



# “다양한 데이터의 통합 분석을 통한 공공정책 의사결정 지원” Data to Actionable Insight

하도훈 / Data사업본부

# #1. 빅데이터 예측 분석

# I. 예측분석

빅데이터 Platform을 통해 데이터 학습을 하고, 예측분석을 한다

## 데이터 수집을 통해



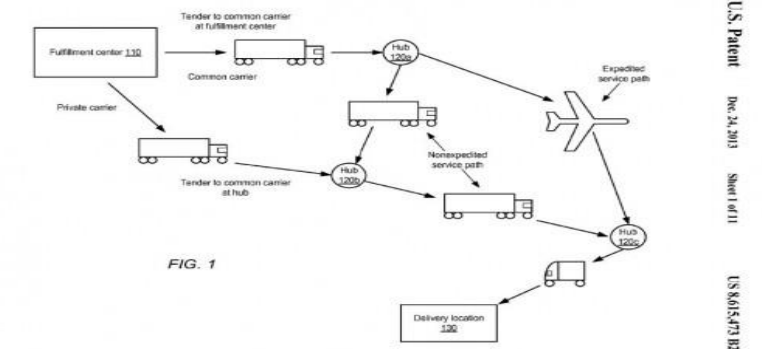
## 학습을 하고



## 예측을 한다



- **소프트뱅크 로봇 페퍼**
  - 고객이 어느 시간에 방문해 어떤 행동을 하는지는 물론 연령, 성별 등을 고려해 고객에게 최적화되도록 맞춤 대응



- **포스코 광양제철소 IoT 적용 공장 '스마트 제철소'**
  - 공장 곳곳에 IoT 센서 설치, 데이터 기반 공정관리, 근로자 건강관리

- **중국 알리바바 타오바오앱**
  - 사용자가 원하는 상품을 이미지만으로 정확하게 찾아내는 앱

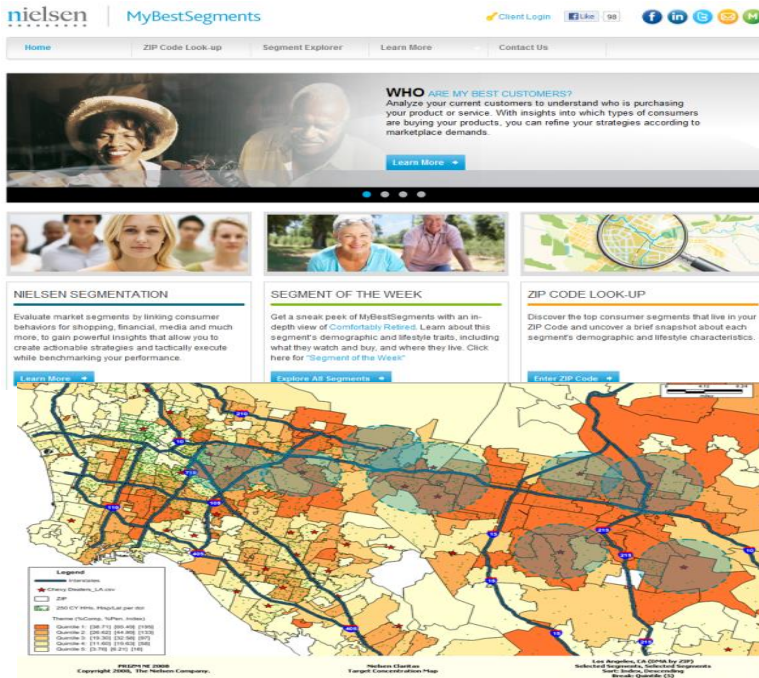
- **아마존 결제 예측 배송**
  - 고객이 이전에 어떤 상품 구매했고, 얼마나 오랜 시간 봤는지 분석



# I. 예측분석

## 다양해지는 시민Needs와특징을 반영할수 있는 형태의DB를 공공정보와 결합

### 지역기반 DB



- 미국 AC-Nielsen: Claritas PRIZM
  - GIS + Census + Segment Data
  - 패널리시지 기관인 AC-Nielsen의 DB서비스

### Single Layer DB

Group	Type	% of GB population
Group 1	Wealthy Achievers, Suburban Areas	Type 1: 3.0%
		Type 2: 2.7%
		Type 3: 2.8%
		Type 4: 3.8%
		Type 5: 2.7%
Group 2	Affluent Greys, Rural Communities	Type 6: 1.5%
		Type 7: 0.6%
Group 3	Prosperous Pensioners, Retirement Areas	Type 8: 1.4%
		Type 9: 1.1%
Group 4	Affluent Executives, Family Areas	Type 10: 2.8%
		Type 11: 1.1%
		Type 12: 0.4%
Group 5	Well-Off Workers, Family Areas	Type 13: 2.1%
		Type 14: 3.2%
		Type 15: 2.0%
Group 6	Affluent Urbanites, Town and City Areas	Type 16: 1.2%
		Type 17: 0.9%
		Type 18: 0.4%
Group 7	Prosperous Professionals, Metropolitan Areas	Type 19: 1.0%
		Type 20: 1.0%
Group 8	Better-Off Executives, Inner City Areas	Type 21: 0.8%
		Type 22: 0.7%
		Type 23: 0.9%
		Type 24: 0.7%
		Type 25: 0.9%
Group 9	Comfortable Middle Ageds, Mature Home Owning Areas	Type 26: 3.0%
		Type 27: 3.4%
		Type 28: 4.4%
		Type 29: 2.3%
		Type 30: 4.1%
Group 10	Skilled Workers, Home Owning Areas	Type 31: 4.0%
		Type 32: 4.7%
		Type 33: 2.8%
Group 11	New Home Owners, Mature Communities	Type 34: 2.6%
		Type 35: 2.7%
		Type 36: 0.9%
Group 12	White Collar Workers, Better-Off Multi-Ethnic Areas	Type 37: 1.8%
		Type 38: 1.3%
		Type 39: 1.9%
Group 13	Older People, Less Prosperous Areas	Type 40: 1.3%
		Type 41: 2.6%
Group 14	Council Estate Residents, Better-Off Homes	Type 42: 2.7%
		Type 43: 1.7%
		Type 44: 0.8%
		Type 45: 2.0%
		Type 46: 1.5%
Group 15	Council Estate Residents, High Unemployment	Type 47: 0.9%
		Type 48: 0.9%
		Type 49: 1.1%
Group 16	Council Estate Residents, Greatest Hardship	Type 50: 1.8%
		Type 51: 0.7%
		Type 52: 0.5%
Group 17	People in Multi-Ethnic, Low-Income Areas	Type 53: 1.1%
		Type 54: 0.5%

- 미국 ACORN
  - A Classification of Residential Neighborhoods: 미국을 22만개 영역, 43개 타입으로 구분하여 판매

### Multi Layers DB



- 미국 AT&T, ACXIOM
  - 3억 명에 달하는 개별 고객 프로파일 정보 판매
  - 미국 CRM 및 마케팅 정보의 표준으로 사용됨: ORACLE 등

# I. 예측분석

## 빅데이터를 통한 360° 고객 분석 (사례)

조직	발견	해석
미국국세청	사기로 의심되는 세금 환급 요청서가 적발된다. 국세청 조사담당자들은 예측 분석을 활용하여 사기로 의심되는 세금 환급 요청서의 순위를 매겨서, 사기로 의심되는 탈세를 찾아내었다.	순위를 매김으로써 조사 건수를 늘리지 않고도 25배나 많은 탈세를 찾아낼 수 있다.
Nokia / Microsoft	휴대폰 위치를 통해서 미래의 사용자 위치를 예측하는 콘테스트 주최. 한 사람이 통화하는 사회적 접촉대 상자를 분석, 스위스의 특정 지역에 거주하는 사람들에게 대해서 <b>평균적으로 그들이 어디 있게 될 것인지를 하루 전에 20미터 범위 내에서 예측 / GPS 데이터에 근거한 사람이 몇 년 후에 어디에 있게 될 것인지 예측</b>	생활 패턴에 의해 위치 정보의 군집 알고리즘에서 나오는 위치 정보를 기반으로 미래 어디에 있게 될 것인지를 예측하게 된다.
항공회사	채식주의자는 비행기를 놓치는 경우가 더 적다. <b>기내식으로 채식을 예약 주문한 항공기 승객들은 예약된 항공편을 놓치지 않고 탈 가능성이 높다.</b>	맞춤식 또는 특정한 음식이 고객을 기다리고 있다는 사실을 인지하고 있는 것은 동기를 부여하거나 책임감을 느끼게 한다.
휴렛팩커드	<b>승진은 퇴사로 이어질 수 있다.</b> HP의 한 조직에서는 승진이 그에 상응하는 충분한 급여 상승을 수반하지 않을 경우 직원들의 퇴사 리스크를 높이는 것으로 나타났다.	금전상으로 보상이 뒤따르지 않는 책임의 증가는 부담의 증가로 인식된다.
Orbitz 여행사	<b>MAC 사용자들은 상대적으로 더 비싼 호텔을 예약한다.</b> 애플의 매킨토시 컴퓨터를 쓰는 오비츠 이용자들은 호텔 방을 예약할 때 윈도우 사용자들 보다 최대 30%까지 더 많은 돈을 지출한다.	매킨토시 컴퓨터는 대체로 윈도우용 보다 더 비싸므로, 사용자의 컴퓨터 운영체제가 어떤 것인가에 따라 선택 옵션을 다르게 제시한다.
자동차보험 회사	<b>신용등급이 낮을수록 자동차 사고가 많이 난다.</b> 만약 당신의 신용등급이 다른 사람들 보다 높다면 자동차 보험회사는 당신의 보험료를 낮추어 줄 것이다..	개인 재무를 책임감 있게 관리하는 사람들은 자기 삶의 다른 측면들에 대해서도 동일한 책임감을 가지고 관리하는 성향이 있다.

## #2. 데이터 확보

# Data as a Service – SKT Big Data



## Real Data 기반 “유동인구, 카드매출정보, 부동산정보 및 GIS 정보” 등이 융합 구성

[illegible]

## II. Life-Style Segment DB

통신사 정보와 제휴사 제공 정보를 DB화 하여, 특성을 반영한 Life Style 구축

### Life Style 구성

활동(A)	관심(I)	의견(O)	인구통계
일	가족	자신	나이
취미	가정	사회적 쟁점	기업체
사회적활동	직업	정치	소득
휴가	공동체	사업	직업
오락	기분전환	경제	가족규모
사교활동	패션	교육	거주지
지역사회활동	음식	기술생산	지리
쇼핑	선호채널	미래	도시규모
스포츠	성취	문화	생활주기

※ AIO(Activities, Interests & Opinions)법 활용 Life Style척도

- 1** 통신사 정보 + 위치정보활용 산정
- 2** 통신사 정보 + 위치정보활용 유추
- 3** 제휴사 정보 + 공공정보활용 유추
- 4** Social 정보 + Survey정보 활용 유추

### Target 지역 특성별 DB 추출

**1**

- . 나이 : 통신 기본정보 수집 활용
- . 거주지 : 위치 정보 활용 특정 장소에 체류 시간 활용
- . 생활주기 : 시간대별 이동 위치 분석 패턴으로 활용

**2**

- . 가족규모 : 23:00~06:00시 특정 지역에 함께 위치하는 집단으로 유추
- . 직업/일 : 업무 시간대 거주하는 위치정보로 유추  
ex) 상주지역/생산직종 등
- . 스포츠 : 주말 활동 시간 대 위치하는 정보 활용  
ex) 골프장/ 야구경기장/축구장 POI 활용
- . 휴가 : 휴가기간 때 이동 위치 파악 이용 빈도 등 계산  
ex) 자동로밍 횟수, 캠핑장 위치 정보 활용

