

# 플로우 빅데이터와 지역 경제지도

---

---

2017.09.18

---

황명화

---

국토연구원 국토정보분석센터

---



## 차례

---

1. 플로우 빅데이터란?
2. 기업간 거래 플로우 빅데이터로 보는 지역 산업지도
3. 신용카드 플로우 빅데이터로 보는 지역간 소비연계 지도
4. 시사점, 해결과제 및 정책제언

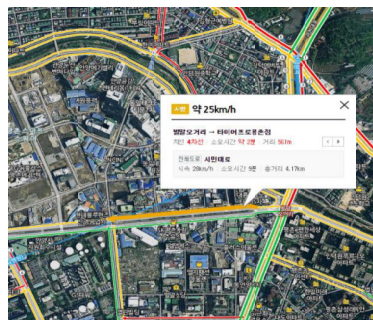
# 1. 플로우 빅데이터란?

## 플로우 빅데이터 (Spatio-temporal Flow Big Data)란?

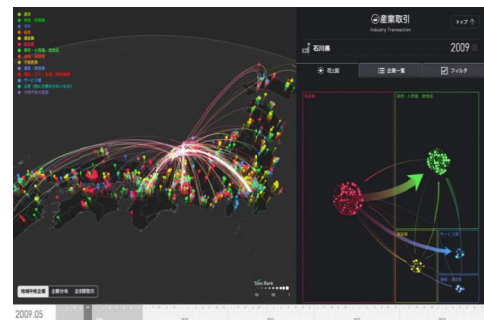
플로우 빅데이터는 빅데이터의 한 유형으로,  
국토공간 상의 다양한 네트워크를 통해 이동하는 사람, 사물, 자금, 정보, 지식 등의  
시공간적 흐름(Flow)과 그 변화를 나타내는 동적인(Dynamic)인 특성을 가진 대규모의 데이터



