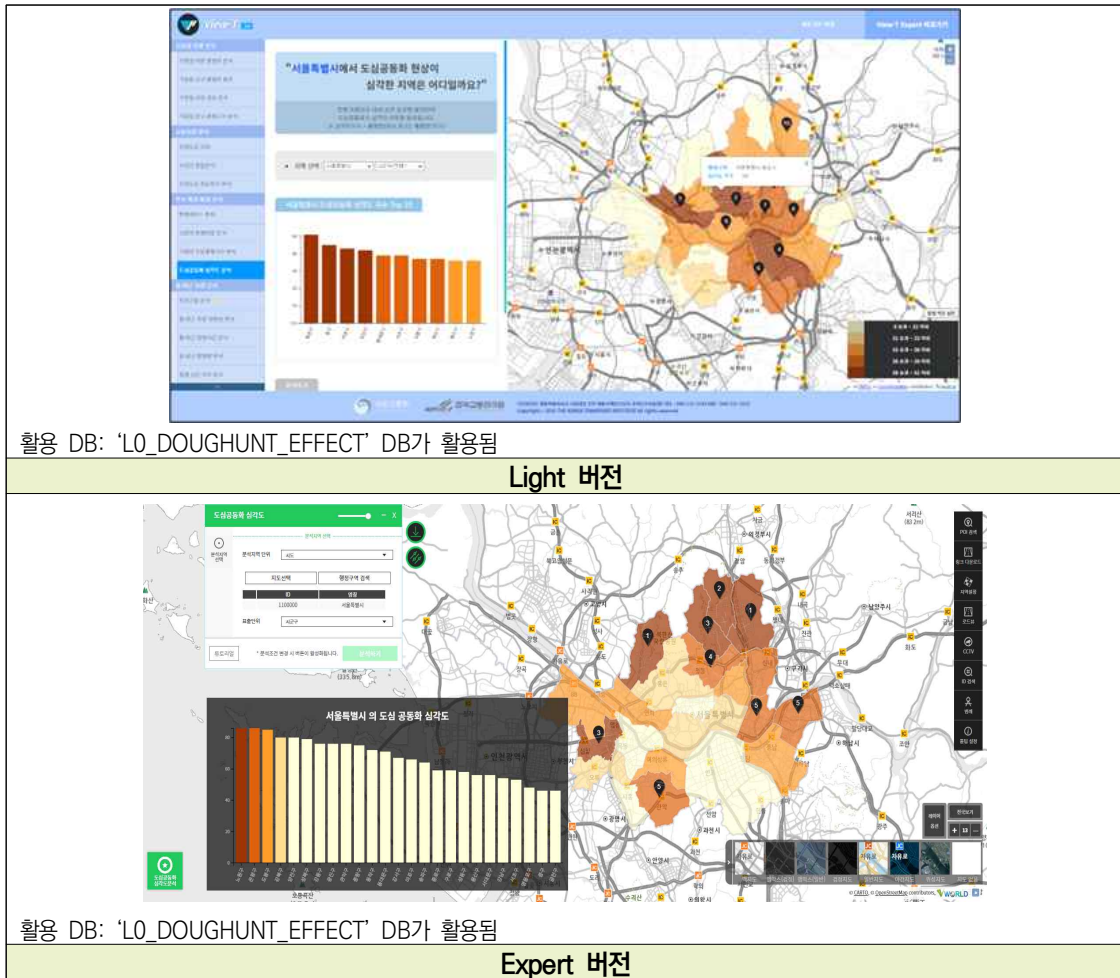
- 지역별 출·퇴근 통행량의 변화를 분석하는 기능으로, 출·퇴근 증감량 DB를 활용하여 분석되도록 설계함
- 전월 대비 기준월의 출·퇴근 통행량의 증가량, 감소량을 측정하여 통행증감도가 높은 지역과 낮은 지역을 그래프와 주제도로 표출함
- 시군구 단위로 분석지역을 선택할 수 있으며, 선택지역에 대한 1년간의 출·퇴근 통행량의 변화를 기준으로 기준월과 출·퇴근 통행량과의 선형관계를 분석하여 $A(\text{통행증감도}) \times (\text{기준월}) \pm \varepsilon$ 의 산식으로 제공함



〈그림 4-17〉 고용인구 변화에 따른 통근통행 분석도구 화면

○ 도심공동화 심각도 분석

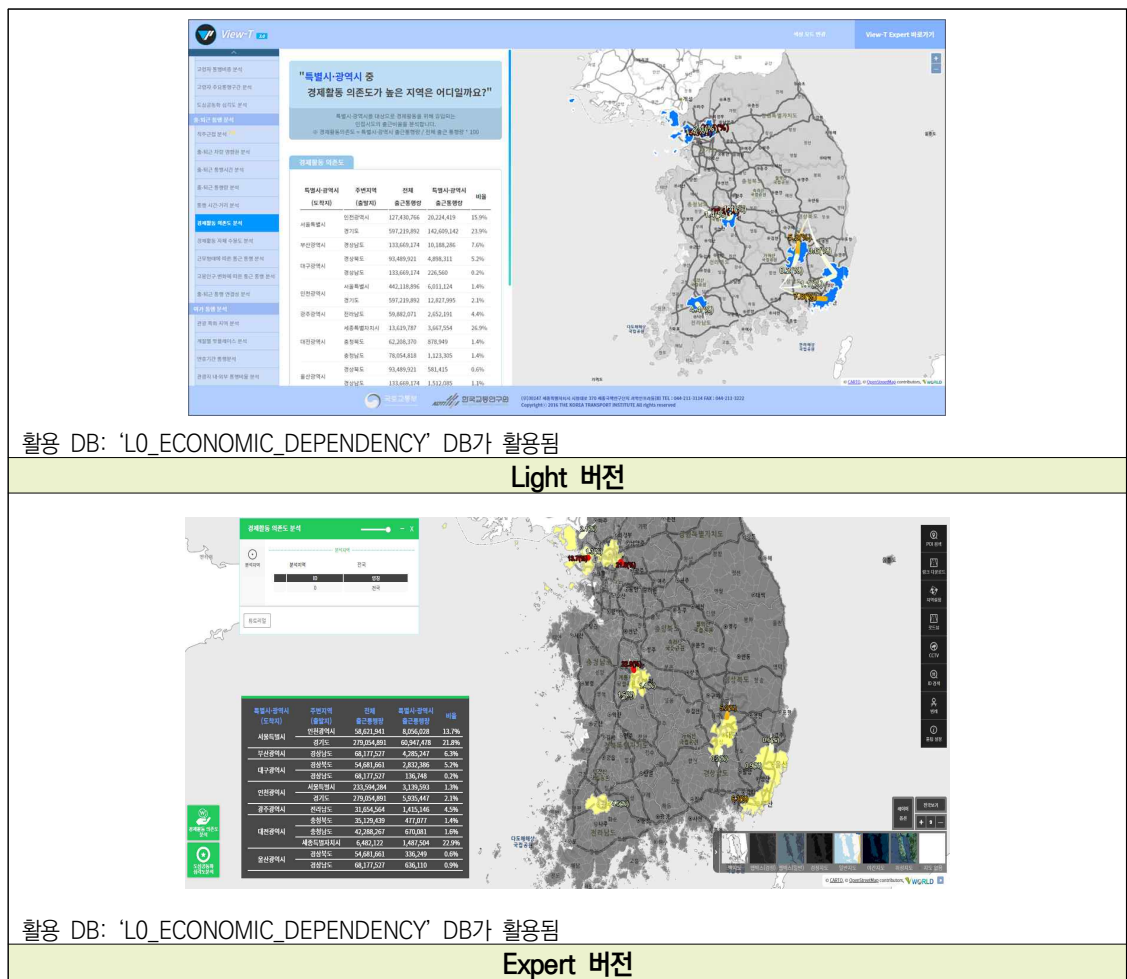
- 도심공동화가 심각한 지역을 분석하는 기능으로, 도심공동화 심각도 DB를 활용하여 분석되도록 설계함
- 최소 시군구 단위로 분석지역을 선택할 수 있으며, 분석지역에 대한 도심공동화 심각도 지수가 단계구분도로 표현됨
- 분석기능은 Light 버전, Expert 버전 모두 동일하나, Expert 버전은 Light 버전보다 표출 방식을 다양하게 설정할 수 있음 (다른 분석 결과와의 중첩, Base map 변경, 범례 조절 등)



