Cancer Atlas
User Manual

Itay Zahor, Shmuel Granot

About the Application

Our application, Cancer Atlas, is an interactive platform for exploring cancer statistics and their relationships with various factors, including socioeconomic, environmental, and behavioral risk factors in the United States. This guide will help you set up and run the application.

Prerequisites

Before you begin, ensure that your computer has the following installed:

- Python 3
- pip (Python's package manager)
- MySQL Server

Installation Steps

1. Clone the Repository
Clone the repository to your local machine using the following command:

```
git clone <https://github.com/itayzahor/Cancer_Atlas_app.git>
```

2. Create a Virtual Environment (Optional)
Creating a virtual environment is recommended for managing dependencies. Run the following command:

```
python -m venv venv
```

3. Activate the Virtual Environment (Optional)
On Windows:

```
.\venv\Scripts\activate
```

On Linux or macOS:

```
source venv/bin/activate
```

4. Install Required Libraries
Install the necessary Python libraries by running:

```
pip install -r requirements.txt
```

If the installation fails, refer to the file:

Python_libraries_installs.txt

This file contains all the libraries downloaded for this project.

Database Setup

1. Ensure that the MySQL Server is running on your machine.
2. Create a database named cancer_atlas in your MySQL Server.
3. Import the necessary tables and data using the data.sql file provided in the repository. On mysql Click on server and then data import and select the data.sql file.
4. The application uses the db_connector.py file to connect to the database. Make sure the connection parameters (e.g., host, user, password, database name) in this file match your MySQL setup.

Running the Application

Start the Flask application by running:

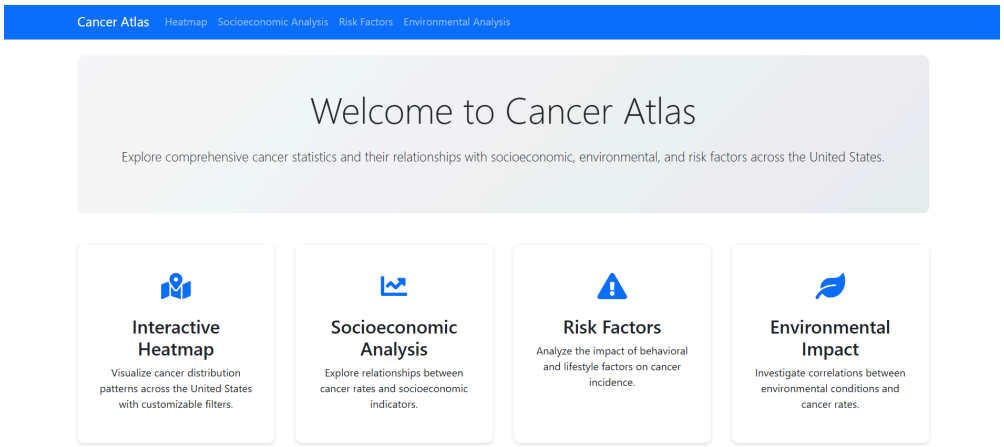
```
python run.py
```

Access the Application:

Once the server is running, the application will be accessible at:

http://127.0.0.1:5000/cancer_atlas

Home Page

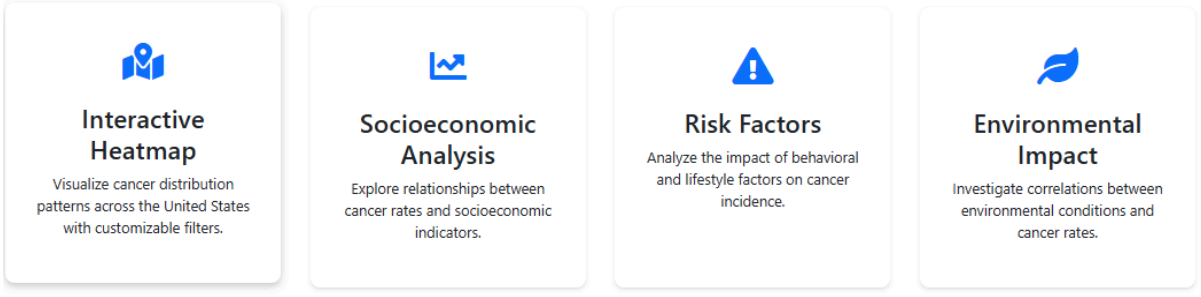


The Home Page serves as the central hub of the application, providing navigation to all the main features and activities. It also lists the data sources used in the application, ensuring transparency and accessibility.

Navigation Links:

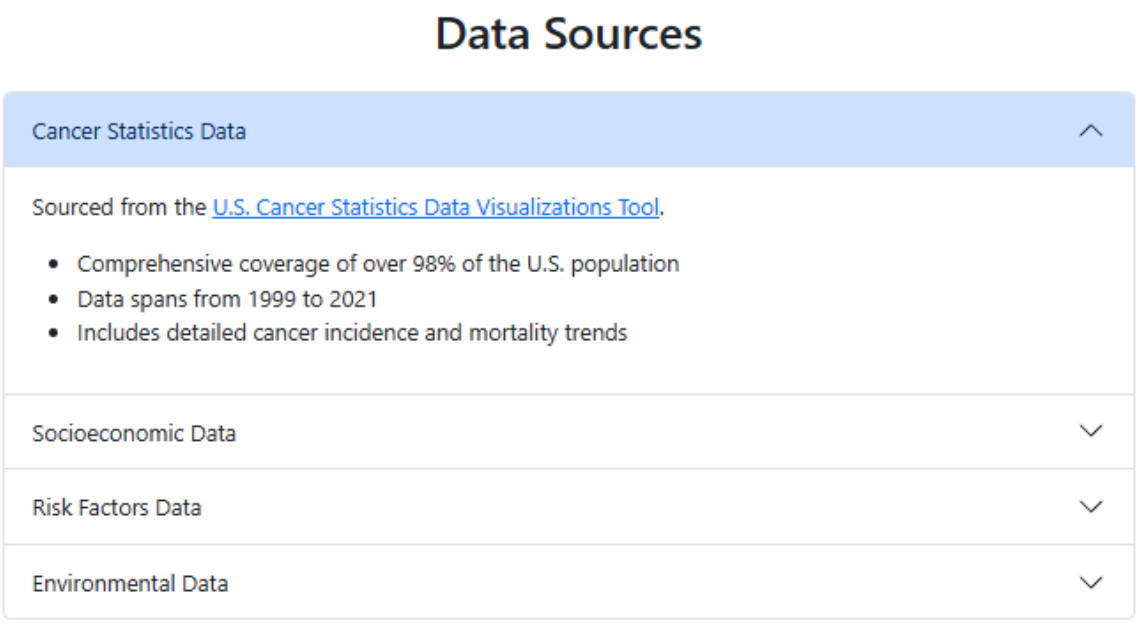
At the top of the page, you will find links to all the main activities of the application:

In the middle of the page as well.



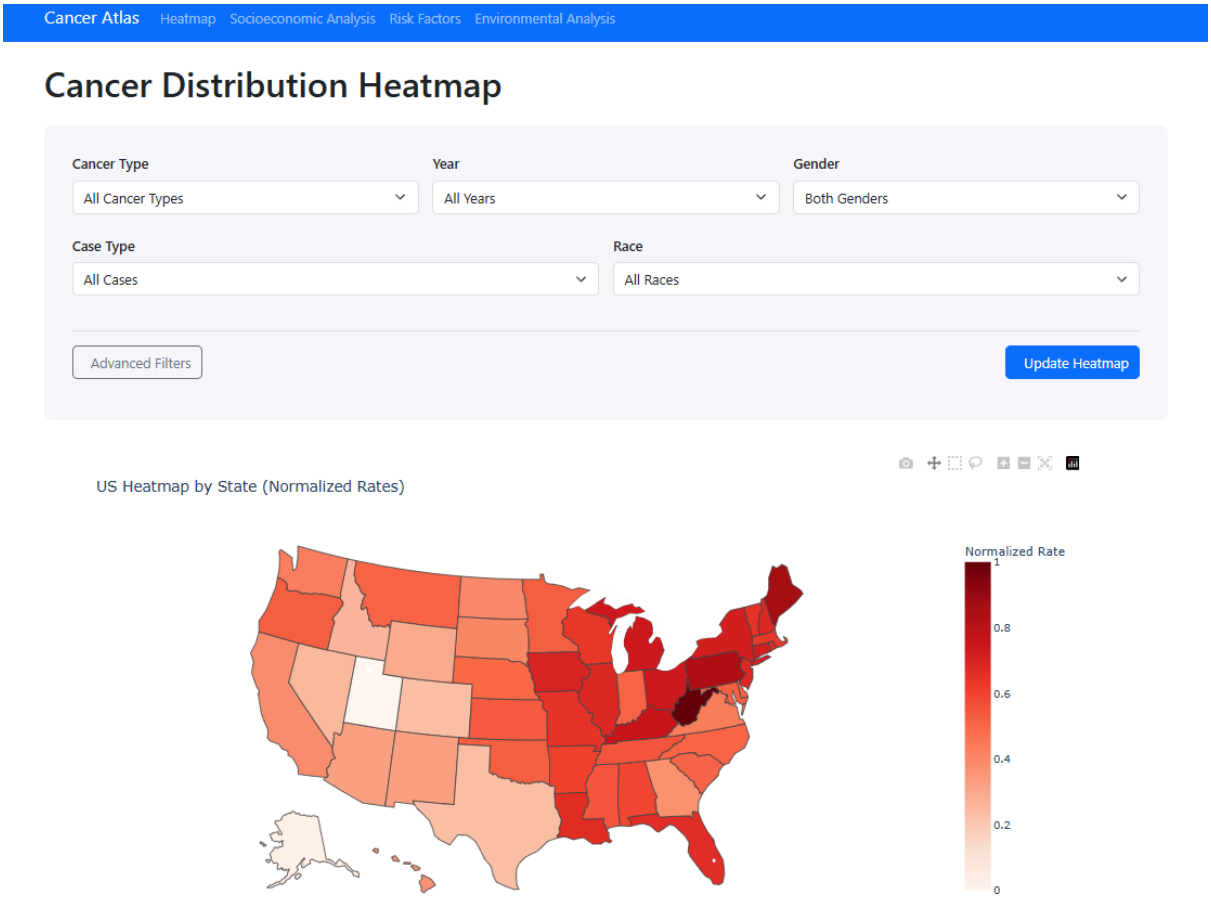
Data Sources:

At the bottom of the page, the Home Page lists all the data sources used in the application. Each source includes:



