DSO 545 Final Report: Agricultural Production in TPP countires

Team Agriculture Research

May 8, 2018

**Team member: Takaaki Kihara, Zichan Han, Tsao Yu, Enhao Zhao, Can Xu**

We did an overview on Global Agricultural Production, with focus on TPP countries, including the United States, Mexico, Canada, South Asia country etc. This report will consist the following part: introduction, data description, data cleaning, exploratory data analysis, and conclusion with future works. Dashboard will show in the end as Appendix.

## Introduction

Agricultural product is highly related to people’s daily life as people need eat food everyday. How price of the food is changing highly affects the agricultural productions and poverty. In this sense, we did some data analysis and try to found out what factors affect the food price and productions, how are more related factors are affected each other in agricultural, and how agriculture and poverty are related.

We first did an overview on Global Agricultural Production, with focus on TPP countries, including the United States, Mexico, Canada, South Asia country etc. Why we choose TPP countries is that we think they play a very important role in the world agriculture market in terms of production and price. We also see a trend that they will continue playing more important role in future. To better understand the dynamic among those countries, we conducted our research on TPP countries.

## Data Description

The data we have used is Crops and agricultural production related data. We found the following data from FAO. The details for each data will show below:

* Food and Agriculture Organization of the United Nations (FAO), with time series data from 1961-2016.This dataset recorded the Area, Crops, Year and the Production, Yield and Growing area of each of the crops.
* Agriculture price data: this is a dataset from FAO, with the time series data from 1991 - 2016, recording the Area, Crops, Year and the crop prices in each of the Area in different years. The unit of the prices is in both local currency and USD, we chose to use the USD as our unit.

## Data Cleaning

After we gather the data in the above section, we chose the data that is useful for our problem statement. We did the couple cleaning work as follow shows detaily:

* We used tidyr to gather the clumns with year information to a value column to reduce columns.
* We used tidyr to spread the different categories of data from a simgle column called“Element” to separate values into individual columns.
* We substitude the spaces in column names to match the appropriate format.
* We used dplyr to filter out the NA values from the dataset so that we can get valid data and use in our analysis.

## Data Analysis

Our problem statement is —————————-, We did the analysis into the following two part:

*Agriculture production among TPP countries* Rice, Wheat, Barley, Maize, Soybeans, Sugar beet, Sugar Cane, Oil crop *Price and production trend analysis on certain corps* Avocados, Blueberries, Quinoa, Wheat, Rice,Sugar cane

First part is to understand an overview of the agriculture production among TPP countries, based on certain corps and time series analysis. Second part is a price and production trend analysis on certain corps and time series analysis. Combine the two part, we are trying to find what factors influence the price of the corps, and the agriculture production.

# Part 1: Agriculture Production among TPP Countries

TPP countries including the United States, Mexico, Canada, South Asia country etc, we first did an overview on Global Agricultural Production. Why we choose TPP countries is that we think they play a very important role in the world agriculture market in terms of production and price. We also see a trend that they will continue playing more important role in future. To better understand the dynamic among those countries, we conducted our research on TPP countries. The following will show the data analysis on certain corps, including Rice, Wheat, Barley, Maize, Soybeans, Sugar beet, Sugar Cane, Oil crop etc.

* RICE

Through the following map time series analysis, we have the following findings.

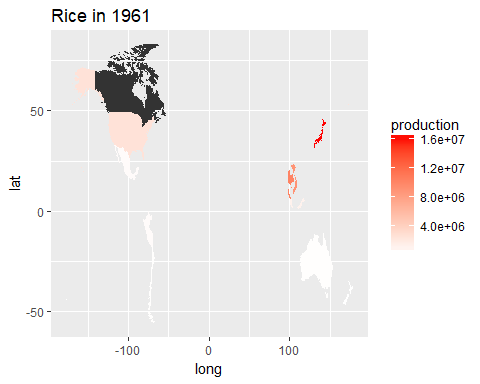
\*Top rice production shift from Japan to Southeast asia (1990)

