

Cancer Atlas

User Manual

Itay Zahor, Shmuel Granot

About the Application.....	3
Prerequisites.....	3
Installation Steps.....	3
Database Setup.....	4
Running the Application.....	4
Home Page.....	5
Navigation Links:.....	5
Data Sources:.....	6
Heatmap Page.....	7
Interactive Heatmap:.....	8
User Filters:.....	8
Advanced filters:.....	9
Update Heatmap:.....	10
Statistics:.....	10
Info:.....	10
Downloadable Data:.....	10
Socioeconomic Analysis Page.....	11
Filter Controls:.....	12
Update Analysis:.....	12
Visualization:.....	13
Analysis Insights:.....	13
Downloadable Data:.....	14
Risk Factors Analysis Page.....	15
Filter Controls:.....	16
Update Analysis:.....	16
Visualization:.....	17
Analysis Insights:.....	18
Downloadable Data:.....	18
Environmental Analysis Page.....	19
Filter Controls:.....	19
Update Analysis:.....	20
Visualization:.....	20
Analysis Insights:.....	21
Downloadable Data:.....	21
Tips.....	22
Troubleshooting.....	22

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:

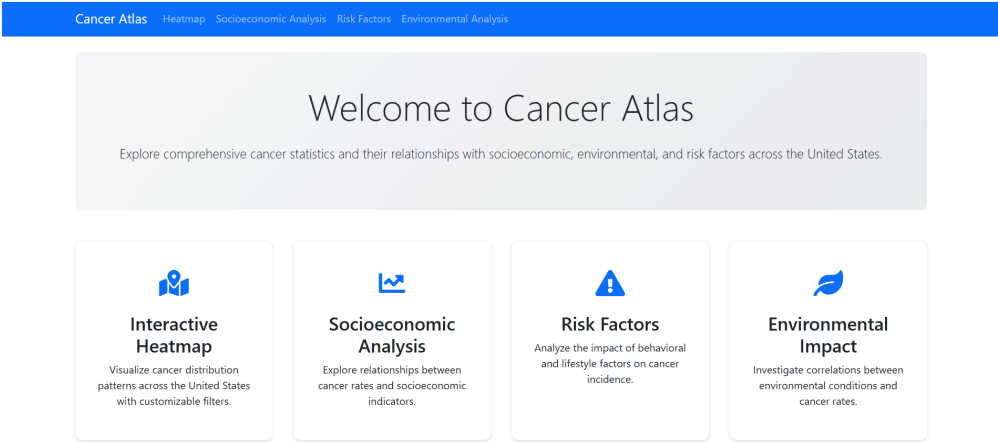
```
python run.py
```

Access the Application:

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

http://127.0.0.1:5000/cancer_atlas

Home Page



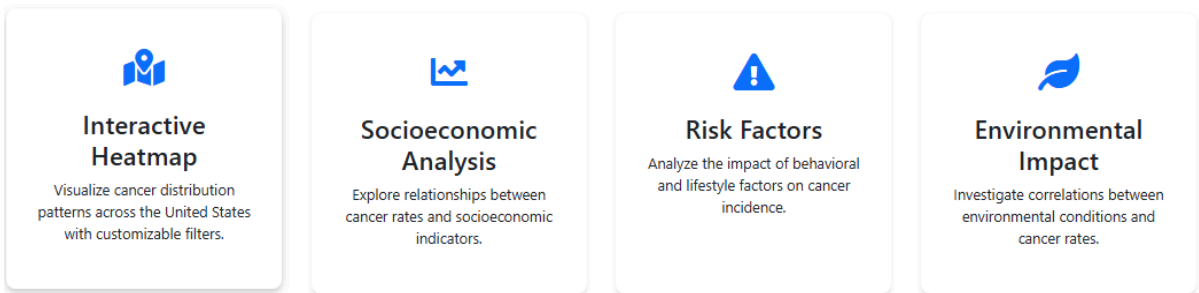
The Home Page serves as the central hub of the application, providing navigation to all the main features and activities. It also lists the data sources used in the application, ensuring transparency and accessibility.

