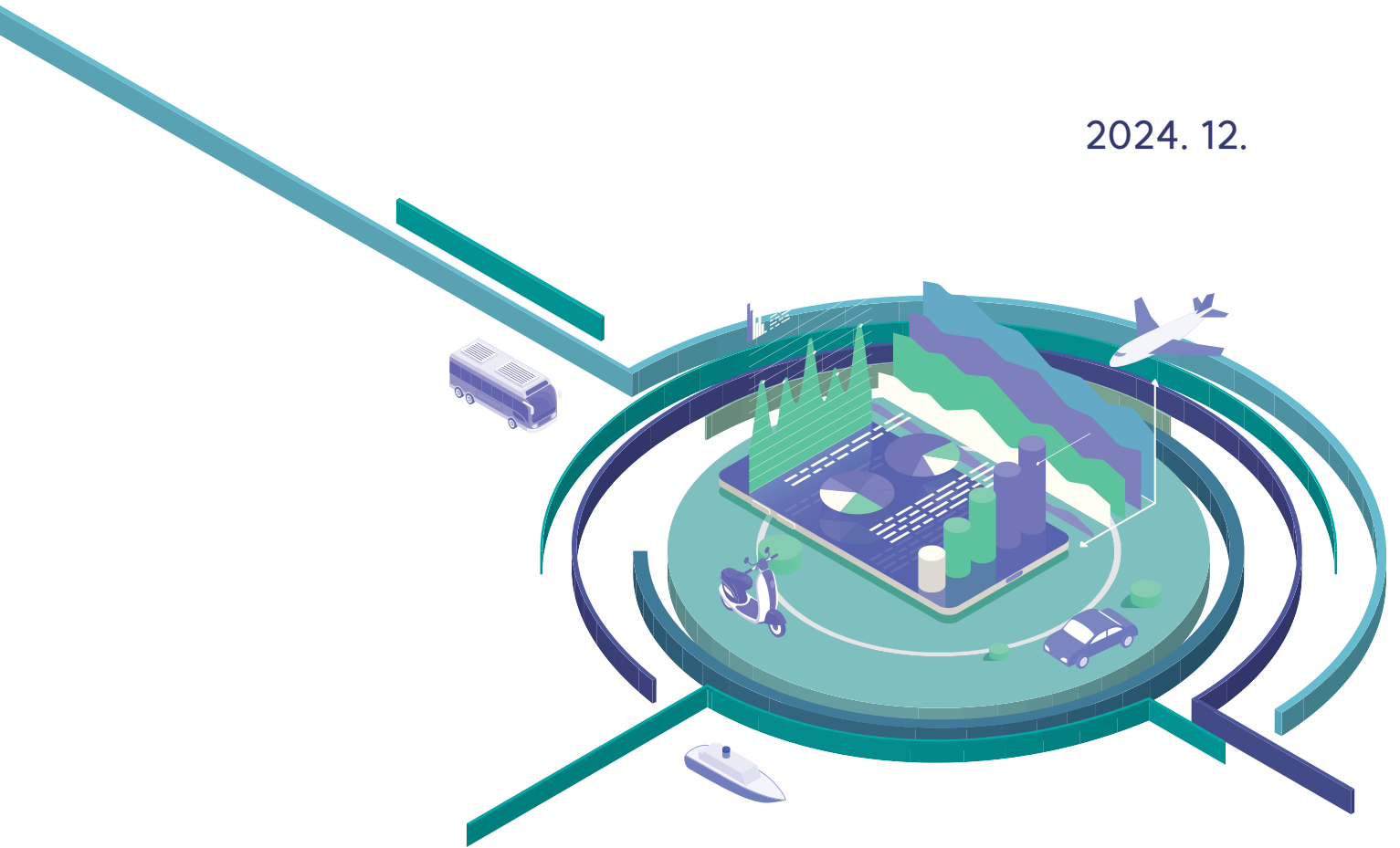


2024년 국가교통조사 및 분석

# 모바일 통신 빅데이터 구축

11

2024. 12.



국토교통부



한국교통연구원  
KOREA TRANSPORT INSTITUTE



# 제 출 문

국토교통부 장관 귀하

본 보고서를 “2024년 국가교통조사 및 분석”의  
최종보고서로 제출합니다.

2024년 12월

한국교통연구원

원장 김 영 찬

본 『2024년 국가교통조사 및 분석』은 다음 연구진에 의해  
수행되었습니다.

### 참여 연구진

<한국교통연구원>	
연구책임자	◦ 김주영 선임연구위원
연구진	◦ 조범철, 김주영 선임연구위원 ◦ 천승훈, 박용일 연구위원 ◦ 황순연, 장동익, 원민수, 이송봉, 이종우 부연구위원 ◦ 김동호, 신영권 책임전문원 ◦ 김규진 주임전문원 ◦ 가보연, 강국수, 광명신, 권순옥, 김관용, 김수아, 김승주, 김호용, 박미란, 박준호, 오연선, 이동엽, 이새봄, 이재영, 이해선, 정정호, 채운혁, 홍성표 연구원 ◦ 조지윤 연구조원
<한국해양수산개발원>	
연구진	◦ 이호춘, 최건우, 황수진 부연구위원 ◦ 류희영 전문연구원 ◦ 박일란 선임사무원
<한국항공협회>	
연구진	◦ 성인영 책임연구원 ◦ 최인영, 김지한, 김창욱, 김진성, 박다영 연구원

# 『2024년 국가교통조사 및 분석』

## 보고서 구성 및 담당연구진

번호	과제명	연구진
제 1권	요약보고서	김주영, 신영권, 가보연
제 2권	전국 여객 기종점통행량 보완갱신	김동호, 김관용, 강국수, 박미란
제 3권	주말 및 침두·비침두 기종점통행량 시범 구축	김동호, 채운혁
제 4권	교통분석용 네트워크 구축	김동호, 김관용, 이동엽, 이새봄
제 5권	항공여객 O/D 조사	한국항공협회
제 6권	전국화물 기종점통행량 보완갱신	원민수, 오연선, 이해선
제 7권	물류거점 화물실태조사	황순연, 김승주, 김호용
제 8권	연안화물 기종점통행량 구축	한국해양수산개발원
제 9권	KTDB 모빌리티 기반지도 구축	천승훈, 정정호
제10권	차량 GPS 빅데이터 구축	이승봉, 이채영
제11권	모바일통신 빅데이터 구축	조범철, 이종우, 곽명신
제12권	국가교통통계DB구축	박용일, 김수아
제13권	특별교통대책기간 통행실태조사	김주영, 김관용, 권순욱
제14권	교통접근성지표 구축	장동익, 박준호

**『2024년 국가교통조사 및 분석』**  
**과제별 공동참여·위탁용역 사업자**

**【공동사업 참여기관】**

- 전국 여객O/D 현행화 공동사업(수도권 부문)
  - 경기연구원, 인천연구원, 서울연구원
  
- 항공O/D 및 특성 조사
  - (사)한국항공협회

**【위탁용역 사업자】**

- 전국여객 O/D 현행화(대구광역권)
  - 홍익대학교 산학협력단
  
- 전국여객 O/D 현행화(대전세종충청권)
  - 신명이앤씨(주)
  
- 전국여객 O/D 현행화(제주권)
  - 홍익대학교 산학협력단
  
- 주말 및 첨두·비첨두 O/D 기초데이터 구축 및 분석
  - (주)모비크리에티브
  
- 교통부문 네트워크 갱신을 위한 GIS 기반 교통망 기초자료 구축
  - 서울시립대학교 산학협력단
  
- 물류거점화물실태조사
  - (주)코리아데이터네트워크

## 【위탁용역 사업자】

- 화물 기종점 통행량 보완·갱신을 위한 교통물류 실증 데이터 수집·가공·전처리 용역  
- ㈜노트스퀘어
- 모빌리티 빅데이터를 활용한 KTDB 기반지도 및 차량통행 데이터셋 구축  
- ㈜큐빅웨어, ㈜유아이네트웍스
- 특별교통대책기간 통행실태조사  
- (주)컨슈머인사이트
- 대중교통 GTFS기반 네트워크 구축  
- ㈜아로정보기술
- Open-source 기반의 교통접근성 산정 최적화 방법론 연구  
- (주)아로정보기술

# 최종보고서 목차

- 제 1권 요약보고서
- 제 2권 전국 여객 기종점통행량 보완갱신
- 제 3권 주말 및 첨두·비첨두 기종점통행량 시범 구축
- 제 4권 교통분석용 네트워크 구축
- 제 5권 항공여객 O/D 조사
- 제 6권 전국화물 기종점통행량 보완갱신
- 제 7권 물류거점 화물실태조사
- 제 8권 연안화물 기종점통행량 구축
- 제 9권 KTDB 모빌리티 기반지도 구축
- 제 10권 차량 GPS 빅데이터 구축
- 제 11권 모바일통신 빅데이터 구축
- 제 12권 국가교통통계DB구축
- 제 13권 특별교통대책기간 통행실태조사
- 제 14권 교통접근성지표 구축

# • 목 차

## 요약

<b>제1장 사업의 개요</b> .....	<b>3</b>
<b>제1절 사업의 배경 및 목적</b> .....	<b>3</b>
1. 사업의 배경 .....	3
2. 사업의 목적 .....	4
<b>제2절 사업의 범위 및 내용</b> .....	<b>5</b>
1. 사업의 범위 .....	5
2. 사업의 내용 .....	6
<b>제2장 모바일 통신데이터 기반 통행 DB 구축 및 결과 점검</b> .....	<b>9</b>
<b>제1절 모바일 통신데이터 기반 통행 DB 구축 개요</b> .....	<b>9</b>
1. 모바일 통신데이터 활용 목적을 고려한 통행 DB 구축 형태 검토 .....	9
2. 통행 DB 구축 형태 설계 .....	11
3. 통행 DB 구축 및 검토 .....	11
<b>제2절 모바일 통신데이터 기반 기종점 통행 DB 구축</b> .....	<b>12</b>
1. 모바일 통신데이터 기반 기종점 통행 DB 형태 설계 .....	12
2. 모바일 통신데이터 기반 통행DB 구축 방법 .....	17
3. 모바일 통신데이터 기반 통행DB 전수화 방안 .....	26
4. 모바일 통신데이터 기반 통행DB 구축 결과 .....	29
<b>제3절 객체단위 모바일 통신데이터 가공 및 통행 DB 구축</b> .....	<b>53</b>
1. 개요 .....	53
2. 객체 단위 모바일 통신데이터 기초분석 .....	56
3. 객체 단위 모바일 통신데이터 오류 제거 및 전처리 .....	59
4. 통행궤적 및 통행정보 데이터 구축 .....	64
5. 객체 단위 모바일 통신데이터 구축 및 활용 관련 검토사항 .....	68

<b>제3장 모바일 통신데이터 기반 통행 DB 검증</b>	<b>73</b>
<b>제1절 개요</b>	<b>73</b>
1. 기본방향	73
2. 검증데이터 수집 및 가공	76
<b>제2절 데이터 간 상호교차검증 결과</b>	<b>84</b>
1. 총통행량 & 통행 비율	84
2. 출근 통행량 & 통행 비율	85
3. 업무 통행량 & 통행 비율	87
4. 귀가 통행량 & 통행 비율	88
<b>제4장 모바일 통신데이터 기반 교통지표 구축</b>	<b>93</b>
<b>제1절 모바일 통신데이터 기반 교통지표 구축</b>	<b>93</b>
1. 개요	93
2. 교통지표 구축	96
<b>제5장 결론 및 차년도 수행계획</b>	<b>121</b>
<b>제1절 결론</b>	<b>121</b>
<b>제2절 차년도 수행계획</b>	<b>122</b>

## ● 표목차

〈표 2-1〉 모바일통신 빅데이터 DB 형태 (전수화 미적용 통행 정보) .....	14
〈표 2-2〉 모바일통신 빅데이터 DB 테이블 정의서 (전수화 미적용) .....	14
〈표 2-3〉 모바일통신 빅데이터 DB 테이블 정의서 (전수화 적용) .....	15
〈표 2-4〉 체류지 식별 기준 .....	23
〈표 2-5〉 인구수 기반 전수화 방법 .....	26
〈표 2-6〉 지역별 총 통행량 .....	30
〈표 2-7〉 시간대별 출근 통행량 (요일별) .....	31
〈표 2-8〉 시간대별 퇴근 통행량 (요일별) .....	33
〈표 2-9〉 시간대별 귀가 통행량 (요일별) .....	35
〈표 2-10〉 시간대별 출근 통행량 (연령대별) .....	37
〈표 2-11〉 시간대별 퇴근 통행량 (연령대별) .....	39
〈표 2-12〉 시간대별 귀가 통행량 (연령대별) .....	41
〈표 2-13〉 시간대별 출근 통행량 (지역별) - 특별시, 광역시, 특별자치시 .....	43
〈표 2-14〉 시간대별 출근 통행량 (지역별) - 도, 특별자치도(계속) .....	44
〈표 2-15〉 시간대별 퇴근 통행량 (지역별) - 특별시, 광역시, 특별자치시 .....	46
〈표 2-16〉 시간대별 퇴근 통행량 (지역별) - 도, 특별자치도(계속) .....	47
〈표 2-17〉 시간대별 귀가 통행량 (지역별) - 특별시, 광역시, 특별자치시 .....	49
〈표 2-18〉 시간대별 귀가 통행량 (지역별) - 도, 특별자치도(계속) .....	50
〈표 2-19〉 지역별 통행시간 (단위: 분) .....	52
〈표 2-20〉 LG유플러스 이동통신 가명정보 데이터 테이블정의서 .....	56
〈표 2-21〉 LG 유플러스 모바일 통신 격자&공간정보 데이터 테이블정의서 .....	57
〈표 2-22〉 LG 유플러스 모바일 통신 전처리 데이터 .....	60
〈표 2-23〉 LG유플러스 모바일 통신 원시데이터 (트립 ID : 720) .....	62
〈표 2-24〉 LG유플러스 모바일 통신 원시데이터 수단 구분 (트립 ID : 720) .....	63
〈표 2-25〉 LG 유플러스 모바일 통신데이터 기반 통행정보DB .....	65
〈표 2-26〉 통행ID별 통행궤적 데이터 구축 결과(트립 ID : 720) .....	67
〈표 3-1〉 LG 데이터 정의서 .....	77
〈표 3-2〉 2023년 11월 시도별 통행량(LG) .....	78
〈표 3-3〉 KT 데이터 정의서 .....	79
〈표 3-4〉 2023년 11월 시도별 통행량(KT) .....	80
〈표 3-5〉 SKT 데이터 정의서 .....	81
〈표 3-6〉 2023년 11월 시도별 통행량(SKT) .....	82
〈표 4-1〉 모바일 통신데이터 기반 교통지표 .....	94
〈표 4-2〉 읍면동-읍면동 단위 통행량 컬렉션의 구성 .....	96
〈표 4-3〉 시군구-시군구 단위 통행량 컬렉션의 구성 .....	97

〈표 4-4〉 시도-시도 단위 통행량 컬렉션의 구성 .....	97
〈표 4-5〉 연휴기간의 통행량 컬렉션의 구성 .....	98
〈표 4-6〉 계절별 통행량 컬렉션의 구성 .....	99
〈표 4-7〉 계절별 기타통행량(도착량) 컬렉션의 구성 .....	100
〈표 4-8〉 통행목적 구분 기준 .....	101
〈표 4-9〉 통행목적별 통행량 컬렉션의 구성 .....	102
〈표 4-10〉 고령자 통행량(발생량 기준) 컬렉션의 구성 .....	103
〈표 4-11〉 고령자 통행량(OD통행량 기준) 컬렉션의 구성 .....	103
〈표 4-12〉 출근 통행시간 컬렉션의 구성 .....	104
〈표 4-13〉 퇴근 통행시간 컬렉션의 구성 .....	105
〈표 4-14〉 내·외부 통행(시군구) 컬렉션의 구성 .....	107
〈표 4-15〉 내·외부 통행(시도) 컬렉션의 구성 .....	107
〈표 4-16〉 심야시간 출퇴근 통행 비율 컬렉션의 구성 .....	109
〈표 4-17〉 월별 통근 통행량 컬렉션의 구성 .....	110
〈표 4-18〉 도심공동화 심각도 컬렉션의 구성 .....	111
〈표 4-19〉 경제활동의존도 컬렉션의 구성 .....	112
〈표 4-20〉 경제활동 자체수용도 컬렉션의 구성 .....	114
〈표 4-21〉 출·퇴근 통행연결성 컬렉션의 구성 .....	116
〈표 4-22〉 초과통근량 컬렉션의 구성 .....	118

## • 그림목차

〈그림 2-1〉 선분 이력으로의 변환	20
〈그림 2-2〉 체류순서 부여 (개인별 이동궤적 형성)	21
〈그림 2-3〉 이동정보와 체류정보 구분 (예시)	22
〈그림 2-4〉 체류지 유형 구분 (예시)	23
〈그림 2-5〉 위치정보, 시간정보, 연령정보 변환 (예시)	24
〈그림 2-6〉 출발/도착 구분 및 통행량 집계 (예시)	25
〈그림 2-7〉 시장점유율 기반 전수화 방법	26
〈그림 2-8〉 데이터 융합	27
〈그림 2-9〉 전수화 과정 도식화	28
〈그림 2-10〉 지역별 총 통행량	29
〈그림 2-11〉 시간대별 출근 통행량 (요일별)	30
〈그림 2-12〉 시간대별 퇴근 통행량 (요일별)	32
〈그림 2-13〉 시간대별 귀가 통행량 (요일별)	34
〈그림 2-14〉 시간대별 출근 통행량 (연령대별)	36
〈그림 2-15〉 시간대별 퇴근 통행량 (연령대별)	38
〈그림 2-16〉 시간대별 귀가 통행량 (연령대별)	40
〈그림 2-17〉 시간대별 출근 통행량 (지역별)	42
〈그림 2-18〉 시간대별 퇴근 통행량 (지역별)	45
〈그림 2-19〉 시간대별 귀가 통행량 (지역별)	48
〈그림 2-20〉 지역별 통행시간 (단위: 분)	51
〈그림 2-21〉 LG 통신 데이터 격자 전국 예시	57
〈그림 2-22〉 서울 → 부산 50m 격자 단위 이동 궤적(2024-10-14(월) 기준)	58
〈그림 2-23〉 서울 → 부산 50m 격자 단위 이동 궤적(2024-10-19(토) 기준)	59
〈그림 2-24〉 시간대별 통행 이동속도(트립 ID : 720)	67
〈그림 3-1〉 상호 검증 방식 예시 - 조사자료 및 통신사별 비교	74
〈그림 3-2〉 데이터 신뢰성 검증 및 보정 절차	75
〈그림 3-3〉 시도 단위 전체 통행의 상호교차검증 결과 (발생량) (단위: 통행/일)	84
〈그림 3-4〉 시도 단위 전체 통행의 상호교차검증 결과 (발생량) (단위: %)	85
〈그림 3-5〉 시도 단위 출근 통행의 상호교차검증 결과 (발생량) (단위: 통행/일)	86
〈그림 3-6〉 시도 단위 출근 통행의 상호교차검증 결과 (발생량) (단위: %)	86
〈그림 3-7〉 시도 단위 업무 통행의 상호교차검증 결과 (발생량) (단위: 통행/일)	87
〈그림 3-8〉 시도 단위 업무 통행의 상호교차검증 결과 (발생량) (단위: %)	88
〈그림 3-9〉 시도 단위 귀가 통행의 상호교차검증 결과 (발생량) (단위: 통행/일)	89
〈그림 3-10〉 시도 단위 귀가 통행의 상호교차검증 결과 (발생량) (단위: %)	89





# 후 여



## 1. 사업의 개요

### 가. 사업의 배경 및 목적

- 개인기기 사용의 보편화, 생활 전반의 디지털화 등 첨단 기술 발전에 따라 개인의 모든 이동이 데이터로 수집되는 모빌리티 빅데이터가 수집되고 있으며, 이를 활용하기 위한 연구가 진행되고 있음
- 모바일통신 빅데이터는 교통 분야에서 활용 중인 기존 빅데이터보다 표본율이 높고, 개별 통행 정보가 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 통행 정보를 제공할 수 있음
- 이동통신단말기를 통해 수집되는 모바일통신 빅데이터를 이용한 통행사슬(Trip Chain) DB를 구축하여 여객 및 화물 O/D 구축을 지원하고, 통행량 등 교통조사 수행 및 기종점 통행량을 구축 및 검증하고자 함

## 나. 사업의 범위 및 내용

### 1) 모바일 통신데이터 기반 기종점 통행DB 구축

- 공간적 범위: 전국
- 시간적 범위: 2024년 10월
  - LG유플러스 최신 원천데이터 기반으로 모바일 통신데이터 통행DB 구축
- 내용적 범위
  - 본 사업의 활용 목적, 통신사의 데이터 특성에 따라 통행 DB의 구조를 설계하고 정의된 로직에 따라 2024년 10월 통행 DB를 생성함
  - 통행 정보(체류지 유형, 통행량, 통행시간, 통행속도, 통행거리 등) 검증 기준 정립하고 정립된 기준에 따라 통행DB 구축 결과를 점검함

### 2) 모바일 통신데이터 기반 통행 DB 검증

- 공간적 범위: 전국 (시도 단위)
- 시간적 범위
  - 모바일 통신데이터: KT, LG, SKT 데이터 (2023년 11월)
  - 조사자료: KTDB 전국 지역 간 목적 OD (2022년)
- 내용적 범위
  - (데이터간 상호 비교를 통한 검증) 데이터 4종(모바일 통신데이터 3종, 조사자료 1종)의 통행목적별 통행량을 상호 비교함
  - 시도 단위로 전체, 통행목적별 발생량, 도착량을 각각 비교하였으며, 발생량과 도착량이 유사하여 본 보고서에서는 발생량 기준으로 검증 결과를 제시하였음

### 3) 개별 객체 단위 모바일 통신데이터 활용 통행지표 개발

- 시간적 범위: 2024년
- 공간적 범위: 전국
- 내용적 범위
  - 기존 사업에서 개발한 교통지표를 2024년 사업에서 구축한 LG유플러스 통신 데이터 기반 통행 DB를 활용하여 구축함
  - 문헌고찰을 통해 모바일 통신데이터에서 산출 가능한 교통지표를 발굴함

## 2. 모바일 통신데이터 기반 기종점 통행DB 구축

### 가. 모바일 통신데이터 기반 기종점 통행DB 형태 설계

#### 1) 모바일 통신데이터 기반 통행DB 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), 종점 정보(도착일자, 도착 시간대, 도착 위치, 도착 체류지 유형), 통행자 정보(성, 연령)를 기준으로 통행 정보가 같은 인구를 집계
    - 일자 정보는 1일 단위로 구축하도록 함
- 출발 시각 및 도착 시각 정보는 집계된 인구가 3명 이하가 될 가능성을 줄이면서, DB 사용자에게 가능한 섬세한 통행 정보를 제공할 수 있도록 한 시간 단위로 입력하도록 함
  - 출발 시각은 출발지에서의 체류 종료시각을 의미하며, 도착시간은 도착지에서의 체류 시작시각을 의미함
- 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 머

무르는지 추정할 수 있는 정보를 제공하기 위한 것으로, 「국가교통조사 및 분석」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

- 통행량 필드에는 전술한 기준에 의해 집계된 인구(단말기 수)를 입력함
  - 통행량 필드에서 전술한 기준에 의해 집계된 통행량을 입력하되, 집계된 통행량이 3 이하의 값을 갖는 경우, 3통행 미만 보정계수를 적용하여 추정된 통행량을 입력하도록 함
- 통행시간 필드에는 집계된 통행 정보를 기준으로 산출한 평균 통행시간, 통행거리 값을 입력함

## 2) 통행DB 구축 형태 검토 및 설계

- 현재 통신사에서 구매 및 확보 가능한 모바일 통신데이터의 공간적·시간적 범위 및 상세도를 고려하여 통행 DB 구축 형태를 설계함
  - 집계 단위 : 500m 격자, 일자별, 1시간, 기종점 단위
- 기종점 통행 DB는 여객 O/D 추정 및 지역간·지역내 통행분포 분석을 목적으로 구축함
  - 전국 읍면동 단위보다 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

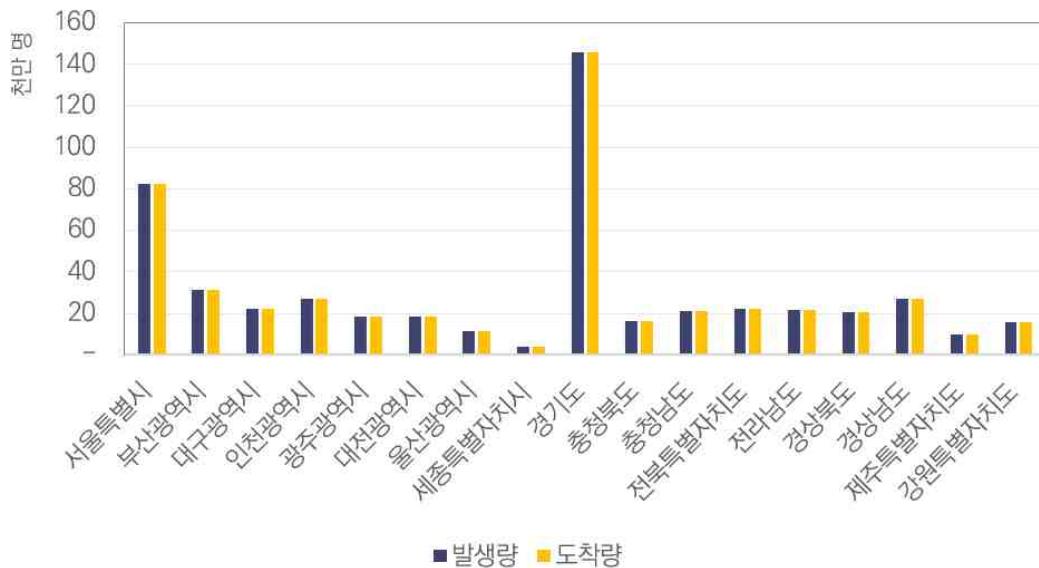
## 3) 통행DB 구축 및 검토

- 기종점 통행 DB 구축 및 검토
  - 2024년 11월에 대한 시간대별/연령별/지역별 기종점 통행 DB 구축 및 검토
- 개인별 통행궤적 DB 구축 및 검토
  - 2024년에 대한 개인별 통행궤적 DB 구축

## 나. 모바일 통신데이터 기반 기종점 통행DB 구축 결과

### ○ 지역별 총 통행량

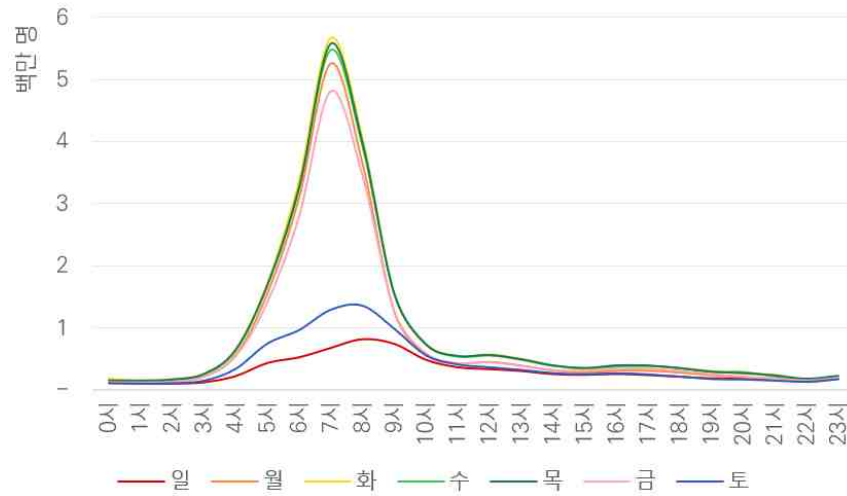
- 모든 시도별로 통행 발생, 도착량 차이가 크지 않은 것으로 나타남
- 시도별 총 통행량은 경기도, 서울특별시, 부산광역시, 인천광역시 순으로 높게 나타나 수도권, 대도시의 통행량이 큰 것으로 나타남
- 하지만 이 결과는 사전에 LG 유플러스 모바일 통신데이터 수집 과정에서 지역별 통행 발생, 도착량 비율을 유사하게 설정했을 가능성이 있으며, 지역별 LG 유플러스 점유율 차이가 있어 해석에 유의해야 함



〈그림 1〉 지역별 총 통행량

### ○ 시간대별 출근 통행량 (요일별)

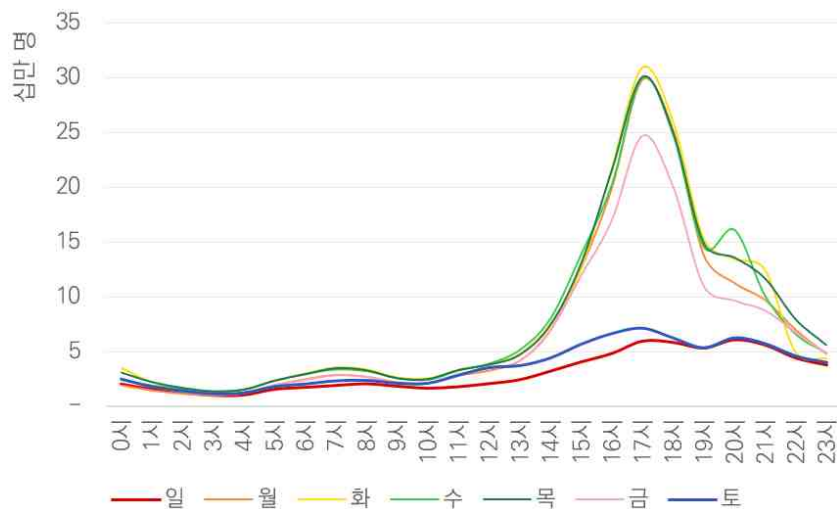
- 주중은 주말 대비 출근통행 발생 비율이 집중되며, 첨두시간은 오전 7시로 나타남
- 주중 월~목 대비 금요일의 출근 통행량이 첨두시(오전 7시)에 가장 낮게 나타남
- 주말은 첨두시간이 오전 8시로 나타나며, 주중에 비해 첨두시 출근 통행량이 최소 3배~최대 6배까지 낮게 나타남



〈그림 2〉 시간대별 출근 통행량 (요일별)

○ 시간대별 퇴근 통행량 (요일별)

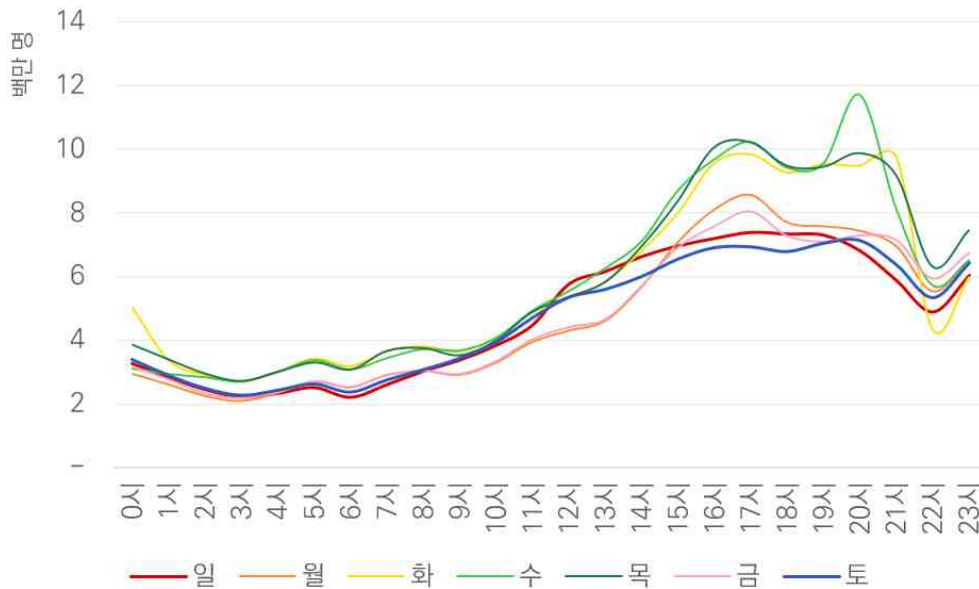
- 퇴근 통행은 주로 늦은 오후 시간에 이루어지며, 주중 퇴근 통행 발생량은 오후 5시에 집중된 것으로 나타남
- 첨두시(오후 5시)와 그 이후 시간대에도 주중 중에서는 금요일의 퇴근 통행량이 가장 낮음



〈그림 3〉 시간대별 퇴근 통행량 (요일별)

○ 시간대별 귀가 통행량 (요일별)

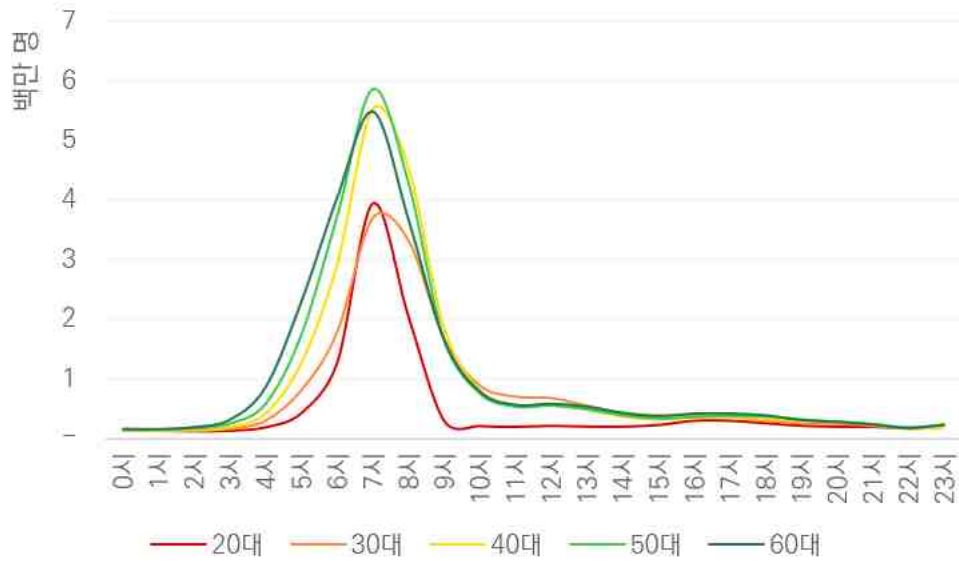
- 귀가 통행 발생량은 오전보다 오후에 집중된 것으로 나타남
- 주중 화~목요일의 오후 8시~9시 통행량이 두드러지는데, 이는 직장에서 업무를 마친 뒤 다른 장소에 방문하고, 이후 집으로 돌아오는 통행량으로 판단됨