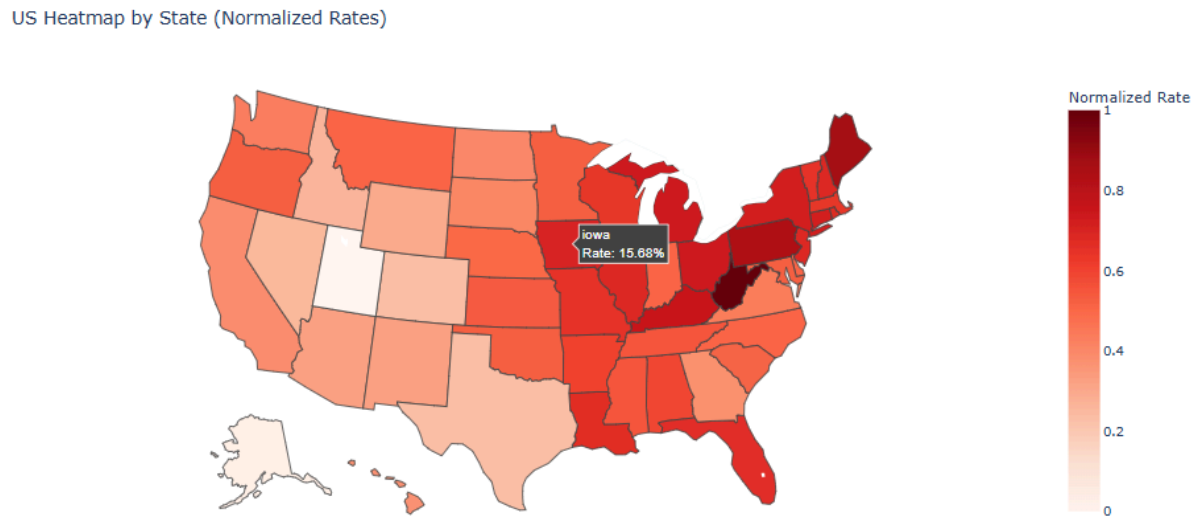
A brief description of the dataset.
A clickable link directing users to the original source for more details or downloads.

Heatmap Page



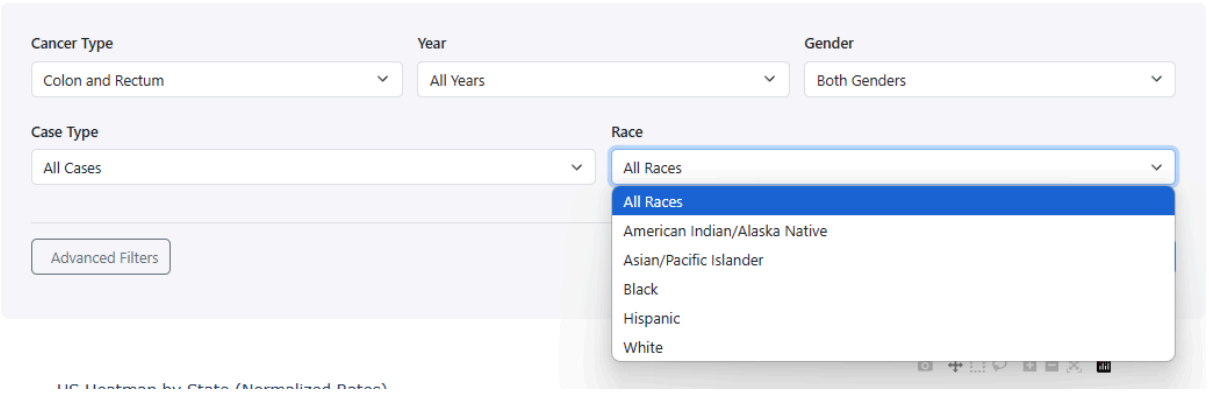
The Heatmap Page provides an interactive map of the United States, displaying cancer rates across states. Users can apply filters to customize the data displayed and gain insights into cancer trends and disparities.

Interactive Heatmap:



Color-coded map showing cancer rates across states.
Hover over a state to view detailed information like cancer rate by the filters and the name of the state

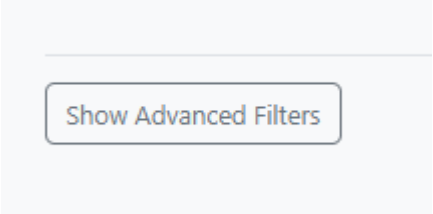
User Filters:



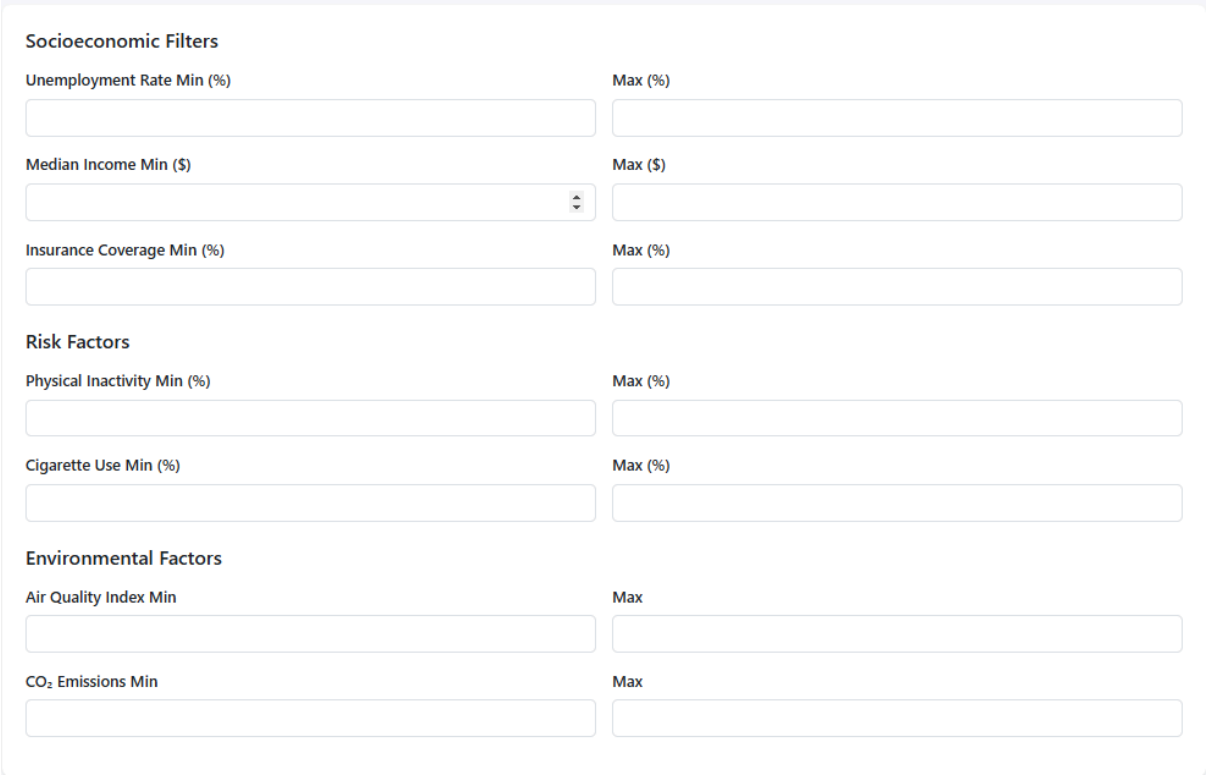
Dropdown that allows the user to filter the heatmap according to his curiosity such as : Cancer type, year, gender, New Cases/Cancer related deaths and race.

