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

According to the fatality data, which type of road user has the most accident?

| Row Labels	Count of Road User
Driver	22931
Motorcycle pillion passenger	351
Motorcycle rider	6300
Passenger	11969
Pedal cyclist	1355
Pedestrian	8014
Grand Total	50920
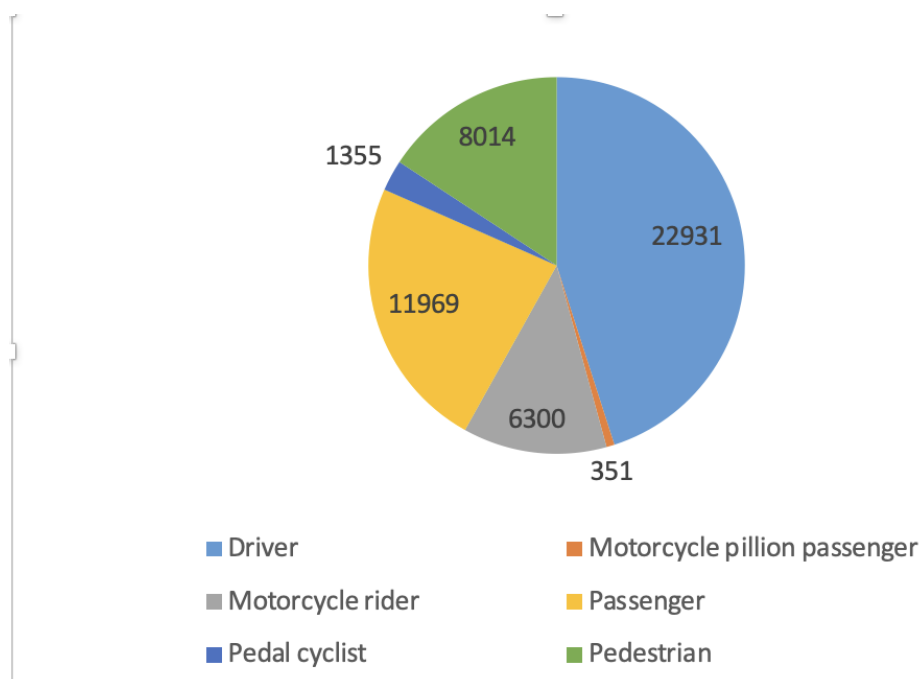
For the first question, the data attributes that I am using will be the Count of Road User. According to the pie chart, we can see that approximately 23 thousands of accident are caused by drivers which has the highest count in the whole pie chart. For the lowest count in the whole pie chart will be motorcycle pillion passenger.

Note: All the null values have been removed from the table as well as the pie chart (eg. -9)

Your Question 2

During the year 1990 to 2000, which year has the highest amount of male accident and which year has the highest amount of female accident?

| Count of Gender	Column Labels		
Row Labels	Female	Male	Grand Total
1990	697	1634	2331
1992	625	1348	1973
1993	560	1393	1953
1994	589	1338	1927
1995	604	1413	2017
1996	564	1406	1970
1997	561	1206	1767
1998	522	1232	1754
1999	546	1216	1762
2000	521	1296	1817
Grand Total	5789	13482	19271
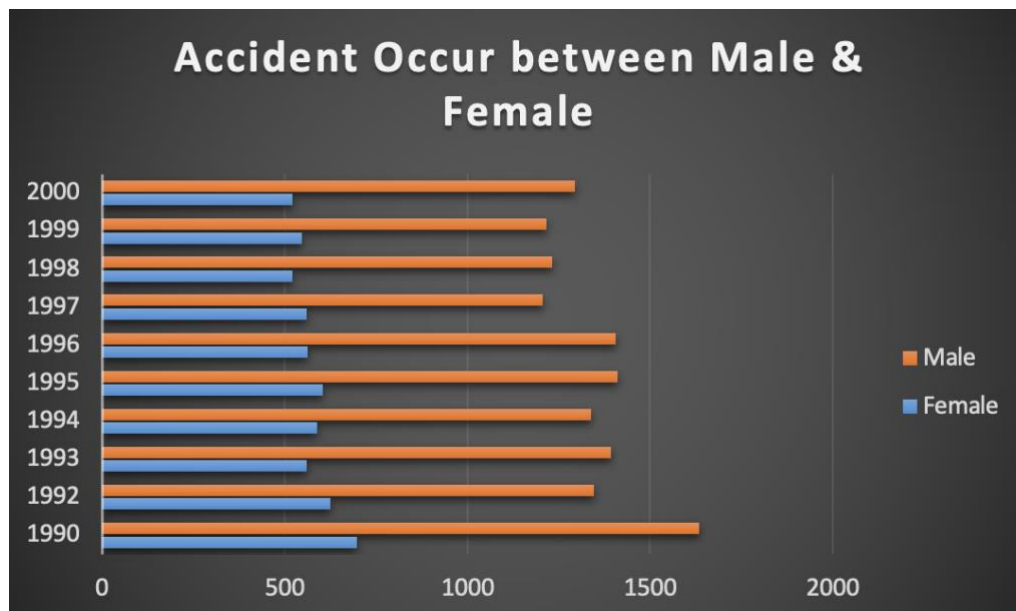
For the second question, I will be using the data attributes of Male, Female and Year. In this Graph, we are comparing the amount of accident occur between Male and Female in year 2000,1999,1998 and etc. During 1990, Male is the highest compare to the other years of Male and the lowest for Female will be during 2000.