〈그림 4〉 시간대별 귀가 통행량 (요일별)

○ 시간대별 출근 통행량 (연령대별)

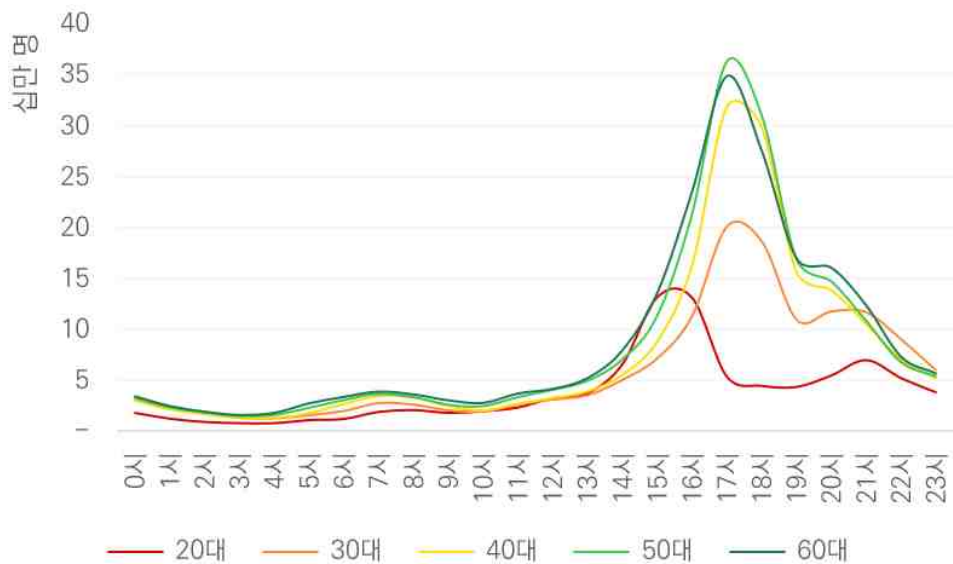
- 모든 연령대의 출근 통행량은 오전 7시 전후로 높게 나타남
- 첨두시(오전 7시)에 40~60대 출근 통행량이 20~30대 대비 높게 나타나는데, 이는 40~60대의 경제활동 종사자 비율이 높거나 혹은 모바일 통신데이터에서 해당 연령대가 차지하는 비율이 높기 때문으로 판단됨
- 연령대가 높을수록 출근 통행이 이른 시간에 발생함



<그림 5> 시간대별 출근 통행량 (연령대별)

○ 시간대별 퇴근 통행량 (연령대별)

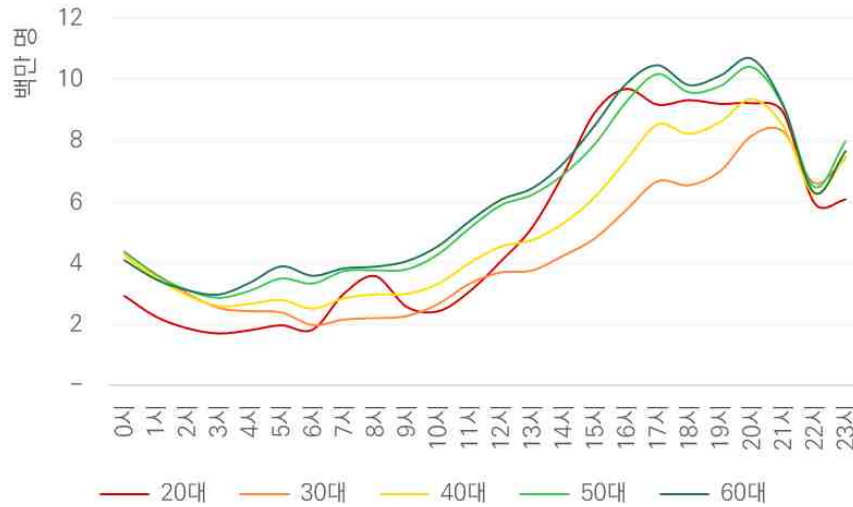
- 20대를 제외한 모든 연령대의 퇴근 통행량은 오후 5~6시에 가장 높게 나타남
- 20대는 오후 3~4시에 퇴근 통행량이 높게 나타나는데, 이는 20대의 경우 학업 또는 파트타임 근로 형태가 통상적인 업무시간(오전 9시~오후 6시)을 갖는 직장인의 근무 패턴과는 다르기 때문으로 판단됨



<그림 6> 시간대별 퇴근 통행량 (연령대별)

○ 시간대별 귀가 통행량 (연령대별)

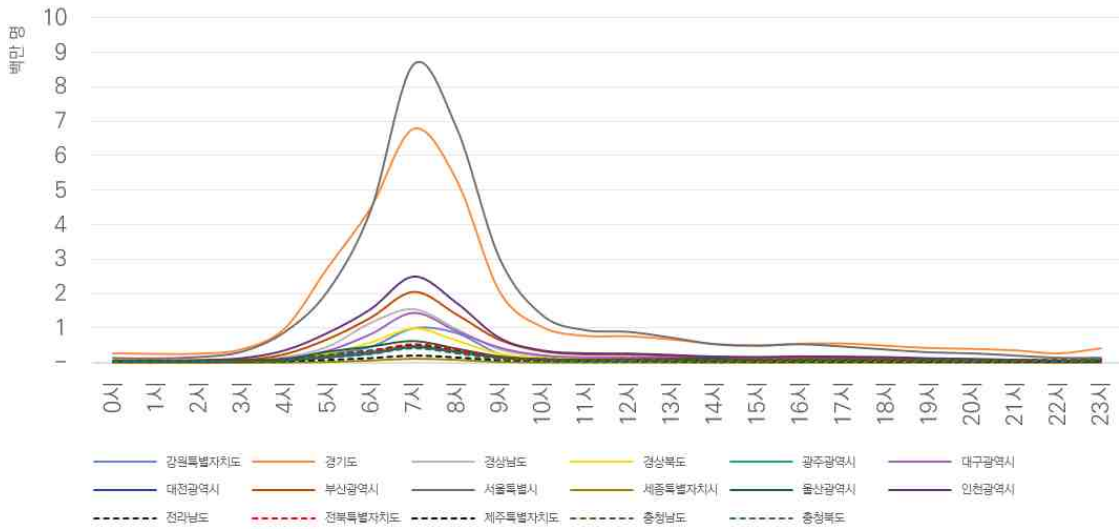
- 시간대별 귀가 통행 발생량은 20대를 제외한 모든 연령대에서 오후 5시, 8시에 높은 양봉형으로 나타남
- 20대 귀가 통행량은 오후 4~9시까지 높게 나타남



〈그림 7〉 시간대별 귀가 통행량 (연령대별)

○ 시간대별 출근 통행량 (지역별)

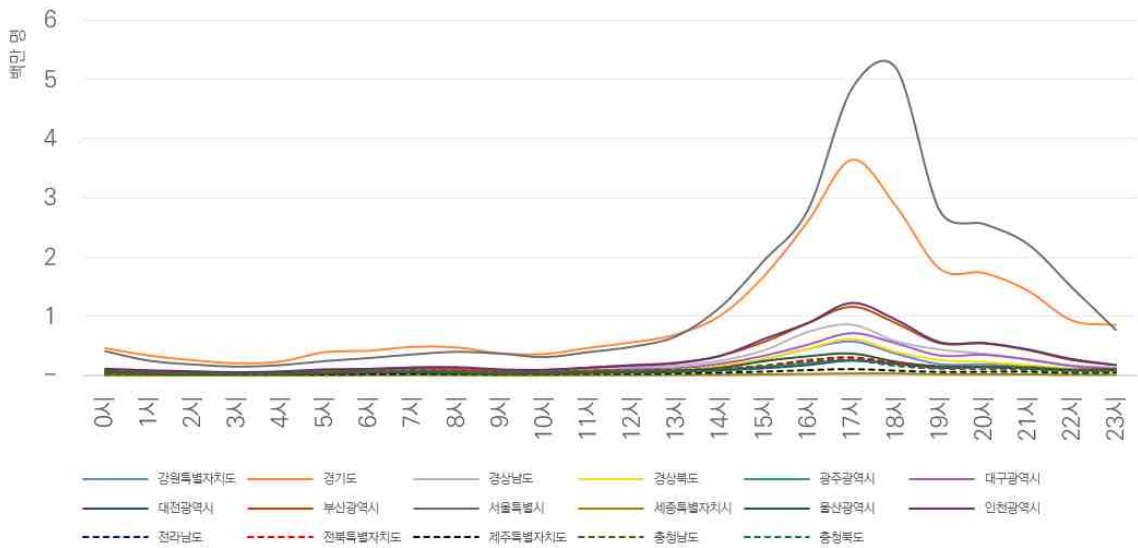
- 시간대별 출근 통행량은 모든 지역에서 오전 7시에 가장 높게 나타나며, 서울특별시, 경기도, 인천광역시, 부산광역시, 경상남도, 대구광역시 순으로 나타남
- 이렇게 지역별 출근 통행량이 수도권, 대도시에서 높게 나타나는 것은 지역의 활발한 취업 기회나 높은 취업자 수에 기인할 수 있음



〈그림 8〉 시간대별 출근 통행량 (지역별)

○ 시간대별 퇴근 통행량 (지역별)

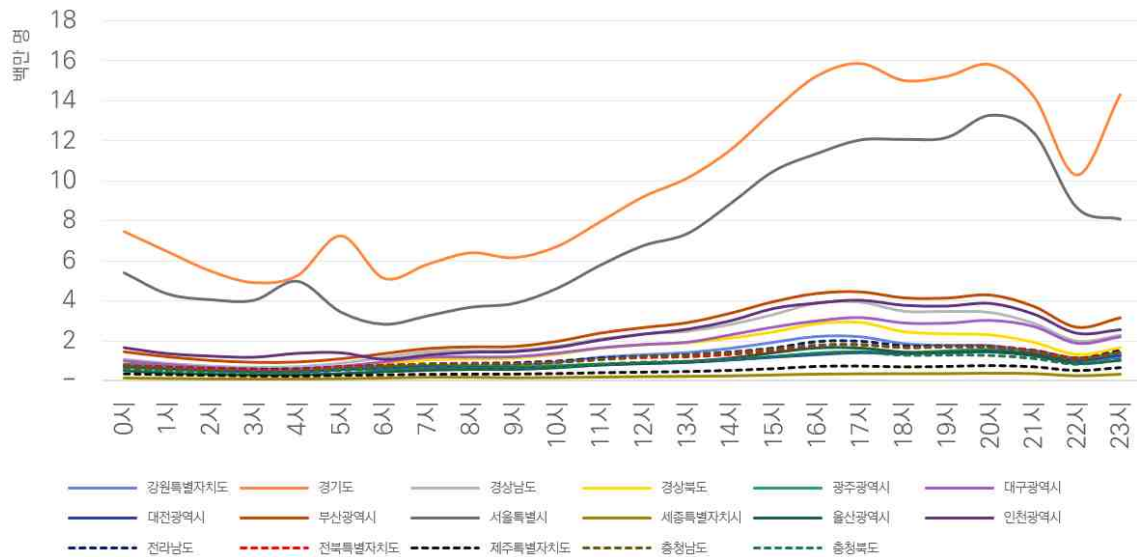
- 시간대별 퇴근 통행량은 출근 통행량과 동일하게 서울특별시, 경기도, 인천광역시, 부산광역시, 경상남도, 대구광역시 순으로 높게 나타남
- 시간대별 퇴근 통행량은 서울특별시를 제외한 모든 지역에서 오후 5시에 가장 높게 나타나며, 서울시는 오후 6시에 가장 높게 나타남



〈그림 9〉 시간대별 퇴근 통행량 (지역별)

○ 시간대별 귀가 통행량 (지역별)

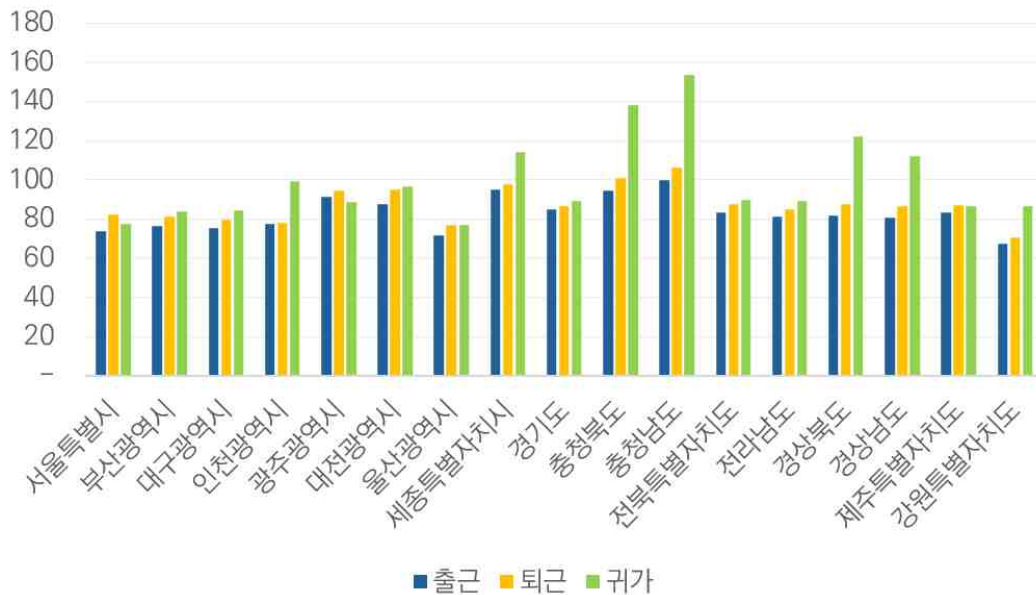
- 대부분 지역에서 귀가 통행량은 오후 5시, 8시에 집중되는 것으로 나타남
- 반면 서울특별시는 오후 8시에 귀가 통행량이 두드러지는데, 이는 업무종료 후 인근의 다양한 활동 기회를 가진 뒤 다른 지역민 대비 늦게 귀가하기 때문으로 판단됨
- 이외에도 수도권(서울, 경기, 인천)의 오전 4~5시 귀가 통행량이 두드러지는데, 이는 수도권에 24시간 운영되는 업종 종사자(병원, 물류센터 등)가 많거나 수도권의 심야버스와 택시 등 교통인프라가 잘 갖춰져 있어 시간에 관계없이 귀가가 가능하기 때문으로 판단됨



〈그림 10〉 시간대별 귀가 통행량 (지역별)

○ 지역별 통행시간

- 서울특별시, 부산광역시, 대구광역시, 인천광역시, 울산광역시와 같은 대도시  
는 통행시간이 다른 지역에 비해 적게 소요되는 것으로 나타남
- 이중 울산은 대표적인 공업도시로 산업단지와 주거지가 밀집해 있어 통행시간  
이 짧은 것으로 판단됨
- 반면 충청북도, 충청남도, 세종특별자치시는 다른 시도 대비 통행시간이 길게  
나타나는데, 이는 대중교통 인프라가 부족하고 교통접근성이 떨어지는 것과  
관련 있을 수 있음



<그림 11> 지역별 통행시간 (단위: 분)

## 다. 객체단위 모바일 통신데이터 가공 및 통행DB 구축

### 1) 객체단위 통행DB 구축 목적

- 모바일 통신데이터는 개인이 휴대전화 단말기를 소지한 상태에서 발생하는 신호를 기반으로 통행 궤적을 분석할 수 있는 중요한 데이터임
- 기존 교통량 데이터는 주로 승용차, 버스, 지하철 등 전통적인 교통수단을 중심으로 데이터를 구축하여 활용하였음
- 보행, 자전거, 개인형 이동수단(PM) 등의 이동정보가 체계적으로 구축되지 않아 해당 수단의 이용행태 분석에 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
- 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 통행DB 구축 범위

- 본 사업에서는 현재 확보 가능한 개별 객체단위 모바일 통신 데이터에 대한 전처리 및 가공, 통행경로 및 통행지표 DB 구축 등 기술적 검토를 수행함
- DB 구축을 위한 객체단위 데이터는 본원에서 확보한 “이동통신 가명정보” 데이터를 활용 예정임
  - 공간적 범위 : 전국
  - 시간적 범위 : 2024. 10. 13. ~ 2024. 10. 20.
  - 집계 단위 : 50m 격자, 일자별, 1분 단위, 개별통행 트립(Trip) 단위 집계
- 본 사업에서 수행한 기술적 검토 사항은 아래와 같음
  - 객체 단위 모바일 통신데이터 오류 검토 및 필터링
  - 모바일 통신데이터 기반 개별 통행경로 DB 구축
  - 모바일 통신데이터 기반 통행지표 구축 관련 기술적 검토

## 3) 객체단위 DB개요 및 데이터 형태

- LG 유플러스 모바일 통신 원시 데이터는 총 21개의 컬럼이며, 기준일자, 트립 ID, 출·도착 시간, 출·도착 시도/시군구/읍면동/50셀 코드, 이동 유형 코드, 평균 속도, 출발-도착 유형으로 구성되어 있음
- LG 유플러스 모바일 통신 원시 데이터의 공간단위는 격자단위로 50m 단위로 구분되어 있으며, 시간 단위는 1분 단위로 구분되어 있음
- 이동거리는 km 단위이며, 속도는 km/h 단위로 되어 있음

《표 1》 LG유플러스 이동통신 가명정보 데이터 테이블정의서

번호	변수명	데이터 유형	설명
1	P_YYYYMMDD	날짜(문자열)	기준 날짜 (YYYY-MM-DD)
2	TRIP_NO	정수	이동 트립 번호
3	DPR_MT1_UNIT_TM	정수	출발 시각 (YYYY-MM-DD HH-MM) 1분 단위
4	DPR_CTDO_CD	정수	출발 시도 코드
5	DPR_CTDO_NM	문자열	출발 시도명
6	DPR_CCW_CD	정수	출발 시군구 코드
7	DPR_CCW_NM	문자열	출발 시군구명
8	DPR_ADNG_CD	정수	출발 읍면동 코드
9	DPR_ADNG_NM	문자열	출발 읍면동명
10	DPR_CELL_ID	정수	출발 셀 ID(50m 격자)
11	ARV_MT1_UNIT_TM	정수	도착 시각 (YYYY-MM-DD HH-MM) 1분 단위
12	ARV_CTDO_CD	정수	도착 시도 코드
13	ARV_CTDO_NM	문자열	도착 시도명
14	ARV_CCW_CD	정수	도착 시군구 코드
15	ARV_CCW_NM	문자열	도착 시군구명
16	ARV_ADNG_CD	정수	도착 읍면동 코드
17	ARV_ADNG_NM	문자열	도착 읍면동명
18	ARV_CELL_ID	정수	도착 셀 ID(50m 격자)
19	DYNA_DYN_KD_CD	정수	이동 유형 코드
20	DYNA_MVMT_SPED	실수	평균 속도
21	OD_TYPE	문자열	출발-도착 유형

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

《표 2》 LG 유플러스 모바일 통신 격자&공간정보 데이터 테이블정의서

순번	필드명	설명	자료형	비고
1	CELL_ID	격자 ID	NUMERIC	-
2	X_AXIS_WGS	중심_X_WGS84	NUMERIC	-
3	Y_AXIS_WGS	중심_Y_WGS84	NUMERIC	-
4	MEGA_NM	시도명	VARCHAR	-
5	MEGA_CD	시도코드	NUMERIC	-
6	CTY_NM	시군구명	VARCHAR	-
7	CTY_CD	시군구코드	NUMERIC	-
8	ADMI_NM	행정동명	VARCHAR	-
9	ADMI_CD	행정동코드	NUMERIC	-

#### 4) 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 21개가 아닌 경우 제거
  - 날짜 오류 제거: 기준일자 필드가 비었거나 형식(YYYY-MM-DD)에 맞지 않은 경우 제거
  - 시간 오류 제거: 출발시간 또는 도착시간 필드가 비었거나 형식(YYYY-MM-DD HH-MM)에 맞지 않은 경우 제거
  - 참조 오류 제거: 출발격자 또는 도착격자 필드가 격자-행정구역 매칭 테이블에 참조되지 않는 경우 제거
- 데이터의 효율적인 처리를 위해 다음과 같은 전처리 과정을 수행함
  - 일자 분할 : 일자별 trip\_id로 파일 분할
  - 시간 정보 변환 : Datetime → Unixtime
  - 통행목적 제외: 무의미한 형태로 포함되어 있는 통행목적 컬럼 제외
- 최종 전처리된 데이터의 형태는 다음과 같음
- 데이터 발생일자와 트립 넘버를 조합한 Trip\_ID로 키값을 설정하였으며, 출발격자ID, 도착격자 ID, 출발 행정동 ID, 도착

#### 5) 통행궤적 및 통행정보 데이터 구축

- 원시데이터에는 개별 통행궤적(Trajectory)에 대한 시퀀스(Sequence) 정보가 존재하지 않음
- 원시데이터는 단순한 위치 정보의 집합으로, 시간 순서(Time Sequence)를 반영하지 않을 경우 이동 경로를 명확히 파악하기 어려움
- 따라서, 출발/도착시각 정보를 활용하여 통행궤적을 재구성할 필요가 있음

- 원시데이터는 단순한 위치 데이터로 저장되며, 시퀀스 정보 없이 개별 위치 좌표만 제공됨
- 원시데이터를 활용하여 연속적인 이동궤적을 구축하고, 이동 구간 간의 관계를 정의하기 위해 연속된 시간대를 기반으로 데이터를 정렬 및 연결이 필요함
- 본 연구에서는 출발지부터 도착지까지의 이동 경로를 분석하기 위하여 연속된 시간 데이터를 기반으로 통행궤적을 구축함

## 라. 객체 단위 모바일 통신데이터 구축 및 활용 관련 검토사항

### 1) 이동 유형별 상세 분석 관련 기술적 검토

- 이동 유형별 분석을 위해 요구되는 기술적 요소
  - 이동 유형을 보다 정밀하게 분석하기 위해서는 위치, 시간, 교통 네트워크 등 데이터 간 융합 분석이 필수적임
  - 기존의 분류 방식은 단순 속도 변수만을 기준으로 이동 유형을 분류하였기 때문에 정밀한 분석이 어려움
  - 예를 들어, 속도만으로 도보, 비보도를 판별할 경우 정체 구간에서 차량 이동이 도보로 잘못 분류될 가능성이 높음
  - 따라서, 속도뿐만 아니라 GPS데이터, 맵매칭 기술 등을 활용하여 도보 이동을 더욱 정확하게 분석할 필요가 있음
- GPS X,Y좌표 데이터 기반 도보 분석
  - 현재 모바일 통신데이터는 GPS 좌표체계가 아닌 격자 단위(Grid Cell)로 이동 궤적을 제공하고 있음
  - 격자 단위는 객체의 이동 흐름을 대략적으로 파악하는데 유용하나, 정확한 도보 경로 추적에는 한계가 있음
- 초(Second) 단위 데이터 수집
  - 본 사업에서 수집된 데이터는 1분 간격의 시간 데이터로 구성되어 있음

- 초 단위의 통신 데이터를 수집하여 통행궤적을 더욱 명확하게 구분할 수 있으며, 실제 이동 구간과 대기 상태 등을 분석할 수 있음
- 가속도 센서 데이터 분석 활용
  - 휴대폰에 내장된 가속도 센서를 활용하여 도보, 달리기, 자전거 등의 이동 수단을 구분할 수 있음
  - 가속도 센서는 이용자의 움직임 강도와 방향을 측정할 수 있으며, 이를 통행도보와 기타 이동 수단을 구분하는 데 활용 가능함
  - 이를 통해, 속도만으로 이동 유형을 판별하는 기존 방식의 한계를 보완할 수 있음
- 도로 네트워크 맵매칭 적용
  - 통행을 정밀하게 분석하기 위해 다양한 교통 네트워크(도로망, 대중교통망, 철도망 등)와 통신데이터를 결합하는 맵매칭 기술 적용이 필요함
  - 맵매칭을 활용하면, 통신데이터를 특정 네트워크에 맵핑하여 도보 통행을 보다 신뢰성 있게 분석할 수 있음
  - 이를 적용할 경우, 이동수단 간 전환 구간(도보, 탑승, 환승 등)을 보다 정밀하게 분석할 수 있으며, 교통 네트워크 기반의 통행 분석이 가능함
  - 향후 이러한 맵매칭 기술을 활용하면 모바일 통신데이터 기반 통행 분석의 정확도를 향상시킬 수 있을 것으로 기대됨

## 2) 통행지표 구축 관련 기술적 검토

- 데이터 정확성 확보를 위한 기술적 검토
  - 데이터의 정확성은 분석 결과의 신뢰성에 직접적인 영향을 미치는 요소임
  - 데이터 수집 과정에서 발생할 수 있는 오류를 최소화하고, 수집된 데이터가 충분한 정확도를 확보하고 있는지 검토하기 위한 기술 개발이 필수적임
  - GPS 신호 오류 보정, 이동속도 및 궤적 검증, 이상치 제거 등
- 데이터 처리 및 분석 환경 구축

- 모바일 통신 원시데이터는 대용량 데이터이며, 이를 효과적으로 처리하기 위해서는 고성능 데이터베이스 설계 및 분석 환경이 필수적임
- 빅데이터 처리 및 분석을 위해서는 데이터베이스, 빅데이터 가공 환경 구축, 인프라 시스템 등 기술적 고려가 필요함
- 통행지표 산출 알고리즘 적용
  - 최근 최적화 알고리즘, 머신러닝 등을 활용한 통행지표 생성 기술이 활발히 연구되고 있음
  - 신뢰성 높은 통행지표를 산출하기 위해서는 최신 알고리즘 적용 및 데이터 셋 구축이 필수적임
- 데이터 속성정보 수집 및 구축
  - 통행지표의 신뢰성을 높이기 위해 이동 궤적 데이터뿐만 아니라, 다양한 속성정보(메타데이터)를 추가로 확보할 필요가 있음
  - 인구통계학적 정보(성별, 연령, 직업 등), 통행목적 정보(출·퇴근, 등·학교, 쇼핑, 여가 등), 교통수단 정보(승용차, 버스, 철도, 해운, 항공, 자전거, PM 등)

### 3. 모바일 통신데이터 기반 통행DB 검증

#### 가. 개요

##### 1) 기본방향

- 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
    - KT 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, KT는 전수화 전후의 통행량을 제공하며, 전수화 통행 정보의 경우 노이즈(noise)를 추가하는 익명화 처리(data anonymization)가 실시됨

#### 나. 검증데이터 수집 및 가공

##### 1) 데이터 수집

###### ① 모바일 통신데이터 기반 통행DB

- 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
    - KT, LG, SKT 데이터를 기반으로 구축한 통행 DB를 수집하였으며, 상호 비교를 위하

여 시간적·공간적·내용적 범위를 동일하게 설정함

- 통신사별로 개인정보보호를 위한 특정 기준 이하의 통행량(3 또는 5 통행량)은 마스킹 처리함

○ LG

- 시간적 범위: 2023년 11월
- 공간적 범위: 전국 (시도 단위)
- 출처: 2024년 DB사업을 통해 구축한 LG 통신데이터 기반 기중점 통행량 데이터

〈표 3〉 LG 데이터 정의서

종류	변수 세부정보	변수명	비고	
기본 정보	기준일자	P_YYYYMMDD	YYYY-MM-DD	
	성별	SEX_DV_CD	MALE, FEMALE	
	연령대	YY10_AGLV_ID	10세 단위	
출발 정보	시간대	DPR_TMZN	00, 01, 02, … , 22, 23	
	셀ID	DPR_CELL_ID	-	
	X좌표	DPR_X_AXIS_WGS	격자 중심점 X좌표	
	Y좌표	DPR_Y_AXIS_WGS	격자 중심점 Y좌표	
	시도코드	DPR_CTDO_CD	격자 중심점 시도코드	
	시도명	DPR_CTDO_NM	격자 중심점 시도명	
	시군구코드	DPR_CCW_CD	격자 중심점 시군구코드	
	시군구명	DPR_CCW_NM	격자 중심점 시군구명	
	행정동코드	DPR_ADNG_CD	격자 중심점 행정동코드	
	행정동명	DPR_ADNG_NM	격자 중심점 행정동명	
	장소유형	DPR_PLC_TYPE	거주지, 근무지, 기타	
	도착 정보	시간대	ARV_TMZN	00, 01, 02, … , 22, 23
		셀ID	ARV_CELL_ID	-
X좌표		ARV_X_AXIS_WGS	격자 중심점 X좌표	
Y좌표		ARV_Y_AXIS_WGS	격자 중심점 Y좌표	
시도코드		ARV_CTDO_CD	격자 중심점 시도코드	
시도명		ARV_CTDO_NM	격자 중심점 시도명	
시군구코드		ARV_CCW_CD	격자 중심점 시군구코드	
시군구명		ARV_CCW_NM	격자 중심점 시군구명	
행정동코드		ARV_ADNG_CD	격자 중심점 행정동코드	
행정동명	ARV_ADNG_NM	격자 중심점 행정동명		
통행 정보	장소유형	ARV_PLC_TYPE	거주지, 근무지, 기타	
	이동자수	MOVE_CUST_CNT	전국민확대계수 적용된 이동자수	
	평균 이동거리	AVG_MOVE_TIME	평균 이동시간	
	평균 이동시간	AVG_MOVE_DIST	평균 이동거리	

- 시도별 통행량

- 2023년 11월 LG사의 시도별 통행 발생량은 경기도(27.7%), 서울특별시(17.2%), 부산광역시(6.5%) 순으로 나타났으며, 수도권은 50.5%에 달함
- 통행량이 가장 적게 나타난 지역은 세종특별자치시이며, 전체의 약 0.7%를 차지함

<표 4> 2023년 11월 시도별 통행량(LG)

(단위: 통행/월, %)

시도	출발		도착	
	통행량	비율	통행량	비율
서울특별시	787,874,234	17.2	787,959,889	17.2
부산광역시	298,479,667	6.5	298,502,833	6.5
대구광역시	209,856,308	4.6	209,890,294	4.6
인천광역시	257,017,671	5.6	257,013,111	5.6
광주광역시	144,826,400	3.2	144,852,747	3.2
대전광역시	145,391,871	3.2	145,391,208	3.2
울산광역시	107,688,679	2.3	107,679,607	2.3
세종특별자치시	31,149,462	0.7	31,135,501	0.7
경기도	1,271,829,252	27.7	1,271,853,585	27.7
충청북도	138,896,775	3.0	138,873,440	3.0
충청남도	171,112,657	3.7	171,095,620	3.7
전북특별자치도	173,427,897	3.8	173,451,190	3.8
전라남도	166,961,056	3.6	166,891,680	3.6
경상북도	200,200,907	4.4	200,170,219	4.4
경상남도	267,070,812	5.8	267,071,794	5.8
제주특별자치도	78,561,311	1.7	78,550,635	1.7
강원특별자치도	138,296,274	3.0	138,257,880	3.0
합 계	4,588,641,233	100.0	4,588,641,232	100.0

○ KT 데이터

- 시간적 범위: 2023년 11월
- 공간적 범위: 전국 (500 × 500m 격자 단위)
- 출처: 국가교통조사사업을 통해 구축한 KT 통신데이터 기반의 2023년 기종점 통행량 데이터

〈표 5〉 KT 데이터 정의서

칼럼 ID	칼럼명	비고
ETL_YMD	기준 일자	
O_ID	출발 격자 ID	500m 격자 ID
O_X	출발 격자 X 좌표	500m 격자 중점 X 좌표
O_Y	출발 격자 Y 좌표	500m 격자 중점 Y 좌표
D_ID	도착 격자 ID	500m 격자 ID
D_X	도착 격자 X 좌표	500m 격자 중점 X 좌표
D_Y	도착 격자 Y 좌표	500m 격자 중점 Y 좌표
O_TIME_CD	출발 시간대 코드	1시간 단위
D_TIME_CD	도착 시간대 코드	1시간 단위
O_TYPE	출발 격자 유형	H:집(야간상주지), W:직장/학교(주간상주지), E:기타
D_TYPE	도착 격자 유형	H:집(야간상주지), W:직장/학교(주간상주지), E:기타
O_SIDO_NM	출발 시도명	
O_SGG_NM	출발 시군구명	
O_ADMI_NM	출발 행정동명	
O_ADMI_CD	출발 행정동 코드	
D_SIDO_NM	도착 시도명	
D_SGG_NM	도착 시군구명	
D_ADMI_NM	도착 행정동명	
D_ADMI_CD	도착 행정동코드	
SEXAGE	성연령 코드	10세 단위 성연령 코드
PURPOSE	이동 목적	출근, 등교, 업무, 귀가, 기타
CNT	전수화 집계 값	
RATIO	전수화 비율	전수화를 위한 비율
AVG_DIST	평균 이동 거리	동일 이동의 평균 이동 거리(10m 단위)
AVG_TIME	평균 이동 시간	동일 이동의 평균 이동 시간(1분 단위)

