프 로 젝 트 1

프로젝트 명	포스트 코로나시디	배 OTT이용률에 영향을 주는 요인 분석	팀 명	다타				
프로젝트 목적		많은 타격을 받았다고 이야기하는 영화계 여 부상한 서비스 중 대표적인 서비스인	개발 기간	4주				
및 개발 동기	OTT서비스를 비교 위한 방법을 생각	l하여 영향을 준 원인과 앞으로의 발전을	개발 인원	5				
	파이썬(3.9.12)	matplotlib, seaborn, Beautifulsoup, okt, pandas, numpy,						
활용 기술	R(4.2.1)	dplyr, pwr,	사용 언어	주 개발언어 : 파이썬(3.9.12)				
			~ 개발 환경	보조 언어 : R(4.2.1)				
개 요	OTT 이용률 / 영화관 이용률 변화 확인 OTT서비스 주 이용층 분석 코로나와 영화관관객수 상관관계 분석 영화관 이용률과 OTT서비스 상관관계 분석 가격에 따른 영화관 이용률 확인							
기대 효과	OTT서비스에서 앞으로 어떤 고객층을 타겟으로 하여 마케팅을 해야하는지를 알 수 있고 가격인상 등에 있어 서 어떻게 접근해야 할 지 예상해 볼 수 있다.							

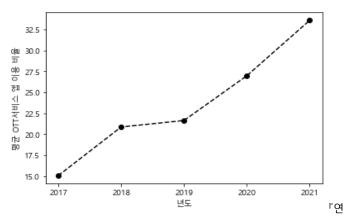
본인 역할	- 데이터 EDA(나이에 따른 OTT서비스, 영화관 이용률 변화 분석 및 시각화, 가격에 따른 영화관 이용률 분석 및 시각화, 코로나 확진자수와 관객수 시각화) - 파이썬과 R을 이용한 상관관계분석 - 코로나에 관한 뉴스 크롤링을 통하여 wordcloud로 시각화
보 완 점	의미가 있는 결과를 낼 수 있는 통계처리를 위해서 더 많은 양의 데이터가 필요

일 정 단 계	1주차			2주차					3주차					4주차									
사전 기획																							
데이서 수집																							
데이터 전처리																							
모델링																							
서비스 구축																							
테스트																							
문서 작업																							
기타																							

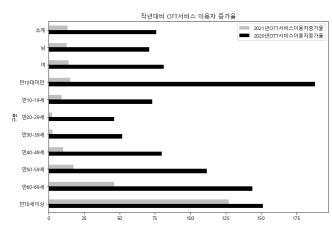
프로젝트1

프로젝트 내용

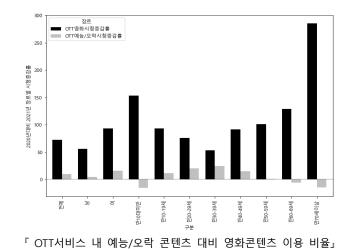
『 OTT서비스가 코로나 기간 동안 증가한 것을 시각적으로 표현하여 확인하고 그에 반해 영화관 관객수는 줄어듦을 시각적으로 확인한다. 추가적으로 OTT서비스에서 영화 콘텐츠를 이용 증가비율은 예능/오락 콘텐츠 이용 증가 비율에 비해 많이 증가함을 시각적으로 표현하여 OTT 서비스를 통해 영화 콘텐츠를 많이 이용했음을 확인했다. 』

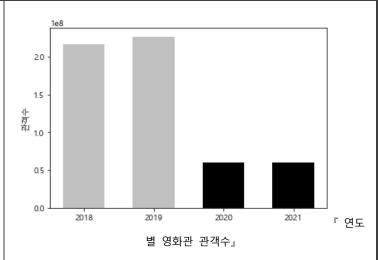


도별 OTT 서비스 관련 어플리케이션 이용 비율 증가』



『 OTT서비스 이용자 설문_OTT 서비스 이용증가 비율』

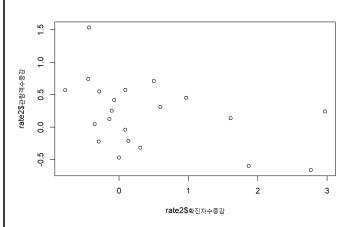




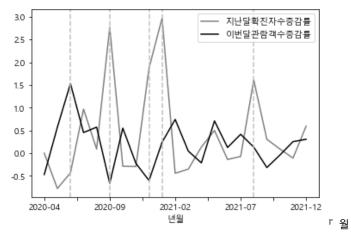
파이썬에서 기본 plot과 matplotlib의 pyplot모듈을 이용하여 시각화

프 로 젝 트 1

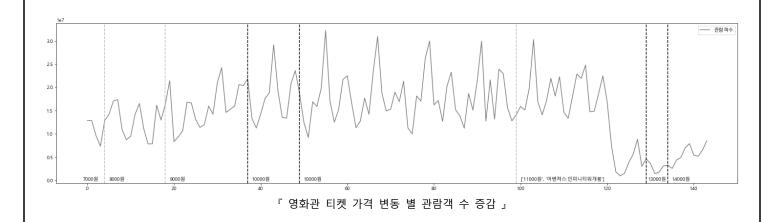
『 코로나 확진자 수가 영화관 관람객 수에 영향을 미치는지 상관관계와 시각적 그래프를 이용하여 확인하고 영화관 티켓가격 또한 영화관 관람객 수 증감에 영향을 미쳤는지 시각적으로 확인하였다. 특히 코로나 확진자 수의 경우 코로나 확진자 수가 다음 날의 영화관 관객수에 영향을 미칠 것이라는 아이디어를 통해 의미 있는 상관관계를 도출해 내었다.』



『 월별 코로나 확진자 수 증감 비율과 관람객수 증감 비율 』



별 코로나 확진자 수와 다음 달 영화관이용객 수』



R의 dplyr을 이용하여 데이터 일부 selection, 통계를 위한 상관관계 분석, 파워검정 및 시각화와 파이썬 기본 plot과 matplotlib을 이용한 시각화

프 로 젝 트 1

『 코로나 확진자 수에 대하여 영화관 관람객 수의 상관관계분석을 통해 2020년은 예상과 같이 코로나 확진자 수에 대한 관람객 수 변화가일어남과 달리 2021년은 관계가 없다고 나타났다. 따라서 2020년과 2021년에 차이가 있다고 보고 그 차이가 코로나에 대한 인식 변화라고생각하여 뉴스 크롤링을 통하여 인식 변화를 확인하고 wordcloud로 표현하였다.』

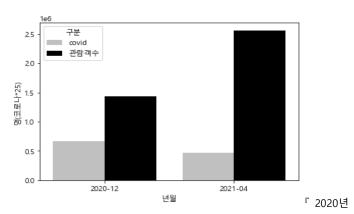
코로나.관람객.다음날20

상관계수	-0.223516
p-value	0.000038

코로나.관람객.다음날21

상관계수	2.946166e-01
p-value	2.569275e-08

『 2020년과 2021년의 코로나 바이러스 확진자 수와 다음 날 영화관 관람객 수의 상관관계 분석 』



과 2021년의 비슷한 코로나 확진자 수에 대한 영화관 관람객 수의 차이 」



『 2020년 12월 코로나에 관한 뉴스 헤드라인 keyword 』



『 2021년 4월 코로나에 관한 뉴스 헤드라인 keyword 』

파이썬에서 피어슨 상관관계분석, 파이썬의 seaborn을 이용한 시각화, BeautifulSoup을 이용한 뉴스 크롤링 및 okt를 이용한 wordcloud생성