Assignment 1

Mallika

8/4/2021

Introduction

This analysis uses a Kaggle dataset - https://www.kaggle.com/anandhuh/latest-covid19-india-statewise-data. This dataset contains latest Covid-19 India state-wise data as on August 03, 2021. This dataset can be used to analyze Covid condition in India.

Load Dataset

```
CovidCases = read.csv("D:/Business Analysis/Semester 2/INTRO TO ANALYTICS/Assignment/Latest Covid-19 In
CovidCases
```

```
##                                 State.UTs Total.Cases  Active Discharged
## 1                         Andaman and Nicobar        7525      14       7382
## 2                              Andhra Pradesh     1952513   22358    1916914
## 3                           Arunachal Pradesh       45703    4465      41025
## 4                                       Assam      557437   15726     536597
## 5                                       Bihar      724390     530     714223
## 6                                  Chandigarh       61922      32      61081
## 7                                Chhattisgarh     1001037    2789     984737
## 8  Dadra and Nagar Haveli and Daman and Diu       10637      48      10585
## 9                                       Delhi     1435844     587    1410216
## 10                                        Goa      170416    1238     166052
## 11                                    Gujarat      824683     342     814265
## 12                                    Haryana      769717     740     759360
## 13                             Himachal Pradesh      205017     931     200573
## 14                             Jammu and Kashmir     320491    1319     314798
## 15                                   Jharkhand      346918     276     341518
## 16                                   Karnataka     2893556   23928    2833276
## 17                                      Kerala     3254064 1386262    3099469
## 18                                      Ladakh       20296      68      20021
## 19                                 Lakshadweep       10127      81       9997
## 20                               Madhya Pradesh      791738     149     781077
## 21                                  Maharashtra     6258079   96833    6029817
## 22                                      Manipur       91460   10251      79744
## 23                                    Meghalaya       60597    4708      54875
## 24                                      Mizoram       31145    7772      23238
## 25                                     Nagaland       27240    1312      25390
## 26                                       Odisha      965715   17201     943069
## 27                                   Puducherry      120227     922     117518
```

```
## 28                                  Punjab      598741     771     581712
## 29                                Rajasthan      953522     328     944242
## 30                                   Sikkim       24823    2849      21644
## 31                                Tamil Nadu     2546689   24025    2488775
## 32                                 Telengana      640659    9625     627254
## 33                                   Tripura       76315    3934      71642
## 34                              Uttar Pradesh     1708152     932    1684471
## 35                                Uttarakhand      341673     611     333703
## 36                                West Bengal     1522833   11891    1492878
##    Deaths Active.Ratio.... Discharge.Ratio.... Death.Ratio....
## 1     129             0.19                98.10            1.71
## 2   13241             1.15                98.18            0.68
## 3     213             9.77                89.76            0.47
## 4    5114             2.82                96.26            0.92
## 5    9637             0.07                98.60            1.33
## 6     809             0.05                98.64            1.31
## 7   13511             0.28                98.37            1.35
## 8       4             0.45                99.51            0.04
## 9   25041             0.04                98.22            1.74
## 10   3126             0.73                97.44            1.83
## 11  10076             0.04                98.74            1.22
## 12   9617             0.10                98.65            1.25
## 13   3513             0.45                97.83            1.71
## 14   4374             0.41                98.22            1.36
## 15   5124             0.08                98.44            1.48
## 16  36352             0.83                97.92            1.26
## 17  15969             4.26                95.25            0.49
## 18    207             0.34                98.65            1.02
## 19     49             0.80                98.72            0.48
## 20  10512             0.02                98.65            1.33
## 21 131429             1.55                96.35            2.10
## 22   1465            11.21                87.19            1.60
## 23   1014             7.77                90.56            1.67
## 24    135            24.95                74.61            0.43
## 25    538             4.82                93.21            1.98
## 26   5445             1.78                97.66            0.56
## 27   1787             0.77                97.75            1.49
## 28  16258             0.13                97.16            2.72
## 29   8952             0.03                99.03            0.94
## 30    330            11.48                87.19            1.33
## 31  33889             0.94                97.73            1.33
## 32   3780             1.50                97.91            0.59
## 33    739             5.15                93.88            0.97
## 34  22749             0.05                98.61            1.33
## 35   7359             0.18                97.67            2.15
## 36  18064             0.78                98.03            1.19
```

Structure of Dataset

```
str(CovidCases)
```

```
## 'data.frame':    36 obs. of  8 variables:
```

```
##  $ State.UTs          : chr  "Andaman and Nicobar" "Andhra Pradesh" "Arunachal Pradesh" "Assam" ...
##  $ Total.Cases        : int  7525 1952513 45703 557437 724390 61922 1001037 10637 1435844 170416 ...
##  $ Active             : int  14 22358 4465 15726 530 32 2789 48 587 1238 ...
##  $ Discharged         : int  7382 1916914 41025 536597 714223 61081 984737 10585 1410216 166052 ...
##  $ Deaths             : int  129 13241 213 5114 9637 809 13511 4 25041 3126 ...
##  $ Active.Ratio....   : num  0.19 1.15 9.77 2.82 0.07 0.05 0.28 0.45 0.04 0.73 ...
##  $ Discharge.Ratio....: num  98.1 98.2 89.8 96.3 98.6 ...
##  $ Death.Ratio....    : num  1.71 0.68 0.47 0.92 1.33 1.31 1.35 0.04 1.74 1.83 ...
```

List of Variables

```
names(CovidCases)
```

```
## [1] "State.UTs"           "Total.Cases"         "Active"
## [4] "Discharged"          "Deaths"              "Active.Ratio...."
## [7] "Discharge.Ratio...." "Death.Ratio...."
```

