**UCD Professional Academy**

**Certificate in Data Analytics (August 2021 intake)**

**Project Report**

**THE SOUTH EAST “BRAIN DRAIN”:**

**HIGHER EDUCATION ENROLMENTS AND BALANCED REGIONAL DEVELOPMENT IN IRELAND**

**Richard Hayes**

**GitHub URL**

<https://github.com/rhayeswaterford/UCDProject>

**Abstract**

Ireland’s national development plan to 2040 places great emphasis on balanced regional growth, that is, on ensuring that population growth, capital investment, and economic development take place in a way that ensures sustainable growth across the entire country not only in Dublin and the East. Higher education is key to this as higher education institutions are critical to the so-called “knowledge economy”: they produce the graduates and intellectual property to drive future prosperity, especially in high tech industries. This project examines higher education student enrolment data to gain insight into the regional profile of higher education in Ireland, with a particular emphasis on the so-called “brain drain” from the counties of the South East of Ireland. It is alleged that these counties (Waterford, Wexford, Kilkenny and Carlow) lose large number of higher education-going students to colleges—especially universities—in other regions and that the migration from the region of students constitutes a loss of intellectual and human capital to the region that does not return. The project in this way constitutes a gap analysis of regional higher education provision and therefore of regional development that will be of use in ensuring the goals of *Ireland 2040* are achieved.

**Introduction**

A keynote in *Ireland 2040*, the Irish National Planning Framework, relates to what is termed “balanced growth” across Ireland’s regions.[[1]](#endnote-1) *Ireland 2040* reads:

From an administrative and planning point of view, Ireland is divided in to three regions: the Northern and Western, Southern, and Eastern and Midland Regional Assembly areas. We need to manage more balanced growth between these three regions because at the moment Dublin, and to a lesser extent the wider Eastern and Midland area, has witnessed an overconcentration of population, homes and jobs. We cannot let this continue unchecked and so our aim is to see a roughly 50:50 distribution of growth between the Eastern and Midland region, and the Southern and Northern and Western regions, with 75% of the growth to be outside of Dublin and its suburbs. (p.11)

This involves a concentration of growth in the cities, according to the plan.

We have five cities in Ireland today in terms of population size (>50,000 people): Dublin, Cork, Limerick, Galway and Waterford. In our plan we are targeting these five cities for 50% of overall national growth between them, with Ireland’s large and smaller towns, villages and rural areas accommodating the other 50% of growth. (p.11)

The cities, and especially the regional cities, will be “places that can foster enterprise and innovation and attract investment and talent.” This can be achieved, the plan goes on, by “the coordination of growth and place making with investment in world class infrastructure, including digital connectivity, and in skills and talent to support economic competiveness and enterprise growth” (p.14).

The OECD, amongst other organisations, recognises that “tertiary education is a major driver of economic competitiveness in an increasingly knowledge-driven global economy”—that is, when it comes to innovation capacity, and economic development, high quality higher education is key, and the higher education institutions are the kind of infrastructure needed to build and sustain economic growth. The OECD declares: “The imperative for countries is to raise higher-level employment skills, to sustain a globally competitive research base and to improve knowledge dissemination to the benefit of society”, and it does this through increased focus and investment in third-level education.[[2]](#endnote-2) If national economic development goals are to be achieved, then, third level institutions will be key to is and, furthermore, it is certain that if national development goals with regard to balanced growth are to be achieved, critical will be the development of strong region-based higher education institutions with the capacity and scale to drive region-based innovation and economic development.

This is recognised in the various Regional Spatial and Economic Strategies (RSESs) published by each region as the framework within which that particular region will develop over the coming decades. The RSESs acknowledge that, in the context of a “knowledge economy”, universities function as critical anchor institutions within regions, not only by attracting talent and investment into the region but also, crucially, by generating the knowledge that drives innovation in the first place, as well as by contributing to the larger economy of the region by their sheer scale. (A study by the Irish Universities Association estimated that each university in Ireland contributes about €1 billion per annum to the economy.[[3]](#endnote-3)) Within regions, third level institutions supply graduates to industry and society, and in this way support regional human capital development; they build the knowledge base; they disseminate that knowledge; and they act as a knowledge custodian and repository.[[4]](#endnote-4) In these ways they act as innovation leaders.

The RSES published by the Southern Regional Assembly highlights the critical importance of higher education to the future Southern region when it comes to the skills:

Access to high quality higher education locally and regionally enhances the quality of life of all the Region’s citizens and facilitates greater levels of third level enrolment and attainment. By improving access to higher education, this can arrest outward migration and encourage significantly greater numbers of people to study and subsequently work and live in the Region.[[5]](#endnote-5)

The RSES regional proposition also depends heavily on higher education institutions to produce not only the skilled graduates needed for regional industry but also to sustain a level of innovation activity through research and development.

But what is the current profile of regional higher education, and how equipped would the Irish higher education systems seem to be—in terms of its spatial distribution—to contribute to balanced growth across Ireland’s regions? This is the focus of this report. The occasion of the report is the impending merger of Waterford Institute of Technology and Institute of Technology, Carlow and their re-designation as South East Technological University. The report looks at the existing profile of student enrolment in the South East region (a sub-region of the Southern Region described in national policy and comprising the counties of Waterford, Wexford, Kilkenny and Carlow). The analysis of that profile raises interesting questions in relation to the current overall balance of provision and points towards what is needed better to achieve national goals.

**Dataset**

The primary dataset used in this project was student enrolment data sourced from the Higher Education Authority (HEA). While the HEA makes myriad datasets available through its website (see https://hea.ie/statistics/data-for-download-and-visualisations/), the dataset with the variables needed for this project was not publically available and was sourced directly from the HEA.

To analyse the SE “brain drain”, it was necessary to have data on (a) the domiciliary origin of students enrolled in higher education in Ireland (by county), (b) the college attended by that student, and (c) the discipline undertaken by that student.

The project also makes use of another dataset, a report on the population of individual counties according to 2016 census data. This data was sourced from a website, http://citypopulation.de/en/ireland/towns. The data used on the website derives from the Central Statistics Office (CSO).

