# SURVEYMASTER PRO: DYNAMIC DATA ANALYSIS & VISUALIZATION HUB FOR IN-DEPTH SURVEY INSIGHT

## Description of the Proposed System

This Python-based statistical system automation is named 'SURVEYMASTER PRO'. This proposed system addresses the challenge of making survey data analysis accessible and understandable to a broad audience, particularly students and researchers with limited statistical knowledge.

The development of this system arises from a compelling need to democratize survey data analysis, making it accessible and comprehensible to researchers, particularly students, who may lack expertise in statistics. This system aims to eliminate reliance on costly statisticians and empower individuals with limited statistical knowledge to conduct meaningful survey data analysis independently.

The proposed system is designed to give researchers an easier way to analyze their survey data with interpretations and graphs. However, this section outlines the boundaries and defines how much this system will provide support while identifying specific processes beyond its purview. Additionally, external factors that may influence the system's performance are discussed.

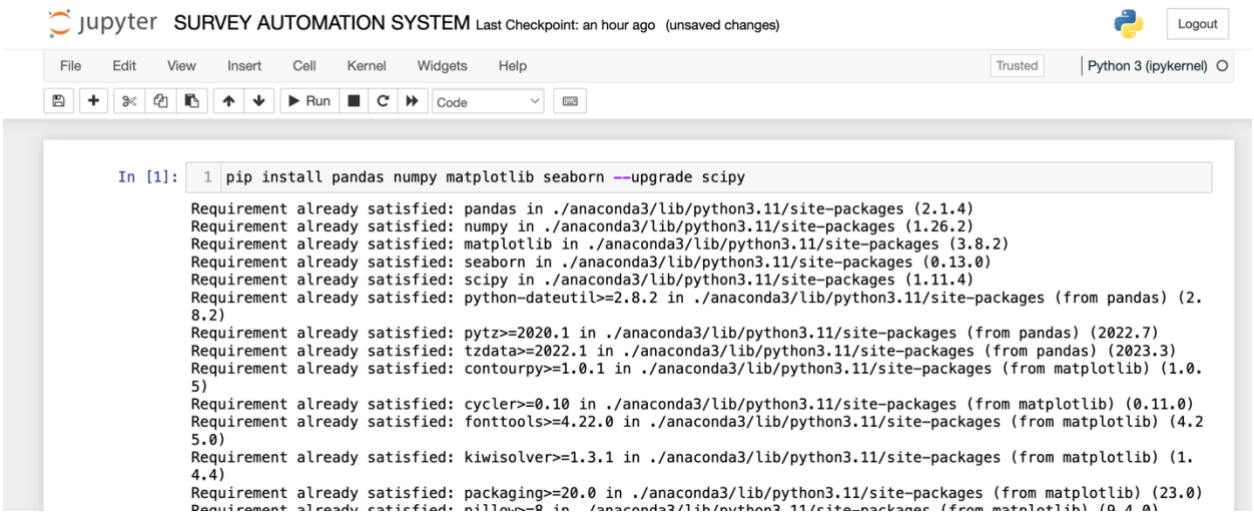
The system encompasses a range of critical functions and features tailored to assist researchers, particularly students, in navigating the intricacies of survey data analysis:

1. **Survey Data Processing**
2. **Descriptive Statistics (Mean, Frequencies, Percentages)**
3. **Graphical Visualization**
4. **Statistical Tests (T-TEST, P-TEST, ANOVA)**
5. **Demographic Analysis:**
6. **Category/Factor Analysis:**
7. **Significant Differences Assessment:**
8. **Correlational Analysis:**

## PROGRAM AUTOMAINTERFACE

### I. Install Python and Libraries:

```
In [ ]: 1 pip install pandas numpy matplotlib seaborn --upgrade scipy
In [ ]: 1 pip install dash dash-core-components dash-html-components dash-table plotly
In [ ]: 1 pip install dash dash-renderer dash-html-components dash-core-components plotly
In [ ]: 1 pip install dash dash-bootstrap-components dash-table
In [ ]: 1 pip install dash plotly
In [ ]: 1 pip install dash dash-bootstrap-components pandas
In [ ]: 1 pip install tabulate
```