Print the top 15 rows of dataset

```
head(CovidCases , 15)
```

```
##                                  State.UTs Total.Cases Active Discharged
## 1                       Andaman and Nicobar        7525     14       7382
## 2                           Andhra Pradesh     1952513  22358    1916914
## 3                        Arunachal Pradesh       45703   4465      41025
## 4                                    Assam      557437  15726     536597
## 5                                    Bihar      724390    530     714223
## 6                                Chandigarh       61922     32      61081
## 7                             Chhattisgarh     1001037   2789     984737
## 8  Dadra and Nagar Haveli and Daman and Diu       10637     48      10585
## 9                                    Delhi     1435844    587    1410216
## 10                                     Goa      170416   1238     166052
## 11                                 Gujarat      824683    342     814265
## 12                                 Haryana      769717    740     759360
## 13                          Himachal Pradesh      205017    931     200573
## 14                         Jammu and Kashmir      320491   1319     314798
## 15                                Jharkhand      346918    276     341518
##    Deaths Active.Ratio.... Discharge.Ratio.... Death.Ratio....
## 1     129             0.19               98.10            1.71
## 2   13241             1.15               98.18            0.68
## 3     213             9.77               89.76            0.47
## 4    5114             2.82               96.26            0.92
## 5    9637             0.07               98.60            1.33
## 6     809             0.05               98.64            1.31
## 7   13511             0.28               98.37            1.35
## 8       4             0.45               99.51            0.04
## 9   25041             0.04               98.22            1.74
## 10   3126             0.73               97.44            1.83
## 11  10076             0.04               98.74            1.22
```

```
## 12   9617             0.10              98.65           1.25
## 13   3513             0.45              97.83           1.71
## 14   4374             0.41              98.22           1.36
## 15   5124             0.08              98.44           1.48
```

Function using any of the variables from the data set

```
TotalActiveCases <- function() { sum(CovidCases$Active)}
 TotalActiveCases()
```

```
## [1] 1655848
```

Data Manipulation Techniques - Filter rows with Active >20000

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
CovidCasesFiltered <- filter (CovidCases, CovidCases$Active > 20000)
CovidCasesFiltered
```

```
##        State.UTs Total.Cases  Active Discharged Deaths Active.Ratio....
## 1 Andhra Pradesh     1952513   22358    1916914  13241             1.15
## 2      Karnataka     2893556   23928    2833276  36352             0.83
## 3         Kerala     3254064 1386262    3099469  15969             4.26
## 4    Maharashtra     6258079   96833    6029817 131429             1.55
## 5     Tamil Nadu     2546689   24025    2488775  33889             0.94
##   Discharge.Ratio.... Death.Ratio....
## 1               98.18            0.68
## 2               97.92            1.26
## 3               95.25            0.49
## 4               96.35            2.10
## 5               97.73            1.33
```

Create new dataset by joining dependent and independent variables

```
CovidCasesnew = cbind(CovidCases$State.UTs, CovidCases$Total.Cases, CovidCases$Active)
CovidCasesnew
```

```
##       [,1]                                          [,2]      [,3]
##  [1,] "Andaman and Nicobar"                         "7525"    "14"
##  [2,] "Andhra Pradesh"                              "1952513" "22358"
##  [3,] "Arunachal Pradesh"                           "45703"   "4465"
##  [4,] "Assam"                                       "557437"  "15726"
##  [5,] "Bihar"                                       "724390"  "530"
##  [6,] "Chandigarh"                                  "61922"   "32"
##  [7,] "Chhattisgarh"                                "1001037" "2789"
##  [8,] "Dadra and Nagar Haveli and Daman and Diu"    "10637"   "48"
##  [9,] "Delhi"                                       "1435844" "587"
## [10,] "Goa"                                         "170416"  "1238"
## [11,] "Gujarat"                                     "824683"  "342"
## [12,] "Haryana"                                     "769717"  "740"
## [13,] "Himachal Pradesh"                            "205017"  "931"
## [14,] "Jammu and Kashmir"                           "320491"  "1319"
## [15,] "Jharkhand"                                   "346918"  "276"
## [16,] "Karnataka"                                   "2893556" "23928"
## [17,] "Kerala"                                      "3254064" "1386262"
## [18,] "Ladakh"                                      "20296"   "68"
## [19,] "Lakshadweep"                                 "10127"   "81"
## [20,] "Madhya Pradesh"                              "791738"  "149"
## [21,] "Maharashtra"                                 "6258079" "96833"
## [22,] "Manipur"                                     "91460"   "10251"
## [23,] "Meghalaya"                                   "60597"   "4708"
## [24,] "Mizoram"                                     "31145"   "7772"
## [25,] "Nagaland"                                    "27240"   "1312"
## [26,] "Odisha"                                      "965715"  "17201"
## [27,] "Puducherry"                                  "120227"  "922"
## [28,] "Punjab"                                      "598741"  "771"
## [29,] "Rajasthan"                                   "953522"  "328"
## [30,] "Sikkim"                                      "24823"   "2849"
## [31,] "Tamil Nadu"                                  "2546689" "24025"
## [32,] "Telengana"                                   "640659"  "9625"
## [33,] "Tripura"                                     "76315"   "3934"
## [34,] "Uttar Pradesh"                               "1708152" "932"
## [35,] "Uttarakhand"                                 "341673"  "611"
## [36,] "West Bengal"                                 "1522833" "11891"
```

Convert dataset from "matrix""array" to "data.frame" after cbind()

```
class(CovidCasesnew)
```

```
## [1] "matrix" "array"
```

```
CovidCasesnew = as.data.frame(CovidCasesnew)
class(CovidCasesnew)
```

```
## [1] "data.frame"
```

Rename column names V1, V2, V3 in CovidCasesnew dataset

```
names(CovidCasesnew)[1] = "State.UTs"
names(CovidCasesnew)[2] = "Total.Cases"
names(CovidCasesnew)[3] = "Active"
CovidCasesnew
```

