

정부의 사회적 거리두기 정책이 배달음식 주문량에 미치는 영향

융합대학원 소셜데이터사이언스 전공

배승예, 이지인, 정승민, 황수현

통계적 연구방법론 final project

2021.12.14 (화)

목차

- 연구배경
- 가설 설정 및 변수 정의
- 데이터 가공
- 분석 및 결과
- 결론

팬데믹은 우리 사회에 어떤 영향을 끼쳤을까

- Covid19가 한국 사회에 끼친 영향 중 **식생활 변화**에 주목
 - 정부의 대표적인 방역 정책 중 하나 ⇒ 사회적 거리두기
 - 사적모임인원제한, 식당영업시간 제한 ⇒ 사회 구성원들의 외식 생활에 변화가 불가피해짐

핵심질문

코로나 19는 국민의 식생활에 어떤 영향을 미쳤을까?



연구가설

정부의 사회적 거리두기 정책은 배달음식 주문량에
영향을 미쳤을 것이다

-
- 연구배경
 - 가설 설정 및 변수 정의
 - 데이터 가공
 - 분석 및 결과
 - 결론

가설 설정 및 변수 정의

- 정부의 사회적 거리두기 정책이 배달음식 주문량에 영향을 미칠 것이라는 가설의 증명을 위한 변수 정의

가설 설정

- 정부의 사회적 거리두기 정책은 배달음식 주문량에 영향을 미칠 것이다.

독립변수

- 체감 거리두기 지수

- 정부의 거리두기 정책이 변화함에 따라 일정 기준을 통해 점수화
- 수도권 지역의 이동량을 반영하여 실제 체감된 거리두기 지수를 산정

통제변수

- 수도권 지역 일 평균 기온

- 배달 주문의 증감에 영향을 미칠 수 있는 요인으로 일 평균 기온을 반영

종속변수

- 수도권 지역 배달음식 주문량

- 수도권 지역에서 나타난 배달 주문량

-
- 연구배경
 - 가설 설정 및 변수 정의
 - 데이터 가공
 - 분석 및 결과
 - 결론

데이터 현황

- 데이터 별 날짜 기준이 상이하여 일별 데이터의 경우 주별로 정리

- 분석 기간: 2020.07.01~ 2021.07.31 (13개월) → 20년 7월 1주차 ~ 21년 7월 4주차
- 분석 지역: 서울, 인천, 경기도

배달량 데이터

- 출처: 경기대학교 빅데이터 센터 (bigdata-telecom.kr)
- 날짜: 일별 → 주별 평균으로 변경
- Columns
 - 지역 / 업종 / 배달 주문량
- 종속변수

이동량 데이터

- 출처: 통계데이터센터 (data.kostat.go.kr)
- 날짜: 주별
- Columns
 - 거주지 (한달동안 00-06시까지 가장 오랜 시간 체류한 기지국 위치)
 - 이동량 (본인이 실거주 하는 시군구 외 지역 30분 이상 체류)

거리두기 데이터

- 출처: 해당 날짜 뉴스를 통해 수집
- 날짜: 일별 → 주별 평균으로 변경
- Columns
 - 사적모임 가능인원, 식당영업시간, 카페영업시간

기온 데이터

- 출처: 기상청 기상자료 개방포털 (data.kma.go.kr)
- 날짜: 일별 → 주별 평균으로 변경
- Columns
 - 지역 / 일평균 기온
- 통제변수

표준 거리두기 지수 (1/2)

- 사회적 거리두기 정책을 독립변수로 정의하기 위한 수치화 작업 진행**
 - 3단계 체제, 5단계 체제 등 거리두기 체제가 계속 변화해 표준화 필요
 - 시간에 따른 정부의 거리두기 정책 변동 사항을 정리하여 수치화 진행
- 거리두기 점수 = $\frac{1}{2} * (\text{실내취식 가능 시간(식당)}) + \frac{1}{2} * (\text{실내취식 가능 시간(카페)}) + \text{사적모임 가능 인원}$**
 - 실내 취식 가능 시간의 영향력이 업종에 따라 과대 반영되는 것을 막기 위해 0.5의 가중치를 둠
 - 실내 취식 가능 시간(식당)
예) 21시까지 영업 가능 시, $(21/24) * \frac{1}{2}$
 - 실내 취식 가능 시간(카페)
예) 21시까지 영업 가능 시, $(21/24) * \frac{1}{2}$
 - 사적모임 가능 인원
예) 50인 집합 가능 시, $50/100$
- 표준 거리두기 지수 = $1 / (\text{거리두기 점수})$**

