



DEPARTMENT OF COMPUTER SCIENCE

MSc Big Data Technologies

MODULE CODE: 7BUIS009W

MODULE TITLE : DATA VISUALISATION AND DASHBOARDING

MODULE LEADER: ASHIF TEJANI

COURSEWORK

SUBMISSION DATE: 26th April 2021

STUDENT ID: w1813148

STUDENT NAME: PATEL BHAUMIKKUMAR SHAILESH

1. Introduction

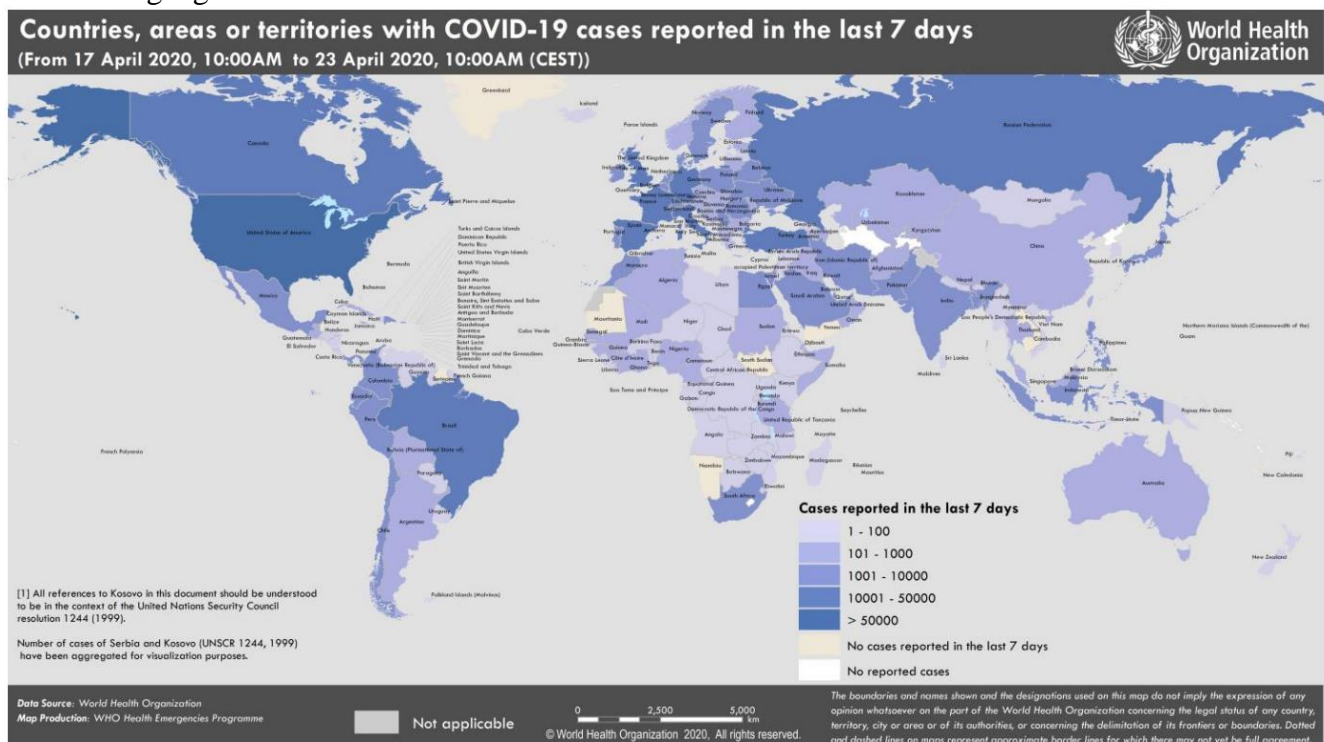
Impact of the Coronavirus infection

As per the WHO, coronavirus (also known as COVID-19) cases of pneumonia of (unknown) cause in Wuhan City, China. A novel coronavirus was identified by Chinese authorities on 7th January 2020 and was named “2019-nCoV”

The WHO Director-General reconvened the Emergency Committee. This was prior to the 10 days period and only 2 days after the primary reports of limited human to human transmission were reported outside China. The Director General accepted the advice and declared the novel coronavirus outbreak a Public Health Emergency of International Concern. This can be the 6th time WHO has declared a PHEIC since the International Health Regulations (IHR) came into force in 2005. WHO’s situation reports for 30 January reported 7818 total confirmed cases worldwide, with the bulk of those in China, and 82 cases reported in 18 countries outside China. WHO gave a risk assessment of very high for China and high at the world level.

