

911 Baltimore Crime Analysis

Contents

→ Data overview

→ Plots

1. Series of plots to describe the temporal pattern (year-to-year, monthly, day-of-week, of the overall crime incidence aggregating from all geo-locations
2. Plot describing the distribution of crime type aggregating from all geo-locations and all time periods
3. Series of plots to illustrate how crime types distributions may vary from year-to-year, monthly, day-of-week aggregating across all locations

→ Interactive dashboard

Data Overview (before cleaning the data)

Overview Alerts 19 Reproduction		
Dataset statistics		Variable types
Number of variables	14	Categorical 11
Number of observations	1000	Numeric 3
Missing cells	1199	
Missing cells (%)	8.6%	
Duplicate rows	35	
Duplicate rows (%)	3.5%	
Total size in memory	640.7 KiB	
Average record size in memory	656.0 B	

There are a total of 14 attributes, 11 being categorical and 3 being numeric. There are 1000 data entries with 8.6% missing cells and 3.5% duplicate rows.

Alerts

Alerts

Total Incidents has constant value "1"

Constant

Dataset has 35 (3.5%) [duplicate rows](#)

Duplicates

CrimeTime has a high cardinality: 419 distinct values

High cardinality

Location has a high cardinality: 840 distinct values

High cardinality

Neighborhood has a high cardinality: 195 distinct values

High cardinality

Premise has a high cardinality: 51 distinct values

High cardinality

Post is highly overall correlated with **Longitude** and [1 other fields](#)

High correlation

Longitude is highly overall correlated with **Post**

High correlation

CrimeCode is highly overall correlated with **Description** and [2 other fields](#)

High correlation

Description is highly overall correlated with **CrimeCode** and [1 other fields](#)

High correlation

Inside/Outside is highly overall correlated with **CrimeCode** and [2 other fields](#)

High correlation

Weapon is highly overall correlated with **CrimeCode**

High correlation

District is highly overall correlated with **Post**

High correlation

Premise is highly overall correlated with **Inside/Outside**

High correlation

Inside/Outside has 167 (16.7%) missing values

Missing

Weapon has 776 (77.6%) missing values

Missing

Neighborhood has 76 (7.6%) missing values

Missing

Premise has 167 (16.7%) missing values

Missing

Location is uniformly distributed

Uniform

Alerts tell the cardinality, correlation and other details like missing values, duplicates etc of the attributes.

After removing the duplicate rows:

skippy summary





Data Summary

dataframe	Values
Number of rows	959
Number of columns	14

Data Types

Column Type	Count
string	10
float64	3
int32	1

number

column_name	NA	NA %	mean	sd	p0	p25	p75	p100	hist
Post	4	0.42	520	260	110	310	730	940	
Longitude	4	0.42	-77	0.043	-77	-77	-77	-77	
Latitude	4	0.42	39	0.028	39	39	39	39	
Total Incidents	0	0	1	0	1	1	1	1	

string

column_name	NA	NA %	words per row	total words
CrimeDate	0	0	1	960
CrimeTime	0	0	1	960
CrimeCode	0	0	1	960
Location	1	0.1	1	960
Description	0	0	1	960
Inside/Outside	160	16	1	960
Weapon	750	78	1	960
District	0	0	1	960
Neighborhood	74	7.7	1	960
Premise	160	16	1	960

End

Missing values

Inside/Outside

Categorical

HIGH CORRELATION

MISSING

Distinct	2
Distinct (%)	0.2%
Missing	167
Missing (%)	16.7%
Memory size	57.0 KiB



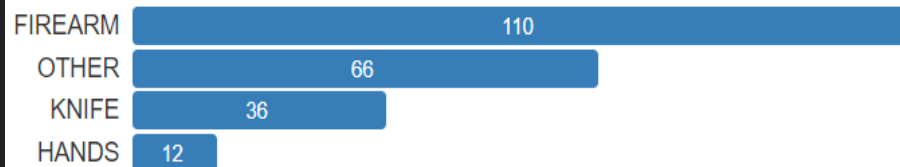
Weapon

Categorical

HIGH CORRELATION

MISSING

Distinct	4
Distinct (%)	1.8%
Missing	776
Missing (%)	77.6%
Memory size	38.2 KiB



Missing values

Premise

Categorical

HIGH CARDINALITY

HIGH CORRELATION

MISSING

Distinct	51
Distinct (%)	6.1%
Missing	167
Missing (%)	16.7%
Memory size	61.6 KiB

Neighborhood

Categorical

HIGH CARDINALITY

MISSING

Distinct	195
Distinct (%)	21.1%
Missing	76
Missing (%)	7.6%
Memory size	65.8 KiB

The amount of missing data in the above 4 attributes is concerning. The most important details in a crime scene like the location, neighbourhood and weapons used should be noted, through which the police department can track the pattern of crimes happening in and around the neighbourhood.

Important attributes that are used in this analysis

District

Categorical

Distinct	10
Distinct (%)	1.0%
Missing	0
Missing (%)	0.0%
Memory size	63.8 KiB

DISTRICT: Eastern, Southwest, Northeast, Northwest, Northern, Southeast, Southern, Central, Western, Unknown

Description

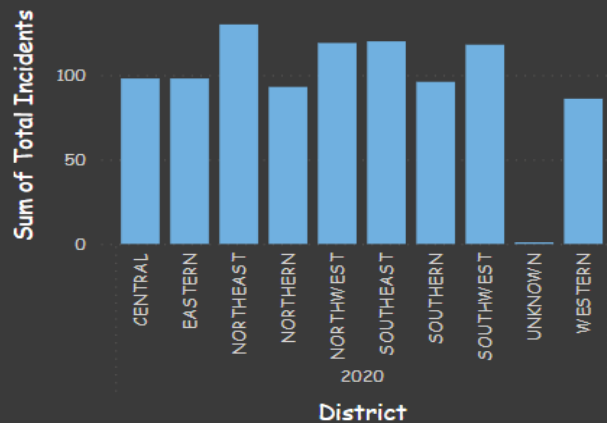
Categorical

Distinct	13
Distinct (%)	1.3%
Missing	0
Missing (%)	0.0%
Memory size	67.4 KiB

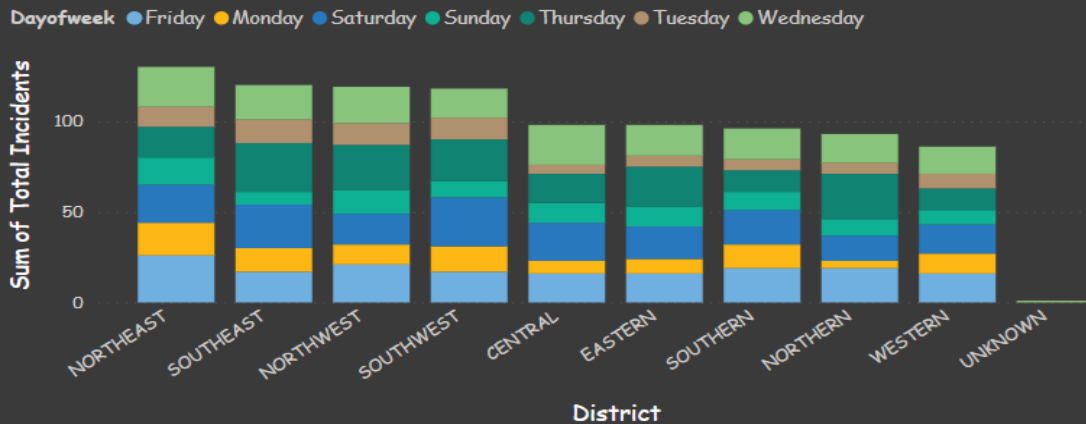
DESCRIPTION: Burglary, Larceny, Robbery - Residence, Auto-theft, Robbery - Commercial, Robbery - Street, Common Assault, Shooting, Agg. Assault, Larceny from auto, Robbery - Carjacking, Homicide, Rape

1. Series of plots to describe the temporal pattern (year-to-year, monthly, and day-of-week) of the overall crime incidence aggregating from all geo-locations