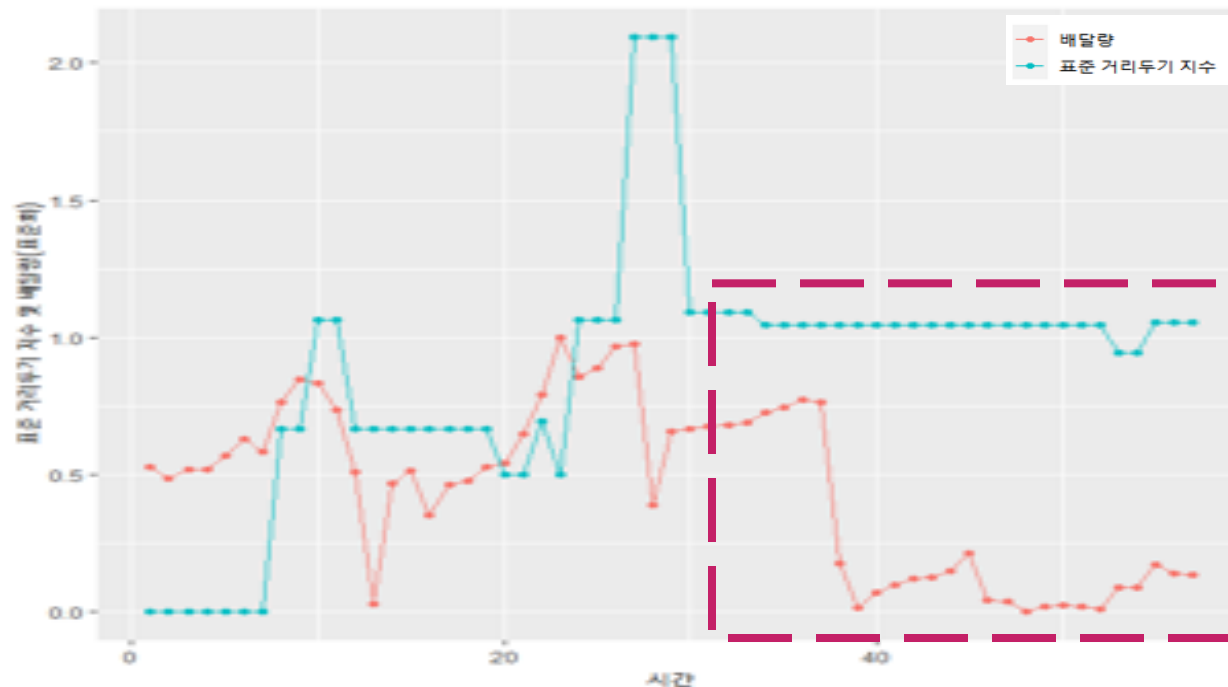
거리두기 지수 정의

지수	내용			
거리두기 점수 (Z)	실내 취식 가능시간	식당	$x = l/24$	$Z = 1/2 * x + 1/2 * y + z$
		카페	$y = m/24$	
	사적 모임 가능인원		$z = n/100$	
표준 거리두기 지수(D)	$D = 1/Z$			
이동 변화율(m)	해당년도 일평균 이동량 / 기준년도 일평균 이동량			
체감 거리두기 지수(PD)	$PD = 1/\{Z(\text{scaled}) + m(\text{scaled})\}$			

표준 거리두기 지수 (2/2)

