Food Safety in Africa

Data Cleaning, Visualization, and Statistical Analysis with R

Yunting Chiu

4/29/2020

Contents

1 - Team Members

- **Yunting Chiu**: tidy data, statistical analysis, project management, data mapping, data visualization.
- **Kingsley Ofoegbu**: data analysis, data-driven hypotheses, data visualization, report writing, tidy data, regression.
- **Shan Lin**: report writing, literature review, data evaluation.
- **Doudou Shi**: report writing, data visualization, data analysis, statistical analysis.

2 - Introduction

Food insecurity is still a major global concern as 1 billion people are suffering from starvation and malnutrition, and the Food and Agriculture Organization of the United Nations (FAO) has concluded that we are still far from reaching millennium development goal (MDG) Number 1: to halve extreme poverty and hunger. Especially in sub-Saharan Africa, where the number of people suffering from hunger is estimated at over 200 million, and this figure could increase in the near future. In this project, we hope to answer certain questions with the data set we have and proffer possible solutions to bring about better Food security in Africa. We know Africa has rich resources when it comes to agriculture and rich soil, but they lack the proper equipment to harvest them is lacking. After this, we should be able to comfortably suggest measures to counter the gaping holes in infrastructure, food supply, hazard control, and questions that need answering.

The analysis is based on Global Food Safety Partnerships (GFSP) dataset from 49 countries between 2006 to 2017. We are interested in quantitative measurement for this project, especially we compared Africa's GDP and total population.

3 - Data Cleaning

3.1 - Loading Libraries

3.2 - Load the data first, and reformating features

3.3 - Reshaping with multiple columns

3.4 - rename the data frame

3.5 - select the columns

```
## # A tibble: 5 x 13
##   Country A8_ImplCat A11_YearInit A12_YearEnd A13_TimeFrame Activity
##   <fct>   <fct>             <int>       <int>         <int> <chr>
## 1 MALAWI  Enterprise         2010        2016             6 Extensi...
## 2 MALAWI  University         2012        2017             5 Researc...
## 3 SOUTH ... NGO                2013        2016             3 Legisla...
## 4 KENYA   NGO                2016        2020             4 Other t...
## 5 MOZAMB... NGO                2009        2015             6 Staff t...
## # ... with 7 more variables: Commodity_cat <chr>, Donor_cat <chr>,
## #   NonDonor_cat <chr>, Total_budget_USD <dbl>, E2_FSBudget <dbl>,
## #   Category <fct>, Hazard_cat <chr>
```

4 - Initial Hypotheses

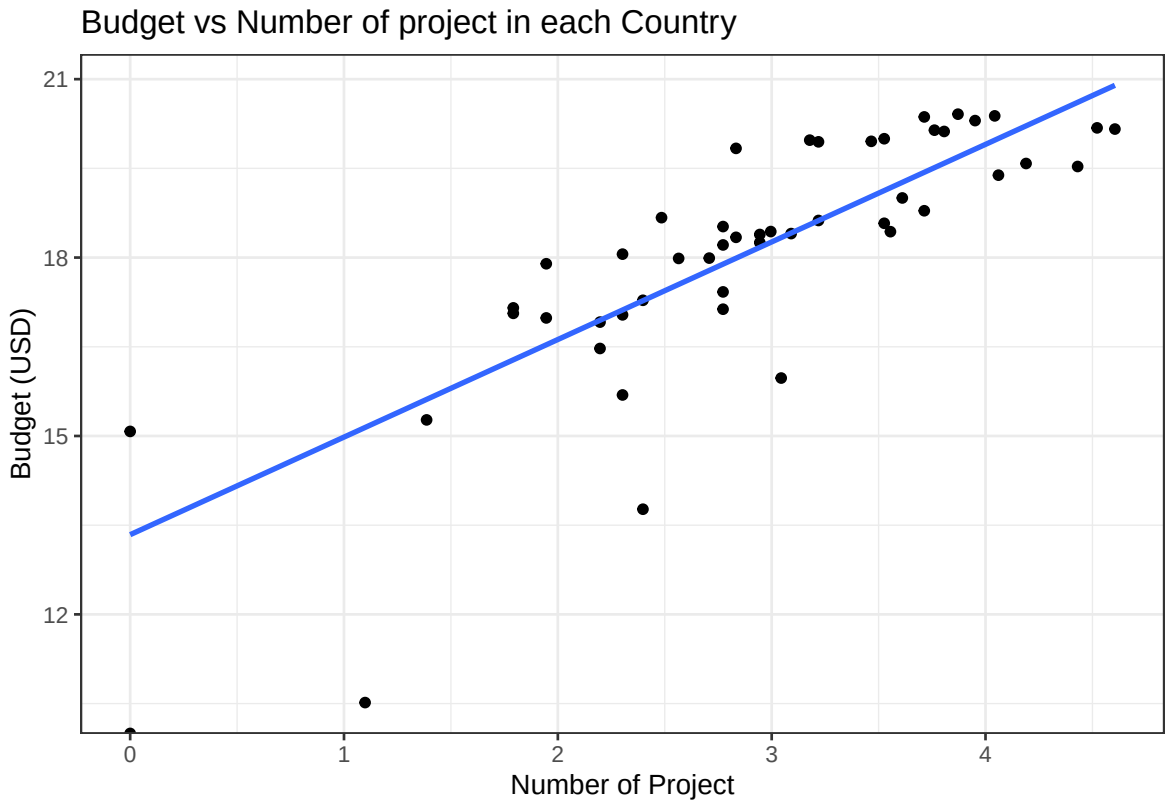
4.1.1 Which countries are funded the most and the least?

```
## # A tibble: 1 x 3
##   Country total_budget project
##   <chr>          <dbl>   <int>
## 1 ZAMBIA    731977868.      48
```

```
## # A tibble: 1 x 3
##   Country   total_budget project
##   <chr>            <dbl>   <int>
## 1 EQ GUINEA            0       1
```

