

Forecasting Air Quality of Boston

A photograph of the Boston skyline across a body of water, likely Boston Harbor. The skyline features several tall skyscrapers, including the Prudential Tower. In the foreground, several sailboats with white and red sails are visible on the water, along with a few small rowing boats. The water is a dark blue-grey color. The sky is overcast and hazy.

Hande Gulbagci Dede

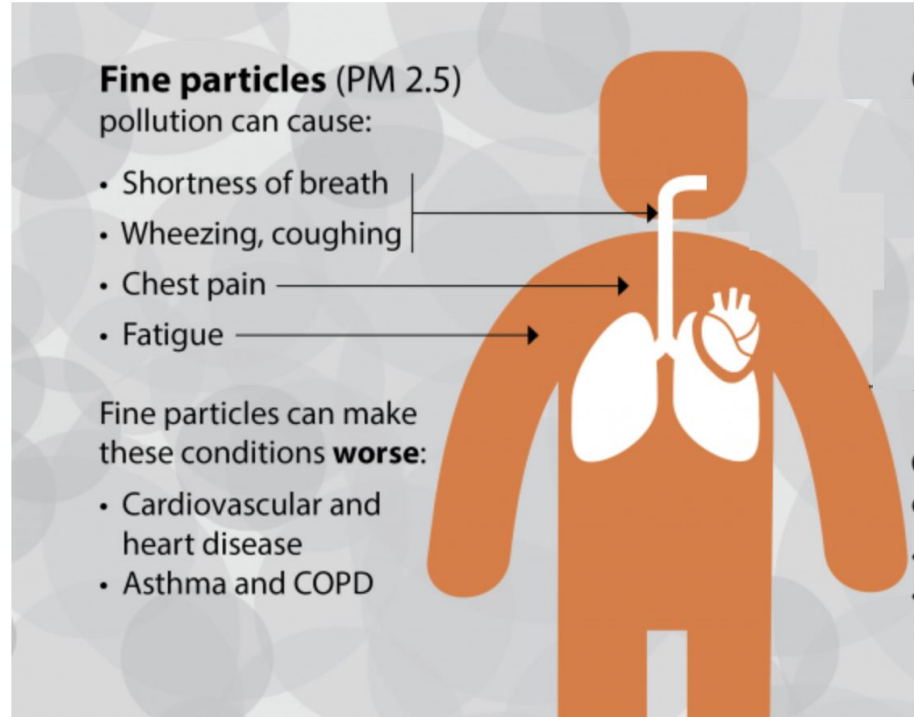
Mentor: Ajith Patnaik

Air Pollution

- Contamination of the environment by any physical, chemical, or biological agent changes the atmosphere's natural characteristics.
- **The Air Quality Index**, the concentration of pollutants in the air in a specific place, is accepted as an indicator of air pollution.

Problem

Air pollution can cause both short-term and long-term effects.



Importance of Air Pollution Forecasting

- To reduce the damage caused by air pollution to human health.
- By forecasting air quality, authorities
 - can take precautions and
 - issue early warnings
- **Aim:** Forecasting the air quality of Boston, the capital and largest city in Massachusetts in the USA, for 2023.



Data Information

1	Air Quality Data	<ul style="list-style-type: none">• Daily average air quality of Boston Station• Source: Quality Historical Data Platform• Parameters: PM25, O3, NO2, SO2, and CO.
2	Weather	<ul style="list-style-type: none">• The weather can significantly impact air quality• Source: visialcrossing.com• 33 attributes (temperature, humidity, wind, etc.)
3	Energy Consumption	<ul style="list-style-type: none">• Source: US Energy Information Administration• The natural gas delivered to consumers and total gasoline all sales in Massachusetts
4	Flight Number and Population	<ul style="list-style-type: none">• Each month's total airport flight number (domestic and international)• Source: Boston Logan International Airport
5	Federal Highway Administration's Reports	<ul style="list-style-type: none">• Source: Office of Highway Policy• Estimated travel on all roads and streets in Massachusetts and the amount of gallons taxed

