

CRIME REPORT



for 3 Major Cities of the U.S.

Group 5

Tanmai Aurangabadkar

Neha Gupta

Era Singh Kajal

Ying Ying Lai

Professor

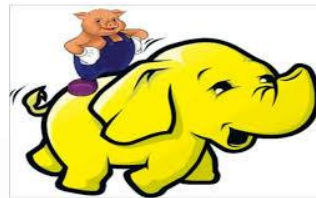
Jongwook Woo

Introduction/ Overview

- Crime situation in the United States has always been an issue
- Reason being: freedom to own weapons, poor economic situation
- Los Angeles, New York and Chicago tops all major U.S. cities in crime issue

Architecture WorkFlow

NYC OpenData



Raw Data
from Source

Basic Reverse
Geocoding

Store
Enriched Data
in HDFS

Analyze Data
using Hive &
Pig

Visualize Data
in Tableau

kaggle



H/W Experimental Specs

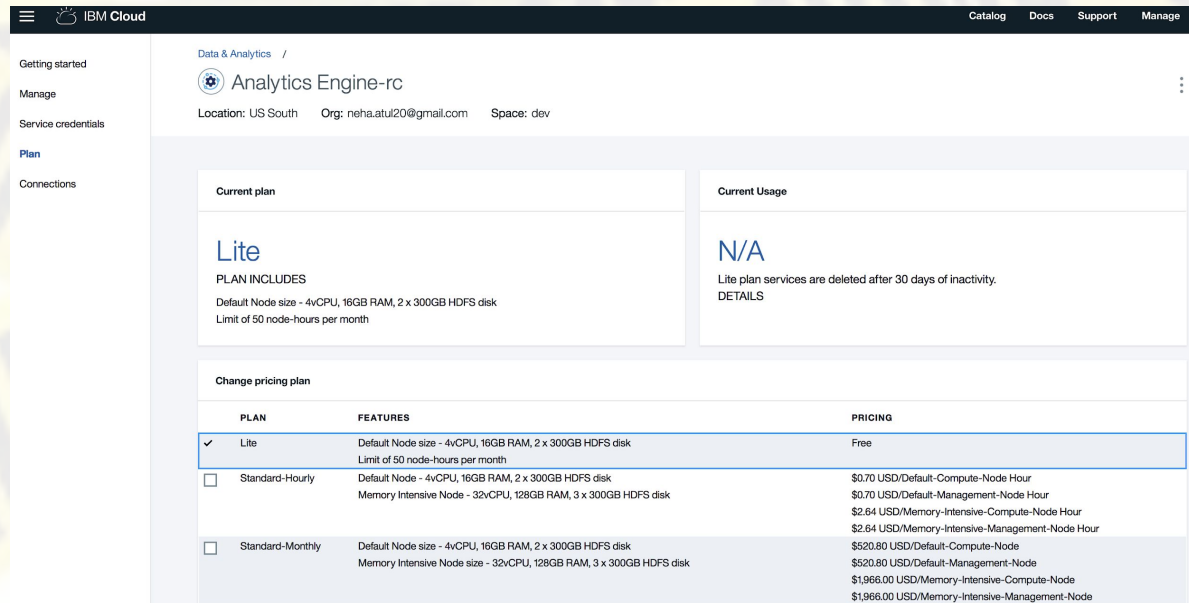
Cluster Version: IBM Analytics Engine

Number of Nodes: 2

Memory Size: 16 GB x2

CPU : 4vcpu x2

HDFS Disk : 600 GB x2



The screenshot displays the IBM Cloud console for the 'Analytics Engine-rc' cluster. The left sidebar contains navigation links: 'Getting started', 'Manage', 'Service credentials', 'Plan', and 'Connections'. The main content area shows the 'Current plan' as 'Lite' with the following details:

- PLAN INCLUDES
- Default Node size - 4vCPU, 16GB RAM, 2 x 300GB HDFS disk
- Limit of 50 node-hours per month

The 'Current Usage' section shows 'N/A' with a note: 'Lite plan services are deleted after 30 days of inactivity. DETAILS'.

Below this, the 'Change pricing plan' section features a table with three columns: PLAN, FEATURES, and PRICING.

PLAN	FEATURES	PRICING
<input checked="" type="checkbox"/> Lite	Default Node size - 4vCPU, 16GB RAM, 2 x 300GB HDFS disk Limit of 50 node-hours per month	Free
<input type="checkbox"/> Standard-Hourly	Default Node - 4vCPU, 16GB RAM, 2 x 300GB HDFS disk Memory Intensive Node - 32vCPU, 128GB RAM, 3 x 300GB HDFS disk	\$0.70 USD/Default-Compute-Node Hour \$0.70 USD/Default-Management-Node Hour \$2.64 USD/Memory-Intensive-Compute-Node Hour \$2.64 USD/Memory-Intensive-Management-Node Hour
<input type="checkbox"/> Standard-Monthly	Default Node size - 4vCPU, 16GB RAM, 2 x 300GB HDFS disk Memory Intensive Node size - 32vCPU, 128GB RAM, 3 x 300GB HDFS disk	\$520.80 USD/Default-Compute-Node \$520.80 USD/Default-Management-Node \$1,966.00 USD/Memory-Intensive-Compute-Node \$1,966.00 USD/Memory-Intensive-Management-Node

Data Size

3.1 GB

Data size > 3GB (GigaByte):

Extra credit 1.5 out of 100% (1.5 points = 3GB x 0.5)

```
-bash-4.1$  
-bash-4.1$ du -sh project-5200/  
3.1G    project-5200/  
-bash-4.1$ du -sh project-5200/*  
354M    project-5200/Crime_Data_from_2010_to_Present.csv  
1.5G    project-5200/Crimes_-_2001_to_present.csv  
1.4G    project-5200/NYPD_Complaint_Data_Historic.csv  
-bash-4.1$
```


