Washington University in St. Louis
Olin Business School

DAT 560E Data Visualization for Business Insights

Homework Assignment 1 (SP2024-MiniA)

Due Date: TBD

Please note that no collaboration with other students will be allowed when working on this assignment.

(120 points)

1. (79 points) Use the Tableau tool to connect to the "Global Superstore 2024A" data.
   (a) Create a Text Table showing Sales, Profit, Quantity, and Shipping Cost by the following dimensions: Category and Sub-Category. (6 points)
   (b) Create a Treemap to show the Profit by the Order Priority and the Region. (6 points)
   (c) Create a pie chart for displaying the Profit (angle) by the Market. Add labels for Market and Profit inside every slide. (6 points)
   (d) Create a Filled Map to show the Profit by Country. (7 points)
   (e) Create a rounded horizon Bar chart to display the Profit by the Market. Add labels for Profit to the end of bars. Add the Market to the Color Marks card. (7 points)
   (f) Create a funnel chart to display the Sales by Region. (7 points)
   (g) Create a Symbol Map to show the Sales by World Cities. Change the map background. (7 points)
   (h) Create a Highlight Table to display the Profit by Segment and Market. (6 points)
   (i) Create a 100% Bar chart to display the Sales by Region (colored by Order Priority). (7 points)
   (j) Create an Area chart showing the Market and the Sales by the Quarters from 2020 to 2023. (6 points)
   (k) Create a Packed Bubble Chart to show the Sales by Country with different colors for the Market. (6 points)
   (l) Create a Line (continuous) Chart to show the Profit by Month (Order Date) and by Market. Add a parameter and a corresponding calculated field for highlighting a market. (8 points)

2. (14 points) Use the Tableau tool to connect to the "MLB-Player-Stats-Batters-2019-2023.xlsx." Make a union of the three tables from 2021 to 2023.
(a) Create a horizon bar chart showing all players' Batting Averages (AVG) in the 2021-2023

seasons for players with an average number of games per year (G per year) played at least 100 (i.e., AVG(G) >= 100). Add a text column to display their corresponding Ranks using the Rank function. Hint: AVG = H/AB. H=hits, AB=at bats. (7 points)
(b) Create a histogram for the HR per year by players with at least an average AB of 330 annually. Add a corresponding parameter for setting up the bin size of the histogram. (7 points)
Data Source: https://www.rotowire.com/baseball/stats.php

3. (7 points) Use the Tableau tool to connect to the "Data for Gantt Chart.xlsx" and create a Gantt Bar Chart. Add a reference line to show the current date.

4. (20 points) Use the Tableau tool to connect to the "DSNY_Monthly_Tonnage_Data.xlsx." Use only 2013-Present (2023) data for the following problems:

(a) Create a calculated field for recycled tons collected (Paper and MGP). Create a calculated field for the Recycle to Refuse ratio. Create a horizontal Bar Chart for the Recycle to Refuse ratio by the borough and the Community District (Do not create a combined field for these two dimensions). Add a text column showing the Rank for each borough & Community District using the table calculation (Rank). Add another text column showing the Rank for each borough & Community District using a calculated field with a RANF function for the Rank. (7 points)

(b) Create a combined field for the borough and community district. Redo (a) using the Rank function. (6 points)

(c) Create a new calculated field with the following code:
"["+STR([Borough])+", "+STR([Communitydistrict])+"]"
Create a Word Cloud for the Recycle to Refuse ratio using this newly created calculated field. (7 points)