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:

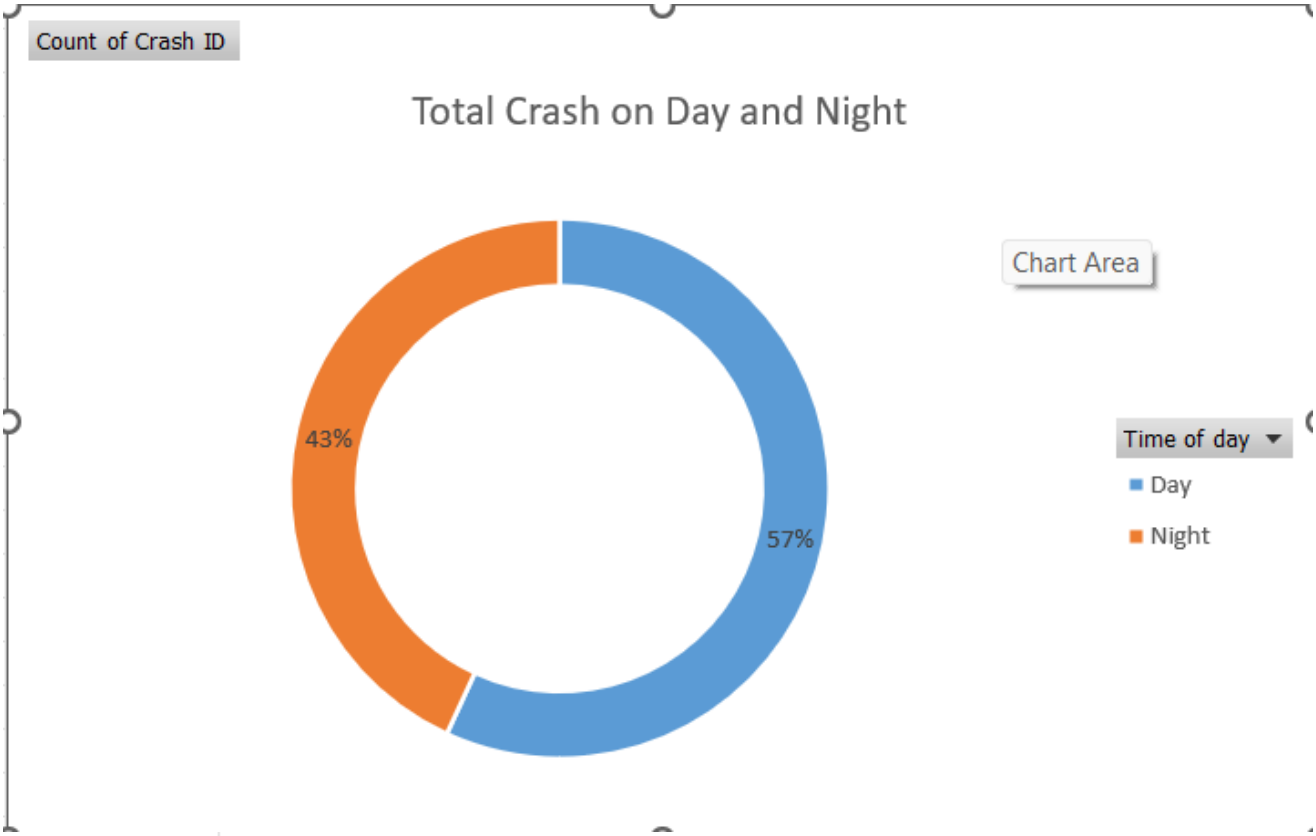
Which data attributes (columns) do you need to answer this question?

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.

How many total crashes occur on day/night time?



