Data Bootcamp Final Project Presentation

HAWA JAMA

APRIL 2022

Department for Education

BLACK DISRUPTOR ACCELERATOR

HAWA JAMA| HEALTHCARE DATA SET REPORT

This dataset was focused on immunisation factors, mortality factors, economic factors, social factors and other health related factors as well. Since the observations in this dataset were based on different countries, it helped with determining the predicting factors which contributed to the lower value of life expectancy. This will help in suggesting a country which area should be given importance in order to efficiently improve the life expectancy of its population. In this project I have focused on analysing the relationship between outlier factors and a countries life expectancy, whether it be related to alcohol consumption, the rate of polio and measles and so on, this had then helped with seeing trends and patterns across multiple countries when it came to overall health.

ABOUT ME

- I am currently a 3rd year International Relations student at Coventry University specialising in cyber conflicts and warfare within my Dissertation.
- In May 2022 I will have completed the Niyo Enterprise Data Analytics & Project Management Bootcamp.
- I have previously worked as a Pension Regulator advisor for HMRC and became interested in tech due to some of the skills I learnt on the job within administration and finance related issues.

WHY DID YOU CHOOSE TO LEARN DATA ANALYTICS, AND WHAT ROLE WOULD YOU LIKE TO WORK IN FOLLOWING WHAT YOU'VE LEARNED?

Why?

I chose to learn Data Analytics due to my enthusiasm when it comes to problem solving and challenging myself. I was also intrigued by a role which implemented real life situations in an everyday job role, which pushes me to think outside of the box. Data analytics and cybersecurity also overlap which is another reason as to why I chose this bootcamp.

What?

I would want to work in a role such as a Data Analyst or Business Intelligence Analyst. I am also interested in learning more about Machine learning and working on my coding as well as continuing on to learn more programming languages.

Throughout this bootcamp I have learnt multiple different things. From using formulas in excel to learning SQL and python, this has helped me gain more of an insight to the skills needed in jobs regarding Data Analysis and visualisation.

ABOUT YOUR PROJECT

The Objective of my project was to employ all the skills I've gained in the duration of this bootcamp to then explore the trends and patterns in my selected dataset, which is a Healthcare dataset, focused on the Factors that influence Life expectancy. I will be showing my findings and how they were implemented using different tools and platforms.

HOW DID YOU APPLY WHAT YOU HAVE LEARNT IN THE BOOTCAMP?

- I firstly used my knowledge of Excel to clean the data and pick out information regarding the life expectancy and health issues of different countries. I used formulas such as VLOOKUP, COUNTIF, MAX&MIN, MEDIAN and SUM.

- I then used SQL to write queries to further analyse my data. Using aggregators such as SUM, AVG, COUNTS, GROUP BY, and ORDER BY. I used WHERE to filter through my data and the CASE WHEN technique to add additional columns.

- Lastly, the bootcamp helped us utilise the skillsets needed for both Tableau and Power bi, and for my last project I used tableau so this time around I chose Power bi to present my data visualisation. I completed this through the use of diverse charts in comparing correlations between the data as well as the trends.

EXCEL DATA ANALYSIS

Question 1 : 1. Use Vlookup to find the country with the highest Life Expectancy after 2001

Italy	Which country having maximum life expectancy

Question 2 Usage of Countif Function

16	How many times was alergias name is in the list
2801	how many countries have a life expectancy over 50
368	how many countries have an ifant death over 50

Question 3 Usage of Max and Min Function

Max	99	Maximum number of cases in a country having hepatities B
Min	1	Minimum number of cases in a country having hepatities B

Pick 3 or more countries and show the median Life expectancy

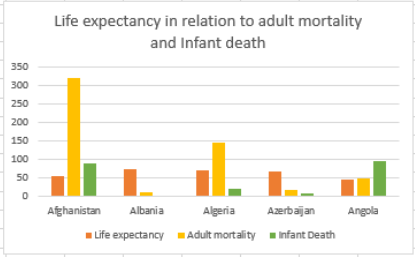
Answer: 71

Countries Picked: Australia,France, Albania, Singapore

I chose 2 developed and 2 developing countries to show that the life expectancy in this case is not majorly linked to the scoial circumstance but more towards health issues.

All 4 countries have a low number of infant deaths aswell, thus increasing the life expectancy.

