DSA8022 – Visual Analytics

Practical Assessment

Dataset

- This dataset (student.csv) describes student achievement in secondary education of two Portuguese schools using data attributes including subjects (mathematics, Portuguese language), student grades, demographic, social and school related features. It was collected by using school reports and questionnaires. Important note: attribute G3 has a strong correlation with attributes G2 and G1 because G3 is the 3rd period (i.e., final year) grade, while G1 and G2 are the 1st and 2nd period grades. It is more difficult to predict G3 without G2 and G1, but such predictions are insightful.

- You have an opportunity to gain insights through visual analytics.

Dashboard Overview

- Using the **student.csv** dataset (available on Canvas), create an interactive dashboard using Shiny and ggplot2. A description of the attributes can be found in student_attributes.txt (also available on Canvas).

- The following slides present the expected dashboard (the theme and functionalities) that you will need to develop.

- Submission: upload your R code (such as app.r) to canvas.

Student Overview Page Bar Chart Tab

Student Dashboard   Options ▾

Student Overview
About

Parameters:

X-Axis:

Grade in Period 1 ▾

Colour by:

Subject ▾

Student Filters:

Random Sample Size:

100 | 522 | 1,044

100 200 300 400 500 600 700 800 900 1,00044

Filter by mother's education:

◉ All
○ 0 - none
○ 1 - primary education
○ 2 - 5th to 9th grade
○ 3 - secondary education
○ 4 - higher education

Filter by age:

15 | 16 | | 22 |

15   16   17   18   19   20   21   22

Apply Filters

Student Overview

| Bar Chart | Scatter plot | Data table |

In a sample of 522 randomly selected students, there are 427 students that satisfy the filter conditions
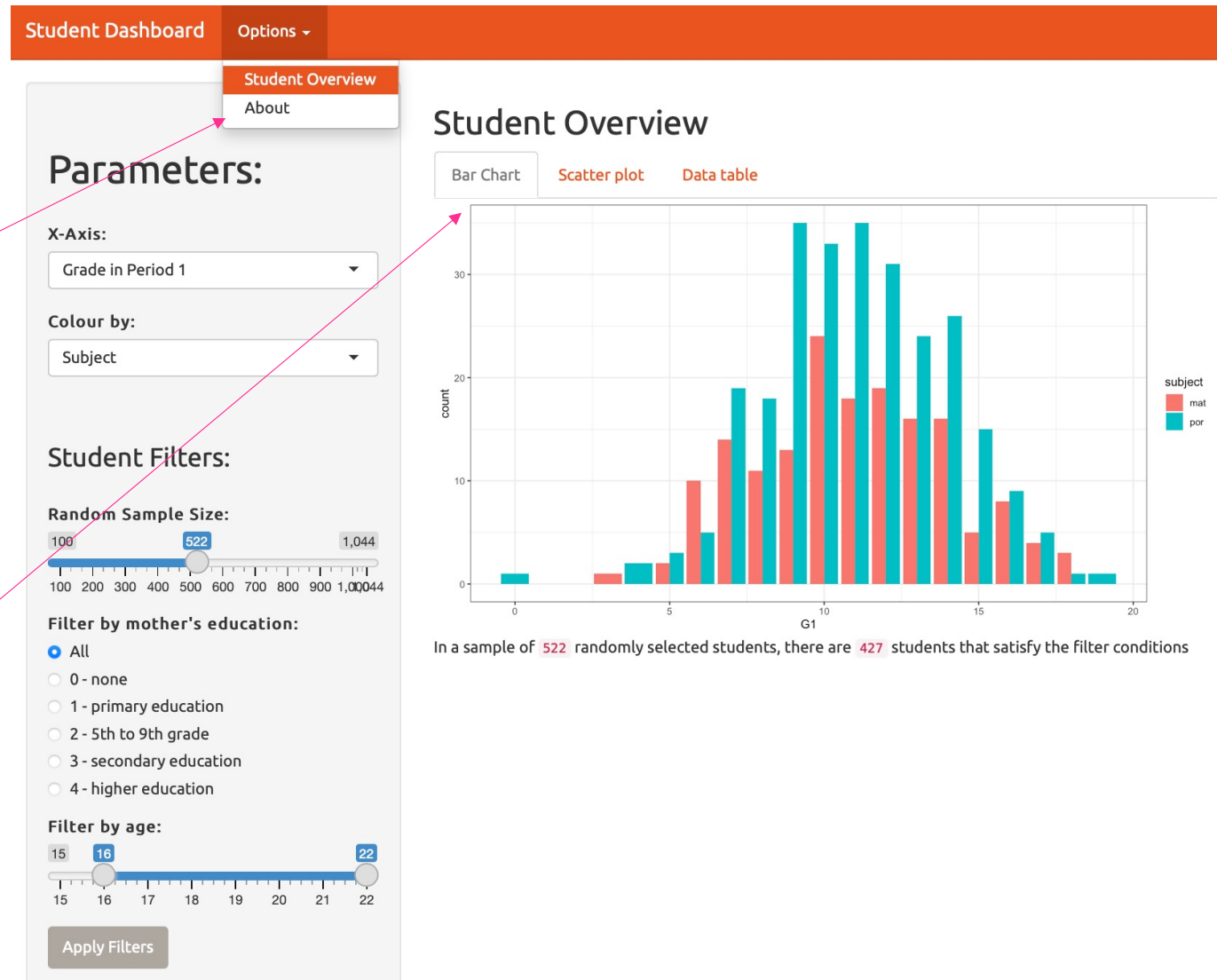
There are two pages:
- Student Overview
- About

