US Income Vs Crime Dashboard

Key Insights and Project Summary

This presentation is divided to highlight the four parts of the project:

- Hypothesis
- Introducing the Datasets
- Key Insights
- Features of the Streamlit app

Table of Contents

Hypothesis

Hypothesis

To analyze the relationship between crime and income, we explore the following hypotheses:

H1: Higher-income cities experience lower crime rates.

H2: Areas with higher household incomes correlates with lower crime rates (at a ZIP code level).

H3: Increasing household income over time within a city leads to a reduction in crime.

Introducing the Datasets

Introducing the Datasets

The project uses the three following datasets:

US_CRIME
Monthly crime statistics for the zip codes of 6 major US cities
(2011-2021)
Source: Snowflake

US_INCOME
Yearly household income statistics for the zip codes of multiple US cities
(2011-2021)
Source: Kaggle

LATLON
Latitude and Longitude values for zip codes across the US.
Source: Kaggle

FINAL_CRIME_WITH_LATLON
Contains data on total monthly crime and yearly household income per zip code, and latitude and longitude values.
Final Dataset used in Streamlit

Why Crime and Income Data?

- **Scope:** Data on US ZIP codes were widely available.
- **Predictability:** Crime trends exhibit patterns over time.

Note: Analysis done on years 2018 - 2021

Cities

1. New York
2. Chicago
3. Los Angeles
4. Seattle
5. Houston
6. San Francisco

Introducing the Datasets

Table Definition: FINAL_CRIME_WITH_LATLON

```
create or replace TABLE US_INCOME.PUBLIC.FINAL_CRIME_WITH_LATLON (
    OFFENSE_CATEGORY VARCHAR(16777216),
    CITY VARCHAR(13),
    YEAR1 NUMBER(4,0),
    MONTH1 NUMBER(2,0),
    TOTAL_CRIMES NUMBER(18,0),
    ZIP NUMBER(38,0),
    HOUSEHOLDS NUMBER(38,0),
    HOUSEHOLDS_LESS_THAN_10K NUMBER(38,1),
    HOUSEHOLDS_10K_15K NUMBER(38,1),
    HOUSEHOLDS_15K_25K NUMBER(38,1),
    HOUSEHOLDS_25K_35K NUMBER(38,1),
    HOUSEHOLDS_35K_50K NUMBER(38,1),
    HOUSEHOLDS_50K_75K NUMBER(38,1),
    HOUSEHOLDS_75K_100K NUMBER(38,1),
    HOUSEHOLDS_100K_150K NUMBER(38,1),
    HOUSEHOLDS_150K_200K NUMBER(38,1),
    HOUSEHOLDS_MORE_THAN_200K NUMBER(38,1),
    HOUSEHOLDS_MEDIAN_INCOME NUMBER(38,0),
    HOUSEHOLDS_MEAN_INCOME NUMBER(38,0),
    YEAR NUMBER(38,0),
    LAT FLOAT,
    LNG FLOAT
);
```

Key Insights

Key Insights – H1

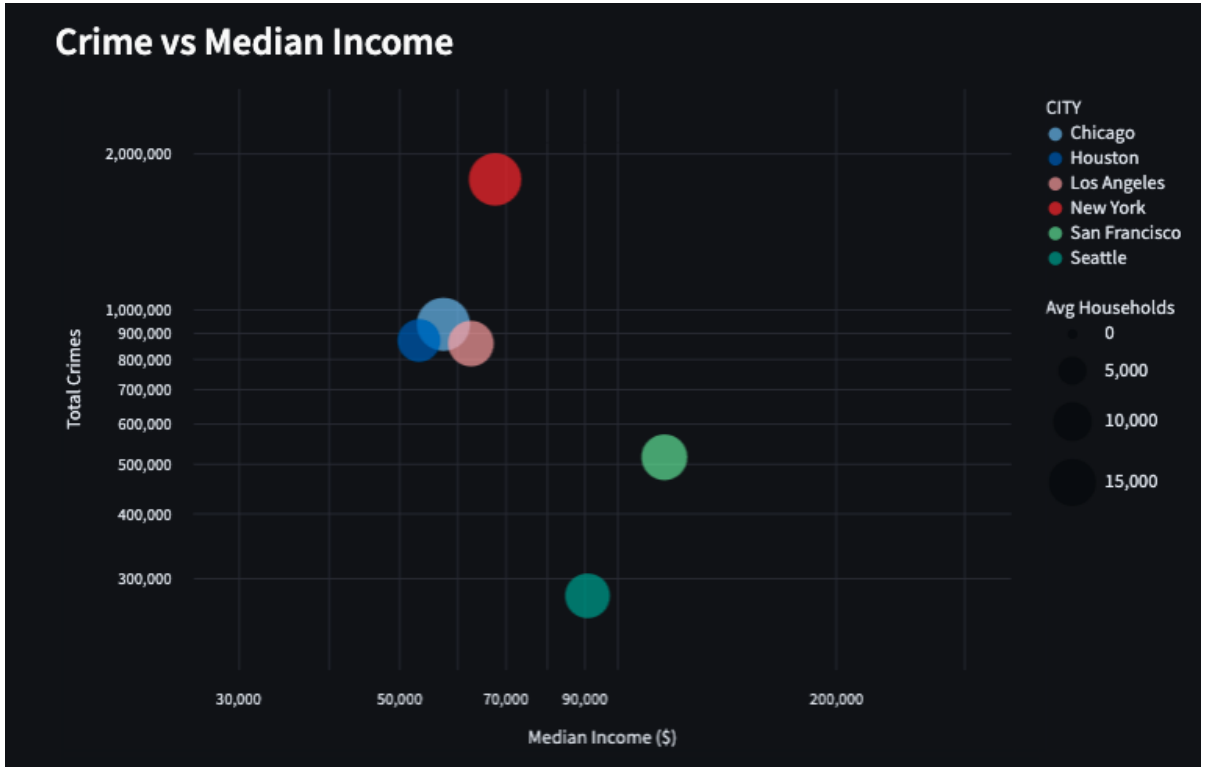
H1: Higher-income cities experience lower crime rates.