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:

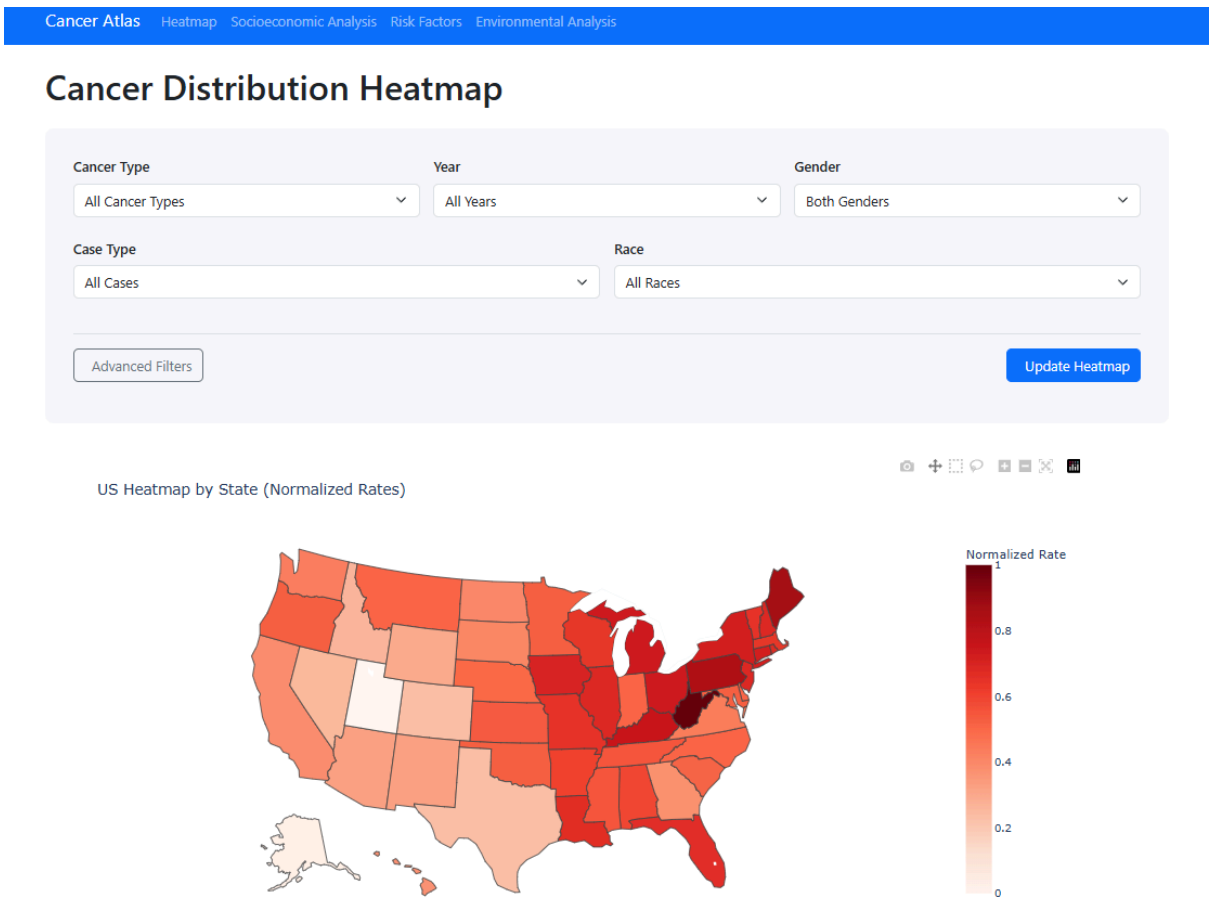
Data Sources

Cancer Statistics Data	^
Sourced from the U.S. Cancer Statistics Data Visualizations Tool .	
<ul style="list-style-type: none">• Comprehensive coverage of over 98% of the U.S. population• Data spans from 1999 to 2021• Includes detailed cancer incidence and mortality trends	
Socioeconomic Data	∨
Risk Factors Data	∨
Environmental Data	∨

A brief description of the dataset.

