

Pollution - The World’s Silent Killer

ZUHA AHMAD
Data Science Institute



Abstract

The objective of the project is to execute a statistical modeling by performing a regression analysis and a time series forecasting of pollution based on traffic sensors. We will use programming tools and techniques, specifically the Zepellin notebook, to import the geo-spatial and environmental data, inputting a superlative algorithm, and processing the data with visualization. The contributing variables in this dataset are ozone index levels, carbon monoxide, nitrogen dioxide, sulfur dioxide, and particulate matter. By observing the trend and seasonality associated with each pollutant, we can determine which chemicals causes more harm, check the low and high peaks of the season, and come up with business ideas to reduce cost.

Introduction

Cities can be regarded as artificial environments created for the benefit of humans, yet none of this liberates people from nature. Cities have to be fed and provided with energy and water in order to exist and sustain its human population. The bigger the city, the more resources and energy must be extracted from the surrounding areas to sustain it. With the growth of cities and industrialization, urban waste and pollution are an important part of history. From the rise of the first urban settlements until the emergence of mega cities in the 20th century, the scale and intensity of pollution has increased.

Air Quality Index

| Air Quality Index Levels of Health Concern | Numerical Value | Meaning |
|--|-----------------|--|
| Good | 0-50 | Air quality is considered satisfactory, and air pollution poses little or no risk. |
| Moderate | 51-100 | Air quality is acceptable; however, for some pollutants there may be a moderate health concern for a very small number of people who are unusually sensitive to air pollution. |
| Unhealthy for Sensitive Groups | 101-150 | Members of sensitive groups may experience health effects. The general public is not likely to be affected. |
| Unhealthy | 151-200 | Everyone may begin to experience health effects; members of sensitive groups may experience more serious health effects. |
| Very Unhealthy | 201-300 | Health alert: everyone may experience more serious health effects. |
| Hazardous | > 300 | Health warnings of emergency conditions. The entire population is more likely to be affected. |

Figure 1: Air Quality Index based on government agencies

Contributing Variables

- Ozone index levels
- Carbon Monoxide
- Nitrogen dioxide
- Sulfur dioxide
- Particulate matter
- Latitude
- Longitude

Methods

Regression Analysis

- Go-to method in analytics
- Explore the forms of these relationships
- Widely used for prediction and forecasting
- Estimating the relationships among variables
- Used to understand which among the independent variables are related to the dependent variable

Results

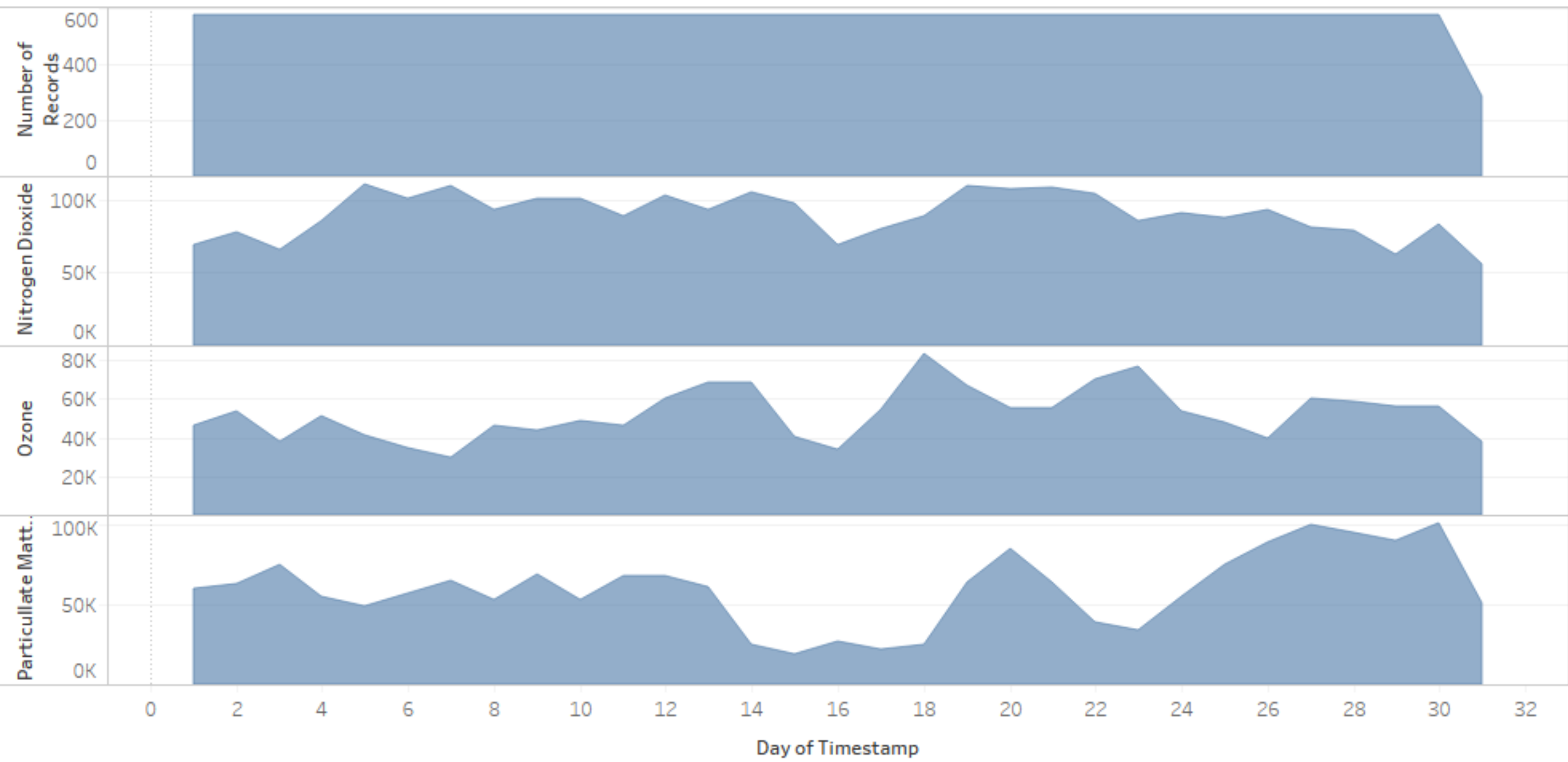


Figure 2: This graph shows four shape distinctions in a timestamp for particulate matter, nitrogen dioxide, and ozone layers.

Conclusions

- Number of records proved to be an all-time high
- Nitrogen Dioxide has an exponential increase
- Particulate matter had a lower peak from 14-18 timestamp

References

[1] Pramod Anantharam, Payam Barnaghi, Krishnaprasad Thirunarayan, and Amit Sheth. Extracting city traffic events from social streams. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 6(4):43, 2015.

[2] <http://iot.ee.surrey.ac.uk:8080/index.html>. Pollution measurements, 2014.

[3] Huamin Qu, Wing-Yi Chan, Anbang Xu, Kai-Lun Chung, Kai-Hon Lau, and Ping Guo. Visual analysis of the air pollution problem in hong kong. *IEEE Transactions on visualization and Computer Graphics*, 13(6), 2007.

[4] Vicki Stone, Helinor Johnston, and Martin JD Clift. Air pollution, ultrafine and nanoparticle toxicology: cellular and molecular interactions. *IEEE transactions on nanobioscience*, 6(4):331–340, 2007.

[5] Yu Zheng, Licia Capra, Ouri Wolfson, and Hai Yang. Urban computing: concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(3):38, 2014.