Crime vs. Median Income:

- Seattle : Higher income, low crime
- San Francisco : High income, lower crime
- New York : Avg income, high crime
- Chicago : Lower income, higher crime
- Los Angeles : Lower income, higher crime
- Houston : Low income, higher crime

General trend : Suggest H1

Outlier : New York (Average income but higher crime)



Key Insights – H1

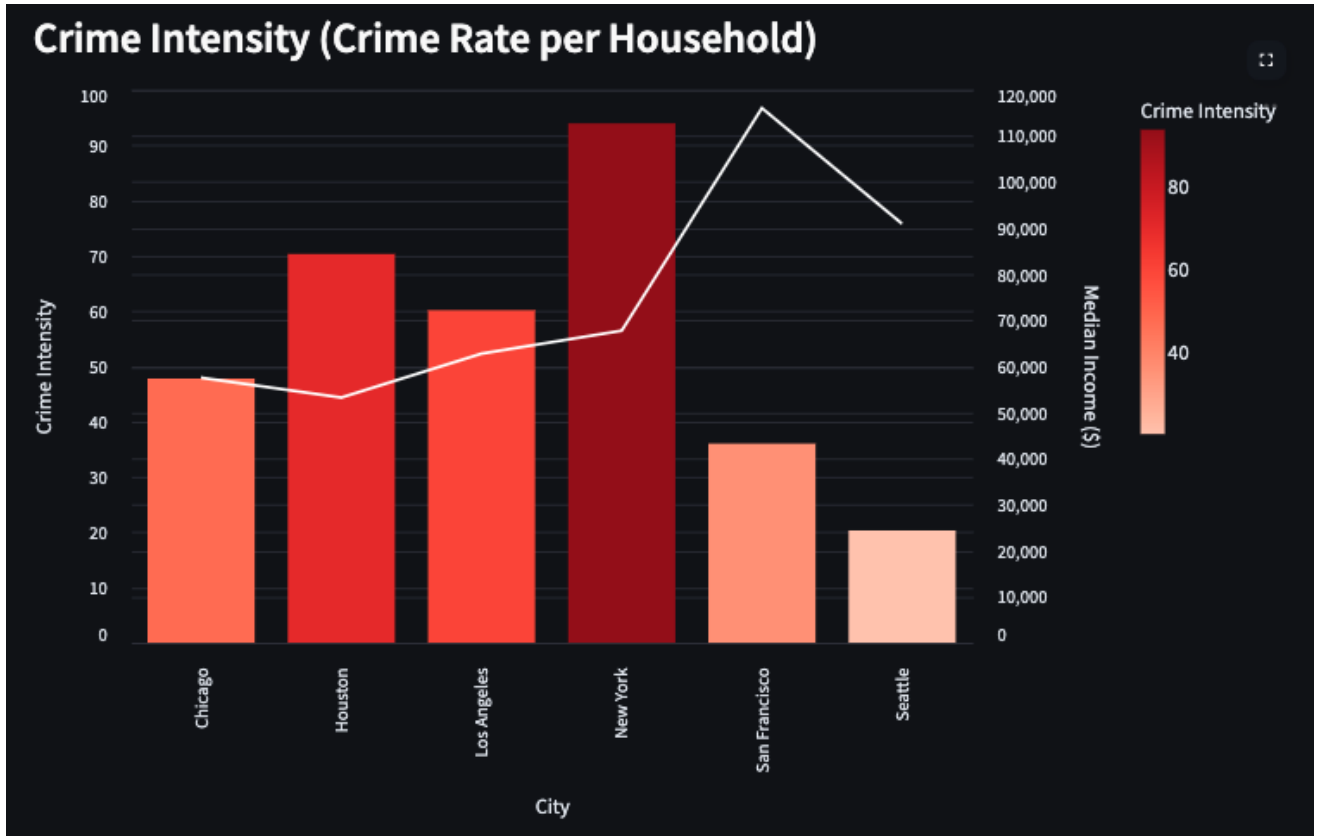
H1: Higher-income cities experience lower crime rates.

Crime Intensity (Crime Rate per Household)

- Seattle : Lowest
- San Francisco : Lower
- New York : Highest
- Chicago : Moderate
- Los Angeles : Moderate
- Houston : Higher

General trend : Suggests H1

Outlier : New York (Average income but higher crime)



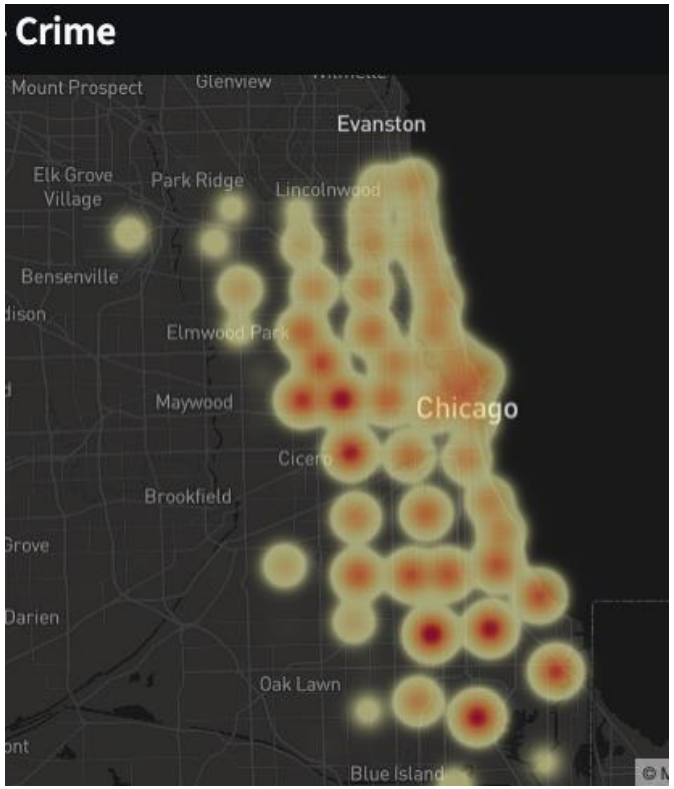
Takeaway: Other factors besides income (eg. population density, law enforcement, socioeconomic factors) may influence crime rates.

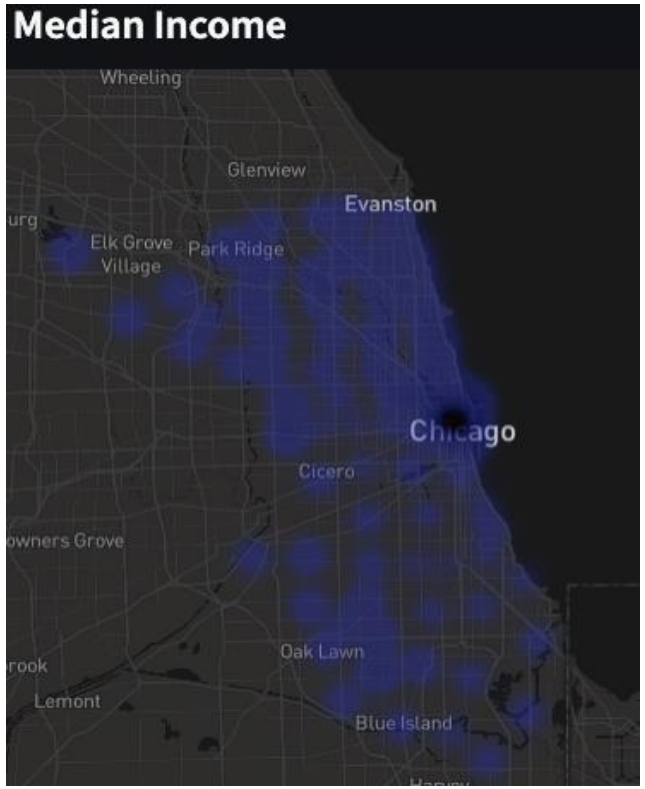
Key Insights – H2

H2: Areas with higher household incomes correlates with lower crime rates (at a ZIP code level).

