

## 2022년 국가교통조사

# 차량 GPS 빅데이터 구축

9

2022. 12



국토교통부  
Ministry of Land, Infrastructure and Transport



한국교통연구원  
KOREA TRANSPORT INSTITUTE

# 제 출 문

국토교통부장관 귀하

본 보고서를 「2022년 국가교통조사」 최종보고서로 제출합니다.

2022년 12월

한국교통연구원

원장 오 재 학

본 『2022년 국가교통조사』는 다음 연구진에 의해  
수행되었습니다.

## 참 여 연 구 진

<한국교통연구원>	
연구책임자	<ul style="list-style-type: none"> <li>◦ 조종석 연구위원</li> </ul>
연 구 진	<ul style="list-style-type: none"> <li>◦ 조범철 연구위원</li> <li>◦ 김주영, 천승훈, 박용일 연구위원</li> <li>◦ 황순연, 장동익, 원민수, 이송봉, 이종우 부연구위원</li> <li>◦ 김동호, 신영권 책임전문원</li> <li>◦ 김규진, 김정은 주임전문원</li> <li>◦ 안덕배 전문연구원</li> <li>◦ 가보연, 강국수, 곽명신, 권기훈, 김운태, 김 현, 박미란, 박준호, 신유선, 양태양, 오연선, 이선아, 이슬기, 이채영, 채정표, 홍성표 연구원</li> <li>◦ 홍연우 연구조원</li> </ul>
<한국해양수산개발원>	
연 구 진	<ul style="list-style-type: none"> <li>◦ 이호춘, 최건우 부연구위원</li> <li>◦ 황수진 전문연구원</li> <li>◦ 박일란 선임사무원</li> <li>◦ 류희영 연구원</li> </ul>
<한국항공협회>	
연 구 진	<ul style="list-style-type: none"> <li>◦ 성인영 실장</li> <li>◦ 손병일 책임연구원</li> <li>◦ 최인영 과장</li> <li>◦ 김지한, 김창욱, 김진성 대리</li> </ul>

**『2022년 국가교통조사』**  
**보고서 구성 및 담당연구진**

번 호	과 제 명	연 구 진
제 1권	요약보고서	조종석, 신영권, 가보연
제 2권	전국 여객 O/D 전수화	조종석, 강국수, 박미란, 채정표
제 3권	교통분석용 네트워크 구축	김동호, 이선아, 이슬기
제 4권	항공여객 O/D 조사	한국항공협회
제 5권	전국화물 O/D 본조사	김주영, 황순연, 권기훈, 김정은, 오연선, 김운태
제 6권	전국화물 O/D 보완갱신	김주영, 김정은
제 7권	해상화물 O/D 본조사	한국해양수산개발원
제 8권	KTDB 플랫폼 기반지도 구축	이송봉, 양태양
제 9권	차량 GPS 빅데이터 구축	천승훈, 이종우, 이채영
제10권	모바일통신 빅데이터 구축	원민수, 신유선
제11권	국가교통통계DB구축	박용일, 곽명신
제12권	특별교통대책기간 통행실태조사	안덕배, 김 현
제13권	교통접근성지표 구축	장동익, 홍성표



**『2022년도 국가교통조사』**  
**과제별 공동참여·위탁용역 사업자**

**【공동사업 참여기관】**

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

**【위탁용역 사업자】**

- 전국여객 O/D 전수화 대전세종충청권
  - (주)신명이엔씨, (주)에스트리
- 전국여객 O/D 전수화 제주권
  - 홍익대학교 산학협력단
- 전국여객 O/D 전수화 대구광역시권
  - 홍익대학교 산학협력단
- 전국여객 O/D 전수화 부산울산권
  - (주)신명이엔씨, (주)에스트리
- 도로 및 철도 교통분석용 네트워크 보완갱신
  - 서울시립대학교 산학협력단
- 전국화물 본조사 사업체 및 화물자동차 표본설계
  - (사)한국교통정책경제학회

## 【위탁용역 사업자】

- 전국화물 본조사 사업체물류현황조사  
- ㈜메트릭스
- 전국화물 본조사 화무자동차 통행실태조사  
- ㈜코리아데이터네트워크
- 전국화물 본조사 물류거점조사  
- ㈜코리아데이터네트워크
- 영업용화물자동차운행기록계 빅데이터를 이용한 화물 기종점통행량 및 운행특성 분석  
- ㈜노트스퀘어
- 모빌리티 빅데이터 DB 구축 및 온라인 서비스 유지보수  
- ㈜큐빅웨어
- 빅데이터 플랫폼사업 감리용역  
- 악티보
- 모바일통신 원시 데이터를 이용한 통행사슬 DB 구축  
- ㈜오픈메이트온
- 모바일통신 원시 데이터 전처리 최적화 및 시스템 연결  
- ㈜ 오픈메이트온
- 특별교통통행실태조사 및 이용자 만족도 조사  
- ㈜컨슈머인사이트

## 최종보고서 목차

- 제 1권 요약보고서
- 제 2권 전국여객 O/D 전수화
- 제 3권 교통분석용 네트워크 구축
- 제 4권 항공여객 O/D 조사
- 제 5권 전국화물 O/D 본조사
- 제 6권 전국화물 O/D 보완갱신
- 제 7권 해상화물 O/D 본조사
- 제 8권 KTDB 플랫폼 기반지도 구축
- 제 9권 차량 GPS 빅데이터 구축
- 제 10권 모바일통신 빅데이터 구축
- 제 11권 국가교통통계DB구축
- 제 12권 특별교통대책기간 통행실태조사
- 제 13권 교통접근성지표 구축

# 목 차

## 요 약

제1장 과업의 개요 .....	3
------------------	---

제1절 과업의 배경 및 목적 .....	3
-----------------------	---

제2절 과업의 범위 및 내용 .....	5
-----------------------	---

제2장 차량 GPS 데이터 가공 및 DB 구축 .....	9
---------------------------------	---

제1절 원시데이터 수집 및 전처리 .....	9
--------------------------	---

제2절 차량 경로데이터 가공 및 기초교통DB 구축 .....	22
-----------------------------------	----

제3절 차량 통행지표 산출 및 DB구축 .....	41
-----------------------------	----

제3장 차량 GPS 데이터 기반 온라인 서비스 개선 및 현행화 .....	47
--	----

제1절 View-T 시스템 개요 .....	47
-------------------------	----

제2절 신규 개발 및 기능 개선 .....	55
-------------------------	----

제3절 통행지표 현행화 .....	67
--------------------	----

제4절 분석도구 현행화 .....	71
--------------------	----

제4절 View-T 운영 및 유지보수 .....	77
----------------------------	----

제4장 결론 및 향후과제 .....	85
---------------------	----

제1절 결론 .....	85
--------------	----

제2절 향후과제 .....	86
----------------	----

## 표 목 차

〈표 2-1〉 차량 GPS 빅데이터 특징 (2020년 데이터 기준) .....	9
〈표 2-2〉 A사의 내비게이션 데이터 형식 .....	10
〈표 2-3〉 A사의 내비게이션 데이터 현황 .....	11
〈표 2-4〉 B사의 내비게이션 데이터 형식 .....	12
〈표 2-5〉 B사의 내비게이션 데이터 현황 .....	12
〈표 2-6〉 DTG 데이터 형식 .....	13
〈표 2-7〉 DTG 데이터 현황 .....	14
〈표 2-8〉 A사의 내비게이션 데이터 시간적 범위 .....	15
〈표 2-9〉 A사의 도로 네트워크 데이터 형식 .....	16
〈표 2-10〉 관측교통량 대비 A사의 내비게이션 데이터 샘플을 분석 결과 .....	18
〈표 2-11〉 교통량 수준별 샘플을 .....	19
〈표 2-12〉 A사 내비게이션 데이터의 조정된 데이터 형식 .....	21
〈표 2-13〉 개별 차량 경로데이터 테이블 구성 .....	30
〈표 2-14〉 시도별 평균속도 증감패턴 분석 .....	33
〈표 2-15〉 상세도로망 Level6 링크 단위 속도프로파일 DB 테이블 구성 .....	34
〈표 2-16〉 지역간 도로 관측교통량 증감패턴 분석 .....	34
〈표 2-17〉 도시부 도로 관측교통량 증감패턴 분석 .....	35
〈표 2-18〉 머신러닝 기반 추정교통량 모형 구축 프로세스 .....	38
〈표 2-19〉 2020년 상시 지점 추정교통량 신뢰도 분석 .....	40
〈표 2-20〉 2020년 주요 도로등급별 추정교통량 신뢰도 분석 .....	40
〈표 2-21〉 추정교통량 DB 테이블 구성 .....	41
〈표 2-22〉 평균속도 DB 테이블 구성 .....	42
〈표 2-23〉 통행지표 DB 생성기준 .....	44
〈표 2-24〉 추정교통량 데이터 테이블 구성(상세도로망 Level6 도로 단위) .....	44
〈표 2-25〉 온라인 서비스 지표 .....	46
〈표 3-1〉 View-T Light 주요 기능 .....	50
〈표 3-2〉 View-T Expert 통행지표 .....	51
〈표 3-3〉 View-T Expert 분석 도구 및 시뮬레이션 .....	54
〈표 3-4〉 View-T 대시보드 주요 기능 설명 .....	55

〈표 3-5〉 View-T 부가 기능 설명 .....	56
〈표 3-6〉 Faster Indicator 데이터 테이블 정의 .....	61
〈표 3-7〉 View-T 웹서비스용 DB 데이터 업데이트 내용 .....	80
〈표 3-8〉 View-T 웹 서비스 성능 테스트 .....	82

## 그림 목차

〈그림 2-1〉 A사의 도로 네트워크와 KTDB 상세도로망 Level6 네트워크 상세도 비교	17
〈그림 2-2〉 개별 차량 속도 플로팅 비교	20
〈그림 2-3〉 속도 제한 상태 비교	21
〈그림 2-4〉 차량 GPS 전처리 및 데이터 가공 프로세스	22
〈그림 2-5〉 링크 기반 데이터 전처리 및 경로 데이터 가공 프로세스	23
〈그림 2-6〉 보정된 데이터 전·후 비교	25
〈그림 2-7〉 1초 단위 궤적(포인트) 데이터 생성 결과	25
〈그림 2-8〉 포인트 기반 데이터 전처리 및 경로 데이터 가공 프로세스	26
〈그림 2-9〉 통행 분리되지 않은 차종별 궤적	28
〈그림 2-10〉 차량 GPS 데이터와 링크 맵 매칭 및 경로 생성 프로세스	29
〈그림 2-11〉 통행 병합 프로세스	29
〈그림 2-12〉 압축된 경로 데이터 형태 예시	31
〈그림 2-13〉 속도프로파일 구축	33
〈그림 14〉 머신러닝 기반 교통량 추정 알고리즘 개발 프로세스	36
〈그림 2-15〉 추정교통량 구축 프로세스	37
〈그림 2-16〉 연도별 신뢰도 분석 결과	40
〈그림 2-17〉 관측교통량과 추정교통량의 패턴 분석	41
〈그림 2-18〉 통행지표 구축 프로세스	43
〈그림 2-19〉 차량 온라인 서비스 제공을 위한 DB 구축 프로세스	45
〈그림 3-1〉 View-T 시스템 구성도	49
〈그림 3-2〉 Faster Indicator 화면	57
〈그림 3-3〉 Faster Indicator 분석 조건 화면	58
〈그림 3-4〉 Faster Indicator 조회 및 다운로드 버튼	58
〈그림 3-5〉 Faster Indicator 신청서 작성 화면	59
〈그림 3-6〉 Faster Indicator 분석 다운로드 결과 화면	59
〈그림 3-7〉 Faster Indicator 분석 결과 화면	60
〈그림 3-8〉 오픈소스 활용 서비스 오픈소스 소프트웨어 화면	62
〈그림 3-9〉 Python, R 프로그램 언어 활용 화면	63
〈그림 3-10〉 오픈소스 소프트웨어 사용자 관리 화면	63

〈그림 3-11〉 관리자 다운로드 현황 화면 .....	64
〈그림 3-12〉 분석도구 차량 유지보수 전후 비교 화면 .....	65
〈그림 3-13〉 홈페이지 로그인, MyPage 삭제 화면 .....	66
〈그림 3-14〉 데이터 다운로드 단계 축소 .....	66
〈그림 3-15〉 메뉴구조도 .....	67
〈그림 3-16〉 Track API 제공 .....	67
〈그림 3-17〉 교통량 지표 화면예시 .....	68
〈그림 3-18〉 속도 지표 화면예시 .....	69
〈그림 3-19〉 혼잡 지표 화면예시 .....	70
〈그림 3-20〉 환경 지표 화면예시 .....	71
〈그림 3-21〉 기종점 차량 통행량 분석 디폴트 변경 화면 .....	72
〈그림 3-22〉 기종점 차량 경로 분석도구 화면예시 .....	73
〈그림 3-23〉 혼잡도로 선정 분석 디폴트 변경 화면 .....	74
〈그림 3-24〉 시·공간 혼잡 분석도구 화면예시 .....	75
〈그림 3-25〉 혼잡도로 통행경로 분석도구 디폴트 변경 화면 .....	76
〈그림 3-26〉 출·퇴근 영향권 분석도구 디폴트 변경화면 .....	77
〈그림 3-27〉 MongoDB 샤딩 구조 .....	78
〈그림 3-28〉 레인지 샤딩의 구성 화면 .....	79
〈그림 3-29〉 해쉬 샤딩의 구성 .....	79
〈그림 3-30〉 DB 장애 대응 전략 .....	80
〈그림 3-31〉 View-T 데이터 가공 성능 테스트 시나리오 .....	81
〈그림 3-32〉 구글에서 제공하는 웹페이지 속도 테스트 화면 .....	83
〈그림 3-33〉 자체 동작 테스트로 발견된 오류 및 추가 개선사항의 목록 .....	84
〈그림 3-34〉 웹 사이트 동작 테스트 프로세스 .....	84





요 약

---



## 요 약

### 제1절 과업의 개요

#### 1. 과업의 배경 및 목적

##### 가. 과업의 배경

- 전국 단위 기초 교통데이터 수집 부족
  - 교통분야의 기초 데이터인 교통량, 속도는 한국건설기술연구원, 각 지자체에서 조사하고 있으나, 교통량, 속도 자료 수집의 공간적 커버리지<sup>1)</sup>가 매우 낮아 전국의 기초 교통 현황을 모니터링 하기에 한계
- 지점 데이터 분석의 한계
  - 공공에서 수집하고 있는 데이터는 지점정보 중심의 데이터로 지점의 단편적 정보만 확인 할 수 있어 시공간적으로 연결된 교통의 흐름을 파악하기에 한계가 있음
- 빅데이터 전처리·가공 환경의 부재
  - 휘발성·단발성 빅데이터 관련 사업은 연속성 있는 빅데이터 전처리·가공환경 구축에 한계가 있으며, 대용량 데이터를 대용량 네트워크에서 효율적으로 전처리·가공하기 위한 환경 부족
- 공급자 중심의 데이터 제공환경
  - 중앙정부, 지자체, 연구기관, 학계 등 다양한 기관과 분야에서 교통 데이터를 요구하고 있으나, 공급자 기반의 데이터 제공환경에서는 대응하기 어려운 한계가 있음
- 전국 단위 일관된 교통DB 및 통행지표 부재
  - 데이터 수집기관, 수집방식 등 지역별로 상이하게 수집되는 데이터 수집 체계는 전국을 일관된 기준으로 평가하고 분석하기에 한계
- 차량이동 분석시스템 부재
  - 과거 차량의 이동의 시공간적 통행패턴과 현상을 분석하기 위해서는 모형 중심의 프로

1) 공공에서 수집하고 있는 교통량, 속도는 전국의 약 3%, 16%정도의 커버리지에 불과

그램을 이용해야 하기 때문에 많은 인력·시간·예산이 소모되었고 분석결과와 현실성이 저하되는 문제 발생

- 차량 GPS 빅데이터 구축 및 활용 필요
  - 전국 도로에서 수집되는 차량 GPS 빅데이터를 활용하여 빅데이터 전처리·가공환경 구축, 전국단위 교통DB 구축, 차량이동 분석 시스템 구축을 통해 과거의 한계를 개선하고 국가교통DB의 신뢰성과 활용성을 높이기 위한 새로운 기반환경 구축 필요

#### 나. 과업의 목적

- 차량 GPS 빅데이터 기반 시계열 전국 단위 데이터 전처리 및 가공
  - 매년 사업을 통해 구축된 교통DB를 통해 연도별 교통변화를 반영한 전국 단위 데이터 전처리 및 가공
- 국가교통데이터베이스 구축 및 교통 현황 모니터링을 위한 기초 교통DB, 통행지표 산출
  - 교통량, 속도 등 교통 분야의 핵심 기초교통 DB 구축
  - 교통혼잡비용, 혼잡강도, 이산화탄소 배출량 등 도로를 다각도로 분석할 수 있는 도로 교통망 성능평가 지표 구축
- 교통수요분석 DB구축 신뢰도 제고
  - 가구통행조사 기반의 인력조사에서 빅데이터로 전환하는 시점에서 교통수요분석DB의 구축 및 검증을 위한 기반 데이터 제공 및 분석지원
- 통행지표 및 데이터 제공 환경 개선
  - 이용자의 요구에 대응할 수 있는 온라인 기반 데이터 제공 환경 개선
  - 차량의 통행행태와 특성을 시공간적으로 분석할 수 있는 분석도구 개선
- 교통분야 주요 정책 및 지자체 지원
  - 예비타당성조사, 교통영향평가 등 주요 정책 지원
  - 지자체 교통현안 문제 해결을 위한 지자체 실증사업 지원

## 2. 과업의 범위 및 내용

### 가. 공간적 범위

- 전국 2차로 이상 도로 및 주요 도로<sup>2)</sup>

### 나. 시간적 범위

- 2020년, 2021년 기준 데이터

### 다. 내용적 범위

- 차량 GPS 데이터 가공 및 DB 구축
  - 원시 데이터 전처리 및 경로데이터 가공
  - 차량 GPS 원시 데이터 검증 및 오류 필터링
  - 차량 GPS 데이터 확대 구축으로 데이터 전처리 가공 알고리즘 개선
  - 기초 교통DB가공 알고리즘 개선
  - 통행지표 구축을 위한 기초 교통 DB구축
  - 구축 데이터 및 DB 검증 및 분석
- 차량 GPS 데이터 기반 온라인 서비스 개선 및 현행화
  - 서비스 현행화 및 이용자 편의성 개선
  - 데이터 개방 확대를 위한 데이터 제공 관련 신규 기능 개발
  - 주요 정책 및 지자체 지원을 위한 데이터 제공
  - View-T 온라인 서비스 운영 및 유지보수

---

<sup>2)</sup> 주요 도로: 편도 1차로 도로 중 데이터 수집 안정성이 확보된 도로

## 제2절 차량 GPS 데이터 가공 및 DB 구축

### 1. 원시데이터 수집 및 전처리

#### 가. 차량 GPS 데이터 수집 개요

- 본 사업을 통해 수집한 차량 GPS 데이터는 A사의 내비게이션 데이터와 B사의 내비게이션 데이터, 한국교통안전공단의 DTG데이터이며, 전국 단위 기초 교통DB 및 통행지표를 산출하기 위한 기반 데이터임
- 차량 GPS 데이터는 도로망 네트워크의 링크 기반으로 수집되는 링크 기반 데이터와 X, Y의 위치좌표인 GPS정보가 수집되는 포인트 기반 데이터로 구분됨
  - A사의 내비게이션 데이터는 도로망 네트워크와 연결된 링크 기반 데이터이며, DTG 데이터와 B사의 내비게이션 데이터는 GPS좌표인 포인트 기반 데이터로 데이터 수집방식이 다르나, 차량의 이동경로에 대해 시간의 연속성과 공간의 연결성이 동시에 수집되는 데이터임

<표 9-1> 차량 GPS 빅데이터 특징 (2020년 데이터 기준)

구분	A사 내비게이션 데이터	B사 내비게이션 데이터	DTG 데이터
제공 파일 구성	18,140개 텍스트 파일	1일 단위 텍스트 파일	한달 400개 파일
OBU ID 생성기준	경로 ID 발생	1일 단위 ID 갱신	차량등록번호로 ID 유지
수집정보 생성기준	내비게이션 실행 시	내비게이션 실행 시	차량 운행 시 수집
데이터 형태	링크 단위	1초 단위 포인트	1초 단위 포인트
좌표체계	EPSG 4301	WGS84	WGS84
용량(년)	7TB	4TB	248TB
이벤트(억/년)	520	2,300	7,300

- 작년 사업에서는 B사의 내비게이션 데이터만 이용하여 기초 교통DB 및 통행지표를 가공·구축하였음
- 시의성 있는 데이터 제공을 위해 최신년도 데이터 수집이 필요하며, 샘플율이 높은 내비게이션 데이터로 데이터의 신뢰도를 높이하고자 A사의 내비게이션 데이터를 추가로 수집하였음

## 1) 차량 내비게이션 데이터

- 차량 내비게이션 데이터는 포인트 기반으로 제공되는 B사의 내비게이션 데이터와 링크 기반으로 제공되는 A사의 내비게이션 데이터로 구성
  - A사의 내비게이션 데이터는 링크 기반으로 경로 데이터가 제공되며, 링크 기반의 진입/진출 시간 정보 속도 등의 정보가 제공됨
  - A사의 내비게이션 주행 데이터는 A사에서 제공하는 네트워크와 매칭해야하며, 시기마다 주행 고유ID가 변경되는 특징을 가지고 있음
  - 주행 고유ID는 목적지를 설정하고 주행할 때마다 생성되어 동일한 사용자라도 주행 별로 다른 ID가 부여됨
  - 휴게소나 주정차 시 내비 종료 상태가 아니라면 정차상태로 데이터가 수집되지만, 종료 상태라면 경로를 임의로 생성하거나 보정하지 않음

&lt;표 9-2&gt; A사의 내비게이션 데이터 형식

항목	설명
데이터 발생 일자	최초 seq가 발생한 날짜
주행 고유 ID	비식별화 처리된 ID, 24WKFL
순번	주행 내 링크 순번
SK MESH 코드	SK Mesh의 앞 4자리
LINK ID	SK내의 LINK ID
LINK 방향	0: 정방향 / 1: 역방향
LINK 길이	단위 : 미터(m)
주행시간	LINK 주행 시간(초)
진입시간	링크 진입시간
진출시간	링크 진출시간
속도	GPS 데이터를 통해 계산된 속도



- B사의 내비게이션 데이터는 1일 단위 텍스트 파일로 제공되며, 수집 주기는 1초 단위로 경로ID, 시간정보, 위치정보, 방향각 등의 정보가 제공됨

<표 9-3> B사의 내비게이션 데이터 형식

항목	설명
단말기 ID	
시간정보	Timestamp
위도	dddmm,mmmm*10,000
경도	dddmm,mmmm*10,000
속도	
방위각	
도로판단	자사지도 기준 - 도로주행여부
경로탐색 플래그	주행(경로탐색주행, 일반운행)
재 탐색 실행	경로이탈로 인한 경로탐색여부
고도 값 부호	양수(0), 음수(1)
고도값	음수의 경우 -> 1값만 존재

## 2) DTG(차량디지털운행기록) 데이터

- DTG 데이터는 영업용 자동차의 운행정보를 실시간으로 저장하여 변화하는 운행상황을 기록하는 디지털 운행기록계를 통해 수집되는 데이터이며, 운전자가 차량디지털운행기록을 제출하는 시기에 따라 데이터의 샘플율이 달라질 수 있음
- DTG 데이터는 자동차의 순간속도, GPS, 분당 엔진회전수, 가속도, 자동차 유형 등의 정보가 제공되며, 1초 단위의 GPS 좌표로 수집됨
- DTG 데이터의 자동차 유형 정보를 이용하여 세분화된 차종의 차량의 이동경로, 통행에 대한 분석을 수행할 수 있음

&lt;표 9-4&gt; DTG 데이터 형식

컬럼명	설명	정보 예시
TRIP_KEY	등록번호 & 정보발생 일시	C-125901568017101206094700
DTG_MDL_NM	단말기 모델명	XDT1000
CHASSIS_NO	차량 고유번호 1	XXXXXXX301795
CAR_TP_CD	차량 유형 구분	11
CAR_REG_NO	차량 고유번호 2	-1259015680
BIZ_REG_NO	운송사업자 등록번호	XXXXXX47349
DRIVER_CODE	운전자 코드	1
DLY_DRIV_DIST	일일 주행거리	0
ACCM_DRIV_DIST	누적 주행거리	268897
OPT_SPD	차량속도	0
RPM	분당 엔진 회전수	393
BREAK_SIG	1: 유, 2: 무	0
GPS_X	차량위치 X (WGS84 경위도)	127075626
GPS_Y	차량위치 Y (WGS84 경위도)	37052681
GPS_AGL	지점별 방위각	0
ACCEL_VX	횡가속	1
ACCEL_VY	종가속	1
CONT_CODE	통신상태코드	11
AREA_CODE	행정기관코드, 대존코드	41
OPT_Time	YYMMDDHHMMSSSS	17101206094700

#### 나. 차량 GPS 원시 데이터 분석

- A사의 내비게이션 데이터는 2020년, 2021년을 24주차로 구분하여 총 180일의 데이터를 수집하였으며, 파일은 37,310개로 분할되어 있음

<표 9-5> A사 내비게이션 데이터의 시간적 범위

구분	2020년	2021년
1월	21일~27일(4주차)	18일~24일(4주차)
2월	10일~16일(3주차)	8일~14일(2주차)
3월	16일~22일(3주차)	15일~21일(3주차)
4월	13일~19일(3주차)	19일~25일(4주차)
5월	18일~24일(4주차)	17일~23일(4주차)
6월	15일~21일(3주차)	14일~20일(3주차)
...	...	...
10월	9일~15일(2,3주차)	8~14일(2,3주차)
11월	16~22일(3주차)	15~21일(3주차)
12월	14일~20일(3주차)	13일~19일(3주차)
명절	- 설날: 1월 24일~26일(3일) - 추석: 9월 29일~10월 5일(7일)	- 설날: 2월 11일~13일(3일) - 추석: 10월 9일~11일(3일)
합계	(7일×12주차) + 6일 = 90일	(7일×12주차) + 6일 = 90일

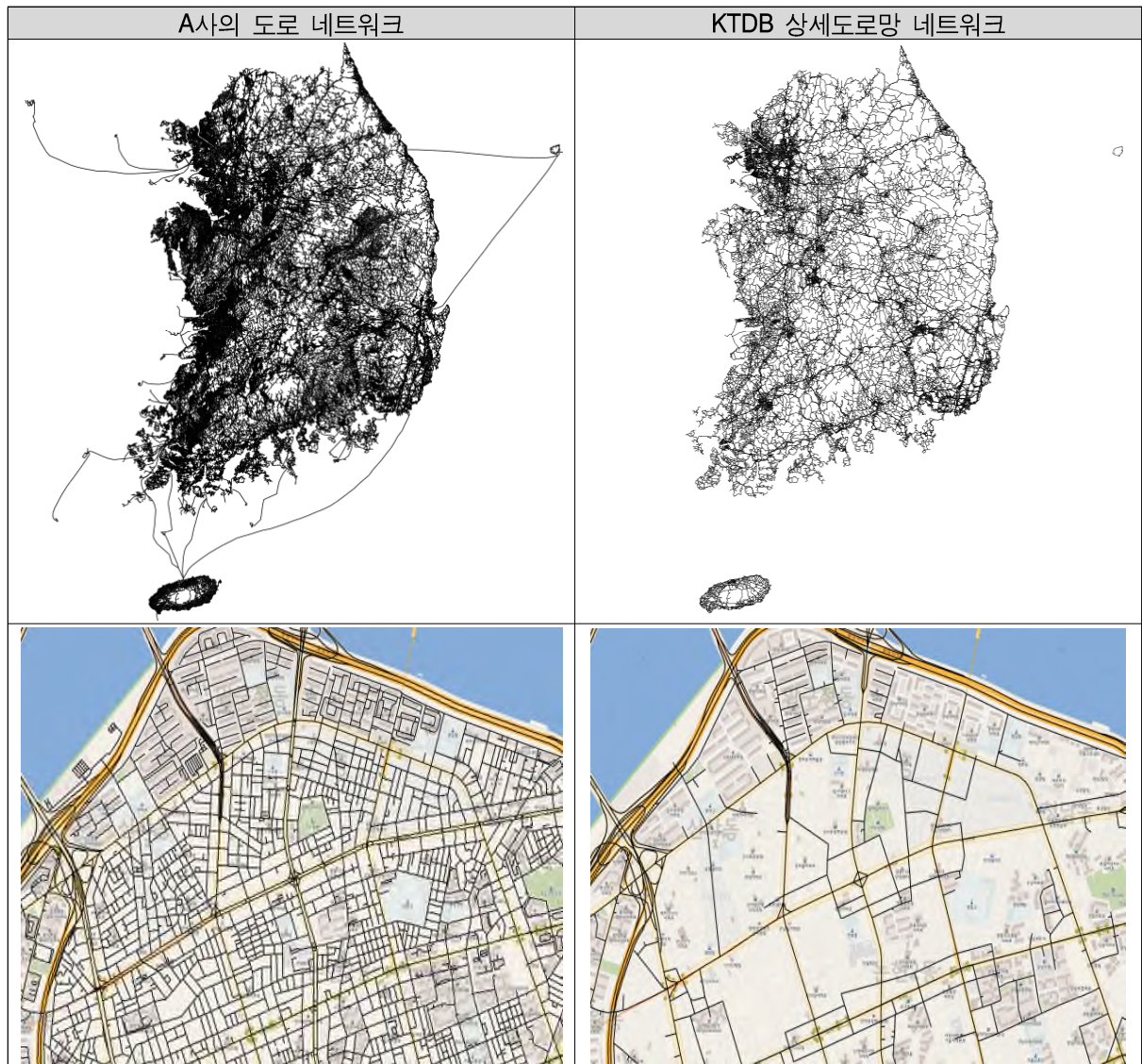
- 수집된 A사 내비게이션 데이터의 전체 용량은 13.6TB이며, 총 천 억건의 이벤트 수(라인수)가 발생되었으며, 개별 차량 경로수는 일 평균 약 600만 개 경로가 나타났음
  - 2020년의 데이터 용량은 6.40TB, 이벤트 수는 약 520억 건, 경로 수는 약 4.6억 건 2021년의 데이터 용량은 7.20TB, 이벤트 수는 약 586억 건, 경로 수는 약 5.4 억건으로 분석됨
  - 코로나 19 영향으로 2020년은 2021년과 비교했을 때 상반기의 이벤트 수, 경로 수가 적게 나타났으며, 2020년 10월에 코로나 19 확진자 수가 감소 추세가 있어 사회적 거리두기가 완화된에 따라 이벤트 수, 경로 수가 급격히 증가한 것으로 분석됨

- A사 내비게이션 데이터는 링크 기반 데이터로 A사의 도로 네트워크 내의 링크ID를 바탕으로 정보를 제공
- A사의 도로 네트워크 데이터는 시기별로 링크 정보를 1개월 2회 업데이트하기 때문에 2020년, 2021년 각 12주차 주행 데이터와 매칭할 수 있도록 네트워크 Shape파일을 24개 버전을 이용해야 함
- 본 사업에서는 KTDB의 상세도로망 Level6 네트워크로 차량 GPS 빅데이터 가공 및 지표를 생성하기 때문에 SK 도로 네트워크의 링크ID로 제공되는 경로를 Level6 네트워크 기반의 경로로 재생성 필요

<표 9-6> A사의 도로 네트워크 데이터 형식

컬럼명	설명
IDXNAME	링크의 MESH 코드
LINK_ID	주행 고유 ID
TLINKIDP1	정방향 교통정보 LINK ID
TLINKIDN1	역방향 교통정보 LINK ID
LENGTH	MESH 코드
ST_DIR	링크의 MESH 내 시작 노드의 NODE ID
ED_DIR	링크의 MESH 내 끝 노드의 NODE ID
ROAD_CATE	도로등급
ROADLEVEL	경로안내를 위해 수정된 링크의 도로등급
ROADSTATE	도로의 포장상태(0: 미조사 / 1: 비포장/ 2: 포장)
...	...
KEY	이종데이터와의 매칭을 위한 값
KS1	정방향 국토부교통정보 LINK ID
KS2	역방향 국토부교통정보 LINK ID

- A사의 도로 네트워크의 링크를 상세도로망 Level6 링크로 경로 재생성 모듈을 개발하기 위해서는 SK 도로 네트워크의 링크와 Level6 링크의 차이를 분석함
- A사의 도로 네트워크의 단방향 링크 연장은 447,610km(2021년 12월 기준)로 상세도로망 Level6 네트워크 연장인 126,040km(2021년 12월 기준)에 비해 3.5배 상세한 수준인 것으로 분석됨



<그림 9-1> A사의 도로 네트워크와 KTDB 상세도로망 Level6 네트워크 상세도 비교

○ A사의 경로 데이터 샘플을 분석

- 전체 교통량 대비 A사의 데이터의 샘플율을 분석하기 위해 관측교통량과 A사의 내비게이션 통행량 자료의 비교를 수행함
- 이를 위해 고속도로, 도시고속도로, 광역시도, 일반국도, 지방도에 대한 관측지점을 선정하였으며, 교통량 수준을 고려하여 선정하였음
- 고속도로와 도시고속도로와 같은 연속류 도로에서는 A사 내비게이션 통행량 샘플율이 약 33%~40% 수준으로 나타나 연속류 도로의 샘플율이 매우 높은 것으로 분석됨
- B사의 내비게이션 통행량은 약 0.1%~1.0% 수준으로 낮은 샘플율을 보이고 있음

- 특히 교통량 수준이 낮은 500대 수준의 도로에서도 A사의 샘플율이 18% 수준으로 나타나 저규격 도로에서도 높은 샘플율을 확보하고 있는 것으로 분석됨

<표 9-7> 관측교통량 대비 A사의 데이터 샘플율 분석 결과

구분	도로명	관측교통량 <sup>3)</sup> (대/일)	A사 통행량 <sup>4)</sup> (대/일)	A사 샘플율(%)	B사 통행량 <sup>5)</sup> (대/일)	B사 샘플율(%)
고속도로	경부고속도로 (관교→신갈)	101,152	33,515	33.1%	129	0.1%
	서해고속도로	68,259	27,500	40.3%	559	0.8%
도시고속 도로	서울외곽순환고속도 로 (청계TG)	80,165	33,571	41.9%	775	1.0%
광역시도	명덕로	6,982	1,561	22.4%	53	0.8%
일반국도	서동대로	3,643	967	26.5%	28	0.8%
	중원대로	5,825	1,100	18.9%	33	0.6%
지방도	한불로	2,062	138	6.7%	5	0.2%
	쌍곡로	545	99	18.2%	2	0.3%

3) 2020년 기준 관측교통량 데이터를 사용하였음

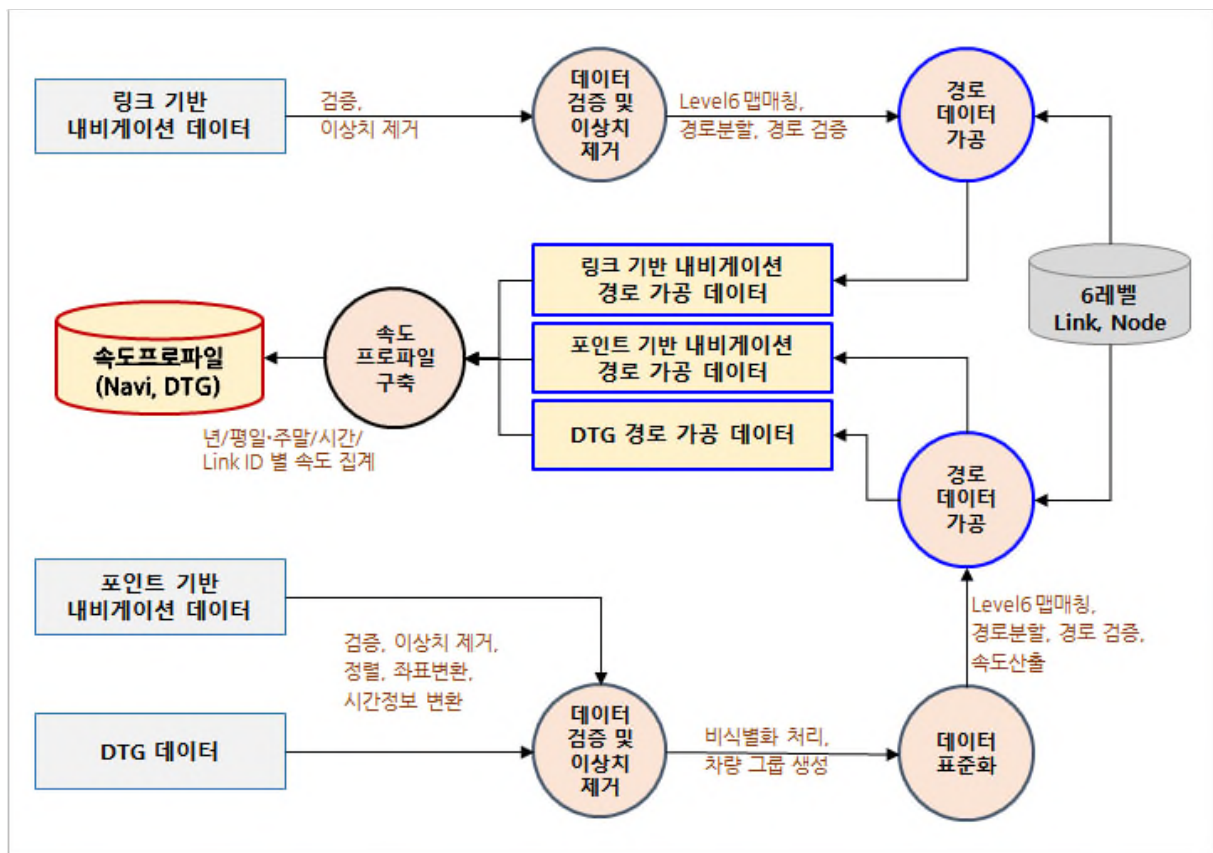
4) A사 통행량 데이터는 2021년 10월 8일 ~ 14일 일주일을 대상으로 산출한 일평균 통행량 데이터임

