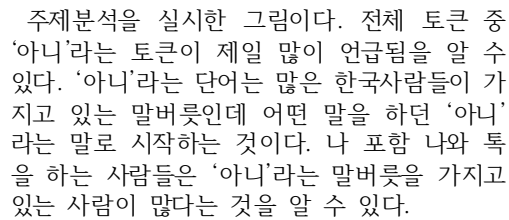
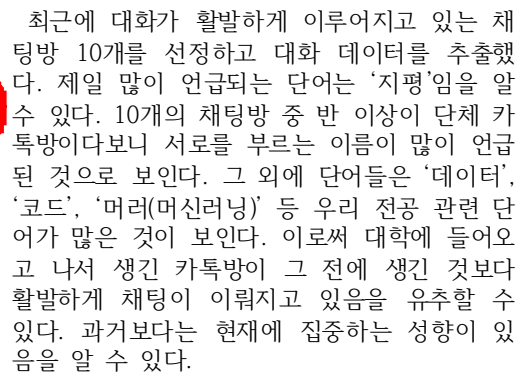
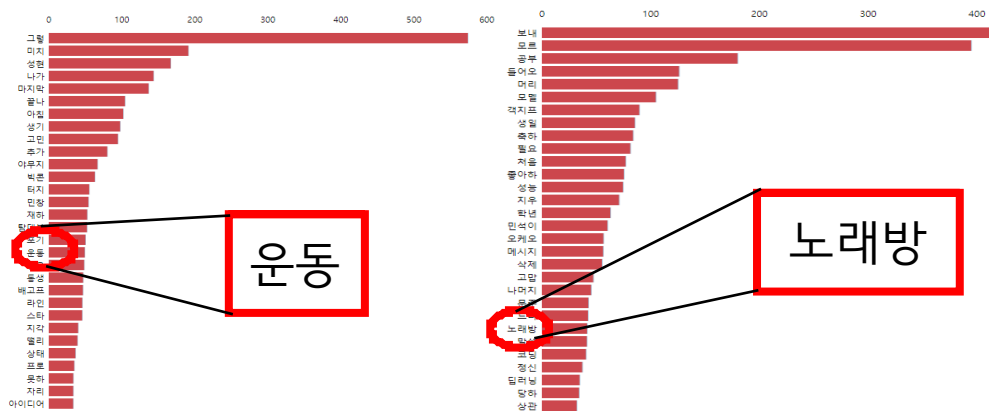


20172848 빅데이터 경영통계 전공  
이지평

1) 카카오톡





10개의 주제로 분석을 진행했는데 대부분 수업 혹은 공모전 관련 용어였다. 그러나 이 중에서 '운동'과 '노래방'이라는 단어가 등장했다. 이 두 단어는 전공 단어만큼 많이 언급되었다는 것을 보여준다.

군대에서 헬스를 시작한 나는 제대한 지 1년이 넘는 지금까지도 동네 헬스장에 꾸준히 다니면서 운동을 하고 있으며, 유튜브에서 관련 영상도 많이 찾아본다. 하루라도 운동을 빼먹으면 불안증이 생길 정도로 열심히 하고 있고 관심도 많다. 그래서 친구들과 운동 관련 이야기를 나누게 되면서 운동이라는 단어가 등장하게 된 것으로 보인다.

스트레스를 받거나 정말 심심할 때 나는 코인노래방을 간다. 그리고 친구들을 만난다면 노래방은 필수 코스이다. 잘하지는 않지만, 유튜브에서 영상을 보며 발성 연습이나 기초를 연습할 정도로 노래에 대해 관심이 많은 편이다. 노래방이라는 단어 또한 이러한 이유로 등장한 것으로 예상된다.

