

분석결과/사례공유

입주자 분석결과

협력 분석사례

서울시 분석사례

공모전 분석결과

입주신청 바로가기 >

공지사항 바로가기 >

빅데이터캠퍼스  
공식 유튜브 바로가기

서울시 분석사례

HOME > 분석결과/사례공유 > 서울시 분석사례



꿈나무카드 분석 보고서는 경희대학교(산학협력단 SK청년비상 빅리더팀, 지도교수 : 전종식 교수)에서 2017년 2월 ~ 2017년 11월에 사각지대에 놓인 결식 우려 아동을 찾아 결식아동 발생을 예방하고, 지역실정에 맞는 급식지원 인프라 확충과 급식지원 아동의 이용 편의성을 도모하기 위하여 아동급식카드 제도 개선과 균형있는 영향섭취 방안 마련을 위한 데이터 분석한 내용을 서울시에 재구성하여 웹 콘텐츠화 하였습니다.

저소득 어린이 급식카드 사용 분석

1. [“활용 데이터 알아보기”](#)
2. [“토픽모델링 분석결과”](#)
3. [“가맹점 분석결과 알아보기”](#)
4. [“사용자 분석결과 알아보기”](#)
5. [“분석결과와 시사점”](#)
6. [“향후 발전 방향”](#)

꿈나무카드

· 꿈나무카드는 1998년에 시작된 결식아동 급식지원 정책의 여러 방안 중 하나로, 급식지원 대상 아동 중 희망하는 아동은 카드를 발급 받아 지정된 가맹점에서 자율적으로 식사할 수 있습니다.

· 결식이 우려되는 아동에게 지역 실정과 아동 욕구에 맞는 급식을 효율적으로 제공하여 건강하고 행복하게 성장할 수 있는 여건을 조성하기 위한 서울시의 아동급식사업으로 연간 45억 원의 예산을 16,732명에게 지원하고 있습니다.



Chapter.1 “활용 데이터 알아보기”

“분석 활용 데이터 및 분석 도구”

본격적으로 분석에 활용할 데이터로 <꿈나무카드 이용내역, 꿈나무카드 거래 데이터, 편의점 구매내역>을 선정하였습니다.

데이터 명분	보유기관	데이터 설명	제공 컬럼명 요약	기준시점
꿈나무카드 이용내역	우리은행	서울시 꿈나무카드 결제 내역 (666,711건)	자치구/카드번호/가맹점명/ 거래금액/거래일시/가맹점주소	2015.7.~2016.12월
서울시내 편의점 점포 현황	편의점 3사 (CU,GS25, 세븐일레븐)	서울 시내 편의점 3사 점포 현황 (6,073건)	자치구/점포명/업종/주소/ 전화번호	2016.6.10
편의점 꿈나무카드 구매내역 분석데이터	CU	꿈나무 카드에 의한 결제 품목 분류별 매출 비율	연월/요일/시간/ 행정동품목분류/매출 비율	2016.7.~2017.6월
자치구별 급식제공기관 현황		자치구별 꿈나무카드 가맹점 및 단체급식소 현황(1,941건/467건)	자치구/상호명/업종/주소/ 전화번호	2016.12월 말
자치구별 아동급식 학년별 현황	아동가족 담당관	서울시 아동급식 지원대상자 취학구분별 인원 현황	자치구/취학구분/인원수	2013~2015년 3개 연도 12월 말
자치구별 아동급식 시비 보조금 정산내역		서울시 연도별 아동급식 예산 집행 금액 내역	자치구/시비/자치구비	2013~2015년

데이터 전처리에 필요한 분석이용 Tool은 “R, JAVA, 엑셀”이었으며, 토픽 모델링과 이상치 분석, 데이터 시각화를 통하여 가  
맹점, 사용자 데이터 비식별화를 하였습니다.

테이블명	키 값	데이터 수	테이블별 컬럼
결제처	가맹점명+주소	1,840	가맹점+주소, 가맹점명, 업종분류, 주소
이용자	카드번호	13,365	카드번호, 고객명
거래		666,711	고객명, 카드번호, 거래일시, 가맹점명, 거래금액, 주소

Chapter.2 “토픽모델링 분석결과”

“데이터 수집 및 형태소 분석”

결식아동급식에 대한 이슈 확인을 위해 토픽모델링을 분석방법으로 정하고 크롤링을 통해 언론 기사와 사용자 후기 데이터를  
수집하였습니다. 수집된 문서들에서 전처리 작업인 형태소 분석을 통한 자연어 처리를 진행 하였습니다.

데이터 수집

결식 아동 급식과 관련된 네이버 뉴스기사 및  
카페 후기 문서 총 6,116건 수집



<JAVA의 'Jsoup' 라이브러리 이용>

형태소 분석

토픽 모델링 분석을 위해 수집된 문서의  
content를 형태소 분석 (명사구 추출)

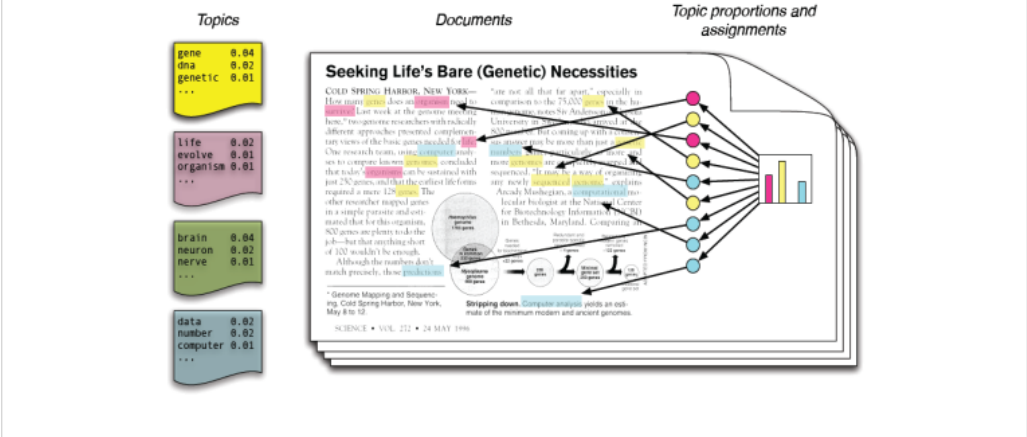


<R의 'KoNLP' package 이용>

형태소 분석 결과를 통합하여 LDA(Latent Dirichlet Allocation) 알고리즘 기반의 토픽모델링을 적용한 토픽 분석을 수행하  
였고 이를 통해 결식 아동 급식과 관련된 주요 키워드와 이슈들을 파악하였습니다.