**Implementation Process**

|  |  |
| --- | --- |
| [1]-[3] | *Setup*  Before any analysis can take place, the project is set up. This involves importing the relevant packages (pandas, matplotlib.pyplot, seaborn and numpy). This also involves importing the main dataset for analysis. This dataset is called simply **raw\_material**. Details include the college the student was registered in and the course the student was on. Duplicates were removed. |
| *Section 1: the Profile of Higher Education*  The first section of the project describes the overall profile of higher education in Ireland and describes pattern of enrolment in higher education in terms of students’ county and regional origin. This section in other words examines regional “balance” in terms of the regions *sending* students to higher education. | |
| [4]-[10] | Section 1.1 asks what Higher Education Institutions in Ireland have the most students.  An extract from **raw\_material**, **student\_population\_colleges**, is taken using *groupby*. This list is sorted and a new column introduced, “Rank”, to generate a list of colleges by student population. This is represented graphically for clarity in the form of a bar chart which allows for simple comparison. The percentage of the total enrolled by each college is also generated and represented in a scatter plot that is generated using Seaborn with three variables (the college, the number of students, and the percentage total) represented. |
| [11]-[16] | A sub-question, 1.1.1—what kinds of institutions do students typically attend, is also examined. Two lists of college types are generated, a list of universities and a list of colleges. These are then concatenated into one list as, generally speaking, colleges and universities are understood to be very similar in their operations. The remaining institutions not on the concatenated list are necessarily the remaining category of institution, namely, Institutes of Technology. Using *isin*, the **student\_population\_colleges** data is sliced into data for each category of institution. The numbers in each category of institution are calculated and this data is used to populate a new DataFrame that summarises the distribution of the student population between Institutes and Universities/Colleges. A simple barchart illustrates the distribution. |
| [17]-[22] | Section 1.2 examines the most popular disciplines for students. Disciplines in **raw\_material** are sorted by International Standard Classification of Education or ISCED codes, a method for determining the broad subject to which the student has been enrolled. The initial dataset, **raw\_material**, is sorted according to the discipline using *groupby* and a set of data, **student\_population\_courses**, extracted. This is then sorted and ranked to determine the most popular courses. The top 20 course areas only are examined. A bar chart shows this in terms of student enrolment and then a second scatterplot that incorporates data on the percentage of the overall enrolment represented by the discipline. The latter scatterplot is shuffled for variety of presentation. |
| [23]-[31] | Section 1.3 investigates the distribution of students by domiciliary origin and looks in the first instance at the county of origin.  Again, a new data set is generated, extracted from **raw\_material**, called **student\_population\_county**. (Counties with null values are removed.) Work is carried out to clean this data as the names of counties Dublin and Antrim are presented with parenthetical comment that is not helpful. This data is sorted and the counties ranked according to which supplies the most students. |
| [32]-[44] | This rather crude method supplies predicable results so in subsection 1.3.1 a participation rate is generated. This involves importing an additional dataset scraped from a website that lists the population of Ireland by county based on the most recent censuses. This is then tidied to remove superfluous data and to bring uniformity to how counties are named: a function is created that passes over the data and identifies the eccentricities in naming convention and returns the standard name into a new column, and then the original name column is removed. The multiple counties listed for Dublin (Fingal, Dun Laoighaire-Rathdown, South Dublin, Dublin City) are thus effectively combined into one entry, Dublin. The Irish name for Laois is also removed. A list of counties is generated then in order of population size. |
| [45]-[52] | This data is then merged with **student\_population\_county** and this allows the generation of a calculation for each county of Students per 1000 population. The output from this is represented graphically. A mean is calculated and included in the graphic. |
| [53] | Section 1.4 then considers regional data and asks what region sends the most students to higher education.  Lists are prepared of the counties in each of the NUTS3 regions of Ireland and, separately, of each National Development Plan region (NDP). |
| [54]-[55] | **Student\_population\_county** data is then interrogated using these lists to generate student population data on each region and both NUTS3 and NDP levels. |
| [56]-[57] | This student population data is then combined into a new dataset for NUTS3 region and a percentage of the total for each region is calculated. |
| [58]-[64] | “Students per 1000” rates are calculated for each region, a mean is determined, and a plot is generated allowing for a comparison of all regions against one another and against the mean. |
| [65]-[71] | This same process is run for NDP regions without the plot. |
| [72] | Additional information is then added to a master table on NRP regions: information on the total regional population and the percentage of the national population in the region, and the number of students per 1000 citizens in each region. This table is re-indexed to tidy up the column order. |
| Section 2 then turns to examining the profile of the South East region when it comes to higher education. | |
| [78]-[81] | An extract is again taken from **raw\_material** to generate a dataset, **SE\_counties\_enrolment**, which is data on students from the South East NUTS3 region enrolled in higher education. A table of the overall number of enrolments by county is generated, along with the total. The SE total and the percentage that total represents of the overall total in higher education is calculated. |
| [82] | In section 2.1, a table of the most popular institutions for students from the SE is generated, **SE\_popular\_colleges**. This is ranked according to the enrolment from the SE of each college. |
| [83] | In section 2.2, a table of the most popular courses for students in the SE is also generated, **SE\_popular\_courses**. This too is ranked according to popularity. Only the top 20 courses are shown. |
| [84]-[89] | 2.2.1 takes that list of popular and examines it against the list of courses that are popular nationally—the attempt here is to identify any anomalies in the SE profile, instances where a course is more or less popular in the region than it is nationally. |
|  | * This is approached first by merging the popular SE list (**SE\_popular\_courses**) with the national list (**student\_population\_course**). The merged list is analysed and the ranks of each of the popular courses examined and classified as either a higher rank in the region in terms of popularity, a lower rank in the region, or the same level of popularity. * This allows for the generation of a table of the top 20 courses that have a higher popularity outside the region than inside the region—meaning courses where students can take the course both in or out of the region but favour studying outside the region. |
|  | Section 2.3 then examines a sub-set of students who are from the SE and study in the SE. |
| [90]-[92] | A list of SE colleges is created. This is used to create two tables, students from the SE who are studying in the SE and students from the SE who are studying outside the SE. The number of students in the SE studying in the SE is tabulated. |
| [93]-[94] | The courses that students from the SE take in the SE are examined and ranked by popularity. A plot is generated to illustrate the choices graphically. |
| [95] | The number of students studying outside the SE is then enumerated and a ranked list of institutions generated. |
| [97] | The table of courses that students from the SE take outside the SE is then generated. |
| [98]-[99] | The tables of courses within and outside the SE are merged in order to be able to see what courses are taken outside the SE that are not taken at all, or are less popular, in the SE. |
| [100]-[102] | In the first instance, the merged table is used to identify courses where there is no SE enrolment. Courses with a NaN value against regional colleges are extracted and tabulated in order of popularity. |
| [103]-[108] | The merged table of courses is also used to compare courses that are more or less popular in the SE and out of it – this in order to be able to generate some information on why students might leave in the first place. The popularity ranking of the courses nationally and regionally are compared and where the score is higher or lower a “popularity match” return is made. This allows for a filtering of courses according to popularity match. |
| [109] | The type of HEIs that students stay or leave the region for are then examined. A table is generated that presents for categories of student: those leaving the SE for university, those staying for university; those leaving for IOTs, those staying for IOTs. |
|  | In section 2.4, comparators are introduced in order to determine whether or not the pattern of student enrolment in the SE is typical. These comparators are drawn from the South West, the West and the Mid West—the other Irish city regions. |
| [110]-[113] | Datasets for the overall enrolment of students from the SW, and the distribution of those students in and out of the region, are generated. |
| [114]-[121] | Similar datasets are generated for the MW and the W. |
| [122]-[127] | This allows for the generation of a table that presents each region and the percentage of students from the region who study in the region vs those who do not. This is represented in a simple bar chart. |
| [128]-[130] | The kind of HEI that recruits students within and outside the region is then examined and tables generated for the SW, W and MW describing the student population. |
| [131]-[132] | Three subplots finally are created showing the pattern of enrolment of the SW, MW and SE side-by-side. |
| [134] | The Technological University case is tested by running the list of courses for which students leave the SE against the courses available in TU Dublin. A function is created to do this by checking one list against the other and returning “HIT” where there is a match. The number of hits is then counted. The return is 0. |

**Results**

***Section 1: the Profile of Irish Higher Education***

1. 160,611 full-time students were enrolled in higher education in Ireland in 2018-19.

*Largest Institutions*

1. The largest Higher Education Institution (HEI) in terms of student enrolment is UCD. The least populace was St Angela’s College, Sligo (figure 1). The relative size of colleges according to the percentage of the overall HE enrolment is represented in Figure 2. More students attend universities or colleges than attend Institutes of Technology (Figure 3).

