



EDS 6333 - EXCEL PROJECT

Functions and Formulas Guide

Group – 1

Group Members:

1. Ruchith Reddy Parnem
2. Manikanta Rudru
3. Janavi Koonamneni
4. Neha Reddy Poreddy

March 2, 2025

Contribution Statement

Each group member independently completed all five questions in the project. To ensure the highest quality, we carefully reviewed each member's work. For the final submission, we selected the best responses for every question, integrating the most accurate calculations, clear explanations, and well-structured visualizations.

Contents

Contribution Statement	1
1 Question 1: Frequency Analysis & Visualization - Manikanta Rudru	3
1.1 Functions and Formulas Used	3
1.2 Visualization Techniques	3
2 Question 2: Advanced Analysis - Manikanta Rudru	3
2.1 Additional Techniques	3
3 Question 3: Descriptive Statistics - Janavi Koonamneni	4
3.1 Descriptive Statistics Functions and Formulas	4
3.2 Box Plot Creation	5
3.3 Correlation and Covariance Analysis	5
4 Question 4: Frequency Analysis - Neha Reddy Poreddy	5
4.1 Functions and Formulas	5
4.2 Univariate Data Analysis Findings	6
4.2.1 Quality Rating:	6
4.2.2 Meal Price Rating:	6
4.3 Bivariate Data Analysis Findings	7
5 Question 5: Statistical Functions - Ruchith Reddy Parnem	8
5.1 Functions and Formulas	8
5.2 Box Plot Creation	8

1 Question 1: Frequency Analysis & Visualization - Manikanta Rudru

1.1 Functions and Formulas Used

Function	Formula	Purpose
Frequency Calculation	=COUNTIF(range, target_value)	Counts occurrences of specific values
Relative Frequency	=G5 / COUNTA(\$B\$2:\$B\$21)	Calculates proportion of each value (G5 contains frequency)
Percentage Frequency	=G5 / COUNTA(\$B\$2:\$B\$21) * 100	Converts relative frequency to percentage
Total Count of Entries	=SUM(data_range)	Sums all entries in a range

Table 1: Frequency Analysis Functions

1.2 Visualization Techniques

1. Bar Chart Creation:

- Create a frequency table
- Select the relevant data
- Go to Insert → Charts → Bar Chart
- Select 2-D Column – Stacked Column Chart

2. Pie Chart Creation:

- Create a frequency table
- Select the necessary data
- Go to Insert → Charts → Pie Chart
- Select 2-D Pie Chart

2 Question 2: Advanced Analysis - Manikanta Rudru

The same functions used in Question 1 (Frequency, Relative Frequency, Percentage Frequency, Sum, Bar Chart, and Pie Chart) are also applied here, with these additional methods:

2.1 Additional Techniques

1. Scatter Plot:

- Select the relevant columns from the dataset

- Navigate to the Insert tab and go to the Charts section
- Choose the Scatter Plot option to visualize the relationship between variables

2. Cross Tabulation (Pivot Table):

- Either select the required columns first and then click on the Insert tab (PivotTable option in top-left corner), or start by clicking Insert, then selecting the data range before pressing OK

Organizing the Pivot Table:

- Rows: Drag the "Type of Customer" field into the Rows section
- Columns: Place the "Net Sales" field into the Columns section
- Values: Drag the "Net Sales" field again into the Values area to calculate the total sales per customer type
- Data can also be grouped into classes of 20 instances for better readability and analysis

