**The Impact of Education on Child Marriage**

**Team Members: Hephzibah, Leon, Olivia and Shibani**

Child marriages, as defined by UNICEF, are those undertaken by men or women under the age of 18 and include unions where a woman and a man live together as if they were married.

Many factors interact to place a child at risk of marriage, including poverty, the perception that marriage will provide ‘protection’, family honor, social norms, customary or religious laws that condone the practice, an inadequate legislative framework and the state of a country’s civil registration system.

Data Questions:

1. Is rate of child marriage decreasing with each year?
2. Correlation of literacy rate and child marriage
3. Geographical impact on child marriage and education
4. Is one gender more susceptible to child marriage than the other?

Data Sources:

1. Child Marriage Dataset from UNICEF

<https://data.unicef.org/topic/child-protection/child-marriage/>

1. Education Dataset from World Bank

<https://data.worldbank.org/topic/education>

1. Countries dataset

<https://developers.google.com/public-data/docs/canonical/countries_csv>

1. Google Maps API
2. World Bank API

http://api.worldbank.org/v2/countries?format=json

Assumptions:

1. Percentage of women married or in union before age 18 and before age 15 is not the same.
2. The correlation developed for this project assumes that other factors affecting child marriage such as the socio-economic factors, the GDP of the country and its population, the religious belief system and other such factors are equal.

Flow Chart:

1. Data Search: Identifying the data sets that would be used for analysis. Developing questions for analysis forming project proposal.
2. Data Munging/Cleaning: Exporting the data sets into Pandas, cleaning it and formatting it into data tables to ensure optimal efficiency for analysis.
3. Data Analysis: Conducting analysis through narrowing down scope to try and answer specific questions pertaining to the data sets.
4. Presentation: Visually represent the data in graphs, charts, heat maps and regression lines to try represent key findings in an optimal manner.

Data Munging Process:

**STEP 1:**

* 1. Exploring raw data trying to identify improvement areas
  2. Removing specific columns and rows irrelevant to the purpose of analysis
  3. Rename columns to more descriptive names
  4. Merge certain datasets to ease analysis
  5. Drop any null values that are meaningless to analysis

**STEP 2:**

* 1. Identifying the specific functions to perform the cleaning
  2. Reasoning the best type of join for merging in order to maintain data quality and integrity
  3. Standardize format of key column to integrate data sets seamlessly (changed to lowercase string)
  4. Creating new data tables containing the clean data

**STEP 3:**

1. Looking over the clean data tables to verify data is clean and prepared for analysis
2. Count rows in data tables to ensure we haven’t lost a significant amount data while cleaning

Extracting Data:

CHILD MARRIAGE DATASET:

1. Percentage of women (aged 20-24 years) married or in union before age 18
2. Percentage of women (aged 20-24 years) married or in union before age 15
3. Percentage of men (aged 20-24 years) married or in union before age 18

EDUCATION DATASET:

1. The literacy rate was chosen as the indicator for education dataset

COUNTRIES DATASET:

1. For information of Latitude and Longitude for each country

COUNTRIES INCOME LEVEL DATASET:

1. For information of income level for each country

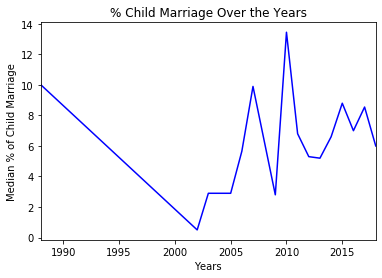
Data Cleaning:

* 1. Dropped columns that were irrelevant to analysis
  2. Made countries column values to lowercase to ensure standardized format for join
  3. Renamed columns to more descriptive names
  4. Created final data frame for analysis with cleaned data

Data Merging:

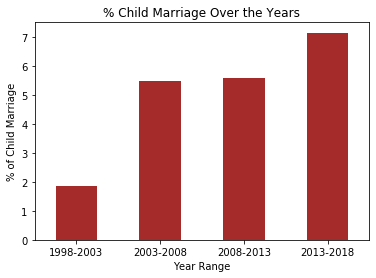
* + Used Inner join
    - Merged Education data set with countries data set ON countries
    - Merged Child Marriage data set with countries data set ON countries
    - This allowed us to get the latitude and longitude values in each data set

