Transport and Telecommunication Institute

Faculty of Engineering Science

Business Intelligence and Data Visualisation

Reflective Diary

Student: Gonzalo, Gamez

Student's ID: st83446

Study Group: 4303MDA

2024

Table of Contents

Initial Data Exploration

As my masters degree is almost done, I need to focus on the skills I need to secure a job here in Latvia. I wish to be a data analyst, scientist, or engineer and since there are more entry points as a data analyst, the skills in the class are the ones I need to showcase to transition my career into Data. Data visualizations have been an afterthought since I have been focusing on manipulating data or applying algorithms in the other class as shown in Figure 1. Since data analyst job postings include Tableau or Power BI, I needed to borrow fiancées Windows laptop to run Power BI since both softwares don't run on Linux or MacOS. The quick interactions and changes with using Power BI made it easier to handle for visualizations since Streamlit with python is more of a challenge to make graphical changes.

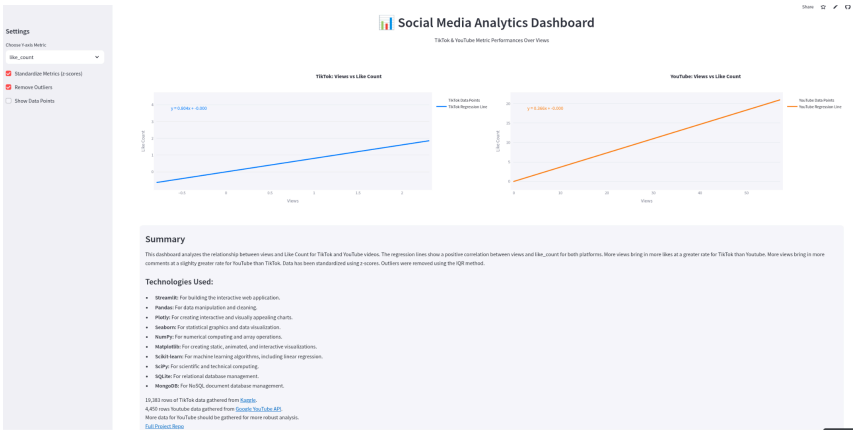
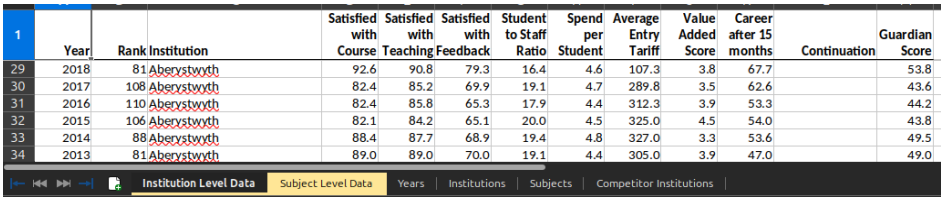


Figure 1. Dashboard created for Big Data Course Project using Streamlit and Jupyter notebooks

Many of my other assignments I have focused on social media since it is related to my masters thesis, though since I was an educator I chose the "Twelve Year Guardian League Table" dataset provided for us. Figure 2 shows an initial view of the dataset where if I was an educator

the metric I would be looking at is Student to Staff Ratio, but since I will be needing to think of stakeholders objectives the metric might be irrelevant.



1	Year	Rank	Institution	Satisfied with Course	Satisfied with Teaching	Satisfied with Feedback	Student to Staff Ratio	Spend per Student	Average Entry Tariff	Value Added Score	Career after 15 months	Continuation	Guardian Score
29	2018	81	Aberystwyth	92.6	90.8	79.3	16.4	4.6	107.3	3.8	67.7		53.8
30	2017	108	Aberystwyth	82.4	85.2	69.9	19.1	4.7	289.8	3.5	62.6		43.6
31	2016	110	Aberystwyth	82.4	85.8	65.3	17.9	4.4	312.3	3.9	53.3		44.2
32	2015	106	Aberystwyth	82.1	84.2	65.1	20.0	4.5	325.0	4.5	54.0		43.8
33	2014	88	Aberystwyth	88.4	87.7	68.9	19.4	4.8	327.0	3.3	53.6		49.5
34	2013	81	Aberystwyth	89.0	89.0	70.0	19.1	4.4	305.0	3.9	47.0		49.0

Institution Level Data | Subject Level Data | Years | Institutions | Subjects | Competitor Institutions

Figure 2. Initial view of data in LibreOffice

Clicking through the pages of the sheet I'm not sure how that tables interact and would not even be sure how it would be handled in a dataframe with jupyter notebooks. Loading data into Power BI, it triggers Power Query to deal with errors as shown in figure 3. Though looking at figure 4, the majority are from the column careers after 15 months which were originally empty. I made the decision to keep rows with the errors since it would remove the other data in the row. Figure 5 shows where it is to remove duplicates for identical rows in Power Query which was done to make sure that no value was being double counted in calculations.

