# DAT 325 Project Three Template

# Data Validation

Replace the bracketed text with the relevant information.

## File Names Table

Perform the following data validation on each of the import, existing, and merge data tables.

* The COUNT of rows in each table. You should reconcile these counts to make sure the number of rows you inserted made it into the merged table.

|  | **Excel Files** | | | | |
| --- | --- | --- | --- | --- | --- |
|  | **Source File** | **Anomalies** | **Import Data** | **Existing Data** | **Merge Data** |
| **Count** | [Insert text.] | [Insert text.] | [Insert text.] | [Insert text.] | [Insert text.] |

* Use the identified variable for **your chosen data set** as listed below to determine MIN, MAX and AVERAGE for each table:

**Data Set:** Twitter

**Variable:** Airline Score by Airline name

**Data Set:** Human Resources

**Variable:** Pay Rate by Department

**Data Set:** Supermarket

**Variable:** Gross Income by location

Determine the MIN, MAX, and AVERAGE in your uploaded table, the MIN, MAX, and AVERAGE in the existing table before you inserted the new rows, and the MIN, MAX, and AVERAGE in the existing table after you insert new rows.

|  | **Excel** | | |
| --- | --- | --- | --- |
|  | **Import Data** | **Existing Data** | **Merge Data** |
| **MIN** | [Insert text.] | [Insert text.] | [Insert text.] |
| **MAX** | [Insert text.] | [Insert text.] | [Insert text.] |
| **AVERAGE** | [Insert text.] | [Insert text.] | [Insert text.] |

## Summary

Describe how the distribution data tells you if the import changes the existing distribution.

[Insert text.]

Why do the validation steps help to ensure that the final import data is clean and of high quality?

[Insert text.]