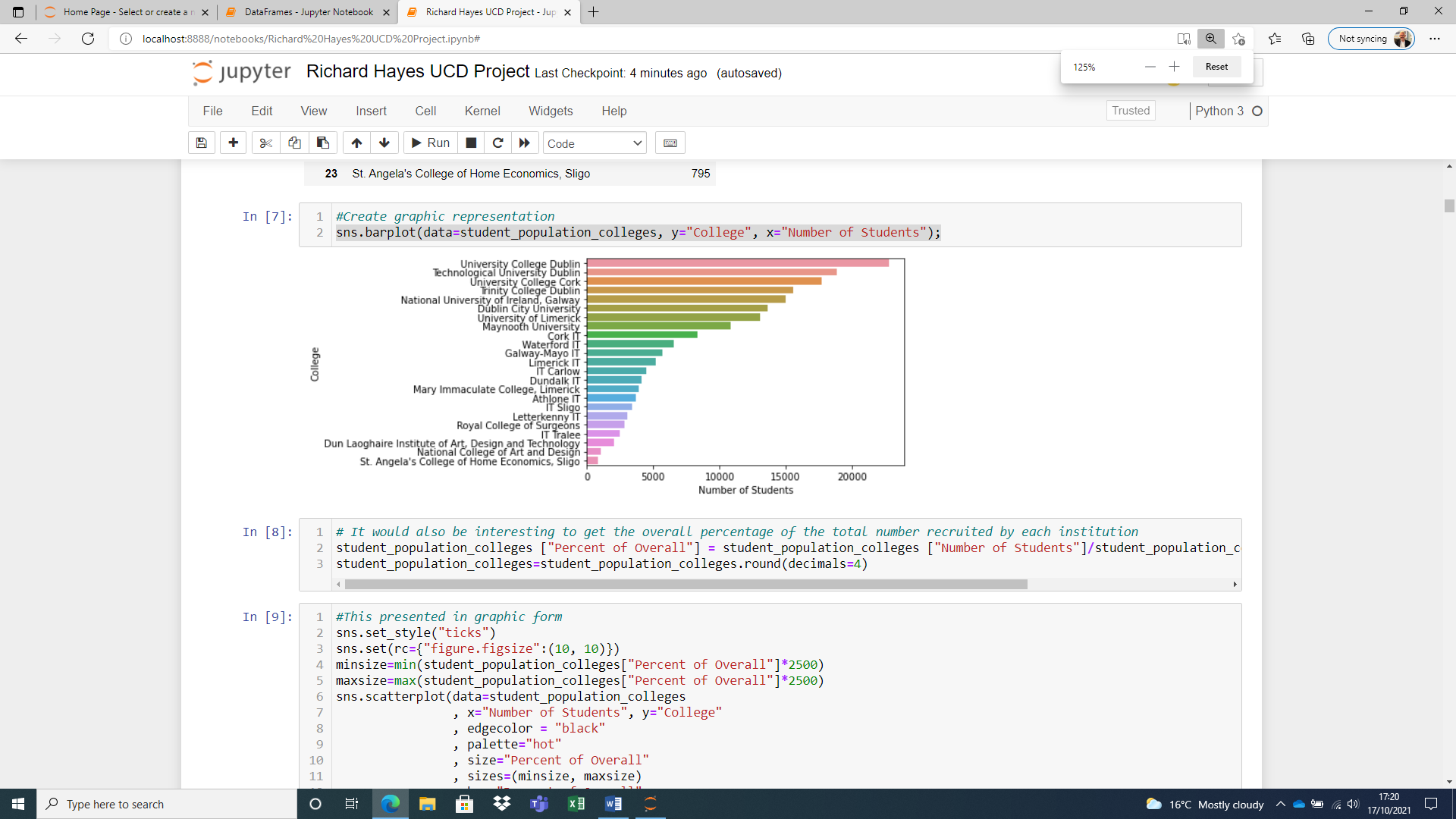


Figure Irish HEIs by Enrolment Count

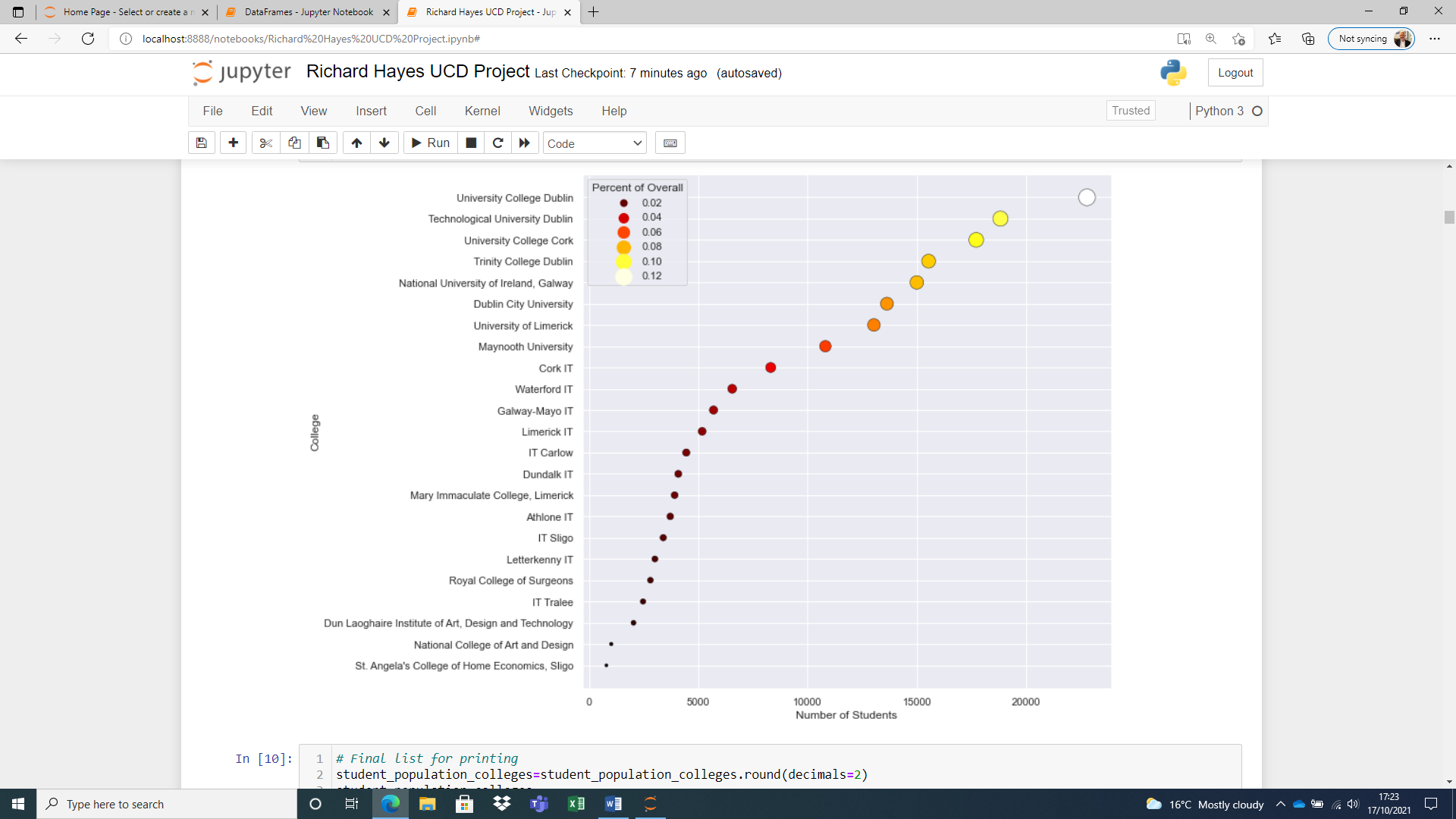


Figure HEI Enrolment Percentages

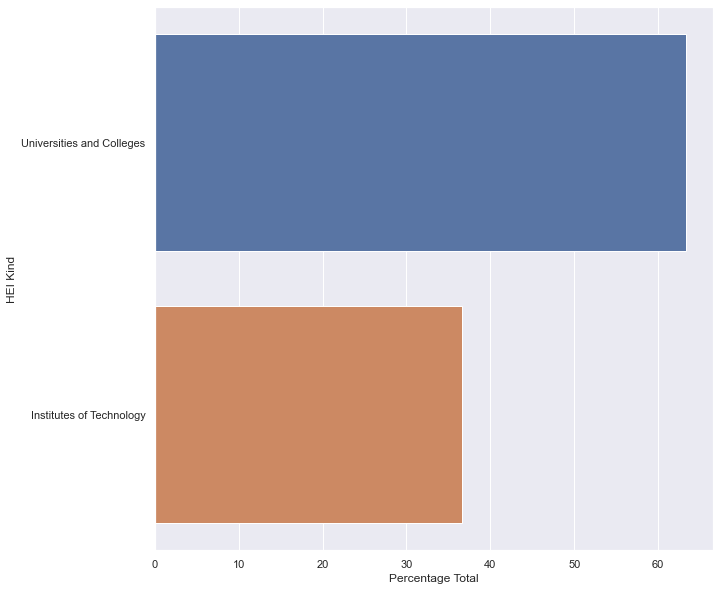
****

Figure Distribution of Enrolment by Institution Type

With the exception of Technological University Dublin (classified in this report with the Institutes of Technology), the largest recruiters of students are the traditional universities.

*Most Popular Courses*

1. The most popular courses in the country are in the areas of Business and Arts, with Nursing, Medicine and Law also enrolling strongly (see Figure 4). Figure 5 shows the relative popularity of disciplines, with the brighter bubbles representing the larger percentage share of the student enrolment.

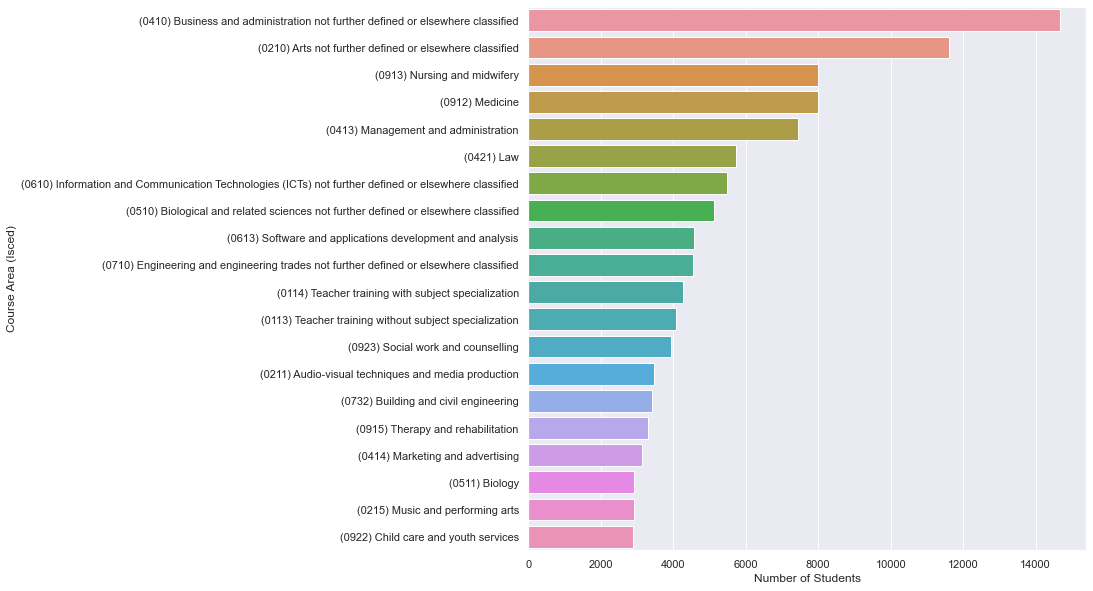
****

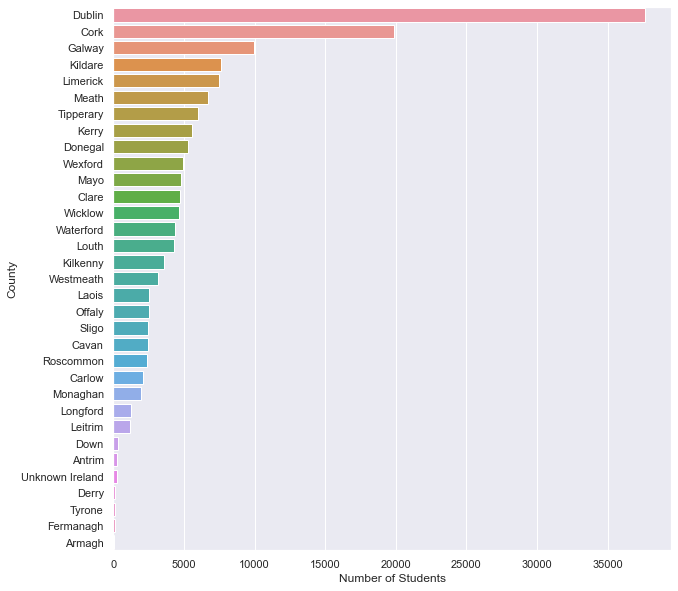
Figure Most Popular Disciplines Nationally

****

Figure Relative popularity of disciplines

