DATA 603
Group project Presentation
on
Denver Crime data

TEAM MEMBERS:

HARSHA DAPARTI

LAXMI PRASANNA CHENNADI

SRINIVASA AKHIL VUTUKURI

Introduction:

- The data we took is crime data of Denver. The reasons of the crime, traffic crimes, place, time of the crime is recorded in the dataset

- This dataset includes criminal offenses in the City and County of Denver for the previous five calendar years plus the current year to date. The data is based on the National Incident Based Reporting System (NIBRS) which includes all victims of person crimes and all crimes within an incident
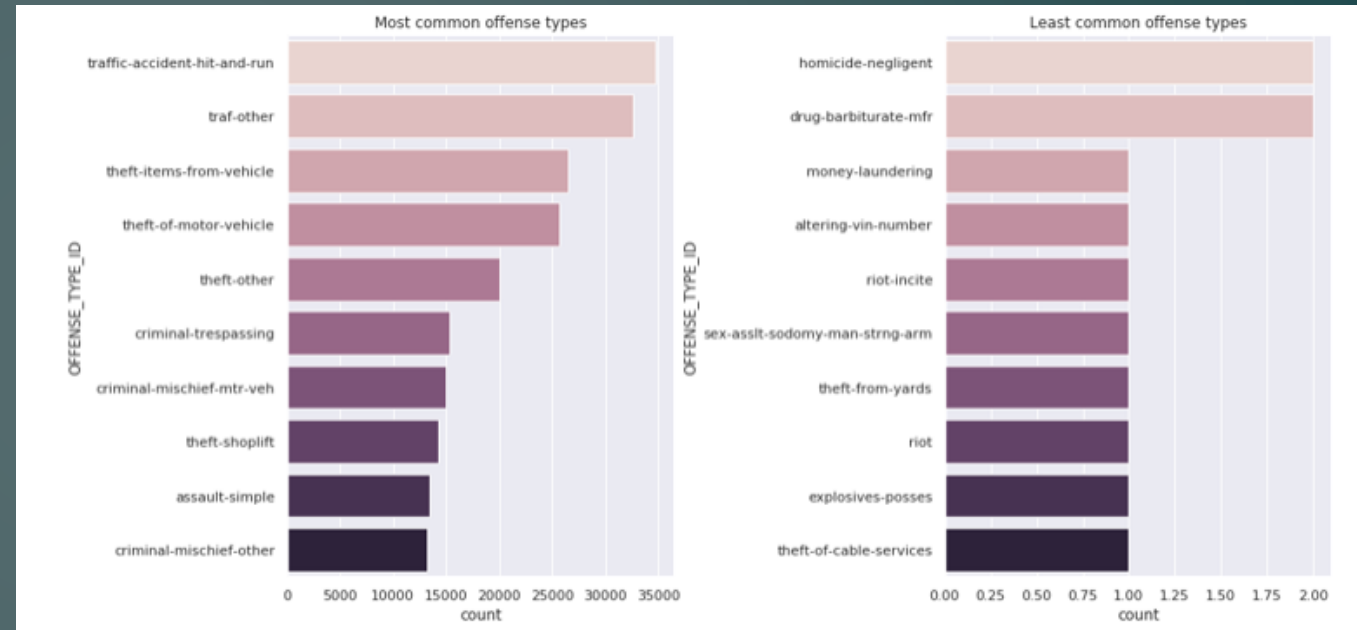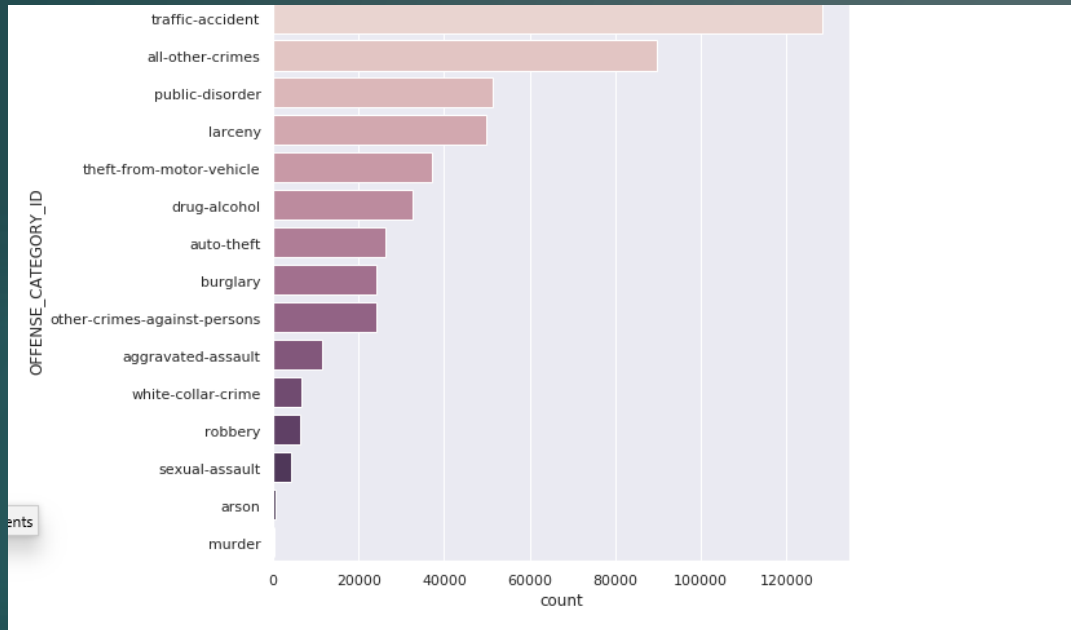
- Data Link : https://www.kaggle.com/paultimothymooney/denver-crime-data

