**MIS 6380.001 - Data Visualization - S18**

**Project Case Study**

**Comparative Analysis of Key Asian Players and USA in the Garments Export Market from 2012 – 2016**

Submitted To:

Prof. Judd Bradbury

Submitted By:

(Group 4)

Annadevara, Divya Sai Srivalli

Karunakaram, Anantha Shashidhar

Seshadri, Sruti

Ammanarul Danasekaran, Thanujhaa Sriee

Bijawat,Sparsh

Contents

[Objective 3](#_Toc512030389)

[Scenario 3](#_Toc512030390)

[Source Data Files 3](#_Toc512030391)

[Data Dictionary: 5](#_Toc512030392)

[Data Cleaning: 7](#_Toc512030393)

[Data Manipulation 10](#_Toc512030394)

[Overview of Visualization Frames and Data Encoding Used 10](#_Toc512030395)

[Explanation of Analysis 19](#_Toc512030396)

[Calculated Measures: 20](#_Toc512030397)

[Insights: 21](#_Toc512030398)

[Data Story: Comparative Analysis of Key Asian Players and USA in the 22](#_Toc512030399)

**DATA VISUALIZATION PROJECT REPORT**

# Objective

To perform a comparative analysis of Key Asian Players and USA in the Garment Exports Market from 2012 - 2016

# Scenario

We are Data Analytics Consultants in the Garments Export Industry entrusted with the responsibility of making recommendations based on the trends in the Garment Exports Market.

# Source Data Files

* **Primary Dataset**
* Export of Garments

This dataset is from the UN Commercial Trade website. This website is the statistical trade data arm of the United Nations. The data in the UN Trade website is accurate and is pooled from official national and customs records.

<https://comtrade.un.org/data>

The data contained columns for Year, Exporter, Importer, Commodity, Net Weight, Trade Value in $ and various individual codes identifying the columns.

* **Secondary Datasets**
* Cotton Production

This dataset is from the Food and Agriculture Organization of the United Nations - FAOSTAT. The data is an official record of the Seed Cotton, Cottonseed and Cotton Lint production across countries in terms of Yield, Production and Area Harvested.

<http://www.fao.org/faostat/en/#data/QC>

The data contained in this dataset contains the Area, Element/Commodity, Year, Value and methods of calculation of the data.

* Labor Wages (Labor Stats)

This dataset is obtained from the “List of Minimum Wages by country” in Wikipedia. It states the official minimum wage rates of the 193 UN member states. The data is normalized across the various sectors in the country.

<https://en.wikipedia.org/wiki/List_of_minimum_wages_by_country>

The data contains columns for Country, Minimum PPP, Nominal Wages for hourly and yearly metrics.

* Labor Productivity (Labor Stats)

This dataset is from the International Labor Organization’s statistical arm - ILOSTAT.

The ILOSTAT Trends Team produces estimates of labor market indicators such as unemployment, employment by sector and status, labor productivity and working poverty, among others, with many estimates disaggregated by age and sex and these marked as ILO modelled estimates.

<http://www.ilo.org/ilostat/faces/oracle/webcenter/portalapp/pagehierarchy/Page3.jspx;ILOSTATCOOKIE=KLHalLUn5gwQxUP0PuCSevFz6CQTjMwgExR1daWNJXi3CUC_0XP1!1506268203?MBI_ID=2&_adf.ctrl-state=1ahpjg95q0_4&_afrLoop=2632410144092190&_afrWindowMode=0&_afrWindowId=null#!%40%40%3F_afrWindowId%3Dnull%26_afrLoop%3D2632410144092190%26MBI_ID%3D2%26_afrWindowMode%3D0%26_adf.ctrl-state%3Doe5mg3j8q_4>

The data contains columns for Country, Year/Time, Output per worker in GDP.

* Unemployment Rates

This dataset is from the International Labor Organization’s statistical arm - ILOSTAT. ILO Modelled estimates are used in this dataset as well.

<http://www.ilo.org/ilostat/faces/oracle/webcenter/portalapp/pagehierarchy/Page3.jspx?MBI_ID=49&_adf.ctrl-state=1ahpjg95q0_72&_afrLoop=2632704605477159&_afrWindowMode=0&_afrWindowId=null#!%40%40%3F_afrWindowId%3Dnull%26_afrLoop%3D2632704605477159%26MBI_ID%3D49%26_afrWindowMode%3D0%26_adf.ctrl-state%3Doe5mg3j8q_21>

The data contains columns for Country, Year, Unemployment rates (%) – ILO modelled estimates, Labor force – ILO modelled estimates, split separately for Male and Female population.

# Data Dictionary:

**Export of Garments**

|  |  |
| --- | --- |
| Column Name | Description |
| Year | Exports during the Years 2012-2016 |
| Country | Origin country that exports the cotton garments |
| Partner | Countries which import cotton garments from the origin country |
| Netweight | Total weight of the garments exported in Kg |
| Trade Value | The total export trade value of the garments in US Dollars |

**Cotton Production**

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| Year | Cotton production data during the Years 2012-2016 |
| Country | Production of cotton in the countries under study |
| Element | Indicates type of raw material  Production - Cotton Lint  Yield - Cotton Seed |
| Item | Total weight of the garments exported in Kg |
| Unit | Tonnes for Cotton Lint, Hectogram for cotton seed |
| Value | The total cotton production in Tonnes |

**Unemployment Rates**

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| Year | Unemployed population data during the Years 2012-2016 |
| Country | Unemployed population in the countries under study |
| Age | Unemployed population in the age ranges of 15-24, above 15, above 25 years |
| Male | Total of unemployed male population |
| Female | Total of unemployed female population |
| Total | Total of female and male unemployed population |

**Labour Stats**

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| Year | Labor statistics data during the Years 2012-2016 |
| Country | Labor statistics in the countries under study |
| Output per Worker | GDP in the country divided by the number of people employed, measure of labor productivity |
| Output Per Worker (PPP) | Standardized labor productivity using Purchasing Power Parity in US Dollars rates in 2010 |

# Data Cleaning:

* Misspelled words, improper cases, trailing spaces were removed in the column header and rows.