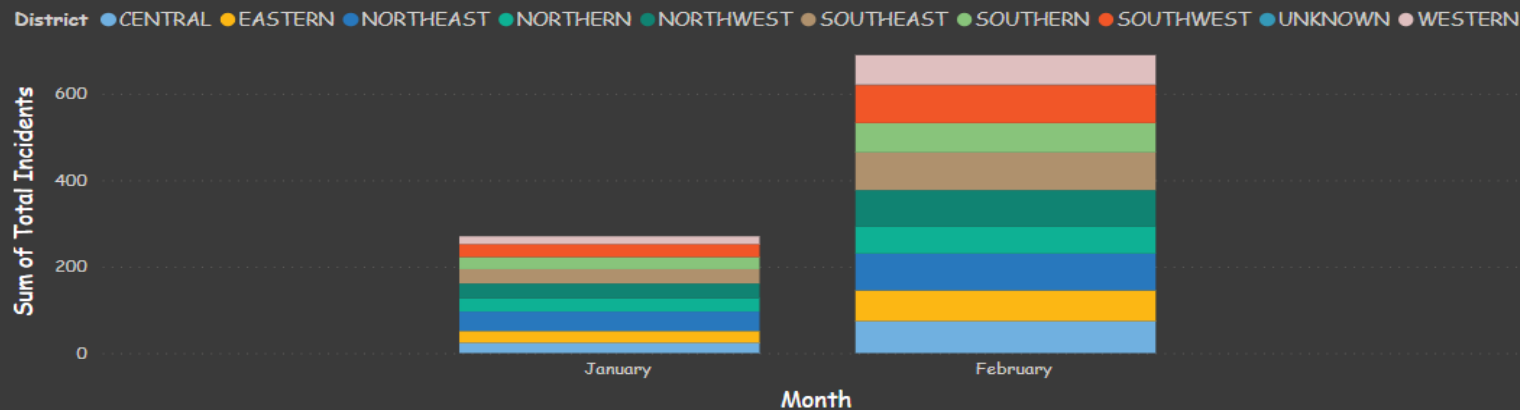
Sum of Total Incidents by Year and District



Sum of Total Incidents by District and Dayofweek

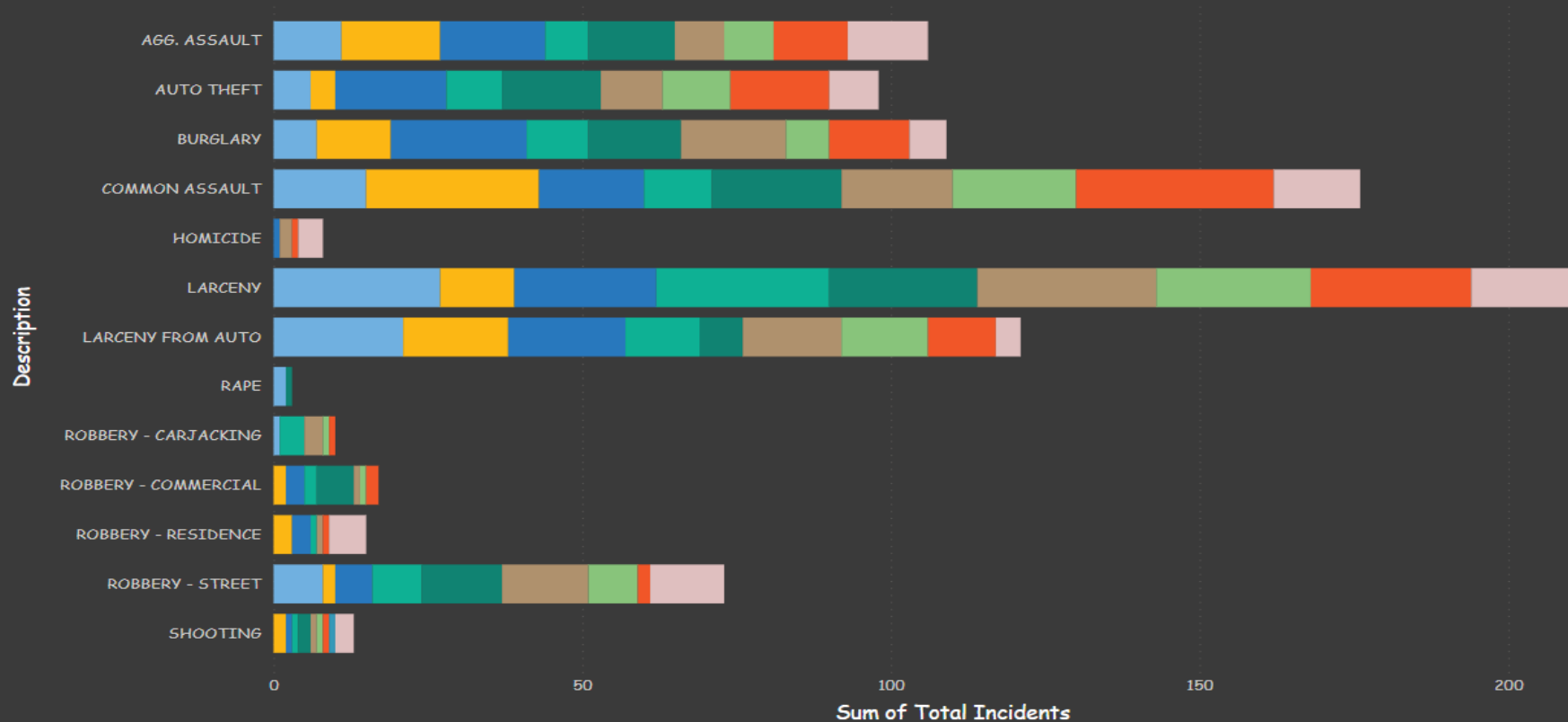


Sum of Total Incidents by Month and District



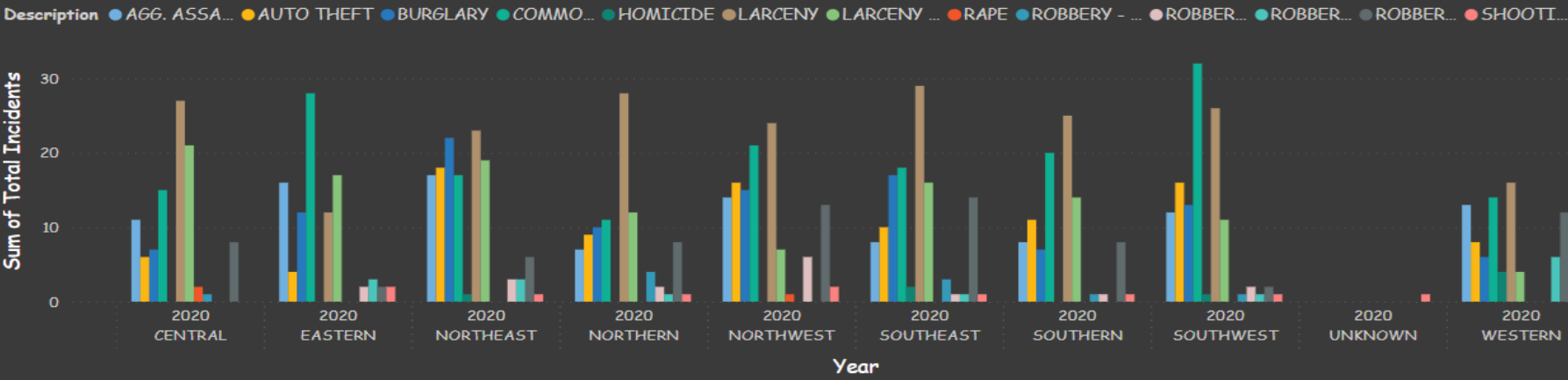
2. Plot describing the distribution of crime type aggregating from **all geo-locations and all time periods**

