FDS Final Project
Variability of UK Higher Education grades

Mrinmoy Sonowal and Kuan Lon Vu

13th April 2021

1 Overview

For several years, grade inflation in higher education degree classification has been a topic of great interest for many stakeholders in the public, government, education sector and industry. It is worried that a continuous increase in the proportion of the top degrees would be unfair for hardworking students from previous years [1]. In this report, we investigated the supposed increase in proportion of 'First class' degrees per year and the odds of receiving First class within different demographic groups. We found that the increase in proportion of 'First class' degrees in 2019/20 is statistically significant. There was no statistically significant increase from 2014/15 to 2016/17 and from 2017/18 to 2018/19. The statistical significance was tested using the 95% confidence interval (CI). We studied the odds being awarded a 'First class' degree by the University of Edinburgh. We found that there is a greater chance of receiving a 'First class' degree from the University of Edinburgh in 2019/20 compare to 2018/2019. We found that there is a disparity between different demographic groups at attaining a 'First class' degree. We measured the odds ratios of being awarded a First class for each demographic group, and tested their statistical significance using 95% CIs. We acquired the data from HESA and performed preliminary analysis to find gradual increases in 'First class' attainment by plotting stacked barchart and boxplots. We later calculated the odds ratios and 95% CIs for mean proportion of 'First class' per year, and plotted these on graphs.

2 Introduction

Context and motivation   This study is an observational longitudinal retrospective cohort study. This is because we have used existing datasets collected from the universities over several years without any interventions. Those datasets are used to detect changes in the number of each degree classification over the sample years and examine whether the odds vary across those years for different demographic groups. Degree classifications in the UK are as follows: First class honours, Upper second class honours, Lower second class honours, Third class honours, and unclassified degrees. As university students, we are interested in understanding the issue of grade inflation as well as understanding our own odds in getting a first, given multiple demographic data points. Degree inflation, caused due to grade inflation, is defined as the decrease in value as a result of more number of students achieving a 'good' degree classification. This can be due to a wide range of factors, including, making exams easier, inflating grades to maintain university reputation or just improvements in learning outcomes [2]. The topic of degree classifications has also been frequently been reported in media and has been a cause of concern for numerous students as jobs get more competitive and the value of the degree diminishes. This study is aimed at finding out if degree inflation is a growing problem by analysing degree classification and achievement data from 2014/15 to 2019/20. [3][4].

Previous work   The 2020 Office for Students (OfS) report investigated and analysed the pattern in the change in the proportion of the First and Upper Second class degrees awarded between 2010/11 to 2018/19 [5]. There is also a related news article from the Guardian reporting on the OfS report [3].

Moreover, research paper from 2015 discussed that the modularisation of degree programmes and the change in assessment methods may have contributed to the increase in First and Upper second class honours degrees [2]. Bachan (2015) also further discussed other possible causes of such increase such as overall increase in student motivation and student ability, and the use of better teaching methods.

The Higher Education Statistics Agency (HESA) report in 2003 found that females gain more first degrees than males when combining first class and upper second class degrees together [6].

Objectives   In this report we would like to explore the following:

1. We are going to investigate if degree inflation in UK universities is still a recurring phenomenon. We would use the data provided by HESA from the years 2014/15 to 2019/20 and study the extent of degree inflation in these years.

2. Next, we study the degree classifications awarded by the University of Edinburgh from the academic years 2014/15 to 2019/20. We analyse if degree inflation is prevalent in this university as well as explore the odds of achieving a first class as compared to other universities.

3. The HESA report in 2003 says that certain demographic groups have greater odds of receiving a First class honour degree [6]. Therefore, we would like to investigate if this is continuing, i.e., the odds of being awarded First class honour is higher in certain groups than in others, in the years 2014/15 to 2019/20.

