DATA VISUALIZATION PROJECT

OPTION 2- TUBERCULOSIS BURDEN BY COUNTRY

PROJECT GOALS

- This project involved turning data into easily consumable visual insights, using Tableau.
- It required creating impactful dashboards that would help stakeholders make decisions, based on a business questions.
- It involved communicating insights with the correct visualizations.

PROCESS

- Download and connect with data in Tableau.
- Analyze the dataset and identify the different data types present.
- Clean the dataset removing null values.
- Build different visualizations to learn more about the dataset.
- Define questions that would be answered using the features in the dataset. Identify trends.

DIFFERENT DATA TYPES IN DATASET

There were only three types of data types within the dataset. They include

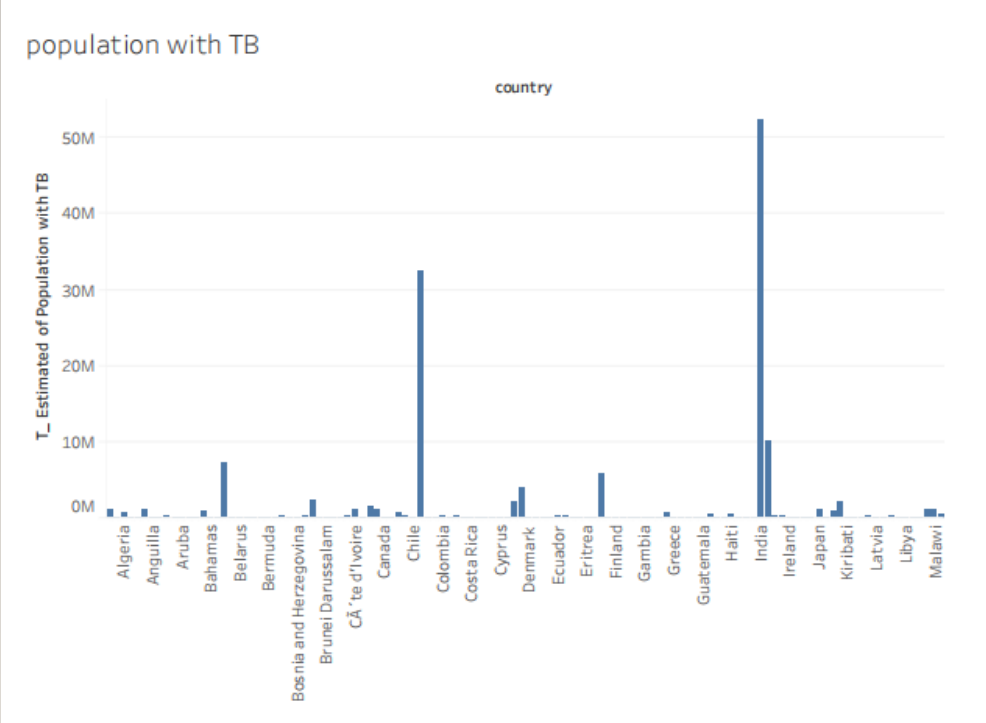
- Number
- String
- String – Geographic role