[사람] 트위터 사용자의 이동패턴

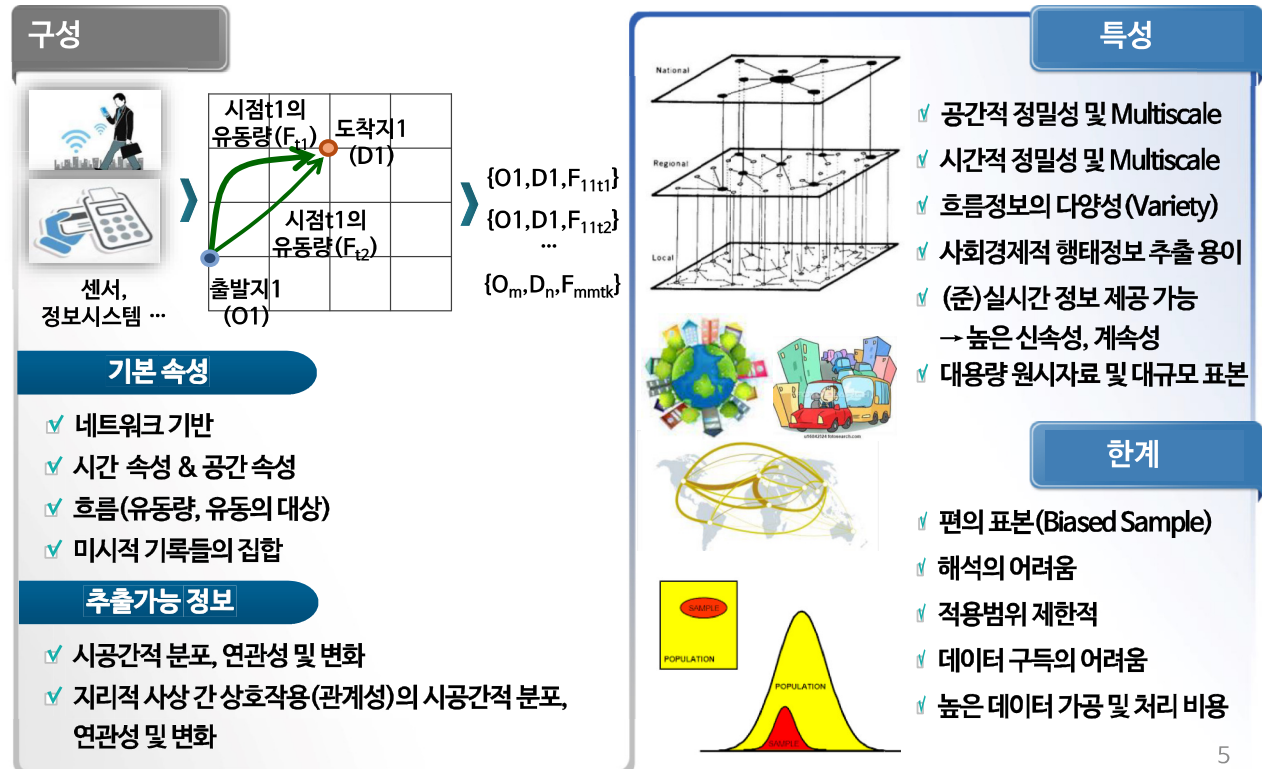


[사물] 도로구간별 교통량



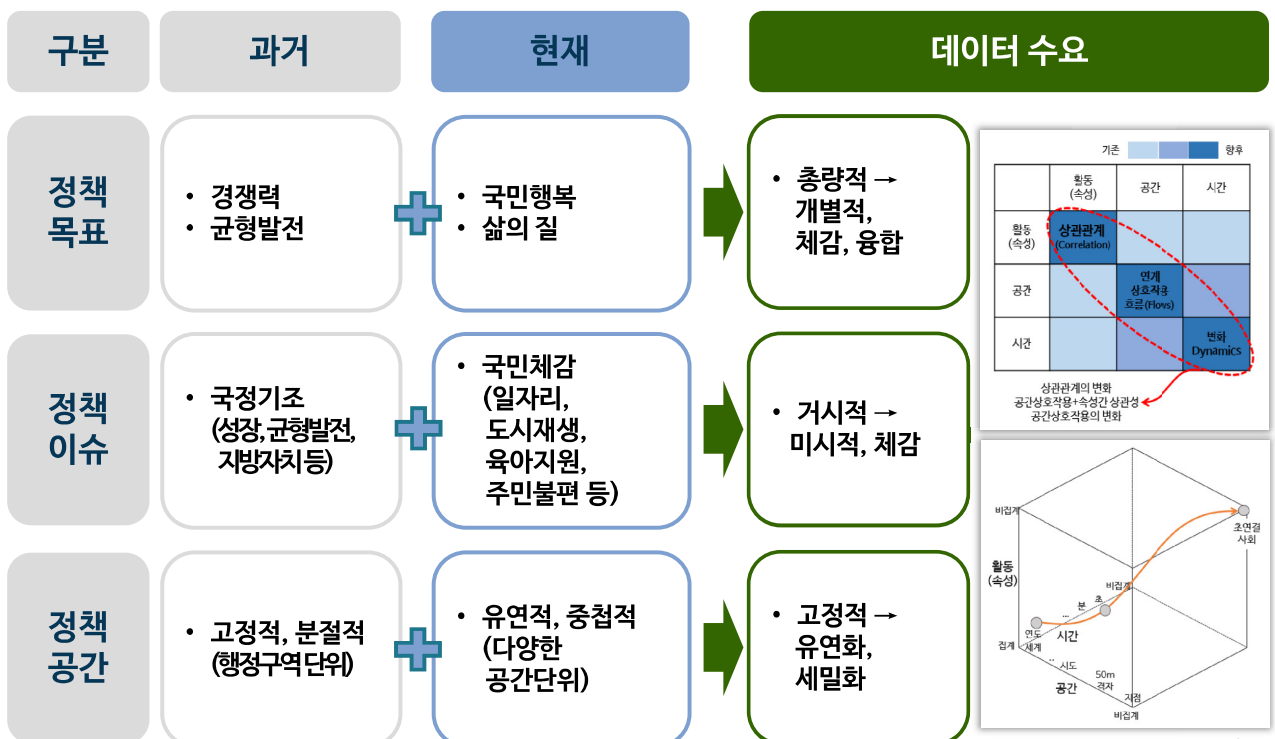
[자금] 일본 기업간 거래패턴 (RESAS)

## 플로우 빅데이터의 구성과 특성



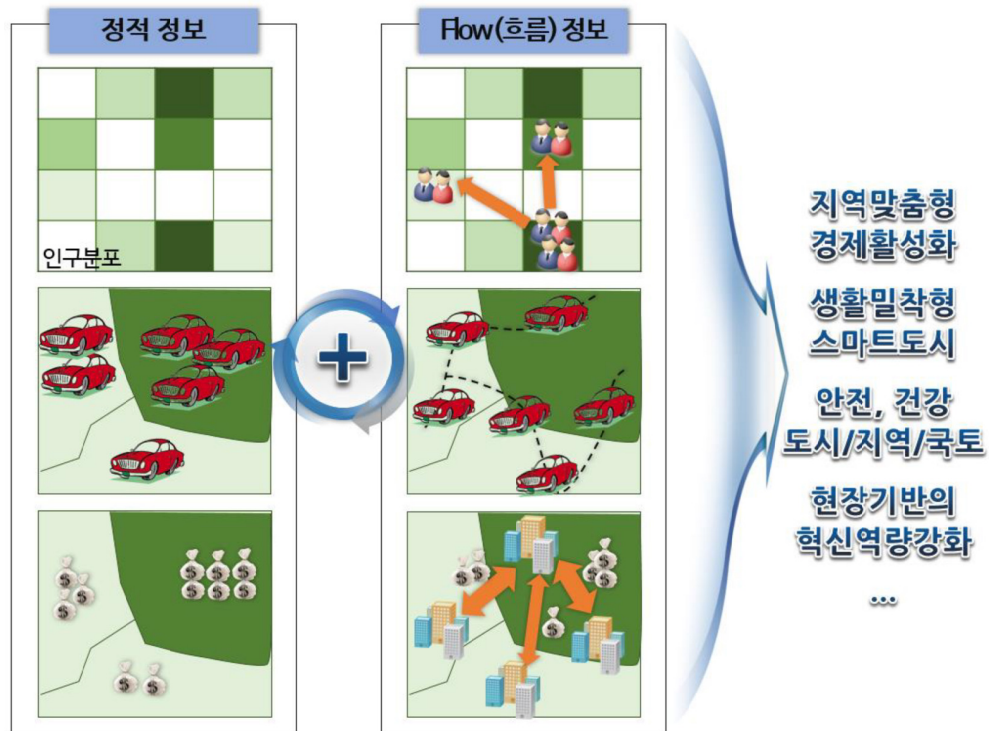
5

## 왜 플로우 빅데이터? 국토정책을 위한 데이터 수요의 변화



6

## 왜 플로우 빅데이터? 국토정책을 위한 데이터 수요의 변화



7

## 왜 플로우 빅데이터?

### 기존 공공 데이터의 한계 - 미시 통계 및 흐름 정보 부족, 부정기성

구분	경제총조사	전국 사업체조사	기업활동조사	지역통계연보 사업체 부분 (나주시 기준)
조사대상	한국표준산업분류 21개대분류중T,U를 제외한산업의모든사업체전수	개인경영농림어업사업체, 국방및가사서비스업, 국제기구및외국기관,간이판매상을 제외한모든사업체전수	국내회사법인중 종사자50인이상이며, 자본금3억원이상기업 12,000~13,000개전수	-
데이터 수집방법	직접조사 (면접 및 직접조사)	직접조사 (면접조사)	면접조사	경제총조사 잠정 데이터 활용
수집(조사/작성) 주기	매 5년	매 1년	매 1년	매 1년
구득가능한 시점	2010년, 2015년	1993년 이후 매년	2006~15년 기간동안 매년	2010~15년 기간동안 매년
구득가능한 지역통계	시도/시군구/ 읍면동 통계 마이크로 데이터를 가공하여 다양한 소지역통계 산출 가능	시도 통계	지역통계 부재, 국내외 거래액 총액 조사	읍면동 통계