〈그림 4-18〉 도심공동화 심각도 분석도구 화면

○ 경제활동의존도 분석

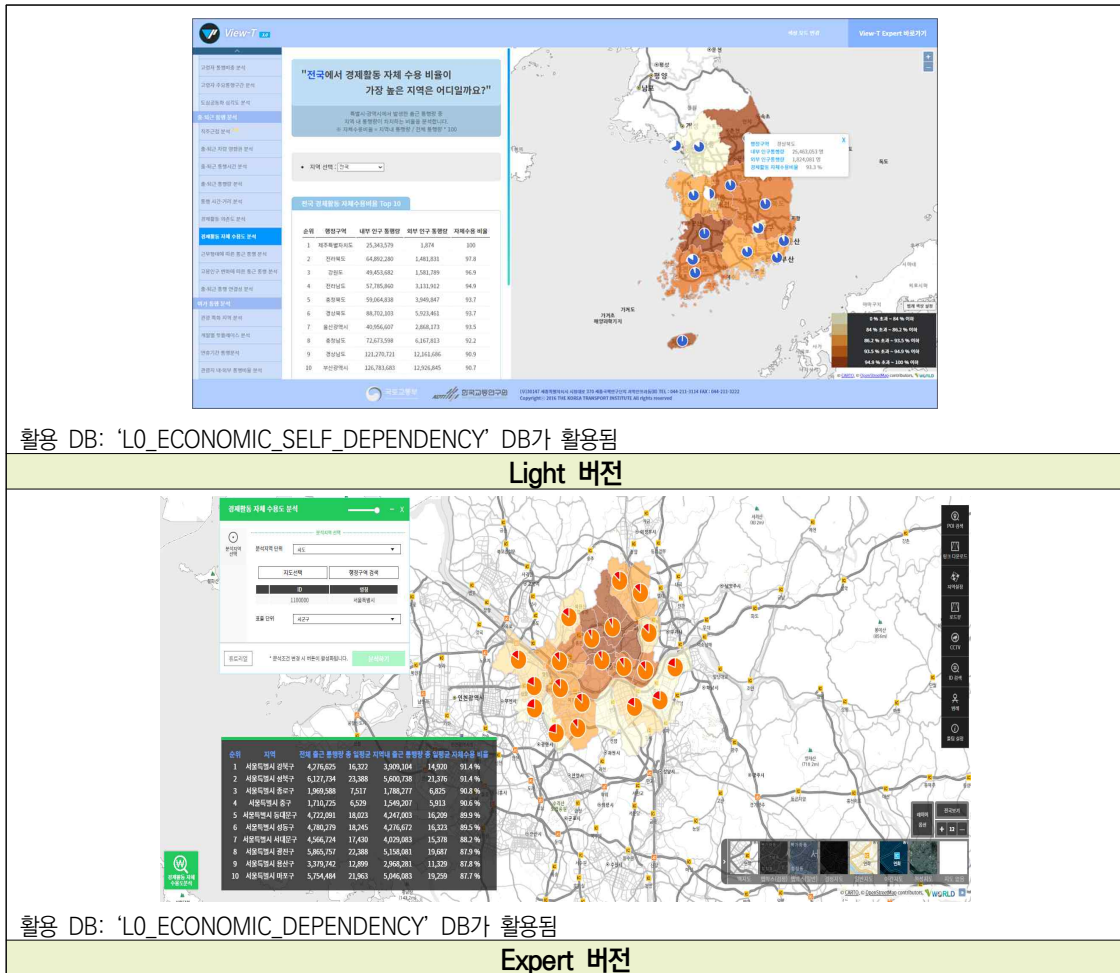
- 특별시·광역시에 대한 경제활동의존도를 분석할 수 있는 기능으로, 경제활동의존도 DB를 활용하여 분석되도록 설계함
- 7개의 특별시·광역시에 대한 경제활동의존도가 지도상에 수치로 표출되며, 특별시·광역시에 의존하고 있는 주변지역이 화살표로 표기됨
- 좌측에는 경제활동의존도 산출에 활용된 통행량 정보가 표로 표출됨
- 분석기능은 Light 버전, Expert 버전 모두 동일하나, Expert 버전은 Light 버전보다 표출 방식을 다양하게 설정할 수 있음 (다른 분석 결과와의 중첩, Base map 변경, 범례 조절 등)



