National University of Computer & Emerging Sciences, Islamabad

Fall-2022

Assignment No. 2

Name: *Syeda Areeba Nadeem*

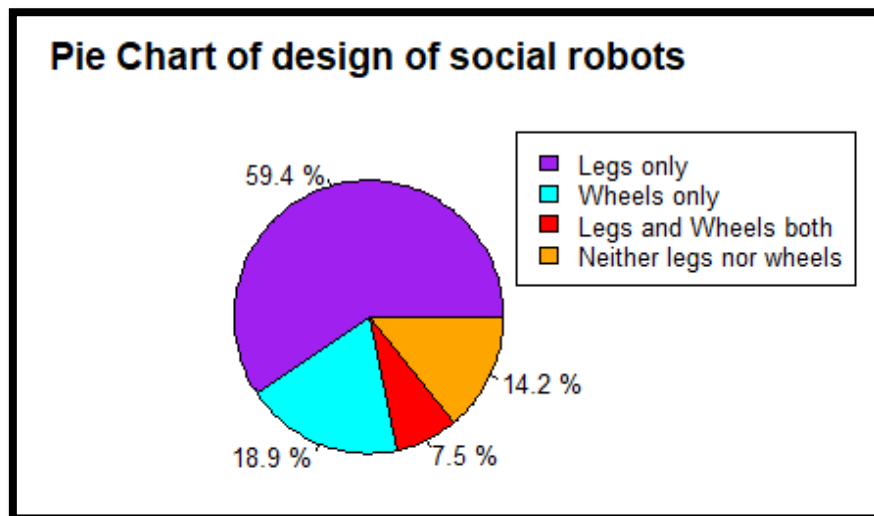**Section:**   *BS(AI) - K*

Roll No: *21I-0307*

Course Name: *Probability and Statistics*

Course Instructor: *Ma'am Sana Saleh*
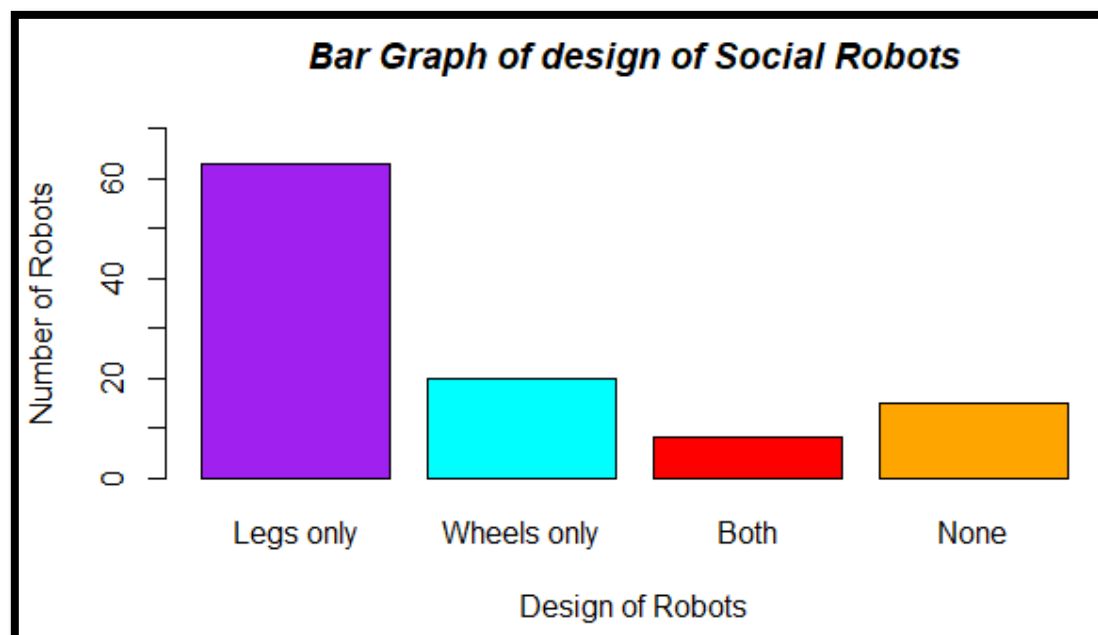
Date of Submission:

09th October, 2022

Question no.1:

Conclusion:

Hence, we can conclude that out of 106 social robots who were selected in the sample to study the trend in their design **59.4%** were built with legs only, **18.9%** with wheels only, **7.5%** with both legs and wheels, and **14.2%** with neither legs nor wheels. Therefore, the majority of the social robots were built with legs only while least were built with both legs and wheels.
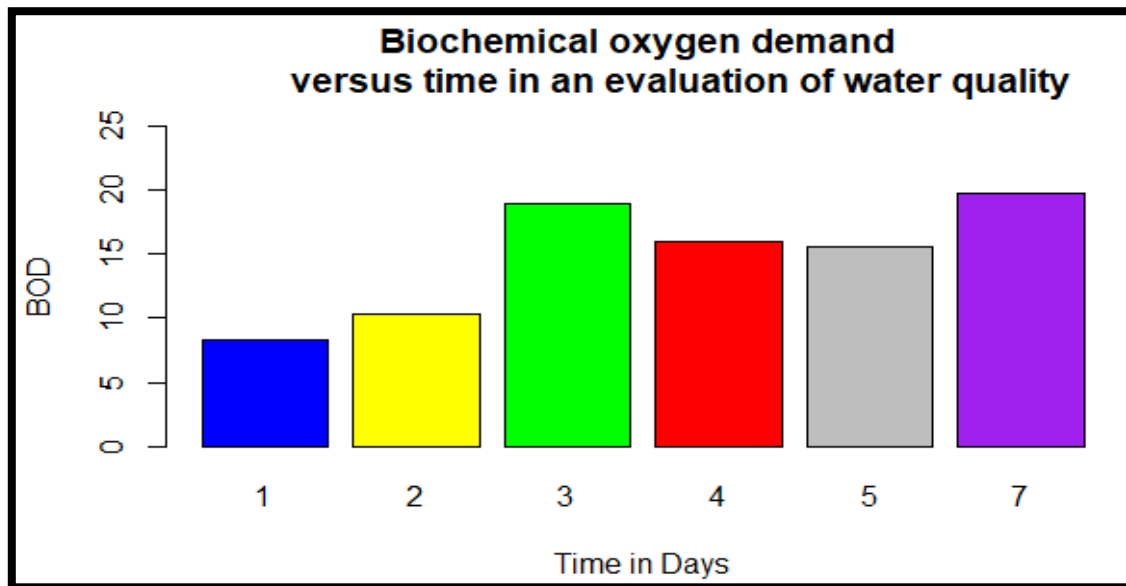
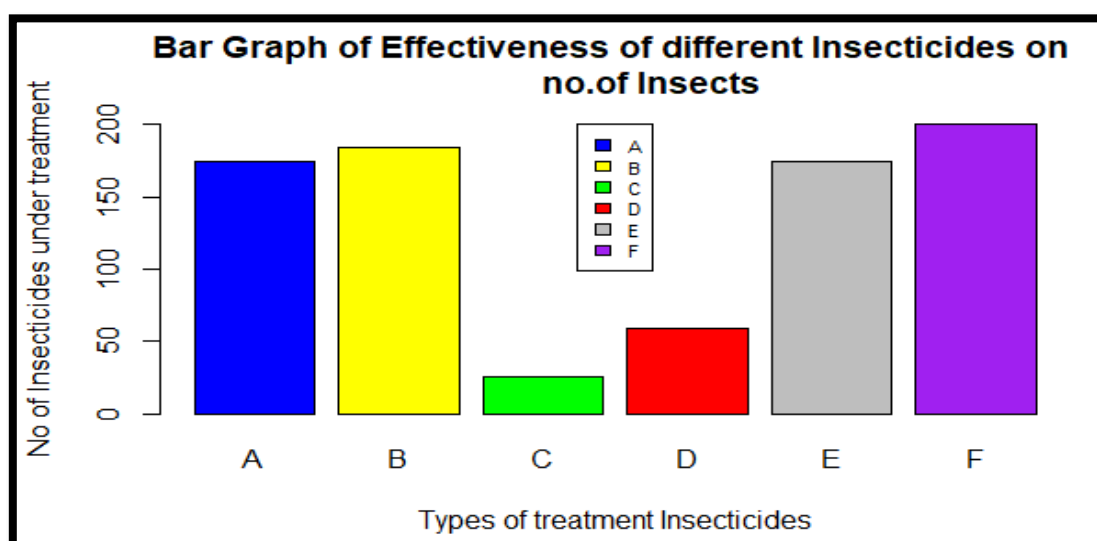The bar plot for above data is:

Question no.2:

Part (a):

Two different built-in datasets have been taken for this part namely, **BOD** and **InsectSprays**.