5) B사 통행량 데이터는 2020년 1년을 대상으로 산출한 일평균 통행량 데이터임

## 2. 차량 경로 데이터 가공 및 기초교통DB 구축

### 가. 경로 데이터 가공 개요

- 원시 데이터는 링크 기반의 내비게이션 데이터, 포인트 기반의 내비게이션 데이터로 구분되며, 각각의 제공되는 데이터 포맷 및 데이터 특성이 서로 상이하여, 이를 고려한 원시 데이터 전처리 및 경로 데이터 가공이 필요함

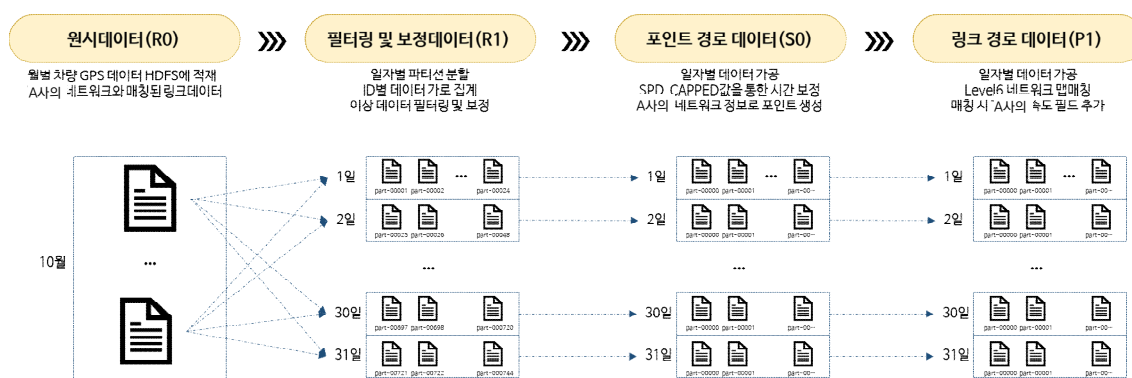


<그림 9-2> 차량 GPS 전처리 및 데이터 가공 프로세스

- 링크 기반으로 제공되는 A사 내비게이션 데이터는 A사 도로 네트워크의 링크관련 정보가 제공되어, 이에 A사 도로 네트워크의 링크와 상세도로망 Level6 네트워크의 링크와의 맵매칭 및 경로가공 모듈을 이용하여 경로데이터를 가공함

## 나. 링크 기반 데이터 전처리 및 경로 데이터 가공 알고리즘 개발

- A사 내비게이션 데이터는 링크 기반 데이터로 기존에 개발된 포인트 기반 데이터 맵 매칭 및 경로가공 모듈을 사용하기에 데이터 형식 및 특징이 다르므로 링크 기반 데이터를 궤적 정보로 생성하여 포인트 기반 데이터 맵매칭 및 경로 가공 모듈을 이용함
- A사 내비게이션 데이터는 B사 내비게이션 데이터와 차이가 있어 오류 유형에 따른 보정 작업을 수행하였으며, 보정된 링크 기반 데이터를 궤적 정보로 만드는 가공 모듈을 개발함
- A사 내비게이션 원시데이터 전처리 및 경로 데이터 가공 프로세스는 다음과 같음



<그림 9-3> 링크 기반 데이터 전처리 및 경로 데이터 가공 프로세스

### 1) 원시데이터(R0) 오류 유형에 따른 보정(R1)

- 유형 1 : DATE 필드 오류
- 유형 2 : 주행 링크 순서 정렬 오류
- 유형 3 : 경로 오류
- 유형 4 : 시간 오류
- 유형 1~4까지의 원시데이터 오류를 보정하여 오류 제거된 데이터(R1) 구축

### 2) 보정데이터(R1)를 궤적 정보데이터(S0)로 생성

- 데이터의 진출입 시간은 캡이 씌워진 속도(제한된 가공속도)로 계산되어있기 때문에



실제 진출입 시간을 구하기 위한 보정이 필요함

### 3) 궤적(S0)데이터의 경로 데이터(P1) 가공

- 차량 GPS 데이터(S0)를 상세도로망 Level6 네트워크 기반으로 경로를 가공함
- Level6 네트워크는 전국 2차선 이상의 도로망으로 노드와 링크로 구성되며, 전처리한 차량 GPS 데이터를 네트워크의 링크와 맵매칭하여 경로를 가공함

### 다. 포인트 기반 데이터 전처리 및 경로 데이터 가공

- DTG데이터는 포인트 기반 데이터로 포인트 기반 데이터 맵매칭 및 경로 가공 모듈을 이용함(해당 모듈은 R&D사업을 통해 개발된 것으로 간단하게 내용을 수록함)
- DTG 원시데이터 전처리 및 경로 데이터 가공 프로세스는 다음과 같음



<그림 9-4> 포인트 기반 데이터 전처리 및 경로 데이터 가공 프로세스

### 라. 경로 데이터 압축 및 통합 경로데이터 구축

#### 1) 차량ID별 링크 구간 통행속도 산출

- 링크와 매칭된 차량 GPS 궤적 정보를 이용하여 활용목적에 따라 다양한 데이터를 구축하기 위해 개별 차량의 링크별 통행속도를 산출함

- 기반 데이터에 따른 각 경로데이터를 도로네트워크와 맵매칭 후 통합 경로 데이터 구축
- 링크 기반 경로데이터는 Level6 네트워크 링크 길이에 따른 속도와 실제 차량 궤적에 따른 속도정보도 포함하여 통행속도를 산출함
- 포인트 기반 경로데이터는 링크와 매칭된 차량 GPS 궤적 정보의 도로구간 길이와 시간정보를 이용하여 통행속도를 산출함

## 2) 일별 차량ID별 경로 데이터 구축

- 링크와 매칭된 경로 데이터를 통행지표 생성을 위한 기초 DB로 활용됨
  - 통행정보, 통행속도, 공간정보가 결합된 경로 데이터 생성하여 표준 포맷으로 경로 DB를 구축함
  - OBU ID별로 1일 단위의 개별 차량의 경로 데이터 구축

<표 9-8> 개별차량 경로데이터 테이블 구성

컬럼명	데이터유형	설명	코드	코드정보
OBUID	Integer	단말기 ID	-	-
Vehtype	Integer	차종 유형	A B AUTO BUS TRUCK	A사 B사 DTG 택시 DTG 버스 DTG 트럭
GroupNum	Integer	경로그룹ID	-	-
Seq	Integer	순서	-	-
Date	DateTime	수집일시	-	-
Vlink	Integer	Level6 가상링크ID	-	-
InTime	Integer	진입시간	-	-
OutTime	Integer	진출시간	-	-
Speed	Double	통행속도	-	
Type	Integer	이상속도유형	1 4	링크속도 0이상 링크속도 미만

### 3) 경로 데이터 압축

- 경로 데이터 압축은 View-T 웹서비스를 위한 목적으로 수행하였음
- 상세도로망 Level6기반 경로 데이터는 대용량 데이터의 검색 속도의 한계로 서비스가 어렵기 때문에 경로 데이터를 주요도로망 Level5.5 단위로 압축하여 온라인 서비스를 수행함
- 주요도로망 Level5.5 기반 경로 데이터 압축 및 통합 경로 데이터 구축 시 610GB에서 150GB로 데이터 용량을 압축함

**<6레벨 기준 경로 데이터>**

필드명	내용	비고
TnzOBUEID	단말기 ID	
TnPathGroup	단말기 별 GroupID	
tnPathSeq	경로 순서	
TnxPathLink	6lev 가상링크ID	
In_Time	진입시간	DATETIME
Out_Time	진출시간	DATETIME
Speed	구간속도	1 ~ 150 km/h
type	보정코드	
emdlID	행정구역 (읍면동)	

**<경로 데이터 압축 및 통합경로 데이터 구축>**

필드	내용	Type	PK
_id	경로 ID (일련번호로 구성)	NumberLong	*
PL	경로 별 분석별 Link 구성 (1001,1002,1003,1004,1005)	Array(NumberInt)	
PT	경로 별 차종 정보 (Auto : 1, Bus : 2, Truck : 3) * 내비게이션은 승용차 '1'	NumberInt	
PM	경로 별 월 정보 (1 ~ 12)	NumberInt	
PD	경로 별 일 정보 (1 ~ 31)	NumberInt	
PW	경로 별 평일/주말 정보 (4개) (주말:0, 평일:1, 특송:10, 공유일:11)	NumberInt	
DW	경로 별 요일 정보 (9개) (월:1, 화:2, 수:3, 목:4, 금:5, 토:6, 일:0, 특송:10, 공유일:11)	NumberInt	
IT	경로 별 진입 시간(초단위로 구성, (시*3600+분*60+초)) 구성 (29460, 29470, 29475, 29485, 29498) (시간정보가 없는 경우 -1 처리함)	Array(NumberInt)	
OT	경로 별 진출 시간(초단위로 구성, (시*3600+분*60+초)) 구성 (29463, 29474, 29483, 29495, 29504) (시간정보가 없는 경우 -1 처리함)	Array(NumberInt)	
DT	행정구역 [3902011, 3902011, 3902011, 3902012, 3902012]	Array(NumberInt)	
DT_SGG	행정구역[3901000, ...]	Array(NumberInt)	
DT_SD	행정구역[3900000, ...]	Array(NumberInt)	

**<그림 9-5> 압축된 경로 데이터 형태 예시**

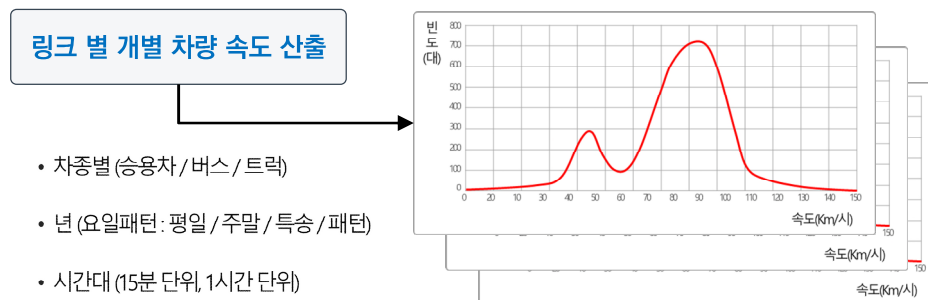
#### 마. 기초교통DB 구축 프로세스 및 데이터 정의

- 전국 교통량 전수화 및 통행지표를 구축하기 위하여, 경로데이터와 관측교통량을 이용하여 기초교통DB 구축
  - 기초교통DB를 구축하기 위해 개별 차량 경로데이터의 속도정보를 이용하여 구축한 속도프로파일과 개별차량 경로데이터를 링크별로 집계한 통행량, 연도별 이상치를 제거한 관측교통량이 필요함
- 속도프로파일 DB 구축 프로세스는 월단위 구축 → 년단위 병합(요일의 패턴별) 순으로 구축. 구축 시 내비게이션 데이터는 승용차로 구분하며, DTG는 버스/화물/택시 데이터로 구분하여 구축
- 관측교통량은 링크별 통행량과 비교하여 이상치가 있거나 연도별 패턴이 다른 관측교통량은 제거하여 구축

- 개별 차량 경로데이터를 이용하여 링크별 통행량을 산출하며, 관측교통량과 링크별 통행량으로 전수화하여 교통량 DB를 구축

#### 바. 속도프로파일 분석 및 DB 구축

- 속도프로파일 DB 구축
  - 교통량, 평균속도 등 통행지표를 구축하기 위한 속도프로파일 DB 구축
  - 링크의 매칭된 궤적 정보를 활용해 개별 차량의 속도 산출 및 산출된 속도 범위에 따른 집계를 수행
  - 링크별 개별차량 속도 프로파일 구축
  - 구축된 속도프로파일은 분포의 특성에 따라 이상치를 제거
  - 시간대가 없는 속도프로파일은 새벽, 비첨두 시간, 첨두 시간을 집계한 결과를 사용
  - 시간대/월/연 단위 대표 통행속도 산출
  - 연평균 평일/주말, 시간대별/전일 속도 프로파일 DB 구축



<그림 9-6> 속도프로파일 구축

<표 9-9> Level6 링크 단위 속도프로파일 DB 테이블 구성

컬럼명	데이터 유형	설명
VlinkID	Integer	가상링크ID
Week_type	Integer	요일코드
Time	Integer	0시 ~ 23시
Speed	Integer	속도 코드 : 0km ~ 150km
Probe_count	Integer	통행량

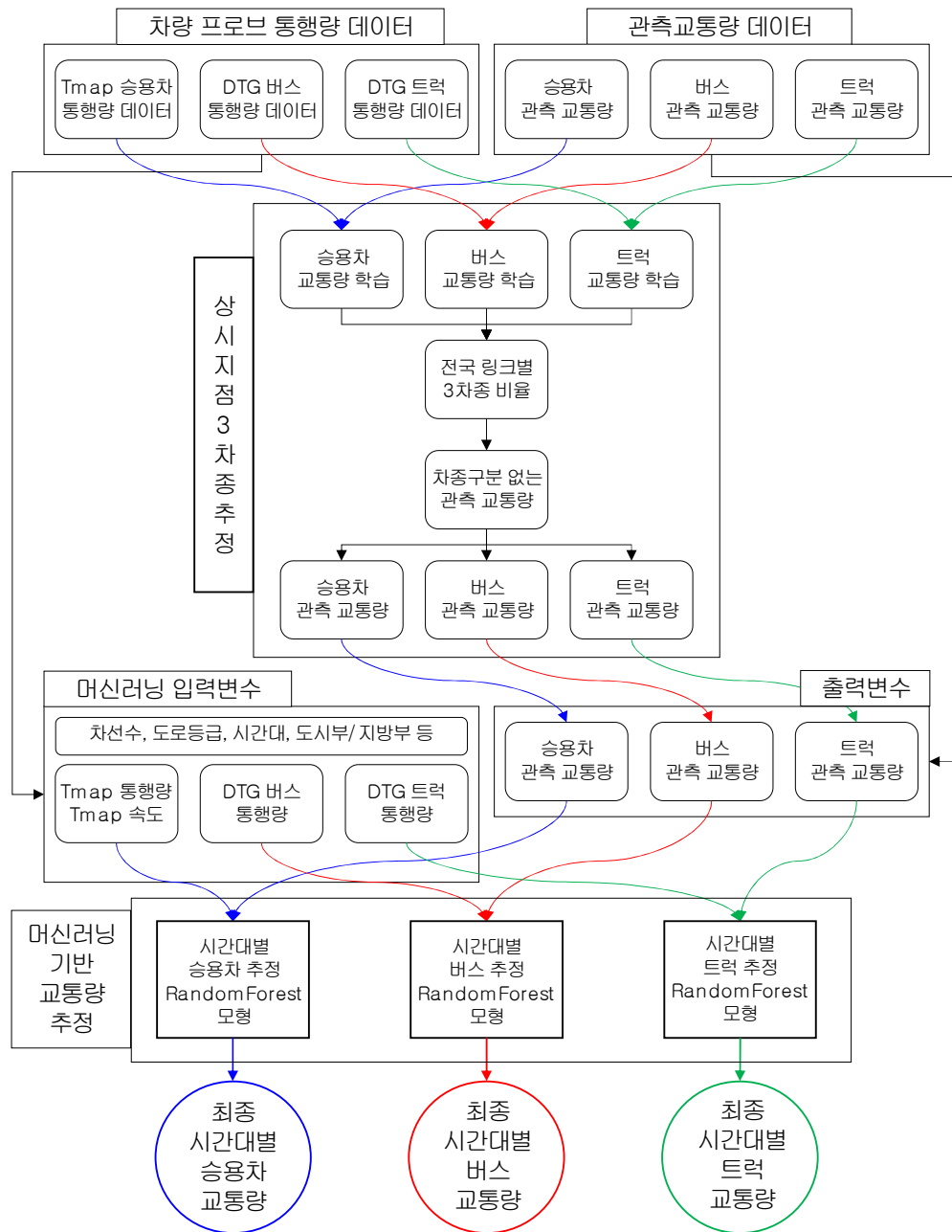
## 사. 관측교통량 데이터 분석 및 DB구축

- 관측교통량은 교통량 전수화를 위한 입력변수 데이터로 이용되며, 추정된 교통량의 검증 데이터로 이용됨
- 지역간 도로의 관측교통량은 국토교통 통계연보에 수록되는 한국건설기술연구원과 한국도로공사의 교통량 데이터임
- 도시부 도로의 관측교통량은 각 지자체에서 조사한 교통량으로 24시간대 관측된 교통량이 부족한 데이터임

## 아. 전국 교통량 전수화를 위한 교통량 추정 알고리즘 개선

- 교통량 추정 알고리즘 개선 목적
  - 전국에서 조사된 교통량은 약 1만 개 지점으로 상세도로망 Level6네트워크의 링크 갯수(약 62만 개)에 비해 매우 낮은 샘플율을 가지고 있으며, 교통량 추정을 위해 사용된 차량 GPS 데이터의 샘플율도 점점 감소됨 → 데이터 샘플율에 영향을 받지 않는 일반화된 알고리즘 개발이 필요함
  - 각 지자체에서 조사한 관측교통량은 차종이 구분되어 있지 않거나, 24시간 전체 시간대에 조사하지 않는 경우가 많아 차종별 시간대별 교통량 추정 시 신뢰도를 높일 방법을 개발함
- 경로 데이터 기반 통행배정을 활용한 관측교통량의 차종별 교통량 추정
  - 경로 데이터가 통과하는 관측교통량을 기준으로 경로 통행량을 전수화 시켜 경로 통행 기반 전국 네트워크에 통행 배정
  - 관측교통량의 차종 추정비율을 산정하며, 해당 결과를 머신러닝 앙상블 기법을 활용한 교통량 추정 알고리즘의 Input데이터로 사용
  - 머신러닝 앙상블 기법을 활용한 교통량 추정 결과의 검증데이터로 이용
- 랜덤 포레스트를 활용한 교통량 추정 알고리즘 개발
  - 최근 회귀 및 분류 문제에서 좋은 성능을 보이고 있는 머신러닝 앙상블 모델 중 랜덤 포레스트 모델을 적용하여 연도별 일관성을 유지하는 일반화된 교통량 추정 알고리즘 개발
  - 도로 속성 정보, 통행량(내비, DTG), 관측교통량, 방법1의 결과데이터, 속도 등 데이터

## 기반의 전국 차종별 시간대별 교통량 추정



&lt;그림 9-7&gt; 추정교통량 구축 프로세스

#### 자. 전국 교통량 '19년, '20년, '21년 추정결과 비교

- 교통량의 신뢰도를 평가하기 위해서는 조사된 교통량 자료가 존재하는 지점에서의 교통량 추정을 통해 현실을 어느 정도로 구현하고 있는지에 대한 검토가 필요함
- 한국개발연구원(2008)에 따르면 추정교통량의 신뢰성을 평가하기 위해 조사교통량과의 오차율 기준으로 추정교통량에 대한 평가를 수행하고 있다. 본 연구에서 분석하고 있는 조사지점의 교통량 수준을 고려하면  $\pm 15\%$  이내의 오차율이 유의한 추정 범위라고 볼 수 있음
- 추정교통량의 신뢰도 평가지표는 ①평균절대비율오차(Mean Absolute Percent Error, MAPE, %)와 ②평균비율오차(Mean Percent Error, MPE, %), ③상관계수(coefficient of correlation,  $R$ ), ④결정계수(coefficient of determination,  $R^2$ )를 활용함

$$MAPE(\%) = \frac{\sum_{i=1}^n \left( \frac{|q_i - \hat{q}_i|}{q_i} \right) \times 100}{n}, \quad MPE(\%) = \frac{\sum_{i=1}^n \left( \frac{q_i - \hat{q}_i}{q_i} \right) \times 100}{n}$$

$$R = \frac{\sum_{i=1}^n q_i \hat{q}_i - \frac{\sum_{i=1}^n q_i \sum_{i=1}^n \hat{q}_i}{n}}{\sqrt{\left[ \sum_{i=1}^n q_i^2 - \frac{(\sum_{i=1}^n q_i)^2}{n} \right] \left[ \sum_{i=1}^n \hat{q}_i^2 - \frac{(\sum_{i=1}^n \hat{q}_i)^2}{n} \right]}}$$

여기서,  $q_i$  = 지점  $i$ 의 조사교통량(대/일)

$\hat{q}$  = 지점  $i$ 의 추정교통량(대/일)

$n$  = 전체 지점 개수(개)

- 국토교통부에서 수집하고 있는 도로교통량조사 자료(국토교통 통계연보)와 지자체에서 수집하고 있는 교통량 조사 자료 중 고속국도, 일반국도, 특별광역시도를 대상으로 신뢰도 분석을 수행하였으며, 선정된 평가대상 지점들은 24시간대가 모두 조사된 지점임
- 고속도로의 상시지점은 6.5%, 일반국도의 상시지점은 11.0%, 서울특별시 상시지점은 6.0% 수준으로 나타나 상시지점에 대해 유의한 결과가 나타났음을 분석하였으

며, 주요 도로등급별 교통량 신뢰도 분석을 수행했을 때 상시 지점과 비슷한 수준으로 나타났음을 확인할 수 있어 신뢰도 높은 교통량 추정이 되었다고 판단됨

<표 9-10> 2020년 상시 지점 추정교통량 신뢰도 분석

구분	신뢰도 평가대상 관측교통량 지점 수	MAPE(%)	MAE(%)	상관계수	결정계수
고속도로	236	6.5	-3.8	0.9980	0.9960
일반국도	1,076	11.0	-5.3	0.9926	0.9852
서울특별시	264	6.0	-4.0	0.9969	0.9939

<표 9-11> 2020년 주요 도로등급별 추정교통량 신뢰도 분석

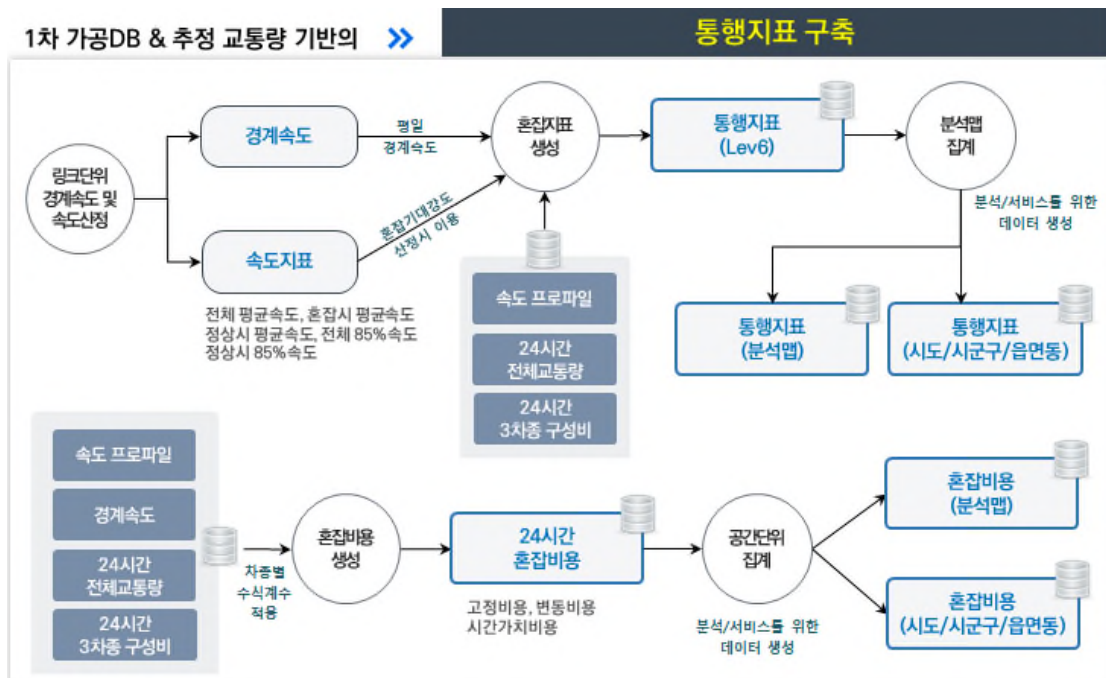
구분	신뢰도 평가대상 관측교통량 지점 수	MAPE(%)	MAE(%)	상관계수	결정계수
고속도로	1,347	7.3	-4.2	0.9904	0.9808
일반국도	3,430	11.6	-6.1	0.9924	0.9850
도시고속도로	60	5.1	-1.8	0.9987	0.9973
특별광역시도	904	10.3	-5.1	0.9767	0.9767



### 3. 차량 통행지표 산출 및 DB구축

#### 가. 통행지표 DB 구축 프로세스

- 통행지표 DB는 가공된 속도프로파일과 추정된 교통량을 기반으로 전국 혼잡지표, 속도지표, 환경지표 등을 구축
- 구축 프로세스는 1. 링크단위 경계속도 산정, 2. 속도프로파일과 교통량 기반의 Level6 단위 통행지표 구축, 3. Level6 단위 통행지표 기반으로 분석맵 단위 집계, 4. 분석맵 단위 기반으로 행정구역 단위 집계 순으로 구축



<그림 9-8> 통행지표 구축 프로세스

- 통행지표 구축을 위한 1차 가공DB 및 기초 교통DB는 다음과 같이 정의함

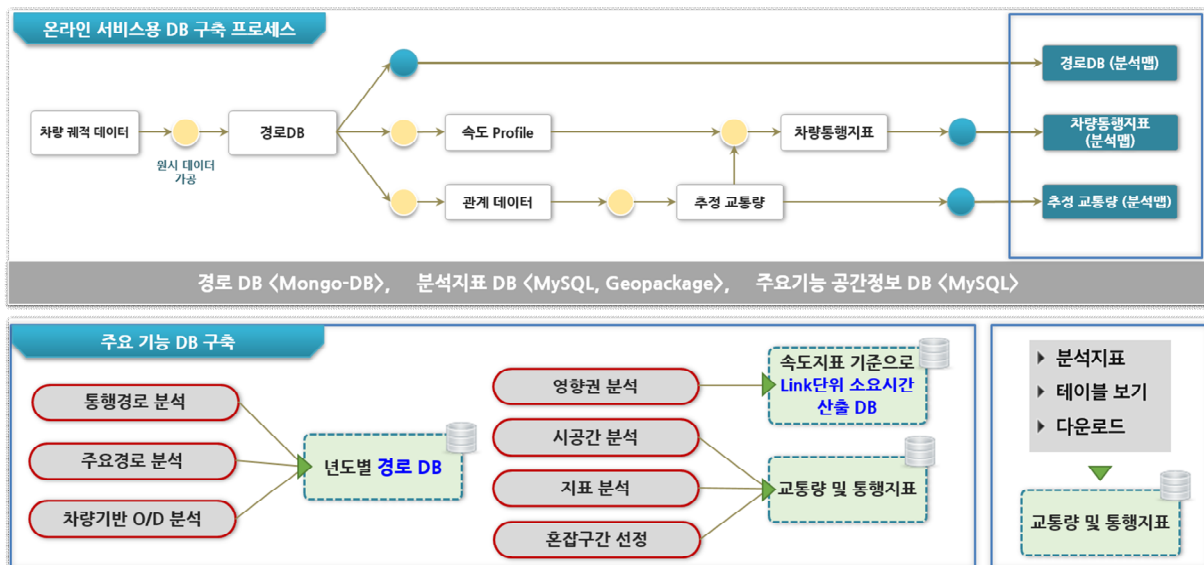
<표 9-12> 통행지표 DB 생성기준

기초교통 DB	공간정보	통행지표 DB 생성기준
<ul style="list-style-type: none"> <li>· 속도프로파일(년단위/요일별/차종별)</li> <li>· Level6 단위 추정 교통량 (년단위/요일패턴 별/시간/차종별)</li> </ul>	<ul style="list-style-type: none"> <li>· Level6 (상세 도로망)</li> <li>· Level5.5 (주요 도로망)</li> <li>· 표준노드링크</li> <li>· 행정구 (시도/시군구/읍면동)</li> </ul>	<ul style="list-style-type: none"> <li>· 공간정보 별 속도지표 및 사고지표 (년단위/요일패턴 별/시간/차종별)</li> <li>· 공간정보 별 혼잡지표 및 혼잡비용 (년단위/요일패턴 별/시간/차종별)</li> </ul>

&lt;표 9-13&gt; 추정교통량 데이터 테이블 구성(상세도로망 Level6 도로 단위)

컬럼명	Type	설명	코드	코드정보
상세도로망_LinkID	Integer	상세도로망Level6네트워크의 LinkID	-	-
도로등급	char	도로의 등급	101 102 103 104 105 106 107 108	고속도로 도시고속도로 일반국도 특별광역시도 국가지원지방도 지방도 시군도 고속도로 연결램프
링크길이	Double	도로구간 길이	-	-
도로명	Varchar2	도로명	-	-
시도코드	Integer	시도명을 나타내는 코드	-	-
시군구코드	Integer	시군구명을 나타내는 코드	-	-
읍면동코드	Integer	읍면동명을 나타내는 코드	-	-
시도명	string	시도의 이름	-	-
시군구명	string	시군구의 이름	-	-
읍면동명	string	읍면동의 이름	-	-
주_유형	Varchar2	평일과 주말 구분	weekday weekend	평일 주말
시간적범위	Varchar2	전일과 시간대 구분	0-23 all	0 : 24-1시 .. 23 : 23-24시 all : 연평균 1일
전체_추정교통량	Integer	모든 차량의 추정교통량	-	-
승용차_추정교통량	Integer	승용차의 추정교통량	-	-
버스_추정교통량	Integer	버스의 추정교통량		
화물차_추정교통량	Integer	화물차의 추정교통량		

## 나. 차량 온라인 서비스 DB 구축



<그림 9-9> 차량 온라인 서비스 제공을 위한 DB 구축 프로세스

- 온라인 서비스를 위한 교통지표는 행정구역별/도로별로 구분하며, 표/그래프/지도표출/다운로드 등으로 서비스
- 온라인 서비스를 위한 교통지표는 교통량, 교통혼잡지표, 교통환경지표, 사고지표로 구성
- 각 분류별 지표는 교통량지표 3종, 교통혼잡지표 6종, 교통환경지표 5종, 사고지표 2종으로 구성되며, 상세 지표별 내용은 다음 표와 같음

&lt;표 9-14&gt; 온라인 서비스 지표

교통지표		지표 설명
교통량	관측 교통량	각 기관에서 조사한 차종별 조사 교통량
	구간 추정 교통량	특정 시간대 동안 해당 도로구간을 통과하는 차량 대수
	차량 주행거리	추정 교통량 기준의 차량주행거리
교통혼잡지표	혼잡강도	전체 차량의 총 통행시간 중 교통 혼잡을 경험한 차량들의 총 통행시간 비율
	평균속도	전체 차량의 속도를 평균한 값
	혼잡시 평균속도	전체 차량 중 교통 혼잡을 경험한 차량들의 평균속도
	정상시 평균속도	전체 차량 중 교통 혼잡을 경험하지 않은 차량들의 평균속도
	지체시간	교통 혼잡으로 인하여 발생하는 차량 1대 당 평균 지체시간
	교통혼잡비용	교통 혼잡에 따른 차량 통행시간 증가로 인하여 발생하는 추가적인 사회적 손실비용(고정비, 변동비, 시간가치비용)
교통환경지표	이산화탄소배출량	특정 시간대 동안 해당 도로구간을 통과하는 차량들로 인하여 발생하는 각 교통 환경지표의 평균 배출량
	미세먼지배출량	
	일산화탄소배출량	
	휘발성 유기화합물 배출량	
	질소산화물 배출량	

○ 차량통행 분석도구 DB 구축

- 차량통행 분석도구 제공을 위한 기반 DB 구축

- 통행경로 분석/차량기반 OD분석/영향권 분석/시공간 분석/주요경로 분석/혼잡구간 분석을 위한 기반 DB 구축

- 서비스 속도 향상을 위한 경로 데이터 압축 및 NoSQL 솔루션 적용을 통해 검색 속도 향상



- 2021년도 월별 퍼센타일 주행 속도의 선택된 조건을 기반으로 결과값을 테이블로 표현
  - 도로의 15%, 25%, 30%, 50%, 75%, 85% 주행속도 분석 데이터 제공
- 2021년도 월별 퍼센타일 주행 속도의 분석 조건을 변경할 수 있는 UI 제공
  - 퍼센타일 주행속도의 연월, 기간, 주요시간대, 권역별, 도로등급 설정할 수 있도록 함

<그림 9-11> Faster Indicator 분석 조건 화면

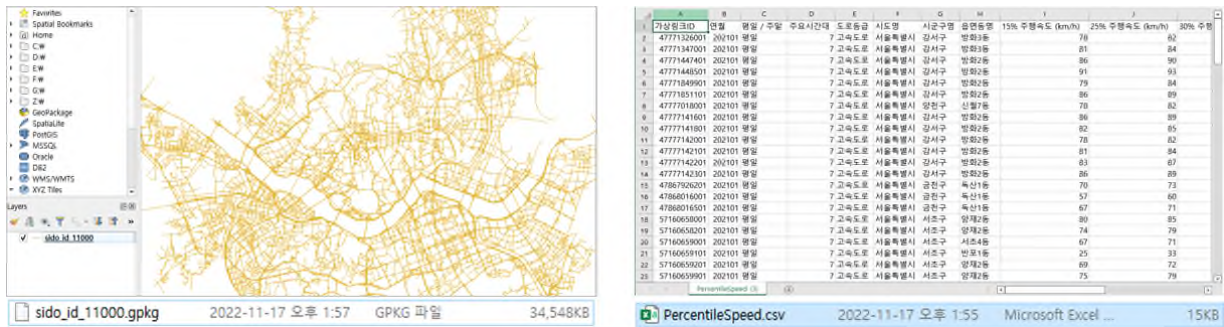
- 사용자가 변경한 분석조건을 적용할 수 있도록 상세 조회 버튼으로 안내함
  - 변경된 분석조건 적용 완료 시 데이터, 네트워크 다운로드 버튼을 활성화 하여 CSV, Geopackage로 내보내기

<그림 9-12> Faster Indicator 조회 및 다운로드 버튼

- 데이터, 네트워크 다운로드 클릭 시 분석도구 사용빈도 및 활용 목적 등을 관리자가 로그를 확인할 수 있도록 신청서 작성할 수 있도록 함

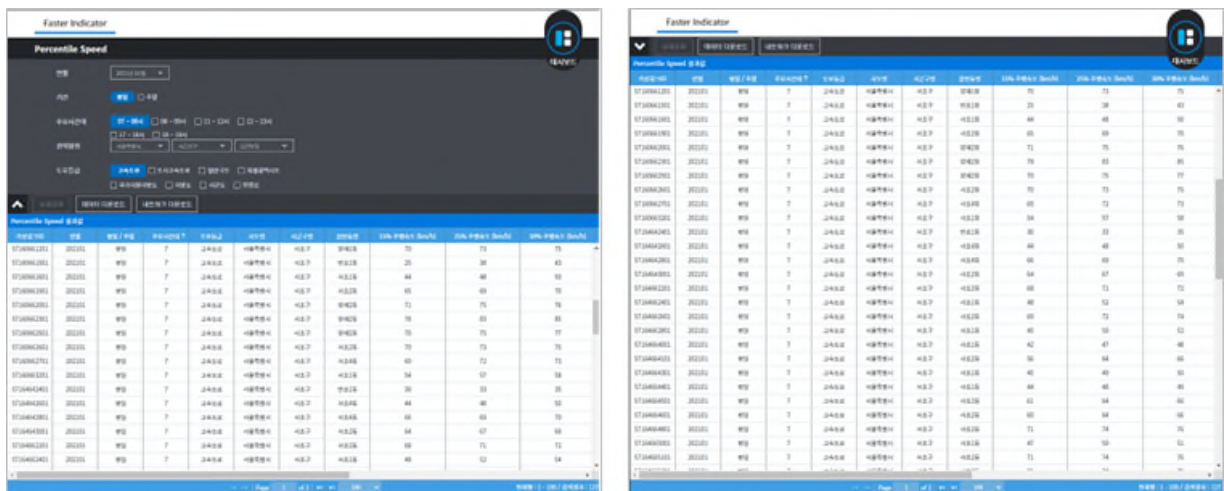
<그림 9-13> Faster Indicator 신청서 작성 화면

- 데이터, 네트워크 다운로드 완료시 형상정보(\*.gpkg), 데이터(\*.csv)로 제공 함
  - 분석조건에 선택된 권역의 시도/시군구 단위의 형상정보(\*.gpkg) 데이터 제공
  - 분석 결과에 대한 \*.csv 파일을 Excel에서도 확인할 수 있도록 제공