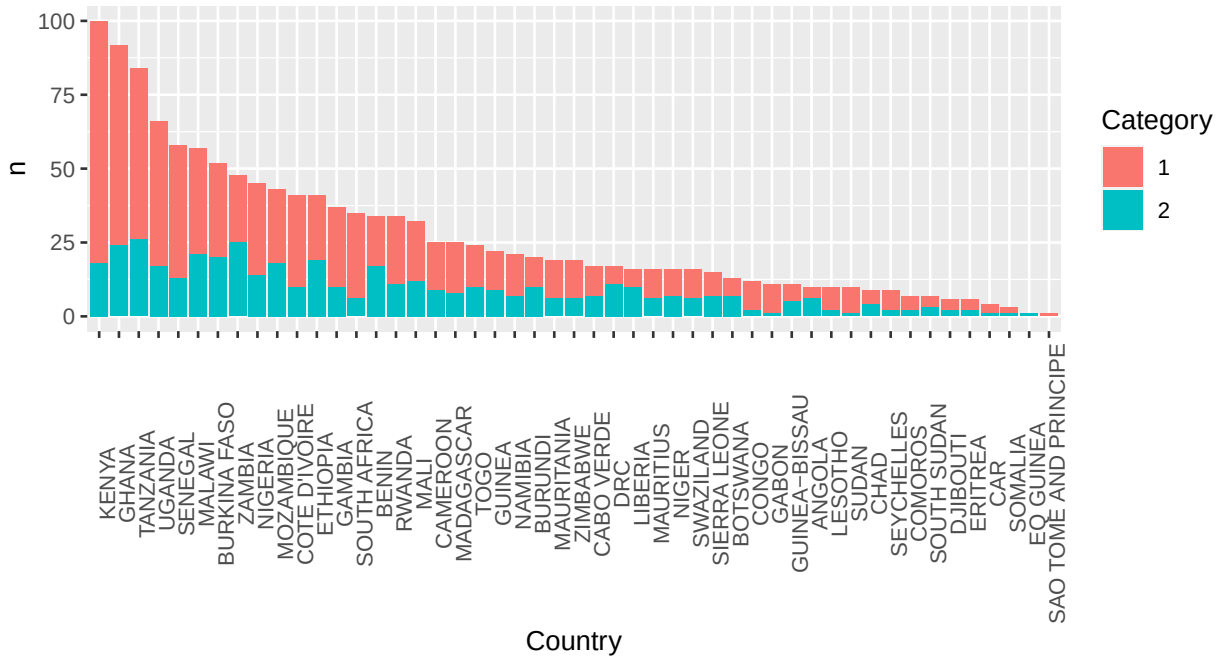
4.1.2 Budget vs Number of project in each Country

- Here we create scatter plots and box plots to ascertain if there is a linear relationship.
- Number of projects done per country and Total Funds allocated per country, from this scatter plot. we can see that the countries that got more fundings had more projects, there is a linear relationship between if a country gets more funding.