```
##                                  State.UTs Total.Cases  Active
## 1                       Andaman and Nicobar        7525      14
## 2                            Andhra Pradesh     1952513   22358
## 3                         Arunachal Pradesh       45703    4465
## 4                                     Assam      557437   15726
## 5                                     Bihar      724390     530
## 6                                Chandigarh       61922      32
## 7                              Chhattisgarh     1001037    2789
## 8  Dadra and Nagar Haveli and Daman and Diu       10637      48
## 9                                     Delhi     1435844     587
## 10                                      Goa      170416    1238
## 11                                  Gujarat      824683     342
## 12                                  Haryana      769717     740
## 13                           Himachal Pradesh      205017     931
## 14                          Jammu and Kashmir      320491    1319
## 15                                  Jharkhand      346918     276
## 16                                  Karnataka     2893556   23928
## 17                                     Kerala     3254064 1386262
## 18                                     Ladakh       20296      68
## 19                                Lakshadweep       10127      81
## 20                              Madhya Pradesh      791738     149
## 21                                 Maharashtra     6258079   96833
## 22                                     Manipur       91460   10251
## 23                                   Meghalaya       60597    4708
## 24                                     Mizoram       31145    7772
## 25                                    Nagaland       27240    1312
## 26                                      Odisha      965715   17201
## 27                                  Puducherry      120227     922
## 28                                      Punjab      598741     771
## 29                                   Rajasthan      953522     328
## 30                                      Sikkim       24823    2849
## 31                                  Tamil Nadu     2546689   24025
## 32                                   Telengana      640659    9625
## 33                                     Tripura       76315    3934
## 34                               Uttar Pradesh     1708152     932
## 35                                 Uttarakhand      341673     611
## 36                                 West Bengal     1522833   11891
```

Remove missing values in CovidCases dataset

```
CovidCases <- na.omit(CovidCases)
CovidCases
```

```
##                                  State.UTs Total.Cases  Active Discharged
```

```
##                                  State.Name Confirmed Recovered  Deceased
## 1                       Andaman and Nicobar      7525        14      7382
## 2                            Andhra Pradesh   1952513     22358   1916914
## 3                         Arunachal Pradesh     45703      4465     41025
## 4                                     Assam    557437     15726    536597
## 5                                     Bihar    724390       530    714223
## 6                                Chandigarh     61922        32     61081
## 7                              Chhattisgarh   1001037      2789    984737
## 8  Dadra and Nagar Haveli and Daman and Diu     10637        48     10585
## 9                                     Delhi   1435844       587   1410216
## 10                                      Goa    170416      1238    166052
## 11                                  Gujarat    824683       342    814265
## 12                                  Haryana    769717       740    759360
## 13                           Himachal Pradesh    205017       931    200573
## 14                           Jammu and Kashmir    320491      1319    314798
## 15                                Jharkhand    346918       276    341518
## 16                                Karnataka   2893556     23928   2833276
## 17                                   Kerala   3254064   1386262   3099469
## 18                                   Ladakh     20296        68     20021
## 19                                Lakshadweep     10127        81      9997
## 20                             Madhya Pradesh    791738       149    781077
## 21                                Maharashtra   6258079     96833   6029817
## 22                                   Manipur     91460     10251     79744
## 23                                 Meghalaya     60597      4708     54875
## 24                                   Mizoram     31145      7772     23238
## 25                                  Nagaland     27240      1312     25390
## 26                                    Odisha    965715     17201    943069
## 27                                Puducherry    120227       922    117518
## 28                                    Punjab    598741       771    581712
## 29                                 Rajasthan    953522       328    944242
## 30                                    Sikkim     24823      2849     21644
## 31                                Tamil Nadu   2546689     24025   2488775
## 32                                 Telengana    640659      9625    627254
## 33                                   Tripura     76315      3934     71642
## 34                             Uttar Pradesh   1708152       932   1684471
## 35                               Uttarakhand    341673       611    333703
## 36                               West Bengal   1522833     11891   1492878
##    Deaths Active.Ratio.... Discharge.Ratio.... Death.Ratio....
## 1     129             0.19                98.10           1.71
## 2   13241             1.15                98.18           0.68
## 3     213             9.77                89.76           0.47
## 4    5114             2.82                96.26           0.92
## 5    9637             0.07                98.60           1.33
## 6     809             0.05                98.64           1.31
## 7   13511             0.28                98.37           1.35
## 8       4             0.45                99.51           0.04
## 9   25041             0.04                98.22           1.74
## 10   3126             0.73                97.44           1.83
## 11  10076             0.04                98.74           1.22
## 12   9617             0.10                98.65           1.25
## 13   3513             0.45                97.83           1.71
## 14   4374             0.41                98.22           1.36
## 15   5124             0.08                98.44           1.48
## 16  36352             0.83                97.92           1.26
## 17  15969             4.26                95.25           0.49
```

```
## 18    207            0.34              98.65           1.02
## 19     49            0.80              98.72           0.48
## 20  10512            0.02              98.65           1.33
## 21 131429            1.55              96.35           2.10
## 22   1465           11.21              87.19           1.60
## 23   1014            7.77              90.56           1.67
## 24    135           24.95              74.61           0.43
## 25    538            4.82              93.21           1.98
## 26   5445            1.78              97.66           0.56
## 27   1787            0.77              97.75           1.49
## 28  16258            0.13              97.16           2.72
## 29   8952            0.03              99.03           0.94
## 30    330           11.48              87.19           1.33
## 31  33889            0.94              97.73           1.33
## 32   3780            1.50              97.91           0.59
## 33    739            5.15              93.88           0.97
## 34  22749            0.05              98.61           1.33
## 35   7359            0.18              97.67           2.15
## 36  18064            0.78              98.03           1.19
```

Create duplicated data in dataset

```
B <- tail(CovidCases, 2)
B
```

```
##      State.UTs Total.Cases Active Discharged Deaths Active.Ratio....
## 35 Uttarakhand      341673    611     333703   7359             0.18
## 36 West Bengal     1522833  11891    1492878  18064             0.78
##    Discharge.Ratio.... Death.Ratio....
## 35               97.67            2.15
## 36               98.03            1.19
```

Create new dataset CovidCaseWITHDUP with duplicated data in dataset

```
CovidCasesWITHDUP = rbind(CovidCases,B)
CovidCasesWITHDUP
```

```
##                                    State.UTs Total.Cases  Active Discharged
## 1                          Andaman and Nicobar        7525      14       7382
## 2                               Andhra Pradesh     1952513   22358    1916914
## 3                            Arunachal Pradesh       45703    4465      41025
## 4                                        Assam      557437   15726     536597
## 5                                        Bihar      724390     530     714223
## 6                                   Chandigarh       61922      32      61081
## 7                                 Chhattisgarh     1001037    2789     984737
## 8   Dadra and Nagar Haveli and Daman and Diu       10637      48      10585
## 9                                        Delhi     1435844     587    1410216
## 10                                         Goa      170416    1238     166052
## 11                                     Gujarat      824683     342     814265
```