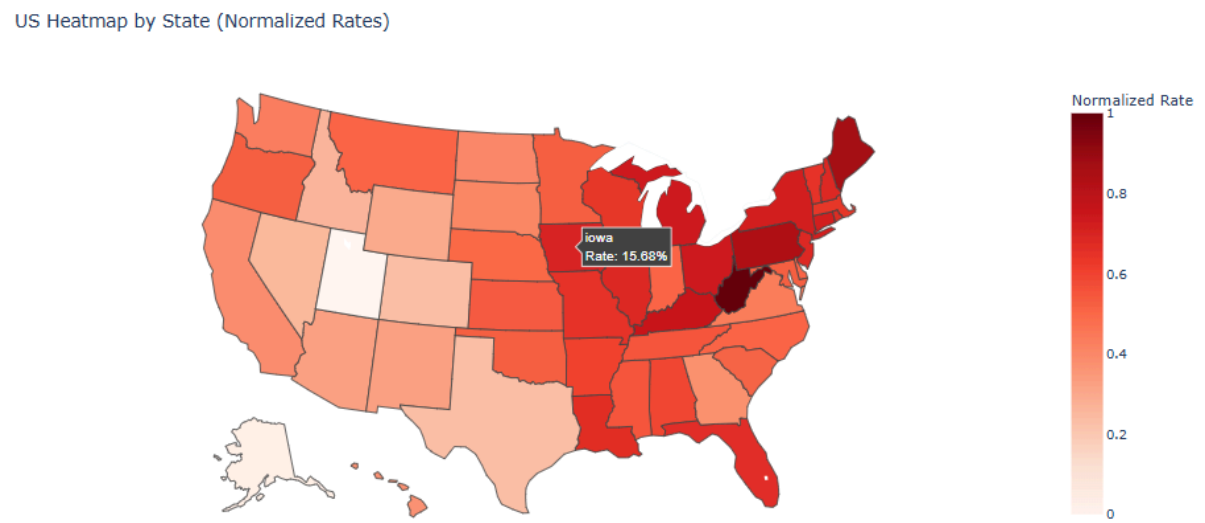
A clickable link directing users to the original source for more details or downloads.

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:

Cancer Type: Colon and Rectum

Year: All Years

Gender: Both Genders

Case Type: All Cases

Race: All Races

Advanced Filters

American Indian/Alaska Native

Asian/Pacific Islander

Black

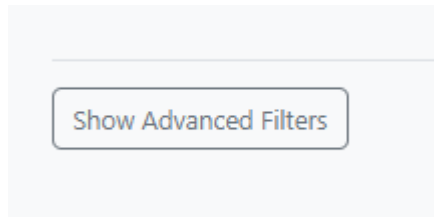
Hispanic

White

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:

Socioeconomic Filters

Unemployment Rate Min (%)

Max (%)

Median Income Min (\$)

Max (\$)

Insurance Coverage Min (%)

Max (%)

Risk Factors

Physical Inactivity Min (%)

Max (%)

Cigarette Use Min (%)

Max (%)

Environmental Factors

Air Quality Index Min

Max

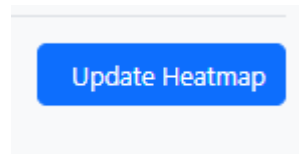
CO₂ Emissions Min

Max

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:

US Cancer Cases	US Average Rate	Highest Rate	Lowest Rate
45389769	13.49%	west virginia: 19.31%	utah: 7.12%

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

Info:

About the Heatmap

This heatmap visualizes cancer cases across the U.S., integrating demographics, socioeconomic factors, behavioral risk factors, and environmental data.

Cancer rates are dynamically calculated based on the selected filters, adjusting the population base accordingly. For example, selecting "Female" and "Asian" uses only the female Asian population for each state.

The normalized rate for each state is calculated using the formula: $\text{Normalized Rate} = (\text{Rate} - \text{Min Rate}) / (\text{Max Rate} - \text{Min Rate})$. This ensures that the color intensity in the heatmap consistently reflects relative differences in cancer rates across states, highlighting areas of concern effectively.

By combining demographic adjustments with socioeconomic, behavioral, and environmental data, the heatmap offers a comprehensive view of health disparities. This visualization helps identify at-risk populations and regions, supporting more targeted and equitable public health initiatives.

