



“AMERICAN WEREWOLVES IN AMERICA”

TEAM PROGRESSBAR

DAVE KINGSLEY, ERIN LAMPA, JUSTIN MILLER, RICKY RAVIN, & KATIE WRIGHT

MOTIVATION AND SUMMARY

- Popular culture has tales of people turning into werewolves during full moons, wreaking havoc on cities through the night
- This led us to ask if there is a noticeable change in crime on days when there is a full moon?
- Our alternate hypothesis (H_a) is, “if werewolves exist, then average crime will be significantly higher during a full moon”.
- Our null hypothesis (H_0) is, “if werewolves don’t exist, then there will be no significant change in average crime during a full moon”.
- Our question to ask ... “Is there an increase in crime during full moons?”
- Looking at 5 major cities in the United States, there is no significant increase (or decrease) in crime when the moon is full or near full.

QUESTION & DATA

- “Is there an increase in crime during full moons?”
- To determine an increase in crime during full moons, we need the total number of crime reports for each day in the period between 1/1/2015 and 12/31/2017 and a chart of moon illumination by date for the same period.
- We found crime data on 5 cities: Denver, CO; Minneapolis, MN; New York City, NY; Portland, OR; and San Francisco, CA. This data was found from the individual city governments’ websites.
- We compiled the moon data from The United States Naval Observatory’s website.

DATA CLEANUP & EXPLORATION

- Each city's data provided unique challenges, but all the data originated from comma-separated values (csv) files.
- Data cleanup was performed in Jupyter Notebooks.
 - The cleanup process included:
 - Removing unnecessary columns
 - Changing date formats
 - Removing cells with blank data
 - Calculating total daily crime and average daily crime

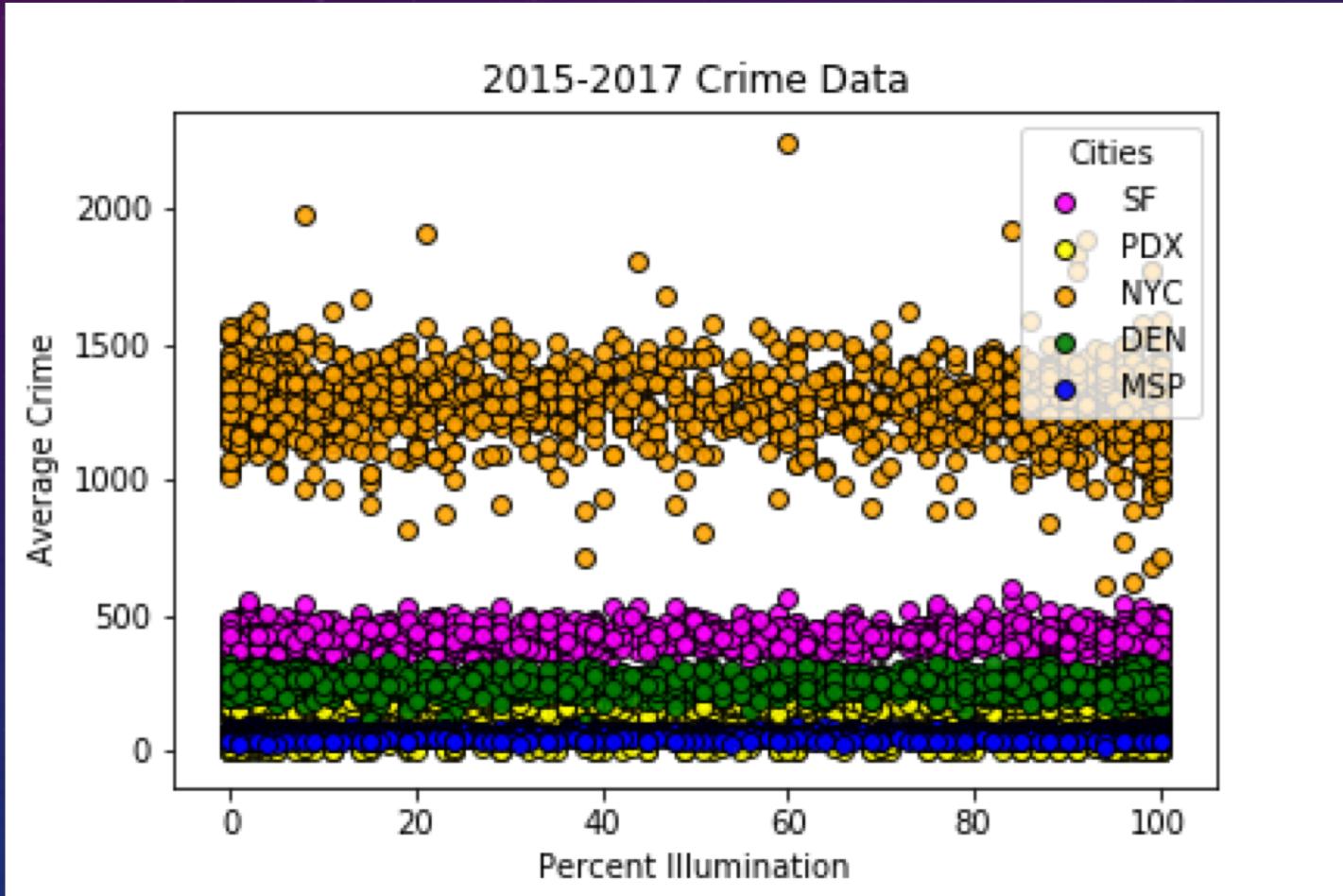
DATA CLEANUP & EXPLORATION CONTINUED:

- Exploration Pros:
 - Crime databases were user friendly
 - Csv files were readily available to download
 - Csv files were mostly complete, well organized and had a generally consistency in formatting across cities
- Exploration Challenges:
 - Determining how to categorize the data.
 - The Moon's illumination cycle is not linear requiring us to divided the frequency of illumination percentages into 10 groups (deciles) to determine the bins to use.
 - Determining the best way to visualize the data.
 - Performing statistical testing on the data and interpreting the result.

