SAMPLES: Final Exam Questions

Sample: Question 1. Multiple choice questions. Select the right choice.

1. Quantitative variables take a predefined type which can be:
   a.
   b. Discrete
   c. Ordinal
   d. Nominal

2. Binary attributes are nominal attributes with
   a. only two values: 0 or 1
   b. ratings
   c. grades

3. Experimental data

   a. is collected from strictly controlled/designed experiments with efforts made to ensure statistical validity.
   b. Is collected from various resources such as internet, survey and external devices.

4. Data is dirty if it has
   a. various data types
   b. incomplete, noisy or inconsistent values

Sample: Question 2. True/false questions.

1. In bivariate exploration, we can use bar chart to explore the relationship between variables.
   a. True
   b. False

2. Pattern discovery is can be used to predict what will happened in future.

a. True
b. False

3. Visual clutter creates excessive cognitive load that can hinder the transmission of our message
   a. True
   b. False

Sample: Question 3. R code questions

1. What is the output of the following code:

```
x <- c (3, NA, NA, 1, 4)
y <- c (2, NA, 1, 2, 2)
x + y
```

Output: 5, NA, NA, 3, 6

2. What is the type of 'a' in the following code: x <- c (1, 2.1, FALSE)

Output: "double"

3. Write a code to read data from file f and print the name of columns.

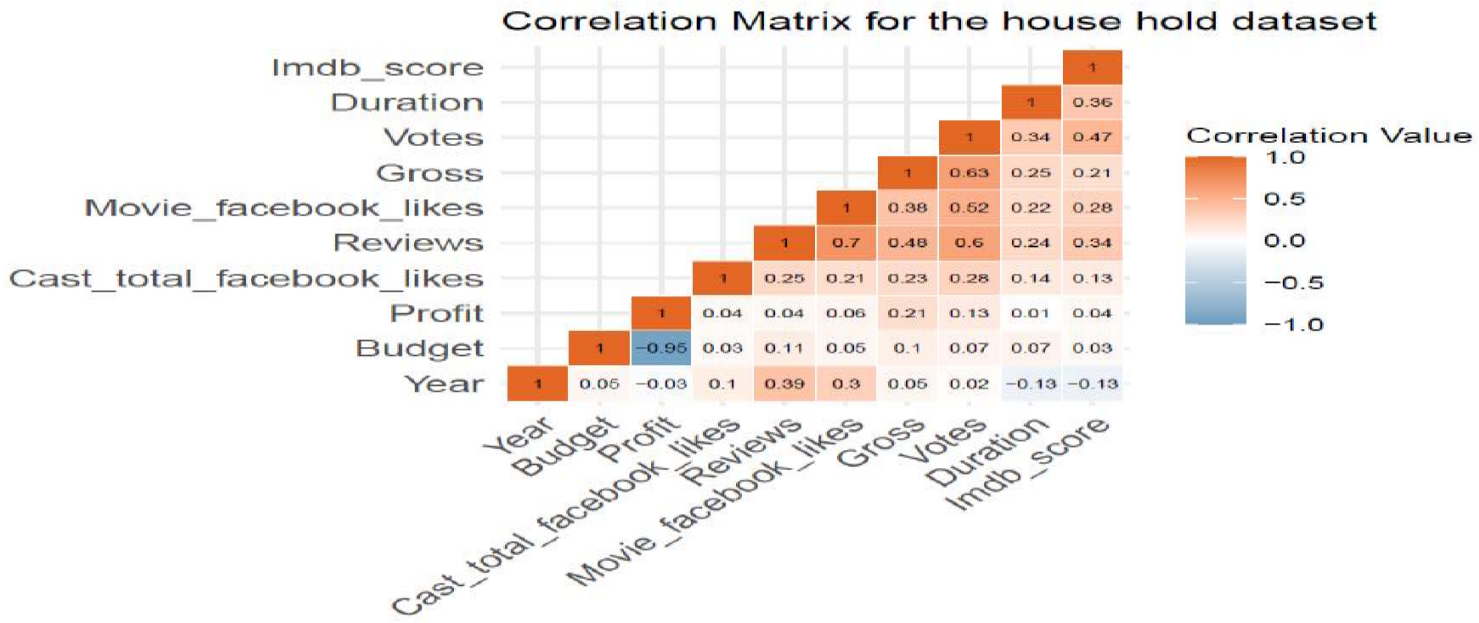
```
dat <- read.csv('f.csv',header=TRUE,sep = ",")
names(dat)
```

4. Write a code to read data from file f1. Use line chart to plot x column versus y column.

```
library("ggplot2")
dat <- read.csv('f1.csv',header=TRUE,sep = ",")
ggplot(dat, aes(x, y))+ geom_line()
```

Sample: Question 4. Short answer question

1. Based on the below Correlations matrix identify strong and weak correlation between variables.



Sample Answer:

In looking at the correlation matrix, it seems that Budget and Profit meet the criteria, with a correlation value of -0.95, which makes sense due to a higher budget, less profit. The second stronger correlation is between Reviews and Movie_Facebook_Likes with 0.7. That means that the slope is relatively close to 1/1.