**Principles of Data Science (5530)-Assignment 1**

**Name: Praveen Kumar Reddy Kadapala**

**ID: 16355737**

**2. Answer:**

**Analysis 1: Box Plot analysis of Math Scores distribution by Lunch type**

Created a box plot graph connecting Math Scores and lunch type, with the x-axis representing the lunch type category and the y-axis representing the Math Scores category.

A graph of a bar chart

Description automatically generated with medium confidence

**Analysis Outcome:** We can infer from this analysis that, the students who are having standard lunch type have obtained more math score compared to the students having free/reduced lunch type.

**Analysis 2: Bar Plot of race/ethnicity distribution**

Created a bar graph connecting Race/Ethnicity and frequency, with the x-axis representing the Race/Ethnicity and the y-axis with the number of instances of each Race/Ethnicity category.

A graph of different colored rectangular bars

Description automatically generated with medium confidence

**Analysis Outcome:** This box plot illustration tells us that compared to all the Race and Ethnicity the count of group C students are more with nearly 300 plus students, followed by group D students with a count of more than 250 plus, followed by group B students with a count of nearly 190 and then group E students with a count of 130 and at last group A students who are less compared other groups with a count of nearly 70.

**Analysis 3: Box plots for math score, reading score and writing score by gender.**

Created three box plot graphs using the math score, reading score and writing scores of students against gender.

A diagram of a graph

Description automatically generated with medium confidence

**Analysis Outcome:**  Above analysis shows that male students have scored good marks in math compared to female students whereas female students have scored good marks compared to male students in reading and writing.

**Visualization 4: Histogram for showing the distribution of Reading Scores**

Plotted a histogram by connecting Reading Scores and frequency, with the x-axis representing the Reading Scores and with the frequency distribution that corresponds to each score on the y-axis.

A graph of a distribution of reading scores

Description automatically generated

**Analysis Outcome:**  The frequency distribution shown above clearly shows that many students are scoring in the range of 65 to 80 in reading. Which indicates that scoring above 80 in reading subject might be difficult.

**Visualization 5: Pie Chart showing the distribution of Parental Education Level**

Plotted a pie chart on the distribution of parental education levels.

A pie chart with text

Description automatically generated

**Analysis Outcome:** The Pie chart shows that many of the student parent’s having studied till college level with 22.6%, followed by associate’s degree level with 22.2%, followed by high school with 19.6% and remaining are having a bachelor's degree (11.8%) or a master’s degree (5.9%).

**Folder Structure:**

**A screenshot of a computer program

Description automatically generated**