정형/비정형 교통 데이터를 통한 서울의 교통 환경 및 만족도 분석

2016년 11월 16일

박 지 영

이화여자대학교 사회과교육과 지리전공 박사과정

목 차

I. 연구의 배경 및 필요성
    I-1. 연구배경 및 목적
    I-2. 연구방법 및 범위

II. 관련 연구
    II-1. 교통 환경 분석에 관한 연구
    II-2. 교통 혼잡도 분석 연구

III. 교통데이터의 지리적 시각화 및 분석
    III-1. 교통데이터를 활용한 비정형데이터의 시각화 및 분석
    III-2. 교통데이터를 활용한 정형데이터의 지리적 시각화

IV. 결론 및 정책적 시사점

참고문헌

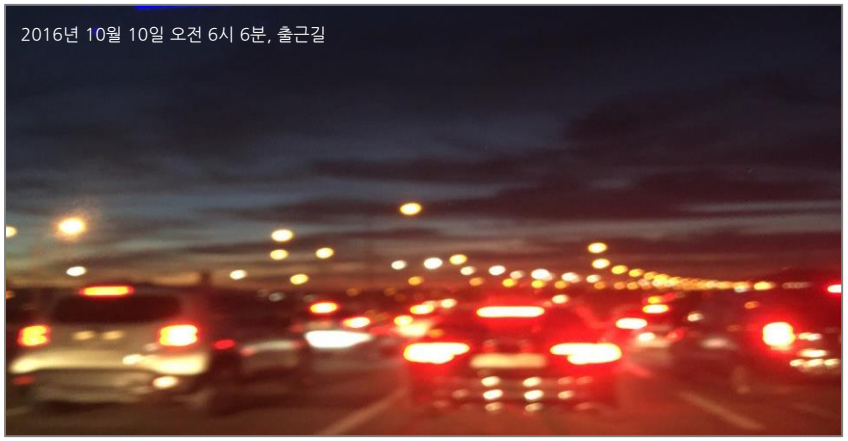
Ⅰ. 연구의 배경 및 필요성

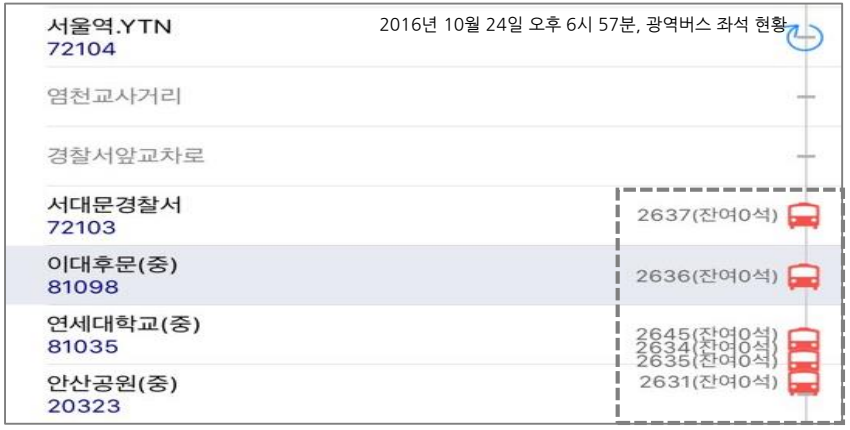
연구 배경

- 일상성과 반복성을 지니는 가장 대표적인 현상인 출퇴근길.
- 교통 환경 특히, 혼잡도나 인구 집중도는 위치, 시간의 영향을 크게 받음 : 특정 시간대(출퇴근 시간대)의 지역별 교통 혼잡에 관한 특성을 분석
- 교통 환경(교통 혼잡, 만족 정도)이란 정량적인 데이터를 분석하는 것에서 그치는 것이 아니라 정성적인 데이터를 분석할 필요가 있음.
- 분석 결과를 토대로 지역별 실질적인 정책적 대안을 제시함.



행복한 출근길을 위한 교통서비스 추진
http://gnews.gg.go.kr/news/news_detail.asp?number=201401101600517055C048


2016년 10월 10일 오전 6시 6분, 출근길



Ⅰ. 연구의 배경 및 필요성

연구 배경

최근 빅 데이터에 대한 연구는 빅 데이터를 어떻게 저장하고 처리하는가를 넘어서 어떻게 **빅 데이터로부터 의미있는 새로운 정보를 도출하고 이를 서비스에 활용**할 것인가가 중요하게 연구됨.

객관적인 수치를 보여주는 정량적인 데이터뿐만 아니라 인간의 감정이나 선호도를 표현하는 텍스트, 사진, 영상 등의 형태의 데이터까지 다양하고 복잡해짐.

인문, 사회, 지리 데이터에 기반한 전통적인 도시구조 분석을 위해 **그동안 교통 부문에서는 교통량 및 수단 분담률과 같은 총량적 데이터가 사용**되어 왔음.

교통 분야와 관련된 IT 기술의 활성화를 위해 두 분야에 대한 연구가 필요하며, 특히 설문 및 전문가 인터뷰와 같은 주관적 데이터 기반 또는 **정성적 분석 방법**과 함께 **정량적 분석 기법이 연구 동향 및 연구 기회 발굴에 객관적인 지표**를 제시해줄 수 있기 때문에 관련 데이터를 분석할 필요가 있음.

주변 맥락을 고려하지 않은 데이터는 절대 완벽한 결과물을 만들어낼 수 없음. 맥락을 이해하기 위해서는 **정성적 데이터에 대한 분석이 필수적이다. 단순 수치를 보여주는 데 그치지 않고 정량적 데이터를 분석하는 정당성을 부여하기도 함.**

최근 교통 부문에서 생성되는 승용차 이용자 통행 자료와 대중교통 이용 정보 등과 같은 개별 통행 행태 관련 자료의 활용은 서울 시민의 통행 활동을 **공간적으로 해석하고 교통 활동의 공간적 집중도를 다각적으로 분석**할 수 있는 좋은 계기가 되고 있음(김원호 외, 2009).

특정 사회적 이슈에 대한 대중들의 의견을 정확히 파악하기 위해서는 해당 이슈에 대한 사람들의 의견을 고려하여야 함(조하나 외, 2013). 비정형 데이터는 시민들의 일상사 중 자발적으로 공유한 의견을 대량으로 포함하고 있어 데이터 분석 시 국민 여론의 경향성을 파악하는 중요한 도구가 될 수 있기 때문에 비정형 데이터로서 SNS 데이터를 활용하고자 함.

기존의 연구에서는 소셜 데이터를 활용하여 대중의 의견을 파악할 때 특정 이슈에 관련된 키워드를 포함하는 메시지들을 수집하였음. 하지만 이러한 키워드가 포함된 데이터만을 이용한 연구 방법은 특정 이슈에 연관되지 않은 데이터까지도 무분별하게 수집하여 대중의 의견을 파악하기 때문에 해당 이슈에 대한 대중의 의견을 분석하는 데 문제점이 있음(조하나 외, 2013).

오준석(2015)은 텍스트 마이닝 기법을 통해 ICT 분야와 교통 관련 연구 동향을 정량적으로 분석하고 교통 분야 IoT 활성화를 위한 연구를 진행한 바 있음. 이를 통해 두 분야가 융합하여 다양한 교통 수단 및 공공 교통에 대한 운영 및 안전 연구를 수행할 필요성을 제시함.