- 시도별 통행량

- 2023년 11월 LG사의 시도별 통행 발생량은 경기도(24.4%), 서울특별시(19.7%), 부산광역시(6.5%) 순으로 나타났으며, 수도권은 49.2%에 달함
- 통행량이 가장 적게 나타난 지역은 세종특별자치시이며, 전체의 약 0.7%를 차지함

〈표 6〉 2023년 11월 시도별 통행량(KT)

(단위: 통행/월, %)

시도	출발		도착	
	통행량	비율	통행량	비율
서울특별시	466,962,643	19.7	469,134,279	19.8
부산광역시	150,262,670	6.5	150,509,838	6.5
대구광역시	108,066,714	4.7	108,074,563	4.7
인천광역시	118,903,620	5.1	119,334,767	5.2
광주광역시	58,830,383	2.5	58,967,283	2.5
대전광역시	69,248,608	3.0	69,301,408	3.0
울산광역시	50,147,715	2.2	50,174,691	2.2
세종특별자치시	16,170,457	0.7	16,174,414	0.7
경기도	565,702,605	24.4	563,377,810	24.3
충청북도	75,535,773	3.3	75,455,837	3.3
충청남도	98,865,329	4.4	98,662,238	4.3
전북특별자치도	80,407,632	3.6	80,381,042	3.6
전라남도	80,408,913	3.6	80,249,376	3.6
경상북도	114,152,916	5.1	114,044,251	5.1
경상남도	135,470,031	5.9	135,311,317	5.9
제주특별자치도	43,944,487	2.0	44,001,010	2.0
강원특별자치도	70,730,311	3.4	70,656,685	3.4
합 계	2,303,810,808	100.0	2,303,810,808	100.0

○ SKT

- 시간적 범위: 2023년 11월
- 공간적 범위: 전국 (시도 단위)
- 출처: 2023년 DB사업을 통해 구축한 SKT 통신데이터 기반 기종점 통행량 데이터

〈표 7〉 SKT 데이터 정의서

칼럼 ID	칼럼명	비고
o_month	기준 월	
o_dow	출발 요일	1~7(월~일)
o_time	출발 시간대	1시간 단위
o_sigun	출발지 시군구	5자리 숫자코드
o_sido	출발지 시도	2자리 숫자코드
o_type	출발 유형	집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)
d_time	도착 시간대	1시간 단위
d_sigun	도착지 시군구	5자리 숫자코드
d_sido	도착지 시도	2자리 숫자코드
d_type	도착 유형	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	평균 통행량	보정된 통행량(5통행 이하 마스킹 처리)
move_time_avg	평균 통행속도	초(Sec)

- 시도별 통행량
  - 2023년 11월 LG사의 시도별 통행 발생량은 서울특별시(22.5%), 경기도(22.4%), 부산광역시(7.2%) 순으로 나타났으며, 수도권의 비율은 50.7%에 달함
  - 통행량이 가장 적게 나타난 지역은 세종특별자치시이며, 발생량은 전체의 약 0.5%를 차지함

〈표 8〉 2023년 11월 시도별 통행량(SKT)

(단위: 통행/월, %)

시도	출발		도착	
	통행량	비율	통행량	비율
서울특별시	668,454,733	22.5	677,344,925	22.8
부산광역시	213,756,736	7.2	216,442,756	7.3
대구광역시	158,427,378	5.3	155,415,696	5.2
인천광역시	170,811,967	5.8	175,260,541	5.9
광주광역시	101,262,500	3.4	99,619,135	3.4
대전광역시	95,544,833	3.2	96,297,654	3.2
울산광역시	66,218,388	2.2	65,651,610	2.2
세종특별자치시	14,785,309	0.5	13,140,298	0.4
경기도	663,268,357	22.4	648,083,790	21.9
충청북도	84,700,537	2.9	84,951,510	2.9
충청남도	105,952,532	3.6	107,806,006	3.6
전라북도	105,450,464	3.6	106,387,101	3.6
전라남도	89,941,418	3.0	91,058,950	3.1
경상북도	131,888,211	4.4	134,429,830	4.5
경상남도	172,163,459	5.8	170,542,868	5.7
제주특별자치도	41,378,058	1.4	41,487,208	1.4
강원특별자치도	82,002,859	2.8	82,087,860	2.8
합 계	2,966,007,739	100.0	2,966,007,739	100.0

## ② 조사자료

- 모바일 통신데이터와 상호교차 검증하기 위한 자료로 기존에 교통 수요 분석의 기초자료로 주로 활용되었던 조사자료를 수집함
  - 조사자료 중에서 통행량 정보를 제공하고 있는 KTDB 전국 지역간 목적 OD 자료를 구득함
- 전국 지역간 목적 OD 자료
  - 시간적 범위: 2022년
  - 공간적 범위: 전국 (시도, 시군구 단위)
  - 내용적 범위: 내부통행 포함, 지역간 목적별 통행량
  - 출처: 2024년 국가교통조사사업을 통해 구축한 자료

## 2) 데이터 간 상호교차검증 결과

### ① 총 통행량 & 통행 비율

○ 총 통행량 측면에서 모바일 통신데이터 3종의 여객 O/D(내부통행 포함)와 SKT 데이터(내부통행 포함)가 가장 유사하게 나타남

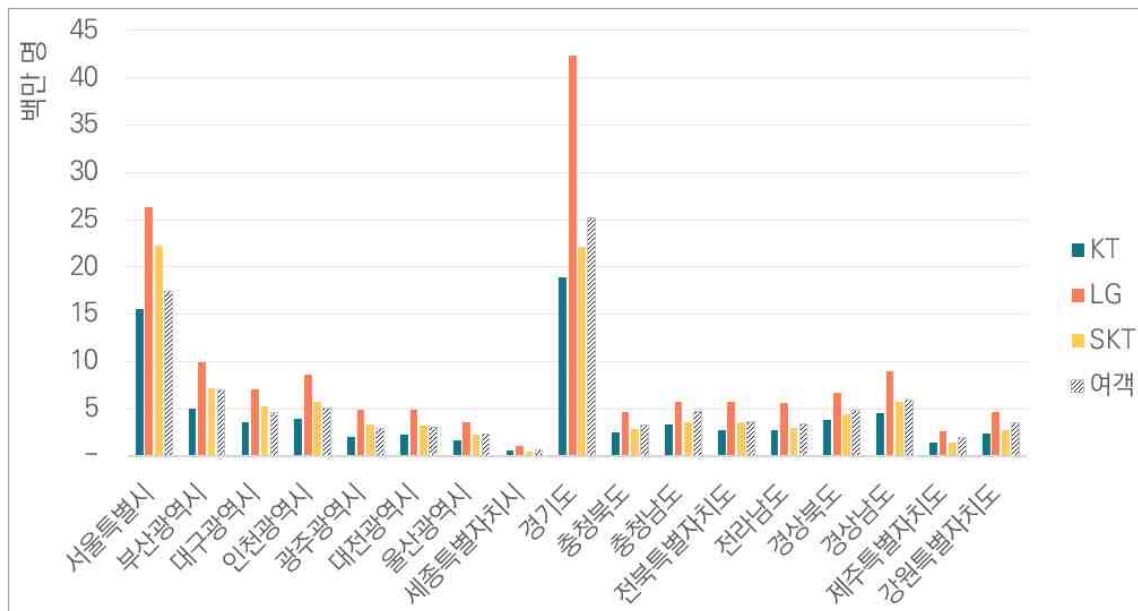
- 총통행량은 3사 중 LG가 가장 높게 나타남

○ 모바일 통신 3사 데이터와 여객 O/D 데이터의 시도별 통행 발생, 도착 비율이 대체로 유사하게 나타나나, 일부 지역은 3사 간 비율 차이가 두드러짐

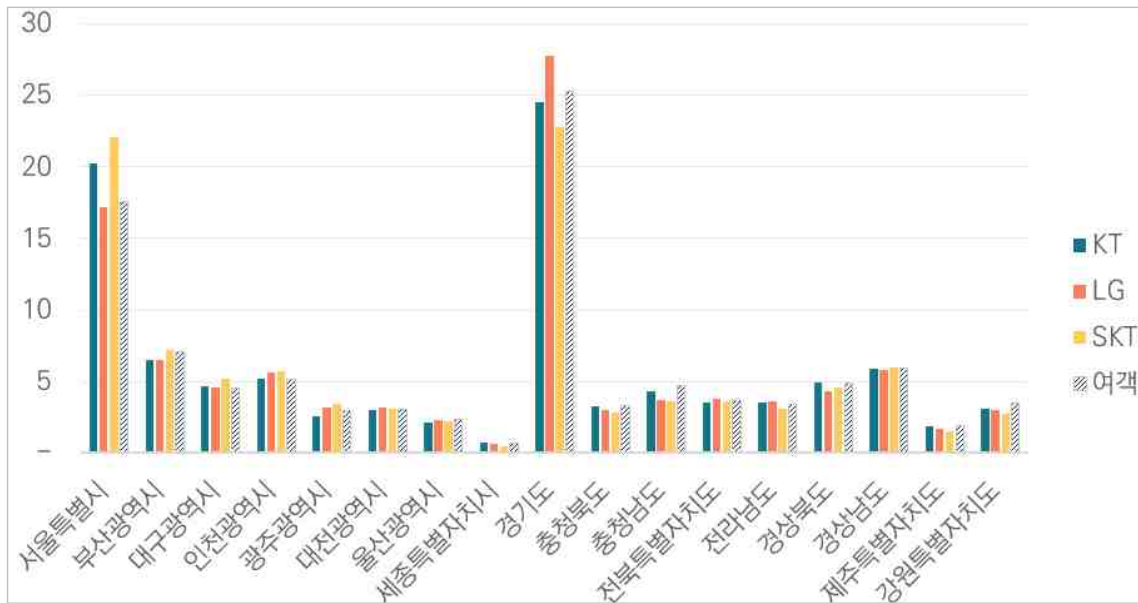
- 서울특별시는 KT, SKT 비율이 높게 나타남

- 반대로 경기도는 LG의 비율이 가장 높게 나타나며, 다음으로 KT, SKT 순임

- 충청도는 KT사의 통행 비율이 타지역 대비 가장 높게 나타남



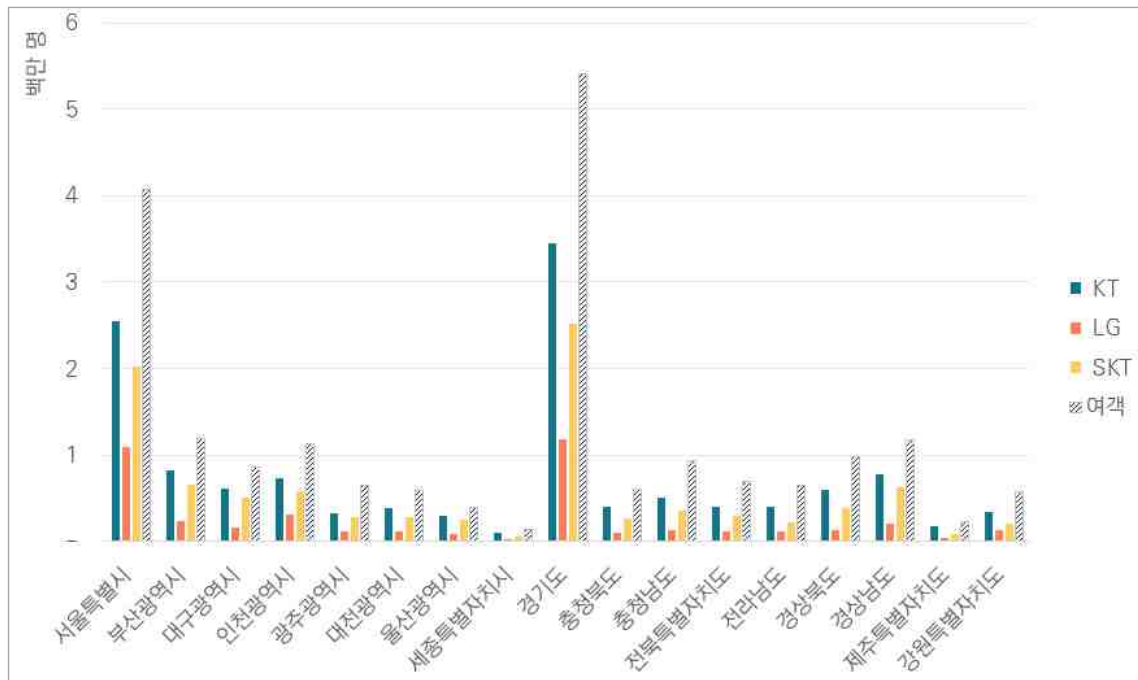
〈그림 12〉 시도 단위 전체 통행의 상호교차검증 결과 (발생량) (단위: 통행/일)



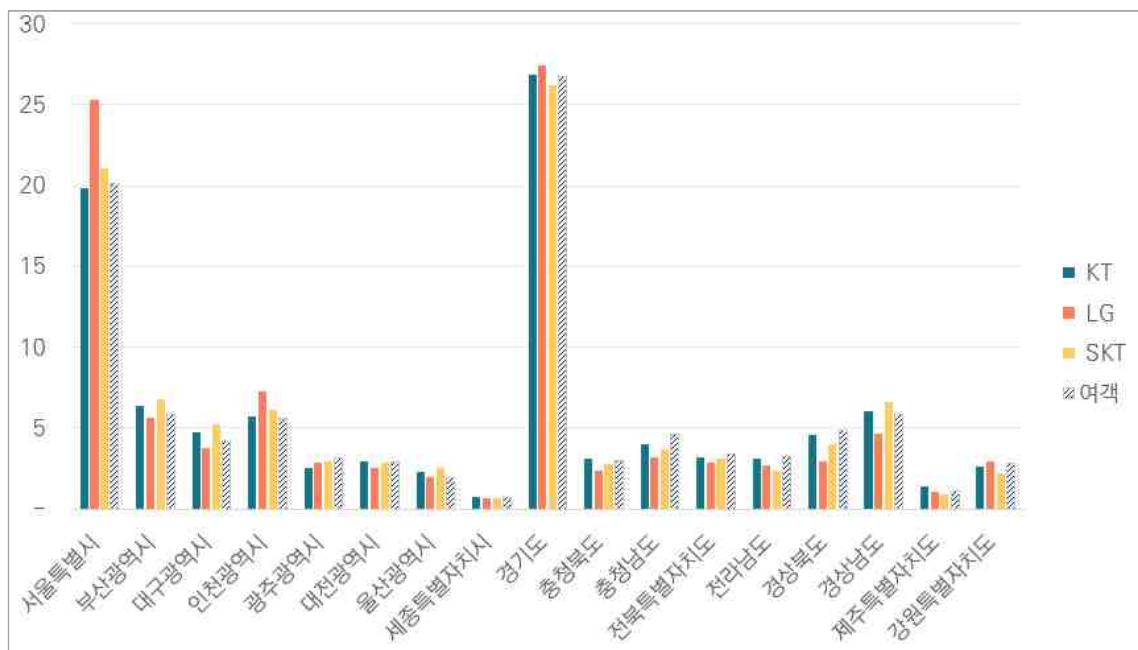
〈그림 13〉 시도 단위 전체 통행의 상호교차검증 결과 (발생량) (단위: %)

## ② 출근 통행량 & 통행 비율

- 출근 통행량은 여객 O/D의 통행량이 모든 지역에서 모바일 통신 3사에 비해 높게 나타났음
  - 모바일 통신 3사 중에서는 KT, SKT, LG 순으로 통행량이 높게 나타남
- 출근 통행 비율은 모바일 통신 3사와 여객 O/D 데이터가 시도별로 유사하게 나타나나, KT 통행 비율이 여객 O/D와 가장 유사하게 나타남
  - 수도권(서울, 경기, 인천)은 LG의 통행 비율이 가장 높게 나타남



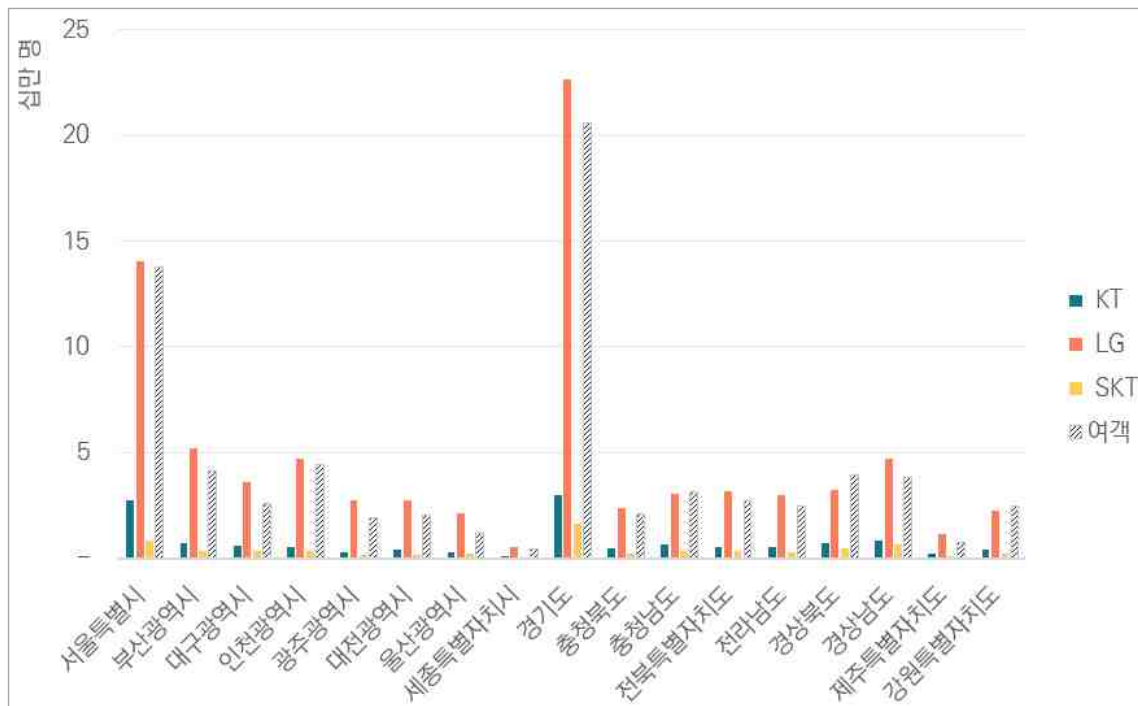
〈그림 14〉 시도 단위 출근 통행의 상호교차검증 결과 (발생량) (단위: 통행/일)



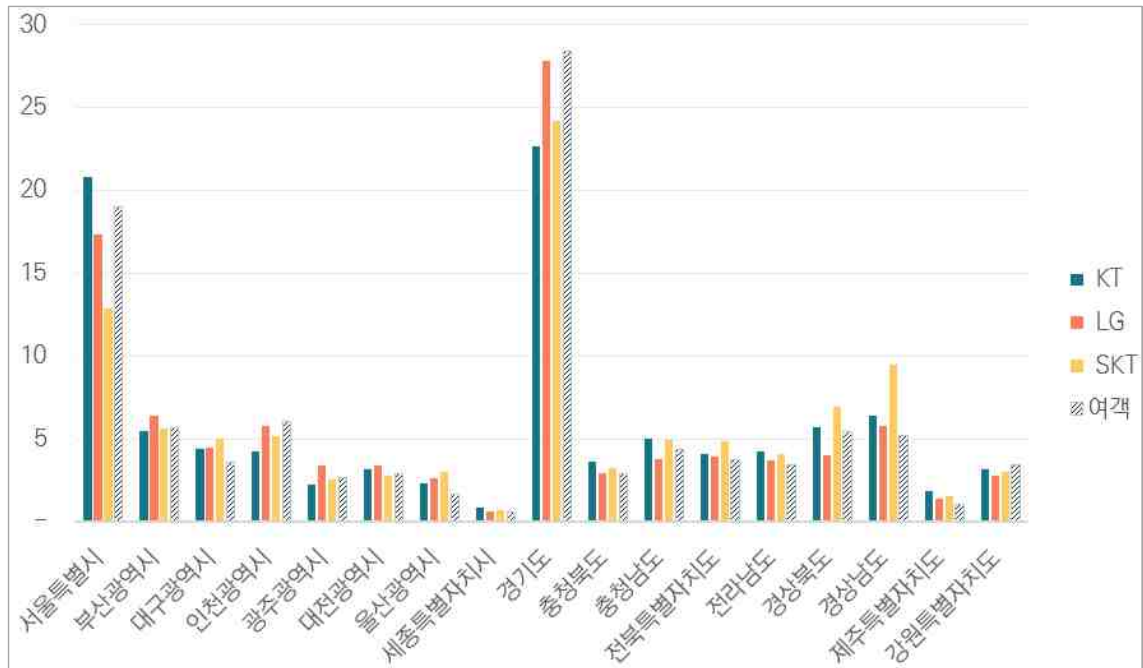
〈그림 15〉 시도 단위 출근 통행의 상호교차검증 결과 (발생량) (단위: %)

### ③ 업무 통행량 & 통행 비율

- 시도별 업무 통행량은 LG가 여객 O/D와 가장 유사하게 나타남
  - KT, SKT는 업무 통행량이 LG에 비해 적으나, 지역별 통행량 경향은 유사함
- 업무 통행 비율 또한 LG가 여객 O/D와 가장 유사하게 나타나며, 지역별로 모바일 3사의 비율이 상이함
  - 서울특별시, 충청도는 KT의 업무통행 비율이 높음
  - 경상북도, 경상남도는 SKT의 시도별 업무 통행 비율이 높게 나타남
  - LG는 경기도, 광역시(부산, 인천, 광주)의 업무 통행 비율이 높게 나타남



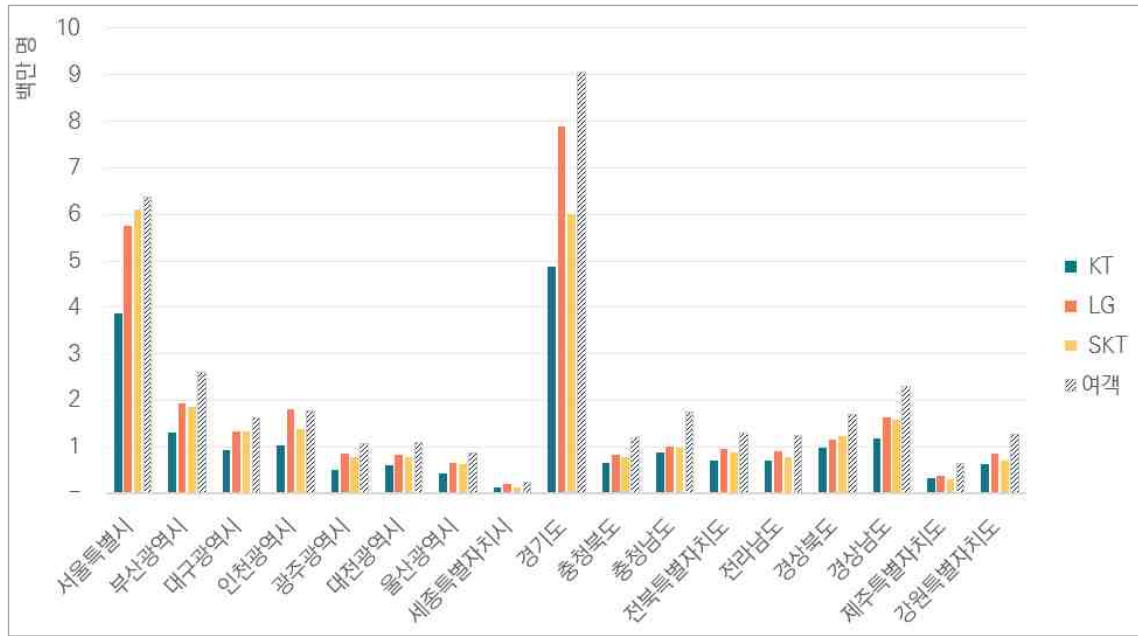
(그림 16) 시도 단위 업무 통행의 상호교차검증 결과 (발생량) (단위: 통행/인)



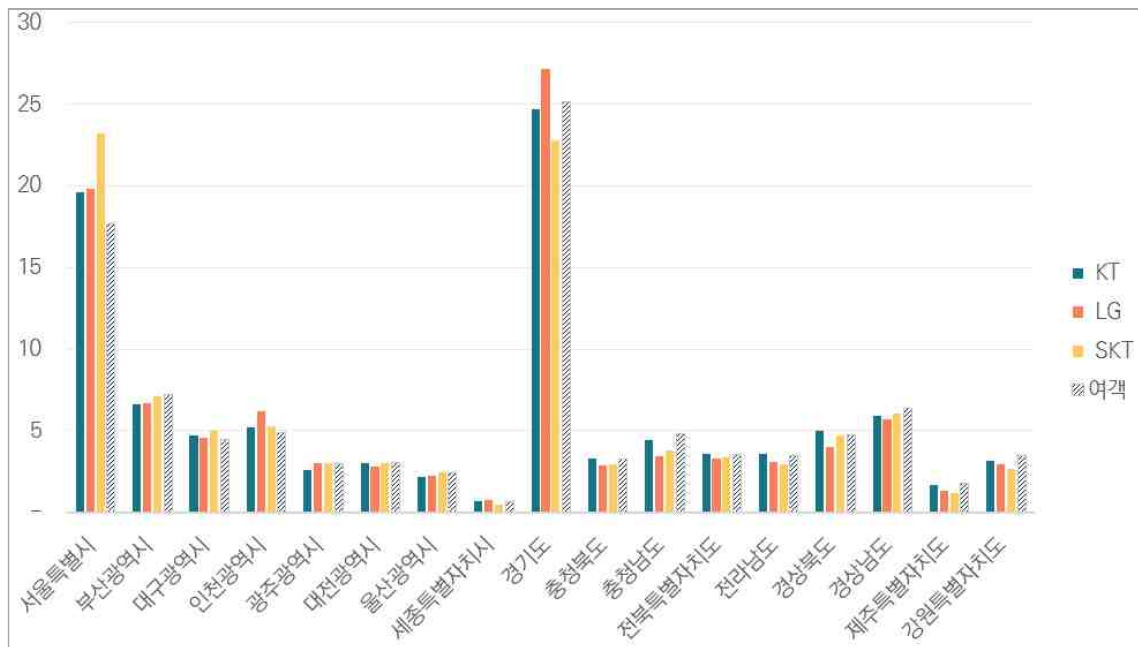
〈그림 17〉 시도 단위 업무 통행의 상호교차검증 결과 (발생량) (단위: %)

#### ④ 귀가 통행량 & 통행 비율

- 귀가 통행량은 여객 O/D이 가장 높게 나타나며, LG 귀가 통행량이 여객 O/D와 유사하게 나타남
  - KT, SKT는 귀가 통행량이 LG에 비해 적으나, 지역별 통행량 경향은 유사함
- 모바일 통신 3사 데이터와 여객 O/D 데이터의 시도별 귀가통행 발생, 도착 비율이 대체로 유사하게 나타나나, 일부 지역은 3사 간 비율 차이가 두드러짐
  - 수도권 중 서울특별시는 SKT의 업무 통행비율이 높게 나타나며, 경기도와 인천광역시는 LG사의 통행비율이 높게 나타남
  - 그 외 충청도, 경상도, 전라도, 제주도, 강원도 지역은 모두 KT사의 귀가 통행 비율이 높게 나타남



〈그림 18〉 시도 단위 귀가 통행의 상호교차검증 결과 (발생량) (단위: 통행/일)



〈그림 19〉 시도 단위 귀가 통행의 상호교차검증 결과 (발생량) (단위: %)

## 4. 모바일 통신데이터 기반 교통지표 구축

### 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 2024년 사업에서 구축한 KT 모바일 통신데이터 기반 통행 DB를 기준으로 아래 표와 같이 모바일 통신데이터로 산출 가능한 교통지표를 DB화 하여 구축함

〈표 9〉 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	설, 가정의 달, 추석, 하계 휴가기간 등 특정 연휴기간을 설정하여 집계한 통행량	출발일자, 도착일자
		계절별 통행량	봄, 여름, 가을, 겨울에 대해 기간을 설정하고 집계한 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	

교통지표		설명	지표 추출 기준 (활용된 통행정보)
대분류	소분류		
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

## 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
  - 모바일 통신 데이터로 구축된 통행DB를 시간대별, 요일별, 연령별 통행량 및 통행목적별 통행시간으로 구분하여 분석함
- 또한, 동일한 기간(2023년 11월)에 생성된 로그 기록에 모바일 통신데이터 가공 알고리즘을 적용하여 통행DB를 검증함
  - 개인별 로그 기록은 「개인정보보호」, 「위치정보의 보호 및 이용 등에 관한 법률」에 저촉되지 않도록 집계한 형태의 DB를 설계하였으며, 이를 기준으로 DB를 구축함
  - 데이터 4종(모바일 통신데이터 3종, 조사자료 1종)의 통행목적별 통행량을 상호 비교함

## 나. 차년도 수행계획

- 시의성 있는 통행 정보를 제공하기 위하여 2025년 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 2024년 기준의 모바일 통신데이터를 기준으로 본 사업에서 진행한 것과 같이 통행량 및 통행패턴에 대한 데이터 간 상호교차 검증을 수행할 예정임
  - 현재는 통신 3사와 전국 지역간 목적 OD의 총통행, 목적별 통행량을 비교하였음
  - 향후에는 통행량과 사회경제지표 간, 지역·권역·도서지역 간 통행량을 비교하며, 음수셀 발생 여부에 대해서도 파악할 예정임
- 모바일 통신데이터를 활용하여 통행지표를 산출 및 현행화할 예정임
  - 개별 객체 단위 모바일 통신데이터를 활용한 통행지표 산출 방안을 검토할 예정임
  - 2024년 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

- 개인기기 사용의 보편화, 생활 전반의 디지털화 등 첨단 기술 발전에 따라 개인의 모든 이동이 데이터로 수집되는 모빌리티 빅데이터가 수집되고 있으며, 이를 활용하기 위한 연구가 진행되고 있음
- 특히 모바일 통신 빅데이터는 교통 분야에서 활용 중인 여타 빅데이터보다 표본율이 높고, 개별통행 정보가 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
  - 모바일 기기와 기지국 간의 송수신 이력인 모바일 통신데이터는 대부분 국민이 이용하고 있는 이동통신 단말기를 통해 축적되기 때문에, 표본율이 매우 높고 개인의 통행궤적이 상세하게 기록됨
- 또한, 기존 인력식 조사 방식에 비해 경제적이고 효율적으로 통행 정보를 취득할 수 있으며, 수집·분석 시간이 단축되어 시의성 있는 통행 정보를 제공할 수 있음

## 2. 사업의 목적

- 국가통합교통체계효율화법 제14조 제1항에 의거 정보통신수단(휴대전화)을 활용하여 교통수단 시설의 운영 실태, 통행량 등 교통조사 수행 및 기종점통행량을 구축 및 검증하고자 함

\* 국가통합체계효율화법 시행령 제8조(국가교통조사의 실시) 제3항 교통수단별 교통시설별 여객 및 화물의 기종점 통행량

- 이동통신단말기를 통해 수집되는 모바일통신 빅데이터를 활용하여 개인단위 통행 정보를 파악할 수 있는 통행DB를 구축하여 교통정책 수립 지원을 위한 각종 교통지표를 구축 및 제공하고자 함

## 제2절 사업의 범위 및 내용

### 1. 사업의 범위

#### 가. 모바일 통신데이터 기반 기종점 통행DB 구축

- 공간적 범위: 전국
- 시간적 범위: 2024년 10월
  - LG유플러스 최신 원천데이터 기반으로 모바일 통신데이터 통행DB 구축

#### 나. 모바일 통신데이터 기반 통행 DB 검증

- 공간적 범위: 전국 (시도 단위)
- 시간적 범위: 2023년 11월
  - KT, LG, SKT 데이터 (2023년 11월) 통행DB, KTDB 전국 지역간 목적 OD 자료 등 4종 데이터 간 비교 분석
    - 시도 단위로 전체, 통행목적별 발생량, 도착량을 각각 비교하였으며, 발생량과 도착량이 유사하여 본 보고서에서는 발생량 기준으로 검증 결과를 제시하였음

#### 다. 개별 객체 단위 모바일 통신데이터 활용 통행지표 개발

- 공간적 범위: 전국
- 시간적 범위: 2024년
  - 2024년 사업에서 구축한 통신데이터 기반 통행DB를 활용함

## 2. 사업의 내용

### 가. 모바일 통신데이터 기반 기종점 통행DB 구축

- 모바일 통신데이터 기반 통행 DB 구축
  - 본 사업의 활용 목적, 통신사의 데이터 특성에 따라 통행 DB의 구조를 설계하고 정의된 로직에 따라 2024년 통행 DB를 생성
  - 2024년 11월에 대한 시간대별/연령별/지역별 기종점 통행 DB 검토
- 개인별 통행계획 DB 구축 및 검토
  - 2024년에 대한 개인별 통행계획 DB 구축

### 나. 모바일 통신데이터 기반 통행 DB 검증

- 모바일 통신데이터 검증
  - (데이터 간 상호 비교를 통한 검증) 데이터 4종(모바일 통신데이터 3종, 조사자료 1종)의 통행목적별 통행량을 상호 비교함
    - 시도 단위로 전체, 통행목적별 발생량, 도착량을 각각 비교하였으며, 발생량과 도착량이 유사하여 본 보고서에서는 발생량 기준으로 검증 결과를 제시하였음

### 다. 개별 객체 단위 모바일 통신데이터 활용 통행지표 개발

- 교통지표 구축
  - 기존 사업에서 개발한 교통지표를 2024년 사업에서 구축한 LG유플러스 통신데이터 기반 통행 DB를 활용하여 구축함
- 신규 교통지표 개발
  - 문헌고찰을 통해 모바일 통신데이터에서 산출 가능한 교통지표를 발굴함