According to the data available on DATA.GOV.UK, up until February 2021, there are 2.2 million people identified as clinically extremely vulnerable in England because of their clinical condition or a clinician’s review, in February, a COVID-19 population risk assessment was undertaken and an additional 1.5 million people were identified as CEV. A total number of 3.7 million clinically extremely vulnerable people noted in England. Half of clinically extremely vulnerable those that were responsive to government advice to shield reported completely following the guidance that’s 49%.

The positivity rate is that the percentage of individuals who have tested positive for COVID-19 at some extent in time. We use current COVID-19 infections to mean testing positive for SARS-CoV-2, with or without having symptoms, on a swab taken from the nose and throat. this can be different to the incidence rate, which may be a measure of only the new polymerase chain reaction (PCR)-positive cases during a given fundamental measure.



2. Acquisition

All the data sets from the data.fov.uk or web links pointed by itself. All the details of how to reach each dataset is given the “Link to resources” section at the end of the document.

Followings are the key datasets used for this coursework.

Dataset Name	Coronavirus (COVID-19) Infection Survey: England
Published by	Department of Health and Social Care
Published date	23 April 2021
URL	https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/coronaviruscovid19infectionsurveydata

Dataset Name	Coronavirus (COVID-19) Infection Survey: Northern Ireland
Published by	Department of Health and Social Care
Published date	23 April 2021
URL	https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/covid19infectionsurveynorthernireland

Dataset Name	Coronavirus (COVID-19) Infection Survey: Scotland
Published by	Department of Health and Social Care
Published date	23 April 2021
URL	https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/covid19infectionsurveyscotland

Dataset Name	Coronavirus (COVID-19) Infection Survey: Wales
Published by	Department of Health and Social Care
Published date	23 April 2021
URL	https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/covid19infectionsurveywales

Dataset Name	Coronavirus (COVID-19) Infection Survey: Technical Data
Published by	Department of Health and Social Care
Published date	23 April 2021
URL	https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/covid19infectionsurveytechnicaldata https://www.ons.gov.uk/visualisations/dvc1322/officialaestimates/datadownload.xlsx

3. Data Preparation

All the data sets used were in report format and most of them were in Excel format with some complex structure like tabs, styles, tables. So, this data not possible to use directly without modify it. A good understanding of the data set can lead to better success in achieving the goal of data visualization.

Data preparation is that the method of cleaning and transforming information before processing and analysis. It is a crucial step before processing and sometimes involves reformatting data, making corrections to data and thus the combining of data sets to complement data. Data preparation is typically a lengthy undertaking for data professionals or business users, but it's essential as a prerequisite to put data in context so on show it into insights and eliminate bias resulting from poor data quality. For example, the data preparation process usually includes standardizing data formats, enriching source data, and removing outliers.

There are several benefits of Data preparations that's helps preparation of efficient, accurate clean data. Data Preparation helps: quickly fixed data errors, Produce top Quality data

