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.

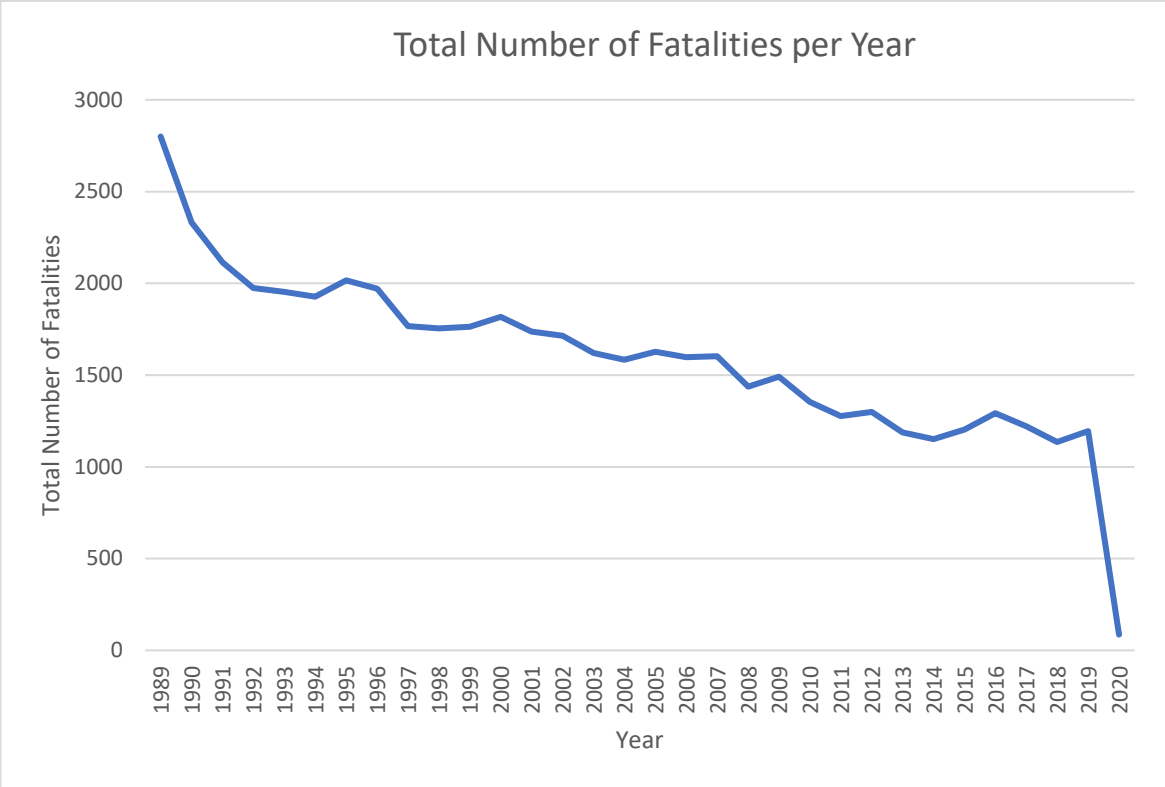
Your Question 1

Your answer here…

Total Number of Fatalities per Year

Year	Sum of Number Fatalities
1989	2800
1990	2331
1991	2113
1992	1974
1993	1953
1994	1928
1995	2017
1996	1970
1997	1767
1998	1755
1999	1764
2000	1817
2001	1737
2002	1715
2003	1621
2004	1583
2005	1627
2006	1598
2007	1603
2008	1437
2009	1491
2010	1353
2011	1277
2012	1300
2013	1187
2014	1151
2015	1204
2016	1292
2017	1221
2018	1135
2019	1194
2020	86
Grand Total	**51001**

Year	Total Number of Fatalities
1989	2800
1990	2331
1991	2113
1992	1974
1993	1953
1994	1928
1995	2017
1996	1970
1997	1767
1998	1755
1999	1764
2000	1817
2001	1737
2002	1715
2003	1621
2004	1583
2005	1627
2006	1598
2007	1603
2008	1437
2009	1491
2010	1353
2011	1277
2012	1300
2013	1187
2014	1151
2015	1204
2016	1292
2017	1221
2018	1135
2019	1194
2020	86



By using the Year and Number of Fatality columns, I am able to create a line graph displaying the total number of fatalities per year. No data transformation is needed as all the data is clear.