## 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 활용 목적을 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) 총여객 O/D 추정 및 지역간·지역내 통행분포 분석을 위한 보완자료로 활용
  - (전국 단위 교통DB 구축) 모바일 통신데이터 기반 사람의 통행을 파악할 수 있는 기종점 인구 통행량 DB 구축
  - (전국 단위 통행지표 구축) 통행목적, 지역 특성, 인적 특성에 따른 전국 전역에 대한 사람 통행 현황 분석을 위한 지표 구축
- 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
  - 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 객체 단위 원천 통행DB 확보 및 활용이 어려우며, 다양한 목적에 대한 분석 및 미시적인 수준의 통행 분석에 어려움이 있음
- 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 설계

- 현재 통신사에서 구매 및 확보 가능한 모바일 통신데이터의 공간적·시간적 범위 및 상세도를 고려하여 통행 DB 구축 형태를 설계함
  - 집계 단위 : 500m 격자, 일자별, 1시간, 기종점 단위
- 기종점 통행 DB는 여객 O/D 추정 및 지역간·지역내 통행분포 분석을 목적으로 구축함
  - 전국 읍면동 단위보다 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 통행패턴 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, 통행패턴, 통행수단, 통행목적 등 상세 정보 포함
  - 대중교통, 보행, 자전거, PM 등 기존 통행DB 미구축 수단에 대한 분석을 위한 통행 DB 설계

## 3. 통행 DB 구축 및 검토

- 기종점 통행 DB 구축 및 검토
  - 2024년 11월에 대한 시간대별/연령별/지역별 기종점 통행 DB 구축 및 검토
- 개인별 통행패턴 DB 구축 및 검토
  - 2024년에 대한 개인별 통행패턴 DB 구축

## 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), 종점 정보(도착일자, 도착 시간대, 도착 위치, 도착 체류지 유형), 통행자 정보(성, 연령)를 기준으로 통행 정보가 같은 인구를 집계
    - 일자 정보는 1일 단위로 구축하도록 함
- 출발 시간 및 도착 시간 정보는 집계된 인구가 3명 이하가 될 가능성을 줄이면서, DB 사용자에게 가능한 섬세한 통행 정보를 제공할 수 있도록 한 시간 단위로 입력하도록 함
  - 출발 시간은 출발지에서의 체류종료시간을 의미하며, 도착시간은 도착지에서의 체류시작시간을 의미함
- 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, 「국가교통조사 및 분석」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
- 통행량 필드에는 전술한 기준에 의해 집계된 인구(단말기 수)를 입력함
  - 통행량 필드에서 전술한 기준에 의해 집계된 통행량을 입력하되, 집계된 통행량이 3 이하의 값을 갖는 경우, 3통행 미만 보정계수를 적용하여 추정된 통행량을 입력하도록 함
- 통행시간 필드에는 집계된 통행 정보를 기준으로 산출한 평균 통행시간, 통행거리 값을 입력함

#### 나. 전수화 미적용 통행 DB 설계

- 전국에 대한 통행 DB 설계(전수화 미적용)
  - 전국 기준 평일/주말 기준 기종점 통행DB 설계 및 구축
  - 공간적 단위 : 500\*500m 격자
  - 시간적 단위 : 시간대(00시~05시 / 06시~08시 / 09시~17시 / 18시~19시 / 20시~23시)
  - 이동통계정보 데이터는 총 7개의 칼럼으로 평일/주말, 시간대, 출발지 정보(시도/시군구/읍면동/500셀 좌표), 도착지 정보(시도/시군구/읍면동/500셀 좌표), 통행량, 이동시간, 이동거리로 구성되어 있음
    - 단, 3통행 미만 통행량에 대해서는 최솟값, 최댓값, 표준편차를 구축 및 제공하지 않음

**<표 2-1> 모바일통신 빅데이터 DB 형태 (전수화 미적용 통행 정보)**

출발 정보								
날짜	시간대	읍면동	시군구	시도	체류지 유형*			
평일/주말	06~08	11010530	11010	11	거주지			
도착 정보								
날짜	시간대	읍면동	시군구	시도	체류지 유형*			
평일/주말	06~08	31091710	31091	31	근무지			
통행 정보**								
총 통행량	이동시간				이동거리			
	평균	최솟값	최댓값	표준편차	평균	최솟값	최댓값	표준편차
6								

주: \* 거주지, 근무지, 기타

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

**<표 2-2> 모바일통신 빅데이터 DB 테이블 정의서 (전수화 미적용)**

종류	변수 세부정보	변수명	비고
기본 정보	기준일자	YYYYMM	기준일자(YYYYMM)
	평일/주말 코드	WEEK_GB	평일, 주말
	시간대	TMZN_GRP	00시~05시 / 06~08시 / 09~17시 / 18시~19시 / 20시~23시
출발 정보	시간대	DPR_TMZN	00, 01, 02, ... , 22, 23
	셀ID	DPR_CELL_ID	-
	X좌표	DPR_X_AXIS_WGS	격자 중심점 X좌표
	Y좌표	DPR_Y_AXIS_WGS	격자 중심점 Y좌표
	시도코드	DPR_CTDO_CD	격자 중심점 시도코드
	시도명	DPR_CTDO_NM	격자 중심점 시도명
	시군구코드	DPR_CCW_CD	격자 중심점 시군구코드
	시군구명	DPR_CCW_NM	격자 중심점 시군구명
	행정동코드	DPR_ADNG_CD	격자 중심점 행정동코드
	행정동명	DPR_ADNG_NM	격자 중심점 행정동명
	장소유형	DPR_PLC_TYPE	거주지, 근무지, 기타
도착 정보	시간대	ARV_TMZN	00, 01, 02, ... , 22, 23
	셀ID	ARV_CELL_ID	-
	X좌표	ARV_X_AXIS_WGS	격자 중심점 X좌표
	Y좌표	ARV_Y_AXIS_WGS	격자 중심점 Y좌표

종류	변수 세부정보	변수명	비고	
	시도코드	ARV_CTDO_CD	격자 중심점 시도코드	
	시도명	ARV_CTDO_NM	격자 중심점 시도명	
	시군구코드	ARV_CCW_CD	격자 중심점 시군구코드	
	시군구명	ARV_CCW_NM	격자 중심점 시군구명	
	행정동코드	ARV_ADNG_CD	격자 중심점 행정동코드	
	행정동명	ARV_ADNG_NM	격자 중심점 행정동명	
	장소유형	ARV_PLC_TYPE	거주지, 근무지, 기타	
통행 정보	개인이동자수	MOVE_CUST_CNT	전국민확대계수 적용된 이동자수	
	이동 시간	평균	MOVE_CUST_CNT	-
		최소	AVG_MOVE_TIME	-
		최대	MIN_MOVE_TIME	-
		표준편차	MAX_MOVE_TIME	-
	이동 거리	평균	STD_MOVE_TIME	-
		최소	AVG_MOVE_DIST	-
		최대	MIN_MOVE_DIST	-
		표준편차	MAX_MOVE_DIST	-

#### 다. 전수화 적용 통행 DB 설계

○ 전국에 대한 통행 DB 설계(전수화 적용)

- DB 구축 대상 기간의 각 월에 대해 요일별/시간대별/통행목적별 출발-도착 통행량 분석을 위한 DB를 설계함

《표 2-3》 모바일통신 빅데이터 DB 테이블 정의서 (전수화 적용)

종류	변수 세부정보	변수명	비고
기본 정보	기준일자	P_YYYYMMDD	YYYY-MM-DD
	성별	SEX_DV_CD	MALE, FEMALE
	연령대	YY10_AGLV_ID	10세 단위
출발 정보	시간대	DPR_TMZN	00, 01, 02, ... , 22, 23
	셀ID	DPR_CELL_ID	-
	X좌표	DPR_X_AXIS_WGS	격자 중심점 X좌표

종류	변수 세부정보	변수명	비고
	Y좌표	DPR_Y_AXIS_WGS	격자 중심점 Y좌표
	시도코드	DPR_CTDO_CD	격자 중심점 시도코드
	시도명	DPR_CTDO_NM	격자 중심점 시도명
	시군구코드	DPR_CCW_CD	격자 중심점 시군구코드
	시군구명	DPR_CCW_NM	격자 중심점 시군구명
	행정동코드	DPR_ADNG_CD	격자 중심점 행정동코드
	행정동명	DPR_ADNG_NM	격자 중심점 행정동명
	장소유형	DPR_PLC_TYPE	거주지, 근무지, 기타
도착 정보	시간대	ARV_TMZN	00, 01, 02, ... , 22, 23
	셀ID	ARV_CELL_ID	-
	X좌표	ARV_X_AXIS_WGS	격자 중심점 X좌표
	Y좌표	ARV_Y_AXIS_WGS	격자 중심점 Y좌표
	시도코드	ARV_CTDO_CD	격자 중심점 시도코드
	시도명	ARV_CTDO_NM	격자 중심점 시도명
	시군구코드	ARV_CCW_CD	격자 중심점 시군구코드
	시군구명	ARV_CCW_NM	격자 중심점 시군구명
	행정동코드	ARV_ADNG_CD	격자 중심점 행정동코드
	행정동명	ARV_ADNG_NM	격자 중심점 행정동명
	장소유형	ARV_PLC_TYPE	거주지, 근무지, 기타
통행 정보	이동자수	MOVE_CUST_CNT	전국민확대계수 적용된 이동자수
	평균 이동거리	AVG_MOVE_TIME	평균 이동시간
	평균 이동시간	AVG_MOVE_DIST	평균 이동거리

## 2. 모바일 통신데이터 기반 통행DB 구축 방법

### 가. LG 유플러스

#### 1) 원천데이터 수집 및 이상치 제거

- 2024년 1월 1일부터 2024년 12월 31일까지 기록된 모바일 단말기 실시간 위치 및 기지국 위치 위경도에 대해 셀ID의 위경도를 맵핑함
- 기지국 좌표가 Null값 이거나 행정동 좌표가 매칭되지 않는 데이터를 제거함
  - 기지국 이상치 기준을 설정하여 기지국별 추정위치 거리 및 패턴, 매칭여부 등을 고려함
  - 셀 유형별 건물 및 도로면적 기준으로 유동인구생성 불가지역을 제외함

#### 2) 분(Minute) 단위 선분이력 변환

- 포인트 단위(밀리세컨 단위)의 로그를 선분 단위(분 단위) 로그로 변환
  - 원천데이터를 분 단위로 분할 생성하기 위하여, 동일 연속 셀ID를 그룹화 및 단일화하여 분 단위별 신호 정확도와 지속시간 등을 고려하여 대표 Raw를 생성함
  - 생성된 대표 Raw데이터를 기반으로 분단위 위치가 없는 시간에 대하여 모바일 단말기 Call이벤트(전화, 문자, 데이터 사용) 발생 위치를 기반으로 보강함

### 3) 통행정보 왜곡 데이터 보완

- Raw 정보를 10분 간격으로 대표위치를 선택하여 위치 변화를 고려하기 때문에 통신상의 핸드오버(Handover)나 시그널 점프(Signal Jump)가 발생할 수 있음
  - 이동초점을 기준으로 불필요한 핸드오버로 나타나는 위치를 보정함
  - 시그널 점프로 추정되는 위치를 이동 관점에서의 분단위 이동 각도(전후 위치 비교)를 통해서 판별하고 이에 대한 보정을 진행함
- 왜곡 데이터가 보정된 최종 데이터를 기반으로 시간대별 이동 거리, 이동시간, 이동패턴, 지형정보, 체류지에 대한 셀ID 등 테이블을 생성함
  - 이동 거리를 분 단위별 셀 중심점 기준으로 계산하고, 이동시간은 대표 셀로 선정된 데이터 간 시간 차이로 계산함
  - 이를 기반으로 이동속도(km/h)는 이전 시간대에서 현 위치로의 진입 이동속도로 셀 간 직선거리/시간으로 계산하여 구축함
  - 또한, 해당 이동속도를 기준으로 체류(0km/h), 보행(0초과 ~ 5km/h 미만), 비보행(5km/h 이상)으로 이동패턴을 구분함

### 4) 체류순서 부여

- 로그 기록 일자, 체류 시작시각, 체류 종료시각을 기준으로 로그 기록 순서에 따라 개인별 데이터를 정렬하여 체류 순서를 부여함
  - 고객식별번호(모바일 단말기 구분코드)와 기준일자를 기준으로 동일한 값을 갖는 데이터를 시간 순서대로 정렬함

### 5) 체류정보 추출

- 이동 중에 기록된 로그 기록을 제외하고 체류 중 로그 기록만을 추출함
  - 트립 타입 필드를 추가하여 체류시간이 30분 미만인 경우는 '경로(Pass-by)'로, 30분 초과인 경우는 '체류(Stay)'로 구분함

## 6) 체류지 유형 구분

- 출발지, 도착지에 대한 체류지는 집, 회사(학교), 관광지, 기타로 구성됨
  - 체류지 유형은 통신사에서 추정하는 고객의 추정거주지 셀ID, 추정 주중 주활동위치 셀ID, 관광지 Poi별 셀ID를 기반으로 구성되며 최종적으로 거주(HH), 근무(CC), 기타(XX)로 분류함
  - 거주지, 근무지 추정은 통신상의 오차 한계로 인하여 특별시와 광역시는 반경 500m, 기타 지역은 반경 1km를 기준으로 추정 거주지와 근무지를 판단함

## 7) 위치정보, 시간정보, 연령정보 범주화

- 개인의 이동궤적을 추적할 수 없도록 원천데이터에서 출발지와 도착지의 위치 정보를 500\*500m 격자 셀, 행정동 단위 등으로 변경한 후, 시간 정보를 한 시간 단위로 재생성하고, 1세 단위의 연령 정보를 10세 단위로 변경함

## 8) 통행량 집계

- 앞서 부여한 체류 순서에 따라 출발과 도착을 구분한 다음, 기점 정보(출발 일자, 출발 시간대(1시간 단위), 출발 위치(500m 격자/읍면동/시군구/시도), 출발 체류지 유형(거주지/근무지/기타)), 종점 정보(도착 시간대(1시간 단위), 도착 위치(500m 격자/읍면동/시군구/시도), 도착 체류지 유형(거주지/근무지/기타)), 통행자 정보(성, 연령)가 동일한 로그 기록을 집계함

## 8) 통행량 보정

- 통행량이 3이하인 경우 모두 3통행으로 치환한 후, 3통행 미만에 대해서는 보정계수를 적용하여 통행량을 재산정함
  - 개인정보보호법에 의거, 3통행 이하의 값은 실제 값이 아닌 추정값을 입력함
- 실제 통행 분석에 활용할 때 통신사의 시장점유율 기준으로 전수화하여 사용함
  - 과학기술정보통신부에 고시되고 있는 무선통신서비스 가입자 통계 정보 중 LTE 가입 정보를 활용하여 LG유플러스 시장점유율 산출 가능

## 나. KT

### 1) 원천데이터 수집 및 이상치 제거

- 2020년 1월 1일부터 2020년 12월 31일까지 기록된 로그를 수집
- 기지국 좌표가 Null값이거나 행정동 정보가 매칭되지 않는 데이터를 제거

### 2) 선분 이력으로 변환

- 포인트 단위의(밀리세컨 단위) 로그를 선분 단위의(분 단위) 로그로 변환
  - 닷(dot) 형태로 기록된 로그를 기록된 시간 순서대로 연결한 후, 최초 기록된 시간(이하 '체류 시작시간')과 마지막으로 기록된 시간(이하 '체류 종료시간') 정보만 추출하여 단말기가 해당 기지국에 식별된 시간(이하 '체류시간')을 산출
  - 체류시간은 체류 종료시간에서 체류 시작시간을 뺀 값이며, 초(sec) 단위는 생략



주: 자료를 참고하여 도식화 함

자료: 박미울·주은정(2017), LTE 시그널 정보를 이용한 위치정보가공 및 유동인구집계 방법, 한국통신학회 2017년도 동계종합학술발표회, p.285.

<그림 2-1> 선분 이력으로의 변환

### 3) 통행 정보 왜곡 데이터 보완

- 데이터 보완 범위를 한정하기 위해 로그 기록을 한 달 단위로 구분
  - 통행 정보 왜곡 데이터를 탐색하고 보완하는 범위를 1개월로 한정 (예: 4월 1일부터 4월 30일까지의 로그 기록을 연결하여 탐색하고 보완)

- 통행 정보를 왜곡할 수 있는 '필요 이상의 핸드오버 데이터(Unnecessary Handover)', '핑퐁 핸드오버(Ping-pong Handover)로 인한 데이터'를 본 과제에서 개발한 전처리 기술을 통해 보정함

#### 4) 체류순서 부여

- 로그 기록 일자, 체류 시작시간, 체류 종료시간을 기준으로 로그 기록 순서에 따라 개인별 데이터를 정렬하여 체류 순서를 구분
  - 고객식별번호(단말기 구분코드)와 데이터 생성일자를 기준으로 KEY 값을 갖는 필드를 형성한 후('기준일-고객식별번호'), 같은 KEY 값을 갖는 데이터 내에서 정렬

ID	그룹 번호	KEY 칼럼 (기준일- 고객식별번호)	가상기지국		체류 시작 시간	체류 종료 시간	연 령 대	체류 시간
			x좌표	y좌표				
1	1	20180520-17232311	126.914	37.545	1400	1750	33	240
2	2	20180520-17232311	126.914	37.549	1751	1752	33	1
3	3	20180520-17232311	126.916	37.548	1754	1809	33	15
4	4	20180520-17232311	126.914	37.549	1810	1813	33	3
5	5	20180520-17232311	126.915	37.551	1813	1814	33	1
6	6	20180520-17232311	126.914	37.549	1814	1815	33	1
		⋮	⋮	⋮				
18	18	20180520-17232311	126.815	37.538	1920	2369	33	279
		⋮	⋮	⋮				
453	1	20180520-1843030	126.934	37.556	0805	1041	25	156
454	2	20180520-1843030	126.936	37.556	1041	1042	25	1
455	3	20180520-1843030	126.935	37.554	1043	1044	25	1
456	4	20180520-1843030	126.933	37.552	1045	1050	25	5
457	5	20180520-1843030	126.929	37.547	1051	1052	25	1
458	6	20180520-1843030	126.923	37.547	1053	1055	25	2
459	7	20180520-1843030	126.918	37.547	1056	1059	25	3
460	8	20180520-1843030	126.914	37.545	1100	1200	25	60
		⋮	⋮	⋮				
679	227	20180522-1843030	126.934	37.556	0003	0830	25	507
680	228	20180522-1843030	126.936	37.556	0831	0833	25	2
		⋮	⋮	⋮				
711	259	20180522-1843030	126.815	37.551	0850	1100	25	130

〈그림 2-2〉 체류순서 부여 (개인별 이동궤적 형성)

### 5) 체류 정보 추출

- 이동 중에 기록된 로그 기록을 제외하고 체류 중에 기록된 로그 기록만을 추출
  - 체류시간이 15분을 초과하는 로그 기록만 추출
    - 트립 타입 필드를 추가하여 체류시간이 15분 이하인 경우는 '경로(Pass-by)'로, 15분 초과인 경우는 '체류(Stay)'로 구분

ID	그룹 번호	KEY 필드 (기준일- 고객식별번호)	가상기지국		체류 시작 시간	체류 종료 시간	연 령 대	체류 시간	트립 타입
			X좌표	Y좌표					
1	1	20160520-17232311	126,914	37,545	1400	1730	33	240	체류
2	2	20160520-17232311	126,914	37,549	1751	1752	33	1	경로
3	3	20160520-17232311	126,916	37,548	1754	1809	33	15	경로
4	4	20160520-17232311	126,914	37,549	1810	1813	33	3	경로
5	5	20160520-17232311	126,915	37,551	1813	1814	33	1	경로
6	6	20160520-17232311	126,914	37,549	1814	1815	33	1	경로
		⋮	⋮	⋮					
18	18	20160520-17232311	126,815	37,538	1900	2369	33	279	체류
		⋮	⋮	⋮					
453	1	20160520-1843030	126,934	37,556	0805	1041	25	156	체류
454	2	20160520-1843030	126,936	37,556	1041	1042	25	1	경로
455	3	20160520-1843030	126,935	37,554	1043	1044	25	1	경로
456	4	20160520-1843030	126,933	37,552	1045	1050	25	5	경로
457	5	20160520-1843030	126,939	37,547	1051	1052	25	1	경로
458	6	20160520-1843030	126,925	37,547	1053	1055	25	2	경로
459	7	20160520-1843030	126,918	37,547	1056	1059	25	3	경로
460	8	20160520-1843030	126,914	37,545	1100	1200	25	60	체류
		⋮	⋮	⋮					
679	227	20160522-1843030	126,934	37,556	0003	0830	25	507	체류
680	228	20160522-1843030	126,936	37,556	0831	0833	25	2	경로
		⋮	⋮	⋮					
711	259	20160522-1843030	126,815	37,561	0850	1100	25	130	체류

<그림 2-3> 이동정보와 체류정보 구분 (예시)

### 6) 체류지 유형 구분

- 체류지 식별 기준(《표 2-4》 참조)에 따라 체류지 유형을 구분하여 필드 추가

《표 2-4》 체류지 식별 기준

체류지 유형		식별 기준			
		체류특성		통행자 연령	비고
		체류 시간	체류 빈도		
주상주 지역	집	알고리즘 기준 최장 체류지 후보 1, 2 중	일주일 기준 최다빈도 체류지	-	알고리즘 기준 제1의 주체류지
	회사	3시간 이상	주 2회 이상	20세 이상	알고리즘 기준 제2의 주체류지 중
	학교			20세 미만	
잠재 활동 지역	정기적	-	주 2회 이상	-	-
	비정기적	집, 회사, 학교, 정기적 잠재활동지역으로 구분되지 않은 그 외의 체류지			

ID	그룹 번호	KEY 값 (기준일- 고객식별번호)	가상기차국		체류 시작 시간	체류 종료 시간	연 령 대	체류 시간	트립 타입	체류지 유형
			X좌표	Y좌표						
1	1	20160520-17232311	126,914	37,545	1400	1750	33	240	체류	집
2	2	20160520-17232311	126,914	37,549	1751	1752	33	1	경로	-
3	3	20160520-17232311	126,916	37,548	1754	1809	33	15	경로	-
4	4	20160520-17232311	126,914	37,549	1810	1813	33	3	경로	-
5	5	20160520-17232311	126,915	37,561	1813	1814	33	1	경로	-
6	6	20160520-17232311	126,914	37,549	1814	1815	33	1	경로	-
18	18	20160520-17232311	126,815	37,538	1920	2069	33	279	체류	회사
453	1	20160520-1843030	126,934	37,556	0805	1041	25	156	체류	집
454	2	20160520-1843030	126,936	37,556	1041	1042	25	1	경로	-
455	3	20160520-1843030	126,935	37,554	1043	1044	25	1	경로	-
456	4	20160520-1843030	126,933	37,552	1045	1050	25	5	경로	-
457	5	20160520-1843030	126,929	37,547	1051	1052	25	1	경로	-
458	6	20160520-1843030	126,923	37,547	1053	1055	25	2	경로	-
459	7	20160520-1843030	126,918	37,547	1056	1059	25	3	경로	-
460	8	20160520-1843030	126,914	37,545	1100	1200	25	60	체류	기타

《그림 2-4》 체류지 유형 구분 (예시)

### 7) 위치정보, 시간정보, 연령정보 변환

- 개인의 이동궤적을 추적할 수 없도록 원천데이터에서 출발지와 도착지의 위치 정보를 분석 맵의 교통폴리곤 ID로 변경한 후, 시간 정보(채류시작시간, 채류종료시간)에서 분 정보를 생략한 시간대 필드를 생성하고, 1세 단위의 연령 정보를 10세 단위로 변경함

ID	그룹 번호	KEY 칼럼 (기준일-고려식별번호)	가상기지국		채류 시작 시간	채류 종료 시간	연령 대	채류 시간	트립 타입	채류지 유형	위치 정보 변환 (교통폴리곤 ID)	채류시작 시간대	채류종료 시간대	연령대 변환
			X좌표	Y좌표										
1	1	20160520-17232311	126.914	37.545	1400	1750	33	240	채류	집	210001	14	17	30
2	2	20160520-17232311	126.914	37.540	1751	1752	33	1	경로	-	-	-	-	
3	3	20160520-17232311	126.916	37.548	1754	1809	33	15	경로	-	-	-	-	
4	4	20160520-17232311	126.914	37.548	1810	1813	33	3	경로	-	-	-	-	
5	5	20160520-17232311	126.915	37.551	1813	1814	33	1	경로	-	-	-	-	
6	6	20160520-17232311	126.914	37.549	1814	1815	33	1	경로	-	-	-	-	
18	18	20160520-17232311	126.815	37.538	1920	2359	33	279	채류	회사	230017	19	23	30
453	1	20160520-1843030	126.934	37.556	0805	1041	25	156	채류	집	311014	08	10	30
454	2	20160520-1843030	126.936	37.556	1041	1042	25	1	경로	-	-	-	-	
455	3	20160520-1843030	126.935	37.554	1043	1044	25	1	경로	-	-	-	-	
456	4	20160520-1843030	126.933	37.552	1045	1050	25	5	경로	-	-	-	-	
457	5	20160520-1843030	126.929	37.547	1051	1052	25	1	경로	-	-	-	-	
458	6	20160520-1843030	126.923	37.547	1053	1055	25	2	경로	-	-	-	-	
459	7	20160520-1843030	126.918	37.547	1056	1059	25	3	경로	-	-	-	-	
460	8	20160520-1843030	126.914	37.545	1100	1200	25	60	채류	기타	312123	11	12	30

<그림 2-5> 위치정보, 시간정보, 연령정보 변환 (예시)

### 8) 통행량 집계

- 앞서 부여한 체류순서에 따라 출발과 도착을 구분한 다음, 기점 정보(출발 일자, 출발 시간대, 출발 위치, 출발 체류지 유형), 종점 정보(도착일자, 도착 시간대, 도착 위치, 도착 체류지 유형), 통행자 정보(성, 연령)가 동일한 로그 기록을 집계함
- 마지막 체류지가 출발 정보로 구분되고, 도착 정보가 없는 경우 이동하지 않은 것으로 보고 '무통행'으로 간주함

ID	그룹 번호	KEY 필드 (기종일- 고려식별번호)	가상기지국		체류 시작 시간	체류 종료 시간	연 령 대	체류 시간	드롭 타입	체류지 유형	위치 정보 변환 (공통플러군 ID)	체류시작 시간대	체류종료 시간대	연령대 분할	
			X좌표	Y좌표											
1	1	20160520-1723311	126.914	37.546	1400	1750	33	240	체류	집	210001	14	17	30	출발
2	2	20160520-1723311	126.914	37.548	1751	1752	33	1	경로	-	-	-	-	-	1통행 ↓ 도착
3	3	20160520-1723311	126.916	37.548	1754	1809	33	15	경로	-	-	-	-	-	
4	4	20160520-1723311	126.914	37.549	1810	1813	33	3	경로	-	-	-	-	-	
5	5	20160520-1723311	126.915	37.561	1813	1814	33	1	경로	-	-	-	-	-	
6	6	20160520-1723311	126.914	37.549	1814	1815	33	1	경로	-	-	-	-	-	
18	18	20160520-1723311	126.815	37.538	1920	2369	33	279	체류	회사	230017	19	23	30	
453	1	20160520-1840030	126.934	37.560	0805	1041	25	150	체류	집	311014	08	10	20	출발
454	2	20160520-1840030	126.936	37.566	1041	1042	25	1	경로	-	-	-	-	-	1통행 ↓ 도착
456	3	20160520-1840030	126.936	37.564	1043	1044	25	1	경로	-	-	-	-	-	
456	4	20160520-1840030	126.933	37.562	1045	1050	25	5	경로	-	-	-	-	-	
457	5	20160520-1840030	126.929	37.547	1051	1052	25	1	경로	-	-	-	-	-	
458	6	20160520-1840030	126.923	37.547	1053	1055	25	2	경로	-	-	-	-	-	
458	7	20160520-1840030	126.918	37.547	1056	1059	25	3	경로	-	-	-	-	-	
460	8	20160520-1840030	126.914	37.545	1100	1300	25	60	체류	기타	312123	11	12	20	도착

〈그림 2-6〉 출발/도착 구분 및 통행량 집계 (예시)

### 9) 통행량 보정

- 통행량이 3 이하(0~3통행)인 경우 모두 3통행으로 변환한 후, 3통행 미만 보정 계수(2장 참고)를 적용하여 통행량을 보정
  - 개인정보보호법에 의거, 3 통행 이하의 값은 실제 값이 아닌, 추정값을 입력
- 실제 통행 분석에 활용할 때에는 통신업체의 시장점유율을 기준으로 전수화 하여 사용
  - 과학기술정보통신부에 고시되고 있는 무선통신서비스 가입자 통계 정보 중 LTE 가입 정보를 활용하여 KT 시장점유율 산출 가능

### 3. 모바일 통신데이터 기반 통행DB 전수화 방안

#### 1) 기존 방식의 전수화 방법

- 시장점유율 기반 전수화 계수 적용
  - 가장 기초적인 전수화 방법으로, 통신사 점유율을 기준으로 [총통행량 X (100/통신사 시장점유율)] 수식을 통해 전수화 계수 값 산출

기준일	출발				도착				성별	연령대	총통행량	평균통행 시간
	요일	시간대	공간	유형	요일	시간대	공간	유형				
20201221	2	10	A	집	2	11	B	학교	F	30	3	337
20201221	2	10	B	집	2	13	C	직장	F	30	2	172
20201221	2	10	C	집	2	13	D	기타	F	30	2	114
...	...	...	...	...	...	...	...	...	...	...	...	...

↓

기준일	출발				도착				성별	연령대	총 통행량	평균통행 시간	전수화 계수 (점유율)	전수화 반영 총 통행량
	요일	시간대	공간	유형	요일	시간대	공간	유형						
20201221	2	10	A	집	2	11	B	학교	F	30	3	337	3	9
20201221	2	10	B	집	2	13	C	직장	F	30	2	172	3	6
20201221	2	10	C	집	2	13	D	기타	F	30	2	114	3	6
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

〈그림 2-7〉 시장점유율 기반 전수화 방법

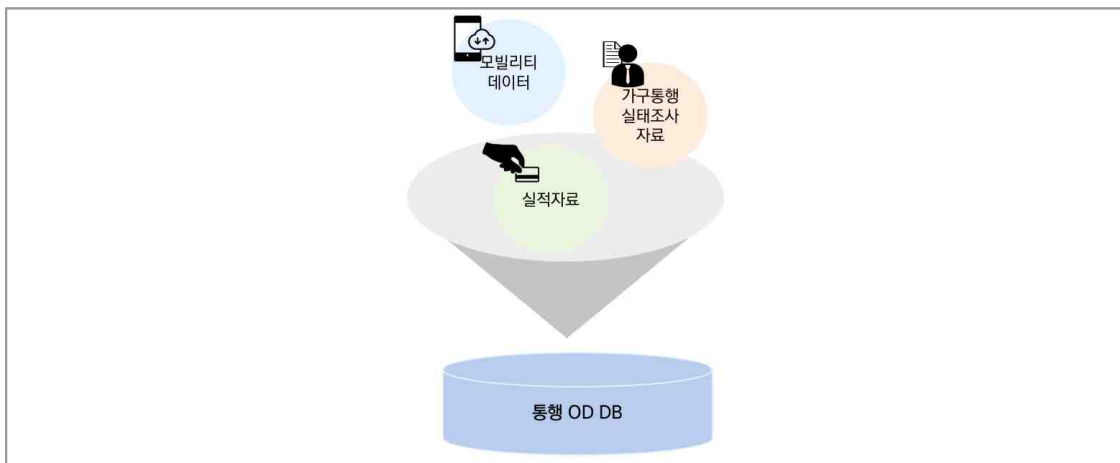
- 인구수 기반 전수화 계수 적용
  - 주민등록인구 정보와 통신사 고객의 거주지 정보를 활용하여 통신사 고객에 대한 통행량을 전국민의 통행량으로 변환
  - 성별, 연령대별, 지역별 속성을 반영한 전수화 방법은 아래 표와 같음

〈표 2-5〉 인구수 기반 전수화 방법

전수화 단계	설명
Step-1	전수화 계수 생성을 위한 주민등록인구 집계 데이터 수집
Step-2	통신데이터에 대하여 주민등록인구와 동년도 시군구(거주지)/성별/연령대 별의 인구 집계
Step-3	주민등록인구, 통신데이터의 집계 값의 비율을 통한 전수화 계수 값 산출 · 시군구/성별/연령대별 (주민등록인구 / 통신데이터 집계 값)을 통해 전수화 계수 값 산출
Step-4	개인 O/D 정보 중 거주지(H) 식별 및 해당 시군구 지역을 개인의 거주지역으로 선정
Step-5	개인별 거주지(H)를 기준으로 모든 개별통행 OD에 전수화 계수 값 반영
Step-6	통행 DB 기준에 따른 500m 단위의 격자 단위로 데이터로 집계 · 500m 단위의 격자 집계 시, 전수화 반영 후 총통행량 값은 개인별 전수화 보정 후 총통행량 값들의 합산 값으로 집계

## 2) 데이터 전수화 방향 설정

- 개인통행실태조사자료와 실적자료를 전수화에 활용
  - 모바일 통신데이터는 시·공간적 해상도가 높아 표본조사인 개인통행실태조사 자료의 한계를 보완할 수는 있지만, 통행에 대한 정보를 추정하는 것이기 때문에 실제 통행 정보와 일치하는지는 확인할 수 없음
  - 이에 반해 개인통행실태조사자료와 실적자료는 모바일 통신데이터에 비해 양적으로는 부족하지만, 실제 조사로 인한 결과값이기 때문에 참고자료로 활용 가능함
  - 따라서 본 사업에서는 개인통행실태조사자료와 실적자료를 활용하여 모바일 통신데이터의 한계를 보완하고 신뢰성을 향상시키고자 함
    - 미국에서는 2020년부터 가구통행실태조사(National Household Travel Survey, NHTS)자료와 모바일 데이터, 사회-경제 데이터를 융합하여 “Next-Generation OD”를 구축 중



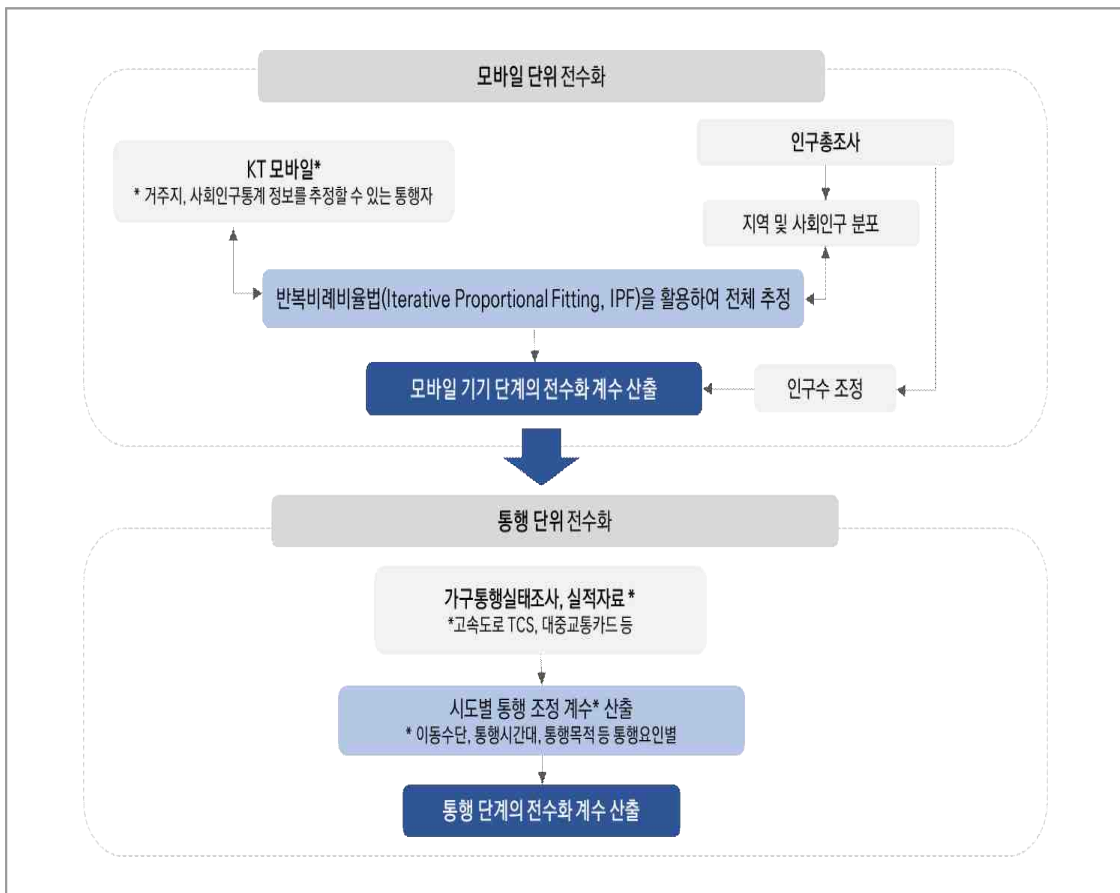
〈그림 2-8〉 데이터 융합

- 타 통신사 데이터에도 적용할 수 있는 보편적 전수화 알고리즘 개발
  - 통신 기술, 물리적인 조건 등으로 인해 원시데이터 형태, 원시데이터로부터 통행 정보(출발지, 도착지, 출발시간, 도착시간 등)를 추출하기까지 가공 과정이 통신사별로 상이함
  - 특정 통신사 데이터를 기준으로 전수화 알고리즘을 개발할 경우, 여타 통신사

데이터로는 전수화 된 통행량을 생산할 수 없다는 문제가 존재하므로 향후 제한이 없도록 보편적인 알고리즘을 개발하고자 함

○ 통행 단위의 전수화 과정 추가

- 기존 전수화 알고리즘은 일부분의 통행량(통신사 이용고객에 대한 통행)을 전체 통행량으로 늘리는 것을 주목적으로 하여 통행 특성(통행목적, 이동수단 등)은 고려되지 않음
- 기존 방식에 통행 단위로 전수화하는 과정을 추가하여 기존 방식의 한계를 보완하고자 함



주: 미국 Next-Generation OD 전수화 과정을 참조하여 재구성하였으며, 실제 활용되는 자료와 구체적인 방식은 추후 변경될 수 있음.

