Collisions in Seattle-USA

IBM Data Science Capstone Project

By. Sanoj Chathuranga Amarasiri

2020/09/05

Introduction

- In this project was conducted to study the relation between collisions and factors associated with those collisions.
- This study been targeted to residents and drivers in Seattle area in USA.
- What are the main factors causing collisions in Seattle?
- Is there ways to reduce collision count?

Data Set

- Respective data was collected from SDOT Traffic Management Division using following link:
  - https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Data-Collisions.csv
- As independent variables
  - Locations of collisions occurred
  - Junction type
  - Weather
  - Road condition
  - Light condition
  - In-attention of the driver
  - Use of drugs
- As dependent variable
  - Collision severity

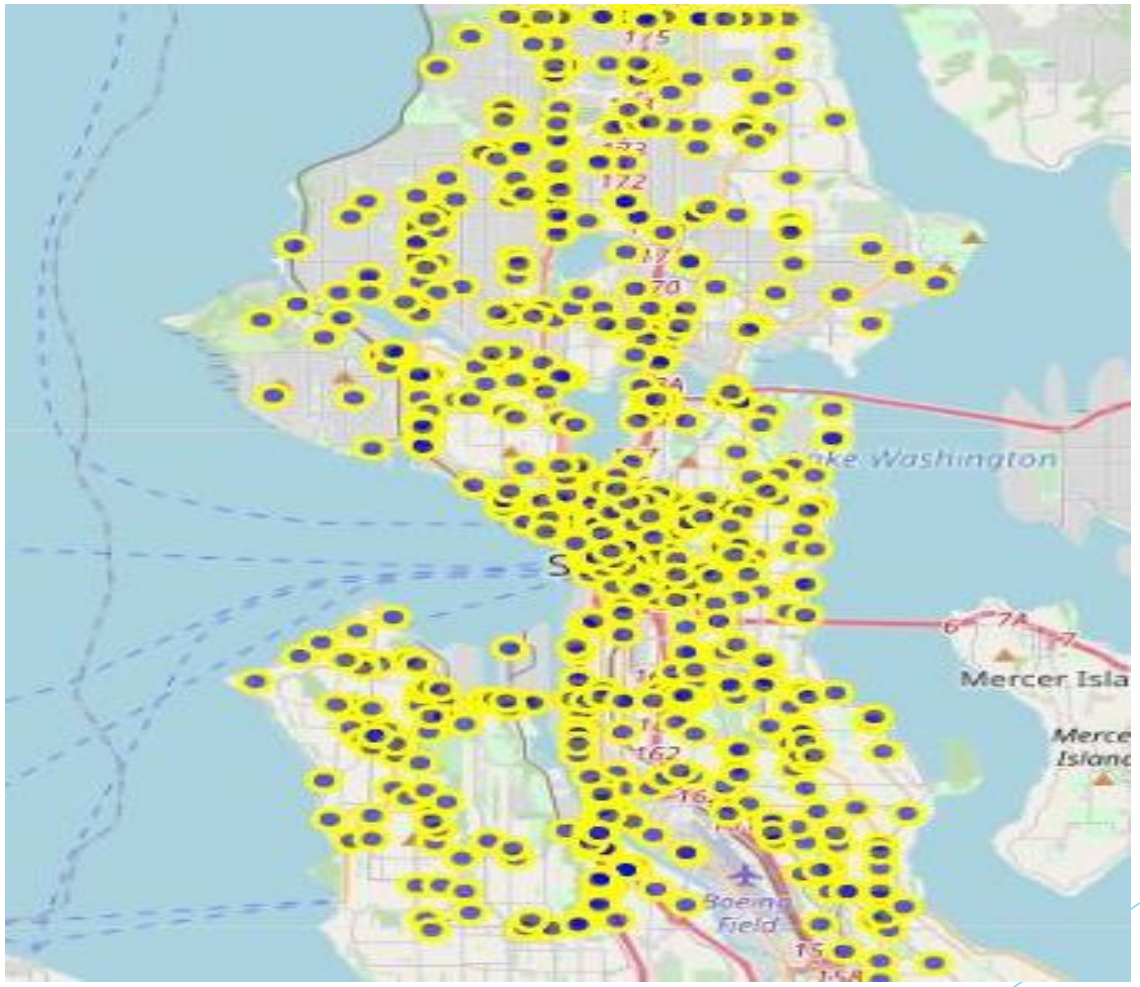
Methodology – Data pre-processing

- Data cleaning
  - Removed all empty rows which gives balanced data set. Furthermore, renamed columns according to requirement.