- **Run the Script**

# SURVEYMASTER PRO: SURVEY DATA ANALYSIS AUTOMATION SYSTEM

## USAGE GUIDELINES:

**Kindly adhere to the following recommendations to minimize errors and achieve the utmost precision in your outcomes.**

1. It is recommended to run this code in Jupyter Notebook within the Anaconda environment for optimal accuracy, as certain features may not be available in other programming environments or languages.
2. File Format: Ensure your survey data is in Excel format (.xlsx).
3. The format of your excel datasets should be presented as follows:

| Age | Gender | Location | RTU is a Good School | The Facility of RTU is promising | BS Satatistics is Easy |
|---|---|---|---|---|---|
| 15 | Male | Manila | Agree | Strongly Agree | Disagree |
| 18 | Female | Cavite | Disagree | Agree | Strongly Agree |

The first columns in your Excel file should contain demographic or personal information, like Age, Gender, Location, etc.. The remaining columns should contain survey questions

In the example data above, the first three (3) columns are reserved for demographic information (Age, Gender, Location). All columns following these initial three columns are dedicated to survey questions.

4. Import Libraries: Make sure to install the following required libraries at the beginning of your code:
5. Categorical Mappings: If your survey contains non-numeric responses, consider creating mappings to convert them to numeric values.
6. Interactive Dashboards: Some analysis options provide interactive dashboards. Use dropdown menus and graphs to explore data interactively.
7. Run the code and follow the prompts and instructions.
8. Enjoy Exploring Your Survey Data: This automation code simplifies survey data analysis, allowing you to gain valuable insights efficiently.

**Have you read and taken into account the provided instructions? Please respond with 'yes' or 'no':**

**Have you read and taken into account the provided instructions? Please respond with 'yes' or 'no':**

YES

## COLLECTION AND CLEANING OF SURVEY DATASET

**INSTRUCTION:** Please ensure you have the directory path of your survey data file.

Your file should be in Excel (.xlsx) format. For example: 'C:/Users/YourName/Documents/survey_data.xlsx'

For guidance on how to find your file's directory, visit: here.

**Enter the directory path of your survey data file:**

`surveydata.csv`

**Error: No data loaded. Please check your file path and ensure it's an Excel file. Please try again.**

**Enter the directory path of your survey data file:**

`/Users/rolanddelarosa/Desktop/FoodVlogData.xlsx`

**Enter the number of columns with demographics of the respondents, for example, age, gender, and/or location (You may refer to USAGE GUIDELINES No.2 for further instruction.):**

6

**Choose how to handle missing data (1: Drop, 2: Fill with mean/mode, 3: Fill with a value):**

`0`

**Error: Please enter a valid choice (1, 2, or 3).**

**Choose how to handle missing data (1: Drop, 2: Fill with mean/mode, 3: Fill with a value):**

2

## III. Selection of Analysis Type (PROGRAM CAPABILITIES):

**Choose how to handle missing data (1: Drop, 2: Fill with mean/mode, 3: Fill with a value):**

2

### PROGRAM CAPABILITIES

**1. DEMOGRAPHICS OF THE RESPONDENTS**

- Create an interactive table displaying demographics data for respondents, including age, gender, and location. Show measures, frequency, and percentages.

**2. CATEGORY/FACTOR ANALYSIS**

- Analyze questions that measure a single factor or category. Provide the titles and questions under each factor/category as an example.

**3. CORRELATIONAL ANALYSIS**

- Investigate the presence of positive relationships between survey questions and between factors/categories. Generate a heatmap and highlight the top 3 strong positive relationships.

**4. SIGNIFICANT DIFFERENCES BETWEEN DATAS**

- Examine if there are significant differences between respondent demographics and their responses within each factor/category. *This analysis should only be conducted after completing the category/factor analysis (Capability No.2).*

**5. INDIVIDUAL ANALYSIS FOR ALL QUESTIONS**

- Perform an individual analysis of all survey questions, presenting them with interactive charts for comprehensive insights.