토픽 모델링

대량의 문서집합에서 ‘단어’나 ‘용어’들과 같은 공통 키워드들을 바탕으로 문서를 주제별로 분류해 주는 기법



“Word Cloud”

비정형 빅데이터 분석을 통해 꿈나무카드 현황을 파악하고 결식아동들이 건강하게 성장할 수 있는 환경을 조성하는 것을 목표 분석을 진행하였습니다.



“R의 LDA 라이브러리 분석결과”

혜택, 동사무소, 서류, 가맹점 등 꿈나무카드의 신청과 사용에 대한 논의가 활발하였고, 예산, 관리, 대책, 복지, 교육 등 정부 정책적 차원의 문제가 논의 되었으며, 건강, 먹거리, 불균형우려 등 영양적 문제가 제시 되었음을 알 수 있었습니다.

N	토픽 모델링 결과 (결식아동 급식)						키워드
	전체	Topic1	Topic2	Topic3	Topic4	Topic5	
	급식	급식	지원	예산	결식아동	단체	지역 교육 저소득 예산 관리 대책 문제
	지원	결식아동	희망	사업	지역	관리	
	예산	정부	대책	무상	한나라당	국회	
	복지	예산	복지	대표	봉사	대통령	
	결식아동	교육	민주당	문제	가정	어린이	
	사업	학교	기부	서울	올해	밥	
	아동	방학	의원	저소득	급식비	겨울	
	지역	정책	카드	경제	학생	초등학교	

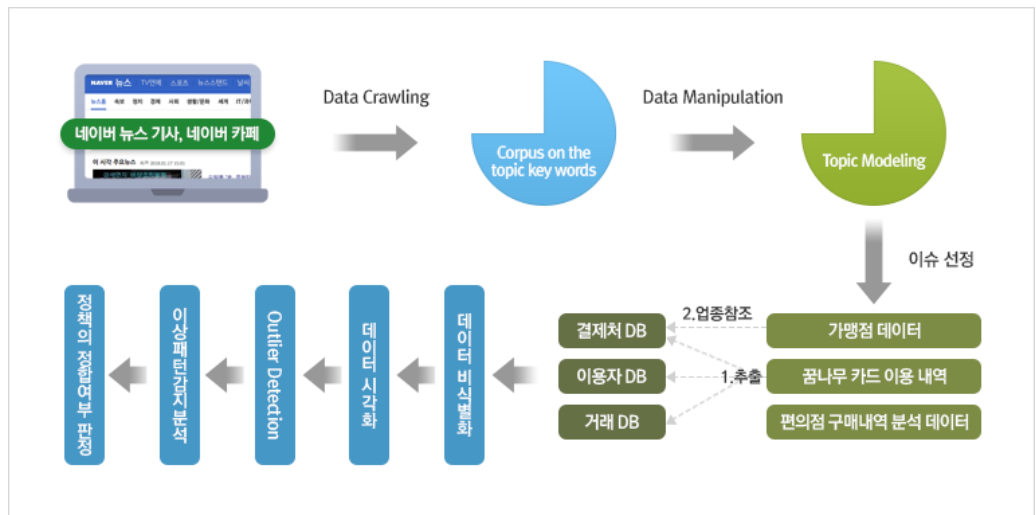
<지역,교육,저소득,예산 등 정책적 차원문제 논의>

토픽 모델링 결과 (꿈나무 카드)						키워드
전체	Topic1	Topic2	Topic3	Topic4	Topic5	결식아동 건강 먹거리 우려 불균형 문제 취약 문제
도시락	급식	도시락	아이들	아동	지원	
지원	센터	문제	카드	배달	아동	
급식	제공	명절	음식	급식	결식아동	
아동	사용	계획	이용	복지	가정	
아이들	주인	사회	양천구	건강	결식	
배달	지역	먹거리	연휴	엄마	청소년	
카드	취약	우려	식사	추석	서울시	
센터	행복	불균형	대상	운영	가구	

<건강,먹거리,불균형,우려 등 영양적 문제 제시>

토픽 모델링 결과 (카페 후기)						키워드
전체	Topic1	Topic2	Topic3	Topic4	Topic5	혜택 동사무소 서류 대상 가맹점 편의점 도시락 중국집
카드	서류	사용	카드	도시락	신청	
사용	꿈나무	지원	센터	중국집	방학	
가구	평일	미성년자	동사무소	부모님	할인	
도시락	혜택	지역	배달	편의점	혜택	
지원	저녁	장애인	생활	배달	편의점	
신청	정보	주인	중국집	가정	구매	
꿈나무	아동	구매	대상	대상	시간	
방학	동사무소	아이들	구매	사회복지	급식	

<혜택,동사무소,서류 등 꿈나무카드의 신청과 사용>

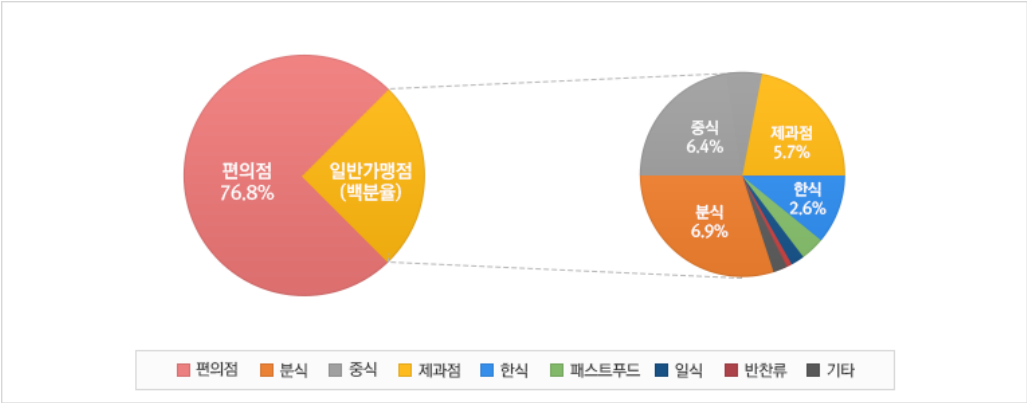


<분석 체계도>

### Chapter.3 “가맹점 분석결과 알아보기”

꿈나무 카드 사용자들이 이용할 수 있는 가맹점 수와 분포를 확인한 결과 전체 가맹점의 76.8%가 편의점이었고, 제과점을 포함할 경우 82.5%로, 일반 음식점에 대한 가맹점 확대 등 다양성 확보 방안이 필요한 것으로 보여집니다.