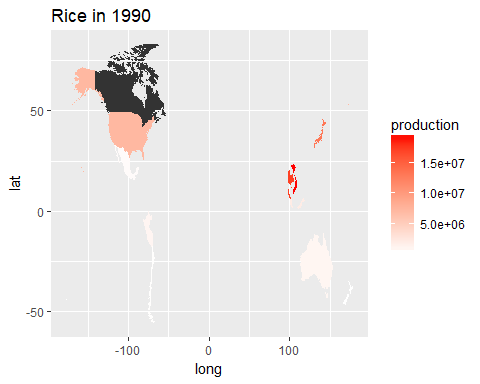
Japan’s rice crisis starts with its older, smaller stomachs. As the population ages, appetites are shrinking. Diets among younger Japanese favour wheat and the country is eating about 20 percent less rice than it did two decades ago. Other sources of demand are also vanishing: Japan drinks about a third as much (rice-based) sake as it did in 1970 and consumption of fish — the traditional accompaniment to rice — is down 30 per cent since 2005. Besides, Japan’s average daily calorie intake has been on the decline, falling to 2,415 due to the population is getting older.

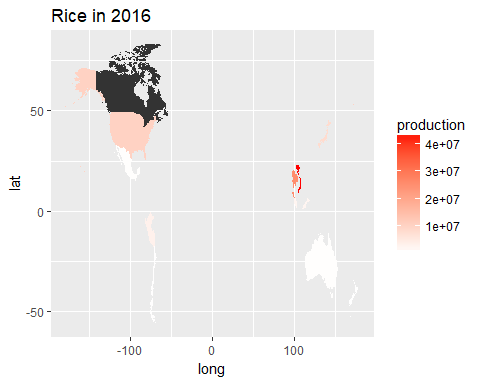
\*Vietnam becomes the main rice production countrie

Vietnam: government support→ The Mekong Delta is the heart of the rice producing region of the country where water, boats, houses and markets coexist to produce a generous harvest of rice. At the end of the Vietnamese war in 1975, the government supported the development of the Mekong delta in South Vietnam. For example, the government supplied to the farmers seed, fertilizer and other essentials. This increase is attributed to planting of modern early maturing rice varieties, better management, appropriate Legal Constructs by the Government.

production = fread("Production\_Crops\_E\_All\_Data.csv")   
   
production = production %>%  
 select(-ends\_with("F")) %>%  
 gather(Y1961:Y2016, key = "Year", value = "production")   
   
 production$Year = str\_sub(production$Year,2)  
 production$Year = as.numeric(production$Year)  
 production$Area = as.character(production$Area)  
   
 production\_TPP = production %>%  
 filter(Area %in% c("United States of America",  
 "Japan",  
 "Canada",  
 "Mexico",  
 "Peru",  
 "Chile",  
 "New Zealand",  
 "Australia",  
 "Brunei Darussalam",  
 "Singapore",  
 "Malaysia",  
 "Viet Nam",  
 "Thailand") &  
 Item %in% c("Rice, paddy",  
 "Wheat",  
 "Barley",  
 "Maize",  
 "Soybeans",  
 "Sugar beet",  
 "Sugar cane",  
 "Oilcrops, Oil Equivalent") &  
 Element == "Production" &  
 Year %in% c(1961, seq(1960,2010,10),2016))   
   
 world\_map = map\_data("world")  
 world\_map$region[world\_map$region == "Brunei"] = "Brunei Darussalam"  
 world\_map$region[world\_map$region == "USA"] = "United States of America"  
 world\_map$region[world\_map$region == "Vietnam"] = "Viet Nam"  
   
 TPP\_map = world\_map %>%  
 filter(region %in% c("United States of America",  
 "Japan",  
 "Canada",  
 "Mexico",  
 "Peru",  
 "Chile",  
 "New Zealand",  
 "Australia",  
 "Brunei Darussalam",  
 "Singapore",  
 "Malaysia",  
 "Viet Nam",  
 "Thailand"))  
  
production\_TPP\_map = right\_join(world\_map, production\_TPP,  
 by = c("region" = "Area"))  
  
production\_TPP\_map\_1961 = production\_TPP\_map %>%  
 filter(Item == "Rice, paddy" & Year == 1961)  
  
 ggplot(TPP\_map, aes(x = long,  
 y= lat,  
 group = group),  
 fill = "white",  
 color = "black") +  
 geom\_polygon() +  
 geom\_polygon(data = production\_TPP\_map\_1961,   
 aes(x = long,  
 y= lat,  
 group = group,  
 fill = production)) +  
 scale\_fill\_gradient(low = "white",  
 high = "red",  
 na.value = "white") +  
 ggtitle(paste("Rice in 1961", sep ="/"))