3   Data

Data provenance   To investigate the objectives, we have used 2 datasets in total. The data from both datasets were collected by the HESA from higher education (HE) providers in the UK. HESA is the official government agency for collecting data about UK higher education. They collect a range of data from HE providers annually which complies to GDPR, and subsequently enabling other government departments and HE funding bodies to base their work on the data provided by HESA [7]. Therefore, data from HESA is a reliable source to base this investigation on. The website cited below contained several csv files that were downloaded and cleaned for the purpose of this project. As this data had an open data licence (CC-BY-4.0 [8]), we were sure that the data can be used liberally as long as the providers are cited. [9]

Data description

1. The first dataset *(providers)* includes the level of qualification data, such as foundation degree, master taught etc. and first degree classification data of 297 HE providers from 2014/15 to 2019/20. This dataset is chosen as the first degree classification data includes data on the different type of degree classifications and degree types by higher education centres in the United Kingdom, which is relevant to the objectives we are exploring. [10]. (Number of records: 114408)

   - **UKPRN** - UK Provider Reference Number [11]
   - **HE provider** - Name of higher education provider
   - **Country of HE provider** - England/Wales/Scotland/NI
   - **Region of HE provider** - Regions of the UK - East Midlands, North East etc.
   - **Academic Year** - From 2014/15 to 2019/20
   - **Qualification/Classification marker**
     - **Level of qualification** - First degree, Postgraduate degree, Foundation, etc.
     - **First degree classification** - First class, second class, etc.
   - **Level of qualification/Degree classification**
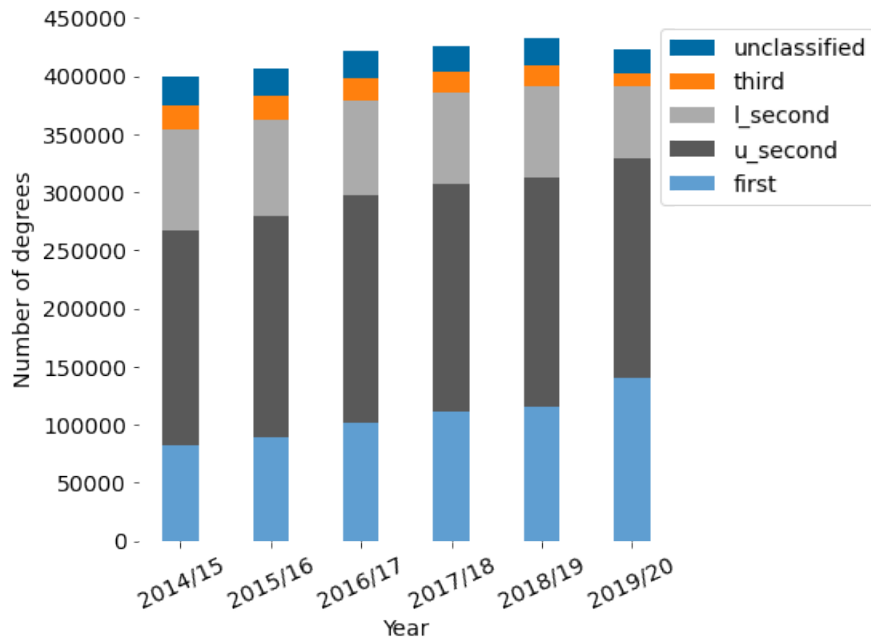   - **Number** - Number of degrees

Figure 1: The distribution of the degree classifications within the total given degrees per year

2. The second dataset *(factors)* contains the data for domiciled students from 2014/15 to 2019/20. The following are the list of data point columns it contains. (Number of records: 8725)

- **Category Marker** - sex, ethnicity, disability status etc.
- **Category** - Name of higher education provider
- **Mode of study** - Full-time or Part-time
- **Country of HE provider** - England/Wales/Scotland/Northern Ireland
- **Academic Year** - From 2014/15 to 2019/20
- **Classification of first degree** -
  - First class, upper second class, lower second class, third class and unclassified
- **Number** - Number of degrees

This data separates the degree class achieved by students by their demographic data and country of origin, as well as their mode of study. A row of the data would give us the number of degrees received by a demographic group, for a particular year, mode of study and country. This dataset is chosen as it provides the data to investigate the odds of First class with the different demographic data. Note that this degree dataset has fewer number of degrees than the first dataset.

