

# #3-1 대회 소개

## 데이콘 Basic 쇼핑몰 지점별 매출액 분석시각화 경진대회

### 목적

쇼핑몰 매장별 프로모션 정보, 주변 날씨, 실업률, 연료 가격 등의 정보를 바탕으로  
쇼핑몰의 각 **지점별 매출액 분석 및 시각화**

**데이터** - train.csv, test.csv

- id : 샘플 아이디
- Store : 쇼핑몰 지점
- Date : 주 단위(Weekly) 날짜
- Temperature : 해당 쇼핑몰 주변 기온
- Fuel\_Price : 해당 쇼핑몰 주변 연료 가격
- Promotion 1~5 : 해당 쇼핑몰의 비식별화된 프로모션 정보
- Unemployment : 해당 쇼핑몰 지역의 실업률
- IsHoliday : 해당 기간의 공휴일 포함 여부
- Weekly\_Sales : 주간 매출액 (목표 예측값)

# #3-2 접근 방식 및 인사이트

## 데이콘 Basic 쇼핑몰 지점별 매출액 분석시각화 경진대회

### 접근 방식

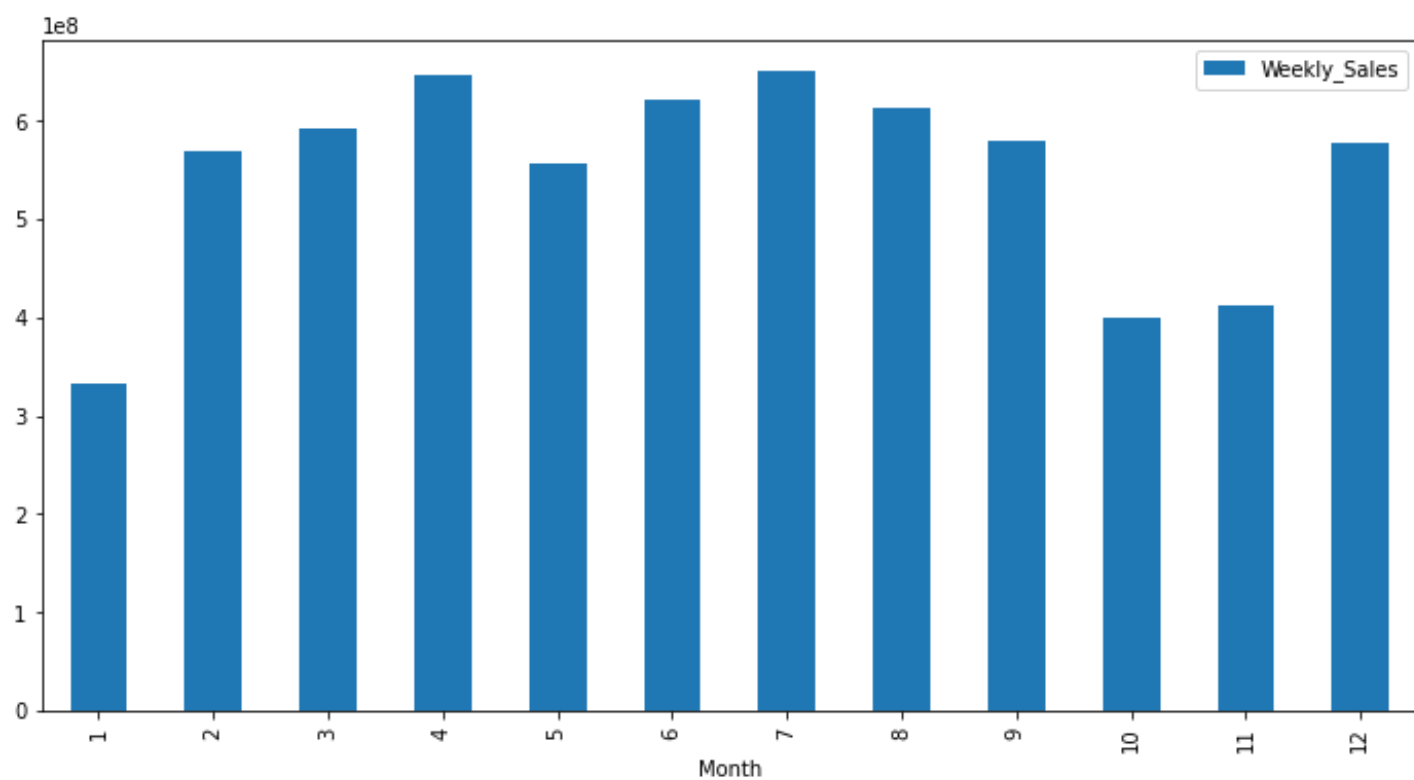
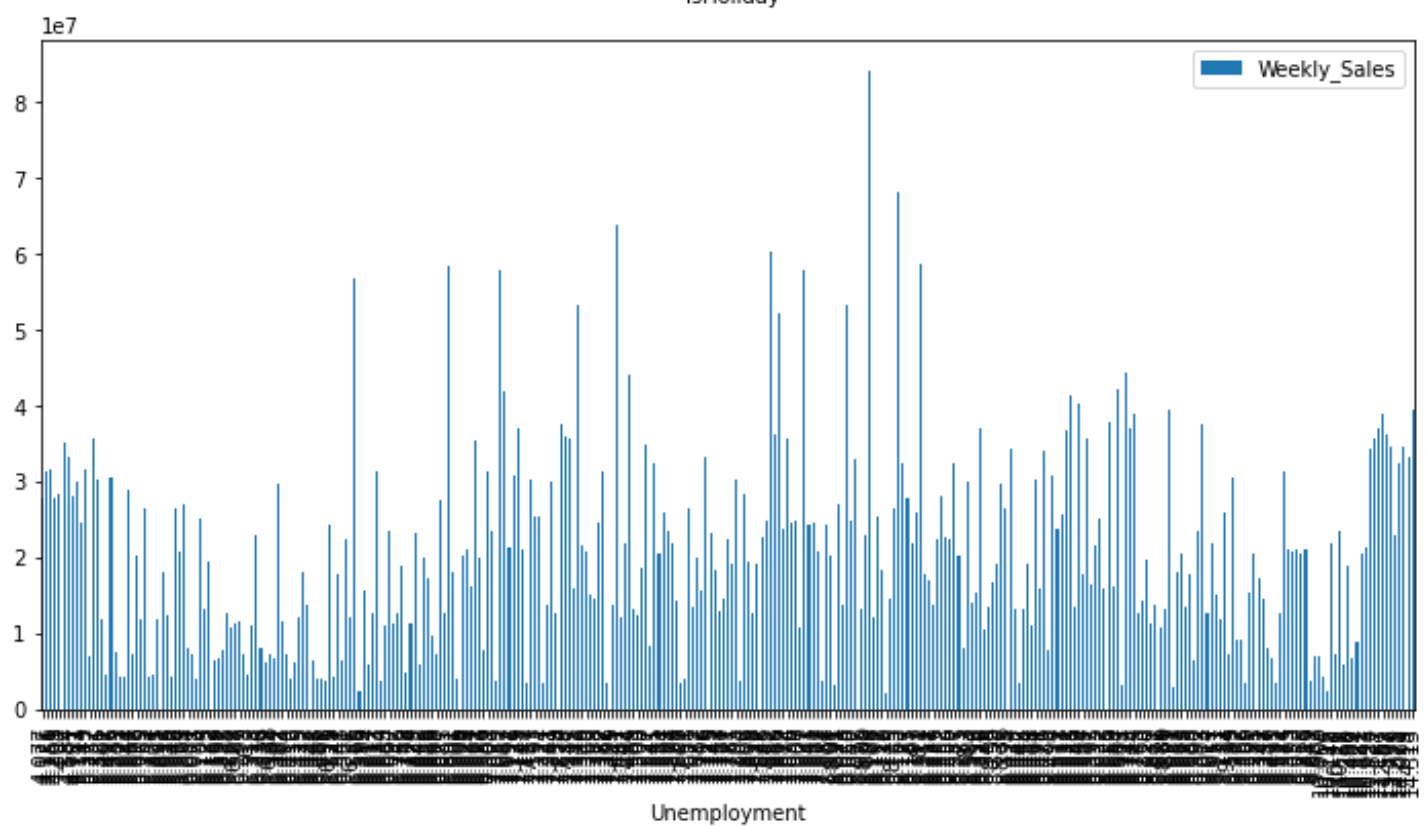
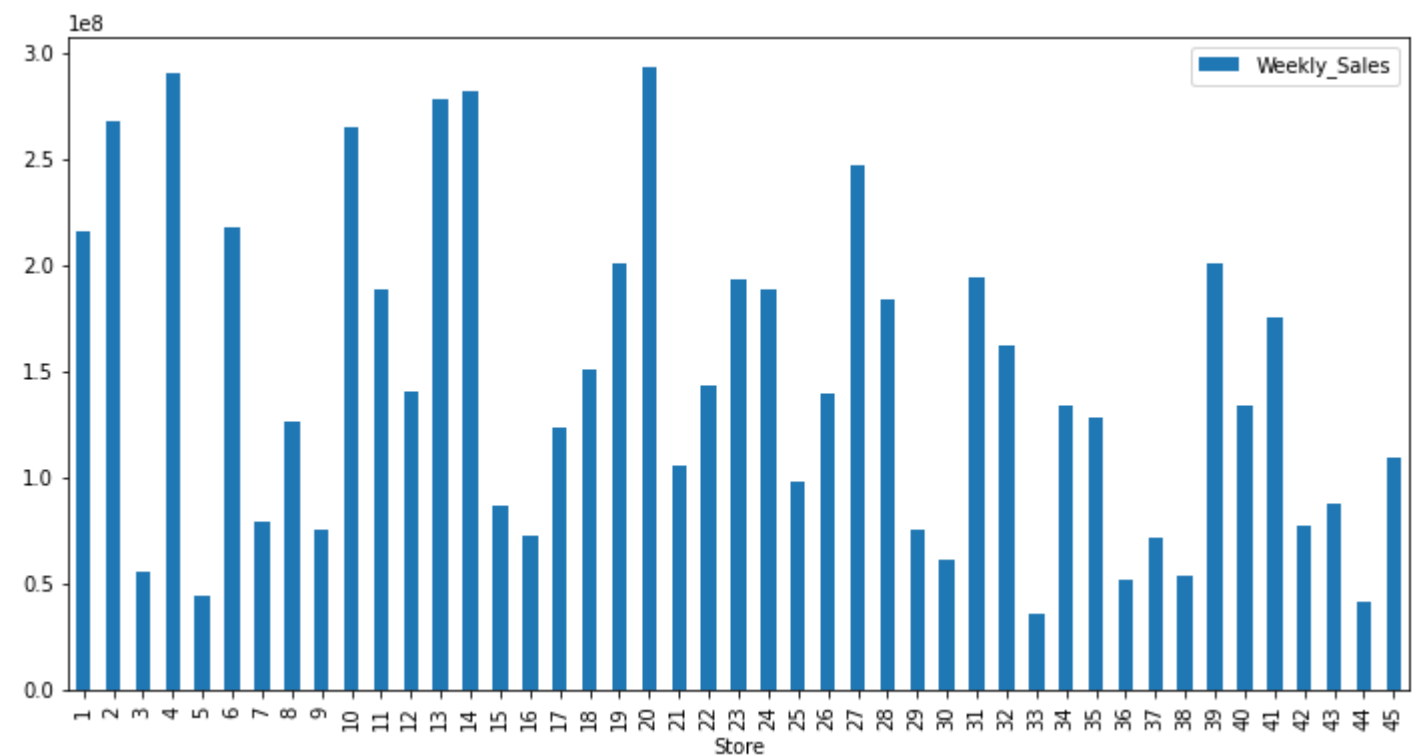
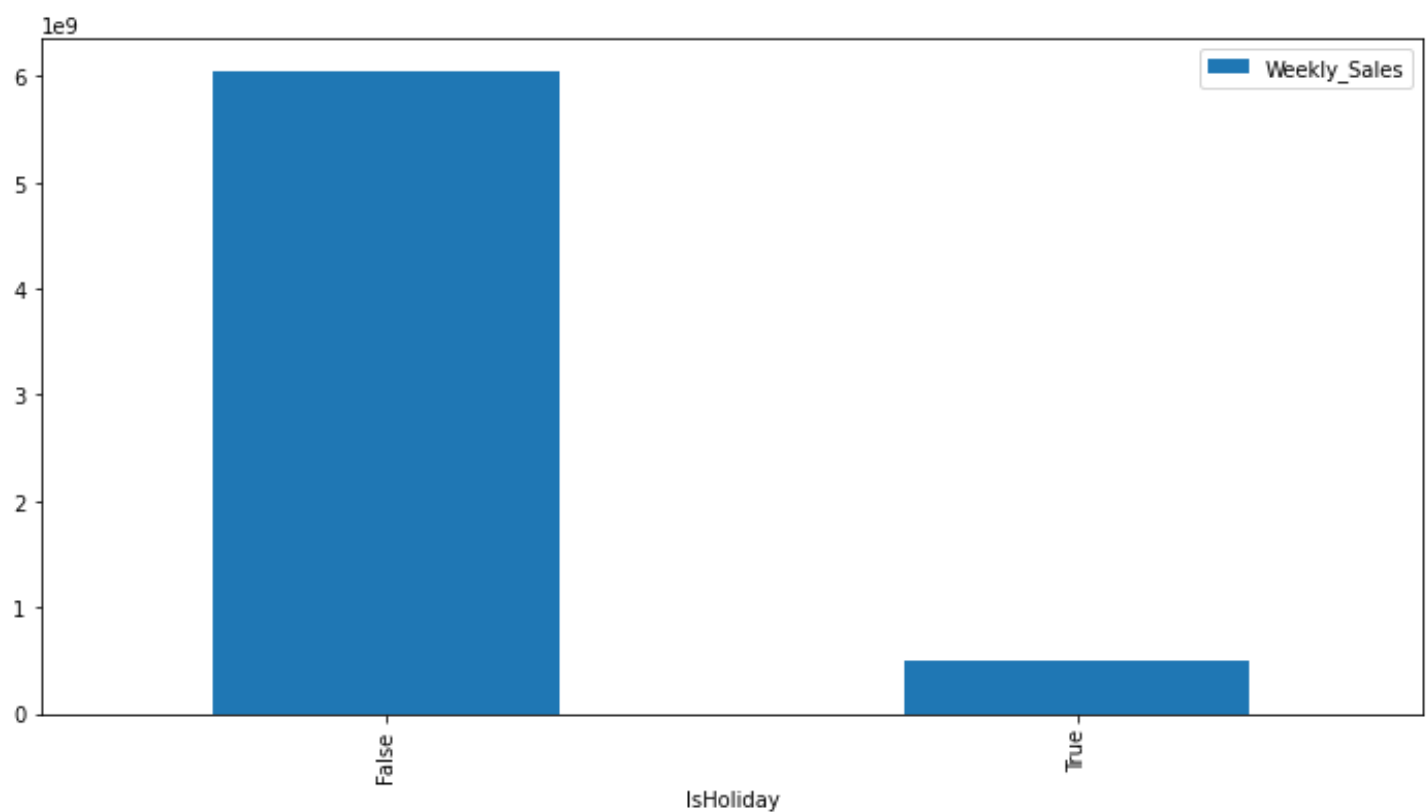
제공된 데이터를 기반으로 다양한 EDA 진행, 여러 feature간의 관계를 파악하고자 함