Raw Data Source: Dataset URLs

Los Angeles

<https://data.lacity.org/A-Safe-City/Crime-Data-from-2010-to-Present/y8tr-7khq>

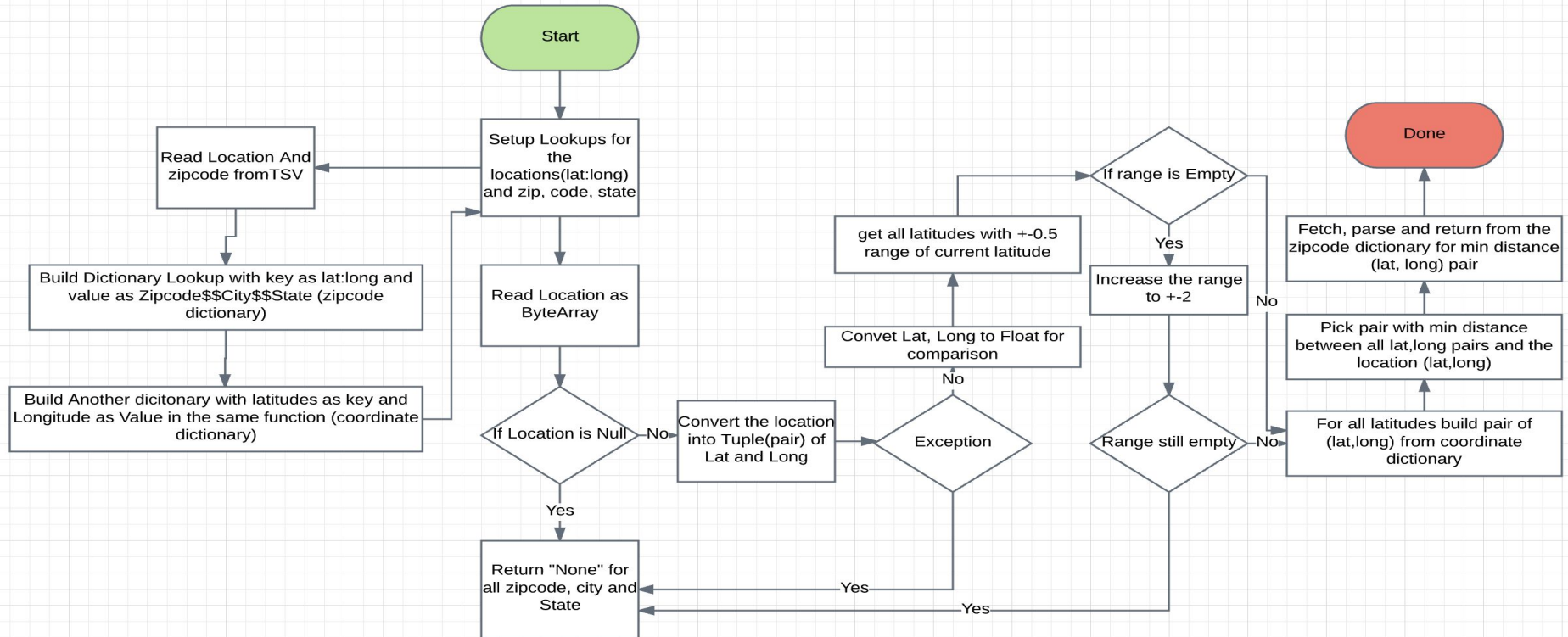
New York

<https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i>

Chicago

<https://www.kaggle.com/currie32/crimes-in-chicago/data>

Basic Reverse Geocoding FlowChart



A background image showing the Statue of Liberty in the foreground on the left, holding a torch and a tablet. Behind her is the dense skyline of New York City, including many skyscrapers like the Chrysler Building. The scene is set against a blue sky with some clouds, and the water of New York Harbor is visible at the bottom.

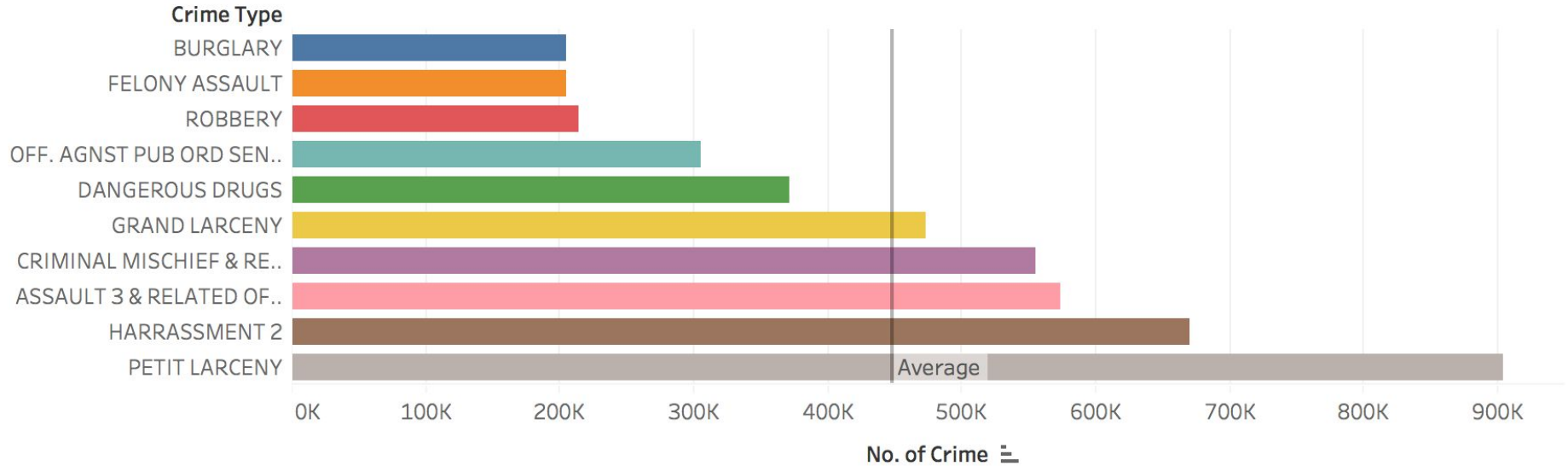
New York

Facts about New York Crime Situation

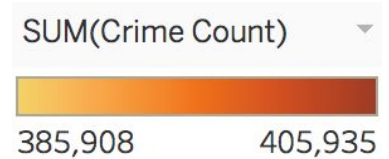
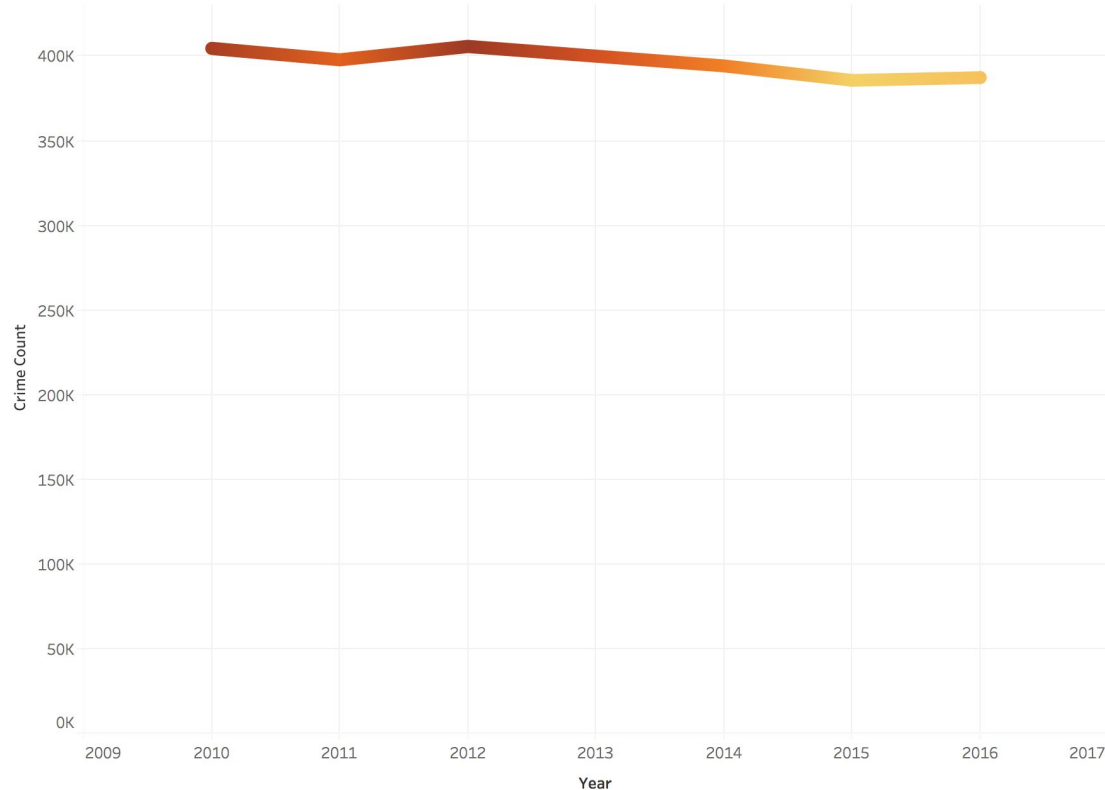
- The overall crime rate in NY is 28% lower than national average
- For every 100,000 people, there are 5.58 daily crimes that occur in NY
- NY is safer than 22% of the cities in the United States
- In NY, you have a 1 in 50 chance of becoming a victim of any crime

(Provide proof/reference in term paper)