The excel functionalities for Spellcheck, Case Conversion Upper/Lower is used for achieving the goal

* The duplicates for the country were all integrated into a single row. The duplicate values were detected using the conditional Formatting option.
* Standardization of all the common Column headers to a single name. This helps us to join the datasheets easily in the Tableau
* The Excel ISBLANK function was used to detect any blank /null values in the rows of the dataset

**Cotton production.xls**

* Removed the excess rows and included data only for Bangladesh, Cambodia, China, India, Indonesia, Myanmar, Pakistan, Philippines, Republic of Korea, Sri Lanka, Thailand, Vietnam
* Fixed spelling errors for the country name Viet Nam to Vietnam
* The dataset had separate rows for North Korea, South Korea which have been integrated into a common row Republic of Korea

**Primary dataset export of garment.xls**

* We have filtered out years and included the export data only for years from 2012 - 2016

**Unemployment Rate.xls**

* Standardized the column name across different datasheets. We have changed the column name "time" to "Year"
* Standardized the column name across different datasheets. We have changed the column name "Reference area" to "Country"
* We have filtered out years and included the export data only for years from 2012 - 2016
* The entire data for unemployment was collected by a model implemented by International Labor Organization (ILO). Hence the column " Labor force modelled ILO " being common to all the rows was deleted

**Labor Productivity.xls**

* The dataset contained separate rows for "Republic of Korea" and "People’s democratic Korea"

Which have been combined as a single column “"Republic of Korea"

* Standardized the column name across different datasheets. We have changed the column name "time" to "Year"

**Method of Approximation of Labor Wages**

**Labor wages.xls**

We consider minimum wages for all countries

**China**

1 CNY =0.16 USD

2300 CNY /MONTH=364.18 USD

Minimum wage is China is set locally

$146 to $667($ 2.78 in shanghai 2.78\*30\*8hrs)

**Bangladesh**

1 BDT =0.012 USD

$ 68 per month cited in wiki for garment workers(https://en.wikipedia.org/wiki/List\_of\_minimum\_wages\_by\_country)

**Cambodia**

140 USD/Month

cited wiki garment and shoe industry

**India**

For example, India has more than 1202 minimum wage rates

Hence considering the maximum wages @2.40\*30   to $11.31\*30

have taken lower estimate (2.40\*30)

**Indonesia**

$100 to $252

**Bangladesh** $ 68

**North Korea** min 5.5 to 11.1 $ per day 5.5\*30= 165$)

**South Korea** $7 to $ 8.5 (7\*30=$210)

**Myanmar** $ 2.70 per day (\*ply by 30) $ 81/Month

**Pakistan** $158

**Sri Lanka** $70.75

**Thailand** $8 per day $240 per month

**Vietnam** $ 121.16 per month

1 USD = 22779.04 VIETNAM DOUG

**Philippines** $ 4.8 \* 30 = $ 144

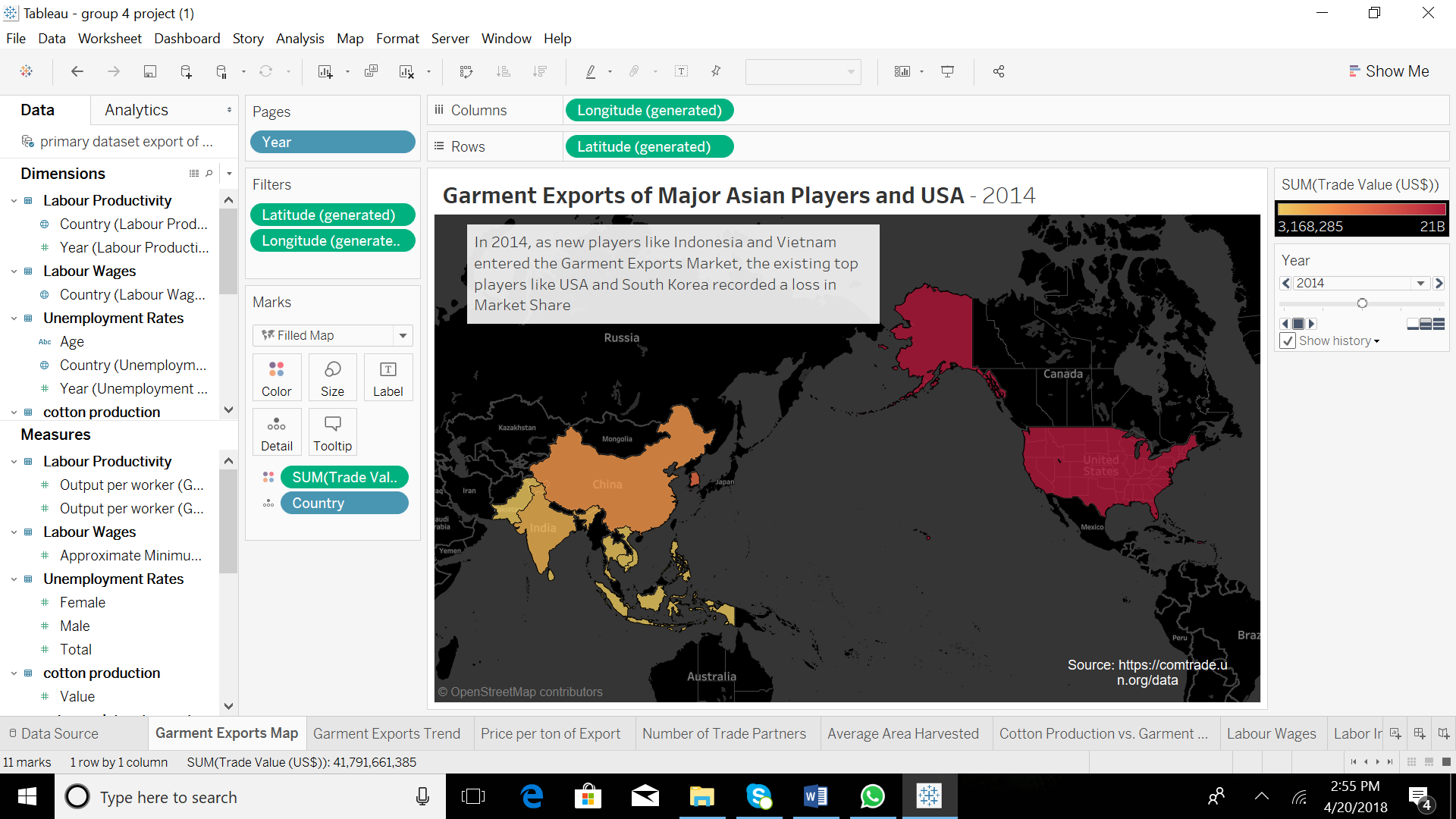
# Data Manipulation

Joins - the primary dataset has been joined with the secondary and tertiary datasets based on matching columns by country and year. The labor wages dataset has been joined with the primary dataset on column country. We have used full outer join to avoid loss of data as the primary dataset has multiple instances of unique country and year combinations

