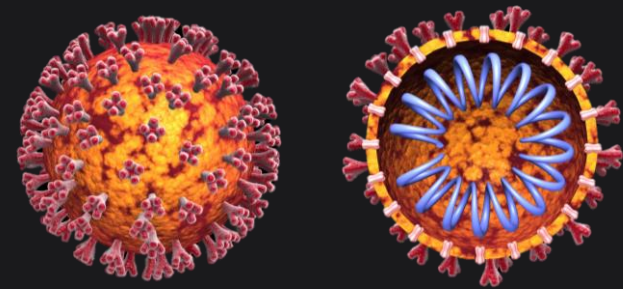


코로나19에 따른 생활 양상 변화



211001 스파게티팀

김범중 정진우 가채원 윤진훈

목차

1

코로나 확진자 변화 추이

2

코로나 확진자 증감에 따른 생활 양상 변화

2-1

코로나 확진자 증감과 공공자전거 이용객 수, 이용 시간

2-2

코로나 확진자 증감과 지하철 승객 수 변화

2-3

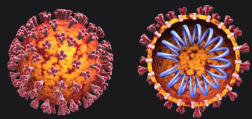
코로나 확진자 증감과 온라인 거래액

3

프로젝트 중 주요 이슈

4

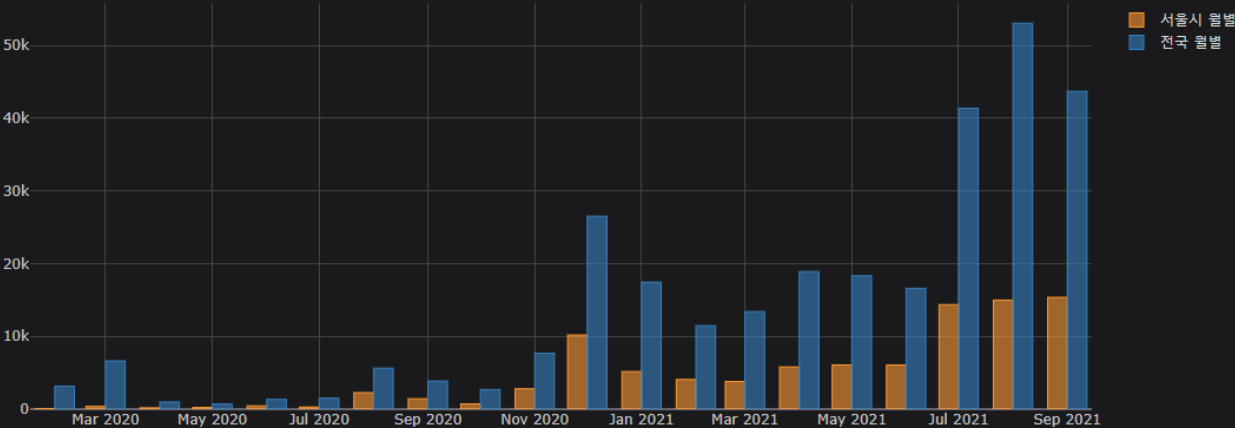
개선 사항



1. 코로나 확진자 변화 추이

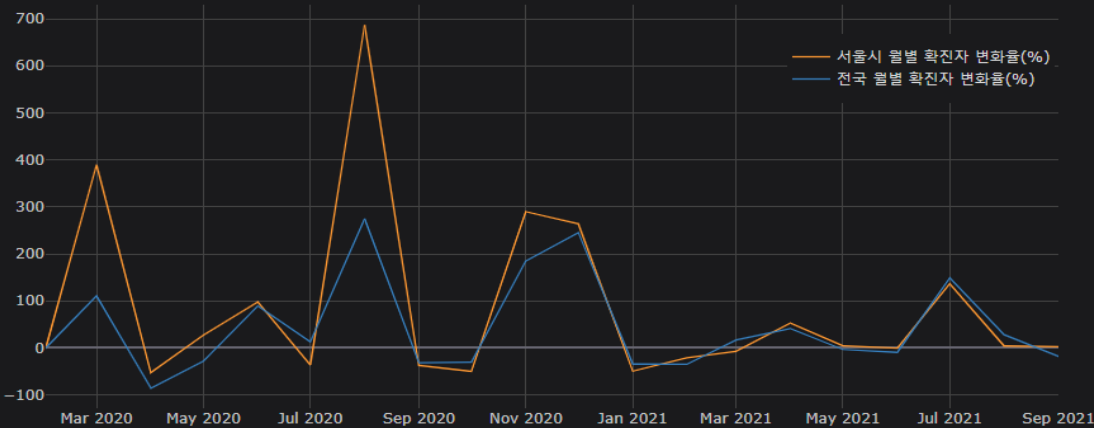
변화량

월별 코로나 확진자 수

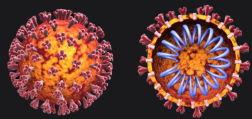


변화율

월별 코로나 확진자 수 변화율(%)