Data Information

Air pollution indicator: $\text{PM}_{2.5}$

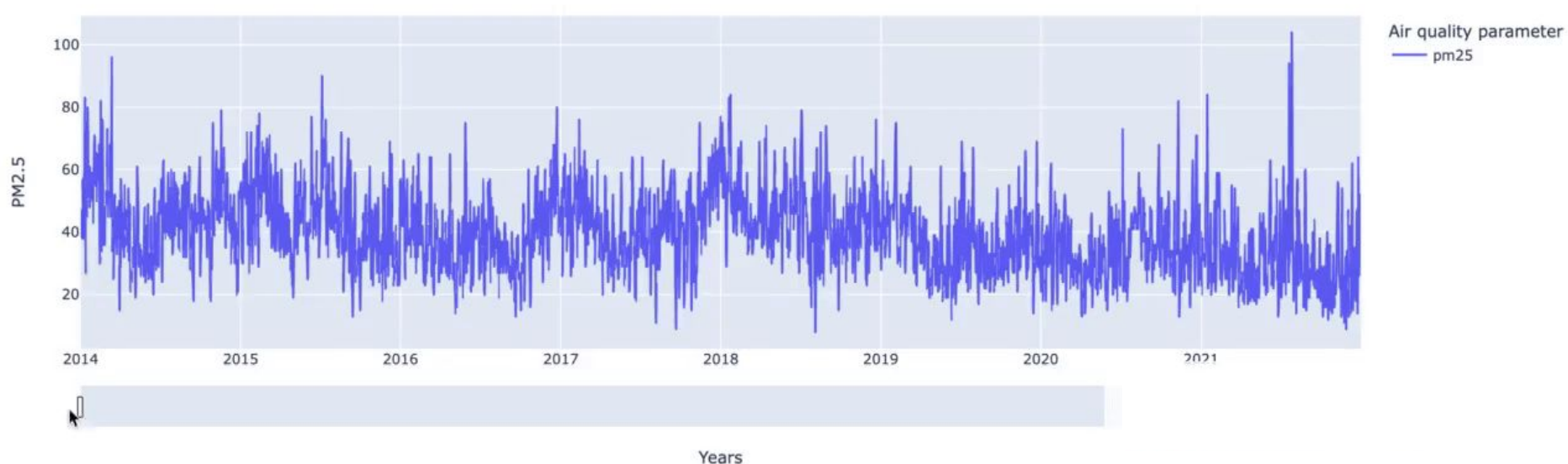
- Tiny particles of solid or liquid suspended in a gas.
- 2.5 refers to the diameter of particles which is 2.5 microns or less.
- Since being smaller they can be transferred to longer distances.

