

Arts Application Programming 2020/21: Coursework

The coursework is worth 60%. There are two submissions:

- **Design decisions (20 marks – 10%):** submission deadline, 10 am, **Friday 30 October (week 6)**, pdf/word file.
- **Final submission (80 marks – 50%):** 10am, **Monday 30 November (week 11)**, zip folder (including code, csv file, any image files and pdf/word file)

Electronic submissions to QMPlus (not email). Allow sufficient time to upload: submissions that are e.g. 1 second late are registered as one-day-late on the system and receive a 5% lateness penalty. Late submissions are given a 5% penalty per day. No work can be submitted more than 7 days after the coursework deadline. Note that work cannot be resubmitted after the coursework deadline, nor after marks are released.

Regulations: Students should read the student handbook to ensure they clearly understand what it meant by plagiarism, and possible penalties involved. This is an individual project and students are expected to present work that is clearly their own.

The aim of this coursework is to generate a visualisation of a dataset.

Students must use ONE of the datasets available on QMPlus (see week 4 block). Note the points about the following datasets:

- Festival headliners (select 1-5 of the festivals – if selecting >1 festival then e.g. match by dates)
- William Blake images (note that if only selecting line engravings then this feature cannot count as one of the 5 columns)
- Wimbledon winners (this will need converting to .csv, but it is currently available as an excel file as there is more than one tab to explore)

This site is helpful for exploring data stored in csv files: <https://www.databasic.io/en/wtfcsv/>

For all data:

- Your csv file MUST contain at least 20 rows (i.e. you do not have to use the entire CSV shown) – note that more interesting visualisations on this module usually show more rows than this, but 20 is the a minimum
- Your csv file MUST contain at least 5 columns (note for the festivals dataset this does not include the festival name unless you are using data from at least 3 of the festivals). Note that if all the data is the same for a column this cannot be included in the column count.
- All the data in your csv MUST be used in the visualisation (i.e. your visualisation must show at least 20 'items', e.g. Wimbledon winners, with 5 features for each item – you cannot use just one or two features of the data)
- Your visualisation must be able to show at least 20 individual items, i.e. rows from the csv, not only e.g. averages of these items
- Note that this data has to be read into your code, not written/hardcoded into your sketch

Part 1. State what dataset you will be using (provide a link or paste the data to the end of the file if you are not using one of the datasets from QMPlus). Note that you cannot change your choice of dataset after this submission.

Part 2: Observation & Analysis

- State your key observations about the data e.g. value ranges, correlations, averages AND HOW THIS IMPACTS VISUALISATION DECISIONS
- If you are not including the entire dataset justify your source of rows and columns
- Define a **numbered list of concrete goals for the visualisation – based on your key observations about this data – then address how you will address/meet each goal.**

e.g. What patterns or relationships can be visualised? Why is it important to understand these patterns or relationships? How will you decide if your visualisation is successful? What should a user be able to do with the visualisation that they could not just by looking at the data in a table? Provide examples. where relevant, justify your choice of data from the full dataset

Do not discuss e.g. graphical forms and layout in this section (use section 3: Design)

[5 marks]

Part 3: Design

Based on the points in part 2, what kinds of visualisation technique are best suited to this

dataset? Explain your design choices for mapping properties of the data to visual properties in your visualisation. Why did you adopt each choice over other possibilities that you considered? What interaction have you considered? Provide examples from the data that inform these decisions. Designs that do not use some values from the dataset will score low marks as the data has to inform your decisions.

What colours, shapes etc? every point in this part needs to be shown as a drawing

Note that the **maximum window size is 1200x800 pixels**, i.e. `max size(1200,800)`.

You must NOT use `fullScreen()`

Provide drawings (not code at this stage). These should form the basis for writing your code

[15 marks]

PART 1

Dataset: Film Rating

PART 2

Key Observation; Data Analysis

1. Year of Film's release

2014;

Top Five (2014)/ Leviathan (2014)/ Unbroken (2014)/

The Imitation Game (2014)/ Night at the Museum: Secret of the Tomb
(2014)/ Selma (2014)/ Wild Tales (2014)/ Annie(2014)

2015;