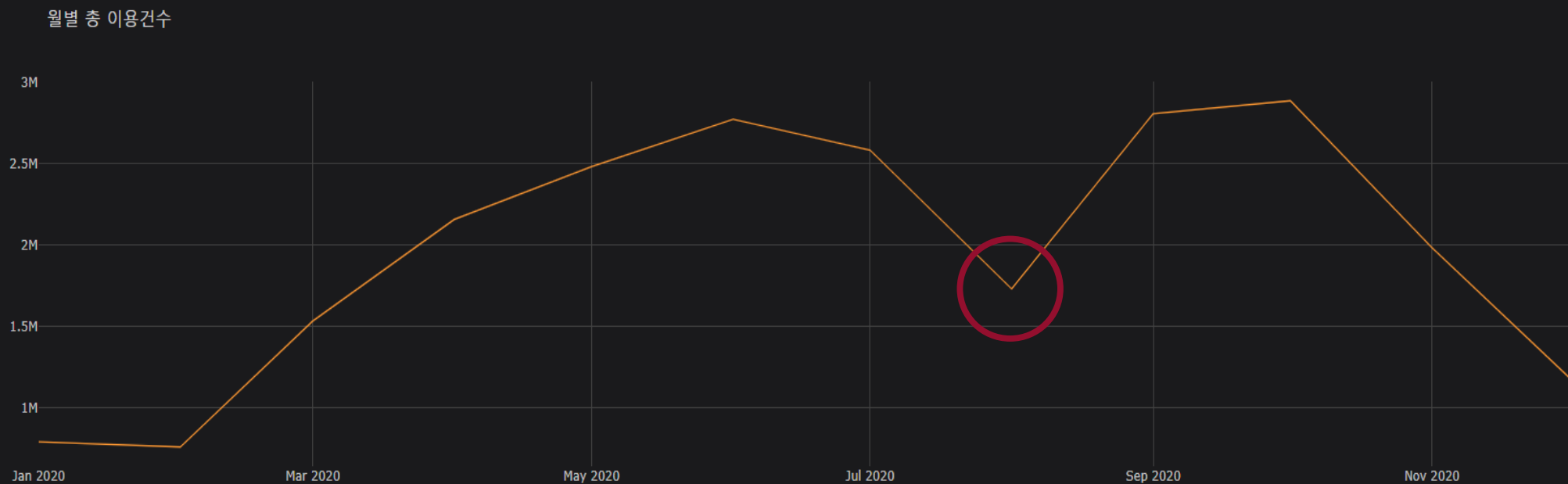


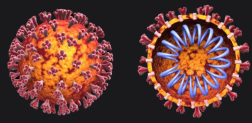
2-1) 코로나 확진자에 따른 공공자전거 이용객 수, 이용 시간



2-1 코로나 확진자 증감과 공공자전거 이용객 수, 이용 시간

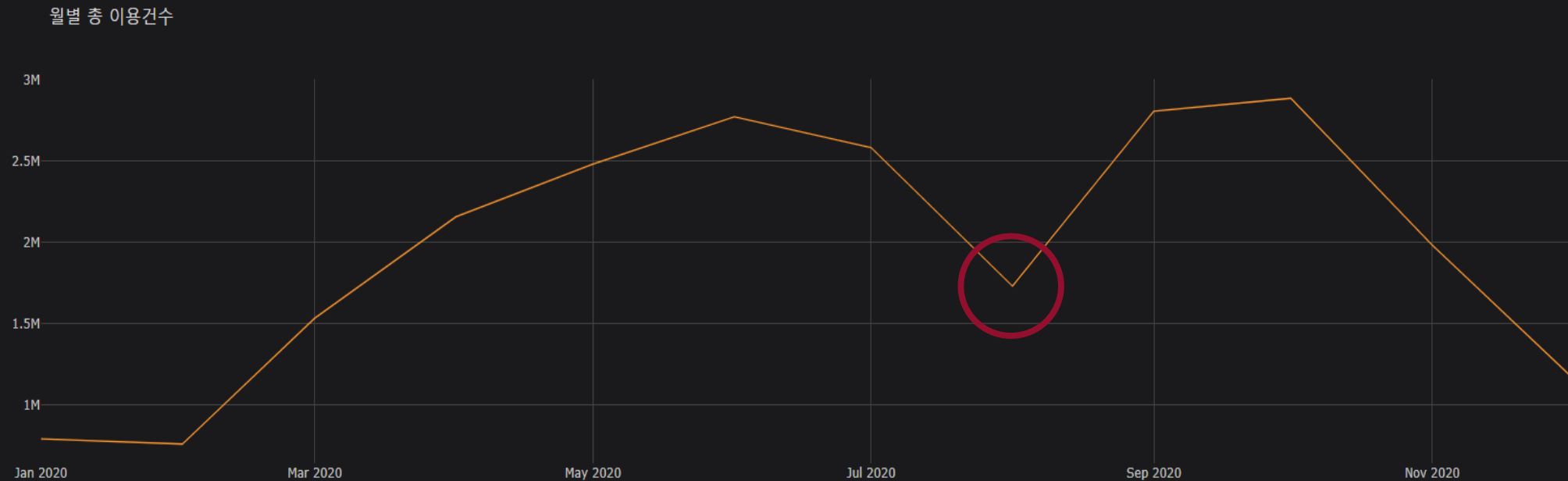
2020년 8월의 이용 건수가 현저히 **적은** 것으로 보아
'사회적 거리두기' 시행으로 공공 자전거 이용이 **감소**한 것으로 판단