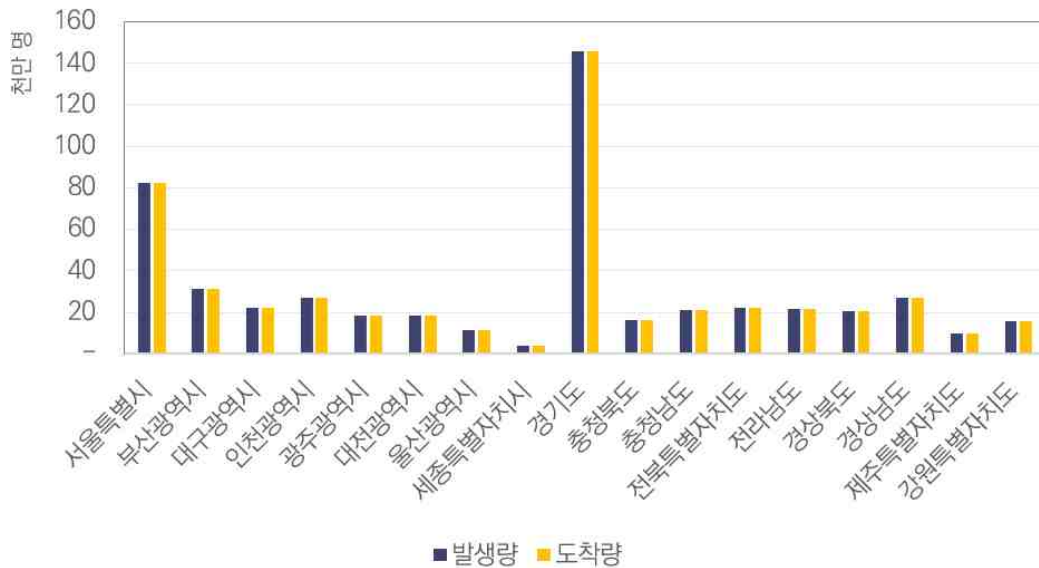
자료: Zhang, L., et al. "Next generation National Household Travel Survey National Origin Destination Data Passenger Origin-Destination Data Methodology Documentation." Federal Highway Administration.[Google Scholar] (2021), p.35.

<그림 2-9> 전수화 과정 도식화

#### 4. 모바일 통신데이터 기반 통행DB 구축 결과

○ 지역별 총 통행량

- 모든 시도별로 통행 발생, 도착량 차이가 크지 않은 것으로 나타남
- 시도별 총 통행량은 경기도, 서울특별시, 부산광역시, 인천광역시 순으로 높게 나타나 수도권, 대도시의 통행량이 큰 것으로 나타남
- 하지만 이 결과는 사전에 LG 유플러스 모바일 통신데이터 수집 과정에서 지역별 통행 발생, 도착량 비율을 유사하게 설정했을 가능성이 있으며, 지역별 LG 유플러스 점유율 차이가 있어 해석에 유의해야 함



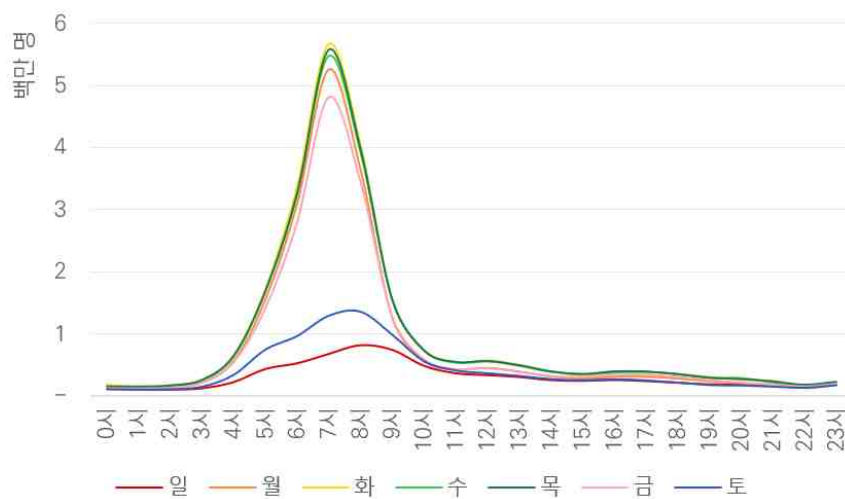
〈그림 2-10〉 지역별 총 통행량

<표 2-6> 지역별 총 통행량

시도	출발량	도착량
서울특별시	824,156,237	824,370,768
부산광역시	311,329,552	311,315,939
대구광역시	220,721,139	220,766,497
인천광역시	268,201,625	268,203,409
광주광역시	183,759,022	183,768,031
대전광역시	181,722,795	181,670,593
울산광역시	113,304,598	113,304,118
세종특별자치시	37,900,922	37,926,860
경기도	1,456,481,771	1,456,484,474
충청북도	161,309,974	161,280,579
충청남도	209,929,368	209,923,795
전북특별자치도	222,858,672	222,870,915
전라남도	216,967,105	216,860,144
경상북도	207,121,117	207,073,146
경상남도	269,707,453	269,698,602
제주특별자치도	97,716,351	97,706,200
강원특별자치도	154,843,088	154,806,721
합계	5,138,030,790	5,138,030,791

○ 시간대별 출근 통행량 (요일별)

- 주중은 주말 대비 출근통행 발생 비율이 집중되며, 첨두시간은 오전 7시로 나타남
- 주중 월~목 대비 금요일의 출근 통행량이 첨두시(오전 7시)에 가장 낮게 나타남
- 주말은 첨두시간이 오전 8시로 나타나며, 주중에 비해 첨두시 출근 통행량이 최소 3배~최대 6배까지 낮게 나타남



<그림 2-11> 시간대별 출근 통행량 (요일별)

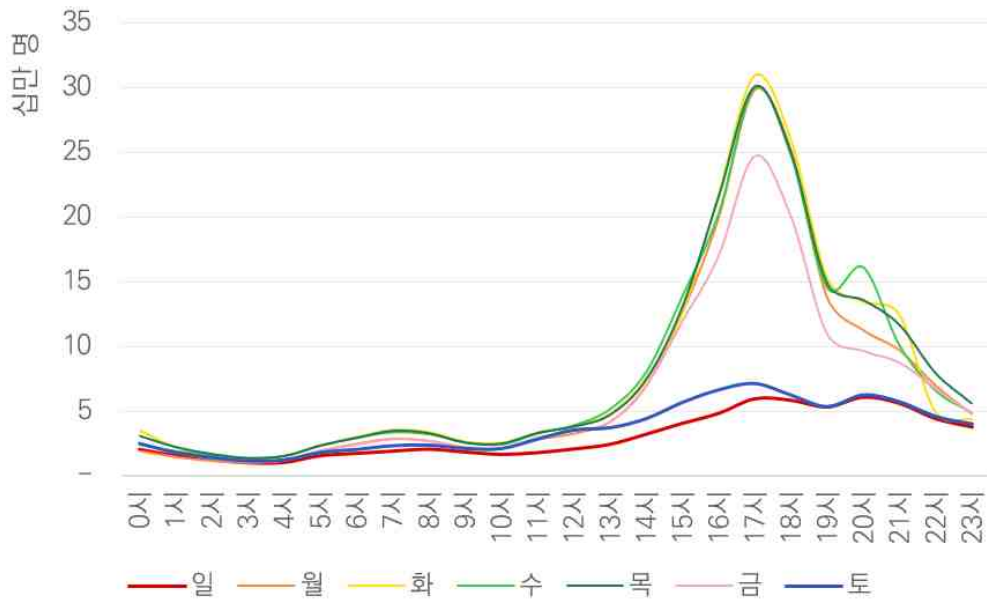
<표 2-7> 시간대별 출근 통행량 (요일별)

단위 : 통행량(인)

구분	일	월	화	수	목	금	토
0시	104,795	121,756	176,959	119,055	148,444	122,979	114,924
1시	98,060	121,496	144,476	126,431	146,395	122,011	106,846
2시	96,179	137,556	164,939	160,186	164,783	134,170	112,042
3시	116,442	218,204	255,036	253,369	253,146	217,869	153,672
4시	218,005	568,347	645,091	641,119	642,785	547,048	343,202
5시	429,645	1,558,404	1,718,407	1,692,709	1,679,967	1,417,725	747,426
6시	521,128	3,101,776	3,366,027	3,281,363	3,245,123	2,798,719	964,986
7시	673,210	5,264,090	5,662,665	5,478,748	5,570,686	4,812,350	1,290,195
8시	813,673	3,648,621	4,038,452	3,916,199	3,972,985	3,447,589	1,354,763
9시	738,225	1,260,058	1,540,063	1,554,335	1,538,555	1,272,789	985,485
10시	480,645	572,350	726,265	740,459	728,228	594,929	560,445
11시	360,616	428,174	535,042	540,657	542,418	434,690	413,518
12시	329,864	454,612	565,806	558,914	558,596	448,891	369,399
13시	300,166	397,687	494,347	490,011	489,997	388,430	327,741
14시	249,133	316,561	386,910	390,707	390,294	323,329	274,907
15시	237,860	286,112	330,701	354,818	348,025	308,982	258,028
16시	251,247	314,809	368,919	380,486	390,855	338,891	274,179
17시	233,547	313,940	373,450	390,513	389,511	345,665	255,271
18시	206,849	282,822	347,705	359,398	342,903	309,878	221,206
19시	181,741	239,694	300,187	295,525	287,689	257,132	179,917
20시	178,979	208,855	255,629	291,593	266,786	216,300	172,978
21시	165,425	186,672	235,383	220,099	231,053	184,337	154,853
22시	147,287	161,350	119,241	154,856	175,966	155,450	135,925
23시	181,199	198,726	177,278	191,709	222,385	194,337	176,513
계	7,313,921	20,362,672	22,928,979	22,583,257	22,727,577	19,394,490	9,948,421

○ 시간대별 퇴근 통행량 (요일별)

- 퇴근 통행은 주로 늦은 오후 시간에 이루어지며, 주중 퇴근 통행 발생량은 오후 5시에 집중된 것으로 나타남
- 첨두시(오후 5시)와 그 이후 시간대에도 주중 중에서는 금요일의 퇴근 통행량이 가장 낮음



<그림 2-12> 시간대별 퇴근 통행량 (요일별)

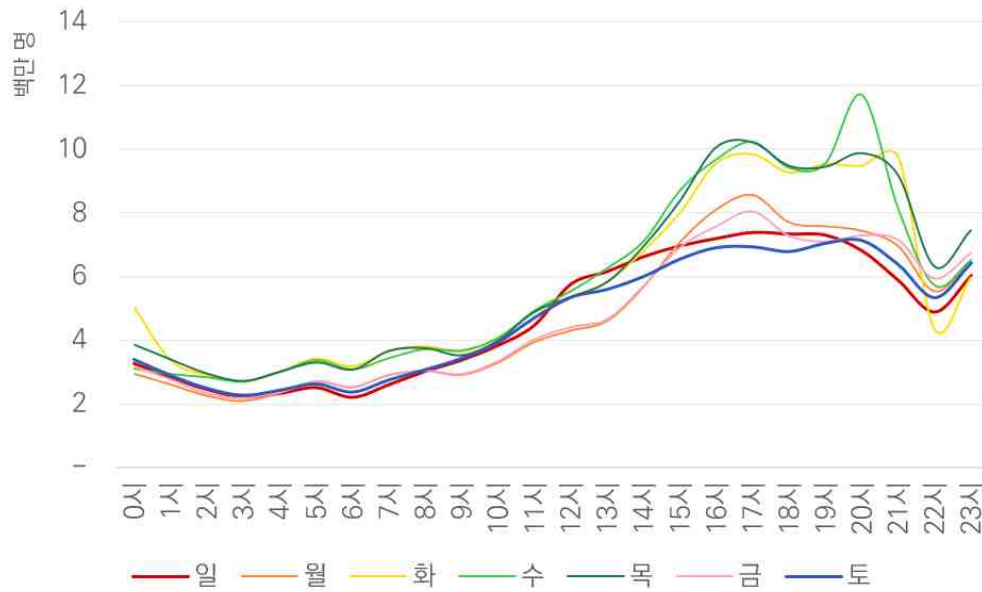
〈표 2-8〉 시간대별 퇴근 통행량 (요일별)

단위 : 통행량(인)

구분	일	월	화	수	목	금	토
0시	208,292	191,352	349,924	241,619	307,883	253,587	251,178
1시	159,791	145,302	217,340	191,024	222,214	179,838	181,924
2시	131,211	114,519	162,931	161,605	168,921	134,951	143,280
3시	105,681	99,835	135,231	136,999	137,586	111,459	116,922
4시	109,382	120,302	153,220	153,889	154,318	123,022	125,578
5시	160,758	194,472	235,792	234,059	235,511	194,853	182,490
6시	178,007	247,532	298,753	295,808	294,043	243,047	205,366
7시	195,094	286,920	353,853	336,638	349,085	284,585	234,439
8시	210,022	270,896	336,012	322,404	329,100	269,571	239,155
9시	188,196	219,226	263,078	259,522	256,473	215,479	212,593
10시	170,541	212,235	254,259	243,745	249,364	212,505	212,304
11시	183,508	281,578	332,842	327,494	332,550	296,305	285,795
12시	211,682	326,258	386,741	390,830	380,777	345,543	355,944
13시	247,299	416,536	477,192	516,564	472,175	417,556	373,295
14시	324,233	696,069	753,169	802,197	744,351	699,189	443,541
15시	407,068	1,261,738	1,263,857	1,386,495	1,304,809	1,193,594	569,384
16시	482,651	1,991,726	2,169,671	2,023,503	2,154,093	1,696,893	665,834
17시	593,881	2,983,878	3,089,449	2,997,714	3,007,840	2,467,401	714,445
18시	581,927	2,520,396	2,595,122	2,478,366	2,503,258	2,003,509	624,730
19시	531,538	1,381,166	1,528,558	1,452,993	1,490,642	1,097,172	536,144
20시	606,346	1,126,807	1,344,116	1,611,798	1,359,596	964,667	626,686
21시	557,337	971,224	1,243,070	1,011,466	1,167,668	875,355	573,073
22시	441,068	699,979	490,265	660,714	794,788	675,508	459,685
23시	378,741	477,352	435,193	488,259	560,135	484,227	402,154
계	7,364,254	17,237,299	18,869,638	18,725,704	18,977,178	15,439,816	8,735,941

○ 시간대별 귀가 통행량 (요일별)

- 귀가 통행 발생량은 오전보다 오후에 집중된 것으로 나타남
- 주중 화~목요일의 오후 8시~9시 통행량이 두드러지는데, 이는 직장에서 업무를 마친 뒤 다른 장소에 방문하고, 이후 집으로 돌아오는 통행량으로 판단됨



<그림 2-13> 시간대별 귀가 통행량 (요일별)

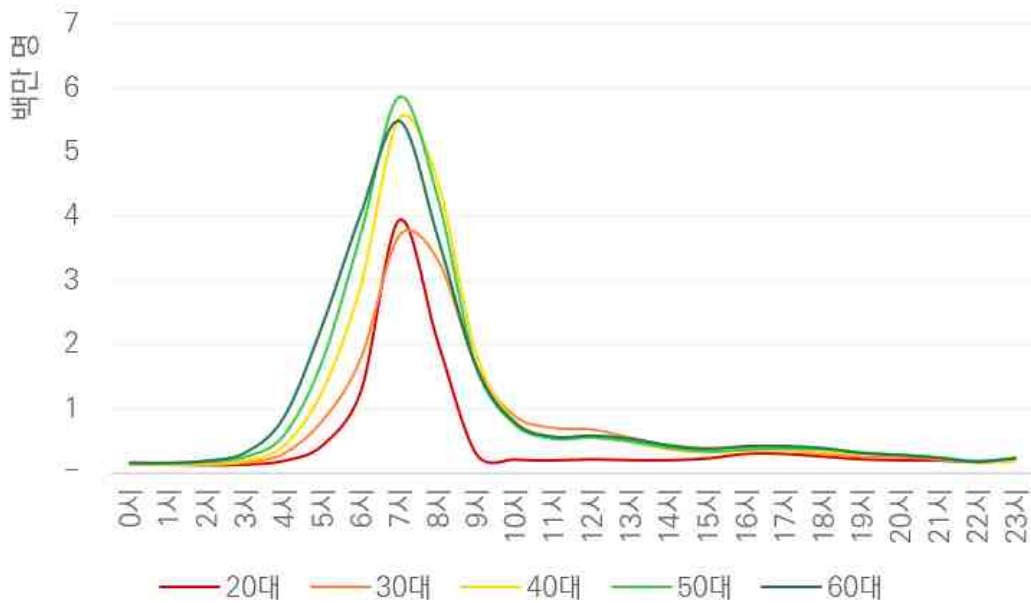
<표 2-9> 시간대별 귀가 통행량 (요일별)

단위 : 통행량(인)

구분	일	월	화	수	목	금	토
0시	3,271,521	2,946,197	5,009,189	3,109,528	3,856,459	3,160,868	3,385,322
1시	2,842,536	2,610,662	3,345,798	2,943,167	3,393,341	2,755,890	2,883,834
2시	2,457,523	2,255,225	2,882,978	2,840,959	2,952,177	2,319,395	2,478,741
3시	2,228,868	2,096,917	2,678,914	2,710,603	2,722,680	2,178,317	2,255,907
4시	2,334,666	2,351,829	2,998,152	3,012,499	3,012,593	2,379,442	2,416,543
5시	2,519,681	2,706,772	3,406,704	3,378,220	3,302,671	2,699,966	2,614,136
6시	2,214,249	2,522,649	3,172,191	3,091,207	3,081,503	2,527,523	2,358,786
7시	2,608,976	2,911,504	3,649,120	3,444,854	3,668,120	2,905,427	2,743,627
8시	3,026,196	3,040,496	3,798,801	3,728,242	3,752,455	3,030,598	3,062,271
9시	3,377,724	2,913,110	3,627,908	3,685,128	3,529,619	2,931,456	3,424,040
10시	3,836,423	3,302,850	4,074,578	4,076,465	3,996,388	3,337,763	3,926,156
11시	4,466,575	3,947,174	4,859,989	4,939,849	4,895,921	4,024,604	4,698,155
12시	5,748,623	4,299,976	5,366,415	5,547,846	5,352,294	4,403,231	5,340,275
13시	6,143,570	4,607,285	5,824,334	6,267,934	5,831,438	4,654,630	5,592,007
14시	6,609,357	5,672,493	6,808,858	7,116,743	6,948,973	5,706,722	5,996,042
15시	6,944,667	7,070,196	8,001,196	8,711,288	8,374,167	6,913,115	6,545,953
16시	7,175,095	8,088,893	9,568,243	9,686,716	10,054,966	7,575,665	6,905,416
17시	7,365,022	8,547,111	9,837,607	10,237,506	10,208,548	8,041,887	6,929,604
18시	7,312,829	7,690,445	9,267,150	9,426,524	9,470,775	7,275,454	6,781,967
19시	7,278,060	7,567,561	9,533,628	9,557,994	9,441,701	7,097,850	7,044,316
20시	6,787,766	7,425,441	9,482,226	11,712,926	9,868,348	7,284,086	7,126,255
21시	5,859,987	6,945,120	9,767,223	8,149,619	9,176,667	7,153,214	6,362,351
22시	4,877,469	5,531,593	4,342,480	5,731,790	6,309,607	5,933,853	5,330,493
23시	6,018,584	6,458,521	5,964,039	6,517,898	7,447,469	6,736,674	6,420,591
계	113,305,966	113,510,021	137,267,721	139,625,504	140,648,879	113,027,631	112,622,787

○ 시간대별 출근 통행량 (연령대별)

- 모든 연령대의 출근 통행량은 오전 7시 전후로 높게 나타남
- 첨두시(오전 7시)에 40~60대 출근 통행량이 20~30대 대비 높게 나타나는데, 이는 40~60대의 경제활동 종사자 비율이 높거나 혹은 모바일 통신데이터에서 해당 연령대가 차지하는 비율이 높기 때문으로 판단됨
- 연령대가 높을수록 출근 통행이 이른 시간에 발생함



<그림 2-14> 시간대별 출근 통행량 (연령대별)

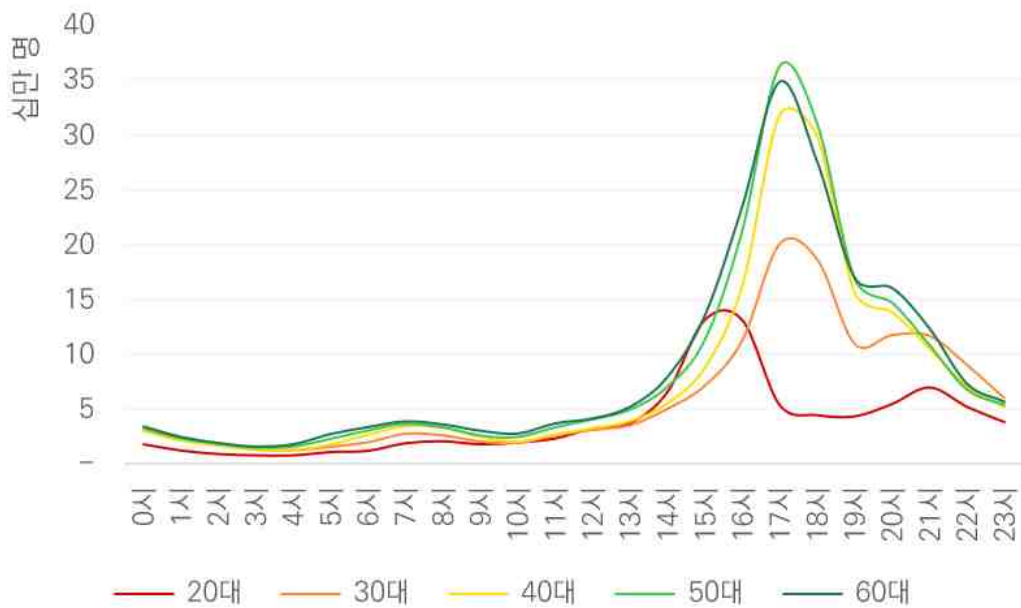
〈표 2-10〉 시간대별 출근 통행량 (연령대별)

단위 : 통행량(인)

구분	20대	30대	40대	50대	60대
0시	133,223	126,057	131,567	155,424	153,684
1시	114,581	120,872	122,245	145,141	153,548
2시	107,948	125,171	128,878	160,837	188,047
3시	118,692	153,154	182,289	241,619	313,633
4시	176,808	294,383	414,760	581,003	864,130
5시	416,422	804,688	1,270,818	1,743,154	2,299,412
6시	1,259,158	1,776,019	2,919,967	3,730,104	4,038,538
7시	3,929,495	3,694,032	5,532,566	5,861,224	5,480,914
8시	2,026,256	3,311,807	4,531,652	4,276,165	3,662,955
9시	280,679	1,655,567	1,836,519	1,636,741	1,640,547
10시	196,396	875,436	800,581	745,500	781,326
11시	183,996	693,313	554,978	527,734	559,092
12시	199,603	668,774	560,859	544,508	579,314
13시	187,909	541,058	475,021	483,591	530,026
14시	187,385	418,146	357,460	383,539	426,174
15시	213,806	377,691	310,952	328,284	372,858
16시	286,652	386,101	337,342	354,745	415,029
17시	285,817	356,332	351,834	374,638	415,008
18시	244,051	292,780	314,995	364,117	384,883
19시	202,657	245,974	280,048	313,289	305,601
20시	186,689	239,671	262,539	277,281	275,045
21시	182,300	202,036	223,496	237,310	232,063
22시	163,891	154,740	156,145	175,890	172,126
23시	214,702	177,417	186,056	224,423	227,182
계	11,499,115	17,691,219	22,243,569	23,866,261	24,471,134

○ 시간대별 퇴근 통행량 (연령대별)

- 20대를 제외한 모든 연령대의 퇴근 통행량은 오후 5~6시에 가장 높게 나타남
- 20대는 오후 3~4시에 퇴근 통행량이 높게 나타나는데, 이는 20대의 경우 학업 또는 파트타임 근로 형태가 통상적인 업무시간(오전 9시~오후 6시)을 갖는 직장인의 근무 패턴과는 다르기 때문으로 판단됨



<그림 2-15> 시간대별 퇴근 통행량 (연령대별)

〈표 2-11〉 시간대별 퇴근 통행량 (연령대별)

단위 : 통행량(인)

구분	20대	30대	40대	50대	60대
0시	172,388	314,904	300,587	321,893	336,786
1시	115,479	222,026	214,086	230,740	241,264
2시	82,938	171,765	166,176	182,177	185,469
3시	71,380	126,578	129,034	146,275	152,809
4시	71,300	120,803	129,127	156,819	174,120
5시	101,470	152,631	176,797	226,812	267,648
6시	112,822	193,184	267,096	303,135	331,189
7시	182,781	271,391	344,260	362,939	383,058
8시	200,070	256,767	332,442	331,009	355,330
9시	173,115	200,613	246,415	256,583	296,998
10시	189,039	189,046	203,234	243,275	271,666
11시	226,341	256,529	270,851	332,395	364,809
12시	317,853	306,060	325,370	403,569	409,587
13시	367,033	348,709	390,413	492,356	515,639
14시	653,454	499,433	553,306	702,793	790,092
15시	1,314,795	712,044	880,160	1,131,222	1,353,486
16시	1,310,051	1,126,239	1,631,222	2,131,866	2,350,585
17시	529,692	2,006,545	3,186,398	3,626,614	3,486,024
18시	440,038	1,862,080	2,979,445	3,095,392	2,751,293
19시	429,179	1,091,813	1,563,657	1,681,549	1,693,975
20시	542,432	1,172,507	1,384,723	1,466,883	1,602,374
21시	693,482	1,164,877	1,052,435	1,079,516	1,231,190
22시	516,278	899,141	706,110	674,554	726,648
23시	375,899	595,216	516,706	536,273	561,716
계	9,189,309	14,260,902	17,950,051	20,116,640	20,833,755

○ 시간대별 귀가 통행량 (연령대별)

- 시간대별 귀가 통행 발생량은 20대를 제외한 모든 연령대에서 오후 5시, 8시에 높은 양봉형으로 나타남
- 20대 귀가 통행량은 오후 4~9시까지 높게 나타남



<그림 2-16> 시간대별 귀가 통행량 (연령대별)

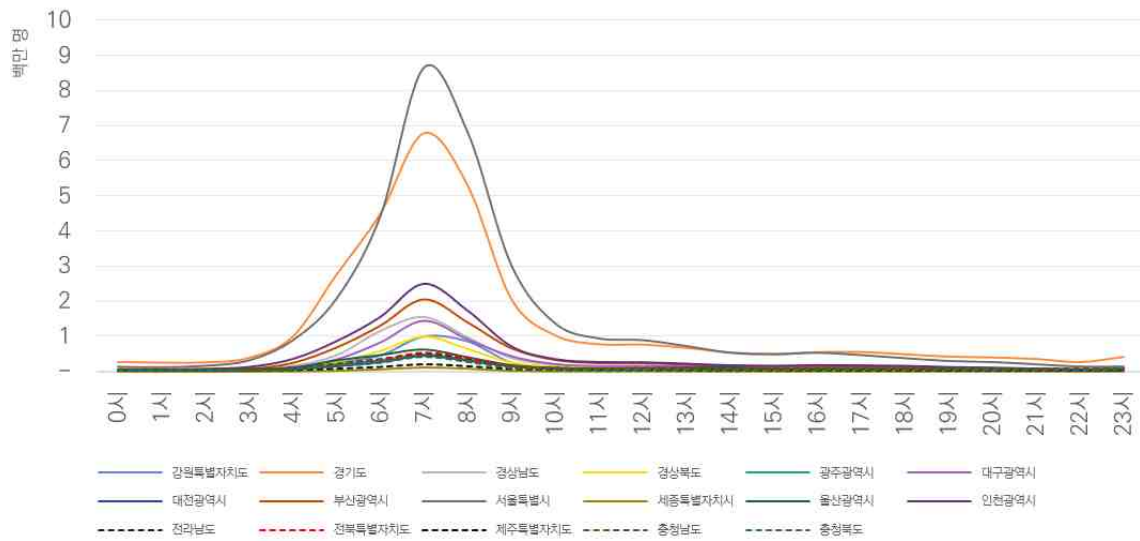
〈표 2-12〉 시간대별 귀가 통행량 (연령대별)

단위 : 통행량(인)

구분	20대	30대	40대	50대	60대
0시	2,930,457	4,329,298	4,252,053	4,379,817	4,072,701
1시	2,254,091	3,635,837	3,510,038	3,620,456	3,450,749
2시	1,875,191	3,000,623	2,943,880	3,125,838	3,097,455
3시	1,705,536	2,526,370	2,593,192	2,867,882	2,952,621
4시	1,811,615	2,418,601	2,668,839	3,101,258	3,340,579
5시	1,972,340	2,380,354	2,791,457	3,500,146	3,875,153
6시	1,826,983	1,963,254	2,512,986	3,335,330	3,566,822
7시	3,007,360	2,145,820	2,851,011	3,742,690	3,809,937
8시	3,577,941	2,192,274	2,977,045	3,766,494	3,867,433
9시	2,575,177	2,259,969	2,995,205	3,799,014	4,044,545
10시	2,428,160	2,654,547	3,327,604	4,287,273	4,536,747
11시	3,073,556	3,319,886	3,998,298	5,139,169	5,350,132
12시	4,058,969	3,687,143	4,541,005	5,890,444	6,055,377
13시	5,166,901	3,751,637	4,750,383	6,235,160	6,431,037
14시	6,914,781	4,238,084	5,307,719	6,901,235	7,255,254
15시	8,923,335	4,800,582	6,163,630	7,890,823	8,495,578
16시	9,695,516	5,726,504	7,368,591	9,264,398	9,855,167
17시	9,180,412	6,670,065	8,534,913	10,179,397	10,462,275
18시	9,330,443	6,540,348	8,232,323	9,579,589	9,813,446
19시	9,207,024	7,001,451	8,616,666	9,787,229	10,122,252
20시	9,229,037	8,157,912	9,366,157	10,410,140	10,679,860
21시	8,938,917	8,302,466	8,493,750	9,139,144	9,179,605
22시	5,970,889	6,618,642	6,264,715	6,491,922	6,291,349
23시	6,083,521	7,417,676	7,499,764	7,982,643	7,650,010
계	121,738,152	105,739,341	122,561,224	144,417,490	148,256,084