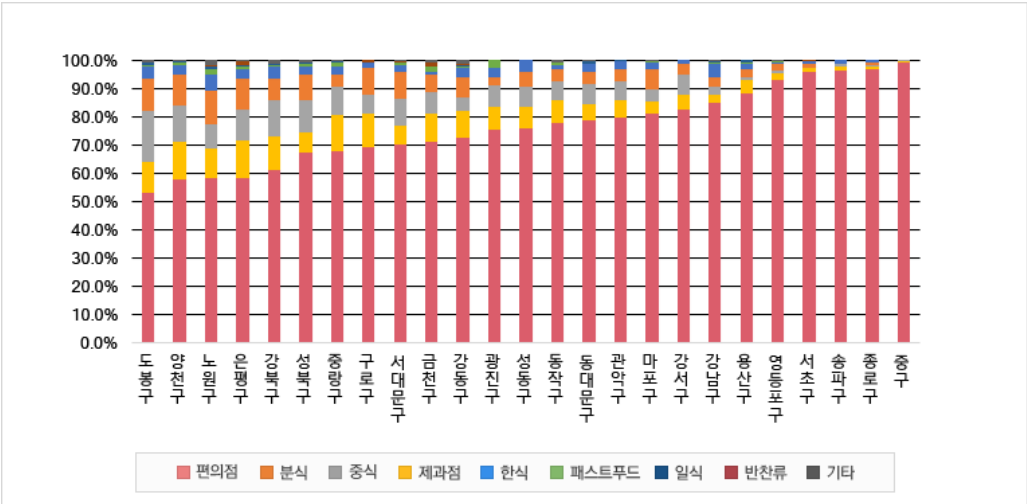
“가맹점 구성 및 업종 비율 (편의점 및 일반 가맹점)”



	편의점	분식	중식	제과점	한식	패스트푸드	일식	반찬류	기타	총계
개수	6076	543	508	449	208	52	29	15	33	7913
비율	76.8%	6.9%	6.4%	5.7%	2.6%	0.7%	0.4%	0.2%	0.4%	100.0%

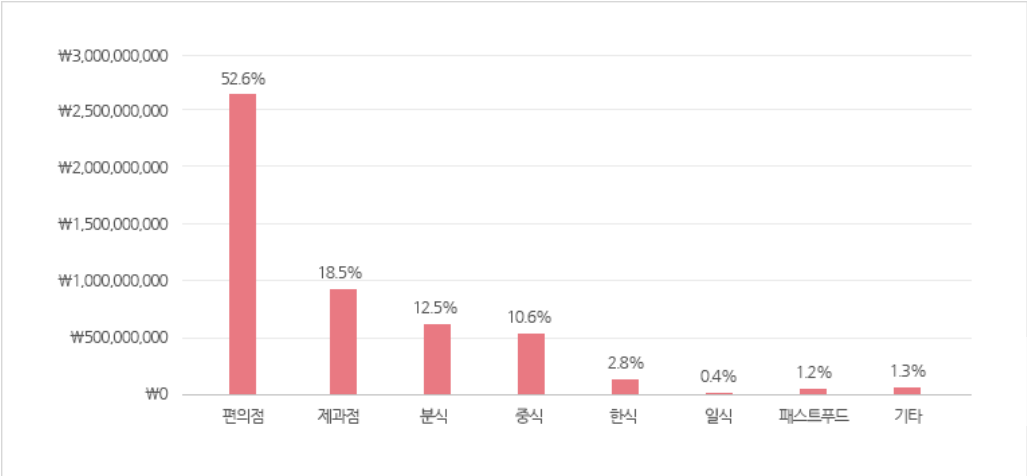
“자치구별 가맹점 업종 비율”

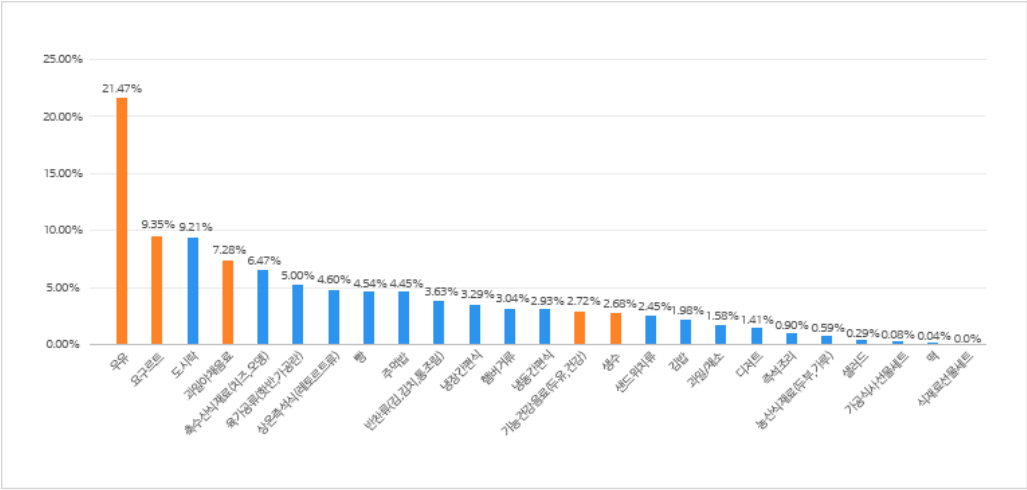
구별로 살펴봐도 일반 가맹점에 비해 편의점의 수가 많습니다 도봉구, 노원구 등도 편의점 비율이 50%를 초과하고 영등포구, 서초구, 송파구, 중로구, 중구는 편의점 비율이 90%를 초과합니다.  
가맹점 업종[한국표준산업분류 10차 개정본(통계청, ‘17.7.1. 시행)]을 구분하고 업종분포를 도출한 결과, 중구 등 일부 자치구에서 편의점 비율이 90% 이상인 것으로 나타나 일반가맹점 확대가 필요한 것으로 분석 되었습니다.