*County and Regional Origin of Students*

1. If we consider the breakdown of student enrolments by County, we see, unsurprisingly, that the county that supplies the most students to higher education institutions is Dublin. Counties with cities (Dublin, Cork, Galway, Limerick) fill ranks 1, 2, 3, and 5 in the county list, the exception (in the Republic) being Waterford.



It is not surprising that the largest population centres supply the most students. A less crude measure might be the number of students per 1000 population, which gives quite a different picture.

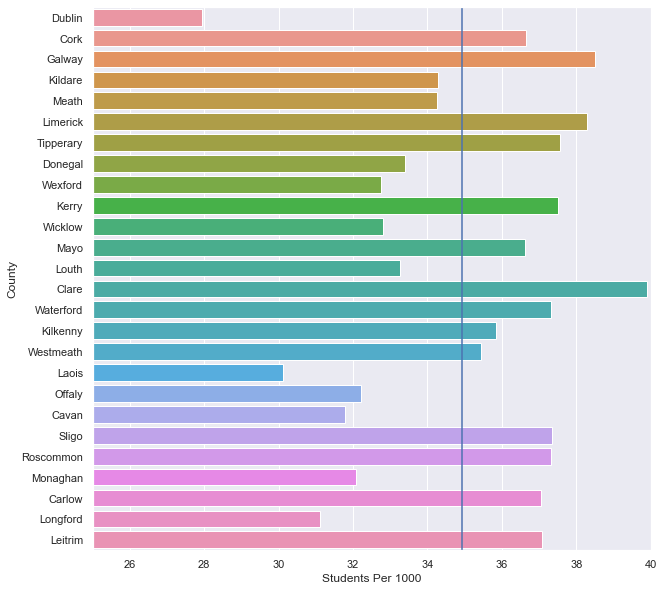


Figure Students Per 1000 by County

We see a very different picture in this graphic, with counties like Clare, Kerry, Sligo, Mayo showing higher levels of students per 1000 than Dublin. There are limitations to this methodology that must be mentioned: the population profiles of each county may well be very different, and this may skew the results.

