## Are Top Universities More Attractive To Tech Startups?

## Introduction

Do we see more tech startups around the top Universities in the UK? There is evidence that tech startups cluster together and become incubators for innovation but does that also correlate with where certain Universities are located?

The purpose of this project is to find out if certain Universities have a higher proportion of tech startups in close proximity and the similarities between those Universities.

## Data

To look at this problem we are going to need two data sets:

- A list of UK universities with ranking data
- A count of tech startups in close proximity to each University

For the university ranking data we will use The Times Higher Education World Rankings data

To get tech startup proximity data we will use the FourSquare API

UK university ranking data

The data is not easy to scrape from the THE website, luckily this data has already been collected by a contributor on kaggle

The data can be found in <a href="https://www.kaggle.com/joeshamen/world-university-rankings-2020/download">https://www.kaggle.com/joeshamen/world-university-rankings-2020/download</a>

Credit must be given to THE for producing this data and Joe Shamen for making it available in this form.

To make the data useful for our purposes we drop columns we are not interested in. Remove the rank field because we will use the index instead.

The data is formatted and then scaled by dividing each number by 100 so we can compare them later on. This leave us with a dataset that looks like this:

	University	International_Students	Percentage_Female	Percentage_Male	Teaching	Research	Citations	Industry_Income	International_Outlook
0	University of Oxford	0.41	0.46	0.54	0.905	0.996	0.984	0.655	0.964
1	University of Cambridge	0.37	0.47	0.53	0.914	0.987	0.958	0.593	0.950
2	Imperial College London	0.56	0.38	0.62	0.845	0.876	0.970	0.699	0.971
3	UCL	0.52	0.57	0.43	0.778	0.887	0.961	0.427	0.962
4 Lo	ondon School of Economics and Political Science	0.71	0.53	0.47	0.690	0.830	0.928	0.351	0.932

## Tech Startup Data

Before we can get the count of tech startups near each University, we need to find the location data for each University. For that we use the geopy package and the Nominatim geocoder.

To ensure the location data is valid, two levels of validation were performed. First we check if any rows have null data for location, and that was found. The missing data was entered manually. Secondly we validate the locations are correct by presenting the Universities on a folium map as shown below.



Using the location data, we use the FourSquare search API to find tech startups within a kilometre of each University. Any Universities that returned an empty venues list were updated with a count of 0. This count is then stored in a new dataframe.