Australia and France have a higher alcohol rate then Singapore and Albania and this can be directly linked to the difference in popualtion of all 4 countries as well as the tourist difference.



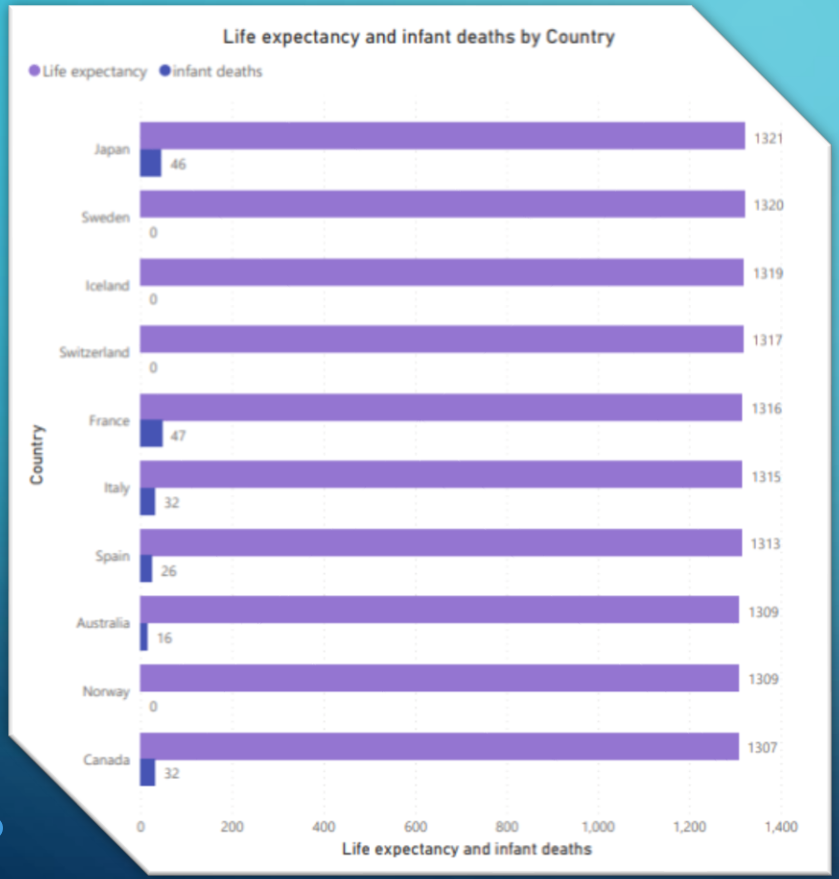
This chart indicates that afghanistan has the highest adult mortality, this could be due to the political cli climate of the country.

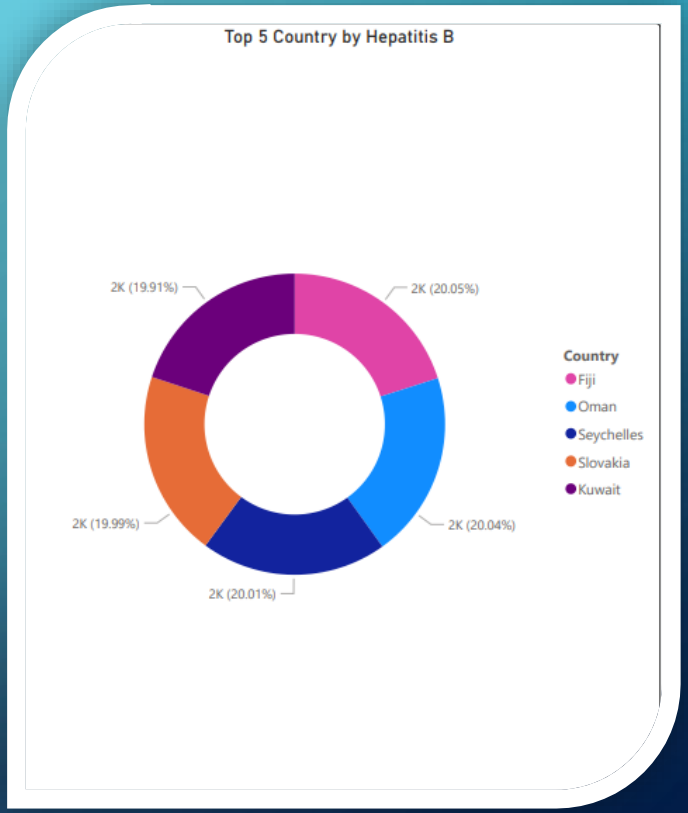
The countries with the higher adult mortality rate also have a high number of infant death rates

