

# 성충권 오존총량과 홍반 자외선 세기의 상관관계

## <소주제> 오존의 역할에 대한 고찰

|       |         |           |                  |
|-------|---------|-----------|------------------|
| 과 목 명 | 컴퓨팅 기초  | 팀 명       | 화씨온도 팀           |
| 지도교수  | 변해선 교수님 | 학 과 / 이 름 | 화학부 박승연, 화학부 최지우 |



### 연구배경 및 필요성

#### 1. 연구배경

**자외선**은 일반적으로 자외선A, 자외선B, 자외선C로 나뉘며, 이 중 자외선A와 자외선B는 오존층에 일부가 흡수되고 그 나머지가 지표에 도달한다. 지표에 도달하는 자외선은 적은 양이지만 인간과 동·식물에게 큰 피해를 줄 뿐만 아니라 광화학 반응에도 영향을 미치면서 대기 환경을 변화시킨다. 기상청은 자외선 복사 중 자외선A(320~400nm)와 자외선B 영역 중 인체에 홍반을 발생시키는 홍반자외선B(280~320 nm)를 관측하고 있다.

성충권에는 지구 대기에 존재하는 오존의 약 90%가 존재하며, 대체로 15~30km 사이에 오존 농도가 높게 존재하는 층을 **오존층**이라고 부른다. 오존층은 태양으로부터 유해 자외선 복사를 흡수하여 자외선으로부터 인간과 생태계가 유지될 수 있도록 보호하고 지구의 기후 조절에 중요한 역할을 하고 있다. 기상청은 세계기상기구(WMO) 지구대기감시(GAW) 프로그램의 일환으로 안면도 기후변화감시소에서 성충권 오존전량을 관측하고 있다.

#### 2. 필요성

최근, 산업 활동과 환경 변화로 인해 오존층에 변화가 생겼으며, 이는 지구 표면의 자외선 노출 증가로 이어질 수 있다. 지표면의 유해 자외선 증가는 인간을 포함한 동식물에게 큰 피해를 야기할 수 있기 때문에, 이 프로젝트는 오존층의 변화가 자외선 강도에 어떤 영향을 미치는지를 분석하는 것을 목표로 한다.

#### 3. 목표

본 프로젝트의 주요 목표는 오존층의 두께 및 구성 변화가 대한민국 표면에서의 자외선 강도와 어떠한 상관관계가 있는지 분석하는 것이다.



### 프로젝트 진행과정

#### 1. 데이터 수집

기상자료개방포털 <https://data.kma.go.kr/cmmn/main.do> -> 데이터 -> 기후 -> 성충권 오존전량과 자외선 세기 데이터 수집  
2014, 2015, 2016 3개년에 대한 데이터를 일 단위로 수집했다. 지역은 안면도(감)으로 통일하였다.

#### 2. 전체코드

```
ozone_data = pd.read_csv('/content/ENV_OZONE_2014_2015_2016_merge.csv',
encoding='unicode_escape', header=0, usecols = [1, 2])
ozone_data.columns = ['date', 'ozone_thickness']
uv_data = pd.read_csv('/content/ENV_UV_2014_2015_2016_merge.csv',
encoding='unicode_escape', header=0, usecols = [1, 3])
uv_data.columns = ['date', 'uv_intensity']
data = pd.merge(ozone_data, uv_data, on='date')
data = pd.DataFrame(data)
data.dropna(inplace=True)

fig, ax1 = plt.subplots()
color_1 = 'tab:blue'
ax1.set_title('Plot with 2 Axes for a dataset with different scales', fontsize=16)
ax1.set_xlabel('Date')
ax1.set_ylabel('Ozone thickness (blue)', fontsize=14, color=color_1)
ax1.plot(data['date'], data['ozone_thickness'], color=color_1)
ax1.tick_params(axis='y', labelcolor=color_1)
ax2 = ax1.twinx()
color_2 = 'tab:red'
ax2.set_ylabel('UV Intensity (red)', fontsize=14, color=color_2)
ax2.plot(data['date'], data['uv_intensity'], color=color_2)
ax2.tick_params(axis='y', labelcolor=color_2)
fig.tight_layout()
plt.show()

linear_regression = linear_model.LinearRegression()
linear_regression.fit(X = pd.DataFrame(data['ozone_thickness']), y =
data['uv_intensity'])
print('a value = ', linear_regression.intercept_)
print('b value = ', linear_regression.coef_)
X = data[['ozone_thickness']] # 독립변수
y = data['uv_intensity'] # 종속변수
X = sm.add_constant(X) # 상수항 추가
model = sm.OLS(y, X).fit()

correlation = data['ozone_thickness'].corr(data['uv_intensity'])
print("상관계수: ", correlation)
print("결정계수 R-squared: ", model.rsquared)
```