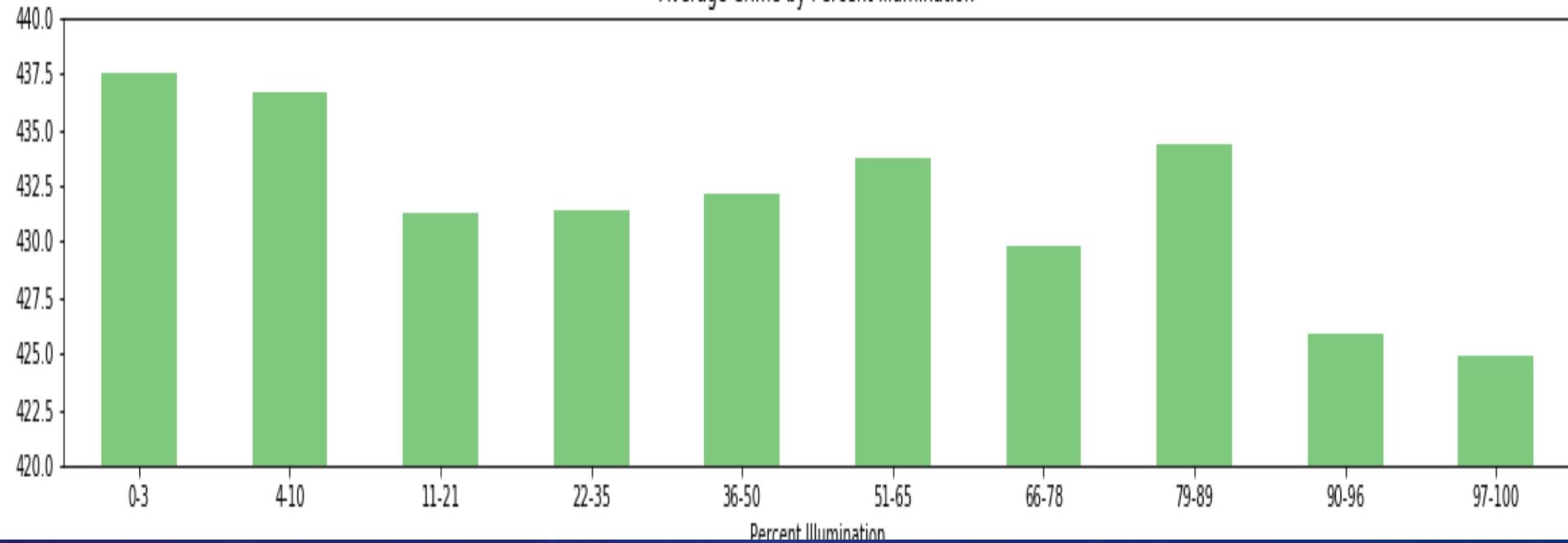
DATA ANALYSIS

- Calculated the total number of crimes for each calendar day in the date range.
- Grouped the dates by corresponding percent illumination.
- Calculated the average total daily crimes for each group.
- Plotted the average total daily crimes per group to determine if there were more crimes when the moon was full or nearly full (97-100%).

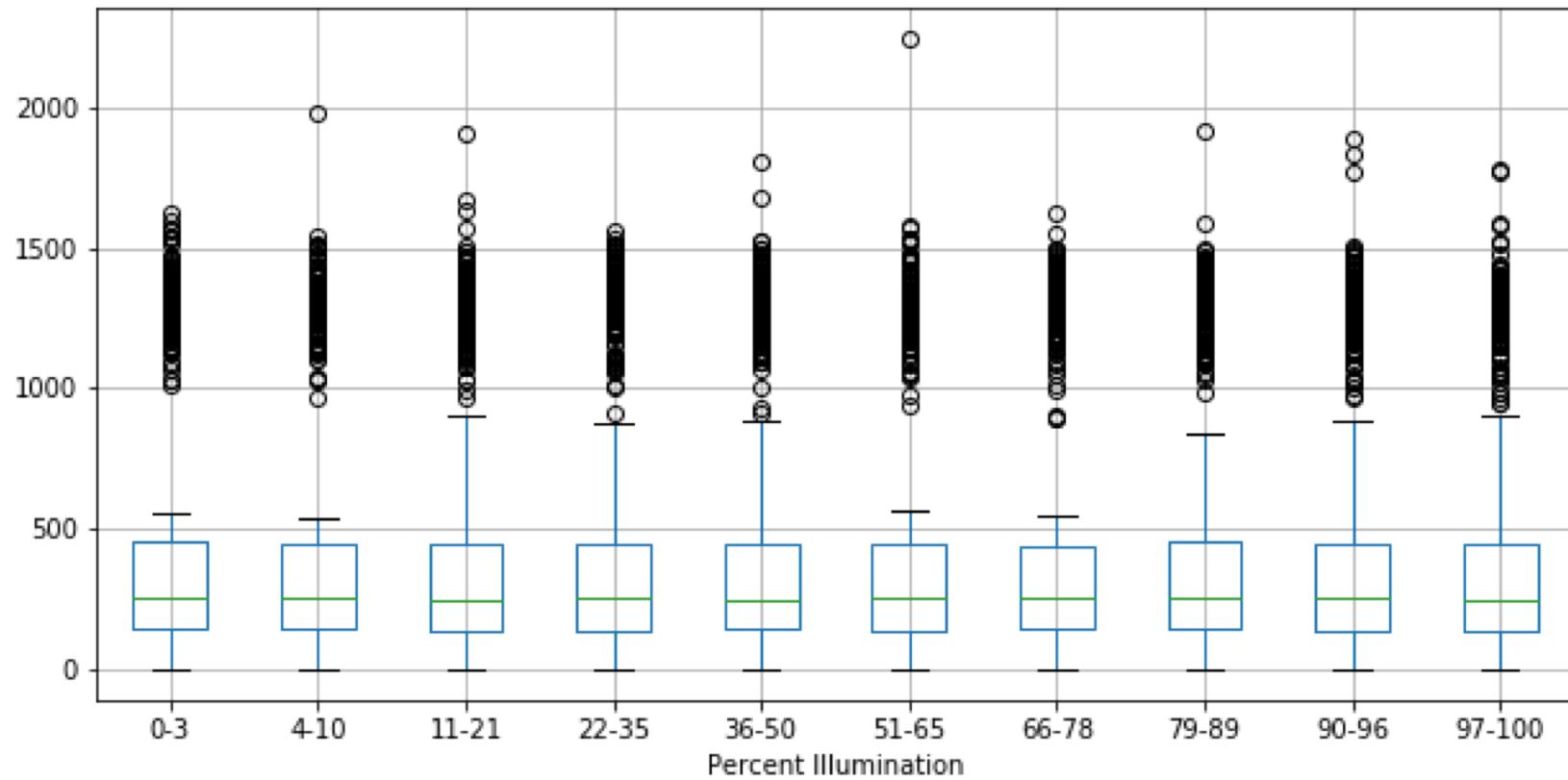
VISUALIZATIONS & DISCUSSION



Average Crime by Percent Illumination



Boxplot grouped by Percent Illumination
Total Crimes



STATISTICAL ANALYSIS

- p-values are all >0.05 , indicating acceptance of the null hypothesis (H_0).
- t-statistics are all between -2 and 2, also indicating that our H_0 is likely true.
- Relatively low F values indicate that for majority of the cities, total crime distribution is not far from the mean.
- All results indicate that there is not much variance from the mean and the probability of observing significant variance from the mean is low.
- The bottom line here is: we have failed to reject our H_0 .

Merged Crime Data for All Cities

Summary

Individual City Statistics:

1. San Francisco:

- Individual t-test: t-statistic = -0.34, pvalue = 0.73
- One-Way ANOVA: F value = 0.39, pvalue = 0.94

2. New York City:

- Individual t-test: t-statistic = -1.80, pvalue = 0.074
- One-Way ANOVA: F value = 1.14, pvalue = 0.33

3. Minneapolis:

- Individual t-test: t-statistic = 0.66, pvalue = 0.51
- One-Way ANOVA: F value = 0.20, pvalue = 0.99

4. Portland:

- Individual t-test: t-statistic = -0.76, pvalue = 0.45
- One-Way ANOVA: F value = 0.34, pvalue = 0.96

5. Denver:

- Individual t-test: t-statistic = -1.33, pvalue = 0.19
- One-Way ANOVA: F value = 1.13, pvalue = 0.34

Overall Crime During Moon Phase Statistics:

- Individual t-test: t-statistic = -0.36, pvalue = 0.72
- One-Way ANOVA: F value = 0.046, pvalue = 0.999

POST MORTEM

- Difficulties we experienced:
 - Open source API's with large, complete and reliable data sets
 - Complete crime data was not as easy to find during our date range as initially expected.

