covid19

Name: Rehana Bensha M

2024-05-05

```
#import libraries
options(repos = c(CRAN = "https://cran.r-project.org"))
install.packages("ggplot2")

library(ggplot2)

covid<-read.csv("C:/Users/abi00/Downloads/dataset/country_wise_latest.csv")
head(covid)

#structure
str(covid)
```

```
## 'data.frame':    187 obs. of  15 variables:
##  $ Country.Region        : chr  "Afghanistan" "Albania" "Algeria"
"Andorra" ...
##  $ Confirmed             : int  36263 4880 27973 907 950 86 167416 37390
15303 20558 ...
##  $ Deaths                : int  1269 144 1163 52 41 3 3059 711 167 713 ...
##  $ Recovered             : int  25198 2745 18837 803 242 65 72575 26665
9311 18246 ...
##  $ Active                : int  9796 1991 7973 52 667 18 91782 10014 5825
1599 ...
##  $ New.cases             : int  106 117 616 10 18 4 4890 73 368 86 ...
##  $ New.deaths            : int  10 6 8 0 1 0 120 6 6 1 ...
##  $ New.recovered         : int  18 63 749 0 0 5 2057 187 137 37 ...
##  $ Deaths...100.Cases    : num  3.5 2.95 4.16 5.73 4.32 3.49 1.83 1.9 1.09
3.47 ...
##  $ Recovered...100.Cases : num  69.5 56.2 67.3 88.5 25.5 ...
##  $ Deaths...100.Recovered: num  5.04 5.25 6.17 6.48 16.94 ...
##  $ Confirmed.last.week   : int  35526 4171 23691 884 749 76 130774 34981
12428 19743 ...
##  $ X1.week.change        : int  737 709 4282 23 201 10 36642 2409 2875 815
...
##  $ X1.week...increase    : num  2.07 17 18.07 2.6 26.84 ...
##  $ WHO.Region            : chr  "Eastern Mediterranean" "Europe" "Africa"
"Europe" ...
```

```
summary(covid)
```

```
##  Country.Region       Confirmed           Deaths           Recovered
##  Length:187         Min.   :     10   Min.   :     0.0   Min.   :      0.0
##  Class :character   1st Qu.:   1114   1st Qu.:    18.5   1st Qu.:    626.5
##  Mode  :character   Median :   5059   Median :   108.0   Median :   2815.0
```

```
##                    Mean   :  88131   Mean   :  3497.5   Mean   :  50631.5
##                    3rd Qu.:  40460   3rd Qu.:   734.0   3rd Qu.:  22606.0
##                    Max.   :4290259   Max.   :148011.0   Max.   :1846641.0
##      Active            New.cases         New.deaths       New.recovered
##  Min.   :      0.0   Min.   :    0.0   Min.   :   0.00   Min.   :    0.0
##  1st Qu.:    141.5   1st Qu.:    4.0   1st Qu.:   0.00   1st Qu.:    0.0
##  Median :   1600.0   Median :   49.0   Median :   1.00   Median :   22.0
##  Mean   :  34001.9   Mean   : 1223.0   Mean   :  28.96   Mean   :  933.8
##  3rd Qu.:   9149.0   3rd Qu.:  419.5   3rd Qu.:   6.00   3rd Qu.:  221.0
##  Max.   :2816444.0   Max.   :56336.0   Max.   :1076.00   Max.   :33728.0
##  Deaths...100.Cases Recovered...100.Cases Deaths...100.Recovered
##  Min.   : 0.000     Min.   :  0.00        Min.   :0.00
##  1st Qu.: 0.945     1st Qu.: 48.77        1st Qu.:1.45
##  Median : 2.150     Median : 71.32        Median :3.62
##  Mean   : 3.020     Mean   : 64.82        Mean   : Inf
##  3rd Qu.: 3.875     3rd Qu.: 86.89        3rd Qu.:6.44
##  Max.   :28.560     Max.   :100.00        Max.   : Inf
##  Confirmed.last.week X1.week.change   X1.week...increase  WHO.Region
##  Min.   :     10     Min.   :   -47   Min.   : -3.840    Length:187
##  1st Qu.:   1052     1st Qu.:    49   1st Qu.:  2.775    Class :character
##  Median :   5020     Median :   432   Median :  6.890    Mode  :character
##  Mean   :  78682     Mean   :  9448   Mean   : 13.606
##  3rd Qu.:  37080     3rd Qu.:  3172   3rd Qu.: 16.855
##  Max.   :3834677     Max.   :455582   Max.   :226.320
```

```
nrow(covid)
```

```
## [1] 187
```

```
ncol(covid)
```

```
## [1] 15
```

