# CRIME AND COMMERCIAL ACTIVITIES IN TORONTO NEIGHBORHOODS

RICHARD HOU
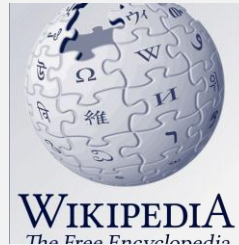
# QUESTION TO BE ANSWERED

- What predicts the type of crimes given the occurrence of a crime in Toronto neighbourhoods?

# SOURCE DATA


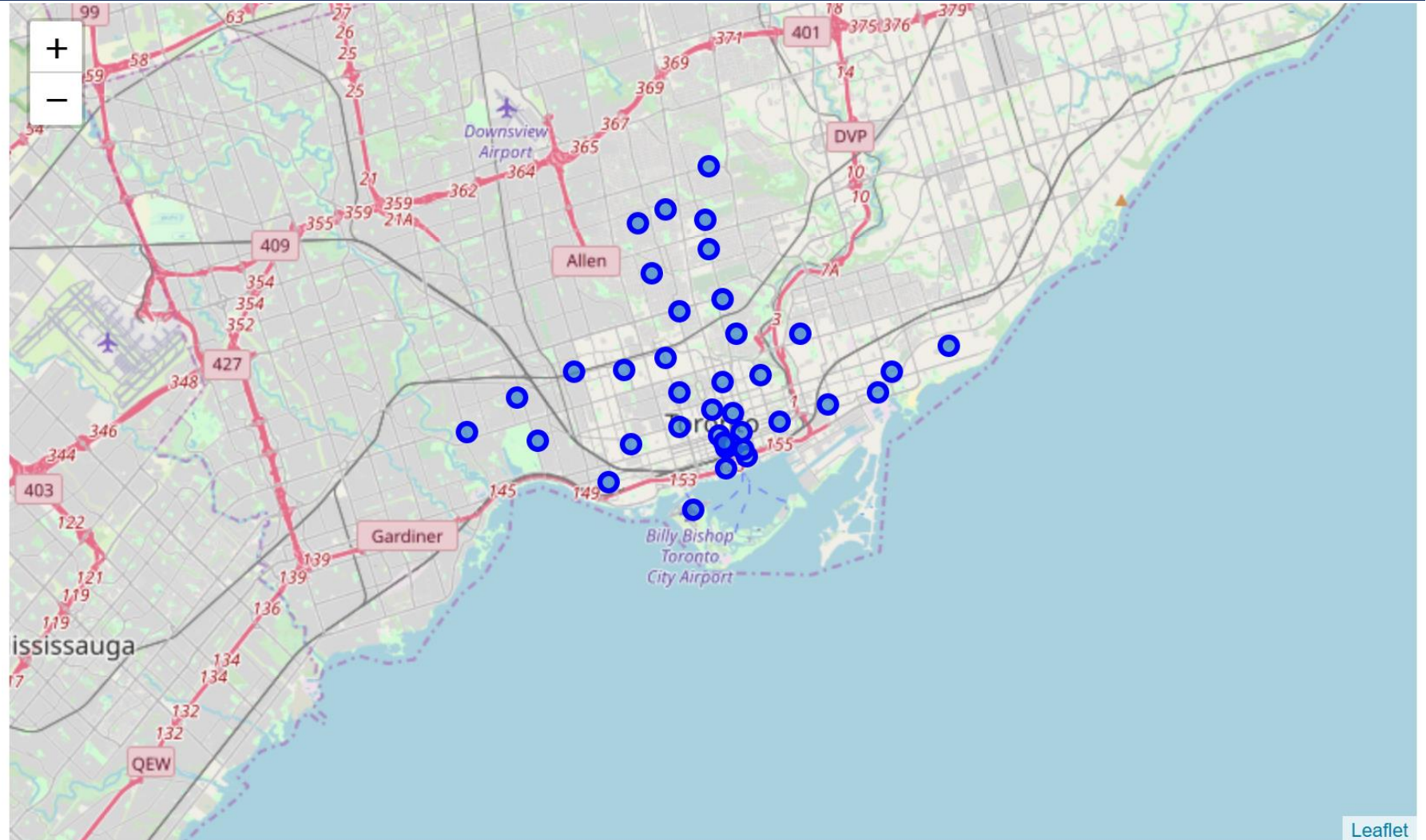
CITY OF TORONTO

WIKIPEDIA

FOURSQUARE

# TORONTO CRIME DATA

## Main information:

Time of crime incidence
Location of crime incidence
Type of crime

# WIKIPEDIA DATA

**Main information:**

List of Toronto neighbourhood Postcode

| Postcode ⬍ | Borough ⬍ | Neighbourhood ⬍ |
|---|---|---|
| M1A | Not assigned | Not assigned |
| M2A | Not assigned | Not assigned |
| M3A | North York | Parkwoods |
| M4A | North York | Victoria Village |
| M5A | Downtown Toronto | Harbourfront |
| M5A | Downtown Toronto | Regent Park |
| M6A | North York | Lawrence Heights |
| M6A | North York | Lawrence Manor |
| M7A | Queen's Park | Not assigned |
| M8A | Not assigned | Not assigned |
| M9A | Etobicoke | Islington Avenue |