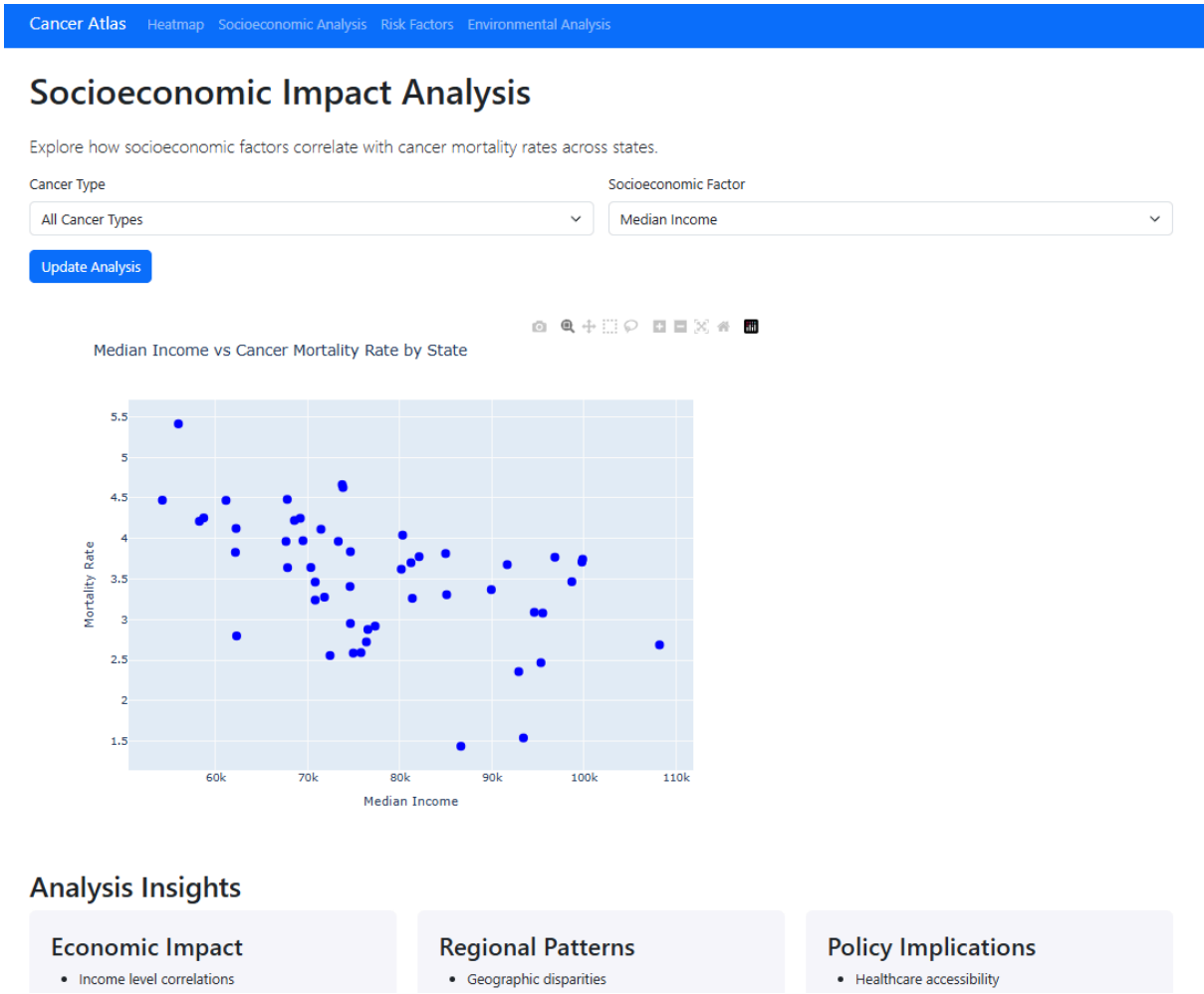
a brief explanation of how the heatmap data is calculated and normalized.

Downloadable Data:



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

Socioeconomic Analysis Page



Filter Controls:

Socioeconomic Impact Analysis

Explore how socioeconomic factors correlate with cancer mortality rates across states.

Cancer Type	Socioeconomic Factor
All Cancer Types	Median Income
<button>Update Analysis</button>	