Data processing
- Cleaning *providers* dataset - We split the csv file as it was very cluttered, for instance, the sum of data and the definition of the data were included in the same file. The different classifications dataset allowed us to see how the number of degrees achieved changed over the years. We then removed all the rows that contained fields such as, 'All' and 'Total' as they were redundant. We only included individual university data for different first degree classifications. We also removed all the NaN values as they couldn't be used. Next, we split our classification pandas' dataframe by each first degree classification, ie., First, Upper Second, Lower Second, Third and Unclassified.
- Cleaning *factors* dataset - This dataset contained all the demographic groups as well as the description in one csv file. First the description was moved to another file. Using pandas, we then split the data set into multiple csv files separated by the category field, i.e., datasets for sex, ethnicity, disability status, etc. were created. Next using pandas we acquired the data for each group
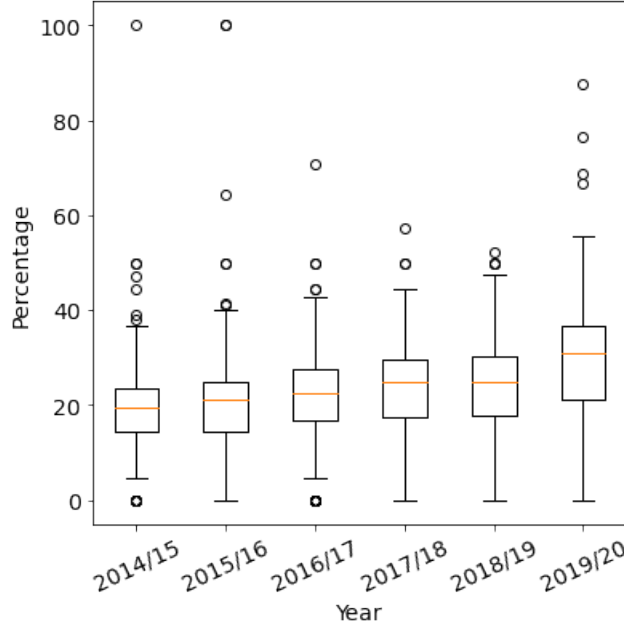
Figure 2: Visual summary of distribution of the proportion of First class degree classification by year

and created a dataframe, and from each dataframe we considered the data-points that achieved a first degree classification. Age Unknown field was removed from this data set, as it didn't have any number of degrees.

4 Exploration and analysis

Figure 1 shows that there is a general increase with the number of First class degrees as the years go on, whilst the number of Upper Second class degrees has remained fairly constant over the sample years. On the complementary, the number of Third class and Lower Second class degrees have been decreasing over the sample years. This increase in First class and decrease in lower honours degree might be contributed by the wider availability of information, allowing easier and quicker access to resources which enables students to gain better grades in their assessments. Moreover, notice the greater increase of First class degrees awarded in 2019/20 compare to rest of the years, and is statistically significant when comparing to 2018/19. This could be influenced by the sudden change of methods of assessment caused by the COVID-19 pandemic, forcing many honours examinations to be taken online instead. To address the sudden disruption of studies, many universities also had the non-detriment policy in-place which ensured students would be awarded a final grade that is no less than their previous in-year assessment [12].

We also looked at the proportion of First class degrees (Figure 2) from the higher education (HE) providers from 2014/15 to 2019/20. This is because by identifying any increase or decrease in the proportion of those degree classifications, it can help detect grade inflation in those years. We chose to investigate the proportion rather than the actual number of degree classification from the HE provider because there are varying number of providers each year. In Figure 2 shows that the distributions are near normal and symmetric distributions.

To measure the uncertainty in the calculated proportions of First class degrees across 2014/15 to 2019/2020, we calculated the 95% confidence intervals.