8



## 2. 기업간 거래 플로우 빅데이터로 보는 지역 산업지도

### 지역산업 활성화, 지역경제의 생산력 증대를 위한 핵심요소

#### 지역산업 관련 정책이슈

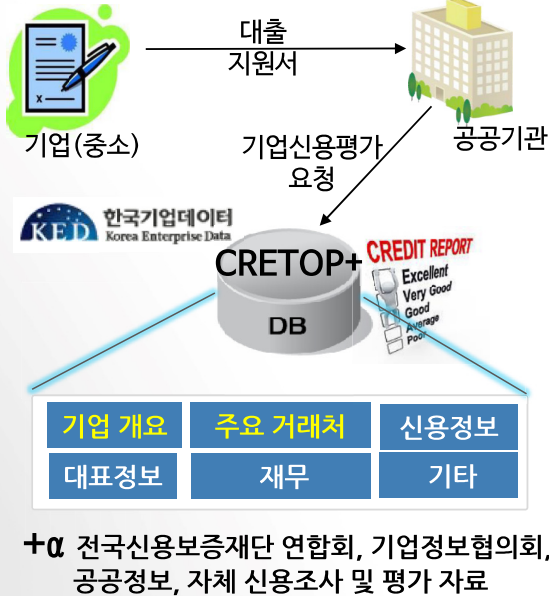


➡ 기업간 거래 플로우 빅데이터로 새롭게 파악할 수 있는 지역산업의 실상은?

## 2016년 기업간 거래 플로우 빅데이터 구축

### 원천자료

#### 민간 데이터



### 2016년 데이터

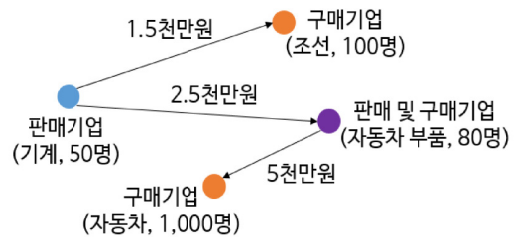
#### 기업(본사) (215,487건)

기업고유번호, 업체명, 사업자번호, 법인번호, 우편번호, 주소, 전화번호, 업종코드 (한국산업분류체계), 업종분류명, 종업원수 ..

#### 주요거래처 (1,255,628건)

기업고유번호, 거래처 기준일, 거래처 구분(판매처/구매처), 거래처 고유번호, 연간거래액, 거래비율 ...

※ 대출지원서, 수거기재자료로 설문조사데이터와 성격 유사



11

## 데이터 가공 및 분석

### 데이터가공 및 분석



### 분석변수

#### 역내 거래액

- 지역 내부 기업들간의 거래액

#### 역외 거래액

- A 지역의 기업들과 다른 지역 기업들간의 거래액

총 거래액 = 역내 거래액 + 역외 거래액

로컬도 =  $100 \times \frac{\text{역내거래액}}{\text{총거래액}}$

A지역의 B지역 의존도

=  $100 \times \frac{(A-B \text{ 지역간 거래액})}{A \text{ 지역의 역외거래액}}$

로컬도가 높으면

경제적 자립도 혹은 고립성이 높은 지역

의존도가 높으면 지역간 경제교류가 높은 지역

12

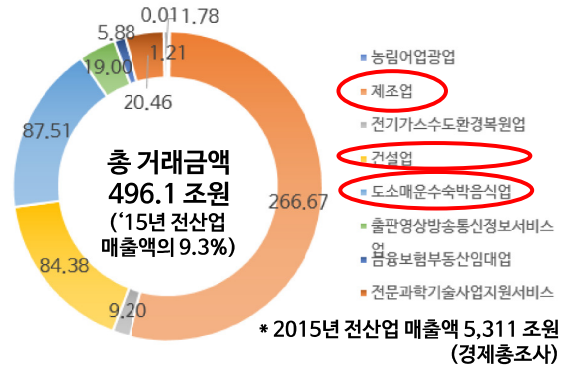
## 판매거래 네트워크 데이터 요약

### 기업(Nodes)

표 4-3 2016년 한국기업데이터 판매거래 네트워크 표본기업의 산업별 분포

업종	2016년 판매거래 네트워크 데이터		사업체총조사 (법인기업 기준)		A/C (%)	B/D (%)
	기업수 (A)	총사자수 (B)	기업수 (C)	총사자수 (D)		
농림어업 광업(01-08)	977	19,304	3,761	45,222	25.98	42.69
제조업(10-33)	65,425	2,840,829	105,319	2,823,261	62.12	100.62
전기가스수도·환경복원업(35-39)	1,632	87,805	6,864	141,502	23.78	62.05
건설업(41-42)	23,246	445,119	63,830	911,901	36.42	48.81
도소매·운수·숙박음식업(45-56)	36,018	905,019	202,568	2,114,796	17.78	42.79
출판영상방송통신·정보서비스업(58-63)	6,765	365,301	26,922	493,860	25.13	71.94
금융보험·부동산임대업(64-69)	2,037	354,746	60,381	884,631	3.37	40.10
전문과학기술·사업지원서비스(70-75)	7,748	700,020	64,603	1,611,031	11.99	43.45
공공행정·교육·복지 기타 서비스(84-96)	2,896	289,055	91,100	2,664,919	3.18	10.85
비분류	6,840	1,822				
계	153,584	5,999,020	625,348	11,691,123	24.48	51.31