**3**

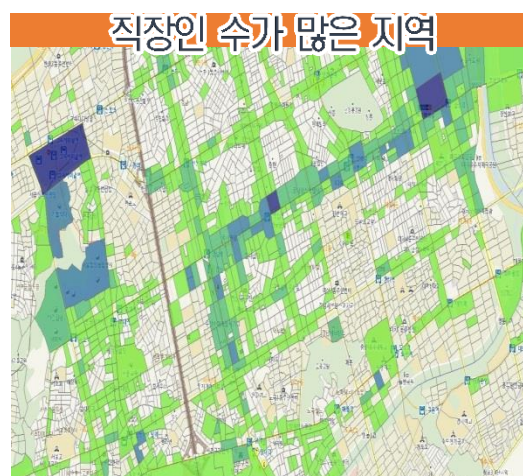
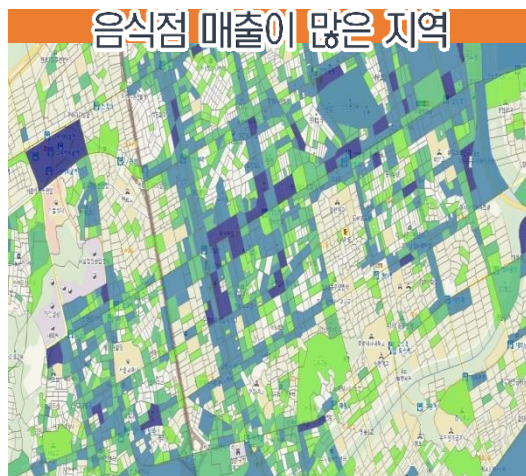
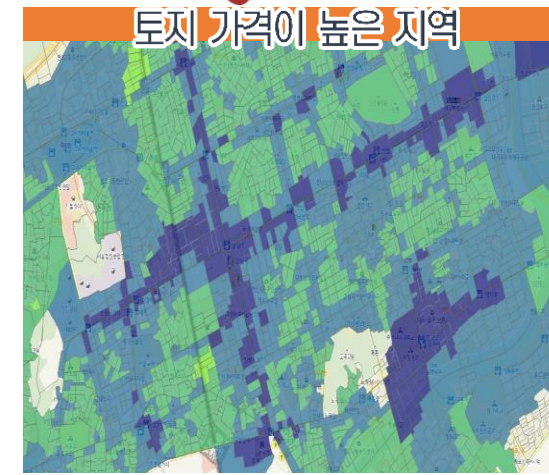
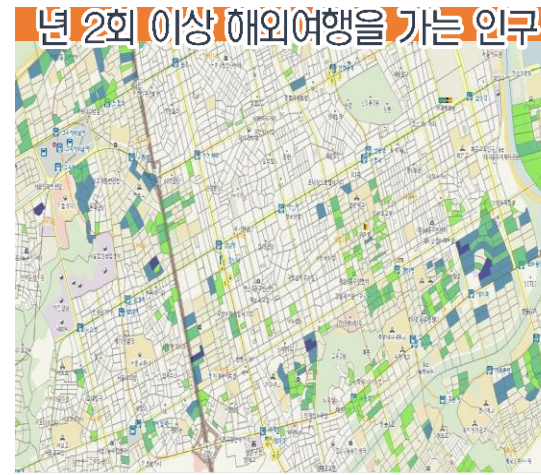
- . 쇼핑/패션 : 카드사 이용 내역 및 멤버십 이용 내용 활용
- . 선호채널 : 모바일, 온라인, 오프라인 내역 활용
- . 음식 : 카드 사용 내역 활용
- . 소득 : 부동산 거래 내역 자료 활용

**4**

- . Survey : 만족도, 목적 등의 의견 조사 정보 활용
- . Social : 온라인상의 주장, 긍부정, 사회적 쟁점 의견 정보 활용



## 개인정보를 보호하면서 통계정보화: MLDB 제공

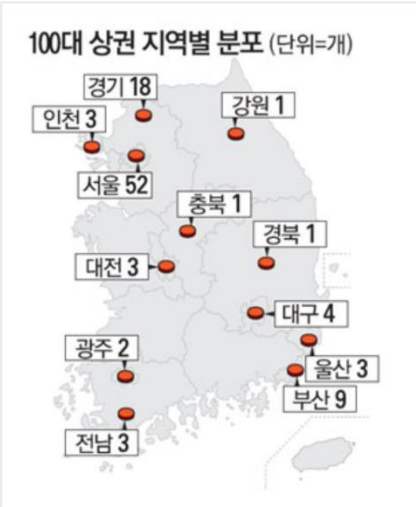






## II. 지역분석 사례

### 매경 '대한민국 100대 상권' 분석, 한경 '대한민국 상권 대전망 대박상권' 분석



#### 연령별로 본 상권 2012년

##### 10대가 많은 상권

서울 화곡역, 천호역, 신림역, 경기 아탑역 등

##### 20대가 많은 상권

서울 신촌역, 혜화역, 이대역, 부산 서면역 등

##### 30대가 많은 상권

서울 강남역, 압구정역, 명동역, 광화문역 등

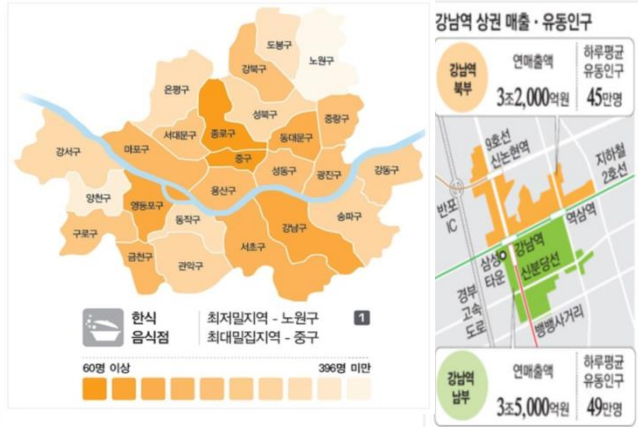
##### 40대가 많은 상권

서울 잠실역, 영등포구청역, 시청역, 부산 범일동역 등

##### 50대가 많은 상권

부산 남포동 자갈치시장 등

\*분석 : SK텔레콤 지오비전



#### 유형별 대한민국 파워상권

##### 20대 상권

상권	하루 유동인구 (20대 비중)
1위 부산 서면역	36만명(29%)
2위 인천 부평시장역	30만명(22%)
3위 대구 대구역	24만명(25%)
4위 대구 반월당사거리	19만명(27%)
5위 서울 서울대입구역	18만명(21%)

##### 직장인 상권

상권	하루 유동인구 (직장인 비중)
1위 서울 강남역 남부	44만명(47%)
2위 서울 선릉역	42만명(49%)
3위 서울 학동역	38만명(49%)
4위 서울 종각역	33만명(41%)
5위 서울 강남역 북부	33만명(45%)

\*유동인구 중 남성 30~40대, 여성 30대를 직장인으로 분류

##### 남성 상권

상권	하루 유동인구 (40~50대 비중)
1위 서울 신당역·동대문역사문화공원역	38만명(24%)
2위 서울 교대역·남부터미널역	31만명(26%)
3위 서울 성수역	30만명(29.36%)
4위 서울 종로5가역	20만명(29%)
5위 서울 잠실역	19만명(25%)

##### 환승 상권

상권	하루 유동인구 (50대 이상 비중)
1위 부산 자갈치·국제시장	40만명(36%)
2위 서울 영등포시장	23만명(34%)
3위 부산 사하구 하단역·달리역	15만명(35%)
4위 부산 범일동역	13만명(37%)
5위 부산 연산동역	12만명(41%)

\*2013년 2월 기준, 자료:SK텔레콤 지오비전

#### 대한민국 100대 상권

매출 순위	지 역	상 권	연매출	유동 인구	매출 순위	지 역	상 권	연매출	유동 인구
1	서울 강남구	강남역 남부	35,289,488,314	45	서울 서초구	충신대		4,183	63,900
2	서울 강남구	강남역 북부	32,264,448,667	52	경기 안양시	안양역		4,027	156,241
3	서울 강남구	압구정역	25,756,149,918	53	서울 강남구	한티역		4,012	62,660
4	서울 종로구	종각역	15,616,453,619	54	서울 마포구	마포역		3,971	114,253
5	서울 강남구	신사역	15,417,217,733	55	서울 송파구	신천역		3,903	39,936
6	서울 강남구	삼성역	12,060,137,229	56	경기 광명시	철산역		3,896	69,997
7	부산 부산진구	서면역	12,040,346,356	57	서울 강남구	양재역		3,714	105,315
8	울산 남구	남구중	10,237,133,833	58	서울 강남구	청담역		3,684	42,259
9	대전 서구	동산2동	10,219,93,155	59	광주 동구	금남로4가역		3,682	86,799
10	서울 강남구	선릉역	9,100,330,971	60	서울 구로구	구로디지털단지역		3,542	81,914
11	서울 강남구	학생사거리	8,608,73,102	61	부산 연제구	연산동역		3,362	90,538
12	서울 성동구	성수역	8,430,321,133	62	경기 부천시	부천역		3,324	105,792
13	경기 부천시	중동사거리	8,317,111,143	63	경기 구리시	구리시		3,307	144,162
14	부산 중구	남포동 자갈치시장 동	8,301,386,280	64	대구 수성구	범어동		3,295	114,840
15	경기 성남시	아람역	8,073,117,481	65	인천 연수구	대동월동		3,263	66,693
16	경기 성남시	서현역	8,049,166,048	66	서울 송파구	삼전동		3,238	119,305
17	서울 강남구	학동역	7,744,297,085	67	경북 구미시	월곡동		3,216	84,901
18	경기 군포시	산본역	7,565,72,266	68	경기 성남시	미금역		3,203	73,211
19	인천 부평구	부평시장역	7,546,252,592	69	인천 남동구	구월로대오거리		3,186	66,670
20	서울 종로구	광화문역	7,411,91,402	70	서울 관악구	건대입구역		3,074	61,857
21	서울 서초구	교대역·남부터미널역	6,996,245,334	71	경기 고양시	주엽역		2,984	55,982
22	서울 중구	신당역·동대문역사문화공원역	6,112,323,993	72	경기 부천시	중1-2-3동		2,923	57,894
23	서울 중구	명동역	6,111,163,411	73	전남 여수시	용곡체육관		2,917	31,365
24	충북 청주시	성안동	5,983,98,134	74	부산 북구	덕천역		2,912	65,520
25	서울 서대문구	신촌역	5,976,192,554	75	울산 남구	삼산동		2,740	57,397
26	서울 노원구	노원역	5,972,127,926	76	서울 은평구	연신내역		2,492	54,738
27	대전 서구	동산 사곡동	5,891,29,460	77	서울 강남구	강남구청		2,478	57,422
28	서울 관악구	서울대입구역	5,792,157,956	78	서울 서초구	방배역		2,469	67,222
29	경기 수원시	수원시청	5,714,133,472	79	서울 영등포구	영등포구청역		2,444	90,690
30	서울 동작구	사당역	5,679,95,462	80	부산 동래구	동래역		2,387	64,300
31	경기 안양시	반계역	5,563,93,736	81	서울 송파구	석촌역		2,248	76,398
32	부산 동구	범일동역	5,545,255,899	82	광명 원주시	중앙로		2,103	53,099
33	경기 고양시	화정동	5,493,91,807	83	서울 종로구	해학역		2,023	51,001
34	서울 서초구	교대역	5,401,174,732	84	서울 동작구	신대방역		2,006	61,984
35	울산 남구	울산시청	5,373,164,215	85	서울 중구	을지로4가역		1,937	191,060
36	경기 안산시	중안역	5,305,100,691	86	서울 성동구	용답역·성동삼거리		1,923	104,216
37	대구 중구	대구역	5,113,236,329	87	경기 하남시	하남시		1,894	116,055
38	서울 마포구	홍대입구역	5,104,126,009	88	서울 종로구	안국역		1,890	57,390
39	경기 이천시	이천	4,957,117,095	89	서울 중구	시청역		1,845	133,634
40	부산 해운대구	장산역	4,765,92,849	90	서울 강동구	길동역		1,820	33,140
41	광주 서구	광주터미널	4,726,45,073	91	서울 서초구	반포동 삼호계곡사거리		1,793	21,853
42	서울 서초구	반포고속터미널	4,682,200,685	92	서울 종로구	종로5가역		1,741	155,077
43	서울 강동구	천호역	4,635,106,315	93	서울 서대문구	이대역		1,536	48,188
44	전남 순천시	조례동	4,618,29,361	94	대구 중구	중앙로역		1,514	58,239
45	서울 영등포구	영등포시장역	4,612,187,118	95	서울 강서구	화곡역		1,384	18,731
46	서울 송파구	잠실역	4,518,144,510	96	대전 서구	갈마동		1,376	43,948
47	경기 부천시	석천사거리	4,301,56,755	97	서울 양천구	오목교역		1,195	16,725
48	서울 관악구	신림역	4,271,154,613	98	부산 부산진구	부전역·서면역		1,104	54,089
49	부산 사하구	하단역·달리역	4,192,125,004	99	서울 강북구	미아삼거리역		922	39,520
50	대구 중구	반월당사거리	4,181,152,301	100	전남 순천시	연향동		913	9,930

