DATA ACTIVITY 3 – COVID-19 IN INDIA JAN-MAR 2020

Task 3.1 – Create binary variable called *has_deaths*, indicating whether any deaths were reported:

```
#Create binary variable has_deaths
Covid19IndiaJan20Mar20$has_deaths <- ifelse(Covid19IndiaJan20Mar20$Deaths > 0,1,0)
```

Task 3.2 – Create frequency table using table() command to show death reporting patterns:

```
#Create frequency table - daily reports of recoveries vs no recoveries
table(Covid19IndiaJan20Mar20$has_deaths)
head(Covid19IndiaJan20Mar20[, c("Deaths", "has_deaths")])
```

```
> table(Covid19IndiaJan20Mar20$has_deaths)

  0   1
245  25
```

Task 3.3 – Convert to a factor variable with appropriate labels: *("No Deaths", "Deaths Reported")*:

```
#Create factor variable with labels
Covid19IndiaJan20Mar20$has_deaths <- factor(
  Covid19IndiaJan20Mar20$has_deaths,
  levels=c(0,1),
  labels=c("No Deaths", "Deaths Reported")
)
table(Covid19IndiaJan20Mar20$has_deaths)
```

```
> table(Covid19IndiaJan20Mar20$has_deaths)

      No Deaths Deaths Reported
            245              25
```

Task 3.4 – Create a categorical variable called *case_level* based on total confirmed cases (Indian + Foreign nationals)

No cases = 0 cases; Low cases = 1-5; Medium cases = 6-15; High cases = 15+

```
#Create categorical variable case_level
  #Compute total confirmed cases
Covid19IndiaJan20Mar20$total_cases <- Covid19IndiaJan20Mar20$ConfirmedIndianNational + Covid19IndiaJan20Mar20$ConfirmedForeignNational
  #Create categorical variable
Covid19IndiaJan20Mar20$case_level <- cut(
  Covid19IndiaJan20Mar20$total_cases,
  breaks=c(-Inf,0,5,15,Inf),
  labels=c("No Cases", "Low Cases", "Medium Cases", "High Cases")
)
  #Check variables
table(Covid19IndiaJan20Mar20$case_level)
head(Covid19IndiaJan20Mar20[, c("total_cases", "case_level")])
```

```
> table(Covid19IndiaJan20Mar20$case_level)

    No Cases    Low Cases Medium Cases   High Cases
           0          177           61           32
> head(Covid19IndiaJan20Mar20[, c("total_cases", "case_level")])
  total_cases case_level
1           1  Low Cases
2           1  Low Cases
3           2  Low Cases
4           3  Low Cases
5           3  Low Cases
6           3  Low Cases
```

Task 3.5 – Create a frequency table for the State/Union Territory variable to see which states had the most frequent reporting:

```
#Create frequency table of reports per state
table(Covid19IndiaJan20Mar20$`State/UnionTerritory`)
  #Sort table in descending order
state_freq <- sort(table(Covid19IndiaJan20Mar20$`State/UnionTerritory`), decreasing=TRUE)
  #View sorted table
state_freq
```

```
> table(Covid19IndiaJan20Mar20$`State/UnionTerritory`)

                    Andhra Pradesh                         Chandigarh                       Chattisgarh
                                10                                  1                                 1
                      Chhattisgarh                            Delhi                            Gujarat
                                 2                                 20                                 2
                           Haryana                   Himachal Pradesh                  Jammu and Kashmir
                                18                                  1                                 1
                         Karnataka                             Kerala                            Ladakh
                                13                                 52                                 1
                    Madhya Pradesh                        Maharashtra                            Odisha
                                 1                                 13                                 6
                        Pondicherry                         Puducherry                            Punjab
                                 2                                  2                                13
                         Rajasthan                         Tamil Nadu                         Telengana
                                19                                 15                                20
      Union Territory of Chandigarh Union Territory of Jammu and Kashmir          Union Territory of Ladakh
                                 2                                 12                                14
                     Uttar Pradesh                        Uttarakhand                       West Bengal
                                18                                  7                                 4

>   #Sort table in descending order
> state_freq <- sort(table(Covid19IndiaJan20Mar20$`State/UnionTerritory`), decreasing=TRUE)
>   #View sorted table
> state_freq

                           Kerala                            Delhi                         Telengana
                               52                               20                                20
                        Rajasthan                          Haryana                      Uttar Pradesh
                               19                               18                                18
                        Tamil Nadu Union Territory of Ladakh                         Karnataka
                               15                               14                                13
                      Maharashtra                           Punjab Union Territory of Jammu and Kashmir
                               13                               13                                12
                   Andhra Pradesh                      Uttarakhand                            Odisha
                               10                                7                                 6
                      West Bengal                       Chhattisgarh                          Gujarat
                                4                                2                                 2
                       Pondicherry                       Puducherry       Union Territory of Chandigarh
                                2                                2                                 2
                       Chandigarh                       Chattisgarh                  Himachal Pradesh
                                1                                1                                 1
                 Jammu and Kashmir                          Ladakh                      Madhya Pradesh
                                1                                1                                 1
```

Task 3.6 – Identify the top 10 states with the highest no. daily reports:

```
#View top 10 states by daily reports
top10states <- head(state_freq, 10)
top10states
```

```
> #View top 10 states by daily reports
> top10states <- head(state_freq, 10)
> top10states

          Kerala            Delhi        Telengana        Rajasthan
              52               20               20               19
         Haryana     Uttar Pradesh       Tamil Nadu Union Territory of Ladakh
              18               18               15               14
       Karnataka      Maharashtra
              13               13
```