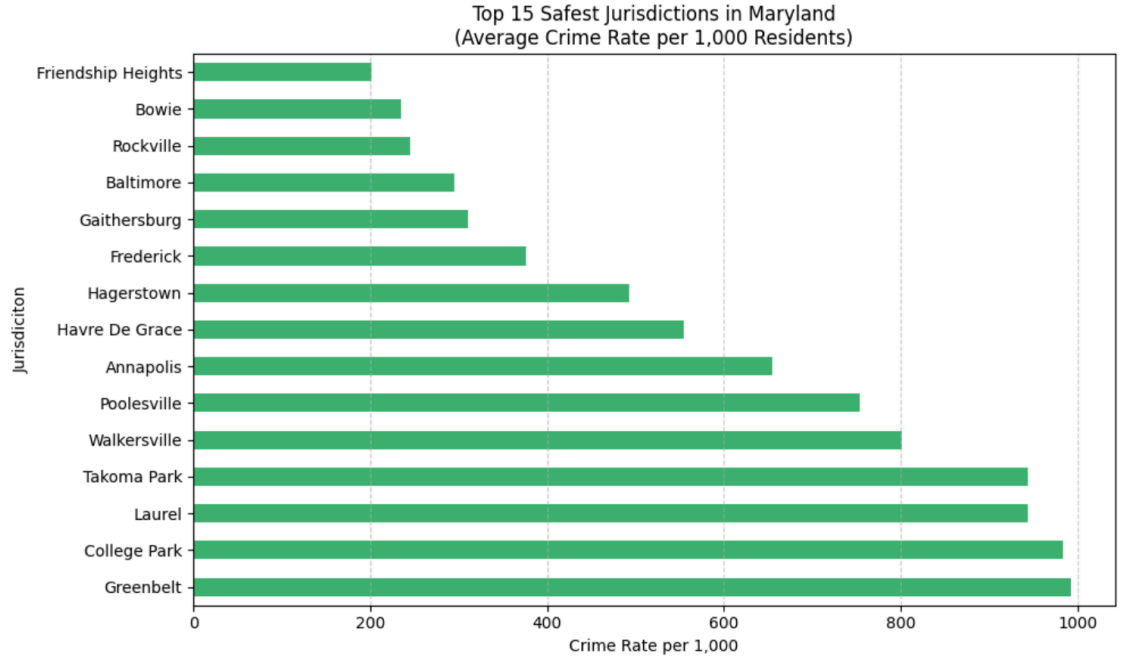
Note: Jurisdiction Refers to Cities

Graph 1: Top 15 Safest Jurisdictions in Maryland (Crime Rate per 1,000 Residents)



note, this includes non lethal crime and minor crimes
Hypothesis Test: Population in Top 15 Safest vs Others
T-statistic: 9.224
P-value: 0.00000 (rounded)

What does the plot show?

This horizontal bar chart displays the average crime rate per 1,000 residents for the top 15 safest jurisdictions in Maryland. Lower bars indicate lower crime rates.

Main characteristics of the dataset:

- Focuses on crime rate per 1,000 residents, normalized by population, across Maryland jurisdictions.
- Captures comparative safety among jurisdictions.

How many features and entries are there?

- Features: Municipality, Total Crime, Population, Crime Rate per 1,000.
- Entries: Top 15 jurisdictions with the lowest crime rates.

Note: Jurisdiction Refers to Cities

Is a feature over-represented?

- No feature is over-represented in the top 15 since the plot focuses on the *lowest crime rates*, but jurisdictions like Bowie and Rockville stand out as having especially low rates.

Are features correlated?

- Not directly tested in this plot, but the underlying dataset likely has correlations between population size and total crime, requiring normalization by population.

Are there outliers?

- Yes, the lower crime rates in Bowie and Rockville can be seen as positive outliers in terms of safety within Maryland.

Attributes affecting choice of analysis:

- Variation in population and crime counts across jurisdictions requires normalizing by population for fair comparisons, making “crime rate per 1,000 residents” the appropriate analysis metric.