<그림 9-14> Faster Indicator 분석 다운로드 결과 화면

- 설정한 분석조건을 기반으로 테이블 형태로 분석 결과값을 표출
  - ▲ 버튼을 이용한 분석결과 테이블 최소 최대화 적용
  - 테이블 컬럼을 클릭하여 해당 값의 오름, 내림차순으로 변경 가능하도록 함



<그림 9-15> Faster Indicator 서비스 결과 화면

- Faster Indicator로 퍼센타일 속도 데이터 테이블 명세서는 다음과 같음

&lt;표 9-15&gt; Faster Indicator 데이터 테이블 정의

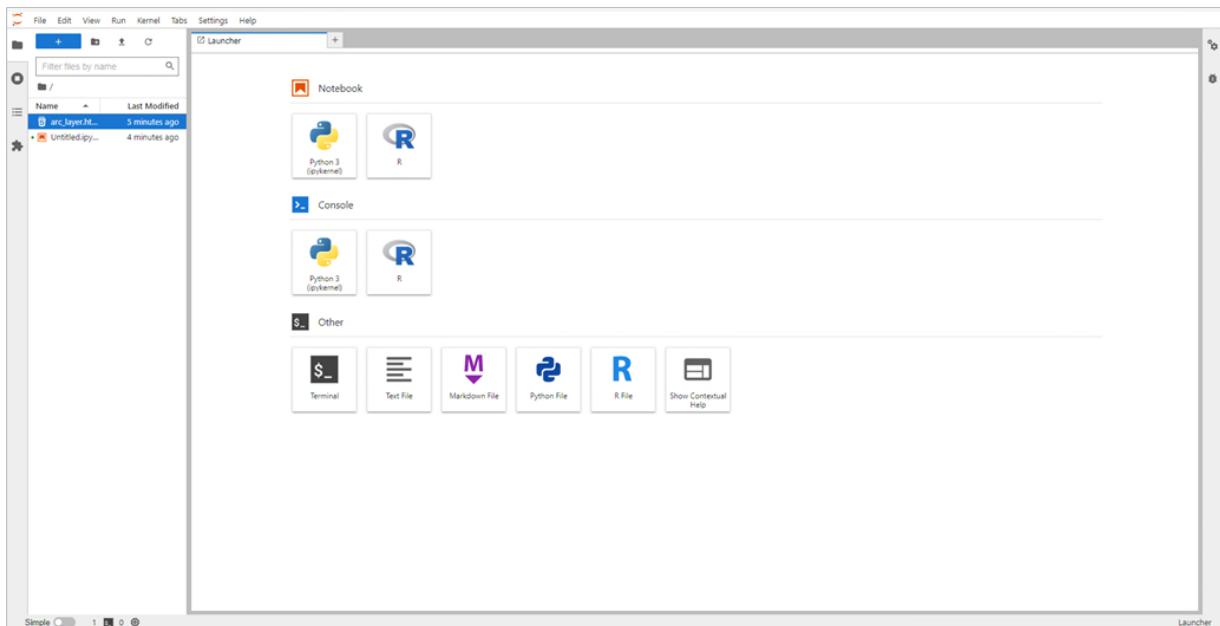
No	Column	설명	Type	자리수	코드정보
1	v_link_id	가상링크 ID	bigint (20)	0	
2	date	날짜	int (6)	0	연월 (000000)
3	week_code	요일코드	int (1)	0	0 : 주말 1 : 평일
4	peak_time	주요시간대	int (2)	0	0 : 오전첨두 (7~9시) 1 : 낮 (11~13시) 2 : 오후첨두 (17~19시)
5	road_rank	도로등급	int (3)		
6	sido_id	시도ID	int (7)		
7	sigungu_id	시군구ID	int (7)		
8	emd_id	읍면동ID	int (7)		
9	speed_15th	15퍼센타일 속도	double		15 백분위 주행속도
10	speed_25th	25퍼센타일 속도	double		25 백분위 주행속도
11	speed_30th	30퍼센타일 속도	double		35 백분위 주행속도
12	speed_50th	50퍼센타일 속도	double		50 백분위 주행속도
13	speed_75th	75퍼센타일 속도	double		75 백분위 주행속도
14	speed_85th	85퍼센타일 속도	double		85 백분위 주행속도
15	speed_avg	평균 속도	double		링크 통행속도 평균 , km/h
16	speed_sd	속도 표준편차	double		링크 통행속도 표준편차, km/h
17	speed_max	최대 속도	double		링크 최대 통행속도, km/h



## 2. 오픈소스 활용 서비스 개발

### 가. 기능 소개

- 이용자가 별도의 분석/시각화 소프트웨어 설치가 없어도 View-T에서 데이터를 다운로드 받아 즉각적으로 분석/시각화 할 수 있는 기반 분석환경을 제공하기 위해 오픈소스 기반 분석환경 개발
- 오픈소스 소프트웨어(JupyterHub, JupyterLab)를 이용하여 데이터 시각화 및 간단한 분석을 할 수 있는 기능
- 다양한 언어의 프로그램을 별도의 설치 없이 오픈소스 소프트웨어를 이용하여 활용할 수 있도록 제공함
- R, Python 등 데이터 시각화 툴을 사용할 수 있도록 제공
- 관리자가 등록한 사용자의 개별적으로 작업할 수 있는 공간 제공



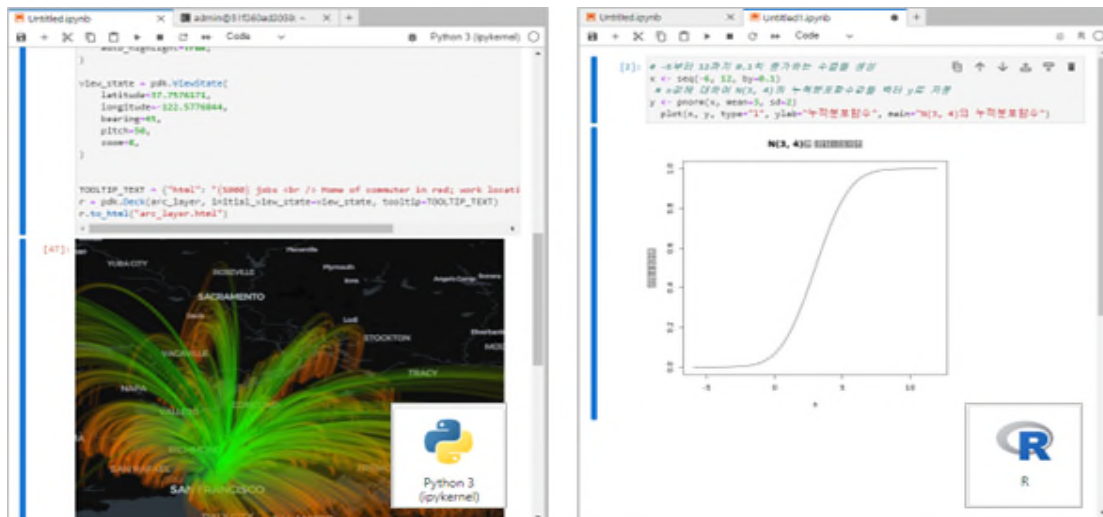
<그림 9-16> 오픈소스 활용 서비스 오픈소스 소프트웨어 화면

### 나. 기능 설명

- 등록된 커널( R, Python )을 이용하여 코드를 직접 작성 및 실행할 수 있도록 함
- python의 에디트 모드, 커맨드 모드, 편집 모드, 셀 경계 인터페이스를 제공하여 사용


자에게 쉽게 코드를 작성할 수 있도록 함

- 사용자 편의를 위한 단축키를 제공함
- Python, R의 코드를 입력 및 실행하여 코드의 결과물을 웹 화면에서 확인 가능 하도록 함



<그림 9-17> Python, R 프로그램 언어 활용 화면

- 사용자 관리 페이지에서 사용자를 등록, 삭제, 제한 등 권한에 대해서 정의할 수 있도록 함
- 사용자의 서버를 관리자가 관리할 수 있도록 재시작, 종료, 시작을 제공함
- 회원 가입한 사용자의 승인 및 거절을 선택할 수 있도록 함


jupyterhub

[집](#)
[토론](#)
[관리자](#)

[관리자](#)
[로그 아웃](#)

사용자 검색

> 그룹 관리

사용자	관리자	성기는 사람	마지막 활동	달리기	행위
<div>사용자 추가</div>			<div>모두 시작</div> <div>모두 중지</div>	<div>종료 허브</div>	
<div>관리자</div>	관리자	[기본]	59초 전	<div>서버 중지</div> <div>액세스 서버</div>	<div>사용자 편집</div>
<div>예드</div>		[기본]	1개월 전	<div>서버 중지</div> <div>액세스 서버</div>	<div>사용자 편집</div>
<div>큐빅</div>		[기본]	절대	<div>서버 시작</div> <div>스톱 페이지</div>	<div>사용자 편집</div>
<div>테스트1</div>		[기본]	절대	<div>서버 시작</div> <div>스톱 페이지</div>	<div>사용자 편집</div>

0 - 4 표시

이전의

다음

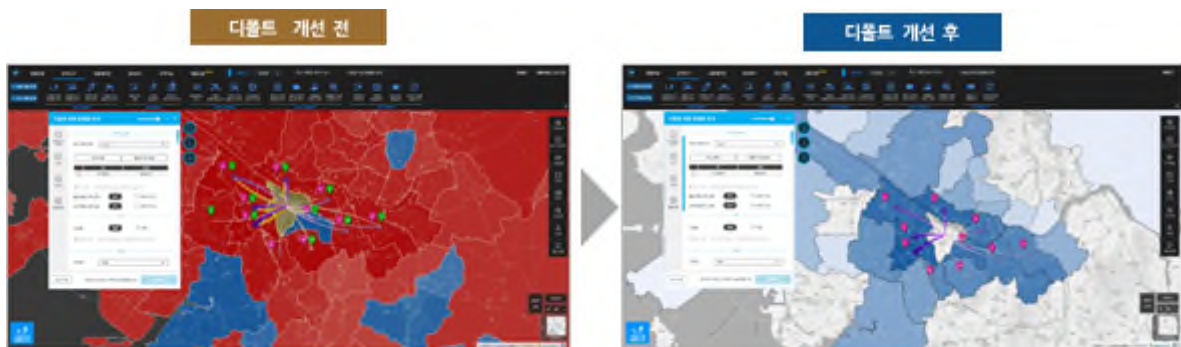
주피터허브 2.3.1 20221017011504

<그림 9-18> 오픈소스 소프트웨어 사용자 관리 화면

### 3. 분석도구 현행화

#### 가. 기종점 차량 통행량 분석도구 현행화

- 출발지에서 각 행정구역으로 가는 차량 통행량과 각 행정구역에서 도착지로 오는 차량 통행량을 분석하는 기능
- 선택한 행정구역(시도, 시군구, 읍면동)단위를 출발지나 도착지로 설정하여 출발지(도착지)에서 도착지(출발지)로 이동하는 차량의 통행량을 분석함
  - 분석지역 단위: 시도, 시군구, 읍면동
  - 선택 진입 행정구역 : 전국, 사용자 정의
  - 선택 진출 행정구역 : 전국, 사용자 정의
  - 기간 및 시간대 : 평일/주말, 전일, 24시간
  - 표출 형태 : 시군구, 읍면동
- 초기화면 주제도 표출 기본값 설정 변경
  - 행정구역 출 도착 변경
  - 행정구역 테두리 색상 변경
- 2020년 데이터 현행화

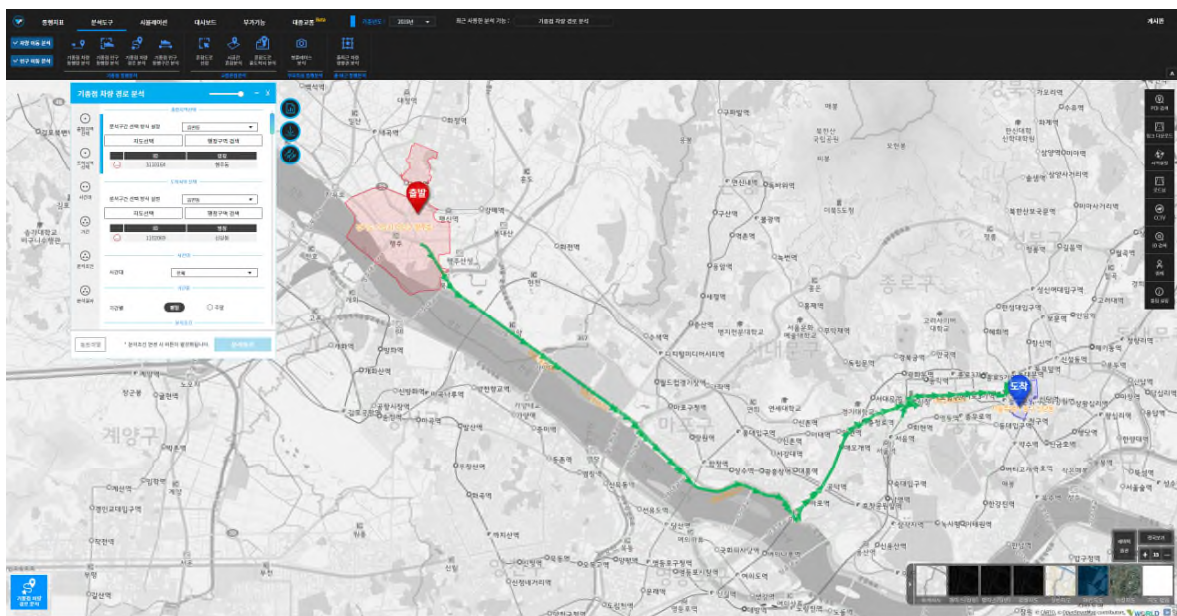


<그림 9-19> 기종점 차량 통행량 분석 디폴트 변경 화면

#### 나. 기종점 차량 경로 분석도구 현행화

- 출발지에서 도착지로 이동하는 차량의 주요경로를 분석하는 기능

- 선택한 행정구역(시도, 시군구, 읍면동)단위를 출발지나 도착지로 설정하여 출발지(도착지)에서 도착지(출발지)로 이동하는 차량의 통행량을 분석함
  - 분석지역 단위: 시도, 시군구, 읍면동
  - 선택 진입 행정구역 : 전국, 사용자 정의
  - 선택 진출 행정구역 : 전국, 사용자 정의
  - 기간 및 시간대 : 평일/주말, 전일, 24시간
  - 표출 형태 : 시군구, 읍면동
- 2020년 데이터 현행화 및 유지보수



<그림 9-20> 기종점 차량 경로 분석도구 화면예시

#### 다. 혼잡도로 선정 분석도구 현행화

- 이용자가 정의하는 혼잡의 기준에 따른 혼잡도로 선정 분석하는 기능
- 선택한 출발지와 도착지로 이동하는 차량 비율로 경로를 순위별, 누적별로 분석함
  - 선택 출발지 : 도로, 행정구역(시군구, 읍면동), 사용자 정의
  - 선택 도착지 : 도로, 행정구역(시군구, 읍면동), 사용자 정의
  - 기간 및 시간대 : 평일/주말, 전일, 24시간
  - 표출 형태 : 순위별 경로, 누적별 경로

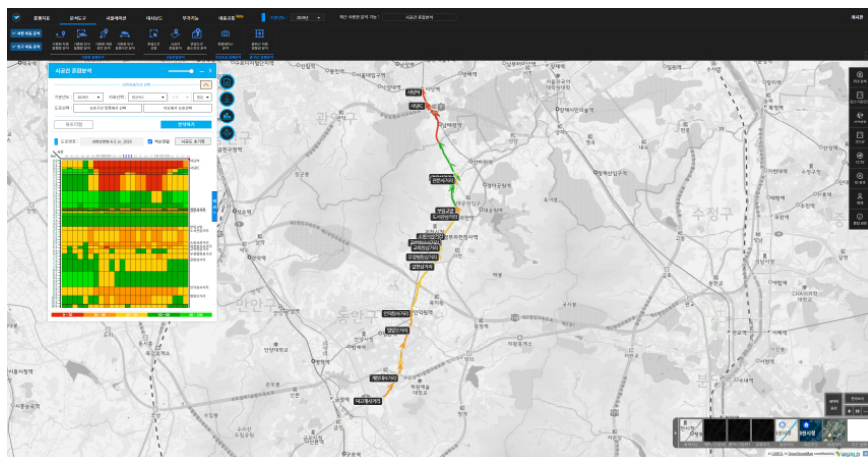
- 초기화면 주제도 표출 기본값 설정 변경
  - 도로 테두리 색상 변경
- 2020년 데이터 현행화 및 유지보수



<그림 9-21> 혼잡도로 선정 분석 디폴트 변경 화면

#### 라. 시·공간 혼잡 분석도구 현행화

- 도로축의 시간과 공간 변화에 따른 혼잡현황을 분석하는 기능
- 도로의 방향, 통행지표(추정교통량, 평균속도, 혼잡시평균속도, 정상시평균속도)를 선택하여 원하는 도로(축)의 시간변화에 따른 혼잡구간을 분석함
  - 기간 : 평일/주말
  - 선택 조건 : 통행지표, 도로구간(축), 도로 방향
- 2020년 데이터 현행화 및 유지보수

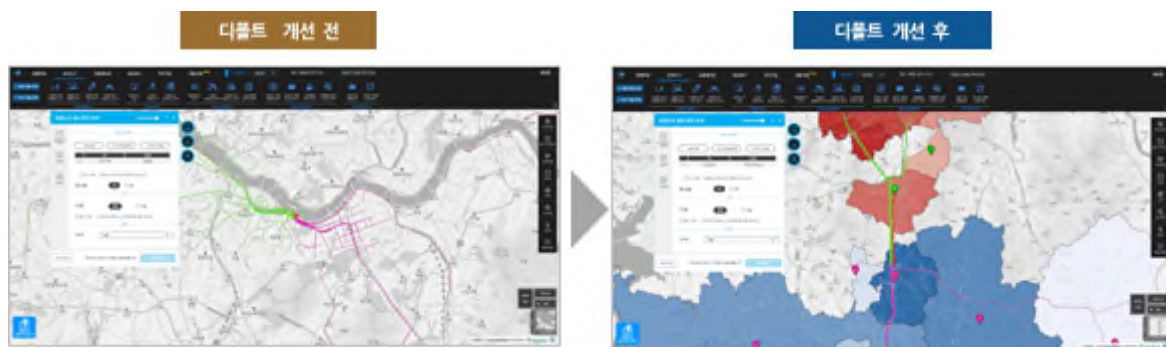


<그림 9-22> 시·공간 혼잡 분석도구 화면예시



#### 마. 혼잡도로 통행경로 분석도구 현행화

- 선택한 도로를 통행하는 차량의 공간적 통행패턴을 분석하는 기능
- 선택한 도로를 통행하는 차량의 출발지와 도착지를 행정구역별(시군구, 읍면동) 또는 도로 단위로 통행량을 분석함
  - 기간 및 시간대 : 평일/주말, 전일, 24시간
  - 선택 공간 조건 : 유턴 on, 유턴 off
- 초기화면 주제도 표출 기본값 설정 변경
  - 행정구역 테두리 색상 변경
  - 행정구역 On 변경
- 2020년 데이터 현행화 및 유지보수

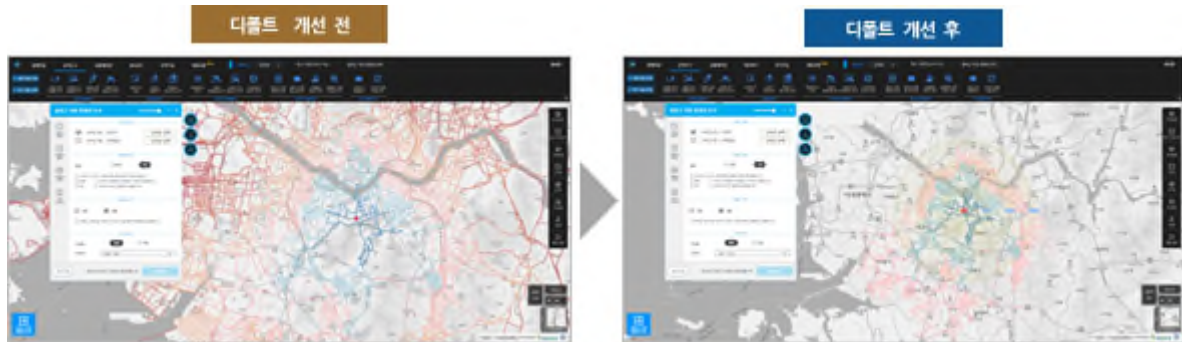


<그림 9-23> 혼잡도로 통행경로 분석도구 디폴트 변경 화면

#### 바. 출·퇴근 차량 영향권 분석도구 현행화

- 선택한 교차로에서 1시간 이내로 오고 가는 차량의 영향권을 10분 단위 간격으로 이동 시간을 분석하는 기능
  - 기간 및 시간대 : 평일/주말, 전일, 24시간
  - 선택 공간 조건 : 중첩되는 영향권, 모든 영향권
  - 선택 시간 조건 : 10분 단위 간격
- 초기화면 주제도 표출 기본값 설정 변경
  - 영향권 표출 범위 변경

- 반경 표출
- 2020년 데이터 현행화 및 유지보수



<그림 9-24> 출·퇴근 영향권 분석도구 디폴트 변경화면

## 제4절 결론 및 향후과제

### 1. 결론

- 본 과업은 교통현황을 모니터링 할 수 있는 차량 GPS 빅데이터 기반 시계열 전국 단위의 기초교통DB 및 통행지표를 산출 및 구축하여 정부 및 지자체 등 다양한 이용자에게 제공하는데 목적이 있음
- 본 과업을 통해 샘플율이 높은 A사의 내비게이션 데이터를 수집하였으며, 해당 데이터를 가공하기 위한 경로데이터 가공 알고리즘 개발, 교통량 전수화 알고리즘 개선을 수행함
- 본 과업을 통해 아래와 같은 기초교통DB 및 18종 통행지표를 구축을 수행함
  - 전국 양방향 2차로 이상 도로에 대한 교통량, 속도DB 구축(약 106만개 링크 대상)
  - 전국 양방향 2차로 이상 도로를 일관성 있게 평가할 수 있는 18종 통행지표 구축(약 106만개 링크 대상)
    - 교통량(3종) : 관측교통량, 추정교통량, 차량주행거리
    - 속도(3종) : 평균속도, 혼잡시평균속도, 정상시평균속도
    - 혼잡(4종) : 혼잡시간강도, 혼잡빈도강도, 혼잡기대강도, 교통혼잡비용
    - 환경(5종) : 이산화탄소 배출량, 미세먼지 배출량, 일산화탄소 배출량, 휘발성유기화합물 배출량, 질소산화물 배출량
    - 안전(3종) : 운전자 피로도, 과속비율, 속도편차(신규지표)
- View-T 온라인 서비스 이용자의 사용성을 높이기 위한 분석도구 고도화 및 이용자의 편의성 개선, 신규 기능 개발 등 편리성과 다양성을 개선하여 온라인 서비스 환경 개선함
  - Faster Indicator 서비스 개발, 오픈소스SW 활용 서비스 개발, 데이터 다운로드 현황 분석 관리자 페이지 개발, 데이터 다운받기 기능 개선 등
- 데이터 기반 실증적 과학적 교통정책 수립을 지원 및 지자체 지원을 통해 교통 빅데이터 기반의 신규 부가가치를 창출함
  - 대도시권 교통혼잡도로 개선사업, 제2차 국가도로망 종합계획수립, 국도시설개량사



업 효과 분석 등 중앙부처, 지자체, 공공기관의 정책기반자료 제공

- 지자체 교통현안 문제 해결을 위한 지자체 실증사업 지원
  - LH토지주택공사의 공모전, 국토연구원의 교통혼잡관련 공공사업, 한국환경공단  
의 교통환경관련 R&D사업 등 여러 기관에서 진행하는 사업을 지원함

## 2. 향후 과제

- 이용자 수요가 높은 신규지표 개발 및 투자 필요
  - 대기행렬 길이나 교차로의 회전 교통량, 고속도로 통행 차량의 일반도로 주요 통행  
경로 DB 구축 등 조금 더 알고리즘의 개발이 필요하거나, 활용도가 높은 가공 DB  
구축에도 향후 투자 필요
- 원시데이터 용량 증가에 따른 알고리즘 및 모듈 구현 시간 단축을 위한 방법론 개발 필요
  - 기존 경로데이터 가공 시 음영 구간에 대한 보정을 위해 경로 탐색을 수행하지만, A  
사의 내비게이션 원시 데이터를 경로데이터로 가공 시 시간이 오래 걸려 음영 구간  
에 대해 보정하지 않았으므로 추후 음영 구간에 대한 경로 탐색을 위한 시간 단축  
방법론 개발 필요 (1일치 데이터에 대해 약 13.8시간 소요되며, 음영구간은 약 2500  
만 건으로 집계됨)
- View-T 온라인 서비스를 국가교통DB시스템에 탑재 하기 위한 작업 필요
  - 현재 View-T 온라인 서비스는 프로토 타입으로 추후 KTDB 시스템의 데이터 개방  
환경 및 이용자 환경 기능으로 개선 및 확대하고자 함
- 기초교통DB 및 통행지표의 시계열성 확보를 위한 현행화 필요
  - 전국 양방향 2차로 이상 도로에 대한 18종 통행지표와 온라인 서비스 제공을 위한  
2021년 기준 데이터로의 현행화 필요



## 제1장 과업의 개요

---

제1절 과업의 배경 및 목적

제2절 과업의 범위 및 내용



## 제1장 과업의 개요

### 제1절 과업의 배경 및 목적

#### 1. 과업의 배경

- 전국 단위 기초 교통데이터 수집 부족
  - 교통분야의 기초 데이터인 교통량, 속도는 한국건설기술연구원, 각 지자체에서 조사하고 있으나, 교통량, 속도 자료 수집의 공간적 커버리지<sup>6)</sup>가 매우 낮아 전국의 기초 교통 현황을 모니터링 하기에 한계
- 지점 데이터 분석의 한계
  - 공공에서 수집하고 있는 데이터는 지점정보 중심의 데이터로 지점의 단편적 정보만 확인 할 수 있어 시공간적으로 연결된 교통의 흐름을 파악하기에 한계가 있음
- 빅데이터 전처리·가공 환경의 부재
  - 휘발성·단발성 빅데이터 관련 사업은 연속성 있는 빅데이터 전처리·가공환경 구축에 한계가 있으며, 대용량 데이터를 대용량 네트워크에서 효율적으로 전처리·가공하기 위한 환경 부족
- 공급자 중심의 데이터 제공환경
  - 중앙정부, 지자체, 연구기관, 학계 등 다양한 기관과 분야에서 교통 데이터를 요구하고 있으나, 공급자 기반의 데이터 제공환경에서는 대응하기 어려운 한계가 있음
- 전국 단위 일관된 교통DB 및 통행지표 부재
  - 데이터 수집기관, 수집방식 등 지역별로 상이하게 수집되는 데이터 수집 체계는 전국을 일관된 기준으로 평가하고 분석하기에 한계
- 차량이동 분석시스템 부재
  - 과거 차량의 이동의 시공간적 통행패턴과 현상을 분석하기 위해서는 모형 중심의 프로그램을 이용해야 하기 때문에 많은 인력·시간·예산이 소모되었고 분석결과의 현실성이 저하되는 문제 발생

<sup>6)</sup> 공공에서 수집하고 있는 교통량, 속도는 전국의 약 3%, 16%정도의 커버리지에 불과

- 차량 GPS 빅데이터 구축 및 활용 필요
  - 전국 도로에서 수집되는 차량 GPS 빅데이터를 활용하여 빅데이터 전처리·가공환경 구축, 전국단위 교통DB 구축, 차량이동 분석 시스템 구축을 통해 과거의 한계를 개선하고 국가교통DB의 신뢰성과 활용성을 높이기 위한 새로운 기반환경 구축 필요

## 2. 과업의 목적

- 차량 GPS 빅데이터 기반 시계열 전국 단위 데이터 전처리 및 가공
  - 매년 사업을 통해 구축된 교통DB를 통해 연도별 교통변화를 반영한 전국 단위 데이터 전처리 및 가공
- 국가교통데이터베이스 구축 및 교통 현황 모니터링을 위한 기초 교통DB, 통행지표 산출
  - 교통량, 속도 등 교통 분야의 핵심 기초교통 DB 구축
  - 교통혼잡비용, 혼잡강도, 이산화탄소 배출량 등 도로를 다각도로 분석할 수 있는 도로 교통망 성능평가 지표 구축
- 교통수요분석 DB구축 신뢰도 제고
  - 가구통행조사 기반의 인력조사에서 빅데이터로 전환하는 시점에서 교통수요분석DB의 구축 및 검증을 위한 기반 데이터 제공 및 분석지원
- 통행지표 및 데이터 제공 환경 개선
  - 이용자의 요구에 대응할 수 있는 온라인 기반 데이터 제공 환경 개선
  - 차량의 통행행태와 특성을 시공간적으로 분석할 수 있는 분석도구 개선
- 교통분야 주요 정책 및 지자체 지원
  - 예비타당성조사, 교통영향평가 등 주요 정책 지원
  - 지자체 교통현안 문제 해결을 위한 지자체 실증사업 지원

## 제2절 과업의 범위 및 내용

### 가. 공간적 범위

- 전국 2차로 이상 도로 및 주요 도로<sup>7)</sup>

### 나. 시간적 범위

- 2020년, 2021년 기준 데이터

### 다. 내용적 범위

- 차량 GPS 데이터 가공 및 DB 구축
  - 원시 데이터 전처리 및 경로데이터 가공
  - 차량 GPS 원시 데이터 검증 및 오류 필터링
  - 차량 GPS 데이터 확대 구축으로 데이터 전처리 가공 알고리즘 개선
  - 기초 교통DB가공 알고리즘 개선
  - 통행지표 구축을 위한 기초 교통 DB구축
  - 구축 데이터 및 DB 검증 및 분석
- 차량 GPS 데이터 개방 확대 및 환경 개선
  - 서비스 고도화 및 이용자 편의성 개선
  - 데이터 개방 확대를 위한 데이터 제공 관련 신규 기능 개발
  - 주요 정책 및 지자체 지원을 위한 데이터 제공
  - View-T 온라인 서비스 운영 및 유지보수

---

<sup>7)</sup> 주요 도로: 편도 1차로 도로 중 데이터 수집 안정성이 확보된 도로





## 제2장 차량 GPS 데이터 가공 및 DB 구축

---

제1절 원시데이터 수집 및 전처리

제2절 경로데이터 가공 및 기초교통DB 구축

제3절 통행지표 산출 및 DB구축



## 제2장 차량 GPS 데이터 가공 및 DB 구축

### 제1절 원시데이터 수집 및 전처리

#### 1. 차량 GPS원시 데이터 수집 및 분석

##### 가. 차량 GPS 데이터 수집 개요

- 본 사업을 통해 수집한 차량 GPS 데이터는 A사의 내비게이션 데이터와 B사의 내비게이션 데이터, 한국교통안전공단의 DTG데이터이며, 전국 단위 기초 교통DB 및 통행지표를 산출하기 위한 기반 데이터임
- 차량 GPS 데이터는 도로망 네트워크의 링크 기반으로 수집되는 링크 기반 데이터와 X, Y의 위치좌표인 GPS정보가 수집되는 포인트 기반 데이터로 구분됨
  - A사의 내비게이션 데이터는 도로망 네트워크와 연결된 링크 기반 데이터이며, DTG 데이터와 B사의 내비게이션 데이터는 GPS좌표인 포인트 기반 데이터로 데이터 수집방식이 다르나, 차량의 이동경로에 대해 시간의 연속성과 공간의 연결성이 동시에 수집되는 데이터임

<표 2-1> 차량 GPS 빅데이터 특징 (2020년 데이터 기준)

구분	A사 내비게이션 데이터	B사 내비게이션 데이터	DTG 데이터
제공 파일 구성	18,140개 텍스트 파일	1일 단위 텍스트 파일	1달 20개 압축파일
OBU ID 생성기준	경로 ID 발생	1일 단위 ID 갱신	차량등록번호로 ID 유지
수집정보 생성기준	내비게이션 실행 시	내비게이션 실행 시	차량 운행 시 수집
데이터 형태	링크 단위	1초 단위 포인트	1초 단위 포인트
좌표체계	EPSG 4301	WGS84	WGS84
용량(년)	7TB	4TB	248TB
이벤트(억/년)	520	2,300	7,300

- 작년 사업에서는 B사의 내비게이션 데이터만 이용하여 기초 교통DB 및 통행지표를 가공·구축하였음
- 시의성 있는 데이터 제공을 위해 최신년도 데이터 수집이 필요하며, 샘플율이 높은 내비게이션 데이터로 데이터의 신뢰도를 높이하고자 A사의 내비게이션 데이터를 추가로 수집하였음

#### 1) 차량 내비게이션 데이터

- 차량 내비게이션 데이터는 포인트 기반으로 제공되는 B사의 내비게이션 데이터와 링크 기반으로 제공되는 A사의 내비게이션 데이터로 구성
  - A사의 내비게이션 데이터는 링크 기반으로 경로 데이터가 제공되며, 링크 기반의 진입/진출 시간 정보 속도 등의 정보가 제공됨
  - A사의 내비게이션 주행 데이터는 A사에서 제공하는 네트워크와 매칭해야하며, 시기마다 주행 고유ID가 변경되는 특징을 가지고 있음
  - 주행 고유ID는 목적지를 설정하고 주행할 때마다 생성되어 동일한 사용자라도 주행 별로 다른 ID가 부여됨
  - 휴게소나 주정차 시 내비 종료 상태가 아니라면 정차상태로 데이터가 수집되지만, 종료 상태라면 경로를 임의로 생성하거나 보정하지 않음

<표 2-2> A사의 내비게이션 데이터 형식

항목	설명
데이터 발생 일자	최초 seq가 발생한 날짜
주행 고유 ID	비식별화 처리된 ID, 24WKFL
순번	주행 내 링크 순번
SK MESH 코드	SK Mesh의 앞 4자리
LINK ID	SK내의 LINK ID
LINK 방향	0: 정방향 / 1: 역방향
LINK 길이	단위 : 미터(m)
주행시간	LINK 주행 시간(초)
진입시간	링크 진입시간
진출시간	링크 진출시간
속도	GPS 데이터를 통해 계산된 속도

&lt;표 2-3&gt; A사의 내비게이션 데이터 현황

년	월	용량(TB)	파일갯수(개)	이벤트 수(개)	경로수(개)
2020년	1월	0.3038	1,009	2,475,362,209	27,706,314
	2월	0.4451	1,009	3,626,788,831	40,594,039
	3월	0.4347	2,017	3,541,753,478	39,642,252
	4월	0.4642	1,009	3,782,195,790	42,333,483
	5월	0.5133	1,009	4,182,042,856	46,808,904
	6월	0.5481	2,018	4,465,560,123	49,982,265
	7월	0.6006	1,383	4,893,739,357	54,774,804
	8월	0.5591	1,643	4,554,879,476	50,982,002
	9월	0.5686	2,000	4,216,918,666	30,900,226
	10월	0.9293	2,016	7,571,185,230	84,743,006
	11월	0.5856	2,018	4,771,456,843	53,406,116
	12월	0.4945	1,009	4,029,315,877	45,099,457
2021년	1월	0.4983	1,008	4,060,092,867	45,443,938
	2월	0.4262	2,018	3,469,051,388	38,867,054
	3월	0.5705	1,009	4,647,969,904	52,023,947
	4월	0.6129	2,018	4,993,206,540	55,888,123
	5월	0.6383	1,147	5,200,922,268	58,213,050
	6월	0.6374	1,880	5,193,370,835	58,128,528
	7월	0.6074	2,018	4,949,108,832	55,394,545
	8월	0.5957	1,009	4,853,204,951	54,321,109
	9월	0.6641	2,018	5,410,366,175	60,557,321
	10월	0.6797	1,010	5,537,683,441	61,982,362
	11월	0.6779	2,017	5,523,193,577	61,820,179
	12월	0.5916	2,018	4,820,278,123	53,952,564
합계		13.60	37,310	107,615,410,309	1,204,559,776