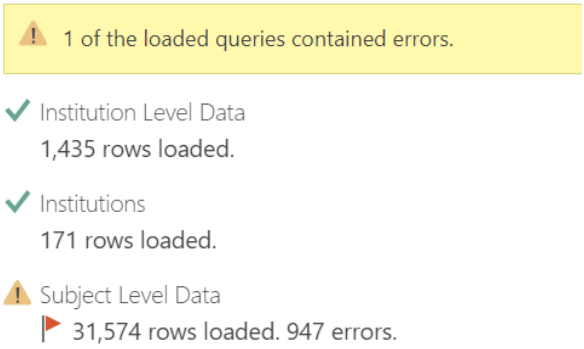


Figure 3. Errors loading into Power BI with Power Query

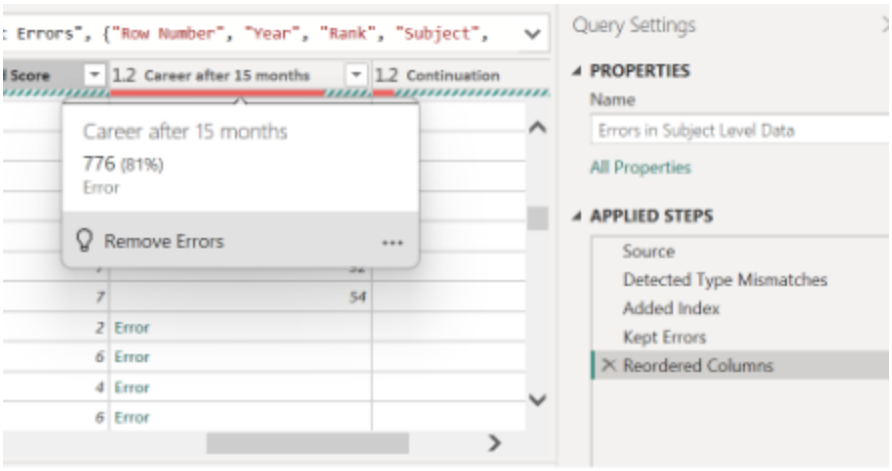


Figure 4. Majority of Errors form career

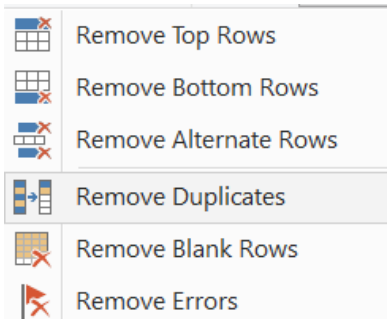


Figure 5. Remove Duplicates

Closing and applying the changes to return to Power BI. As suggested from our guest lecture by Vjačeslavs Matvejevs, I went to check the model view to get a better understanding of the data relationship as shown from figure 6. To prevent a many to many relationship between a Subject Level Data to Institution Level Data table, there was a years table created. Table of Subjects is not connected and the table of all Institutions has a one to many relationship to Competitor Institutions. Figure 7 shows what items are in Competitor Institutions, so columns are renamed to directly connect to Institution Level Data. Figure 8, shows the adjustments made for their competing institutions to be directly connected to Institution Level Data.

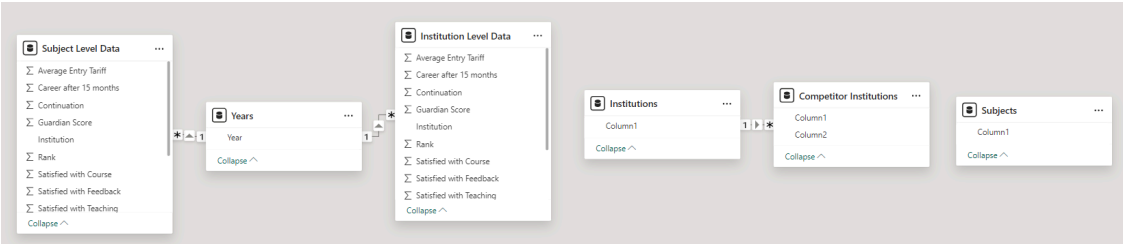


Figure 6. Initial Model View

Figure 7. Table view of competitor institutions

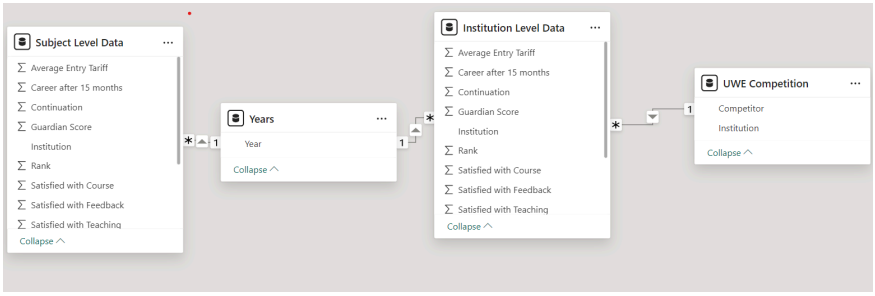


Figure 8. Model View Relationship adjustments

Understand the Business Requirements and Identify Stakeholders

As previously stated, I could take the viewpoint as an educator looking to teach at a university. I might be interested in it but this is meant to present to a stakeholder that may be for a university admin or a student looking to enroll in a university. The data set is set up to compare competitors of University of the West of England (UWE Bristol, UWE) as shown in figure 9. I'm assuming that students are asked to rank the universities they want to attend. Highly desired institutions are on the shortlist, longlist are those that are a lower priority, and UWE is included

on their list. Though this should have been a meeting with the stakeholder, I went with ChatGPT to decide what the data meant. Let's make the stakeholder is admin for UWE and a requirement is how and to whom the school should advertise or how it compares to the its competition.

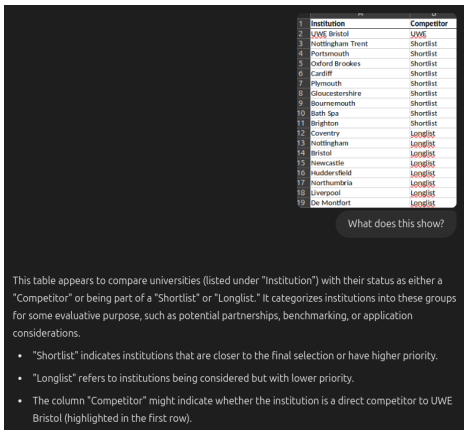


Figure 9. Asking about competitors ChatGPT(2024)

My initial assumptions would be that rank is better across the board, along with cost, and satisfied with teaching a key performance metric, but researching is required since my assumptions are US based and biased. Research was conducted for student enrollment into UK institutions. According to a study (Wilkins and Huisman, 2010) 160 international students studying at a UK university were surveyed on their choices for choosing to study overseas. Improved employment prospects was a dominant factor to study overseas by 91% answered by the students as shown from figure 10. The same study also had 98% as the reputation of the university was a factor for students' choice of institution.

Table 2. Factors Influencing Students' Decision to Study Overseas (as Percentages)

Factor	U/G ($n = 28$)	P/G ($n = 132$)	Chinese ($n = 60$)	Indian ($n = 21$)	European ($n = 40$)	All students ($n = 160$)
Difficult to gain place at home	7	5	7	0	5	6
Course not available at home	4	11	12	19	0	10
Lower quality of education at home	11	24	31	19	10	22
Experience a different culture	79	93	86	100	98	90
Improve my English	75	93	83	9	95	89
Improve employment prospects	86	92	86	100	95	91
Higher quality education overseas	39	73	69	81	48	67
Parental decision/ influence	75	61	67	57	58	63
Improve prospects for emigration	29	32	28	33	43	32

Note: U/G = undergraduate students; P/G = postgraduate students.

Figure 10. Table 2 from Wilkins and Huisman research (2010)

Another study (McManus, Haddock-Fraser and Rands, 2017) was conducted to investigate what factors are most important when choosing a university in the UK. This study found that factors such as league table ranking, employment prospects after graduation, and tuition fees are all important considerations for students. However, the importance of these factors can vary depending on whether the student is male or female, whether or not their parents attended university and if an international student. A case study (Rembielak, Rashid and Parlińska, 2020) about Polish students studying in British higher education institutes choose pull factors (reasons for choosing UK) compared to push factors (reason for leaving Poland). The pull factors noted were multicultural environment, reputation, improved employability, facilities, and scholarships.