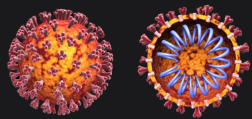




2-1 코로나 확진자 증감과 공공자전거 이용객 수, 이용 시간

2020년 8월의 이용 건수가 현저히 **적은** 것으로 보아
'사회적 거리두기' 시행으로 공공 자전거 이용이 **감소**한 것으로 판단





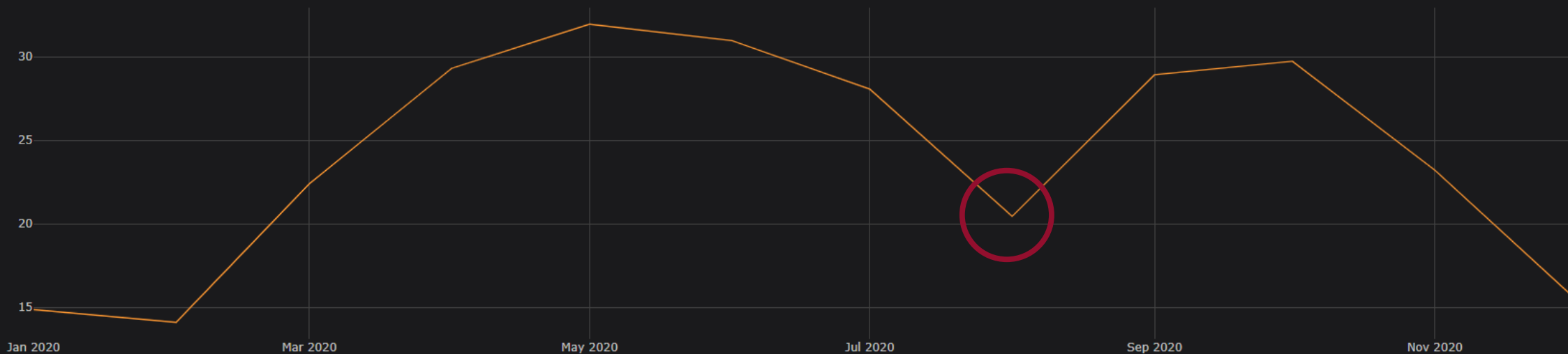
2-1 코로나 확진자 수 증감과 공공자전거 이용객 수, 시간

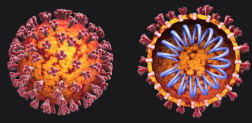
2020년 8월 총 이용 건수 감소, 1인 평균 이용 건수 감소