### 인사이트

- feature별로 Weekly Sales 관찰
- Heatmap 등의 시각화를 통해 feature간의 연관관계 추론
- AutoEDA를 통해 다양한 데이터 분포, 관계를 효율적으로 파악)

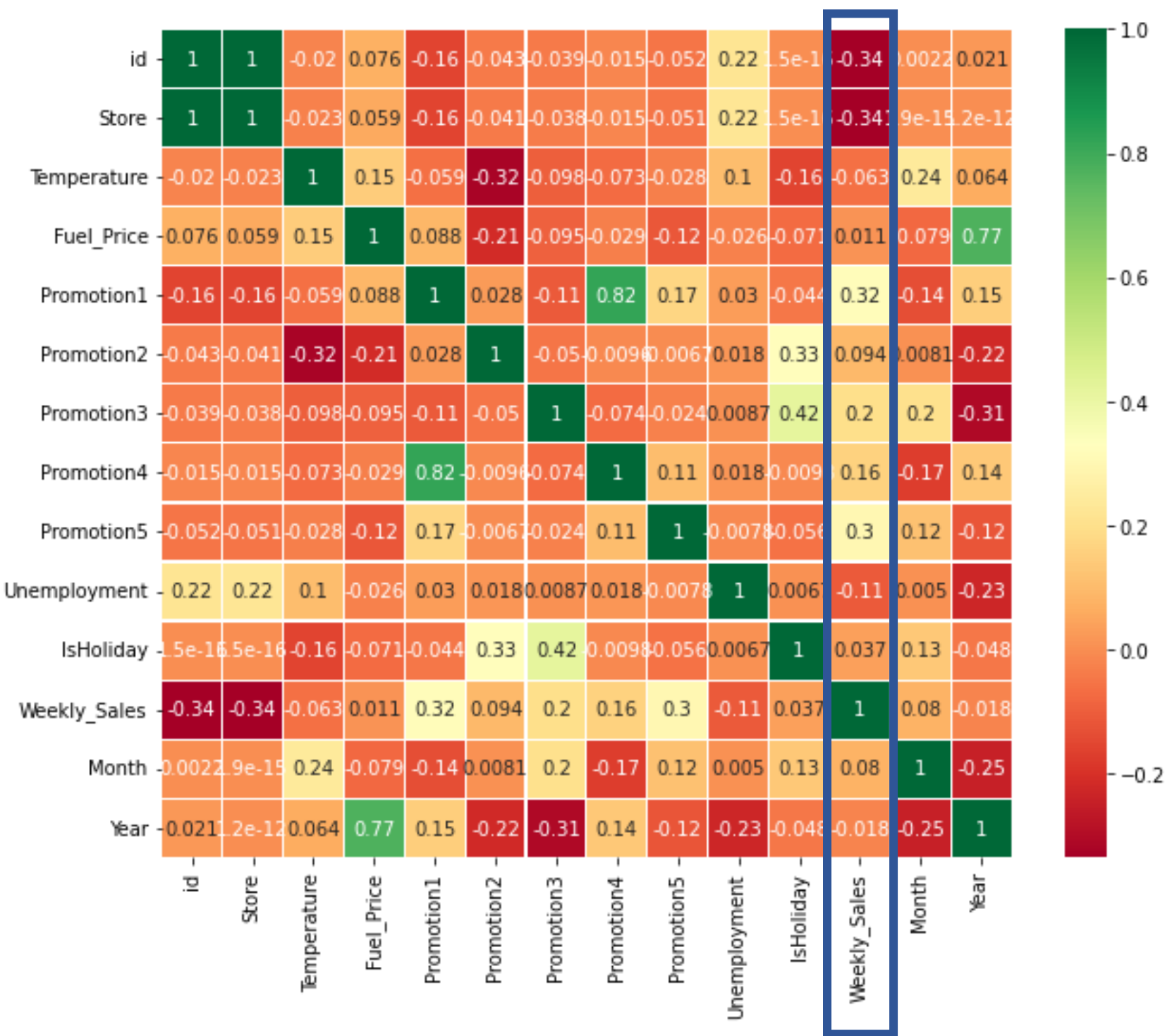
