**6.1 Sourcing Open Data**

**Data Source**

Global Financial Inclusion (Global Findex) Database 2021

Microdata from the World Bank

**Data Sourcing & Collection**

Open data provided by the World Bank made up of external data collected from national representative surveys of almost 145,000 adults in 139 economies representing 97 percent of the world’s population.

In most developing economies, Global Findex data have traditionally been collected through face-to-face interviews.

* In economies where telephone coverage represents less than 80 percent of the population or where in-person surveying is the customary methodology.
* *However, because of ongoing COVID-19–related mobility restrictions, face-to-face interviewing was not possible in some of these economies in 2021.*
* *Phone-based surveys were therefore conducted in 67 economies that had been surveyed face-to-face in 2017.*

In economies where face-to-face surveys are conducted, the first stage of sampling is the identification of primary sampling units.

* *These units are stratified by population size, geography, or both, and clustering is achieved through one or more stages of sampling.*
* *Where population information is available, sample selection is based on probabilities proportional to population size; otherwise, simple random sampling is used.*

Respondents are randomly selected within the selected households.

* *For paper surveys, the Kish grid method is used to select the respondent.*
* *In economies where cultural restrictions dictate gender matching, respondents are randomly selected from among all eligible adults of the interviewer's gender.*

In traditionally phone-based economies, respondent selection follows the same procedure as in previous years, using random digit dialing or a nationally representative list of phone numbers.

* *In most economies where mobile phone and landline penetration is high, a dual sampling frame is used.*

**Reason for study**

I chose this dataset because it is broadly useful information to understand the world’s developing economies. It has relevant information that ties into all my projects in terms of understanding global markets for banking, disease, and marketing ventures for consumer goods. I like that it has both a research aspect in understanding how the world is scaling in advancements and technology that make their lives more convenient and opens opportunities to participate in the global economy, as well as it can be useful to know in business to track emerging markets to expand into.

**Data Limitations**

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.

**Ethics**

Respecting the privacy of responders who may not feel comfortable divulging their personal finance information in terms of where their funding comes from, and how the survey responders save and borrow with these funds. Also, they are not being paid for their data and may not be fully aware of the future uses of it for sales and marketing purposes outside their countries.

**Contents**

The Global Findex is the world’s most comprehensive database on financial inclusion. It is also the only global demand-side data source allowing for global and regional cross-country analysis to provide a rigorous and multidimensional picture of how adults save, borrow, make payments, and manage financial risk.

\**The latest edition follows the 2011, 2014, and 2017 editions, and it includes a number of new series measuring financial health and resilience and contains more granular data on digital payment adoption, including merchant and government payments.*

The Global Findex is an indispensable resource for financial service practitioners, policy makers, researchers, and development professionals.

The survey consists of questions relating to if the respondents:

* have bank accounts,
* household sources of funding.
* Household use of credit cards and digital technologies to make payments on mobile or internet transactions.
* As well as saving and borrowing habits for financial resiliency.

**Data Relevancy**

**Objective:** Understand how education levels are related to income quintile levels across all geographies.

**Hypothesis:** There will be a strong correlation between populations with higher penetration of higher levels of education with higher % penetrations of the population in higher income quintiles as well as if they are in the workforce.

It would be interesting to see if it’s the same of different for males vs. females by country/region.

**Objective:** Understand how age, income, and urbanity correlates to adaptivity to financial products like having a financial account, mobile money account, debit card or credit cards and usage of these products.

**Hypothesis:** Younger survey respondents are more likely to be more frequent users of mobile money, and debit/credit cards and take advantage of different uses.

How is this also influenced by income and region/ cultural norms?

**Objective:** Understand what drives biggest financial concerns across demographics: age, sex, education, urban/rural, and region and how it relates to borrowing and saving.

**Hypothesis:** Those in lower income quintiles and who are dependent on government transfers of agricultural transfers are more likely to be concerned about finances.