4.1.3 - Distribution projects in respect of the country

- we can see each country of project type.



5 - Exploratory Data Analysis

5.1 - Datasets

5.1.1 - which countries are involved in the project?

- 49 countries

```
## [1] 49
```

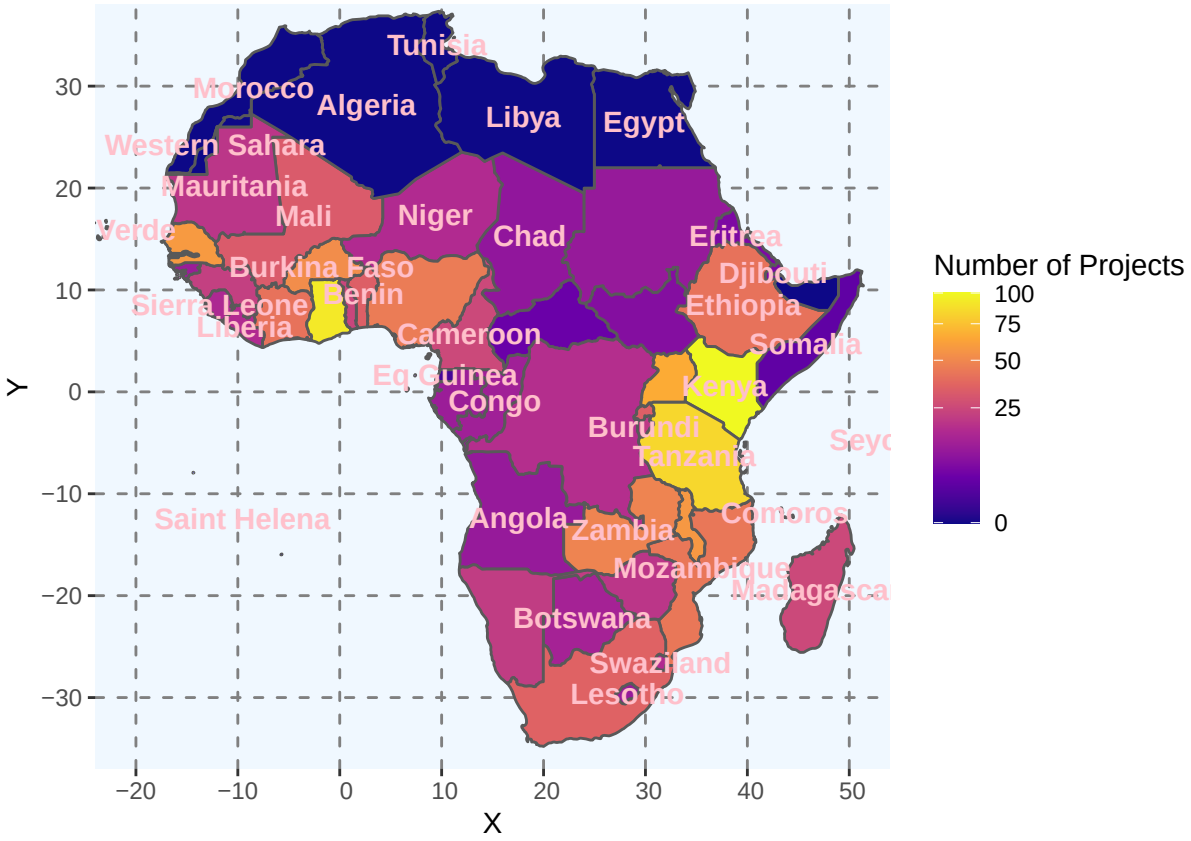
5.2 - Data Visualization

5.2.1 - Drawing Africa maps programmatically

- Obviously, the Africa map is easy to let readers know the association between the number of aid programs and each country.

