Big data COVID.
COVID-19's impacts in relation to the North-South divide of England.

Forge Valley School   Dylan   Holly
Heidi   Amber

Summary of project:

The focus of the big data COVID project was to learn to appreciate, understand and apply the uses of data and modelling on planning for civil disaster, as well as monitoring and limiting the impacts of disaster when it occurs.

- **Data science** is using mathematical techniques and literacy to learn from data.
- **Big data** refers to complex data gathered and updated quickly, often requiring more than normal computing methods to store, the largest data sets being stored in scalable warehouses; big data within COVID includes track and trace data and treatment history.
- **Data Modelling** includes using big data, as well as making plausible assumptions to create meaningful predictive models for future changes to the data.

In this project we researched and analysed our own data using the skills in data comprehension and analysis we gained from the excel training and seminar 3, to interpret the data and create resources based on this.

Background information:

COVID-19 is an infectious disease caused by SARS (severe acute respiratory syndrome) -COV-2 virus which is transmitted by airborne particles that contain the virus. This disease affects the respiratory system and symptoms include: a continuous cough, shortness of breath, fatigue and headaches. The virus was first identified in Wuhan, China in December 2019. Following this, the virus spread worldwide causing a pandemic which is still prevalent today with 519 million cases around the world and counting. Since this first case, the virus has mutated many times into new variants including: alpha, beta, delta and omicron. Each of these have different R_0 (the reproduction number), showing the average number of people that one infected person will infect.
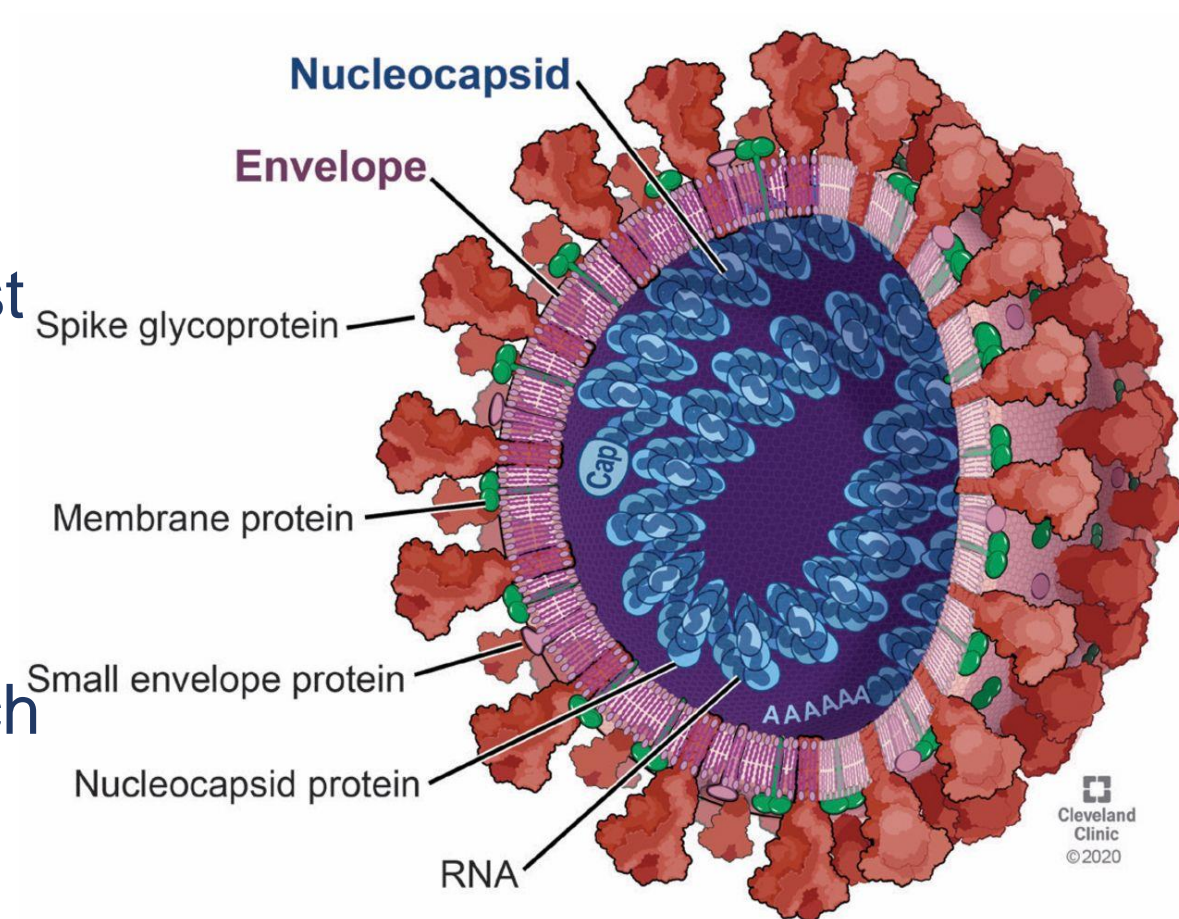
Image above: labelled diagram of the structure of a coronavirus. From: https://www.ccjm.org/content/87/6/321