Sum of Total Incidents by Description and District

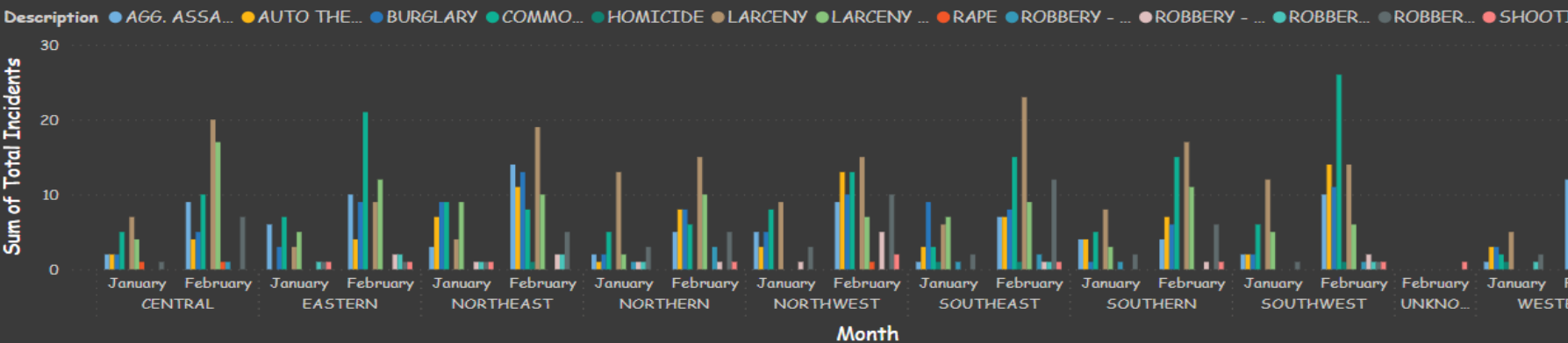


3. Series of plots to illustrate how crime types distributions might vary from year-to-year, month-to-month, or by day-of-week aggregating across all locations

Sum of Total Incidents by District, Year and Description

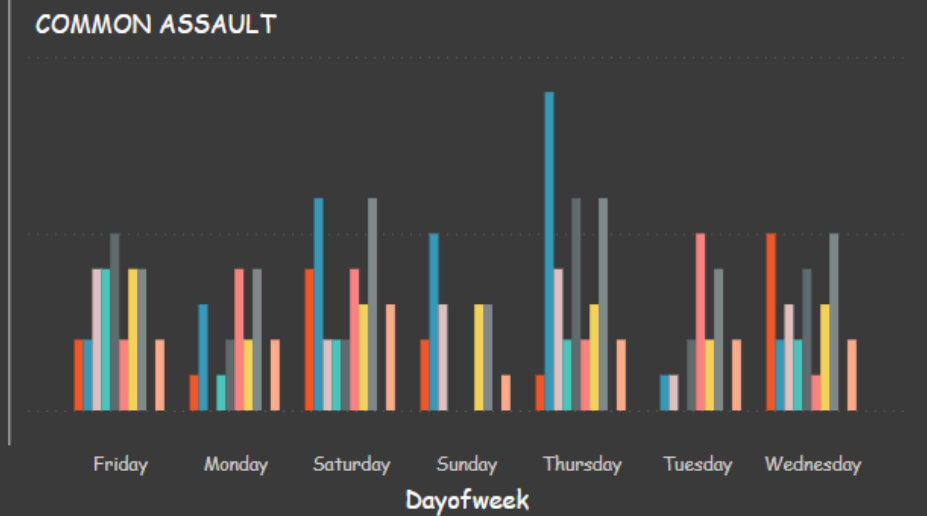
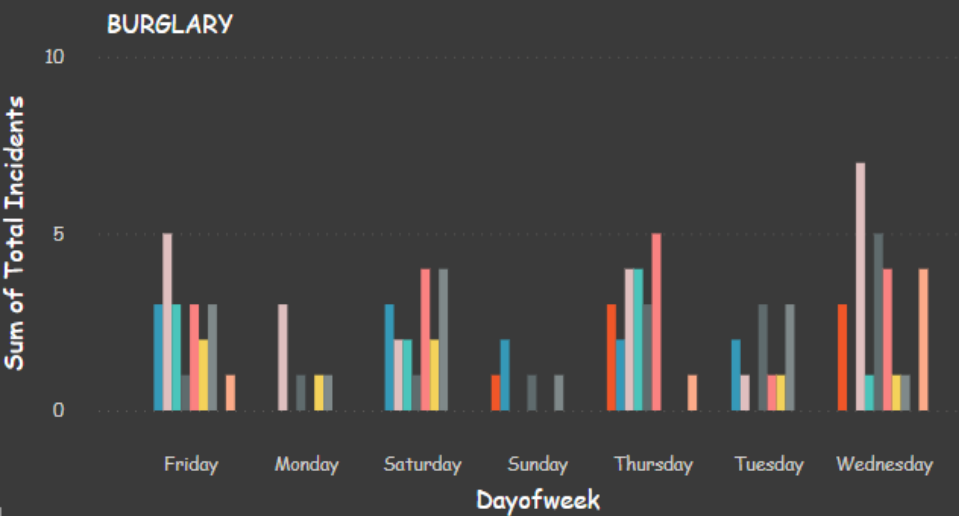
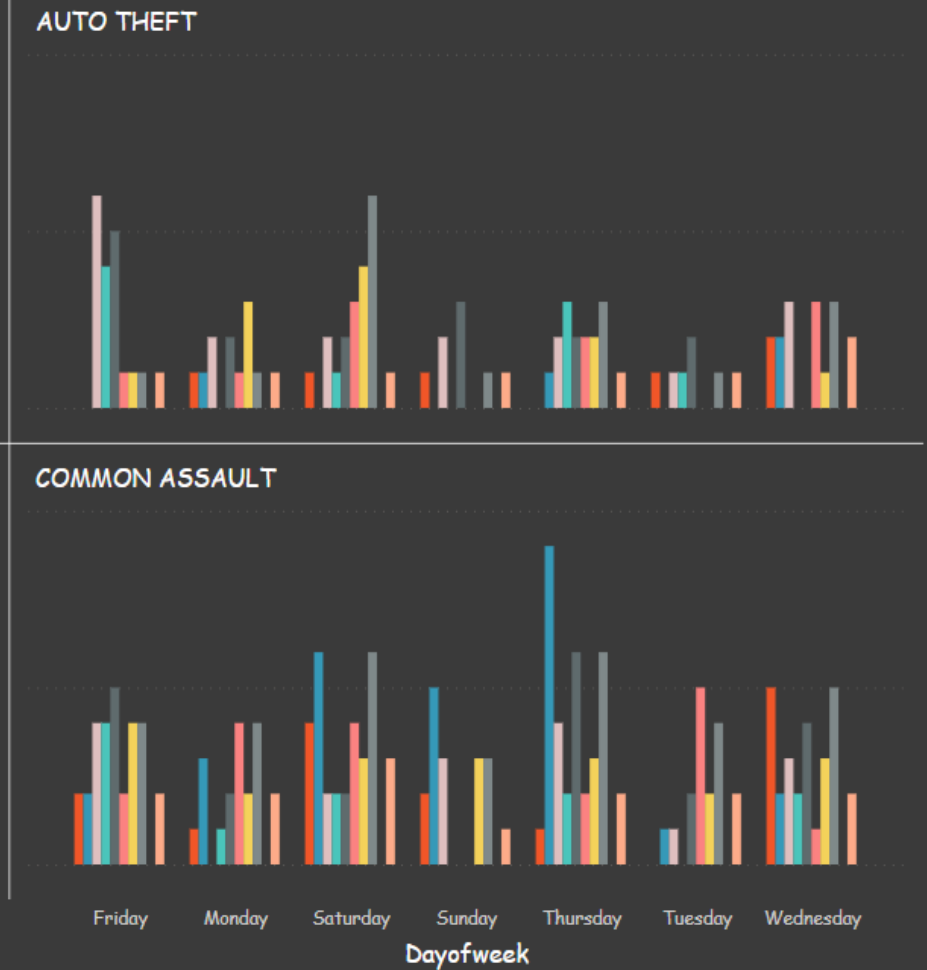
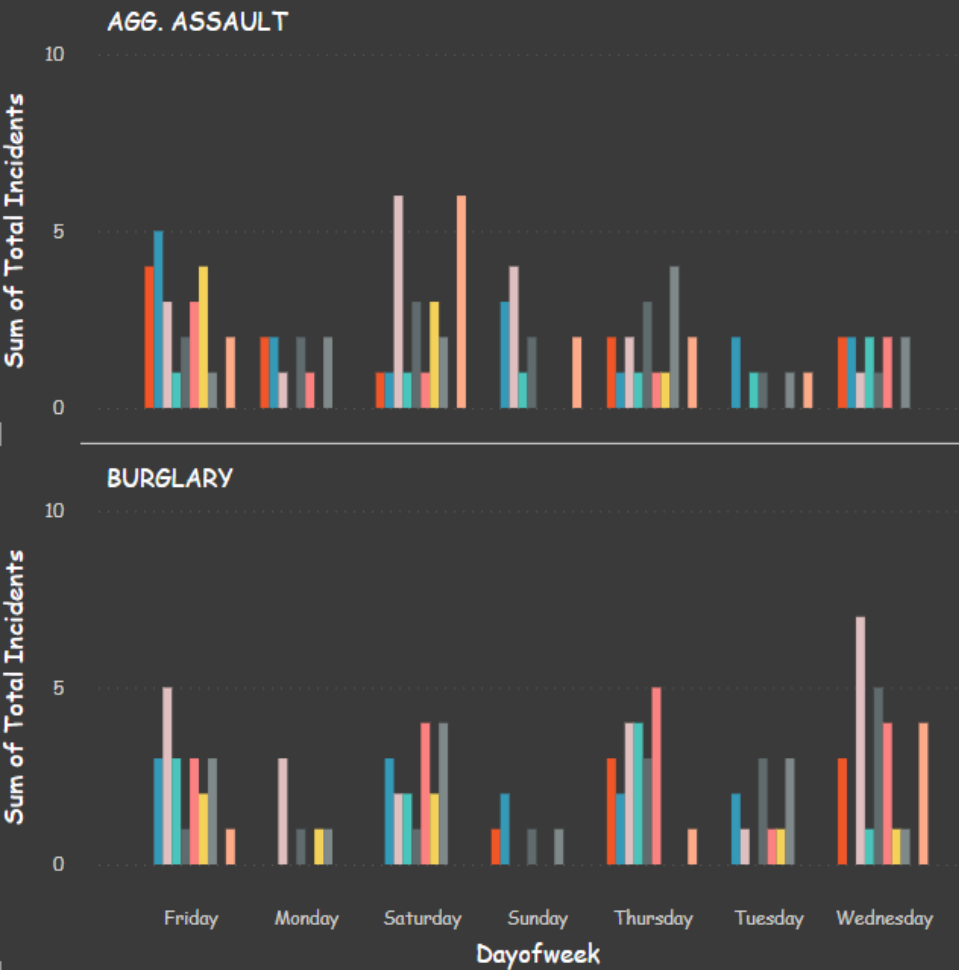