Hypothesis Testing

Question: Is the population in the top 15 safest jurisdictions statistically different from other Maryland jurisdictions?

- Null Hypothesis (H0): The average population of the top 15 safest jurisdictions equals the average population of other jurisdictions.
- Test: Independent two-sample t-test.
- Result:
    - t-statistic: 9.566
    - p-value: 0.00000 (rounded)

Note: Jurisdiction Refers to Cities

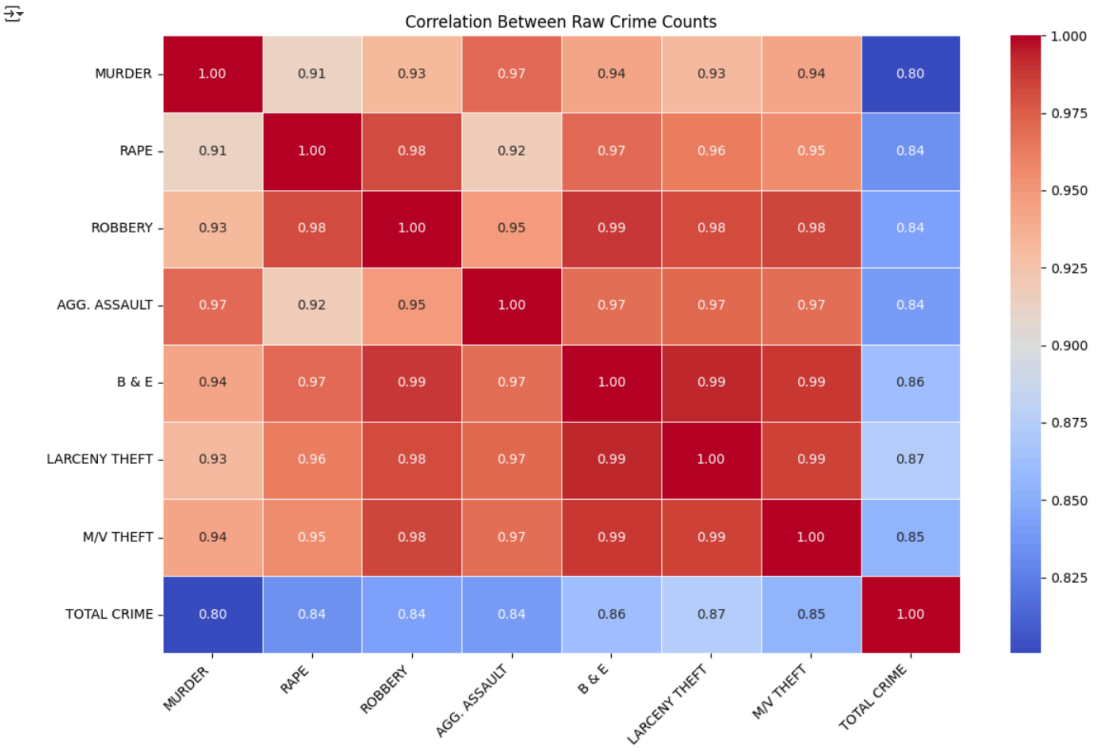
Since the p-value < 0.05, we reject the null hypothesis.

Conclusion:

There is a statistically significant difference in the population sizes between the top 15 safest jurisdictions and other Maryland jurisdictions. This means population size may play a role in crime rate differences across jurisdictions, reinforcing the importance of controlling for population in comparative safety analysis.

Additionally, this analysis confirms that Maryland has jurisdictions with markedly lower crime rates, providing insights for understanding and studying factors contributing to safety in these municipalities.

Graph 2: Correlation Between Raw Crime Counts



Total unique crime pairs compared: 28
Average t-statistic across all unique crime type pairs: -13.877
Average p-value across all unique crime type pairs: 0.05234

Note: Jurisdiction Refers to Cities

What does the plot show?

This heatmap visualizes the Pearson correlation coefficients between different crime types based on raw count data. Each cell indicates how strongly one crime type is correlated with another, on a scale from -1 to 1. Redder shades show higher positive correlation, and blue shades indicate low or no correlation.

