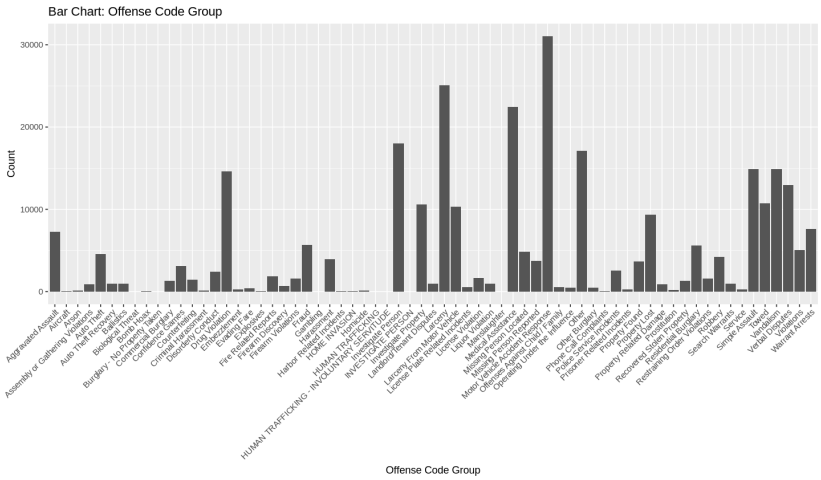
```
!pip install ggplot2 lubridate dplyr
```

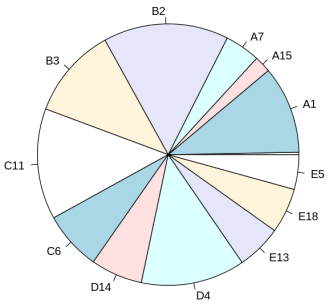
```
Error in parse(text = input): <text>:1:6: unexpected symbol
1: !pip install
         ^
Traceback:
```

```
# Install and load required packages
install.packages(c("ggplot2", "lubridate", "dplyr"))
library(ggplot2)
library(lubridate)
library(dplyr)

options(repr.plot.width = 12)

# Read the CSV data
data <- read.csv("crime.csv")

# Convert OCCURRED_ON_DATE to Date type
data$OCCURRED_ON_DATE <- as.Date(data$OCCURRED_ON_DATE, format="%d-%m-%Y")

# Preprocess latitude and longitude
data <- data %>%
  mutate(
    Lat = round(as.numeric(Lat), 3),
    Long = round(as.numeric(Long), 3)
  ) %>%
  filter(
    grepl("^42", as.character(Lat)),
    grepl("^-71", as.character(Long))
  )

# 1. Bar Chart: Offense Code Group
ggplot(data, aes(x=OFFENSE_CODE_GROUP)) +
  geom_bar() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(title="Bar Chart: Offense Code Group", x="Offense Code Group", y="Count")

# 2. Pie Chart: District
district_counts <- table(data$DISTRICT)
pie(district_counts, main="Pie Chart: Incidents by District", labels=names(district_counts))

# 3. Histogram: Hour of Incident
ggplot(data, aes(x=HOUR)) +
  geom_histogram(binwidth=1, fill="blue", color="black") +
  labs(title="Histogram: Hour of Incident", x="Hour", y="Count")

# 4. Time line chart: Incidents over time
ggplot(data, aes(x=OCCURRED_ON_DATE)) +
  geom_line(stat="count") +
  labs(title="Time Line Chart: Incidents over Time", x="Date", y="Number of Incidents")

# 5. Scatter plot: Latitude vs Longitude
ggplot(data, aes(x=Long, y=Lat)) +
  geom_point() +
  labs(title="Scatter Plot: Incident Locations", x="Longitude", y="Latitude") +
  scale_x_continuous(limits = c(-71.2, -71.0)) +
  scale_y_continuous(limits = c(42.2, 42.4))

# 6. Bubble plot: Offense Code Group, Hour, and Count
data_summary <- data %>%
  group_by(OFFENSE_CODE_GROUP, HOUR) %>%
  summarise(count = n(), .groups = 'drop')

ggplot(data_summary, aes(x=HOUR, y=OFFENSE_CODE_GROUP, size=count)) +
  geom_point(alpha=0.5) +
  labs(title="Bubble Plot: Offense Code Group by Hour", x="Hour", y="Offense Code Group", size="Count")
```

