Data Bootcamp Final Project Presentation

Life Expectancy WHO(World Health Organisation) Dataset

Kamilah McCarthy

https://github.com/kamkodes/data_analytics_project

About Me

Personal & Professional Development

Future Proofing

Opportunity

Community

Networking

Actively seeking Project Management opportunities

NIYO Bootcamps

Department for Education

SKILLS FOR LIFE

My Project - Objectives

- Healthcare - Life Expectancy WHO (World Health Organisation) Dataset

- Using the data analysis skill I have developed over the past 16 weeks to understand what factors impact life expectancy

Excel
Power BI
SQL

My Project - Excel Findings

=MEDIAN(E1576,E1574)

country	year	Original adult_mortal	Cleaned adult_mortal
Malawi	2015	365	365
Malawi	2014	377	377
Malawi	2013	394	394
Malawi	2012	42	418
Malawi	2011	441	441
Malawi	2010	462	462
Malawi	2009	491	491
Malawi	2008	525	525
Malawi	2007	559	559
Malawi	2006	587	587
Malawi	2005	66	601
Malawi	2004	615	615
Malawi	2003	613	613
Malawi	2002	67	606
Malawi	2001	599	599
Malawi	2000	588	588

I was interested in using the Adult Mortality metric for Malawi, but noticed that there were obvious errors as the overall trend showed that there was a steady decline.

I identified the anomalies and decided to replace them with the median using the values either side.

If this something like this had occurred in the workplace I would seek to understand how and why the anomalies occurred and try to rectify it with correct data before making assumptions based on trends.

My Project - Excel Findings

Developed vs Developing Countries

- I wanted to draw a distinction between Developed and Developing Countries
- By visualising my data I also noticed that there were some anomalies. E.g. Sweden was labelled as Developing Country
- If I was in the work place there would be clear definitions as to what is considered Developed vs Developing. So I decided to do just that using my own definitions

- However, I soon came to realise that the GDP for some countries was potentially wrong as it stated that

Some economists consider **$12,000 to $15,000 per capita GDP** to be sufficient for developed status while others do not consider a country developed unless its per capita GDP is above $25,000 or $30,000.

https://www.investopedia.com › ... › Macroeconomics

What Is a Developed Economy? Definition, How It Works, HDI ...

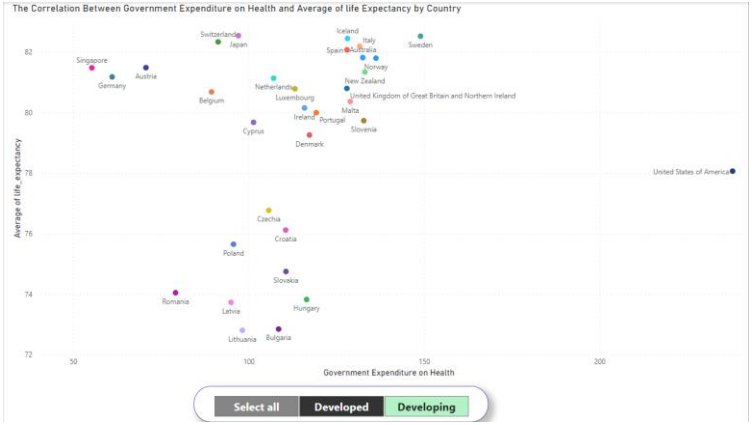
```
=IF(R2<20001,"Developing","Developed")
```

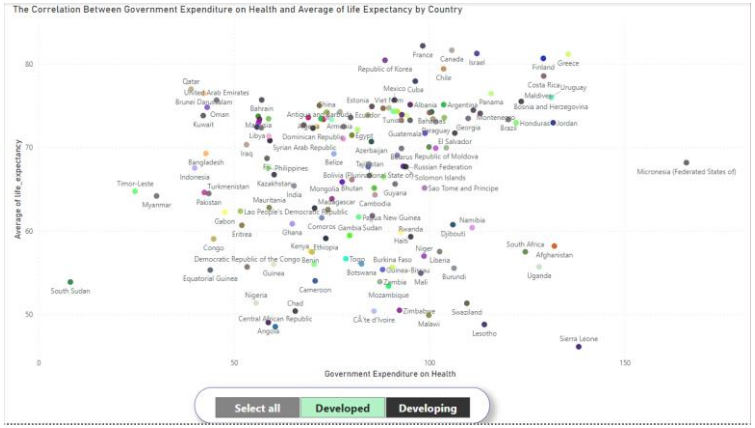
NIYO Bootcamps

Department for Education

SKILLS FOR LIFE

My Project - Dashboard





Filters

Search

Filters on this visual

status
is (All)