Advanced filters:

To use the advanced filters press on:



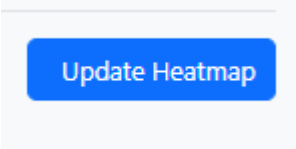
Then this options will come up:



The advanced filters allow users to refine the heatmap by applying state-specific statistics. For example, users can display only states with a median income below 90,000.

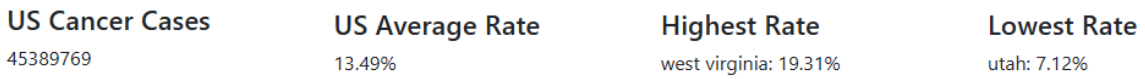
Users can select a range (e.g., minimum and maximum values) or specify only one boundary (e.g., maximum value only). The filters will apply the appropriate limits based on the input.

Update Heatmap:



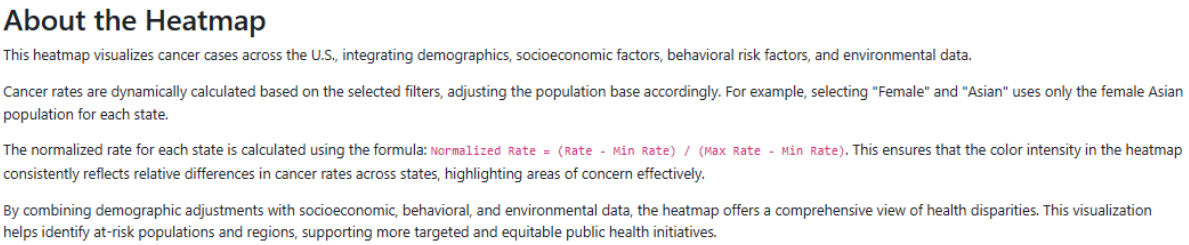
To see the updated heatmap according to the filters just press on “Update Heatmap” button and it will show

Statistics:



According to the filters you can see key statistics from the data

Info:

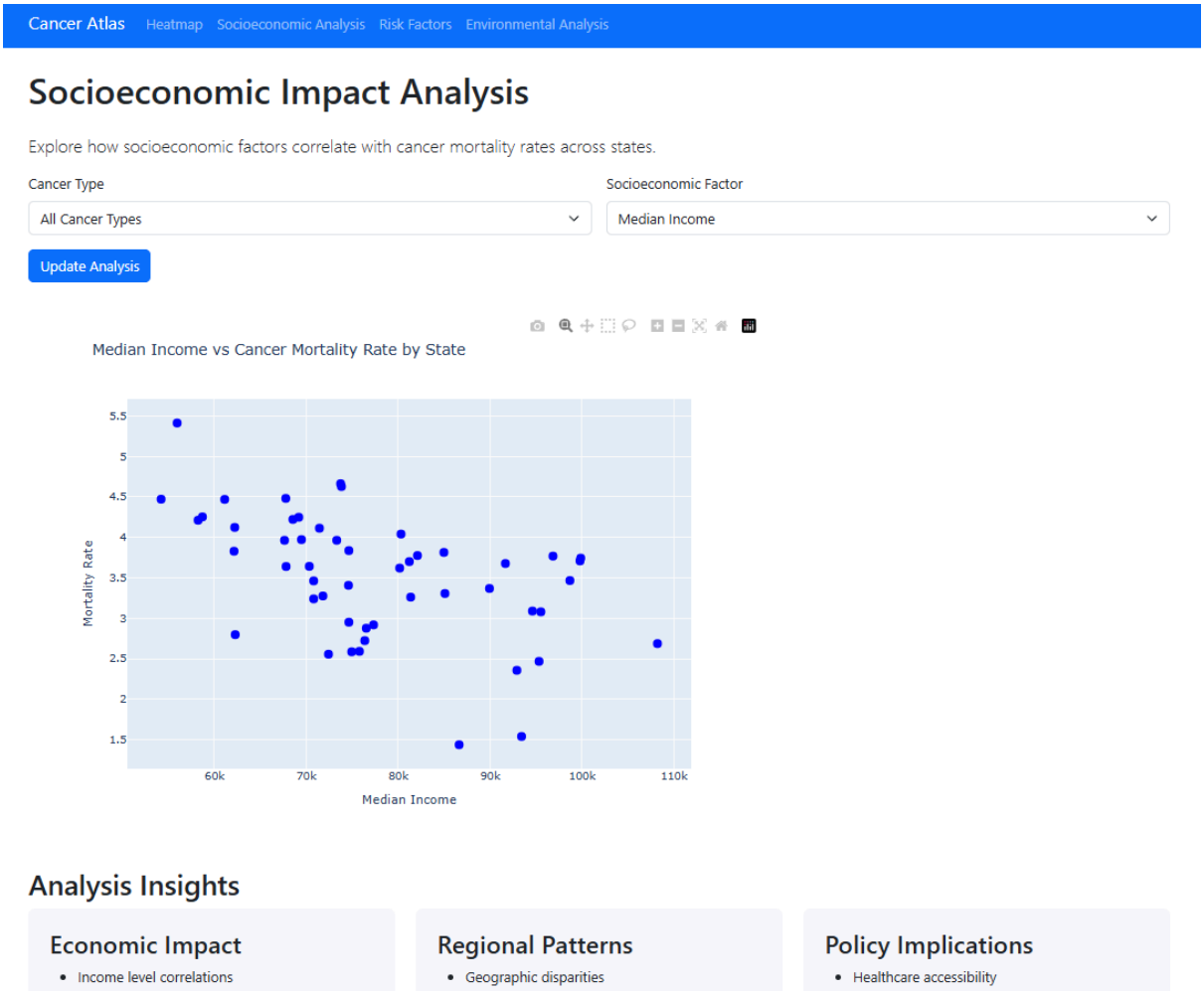


a brief explanation of how the heatmap data is calculated and normalized.

Downloadable Data:

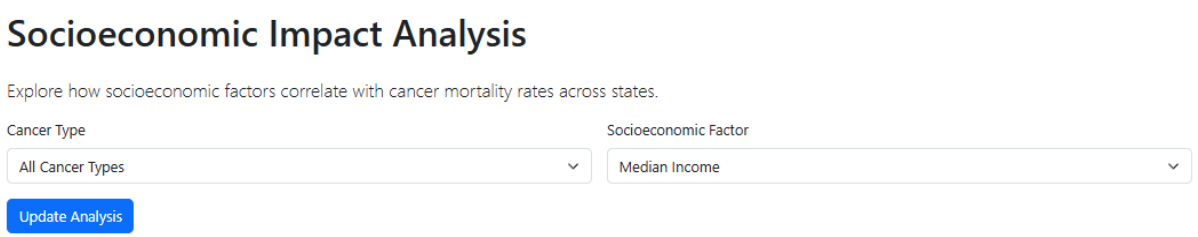
Export the data displayed on the heatmap as a CSV file for further analysis.

Socioeconomic Analysis Page



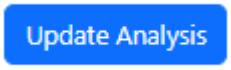
The Socioeconomic Analysis page allows users to explore how various socioeconomic factors correlate with cancer mortality rates across U.S. states. It provides an interactive interface to select filters, visualize data through scatter plots, and gain insights into regional and policy-related patterns.

Filter Controls:



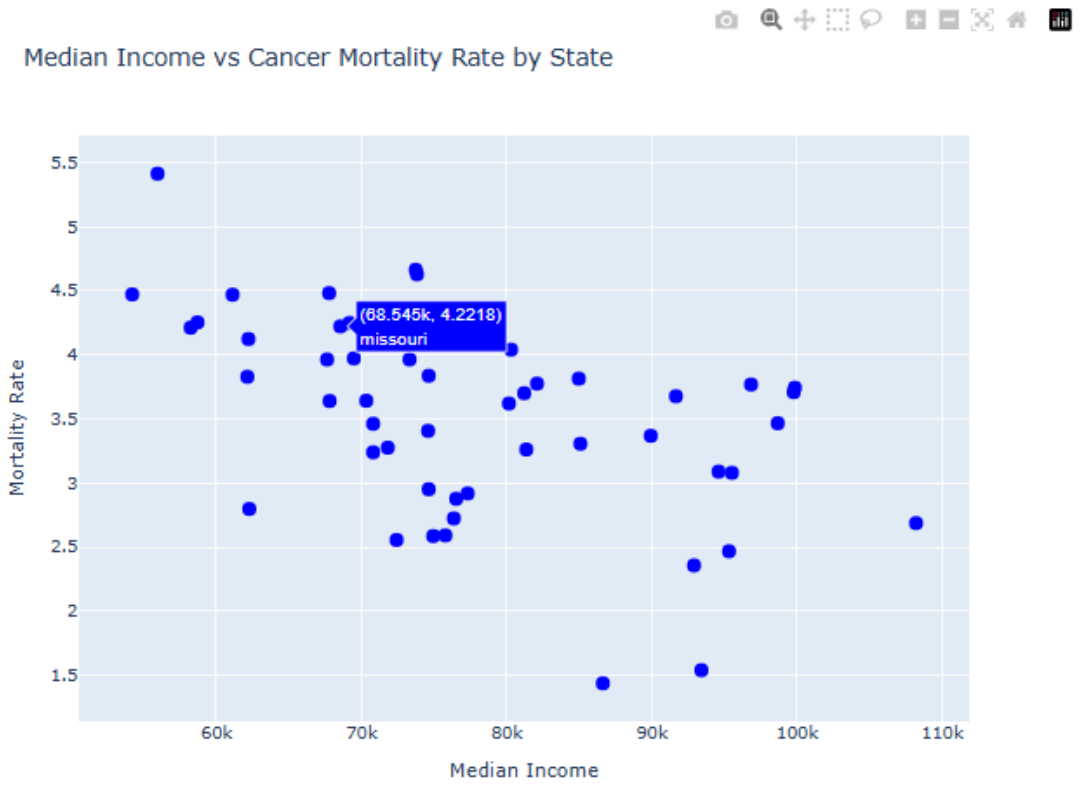
Users can select specific cancer types and a selection of a socioeconomic factor, such as:
Median Income: Correlation between income levels and cancer mortality.
Filters dynamically adjust the scatter plot visualization when submitted.

