

1. DATA 전처리

OTT Service

1. 데이터 전처리

데이터 출처

- 미디어 통계 포털
 - 통계청
- 영화관입장권 통합 전산망
 - 한국 데이터 거래소

- 가장 많이 이용하는 OTT 서비스
- 가장 많이 이용하는 OTT 서비스의 주 이용 콘텐츠
- 자주 사용하는 스마트 기기 애플리케이션
- 최근 3개월 OTT 서비스 이용 경험 여부
- 1년 동안 가장 많이 이용한 여가 공간
- OTT 서비스 유료 결제 여부
- OTT T 서비스 월 평균 지출 금액
- OTT 서비스 이용빈도
- 총 영화관 관객수 및 매출액
- 전국 코로나 확진자

1. 데이터 전처리

데이터 전처리 방향

- **코로나 기간(2020-2021) 2년간의 데이터 사용**

2년 동안 코로나 방역 정책 및 사람들의 인식 변화에 따른 변동 사항을 확인하고자함

- **하나의 FINAL 데이터셋을 만드는 것이 아닌 분석 목적에 맞는 새로운 데이터셋 생성**

OTT 서비스의 전반적인 이용 추이와 유저의 이용 패턴 분석

코로나 기간 중 OTT 서비스 이용에 영향을 주는 요인 분석

1. 데이터 전처리

데이터 소개

예시)

- ▶ 가장 많이 이용하는 ott service
- ▶ 가장 많이 이용하는 ott service의 주 이용 콘텐츠
- ▶ 자주 사용하는 스마트 기기 애플리케이션
- ▶ 최근 3개월 OTT 서비스 이용 경험 여부

구분
별(2)

2020

... 2021

	구분 별(2)	사례수 (명)	티빙 (TvING) (%)	웨이브 (Wavve) (구 폭 pooq), 옥 수수) (%)	KT 시즌 (Seezn) (구 olleh tv 모바일) (%)	U+ 모 바일(구 U+비디 오포털) (%)	에브 리온 TV (%)	네이 버시 리즈 (구 네 이버 N스 토어) (%)	공 TV (%)	넷플 릭스 (%)	카 카 오 TV (%)	...	네이 버시 리즈 (구 네 이버 N스 토어) (%)	공 TV (%)	넷플 릭스 (%)	카 카 오 TV (%)	유튜 브 (%)	아마 존 프 라임 비디 오 (%)	왓차 플레 이 (%)	쿠팡 플레 이 (%)	아프 리카 TV (%)	기 타 (%)
전 제	소계	7434.0	4.4	4.0	1.0	1.3	0.0	0.7	0.2	7.1	1.6	...	0.3	0.2	8.8	1.8	81.5	0.0	0.2	0.2	0.3	0.0
성 별	남	3872.0	3.9	3.9	1.1	1.3	0.0	0.7	0.2	6.5	1.6	...	0.3	0.2	8.1	1.6	82.6	0.0	0.1	0.1	0.5	0.1
	여	3562.0	4.9	4.0	0.8	1.3	0.1	0.6	0.1	7.8	1.7	...	0.4	0.1	9.5	2.0	80.3	0.0	0.3	0.3	0.1	0.0
연 령	만10 대미 만	173.0	1.2	0.0	2.7	0.0	0.0	0.0	0.0	0.9	0.0	...	0.0	0.0	6.2	0.0	90.2	0.0	0.0	0.0	0.0	0.4
	만10- 19세	825.0	3.8	1.0	0.5	0.5	0.0	0.4	0.1	4.6	1.3	...	0.1	0.2	6.8	3.5	86.2	0.0	0.4	0.0	0.1	0.1
	만20- 29세	1378.0	6.3	3.8	0.7	0.9	0.0	0.6	0.1	13.1	1.9	...	0.5	0.2	17.1	1.4	70.3	0.0	0.6	0.1	1.5	0.1
	만30- 39세	1420.0	4.3	8.1	0.6	1.5	0.0	0.7	0.3	9.5	1.2	...	0.3	0.3	11.3	2.1	74.2	0.0	0.3	0.5	0.0	0.0
	만40- 49세	1481.0	5.2	4.3	1.2	2.5	0.2	0.9	0.2	7.7	1.0	...	0.5	0.1	11.1	2.0	77.4	0.0	0.1	0.4	0.1	0.0
	만50- 59세	1373.0	3.8	2.0	0.9	0.8	0.0	1.1	0.1	3.8	2.1	...	0.3	0.3	4.7	1.6	87.8	0.0	0.0	0.0	0.1	0.0
	만60- 69세	666.0	2.2	3.8	1.9	1.1	0.0	0.4	0.0	1.2	2.7	...	0.3	0.1	1.3	0.3	94.2	0.0	0.0	0.0	0.0	0.0
만70 세이 상	119.0	1.1	4.1	2.5	1.8	0.0	0.0	0.0	0.0	3.6	...	0.0	0.0	0.4	2.3	94.0	0.0	0.0	0.0	0.0	0.0	

