

# 경기도 코로나19 예측

**Team BITAmin**

# 목차

---



참가 동기 및 목적



모델 구현 방법



참가정보 및 예측결과



정책 제언



활용데이터 및 가공방법

# 참가 동기 및 목적

---



코로나바이러스감염증-19(이하 코로나)로 인하여 전 세계적인 팬데믹 현상을 겪는 가운데, 다가올 국내 최대 명절인 추석에 대한 대비가 필요한 상황이다. 객관적인 데이터를 바탕으로 한 감염자예측을 통해 방역에 대한 경각심을 일깨우고, 자체적인 정책 제언으로 도민 안전에 이바지하기 위해 공모전에 참가하였다.



합리적인 가설을 세우고 구체적으로 데이터를 분석하여 기간내 실제 감염자수와 근사한 수치를 예측한다. 또한 다양한 각도에서 데이터를 수집하여 코로나 감염의 피해를 막기 위한 실용적인 정책을 제언한다.

# 참가정보 및 예측결과

참가정보					
팀명 및 대표자 성명	BITAmin / 김규진		참여 구분	팀(5명)	
팀 구성원 성명	김동현 / 이정민 / 이주희 / 임세종		소속 (학교/연구소/기업명)	성균관대학교 시스템경영공학과	
대표자 전화번호	010-2006-8248		대표자 email 주소	<a href="mailto:gjkim0113@gmail.com">gjkim0113@gmail.com</a>	
모델링 정보 요약	개발 환경	Python Jupyter Notebook / 3.7.4			
	활용 데이터	경기도 확진자 수 데이터 / 서울시 코로나19 확진자 데이터 / 한국도로공사 교통량 데이터(19,20년) / 서울시 카드이용 데이터			
	모델 소개	교통량, 가맹점 이용건수, 거리두기 단계별 확진자 변수를 통해 코로나 위험도 측정 지수 R을 계산하여 미래에 나타날 확진자를 예측한다.			
유입환자 예측 결과					
일차	9/30	10/1	10/2	10/3	10/4
환자 수	18	17	16	16	15

# 가설 및 변수설정(1)

가설 1

실내활동이 증가함에 따라, 확진자 수가 증가할 것이다.

강력한 거리두기 기간  
(3월22일 ~ 4월 19일)



정부에서 강력한 거리두기 정책을 시행한 결과, 수도권 지역의 확진자 수가 500여 명에서 100여명으로 80%가량 감소했다. 이 때, 실내 활동이 감소함

생활 속 거리두기 기간  
(5월 15일 ~ 6월 15일)



정부에서 생활 속 거리두기 정책을 시행한 결과, 수도권 지역의 확진자 수가 160여 명에서 1000여명으로 500%가량 증가했다. 이 때, 실내 활동이 증가



카드 데이터를 활용하여, 거리두기 단계별로 실내활동 증감을 분석

→ 실내 활동이 증가하면, 확진자 수도 확연히 증가하는 것을 확인

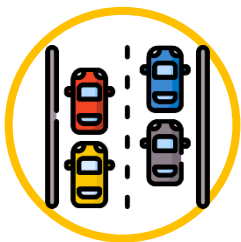
# 가설 및 변수설정(2)

---

## 가설 2

추석 연휴기간 동안 사람들의 움직임에 변화가 있을 것이다

### 명절기간 교통량 및 휴게소 이용증가



추석 연휴기간 유동인구 증가에 따라 고속도로 및 휴게소 이용이 늘어날 것이다.

### 코로나로 인한 작년대비 교통량 감소



코로나 감염위험 증가로 인해 많은 사람들이 작년 추석과 비교하여 귀성을 자제할 것이다.

### 추석기간 특정 실내 점포이용 증가



추석 연휴기간에 특정 업종 점포의 이용율이 연평균보다 높아질 것이다.

# 가설 및 변수설정

---

## 최종가설

교통량 증가율과 점포이용율을 통해서 확진자 수를 예측할 수 있다

- 가설 1에 따르면 추석 연휴기간의 교통량 증가에 따라 휴게소 이용이 증가하고, 휴게소 내에 많은 사람들이 몰리게 되면서 코로나 감염위험이 높아질 것이다.
- 가설 1에 따르면 추석기간에는 특정 점포이용이 증가한다. 이는 설연휴 기간의 점포이용율을 통해 확인할 수 있다. 또한 가설 2에 따르면 생활 속 거리두기 단계는 실내활동 증감에 영향을 미친다. 따라서 추석기간에는 특정 점포이용이 증가할 것이므로 실내활동이 증가하여 코로나 감염위험이 높아질 것이다.

