School of Science, Computing and Engineering Technologies

COS30045

LAB 4.1 Design Studio

Overview

In this lab you will be given a sample data set and asked to identify the different data and attribute types. You will also think about some questions about this data set that might be answered by a visualisation.

ardd\_fatalities\_Jan2020\_0.xlsx (download from Canvas)

Download and review this data set before attempting this exercise.

1 Interpreting the data set

Complete the LAB 4.1 Quiz.

2 Visualisation Design

Think of three questions you would like to answer with that require a data visualistion.

For each data question you will need to consider the following:

Which data attributes (columns) do you need to answer this question?

Do you need to transform any of the data?

Does the data type change when you transform the data? If so how.

Make a sketch of how you think your visualisation might look and add to this document.

Question 1: Is there a trend in the number of fatalities over time (yearly or monthly)?

* **Data attributes:**

Columns: Number Fatalities, Year, and optionally Month

* **Data transformation:**

We need to aggregate the fatalities by year or by both year and month.

* **Data type change:**

No major transformation needed, but we'll need to aggregate the data by summing fatalities by year or month.

* **Visualization:**

A line chart is suitable for showing the trend over time. The x-axis represents the year (or month/year combination), and the y-axis represents the total number of fatalities.

* **Sketch:**

A line chart with a continuous line that shows whether fatalities are increasing, decreasing, or staying constant over time.

Question 2: What is the distribution of fatalities by road user type (e.g., Driver, Passenger, Pedestrian)?

* **Data attributes:**

Columns: Crash ID, Road User

* **Data transformation:**

Aggregation by Road User is required to sum up fatalities for each type of road user.

* **Data type change:**

No major transformation needed beyond summing up.

* **Visualization:**

A pie chart or a bar chart that shows the proportion of fatalities for each type of road user.

* **Sketch:**

A pie chart with different segments for each road user type (Driver, Passenger, etc.) showing their respective percentages of the total fatalities, or a bar chart with road user types on the x-axis and fatality counts on the y-axis.

Question 3: What is the relationship between crash type and the involvement of heavy vehicles (e.g., buses, heavy trucks)?

* **Data attributes:**

Columns: Crash Type, Bus Involvement, Heavy Rigid Truck Involvement, Articulated Truck Involvement

* **Data transformation:**

We need to group the data by Crash Type and then calculate the proportion of crashes that involved each heavy vehicle type.

* **Data type change:**

No data type change required, but we need to convert involvement from categorical (Yes, No) to numerical (1 for Yes, 0 for No) for summarization purposes.

* **Visualization:**

A stacked bar chart where the x-axis represents the different crash types (e.g., Single, Multiple, Pedestrian) and the y-axis shows the proportion of heavy vehicle involvement. Each bar can be divided (stacked) into different vehicle types (Bus, Heavy Rigid Truck, Articulated Truck) to show how often each type of vehicle was involved in crashes.

* **Sketch:**

A stacked bar chart where each crash type has a segmented bar representing bus involvement, heavy rigid truck involvement, and articulated truck involvement. We can see at a glance how each crash type correlates with the involvement of heavy vehicles.