Team Name: TBD

Team Members:
Rohit Vakkalagadda (rv324@cornell.edu)
Kaitlyn Lu (kl995@cornell.edu)
Lucas He (th689@cornell.edu)
Project Manager: Andy He

What, if any, insights did you gain from the Oakley reading?
Ensuring high-quality work early will prevent workload-based resentment within a group.

What, if any, insights did you get from your PM meeting?
Our PM was sick this week, so we were unable to meet.

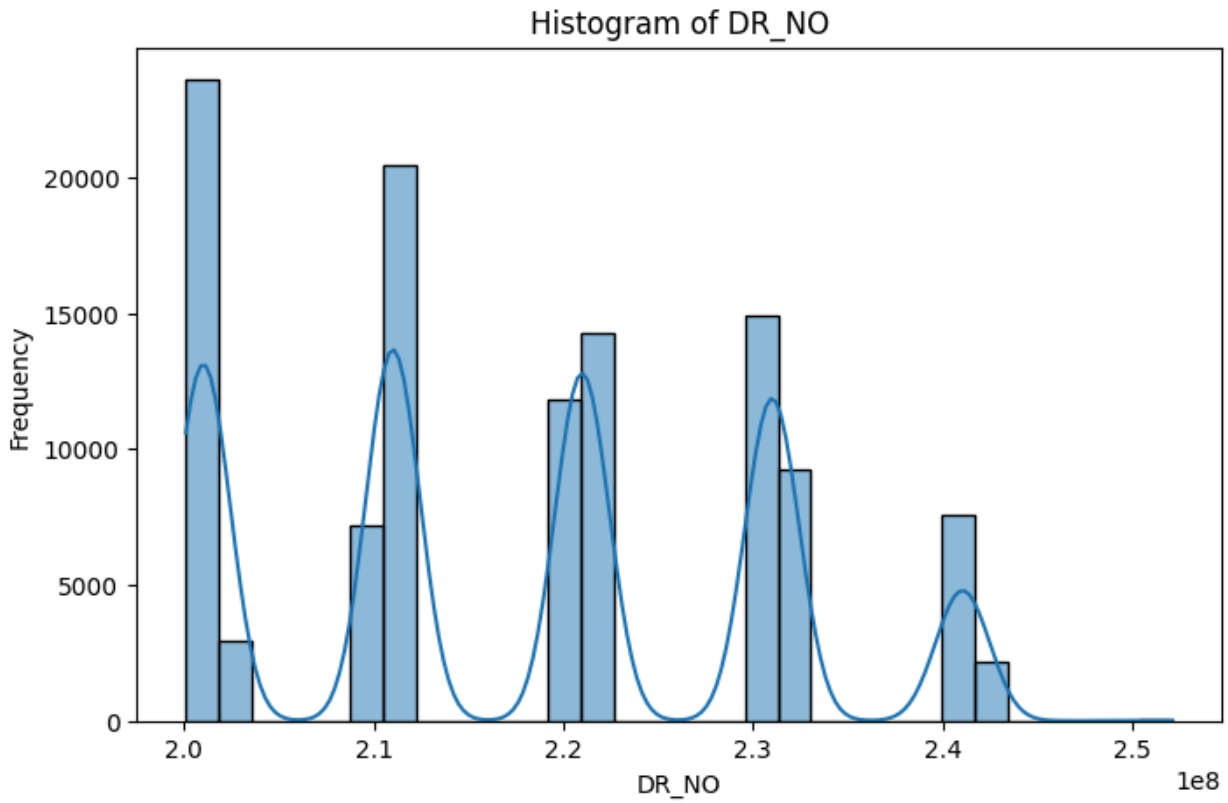
Report summary statistics about each column in your dataset.

- Summary for AREA NAME:
  Unique values: 21
- Summary for Crm Cd Desc:
  Unique values: 121
- Summary for Mocodes:
  Unique values: 68816
- Summary for Vict Sex:
  Unique values: 4
- Summary for Premis Desc:
  Unique values: 280
- Summary for LOCATION:
  Unique values: 8138
- Summary for Cross Street:
  Unique values: 8926
- Summary for DR_NO:
  Stats: mean 2.177692e+08, std 1.274684e+07, min 2.001005e+08, 50% 2.201234e+08, max 2.521040e+08
- Summary for TIME OCC:
  Stats: mean 1369.055331, std 662.537589, min 1.000000, 50% 1500.000000, max 2359.000000
- Summary for AREA:
  Stats: mean 10.033279, std 6.337220, min 1.000000, 50% 11.000000, max 21.000000
- Summary for Rpt Dist No:

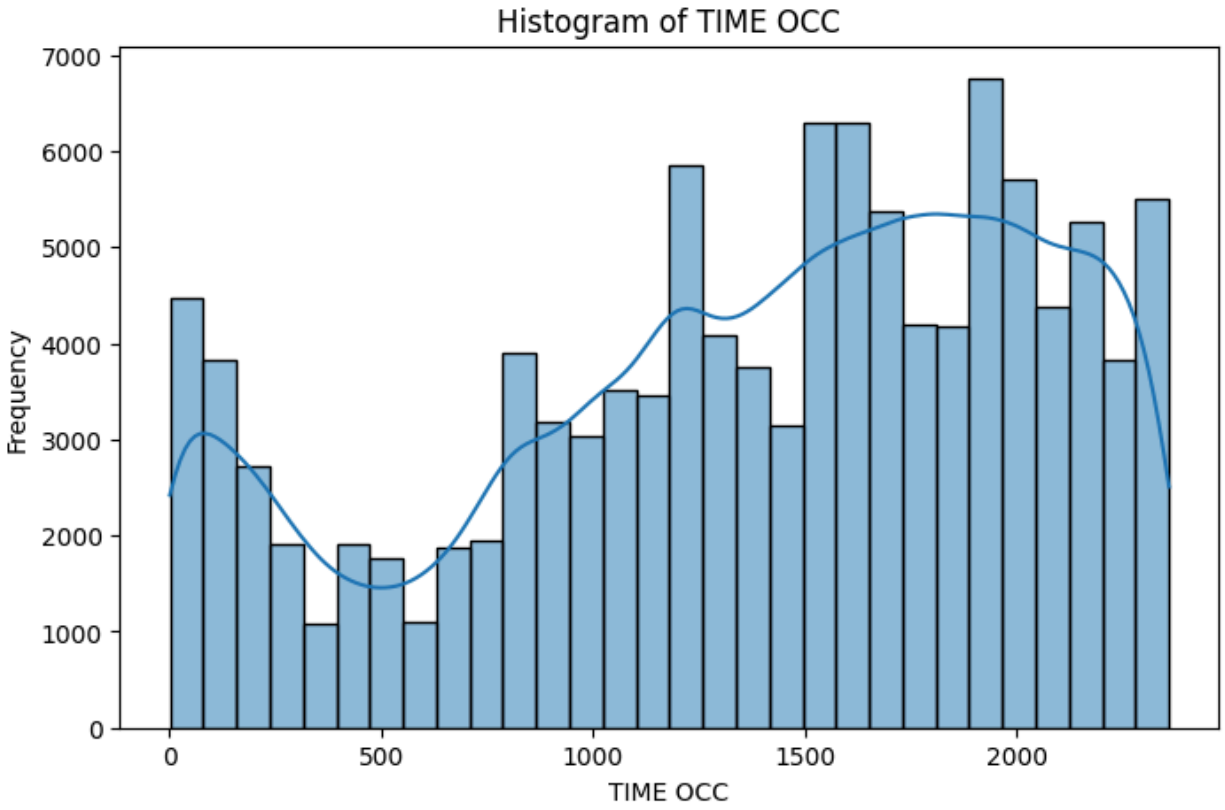
Stats: mean 1048.992317, std 631.355909, min 101.000000, 50% 1117.000000, max 2199.000000