# FOURSQUARE DATA

## Main information:

Venues information
**Venue category**

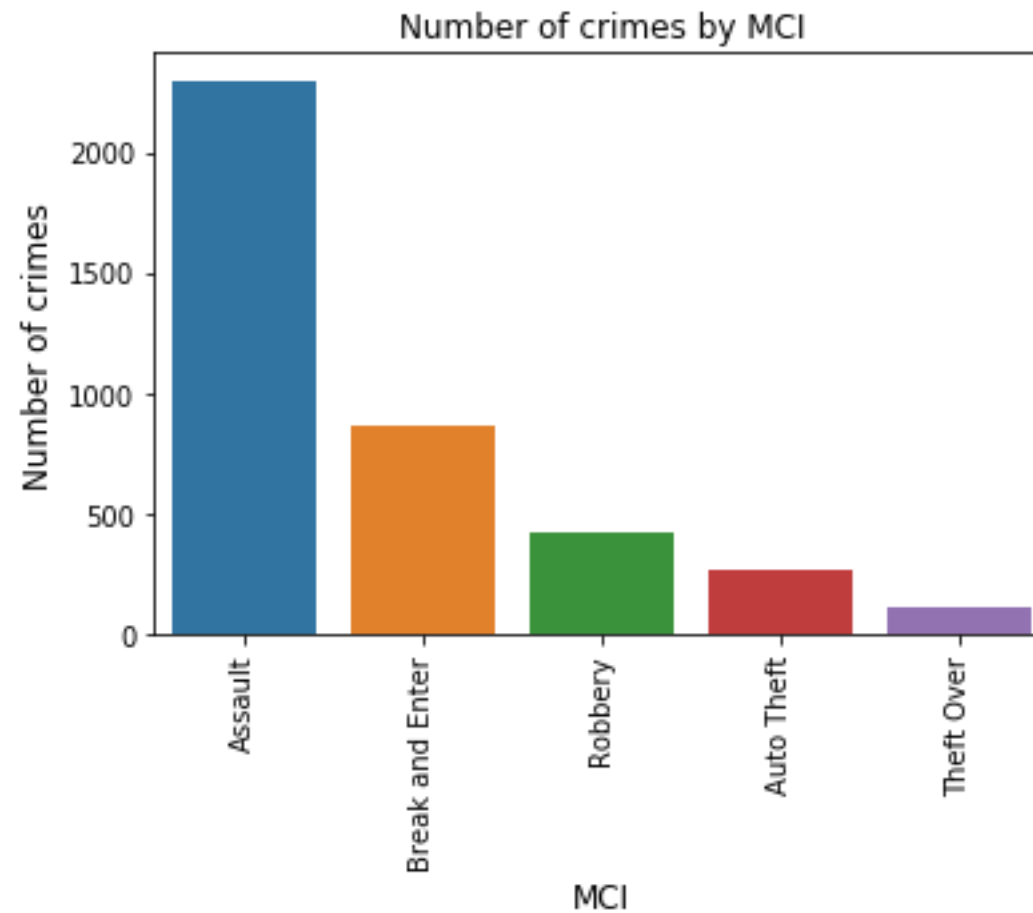| Venue | Venue Latitude | Venue Longitude | Venue distance | Venue ID | Venue Category |
|---|---|---|---|---|---|
| Roselle Desserts | 43.653447 | -79.362017 | 143 | 4bf58dd8d48988d16a941735 | Bakery |
| Tandem Coffee | 43.653559 | -79.361809 | 122 | 4bf58dd8d48988d1e0931735 | Coffee Shop |
| Toronto Cooper Koo Family Cherry St YMCA Centre | 43.653191 | -79.357947 | 247 | 4bf58dd8d48988d175941735 | Gym / Fitness Center |
| The Distillery Historic District | 43.650244 | -79.359323 | 459 | 4deefb944765f83613cdba6e | Historic Site |
| Impact Kitchen | 43.656369 | -79.356980 | 376 | 4bf58dd8d48988d1c4941735 | Restaurant |