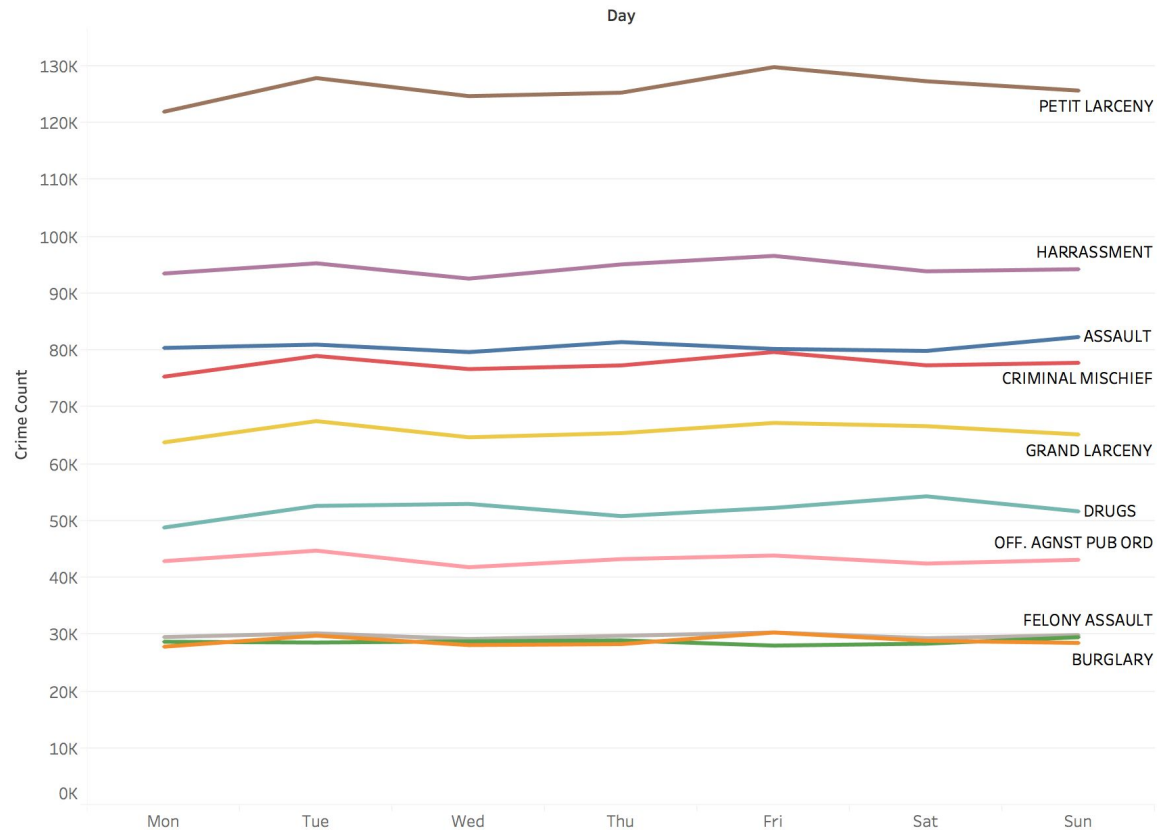
New York Top 10 Crime Types



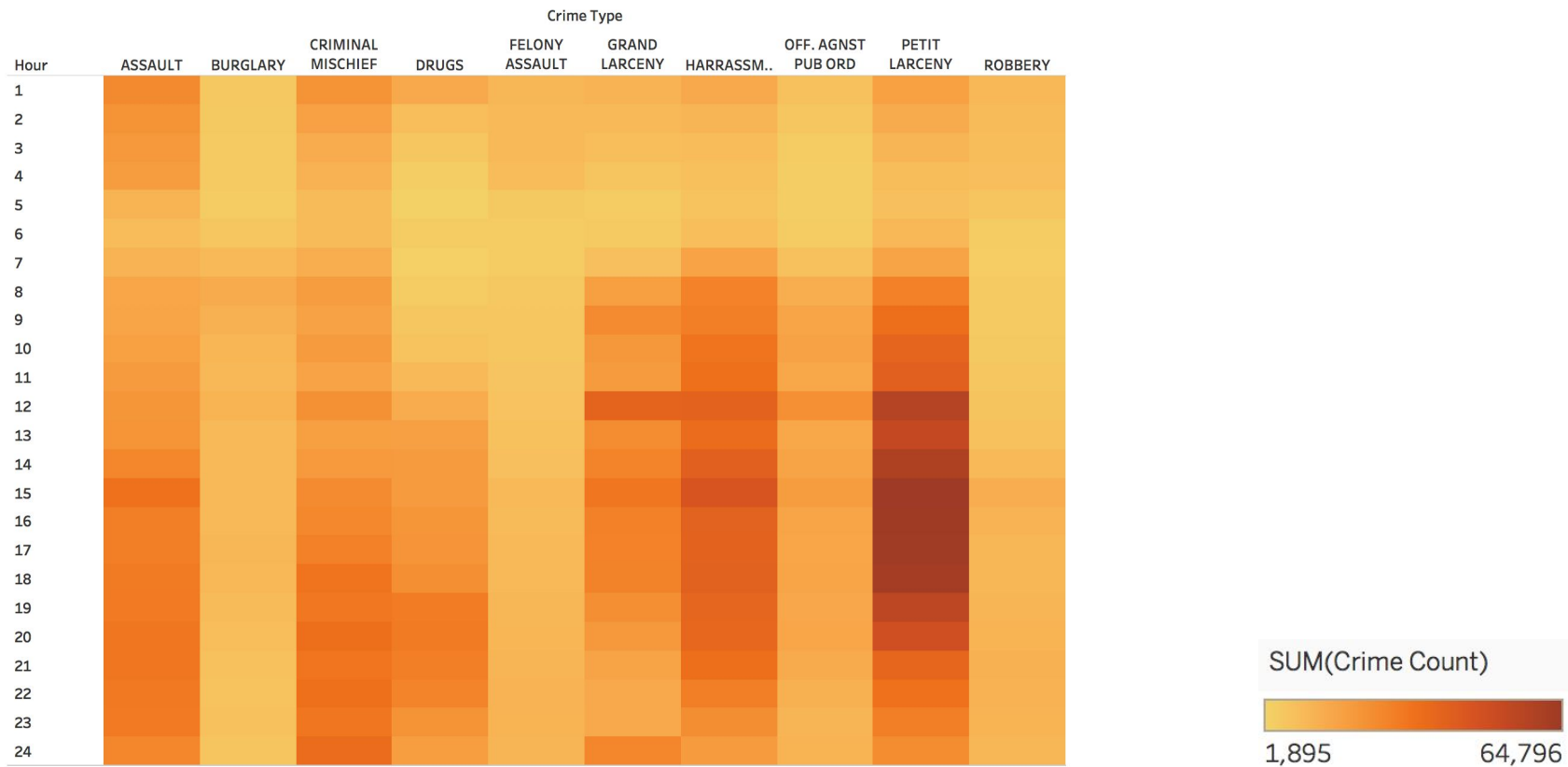
New York Crime Overview, by Year



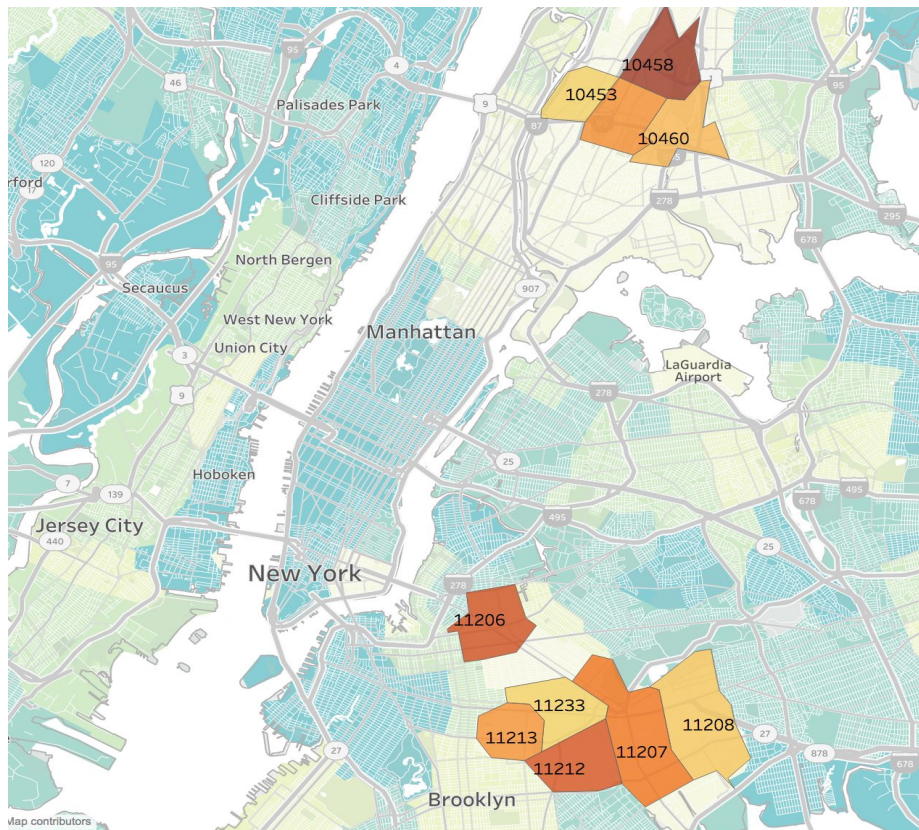
New York Crime Overview, by Day



New York Crime Overview, by Hour



New York Crime Overview, by ZipCode



SUM(Crime Count)



71,578

99,800

2017 Household Income...

- 0 to 40,800
- 40,800 to 48,200
- 48,200 to 56,600
- 56,600 to 68,800
- 68,800 to 501,000