➔ 이에 따라 교통량 증가율과 점포이용율을 통해서 감염위험정도와 확진자 수를 예측할 수 있다.

## 변수 설정

$R(\text{코로나 위험도 측정지수}) = F(\text{확진자}) * C(\text{카드이용건수}) * T(\text{교통량})$

# 활용데이터 및 가공방법(1)

## 경기도 확진자 수 데이터

- 경기도 감염병관리지원단

주어진 데이터에서 확진일을 기준으로 일일 확진자 수 데이터를 가공하였음.

→ 경기도 날짜별 확진자와 누적 확진자 수 도출.

```
numbers['누적 확진자수']=np.cumsum(numbers['일일 확진자수'],axis=0)
numbers.head()
```

	날짜	일일 확진자수	누적 확진자수
0	2020-01-26	1	1
1	2020-01-27	1	2
2	2020-02-01	1	3
3	2020-02-02	2	5
4	2020-02-05	2	7

## 서울시 코로나19 확진자 현황

- 서울특별시 감염병관리지원단

원본데이터에서 확진일을 기준으로 환자번호를 추출하여, 확진일별 환자수를 count.

→ 서울시 날짜별 확진자와 누적 확진자 수 도출

```
seoul_number=seoul['환자번호'].groupby(seoul['확진일']).count()
seoul_number=pd.DataFrame(seoul_number)
seoul_number.head()
```

환자번호	
확진일	
1.24.	1
1.30.	3
1.31.	3
2.16.	2
2.19.	2



# 활용데이터 및 가공방법(2)

## 주간 교통분석(교통량)

- 한국도로공사 고속도로 공공데이터 포털

19년 추석연휴 9/12~15일의 날짜 별 당일교통량을 각각 전주교통량, 전전주교통량, 전월평균 교통량과 비교. →

추석 연휴기간의 교통량 증가율 도출

19년과 20년의 3~8월 교통량을 비교.

→ 전년대비 올해 교통량 증가율 도출

## 영업소별 교통량

- 한국도로공사 고속도로 공공데이터 포털

: 여름휴가 시즌을 7월말~8월초 2주를 기준으로 두고, 19년과 20년 여름휴가시즌의 평균 교통량(1,2종)을 비교.

여름 휴가 시즌: 2019년 7/27~8/11, 2020년 7/25~8/9

→ 2020년 여름휴가 교통량 감소율 도출

# 추석의 교통량 증가율

```
chuseok19['전주rate']=(chuseok19['당일교통량']-chuseok19['전주교통량'])/chuseok19['전주교통량']*100
chuseok19['전전주rate']=(chuseok19['당일교통량']-chuseok19['전전주교통량'])/chuseok19['전전주교통량']*100
chuseok19['전월평균rate']=(chuseok19['당일교통량']-chuseok19['전월평균교통량'])/chuseok19['전월평균교통량']*100
chuseok19
```

	당일교통량	전주교통량	전전주교통량	전월평균교통량	전주rate	전전주rate	전월평균rate
일자							
20190911	5773125	4950710	5322642	5447679	16.612062	8.463522	5.974030
20190912	5764190	5052335	5299278	5447679	14.089624	8.773120	5.810016
20190913	7017734	5587281	5882335	5447679	25.601952	19.301842	28.820623
20190914	5858019	3602932	6138893	5447679	62.590329	-4.575320	7.532382

```
summer20=pd.merge(Traffic20_08,Traffic20_07, left_index=False, right_index=False, how='outer' )
summer20
```

	1종교통량	2종교통량	3종교통량	4종교통량	5종교통량	6종교통량	총교통량
0	8174573	122308	173969	111532	149633	324264	9056279
1	7363218	60708	124290	53644	42670	279134	7923664
2	7242527	191242	239999	179154	242652	342241	8437815
3	7132665	206295	253075	202874			

```
summer20=summer20.mean()
summer20
```

