ISM 465: Data Acquisition and Analytics Development
Mastery Assignment
300 points

1. **Data Analytics:**

Some business analysts and managers spend a majority their time finding solutions to problems. Sometimes, they even jump the gun by solving the problem before they know what the problem is.

   a. Albert Einstein said, "if I had an hour to solve a problem I'd spend 55 minutes thinking about the problem and five minutes thinking about solutions". What potential dangers are the business analysts and managers involved with?

   b. Do you agree with the quote from Albert Einstein? Before answering, consider what we learned about data analytics so far. Please discuss how this quote may help you understand how to preform data analytics correctly and avoid some potential dangers or mistakes.

2. **Mini Case Study:**

A bike company recently designed a new mountain bike. In order to promote this new bike, this company ran a three-month marketing campaign on Facebook from March 2016 to May 2016 in the United States. Four social marketing methods were used in this campaign.

   - First, a product page was created on Facebook, and FB users can "Like" the page and then receive updates from the page in their news feed.
   - Second, a FB group was created to allow group members to have high levels of engagement.
   - Third, a short video about this bike was created and then posted on the company page and the product page.
   - Finally, a $100 off digital coupon was shared with a few targeted customers and allowing them to share the coupon with their friends.

In this campaign, a FB user can enter a sweepstakes lottery session to win a bike by clicking "Like" the product page, posting comments on the product group page, sharing the product page with friends, or inviting friends to the product group. In addition, some target users can also share the $100 off digital coupon on Facebook in the first three weeks.

This is the first time that this company ran a marketing campaign on Facebook. A marketing manager in the bike company was asked to evaluate the performance of the marketing campaign. Based on the evaluation, the manager was required to decide whether this campaign was successful in terms of promoting the new bike, increasing the bike sales, and enhancing the company brand. In addition, the manager needed to justify the conclusions. All the findings will be reported at the executive meeting.

   - Based on the mini case given above, please answer the following questions (a-c):

   a. What specific variables could be used to evaluate the performance of this marketing campaign? Please list at least three variables (e.g. "likes") and briefly explain why you selected each of them.

b. Please indicate the unit of analysis (UoA) (see table below) in the following three scenarios and then answer the questions. (e.g. A researcher wants to investigate how knowledge sharing in supply chains impact their performance, UoA: supply chain)

- Are the units of analysis in the three scenarios the same? Why?
- Please briefly discuss why it is important to understand the unit of analysis in analytics.

Scenario	Unit of Analysis
This manager wants to compare the effectiveness of different social marketing methods in promoting the new bike.	
The manager wants to know which types of customers are likely to purchase the new bike.	
The manager finds that a person's tier 1 network (e.g. all friends that a person is directly connected to) on Facebook is very important for promoting the new bike. Therefore, he wants to investigate why certain networks have higher sharing rates.	

c. In addition to the data generated by Facebook, what other data will you recommend this manager to collect? Please provide one additional data source and briefly explain why this data source is needed. (e.g. think big data).

3. **Data Integration:**
    Instructions:
    - Be sure to type each query (e.g. customername) versus a direct copy and paste into the SQL tool. When a query is copied and pasted, it may not work due to a format issue. If you do not know what each statement means, please refer to https://www.w3schools.com/sql/default.asp
    - For each question, copy and paste both the query and screen shot of the results page (see example in question one).

SQL Statement:

```
SELECT
Firstname, Lastname
FROM Employees
```

Your Database:

Tablenames	Records
Customers	91
Categories	8
Employees	10
OrderDetails	518
Orders	196
Products	77
Shippers	3
Suppliers	29

Edit the SQL Statement, and click "Run SQL" to see the result.

Run SQL »

Result:

Nancy	Davolio
Andrew	Fuller
Janet	Leverling
Margaret	Peacock
Steven	Buchanan
Michael	Suyama
Robert	King
Laura	Callahan
Anne	Dodsworth
Adam	West

   a. **Show first names and last names of employees in the database. See example below.**

   b. **Find the average price for each supplier.**
   c. **Find supplier with an average price greater than $30.**
   d. **Find suppliers with an average price greater than $30 with supplier IDs greater than four. Then sort the SupplierIDs in descending order.**
   e. **Pull all the attributes of the product with product ID greater than 6, categoryID equal to 3, and price greater than $15 from the products table.**
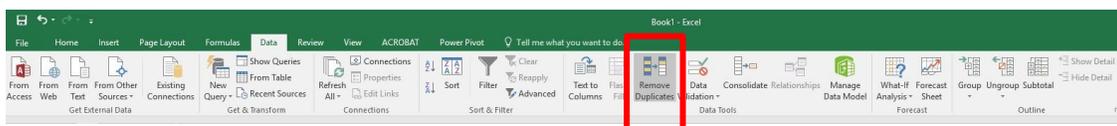
4. Manipulating data in excel.

Find the Excel file "Northwinds" under Unit 6: Assigned readings and preparation activities. Identify duplicated values in this dataset using Pivot Table and remove duplicated data using the function "Remove Duplicates" in Excel. Please answer the following questions:

Insert PivotTable:
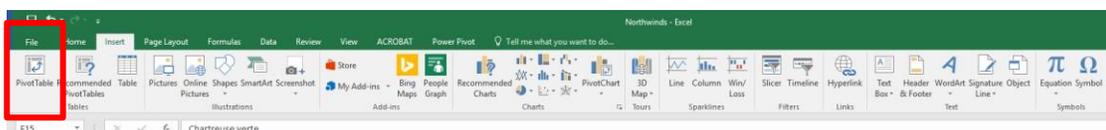
Remove Duplicates:

   a. One record appears three times in this dataset. Indicate the ID number of this record.
   b. How many records appears twice in this dataset?

    c.   How many duplicated records are removed from the dataset after you use the function "remove duplicates" in Excel?

5. Data Transformation & Basic Analysis:

Open and download the Excel file (zipped file) "Amazon_ratings_Musical_Instruments" under Unit 6: Assigned readings and preparation activities. It may take a while because this dataset includes 500,176 records. This dataset holds ratings from online reviews in the category of musical instruments on AMAZON. It includes four attributes:
1. Reviewer ID
2. Product ID
3. Rating Star (from 1 to 5)
4. Standard Time

After completing this data transformation, insert a Pivot Table to answer a few questions:

1. Which year generated the highest average ratings of online reviews in the musical instrument category? What is the average (round to two decimal places)?
2. Which year generated the largest number of online reviews in the musical instrument category? How many reviews did this year generate (please type the number here, please do not include any symbols such as commas)?
3. In the musical instrument category, which product had the highest number of online reviews? Copy and Paste the product ID here. How many online reviews did this product receive (please type the number here, please do not include any symbols such as commas)?
4. In year 2012, which product received the highest number of online reviews? How many reviews did it receive?