- B사의 내비게이션 데이터는 1일 단위 텍스트 파일로 제공되며, 수집 주기는 1초 단위로 경로ID, 시간정보, 위치정보, 방향각 등의 정보가 제공됨

<표 2-4> B사의 내비게이션 데이터 형식

항목	설명
단말기 ID	
시간정보	Timestamp
위도	ddmm,mmmm*10,000
경도	ddmm,mmmm*10,000
속도	
방위각	
도로판단	자사지도 기준 - 도로주행여부
경로탐색 플래그	주행(경로탐색주행, 일반운행)
재 탐색 실행	경로이탈로 인한 경로탐색여부
고도 값 부호	양수(0), 음수(1)
고도값	음수의 경우 -> 1값만 존재

<표 2-5> B사의 내비게이션 데이터 현황

년	월	용량(GB)	이벤트 수(개)	경로수(개)
2020년	1월	375.68	4,655,180,014	3,299,350
	2월	302.04	3,742,976,724	2,751,203
	3월	305.64	3,787,726,686	2,764,191
	4월	314.33	3,895,683,909	2,760,843
	5월	335.97	4,164,096,643	2,871,270
	6월	301.93	3,742,878,043	2,562,221
	7월	391.59	4,854,097,586	3,394,566
	8월	381.03	4,723,417,409	3,196,442
	9월	348.42	4,319,688,483	2,907,783
	10월	425.73	5,277,656,376	3,455,275
	11월	417.89	5,180,761,084	3,430,335
	12월	384.23	4,763,538,378	3,233,203
합계		4.28TB	53,107,701,335	36,626,682

## 2) DTG(차량디지털 운행기록) 데이터

- DTG 데이터는 영업용 자동차의 운행정보를 실시간으로 저장하여 변화하는 운행상황을 기록하는 디지털 운행기록계를 통해 수집되는 데이터이며, 운전자가 차량 디지털 운행기록을 제출하는 시기에 따라 데이터의 샘플율이 달라질 수 있음
- DTG 데이터는 자동차의 순간속도, GPS, 분당 엔진회전수, 가속도, 자동차 유형 등의 정보가 제공되며, 1초 단위의 GPS 좌표로 수집됨
- DTG 데이터의 자동차 유형 정보를 이용하여 세분화된 차종의 차량의 이동경로, 통행에 대한 분석을 수행할 수 있음

&lt;표 2-6&gt; DTG 데이터 형식

컬럼명	설명	정보 예시
TRIP_KEY	등록번호 & 정보발생 일시	C-125901568017101206094700
DTG_MDL_NM	단말기 모델명	XDT1000
CHASSIS_NO	차량 고유번호 1	XXXXXX301795
CAR_TP_CD	차량 유형 구분	11
CAR_REG_NO	차량 고유번호 2	-1259015680
BIZ_REG_NO	운송사업자 등록번호	XXXXXX47349
DRIVER_CODE	운전자 코드	1
DLY_DRIV_DIST	일일 주행거리	0
ACCM_DRIV_DIST	누적 주행거리	268897
OPT_SPD	차량속도	0
RPM	분당 엔진 회전수	393
BREAK_SIG	1: 유, 2: 무	0
GPS_X	차량위치 X (WGS84 경위도)	127075626
GPS_Y	차량위치 Y (WGS84 경위도)	37052681
GPS_AGL	지점별 방위각	0
ACCEL_VX	횡가속	1
ACCEL_VY	종가속	1
CONT_CODE	통신상태코드	11
AREA_CODE	행정기관코드, 대존코드	41
OPT_Time	YYMMDDHHMMSSSS	17101206094700

&lt;표 2-7&gt; DTG 데이터 현황

년	차종	월	용량(TB)	이벤트 수(개)	경로수(개)
2020 년	버스	1월	11.82	64,825,403,057	4,405,231
		2월	10.92	59,849,331,353	3,921,881
		3월	10.03	51,774,840,281	3,506,065
		4월	9.44	54,936,262,280	3,566,959
		5월	15.88	87,364,307,252	8,937,639
		6월	10.85	59,507,290,485	4,411,570
		7월	11.17	61,283,188,783	4,538,936
		8월	15.42	84,856,149,429	7,973,147
		9월	10.24	57,720,550,259	4,325,916
		10월	10.52	60,284,604,649	4,478,775
		11월	10.99	46,004,780,734	4,332,991
		12월	9.46	51,925,716,748	3,332,319
		소계	136.74	740,332,425,310	57,731,429
	트럭	1월	2.73	14,806,926,554	2,394,148
		2월	2.68	14,545,982,508	2,414,587
		3월	2.96	16,043,572,733	2,799,530
		4월	2.96	16,109,259,441	2,761,493
		5월	3.04	17,283,727,735	2,930,017
		6월	3.28	18,593,834,682	3,091,236
		7월	3.32	18,356,400,044	3,088,544
		8월	3.18	17,347,996,611	2,751,441
		9월	3.46	18,891,484,604	3,095,887
		10월	3.56	19,414,422,363	3,199,016
		11월	3.82	20,842,871,427	3,478,150
		12월	3.79	20,688,833,671	3,187,643
		소계	38.78	212,925,312,373	35,191,692
	택시	1월	6.73	37,433,361,409	5,308,625
		2월	6.08	33,851,139,308	4,912,292
		3월	5.46	30,429,403,508	4,733,337
		4월	5.77	32,149,322,402	5,077,446
		5월	6.29	35,113,244,132	5,594,653
		6월	6.97	41,039,652,013	6,548,458
		7월	6.97	39,884,074,872	6,230,267
		8월	6.69	37,427,220,873	5,596,802
		9월	6.23	34,811,983,566	5,523,908
		10월	7.21	40,380,292,340	6,537,490
		11월	2.65	14,840,254,731	2,272,113
		12월	5.81	32,608,872,783	5,201,819
		소계	72.86	409,968,821,937	63,537,210
	합계		248.38	1,363,226,559,620	156,460,331



### 나. 차량 GPS 원시 데이터 분석

- A사의 내비게이션 데이터는 2020년, 2021년을 24주차로 구분하여 총 180일의 데이터를 수집하였으며, 파일은 37,310개로 분할되어 있음

<표 2-8> A사의 내비게이션 데이터 시간적 범위

구분	2020년	2021년
1월	21일~27일(4주차)	18일~24일(4주차)
2월	10일~16일(3주차)	8일~14일(2주차)
3월	16일~22일(3주차)	15일~21일(3주차)
4월	13일~19일(3주차)	19일~25일(4주차)
5월	18일~24일(4주차)	17일~23일(4주차)
6월	15일~21일(3주차)	14일~20일(3주차)
...	...	...
10월	9일~15일(2,3주차)	8~14일(2,3주차)
11월	16~22일(3주차)	15~21일(3주차)
12월	14일~20일(3주차)	13일~19일(3주차)
명절	- 설날: 1월 24일~26일(3일) - 추석: 9월 29일~10월 5일(7일)	- 설날: 2월 11일~13일(3일) - 추석: 10월 9일~11일(3일)
합계	(7일×12주차) + 6일 = 90일	(7일×12주차) + 6일 = 90일

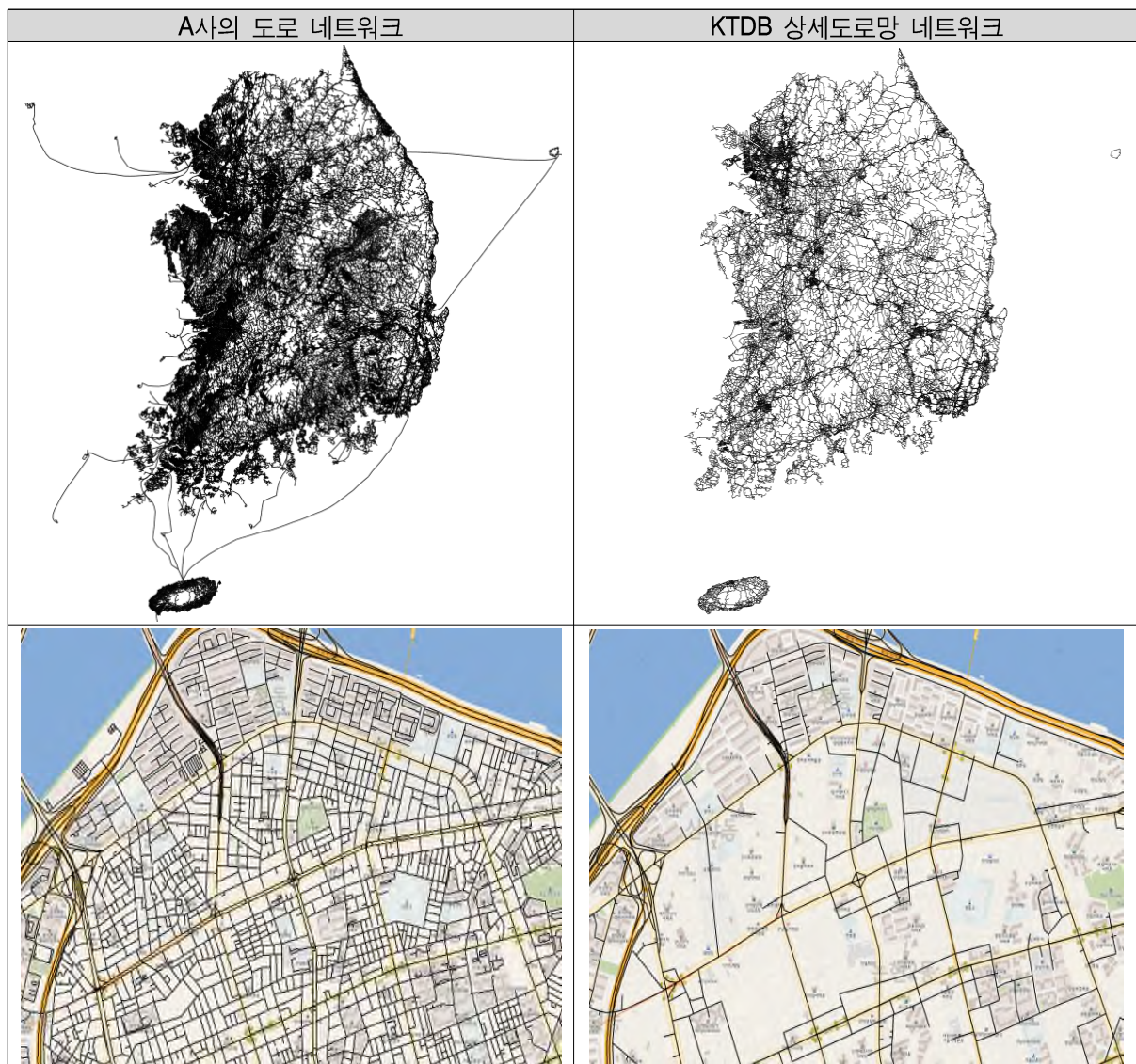
- 수집된 A사 내비게이션 데이터의 전체 용량은 13.6TB이며, 총 천 억건의 이벤트 수(라인수)가 발생되었으며, 개별 차량 경로수는 일 평균 약 600만 개 경로가 나타났음
  - 2020년의 데이터 용량은 6.40TB, 이벤트 수는 약 520억 건, 경로 수는 약 4.6억 건 2021년의 데이터 용량은 7.20TB, 이벤트 수는 약 586억 건, 경로 수는 약 5.4 억건으로 분석됨
  - 코로나 19 영향으로 2020년은 2021년과 비교했을 때 상반기의 이벤트 수, 경로 수가 적게 나타났으며, 2020년 10월에 코로나 19 확진자 수가 감소 추세가 있어 사회적 거리두기가 완화된에 따라 이벤트 수, 경로 수가 급격히 증가한 것으로 분석됨

- A사 내비게이션 데이터는 링크 기반 데이터로 A사의 도로 네트워크 내의 링크ID를 바탕으로 정보를 제공
- A사의 도로 네트워크 데이터는 시기별로 링크 정보를 1개월 2회 업데이트하기 때문에 2020년, 2021년 각 12주차 주행 데이터와 매칭할 수 있도록 네트워크 Shape파일을 24개 버전을 이용해야 함
- 본 사업에서는 KTDB의 상세도로망 Level6 네트워크로 차량 GPS 빅데이터 가공 및 지표를 생성하기 때문에 SK 도로 네트워크의 링크ID로 제공되는 경로를 Level6 네트워크 기반의 경로로 재생성 필요함

<표 2-9> A사의 도로 네트워크 데이터 형식

컬럼명	설명
IDXNAME	링크의 MESH 코드
LINK_ID	주행 고유 ID
TLINKIDP1	정방향 교통정보 LINK ID
TLINKIDN1	역방향 교통정보 LINK ID
LENGTH	MESH 코드
ST_DIR	링크의 MESH 내 시작 노드의 NODE ID
ED_DIR	링크의 MESH 내 끝 노드의 NODE ID
ROAD_CATE	도로등급
ROADLEVEL	경로안내를 위해 수정된 링크의 도로등급
ROADSTATE	도로의 포장상태(0: 미조사 / 1: 비포장/ 2: 포장)
...	...
KEY	이종데이터와의 매칭을 위한 값
KS1	정방향 국토부교통정보 LINK ID
KS2	역방향 국토부교통정보 LINK ID

- A사의 도로 네트워크의 링크를 상세도로망 Level6 링크로 경로 재생성 모듈을 개발하기 위해서는 SK 도로 네트워크의 링크와 Level6 링크의 차이를 분석함
- A사의 도로 네트워크의 단방향 링크 연장은 447,610km(2021년 12월 기준)로 상세도로망 Level6 네트워크 연장인 126,040km(2021년 12월 기준)에 비해 3.5배 상세한 수준인 것으로 분석됨



<그림 2-1> A사의 도로 네트워크와 KTDB 상세도로망 Level6 네트워크 상세도 비교

○ A사의 경로 데이터 샘플을 분석

- 전체 교통량 대비 A사의 데이터의 샘플율을 분석하기 위해 관측교통량과 A사의 내비게이션 통행량 자료의 비교를 수행함
- 이를 위해 고속도로, 도시고속도로, 광역시도, 일반국도, 지방도에 대한 관측지점을 선정하였으며, 교통량 수준을 고려하여 선정하였음
- 고속도로와 도시고속도로와 같은 연속류 도로에서는 A사 내비게이션 통행량 샘플율이 약 33%~40% 수준으로 나타나 연속류 도로의 샘플율이 매우 높은 것으로 분석됨
- B사의 내비게이션 통행량은 약 0.1%~1.0% 수준으로 낮은 샘플율을 보이고 있음

- 특히 교통량 수준이 낮은 500대 수준의 도로에서도 A사의 샘플율이 18% 수준으로 나타나 저규격 도로에서도 높은 샘플율을 확보하고 있는 것으로 분석됨

<표 2-10> 관측교통량 대비 A사의 내비게이션 데이터 샘플율 분석 결과

구분	도로명	관측교통량 <sup>8)</sup> (대/일)	A사 통행량 <sup>9)</sup> (대/일)	A사 샘플율(%)	B사 통행량 <sup>10)</sup> (대/일)	B사 샘플율(%)
고속도로	경부고속도로 (관교→신갈)	101,152	33,515	33.1%	129	0.1%
	서해고속도로	68,259	27,500	40.3%	559	0.8%
도시고속 도로	서울외곽순환고속도 로 (청계TG)	80,165	33,571	41.9%	775	1.0%
광역시도	명덕로	6,982	1,561	22.4%	53	0.8%
일반국도	서동대로	3,643	967	26.5%	28	0.8%
	중원대로	5,825	1,100	18.9%	33	0.6%
지방도	한불로	2,062	138	6.7%	5	0.2%
	쌍곡로	545	99	18.2%	2	0.3%

- 교통량 수준에 따른 A사의 내비게이션 통행량 샘플율 분석
  - 관측교통량이 1만대를 기준으로 샘플율이 약 10%가 확보되는 것으로 나타남
  - 관측교통량이 6.5만대를 기준으로 샘플율이 약 20%가 확보되는 것으로 나타남

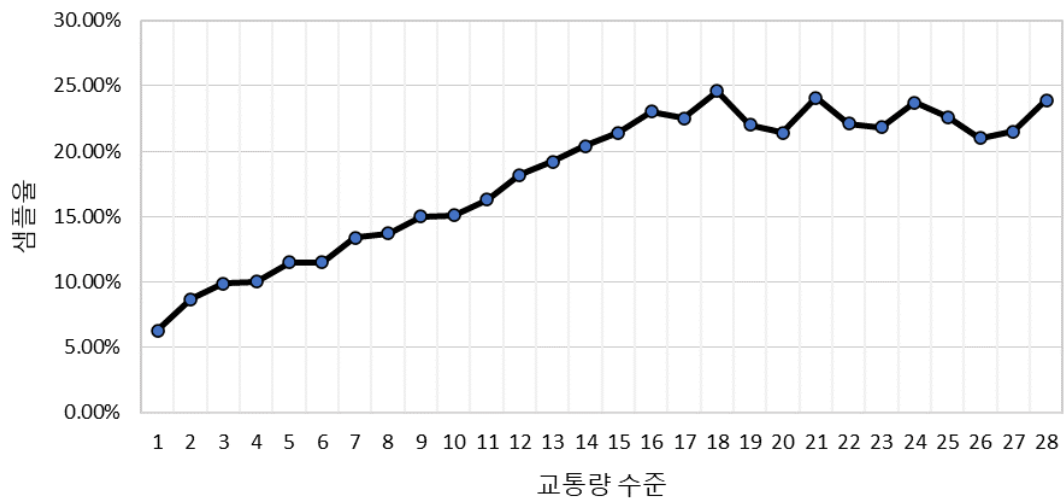
8) 2020년 기준 관측교통량 데이터를 사용하였음

9) A사 통행량 데이터는 2021년 10월 8일 ~ 14일 일주일을 대상으로 산출한 일평균 통행량 데이터임

10) B사 통행량 데이터는 2020년 1년을 대상으로 산출한 일평균 통행량 데이터임

&lt;표 2-11&gt; 교통량 수준별 샘플율

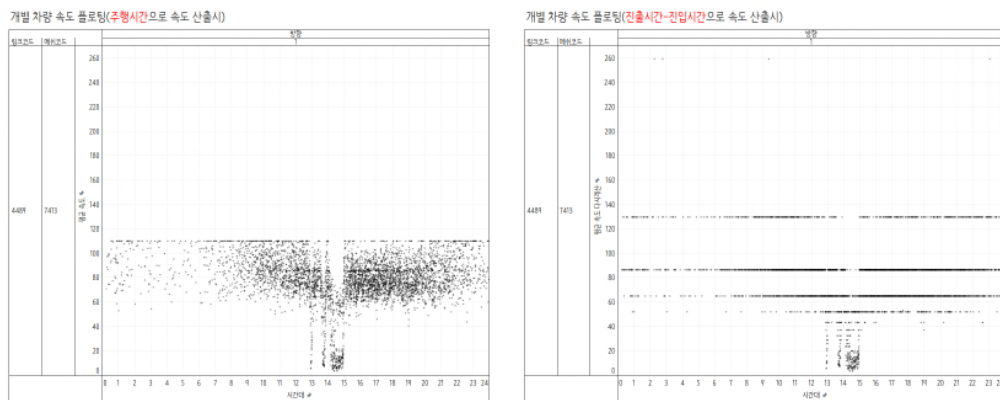
교통량 수준	교통량 범위		관측지점수	관측교통량(대/일)	A사 통행량(대/일)	샘플율
1	0K	5K	4,763	9,465,116	592,693	6.3%
2	5K	10K	1,616	11,735,501	1,026,755	8.7%
3	10K	15K	998	12,324,726	1,224,000	9.9%
4	15K	20K	751	12,948,360	1,290,957	10.0%
5	20K	25K	525	11,748,989	1,347,362	11.5%
6	25K	30K	318	8,694,657	1,004,045	11.5%
7	30K	35K	231	7,463,739	998,341	13.4%
8	35K	40K	155	5,764,780	791,667	13.7%
9	40K	45K	116	4,917,777	738,588	15.0%
10	45K	50K	82	3,874,575	584,516	15.1%
11	50K	55K	59	3,112,691	508,680	16.3%
12	55K	60K	29	1,666,707	303,305	18.2%
13	60K	65K	35	2,196,546	422,383	19.2%
14	65K	70K	24	1,617,700	329,702	20.4%
15	70K	75K	33	2,378,583	509,353	21.4%
17	75K	80K	35	2,709,083	624,200	23.0%
18	80K	85K	14	1,158,694	260,325	22.5%
19	85K	90K	20	1,745,220	430,008	24.6%
20	90K	95K	9	830,423	182,839	22.0%
21	95K	100K	18	1,744,302	372,445	21.4%
22	100K	105K	12	1,228,387	296,072	24.1%



## 2. 차량 GPS원시 데이터 오류 검토

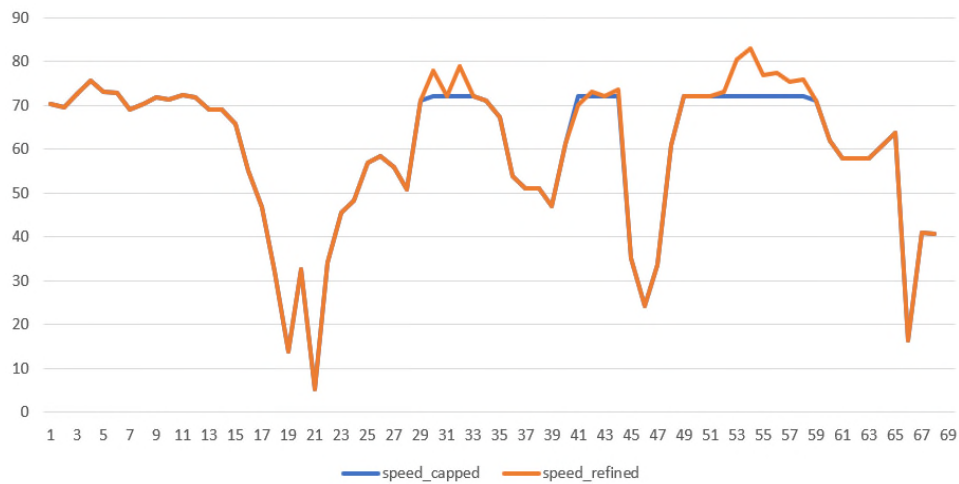
### 가. A사 내비게이션 원시 데이터의 오류 검토

- A사 내비게이션 데이터의 속도 정보 검토
  - 개별 차량 속도 형태를 플로팅 했을 때 ‘주행시간’으로 속도 산출과 ‘진출시간-진입 시간’으로 속도 산출 시 속도 분포의 형태가 다름을 분석함
  - 주행시간으로 속도 산출 시 개별 차량 속도 형태를 그래프화 (왼쪽 그림)
    - 최대속도에 대해 제한을 두어 특정 상한 속도 이상을 넘어갈 수 없음
  - 진출시간-진입시간으로 속도 산출 시 개별 차량 속도 형태를 그래프화 (오른쪽 그림)
    - 진출입 시간의 최소단위가 초이기 때문에 속도 분포가 일반화되는 현상이 발생



<그림 2-2 > 개별 차량 속도 플로팅 비교

- A사 내비게이션 데이터 제공 필드 조정
  - A사 내비게이션 데이터를 적재할 때 속도 제한이 없는 상태의 진출·입 시간이 초 단위로 반올림되도록 가공됨
  - 반올림되지 않은 필드를 포함하여 데이터를 재추출하는 것으로 조정함
  - 속도 제한이 없는 상태에서 속도를 계산 시 이상 값이 나타나는 경우 “단말기에서 올려준 속도 값”을 활용하여 보정
  - GPS가 미 존재하여 실제 속도 값이 계산되지 않은 경우 제한이 걸린 속도 사용



&lt;그림 2-3&gt; 속도 제한 상태 비교

&lt;표 2-12&gt; A사 내비게이션 데이터의 조정된 데이터 형식

컬럼명	설명
TRIP_INIT_SESSION_ID	주행 고유 ID
LINK_SEQ	주행 내 링크 시퀀스
MESH_CD	MESH 코드
LINK_ID	LINK 코드
LINK_DIR	LINK 방향 (0 : 정방향 / 1 : 역방향)
LENGTH	LINK 길이
DRIVING_TIME	LINK 주행시간 (초)
ENTRY_TIME	진입시간(연월일시분초)
EXIT_TIME	진입시간(연월일시분초)
SPEED_CAPPED	CAP이 찍워진 상태로 계산된 속도값
SPEED_GPS	단말기에서 올려준 속도값의 평균
SPEED_REFINED	보완된 최종 속도값
CAPPED_YN	CAP 쓰워짐 여부

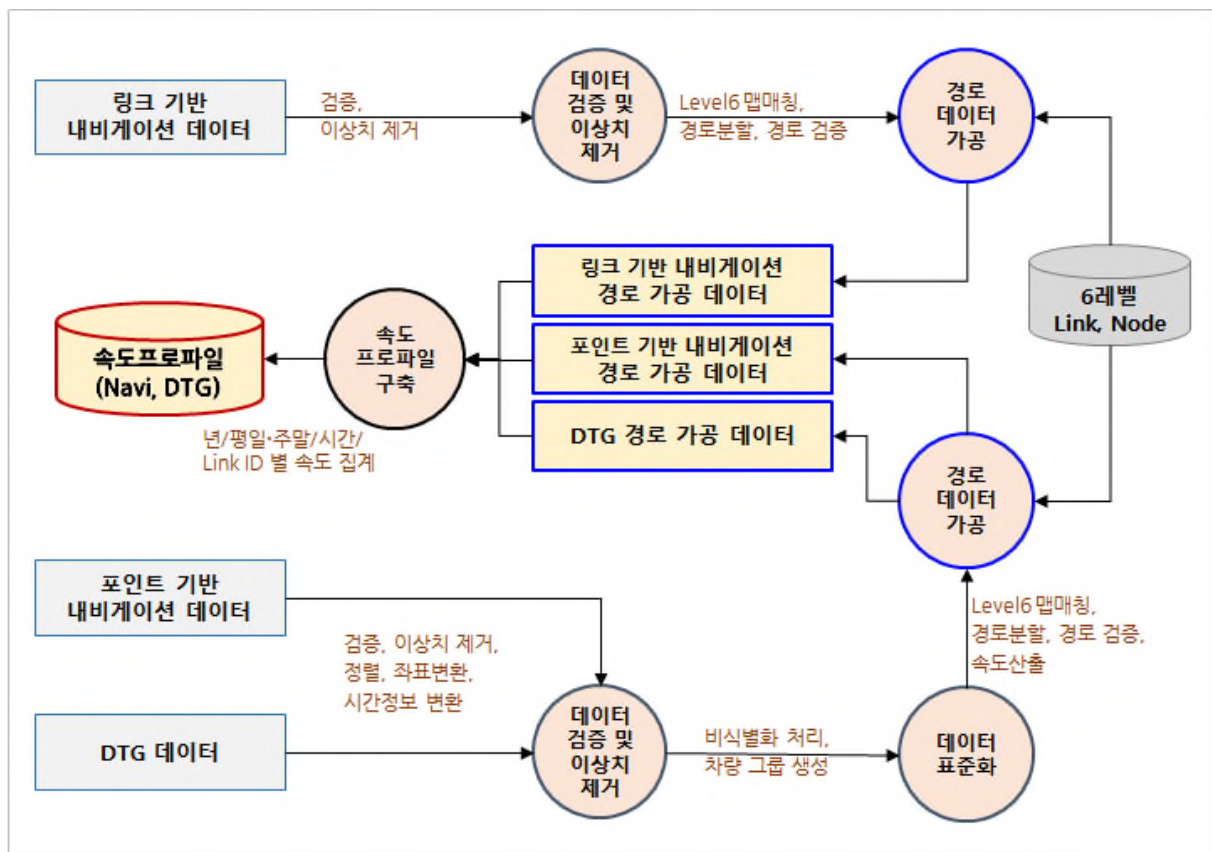


## 제2절 차량 경로데이터 가공 및 기초교통DB 구축

### 1. 차량 경로 데이터 가공

#### 가. 경로 데이터 가공 개요

- 차량 GPS 원시 데이터는 링크 기반의 내비게이션 데이터, 포인트 기반의 내비게이션 데이터로 구분되며, 각각의 제공되는 데이터 포맷 및 데이터 특성이 서로 상이하여, 이를 고려한 데이터 전처리 및 경로 데이터 가공이 필요함



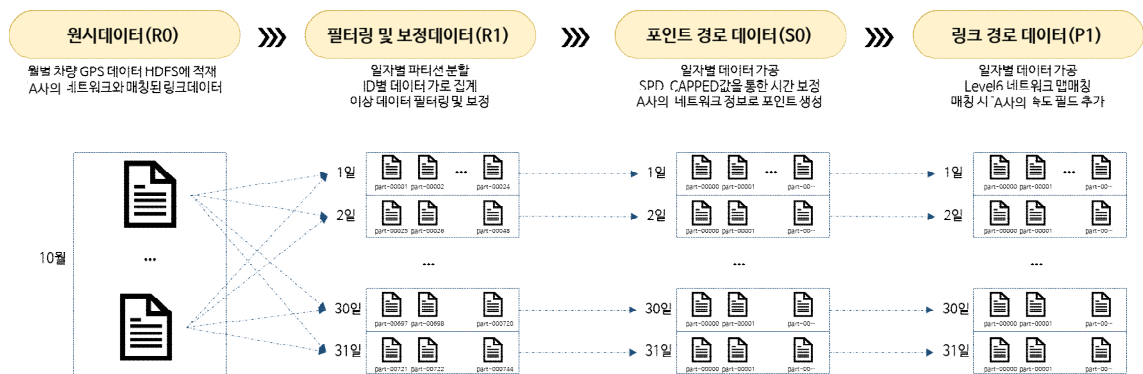
<그림 2-4> 차량 GPS 전처리 및 데이터 가공 프로세스

- 링크 기반으로 제공되는 A사 내비게이션 데이터는 A사 도로 네트워크의 링크관련 정보가 제공되어, 이에 A사 도로 네트워크의 링크와 상세도로망 Level6 네트워크의 링크와의 맵매칭 및 경로가공 모듈을 이용하여 경로데이터를 가공함



## 나. 링크 기반 데이터 전처리 및 경로 데이터 가공 알고리즘 개발

- A사 내비게이션 데이터는 링크 기반 데이터로 기존에 개발된 포인트 기반 데이터 맵 매칭 및 경로가공 모듈을 사용하기에 데이터 형식 및 특징이 다르므로 링크 기반 데이터를 궤적 정보로 생성하여 포인트 기반 데이터 맵매칭 및 경로 가공 모듈을 이용함
- A사 내비게이션 데이터는 B사 내비게이션 데이터와 차이가 있어 오류 유형에 따른 보정 작업을 수행하였으며, 보정된 링크 기반 데이터를 궤적 정보로 만드는 가공 모듈을 개발함
- A사 내비게이션 원시데이터 전처리 및 경로 데이터 가공 프로세스는 다음과 같음



<그림 2-5> 링크 기반 데이터 전처리 및 경로 데이터 가공 프로세스

### 1) 원시데이터(R0) 오류 유형에 따른 보정(R1)

- 유형 1 : DATE 필드 오류
- DATE 필드의 날짜와 링크 진입시간의 일자가 일치하지 않음 → 링크 순서 필드로 정렬 후 첫 링크의 진입시간 필드를 통해 일자별 분할

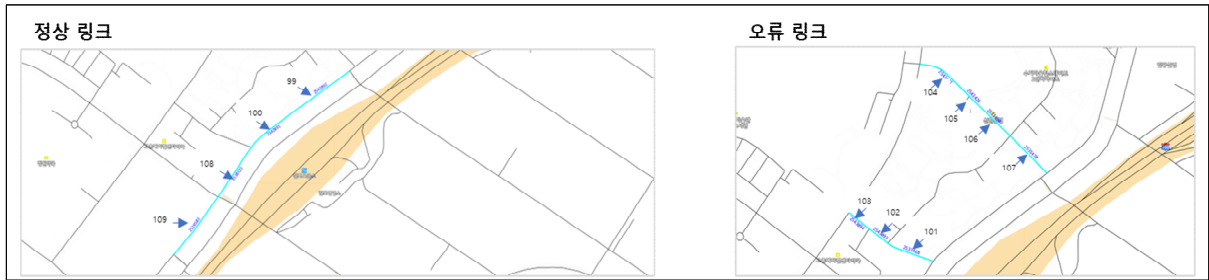
20210208	105520210208235734644500	...	1	4787	16033	0	63.894627	6.790106	20210208000226	20210209000232	33.8758566066568	WN	...	33.8758566066568	n	...
20210208	105720210208235718506500		1	5631	114	0	121	44.565998	20210208000653	20210209000737	9.77426781736157	WN		9.77426781736157	n	
20210208	108120210208232709308500		1	4777	3744	1	26	2.8363636	20210208073336	20210209073339	33.0000004230769	32.0		32	y	
20210209	105020210209175213265501		1	7413	28853	1	84	27.49091	20210210030545	20210210030559	10.9999996362434	WN		10.9999996362434	y	
20210209	105020210209210128425500		1	4775	9789	1	231.51741	11.883713	20210210084021	20210210084032	70.1348708101584	WN		70.1348708101584	n	

- 유형 2 : 주행 링크 순서 정렬 오류
- LINK\_SEQ의 진입시간이 이전 LINK\_SEQ의 진출시간과 불일치 오류 → 진출시간, 진입시간, 주행 링크 순서로 정렬(네트워크 형상에 맞춘 정렬)

20210208	105020210208000357760500	23	4778	10070	0	63	3.8860536	20210208001414	20210208001418	58.3625506349166	#N	58.3625506349166	n
20210208	105020210208000357760500	(1) 24	4778	20108	0	440	26.829266	20210208001418	20210208001445	59.0400050452368	#N	59.0400050452368	n
20210208	105020210208000357760500	(2) 25	4778	10082	1	18	1.0813245	20210208001445	20210208001446	59.9265178818233	#N	59.9265178818233	n
20210208	105020210208000357760500	(3) 26	4778	20109	1	9	0.5406622	20210208001446	20210208001447	59.9265012559596	#N	59.9265012559596	n
20210208	105020210208000357760500	(4) 27	4778	10089	0	330	19.592194	20210208001447	20210208001450	60.6363942700853	#N	60.6363942700853	n
20210208	105020210208000357760500	(5) 28	4778	10079	0	7	0.4235841	20210208001447	20210208001447	59.4923180544312	#N	59.4923180544312	n
20210208	105020210208000357760500	29	4778	10090	1	34	2.0437646	20210208001507	20210208001509	59.8894804225497	#N	59.8894804225497	n
20210208	105020210208000357760500	30	4778	10083	0	37	2.1552153	20210208001509	20210208001511	61.8035701583967	#N	61.8035701583967	n

○ 유형 3 : 경로 오류

- 비정상적인 경로로 경로 순서가 되어 있는 경우 → 진출 시간보다 다음 진입시간이 더 작은 경우 현재 링크의 진출 시간을 다음 링크의 진입시간으로 보정



○ 유형 4 : 시간 오류

- 오류 유형 3까지 보정했음에도 발생하는 시간 오류들 → 오류들을 필터링하여 제거

1) IN_TIME 오류 : 이전 레코드의 진출시간보다 현재 레코드의 진입시간이 작은 경우													
										id=105020211008000610512S00			
20211008	105020211008000610512S00	27	5724	3643	0	248	23.28551	20211008001349	20211008000746	36.34144	-1	36.34144	n
20211008	105020211008000610512S00	1	5724	5522	0	86	12.39884	20211008000746	20211008000758	33.96649	-1	33.96649	n
20211008	105020211008000610512S00	2	5724	5521	1	50	7.283177	20211008000758	20211008000806	29.94573	-1	29.94573	n
20211008	105020211008000610512S00	3	5724	3243	0	31	3.001319	20211008000806	20211008000809	41.02395	-1	41.02395	n
2) OUT_TIME 오류 : 현재 레코드의 진입시간보다 현재 레코드의 진출시간이 작은 경우													
										id=105020211008140158359S00			
20211008	105020211008140158359S00	90	4862	5506	1	25	1.944433	20211008143054	20211008143056	46.28599	-1	46.28599	n
20211008	105020211008140158359S00	91	4862	1329	0	46	4.702611	20211008143056	20211008143101	35.21448	-1	35.21448	n
20211008	105020211008140158359S00	29	4862	5512	1	232	51.78154	20211008141256	20211008143237	17.80464	-1	17.80464	n
20211008	105020211008140158359S00	94	4862	6369	0	8	1.033302	20211008143237	20211008143238	27.8718	-1	27.8718	n
20211008	105020211008140158359S00	95	4862	6366	1	7	0.967499	20211008143239	20211008143240	26.04655	-1	26.04655	n