- 표준 거리두기 지수의 경우 계단모양의 그래프가 나타나며, 일부 구간에서 거리두기 지수와 배달 주문량이 서로 다른 양상을 보임
 - 보다 실질적 의미의 거리두기 지표의 필요성

시간에 따른 표준 거리두기 지수와 배달량



체감 거리두기 지수 (1/2)

- 체감 거리두기 지수의 경우 수도권 인구 이동량 지표를 활용하여 계산
 - 인구 이동 변화율을 표준거리두기지수에 더함으로써 사람들이 거리두기에 어떻게 반응했는지를 지수에 반영함
- 이동 변화율 = 해당년도 일평균 이동량 / 기준년도 일평균 이동량
 - 이동 = 본인이 실거주 하는 시군구 외 타 시군구의 행정동을 방문하여 30분 이상 체류한 경우
 - 기준년도: 코로나19 발생 전인 2019년을 기준 년도로 삼음
평년(2019년) 대비 얼마나 많은 이동량 증감이 있었는가 측정함

- 체감 거리두기 지수
= 1 / (표준화된 거리두기 점수 + 표준화된 이동 변화율)
 - 표준화 방법: Min.Max.Scaling

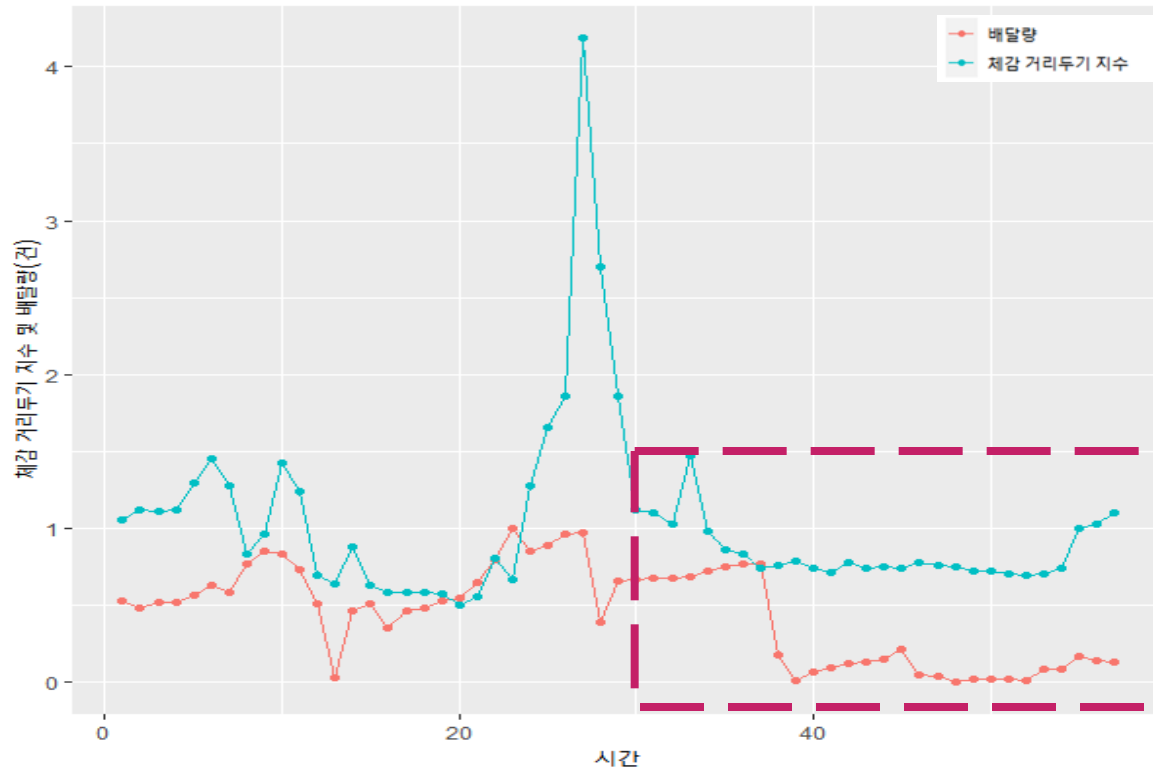
거리두기 지수 정의

지수	내용			
거리두기 점수 (Z)	실내 취식 가능시간	식당	$x = l/24$	$Z = 1/2 * x + 1/2 * y + z$
		카페	$y = m/24$	
	사적 모임 가능인원		$z = n/100$	
표준 거리두기 지수(D)	$D = 1/Z$			
이동 변화율(m)	해당년도 일평균 이동량 / 기준년도 일평균 이동량			
체감 거리두기 지수(PD)	$PD = 1/\{Z(\text{scaled}) + m(\text{scaled})\}$			

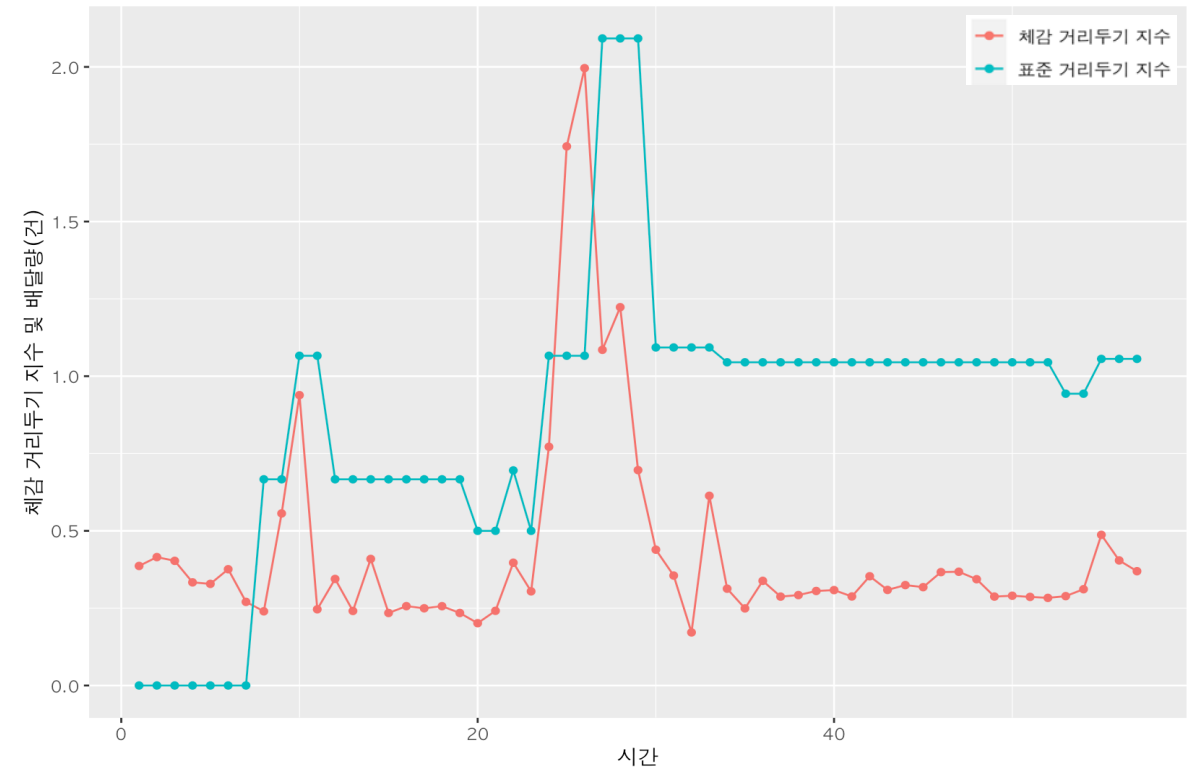
체감 거리두기 지수 (2/2)