#1  
데이터 불러오기,  
전처리, 합치기

#3  
단순회귀분석

#4  
상관계수  
결정계수

```
plt.scatter(model.predict(X), model.resid)
plt.xlabel('Predicted Values')
plt.ylabel('Residuals')
plt.title('Residuals vs Predicted')
plt.axhline(y=0, color='r', linestyle='--')
plt.show()

fig, ax = plt.subplots(figsize=(6, 4))
stats.probplot(model.resid, dist="norm", plot=ax)
plt.title("Q-Q Plot of Residuals")
plt.show()

plt.figure(figsize=(10, 6))
plt.scatter(data['ozone_thickness'], data['uv_intensity'], alpha=0.5,
label='Actual Data')
prediction = linear_regression.predict(X = pd.DataFrame(data['ozone_thickness']))
plt.plot(data['ozone_thickness'], prediction, color='red', label='Regression
Line')
plt.xlabel('Ozone Thickness')
plt.ylabel('UV Intensity')
plt.title('Ozone Thickness vs UV Intensity with Regression Line')
plt.legend()
plt.show()
```

#5  
잔차분석

#6  
산점도  
회귀선추가

#7  
결과시각화

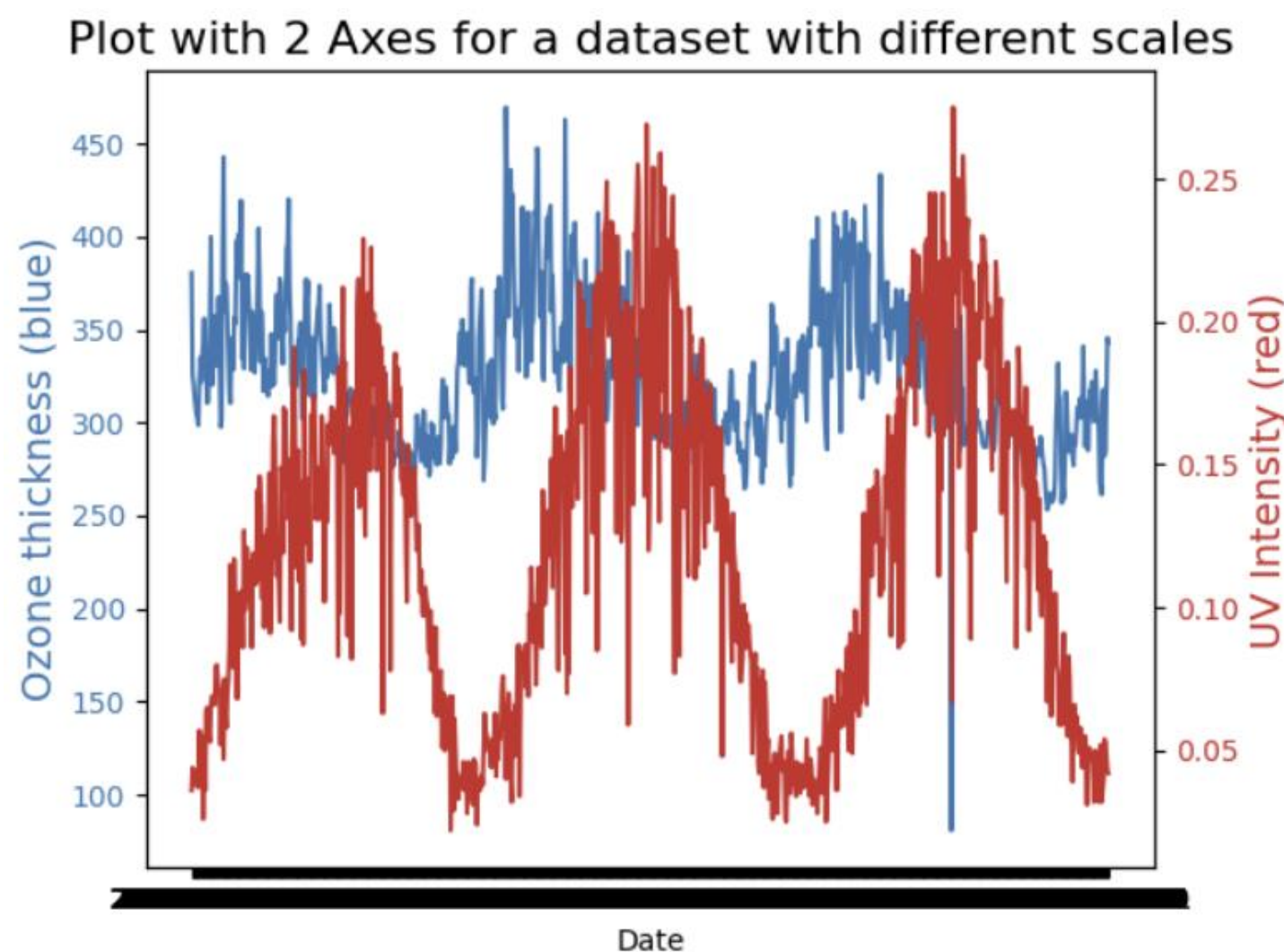


### 프로젝트 성과

#### 1. 시계열 분석

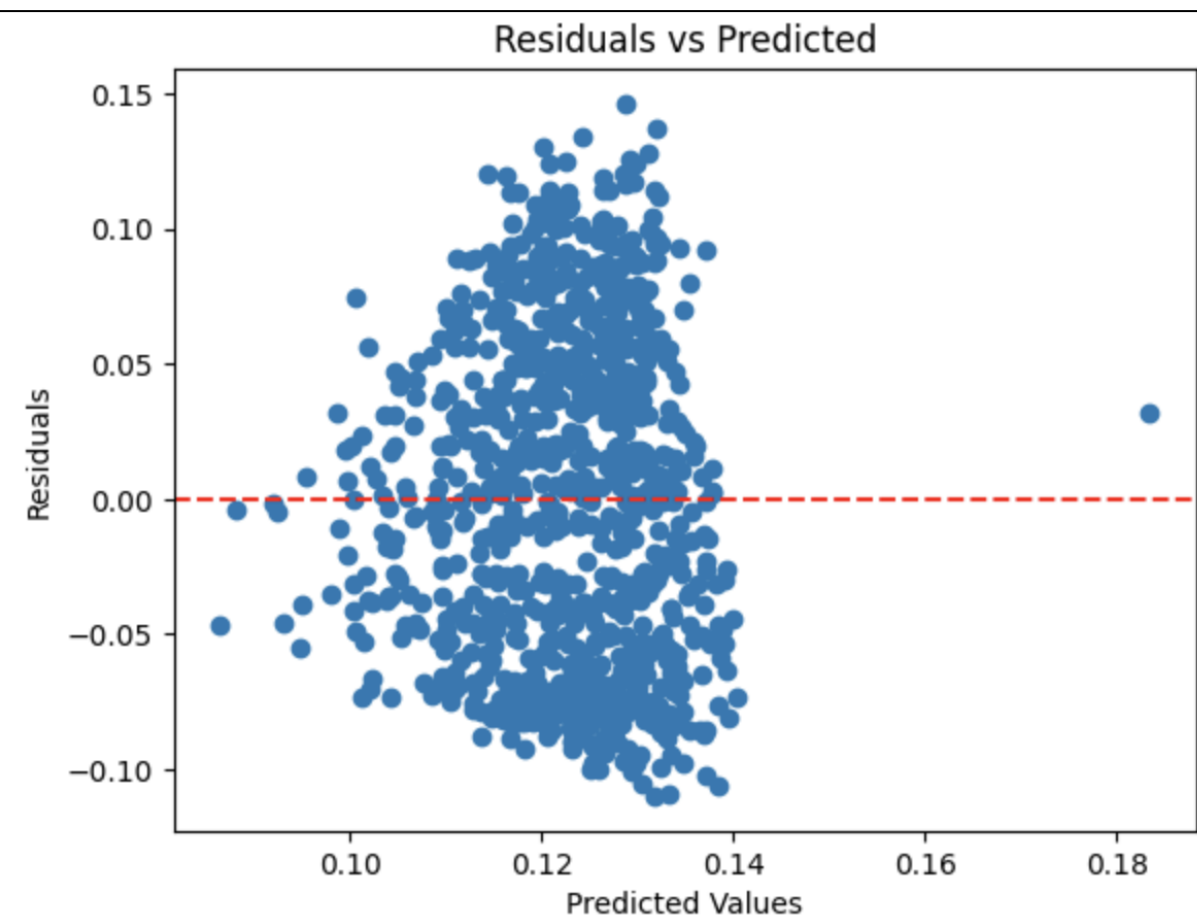
- Date 라는 공통 열을 중심으로 오존총량과 자외선 세기를 y축에 나타냈다. 이때 두 데이터의 숫자 scale이 달라 axis를 추가하는 코드를 작성했다.  
(x축: Date, y축: Ozone thickness, UV intensity)

- 분석 결과  
3개년을 분석한 결과가 다음과 같은 양상을 보이는 것으로 보아 1년을 주기로 오존총량과 자외선 세기가 주기를 가진다고 분석했다. 대체로 증가 추세가 유사하나 감소 추세는 자외선 세기가 오존총량보다 느렸다.