Sum of Total Incidents by District, Month and Description

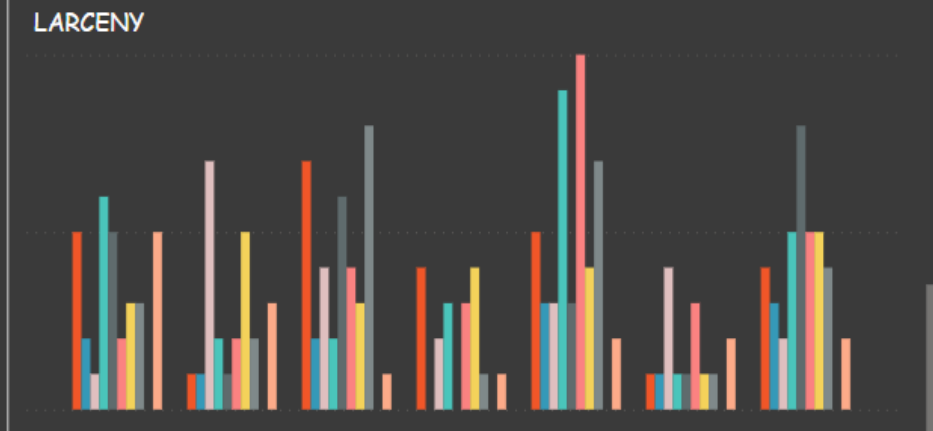


Sum of Total Incidents by Dayofweek, District and Description

District CENTRAL EASTERN NORTHEAST NORTHERN NORTHWEST SOUTHEAST SOUTHERN SOUTHWEST UNKNOWN WESTERN



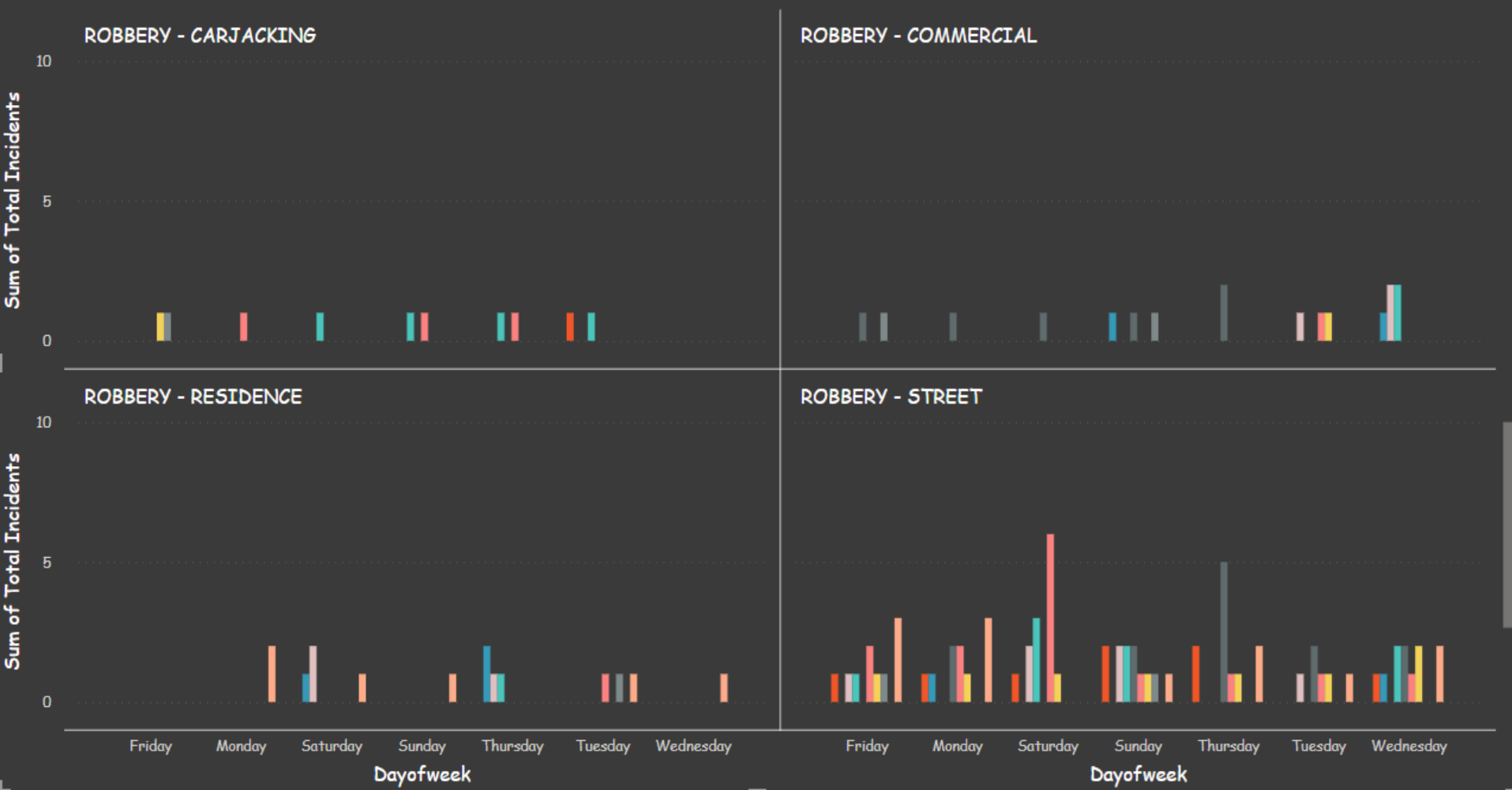
District ● CENTRAL ● EASTERN ● NORTHEAST ● NORTHERN ● NORTHWEST ● SOUTHEAST ● SOUTHERN ● SOUTHWEST ● UNKNOWN ● WESTERN



