Much of the analysis about why the UK voted to leave the European Union in June 2016 has been done by looking at individual factors in isolation, or using opinion poll data from both before and after the vote.

In a new research paper, I applied two statistical analyses to the actual referendum voting data obtained from the Electoral Commission and the UK’s latest census data. I found that while voters’ level of higher education was the most important factor, the gender of voters and the turnout level also had parts to play in the victory for the Leave campaign.

The first method I used was called a “multivariate regression” analysis – a powerful statistical technique used for predicting the unknown value of a dependent variable, such as the percentage of Leave votes, from variables that can explain it, such as education, turnout, age and gender. This method can answer the question: “How much does the percentage of Leave votes change when we alter a significant factor, keeping other factors unchanged?”

The other method I used, called a “logit regression”, is used to predict the odds ratio – the relative chance of an event occurring – of a majority of people voting to leave the EU. It can answer the question: “How much does the probability of voting for Leave change when we alter a significant factor, keeping other factors unchanged?”
The statistical analyses took into account various potential factors, such as turnout, income level, gender, education, unemployment, age, whether the person was UK-born and their social grades. They also took into account the change in British adult population between 2011, when the last census was conducted, and the referendum in 2016 to identify statistically significant factors that could have had an influence on the result and then to quantify their impact.

The influence of university

I found higher education was the predominant factor dividing the nation — something a number of previous analyses have also focused on. A heat map I produced with the data tells the story.

The two maps show which areas voted in the majority for Leave and Remain in England and Wales. The national average in Britain is for 26% of British adults to have a university degree and above. The colours of the map show which areas are higher (green) or lower (red) than this national average. Most areas in which a majority of people voted for Remain (the right-hand side of the map) have a higher proportion of British adults with higher education than the national average. Areas with in darker red have a lower proportion of British adults with higher education.

I found that higher education alone explains about 77% of the total variation in the referendum vote across areas in England and Wales.

Areas in Scotland with a higher proportion of university-educated British adults also showed a higher turnout in favour of Remain.

Turnout, jobs and gender

The statistical analysis can also be used to suggest that, keeping other things equal, a decrease of about 7% in turnout in England and Wales could have reduced the total UK Leave vote by around 4%. Theoretically, this could have resulted in a Remain victory overall because of the close nature of the result, in which Leave got 51.9% of the vote and Remain 48.1%.
An increase of about 3% in the proportion of British adults accessing higher education in England and Wales could also have reversed the referendum result, provided they voted in a way consistent with the data.

As the graph below shows, areas with a lower unemployment rate tended to have a higher turnout in England and Wales, regardless of whether the majority voted Leave or Remain.

This phenomenon supports arguments by the American political scientist Benjamin Radcliff that a poor economy discourages the participation in democracy of those voters most adversely affected by it.

The right upper side of the graph shows that areas in England and Wales with a lower unemployment rate tended to have a higher turnout in support of Leave.

There was also highly significant evidence suggesting that demographic changes of male adults in individual areas had an effect on the referendum outcome. Areas with a higher proportion of British adult men (who were eligible to vote in the referendum) were associated with a higher percentage of Leave votes. Those areas with a higher percentage of British adult women had a lower percentage of Leave votes.

By my estimate, if there had been an extra 1% of men in England and Wales this could have pushed the proportion of Leave votes up by 1.18%. For areas in Scotland and Northern Ireland, an extra 1%
would have pushed the Leave vote percentage up by 4.04%.

In my multivariate regression analysis, the age of voters appeared to be less significant than education, turnout and gender. Although a higher proportion of elderly people in Britain contributed to a relatively higher percentage of Leave votes, these extra Leave votes did not lead to a Leave outcome in individual areas on their own, as the factor of elderly British became insignificant in the logit regression analysis I did for the whole UK. This implies that reports in the wake of the referendum that it was older people who won the vote for Leave, have been overstated.