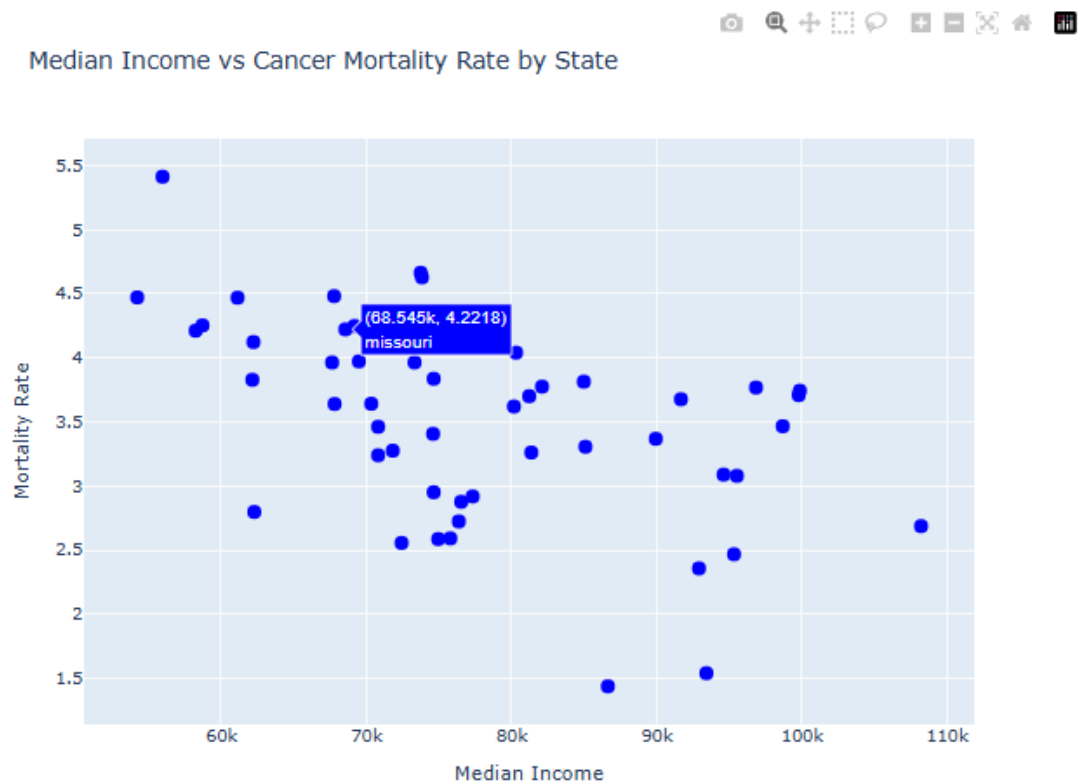
Users can select specific cancer types and a selection of a socioeconomic factor, such as:
Median Income: Correlation between income levels and cancer mortality.
Filters dynamically adjust the scatter plot visualization when submitted.

Update Analysis:

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:

Analysis Insights

Economic Impact

- Income level correlations
- Employment status effects
- Healthcare access patterns