RAPE

Sum of Total Incidents by Dayofweek, District and Description

District CENTRAL EASTERN NORTHEAST NORTHERN NORTHWEST SOUTHEAST SOUTHERN SOUTHWEST UNKNOWN WESTERN



Interactive Dashboard

Video recording explaining the dashboard: <https://icecreamapps.com/v/ghrxnpx>

911 Baltimore Crime Data Analysis

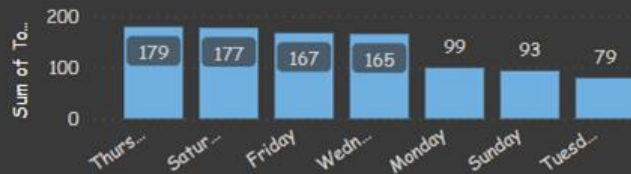
Month

All

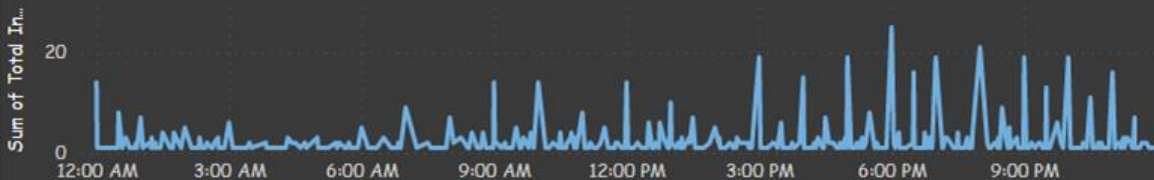
959

Total incidents

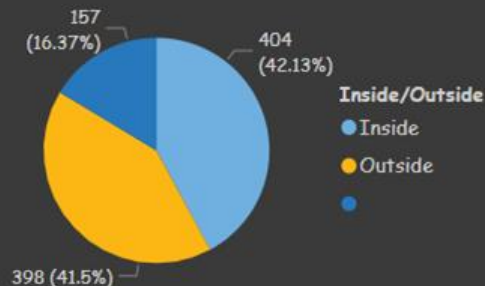
Sum of Total Incidents by Day of week



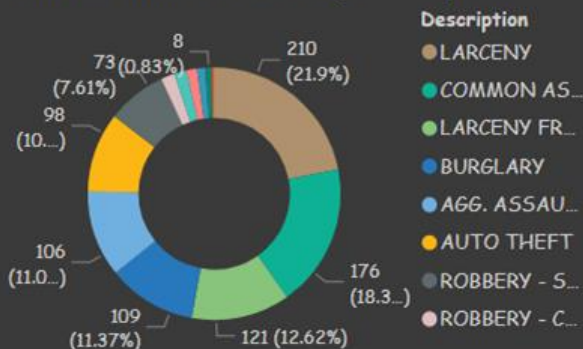
Sum of Total Incidents by Crime Time



Sum of Total Incidents by Inside/Outside



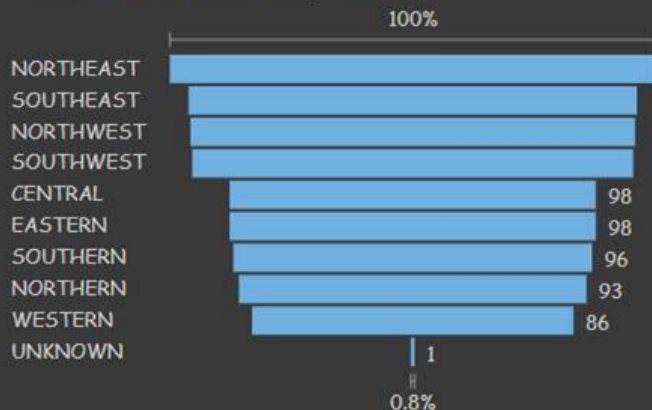
Sum of Total Incidents by Description



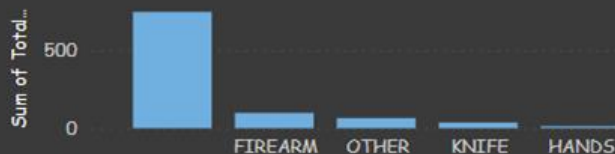
Number of incidents by Crime Date



Number of incidents by District



Weapons Used



Larceny

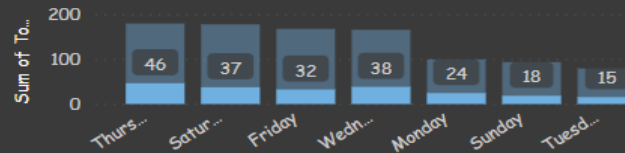
911 Baltimore Crime Data Analysis

Month
All

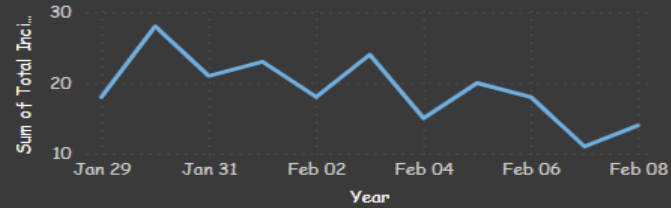
210

Total incidents

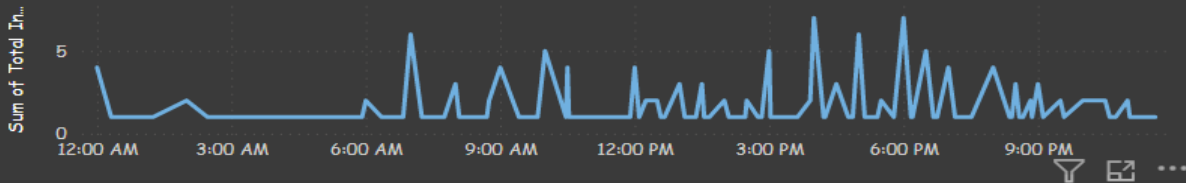
Sum of Total Incidents by Dayofweek



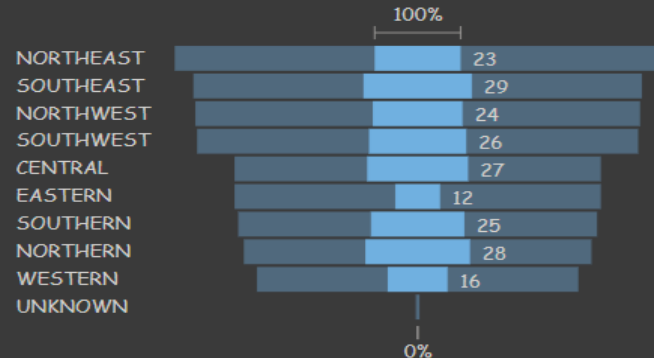
Number of incidents by Crime Date



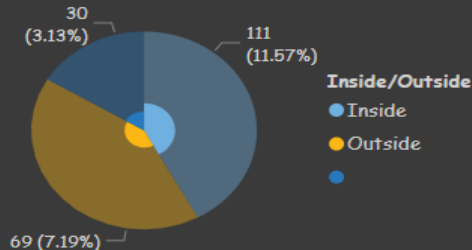
Sum of Total Incidents by CrimeTime



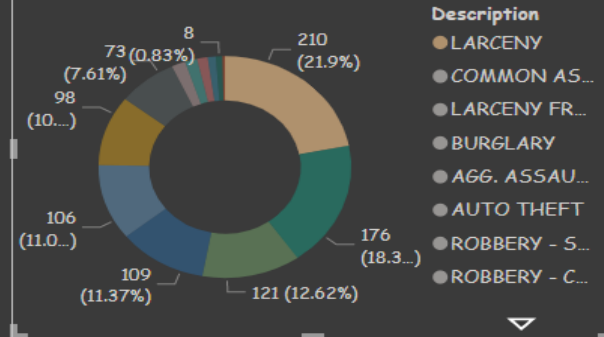
Number of incidents by District



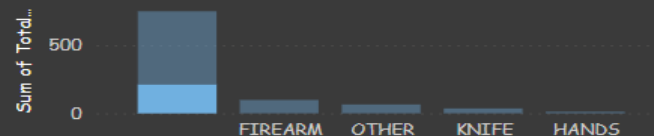
Sum of Total Incidents by Inside/Outside