\* 2013년 기준, 자료:SK텔레콤 지오비전 빅데이터 서비스



\*분석 : SK텔레콤 지오비전

# **#3. 데이터 활용**

## **Business Cases**



### III. 데이터 활용

#### 현업에서 겪고있는 Pain Points

#### 경험과 직관에 따른 분석의 한계

뉴스 관련도순 | 최신순



[컬러풀대구축제 130만명 참여? 대구시, 근거 없는 뺑뺑기](#)

뉴스민 | 2017.05.30. | [🔗](#)

대구시는 지난 27일, 28일 열린 '2017컬러풀대구페스티벌'에 시민과 **관광객** 130만 명이 참여해 "글로벌... 대구시는 왜 이렇게 **뺑뺑기**했을까? 시민을 대상화하는 행정 탓에 '숫자'에만 쫓려 대구지역 한 일간지...



[대구시는 뺑뺑기 공장? '컬러풀대구페스티벌' 폐북 라이브 53만 시청의 진실](#)

뉴스민 | 2017.05.31. | [🔗](#)

대구시는 시민과 **관광객** '130만명'명 축제 참석'이라며 근거 없는 인원 추산에 미쳐, 축제 생중계 시청자... 그러나 대구시가 자화자찬한 53만 명 시청은 **뺑뺑기**에 불과하다. 대구시는 53만 명 시청자 추산 근거로...

[해운대 피서객數 빅데이터 집계... '뺑뺑기'없애](#)

문화일보 | 14면3단 | 2017.05.25. | 네이버뉴스 | [🔗](#)

수 '뺑뺑기' 논란이 개선될지 주목된다. 해운대구는 피서객의 정확한 정보 **분석** 결과를 관광분야 등 구정... 실제로 부산시가 지난해 연간 장소별 **관광객**을 빅데이터 방식으로 처음 집계한 결과 예상을 뒤엎고 해운대는...



["관광수요 예측 엉터리"...팔공산 구름다리 \*\*관광객 뺑뺑기\*\* 논란](#)

경향신문 | 2017.04.06. | 네이버뉴스 | [🔗](#)

팔공산 구름다리 조감도 또 시가 이 사업으로 동구 주민 90%인 34만9000명이 이익을 본다고 **분석**했고 1만5000명 가량의 간접고용 효과가 있다고 주장했다. 아울러 팔공산을 찾는 **관광객** 체류시간도 1~2시간 늘릴 수...

▶ 팔공산 구름다리로 **관광객** 177만→5... 연합뉴스 | 2017.04.06. | 네이버뉴스

[뉴스 더보기 >](#)

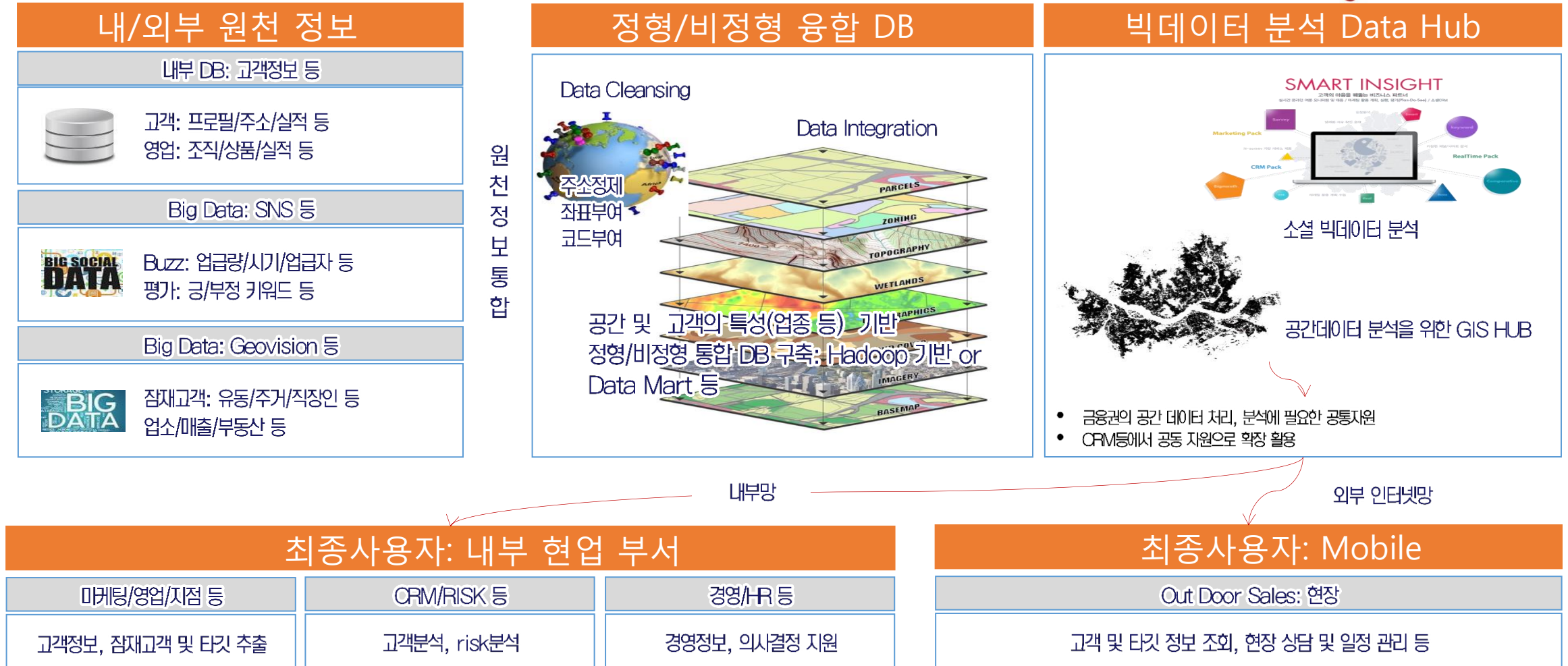
#### + 외부 데이터 & Life Style / Social 분석

"4차 산업혁명 시대에 발맞춰  
데이터 기반 과학적 행정 구현"

- **공공 데이터의 시계열 분석, 원인 파악**  
공공 보유 데이터의 활용을 통한 현업 업무 적용
- **누구나 한대씩 가지고 있는 휴대폰을 이용한 활동인구**  
방대한 데이터 결합 휴대폰 위치정보 활용
- **신용카드 결제 정보를 활용한 업종별 매출 증감**  
업종별, 성별, 시간대별, 장소별 카드 결제 정보 활용
- **소셜(뉴스, 블로그, SNS, 게시판 등) 상의 BUZZ 분석**  
인터넷 상에 공유되는 감성, 긍정/부정 분석

### III. 데이터 활용

#### 내외부 Data 결합 분석(예시)



### Ⅲ. 민간/공공 활용

#### 빅데이터 활용 전제 조건

세상에 완벽한 데이터는 없기 때문에 “차선의 데이터로 최선의 가치를 도출” 하는 것

1+2=3 인과관계가 아닌 “A와 B를 결합하여 C의 연관관계 패턴”을 찾는 것

### III. B2G 활용 : 관광객 분석

## 관광객, 축제, 외국인 분석 등 공공정책 의사결정 지원

### 내국인 유입인구 분석

- 지리적인 여건 때문에 부산 경남 지역에서 많은 인구가 유입되고, 경기도 평택시와 강원도 동해시 인구가 특징적인 부분임.
- 진해 벚꽃축제 유입 지역중 평택 2함대, 동해 1함대 해군 관계자 유입이 많았던 것으로 확인됨.

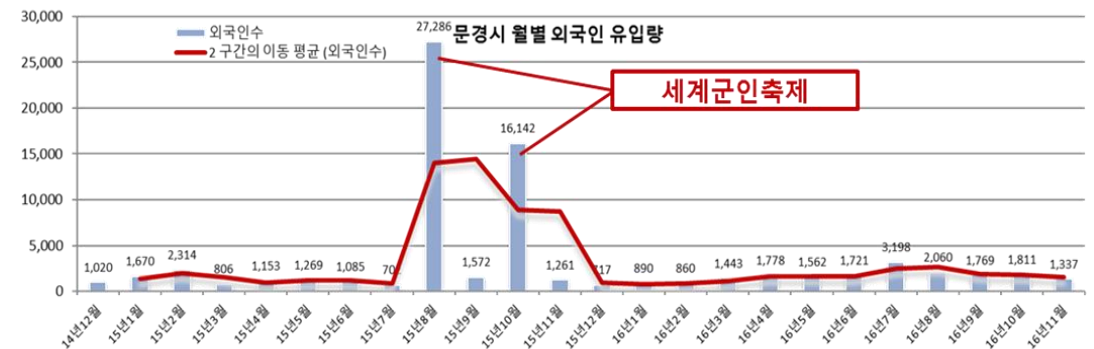
#### 시군구별 유입인구 현황 - 행사기간

순위	지역	인구수	점유율
1	경상남도 김해시	144,921	8.6%
2	부산광역시 강서구	105,680	6.3%
3	부산광역시 사하구	100,936	6.0%
4	부산광역시 남구	54,691	3.3%
5	부산광역시 부산진구	52,662	3.1%
6	부산광역시 사상구	48,286	2.9%
7	경상남도 거제시	46,291	2.8%
8	부산광역시 북구	42,189	2.5%
9	부산광역시 해운대구	33,528	2.0%
10	경기도 평택시	29,842	1.8%
11	경상남도 진주시	27,204	1.6%
12	경상남도 양산시	25,813	1.5%
13	부산광역시 동래구	23,942	1.4%
14	경상남도 통영시	20,464	1.2%
15	부산광역시 영도구	19,897	1.2%
16	부산광역시 연제구	18,772	1.1%
17	대구광역시 달서구	18,489	1.1%
18	강원도 동해시	17,949	1.1%
19	부산광역시 금정구	17,656	1.1%
20	경상남도 함안군	17,437	1.0%

#### 시군구별 유입인구 주제도 - 행사기간



### 외국인 로밍인구 분석



#### 문경시 국적별 외국인 방문객 및 점유율

순위	국적	외국인수	문경시 점유율	전국 점유율
1	중국	15,698	20.8%	46.98%
2	러시아	8,936	11.8%	1.41%
3	일본	6,902	9.1%	13.85%
4	네팔	6,404	8.5%	0.13%
5	미국	2,692	3.6%	5.47%
6	라오스	2,552	3.4%	0.05%
7	대만	2,528	3.4%	4.51%
8	폴란드	1,588	2.1%	0.10%
9	키르기스스탄	1,474	2.0%	0.04%
10	태국	1,414	1.9%	2.82%
11	필리핀	970	1.3%	3.19%
12	나이지리아	963	1.3%	0.02%
13	모로코	880	1.2%	0.02%
14	인도네시아	866	1.1%	1.62%
15	캐나다	819	1.1%	1.07%

- 세계 군인 축제에 영향으로 외국인 방문지의 경우 선수촌으로 활용된 숙소 주변에 많은 것으로 나타남.
- 세계 군인 축제에 러시아 군인들의 참여율이 높아 러시아 방문인구가 급증한 것으로 보임.