# #3-2 접근 방식 및 인사이트

- feature별로 Weekly Sales 관찰

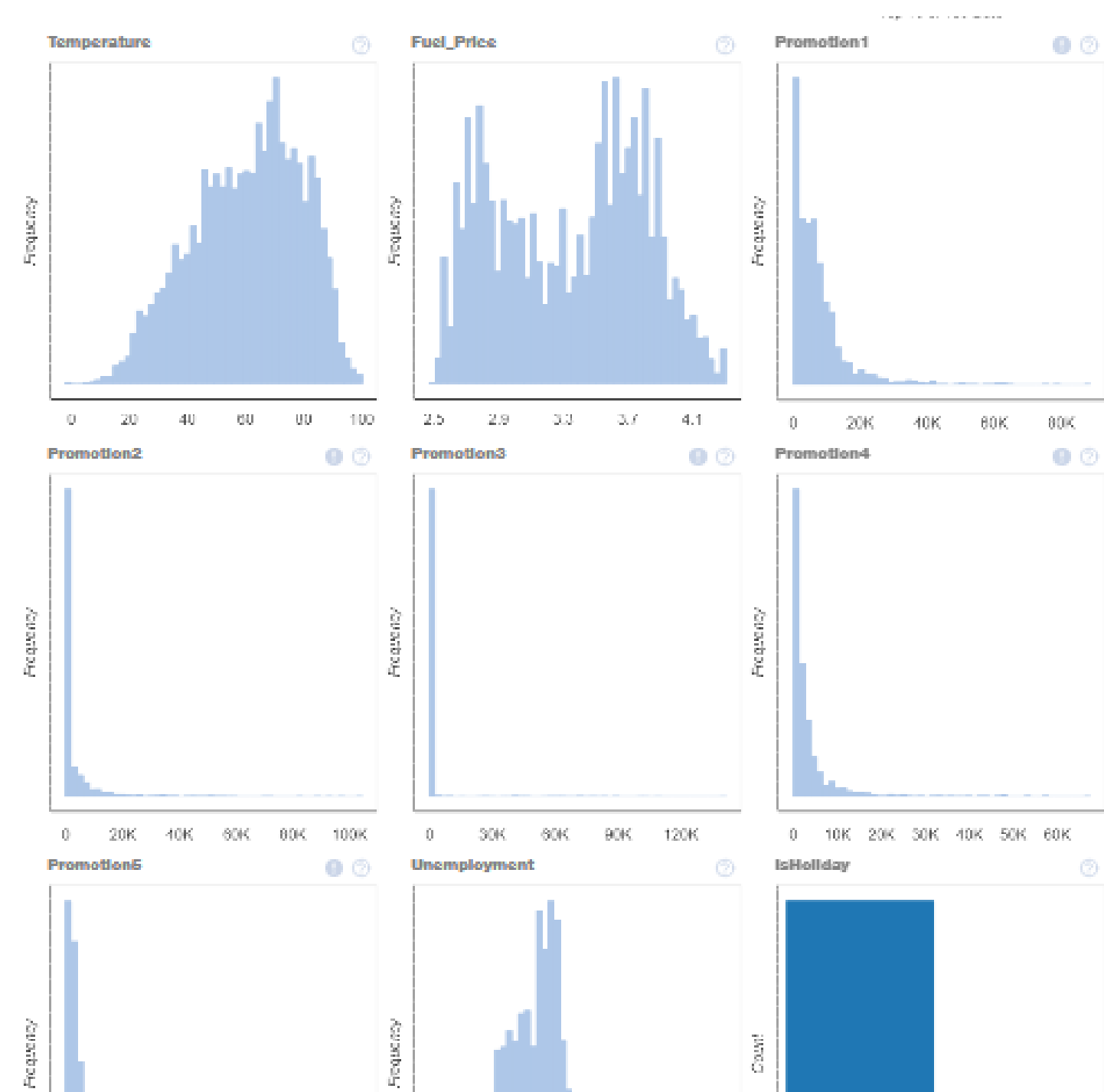


# #3-2 접근 방식 및 인사이트

- Heatmap 등의 시각화를 통해 feature간의 연관관계 추론



- AutoEDA를 통해 다양한 데이터 분포, 관계를 효율적으로 파악)