Longitude	Latitude	JUNCTIONTYPE	INATTENTIONIND	UNDERINFL	WEATHER	ROADCOND	LIGHTCOND	SPEEDING	SEVERITYDESC
-122.344947	47.720482	Driveway Junction	Y	N	Clear	Dry	Daylight	Y	Injury Collision
-122.328913	47.613466	At Intersection (intersection related)	Y	N	Clear	Dry	Daylight	Y	Property Damage Only Collision
-122.374417	47.519289	Mid-Block (not related to intersection)	Y	N	Clear	Dry	Daylight	Y	Injury Collision
-122.358307	47.653110	Mid-Block (not related to intersection)	Y	N	Snowing	Snow/Slush	Daylight	Y	Injury Collision
-122.321917	47.595831	Mid-Block (not related to intersection)	Y	N	Raining	Wet	Dark - Street Lights On	Y	Injury Collision

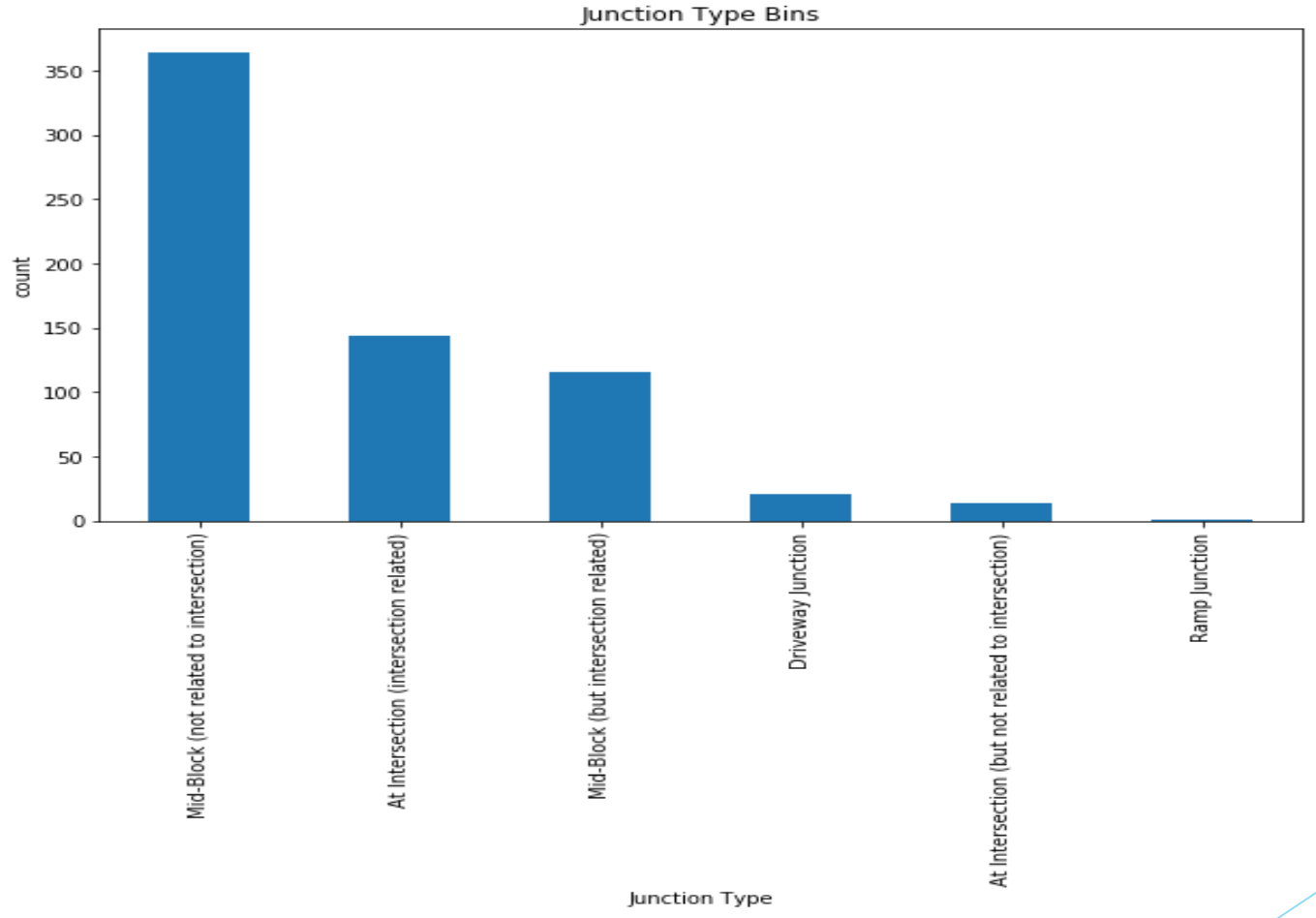
Methodology – Data Analysis

- Mapped collision coordinates in Seattle map in order to identify high intensity collision zones.