DATA VISUALIZATION

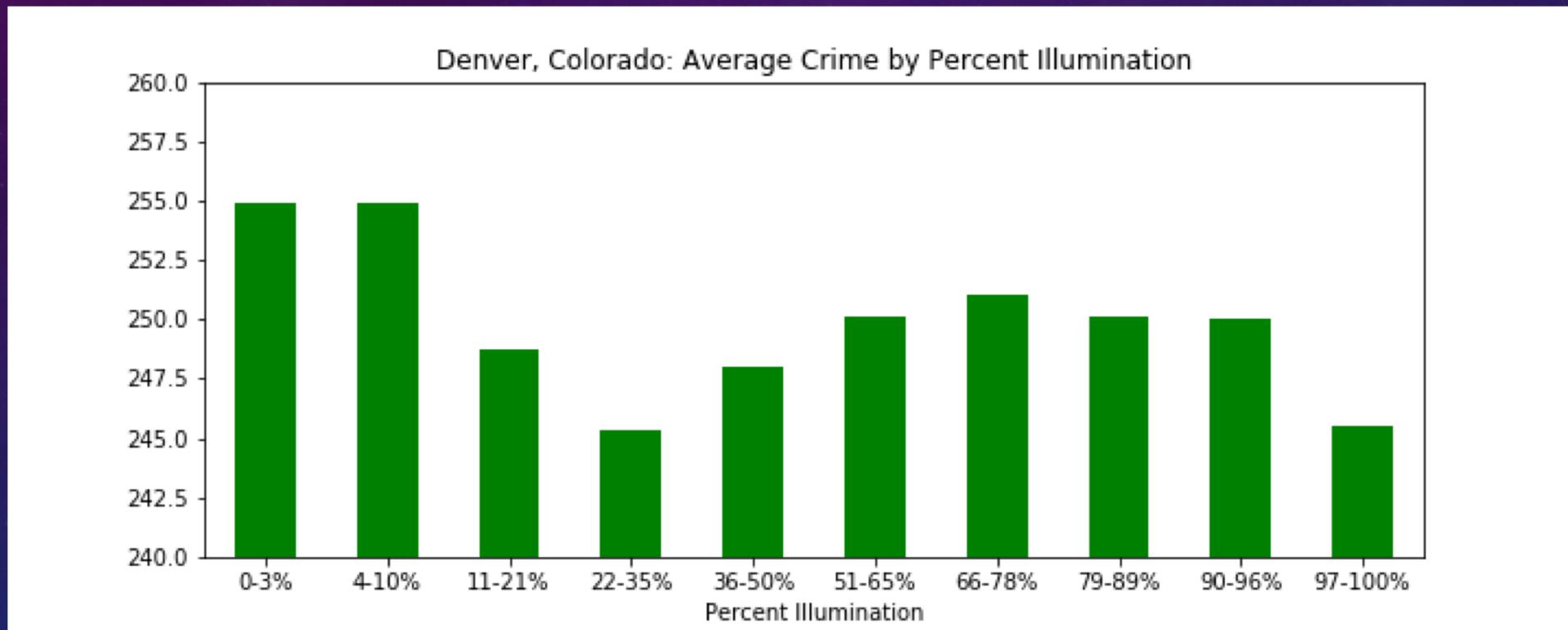


SO HOT RIGHT NOW

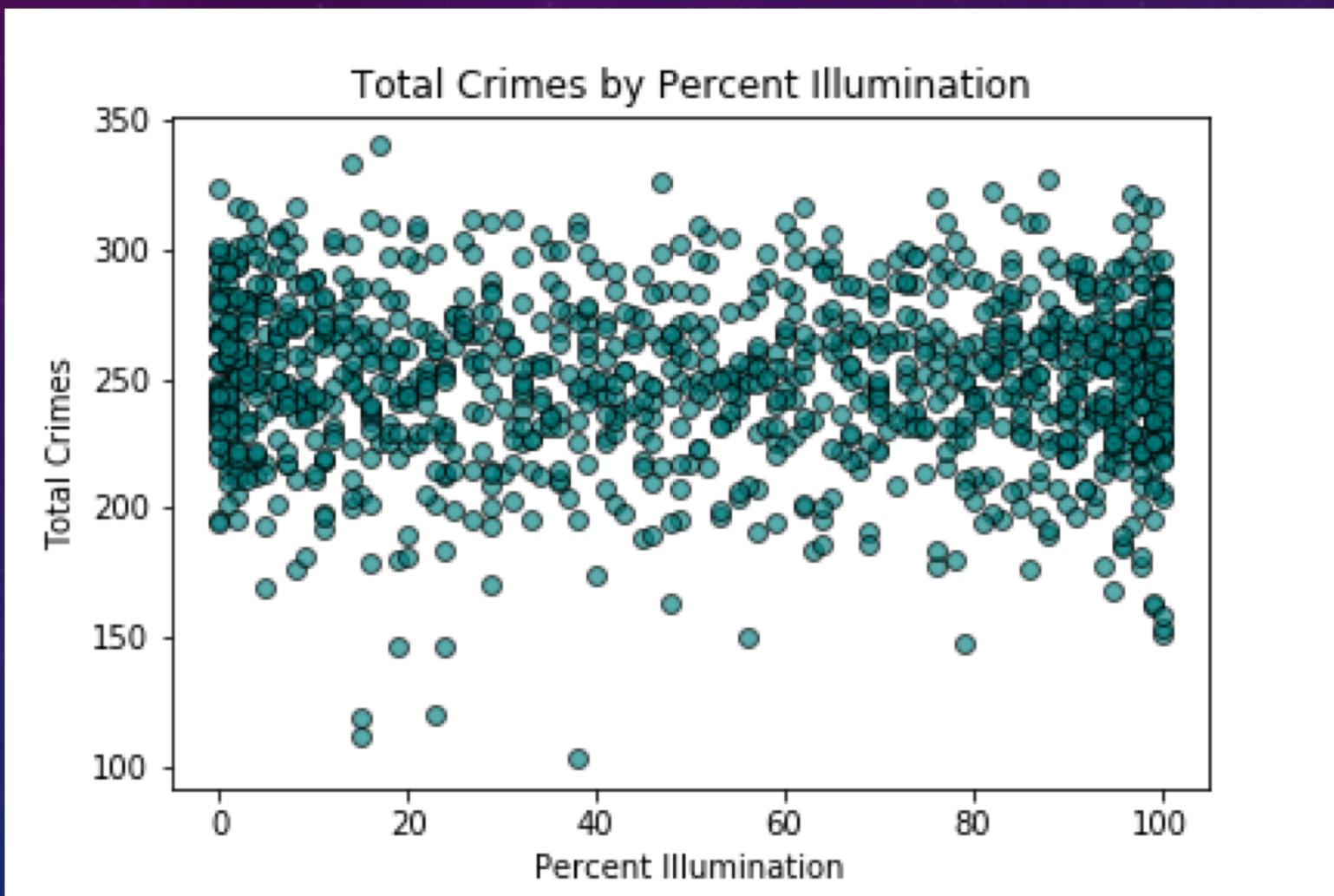
memegenerator.net

QUESTIONS?