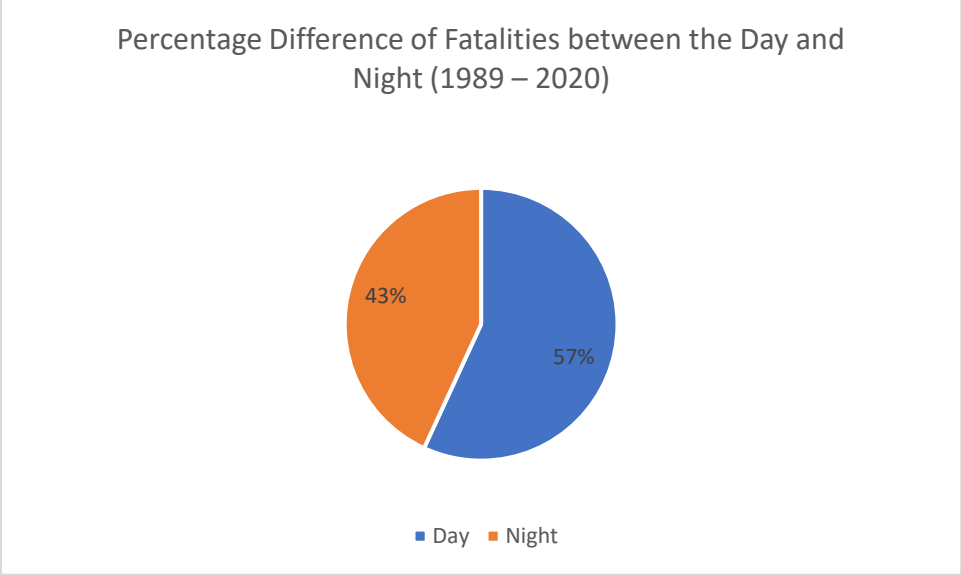
Your Question 2

Your answer here

Percentage Difference of Fatalities between the Day and Night (1989 – 2020)

Row Labels	Count of Crash ID
Day	29004
Night	21997
Grand Total	**51001**

Time of Day	Percentage of Fatalities
Day	29004
Night	21997



By using the Time of Day column and the Count of Crash IDs from the year 1989 to 2020, I am able to create a pie chart displaying the percentage of road accidents that occur during different times of the day. No data transformation is needed as we use the count (no misleading / confusing data placeholders present).

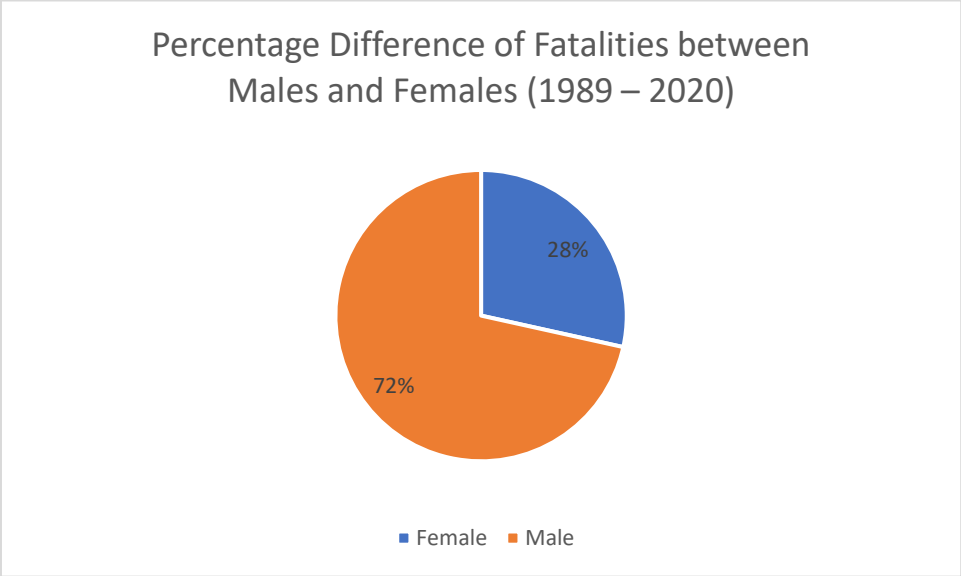
Your Question 3

Your answer here

Percentage Difference of Fatalities between Males and Females (1989 – 2020)

Row Labels	Count of Crash ID
-9	22
Female	14512
Male	36466
Unspecified	1
Grand Total	**51001**

Gender	Total Number of Fatalities
Female	14512
Male	36466



By using the Gender column and the Count of Crash IDs from the year 1989 to 2020, I am able to create a pie chart displaying the percentage of road accidents that occur between genders. Data transformation includes the destructive removal of "-9" unknown data and unspecified gender in the data given. This is because those data are unknown and do not contribute to the study.