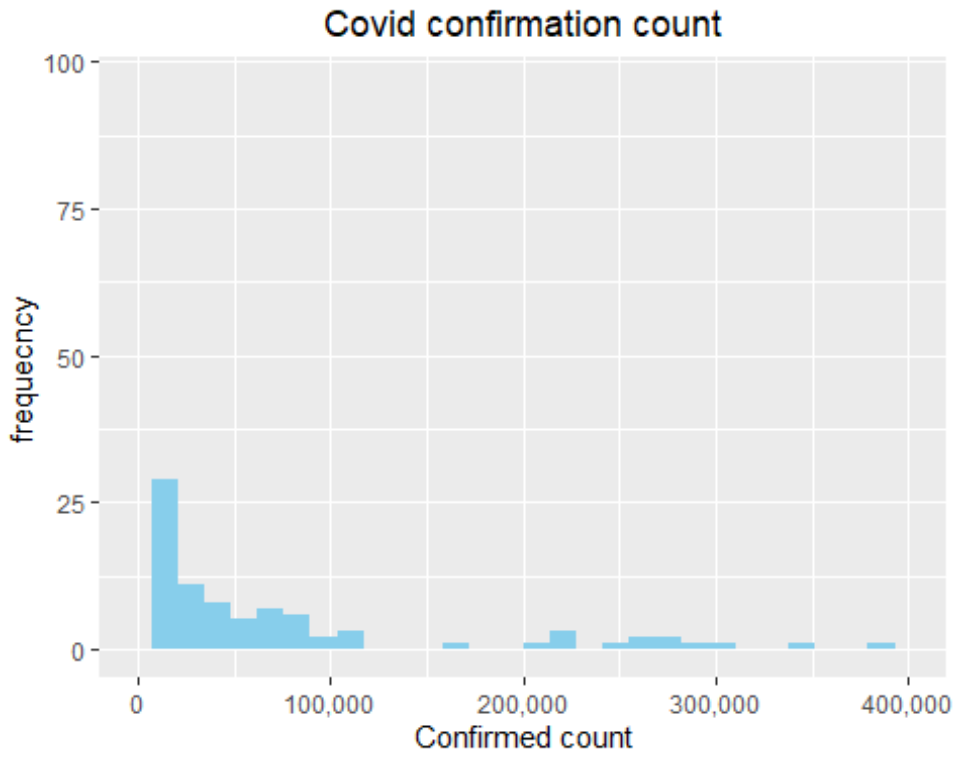
```
#null values
sum(is.na(covid))
```

```
## [1] 0
```