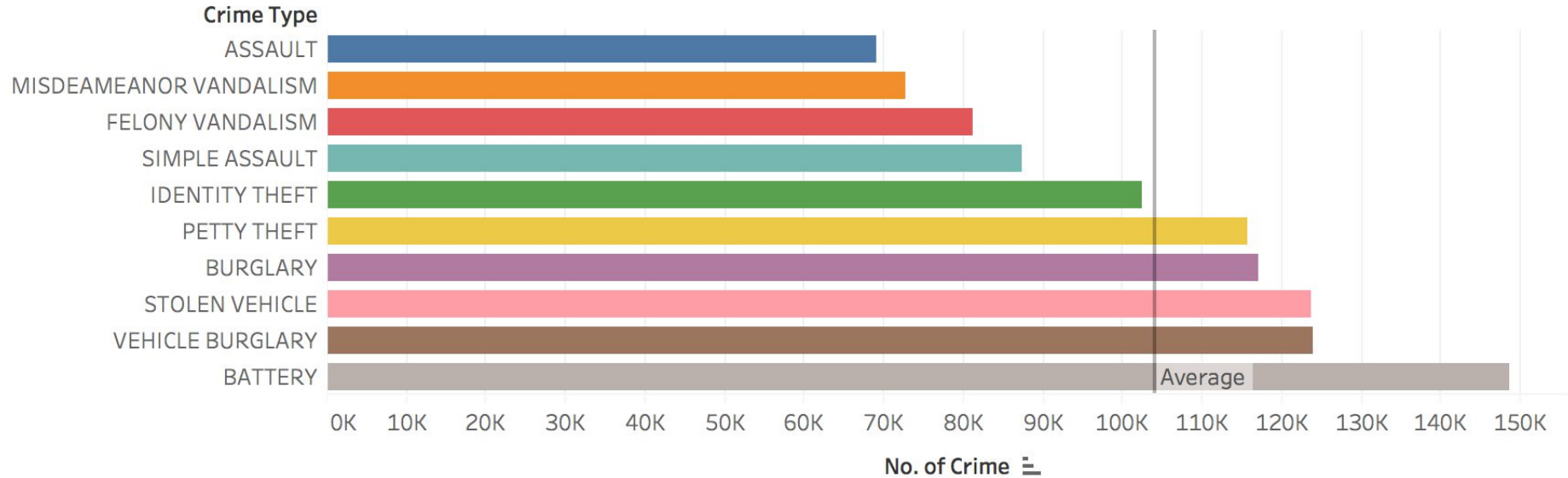
Los Angeles

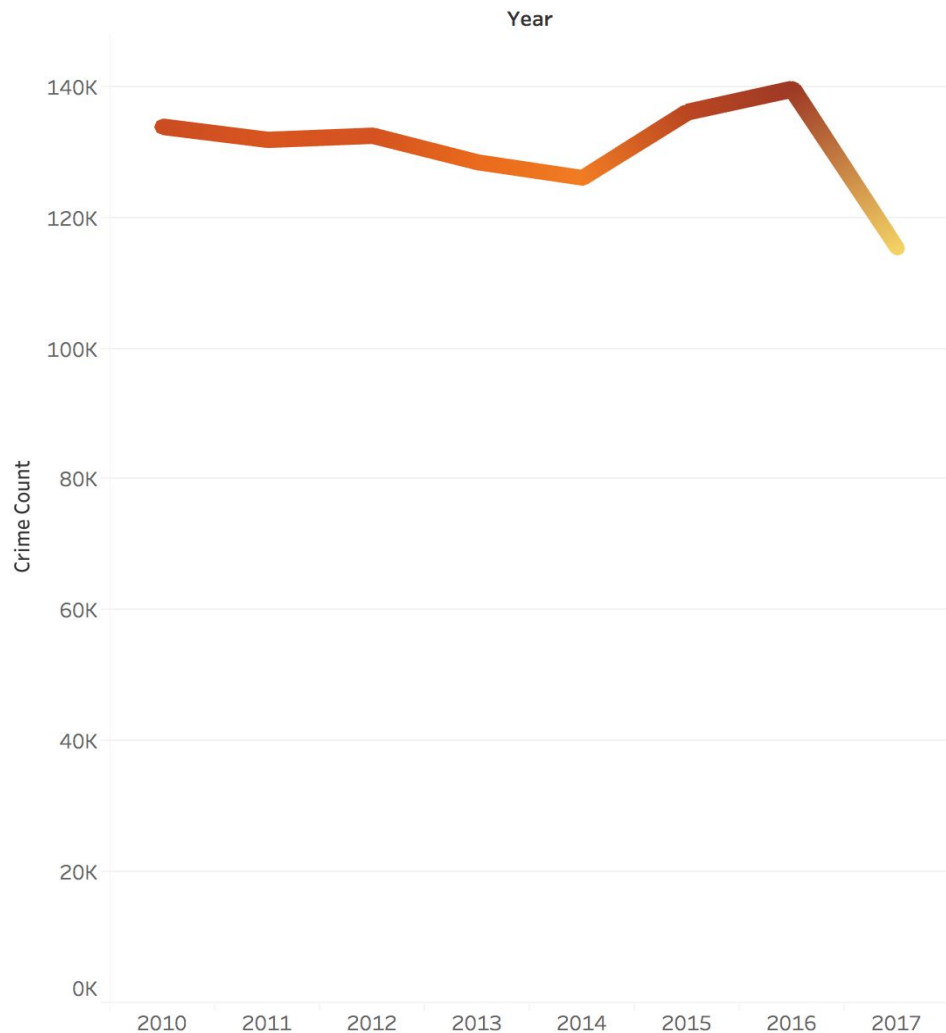


Facts about Los Angeles Crime Situation

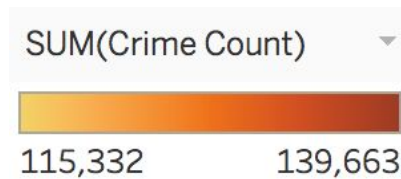
- The overall crime rate in LA is 13% higher than national average
- For every 100,000 people, there are 8.75 daily crimes that occur in LA
- In LA, you have a 1 in 32 chance of becoming a victim of any crime
- The number of total year over year crimes has increased by 7%