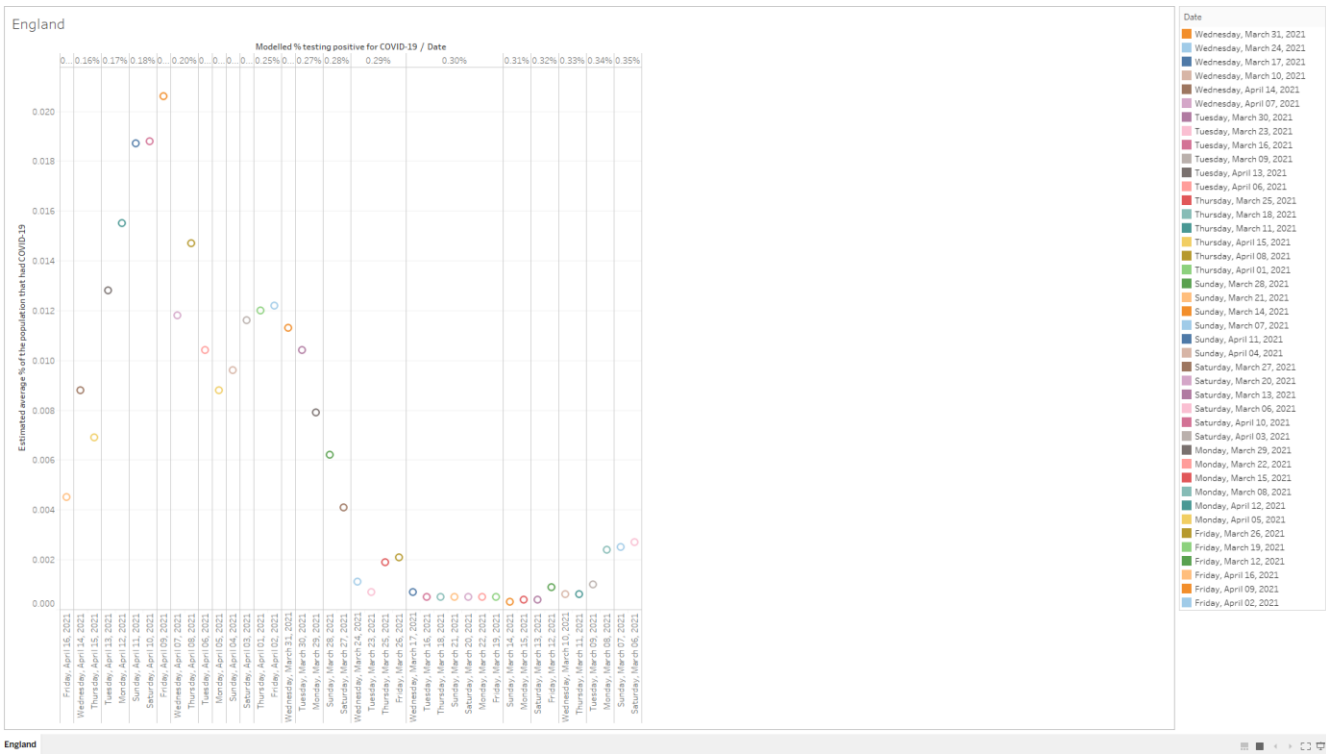
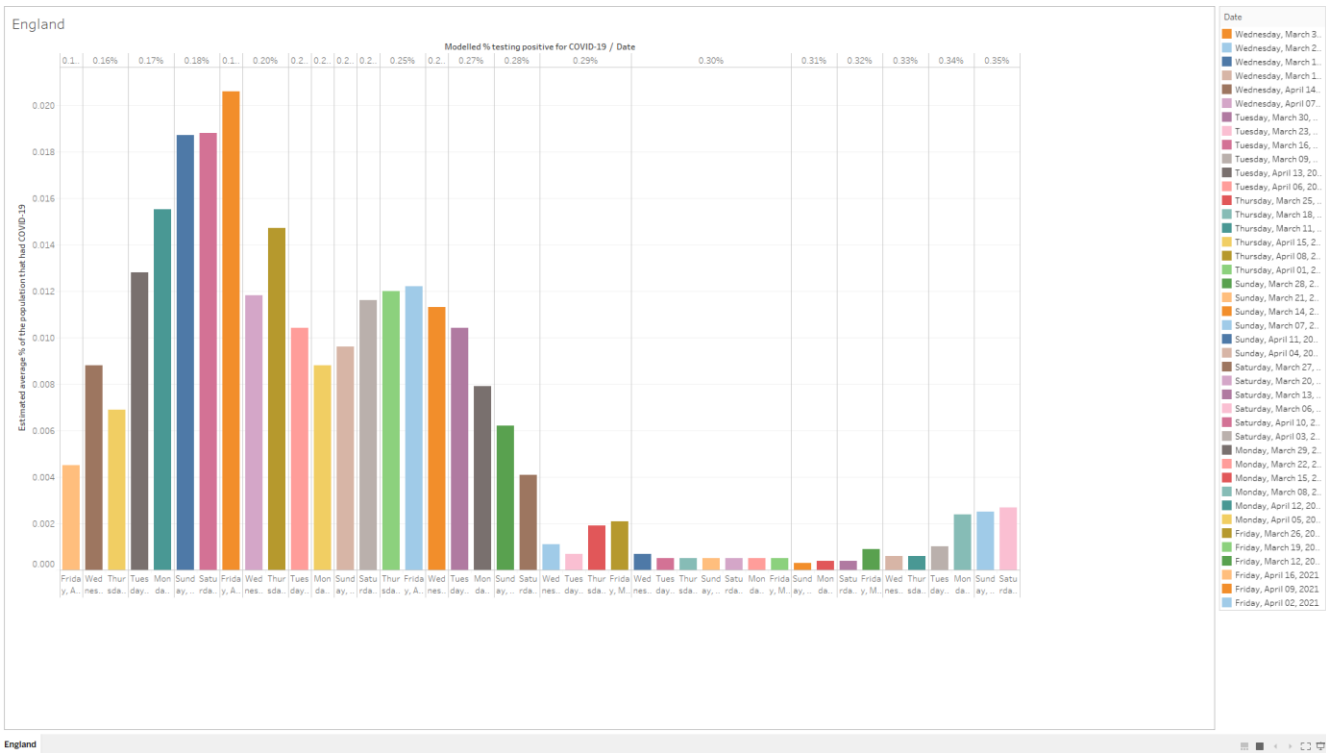
This data required some types of formatting so we can analyse easily in Tableau or R studio. I delete some inappropriate & irrelevant data from rows and column. This was a vital in the case of result of the column. Few columns contain percentile of the people's infection rates from region of UK. I also edit the heading of each column with the appropriate name display in Tableau. Filter Option in MS Excel aiding in detecting and removing all irrelevant data fields from the datasets.

