Survey Completion Analysis Report

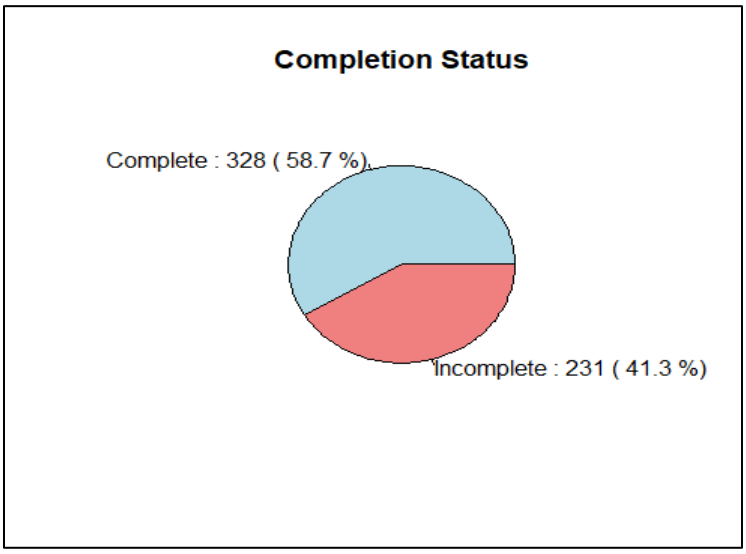
1. Overview of Data Collection

The survey dataset contained responses from a total of 577 entries, including 18 pilot test responses. After excluding the pilot test entries, the final dataset consisted of **559 valid responses** for analysis. The responses were categorized as either **complete** or **incomplete**, with a focus on understanding the completion rates of incomplete cases.

2. Completion Status Breakdown

Out of the 559 valid responses:

- **328 (58.7%)** were marked as **complete**.
- **231 (41.3%)** were marked as **incomplete**.



This indicates that more than half of the respondents completed the survey in its entirety, which is a positive indicator for the overall survey response quality.

3. Analysis of Incomplete Cases

Among the 231 incomplete responses, the analysis revealed the following completion rates:

- **Minimum completion rate:** 39.6%.
- **Maximum completion rate:** 92.2%.

This demonstrates significant variation in how much of the questionnaire was completed by respondents who did not finish the survey.

4. Acceptable Completion Rate Assessment

Research and survey guidelines suggest that acceptable completion rates depend on several factors, including the length of the questionnaire and the target audience. For long questionnaires, an **acceptable completion rate is widely benchmarked at 50% or higher**.

Using this benchmark:

- **133 (52%)** of the incomplete responses had a completion rate of 50% or above.
- **98 (48%)** of the incomplete responses had a completion rate below 50%.

This indicates that slightly more than half of the incomplete responses met the acceptable threshold for completion, even if they did not finish the survey in its entirety.

5. Insights and Recommendations

Based on the analysis, the following key insights and recommendations can be drawn:

1. **Overall Survey Performance:**
    - The survey achieved a **completion rate of 58.7%** for complete responses, which aligns with reasonable outcomes for similar survey efforts and industry standards.
    - The **41.3% of responses marked as incomplete** provides an opportunity to gain additional insights from partially completed data, especially given the observed variation in completion rates.
2. **Incomplete Responses:**
    - Among the incomplete responses, a wide range of completion rates was observed, with progress ranging from **39.6% to 92.2%.**
    - The variation in these rates highlights the potential for meaningful analysis of partially completed responses, particularly those that approached high levels of completion.
3. **Completion Rate Benchmark:**

- Using a benchmark of **50% as an acceptable completion rate** for long questionnaires, **52% of the incomplete responses** met or exceeded this threshold.
- The remaining **48% of incomplete responses** fell below this threshold, but the partially completed data may still offer valuable insights depending on the sections completed.
- Proceeding with the analysis of these responses can help maximize the value of the dataset without necessitating changes to the questionnaire.

6. Conclusion

The survey achieved a respectable completion rate of 58.7% for complete responses and demonstrated that among incomplete cases, a majority (52%) still met the acceptable 50% benchmark.

Annexes: R codes

```
# Set working directory
setwd("C:\\Users\\albert.orwa\\Downloads")
# Load required libraries
library(dplyr)
# Read the dataset
df <- read.csv("TrainingNeeds.csv")
# Exclude the first 18 rows (pilot)
df <- df[-c(1:18), ]
# Create a table of counts
summary_table <- table(df$Completed)
# Print the table
print(summary_table)
# Calculate percentages for the pie chart
percentages <- round(summary_table / sum(summary_table) * 100, 1)
# Create labels for the pie chart (include counts and percentages)
labels <- paste(names(summary_table), ":", summary_table, "(", percentages, "%)")
# Create the pie chart
pie(summary_table,
    labels = labels,
    main = "Completion Status",
    col = c("lightblue", "lightcoral"))
#filter for "Incomplete"
df <- df[df$Completed == "Incomplete", ]
# Assuming the data is in a dataframe called df
# Remove non-response or non-relevant columns
df_clean <- df[,-c(1, 2)]  # Removing RecordID and Timestamp columns
# Define a function to calculate completion rate for each row
calculate_completion_rate <- function(row) {
  # Count non-missing values
```

```
  non_missing <- sum(!is.na(row) & row != "")
  # Calculate the total number of relevant columns (excluding RecordID and Timestamp)
  total_columns <- ncol(df_clean)
  # Calculate completion rate as the ratio of non-missing values to total columns
  completion_rate <- non_missing / total_columns
  return(completion_rate)
}
# Apply the function to each row and create a new column for the completion rate
df$CompletionRate <- apply(df_clean, 1, calculate_completion_rate)
df$CompletionRate <- df$CompletionRate * 100
# Save the updated dataset with completion rates
write.csv(df, "TrainingNeeds_with_completion_rate.csv", row.names = FALSE)
```