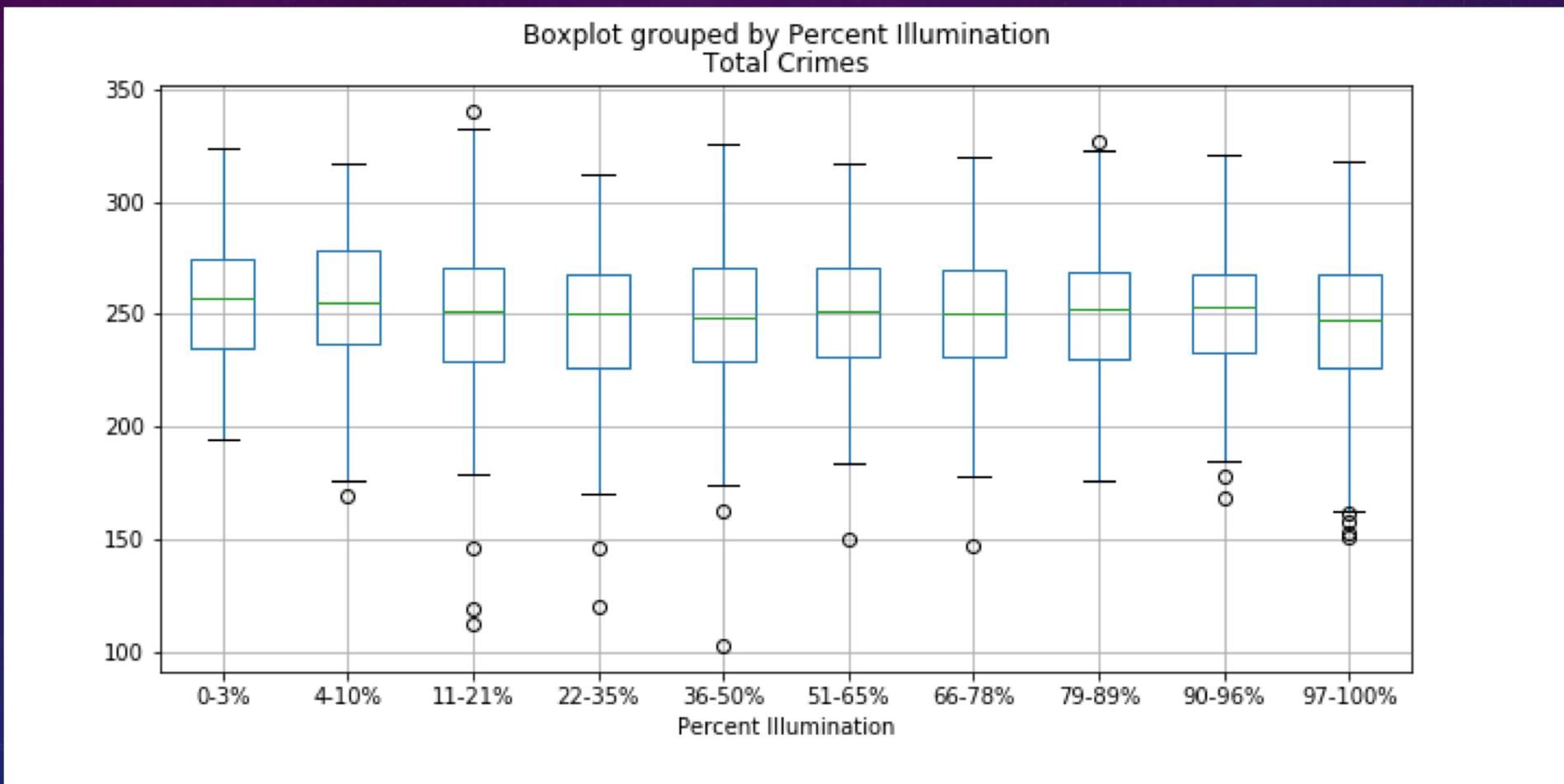
APPENDIX: ADDITIONAL VISUALIZATIONS FOR EA. CITY



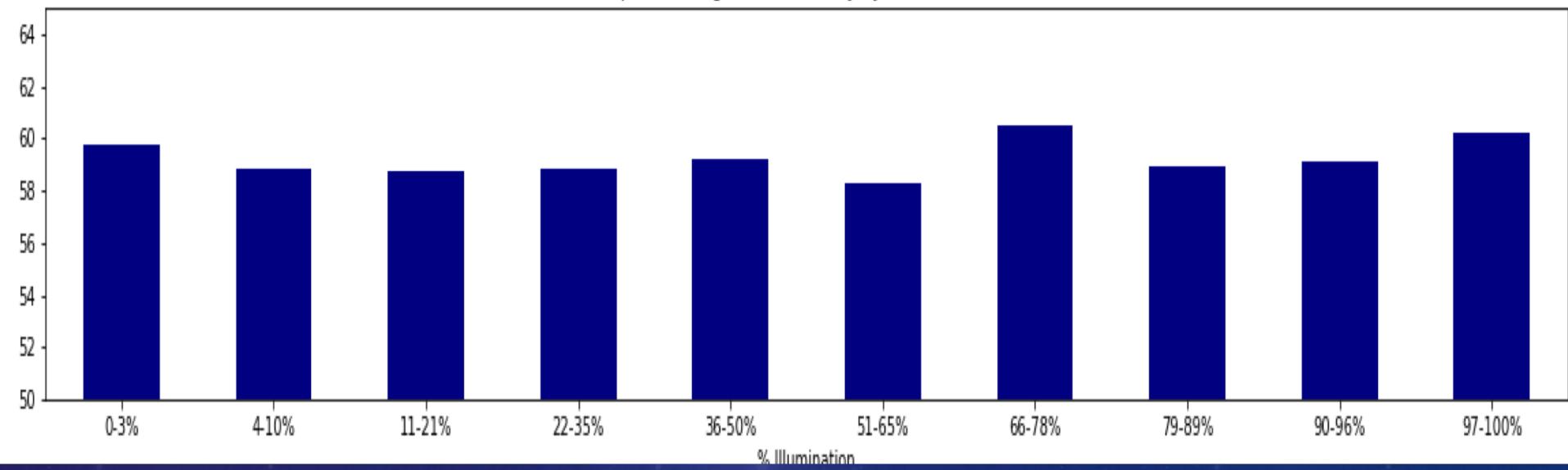
Denver, Colorado Total Crimes by Percent Illumination



