

Visualization of Indian Trade Statistics

CSE3020 – DATA VISUALISATION

PROJECT BASED COMPONENT REPORT

By

DEVA DHARSHAN D

20BCE0282

School of Computer Science and Engineering



April 2022

DECLARATI **ON**

I hereby declare that the report entitle “Visualisation of Indian Trade Statistics” submitted by me, for the CSE3020 DATA VISUALISATION (EPJ) to VIT is a record of bonafide work carried out by me under the supervision of Dr.S.VENGADESWARAN .

I further declare that the work reported in this report has not been submitted and will not be submitted, either in part or in full, for any other courses in this institute or any other institute or university.

Place: Vellore

Date: 29 April 2022

Signature of the Candidate

CONTENTS

	P.No
1. ABSTRACT	4
2. INTRODUCTION TO THE PROJECT	
• OBJECTIVE	4
• PROBLEM STATEMENT	4
• FUNCTIONAL REQUIREMENTS	5
3. DATA ABSTRACTION	5
4. TASK ABSTRACTION	5
5. DASHBOARD IMPLEMENTATION	5
6. RESULT ANALYSIS	17
7. CONCLUSION	18
8. APPENDIX	18
• SCREEN SHOTS	
• SAMPLE CODING	

ABSTRACT:

Trade is a vital aspect for any country to keep its economy healthy and growing. Exports and imports done by a country are a testimony to the success of country's manufacturing power and is a critical factor in analysing the self-reliance in sustainability. This data needs to be understood by the general public in order to make sound conclusions and educate themselves about the economy in a transparent way. Therefore, visualisation emerges as a solution to comprehend these enormous trade statistics.

INTRODUCTION TO THE PROJECT:

OBJECTIVE:

The visualisation will be of great use to investors and business groups. They can understand the ongoing trade patterns in the market and therefore decide on where and what to invest and trade. The data would also grant a political insight to each country and individual state in India. The amount of trade undertook from a particular region is a reflection of the economic policies of the government and also a probable parameter to estimate standard of living of people. Hence the visualisation presents a useful case study for social scientists as well as for political analysts.

PROBLEM STATEMENT:

The aim of the project is to visualise and monitor the import and export of merchandise and services trade operations undertaken by the Government of India. The task is to break down the massive data in order to show users the relevant information. It reflects the trade analysis of our country in terms of top principal commodities, the volume of trade performed with different countries and regions of the world, contribution of each individual state and territory in India as well as yearly comparison of trade data.

FUNCTIONAL REQUIREMENTS:

The dataset for the project was collected from the official website of the Department of Commerce, Ministry of Commerce and Industry (GoI). The website can be viewed by clicking [here](#)! The dataset contains merchandise and services trade data grouped into various domains across different years and can be viewed in a couple of currencies of exchange. The dataset also gives insight to India's logistics performance indices and budget expenditures.

DATA ABSTRACTION:

The data can be represented in form of tables where the rows represent the item data and the columns represent the column attributes like trading partner country, commodity and the trade amount. The name of countries and commodities are qualitative and nominal data. The year of trade is discrete and interval type. The trade amount is ratio. The countries occupy an explicit spatial position on the world map and thus can be plotted as 2D surfaces under geometry dataset type. The trade amount can be visualized in two currencies: rupees and dollars, thus a tree dataset.

TASK ABSTRACTION:

We can consume and produce the data. We can present the data in forms of graphs, plots and maps. We can derive trade balance and trade deficit from the data. We can lookup for trade statistics of a particular item with a particular country, else we can browse and locate if the information is not known, respectively. We can compare trades with different countries and then summarize our results. We can target trends in data to understand peaks in imports/exports and we can feature a particular state or district to the interest of the user.

DASHBOARD IMPLEMENTATION:

The project uses flexdashboard for implementation. Additionally various R libraries are installed and imported for executing different visualisation plots. Some themes have

been created to provide an overall clean and aesthetic appearance to the dashboard. Therefore, various data files that have been stored and saved in the local system as

.csv files have been read in R console for successfully plotting the corresponding data in our graphs. This has been achieved using the readr library in R which provides a read_csv function.