All data sets were in different sources and formats therefore merged then together creating one dataset to support the visualization needs.

4. Discovery

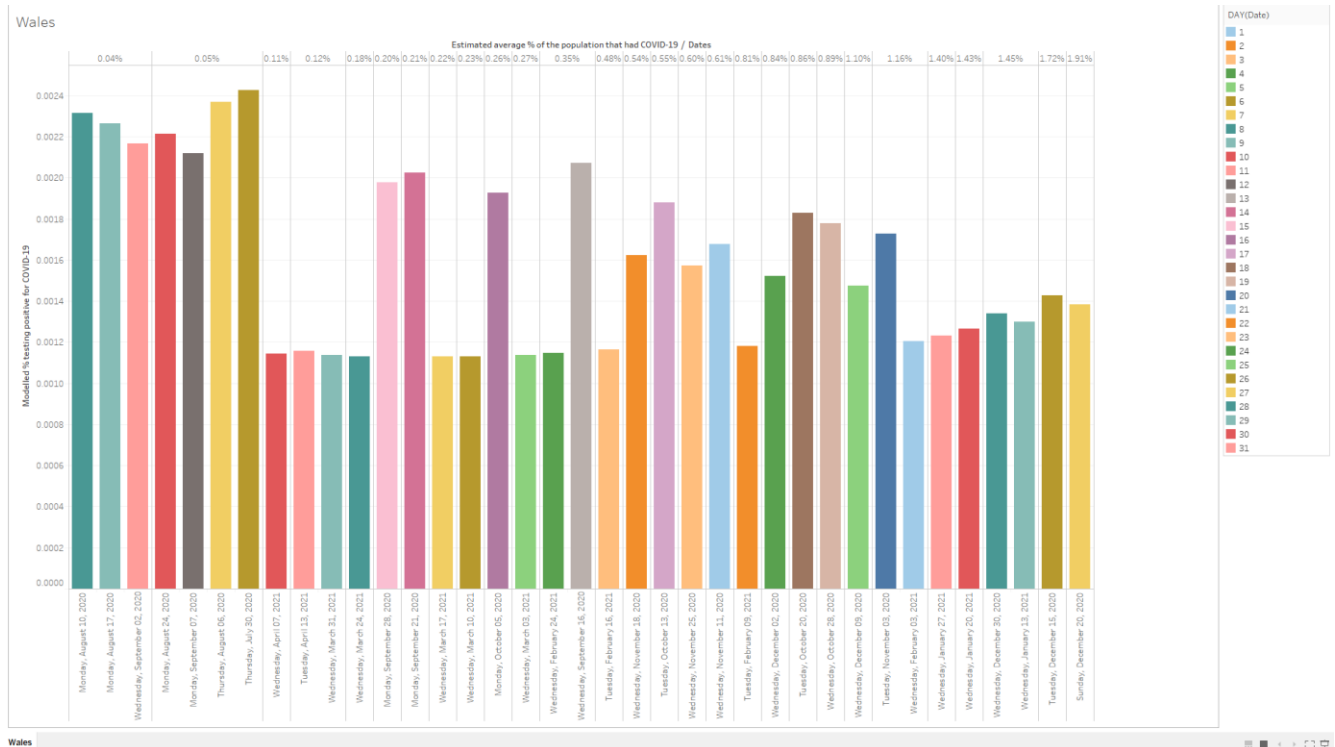
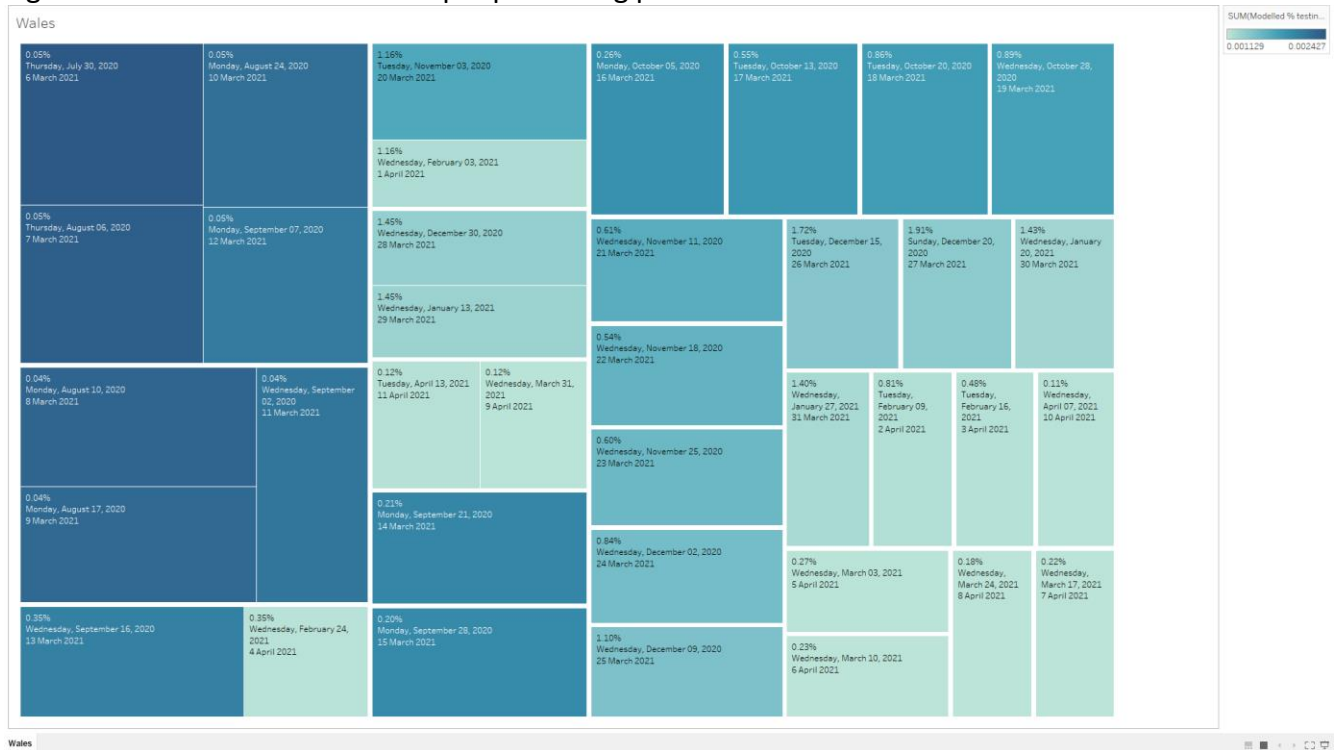
As per the Data report, the estimate of 90,000 people in England had COVID-19. This equates to 0.17% of the England's populations. Another report suggests that the percentile of the positive test in England has declined continues during week ending 16 April 2021

