

## 1 Introduction

The growing alcohol consumption among teenagers has become a matter of concern in the past few years. Therefore, the main objective of this visualisation focuses on finding out how factors like age and students going out with their friends have an impact on their alcohol consumption. Another objective of the visualisation is to analyse/identify how the health status of the students varies with their alcohol consumption.

## 2 Description

The data set that has been used in this visualization has been collected from a survey of students studying maths and Portuguese language courses in secondary schools. It contains lots of interesting information about these students. The main information that has been visualised from this data set is how the factors like age and going out with friends have impacts on the alcohol consumption of these students. It has also been visualised how the health status varies with alcohol consumption of these students.

The data that has been visualised from the data set is using weekend alcohol consumption (numeric: from 1-very low to 5-very high) by the name of Walc, going out with friends (numeric: from 1-very low to 5-very high) by the name of gout, current health status (numeric: from 1-very low to 5-very high) by the name of health in the data set and H1-H5 in the visualisation, student's age (numeric: 15 to 22) by the name of age.

There is a total of three visualisations used to interpret the above-mentioned data. These three visualisations together present simultaneous views, animated views and interactive manipulation of the view.

The very **first visualization** aims to demonstrate the relationship between the weekend alcohol consumption of students and the how they go out with their friends. Since the aim of this visualisation is to demonstrate a relationship therefore a chord diagram has been used, there are a number of arcs each representing one of the factors of the relationship (includes go\_out 1-5 and Walc1-5), there arcs are arranged in such a fashion that they represent a circle. There are several chords connecting the arcs to one another, these chords signify how many students demonstrate that particular relationship. In order to enhance the visualisation, moving animations in the form of small balls have been used to along the path of the chords. The colours for this visualisation have been chosen such that each arc representing an element has a different colour and the balls/animations have the colours from according to their origin (the arc that they emerge from). The user can click on any arc/element to show/hide it from the visualisation, this demonstrates **interactive manipulation** of the view by the user. The orientation of the visualisation can be chosen by the user, by simply dragging the arcs/elements. This visualisation helps us analyse that maximum students (54 out of 375) fall in category of going out 2 times with their friends and having a weekend alcohol consumption of 1 day. Then there are 47 students who go out three times and have an alcohol consumption of one on the weekends. Therefore, this visualisation can be said to be **analyst centric** as it exposes some important insight about the data and this also helps us analyse the trends observed in the data. This visualisation displays the technical element of **animated view** of data.

The **second visualisation** used, represents the relationship between the age of the students and their alcohol consumption on weekends. Again a chord diagram has been chosen to represent this relation but this chord diagram is visually different from the one used in the first part. Arcs have been used to represent the different elements of the relationship weekly alcohol consumption (Walc1-5) and age (age15-22). Initially, the arcs/elements and the chords are coloured in lighter shades but when we hover over to any of them the colour shifts to a darker shade of the original colour and the labels also darken representing the current selection. Pop up boxes can be observed when we hover over to any of the chords, the thickness of the chords represents the number of students exhibiting that relationship, this demonstrates **interactive manipulation** by the user. This visualisation helps us analyse that the maximum students (46/375) have a weekend alcohol consumption of 1 and are 15

years old. This visualisation also helps us know that maximum students have an alcohol consumption of 1 followed by the second maximum number of students having an alcohol consumption of 2 for all the age groups in common and only a very small amount of students have alcohol consumption of 5. Therefore, this information can be said to be **Analyst centric** as this visualisation seeks to explore some important insight about the data.

The **third visualisation** an interactive radial histogram that represents the number of students based on the relationship between their current health status (H1-5) and their weekend alcohol consumptions (W1-5). Just above this visualisation, a slider bar has been provided that helps us zoom in this visualisation. Another button is provided just below the right end of this scroll bar that helps in returning to the original position of the visualisation. By hovering over the histograms one can find out the number of students exhibiting that relationship. Another form of **interactive visualisation** has been implemented here; we can select a particular sector of this circular visualisation with the help of a mouse click that helps us zoom in the visualisation. This visualisation helps us find out that maximum students that belong to a particular relation that is 47/375 students have current health status as 5 and weekend alcohol consumption as 1. This graphs also helps us analyse that if the studs have less alcohol consumption then they usually tend to display a good health status. Therefore, this visualisation can be said to be **analyst centric** as it seeks to explore the important information in the data.

R scripts were run on the original data in R studio to filter the data according to the requirements. As the original data contains many other variables that have not been utilised here.

### 3 Citing third party resources

amCharts 4 library is a java script library that has been used in this assignment to enhance the visual effects produced by using java script, HTML and CSS.

### References

- [1] <https://www.kaggle.com/uciml/student-alcohol-consumption#student-mat.csv>
- [2] <https://www.amcharts.com/docs/v4/getting-started/using-javascript/>
- [3] [https://en.wikipedia.org/wiki/Chord\\_diagram](https://en.wikipedia.org/wiki/Chord_diagram)