Formula for the 95% confidence interval for the mean proportion of First class degree classification:

$$CI = \exp\left[\bar{x} \pm \overbrace{1.96}^{\text{z-score}} \times \overbrace{\frac{\sigma}{\sqrt{n}}}^{\text{SE of mean}}\right] \qquad (1)$$
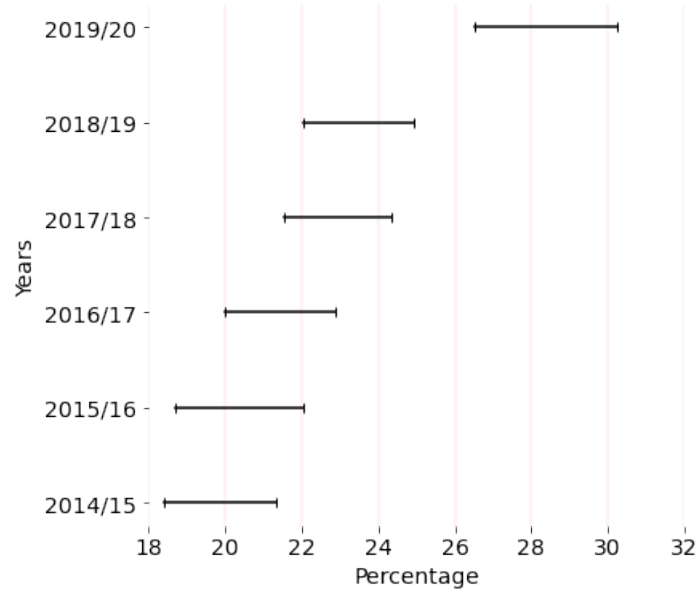
4

Figure 3: Confidence interval of the mean proportion of First class degrees per year

Where:
$\bar{x}$ = the mean proportion of First class degrees in that year
σ = the standard deviation
n = the number of universities concerned

As shown in Figure 3, we for example, can be 95% confident that the proportion of the first class degrees in 2014/15 is 18.41 and 21.32, and in 2019/20 is between 26.51 and 30.26 (all values quoted to 2 d.p.). Figure 3 also shows that it the proportion of degrees with First class honours in 2019/20 are statistically significant as there is no overlap when comparing 2019/20 to the rest of the years. The overlaps from 2014/15 to 2016/17 and 2017/18 to 2018/19 indicate that whilst the ranges of proportion of first class degrees have shifted to the right (increased) over those years, the true proportion of the population may not have increased, hence they are not statistically significant. Moreover, it should be noted that the increase in proportion of first class degrees in 2017/18 from 2014/15 is statistically significant.

Figure 4 shows that for years 2015/16, 2016/17 and 2019/20, there is a greater chance of being awarded a First class degree by the University of Edinburgh. The greater increase in the odds in year 2019/20 could be due to examinations with adapted format to allow remote assessment, and the University's participation in the non-detriment policy to support students [13]. Therefore the average grades of the final year students could only be equal to or higher than the grades from the assessments already taken that year. Moreover, this increase in 2019/20 matches the increase in total number of First class degrees in 2019/20 shown in Figure 1.

4.1 Odds ratios of receiving First class degrees for different demographic groups

Odds ratios (OR) are used to compare the odds of receiving First class degrees, given exposure to each demographic group in Table 1. This is because when the OR of a particular demographic data is greater than 1, it means that demographic group is associated with greater chance of First class degrees.

