**Exercise 6.7**

**Storyboard Development for Global Financial Inclusion Dataset**

1. **Who** is it for?

The client is a non-profit group in the U.S looking to gain insights to the economic progress in developing countries.

1. **Why** is it being built?

The client is interested in identifying the attributes and relationships that lend to greater financial inclusion for people in developing countries and lend to their greater participation in the world economy and their higher standards of living.

1. **What** will it consist of?

The storyboard will illustrate financial quintiles across persons across 140 countries and the attributes that correlate such as education, sex, urbanity, employment, and access to financial products.

1. **When** will it be used?

The findings will be used to identify which countries have higher counts of the population with demographic groups who need greater assistance to improve their participation in financial independence and higher standards of financial inclusion.

1. **Where** will it be hosted?

Tableau Public

What analysis is relevant? How will you introduce and wrap up the story? For this example, the following structure has been developed to ensure the analysis and its impact on real estate pricing is clear to users:

1. Income Quintile map by country.
2. Exploratory Analysis: Which attributes have stronger correlations to higher Income quintiles? (Correlation heatmap)
3. Linear Regression: what’s the relationship between education and income quintile?  (Scatterplot)
4. Cluster Analysis: What is the age distribution across all 7 region codes and which have highest longevity (age) distribution? (LARGEST grouping IN 20-40)
5. Cluster Analysis Results: Financial worries and attributes- how is Age, Income and education related?
6. Final Results and Recommendations: answering key client questions about predictors of income mobility, as well as limitations of the analysis and proposed next steps.

* **Create a summary that states how your results are useful based on the initial research questions and hypothesis.**
* The data findings are inconclusive in in illustrating a strong correlation between income mobility and education, neither age, sex, or urbanity/rural locale correlates to higher income levels.
* The data is so large and is intended to reflect 97% of the world’s population and is hard to find strong relationships within such a large pool.
* There were moderately strong correlations between financial concerns as education rises, and higher wages also had a moderately strong correlation to education.
* Concerns about education costs lower as age rises- however financial worries are prevalent across ages, sex, education levels and across all income quintiles.
* There was also a moderate strong correlation between region and financial concerns.
* **Propose next steps for further analysis.**

Compare survey data from 2011, 2014 and 2017 to track progress by country and any setbacks that occurred post 2020 and isolate smaller pools within countries to be able to find more significant trends in lower levels of data.

**•••**

**Include some limitations of your study**.

Data weighting is used to ensure a nationally representative sample for each economy.

* Final weights consist of the base sampling weight, which corrects for unequal probability of selection based on household size, and the poststratification weight, which corrects for sampling and nonresponse error.
* Poststratification weights use economy-level population statistics on gender and age and, where reliable data are available, education or socioeconomic status.

The accuracy of the data depends on participation and getting a selection of responders that proportionately represents the entire population accurately. As well as getting consent from participants and truthful responses.