### 기업간 거래(Links)

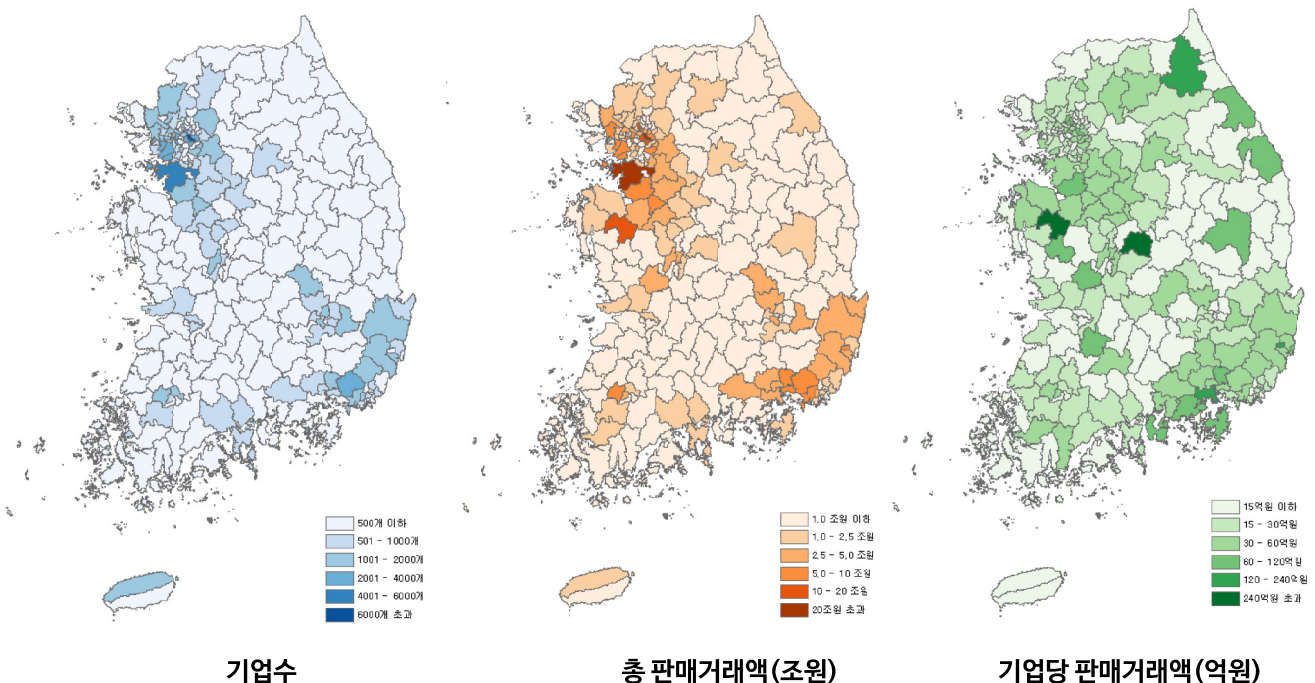


- 총거래관계수: 251,772개
- 거래관계당 거래액: 19.7억원
- 제조업내 거래액 187.2조원(53.75%), 가장 높음
- 업종간 거래액은 제조업 → 도소매·운수·숙박업이 가장 많음(51.64조원)