<그림 4-19> 경제활동의존도 분석도구 화면

○ 경제활동자체수용도 분석

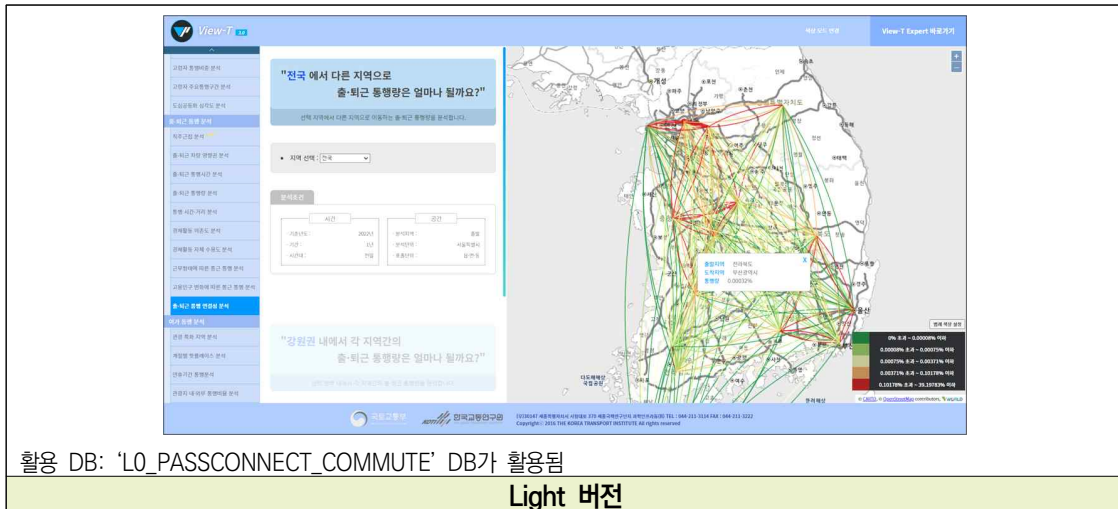
- 지역별 경제활동 자체수용도를 분석할 수 있는 기능으로, 경제활동 자체수용도 DB를 활용하여 분석되도록 설계함
- 시도 단위로 분석지역을 선택할 수 있으며, 분석지역에 대한 시군구별 경제활동자체수용도(통근 통행에 대한 내부통행 비율)가 파이차트 형식으로 지도에 표기됨
- 좌측에는 분석결과(내부 통행량, 외부통행량, 경제활동 자체수용도)가 표로 표출됨
- 분석기능은 Light 버전, Expert 버전 모두 동일하나, Expert 버전은 Light 버전보다 표출 방식을 다양하게 설정할 수 있음 (다른 분석 결과와의 중첩, Base map 변경, 범례 조절 등)



〈그림 4-20〉 경제활동 자체수용도 분석도구 화면

○ 출·퇴근 통행연결성 분석 (Light 버전 한정)

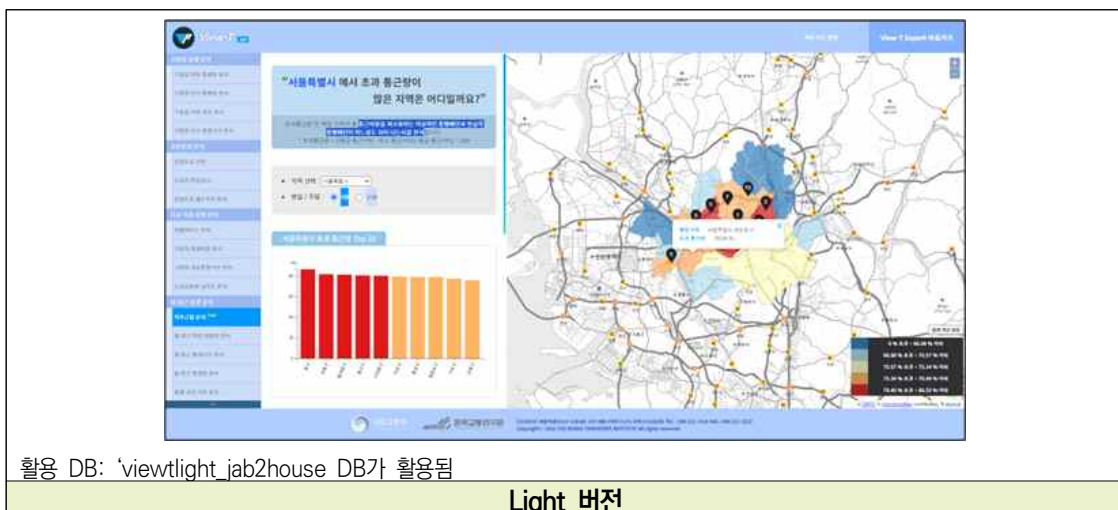
- 기종점간 통근 통행 연결성을 분석하는 기능으로, 출·퇴근 통행 연결성 DB를 활용하여 분석되도록 설계함
- 시도 단위 또는 권역 단위로 분석지역을 선택할 수 있으며, 분석지역 내 출·퇴근 통행 구간(기종점 연결)에 대한 연결성이 우선도로 지도에 표출됨



〈그림 4-21〉 출·퇴근 통행연결성 분석도구 화면

○ 직주 근접 분석 (Light 버전 한정)

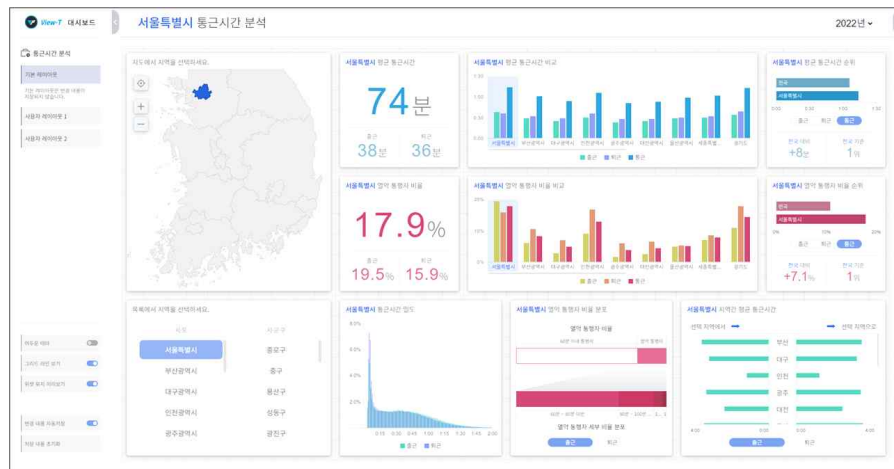
- 통근비용을 최소화하는 이상적인 통행패턴과 현실의 통행패턴이 어느정도 차이 나는지를 분석하는 기능으로, 통근통행 효율성 DB를 활용하여 분석되도록 설계함
- 분석지역(최소 시도 단위)과 분석기준(4월, 10월)을 선택할 수 있으며, 선택조건에 대한 초과통근량을 기준으로 통근 통행 효율성이 떨어지는 지역이 단계구분도로 표출됨
- 좌측에는 통근 통행 효율성이 낮은 곳 10곳에 대한 분석결과(실제통근거리, 최소통근거리, 최대통근거리, 초과통근량, 통근잠재력)가 그래프와 표로 표출됨



