Interactive Visualisation
Presentation

Team Members

ATHIRA, SATHEESAN (W9500883)

DANIEL, HARDING (Q5038939)

MIDHUN TEJA, RAVIPATI (W9506023)

OLUWASEGUN, SALAMI (A0341309)

YU SHING, LUI (A0201746)

Introduction

- Students' attendance at lectures and active engagement with other academic fixtures is critical to the success of both the students and the university alike

- It is therefore important to draw insight from students' past attendance data

Motivation

- To demonstrate our group's capacity and expertise at handling all the university's web-based interactive data visualisation work

- The choice of charts has been carefully made with the audience in mind (i.e. the university community and the public in general)

Introduction

Visualisation theme:

The effect of timetabling on student attendance

Elements of Timetabling Investigated:

- Lecture start time
- The day of the week
- Number of classes per day

Introduction

Data Cleaning Approach

Records with the following were removed from the visualisations:

- Missing values

- Invalid values e.g integers under the postal code field

- Illogical values:
    - Students with zero attendance (50 students)
    - Students with less than 13 lecture fixtures (17 students)
    - Clashing classes (approx. 14,000)

Data Cleaning

- The procedures of finding average distance.

- Sort out the valid postcode

- Search UK postal code of Longitude and Latitude from website.

- (Ref: https://gridreferencefinder.com/postcodeBatchConverter/)

Data Cleaning

- Calculate the distance of each postal code by excel

- (Ref: https://www.techwalla.com/articles/the-haversine-equation-in-excel)

- Cross out data if the distance is over 60 miles as too far from Uni (Approx. distance from Middlesbrough to York)

- 100,842 of 168,614 can be used for analysis

	L	M	N	O	P	Q	R	S
	Postal Code	Latitude_1	Longitude_1	Postal Code	Latitude_2	Longitude_2	Distance 1 (miles	Distance 2 (KM)
	TS1	54.575852	-1.2449419	YO8	53.784935	-1.0676994	55.17409058	88.7964546
				YO7	54.233504	-1.3423004	24.00152242	38.6277341
	TS1 3BX	54.5704654	-1.2369515	YO62	54.164803	-0.9330358	31.08493876	50.027691
				YO61	54.101648	-1.1103757	33.24526166	53.5044862
				YO42	53.925491	-0.8010515	48.42808765	77.9395264
				YO32	54.01889	-1.0914916	39.01835202	62.7956218
				YO31	53.967018	-1.0607451	42.76354925	68.8230929
				YO30	54.022093	-1.1572699	38.46518297	61.9053588
				YO25	53.999425	-0.4591241	50.95223433	82.0018548
				YO22	54.482391	-0.6093807	26.31581472	42.3523256
				YO21	54.485233	-0.6268204	25.58817501	41.1812718
				YO19	53.920829	-1.0654973	45.8828971	73.8433302
				YO18	54.245996	-0.7774743	29.57382519	47.5957247
				YO17	54.136037	-0.7971226	35.37274238	56.9284257
				YO16	54.089866	-0.2090223	53.62146977	86.2976872
				YO15	54.113452	-0.1501593	54.50793587	87.724354
				YO14	54.208928	-0.293673	45.95237047	73.9551398
				YO13	54.315349	-0.4485119	36.75145128	59.1473016
				YO12	54.276028	-0.4175671	39.22261065	63.124353
				YO11	54.282343	-0.4023018	39.5169135	63.5980001
				WV7	52.642739	-2.2812253	140.3038558	225.803178
				WV6	52.593405	-2.1457914	142.0189012	228.563349
				WV2	52.581437	-2.1246181	142.5970277	229.493778
				WV16	52.534316	-2.4197433	149.2249417	240.160656
				WV14	52.565208	-2.0735722	143.1724316	230.419826
				WS15	52.74699	-1.9272063	129.5497464	208.495656
				WS14	52.672922	-1.8348838	133.8256854	215.377295
				WS13	52.684572	-1.8281831	132.9829657	214.021033
				WS12	52.708784	-1.9690677	132.5077575	213.25624
				WS11	52.687057	-2.0284202	134.5346879	216.518355
				WR10	52.110565	-2.0747025	173.9226506	279.908823

Graph 1– Cluster Bar Graph

Introduction

- Percentage of present and absent vs start time graph by cluster bar graph

Analysis

- Overall student present or absent during the starting time of the course class
- During this implementation that I came to know that averagely, percentage of present is more than percentage of absent during module lecture
- Although at a particular time; percentage of absent is more compared to the percentage of present

Graph 1– Cluster Bar Graph

Data Cleaning

- I filter the data in excel by counting the percentage of students present and absent during the start time

- Sorting the data in csv converted to JSON and Percentage present and percentage absence vs start time graph by cluster bar graph

- Arranging the data in csv file for uploading into the D3jS and plotting the cluster bar graph

