**The Maharaja Sayajirao University of Baroda**
**Department of Statistics, Faculty of Science**
**M. Sc. Previous(Sem-1)**
**Data Visualization**
**Assignment-1**

**Date: 31-07-2025**

**Consider the diamonds dataset in ggplot package in R. This dataset contains contain the prices and other attributes of almost 54,000 of diamonds.**

**Q.1** Create a bar chart showing the count of diamonds for each cut. Use fill to differentiate the colour

**Q.2** Modify the above chart to show proportions.

**Q.3** Create a boxplot to compare the distribution of diamond price across different clarity levels. Use a custom colour.

**Q.4** Create a scatter plot to visualize the relationship between carat and price. Use colour to differentiate the cut quality. Add custom legend to the plot also change the label of x- axis and y-axis.

**Q.5** For above scatter plot, create subplots corresponding to different qualities of cut in an appropriate configuration.

**Q.6** Create a box plot showing prices for diamonds with different colours of diamonds.

**Q.7** Create a histogram of price with density curve superimposed on it.

**Q.8** Create a bar chart showing average prices for diamonds with different qualities of cut. Display the average price on each bar.

**Q.9** Use log transformation for price to reduce skewness and visualize its relationship with carat. Use facet to compare. Also, break the axis tick of x-axis and add the title and axis label on the graph.

**Q.10** Create a scatter plot of depth and price and use reverse transformation to flip the y-axis in a plot.