**환경적 요인과 매출액의 인과관계**

**Part A Introduction**  
 2019년 9월 1일부터 2022년 2월 1일까지 총 880일의 진영 Subway 지점 매출 상세와 북 창원 관측소의 날씨 정보를 기반으로 환경적인 요인이 Subway 진영점의 매출액에 어떠한 영향을 주는지 분석한 리포트이다.

**Part B Methodology**  
 데이터는 Python 프로그램을 통해 1차적으로 가공이 되었고 그 중 필요한 정보인 일시, 기온, 강수량, 풍속, 습도, 적설량, 지면온도, 대기온도, 요일, 총 매출, 매장 매출, 배달 매출로 이루어져 있다. 첫번째로 종속 변수의 값인 총 매출의 데이터를 히스토그램을 이용하여 데이터의 분포도를 확인한다. 이후 선형모형, 일반화선형모형을 이용하여 독립 변수들의 계수를 계산하고 Diagnostic Test를 통해 모형의 적합성을 판단한다.

**Part C Results**  
 3가지 다른 방식의 분석을 실행하였다. 1번은 총매출과 환경적인 요인의 인과관계, 2번은 주말을 제외한 평일 총매출과 환경적인 요인의 인과관계, 3번은 배달 매출 비율과 환경적인 요인의 인과관계이다. 모든 Model 모두 회귀모형의 가정을 위반하는 특이점은 찾지 못했다.

총매출과 환경적인 요인의 인과관계에서는 총매출의 기준점을 금요일 그리고 모든 환경적 요인이 0일 때 매출이 1,131,000원이고 온도가 1도 오를 때마다 +25,000원, 습도가 1퍼센트 오를 때마다 -2,400원, 대기 중 온도가 1도 오를 때마다 -14,800원 그리고 토, 일 매출이 22~26만 원 더 높았고 화, 수 매출은 6~7만 원 정도 낮았다.

주말을 제외한 평일 총매출에서도 비슷한 양상을 띄었는데, 총매출의 기준점을 금요일 그리고 모든 환경적 요인이 0일 때 매출은 1,049,000이고 온도가 1도 오를 때마다 +28,500원 바람이 1 m/s 더 세게 불면 +35,000원 그리고 대기 중 온도가 1도 오를 때마다 -19,000원 그리고 화, 수 매출이 약 6~7만 원 정도 낮았다.

배달 매출 비율과 환경적 요인에서는 종속 변수에 통계적으로 유의미하게 영향을 주는 독립 변수를 찾을 수 없었다. 날이 거듭할수록 배달 매출 비율이 높아지는 것이 산포도를 통해 확인되었으나 이 역시 통계적으로 유의미한 수치는 아니었다.

**Part D Discussion**  
 결과를 통해서 우리는 대체적으로 기온이 높을 때 사람들이 서브웨이를 즐겨 먹는다는 것을 알 수 있었다. 평균적으로 화, 수요일의 매출이 유의미하게 적었고 이를 통해서 1인 생산성이 낮은 화, 수요일의 직원들의 근무시간을 줄이는 것이 이익을 높이는 것에 도움이 될 것이라고 생각한다. 예상과 달리 비나 눈의 유무는 배달 매출에 유의미한 영향을 주지 않았다. 모순적이게도 대기 중 온도의 상승은 매출 감소에 영향을 주었는데 이는 추가적인 분석이 필요하다고 생각된다.

\* 처음 가설은 ‘여름은 성수기 그리고 월요일은 평일 중 가장 바쁘다’ 였고 80% 정도 맞는 추측이었다고 생각한다.  
\* 점장님에게 화, 수 알바 쓰는 시간을 줄여보라고 권유했으나 그건 어렵다는 답변을 받았다.  
\* 분석에 활용된 데이터 그리고 프로그래밍 언어인 R과 Python의 Appendix(부록)은 <https://github.com/parky0113/Side-Project-Toys-/tree/main/Subway%20Research> 에서 찾아볼 수 있다.