In "Student Overview" page, there are three tab panels:
- Bar chart
- Scatter plot
- Data table
These panels have their own parameters, and are reactive to changes to the parameters. These panels are reactive to changes to filters only when Apply Filters button is pressed.

Student Overview Page   Scatter plot Tab

Student Dashboard   Options ▾

Parameters:

X-Axis:

Grade in Period 1   ▾

Y-Axis:

Grade in Period 3   ▾

Colour by:

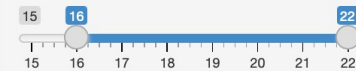Subject   ▾

Student Filters:

Random Sample Size:

| 100 | | 522 | | 1,044 |

100 200 300 400 500 600 700 800 900 1,00044

Filter by mother's education:

- ⦿ All
- ○ 0 - none
- ○ 1 - primary education
- ○ 2 - 5th to 9th grade
- ○ 3 - secondary education
- ○ 4 - higher education

Filter by age:

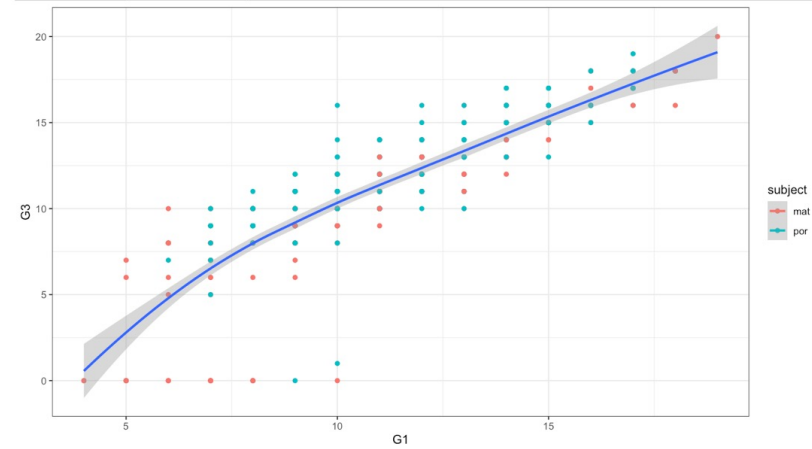| 15 | 16 | | 22 |

15   16   17   18   19   20   21   22
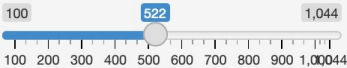
Apply Filters

Student Overview

Bar Chart   Scatter plot   Data table

In a sample of 522 randomly selected students, there are 415 students that satisfy the filter conditions

Student Overview Page   Data table Tab

Parameters:

Student Filters:

Random Sample Size:

100   522   1,044

100 200 300 400 500 600 700 800 900 1,00 1044

Filter by mother's education:
- ● All
- ○ 0 - none
- ○ 1 - primary education
- ○ 2 - 5th to 9th grade
- ○ 3 - secondary education
- ○ 4 - higher education

Filter by age:

15   16   22

15 16 17 18 19 20 21 22

[Apply Filters]