Studies from the Quacquarelli Symonds (QS) a British company specializing in education and studying abroad also known for QS World University Rankings have recent surveys about students. Figure 11 shows that around the time the Covid Vaccine was released, some of the

most important factors for considering which country to study in were employment prospects and reputation of higher education institutions. An insight provided from 10,000 international students who graduated from 37 UK universities provided by QS (2024) over the past six years is shown in figure 12. 53% of respondents identified UK universities could improve in employment support.

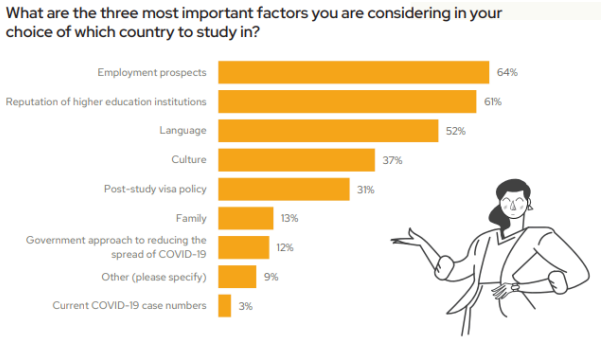


Figure 11. Survey of students from QS(2022)

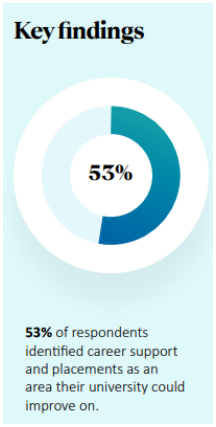


Figure 12. Insights of International Graduate Outcomes 2024 from QS

After this research, I will be using rank and score as reputation and looking at the metric Career after 15 months to help target advertisements to international students against UWE competition. I will compare how UWE has fared against the competition from 2011 to 2024 to advertise and if/what improvements need to be made. Figure 8 allows the comparison to be done in detail by subject, though I will only look at institutions data. Since data was only up to 2022,

figure 13 shows the missing data of UWE and competition. Figure 14 shows it loaded into Power BI with it, then with Power Query the Table was stacked onto Institution Level Data.

	Year	Rank	Institution	Satisfied with Course	Satisfied with Teaching	Satisfied with Feedback	Student to Staff Ratio	Spend per Student	Average Entry Tariff	Value Added Score	Career after 15 months	Continuation	Guardian Score
2	2024	94	Bath Spa		79.9	75.7	19.6	3.1	123	6.1	70	91.3	52.4
3	2024	99	Bournemouth		76.1	64.3	18.3	3.3	119	5.6	81	89.9	51
4	2024	70	Brighton		74.9	63.3	17	4.8	122	6.8	79	89.2	56.7
5	2024	17	Bristol		79.8	58.2	13.9	7.4	174	5.5	88	96.6	69.8
6	2024	29	Cardiff		80	62.6	13.9	6.4	153	4	87	94.5	64.2
7	2024	46	Coventry		83.6	70.3	13	n/a	126	6.9	78	85.5	60.2
8	2024	117	De Montfort		73.3	65.3	18.3	3.3	114	4.9	75	87.4	40.6
9	2024	104	Gloucestershire		76.8	66.3	15.4	5.7	121	2.3	75	89.7	48.2
10	2024	79	Huddersfield		76.4	73	14.6	5.1	130	6.2	75	87.9	55.2
11	2024	36	Liverpool		78.4	58.9	13.6	7.4	147	6.1	84	95.9	62.1
12	2024	67	Newcastle		76	57	14.3	6.4	150	4.4	86	95.3	57.3
13	2024	38	Northumbria		76.4	70.6	15.5	4.8	143	6.6	82	89.2	61.3
14	2024	59	Nottingham		79.5	59.2	15.7	6	155	4.5	87	95.7	58
15	2024	42	Nottingham Trent		80.8	73.8	15.1	4.4	128	4.1	77	92.5	60.9
16	2024	51	Oxford Brookes		76.2	66.4	13	4.8	118	5	80	93.6	59.3
17	2024	69	Plymouth		79.8	67.6	16.7	5.2	135	5.3	83	91.5	57
18	2024	33	Portsmouth		80.2	71.6	15.8	7.5	127	5.6	79	90.7	62.7
19	2024	43	UWE Bristol		80.2	68.5	16.8	4.5	129	5.7	80	90.2	60.7
20	2023	89	Bath Spa	74.3	79.2	73.6	16.6	5.5	125	4.6	66	92.7	64.2
21	2023	108	Bournemouth	66.2	73.8	62.2	19.2	4.7	116	4.7	78	93.5	60
22	2023	92	Brighton	64.5	74	62.9	18.1	5.3	116	6.2	77	91.4	63.9
23	2023	16	Bristol	74.6	81.1	58.8	14.1	7	167	5.7	85	98	76.2
24	2023	35	Cardiff	74	79.4	61.2	14.5	6.6	148	4.5	85	96.1	71.4
25	2023	51	Coventry	77.6	82.6	70.3	13.8	6.5	119	6.6	76	88.6	69.4
26	2023	120	De Montfort	67.2	72.1	63.5	18.5	3.1	109	4.9	72	89.6	53.4
27	2023	88	Gloucestershire	73.7	77.9	69.1	17.2	5.6	120	4.4	71	90.2	64.3
28	2023	78	Huddersfield	71.7	75.1	70.8	14.7	4.5	126	5.9	71	92.3	65.9
29	2023	40	Liverpool	75.2	80.2	63	13.8	7.6	144	5.7	80	96.9	71.1
30	2023	63	Newcastle	68.8	76.5	56.9	14.7	5.8	148	5.1	82	97.1	68.2
31	2023	46	Northumbria	70.9	75.8	68.9	15.7	4.2	135	6.5	79	92.4	70
32	2023	63	Nottingham	75.2	80.5	60.2	15.8	6.5	152	4.2	85	97	68.2
33	2023	62	Nottingham Trent	77.5	79	71.8	15.3	4.3	123	3.5	72	94.6	68.4
34	2023	42	Oxford Brookes	73.1	76.4	64.6	14	4.8	122	4.9	78	95.6	70.8
35	2023	63	Plymouth	75.2	80.2	69	17.3	5.1	129	6.6	79	94.5	68.2
36	2023	67	Portsmouth	74.7	77.9	69.4	16.8	6.7	116	5.9	75	93	67.8
37	2023	24	UWE Bristol	76.9	80.5	69.2	15.7	4	123	6.5	79	92.7	72.8

Figure 13. LibreOffice of data missing for 2023 and 2024

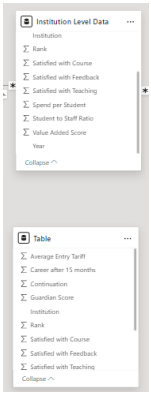


Figure 14. Table created with missing data

After talking with the stakeholder and showing the research that was conducted, we agreed to table 1 as stakeholder concerns. Assuming that I can use rank and score as reputation.