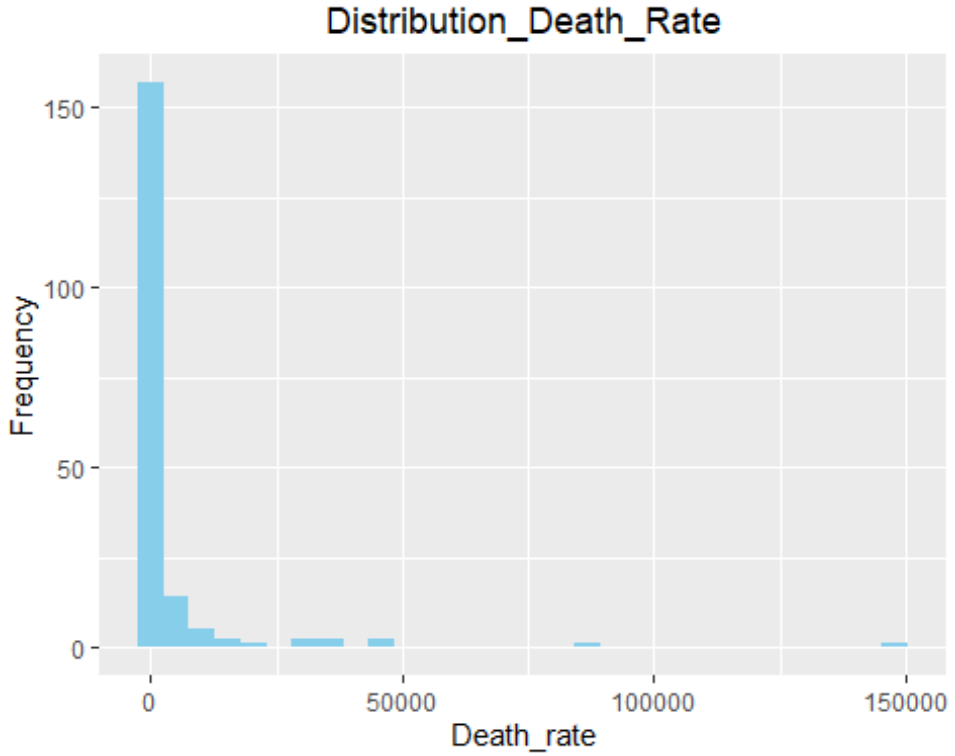
```
factor(covid$Country.Region)
```

```
## 187 Levels: Afghanistan Albania Algeria Andorra Angola ... Zimbabwe
```

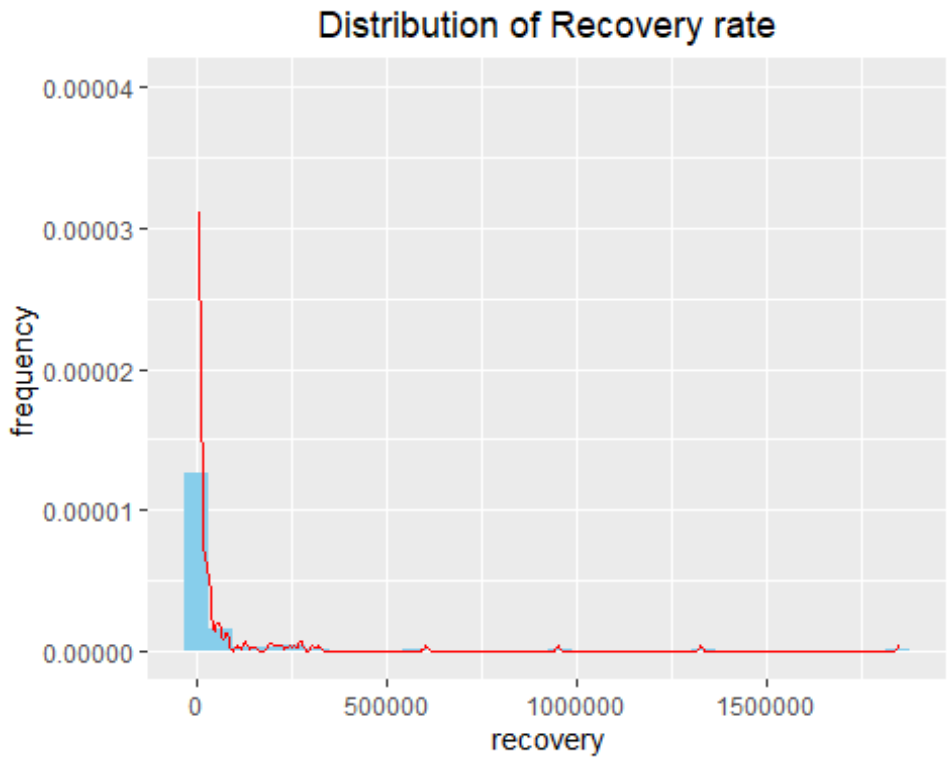
```
#univariate analysis
ggplot(covid,aes(covid$Confirmed))+geom_histogram(fill='skyblue')+labs(title=
"Covid confirmation count")+theme(plot.title=element_text(hjust =
0.5))+xlab('Confirmed count')+ylab('frequecncy') +scale_x_continuous(labels
=scales::comma_format(),limit=c(0,400000))
```