Methodology – Data Analysis Junction Type

According to bar chart Mid-Block (not related to intersection) & At Intersection (intersection related) cause more collisions.

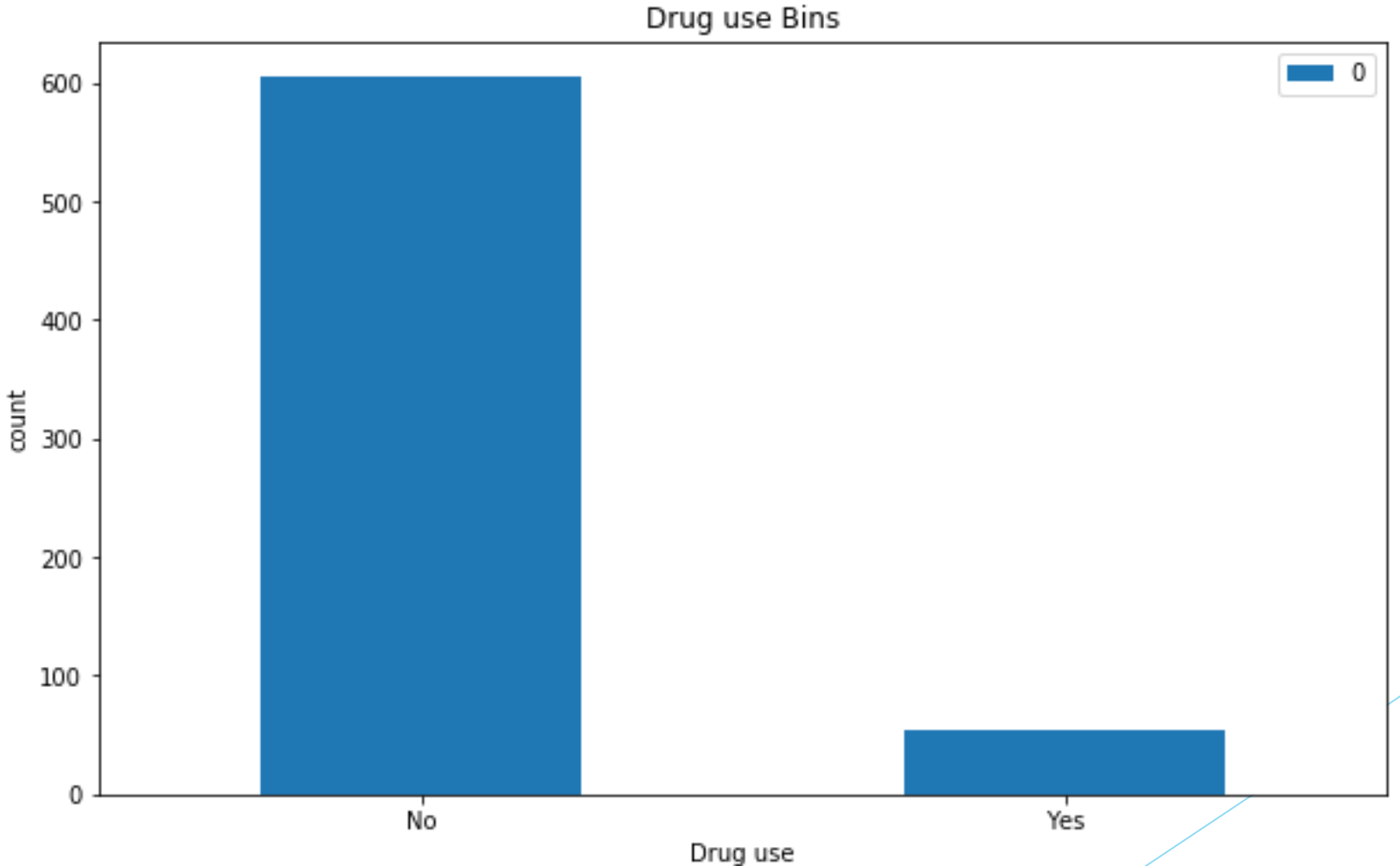


Methodology – Data Analysis In-attention

- Analyzed data set for collision count with regard to the in-attention of driver.
- According to the result all the collisions were occurred due to in-attention of the driver.