Sum of Total Incidents by Description



Weapons Used



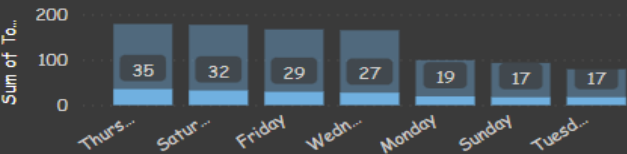
Common Assault

911 Baltimore Crime Data Analysis

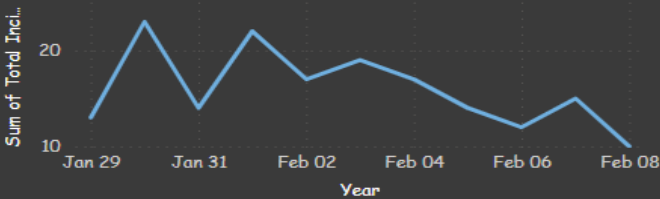
Month
All

176
Total incidents

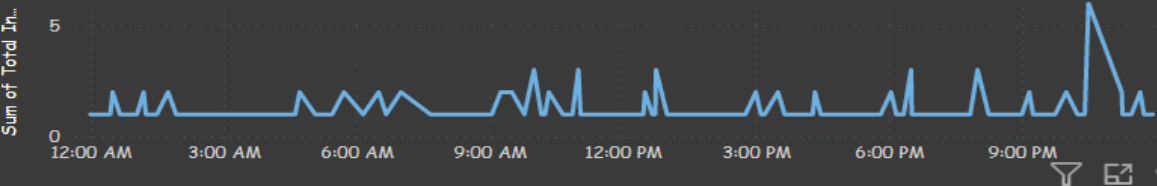
Sum of Total Incidents by Dayofweek



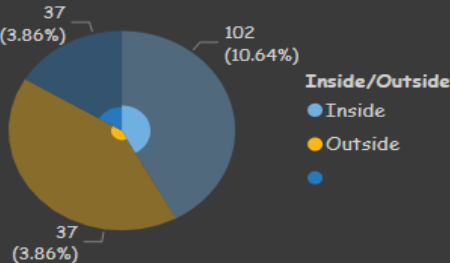
Number of incidents by Crime Date



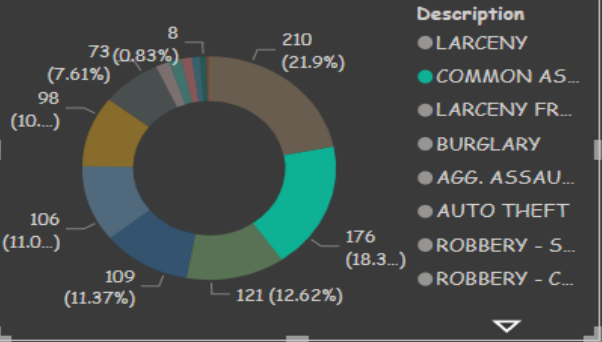
Sum of Total Incidents by CrimeTime



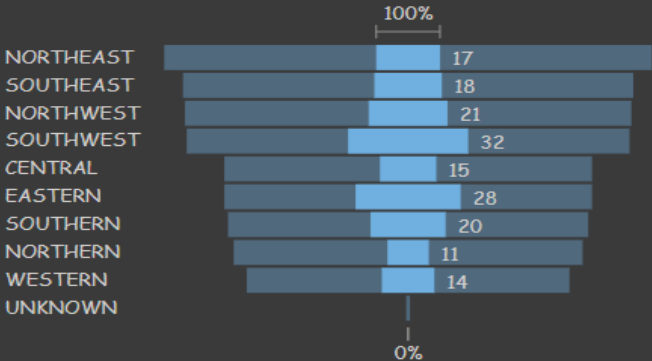
Sum of Total Incidents by Inside/Outside