About the Dataset

▶ 19 Columns in the data set namely –

▶ INCIDENT_ID, OFFENSE_ID, OFFENSE_CODE, OFFENSE_CODE_EXTENSION, OFFENSE_TYPE_ID, OFFENSE_CATEGORY_ID, FIRST_OCCURRENCE_DATE, LAST_OCCURRENCE_DATE, REPORTED_DATE, INCIDENT_ADDRESS, GEO_X, GEO_Y, GEO_LON, GEO_LAT, DISTRICT_ID, PRECINCT_ID, NEIGHBORHOOD_ID, IS_CRIME, IS_TRAFFIC

▶ Around half a million records

▶ For the EDA, the incidents were analyzed using First_Occurrence_Date and were categorized using Offense_Type_ID and Offense_Category_ID

Tools used

- Mongo DB
- Apache spark SQL
- Python
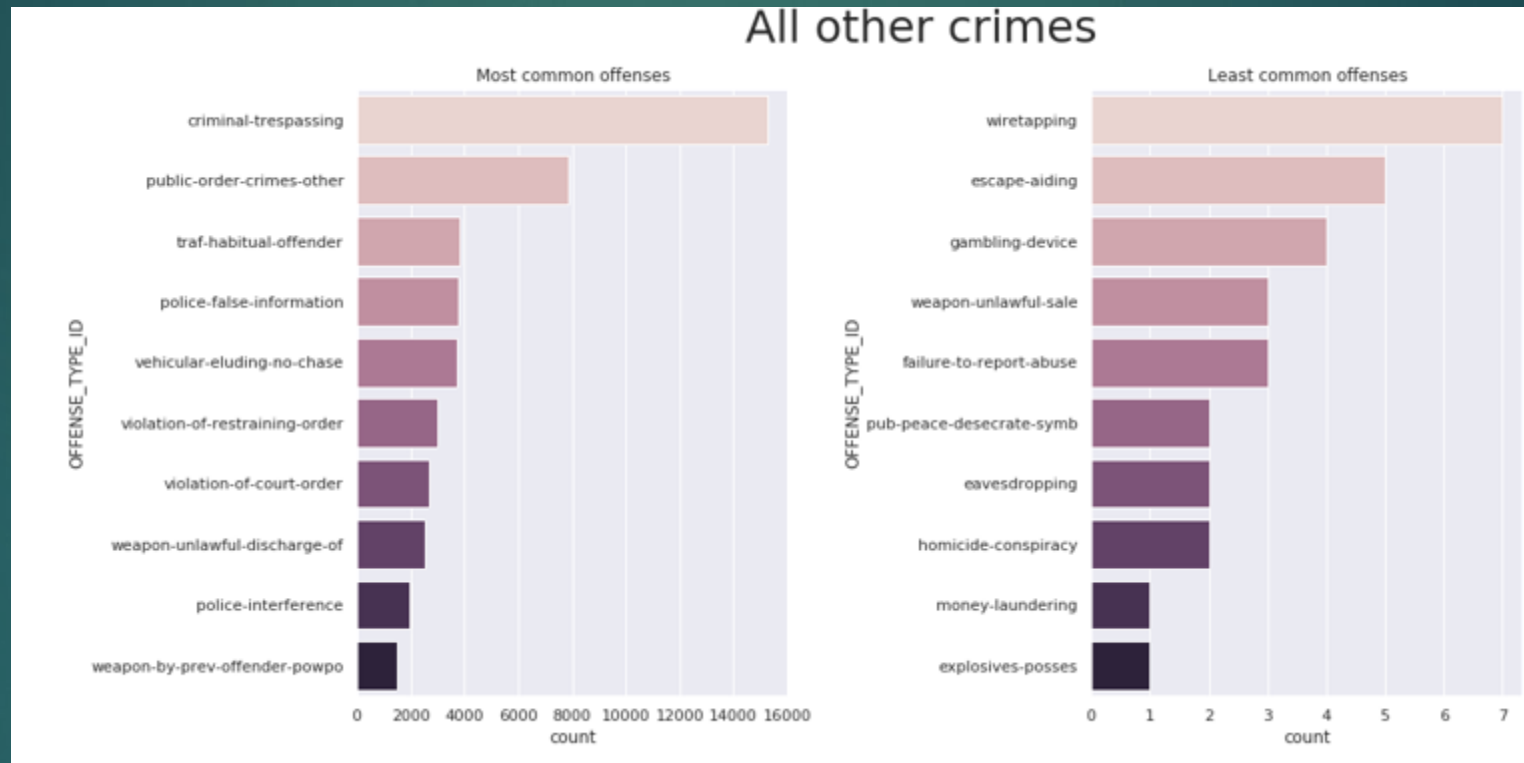- Anaconda (integration with jupyter notebook)