〈그림 4-22〉 직주 근접 분석도구 화면

3. 대시보드 및 월간리포트를 통한 교통지표 제공

- 연구, 정책 등에 자주 활용되는 통근시간은 대시보드와 월간리포트로도 정보를 제공함
- 여러 정보를 한 눈에 파악할 수 있는 대시보드로 정보를 제공함으로써, 의사 결정 지원을 도모함
 - 선택지역에 대한 출근, 퇴근, 통근(출근+퇴근)의 평균통행시간, 열악통행자 비율 정보와 함께, 통행시간 밀도, 열악통행자 분포, 순위 정보를 종합적으로 제공함



〈그림 4-23〉 대시보드 화면

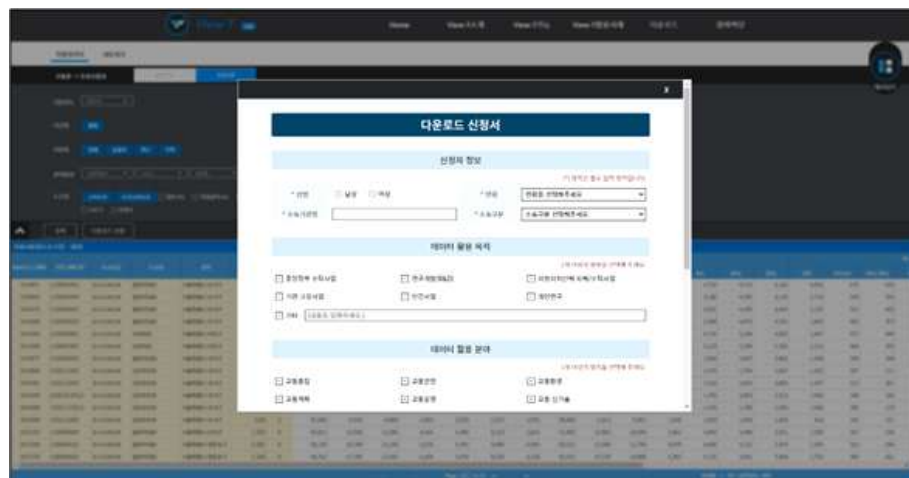
- 대시보드의 정보를 기반으로 전국 시도별 통근시간을 비교하여 월간리포트로 발간함



〈그림 4-24〉 모빌리티 월간리포트 화면

4. 다운로드 기능을 통한 교통지표 제공

- Expert 분석도구 내 구축된 교통지표는 다운받아 사용할 수 있도록 함
 - 기종점 인구 통행량 분석, 기종점 인구 통행구간 분석, 핫플레이스 분석, 고령자 통행비중 분석, 고령자 주요 통행구간 분석, 연휴기간 통행분석, 내·외부 통행 비율 분석에 활용된 교통지표 DB를 제공함



〈그림 4-25〉 다운로드 화면



제5장 결론 및 차년도 수행계획

제1절 결론

제2절 차년도 수행계획

제5장 결론 및 차년도 수행계획

제1절 결론

- 본 사업에서는 모바일 통신데이터 구축 및 활용 목적을 고려하여 통행 DB 구축 형태를 설계하고 통행 DB를 구축하였음
 - 국가교통조사사업의 주요 목적인 여객 O/D 추정 및 지역간·지역내 통행분포 분석을 목적으로 기점 정보, 도착 정보, 통행자 정보를 중심으로 기종점 통행 DB를 설계하고 구축하였음
 - 2022년 11월에 대한 일자별 기종점 통행 DB를 구축함
 - 2023년에 대한 월별 기종점 통행 DB를 구축함
 - 이용자의 분석 및 활용 목적에 따른 데이터 가공 및 재생산이 가능한 체계를 마련하기 위해 개인별 통행궤적 DB를 구축하고 기술적 검토를 수행하였음
 - 2023년 10월 중 8일에 대한 개인별 통행궤적 DB를 구축함
 - 도보-비도보 구분 등 개인별 통행궤적 DB 관련 기술적 검토를 수행함
 - 「개인정보보호법」, 「위치정보의 보호 및 이용 등에 관한 법률」에 저촉되지 않도록 통행 DB를 설계 및 구축하였음
- 또한, 모바일 통신데이터 기반의 통행 DB를 검증하고, 통행량 차이가 나타나는 지역, 구간에 대한 특성을 파악한 다음, 이를 보정할 수 있는 방법에 대하여 기술하였음
 - 모바일 통신데이터로 구축한 통행량이 실제 통행을 잘 반영하고 있는지 검증하기 위하여 조사자료, 사회경제지표, 실적자료와 상호 비교하여 검증하였음