○ 유형 1~4까지의 원시데이터 오류를 보정하여 오류 제거된 데이터(R1) 구축

2) 보정데이터(R1)를 궤적 정보데이터(S0)로 생성

- 데이터의 진출·입 시간은 캡이 씌워진 속도(제한된 가공 속도)로 계산되어있기 때문에 실제 진출·입 시간을 구하기 위한 보정이 필요함
  - 보완된 속도를 통한 링크 UP\_TIME 계산
  - 이전 레코드와 현재 레코드 간 GAP\_TIME을 계산
  - 첫 레코드의 링크 진입 시간을 기준으로 링크의 시간 재계산
  - 현재 레코드의 진입 시간은 이전 레코드의 진출 시간 + GAP\_TIME

- 현재 레코드의 진출 시간은 현재 레코드의 진입시간 + UP\_TIME
- 1초 단위로 궤적(포인트) 생성

시간 보정 전						시간 보정 후		
순번	진입시간 (sec)	진출시간 (sec)	주행시간 (sec)	링크길이 (m)	티맵_속도 (km/h)	계산된_주행시간 (millisec)	순번	진입시간 (millisec)
1	1633657251	1633657261	10	69	22.9820676	10808.427	1	1633657251000
2	1633657261	1633657281	20	67	12.52166762	19262.61	2	1633657261808
3	1633657281	1633657285	4	21	16.08886947	4698.9007	3	1633657281071
4	1633657285	1633657307	22	78	13.08372128	21461.784	4	1633657285770
5	1633657307	1633657316	9	38	14.21919766	9620.796	5	1633657307232
6	1633657316	1633657325	9	57	23.73010544	8647.2435	6	1633657316853
7	1633657325	1633657327	2	15	29.96712773	1801.9745	7	1633657325500
8	1633657327	1633657333	6	57	35.53801577	5774.0984	8	1633657327302
9	1633657333	1633657337	4	45	34.59832941	4682.307	9	1633657333076
10	1633657337	1633657344	7	36	18.56051359	6982.5654	10	1633657337758
11	1633657344	1633657354	10	52	19.72996923	9488.104	11	1633657344741
12	1633657354	1633657377	23	293	46.28127458	22791.075	12	1633657354229
13	1633657377	1633657386	9	137	51.88135664	9506.305	13	1633657377020
14	1633657386	1633657394	8	104	45.49400364	8229.656	14	1633657386526
15	1633657394	1633657501	107	166	4.550305048	131331.854	15	1633657394756
16	1633657547	1633657560	13	107	31.50285292	12227.464	16	163365752088
17	1633657560	1633657579	19	47	8.596289878	19682.91	17	1633657584315

&lt;그림 2-6&gt; 보정된 데이터 전·후 비교



&lt;그림 2-7&gt; 1초 단위 궤적(포인트) 데이터 생성 결과

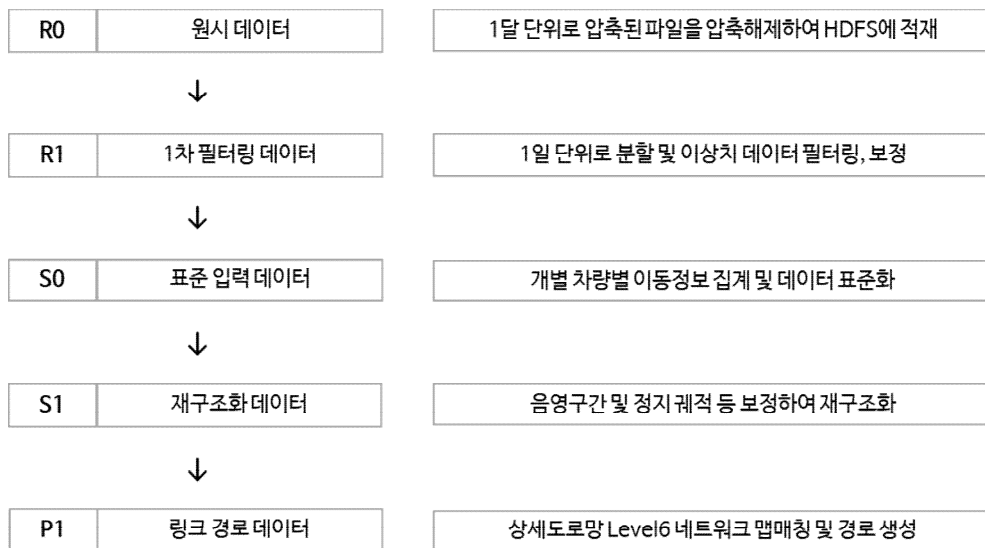
### 3) 궤적(S0)데이터의 경로 데이터(P1) 가공

- 차량 GPS 데이터(S0)를 상세도로망 Level6 네트워크 기반으로 경로를 가공함
- Level6 네트워크는 전국 2차선 이상의 도로망으로 노드와 링크로 구성되며, 전처리한 차량 GPS 데이터를 네트워크의 링크와 맵매칭하여 경로를 가공함
  - 맵매칭 및 경로생성: 포인트와 링크 간의 맵 매칭을 통해 경로를 생성하며, 포인트 위치에서 근접한 링크를 검색 후 최대 비용 탐색 방법을 이용하여 최대 비용을 나타내는 경로를 선정하여 링크와 맵 매칭함
  - 차량ID별 링크 구간 통행속도 산출 : 링크와 매칭된 궤적정보의 도로구간 길이와 진입시간정보를 이용하여 개별차량의 방향별 통행속도 산출
  - 통행병합: 고속도로 구간, 휴게소에서 종료된 경우 분리된 경로를 하나의 경로로 생

성하는 것으로 휴게소-휴게소, 고속도로-고속도로, 도시고속도로-고속도로일 때 진행함, 병합 시 시간 정보 및 구분 코드 정보를 추가함

#### 나. 포인트 기반 데이터 전처리 및 경로 데이터 가공

- DTG데이터는 포인트 기반 데이터로 포인트 기반 데이터 맵매칭 및 경로 가공 모듈을 이용함(해당 모듈은 R&D사업을 통해 개발된 것으로 간단하게 내용을 수록함)
- DTG 원시데이터 전처리 및 경로 데이터 가공 프로세스는 다음과 같음



<그림 2-8> 포인트 기반 데이터 전처리 및 경로 데이터 가공 프로세스

##### 1) 데이터 검증 및 이상치 제거(R1 데이터 생성)

- 1초 단위 정보가 생성되는 차량 GPS 데이터는 통행정보 이외에도 부가 정보가 많아 1년치 데이터를 수집하면 약 100TB 이상의 데이터 저장소가 필요하여 이를 빠르게 가공하기 위해서는 불필요한 부가 정보를 제외한 데이터 적재가 필요함
- 데이터 검증 및 필터링 : 오류 발견, 보정, 삭제 및 중복성 확인 등의 과정을 통해 데이터 품질을 향상시킴
  - 차량 GPS 데이터 좌표 오류 필터링
  - 차량 GPS 데이터 중복 제거
  - 시간정보 오류 필터링 및 Timestamp 형식의 데이터 저장

## 2) 데이터 표준 DB 변환(S0 데이터 생성)

- 차량 GPS 데이터의 처리 속도 향상을 위해 데이터 구조를 변경하고 개인정보보호를 위한 ID변환 등 표준 DB 포맷으로 변환이 필요함
  - 내비게이션 데이터와 DTG 데이터의 부가정보가 다르게 구성되어 있으므로 향후 다양한 목적에 따라 이용할 수 있도록 데이터 구분 코드를 부여함
  - 기관마다 제공하고 있는 데이터의 좌표체계가 다르므로 동일한 좌표로 변환하는 작업을 통해 UTM-K 좌표로 변환함
  - 이벤트에 따라 세로로 나열된 정보를 OBU ID별로 시간을 정렬하여 가로로 나열하는 구조로 DB를 재구성함

## 3) 출·도착 궤적 분리(S1 데이터 생성)

- 차량의 궤적 데이터는 1일 기준으로 데이터가 연속 수집되기 때문에 출도착 기준의 통행분리가 필요함. 통행분리를 위하여, 개별 통행에 대한 형태 분석 및 분석결과를 토대로 출도착 분할에 대한 방안 필요
- 출발/경유/도착에 대한 판단 및 통행분리를 위하여 데이터를 분석한 결과, 1.목적지까지 정차없이 도착하는 경우와, 2.장거리 운행 중 휴게소에 정차/주유하기 위한 정차/배송업무 등으로 인하여 반복 정차하는 경우 등과 같이 통행 중에 경유지를 거쳐 목적지에 도착하는 통행 등이 분석되었음. 또한 차종별 통행 패턴 분석 시 화물은 차고지에서 시동을 끄지 않고 정차 및 차고지를 계속 주행하는 경우가 있는데, 운반한 물건을 적재하는 과정에서 생기는 특징으로 추정. 또한 버스는 회차지에서 정차하지 않고 주행하거나, 몇시간 이상 정차하는 등 다양한 경우가 발생하여, 이에 대한 사항을 고려한 통행분리 과정 필요



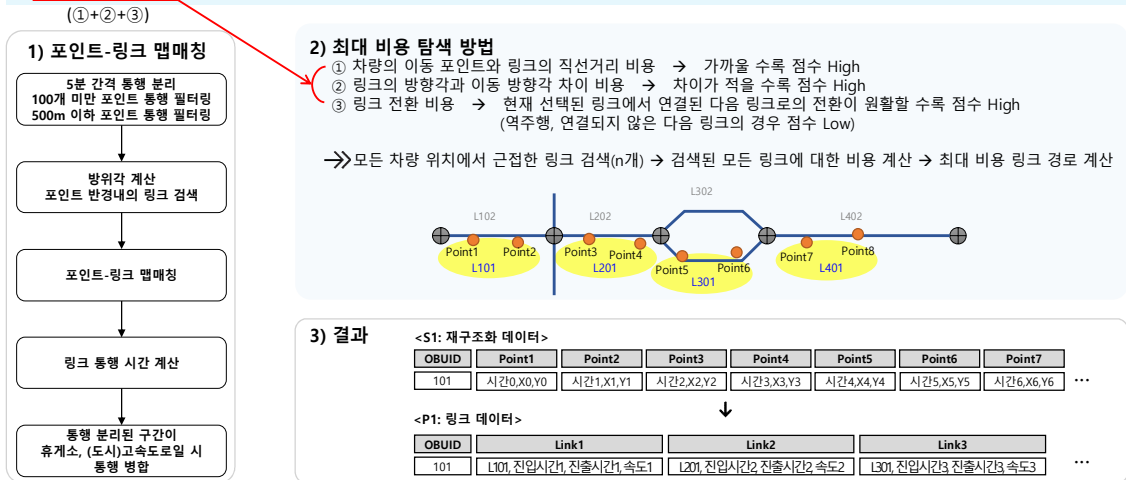
<그림 2-9> 통행 분리되지 않은 차종별 궤적

- GPS 데이터가 특정 시간 이상 위치정보를 수집하지 않을 시 해당 구간을 기준으로 통행을 분리
- 이전 포인트와 현재 포인트를 비교하여 좌표가 동일할 시, 현재 포인트로부터 50m 이내의 포인트 개수를 카운트. 카운트된 포인트의 개수가 300개 이상(5분 이상)일 시, 해당 지점으로부터 통행을 분리
- 버스, 트럭 데이터는 노선을 반복해서 통행하거나, 출발지로 회귀하는 특성이 존재. 차고지, 회차지를 판별 및 통행분리를 위해 시간/공간적 조건을 통한 통행분리 지점을 집계하여, 집계 결과를 이용한 통행분리 지점 판단 및 분리

#### 4) 포인트 기반 데이터와 Level6 네트워크 맵 매칭 및 경로 생성(P1 데이터 생성)

- 본 사업에서 구축하는 차량 GPS 데이터는 기초 교통 및 통행지표 구축 시 사용되는 주요 기반 데이터로 활용하기 때문에, 지표 생성 단위와 동일한 Level6 네트워크 기준으로 맵 매칭 후 링크 기준의 경로 데이터 생성 필요
  - 차량 GPS 데이터의 위치 정보를 기준으로 각 포인트의 진행 방향각 정보를 계산함. 방향각은 진북을 기준으로 시계방향인 우측 방향 각도를 계산
  - 포인트 자료 주변 도로 네트워크(링크) 검색. 검색된 링크를 대상으로 포인트와 링크와의 최단 거리, 링크에서의 방위각 계산, 링크를 따라서 이동한 거리 정보를 계산함
  - 모든 포인트는 주변에 검색된 링크를 대상으로 거리, 방위각 비용, 현재 링크와 다음 링크와의 전환 점수를 계산하게 되며 모든 점수가 최대 비용을 나타내는 경로를 선정 및 적재

• 최대 비용 탐색 방법을 통한 포인트-링크 맵매칭 → 링크 매칭의 정확도 향상

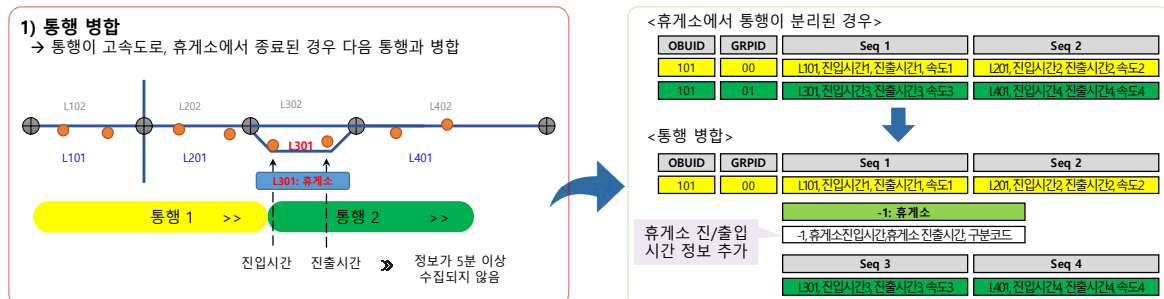


<그림 2-10> 차량 GPS 데이터와 링크 맵 매칭 및 경로 생성 프로세스

5) 통행 병합(P1 데이터 생성)

- 내비게이션 데이터는 사용자가 내비게이션 실행 시 수집이 진행되며, 종료시 수집이 종료됨. 이에 사용자가 휴게소에서 내비게이션 종료 또는 실행 시 휴게소를 기점으로 통행이 분리될 수 있기 때문에 통행의 시종점으로 적합하지 않는 구간은 연속된 통행을 유지할 수 있도록 처리함
- 통행이 병합되는 경우는 휴게소-휴게소, 고속도로-고속도로, 도시고속도로-고속도로 조합의 경우 발생
- 통행 병합 시 시간 정보 및 구분 코드 정보를 추가하여, 데이터 분석 및 검증 시 활용 수 있도록 함

• 통행 병합 → 휴게소, (도시)고속도로에서의 연속된 통행 유지



<그림 2-11> 통행 병합 프로세스



## 다. 경로 데이터 압축 및 통합 경로데이터 구축

### 1) 차량ID별 링크 구간 통행속도 산출

- 링크와 매칭된 차량 GPS 궤적 정보를 이용하여 활용목적에 따라 다양한 데이터를 구축하기 위해 개별 차량의 링크별 통행속도를 산출함
  - 기반 데이터에 따른 각 경로데이터를 도로네트워크와 맵매칭 후 통합 경로 데이터 구축
  - 링크 기반 경로데이터는 Level6 네트워크 링크 길이에 따른 속도와 실제 차량 궤적에 따른 속도정보도 포함하여 통행속도를 산출함
  - 포인트 기반 경로데이터는 링크와 매칭된 차량 GPS 궤적 정보의 도로구간 길이와 시간정보를 이용하여 통행속도를 산출함

### 2) 일별 차량ID별 경로 데이터 구축

- 링크와 매칭된 경로 데이터를 통행지표 생성을 위한 기초 DB로 활용됨
  - 통행정보, 통행속도, 공간정보가 결합된 경로 데이터 생성하여 표준 포맷으로 경로 DB를 구축함
  - OBU ID별로 1일 단위의 개별 차량의 경로 데이터 구축

<표 2-13> 개별 차량 경로데이터 테이블 구성

컬럼명	데이터유형	설명	코드	코드정보
OBUID	Integer	단말기 ID	-	-
Vehtype	Integer	차종 유형	A B AUTO BUS TRUCK	A사 B사 DTG 택시 DTG 버스 DTG 트럭
GroupNum	Integer	경로그룹ID	-	-
Seq	Integer	순서	-	-
Date	DateTime	수집일시	-	-
Vlink	Integer	Level6 가상링크ID	-	-
InTime	Integer	진입시간	-	-
OutTime	Integer	진출시간	-	-
Speed	Double	통행속도	-	-
Type	Integer	이상속도유형	1 4	링크속도 0이상 링크속도 미만



## 3) 경로 데이터 압축

- 경로 데이터 압축은 View-T 웹서비스를 위한 목적으로 수행하였음
- 상세도로망 Level6기반 경로 데이터는 대용량 데이터의 검색 속도의 한계로 서비스가 어렵기 때문에 경로 데이터를 주요도로망 Level5.5 단위로 압축하여 온라인 서비스를 수행함
- 주요도로망 Level5.5 기반 경로 데이터 압축 및 통합 경로 데이터 구축 시 610GB에서 150GB로 데이터 용량을 압축함

<6레벨 기준 경로 데이터>			<경로 데이터 압축 및 통합경로 데이터 구축>			
필드명	내용	비고	필드	내용	Type	PK
TnzOBUEID	단말기 ID		_id	경로 ID (일련번호로 구성)	NumberLong	*
TnPathGroup	단말기 별 GroupID		PL	경로 별 분석별 Link 구성 (1001,1002,1003,1004,1005)	Array(NumberInt)	
tnPathSeq	경로 순서		PT	경로 별 차종 정보 (Auto : 1, Bus : 2, Truck : 3) * 내비게이션은 승용차 '1'	NumberInt	
TnxPathLink	6lev 가상링크ID		PM	경로 별 월 정보 (1 ~ 12)	NumberInt	
In_Time	진입시간	DATETIME	PD	경로 별 일 정보 (1 ~ 31)	NumberInt	
Out_Time	진출시간	DATETIME	PW	경로 별 평일/주말 정보 (4개) (주말:0, 평일:1, 특송:10, 공유일:11)	NumberInt	
Speed	구간속도	1 ~ 150 km/h	DW	경로 별 요일 정보 (9개) (월:1, 화:2, 수:3, 목:4, 금:5, 토:6, 일:0, 특송:10, 공유일:11)	NumberInt	
type	보정코드		IT	경로 별 진입 시간(초단위로 구성, (시*3600+분*60+초)) 구성 (29460, 29470, 29475, 29485, 29498) (시간정보가 없는 경우 -1 처리함)	Array(NumberInt)	
emdlID	행정구역 (읍면동)		OT	경로 별 진출 시간(초단위로 구성, (시*3600+분*60+초)) 구성 (29463, 29474, 29483, 29495, 29504) (시간정보가 없는 경우 -1 처리함)	Array(NumberInt)	
			DT	행정구역 [3902011, 3902011, 3902011, 3902012, 3902012]	Array(NumberInt)	
			DT_SGG	행정구역[3901000, ...]	Array(NumberInt)	
			DT_SD	행정구역[3900000, ...]	Array(NumberInt)	

&lt;그림 2-12&gt; 압축된 경로 데이터 형태 예시

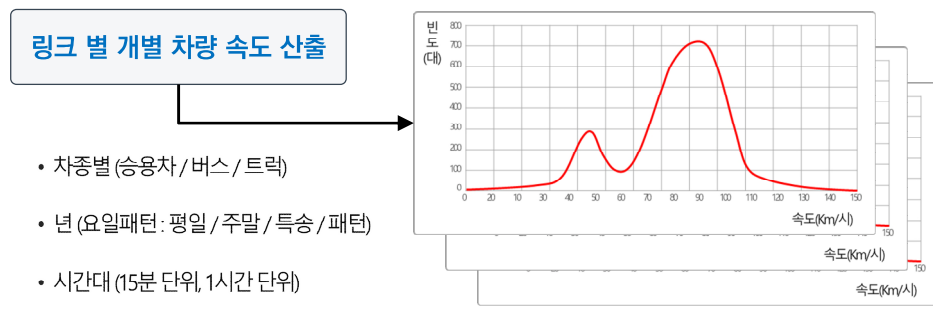
## 2. 기초교통DB 구축

### 가. 기초교통DB 구축 프로세스 및 데이터 정의

- 전국 교통량 전수화 및 통행지표를 구축하기 위하여, 경로데이터와 관측교통량을 이용하여 기초교통DB 구축
  - 기초교통DB를 구축하기 위해 개별 차량 경로데이터의 속도정보를 이용하여 구축한 속도프로파일과 개별차량 경로데이터를 링크별로 집계한 통행량, 연도별 이상치를 제거한 관측교통량이 필요함
- 속도프로파일 DB 구축 프로세스는 월단위 구축 → 년단위 병합(요일의 패턴별) 순으로 구축. 구축 시 내비게이션 데이터는 승용차로 구분하며, DTG는 버스/화물/택시 데이터로 구분하여 구축
- 관측교통량은 링크별 통행량과 비교하여 이상치가 있거나 연도별 패턴이 다른 관측교통량은 제거하여 구축
- 개별 차량 경로데이터를 이용하여 링크별 통행량을 산출하며, 관측교통량과 링크별 통행량으로 전수화하여 교통량 DB를 구축

### 나. 속도프로파일 분석 및 DB 구축

- 속도프로파일 DB 구축
  - 교통량, 평균속도 등 통행지표를 구축하기 위한 속도프로파일 DB 구축
  - 링크의 매칭된 궤적 정보를 활용해 개별 차량의 속도 산출 및 산출된 속도 범위에 따른 집계를 수행
  - 링크별 개별 차량 속도 프로파일 구축
  - 구축된 속도프로파일은 분포의 특성에 따라 이상치를 제거
  - 시간대가 없는 속도프로파일은 새벽, 비첨두 시간, 첨두 시간을 집계한 결과를 사용
  - 시간대/월/연 단위 대표 통행속도 산출
  - 연평균 평일/주말, 시간대별/전일 속도 프로파일 DB 구축



&lt;그림 2-13&gt; 속도프로파일 구축

- 속도 프로파일을 이용하여 구축되는 평균속도 자료를 활용하여 2019년과 2020년 평균속도의 증감패턴 분석을 수행
  - 서울시는 평균속도가 28.5km/h로 전체 시도 중 가장 낮게 나타났으며, 강원도는 73.8km/h로 전체 시도 중 평균속도가 가장 높게 나타남
  - 특히, 2019년 대비 2020년은 모든 시도에서 평균속도가 증가한 것으로 분석되었으며, 기타시도는 특광역시에 비해 평균속도 증감율이 더 높게 나타남

&lt;표 2-14&gt; 시도별 평균속도 증감패턴 분석

(단위 : km/h)

시도명	2019년	2020년	증감율
서울특별시	27.4	28.5	4.0%
부산광역시	35.3	38.3	8.3%
대구광역시	39.0	42.3	8.4%
인천광역시	37.2	40.1	7.9%
광주광역시	34.7	40.1	15.4%
대전광역시	39.7	45.8	15.3%
울산광역시	45.1	53.1	17.6%
세종특별자치시	52.9	55.7	5.2%
경기도	49.0	51.7	5.5%
강원도	65.6	73.8	12.5%
충청북도	65.3	73.0	11.7%
충청남도	66.6	72.9	9.4%
전라북도	60.5	72.3	19.5%
전라남도	64.1	70.5	10.1%
경상북도	68.3	77.8	14.0%
경상남도	60.2	66.5	10.4%
제주특별자치도	37.8	42.7	12.9%

주: 2019년은 B사 데이터이고, 2020년은 A사 데이터로 산출하였으며, 연도별 동일 링크 대상으로 분석

- 상세도로망 Level6 링크 단위 속도프로파일 테이블 정의서
  - 평일·주말/1시간 단위 or 15분 단위/차종별 DB생성. 이에 대한 테이블 정의서는 아래와 같음

<표 2-15> 상세도로망 Level6 링크 단위 속도프로파일 DB 테이블 구성

컬럼명	데이터 유형	설명
VlinkID	Integer	가상링크ID
Week_type	Integer	요일코드
Time	Integer	0시 ~ 23시
Speed	Integer	속도 코드 : 0km ~ 150km
Probe_count	Integer	통행량

#### 다. 관측교통량 데이터 분석 및 DB구축

- 관측교통량은 교통량 전수화를 위한 입력변수 데이터로 이용되며, 추정된 교통량의 검증 데이터로 이용됨
  - 지역간 도로의 관측교통량은 국토교통 통계연보에 수록되는 한국건설기술연구원과 한국도로공사의 교통량 데이터임
  - 도시부 도로의 관측교통량은 각 지자체에서 조사한 교통량으로 24시간대 관측된 교통량이 부족한 데이터임

<표 2-16> 지역간 도로 관측교통량 증감패턴 분석

시도명	지점수(개)	평균 관측교통량(대/일)		관측교통량 증감률
		2019년	2020년	
고속도로	623	49,281	48,225	-2.14%
일반국도	1,593	13,185	13,093	-0.70%
지방도	1,213	5,411	5,527	2.14%
국지도	364	8,834	8,761	-0.83%

자료1: 국토교통부(2021), 2020 도로교통량 통계연보, pp.23-24

〈표 2-17〉 도시부 도로 관측교통량 증감패턴 분석

시도명	지점수(개)	평균 관측교통량(대/일)		관측교통량 증감률
		2019년	2020년	
서울특별시	135	78,415	74,748	-4.7%
부산광역시	90	69,381	67,213	-3.1%
대구광역시	108	31,775	31,314	-1.4%
인천광역시	162	62,471	61,682	-1.3%
광주광역시	72	86,022	82,312	-4.3%
대전광역시	106	17,938	17,467	-2.6%
울산광역시	49	42,557	38,901	-8.6%
세종특별자치시	57	5,603	5,858	4.6%

자료1: 서울특별시(2021), 2020 서울특별시 교통량 조사자료, p.3

자료2: 부산광역시(2021), 2020서울특별시 교통량 조사자료, p.17

자료3: 대구광역시(2021), 2020서울특별시 교통량 조사자료, pp.61-73

자료4: 인천광역시(2021), 2020 도시교통 기초조사, p.17

자료5: 광주광역시(2021), 2020년 차량속도, 교통량 조사결과, pp.1-3

자료6: 대전광역시(2021), 대전광역시 2020년도 교통조사 및 분석 보고서, p.23

자료7: 울산광역시(2021), 2020 울산광역시 교통량 조사자료, pp.3-10

자료8: 세종특별자치시(2021), 세종시 2020년 교통 기초조사 및 분석, pp.16-21

## 라. 전국 교통량 전수화를 위한 교통량 추정 알고리즘 개선

### ○ 교통량 추정 알고리즘 개선 목적

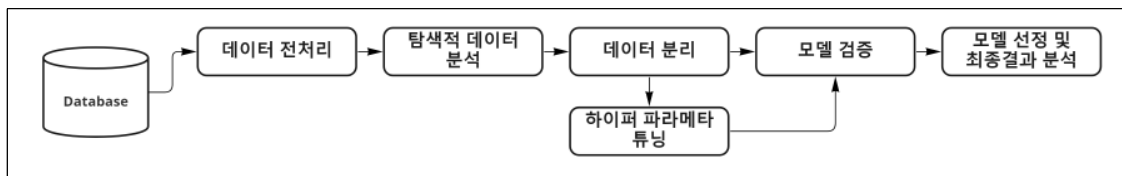
- 전국에서 조사된 교통량은 약 1만 개 지점으로 상세도로망 Level6네트워크의 링크 갯수(약 62만 개)에 비해 매우 낮은 샘플율을 가지고 있으며, 교통량 추정을 위해 사용된 차량 GPS 데이터의 샘플율도 점점 감소됨 → 데이터 샘플율에 영향을 받지 않는 일반화된 알고리즘 개발이 필요함
- 각 지자체에서 조사한 관측교통량은 차종이 구분되어 있지 않거나, 24시간 전체 시간대에 조사하지 않는 경우가 많아 차종별 시간대별 교통량 추정 시 신뢰도를 높일 방법을 개발함

### ○ 경로 데이터 기반 통행배정을 활용한 관측교통량의 차종별 교통량 추정

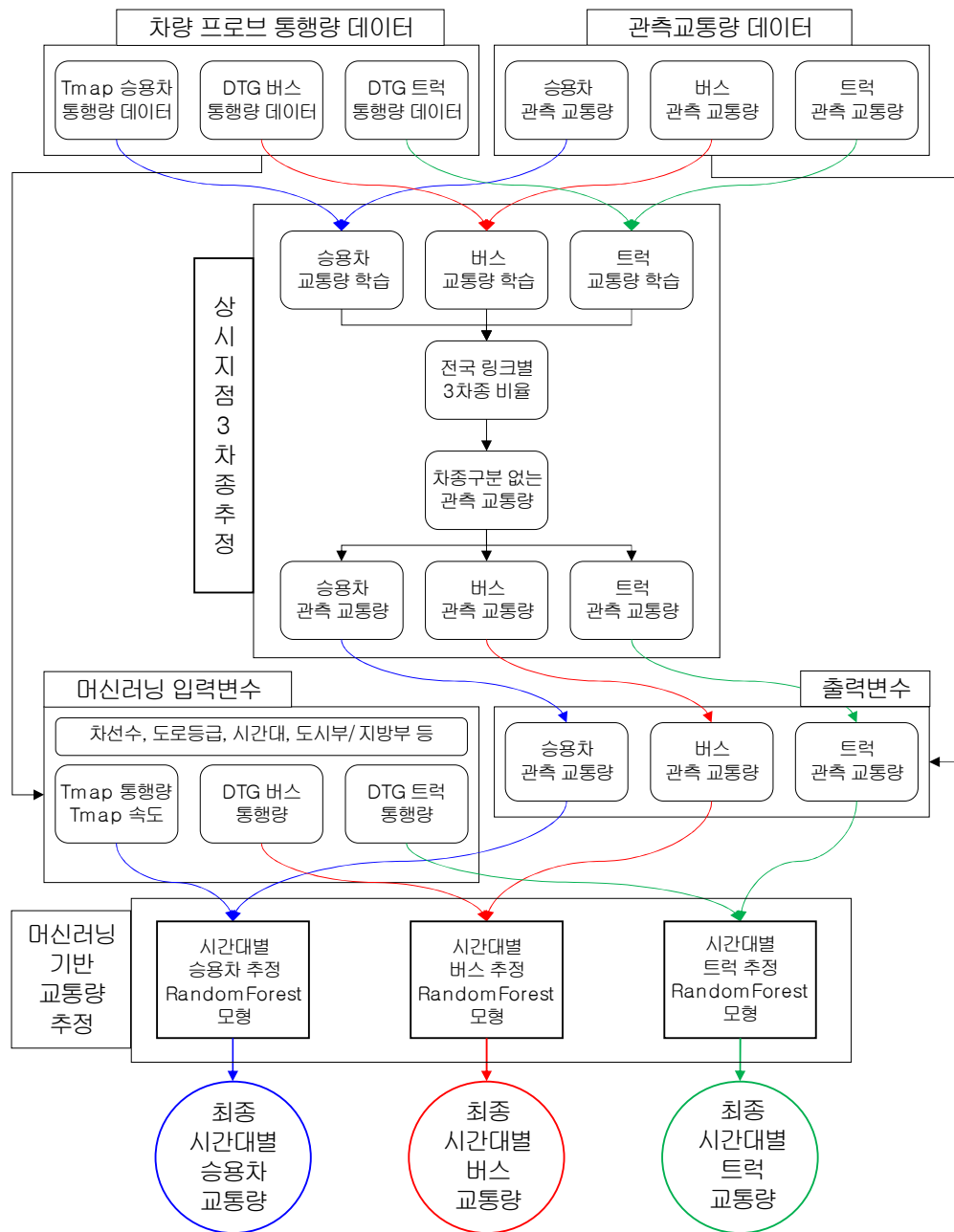
- 경로 데이터가 통과하는 관측교통량을 기준으로 경로 통행량을 전수화시켜 경로 통행 기반 전국 네트워크에 통행 배정
- 관측교통량의 차종 추정비율을 산정하며, 해당 결과를 머신러닝 앙상블 기법을 활용한 교통량 추정 알고리즘의 Input데이터로 사용
- 머신러닝 앙상블 기법을 활용한 교통량 추정 결과의 검증데이터로 이용

### ○ 랜덤 포레스트를 활용한 교통량 추정 알고리즘 개발

- 최근 회귀 및 분류 문제에서 좋은 성능을 보이고 있는 머신러닝 앙상블 모델 중 랜덤 포레스트 모델을 적용하여 연도별 일관성을 유지하는 일반화된 교통량 추정 알고리즘 개발
- 도로 속성 정보, 통행량(내비, DTG), 관측교통량, 방법1의 결과 데이터, 속도 등 데이터 기반의 전국 차종별 시간대별 교통량 추정



<그림 14> 머신러닝 기반 교통량 추정 알고리즘 개발 프로세스



&lt;그림 2-15&gt; 추정교통량 구축 프로세스

&lt;표 2-18&gt; 머신러닝 기반 추정교통량 모형 구축 프로세스

구분	설명	예시																																
1차 가공DB를 활용한 입력변수 구축	관측교통량, 차량 프로브 통행량, 평균속도 등 교통량 추정을 위한 입력 변수 구축	<table><tr><th>변수명</th><th>유형</th><th>설명</th></tr><tr><td>road_rank</td><td>정수형</td><td>도로등급</td></tr><tr><td>max_speed</td><td>실수형</td><td>제한속도</td></tr><tr><td>length</td><td>정수형</td><td>링크연장</td></tr><tr><td>lanes</td><td>정수형</td><td>차로수</td></tr><tr><td>emd_type</td><td>문자열</td><td>도시/지방 코드</td></tr><tr><td>total_probe</td><td>정수형</td><td>전체 프로브</td></tr><tr><td>tm_spd</td><td>실수형</td><td>티맵 속도</td></tr></table>	변수명	유형	설명	road_rank	정수형	도로등급	max_speed	실수형	제한속도	length	정수형	링크연장	lanes	정수형	차로수	emd_type	문자열	도시/지방 코드	total_probe	정수형	전체 프로브	tm_spd	실수형	티맵 속도								
변수명	유형	설명																																
road_rank	정수형	도로등급																																
max_speed	실수형	제한속도																																
length	정수형	링크연장																																
lanes	정수형	차로수																																
emd_type	문자열	도시/지방 코드																																
total_probe	정수형	전체 프로브																																
tm_spd	실수형	티맵 속도																																
데이터 전처리	수집된 데이터의 이상치 검토 및 전처리																																	
탐색적 데이터분석	변수별 상관관계 분석을 통한 입력 변수 선정																																	
입력변수 선정	도로등급, 제한속도, 도로연장, 차로수, 도시부/지방부, 차량프로브 통행량, 평균속도	<table><tr><th>번호</th><th>변수명</th><th>한글명</th><th>설명</th></tr><tr><td>1</td><td>road_rank</td><td>도로등급</td><td>고속도로, 도시고속도로, 일반국도, 특별광역시도, 국지도, 지방도, 시군도, 연결로</td></tr><tr><td>2</td><td>max_speed</td><td>제한속도</td><td></td></tr><tr><td>3</td><td>length</td><td>링크연장</td><td>미터</td></tr><tr><td>4</td><td>lanes</td><td>차로수</td><td>단방향 차로수</td></tr><tr><td>5</td><td>emd_type</td><td>도시/지방 코드</td><td>도시(0)1, 지방(0), 면(0)</td></tr><tr><td>6</td><td>total_probe</td><td>전체 프로브</td><td>티맵 통행량+링크웨이 통행량+트랙 통행량+버스 통행량+택시 통행량</td></tr><tr><td>7</td><td>tm_spd</td><td>티맵 속도</td><td>티맵 연평균 평균속도 티맵 월별 평균속도</td></tr></table>	번호	변수명	한글명	설명	1	road_rank	도로등급	고속도로, 도시고속도로, 일반국도, 특별광역시도, 국지도, 지방도, 시군도, 연결로	2	max_speed	제한속도		3	length	링크연장	미터	4	lanes	차로수	단방향 차로수	5	emd_type	도시/지방 코드	도시(0)1, 지방(0), 면(0)	6	total_probe	전체 프로브	티맵 통행량+링크웨이 통행량+트랙 통행량+버스 통행량+택시 통행량	7	tm_spd	티맵 속도	티맵 연평균 평균속도 티맵 월별 평균속도
번호	변수명	한글명	설명																															
1	road_rank	도로등급	고속도로, 도시고속도로, 일반국도, 특별광역시도, 국지도, 지방도, 시군도, 연결로																															
2	max_speed	제한속도																																
3	length	링크연장	미터																															
4	lanes	차로수	단방향 차로수																															
5	emd_type	도시/지방 코드	도시(0)1, 지방(0), 면(0)																															
6	total_probe	전체 프로브	티맵 통행량+링크웨이 통행량+트랙 통행량+버스 통행량+택시 통행량																															
7	tm_spd	티맵 속도	티맵 연평균 평균속도 티맵 월별 평균속도																															
모형 검토	선형회귀모형, 랜덤포레스트모형, XGboost 모형 선정																																	
교통량 추정모형 구축	전처리된 데이터와 선정된 모형을 활용하여 교통량 추정모형 구축																																	
교통량 추정 및 평가	구축된 교통량 추정의 추정교통량 신뢰도 분석																																	
교차검증	최종 선정된 랜덤포레스트 모형에 대한 교차검증을 통해 일반화 성능 검토																																	
변수 분석	교통량 추정에 영향을 끼치는 변수별 영향 분석																																	