Stakeholder	Concern
UWE admin	How do we compare to our competing universities in rank, score, and career over the years 2011-2024? If/what needs improvement against competition?
International Students	Why should I choose to attend UWE?

Table 1. Stakeholder Concerns

Visualize the Dataset to Answer Business Questions

I will be using averages across the years to make a numerical comparison against UWE and the competition. Creation of columns in the dataset is possible for aggregation, but Data Analysis Expressions (DAX) is able to do aggregation in Power BI. Figure 15 shows the start of the use of many aggregations since they can provide measures for the year.

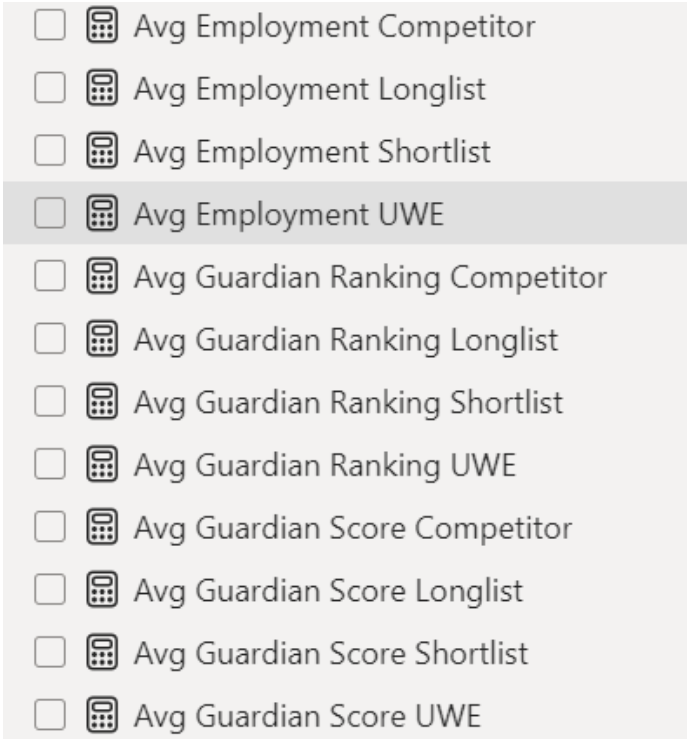


Figure 15. Several metrics created

Referring to course materials, line charts are used to show trends across continuous measurements. Since we will compare across years of the competition, line charts will be used. Figure 16 shows all competitors, those on the longlist and those on the shortlist against on the line all together. It is possible to focus the audience's attention while adjusting color and dashes, but there are lines overlapping and very close to each other. The choice was made to only compare one to one to have visual clarity along with DAX to enable toggling over choices.

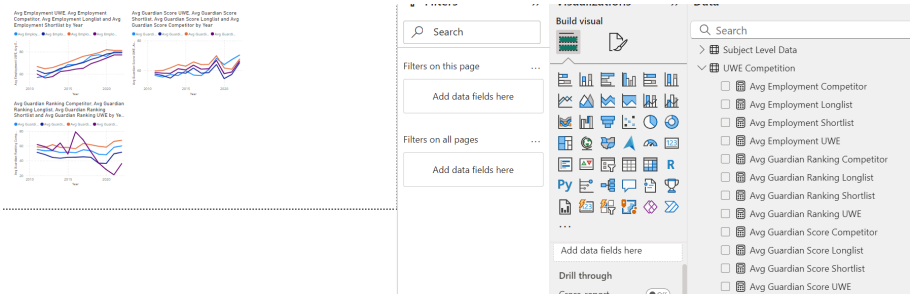


Figure 16. Comparing with all metrics

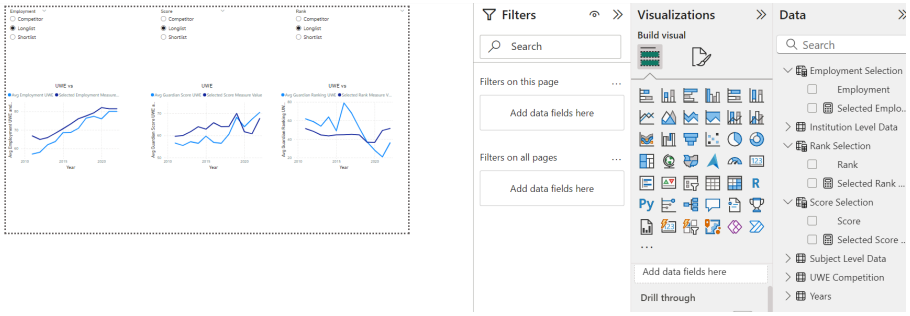


Figure 17. Adjustments to compare 1 to 1 dynamically

Since UWE colors are red, I could have used it as the UWE color of red but it may not be a great color throughout the whole dashboard since red does signal implicit themes. For accessibility, accessible default will be used throughout the dashboard. To show the competitor institutions, the competitor institution table was included as shown in figure 18. Figure 19 shows a multi card would not be suitable since the lack of clarity about the institution.

Figure 18. Table inserted to show competitors along with graph adjustments

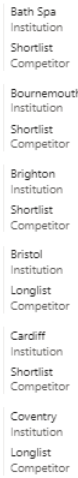


Figure 19. Multicard for competitors

Figure 20 shows cards being constructed to display averages over the years depending on whichever clicker is chosen. A couple of Gestalt Principles and preattentive attributes were considered to adjust information to for presentation on the dashboard. Proximity, connection, and symmetry were to lay out employment, rank, and score to show cards and graphs are related to each other. Figure 21 using similarity and keeping accessibility in consideration, competition is displayed as a dashed line throughout the graphs. Titles, legend, and clicker are adjusted to be descriptive when choices are changed from DAX measurements. Spatial positioning was adjusted for cards with UWE and competition with left and right as shown in figure 22.