• 통신사 보정비율을 적용하여 시군 전체 방문 외국인 추정함



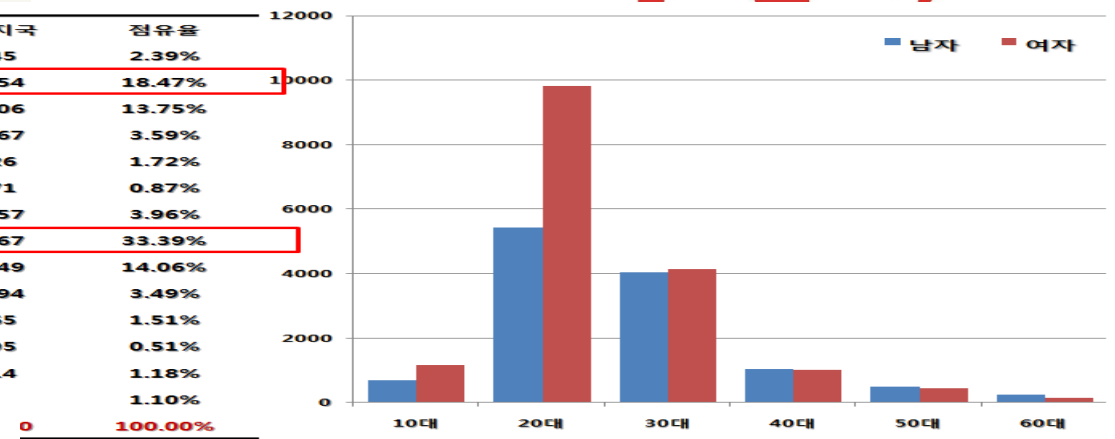
# I. B2G 활용 : 축제분석

## 축제 분석: 입장객 통계비교



인천시 입장객 통계 : 8만 6000명  
 SKT 인구 집계 : 8만1358명  
 SKT 통신인구의 경우 휴대폰 미소지자는 제외된 수치임

성/연령	이동기지국	점유율
남자_10대	702.45	2.39%
남자_20대	5428.54	18.47%
남자_30대	4039.06	13.75%
남자_40대	1053.67	3.59%
남자_50대	505.26	1.72%
남자_60대	255.71	0.87%
여자_10대	1164.57	3.96%
여자_20대	9812.67	33.39%
여자_30대	4131.49	14.06%
여자_40대	1025.94	3.49%
여자_50대	443.65	1.51%
여자_60대	150.95	0.51%
외국인	348.14	1.18%
		1.10%



기지국	날짜	방문자_총합
이동기지국	20160812	17,931
이동기지국	20160813	34,044
이동기지국	20160814	29,386

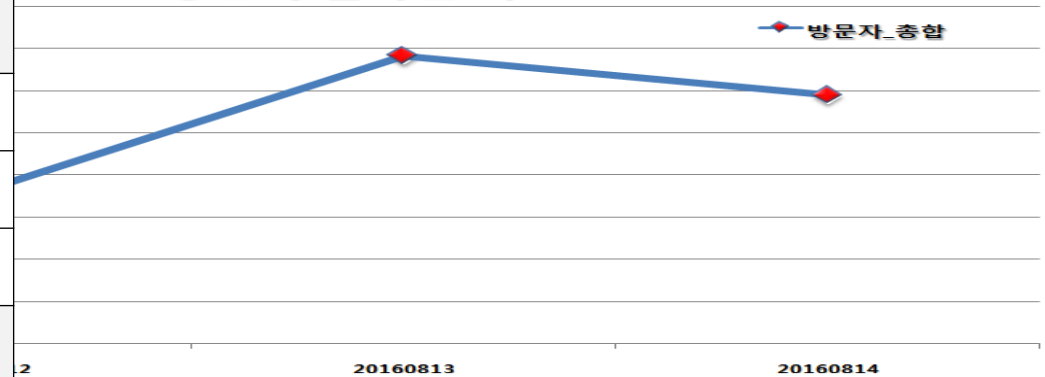
### ▶ 뉴스 발표 자료

연예	펜타포트 입장객 8만 6000명 '뜨거웠던 3일간'
기사본문	댓글 바로가기 등록 : 2016-08-16 15:36
스팟뉴스팀 (spotent@dailian.co.kr) 기사더보기 +	
집객수로는 8월 12일 1만 7000명, 13일 3만 8000명, 14일 3만 1000명이 운집해 총 8만 6000명이 행사장을 찾았으며, 펜타포트 기간에 연일 폭염특보가 발효되는 상황이었음에도 불구하고, 단 한 건의 안전사고는 발생하지 않았을 정도로 안전한 운영 노하우를 엿볼 수 있었다.	

### ▶ SKT 분석자료

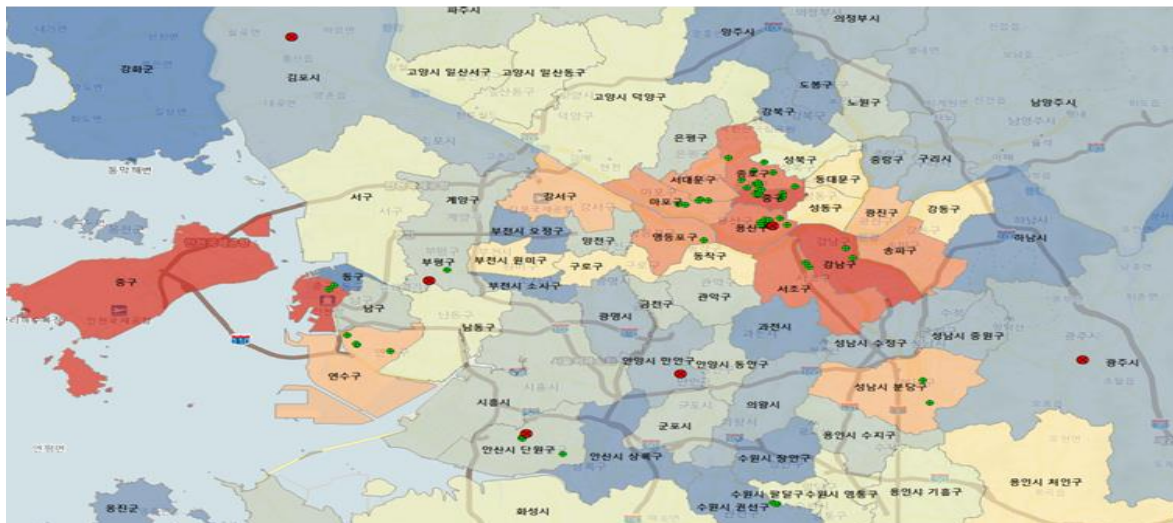
날짜	방문자_총합
20160812	17,930
20160813	34,043
20160814	29,385
총합	81,358

### 방문객 일자별 비교



### III. B2B 활용: 외국인 분석

#### 이슬람 외국인 유동인구 분석



순위	시군구명	무슬림 인구비율	이슬람 성원	무슬림 친화식당	무슬림 기도실
1	서울 중구	20.8%	0	8	1
2	서울 강남구	15.0%	0	3	1
3	인천 중구	14.1%	0	2	1
4	서울 종로구	5.7%	0	15	0
5	서울 용산구	5.0%	1	31	0
6	서울 서초구	4.8%	0	2	1
7	서울 송파구	3.8%	0	0	0
8	서울 영등포구	3.1%	0	1	1
9	서울 마포구	2.9%	0	2	1
10	서울 서대문구	2.4%	0	3	1

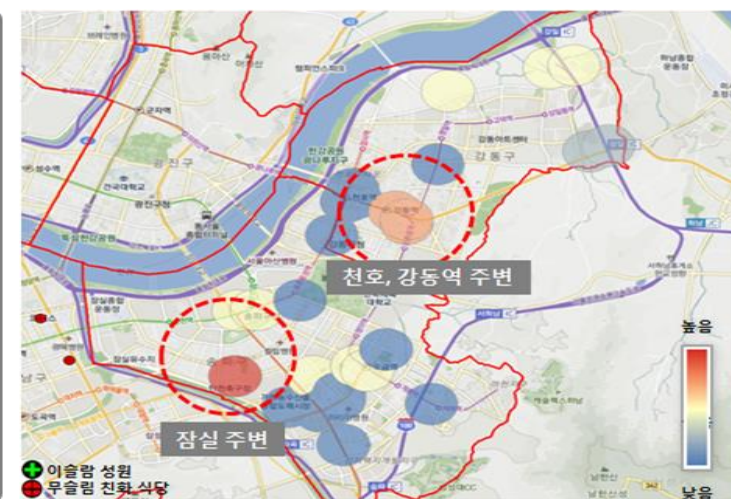


#### 서울 강남구

- 전체 점유율이 15%로 2위인 강남구의 경우, 역삼동과 논현동 주변에 무슬림 친화식당이 2곳에 불과하며, 이슬람 성원은 없고, 기도실만 1개 존재함
- 무슬림 방문방문이 높은 지역이므로 보다 많은 무슬림 친화식당이 필요할 것으로 보임