- Summary for Part 1-2:
  Stats: mean 1.367838, std 0.482219, min 1.000000, 50% 1.000000, max 2.000000
- Summary for Crm Cd:
  Stats: mean 454.168793, std 224.573729, min 110.000000, 50% 350.000000, max 956.000000
- Summary for Vict Age:
  Stats: mean 37.738510, std 14.656727, min -3.000000, 50% 35.000000, max 99.000000
- Summary for Premis Cd:
  Stats: mean 167.435560, std 175.756263, min 101.000000, 50% 101.000000, max 971.000000
- Summary for LAT:
  Stats: mean 33.971073, std 1.737018, min 0.000000, 50% 34.048100, max 34.334300
- Summary for LON:
  Stats: mean -118.024601, std 6.025506, min -118.666600, 50% -118.297000, max 0.000000

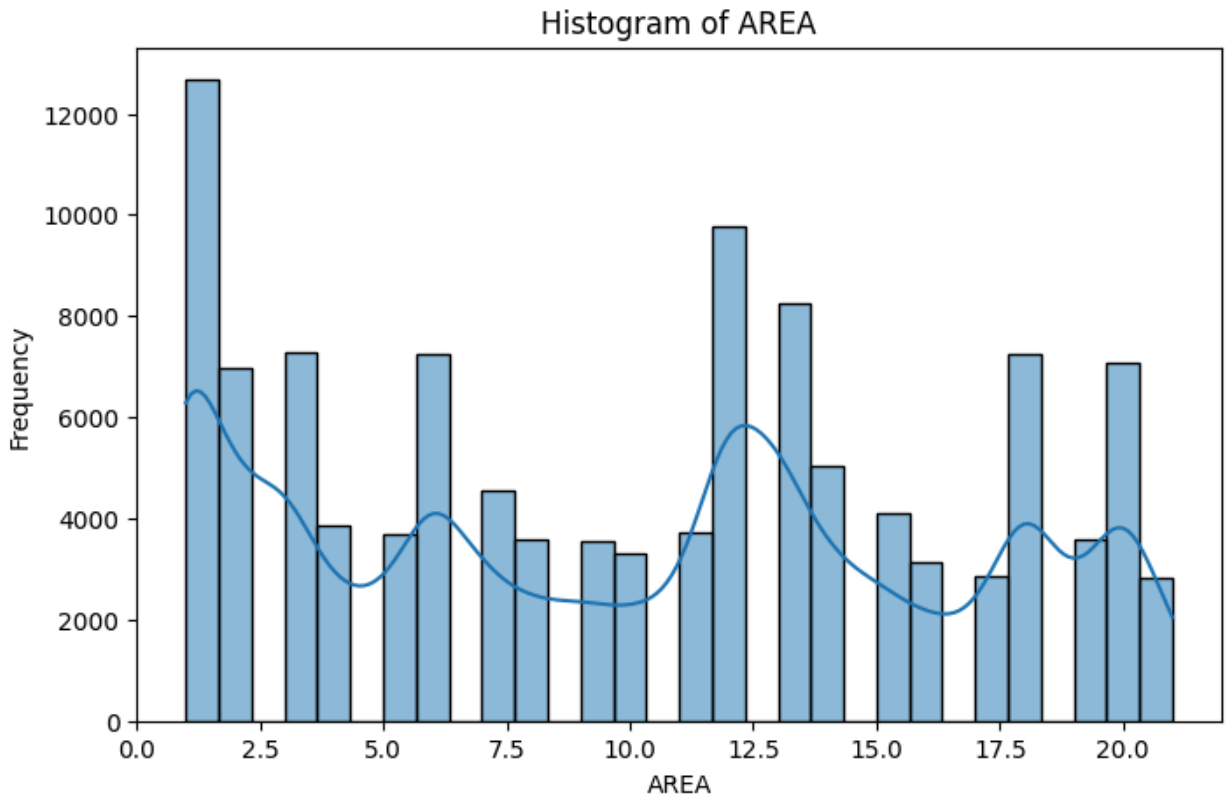
Include all plots with titles, labels, units, accompanied by short descriptions about what each plot tells you.



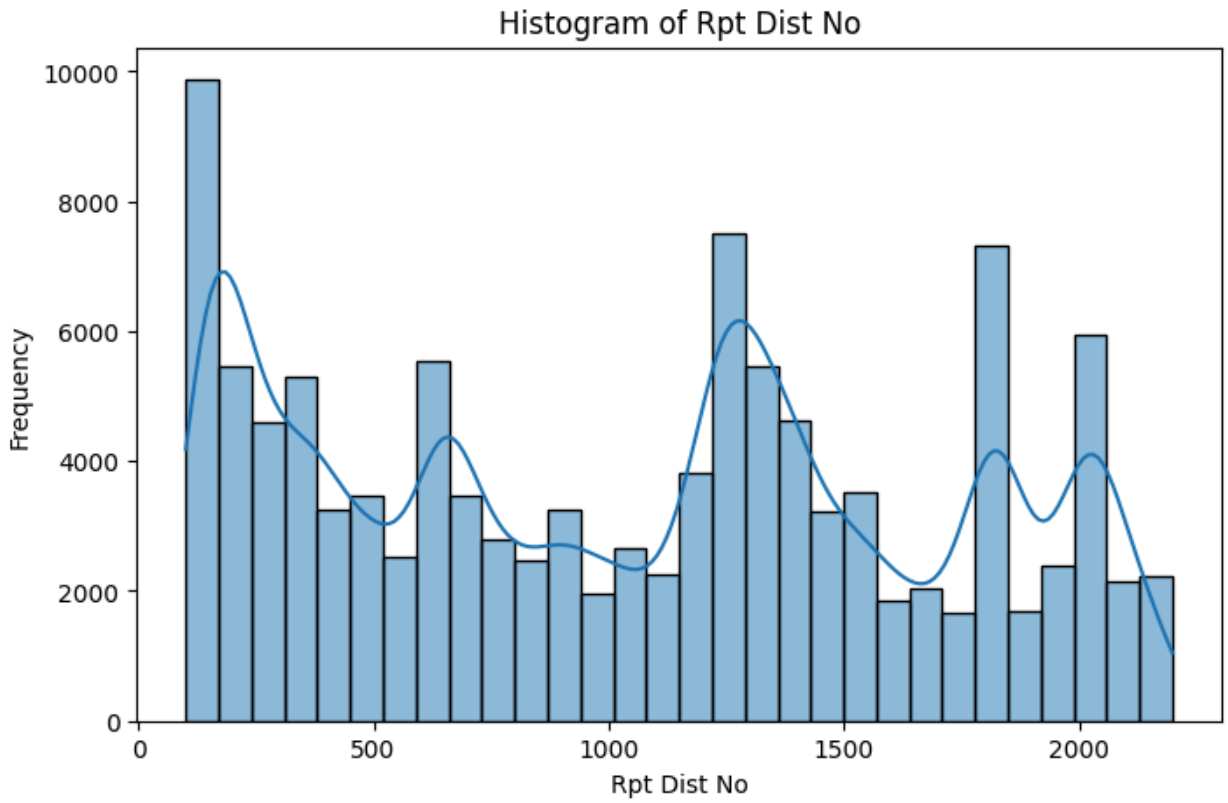
This plot shows the distribution of DR_NO (Division of Records Number), which is the specific case number associated with a case.