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

which data attributes (columns) do u need to answer this question

-Crash ID and time of the day (day/night)

Do you need to transform any of the data

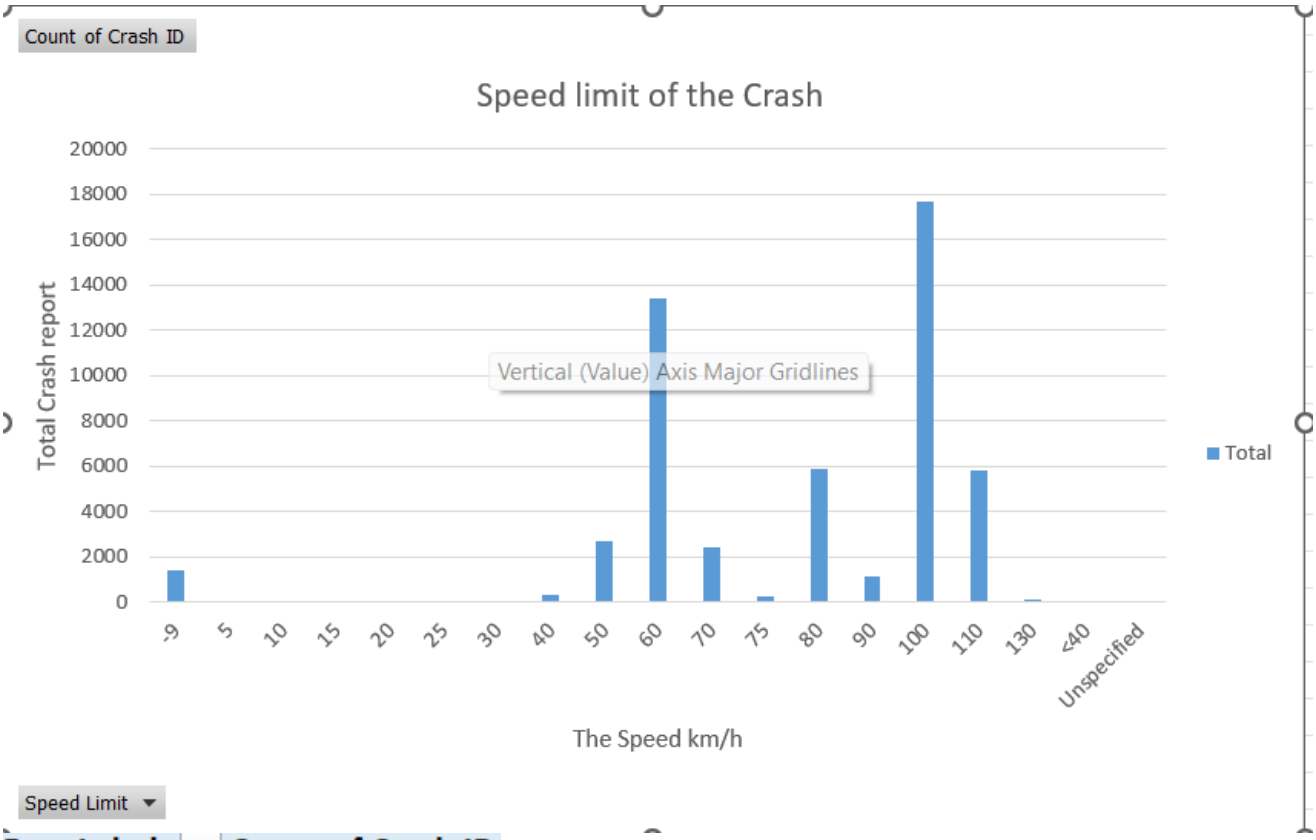
-We need to aggregate the data and also to calculate the percentages

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

-the data type for the Crash ID remains as a count when aggregate

-during calculation of percentage a transformation is needed to change the integer to percentage value as seen (57% for day and 43% as night)

What is the speed limit of all average crashes occur all time?