## 2) 카카오뱅크

### \* 거래 term

소비패턴이나 성향을 보고도 그 사람에 대해 알 수 있다. 카카오뱅크에서 최근 1년간의 카드 결제 기록을 추출해서 분석했다.

```
card.loc[card['term'] == card['term'].min()]
```

count	mean	std	min	25%	50%	75%	max
396	0 days 21:59:46.792929292	1 days 00:55:58.076460314	0 days 00:00:00	0 days 03:56:31	0 days 16:32:02.500000	1 days 04:31:23	6 days 14:55:18

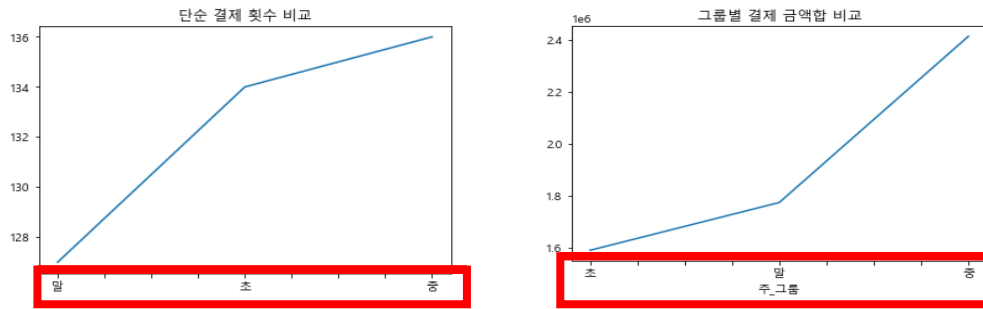
Name: term, dtype: object

```
card.loc[card['term'] == card['term'].max()]
```

거래일시	취소여부	구분	금액	가맹점명	사업자번호	과세유형	term
237 2021-05-28 10:03:45	정상	일시불	6000	큰집닭강정가능점	450-36-00219	일반과세자	6 days 14:55:18

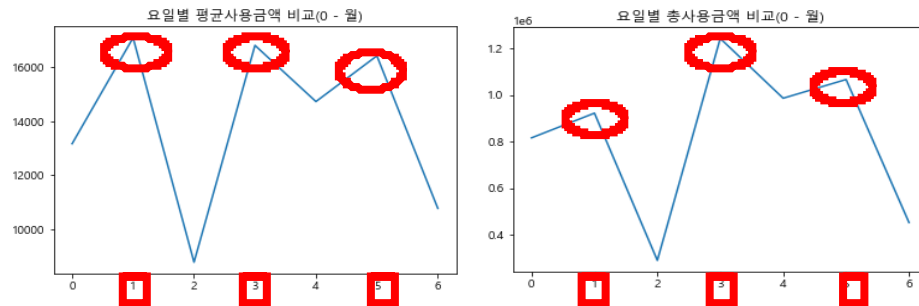
다음은 거래 사이의 term의 관한 정보이다. 최근 1년간 카드 거래 term은 총 396건이고, 평균적으로 약 22시간에 한 번 결제를 했다는 점을 알 수 있다. 가장 큰 값은 약 6일 15시간이고, 가장 작은 값은 0이라는 것이 보이는데, term 값이 0인 것은 어떤 것인지 궁금해서 행들을 뽑아보니 매달 자동 결제되는 교통비 내역임을 알 수 있었다. 반면, 최댓값인 행을 뽑아보니 21년 5월 28일 '큰집닭강정가능점'에서 결제한 내역이었음을 알 수 있다. 또한, term의 3분위수 값(75%)이 1일 4시간대인 것으로 보아, 대부분의 결제는 1일 이내의 term을 가지고 있음을 생각해볼 수 있다. 표준편차 또한 1일대 인 것을 보면, 카드결제를 굉장히 자주하는 편이라는 것을 유추할 수 있다.