3 Question 3: Descriptive Statistics - Janavi Koonamneni

3.1 Descriptive Statistics Functions and Formulas

Function	Formula	Example
Mean	=AVERAGE(data_range)	=AVERAGE(C5:C44)
Median	=MEDIAN(data_range)	=MEDIAN(D5:D44)
Mode	=MODE(data_range)	=MODE(C5:C44)
First Quartile (Q1)	=QUARTILE.INC(data_range, 1)	=QUARTILE.INC(D5:D44, 1)
Median (Q2) or =MEDIAN(data_range)	=QUARTILE.INC(data_range, 2)	=QUARTILE.INC(D5:D44, 2)
Third Quartile (Q3)	=QUARTILE.INC(data_range, 3)	=QUARTILE.INC(D5:D44, 3)
Range	=MAX(data) - MIN(data)	=MAX(D5:D44) - MIN(D5:D44)
Interquartile Range (IQR)	=QUARTILE.INC(data, 3) - QUARTILE.INC(data, 1)	=QUARTILE.INC(D5:D44, 3) - QUARTILE.INC(D5:D44, 1)
Sample Variance	=VAR.S(data_range)	=VAR.S(H56:H95)
Standard Deviation (SD)	=SQRT(VAR.S(data))	=SQRT(VAR.S(H56:H95))
Minimum (Smallest Value)	=MIN(data_range)	=MIN(C5:C44)
Maximum (Largest Value)	=MAX(data_range)	=MAX(D5:D44)

Table 2: Descriptive Statistics Functions

3.2 Box Plot Creation

To create a box plot in Excel:

1. Organize data in a single column, sorted in ascending order
2. Select the range to include in the plot
3. Go to the Insert tab on the Ribbon and click on Insert Statistic Chart
4. Choose the Box and Whisker chart option
5. Excel automatically generates the box plot showing minimum, Q1, median, Q3, and maximum values, along with any outliers
6. Customize the chart by adding titles and adjusting labels for clarity

3.3 Correlation and Covariance Analysis

Function	Formula	Example
Covariance	=COV.S(array1, array2)	=COV.S(F182:F221, G182:G221)
Standard Deviation (SD)	=SQRT(VAR.S(array))	=SQRT(VAR.S(F182:F221))
Correlation Coefficient (r)	= COV.S(x, y) / (SD(x) * SD(y))	=T186 / (T189 * T190)

Table 3: Correlation and Covariance Functions

Findings: The covariance formula calculates how two variables vary together, and the correlation coefficient standardizes this measure to show the strength of the relationship. A correlation coefficient of 0.961 suggests a very strong positive linear relationship between the two variables, meaning that as one variable (e.g., Revenue) increases, the other variable (e.g., Estimated Team Value) also tends to increase consistently and predictably.

4 Question 4: Frequency Analysis - Neha Reddy Poreddy

4.1 Functions and Formulas

Function	Formula	Example
Count	=COUNTIF(range, value)	=COUNTIF(B2:B301, 2)
Relative Frequency	=Individual grade frequency/total frequency	=H12/COUNTA(B2:B301)
Percentage	=(Relative frequency)*100	=(H12/COUNTA(B2:B301))*100
Sum	=SUM(range)	=SUM(H12:H14)

Table 4: Frequency Analysis Functions for Question 4

4.2 Univariate Data Analysis Findings

4.2.1 Quality Rating:

- **Distribution Overview:**
 - Rating 2 (moderate quality): 150 instances (50%)
 - Rating 1 (low quality): 84 instances (28%)
 - Rating 3 (high quality): 66 instances (22%)
- **Visual Analysis:**
 - Bar charts confirm rating 2 as the predominant category
 - Frequency and relative frequency patterns align consistently
 - Pie chart visualization shows rating 2 comprising exactly half of all ratings
- **Key Observations:**
 - Ratings cluster around the middle value, indicating average quality performance
 - Lower ratings exceed higher ratings by 6 percentage points
 - The skew toward lower quality suggests room for improvement initiatives
- **Strategic Implications:**
 - Focus should be on converting moderate ratings to high ratings
 - Addressing quality deficiencies could shift the distribution toward higher ratings
 - Current pattern provides a baseline for measuring improvement efforts