“업종별 매출액 및 가맹점 수 비율 비교”

2015년 7월부터 2016년 12월말까지의 꿈나무카드 사용자 매출데이터로 전체 가맹점의 매출 규모는 편의점이 52.6%로 가장 높았고, 다음으로 제과점, 분식, 중식 순이었으며, 편의점(CU)에서 꿈나무카드 사용 구매품목 비율은 우유가 21.47%로 가장 높았고, 다음으로 요구르트, 도시락 순이었습니다.

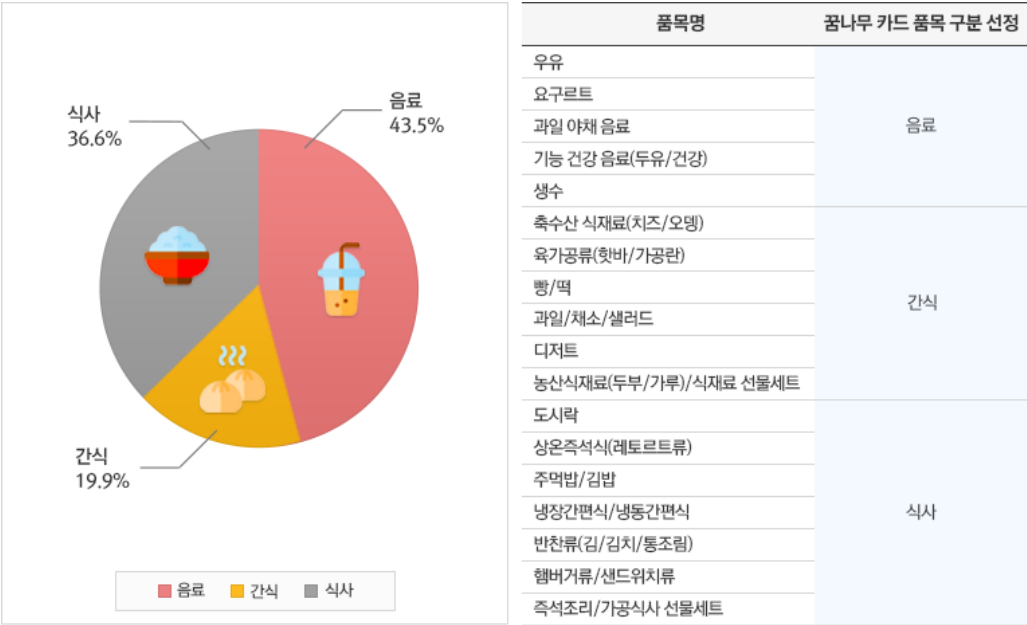




〈편의점 CU 편의점 카드 구매 분류별 월평균 매출 비율 상세 내역〉

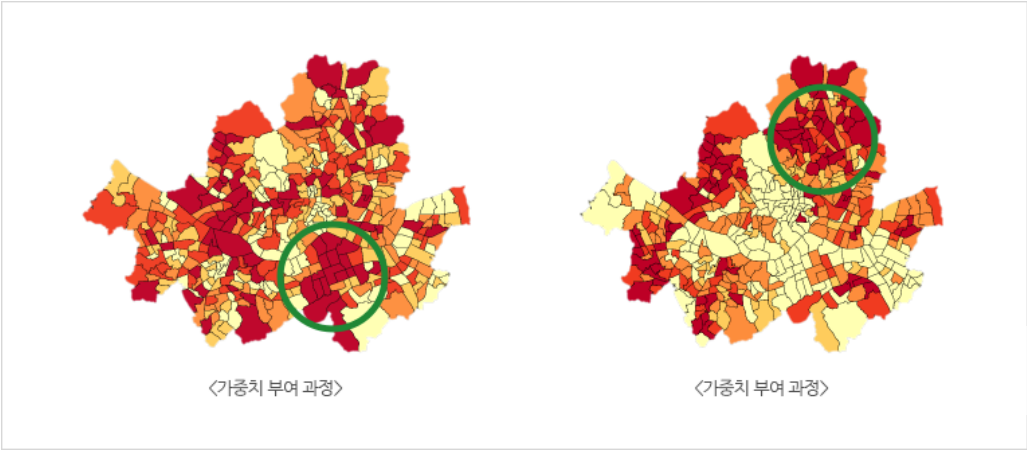
“편의점 품목 구분별 매출 비율”

매출 비율로 음료가 43.5%, 식사가 36.6%, 간식이 19.9%로 도시락류(식사류)/간식류와 우유(음료류)를 같이 섭취하지 않는 경우 영양적으로 불균형 섭취 가능성이 높을 것으로 보여집니다.