Update Analysis:



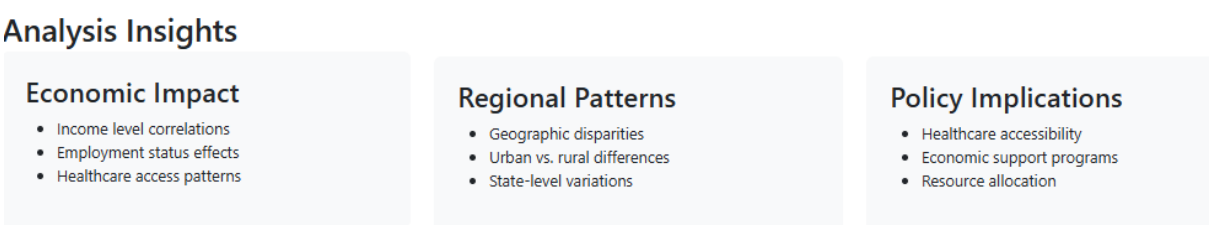
To update the analysis according to the filters just press on the button “Update Analysis”.

Visualization:



Displays a scatter plot that illustrates the correlation between the selected socioeconomic factor and cancer mortality rates for each state.
Hovering over a data point reveals detailed information about the state, including:
The name of the state.
The selected socioeconomic factor value.
The cancer mortality rate for that state.

Analysis Insights:



Provides contextual information about how the data can help users understand trends and correlations between socioeconomic factors and cancer outcomes.

Downloadable Data:

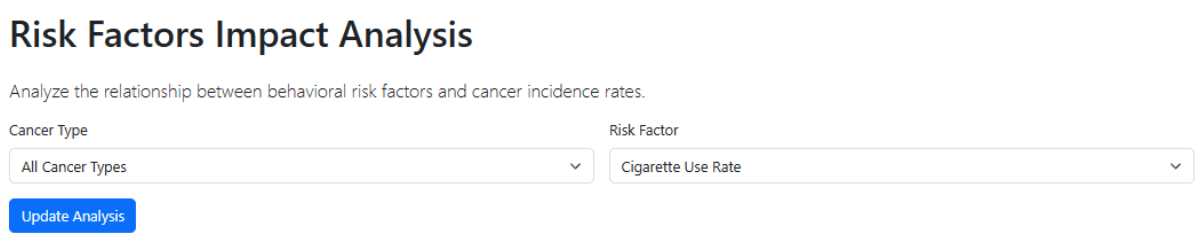
Download Data

Users can download the analyzed data as a CSV file for further research or offline analysis.

Risk Factors Analysis Page

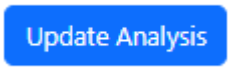
The Risk Factors Analysis page allows users to examine how behavioral risk factors, such as smoking and physical inactivity, correlate with cancer incidence rates across U.S. states. It provides an interactive interface to apply filters, visualize data through scatter plots, and gain valuable insights into the relationship between lifestyle choices and cancer outcomes.

Filter Controls:



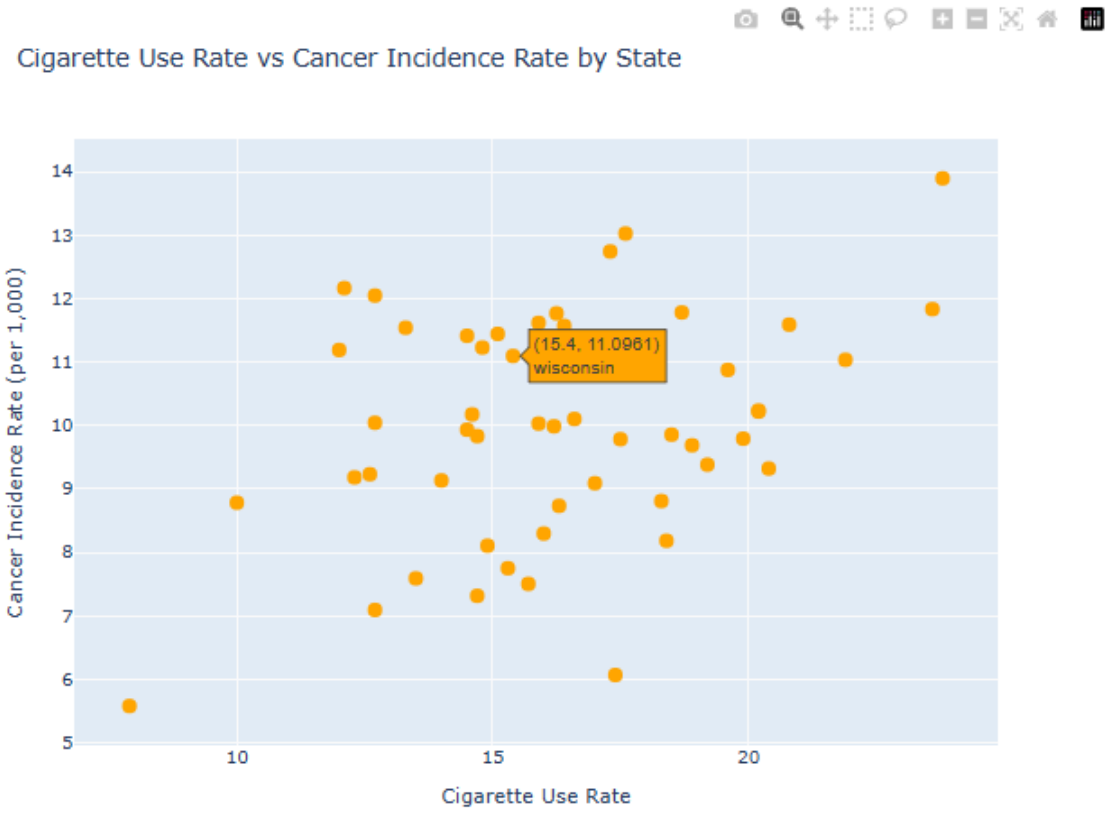
Users can select specific cancer types and a behavioral risk factor, such as: Smoking Rates and Physical Inactivity to Analyze the effect of sedentary lifestyles on cancer outcomes. Filters dynamically adjust the scatter plot visualization when submitted.