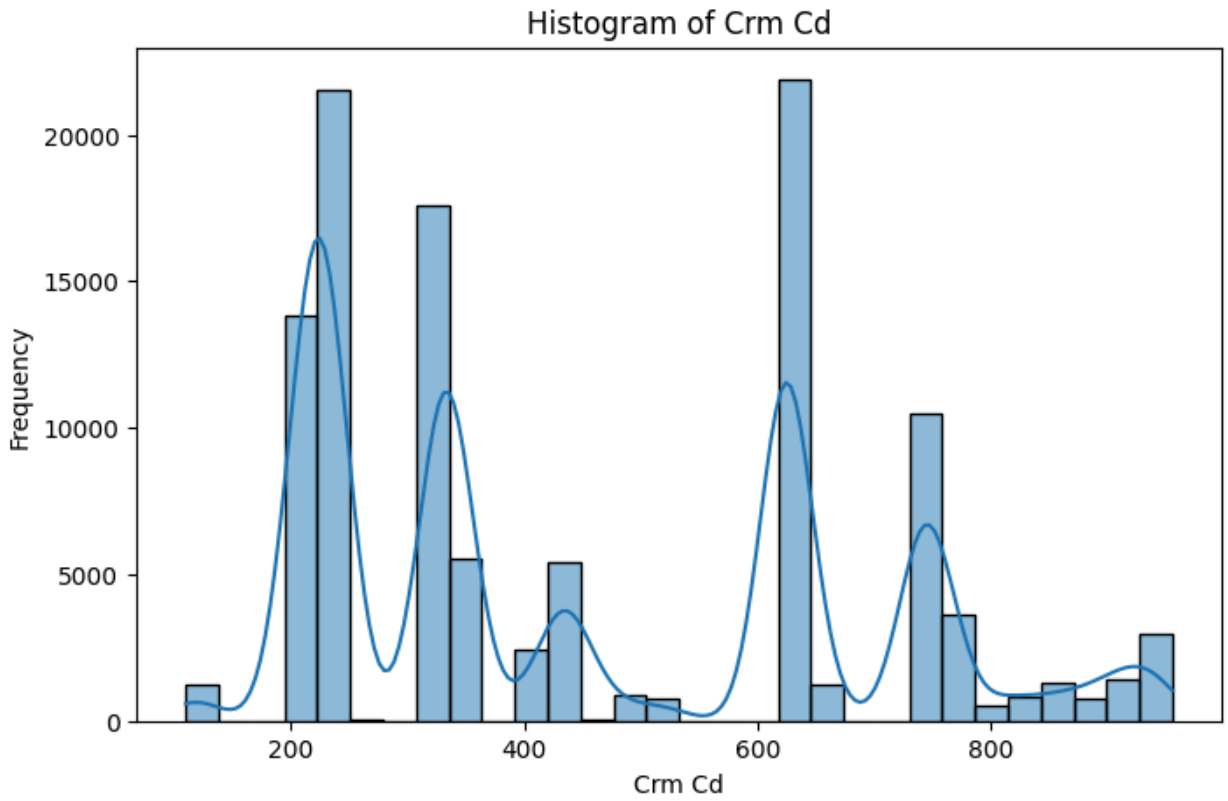
This plot shows the distribution of the times that crimes occur, indicating that less crimes happen at 5 AM (500) and more happen in the evening.



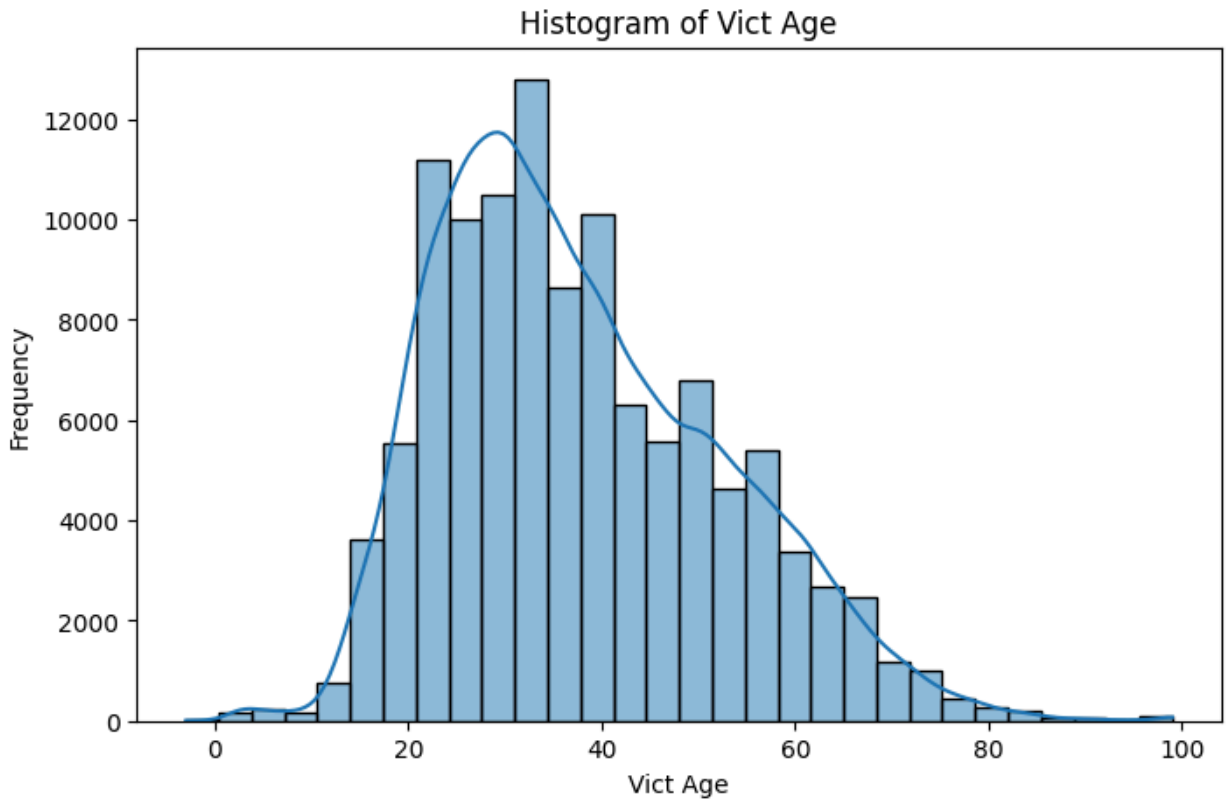
This plot shows the distribution of areas where crimes occur, where geographic location is indicated by a predetermined index.