○ 시간대별 출근 통행량 (지역별)

- 시간대별 출근 통행량은 모든 지역에서 오전 7시에 가장 높게 나타나며, 서울특별시, 경기도, 인천광역시, 부산광역시, 경상남도, 대구광역시 순으로 나타남
- 이렇게 지역별 출근 통행량이 수도권, 대도시에 높게 나타나는 것은 지역의 활발한 취업 기회나 높은 취업자 수에 기인할 수 있음



〈그림 2-17〉 시간대별 출근 통행량 (지역별)

〈표 2-13〉 시간대별 출근 통행량 (지역별) - 특별시, 광역시, 특별자치시

단위 : 통행량(인)

구분	서울	부산	대구	인천	광주	대전	울산	세종
0시	139,718	37,945	26,297	45,620	48,403	39,254	14,191	7,436
1시	129,167	38,936	24,777	46,094	45,925	37,056	13,145	7,012
2시	170,291	50,039	29,589	64,393	47,147	36,454	17,010	6,529
3시	323,427	85,405	50,371	118,088	52,245	42,339	29,174	7,911
4시	896,757	256,287	121,380	346,712	85,602	68,379	87,234	13,914
5시	2,084,187	692,085	340,713	857,711	198,624	158,486	303,943	30,893
6시	4,424,832	1,310,350	820,035	1,550,849	315,299	270,997	452,445	70,131
7시	8,662,044	2,058,123	1,433,666	2,501,309	476,957	445,350	611,007	118,460
8시	6,808,907	1,403,734	904,556	1,729,063	330,578	307,352	392,121	87,464
9시	3,027,741	668,984	431,340	705,799	148,064	149,132	166,954	29,475
10시	1,376,604	338,353	209,153	341,876	86,292	83,314	83,001	16,773
11시	949,597	249,340	151,107	259,596	70,696	65,841	64,674	13,821
12시	900,821	245,070	151,457	252,078	74,660	68,727	72,232	16,385
13시	732,481	205,317	135,140	211,578	71,621	61,001	90,066	13,387
14시	541,547	157,788	102,928	164,098	64,034	54,112	104,192	10,999
15시	496,779	148,297	95,224	152,435	60,344	52,502	49,859	11,537
16시	541,544	158,767	95,841	169,715	65,933	58,219	52,262	13,308
17시	473,448	152,711	99,696	163,187	68,340	58,646	53,199	12,685
18시	387,307	136,247	98,363	149,472	67,732	59,036	46,703	12,345
19시	306,471	107,729	82,196	116,242	67,422	58,810	37,206	12,256
20시	276,659	97,622	67,682	97,916	67,713	58,403	33,886	11,671
21시	212,730	72,674	54,674	71,042	65,039	55,399	32,859	11,690
22시	146,844	52,211	40,547	46,728	55,365	48,117	21,749	9,775
23시	149,503	64,781	49,437	53,786	67,912	58,694	23,925	13,606
계	34,159,407	8,788,798	5,616,169	10,215,386	2,701,948	2,395,619	2,853,039	559,465

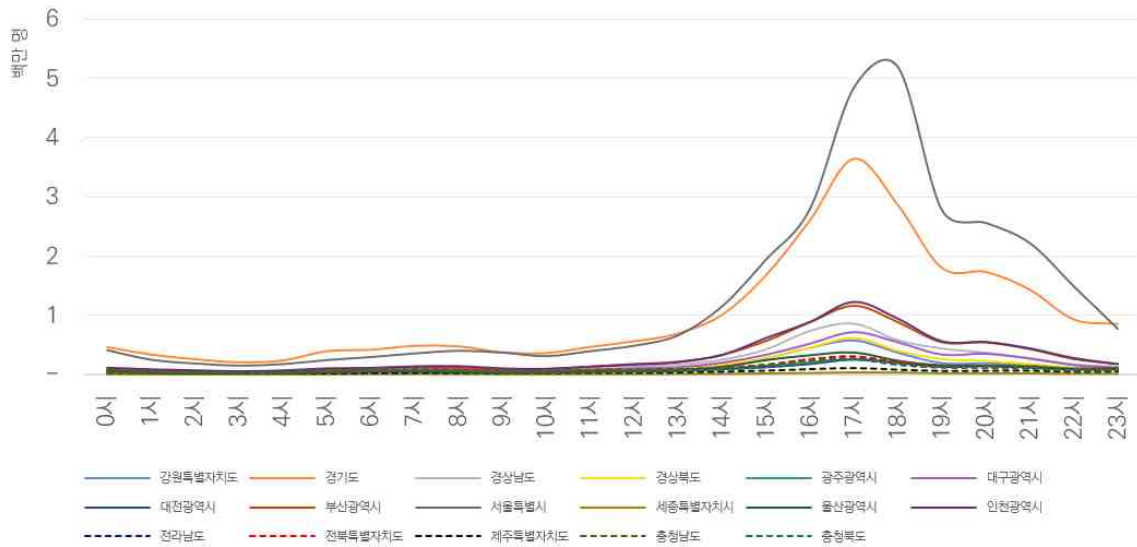
<표 2-14> 시간대별 출근 통행량 (지역별) - 도, 특별자치도(계속)

단위 : 통행량(인)

구분	경기	충북	충남	전북	전남	경북	경남	제주	강원
0시	279,643	32,080	45,923	43,829	44,882	25,386	34,846	17,822	25,635
1시	265,801	29,881	43,534	42,157	43,812	24,271	33,618	17,961	22,565
2시	274,163	31,417	42,636	43,113	45,638	26,401	40,913	18,167	25,953
3시	401,089	36,104	51,265	50,308	55,237	38,470	65,044	21,456	39,805
4시	1,006,134	66,032	86,177	80,376	93,621	92,998	167,835	37,522	98,638
5시	2,766,642	150,730	203,334	188,179	209,105	267,238	464,278	75,434	252,699
6시	4,495,716	248,373	317,085	331,838	350,920	584,766	1,148,974	132,250	454,261
7시	6,784,948	414,625	463,478	530,065	512,212	993,672	1,543,622	210,006	992,399
8시	5,292,925	268,302	314,045	374,875	362,225	640,407	971,723	151,367	852,636
9시	2,069,247	103,292	134,127	147,634	142,092	253,739	389,615	67,479	254,796
10시	1,039,641	62,812	83,077	89,023	88,363	133,652	197,941	42,321	131,126
11시	788,817	52,638	69,227	72,470	74,806	98,916	149,033	34,798	89,735
12시	779,037	59,774	77,351	82,515	82,552	115,289	162,004	35,297	110,835
13시	686,041	55,218	75,729	78,145	80,445	111,562	154,143	30,607	95,898
14시	553,773	47,553	68,097	69,501	71,554	90,042	130,948	26,261	74,415
15시	515,310	47,890	67,259	62,378	69,427	82,382	117,710	25,902	69,289
16시	561,803	54,211	73,086	71,890	75,921	94,739	130,226	27,290	74,630
17시	568,194	59,012	77,985	75,383	84,093	119,264	134,611	27,391	74,052
18시	502,024	58,504	74,268	75,875	76,847	98,973	134,184	26,456	66,426
19시	435,401	56,205	73,450	70,818	71,412	75,416	96,218	25,502	49,131
20시	411,280	51,692	71,844	69,055	70,671	61,318	79,377	25,418	38,913
21시	369,339	49,004	68,493	65,284	68,982	54,643	66,085	25,345	34,540
22시	281,973	41,671	56,484	53,334	56,812	42,103	48,761	21,812	25,788
23시	425,228	50,284	72,868	68,713	75,722	50,063	57,578	28,430	31,617
계	31,554,171	2,127,304	2,710,821	2,836,759	2,907,350	4,175,711	6,519,289	1,152,294	3,985,785

○ 시간대별 퇴근 통행량 (지역별)

- 시간대별 퇴근 통행량은 출근 통행량과 동일하게 서울특별시, 경기도, 인천광역시, 부산광역시, 경상남도, 대구광역시 순으로 높게 나타남
- 시간대별 퇴근 통행량은 서울특별시를 제외한 모든 지역에서 오후 5시에 가장 높게 나타나며, 서울시는 오후 6시에 가장 높게 나타남



〈그림 2-18〉 시간대별 퇴근 통행량 (지역별)

〈표 2-15〉 시간대별 퇴근 통행량 (지역별) - 특별시, 광역시, 특별자치시

단위 : 통행량(인)

구분	서울	부산	대구	인천	광주	대전	울산	세종
0시	423,931	98,884	64,191	111,364	70,749	57,345	69,677	9,543
1시	263,188	76,370	49,438	84,096	52,704	46,420	26,408	7,657
2시	199,741	60,994	35,793	67,353	44,499	36,474	20,099	6,378
3시	163,038	49,255	28,856	52,139	38,214	32,047	16,098	5,350
4시	188,711	50,670	29,399	65,206	39,754	33,061	17,903	5,965
5시	256,992	84,375	50,148	100,166	56,639	46,027	33,544	7,618
6시	307,745	108,525	73,693	108,444	65,679	56,935	50,853	10,732
7시	368,915	129,250	100,127	134,986	68,114	65,311	41,702	12,766
8시	413,058	119,954	94,503	138,001	66,656	59,159	37,839	11,567
9시	382,487	94,171	69,790	99,292	54,630	49,419	30,176	8,739
10시	323,576	98,382	68,833	92,840	54,592	48,637	33,158	8,477
11시	408,454	139,292	92,252	130,892	61,004	58,718	47,378	11,581
12시	499,589	184,217	112,792	165,861	68,566	62,461	60,889	11,510
13시	675,839	222,272	129,542	205,723	80,358	69,576	70,982	14,374
14시	1,166,221	339,577	207,229	328,516	103,318	95,000	126,709	17,921
15시	1,954,990	572,216	340,010	617,334	146,908	132,362	247,560	25,236
16시	2,814,250	886,594	528,873	891,515	210,159	183,019	331,787	34,468
17시	4,854,743	1,162,071	722,218	1,230,005	253,236	257,744	379,292	44,617
18시	5,193,815	885,313	549,906	942,004	192,263	200,638	235,019	42,868
19시	2,783,029	549,109	343,968	560,313	143,572	137,428	152,467	28,357
20시	2,568,982	546,991	357,591	548,229	150,120	147,530	148,882	27,506
21시	2,227,624	434,797	274,175	440,057	143,163	126,807	124,686	26,083
22시	1,500,514	267,856	165,933	276,190	102,520	90,343	88,753	19,544
23시	782,474	185,059	115,907	172,844	110,998	97,023	100,401	19,540
계	30,721,905	7,346,193	4,605,167	7,563,371	2,378,414	2,189,484	2,492,262	418,398

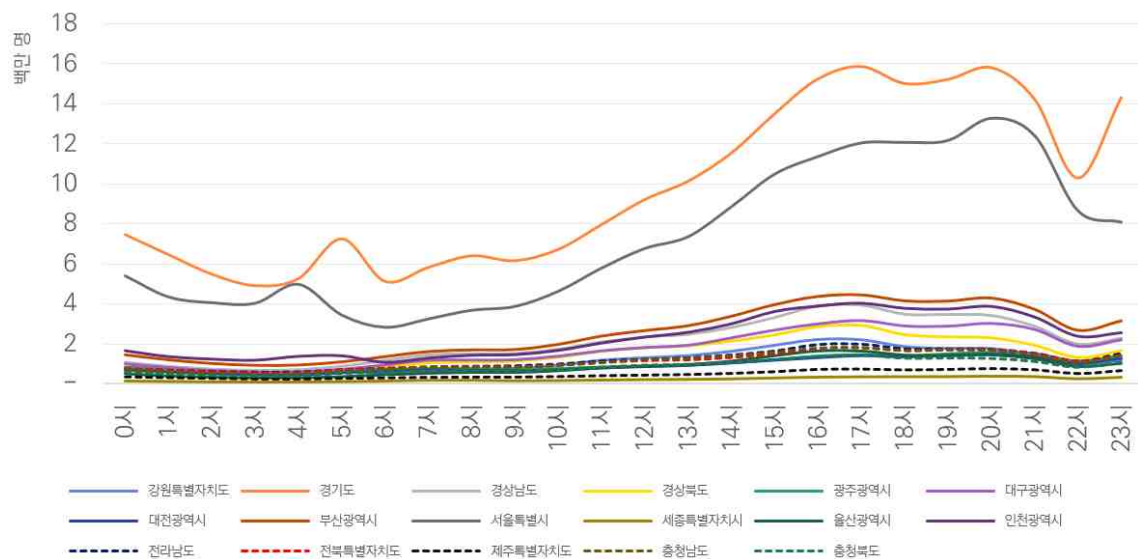
<표 2-16> 시간대별 퇴근 통행량 (지역별) - 도, 특별자치도(계속)

단위 : 통행량(인)

구분	경기	충북	충남	전북	전남	경북	경남	제주	강원
0시	472,124	43,148	58,625	59,043	58,318	50,976	77,234	26,522	52,159
1시	348,271	34,942	47,784	49,459	49,619	39,579	65,036	21,470	34,990
2시	270,934	29,618	40,861	41,751	42,991	29,626	46,642	17,095	26,569
3시	219,241	27,804	36,751	37,186	40,267	24,059	36,781	14,562	22,067
4시	249,437	28,602	40,135	38,342	43,821	28,476	40,980	14,688	24,562
5시	406,227	45,688	57,161	55,413	58,343	49,088	69,023	17,295	44,188
6시	428,785	58,259	77,080	70,303	73,291	101,829	102,661	20,996	46,744
7시	495,641	66,168	85,802	78,655	75,578	111,596	122,000	24,890	59,114
8시	486,295	55,239	71,288	73,804	71,341	83,940	101,903	24,566	68,049
9시	385,998	42,255	59,230	59,311	60,773	61,830	78,552	21,650	56,265
10시	374,169	42,370	58,467	60,885	60,862	64,796	84,252	22,084	58,575
11시	478,144	52,324	69,259	77,629	77,460	95,405	124,203	26,841	89,237
12시	571,177	58,174	77,125	78,522	79,339	108,497	142,768	27,500	88,789
13시	703,498	68,051	89,050	82,214	89,839	120,285	163,303	33,733	101,978
14시	1,019,281	88,802	113,495	110,812	120,480	175,809	249,676	45,337	154,566
15시	1,685,580	124,286	162,229	164,054	165,863	281,884	427,324	64,631	274,476
16시	2,610,529	184,008	235,305	256,786	269,079	464,881	739,974	92,528	450,619
17시	3,637,426	253,129	266,830	303,051	313,322	630,439	861,274	112,032	573,177
18시	2,855,750	161,958	192,296	211,716	203,737	410,046	580,343	82,831	366,805
19시	1,801,947	117,550	147,840	151,224	140,516	268,519	439,783	59,402	193,191
20시	1,735,245	116,226	145,597	146,976	140,525	237,607	368,519	63,119	190,370
21시	1,438,007	97,284	120,836	126,688	128,305	185,632	278,332	63,220	163,499
22시	938,137	69,337	87,573	92,236	93,145	116,969	172,582	42,790	97,584
23시	862,763	76,301	111,275	106,244	111,986	102,840	132,729	48,411	89,266
계	24,474,606	1,941,524	2,451,894	2,532,301	2,568,797	3,844,607	5,505,873	988,192	3,326,842

○ 시간대별 귀가 통행량 (지역별)

- 대부분 지역에서 귀가 통행량은 오후 5시, 8시에 집중되는 것으로 나타남
- 반면 서울특별시는 오후 8시에 귀가 통행량이 두드러지는데, 이는 업무종료 후 인근의 다양한 활동 기회를 가진 뒤 다른 지역민 대비 늦게 귀가하기 때문으로 판단됨
- 이외에도 수도권(서울, 경기, 인천)의 오전 4~5시 귀가 통행량이 두드러지는데, 이는 수도권에 24시간 운영되는 업종 종사자(병원, 물류센터 등)가 많거나 수도권의 심야버스와 택시 등 교통인프라가 잘 갖춰져 있어 시간에 관계없이 귀가가 가능하기 때문으로 판단됨



〈그림 2-19〉 시간대별 귀가 통행량 (지역별)

〈표 2-17〉 시간대별 귀가 통행량 (지역별) - 특별시, 광역시, 특별자치시

단위 : 통행량(인)

구분	서울	부산	대구	인천	광주	대전	울산	세종
0시	5,390,489	1,453,360	997,186	1,662,581	831,597	717,311	517,263	158,207
1시	4,334,989	1,202,219	809,201	1,366,864	717,712	617,217	404,212	135,472
2시	4,046,354	1,020,536	663,551	1,244,587	625,940	531,250	345,602	114,905
3시	4,019,134	928,505	591,713	1,189,792	549,242	476,287	307,206	106,721
4시	4,967,362	942,798	599,231	1,384,666	546,370	468,493	313,085	111,272
5시	3,425,676	1,101,352	722,721	1,411,635	686,343	578,259	373,771	126,774
6시	2,815,369	1,360,351	948,199	1,087,255	723,489	646,508	467,428	156,273
7시	3,232,913	1,614,896	1,173,734	1,300,148	756,974	714,030	541,202	194,216
8시	3,669,596	1,698,819	1,185,694	1,435,874	754,983	699,539	583,326	196,053
9시	3,858,402	1,717,198	1,203,859	1,467,093	725,137	671,035	580,856	168,663
10시	4,619,286	1,973,553	1,380,295	1,679,437	754,386	712,409	663,186	179,197
11시	5,780,571	2,388,950	1,642,698	2,043,573	829,146	808,161	805,449	205,183
12시	6,764,742	2,665,123	1,805,563	2,337,734	877,465	875,023	894,893	228,552
13시	7,353,317	2,907,575	1,920,861	2,576,562	936,947	931,404	963,307	239,543
14시	8,856,600	3,381,264	2,285,985	2,996,301	1,037,625	1,057,869	1,131,641	265,874
15시	10,491,673	3,963,684	2,673,192	3,610,783	1,190,459	1,192,185	1,400,688	312,481
16시	11,367,591	4,370,611	2,979,428	3,880,186	1,374,373	1,316,066	1,634,735	355,657
17시	12,052,018	4,449,328	3,143,548	4,018,621	1,440,702	1,413,609	1,626,926	373,610
18시	12,078,005	4,149,917	2,875,607	3,765,496	1,380,954	1,351,727	1,426,823	376,821
19시	12,172,131	4,136,658	2,864,042	3,726,377	1,480,910	1,422,507	1,404,816	381,912
20시	13,285,228	4,289,946	3,002,354	3,859,689	1,556,524	1,483,771	1,438,614	401,742
21시	12,415,892	3,727,517	2,703,608	3,332,624	1,492,694	1,359,203	1,252,214	380,997
22시	8,654,127	2,678,984	1,869,618	2,380,856	1,127,290	1,030,572	869,377	271,816
23시	8,082,524	3,150,800	2,186,160	2,556,390	1,397,418	1,284,571	1,025,479	347,742
계	173,733,988	61,273,945	42,228,051	56,315,124	23,794,681	22,359,008	20,972,099	5,789,683

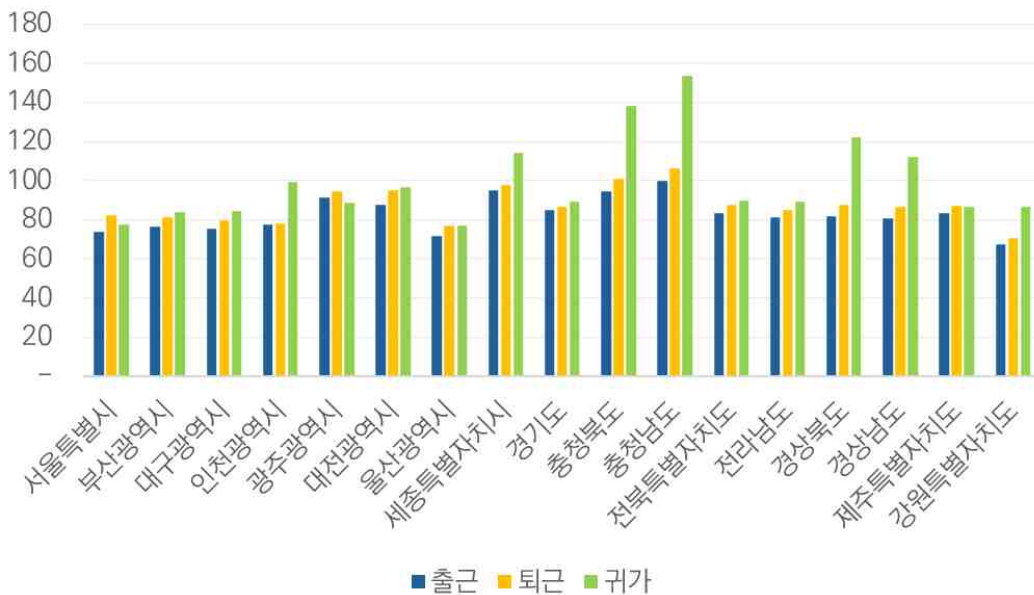
〈표 2-18〉 시간대별 귀가 통행량 (지역별) - 도, 특별자치도(계속)

단위 : 통행량(인)

구분	경기	충북	충남	전북	전남	경북	경남	제주	강원
0시	7,447,318	575,326	724,496	774,790	724,186	662,402	1,058,309	341,809	702,454
1시	6,446,161	491,671	627,390	685,436	671,098	548,706	857,280	304,021	555,578
2시	5,472,058	430,254	540,017	605,547	604,538	458,823	719,643	264,703	498,689
3시	4,897,640	392,024	502,776	552,847	567,423	424,261	655,830	236,216	474,586
4시	5,252,419	396,708	520,668	556,707	576,375	443,284	684,046	232,465	509,772
5시	7,237,310	522,894	650,862	706,253	688,112	585,914	856,970	266,849	686,456
6시	5,113,223	568,862	742,058	770,919	760,375	824,563	1,153,869	290,845	538,521
7시	5,811,670	651,418	838,485	837,067	833,924	1,043,183	1,435,159	319,415	633,192
8시	6,388,111	655,512	849,715	855,314	866,948	1,084,005	1,485,092	330,790	699,687
9시	6,142,989	647,434	844,943	849,972	887,815	1,128,086	1,497,442	331,393	766,670
10시	6,712,060	716,898	928,749	927,527	975,194	1,324,844	1,720,474	358,701	924,429
11시	7,949,088	834,672	1,066,113	1,065,457	1,127,703	1,629,406	2,081,442	400,992	1,173,664
12시	9,202,598	930,921	1,211,157	1,162,784	1,219,408	1,817,916	2,336,568	431,476	1,296,736
13시	10,117,917	976,813	1,273,191	1,197,784	1,290,645	1,900,295	2,471,494	461,865	1,401,677
14시	11,528,848	1,082,589	1,394,060	1,315,068	1,436,644	2,139,485	2,813,932	517,632	1,617,771
15시	13,491,823	1,213,177	1,568,648	1,512,061	1,637,115	2,458,455	3,322,133	604,656	1,917,368
16시	15,228,756	1,358,900	1,764,575	1,767,675	1,962,389	2,868,556	3,896,304	723,122	2,206,069
17시	15,843,353	1,429,650	1,785,796	1,801,853	1,980,443	2,934,844	3,943,783	739,214	2,189,986
18시	14,993,889	1,279,957	1,646,177	1,654,323	1,747,793	2,464,853	3,487,664	696,320	1,848,819
19시	15,201,072	1,290,945	1,673,037	1,710,673	1,741,530	2,357,413	3,472,967	720,883	1,763,238
20시	15,786,018	1,269,433	1,651,249	1,702,338	1,719,443	2,302,672	3,422,041	764,074	1,751,914
21시	14,209,053	1,116,595	1,438,678	1,515,980	1,527,180	1,935,239	2,879,091	702,568	1,425,047
22시	10,270,471	827,689	1,059,680	1,128,081	1,122,352	1,311,626	1,985,667	512,113	956,966
23시	14,289,946	1,089,150	1,420,460	1,437,224	1,515,032	1,655,477	2,269,490	661,787	1,194,125
계	235,033,791	20,749,492	26,722,976	27,093,680	28,183,664	36,304,309	50,506,691	11,213,909	27,733,416

○ 지역별 통행시간

- 서울특별시, 부산광역시, 대구광역시, 인천광역시, 울산광역시와 같은 대도시  
는 통행시간이 다른 지역에 비해 적게 소요되는 것으로 나타남
- 이중 울산은 대표적인 공업도시로 산업단지와 주거지가 밀집해 있어 통행시간  
이 짧은 것으로 판단됨
- 반면 충청북도, 충청남도, 세종특별자치시는 다른 시도 대비 통행시간이 길게  
나타나는데, 이는 대중교통 인프라가 부족하고 교통접근성이 떨어지는 것과  
관련 있을 수 있음



〈그림 2-20〉 지역별 통행시간 (단위: 분)

〈표 2-19〉 지역별 통행시간 (단위: 분)

시도	출근	퇴근	귀가
서울특별시	73	82	78
부산광역시	77	81	84
대구광역시	75	79	84
인천광역시	78	78	99
광주광역시	91	94	89
대전광역시	88	95	97
울산광역시	72	77	77
세종특별자치시	95	97	114
경기도	85	86	89
충청북도	95	101	138
충청남도	100	106	153
전북특별자치도	84	87	90
전라남도	81	85	89
경상북도	82	88	122
경상남도	80	87	112
제주특별자치도	83	87	87
강원특별자치도	67	71	87
평균	83	87	99

## 제3절 객체단위 모바일 통신데이터 가공 및 통행 DB 구축

### 1. 개요

#### 가. 객체단위 DB 구축 목적

- 모바일 통신데이터는 개인이 휴대전화 단말기를 소지한 상태에서 발생하는 신호를 기반으로 통행 궤적을 분석할 수 있는 중요한 데이터임
  - 공간적, 시간적 이동 패턴을 분석할 수 있으며, 교통 및 관련 분야에서 활용 가치가 높음
  - 기존 교통조사 방식(설문조사, 교통량 조사 등)과 비교하였을 때, 더욱 광범위한 지역에서 대규모 데이터를 수집하고 분석할 수 있는 장점이 있음
  - 도시 교통망 최적화, 교통 혼잡 완화, 대중교통 운영 효율성 증대, 신교통수단 도입 효과 분석 등 다양한 교통 분야에서 활용 가치가 높음
- 기존 교통량 데이터는 주로 승용차, 버스, 지하철 등 전통적인 교통수단을 중심으로 데이터를 구축하여 활용하였음
- 보행, 자전거, 개인형 이동수단(PM) 등의 이동정보가 체계적으로 구축되지 않아 해당 수단의 이용행태 분석에 한계가 존재함
- 기존 데이터는 주로 집계 수준에서 제공되며, 개별 객체(Object-Level)의 통행 경로 및 특성을 반영하기 어려움
  - 이동 속도, 목적지 도착 시간, 정체 구간, 교통수단 간 환승 등 세부적인 분석이 미흡하여 정책 수립 및 모빌리티 서비스 개선에 필요한 세부 정보 제공이 부족한 실정임
- 보다 정밀한 통행 행태 분석을 위해 객체 단위의 출발지(Origin), 도착지(Destination), 통행 궤적, 이동수단, 통행목적 등의 상세 정보를 포함한 데이터

구축이 필수적임

- 이동 행태별 맞춤형 분석이 가능하도록 객체 단위의 통행 궤적을 추적하고, 리를 다양한 교통 정책 및 연구에 활용할 수 있도록 체계 마련이 필요함
- 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 성능평가 지표 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 데이터에 대한 전처리 및 가공, 통행경로 및 통행지표 DB 구축 등 기술적 검토를 수행함
- DB 구축을 위한 객체 단위 데이터는 본원에서 확보한 “이동 동선 가명정보” 데이터를 활용 예정임
  - 공간적 범위 : 전국
  - 시간적 범위 : 2024. 10. 14. ~ 2024. 10. 20.
  - 집계 단위 : 50m 격자, 일자별, 1분 단위, 개별통행 트립(Trip) 단위 집계
- 본 사업에서 수행한 기술적 검토 사항은 아래와 같음
  - 객체 단위 모바일 통신데이터 오류 검토 및 필터링
  - 모바일 통신데이터 기반 개별 통행경로 DB 구축
  - 모바일 통신데이터 기반 통행지표 구축 관련 기술적 검토

## 2. 객체 단위 모바일 통신데이터 기초분석

### 가. 객체 단위 DB 개요 및 데이터 형태

- LG 유플러스 모바일 통신 원시데이터는 총 21개의 컬럼이며, 기준일자, 트립 ID, 출·도착 시각, 출·도착 시도/시군구/읍면동/50셀 코드, 이동 유형 코드, 평균 속도, 출발-도착 유형으로 구성되어 있음
- LG 유플러스 모바일 통신 원시데이터의 공간 단위는 격자 단위로 50m 단위로 구분되어 있으며, 시간 단위는 1분 단위로 구분되어 있음
- 이동 거리는 km 단위이며, 속도는 km/h 단위로 되어 있음

〈표 2-20〉 LG유플러스 이동통신 가명정보 데이터 테이블정의서

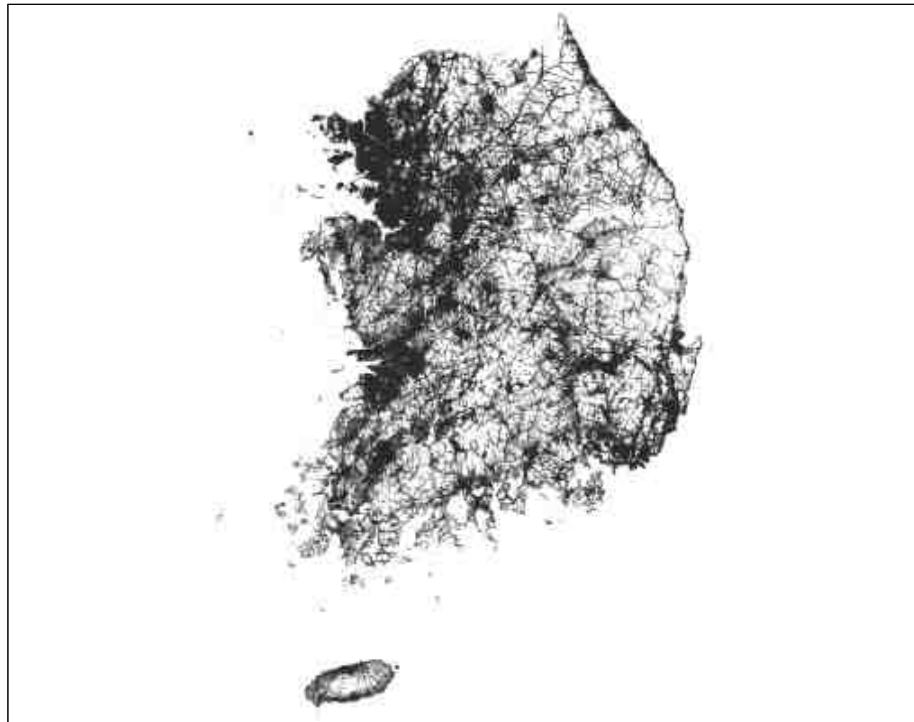
번호	변수명	데이터 유형	설명
1	P_YYYYMMDD	날짜(문자열)	기준 날짜 (YYYY-MM-DD)
2	TRIP_NO	정수	이동 트립 번호
3	DPR_MT1_UNIT_TM	정수	출발 시각 (YYYY-MM-DD HH-MM) 1분 단위
4	DPR_CTDO_CD	정수	출발 시도 코드
5	DPR_CTDO_NM	문자열	출발 시도명
6	DPR_CCW_CD	정수	출발 시군구 코드
7	DPR_CCW_NM	문자열	출발 시군구명
8	DPR_ADNG_CD	정수	출발 읍면동 코드
9	DPR_ADNG_NM	문자열	출발 읍면동명
10	DPR_CELL_ID	정수	출발 셀 ID(50m 격자)
11	ARV_MT1_UNIT_TM	정수	도착 시각 (YYYY-MM-DD HH-MM) 1분 단위
12	ARV_CTDO_CD	정수	도착 시도 코드
13	ARV_CTDO_NM	문자열	도착 시도명
14	ARV_CCW_CD	정수	도착 시군구 코드
15	ARV_CCW_NM	문자열	도착 시군구명
16	ARV_ADNG_CD	정수	도착 읍면동 코드
17	ARV_ADNG_NM	문자열	도착 읍면동명
18	ARV_CELL_ID	정수	도착 셀 ID(50m 격자)
19	DYNA_DYN_KD_CD	정수	이동 유형 코드
20	DYNA_MVMT_SPED	실수	평균 속도
21	OD_TYPE	문자열	출발-도착 유형

