Results Discussions

**Part 1: Regression!**

**Overview:**

* The regression analysis aims to model the relationship between decibel levels (independent variable, X) and quiz marks (dependent variable, y). It uses Ordinary Least Squares (OLS) regression to estimate the linear relationship.

**Model Summary:**

* **R-squared**: The R-squared value is 0.878, indicating that approximately 87.8% of the variability in quiz marks can be explained by the decibel levels. This suggests a strong linear relationship between noise level and quiz performance.
* **Adjusted R-squared**: The adjusted R-squared is 0.877, slightly lower than the R-squared, reflecting minimal shrinkage, which is expected due to the single predictor variable.
* **F-statistic**: The model's F-statistic is 705.1, with a very low p-value (1.47e-46), indicating that the overall regression model is statistically significant. This implies that the decibel levels significantly affect quiz marks.
* **Coefficients**:
  + **Intercept (const)**: The intercept is approximately 129.1, suggesting that if the decibel level were zero (which is hypothetical in this context), the expected quiz marks would be around 129.1.
  + **Slope (x1)**: The slope coefficient for decibel levels is approximately -0.7767, indicating a negative relationship between noise level and quiz marks. Specifically, for every one-unit increase in decibel level, quiz marks decrease by approximately 0.7767.
  + **P-value for coefficients**: Both coefficients (intercept and slope) have p-values near zero, showing that they are statistically significant.

**Other Diagnostics:**

* **Log-Likelihood**: The log-likelihood value is -256.49, which serves as a baseline to compare this model to alternative models.
* **AIC/BIC**: The Akaike Information Criterion (AIC) is 517.0, and the Bayesian Information Criterion (BIC) is 522.2, both of which are used for model selection and penalize the number of parameters in the model.
* **Residual Analysis**:
  + **Omnibus Test and Jarque-Bera Test**: The p-values for these tests are low, suggesting that the residuals may not be normally distributed.
  + **Durbin-Watson Test**: The Durbin-Watson statistic is approximately 1.979, indicating no significant autocorrelation in the residuals.
  + **Skew and Kurtosis**: The skew is close to zero, and kurtosis is around three, which suggests the distribution of residuals is nearly symmetric but slightly leptokurtic (higher peak).

**Part 2: One-way ANOVA**

The ANOVA compares quiz marks across three noise categories:

* **Low (decibels < 55)**
* **Moderate (55 ≤ decibels < 75)**
* **High (decibels ≥ 75)**

**Results:**

* **F-value**: The F-value obtained from the ANOVA test indicates the ratio of variance between groups to the variance within groups. In this case, a higher F-value would suggest a significant difference in quiz marks across the different noise levels.
* **P-value**: A low p-value indicates that the differences in quiz marks across the three categories are statistically significant.

**Part 3: T-Test Analysis**

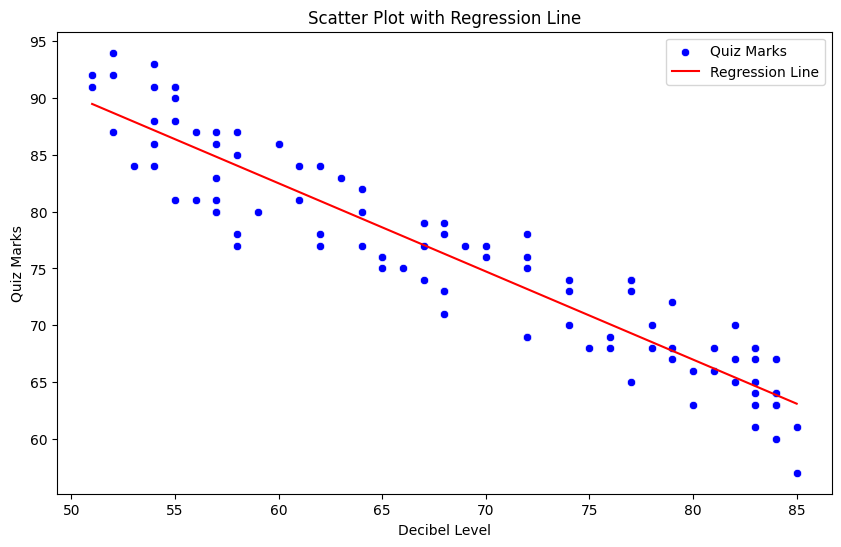
A t-test was conducted to compare quiz marks between the Low and High noise categories.

