Analyze any csv dataset using R.

```r
library (dplyr)
library (ggplot2)

setwd ("M: /rlanguage /rp")
getwd ()

data <- read.csv (" murd1.csv ")
  View (data)
  head (data)
  tail (data, 10)
  str (data)
  summary (data)
  data $ state.length
  sum (is.na(data))
  ggplot (data, aes(y = state, n = murders)) + geom_bar(
                    stat = "identity")

  ggplot (data, aes(y = state, n = gunmurders)) +
  geom_bar (stat = "identity")

  ggplot (data, aes(y = PopulationDensity, n = murders,
      group = PopulationDensity, colour = murders))
      + geom_line() + goem_point()
```

Name - Gungan Panday   University-Roll no - 2101074

```
ggplot ( data, aes (n = murders, y = region)) + geom_boxplot()

    data_size <- factor(data)
    Shr ( data_size )
    summary (data_size)
    levels ( data - size )
    data - table ( <- table( data_size)
    pie ( data_table)
    ggplot ( data, aes(y= murders, fill = region, n=region)) +
    geom_bar ( width = 1, stat = "identity") + coord_polar(
        "n", start = 0)
```
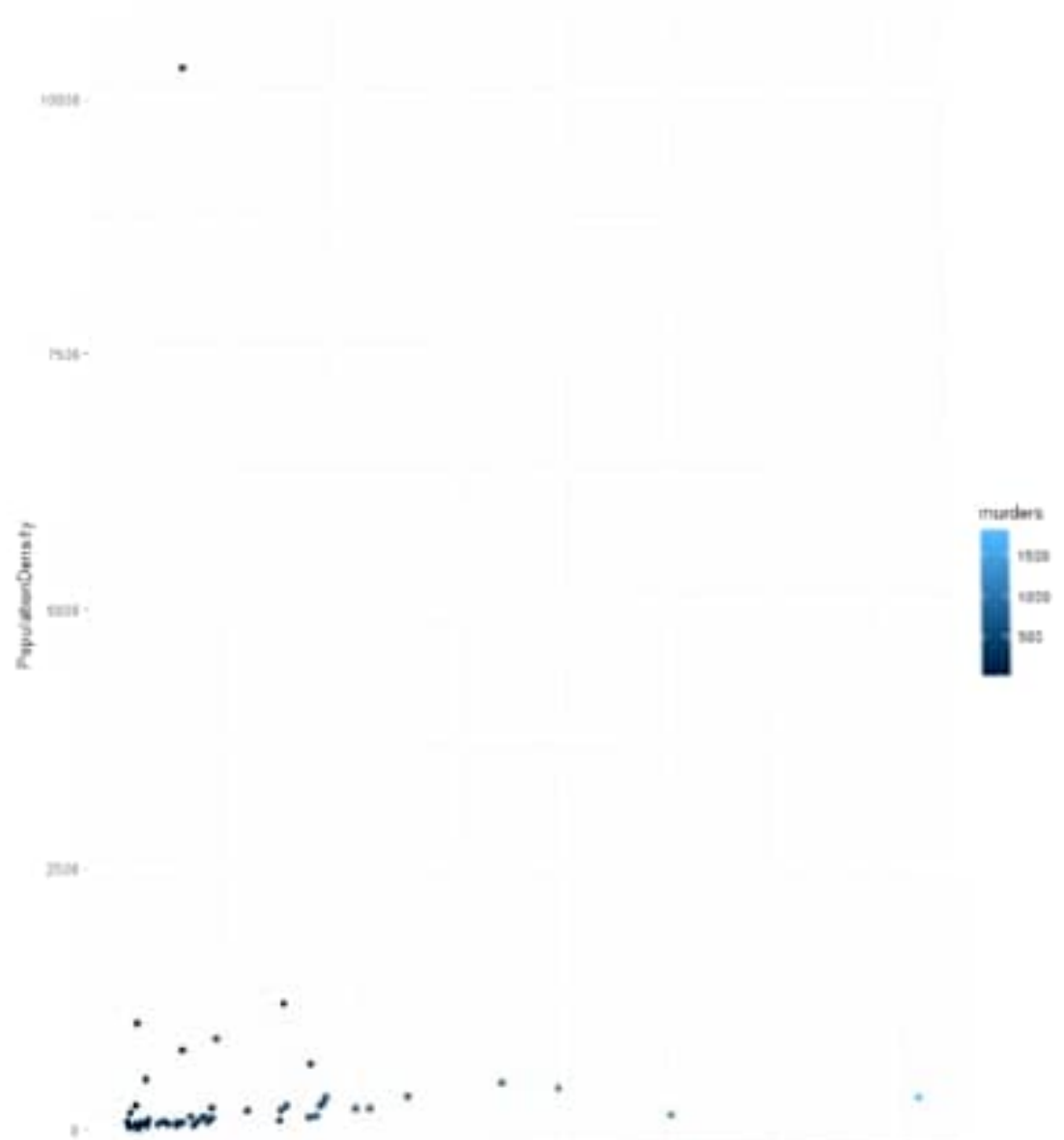
Name - Gungan Panday   University - Roll no - 2101074