# DATA EXPLORATION

|   | Postcode | Borough | Neighborhood |
|---|----------|---------|--------------|
| 0 | M3A | North York | Parkwoods |
| 1 | M4A | North York | Victoria Village |
| 2 | M5A | Downtown Toronto | Harbourfront,Regent Park |
| 3 | M6A | North York | Lawrence Heights,Lawrence Manor |
| 4 | M7A | Queen's Park | Queen's Park |

# DATA EXPLORATION

| | Venue value counts | Venue Category |
|---|---|---|
| **Coffee Shop** | 159 | Coffee Shop |
| **Café** | 123 | Café |
| **Italian Restaurant** | 71 | Italian Restaurant |
| **Park** | 66 | Park |
| **Bakery** | 63 | Bakery |

# DATA EXPLORATION



Number of crimes by MCI

# DATA EXPLORATION



Number of crimes by year

# DATA EXPLORATION



Number of crimes by month of the year

# DATA EXPLORATION



Number of crimes by Day
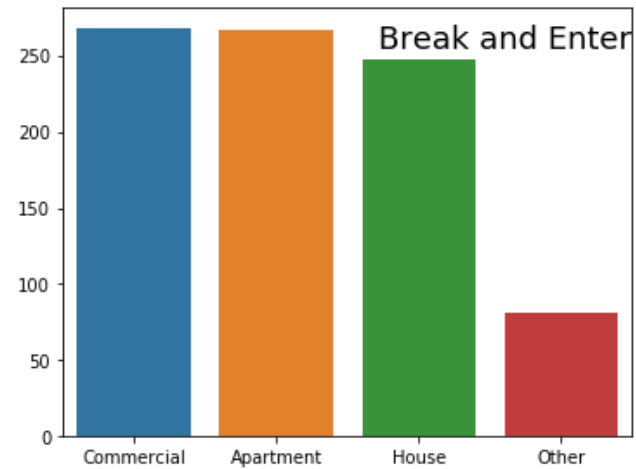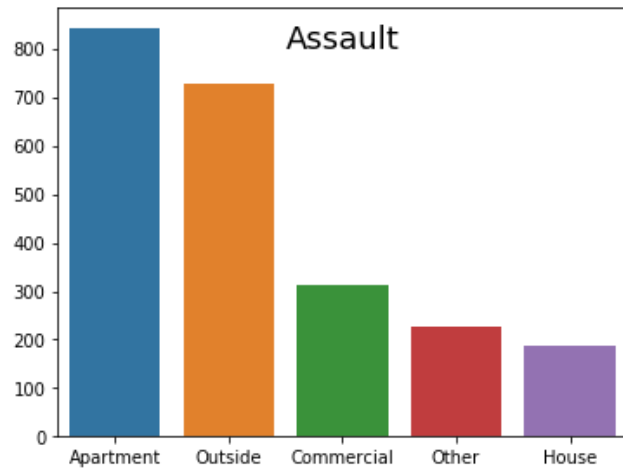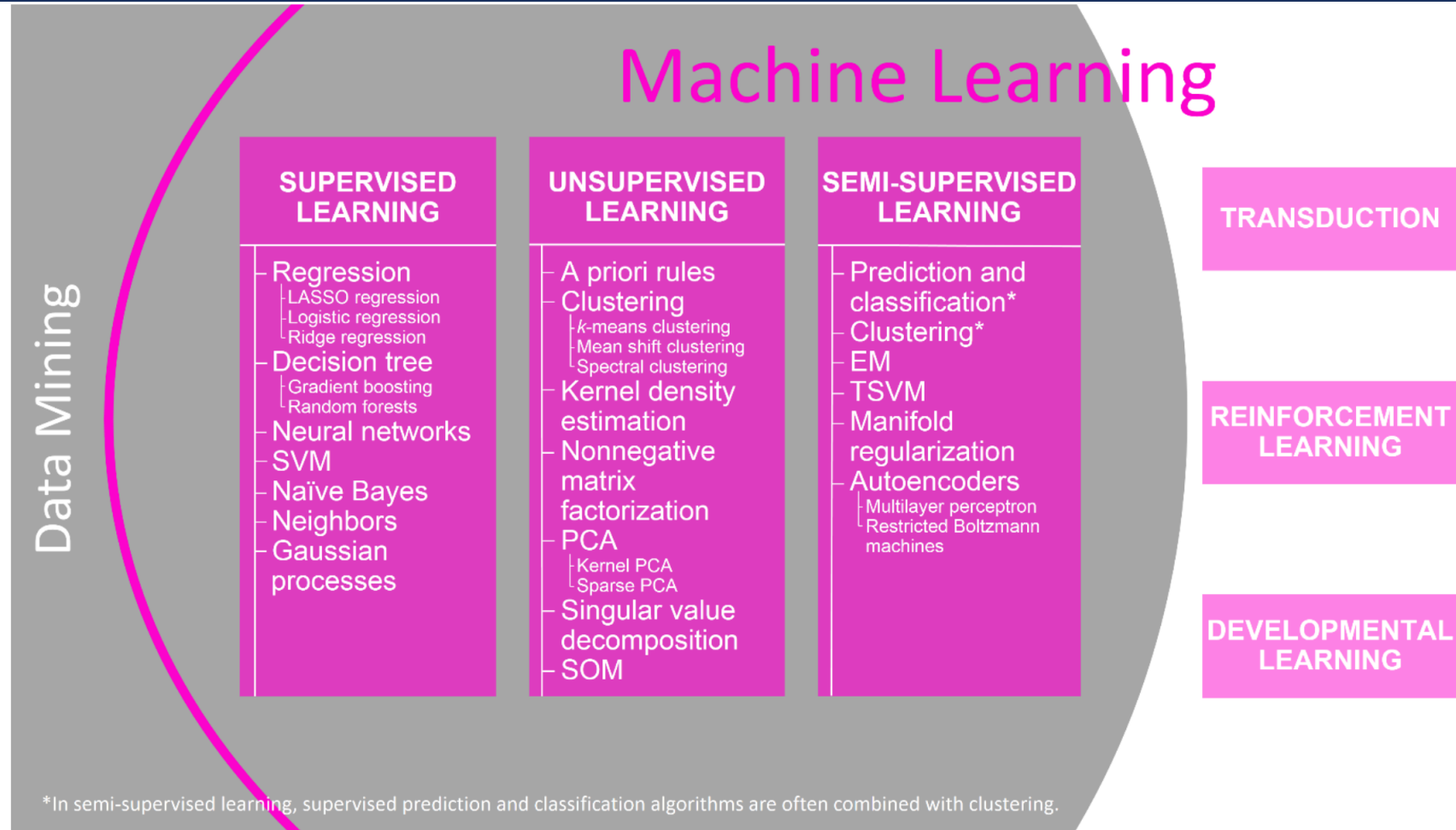
# DATA EXPLORATION
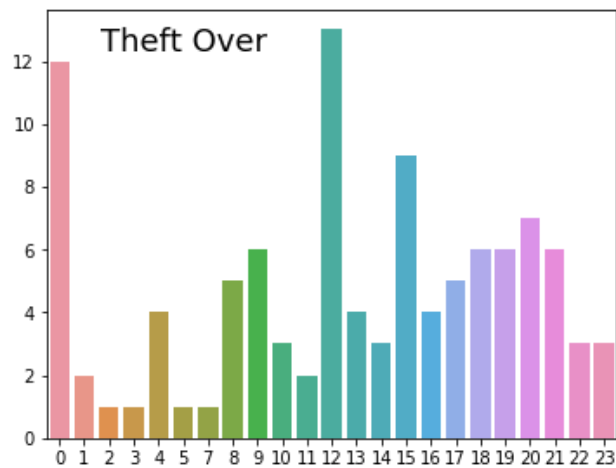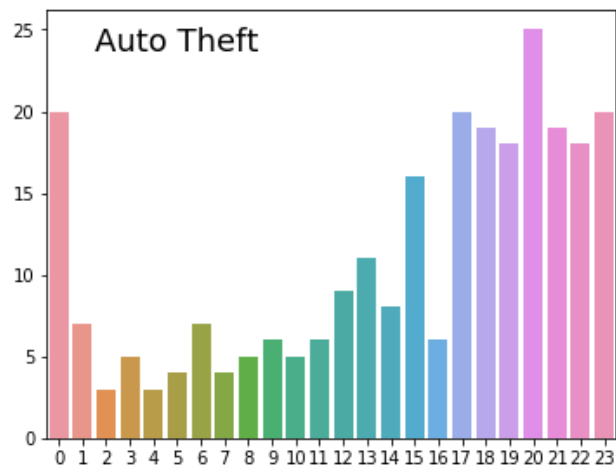