Speed Limit

Row Labels	Count of Crash ID
-9	1381
5	3
10	17
15	1
20	26
25	2
30	13
40	321
50	2685
60	13371
70	2424
75	254
80	5843
90	1097
100	17653
110	5796
130	106
<40	6
Unspecified	2
Grand Total	**51001**

which data attributes (columns) do u need to answer this question

-Crash ID and speed Limit(km/h)

Do you need to transform any of the data

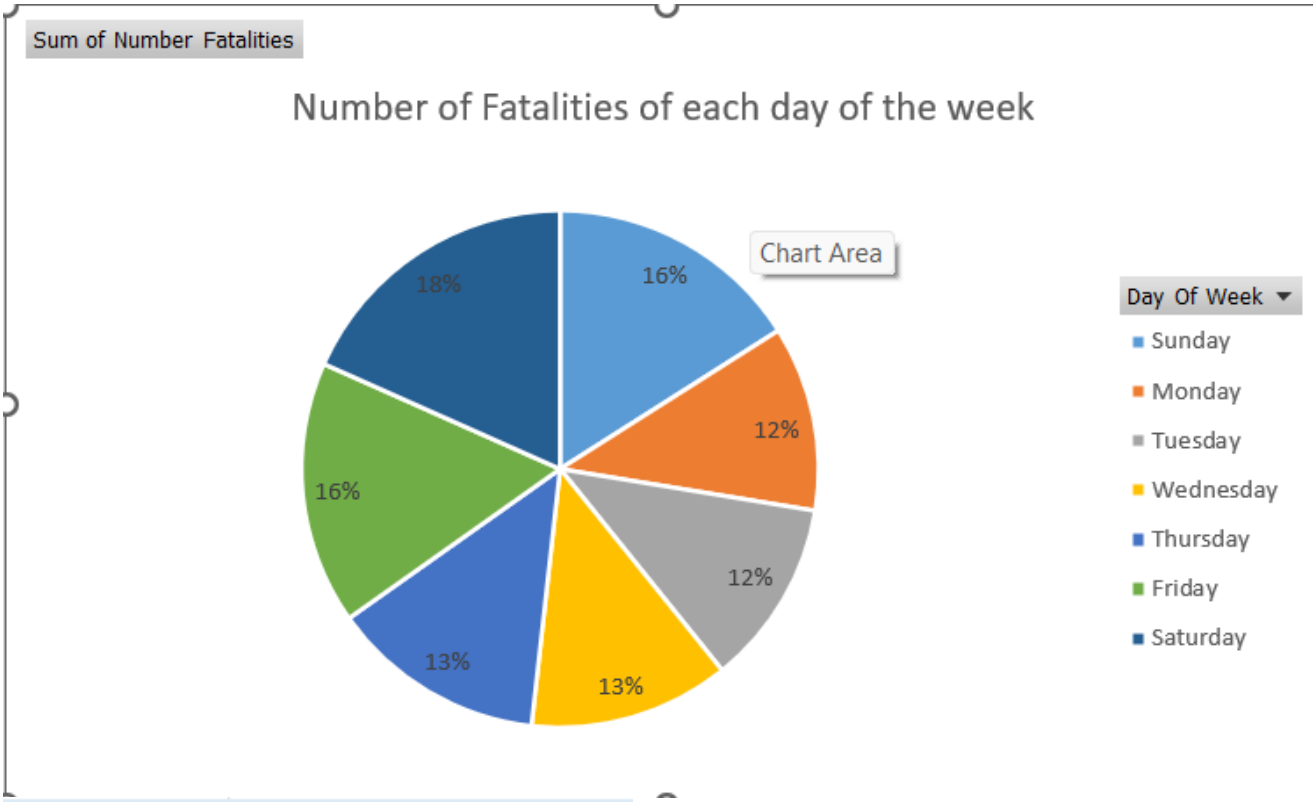
- Yes, it is needed to cleaning , sorting and filtering last but not least to also aggregate each data

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

-Yes it includes using speed range like <40/ unspecified that would be needed to categorized

-when aggregate speed categories, transform the limit attributes from a numerical value to a range value (bar chart)

What is the average number of crashes each day of the week?



Row Labels	Sum of Number Fatalities
Sunday	8163
Monday	5905
Tuesday	5929
Wednesday	6405
Thursday	6855
Friday	8390
Saturday	9354
Grand Total	**51001**

which data attributes (columns) do u need to answer this question

-Day of the week and sum of fatalities

Do you need to transform any of the data

-No, because we didn't transform any value at all

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

-You convert percentage into actual fatality counts, from float to integer

-each individual day, data becomes categorical instead of individual days

Include this file as evidence for your Demonstration 2