Figure 2 & 3. England Percentile of people testing positive



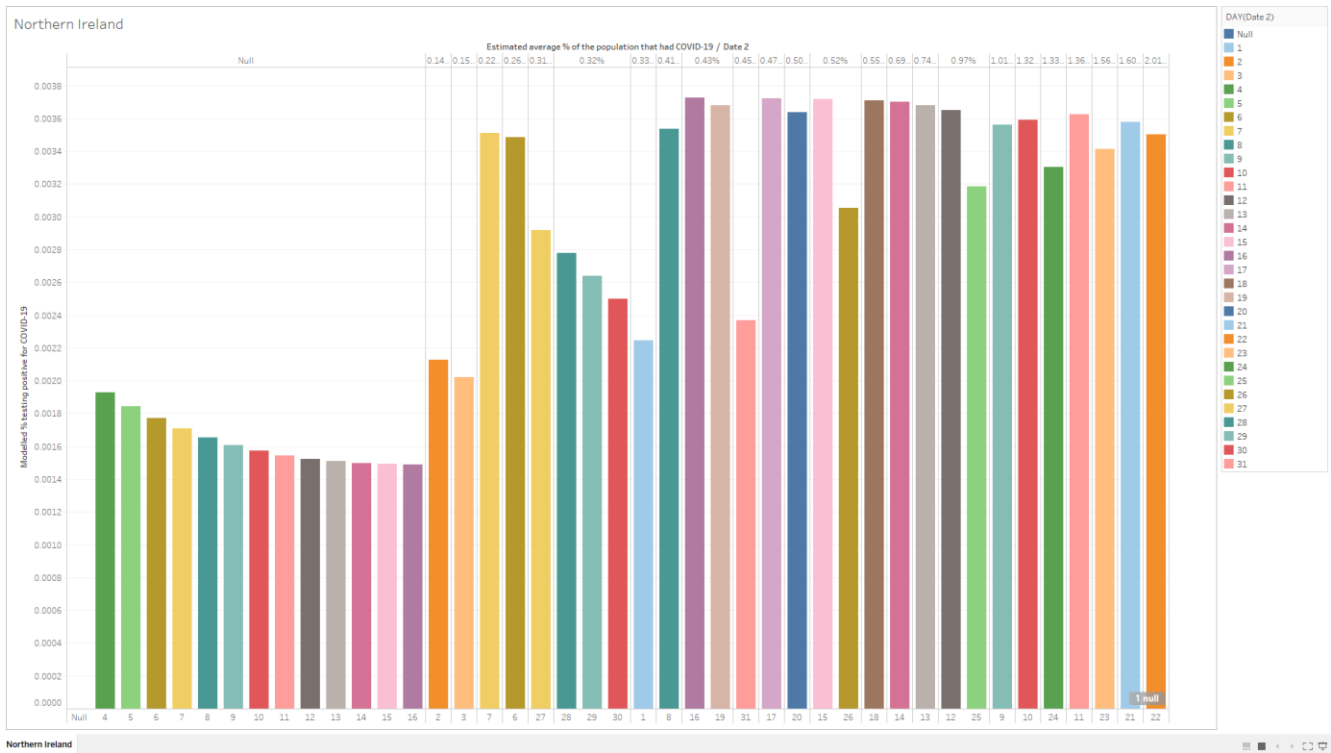
In Wales, the estimate of 3,600 people had coronavirus over the same time. This equates to 0.12% of wales populations. The cases are level during week ending 16 April 2021.

Figure 4 & 5. Wales Percentile of people testing positive



The estimate of 2,800 people in Northern Ireland had coronavirus. The percentages are declined in the two weeks up to 16 April 2021

Figure 6. Northern Ireland Percentile of people testing positive



In Scotland, the 9,300 people had covid-19 and that's equates to 0.18% of the population. So, that's decrease in the week ending 16 April 2021.

