# DAT 325 Project Two Template

# Executive Summary Report

Human Resources Data

## Data Set Anomalies

The first row is an example.

| **Key Value** | **Description of Anomaly** | **Plan for Resolution** |
| --- | --- | --- |
| *Example:*  *101339052 (Employee Number field)* | *Example: Last name missing* | *Example: Delete row* |
| 1106026572 - Employee Number Field | Missing either first/last name  (One present) | Delete Row |
| 1312063507 - Employee Number Field | Missing Marital Status (Not Null) | Delete Row |
| 1104025243 – Employee Number Field | Missing Marital Status (Not Null) | Delete Row |
| 1308060959 – Employee Number Field | Negative Pay Rate Value (Logical Error) | Delete or Edit Issue if all that is needed is to remove “-” |

## Data Types

You may need to add rows. The first three rows provide example answers. You will need to assess all header names from your chosen data set.

| **Header Name From File** | **Data Types Note** |
| --- | --- |
| *Example: State* | *Example: OK* |
| *Example: DOB* | *Example: Won’t Use* |
| *Example: Dates Employed* | *Example: Needs to be rounded* |
| Employee Number | OK |
| Employee Name | Separate Into First/Last Name Fields |
| Age | Not needed |
| Pay Rate | OK |
| State | OK |
| Zip | OK |
| DOB | OK |
| MaritalStatus | Change to INT to assign status |
| Sex | Change to INT to assign gender |
| CitizenDesc | Change to INT to assign citizenship |
| Hispanic/Latino | Not needed |
| RaceDesc | Change values to INT to look up in RaceDesc Dictionary |
| Date of Hire | OK |
| Days Employed | Not Needed |
| Date of Termination | OK |
| Reason For Term | OK |
| Employment Status | OK |
| Department | OK |
| Position | OK |
| Manager Name | Not Needed |
| Employee Source | Not Needed |
| Performance Score | Not Needed |

## Specific Transformations Needed to Join the Data

You may need to add rows. The first row is an example. Responses may have one to three Excel functions.

| **Header Name From File** | **Excel Function One** | **Excel Function Two** | **Excel Function Three** |
| --- | --- | --- | --- |
| *Example: Smoker Status* | *Example:*  *IF(A1:A52=”Yes”,1,0)* |  |  |
| MaritalStatus | IF(H2:H106=”Married”,1,0) | IF(H2:H106=”NotMarried”,0,1) | IF(H2:H106=”N/A”,2,0) |
| Sex | IF(I1:I106=”Male”,0,1) | IF(I1:I106=”Female”,1,0) | IF(I1:I106=”N/A”,2,0) |
| CitizenDesc | IF(J2:J106=”Non-Citizen”,0,1) | IF(J2:J106=”US Citizen”,1,0) | IF(J1:IJ06=”N/A”,2,0) |
| RaceDesc | Need Values for each race to create functions |  |  |

## Executive Summary

In order to combine these data sets, we need to ensure all the data values and types match between the two tables. When analyzing the data, we learned of a few data values that were not needed to match with the data from Bruce Enterprises. The Wayne Enterprises data values that were not needed in the transfer were, Age, Hispanic/Latino, Days Employed, Manager Name, Employee, Source, and Performance Score. On top of that a few of the data types for a few of the values needed to be altered using excel functions. The MaritalStatus, Sex, CitizenDesc, and RaceDesc all needed to have their character values changed to integers to assign their values. As for RaceDesc, to complete this we need to acquire the Lookup table to get the proper value for each race. All the other name data values are given a 0,1,2 key to assign their values. Doing this will ensure that our data is clean and accurate and will be able to efficiently and effectively join the two data sets.