Divieratly of Cultorid  University of Clambridge  I University of Clambridge  I University of Clambridge  I University of Clambridge  I University of Cambridge  I University of Cambridge  I University of Clambridge  I Universi	Tech_Cnt
1	6
2   Imperial College London	17
3	15
5 4 January of Pathical Science 6 5 University of Editoropy 7 6 Nog•5 Childege Lordon 7 University of Banchester 9 6 University of Manchester 10 9 University of Manchester 11 10 University of Ollappe 12 11 Many University of London 13 12 University of Ollappe 14 13 University of Ollappe 15 14 University of Shaffield 16 15 University of Shaffield 17 University of Shaffield 18 17 Lancaster University 19 19 University of Editor 20 University of Editor 21 University of Editor 22 University of Editor 23 University of Editor 24 University of London 25 University of London 26 University of London 27 28 University of Lancaster 28 29 University of Lancaster 29 20 University of Shaffield 20 21 University of Shaffield 20 22 University of Shaffield 21 University of Shaffield 22 University of Shaffield 23 24 University of Shaffield 24 University of Shaffield 25 25 University of Shaffield 26 27 University of Shaffield 27 28 Newcastle University 28 29 University of Manches 29 20 University of Shaffield 30 30 University of Floading	47
6   University of Estitution 7   0   King-Spt College London 9   University of Manchester 9   0   University of Manchester 10   0   University of Manchester 11   10   University of Ossagou 12   11 May University of Collegou 13   12   University of Collegou 14   13   University of Standard 15   14   University of Standard 16   15   University of Standard 17   16   Durham University 18   17   Lancaster University 19   18   University of Standard 20   19   University of Standard 21   University of University 22   University of University 23   University of Lesder 24   25   University of Lesder 25   26   University of Lesder 26   27   University of Lesder 27   28   University of Lesder 28   29   University of Lesder 29   20   University of Lesder 20   21   University of Lesder 21   22   University of Lesder 22   23   University of Lesder 24   25   University of Lesder 26   27   University of Estal Angla 27   University of Standards 28   Newcastle University 29   20   University of Standards 30   20   University of Reading 31   University of Standards	45
7 6 Noge College London 8 7 Velowestry of Watercalet 9 8 University of Watercalet 10 9 University of Watercalet 11 10 University of Entered 12 11 Many University of Entered 13 12 University of Standard 14 13 University of Standard 15 14 University of Standard 16 15 University of Standard 17 16 University of Standard 18 17 Lancaster University 19 10 University of Entered 20 10 University of Entered 21 20 University of Entered 22 21 University of Extered 23 22 University of Lenedard 24 23 University of Lancaster 25 24 University of Lancaster 26 25 University of Lancaster 27 28 Cardiff University 28 29 University of Extered 29 27 University of Extered 29 28 Newcaster University 29 29 University of Extered 29 29 University of Extered 29 20 University of Extered 20 21 University of Extered 20 22 University of Extered 21 23 University of Extered 22 23 University of Extered 23 24 University of Extered 24 25 University of Extered 25 26 University of Extered 26 27 University of Extered 27 28 University of Extered 28 29 University of Extered 29 20 University of Extered 20 21 University of Extered 20 22 University of Extered 21 22 University of Extered 22 23 University of Extered 23 24 University of Extered	30
8 7. University of Manchester 9 8 University of Valencia 10 9 University of Gisepow 11 10 University of Gisepow 12 11 Mary University of Casepow 13 12 University of Casepow 14 13 University of Samengham 15 14 University of Samengham 16 15 University of Samengham 17 16 Duraham University 18 17 University of Suntamenton 18 19 University of Suntamenton 19 University of Suntamenton 19 University of Suntamenton 20 19 University of Suntamenton 21 22 University of Suntamenton 22 23 University of University 23 24 University of University 25 26 University of Lexedor 26 27 University of Lexedor 27 28 University of Lexedor 28 29 University of Lexedor 29 20 University of Sandrese 20 20 University of Sandrese 21 22 University of Sandrese 22 23 24 University of Sandrese 23 25 University of Sandrese 24 26 Caseful University 25 26 Caseful University 26 27 University of Sandrese 27 28 Caseful University of Sandrese 28 29 University of Sandrese 39 20 University of Sandrese 30 20 University of Sandrese 31 30 University of Dundee	44
9 8 University of Various: 10 9 University of Bottot 11 10 University of Classper 12 11 Mary University of Classper 13 12 University of Classper 14 13 University of Standard 15 14 University of Standard 16 15 University of Standard 17 16 University of Standard 18 17 Lancaster University 19 10 University of Exister 20 10 University of Exister 21 20 University of Exister 22 21 University of Exister 23 22 University of Lesses 24 23 University of Lesses 25 24 University of Lesses 26 25 University of Lesses 27 20 University of Standard 28 29 University of Standard 29 20 University of Standard 29 20 University of Standard 29 20 University of Standard 20 20 University of Standard 20 20 University of Standard 21 University of Standard 22 20 University of Standard 23 21 University of Standard 24 University of Standard 25 26 University of Standard 26 27 University of Standard 27 28 University of Standard 28 31 University of Standard	39
10	4
10	30
11 May University of London 13 12 University of Smitheld 14 University of Smitheld 15 4 University of Smitheld 16 15 University of Smitheld 17 15 University of Smitheld 18 17 Lancaste University 19 18 17 Lancaste University 19 19 University of Existe 20 19 University of Smitheld 21 20 University of Smitheld 22 21 University of London 23 22 University of London 24 23 University of London 25 24 University of Lendon 26 25 University of Lendon 27 26 University of Lendon 28 29 20 University of Lendon 29 20 Candidate 20 21 University of Lendon 20 22 University of Smitheld 20 22 University of Smitheld 21 22 University of Smitheld 22 23 University of Smitheld 23 24 University of Smitheld 24 25 University of Smitheld 25 26 University of Smitheld 26 27 University of Smitheld 27 28 University of Smitheld 28 29 20 University of Smitheld 29 20 University of Smitheld 20 University of Smitheld 20 31 University of Dmitheld	11
12	11
15	0
15	4
15	- 1
15	2
19	0
19	- 1
10	2
21 20 University of Notingquam 22 21 University of Leeds 23 22 University of Leeds 24 23 University of Levespool 25 24 University of Accretice 26 25 University of Accretice 27 26 Card University of Set Angle 28 27 University of Set Angle 29 20 Newcastle University 30 20 ⊔een-© Lutiversity 31 30 University of Reading 31 31 University of Dundee	2
22 21 University of Leeds 23 22 University of Leeds 24 22 University of Leeds 25 24 University of Leeds 26 25 University of Each Angla 27 26 Cardff University 28 27 University of Start Angla 29 27 University of Start Angla 20 28 Newcastle University 30 29 University Only of Reading 31 30 University of Reading 32 31 University of Disordes	2
22 22 University of Liversod 24 23 University of Liversod 25 24 University of Alexdose 26 25 University of Set Angla 27 26 Conf University 28 27 University of St Andrews 29 20 Newcaste University 30 29 usen \$\tilde{\text{Ps}}\$ University of Reading 31 30 University of Reading 32 31 University of University	3
24 2) University of Laccester 25 24 University of Laccester 26 25 University of East Anglia 27 26 Cardiff University 28 27 University of St Andrews 29 28 Newcastle University 30 29 University Of Medicing 31 30 University of Reading 32 31 University of Officeden	4
25 24 University of Aberdeen 26 25 University of East Angla 27 26 Card University 28 27 University of St Andrews 29 20 Newcaste University 30 20 ween ◆ University of Reading 31 30 University of Dundee	10
26	1
27         26         Cardiff University           28         27         University of SI Andrews           29         28         Neversity Of SI Andrews           30         20 users@u University of Selfast           31         30         University of Reading           32         31         University of Dundee	- 1
28         27         University of St Andrews           29         28         Newcastle University           30         29 usen-6s: University Beflast           31         30         University of Reading           32         31         University of Dundee	13
29 28 Newcastle University 30 29 user⊕s University Beffast 31 30 University of Reading 32 31 University of Dundee	- 1
30         29 ueen●s University Belfast           31         30         University of Reading           32         31         University of Dundee	25
31 30 University of Reading 32 31 University of Dundee	17
32 31 University of Dundee	4
	9
22 aAe Univareity of London	2

We now have the datasets required in formats that we can use in the analysis phase. To avoid having to repeat the exercise the datasets were saved to csv files.