Student Overview

Bar Chart   Scatter plot   Data table

Show [10 ⬍] entries   Search: [＿＿＿＿＿＿]

	subject ⬍	school ⬍	sex ⬍	famsize ⬍	Medu ⬍	Fedu ⬍	age ⬍	studytime ⬍
1	por	GP	M	LE3	4 - higher education	4 - higher education	18	2
2	mat	GP	F	LE3	4 - higher education	4 - higher education	16	2
3	por	GP	F	LE3	1 - primary education	1 - primary education	18	2
4	por	GP	M	GT3	3 - secondary education	2 - 5th to 9th grade	17	1
5	por	MS	M	LE3	1 - primary education	1 - primary education	18	2
6	por	MS	M	GT3	2 - 5th to 9th grade	2 - 5th to 9th grade	17	1
7	por	MS	M	GT3	2 - 5th to 9th grade	1 - primary education	18	1
8	por	MS	F	GT3	1 - primary education	3 - secondary education	16	1
9	por	GP	F	LE3	4 - higher education	4 - higher education	21	3
10	por	MS	F	GT3	4 - higher education	4 - higher education	17	2

Showing 1 to 10 of 416 entries   Previous [1] 2 3 4 5 … 42 Next

In a sample of 522 randomly selected students, there are 416 students that satisfy the filter conditions

Student Dashboard   Options ▾

About Data: A Summary

Size of the entire dataset is: `1044` , dimensionality is: `15` !
Breakdown of data, grouped by school and sex and summarised
by mean and standard deviation of absences, is given below:

school	sex	avg_absences	sd_absences
GP	F	5.38	7.71
GP	M	4.56	5.64
MS	F	2.36	3.19
MS	M	3.57	3.55

About Author

My name is *<insert_your_name_here>*

A Queen's University Belfast student

SHAPING
A BETTER
WORLD
SINCE 1845

This is our brand driver.

QUEEN'S
UNIVERSITY
BELFAST
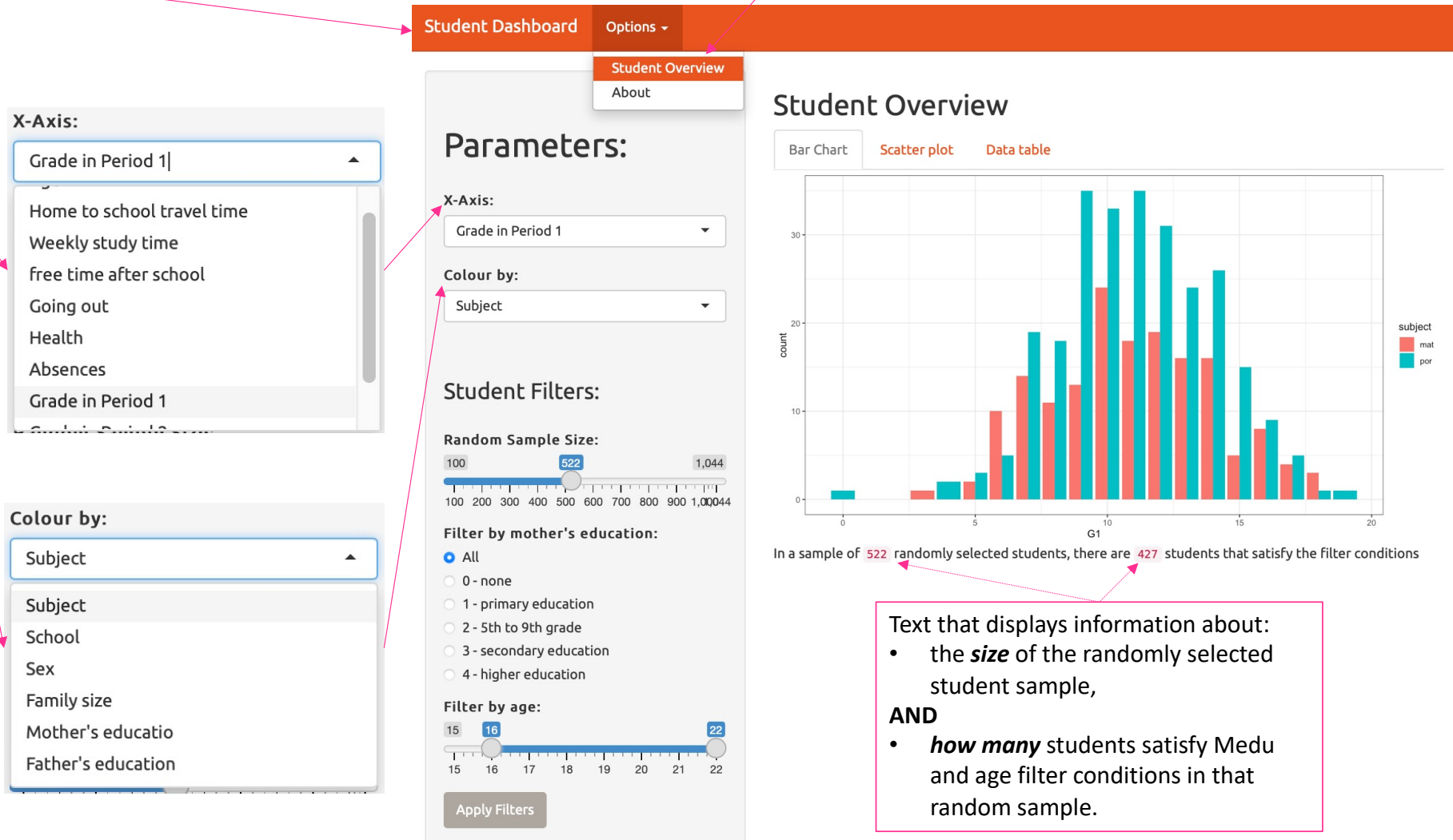
The page *theme* should be set properly to match the colours. **HINT**: consider using *shinytheme* function with a proper value.