**Select the analysis you want to perform (e.g., '1' for Profile of the Respondents, '1,2,3' for multiple analyses, or 'all' for all analyses):**

4

**Error: You cannot select '4' without selecting '2'.**

**Select the analysis you want to perform (e.g., '1' for Profile of the Respondents, '1,2,3' for multiple analyses, or 'all' for all analyses):**

er

**Error: Invalid input. Please enter a digit from 1 to 5, digits from 1 to 5 separated by a comma, or 'all'. Example: '1', '2,3', 'all'.**

**Select the analysis you want to perform (e.g., '1' for Profile of the Respondents, '1,2,3' for multiple analyses, or 'all' for all analyses):**

ALL

## IV. NON-NUMERIC RESPONSE CONVERSION:

**Select the analysis you want to perform (e.g., '1' for Profile of the Respondents, '1,2,3' for multiple analyses, or 'all' for all analyses):**

ALL

### NON-NUMERIC RESPONSES CONVERSIONS:

**INSTRUCTION:** Based on checking, your survey contains non-numeric responses: **Strongly Agree**, **Strongly Disagree**, **Agree**, **Disagree**

Please convert these non-numeric responses into numeric values so that the system can calculate them.

**Non-numeric Response Conversion Table Example:**

| Non-numeric | Numeric Equivalent |
|---|---|
| Very Satisfied | 4 |
| Satisfied | 3 |
| Dissatisfied | 2 |
| Very Dissatisfied | 1 |

Please enter the numeric value for **Strongly Agree**:

0

**Error:** Invalid input. Please enter a valid numeric value greater than 0.

Please enter the numeric value for **Strongly Agree**:

4

Please enter the numeric value for **Strongly Disagree**:

1

Please enter the numeric value for **Agree**:

3

Please enter the numeric value for **Disagree**:

2

## V. Profile of the Respondents:

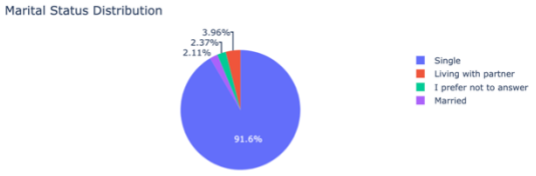
Please enter the numeric value for **Disagree**:

2

### I. PROFILE OF THE RESPONDENTS

Marital Status

| Measure | Frequency | Percentage |
|---|---|---|
| Single | 347 | 92% |
| Living with partner | 15 | 4% |
| I prefer not to answer | 9 | 2% |
| Married | 8 | 2% |




**VERBAL INTERPRETATION [Demographic 1 : Marital Status]**

**According to the result of 'Marital Status' demographics, 'Single' has the highest number of respondents with 92% while the lowest is 'I prefer not to answer' with 2%**

### I. PROFILE OF THE RESPONDENTS

Employment Status
Age
Gender
Marital Status
Educational Attainment
**Employment Status**
District location in Metro Manila


**VERBAL INTERPRETATION [Demographic 1 : Employment Status]**

**According to the result of 'Employment Status' demographics, 'Student' has the highest number of respondents with 68% while the lowest is**

## VI. CATEGORY/FACTOR ANALYSIS

### II. CATEGORY/FACTOR ANALYSIS

**DESCRIPTION:** This program capability enables you to analyze your survey questions grouped under a single category. This feature is particularly useful when you want to assess questions that measure a specific factor or category.

In your survey data file, each question is assigned a unique number for your convenience (Q1, Q2, Q3..). This means you no longer need to copy and paste questions. Instead, you can simply input the corresponding question number."

Q1. The food vlogger is interesting, relatable and entertaining to watch
Q2. The food vlogger's non-verbal gestures looks real and natural
Q3. The food vlogger present the information in cheerful and energetic manner
Q4. The food vlogger's content is well-organized and cited.
Q5. The food vlogger's video received generally good feedback and comments.
Q6. The food vlogger influenced me to check out and visit the food hub.
Q7. The food vlogger has culinary background.
Q8. The food vlogger present detailed information about Ugbo as food hub destination.
Q9. The food vlogger has knowledge about the culture behind the featured foods.

**INSTRUCTION:** Please specify the titles for each of your categories and their respective questions by entering the corresponding question numbers from the list above. Once you have listed all of the categories, simply type 'done' to continue.