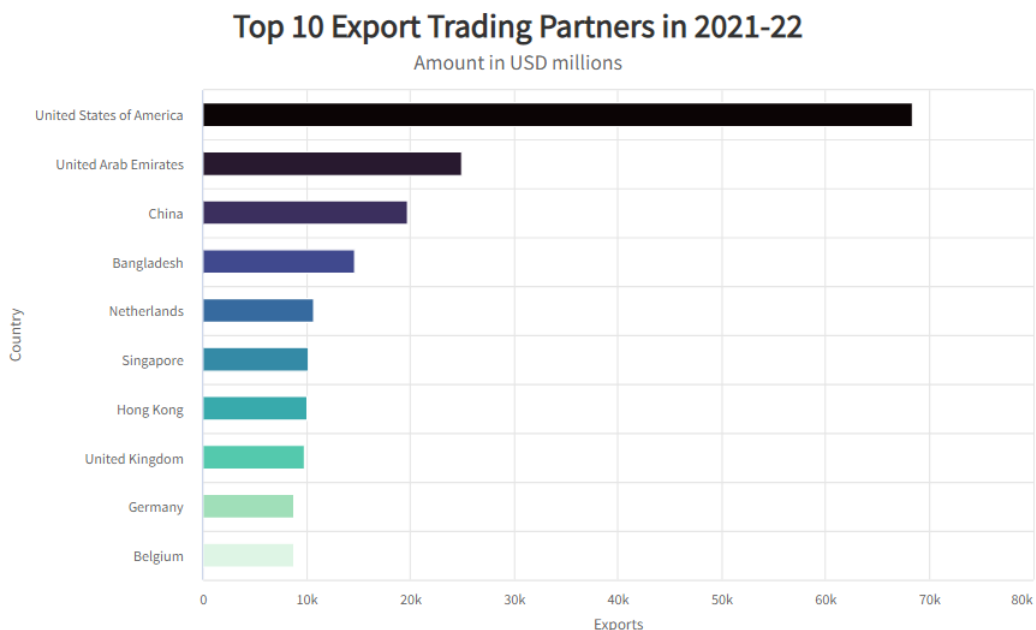
```
1  ---
2  title: "Data Visualization J Component"
3
4  output:
5
6    flexdashboard::flex_dashboard:
7      orientation: columns
8      vertical_layout: fill
9  ---
10
11  ```{r setup, include = FALSE}
12
13  library(flexdashboard)
14  library(tidyverse)
15  library(highcharter)
16  library(gt)
17  library(htmltools)
18  library(viridis)
19  library(readr)
20  library(dplyr)
21  library(plotly)
22  library(sf)
23  library(rnaturalearth)
24  library(countrycode)
25  library(ggmap)
26  library(RColorBrewer)
27  library(wordcloud)
28  library(tm)
29  library(ggplot2)
30  library(tmap)
31  library(tmaptools)
32  library(leaflet)
33  library(treemap)
34  library(DT)
35  library(plotrix)
36
37  thm <-
38    hc_theme(
39      colors = c("#1a6ecc", "#434348", "#90ed7d"),
40      chart = list(
41        backgroundColor = "transparent",
42        style = list(fontFamily = "Source Sans Pro")
43      ),
44      xAxis = list(
45        gridLineWidth = 1
46      )
47    )
48
49  custom_colors <- viridis::mako(n = 10)
50
51  custom_colors2 <- viridis::mako(n = 37)
52
53  fontStyle = list(
54    family = "Source Sans Pro",
55    size = 15,
56    color = "black"
57  )
58
59  label = list(
60    bgcolor = "#EEEEEE",
61    bordercolor = "transparent",
62    font = fontStyle
63  )
64
65  ```
66  ```{r}
67
68  country_wise <- read_csv("countrywise_data_2021_22.csv")
69  five_years <- read_csv("countrywise_data_2017_22.csv")
70  total <- read_csv("total_figures_2017_22.csv")
71  product <- read_csv("top_export_commodities_2021_22.csv")
72  commodity <- read_csv("commodities_data_2021_22.csv")
73  state_wise <- read_csv("statewise_exports_2021_22.csv")
74
75  ```
```

```

78 Column {data-width=300}
79 -----
80
81 ### Top Export Partners
82
83 ```{r}
84
85 country_wise %>%
86   group_by(Country) %>%
87   summarise(Exports) %>%
88   arrange(desc(Exports)) %>%
89   head(10) %>%
90   hchart('bar', hcaes(x = Country, y = Exports, color = custom_colors)) %>%
91   hc_add_theme(thm) %>%
92   hc_tooltip(pointFormat = '<b>Exports: </b> {point.y} <br>' ) %>%
93   hc_title(text = "Top 10 Export Trading Partners in 2021-22",
94           style = list(fontSize = '25px', fontWeight = 'bold')) %>%
95   hc_subtitle(text = "Amount in USD millions",
96              style = list(fontSize = '16px'))
97
98 ```

```

The following horizontal bar chart is generated from the above code. The dataset has been grouped according to country names and the top ten countries have been selected based on the amount of export done for financial year 2021-22. A colour channel has been added to the bars which corresponds to the amount. Thereafter a theme has been added and appropriate title and subtitle has been placed. Moreover, the plot is made interactive with a hovering property provided to each bar in the graph. Hovering over a plot would display would display the name of country and the exact export amount in USD millions. This has been plotted using the highcharter library in R.