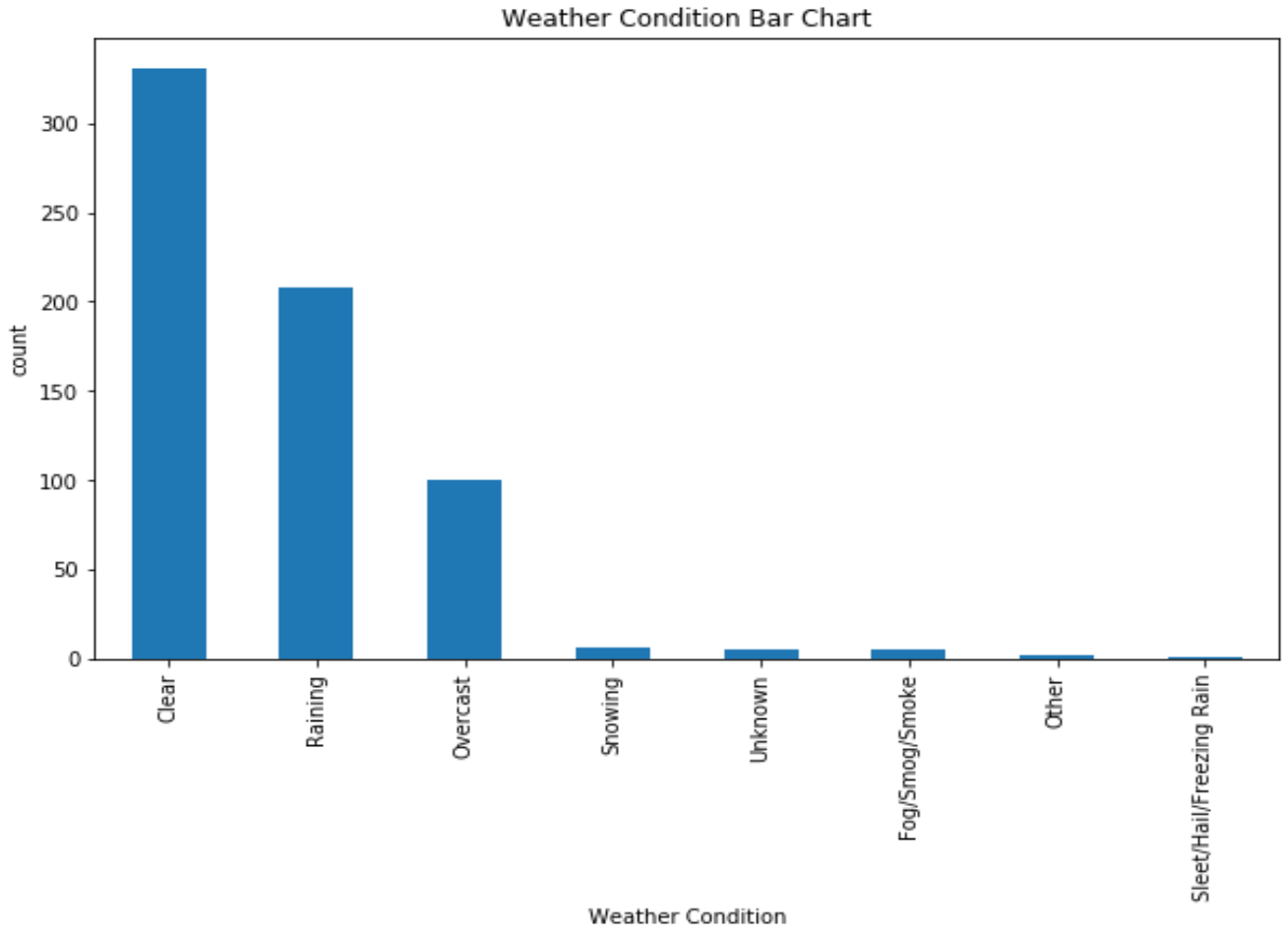
Methodology – Data Analysis-Drug Use

- It appears to be there isn't much impact of drug use for collisions in this dataset.