Ⅰ. 연구의 배경 및 필요성

연구 목적

본 연구의 목적은 정량적 데이터와 정성적인 데이터를 고려하여 서울의 인구 집중에 따르는 교통 환경을 분석하고 시각화하기 위함.

첫째

- 교통 혼잡도를 지수화하기 위해 도로 차선, 도로 속도, 도로길이 등의 **도로요인**과 유동인구, 상주인구 등의 **인구요인**을 중심으로 **교통 혼잡도**를 산정

둘째

- 교통과 관련된 정형 데이터를 대상으로 다차원적인 시각화와 분석을 시도하고 **특정 시간대별 교통혼잡도**를 분석

셋째

- 기존 구축된 정형데이터와 더불어 소셜 미디어에서 생산하는 비정형 데이터를 분석함. '교통 환경(차 막힘, 교통체증)'이라는 키워드를 중심으로 비정형 데이터를 분석하여 정량적 데이터를 분석하여 찾지 못하는 의미를 분석하고 실질적인 교통 분야의 정책도 제안하고자 함.

Ⅰ. 연구의 배경 및 필요성

연구 방법 및 범위

서울의 교통 환경관련 정형, 비정형 데이터의 다각적인 분석을 통한 예측

- 도로속도, 차선 수, 도로길이 등 교통 데이터, 유동인구(출퇴근 시간대), 상주인구 데이터 분석
- SNS에 드러난 교통 환경(출퇴근길) 관련 비정형 데이터 분석

서울의 교통환경, 인구밀집에 관한 지리적 시각화 및 분석을 통한 정책적 제안

Ⅰ. 연구의 배경 및 필요성

교통환경분석

1 정형데이터분석

도로요인
- 구별 도로 속도
- 도로차선
- 도로길이

인구요인
- 유동인구
- 상주인구

심리적 요인
- 만족도 (서울 서베이 교통부문)

→ **혼잡도 도출**

- ➢ 도로요인은 낮을수록 : (-) 상관
- ➢ 인구요인은 높을수록 : (+) 상관

→ 예측 : 혼잡도가 높은 지역

혼잡도가 높은 지역은 교통 환경 만족도가 낮다.

2 비정형데이터분석

SNS data 활용
- 수집기간 : 2016.04.01 ~ 2016.06.30.
- 대상 : 트위터에 게시된 트윗 중 키워드를 포함하는 총 7080개 트윗 추출
- 키워드 : 차 막힘(차 막힌다, 버스 막힌다, 교통 체증 포함)

→ Text-mining

교통환경과 관련된 대중의 여론 파악 및 시간대별 관련 트윗 분포 분석

* 분석 및 시각화 툴 : QGIS 2.16 버전, R project 3.1.3

II. 관련 연구

교통 환경 관련 연구

- 김경환(2011)은 설문조사를 통해 요금수준이나 버스의 불규칙한 운행 간격 등과 같은 문제, **대중교통과 관련된 시설물 개선 방안 등을 제시**하고 ***이용자들의 대중교통 서비스에 대한 의식*** 분석

- 이원도,조창현(2011)은 ***유동인구의 특성을 시계열적인 변화***를 중심으로 분석. 이 연구에 따르면, 교통수요예측 및 도시계획은 상주 인구에만 국한됨으로써 태생적 한계를 지닌다고 보았고 이러한 한계점을 보완하기 위해 **상주인구와 더불어 유동인구에 대한 연구가 시행**되어야 한다고 봄.

- 김원길(2009)은 버스 운영 계획 수립 모형을 기반으로 서비스 평가지표 중 **운행 서비스 측면에서 이용자 접근 시간, 대기시간, 차내 혼잡, 교통 혼잡, 돌발 상황, 교통 법규 위반, 운행 규정** 등을 고려해서 승객 불편 등을 ***비용함수로 정량화***하여 노선 및 정류소별 ***대중교통 이용자 서비스*** 수준을 상세하게 평가

- 10년(2000-2010년)동안 수도권 도시를 대상으로 ***도시 특성을 따라 지역 간 통근 행태***가 어떻게 이루어졌는지 회귀분석을 통해 분석한 연구(진은애 외, 2013), **도시 간 버스, 승용차, 철도의 3개 부분으로 세분화시킨** ***통근통행 데이터***를 기반으로 수도권 지역 내의 도시 간 상호작용 및 공간구조를 분석한 연구도 있음(노진아 외, 2010).

II. 관련 연구

교통 혼잡도 분석 연구

- ***지역별 혼잡비용 산정***과 관련하여 경기도 시군별 km(차로)당 혼잡비용을 책정한 연구(우왕희 외, 2008)

  - 수도권의 시군별, 도로 위계별로 운행속도 편익비용과 기준 속도 편익비용의 차를 이용하여 혼잡 비용을 산정
  - 미국의 경우, 지체비용(승용차의 손실시간 추정치, 혼잡시 상업용 차량의 증가된 운영비용 추정치)와 연료비용(첨두시 혼잡통행속도, 평균 연료 소모량 및 차량의 지체시간)으로 산정
  - 일본의 경우, 시간대별 교통량과 혼잡시의 여행속도 등으로 측정
  - 캐나다의 경우, 첨두 시간대에 발생하는 혼잡에 대해 이용자가 어느 정도의 혼잡은 수용하나 그 정도가 심해져 더 이상 수용할 수 없는 한계가 존재한다고 가정하여 이에 따른 혼잡 비용 추정

- 박준식 외(2014)는 ***출퇴근 통행에 초점을 두고 수급불균형을 해소 하거나 완화*** 하기 위한 대중교통 서비스 개선방안을 제시한 바 있으며. 서울시 강남구 지역 일대를 대상으로 GIS 기능 중의 하나인 공간결합 기능을 활용한 ***실시간 교통 정보를 위한 핵심 요소***를 제시한 연구도 있음(김시곤 외, 1999).

- 교통 혼잡 관련 연구 사례에서는 혼잡지표의 개발이 중심을 이루고 있으며 Lindlrya(1986)의 혼잡심각도지표(CSI), Lomax et al.(1982)의 도로혼잡지표(RCI), Cottrell(1992)의 도로차선연장 체류시간지표(LM야),Thurgood(1994)의 혼잡지표 정규화 지수(FCI), 최기주 외(2004)의 GPS Prove 차량을 이용한 간선도로의 실시간 혼잡지표(RACI) 등이 있음.

- GPS Probe 차량을 이용하여 실시간으로 간선도로상의 혼잡 유무를 가릴 수 있는 모형을 개발하여 ***교통 상황에 대한 정보를 링크혼잡정보와 링크 통행 속도 정보를 구하여 도로 이용자들에게 제공하는 방안을 연구***한 바도 있음(최기주 외, 2004).