#### 마. 전국 교통량 '19년, '20년, '21년 추정결과 비교

- 교통량의 신뢰도를 평가하기 위해서는 조사된 교통량 자료가 존재하는 지점에서의 교통량 추정을 통해 현실을 어느 정도로 구현하고 있는지에 대한 검토가 필요함
- 한국개발연구원(2008)에 따르면 추정교통량의 신뢰성을 평가하기 위해 조사교통량과의 오차율 기준으로 추정교통량에 대한 평가를 수행하고 있다. 본 연구에서 분석하고 있는 조사지점의 교통량 수준을 고려하면  $\pm 15\%$  이내의 오차율이 유의한 추정 범위라고 볼 수 있음
- 추정교통량의 신뢰도 평가지표는 ①평균절대비율오차(Mean Absolute Percent Error, MAPE, %)와 ②평균비율오차(Mean Percent Error, MPE, %), ③상관계수(coefficient of correlation,  $R$ ), ④결정계수(coefficient of determination,  $R^2$ )를 활용함

$$MAPE(\%) = \frac{\sum_{i=1}^n \left( \frac{|q_i - \hat{q}_i|}{q_i} \right) \times 100}{n}, \quad MPE(\%) = \frac{\sum_{i=1}^n \left( \frac{q_i - \hat{q}_i}{q_i} \right) \times 100}{n}$$

$$R = \frac{\sum_{i=1}^n q_i \hat{q}_i - \frac{\sum_{i=1}^n q_i \sum_{i=1}^n \hat{q}_i}{n}}{\sqrt{\left[ \sum_{i=1}^n q_i^2 - \frac{(\sum_{i=1}^n q_i)^2}{n} \right] \left[ \sum_{i=1}^n \hat{q}_i^2 - \frac{(\sum_{i=1}^n \hat{q}_i)^2}{n} \right]}}$$

여기서,  $q_i$  = 지점  $i$ 의 조사교통량(대/일)

$\hat{q}$  = 지점  $i$ 의 추정교통량(대/일)

$n$  = 전체 지점 개수(개)

- 국토교통부에서 수집하고 있는 도로교통량조사 자료(국토교통 통계연보)와 지자체에서 수집하고 있는 교통량 조사 자료 중 고속국도, 일반국도, 특별광역시도를 대상으로 신뢰도 분석을 수행하였으며, 선정된 평가대상 지점들은 24시간대가 모두 조사된 지점임
- 고속도로의 상시지점은 6.5%, 일반국도의 상시지점은 11.0%, 서울특별시 상시지점은 6.0% 수준으로 나타나 상시지점에 대해 유의한 결과가 나타났음을 분석하였으

며, 주요 도로등급별 교통량 신뢰도 분석을 수행했을 때 상시 지점과 비슷한 수준으로 나타났음을 확인할 수 있어 신뢰도 높은 교통량 추정이 되었다고 판단됨

<표 2-19> 2020년 상시 지점 추정교통량 신뢰도 분석

구분	신뢰도 평가대상 관측교통량 지점 수	MAPE(%)	MAE(%)	상관계수	결정계수
고속도로	236	6.5	-3.8	0.9980	0.9960
일반국도	1,076	11.0	-5.3	0.9926	0.9852
서울특별시	264	6.0	-4.0	0.9969	0.9939

<표 2-20> 2020년 주요 도로등급별 추정교통량 신뢰도 분석

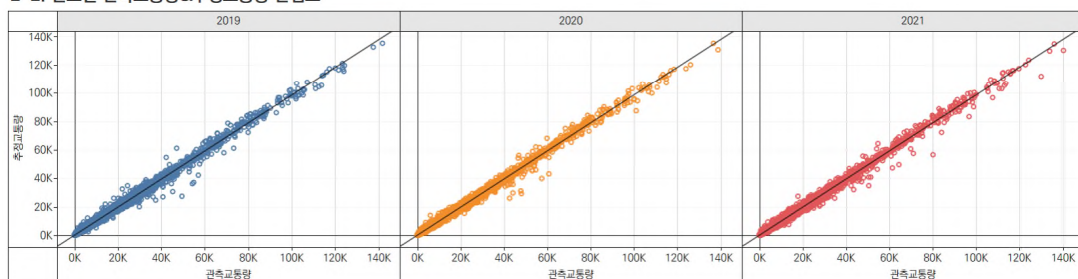
구분	신뢰도 평가대상 관측교통량 지점 수	MAPE(%)	MAE(%)	상관계수	결정계수
고속도로	1,347	7.3	-4.2	0.9904	0.9808
일반국도	3,430	11.6	-6.1	0.9924	0.9850
도시고속도로	60	5.1	-1.8	0.9987	0.9973
특별광역시도	904	10.3	-5.1	0.9767	0.9767

- '20년, '21년 교통량 추정을 위해 '19년 데이터와 연도별 교통량 신뢰도 분석 수행
  - '19년 데이터는 B사의 내비게이션 데이터로 추정된 교통량이고, '20년과 '21년 데이터는 A사의 내비게이션 데이터로 추정된 교통량을 사용하였음
  - 각 연도별 입력변수는 동일하게 '19년 DTG데이터를 사용함(추후 이후년도 DTG데이터 수집하여 재추정 예정)
  - 연도별로 동일한 네트워크를 가지고 있는 전국 관측교통량 링크에 대해 신뢰도 분석을 수행하였고 약 12~13% 수준으로 크게 차이가 없었음

2-1. 연도별 MAPE

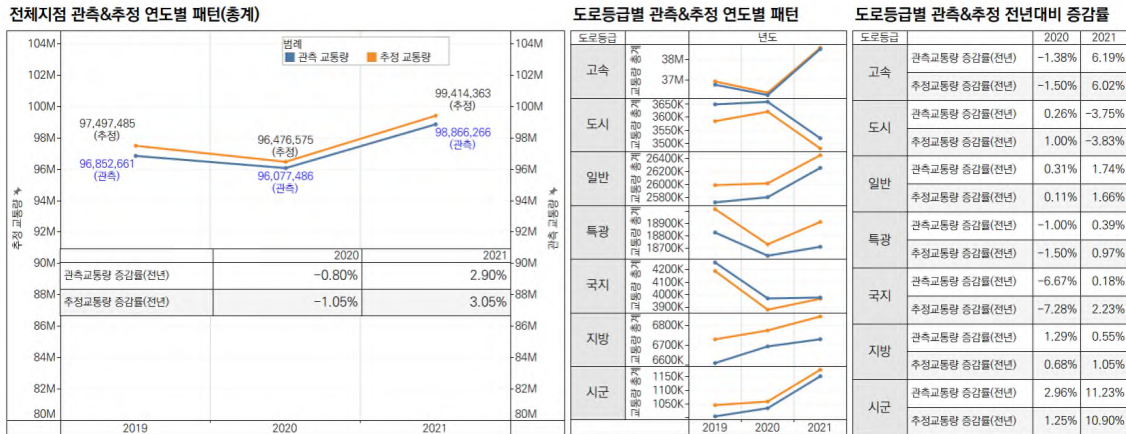
2019	2020	2021
13.5% (8,582지점)	11.8% (8,582지점)	12.1% (8,580지점)

2-2. 연도별 관측교통량&추정교통량 산점도



<그림 2-16> 연도별 신뢰도 분석 결과

- 관측교통량과 추정교통량의 연도별 패턴은 동일하게 나타났으며, 추정교통량이 관측교통량 대비 약 0.5% 높게 추정됨을 분석하였음



<그림 2-17> 관측교통량과 추정교통량의 패턴 분석

#### 바. 추정교통량 및 평균속도 구축

- 추정교통량 구축 범위는 다음과 같음
  - 교통량 전수화를 통해 전국 2차로 이상 도로에 대한 추정교통량DB 구축
    - 구축범위는 차종(승용차, 버스, 트럭), 요일패턴별(평일, 주말), 시간(전일, 시간대별)단위 구축
    - 집계단위 : 상세도로망, 주요도로망, 행정구역(시도/시군구/읍면동)

<표 2-21> 추정교통량 DB 테이블 구성

컬럼명	Type	설명	코드	코드정보
V_link_id	Integer	상세도로망 Level6 링크ID	-	-
Week_code	Integer	요일코드	0 1	주말 평일
Time_code	Integer	시간코드	0-23	0 : 1-2시 1 : 2-3시 ... 23 : 24시-1시
B_total_veh	Double	교통량(전체)	-	-
B_auto	Double	교통량(승용차)	-	-
B_bus	Double	교통량(버스)	-	-
B_truck	Double	교통량(화물차)	-	-

- 평균속도 DB 구축 범위는 다음과 같음
  - 가공한 속도프로파일 기반으로 평균속도 DB 구축
    - 구축범위는 차종(승용차, 버스, 트럭), 요일패턴별(평일, 주말), 시간(전일, 시간대별) 구축
    - 집계단위 : 상세도로망, 주요도로망, 행정구역(시도/시군구/읍면동)

<표 2-22> 평균속도 DB 테이블 구성

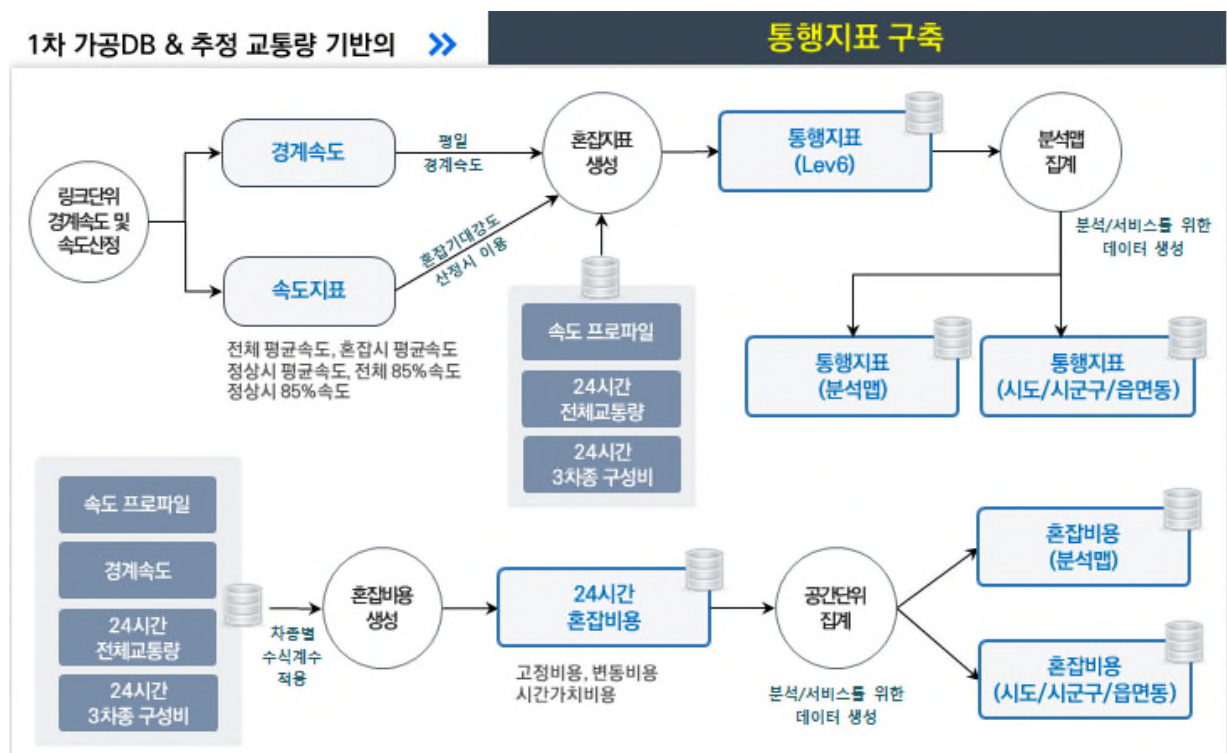
Column	Type	설명	코드	코드정보
V_link_id	Integer	가상링크ID	-	-
Week_code	Integer	요일코드	0 1	주말 평일
Time_code	Integer	시간코드	0-23	0 : 1-2시 1 : 2-3시 ... 23 : 24시-1시
Speed	Double	평균속도	-	

### 제3절 차량 통행지표 산출 및 DB구축

#### 1. 통행지표 DB 구축 및 데이터 정의

##### 가. 통행지표 DB 구축 프로세스

- 통행지표 DB는 가공된 속도프로파일과 추정된 교통량을 기반으로 전국 혼잡지표, 속도지표, 환경지표 등을 구축
- 구축 프로세스는 1. 링크단위 경계속도 산정, 2. 속도프로파일과 교통량 기반의 Level6 단위 통행지표 구축, 3. Level6 단위 통행지표 기반으로 Level5.5 단위 집계, 4. Level5.5 단위 기반으로 행정구역 단위 집계 순으로 구축



<그림 2-18> 통행지표 구축 프로세스

- 통행지표 구축을 위한 1차 가공DB 및 기초 교통DB는 다음과 같이 정의함

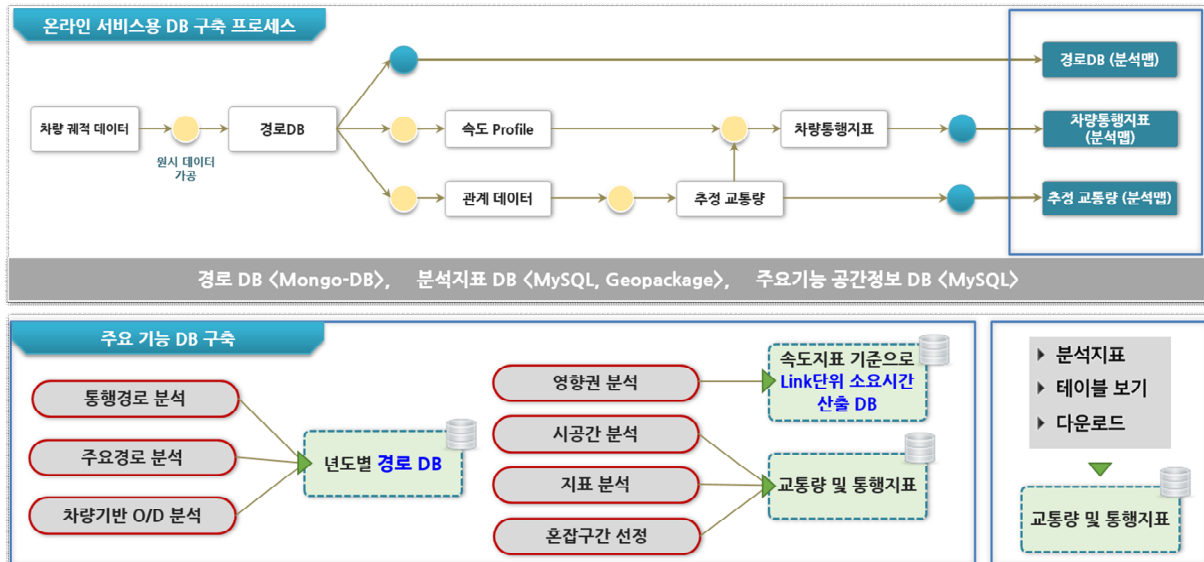
<표 2-23> 통행지표 DB 생성기준

기초교통 DB	공간정보	통행지표 DB 생성기준
<ul style="list-style-type: none"> <li>· 속도프로파일(년단위/요일별/차종별)</li> <li>· Level6 단위 추정 교통량 (년단위/요일패턴 별/시간/차종별)</li> </ul>	<ul style="list-style-type: none"> <li>· Level6 (상세 도로망)</li> <li>· Level5.5 (주요 도로망)</li> <li>· 표준노드링크</li> <li>· 행정구 (시도/시군구/읍면동)</li> </ul>	<ul style="list-style-type: none"> <li>· 공간정보 별 속도지표 및 사고지표 (년단위/요일패턴 별/시간/차종별)</li> <li>· 공간정보 별 혼잡지표 및 혼잡비용 (년단위/요일패턴 별/시간/차종별)</li> </ul>

<표 2-24> 추정교통량 데이터 테이블 구성(상세도로망 Level6 도로 단위)

컬럼명	Type	설명	코드	코드정보
상세도로망_LinkID	Integer	상세도로망Level6네트워크의 LinkID	-	-
도로등급	char	도로의 등급	101 102 103 104 105 106 107 108	고속도로 도시고속도로 일반국도 특별광역시도 국가지원지방도 지방도 시군도 고속도로 연결램프
링크길이	Double	도로구간 길이	-	-
도로명	Varchar2	도로명	-	-
시도코드	Integer	시도명을 나타내는 코드	-	-
시군구코드	Integer	시군구명을 나타내는 코드	-	-
읍면동코드	Integer	읍면동명을 나타내는 코드	-	-
시도명	string	시도의 이름	-	-
시군구명	string	시군구의 이름	-	-
읍면동명	string	읍면동의 이름	-	-
주_유형	Varchar2	평일과 주말 구분	weekday weekend	평일 주말
시간적범위	Varchar2	전일과 시간대 구분	0-23 all	0 : 24-1시 .. 23 : 23-24시 all : 연평균 1일
전체_추정교통량	Integer	모든 차량의 추정교통량	-	-
승용차_추정교통량	Integer	승용차의 추정교통량	-	-
버스_추정교통량	Integer	버스의 추정교통량		
화물차_추정교통량	Integer	화물차의 추정교통량		

## 나. 차량 온라인 서비스 DB 구축



&lt;그림 2-19&gt; 차량 온라인 서비스 제공을 위한 DB 구축 프로세스

- 온라인 서비스를 위한 교통지표는 행정구역별/도로별로 구분하며, 표/그래프/지도표출/다운로드 등으로 서비스
  - 온라인 서비스를 위한 교통지표는 교통량, 교통혼잡지표, 교통환경지표, 사고지표로 구성
  - 각 분류별 지표는 교통량지표 3종, 교통혼잡지표 6종, 교통환경지표 5종, 사고지표 2종으로 구성되며, 상세 지표별 내용은 다음 표와 같음

&lt;표 2-25&gt; 온라인 서비스 지표

교통지표		지표 설명
교통량	관측 교통량	각 기관에서 조사한 차종별 조사 교통량
	구간 추정 교통량	특정 시간대 동안 해당 도로구간을 통과하는 차량 대수
	차량 주행거리	추정 교통량 기준의 차량주행거리
교통혼잡지표	혼잡강도	전체 차량의 총 통행시간 중 교통 혼잡을 경험한 차량들의 총 통행시간 비율
	평균속도	전체 차량의 속도를 평균한 값
	혼잡시 평균속도	전체 차량 중 교통 혼잡을 경험한 차량들의 평균속도
	정상시 평균속도	전체 차량 중 교통 혼잡을 경험하지 않은 차량들의 평균속도
	지체시간	교통 혼잡으로 인하여 발생하는 차량 1대 당 평균 지체시간
	교통혼잡비용	교통 혼잡에 따른 차량 통행시간 증가로 인하여 발생하는 추가적인 사회적 손실비용(고정비, 변동비, 시간가치비용)
교통환경지표	이산화탄소배출량	특정 시간대 동안 해당 도로구간을 통과하는 차량들로 인하여 발생하는 각 교통 환경지표의 평균 배출량
	미세먼지배출량	
	일산화탄소배출량	
	휘발성 유기화합물 배출량	
	질소산화물 배출량	

○ 차량통행 분석도구 DB 구축

- 차량통행 분석도구 제공을 위한 기반 DB 구축

- 통행경로 분석/차량기반 OD분석/영향권 분석/시공간 분석/주요경로 분석/혼잡구간 분석을 위한 기반 DB 구축

- 서비스 속도 향상을 위한 경로 데이터 압축 및 NoSQL 솔루션 적용을 통해 검색 속도 향상





## **제3장 차량 GPS 데이터 기반 온라인 서비스 개선 및 현행화**

---

**제1절 View-T 시스템 개요**

**제2절 신규 개발 및 기능개선**

**제3절 통행지표 현행화**

**제4절 분석도구 현행화**

**제5절 View-T 운영 및 유지보수**

## 제3장 차량 GPS 데이터 기반 온라인 서비스 개선 및 현행화

### 제1절 View-T 시스템 개요

#### 1. View-T 시스템 구성

- View-T 시스템은 다양한 분석도구와 지표, 및 정보를 이용자에게 제공하는 서비스 기능과 서비스 운영, 모니터링 등을 위한 관리자 기능으로 구성



<그림 3-1> View-T 시스템 구성도

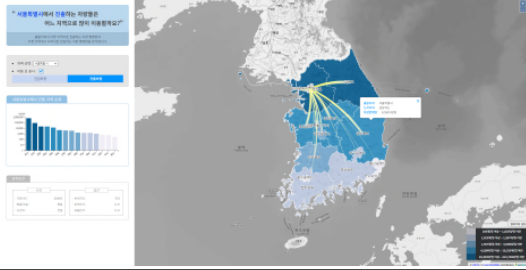

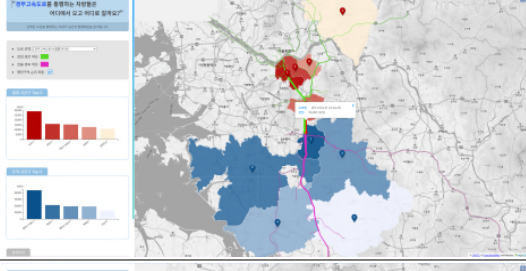
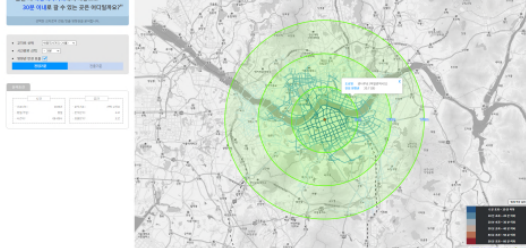
- 서비스 기능 : 교통 및 통신 데이터를 시각적으로 분석할 수 있는 Expert & Light 분석도구와 대시보드 및 분석 결과와 교통 지표 다운로드 서비스 등을 제공
- 관리자 기능 : 회원의 서비스 이용 권한을 관리하며, 회원의 서비스 이용 현황 및 다운로드 데이터 통계, 그리고 회원의 요청이나 문의에 대응하는 시스템을 제공

## 2. View-T 시스템 주요 기능 소개

### 가. View-T Light

- View-T Light는 일반 사용자들도 손쉽게 모빌리티 데이터의 시공간적 특성을 다양한 분석기능을 통해 빠르고 직관적으로 분석할 수 있게 구축한 서비스임
- View-T Light는 총 24개의 기능으로 구분되어 있으며, 기종점 통행분석, 교통 혼잡 분석 등 차량과 통신 데이터에 관련된 주제별 분석을 수행할 수 있음

<표 3-1> View-T Light 주요 기능

구분		화면예시	설명
기종점 통행분 석	기종점 차량 통행량 분석		출발지에서 다른 지역으로 진출하는 차량 통행량과 다른 지역에서 도착지로 진입하는 차량 통행량을 분석
	기종점 차량 경로 분석		출발지역에서 도착지역으로 이동하는 차량의 주요경로를 분석
교통 혼잡 분석	혼잡 도로 출도착 지 분석		선택된 도로를 통행하는 차량의 공간적 통행패턴을 분석
출·퇴근 통행분 석	출·퇴근 차량 영향권 분석		선택한 교차로의 진입/진출 영향권을 분석


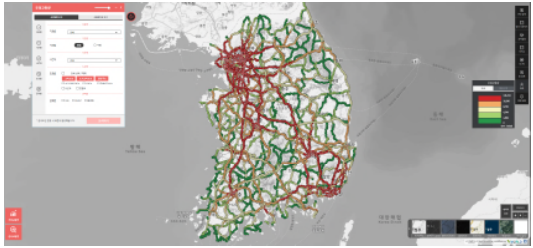
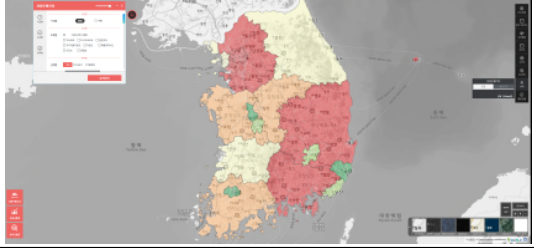
### 나. View-T Expert


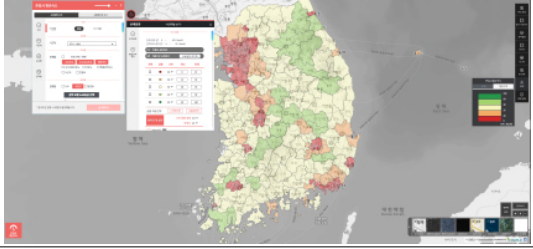

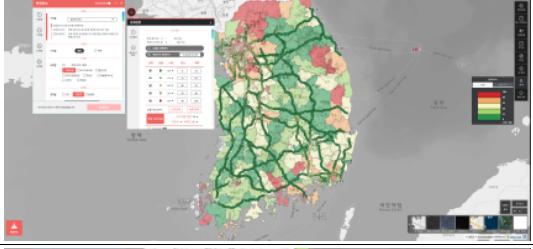
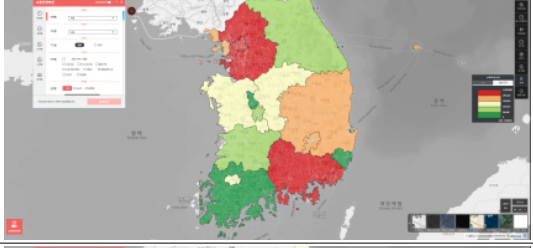
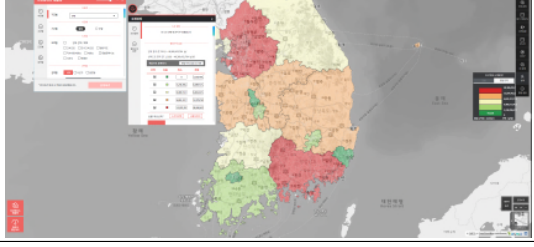
- View-T Expert는 빅데이터를 활용하여 상세 분석이 가능한 서비스로 사용자가 직접 분석 조건과 표출 설정을 통해 심도 있는 분석이 가능함
- View-T Expert는 크게 분석의 기초 교통현황을 모니터링 할 수 있는 통행지표와 모빌리티 빅데이터의 시·공간적 특성과 행태를 전국적으로 분석할 수 있는 분석도구로 구성됨

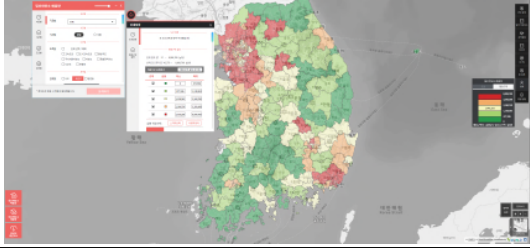
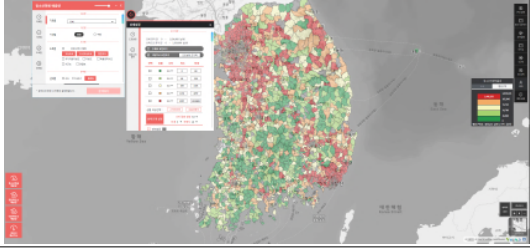

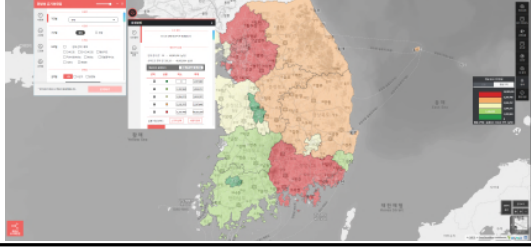
#### 1) 통행지표

- 통행지표는 모빌리티 데이터를 이용한 다양한 분석방법의 기준이 되는 표지인 기초 정보로 전국 도로망과 공간에 대한 차량 및 인구 이동 정보를 시각화하여 제공

<표 3-2> View-T Expert 통행지표

구분		화면예시	설명
교통량 지표	관측 교통량		현장조사 및 검지기를 통해 조사한 교통량
	추정 교통량		차량GPS데이터를 활용하여 미관측 도로를 대상으로 추정된 교통량
	차량 주행거리		도로를 주행하는 모든 차량들의 이동한 총 주행거리

속도 지표	평균속도		도로를 주행하는 모든 차량의 평균 속도
	혼잡시 평균속도		정상시 평균속도:혼잡을 경험하지 않은 차량들의 평균 속도
	정상시 평균속도		혼잡시 평균속도:혼잡을 경험한 차량의 평균속도
혼잡 지표	혼잡강도		혼잡시간강도:모든 차량들의 통행시간 대비 혼잡을 경험한 차량의 총 통행시간의 비율 혼잡빈도강도:모든 차량들 중 혼잡을 경험한 차량의 비율
	교통 혼잡비용		교통 혼잡으로 인해 발생하는 손실을 화폐가치로 환산한 비용
환경지 표	이산화 탄소 배출량		차량 주행시 발생하는 이산화탄소(CO2) 배출량

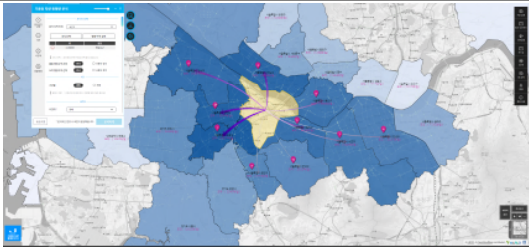
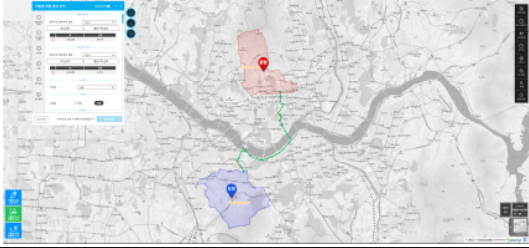
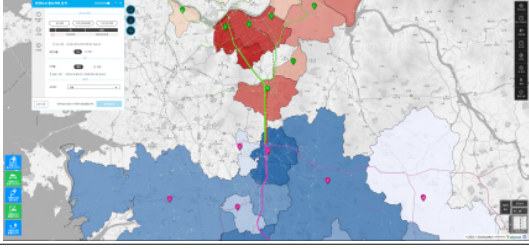

일산화 탄소 배출량		차량 주행시 발생하는 일산화탄소(CO) 배출량
질소 산화물 배출량		차량 주행시 발생하는 질소산화물(NOx) 배출량
미세먼지 배출량		차량 주행시 발생하는 미세먼지 (PM10) 배출량
휘발성 유기 화합물		차량 주행시 발생하는 휘발성 유기 화합물(VOC) 배출량



## 2) 분석도구

- 분석도구는 전문 이용자가 다양한 조건을 세밀하게 설정하여 View-T Light 보다 전문적이고 자세한 분석이 가능함
- 분석 사용자가 원하는 방식으로 시각화할 수 있는 다양한 시각화 옵션 및 데이터를 통한 분석이 필요한 다양한 데이터 다운로드 기능 제공

&lt;표 3-3&gt; View-T Expert 분석 도구 및 시뮬레이션

구분		화면예시	설명
기종점 통행 분석	기종점 차량 통행량 분석		출발지에서 다른 지역으로 진출하는 차량 통행량과 다른 지역에서 도착지로 진입하는 차량 통행량을 분석하는 기능
	기종점 차량 경로 분석		출발지역에서 도착지역으로 이동하는 차량의 주요경로를 분석하는 기능
교통 혼잡 분석	혼잡 도로 출도착 지 분석		선택된 도로를 통행하는 차량의 공간적 통행패턴을 분석하는 기능
출·퇴근 통행 분석	출퇴근 차량 영향권 분석		선택한 교차로의 진입/진출 영향권을 분석하는 기능

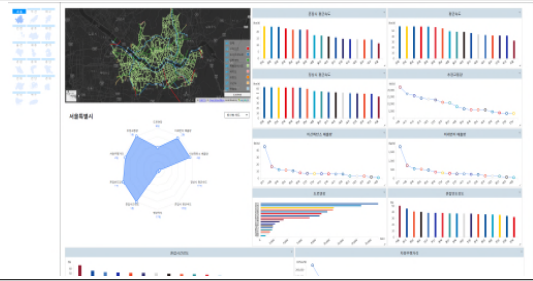
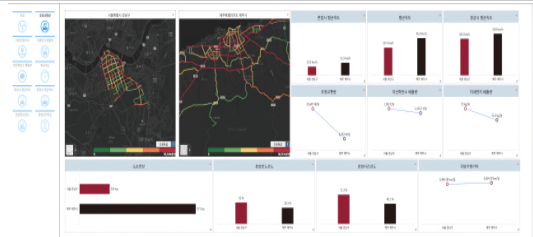



시물레이션	개별 차량 통행 시물 레이션		선택된 행정구역을 출발 또는 도착하는 개별 차량통행 시물레이션
-------	-----------------------------	---	------------------------------------

#### 다. View-T 대시보드

- View-T의 다양한 데이터를 한 눈에 알아볼 수 있도록 종합적인 형태로 제공하는 대시보드 서비스



<표 3-4> View-T 대시보드 주요 기능 설명

구분	화면예시	설명
시도별 차량 통행지표		시도별 단위로 View-T Expert에서 서비스하는 다양한 차량 통행지표를 종합적인 형태로 제공
시군구별 차량 통행지표		시군구별 단위로 View-T Expert에서 서비스하는 다양한 차량 통행지표를 종합적인 형태로 제공
연도별 차량 통행지표		연도별 단위로 View-T Expert에서 서비스하는 다양한 차량 통행지표를 종합적인 형태로 제공

#### 라. View-T 데이터 다운로드 기능

- 사용자가 직접 조건을 설정하여 필요한 View-T 데이터를 직접 수집할 수 있도록 하는 서비스

&lt;표 3-5&gt; View-T 부가 기능 설명

구분	화면예시	설명
다운로드		View-T에서 서비스하는 형상정보와 지표 정보를 사용자에게 다운로드 제공
Open API		View-T에서 사용하는 데이터를 Open API 서비스로 사용자에게 제공

## 제2절 신규 개발 및 기능개선

### 1. Faster Indicator 서비스 개발

#### 가. 개발 배경

- 기존 View-T는 현재년도 대비 2년전의 데이터가 제공되고 있어 시의성 측면의 한계가 존재하였음
- 데이터의 시의성을 증진 시키기 위해 선제적으로 구축 가능한 지표를 대상으로 데이터를 빠르게 제공하여 이용자가 시의성 있는 데이터를 활용할 수 있도록 Faster indicator 서비스를 개발함

#### 나. 기능 소개

- 2021년도 월별 전국 퍼센타일 주행속도 제공 및 사용자가 분석 조건을 설정하여 결과 테이블을 표출 및 데이터와 네트워크 정보를 다운받을 수 있음
- View-T 홈페이지 상단 메뉴에서 “다운로드 > Faster Indicator” 메뉴를 클릭시 페이지 이동