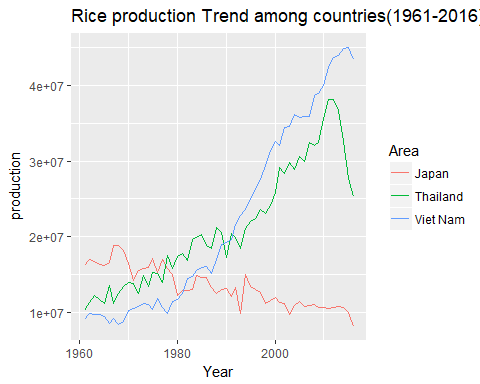






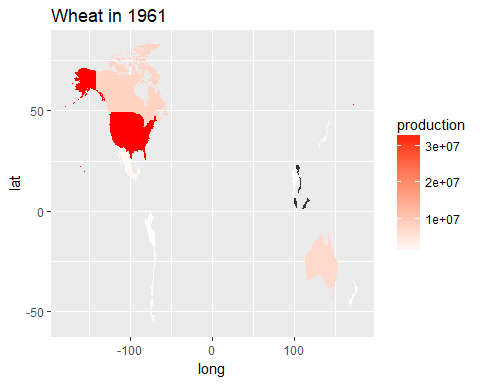
The following chart showed how the top-rice-production country of Rice shifted form Japan, to Thailand and Vietnam from 1961 to 2016

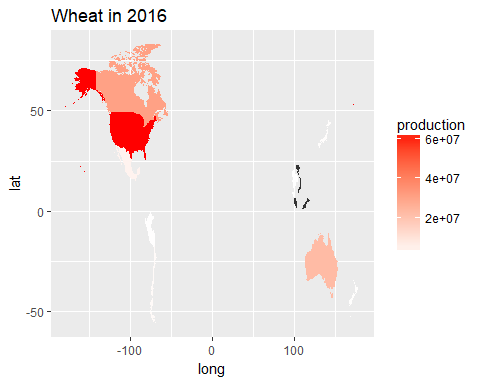
production = fread("Production\_Crops\_E\_All\_Data.csv")   
  
production = production %>%  
 select(-ends\_with("F")) %>%  
 gather(Y1961:Y2016, key = "Year", value = "production")   
  
production$Year = str\_sub(production$Year,2)  
production$Year = as.numeric(production$Year)  
production$Area = as.character(production$Area)  
production$production = as.numeric(production$production)  
  
production\_TPP\_rice = production %>%  
 filter(Area %in% c("Japan",  
 "Viet Nam",  
 "Thailand") &  
 Item == "Rice, paddy" &  
 Element == "Production")   
  
ggplot(production\_TPP\_rice, aes(x = Year))+  
 geom\_line(aes(y = production, colour = Area))+  
 ggtitle("Rice production Trend among countries(1961-2016)")

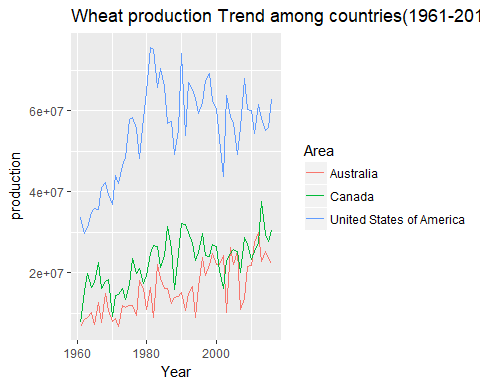


* Wheat

*The United States is ranked first in production volume of wheat in TPP* 1990-2016 Canada and Australia have been increasing production Australia–>1996 and 2010 Australia’s wheat growers achieved 49% of their yield potential because wheat growers have adopted advances in technology and adapted them to their needs. They have adopted improved varieties as well as improved practices, including reduced cultivation (or “tillage”) of their land, controlled traffic to reduce soil compaction, integrated weed management and seasonally targeted fertiliser use.