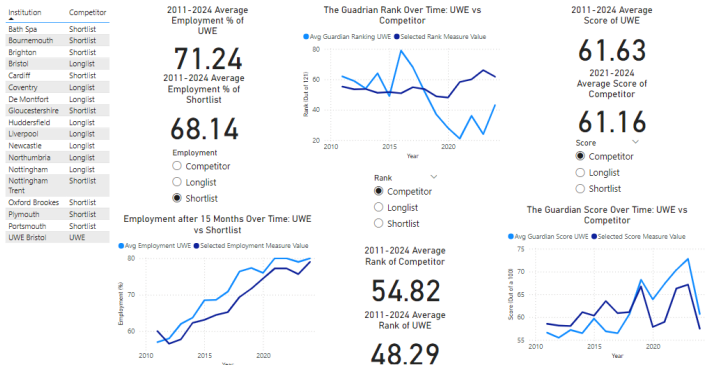


Figure 20. Cards to show average across all years

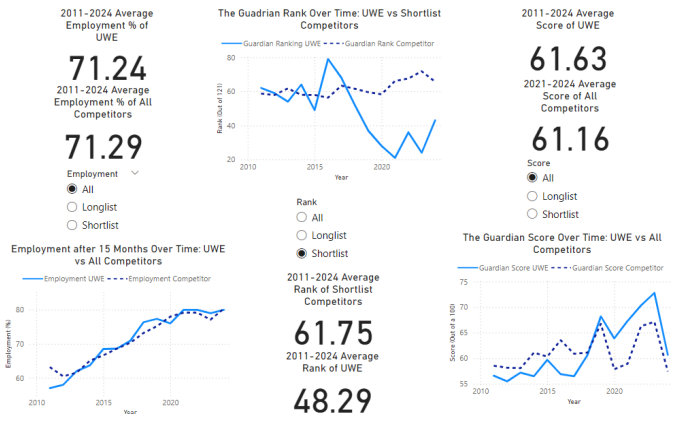


Figure 21. Dashed lines to show competitors

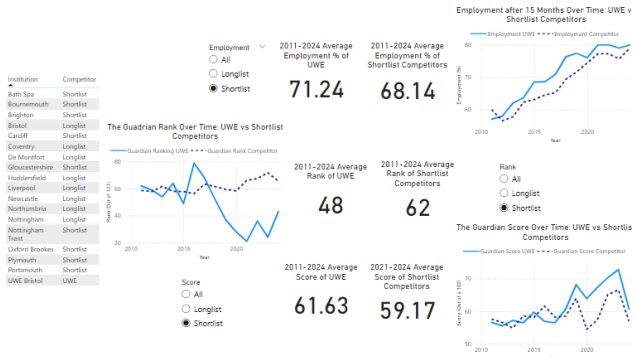


Figure 22. Graphs and cards adjusted to make easier comparison

Refinement with the visuals continues with colors with same length and width of the corresponding cards and clickers grouping together in the same location as seen in figure 23. UWE logo added to give a quick visual que it is to be used by UWE. Figure 24 objects are enclosed to show that they are their own object with their proximity showing they are related.

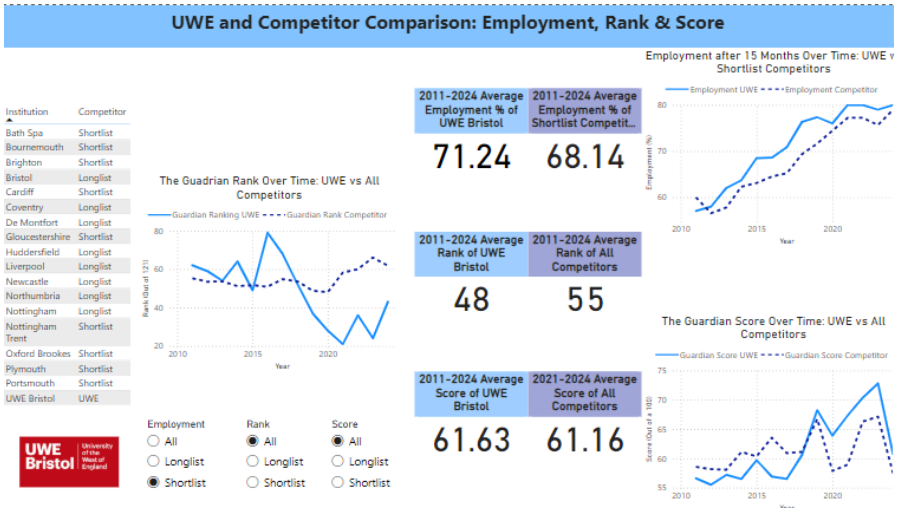


Figure 23. Title, colors, and UWE logo added, clickers aligned

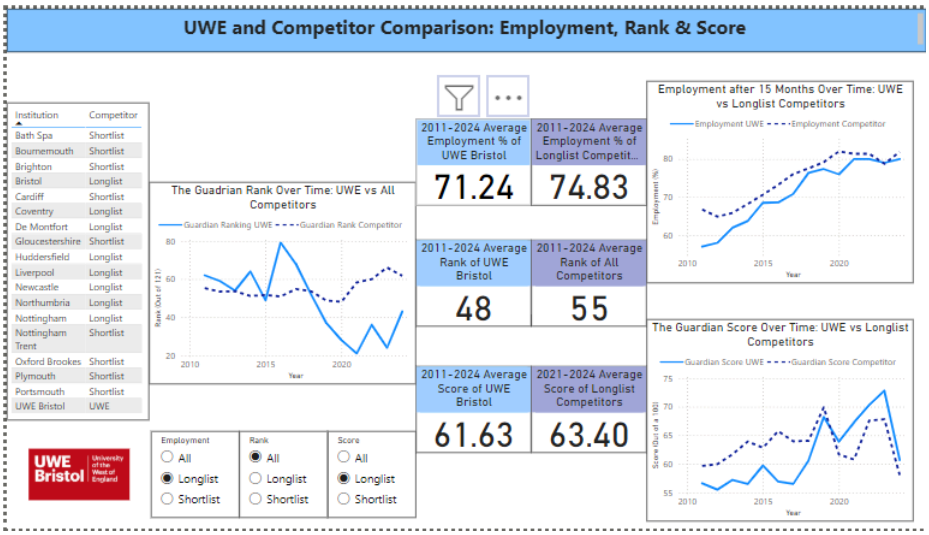


Figure 24. Enclosing objects

Figure 25 shows the background changed to be gray to reduce eye strain and improve contrast for the graphs with the colors present on it. All graphs were giving 4:3 ratios with the same width and length since not one singular graph is the focus.

