**TrueNorth Data Scientist Assignment**

You are currently employed as a data scientist at FalseSouth.

FalseSouth is in the transportation industry, and we help drivers book and find loads with brokers as well as take care of drivers’ back office operations. FalseSouth usually gets paid directly by brokers for the service of hauling loads.

However, FalseSouth is having trouble assessing if our collections process has been successful and how much money we are receiving from the brokers. Therefore, the FalseSouth team has assigned to you the task of digging into the accounts receivable data to help with the collections.

Attached is a dataset with all of FalseSouth’s invoicing data with the following columns: ● **COMPANY\_NAME**: This is the Marvel broker company name

● **LOAD\_NUM**: This is the unique ID from the Marvel broker for them to identify loads that are being hauled.

● **INVOICE\_ID**: This is FalseSouth’s unique ID to identify invoices we sent to brokers for payment of loads hauled.

● **INVOICE\_DATE**: This is the date the invoice was sent out to the broker. ● **DUE\_DATE**: This is the date the invoice is due for payment.

● **BALANCE**: This is the amount on the invoice that needs to be paid by the broker. ● **TOTAL\_AMOUNT**: This is the amount originally invoiced to the broker.

Please address the following questions and deliverables:

● Please identify and list out any potential data errors and/or issues.

● How much has FalseSouth invoiced to all brokers?

● How much has FalseSouth collected?

● What balance remains to be collected? How much of that is overdue? ● Please create a chart by month showing invoiced amount split by paid vs. not paid ● Please provide some descriptive statistics with regards to invoicing data provided. In your summary statistics, please answer the following questions:

○ What is the median invoice amount?

○ What is the mean/median/mode number of invoices per broker?

● To support our collections effort, please create a separate statement for each of the brokers of all their invoices and balances. A sample template for the statement has been provided in the attached Excel. (Hint: Please create a script that generates the statements as opposed to manually creating them. We are looking to assess your technical skills in the completion of the assignment deliverables.)

Please provide the following when turning in the assignment:

● R or Python script which shows all of your work

● PDF, word document, or any other document form of your choice which shows the outputs from the above questions (excluding the statements; refer to next bullet) ● Folder with all of the statements to the brokers