#### 서울 송파구

- 무슬림을 위한 시설이 전무한 상태로, 무슬림 친화식당과 기도실 등 시설 확충이 필요할 것으로 보임

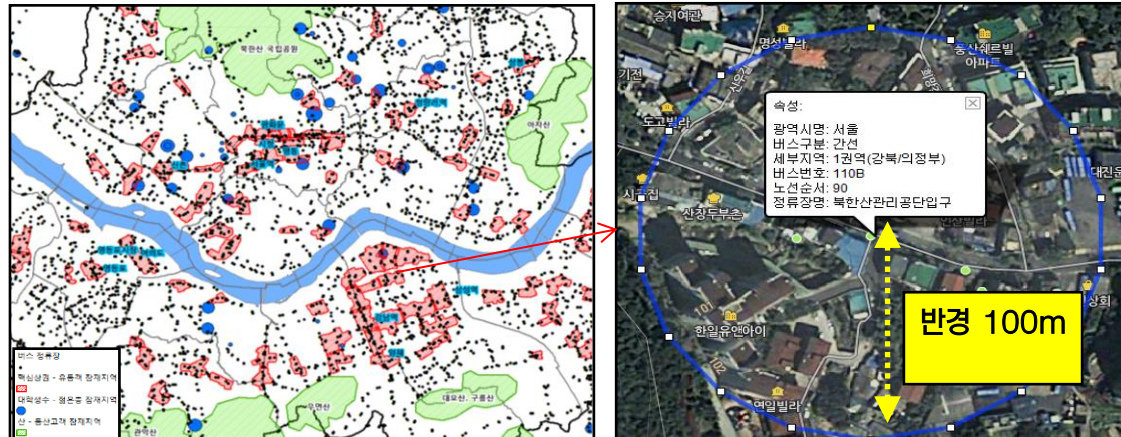




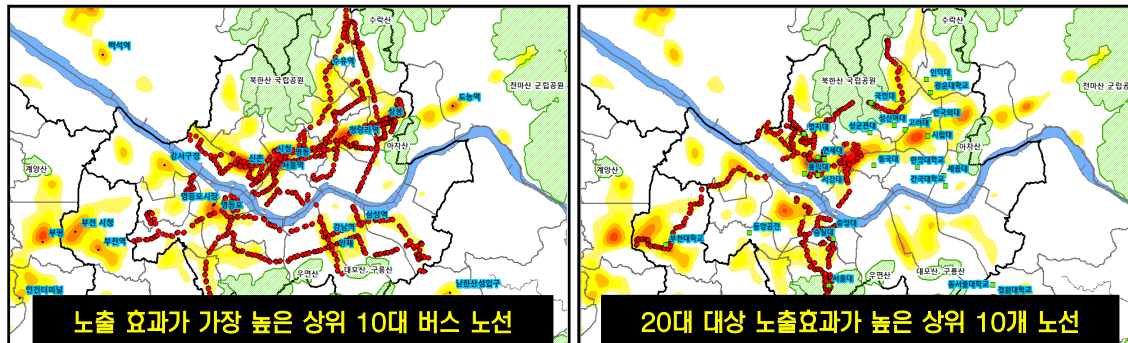
### III. B2B 활용 : 광고 효과성

## 버스 정류장 광고 주목도 효과성 분석을 위한 유동인구 분석

### 특정 상품의 Target 밀집지역 분류



110B 버스 노선 전체 평가 - 거주인구: 68,997명, 직장인: 27,316명, 대학생수: 6개 대학  
94,929 명 등산로 포함 비율: 10%

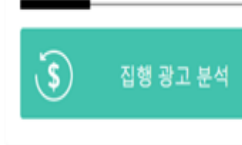


### 판매 잠재력 기반 영업 지원

#### 유동인구 분석



#### 광고효과 분석



#### 광고예측 분석



### 버스 정류장/노선별 광고 효과 분석

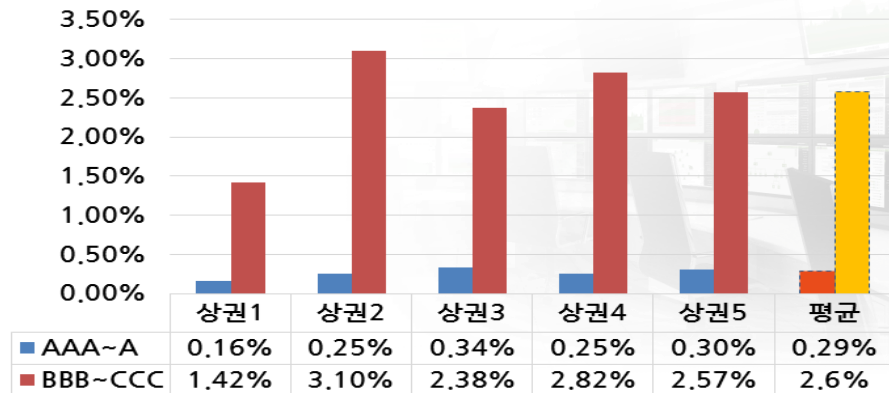


### III. B2B 활용: 자영업 신용도

## 신용평가 분야의 Big Data 활용한 상권정보, 자영업 대출

#### 자영업자 상권 유용성(사고율 기준) 분석

상권 1등급(SKT 분류) 내 자영업자의 사고율: 전체 평균의 절반 수준

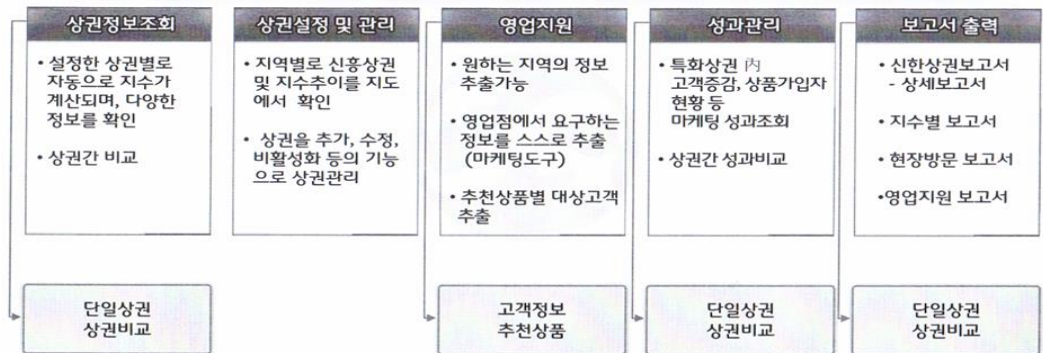


신용보증재단중앙회 검증 결과: 모형의 AR값 0.10이상 향상 가능

sampling	기존 모형	SKT 모형
1	0.626	0.758
10	0.619	0.724
평균	0.626	0.739

#### 국내 은행 구축 및 활용 사례

S은행 '특화 상권관리 시스템' : 2013~2017년 지속적 운영 중

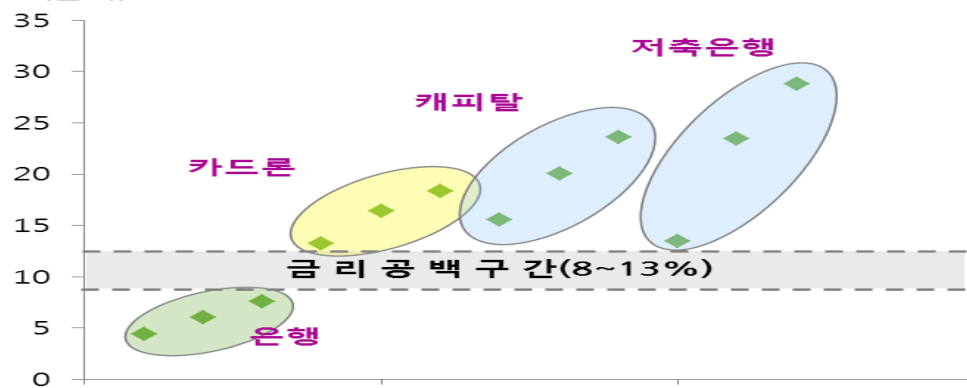
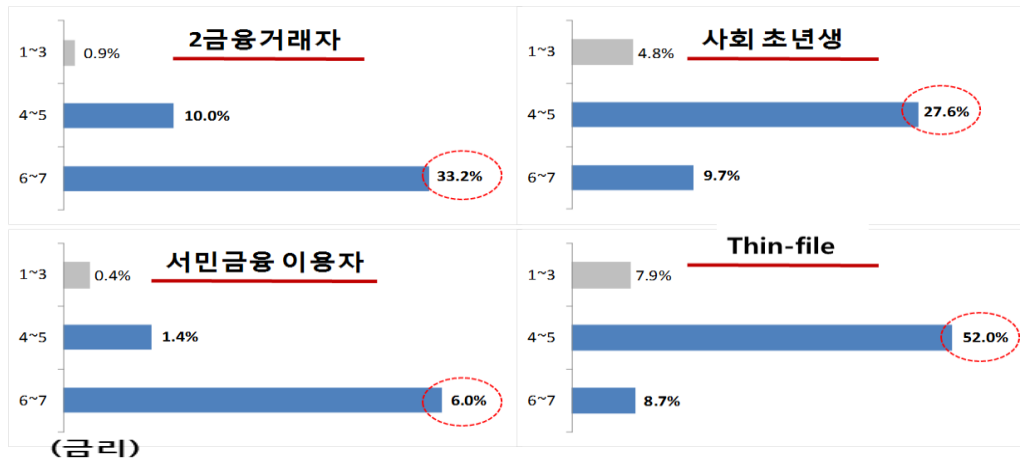




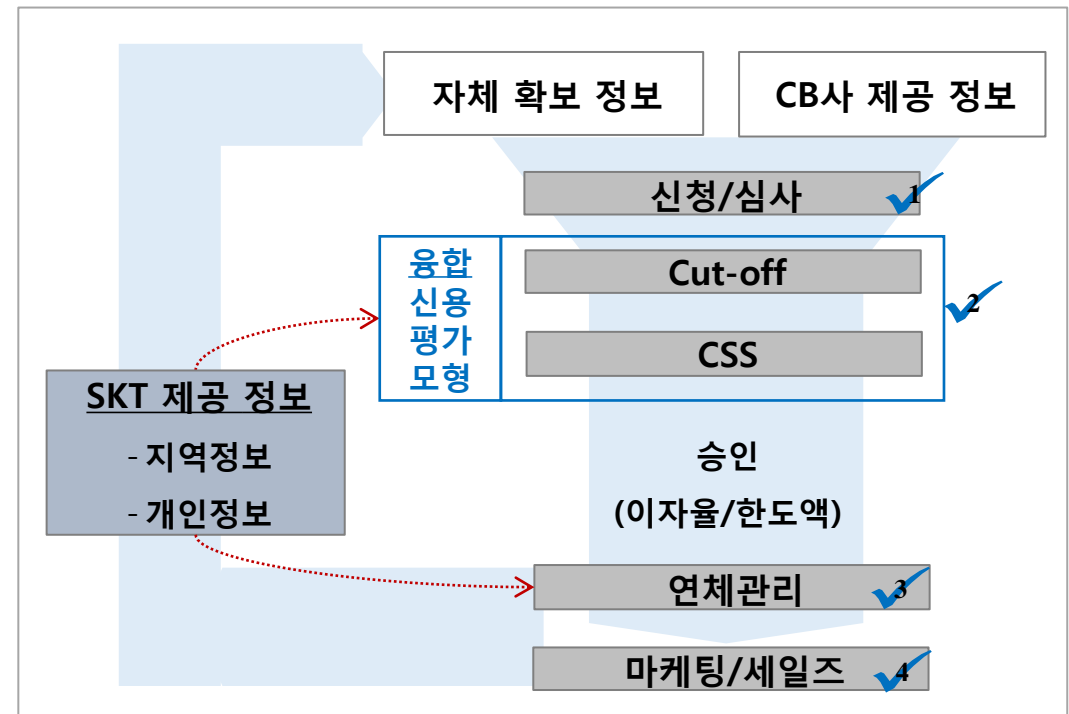
## II. B2B 활용 : 개인신용

Thin Filer : 개인동의기반 이자혜택, 대출혜택 부여

### 중신용 계층의 금융 서비스 불이익 해소 필요



### Telco Score 용 요약정보 제공



### III. B2B 활용: CCTV 입지

#### CCTV 설치 요구 등의 민원 응대: 여성/어린이 생활 안전 강화

##### CCTV 설치 조건 정의

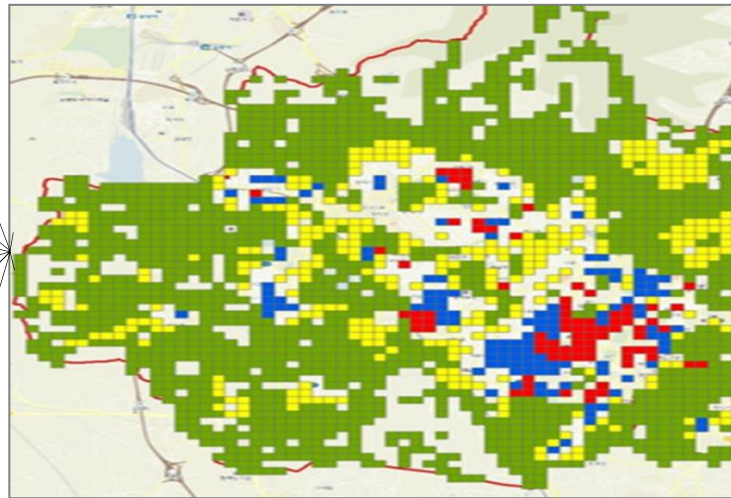
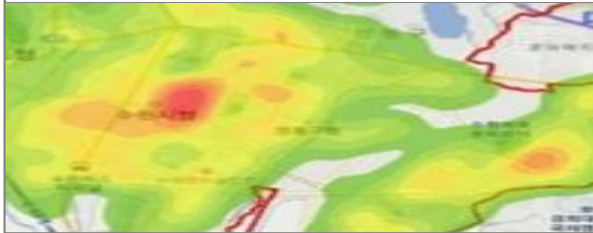
현 CCTV 설치 장소