StartTime	Present_percent	Percentage Of ABS
09:00	63.15452904	36.84547096
10:00	64.24128937	35.75871063
11:00	64.09048652	35.90951348
12:00	55.64083519	44.35916481
13:00	64.0569395	35.9430605
14:00	59.4986554	40.5013446
15:00	65.67566464	34.32433536
16:00	54.101162	45.898838
17:00	41.07526882	58.92473118
18:00	69.81132075	30.18867925
19:00	64.22018349	35.77981651

Graph 1– Cluster Bar Graph

Link for Graph

https://q5038939.scedt.tees.ac.uk/ICA/ICA/MidhunTejaRavipati/Graph1.html

Conclusion

- As observed in graph, It is understandable that overall percentage present is greater percentage absent
- However, the absent percentage is more than present during the time 5pm
- Although, the highest percentage of present is marked at 6pm.

Graph 2 – Grouped Bar Chart

- Shows how number of classes per day impact attendance across different age groups

- Grouped bar chart chosen because:
  - It allows plotting 2 categorical variables against a numeric variable
  - The levels of each group are easily perceivable and comparable

- **Note:** Since timetable is based on individual students, the analysis used students as the unit and were aggregated across groups.

Graph 2 – Grouped Bar Chart

Key Findings

- The number of classes per day tend to affect attendance differently across age groups

- Single lecture days tend to have the least attendance

- Attendance tends to increase with increasing number of classes in the day

- Days with 3 and 4 lectures appear to be the most suitable across all age groups

- Days with more than 4 lectures per day show sharp decline in attendance for age groups above 34 years.

Link to the graph

Graph 3 – Sunburst Graph

- This graph shows the attendance of all students on all courses broken down by weeks and then days.

- This data was chosen to be displayed as it allow for a macro-micro view of which days/weeks are less/more attended.

- A sunburst chart was chosen to represent this data as it allows for a hierarchical view of the attendance for the days and weeks.

https://q5038939.scedt.tees.ac.uk/ICA/ICA/DanielHarding/index.html

Graph 3 – Sunburst Graph

- The notable areas of d3 used were:
  - The use of an arc generator in order to produce the path needed to draw the sections.
  - The use of ScaleOrdinal in order to handle colour

- Reason for the red segments:
  - Editorial salience, it was used to highlight any areas where attendance was lower than 1000.

Graph 3 – Sunburst Graph

- **What the graph shows:**
  - Wednesdays typically have the worst attendance.
  - Attendance overall decreases as the weeks advance into the semester.
  - Weeks 11 and 12 have low attendance with week 12 attendance being very low.

Graph 4 – Bubble Chart

- Plots number of classes per day and the total number of classes per student against their attendance level.
- This investigates whether courses that are more engaging have higher attendance level or not.

The choice of bubble chart based on its ability to:

- show three dimensions of data
- display lots of classes in a relatively smaller space
- enhance the recognition of associations between the variables

Graph 4 – Bubble Chart

Features of the chart:

- Animation

- Checkboxes to filter based on departments

- Mouse-hover highlights

- Tooltips

Graph 4 – Bubble Chart

Key Findings

- Programs with more classes have higher number of classes per day

- Programs with highest number of classes have lower attendance

- Computer science programs have the highest number of classes

- There doesn't seem to be a significant difference between undergraduate and postgraduate

- Computer Media, and Arts students tend to have the highest attendance

Link to the graph

Graph 5 – Line Chart

- This graph shows the average attendance percentage over the weeks.

- This data was chosen to give an overall view of attendance as the weeks progressed

- Attendance is highest at the beginning and in the 11th week of the semester

- Attendance is lowest at the week 2,week 7 and week 12 of the semester.

https://q5038939.scedt.tees.ac.uk/ICA/ICA/AthiraSatheesan/Graph1.html

Graph 5 – Line Chart

This graph was created using Plotly and as such, no difficulties using D3 were encountered

Average Attendance percentage by Weeks
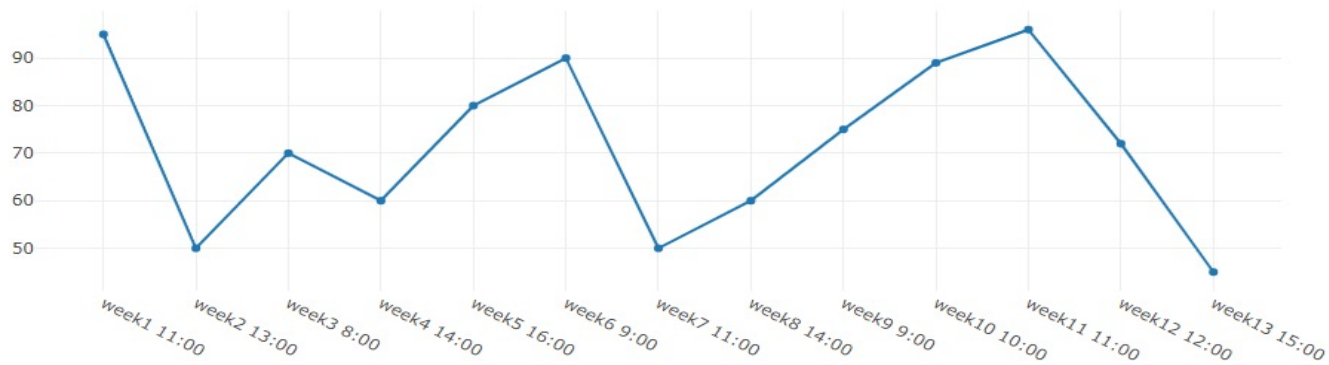
Graph 6 – Bar Graph