#scatterplot

```
ggplot(data, aes(y =PopulationDensity, x
=murders, group=PopulationDensity,
colour=murders)) +geom_line() + geom_point()
```
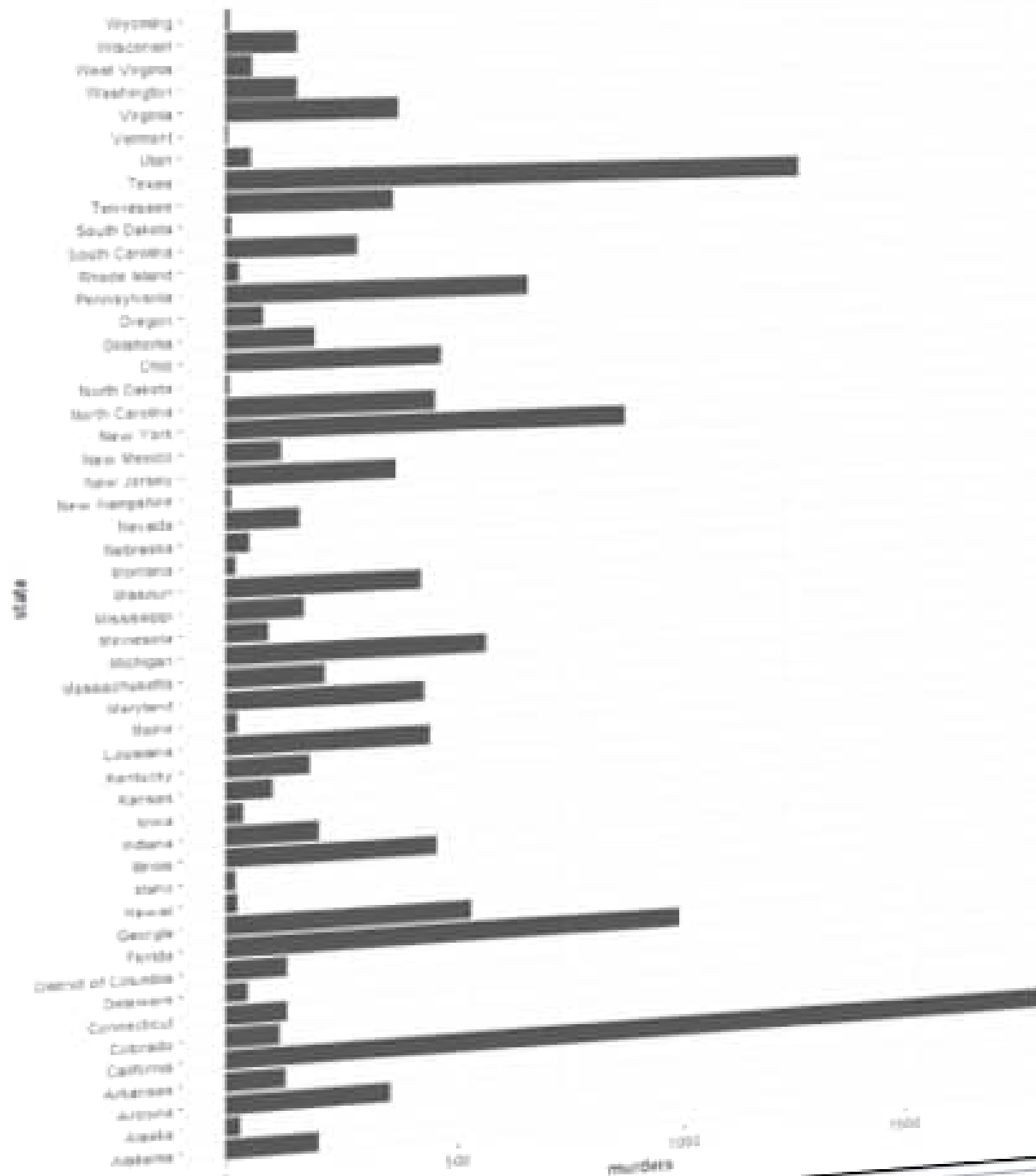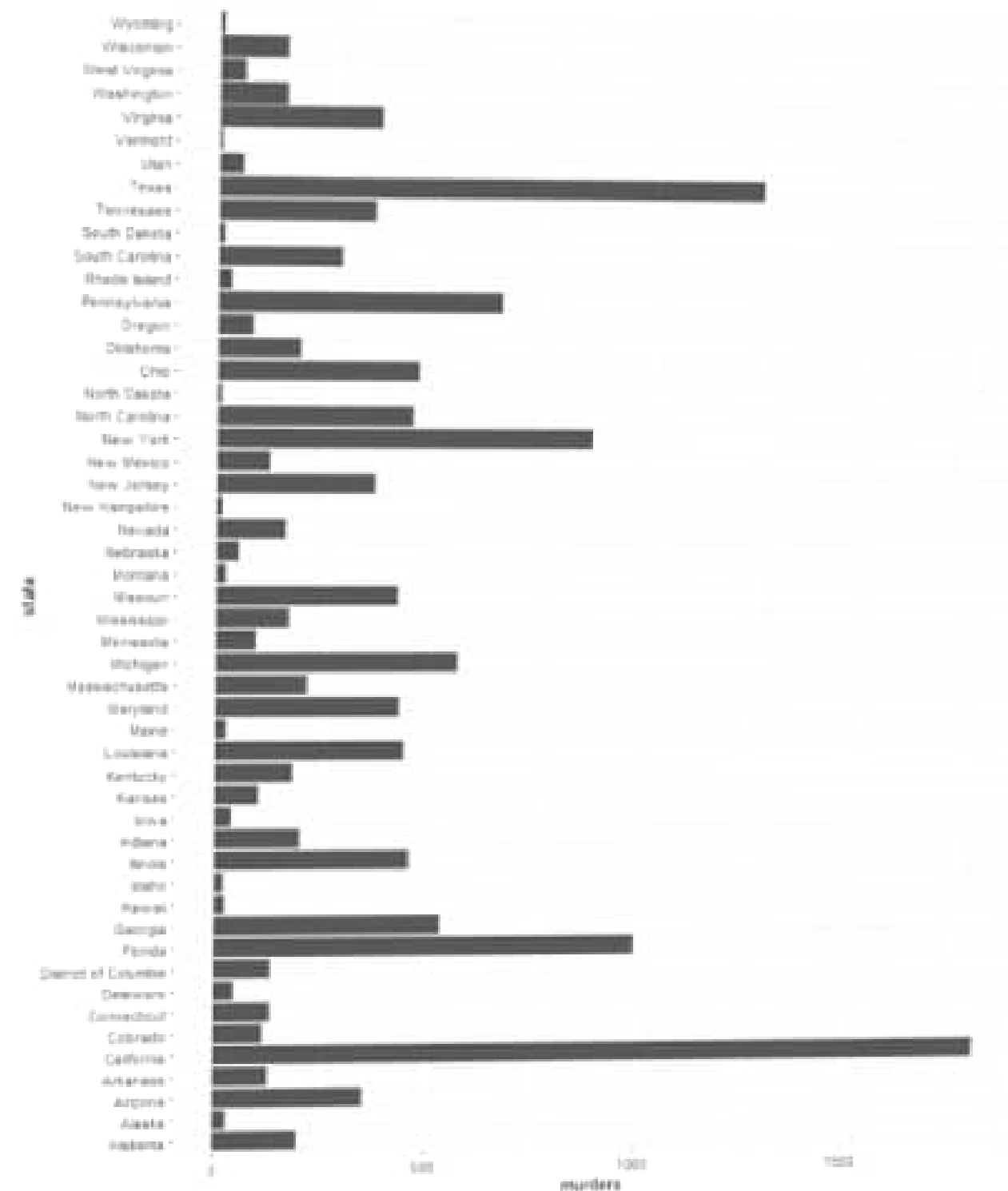
#piechart

```
ggplot(data,aes(y=murders,fill=region,x=region)) +
geom_bar(width=1,stat="identity")+coord_polar("x",s
tart=0)
```