1. As the National Development Plan operates at regional level, however, it is important to examine *regional* participation insofar as that is possible within the limitations of the data. In terms of the National Development Plan regions, the following table shows the distribution of students and population and offers some insight into the “balance” or otherwise of the regions currently.

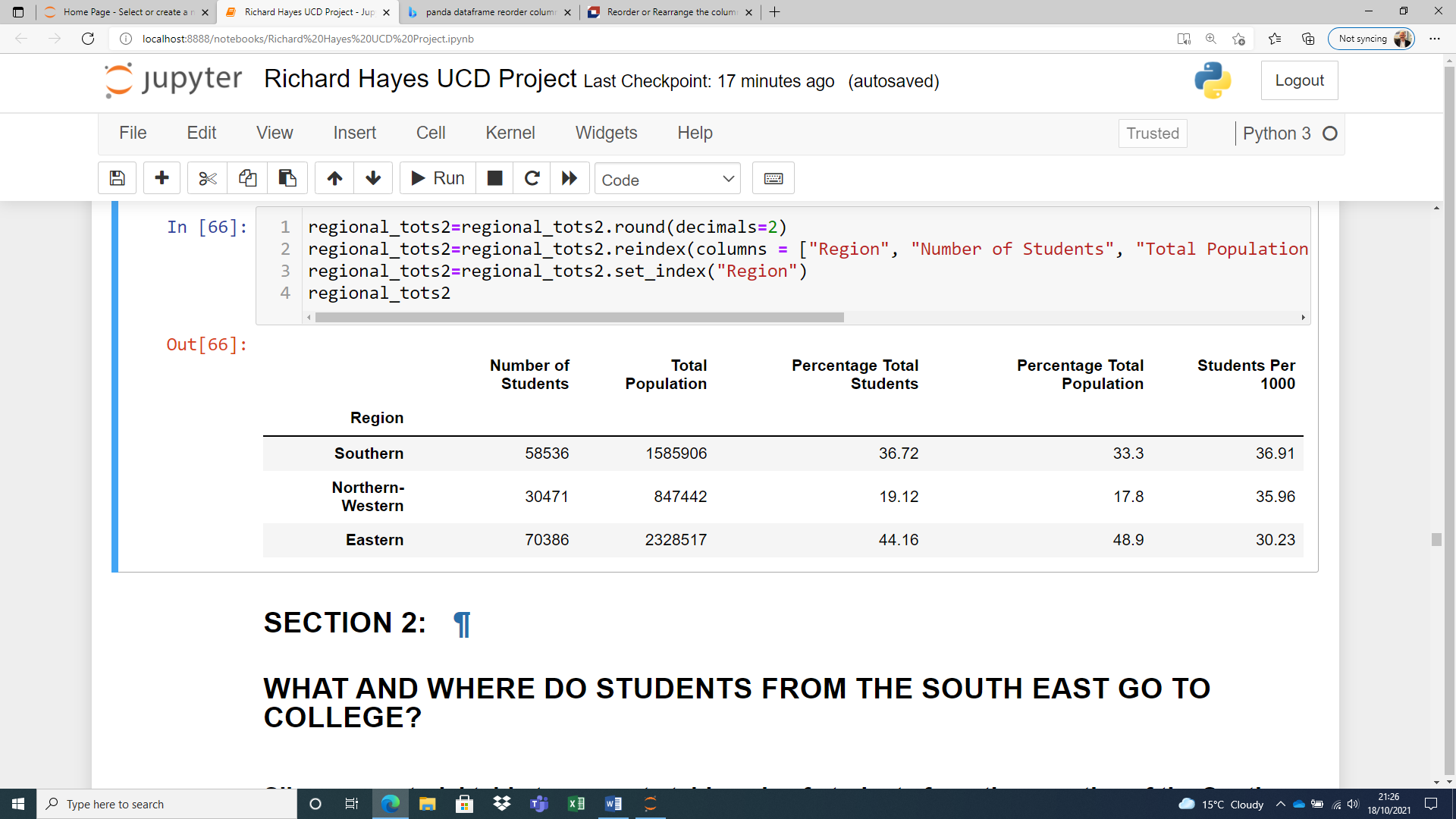


Table Regional Distribution of Students and Population (NDP Regions)

1. At the NUTS3 regional level, that is, at the sub-regional level in the NDP, we see interesting variation if we consider the Students Per 1000 method as an indication of overall participation rates:

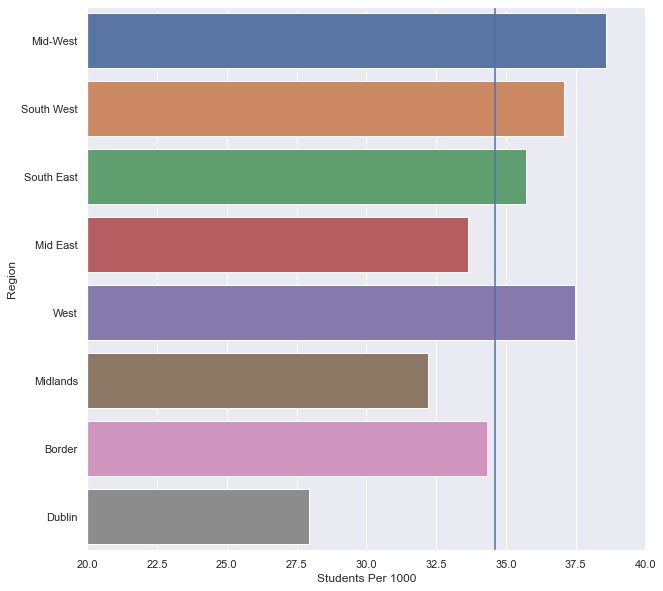


Figure Students per 1000 NUTS3 Region

Again the limitations of this methodology are pointed out.

**Section 2: Higher Education in the South East**

1. The second part of this report examines in detail the case of the South East of Ireland. 14,907 full-time students attend third level institutions with a domiciliary origin of one of the four South East counties (Carlow, Kilkenny, Waterford, Wexford). This represents just over 8% of the overall student population of the country.

*Most Popular Colleges*

1. The most popular colleges for students from the South East are in the South East—Waterford IT and IT Carlow. These colleges enrol, respectively, 25.7% and 11.3% of the total population of students from the region. The least popular is, unsurprisingly, the most distant, Letterkenny IT. The following table lists the colleges in order of popularity:

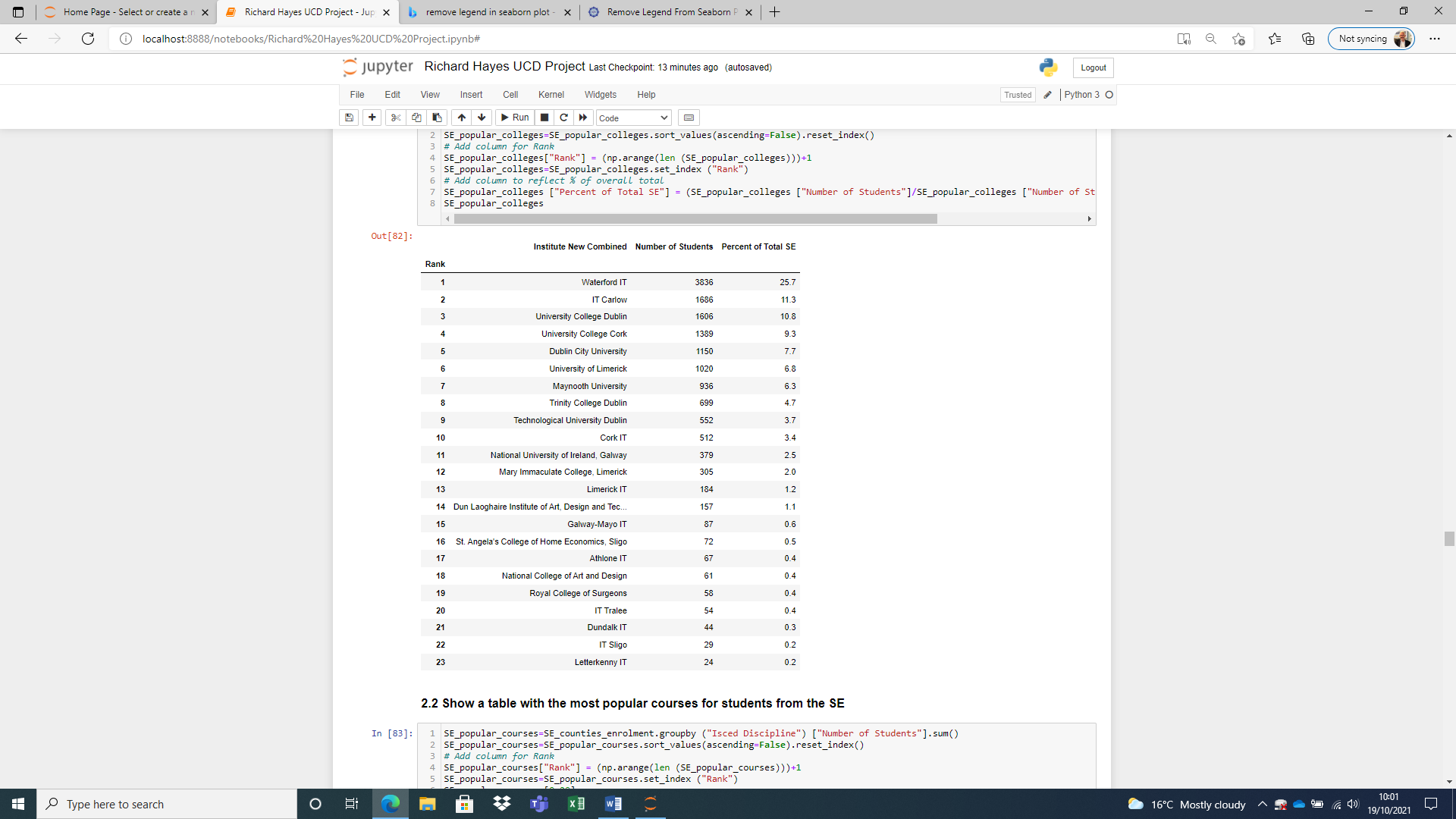


Table Institutions enrolling SE students

*Most Popular Courses*

1. The most popular courses for students from the South East are given in this table:

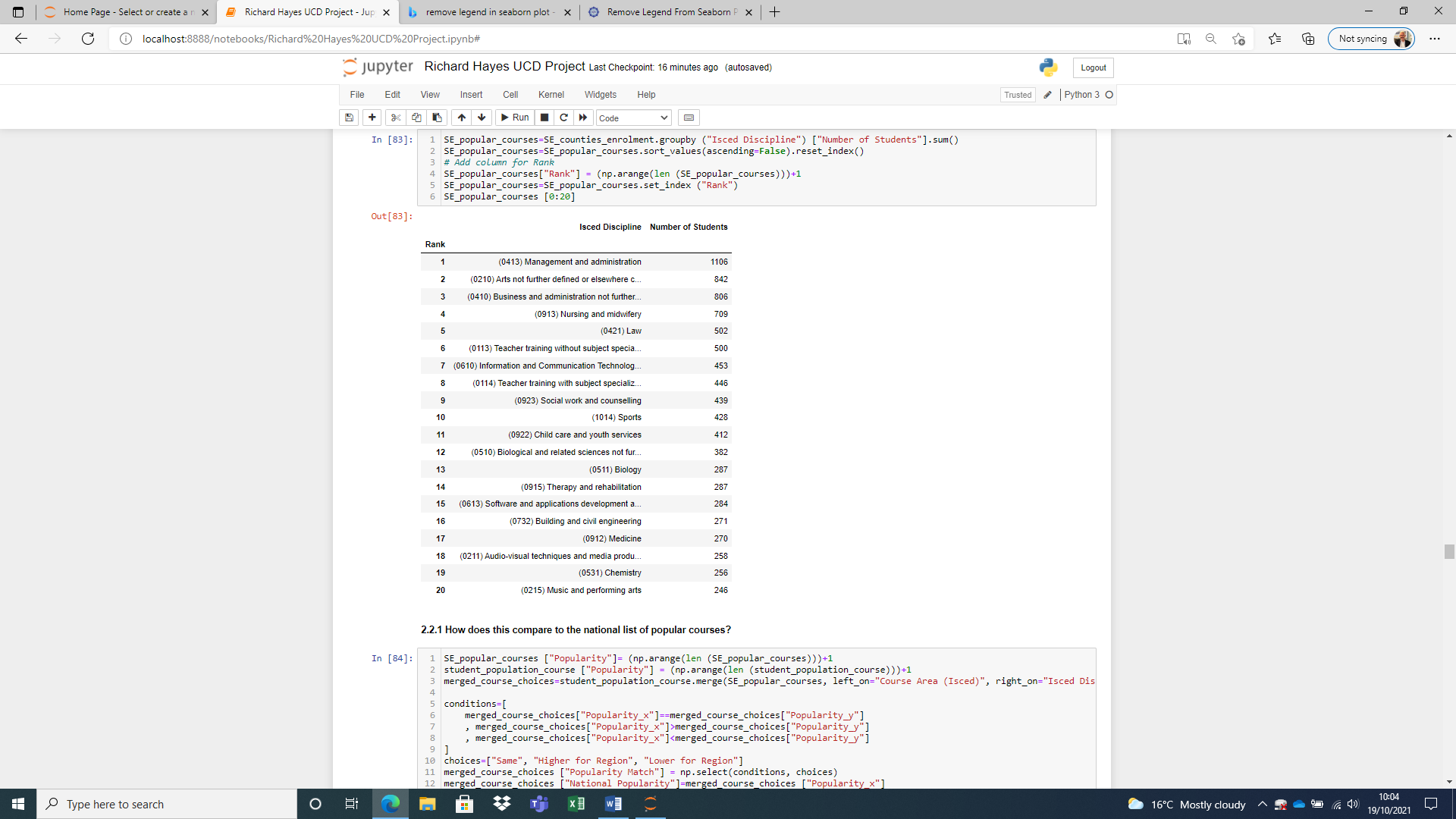


Table Most popular courses with SE students

*Students from the South East Studying in and outside of the South East*

1. As has been indicated above, the most popular colleges for students from the South East are the two colleges in the South East, Waterford IT and IT Carlow. 5522 students attend one or other college, 1686 to Carlow, 3836 to Waterford. A significant number of students from the South East, however, study outside the South East, 9385 in total. In percentage terms, 37% of students from the South East study in the South East, 62.96% study outside the South East.
2. Students from the South East study a range of courses. A comparison of the relative popularity of courses for students staying in the region and leaving the region is interesting. The following table lists course areas that are popular with students leaving the region as against courses that are popular with students retained in the region. This comparison can help identify the rationale for students leaving to study outside the region.