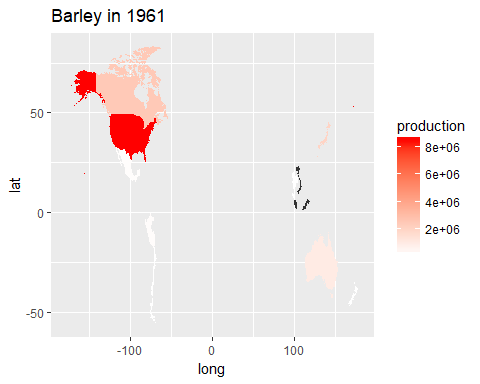


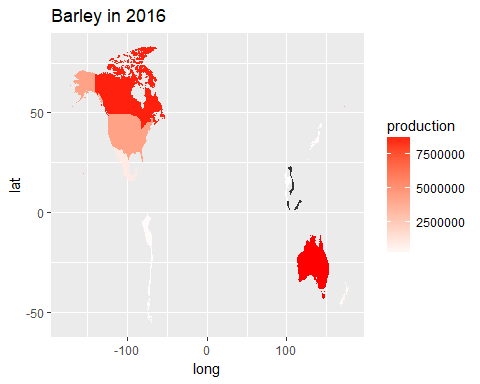


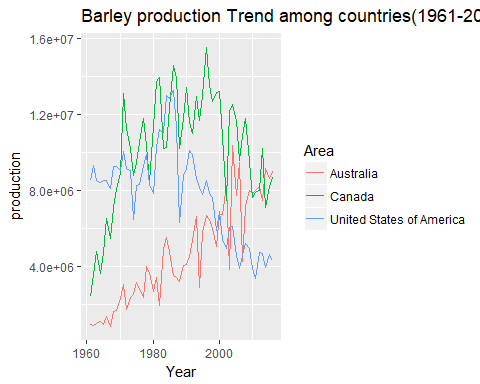
From the chart below we can see the production country of wheat did not shift. 

* Barley

1. top barley production from the State to Canada–>The pattern of crop receipts has changed considerably over time, Australia has been increasing from 2000 and right not the as much as Canada





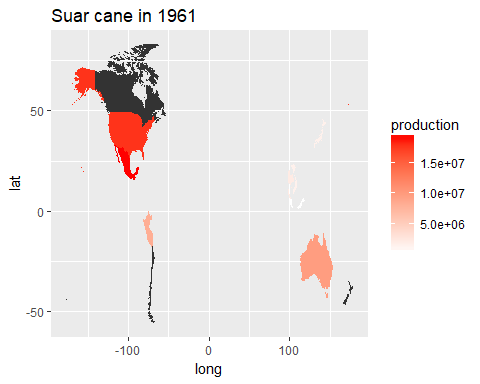
From the chart below we can see the production country of Barley shifted a little. 

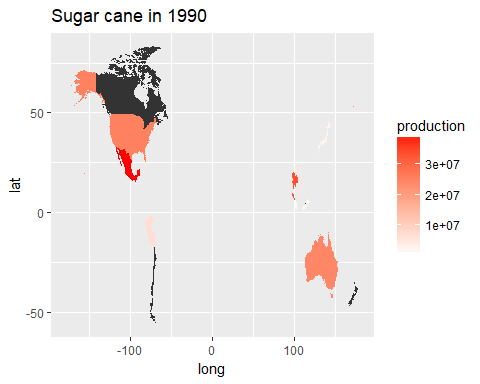
* Sugar Cane
* Mexico had been ranked the top of sugar cane production country from 1961-1990. However, Mexican sugar industry is recognized as uncompetitive in nature after 1990, mainly due to stagnating low sugarcane and sucrose yield and highly fiber in stalks per hectare and variability in the production process.

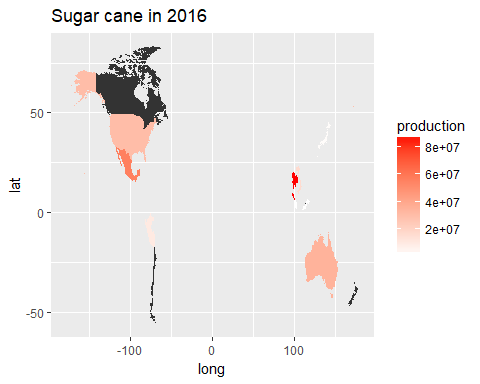
Half of all Mexican sugarcane production occurs on ranchos that are 2 hectares or less in area. This small size of sugarcane farms effectively prevents the use of automated equipment, hence keeping sugarcane a manual labor operation. Manual harvesting requires skilled laborers as incorrectly harvested cane leads to loss of cane and lower sugar yield, poor juice quality, and problems in milling due to extraneous matter.

