

Name:	Chaitya Arun Dobariya
UID:	2021600017
Course:	Advanced Data Visualization

Experiment 4

Aim:	<p>Create basic charts using R programming language on dataset Crime or Police / Law and Order</p> <p>Basic - Bar chart, Pie chart, Histogram, Time line chart, Scatter plot, Bubble plot</p> <p>Write observations from each chart</p>
-------------	---

1. Dataset

You can find the dataset [here](#).

Description

This dataset reflects incidents of crime in the City of Los Angeles dating back from 2010 to 2020. This data is transcribed from original crime reports that are typed on paper and therefore there may be some inaccuracies within the data. Some location fields with missing data are noted as (0°, 0°). Address fields are only provided to the nearest hundred block in order to maintain privacy. This data is as accurate as th the data in the database.

Fields

DR_NO: Division of Records Number (Integer)

API Field Name: MM/DD/YYYY. (DateTime)

DATE OCC: MM/DD/YYYY. (DateTime)

TIME OCC: In 24 hour military time. (Integer)

AREA: The LAPD has 21 Community Police Stations referred to as Geographic Areas within the department. These Geographic Areas are sequentially numbered from 1-21. (Integer)

AREA NAME: The 21 Geographic Areas or Patrol Divisions are also given a name designation that references a landmark or the surrounding community that it is responsible for. (String)

Rpt Dist No: A four-digit code that represents a sub-area within a Geographic Area. All crime records reference the "RD" that it occurred in for statistical comparisons. (Integer)

Crm Cd: Indicates the crime committed. (Same as Crime Code 1) (Integer)

Crm Cd Desc: Defines the Crime Code provided. (String)

Mocodes: Modus Operandi: Activities associated with the suspect in commission of the crime. (Integer)

Vict Age: Two character numeric. (Integer)

Vict Sex: F - Female M - Male X - Unknown. (String)

Vict Descent: Descent Code: A - Other Asian B - Black C - Chinese D - Cambodian F - Filipino G - Guamanian H - Hispanic/Latin/Mexican I - American Indian/Alaskan Native J - Japanese K - Korean L - Laotian O - Other P - Pacific Islander S - Samoan U - Hawaiian V - Vietnamese W - White X - Unknown Z - Asian Indian. (String)

Premis Cd: The type of structure, vehicle, or location where the crime took place. (Integer)

Premis Desc: Defines the Premise Code provided. (String)

Weapon Used Cd: The type of weapon used in the crime. (Integer)

Weapon Desc: Defines the Weapon Used Code provided. (String)

Status: Status of the case. (IC is the default). (Integer)

Status Desc: Defines the Status Code provided. (String)

Crm Cd 1: Indicates the crime committed. Crime Code 1 is the primary and most serious one. Crime Code 2, 3, and 4 are respectively less serious offenses. Lower crime class numbers are more serious. (Integer)

Crm Cd 2: May contain a code for an additional crime, less serious than Crime Code 1. (Integer)

Crm Cd 3: May contain a code for an additional crime, less serious than Crime Code 1. (Integer)

Crm Cd 4: May contain a code for an additional crime, less serious than Crime Code 1. (Integer)

LOCATION: Street address of crime incident rounded to the nearest hundred block to maintain anonymity. (String)

Cross Street: Cross Street of rounded Address. (Integer)

LAT: Latitude. (Float)

LON: Longitude. (Float)

Importing the libraries and data:

```
data_visualization.R x
1 # Load required libraries
2 library(ggplot2)
3 library(dplyr)
4 library(lubridate)
5 library(stringr)
```

Reading the dataset:

```
5
6 # Load the dataset
7 crime_data <- read.csv("Crime_Data_from_2020_to_Present.csv")
```

