DSO 545 Final Report: Agricultural Production in TPP countries

Team Agriculture Research, May 8, 2018

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

Executive Summary

Agricultural product is highly related to people's daily life as people need to eat food everyday. In this sense, 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.

We first build shinyr dashboard to see the dynamic change of the crops from year to year. In this way, we get the idea of how the production of each product changes. Then, we focus on the most basic as well as important crops to see their production trend and try to find the reason behind it. With that in mind, we move on to our second step of research and try to see the price change with the production variation. This time, we choose some high value food such as blueberry, avocado etc. to try to dig more information about the change and causes.

Problem Statement

We want to explore how the dynamic of the production and price change among the major products with the focus in TPP countries. We all know the importance of agriculture to the world, we tried to find out what are the reason affect the production and thus can guide us make more predictions in future. By doing so, we could better plan what we want to grow to better serve the society.

Key Findings

In the first part, we could see that the production among countries are highly affected by the government policy, and what other countries production amount on the same corps, through the following findings:

- Main Rice production area shifted from Japan to Southeast Asia countries.
- The United States is ranked first in production volume of wheat in TPP
- 1990-2016 Canada and Australia have been increasing production
- In 1990, Thailand became the biggest sugar cane production while Mexico declined significantly.

In the second part , We could see that price is not the main and not the only factors affect production. Sometimes the price is positive correlated with productions, but sometimes dont. We try to dig the reasons behind the scene, and we have the following findings:

- The price of avocado has increased steadily over the years as people see it as a healthy food.
- 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.
- The demand of Quinoa kept increasing which drove the price and production soaring. However, after reached its peak, the price and production slumped.

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 columns with year information to a value column to reduce columns.

- We used tidyr to spread the different categories of data from a single column called"Element" to separate values into individual columns.

- We substitute 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 that we want to explore how the dynamic of the production and price change among the major products with the focus in TPP countries. Our analysis is conducted in the following two parts:

1. Agriculture production among TPP countries.
    ➢ Rice, Wheat, Barley, Sugar Cane production trend analysis.
2. Price and production relationship
    ➢ Avocados, Blueberries, Quinoa, Wheat, Rice, Sugarcane.

First part is to understand an production 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 with time series analysis. Combining the two parts, we are trying to find the trend and factors influence the production of the corps.

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 and more important role in future. To better understand the dynamics among those countries, we conducted our research on TPP countries. The following shows the data analysis on certain corps, including Rice, Wheat, Barley, Soybeans, Sugar Cane, etc.

Rice

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

❏ Top rice production shift from Japan to Southeast asia

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 country

With 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.

```r
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",
```

```r
                            "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 from 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.

United States has been on the top list of wheat producer for a long time. also, in 2016, 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

❏ Top barley production shift from the United State to Canada and Australia
The pattern of Barley yields has changed considerably over time, Australia has been increasing from 2000 and right now the same as much as Canada.

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

Sugar Cane

- ❏ Mexico had been ranked the top of sugar cane production country from 1961-1990. However, Mexican sugar industry is recognized as non competitive 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. Another 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 production 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, Sugarcane 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.

```r
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 %>%
```

```r
  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 fallen the 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

•   The early 1990s, international rice prices have followed a falling tendency.
This 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.

Sugar Cane

- Sugar cane 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

To sum up, we did an overview on Agriculture production among TPP countries. We understand the corps production shift among countries on a yearly basis. We also did price and production trend analysis on certain corps and found that the price is not a main factor affect the corps production. Different types of the corps have different factors to affect the production. Such as fruit, it affect more by the demand, and for corps, it is more affected by the government policy.

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.

Last but not least, we recommend keep tracking the price and production data for certain types of corp such as wheat. After we get better understanding of the factors affecting the production and the price relationship, we could better predict the future price trend, or production trend for certain corps.

Appendix

Data Dashboard

For the exploratory research, we built two shiny R App(Screenshot of the Shiny R App UI is shown below).

The first App is for overview of TPP countries Agricultural Production. On the left, we have a drop down menu that enable the users to 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.

The Second App is for overview of crop price and production trends over time. On the left, we have a drop down menu that enable the users to choose the specific corps, and the countries to display. On the right, you could see the trends displayed in line chart.

For the specific analysis, we used two types of graphs, 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.

Reference

Data Source

http://www.fao.org/docrep/005/X0513E/x0513e24.htm#PRODUCTION
https://data.worldbank.org/indicator/NY.GDP.MKTP.CD

Resources

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