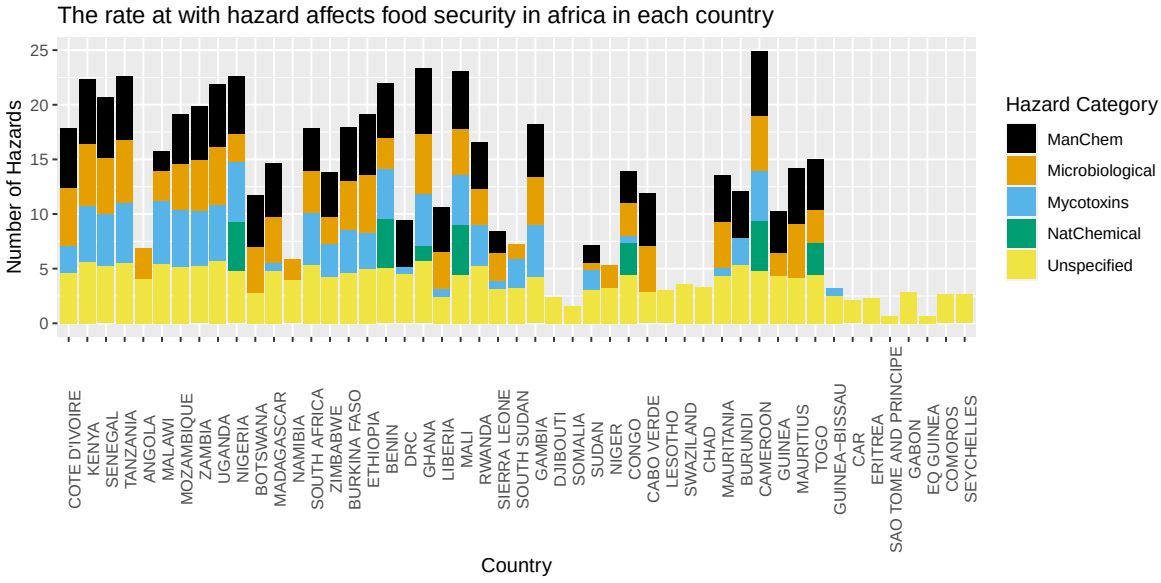
```
## # A tibble: 24 x 2
##    Country                          name
##    <chr>                            <chr>
##  1 ALGERIA                          Algeria
##  2 CABO VERDE                       <NA>
##  3 CAPE VERDE                       Cape Verde
##  4 CAR                              <NA>
##  5 CENTRAL AFRICAN REPUBLIC         Central African Rep.
##  6 CONGO                            <NA>
##  7 COTE D'IVOIRE                    <NA>
##  8 DEMOCRATIC REPUBLIC OF THE CONGO Dem. Rep. Congo
##  9 DRC                              <NA>
## 10 EGYPT                            Egypt
## # ... with 14 more rows
```

```
## # A tibble: 8 x 2
##   Country        name
##   <chr>          <chr>
## 1 ALGERIA        Algeria
## 2 EGYPT          Egypt
## 3 LIBYA          Libya
## 4 MOROCCO        Morocco
## 5 SAINT HELENA   Saint Helena
## 6 SOMALILAND     Somaliland
## 7 TUNISIA        Tunisia
## 8 WESTERN SAHARA W. Sahara
```