Crime_Data_from_2020_to_Present.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	DR_NO	Date Rptd	DATE OCC	TIME OCC	AREA	AREA NAME	Rpt Dist N	Part 1-2	Crm Cd	Crm Cd De Mocodes	Vict Age	Vict Sex	Vict Descent	Premis Cd	
2	10304468	01-08-2020 00:00	01-08-2020 00:00	2230	3 Southwest	377	2	624 BATTERY - 0444 0913		36 F	B			501 :	
3	190101086	01-02-2020 00:00	01-01-2020 00:00	330	1 Central	163	2	624 BATTERY - 0416 1822		25 M	H			102 :	
4	201220752	09/16/2020 12:00:00 AM	09/16/2020 12:00:00 AM	1230	12 77th Street	1259	2	745 VANDALIS 2004 1820		62 M	B			502 I	
5	191501505	01-01-2020 00:00	01-01-2020 00:00	1730	15 N Hollywood	1543	2	745 VANDALIS 0329 1402		76 F	W			502 I	
6	191921269	01-01-2020 00:00	01-01-2020 00:00	415	19 Mission	1998	2	740 VANDALIS 329		31 X	X			409 I	
7	200100501	01-02-2020 00:00	01-01-2020 00:00	30	1 Central	163	1	121 RAPE, FOR 0413 1822		25 F	H			735 I	
8	200100502	01-02-2020 00:00	01-02-2020 00:00	1315	1 Central	161	1	442 SHOPLIFT 1402 2004		23 M	H			404 I	
9	200100504	01-04-2020 00:00	01-04-2020 00:00	40	1 Central	155	2	946 OTHER MI: 1402 0392		0 X	X			726 I	
10	200100507	01-04-2020 00:00	01-04-2020 00:00	200	1 Central	101	1	341 THEFT-GR, 1822 0344		23 M	B			502 I	