**EXAMPLE:**
Enter the 1st category title (or 'done' to finish): ENTERTAINMENT
Enter the question numbers for the 'ENTERTAINMENT category, separated by commas. (Please refer to the numbered questions above): Q1, Q2, Q3, Q4

```
Enter the 1st category title (or 'done' to finish): ENTERTAINMENT
Enter the question numbers for the 'ENTERTAINMENT' category, separated by commas. (Please refer to the numbered que
stions above. Enter as Q1, Q2, Q3, Q4):  Q1, Q2, Q3
Enter the 2nd category title (or 'done' to finish): CREDIBILITY
Enter the question numbers for the 'CREDIBILITY' category, separated by commas. (Please refer to the numbered quest
ions above. Enter as Q1, Q2, Q3, Q4):  Q4,Q5,Q6
Enter the 3rd category title (or 'done' to finish): EXPERTISE
Enter the question numbers for the 'EXPERTISE' category, separated by commas. (Please refer to the numbered questio
ns above. Enter as Q1, Q2, Q3, Q4):  Q7, Q8, Q9
Enter the 4th category title (or 'done' to finish): DONE
```

```
Enter the question numbers for the 'EXPERTISE' category, separated by commas. (Please refer to the numbered questio
ns above. Enter as Q1, Q2, Q3, Q4):  Q7, Q8, Q9
Enter the 4th category title (or 'done' to finish): DONE
```

ENTERTAINMENT

**ENTERTAINMENT**

CREDIBILITY

EXPERTISE

| | | | | |
|---|---|---|---|---|
| The food vlogger present the information in cheerful and energetic manner | 379 | 3.53 | 0.54 | Very Hi |
| ENTERTAINMENT OVERALL | 379 | 3.40 | 0.56 | Hi |

## VERBAL INTERPRETATION

**In 'ENTERTAINMENT', the highest scoring question is 'The food vlogger present the information in cheerful and energetic manner' with a mean score of 3.53 (Very High), while the lowest is 'The food vlogger's non-verbal gestures looks real and natural' with a score of 3.18 (High),**

**Overall, this category rates as High with an average score of 3.40,**

### Overall Analysis

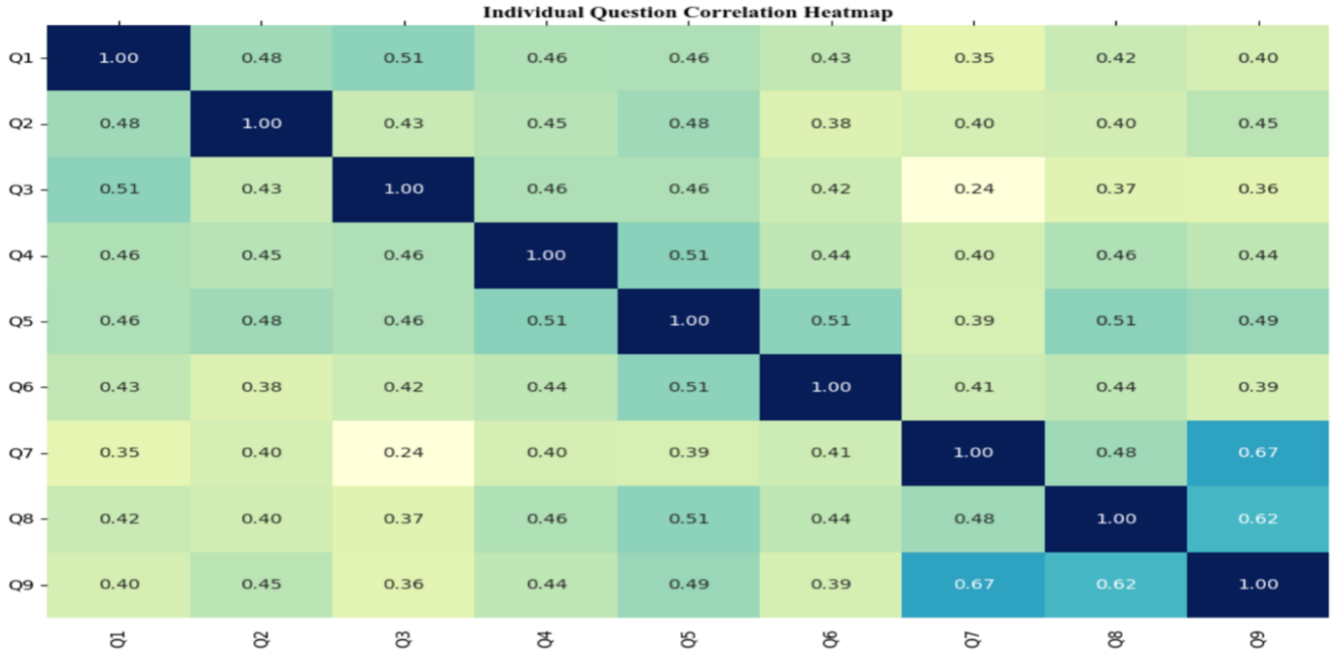
| Category | N | Mean | Std Dev | Interpretation |
|---|---|---|---|---|
| ENTERTAINMENT | 379 | 3.40 | 0.56 | High |
| CREDIBILITY | 379 | 3.33 | 0.60 | High |
| EXPERTISE | 379 | 3.04 | 0.74 | High |
| OVERALL | 379 | 3.26 | 0.63 | High |

