MTH208: Endsem Exam - 2024

Instructions

1. Follow the instructions for each question **exactly**.
2. Make sure you DO NOT have command `rm(list = ls())` in any R Scripts you submit.
3. Do not install any packages in your submitted scripts.
4. If your code is not properly commented or horizontal/vertical spacing is missing, points will be deducted.
5. There are three questions in this exam overall 100 points.

Questions

1. (40 points) Approximating the area of a circle with Rcpp

   a. Your goal in this problem will be to approximate the area of a circle of radius r with center (a, b) using the methods described below, with Rcpp. Recall, the equation of this circle for $(x, y) \in \mathbb{R}^2$ is

   $$(x-a)^2 + (y-b)^2 \leq r^2 \,.$$

   In order to approximate the area, you have to do the following steps:

   i. Consider a square that exactly contains the circle and draw a point uniformly from the square. This can be done by drawing a point t uniformly from the bottom side of the square and a point s uniformly from the left side of the box. Then (t, s) is a point uniformly generated from the square.

   ii. Check whether the point (t, s) is inside the circle.

   iii. If not inside the circle repeat i. and ii. until you get a point inside the circle.

   Repeat i., ii., iii, in order to get 1000 points inside the circle. Mathematically it is known that

   $$\Pr(\text{Point } (t,s) \text{ is inside the circle}) = \frac{\text{Area of circle}}{\text{Area of square}} \,.$$

   Using the above equation and the Steps i. - iii., your Rcpp function should return the approximated area of the circle (single numeric value)

   I will compile the Rcpp code using `sourceCpp` and the name of the function should be `area_circle`. Below is what the structure of the function should look like

```
// r: the radius of the circle
// center: coordinates for the center of the circle

... area_circle(double r, NumericVector center)
{
  ...
  ...
  ...
  return area;
}
```

Instructions:

- Copy and paste **ONLY** your final function into the file `ans1.cpp` in your exam repository. **DO NOT PASTE ANY OTHER CODE.**

2. (30 points) Numerical Stability

You are provided with an `ques2.Rdata` file that contains two numeric vectors `x_vec` and `y_vec`. Mathematically, I will denote `x_vec` as $x = (x_1, \dots, x_n)$ and `y_vec` as $y = (y_1, \dots, y_n)$. Your goal is to output a numeric vector of probabilities:

$$p_i = \frac{e^{x_i}}{e^{x_i} + e^{y_i}} \qquad \text{for } i = 1, \dots, n\,.$$

Denote $p = (p_1, \dots, p_n)$. Save p as a numeric vector `prob_vec`. Your code should look like the following:

```
...
...

prob_vec <- ...
```

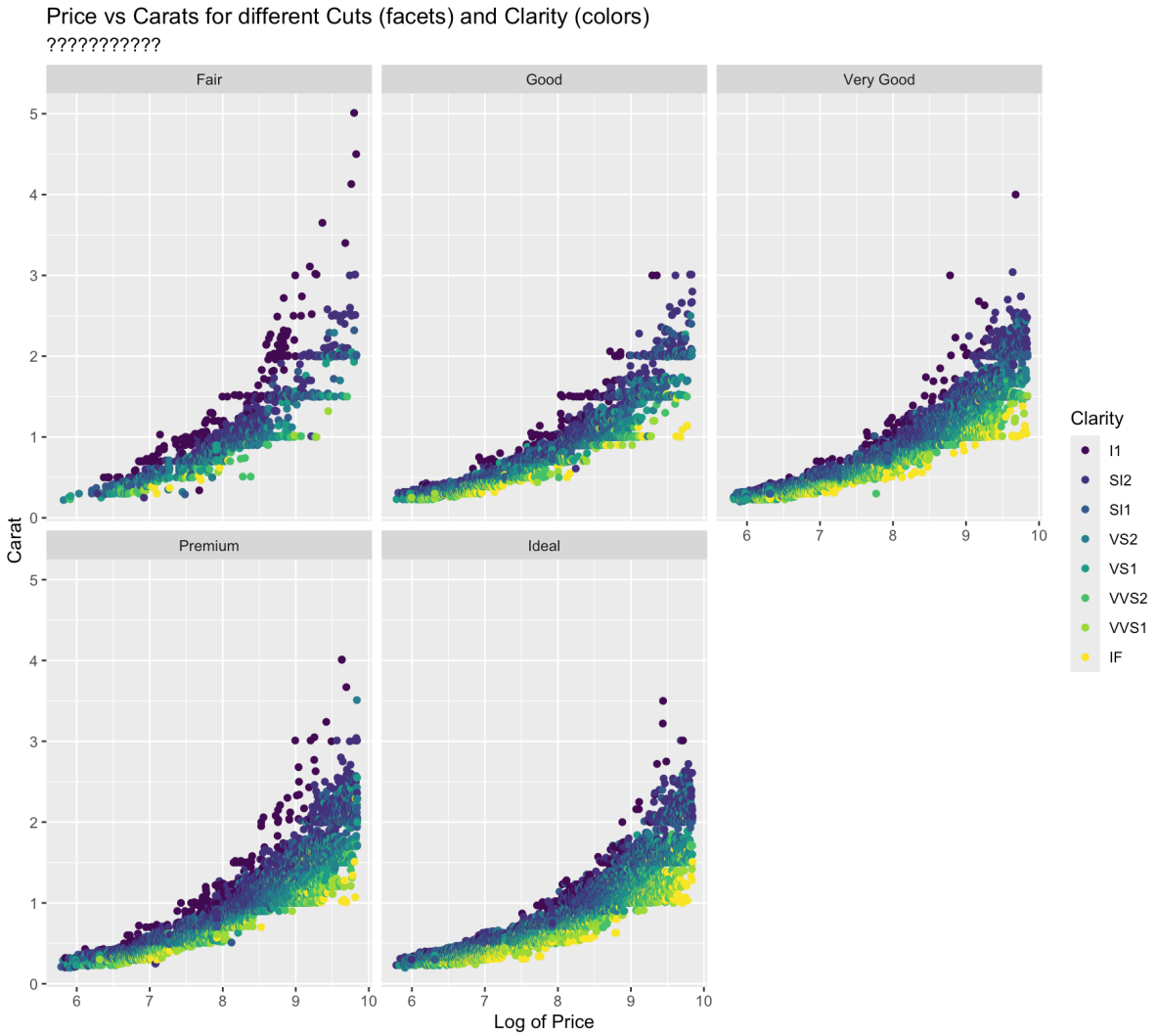
Instructions:

- Copy all and only relevant code and paste in the file `ans2.R`.

3. (30 points) In the library `ggplot2` is the dataset `diamonds` which contains data on the prices and other attributes of almost 54,000 diamonds. Looking at the data, I make the claim that

   The better the clarity of the diamond (IF), the cheaper it is.

   You, the data analyst produces the following plot to me, that you assert disproves my claim above.



   a. Recreate the above plot EXACTLY. You may find the following commands useful at some point:

      - `geom_point()`
      - `facet_wrap()`

   b. Replace the ?????? in the subtitle of the plot with a short explanation of how my claim is disproved. It must be a 1-line explanation.

   Save you ggplot in `diamondPlot` object and the last line of your code should must be `diamondPlot`. For example, your code should look like:

```
library(ggplot2)
data(diamonds)
...
...
...

diamondPlot <- ....
diamondPlot
```

Instructions:

- Copy all and only relevant code and paste in the file `ans3.R`.
- Only paste code required to produce your plot.