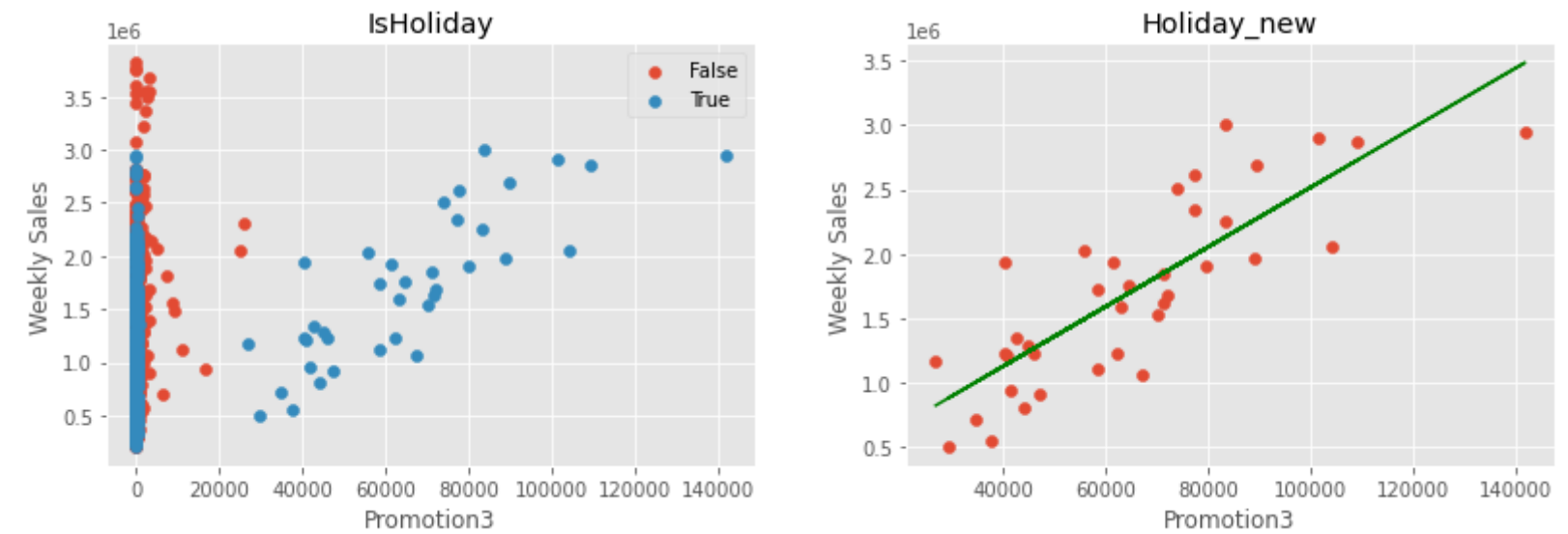


# #3-3 1등 코드 분석

## 1) IsHoliday와 Promotion3, Weekly Sales의 관계

IsHoliday가 True일 때 Promotion3의 값이 20000 이상이면, Weekly\_Sales가 우상향으로 증가

20000 이상인 값들만 추출 비례하는지 확인하기 위해 선형회귀를 사용하여 예측용 직선 그림