```
## 12                                Haryana      769717     740     759360
## 13                       Himachal Pradesh      205017     931     200573
## 14                       Jammu and Kashmir     320491    1319     314798
## 15                              Jharkhand      346918     276     341518
## 16                              Karnataka     2893556   23928    2833276
## 17                                 Kerala     3254064 1386262    3099469
## 18                                 Ladakh       20296      68      20021
## 19                              Lakshadweep      10127      81       9997
## 20                          Madhya Pradesh      791738     149     781077
## 21                             Maharashtra     6258079   96833    6029817
## 22                                 Manipur       91460   10251      79744
## 23                               Meghalaya       60597    4708      54875
## 24                                 Mizoram       31145    7772      23238
## 25                                Nagaland       27240    1312      25390
## 26                                  Odisha      965715   17201     943069
## 27                              Puducherry      120227     922     117518
## 28                                  Punjab      598741     771     581712
## 29                               Rajasthan      953522     328     944242
## 30                                  Sikkim       24823    2849      21644
## 31                              Tamil Nadu     2546689   24025    2488775
## 32                               Telengana      640659    9625     627254
## 33                                 Tripura       76315    3934      71642
## 34                            Uttar Pradesh     1708152     932    1684471
## 35                              Uttarakhand      341673     611     333703
## 36                              West Bengal     1522833   11891    1492878
## 351                             Uttarakhand      341673     611     333703
## 361                             West Bengal     1522833   11891    1492878
##     Deaths Active.Ratio.... Discharge.Ratio.... Death.Ratio....
## 1      129             0.19                98.10            1.71
## 2    13241             1.15                98.18            0.68
## 3      213             9.77                89.76            0.47
## 4     5114             2.82                96.26            0.92
## 5     9637             0.07                98.60            1.33
## 6      809             0.05                98.64            1.31
## 7    13511             0.28                98.37            1.35
## 8        4             0.45                99.51            0.04
## 9    25041             0.04                98.22            1.74
## 10    3126             0.73                97.44            1.83
## 11   10076             0.04                98.74            1.22
## 12    9617             0.10                98.65            1.25
## 13    3513             0.45                97.83            1.71
## 14    4374             0.41                98.22            1.36
## 15    5124             0.08                98.44            1.48
## 16   36352             0.83                97.92            1.26
## 17   15969             4.26                95.25            0.49
## 18     207             0.34                98.65            1.02
## 19      49             0.80                98.72            0.48
## 20   10512             0.02                98.65            1.33
## 21  131429             1.55                96.35            2.10
## 22    1465            11.21                87.19            1.60
## 23    1014             7.77                90.56            1.67
## 24     135            24.95                74.61            0.43
## 25     538             4.82                93.21            1.98
## 26    5445             1.78                97.66            0.56
```

```
## 27    1787            0.77             97.75           1.49
## 28   16258            0.13             97.16           2.72
## 29    8952            0.03             99.03           0.94
## 30     330           11.48             87.19           1.33
## 31   33889            0.94             97.73           1.33
## 32    3780            1.50             97.91           0.59
## 33     739            5.15             93.88           0.97
## 34   22749            0.05             98.61           1.33
## 35    7359            0.18             97.67           2.15
## 36   18064            0.78             98.03           1.19
## 351   7359            0.18             97.67           2.15
## 361  18064            0.78             98.03           1.19
```

Remove duplicated data in dataset CovidCasesWITHDUP

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.2     v stringr 1.4.0
## v tidyr   1.1.3     v forcats 0.5.1
## v readr   2.0.0

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
CovidCasesWITHDUP <- distinct(CovidCasesWITHDUP)
CovidCasesWITHDUP
```

```
##                                State.UTs Total.Cases  Active Discharged
## 1                       Andaman and Nicobar        7525      14       7382
## 2                            Andhra Pradesh     1952513   22358    1916914
## 3                         Arunachal Pradesh       45703    4465      41025
## 4                                     Assam      557437   15726     536597
## 5                                     Bihar      724390     530     714223
## 6                                Chandigarh       61922      32      61081
## 7                              Chhattisgarh     1001037    2789     984737
## 8  Dadra and Nagar Haveli and Daman and Diu       10637      48      10585
## 9                                     Delhi     1435844     587    1410216
## 10                                      Goa      170416    1238     166052
## 11                                  Gujarat      824683     342     814265
## 12                                  Haryana      769717     740     759360
## 13                           Himachal Pradesh      205017     931     200573
## 14                           Jammu and Kashmir      320491    1319     314798
## 15                                 Jharkhand      346918     276     341518
## 16                                 Karnataka     2893556   23928    2833276
## 17                                    Kerala     3254064 1386262    3099469
## 18                                    Ladakh       20296      68      20021
```

```
## 19                            Lakshadweep       10127      81       9997
## 20                         Madhya Pradesh      791738     149     781077
## 21                            Maharashtra     6258079   96833    6029817
## 22                                Manipur       91460   10251      79744
## 23                              Meghalaya       60597    4708      54875
## 24                                Mizoram       31145    7772      23238
## 25                               Nagaland       27240    1312      25390
## 26                                 Odisha      965715   17201     943069
## 27                             Puducherry      120227     922     117518
## 28                                 Punjab      598741     771     581712
## 29                              Rajasthan      953522     328     944242
## 30                                 Sikkim       24823    2849      21644
## 31                             Tamil Nadu     2546689   24025    2488775
## 32                              Telengana      640659    9625     627254
## 33                                Tripura       76315    3934      71642
## 34                           Uttar Pradesh     1708152     932    1684471
## 35                             Uttarakhand      341673     611     333703
## 36                             West Bengal     1522833   11891    1492878
##    Deaths Active.Ratio.... Discharge.Ratio.... Death.Ratio....
## 1     129             0.19                98.10            1.71
## 2   13241             1.15                98.18            0.68
## 3     213             9.77                89.76            0.47
## 4    5114             2.82                96.26            0.92
## 5    9637             0.07                98.60            1.33
## 6     809             0.05                98.64            1.31
## 7   13511             0.28                98.37            1.35
## 8       4             0.45                99.51            0.04
## 9   25041             0.04                98.22            1.74
## 10   3126             0.73                97.44            1.83
## 11  10076             0.04                98.74            1.22
## 12   9617             0.10                98.65            1.25
## 13   3513             0.45                97.83            1.71
## 14   4374             0.41                98.22            1.36
## 15   5124             0.08                98.44            1.48
## 16  36352             0.83                97.92            1.26
## 17  15969             4.26                95.25            0.49
## 18    207             0.34                98.65            1.02
## 19     49             0.80                98.72            0.48
## 20  10512             0.02                98.65            1.33
## 21 131429             1.55                96.35            2.10
## 22   1465            11.21                87.19            1.60
## 23   1014             7.77                90.56            1.67
## 24    135            24.95                74.61            0.43
## 25    538             4.82                93.21            1.98
## 26   5445             1.78                97.66            0.56
## 27   1787             0.77                97.75            1.49
## 28  16258             0.13                97.16            2.72
## 29   8952             0.03                99.03            0.94
## 30    330            11.48                87.19            1.33
## 31  33889             0.94                97.73            1.33
## 32   3780             1.50                97.91            0.59
## 33    739             5.15                93.88            0.97
## 34  22749             0.05                98.61            1.33
## 35   7359             0.18                97.67            2.15
```

```
## 36  18064             0.78              98.03           1.19
```