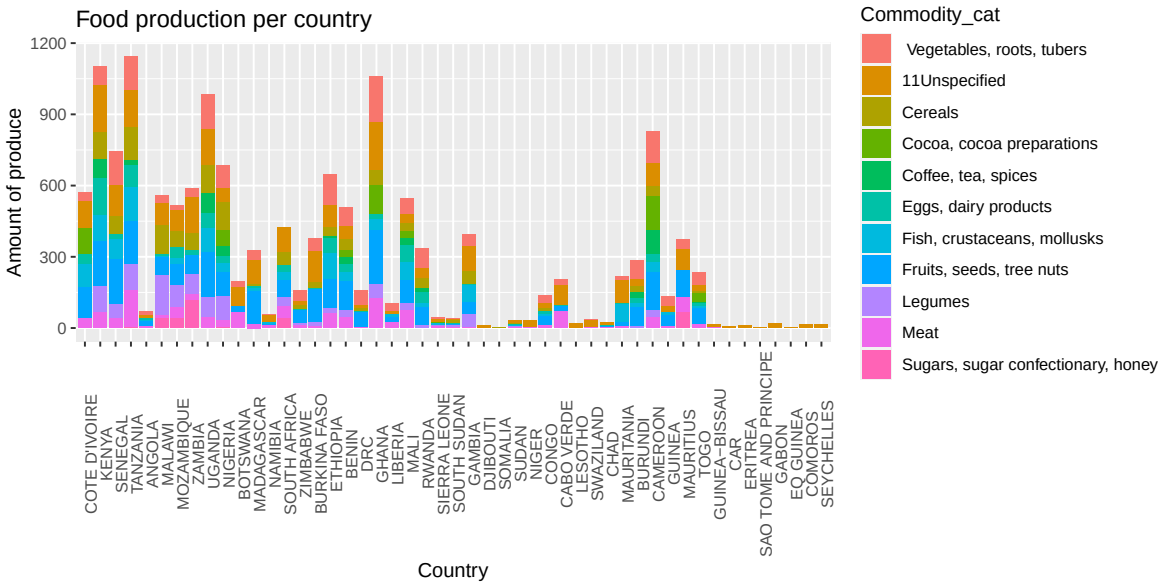
5.2.2 - The rate at with hazard affects food security in africa in each country

- Here we know that different hazards affects different crops but what are those hazards and how much do they affect security in each country, this shows just that and with it organizations would know how to tailor their plans for specific countries.



5.2.3 - Which country is vegetarian and non-vegetarian?

- Here we want to find out which country receives what type of commodity the most.



5.3 - Statistical Inference

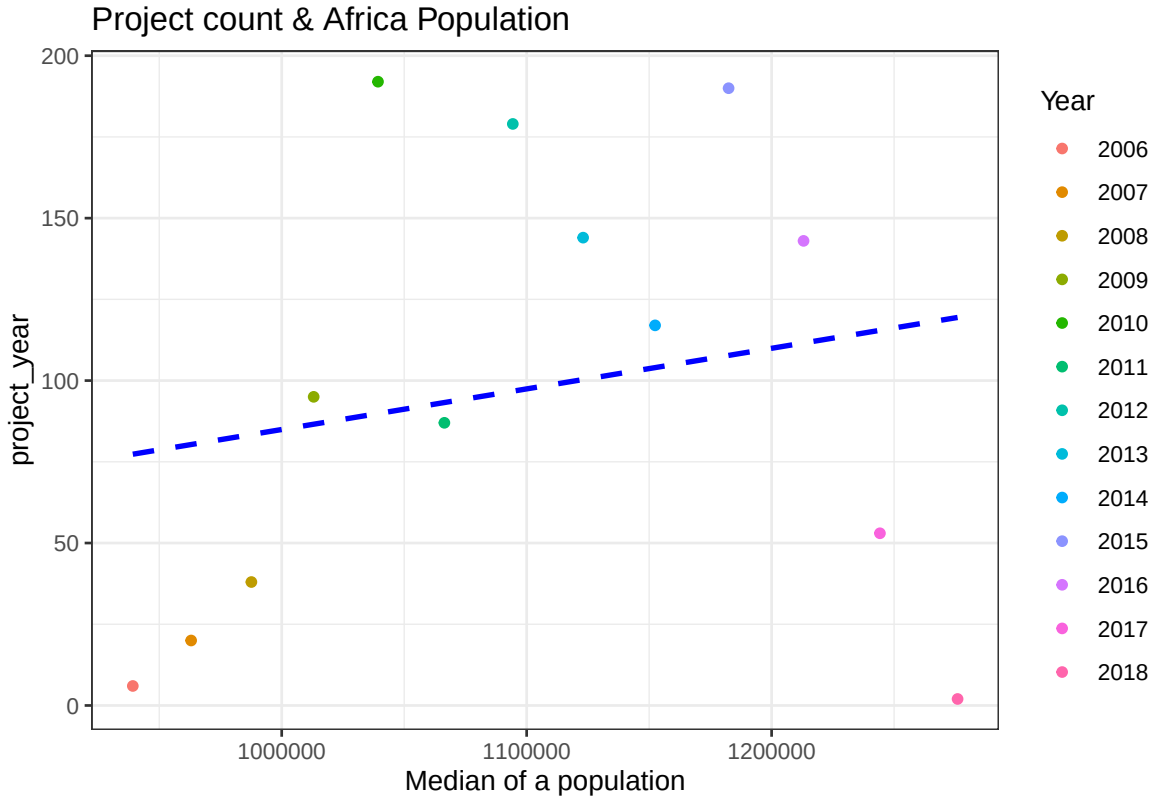
5.3.1 - Calculate the positive correlation with medians of Africa population and the Food Safety project.