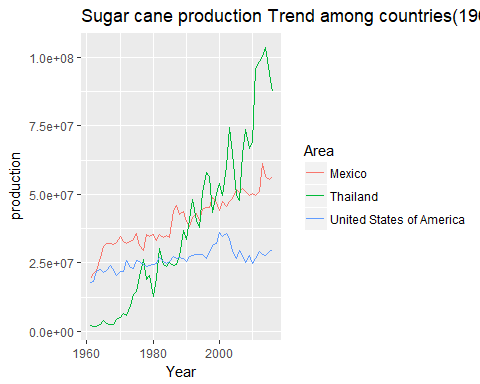
\*Since 1990 Thailand has became the biggest production–>

Following a decade and a half of sustained expansion, sugarcane areas and production increased sharply in recent years, with a record output achieved in 1996. The recent success of the industry can be attributed to several key factors, including attractive sugarcane prices, sugar factory relocation and capacity expansion policies which have successfully encouraged the extension of sugarcane areas. A third factor has been favourable weather. Since less than 10 percent of sugarcane area, now over one-million ha, is irrigated, favourable rainfall distribution has been an important factor in improved yields.







From the chart below we can see the the main prodcution country of sugar cane shifted from US and Mexico to Thailand. 

# Part 2: Price and Production Trend Analysis on Certain Corps

Price usually have a direct effect on production, for example if the price is high, the production should be high, because people want to earn the high price of production to increase the profit.

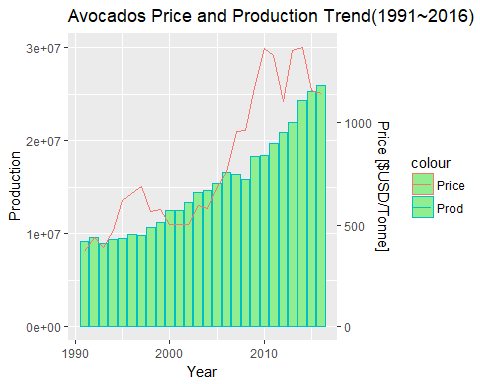
However, through our data analysis we have different foundings. The price is not the mainly influential factors. It happens more often when price go down, the production keep going up, and for certain corps, it seems that the price does not related to the production amount. The main reason should be the agriculture product is a necessary food for everyday, and the production is affect a lot by the government policy and the technique improvements. We try to give a detailed explanation in the following for certain corps, including Avocados, Blueberries, Quinoa, Wheat, Rice, Sugar cane etc.

### Avocados

* Demand for avocados has grown exponentially

People realise its health benefits worldwide. It is also driving prices up. Prices are stable from 2010 and even show a slightly rising trend. Several countries want to benefit from this growth, so a lot is being planted.

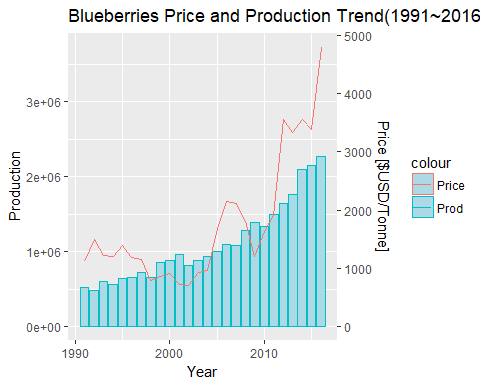
harvest = fread("Production\_Crops\_E\_All\_Data.csv")  
price = fread("Prices\_E\_All\_Data.csv")  
  
  
harvest = harvest %>%  
 select(-ends\_with("F")) %>%  
 gather(Y1961:Y2016, key = "Year", value = "production")  
  
harvest$Year = str\_sub(harvest$Year,2)  
harvest$Year = as.numeric(harvest$Year)  
harvest$Area = as.character(harvest$Area)  
  
harvest <- harvest %>%  
 select("Area", "Item","Element","Year","production") %>%  
 spread(key = Element, value = production)  
  