Regional Patterns

- Geographic disparities
- Urban vs. rural differences
- State-level variations

Policy Implications

- Healthcare accessibility
- Economic support programs
- Resource allocation

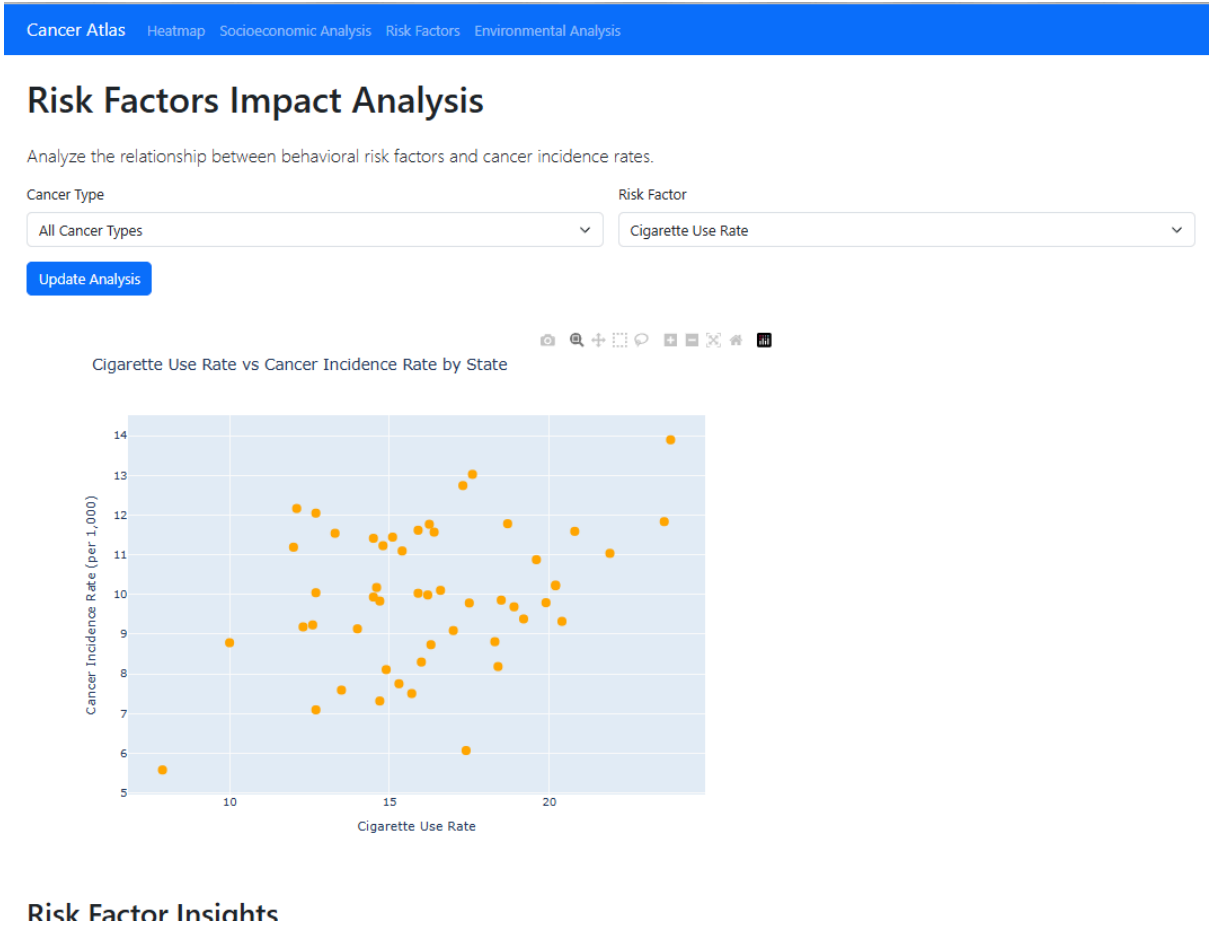
Provides contextual information about how the data can help users understand trends and correlations between socioeconomic factors and cancer outcomes.

Downloadable Data:

Download Data

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

Risk Factors Analysis Page



Filter Controls:

Risk Factors Impact Analysis

Analyze the relationship between behavioral risk factors and cancer incidence rates.

Cancer Type

All Cancer Types

▼

Risk Factor

Cigarette Use Rate

▼

Update Analysis

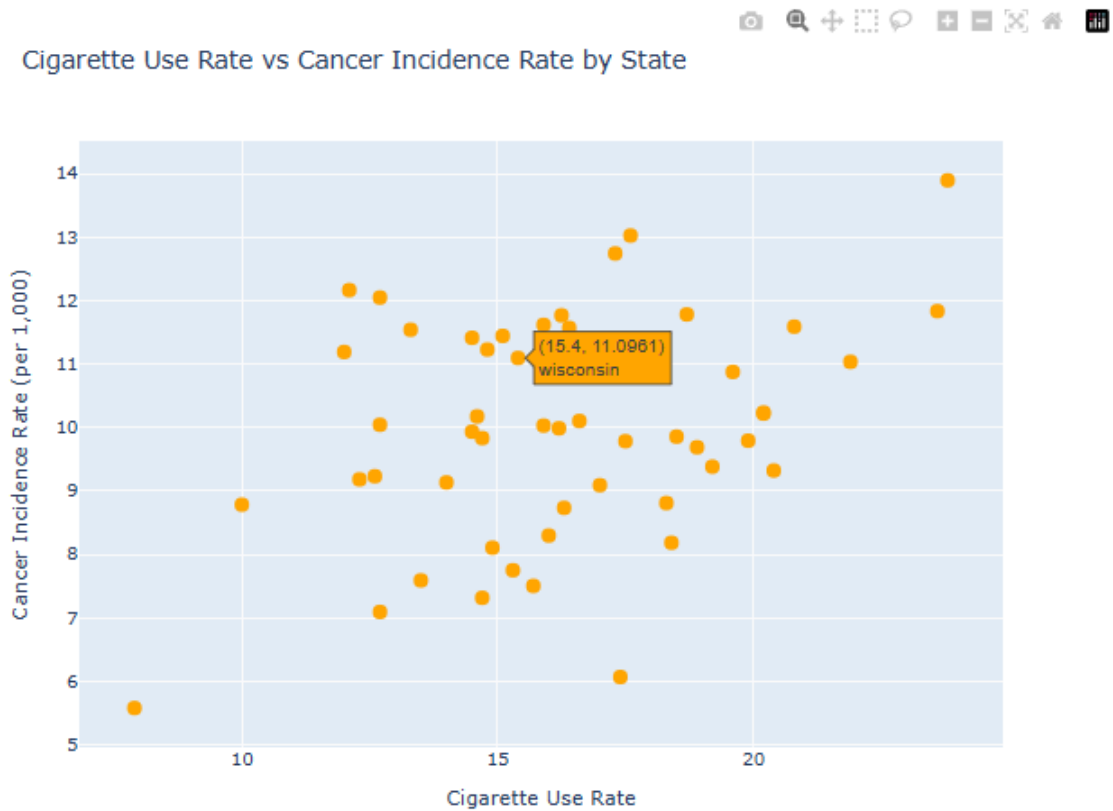
Users can select specific cancer types and a behavioral risk factor, such as: Smoking Rates and Physical Inactivity to Analyze the effect of sedentary lifestyles on cancer outcomes. Filters dynamically adjust the scatter plot visualization when submitted.