II. 관련 연구

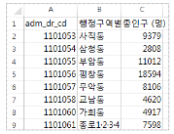
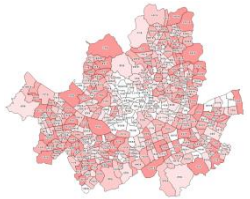
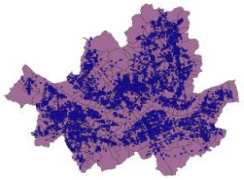
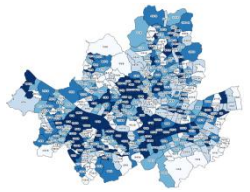
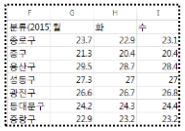
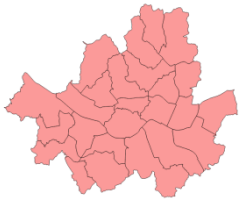
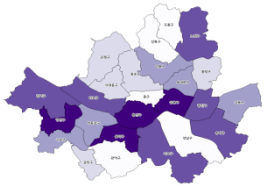
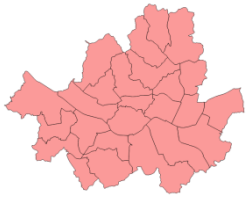
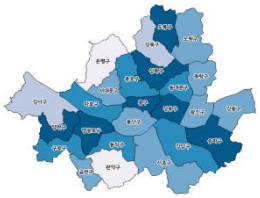
교통 혼잡의 정의

- '교통 혼잡' 정의
  - 혼잡도: (교통 분야에서의 사전적 의미) ***차량이 여럿이 한데 뒤섞이어 어수선한 상태를 말하며 이러한 도로 교통의 혼잡한 정도를 나타내는 지표***임(조순기, 2008). 혼잡도의 결과는 그 정도에 따라 '정체극심', '정체', '지체', '서행', '지체서행', '차단' 등으로 표현
  - 이상건 외(1997)는 도로의 혼잡을 용량을 초과하는 과다한 교통 수요 등의 다양한 원인에 의해 발생하는 ***차량 흐름의 지정체 및 대기행렬 현상***으로 설명

- 본 연구의 차별점
  - 대다수의 교통 혼잡에 관한 연구를 살펴보면 ***도로 교통에 관한 요인을 기준으로 지표를 선정*** 한 경우가 많음. 실제 '교통이 혼잡하다', '교통환경이 열악하다', 교통 환경이 만족스럽다. 그렇지 못하다'는 ***도로 교통 요인뿐만 아니라 인구가 집중*** 되는 버스 정류장이나 지하철역의 분포, 인구 집중 정도 등이 영향을 미친다고 생각됨.
  - 그러므로 도로 요인 이외에 인구가 집중되는 교통 환경을 반영하여 ***인구요인을 추가적으로 요인으로 고려하여 교통 환경에 관한 분석***을 실시
  - 교통과 관련된 ***정량적 데이터에 대한 분석 또는 설문지를 통한 의식 분석 정도의 연구*** 가 주를 이루었지만 본 연구에서는 이를 개선하는 방안으로 ***정량적 데이터와 정성적인 데이터를 고려한 분석***을 실시하고자 함.

III. 교통데이터의 지리적 시각화

활용 데이터 및 요인 구축 방법

요인		구축과정				활용데이터
인구요인	상주인구		속성결합			
	유동인구		공간결합			
도로요인	도로차선		불필요한 값 제거		공간결합	
	도로속도		속성결합			
	도로길이		공간결합		길이/면적 (비거주지 제거)	

III. 교통데이터의 지리적 시각화

교통 데이터를 활용한 비정형데이터의 시각화

- 교통 환경 관련 SNS 데이터 분석

1) 비정형 데이터를 분석하기 위해 SNS 상의 교통 혼잡도와 관련된 키워드인 ‘ 차 막힘’, ‘교통 체증’ 등의 단어가 포함된 트윗을 수집
2) 해석이 불가하거나 불필요한 노이즈 제거
   ① 명사/ 상태동사 등 설정된 특징 어휘만 남김
   ② 한 글자 어휘는 모두 제거, 2음절 이상의 단어 추출
   ③ 숫자는 의미 파악이 어려우므로 모두 제거
3) 텍스트 자료 내 단어들의 사용 빈도를 통계적으로 분석 : 교통과 관련된 대중들의 감정, 단순 교통 관련 텍스트 중 주제들 간의 연관관계나 부정적인 교통 관련 단어들이 비교적 빈번하게 나타나는 시간대가 언제인지 살펴보는 텍스트 마이닝을 통해 교통 환경과 관련된 대중의 여론을 살펴봄.