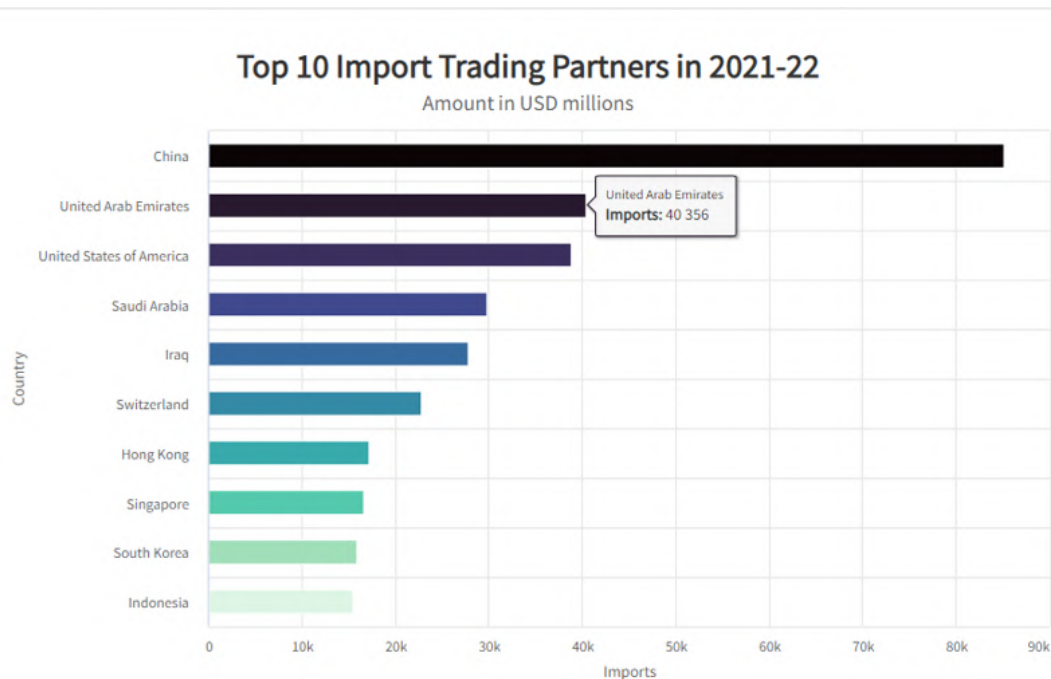


```

99
100 ▾ ### Top Import Partners
101
102 ▾ ```{r}
103
104 country_wise %>%
105   group_by(Country) %>%
106   summarise(Imports) %>%
107   arrange(desc(Imports)) %>%
108   head(10) %>%
109   hchart('bar', hcaes(x = Country, y = Imports, color = custom_colors)) %>%
110   hc_add_theme(thm) %>%
111   hc_tooltip(pointFormat = '<b>Imports: </b> {point.y} <br>' ) %>%
112   hc_title(text = "Top 10 Import Trading Partners in 2021-22",
113           style = list(fontSize = '25px', fontWeight = 'bold')) %>%
114   hc_subtitle(text = "Amount in USD millions",
115              style = list(fontSize = '16px'))
116
117 ▴ ```
118
119

```

A similar bar chart has been plotted for import data as well with the same properties.



```

120 Column {data-width=200}
121 ▾ -----
122
123 ▾ ### Exports Data as Text
124
125 ▾ ```{r}
126
127 wordcloud(c(country_wise$Country), freq = c(c(floor(country_wise$Exports))), random.color = TRUE, col =
128   brewer.pal(9, "Reds"), scale = c(2,.5), min.freq = 1)
129 ▴ ```
130
131 ▾ ### Imports Data as Text
132
133 ▾ ```{r}
134
135 wordcloud(c(country_wise$Country), freq = c(c(floor(country_wise$Imports))), random.color = TRUE, col =
136   brewer.pal(9, "Reds"), scale = c(2,.5), min.freq = 1)
137 ▴ ```
138
139

```

Next, a text visualisation is created for the same data. The values in text have been given as country names and they are repeated in accordance to the export or import amount to create a varying frequency in the text. This has been done using wordcloud

library in R and the colour scheme is taken from colour palettes available in the RColorBrewer library.

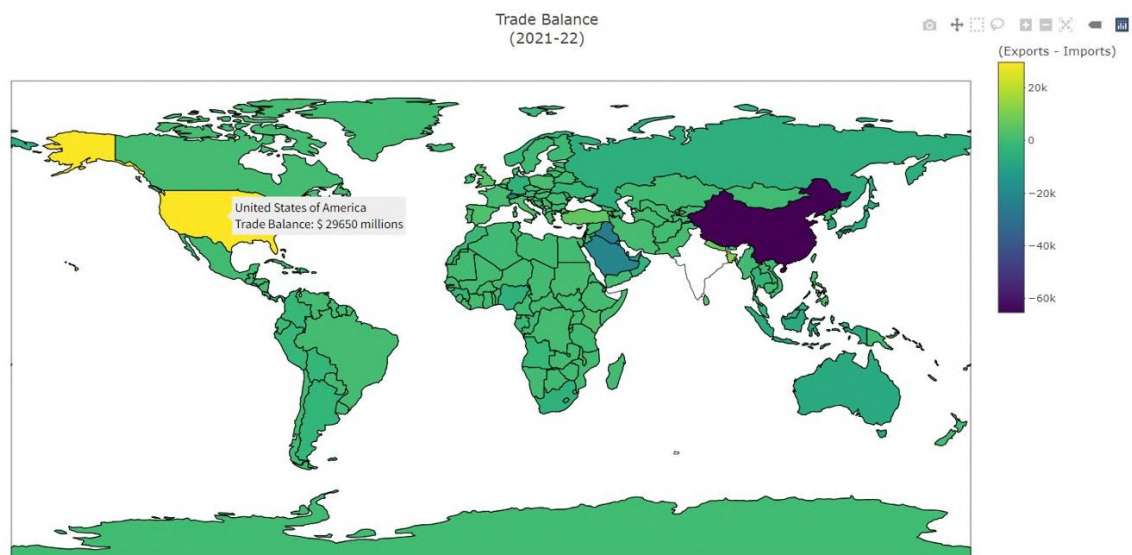


```

139
140 Column {.tabset .tabset-fade data-width=500}
141 -----
142
143 ### Trade Balance Data
144
145 ```{r}
146
147 data <- country_wise %>% mutate(hover = paste0(Country, "\nTrade Balance: $ ", (Exports-Imports), " millions"))
148
149 map <- plot_geo(data, locationmode = "ISO-3") %>% add_trace(locations = ~Code, z = ~(Exports-Imports), zmin =
min(data$Exports - data$Imports), zmax = max(data$Exports - data$Imports), color = ~(Exports - Imports), text =
~hover, hoverinfo = 'text') %>% layout(title = "Trade Balance\n(2021-22)") %>% style(hoverlabel = label)
150
151 map
152
153 ```
154
155

```

A map has been plotted using plotly library in R which provides the `plot_geo` function to plot world maps. The location mode used by the function is ISO-3. Thus, the dataset is inner joined such that the standardised three letter country code is joined with each country's name in the data. The aim of the map is to colour the countries according to the trade balance for the given financial year. The trade balance is calculated by subtracting the import amount from the export amount. The user can then visualise with which country India has a trade surplus and with which it has a trade deficit. The map is interactive as when we hover over a country in the map it shows us a text label containing the country name and the balance amount.