지역구분	월	속도구분	5th	10th	25th	50th	75th	90th	95th
477101000001	2021.01	일반	78	72	75	85	95	100	100
477101000002	2021.01	일반	78	72	75	85	95	100	100
477101000003	2021.01	일반	78	72	75	85	95	100	100
477101000004	2021.01	일반	78	72	75	85	95	100	100
477101000005	2021.01	일반	78	72	75	85	95	100	100
477101000006	2021.01	일반	78	72	75	85	95	100	100
477101000007	2021.01	일반	78	72	75	85	95	100	100
477101000008	2021.01	일반	78	72	75	85	95	100	100
477101000009	2021.01	일반	78	72	75	85	95	100	100
477101000010	2021.01	일반	78	72	75	85	95	100	100
477101000011	2021.01	일반	78	72	75	85	95	100	100
477101000012	2021.01	일반	78	72	75	85	95	100	100
477101000013	2021.01	일반	78	72	75	85	95	100	100
477101000014	2021.01	일반	78	72	75	85	95	100	100
477101000015	2021.01	일반	78	72	75	85	95	100	100
477101000016	2021.01	일반	78	72	75	85	95	100	100
477101000017	2021.01	일반	78	72	75	85	95	100	100
477101000018	2021.01	일반	78	72	75	85	95	100	100
477101000019	2021.01	일반	78	72	75	85	95	100	100
477101000020	2021.01	일반	78	72	75	85	95	100	100
477101000021	2021.01	일반	78	72	75	85	95	100	100
477101000022	2021.01	일반	78	72	75	85	95	100	100
477101000023	2021.01	일반	78	72	75	85	95	100	100
477101000024	2021.01	일반	78	72	75	85	95	100	100
477101000025	2021.01	일반	78	72	75	85	95	100	100
477101000026	2021.01	일반	78	72	75	85	95	100	100
477101000027	2021.01	일반	78	72	75	85	95	100	100
477101000028	2021.01	일반	78	72	75	85	95	100	100
477101000029	2021.01	일반	78	72	75	85	95	100	100
477101000030	2021.01	일반	78	72	75	85	95	100	100
477101000031	2021.01	일반	78	72	75	85	95	100	100
477101000032	2021.01	일반	78	72	75	85	95	100	100
477101000033	2021.01	일반	78	72	75	85	95	100	100
477101000034	2021.01	일반	78	72	75	85	95	100	100
477101000035	2021.01	일반	78	72	75	85	95	100	100
477101000036	2021.01	일반	78	72	75	85	95	100	100
477101000037	2021.01	일반	78	72	75	85	95	100	100
477101000038	2021.01	일반	78	72	75	85	95	100	100
477101000039	2021.01	일반	78	72	75	85	95	100	100
477101000040	2021.01	일반	78	72	75	85	95	100	100
477101000041	2021.01	일반	78	72	75	85	95	100	100
477101000042	2021.01	일반	78	72	75	85	95	100	100
477101000043	2021.01	일반	78	72	75	85	95	100	100
477101000044	2021.01	일반	78	72	75	85	95	100	100
477101000045	2021.01	일반	78	72	75	85	95	100	100
477101000046	2021.01	일반	78	72	75	85	95	100	100
477101000047	2021.01	일반	78	72	75	85	95	100	100
477101000048	2021.01	일반	78	72	75	85	95	100	100
477101000049	2021.01	일반	78	72	75	85	95	100	100
477101000050	2021.01	일반	78	72	75	85	95	100	100
477101000051	2021.01	일반	78	72	75	85	95	100	100
477101000052	2021.01	일반	78	72	75	85	95	100	100
477101000053	2021.01	일반	78	72	75	85	95	100	100
477101000054	2021.01	일반	78	72	75	85	95	100	100
477101000055	2021.01	일반	78	72	75	85	95	100	100
477101000056	2021.01	일반	78	72	75	85	95	100	100
477101000057	2021.01	일반	78	72	75	85	95	100	100
477101000058	2021.01	일반	78	72	75	85	95	100	100
477101000059	2021.01	일반	78	72	75	85	95	100	100
477101000060	2021.01	일반	78	72	75	85	95	100	100
477101000061	2021.01	일반	78	72	75	85	95	100	100
477101000062	2021.01	일반	78	72	75	85	95	100	100
477101000063	2021.01	일반	78	72	75	85	95	100	100
477101000064	2021.01	일반	78	72	75	85	95	100	100
477101000065	2021.01	일반	78	72	75	85	95	100	100
477101000066	2021.01	일반	78	72	75	85	95	100	100
477101000067	2021.01	일반	78	72	75	85	95	100	100
477101000068	2021.01	일반	78	72	75	85	95	100	100
477101000069	2021.01	일반	78	72	75	85	95	100	100
477101000070	2021.01	일반	78	72	75	85	95	100	100
477101000071	2021.01	일반	78	72	75	85	95	100	100
477101000072	2021.01	일반	78	72	75	85	95	100	100
477101000073	2021.01	일반	78	72	75	85	95	100	100
477101000074	2021.01	일반	78	72	75	85	95	100	100
477101000075	2021.01	일반	78	72	75	85	95	100	100
477101000076	2021.01	일반	78	72	75	85	95	100	100
477101000077	2021.01	일반	78	72	75	85	95	100	100
477101000078	2021.01	일반	78	72	75	85	95	100	100
477101000079	2021.01	일반	78	72	75	85	95	100	100
477101000080	2021.01	일반	78	72	75	85	95	100	100
477101000081	2021.01	일반	78	72	75	85	95	100	100
477101000082	2021.01	일반	78	72	75	85	95	100	100
477101000083	2021.01	일반	78	72	75	85	95	100	100
477101000084	2021.01	일반	78	72	75	85	95	100	100
477101000085	2021.01	일반	78	72	75	85	95	100	100
477101000086	2021.01	일반	78	72	75	85	95	100	100
477101000087	2021.01	일반	78	72	75	85	95	100	100
477101000088	2021.01	일반	78	72	75	85	95	100	100
477101000089	2021.01	일반	78	72	75	85	95	100	100
477101000090	2021.01	일반	78	72	75	85	95	100	100
477101000091	2021.01	일반	78	72	75	85	95	100	100
477101000092	2021.01	일반	78	72	75	85	95	100	100
477101000093	2021.01	일반	78	72	75	85	95	100	100
477101000094	2021.01	일반	78	72	75	85	95	100	100
477101000095	2021.01	일반	78	72	75	85	95	100	100
477101000096	2021.01	일반	78	72	75	85	95	100	100
477101000097	2021.01	일반	78	72	75	85	95	100	100
477101000098	2021.01	일반	78	72	75	85	95	100	100
477101000099	2021.01	일반	78	72	75	85	95	100	100
477101000100	2021.01	일반	78	72	75	85	95	100	100

<그림 3-2> Faster Indicator 화면

- 2021년도 월별 퍼센타일 주행 속도의 선택된 조건을 기반으로 결과값을 테이블로 표현
  - 도로의 15%, 25%, 30%, 50%, 75%, 85% 주행속도 분석 데이터 제공
- 2021년도 월별 퍼센타일 주행 속도의 분석 조건을 변경할 수 있는 UI 제공
  - 퍼센타일 주행속도의 연월, 기간, 주요시간대, 권역별, 도로등급 설정할 수 있도록 함

<그림 3-3> Faster Indicator 분석 조건 화면


- 사용자가 변경한 분석조건을 적용할 수 있도록 상세 조회 버튼으로 안내함
  - 변경된 분석조건 적용 완료 시 데이터, 네트워크 다운로드 버튼을 활성화 하여 CSV, Geopackage로 내보내기

<그림 3-4> Faster Indicator 조회 및 다운로드 버튼

- 데이터, 네트워크 다운로드 클릭 시 분석도구 사용빈도 및 활용 목적 등을 관리자가 로그를 확인할 수 있도록 신청서 작성할 수 있도록 함



- 

- 설정한 분석조건을 기반으로 테이블 형태로 분석 결과값을 표출
  -  버튼을 이용한 분석결과 테이블 최소 최대화 적용
  - 테이블 컬럼을 클릭하여 해당 값의 오름, 내림차순으로 변경 가능하도록 함

Factor Indicator										
▼ <span>SELECT</span> <span>MORE CITIES</span> <span>VIEW ALL CITIES</span>										
Percentage (year 2018)										
Age group	AGE	WGS / 1000	Population	Gender	City	City Code	Country	City (WGS) Score	City (WGS) Score	City (WGS) Score
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	71	71
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	70	69
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	68	68
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	66	66
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	65	65
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	64	64
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	63	63
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	62	62
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	61	61
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	60	60
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	59	59
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	58	58
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	57	57
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	56	56
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	55	55
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	54	54
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	53	53
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	52	52
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	51	51
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	50	50
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	49	49
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	48	48
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	47	47
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	46	46
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	45	45
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	44	44
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	43	43
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	42	42
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	41	41
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	40	40
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	39	39
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	38	38
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	37	37
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	36	36
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	35	35
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	34	34
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	33	33
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	32	32
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	31	31
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	30	30
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	29	29
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	28	28
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	27	27
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	26	26
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	25	25
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	24	24
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	23	23
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	22	22
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	21	21
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	20	20
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	19	19
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	18	18
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	17	17
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	16	16
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	15	15
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	14	14
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	13	13
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	12	12
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	11	11
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	10	10
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	9	9
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	8	8
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	7	7
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	6	6
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	5	5
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	4	4
0-14 years old	0-14	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	3	3
15-64 years old	15-64	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	2	2
65+ years old	65+	WGS	7	2018.0	0-14 years old	0-14	0-14	0-14	1	1

<그림 3-7> Faster Indicator 분석 결과 화면

## 다. 데이터 테이블 명세서

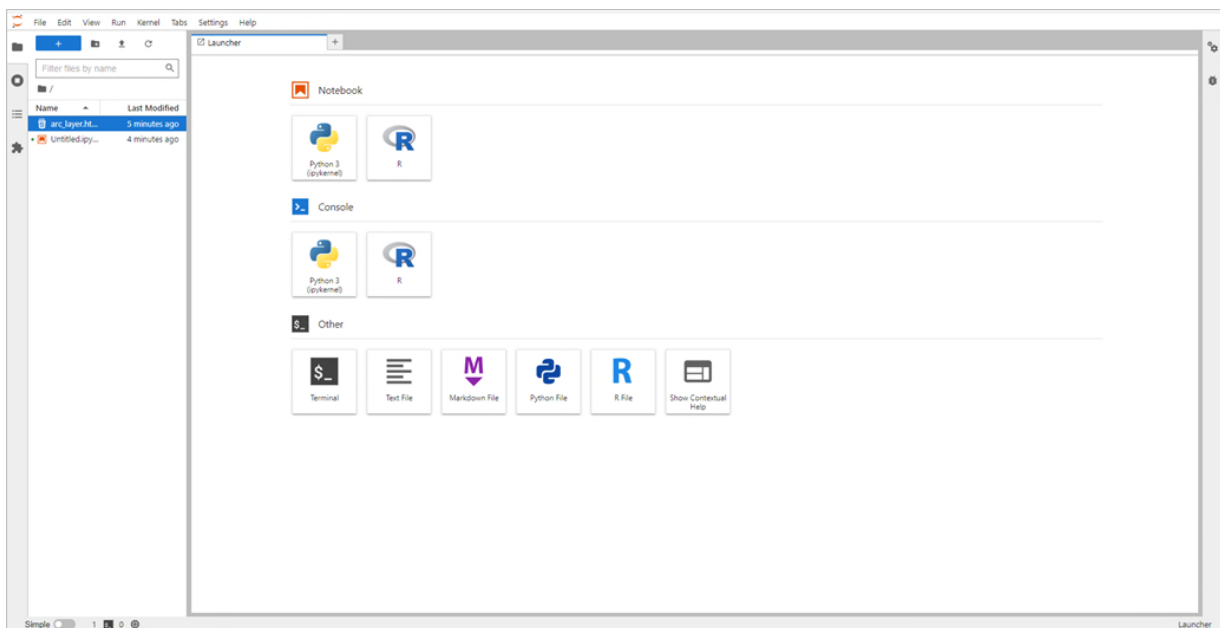
&lt;표 3-6&gt; Faster Indicator 데이터 테이블 정의

No	Column	설명	Type	자리수	코드정보
1	v_link_id	가상링크 ID	bigint (20)	0	
2	date	날짜	int (6)	0	연월 (000000)
3	week_code	요일코드	int (1)	0	0 : 주말 1 : 평일
4	peak_time	주요시간대	int (2)	0	0 : 오전첨두 (7~9시) 1 : 낮 (11~13시) 2 : 오후첨두 (17~19시)
5	road_rank	도로등급	int (3)		
6	sido_id	시도ID	int (7)		
7	sigungu_id	시군구ID	int (7)		
8	emd_id	읍면동ID	int (7)		
9	speed_15th	15퍼센타일 속도	double		15 백분위 주행속도
10	speed_25th	25퍼센타일 속도	double		25 백분위 주행속도
11	speed_30th	30퍼센타일 속도	double		35 백분위 주행속도
12	speed_50th	50퍼센타일 속도	double		50 백분위 주행속도
13	speed_75th	75퍼센타일 속도	double		75 백분위 주행속도
14	speed_85th	85퍼센타일 속도	double		85 백분위 주행속도
15	speed_avg	평균 속도	double		링크 통행속도 평균 , km/h
16	speed_sd	속도 표준편차	double		링크 통행속도 표준편차, km/h
17	speed_max	최대 속도	double		링크 최대 통행속도, km/h

## 2. 오픈소스 활용 서비스 개발

### 가. 기능 소개

- 이용자가 별도의 분석/시각화 소프트웨어 설치가 없어도 View-T에서 데이터를 다운로드 받아 즉각적으로 분석/시각화 할 수 있는 기반 분석환경을 제공하기 위해 오픈소스 기반 분석환경 개발
- 오픈소스 소프트웨어(JupyterHub, JupyterLab)를 이용하여 데이터 시각화 및 간단한 분석을 할 수 있는 기능
- 다양한 언어의 프로그램을 별도의 설치 없이 오픈소스 소프트웨어를 이용하여 활용할 수 있도록 제공함
- R, Python 등 데이터 시각화 툴을 사용할 수 있도록 제공
- 관리자가 등록한 사용자의 개별적으로 작업할 수 있는 공간 제공



<그림 3-8> 오픈소스 활용 서비스 오픈소스 소프트웨어 화면

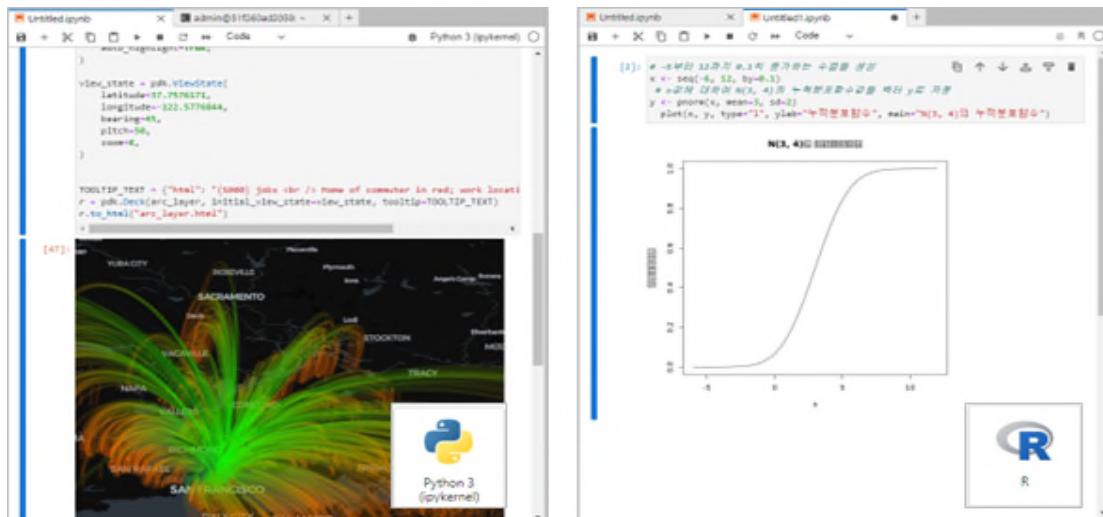
### 나. 기능 설명

- 등록된 커널( R, Python )을 이용하여 코드를 직접 작성 및 실행할 수 있도록 함
- python의 에디트 모드, 커맨드 모드, 편집 모드, 셀 경계 인터페이스를 제공하여 사용



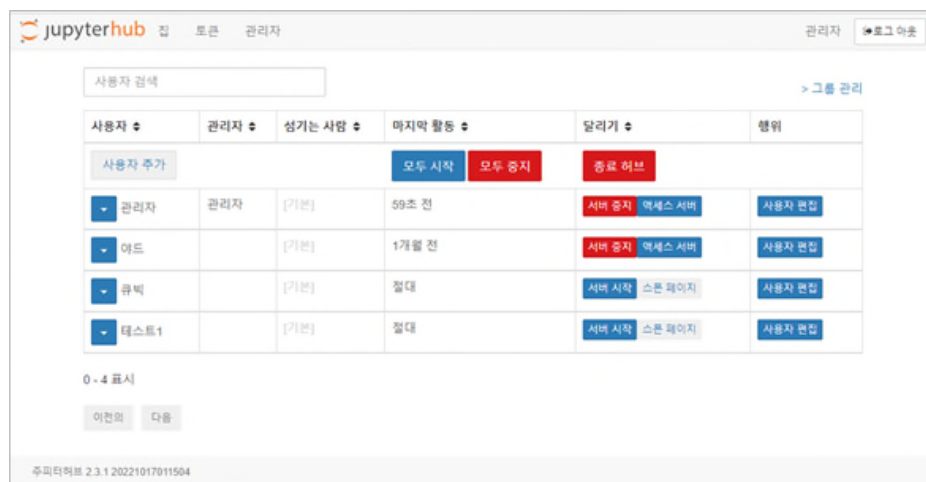
자에게 쉽게 코드를 작성할 수 있도록 함

- 사용자 편의를 위한 단축키를 제공함
- Python, R의 코드를 입력 및 실행하여 코드의 결과물을 웹 화면에서 확인 가능 하도록 함



<그림 3-9> Python, R 프로그램 언어 활용 화면

- 사용자 관리 페이지에서 사용자를 등록, 삭제, 제한 등 권한에 대해서 정의할 수 있도록 함
- 사용자의 서버를 관리자가 관리할 수 있도록 재시작, 종료, 시작을 제공함
- 회원 가입한 사용자의 승인 및 거절을 선택할 수 있도록 함



<그림 3-10> 오픈소스 소프트웨어 사용자 관리 화면

### 3. 데이터 다운로드 현황 분석 서비스 개발

#### 가. 기능 소개

- View-T에서 제공하는 Faster Indicator, 데이터 다운로드 기능(차량 데이터, 형상정보 데이터)을 사용자의 이용 정보를 확인할 수 있는 관리자 기능 개발함
- 관리자 기능의 데이터 다운로드 현황 페이지는 기능별 다운로드 횟수 및 다운로드 서비스 이용자의 소속 정보를 확인할 수 있도록 개발함

기능명	공간적 범위	시간적 범위	이용자	이용 횟수	이용 일자	이용 시간
PERCENTILE SPEED	전국	전체	이용자	1	2023.01.01	10:00:00
PERCENTILE SPEED	전국	전체	이용자	1	2023.01.01	10:00:00
PERCENTILE SPEED	전국	전체	이용자	1	2023.01.01	10:00:00
PERCENTILE SPEED	전국	전체	이용자	1	2023.01.01	10:00:00
PERCENTILE SPEED	전국	전체	이용자	1	2023.01.01	10:00:00
PERCENTILE SPEED	전국	전체	이용자	1	2023.01.01	10:00:00

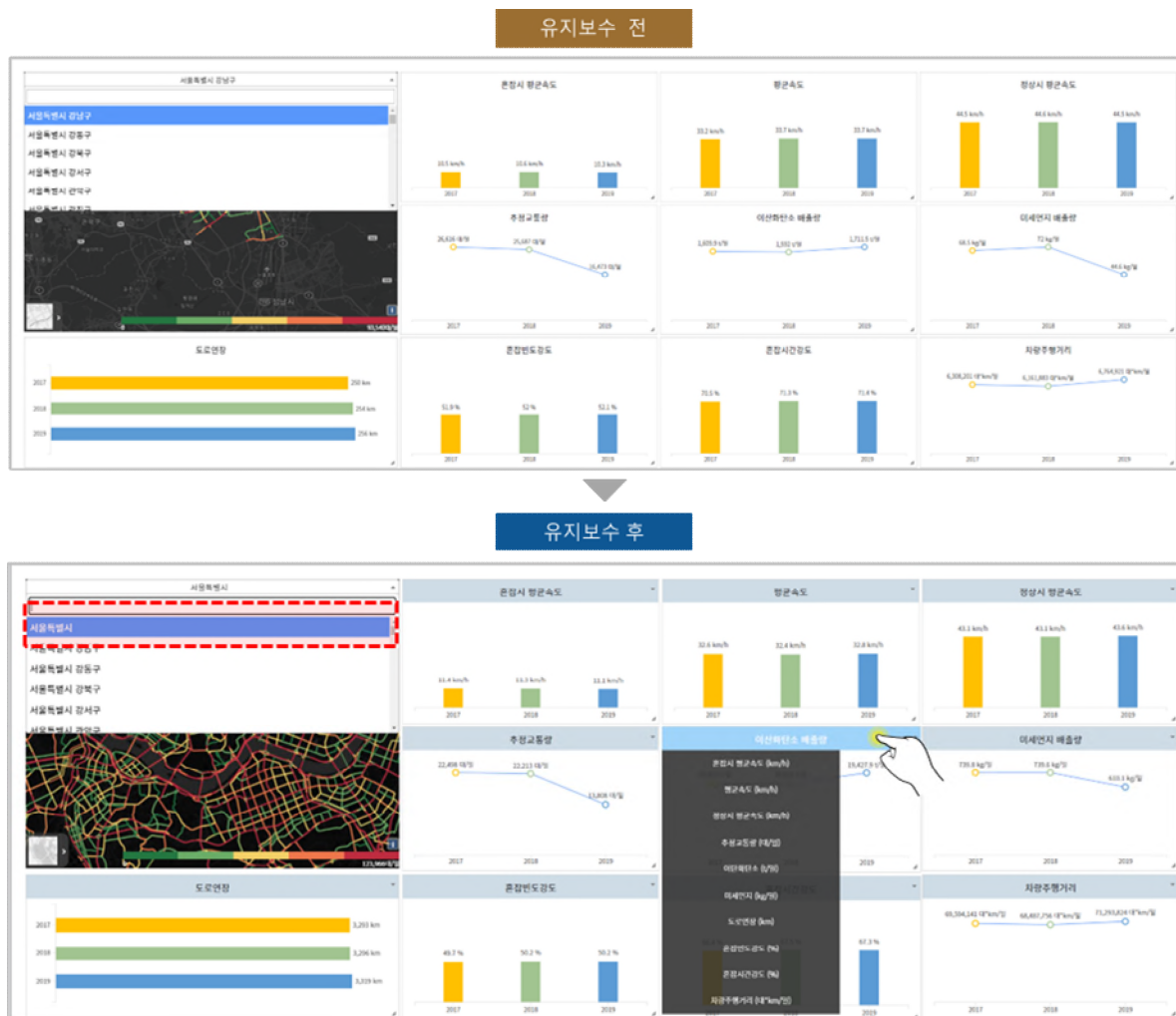
<그림 3-11> 관리자 다운로드 현황 화면

#### 나. 기능 설명

- 관리자 페이지에서 '데이터 다운로드 현황' 버튼을 클릭하여 페이지 이동
- 데이터 다운로드 현황 또는 사용자 이용 현황을 클릭하면 해당 페이지가 호출
- 데이터 다운로드 현황은 '요청 데이터명, 공간적 범위, 시간적 범위'가 같은 내역 정보의 다운로드 횟수의 합계를 호출
- 사용자 이용 현황은 '소속기관, 데이터 활용 목적, 데이터 구분'이 같은 내역 정보의 다운로드 횟수의 합계를 호출

#### 4. 분석도구 차량 대시보드 기능 개선

- 시도별, 시군구별, 연도별 차량 통행지표 대시보드 유지보수
  - 그래프와 주제를 구분하기 쉽도록 색상의 차이를 주어 가독성 및 가시성을 높임
  - 주제부는 그래프의 주제를 알림과 동시에 주제 변경 기능을 가진 메뉴 버튼을 쉽게 이해할 수 있도록 화살표 아이콘을 추가



<그림 3-12> 분석도구 차량 유지보수 전후 비교 화면

- 연도별 차량 대시보드는 기존에 시군구 단위로 서비스를 제공하였으나 시도 단위의 분석을 추가하여 분석의 폭을 넓힘

## 5. View-T 회원 시스템 개선

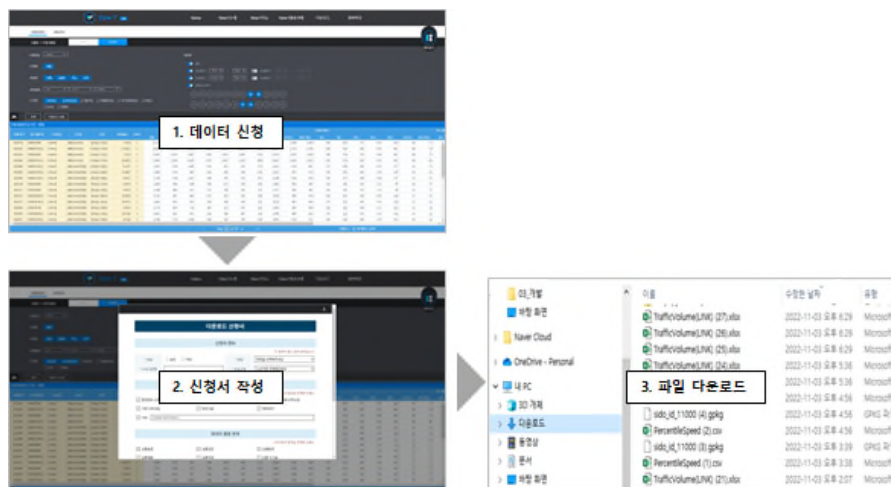
- 홈페이지 회원 관련 UI 개선
  - 회원정보의 수집을 최소화 하여 사용자 편의성 개선



<그림 3-13> 홈페이지 로그인, MyPage 삭제 화면

## 6. 데이터 다운받기 기능 개선

- 데이터 다운받기 단계 축소
  - 기존 단계는 데이터 다운로드 요청 -> 신청서 작성 -> 관리자 데이터 제공 수락 또는 거절 -> MyPage의 데이터 신청 상황 표시 -> 수락된 데이터 다운로드
  - 변경된 단계는 데이터 다운로드 선택 -> 신청서 작성 -> 데이터 다운로드



<그림 3-14> 데이터 다운로드 단계 축소

## 7. 분석도구 메뉴 통합 및 메뉴 위치 변경

- 분석도구 통합 및 메뉴 위치 변경

기존 메뉴 구조도		개선 메뉴 구조도	
통행지표	분석도구	통행지표	분석도구
관측 교통량	기종점 차량 통행량 분석	관측 교통량	기종점 차량 통행량 분석
추정 교통량	기종점 인구 통행량 분석	추정 교통량	기종점 인구 통행량 분석
차량 주행거리	기종점 차량 경로 분석	차량 주행거리	기종점 차량 경로 분석
평균속도	기종점 인구 통행구간 분석	평균속도	기종점 인구 통행구간 분석
혼잡시 평균속도	혼잡도로 선정	혼잡시 평균속도	혼잡도로 선정
정상시 평균속도	시공간 혼잡 분석	정상시 평균속도	시공간 혼잡 분석
혼잡강도	혼잡도로 출도착지 분석	혼잡강도	혼잡도로 출도착지 분석
교통 혼잡비용	핫플레이스 분석	교통 혼잡비용	핫플레이스 분석 <b>통합</b>
이산화탄소 배출량	고령자 통행비중 분석	이산화탄소 배출량	고령자 통행비중 분석
일산화탄소 배출량	고령자 주요 통행구간 분석	일산화탄소 배출량	고령자 주요 통행구간 분석
질소산화물 배출량	도시공동화 심각도 분석	질소산화물 배출량	출퇴근 차량 영향권 분석
미세먼지 배출량	출퇴근 차량 영향권 분석	미세먼지 배출량	연휴기간 통행 분석
휘발성 유기화합물	출퇴근 통행 접근성 분석	휘발성 유기화합물	내외부 통행 비율 분석
	경제활동 의존도 분석	도시공동화 심각도 분석	이동
	경제활동 지체 수용도 분석	경제활동 의존도 분석	이동
	관광 특화 지역 분석	경제활동 지체 수용도 분석	이동
	계절별 핫플레이스	통근통학 시간 분석	이동 기존 명칭 '출퇴근 접근성 분석'
	연휴기간 통행 분석		
	내외부 통행 비율 분석		

<그림 3-15> 메뉴구조도

## 10. Open API 연계 서비스 개발

- 외부 Track 홈페이지와의 데이터 연계를 통한 추가적인 API 제공



<그림 3-16> Track API 제공



## 제3절 통행지표 현황화

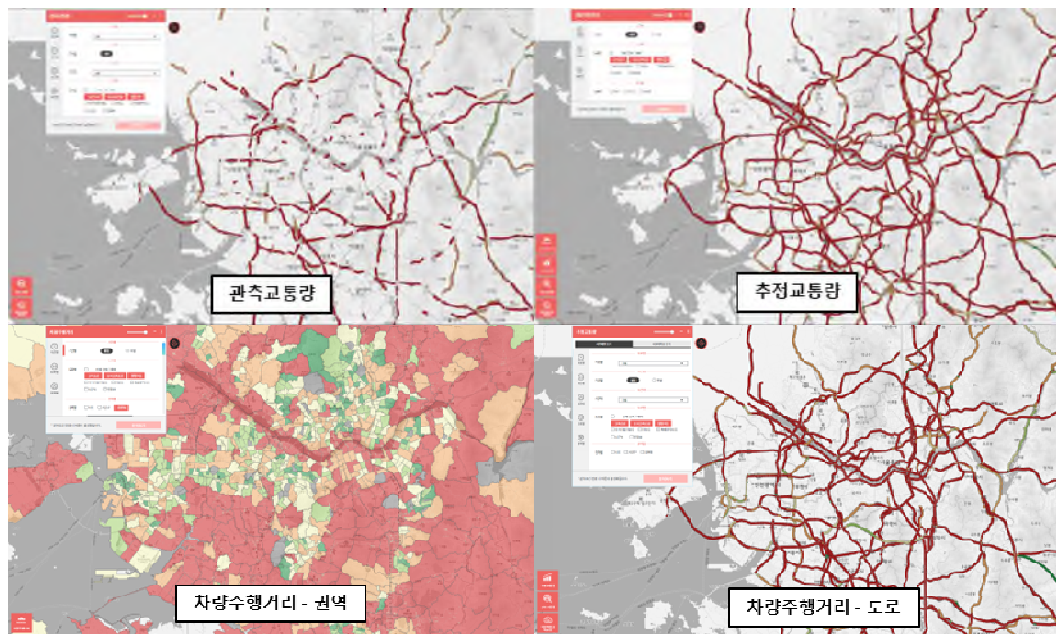
### 1. 교통량 지표 현황화

#### 가. 지표의 종류 및 설명

- 관측교통량 : 현장조사 및 검지기를 통해 조사한 교통량
  - 차종별(버스/승용차/트럭), 기간별(평일/주말), 시간대별, 도로별
- 추정교통량: 차량 GPS 데이터를 활용하여 미관측 도로를 대상으로 추정된 교통량
  - 차종별(버스/승용차/트럭), 기간별(평일/주말), 시간대별, 도로별, 권역별
- 차량주행거리: 도로를 주행하는 모든 차량들의 이동한 총 주행거리
  - 기간별(평일/주말), 도로별, 권역별

#### 나. 현황화 내용

- 2020년 데이터 현황화



<그림 3-17> 교통량 지표 화면예시

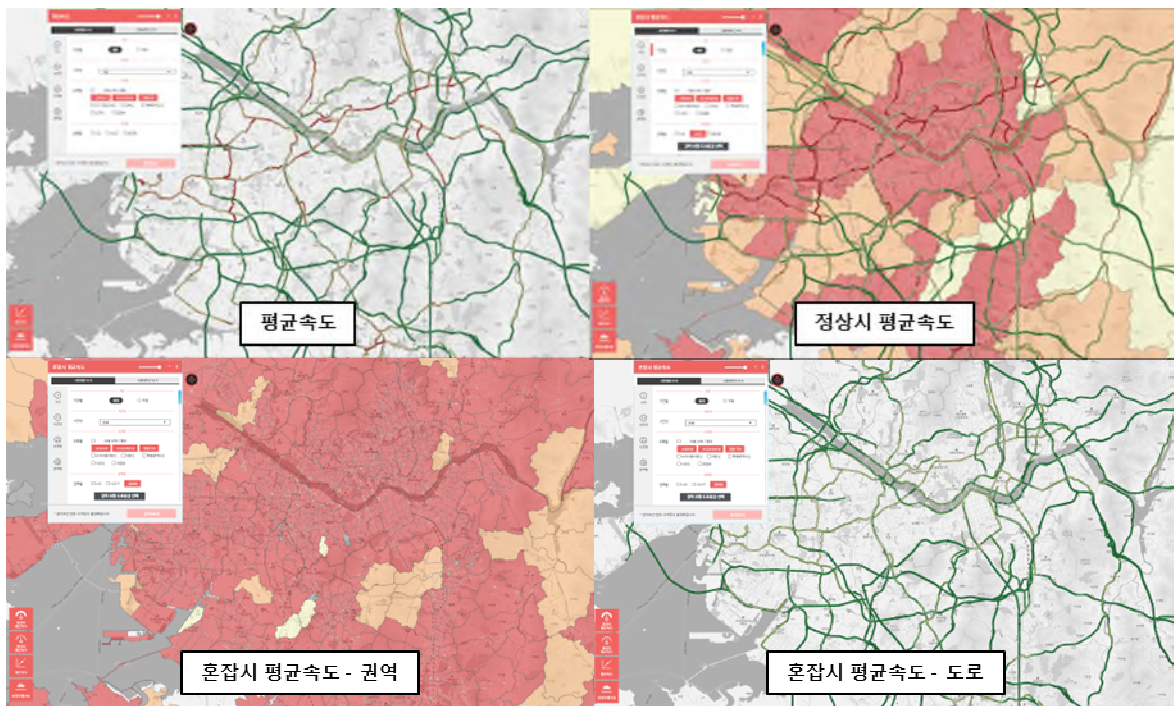
## 2. 속도 지표 현행화

### 가. 지표의 종류 및 설명

- 평균속도 : 도로를 주행하는 모든 평균 차량들의 속도를 평균한 것
  - 기간별(평일/주말), 시간대별, 도로별, 권역별
- 정상시 평균속도 : 혼잡을 경험하지 않은 차량들의 속도를 평균한 것
  - 기간별(평일/주말), 시간대별, 도로별, 권역별
- 혼잡시 평균속도 : 혼잡을 경험한 차량들의 속도를 평균한 것
  - 기간별(평일/주말), 시간대별, 도로별, 권역별

### 나. 현행화 내용

- 2020년 데이터 현행화



<그림 3-18> 속도 지표 화면예시

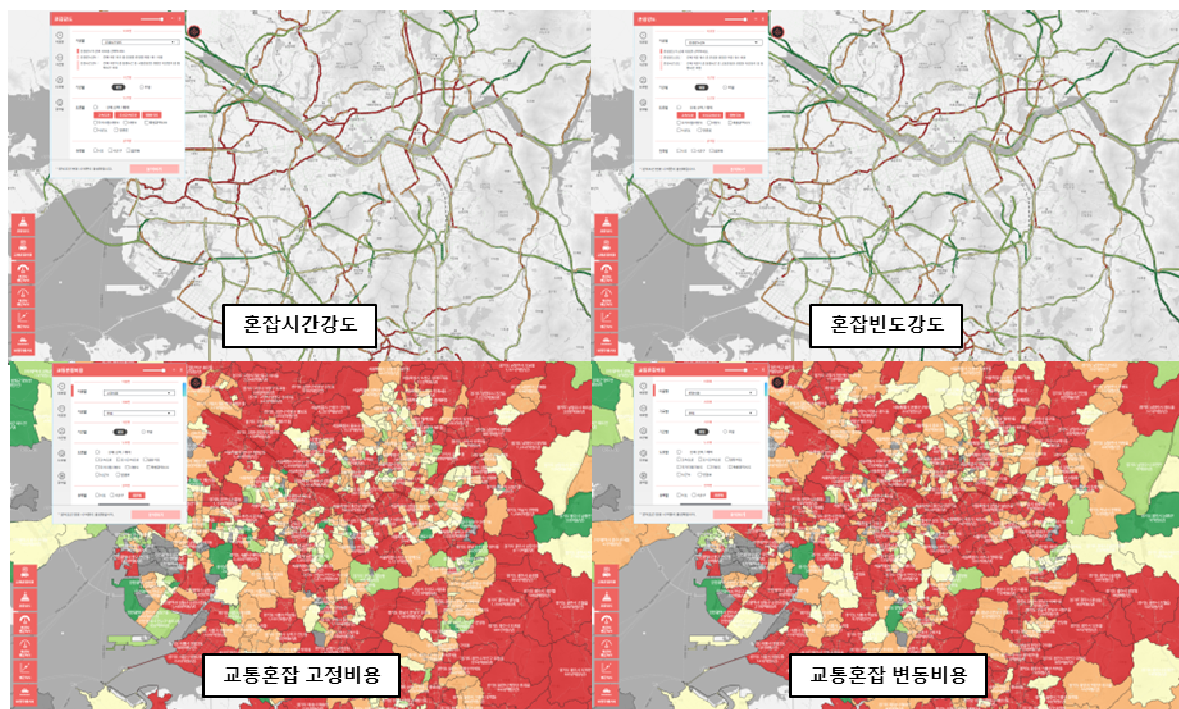
### 3. 혼잡 지표 현행화

#### 가. 지표의 종류 및 설명

- 혼잡시간강도 : 모든 차량들의 총 통행시간 대비 혼잡을 경험한 차량의 총 통행시간 비율
  - 기간별(평일/주말), 도로별, 권역별
- 혼잡빈도강도 : 모든 차량들 중 혼잡을 경험한 차량의 비율
  - 기간별(평일/주말), 도로별, 권역별
- 교통혼잡비용 : 교통혼잡으로 인해 발생하는 손실을 화폐가치로 환산한 비용
  - 비용(고정비용, 변동비용, 시간가치비용), 차종별(버스/승용차/트럭), 기간(평일/주말)별, 도로별, 권역별