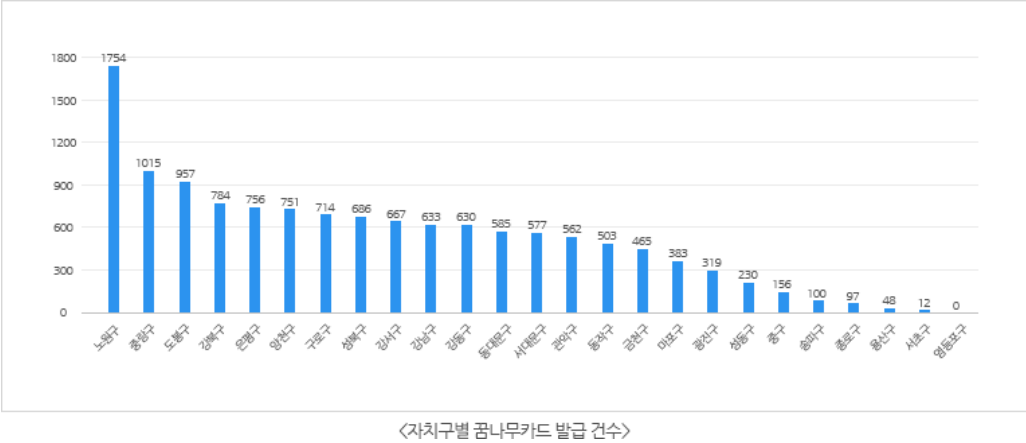
“행정구역별 가맹점 현황”

서울시 424개 행정동을 가맹점의 수에 따라 5분위로 나누어 온도지도로 시각화한 결과 각 구청과 행정동 별 편의점 카드 가맹점에 대한 정책 현황을 알 수 있습니다.  
강남지역은 편의점, 강북지역은 일반가맹점 비율이 높았습니다.



자치구별 공나무카드 결제를 기준으로 공나무 카드 발급건수를 계산한 결과, 전체 인원을 13,365명으로 정의하였고, 이중 노원구의 발급건수가 1,754명으로 가장 많은 것으로 분석 되었습니다. (영등포구는 자료가 없어 분석에서는 제외)

카드 사용자에 대한 명확한 데이터가 없으므로 2015년 7월~2016년 12월까지의 거래내역에서 사용된 카드 한 장을 사용자 1명으로 가정하고 분석을 진행하여, 전체 공나무카드 사용자는 13,365명으로 가정하고 공나무카드 거래 내역의 결제 시간과 결제처에 따른 업종 분류 데이터를 이용하였습니다.



“분석대상 정의”

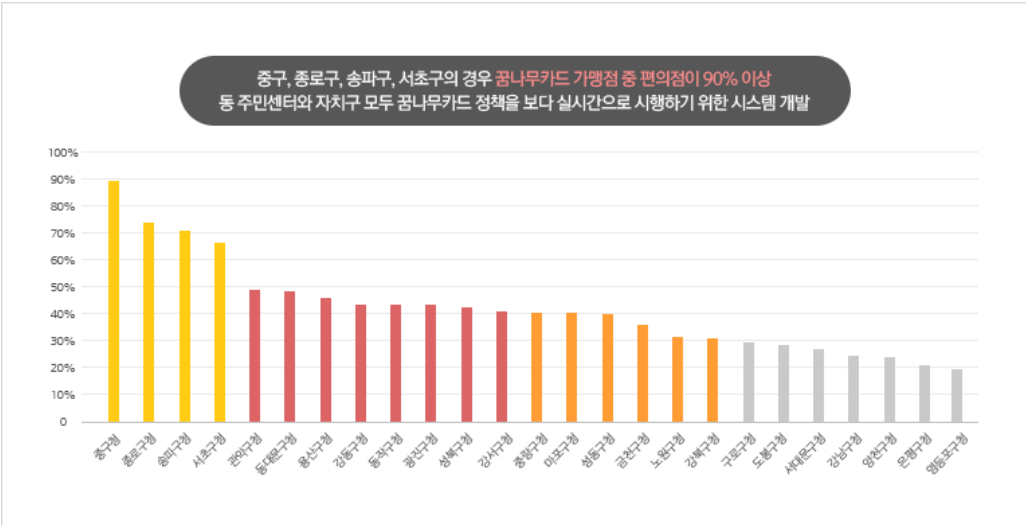
사용자 이상패턴을 감지하기 위해 Outlier Detection 알고리즘을 활용하였습니다. 공나무카드 통합시스템의 목적에 맞게 ‘이상패턴 감지 시스템’으로 명명하고, 알고리즘에 의해 관심 필요 대상 집단을 3가지 유형으로 구분하여 아래와 같이 정의 하였습니다.

이상패턴 사용자 집단	내용(2015년 07월~2016년 12월의 18개월 간 공나무카드 전체 이용 내역)
1 새벽 시간대 위주 사용자	공나무카드 전체 이용시간 중 새벽 시간대(자정 0시~새벽 6시)에 50% 이상 사용한 사용자
2 편의점 위주 사용자	공나무카드 전체 결제 중 편의점에서 90% 이상 사용한 사용자
3 식사 시간대 비 사용자	18개월 동안 식사 시간대 결제비율 50% 미만 사용자 (아침 07~11시/점심 11~15시/저녁 17~21시)

〈관심 필요 대상(이상패턴 사용자-중복 포함) 구분〉

“자치구별 관심 필요 대상자 비율 현황”

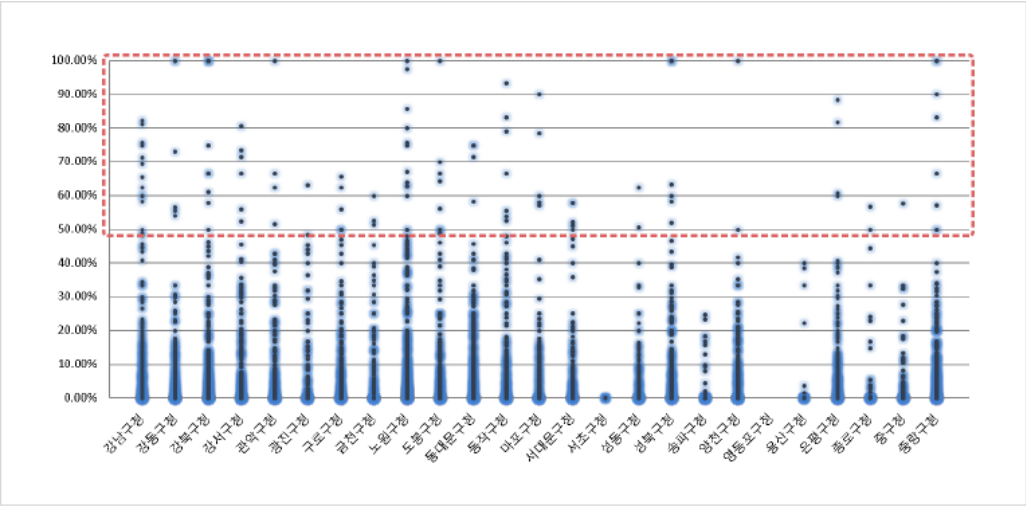
자치구별 공나무카드 가맹점과 사용자에 대한 분석결과, 관심 필요 대상 비율의 차이가 커 자치구 실정에 맞게 모니터링을 통해 피드백이 가능한 시스템이 필요할 것으로 분석 되었습니다.