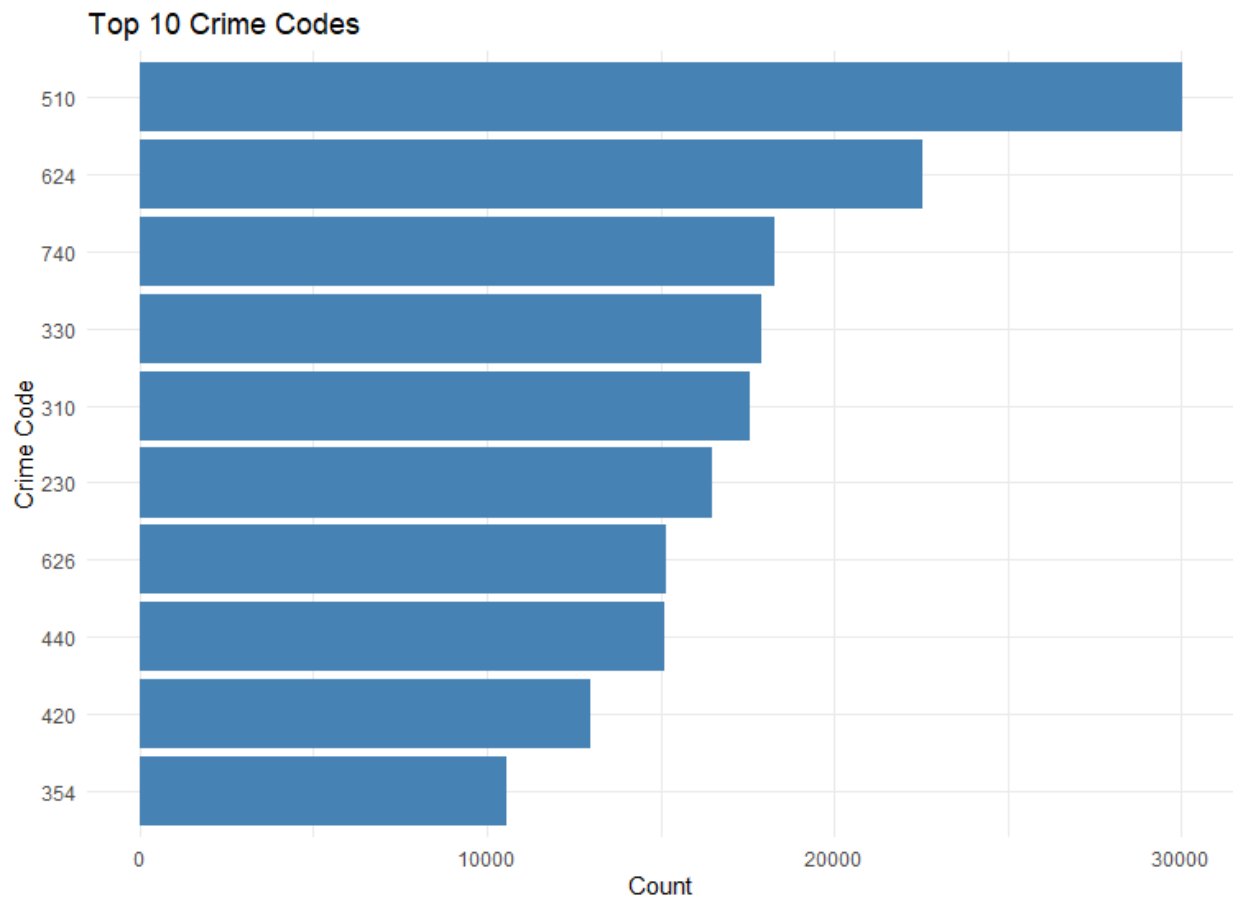
	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
1	Premis Cd	Premis De	Weapon U	Weapon D	Status	Status Des	Crm Cd 1	Crm Cd 2	Crm Cd 3	Crm Cd 4	LOCATION	Cross Stre	LAT	LON
2	501	SINGLE FA	400	STRONG-F	AO	Adult Othe	624				1100 W 39TH		34.0141	-118.298
3	102	SIDEWALK	500	UNKNOWN	IC	Invest Cor	624				700 S HILL		34.0459	-118.255
4	502	MULTI-UNIT DWELLING (APART	IC			Invest Cor	745				700 E 73RD		33.9739	-118.263
5	502	MULTI-UNIT DWELLING (APART	IC			Invest Cor	745	998			5400 CORTEEN		34.1685	-118.402
6	409	BEAUTY SUPPLY STORE		IC		Invest Cor	740				14400 TITUS		34.2198	-118.447
7	735	NIGHT CLU	500	UNKNOWN	IC	Invest Cor	121	998			700 S BROADWAY		34.0452	-118.253
8	404	DEPARTMENT STORE		IC		Invest Cor	442	998			700 S FIGUEROA		34.0483	-118.263
9	726	POLICE FACILITY		IC		Invest Cor	946	998			200 E 6TH		34.0448	-118.247
10	502	MULTI-UNIT DWELLING (APART	IC			Invest Cor	341	998			700 BERNARD		34.0677	-118.24

2. Data Preprocessing

```
8  
9 # Convert DATE_OCC to Date type  
10 crime_data$DATE_OCC <- as.Date(crime_data$DATE_OCC, format = "%m/%d/%Y")  
11
```

3. Charts & Plots

3.1 Bar Chart:

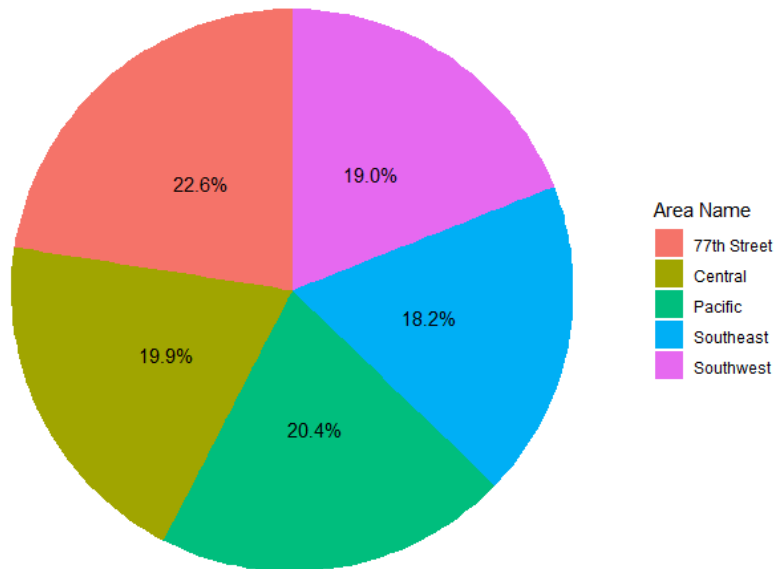


Observation:

This chart reveals the most frequent crime codes, helping identify the most common types of crimes reported. It allows for quick comparison of crime frequencies and highlights areas that may require more attention from law enforcement.

3.2 Pie Chart

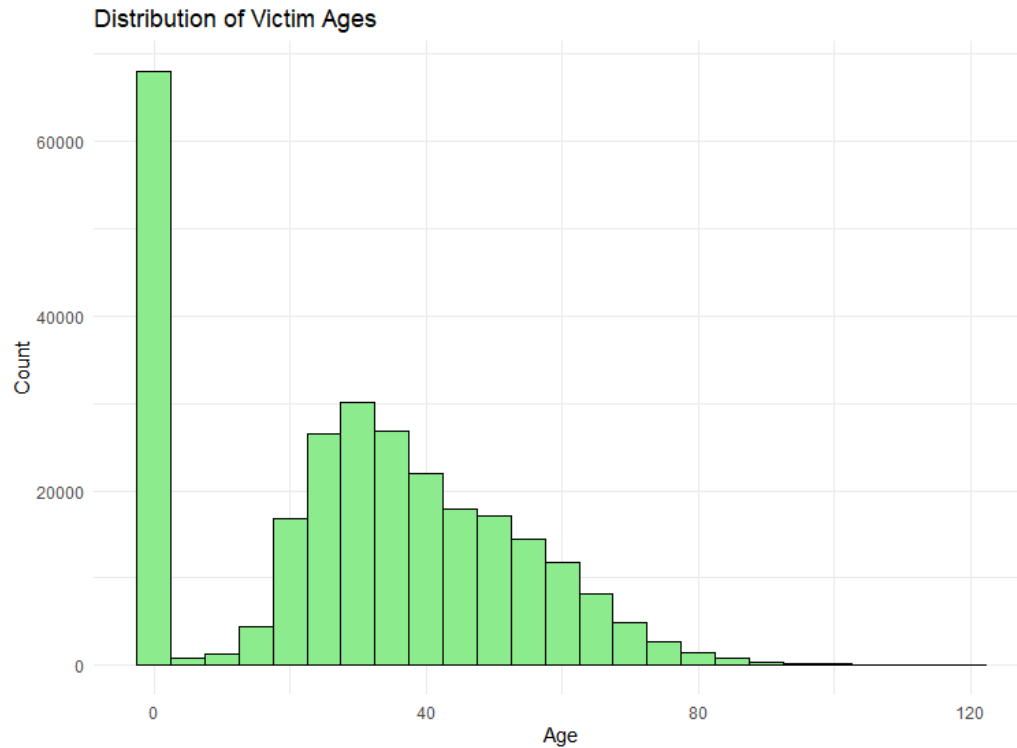
Crime Distribution by Top 5 Areas



Observation:

This visualization shows the distribution of crimes across the top 5 areas, allowing for easy comparison of crime rates between different regions. It helps identify which areas are most affected by crime and may need increased police presence or community intervention programs.