CCTV 설치 민원 위치



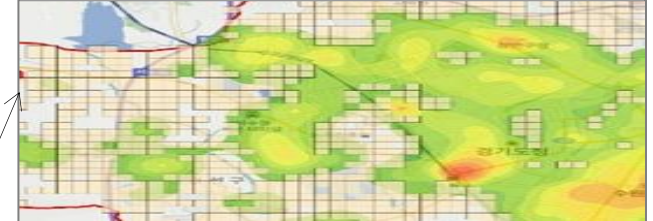
유동/주거인구



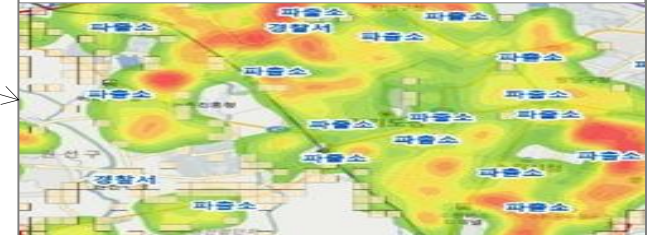
- 인구 많고 CCTV 없는 곳
- 인구 많고 CCTV 있는 곳
- 인구 적고 CCTV 없는 곳
- 인구 적고 CCTV 있는 곳

##### 민원 지역 대상 우선순위 부여

유동인구 100명 이상 제외



지구대 인근 지역 제외



여성/청소년 많은 지역



### III. B2B 활용: 야구장 입지

#### 새 야구장 입지 선정 후보지 대상 행안부/국토부/통신사 데이터 분석

##### 마산종합운동장

→ 1순위



구분	주거인구	직장인구
15 Min	412,540	189,407
30 Min	605,786	179,799
45 Min	462,112	162,003
60 Min	853,513	324,505

##### 창원종합운동장

→ 2순위



구분	주거인구	직장인구
15 Min	60,538	15,030
30 Min	510,673	247,052
45 Min	619,986	150,912
60 Min	555,502	265,495

##### 진해 육군대학 부지

→ 3순위



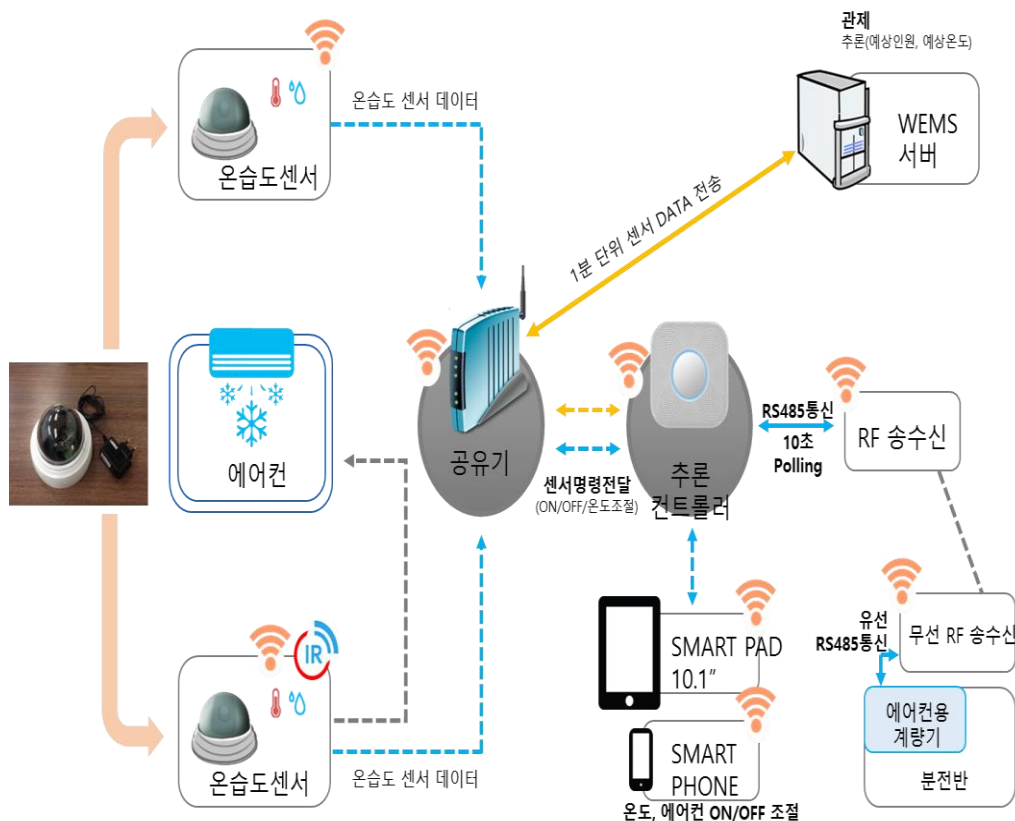
구분	주거인구	직장인구
15 Min	70,778	11,798
30 Min	209,766	121,730
45 Min	833,425	234,118
60 Min	438,624	227,880



## II. B2B 활용 사례 : 에너지 절감

### 이용 패턴을 활용한 에너지 절감

#### 소상공인(24시 편의점 등) 에너지 절감



#### S\_ARIMA Gravity Model

- 기상상태 및 월/요일/시간대에 따른 해당 지역, 업종의 총 잠재고객 수 및 총 매출 계산에 활용
- 계산된 잠재고객 수 및 총 매출을 이용, 대상 매장의 시장점유율을 계산하여 매출에 영향을 주지 않는 에너지 절감 모델 개발, 적용에 활용: 고객사 POS 정보를 감안하여 필요한 경우 성/연령에 따른 잠재고객 및 매출 규모 산정 가능

#### 시장잠재력 예측: 중력(Gravity) 모델

$$PNC_{SKT}(x) = \left[ \sum_{i=1}^3 ((NR_i W_r + NW_i W_w) P_{ij} / (1Y1M)) + (NI_i W_i P_{ij}) \right] \alpha_{avg} \beta_{avg}$$

$$PNC_{NCODI}(x) = \sum_{y=current\ year-4}^{current\ year-1} \left( \begin{matrix} \text{If Today is Normal Day : } NVC_y(\text{Compare to } N_{weather}) \\ \text{If Today is Special Day : } NVCS_y(\text{Compare to } N_{special}) \end{matrix} \right) \left( \begin{matrix} \alpha_{weather} \\ \beta_{weather} \\ \delta_{weather} \end{matrix} \right) + C_{avg}$$

#### POS 및 기상데이터 분석: Seasonal ARIMA

- 기상상태와 고객사의 POS에 있는 고객 수 및 매출액을 계산하여 기상상태 및 월/요일/시간대에 따른 고객 수 및 매출간의 상관관계를 계산, 중력모델을 이용한 잠재시장 규모와 비교, 대상 매장의 시장점유율을 파악
- S-ARIMA는 계절적 변화를 감안하여 시계열 데이터 기반 예측이 가능한 방법론: 통계 방법론



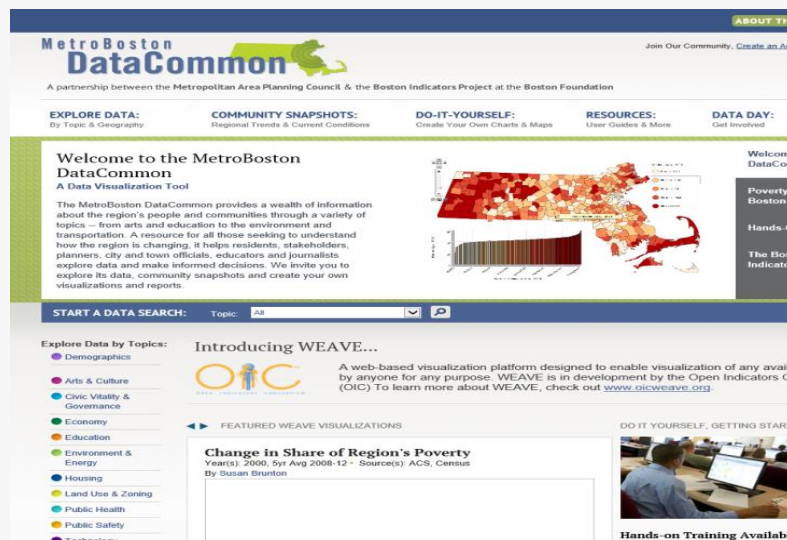


## #4. 데이터 미래 Data Driven

## IV. To Be : Data 융복합

### Keyword 1: 데이터의 융복합 - 활용 가치 확대 (데이터 결합, 활용 가치 발굴)

#### 시민 참여형 지도 작성 서비스



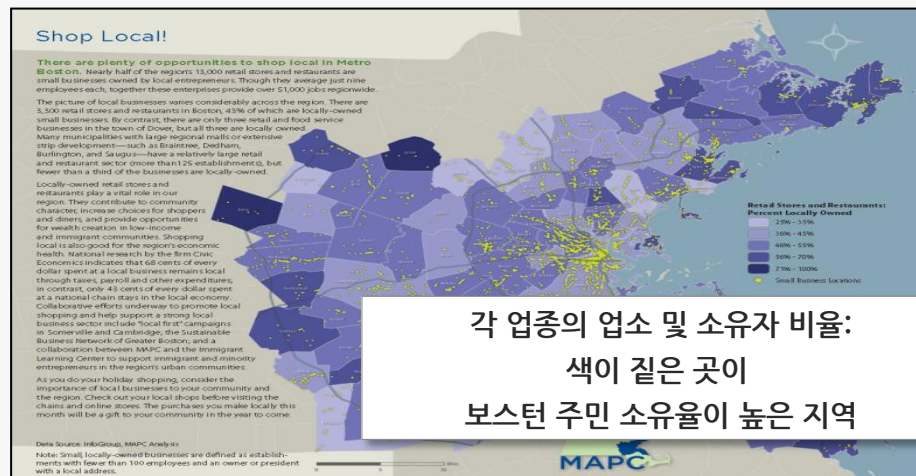
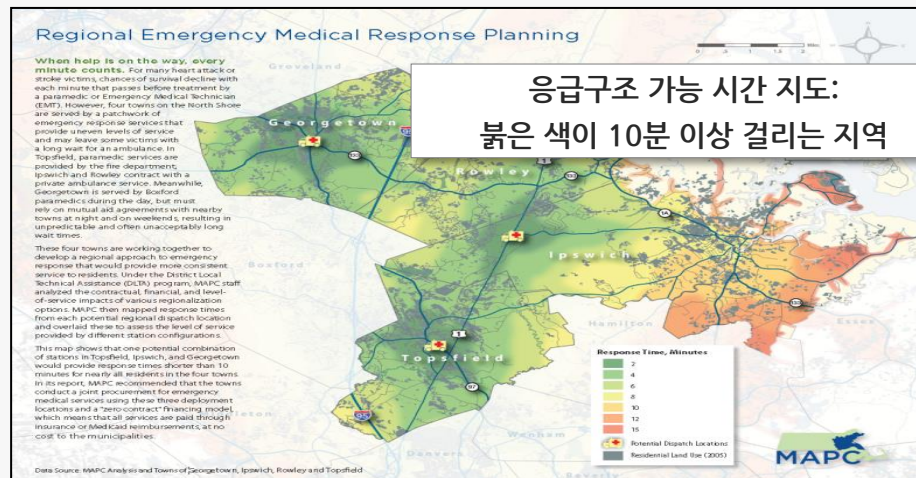
#### REGIONAL MAP GALLERY:

Explore MAPC's Annual Report and Calendar maps, illustrating conditions in the region and see whether the map you want may already exist.