Results

- Excluding accidents most recorded crimes were public disorders
- Then its larceny (burglaries or thefts)
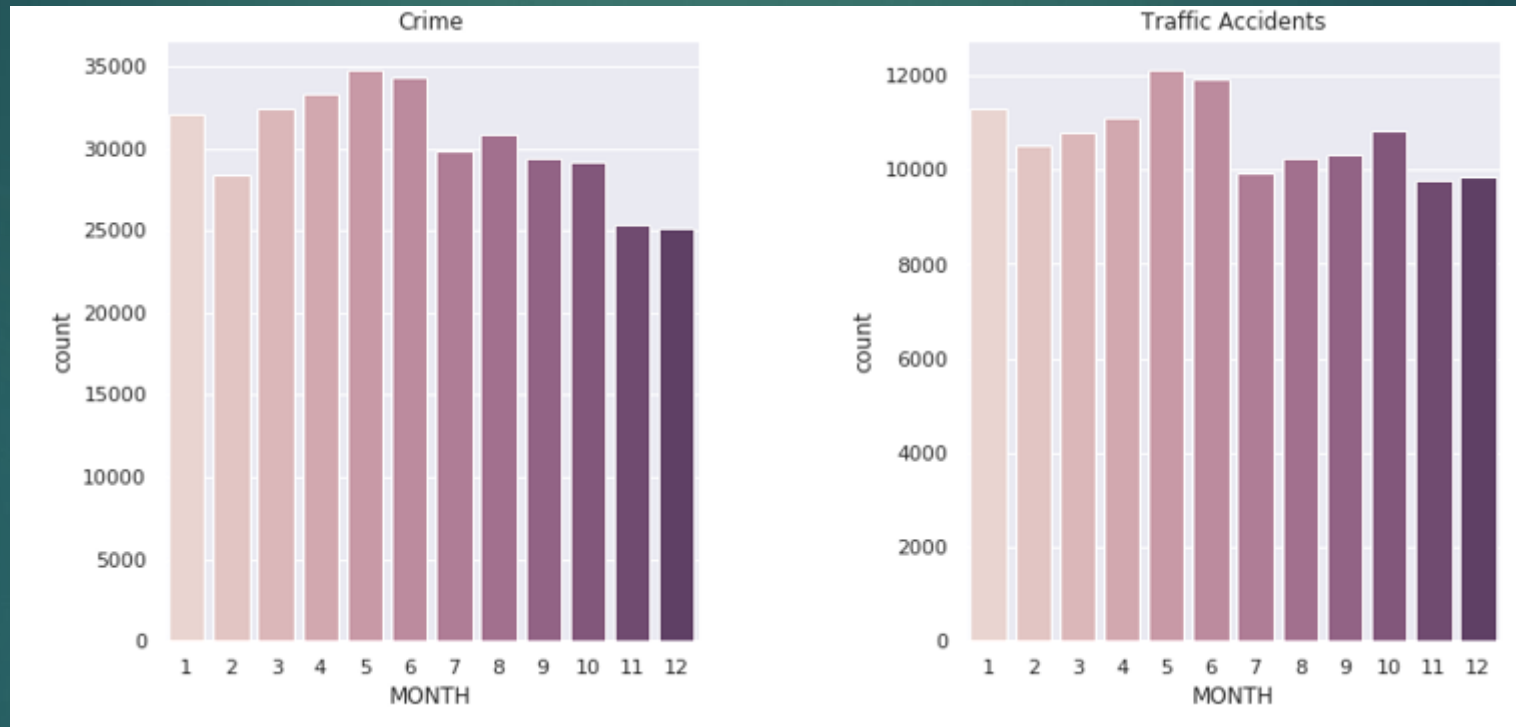- Murder is very rare in the chosen data