This virus has affected millions around the globe, but our research is focused on the differences between the North and the South of England due to differences in treatment from the government towards regions of the UK, for example differences in funding and resource allocation as well as differences in culture.

Image left: Map of the UK with its regions labelled. From: https://ukmap360.com/united-kingdom-(uk)-region-map

Research aims:

- To compare how COVID impacted the North and South of England.
- To identify the reasons for these differences.

To ensure a fair comparison between regions a Northern region and a Southern region with similar populations were chosen: Yorkshire & Humber and the South West.

Topics we were interested in ranged from the impact on education to the impact on healthcare and the NHS. We also had an interest in the virology of COVID and how the different variants spread in different regions.

Key questions:
- Why may there be a differences between regions?
- Were some variants more widespread and severe in some regions compared to others?
- Was there a political bias in the allocation of budgets to different hospitals and schools which had a direct impact on difference in performance between regions?
- Did some regions have more access to resources that others didn't?
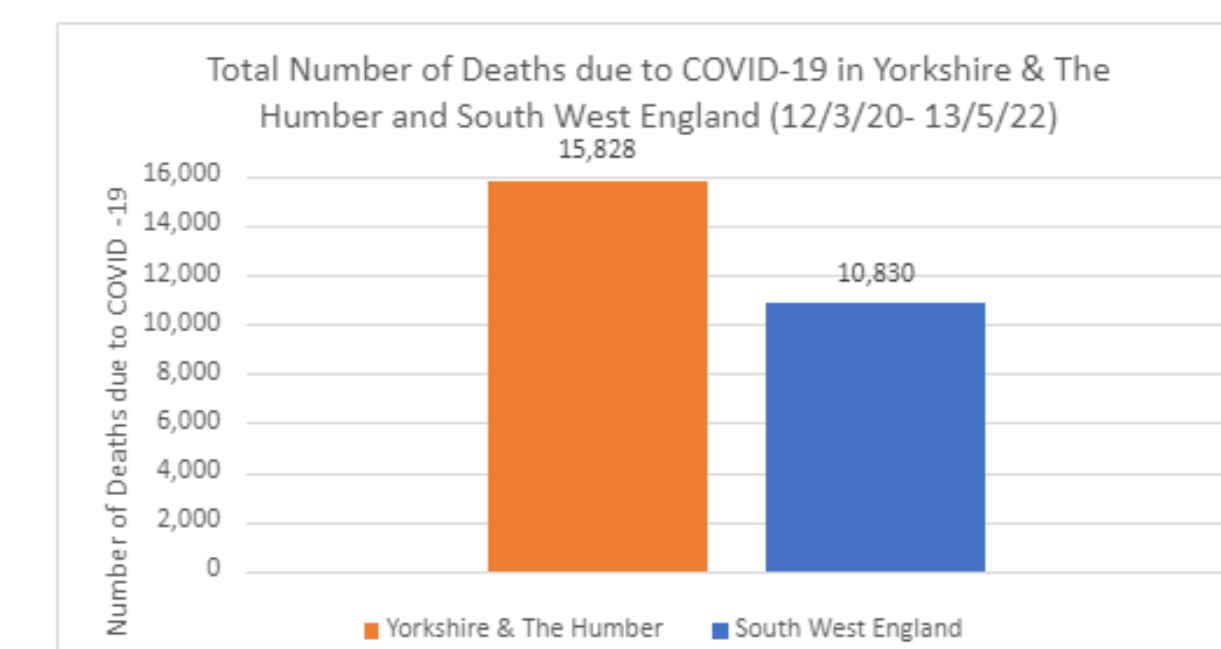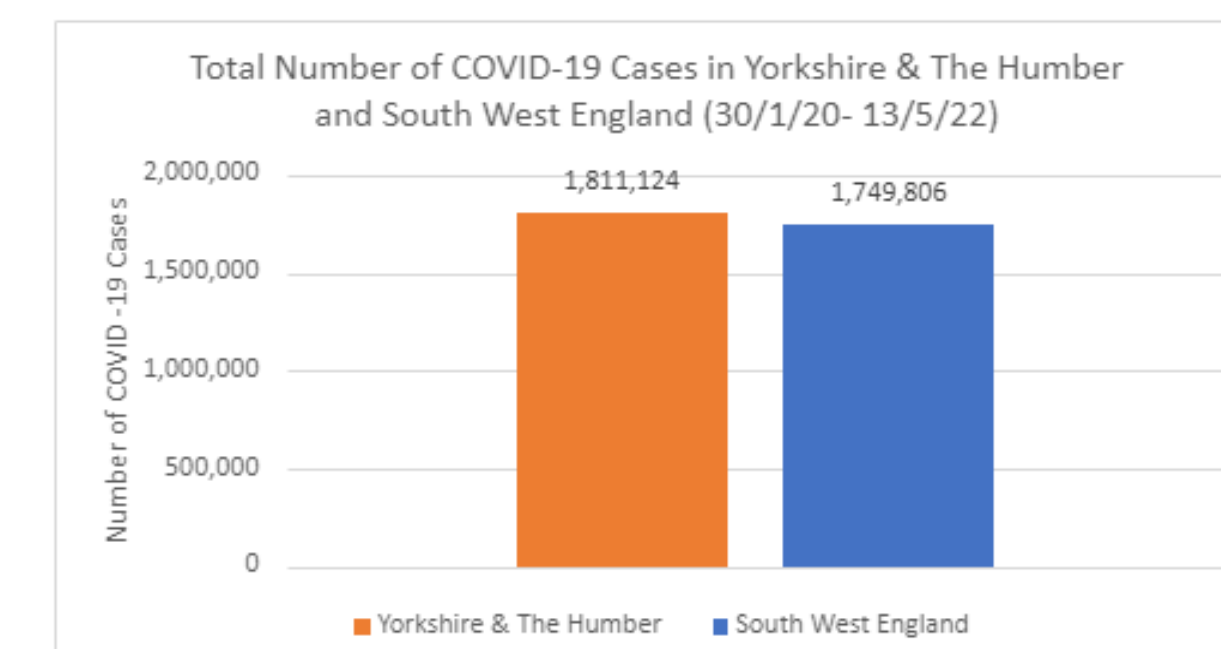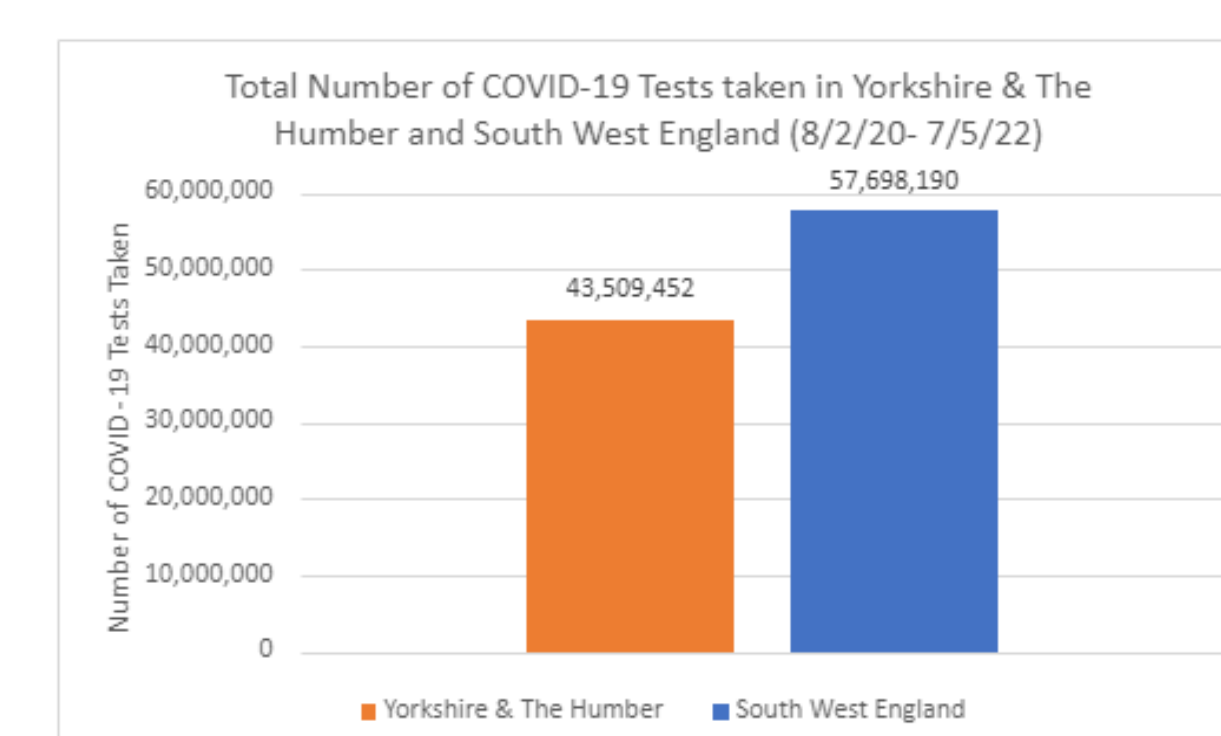
Experimental Method:

To obtain our data, we looked at various government websites for accurate and reliable statistics about the impact of COVID-19 on Yorkshire & The Humber and South West England. The data you see about the effect on healthcare was appropriated from https://coronavirus.data.gov.uk/ and processed into our own graphs on Microsoft Excel. Furthermore, for the education research, we used the government website's create your own table function to get the data on absences. We also used models to simulate data.

Predicted results:

We would predict there to be more cases, deaths and hospitalisations in the Northern regions of the UK. We would predict this because of expected allocations of PPE and funding. We would expect more cases and deaths in Yorkshire and the Humber due to higher population density.

Research into healthcare:

South West England – 5.616 million (2019)
Vaccinations (8 Dec 2020 – 11 May 2022): 12,725,627
Yorkshire & The Humber – 5.486 million (2019)
Vaccinations (8 Dec 2020 – 11 May 2022): 11,343,889

Total Number of COVID-19 Tests taken in Yorkshire & The Humber and South West England (8/2/20- 7/5/22)

Total Number of COVID-19 Cases in Yorkshire & The Humber and South West England (30/1/20- 13/5/22)

Total Number of Deaths due to COVID-19 in Yorkshire & The Humber and South West England (12/3/20- 13/5/22)
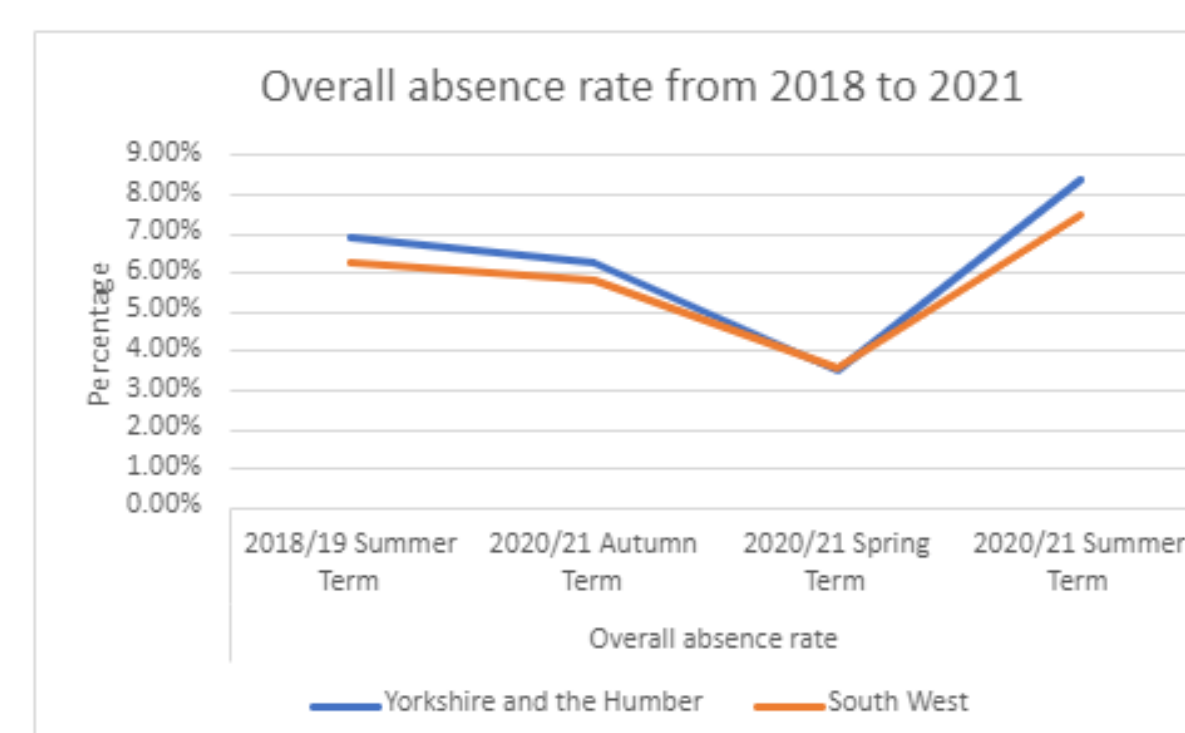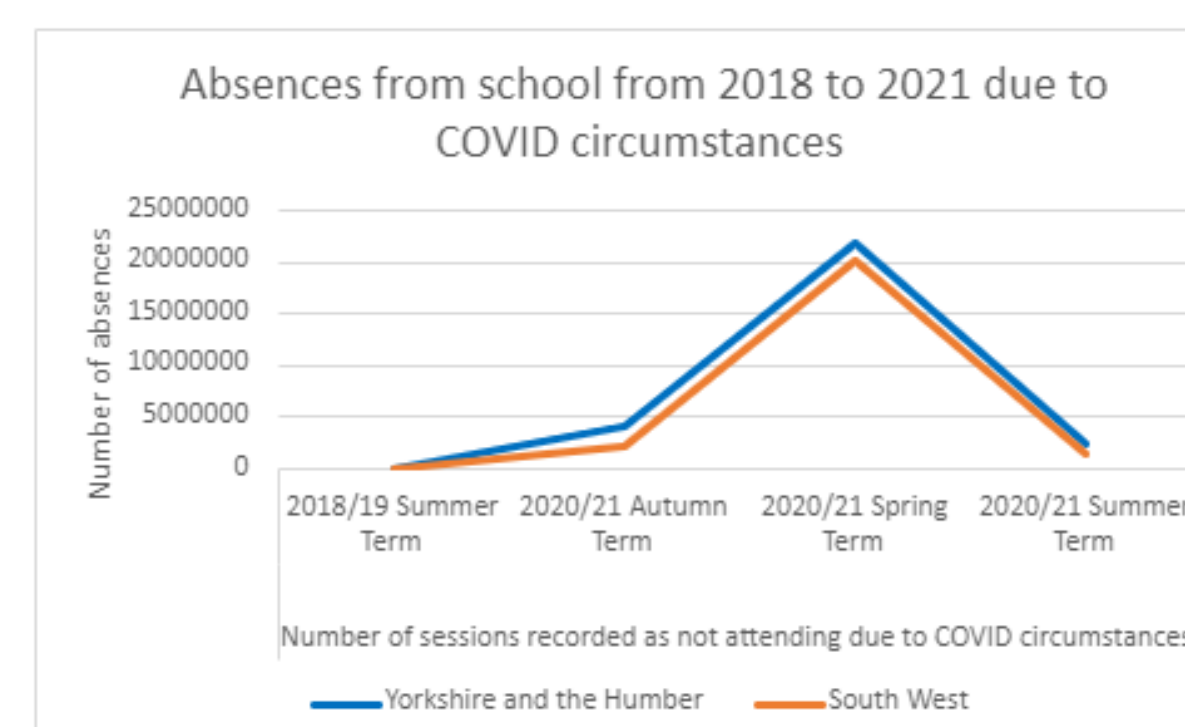
Conclusions:

Our research into the impact of COVID-19 on healthcare in Yorkshire & The Humber vs South West England shows that the increase in number of vaccinations administered could have led to a lower number of COVID cases and 5,000 less deaths. We can also see that there was a much higher amount of testing in South West England implying the data obtained is more accurate. Also, the lower amount of testing in Yorkshire & The Humber suggests that there may be a significantly larger amount of unreported COVID-19 cases.

Research into education:

Comparing absences due to COVID-19 between Yorkshire and the Humber and the South West. Also comparing the overall absence rate between the two regions.

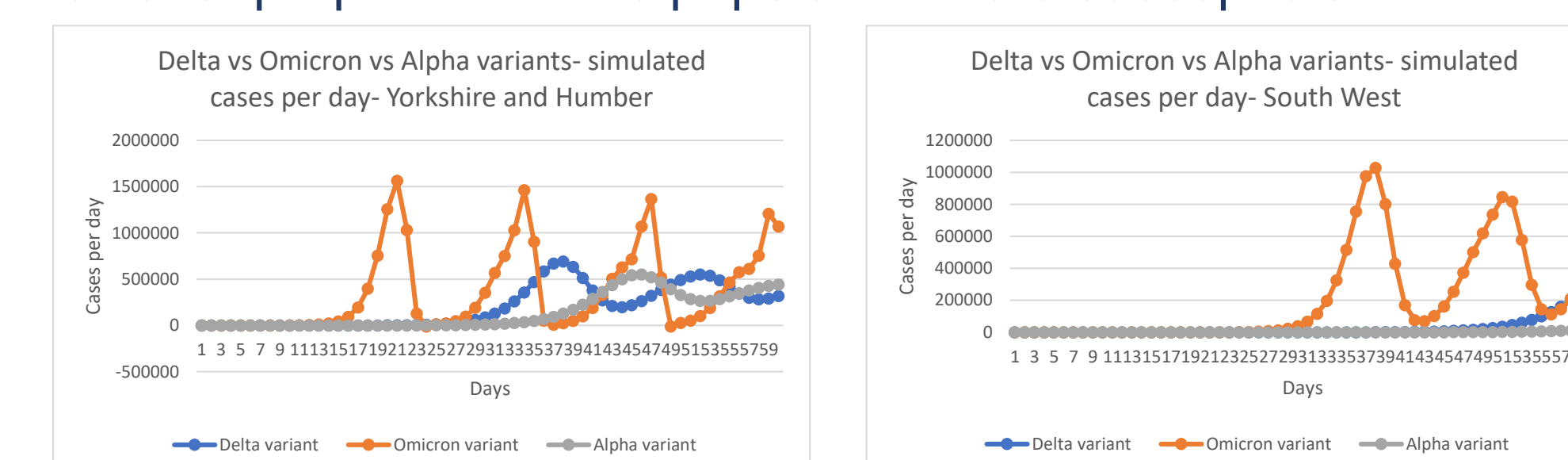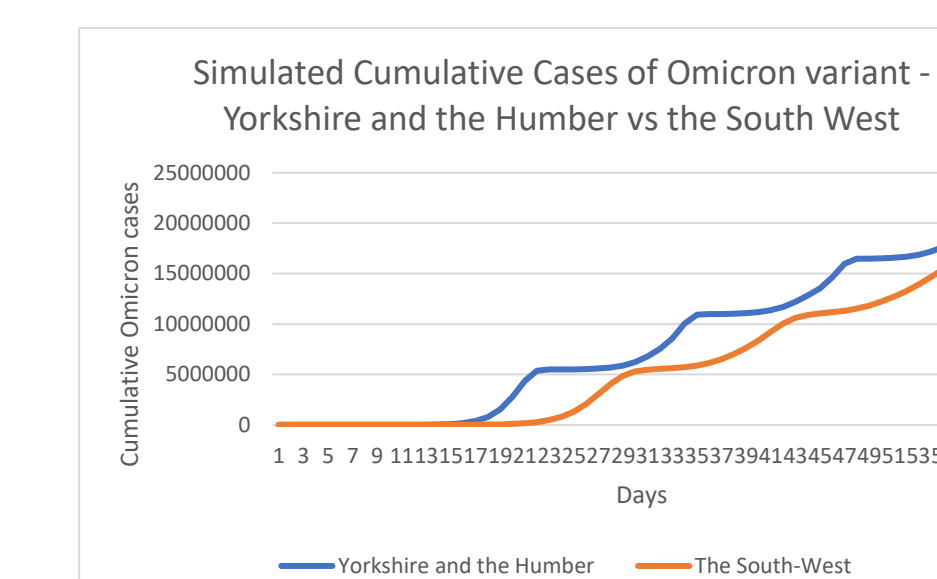Absences from school from 2018 to 2021 due to COVID circumstances

