

Saveetha School of Engineering + B

class test.3

ITA04 - Statistics With R programming

1. Program to perform the operations with the data.

a, Create data frame.

b, Extract score and attempts column by their positions.

c, Extract name and quality column by their positions.

d, To extract 3rd & 5th rows with 1st & 3rd columns.

e, To extract 'James' details.

Code:

```
exam_data <- data.frame(name = c('Anastasia', 'Dima', 'Katherine',  
  'James', 'Emily', 'Michael', 'Matthew', 'Laura'),  
  score = c(15.9, 9, 16.5, 12, 9, 20, 14.5, 13.5), attempts =  
  c(1, 3, 2, 3, 2, 3, 1, 1), qualify = c('Yes', 'No', 'Yes', 'no', 'no',  
  'yes', 'yes', 'no'))
```

```
score_column <- exam_data$score.
```

```
attempts_column <- exam_data$attempts.
```

```
name_column <- exam_data[, "name"]
```

```
qualify_column <- exam_data[, "qualify"]
```

```
subset_data <- exam_data[c(3,5), c(1,3)]
```

```
james_details <- exam_data[exam_data$name == "James",]
```

```

print(exam_data)
print(score_column)
print(name_column)
print(qualify_column)
print(subset_data)
print(james_details)

```

Output:

| name | score | attempts | qualify |
|-----------|-------|----------|---------|
| Anastasia | 12.5 | 1 | Yes |
| Dima | 9.0 | 3 | no |
| Katherine | 16.5 | 2 | Yes |
| James | 12.0 | 3 | no |
| Emily | 9.0 | 2 | no |
| Michael | 20.0 | 3 | Yes |
| Matthew | 14.5 | 1 | Yes |
| Laura | 13.5 | 1 | no |

12.5 9.0 16.5 12.0 9.0 20.0 14.5 13.5

1 3 2 3 2 3 1 1

Yes no Yes no no Yes Yes no

Katherine 2

Emily 2

James 12 3 no

7 1 4 5
8 2 4 5

2. Create Dataframe for the data.

i, R code to melt the data & display as a long format data?

ii, R code to use cast function appropriately to compute the average of x & y with respect to "time".

Code:

```
n <- c(1,1,2,2)
```

```
time <- c(1,2,1,2)
```

```
x <- c(6,3,2,5)
```

```
y <- c(1,4,6,9)
```

```
data <- data.frame(n, time, x, y)
```

```
melted_data <- melt(data, id.vars=c("n", "time"))
```

```
cast_data <- dcast(melted_data, time ~ variable, mean)
```

```
print(data)
```

```
print(melted_data)
```

```
print(cast_data)
```

Output:

| n | time | x | y |
|---|------|---|---|
| 1 | 1 | 6 | 1 |
| 1 | 2 | 3 | 4 |
| 2 | 1 | 2 | 6 |
| 2 | 2 | 5 | 9 |

| n | time | variable | value |
|---|------|----------|-------|
| 1 | 1 | x | 6 |
| 2 | 1 | y | 1 |
| 3 | 2 | x | 3 |
| 4 | 2 | y | 4 |
| 5 | 1 | x | 2 |
| 6 | 1 | y | 6 |
| 7 | 2 | x | 5 |
| 8 | 2 | y | 9 |

time

x

y

1

4.0

3.5

2

4.0

6.5