# Overview of Visualization Frames and Data Encoding Used

1. Establisher Frame:

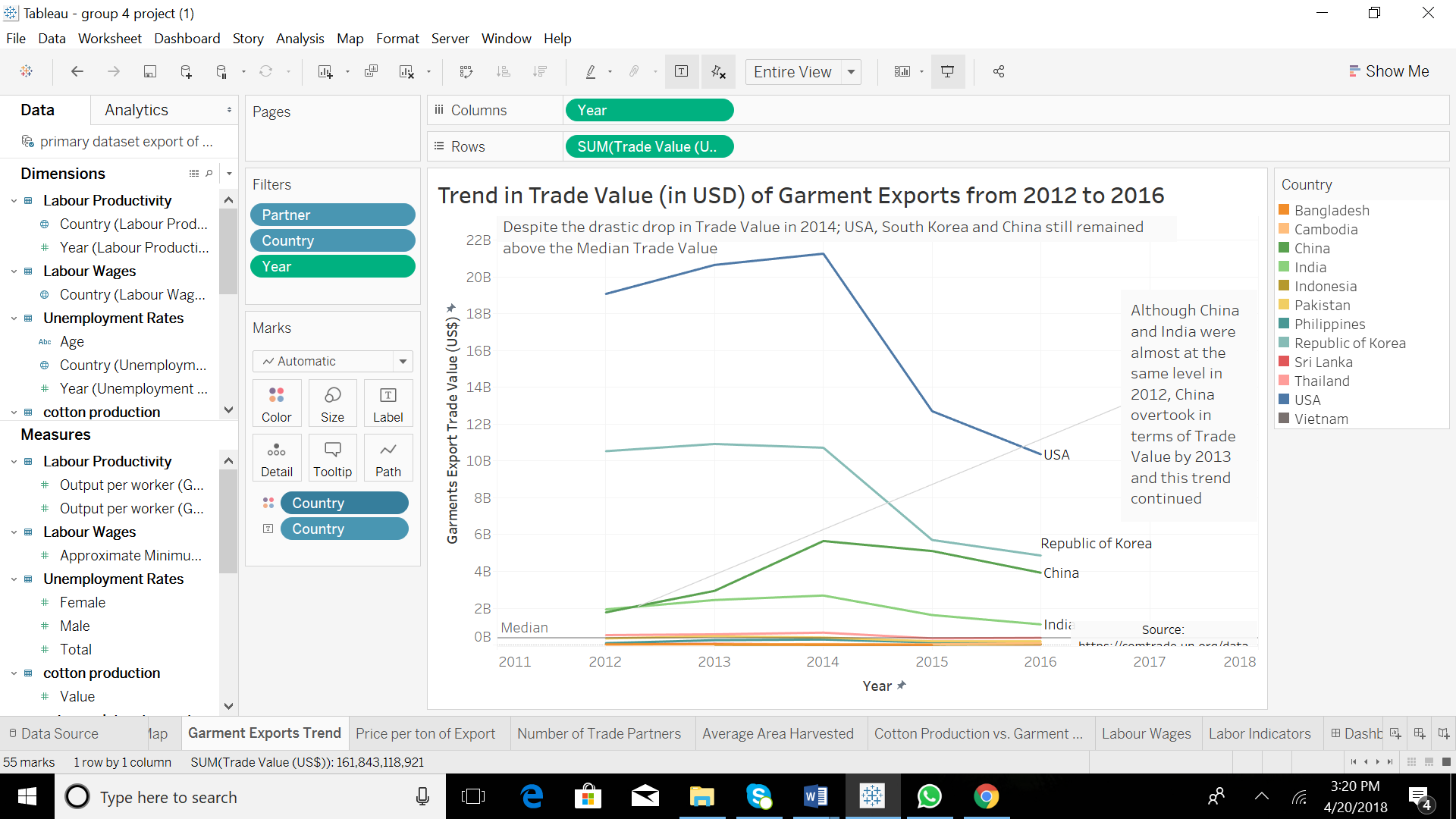
**a. Garment Exports of Major Asian Players and USA (2012 – 2016):** By using the year in play axis, the Trade Value (in US$) of Garment Exports is tracked for each country under consideration from 2012 – 2106 in the Garment Exports Map.



|  |  |  |
| --- | --- | --- |
| **Dimension/Measure** | **Encoding** | **Significance** |
| Latitude | Position X Rows | Used for generating Maps |
| Longitude | Position Y Columns | Used for generating Maps |
| Trade Value in US $ | Color | Darker Intensity of color encodes higher trade value |
| Country | Detail | Shown for country |

2. Rising Action Frame:

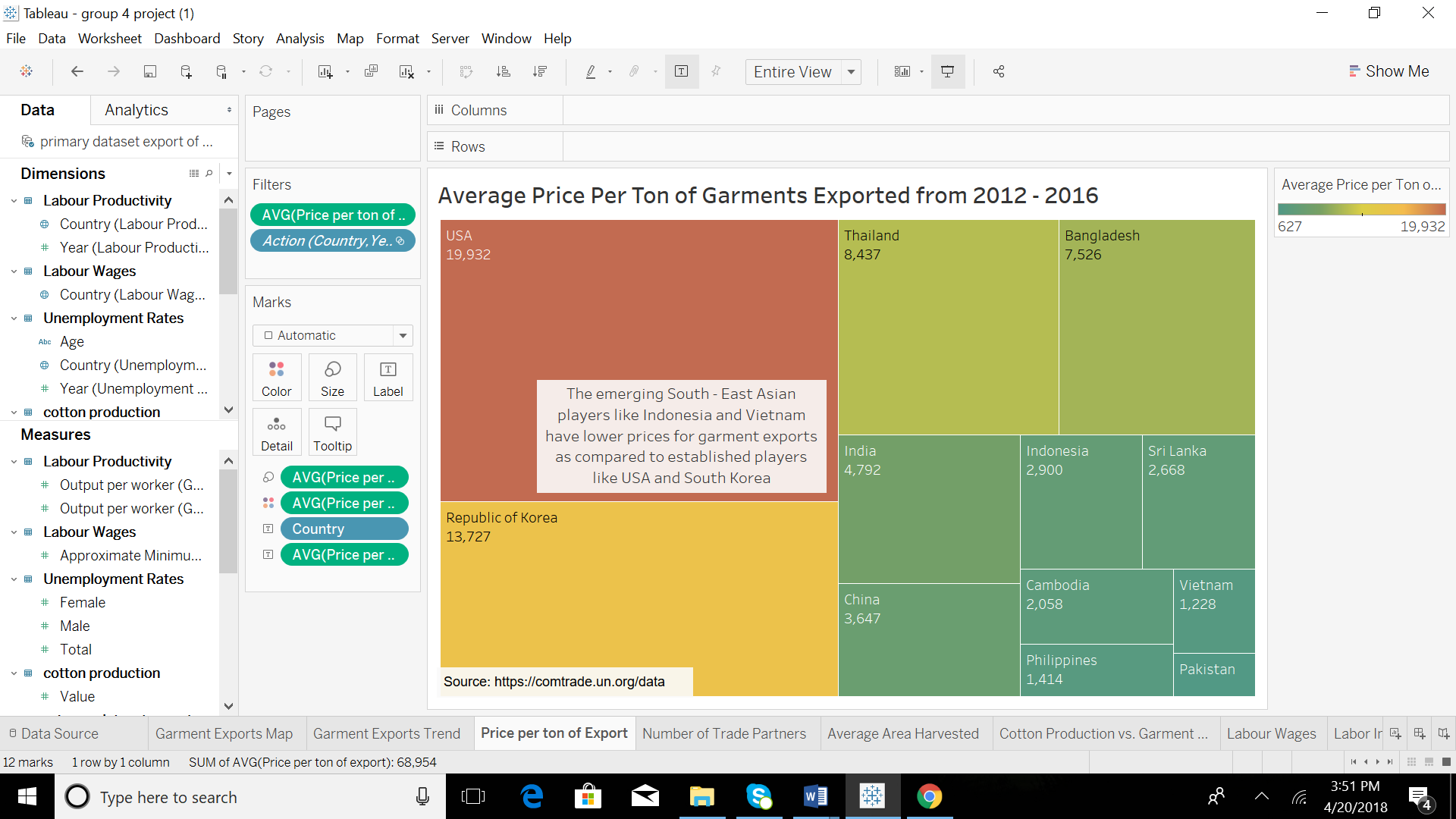
**a. Initial Frame: Trend in Trade Value (in USD) of Garment Exports from 2012 - 2016**: This line chart shows the trends in Trade Value of Garment Exports. This visualization explores the patterns in the Garment Exports market. The performance of each country in terms of Trade value over the years can be observed from this visualization.