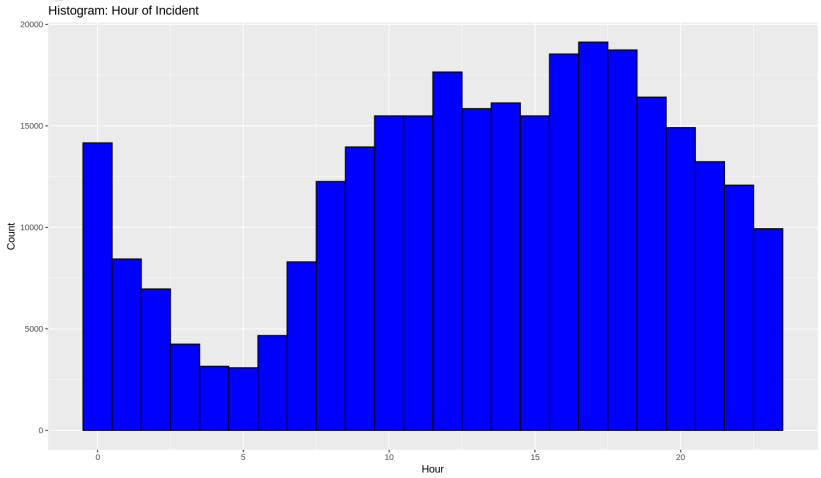
```
Installing packages into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)
```

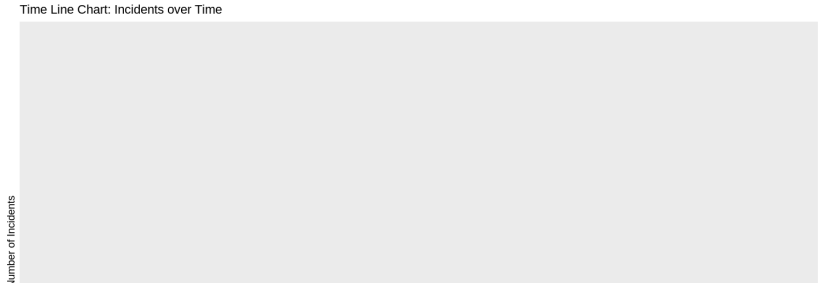


Pie Chart: Incidents by District

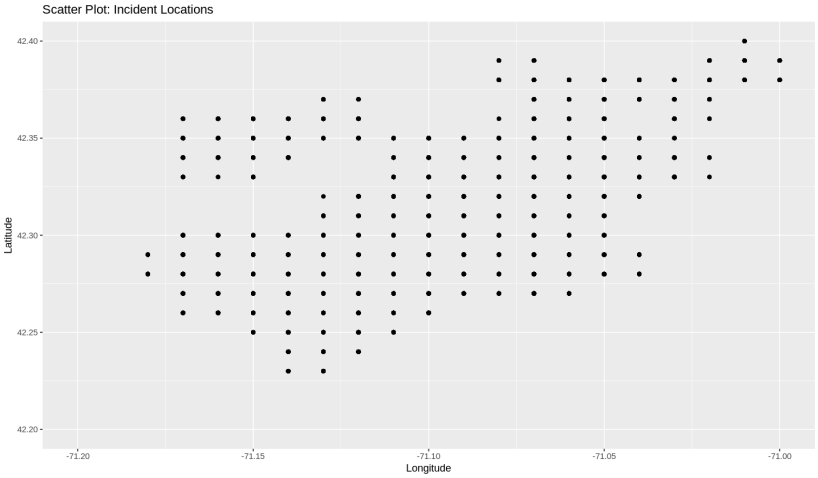


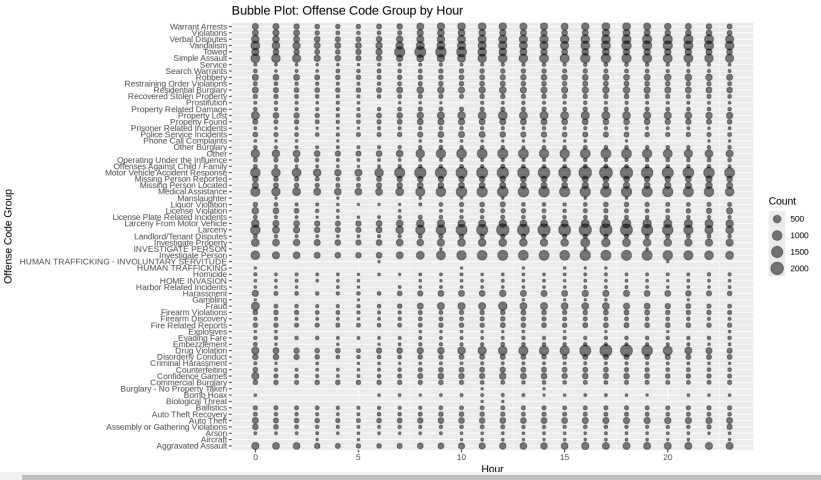

```
Warning message in min(x):
“no non-missing arguments to min; returning Inf”
Warning message in max(x):
“no non-missing arguments to max; returning -Inf”
Warning message in min(d[d > tolerance]):
“no non-missing arguments to min; returning Inf”
Warning message:
“Removed 298325 rows containing non-finite outside the scale range
(`stat_count()`).”
```