For BOD:

We can conclude that highest BOD was observed when the time taken for the evaluation of water quality was 7 days.

For InsectSprays:

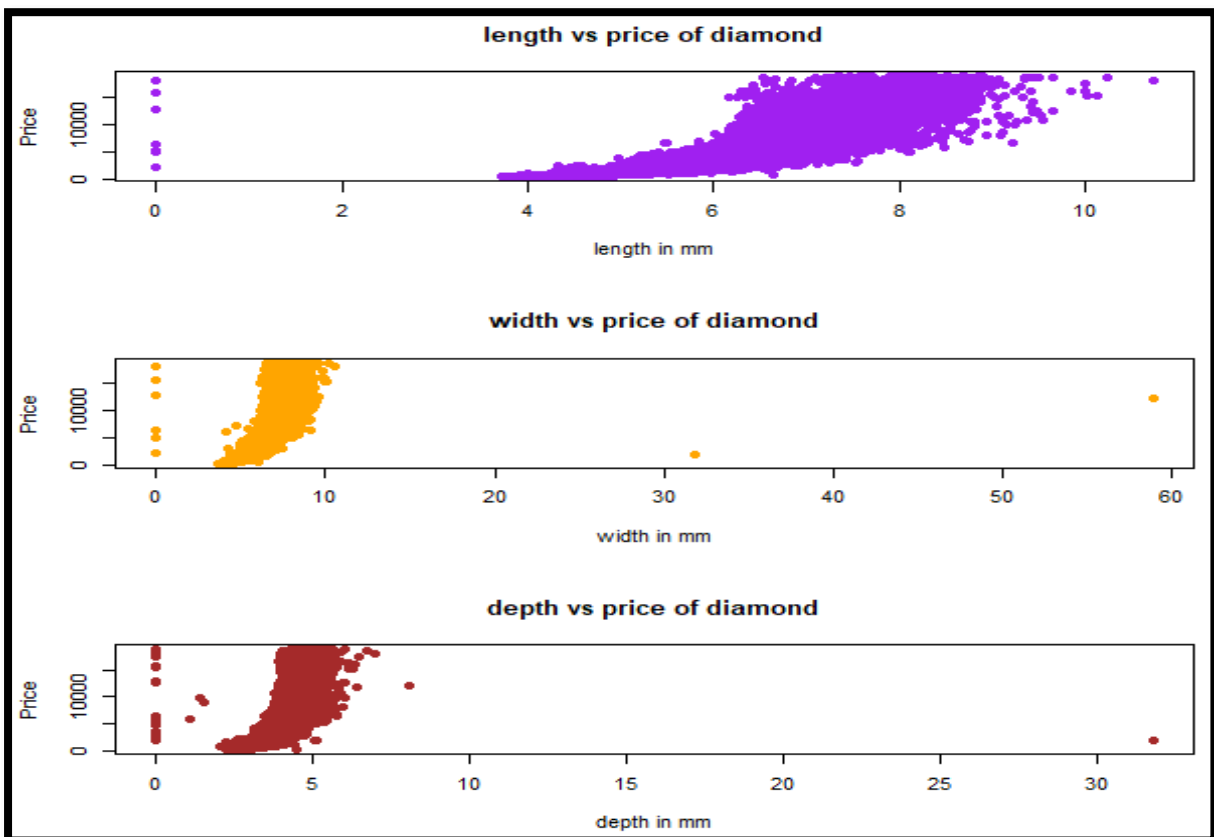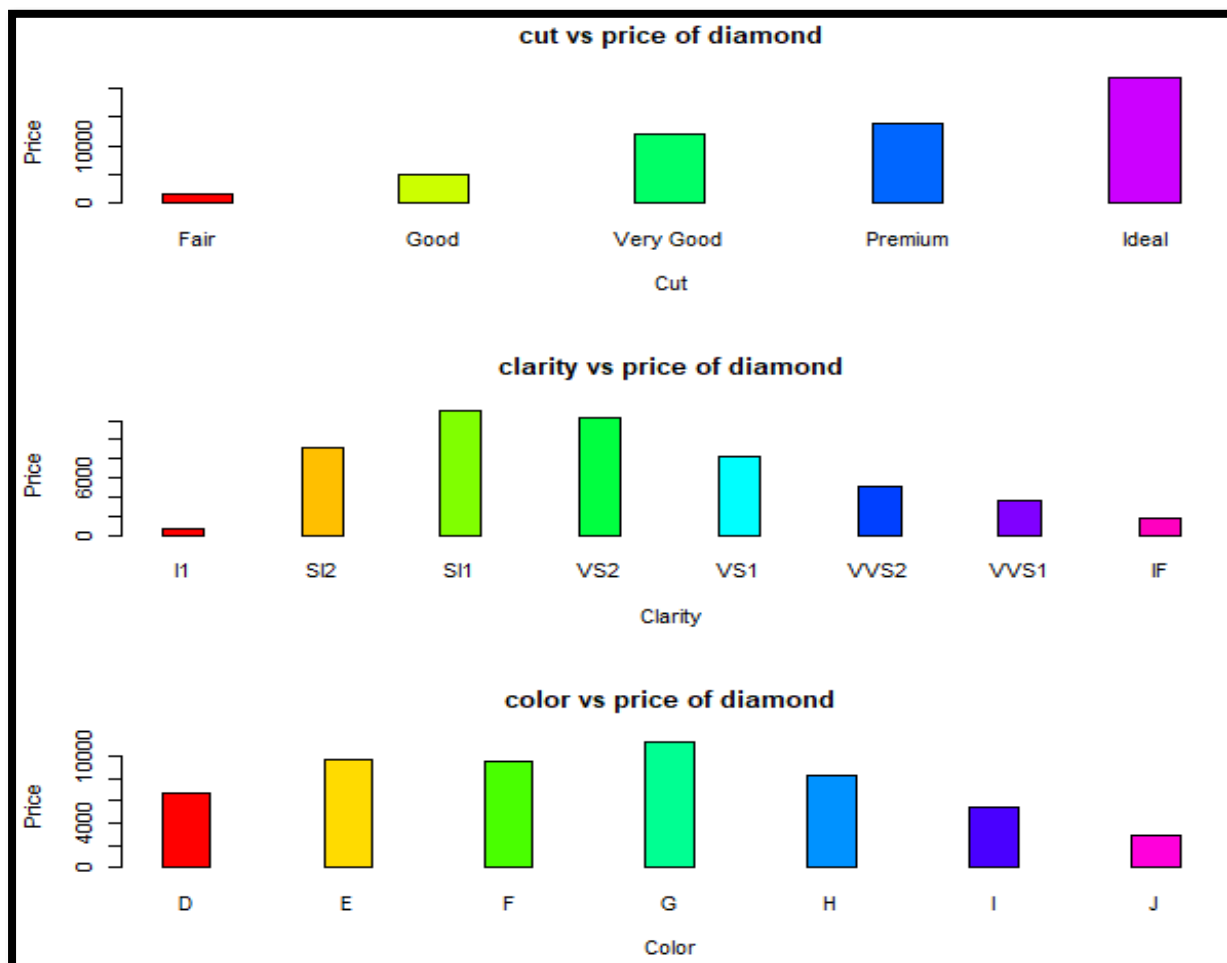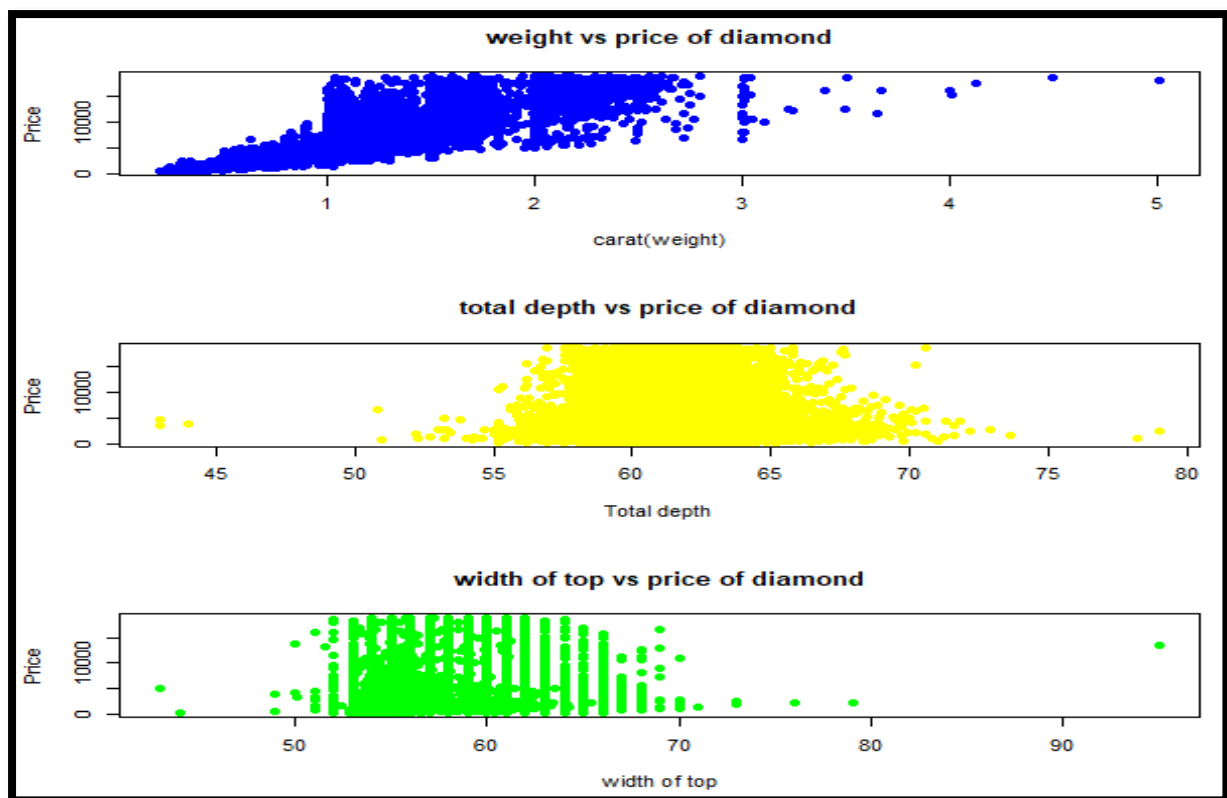We can conclude that type **F** insecticide was the most effective.