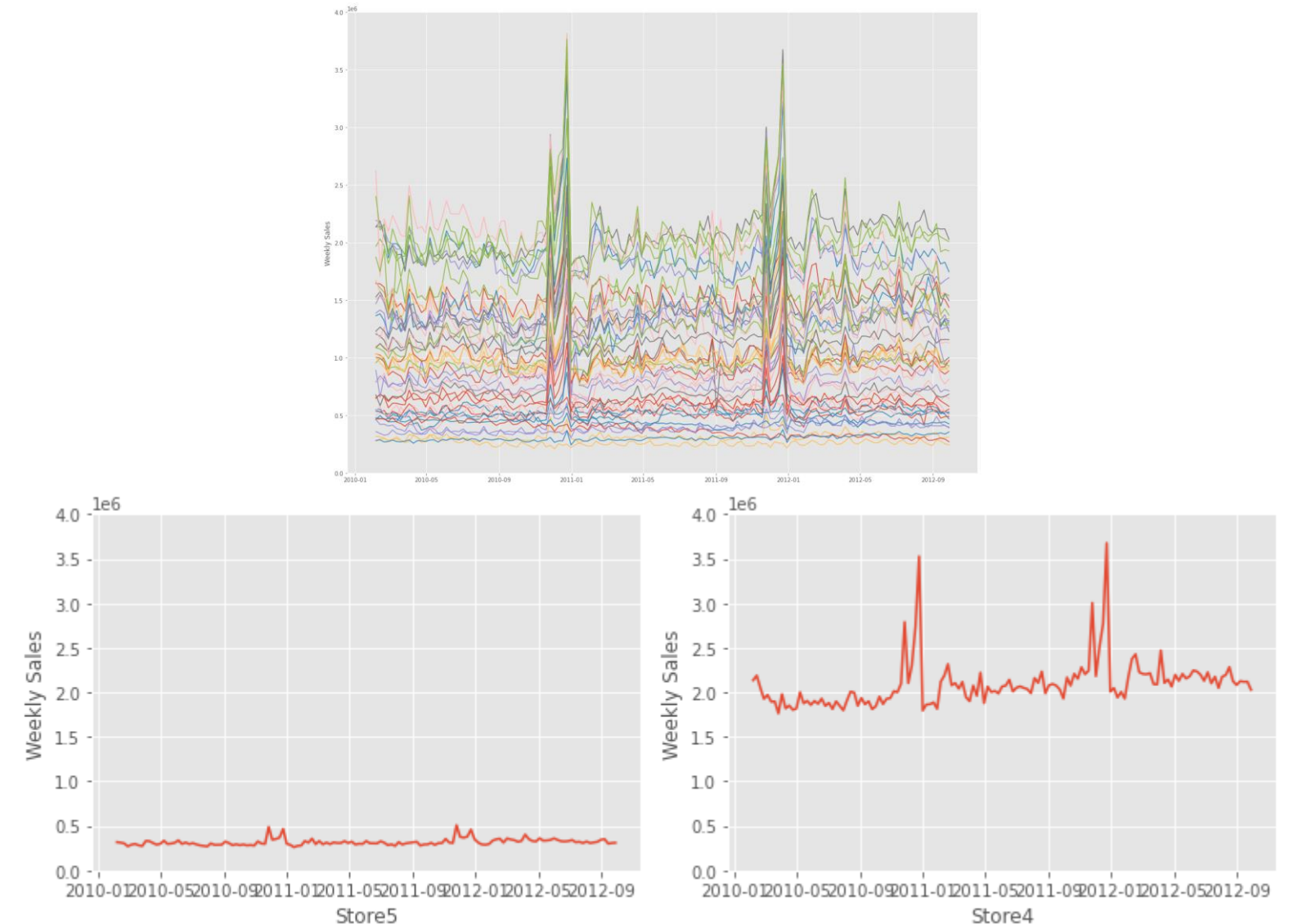
## 2) 1년 중 언제가 매출이 가장 큰 시기

위 그래프들을 분석해보면 크게 두 가지 유형으로 나눌 수 있음

- 1) 전체적인 매출이 적고 매출의 변화가 평이한 매장 ex)3, 5, 33
- 2) 전체적인 매출이 크고 특정 구간에 매출이 치솟는 매장 ex)1, 4, 14