|  |  |  |
| --- | --- | --- |
| **Dimension/Measure** | **Encoding** | **Significance** |
| Country | Color | Used to distinguish the countries that exports garments |

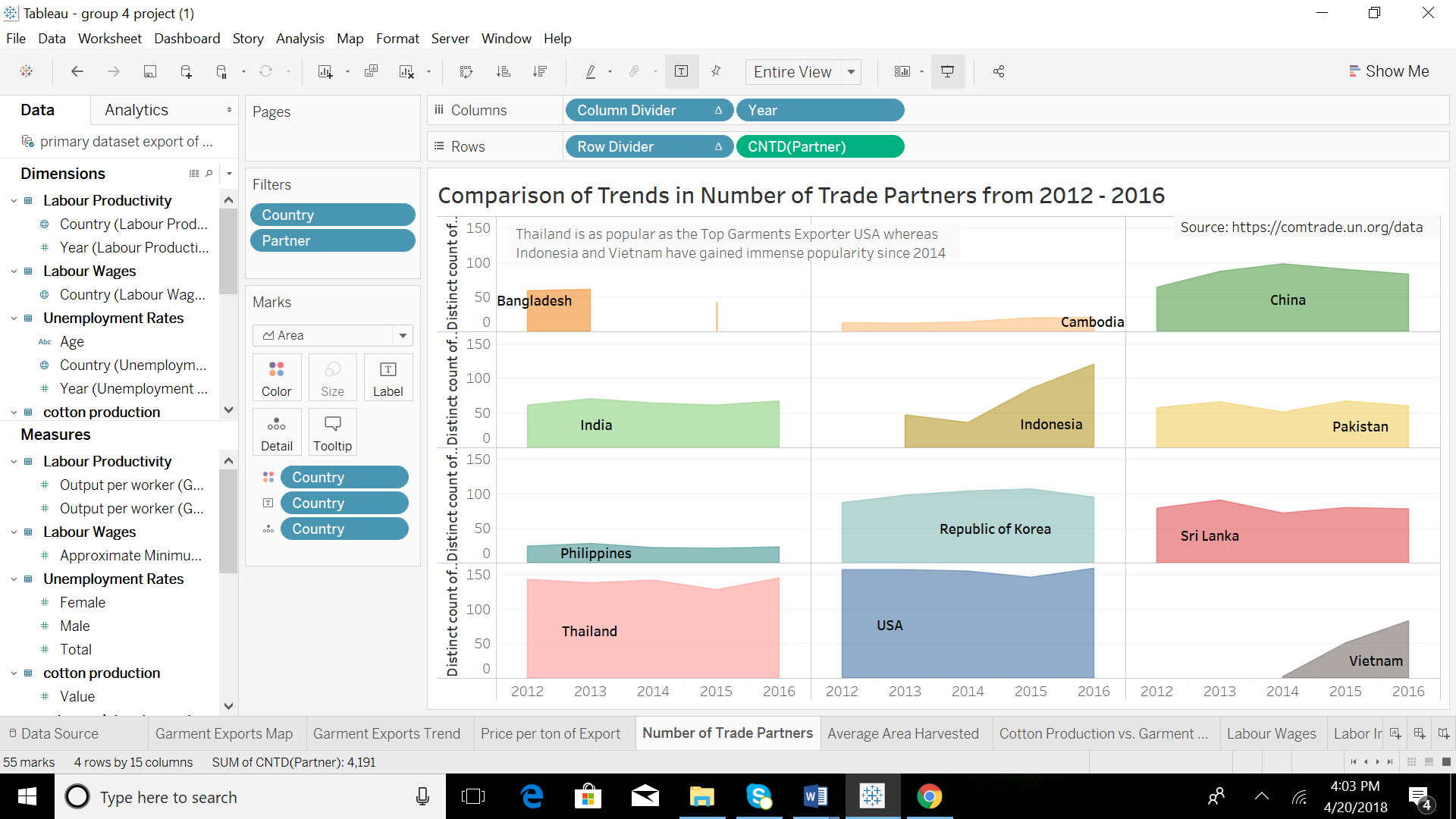
3. Peak Frames:

**a. Conflict 1: Average price Per Ton of Garments Exported from 2012 – 2016:** This visualization compares the average price per ton of garments. By using the calculated measure, Price per ton of export, which is calculated by using the formula: Trade value in USD / (Netweight in kg/1000), the countries are analyzed from a price point of view.



|  |  |  |
| --- | --- | --- |
| **Dimension/Measure** | **Encoding** | **Significance** |
| Cost per ton of Export | Filters | Used for filtering Null values if any |
| Cost per ton of Export | Color | Darker Intensity of color encodes higher |
| Cost per ton of Export | Size | Larger Size of the country in the tree map encodes Cost per ton of Export in the country |