총 이용 감소율 < 1인 평균 이용 감소율

감소 유저 중 헤비 유저들의 비중이 많은 것을 알 수 있음

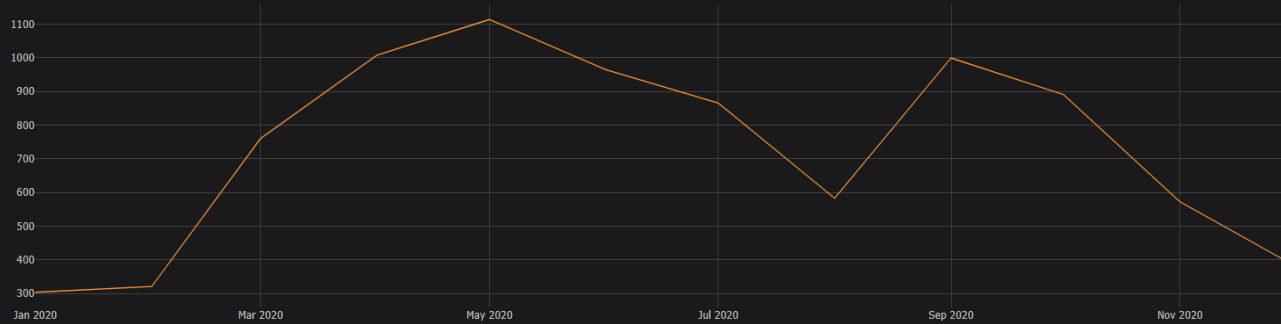
월별 한사람 평균 이용건수



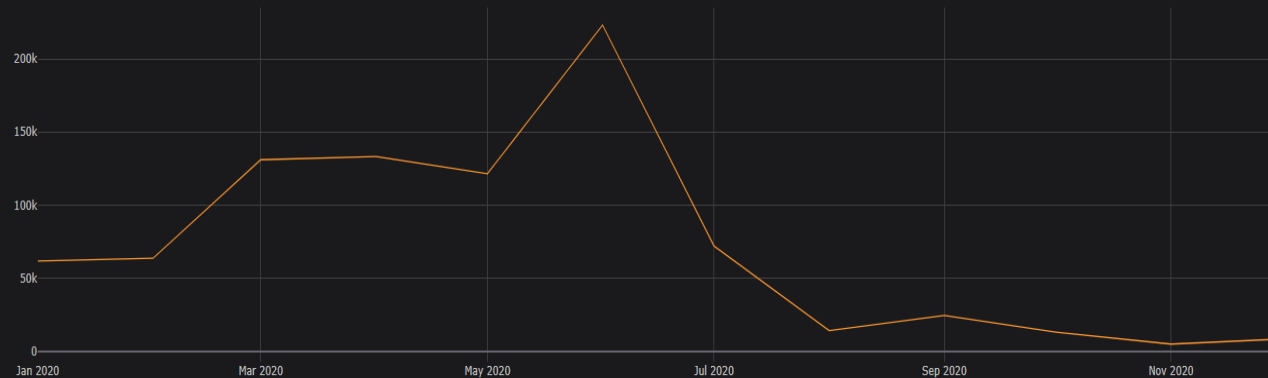


2-1 코로나 확진자 수 증감과 공공자전거 이용객 수, 시간

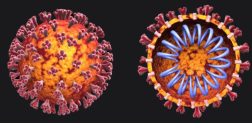
월별 한사람 평균 이용시간



월별 한사람 평균 이동거리

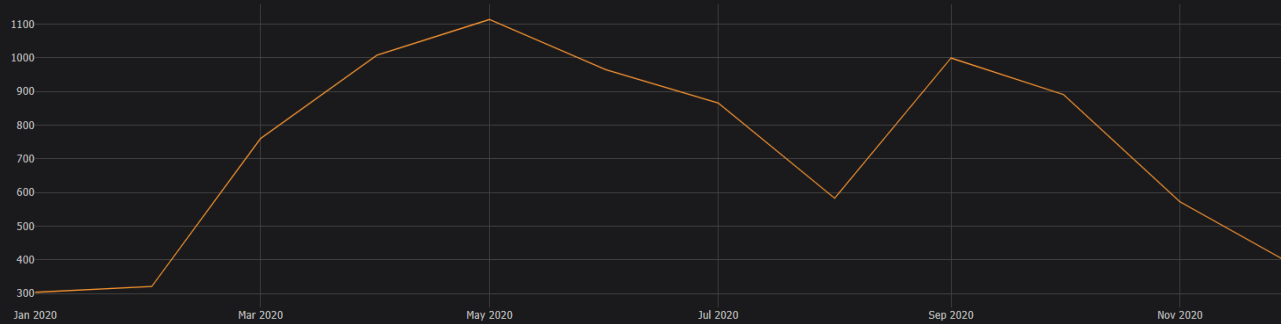


1인당 평균 이용 건수가 감소한 만큼