Country	Life expectancy	Adult mortality	Infant Death
Afghanistan	54.8	321	88
Albania	72.6	11	1
Algeria	71.3	145	21
Azerbaijan	66.6	16	9
Angola	45.3	48	97

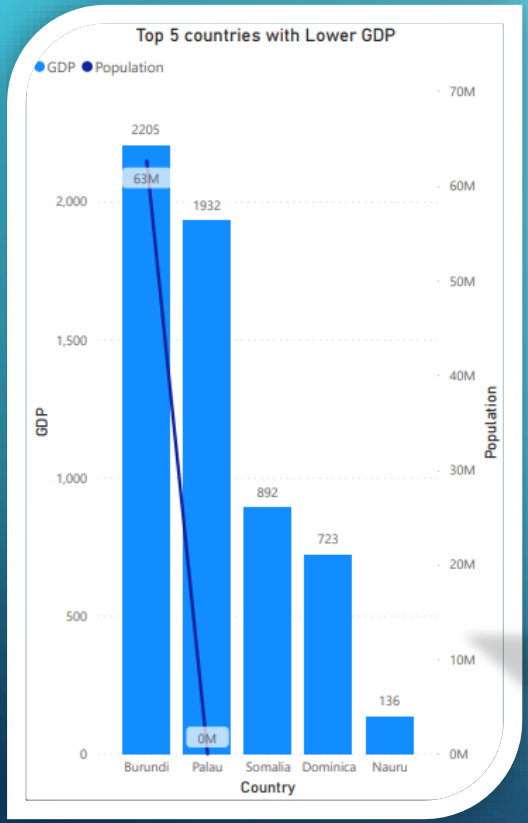
Here I came up with own questions regarding the dataset and used different formulas taught in class to get an analysis