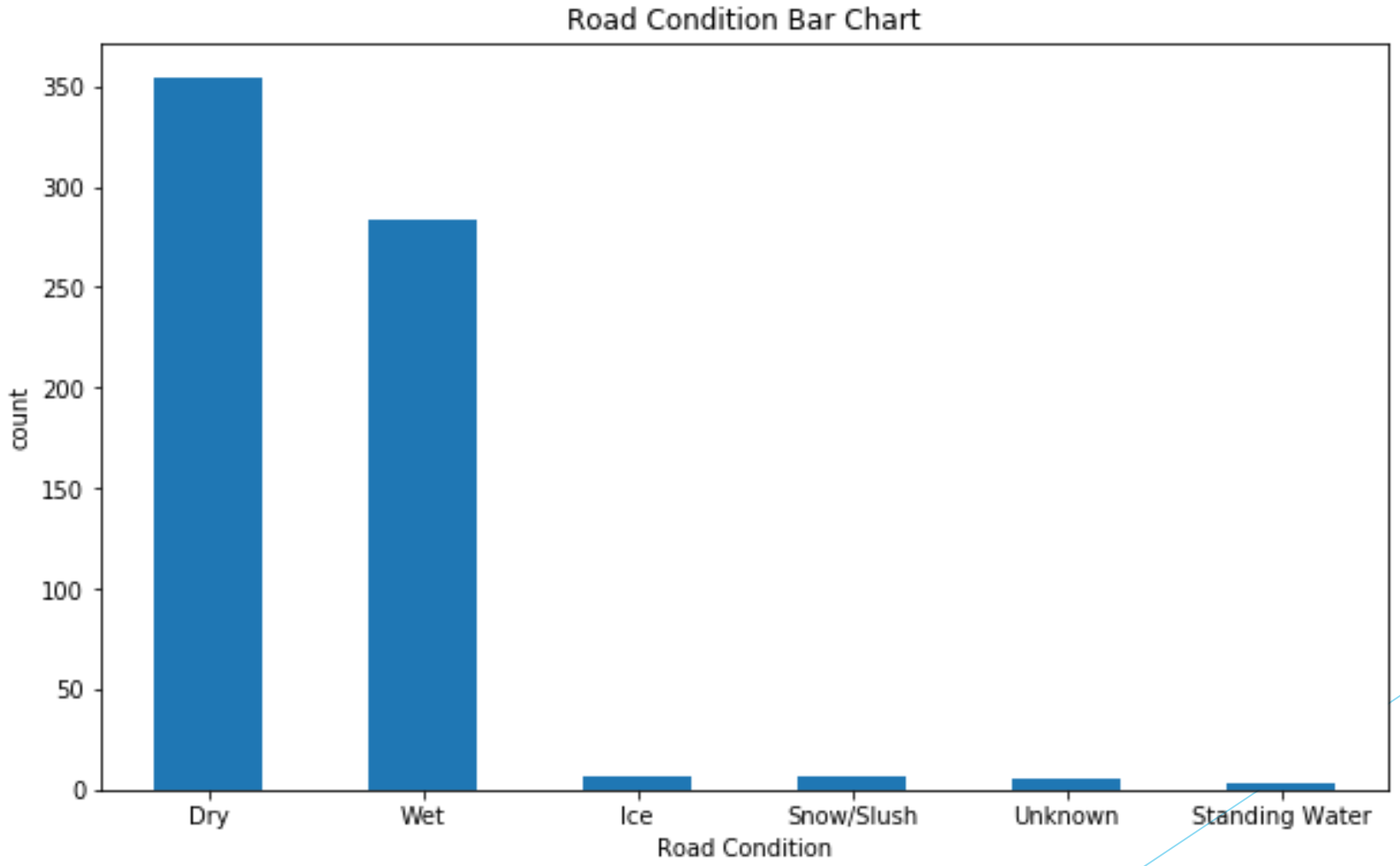
Methodology – Data Analysis-Weather

- Most collisions occurred in Clear & Raining weather conditions.