Los Angeles Top 10 Crime Types

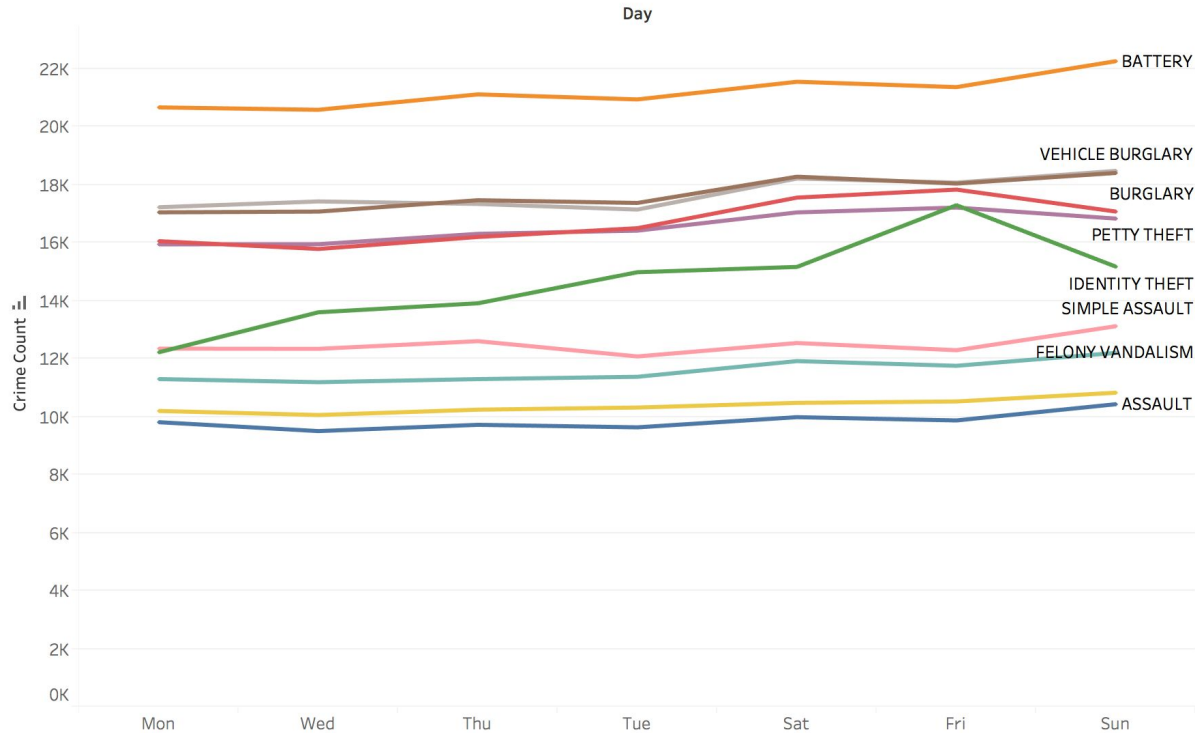




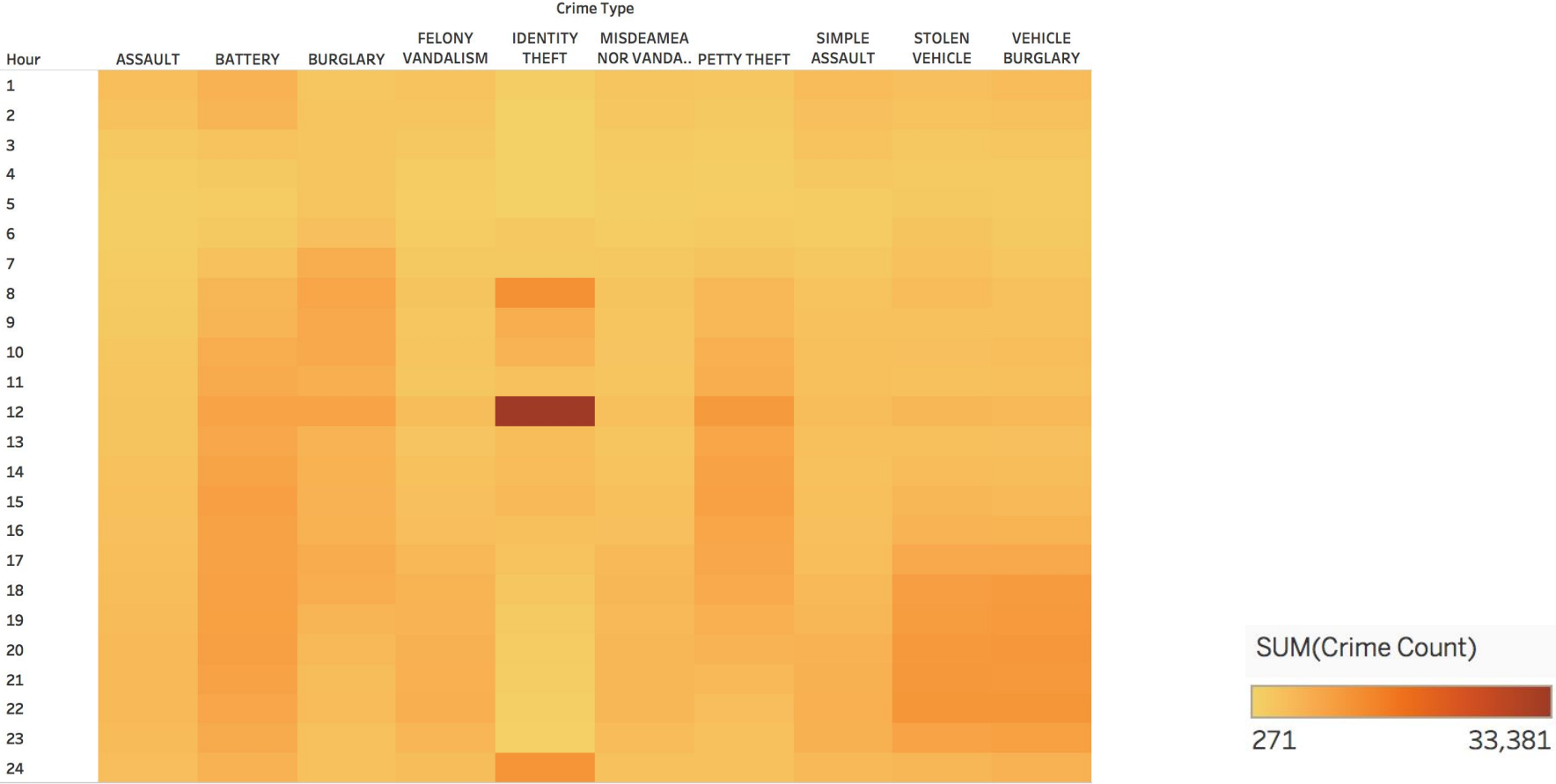
Los Angeles Crime Overview, by Year



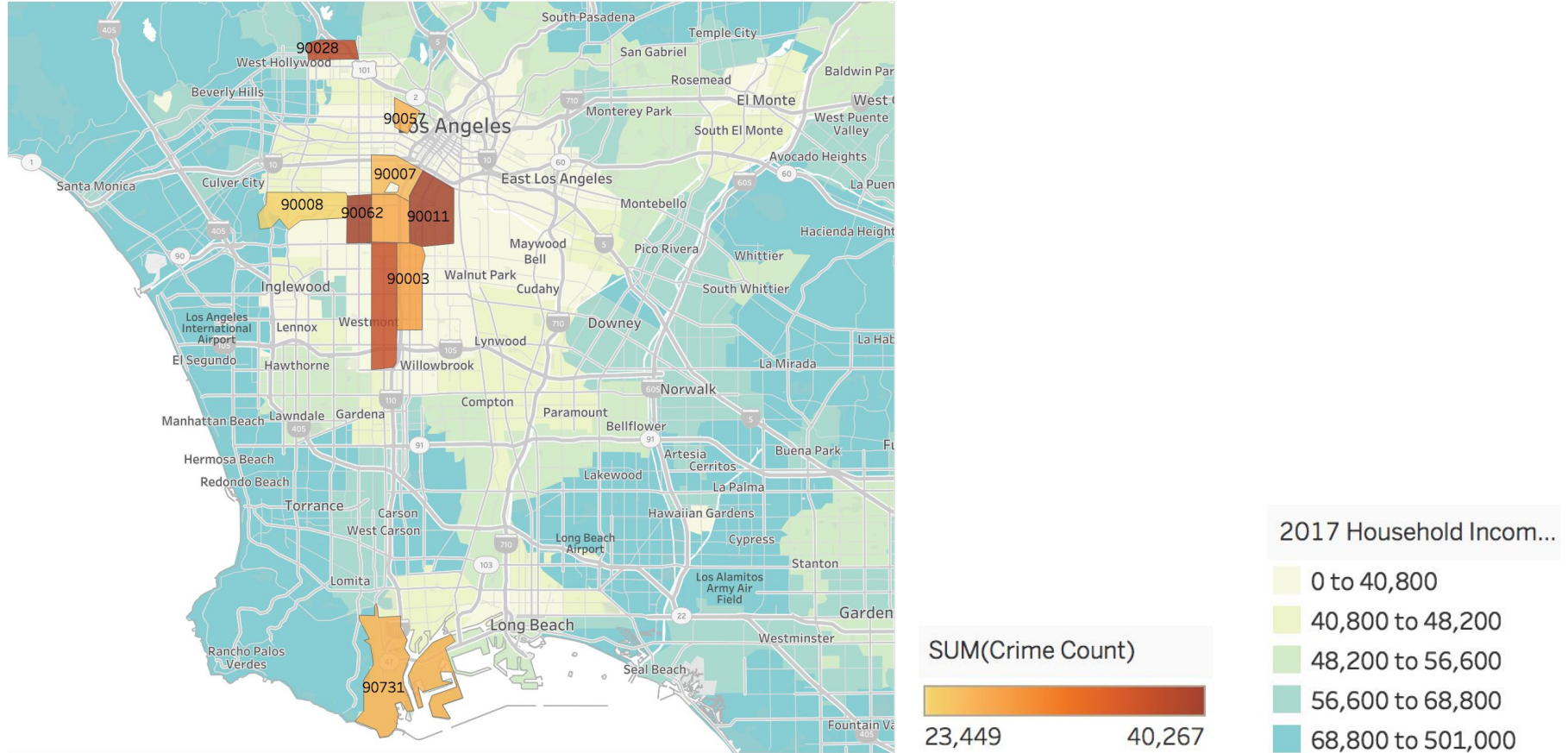
Los Angeles Crime Overview, by Day



Los Angeles Crime Overview, by Hour



Los Angeles Crime Overview, by ZipCode



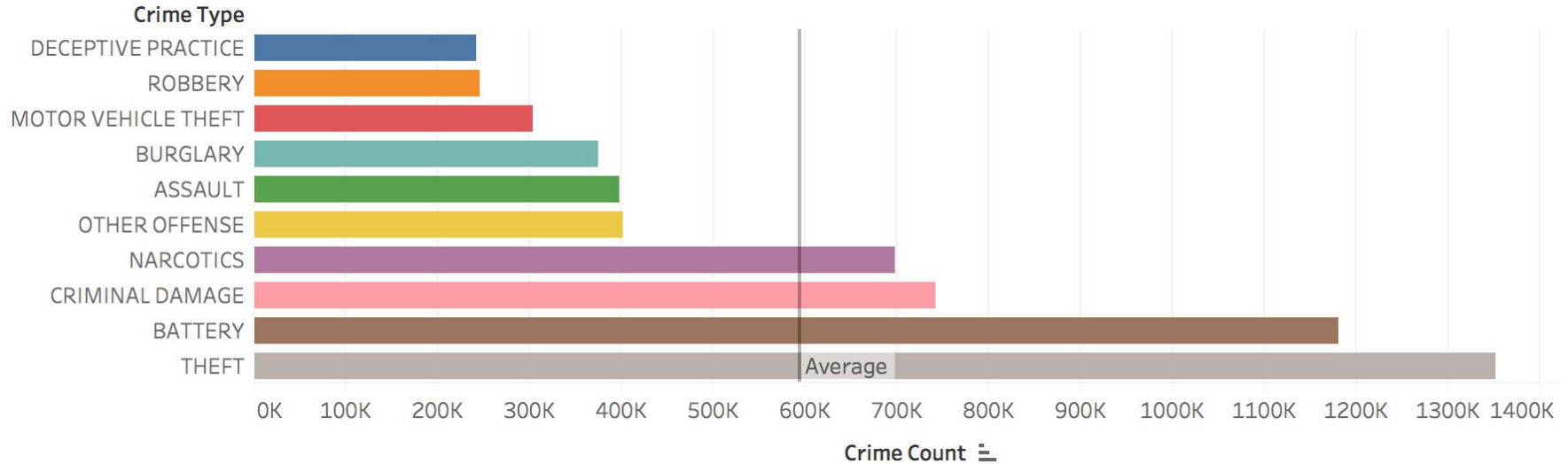
A wide-angle photograph of the Cloud Gate sculpture, also known as 'The Bean', in Chicago. The sculpture is a