Date

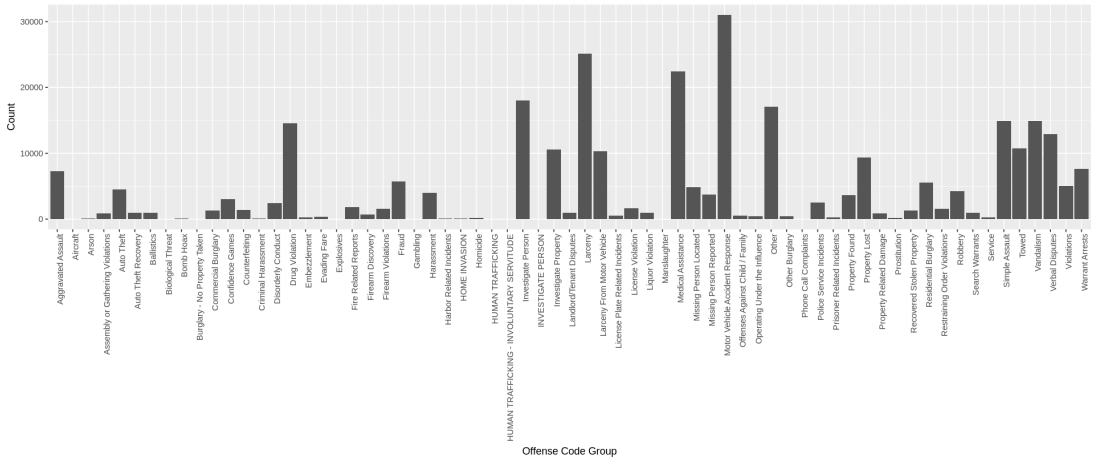

Scatter Plot: Incident Locations
Latitude
42.40
42.35
42.30
42.25
42.20
-71.20
-71.15
-71.10
-71.05
-71.00
Longitude


Bubble Plot: Offense Code Group by Hour
Offense Code Group
Warrant Arrests
Violations
Verbal Disputes
Vandalism
Towed
Simple Assault
Service
Search Warrant
Robbery
Restraining Order Violations
Residential Burglary
Recovered Stolen Property
Prostitution
Property Related Damage
Property Lost
Property Found
Prisoner Related Incidents
Police Service Incidents
Phone Call Complaints
Other Burglary
Other
Operating Under the Influence
Offenses Against Child / Family
Motor Vehicle Accident Response
Missing Person Reported
Missing Person Located
Medical Assistance
Manslaughter
Liquor Violation
License Plate Related Incidents
Larceny From Motor Vehicle
Larceny
Landlord/Tenant Disputes
Investigate Property
INVESTIGATE PERSON
Investigate Person
HUMAN TRAFFICKING - INVOLUNTARY SERVITUDE
HUMAN TRAFFICKING
Homicide
HOME INVASION
Harbor Related Incidents
Harassment
Gambling
Fraud
Firearm Violations
Firearm Discovery
Fire Related Reports
Explosives
Evading Fare
Embezzlement
Drug Violation
Disorderly Conduct
Criminal Harassment
Counterfeiting
Confidence Games
Commercial Burglary
Burglary - No Property Taken
Bomb Hoax
Biological Threat
Auto Theft
Auto Theft Recovery
Assembly or Gathering Violations
Arson
Aircraft
Aggravated Assault
0
5
10
15
20
Hour
Count
500
1000
1500
2000