구분
별(2)

2020

2021

구 분 별 (2)	사례수 (명)	TV 방송 프로그램 (VOD/다시 보기) (%)	해외 TV 방 송 프 로그램 (%)	실시 간 TV (%)	영화 (%)	OTT 사업자 자체 제작 프로그램(넷 플릭스 오리 지널 등) (%)	예능/오 락(개인 방송 콘 텐츠 포 함) (%)	음악 (%)	기 타 (%)	사례수 (명)	영화 (%)	예능/오 락(개인 방송 콘 텐츠 포 함) (%)	드라 마 (%)	교양/ 다큐 멘터 리 (%)	교육/ 학습 (%)	애니메 이션 (만화) (%)	스포 츠 (%)	뉴스 (%)	기 타 (%)	
전 체	소 계	7434.0	57.0	5.3	24.5	22.3	19.8	73.1	32.1	0.3	8293.7	38.5	80.5	47.0	19.6	9.2	10.2	23.4	25.8	2.0
성 별	남	3872.0	54.9	5.5	24.9	23.9	20.4	75.2	30.3	0.4	4283.5	37.4	78.6	33.2	20.6	8.4	9.3	39.9	29.7	1.7
	여	3562.0	59.2	4.9	24.0	20.6	19.1	70.8	34.1	0.3	4010.2	39.8	82.5	61.7	18.6	10.1	11.2	5.8	21.6	2.2
연 령	만 10 대 미 만	173.0	24.9	0.0	12.3	3.2	9.7	85.7	19.8	1.0	174.5	8.1	71.3	7.1	9.3	41.8	72.9	0.3	0.9	0.0
	만 10- 19 세	825.0	48.6	2.6	18.3	15.9	12.9	79.6	40.4	0.9	880.1	30.7	88.8	31.1	12.4	36.1	39.1	9.9	5.1	1.6

1. 데이터 전처리

데이터 소개

예시)

- ▶ 가장 많이 이용하는 ott service
- ▶ 가장 많이 이용하는 ott service의 주 이용 콘텐츠
- ▶ 자주 사용하는 스마트 기기 애플리케이션
- ▶ 최근 3개월 OTT 서비스 이용 경험 여부

구분 별 (1)	구분 별 (1)	2017										2018 ...										2020										2021																																																																																																															
		소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)	인스턴 트메신 저(카 카오톡 라인 페이 스 등) SNS 자재 메신저 등) (%)	소셜 네트 워킹 (카카 오리 페이 스위 터 등) (%)

구분	구분별 (2)	2019			2020			2021		
		사례수 (명)	이용경험 있음 (%)	이용경험 없음 (%)	사례수 (명)	이용경험 있음 (%)	이용경험 없음 (%)	사례수 (명)	이용경험 있음 (%)	이용경험 없음 (%)
전체	소계	10864.0	41.0	59.0	10302.0	72.2	27.8	10154.0	81.7	18.3
성별	남	5401.1	44.0	56.0	5146.0	75.2	24.8	5061.2	84.6	15.4
	여	5462.9	38.1	61.9	5156.0	69.1	30.9	5092.8	78.7	21.3
연령	만10대미만	346.5	23.3	76.7	258.0	67.1	32.9	226.1	77.2	22.8
	만10-19세	1036.2	48.4	51.6	984.0	83.8	16.2	962.7	91.4	8.6
	만20-29세	1564.4	65.4	34.6	1440.0	95.7	4.3	1411.4	97.9	2.1
	만30-39세	1578.3	62.0	38.0	1509.0	94.1	5.9	1441.2	96.5	3.5
	만40-49세	1838.3	47.2	52.8	1745.0	84.9	15.1	1694.0	93.6	6.4
	만50-59세	1924.9	35.5	64.5	1827.0	75.2	24.8	1788.0	88.4	11.6
	만60-69세	1383.8	19.8	80.2	1377.0	48.3	51.7	1453.5	70.6	29.4
	만70세이상	1191.6	4.1	95.9	1162.0	10.3	89.7	1177.2	23.4	76.6