Dropdown menu – to select page

The X-Axis selectInput should include *all ordinal numeric variables* in the student dataset

The "Colour by" selectInput should include *all categorical variables* in student dataset

Student Dashboard Options ▾

Student Overview
About

Student Overview

Bar Chart   Scatter plot   Data table

X-Axis:

Grade in Period 1

Home to school travel time
Weekly study time
free time after school
Going out
Health
Absences
Grade in Period 1

Parameters:

X-Axis:

Grade in Period 1

Colour by:

Subject

Student Filters:

Random Sample Size:

100   522   1,044
100 200 300 400 500 600 700 800 900 1,000 1,044

Filter by mother's education:
- ◉ All
- ○ 0 - none
- ○ 1 - primary education
- ○ 2 - 5th to 9th grade
- ○ 3 - secondary education
- ○ 4 - higher education

Filter by age:
15   16   22
15 16 17 18 19 20 21 22

Apply Filters

Colour by:

Subject

Subject
School
Sex
Family size
Mother's educatio
Father's education

count
30
20
10
0

0   5   10   15   20
G1

subject
mat
por

In a sample of 522 randomly selected students, there are 427 students that satisfy the filter conditions

Text that displays information about:
- the *size* of the randomly selected student sample,

AND
- *how many* students satisfy Medu and age filter conditions in that random sample.

Default values. **NB**: The input controls are *conditional* on plotOutput type (bar chart, scatter plot or data table). **HINT**: Use *conditionalPanel* function.

A dashboard user should be able to select desired number of *random* student samples

Perform the filtering procedure *on the selected random sample* to separate the observations where:
- **Mother's education** is equal to one of the labels that appear in the **Medu** column of the student dataset.
  NB*:* if the radio button *'All'* is selected, then all Medu labels should be included in the filtered dataset.

AND
- The values of the **age** variable are in the specified range

Data filtering process should be performed **ONLY** when the **'Apply Filters'** button is pressed.

Student Dashboard   Options ▾

Parameters:

X-Axis:
Grade in Period 1   ▼

Y-Axis:
Grade in Period 3   ▼

Colour by:
Subject   ▼

Student Filters:

Random Sample Size:
100   [522]   1,044
100 200 300 400 500 600 700 800 900 1,00044

Filter by mother's education:
- ● All
- ○ 0 - none
- ○ 1 - primary education
- ○ 2 - 5th to 9th grade
- ○ 3 - secondary education
- ○ 4 - higher education

Filter by age:
15   [16]   [22]
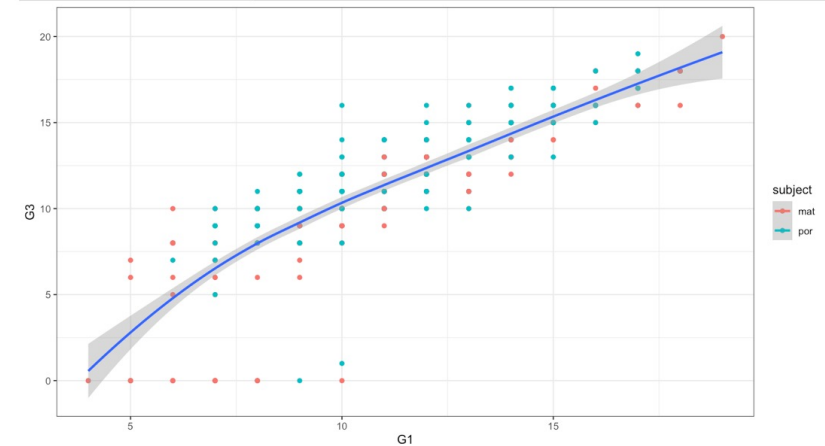15 16 17 18 19 20 21 22

Apply Filters

Student Overview

Bar Chart   Scatter plot   Data table

In a sample of 522 randomly selected students, there are 415 students that satisfy the filter conditions

HINT: Use *uiOutput* function to display this textual information

- *Minimum* random sample size should be manually set to 100.
- *Maximum* sample size should be equal to the student dataset size (which should be determined programmatically).
- *Starting position* of a slider should be a half of the student dataset size (which should be determined *programmatically*).