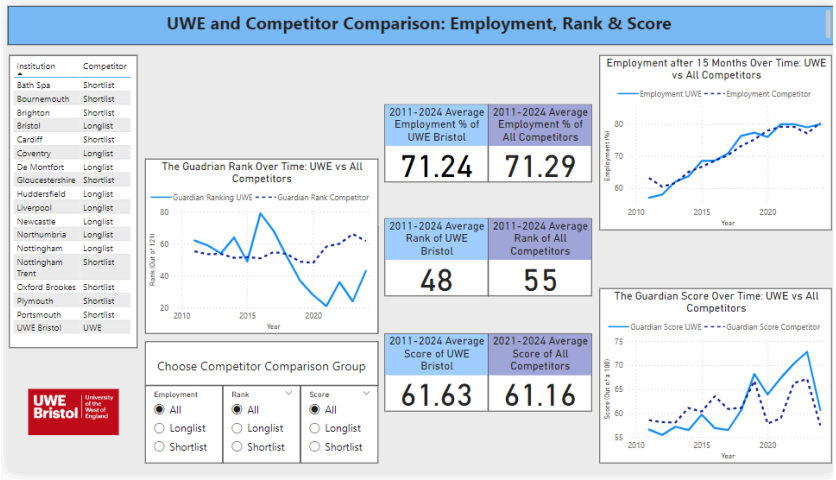


Figure 25. Background colors and making graph 4:3 ratios and title of clickers

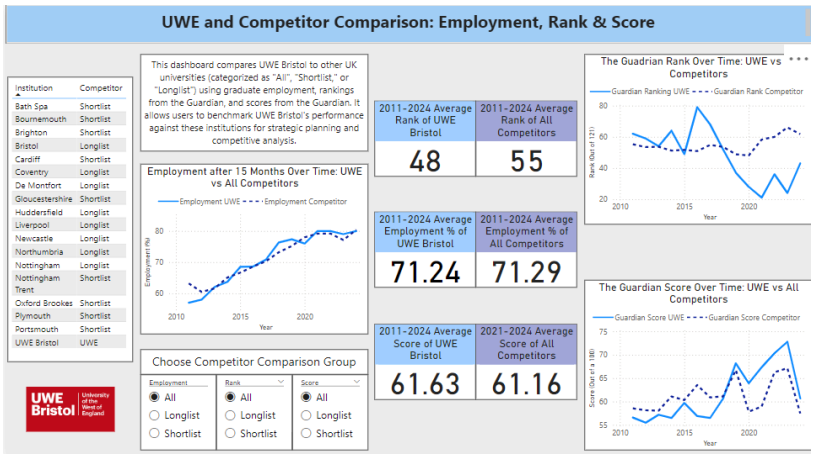


Figure 26. Possible final layout

Figure 26 shows a possible final layout of the dashboard. With it, it is time to think like a designer and consider the affordances with clicking on the dashboard. Clicking on graphs and tables has unintended filtering across graphs and cards. Figure 27 shows a portion of adjusting of interactions and only making sure clickers gives the user control to change. For the table, an invisible rectangle was placed on the table to still give the user a chance to reorder the list only as shown in Figure 28. Annotations may be put in the final version to visually put for positive and negative difference between competition, but due to time constraints it may be done later.

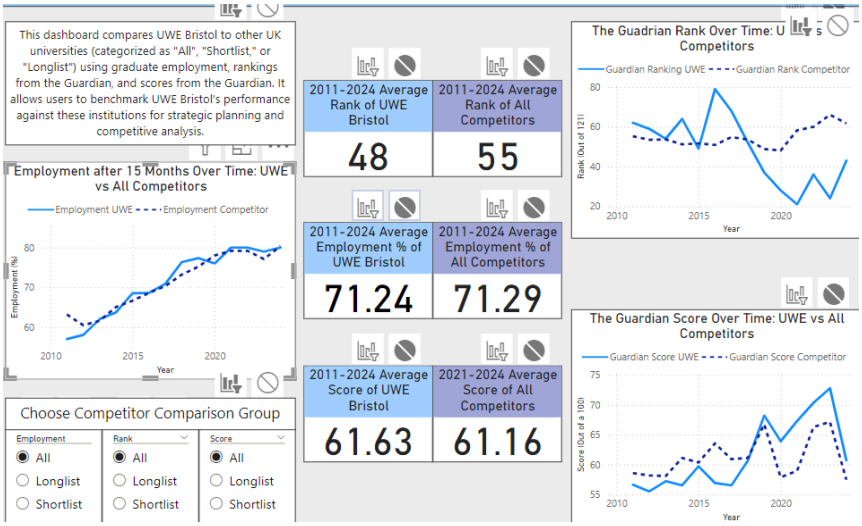
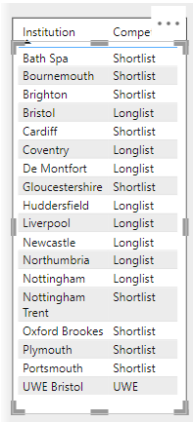


Figure 27. Disabling interactions



Institution	Compe
Bath Spa	Shortlist
Bournemouth	Shortlist
Brighton	Shortlist
Bristol	Longlist
Cardiff	Shortlist
Coventry	Longlist
De Montfort	Longlist
Gloucestershire	Shortlist
Huddersfield	Longlist
Liverpool	Longlist
Newcastle	Longlist
Northumbria	Longlist
Nottingham	Longlist
Nottingham Trent	Shortlist
Oxford Brookes	Shortlist
Plymouth	Shortlist
Portsmouth	Shortlist
UWE Bristol	UWE

Figure 28. Invisible rectangle over competitor table

At this point I was ready to show the client the dashboard from figure 29, but after writing my reflection and looking back on coursework I could not make a justification with why I have the graphs in their position. The original layout had me having the graphs stack, but figure 30 shows me a different layout that is more symmetrical. Figure 31 dashboard was created for symmetry along with grouping of values to corresponding graphs. Visually the table was adjusted to make it fit on the dashboard with all values. Later adjustments may be made for clarity for separation of values.