13

## 시군구별 기업수, 판매거래액 분포

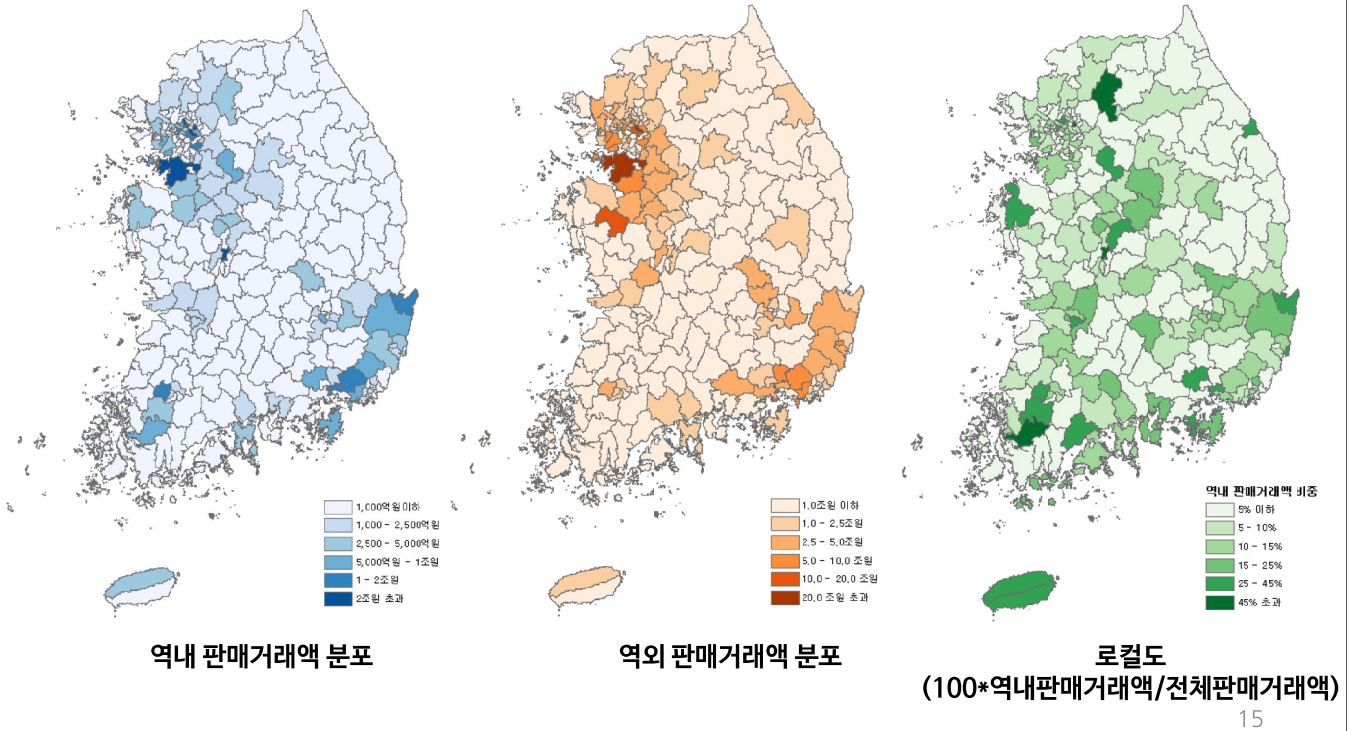
※ 본사기업의 분포 및 본사기업간 거래



14

## 시군구별 역내/역외 판매거래액 분포

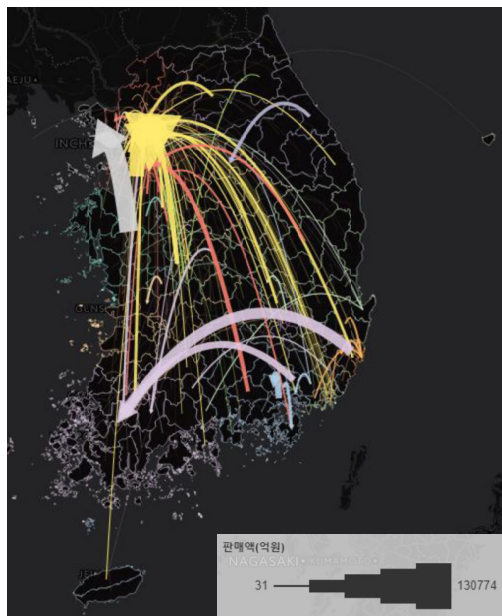
※ 본사기업의 분포 및 본사기업간 거래



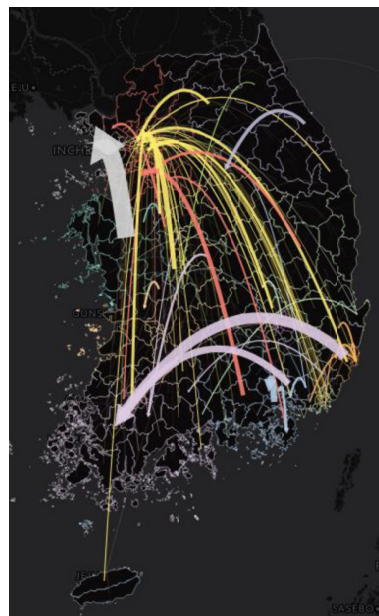
15

## 시군구간 판매거래액 분포

개별 시군구 역외 판매거래액의 구매지역별 구성비율(지역간 의존도)을 계산하여,  
(판매지역, 최대 구매지역, 역외판매거래액)을 화살표로 표시, 화살표 색상은 구매지역의 시도



시군구별 1순위 고객지역  
(전국 시군구, 화살표 종점 - 고객지역)



시군구별 1순위 고객지역  
(비수도권 시군구)

비수도권 시군구의  
주요 1순위 고객지역

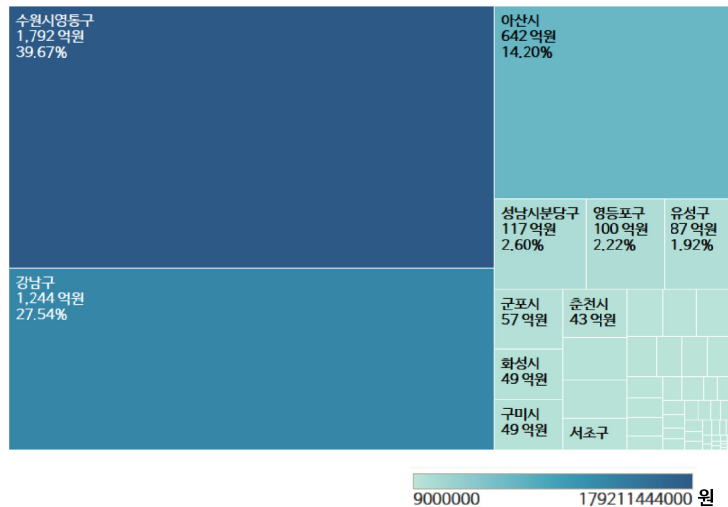
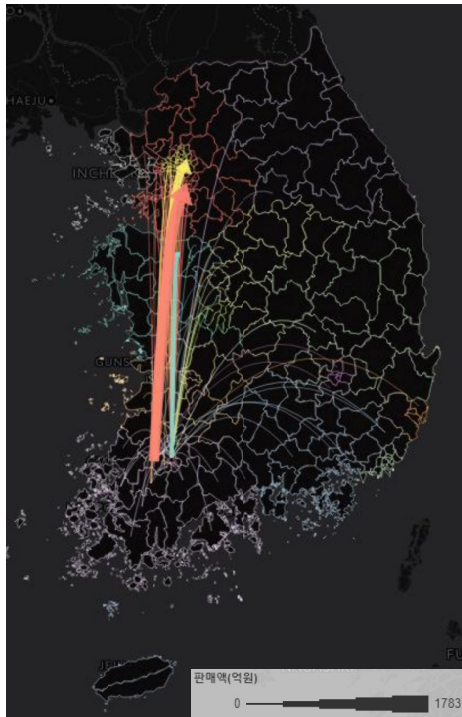
1. 서울시 강남구 (20개 시군구)
2. 서울시 종로구 (18개 시군구)
3. 서울시 서초구 (12개 시군구)
4. 서울시 영등포구 (9개 시군구)
5. 울산시 동구 (9개 시군구)
6. 전남 나주시 (8개 시군구)