〈교통 환경 관련 텍스트 마이닝 결과 (워드클라우드)〉

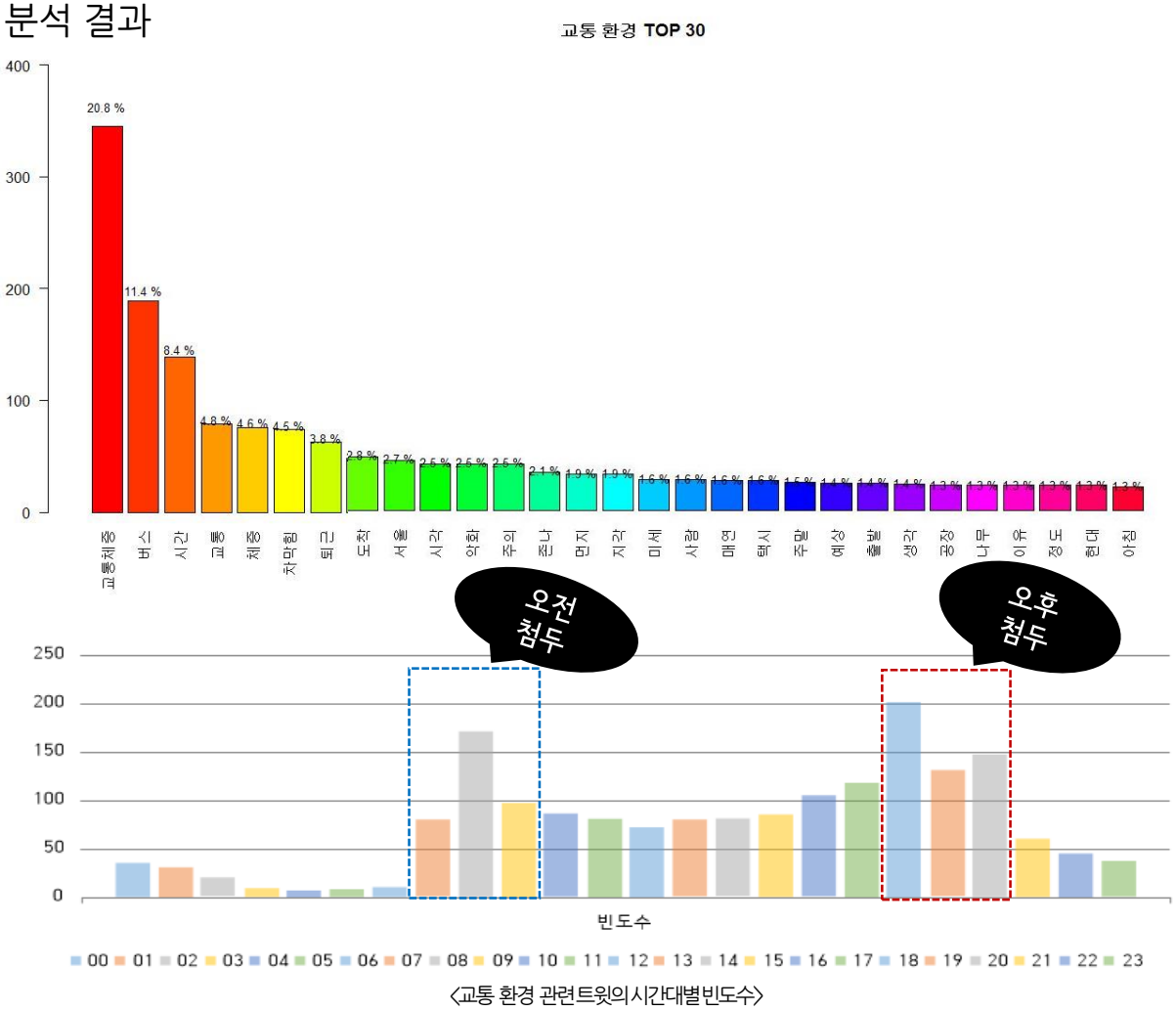


〈교통 환경 관련 상위 10개 텍스트〉

교통 데이터를 활용한 비정형데이터의 시각화

- 교통 환경 관련 SNS 데이터 분석

2) 분석 결과



〈교통 환경 관련트윗의시간대별빈도수〉

- 비정형데이터를 분석한 결과, 교통환경(교통체증, 차막힘)과 관련된 SNS 데이터에서 주제어인 '교통체증'을 제외한 텍스트 중 **버스**가 차지하는 비율이 상당히 높았음.
- **퇴근**길이라는 텍스트와 시간대별 그래프에서의 **오후 첨두 시간대(18-20시)의 빈도수가 높다**는 점은 서로 상통하는 결과라 볼 수 있음.
- 또한 '아침'의 빈도수와 SNS의 시간대별 분포에서 본격적인 출퇴근 러시아워 시간대인 7시에 급격히 SNS 빈도수가 증가한다는 점은 향후 교통 환경 분석 시, 비정형 데이터가 객관적인 수치에서 찾아내지 못하는 숨겨진 의미를 제시할 수 있음을 보여줌.

교통 데이터를 활용한 정형 데이터의 지리적 시각화

- 교통 혼잡도 요인

- 본 연구의 주요 대상인 혼잡도를 도출하기 위해 다음의 표와 같은 요인을 활용
- 교통 환경과 관련된 다양한 요인이 존재하지만 그 중 도로 요인과 인구 요인을 고려하여 도출함. 관련 연구를 살펴보면 김원호 외(2009)는 혼합율 계산시, 재차인원/버스의 용량(운행횟수*정원)으로 도출하기도 함.

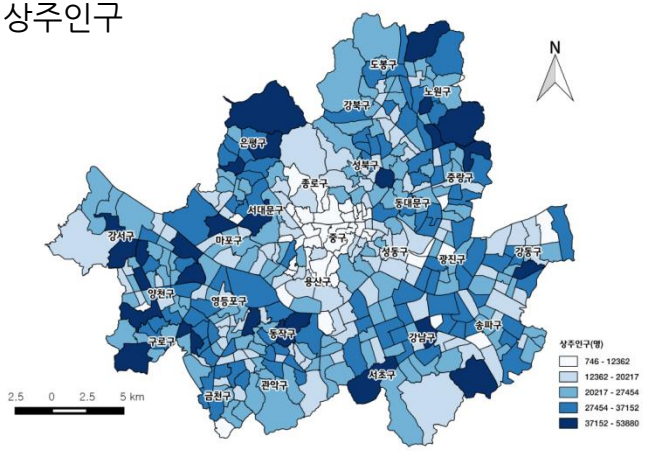
요인명	설명	비고
상주인구	• 기본적으로 인구가 많이 상주하는 지역의 교통 혼잡도가 높을 것으로 전제함. • 실제 상주인구가 많은 지역에서 인구이동이 많이 나타나며 특히, 출퇴근시간대는 거주로부터/거주지로의 이동이 많은 시간대임.	(+)
유동인구	• 유동인구가 많은 지역은 교통량이 집중되는 지역으로 예상 가능하므로 유동인구가 많은 지역일수록 혼잡도가 높음. • 연구 대상인 출퇴근 시간대에 해당되는 오전 첨두 시간(7-9시)와 오후 첨두 시간(18-20시)에는 유동인구가 가장 많을 것으로 예측함.	(+)
도로속도	• 도로속도가 빠른 지역이 혼잡도가 낮을 것으로 전제함. • 수용 가능한 도로의 용량에 비해 실제 교통량이 적정 이상인 것으로 판단되므로 도로 속도와 혼잡도는 음의 상관 관계를 가진다고 예측함.	(-)
도로차선	• 도로차선의 수가 많을수록 혼잡도가 낮을 것으로 전제함. • 도로차선이 넓은 지역은 기존의 교통 수요가 많은 지역으로 예측되고 그에 따른 교통 서비스가 이루어진 것으로 예측함.	(-)
도로길이	• 도로길이가 길수록 혼잡도가 낮을 것으로 전제함. • 면적 대비 도로가 잘 발달된 지역은 도로접근성이 좋은 지역으로 교통 서비스가 이루어진 것으로 예측함.	(-)

주: 혼잡도 = (도로차선 + 도로속도 + 도로길이 + 유동인구 + 상주인구)를 각각 5점 척도로 지수화하여 계산함.
자료출처 : 도로차선, 도로길이(국가공간정보포털 제공), 도로속도. 유동인구(서울시 제공), 상주인구(통계청 제공)
공간단위 : 서울

교통 데이터를 활용한 정형 데이터의 지리적 시각화

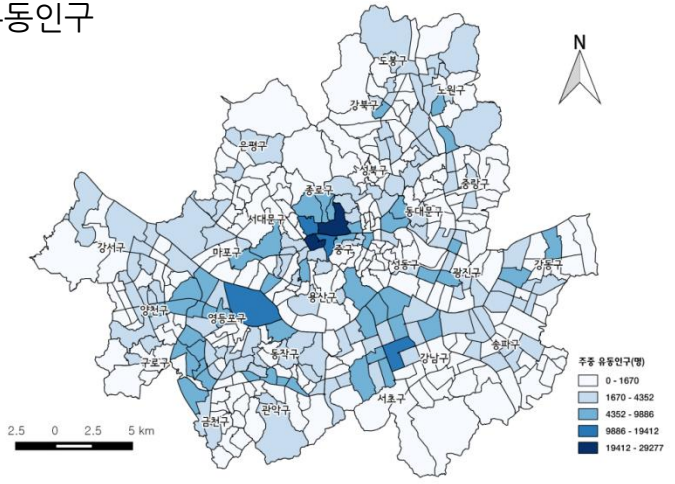
- 교통 관련 정형데이터 분석 요인

1) 상주인구



〈서울시 동별 인구 현황(2015년 기준)〉

2) 유동인구



〈서울시 주중 평균 유동 인구 현황(2012년 기준)〉

[단위: 명]

연도	수도권	서울특별시	인천광역시	경기도
2005	22,621,232	9,762,546	2,517,680	10,341,006
2010	23,459,570	9,631,482	2,632,035	11,196,053
2015	25,273,824	9,904,312	2,890,451	12,479,061
인구증감율	11.726%	1.452%	14.806%	20.676%

