

Name :- Aman deep Rawat

Roll No :- 21711063 / 36

Sec :- 6

Ans 2 :- we are using here Titanic dataset to analyze.
load data.

=> titanic <- read.csv("C:/Users/Desktop/titanic.csv",
header = TRUE, sep = ",")

Peek at your data

=> view(titanic)

This helps in for familiarising with the data set

=> tail(titanic, 10)

return Bottom, 10, rows

=> names(titanic)

This helps us in checking out the variables
in the data set.

=> summary(titanic)

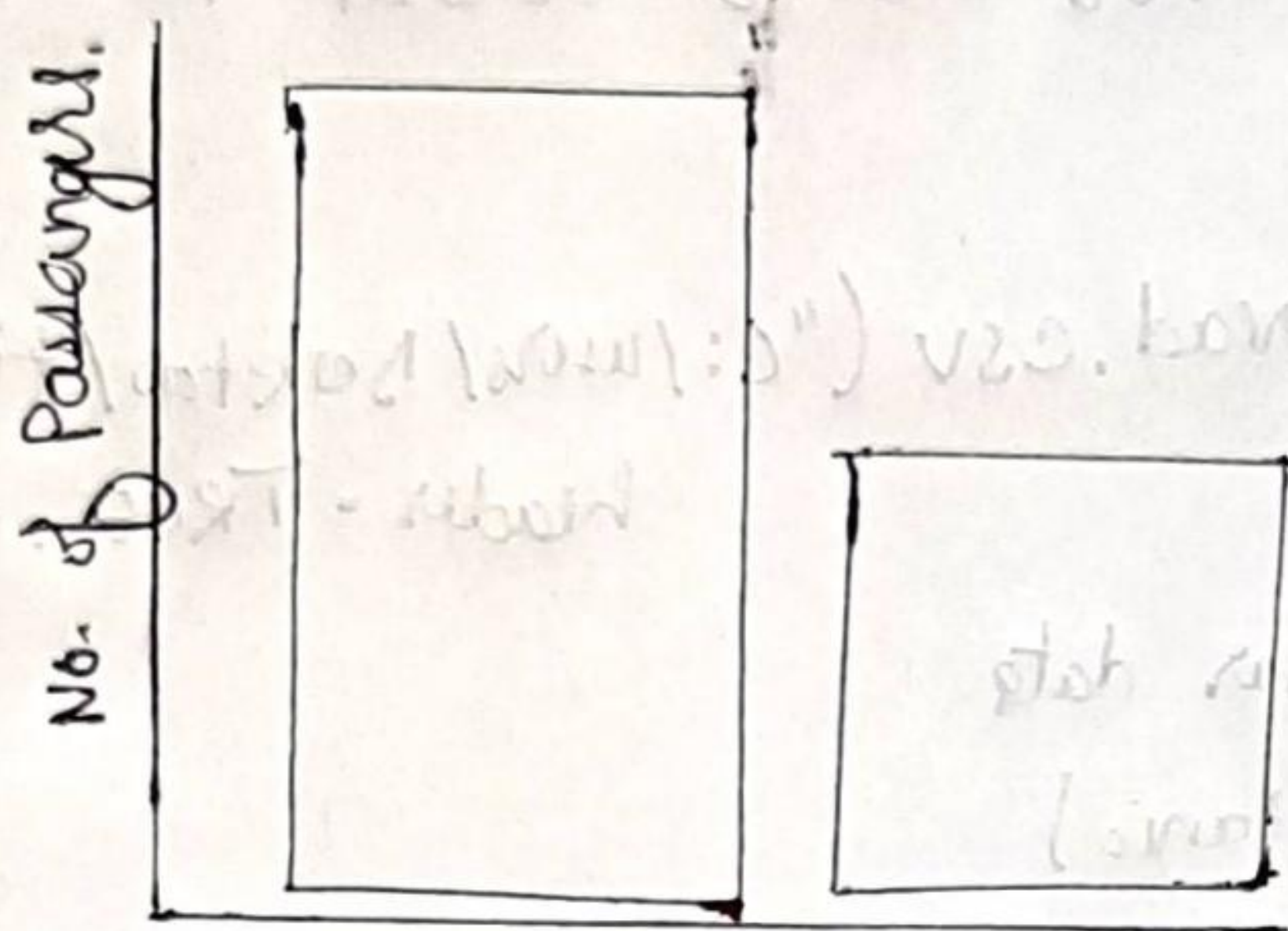
It is one of the most important function
that helps in summarising each attribute
in the dataset.

It gives the descriptive statistics of true
data.

