**Team Research and Development**

**GitHub Repository URL**:

<https://github.com/M-ismail-UH/Team-Research-and-Development>

**Group Number:** 174A

**Student ID’s:**

<https://github.com/Aktheysay> 23036945

<https://github.com/kashif-shahzad-ai> 23015843

<https://github.com/tarikuk> 23010627

<https://github.com/M-ismail-UH> 23001751

<https://github.com/ammmar-hash> 23031629

**Dataset ID:**

DS060 (Top 240 restaurants in LA )

This dataset is interesting to us because this dataset is interesting to us because this dataset allows us to explore how restaurant ratings relate to price levels and cuisine popularity across neighborhoods in Los Angeles  
  
**Independent Variable:** Style (type of cuisine) - Nominal/Categorical

**Dependent Variable**: Star Rating - Interval/Measurement

# **Statistical Test**:

**Parametric Test:** Use ANOVA if normality holds.

**Non-Parametric Alternative**: Use Kruskal-Wallis if normality is violated. Hypothesis Testing Framework:

# **Research Question**

Is there a difference in the mean of StarRating among different Styles of cuisine”.

# **Hypothesis**

Null hypothesis (H0): There is no significant difference in the mean StarRating among the different Styles of cuisine.

Alt hypothesis (H1): There is a significant difference in the mean StarRating among the different Styles of cuisine.

A screenshot of a computer

Description automatically generated

# **Histogram:**

A graph with a curve

Description automatically generated

The normal curve overlaid on the histogram **does not** track the shape of the underlying data closely, and this is expected given that there are clear skew and tail deviations from normality in these distributions. **Spearman’s Rho or Kendall’s Tau** in view of the fact that these tests do not assume normality and therefore are more appropriate for data that is other than normal

## **X and Y Axes:**

The x-axis is titled **Star Rating**. The y-axis reads "Frequency," which is correct, Since the task emphasizes direct counts (how many restaurants fall within a specific star rating), labeling the y-axis as "Frequency" makes the chart more intuitive for this analysis.

## **Title:**

It has a title that describes the plot: **Histogram of Star Ratings with Normal Curve Overlay**.

## **Normal Bell Curve:**

The smooth normal bell curve on top of the histogram

## **Data Visualization: For Comparison of Means/median and Correlation questions only. Is the Histogram correct?**

Yes

## **Dependent variable:**

The histogram describes the dependent variable (**StarRating**) appropriately, as it presents a frequency distribution.

## **Normal Curve Overlay:**

Having a normal curve overlay thus helps you to visually check for the data distribution if its normal or not, which is important in choosing the right statistical test to perform.

## **Axes and Labels:**

Axes are correctly labelled, and the plot has a sensible title fulfilling our third requirement for clear and informative visualization.

**Suitability for Analysis:**

The histogram accurately enables

## **Comparison of Means/Medians:**

This depends on the shape of the distribution, to choose a part between parametric and non-parametric tests.

## **Correlation assessment:**

The difference is the main part, but the distribution of dependent variable cannot be neglected as well.

## **Data Visualisation: Does response state whether dep variable appears normally distributed or not? (N/A for comp of proportions)?**

The distribution of our dependent variable, **StarRating**, shows a small right skew away from the normal curve. It does not quite fit the bell-shaped curve of a normal distribution. The data does not follow perfect normality. Although the histogram displays near-symmetry, deviations suggest that this distribution does not meet the parametric assumption of normality.

# **Rscript**

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a graph

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated

A screenshot of a computer

Description automatically generated

An analysis of variance (ANOVA) was conducted to determine the effect of restaurant style on star ratings. The results revealed a highly significant effect of style on star ratings, F(5, 120) = 88.85, p < 0.001.

# **Results Section:**

A one-way ANOVA was conducted to evaluate the effect of restaurant style on star ratings. The analysis revealed a statistically significant effect of style on star ratings, F(5, 120) = 88.85, p < 2e-16. This indicates that the mean star ratings differ significantly across restaurant styles

# **Discussion Section:**

The results show that restaurant style plays a crucial role in determining customer ratings. Styles such as [insert specific styles if post-hoc analysis was conducted] might appeal more to customers, leading to higher ratings. These findings highlight the importance of considering restaurant style as a factor in marketing and business strategies.

F value = 88.85

P value < 2e-16