Avengers: Age of Ultron (2015)/ Cinderella (2015)/ Ant-Man (2015)/ Do
You Believe? (2015)/ Hot Tub Time Machine 2 (2015)/ The Water
Diviner (2015)/ Irrational Man (2015)/ Shaun the Sheep Movie (2015)/
Love & Mercy (2015)/ Far From The Madding Crowd (2015)/ Black Sea
(2015)/ Taken 3 (2015)/ Ted 2 (2015)/ Southpaw (2015)/ Pixels (2015)/
McFarland, USA (2015)/ Insidious: Chapter 3 (2015)/ The Man From
U.N.C.L.E. (2015)/ Run All Night (2015)/ Trainwreck (2015)/ Ex
Machina (2015)/ Still Alice (2015)/ The End of the Tour (2015)/ Red
Army (2015)/ When Marnie Was There (2015)/ The Hunting Ground
(2015)/ The Boy Next Door (2015)/ Aloha (2015)/ The Loft (2015)/ 5
Flights Up (2015)/ Welcome to Me (2015)/ Saint Laurent (2015)/ Maps
to the Stars (2015)/ I'll See You In My Dreams (2015)/ Timbuktu (2015)/
About Elly (2015)/ The Diary of a Teenage Girl (2015)/ Kingsman: The

Secret Service (2015)/ Tomorrowland (2015)/ The Divergent Series:
Insurgent (2015)/ Fantastic Four (2015)/ Terminator Genisys (2015)/
Pitch Perfect 2 (2015)/ Entourage (2015)/ The Age of Adaline (2015)

**2. Rotten Tomatoes score/ Rotten Tomatoes Users Score/ Metacritic/ Averages
of Rotten Tomatoes, Rotten Tomatoes Users Score and Metacritic**

0-10;

10-20;

20-30;

30-40;

40-50;

50-60;

60-70;

70-80;

80-90;

90-100;

3. IMDB (/10)

Below 5.0/ Above 5.0

4. Fandango Stars (/5)

Below 3.0/ Above 3.0

Distribution:

The basic body is the year, and the branches are within the range of the average score. And the branches are subdivided and the scores of each site are expressed using the range.

Concrete Goals for the Visualisation

1. What patterns or relationships can be visualised?

The biggest frame is divided into the opening year; 2014 and 2015. It can be further divided, the scope of the scores can be subdivided by the sites that evaluate the movie and visualized.

For example, movies released in 2014 are visualized using rotten tomatoes site scores (from 0 to 10, 10 to 20, 20 to 30, 30 to 40, 40 to 50, 50 to 60, 60 to 70, 70 to 80, 80 to 90, and 90 to 100).

2. Why is it important to understand these patterns or relationships?

This is because movie rankings are basically divided by scores and each site has a different standard of rating. The different criteria mean that some sites have a perfect score of 100 (Rotten Tomatoes users and Metacritical) and others have a decimal point in overall evaluation with a perfect score of 10 or 5 (IMDB and Fantango Star). However, the difference between these two representative distinctions is that each of them has to be paid separately when averaging because the perfect score is different. Especially, the reason why we get the average is because we can see the movie's evaluation more easily. This is a standard by which

users can make judgments based on average results without having to go around various sites to get an assessment of a movie.

For example, Rotten Tomatoes, Rotten Tomatoes users and Metacritical, the average of the ratings of the three sites is divided by three. Then, the scores can be divided into details and visualized again, such as those range (from 0 to 10, 10 to 20, 20 to 30, 30 to 40, 40 to 50, 50 to 60, 60 to 70, 70 to 80, 80 to 90, and 90 to 100).