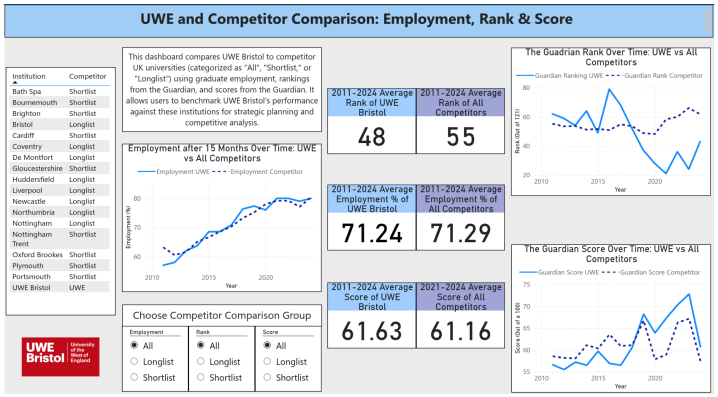


Figure 29. Possible final draft of dashboard

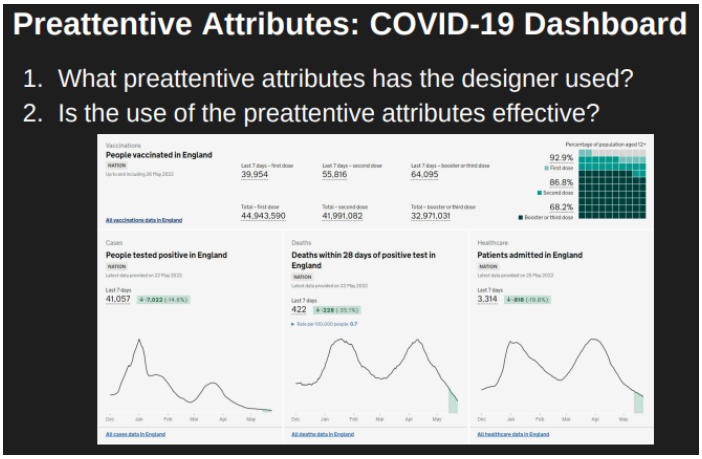


Figure 30. Slide from class presentation (Spiridovska, 2024)

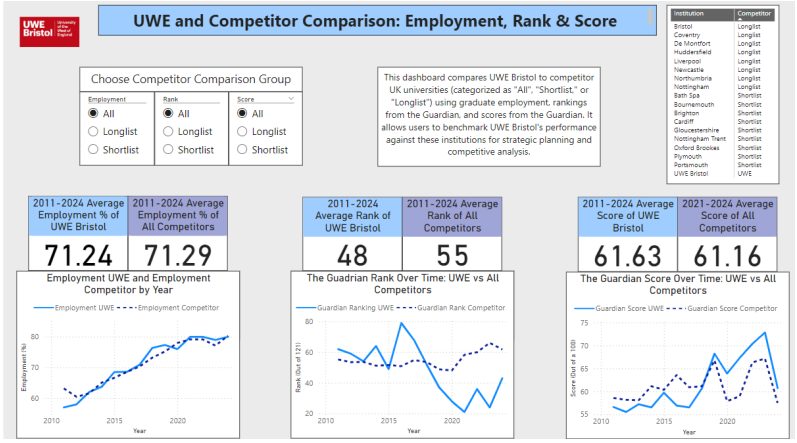


Figure 31. 2nd possible final draft of dashboard

Dashboards for figure 29 and 31 were both shown to the stakeholder, with the 2nd being acceptable in using.

Develop a Narrative Telling a Story Around the Data

The presentation will be following the discovery journey method going through what I have reflected so far. The powerpoint will be developed and presented in later time so the final presentation may be slightly different. The following is the general flow of how the story will unfold and connect with thoughts being written down in a loose structure.

Intro: Prompting and asking why the audience chose this university or for undergrad. Might give possible reasons or even some of my own.

Setup scenario: How would UWE attract you (a UK university). UWE has a list of competitions which is shown from figure 28. May add how shortlist and longlist may be identified. Along with talking about data about the best UK universities ranking provided by The Guardian. We have data from 2011 to 2024, but first does it matter to use this data?

Research: Talk about research that is specific from the UK since we may have assumptions for Latvia and the US. Describe over the years from how reputation and employment matters across to international students, might be presented in a timeline. Using table 1 as what are stakeholder concerns after presenting to them.

Talking about data: Now looking at it, I will use annotations for the final draft to show the comparison of positive and negative differences. Descriptions will be added at bottom of the presentation for accessibility. Figure 32 shows vs shortlist competitors, UWE is better in these metrics. Figure 33 shows against all competition we are not the best. Figure 34 shows that actually those on the longlist are the better performing schools. This goes against an assumption I had that competition on the longlist were on lower ranking since they weren't as good, but might lower since they are a long shot of being accepted.

Recommendations for Future Improvements: In final section and show table 1 once again.

QA

A practice run through will be done after the presentation is created to see if the story is coherent and within the 10 minute time limit.