1. 데이터 전처리

변수 소개

- ▶ 가장 많이 이용하는 ott service
티빙 ~ 쿠팡 플레이
- ▶ 가장 많이 이용하는 ott service의 주 이용 콘텐츠
다시보기 ~ 예능/오락
- ▶ 자주 사용하는 스마트 기기 애플리케이션
소셜 네트워킹 ~ 게임
- ▶ 최근 3개월 OTT 서비스 이용 경험 여부
있다 / 없다
- ▶ 1년 동안 가장 많이 이용한 여가 공간
2016 '영화관' ~ 2021 '영화관'
- ▶ OTT 서비스 유료 결제 여부
있다 / 없다
- ▶ OTT 서비스 월 평균 지출 금액
5천원 미만 ~ 3만5천원 미만
- ▶ OTT 서비스 이용빈도
하루에도 여러 번 ~ 분기/연간 1회

이상치 및 결측치 처리 X

일정 기준 미만인 변수는 기타 컬럼을 병합 처리
연도별 증감율에 대한 파생변수 생성

1. 데이터 전처리

변수 소개

▶ 코로나 확진자와 영화관 관람객수 (월별)

- 년월
- 관람객수
- Covid
- 확진자수 증감
- 관람객수 증감

[파생 변수] 생성

전월 대비

확진자수증감

관람객수증감

	년월	관람객수	covid	확진자수증감	관람객수증감
0	2020-03	1834722	5560.0	0.000000	0.000000
1	2020-04	972572	1241.0	-0.776799	-0.469908
2	2020-05	1526236	703.0	-0.433521	0.569278
3	2020-06	3864533	1385.0	0.970128	1.532068
4	2020-07	5618701	1509.0	0.089531	0.453915
5	2020-08	8834699	5685.0	2.767396	0.572374
6	2020-09	2988647	4051.0	-0.287423	-0.661715
7	2020-10	4633609	2864.0	-0.293014	0.550404
8	2020-11	3595310	8219.0	1.869763	-0.224080
9	2020-12	1438873	32638.0	2.971043	-0.599792
10	2021-01	1786117	18215.0	-0.441908	0.241331
11	2021-02	3111920	11827.0	-0.350700	0.742282
12	2021-03	3256510	13430.0	0.135537	0.046463
13	2021-04	2562285	20148.0	0.500223	-0.213181
14	2021-05	4379476	17356.0	-0.138575	0.709207
15	2021-06	4927990	16109.0	-0.071848	0.125246
16	2021-07	6976451	42035.0	1.609411	0.415679
17	2021-08	7910082	54897.0	0.305983	0.133826
18	2021-09	5412319	59899.0	0.091116	-0.315770
19	2021-10	5192139	53450.0	-0.107665	-0.040681
20	2021-11	6512815	85232.0	0.594612	0.254361
21	2021-12	8502983	183680.0	1.155059	0.305577

1. 데이터 전처리

변수 소개

▶ 코로나 확진자와 영화관
관람객수 (일별)