To find out whether the odds ratios of being awarded First class across 2014/15 to 2019/20 for each demographic group are statistically significant or rather due to chance, we calculated the 95% confidence intervals for those odds ratios as shown in Table 1. This is because if the 95% CI of an OR includes 1, then that OR is not statistically significant.
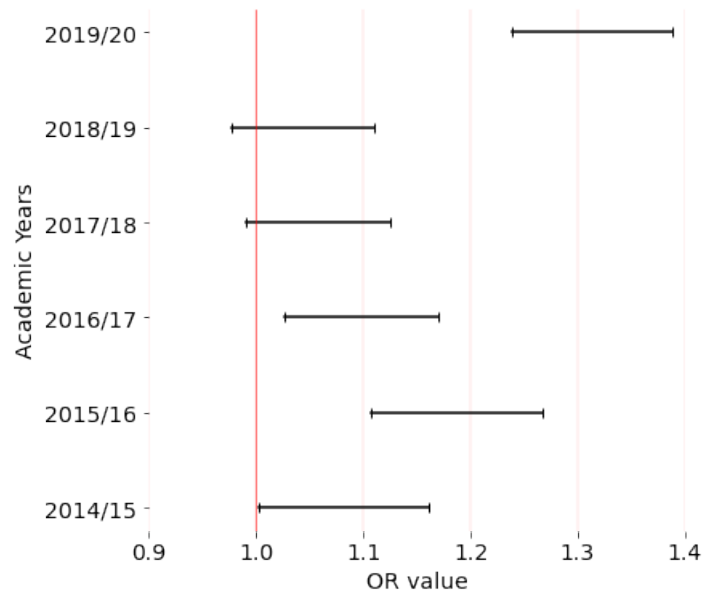
Figure 4: Confidence intervals of the odds ratio for First class degrees awarded in the University of Edinburgh by year.

Formula for calculating the odds ratio for each demographic category:

$$OR = \frac{a/c}{b/d} = \frac{ad}{bc} \tag{2}$$

Formula for 95% Confidence Interval for the odds ratio [14]:

$$CI = \exp\left[\ln\left(OR\right) \pm \overbrace{1.96}^{\text{z-score}} \times \underbrace{\sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}}_{\text{SE of ln(OR)}}\right] \tag{3}$$

Where:
a = Number for firsts for that group b = Number of all other degrees for that group
c = Number of firsts for other groups d = Number of other degrees for other groups

Therefore, the following is the breakdown of the odds ratio for each demographic groups (Table 1).

- For the category of sex:
  - $OR > 1$: Female, Other
  - $OR < 1$: Male
  - Statistically Insignificant: Not known

- For the category of ethnicity:
  - $OR > 1$: White
  - $OR < 1$: Black, Asian, Mixed, Other, Not known

- For the category of religious belief:
  - $OR > 1$: No religion, Jewish
  - $OR < 1$: Buddhist, Christian, Hindu, Muslim, Sikh
  - Statistically Insignificant: Spiritual, Any other religion or belief, Not known

- For the category of age group:
  - $OR > 1$: 21-24 years
  - $OR < 1$: 20 and under, 25-29 years
  - Statistically Insignificant: 30 years and over

- For the category of disability status
  - $OR > 1$: No known disability
  - $OR < 1$: Known disability

Table 1: Confidence Interval Table for First class degree odd ratios for each demographic

Demographic group	Odds ratio	Lower 95% CI	Upper 95% CI	Standard error of ln(OR)
Female	1.0425	1.0357	1.0493	0.0033
Male	0.9586	0.9524	0.9649	0.0033
Other	1.4413	1.2472	1.6655	0.0738
Not known	0.9178	0.5533	1.5224	0.2582
White	1.6630	1.6496	1.6765	0.0041
Black	0.4202	0.4133	0.4272	0.0085
Asian	0.7293	0.7213	0.7374	0.0056
Mixed	0.9100	0.8946	0.9257	0.0087
Other	0.7201	0.6986	0.7423	0.0155
Not known	0.5569	0.5407	0.5736	0.0150
No religion	1.2595	1.2463	1.2728	0.0054
Buddhist	0.8287	0.7616	0.9018	0.0431
Christian	0.8869	0.8767	0.8972	0.0059
Hindu	0.8410	0.8065	0.8770	0.0214
Jewish	1.2216	1.1325	1.3177	0.0386
Muslim	0.6440	0.6305	0.6578	0.0108
Sikh	0.8272	0.7836	0.8733	0.0276
Spiritual	0.9988	0.9506	1.0495	0.0252
Any other religion or belief	1.0272	0.9814	1.0751	0.0232
Not known	1.0161	0.9966	1.0359	0.0099
20 and under	0.8553	0.8385	0.8725	0.0101
21-24 years	1.0916	1.0826	1.1007	0.0042
25-29 years	0.8904	0.8796	0.9013	0.0062
30 years and over	0.9885	0.9770	1.0001	0.0059
Known disability	0.8993	0.8911	0.9075	0.0046
No known disability	1.1119	1.1018	1.1221	0.0046