Introduction

- Total number of students for top 10 course by degree

Analysis

- After analysis, I have taken the top 10 courses in MSc and BSc where the student population is high.

- Majority of students in MSc have opted to Concept Art.

- All courses that are listed in top 10 belongs to Computer and Digital Field

Graph 6 – Bar Graph

Link for Graph

https://q5038939.scedt.tees.ac.uk/ICA/ICA/MidhunTejaRavipati/Graph2.html

Conclusion

- In MSc majority of students opted to Concept Arts
- In BSc for non-Computer course majority students opted to Fine Arts
- Comparative to other streams in arts majority of courses that are opted by students are related to Animation Graphics and Music technology

Graph 7 – Bar Graph

- The bar graph displaying the average student distance to the university by each course

- Choose a bar graph is based on:
  - Able to summarise representation of the whole topic easily
  - Made to simplify the comparison between set of data make it understandable to the target audience.
  - Very straightforward when the target audience read it.

- The link of the graph:
- https://q5038939.scedt.tees.ac.uk/ICA/ICA/YuShingLui/index.html

Graph 7 – Bar Graph

- Cause of using animation:
    - Draw the audience attention before explaining the graph

- Reasons to use tooltip:
    - Appear text labels when the user hovers over or touches the bar.
    - Provide specific information when the user looks at the specific bar.

Graph 7 – Bar Graph

- Result of the bar graph:
  - The average distance of MA Animation is incredibly high compare with other course
  - Mainly those courses of average distance in between 10 – 20 miles.
  - Over 50 miles differences of average distance compare with the highest and the lowest one

Graph 8 – Clustered Bar Graph

- The clustered bar graph displaying average distance from university and attendance against with start time

- Pros of clustered bar graph :
  - Able to make comparison of multiple data series per category
  - Can show changes over average distance

- Reasons to use mouse-hover highlights
  - Easy to draw the user attention
  - More interactive a graph

- The link of the graph:

https://q5038939.scedt.tees.ac.uk/ICA/ICA/YuShingLui/index.html

Graph 8 – Clustered Bar Graph

- Result of the clustered bar graph :
    - Almost the average distance of absent is higher than the average distance of present
    - Nearly the same of average distance at 14:00
    - Have a huge differences of these two status at 18:00 and 19:00
    - At 18:00 and 19:00, the average distance of present is significantly lower than the other time slot

Graph 9 – Clustered Bar Chart

- This graph shows the attendance and absence of the rooms used for teaching.

- This data was chosen to be displayed as it would highlight any rooms that have exceptional attendance or absence and this would allow for further investigation.

- A clustered bar chart was used as it allowed for the display of both the absence and attendance value of a room at the same time allowing for easy comparison.

https://q5038939.scedt.tees.ac.uk/ICA/ICA/DanielHarding/index.html

Graph 9 – Clustered Bar Chart

- The notable areas of d3 used were:
  - The use of two ScaleBands in order to properly scale the clusters of bars
  - The use of bbox in order to provide the dimensions of the text in the tooltip

Graph 9 – Clustered Bar Chart

- **What the graph shows:**
  - The main lecture halls OL1, G0.57 G0.54 and OL7 have the most attendance and absences.
  - That ICT1.01b and ICT1.01c have relatively low absences for their attendance levels
  - That A2.01, AG0.6 and T2.06 have no attendance

Graph 10 – Sunburst Chart

The sunburst chart showing average attendance of courses further broken down by modules.

This was chosen to provide a micro to macro view of attendance for specific courses and to allow for the identification of any potential problematic courses or modules within a course.

https://q5038939.scedt.tees.ac.uk/ICA/ICA/AthiraSatheesan/Graph2.html

Graph 10 – Sunburst Chart

This graph was created using Plotly and as such, no difficulties using D3 were encountered

INSTITUTION ACADEMIC PERCENTAGE

Modules and Courses attendance percentage

Sunburst Chart

Conclusion

In conclusion, by producing these interactive visualisations for the data provided we have highlighted the following areas for further investigation:

- Wednesday is the least attended day

- Attendance trails off the further into a semester you get

- There is a 'sweet spot' in the number of classes per day for encouraging attendance

- MA Animation students have the highest average distances compared with others

- Only the student living within 10 miles attend the class at 18.00 and 19.00

- It might be worth considering the living locations of students before scheduling classes

Q & A