Note: All the null values have been removed from the table as well as the graph (eg. -9)

Your Question 3

In which period has the most accident happened between Christmas Period and Easter Period?

Row Labels	Column Labels	No		Total Count of Christmas Period	Total Count of Easter Period
	Count of Christmas Period	Count of Easter Period			
No	49159	49159		49159	49159
Yes	1603	1603		1603	1603
Grand Total	**50762**	**50762**		**50762**	**50762**
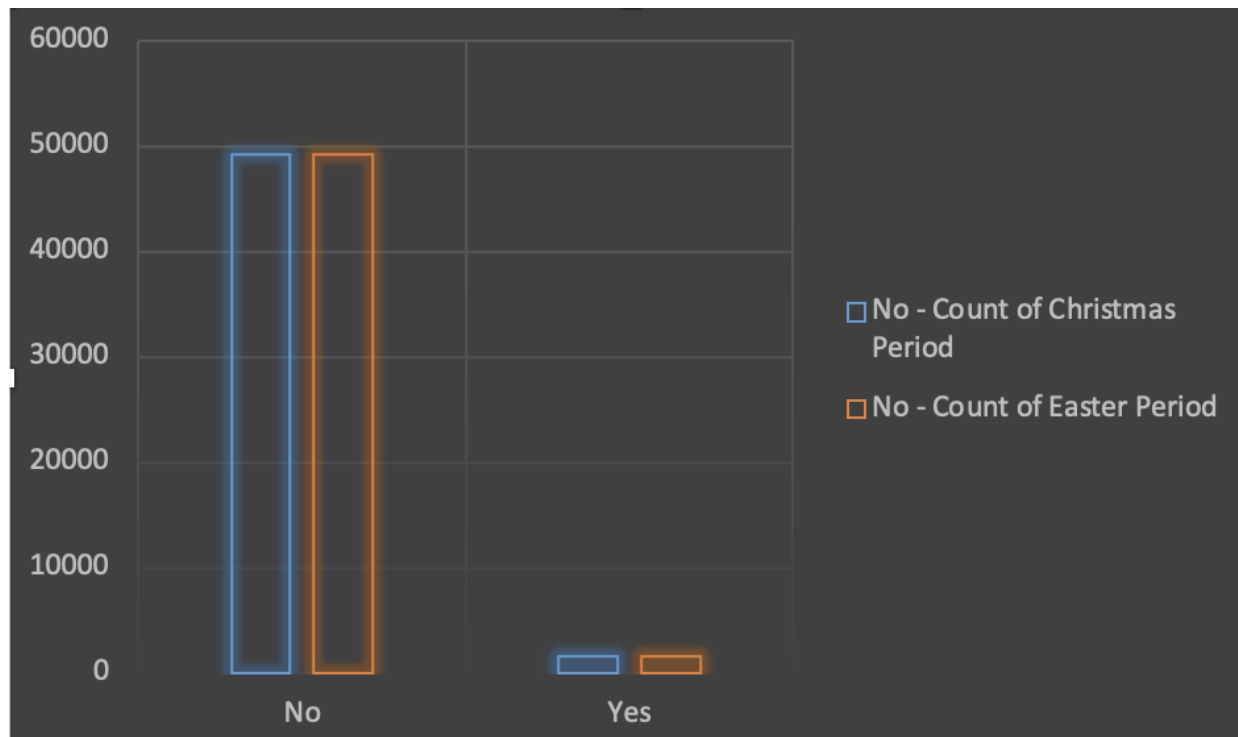
For the third question, I will be using Christmas Period and Easter Period as the data attribute. In this graph, we are comparing which period has the most accident between Christmas Period and Easter Period. According to the graph, we can see that both periods have the same amount of accident.

Note: All the null values have been removed from the table as well as the graph (eg. -9)

Include this file as evidence for your Demonstration 2