Update Analysis:

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:

Risk Factor Insights

Behavioral Patterns	Prevention Strategies
<ul style="list-style-type: none">Smoking prevalence impactPhysical activity correlationRegional behavior variations	<ul style="list-style-type: none">Public health interventionsLifestyle modification programsCommunity education initiatives

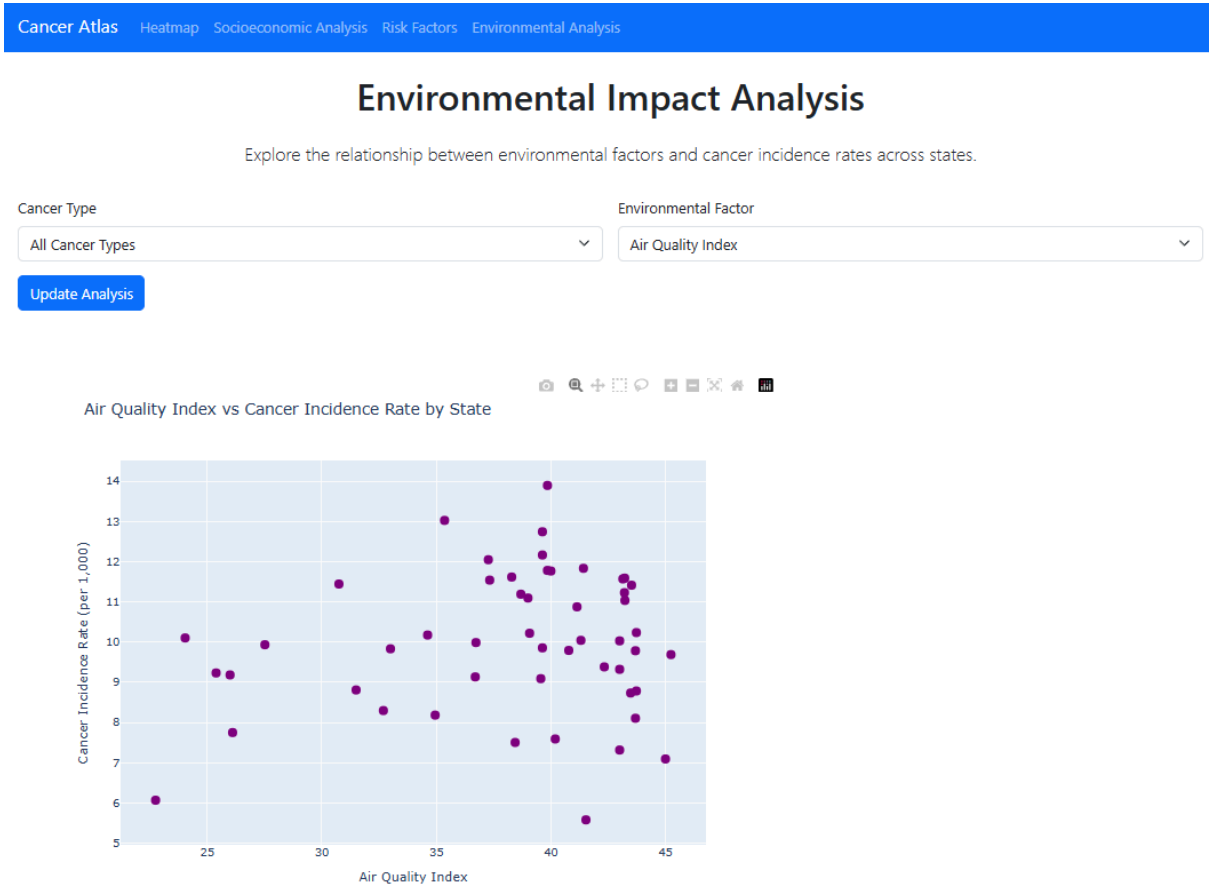
Provides contextual information about how the data can help users understand trends and correlations between risk factors and cancer outcomes.