* **T-statistic**: The calculated t-statistic measures the difference in means relative to the variation in the quiz marks.
* **P-value**: A low p-value (usually <0.05) indicates a statistically significant difference in quiz marks between the Low and High noise groups.

**Conclusion**

The results indicate that noise levels (decibel levels) have a statistically significant negative effect on quiz performance. As the noise level increases, quiz marks tend to decrease. The ANOVA and t-test further confirm that significant differences exist in quiz marks between different noise level categories. Thus, minimizing noise could be beneficial for improving performance in tasks requiring concentration, such as quizzes.

Part 4: Data Visualizations



**Scatter Plot with Regression Line Analysis**

The scatter plot displays the relationship between **Decibel Level** (x-axis) and **Quiz Marks** (y-axis). Each blue dot represents an individual observation, indicating the quiz marks scored at a given noise level.

**Key Observations:**

1. **Negative Relationship**: The scatter plot shows a clear negative correlation between decibel levels and quiz marks. As the noise level (decibel level) increases, the quiz marks tend to decrease. This suggests that higher noise levels are associated with poorer quiz performance.
2. **Regression Line**: The red regression line represents the best-fit linear relationship between the two variables. The downward slope of the line confirms the negative association, indicating that on average, quiz marks drop as noise levels rise.
3. **Data Spread**: Although there is some spread around the regression line, the data points are relatively close to the line, indicating a strong linear relationship. This visual observation is consistent with the high R-squared value (0.878), which signifies that approximately 87.8% of the variation in quiz marks is explained by the decibel level.

**Short Report**

The scatter plot with the regression line highlights a significant negative relationship between noise levels and quiz performance. The analysis suggests that as noise levels increase, there is a noticeable decline in quiz marks. This pattern is consistent across the range of decibel levels, with higher noise levels resulting in lower average scores. The regression line serves as a visual representation of this trend, reinforcing the conclusion that reducing noise could potentially improve cognitive task performance, such as taking quizzes.

**Description of the Box Plot**

A diagram of a quiz marks

Description automatically generated with medium confidence

The box plot visualizes the distribution of quiz marks across three different noise level categories: Low, Moderate, and High.

* **Low Noise Level**:
  + **Median**: Around 90
  + **Interquartile Range (IQR)**: Approximately 87 to 93
  + **Whiskers**: Extend from around 85 to 95
  + **Color**: Blue
* **Moderate Noise Level**:
  + **Median**: Around 80
  + **IQR**: Approximately 75 to 85
  + **Whiskers**: Extend from around 70 to 90
  + **Color**: Grey
* **High Noise Level**:
  + **Median**: Around 67
  + **IQR**: Approximately 64 to 71
  + **Whiskers**: Extend from around 60 to 75
  + **Color**: Red
  + **Outliers**: One point below 60

**Short Report on the Diagram**

This box plot illustrates the relationship between noise level categories and quiz marks, suggesting that higher noise levels might negatively affect quiz performance. Here are the key observations:

1. **Low Noise Level**:
   * Students exposed to low noise levels achieved the highest quiz marks, with a median score around 90.
   * The distribution is relatively tight, indicating consistent performance among students.
2. **Moderate Noise Level**:
   * The median quiz marks drop to around 80.
   * There is more variability in scores compared to the low noise level category, as indicated by a wider interquartile range and whiskers.
3. **High Noise Level**:
   * The median quiz marks further drop to around 67.
   * This category shows the lowest quiz performance, with a broader spread of scores and an outlier below 60.

