Assignment A01

A professor of machine learning is planning to get married to his long-time girlfriend. He has never shopped for diamonds before. In the mall, he was confronted with a dizzying array of diamond characteristics, configurations, and pricing. His quick search revealed that diamonds are primarily characterized by 4C viz. Color, Cut, Carat Weight and Clarity besides Polish, Symmetry, and certification. He scrapped the web to collect information from three different wholesaler websites to build his pricing model to ensure he does not get cheated while purchasing the diamond ring. Build a Linear Regression Model to predict the price of the diamond ring of his interest.

Professor's Diamond Ring

Price	Carat	Cut	Color	Clarity	Polish	Symmetry	Certification
$3100	0.9	Very Good	J	S12	Good	Very Good	GIA

i. Perform and upload your RMD file / Python Script on D2L

ii. Use suitable transformations to improve the variables predictive power.

iii. Use ggplot2 / Matplotlib and its extensions.

Consider exploring https://exts.ggplot2.tidyverse.org/gallery/

Evaluation Metric	**Marks**
Summary of Findings: Identify Determinants of Price Comments on distribution of data	2
Univariate Analysis of Metric and Non-Metric Data: Price, Carat, Clarity Color, Cut, Certification, Polish, Symmetry	2
Bivariate Analysis: Correlation, Covariance	2
Regression Model Simple and Multiple Linear Regression	2
Timely Submission	2