**HR Survey Analysis project**

**THE ASSIGNMENT**

To analyze the survey response data, and prepare a visual summary for HR leadership team.

**THE OBJECTIVES**

1. Explore and profile the data to correct any quality issues
2. Prepare and reformat the data for visualization
3. Visualize the data and identify key insights and recommendations

**DATA SOURCE**

Maven Analytics, Survey response from ~1500 employees in Washington state. File type: Excel; data structure: single table; # of records: 14725; *# of fields*: 10.

*PROGRESS OF THE PROJECT*

Objective 1

Identification of the seasonal patterns. The first objective is to calculate the number of collisions by month and year and visualize them using line charts.

The following tasks were set to achieve the objective:

1. Calculation of the count of 'Collision ID' by year and month, and visualizing the results using a line chart
2. Filtering out the incomplete month of April 2023
3. Modifying the line chart so that each year is shown as a separate line
4. Applying formatting to finalize the chart and summarize insights from the analysis

Objective 2

Visualize weekly trends. The second objective is to calculate the number of collisions by time of day and day of week and visualize the data using a heatmap.

The following tasks were set to achieve the objective:

1. Extracting the Weekday and Hour from the 'Date' and 'Time' columns
2. Calculating the count of 'Collision ID' by Weekday and Hour
3. Creating a heatmap to visualize collision hotspots by time of day and day of week
4. Modifying the heatmap to a "white-white-red" 3-color scale to keep focus on the most dangerous periods in the week and summarize the insights from your analysis

Objective 3

Analyze contributing factors. To find the top 10 contributing factors by number of collisions and calculate the percentage of the collisions involving injuries or fatalities.

The following tasks were set to achieve the objective:

1. Calculating the count of 'Collision ID' by 'Contributing Factor'
2. Filtering the top 10 contributing factors by collisions and sorting them in descending order
3. Calculating '% of Dangerous Collisions' for each contributing factor by taking the number of collisions with an injury or fatality and dividing them by the total
4. Adding data bars to the '% of Dangerous Collisions' values to visualize the results, and summarize the insights from your analysis

**INSIGHTS**

The key insights of the project are:

1. In 2022, the number of collisions saw a gradual reduction
2. The highest number of collisions occurs on working days, between 3 p.m. and 6 p.m.
3. The contributing factors with the highest number of fatalities and injuries are: Failure of Yield Right-of-Way (64%) and Traffic Control Disregarded (59%)