Default **Medu** Filter selection

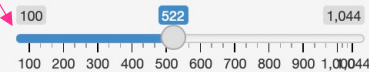*Programmatically* determine minimum and maximum age values.

Student Dashboard   Options ▾

Parameters:

Student Filters:
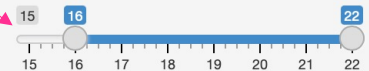
Random Sample Size:

| 100 | **522** | 1,044 |

100 200 300 400 500 600 700 800 900 1,000 1,044

Filter by mother's education:

- ● All
- ○ 0 - none
- ○ 1 - primary education
- ○ 2 - 5th to 9th grade
- ○ 3 - secondary education
- ○ 4 - higher education

Filter by age:

| 15 | **16** | | **22** |

15 16 17 18 19 20 21 22

Apply Filters

Student Overview

Bar Chart   Scatter plot   Data table

Show 10 ⬍ entries   Search: [   ]

	subject ⬍	school ⬍	sex ⬍	famsize ⬍	Medu ⬍	Fedu ⬍	age ⬍	studytime ⬍
1	por	GP	M	LE3	4 - higher education	4 - higher education	18	2
2	mat	GP	F	LE3	4 - higher education	4 - higher education	16	2
3	por	GP	F	LE3	1 - primary education	1 - primary education	18	2
4	por	GP	M	GT3	3 - secondary education	2 - 5th to 9th grade	17	1
5	por	MS	M	LE3	1 - primary education	1 - primary education	18	2
6	por	MS	M	GT3	2 - 5th to 9th grade	2 - 5th to 9th grade	17	1
7	por	MS	M	GT3	2 - 5th to 9th grade	1 - primary education	18	1
8	por	MS	F	GT3	1 - primary education	3 - secondary education	16	1
9	por	GP	F	LE3	4 - higher education	4 - higher education	21	3
10	por	MS	F	GT3	4 - higher education	4 - higher education	17	2

Showing 1 to 10 of 416 entries   Previous [1] 2 3 4 5 … 42 Next

In a sample of 522 randomly selected students, there are 416 students that satisfy the filter conditions

About Data: A Summary

Size of the entire dataset is: 1044 , dimensionality is: 15 !
Breakdown of data, grouped by school and sex and summarised
by mean and standard deviation of absences, is given below:

school	sex	avg_absences	sd_absences
GP	F	5.38	7.71
GP	M	4.56	5.64
MS	F	2.36	3.19
MS	M	3.57	3.55

Width: *5 cells* in the grid

Heading 1

Text about the *size* of the dataset, its *dimensionality* (i.e., no of columns) and the summary table

Breakdown of data grouped by *school* and *sex* and summarised by *mean* and *standard* deviation of *absences*

There is a *horizontal line* here in black colour.

About Data and **About Author** need to be in *different rows*

About Author

Heading 1

My name is *<insert_your_name_here>*

A Queen's University Belfast student

Bold and Italic

About Author text and the "Shaping …" image need to be in a *same row*

SHAPING
A BETTER
WORLD
SINCE 1845

This is our brand driven.

QUEEN'S
UNIVERSITY
BELFAST

Width: *5 cells* in the grid

Empty space width: *4 cells*

Width: *3 cells* in the grid

Image URL:
https://bit.ly/3pQpS4b

Hyperlink to: http://www.qub.ac.uk

Grading
(based on implementation)

- Dashboard Layout Design (overall) **[15 marks]**

- "Student Overview" Page:
  - Bar Chart Functionality **[10 marks]**
  - Scatter Plot Functionality **[10 marks]**
  - Data Table Functionality **[10 marks]**
  - Conditional Panel **[10 marks**]
  - Textual Information about selected/filtered data **[5 marks**]
  - Student Filtering implementation **[10 marks]**
  - Apply Filters Button implementation **[10 marks]**

- "About" Page **[20 marks]**

You may be invited to attend an interview session during marking.