Homework Grading Report

Student Name:	Logan Balfour
Assignment:	a1
Graded On:	September 22, 2025 at 02:13 PM
Final Score:	28.8 / 37.5 points (76.7%)

Score Summary

Overall Performance: Satisfactory (76.7%)

Component Scores:

- Working Directory: 2.0 points

- Package Loading: 4.0 points

- Data Import: 11.0 points

- Data Inspection: 8.0 points

- Reflection Questions: 3.8 points

Performance by Category

■ Excellent **Working Directory:** 2.0/2 points (100%)
■ Excellent **Package Loading:** 4.0/4 points (100%)
■ Excellent **Data Import:** 11.0/5 points (220%)
■ Excellent **Data Inspection:** 8.0/8 points (100%)
■ Needs Work **Reflection Questions:** 3.8/12.5 points (30%)

Detailed Analysis:

- Working Directory (2/2 points): Correctly used getwd() and showed output

- Package Loading (4/4 points): tidyverse loaded and executed successfully | readxl loaded and executed successfully

- ■ CSV Import (5/5 points): sales_df variable created with read_csv | Correct filename (sales_data.csv)

- ■ Excel Import (6/6 points): ratings_df created with read_excel | comments_df created with read_excel

- Data Inspection (8.0/8 points): head() used and executed | str() used and executed | summary() used and executed | sales_df properly analyzed | ratings_df properly analyzed | comments_df properly analyzed

- Reflection Questions (3.8/12.5 points): Data Types Analysis (1.6/4 points) ■ Focus on the Date and Amount columns from sales_df - what data types are they? | Good thinking about appropriateness - try connecting this to business needs | You answered the question - try adding more detail next time

Data types matter more than you might think. If your dates are stored as text ("2023-01-15"), you can't calculate time differences or trends. If amounts have dollar signs ("$1,234.56"), you can't do math with them. When I see dates stored properly as date objects, I know you can calculate things like "days between orders" or "monthly sales patterns." When amounts are numeric (1234.56), you can sum, average, and analyze them. This isn't just technical nitpicking - it's about what analysis you can actually do with your data. Check this first, always. It'll save you headaches later. Data Quality Assessment (0.6/4 points) ■ Look at your data outputs - do you see any missing values (NA's) or unusual patterns? | ■ Think about this: how would missing data or errors affect your business conclusions? | Nice response - you could add more specific examples

Look for problems that will mess up your analysis. Missing values can throw off your totals. Inconsistent formatting (like "North" vs "NORTH" vs "north") will split your data when you try to group it. Watch for things that don't make business sense - negative sales amounts, future dates, or someone buying 999,999 keyboards (probably a data entry error). I also want to see you think about impact. If 5% of values are missing, that's different from 50% missing. If you have weird outliers, will they skew your averages? This isn't busy work - bad data leads to bad decisions. Spend time here and your analysis will be much more reliable. Analysis Readiness (1.6/4.5 points) You mentioned the datasets - now compare which is most ready for analysis | ■ Think about what you'd need to do to make the messiest dataset analysis-ready | Good effort - nice reasoning

Compare the datasets and tell me which one you'd start analyzing first. Think practically - which has fewer missing values? Which has cleaner, more consistent formatting? Which one can answer your most important business questions? For example, if your sales data is mostly complete but your feedback data has lots of gaps and messy text, you'd probably start with sales data to get quick insights, then clean up the feedback data later. In real work, you rarely get perfect data. You have to prioritize where to spend your time. Show me you can think strategically about this - it's a key skill. ■ Overall Reflection Quality: Needs Development The reflection questions are where you really develop your analytical thinking skills. Take more time with these - they're not just busy work! Look carefully at your data outputs, think about what you observe, and explain your reasoning. This kind of thinking is what separates good analysts from great ones.