Add data fields here

Filters on this page

Add data fields here

Filters on all pages

Add data fields here

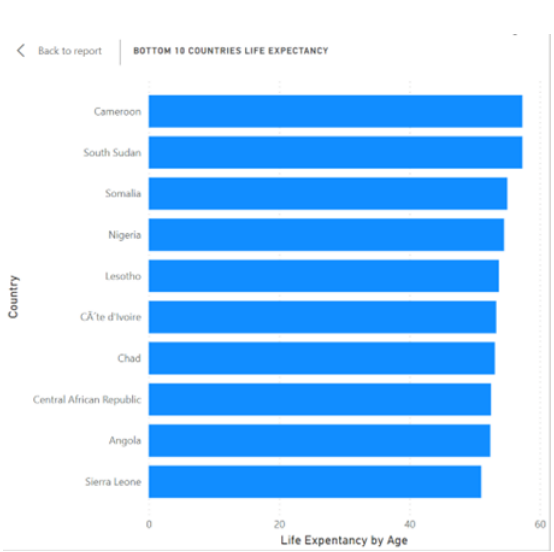
NIYO Bootcamps

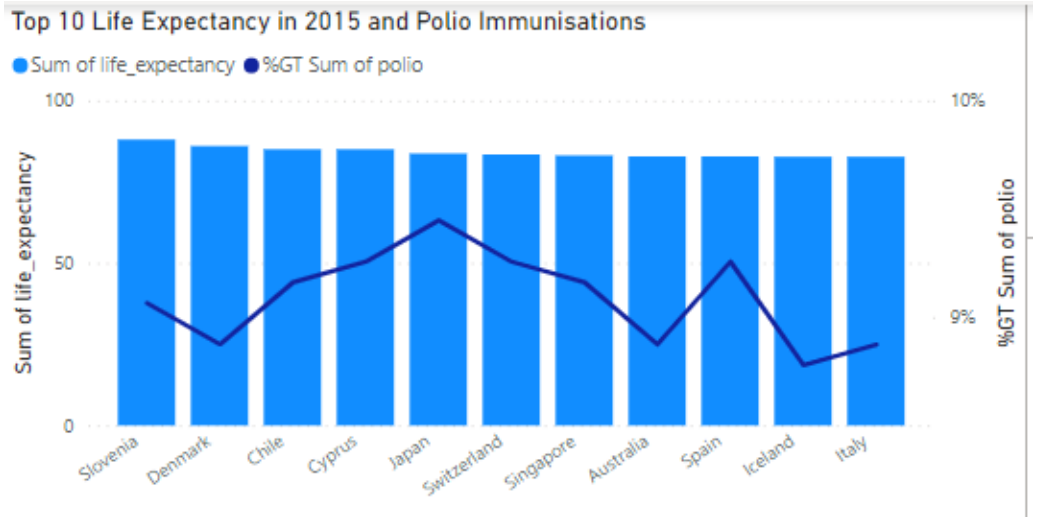
Department for Education

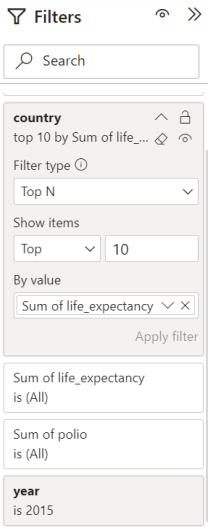
SKILLS FOR LIFE

My Project - Dashboard

Using Filters to categorise data into Top/Bottom 10 and overlaying with additional data to identify any correlation







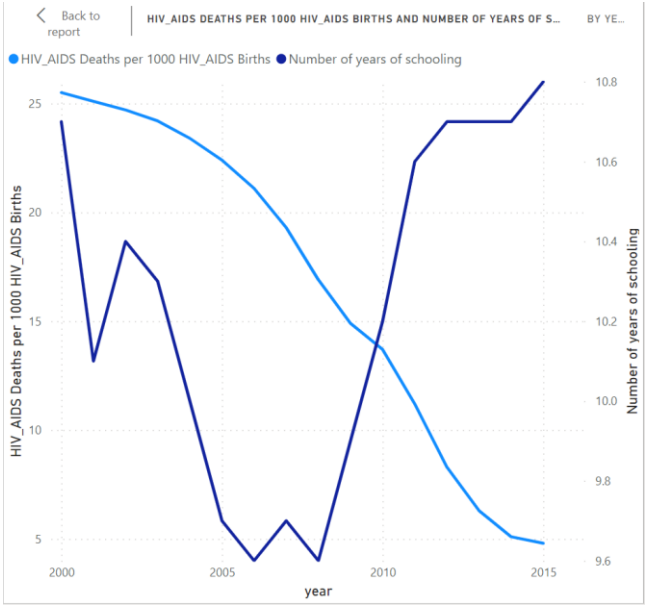
NIYO Bootcamps

Department for Education

SKILLS FOR LIFE

My Project - Dashboard

Leading cause of Death in Malawi is HIV. The obvious thing to do would be to draw a distinction between adult mortality and HIV deaths, but I wanted to see if there was a relationship Education and HIV.