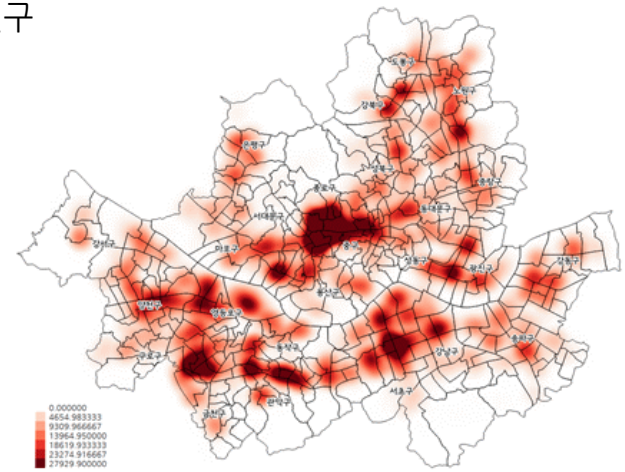
주: 인구증감율(%) = (2015년 인구-2005년 인구)/2005년 인구 * 100
자료출처: 행정동별 인구통계(통계청, 2005-2015)

- 상주인구를 살펴보면 서울시의 인구 증감율은 크지 않지만 인천과 경기의 인구 증감율이 매우 크다는 점을 알 수 있음.
- 이는 서울 내부에서의 이동보다는 서울과 외부지역에서의 진출입 인구가 많음을 시사함.
- 실제 광화문이나 종로 일대의 출근 인구는 일산, 파주 등 경기 북부지역 거주 인구에서의 출근 양상이 많이 나타나며 경기 남동부지역에서 강남 지역으로의 인구 진입이 많음.
- 이는 출퇴근 시간대에 서울의 중심부보다 외곽지역에서의 혼잡도가 높을 것임을 예상할 수 있음.

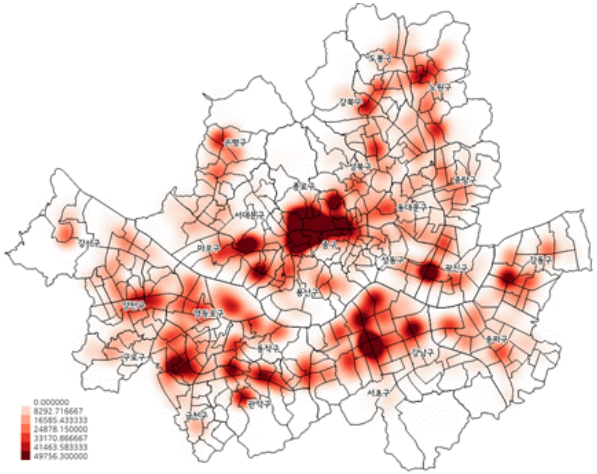
교통 데이터를 활용한 정형 데이터의 지리적 시각화

- 교통 관련 정형데이터 분석 요인

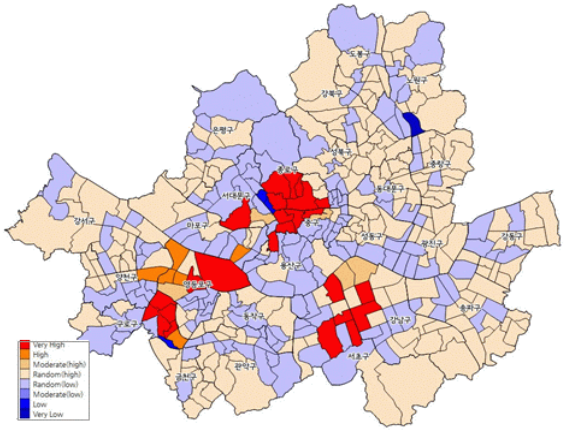
2) 유동인구



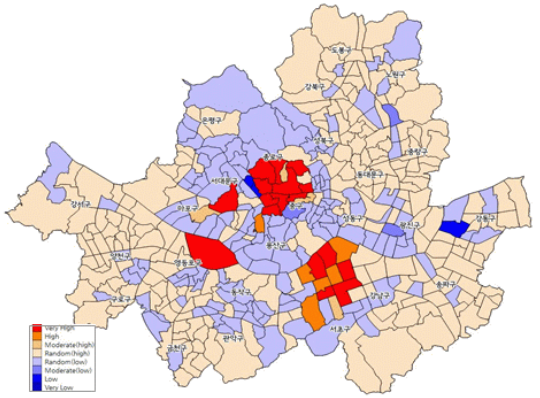
〈서울시 동별 평일 오전첨두 유동인구 현황〉



〈서울시 동별 평일 오후첨두 유동인구 현황〉



〈서울시 동별 평일 오전첨두 유동인구 현황(Moran's I 통계)〉



〈서울시 동별 평일 오후첨두 유동인구 현황(Moran's I 통계)〉

III. 교통데이터의 지리적 시각화

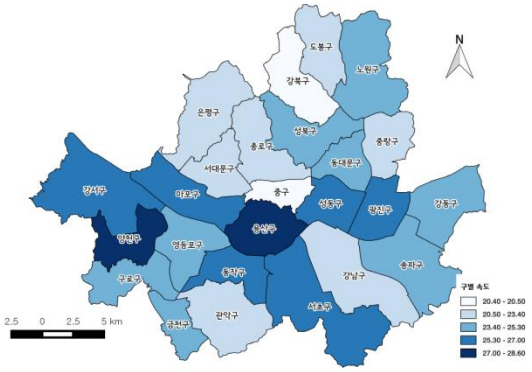
교통 데이터를 활용한 정형 데이터의 지리적 시각화

- 교통 관련 정형데이터 분석 요인

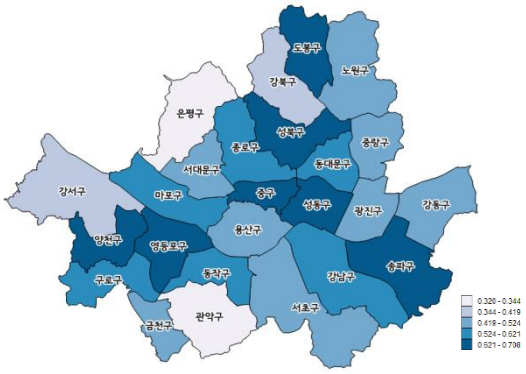
3) 도로 차선



4) 도로 속도



5) 도로 길이



- 본 연구에서는 서울에 구축된 전 도로를 모두 반영하여 보도와 차도가 구분되지 않은 이면도로 등 차선이 명확치 않은 도로를 제외하고 분석에 활용
- 혼잡도를 계산시 도로 차선이 높을수록 혼잡도가 낮다는 부적인 상관관계로 전제하였는데 도로차선이 높은 지역은 기존의 교통 수요가 많은 지역으로 예측되고 그에 따른 교통 서비스가 이루어진 것으로 예측함. 실제로 교통 서비스가 이루어진 지역의 경우 시민들이 느끼는 만족도는 혼잡도에 비해 높았음.
- 도로길이는 각 구별 면적 대비 도로의 길이를 산출하여 혼잡도 계산에 활용함. 도로의 발달은 곧, 도로로의 접근성이 용이함을 의미하므로 도로 차선과 더불어 상대적으로 수요에 따른 교통 서비스가 이루어진 지역으로 가정함.
- 도로속도는 곧 교통량과 관련이 깊음. 교통량이 많을수록 속도는 낮아지므로 혼잡도 계산시, 중요한 고려요인이 됨.
- 본 연구에서 활용한 속도 데이터는 구별 속도 데이터로서 지역별 정확성이 떨어지는 한계점을 가지고 있으며 주말을 제외한 특정 시간대에 출퇴근이 주로 나타나는 주중의 평균 속도를 기준으로 분석하였음.