Crime vs. Median Income Heatmaps
Ex: Chicago

General trend : Suggests H2
Exception : City Centres



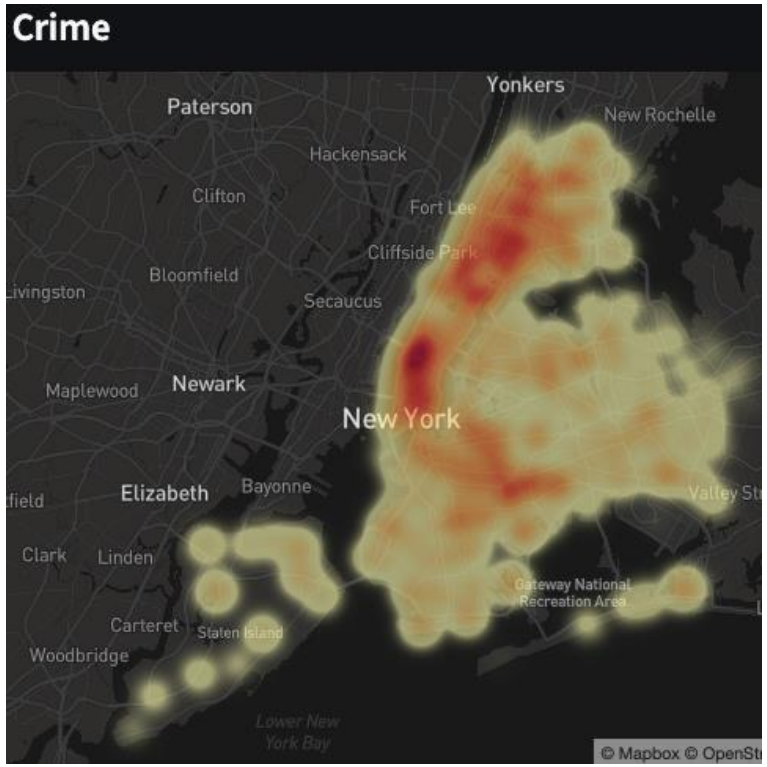


Key Insights – H2 Exception

H2: Areas with higher household incomes correlates with lower crime rates (at a ZIP code level).

Crime vs. Median Income Heatmaps
Ex: New York

Trend at City Centres: High crime
Probable Reason:
Tourism
Commercial activity

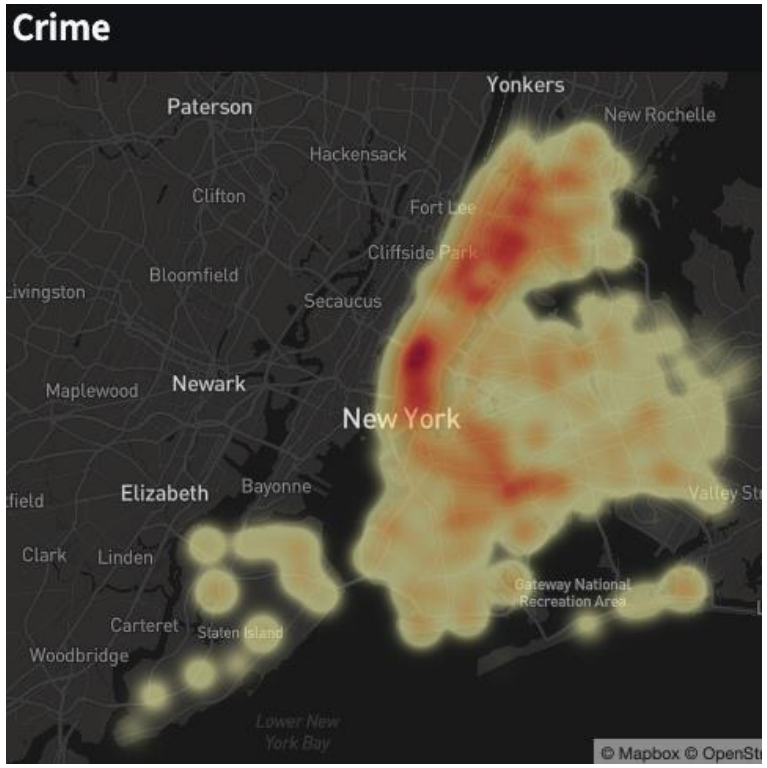


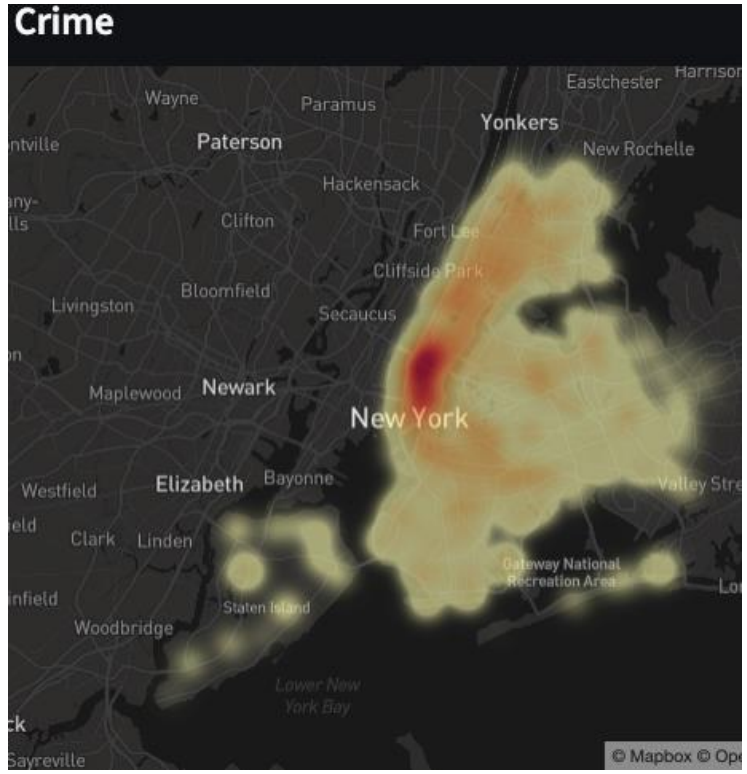
Key Insights – H2 Exception Example (Theft)

H2: Areas with higher household incomes correlates with lower crime rates (at a ZIP code level).