- 모바일 통신데이터는 통신사별 통신설비의 차이, 가공 알고리즘의 차이에 따라 구축되는 통행량의 차이가 크기 때문에, 3종의 모바일 통신데이터(KT, SKT 지라프, SKT 리트머스 데이터)를 수집하여 상호 비교함
- 조사자료 중에서 통행 정보를 담고 있는 통행실태조사(2019년, 2021년)와 인구총조사를 수집하여 모바일 통신데이터와 비교 분석함
- 전체 통행, 통행목적별(통근, 통학, 귀가, 기타) 통행에 대하여 통행 유발, 유인과 영향이 있을 것으로 예상되는 사회경제지표(인구수, 학령인구수 등)와 비교함
- 또한, 조사자료도 추정치이기에 참값으로 볼 수 있는 실적자료를 취득하여 모바일 통신데이터에서 추출한 통행량과 비교함
- 의사결정나무분석을 통해 통행량 차이가 나타나는 구간의 특성을 제시하고, 분석 결과를 토대로 보정기준과 보정 방법을 제시하였음
 - 전체 통행, 통행목적별(통근, 통학, 귀가, 기타)로 데이터간 통행량 차이가 나타나는 유형을 제시하고, 의사결정나무분석 결과에서 나온 절단값을 기준으로 통행량을 보정할 수 있는 방법을 제시하였음
 - 향후, 이러한 보정 대상과 가중치를 이용하여 각 데이터 간 산술평균 또는 가중평균하여 구축된 기준점 통행량의 신뢰성을 개선할 수 있을 것으로 판단함
- 아울러, 2022년 사업에서 구축한 KT 모바일 통신데이터를 활용하여 다양한 교통지표를 구축하고, 온라인 서비스의 분석도구, 대시보드, 월간 리포트 등을 통해 구축한 교통지표를 시각화하여 제공하였음
 - 기존 사업에서 개발한 교통지표와 새롭게 개발한 지표를 포함하여 다양한 교통지표를 구축함 (통행량, 통행시간, 통행거리, 공동화지수, 직주균형지표 등)
 - 일반인을 대상으로 개발한 Light 버전의 분석도구, 전문가를 대상으로 개발한 Expert 버전의 분석도구의 교통지표를 모두 2022년 기준으로 현행화함
 - 사용자, 관리자 편의를 고려하여 데이터 다운로드 기능을 업데이트하였음

제2절 차년도 수행계획

- 시의성 있는 통행 정보를 제공하기 위하여 2024년 기준의 모바일 통신데이터를 구축할 예정이며, 구축 결과에 대한 점검을 통해 이상치 여부를 파악할 예정임
 - 본 사업의 활용 목적, 통신사의 데이터 특성을 고려하여 통행 DB 구조를 설계하고 이에 따라 통행 DB를 생성할 예정임
- 본 사업을 통해 구축한 2023년 기준의 모바일 통신데이터를 기준으로 본 사업에서 진행한 것과 같이 통행량 및 통행패턴에 대한 데이터간 상호 교차 검증을 수행할 예정임
 - 시공간별·목적별 개인통행실태조사, 인구총조사, 3개 통신사 자료, 차량 GPS, 대중교통카드, TCS, 사회경제지표 등의 자료를 대상으로 교차 검증할 예정임
- 모바일 통신데이터를 활용하여 통행지표를 산출 및 현행화할 예정임
 - 개별 객체 단위 모바일 통신데이터를 활용한 통행지표 산출 방안을 검토할 예정임
 - 2023년 모바일 통신데이터를 이용하여 교통지표(통행량, 통행시간 등)를 현행화할 예정임



부 록

제1절 직주균형지표 산출 결과

부 록

제1절 직주균형지표 산출 결과

1. 직주균형지표 개념

- 거주지와 근무지간 교통비용(통행시간, 통행거리)이 최소화될 때 가장 효율적인 통근 통행을 기대할 수 있으나 실제로는 이보다 더 큰 교통비용이 발생함
- 이처럼 최소 비용을 가정하였을 때 나타나는 이상적인 최소통근거리(Minimum Required Commute, MRC)와 실제 통근거리(Actual Required Commute, ARC)간의 차이를 초과통근(Excess commute)이라고 부르며, 이를 통해 직주 균형 정도를 확인할 수 있음

2. 직주 균형 지표 산출 과정

가. 모바일 통신데이터에서 출근 통행 데이터 추출

- 한국교통연구원에서 구축한 KT 모바일 통신데이터 기반의 통행 DB에서 통행자의 출발지 정보가 거주지(H)로 기록되어 있으면서, 도착지 정보가 근무지(C)로 기록되어 있는 데이터를 추출함
 - 시간적 범위: 2022년 4월 11일 ~ 2022년 4월 17일 (일주일)
 - 공간적 범위: 전국, 시군구 단위

〈부표 1〉 KT 모바일 통신데이터 기반의 통행 DB 구조

출발지 정보								도착지 정보								통행자 정보		통행 정보		
일자	요일	시간	격자	읍면동	시군구	시도	채류지 유형	일자	요일	시간	격자	읍면동	시군구	시도	채류지 유형	성별	연령	총 통행량	보정 계수	...
221001	2	10	가나89b14a	...	11220	11	H	221001	2	11	나바66b60b	...	31091	31	C	M	30	6
221004	4	07	바사38b38a	...	37530	37	C	221004	4	09	라마46b43a	...	31104	31	H	M	20	14

나. 지역별 통근량 및 통근거리 산출

- 500m 단위의 격자로 구축되어있는 통행량 정보를 시군구 단위로 재집계함

〈부표 2〉 지역별 통근량 통근거리 산출 예시

통근량	근무지						
거주지	1	2	3	4	5	...	3518
1	62,266	28,259	5,815	6,290	2,110	...	
2	15,713	54,455	8,762	12,395	3,084	...	
3	16,040	34,750	91,843	8,516	3,053	...	
4	22,117	43,352	12,114	123,803	25,083	...	
...	
3518							

- 통근거리는 시군구별 중심점간의 직선거리로 계산하며, 내부통행인 경우에는 해당 지역의 면적을 π 로 나눈 후 루트를 씌우는 방식으로 계산함

〈부표 3〉 지역별 통근거리 산출 예시

통근거리	근무지						
거주지	1	2	3	4	5	...	3518
1	2.76	4.19	7.05	7.44	10.97	...	
2	4.19	1.78	3.49	4.11	8.07	...	
3	7.05	3.49	2.64	5.83	9.50	...	
4	7.44	4.11	5.83	2.31	3.98	...	
...	
3518							