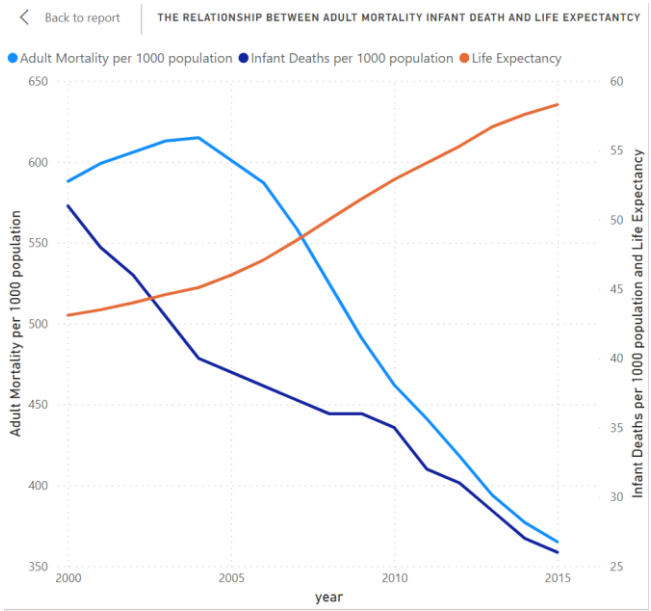


NIYO Bootcamps

Department for Education

SKILLS FOR LIFE

My Project - Dashboard



How does infant deaths and adult mortality impact life expectancy?

The visualisation will show that as both Infant Deaths and Adult Mortality decrease, Life Expectancy in Malawi increases

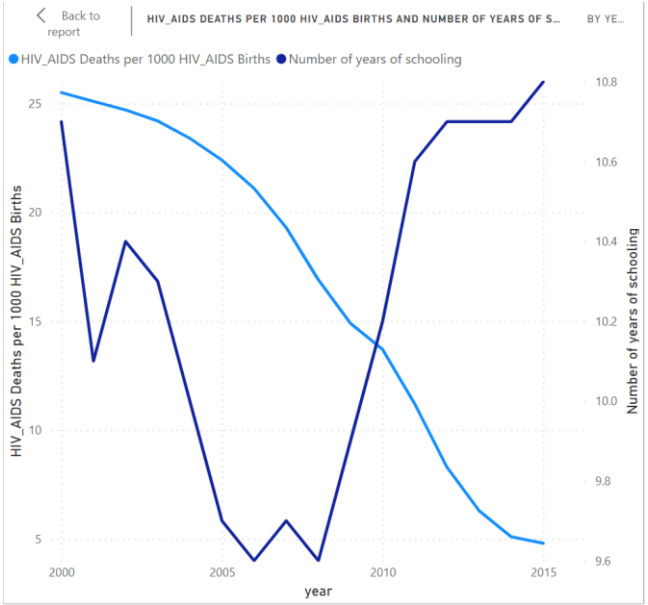
NIYO Bootcamps

Department for Education

SKILLS FOR LIFE

My Project - Dashboard

Leading cause of Death in Malawi is HIV. The obvious thing to do would be to draw a distinction between adult mortality and HIV deaths, but I wanted to see if there was a relationship Education and HIV.



NIYO Bootcamps

Department for Education

SKILLS FOR LIFE

My Project - SQL Analysis

```
-- I want to be able to see all table of all countires that have a stats of 'Developing' in 2014
SELECT * FROM life_expectancy.healthdata
WHERE year = 2014 and status = 'Developing';
```

country	year	status	life_expectancy	adult_mortality	infant_deaths	alcohol	percentage_expenditure	hepatitis_B	me
Afghanistan	2014	Developing	59.9	271	64	0.01	73.52358168	62	492
Albania	2014	Developing	77.5	8	0	4.51	428.7490668	98	0
Algeria	2014	Developing	75.4	11	21	0.01	54.2373183	95	0
Angola	2014	Developing	51.7	348	67	8.33	23.965612	64	116
Argentina	2014	Developing	76.2	118	8	7.93	847.3717463	94	1
Armenia	2014	Developing	74.6	12	1	3.91	295.6087143	93	13
Azerbaijan	2014	Developing	72.5	119	5	0.01	306.1824313	94	0
Bangladesh	2014	Developing	71.4	132	98	0.01	10.44640334	97	289

