

NUM2 – NUM2 TASK 1: DATA CLEANING

DATA CLEANING – D206

PRFA – NUM2

TASK OVERVIEW

SUBMISSIONS

EVALUATION REPORT

EVALUATION REPORT – ATTEMPT 1 – PASSED

Overall Evaluator Comments

EVALUATOR COMMENTS

Your submission provided some excellent data cleanup and principal component analysis visuals. These will help tell the data story of the process works.

A. Question or decision

Competent The description states a question or decision that can be addressed through analysis of the chosen data set. The question or decision is relevant to a realistic organizational need or situation.

There are no comments for this aspect.

B. Required Variables

Competent The description includes the variables in the data set and indicates the specific type of data being described, and includes examples from the data set to support claims.

There are no comments for this aspect.

C1. Plan to Find Anomalies

Competent The proposal includes a detailed description of the techniques and steps needed for identifying anomalies in the selected data set.

There are no comments for this aspect.

C2. Justification of Approach

Competent The justification includes the characteristics of the data being assessed and references the approach used to assess the quality of the data. The justified approach aligns with the selected data set.

There are no comments for this aspect.

C3. Justification of Tools

Competent The justification describes the benefits of using the programming language, including any libraries and packages used to clean the data, and includes specific examples of how these tools are ideal in this scenario as opposed to other available tools.

There are no comments for this aspect.

C4. Provide the Code

Competent The submission provides the complete and executable code, which could be used to identify anomalies in the data set.

There are no comments for this aspect.

D1. Cleaning Findings

Competent The description accurately includes *all* of the anomalies found by running the code from part C4.

EVALUATOR COMMENTS: ATTEMPT 1

The submission includes an adequate description of the code used to clean the anomalies.

D2. Justification of Mitigation Methods

Competent The justification includes the specific mitigation methods for each type of anomaly listed in part D1.

There are no comments for this aspect.

D3. Summary of the Outcomes

Competent The summary details the outcome from the implementation of each data-cleaning step. The summarized expected outcomes are plausible given the interventions.

There are no comments for this aspect.

D4. Mitigation Code

Competent The submission provides complete and executable code that could be used to mitigate the anomalies.

There are no comments for this aspect.

D5. Clean Data

Competent The submission includes a clean data set created from the raw data. The provided data set includes the complete list of variables from the chosen data set in part A.

There are no comments for this aspect.

D6. Limitations

Competent The submission accurately summarizes the limitations of the implemented data-cleaning process.

There are no comments for this aspect.

D7. Impact of the Limitations

Competent The submission includes a discussion of the impact of the limitations from part D6. The discussion logically aligns with the question or decision from part A.

There are no comments for this aspect.

E1. Principal Components

Competent The submission lists *all* principal components of the data set.

There are no comments for this aspect.

E2. Criteria Used

Competent The description of how the principle components of the data set were identified is accurate and complete.

There are no comments for this aspect.

E3. Benefits

Competent The description of how the organization can benefit from the results of the PCA is logical and accurate.

There are no comments for this aspect.

F. Video

Competent The Panopto video recording demonstrates the warning-and error-free functionality of the code used to support the discovery of anomalies and the data cleaning process. An accurate summary of the programming environment is provided in the video.

There are no comments for this aspect.

g. Sources for Third-Party Code

Competent The submission records all web sources used to acquire data or third-party code and all of the web sources are reliable.

There are no comments for this aspect.

H. Sources

Competent The submission includes in-text citations for sources that are properly quoted, paraphrased, or summarized and a reference list that accurately identifies the author, date, title, and source location as available.

There are no comments for this aspect.

I. Professional Communication

Competent Content reflects attention to detail, is organized, and focuses on the main ideas as prescribed in the task or chosen by the candidate. Terminology is pertinent, is used correctly, and effectively conveys the intended meaning. Mechanics, usage, and grammar promote accurate interpretation and understanding.

There are no comments for this aspect.