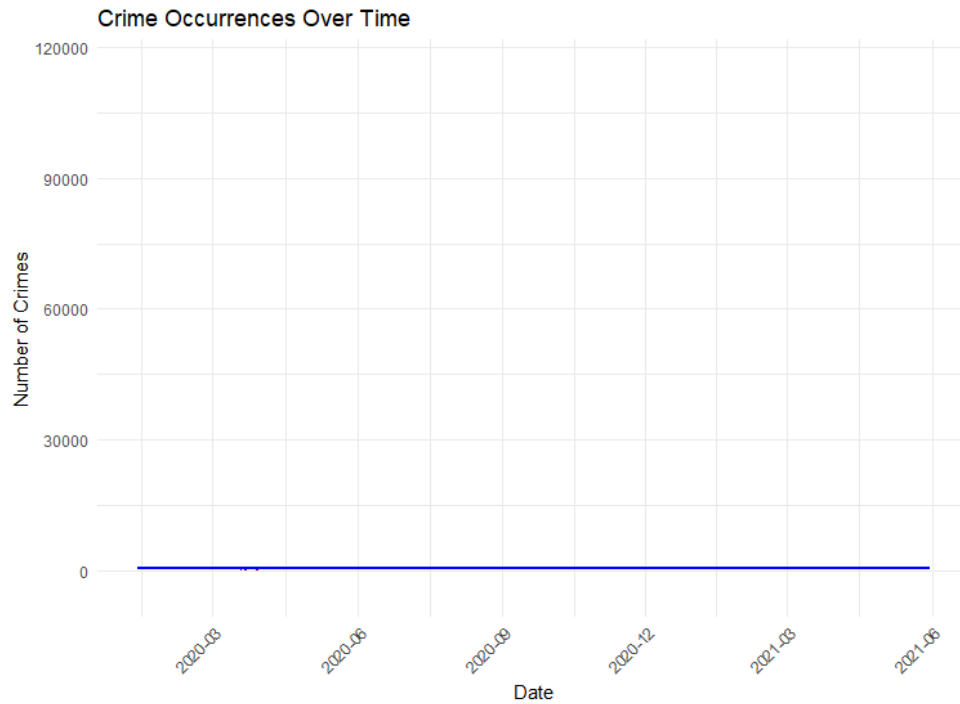
3.3 Histogram



Observation:

This histogram displays the age distribution of crime victims, revealing patterns in victimization across different age groups. It can help identify vulnerable populations and inform targeted crime prevention strategies for specific age demographics.

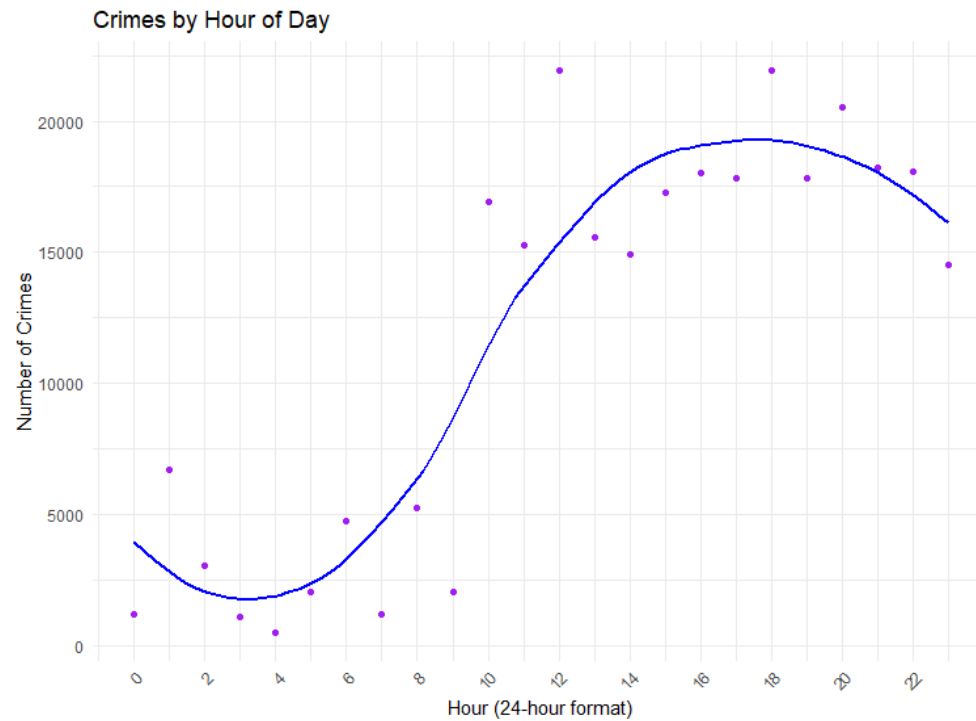
3.4 Time Line Chart



Observation:

This chart illustrates crime trends over time, showing seasonal patterns, long-term trends, or sudden spikes in criminal activity. It's useful for understanding how crime rates fluctuate and for planning resource allocation based on historical patterns.