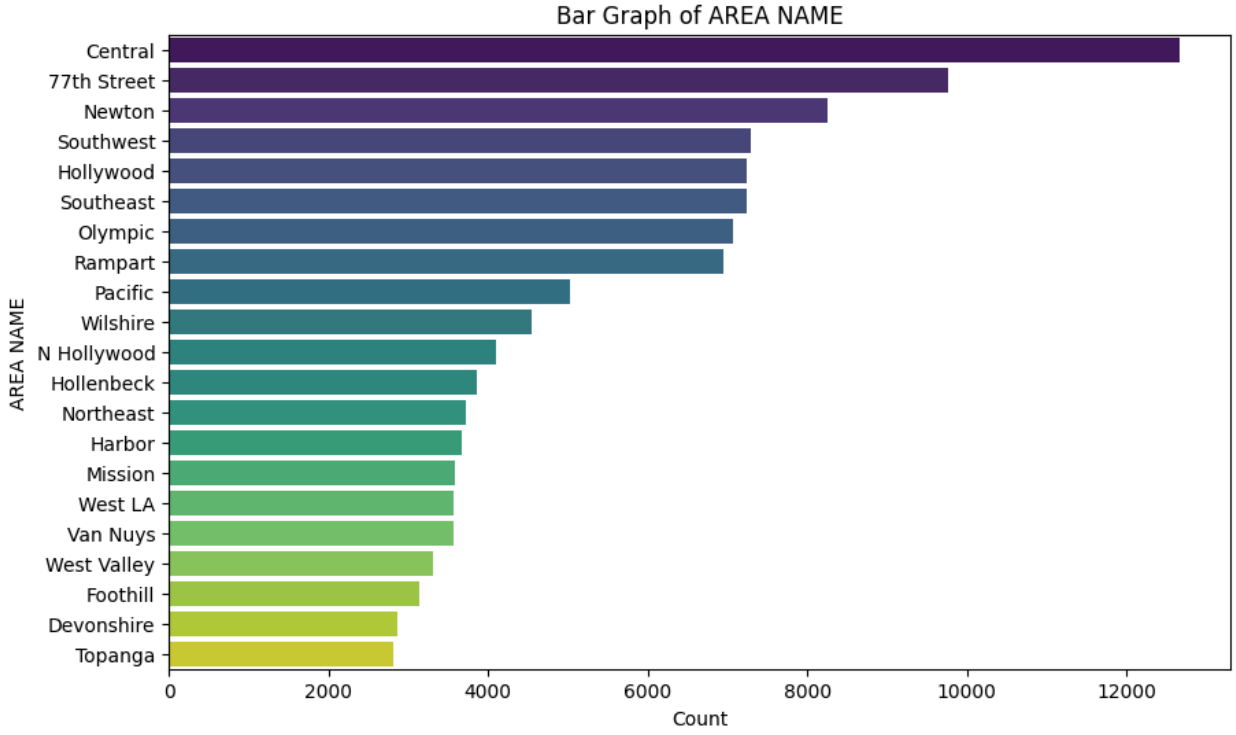
This plot shows the distribution of Rpt Dist No (reporting district number), which is the number code that represents a geographical location within a LA Police Department division.



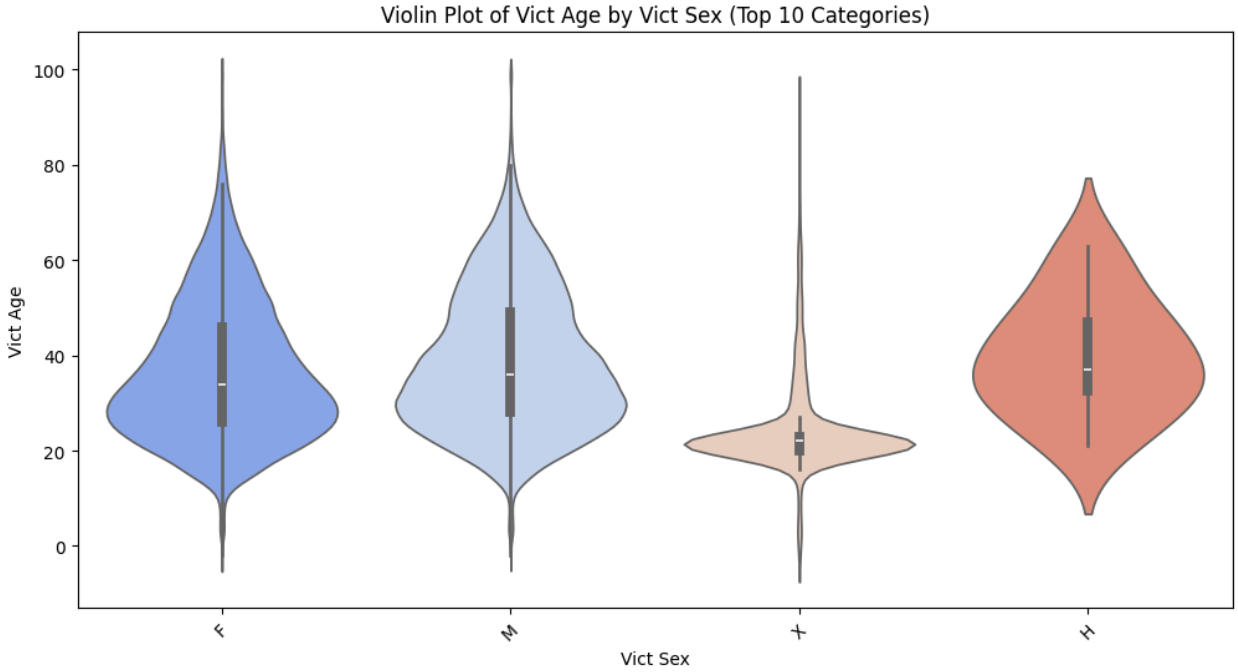
This plot shows the distribution of of Crm Cd, which is the numerical code of a crime committed; this shows that most crimes committed have codes in the 200s or 600s.