My Project - SQL Analysis

```sql
-- Find the Average Life Expectancy by country and order from largest to smallest
select country, AVG(life_expectancy) AS ' Average Life Expectancy'
FROM life_expectancy.healthdata
WHERE status = 'Developed'
GROUP BY country
ORDER BY 2 desc;
```

	country	Average Life Expectancy
▶	Australia	81.90714285714286
	Austria	81.47999999999998
	Belgium	80.65333333333332

My Project - SQL Analysis

```sql
-- Find the Average Life Expectancy by  For Developing Countries and order from smallest to largest
select country, AVG(life_expectancy) AS ' Average Life Expectancy'
FROM life_expectancy.healthdata
WHERE status = 'Developing'
GROUP BY country
ORDER BY 2 asc;
```

	country	Average Life Expectancy
▶	Angola	50.675000000000004
	Benin	57.70769230769231
	Afghanistan	58.193749999999994
	Botswana	64.65
	Bhutan	65.92
	Belize	69.15333333333335
	Belarus	69.74666666666668
	Bangladesh	69.96666666666665

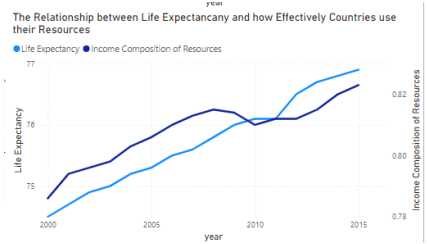
My Project - SQL Analysis

```sql
-- Catergorise Coutries by first letter
SELECT country,
CASE
WHEN country LIKE 'a%' THEN 'A'
WHEN country LIKE '%b%' THEN 'B'
ELSE 'Other Countries'
END AS 'Country 1st Letter'
FROM life_expectancy.healthdata;
```

country	Country 1st Letter
Azerbaijan	A
Azerbaijan	A
Azerbaijan	A
Azerbaijan	A
Bangladesh	B
Bangladesh	B
Bangladesh	B
Bangladesh	B

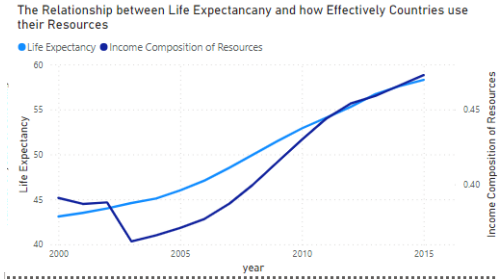
My Project - Recommendations/Insights

Life Expectancy Increases when countries utilise their resources more productively. Further research should be focused in this area

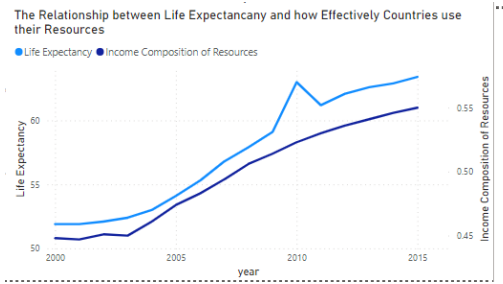


Bahrain

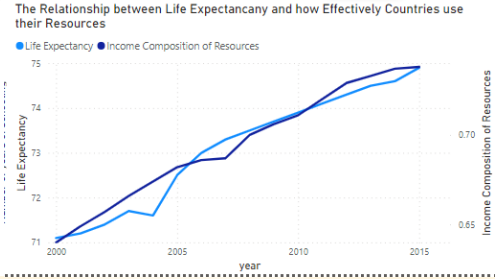
Kenya



Malawi



Thailand

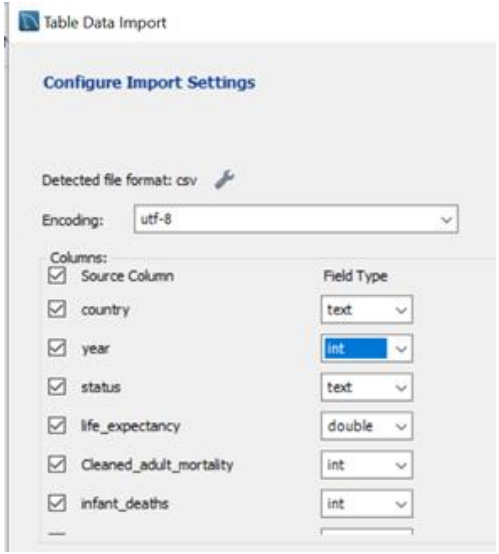


NIYO Bootcamps

Department for Education

SKILLS FOR LIFE

Challenges



NIYO Bootcamps

Department for Education

SKILLS FOR LIFE

Conclusion and Key Learnings

Data must be clean - You cannot generate accurate insights with dirty data

Lean on your network - utilise your support system and actively share knowledge

Data Visualisation!