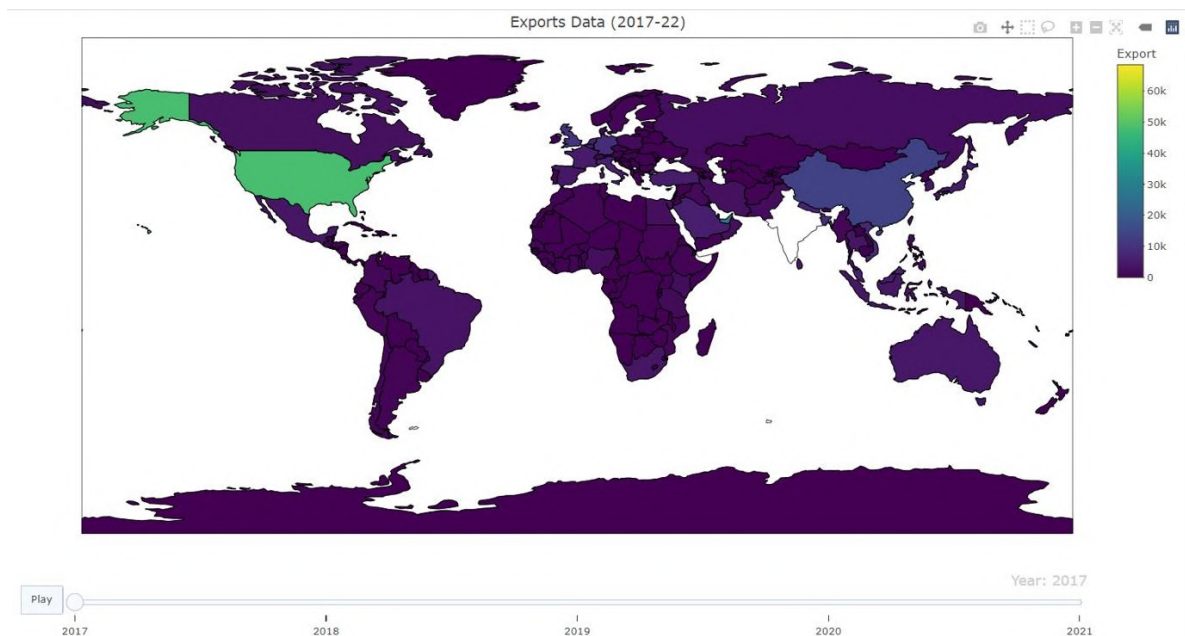


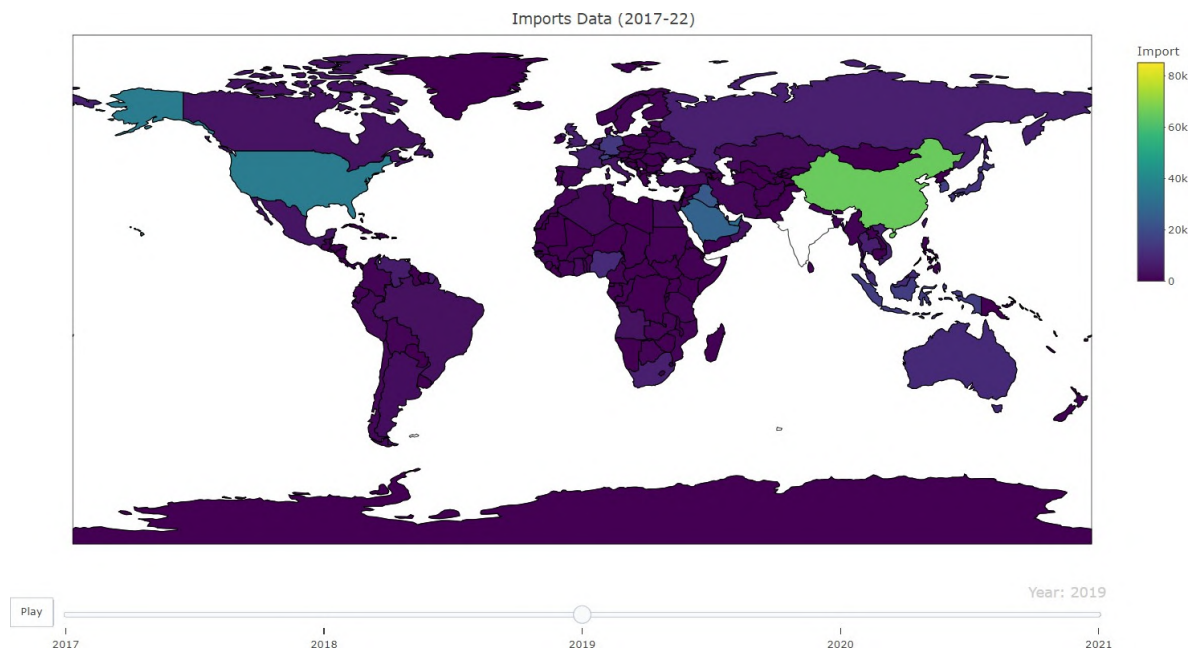
```

155
156 ▾ ### Exports Data (2017-22)
157
158 ▾ ```{r}
159
160 data <- five_years[,c("Export", "Code", "Country", "Year")] %>% mutate(hover = paste0(Country, "\n$", Export, "
millions"))
161
162 map <- plot_geo(data, locationmode = "ISO-3", frame = ~Year) %>% add_trace(locations = ~Code, z = ~Export, zmin
= 0, zmax = max(data$Export), color = ~Export, text = ~hover, hoverinfo = 'text') %>% layout(title = "Exports
Data\n(2017-22)") %>% style(hoverlabel = label)
163
164 map
165
166 ▾ ```
167
168 ▾ ### Imports Data (2017-22)
169
170 ▾ ```{r}
171
172 data <- five_years[,c("Import", "Code", "Country", "Year")] %>% mutate(hover= paste0(Country, "\n$", Import, "
millions"))
173
174 map <- plot_geo(data, locationmode = "ISO-3", frame = ~Year) %>% add_trace(locations = ~Code, z = ~Import, zmin
= 0, zmax = max(data$Import), color = ~Import, text = ~hover, hoverinfo = 'text') %>% layout(title = "Imports
Data\n(2017-22)") %>% style(hoverlabel = label)
175
176 map
177
178 ▾ ```
179