- As the project start years are between 2006 to 2018(removed NA). In this case, we download the 2006-2018 Africa area total population from UN data to compare it.

```
## # A tibble: 13 x 4
##    `Country or Area` Year  Variant    Value
##    <chr>             <fct> <chr>      <dbl>
##  1 Africa            2018  Medium  1275921.
##  2 Africa            2017  Medium  1244222.
##  3 Africa            2016  Medium  1213041.
##  4 Africa            2015  Medium  1182439.
##  5 Africa            2014  Medium  1152434.
##  6 Africa            2013  Medium  1123045.
##  7 Africa            2012  Medium  1094343.
##  8 Africa            2011  Medium  1066410.
##  9 Africa            2010  Medium  1039304.
## 10 Africa            2009  Medium  1013046.
## 11 Africa            2008  Medium   987623.
## 12 Africa            2007  Medium   963022.
## 13 Africa            2006  Medium   939210.
```

- Choose project start year from Food Safety In Africa data ,and plot

- The plot seems Food Safety project not deeply help Africa population to increase



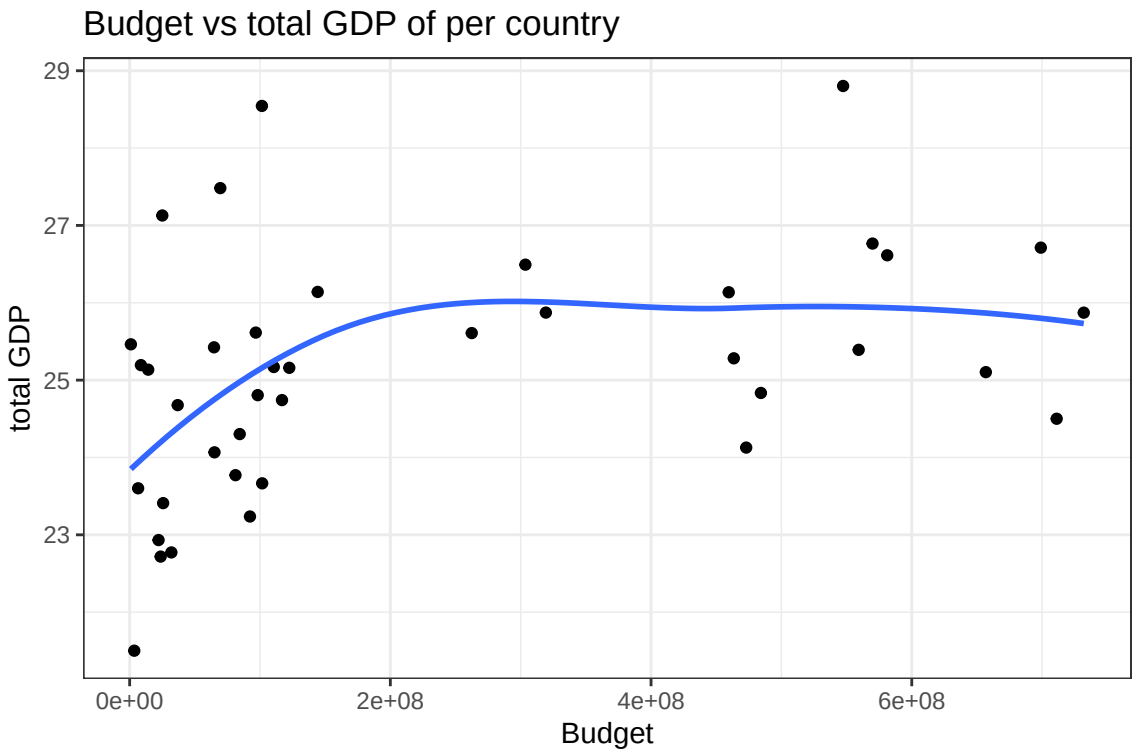
5.3.2 - Does international investment in food in African countries affect local GDP?