Crime vs. Median Income Heatmaps
Ex: New York

Left: Total crime
Right: Theft





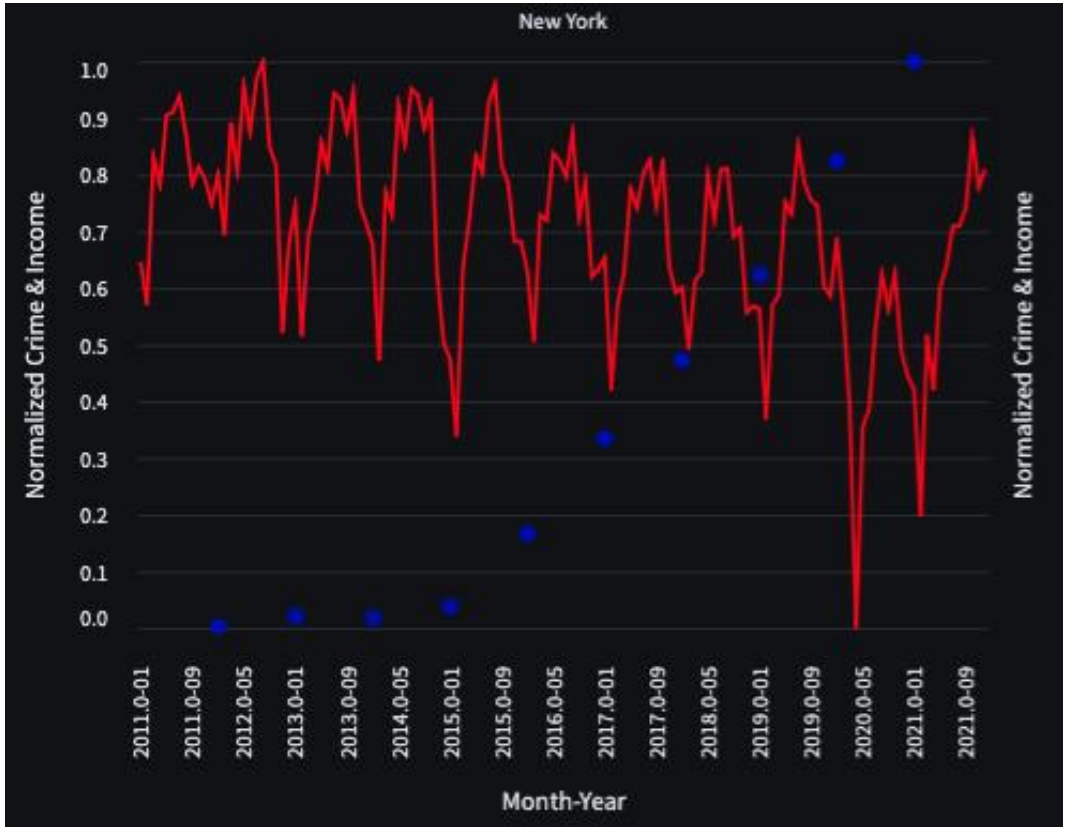
Takeaway: Generally, H2 holds but certain areas deviate due to crimes affected by other factors like population density, law enforcement, socioeconomic factors.

Key Insights – H3

H3: Increasing household income over time within a city leads to a reduction in crime.

Time Series Crime & Income Trends
Ex: New York

General Trend: Crime fluctuates independently of income



Takeaway: Increasing income alone does not reduce crime. Crime trends are more complex; other social, economic, and policy factors likely influence crime rates (Ex: COVID in 2020).

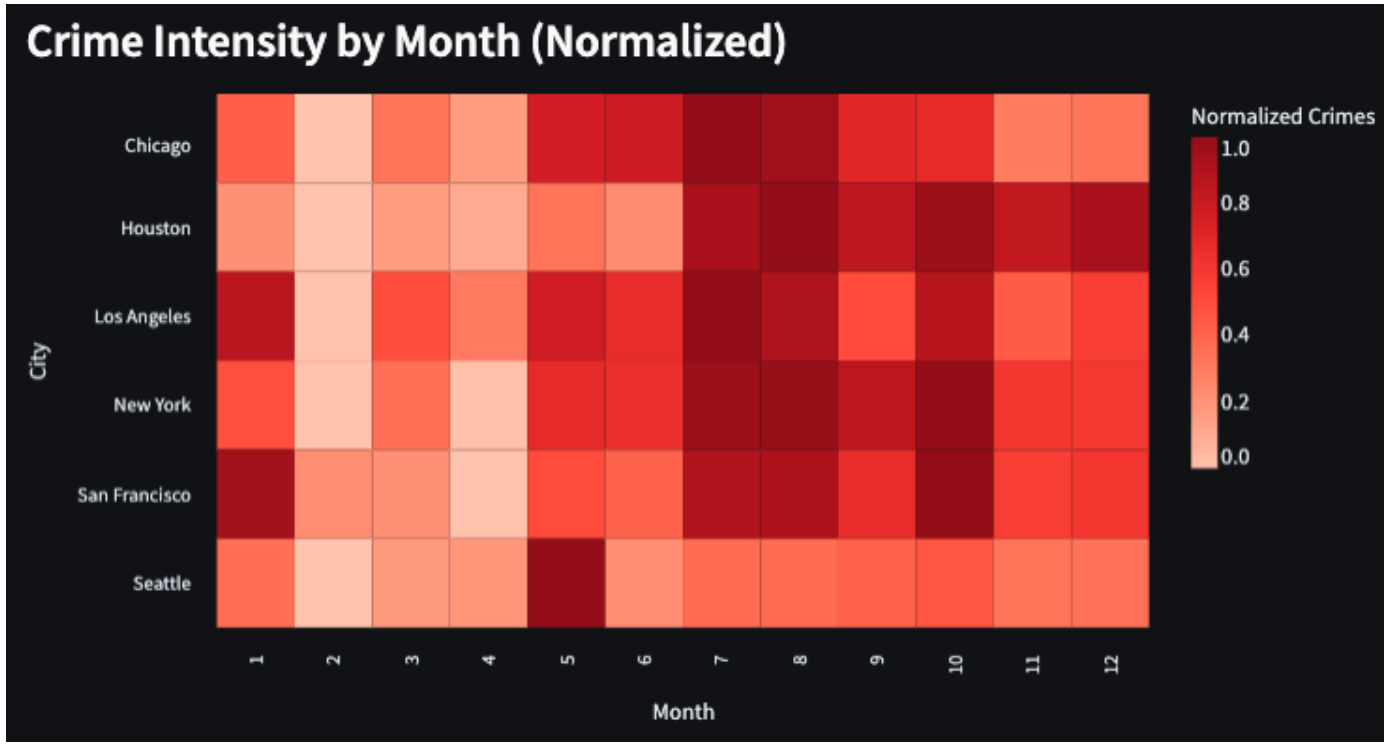
Other Insights

Other Insights

Why does crime increase in Summer and drop in February?