Reorder multiple rows in descending alphabetical order of State and Total Cases

```
CovidCases %>% arrange(desc(CovidCases$State.UTs, Total.Cases))
```

```
##                                    State.UTs Total.Cases  Active Discharged
## 1                                West Bengal     1522833   11891    1492878
## 2                                Uttarakhand      341673     611     333703
## 3                              Uttar Pradesh     1708152     932    1684471
## 4                                    Tripura       76315    3934      71642
## 5                                  Telengana      640659    9625     627254
## 6                                  Tamil Nadu     2546689   24025    2488775
## 7                                     Sikkim       24823    2849      21644
## 8                                  Rajasthan      953522     328     944242
## 9                                     Punjab      598741     771     581712
## 10                                Puducherry      120227     922     117518
## 11                                    Odisha      965715   17201     943069
## 12                                 Nagaland       27240    1312      25390
## 13                                 Mizoram        31145    7772      23238
## 14                                Meghalaya       60597    4708      54875
## 15                                  Manipur       91460   10251      79744
## 16                                Maharashtra     6258079   96833    6029817
## 17                             Madhya Pradesh      791738     149     781077
## 18                                Lakshadweep       10127      81       9997
## 19                                     Ladakh       20296      68      20021
## 20                                     Kerala     3254064 1386262    3099469
## 21                                  Karnataka     2893556   23928    2833276
## 22                                  Jharkhand      346918     276     341518
## 23                            Jammu and Kashmir      320491    1319     314798
## 24                             Himachal Pradesh      205017     931     200573
## 25                                     Haryana      769717     740     759360
## 26                                     Gujarat      824683     342     814265
## 27                                         Goa      170416    1238     166052
## 28                                       Delhi     1435844     587    1410216
## 29 Dadra and Nagar Haveli and Daman and Diu       10637      48      10585
## 30                                 Chhattisgarh     1001037    2789     984737
## 31                                   Chandigarh       61922      32      61081
## 32                                        Bihar      724390     530     714223
## 33                                        Assam      557437   15726     536597
## 34                            Arunachal Pradesh       45703    4465      41025
## 35                               Andhra Pradesh     1952513   22358    1916914
## 36                           Andaman and Nicobar        7525      14       7382
##    Deaths Active.Ratio.... Discharge.Ratio.... Death.Ratio....
## 1   18064             0.78               98.03           1.19
## 2    7359             0.18               97.67           2.15
## 3   22749             0.05               98.61           1.33
## 4     739             5.15               93.88           0.97
## 5    3780             1.50               97.91           0.59
## 6   33889             0.94               97.73           1.33
## 7     330            11.48               87.19           1.33
## 8    8952             0.03               99.03           0.94
```

```
## 9   16258             0.13              97.16            2.72
## 10   1787             0.77              97.75            1.49
## 11   5445             1.78              97.66            0.56
## 12    538             4.82              93.21            1.98
## 13    135            24.95              74.61            0.43
## 14   1014             7.77              90.56            1.67
## 15   1465            11.21              87.19            1.60
## 16 131429             1.55              96.35            2.10
## 17  10512             0.02              98.65            1.33
## 18     49             0.80              98.72            0.48
## 19    207             0.34              98.65            1.02
## 20  15969             4.26              95.25            0.49
## 21  36352             0.83              97.92            1.26
## 22   5124             0.08              98.44            1.48
## 23   4374             0.41              98.22            1.36
## 24   3513             0.45              97.83            1.71
## 25   9617             0.10              98.65            1.25
## 26  10076             0.04              98.74            1.22
## 27   3126             0.73              97.44            1.83
## 28  25041             0.04              98.22            1.74
## 29      4             0.45              99.51            0.04
## 30  13511             0.28              98.37            1.35
## 31    809             0.05              98.64            1.31
## 32   9637             0.07              98.60            1.33
## 33   5114             2.82              96.26            0.92
## 34    213             9.77              89.76            0.47
## 35  13241             1.15              98.18            0.68
## 36    129             0.19              98.10            1.71
```