**Conclusion**

The data suggests a clear trend where increased noise levels correspond to lower quiz marks. Students performed best in low noise environments and worst in high noise environments. The consistency of scores also decreases as noise levels increase, indicating that higher noise levels not only lower performance but also lead to greater variability in student scores.

**Histogram**

A graph with a green line

Description automatically generated

**Description of the Histogram**

The histogram visualizes the distribution of quiz marks among students. The data is represented in bins, and a KDE (Kernel Density Estimate) line overlays the histogram to show the smoothed distribution.

* **Bins**:
  + The quiz marks are distributed into intervals (bins) from approximately 55 to 95.
  + The height of each bar represents the frequency of quiz marks within that interval.
* **Frequency**:
  + The most frequent quiz marks fall in the range of 75 to 80, with a frequency of around 20.
  + The range of 65 to 70 also has a relatively high frequency, just below 20.
  + Other intervals have lower frequencies, with the lowest frequencies observed at the extremes (55-60 and 90-95).
* **KDE Line**:
  + The KDE line provides a smooth estimate of the distribution, showing peaks and troughs corresponding to the histogram bars.
  + The peak of the KDE line is around the 75 mark, indicating the most common quiz score.

**Short Report on the Diagram**

The histogram provides a detailed view of the distribution of quiz marks, revealing several key insights:

1. **Central Tendency**:
   * The majority of quiz marks fall between 65 and 85, indicating that most students scored within this range.
   * The peak around 75-80 suggests that this is the most common score range.
2. **Spread**:
   * Quiz marks are spread across a range from approximately 55 to 95.
   * There is a noticeable drop in frequency at the extremes, with fewer students scoring below 60 and above 90.
3. **Distribution Shape**:
   * The distribution is approximately unimodal, with a single peak around the 75 mark.
   * The KDE line helps to visualize the smooth distribution of the data, highlighting the central tendency and variability.

**Conclusion**

The distribution of quiz marks indicates that most students scored between 65 and 85, with a peak around 75-80. The data shows a relatively balanced spread, with fewer students achieving very low or very high scores. This visualization is useful for understanding the overall performance of students on the quiz and can inform further analyses or educational interventions to address any observed patterns.

**Bar Plot**

A bar graph with different colored bars

Description automatically generated

**Introduction**

This report examines the relationship between noise levels and students' quiz performance. The study categorizes noise levels into three distinct groups: Low, Moderate, and High. The primary objective is to determine how varying noise levels impact students' average quiz marks.

**Data Visualization**

The bar plot provided illustrates the average quiz marks for each noise level category. The categories and their corresponding average quiz marks are as follows:

1. **Low Noise Level (Light Blue Bar):**
   * Average Quiz Marks: Approximately 85
2. **Moderate Noise Level (Pink Bar):**
   * Average Quiz Marks: Slightly below 80
3. **High Noise Level (Yellow Bar):**
   * Average Quiz Marks: Around 70

**Analysis**

The data reveals a clear trend indicating that higher noise levels negatively impact students' quiz performance. Specifically:

* **Low Noise Level:**
  + Students exposed to low noise levels achieved the highest average quiz marks, just above 85. This suggests that a quiet environment is highly conducive to better academic performance.
* **Moderate Noise Level:**
  + In a moderately noisy environment, the average quiz marks drop to slightly below 80. While still relatively high, this category shows a noticeable decrease in performance compared to the low noise level category.
* **High Noise Level:**
  + Students subjected to high noise levels scored the lowest, with average quiz marks around 70. This significant reduction highlights the detrimental effect of loud environments on students' ability to perform well in quizzes.

**Conclusion**

The bar plot clearly demonstrates the inverse relationship between noise levels and quiz performance. As noise levels increase, the average quiz marks decrease. These findings emphasize the importance of maintaining a low-noise environment to enhance students' academic performance. Educational institutions and students should consider minimizing noise to create optimal learning conditions.