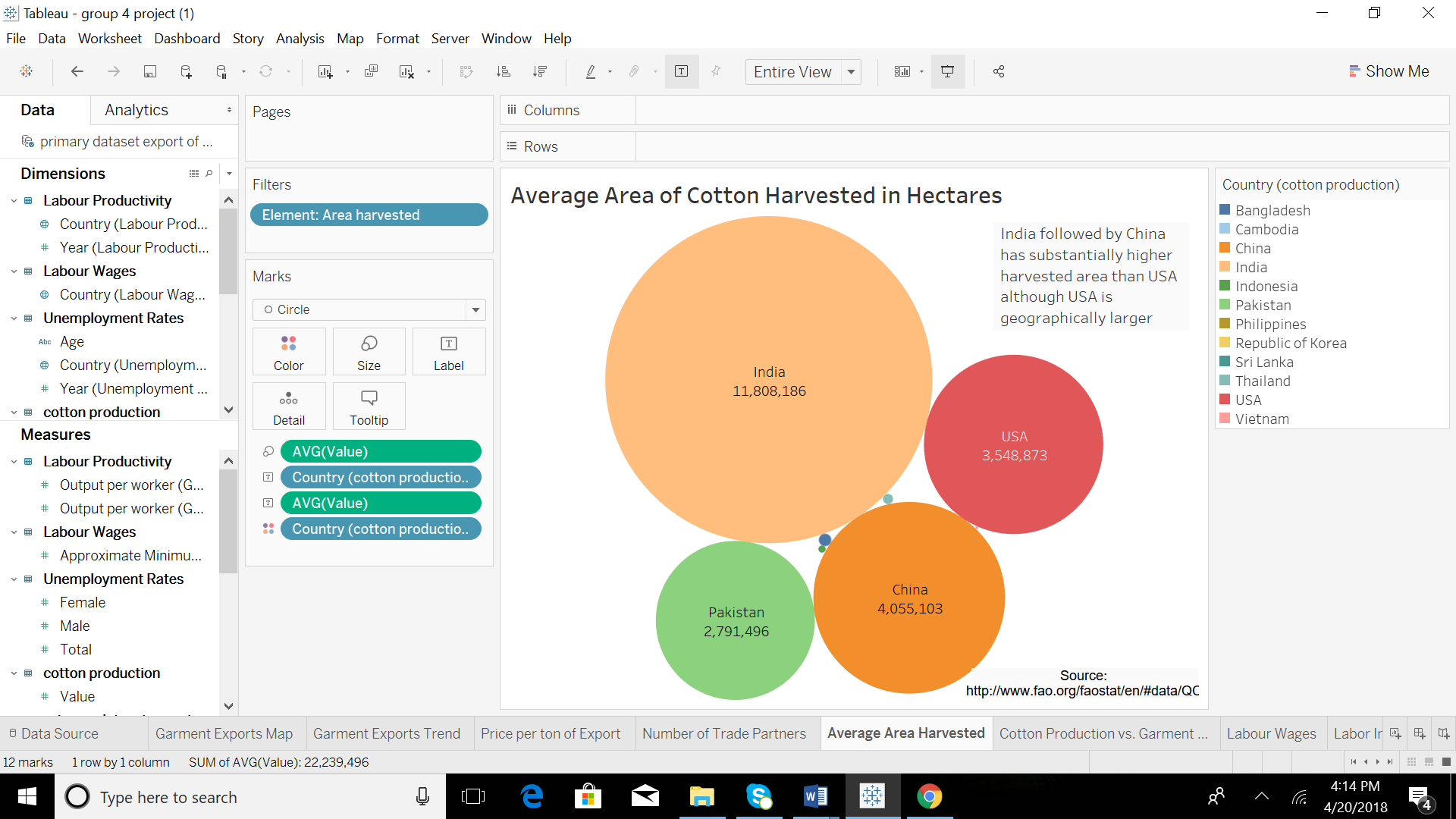
**b. Conflict 2: Comparison of Trends in Number of Trade Partners form 2012 – 2016:** This small multiple supports a comparative analysis of trends in number of distinct Trade Partners for each country in the Garments Export Industry.



|  |  |  |
| --- | --- | --- |
| **Dimension/Measure** | **Encoding** | **Significance** |
| Country | Color | Used to distinguish the countries that exports garments from 2012 -2016 |
| Country | Detail | Displays the countries and their corresponding  export partners |

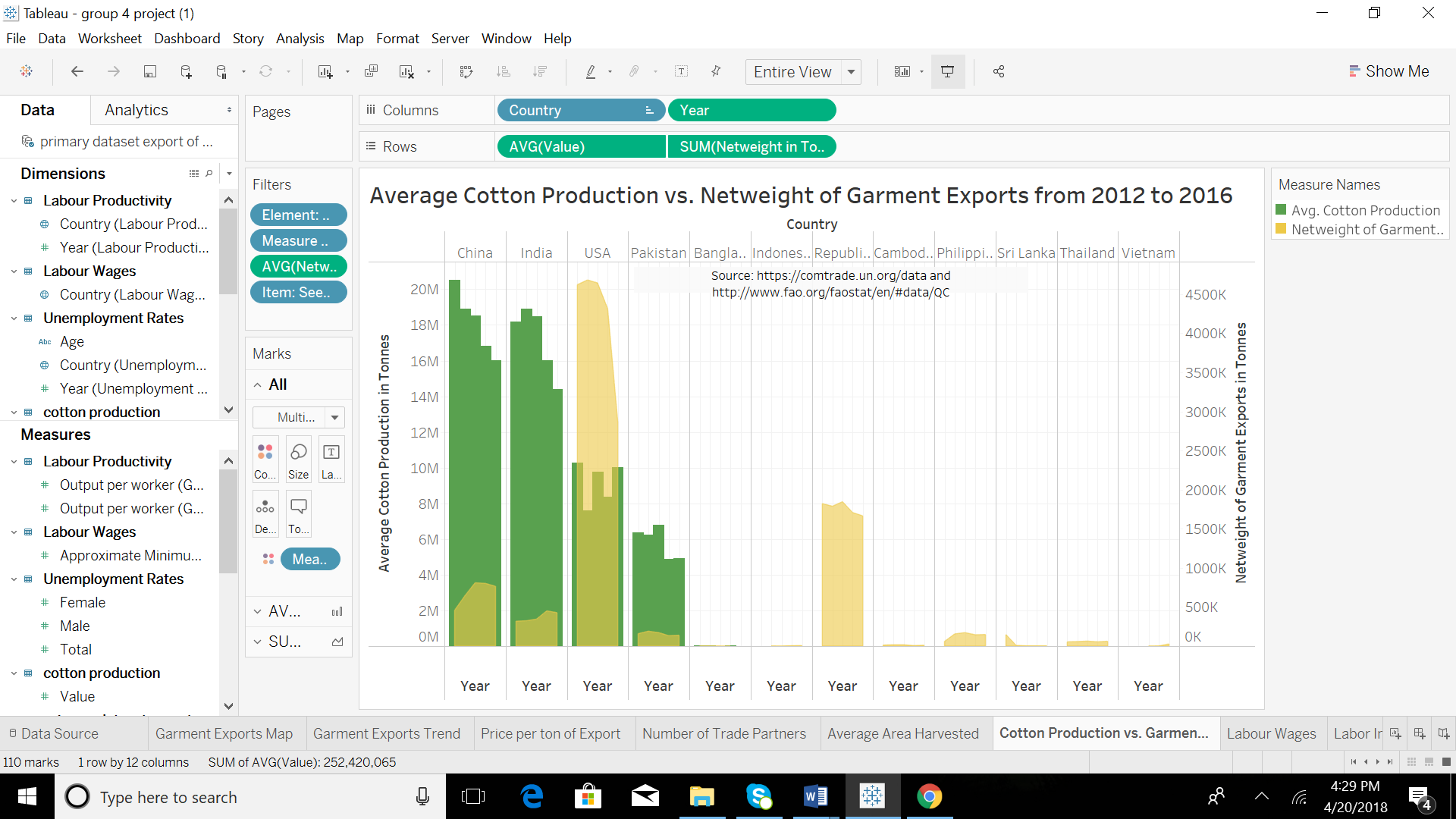
4. Falling Action Frames:

**a. Release 1: Average Area of Cotton harvested in Hectares:** This bubble chart shows the average area of cotton harvested area using size encoding. This visualization shows the area utilized for the production of the primary input (cotton) for garments industry in each country under consideration.



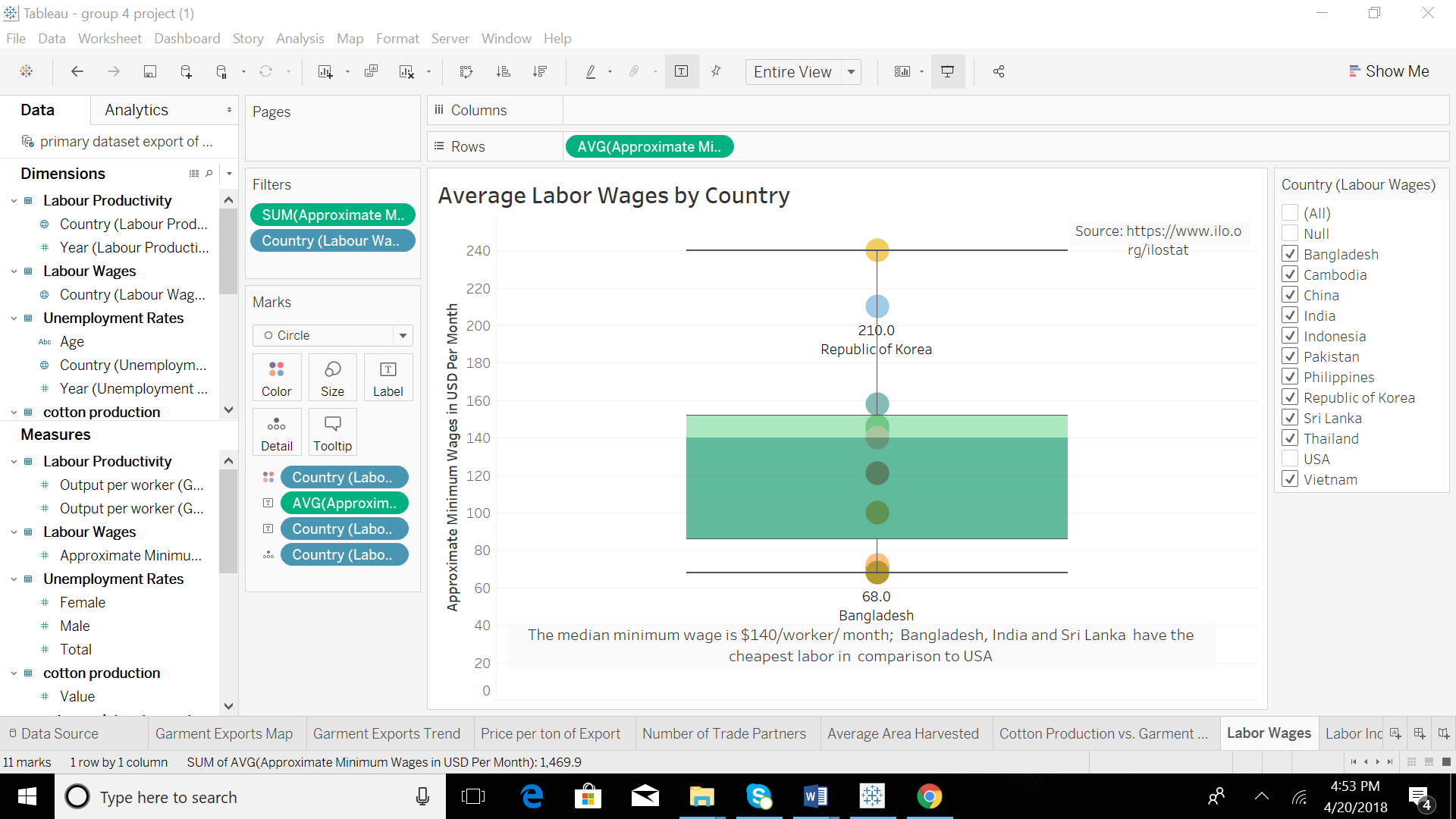
|  |  |  |
| --- | --- | --- |
| **Dimension/Measure** | **Encoding** | **Significance** |
| Country | Size | Larger size indicates larger area of cotton harvested in Hectares |
| Country | Color | Used to distinguish the countries that harvests cotton |

**b. Release 2: Average Cotton Production vs. Net weight of Garment Exports from 2012-2016:** This dual axis chart has Average Cotton Production for each country is represented by green bar chart and areachart represents the Netweight of Garment Exports over the years. The relationship between the production of cotton (major input) and netweight exported is being explored in this visualization.