- Large number of vehicle related crimes
- Public disorder is seen quite a bit and there was one riot
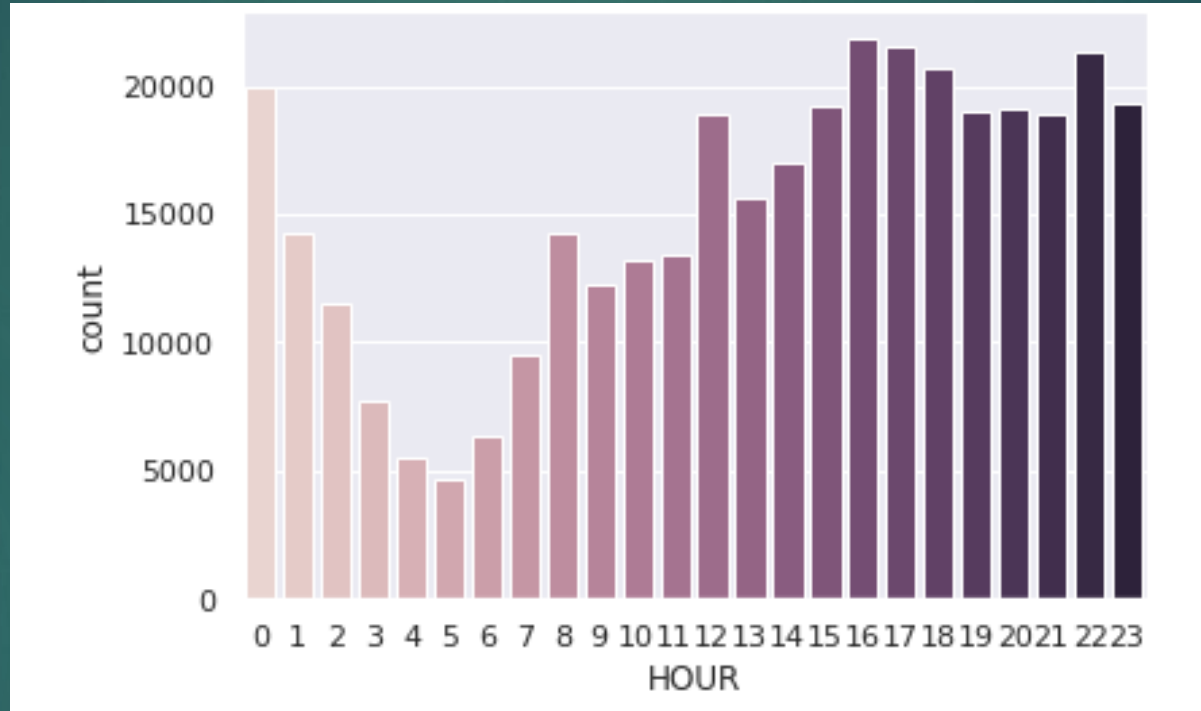- Theft of cable services?

