The Well Being of Women

By: Betsy Kalkwarf, Roopa Raghav, Jessica Martin, and Paola Rivas

Reason Topic was Selected

As women, we care about women's well-being and strive to figure out what aspects of life most impact if women live long and happy lives.

Source of data

LivWell: A global Longitudinal Database with data related to Women

- "LivWell is a global longitudinal database which provides a range of key indicators related to women’s socioeconomic status, health and well-being, access to basic services, and demographic outcomes.
- Data are available at the sub-national level for 52 countries and 447 regions.
- A total of 134 indicators are based on 199 Demographic and Health Surveys for the period 1990-2019, supplemented by extensive information on socioeconomic and climatic conditions in the respective regions for a total of 190 indicators.
- The resulting data offer various opportunities for policy-relevant research on gender inequality, inclusive development, and demographic trends at the sub-national level." Source

Questions the team hopes to answer with data

- Do different aspects of life (education, household factors, precipitation, fertility, etc.) impact overall women's well-being in similar ways?
- As we advance through the years do we see an increase in healthier women?
- What is the relationship between women's well-being in countries with different levels of development?

Data Exploration and Analysis

Data Exploration

Machine Learning:

- SciKitLearn is the machine learning library we'll be using to create a classifier.
- Our training and testing setup is unsupervised.
- From there a PCA model is initialized and transformed into a dataframe
- The null values were dropped and a scatter plot based on the Elbow Curve to decide the best K-value
- Link: https://github.com/Betsy-Kalkwarf/Women-Well-Being/blob/main/ML_Indicators.ipynb

Database Storage:

- AWS RDS is the database setup we intend to use
- We will integrate Postgres sql to display the ETL process for country demographics.
- Creating S3 buckets
- Link: https://github.com/Betsy-Kalkwarf/Women-Well-Being/blob/main/Analysis_startercode.ipynb

Data Analysis

Data Cleaning and Analysis:

- Pandas will be used to clean the data and perform an exploratory analysis.
- Further analysis will be completed using Python

Dashboard:

- The dashboard will be hosted on Tableau.
- There is no interaction with the dashboard
- Link:

Storyboard

Thanks!