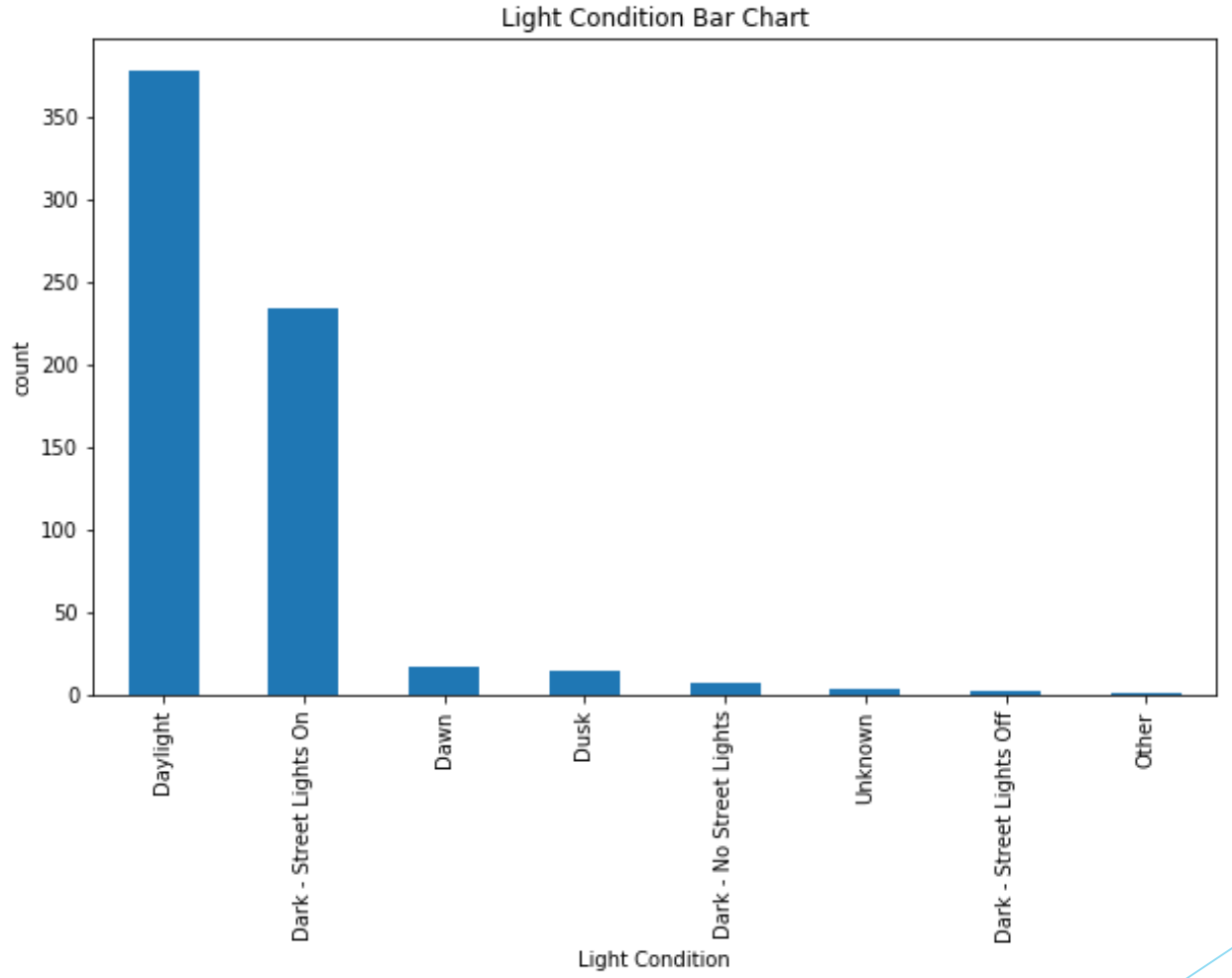
Methodology – Data Analysis-Road

- Most collisions occurred during Dry and followed by Wet conditions.



Methodology – Data Analysis-Light

- Most collisions occurred during Daylight and Dark-Street Lights On conditions



Results & Discussion

- Most of cases occurred in Mid-block junctions
- All collisions have been occurred due to in-attention when driving
- There isn’t much impact of drug use for collisions in this dataset
- Most of the collisions occurred during Clear weather conditions and there is an impact by Rainy conditions
- Most of collisions occurred during the dry and followed by wet road conditions
- Most of collisions occurred during day light and followed by Dark and Street lights on conditions rather than completely dark conditions

Conclusion

- All the collisions occurred mainly due to Speeding and In-attention
- Rather than bad Weather, Road and Light conditions most collisions occurred in good conditions
- In good driving conditions drivers pay less attention to driving and tend to speed which cause more venerability for collision
- By taking necessary actions to notify drivers about the collision risk at Mid-block junctions which tend to reduce almost 50% of collisions.

Thank You