

Preprocess Time Series

Data

2017 - 2021 Active Population Data

Goal: forecast 2month data

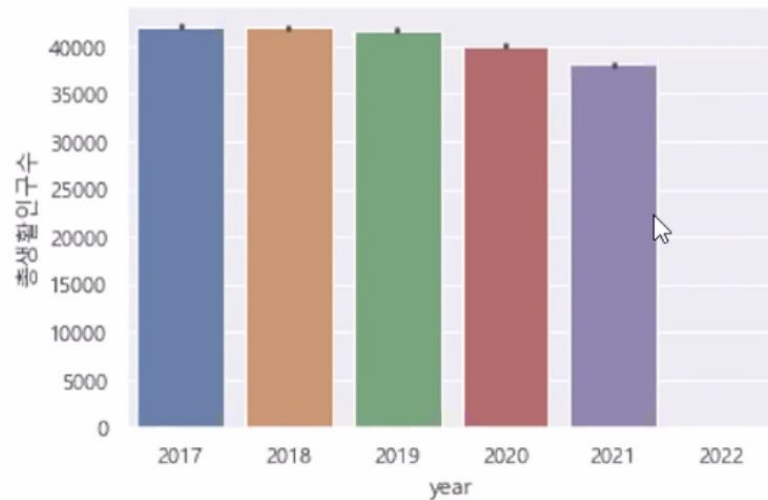
[illegible]

Domain knowledge

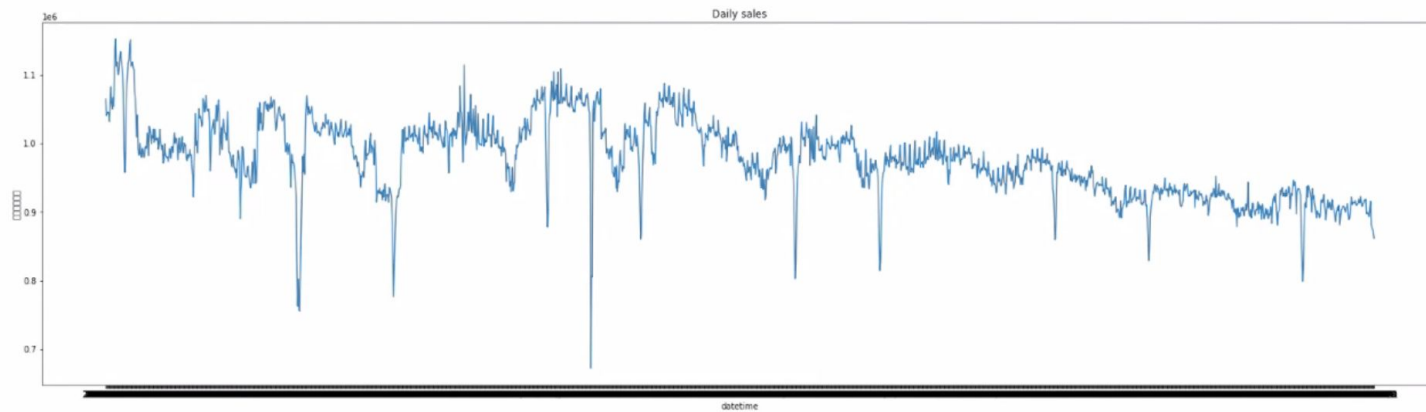
2019-2020 appearance of Covid-19

2021 infection rates peak

2022 ???