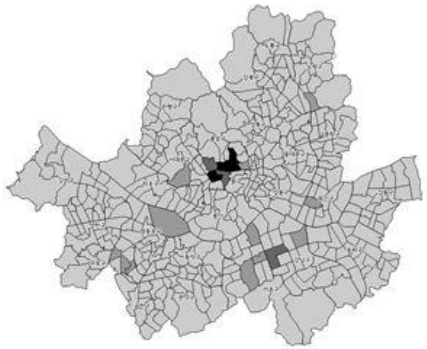
교통 환경에 대한 예측

- 교통 혼잡도 결과 도출

2) 혼잡도 표현을 위한 각 요인별 척도화(5점)

- 유동인구와 상주인구는 양의 상관관계를 가지므로 인구가 많을 수록 5점, 작을수록 1점에 가까운 점수 부여
- 도로속도와 도로차선, 도로길이는 음의 상관관계를 가지므로 속도가 느리거나 차선이 낮을수록 5점, 빠르거나 차선이 높을수록 1점을 부여

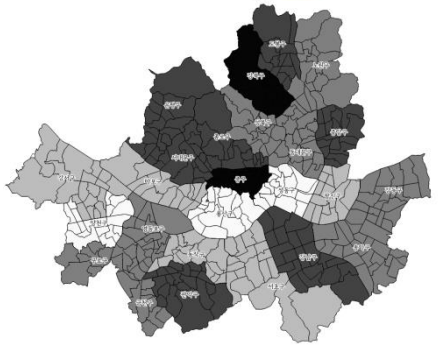
요인	범위	요인	범위
유동 인구	0<=mob_pop<9561	도로 차선	2<=lanes<5
	9561<=mob_pop<19122		5<=lanes<8
	19122<=mob_pop<28683		8<=lanes<11
	28683<=mob_pop<38244		11<=lanes<14
	38244<=mob_pop<47806		14<=lanes<17
상주 인구	746<=settle_pop<11373	도로 속도	20.4<=speed<22
	11373<=settle_pop<22000		22<=speed<23.7
	22000<=settle_pop<32626		23.7<=speed<25.3
	32626<=settle_pop<43253		25.3<=speed<27
	43253<=settle_pop<53881		27<=speed<28.7
		도로 길이	0.320<=road<0.398
			0.398<=road<0.475
			0.475<=road<0.553
			0.553<=road<0.631
			0.631<=road<0.709



〈유동인구〉



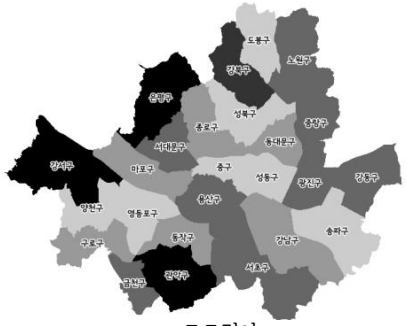
〈상주인구〉



〈도로속도〉



〈도로차선〉



<도로길이>

교통 환경에 대한 예측

- 교통 혼잡도 결과 도출

  3) 혼잡도 분석 결과



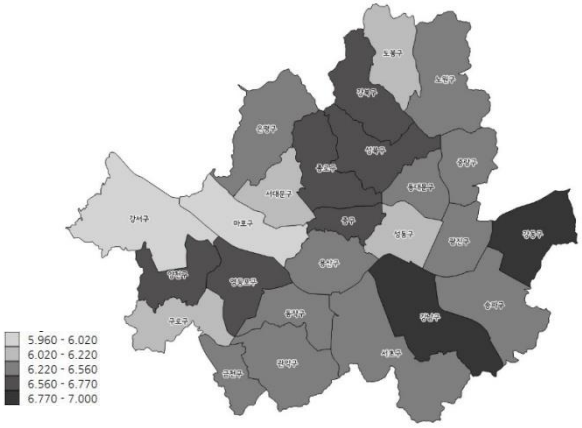
〈교통혼잡도〉

- 분석 결과, 어두울수록 혼잡도가 높은 지역임.
- 업무지구나 상업지구가 많이 분포하는 **강남구**와 종로구, 중구 일부 지역의 혼잡도가 매우 높은 것으로 나타났으며 그 밖에도 **은평구, 도봉구** 일대도 혼잡도가 높음을 알 수 있음.
- 상대적으로 주거지가 많이 분포하는 **양천구, 용산구, 성동구, 성북구** 지역이 비교적 혼잡도가 낮게 나타남.
- 이는 유동인구가 특정지역에 집중되어 높게 나타나는 강남구와 중구 일대에서 혼잡도에 크게 반영되었음을 알 수 있음.
- 혼잡도와는 별도로 강남구의 경우 다른 지역에 비해 다양한 버스 노선이나 택시 등 교통 서비스가 잘 이루어지고 있어 만족도가 높음을 추측할 수 있음.