#### 나. 현행화 내용

- 2020년 데이터 현행화



<그림 3-19> 혼잡 지표 화면예시



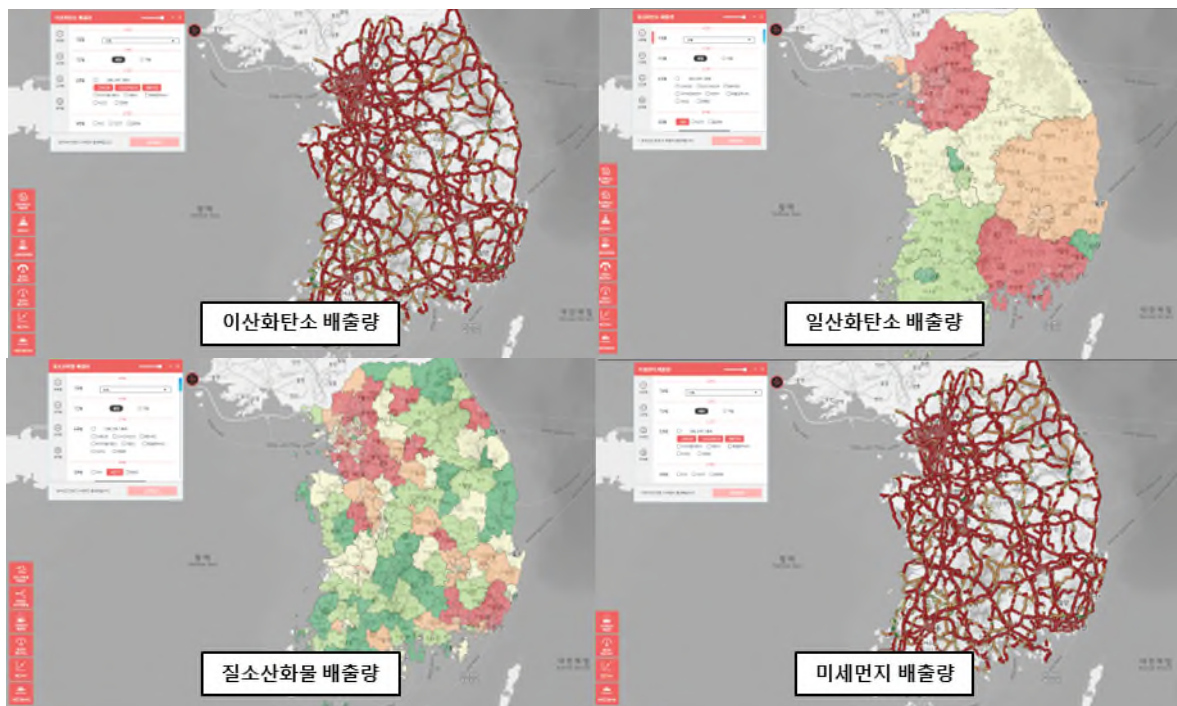
#### 4. 환경 지표 현행화

##### 가. 지표의 종류 및 설명

- 이산화탄소 배출량 : 차량 주행시 발생하는 이산화탄소 배출량
- 미세먼지 배출량 : 차량 주행시 발생하는 미세먼지 배출량
- 일산화탄소 배출량 : 차량 주행시 발생하는 일산화탄소 배출량
- 휘발성 유기 화합물 배출량 : 차량 주행시 발생하는 휘발성 유기 화합물 배출량
- 질소산화물 배출량 : 차량 주행시 발생하는 질소산화물 배출량
  - (공통)차종별(버스/승용차/트럭), 기간별(평일/주말), 도로별, 권역별

##### 나. 현행화 내용

- 2020년 데이터 현행화



<그림 3-20> 환경 지표 화면예시

## 제4절 분석도구 현행화

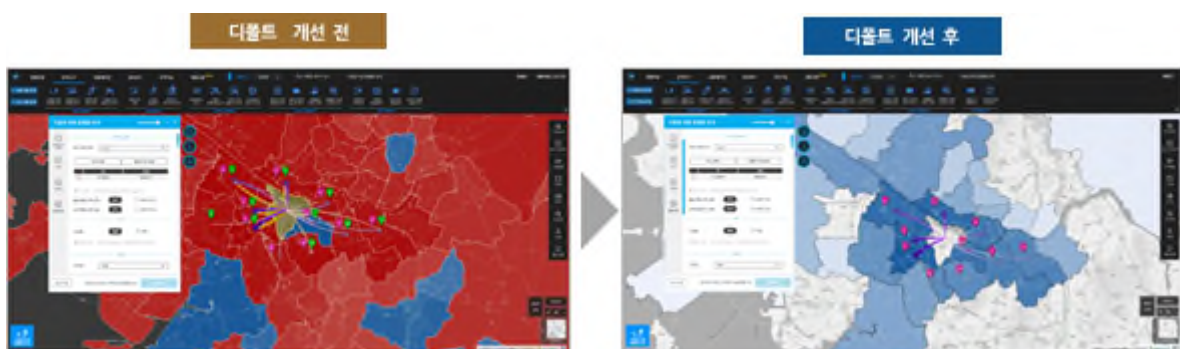
### 1. 기종점 차량 통행량 분석도구 현행화

#### 가. 기능 소개

- 출발지에서 각 행정구역으로 가는 차량 통행량과 각 행정구역에서 도착지로 오는 차량 통행량을 분석하는 기능
- 선택한 행정구역(시도, 시군구, 읍면동)단위를 출발지나 도착지로 설정하여 출발지(도착지)에서 도착지(출발지)로 이동하는 차량의 통행량을 분석함
  - 분석지역 단위: 시도, 시군구, 읍면동
  - 선택 진입 행정구역 : 전국, 사용자 정의
  - 선택 진출 행정구역 : 전국, 사용자 정의
  - 기간 및 시간대 : 평일/주말, 전일, 24시간
  - 표출 형태 : 시군구, 읍면동

#### 나. 현행화 내용

- 초기화면 주제도 표출 기본값 설정 변경
  - 행정구역 출 도착 변경
  - 행정구역 테두리 색상 변경
- 2020년 데이터 현행화



<그림 3-21> 기종점 차량 통행량 분석 디폴트 변경 화면

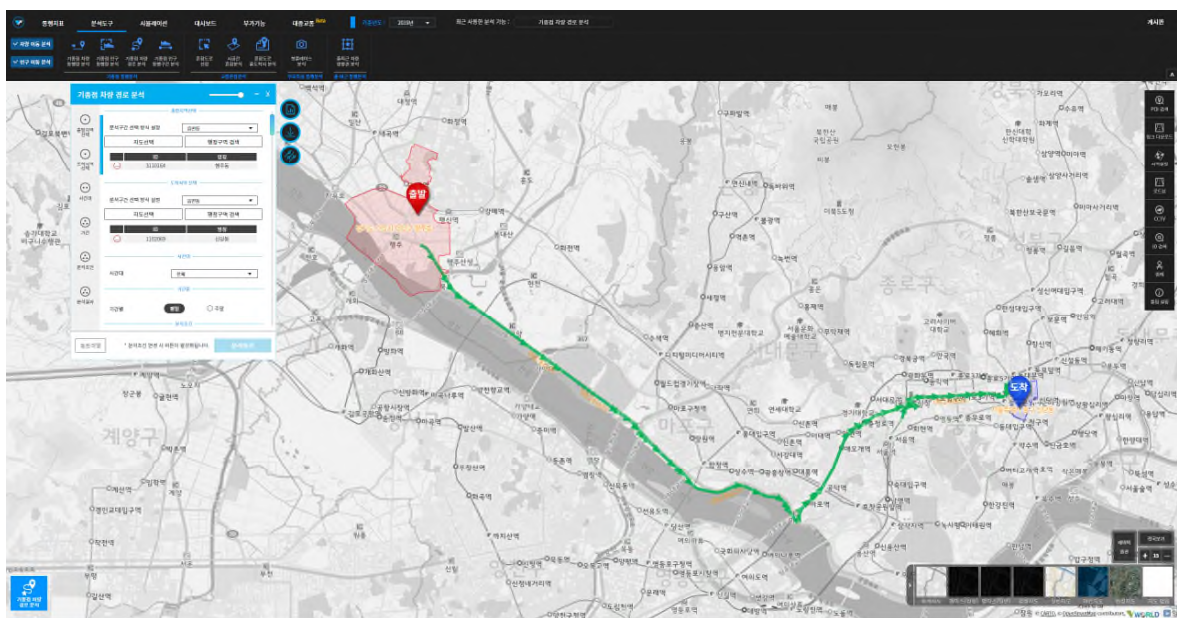
## 2. 기종점 차량 경로 분석도구 현행화

### 가. 기능 소개

- 출발지에서 도착지로 이동하는 차량의 주요경로를 분석하는 기능
- 선택한 행정구역(시도, 시군구, 읍면동)단위를 출발지나 도착지로 설정하여 출발지(도착지)에서 도착지(출발지)로 이동하는 차량의 통행량을 분석함
  - 분석지역 단위: 시도, 시군구, 읍면동
  - 선택 진입 행정구역 : 전국, 사용자 정의
  - 선택 진출 행정구역 : 전국, 사용자 정의
  - 기간 및 시간대 : 평일/주말, 전일, 24시간
  - 표출 형태 : 시군구, 읍면동

### 나. 현행화 내용

- 2020년 데이터 현행화 및 유지보수



<그림 3-22> 기종점 차량 경로 분석도구 화면예시

### 3. 혼잡도로 선정 분석도구 현행화

#### 가. 기능 소개

- 이용자가 정의하는 혼잡의 기준에 따른 혼잡도로 선정 분석하는 기능
- 선택한 출발지와 도착지로 이동하는 차량 비율로 경로를 순위별, 누적별로 분석함
  - 선택 출발지 : 도로, 행정구역(시군구, 읍면동), 사용자 정의
  - 선택 도착지 : 도로, 행정구역(시군구, 읍면동), 사용자 정의
  - 기간 및 시간대 : 평일/주말, 전일, 24시간
  - 표출 형태 : 순위별 경로, 누적별 경로

#### 나. 현행화 내용

- 초기화면 주제도 표출 기본값 설정 변경
  - 도로 테두리 색상 변경
- 2020년 데이터 현행화 및 유지보수



<그림 3-23> 혼잡도로 선정 분석 디폴트 변경 화면



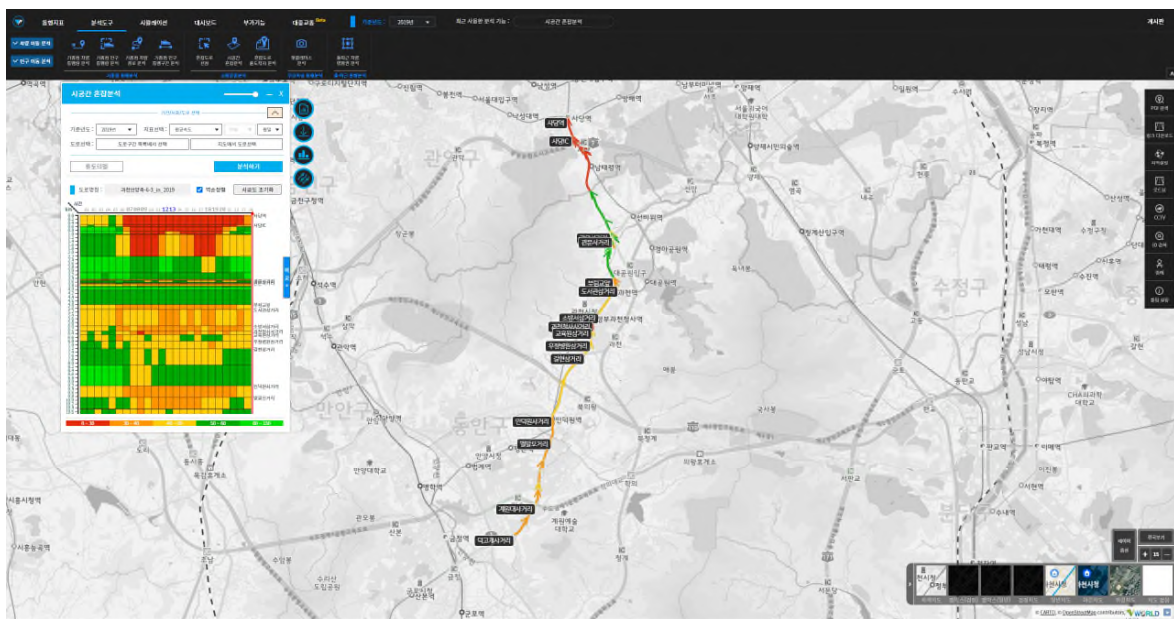
#### 4. 시·공간 혼잡 분석도구 현행화

##### 가. 기능 소개

- 도로축의 시간과 공간 변화에 따른 혼잡현황을 분석하는 기능
- 도로의 방향, 통행지표(추정교통량, 평균속도, 혼잡시평균속도, 정상시평균속도)를 선택하여 원하는 도로(축)의 시간변화에 따른 혼잡구간을 분석함
  - 기간 : 평일/주말
  - 선택 조건 : 통행지표, 도로구간(축), 도로 방향

##### 나. 현행화 내용

- 2020년 데이터 현행화 및 유지보수



<그림 3-24> 시·공간 혼잡 분석도구 화면예시

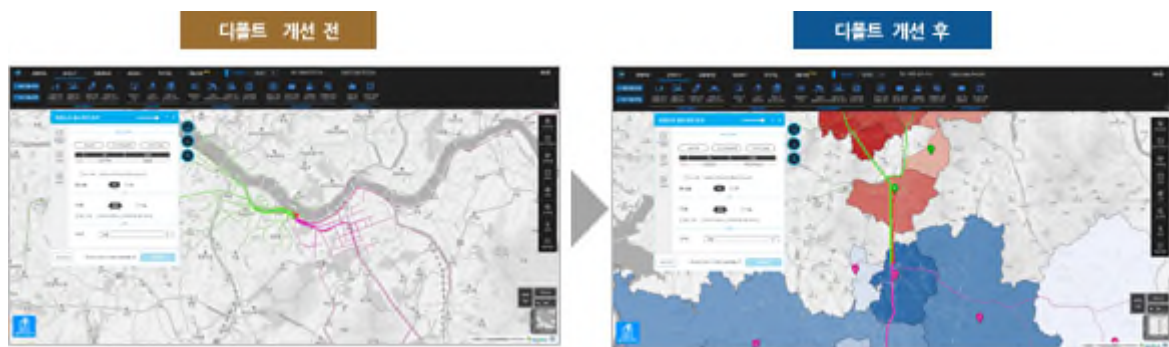
## 5. 혼잡도로 통행경로 분석도구 현행화

### 가. 기능 소개

- 선택한 도로를 통행하는 차량의 공간적 통행패턴을 분석하는 기능
- 선택한 도로를 통행하는 차량의 출발지와 도착지를 행정구역별(시군구, 읍면동) 또는 도로 단위로 통행량을 분석함
  - 기간 및 시간대 : 평일/주말, 전일, 24시간
  - 선택 공간 조건 : 유턴 on, 유턴 off

### 나. 현행화 내용

- 초기화면 주제도 표출 기본값 설정 변경
  - 행정구역 테두리 색상 변경
  - 행정구역 On 변경
- 2020년 데이터 현행화 및 유지보수



<그림 3-25> 혼잡도로 통행경로 분석도구 디폴트 변경 화면

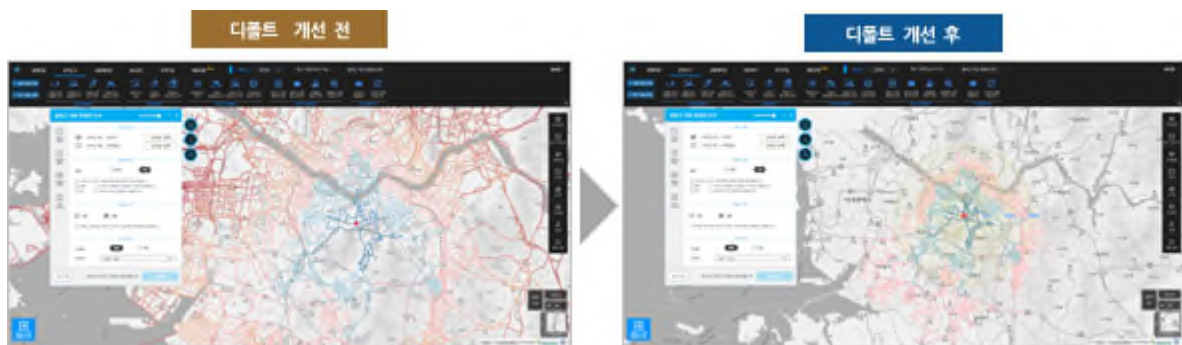
## 6. 출·퇴근 차량 영향권 분석도구 현행화

### 가. 기능 소개

- 선택한 교차로에서 1시간 이내로 오고 가는 차량의 영향권을 10분 단위 간격으로 이동 시간을 분석하는 기능
  - 기간 및 시간대 : 평일/주말, 전일, 24시간
  - 선택 공간 조건 : 중첩되는 영향권, 모든 영향권
  - 선택 시간 조건 : 10분 단위 간격

### 나. 현행화 내용

- 초기화면 주제도 표출 기본값 설정 변경
  - 영향권 표출 범위 변경
  - 반경 표출
- 2020년 데이터 현행화 및 유지보수



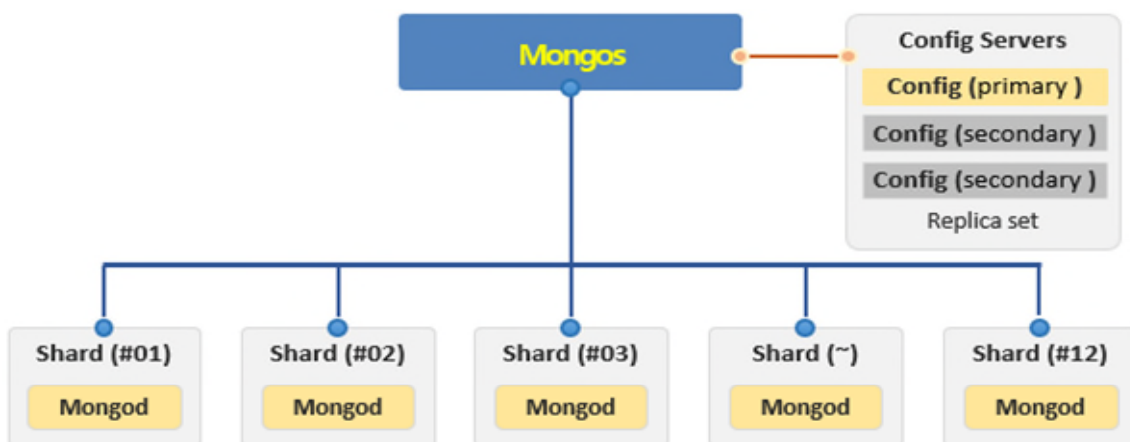
<그림 3-26> 출·퇴근 영향권 분석도구 디폴트 변경화면

## 제5절 View-T 운영 및 유지보수

### 1. 빅데이터 서비스를 위한 최적화 시스템 구성

#### 가. 대용량 통신 데이터 분산시스템 최적화

- 빅데이터 서비스를 위한 최적화의 필요성
  - 대용량의 통신 데이터를 신속하고 안전하게 서비스하기 위해서는, S/W와 H/W에 대한 최적화된 시스템 구성이 필요함
  - 대용량의 데이터를 물리적으로 한 대의 서버에 저장하고 처리하게 된다면 서버의 부하와 하드웨어적 한계가 발생하게 되므로 이를 해결하는 방안이 필요함
- 빅데이터 서비스를 위한 최적화 방안
  - 대용량의 데이터를 빠르게 처리하고 서버와 하드웨어에 부하가 가지 않도록 분산 저장하는 MongoDB의 샤딩 기술을 사용함
- MongoDB 샤딩 정의
  - 샤딩은 다음의 3가지 구성요소로 구성되며 각각의 역할이 부여됨
    - Mongos : 중계자 역할로 Application의 질의를 받아 Shard 서버의 응답을 중계
    - Config : 저장된 데이터가 Shard중에 위치 등의 Shard Meta 정보를 저장
    - Shard : 실제 데이터가 저장되는 서버



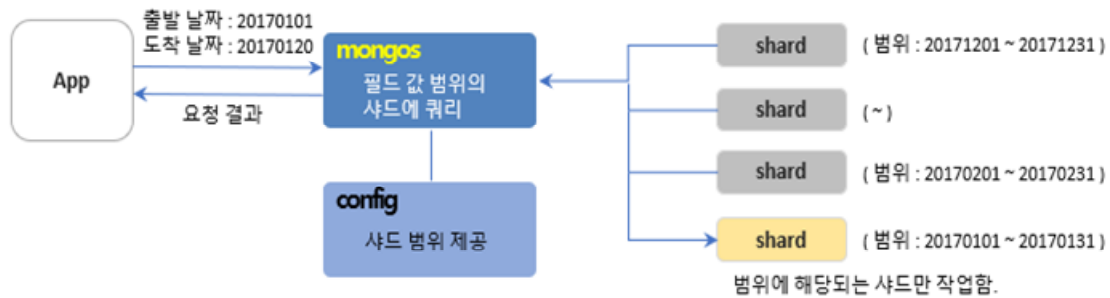
<그림 3-27> MongoDB 샤딩 구조



○ DB 특성별 샤딩 알고리즘 적용

- 행정구역이 매칭된 통신 데이터

- View-T의 서비스 중 데이터 테이블 다운받기에서 사용되며 기간의 범위로 검색하여 서비스하는 데 사용되므로 레인지 샤딩과 지역 기반 샤딩을 사용함

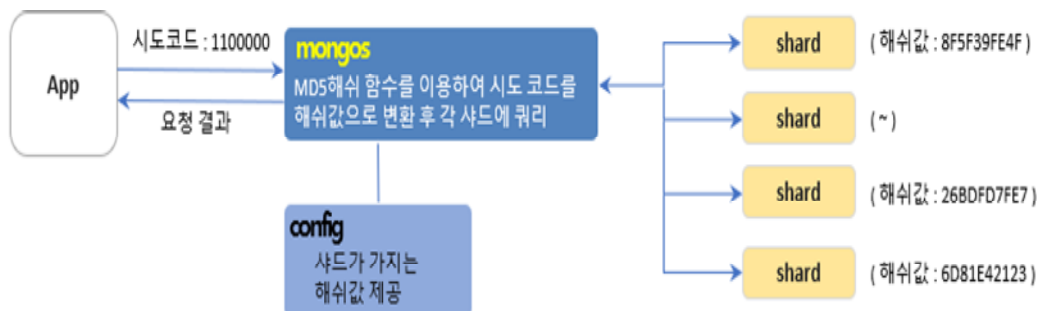


<그림 3-28> 레인지 샤딩의 구성 화면

- 레인지 샤딩을 적용하여 범위 검색 시 해당 범위에 속한 Shard에서만 작업이 이루어지기 때문에 빠르고 서버에 부하를 주지 않음
- 지역 기반 샤딩을 적용하여 월별로 Shard에 분배하기 때문에 분산 시 고르게 분배됨

- 교통폴리곤 기준 통행량, 행정구역 단위의 출발/도착 통행량

- View-T의 서비스 중 특정 행정구역을 선택 후 사용하는 주요통행지역분석, 주요통행공간분석, hotplace분석에 사용되며 행정구역으로 검색하여 분석하므로 해쉬 샤딩을 사용함



<그림 3-29> 해쉬 샤딩의 구성

- 해쉬 함수를 사용함으로 각 Shard에 데이터 분배가 고르게 되고 특정 행정구역 검색에 대한 처리가 빠르며 기준 필드들을 전부 샤드 키로 구성하여 기준 필드들로 검색하여도 신속한 처리를 할 수 있음

## 2. 운영 및 DB유지보수

### 가. 신규 구축 DB 업데이트

- 구축된 통행지표DB, 분석도구 DB를 온라인 서비스 업데이트
  - 업데이트 대상이 되는 데이터는 네트워크 데이터, 웹 분석용 테이블, 형상정보 및 경로 빅데이터 이관
  - 데이터의 적용 시스템 및 적용 내용은 다음과 같음

<표 3-7> View-T 웹서비스용 DB 데이터 업데이트 내용

대상 데이터	업데이트 대상 시스템	업데이트 내용
네트워크 데이터	RDBMS Server (MySQL)	신규 구축될 네트워크 데이터
웹 분석용 데이터	RDBMS Server (MySQL)	차량 GPS 기반 웹 분석용 데이터
형상정보	GIS Server (GeoServer)	업데이트된 차량 모빌리티 기반지도 업데이트된 사람 모빌리티 기반지도
경로 데이터	BigData Solution (MongoDB)	차량 GPS 기반 개별차량 이동 궤적 데이터 모바일 데이터 기반 웹 분석용 데이터

- DB 장애 대응 전략
  - DB에 대한 정기적인 백업 정책을 수립하여 항상 데이터 유실을 미리 방지하고자 함



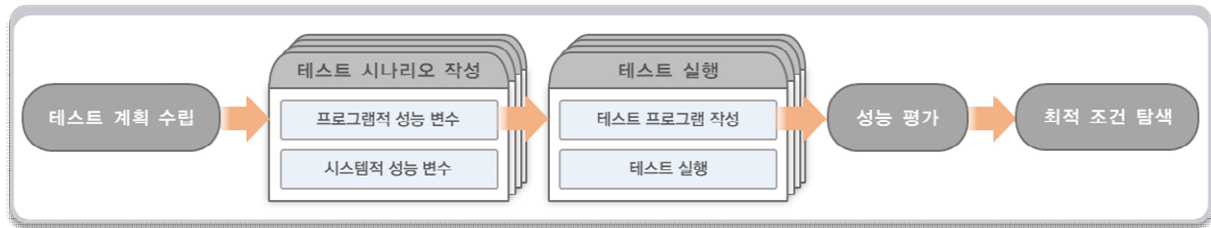
<그림 3-30> DB 장애 대응 전략

## 3. 성능 테스트

### 가. 데이터 가공 성능 테스트

- 대용량의 데이터를 빠르게 처리하고 서버와 하드웨어에 부하가 가지 않도록 분산 저

장하는 MongoDB의 샤딩 기술을 사용함



<그림 3-31> View-T 데이터 가공 성능 테스트 시나리오

- View-T 서비스에서 이용자 특성을 고려한 고급 사용자에게는 상세 분석이 가능한 교통 폴리곤 단위 집계 통행량 DB를 사용하고, 일반 사용자에게는 분석 특성이 있는 분석 특성 집계 통행량 DB를 사용하도록 설계한 내용으로 검색 평균 속도 비교 테스트를 진행
- 조건은 통행목적을 기준으로 출근, 등교, 퇴근, 하교, 점심시간, 귀가, 여가로 하여 크게 시도단위에서는 서울시와 경기도, 시군구 단위에서는 강남구로 진행한 결과 전체 평균 속도는 교통 폴리곤 단위 집계 통행량 DB에서 약 20분, 분석 특성(요일, 기간, 시간, 연령대) 집계 통행량 DB에서 약 5분, 분석 특성(통행목적) 집계 통행량 DB에서 약 1.5분 소요됨

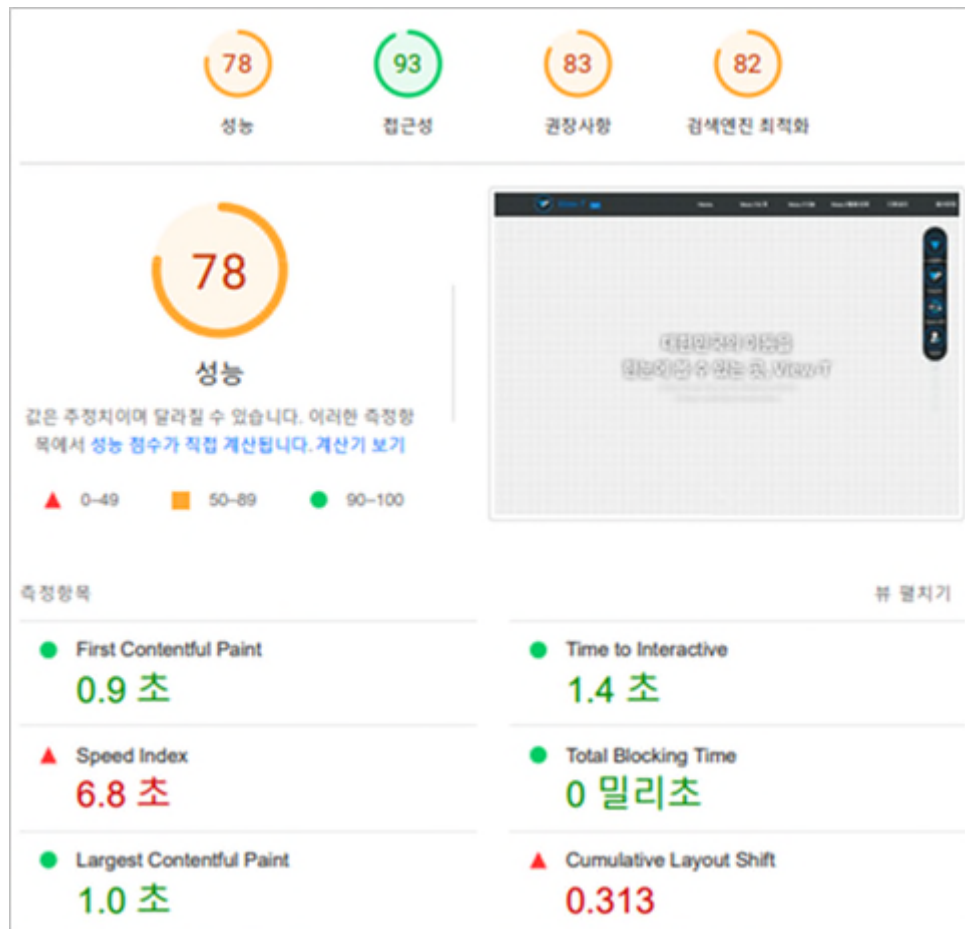
#### 나. 웹 서비스 성능 테스트

- 웹 서비스의 성능 테스트는 시스템 개발 종료 시점에 수행함
- 응답시간과 처리량, 병목 구간 등을 확인하여 서비스나 서비스 시스템의 문제점을 개선함

&lt;표 3-8&gt; View-T 웹 서비스 성능 테스트

테스트 항목	테스트 내용	충족 기준
Load 테스트	부하(Load)를 순차적으로 증가시키면서 테스트 진행 응답시간이 급격히 증가하거나 더는 처리량이 증가하지 않는 구간을 탐색 해당 구간의 CPU, 메모리 등의 점유 값이 기준값을 상회하는 임계 값을 찾아 튜닝 및 테스트 반복	CPU 점유율 90% 미만 메모리 점유율 80% 미만 결과 응답 3초 이내
Stress 테스트	임계 값 이상의 요청이나 비정상적인 요청을 보내 비정상적인 상황의 처리상태를 테스트함 시스템의 최고 성능 한계를 측정하기 위한 테스트	CPU 점유율 100% 인 상태로 3시간 이상 정상작동
Spike 테스트	갑자기 사용자가 몰렸을 경우를 가정하여 요청이 정상적으로 처리되는지를 테스트함 부하가 줄어든 때 정상적으로 반응하는지를 테스트함	동시접속 100개에서 3000회 접속 시 정상작동 결과 응답 3초 이내
Stability 테스트	긴 시간 동안 테스트를 진행하여 테스트 시간에 따른 시스템의 메모리 증가, 성능 정보의 변화 등을 확인하기 위한 테스트를 진행	일주일간 일 평균 동시접속 50개 테스트 결과 응답 3초 이내

- 페이지에 관한 진단 정보를 수집하고 분석하여 각 카테고리의 점수가 표시됩니다. 90 점 이상의 점수는 좋음으로 간주됩니다. 50~89점은 개선이 필요한 점수이며 50점 미만은 좋지 않은 것으로 간주 함
- PageSpeed Insights에서는 Chrome 사용자 환경 보고서의 데이터를 통합하여 FCP 및 DCL에서 캡처한 사이트의 실제 성능 정보를 제공 함



<그림 3-32> 구글에서 제공하는 웹페이지 속도 테스트 화면

#### 다. 웹 사이트 동작 테스트

- 서비스 개선 및 신규 추가 구현된 기능을 중심으로 웹 사이트의 동작 테스트를 자체적으로 진행하여 오류 및 추가 개선점을 해결함
- 사이트 동작 테스트는 기능 구현 담당자 간의 구현 내용을 서로 교환하여(크로스체크)함으로써 세부 내용을 인지하지 못하는 일반 사용자의 관점에서 조작에 불편함이 나 의도치 않은 작동 Case를 수집하여 별도로 관리되는 일정에 맞춰 수정을 진행함



## 제4장 결론 및 향후과제

---

### 제1절 결론

### 제2절 향후과제

## 제4장 결론 및 향후과제

### 제1절 결론

- 본 과업은 교통현황을 모니터링 할 수 있는 차량 GPS 빅데이터 기반 시계열 전국 단위의 기초교통DB 및 통행지표를 산출 및 구축하여 정부 및 지자체 등 다양한 이용자에게 제공하는데 목적이 있음
- 본 과업을 통해 샘플율이 높은 A사의 내비게이션 데이터를 수집하였으며, 해당 데이터를 가공하기 위한 경로데이터 가공 알고리즘 개발, 교통량 전수화 알고리즘 개선을 수행함
- 본 과업을 통해 아래와 같은 기초교통DB 및 18종 통행지표를 구축을 수행함
  - 전국 양방향 2차로 이상 도로에 대한 교통량, 속도DB 구축(약 106만개 링크 대상)
  - 전국 양방향 2차로 이상 도로를 일관성 있게 평가할 수 있는 18종 통행지표 구축(약 106만개 링크 대상)
    - 교통량(3종) : 관측교통량, 추정교통량, 차량주행거리
    - 속도(3종) : 평균속도, 혼잡시평균속도, 정상시평균속도
    - 혼잡(4종) : 혼잡시간강도, 혼잡빈도강도, 혼잡기대강도, 교통혼잡비용
    - 환경(5종) : 이산화탄소 배출량, 미세먼지 배출량, 일산화탄소 배출량, 휘발성유기화합물 배출량, 질소산화물 배출량
    - 안전(3종) : 운전자 피로도, 과속비율, 속도편차(신규지표)
- View-T 온라인 서비스 이용자의 사용성을 높이기 위한 분석도구 고도화 및 이용자의 편의성 개선, 신규 기능 개발 등 편리성과 다양성을 개선하여 온라인 서비스 환경 개선함
  - Faster Indicator 서비스 개발, 오픈소스SW 활용 서비스 개발, 데이터 다운로드 현황 분석 관리자 페이지 개발, 데이터 다운받기 기능 개선 등
- 데이터 기반 실증적 과학적 교통정책 수립을 지원 및 지자체 지원을 통해 교통 빅데이터 기반의 신규 부가가치를 창출함



- 대도시권 교통혼잡도로 개선사업, 제2차 국가도로망 종합계획수립, 국도시설개량사업 효과 분석 등 중앙부처, 지자체, 공공기관의 정책기반자료 제공
- 지자체 교통현안 문제 해결을 위한 지자체 실증사업 지원
  - LH토지주택공사의 공모전, 국토연구원의 교통혼잡관련 공공사업, 한국환경공단 의 교통환경관련 R&D사업 등 여러 기관에서 진행하는 사업을 지원함

## 제2절 향후과제

- 이용자 수요가 높은 신규지표 개발 및 투자 필요
  - 대기행렬 길이나 교차로의 회전 교통량, 고속도로 통행 차량의 일반도로 주요 통행 경로 DB 구축 등 조금 더 알고리즘의 개발이 필요하거나, 활용도가 높은 가공 DB 구축에도 향후 투자 필요
- 원시데이터 용량 증가에 따른 알고리즘 및 모듈 구현 시간 단축을 위한 방법론 개발 필요
  - 기존 경로데이터 가공 시 음영 구간에 대한 보정을 위해 경로 탐색을 수행하지만, A사의 내비게이션 원시 데이터를 경로데이터로 가공 시 시간이 오래 걸려 음영 구간에 대해 보정하지 않았으므로 추후 음영 구간에 대한 경로 탐색을 위한 시간 단축 방법론 개발 필요 (1일치 데이터에 대해 약 13.8시간 소요되며, 음영구간은 약 2500만 건으로 집계됨)
- View-T 온라인 서비스를 국가교통DB시스템에 탑재하기 위한 작업 필요
  - 현재 View-T 온라인 서비스는 프로토타입으로 추후 KTDB 시스템의 데이터 개방 환경 및 이용자 환경 기능으로 개선 및 확대하고자 함
- 기초교통DB 및 통행지표의 시계열성 확보를 위한 현행화 필요
  - 전국 양방향 2차로 이상 도로에 대한 18종 통행지표와 온라인 서비스 제공을 위한 2022년 기준 데이터로의 현행화 필요

2022년 국가교통조사

9

## 차량 GPS 빅데이터 구축



국토교통부  
Ministry of Land, Infrastructure and Transport



한국교통연구원  
KOREA TRANSPORT INSTITUTE