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.

Your Question 1: What are the most common crash types and their distribution throughout the year?

1. Data attribute

* “Month”, “Crash Type”

1. Data transformation

* Count how many times each kind of crash happens every month.
* Changing the list of crashes into simple numbers for to count them easily.

1. Data type change

* “Crash Type” start as names, turn into numbers for easy counting.
* “Month” is the label of time, change in to order numbers from January to December to see changes over the year.

1. Visualization sketch

A graph of a number of vehicles

Description automatically generated

Your Question 2: How does the severity of accidents change throughout the day?

1. Data attribute

* “Time of Day”, “Total Severity Score”, “Number of Accidents”, “Average Severity”

1. Data transformation

* Calculate 'Average Severity' For each time segment, divide the 'Total Severity Score' by the 'Number of Accidents' to get the 'Average Severity'.
* Group the accident times into categories like 'Morning', 'Afternoon', 'Evening', or 'Night'.

1. Data type change

* Time of Day: a way to split the day into parts like morning or night, stays the same.
* Total Severity Score: all the severity scores added up. It's a number and doesn't change.
* Number of Accidents: how many accidents happened. It's a count and doesn't change.
* Average Severity: a number we get from dividing the total severity by the number of accidents. It's still a number but shows an average for each part of the day.

1. Visualization sketch

A graph with blue and orange bars

Description automatically generated

Your Question 3: Which age groups are most affected by road fatalities?

1. Data attribute

* “Age Group”, “Count of Fatalities”

1. Data transformation

* Add together the death of each age group.

1. Data type change

* Keep the age groups as they are just adding up numbers, no change.

1. Visualization sketch

A pie chart with different colored circles

Description automatically generated

Include this file as evidence for your Demonstration 2