Conclusions:

Interestingly, the number of overall absences decreased as the number of absences due to COVID increased during the spring term of 2020/21.

Overall absence rate from 2018 to 2021

Modelling and Simulation:

We created a model based on SIR modelling for COVID growth of three variants of concern in order to compare the impacts of each variant in Yorkshire and the Humber versus the South West. The model took into account the differing R_0 of each different variant. It also considered the impact of population density on the spread of COVID and how R_0 change when a smaller proportion of the population were susceptible.

Delta vs Omicron vs Alpha variants- simulated cases per day- Yorkshire and Humber

Delta vs Omicron vs Alpha variants- simulated cases per day- South West

The simulated results show clearly that the impact of each COVID variant is worse in Yorkshire and the Humber, the number of cases increases more quickly, and the peaks in daily cases are higher. This is expected due to the higher population density in Yorkshire and the Humber (355 people km^{-2} compared to 236.07 people km^{-2}). This effect is even greater when considering Omicron variant due to the naturally higher R_0, the change in population density had an even larger impact on the daily cases and cumulative cases then the other variants.

Simulated Cumulative Cases of Omicron variant - Yorkshire and the Humber vs the South West

Further research:

With further time and resources we would:
- Discover how the impact of different variants of concern spread through the populations of Yorkshire and the Humber versus the South West. This could be linked to differences in population density and the reproductive number of the different variants. Although we modelled this, we would compare our models to real data.
- Research into the allocation of funds and resources in hospitals, such as PPE, in the different regions to decipher whether this was a factor into the mortality rates.
- More research into COVID's impact on education and how it differed in the North-south divide. We would've gone into detail about grades achieved vs expected grades, cases reported, absences due to COVID and support that schools received from the government.
- Continue to improve the simulated models, taking into account vaccinations, attitudes to restrictions, and social distancing.

FORGE VALLEY

IRIS
The Institute for Research in Schools