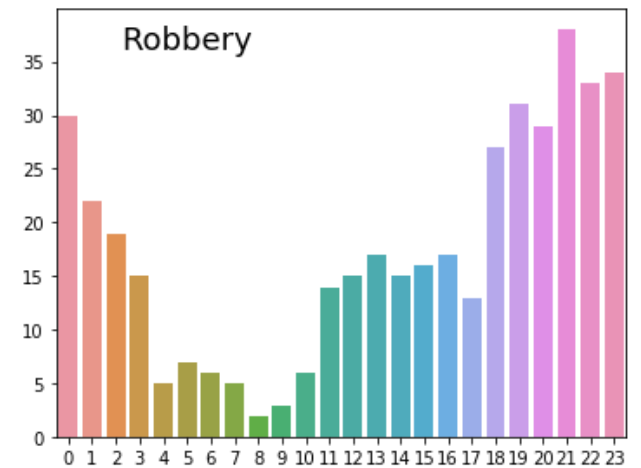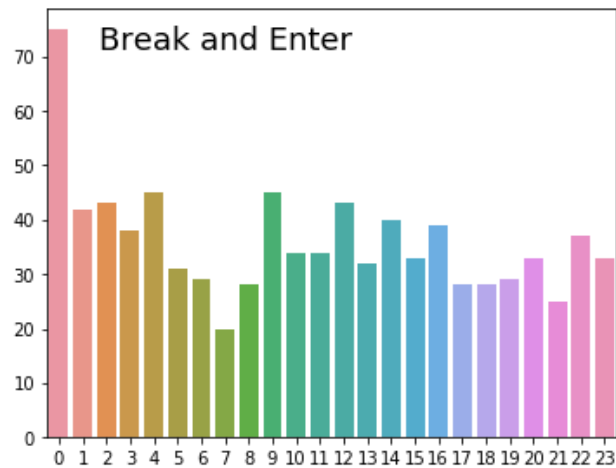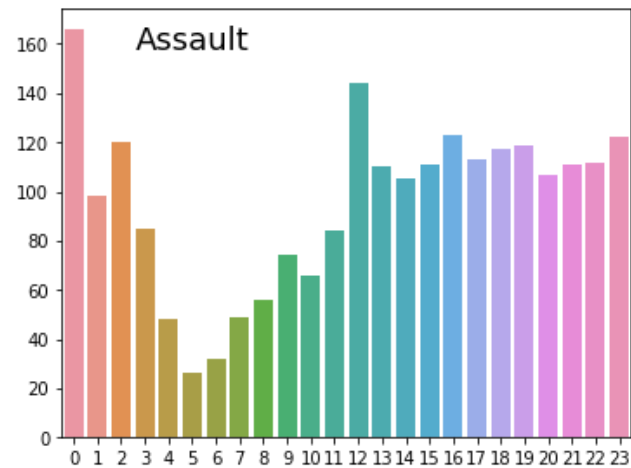
# DATA EXPLORATION
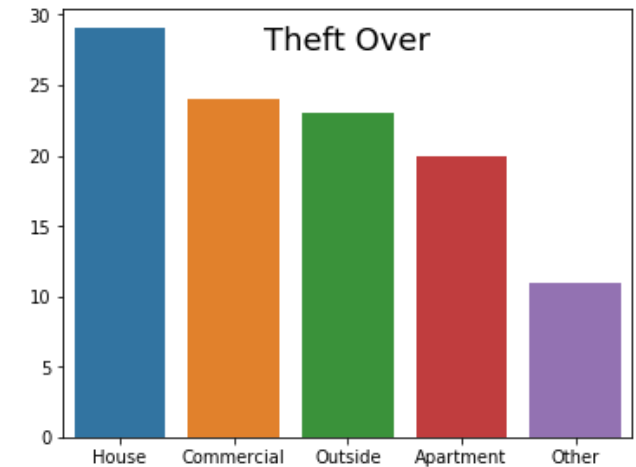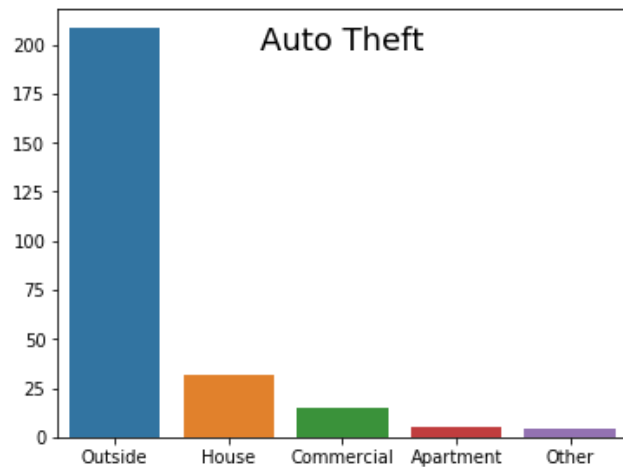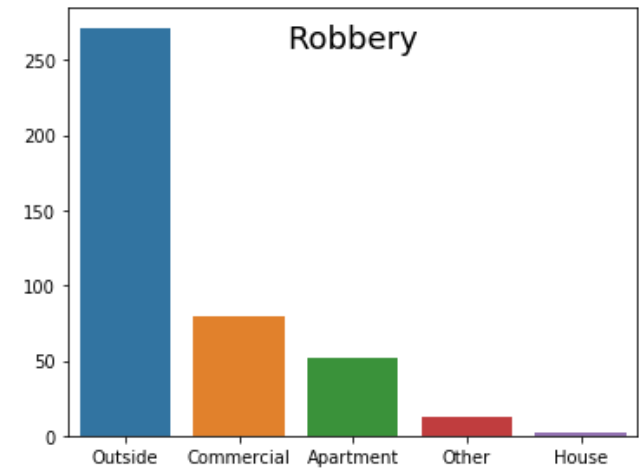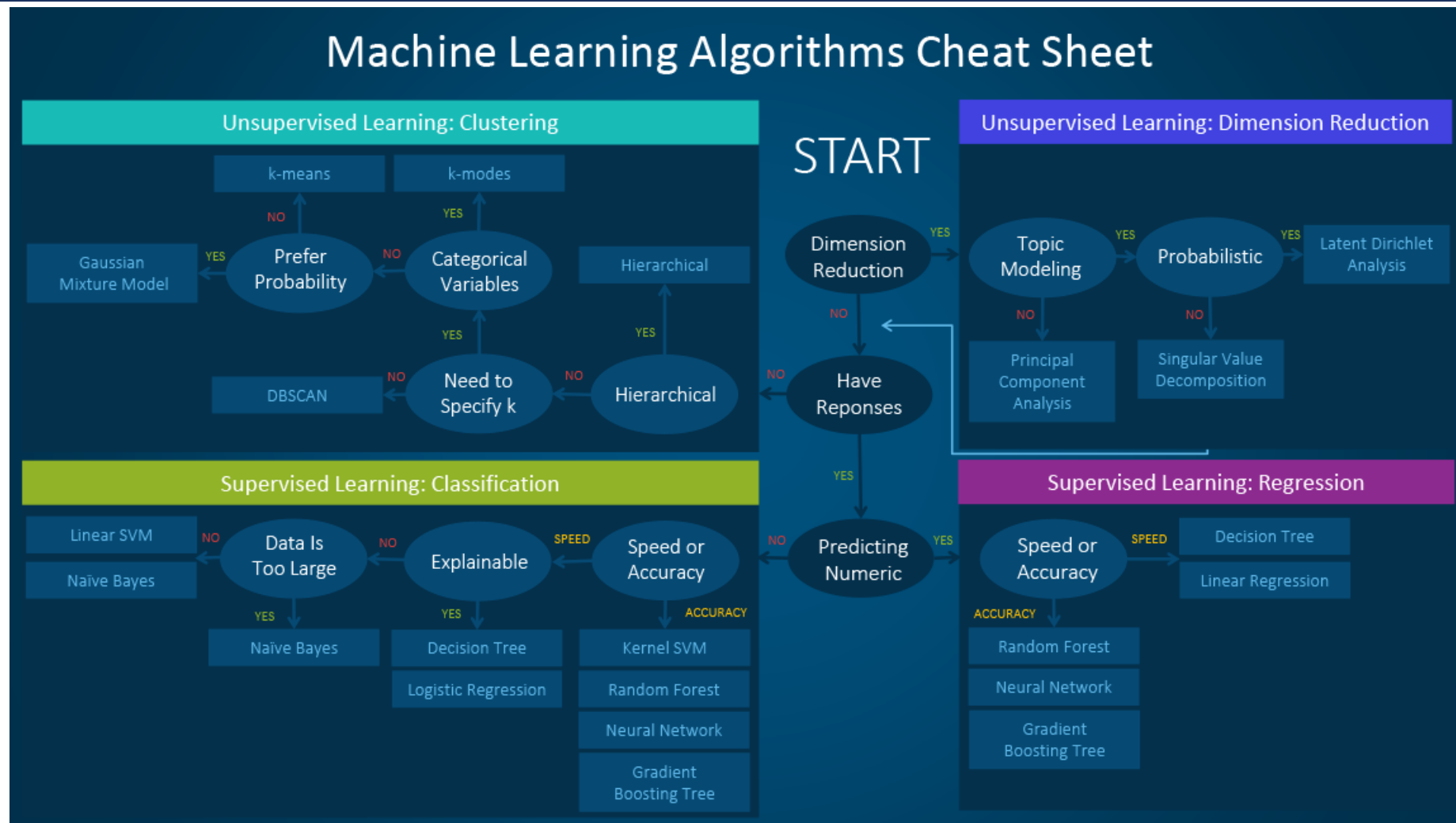
# MACHINE LEARNING



