# Examination of continuation of STEM studies by gender

I have chosen to investigate the continuation of study of STEM subjects in secondary school by each gender. I have performed this investigation by using the data available via the [SQA](https://www.sqa.org.uk/sqa/57518.8313.html), which tracks the number of students each year, and level of achievement, by gender. I have initially tracked the number of students studying each subject from National 5 level in 2017 to Advanced Higher in 2019. I have charted the percentage of each gender’s progression from National 5 to Higher, and Higher to Advanced Higher.

The data shows a notably higher continuation in study, from National 5 to Higher, in Males in Administration and IT, Computing Science and Mathematics. There is also a notably higher continuation in study in Females in the subjects of Business Management, Design and Manufacture, Engineering Science, Environmental Science and Graphic Communication. The continuation of study to Advanced Higher level sees Males notably higher in Mathematics and Physics, whereas Biology, Chemistry and Graphic Communication sees a higher percentage continuation by Females.

After this initial examination I thought it would be prudent to examine whether achievement of the precursor level had influence on the percentage of students continuing study. I have taken the number of students who achieved a grade between A to C grade and found the difference between this number and those who have continued study in each subject, giving the “Voluntary Drop Off” of students in each subject.

The is notably higher “Voluntary Drop Off” in National 5 to Higher in Females in the following subjects: Physics, Mathematics, Computing Science and Administration and IT. This means that Females achieved a passing mark in the following subjects but chose not to continue study in the next level. In Higher to Advance Higher, Females continued to drop Physics, Mathematics and Computing Science at a higher percentage than the Males. In National 5 Males chose not to continue the following subjects: Engineering Science and Design and Manufacture. This is contrasted by a significant difference in pick-up of Administration and IT. In Advanced Higher Males had notably higher levels of drop off in Chemistry and Biology.

Upon seeing the significant difference level of drop off in Computing Science between Males and Females I felt it prudent to investigate whether this is a reoccurring trend or a one-off abnormality. This led me to charting the percentage of drop off from National 5 to Higher over the years 2015-2019. This chart shows a yearly difference of roughly 10% in each gender dropping Computing Science from National 5 to Higher.

This investigation shows there are certain subjects which appear to have gendered preference for continuation of study, such as Physics, Mathematics and Computing Science with Males and Biology and Chemistry with Females. This is especially apparent when considering the “Voluntary Drop Off” of Females in the subjects: Physics and Computing Science. This raises questions on why Females are choosing not to continue in these subjects at a higher level than Males as Females are already underrepresented in these subjects. I believe these concerns would warrant further study along with investigation into each subject to see if there is a concurrent difference in drop off, as there is in Computing Science.

Notes on bias in data and visualisation

The datasets used have been pulled directly from the SQA website and contain no algorithmic calculations or geographical bias, as this is the uninfluenced data for all Secondary Students within Scotland studying at each level. The one concern with the provided datasets is that the percentages are independently rounded, therefore can minimally affect calculation, as they may not add to 100%. The visualisation I have created concern the percentage continuation of each gender and do not reference the number of students studying each subject, therefore are not influenced by the representation numbers of each gender, making this a fairer analysis.

Ethics of examination

I do not believe there are any ethical concerns in the investigation of the continuation of STEM studies by gender, as each gender are compared by percentage rather than student representation in each subject. This study could by completed by referencing each gender anonymously to ensure there is no unconscious bias or ethical concerns. This study shows that there is gendered preference of continuation is certain subjects by both genders, which warrants further investigation.

Cleaning and Wrangling of Data

I have downloaded the relevant datasets directly from the SQA, from these datasets I have copied the relevant sheets to a new workbook in Excel. In the new workbook, I have cleaned the data by removing the rows for each none STEM subject. This makes the data more manageable and visually appealing. I have then created a new worksheet to complete all workings and calculations, along with the creation of data visualisations. I have kept each worksheet independent and reference the data directly to ensure accuracy and remove chance of contamination.