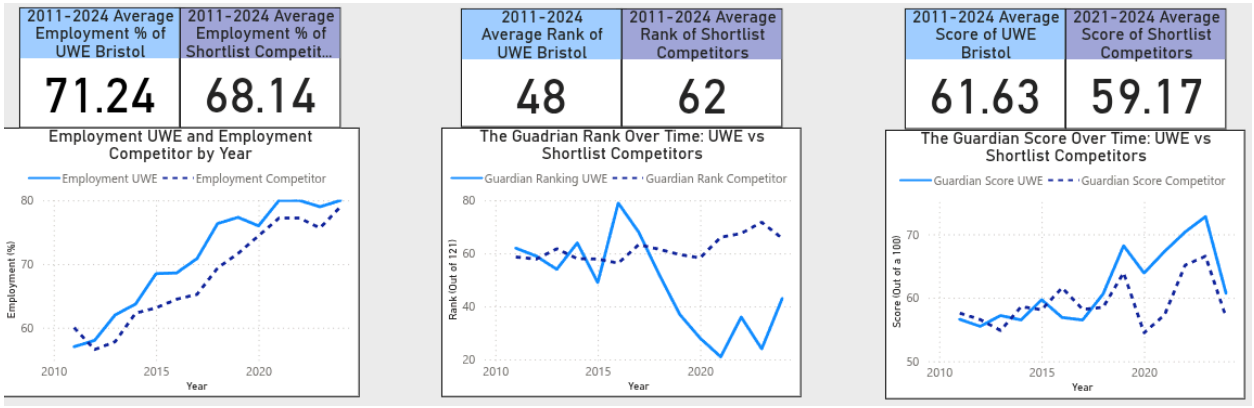


Figure 32. UWE vs shortlist Competitors

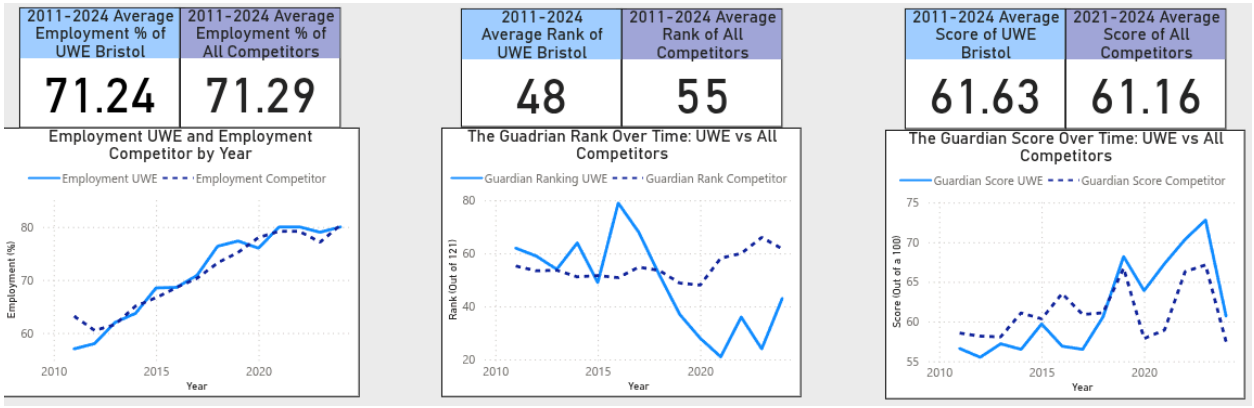


Figure 33. UWE vs All Competitors

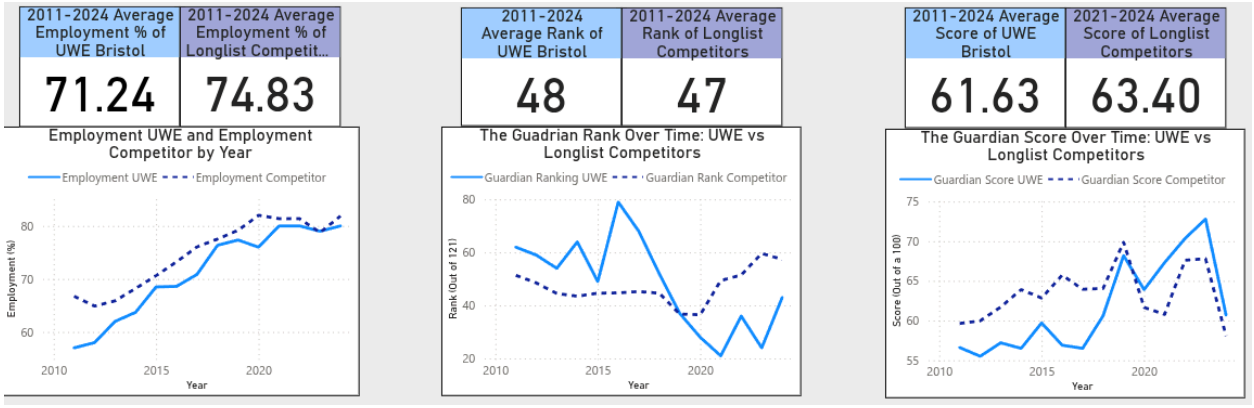


Figure 34. UWE vs shortlist Competitors

Identify Recommendations for Future Improvements

Looking at the data, employment is the metric that needs improving against our competitors since year over year we trade positions between the higher percentage. UWE should compare against longlist competitors since they are the better performing schools that are in consideration of prospective students. Dashboards could be created for employment and ranking based on subjects to see those below average of the institution and need improvement.

Targeting international students, advertising that for the past 7 years UWE has ranked higher against competing schools. On average UWE has better employment outcomes compared to institutions that were shortlisted since 2011. Though having a marketing team do focus testing to see what makes it sound more appealing.

Final reflection

After reflecting about the final presentation and writing my thoughts down, colors can be different for “All”, “shortlist”, and “longlist”. An annotation of shaded areas between UWE and competitors to show when and by how much better/worse UWE is doing. Title needs to be adjusted for one graph and table may be adjusted to keep it small but show entries as distinct or might split into two tables. All these adjustments should make it more apparent for the viewer to understand the data quicker. Further iterations will be done to the dash by the time the presentation is done.

Dashboard Link

Reference list

Chatgpt.com. (2024). *ChatGPT*. [online] Available at: https://chatgpt.com/c/67701ae3-137c-8008-80f4-4b801c2a24da [Accessed 22 Dec. 2024].

McManus, R., Haddock-Fraser, J. and Rands, P. (2017). A methodology to understand student choice of higher education institutions: the case of the United Kingdom. *Journal of Higher Education Policy and Management*, 39(4), pp.390–405. doi:https://doi.org/10.1080/1360080x.2017.1330806.

QS (2022). *QS Higher Education Briefing: Is COVID-19 still impacting student decision-making?* [online] *QS.com*. Available at: https://insights.qs.com/hubfs/Reports/CRE-904%20_May%202022-Briefing_V2%20(2).pdf?utm_medium=email&_hsmi=65742003&utm_content=65742003&utm_source=hs_automation.

QS (2024). *International Graduate Outcomes 2024*. [online] Available at: https://26055784.fs1.hubspotusercontent-eu1.net/hubfs/26055784/Reports/International%20Graduate%20Outcomes%20Report%202024.pdf [Accessed 2 Jan. 2025].

Rembielak, G., Rashid, T. and Parlińska, A. (2020). FACTORS INFLUENCING STUDENTS' CHOICES AND DECISION- -MAKING PROCESS: A CASE STUDY OF POLISH STUDENTS STUDYING IN A BRITISH HIGHER EDUCATION INSTITUTION. *Acta Scientiarum Polonorum. Oeconomia*, 19(3), pp.85–95. doi:https://doi.org/10.22630/aspe.2020.19.3.31.

Spiridovska, N. (2024). *BIDV_UWE | e-TSI*. [online] E.tsi.lv. Available at: https://e.tsi.lv/pluginfile.php/204168/mod_resource/content/9/5_Preattentive%20Attributes_BI_DV.pdf [Accessed 4 Jan. 2025].

Wilkins, S. and Huisman, J. (2010). Student Recruitment at International Branch Campuses: Can They Compete in the Global Market? *Journal of Studies in International Education*, 15(3), pp.299–316. doi:https://doi.org/10.1177/1028315310385461.