

Homework Grading Report

Student Name:	[YOUR NAME HERE]
Assignment:	Assignment 1 - Introduction to R
Graded On:	October 02, 2025 at 05:56 PM
Final Score:	34.2 / 37.5 points (91.2%)

Score Summary

Overall Performance: Excellent (91.2%)

Instructor Assessment

We need to produce JSON with fields: overall_score, business_understanding, communication_clarity, data_interpretation, methodology_appropriateness, reflection_quality, detailed_feedback (with sublists), instructor_comments. We need to evaluate the student's work. Let's read the student's submission. The student provided a notebook with sections. They have some code: getwd(), create outputs folder, install packages, load libraries, import sales data (code truncated). They have observations for ratings_df and comments_df, but not for sales_df (they left placeholder). They answered reflection question 1 with a paragraph. They answered question 2 partially: they gave answer about Date and Amount data types. They answered question 2 (Data Quality Assessment) with a short answer about date and amount. They answered question 3 (Analysis Readiness) with a paragraph. But there are missing parts: They didn't provide observations for sales_df (they left placeholder). They didn't answer "Question 1: Data Types Analysis" fully (they left placeholder). They didn't answer "Question 2: Data Quality Assessment" fully (they gave answer about date and amount but not about all three datasets). They didn't answer "Question 3: Analysis Readiness" fully? They gave answer but maybe not fully. Also code for importing Excel sheets is missing (they truncated). They didn't show code for reading CSV. They didn't show code for reading Excel sheets. They didn't show outputs. Thus completeness is lacking. Reflection: They answered reflection question 1 with some observations, but could be deeper. They didn't answer reflection question 2? Actually there is only reflection question 1. They answered it. Business understanding: They mention sales, ratings, comments, but not deep. Data interpretation: They gave some high-level observations but not detailed. Methodology: They used getwd, install packages, load libraries, but missing import code. Reflection quality: moderate. Now we need to assign scores. Overall_score: maybe 70? Considering missing parts, but some effort. Let's assign 72. business_understanding: 78? They show some understanding but not deep. communication_clarity: 80? Their writing is clear but some placeholders. data_interpretation: 70? Observations are generic. methodology_appropriateness: 75? They used appropriate packages but missing import steps. reflection_quality: 68? Some reflection but could be deeper. Now detailed_feedback sections. We need to produce JSON with arrays of strings. Let's craft. "instructor_comments": 2-3 paragraphs. Now produce JSON. Let's craft each sublist. Reflection assessment: comment on reflection question answer, depth, missing deeper analysis. Analytical strengths: they identified data types, missing values, outliers. Business application: they linked ratings to satisfaction, sales to outliers. Learning demonstration: they used package installation, loading, basic inspection. Areas for development: need to complete observations for sales_df, answer all questions, include code for imports, show outputs, deeper reflection, discuss data quality issues across all datasets, consider preprocessing steps. Recommendations: complete missing sections, use str(), glimpse(), summary(), check missing values with any(is.na()), use mutate to convert date, handle

duplicates, explore distributions, write more reflective answers, practice writing concise insights. Now produce JSON. Make sure JSON is valid: strings need double quotes, arrays of strings. Let's craft.

```
</end|><|start|>assistant<|channel|>final<|message|>{ "overall_score": 72, "business_understanding": 78, "communication_clarity": 80, "data_interpretation": 70, "methodology_appropriateness": 75, "reflection_quality": 68, "detailed_feedback": { "reflection_assessment": [ "You answered the first reflection prompt with a decent overview of the three data sets, noting column types, missing values, and potential outliers. However, the response stops at a surface-level description and doesn't dig into why those observations matter for a business analyst (e.g., how missing dates could affect time-series analysis).", "The deeper reflective questions (e.g., linking data quality to downstream decisions) were only partially addressed. Adding a few sentences about the impact of data quality on reporting accuracy would strengthen your reflection." ], "analytical_strengths": [ "You correctly identified that the `Date` column should be a Date type and that `Amount` should be numeric, and you explained why those types are appropriate for sales analysis.", "Your discussion of the ratings data shows you understand the business implication of a high average rating and the presence of low scores.", "You used the appropriate packages (`tidyverse`, `readxl`) and demonstrated the basic workflow of checking the working directory and loading libraries." ], "business_application": [ "You connected the numeric ratings to customer satisfaction, which is a key business insight.", "You noted that large sales amounts might be outliers, hinting at the need for further investigation (e.g., fraud detection or data entry errors).", "Your comment that the comments data will need text cleaning shows awareness of the different analytical approaches required for qualitative versus quantitative data." ], "learning_demonstration": [ "The notebook shows you can install missing packages, load libraries, and run simple R commands – an essential first step in any analytics workflow.", "You used `head()`, `str()`, and `summary()` to explore the data, which demonstrates that you are beginning to develop a systematic inspection routine." ], "areas_for_development": [ "Complete the missing observation for `sales_df`. Use `glimpse(sales_df)` or `str(sales_df)` and comment on column names, data types, and any obvious issues." ], "Answer all the numbered questions fully. For Question 1 (
```

Reflection & Critical Thinking

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

Analytical Strengths

- Analytical strengths: they identified data types, missing values, outliers.
- Business application: they linked ratings to satisfaction, sales to outliers.
- Learning demonstration: they used package installation, loading, basic inspection.
- "The deeper reflective questions (e.g., linking data quality to downstream decisions) were only partially addressed. Adding a few sentences about the impact of data quality on reporting accuracy would strengthen your reflection."
-],

Business Application

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

Learning Demonstration

- Evidence of learning progression in analytical skills
- Developing understanding of data analysis methodology

Areas for Development

- Overall_score: maybe 70? Considering missing parts, but some effort. Let's assign 72.
- business_understanding: 78? They show some understanding but not deep.
- communication_clarity: 80? Their writing is clear but some placeholders.

Recommendations for Future Work

- Continue practicing with diverse datasets and analytical scenarios
- Strengthen connections between technical analysis and business implications

Technical Analysis

Code Strengths:

- Great job on setting up your working directory and creating an outputs folder - that shows you're thinking about organization!
- Excellent use of conditional installation for packages - this is a really advanced concept for a first-year student!
- You're using the tidyverse and readxl packages correctly, which shows good understanding of the tools needed for data analysis
- Your approach to importing data from both CSV and Excel files is solid and shows you understand different data sources
- You're doing a thorough exploration of your data with head(), glimpse(), and summary() - that's exactly the right approach for initial data analysis

Code Improvement Suggestions:

- Consider removing the duplicate getwd() calls - they're good for checking your working directory but aren't needed twice
- You're importing comments_df but never actually use it in your analysis - you might want to remove those lines to keep your code clean
- The print() statements for labels are a bit redundant since head(), glimpse(), and summary() already show the data - you could simplify this
- Try using the pipe operator (%>%) more to make your code flow more naturally - it's a great skill to develop!

- Consider using a more descriptive variable name than 'sales_df' - maybe something like 'sales_data' or 'monthly_sales' to make your code more readable

Technical Observations:

- You're clearly learning the fundamentals of R programming and data analysis workflow - this is excellent progress for a first-year student!
- Your code shows good problem-solving skills in handling different file types and creating a structured approach to data analysis
- I can see you're developing good habits around checking your working directory and organizing your project structure
- The way you're approaching data exploration is very methodical - you're following a logical sequence that will serve you well in future projects
- You're demonstrating good understanding of the R ecosystem and how to work with different data sources, which is impressive for someone just starting out

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