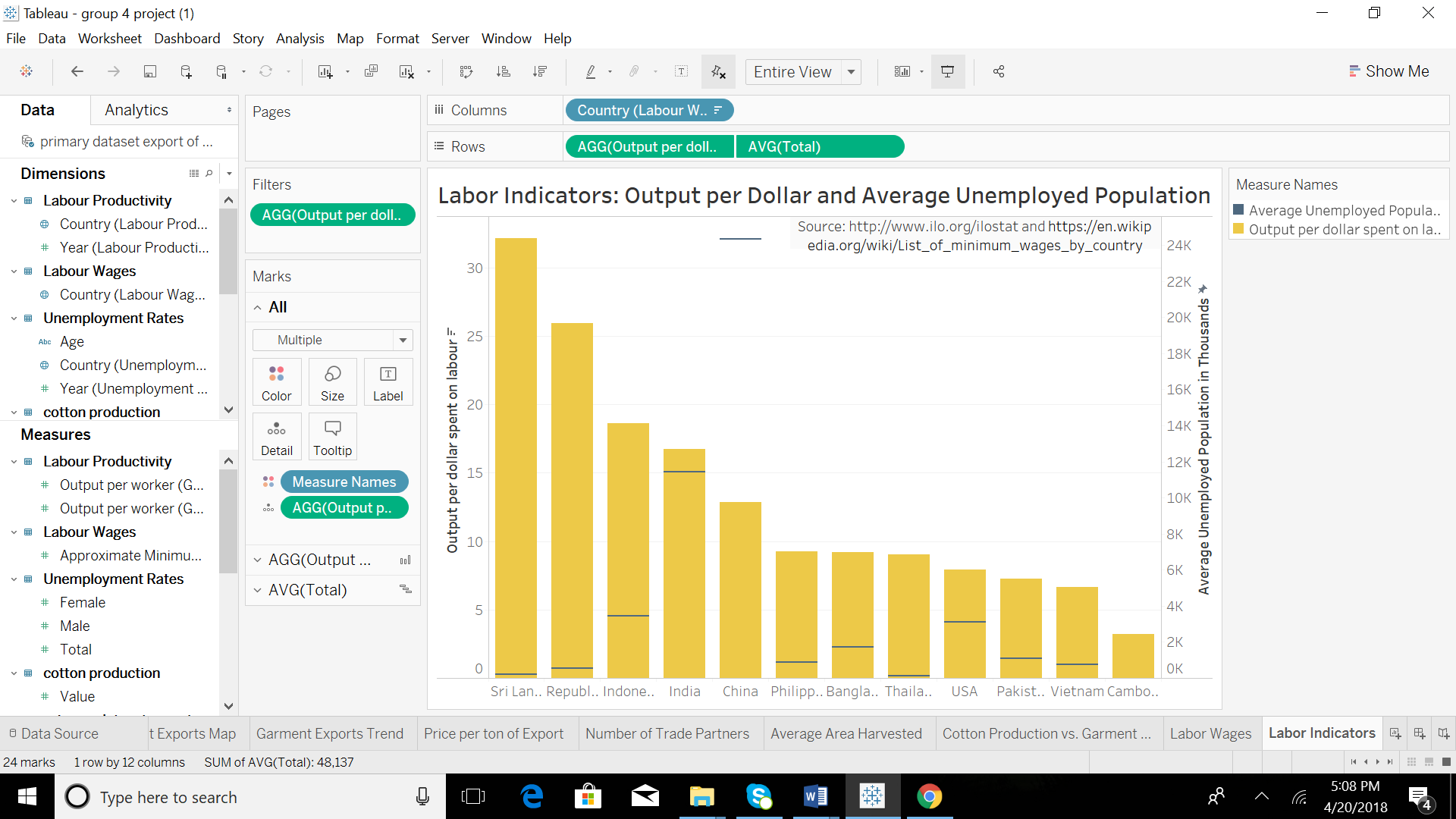
|  |  |  |
| --- | --- | --- |
| **Dimension/Measure** | **Encoding** | **Significance** |
| Value | Position X Rows | Average value of cotton production in Tonnes |
| Country | Position Y Columns | Used to represent cotton production in different countries |
| Value | Color | Average value of cotton production in Tonnes indicated with darker intensity of color |
| Net weight | Color | Sum of net weight of exports in Tonnes indicated with lighter intensity of color |

**c. Release 3: Average Labor Wages by Country:** This box plot shows the distribution average minimum labor wages for the years under consideration for each country. This visualization aids to analyze the countries with cheap labor that will in turn help achieve low cost of production for garments.



|  |  |  |
| --- | --- | --- |
| **Dimension/Measure** | **Encoding** | **Significance** |
| Country | Color | Used to distinguish the countries by their labor wages |
| Country | Detail | Detailed view of countries and their average labor wages in the box plot |

**d. Release 4: Labor indicators: Output per Dollar and Average Unemployed Population:** This dual axis chart shows the Output per Dollar spent and Average Unemployed Population in thousands for each country. These labor indicators are being analyzed to examine the labor potential of each country under consideration.



|  |  |  |
| --- | --- | --- |
| **Dimension/Measure** | **Encoding** | **Significance** |
| Output Per Dollar spent on Labor | Position X Rows | Average of Output per Dollar |
| Country | Position Y Columns | Used to represent countries |
| Output Per Dollar spent on Labor | Color | Used for coloring the bar charts representing Output Per Dollar spent on Labor in each country |
| Unemployed Population | Detail | Aggregation of the average rate of unemployed population in each country |

# Explanation of Analysis

Small Multiples

We have used small multiples to create a group of charts/graphs that share the same axes and scales. This allows the user to compare trends across dimensions in a single view.

The small multiples represent a descriptive view of data and is used to answer question such as “Which country is the largest exporter of cotton garments?” . Also, it provides a comparative view by answering the question “Who among the given countries is the largest exporter of cotton garments in the years 2102-2016?”.