Outliers



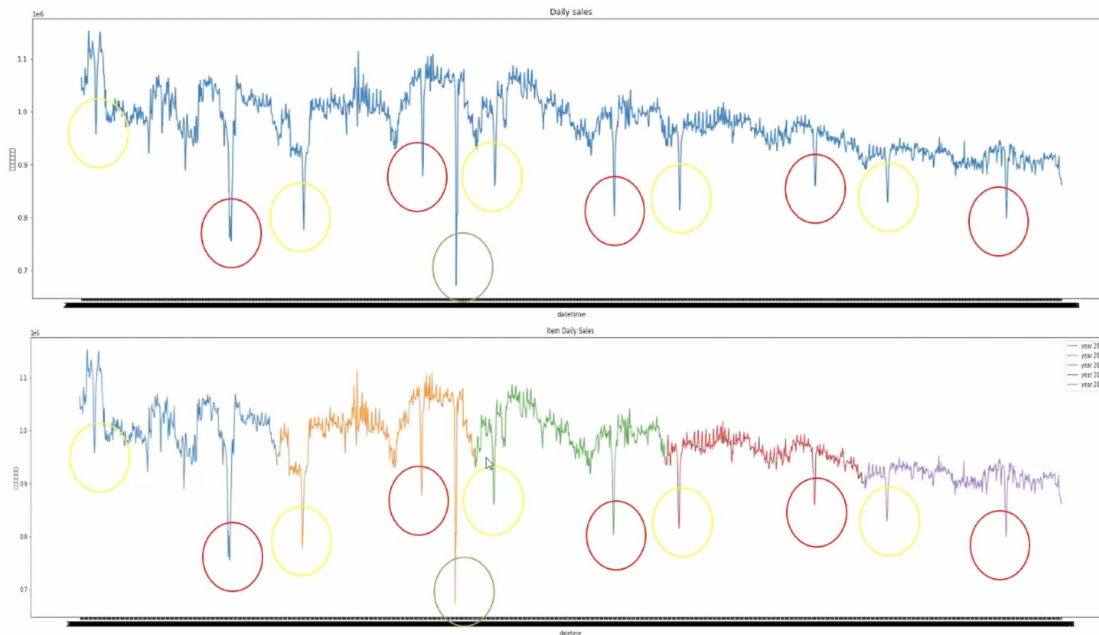
Getting Trend from Outliers

red: 추석

yellow: 설날

green: lost

(fire on data center)



Shift and Rolling

```
for i in range(1416,1416+(24*7)):
    df_total[f'총생활인구수_{i}lag'] = df_total['총생활인구수'].shift(i)
```

```
for i in range(1416,1416+(24*7)):
    df_total[f'총생활인구수_{i}이동평균'] = df_total[f'총생활인구수_{i}lag'].rolling(window=1416).mean()
```

시간대 구분	총생활인구수	년	월	일	weekday_0	weekday_1	weekday_2	weekday_3	weekday_4	..	총생활인구수_1574이동평균	총생활인구수_1575이동평균	총생활인구수_1576이동평균	총생활인구수_1577이동평균	총생활인구수_1578이동평균	총생활인구수_1579이동평균	총생활인구수_1580이동평균	총생활인구수_1581이동평균	총생활인구수_1582이동평균	총생활인구수_1583이동평균
2998	22	38563.1827	2017	5	5	0.0	0.0	0.0	0.0	1.0	--	44198.816904	44200.220788	44201.163320	44202.154401	44202.451685	44202.743557	44202.961268	44203.568294	44204.293558
2999	23	38857.9001	2017	5	5	0.0	0.0	0.0	0.0	1.0	--	44197.611944	44198.816904	44200.220788	44201.163320	44202.154401	44202.451685	44202.743557	44202.961268	44203.568294
3000	0	39289.3581	2017	5	6	0.0	0.0	0.0	0.0	0.0	--	44196.473656	44197.611944	44198.816904	44200.220788	44201.163320	44202.154401	44202.451685	44202.743557	44202.961268
3001	1	39498.6305	2017	5	6	0.0	0.0	0.0	0.0	0.0	--	44195.265283	44196.473656	44197.611944	44198.816904	44200.220788	44201.163320	44202.154401	44202.451685	44202.743557
3002	2	39392.0822	2017	5	6	0.0	0.0	0.0	0.0	0.0	--	44194.117890	44195.265283	44196.473656	44197.611944	44198.816904	44200.220788	44201.163320	44202.154401	44202.451685
--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
43531	19	33304.5400	2021	12	31	0.0	0.0	0.0	0.0	1.0	--	38004.516596	38004.197890	38003.628703	38003.203196	38002.837111	38002.590300	38002.462235	38001.357871	38000.904904
43532	20	33308.7201	2021	12	31	0.0	0.0	0.0	0.0	1.0	--	38004.932074	38004.516596	38004.197890	38003.628703	38003.203196	38002.837111	38002.590300	38002.462235	38001.357871
43533	21	34063.1721	2021	12	31	0.0	0.0	0.0	0.0	1.0	--	38006.025245	38004.932074	38004.516596	38004.197890	38003.628703	38003.203196	38002.837111	38002.590300	38001.357871
43534	22	34380.0806	2021	12	31	0.0	0.0	0.0	0.0	1.0	--	38006.084813	38006.025245	38004.932074	38004.516596	38004.197890	38003.628703	38003.203196	38002.837111	38002.590300
43535	23	37074.2785	2021	12	31	0.0	0.0	0.0	0.0	1.0	--	38005.395813	38006.084813	38006.025245	38004.932074	38004.516596	38004.197890	38003.628703	38003.203196	38002.837111

result

RMSE 982