4.2.2 Meal Price Rating:

- **Distribution Overview:**
 - Moderate pricing (Rating 2): 117 instances (39%)
 - Higher pricing (Rating 3): 105 instances (35%)
 - Lower pricing (Rating 1): 78 instances (26%)
- **Visual Analysis:**
 - Bar charts show moderate pricing as the most prevalent category
 - Higher pricing follows closely behind the moderate category
 - Pie chart confirms moderate pricing represents the largest segment
- **Key Observations:**
 - Combined moderate and high pricing ratings account for 74% of all responses
 - Distribution shows a clear tilt toward mid-to-high price perception
 - Lower price perceptions represent the smallest portion of feedback

- **Strategic Implications:**

- Customer base generally perceives prices in the moderate to higher range
- Lower affordability perceptions are less common among respondents
- Businesses may need to evaluate if current pricing strategy aligns with desired market positioning

4.3 Bivariate Data Analysis Findings

Overview: This analysis examines the relationship between Quality Ratings and Meal Price Ratings using cross-tabulated data visualized in a bar chart.

- **Data Structure:**

- X-axis: Quality Rating (1-3)
- Y-axis: Meal Price Rating (1-3, color-coded)
- Values represent count frequencies for each rating combination

- **Count Distribution:**

- **Low Quality (Rating 1):**
 - * Low Price (1): 42 instances
 - * Medium Price (2): 39 instances
 - * High Price (3): 3 instances
- **Medium Quality (Rating 2):**
 - * Low Price (1): 33 instances
 - * Medium Price (2): 63 instances
 - * High Price (3): 54 instances
- **High Quality (Rating 3):**
 - * Low Price (1): 3 instances
 - * Medium Price (2): 15 instances
 - * High Price (3): 48 instances

- **Key Patterns:**

- Medium quality ratings show the highest overall frequency
- High prices rarely coincide with low quality ratings (only 3 instances)
- Low prices rarely coincide with high quality ratings (only 3 instances)
- Medium quality most commonly aligns with medium pricing
- High quality predominantly corresponds with high pricing

5 Question 5: Statistical Functions - Ruchith Reddy Parnem

5.1 Functions and Formulas

Function	Formula	Description
Mean	=AVERAGE(Range of Data)	Calculates the average of a data set
Median	=MEDIAN(Range of Data)	Finds the middle value of a data set
Mode	=MODE(Range of Data)	Identifies the most frequent value
Quartiles (Q1, Q2, Q3)	=QUARTILE.INC(Range of Data, k)	Finds the kth quartile (where k = 1, 2, or 3)
Range	=MAX(Range of Data) - MIN(Range of Data)	Calculates the difference between maximum and minimum values
Inter Quartile Range (IQR)	=QUARTILE.INC(Range of Data, 3) - QUARTILE.INC(Range of Data, 1)	Calculates the difference between Q3 and Q1
Lower limit	=QUARTILE.INC(Range of Data, 1) - 1.5*(Q3 - Q1)	Determines lower boundary for outliers
Upper limit	=QUARTILE.INC(Range of Data, 3) + 1.5*(Q3 - Q1)	Determines upper boundary for outliers
Sample Variance	=VAR.S(Range of Data)	Calculates variance where $\text{variance}(x) = (x - \text{mean})^2 / (n - 1)$
Standard Deviation	=SQRT(VAR.S(Range of Data))	Square root of variance
Coefficient of Variation	=(SQRT(VAR.S(Range of Data)))/AVERAGE(Range of Data)*100	(Standard Deviation/Mean) * 100%
Maximum	=MAX(Range of Data)	Finds the largest value in a data set
Minimum	=MIN(Range of Data)	Finds the smallest value in a data set
Sample Covariance	(X-Mean(x))*(Y-Mean(y))/Total(n)-1	Measures how two variables change together
Sample Correlation Coefficient	Sample Covariance(x,y)/(Standard Deviation(x)*Standard Deviation(y))	Standardized measure of correlation between variables

Table 5: Statistical Functions for Question 5

5.2 Box Plot Creation

Select all the values of the dataset, then open the Insert tab. In the chart section, select "Box and Whisker" to create a box plot visualization.