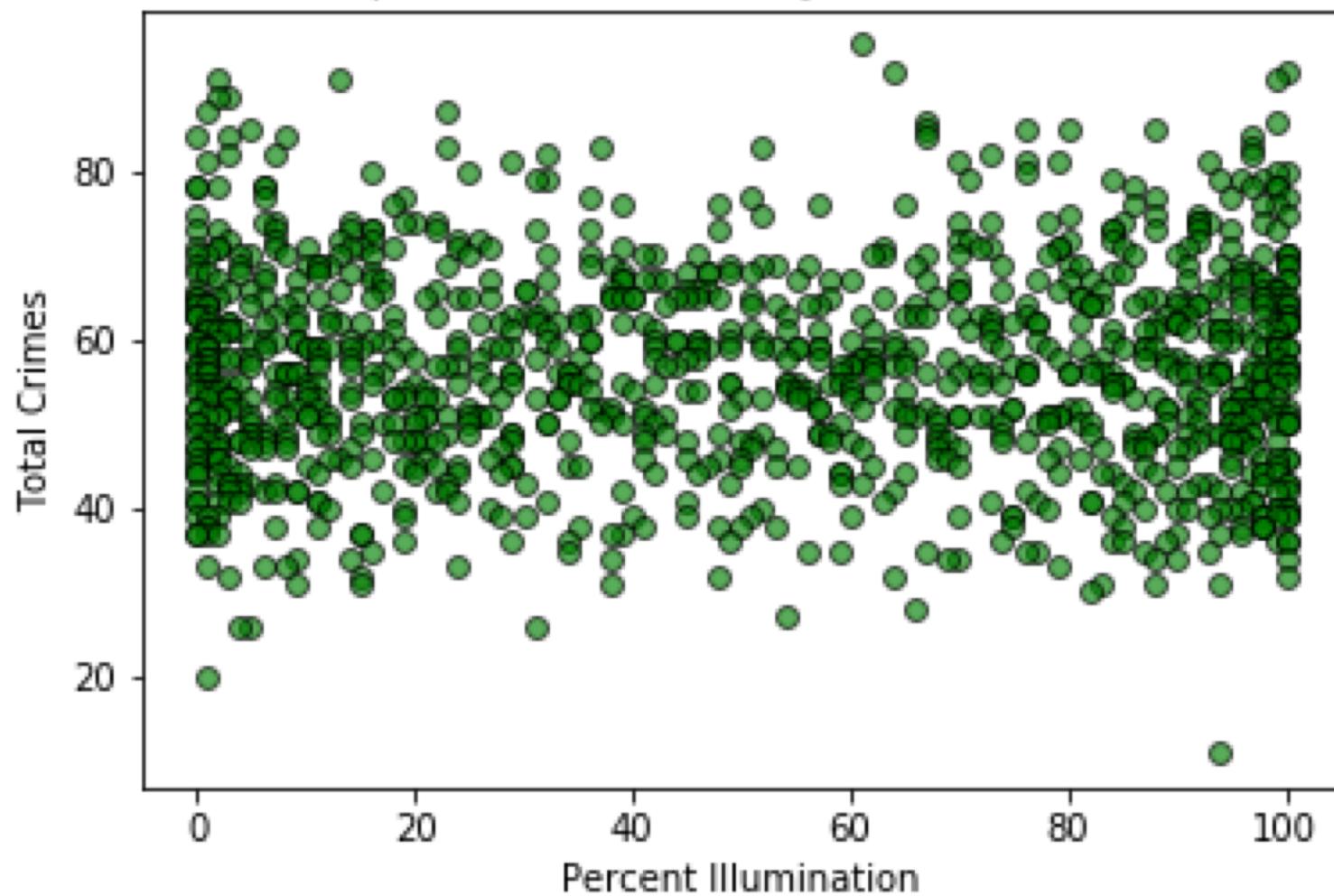
DENVER, COLORADO BOX PLOT



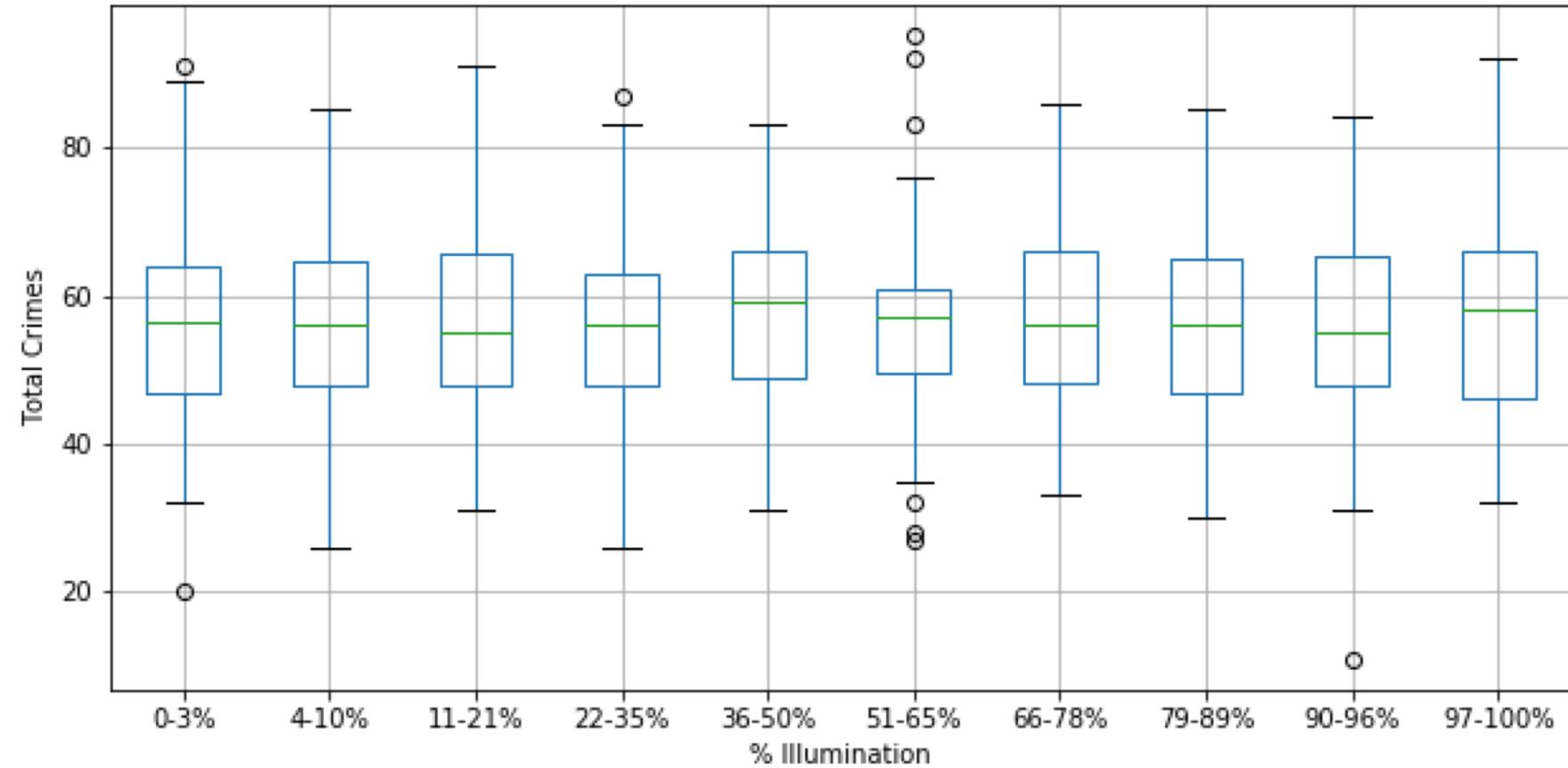
Minneapolis Average Crime Per Day by Percent Illumination



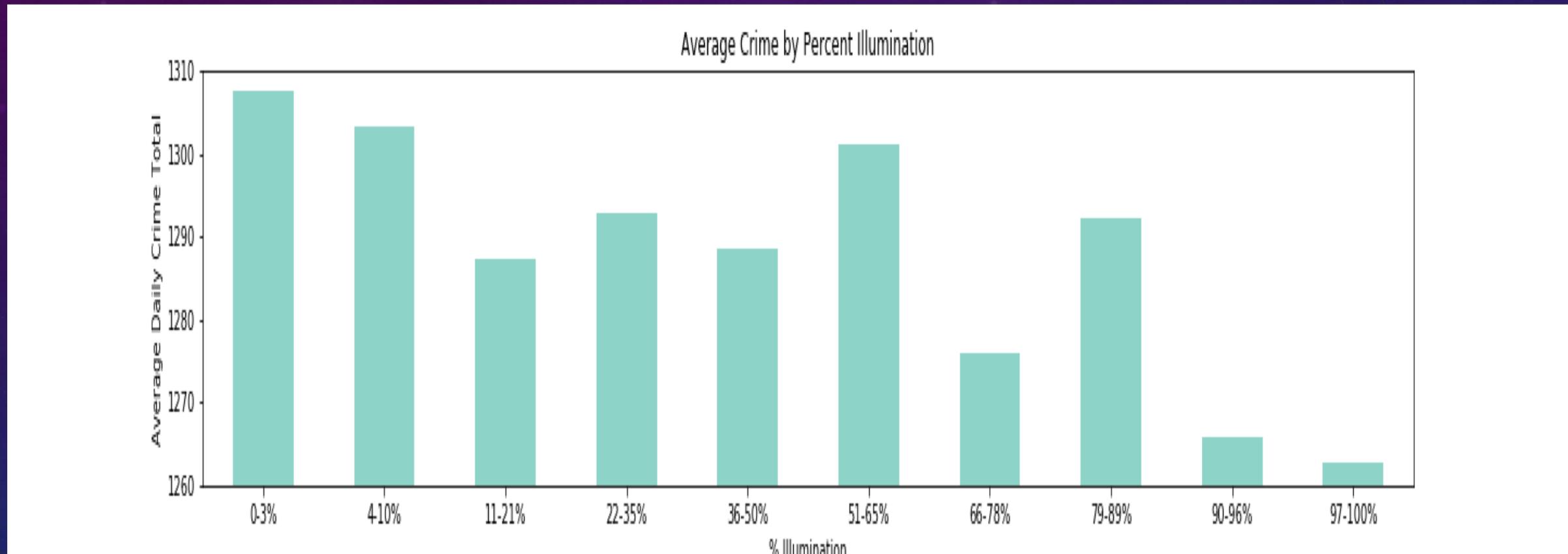
Minneapolis Total Crimes by Percent Illumination



