Data Science and R – Lab 14

Exercises with dplyr. Learning outcomes:

`select(), filter(), arrange(), mutate(), summarise(), group_by(), %>%`

0) Ensure the '`dplyr`' and '`ggplot2`' packages are installed and loaded. We'll be working with the `diamonds` dataset loaded in ggplot2.

1) dplyr verbs

`Use dplyr verbs to answer question 1`

a) Find all diamonds with `clarity` `"I1"` (10 out of 741 rows should be displayed. Enclose the R command inside `View()` to view all the results)

`Code:`_____

`Answer:`_____

b) Find all diamonds priced between $10,000 and $12,000 (inclusive)

Code:_____

c) How many diamonds have weight more than 4 `carat`? Display their weight (carat), cut, color and clarity

`Code:`_____

`Answer:`_____

d) Find all diamonds that have the best `cut` or the best `clarity`. Display their weight, cut, color and clarity

`Code:`_____

e) Order the diamonds from d) according to their weight, `carat`. What are the cuts of the 3 heaviest diamonds?

`Code:`_____

`Answer:`_____

f) Create a new column `price.per.weight` which is the price of the diamond per carat.

`Code:`_____

g) Arrange diamonds by the lowest price per weight and then by the highest `depth` and view the results using `View()`.

`Code:`_____

h) Try to think of ways to summarise the `price` or weight `(carat)` of diamonds based on the different `color` groups. Provide 3-5 summary functions.

`Code:`_____

_____

_____

2) Piping (%>%)

`Ensure the 'magrittr' package is installed and loaded`

Use the piping operator to answer the following questions. Repeat questions 1a) - g) using the piping operator.

a) Find all diamonds with `clarity` `"I1"`

`Code:`_____

b) Find all diamonds priced between $10,000 and $12,000.

`Code:`_____

c) How many diamonds have weight more than 4 `carat` in weight? Display their weight (carat), cut, color and clarity.

`Code:`_____

`Answer:`_____

d) Find all diamonds that have the best `cut` or the best `clarity`. Display their weight, cut, color and clarity.

`Code:`_____

e) Order the diamonds from d) according to their weight, `carat`. What are the cuts of the 3 lightest diamonds?

`Code:`_____

`Answer:`_____

f) Create a new column `price.per.weight` which is the price of the diamond per carat.

`Code:`_____

g) Arrange diamonds by the lowest price per weight and then by the highest `depth` and view the results using `View()`.

`Code:`_____

h) Summarise the weight of diamonds based on the different `cuts` of diamonds. Provide 3-5 summary functions (one being the count of each cut, `n()`), and then exclude those with counts less than 5000 and order by `cut` in descending order.

`Code:`_____