To understand the dataset better several visualizations were done. They include

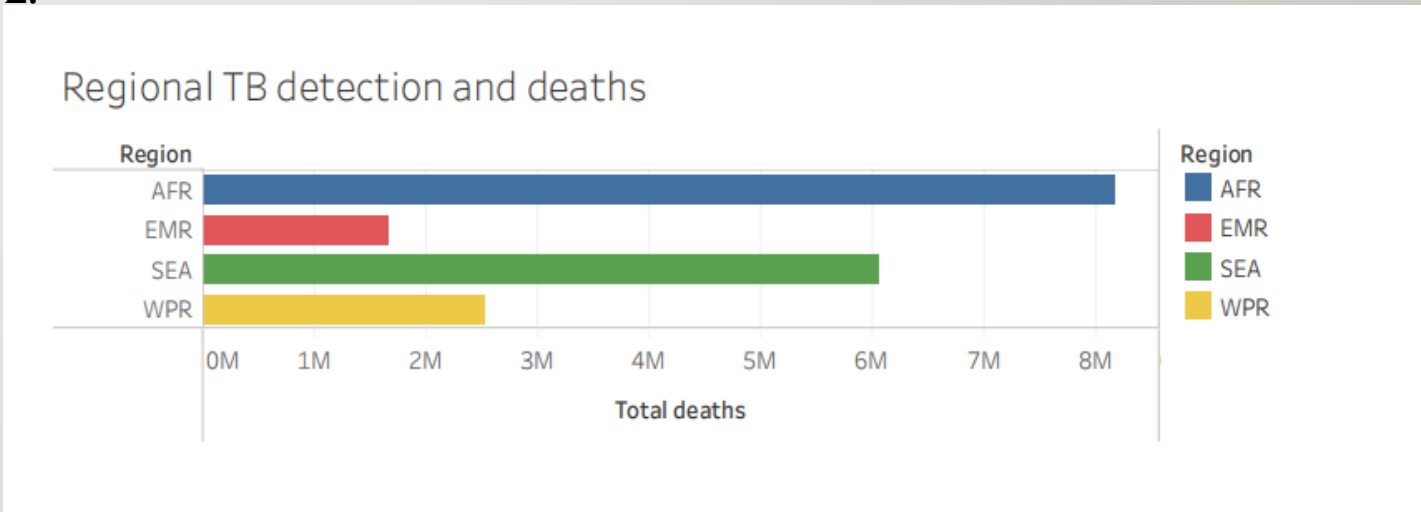
- Horizontal Bars
- Pie Charts
- Lines
- Heat Maps
- Geographical Maps
- Packed Bubbles
- Scattered plots

VISUALIZATIONS BUILT TO UNDERSTAND DATASET

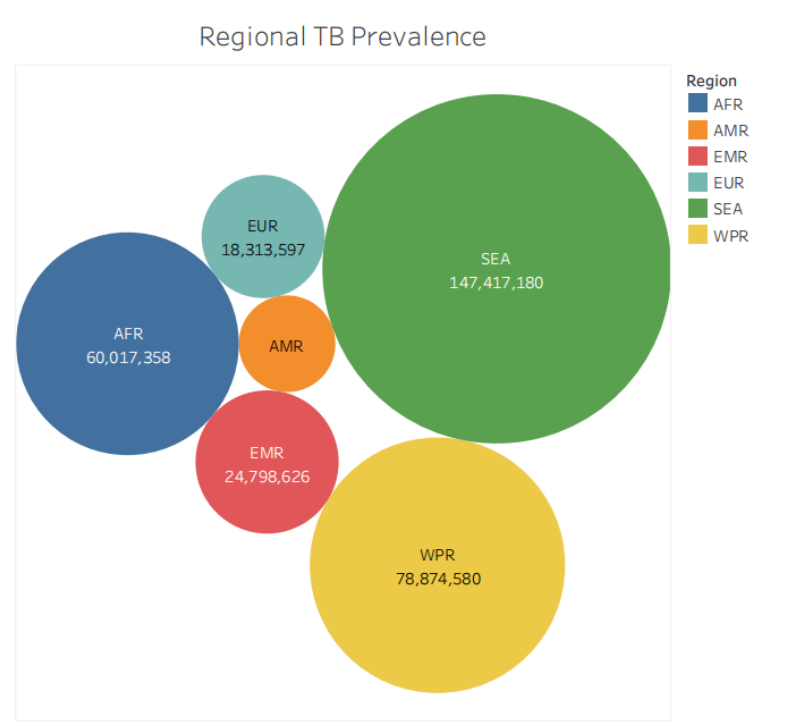
1.