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

〈표 2-21〉 LG 유플러스 모바일 통신 격자&공간정보 데이터 테이블정의서

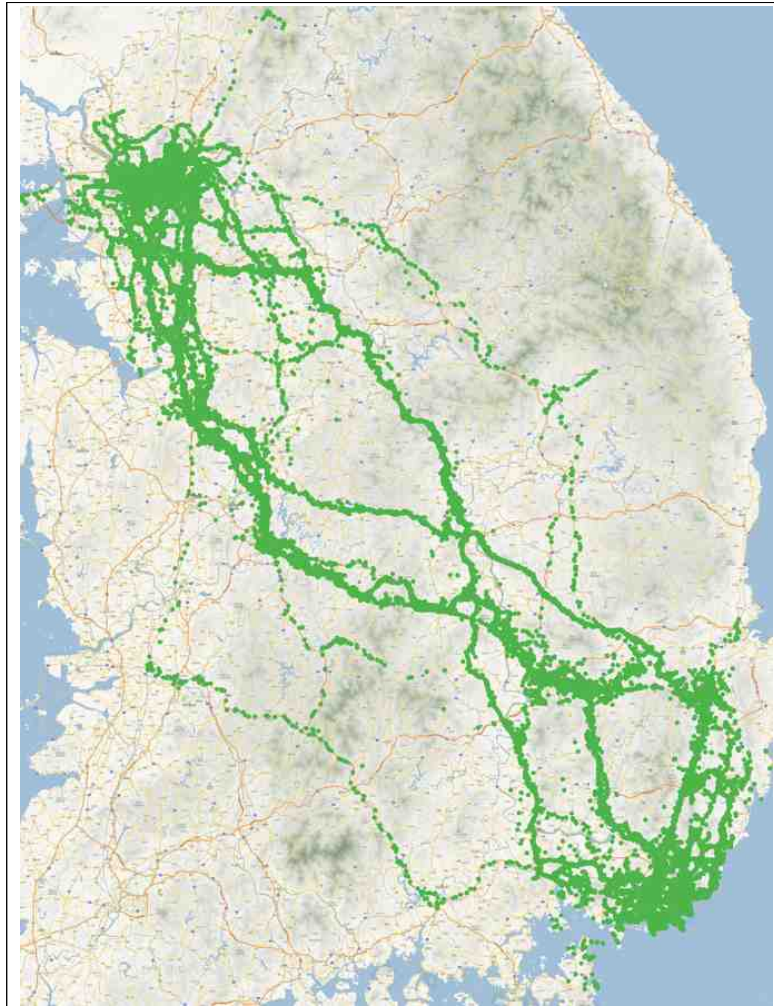
순번	필드명	설명	자료형	비고
1	CELL_ID	격자 ID	NUMERIC	-
2	X_AXIS_WGS	중심_X_WGS84	NUMERIC	-
3	Y_AXIS_WGS	중심_Y_WGS84	NUMERIC	-
4	MEGA_NM	시도명	VARCHAR	-
5	MEGA_CD	시도코드	NUMERIC	-
6	CTY_NM	시군구명	VARCHAR	-
7	CTY_CD	시군구코드	NUMERIC	-
8	ADMI_NM	행정동명	VARCHAR	-
9	ADMI_CD	행정동코드	NUMERIC	-

- 공간정보 시각화 분석을 위해 별도로 격자 shp을 구축하였으며, 격자의 개수는 전국 총 40,576,580개임

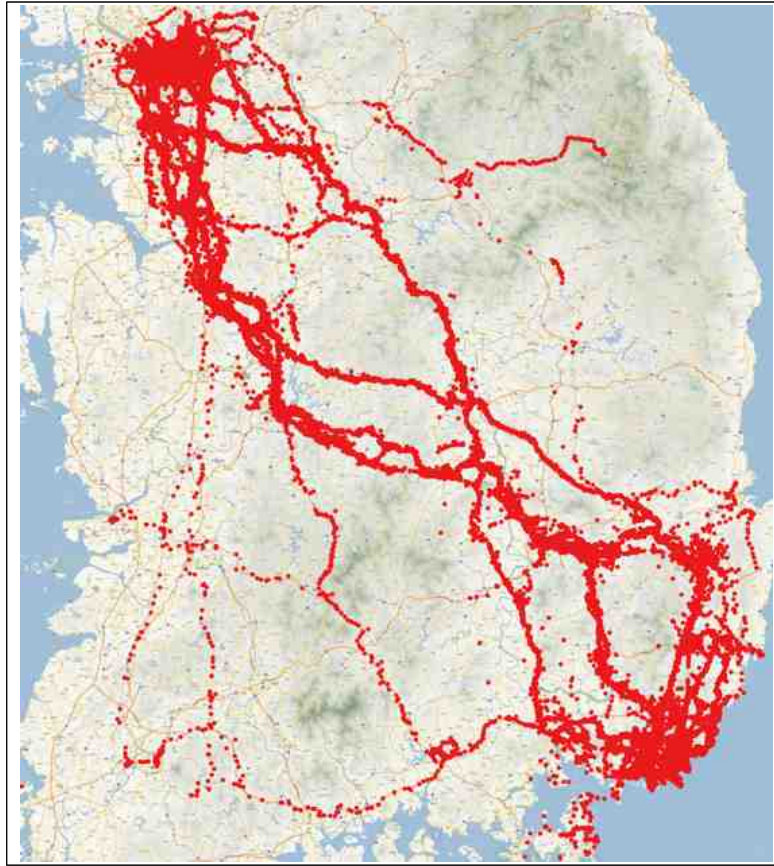


〈그림 2-21〉 LG 통신 데이터 격자 전국 예시

- 50m 격자 단위 이동 궤적데이터 시각화 결과는 아래와 같음



〈그림 2-22〉 서울 → 부산 50m 격자 단위 이동 궤적  
(2024-10-14(월) 기준)



〈그림 2-23〉 서울 → 부산 50m 격자 단위 이동 궤적  
(2024-10-19(토) 기준)

### 3. 객체 단위 모바일 통신데이터 오류 제거 및 전처리

#### 가. 객체 단위 데이터 오류 제거 및 전처리

- 객체 단위 모바일 통신데이터의 오류를 제거하고 대용량 데이터의 처리를 위해 다음과 같은 과정을 수행함
- 오류 제거
  - 파싱 오류 제거: 한 라인의 필드 개수가 21개가 아닌 경우 제거
  - 날짜 오류 제거: 기준일자 필드가 비었거나 형식(YYYY-MM-DD)에 맞지 않은 경우 제거

- 시간 오류 제거: 출발시간 또는 도착시간 필드가 비었거나 형식 (YYYY-MM-DD HH-MM)에 맞지 않은 경우 제거
  - 참조 오류 제거: 출발 격자 또는 도착 격자 필드가 격자-행정구역 매칭 테이블에 참조되지 않는 경우 제거
- 데이터의 효율적인 처리를 위해 다음과 같은 전처리 과정을 수행함
- 일자 분할 : 일자별 trip\_id로 파일 분할
  - 시간 정보 변환 : Datetime → Unixtime
  - 통행목적 제외: 무의미한 형태로 포함되어 있는 통행목적 컬럼 제외
- 최종 전처리된 데이터의 형태는 다음과 같음
- 데이터 발생 일자와 트립 넘버를 조합한 Trip\_ID로 키값을 설정하였으며, 출발 격자ID, 도착격자 ID, 출발 행정동 ID, 도착 행정동 ID, 출발시각, 도착시각, 이동거리, 속도로 컬럼을 구성하였음

〈표 2-22〉 LG 유플러스 모바일 통신 전처리 데이터

순번	K-V	필드명	설명	자료형	비고
1	KEY	TRIP_ID	데이터 발생일자	정수	P_YYYYMMDD + TRIP_NO
2	VALUE (반복)	DPR_MT1_UNIT_TM	출발 시각	정수	YYYY-MM-DD hh-mm
3		DPR_CELL_ID	출발 50셀 ID	정수	
4		ARV_MT1_UNIT_TM	도착 시각	정수	
5		ARV_CELL_ID	도착 50셀 ID	정수	
6		DYNA_DYN_KD_CD	이동 유형 코드	정수	1=정지, 2=도보, 3=비도보
7		DYNA_MVMT_SPED	평균 속도	실수	km/h

## 나. 도보·비도보 수단 구분

- 개별 ID의 이동(트립)에는 도보, 승용차, 버스, 철도, 항공 등 다양한 교통수단이 포함될 수 있음
- 따라서, 개별 객체의 이동 데이터를 활용하여 도보와 비도보 수단을 명확히 구분하는 기준을 설정할 필요가 있음
- 이에 따라 평균 이동속도를 기준으로 도보와 도보 외의 통행을 구분함
- 일반적인 성인의 평균 걸음 속도를 고려하여 5km/h 미만의 이동을 도보로 정의함
  - 5km/h 이상의 속도를 보이는 이동 구간은 비도보 수단(승용차, 버스, 철도 등)으로 분류함
  - 일반적인 보행 속도 범위(3~5km/h)를 고려하여 도보 구간을 설정하고, 이동 속도가 변화하는 구간에서는 추가적인 필터링 적용 가능함
- 전체 트립을 단순히 도보와 도보 외로 구분하는 것이 아니라, 각 단위 이동(세그먼트)별로 도보와 비도보 수단을 구분하여 분석함
- 그러나, 차량이 혼잡하거나 정체가 심한 구간에서는 차량 속도가 5km/h 미만으로 낮아질 가능성이 있음
  - 단순 이동속도 기준으로 도보와 도보 외 이동을 구분할 경우, 오류가 발생할 수 있음
  - 단순 속도 기준뿐만 아니라, 이동 거리, 정지 여부, 연속된 속도 변화 패턴 등을 고려한 다중 조건 기반 도보·비도보 구분이 필요함
- GPS신호 수신 불량지역(지하, 터널 등)에서는 일시적으로 데이터가 왜곡될 수 있으며, 객체 단위별 트립 정보는 1분 단위로 구축되지 않을 수 있음

〈표 2-23〉 LG유플러스 모바일 통신 원시데이터 (트립 ID : 720)

출발 셀 ID	도착 셀 ID	출발 시각	도착 시각	이동 유형	속도
79432038	79160028	2024-10-14 10:05	2024-10-14 10:06	3	95.72
79160028	79080023	2024-10-14 10:06	2024-10-14 10:07	3	33.54
79080023	79000022	2024-10-14 10:07	2024-10-14 10:08	3	23.43
79000022	78904010	2024-10-14 10:08	2024-10-14 10:09	3	56.92
78904010	78888004	2024-10-14 10:09	2024-10-14 10:10	3	21.21
78888004	78888004	2024-10-14 10:10	2024-10-14 10:11	1	0
78888004	78775991	2024-10-14 10:11	2024-10-14 10:12	3	63.57
78775991	78791988	2024-10-14 10:12	2024-10-14 10:13	3	6.71
...	...	...	...	...	...

- 이동속도를 기반으로 이동 유형에 대한 변수를 설정함
  - 속도 구간에 따라 세 가지 변수 값(1=정지, 2=도보, 3=비도보)으로 구분함
    - 정지 : 평균 이동속도가 0km/h 값으로, 이동이 거의 없거나 신호대기 등으로 정지인 상태
    - 도보 : 평균 이동속도가 5km/h 미만 값으로, 일반적인 성인의 도보 속도 범위에 해당하는 구간
    - 비도보 : 평균 이동속도가 5km/h 이상으로 승용차, 버스, 철도 등 도보가 아닌 교통수단을 이용하는 구간

〈표 2-24〉 LG유플러스 모바일 통신 원시데이터 수단 구분 (트립 ID : 720)

출발 셀 ID	도착 셀 ID	출발 시각	도착 시각	이동 유형	속도	수단구분
79432038	79160028	2024-10-14 10:05	2024-10-14 10:06	3	95.72	비도보
79160028	79080023	2024-10-14 10:06	2024-10-14 10:07	3	33.54	비도보
79080023	79000022	2024-10-14 10:07	2024-10-14 10:08	3	23.43	비도보
79000022	78904010	2024-10-14 10:08	2024-10-14 10:09	3	56.92	비도보
78904010	78888004	2024-10-14 10:09	2024-10-14 10:10	3	21.21	비도보
78888004	78888004	2024-10-14 10:10	2024-10-14 10:11	1	0	정지
78888004	78775991	2024-10-14 10:11	2024-10-14 10:12	3	63.57	비도보
78775991	78791988	2024-10-14 10:12	2024-10-14 10:13	3	6.71	비도보
...	...	...	...	...	...	

## 4. 통행궤적 및 통행정보 데이터 구축

### 가. 통행궤적 및 통행정보 데이터 구축

- 앞서 살펴본 바와 같이, 원시데이터에는 개별 통행궤적(Trajectory)에 대한 시퀀스(Sequence) 정보가 존재하지 않음
- 원시데이터는 단순한 위치 정보의 집합으로, 시간 순서(Time Sequence)를 반영하지 않을 경우 이동 경로를 명확히 파악하기 어려움
- 따라서, 출발/도착시각 정보를 활용하여 통행궤적을 재구성할 필요가 있음
- 트립ID, 출발/도착시각 등의 정보가 시간 순서대로 정렬되어야만 객체의 통행 패턴을 분석할 수 있음
- 원시데이터는 단순한 위치 데이터로 저장되며, 시퀀스 정보 없이 개별 위치 좌표만 제공됨
- 원시데이터를 활용하여 연속적인 이동궤적을 구축하고, 이동 구간 간의 관계를 정의하기 위해 연속된 시간대를 기반으로 데이터를 정렬 및 연결이 필요함
- 본 연구에서는 출발지부터 도착지까지의 이동 경로를 분석하기 위하여 연속된 시간 데이터를 기반으로 통행궤적을 구축함

〈표 2-25〉 LG 유플러스 모바일 통신데이터 기반 통행정보DB

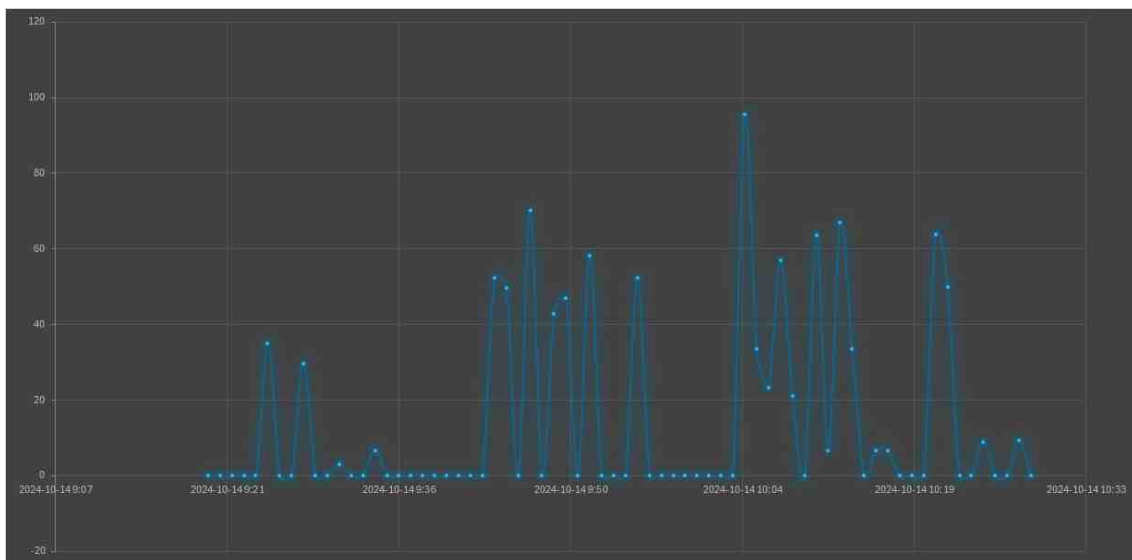
순번	K-V	필드명	설명	자료형	비고
1	KEY	TRIP_ID	데이터 발생일자	정수	P_YYYYMMDD + TRIP_NO
2	VALUE (반복)	DPR_MT1_UNIT_TM	출발 시각	정수	YYYY-MM-DD hh-mm
3		DPR_CTDO_CD	출발 시도 코드	정수	
4		DPR_CCW_CD	출발 시군구 코드	정수	
5		DPR_ADNG_CD	출발 읍면동 코드	정수	
6		DPR_CELL_ID	출발 50셀 ID	정수	
7		ARV_MT1_UNIT_TM	도착 시각	정수	
8		ARV_CTDO_CD	도착 시도 코드	정수	
9		ARV_CCW_CD	도착 시군구 코드	정수	
10		ARV_ADNG_CD	도착 읍면동 코드	정수	
11		ARV_CELL_ID	도착 셀 ID	정수	
12		DYNA_DYN_KD_CD	이동 유형 코드	정수	1=정지, 2=도보, 3=비도보
13		DYNA_MVMT_SPED	평균 속도	실수	km/h
14		OD_TYPE	출발-도착 유형	문자	출발지, 경유지, 도착지

## 나. 개별 통행 ID별 통행 및 통행정보 프로파일 산출 및 검토

- 이동 객체는 고유한 통행 ID가 부여되며, 시간 순서에 따라 이동 경로 분석이 가능함
- 예를 들어, 통행ID 720번에 해당하는 이동 경로는 다음과 같이 출발지부터 도착지까지 총 70개의 연속적인 이동 구간이 포함되어 있음
- 각 이동 구간별 출발 셀ID, 도착 셀ID, 출발 시각, 도착 시각, 이동속도, 이동 유형(정지, 도보, 비도보), 출발-도착 유형(출발지, 경유지, 도착지) 등의 정보를 포함하여 상세한 이동 행태를 구축함
  - 첫 번째 행(순서1)은 출발지로 정의되며, 정지 상태에서 시작됨
  - 이후 2~69번째 값은 출발지 이후부터 도착지 이전까지의 모든 이동을 경유지로 설정하고, 해당 구간에서 이동 유형 및 수단 구분을 분류함
  - 마지막 이동 데이터는 도착지로 정의되며, 최종적으로 이동을 완료한 시점임
- 출발지, 경유지, 도착지로 이어지는 총 70개의 연속적인 이동 기록을 기반으로, 총 이동시간, 이동속도 등 관련 통행 정보를 산출함
  - 출발지부터 도착지까지의 총 이동 시간은 약 70분으로 산정됨
  - 이동 구간의 속도를 살펴보면 최저 속도는 0km/h(정지 상태), 평균 속도는 14.2km/h, 최고속도는 95.7km로 나타남
  - 최고속도(95.7km/h)는 비도보 구간에서 관측되었으며, 이를 통해 이동 수단이 차량, 버스, 철도 등의 교통수단을 포함하는 것으로 판단됨
  - 이동 유형(정지, 도보, 비도보)을 구분하여 빈도를 분석한 결과, 정지 횟수 44회, 도보 1회, 비도보 25회로 집계됨
  - 정지 상태가 44회로 가장 높은 비중을 차지하고 있으며, 이동 유형이 반복적으로 변화하는 특징을 가짐
  - 이를 통해, 교통체증, 신호대기, 탑승 지점 등에서의 정체 현상이 빈번하게 발생했을 가능성이 높음

〈표 2-26〉 통행ID별 통행계적 데이터 구축 결과(트립 ID : 720)

순서	출발 셀 ID	도착 셀 ID	출발 시각	도착 시각	이동 유형	속도	수단 구분	이동 유형
1	7922 4162	7922 4162	2024-10-14 9:20	2024-10-14 9:21	1	0	정지	출발지
...	...	...	...	...	...	...	...	...
46	7943 2038	7916 0028	2024-10-14 10:05	2024-10-14 10:06	3	95.72	비도보	경유지
47	7916 0028	7908 0023	2024-10-14 10:06	2024-10-14 10:07	3	33.54	비도보	경유지
48	7908 0023	7900 0022	2024-10-14 10:07	2024-10-14 10:08	3	23.43	비도보	경유지
49	7900 0022	7890 4010	2024-10-14 10:08	2024-10-14 10:09	3	56.92	비도보	경유지
50	7890 4010	7888 8004	2024-10-14 10:09	2024-10-14 10:10	3	21.21	비도보	경유지
51	7888 8004	7888 8004	2024-10-14 10:10	2024-10-14 10:11	1	0	정지	경유지
52	7888 8004	7877 5991	2024-10-14 10:11	2024-10-14 10:12	3	63.57	비도보	경유지
53	7877 5991	7879 1988	2024-10-14 10:12	2024-10-14 10:13	3	6.71	비도보	경유지
...	...	...	...	...	...	...	...	...
70	7853 5943	7853 5943	2024-10-14 10:29	2024-10-14 10:30	1	0	정지	도착지



〈그림 2-24〉 시간대별 통행 이동속도(트립 ID : 720)

## 5. 객체 단위 모바일 통신데이터 구축 및 활용 관련 검토사항

### 가. 이동 유형별 상세 분석 관련 기술적 검토

- 이동 유형별 분석을 위해 요구되는 기술적 요소
  - 이동 유형을 보다 정밀하게 분석하기 위해서는 위치, 시간, 교통 네트워크 등 데이터 간 융합 분석이 필수적임
  - 기존의 분류 방식은 단순 속도 변수만을 기준으로 이동 유형을 분류하였기 때문에 정밀한 분석이 어려움
  - 예를 들어, 속도만으로 도보, 비도보를 판별할 경우 정체 구간에서 차량 이동이 도보로 잘못 분류될 가능성이 높음
  - 따라서, 속도뿐만 아니라 GPS데이터, 맵매칭 기술 등을 활용하여 도보 이동을 더욱 정확하게 분석할 필요가 있음
- GPS X,Y좌표 데이터 기반 도보 분석
  - 현재 모바일 통신데이터는 GPS 좌표체계가 아닌 격자 단위(Grid Cell)로 이동 궤적을 제공하고 있음
  - 격자 단위는 객체의 이동 흐름을 대략적으로 파악하는데 유용하나, 정확한 도보 경로 추적에는 한계가 있음
- 초(Second) 단위 데이터 수집
  - 본 사업에서 수집된 데이터는 1분 간격의 시간 데이터로 구성되어 있음
  - 초 단위의 통신 데이터를 수집하여 통행궤적을 더욱 명확하게 구분할 수 있으며, 실제 이동 구간과 대기 상태 등을 분석할 수 있음
- 가속도 센서 데이터 분석 활용
  - 휴대폰에 내장된 가속도 센서를 활용하여 도보, 달리기, 자전거 등의 이동 수단을 구분할 수 있음

- 가속도 센서는 이용자의 움직임 강도와 방향을 측정할 수 있으며, 이를 통행 도보와 기타 이동 수단을 구분하는 데 활용 가능함
  - 이를 통해, 속도만으로 이동 유형을 판별하는 기존 방식의 한계를 보완할 수 있음
- 도로 네트워크 맵매칭 적용
- 통행을 정밀하게 분석하기 위해 다양한 교통 네트워크(도로망, 대중교통망, 철도망 등)와 통신데이터를 결합하는 맵매칭 기술 적용이 필요함
  - 맵매칭을 활용하면, 통신데이터를 특정 네트워크에 맵핑하여 도보 통행을 보다 신뢰성 있게 분석할 수 있음
  - 이를 적용할 경우, 이동수단 간 전환 구간(도보, 탑승, 환승 등)을 보다 정밀하게 분석할 수 있으며, 교통 네트워크 기반의 통행 분석이 가능함
  - 향후 이러한 맵매칭 기술을 활용하면 모바일 통신데이터 기반 통행 분석의 정확도를 향상시킬 수 있을 것으로 기대됨

#### 나. 통행지표 구축 관련 기술적 검토

- 데이터 정확성 확보를 위한 기술적 검토
- 데이터의 정확성은 분석 결과의 신뢰성에 직접적인 영향을 미치는 요소임
  - 데이터 수집 과정에서 발생할 수 있는 오류를 최소화하고, 수집된 데이터가 충분한 정확도를 확보하고 있는지 검토하기 위한 기술 개발이 필수적임
  - GPS 신호 오류 보정, 이동속도 및 궤적 검증, 이상치 제거 등
- 데이터 처리 및 분석 환경 구축
- 모바일 통신 원시데이터는 대용량 데이터이며, 이를 효과적으로 처리하기 위해서는 고성능 데이터베이스 설계 및 분석 환경이 필수적임
  - 빅데이터 처리 및 분석을 위해서는 데이터베이스, 빅데이터 가공 환경 구축, 인프라 시스템 등 기술적 고려가 필요함

○ 통행지표 산출 알고리즘 적용

- 최근 최적화 알고리즘, 머신러닝 등을 활용한 통행지표 생성 기술이 활발히 연구되고 있음
- 신뢰성 높은 통행지표를 산출하기 위해서는 최신 알고리즘 적용 및 데이터 셋 구축이 필수적임

○ 데이터 속성정보 수집 및 구축

- 통행지표의 신뢰성을 높이기 위해 이동 궤적 데이터뿐만 아니라, 다양한 속성정보(메타데이터)를 추가로 확보할 필요가 있음
- 인구통계학적 정보(성별, 연령, 직업 등), 통행목적 정보(출·퇴근, 등·학교, 쇼핑, 여가 등), 교통수단 정보(승용차, 버스, 철도, 해운, 항공, 자전거, PM 등)



## 제3장 모바일 통신데이터 기반

### 통행 DB 검증

제1절 개요

---

제2절 검증 결과

---



## 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)이 있어야 하는데, 이에 대한 논란이 존재함
  - 대표적인 통계인 한국교통연구원의 KTDB와 통계청 인구총조사의 자료는 검증하고자 하는 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 사회경제적 지표를 활용한 통행 OD를 검증하고 있으나, 이는 공간 및 집단 간 동질성이 전제되어야 하나 실제로는 그렇지 않을 수 있음 (e.g., 지가, 상호교차 검증역간 통행패턴의 이질성, 직업 간 재택근무 비율 상이 등)
  - 또한, 모바일 통신데이터와 같은 빅데이터에서는 기존 조사에서는 확인할 수

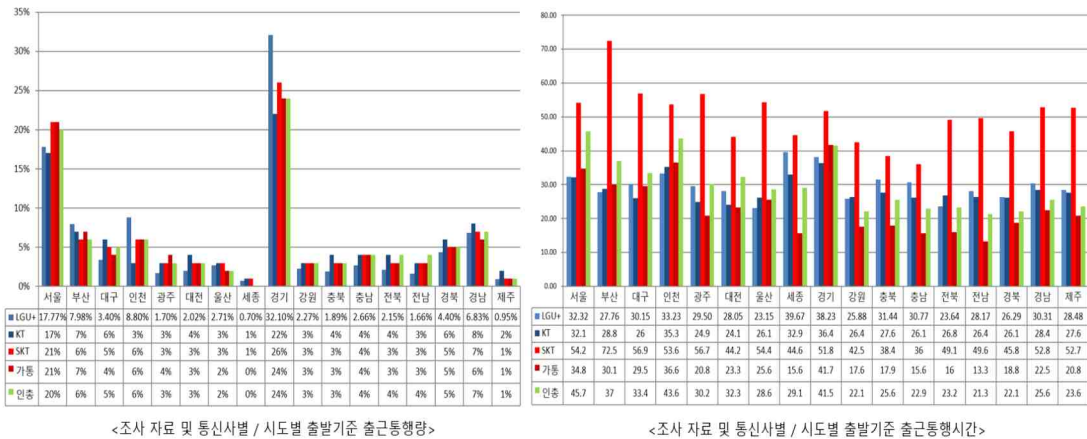
없었던 통행패턴을 확인할 수 있음에도 불구하고, 평균적 검증 접근방법이 적  
합하지 않을 수 있음

○ 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 데이터별 보정 방법을 제시할 수 있음

- ① 상호교차 검증

· <그림 3-1>과 같이 KTDB, 인구총조사, 3개 통신사, 차량 GPS, 대중교통카드, 수송  
실적, 사회경제지표 등을 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 활용  
할 수 있음

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

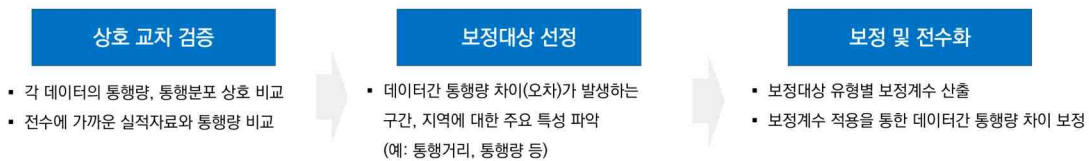
여기서, 데이터 종류 n(1~m)에 대한 i에서 j까지 통행량  $T_{ijn}$  과 가중치  $a_{ijn}$

- 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 기준점으로 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

- 본 사업에서는 모바일 통신데이터와 조사자료를 상호교차 방식을 통해 검증하고자 함
  - 모바일 데이터(LG, KT, SKT 3사 자료)와 조사자료(전국 지역간 목적 OD 자료)의 통행량을 상호교차 검증하고자 함
  - 상호교차 검증 대상은 본 사업을 통해 수집 가능한 데이터로 한정하고 공간 단위 변화에 따른 오차 변화를 확인하기 위해 분석 단위는 시도 단위로 설정하고자 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
    - LG, KT, SKT 데이터를 기반으로 구축한 통행 DB를 수집하였으며, 상호 비교를 위하여 시간적·공간적·내용적 범위를 동일하게 설정함
    - 통신사별로 개인정보보호를 위한 특정 기준 이하의 통행량(3 또는 5 통행량)은 마스킹 처리함
    - 단, KT는 전수화 전후의 통행량을 제공하며, 전수화 통행 정보의 경우 노이즈(noise)를 추가하는 익명화 처리(data anonymization)가 실시됨

#### ① LG

- 시간적 범위: 2023년 11월
- 공간적 범위: 전국 (시도 단위)
- 출처: 2024년 DB사업을 통해 구축한 LG 통신데이터 기반 기종점 통행량 데이터

〈표 3-1〉 LG 데이터 정의서

종류	변수 세부정보	변수명	비고	
기본 정보	기준일자	P_YYYYMMDD	YYYY-MM-DD	
	성별	SEX_DV_CD	MALE, FEMALE	
	연령대	YY10_AGLV_ID	10세 단위	
출발 정보	시간대	DPR_TMZN	00, 01, 02, … , 22, 23	
	셀ID	DPR_CELL_ID	-	
	X좌표	DPR_X_AXIS_WGS	격자 중심점 X좌표	
	Y좌표	DPR_Y_AXIS_WGS	격자 중심점 Y좌표	
	시도코드	DPR_CTDO_CD	격자 중심점 시도코드	
	시도명	DPR_CTDO_NM	격자 중심점 시도명	
	시군구코드	DPR_CCW_CD	격자 중심점 시군구코드	
	시군구명	DPR_CCW_NM	격자 중심점 시군구명	
	행정동코드	DPR_ADNG_CD	격자 중심점 행정동코드	
	행정동명	DPR_ADNG_NM	격자 중심점 행정동명	
	장소유형	DPR_PLC_TYPE	거주지, 근무지, 기타	
	도착 정보	시간대	ARV_TMZN	00, 01, 02, … , 22, 23
		셀ID	ARV_CELL_ID	-
		X좌표	ARV_X_AXIS_WGS	격자 중심점 X좌표
Y좌표		ARV_Y_AXIS_WGS	격자 중심점 Y좌표	
시도코드		ARV_CTDO_CD	격자 중심점 시도코드	
시도명		ARV_CTDO_NM	격자 중심점 시도명	
시군구코드		ARV_CCW_CD	격자 중심점 시군구코드	
시군구명		ARV_CCW_NM	격자 중심점 시군구명	
행정동코드		ARV_ADNG_CD	격자 중심점 행정동코드	
행정동명		ARV_ADNG_NM	격자 중심점 행정동명	
통행 정보	이동자수	MOVE_CUST_CNT	전국민확대계수 적용된 이동자수	
	평균 이동거리	AVG_MOVE_TIME	평균 이동시간	
	평균 이동시간	AVG_MOVE_DIST	평균 이동거리	

○ 시도별 통행량

- 2023년 11월 LG사의 시도별 통행 발생량은 경기도(27.7%), 서울특별시(17.2%), 부산광역시(6.5%) 순으로 나타났으며, 수도권의 비율은 50.5%에 달함
- 통행량이 가장 적게 나타난 지역은 세종특별자치시이며, 전체의 약 0.7%를 차지함

〈표 3-2〉 2023년 11월 시도별 통행량(LG)

(단위: 통행/월, %)

시도	출발		도착	
	통행량	비율	통행량	비율
서울특별시	787,874,234	17.2	787,959,889	17.2
부산광역시	298,479,667	6.5	298,502,833	6.5
대구광역시	209,856,308	4.6	209,890,294	4.6
인천광역시	257,017,671	5.6	257,013,111	5.6
광주광역시	144,826,400	3.2	144,852,747	3.2
대전광역시	145,391,871	3.2	145,391,208	3.2
울산광역시	107,688,679	2.3	107,679,607	2.3
세종특별자치시	31,149,462	0.7	31,135,501	0.7
경기도	1,271,829,252	27.7	1,271,853,585	27.7
충청북도	138,896,775	3.0	138,873,440	3.0
충청남도	171,112,657	3.7	171,095,620	3.7
전북특별자치도	173,427,897	3.8	173,451,190	3.8
전라남도	166,961,056	3.6	166,891,680	3.6
경상북도	200,200,907	4.4	200,170,219	4.4
경상남도	267,070,812	5.8	267,071,794	5.8
제주특별자치도	78,561,311	1.7	78,550,635	1.7
강원특별자치도	138,296,274	3.0	138,257,880	3.0
합 계	4,588,641,233	100.0	4,588,641,232	100.0

② KT

- 시간적 범위: 2023년 11월
- 공간적 범위: 전국 (500 × 500m 격자 단위)
- 출처: 국가교통조사사업을 통해 구축한 KT 통신데이터 기반의 2023년, 2024년 기종점 통행량 데이터