Part (b):

```
> summary(diamonds)
     carat                cut           color        clarity          depth
 Min.   :0.2000   Fair     : 1610   D: 6775   SI1    :13065   Min.   :43.00
 1st Qu.:0.4000   Good     : 4906   E: 9797   VS2    :12258   1st Qu.:61.00
 Median :0.7000   Very Good:12082   F: 9542   SI2    : 9194   Median :61.80
 Mean   :0.7979   Premium  :13791   G:11292   VS1    : 8171   Mean   :61.75
 3rd Qu.:1.0400   Ideal    :21551   H: 8304   VVS2   : 5066   3rd Qu.:62.50
 Max.   :5.0100                     I: 5422   VVS1   : 3655   Max.   :79.00
                                    J: 2808   (Other): 2531
     table           price            x                y                z
 Min.   :43.00   Min.   :  326   Min.   : 0.000   Min.   : 0.000   Min.   : 0.000
 1st Qu.:56.00   1st Qu.:  950   1st Qu.: 4.710   1st Qu.: 4.720   1st Qu.: 2.910
 Median :57.00   Median : 2401   Median : 5.700   Median : 5.710   Median : 3.530
 Mean   :57.46   Mean   : 3933   Mean   : 5.731   Mean   : 5.735   Mean   : 3.539
 3rd Qu.:59.00   3rd Qu.: 5324   3rd Qu.: 6.540   3rd Qu.: 6.540   3rd Qu.: 4.040
 Max.   :95.00   Max.   :18823   Max.   :10.740   Max.   :58.900   Max.   :31.800
```


