Question 1

You can access datasets from the R datasets package by using

```
data(NAME_OF_DATASET)
```

For this question, we will use the dimaonds data from the `ggplot2` library.

```
library(tidyverse) # Note the tidyverse package loads the ggplot2 library
data(diamonds)
```

Note you can learn about this dataset by using

```
help(diamonds)
```

a. Determine the (i) mode and (ii) class of the `diamonds` data object.

b. How would you find how many rows and columns the object has by using R functions `nrow` and `ncol` ? Give the code and the result.

c. What is the value contained in row 12345 and the `depth` column (which contains the depth percentage)?

d. Write a line of code that creates a new data object called `diamonds_imp` which is of the same mode and class as the original `diamonds` data object and contains the same columns as the original, but also contains three new columns: `x_imp` , `y_imp` , `z_imp` where each of these measurements are Imperial measurements in inches, i.e. `x_imp` is equal to `x` divided by 25.4, as there are 25.4 mm in 1 inch. Show the first 6 rows of the resulting data object.

e. Write a line of code that adds a column named `over_under` to the `diamonds_imp` data object that contains the difference between the price of the diamond in that row and the `median` of the prices of other diamonds with the same `color` .

f. Write a line of code that creates a new data object from the original `diamonds` data object named `Expensive` that contains only the diamonds whose price is *strictly* greater than $18800 and show the contents of that data object.