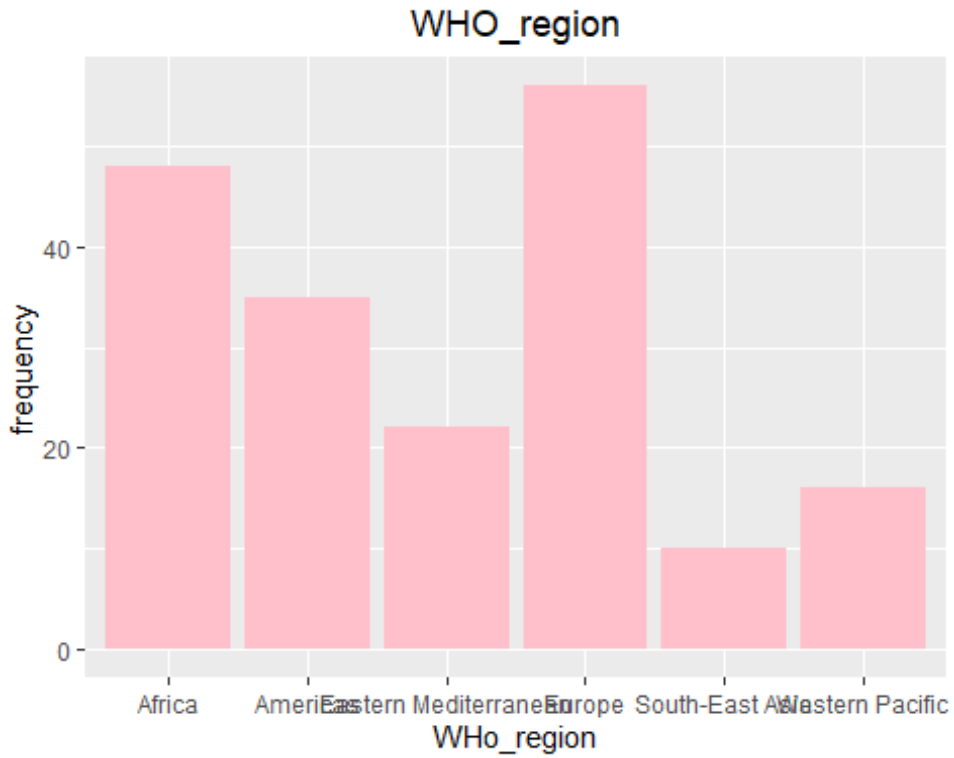

```
ggplot(covid,aes(covid$Deaths))+geom_histogram(fill='skyblue')+labs(title='Distribution_Death_Rate')+xlab('Death_rate')+ylab('Frequency')+theme(plot.title = element_text(hjust = 0.5))
```



```
ggplot(covid,aes(covid$Recovered))+geom_histogram(aes(y=..density..),fill='sk
yblue')+xlab('recovery')+ylab('frequency')+labs(title="Distribution of
Recovery rate" )+theme(plot.title =
element_text(hjust=0.5))+geom_density(col='red')+scale_y_continuous(labels =
scales::comma_format(),limits = c(0,0.00004))
```




```
ggplot(covid,aes(x=factor(covid$WHO.Region),fill=covid$Country.Region))+geom_
bar(fill='pink')+xlab('WHo_region')+labs(title='WHO_region')+ylab('frequency'
)+theme(plot.title = element_text(hjust=0.5))
```


WHO_region
frequency
40
20
0
Africa
Americas
Eastern Mediterranean
Europe
South-East Asia
Western Pacific
WHo_region