1종교통량	6.899330e+06	1종교통량	-6.282803
2종교통량	1.747734e+05	2종교통량	-5.993745

# 19년 여름휴가대비 20년 여름휴가 교통량 변화  
(summer20-summer 19)/summer 19\*100

# 활용데이터 및 가공방법(3)

## 명절(전,중,후)기간 영업소 매출분석

### - 데이콘 서울시 카드이용데이터

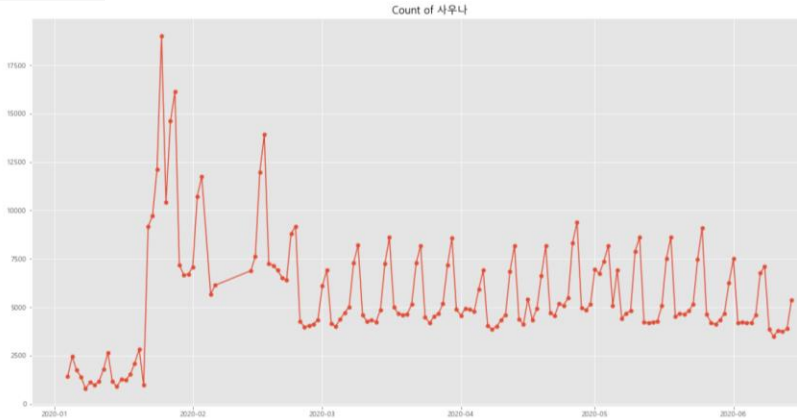
서울시 카드데이터분석을 통해 올해 설연휴 1/23~27  
기간을 전,중,후로 나누어 연 평균매출과 비교.

→ 명절기간의 가맹점 이용증가율 도출

```
In [28]: #설중반기간 카드데이터건수가 높은 지수 상위 20  
card_vs_m.head(20)
```

Out [28]:

	Store	Count	Count_Mean	Count_Seol	Count_Seol_Mean
106	티켓	141856	915.200000	8745	2915.000000
53	사우나	874782	5643.754839	44047	14682.333333
60	특급 호텔	559507	3609.722581	26110	8703.333333
115	기타가구	109527	706.625806	4991	1663.666667
82	유흥주점	258671	1668.845161	10688	3562.666667
51	영화관	940426	6067.264516	38675	12891.666667
38	고속 버스	1882166	12143.006452	74339	24779.666667
84	볼 링 장	246848	1592.567742	8112	2704.000000
95	사무통신기기수리	186809	1205.219355	5367	1789.000000
18	기타레저업	7630200	49227.096774	217120	72373.333333
78	여 객 선	297454	1919.058065	8358	2786.000000
103	일반백화점	152701	985.167742	4107	1369.000000



## 사회적 거리두기 기간 영업소 매출분석

### - 데이콘 서울시 카드이용데이터

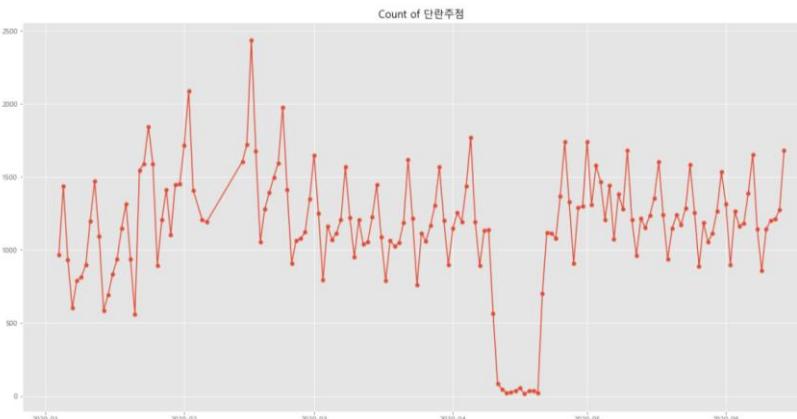
1차, 2차 사회적거리두기 기간동안 카드지출 횟수가 연  
평균과 비교하여 감소한 상위리스트를 이끌어 냄.

→ 사회적거리두기 기간동안 가맹점 이용 감소율 도출

```
In [90]: #1차 거리두기간동안 매출이 평균보다 낮았던 상위리스트  
card_vs_first = card_vs_first.sort_values(by=['Std'])  
card_vs_first.head(20)
```

Out [90]:

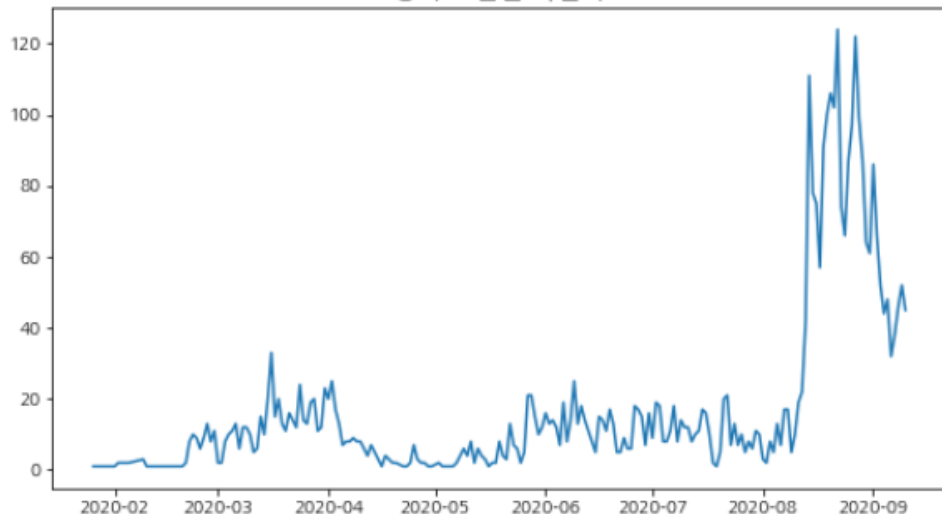
	Store	Count	Count_Mean	Count_Seol	Count_S
52	영화관	940426	6067.264516	58629	202
68	항 공 사	439718	2836.890323	40243	138
83	유흥주점	258671	1668.845161	26588	97
100	헬스 클럽	172221	1111.103226	18413	63
114	예체능학원	113355	731.322581	13018	44
60	유아원	571912	3689.754839	66930	230
99	기타응역서비스	176173	1136.600000	20839	77
97	단란주점	179581	1158.587097	24019	82
107	티켓	141856	915.200000	19543	67
53	노래방	878709	5669.090323	129083	445



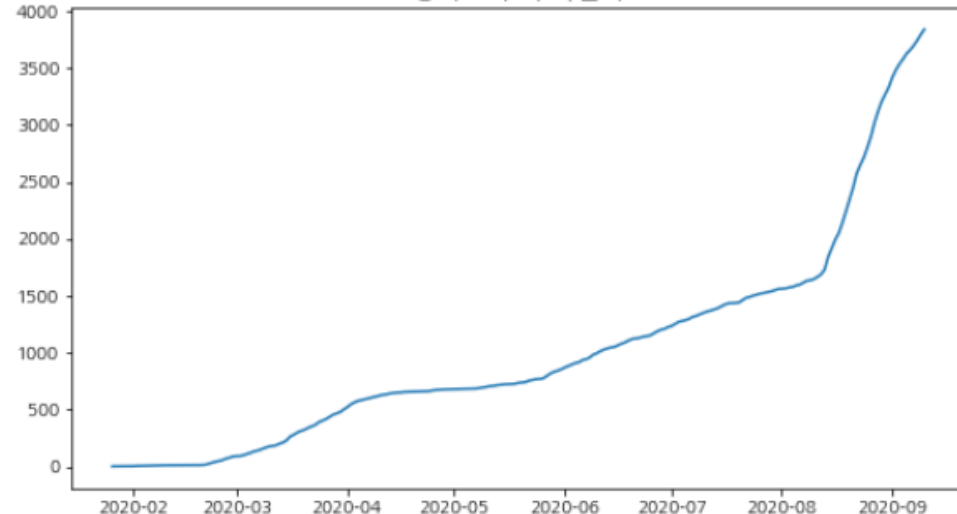
# 모델 구현 방법(1)

---

경기도 일별 확진자



경기도 누적 확진자



**경기도 확진자 수의 추세**

# 모델 구현 방법(2)

---

지수평활법 : 가장 최근 데이터에 가장 큰 가중치가 주어지고 시간이 지남에 따라 가중치가 기하학적으로 감소되는 가중치 이동 평균 예측 기법

$$F_t = \alpha A_{t-1} + \alpha(1-\alpha)A_{t-2} + \alpha(1-\alpha)^2 A_{t-3} + \dots$$

( $F_t$ : t시점의 예측값,  $\alpha$ : 평활상수(0.05~0.30),  $A_{t-1}$ : (t-1)시점의 실제 값)

```
#a는 0.05에서 0.3사이
a=0.2
def 지수평활(list):
    pre=[]
    for i in range(len(list)):
        p=a*((1-a)**(len(list)-1-i))*list[i]
        pre.append(p)
    return sum(pre)
```

경기도 감염병 관리 지원단의 경기도 확진자 수 데이터를 리스트화 해서 사용

# 모델 구현 방법(3)

## 정책시행에 따른 R값의 변화

- 강력한 사회적 거리두기(R1,R2) [3월 22일 - 4월 19일]
- 생활 속 사회적 거리두기(R3,R4) [5월 6일 - 6월 28일]