“새벽시간대 위주 사용자”

각 구청에 등록된 개별 공나무카드가 자정(0시)부터 새벽(6시) 사이에 결제한 횟수를 전체 결제 횟수로 나눈 비율 값으로 산포도를 그렸습니다.

공나무카드 새벽시간 결제비율 50% 이상은 116명이었으며, 야간 결제 횟수의 규모가 크고 지속적으로 발생하는 경우에 대한 지속적인 관심이 필요할 것으로 분석 되었습니다.



<자치구별 새벽시간(00:00~06:00)대 결제 비율 분포도>

전체 공나무카드 사용자 중 새벽시간 결제비율이 50% 이상인 경우가 116명으로 나타났습니다.  
새벽 결제 횟수의 규모가 크고 지속적으로 발생하는 경우, 이용자의 상황을 주기적으로 확인하여 필요한 복지 서비스를 안내하고 지원해야 합니다.

자치구	강남구	강동구	강북구	강서구	관악구	광진구	구로구	금천구	노원구
50% 이상 90% 미만	12	4	6	6	3	1	6	3	11
90% 이상	0	2	4	0	1	0	0	0	2
합계 인원 수	12	6	10	6	4	1	6	3	13
자치구	도봉구	동대문구	동작구	마포구	서대문구	서초구	성동구	성북구	
50% 이상 90% 미만	7	4	6	4	5	0	2	4	
90% 이상	1	0	1	1	0	0	0	3	
합계 인원 수	8	4	7	5	5	0	2	7	
자치구	송파구	양천구	영등포구	용산구	은평구	종로구	중구	중랑구	
50% 이상 90% 미만	0	1	0	0	4	2	1	5	
90% 이상	0	1	0	0	0	0	0	3	
합계 인원 수	0	2	0	0	4	2	1	8	

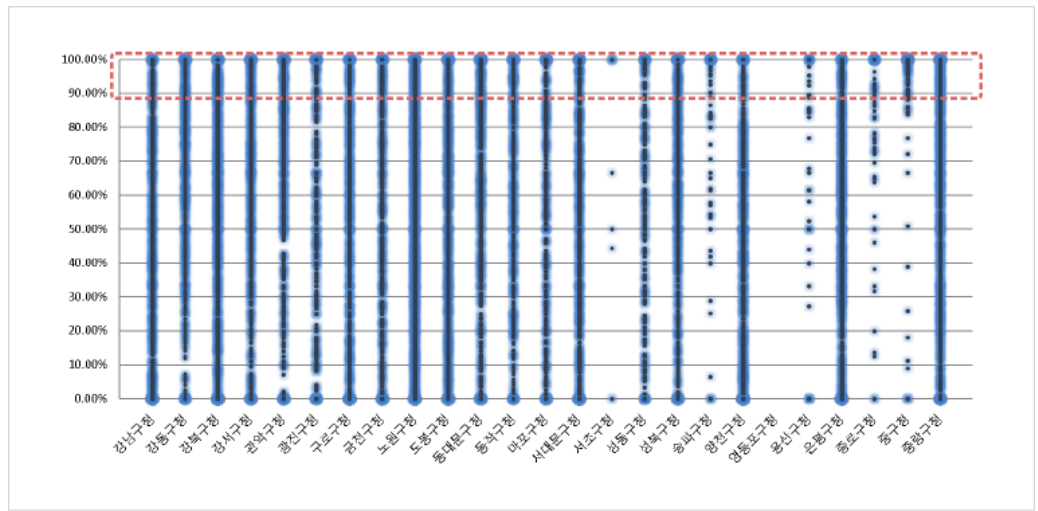
<자치구별 새벽시간대(00:00~06:00)대 결제를 50%이상 현황>

“편의점 위주 사용자”

전체 결제 횟수 대비 편의점 결제 횟수의 비율을 카드별로 계산하여 산포도를 그렸습니다. 편의점 결제 비율은 각 구청별 분포가 상 당히 고르게 나타났습니다.

각 자치구에 등록된 개별 공나무카드에서 편의점 결제비율이 90% 이상인 경우가 3,391명으로 나타나, 편의점에서도 영양을 충분히 채길 수 있는 방법에 대한 정보를 카드 이용자에게 제공할 필요가 있습니다.





〈자치구별 편의점 결제 비율 분포도〉

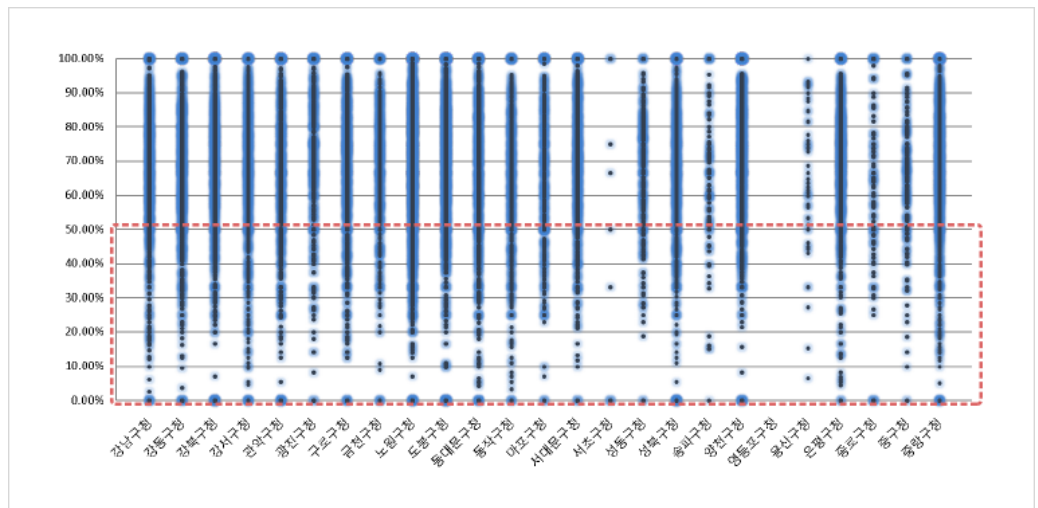
편의점 비율이 높은 이유가 카드 이용자 주변에 일반 가맹점이 부족하거나, 가맹점 위치를 모르는 경우도 있을 수 있어 이용자의 위치를 확인할 수 있는 경우, 그 원인을 추정하여 가맹점을 추가 선정하거나 가맹점에 대한 정보를 사용자에게 알려주는 방식으로 다양한 영양소 섭취를 가능케 하는 조치가 필요한 것으로 분석됩니다.