16





## 광주전남혁신도시 전기전자 산업의 타지역 동종산업 연계현황



광주광역시, 나주시 이외 전기전자업종 기업에서  
타 시군구 전기전자업종으로의 판매거래액

19

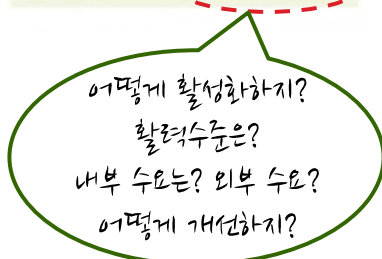
## 기업간 거래 플로우 빅데이터, 지역산업 진단 지원도구로 활용가능

- 지역산업에 관한 새로운 현황정보를 제공하여 공인통계 보완자료로 활용가치 높음
  - 지역별 판매거래액, 역내외 판매거래액, 기업거래의 로컬도 및 지역간 의존도 등
- 지역간 산업연계 특성과 업종별 차이 등에 관한 세밀한 정보 제공
  - 지역간 산업연계의 필요성과 중요성에 관한 정책적 논의는 많으나, 실태를 보여주는 데이터 근거는 부족한 실정
  - 기업간 거래 플로우 데이터는 지역간 산업연계 및 산업간 연계 실정에 관한 근거자료로 활용 가능
  - 활용잠재력이 높은 정책분야로는 경제협력권 설정, 지역산업 판로개척 지원, 지역산업지원 및 지역개발(혁신도시 등) 사업의 효과 평가 등
  - 읍면동/시군구 등 행정구역을 넘어선 실제 산업클러스터의 분포와 구조의 변화를 모니터링하는데도 적용 가능할 것으로 판단

20

### 3. 신용카드 플로우 빅데이터로 보는 지역간 소비연계 지도

지역주민의 소비, 지역경제 활력을 보여주는 바로미터  
지역주민의 소비 장소와 행태로 지역상권의 활성화 방향 파악 가능?



2017.5.24 중앙일보



- 신용카드 플로우 빅데이터로,
- 지역내부 소비수요를 파악할 수 있을까?
- 새로운 경제활력 지표를 도출할 수 있을까?
- 지역상권의 매력도 개선전략을 도출할 수 있을까?



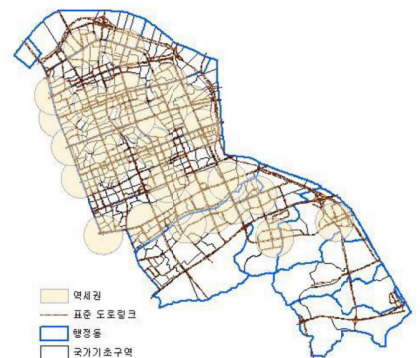
## 데이터 수집 및 가공

### ■ 데이터 수집

구분	자료출처	속성	건수
신용카드 매출 데이터	신한카드 (강남구 379개 국 가기초구역)	국가기초구역 번호, 주중/주말 구분, 시간대 구분, 업종구분, 매출액, 매출건수, <b>시군구(서울/경기) 및 시도(서울/경기 이외 지역) 단위 소비자 유입지</b>	817,775건 (15년 10월)  820,141건 (16년 4월)
택시운행 분석 데이터	서울 열린데이터 광장	도로 표준링크 ID, 요일(월,화,수,목,금,토, 일), 시간대, 날씨, 시군구 단위 목적지(승차 의 경우), 승차횟수, 하차횟수, 공차운행 횟수	10,431,016건 (강남구, 14년 11월 ~15년 10월)

### ■ 데이터 가공

- 신용카드데이터를 역세권(31개지하철역500m반경),  
행정동단위로 집계
- 택시운행분석데이터에포함된 150m표준 도로링크  
단위의승차하차횟수를국가기초구역및 역세권단위로 집계
- \* 역세권및국가기초구역과교차하는도로링크의속성값집계

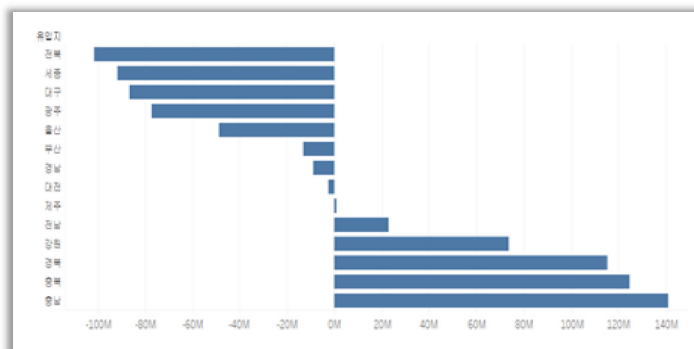


23

## 지방 거주자의 강남에서의 소비행태로 본 고급 상업수요

### ■ 신용카드 소비의 강남 의존성이 높은 지역

- 강원, 부산, 충북, 충남 (월평균 약 21~26억원 지출)
- 전북/세종/대구/광주에서는 '15년 10월에서 '16년 4월 사이 지출액 감소
- 충남/충북/강원에서는 동 기간 지출액 증가



### ■ 지방 소비자의 강남 의존성이 높은 업종

- 의료서비스, 일반음식점, 의류 및 직물업체, 학원(9위) 등

24

## 지방 거주자가 강남의 의료 및 학원교육 서비스에 지출한 금액

### 의료서비스지출 (16년4월)

강원도 998,447,186	경상북도 711,182,989	경상남도 817,798,060	부산광역시 521,631,470
충청북도 894,813,830	전라남도 460,811,156	대구광역시 404,576,380	광주광역시 349,189,360
충청남도 870,322,600	대전광역시 432,666,530	전라북도 377,194,914	울산광역시

### 학원교육 서비스지출

'15년  
10월

대전광역시 75,225,100	경상남도 65,452,955	대구광역시 56,783,690	전라북도
부산광역시 72,068,290	충청남도 61,867,860	경상북도 38,645,520	
강원도 69,190,780	충청북도 59,923,080	광주광역시 32,889,526	전라남도

'16년  
4월

대전광역시 85,246,090	충청북도 72,608,345	대구광역시 54,908,089	전라북도
경상남도 79,397,010	충청남도 63,974,870	광주광역시 47,644,380	
부산광역시 75,121,265	강원도 56,467,830	전라남도 33,418,520	울산광역시

25

## 신용카드 매출처럼 택시 승하차횟수를 경제활력 지표로 활용가능?

<표 4-22> 신용카드 매출통계와 택시 승하차횟수의 상관분석 결과(국가기초구역단위)