- Now, we want to analyze whether food aid to African countries has an impact on their GDP. First, take the GDP data of all the countries in Africa, and then select the data from 2006 to 2017.

```
## # A tibble: 6 x 14
##   Country  `2006`  `2007`  `2008`  `2009`  `2010`  `2011`  `2012`  `2013`
##   <chr>     <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
## 1 ARUBA   2.42e 9 2.62e 9 2.75e 9 2.50e 9 2.39e 9 2.55e 9 2.53e 9 2.58e 9
## 2 AFGHAN... 6.97e 9 9.75e 9 1.01e10 1.24e10 1.59e10 1.78e10 2.00e10 2.06e10
## 3 ANGOLA  5.24e10 6.53e10 8.85e10 7.03e10 8.38e10 1.12e11 1.28e11 1.37e11
## 4 ALBANIA 8.90e 9 1.07e10 1.29e10 1.20e10 1.19e10 1.29e10 1.23e10 1.28e10
## 5 ANDORRA 3.54e 9 4.02e 9 4.01e 9 3.66e 9 3.36e 9 3.44e 9 3.16e 9 3.28e 9
## 6 ARAB W... 1.40e12 1.64e12 2.08e12 1.80e12 2.11e12 2.50e12 2.79e12 2.87e12
## # ... with 5 more variables: `2014` <dbl>, `2015` <dbl>, `2016` <dbl>,
## #   `2017` <dbl>, tot_gdp <dbl>
```

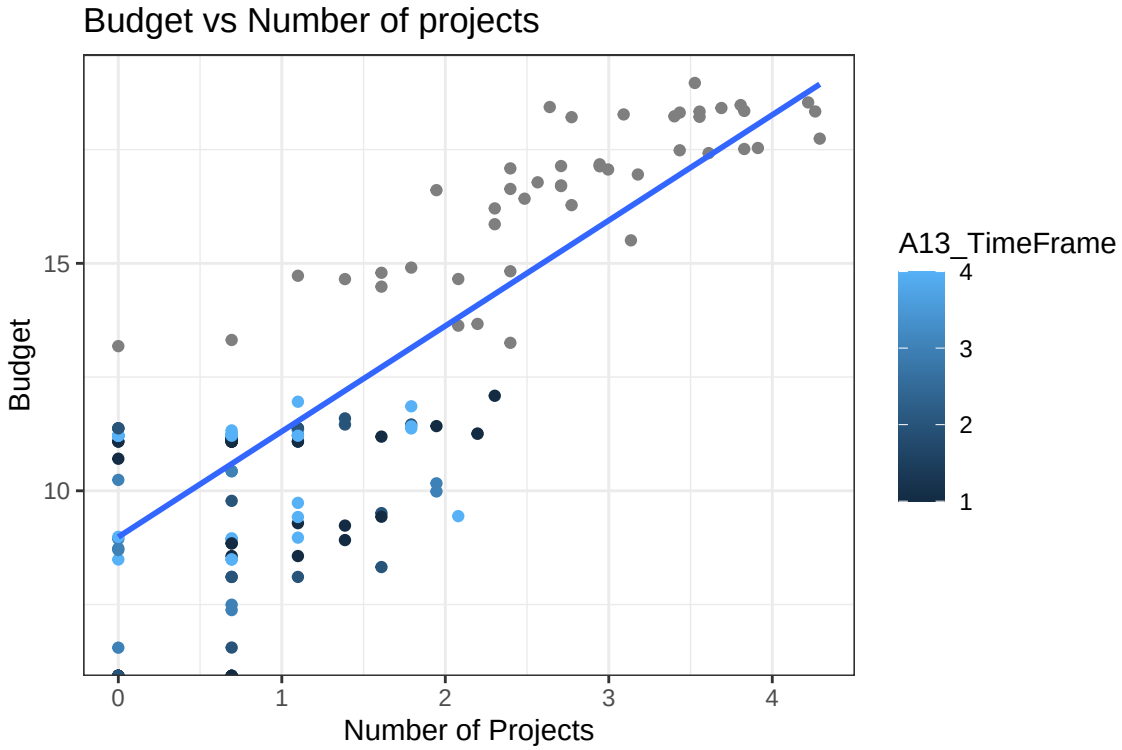
- Then, we combine the investment table with the GDP table, and select the countries for which investment data are available.At the same time, rows with missing values and duplicate rows are cleared from the data

- Finally, as shown in the figure below, the data does not show a significant correlation. Thus, it can be concluded that food investment has no direct bearing on the GDP growth of the countries concerned.