다. 실제통근거리(ARC)와 최소통근거리(MRC) 산출

- 실제통근거리(ARC)는 앞서 산출한 통근량과 통근거리 값을 아래 산식에 적용하여 계산함

$$ARC = \frac{\sum_h \sum_j t_{hj} * d_{hj}}{T}$$

여기서,

t_{hj} : 해당 지역의 통근자 수 (거주지에서 근무지로 이동하는 통행)

d_{hj} : 거주지와 근무지의 거리

T : 전체 통근자 수

- 최소통근거리(MRC)는 선형계획법(Linear Programming, LP)으로 제약조건을 만족하는 가능해들 중에서 아래 목적함수를 최소화하는 값으로 산출함

$$MRC = \min \frac{1}{T} * \left(\sum_h \sum_j t_{hj} * d_{hj} \right)$$

- 제약조건:

$$\sum_j t_{hj} = O_h$$

$$\sum_h t_{hj} = D_j$$

$$t_{hj} \geq 0$$

여기서,

O_h : 해당 지역에 거주하는 전체 통근자의 수

D_j : 해당 지역에 근무하는 통근자의 수

라. 초과통근량(E) 계산

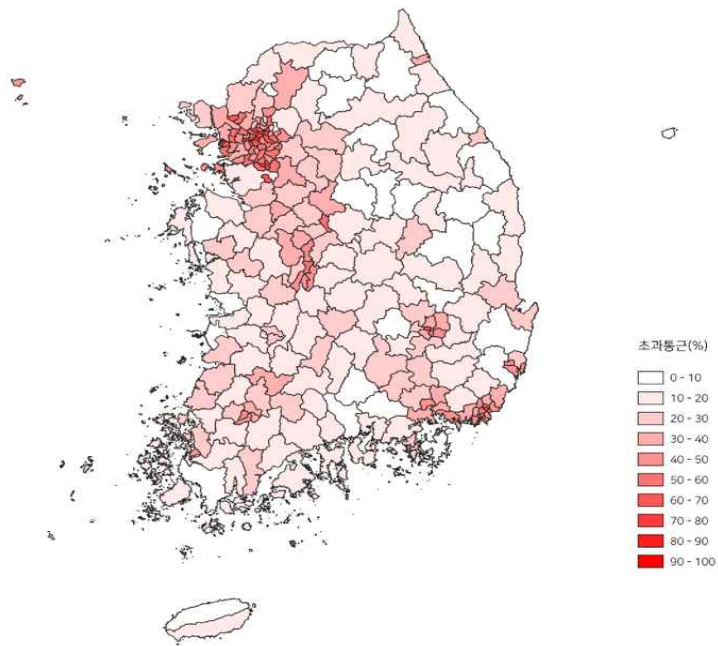
- 초과통근량(E)은 실제통근거리(ARC)와 최소통근거리(MRC)와의 차이로서, 아래 산식에 의해 산출함

$$E = \left(\frac{ARC - MRC}{ARC} \right) 100$$

- 초과통근량은 실제 통행이 통행비용을 최소화했을 때 나타나는 이상적인 통행과 일치할 경우 0%로 나타나며, 100%에 가까운 값일수록 비효율적인 통행이 이루어지고 있다고 해석할 수 있음

3. 직주 균형 지표 산출 결과

- 직주 균형이 맞지 않아 통근이 열악한 지역 중 상위 10%는 대체로 수도권(서울, 경기, 인천)으로 나타남
 - 그 중 서울특별시 강남구, 영등포구, 중구 지역과 같은 경우는 상대적으로 근로자 수가 많은데 지가도 높아, 해당 지역에서 근무하는 사람들은 주변에서 주거를 마련하지 못하고 출퇴근을 위해 많은 시간을 소요하는 것으로 판단됨
 - 서울특별시 강남구 (초과통근량 기준 14위, 지가 순위 2위, 종사자 순위 1위)
 - 서울특별시 영등포구 (초과통근량 기준 16위, 지가 순위 7위, 종사자 순위 4위)
 - 서울특별시 중구 (초과통근량 기준 18위, 지가 순위 1위, 종사자 순위 3위)
- 반면 통근 여건이 상대적으로 좋은 하위 10%는 강원, 충남, 전남, 경북 등 지방 지역으로 나타남
 - 하위 10%에 포함된 대다수의 지역은 종사자 수가 많더라도 지가가 전국 하위 10% 수준으로 형성되어 있어, 해당 지역에서 근무하는 사람들은 직장 근처에서 주거를 마련하여 비교적 가까운 거리에서 통근하는 것으로 사료됨
 - 경상남도 산청군 (초과통근량 기준 226위, 지가 순위 234위, 종사자 순위 19위)
 - 경상북도 영양군 (초과통근량 기준 228위, 지가 순위 243위, 종사자 순위 22위)
 - 강원도 인제군 (초과통근량 기준 244위, 지가 순위 228위, 종사자 순위 4위)



〈부도 1〉 전국 시군구별 초과통근 정도

〈부표 4〉 초과통근량 기준 상위 10%, 하위 10% 지역

상위 10%			하위 10%		
순위	지역	초과통근량(E)	순위	지역	초과통근량(E)
1	인천광역시 동구	82.25	226	경상남도 산청군	9.43
2	경기도 수원시 팔달구	76.21	227	충청북도 충주시	9.41
3	서울특별시 서대문구	75.29	228	경상북도 영양군	9.31
4	서울특별시 동작구	74.56	229	충청북도 제천시	9.17
5	서울특별시 성동구	74.24	230	충청남도 서산시	8.64
6	서울특별시 용산구	72.08	231	강원도 춘천시	8.18
7	서울특별시 구로구	70.51	232	경상북도 안동시	8.12
8	서울특별시 동대문구	70.05	233	경상남도 진주시	8.00
9	서울특별시 금천구	69.98	234	경상북도 경주시	7.62
10	경기도 수원시 영통구	69.79	235	경상남도 거제시	7.62
11	경기도 오산시	69.72	236	강원도 횡성군	7.02
12	부산광역시 중구	69.02	237	울산광역시 울주군	6.89
13	부산광역시 서구	67.32	238	제주특별자치도 제주시	6.71
14	서울특별시 강남구	67.22	239	충청북도 단양군	6.18
15	부산광역시 영도구	67.18	240	충청남도 서천군	5.65
16	서울특별시 영등포구	66.75	241	경상남도 하동군	5.30
17	경기도 고양시 일산서구	65.53	242	경상북도 군위군	5.07
18	서울특별시 중구	65.04	243	강원도 강릉시	5.05
19	대구광역시 서구	65.02	244	강원도 인제군	4.90
20	서울특별시 관악구	62.80	245	경기도 동두천시	4.01
21	인천광역시 미추홀구	62.23	246	경상북도 봉화군	3.35
22	부산광역시 연제구	60.61	247	경상북도 울릉군	2.50
23	충청북도 증평군	59.96	248	경상북도 성주군	2.05
24	경기도 수원시 장안구	59.87	249	강원도 화천군	1.98
25	경기도 수원시 권선구	59.31	250	경기도 구리시	0.00

<부표 5> 시군구별 직주균형지표 결과 (전체)