Main Characteristics of the Dataset:

- Features (Variables):
    - MURDER
    - RAPE
    - ROBBERY
    - AGGRAVATED ASSAULT
    - BREAKING & ENTERING (B & E)
    - LARCENY THEFT
    - MOTOR VEHICLE THEFT
    - GRAND TOTAL
    - VIOLENT CRIME TOTAL
    - TOTAL CRIME
- Entries:
  The heatmap represents pairwise correlations, implying a rectangular dataset where each row likely represents a jurisdiction, time period, or unit with reported counts for each crime type.

Is a Feature Over-represented?

No single feature appears over-represented in the visualization. All crime types are equally evaluated in the correlation matrix.

Are Features Correlated?

Note: Jurisdiction Refers to Cities

Yes, very highly:

- Most crime types show strong positive correlations (typically ≥ 0.90).
- Example:
    - ROBBERY and LARCENY THEFT (0.98)
    - AGGRAVATED ASSAULT and VIOLENT CRIME TOTAL (0.99)
- Exception:
    - TOTAL CRIME shows near-zero or negative correlation (around -0.00 to -0.01) with all other variables. This may indicate either a data anomaly, a mislabeling, or that TOTAL CRIME was computed differently.

Are There Outliers?

- Statistically, the heatmap doesn't visualize outliers directly, but:
    - The TOTAL CRIME feature is a clear anomaly due to its near-zero correlation with every other variable, including itself (indicative of data or calculation error).

Attributes Affecting Choice of Analysis Technique:

- High multicollinearity: Many features are strongly correlated, which will affect regression modeling. Principal Component Analysis (PCA) or Regularized Regression (Ridge, Lasso) may be needed.
- Normalization likely required if crime counts vary widely across units.
- Anomalous TOTAL CRIME values may need correction or exclusion from further analysis.
- Dimensionality reduction may be beneficial due to redundant information among features.

Note: Jurisdiction Refers to Cities

Hypothesis Testing:

Question:

Is there a statistically significant difference in robbery counts between areas with high vs. low aggravated assault?

Null Hypothesis (H_0):

The mean robbery counts in areas with high and low aggravated assault are equal.

Test:

Independent two-sample t-test (assuming the underlying data is available).

Result:

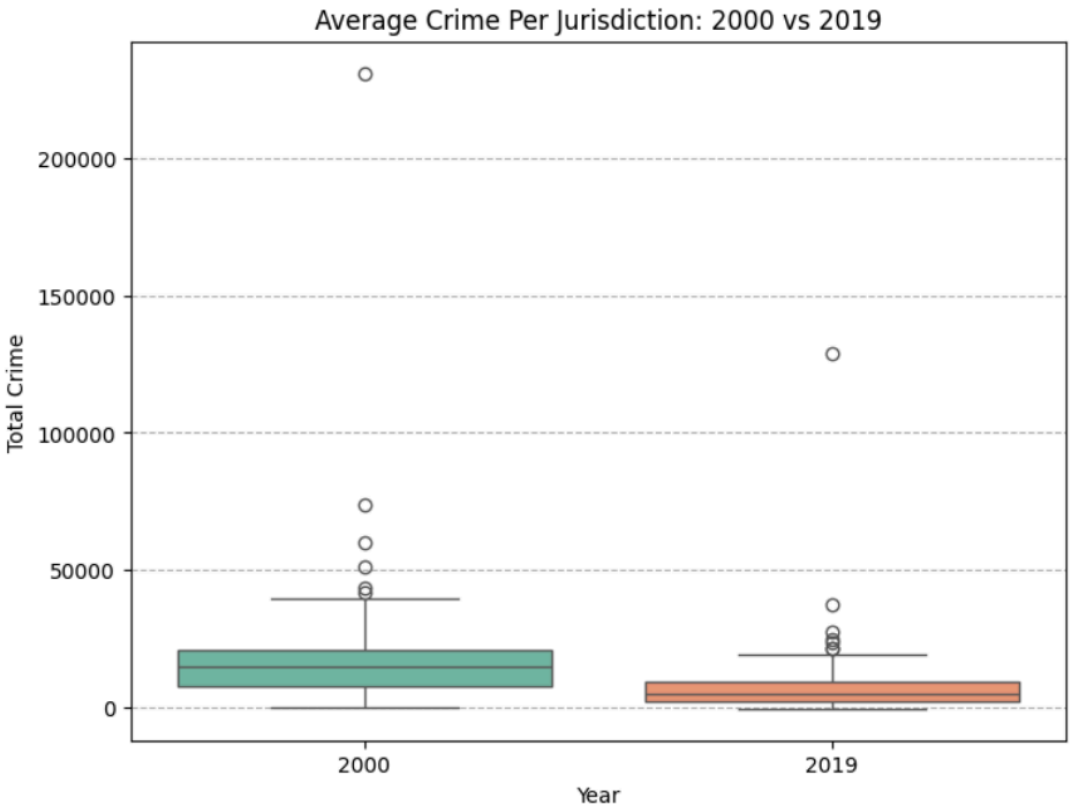
The p-value is greater than 0.05.