Possible Reasons:

- More outdoor activities in Summer
- Summer vacation - more unsupervised youth
- Longer daylight hours
- Tourism peaks (seasonal crime spikes)
- Winter: shorter days, colder weather



Potential:

Is crime dependent on temperature, weather, and activity levels (public holidays, summer vacation, tourism)?

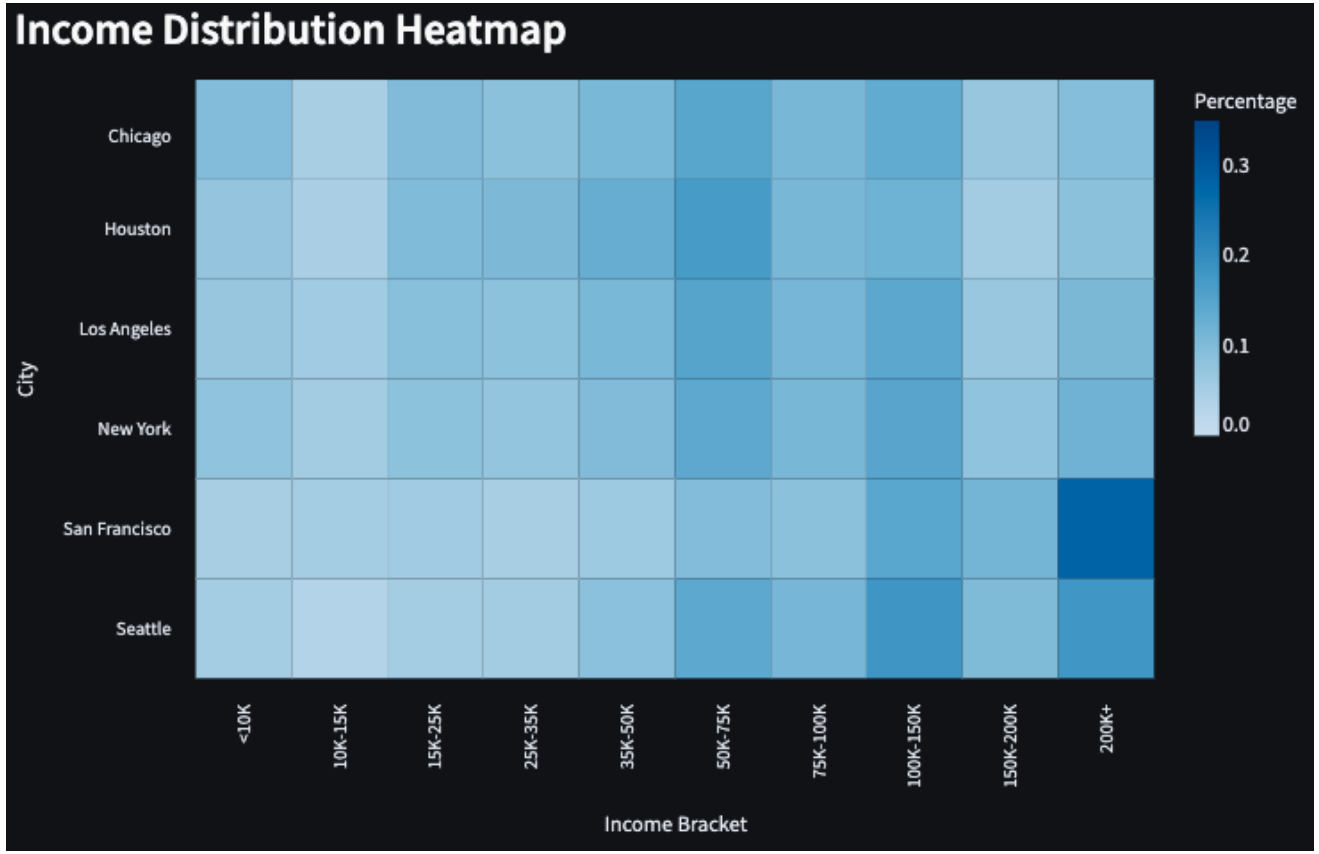
Do different crimes peak during different seasons?

Other Insights

Does higher income mean ‘wealthy’ in the city?

Key Insights:

- City’s income distribution should be considered alongside cost of living.
- The definition of “middle-class” varies greatly across cities.
- Data skews due to migration patterns and affordability crisis.