large, bean-shaped, highly reflective metallic structure that mirrors the surrounding city skyline and the sky. It sits on a paved plaza. In the background, several tall skyscrapers of the Chicago skyline are visible under a blue sky with scattered white clouds. The word 'Chicago' is superimposed in a large, black, sans-serif font over the center of the image.

Chicago

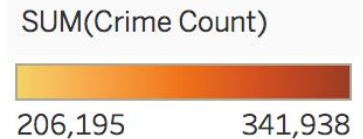
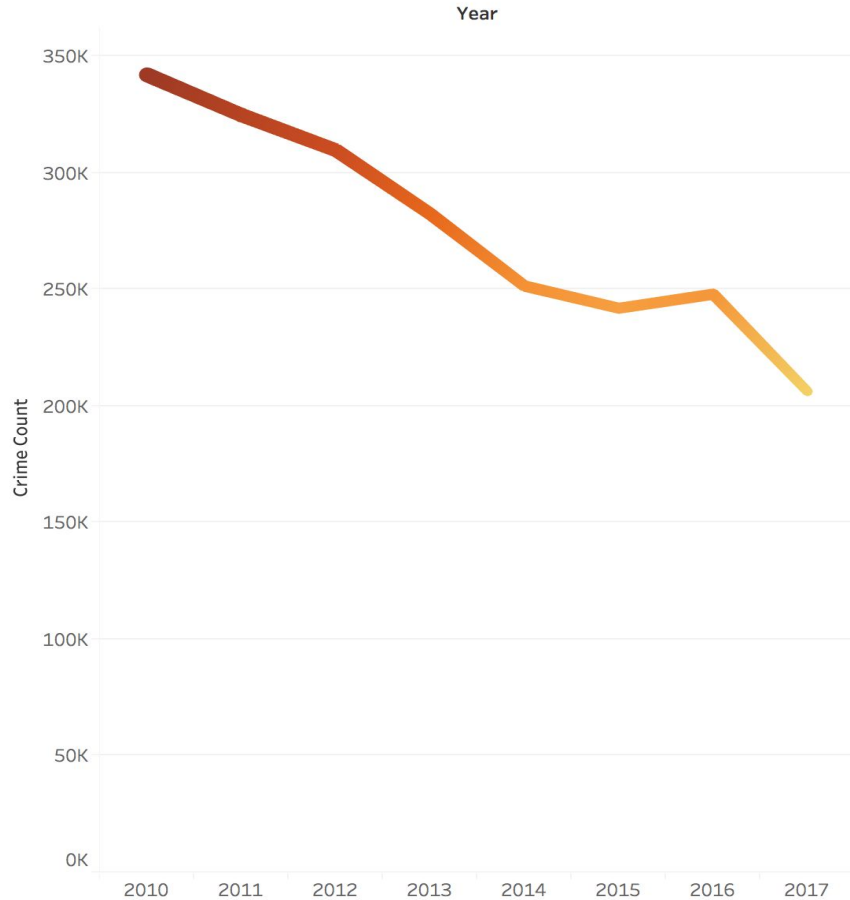
Facts about Chicago Crime Situation