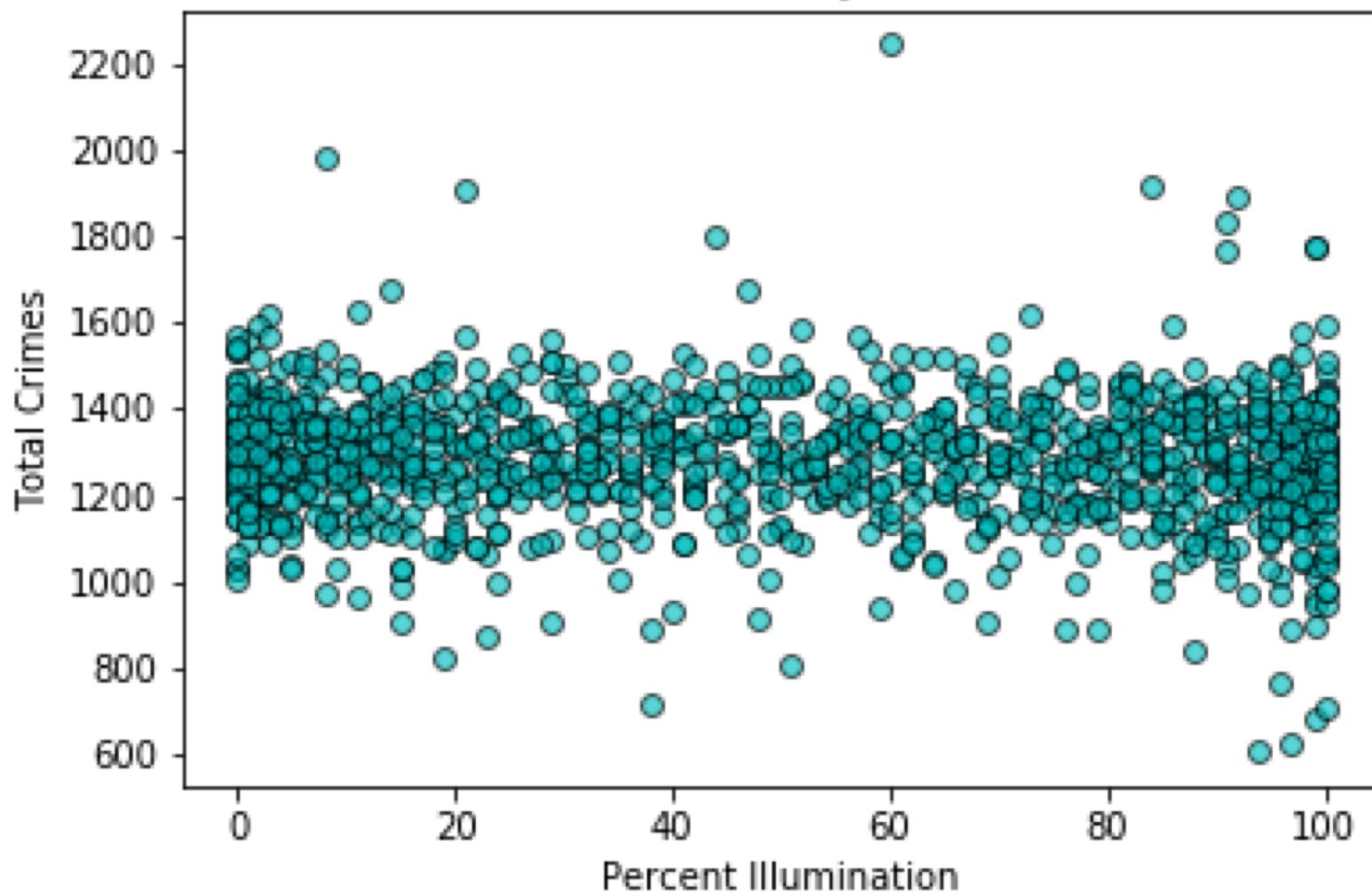
Boxplot grouped by % Illumination
Minneapolis