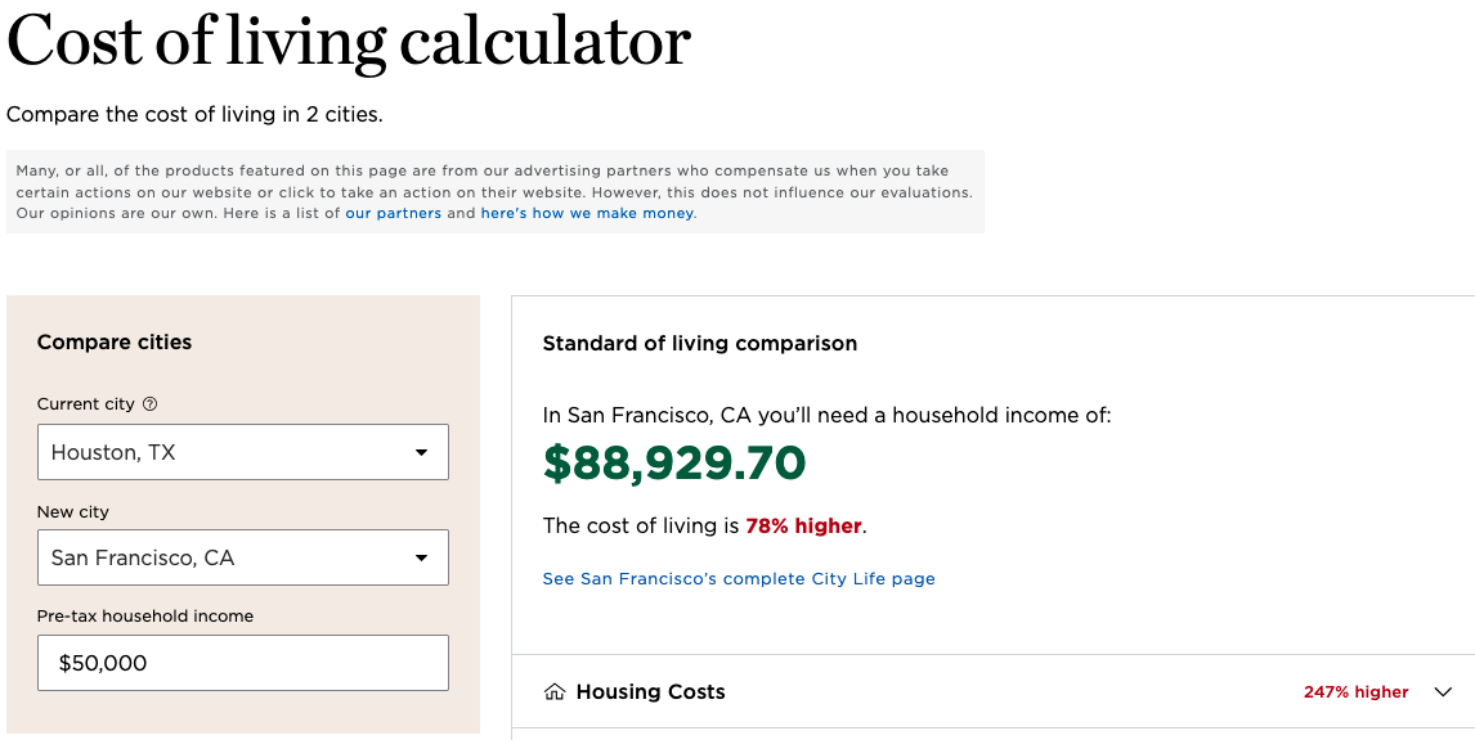
Potential:

How would income brackets change if adjusted for cost of living?

Other Insights

Does higher income mean ‘wealthy’ in the city?

Cost of living example comparison



Cost of living calculator

Compare the cost of living in 2 cities.

Many, or all, of the products featured on this page are from our advertising partners who compensate us when you take certain actions on our website or click to take an action on their website. However, this does not influence our evaluations. Our opinions are our own. Here is a list of our partners and here's how we make money.

Compare cities

Current city

Houston, TX

New city

San Francisco, CA

Pre-tax household income

$50,000

Standard of living comparison

In San Francisco, CA you'll need a household income of:

$88,929.70

The cost of living is 78% higher.

See San Francisco's complete City Life page

Housing Costs 247% higher

Crime Prediction

Predict future monthly crime trend for 4 cities for 2022 *(not a part of main streamlit app).*

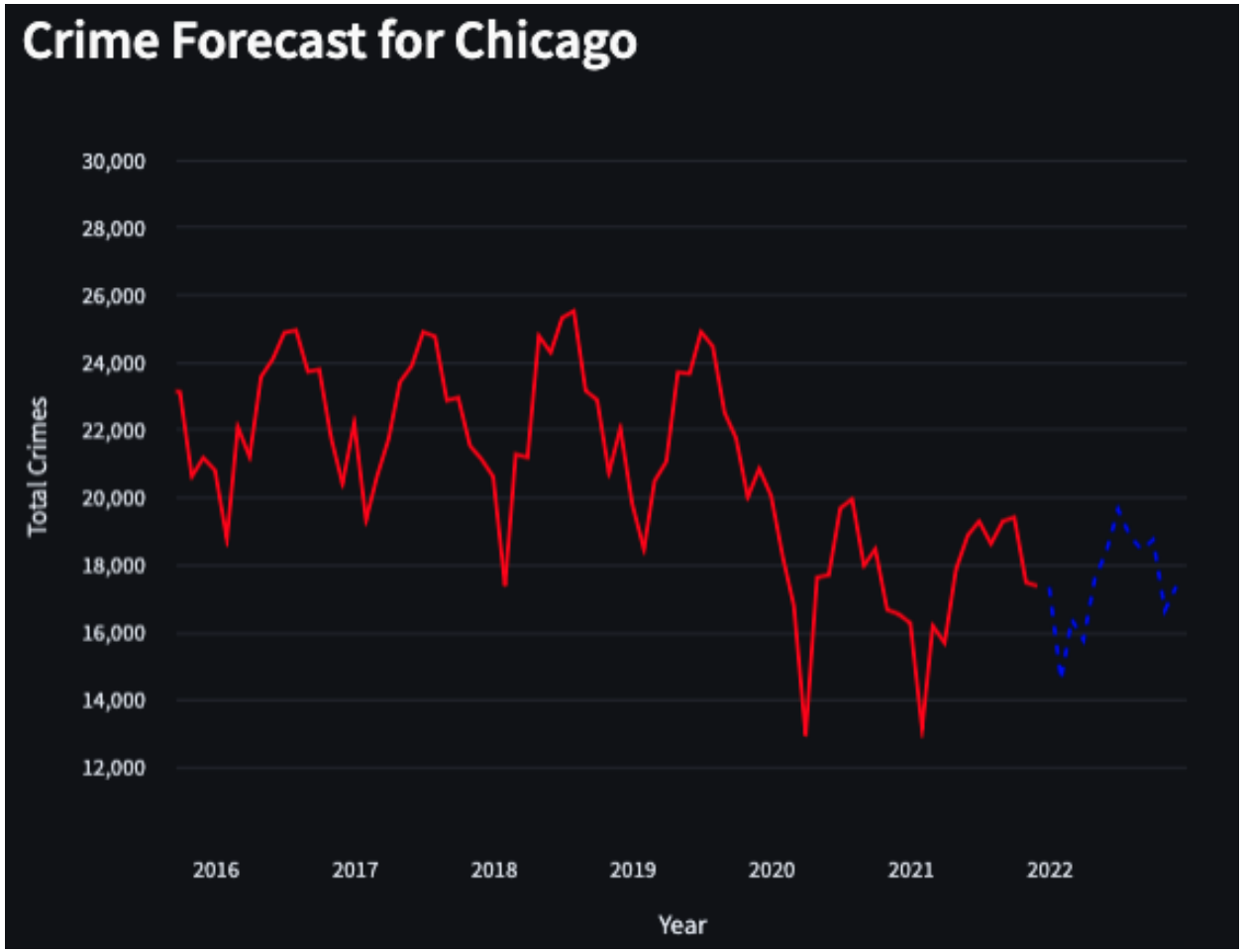
Methodology:

- Used Snowflake AI & ML Studio - Time Series Forecasting
- Model considers seasonality, trends in historical crime data
- Trained data: 2011-2021 (monthly – red solid)
- Predicted: 2022 (monthly – blue dashed)

Key Features:

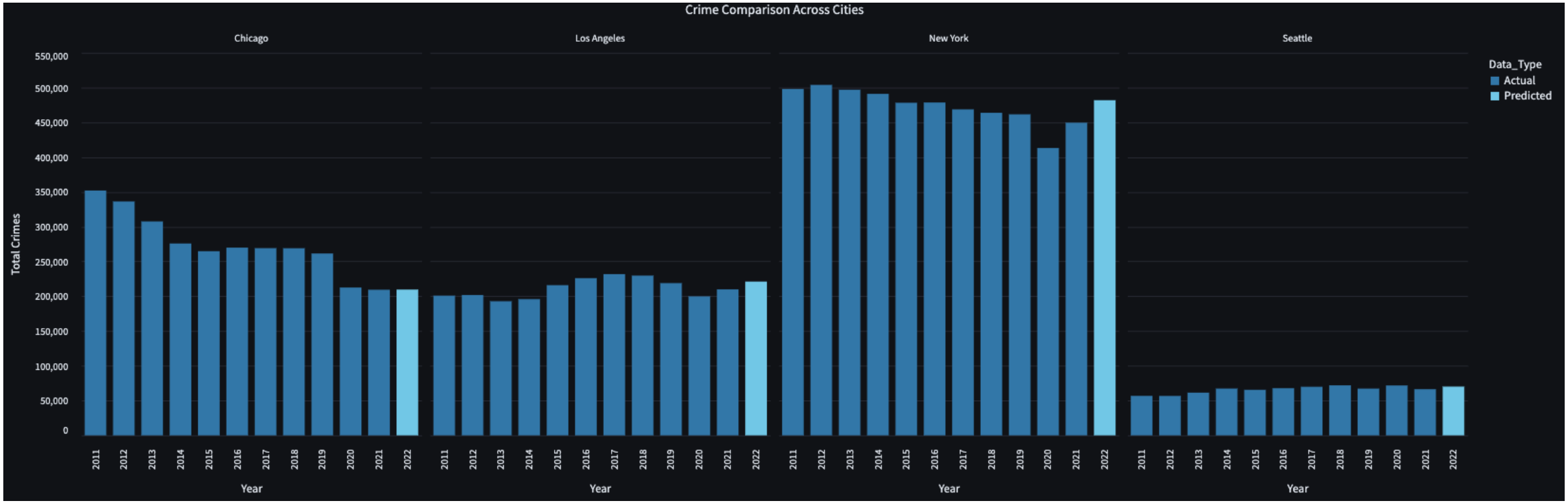
- Automated feature engineering
- City Specific Forecasts

Potential:
Adding more features (temperature, tourism fluctuations, income)
Why Predict? Enable targeted actions to reduce crime



Crime Prediction

Predict future monthly crime trend for 4 cities for 2022.

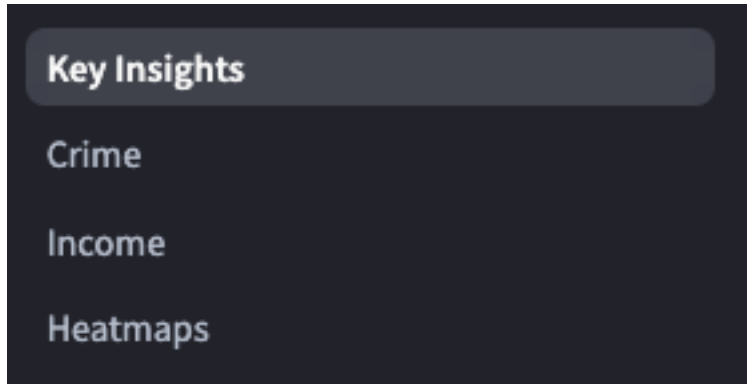


Features of Streamlit App

Key Features of App

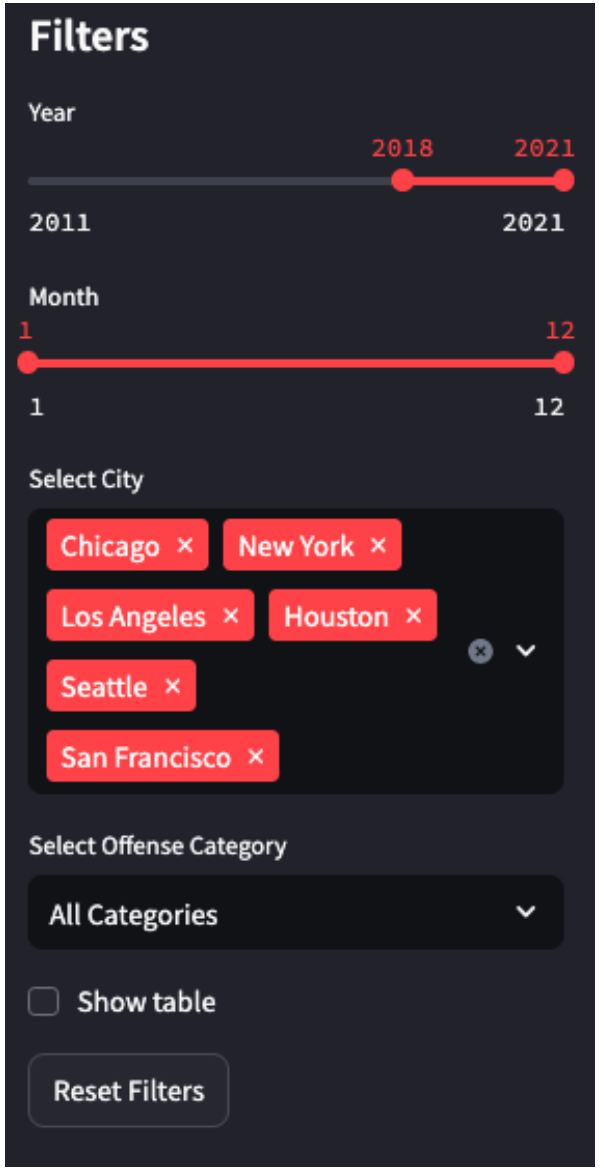
Multipage Navigation

1. **Key Insights**
   Analyze relationship between crime and income
2. **Crime**
   Explore yearly and monthly crime patterns
3. **Income**
   Compare income distributions between cities
4. **Heatmaps**
   Visualize crime and income density within each city



