DATA-332-02 Applied Business Analytics (20243SP)

**Whole Game Project**

Exercising the skills learned in the first eight chapters of R for Data Science, Hadley et. Al (2e)

Working with the data in R requires practice with purpose. This exercise will help you articulate and execute a plan based on a set of data of which you have no experience.

The SOAR Methodology will be the framework for you to work and communicate your efforts to complete the project

**Specify the Question (each question is 5 points)**

Question(s) to be answered

“Are shorter Delivery Times actually more expensive?”

“How many shipments are completed by each Carrier?”

"What is average cost per shipment from each Carrier?”

“What is cost per KG-KM for all Carriers?”

“What is the rank for each Carrier by Cost per KG-KM?”

“How does Longest Diagonal affect Cost?”

In three paragraphs answer these questions:

What is your plan? **(5 points)**

What are the steps you will take to answer the questions with data and report the results? **(5 points)**

**Obtain the Data**

**Import**

For each for the thirteen columns in the data articulate your understanding of the data. Prepare a dictionary for review with the source of the data to verify your understanding.

Your data dictionary should include answers to these questions for each column of the data:

Column Title as obtained:

Data Type:

Statistical Summary:

Missing and Outlier Data:

**Analyze the Data**

**Tidy**

Assign the data types to each column

**Transform**

Utilize each function listed below when transforming rows of your data:

Filter, Arrange, and Distinct

Utilize each function listed below when transforming columns of your data:

Mutate, Select, Rename, Relocate

**Report the Results**

**Visualize**

Present your final data table used in the analysis (the outcome of Import, Tidy, and Transform Steps)

Answer each specific question with supporting data and visualizations