Tableau - Lab

**Part I: Tuberculosis**

Worksheet and dashboard creation

**Understand the data**

* Load TB\_Burden\_Country.csv into a new workbook
* Create a new worksheet
* Note all of the measures in this data set about tuberculosis and HIV
* Incidence - number of new cases per year
* Prevalence - number of people living with disease during a particular year
* Mortality - normalized by population
* Number of deaths - not normalized by population

**Explore the data set, what are the Dimensions and Measures?**

* This data set divides the countries up into six regions
* There is 24 years of data, from 1990 - 2013
* There are upper and lower bounds for each measured data
* Data for “all forms” of TB
  + Incidence/100,000
  + Prevalence/100,000
  + Incidence
  + Prevalence
* Data broken into “with” and “without” HIV
  + Number of deaths
  + Mortality/100,000
* Data “with” HIV
  + Incidence
  + Incidence/100,000
  + Estimated HIV in incident TB (percent)

**CHECK IN WITH THE CLASS, ARE THERE ANY QUESTIONS ABOUT THIS DATASET?**

**Think of a starting point for how to explore data, where might you end up?**

* ~~Hide measures that are not relevant for ease of use (perhaps upper and lower bound, to simplify lab), right click on Measure and click Hide, can unhide by right-clicking in the Measures pane~~
* Start by creating a few map visualizations (click Show Me to choose between symbol and fill map). Format each visualization to your liking and in a new worksheet, qualitatively answer a few questions about the data, e.g.:
  + How are the countries organized into regions?

Africa, Americas, Europe, South east asia, WPR, EMR

* + Which regions seem to have most/least cases of TB? What if you normalize by population?

Most :SEA

Least: AMR

Normalized:

MOST: SEA

LEAST: AMR

* + In which parts of the world is it common to be HIV+ when you are diagnosed with TB? Is that different from above?
  + AFR , different than above.
* Save your workbook

**CHECK IN WITH YOUR NEIGHBORS, WHAT ARE YOUR INITIAL THOUGHTS, WHAT WAS CHALLENGING/SURPRISING?**

**Explore the data further, what do you find interesting?**

* Create enough scatter plots to explore relationships between prevalence, incidence and mortality/deaths. Save each plot in a new worksheet. Format all plots to your liking.
* Explore how being HIV+ (or not) affects those relationships. (NOTE: Add two measures to Rows Shelf and then drag name on top of other measure to plot on same axis).

Less deaths per prevalence if HIV + compared to all forms of TB

* After this, try adding region to color the scatter plots.
  + Note what you find interesting about the measure relationships
  + Note what you find interesting about HIV+ relationship to TB
  + Note what you find interesting about regions
  + Save your workbook

**DOES ANYONE HAVE A REALLY NICE GRAPH THAT YOU WANT TO SHARE WITH THE CLASS?**

**Explore the data further by region, what do you find interesting?**

* Create enough bar charts and box plots to explore the Measures data by **region**. Think about the years that you are considering; do you want to use sum or averaged Measures data in the bar charts? Create year filters as you see it. Try adding year as an additional dimension to bar charts or box plots. Format each visualization to your liking and in a new worksheet, qualitatively answer a few questions about the data, e.g.:
  + Are you surprised at any of the regional results based on what you learned from the map graphics? You could dive into those unexpected results in the next part of the lab.

More people die from HIV in AFRICA than any other kind of TB

AFRICA AND SEA have a good proportion of outliers

* + Do you see any interesting trends by Year? Do these trends seem most interesting for a particular region?
  + Save your workbook

**CHECK IN WITH YOUR NEIGHBORS, WHAT HAVE YOU LEARNED THAT IS DIFFERENT THAN WHAT THEY HAS LEARNED?**

**Create a dashboard and tell your story**

* Create a dashboard to showcase your best work so far
* (Depending on time) Create a story with your work so far which explains how you’re thinking about the data (click boxed icon, drag sheet, add a brief caption to tell someone else about the progression in your understanding of the data, add a new point and continue). Explain what is interesting to you about the data/process in the story, and set up background for what you will explore next.
* Save your workbook



**DO ANY OF YOU HAVE A REALLY NICE DASHBOARD THAT YOU WANT TO SHARE WITH THE CLASS?**

**Explore a subset of data vs. year, what do you find interesting?**

* Pick a region of interest and create a subset of those countries. Explore trends in the measures you find most interesting vs. year. Compare to the rest of the world. Explore trends over time for the individual countries. Save each in a new worksheet, and format each chart.
  + What is interesting about these trends?
  + What conclusions can you draw using the regional data and the country-level data?
  + Create some interesting Bubble Charts, Tree Maps or Heat Maps to highlight your favorite results, do these help visualize that data? Save each in a new worksheet and format. Does they tell you anything new?
* Create a new dashboard with your best results from the entire lab
* (Depending on time) Finish your story, adding relevant worksheets and captions that lead you to the conclusions you found above

**CHECK IN WITH YOUR NEIGHBORS, WHAT HAVE YOU LEARNED THAT IS DIFFERENT THAN WHAT THEY HAS LEARNED?**

**Create a final dashboard and finish your story**

* Create a new dashboard with your best results from the entire lab
* (Depending on time) Finish your story, adding relevant worksheets that lead to the conclusions you found above
* Save your workbook
* Explore other students' dashboards (and stories), are they more similar or different than you expected given the same data set, but open ended instructions?

**DO ANY OF YOU HAVE A REALLY NICE FINAL DASHBOARD THAT YOU WANT TO SHARE WITH THE CLASS?**

**Part II: Burning Glass xCase**

Create relevant visualizations