1인당 평균 이용 시간도 비슷한 폭으로 감소

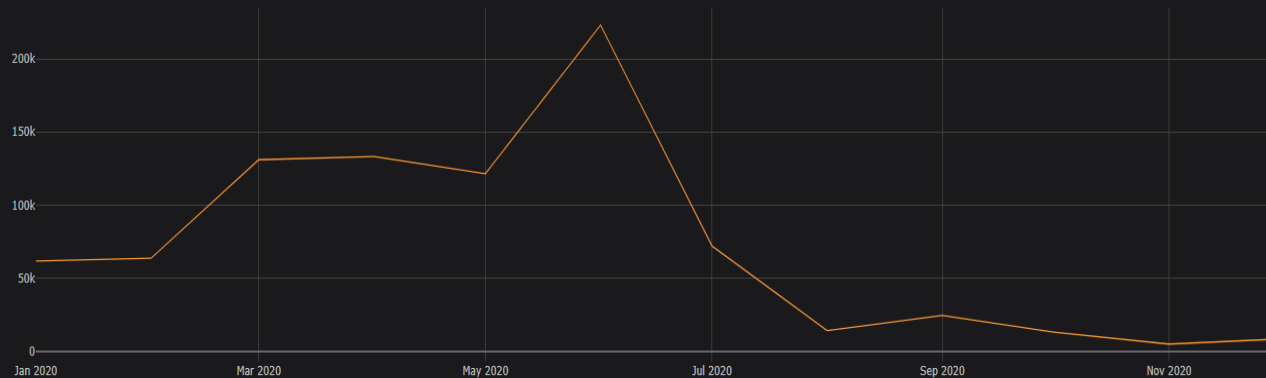


2-1 코로나 확진자 수 증감과 공공자전거 이용객 수, 시간

월별 한사람 평균 이용시간



월별 한사람 평균 이동거리

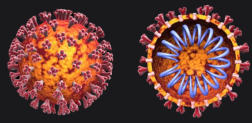


이동 거리는 이용 건수와 무관히
2020년 7월부터 급격 감소, 이후 저값 유지

데이터 문제가 아닐 시, 계절적 사회적 영향으로 볼 수 있으나
이용 시간이 인지되었음에도 측정된 이동 거리가 없다는 의문