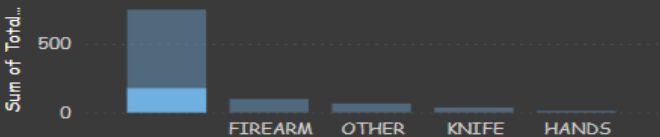
Sum of Total Incidents by Description



Number of incidents by District



Weapons Used



Rape

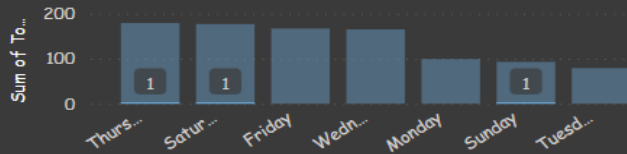
911 Baltimore Crime Data Analysis

Month
All

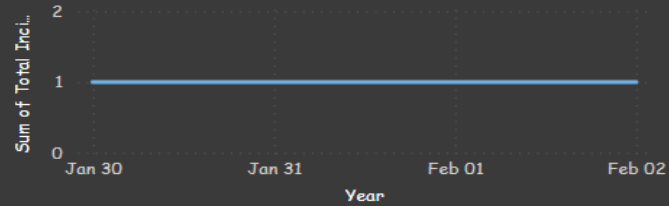
3

Total incidents

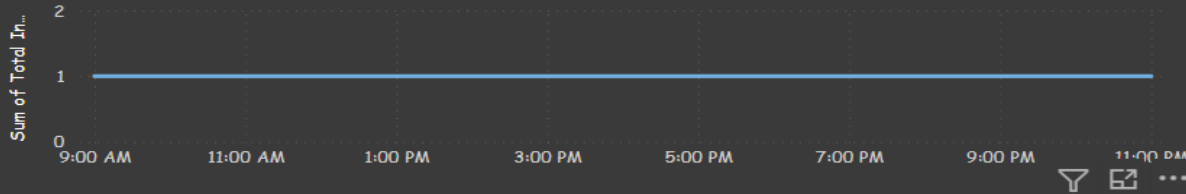
Sum of Total Incidents by Dayofweek



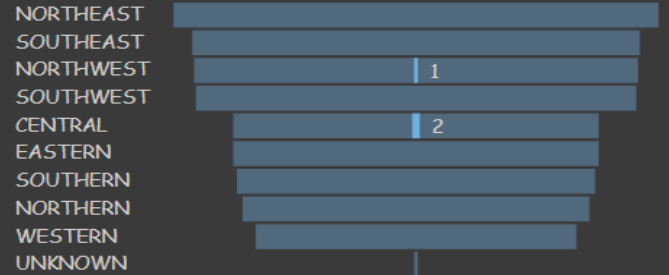
Number of incidents by Crime Date



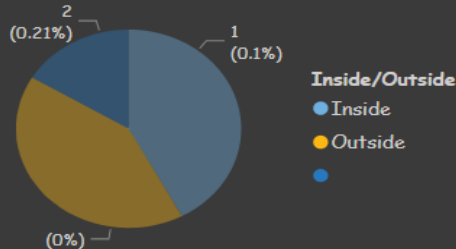
Sum of Total Incidents by CrimeTime



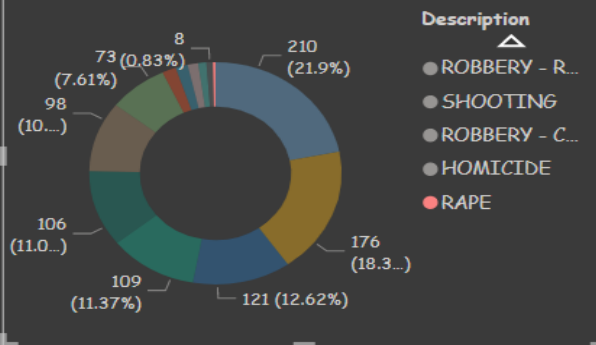
Number of incidents by District



Sum of Total Incidents by Inside/Outside



Sum of Total Incidents by Description



Weapons Used



Northeastern Region

911 Baltimore Crime Data Analysis

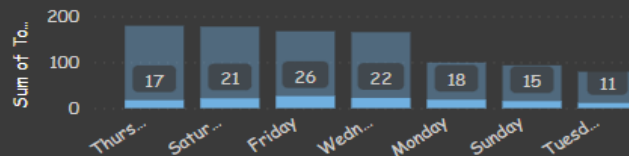
Month

All

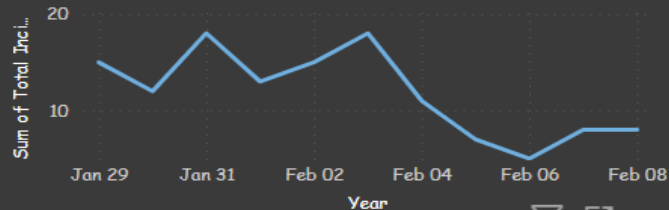
130

Total incidents

Sum of Total Incidents by Dayofweek



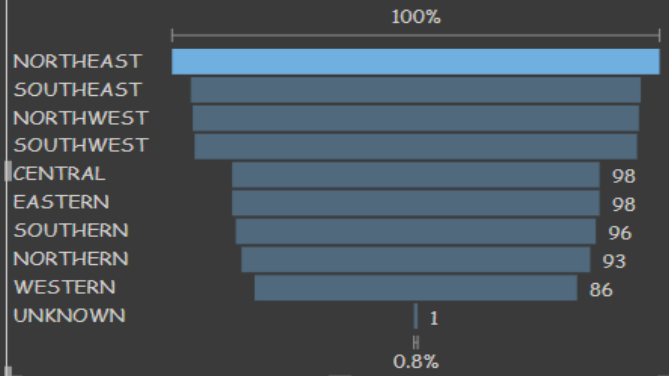
Number of incidents by Crime Date



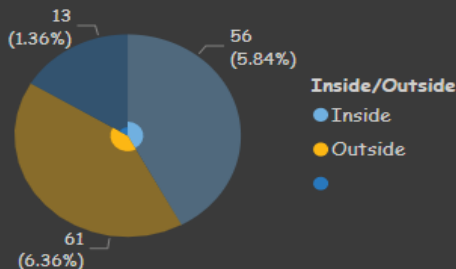
Sum of Total Incidents by CrimeTime



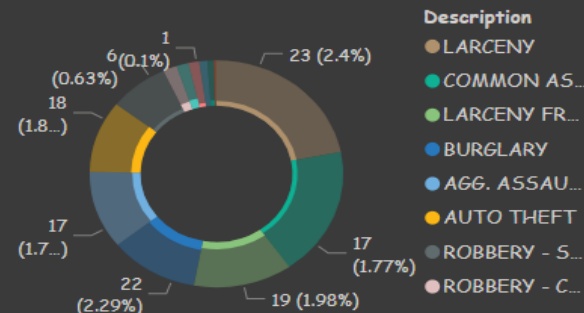
Number of incidents by District



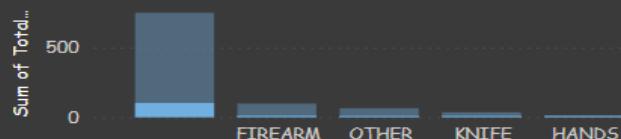
Sum of Total Incidents by Inside/Outside



Sum of Total Incidents by Description



Weapons Used



Western Region

911 Baltimore Crime Data Analysis

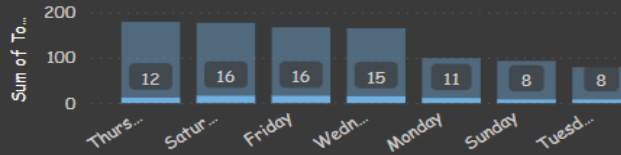
Month

All

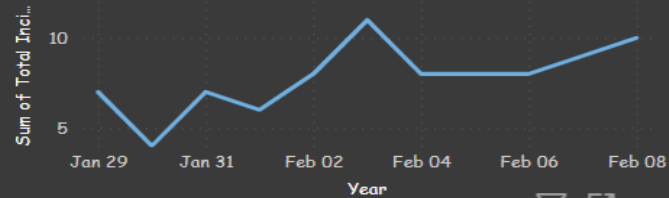
86

Total incidents

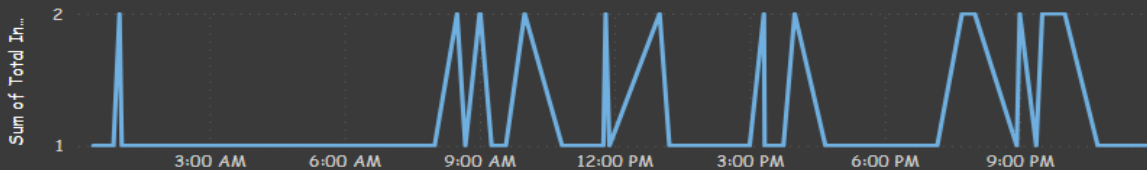
Sum of Total Incidents by Dayofweek



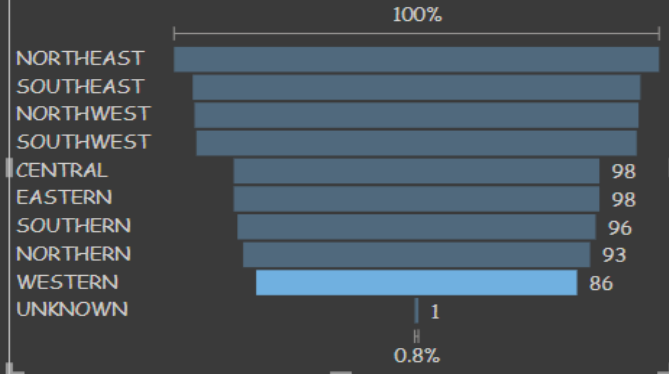
Number of incidents by Crime Date



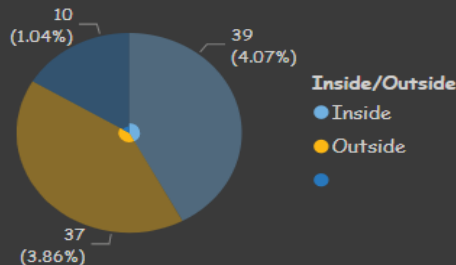
Sum of Total Incidents by CrimeTime



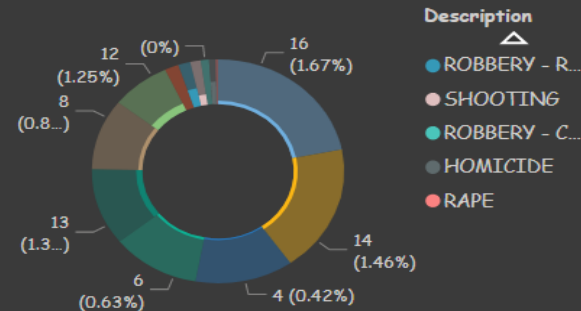
Number of incidents by District



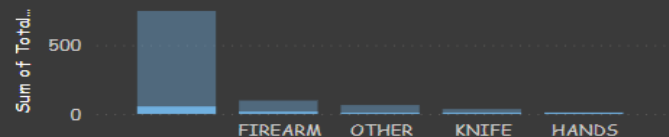
Sum of Total Incidents by Inside/Outside