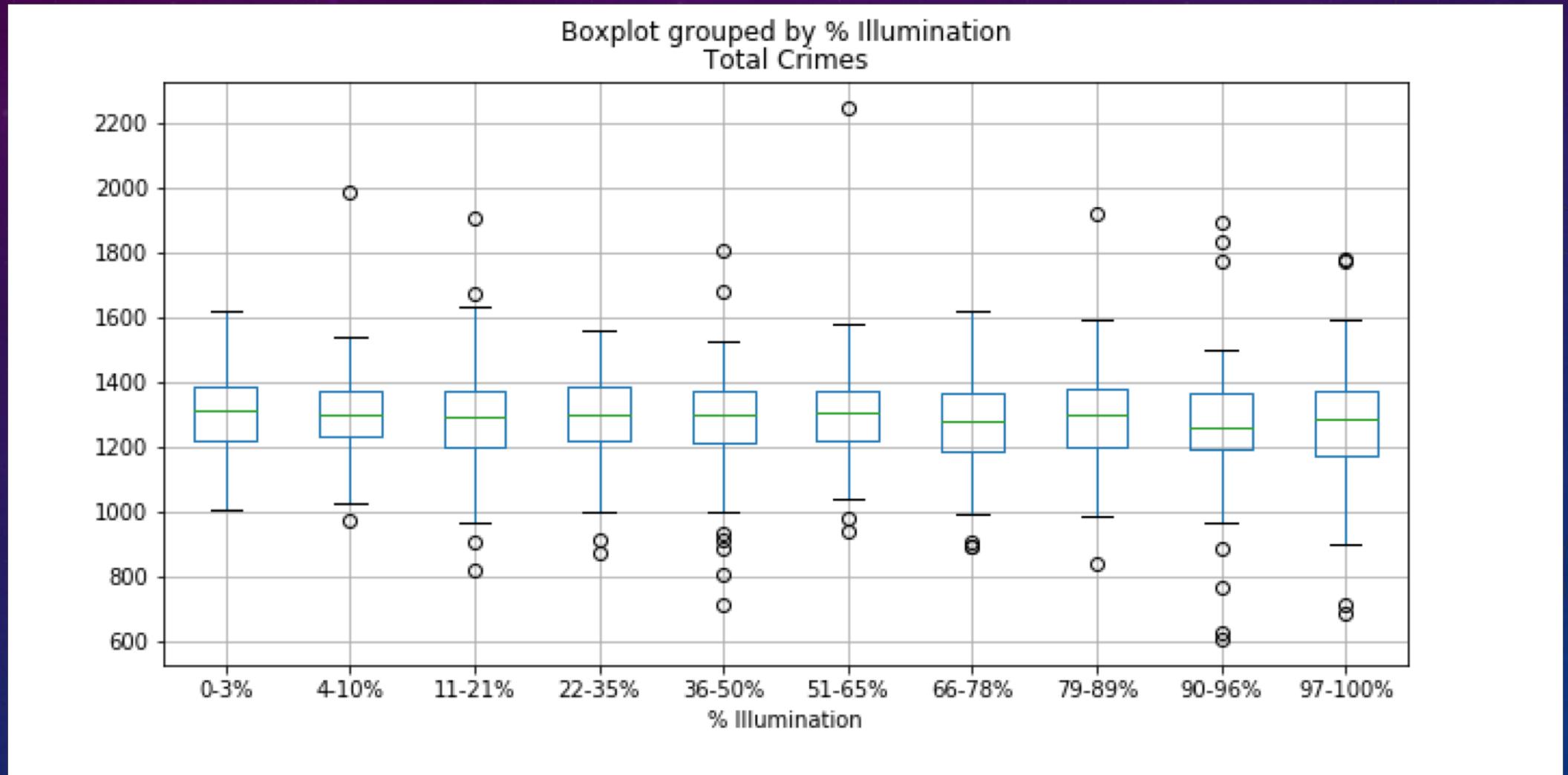
NEW YORK CITY AVERAGE CRIME BY PERCENT ILLUMINATION



Total Crimes in New York by Percent Illumination



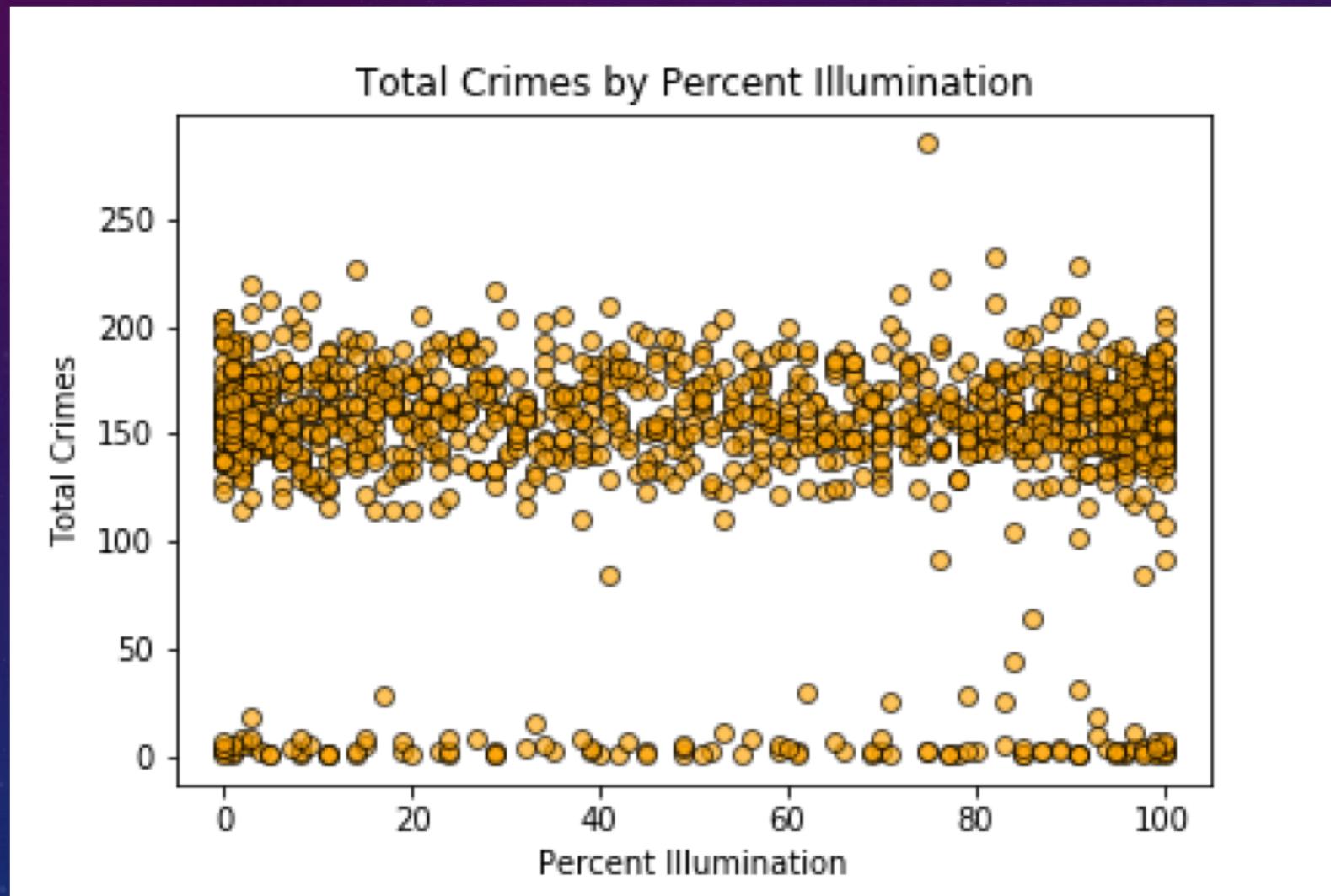
NEW YORK CITY BOXPLOT



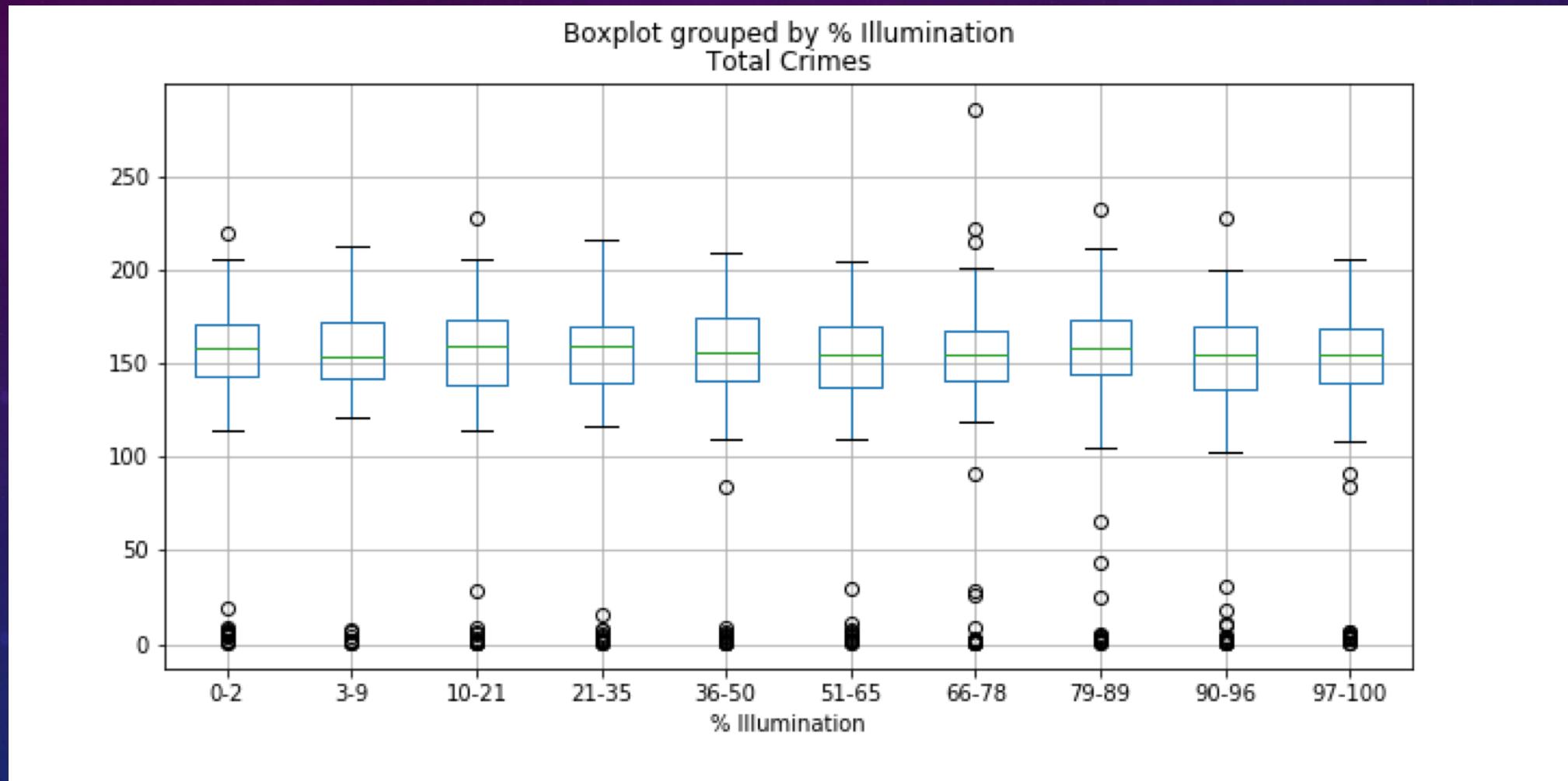
PORLAND, OR AVERAGE CRIME BY PERCENT ILLUMINATION