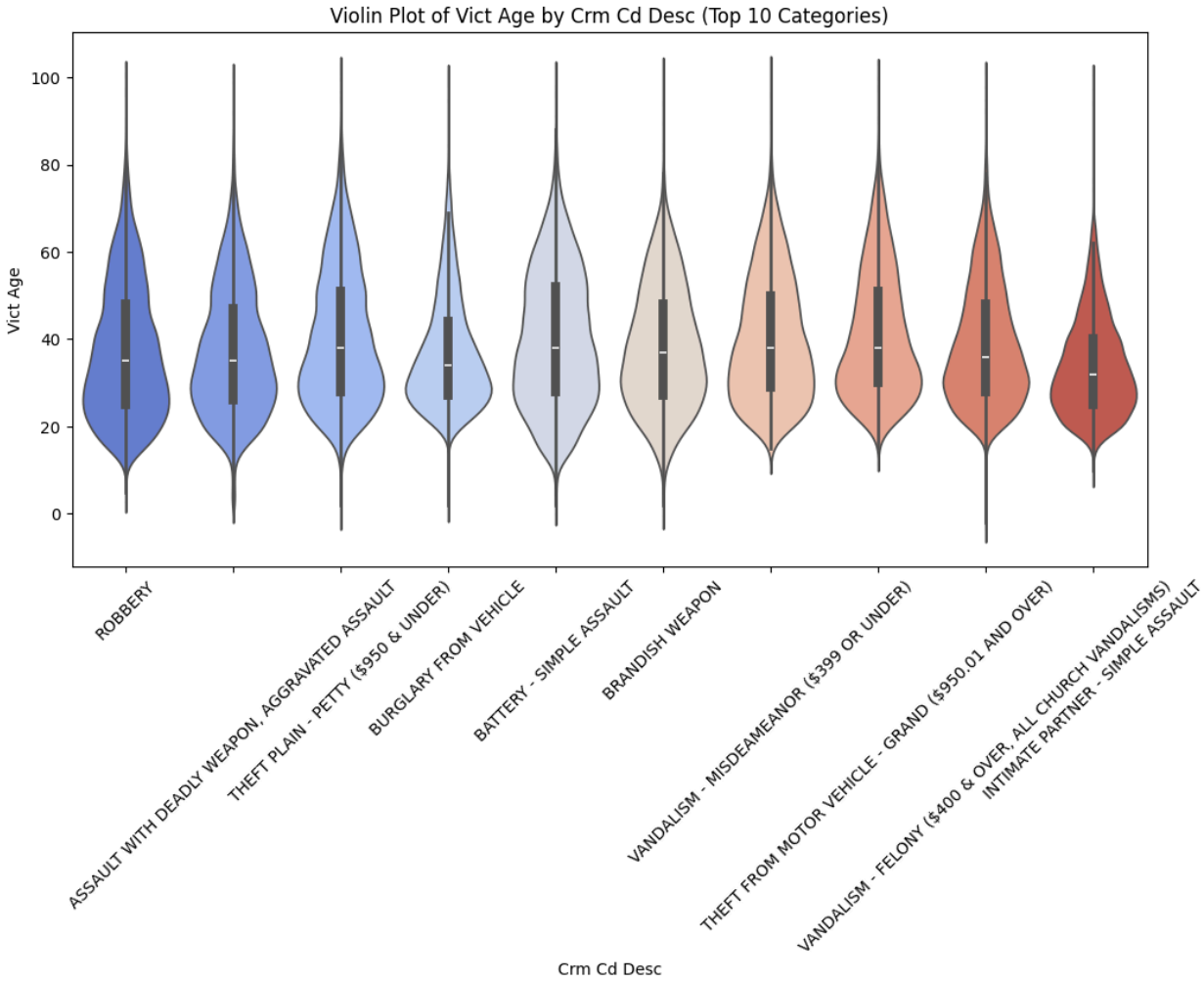
This plot shows the distribution of ages of people who commit crimes, with the peak happening around 20-40 years old.



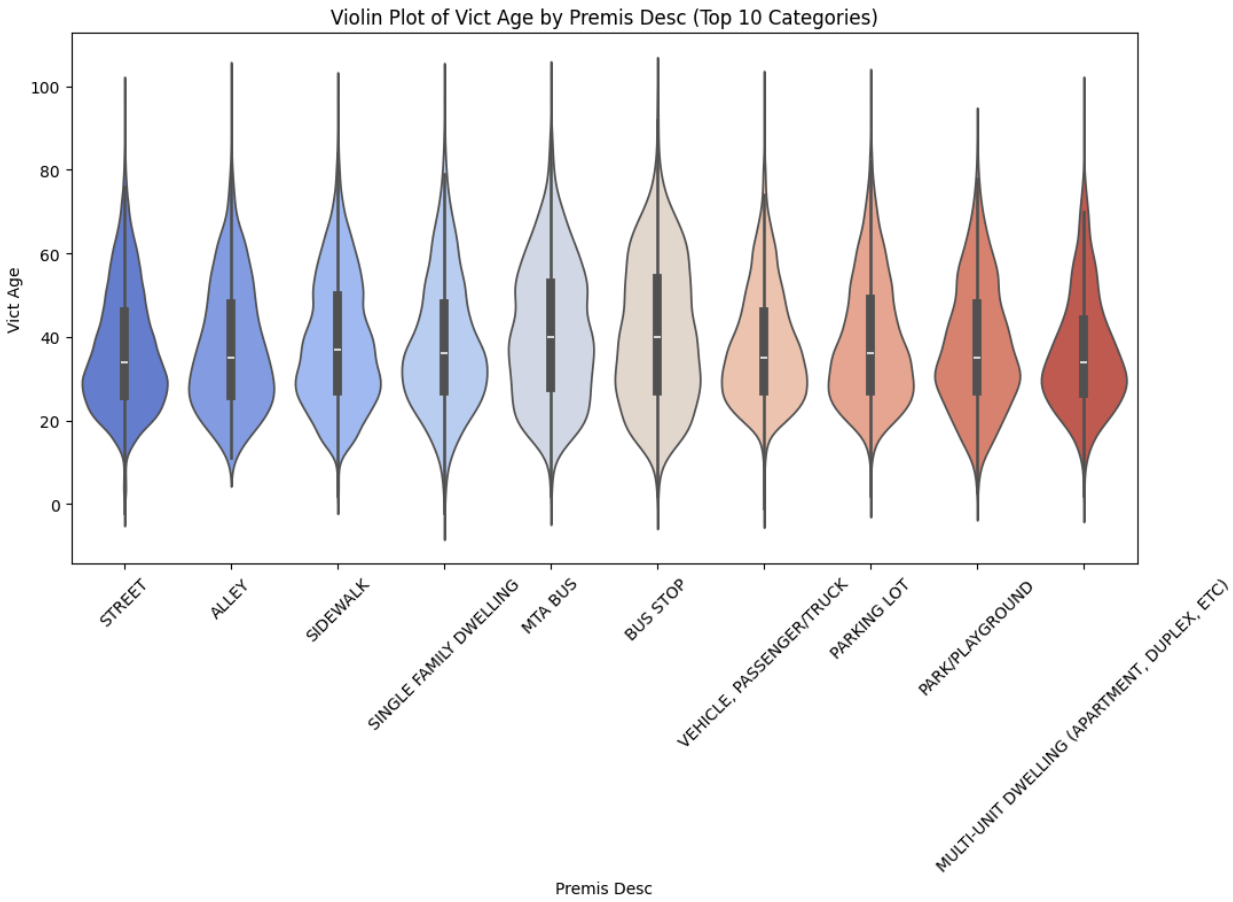
This plot how many reported crimes occur per area (it can be seen that the highest number of reported crimes occur in the Central area).