\* 주\_그룹별 결제 현황



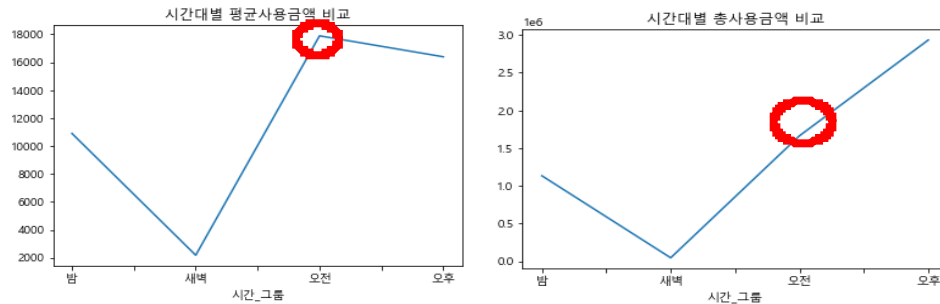
다음은 한달을 초(1~10일) / 중(11~20일) / 말(21~31일) 세 그룹으로 나눠서 결제 내역을 나타낸 것이다. 첫 번째 그래프는 단순 결제횟수를 비교한 것이다. 말<초<중 순으로 증가하는 것을 볼 수 있다. 두 번째 그래프는 그룹별 결제 금액을 비교한 것이다. 이는 초<말<중 순인 것을 확인할 수 있다. 결제횟수는 '초'그룹이 '말'그룹보다 훨씬 많지만, 결제금액합은 '말'그룹이 근소하게 앞서는 것으로 보아, '말'그룹 시기에는 '초'그룹보다 상대적으로 고가의 결제를 한다는 것을 유추할 수 있다.

\* 요일별 결제 내역



요일별 결제 내역이다. 평균 사용금액과 총사용금액 모두 화, 목, 토요일이 높다는 것을 알 수 있다. 이는 21년 4월까지 월, 수, 금에 알바를 했던 패턴이 반영된 것으로 보인다. 월수금이 알바이므로, 약속을 화, 목요일에 잡을 수 밖에 없었기 때문에 화, 목의 사용금액이 높게 나오는 것을 볼 수 있다. 총 사용금액에서의 차이를 보면 화요일이 많이 적어지는 것을 볼 수 있다. 이는 개강을 하게 되면서 약속을 주로 목, 토요일에 잡기 때문인데, 이 패턴이 반영된 것으로 볼 수 있다.

\* 시간대별 결제 내역



시간대별 결제 내역이다. 시간대는 0시 기준으로 6시간씩 잘라서 새벽(0~6시) / 오전(6~12) / 오후(12~18) / 밤(18~24) 네 그룹으로 나눴다. 평균사용금액을 보면 오전>오후>밤>새벽 순이지만, 총사용금액을 본다면, 오후>오전>밤>새벽 순임을 알 수 있다. 평균사용금액은 오전이

오후보다 더 높지만, 총 사용금액은 오후가 더 높은 것을 미루어 보았을 때, 오후의 결제량이 오전보다 훨씬 많다는 것을 유추할 수 있다.

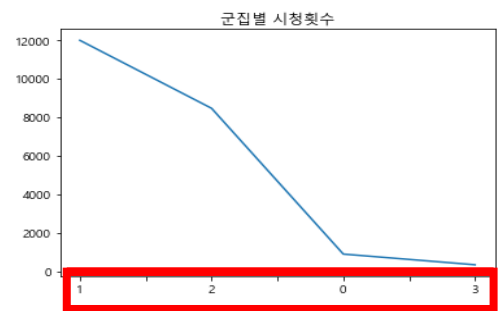
### 3) 유튜브

하루 중 가장 많은 시간 사용하는 플랫폼이다. 16년도부터 시청내역을 추출했지만, 아쉽게도 군대에 있던 시절(18년도 말 ~ 20년도 중반)의 데이터는 다른 계정을 써서 그런지 구할 수 없었다. 시청내역 위주로 분석을 진행했다.

시청내역의 title을 한글제목만 뽑아서 임베딩하고 군집분석을 실시했다. 총 4개의 군집으로 분류했다.