A country is considered the largest garment exporter if it has the highest count (number)of trade partner countries.Hence, we have used the Count Distinct measure on the column named “Partner”

Our analysis includes 12 countries and we have displayed the visualizations of the number of trade partners for 12 countries in a 3\*4 matrix format.Row divider and Column Divider measures have been created to achieve this outcome

**Row Divider**  
int((index()-1)/(round(sqrt(size()))))

**Column Divider**  
(index()-1) %(round(sqrt(size())))

The trends such as the rise or fall in the number of trade partners for a country over the given years is clearly evident in the given charts

# Calculated Measures:

**Worksheet: Labor Wages**

|  |  |
| --- | --- |
| Measure | Formula |
| Approximate annual wage | Approximate minimum wages in USD per month \* 12 |

**Worksheet: Number of Trade Partners**

|  |  |
| --- | --- |
| Measure | Formula |
| Column divider | (index()-1) %(round(sqrt(size()))) |
| Row divider | int((index()-1/(round(sqrt(size())))) |

**Worksheet: Price per Ton of Export**

|  |  |
| --- | --- |
| Measure | Formula |
| Price per ton of export | trade value in USD / (netweight in kg/1000) |
| Row divider | int((index()-1/(round(sqrt(size())))) |

**Worksheet: Cotton Production vs. Garment Exports**

|  |  |
| --- | --- |
| Measure | Formula |
| Netweight in tonnes | netweight in kg/1000 |

**Worksheet: Labor Indicators**

|  |  |
| --- | --- |
| Measure | Formula |
| Output per dollar spent on labor | avg(output per worker in ppp)/ avg( approximate minimum wages in USD per month \*12) |

# Insights:

* From 2014 onwards, as new players like Indonesia and Vietnam entered the Garment Exports Market, the existing top players like USA and South Korea recorded a loss in Trade Value. The trade value of USA, the largest exporter in business reduced by 46% from 2012 to 2016.
* Despite the drastic drop in Trade Value in 2014; USA, South Korea and China still remained way above the Median Trade Value of US$ 359,529,930. Although China and India were almost at the same level in 2012, China overtook in terms of Trade Value by 2013 and this trend continued. In 2016, the Trade value of China was almost 3.5 times of that of India.
* The emerging South - East Asian players like Indonesia and Vietnam have lower prices for garment exports as compared to established players like USA and South Korea. The average price per ton of garments of USA, the largest exporter is almost 7 times the price of exports from the emerging players in the garment exports industry.
* Thailand is as popular as the Top Garments Exporter USA with a difference of just 14 client countries as of 2016 whereas Indonesia and Vietnam have gained immense popularity since 2014 with a clientele of almost 100 countries as of 2016.
* India followed by China has substantially higher harvested area than USA although USA is geographically larger. Cotton harvested area in India is 3 times that of USA.
* Despite comparatively low cotton production, USA and South Korea are the top two exporters by Netweight of Garments exported.
* The median minimum wage is $140 per worker per month; Bangladesh, India and Sri Lanka have the cheapest labor in comparison to USA. The minimum wages in these Asian countries is way below the median.
* Sri Lanka and South Korea have the highest output per dollar spent on labor whereas China and India have bigger labor pools because of persistent unemployment rates. The output per dollar of Sri Lanka is 4 times that of USA. China has almost 3000 times the Average Unemployed population of USA.

Data Story: Comparative Analysis of Key Asian Players and USA in the Garments Export Market from 2012 – 2016

**Setting of the Story:** The Garments Export Market and its key players (Bangladesh, Cambodia, China, India, Pakistan, South Korea, Philippines, Indonesia, Sri Lanka, Thailand, Vietnam and USA) during the period of 2012 – 2016.

**Characters of the Story:** Key Asian Players (Bangladesh, Cambodia, China, India, Pakistan, South Korea, Philippines, Indonesia, Sri Lanka, Thailand, Vietnam) and USA

**Storyline used:** Freytag's pyramid has been used for the narrative. The plot of our story consists of five parts:

* Exposition: A brief introduction
* Rising Action: The conflict central to the plot is introduced
* Climax: The turning point of the story
* Falling Action: Moving towards resolution of the conflict
* Resolution: Conflicts are concluded and outcomes are revealed

**Rhetoric used:** A logical rhetoric has been used for the analysis under consideration. The trends and patterns in the garment exports industry have been used to make suitable recommendations. Logic or reason has been used to perform the comparative analysis of Countries under consideration.

**Narrative:**

**1. Exposition:**

In 2012, USA had the highest trade value at almost $20 billion. In 2013, Indonesia became a significant market player. In 2014, Vietnam entered the market. From 2014 onwards Indonesia and Vietnam continued to grow. By 2016 Indonesia doubled its trade value as compared to 2014 whereas the trade value of USA had declined by almost 50% to $10 billion. Thus, the two Asian players – Indonesia and Vietnam became significant exporters in the time frame of 2014 – 2016 and USA experienced decline in the trade value of its garment exports.

**2. Rising Action:**

Between the years 2014 and 2015, the 4 four major players – USA, South Korea, China and India consistently remained above the median trade value of $350mn. However, USA and South Korea experienced a steep decline in trade values and India and China recorded a comparatively less steep one, but significant declines.

**3. Climax:**

When the price per ton of garment exports was compared for the countries under consideration, USA is the clear winner and nets approximate $20000/ton of export. South Korea is the next highest netter – making approximate $14000/ton of export. The middle players in this arrangement are India, China and Indonesia. Their price/ton of export is approximately 60-80% lower than that of USA. Also, when a comparison of the distinct number of trade partners for each major player across the years was conducted. To determine the strength of each country’s trade network. It shows that USA has consistently maintained the largest number of trade partners. Thailand has the next highest number of trade partners. India and China have a steady network of trade partners. Remarkably noticeable is the growth of Indonesia and Vietnam post 2014.

**4. Falling Action:**

When factors that influence garment exports such as raw material (cotton) and labor costs were analyzed. India and China have the largest cotton area harvested. USA and Pakistan hold the third and fourth position. Interestingly, the geographical sizes of the countries are in completely different in order.

Examining further, the relationship between the production of cotton and netweight exported indicates that South Korea’s garment industry is completely independent of homegrown cotton production. The USA has low dependence on home-grown cotton whereas India and China have cotton production that far outstrips their garment exports requirements.

Another major industry influencer is the cost of labor. USA at $11.60 minimum wages per month is a hefty outlier when compared to other countries. Countries like India and Bangladesh are well below the median minimum wages whereas countries like Indonesia and China are moderately placed. Thailand and South Korea are well above the median and have high labor costs. Thus, a drastic differentiation emerges based on the labor wages.

When the relationship between the most important labor metrics output per dollar spent

and average unemployed population in each country was examined. Sri Lanka records the highest output per dollar. Indonesia, India and China are among the top performers here as well. China has the highest unemployed population and so provides a big pool of unskilled labor.

**5. Resolution:**

Taking into consideration the performance of the countries from 2012 – 2014, their cotton production (raw material) and labor metrics; Indonesia, China and India are the recommended exporters in the Garments industry. India and China are established players in the Garment Exports market. They produce ample cotton to satisfy the future demand. India and China also form part of the top ranks in the labor indexes. Indonesia even with low cotton production has highly favorable labor indexes. Additionally, it has displayed fast paced and consistent growth in the market. Thailand and South Korea lose due to high labor costs, while the other Asian players perform poorly in labor metrics. Thus, rising popularity of Asian countries and declining market share of USA can be explained by our analysis.