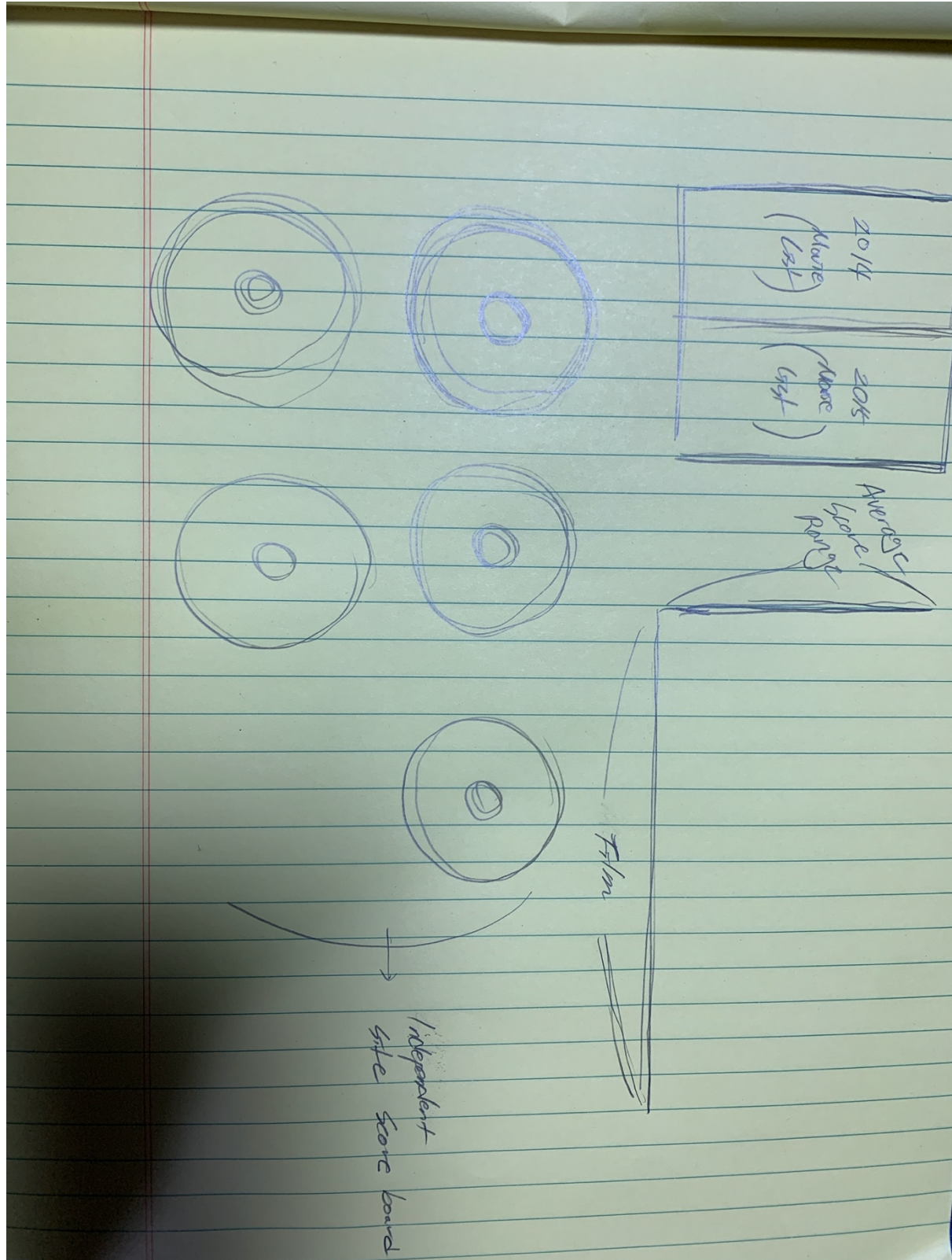
3. How will you decide if your visualisation is successful?

This visualisation will be successful because it has analysed and used the simplest and most basic information in a given piece of information and is unlikely to create much difficulty and complexity in drawing or using code.

4. What should a user be able to do with the visualisation that they could not just by looking at the data in a table?

Visualization of data is a graphical representation of information and data. Users should provide trends, outliers, and patterns in the data by using visual elements such as graph, table and maps. Our eyes are sensitive to colors and patterns. In other words, data visualization should allow us to see trends and singularities quickly and concisely when we look at charts. When users visualize data, they need to be able to deliver information accurately and quickly, rather than simply making it flashy and big.

PART 3



I decided to use the visualisation technology of the icon- based technique. Because there is a lot of information, which is the year of movie release and the average score of each site, to contain. I thought that the information I chose is difficult to put in a single graph or table. Therefore, I believe that if I use icons to separate and express each piece of information, you will find much more information quickly and easily. And I think, in this case, people who look at my visualised drawing can simply find the information they want. In fact, if you look at movie brochures, there are cases where visualised pictures or phrases are attached that are scored on each site. I came up with this and organized my visualisation. The same shape as the movie poster, half of which was released in 2014, is provided with a list of films released in 2015, and the average visualisation of the three sites in dot graphs. Provides movie-specific scores for each site above a circle graph, star-shaped or horizontal graph. Unlike the average graph, only the lower and upper criteria are used for concise and fast transmission. For example, like the standard 3.0 points that I choose from Fandango Stars site, some films mark 3.0 points below and others above 3.0, making it simple and easy for people to grasp the information.