- 체감 거리두기 지수와 배달량은 비슷한 양상을 보임

시간에 따른 체감 거리두기 지수와 배달량



시간에 따른 표준 및 체감 거리두기 지수

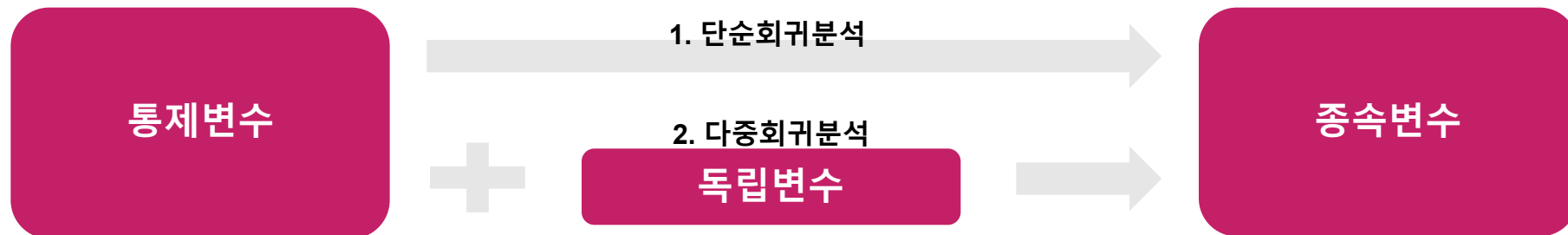


-
- 연구배경
 - 가설 설정 및 변수 정의
 - 데이터 가공
 - 분석 및 결과
 - 결론

통제변수 반영을 위한 위계적 회귀분석 진행

- 위계적으로 회귀분석을 실시하여 독립변수가 종속변수에 미치는 영향에서 통제변수의 영향을 보정한다.
- 1차 분석: 통제변수와 종속변수 간 단순회귀분석
 - H0: 일 평균 기온이 배달음식 주문량에 미치는 영향이 없다
 - H1: 일 평균 기온이 배달음식 주문량에 미치는 영향이 있다
- 2차 분석: 독립변수+통제변수와 종속변수 간 다중회귀분석
 - H0: 체감거리두기지수가 배달음식 주문량에 미치는 영향이 없다
 - H1: 체감거리두기지수가 배달음식 주문량에 미치는 영향이 있다

통제변수 투입에 따른 위계적 회귀분석 모델



회귀분석 결과 일 평균 기온은 배달 주문량에 유의미한 영향을 미침

- 배달 주문량 = $30337.63 - 333.16 * \text{평균기온}$
- 회귀분석 결과 일 평균 기온으로 나타난 날씨가 배달 주문량에 유의미한 영향을 미침
 - p-value = 0.000439, Advanced R-square = 0.1884

일평균 기온과 배달 주문량 회귀분석결과

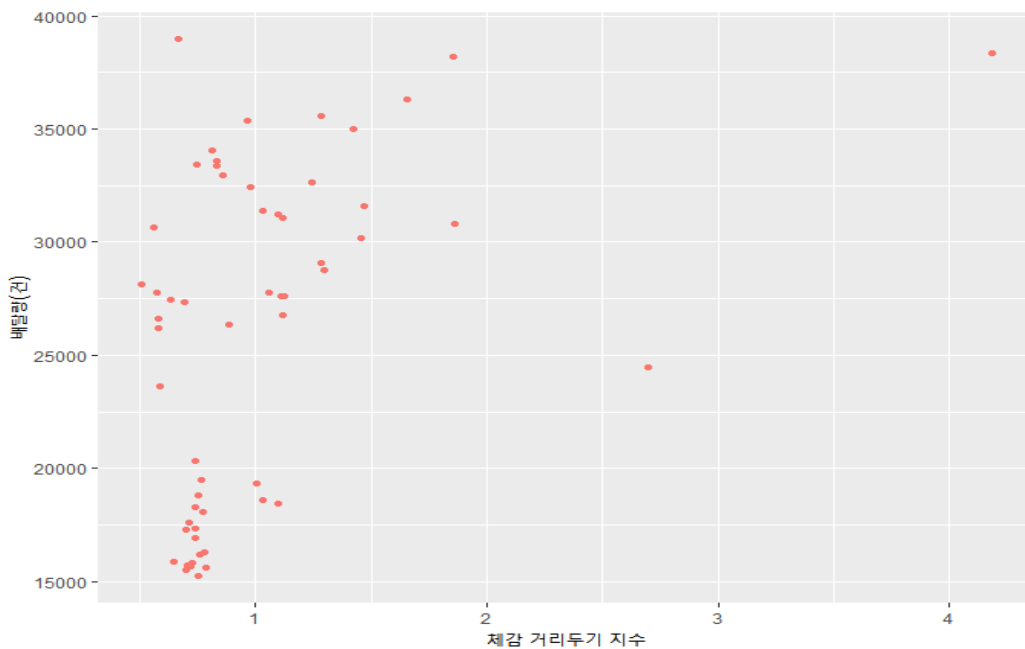
```
Coefficients:
              Estimate Std. Error t value      Pr(>|t|)
(Intercept) 30337.63    1483.59   20.449 < 0.0000000000000002 ***
avg.temp    -333.16      89.05    -3.741    0.000439 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6675 on 55 degrees of freedom
Multiple R-squared:  0.2028,    Adjusted R-squared:  0.1884
F-statistic: 14 on 1 and 55 DF,  p-value: 0.000439
```

