

**Assessment workshop:**

**Exploring examination outcome statistics**

**Student Workshop**

Every August, examination results for GCSEs and A levels are released in the UK. The outcomes are of enormous personal significance to the learners who are receiving their results, along with their parents and teachers. There is also a great deal of interest at a national level, as people seek to compare the results for different groups of students and different geographical regions between years. Interpreting national examination outcome statistics can, however, be surprisingly perilous.

**Learning outcomes**

* To learn how to interpret examination statistics.
* To observe the variation in outcomes by different variables (e.g., by year, subjects, geographical area, gender and subject).
* To learn to question the causes of these variations.
* To consider the meaning of examination standards.

**Materials**

Ofqual publishes a series of interactive visualisations for examination results (<https://analytics.ofqual.gov.uk/>), which form the materials for the workshop. Similar but static information is published by the Joint Council for Qualifications here: <https://www.jcq.org.uk/examination-results/>. The latter includes results for Wales and Northern Ireland[[1]](#footnote-2).

**The activities**

1. Begin with [A level outcomes in England](https://analytics.ofqual.gov.uk/apps/Alevel/Outcomes/). <https://analytics.ofqual.gov.uk/apps/Alevel/Outcomes/>

Make sure ‘all subjects’ is selected and take a look at the graph entitled “All subjects: cumulative percentage outcomes for all candidates”. Note that if you click on a point on the graph, it gives you the exact statistic.

* 1. What does the graph tell you about exam outcomes?
	2. The passing grade is grade E. Are these exams too easy? What other information would you need before you could answer this question?
	3. Why do the results for 2020 to 2022 look higher than for other years? Is this a reduction in standards?

Now compare the following two graphs, which show the results for male and female students respectively.

* 1. Are the results by gender different? What does this tell you about fairness of the qualifications?

Take a look at the graph showing the number of A level entries, entitled ‘All subjects: number of certificates’.

* 1. Why might the number of students taking the qualifications change in different years?
	2. Why would the number of entries be different for males and females?
	3. What does this mean for comparisons between male and female exam outcomes?
1. Now let’s take a look at outcomes by centre type: <https://analytics.ofqual.gov.uk/apps/Alevel/CentreType/>

You can select all subjects combined or a particular subject of interest, to look at the proportion of grades attained in each centre type over a period of some years.

* 1. What do you notice about the differences between outcomes by centre type? What might the causes of this be?
	2. Are the same patterns shown every year? Why would they change?
	3. What group of students benefited most when the results went up overall? Why would this happen?
	4. Which group was disadvantaged most when results when down? What might the likely causes be?
	5. What would your predictions be for the likely pattern of results in the next year?
	6. What do these patterns tell you about the fairness of these assessments?
1. Look at how results change for individual schools and colleges: <https://analytics.ofqual.gov.uk/apps/Alevel/CentreVariability/>

Here, you must select a particular subject. You can choose to look at all centres or just those with a stable number of candidates in the years of comparison. Begin by looking at the current year statistics, though you can also look at previous years and ‘longer-term’, over a period of years.

* 1. What are your expectations about how much the percentage of candidates getting the grades would change in your school or college, year on year? Do these figures in the graph surprise you?
	2. Why might there be changes year on year within schools or colleges, even if there is little change at a national level?

**Tutor notes:**

These notes are intended to support you in working with students to interpret the graphs and to save you time. They have been written bearing in mind that not all students will be from the UK, so some points may seem self-evident to those who are familiar with the UK system.

**Activity 1**

1. The horizontal, x-axis shows the year. The vertical, y-axis, shows the percentage of entries. Each line shows a particular grade and the grades for A levels are A\* to E.

The proportion of students who took A levels gaining each grade (or above) is given in the cumulative percentages view. This is much easier to interpret than the proportion who simply got each grade.

This graph relates to entries for the examination, not candidates. Candidates typically take 3 A-levels and achieving 3 Bs will get people into a good university course.

Note that grade A\* was introduced in 2010 so that universities could discriminate better between the candidates who were achieving the very highest performances in the examinations.

1. Approximately 60% of the 18-year-old cohort take A-levels; those who do not think they would be suitable take other qualifications.

Here, students might say that they would need to know things about the quality of the education system, teacher qualifications, teacher vacancies, resources available, the content of the curriculum, question papers and marking schemes and so on. They are less likely to mention that there are different subjects underlying these data and the entries for those can vary across years.

1. Different assessment arrangements were made during the pandemic, involving teacher assessments or judgments. This resulted in an increase in outcomes.
2. The results by gender are not, in fact, very different. However, underlying these figures are large gender differences in subject choice. This opens the question of whether all A level subjects are the same standard (which may lead to discussion of what it might mean to say that different A level subjects are ‘comparable’).
3. The size of the cohort changes across years, due to changes in the birth rate and immigration and migration. Additionally, other qualifications may become more or less popular over time. The pandemic also meant that fewer entries were submitted in 2020.
4. There are gendered choices of subjects in these examinations, but these patterns are prevalent around the globe unless there is no choice about which subjects are taken.

**Activity 2**

1. The horizontal, x-axis shows the year. The vertical, y-axis, shows the percentage of candidates gaining that grade or above. Each line shows a different centre type.

Independent schools tend to do better than the other centre types.

Often Further Education colleges have weaker results.

Socioeconomic advantages of the students and the schools affect the outcomes.

1. The results rose over the pandemic for all centre types and went back down in a staged manner after it. This was a matter of government policy on standards.
2. Over the pandemic, the outcomes went up overall and the gap between independent schools and other centres widened. This could be due to socioeconomic advantages. However, teacher assessments were used during the pandemic and it may be that independent schools were more generous in their grading.
3. The gap between FE colleges and other centres widened in 2022. This may be due to socioeconomic differences between learners and the institutions which impacted on student performances in the examinations that year.
4. This is an opportunity to discuss the effect of standard setting policy on outcomes. If there is a policy of maintaining the national level statistics, then we might expect a flat line, for example, as we see (more or less) between 2018 and 2019. If there is a policy to fund FE colleges to a larger extent, we might expect a reduction of the gap between their outcomes and other centres.
5. Assessments are supposed to produce a meritocratic society, yet we see these sociological patterns in the outcomes. This is an opportunity to discuss whether fairness should be restricted to the standardisation of the assessments or a broader definition should be considered.

**Activity 3**

1. The horizontal, x-axis shows the change in the percentage of candidates in a school getting a grade A and above. The vertical, y-axis, shows the number of centres with that percentage of change.

Most centres have a small percentage of change (bunching around zero) and a small number of centres have large fluctuations, year on year.

If the centre size is set to large centres (60+), then we see fewer centres (of course) and the changes are far smaller.

Small centres represent small samples of students. Statistically, we are more likely to get fluctuations with small samples because a few people can influence the statistics more and they could be extreme cases (very good or very weak students).

1. Although there are changes in the percentages for each centre every year, the overall national statistics could be almost identical. Rises in outcomes in one centre can be offset by a fall in another centre.
1. Other relevant statistics include the Qualifications Wales Official Statistics release ‘[Variation in GCSE, AS, and A level qualification results](https://qualificationswales.org/publications-resources/variation-in-gcse-as-and-a-level-qualification-results-2023/)’, which provides information about the distribution of changes in results at a centre level. [↑](#footnote-ref-2)