자치구명	강남구	강동구	강북구	강서구	관악구	광진구	구로구	금천구	노원구
인원수	81	205	157	209	204	106	137	119	347
자치구명	도봉구	동대문구	동작구	마포구	서대문구	서초구	성동구	성북구	
인원수	163	222	163	106	102	7	70	208	
자치구명	송파구	양천구	영등포구	용산구	은평구	종로구	중구	중랑구	
인원수	68	88	0	20	92	60	133	324	

〈자치구별 새벽시간대(00:00~06:00)대 결제율 50%이상 현황〉

#### “식사시간 이외의 이용자”

식사시간을 아침은 07:00~11:00, 점심은 11:00~15:00, 17:00~21:00로 정의하고, 이후 앞서 정의한 식사시간 사이에 해당하는 횟수를 각각 계산한 다음 전체 결제 횟수로 나눈 비율 값을 점으로 찍어 시각화 하였습니다.



〈자치구별 식사시간대 결제 비율 분포도〉

자치구에 등록된 개별 꿈나무카드에서 식사시간 결제비율이 50% 미만인 전체 13,365명 대비 16.5%인 2,203명으로 수준으로 적지 않아 아동 및 청소년기는 식습관 형성 결정기로 성인까지 식습관에 영향을 미칠 수 있어 식습관 정립에 대한 교육과 관리 등의 대책이 필요합니다.

자치구명	강남구	강동구	강북구	강서구	관악구	광진구	구로구	금천구	노원구
20% 미만	22	16	30	28	25	8	20	10	32
20% 이상 50% 미만	78	95	93	91	87	36	90	64	255
인원 수	100	111	123	119	112	44	110	74	287

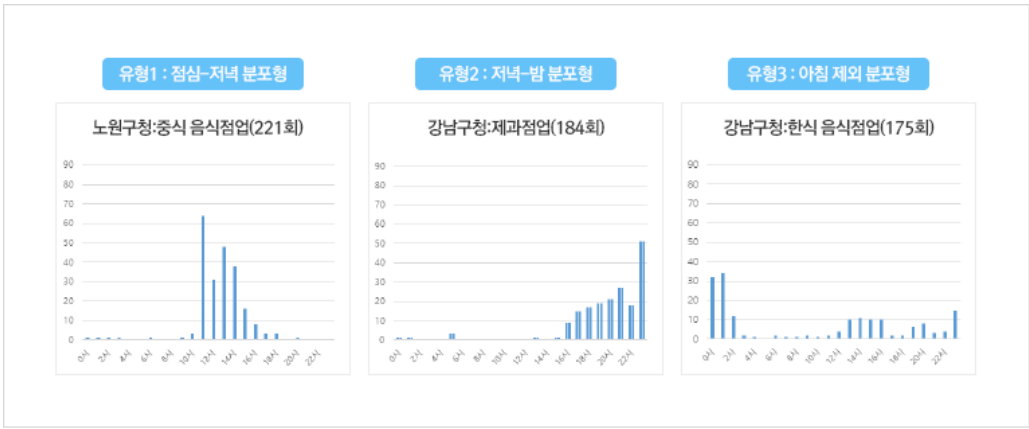
자치구명	도봉구	동대문구	동작구	마포구	서대문구	서초구	성동구	성북구
20% 미만	37	25	18	8	9	5	4	36
20% 이상 50% 미만	112	95	85	62	69	1	35	103
인원 수	149	120	103	70	78	6	39	139

자치구명	송파구	양천구	영등포구	용산구	은평구	종로구	중구	중랑구
20% 미만	5	17	0	2	17	7	5	29
20% 이상 50% 미만	12	88	0	7	69	22	26	113
인원 수	17	105	0	9	86	29	31	142

〈자치구별 편의점 결제율 90% 이상 인원수 현황〉

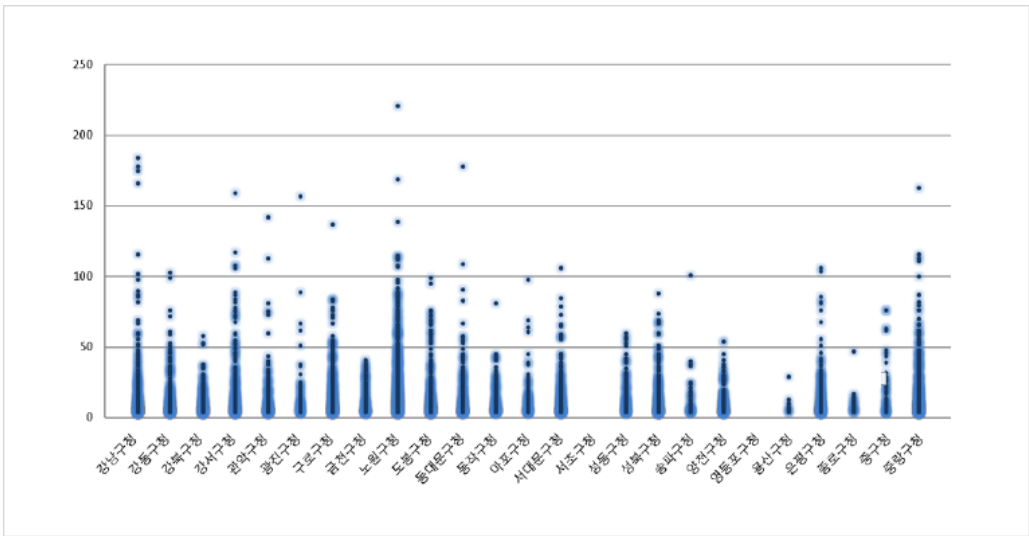
“연속결제 이용자”

동일 가맹점에서 동일한 금액의 연속 결제 패턴을 추출하고, 그 중 상위 10개를 결제 시간별로 시각화 한 결과, 식사시간이 저녁 - 새벽에 분포하는 등의 이상 패턴도 있어 이에 대해 신규 복지서비스이나 집중관리 대상으로서 관심이 필요합니다.