시군구명	최소통근거리(MRC)	실제통근거리(ARC)	초과통근량(E)
서울 종로구	5.31	12.21	56.49
서울 중구	4.50	12.88	65.04
서울 용산구	3.43	12.27	72.08
서울 성동구	2.65	10.28	74.24
서울 광진구	5.70	9.09	37.25
서울 동대문구	2.54	8.47	70.05
서울 중랑구	5.16	7.05	26.86
서울 성북구	4.90	7.77	36.97
서울 강북구	3.40	6.27	45.75
서울 도봉구	5.21	6.35	17.89
서울 노원구	3.36	7.34	54.18
서울 은평구	5.28	7.23	27.04
서울 서대문구	2.77	11.22	75.29
서울 마포구	5.26	11.00	52.21
서울 양천구	6.32	7.33	13.77
서울 강서구	4.61	9.83	53.14
서울 구로구	2.90	9.84	70.51
서울 금천구	3.13	10.42	69.98
서울 영등포구	3.76	11.32	66.75
서울 동작구	2.28	8.97	74.56
서울 관악구	3.30	8.88	62.80
서울 서초구	7.65	13.53	43.45
서울 강남구	4.49	13.71	67.22
서울 송파구	4.81	10.48	54.13
서울 강동구	4.75	8.32	42.96
부산 중구	2.18	7.04	69.02
부산 서구	2.12	6.49	67.32
부산 동구	4.14	7.24	42.72
부산 영도구	2.16	6.57	67.18
부산 부산진구	3.32	6.69	50.29
부산 동래구	2.30	5.62	59.02
부산 남구	3.38	6.40	47.19
부산 북구	3.56	6.96	48.94
부산 해운대구	4.08	7.30	44.13
부산 사하구	3.89	7.43	47.69
부산 금정구	5.19	8.22	36.90
부산 강서구	8.50	13.43	36.70
부산 연제구	2.26	5.75	60.61
부산 수영구	2.43	5.30	54.14
부산 사상구	5.58	8.07	30.83
부산 기장군	8.72	12.69	31.30
대구 중구	4.31	7.91	45.57
대구 동구	7.60	11.09	31.46
대구 서구	2.35	6.73	65.02
대구 남구	2.97	6.87	56.73
대구 북구	5.46	8.78	37.80
대구 수성구	4.94	8.84	44.10
대구 달서구	4.80	7.44	35.54
대구 달성군	10.13	12.81	20.87

시군구명	최소통근거리(MRC)	실제통근거리(ARC)	초과통근량(E)
인천 중구	11.19	17.50	36.06
인천 동구	1.55	8.72	82.25
인천 연수구	4.71	9.66	51.29
인천 남동구	4.28	7.81	45.16
인천 부평구	3.20	7.36	56.56
인천 계양구	3.81	7.73	50.79
인천 서구	6.50	10.74	39.52
인천 미추홀구	2.82	7.46	62.23
인천 강화군	12.32	16.06	23.26
인천 옹진군	15.99	27.28	41.37
광주 동구	5.99	9.39	36.21
광주 서구	3.95	7.09	44.28
광주 남구	4.39	7.40	40.69
광주 북구	6.16	9.39	34.35
광주 광산구	8.49	11.23	24.38
대전 동구	6.61	11.15	40.68
대전 중구	4.44	8.07	44.90
대전 서구	5.51	9.24	40.31
대전 유성구	7.96	11.11	28.33
대전 대덕구	5.19	11.47	54.71
울산 중구	3.43	6.79	49.49
울산 남구	5.03	8.87	43.29
울산 동구	3.50	6.45	45.81
울산 북구	7.19	9.74	26.15
울산 울주군	16.63	17.86	6.89
세종 세종시	12.16	17.58	30.80
경기 수원시 장안구	3.26	8.11	59.87
경기 수원시 권선구	3.87	9.52	59.31
경기 수원시 팔달구	2.02	8.50	76.21
경기 수원시 영통구	2.98	9.85	69.79
경기 성남시 수정구	5.82	11.31	48.60
경기 성남시 중원구	4.34	10.32	58.00
경기 성남시 분당구	5.20	12.05	56.83
경기 의정부시	5.10	9.04	43.56
경기 안양시 만안구	3.41	8.20	58.42
경기 안양시 동안구	5.67	8.84	35.81
경기 부천시	4.54	7.57	40.07
경기 광명시	4.49	8.64	48.02
경기 평택시	12.37	16.93	26.92
경기 동두천시	10.94	11.40	4.01
경기 안산시 상록구	4.31	9.87	56.35
경기 안산시 단원구	8.20	15.13	45.81
경기 고양시 덕양구	7.31	10.82	32.47
경기 고양시 일산동구	4.89	10.20	52.05
경기 고양시 일산서구	3.63	10.52	65.53
경기 과천시	5.59	11.87	52.91
경기 구리시	12.16	9.53	-27.54
경기 남양주시	12.09	14.81	18.35
경기 오산시	3.68	12.16	69.72
경기 시흥시	6.74	10.70	36.97
경기 군포시	4.71	8.72	45.98

시군구명	최소통근거리(MRC)	실제통근거리(ARC)	초과통근량(E)
경기 의왕시	4.17	9.84	57.64
경기 하남시	5.43	10.62	48.82
경기 용인시 처인구	12.57	18.70	32.79
경기 용인시 기흥구	5.10	12.12	57.93
경기 용인시 수지구	5.33	9.04	41.03
경기 파주시	14.81	19.37	23.55
경기 이천시	14.30	20.00	28.52
경기 안성시	14.32	20.10	28.74
경기 김포시	9.45	14.65	35.49
경기 화성시	15.82	19.69	19.62
경기 광주시	11.69	16.57	29.43
경기 양주시	9.95	13.60	26.83
경기 포천시	16.23	24.38	33.45
경기 여주시	14.95	19.75	24.29
경기 연천군	17.53	21.61	18.86
경기 가평군	17.58	21.75	19.17
경기 양평군	16.70	23.12	27.78
강원 춘천시	18.84	20.51	8.18
강원 원주시	16.61	18.47	10.11
강원 강릉시	18.20	19.17	5.05
강원 동해시	7.60	10.11	24.86
강원 태백시	9.82	11.38	13.69
강원 속초시	5.80	8.66	32.95
강원 삼척시	19.42	21.46	9.49
강원 홍천군	25.51	28.97	11.96
강원 횡성군	20.65	22.21	7.02
강원 영월군	19.75	24.19	18.36
강원 평창군	23.62	26.35	10.34
강원 정선군	21.87	26.27	16.74
강원 철원군	16.83	19.79	14.97
강원 화천군	20.14	20.55	1.98
강원 양구군	15.91	18.33	13.21
강원 인제군	23.97	25.20	4.90
강원 고성군	16.58	19.15	13.45
강원 양양군	16.51	19.04	13.30
충북 충주시	17.67	19.51	9.41
충북 제천시	16.76	18.45	9.17
충북 청주시 상당구	11.33	14.48	21.79
충북 청주시 서원구	6.26	12.24	48.89
충북 청주시 흥덕구	8.52	13.17	35.34
충북 청주시 청원구	10.27	13.62	24.58
충북 보은군	15.23	17.87	14.79
충북 옥천군	14.44	17.58	17.90
충북 영동군	17.06	20.46	16.60
충북 진천군	14.11	18.80	24.96
충북 괴산군	19.96	23.03	13.33
충북 음성군	15.63	22.73	31.24
충북 단양군	18.37	19.58	6.18
충북 증평군	5.10	12.75	59.96
충남 천안시 동남구	11.81	16.83	29.81
충남 천안시 서북구	7.94	13.22	39.96