- Demographics
- Arts & Culture
- Civic Vitality & Governance
- Economy
- Education
- Environment & Energy
- Housing
- Land Use & Zoning
- Public Health
- Public Safety
- Technology
- Transportation
- Geographic Boundaries

보스턴시에서 시가 보유하고 있는 데이터 공개, 목적 별로 다양한 방식으로 데이터 조회 가능, 시민들이 지도를 작성/공유할 수 있는 기능 제공

- 공공정보를 제공하는 방식: 정보 공개 및 시민 참여형 서비스
- 시 정부 내 각 부서에서 정보를 제공하고, 해당 지도를 이용하여 정책을 홍보하고 정보를 제공하는 형태로 운영: 대시민 정책 및 정보 제공 채널

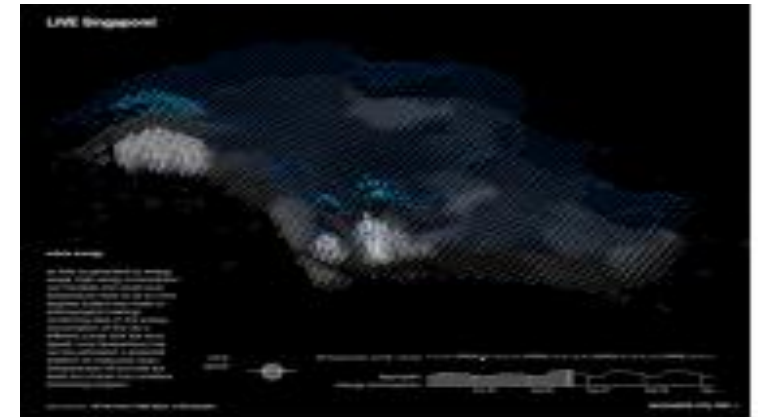
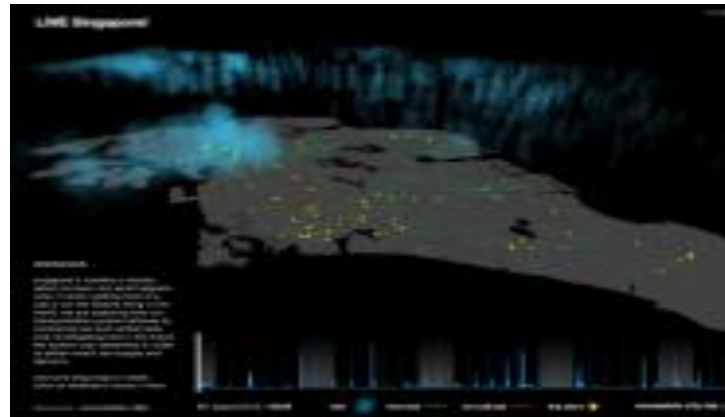
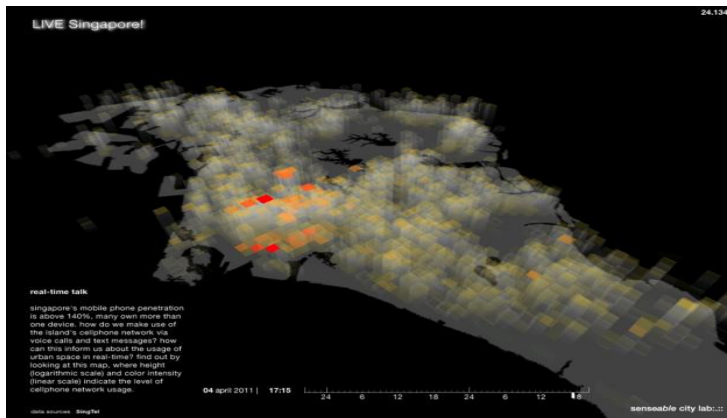




## IV. To Be: 실시간 IoT 데이터

### Key word 2: Real Time - IoT기반 데이터 분석 (실시간 데이터 Visualization)

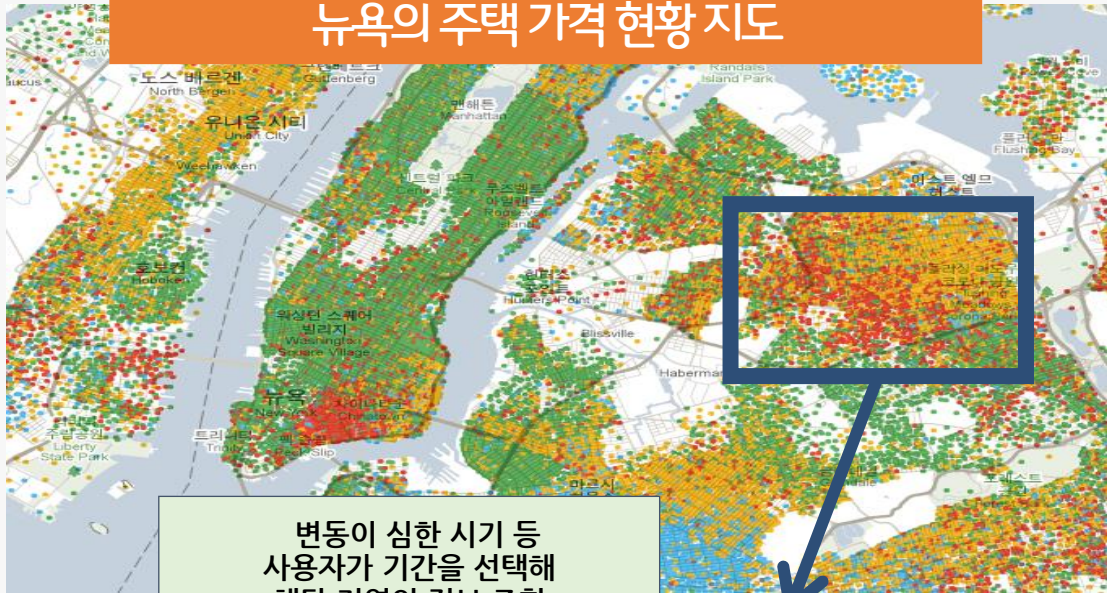
- 휴대폰 데이터 사용량을 시공간적으로 분석, Real Time으로 인구이동 파악, 도시/상권관리
- 기후/기온의 변화에 따른 에너지 사용량의 변화 파악: 온도 상승 시 에너지 제어
- 우천 시 기상정보, GPS 위치정보 등을 이용 일반차량, 택시 이용 현황 분석, 정보 제공
- 축제/행사/도로통제 또는 교통사고 등 이벤트 발생 시 교통 상황 분석, 우회로 제공



## IV. To Be : 시계열 Data 처리/분석

Keyword 3 : 시계열 분석 – 데이터의 생성/관리/분석 체계 구축 및 활용 (시계열 패턴 분석)

뉴욕의 주택 가격 현황 지도



변동이 심한 시기 등  
사용자가 기간을 선택해  
해당 지역의 정보 조회



퀸즈 지역의 주택가격 증감률 변화 차트

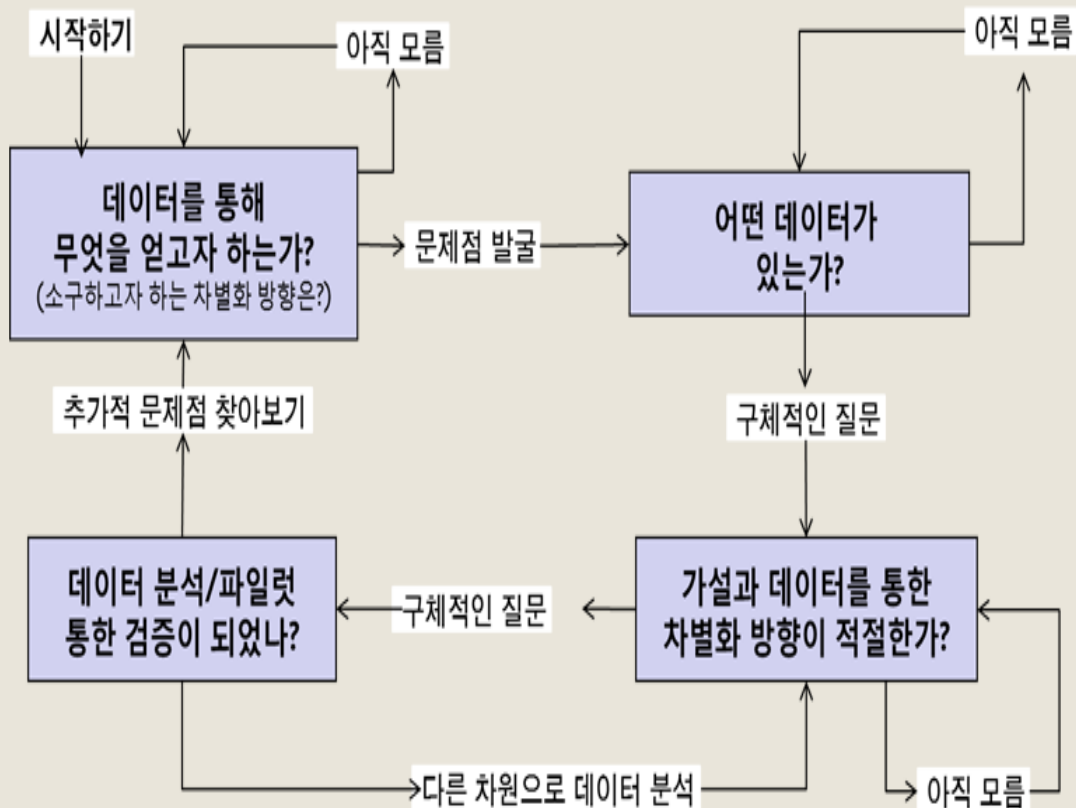
美 미네아폴리스, 경전철 건설에 따른  
토지 이용 변화 추이(5년 간)