→ 매출이 작은 지점은 매출변화율이 매우 작음

→ 매출이 큰 지점은 연말, 특히 크리스마스에 매출이 폭발적으로 증가



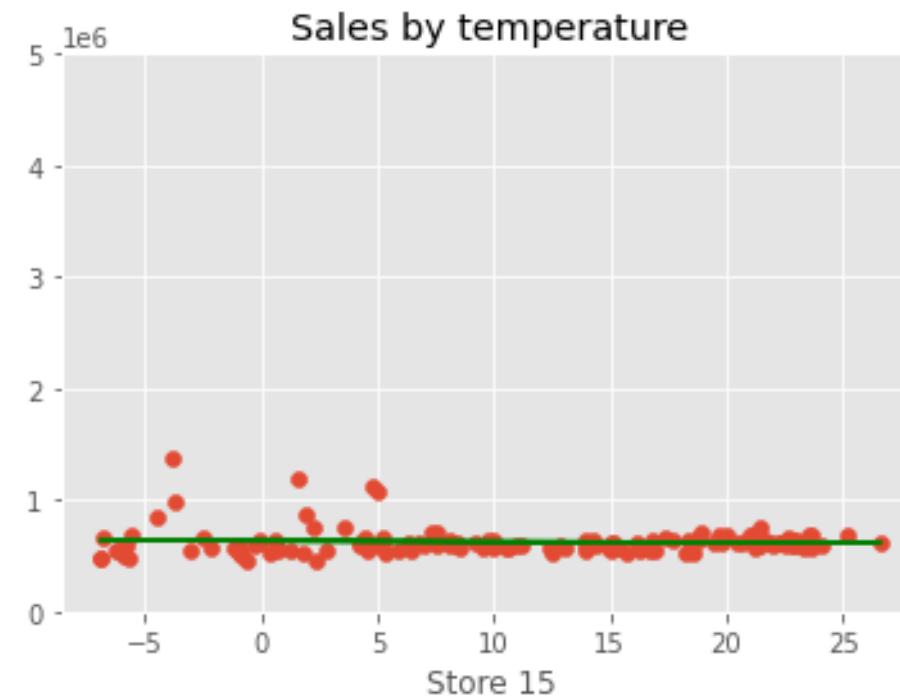
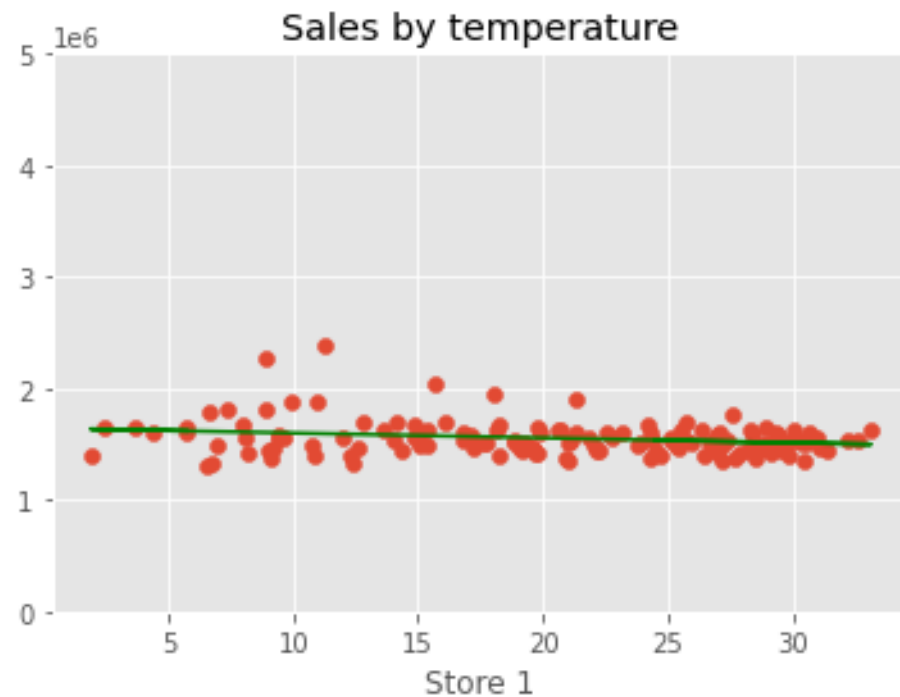
# #3-3 1등 코드 분석

## 3) Promotion과 매출의 관계

- Promotion1의 경우 **매출에 주는 영향이 없다.**
- Promotion2의 경우 **12월 30일에 가장 컸지만 오히려 크리스마스가 지나 매출이 급감하였다.**
- Promotion3의 경우 앞선 분석과 같이 **IsHoliday가 True이고 Promotion3가 큰 값이면 매출이 값에 비례하며 크게 나왔다.**
- Promotion4의 경우 **Promotion1과 행사일이 대체로 같다.**
- Promotion5의 경우 **매출에 주는 영향이 없다.**

## 4) 기온과 매출의 관계

기온과 매출은 관련이 없음



# THANK YOU