Update Analysis:



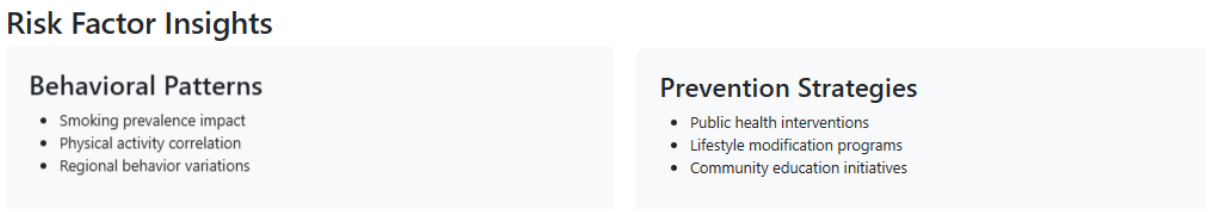
To update the analysis according to the selected filters, users can click the "Update Analysis" button.

Visualization:



The page displays a scatter plot that illustrates the correlation between the selected risk factor and cancer incidence rates for each state.
Hovering over a data point reveals detailed information about the state, including:
The name of the state.
The selected risk factor value.
The cancer incidence rate for that state.

Analysis Insights:

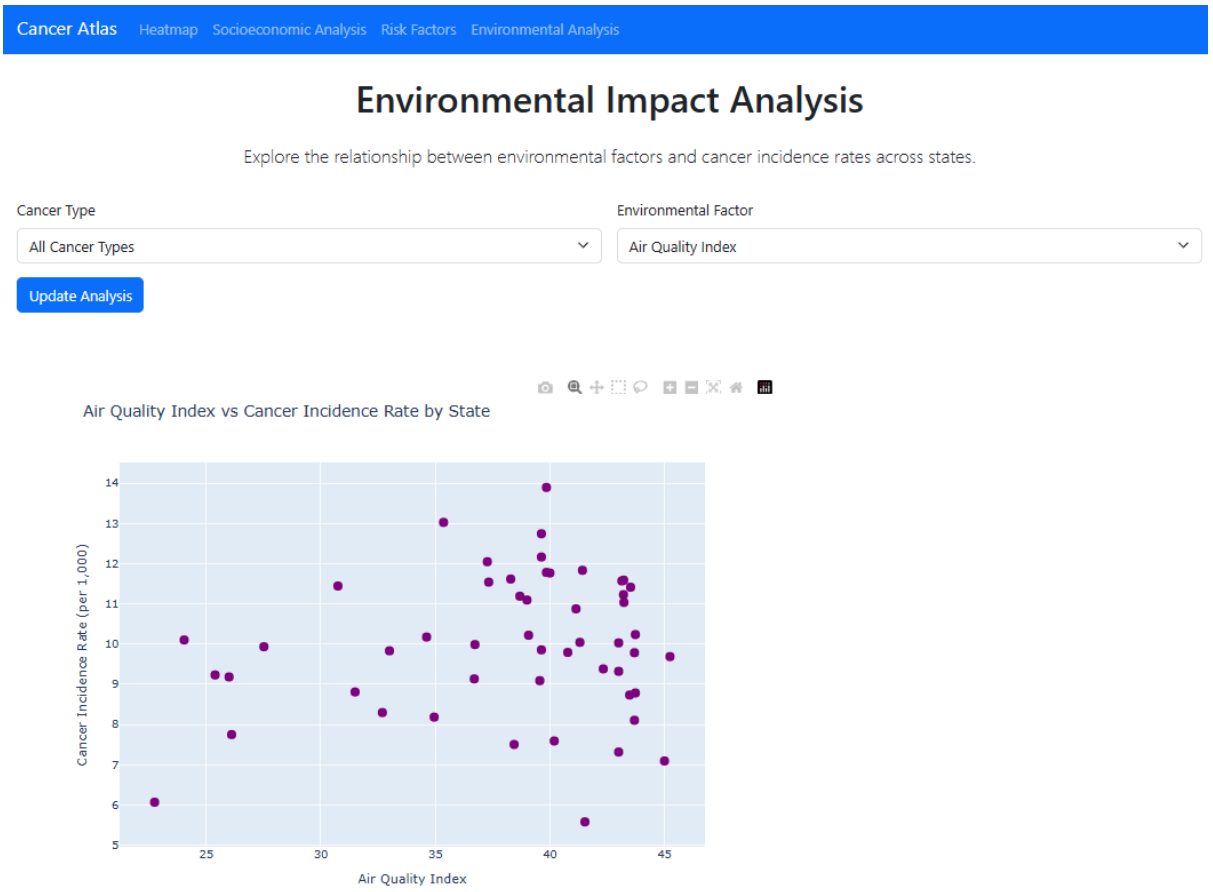


Provides contextual information about how the data can help users understand trends and correlations between risk factors and cancer outcomes.

Downloadable Data:

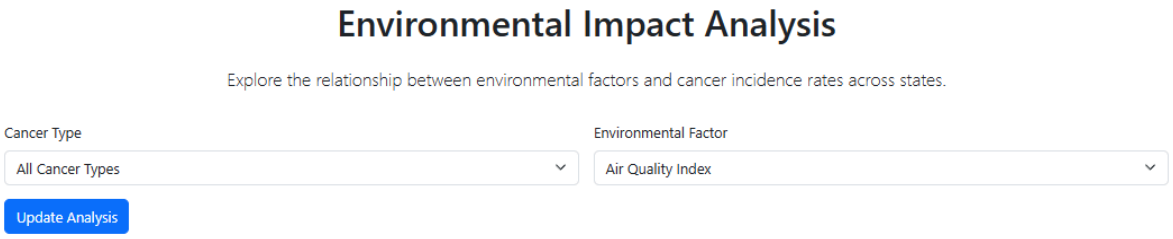
Users can download the analyzed data as a CSV file for further research or offline analysis.