- cov.2 covid
- cov.3 covid
- mov.2 movie
- mov.3 movie

Wide형 데이터
Long형으로

결측치 처리

jan21	feb21	mar21	apr21	may21	jun21	jul21	aug21	sep21	oct21	nov21	dec21
60747	60260	162608	65824	204635	76466	118656	409792	224163	115591	81511	138340
49018	59759	47070	68983	159691	77178	164475	225973	155882	242601	83333	118661
40186	107395	77666	187487	57398	121954	306639	217963	173685	264556	337027	123887
14518	82535	79700	160755	77940	148388	286630	230747	396214	212283	223146	236162
15741	84352	87069	43160	326744	347457	93561	203358	364069	58414	277533	209509
16591	179138	222971	40889	61494	307184	109637	189014	110139	72805	592547	68607
15393	162214	214664	56611	69404	91282	256487	358920	92777	62562	515982	67631
15419	61927	58559	52712	113036	78851	224746	347158	100032	77907	143154	84529
34658	64470	55426	60001	122405	98863	249764	135384	106840	170800	125358	90725
30662	108568	58456	119307	40077	89855	527019	129620	110278	195155	180743	89414
10776	167142	71930	108861	39008	124331	462065	259735	251127	151081	148272	154135
11659	167858	79767	39056	53654	270495	125104	226155	245787	49925	179499	145378

...

198648	NA	61726	118525	286486	185280	221926	97131	122847	240786	123932	232871
185971	NA	110776	NA	96530	NA	376778	112342	NA	228328	NA	313716

	cov2.covid	mov2.movie
1	586	63441
2	0	60084
3	0	61515
4	516	65692
5	438	68790
6	518	123977
7	483	106888
8	367	51615
9	248	51426
10	131	57005
11	242	50020

...

294	1053	38944
295	1097	18550
296	926	20257
297	869	65196
298	1092	65960
299	985	141746
300	1241	87197
301	1132	79216
302	970	34456
303	808	36828
304	1046	52810
305	1050	65097

	cov3.covid	mov3.movie
1	1029	49018
2	824	40186
3	657	14518
4	1020	15741
5	715	16591
6	840	15393
7	870	15419
8	674	34658
9	641	30662
10	665	10776
11	451	11659

...

349	7850	415297
350	7622	438838
351	7435	679040
352	7314	715005
353	6236	231784
354	5318	216070
355	5202	294007
356	7456	252592
357	6919	363934
358	6233	814717
359	5842	592407
360	5419	205300
361	4207	184211
362	3865	257281
363	5409	232871
364	5037	313716



1. 데이터 전처리

변수 소개

▶ OTT 영화 이용과 영화관 이용

	OTT영화시청증감률	여가영화관증감률
구분		
10대	97.080292	-27.868852
20대	76.724138	-35.483871
30대	53.267974	-2.325581
40대	91.101695	-54.838710
50대	101.351351	-33.333333
60대	129.113924	-50.000000
70세이상	285.714286	33.333333
남성	56.485356	-30.000000
소계	72.645740	-30.000000
여성	93.203883	-33.333333

	통계분류(1)	2020	2021	여가영화증감률
	통계분류(1)	영화관	영화관	
전체	소계	3.0	2.1	-30.000000
성별	남성	3.0	2.1	-30.000000
	여성	3.0	2.0	-33.333333
연령별	15~19세	6.1	4.4	-27.868852
	20대	6.2	4.0	-35.483871
	30대	4.3	4.2	-2.325581
	40대	3.1	1.4	-54.838710
	50대	1.8	1.2	-33.333333
	60대	0.8	0.4	-50.000000
	70세이상	0.3	0.4	33.333333

		2020	2021	OTT영화시청증가율
전체	소계	22.3	38.5	72.645740
성별	남	23.9	37.4	56.485356
	여	20.6	39.8	93.203883
연령	만10대미만	3.2	8.1	153.125000
	만10-19세	15.9	30.7	93.081761
	만20-29세	34.8	61.5	76.724138
	만30-39세	30.6	46.9	53.267974
	만40-49세	23.6	45.1	91.101695
	만50-59세	14.8	29.8	101.351351
	만60-69세	7.9	18.1	129.113924
	만70세이상	3.5	13.5	285.714286

MERGE

[파생변수] 생성

- OTT영화시청증감률
- 여가영화관증감률

1. 데이터 전처리

워드 클라우드

▶ 2020년과 2021년 코로나에 대한 인식변화

상관분석 내용을 보충하기 위한 시각적 자료

BeautifulSoup
Urlopen

뉴스
21년도 4월
20년도 12월

WordCloud
Okt

불용어 처리 및
워드 클라우드 생성