R1 대비 R2는 약 1/4배 감소

→ 강력한 사회적 거리두기를 시행한 결과 위험도가 감소

R3 대비 R4는 약 2배 증가

→ 생활 속 사회적 거리두기를 시행한 결과 위험도가 증가

$R(\text{코로나 위험도 측정지수}) = F(\text{확진자}) * C(\text{카드이용건수}) * T(\text{교통량})$

$F_t = (t/2)\text{기간 확진자수} / t\text{기간 확진자수}$

$C_t = (t/2)\text{기간 카드이용건수} / t\text{기간 카드이용건수}$

$T_t = (t/2)\text{기간 교통량} / t\text{기간 교통량}$

```
R1 = st1 * sc1 * sf1
R2 = st2 * sc2 * sf2
R3 = nt1 * nc1 * nf1
R4 = nt2 * nc2 * nf2
```

```
print(R1,R2)
```

```
0.19865491578256694 0.050294069390067114
```

```
print(R3,R4)
```

```
0.08048648083660012 0.17265552040293805
```

가정)

1. 특정 정책이 시행한 뒤 2주후부터 정책의 효과가 나타난다.
2. 정책시행기간은 한달이며, 가운데 날짜를 기준으로 2주전후로 나눈다.  
ex) R1은 강력한 사회적 거리두기 가운데 날짜 기준 2주전, R2는 2주후이다.

# 정책 제언(1)



휴게소 QR Pass

## QR패스를 이용한 휴게소 이용객 분산 정책

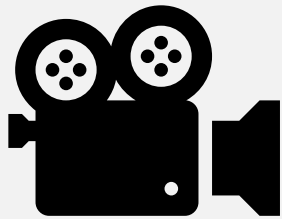
: 교통량을 임의로 억제하는 방법보다 이용객이 많은 휴게소(안성, 죽전, 화성, 기흥 등)를 분산시키는 정책을 펼침.

경기도내 휴게소마다 QR Pass를 통해 출입, 나감을 체크한 뒤 실시간으로 휴게소 인원을 확인하는 어플을 개발함.

귀성길에 가더라도 이용객이 많은 휴게소를 가급적 지양할 수 있도록 안내문구를 띄우고 다음에 거쳐갈 휴게소와 비교해줌.

이는 또한 휴게소내 코로나 감염자가 발생할 경우 QR패스 로그데이터를 바탕으로 진료대상자 파악을 빠르게 진행할 수 있음

## 정책 제언(2)



가맹점 QR Pass

2.

### QR패스를 이용한 가맹점 이용객 분산 정책

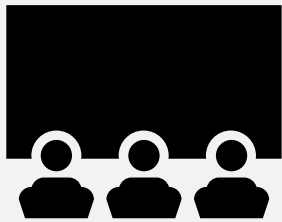
: 휴게소 인원분산 정책과 더불어 일반 가맹점 이용객의 현황을 확인할 수 있도록 하여 추석직전 소비자가 몰리지 않도록 도움.

3월에 공적마스크 재고 알림이 어플과 유사한 알고리즘으로, 지도에 실시간 이용객 현황을 확인하여 과도하게 밀집된 상황을 방지할 수 있도록 함.

2.5단계 혹은 3단계 거리두기로 전염병 예방을 위한 조치를 취할 수 있지만, 명절과 연휴에 완전히 영업을 중지하지 않고 적절한 이용객 분산을 통해 상권과 방역 두 마리 토끼를 잡을 수 있도록 함.

이 정책 역시 가맹점 내 코로나 감염자가 발생할 경우 QR패스 로그데이터를 바탕으로 진료대상자 파악을 빠르게 진행할 수 있음.

# 정책 제언(3)



가맹점 QR Pass

3.

## 경기도 내 언택트 문화생활 지원

: 경기도 내에서 추석기간에 많은 언택트 문화행사들을 지원하여, 추석을 안전하고 즐겁게 보낼 수 있도록 돕는다.

온라인 미술관(VR, 유튜브 채널)

온라인 공연

: 경기아트센터의 공연과 전시를 온라인으로 진행

전자도서관(E-book)

: 경기도 내 도서관에서 E-book을 대여하거나, 드라이브 스루 형식으로 책을 대여

드라이브 인 콘서트 (콘서트 또는 자동차 극장)

드라이브 스루 국내여행

: 경기도 내 관광명소 "차박캠핑"(차안에서 숙식을 해결), 드라이브