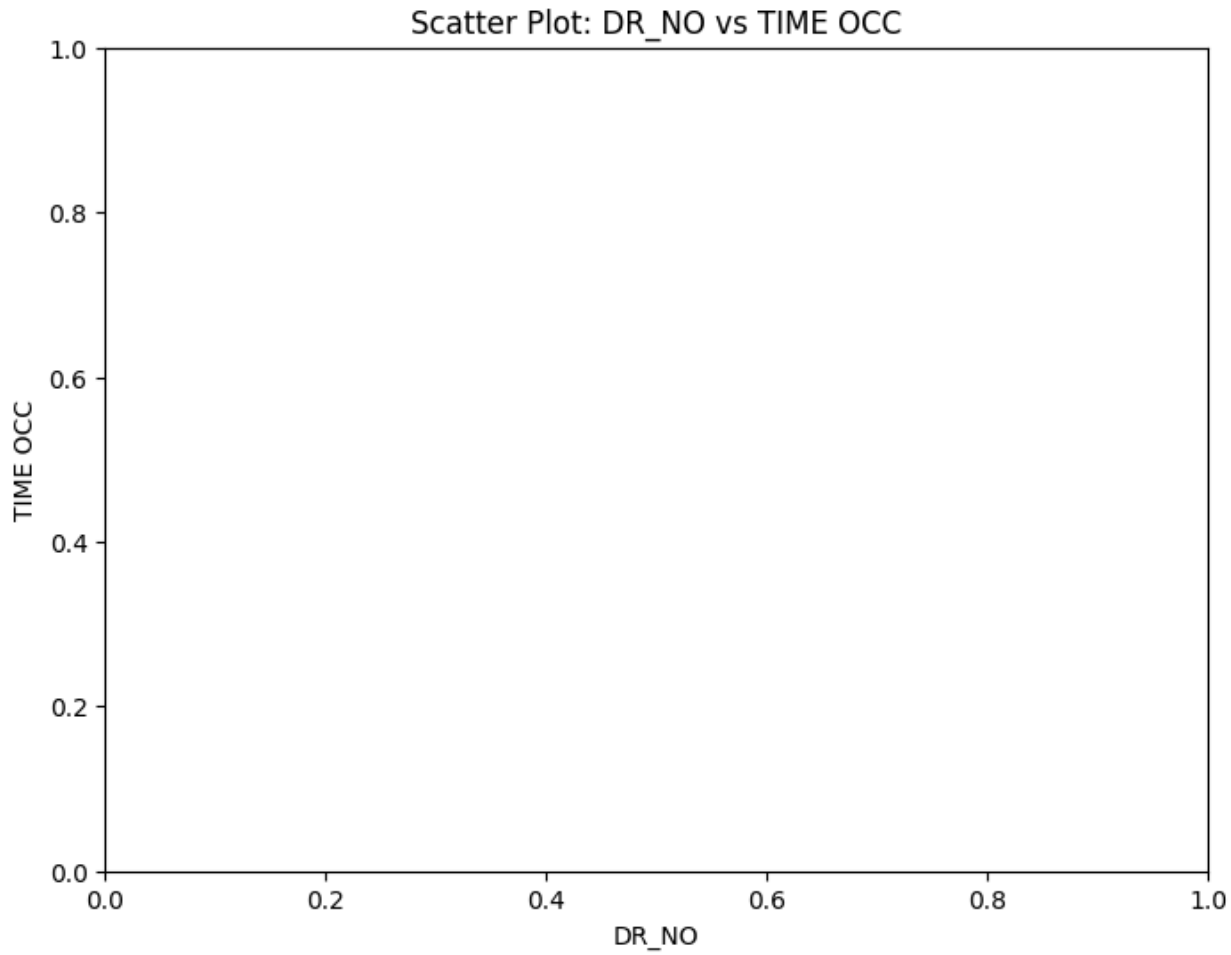
This plot shows the distribution per sex of the age of the person who committed the crime.



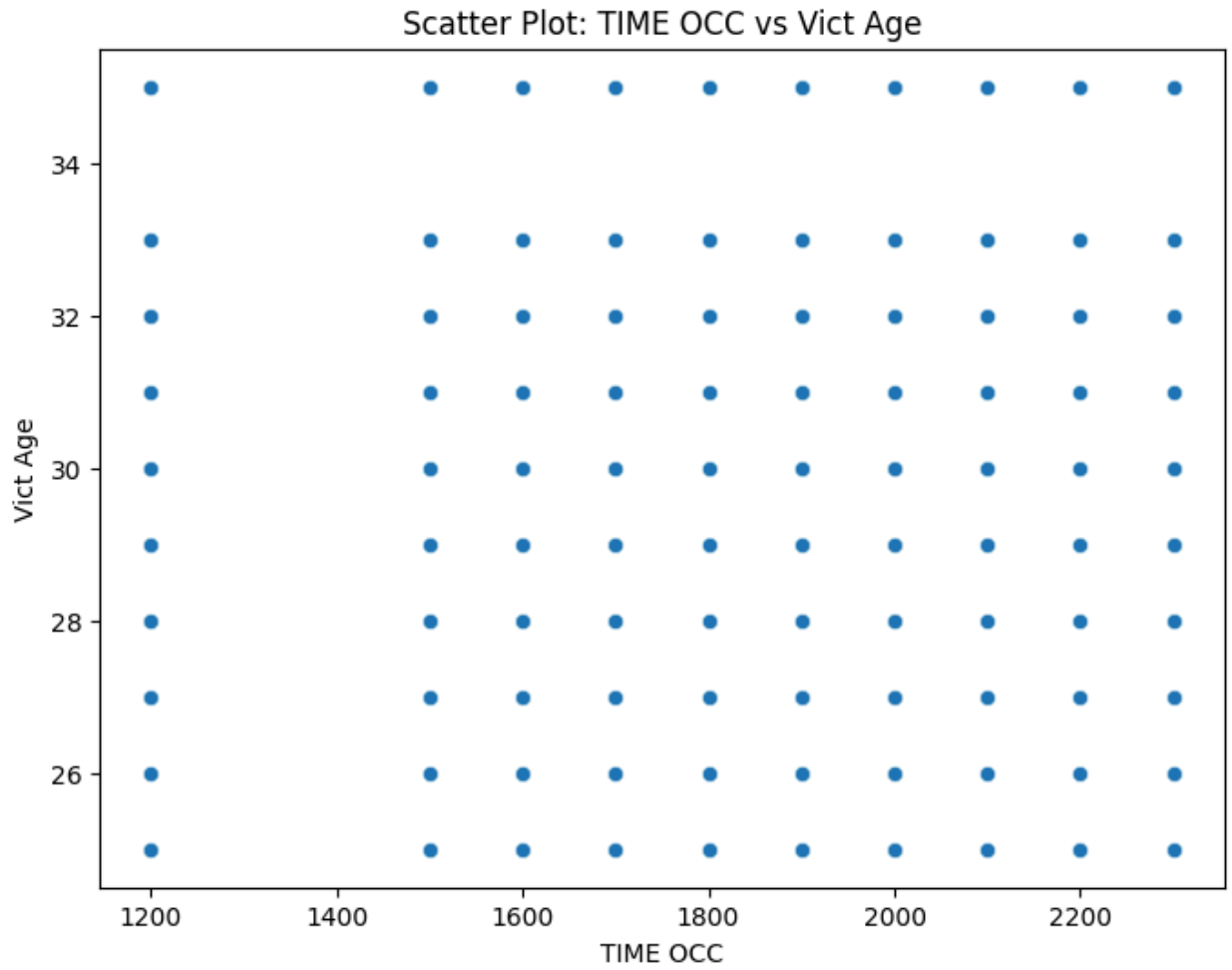
This plot shows the distribution per crime description of the age of the person who committed the crime.