4

weight vs price of diamond

total depth vs price of diamond

width of top vs price of diamond

cut vs price of diamond
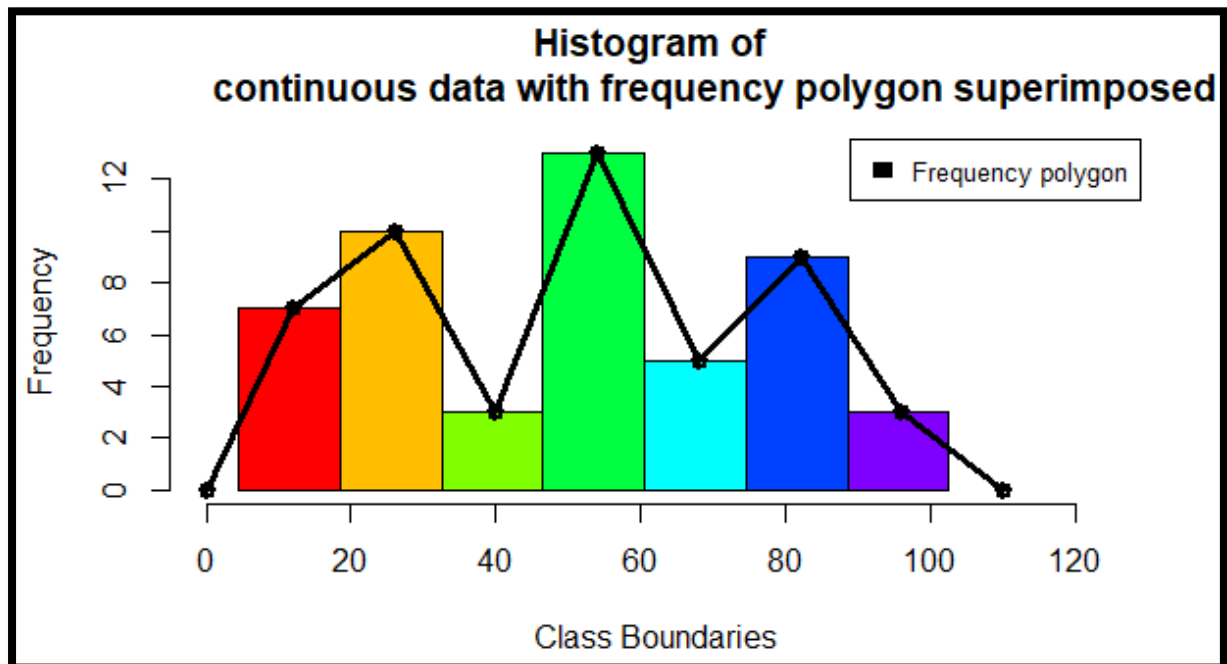
clarity vs price of diamond

color vs price of diamond

Best variable to predict price:

Among all the variables of diamonds that forms its attributes, **cut** is the best variable to predict the price of a diamond as they are directly proportional. From the bar plot it can be seen price of the diamonds increases as the quality of cut gets better. Thus, those diamonds which have the highest prices are of the best cut i.e., **ideal**. On the other hand, diamonds with worst cut i.e., **fair**, have lowest prices. Second best variable to predict the price of the diamond can be its length (x) given in mm as the price tends to increase with the length of the diamond to some extent. It is not as good as cut variable yet a considerable option after cut variable.

All other variables do not have a definite or well-defined relation with the price of diamonds. Therefore, these variables cannot be accounted for being the best variables for predicting the price of diamonds.

Question no.3:

Question no.4:

Part (a):

Total 229 countries are included in the given study of Covid-19.

