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
  - Image below shows the main stats from the LivWell data

In [19]:
```
livewell_etl_df = livewell_etl_df.loc[livewell_etl_df['year']>= 2000]
livewell_etl_df.describe()
```

Out[19]:

	year	DM_age_15.19_p_se	DM_age_20.24_p_se	DM_age_25.29_p_se	DM_age_30.34_p_se	DM_age_35.39_p_se	DM_age_40.44_p_se
count	5967.000000	5967.000000	5967.00000	5967.000000	5967.000000	5967.000000	5967.000000
mean	2007.827216	1.243674	1.23557	1.239853	1.150581	1.063693	0.969010
std	4.755202	0.576403	0.48798	0.463376	0.411243	0.381618	0.374867
min	2000.000000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000
25%	2004.000000	0.876000	0.92000	0.935167	0.877321	0.808000	0.718000
50%	2008.000000	1.226667	1.18200	1.184000	1.104000	1.020000	0.920000
75%	2012.000000	1.557750	1.46775	1.457321	1.360000	1.262678	1.141833
max	2019.000000	7.924000	6.15400	6.020000	4.656000	4.016000	4.756000

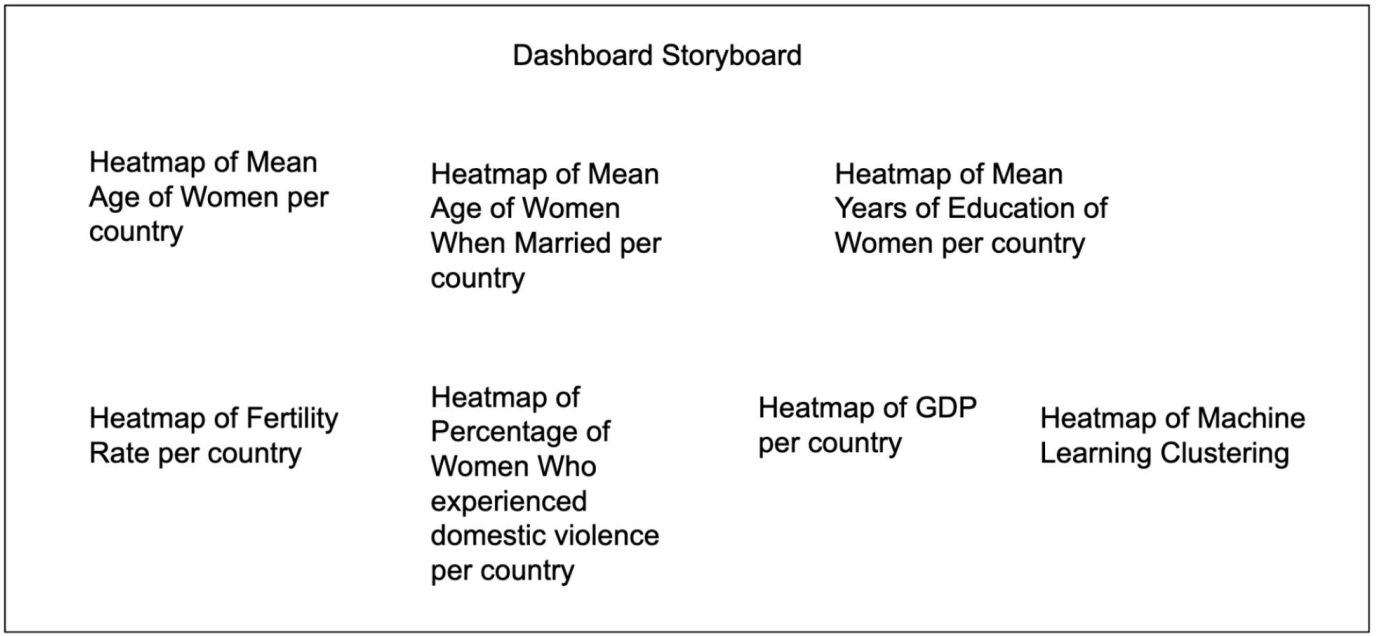
8 rows × 38 columns

Data Analysis

Dashboard:

- The dashboard will be hosted on Tableau.
- Interactive elements include scrolling over country name and cluster number
- Link: https://github.com/Betsy-Kalkwarf/Women-Well-Being/blob/Betsy_branch/Dashboard%20storyboard.pdf

Storyboard


Dashboard Storyboard
Heatmap of Mean Age of Women per country
Heatmap of Mean Age of Women When Married per country
Heatmap of Mean Years of Education of Women per country
Heatmap of Fertility Rate per country
Heatmap of Percentage of Women Who experienced domestic violence per country
Heatmap of GDP per country
Heatmap of Machine Learning Clustering

Thanks!