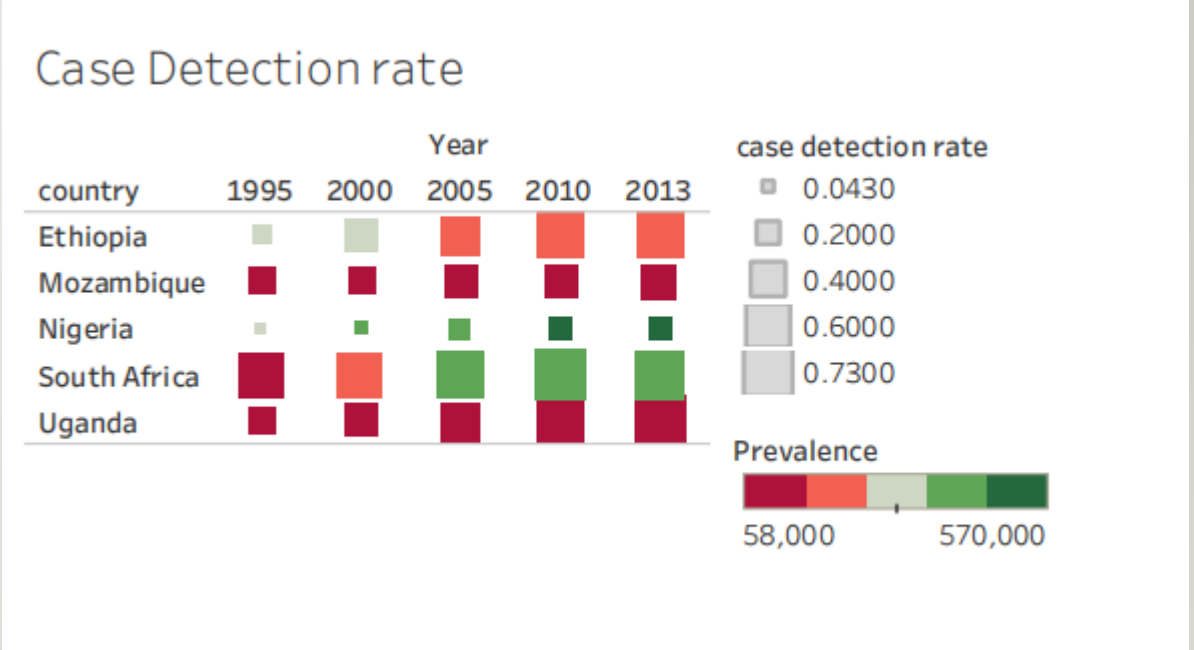


2.



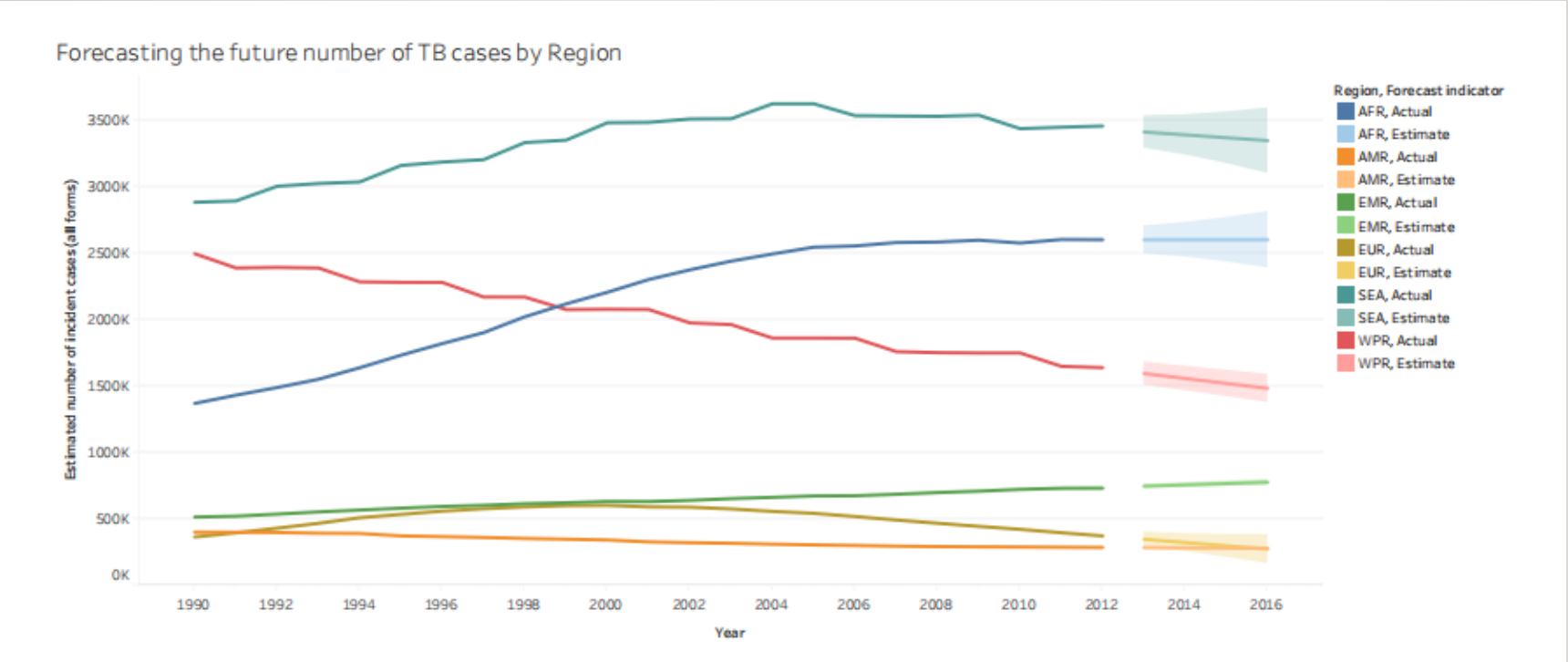
VISUALIZATIONS BUILT TO UNDERSTAND DATASET