신용카드 매출 지표	택시승차횟수	택시하차횟수
총 매출액	.635**	.682**
총 매출건수	.648**	.629**
주중	.589**	.632**
매출액	.589**	.632**
매출건수	.594**	.576**
주말	.634**	.681**
매출액	.634**	.681**
매출건수	.657**	.631**
0~3시	.585**	.549**
매출액	.585**	.549**
매출건수	.605**	.495**
3~6시	.500**	.479**
매출액	.500**	.479**
매출건수	.508**	.421**
6~9시	.238**	.315**
매출액	.238**	.315**
매출건수	.525**	.502**
9~12시	.368**	.450**
매출액	.368**	.450**
매출건수	.541**	.566**
12~15시	.598**	.662**
매출액	.598**	.662**
매출건수	.609**	.623**
15~18시	.577**	.630**
매출액	.577**	.630**
매출건수	.577**	.583**
18~21시	.606**	.621**
매출액	.606**	.621**
매출건수	.622**	.579**
21~0시	.700**	.645**
매출액	.700**	.645**
매출건수	.663**	.549**
일반음식점	.706**	.660**
매출액	.706**	.660**
매출건수	.646**	.589**
강남구민	.423**	.498**
매출액	.423**	.498**
매출건수	.399**	.414**
강남구 외지인	.658**	.694**
매출액	.658**	.694**
매출건수	.666**	.638**

<표 4-23> 신용카드 매출통계와 택시 승하차횟수의 상관분석 결과(역세권 단위)

신용카드 매출 지표	택시승차횟수	택시하차횟수
총 매출액	.790**	.838**
총 매출건수	.821**	.809**
편의점	.909**	.865**
매출액	.909**	.865**
매출건수	.902**	.848**
일반음식점	.942**	.918**
매출액	.942**	.918**
매출건수	.886**	.850**
숙박업	.647**	.649**
매출액	.647**	.649**
매출건수	.784**	.746**
레저시설 및 판매업	.778**	.810**
매출액	.778**	.810**
매출건수	.441**	.409**
노래방	.701**	.634**
매출액	.701**	.634**
매출건수	.623**	.493**
유흥 및 사치업	.751**	.650**
매출액	.751**	.650**
매출건수	.611**	.515**
강남구민	-	-
매출액	-	-
매출건수	-	-
강남구 외지인	.848**	.877**
매출액	.848**	.877**
매출건수	.853**	.820**
강남구민	-	-
매출액	-	-
매출건수	-	-
강남구 외지인	.807**	.872**
매출액	.807**	.872**
매출건수	.795**	.781**
강남구민	-	-
매출액	-	-
매출건수	-	-
강남구 외지인	.847**	.867**
매출액	.847**	.867**
매출건수	.858**	.833**

26

## 4. 시사점, 해결과제 및 정책제언

### 정책활용 측면의 시사점

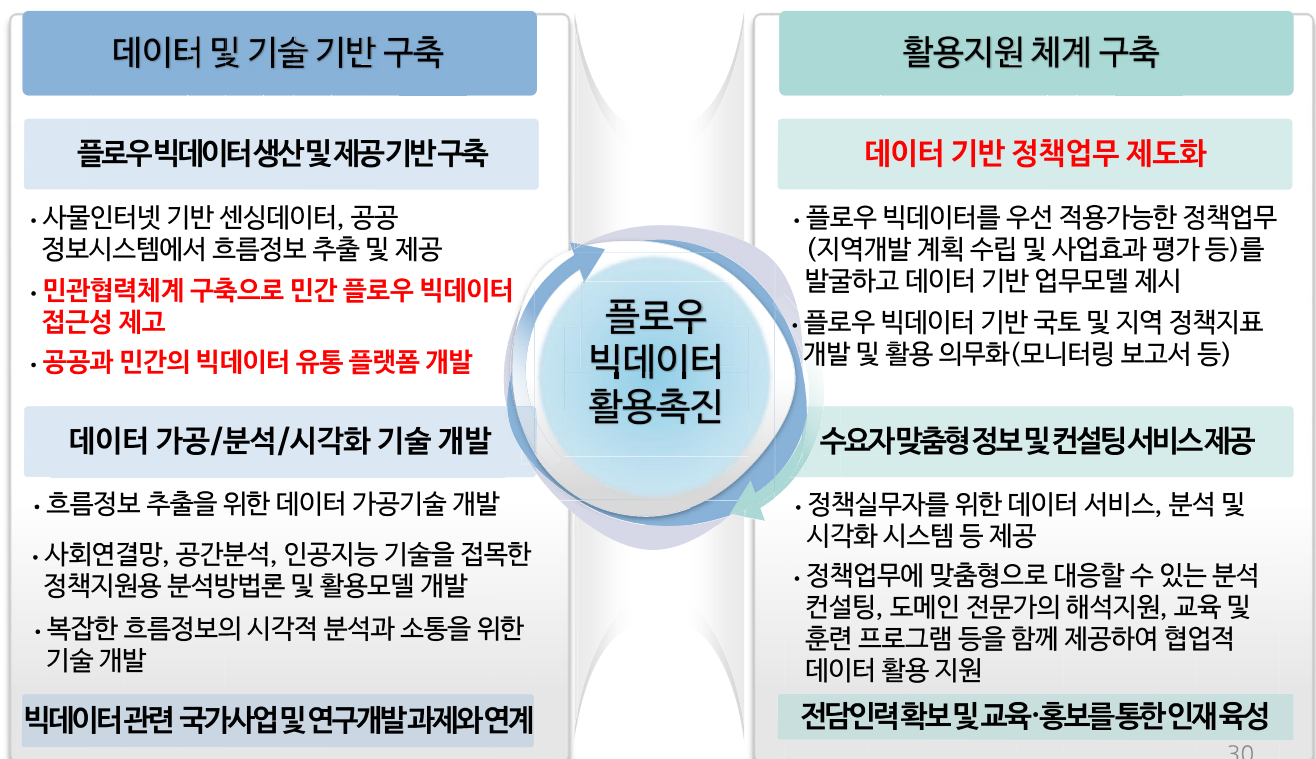
- 지역경제 정책지원 도구로서 플로우 빅데이터의 활용가능성 확인
- 지역경제 정책지원을 위한 플로우 빅데이터 활용방향 제언