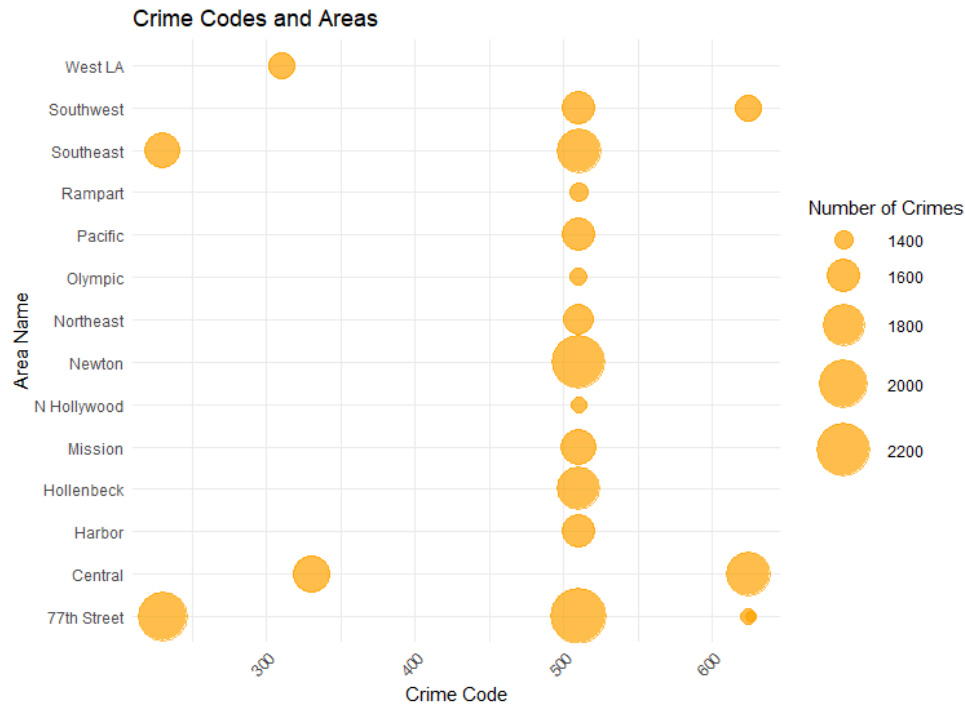
3.5 Scatter Chart



Observation:

This plot reveals how crime frequency varies throughout the day, highlighting peak hours for criminal activity. It can inform decisions about police shift scheduling and resource allocation to better address crime patterns at different times of day.

3.6 Bubble Chart



Observation:

This visualization combines crime types, areas, and frequencies in a single plot, allowing for the identification of specific crime-area combinations that are particularly problematic. It helps in understanding which types of crimes are more prevalent in certain areas, aiding in targeted crime prevention strategies.

Conclusion

This comprehensive analysis of crime data through various visualizations has provided valuable insights into the patterns and trends of criminal activity. The bar chart of top crime codes highlighted the most prevalent offenses, while the pie chart revealed the distribution of crimes across different areas. The histogram of victim ages shed light on vulnerable demographics, and the timeline chart exposed temporal trends in crime occurrences. The scatter plot of crimes by hour of day uncovered daily patterns in criminal activity, and the bubble plot effectively combined crime types, areas, and frequencies. Together, these visualizations offer a multi-faceted view of the crime landscape, enabling law enforcement and policymakers to make data-driven decisions for resource allocation, targeted interventions, and community safety initiatives. This analysis demonstrates the power of data visualization in transforming raw crime statistics into actionable insights for crime prevention and public safety strategies.