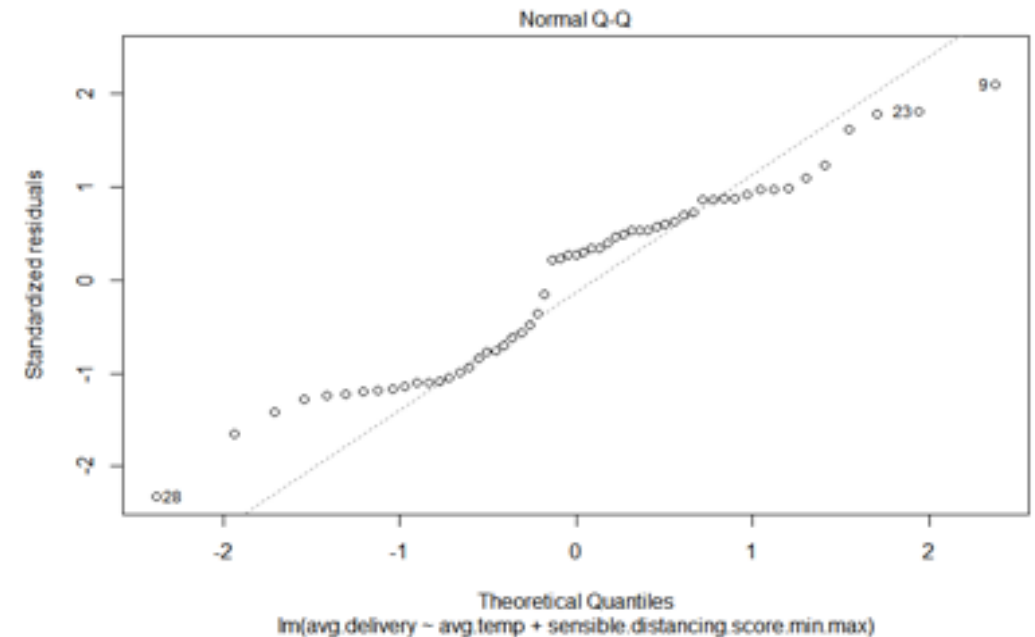
체감 거리두기 지수와 배달 주문량 사이에 약한 선형성이 존재함

- 사전검증 결과 체감거리두기 지수와 배달 주문량의 회귀분석이 가능
 - Correlation 0.42로 체감 거리두기 지수와 배달 주문량 간 선형성이 존재