# cluster_title == 0 : 개그, 게임 ytb_hist.query('cluster_title == 0')['title'].sample(10)		# cluster_title == 1 : 스포츠, 운동, 예능 ytb_hist.query('cluster_title == 1')['title'].sample(10)	
8481 공포물카 혼자 있는 방 침대 아래에 괴물이 있다면 진실 기록기~~~~~	21835 최영준의 교과서급 디펜스 그해 그 경기 커트라이더 리그 시즌	13649 진경한 간지 투우사 최영준 커트라이더 리그 시즌 개막전 하이라이트	
18389 난전 대장 커트라이더 하는 데스매치 경전도만 봤습니다 ㅋㅋ	9418 마선호 선수의 대보형에 나가지 않는 이유   내추럴 보디빌드 마선호 인터뷰	6621 송중기 한껏 꾸민 전여빈의 연수원 동기모임 행차에 참가불만ㅋㅋ 변산조 방송	
13783 신 코드네임 리스트를 보았습니다~~~~~ 커트라이더 러쉬롤러스	2461 역 장전 발드럼 우승 그시절 스포인 무적함대 완성 비아 부활시작 파파	2293 기안 최초의 출판 박나래와의 불화를 예제 외면하는 이유 최혁만이 재년 개역 비밀	
5257 김민준 뽐있다고 뭐라하는 최혜나 ㅋㅋ	4618 결핵시를 쓴다면 무조건 무료 인보면 손해인 역종기능 가지 설정방법 결핵시 노트 세팅하기	15488 오마이걸 효정 현아도 본 효정의 버블링	
1641 이만다 사이프라이드 백치연기모음~~~~~	1938 직접 겪은 김준호 선수의 황제헬스 실제로 존난 헬	389 봄 빼빼리는 장영한 웃겨서 숨 넘어가는 대연 놀라운토요일	
18153 본인미 왜 살찌는지 모르겠다는 친구 왜 살찌는지 오목조목 알려주기~~~~~			
2275 온한 대학교 기숙사의 캠퍼스 ~~~~~ 유머			
3981 명나니 시즌동생 성광국 하는 아이유인나~~~~~			
13884 프랑스 선수들 김일 화나게 만든 황의조 ~~~~			
11842 임술 리얼 뽐숭이 조유리 ~~~~			
Name: title, dtype: object	Name: title, dtype: object		
# cluster_title == 2 : 아이즈원 콘텐츠, 음악콘텐츠 ytb_hist.query('cluster_title == 2')['title'].sample(10)		# cluster_title == 3 : 영어재능, 외제어 ytb_hist.query('cluster_title == 3')['title'].sample(10)	
12647 아케이드 도망가자 선우정아	5582		
11129 게스트할한한책	12972		
8827 빈센 플   바래 마이크스웨어무스	3813	망	
19436 머쉬메틀 고등래퍼 지원영상	18059		
18565 놀라운토요일 비글이 폭발 걸스데이 웃을지외 모을 의 팬클럽	16896		
7157 아이즈원 브이라이브 하이라이트 조유리편	4327		
1970 바와이 마에스트로 쇼츠 데자부	731	개월	
16538 졸업에 드디어 오셨습니다	1854		
18306 아이즈원 잔망 참아치 차원 밀구리	7062		
16619 아이즈원 최예나 허세부리는 오리 사냥	4675	망	
Name: title, dtype: object	Name: title, dtype: object		

0번 군집은 주로 개그, 게임 콘텐츠의 제목이 많았고, 1번 군집은 스포츠, 헬스, 예능 관련 콘텐츠의 제목이 많았다. 2번 군집은 아이즈원, 음악 관련 콘텐츠가 다수를 차지했고, 마지막 3번 군집은 영어제목, 혹은 맞춤법이 맞지 않게 작성된 콘텐츠들이다.



각 군집별 시청횟수를 비교해본 결과, 1>2>0>3 순으로 시청횟수가 많은 것으로 나타났다. 1, 2번 군집의 영상이 대부분을 차지하는 것으로 보아, 내가 주로 보는 영상의 콘텐츠는 '스포츠(헬스), 예능, 아이즈원, 음악'을 포함하는 콘텐츠임을 알 수 있다.

나에게 특별한 장소란 무엇일까 계속 고민을 해왔다. 분석을 진행할수록, 일상 속에서 내 삶을 특별하게 해주는 요소들을 발견했다. 바로 '운동'과 '노래'이다. 남들처럼 거창하진 않지만, 나에게 특별한 장소는 '헬스장'과 '(코인)노래방'이라고 생각한다. 근거는 위에서 자주 사용하는 앱들을 분석해보며 나온 결과들이다.