Part (b):

```
[1] "Covid cases were found in the following countires of China:"
> china_provinces
 [1] "Anhui"          "Beijing"       "Chongqing"      "Fujian"
 [5] "Guangdong"      "Guangxi"       "Guizhou"        "Hainan"
 [9] "Hebei"          "Henan"         "Hubei"          "Hunan"
[13] "Jiangsu"        "Jiangxi"       "Liaoning"       "Ningxia"
[17] "Shandong"       "Shanghai"      "Shanxi"         "Sichuan"
[21] "Tianjin"        "Yunnan"        "Zhejiang"       "Gansu"
[25] "Heilongjiang"   "Jilin"         "Shaanxi"        "Xinjiang"
[29] "Inner Mongolia" "Qinghai"       "Tibet"
> length(china_provinces)
[1] 31
```
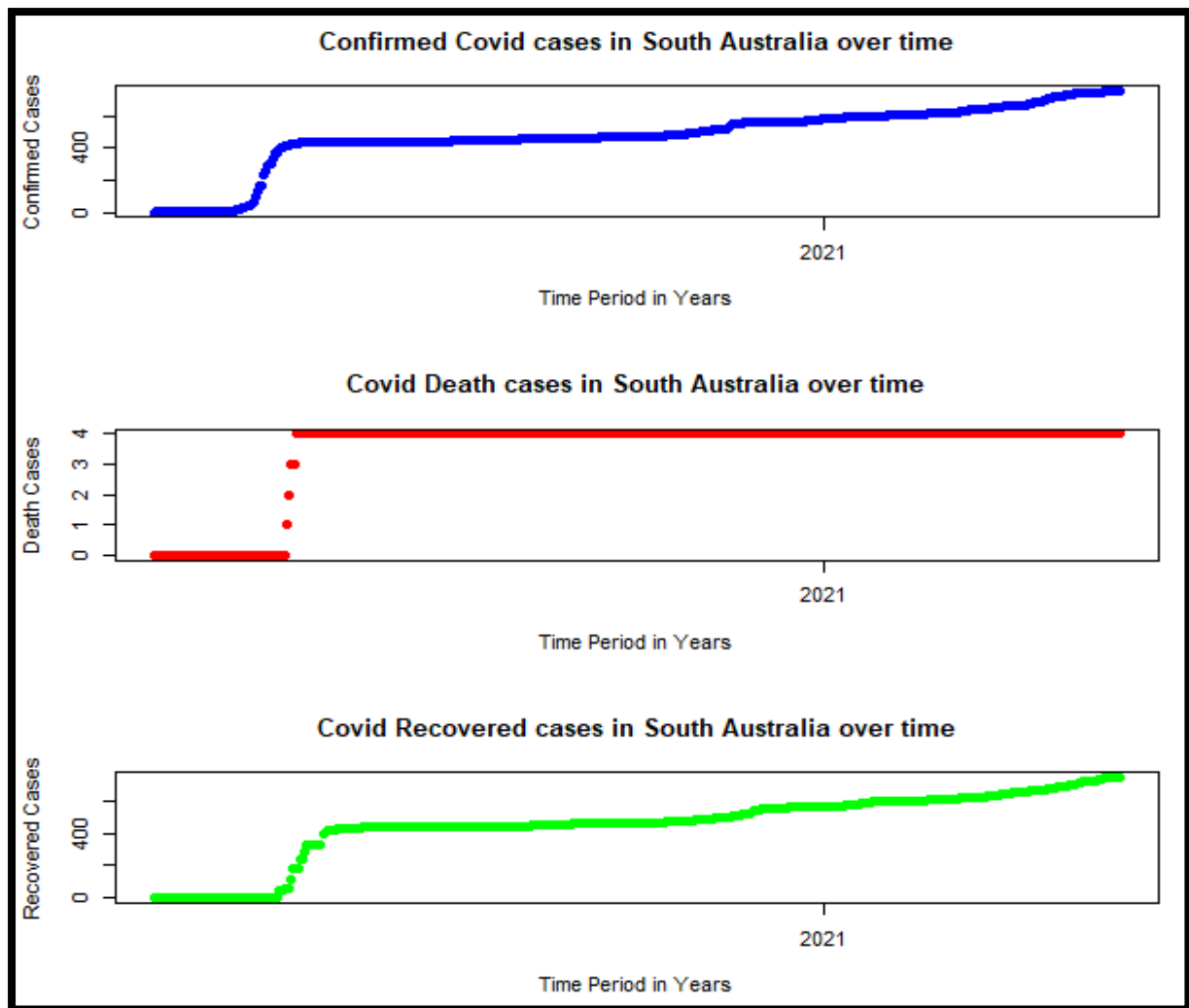
<u>Cumulative Death cases in *'Mainland China'*:</u>   **2089872**

Part (c):

```
> #--------------------------------PART (c)-----------------------------------#
> summary(covid_data)
     SNo           ObservationDate     Province.State      Country.Region
 Min.   :     1    Length:306429       Length:306429       Length:306429
 1st Qu.: 76608    Class :character    Class :character    Class :character
 Median :153215    Mode  :character    Mode  :character    Mode  :character
 Mean   :153215
 3rd Qu.:229822
 Max.   :306429
 Last.Update           Confirmed           Deaths             Recovered
 Length:306429     Min.   :-302844     Min.   :  -178     Min.   :-854405
 Class :character  1st Qu.:   1042     1st Qu.:    13     1st Qu.:     11
 Mode  :character  Median :  10375     Median :   192     Median :   1751
                   Mean   :  85671     Mean   :  2036     Mean   :  50420
                   3rd Qu.:  50752     3rd Qu.:  1322     3rd Qu.:  20270
                   Max.   :5863138     Max.   :112385     Max.   :6399531
> |
```

Part (d):

Confirmed Covid cases in South Australia over time

Covid Death cases in South Australia over time

Covid Recovered cases in South Australia over time

Part (e):

File is inside the zip folder.