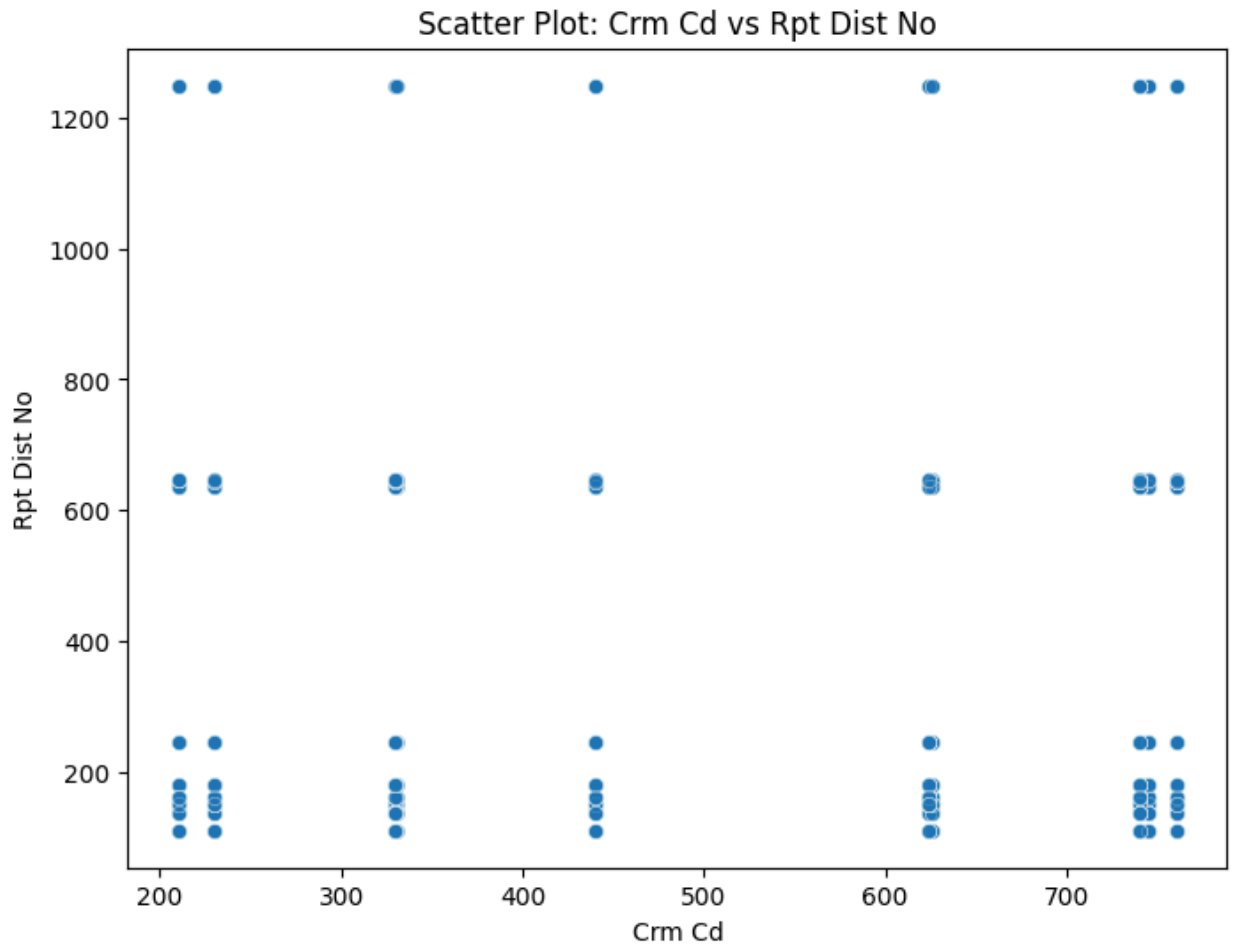
This plot shows the distribution per crime location of the age of the person who committed the crime.



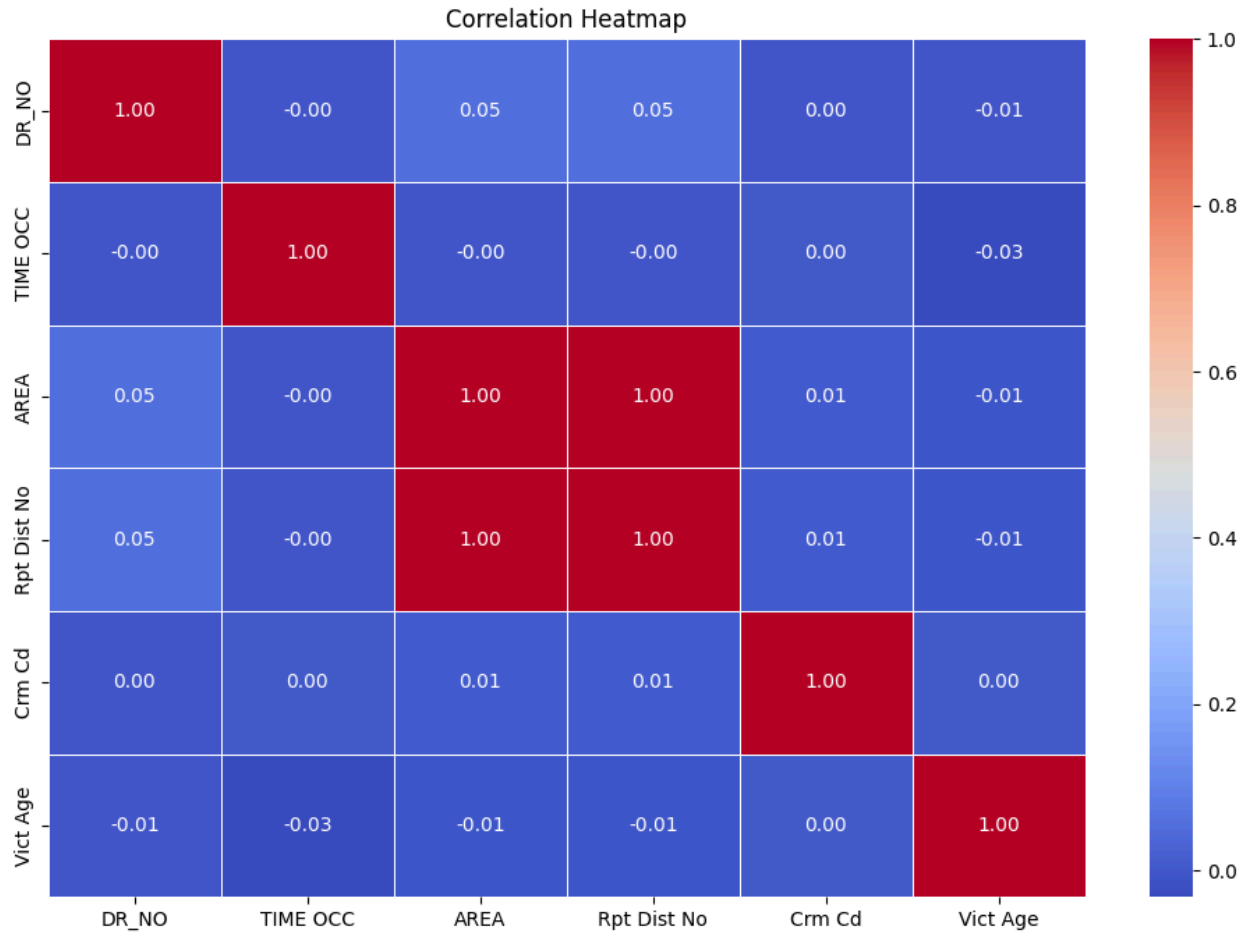
This plots DR_NO of a case against the time the crime occurs; evidently, there is no relationship.