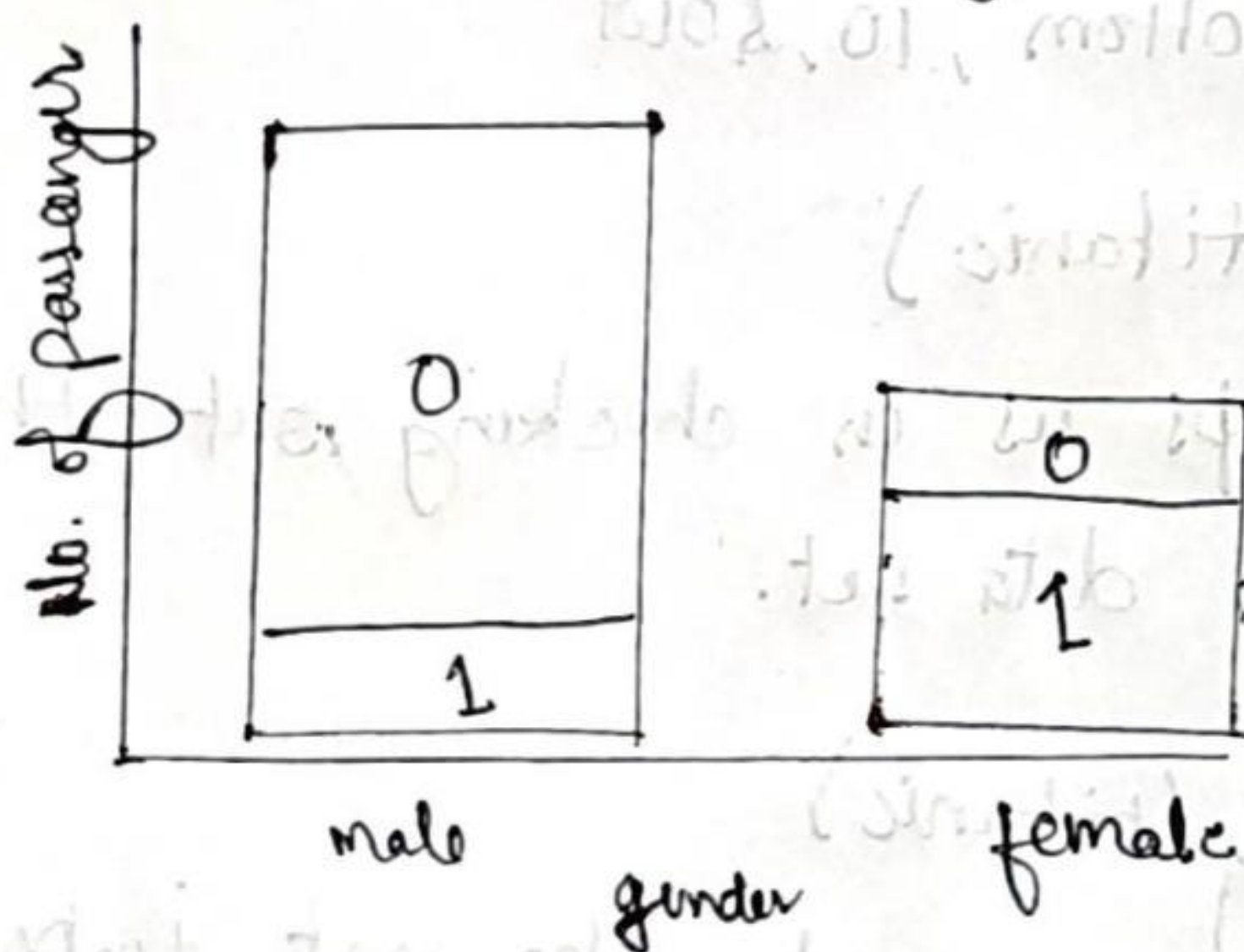
Analysis & visualization

• Survival rate

```
ggplot(titanic, aes(x=Survival)) + geom_bar()
```



• Survival rate based gender



```
ggplot(titanic, aes(x=Sex, fill=Survival)) +  
  theme_bw() + geom_bar() +  
  labs(y="Number of Passengers",  
       title="Survival Rate by Gender")
```


Ans-4).

Descriptive Statistics :-

Summary :- Gives up the descriptive stats like

In case of Numerical data :-

Gives mean, mode, median, Range

Measures of Central Tendency

⇒ mean (Titanic & fare) $\left[\begin{array}{l} \text{on Average Person} \\ \text{spent \$32 to board the} \\ \text{Titanic} \end{array} \right]$
32.20421

⇒ mode (Titanic & Age) $\left[\begin{array}{l} \text{mode Common Age on} \\ \text{Titanic} \end{array} \right]$
24

⇒ median (train & fare)
14.542

Measure of spread :

Range (Titanic & fare) $\left[\begin{array}{l} \text{It shows lowest \&} \\ \text{highest Value of fare} \end{array} \right]$
0.600 512.3292

⇒ $\text{var}(\text{ Titanic } \& \text{ fare})$

2469.437

⇒ $\text{sqrt}(\text{var}(\text{ Titanic } \& \text{ fare}))$

49.69343

Inferential Statistics :-

Hypothesis Testing :-

`new.data <- subset (Titanic, $ pclass == 1)`

⇒ `test2 = function (a, b, n) {`

`sample-mean = mean (a)`

`pop-mean = mean (b)`

`c = nrow = (n)`

`var b = var (b)`

`data = (sample-mean, pop-mean) / sqrt (var-b / c)`

`return data .`

call function as

`z.test2 (new.data $ survived, Titanic $ survived,
new.data)`

7.423820