names(harvest) <- gsub(" ", "\_", names(harvest))  
  
harvest <- harvest %>%  
 select(-Feed, -Area\_harvested, -Yield) %>%  
 filter(!is.na(Production),  
 Year >= 1991)  
  
  
price = price %>%  
 select(-ends\_with("F")) %>%  
 gather(Y1991:Y2016, key = "Year", value = "Price")  
  
price$Year = str\_sub(price$Year,2)  
price$Year = as.numeric(price$Year)  
price$Area = as.character(price$Area)  
  
price <- price %>%  
 select("Area", "Item","Element","Year","Price") %>%  
 spread(key = Element, value = Price)  
  
names(price) <- gsub(" ", "\_", names(price))  
  
price <- price %>%  
 filter(!is.na(`Producer\_Price\_(USD/tonne)`))  
  
  
yearly\_production <- harvest %>%  
 group\_by(Year, Item) %>%  
 summarise(total\_prod = sum(Production))  
  
yearly\_price <- price %>%  
 group\_by(Year, Item) %>%  
 summarise(avg\_price = median(`Producer\_Price\_(USD/tonne)`))  
   
  
price\_prod <- inner\_join(yearly\_price, yearly\_production,  
 by = c("Item","Year"))  
  
price\_prod\_index = price\_prod %>%  
 filter(Item == "Avocados" ) %>%  
 mutate(index = total\_prod/avg\_price)   
  
multiple = mean(price\_prod\_index$index) \*1.1  
  
price\_prod\_index %>%  
 ggplot(aes(x = Year)) +  
 geom\_col(aes(y = total\_prod, colour = "Prod"), fill = "lightgreen") +  
 geom\_line(aes(y = avg\_price\*multiple, colour = "Price")) +  
 scale\_y\_continuous(sec.axis = sec\_axis(~./multiple, name = "Price [$USD/Tonne]"))+  
 ggtitle("Avocados Price and Production Trend(1991~2016)")+  
 xlab("Year")+  
 ylab("Production")



### Blueberries

* Expecting increased consumption

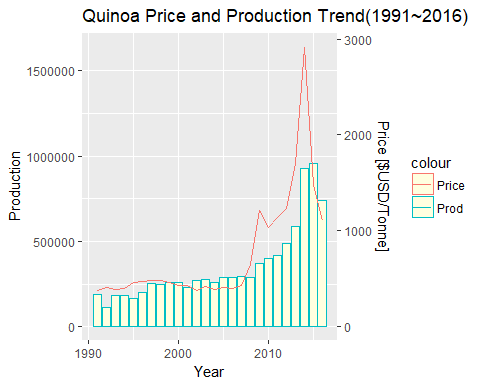
Among the top fruits – including bananas, strawberries, apples, grapes and citrus – blueberries are the only one expecting increased consumption and there is still further room for growth.



### Quinoa

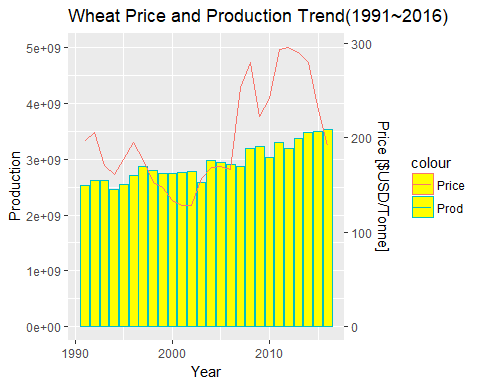
* Demand ourstripped supply

As the rich world discovered the grain, demand outstripped supply. The soaring price transformed global production, prompting a large increase in supply, which ultimately pricked the quinoa price. The existing farmers increasingly used their quinoa crop to sell, and new growers were tempted in. Today more than 50 countries around the world are growing it. Many countries are holding on to unsold quinoa stock. As a result, the price of quinoa has fallenthe price of quinoa has fallen dramatically.