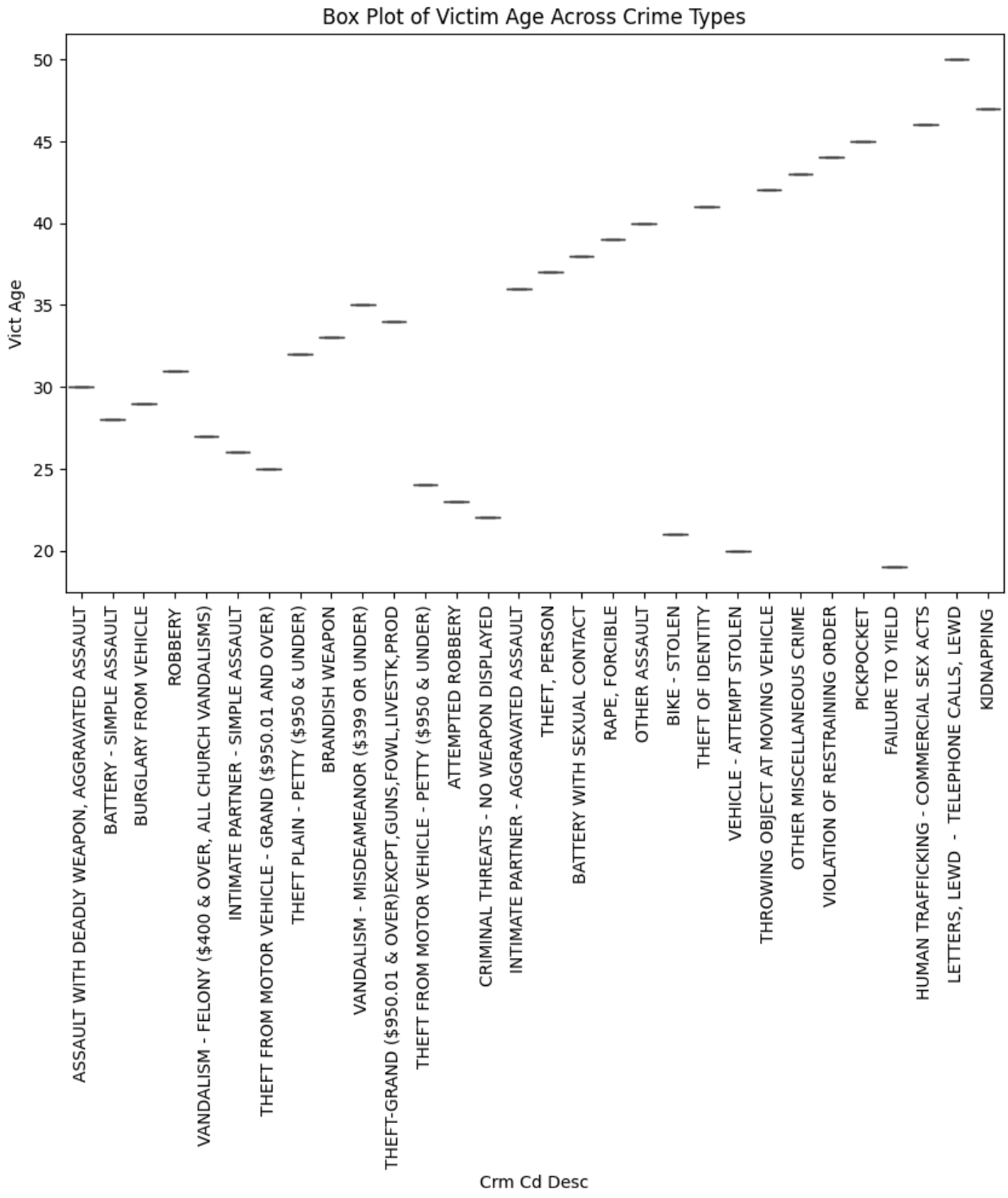
This plots the time the crime occurs against the age of the person who committed the crime.



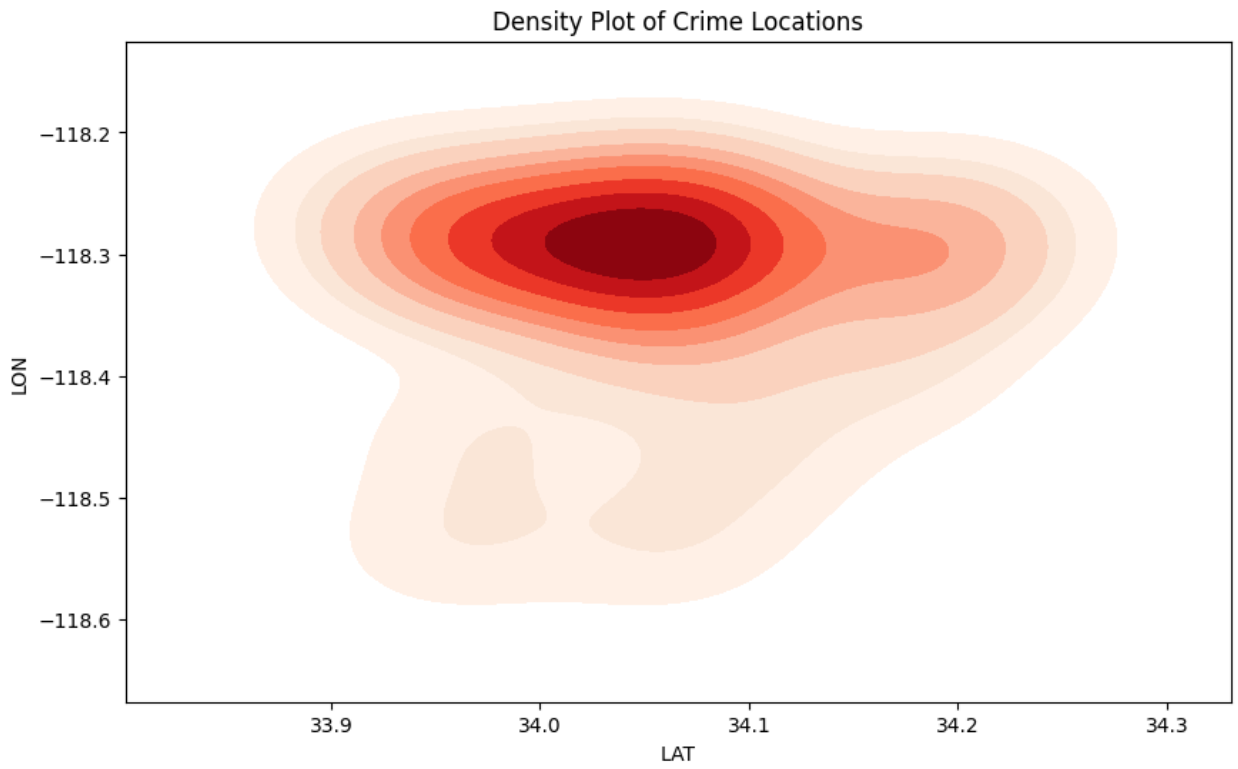
This plots the numerical code of the crime committed against the reporting district number of the crime.



This plot shows the correlation between DR_NO, crime occurrence time, crime occurrence area, reporting district number, code of crime committed, and the age of the person who committed the crime.



This plot shows the boxplot for the age of the person who committed the crime, organized by the type of crime committed.



This plot shows the distribution of crimes over the geographical region represented by the latitude and longitude scales.

Did you find any interesting patterns or directions you want to explore? We thought it was interesting that the center of the density plot for crime locations roughly aligned with downtown LA. Also, the scatterplots showed us no information.

What insights do you want to share about the project work? N/A

What questions do you have? N/A

Team Signatures:
Rohit Vakkalagadda
Kaitlyn Lu
Lucas He