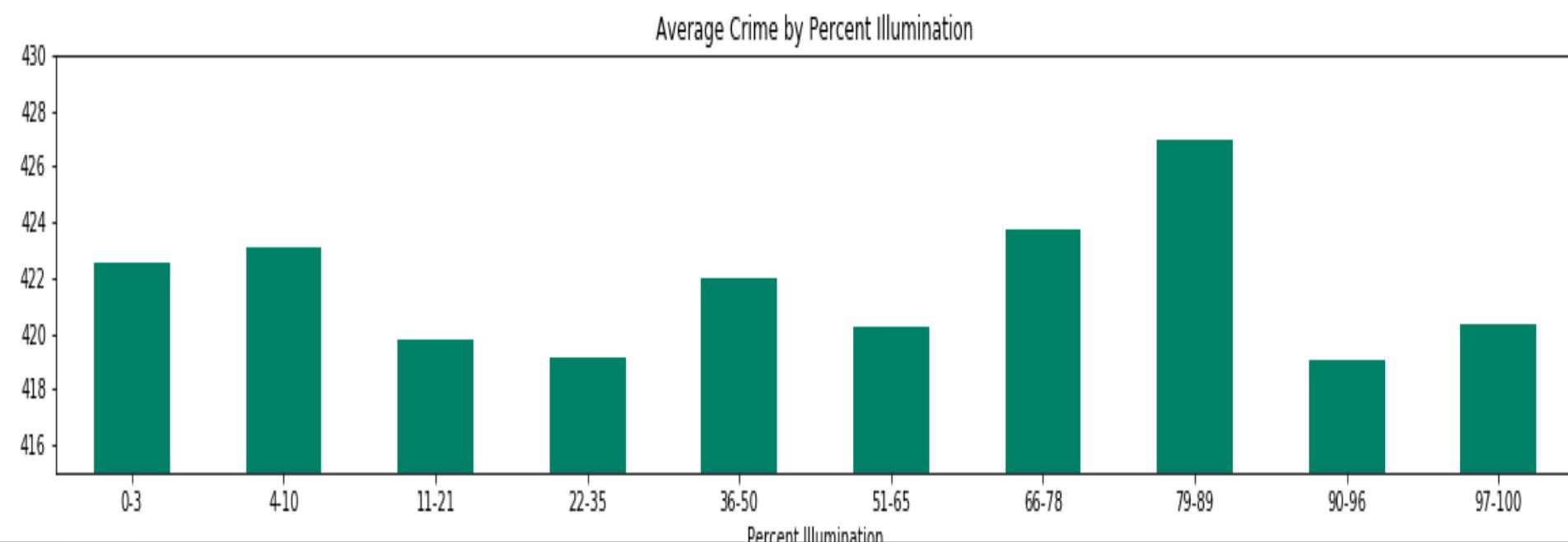
PORLAND, OR TOTAL CRIMES BY PERCENT ILLUMINATION



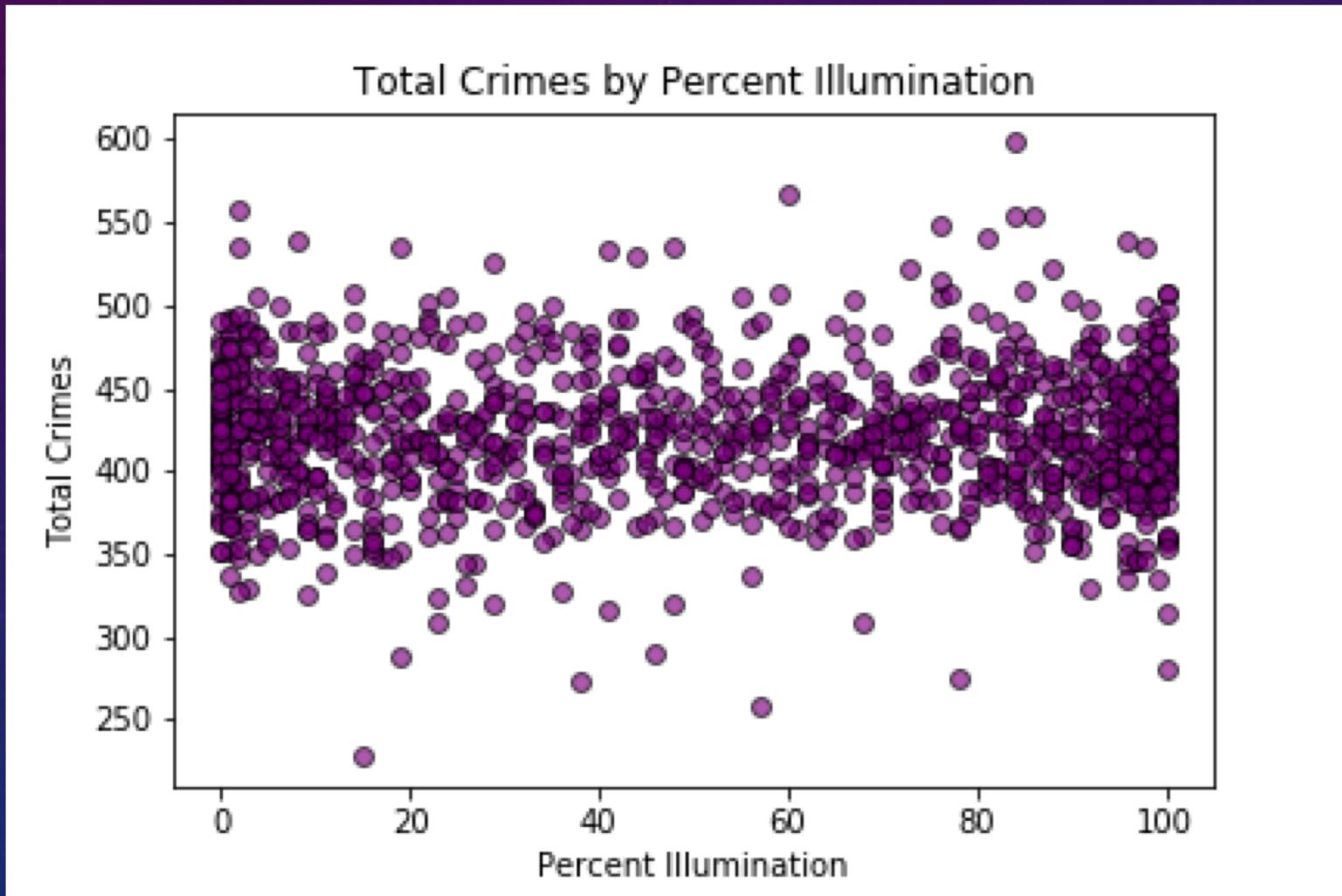
PORLAND, OR BOXPLOT



SAN FRANCISCO AVERAGE CRIME BY PERCENT ILLUMINATION



SAN FRANCISCO TOTAL CRIMES BY PERCENT ILLUMINATION



SAN FRANCISCO BOXPLOT