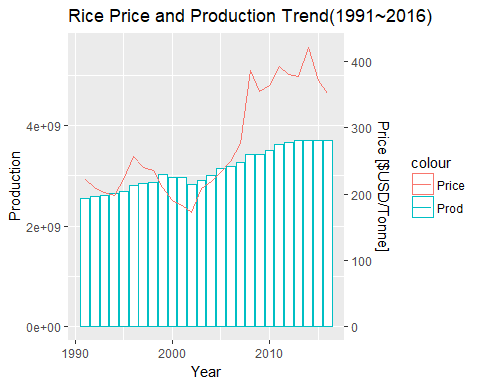
* Wheat

Wheat is mostly produced for human consumption, although between 30-40% of total wheat output goes directly into animal feed. Wheat prices correlate to the prices of corn, barley, and sorghum as they are substitutes in animal feed compound. Beside food and feed, a small percentage of worldwide wheat production will go into the production of wheat ethanol. The by-product of wheat ethanol, wheat ddgs, is a good source of animal feed protein and competes with rapemeal and soymeal. Wheat prices of ddgs tend to be lower than other feed ingredients since it is mainly composed of a by-product. As the price of wheat is affected by lots of reasons while wheat plays an important role as a critical food resource, the price of wheat fluctuates a lot while the the production keeps growing constantly.



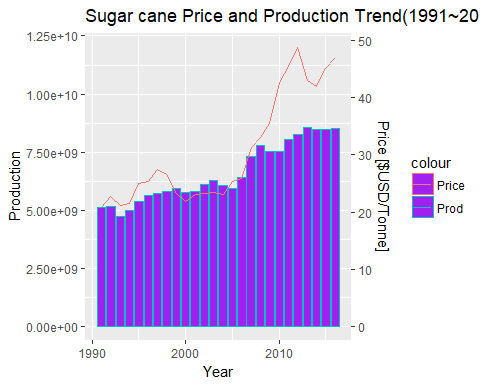
* Rice

Since the early 1990s, international rice prices have followed a falling tendency, mainly reflecting yield gains, bumper crops, changes in policies in major producing countries, and exchange rate movements. Presently, there are numerous international price quotations for rice. In 2008, during the global financial crisis, the price of rice more than tripled, which had a substantially negative impact on the poor, especially in Asia. Although rice prices have fallen back, they have remained significantly above historical price levels, with the price in 2011–2012 between 61% and 71% higher than the average prices in 2006–2007.



* Suger cane

Presently sugar is one of the most regulated commodities being subject to quotas, subsidies and import duties in almost every country. Since 1990 the consumption of sugar has been growing at a rate of around 5% per year in Asia and the Middle East and 4% in Africa. World consumption at the moment is increasing. While consumption of sugar has gone up, sugar prices in recent years have been under pressure because of large supplies in the world market.



# Conclusion & Future Works

In conclusion, we did an overview on Agriculture production among TPP countries, we could understand the corps production shift among countries from years, because of government policy, country economy growth (brief write down the reason found in the previous data analysis part). We also did price and production trend analysis on certain corps, we could found that the price is not a main factor affect the corps production. Different type of the corps have different factors to affect the production. Such as fruit, it affect more by the demand, and for corps, it more affect by the government policy.(just write what found and reasons on the previous data analysis part. Showed above is just example.

For the future work, we expect the TPP country will keep the current production of certain types of corps in few decades. For example Vietnam will keep produce Rice. However, we could also keep an eye on the government policy change and the international relationship among these countries,in this way, we could predict the types of corps of future agricultural production shift among these countries.

Besides, we recommend keep tracking the price and production data for certain type corps. after we know the factors affect the production and the price relationship, we could better predict the future price trend, or production trend for certain corps.

# Appendix

Data Dashboard

The first part analysis is using the shiny R app. Show the following as an example, the left have drop down that you could choose the specific corps, and the specific year to display. On the right, you could see the color change, as deeper red color means more production.

* Insert screenshot 1

The second part have two types of graph, first is line chart, second is histogram combined with line chart.It shows year on the x-axis with the total production on the y-axis. (could add more)

* Insert screenshot 2

### Reference

Data Source(data source website)

Resources (all wiki and other website)

<http://www.solmexiconews.com/mexican-sugar-cane-industry-swing-that-machete/> <http://www.fao.org/docrep/005/X0513E/x0513e24.htm#PRODUCTION> <https://www.economist.com/blogs/economist-explains/2016/05/economist-explains-17> <https://www.commoditybasis.com/wheat_prices> <https://www.commoditybasis.com/sugar_prices>