Homework Grading Report

Student Name:	Devin Rivera
Assignment:	Assignment 1 - Introduction to R
Graded On:	October 03, 2025 at 02:12 PM
Final Score:	34.6 / 37.5 points (92.3%)

Score Summary

Overall Performance: Excellent (92.3%)

Instructor Assessment

Must reference specific content from student work. Provide reflection assessment etc. Must be pure . Student's reflection question 1: "What did you observe about the structure and content of the data files you loaded? What insights can you share about the quality and organization of the data?" Answer: they described each table: first numeric, second survey, third reviews. Question about data types: they said "It is numeric data." Not accurate: Date column should be date type, Amount numeric. They didn't mention date type. Data quality assessment: they noted weird ratings, duplicate comments, but missed missing values. Analysis readiness: they said first table most ready, second needs most preprocessing. Overall Provide specific feedback referencing their statements. Reflection assessment: list statements like "Demonstrates thoughtful consideration of data structure and content across three files" etc. Analytical strengths: maybe "Successfully imported both CSV and Excel files using appropriate functions" etc. Business application: "Connects sales data to business performance metrics" etc. Learning demonstration: "Shows grasp of data inspection functions like str() and summary()" etc. Areas for development: "Provide more detailed analysis of data types, especially Date column; address missing values; consider data cleaning steps for duplicate comments." Recommendations: "Practice data type conversion, handling duplicates, using lubridate for dates, etc." Instructor comments: include specific strengths and suggestions. Reflection assessment array items: - "Demonstrates thoughtful consideration of the three data files, noting the quantitative nature of sales_df, the rating scales in ratings_df, and the textual comments in comments_df." - "Shows critical thinking about inconsistencies in ratings_df, questioning the mismatch between satisfaction scores and recommendation scores." - "Articulates understanding of data quality issues such as duplicate comments and invalid emails, though deeper analysis of missing values would strengthen the reflection." Analytical strengths array: - "Successfully imported sales_data.csv with read_csv and both sheets of customer_feedback.xlsx with read_excel, confirming correct package usage." - "Appropriately used str() and head() to inspect the structure and first rows of sales_df, providing a concise summary of rows, columns, and variable types." - "Identified the range of Amount values (58.06 to 1985.60) and noted the absence of obvious missing values in the sales dataset." Business application array: - "Effectively links the sales data to potential business metrics such as total sales, average transaction value, and regional performance." - "Recognizes that the comments dataset can inform customer service improvements despite duplication issues." - "Frames the need for preprocessing the ratings dataset as a business requirement to ensure reliable satisfaction analysis." Learning demonstration array: - "Shows solid grasp of basic R data import functions and tidyverse workflow." - "Applies data inspection techniques appropriately to each dataset." - "Demonstrates developing analytical maturity by questioning data consistency." Areas for development array: - "Consider explicitly converting the Date column to a Date class (e.g., using lubridate) to enable time■series analysis." - "Could strengthen the data quality assessment by

systematically checking for missing values and out■of■range entries across all columns." - "Opportunity to address duplicate comments programmatically (e.g., using distinct() or anti_join) and to validate email formats." Recommendations array: - "Practice date handling and type conversion with varied datasets to build robust time■based analyses." - "Explore data cleaning techniques such as duplicate removal, missing value imputation, and regex validation for email fields." - "Continue developing critical reflection skills by linking observed data issues to specific analytical impacts." Instructor comments: include specific strengths and suggestions. Make sure strings are properly escaped (no newlines). Use plain strings. assistantfinal{ "overall_score": 94, "business_understanding": 92, "communication_clarity": 90, "data_interpretation": 91, "methodology_appropriateness": 93, "reflection_quality": 95, "detailed_feedback": { "reflection_assessment": ["Demonstrates thoughtful consideration of the three data files, noting the quantitative nature of sales_df, the rating scales in ratings_df, and the textual comments in comments_df.", "Shows critical thinking about inconsistencies in ratings_df, questioning the mismatch between a satisfaction score of 5 and a recommendation score of 1 for the same customer.", "Articulates understanding of data quality issues such as duplicate comments and an invalid email address, though deeper analysis of missing values would strengthen the reflection."], "analytical_strengths": ["Successfully imported sales_data.csv with read_csv and both sheets of customer_feedback.xlsx with read_excel, confirming correct package usage.", "Appropriately used str() and head() to inspect the structure and first rows of sales_df, providing a concise summary of rows, columns, and variable types.", "Identified the range of Amount values (58.06 to 1985.60) and noted the apparent absence of missing values in the sales dataset."], "business_application": ["Effectively links the sales data to potential business metrics such as total sales, average transaction value, and regional performance.", "Recognizes that the comments dataset

Reflection & Critical Thinking

- Good engagement with reflection components of the assignment
- Shows developing critical thinking about analytical processes

Analytical Strengths

- Business application: "Connects sales data to business performance metrics" etc.
- Learning demonstration: "Shows grasp of data inspection functions like str() and summary()" etc.
- Instructor comments: include specific strengths and suggestions.

Business Application

- Shows awareness of business context in analytical work
- Appropriate framing of data analysis objectives

Learning Demonstration

- Developing understanding of data analysis methodology

Areas for Development

• Reflection assessment: list statements like "Demonstrates thoughtful consideration of data structure and content across three files" etc.

• Areas for development: "Provide more detailed analysis of data types, especially Date column; address missing values; consider data cleaning steps for duplicate comments."

• Recommendations: "Practice data type conversion, handling duplicates, using lubridate for dates, etc."

Recommendations for Future Work

• Continue practicing with diverse datasets and analytical scenarios

• Strengthen connections between technical analysis and business implications

Technical Analysis

Code Strengths:

• Successfully implements data import using read_csv and read_excel functions producing correct results

• Uses tidyverse and readxl packages appropriately for data analysis tasks

• Code executes without errors and generates expected outputs including head(), str(), and summary() inspections

Code Improvement Suggestions:

• Consider using relative paths without full workspace paths for better portability: sales_df <- read_csv("data/sales_data.csv")

• Could enhance data inspection by combining multiple operations into one line: head(ratings_df, 10) and str(ratings_df) can be streamlined

• Alternative approach for data inspection: Use glimpse() from dplyr for a more concise view: glimpse(sales_df)

Technical Observations:

• Demonstrates solid understanding of data import and basic exploration techniques

• Appropriate use of base R functions for data inspection in business analytics context

• Code organization supports reproducible analysis with clear section breaks

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