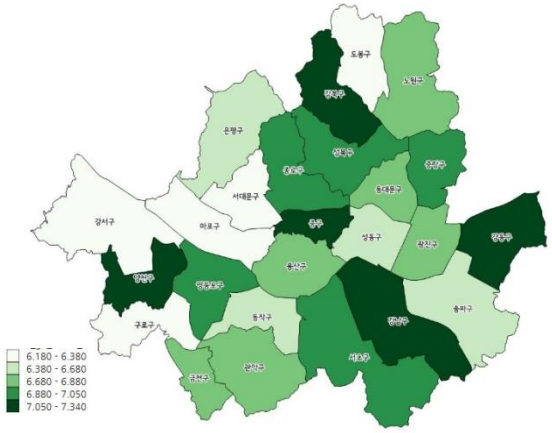
III. 교통데이터의 지리적 시각화

교통 환경에 대한 예측

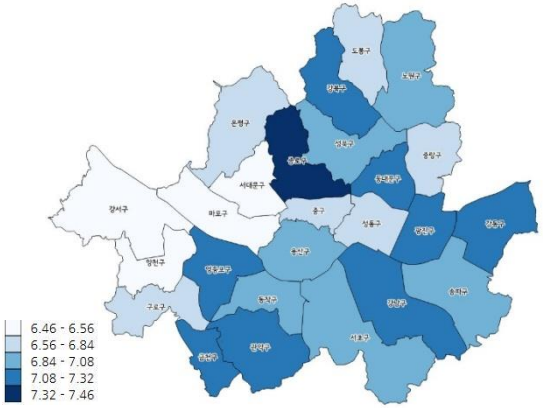
- 교통 만족도 조사 결과



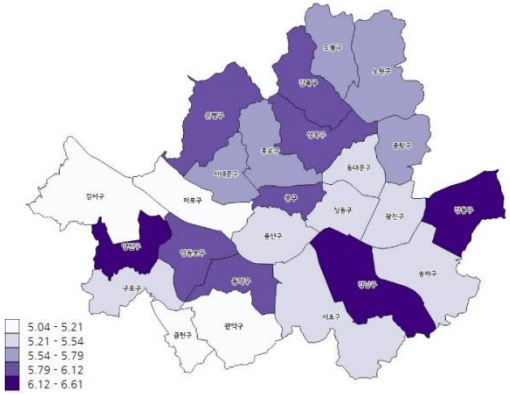
〈교통부문에대한 평균 만족도〉



〈버스 이용에대한 만족도〉



〈지하철 이용에대한 만족도〉

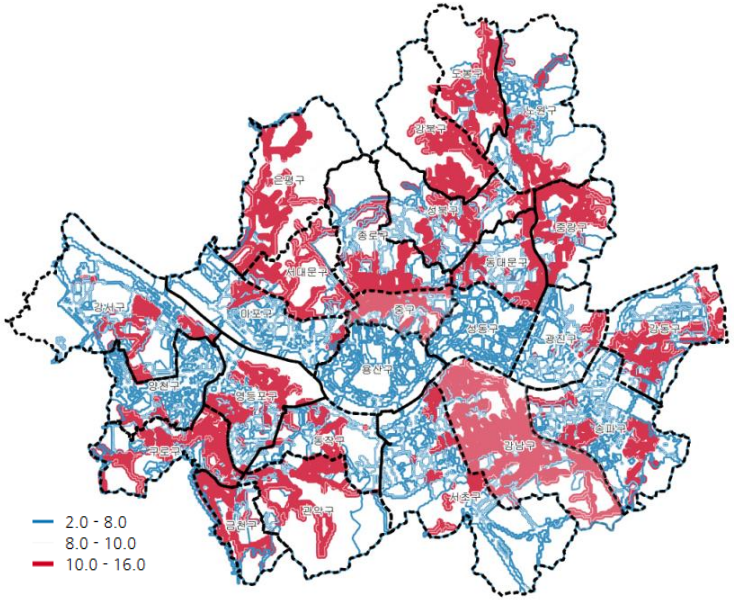


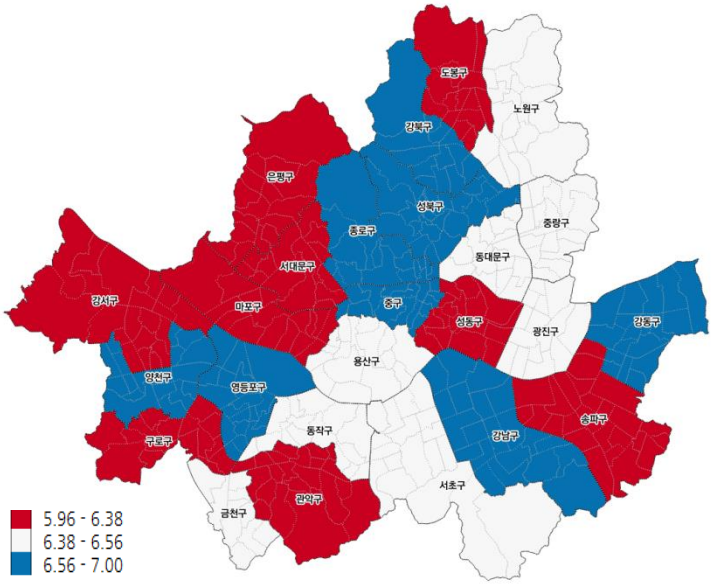
〈택시 이용에대한 만족도〉

- 교통 환경에 대한 만족도는 2014년을 기준으로 서울 시민들이 느끼는 교통 부문 만족에 대한 조사 결과를 지도화한 것.
- 전반적인 만족도는 **강남구, 강동구, 양천구, 영등포구** 등이 높게 나타났음.
- 교통수단별로 차이점이 나타났는데 지하철이용과 관련해서는 종로구에서 가장 높게 나타났으며 버스이용이나 택시이용 부분에서는 양천구와 강남구 지역의 만족도가 높게 나타났음.
- 강남구 지역은 다수의 버스 노선이 발달하고 택시 운행 서비스 등 교통서비스가 발달한 점이 만족에 영향을 준 것으로 보임.

교통 혼잡도와 만족도 비교 결과

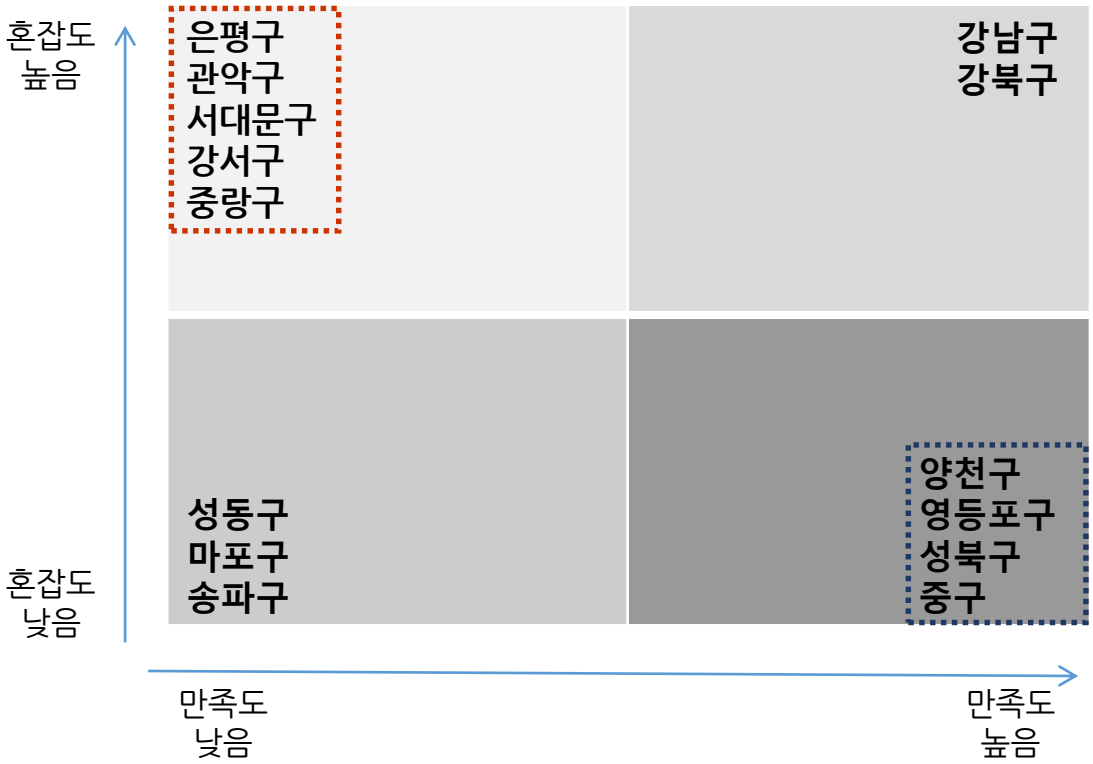
- 지역별 차별적인 교통 문제 정책적 대안 제시
  - 강남구는 다수의 버스 노선이 분포하고 조밀한 지하철망과 상대적으로 집중된 택시운행 등 우수한 교통 서비스가 제공되므로 혼잡도에 상관없이 교통 부문 만족도에서 높은 점수를 받음.
  - 양천구와 영등포구는 대체로 혼잡도가 낮은 지역이 대다수이며 만족도 또한 높게 나타났음.
  - 은평구, 관악구, 서대문구, 강서구 지역은 혼잡도가 높은 지역이 대다수이며 불만족도 또한 상대적으로 높게 나타났음. 대부분 위와 같은 교통 서비스 혜택을 받지 못하므로 교통 문제에 있어서 소외지역으로 해석될 수 있음.
  - 비교적 혼잡도가 높은 지역과 낮은 지역이 섞여 나타나는 노원구, 광진구, 동대문구 지역은 만족도에 있어서도 상대적으로 중간 수준인 것으로 나타남.