〈자치구별 식사시간대 결제 비율 분포도〉

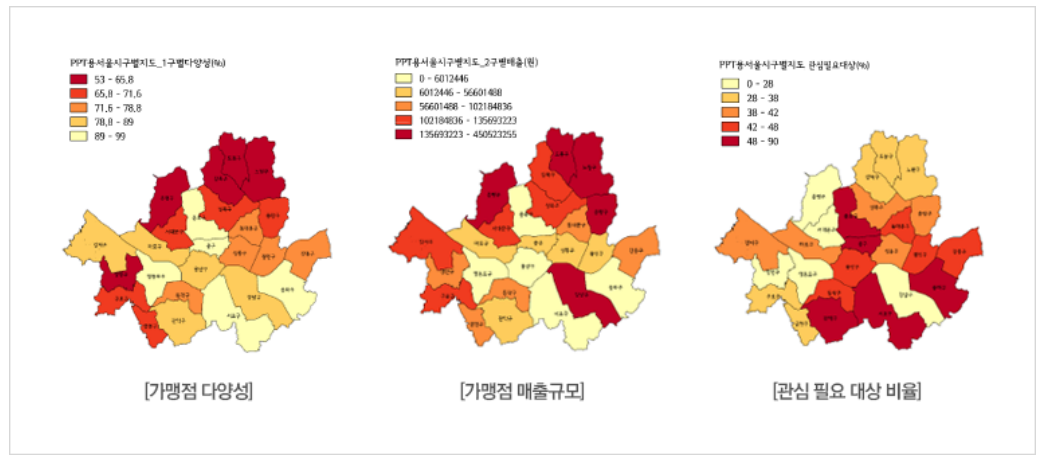
연속 결제는 지속적으로 동일 유형의 메뉴로 식사를 할 가능성이 높고, 이는 상대적으로 영양성분의 불균형 섭취 가능성이 높 아 카드 이용자의 개별 상황 확인과 식습관 정립에 대한 교육부터 생활에 대한 상담과 보호 관리 등의 복지서비스도 검토할 필요가 있는 것으 로 나타났습니다.



〈자치구별 연속결제 횟수 분포도〉

Chapter.5 “분석결과와 시사점”

각각의 온도 지도는 유사한 패턴을 보이 나 자치구별 현황에 따른 세부 패턴은 다르게 나타납니다. 따라서 지역 실정에 맞는 꿈나무 카드 제도의 개선이 필요합니다.



〈자치구별 관심 필요 대상자 비율〉

### “사용자 이상패턴 감지 시스템”

분석결과, 가맹점과 사용자 그리고 정책을 집행하는 동 주민센터와 자치구에게 실질적인 활용토록 하기 위해 일회성 분석이 아닌 주 기적으로 모니터링 할 수 있는 피드백 시스템이 필요합니다.



### “사용자 참여형 어플리케이션 개발”

결식아동급식, 꿈나무카드와 관련된 뉴스 데이터와 네이버 카페후기에서 수집한 데이터의 토픽모델링을 통해 키워드 분석결과, 가맹점 안내부족, 영양 불균형 등 지속적 이슈 제기로 이에 대한 대책이 필요합니다.

꿈나무카드 관련 다양한 정보 제공으로 사용자의 편의를 확보하기 위해 카드 사용자 위치 정보, 소비패턴 데이터 수집, 영양 정보와 가맹점 추천 등의 기능을 수행하는 어플리케이션 통합정보망이 개발되어야 합니다.




빅데이터캠퍼스 : 02-2124-2950 | 관리부서 : 정보기획관 통계데이터담당관  
주소 : 마포구 매봉산로 31 에스플렉스센터 스마티움동 15층  
Copyright © 2016 Seoul Metropolitan Government, all rights reserved.

권원사이트

서울열린데이터광장

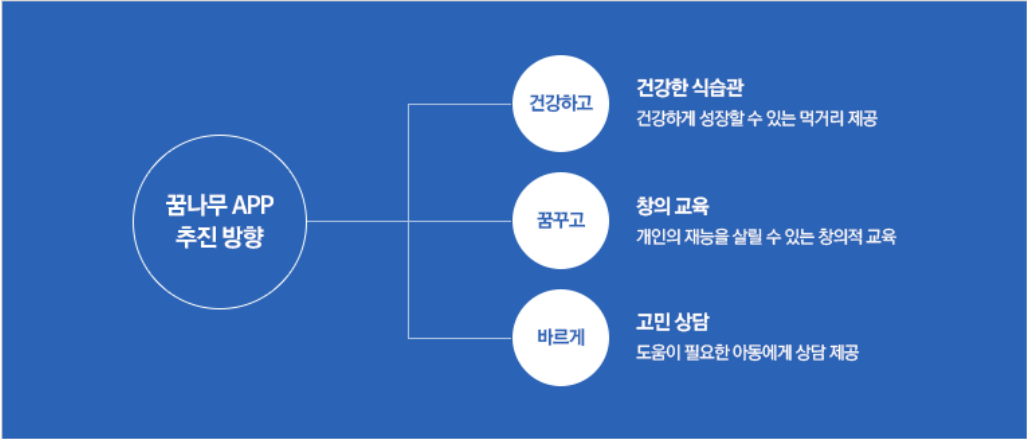
이동

## Chapter.6 “향후 발전 방향”



### 꿈나무 어플리케이션은,

꿈나무카드 사용자들이 바르고 건강하게 성장할 수 있도록 차별화된 교육과 심리상담 프로그램을 제공 하고자 합니다.



목록