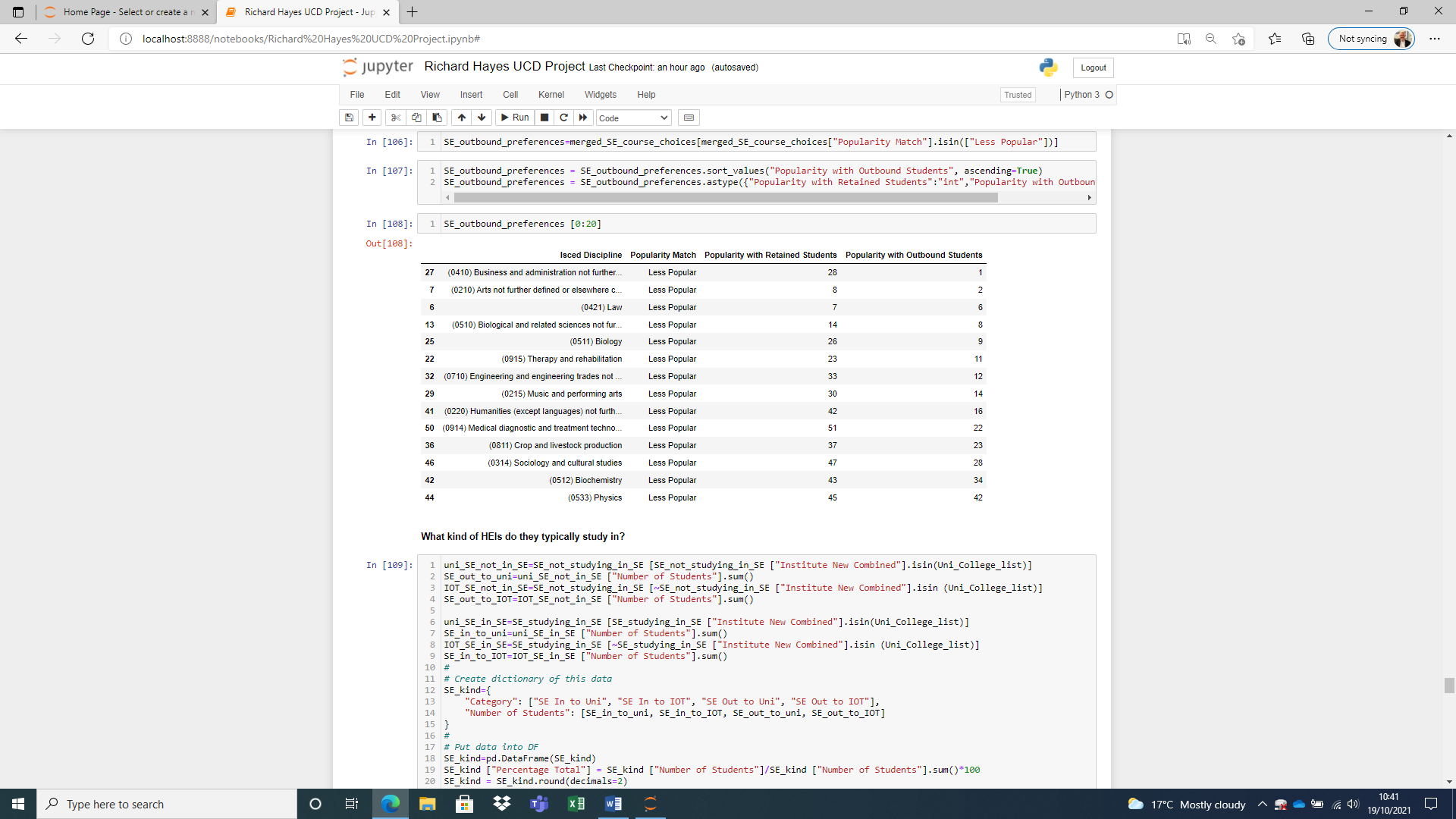


Table Relative Popularity of Courses for Retained/Outbound Students

1. This table, however, does not recognise course areas that might be popular with outbound students and for which there is *no* popularity at all amongst retained students, because these courses do not exist within the region: they have no enrolment in the region. The following table lists the top 20 of these course areas in order of popularity.

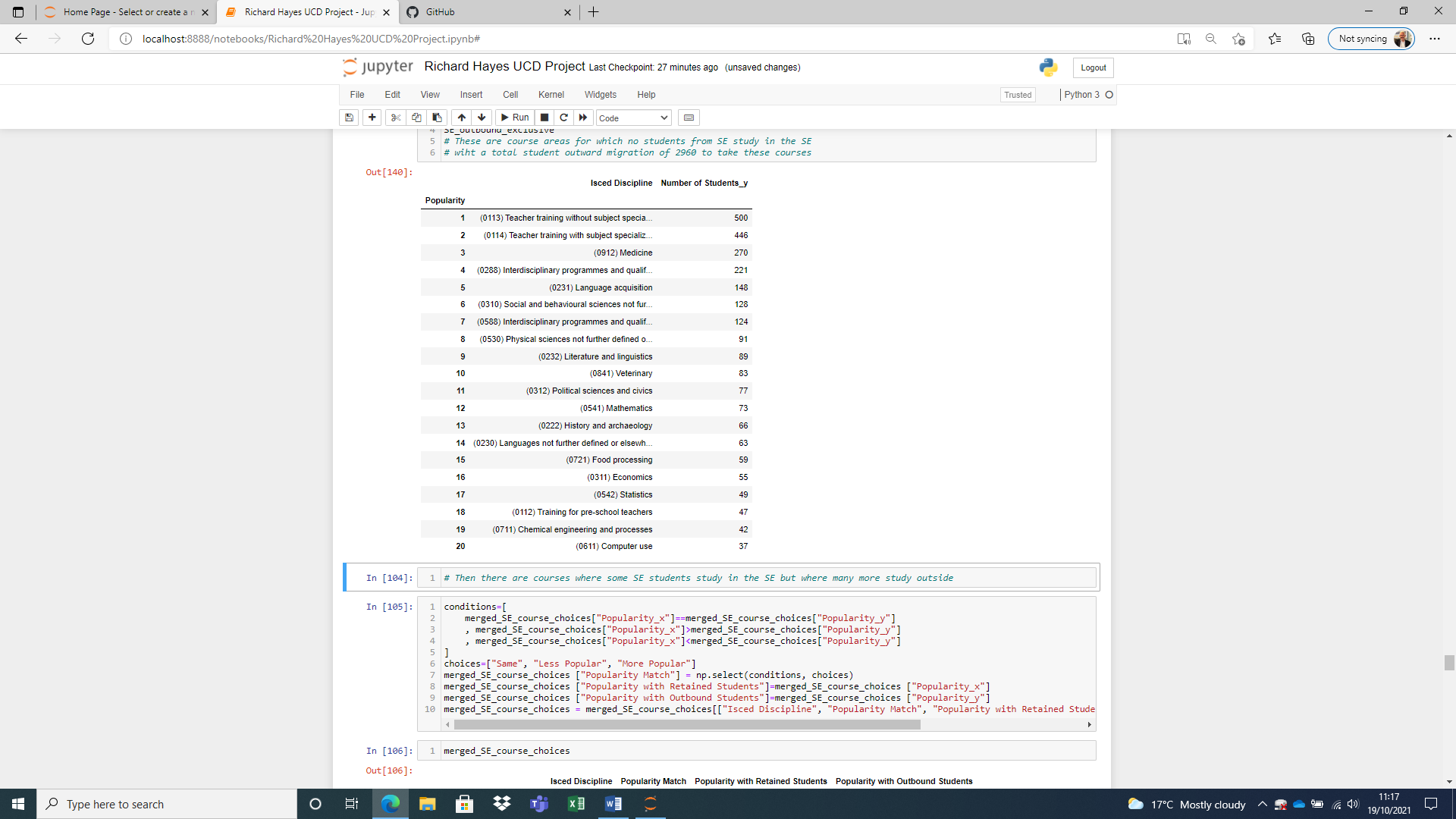


Table Course areas with no enrolment in SE from SE students

1. Outbound students enrol in either universities and colleges or Institutes of Technology outside the region. Examining the distribution of outbound South East students to these different institutional types shows 51.5% of outbound students enrol in a university or college, 11.5% in an Institute of Technology. Retained students all register in an Institute of Technology as the South East does not have a university.

