

CSE-320

①

CA-1

E0119035

Country Region - Nominal

Lat - Continuous

Long - Continuous

Age - Continuous

Confirmed - discrete

Deaths - discrete

Recovered - discrete

Active - discrete

2)

(a) State 1 : Age and death

Statement 2 : Gender and death

b)

Variable

Summary Statistics

Gender

Mode, frequency

Age

Mean, median, mode,  
Std. deviation, range,  
Variance, min, max,  
frequency

Death

Mean, Median, mode, freq  
std deviation, range  
Variance, min, max

3)

For missing Value

```
data = read.csv("covid.csv")
```

# For numeric data

```
data$age = ifelse(is.na(data$age),  
  ave(data$age, FUN = function(x)  
    mean(x, na.rm = TRUE)), data$age)
```

# For categorical data - encoding

```
data$country = Factor(data$country,  
  levels = c("India", "USA"),  
  labels = c(1, 2))
```