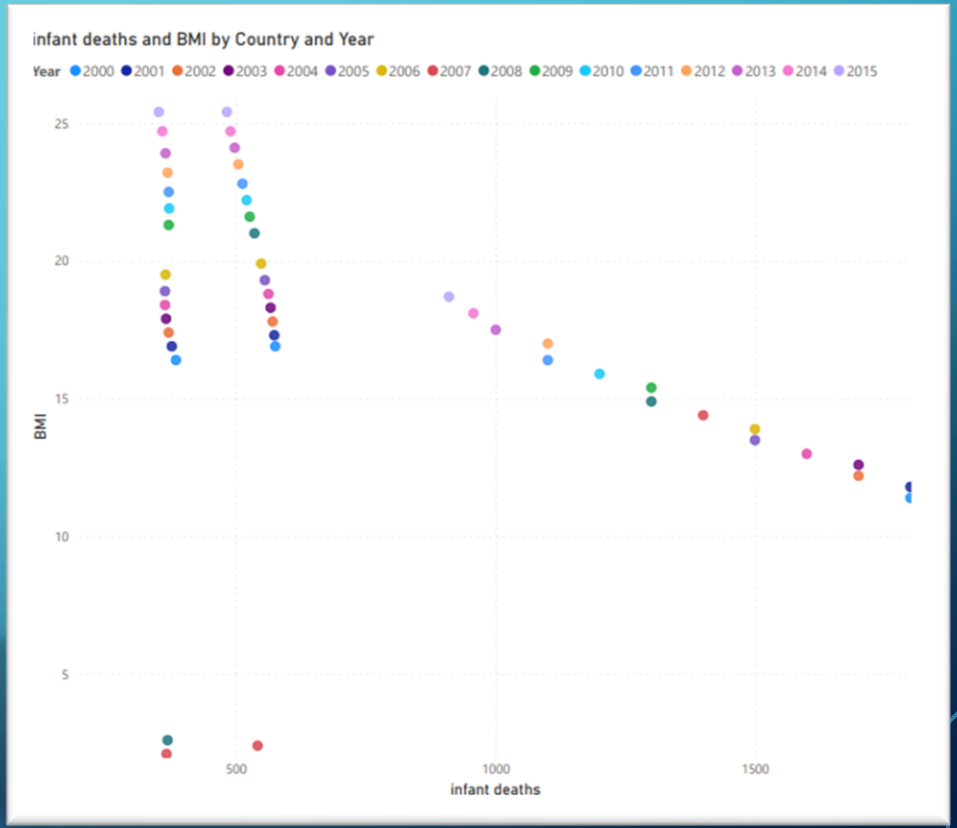
FINDINGS: POWER BI DATA VISUALISATION



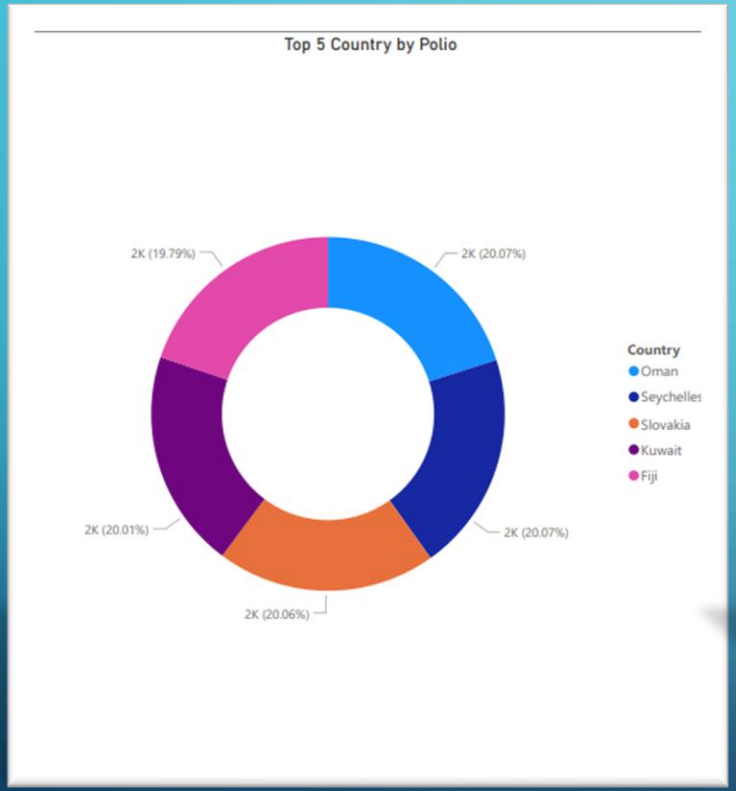


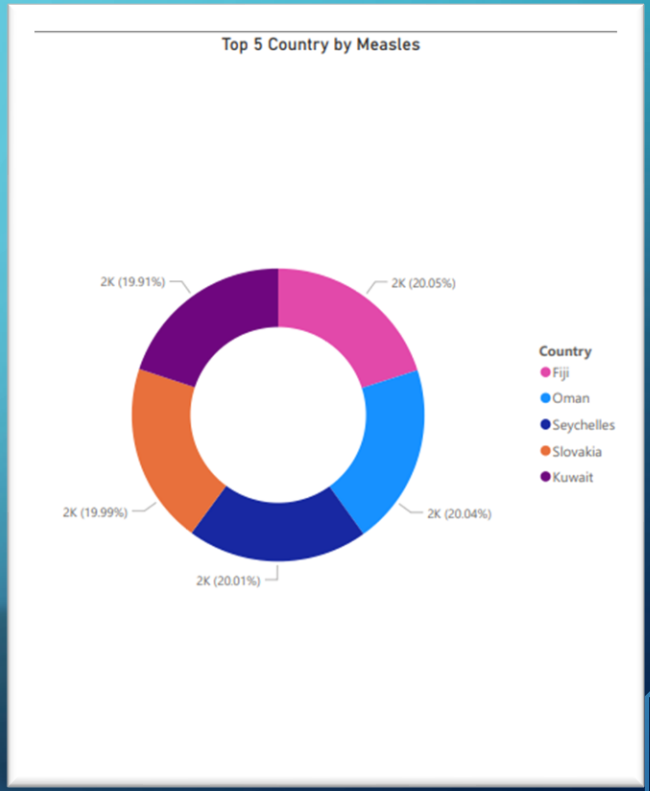
FINDINGS: POWER BI DATA VISUALISATION



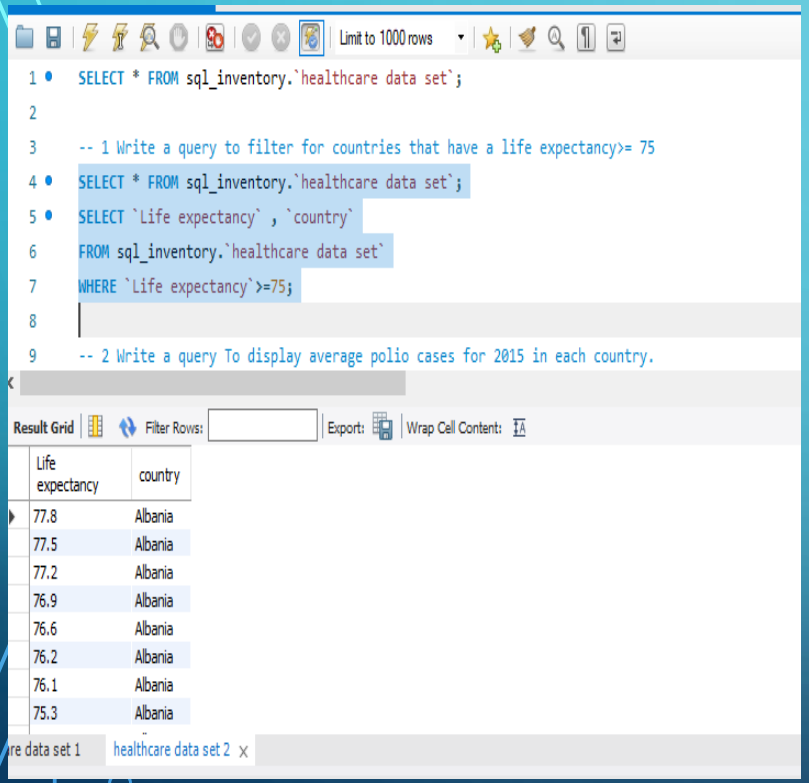


FINDINGS: POWER BI DATA VISUALISATION





SQL DATA ANALYSIS

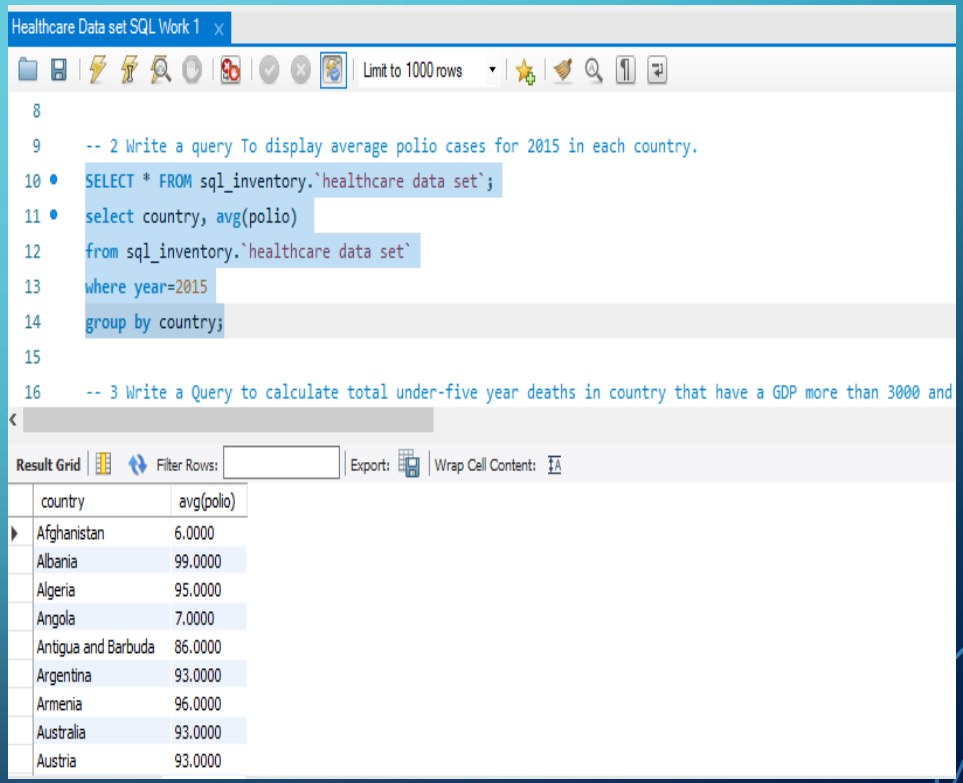


```
1  SELECT * FROM sql_inventory.`healthcare data set`;
2
3  -- 1 Write a query to filter for countries that have a life expectancy>= 75
4  SELECT * FROM sql_inventory.`healthcare data set`;
5  SELECT `Life expectancy` , `country`
6  FROM sql_inventory.`healthcare data set`
7  WHERE `Life expectancy`>=75;
8
9  -- 2 Write a query To display average polio cases for 2015 in each country.
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Life expectancy	country
77.8	Albania
77.5	Albania
77.2	Albania
76.9	Albania
76.6	Albania
76.2	Albania
76.1	Albania
75.3	Albania

re data set 1 | healthcare data set 2



Healthcare Data set SQL Work 1