Machine Learning Taxonomy

Source: An Overview of SAS® Visual Data Mining and Machine Learning on SAS® Viva. SAS Institute Inc.

# MACHINE LEARNING

Source:
https://blogs.sas.com



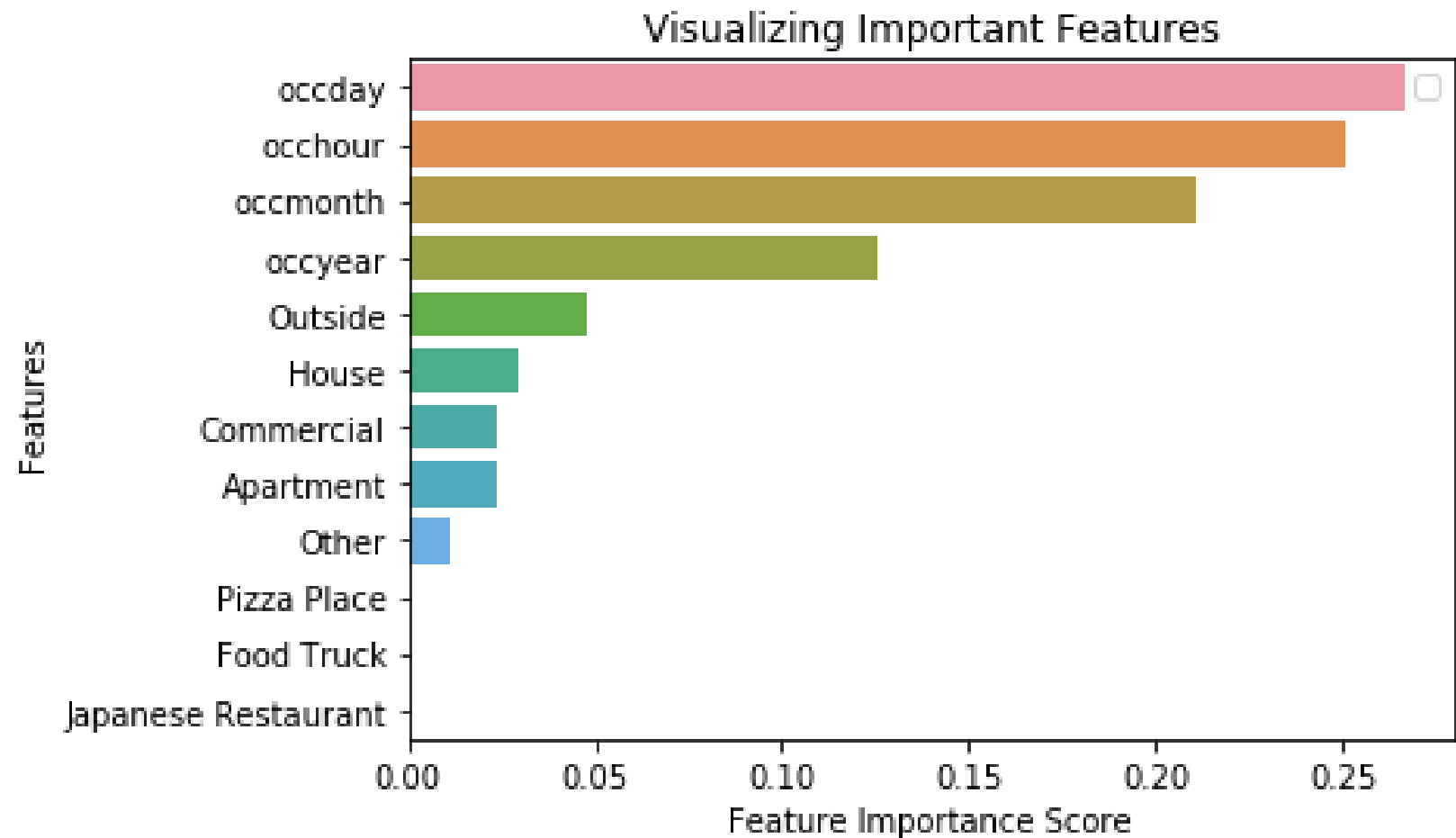Machine Learning Algorithms Cheat Sheet

# MACHINE LEARNING

- K-Nearest Neighbour (KNN)

- Logistic regression (LR)

