Homework Grading Report

Student Name:	Bailey Wright
Assignment:	Assignment 1 - Introduction to R
Graded On:	October 03, 2025 at 05:11 PM
Final Score:	33.1 / 37.5 points (88.3%)

Score Summary

Overall Performance: Good (88.3%)

Instructor Assessment

Your submission meets all technical requirements and you have thoughtfully linked data■cleaning decisions to business scenarios. The reflections are present and show basic critical thinking, though they could be expanded with more depth and quantitative support. Your documentation of ethical considerations is a strong point. For future work, aim to provide more detailed justification for methodological choices and polish the written communication. Overall, solid progress in developing both analytical and business■oriented skills.

Technical Analysis

Code Strengths:

- Successfully implements data quality assessment and missing value treatment
- Uses tidyverse package appropriately for data manipulation and visualization
- Code executes without errors and generates expected outputs for data cleaning tasks

Code Improvement Suggestions:

- Consider using `colSums(is.na(sales_imputed))` instead of `sapply(sales_imputed, function(x) sum(is.na(x)))` for cleaner missing value checking

- Could enhance outlier detection by adding a check for negative values in Quantity column: `sales_df %>% filter(Quantity < 0)`

- Alternative approach for imputing Customer_Name with mode: `sales_imputed$Customer_Name <- names(sort(table(sales_imputed$Customer_Name), decreasing = TRUE))[1]`

Technical Observations:

• Demonstrates solid understanding of data cleaning processes including missing value treatment and outlier detection

• Appropriate use of IQR method for outlier detection in business analytics context

• Code organization supports reproducible analysis with clear section breaks and comments

Additional Code Enhancement Examples:

Data Exploration Enhancement:

```
# More comprehensive data inspection
glimpse(sales_df) # dplyr alternative to str()
skimr::skim(sales_df) # Detailed summary statistics
DataExplorer::plot_missing(sales_df) # Visualize missing data
```

Data Visualization:

```
# Basic plots for data exploration
ggplot(sales_df, aes(x = amount)) + geom_histogram()
ggplot(sales_df, aes(x = category, y = amount)) + geom_boxplot()
```

Data Cleaning:

```
# Handle missing values
sales_df <- sales_df %>%
filter(!is.na(amount)) %>%
mutate(amount = ifelse(amount < 0, 0, amount))
```

Performance by Category