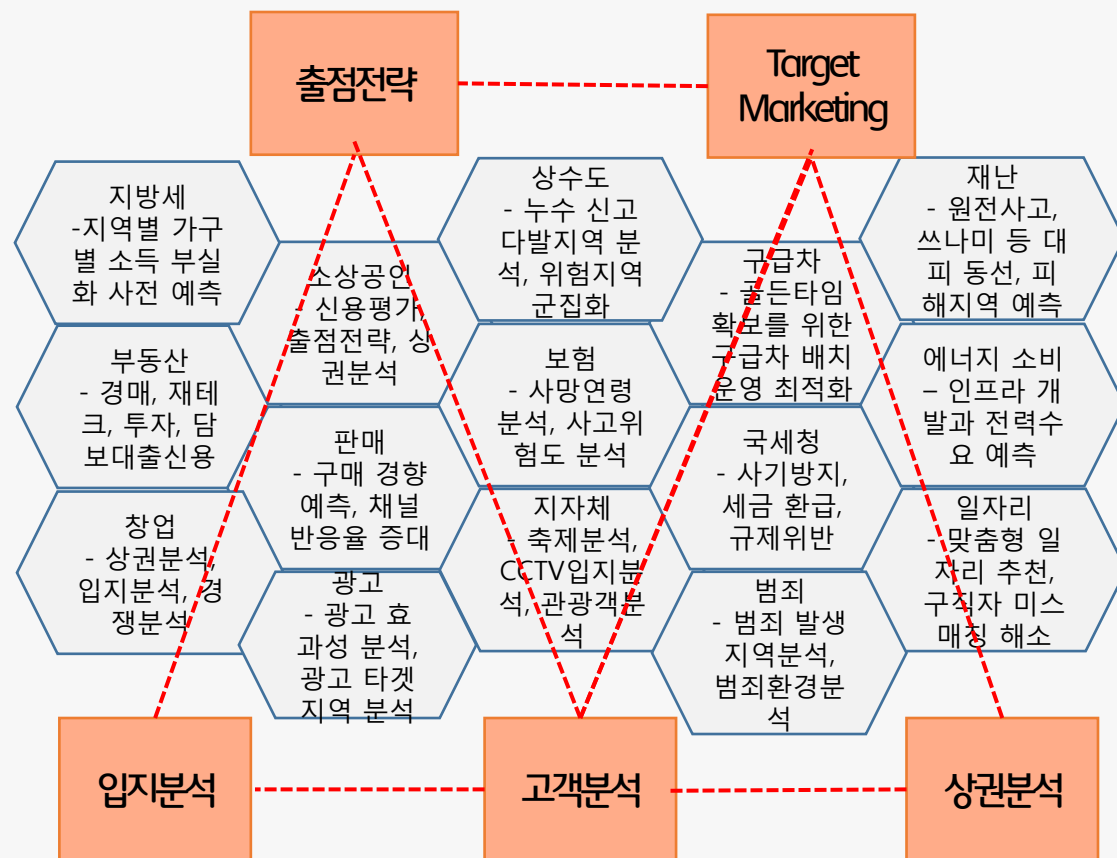
## IV. To Be : 4차산업으로 가는 예측 분석

4차산업의 기반은 빅데이터 분석에서 출발, 어떻게 시작할 것인가? 함께!!!

### 4차 산업 혁명, 함께 찾는 데이터 활용 아이디어



### 고객 분석을 통한 산업 분야별 Success Story 발굴





## 부서별 빅데이터 후보과제 목록

부서명	과제명	내용	관련데이터
사회복지과	효율적인노인의료시설, 재가노인복지시설운영	-효율적인 노인의료시설 및 재가 노인 시설운영관리 -시설이용자들에게 체계적인 정보제공으로 이용 편의제공 -장애인 밀집 지역 대상 복지사 배치 효율화, 장애인 대상 택배 배달, 가정 전등 교체 등 연계	-노인의료시설현황 -재가노인복지시설현황 -이통사 85세이상 노인 체류지 및 유동인구 -이통사 장애인요금제 가입 분포 데이터
도시정책과	무인택배함 적정 설치위치분석	-단독주택 및 1인가구 거주 밀집지역등 무인택배함이 우선적으로 필요할 것으로 예상되는 지역분석	-주택유형 -거주인구및가구형태등 -이통사 체류지 인구
안전총괄과	생활안전지도 자료 보완으로 활용성 증대	-생활 안전지도의 교통안전지도, 재난안전지도, 치안안전지도, 여성밤길안전지도, 노인건강지도 등 분야별 유동인구 보완 -폭염에 대한 대피안내	-이동통신 사용량에 따른 유동인구등 -이통사 연령대별 유동인구 등
문화관광과	지역관광 자원개발	-관광객의 선호 및 성향파악으로 관광정책 수립지원 -외국인 관광객 선호 및 성향파악으로 관광정책 수립지원 -축제 분석을 통한 축제 경제성 분석 -연계관광지 분석을 통한 연계 관광정책 수립지원	-관광객소비패턴분석 (이통사 상관분석 매출액 데이터) -관광지 방문객 분석 (이통사 유동인구, 유입지 데이터) -관광객이동경로파악 (교통카드사용정보등) -관광사이트 및 소셜 데이터접속자분석 -이용자후기파악 (블로그,카페,소셜네트워크등) -이통사 외국인 로밍 데이터
세정과	납세자의 세대구성에 맞는 홍보수단 및 편의시책결정 (선정및평가)	-기존 : 홍보매체 선정시 세대별 호응수준을 고려하여 효과성을 높이는 전략적 접근 방안을 모색하지 않고 전세대를 대상으로 일괄 적용 -개선 : 세대별 IT 기기 활용수준, 이해도등을 분석하여 이를 바탕으로 세대별로 홍보	-세대별 홍보수단 선호도분석자료 (IT기기, 안내문, 전화 등) -세대별인구구성현황 -연령대별 납세자분포 (이통사 연령대별 거

기업지원과	기업체 증감현황분석	-관내기업체 증감현황을 분석하여 각분야 정책사업에 활용 -관내 창업/업종변경 등 소상공인 지원 정책에 활용	-공장등록·취소현황 -기업체통계조사 -사업자등록·폐업현황 -이통사 상관분석 데이터(인허가 가계정보)
교통정책과	장애인 콜택시 운영분석	-택시 대기시간을 단축하고 조건(계절, 시간대 등) 별로 적절한 운영 규모를 산출	-장애인 콜택시 운영 데이터 -이통사 대리운전 요청 콜 분석데이터
도로과	관내 도로명을 통한 주소개념 정립	-관내 도로명과 각 도로명의 길이, 시점, 중점등을 파악하여 위치찾기 지원 -관내 도로 증설, 신설 정책 수립	-도로명, 도로길이, 도로의 시점, 중점등 -이통사 GIS 데이터(TMAP) -이통사 T-MAP의 센서데이터를 활용 포트폴리오 위치 수집
생명농업기술센터	도시텃밭 계획수립	-시민농원 이용 시민 세대조사, 이용시민 지역조사 실시해 도시 텃밭 계획시 텃밭 적절히 분배	-시민 전체 거주 지역조사 -시민농원 이용시민 거주 지역조사 -시민농원 이용 세대 조사 등 -이통사 연령대별 유동인구데이터
상수도과	믿을수 있는 맑은물 관리	-수돗물수질검사결과 -먹는물수질검사안내 -지역별 상수도 수요대상 파악	-수돗물 수질검사 결과 매월 데이터 -수돗물 공급지역별 이통사 인구분포 데이터
도시정비과	도시정비사업 주민참여 확대 시스템 (홈페이지구축) 관내 도로 포트홀 파	-주민□관의 정비사업 정보교류 -SNS 실현으로 주민과 공공이 상호 소통하는 문화정착 -도시정비사업 최신정보를 신속□정확하게 제공	-도시정비사업 구역별기본계획 현황 -추진현황 -주민의견게시판

평생학습과	평생학습 자원현황제공	-관내 평생학습 자원현황을 제공함으로써 내 주변의 평생학습현황을 쉽게 알 수 있도록 제공	-평생 교육기관 주소 및 버스정거장, 버스번호 -평생학습 프로그램 현황 -평생교육 현장 실습 가능 기관 -평생학급 주 연령층 이동통신 상주인구데 이터
중앙도서관	도서관 장서 구성 및 특성화 정책 수립	-장서이용도 파악으로 장서선택 및 특성화 사업 지원 -도서관 이용자 체류지 분석을 통한 작은 도서관 신설 통합 등의 입지 선정	-자료구입 및 특성화 자료분야 -이용자의 유입지역 데이터를 이동사 체류 지데이터로 분석
차량등록 사업소	교통관련 정책수립지원	-각종 교통관련 정책수립 지원 -관내 공영주차장 입지 선정	-자동차(건설기계, 이륜차포함) 등록현황 -이동사 유동인구 데 이터, 상주인구데이터
도시정책과	시내버스 와이파이 운영 활성화 공공 와이파이 활성화	-관내 시내버스 와이파이 이용 효율 증대 -와이파이 시민 이용확대에 따른 시민 통 신요금 인하 효과 증대	-관내 시내버스 와이파이 운영 데이터 -이동사 유동인구 데 이터 -이동사 인터넷 이용 데이터
도시정책과	운동장 이용 연령 다양화	-운동장 이용 고객 연령을 파악, 연령 다양화를 통한 이용객 증대	-수업별 이용 연령 파악 -운동장 셔틀버스 이용 현황 -유동인구등
도시정책과	주택구매 수요에 맞는 맞춤형개발사업	-주택별(공동주택,단독주택,원룸등) 구매 수요를 활용한 맞춤형 주택개발	-주택 보유내역 (자가, 전세, 월세) -주택별 구매수요 등 -이동사 상주인구 데 이터

도시정책과	주민이 원하는 맞춤형 공공시설 건립	-지역별 제반여건 분석 후 지역에 맞는 공공시설건립	-공공시설물 분포 및 이용현황 -지역단위 세대별 가족 구성원 형태 -지역별 연령분포 등 -이동사 유동인구 등
경제활성화 전략본부, 균형발전사 업단	관내 기업현황 정보서비스	-관내 기업 경영정보등 현황이 체계화 되어 서비스 필요 -관내기업의 업종, 경영현황, 매출, 주요제품, 사업장 위치 등 데이터 활용 -서비스 제공시 기업이 사업전략 및 마케팅에 활용가능	-사업자 등록현황 -재무제표 현황 -근로자 현황 -이동사 유동인구, 유 입인구, 외국인 인구 등
경제활성화 전략본부, 균형발전사 업단	중소기업 지원 사업현황 서비스	-기업에게 체계적이고 종합적인 기업지원 사업현황 제공서비스 필요 -기업이 필요한 지원사업을 선택가능 -기업에게 적절한 시기에 맞춤형 지원 서비스 가능	-중앙부처 지원사업 현황(미래부, 문화부, 중기청 등) -광역시초 지자체 지원 사업현황 -산하기관 포함 -이동사 외부감사기업 정보 등 기업체 데이 터
경제활성화 전략본부, 균형발전사 업단	관내기업 창업 정보서비스	-관내기업의 창업한 지역정보 및 기업정보 -산업별, 지역별 창업 기업정보를 구축하여 창업 활성화 및 일자리 정책에 반영	-재단 보유 기업DB -통계청이 보유한 사업체 통계 -창업 관련 지원 사업 정보 등 -창업 아이템 활용가 능한 이동통신 데이터 제공
문화재단	미술관람객 개발 및 만족도 향상 방안	-관심 분야를 분석하여 관람객 증진 및 만족도 향상에 기여	-성별, 나이, 수입, 미술관방문횟수, 관심분야, 방문시간대 -이동사 미술관 주변 거주인구 유동인구 분 석 데이터
교육청소년 과	청소년 시설 이용 고객확대	-청소년 밀집지역을 분석하여 청소년 야외 행사 및 교육 사업 추진	-대중교통(카드) 이용 데이터 -이동사 청소년 인구 밀집지역 데이터

# Q & A

The background of the slide is a light gray illustration. It features a map of South Korea with several location pins. To the left of the map is a bar chart with three bars of increasing height. Below the map is a stylized city skyline with various building shapes. To the right of the map, there are some abstract shapes resembling balloons or clouds. The overall theme is data and geography.