〈표 3-3〉 KT 데이터 정의서

칼럼 ID	칼럼명	비고
ETL_YMD	기준 일자	
O_ID	출발 격자 ID	500m 격자 ID
O_X	출발 격자 X 좌표	500m 격자 중점 X 좌표
O_Y	출발 격자 Y 좌표	500m 격자 중점 Y 좌표
D_ID	도착 격자 ID	500m 격자 ID
D_X	도착 격자 X 좌표	500m 격자 중점 X 좌표
D_Y	도착 격자 Y 좌표	500m 격자 중점 Y 좌표
O_TIME_CD	출발 시간대 코드	1시간 단위
D_TIME_CD	도착 시간대 코드	1시간 단위
O_TYPE	출발 격자 유형	H:집(야간상주지), W:직장/학교(주간상주지), E:기타
D_TYPE	도착 격자 유형	H:집(야간상주지), W:직장/학교(주간상주지), E:기타
O_SIDO_NM	출발 시도명	
O_SGG_NM	출발 시군구명	
O_ADMI_NM	출발 행정동명	
O_ADMI_CD	출발 행정동 코드	
D_SIDO_NM	도착 시도명	
D_SGG_NM	도착 시군구명	
D_ADMI_NM	도착 행정동명	
D_ADMI_CD	도착 행정동코드	
SEXAGE	성연령 코드	10세 단위 성연령 코드
PURPOSE	이동 목적	출근, 등교, 업무, 귀가, 기타
CNT	전수화 집계 값	
RATIO	전수화 비율	전수화를 위한 비율
AVG_DIST	평균 이동 거리	동일 이동의 평균 이동 거리(10m 단위)
AVG_TIME	평균 이동 시간	동일 이동의 평균 이동 시간(1분 단위)

○ 시도별 통행량

- 2023년 11월 LG사의 시도별 통행 발생량은 경기도(24.4%), 서울특별시(19.7%), 부산광역시(6.5%) 순으로 나타났으며, 수도권 비율은 49.2%에 달함
- 통행량이 가장 적게 나타난 지역은 세종특별자치시이며, 전체의 약 0.7%를 차지함

〈표 3-4〉 2023년 11월 시도별 통행량(KT)

(단위: 통행/월, %)

시도	출발		도착	
	통행량	비율	통행량	비율
서울특별시	466,962,643	19.7	469,134,279	19.8
부산광역시	150,262,670	6.5	150,509,838	6.5
대구광역시	108,066,714	4.7	108,074,563	4.7
인천광역시	118,903,620	5.1	119,334,767	5.2
광주광역시	58,830,383	2.5	58,967,283	2.5
대전광역시	69,248,608	3.0	69,301,408	3.0
울산광역시	50,147,715	2.2	50,174,691	2.2
세종특별자치시	16,170,457	0.7	16,174,414	0.7
경기도	565,702,605	24.4	563,377,810	24.3
충청북도	75,535,773	3.3	75,455,837	3.3
충청남도	98,865,329	4.4	98,662,238	4.3
전북특별자치도	80,407,632	3.6	80,381,042	3.6
전라남도	80,408,913	3.6	80,249,376	3.6
경상북도	114,152,916	5.1	114,044,251	5.1
경상남도	135,470,031	5.9	135,311,317	5.9
제주특별자치도	43,944,487	2.0	44,001,010	2.0
강원특별자치도	70,730,311	3.4	70,656,685	3.4
합 계	2,303,810,808	100.0	2,303,810,808	100.0

③ SKT

- 시간적 범위: 2023년 11월
- 공간적 범위: 전국 (시도 단위)
- 출처: 2023년 국가교통조사사업을 통해 구축한 SKT 통신데이터 기반 기종점 통행량 데이터

〈표 3-5〉 SKT 데이터 정의서

칼럼 ID	칼럼명	비고
o_month	기준 월	
o_dow	출발 요일	1~7(월~일)
o_time	출발 시간대	1시간 단위
o_sigun	출발지 시군구	5자리 숫자코드
o_sido	출발지 시도	2자리 숫자코드
o_type	출발 유형	집(0), 회사(1), 학교(2), 쇼핑(레저)(3), 기타(4), 관광(5)
d_time	도착 시간대	1시간 단위
d_sigun	도착지 시군구	5자리 숫자코드
d_sido	도착지 시도	2자리 숫자코드
d_type	도착 유형	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	평균 통행량	보정된 통행량(5통행 이하 마스킹 처리)
move_time_avg	평균 통행속도	초(Sec)

- 시도별 통행량
  - 2023년 11월 LG사의 시도별 통행 발생량은 서울특별시(22.5%), 경기도(22.4%), 부산광역시(7.2%) 순으로 나타났으며, 수도권의 비율은 50.7%에 달함
  - 통행량이 가장 적게 나타난 지역은 세종특별자치시이며, 발생량은 전체의 약 0.5%를 차지함

〈표 3-6〉 2023년 11월 시도별 통행량(SKT)

(단위: 통행/월, %)

시도	출발		도착	
	통행량	비율	통행량	비율
서울특별시	668,454,733	22.5	677,344,925	22.8
부산광역시	213,756,736	7.2	216,442,756	7.3
대구광역시	158,427,378	5.3	155,415,696	5.2
인천광역시	170,811,967	5.8	175,260,541	5.9
광주광역시	101,262,500	3.4	99,619,135	3.4
대전광역시	95,544,833	3.2	96,297,654	3.2
울산광역시	66,218,388	2.2	65,651,610	2.2
세종특별자치시	14,785,309	0.5	13,140,298	0.4
경기도	663,268,357	22.4	648,083,790	21.9
충청북도	84,700,537	2.9	84,951,510	2.9
충청남도	105,952,532	3.6	107,806,006	3.6
전라북도	105,450,464	3.6	106,387,101	3.6
전라남도	89,941,418	3.0	91,058,950	3.1
경상북도	131,888,211	4.4	134,429,830	4.5
경상남도	172,163,459	5.8	170,542,868	5.7
제주특별자치도	41,378,058	1.4	41,487,208	1.4
강원특별자치도	82,002,859	2.8	82,087,860	2.8
합 계	2,966,007,739	100.0	2,966,007,739	100.0

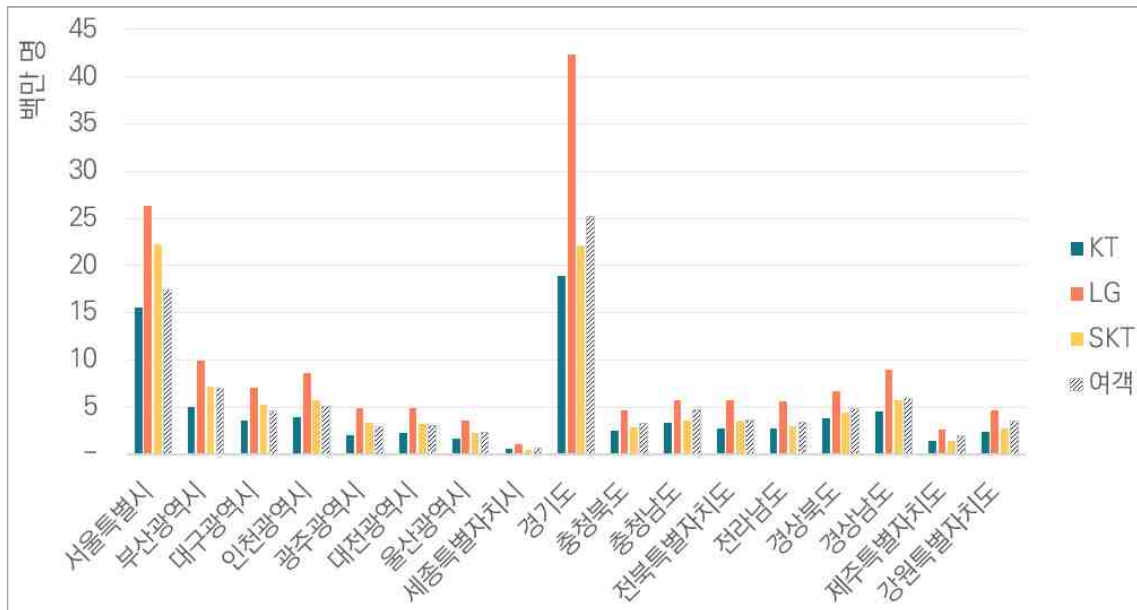
## 2) 조사자료

- 모바일 통신데이터와 상호교차 검증하기 위한 자료로 기존에 교통 수요 분석의 기초자료로 주로 활용되었던 조사자료를 수집함
  - 조사자료 중에서 통행량 정보를 제공하고 있는 KTDB 전국 지역간 목적 OD 자료를 구득함
- 전국 지역간 목적 OD 자료
  - 시간적 범위: 2022년
  - 공간적 범위: 전국 (시도, 시군구 단위)
  - 내용적 범위: 내부통행 포함, 지역간 목적별 통행량
  - 출처: 2024년 국가교통조사사업을 통해 구축한 자료

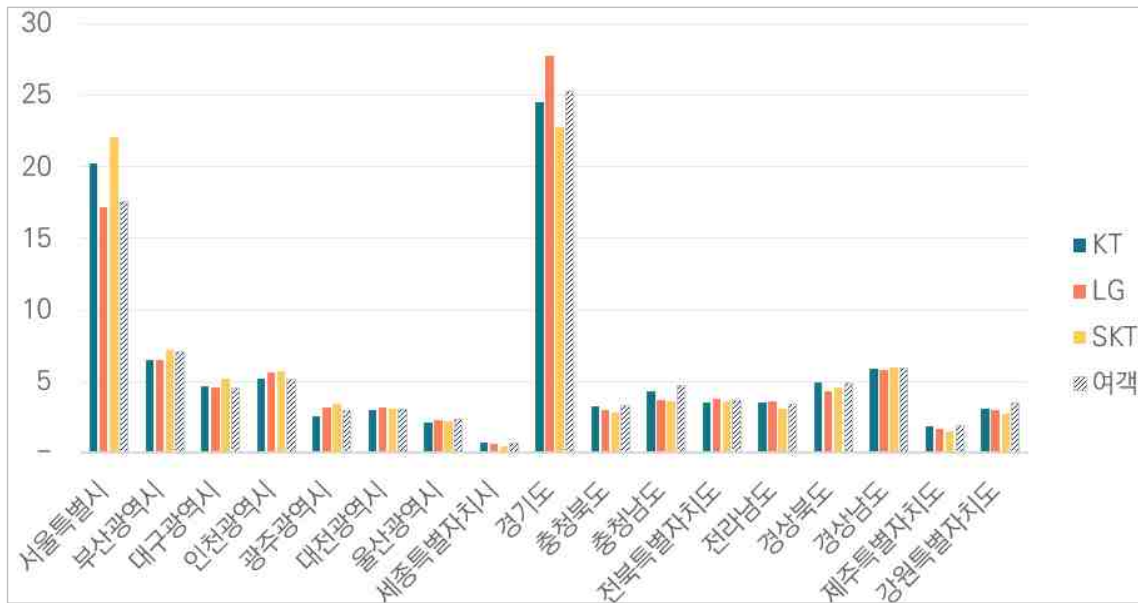
## 제2절 데이터 간 상호교차검증 결과

### 1. 총통행량 & 통행 비율

- 총 통행량 측면에서 모바일 통신데이터 3종의 여객 O/D(내부통행 포함)와 SKT 데이터(내부통행 포함)가 가장 유사하게 나타남
  - 총통행량은 3사 중 LG가 가장 높게 나타남
- 모바일 통신 3사 데이터와 여객 O/D 데이터의 시도별 통행 발생, 도착 비율이 대체로 유사하게 나타나나, 일부 지역은 3사 간 비율 차이가 두드러짐
  - 서울특별시는 KT, SKT 비율이 높게 나타남
  - 반대로 경기도는 LG의 비율이 가장 높게 나타나며, 다음으로 KT, SKT 순임
  - 충청도는 KT사의 통행 비율이 타지역 대비 가장 높게 나타남



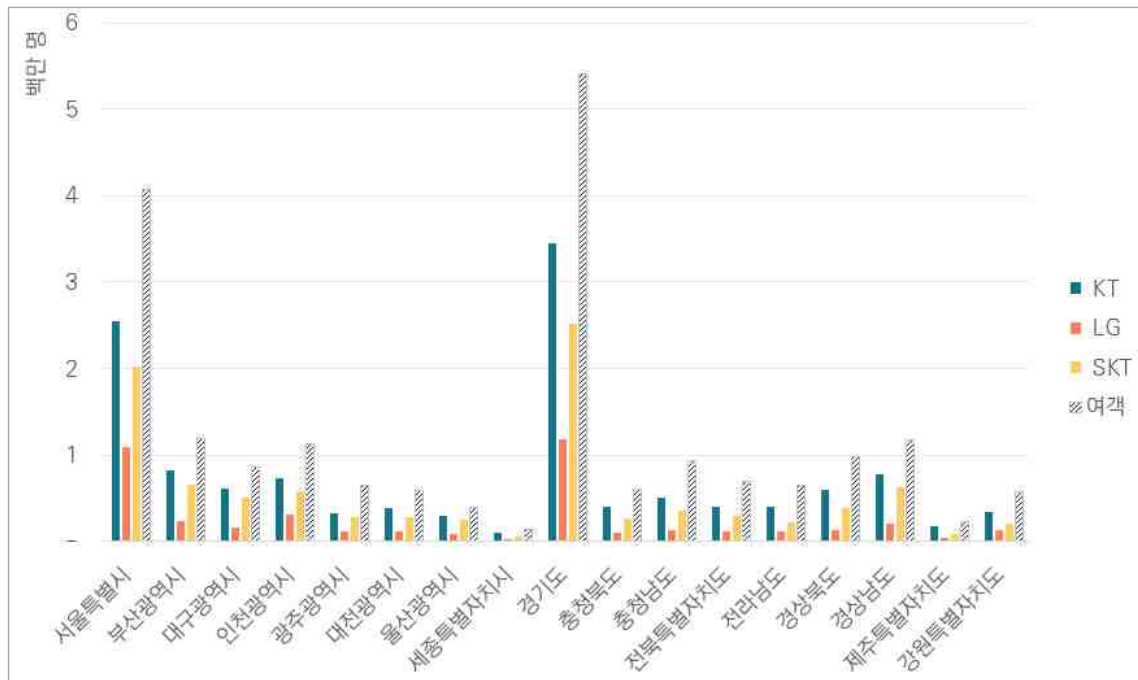
<그림 3-3> 시도 단위 전체 통행의 상호교차검증 결과 (발생량) (단위: 통행/일)



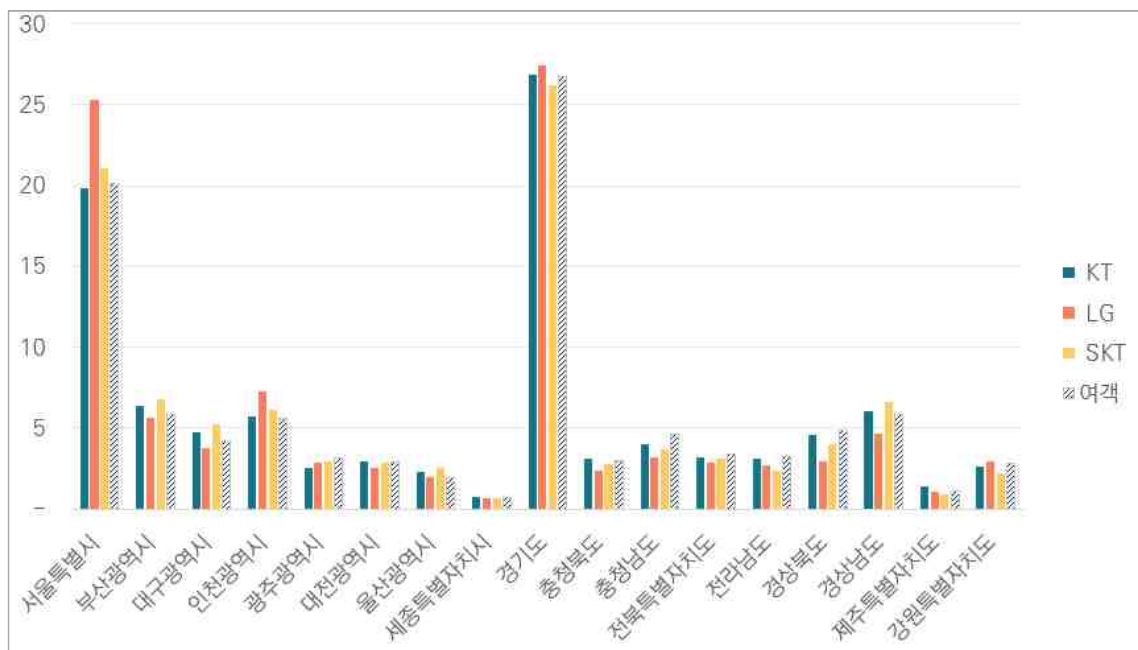
<그림 3-4> 시도 단위 전체 통행의 상호교차검증 결과 (발생량) (단위: %)

## 2. 출근 통행량 & 통행 비율

- 출근 통행량은 여객 O/D의 통행량이 모든 지역에서 모바일 통신 3사에 비해 높게 나타났음
  - 모바일 통신 3사 중에서는 KT, SKT, LG 순으로 통행량이 높게 나타남
- 출근 통행 비율은 모바일 통신 3사와 여객 O/D 데이터가 시도별로 유사하게 나타나나, KT 통행 비율이 여객 O/D와 가장 유사하게 나타남
  - 수도권(서울, 경기, 인천)은 LG의 통행 비율이 가장 높게 나타남



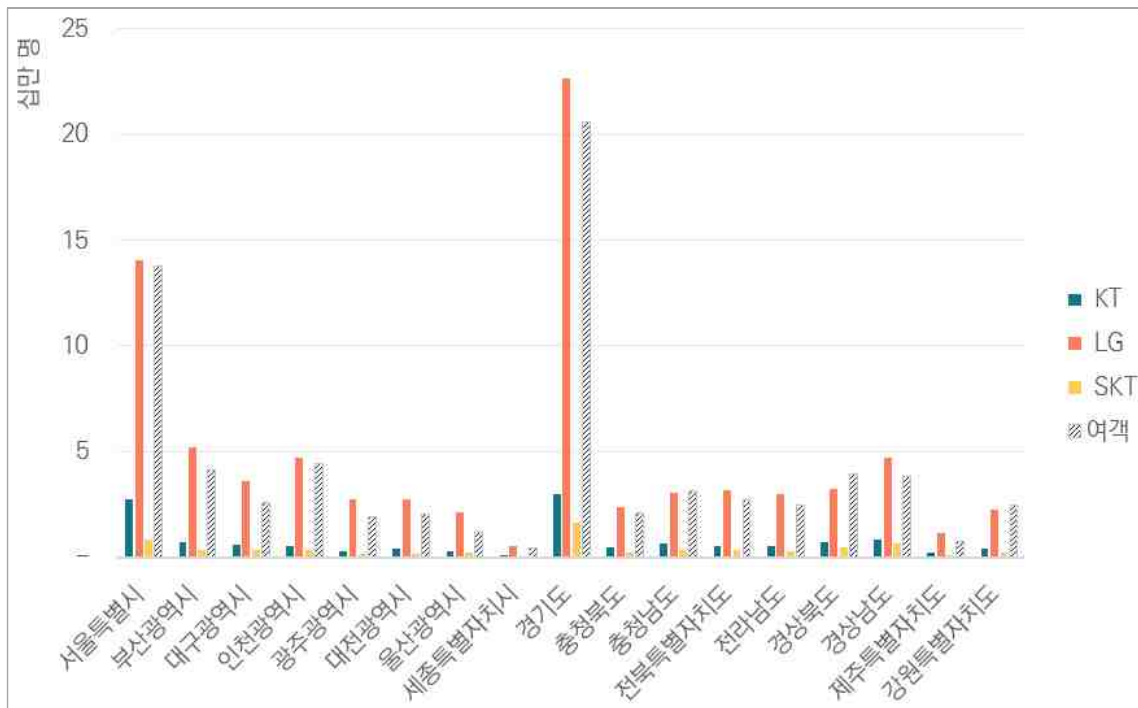
<그림 3-5> 시도 단위 출근 통행의 상호교차검증 결과 (발생량) (단위: 통행/일)



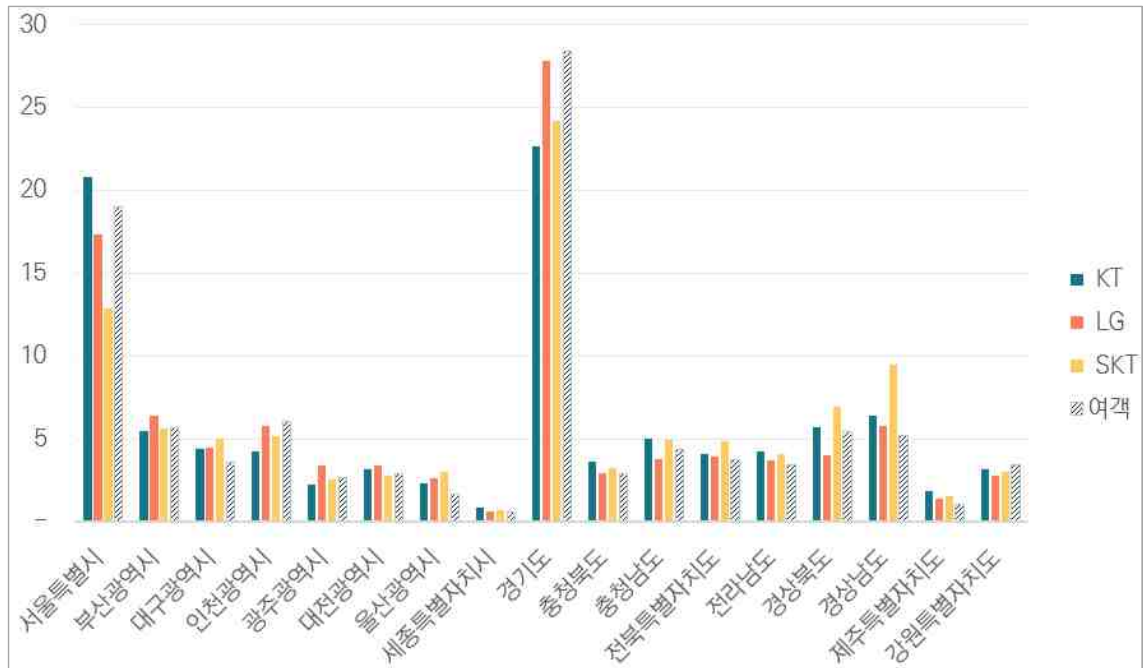
<그림 3-6> 시도 단위 출근 통행의 상호교차검증 결과 (발생량) (단위: %)

### 3. 업무 통행량 & 통행 비율

- 시도별 업무 통행량은 LG가 여객 O/D와 가장 유사하게 나타남
  - KT, SKT는 업무 통행량이 LG에 비해 적으나, 지역별 통행량 경향은 유사함
- 업무 통행 비율 또한 LG가 여객 O/D와 가장 유사하게 나타나며, 지역별로 모바일 3사의 비율이 상이함
  - 서울특별시, 충청도는 KT의 업무통행 비율이 높음
  - 경상북도, 경상남도는 SKT의 시도별 업무 통행 비율이 높게 나타남
  - LG는 경기도, 광역시(부산, 인천, 광주)의 업무 통행 비율이 높게 나타남



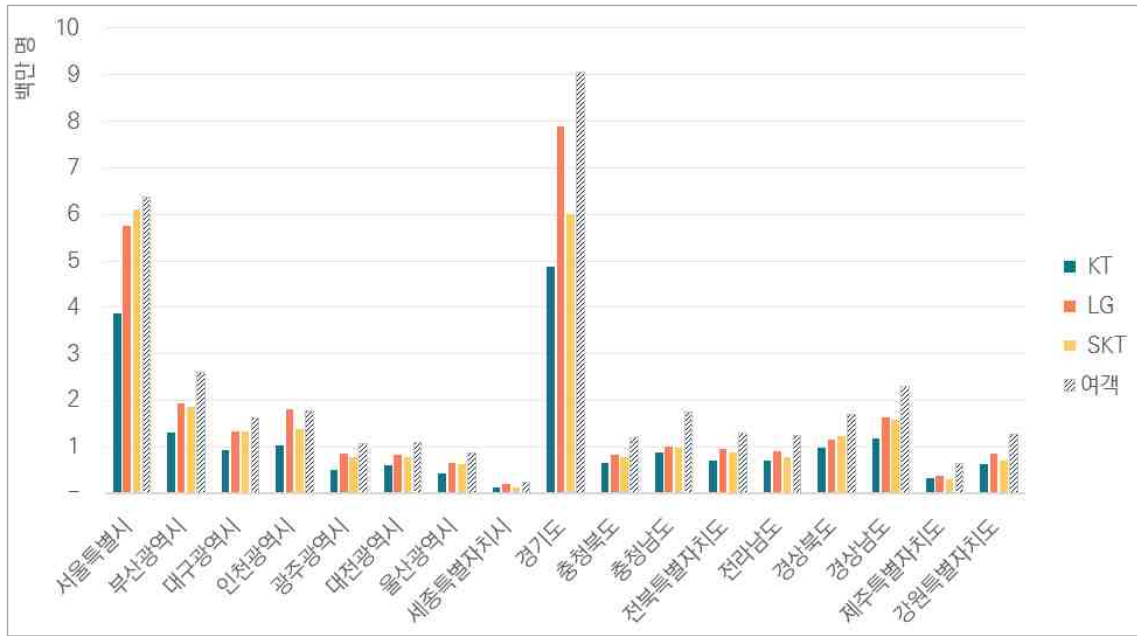
〈그림 3-7〉 시도 단위 업무 통행의 상호교차검증 결과 (발생량) (단위: 통행/일)



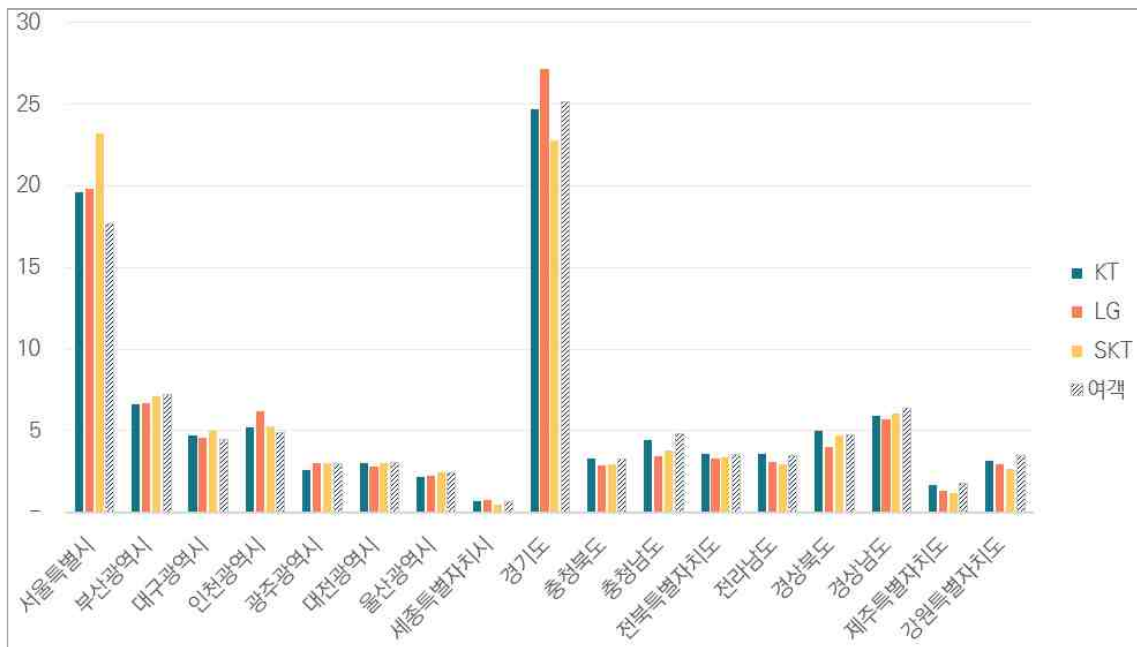
<그림 3-8> 시도 단위 업무 통행의 상호교차검증 결과 (발생량) (단위: %)

#### 4. 귀가 통행량 & 통행 비율

- 귀가 통행량은 여객 O/D이 가장 높게 나타나며, LG 귀가 통행량이 여객 O/D와 유사하게 나타남
  - KT, SKT는 귀가 통행량이 LG에 비해 적으나, 지역별 통행량 경향은 유사함
- 모바일 통신 3사 데이터와 여객 O/D 데이터의 시도별 귀가통행 발생, 도착 비율이 대체로 유사하게 나타나나, 일부 지역은 3사 간 비율 차이가 두드러짐
  - 수도권 중 서울특별시는 SKT의 업무 통행비율이 높게 나타나며, 경기도와 인천광역시는 LG사의 통행비율이 높게 나타남
  - 그 외 충청도, 경상도, 전라도, 제주도, 강원도 지역은 모두 KT사의 귀가 통행 비율이 높게 나타남



<그림 3-9> 시도 단위 귀가 통행의 상호교차검증 결과 (발생량) (단위: 통행/일)



<그림 3-10> 시도 단위 귀가 통행의 상호교차검증 결과 (발생량) (단위: %)





## 제4장 모바일 통신데이터 기반

### 교통지표 구축

제1절 모바일 통신데이터 기반 교통지표 구축

---



## 제4장 모바일 통신데이터 기반 교통지표 구축

### 제1절 모바일 통신데이터 기반 교통지표 구축

#### 1. 개요

- 모바일 통신데이터는 기존 조사 기반 자료에 비해 높은 해상도의 시공간 정보 (1년 365일 1시간 단위 50미터 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	설, 가정의 달, 추석, 하계 휴가기간 등 특정 연휴기간을 설정하여 집계한 통행량	출발일자, 도착일자
		계절별 통행량	봄, 여름, 가을, 겨울에 대해 기간을 설정하고 집계한 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 2024년 사업에서 구축한 통신데이터 기반의 통행DB를 활용함
    - 시간적 범위: 2024년 1월 1일~ 2024년 12월 31일(1년)
    - 공간적 범위: 전국

## 2. 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 4-2>, <표 4-3>, <표 4-4>과 같이 각각 구축함

<표 4-2> 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 4-3〉 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 4-4〉 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 4-5> 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, 하계휴간 기간을 연휴 기간으로 설정함
  - 단, 집계시 출발 시도가 특별시, 광역시가 아닌 경우, 도착 시군구가 시 단위가 아닌 경우는 제거함

**<표 4-5> 연휴기간의 통행량 컬렉션의 구성**

컬렉션명	L0_HOLIDAY_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 4-6> 참조)
    - 성별, 연령대별 통행 차이가 있을 것으로 예상되어 성별, 연령대별 통행량을 구분하여 집계한 필드를 추가함

<표 4-6> 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 4-7> 참조)

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

**<표 4-7> 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 4-8> 과 같이 출발지의 체류유형과 도착지의 체류유형 조합으로 통행목적을 구분 할 수 있음
  - 단, 집(H)→기타 잠재(X)와 집(H)→기타(O) 통행 중 시군구간 이동만 “관광” 목적으로 구분함

<표 4-8> 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 4-9>와 같이 통행목적별 통행량을 구축함

〈표 4-9〉 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), 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 4-10>과 같이 출발지를 기준으로 통행량을 집계한 발생량 DB와 <표 4-11>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 4-10> 고령자 통행량(발생량 기준) 컬렉션의 구성

컬렉션명	viewlight_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 4-11> 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 4-9> 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 4-12>, 퇴근 통행은 <표 4-13> 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 4-12> 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 통행량)
통행량	V	Integer	"V" : 전체 통행량	출근 통행

〈표 4-13〉 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 4-9> 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
  - 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 4-14〉 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 4-15〉 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 4-10>, <표 4-11> 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

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

- 심야시간 출퇴근 통행 비율 DB는 출근 통행 비율과 퇴근 통행 비율을 구분하여 구축함 (<표 4-16> 참조)
  - 단, DB 구축시 통행자의 연령대는 주요 경제활동인구인 20대~60대로 한정함

〈표 4-16〉 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 측정하는 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 4-17> 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, 유동 인구 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 4-18>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 4-18> 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 4-19>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 4-19> 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 4-20〉 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 4-21>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 4-21〉 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 기능해들 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 4-22>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  $\pi$ 로 나눈 후 루트를 씌워서 계산함

〈표 4-22〉 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)



## 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 기존점 통행 DB를 설계하고 구축하였음
  - 모바일 통신 데이터로 구축된 통행DB를 시간대별, 요일별, 연령별 통행량 및 통행목적별 통행시간으로 구분하여 분석함
- 또한, 동일한 기간(2023년 11월)에 생성된 로그 기록에 모바일 통신데이터 가공 알고리즘을 적용하여 통행DB를 검증함
  - 개인별 로그 기록은 「개인정보보호」, 「위치정보의 보호 및 이용 등에 관한 법률」에 저촉되지 않도록 집계한 형태의 DB를 설계하였으며, 이를 기준으로 DB를 구축함
  - 데이터 4종(모바일 통신데이터 3종, 조사자료 1종)의 통행목적별 통행량을 상호 비교함

## 제2절 차년도 수행계획

- 시의성 있는 통행 정보를 제공하기 위하여 2025년 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 2024년 기준의 모바일 통신데이터를 기준으로 본 사업에서 진행한 것과 같이 통행량 및 통행패턴에 대한 데이터 간 상호교차 검증을 수행할 예정임
  - 현재는 통신 3사와 전국 지역간 목적 OD의 총통행, 목적별 통행량을 비교하였음
  - 향후에는 통행량과 사회경제지표 간, 지역·권역·도서지역 간 통행량을 비교하며, 음수셀 발생 여부에 대해서도 파악할 예정임
- 모바일 통신데이터를 활용하여 통행지표를 산출 및 현행화할 예정임
  - 개별 객체 단위 모바일 통신데이터를 활용한 통행지표 산출 방안을 검토할 예정임
  - 2024년 모바일 통신데이터를 이용하여 교통지표(통행량, 통행시간 등)를 현행화할 예정임

2024  
국가교통조사  
및 분석

11



모바일 통신 빅데이터 구축