Figure 7. Scotland Percentile of people testing positive

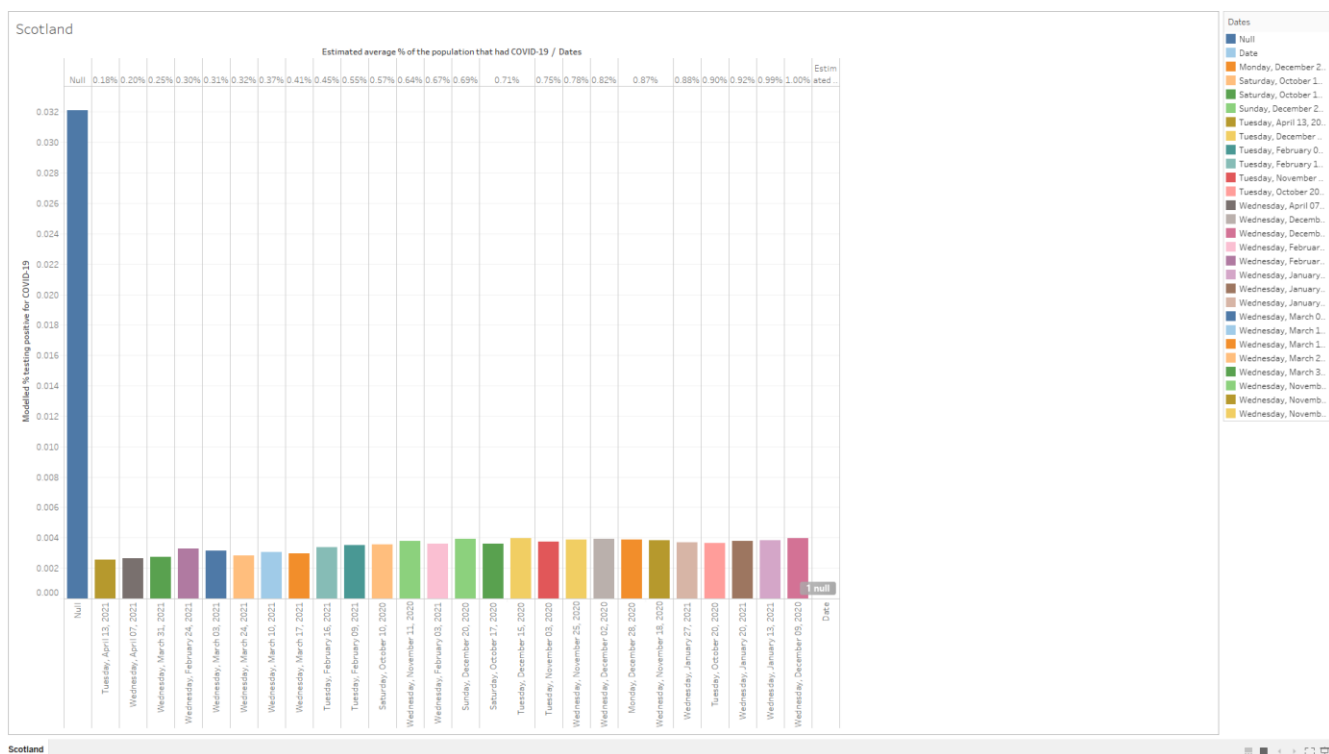
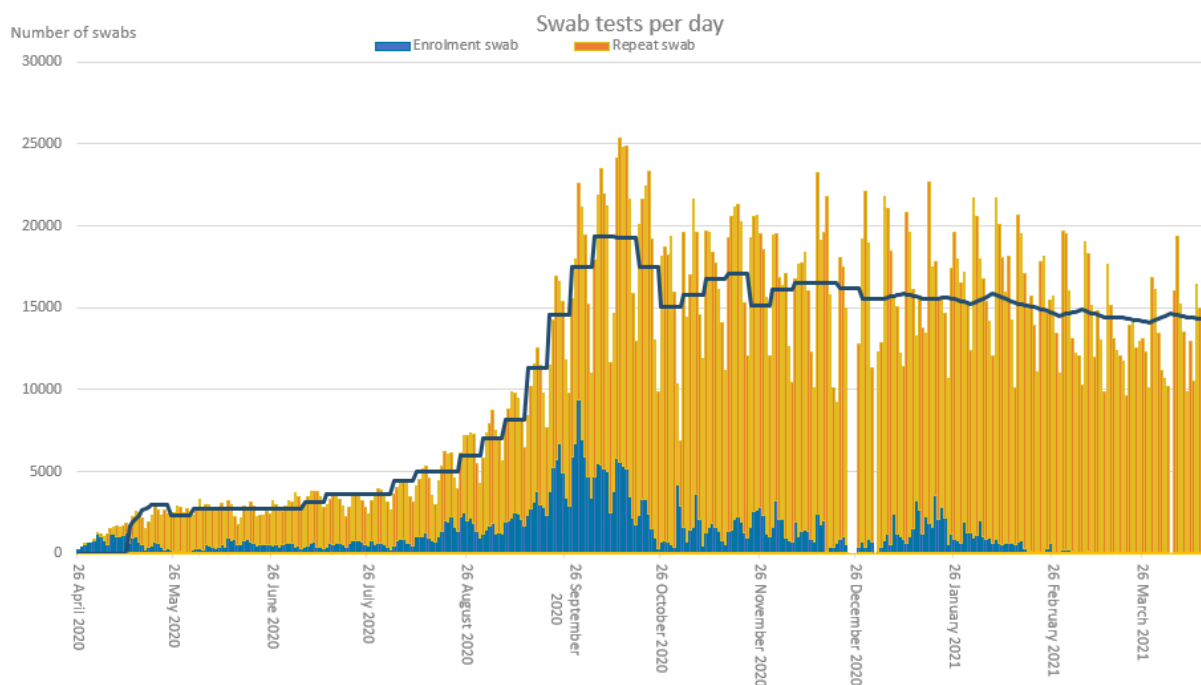