6 - Data-driven Hypotheses

6.1 - Here we want to test if our initial hypothesis is true (fail to reject) or not(reject)



6.2 - Multiple regression on the hypothesis:

```
## # A tibble: 131 x 4
## # Groups:   Country [48]
##    Country      A13_TimeFrame Budget Number_of_project
##    <chr>                <dbl>  <dbl>             <dbl>
##  1 ANGOLA                   1  11.1              0.693
##  2 ANGOLA                  NA  16.6              1.95
##  3 BENIN                    1   9.29             1.10
##  4 BENIN                   NA  18.3              3.43
##  5 BOTSWANA                 2  11.4              0
##  6 BOTSWANA                NA  16.6              2.40
##  7 BURKINA FASO             1   8.57             0.693
##  8 BURKINA FASO             2  11.6              1.39
##  9 BURKINA FASO            NA  18.5              3.81
## 10 BURUNDI                  1   8.92             1.39
## # ... with 121 more rows
```

```
##
## Call:
## lm(formula = A13_TimeFrame ~ Number_of_project + Budget, data = the_model)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.91334 -1.25797 -0.01318  1.09562  2.03850
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         2.68104    1.05466   2.542   0.0129 *
## Number_of_project -0.44121    0.21857  -2.019   0.0468 *
## Budget              0.02096    0.10232   0.205   0.8382
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.259 on 81 degrees of freedom
##   (47 observations deleted due to missingness)
## Multiple R-squared:  0.0479, Adjusted R-squared:  0.02439
## F-statistic: 2.037 on 2 and 81 DF,  p-value: 0.137
```

7 - Discussion

The project of Food Safety in Africa did not magnificently help the Africa population increasing. Therefore, it is recommended that Africa not focus on projects that are not directly linked to food, such as employee training, because this does not play a positive role in their food growth. The investment in food is conducive to the GDP growth of relevant countries.

However, if you directly invest money in food, you can increase crop output, stimulate people to increase food consumption, and thus steadily increase GDP. According to the country's level and conditions, it is necessary to formulate a strategy that adapts to its own country. African countries can increase their investment in food and reduce disasters, thereby increasing food production and solving Africa's food problems.

8 - References

- Breman, H., & Debrah, S. (2003). Improving African food security. *SAIS Review, 23(1), 153-170.*
  Artilce for our reference.

- Food Safety in Africa: Past Endeavors and Future Directions.(2019, February 5). Retrieved from https://datacatalog.worldbank.org/dataset/food-safety-africa-past-endeavors-and-future-directions
  This data is the open-source from The World Bank Group collected. The Global Food Safety Partnership's (GFSP) Food Safety in Africa provides an approach to illustrative information on 518 food safety investments in sub-Saharan Africa from 2010 to early 2017.

- List of sovereign states and dependent territories in Africa. (2020, April 27). Retrieved from https://en.wikipedia.org/wiki/List_of_sovereign_states_and_dependent_territories_in_Africa
  Africa Countries renamed.

- Lovelace, R., Nowosad, J., & Muenchow, J. (2020, April 21). Chapter 8: Making maps with R. Retrieved from https://geocompr.github.io/geocompkg/articles/solutions08.html
  How to create Africa map with the existing dataset.

- Mwaniki, A. (2006). Achieving food security in Africa: Challenges and issues. *UN Office of the Special Advisor on Africa (OSAA).*
  Artilce for our reference.

- New GFSP Report Quantifies Food Safety Investment in sub-Saharan Africa. (n.d.). Retrieved from https://www.gfsp.org/new-gfsp-report-quantifies-food-safety-investment-sub-saharan-africa
  Global Food Safety Partnership's website was announced the official information for this data frame. We can download reading guide via this association.

- UNdata | explorer. (n.d.). Retrieved from http://data.un.org/Explorer.aspx?d=PopDiv
  Get some quantitative data such as death rates or total population for the countries.