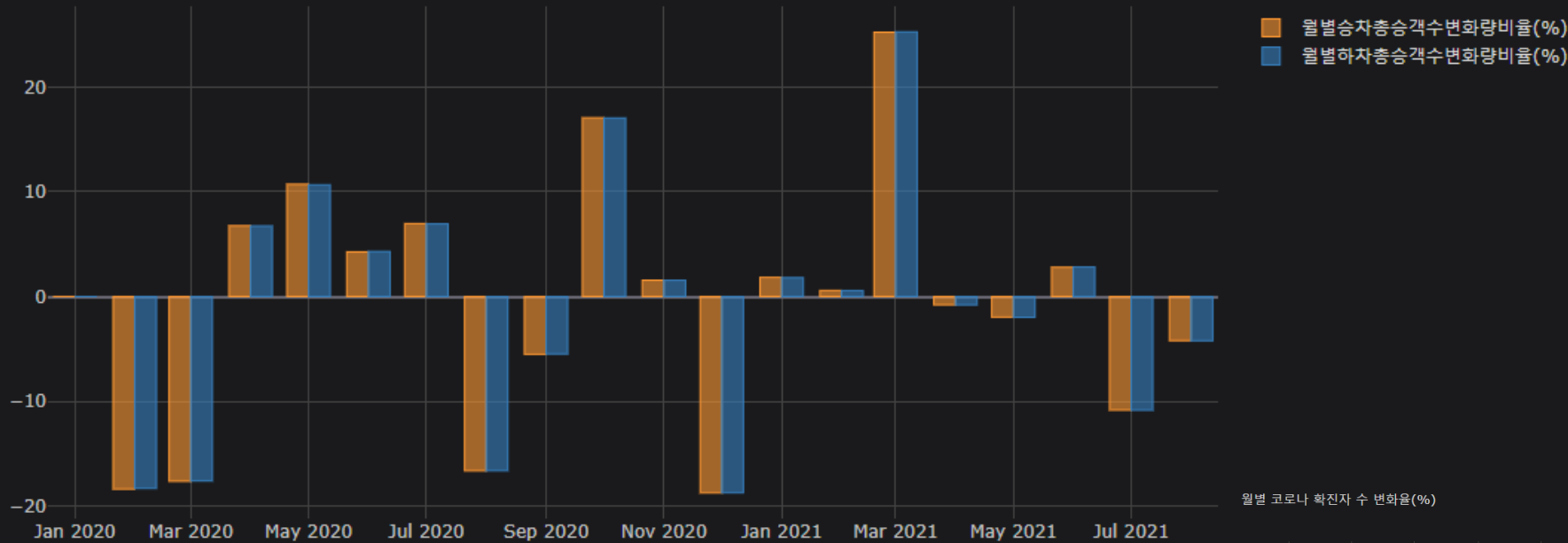
따라서 공공데이터의 이동거리 컬럼의 문제로 판단됨

2-2 코로나 확진자 증감과 지하철 승객수 변화

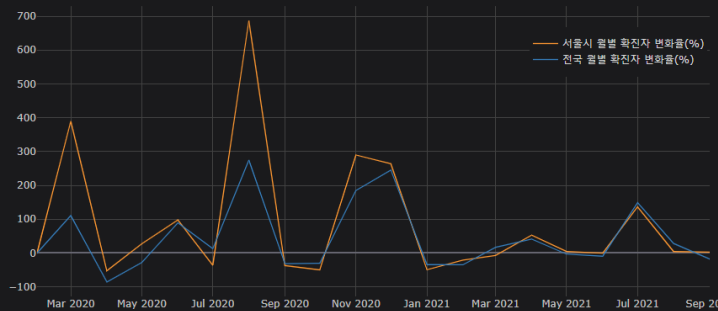


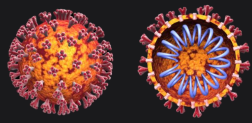
2-2 코로나 확진자 수 증감에 따른 지하철 승객 수 변화

월별 승하차승객수 변화 비율



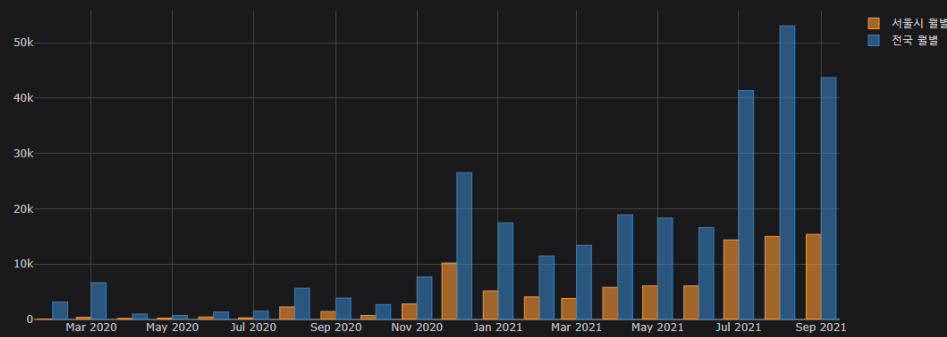
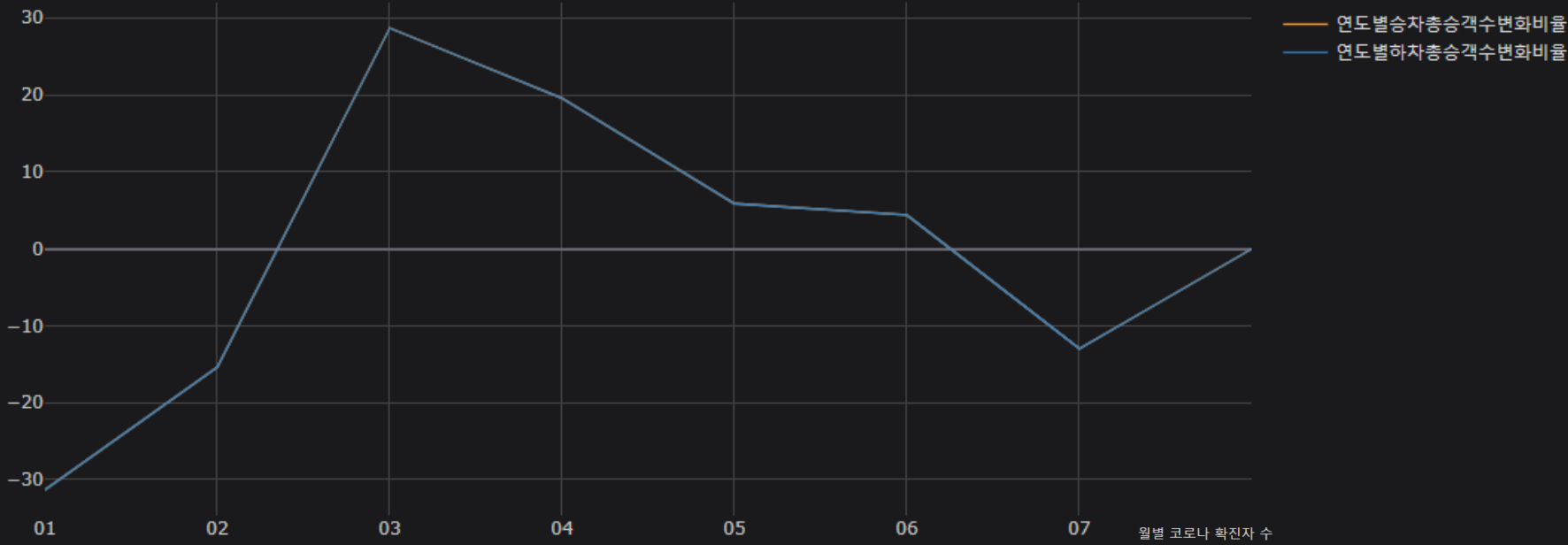
월별 코로나 확진자 수 변화율(%)