주요 정책	플로우 빅데이터 활용 방향	관련 데이터 주제
지역산업 육성	<ul style="list-style-type: none"> <li>• 산업입지 전략 및 산업지원책 수립 지원</li> <li>• 산업단지 쇠퇴수준 및 연계구조 다양성 진단</li> <li>• 경제 (연계) 권역 및 지구 설정</li> <li>• 산업육성 관련제도와사업의효과평가&amp;모니터링</li> </ul>	<ul style="list-style-type: none"> <li>• 기업간 거래</li> <li>• 사업체 이주</li> <li>• 특허및문헌인용망등</li> </ul>
지역개발 및 계획 수립	<ul style="list-style-type: none"> <li>• 지역/생활권/도시 연계 등을 요하는 각종 지역개발 사업의 전략 수립 지원</li> <li>• 각종 지역개발 계획 및 사업의 타당성 &amp; 효과 평가 및 모니터링</li> </ul>	<ul style="list-style-type: none"> <li>• 기업간 거래</li> <li>• 모바일 유동인구</li> <li>• 신용카드 소비</li> <li>• 지역간통행및교통량등</li> </ul>
상권 활성화	<ul style="list-style-type: none"> <li>• 상권활성화수준진단지표및관련자료제공</li> <li>• 시장및상권활성화전략도출지원</li> <li>• 시장및상권활성화관련각종지원사업의효과평가및모니터링</li> <li>• 상권 관련지역마케팅및판로개척 전략도출지원</li> </ul>	<ul style="list-style-type: none"> <li>• 기업간 거래</li> <li>• 모바일 유동인구</li> <li>• 신용카드 소비</li> <li>• 택시 및 대중교통 이용</li> <li>• 점포이용SNS&amp;앱자료</li> </ul>
관광지역 개발 및 활성화	<ul style="list-style-type: none"> <li>• 관광코스, 자원 및 프로그램 발굴 지원</li> <li>• 관광객 유치전략 수립 지원</li> <li>• 관광 관련 지역협력권 설정 지원</li> </ul>	<ul style="list-style-type: none"> <li>• 모바일 유동인구</li> <li>• 신용카드 소비</li> <li>• 지역간통행및교통량등</li> </ul>

## 플로우 빅데이터 활용을 위한 해결 과제

- 데이터 접근 및 구득의 어려움
  - 이용가능한 데이터 목록, 가격, 구매/구득 방법, 품질 등에 관한 정보 제한적
- 데이터 신뢰성 검증 및 향상(데이터의 대표성, 시간적/지역적 완결성 등)
  - 대표성을 갖추고 시공간적으로 비교가능한 정보추출을 위한 연구개발 노력 필요
- 가공되지 않은 원석, Passive 데이터를 Actionable 정보로 변환 필요!
  - 기존 통계, 행정정보, 다른 빅데이터와의 융합, 정제, 가공, 비교 등을 통해 유의미하고 신뢰성 높은 새로운 사회경제지표로 재창출 필요
  - 공공과 민간의 데이터가 호환될 수 있도록 표준화된 용어, 기준, 분류체계 등 정립 필요
- 활용가능성 진단을 넘어선 실용화로의 도약 필요
  - 정책개발 및 평가사업 등에 직접 활용될 수 있는 정책실용화 방안 마련 필요
- 도구 및 전문인력의 부재
  - 데이터 가공, 분석 및 시각화 방법/도구 개발, 인력양성 및 활용 등의 투자 필요

## 플로우 빅데이터 활용촉진을 위한 정책제언



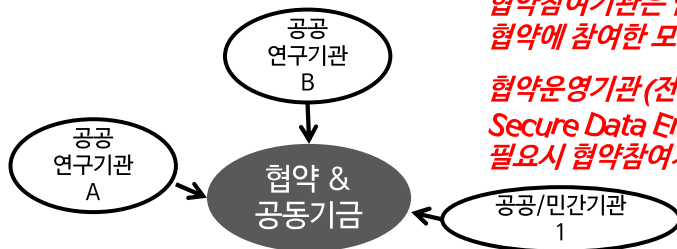
## 국책 및 지역 연구원이 함께하는 지역경제 신지식 창출 체계 마련 필요

### ■ 플로우 빅데이터와 로컬경제 지식을 연계한 국토공간 신지식 창출 협업R&D

- 4차산업혁명 시대, 국토와 지역의 경제공간은 어떻게 변화하는가?
- 국가혁신클러스터, 어떻게 만들어지고 있고 어떤 효과를 내는가?
- 상기 현안에 대응할 수 있는 새로운 정보와 지식을 창출하려면, 지역 내외부 사람, 사물, 자금, 지식 등의 흐름과 연계에 관한 글로벌 데이터의 발굴, 융합, 가치창조의 노력 필요
- 정부기관과 함께, 국책 및 지역 연구원이 신지식 창출의 리더이자 파트너로 역할 가능

### ■ 연구데이터 공유 및 공동구매 협약제도 운영을 통한 데이터 협업체계 마련

- 국책 및 지역 연구원이 보유한 데이터를 하나의 채널을 통해 공유
- 협약에 참여하는 연구기관은 일정 연회비를 내고, 협약기관들이 공유하는 모든 데이터 이용
- 연회비 중 일부는 공동 데이터 구매자금으로 적립하여, 활용수요가 높은 데이터 공동 구매 후 모든 협약기관에서 공유



협약참여기관은 연회비 납부 및 Clean (Raw) 데이터 공유, 협약에 참여한 모든 기관의 데이터 접근

협약운영기관(전문기관이나 대학) 데이터유통 플랫폼 및 Secure Data Environment 제공, 공동자금 운영 및 데이터 구매, 필요시 협약참여기관간 협업연구를 통해 2차데이터 생산/판매

31

# 감사합니다!

mhhwang@krihs.re.kr