Sum of Total Incidents by Description



Weapons Used



Inside

911 Baltimore Crime Data Analysis

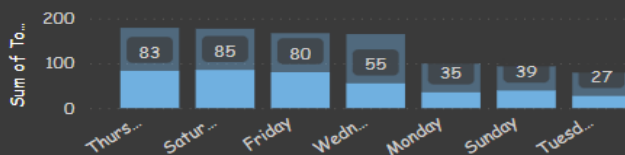
Month

All

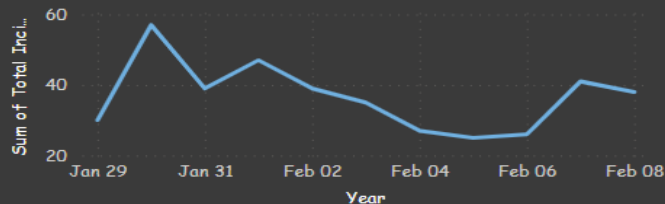
404

Total incidents

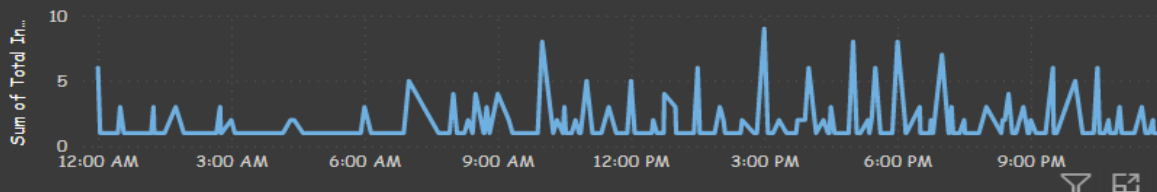
Sum of Total Incidents by Dayofweek



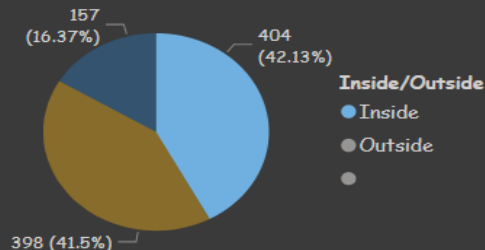
Number of incidents by Crime Date



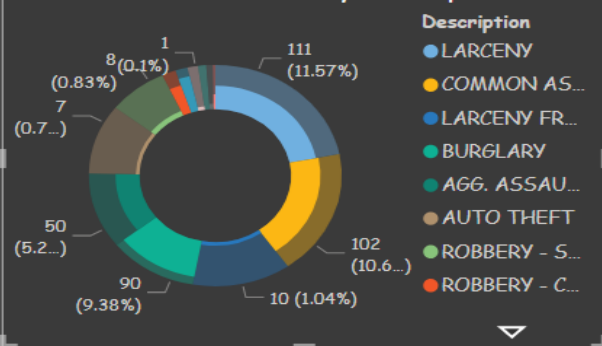
Sum of Total Incidents by CrimeTime



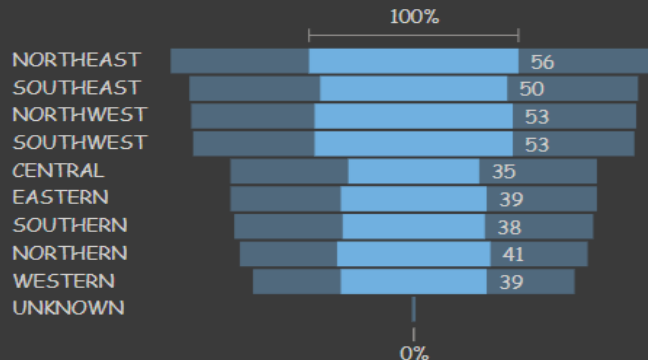
Sum of Total Incidents by Inside/Outside



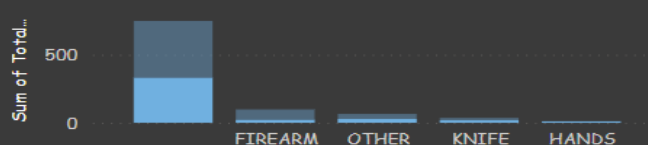
Sum of Total Incidents by Description



Number of incidents by District



Weapons Used

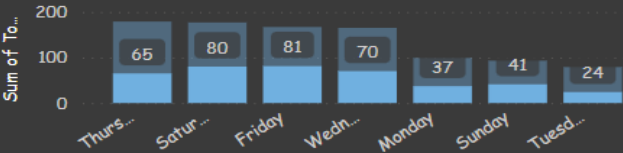


911 Baltimore Crime Data Analysis

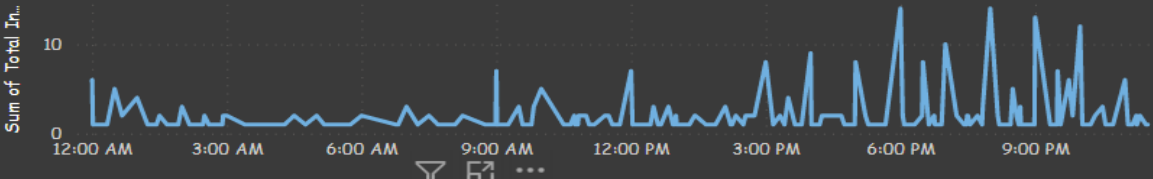
Month
All

398
Total incidents

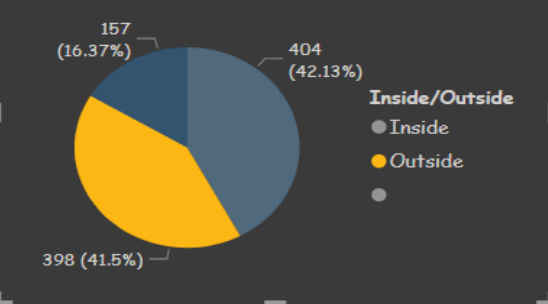
Sum of Total Incidents by Dayofweek



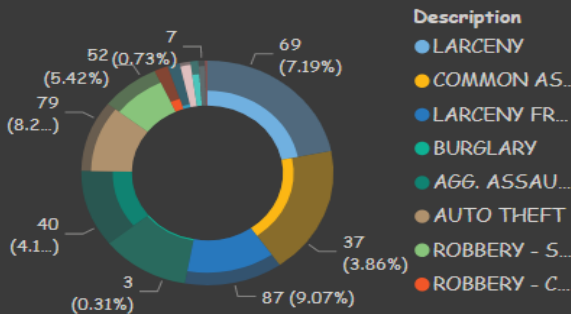
Sum of Total Incidents by CrimeTime



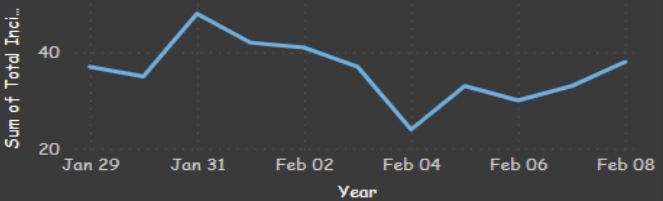
Sum of Total Incidents by Inside/Outside



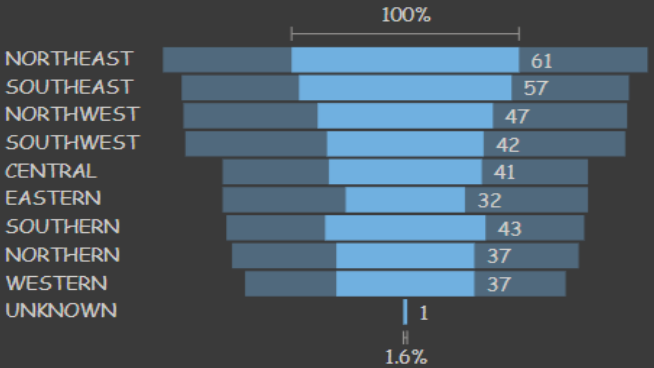
Sum of Total Incidents by Description



Number of incidents by Crime Date



Number of incidents by District



Weapons Used