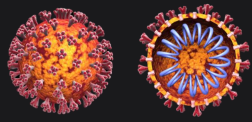


2-2 코로나 확진자 수 증감에 따른 지하철 승객 수 변화

2020년대비 2021년 총 승하차승객수 비율

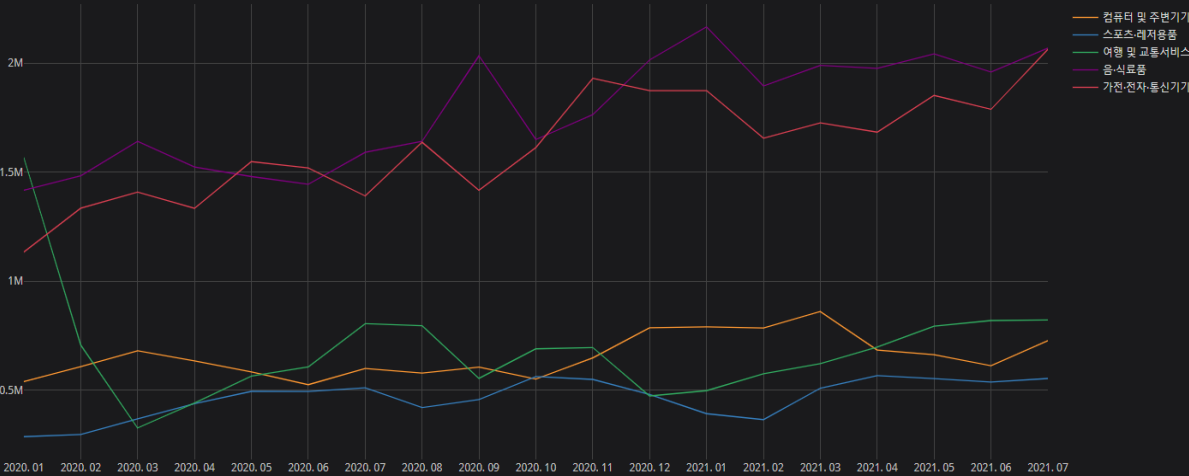


2-3 코로나 확진자 증감과 온라인 거래액

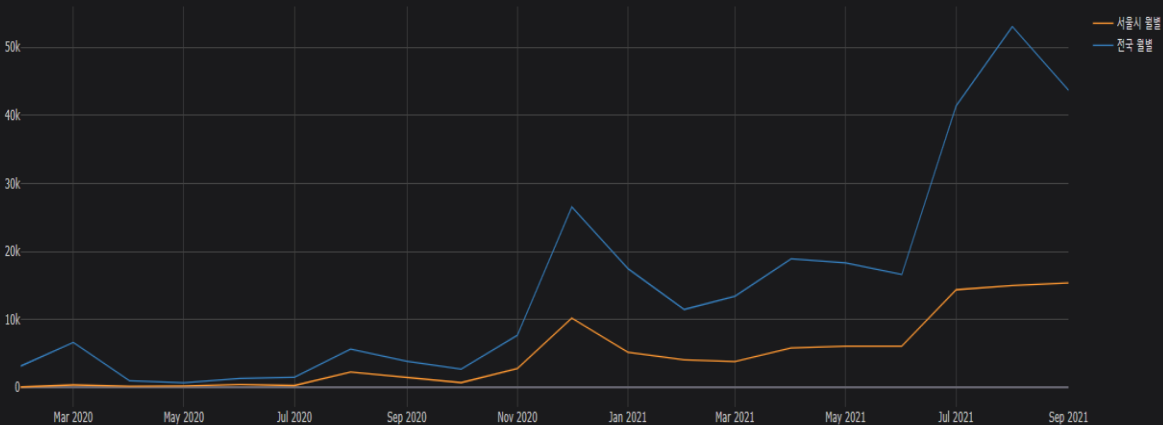


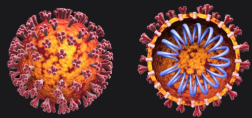
2-3 코로나 확진자 증감과 온라인 거래액

인터넷쇼핑몰 상품군별 거래액



코로나 확진자수





3 프로젝트 중 주요 이슈

코로나 확진자 수

- 확진자 수 일별로 존재, 데이터 상 숫자가 누적값이라 처리 필요.
- 해결) 해당 월 마지막 일을 잡아 차로 매달 확진자수 계산해 데이터 셋을 만들어 해결.
- 만들어진 데이터셋 값을 코로나 라이브의 월별 확진자 수와 대조 비교 확인.

공공 자전거

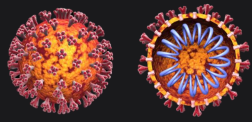
- 데이터셋에 결측값과 //N으로 표현된 값이 존재해 처리 어려.
- 해결) map을 이용해 특정 문자만 처리시 지정되지 않은 자료는 전부 NAN값으로 변하는 문제 발생. replace를 이용해 해결.

지하철 승객 수

- + 지하철 승객수
 - 데이터가 일별로 정리되어 월별로 정리가 필요.
- 처 - 해결) 날짜를 조건으로 지정하여 새로운 데이터셋을 만들어 해결.

온라인 거래액

- + 온라인 거래액
 - 그래프 y값이 계속 증가하는 그래프만 그려지는 문제.
- 해결) y축 값을 확인하니 오름차순이 아님을 발견. 문자형으로 인식된다 판단. info, dtype을 이용해 확인한 뒤 자료형을 변환해 해결.



4 개선 사항 (210930 기준)

변화를 파악하기 쉽도록 코로나 확진자 그래프와 각 항목 그래프를 한 그래프로 완성.

각 항목별 더 많은 분석 요인 추가.

코로나와 매출 품목별 변화 요인을 분석할 수 있는 데이터 추가.

그래프 변화 시점 주요 뉴스 제목을 시각화

Q&A