Lab-02

Write your name here

February 04, 2022

Preface

The goal of this assignment is to gain more familiarity with data frames – think "spreadsheets" – and how to work with them in R. While the examples so far may seem fairly trivial, the value of using a tool like R to work with data frames should become evident as we work with larger datasets and do more complicated things. Bear with us!

In this lab we are providing some code snippets to serve as "scaffolding" to help guide you through each step. Over the course of the semester we will provide less scaffolding and more open-ended questions. As always, please come to office hours and reach out to your teaching staff if you have any questions.

Diamonds

We will work with the data frame `diamonds2` that has been derived from the `diamonds` data provided in the package **ggplot2**. Details of the original data are available here. The data frame we will work with has 7 variables:

```
head(diamonds2)
```

```
## # A tibble: 6 x 7
##   carat cut       color clarity depth table price
##   <dbl> <chr>     <chr> <chr>   <dbl> <dbl> <int>
## 1  0.23 Ideal     E     SI2      61.5    55   326
## 2  0.21 Premium   E     SI1      59.8    61   326
## 3  0.23 Good      E     VS1      56.9    65   327
## 4  0.29 Premium   I     VS2      62.4    58   334
## 5  0.31 Good      J     SI2      63.3    58   335
## 6  0.24 Very Good J     VVS2     62.8    57   336
```

1. Use both of the indexing operators [] and $ to get the price variable and assign it to a new object `price`. Print the head of `price`.

```
price <- FALSE # Replace FALSE with your code for the [] version
price <- FALSE # Replace FALSE with your code for the $ version
head(price) # After you complete the preceding lines, this will not return FALSE
```

```
## [1] FALSE
```

2. What are the mininum, maximum, and average prices of the diamonds?

Edit the inline code snippets to weave your answers to these questions into the provided text.

The minimum price is 0 dollars, the maximum price is `insert your inline code here` dollars, and the average price is `insert your inline code here` dollars.

3. Use the indexing operator [] to find the most expensive diamond and assign it to a new object `bling`. What are the attributes of this diamond?

```
bling <- diamonds2[FALSE, ] # Replace FALSE with your code
```

The most expensive diamond is numeric(0) carats, the color character(0), clarity `inline code`, depth = `inline code`, and table = `inline code`.

4. How many diamonds are included in this data set?

There are `inline code` diamonds in this data set.

5. What are the different types of diamond cut?

There are FALSE different types of diamond cut.

6. How many diamonds are there in the level of premium cut?

```
premium <- FALSE # Replace FALSE with your code
```

There are diamonds in the level of premium cut.

7. What fraction of diamonds are more than 1 carat?

```
# Your code goes here
```

Among all diamonds included in this dataset, a fraction of `inline code` are more than 1 carat.

8. What is the average price of the diamonds with Very Good cut and color I?

```
# Your code goes here
```

The average price of the diamonds with Very Good cut and color I is `inline code` dollars.