Assignment 1

Part 1: Exercises for Tableau data connection, dimension, trend line, annotation and dashboard

- Connect to the BeerAndPretzels dataset (extract the data into Tableau).
- Create a scatter graph that plots Non-Pretzels against Pretzels. (The vertical - or "Y" - axis should be Non-Pretzels. The Horizontal - or "X" - axis should be Pretzels.) Each point on the plot should be depicted by a solid circle.
- Create a second Worksheet with a scatter plot that is identical to the plot created in Step 2. Set the size of each circle to depict the value of Beer.
- Create a third Worksheet with a scatter plot that is identical to the scatter plot created in Step 3. Set the color of each point on the plot to depict Region.
- Create a fourth Worksheet in the Tableau workbook that plots Pretzels (Y) against Beer (X) and Non-Pretzels (Y) against Beer (X) in the same sheet.
- Add each of the four Worksheets to a single dashboard in the order in which they were created.
- Annotate your dashboard to describe the relationship between Pretzels and Non-Pretzels, Pretzels and Beer, and Non-Pretzels and Beer that is depicted in the dashboard contents. Highlight whether the relationship between Beer and Pretzels/Non-Pretzels is depicted more clearly in Worksheet 2 (Step 3 above) or Worksheet 4 (Step 5 above). Express whether there is any apparent difference in the relationship between Non-Pretzels and Pretzels across Regions (Worksheet 3, Step 4). Note: It is import to show the trend line when analyzing the relationship.

Part 2: Distribution of continuous and categorical variables, Tableau Story

- Extract the Titanic dataset into the same Tableau Workbook. (Worksheet -> Data -> add data source) Exclude any rows that are missing the Age of the ship passenger (Add filter by the add button in the top-right of the data panel).
- Create a new column in the Titanic dataset that discretized the Age variable, with Age values > 16 reflected as "Adult" and those not > 16 as "Child" (Group the data, or do it in Excel by using the IF function). Please carefully handle the null variables.
- Create a histogram that depicts the distribution of passengers' ages.
- Create a Mosaic graph (Treemap in Tableau) that depicts the total distribution of passengers (CNT) across gender, cabin class, and age group.
- Create a Mosaic graph that depicts the total distribution of survivors (SUM) across gender, cabin class, and age group.
- Create a Mosaic graph that depicts the *number* of passengers from each combination of gender, cabin class, and age group as a percent of all passengers ("Quick Table Calucation" -> "Percent of Total") on board the Titanic. (Show percentage as the label)
- Create a Mosaic graph that depicts the *percentage* of passengers that survived from each combination of gender, cabin class, and age group as the survival rate (consider using AVG(survived) to show the survival rate. You could also do some math to prove it if you are interested.) for each of those combinations of metrics. (Show percentage as the label)

- Using each visualization (Mosaic graph) of the Titanic dataset, create a Tableau story that describes the separate distribution of passengers (may consider using CNT(survived)) across characteristics gender, cabin class, and age group; also include the visualization that describes the survival experience of Titanic passengers across combinations of gender, cabin class, and age group.

Part 3

- Save the above (including the BeerAndPretzels data extract and all graphics and the Titanic dataset and all graphics) in a single Tableau Packaged Workbook (with a **.twbx** file extension) and submit on Blackboard. (Note: Wrong format would lead to 10 points deduction)