Downloadable Data:



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

Environmental Analysis Page

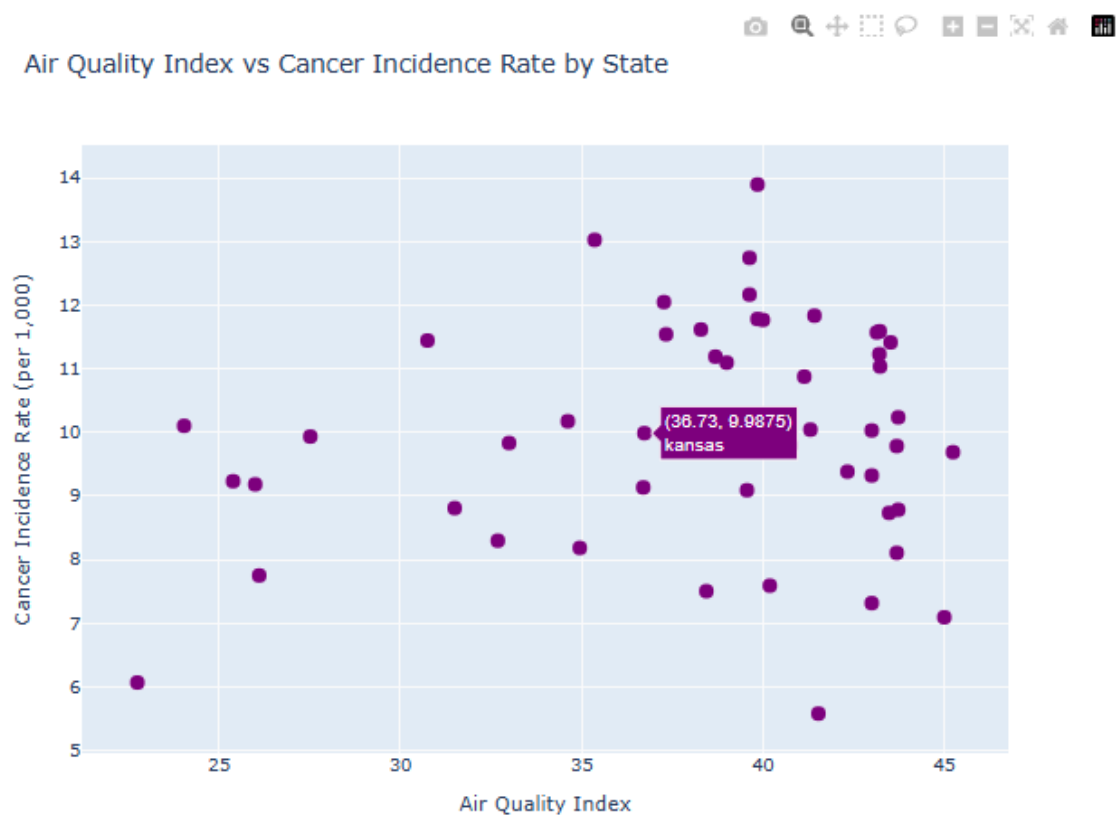


Update Analysis:

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:

Analysis Insights

Key Findings

- Correlation between environmental factors and cancer rates
- Regional variations in environmental impact
- Temporal trends in environmental conditions

Recommendations

- Policy implications for environmental protection
- Public health interventions
- Areas for further research

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) .