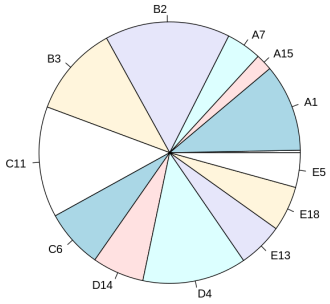
```r
# Install and load required packages
install.packages(c("ggplot2", "dplyr", "lubridate"))
library(ggplot2)
library(dplyr)
library(lubridate)

options(repr.plot.width = 16)

# Read the CSV data
data <- read.csv("crime.csv")

# Preprocess latitude and longitude
data <- data %>%
  mutate(
    Lat = round(as.numeric(Lat), 2),
    Long = round(as.numeric(Long), 2)
  ) %>%
  filter(
    grepl("^42", as.character(Lat)),
    grepl("^-71", as.character(Long))
  )

# 1. Bar Chart: Offense Code Group
ggplot(data, aes(x=OFFENSE_CODE_GROUP)) +
  geom_bar() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
  labs(title="Bar Chart: Offense Code Group", x="Offense Code Group", y="Count")

# 2. Pie Chart: District
district_counts <- table(data$DISTRICT)
pie(district_counts, main="Pie Chart: Incidents by District", labels=names(district_counts))

# 3. Histogram: Hour of Incident
ggplot(data, aes(x=HOUR)) +
  geom_histogram(binwidth=1, fill="blue", color="black") +
  labs(title="Histogram: Hour of Incident", x="Hour", y="Count")

# 4. Time line chart: Incidents by Month | TO BE CHANGED
data %>%
  group_by(YEAR, MONTH) %>%
  summarise(count = n(), .groups = 'drop') %>%
  mutate(date = as.Date(paste(YEAR, MONTH, "01", sep="-"))) %>%
  ggplot(aes(x=date, y=count)) +
  geom_line() +
  labs(title="Time Line Chart: Incidents by Month", x="Date", y="Number of Incidents")

# 5. Scatter plot: Latitude vs Longitude
ggplot(data, aes(x=Long, y=Lat)) +
  geom_point() +
  labs(title="Scatter Plot: Incident Locations", x="Longitude", y="Latitude") +
  scale_x_continuous(limits = c(-71.2, -71.0)) +
  scale_y_continuous(limits = c(42.2, 42.4))

# 6. Bubble plot: Day of Week, Hour, and Count
data %>%
  group_by(DAY_OF_WEEK, HOUR) %>%
  summarise(count = n(), .groups = 'drop') %>%
  ggplot(aes(x=HOUR, y=DAY_OF_WEEK, size=count)) +
  geom_point(alpha=0.5) +
  labs(title="Bubble Plot: Incidents by Day of Week and Hour", x="Hour", y="Day of Week", size="Count")

# 7. Heat map: Month vs Hour
data %>%
  group_by(MONTH, HOUR) %>%
  summarise(count = n(), .groups = 'drop') %>%
  ggplot(aes(x=HOUR, y=factor(MONTH), fill=count)) +
  geom_tile() +
  scale_fill_gradient(low="white", high="red") +
  labs(title="Heat Map: Incidents by Month and Hour", x="Hour", y="Month", fill="Count")


# 6. Bubble plot: Offense Code Group, Hour, and Count
data_summary <- data %>%
  group_by(OFFENSE_CODE_GROUP, HOUR) %>%
  summarise(count = n(), .groups = 'drop')

ggplot(data_summary, aes(x=HOUR, y=OFFENSE_CODE_GROUP, size=count)) +
  geom_point(alpha=0.5) +
  labs(title="Bubble Plot: Offense Code Group by Hour", x="Hour", y="Offense Code Group", size="Count")
```

```
Installing packages into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)
```

