Global Food Prices and population Trends

Data Science 2 – Statistics for Data Science Summary Report

**Submitted by: Group 2**

Jobbin Samuel

Ricky Lam

Maxim Smetanin

Ramila Mudarth

Lizhong Wang

James Lising

# Objective

The objective in our assignment was to determine the impact of food prices of specific commodity in developing countries against its GDP and Population Trends – Birth Rate, Death Rate and Child Mortality to determine any kind of correlation that might exists using last twenty years of data.

We predefined the two hypotheses prior to our analysis as follows

**Hypothesis 1:** How much food price influence Population? Null Hypothesis is food price isn’t a key driver of population. The alternative Hypothesis is that food price somewhat affect population.

**Hypothesis 2:** How much do food prices impact all population trends and GDP parameters? Null Hypothesis is that food prices impact all parameters equally. The alternative hypothesis is that there are some differences between some parameters affected by food prices.

# Data Source

The preparation of the data set for modelling required combination of Global Food Prices Data from [WFP (The World Food Program)](https://www.wfp.org/),  GDP Trends, Unit of Measure Equivalization Table, Exchange Rate, Birth Rate, Death Rate and Child Mortality. The final table had twenty years of Food prices in US dollars as well as the corresponding Birth, Death, Child Mortality and GDP. The US dollar conversion we used the current rate for simplicity.

# Data Analysis

When a categorical correlation matrix was ran for different commodities type, the values were close to zero as a result, the conclusion was that there could be no relationship between the different commodities.

When a correlation matrix was run using one commodity – Rice, initial high-level observations from the Heat Map shows strong positive correlation between the following parameters

* Birth Rate, Fertility, Death Rate and Child Mortality – suggesting that they are colinear. This is a natural observation, for example, we would expect birth rate to depend on fertility.

The negative correlation exists in the following parameters

* GDP per capita and birth rate, child mortality, fertility, death rate
* Price of Rice versus Child Mortality, Birth, Fertility and Death rates

# Modelling

## Linear Regression analysis was used to analyze relationship between rice price and birth rate.

Scatter plot of Rice prices and birth rate shows no correlation on global and continent levels (Figure 1A). Linear regression analysis shows no relationship between birth rate and rice on the continent level. Pearson squared coefficient is equal to 0.07, p-value for rice coefficient is 20%. We fail to reject null hypothesis – there are no influence of rice price on the continent level.

However, when we increase granularity and examine by Asian regions (Figure 1B), countries (Figure 1C) and country level (Figure 1D), the correlations becomes more defined. For example, Afghanistan (Figure 1D) R squared is equal to 0.62, the coefficient is -45.2 and p-value 1%. Therefore, with 99% confidence we reject null hypothesis and conclude that there is an impact of the food price on the birth rate in Afghanistan.

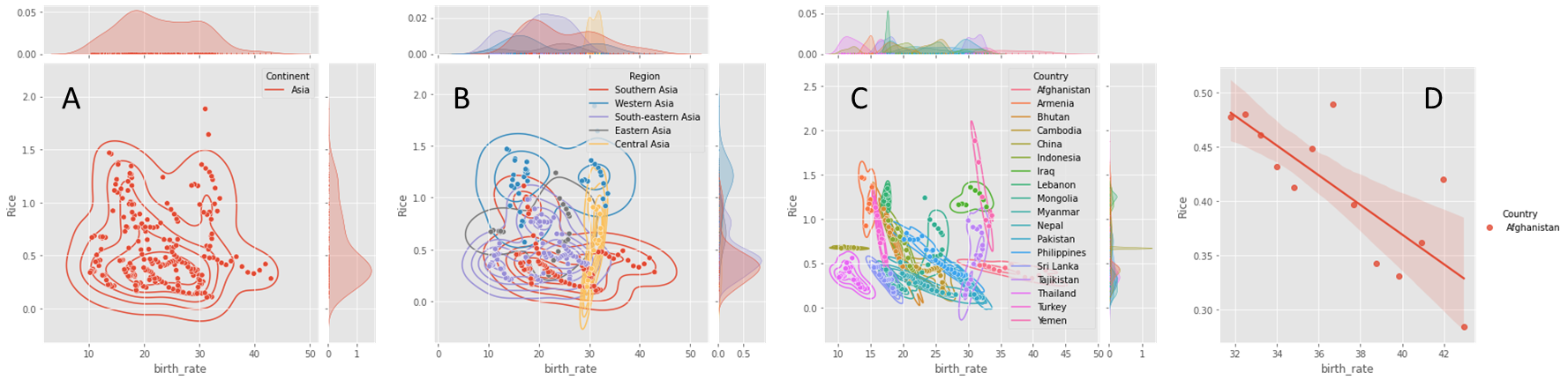


Figure 1. Average rice price and birth rate in Asia (A), Asia regions (B), Asian countries (C) and Afghanistan(D)

## Multi Variate Linear regression modelling was conducted for the Rice commodity Price against

**Boot strapping Results**

# Conclusion