ggplot(data, aes(y=state, x=murders))+
geom_bar(stat = "identity")

ggplot(data, aes(y=state, x=murders))+
geom_bar(stat = "identity")

Gungan Pandey   University rull no — 2101074

Disscus Descriptive and Infrential Statistics of above
dataset.

Descriptive Statistics —

Descriptive statistics are a part of statistics that
can be used to describe data. It is used to
summarise the attributes of a sample in such a
way that a pattern can be drawn from the group.
It enables ref researchers to present data in a
more meaningful way such that easy interpretation
can be made.

population —

    min —   563626

    mean —   6075769

    median —   4339362

populationdensity —

    min —   1.264

    mean —   394.549

    median   102.600

    Man —   10298.000

    1st Qu —   46.105

    3rd Qu —   224.350

Inferential statistics -

Inferential statistics is a branch of statistics that is used to make inferences about the population by analyzing a sample. When the population data is very large it becomes difficult to use it. In such cases, certain samples are taken that are representative of the entire population. Inferential statistics draws conclusions regarding the population using these samples.

Sampling stategies such as simple random sampling, cluster sampling, stratified sampling, and systematic sampling need to be used in order to choose correct samples from the population. Some methodlogies used in inferential statistic's are as follows -

Hypothesis Testing   (z tent, y tent)

Regression Analysis (check relationship b/w
                        dependent & independent)