Rename column name State.UTs = X in CovidCases dataset

```
CovidCases %>% rename(X=State.UTs)
```

```
##                                         X Total.Cases  Active Discharged
## 1                      Andaman and Nicobar        7525      14       7382
## 2                           Andhra Pradesh     1952513   22358    1916914
## 3                        Arunachal Pradesh       45703    4465      41025
## 4                                    Assam      557437   15726     536597
## 5                                    Bihar      724390     530     714223
## 6                                Chandigarh       61922      32      61081
## 7                             Chhattisgarh     1001037    2789     984737
## 8  Dadra and Nagar Haveli and Daman and Diu       10637      48      10585
## 9                                    Delhi     1435844     587    1410216
## 10                                     Goa      170416    1238     166052
## 11                                 Gujarat      824683     342     814265
## 12                                 Haryana      769717     740     759360
## 13                          Himachal Pradesh      205017     931     200573
## 14                          Jammu and Kashmir      320491    1319     314798
## 15                                Jharkhand      346918     276     341518
## 16                                Karnataka     2893556   23928    2833276
## 17                                   Kerala     3254064 1386262    3099469
## 18                                   Ladakh       20296      68      20021
```

```
## 19                            Lakshadweep       10127      81       9997
## 20                         Madhya Pradesh      791738     149     781077
## 21                            Maharashtra     6258079   96833    6029817
## 22                                Manipur       91460   10251      79744
## 23                              Meghalaya       60597    4708      54875
## 24                                Mizoram       31145    7772      23238
## 25                               Nagaland       27240    1312      25390
## 26                                 Odisha      965715   17201     943069
## 27                             Puducherry      120227     922     117518
## 28                                 Punjab      598741     771     581712
## 29                              Rajasthan      953522     328     944242
## 30                                 Sikkim       24823    2849      21644
## 31                             Tamil Nadu     2546689   24025    2488775
## 32                              Telengana      640659    9625     627254
## 33                                Tripura       76315    3934      71642
## 34                           Uttar Pradesh     1708152     932    1684471
## 35                             Uttarakhand      341673     611     333703
## 36                             West Bengal     1522833   11891    1492878
##    Deaths Active.Ratio.... Discharge.Ratio.... Death.Ratio....
## 1     129             0.19                98.10            1.71
## 2   13241             1.15                98.18            0.68
## 3     213             9.77                89.76            0.47
## 4    5114             2.82                96.26            0.92
## 5    9637             0.07                98.60            1.33
## 6     809             0.05                98.64            1.31
## 7   13511             0.28                98.37            1.35
## 8       4             0.45                99.51            0.04
## 9   25041             0.04                98.22            1.74
## 10   3126             0.73                97.44            1.83
## 11  10076             0.04                98.74            1.22
## 12   9617             0.10                98.65            1.25
## 13   3513             0.45                97.83            1.71
## 14   4374             0.41                98.22            1.36
## 15   5124             0.08                98.44            1.48
## 16  36352             0.83                97.92            1.26
## 17  15969             4.26                95.25            0.49
## 18    207             0.34                98.65            1.02
## 19     49             0.80                98.72            0.48
## 20  10512             0.02                98.65            1.33
## 21 131429             1.55                96.35            2.10
## 22   1465            11.21                87.19            1.60
## 23   1014             7.77                90.56            1.67
## 24    135            24.95                74.61            0.43
## 25    538             4.82                93.21            1.98
## 26   5445             1.78                97.66            0.56
## 27   1787             0.77                97.75            1.49
## 28  16258             0.13                97.16            2.72
## 29   8952             0.03                99.03            0.94
## 30    330            11.48                87.19            1.33
## 31  33889             0.94                97.73            1.33
## 32   3780             1.50                97.91            0.59
## 33    739             5.15                93.88            0.97
## 34  22749             0.05                98.61            1.33
## 35   7359             0.18                97.67            2.15
```

```
## 36  18064             0.78               98.03            1.19
```

Add new variables/column in dataset by using a mathematical function

```
CovidCases$Active_Per = (CovidCases$Active.Ratio)*100
CovidCases
```

```
##                                  State.UTs Total.Cases  Active Discharged
## 1                      Andaman and Nicobar        7525      14       7382
## 2                           Andhra Pradesh     1952513   22358    1916914
## 3                        Arunachal Pradesh       45703    4465      41025
## 4                                    Assam      557437   15726     536597
## 5                                    Bihar      724390     530     714223
## 6                                Chandigarh       61922      32      61081
## 7                              Chhattisgarh     1001037    2789     984737
## 8  Dadra and Nagar Haveli and Daman and Diu       10637      48      10585
## 9                                    Delhi     1435844     587    1410216
## 10                                     Goa      170416    1238     166052
## 11                                 Gujarat      824683     342     814265
## 12                                 Haryana      769717     740     759360
## 13                          Himachal Pradesh      205017     931     200573
## 14                          Jammu and Kashmir      320491    1319     314798
## 15                                Jharkhand      346918     276     341518
## 16                                Karnataka     2893556   23928    2833276
## 17                                   Kerala     3254064 1386262    3099469
## 18                                   Ladakh       20296      68      20021
## 19                                Lakshadweep       10127      81       9997
## 20                             Madhya Pradesh      791738     149     781077
## 21                                Maharashtra     6258079   96833    6029817
## 22                                    Manipur       91460   10251      79744
## 23                                  Meghalaya       60597    4708      54875
## 24                                    Mizoram       31145    7772      23238
## 25                                   Nagaland       27240    1312      25390
## 26                                     Odisha      965715   17201     943069
## 27                                  Puducherry      120227     922     117518
## 28                                     Punjab      598741     771     581712
## 29                                  Rajasthan      953522     328     944242
## 30                                     Sikkim       24823    2849      21644
## 31                                  Tamil Nadu     2546689   24025    2488775
## 32                                  Telengana      640659    9625     627254
## 33                                    Tripura       76315    3934      71642
## 34                               Uttar Pradesh     1708152     932    1684471
## 35                                 Uttarakhand      341673     611     333703
## 36                                 West Bengal     1522833   11891    1492878
##    Deaths Active.Ratio.... Discharge.Ratio.... Death.Ratio.... Active_Per
## 1     129             0.19               98.10            1.71         19
## 2   13241             1.15               98.18            0.68        115
## 3     213             9.77               89.76            0.47        977
## 4    5114             2.82               96.26            0.92        282
## 5    9637             0.07               98.60            1.33          7
## 6     809             0.05               98.64            1.31          5
## 7   13511             0.28               98.37            1.35         28
```

```
## 8       4             0.45             99.51            0.04         45
## 9   25041             0.04             98.22            1.74          4
## 10   3126             0.73             97.44            1.83         73
## 11  10076             0.04             98.74            1.22          4
## 12   9617             0.10             98.65            1.25         10
## 13   3513             0.45             97.83            1.71         45
## 14   4374             0.41             98.22            1.36         41
## 15   5124             0.08             98.44            1.48          8
## 16  36352             0.83             97.92            1.26         83
## 17  15969             4.26             95.25            0.49        426
## 18    207             0.34             98.65            1.02         34
## 19     49             0.80             98.72            0.48         80
## 20  10512             0.02             98.65            1.33          2
## 21 131429             1.55             96.35            2.10        155
## 22   1465            11.21             87.19            1.60       1121
## 23   1014             7.77             90.56            1.67        777
## 24    135            24.95             74.61            0.43       2495
## 25    538             4.82             93.21            1.98        482
## 26   5445             1.78             97.66            0.56        178
## 27   1787             0.77             97.75            1.49         77
## 28  16258             0.13             97.16            2.72         13
## 29   8952             0.03             99.03            0.94          3
## 30    330            11.48             87.19            1.33       1148
## 31  33889             0.94             97.73            1.33         94
## 32   3780             1.50             97.91            0.59        150
## 33    739             5.15             93.88            0.97        515
## 34  22749             0.05             98.61            1.33          5
## 35   7359             0.18             97.67            2.15         18
## 36  18064             0.78             98.03            1.19         78
```

