U.S. Air Pollution Insight Platform

Team-089: Ravi Thakkar, Guneet Sachdeva, Mitchell Gilmer

Project Summary

The U.S. Air Pollution Insight Platform is a web application dedicated to visualizing, analyzing, and interacting with pollution data from the U.S. spanning from 2000-2016. The platform will serve as a centralized hub to easily access insights into historical pollution trends for researchers, environmentalists, and concerned citizens.

Dataset Background

The Dataset is one of the TA-Approved choices provided to us. This dataset contains daily pollution data on four major pollutants (Nitrogen Dioxide, Sulphur Dioxide, Carbon Monoxide and Ozone) from the USA's Environmental Protection Agency (EPA). The dataset has been publicly made available on Kaggle by Brenda So and contains a total of 28 fields and 1.4 million observations. Many columns seem to deal with the detailed physical location of pollution, the exact amount (in parts per billion) of emissions, and recorded dates of pollution.

Here is a link to the dataset: https://www.kaggle.com/datasets/sogun3/uspollution

Application Description

After initially researching different ideas and datasets to use for the project, this one piqued our interest due to its increasing relevance and impact for sustainability. Our initial interest in this project led us to comb through the EPA's website to find comprehensive information on pollution data, but what we found was a hard-to-navigate and unintuitive platform. For researchers and everybody else, gathering the data you need through the EPA's website seems like a headache-inducing task.

Our idea is to make it much easier for anybody to access US pollution data and use it to gain insights into the U.S.'s pollution patterns. We want to provide a user-friendly interface to interact with the data and offer tools for filtering, visualizing, and deriving actionable insights. By converting raw data into intuitive graphs, heatmaps, and trend analysis, we aim to foster environmental awareness and promote research initiatives.

Usefulness

While there are other websites that provide environmental data, such as the U.S. EPA's official site, they often do not have user-friendly tools for quick visualization and analysis. Our website differs by offering a streamlined experience, with features like comparison tools, predictive modeling, and location-specific breakdowns. This caters to both experts who want a deep dive and laypeople looking for an overview.

Realness:

The data driving our project is sourced directly from the U.S. EPA's database. With over 1.4 million observations spanning 28 fields, the authenticity and comprehensiveness of the dataset are beyond question.

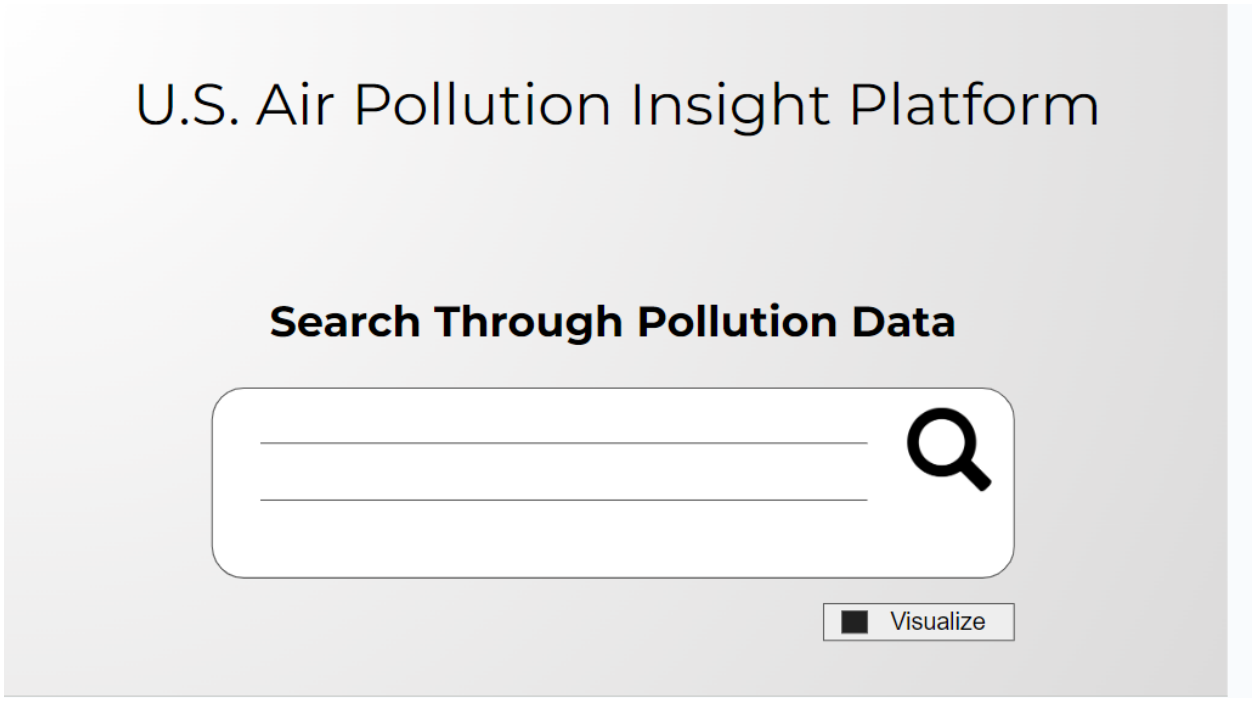
Functionality:

- Visualization Tools: Users can visualize pollutant levels over time using dynamic graphs, compare different pollutants, and examine trends year-on-year.
- Location Filters: Data can be filtered by state, city, or specific regions to allow users to hone in on areas of interest.
- Comparison Tool: Users can compare pollution levels between different locations, periods, or pollutants.
- Search Function: Enables users to quickly find specific data points or periods.

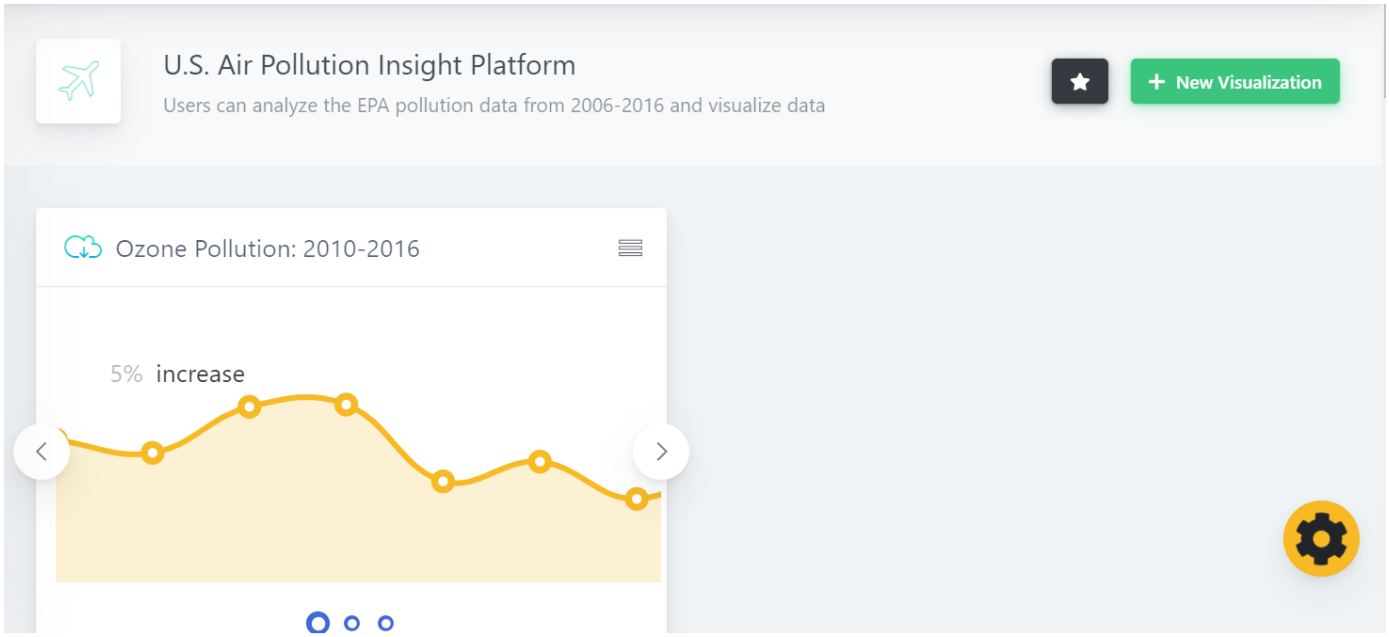
Project Work Distribution:
Frontend: Ravi
Setting up database: Ravi, Guneet
Connecting database to application: Varik
Backend Stored Procedures/algorithm: Varik, Guneet, Ravi

UI Mockup

Search / Retrieve Data Page:



Visualizations Page :



I'm planning to use a react dashboard library to conveniently allow for all functionality with a nice looking website.