앞서 '나는 누구인가'에 대해 분석을 진행해봤다. 자주 사용하는 어플 3개로 나를 분석해보자면, 나는 과거보다는 현재(대학교)의 특방이 더 활발하며, '아니'라는 말버릇을 사용하고 '운동'과 '노래방'을 좋아하는 사람이다. 또 평균 22시간 간격으로 카드 결제를 하고, 월말에 월초보다 과소비를 하는 경향이 있는 사람이며, 화·목·토요일 / 오후에 주로 카드 결제를 하는 사람이다. '스포츠(헬스), 예능, 아이즈원, 음악'을 포함하는 콘텐츠를 주로 시청하는 사람이기도 하다. '내가 몰랐던 나'에 대해 알게 되어서 굉장히 재미있었고, 나를 돌아보면서 스스로에 대한 생각을 정리해볼 수 있는 시간이 많이 주어져서 의미있는 프로젝트였다.

이제는 '무엇을 할 것인가'에 대해 고민해봐야 한다. 사실 나는 아직 무엇을 할 것이라는 확실한 비전이 없다. 막연하게 현재 하고있는 전공 쪽으로 취업을 해야겠다는 생각만 가지고 있을 뿐이다. 그래서 확실하게 하고 싶은 것이 생기면 기초적인 부분에서 부족함이 없도록 공부하고 있다.

순번	구분	이수구분	이수요구 학점	이수확정			이수해야 할 학점
				실 이수	등록인정	계	
1	제1전공심화	빅데이터경영통계전공	빅데이터경영통계전공	0	0	0	0
2	제2전공		2전공(계열교양)	0	0	0	0
3	제2전공		2전공(전공)	0	0	0	0
4	제3전공		3전공(계열교양)	0	0	0	0
5	제3전공		3전공(전공)	0	0	0	0
6	연계융합전공	데이터사이언스융합전공	연계융합(전공)	36	6	15	21

	월	수	금
9	선형대수 미래관6층11호실	선형대수 미래관6층11호실	
10			
11		고급통계실무 경영관1층3호실	고급통계실무 경영관1층3호실
12	딥러닝 경영관1층3호실		빅데이터프로젝트:캡스톤 디자인II 경영관1층3호실
1			
2			
3	경영정보학원 본 경영관5층6호실	공간빅데이터 분석 경영관1층3호실	마케팅에널리틱스 국제관1층2호실
4			
5			
6	D&A 컨퍼런스 세션		
7			

현재는 빅데이터 경영통계 전공과 소프트웨어 융합 대학의 과목을 같이 이수하는 '데이터사이언스 융합 전공'을 하고 있다. 데이터 사이언티스트로 진출하려면 단순히 분석뿐만 아니라 컴퓨터를 다루는 역량 또한 필요하다고 들었기 때문이다. 그래서 이번 학기에는 '선형대수'라는 과목을 듣고 있다. 교수님께서 선형대수를 컴퓨터 자료구조에 대한 관점으로 가르쳐주셔서 굉장히 흥미있게 공부 중이다.

전공을 더 깊게 파고드는 관점에서는 현재 전공에 유일하게 있는 학회 'D&A'에서 컨퍼런스를 준비하고 있다. '자연어 처리'를 포커스로 잡고 주제를 '사투리 번역기'로 하며 진행 중에 있다.

외부 활동으로는 여러 공모전에 참여해왔다. 데이터나 캐글처럼 모델링에 집중하는 대회나, 빅콘같은 처음부터 끝까지 하나의 프로젝트를 완성하는 대회까지, 골고루 참여하고 있다. 아직까지 수상은 하지 못했지만, 대회를 거듭할수록 배울 것이 많아지고 잘하는 사람도 많이 보다보니 시야도 넓어진다. 나는

아직 멀었다는 생각이 들면서도 언젠가 저들처럼 되겠다는 의지가 생긴다.

그동안 주어진 데이터에서만 분석을 해왔는데, 이번에는 스스로 데이터를 구하고 가공해서 분석해볼 수 있는 아주 좋은 기회였다. '나'라는 사람에 대해 오롯이 생각해볼 수 있는 시간을 가져보며 나에 대해 조금 더 깊은 이해를 하기도 했다. '나'를 분석해보라는 그동안 접해보지 못했던 미션을 주신 교수님께 감사드리며, 앞으로도 꾸준히 '나'를 분석하며 발전해나가는 사람이 될 것이다.