```

Again, two world maps have been plotted using the same function to generate views for exports and imports data. The difference lies that this time we are visualising the data from financial year 2017-18 to 2021-22 in the same map. A slider is provided below the map that changes the frame according to year.



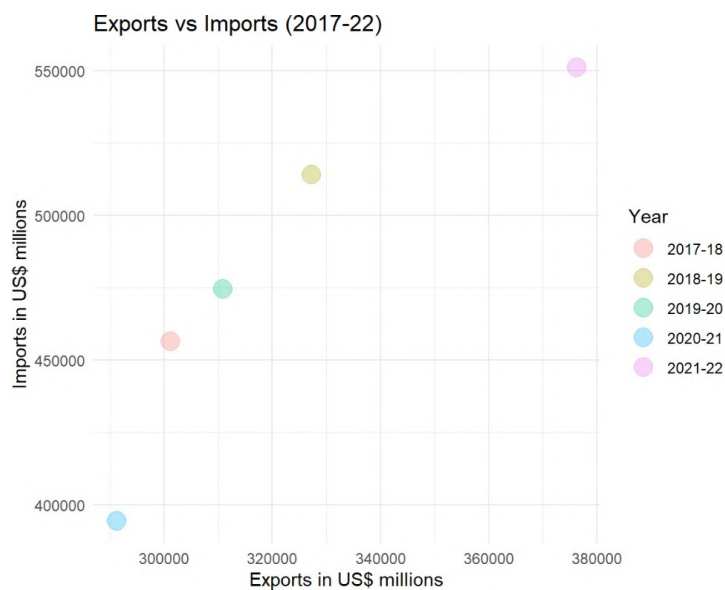


```

179
180 ### Trade Comparsion (2017-22)
181
182 {r}
183
184 ggplot(total, aes(x = Export, y = Import, color = Year)) +
185   geom_point(size = 5, alpha = 0.3) +
186   labs(x = 'Exports in US$ millions', y = 'Imports in US$ millions', title = 'Exports vs Imports (2017-22)') +
187   theme_minimal()
188
189
190

```

A scatterplot is plotted to visualise the export versus import data for last five years. A ggplot has been used for implementation. Colour channel has been used on point marks to display different years.

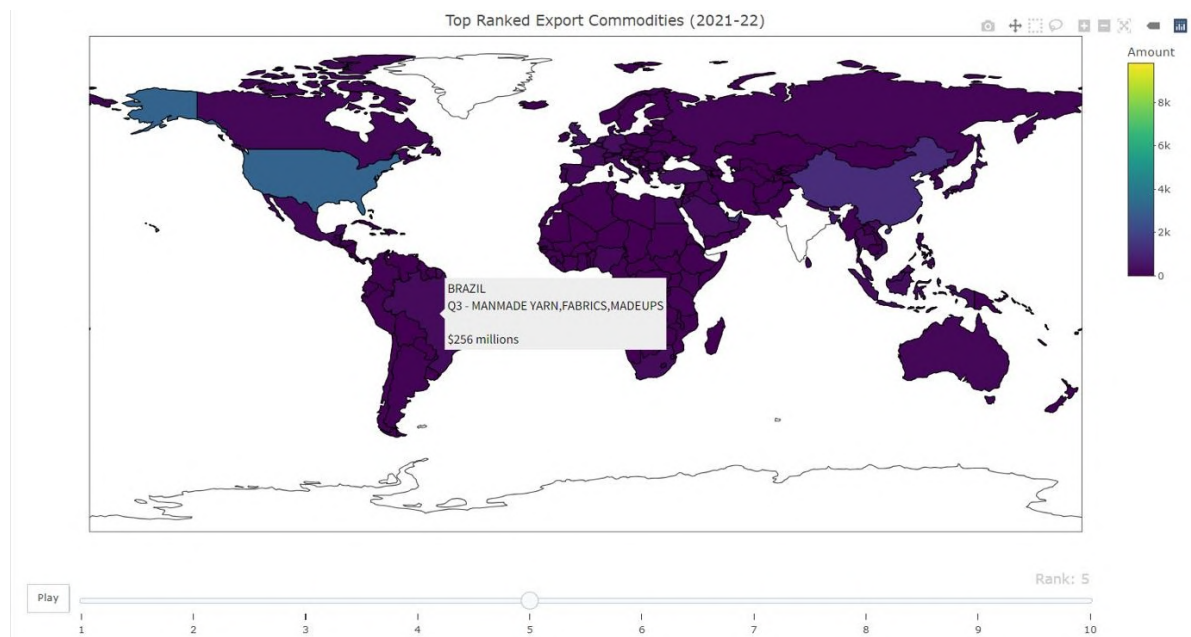



```

191 ▾ ### Top Exports
192
193 ▾ ```{r}
194
195 data <- product[,c("Product", "Code", "Country", "Rank", "Amount")] %>% mutate(hover = paste0(Country, "\n",
196   Product, "\n", "\n$", Amount, " millions"))
197
198 map <- plot_geo(data, locationmode = "ISO-3", frame = ~Rank) %>% add_trace(locations = ~Code, z = ~Amount, zmin
199   = 0, zmax = max(data$Amount), color = ~Amount, text = ~hover, hoverinfo = 'text') %>% layout(title = "Top Ranked
200   Export Commodities\n(2021-22)") %>% style(hoverlabel = label)
201 ▴ ```