The odds are 1.6630 (1.6496-1.6765) times higher that a white student would receive a First class degree compared to a non-white student. White is the only demographic group on the category of ethnicity that has a increased odds of receiving a First class degree. Moreover, the odds ratios for Black is 0.4202 (0.4133-0.4272) and Asian is 0.7293 (0.7213-0.7374) with Black and White having the largest difference in odds ratio. This reflects the degree attainment gap for the minority ethnic groups in the UK HE student population which universities are trying to tackle [15]. As the standard error (SE) of the log OR increases, the greater the range shown in the CI. Notice for the category of sex, the SE of the log OR for 'Not known' (OR=0.9178, 0.5533-1.5224) is 0.2582 - greater than the SEs of ln(OR) for the other 3 groups. Therefore 'Not known' has the largest range for its 95% CI of OR. This is because the smaller the number of data (i.e. number of Sikhs receiving a First class) we have on a particular group, the larger their SE of ln(OR), resulting in a increased CI range. This is also true for '20 and under' as there are few UK-domiciled students who graduate age 20 or under.

5   Discussion and conclusions

Summary of findings

1. We found that from the sample years of 2014/15 to 2019/20, the sample proportion of First class degrees have a **limited increase** between 2014/15 and 2018/19. The only statistically significant increase was comparing between 2014/15 and 2017/18. However, in 2019/20 we find that there is statistically significant increase of proportion of First class degrees awarded, due to the COVID-19 pandemic. The universities modified their assessment policies to accommodate the exceptional circumstance, hence the hike in 2019/20 was expected.

2. In the University of Edinburgh, we found that degree inflation was similar to most other universities. In the academic year 2019/20, due to the ongoing pandemic, the university took a non-detriment approach, because of which it expected students to get better or similar grades then they used to. We also found the odds of getting a first at the university were statistically higher than other universities in the years: 2014/15, 2015/16, 2016/17, 2019/20 .

3. This report clearly shows the **statistically significant difference in odds between the different demographic groups** as it is a topic that is closely watched by the education sector. We suggest that there is still room for improvement for the HE providers in terms of reducing the degree attainment gap between the demographic groups.

Evaluation of own work: strengths and limitations   This report has used **reliable and legitimate sources** for datasets and references, citing from publicly available sources such as governmental departments, and broadsheet media. We used statistical methods to support its findings to the stated objectives. The visualisations take into account colourblindness. To acknowledge the limitations of the report, the datasets used only covered six recent years. Moreover, when investigating the impact of each demographic group on the odds of first class degrees, we calculated the odds ratios based on the total number of First and non-First class degrees across the sample years. However, there may be some hidden patterns in the data, for example any improvements in terms of odds ratio difference between white and other minority ethnic groups. Therefore to improve, we could calculate the odds ratios for each demographic group per year. Currently the demographic dataset *factors* does not include individual data from the universities, this would be useful to find out which universities have the highest disparities, as well as how diverse different universities are in general. Further limited by the small sample years of the datasets, the odds ratios in Table 1 may not be fully reflecting any solutions or strategies implemented to reduce the degree attainment gap between the different demographic groups since the significance of those solutions may only be reflected in the long-term.

Comparison with any other related work   HESA report on the proportion of First class degree within each ethnicity group from 2018/19 demonstrates a similar pattern to the odds ratios of the ethnicity groups in Table 1 [16]. Notice that the HESA report evaluates the proportions yearly whereas we investigated it in this report with all the sample years combined.