시군구명	최소통근거리(MRC)	실제통근거리(ARC)	초과통근량(E)
충남 공주시	16.98	20.72	18.04
충남 보령시	13.64	15.64	12.82
충남 아산시	14.24	18.20	21.77
충남 서산시	15.45	16.91	8.64
충남 논산시	13.25	16.76	20.94
충남 계룡시	7.02	9.71	27.74
충남 당진시	16.05	18.13	11.45
충남 금산군	15.48	18.70	17.21
충남 부여군	15.25	18.26	16.49
충남 서천군	13.17	13.96	5.65
충남 청양군	13.87	18.06	23.17
충남 홍성군	11.88	16.21	26.68
충남 예산군	14.04	17.23	18.51
충남 태안군	13.64	15.24	10.51
전북 전주시 완산구	5.42	7.67	29.37
전북 전주시 덕진구	6.10	8.31	26.63
전북 군산시	11.56	14.07	17.81
전북 익산시	12.92	15.37	15.98
전북 정읍시	15.67	18.36	14.68
전북 남원시	15.79	19.36	18.41
전북 김제시	15.69	18.04	13.04
전북 완주군	13.97	16.67	16.17
전북 진안군	18.87	21.71	13.10
전북 무주군	15.08	19.02	20.70
전북 장수군	15.01	20.10	25.32
전북 임실군	18.05	21.32	15.34
전북 순창군	13.56	20.37	33.40
전북 고창군	14.57	19.08	23.65
전북 부안군	15.08	18.96	20.49
전남 목포시	4.15	7.80	46.84
전남 여수시	14.60	16.30	10.45
전남 순천시	17.03	19.12	10.95
전남 나주시	14.83	17.43	14.89
전남 광양시	13.33	16.41	18.79
전남 담양군	12.28	15.64	21.50
전남 곡성군	17.81	23.17	23.12
전남 구례군	14.12	17.17	17.77
전남 고흥군	17.21	20.47	15.94
전남 보성군	17.66	20.92	15.57
전남 화순군	16.11	18.44	12.63
전남 장흥군	14.14	18.38	23.04
전남 강진군	14.63	18.00	18.75
전남 해남군	19.57	21.63	9.51
전남 영암군	18.08	20.09	10.04
전남 무안군	12.14	17.26	29.68
전남 함평군	16.00	18.69	14.41
전남 영광군	13.40	17.76	24.58
전남 장성군	15.26	17.80	14.31
전남 완도군	11.80	13.85	14.80
전남 진도군	12.86	15.42	16.64
전남 신안군	18.78	22.25	15.59

시군구명	최소통근거리(MRC)	실제통근거리(ARC)	초과통근량(E)
경북 포항시 남구	14.04	18.57	24.37
경북 포항시 북구	15.31	19.26	20.51
경북 경주시	21.73	23.52	7.62
경북 김천시	18.20	21.67	16.02
경북 안동시	22.04	23.99	8.12
경북 구미시	14.02	17.10	17.98
경북 영주시	14.63	16.48	11.21
경북 영천시	18.66	21.32	12.48
경북 상주시	20.54	23.38	12.15
경북 문경시	17.08	20.24	15.62
경북 경산시	11.72	14.31	18.06
경북 군위군	21.06	22.19	5.07
경북 의성군	19.78	24.35	18.77
경북 청송군	18.25	22.50	18.87
경북 영양군	18.51	20.41	9.31
경북 영덕군	17.95	20.71	13.34
경북 청도군	15.79	19.10	17.34
경북 고령군	15.54	19.30	19.47
경북 성주군	21.91	22.37	2.05
경북 칠곡군	12.71	17.65	28.00
경북 예천군	14.54	20.38	28.66
경북 봉화군	22.82	23.61	3.35
경북 울진군	17.93	20.19	11.20
경북 울릉군	4.85	4.97	2.50
경남 진주시	15.06	16.37	8.00
경남 통영시	8.90	11.65	23.56
경남 사천시	12.58	14.95	15.87
경남 김해시	12.21	14.71	17.01
경남 밀양시	17.15	20.03	14.38
경남 거제시	11.69	12.66	7.62
경남 양산시	13.08	15.99	18.19
경남 창원시 의창구	8.06	13.28	39.30
경남 창원시 성산구	7.03	10.65	34.00
경남 창원시 마산합포구	8.79	12.90	31.85
경남 창원시 마산회원구	5.38	10.60	49.28
경남 창원시 진해구	6.36	11.22	43.29
경남 의령군	14.95	19.48	23.27
경남 함안군	11.54	16.18	28.67
경남 창녕군	15.84	20.30	21.96
경남 고성군	15.35	18.90	18.81
경남 남해군	12.09	14.71	17.77
경남 하동군	19.25	20.33	5.30
경남 산청군	19.50	21.53	9.43
경남 함양군	15.70	18.96	17.19
경남 거창군	15.98	18.32	12.77
경남 합천군	18.47	23.18	20.30
제주 제주시	16.44	17.62	6.71
제주 서귀포시	14.02	16.24	13.68

9

모바일 교통 빅데이터 구축 및 활용

2023년 국가교통조사 및 분석