```

Next, we have again plotted a world map. This time the aim is to visualise the top ten products or commodities we have exported to that country in this year. The slider in the bottom changes the frame from rank one to rank ten.

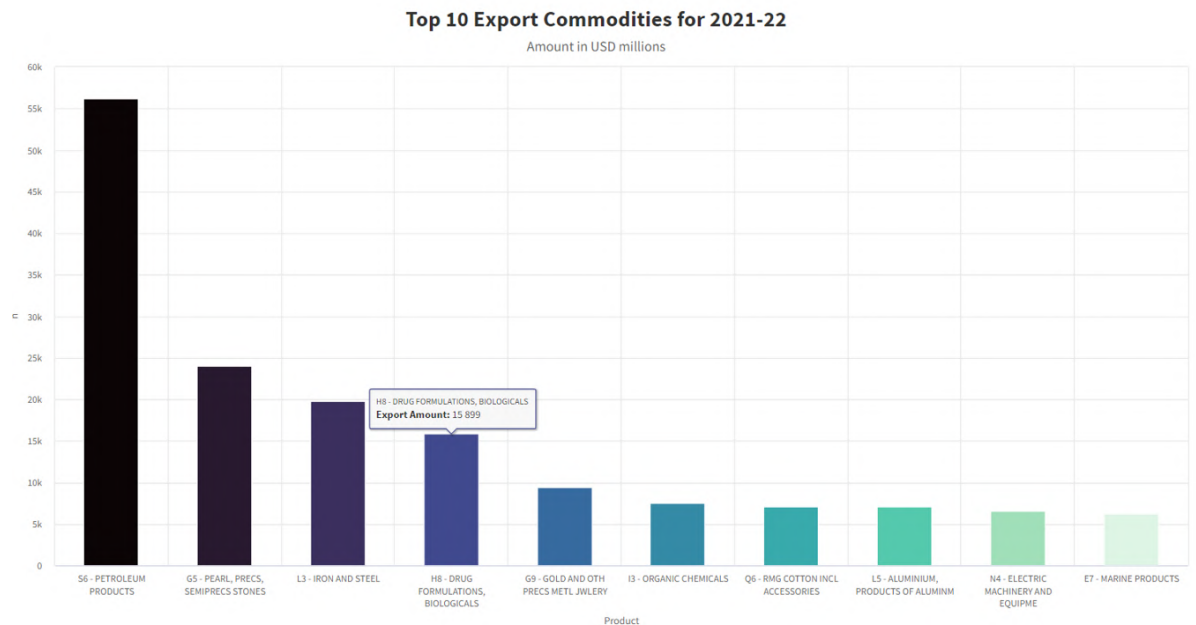


```

203 ▾ ### Top Export Commodities
204
205 ▾ ```{r}
206
207 data <- product %>% group_by(Product) %>% summarize(n = sum(Amount)) %>% arrange(desc(n)) %>% slice(1:10,)
208
209 data %>%
210   hchart('column', hcaes(x = Product, y = n, color = custom_colors)) %>%
211   hc_add_theme(thm) %>%
212   hc_tooltip(pointFormat = '<b>Export Amount: </b> {point.y} <br>' ) %>%
213   hc_title(text = "Top 10 Export Commodities for 2021-22",
214     style = list(fontSize = '25px', fontweight = 'bold')) %>%
215   hc_subtitle(text = "Amount in USD millions",
216     style = list(fontSize = '16px'))
217
218 ▴ ```
219

```

Again, a bar chart has been implemented to visualise which are the top ten commodities overall for export in 2021-22. The difference to be noted here is the argument column which has been used to plot vertical chart whereas the argument bar is used to plot horizontal chart.



```

220 ### Data Table of Commodities (2021-22)
221
222 ```{r}
223
224 data <- commodity %>% group_by(Product) %>% arrange(desc(Amount)) %>% summarise(Country, Amount)
225
226 datatable(data, filter = 'top', options = list(pageLength = 15), class = 'hover cell-border stripe')
227
228 ```
229

```

A data table has been plotted using the datatable function in DT library of R. The purpose of the data is to understand that for which commodity which country is the top export partner and the corresponding amount is displayed alongside. The table contains more than 2000 entries and at one moment a maximum of 15 entries are displayed on the page to respect the dimensions of webpage. Each column in the data table is sortable and we can also search for a particular commodity or a country using the search box which presents us with a drop-down menu.