*The South East in Comparison with Other Regions*

1. Some insights, particular with regard to “balanced growth”, can be gained by comparing the profile of the enrolment of students with an origin in the South East with the profiles of student cohorts from other Irish regions. Given the priority in the national development plan accorded to regional cities, the four “city regions” are taken together: the South West (where the city is Cork), the South East (where the city is Waterford), the Mid West (where the city is Limerick), and the West (where the city is Galway).
2. The following shows the distribution of students from each region who study in or out of the region.

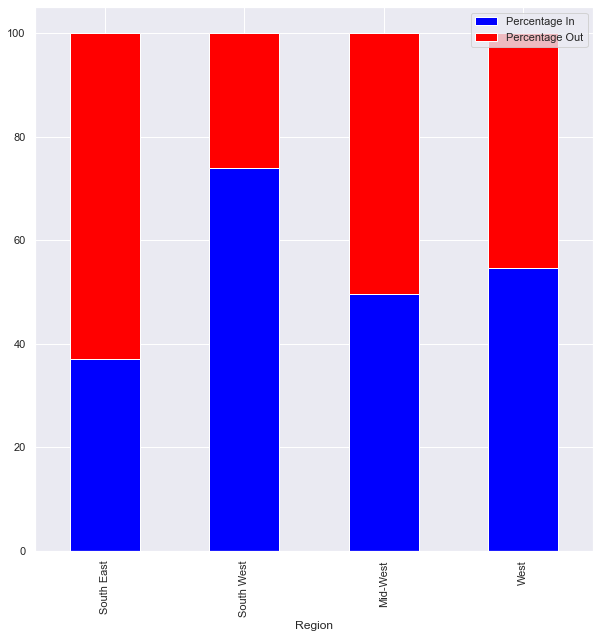


Figure Regions Compared: Students In or Out

The following figure considers the regional profile in much more detail, describing the regional students who study in universities in the region, universities outside the region, Institutes in the region, and Institutes outside.

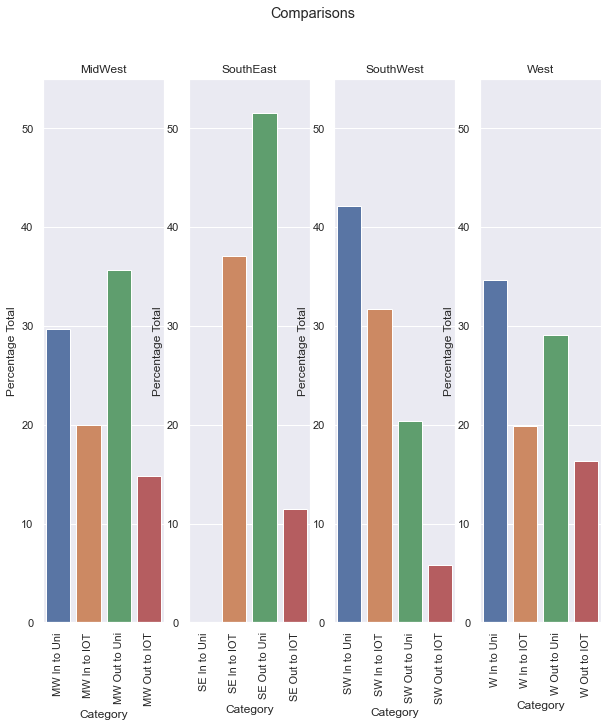


Figure Regional Comparisons by Type

Technological Universities

1. The creation of a new type of university, a Technological University, in the South East might be seen as a means to address the imbalances evident above. At the time the data examined here was published there was one Technological University in Ireland, Technological University Dublin (there are now two more). To test whether a university like TU Dublin would be sufficient to address the needs of outbound South East students, the list of courses available in TU Dublin was tested to see if that institution covers the course areas not available in the South East (Table 5 above). This analysis showed that none of the areas not available in the South East is available in TU Dublin.

**Insights**

This report examines student enrolment data in order to investigate the nature of higher education provision and its relationship with regional “balance”, so critical to Ireland’s development plans.

When considered from the point of view of student domiciliary origins, the data suggests that:

1. There is relative balance between the different regions in Ireland, at both NUTS3 and National Development Plan levels. As Table 1 and Figure 7 above indicate, relatively speaking, all regions seem to send more or less the same proportion of their populations to college with, if anything, Dublin demonstrating a lower performance than the other regions.

However, when considered from the point of view of student *destination*, that is, from the point of view of the colleges where the student ends up studying, then quite a different picture emerges. The data suggests that:

1. There is considerable variation in the profile of students retained within a region and departing a region, with one region showing 75% retention (the South West) while another shows less than 40% retention (the South East);
2. The South East compares unfavourably to the other regions (Figure 8), with a significantly larger percentage than any other region departing the region to study;
3. The departing students opt in the main to study at universities (Figure 9), and the volume of students leaving the region to study at a university is very significantly higher than in other regions. The percentage of students leaving each region to study in an IOT is similar across all regions;
4. Balanced regional development would therefore require addressing the absence of a university in the South East which would seem to be the ingredient that retains significant populations of students in the other regions;
5. The university that should be created in the South East would need to offer courses in areas not currently addressed in the region. Students depart the region for courses that are not available in the region (evidenced by the fact that there are no enrolments in these courses in regional colleges). These courses include, surprisingly, Teacher Training as well as more obvious choices such as Medicine and Veterinary studies. Meriting further inquiry are courses that *are* available in the South East but are more popular outside the South East, such as Arts. One might suggest that the range and variety of Arts subjects outside the region is perhaps wider than within—but this would need to be investigated;
6. TU Dublin offers none of the courses for which university-bound South East students leave, so the Technological University formula, if TU Dublin is typical of it, would not be adequate to retain students in the South East.

**References**

1. See <https://www.gov.ie/pdf/?file=https://assets.gov.ie/166/310818095340-Project-Ireland-2040-NPF.pdf#page=1> [↑](#endnote-ref-1)
2. OECD (2008), *Tertiary Education for the Knowledge Society: Volume 1 and Volume 2*, OECD Reviews of Tertiary Education, OECD Publishing, Paris, <https://doi.org/10.1787/9789264046535-en>, p.13. [↑](#endnote-ref-2)
3. See <https://www.iua.ie/wp-content/uploads/2019/08/Indecon-Independent-Assessment-of-the-Economic-and-Social-Impact-of-the-Irish-Universities_full-report-4.4.19.pdf>, p.54. [↑](#endnote-ref-3)
4. See OECD p.13. [↑](#endnote-ref-4)
5. Southern Region: Regional Spatial and Economic Strategy <http://www.southernassembly.ie/regional-planning/rses> p.194. [↑](#endnote-ref-5)