교통 혼잡도와 만족도 비교 결과

- 혼잡도와 만족도와의 관계를 기준으로 4가지 유형으로 분류



구분	혼잡도순위	만족도순위
1	은평구	강동구
2	관악구	강남구
3	강북구	양천구
4	서대문구	영등포구
5	강남구	강북구
6	중랑구	종로구
7	강서구	성북구
8	노원구	중구
9	강동구	동작구
10	종로구	노원구
11	서초구	서초구
12	구로구	광진구
13	동작구	동대문구
14	도봉구	중랑구
15	광진구	용산구
16	동대문구	금천구
17	중구	은평구
18	금천구	관악구
19	영등포구	송파구
20	송파구	성동구
21	성북구	도봉구
22	마포구	서대문구
23	용산구	구로구
24	양천구	마포구
25	성동구	강서구

Ⅳ. 결론 및 정책적 시사점

결론 및 정책적 제안

- 정형 데이터와 비정형 데이터를 함께 고려하여 분석함으로써 교통에 관한 지역별 실질적인 대안을 제시할 수 있을 것으로 기대됨.
- 교통 분야 만족도와 비교하여 혼잡도와 시민들이 느끼는 만족과의 경향성을 파악할 수 있을 것으로 사료됨.

❖ 비정형 데이터를 분석한 결과, 대중교통 중에서도 버스와 같은 특정 교통수단을 이용하거나 퇴근 시간과 같은 특정 시간대에 교통에 대한 부정적인 여론이 나타남을 확인할 수 있음.

❖ 비정형 데이터를 분석한 결과, 교통 환경 관련 트윗이 특정 시간대, 즉 오전 첨두 시간(7시-9시대)와 오후 첨두 시간(18시-20시대)에 나타남을 확인할 수 있었으며 교통 환경에 대한 만족도가 특정 시간대의 교통 환경과 관련성이 있음을 확인할 수 있음.

❖ 교통 환경을 분석하기 위한 요인으로 도로 속도나 차선, 도로길이와 같은 도로요인뿐만 아니라 유동인구와 상주인구와 같은 인구학적 요인을 고려할 필요가 있음.

❖ 정형데이터를 활용하여 출퇴근 시간대에 교통의 혼잡도를 도출한 결과, 교통이 혼잡한 지역으로 나타난 지역과 교통 부문 만족도 조사에서 만족도가 비교적 낮게 나타난 지역에 대한 상관관계가 있음을 확인할 수 있음. 지역별로 차별화된 교통 정책을 시행할 필요가 있음.

Ⅳ. 결론 및 정책적 시사점

연구의 한계점 및 향후 과제

- 연구 지역을 서울 지역으로 한정시킨 점
  - 실제 교통 혼잡은 서울 내부에서의 이동만이 아닌 서울과 경기도간의 유출입 인구의 영향을 받음. 향후 수도권 전지역으로 확대 적용하여 분석
- 구별, 동별 데이터를 대상으로 분석
  - 향후 그리드(grid) 기반의 데이터로 분석하는 방법을 제안함.
- 각 요인에 따라 가중치를 두어 차등화할 필요가 있음에도 불구하고 동일하다고 전제한 점이나 활용된 데이터의 년도가 상이함.
  - 경향성을 파악할 수 있는 자료로 활용 가능

참고문헌

김경환·박병호, 2011, “대중교통 이용실태 분석을 통한 활성화 방안 연구 : 청주시를 대상으로”, 대한교통학회 학술대회지, 2011(1), 656-661.

김광섭·남두희, 2014, “Focus+Context 시각화 기법을 사용한 교통정보 웹 서비스“, 한국인터넷방송통신학회, 14(2), 101-106.

김시곤·배상훈, 1999, “GIS와 GPS기법을 활용한 실시간 교통정보산출핵심요소”, 한국공간정보시스템학회 논문지, 1(1), 69-78.

김원길·노창균·손봉수, 2012, “대중교통 이용자 관점의 서비스 평가 모형 개발”, 대한교통학회지, 30(1), 149-159.

김원호·이신해·김승준, 2009, 교통데이터로 본 서울시의 중심성, 정책 리포트, 51, 1-18.

김정훈, 2016, “비즈니스 분석에 정형 데이터와 비정형 데이터가 미치는 영향에 대한 비교 연구”, 경희대학교 대학원 석사논문.

노진아·장훈·김지소, 2010, “수단별 통근통행에 따른 수도권의 공간구조 특성에 관한 연구”, 한국지형공간정보학회 학술대회, 69-76.

박준식·박상우·김영국·정동우·문지혜, 2014, 출퇴근 수급 불균형 해소를 위한 대중교통 서비스 정책, 한국교통연구원 기본연구 보고서, 1-134.

방승열·하효동·김창재, 2015, “공공개방 데이터 활용을 위한 빅 데이터 기반의 소프트웨어 아키텍처 설계에 관한 연구”, 한국정보기술학회논문지, 13(10), 99-107.

오준석, 2015, “텍스트 마이닝 방법을 통한 국내 교통·ICT 융합 분야 연구 기회 발견”, 교통연구, 22(4), 93-110.

우왕희·강진동·조중래, 2008, “시군별 혼잡비용 산정”, 대한교통학회 학술대회지, 2008(3), 189-193.

이경준·노윤환 ·윤상경 ·조영석, 2014, 부산지역 교통관련 기사를 이용한 비정형 빅 데이터의 정형화와 시각적 해석“, 2014, 1431-1438.

이만재, 2013, “공공 데이터 어낼리틱스와 비주얼라이제이션”, 정보과학회지, 31(3), 101-107.

이원도·조창현, 2011, “유동인구의 시계열적 변화분석: 서울시를 사례지역으로”, 대한교통학회 학술대회지, 2011(3), 95-99.

조순기, 2008, “고속도로 혼잡에 대한 운전자 인지도 분석“, 서울시립대학교 도시과학대학원 석사논문.

조하나·정연오·이재동·이지형, 2013, “인터넷 뉴스 댓글의 감성 분석을 통한 오피니언 마이닝“, 한국지능시스템학회 학술발표 논문집, 23(1), 149-150

조혜진·김강수, 2007, “수도권 통근통행의 접근도 변화패턴 분석”, 대한지리학회지, 42(6), 914-929.

진은애·주자훈·이우종, 2013, “수도권 도시특성 변화에 따른 지역간 통근행태 분석”, 국토계획, 48(7), 71-91.

최기주·장정아·정재영·심상우, 2004, “교통정보제공을 위한 GPS/GIS 기반의 실시간 혼잡지표개발”, 12(4), 53-60.

Martins, B. E., G., 2008, “Geographically Aware web text mining”, Ph.D. thesis, Departamento de Information, Faculdade de Ciencias, Universidade de Lisboa.

Glen S. Thurgood, 1994, “Development of a Freeway Congestion Index Using an Instrumented Vehicle”, TRB 1494.

Cottrell W.D, 1991, “Measurement of the extent and durraion of freeway congestion in urbanized areas”, In Compendium of Technical Papers, 61st Annual Meeting, Institute of Transportation Engineers, Washington D.C.

Herbert S. Levinson, Timothy J. Lomax, 1996, “ Developing A Travel Time Congestion Index”, presented at 75th TRB Meeting,

Jeffrey A, Lindley, 1987, “Urban Freeway Congestion: Quantification of the problem and Effectiveness of Potential Solutions”, ITE Journal.

감사합니다.