Moreover, in terms of the pattern of the proportion of First class degrees, BBC News report that says the **increase in proportion of First class degrees seemed to have ceased in 2018/19** [17]. Our report matches that finding in Figure 3 where it shows that the increase in proportion of First class, except from 2014/15, are statistically insignificant. The statistically significant increase in 2019/20 found in Figure 3 also fits the report from the Guardian which states a record number of First class degrees in 2019/20.

Improvements and extensions   To improve the findings in this report, we could investigate the objectives with additional datasets that includes more sample years, i.e. before 2014/15, to check for any significant patterns. We could use one-hot encoding to convert our categorical dataset perform logistical regression to find a potential regressor. A third dataset on degree classification achievement separated by subjects was also available. A similar observational study on that dataset could also be conducted. Moreover, as an extension, we could investigate the proportion of degree classifications of other countries i.e. the US, and compare the findings in this report to those of the US for any similarities or differences.

References

[1] D. Hinds. Universities told to end grade inflation. *GOV.UK*, 2019. `https://www.gov.uk/government/news/universities-told-to-end-grade-inflation`.

[2] R. Bachan. Grade inflation in UK higher education. *Studies in Higher Education*, 2015. `http://dx.doi.org/10.1080/03075079.2015.1019450`.

[3] S. Weale. Proportion of students in England awarded first-class degrees soars. *The Guardian*, 2020. `https://www.theguardian.com/education/2020/nov/19/students-england-awarded-first-class-degrees-grade-inflation`.

[4] S. Coughlan. https://www.bbc.co.uk/news/education-48951653. *BBC News*, 2019. `https://www.bbc.co.uk/news/education-48951653`.

[5] Office for Students. Analysis of degree classifications over time - changes in graduate attainment from 2010-11 to 2018-19. `https://www.officeforstudents.org.uk/media/55b365fd-2d77-46c0-a8ad-45304047a0be/analysis-of-degree-outcomes-over-time-2020.pdf`.

[6] Higher Education Statistics Agency. Women gain more first and upper second class first degrees than men. *Higher Education Statistics Agency*, 2003. `https://www.hesa.ac.uk/news/08-04-2003/qualifications-achieved`.

[7] Imperial College London. About HESA and the use of your data. *Imperial College London*, no date. `https://www.imperial.ac.uk/admin-services/secretariat/information-governance/data-protection/hesa/`.

[8] Attribution 4.0 international (cc by 4.0). "You are free to share and adapt".

[9] Introduction to data on HE students. `https://www.hesa.ac.uk/data-and-analysis/students/introduction`.

[10] UKPRN data report on HE qualifies by providers. `https://www.hesa.ac.uk/data-and-analysis/students/table-16`.

[11] UKPRN. `https://www.hesa.ac.uk/collection/c19041/a/ukprn`.

[12] S. Weale. Record one in three uk students gained first-class degree in 2020. *The Guardian*, 2021. `https://www.theguardian.com/education/2021/jan/27/record-one-in-three-uk-students-first-class-degree-2020`.

[13] C. Harmon. Coronavirus (Covid-19): exams and assessments update. *The University of Edinburgh*, 2020. `https://www.ed.ac.uk/files/atoms/files/coronavirus_update_-_exams_and_assessments_-_26_march_20.pdf`.

[14] Magdalena Szumilas. Explaining Odds Ratios. *PubMed Central*, 2010. `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2938757/`.

[15] Universities UK. BAME attainment gap. *Universities UK*, 2019. `https://universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2019/bame-student-attainment-uk-universities-closing-the-gap.pdf`.

[16] Higher Education Statistics Agency. Undergraduate degree results. *Ethnicity facts and figures GOV.UK*, 2020. `https://www.ethnicity-facts-figures.service.gov.uk/education-skills-and-training/higher-education/undergraduate-degree-results/latest`.

[17] S. Coughlan. 'Grade inflation' in top degree grades stopping. *BBC News*, 2020. `https://www.bbc.co.uk/news/education-51136349`.