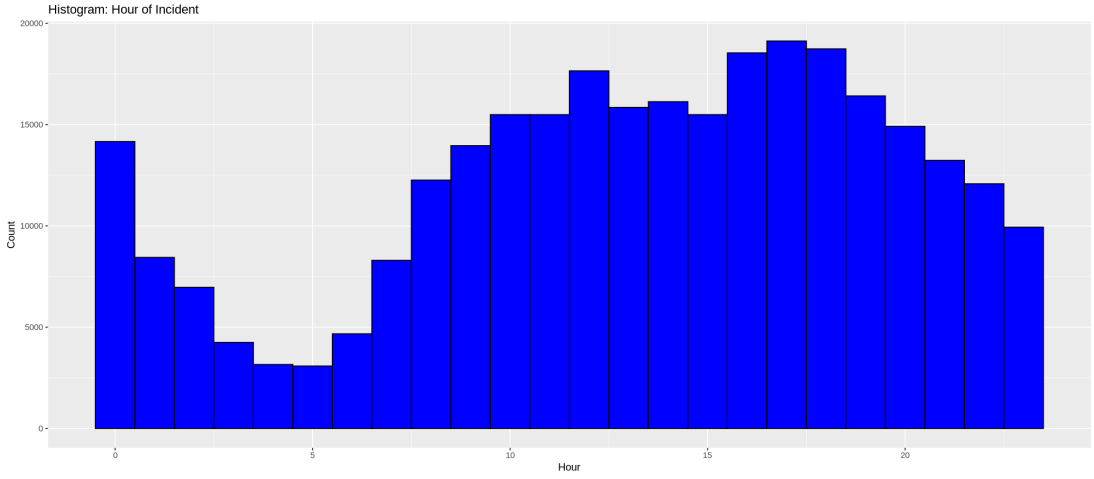
Bar Chart: Offense Code Group



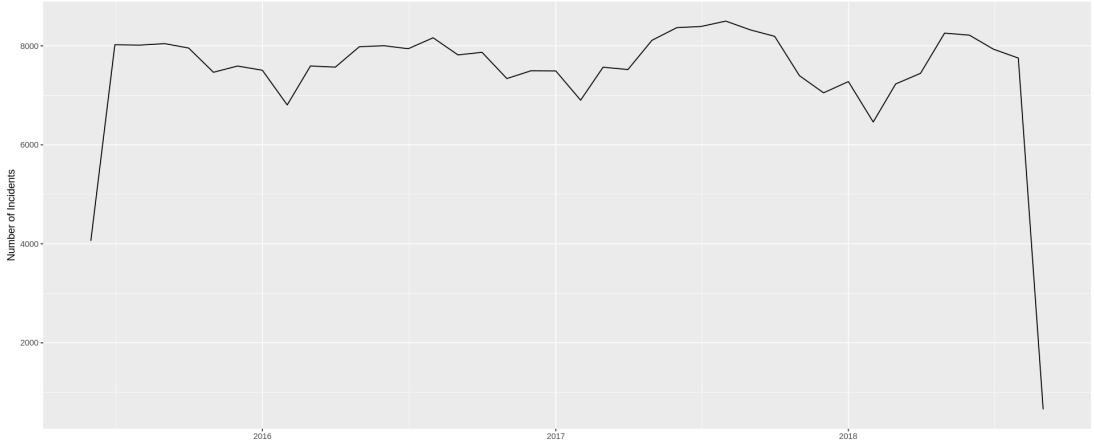
Pie Chart: Incidents by District



Histogram: Hour of Incident

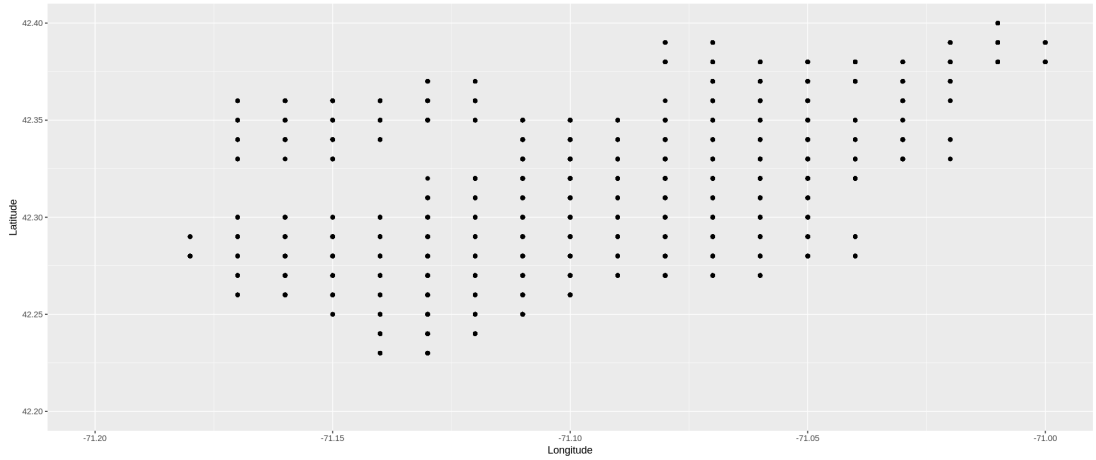


Time Line Chart: Incidents by Month



Date

Scatter Plot: Incident Locations



Bubble Plot: Incidents by Day of Week and Hour