**Question 1: Is child marriage decreasing over time?**

****

Drastic drops and rises in the graph indicating lack of data in certain years for some countries which is essentially skewing the median calculation

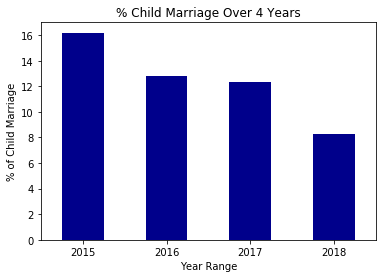
This analysis showed us that it might be more insightful to create ranges of the years to try and identify a trend in the data and avoid the significant drops and rises



We see a significant average increase over the years in child marriage percentage

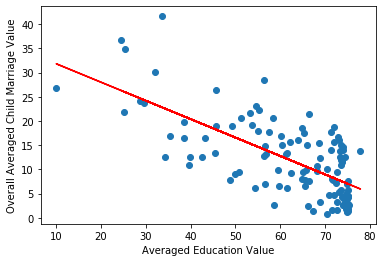
Lack of observations/data points in historic years may be cause of low median initially

In the range 1998-2003; there were 3 observation values in comparison to the 186 observation values in the years ranging from 2013-2018



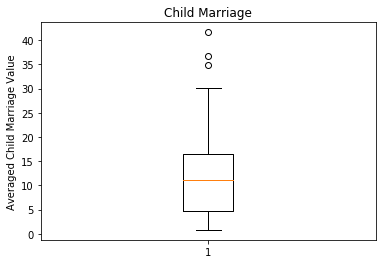
We had sufficient data points from year 2015 to 2018. When this data is represented, a steady decrease in the rates of child marriage is seen with each passing year.

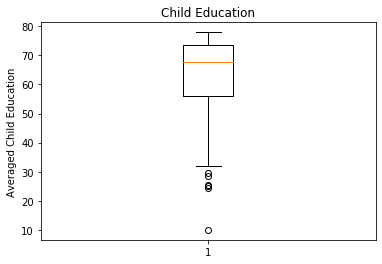
Question 2: What is the correlation of literacy rate and child marriage?

****

The graph shows negative correlation indicating that the average of child marriage value goes on decreasing as the level of literacy rate increases.

The r-square value -0.6880431291174198 indicating a moderate correlation between child marriage and literacy rate.

****

****

OUTLIERS:

Burkina Faso (25.20), Chad (25.40), Guinea (28.70), Mali (29.71), Niger (24.50), Somalia (10.03)

OUTLIERS:

Central African Republic (41.63), Chad (34.86), Niger (36.66)

Question 3: What is the geographical impact on child marriage and education?

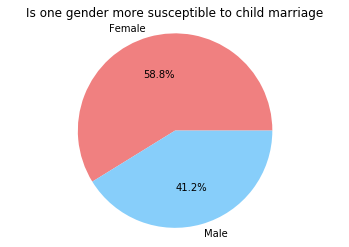


CHILD MARRIAGE



Literacy Rate

Question 4: Is one gender more susceptible to child marriage than the other?



To provide a non-biased distribution of the genders susceptible to child-marriage, only the indicators of child marriage for men and women below 18 were considered.

Null and Alternative Hypothesis

* When calculating the p-value; we chose to calculate the statistical significance of two populations:
  1. countries with low education/child marriage rates
  2. countries with high education/child marriage rates

P-Value = 1.835636112113863e-09

* As the p-value is lower than our statistical significance of 5%, we therefore reject the null hypothesis that education rates has no effect on child marriage rates.

CONCLUSION:

It is seen that increase in literacy rate is correlated with the reduction of child marriages.

Child marriages are one of the major factors contributing to complications during pregnancies and a leading cause of death of girls in low- and middle-income countries. Child marriage often compromises a girl’s development by resulting in early pregnancy and social isolation, limiting her opportunities for career and vocational advancement and placing her at increased risk of domestic violence. Marriage may similarly place boys in an adult role for which they are unprepared, and may place economic pressures on them and curtail their opportunities for further education or career advancement.

Hence, it is vital to boost the literacy rate to bring down the rate of child marriage.