Add new variables/column in dataset by using a mathematical function

```
CovidCases$Active_DBL = data.frame(CovidCases$Active.Ratio)*2
CovidCases$Active_DBL
```

```
##    CovidCases.Active.Ratio
## 1                     0.38
## 2                     2.30
## 3                    19.54
## 4                     5.64
## 5                     0.14
## 6                     0.10
## 7                     0.56
## 8                     0.90
## 9                     0.08
## 10                    1.46
## 11                    0.08
## 12                    0.20
## 13                    0.90
## 14                    0.82
## 15                    0.16
## 16                    1.66
```

```
## 17                   8.52
## 18                   0.68
## 19                   1.60
## 20                   0.04
## 21                   3.10
## 22                  22.42
## 23                  15.54
## 24                  49.90
## 25                   9.64
## 26                   3.56
## 27                   1.54
## 28                   0.26
## 29                   0.06
## 30                  22.96
## 31                   1.88
## 32                   3.00
## 33                  10.30
## 34                   0.10
## 35                   0.36
## 36                   1.56
```

Extract 5 random rows without replacement using random number generator engine

```
set.seed(1234)
CovidCases %>% sample_n(5, replace = FALSE)
```

```
##   State.UTs Total.Cases Active Discharged Deaths Active.Ratio....
## 1    Punjab      598741    771     581712  16258             0.13
## 2 Karnataka     2893556  23928    2833276  36352             0.83
## 3   Manipur       91460  10251      79744   1465            11.21
## 4     Delhi     1435844    587    1410216  25041             0.04
## 5     Bihar      724390    530     714223   9637             0.07
##   Discharge.Ratio.... Death.Ratio.... Active_Per CovidCases.Active.Ratio
## 1               97.16            2.72         13                    0.26
## 2               97.92            1.26         83                    1.66
## 3               87.19            1.60       1121                   22.42
## 4               98.22            1.74          4                    0.08
## 5               98.60            1.33          7                    0.14
```

Summary Statistics Dataset

```
summary(CovidCases)
```

```
##   State.UTs          Total.Cases          Active          Discharged
##  Length:36          Min.   :   7525   Min.   :     14   Min.   :   7382
##  Class :character   1st Qu.:  61591   1st Qu.:    483   1st Qu.:  59530
##  Mode  :character   Median : 452178   Median :   1275   Median : 439058
##                     Mean   : 871442   Mean   :  45996   Mean   : 848420
```

```
##                    3rd Qu.: 974546   3rd Qu.:   9782   3rd Qu.: 954366
##                    Max.   :6258079   Max.   :1386262   Max.   :6029817
##      Deaths         Active.Ratio....  Discharge.Ratio.... Death.Ratio....
##  Min.   :     4.0   Min.   : 0.0200   Min.   :74.61       Min.   :0.040
##  1st Qu.:   791.5   1st Qu.: 0.1225   1st Qu.:96.33       1st Qu.:0.935
##  Median :  5119.0   Median : 0.7500   Median :97.92       Median :1.330
##  Mean   : 11682.0   Mean   : 2.6658   Mean   :96.07       Mean   :1.260
##  3rd Qu.: 13308.5   3rd Qu.: 2.0400   3rd Qu.:98.60       3rd Qu.:1.617
##  Max.   :131429.0   Max.   :24.9500   Max.   :99.51       Max.   :2.720
##    Active_Per      Active_DBL.CovidCases.Active.Ratio
##  Min.   :   2.00   Min.   : 0.04000
##  1st Qu.:  12.25   1st Qu.: 0.24500
##  Median :  75.00   Median : 1.50000
##  Mean   : 266.58   Mean   : 5.33167
##  3rd Qu.: 204.00   3rd Qu.: 4.08000
##  Max.   :2495.00   Max.   :49.90000
```