Data period: From 2014 January to 2021 December

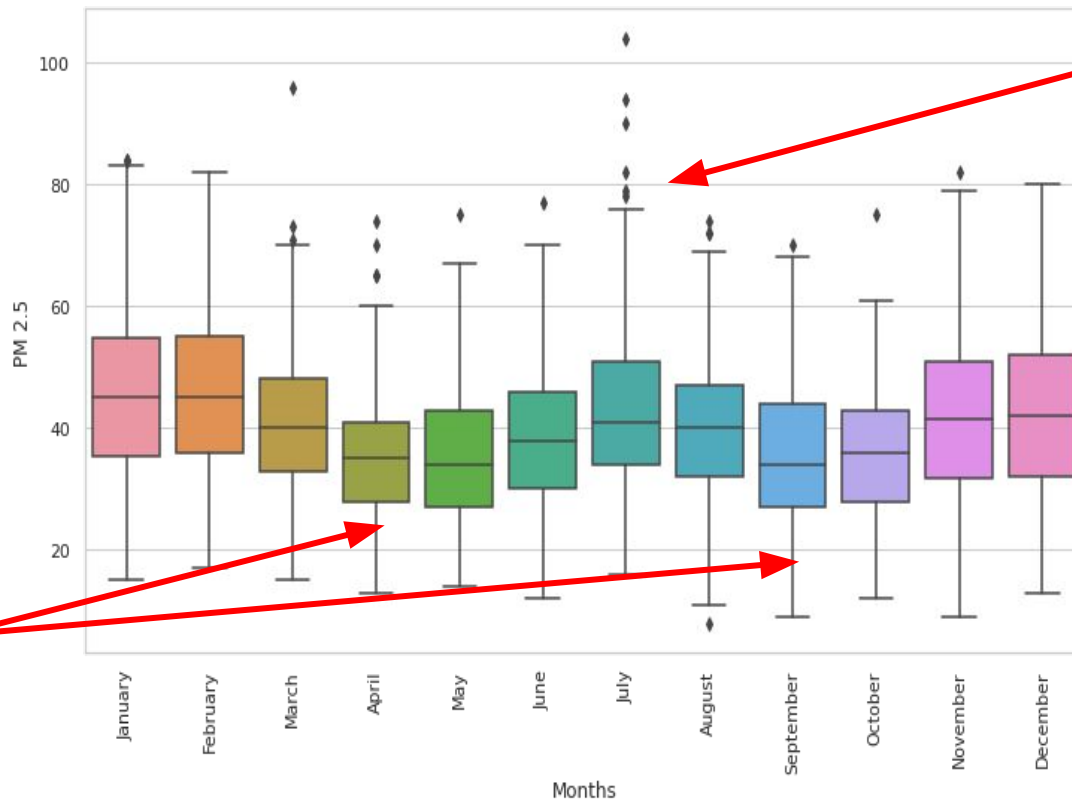
A total of 33 attributes

Changes in PM_{2.5} between 2014 and 2021

Change in PM2.5 between 2014-2021



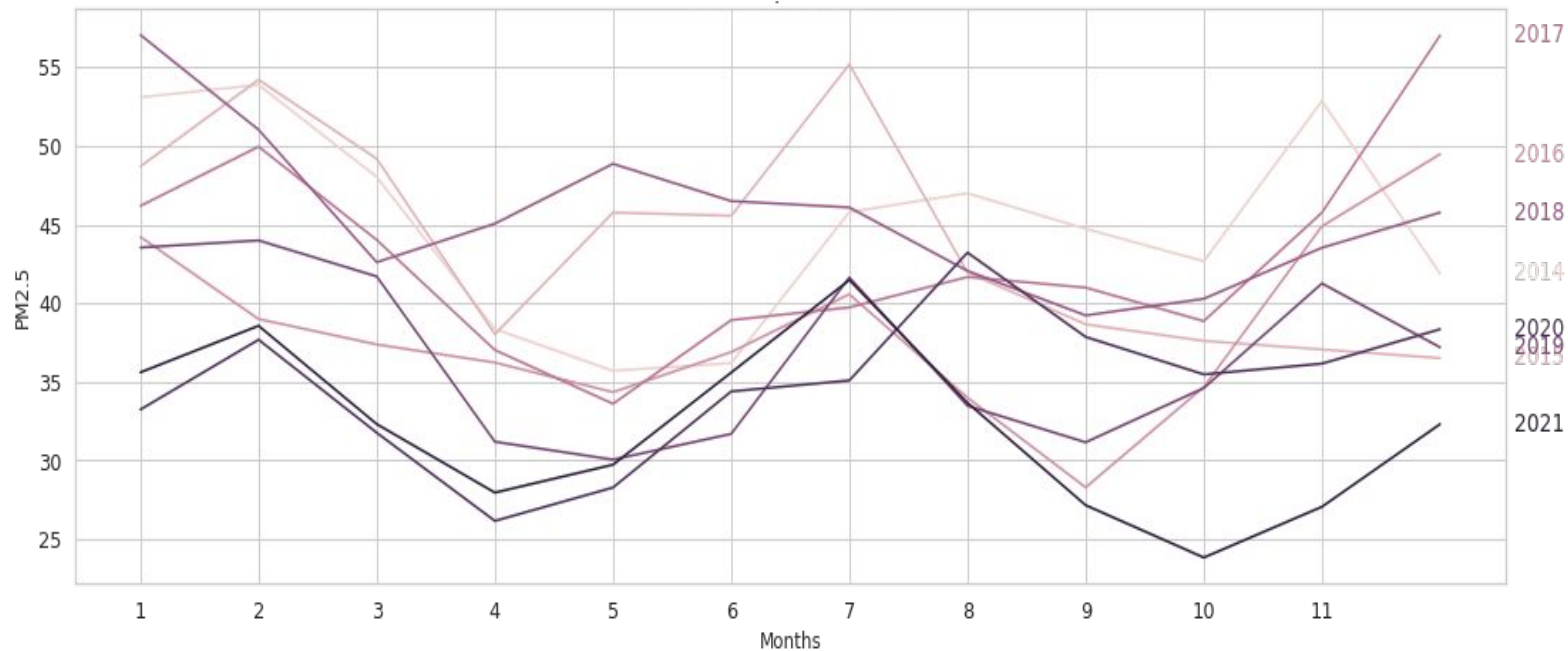
Monthly Average of PM_{2.5}



**Decrease in
spring and
fall**

**Increase in
summer
-JULY-**

Seasonal Plot of PM_{2.5}



Modeling Overview

I tested **univariate** time series models:

- Naive, Seasonal Naive, Simple Average as simple methods
- Simple Exponential Smoothing, Holt-Winters,
- ARMA, ARIMA, SARIMA,
- FbProphet,
- PyCaret
- LSTM.