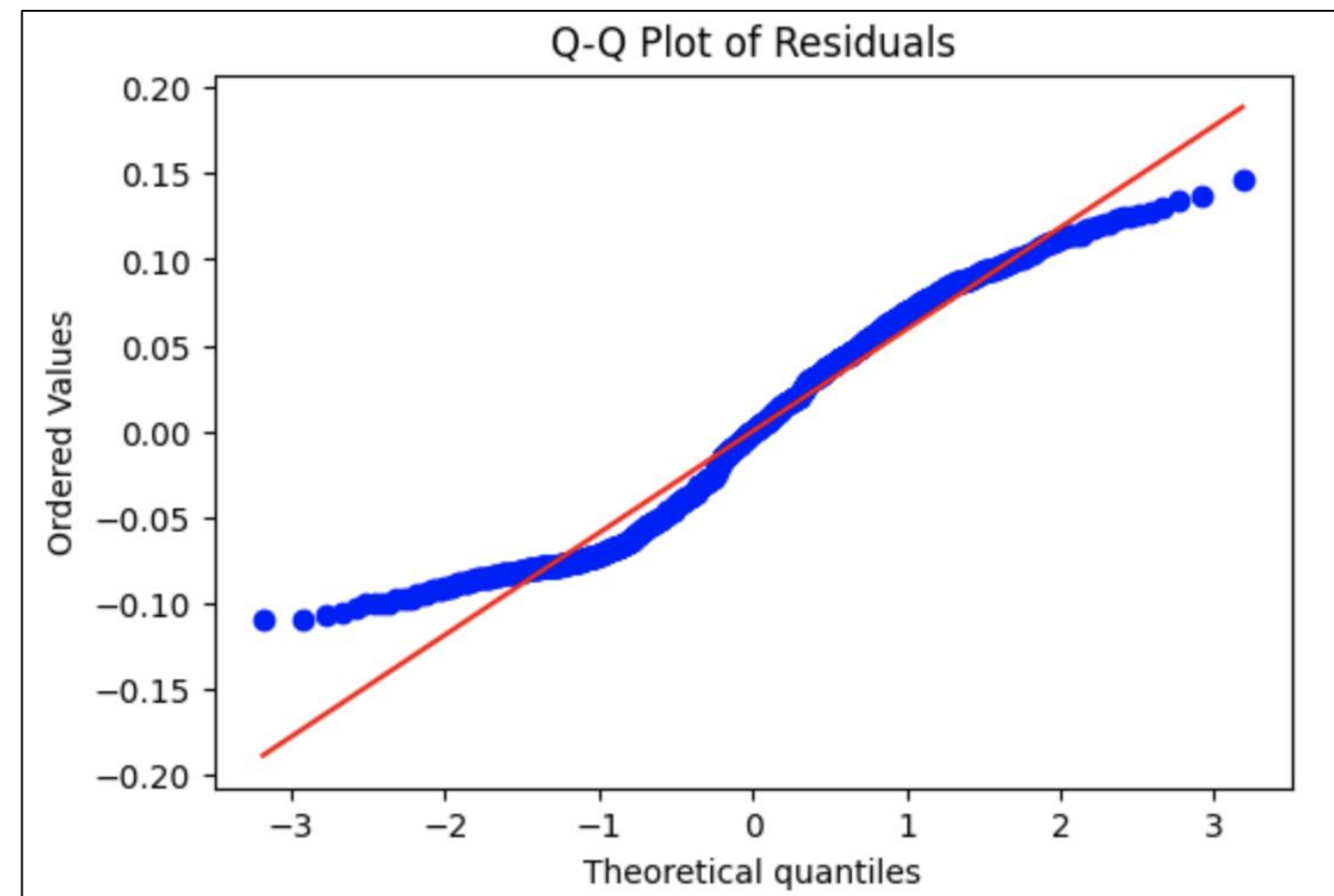
#### 2. 잔차 분석

잔차도



잔차란 실제값과 예측값의 차이를 뜻하며 이를 산점도로 나타낸 것이 잔차도인 것이다.  
1. 잔차가 무작위하게 분포되어 있는 것을 확인할 수 있다.  
2. 1.을 통해 잔차들이 독립이라는 사실을 알 수 있다.  
3. 0을 중심으로 약한 평행한 띠를 형성하여 등분산성을 검토했다.