```
8
9  -- 2 Write a query To display average polio cases for 2015 in each country.
10 SELECT * FROM sql_inventory.`healthcare data set`;
11 select country, avg(polio)
12 from sql_inventory.`healthcare data set`
13 where year=2015
14 group by country;
15
16 -- 3 Write a Query to calculate total under-five year deaths in country that have a GDP more than 3000 and
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

country	avg(polio)
Afghanistan	6.0000
Albania	99.0000
Algeria	95.0000
Angola	7.0000
Antigua and Barbuda	86.0000
Argentina	93.0000
Armenia	96.0000
Australia	93.0000
Austria	93.0000

SQL DATA ANALYSIS

```
16  -- 3 Write a Query to calculate total under-five year deaths in country that have a GDP more than 3000 and show resul
17  select * from sql_inventory.`healthcare data set`;
18  select sum(`under-five deaths`) , `country`,`year`
19  from sql_inventory.`healthcare data set`
20  group by Country
21  having sum(`GDP`)>3000
22  order by sum(`under-five deaths`) asc;
23
24  -- 4 Write a Query to show  each country that has a total infant death greater than 5000 and also show number of meas
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

sum(`under-five deaths`)	country	year
22	Australia	2015
27	Panama	2015
30	Djibouti	2015
30	Spain	2015
31	Mongolia	2015
32	Canada	2015
32	Comoros	2015
35	Chile	2015

healthcare data set 5 | Result 6

Output

Action Output

#	Time	Action	Message
5	17:46:48	select * from sql_inventory.`healthcare data set` LIMIT 0, 1000	1000 row(s) returned
6	17:46:48	select sum(`under-five deaths`) , `country`,`year` from sql_inventory.`healthcare data set` grou...	158 row(s) returned

Limit to 1000 rows

```
23
24  -- 4 Write a Query to show  each country that has a total infant death greater than 5000 and also show number of meas
25  select * from sql_inventory.`healthcare data set`;
26  select `country`,sum(`infant deaths`) , count(`Measles`)
27  from sql_inventory.`healthcare data set`
28  group by `country`
29  having sum(`infant deaths`)>5000
30  order by sum(`infant deaths`) asc ;
31
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

country	sum(`infant deaths`)	count(`Measles`)
Pakistan	5887	16
Nigeria	8571	16
India	21867	16

healthcare data set 7 | Result 8

Output

Action Output

#	Time	Action	Message
7	17:47:47	select * from sql_inventory.`healthcare data set` LIMIT 0, 1000	1000 row(s) returned
8	17:47:47	select `country`,sum(`infant deaths`) , count(`Measles`) from sql_inventory.`healthcare data se...	3 row(s) returned

TOP 3 THINGS YOU HAVE LEARNT ON THE BOOTCAMP

Educationally: The need for coding in data, and the importance of understanding why patience and resilience is required to complete an intensive bootcamp. I also learnt how to utilise my personal time effectively and the daily practical's in lessons aided my learning further.

Personally: It helped me become more confident in my technical skills especially in relation to coding, and helped me to challenge myself and reach for higher expectations and goals.

Career Wise: I am now more sure of the route I want to take in data analytics and working more on other programming languages to increase personal development.

Department for Education