Cross-validation and hypermeter tuning

Modeling Overview

Feature Selection for exogenous variables

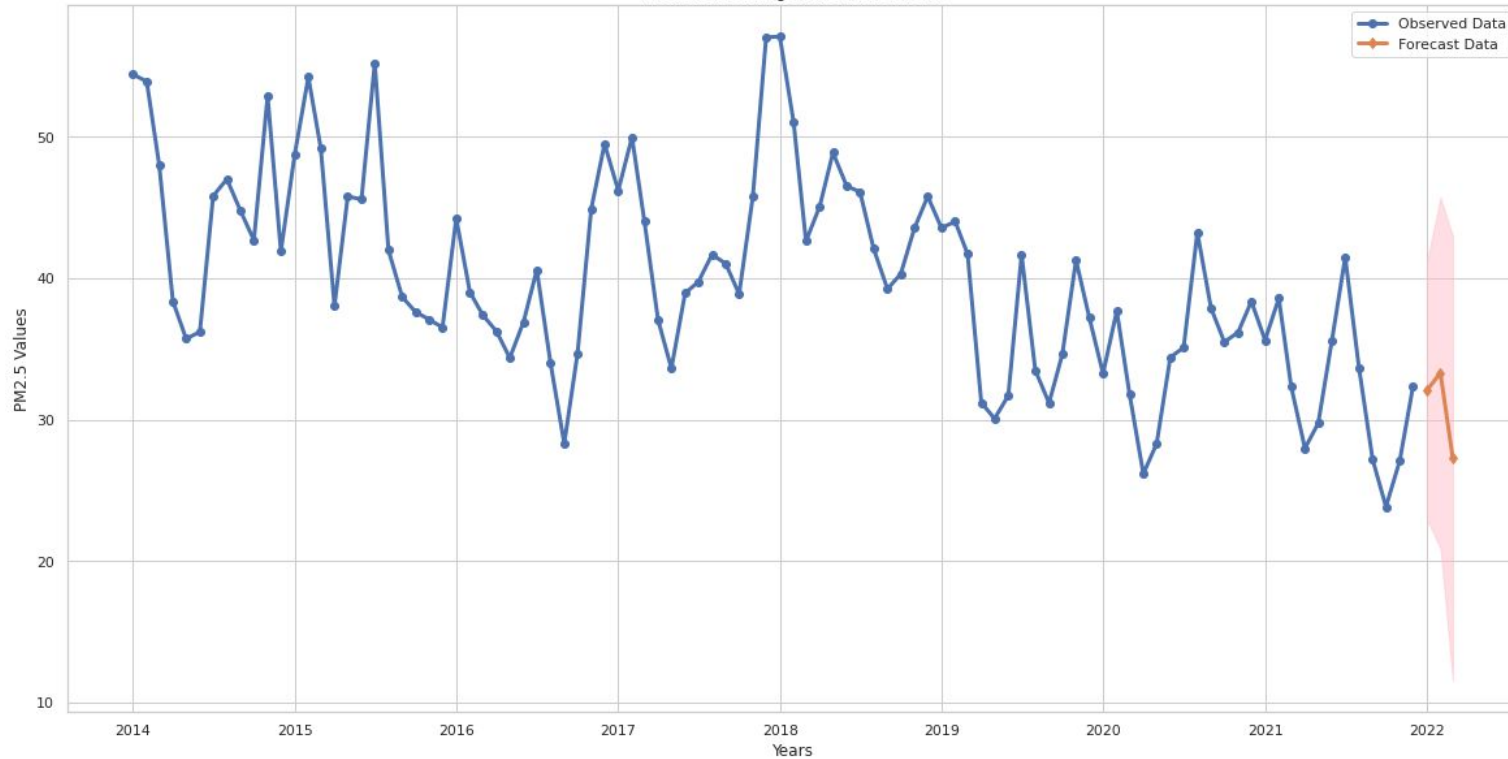
- Random Forest Regression,
- Stepwise Regression - Backwards elimination approach.

Multivariate time series models:

- ARIMAX,
- SARIMAX,
- FbProphet,
- PyCaret

Forecasting

Forecasting $PM_{2.5}$ values in 3 months of 2022 with SARIMA Model



Future Improvements

- Using observed data of January, February, and March of 2022 for assessing performance of champion and challenger models.
- Univariate performed better than multivariate models. Some exogenous features were not statistically significant.
 - Examine different features such as wildfires and vehicle/human mobility
- Predicting air quality for smaller areas
- Gathering data from stations much closer to each other.

A scenic view of a city skyline across a body of water. The skyline features several tall, modern skyscrapers and older buildings. In the foreground, the water is dark blue with several sailboats and small boats. The text "Thank you" is overlaid in a large, teal font.

Thank you

Hande Gulbagci Dede
e-mail: handeguldede@gmail.com
Github: <https://github.com/hangulde>