For instance, in the below snippet, we can view that for tea (A1 is the product code that has been used for international trade as standardised by the World Trade Organisation), the top exporters are Iran, Russia, Kazakhstan, Niger, Azerbaijan and Kyrgyzstan followed by Tajikistan. The value of export is also shown in USD millions.

Show **15** entries Search:

	Product	Country	Amount
	All	All	All
1	A1 - TEA	IRAN	99
2	A1 - TEA	RUSSIA	75
3	A1 - TEA	KAZAKHSTAN	14
4	A1 - TEA	NIGER	3
5	A1 - TEA	AZERBAIJAN	1
6	A1 - TEA	KYRGYZSTAN	1
7	A1 - TEA	TAJIKISTAN	1
8	A2 - COFFEE	JORDAN	45
9	A2 - COFFEE	LIBYA	19
10	A2 - COFFEE	UKRAINE	15
11	A2 - COFFEE	TUNISIA	13
12	A2 - COFFEE	SYRIA	10
13	A2 - COFFEE	BENIN	7
14	A2 - COFFEE	CROATIA	7
15	A2 - COFFEE	MAURITANIA	6

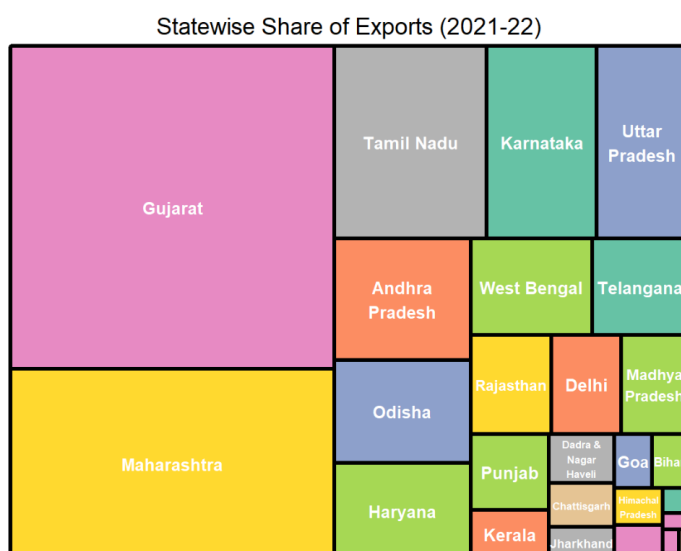
Showing 1 to 15 of 2,085 entries Previous **1** 2 3 4 5 ... 139 Next

```

230 ### State Wise Treemap
231
232 {r}
233
234 treemap(state_wise, index = "States", vSize = "Amount", type = "index", fontsize.labels = 10, fontcolor.labels =
  "white", fontface.labels = 2, bg.labels = "transparent", align.labels =c("center","center"), border.col =
  "black", border.lwds = 3, title = "Statewise Share of Exports (2021-22)", fontsize.title = 14, palette = "Set2")
235
236
237

```

The next element in our dashboard is a treemap which gives us the information about export performed by each state and union territory in India during the given financial year. The size of box is equivalent to the value of exports done and labels with states names have been provided for clear understanding.