Environmental Analysis Page



The Environmental Analysis page allows users to explore how environmental factors, such as air quality and CO2 emissions, correlate with cancer incidence rates across U.S. states. It provides an interactive interface for applying filters, visualizing data through scatter plots, and gaining insights into how environmental conditions impact cancer outcomes.

Filter Controls:



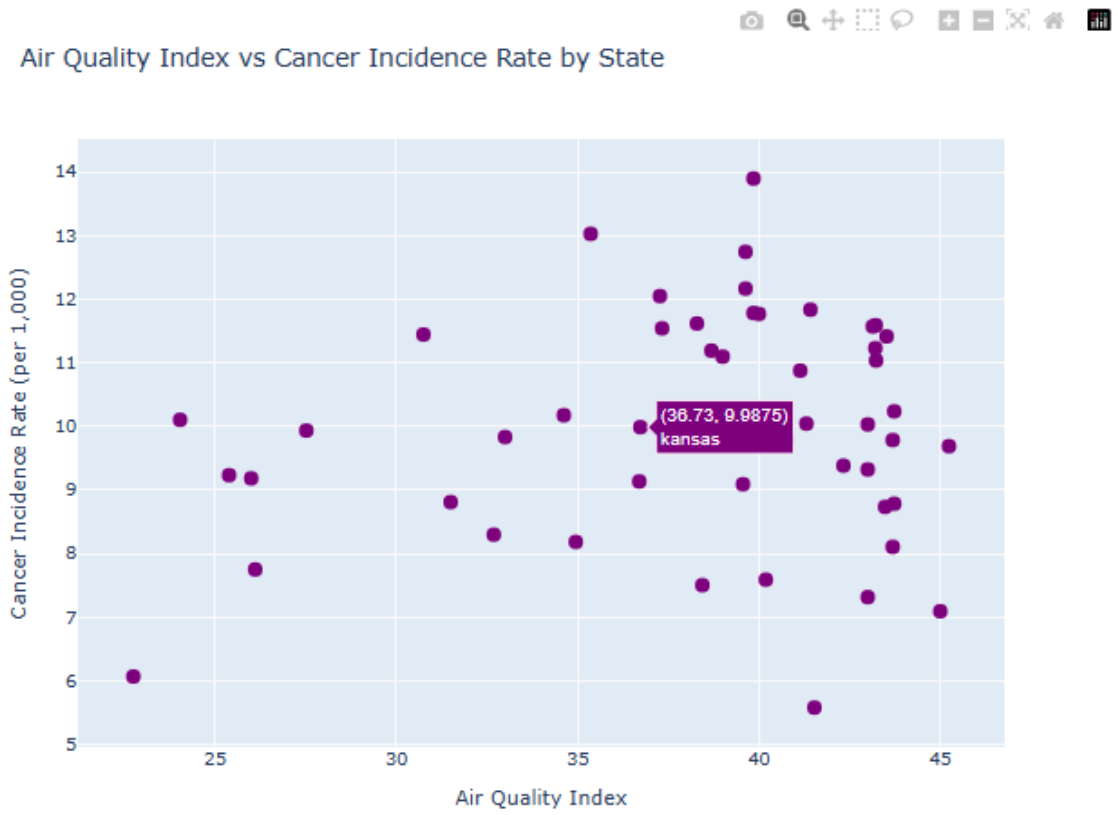
Users can select specific cancer types and an environmental factor to analyze, such as:
Air Quality Index (AQI): Assess how air pollution levels relate to cancer incidence.
CO2 Emissions: Examine the influence of carbon dioxide emissions on cancer outcomes.
Filters dynamically update the scatter plot visualization when submitted.

Update Analysis:

Update Analysis

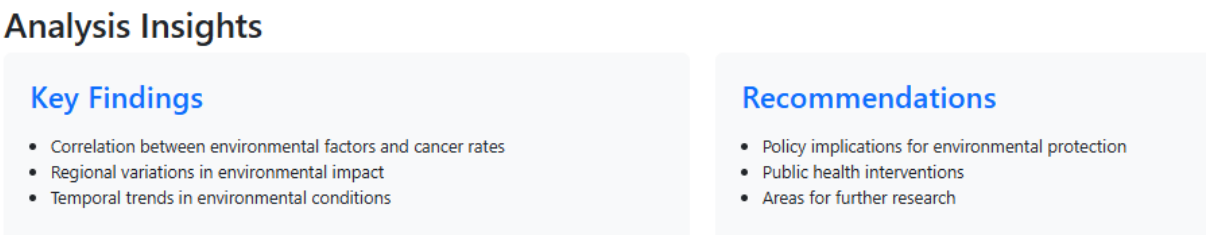
To apply the selected filters, users can click the "Update Analysis" button, which refreshes the visualization.

Visualization:



The page displays a scatter plot that highlights the correlation between the chosen environmental factor and cancer incidence rates for each state.
Hovering over a data point reveals:
The name of the state.
The selected environmental factor value.
The cancer incidence rate for that state.

Analysis Insights:



Offers information about how the data helps users understand the relationship between environmental conditions and cancer rates, providing valuable context for research and policymaking.

Downloadable Data:

Users can download the analyzed data as a CSV file for further research or offline analysis.

Tips

- Use the dropdown menus to customize the parameters to your needs.
- You can return to the Home Page at any time by clicking on the logo.
- Follow the instructions on each page to ensure a smooth user experience.

Troubleshooting

404 Error:
If you encounter a 404 error, ensure you are accessing the correct URL and that the application is running.

Database Connection Issues:
Verify that MySQL Server is active and the connection details in db_connector.py are correct (e.g the host, user, password and database) .