Figure 8. COVID-19 Infection survey Swab Test Per Day



As per this figure, after 26 April 2020 there will be swab test started and near to the 26th August it's noted up till 20,000. After that it's continue fluctuate near to the 15,000. This swab tests cover the whole United Kingdom.

5. Visualization

For amount spending shown in Figure 1 to 7, a graph and visuals are found appropriate because of its efficient in reflecting quantitative information and particularly in depicting trends and complex dataset. Additionally, it's a graphical display is easy to understand and do not required additional technical skill for doing these activities.

The last figure represents the COVID-19 infection survey tests per day and this figure help in visually ranking the differences between Enrolment and repeat tests well suitable for encoding in this scenario compared to others. It's displayed each day number individually, which display and understand the situation about the infection on day-to-day basis.

The data are often used for estimating the quantity of current positive cases within the community, including cases where people don't report having any symptoms identifying differences in numbers of positive cases between different regions estimating the number of latest cases and alter over time in positive cases

6. Conclusion

The study gives us the estimate the total number of people with covid-19 symptom and the daily number of fresh cases based on the data and swab tests collected by Department of Health and Social Care. This is a predictive power of symptom who are infected with covid-19 but no symptoms displayed.

Figure 7. How our coronavirus antibody testing study works

Real-time Assessment of Community Transmission

REACT 2 Studies

The REACT 2 programme is a series of studies assessing a number of antibody tests to see how accurate they are and how easily people can use them at home. This work is helping scientists find the best home test and estimate how far the virus has spread.

1

Evaluating various finger-prick antibody tests in the lab and clinic.



300 NHS staff who previously had COVID-19 and 500 pre-pandemic samples from key workers

Why?

To understand and compare how different tests perform in the clinic and the lab



2

Piloting the acceptability and usability of home-based antibody testing.



300 public volunteers

Why?

To identify issues and concerns with carrying out finger-prick antibody tests at home in order to improve the design and experience of the larger-scale studies.



3

Further exploring the usability of home-based antibody testing.



14,000 randomly selected members of the public

Why?

To explore whether people can perform the finger prick test at home and read the results correctly, as well as identifying any issues in delivering the antibody tests to a large group of people.



4

Comparing the accuracy of different types of tests for past and current infection



5,500 key workers

Why?

To look for evidence of infection in key workers and assess how finger-prick tests perform compared to lab-based tests. Also, to explore the potential role of using different types of samples (e.g. dried blood spots, saliva)



5

A large-scale national study using home antibody tests



Up to 200,000 randomly selected members of the public

Why?

To estimate how many people have already had COVID-19 in England and better understand how the virus is spreading through the population.



Research & study implementation by

Imperial College
London

Ipsos MORI



NIHR | National Institute
for Health Research

Funded by



Department
of Health &
Social Care

In support of



7. References

[1] World Health Organization

~<https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/novel-coronavirus-2019-ncov>

~ <https://www.who.int/news/item/27-04-2020-who-timeline---covid-19>

[2] Office for National Statistics

~<https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronavirusandclinicallyextremelyvulnerablepeopleinengland/22marchto31march2021#toc>

[3] Figure 7.

<https://www.imperial.ac.uk/medicine/research-and-impact/groups/react-study/the-react-2-programme/>

[4] Figure 1.

<https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200423-sitrep-94-covid-19.pdf>