Results extended…

- Among most common offenses criminal trespassing was prominent
- Possessing explosives and money laundering are the least common offenses
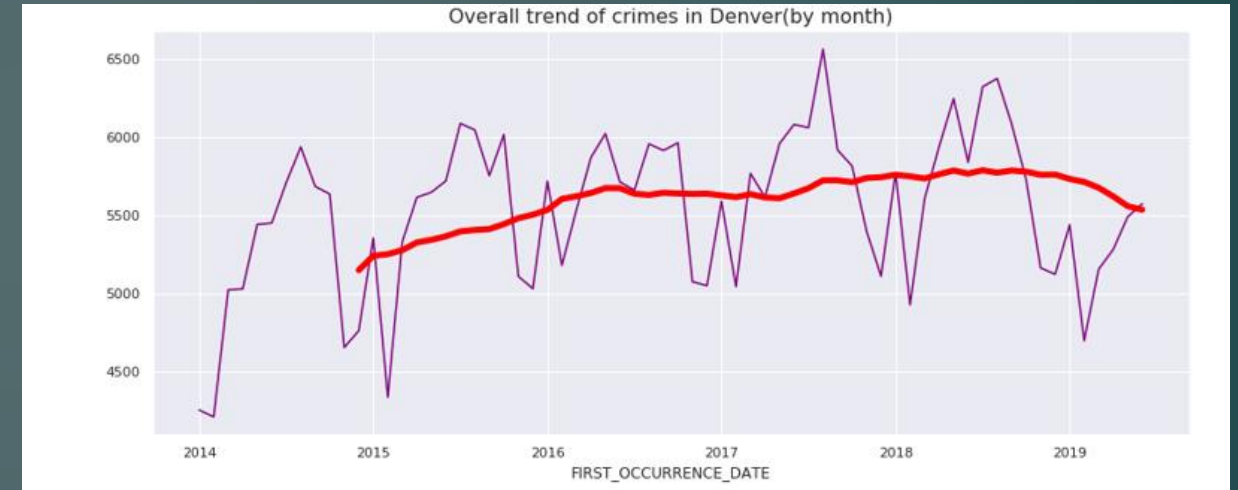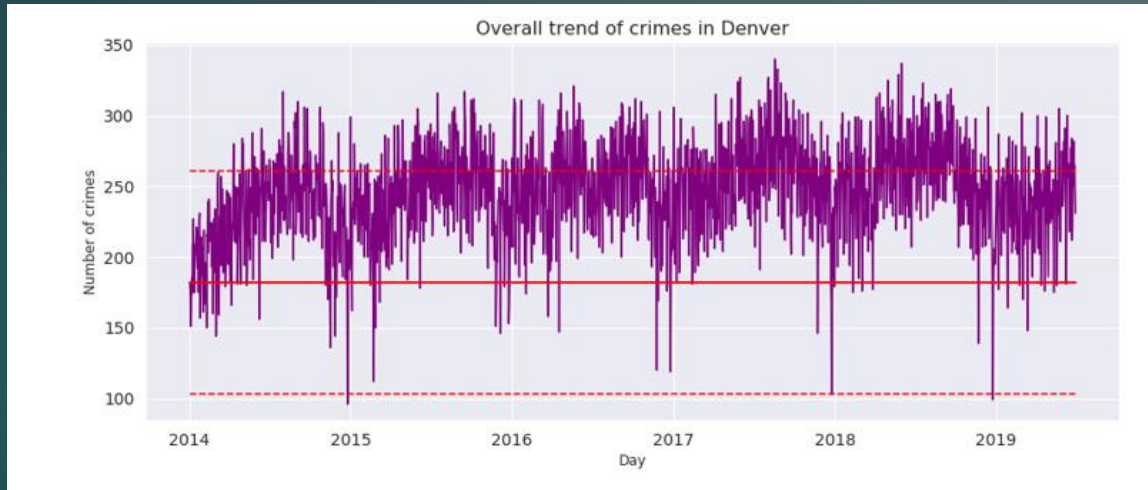
Results extended...

- The above visualizations were made using the binary attributes IS_CRIME and IS_TRAFFIC
- It can be observed that May and June have a spike in both traffic accidents and crimes.

Results extended…

- The most crimes occurred were recorded around 4-6PM and a spike is seen again at midnight.

Results extended…

- The overall trend seems to be increasing, and there's a slightly decrease in 2019.
- February 2014 is the best month with 4211 crimes. Also February of every year is observed to be the month with least number of incidents.

Learnings

▶ Using all the technologies together and creating a suitable environment was the biggest hurdle in this project.