## OVERALL VERBAL INTERPRETATION

**In the overall analysis, the highest scoring category is 'ENTERTAINMENT' with a mean score of 3.40 (High), while the lowest is 'EXPERTISE' with a score of 3.04 (High),**

# VII. Correlational Analysis*Implications:*

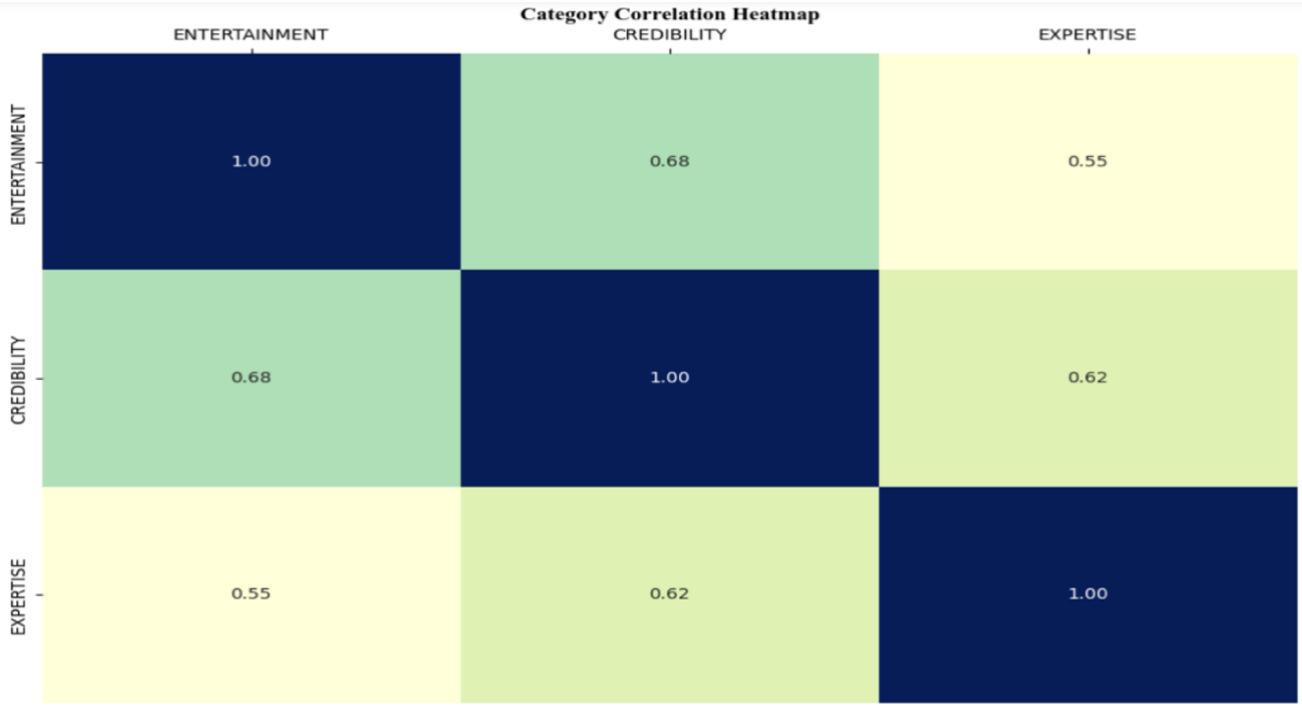
## III. CORRELATIONAL ANALYSIS




## VERBAL INTERPRETATION - Top 3 Strong Positive Relationships

**The variables 'The food vlogger has knowledge about the culture behind the featured foods.' and 'The food vlogger has culinary background.' show a strong positive correlation: 0.67**

**The variables 'The food vlogger has knowledge about the culture behind the featured foods.' and 'The food vlogger present detailed information about Ugbo as food hub destination.' show a strong positive correlation: 0.62**