- The overall crime rate in Chicago is 51% higher than national average
- For every 100,000 people, there are 11.77 daily crimes that occur in Chicago
- Chicago is safer than 4% of the cities in the United States
- In Chicago, you have a 1 in 24 chance of becoming a victim of any crime

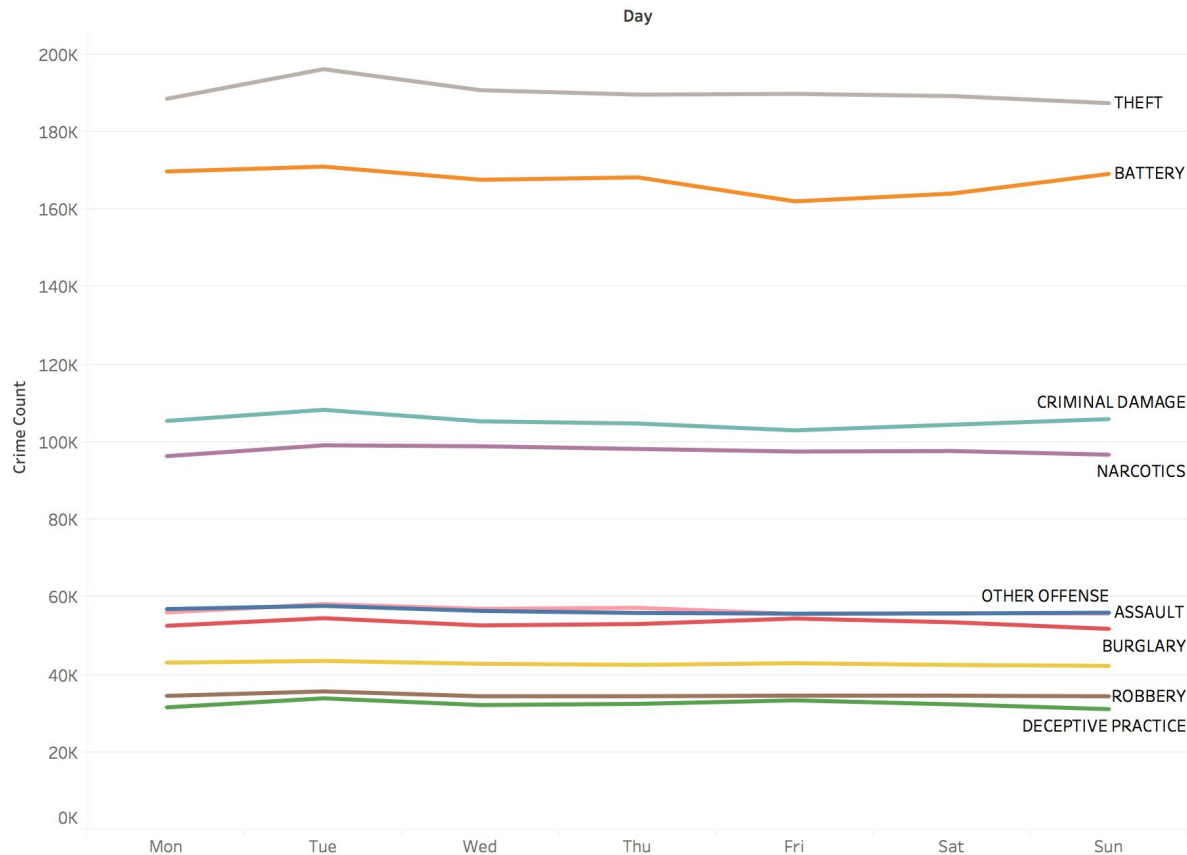
Chicago Top 10 Crime Types



Crime Overview, by Year

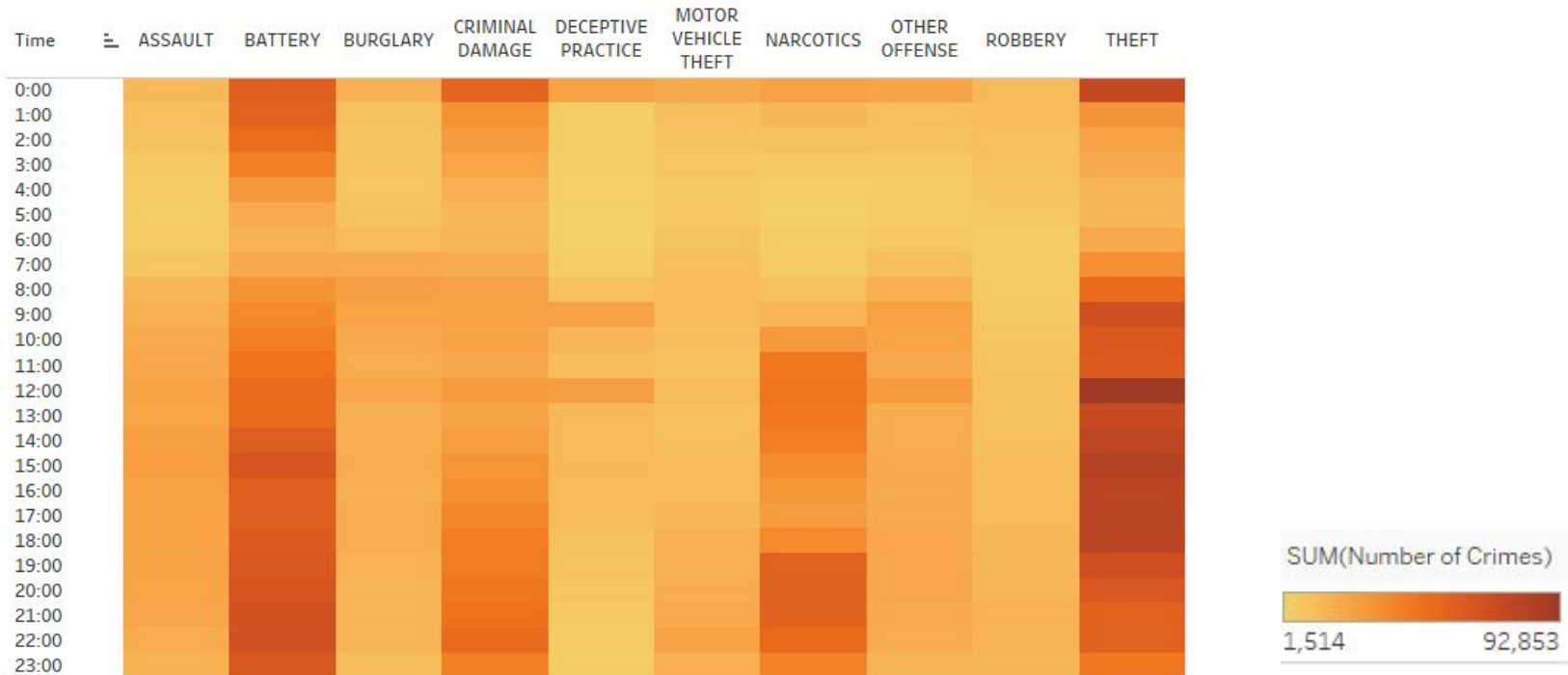


Chicago Crime Overview, by Day

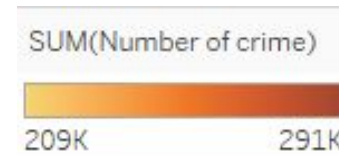
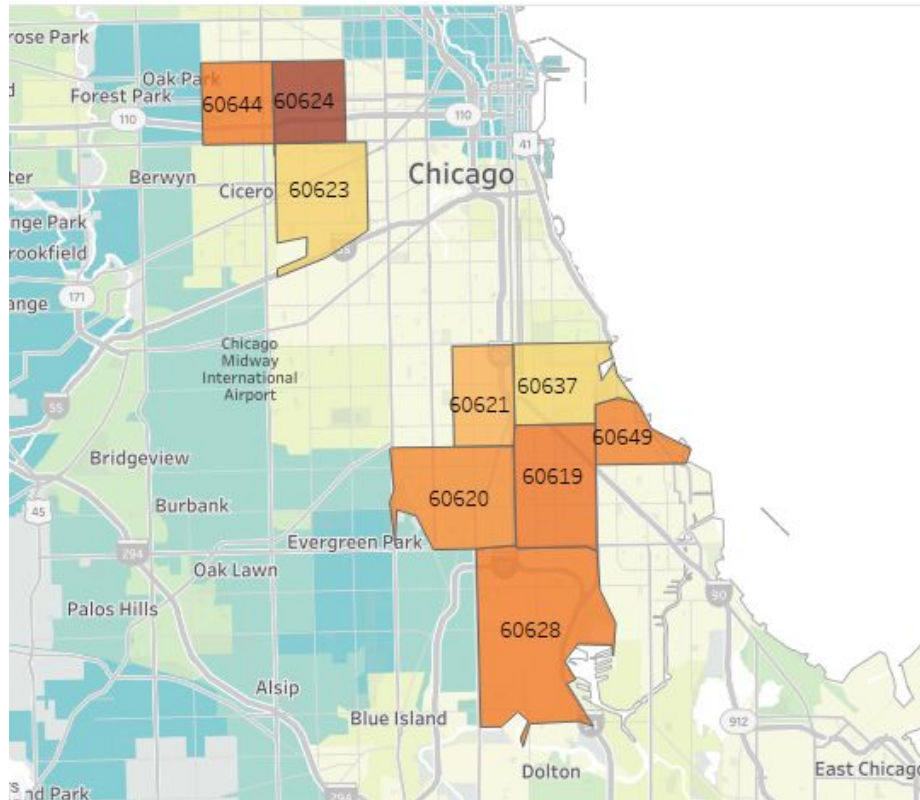


Chicago Top 10 Crime Types - Hourly

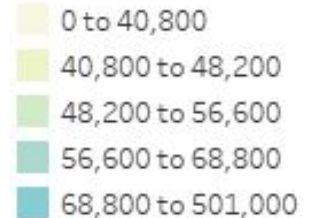
Top 10 Crime Types - Hourly



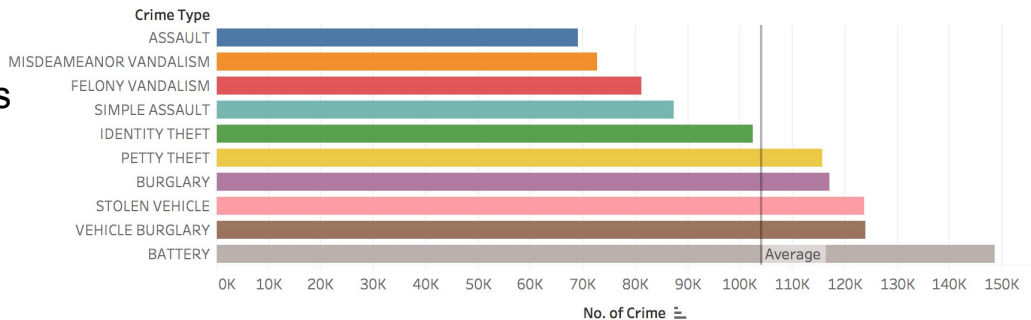
Crime Overview in Chicago Top 10 Zipcodes



2017 Household Incom...

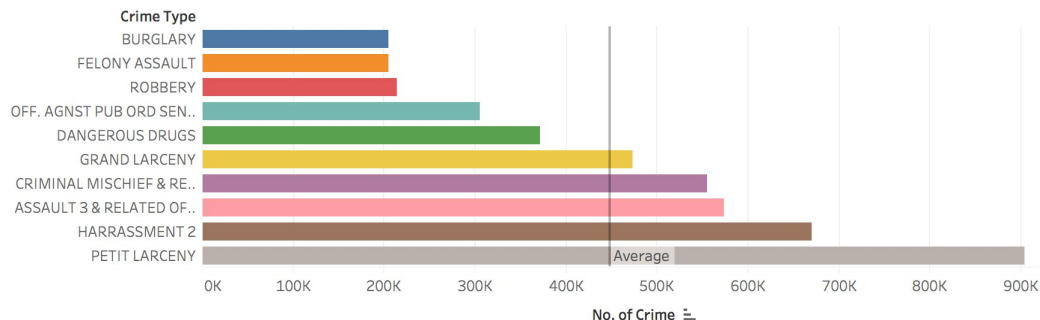


Los Angeles



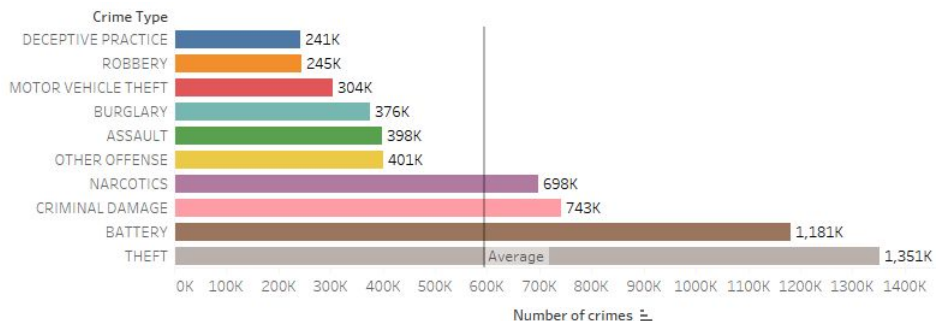
Average around
105K

New York



Average around
440K

Chicago



Average around
600K

In Comparison

- Crime Severity (in average):
Chicago (most dangerous) > New York > Los Angeles > (safest)
- However, theft-related crime is the common crime type among 3 cities
- Crimes can be committed anytime throughout the day
- Economic situation is undeniably one of the most important factor that caused people to commit crimes
- Generally, crime will always present when population density is high

Github Link

<https://github.com/ngupta8/5200-Project>

References

Crime Facts:

<http://www.areavibes.com/los+angeles-ca/crime/>

<http://www.areavibes.com/new+york-ny/crime/>

<http://www.areavibes.com/chicago-il/crime/>

Basic Reverse Geocoding :

<http://download.geonames.org/export/zip/>

<https://stackoverflow.com/questions/5981502/select-the-closest-pair-from-a-list>

<https://stackoverflow.com/questions/27027885/slice-list-of-floats-by-value-in-python>



Thank You!