VISUALIZATIONS BUILT TO UNDERSTAND DATASET

VISUALIZATIONS BUILT TO UNDERSTAND DATASET



BUSINESS QUERIES/QUESTIONS

There were three (4) main data queries/questions

1. Which country had the most TB prevalence and deaths worldwide
2. Which Region/country had the most prevalence and deaths associated with TB
3. Which Countries in Africa are most affected?
4. What does the future hold for the world concerning TB?

RESULTS

- To answer the first query a map of the world highlighting the different countries, with their associated TB prevalence and deaths was made. It identified India as having the most TB prevalence with over 97 million estimated cases and about an estimated 10 million deaths during the time frame. China also had a very high prevalence of TB.
- Scatter plots were generated to understand the relationship between TB deaths and the country's population. It was determined that India had the highest population and also the highest deaths but South Africa had the highest TB-HIV related deaths worldwide.

QUERY 2

- To know the Region that had the most TB Prevalence and deaths, a packed bubbles plot alongside a pie chart was made. These plots showed that South East Asia (SEA), had over 147 million estimated prevalent cases and had a death profile of just over an estimated 17 million. However, Africa (AFR) which had a lower TB prevalence of about an estimated 60 million had an almost equivalent death profile of about 16.7 million cases.

QUERY 3

- To understand the countries in Africa that are most affected with TB, a geographic map, bars and a heat map was used. These visualizations identified Nigeria as having the highest TB prevalence and deaths with South Africa coming a close second. Nigeria had an estimated 10 million people with TB and an estimated 4million deaths while South Africa had an estimated 2 million deaths and an estimated 7 million cases. A relationship was observed between TB deaths and HIV as higher deaths were associated with cases that had HIV positive status (e.g South Africa).

QUERY 4

- With the aid of line charts and forecasting, it is estimated that less than 5million deaths would be recorded in the span of 3 years. Whereas TB prevalence worldwide is estimated to drop to less than 10 million. Regionally, each region's incident cases would continue on the same trajectory as presently.

RECOMMENDATIONS

- An investigation into the reasons why the death profile in Africa is as high as that of South East Asia even though their prevalence of TB is less than half that of South East Asia. This would help to further reduce the death profile worldwide.
- Incident cases are far lower than the death profile, indicating that there is a gap in cases being identified. Which can be an indication of treatment not getting to the sick. The gap should thus be investigated.

CHALLENGES AND FUTURE GOALS

- Identifying the right visualization to use to adequately extract the results appropriately.
- Deciding what visualization passed the necessary information in the simplest manner and can be put into the dashboard.
- Future Goals
- Arrange the dashboards better.
- Look for better visual combinations to convey the results appropriately.