## VERBAL INTERPRETATION - Top 3 Strong Positive Relationships

**The variables 'ENTERTAINMENT' and 'CREDIBILITY' show a strong positive correlation: 0.68**

**The variables 'CREDIBILITY' and 'EXPERTISE' show a strong positive correlation: 0.62**

**The variables 'ENTERTAINMENT' and 'EXPERTISE' show a strong positive correlation: 0.55**

# VIII. Significant Differences Analysis

## IV. SIGNIFICANT DIFFERENCES BETWEEN PROFILE OF THE RESPONDENTS AND SURVEY RESPONSES

Age

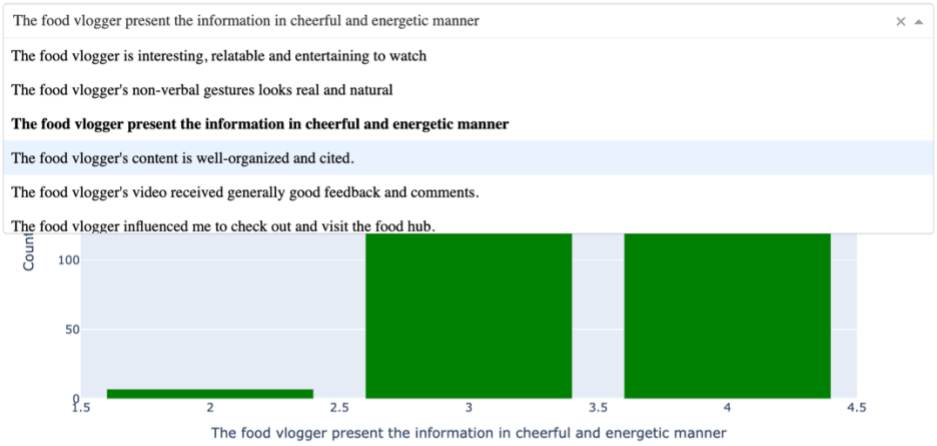
| Category | Demographic | Group | N | Mean | Std Dev | T-Value | P-Value | Significant Difference |
|---|---|---|---|---|---|---|---|---|
| ENTERTAINMENT | Age | 18 to 39 | 352 | 3.39 | 0.4544 | -0.0759 | 0.9395 | No |
| ENTERTAINMENT | Age | 40 to 59 | 4 | 3.08 | 0.5000 | -1.2498 | 0.3000 | No |
| ENTERTAINMENT | Age | 17 or below | 21 | 3.52 | 0.3883 | 1.5109 | 0.1465 | No |
| ENTERTAINMENT | Age | I prefer not to answer | 2 | 3.00 | 0.4714 | -1.1873 | 0.4456 | No |
| CREDIBILITY | Age | 18 to 39 | 352 | 3.33 | 0.4924 | 0.0000 | 1.0000 | No |
| CREDIBILITY | Age | 40 to 59 | 4 | 3.17 | 0.4303 | -0.7746 | 0.4950 | No |
| CREDIBILITY | Age | 17 or below | 21 | 3.38 | 0.3842 | 0.5680 | 0.5764 | No |
| CREDIBILITY | Age | I prefer not to answer | 2 | 3.17 | 0.2357 | -1.0000 | 0.5000 | No |
| EXPERTISE | Age | 18 to 39 | 352 | 3.03 | 0.6379 | -0.0858 | 0.9317 | No |
| EXPERTISE | Age | 40 to 59 | 4 | 2.58 | 0.4194 | -2.1587 | 0.1197 | No |
| EXPERTISE | Age | 17 or below | 21 | 3.16 | 0.5639 | 0.9969 | 0.3307 | No |
| EXPERTISE | Age | I prefer not to answer | 2 | 3.17 | 0.2357 | 0.7836 | 0.5768 | No |

## VERBAL INTERPRETATION

**No significant differences found for the demographic: Age.**

## IX. Individual Question AnalysisVerbal Interpretation:

### V. INDIVIDUAL ANALYSIS FOR ALL QUESTIONS




**VERBAL INTERPRETATION: 'The food vlogger present the information in cheerful and energetic manner**

**The data of 'The food vlogger present the information in cheerful and energetic manner' shows that the option with the highest responses is '4' with 207 respondents, accounting for 54.62% of the total. The calculated mean for 'The food vlogger present the information in cheerful and energetic manner' is 3.53.**