- Decision Tree (Tree)
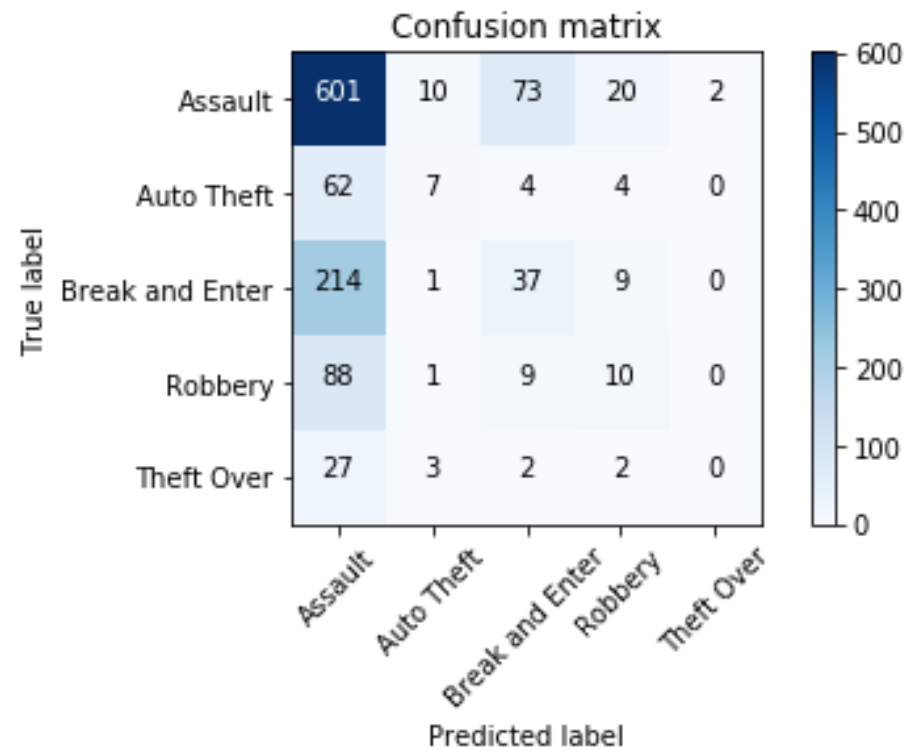
- SVM

- Neural Network (NN)

- RandomForest (RF)
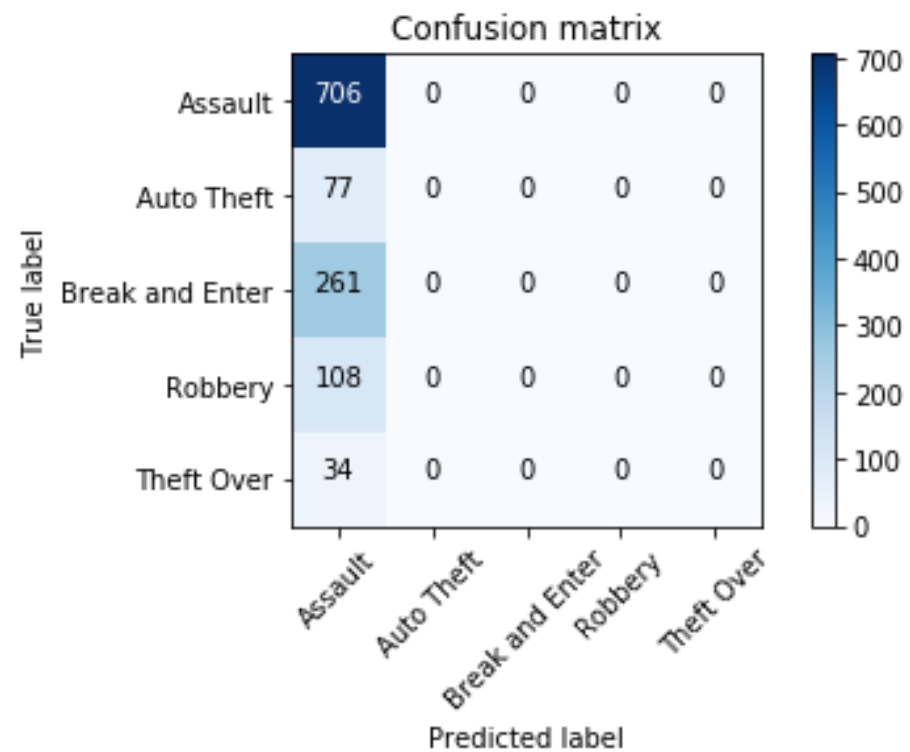
# MACHINE LEARNING: RANDOMFOREST

Accuracy ratio: 62%



Visualizing Important Features

# MACHINE LEARNING: KNN

Accuracy ratio: 55%



Confusion matrix

# MACHINE LEARNING: LR, TREE, SVM, NN

Accuracy ratio: 60%



Confusion matrix

# CONCLUSION

- Not so satisfactory result
- Not all data science project will lead to success
- Many insights from the data exploration
- Many insights from RandomForest method
- Great hands-on experience

THANK YOU