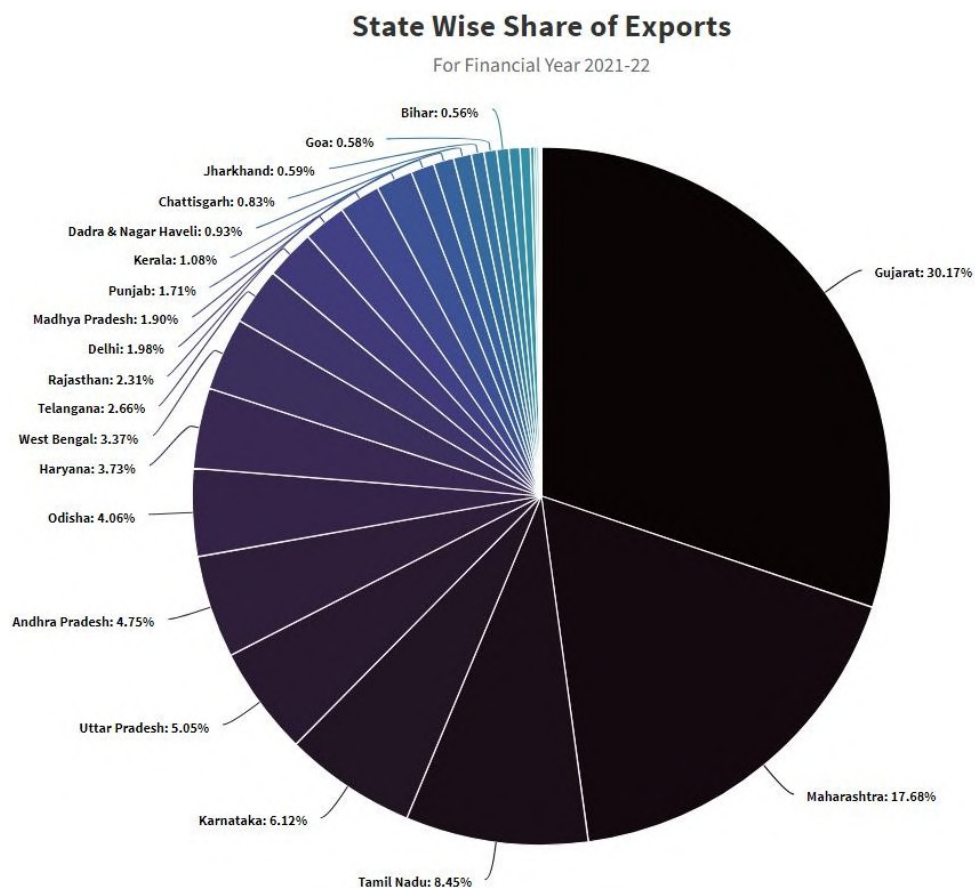


```

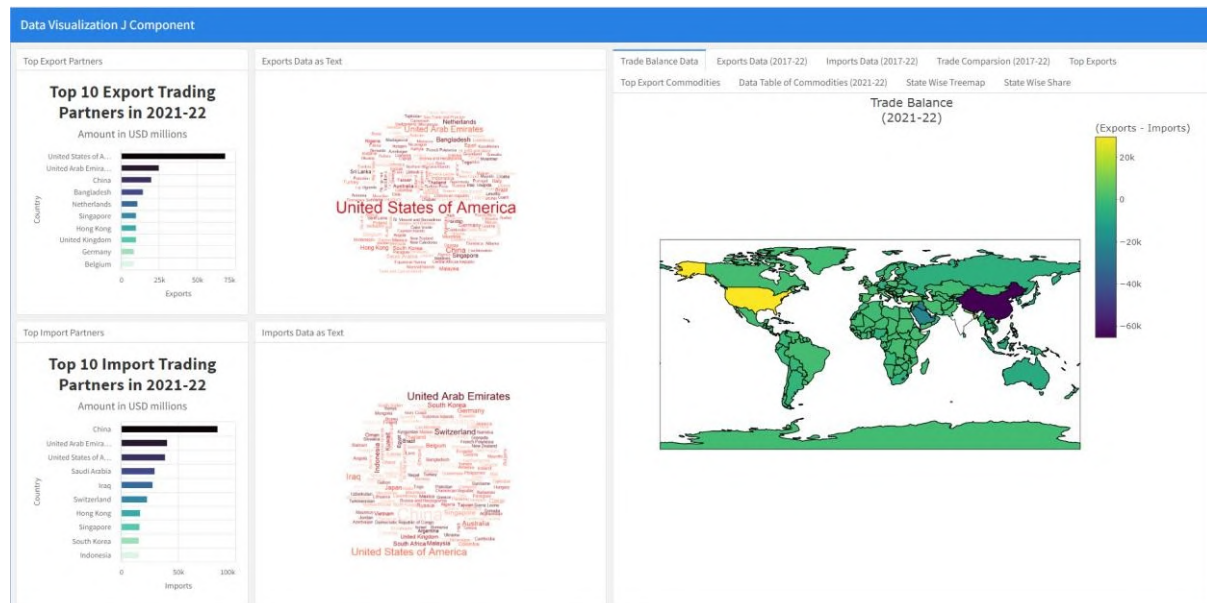
238- ### State Wise Share
239
240- ```{r}
241
242- pie_data <- state_wise[,c("States", "Amount")]
243
244
245- pie_inter <- pie_data %>% hchart("pie", hcaes(x = States, y = Amount, color = custom_colors2), dataLabels =
246-   list(enabled = TRUE, format = '{point.name}: {point.percentage:.2f}%')) %>% hc_title(text = "State Wise Share of
247-   Exports", style = list(fontSize = '25px', fontWeight = 'bold')) %>% hc_add_theme(thm) %>%
248-   hc_tooltip(pointFormat = 'Amount: {point.y} CR RS' ) %>%
249-   hc_subtitle(text = "For Financial Year 2021-22",
250-     style = list(fontSize = '16px'))
251
252- pie_inter
253- ```

```

Lastly, a pie chart has been plotted using highcharter library. The pie is divided into sectors corresponding to share of each state and territory in exports. The chart is made interactive as when we hover over a sector, it is highlighted and the related information is reflected in a label.



The following dashboard will emerge as a result of execution of the above code chunks. The first two bar plots and two text visualisations have been assigned to two separate columns while the other visualisations are put in the third column. The tab switch feature has been provided in that column to switch views and navigate to other plots on the same screen.



RESULT ANALYSIS:

Various analysis can be drawn from the plotted visualisations such as in regard to India's top trading partners and top trade commodities. A major conclusion which can be drawn is that we were able to understand the effect of pandemic on international trade. The scatterplot gave a glimpse of how the data was at its lowest in 2020-21 during the shown five-year period and how the country performed in coping up with the pandemic with respect to the trade volume carried out in next financial year. The treemap and pie chart can also be analysed to understand which states have the most favourable policies for economy and trade. One can also comprehend that effect of sea in trade which can be understood from the fact that the top states involved in exports were largely coastal states and this implies their access to ports for trade.

CONCLUSION:

The required dashboard was successfully implemented using R and provides a number of different plots to satisfy the aim of our project and visualisation.

APPENDIX:

SCREENSHOTS:

The relevant screenshots can be accessed by clicking [here](#)!

SAMPLE CODING:

The code of the project can be viewed by clicking [here](#)!