Dynamic Sidebar Filters

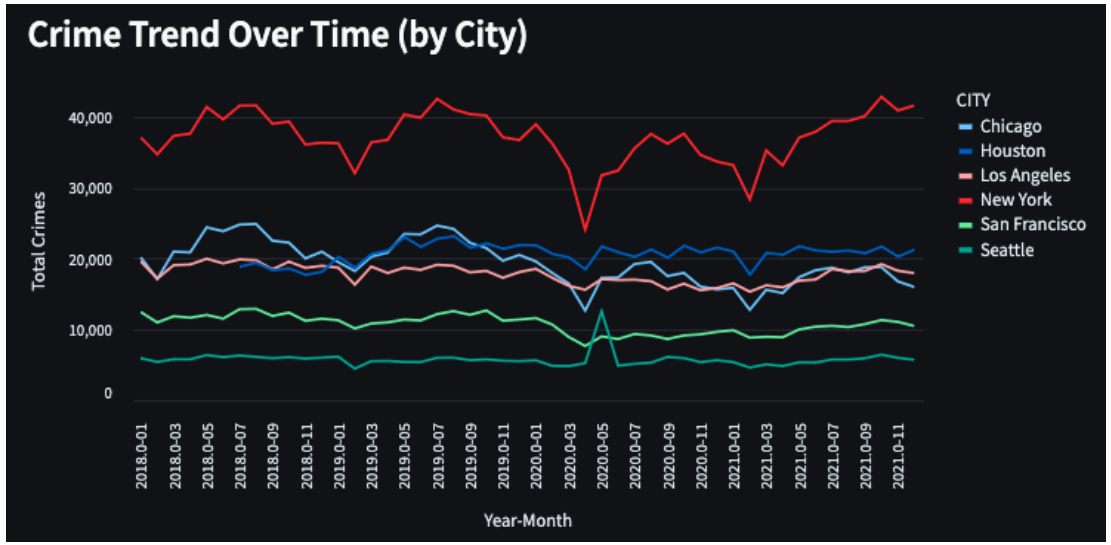
1. **Year Range**
   Default: 2018 to 2021
2. **Month Range**
   Default: 1 to 12
3. **City Selection (Multiselect)**
   Compare across multiple cities
4. **Offense Type Dropdown**
   Analyze crime for each category
5. **Show Table**
   Choose when to show selected table
6. **Reset Filters**



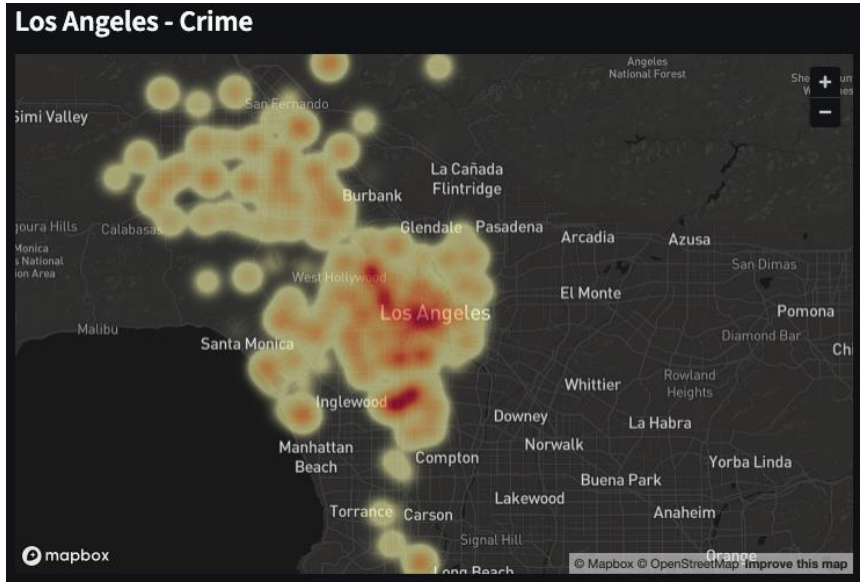
Key Features of App

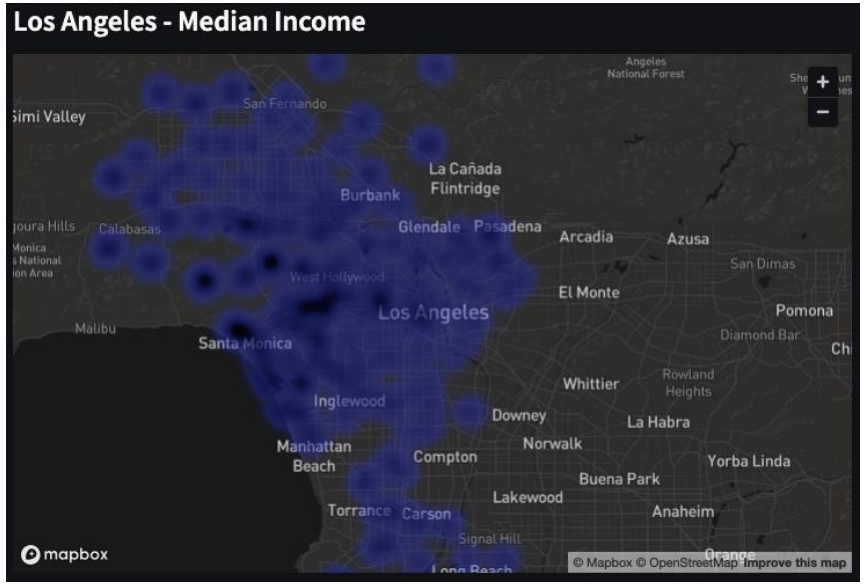
Interactive Data Visualizations

1. **Altair-Powered Charts**
   Analyze crime and income trends and comparisons. Explore trend over time.
2. **City-Specific Heatmaps**
   Displays geographic distributions of crime and income




```
col1, col2 = st.columns(2)
with col1:
    st.subheader("Crime Trend Over Time (by City)")
    st.altair_chart(chart1, use_container_width=True)
```





Key Features of App

Seamless Data Integration

Connected to Snowflake

Real-time access of datasets through snowflake.

```python
st.set_page_config(
    page_title="US Income vs Crime Dashboard",
    layout="wide"
)
st.markdown("<h1 style='text-align: center;'>US Income vs Crime Dashboard</h1>", unsafe_allow_html=True)


@st.cache_resource
def create_session():
    return Session.builder.configs(st.secrets.snowflake).create()

session = create_session()


# Connect to Snowflake


sql_query = "SELECT * FROM US_INCOME.PUBLIC.FINAL_CRIME_WITH_LATLON"
df = session.sql(sql_query).to_pandas()
```

Streamlit & Docker Deployment

Fully containerized using Docker, ensuring easy setup and deployment.

Questions?
Suggestions?