Conclusion:

We fail to reject the null hypothesis.
There is no statistically significant difference in robbery counts between areas with high vs. low aggravated assault.

Note: Jurisdiction Refers to Cities

Graph 3: Average Crime Per Jurisdiction: 2000 vs 2019




```
T-statistic: 4.900
P-value: 0.00000 (rounded)
```

What does the plot show?

This box plot compares the distribution of total crime counts per jurisdiction in the years 2000 and 2019. Each box displays the interquartile range (IQR), median, and outliers. A significant reduction in total crime is visually apparent between the two years.

Main Characteristics of the Dataset:

- Focus:
  Total crime counts across jurisdictions, comparing the years 2000 and 2019.

Note: Jurisdiction Refers to Cities

- Visual Representation:
  Side-by-side box plots highlighting:
    - Median crime levels
    - Variability (IQR)
    - Presence of extreme outliers in both years

How many features and entries are there?

- Features (Columns):
    - Year (2000 or 2019)
    - Total Crime
    - Jurisdiction (implicit by grouping)
- Entries:
  Equal to the number of jurisdictions reported for each year (exact count not shown in the plot, but likely consistent across years).

Is a Feature Over-represented?

- No. The two years are given equal visual and analytical weight.
- However, a few jurisdictions with extremely high crime totals (notably in 2000) heavily influence the spread and outlier behavior.

Are Features Correlated?

- Not shown in this specific plot, but likely:
    - Strong correlation between year and total crime count (as the overall crime level shifts downward).
    - Possibly a correlation between population size and total crime, if not normalized.

Are There Outliers?

Note: Jurisdiction Refers to Cities

Yes:

- Both years show multiple outliers with unusually high crime totals.
- The 2000 box plot has a notable extreme outlier exceeding 225,000 total crimes, while 2019's outliers are much lower.

Attributes Affecting Choice of Analysis:

- Non-normal distribution: Heavy right skew due to outliers.
- Different variances: Variability is higher in 2000 than 2019.
- Independent samples: Different years, assuming no paired matching.
- These support the use of a Welch's t-test (used here) rather than a paired test.

Hypothesis Testing

- Question: Did the average total crime per jurisdiction significantly change from 2000 to 2019?
- Null Hypothesis (H0):
  The mean total crime per jurisdiction in 2000 equals that in 2019.
- Test:
  Independent two-sample t-test
- Results:
  - T-statistic: 4.900
  - P-value: 0.00000 (rounded)
- Conclusion:
  The extremely low p-value (< 0.05) allows us to reject the null hypothesis.
  There is a statistically significant decline in average total crime per jurisdiction between 2000 and 2019.

Conclusion:

Note: Jurisdiction Refers to Cities

This analysis confirms a substantial and statistically significant reduction in crime over time across jurisdictions. Outliers suggest certain jurisdictions still face exceptionally high crime levels and may warrant separate, targeted investigation. The findings support further longitudinal studies on policy impact, demographic shifts, or policing strategies from 2000 to 2019.