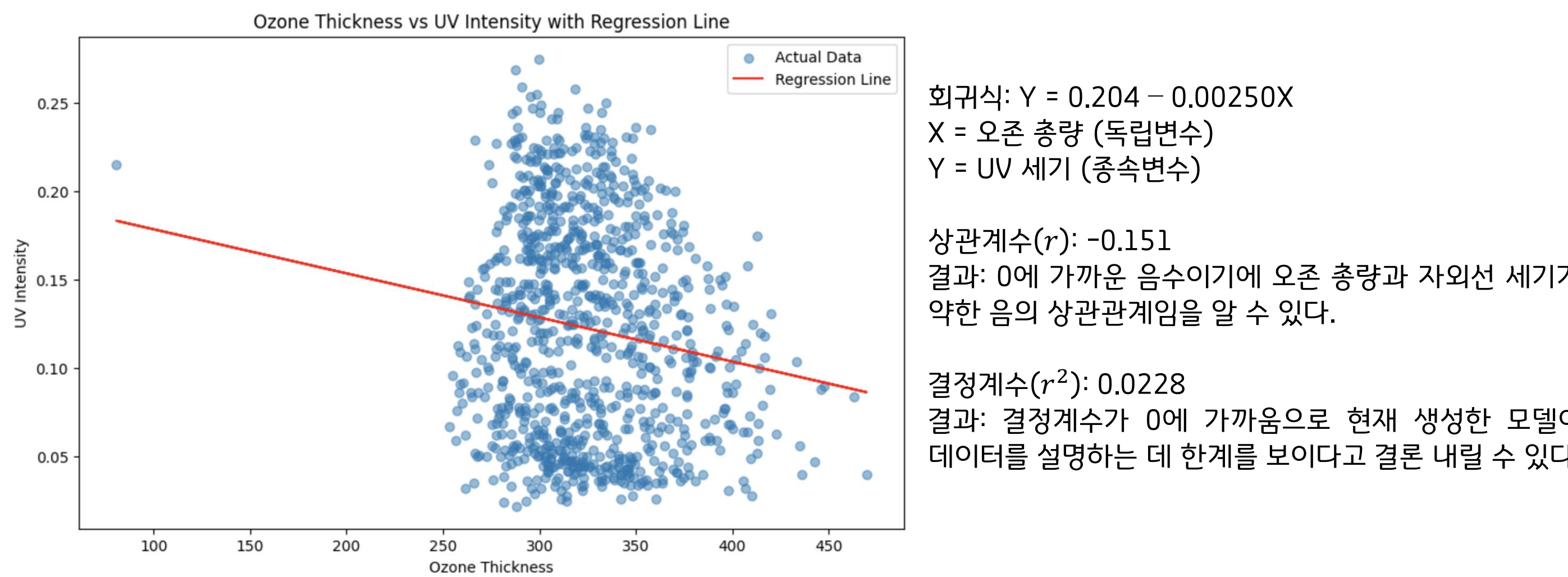
Q-Q 플롯



Q-Q플롯이란 "Quantile"- "Quantile" 플롯의 줄인 말으로 두 확률 분포의 분위수를 비교하는 그래프이다. 이를 통해 잔차가 정규 분포를 따르는지 확인 할 수 있다.  
1. 잔차가 대각선(45도) 근처에 대략적으로 위치하고 있다는 것을 볼 수 있다.  
2. 1.을 통해 잔차는 정규 분포를 따른다고 결론 지을 수 있다.

∴ 단순회귀분석을 위한 잔차의 검토를 마무리하였으며 결과적으로 잔차에 대한 가정에 적절히 들어 맞는다.

#### 3. 데이터의 산점도와 회귀선의 시각화



### 프로젝트 기대효과

#### 1. 오존 총량과 UV세기의 상관관계 분석

선형회귀선을 그려 오존 총량과 UV 세기의 대략적인 관계를 확인 할 수 있었다. 강한 음의 관계는 아니지만 오존 총량이 늘어나면 UV 세기가 줄어드는 추세를 살펴볼 수 있었다. 그렇기에 추가적인 연구를 한다면 시계열 데이터를 참고해 오존 총량의 변화가 UV 세기 반영되기까지 어느 정도의 시간차를 가지고 있음을 고려해 미래의 연구를 진행할 수 있을 것이다.

#### 2. 측정되는 UV 세기에 영향을 줄 수 있는 다양한 요인 제언

이번 연구는 단일회귀분석을 활용해 하나의 독립변수를 이용해 종속변수를 설명하려고 했다. 하지만 UV 세기는 다양한 요인에 영향을 받을 수 있다. 예를 들어 태양의 고도, 계절, 구름의 양과 종류, 날씨, 등 여러 요인이 UV 세기에 영향을 줄 수 있다. 만약 다중회귀분석을 이용한다면 다양한 독립변수를 설정해 UV 세기에 주는 요인을 파악 및 분석할 수 있기에 더 효과적인 분석이 될 수 있다.