체감거리두기 지수와 배달량의 선형성 검증



체감거리두기 지수와 배달량의 정규성 검증



체감 거리두기 지수 1단위 증가할 때 배달 주문량은 3679.83만큼 증가함

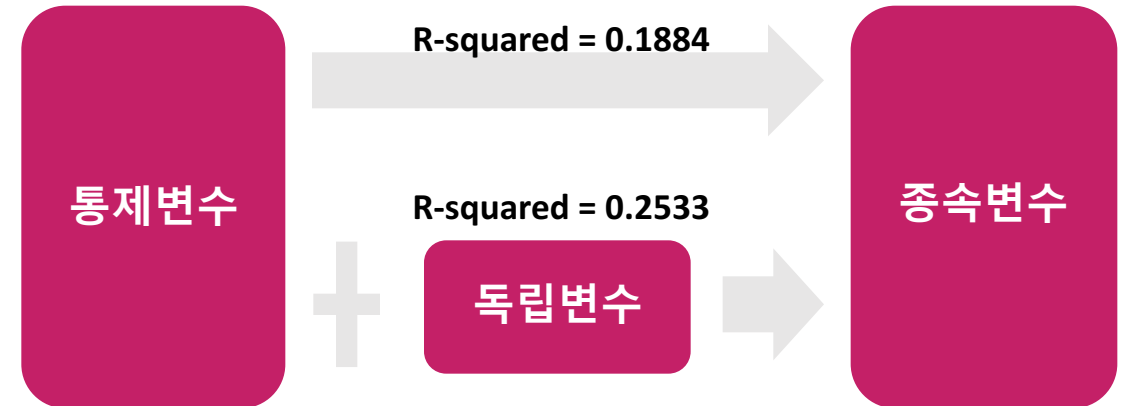
- 배달 주문량 = $25508.89 + 3679.83 * \text{체감 거리두기 지수} - 252.23 * \text{일평균기온}$
- 체감 거리두기 지수가 배달 주문량에 유의미한 영향을 미침 (p-value = 0.025)
 - 회귀계수에 대한 p-value값 0.025로 표본을 통해 얻은 체감 거리두기 지수의 회귀계수가 표본의 모집단에서도 통계적으로 유의미하다고 볼 수 있다.

체감거리두기 지수와 배달 주문량 회귀분석결과

```
Coefficients:
              Estimate Std. Error t value      Pr(>|t|)
(Intercept)    25508.89    2539.51  10.045 0.0000000000000584 ***
avg.temp       -252.23     92.71  -2.721  0.00875 **
sensible.distancing.score.min.max 3679.83    1599.86   2.300  0.02533 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6429 on 54 degrees of freedom
Multiple R-squared:  0.274,    Adjusted R-squared:  0.2471
F-statistic: 10.19 on 2 and 54 DF,    p-value: 0.0001761
```

통제변수 투입에 따른 위계적 회귀분석 모델에서 R-square 값의 변화



-
- 연구배경
 - 가설 설정 및 변수 정의
 - 데이터 가공
 - 분석 및 결과
 - 결론

체감 거리두기 지수는 배달량에 유의미한 영향을 미친다

▪ 의의

- 분석 결과 정부에서 거리두기 정책을 강화할 때마다 **사람들의 이동량에 변화가 있었으며**, 이를 반영한 **체감 거리두기 지수는 배달 주문량에 유의미한 영향을 미치는 것으로** 나타남.
- 거리두기가 배달 주문량에 미치는 영향이 유의 하다면, 정부가 이를 고려하여 강도를 조정함으로써 자영업자, 특히 **소상공인의 피해를 최소화**하는 데 활용할 수 있음.
- 거리두기 정책을 단순 지표화하지 않고, 이동량을 반영하여 **체감 거리두기 지수**를 사용했다는 데 의의가 있음.

▪ 한계점 및 향후 연구방향

- 기온 외 음식 배달 주문에 **영향을 미치는 또 다른 요인**을 통제변수로 추가한다면 더 정확한 예측이 가능할 것으로 보임.
- **표준 거리두기 지수와 체감 거리두기 지수 간 차이**에 대해 추가 연구도 가능.
시간의 흐름에 따라 체감 거리두기 지수가 더 큰 폭으로 감소하는데, 이를 방역 피로도 혹은 방역정책 효과 시간에 따른 변경에 대한 연구로 확장 가능

감 사 합 니 다