Mean

```
mean(CovidCases$Total.Cases)
```

```
## [1] 871441.7
```

Median

```
median(CovidCases$Total.Cases)
```

```
## [1] 452177.5
```

Mode

```
mode(CovidCases$Total.Cases)
```

```
## [1] "numeric"
```

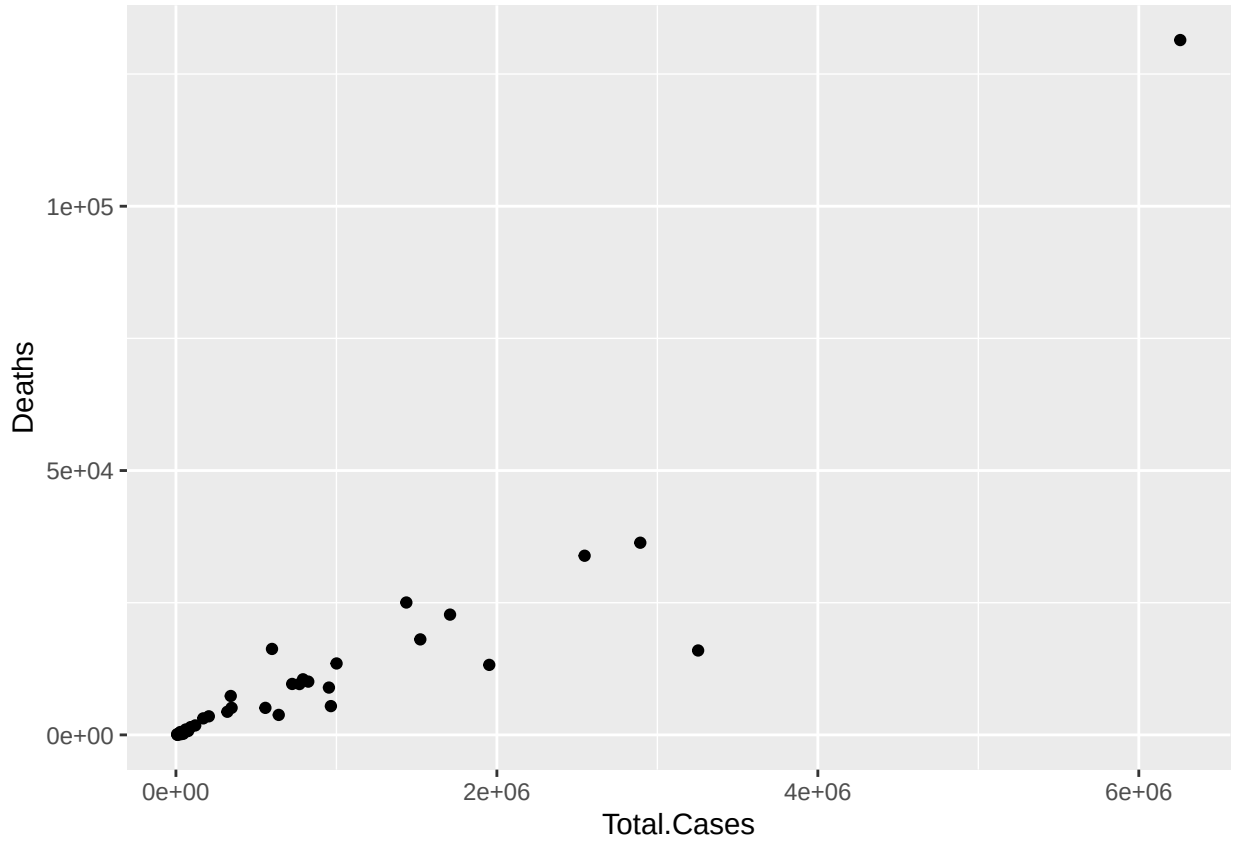
Range

```
range(CovidCases$Total.Cases)
```

```
## [1]    7525 6258079
```

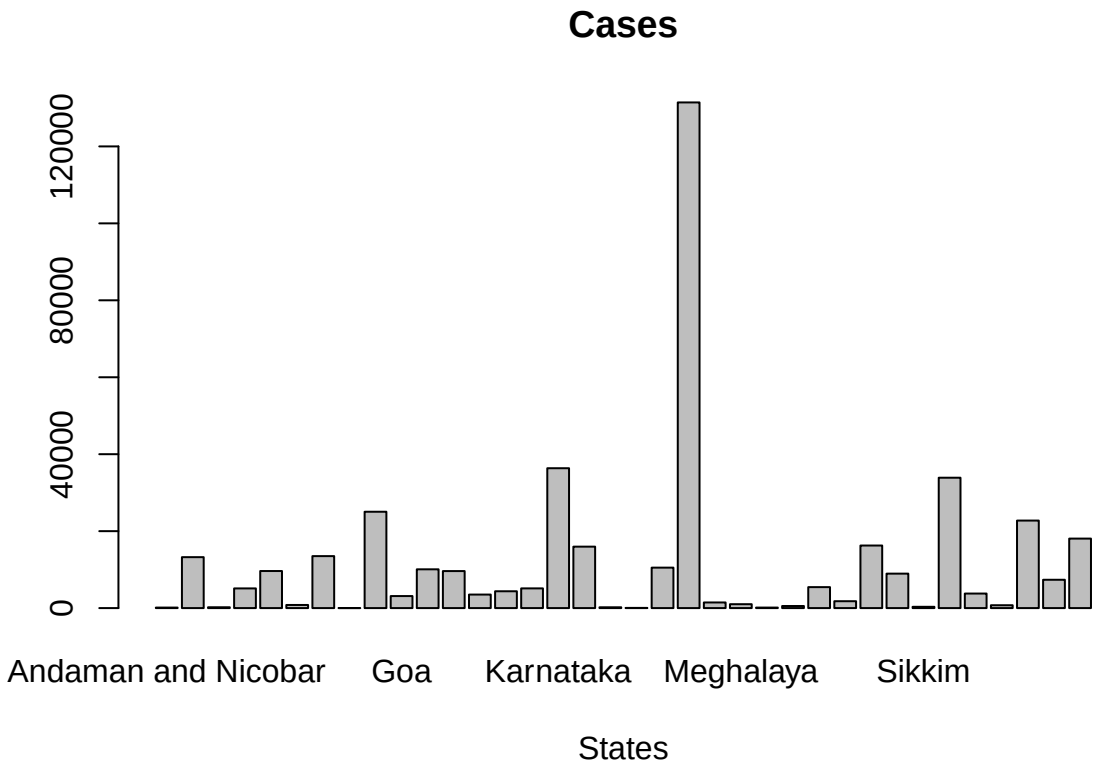
Scatter Plot for Total and Death Ratio

```
ggplot(CovidCases, aes(x=Total.Cases, y=Deaths))+geom_point()
```



Bar Plot

```
barplot(CovidCases$Deaths, main="Cases",   xlab="States",names.arg = CovidCases$State.UTs)
```



GIT Repository

https://github.com/mallikashakya/R-prog-assignment-group7