Code Issues & Fixes

Issues Found:

- ERROR: Error: '/workspaces/assignment-1-logan3941/data/sales_data.csv' does not exist.
- ERROR: Error: `path` does not exist: '/workspaces/assignment-1-logan3941/data/customer_feedback.xlsx'
- ERROR: Error: object 'sales_df' not found
- ERROR: Error: object 'ratings_df' not found
- ERROR: Error: object 'comments_df' not found

Specific Code Solutions:

Data Import Fix - CSV File Not Found

**

Working Directory Solutions:
Option 1: If your working directory is set to the data folder:
• Use: read_csv("sales_data.csv") - just the filename
Option 2: If your working directory is the project root:
• Use: read_csv("data/sales_data.csv") - include the data/ folder
Check your setup: Run getwd() to see where you are, then adjust your file paths accordingly.

```
# Check your working directory and file location
getwd() # See where R is currently looking
list.files() # See what files are in current directory
list.files("data/") # See what's in the data folder
# For CSV files, use:
sales_df <- read_csv("data/sales_data.csv")
# NOT: read_csv("../data/sales.csv") or read_csv("sales.csv")
# Make sure:
# 1. File is named exactly "sales_data.csv" (check spelling!)
# 2. File is in a "data" folder in your project
# 3. You're running from the correct working directory
```

Variable Fix - sales_df not found

**

```
# You're trying to use sales_df before creating it
# Make sure you run this cell first:
sales_df <- read_csv("data/sales_data.csv")
# Then you can use it:
head(sales_df)
str(sales_df)
summary(sales_df)
```

Variable Fix - ratings_df not found

**

```
# You're trying to use ratings_df before creating it
# Make sure you run this cell first:
ratings_df <- read_excel("data/ratings_data.xlsx", sheet = "ratings")
# Then you can use it:
head(ratings_df)
```

Variable Fix - comments_df not found

You're making progress. Each assignment builds on the previous one, so nail down these fundamentals.

```
# You're trying to use comments_df before creating it
# Make sure you run this cell first:
comments_df <- read_excel("data/ratings_data.xlsx", sheet = "comments")
# Then you can use it:
head(comments_df)
```

Reflection Questions Feedback

Data Types: 1.6/4 points (Needs Improvement)

Data Quality: 0.6/4 points (Needs Improvement)

Analysis Readiness: 1.6/4.5 points (Needs Improvement)

Next Steps

Nice Progress! (28.8/37.5 points - 76.7%) You're learning the fundamentals well. With some attention to the details below, you'll be ready for more advanced analysis. Here's what to focus on for next time: Reflection Questions: Take more time with these. Look at your data outputs and explain what you see. These aren't just busy work - they help you think analytically. Code Execution: Fix any error messages before submitting. Red error text means something went wrong - don't ignore it. Data Import Fix - CSV File Not Found: ```r # Check your working directory and file location getwd() # See where R is currently looking list.files() # See what files are in current directory list.files("data/") # See what's in the data folder # For CSV files, use: sales_df <- read_csv("data/sales_data.csv") # NOT: read_csv("../data/sales.csv") or read_csv("sales.csv") # Make sure: # 1. File is named exactly "sales_data.csv" (check spelling!) # 2. File is in a "data" folder in your project # 3. You're running from the correct working directory ``` Variable Fix - sales_df not found: ```r # You're trying to use sales_df before creating it # Make sure you run this cell first: sales_df <- read_csv("data/sales_data.csv") # Then you can use it: head(sales_df) str(sales_df) summary(sales_df) ``` Variable Fix - ratings_df not found: ```r # You're trying to use ratings_df before creating it # Make sure you run this cell first: ratings_df <- read_excel("data/ratings_data.xlsx", sheet = "ratings") # Then you can use it: head(ratings_df) ``` Variable Fix - comments_df not found: ```r # You're trying to use comments_df before creating it # Make sure you run this cell first: comments_df <- read_excel("data/ratings_data.xlsx", sheet = "comments") # Then you can use it: head(comments_df) ``` You're making progress. Each assignment builds on the previous one, so nail down these fundamentals.

Study Tips:

- Good foundation! Focus on providing more detailed explanations in reflection questions
- Practice connecting technical concepts to business applications