SEA FOOD TRADE OF US

**The report encompasses the graphics created using Tableau**

**Data exploration using R Studio and Tableau**

Contents

[EXECUTIVE SUMMARY 2](#_Toc64036353)

[PURPOSE OF THE PROJECT 3](#_Toc64036354)

[DATA 3](#_Toc64036355)

[About the data 3](#_Toc64036356)

[Data Preprocessing 3](#_Toc64036357)

[Preprocessing of Export data 3](#_Toc64036358)

[Preprocessing of Import data 1](#_Toc64036359)

[DASHBOARD 1](#_Toc64036360)

[INSIGHTS FROM DATA 2](#_Toc64036361)

# EXECUTIVE SUMMARY

The purpose of this project is to create visuals using Tableau software. The data is preprocessed in RStudio to gain insights into the trade pattern of US in aquaculture.

# PURPOSE OF THE PROJECT

This study has the following purpose:

* Showcase the flexibility of Tableau to create Geospatial visualizations
* Trend analysis in Tableau

# DATA

## About the data

The datasets used for this analysis are “Aquaculture\_Aquaculture\_Exports.csv” and “Aquaculture\_Aquaculture\_Imports.csv” which contains the information the seafood export and import of US to the rest of the word from 1989 to 2016. Each data has 3540083 rows and 10 columns, containing information like destination of Import / export, Month of import/ export year of transaction, name of the commodity traded, quantity traded, and the value of Import/ export in US$. These datasets were taken from Kaggle.

## Data Preprocessing

The closer inspection revealed that, these dataset have information of the quantity traded in KG, export / import value in USD all recorded in single column named “Attribute\_DESC” .Also for specific location of order, geography numeric codes (like 1,2,3..) . In addition to this, there is another redundant column namely Unit\_Desc which measures the unit of the attribute (if the attribute is import/ export value, “Unit\_DESC” shows “US$”. Even though, the data have no missing values, it was not fully organized and hence it was necessary to straighten the data prior to any serious analysis.

### Preprocessing of Export data

Further probe into the data revealed too many inconsistencies in the data. While majority of them are corrected in excel, a few complex corrections were made in R. The coloum COMMODITY\_DESC shows the type of good traded. It contains some valuable yet haphazardly arranged information regarding types of fish, whether it is fresh or frozen, variety of the fish etc. It is mostly in this format*: “atlantic salmon (salmo salar) fillet/oth meat,fr/ch”.* The coloum showed 173 unique items. The first and foremost step in the analysis was to identify the information from this column and categorize them accordingly.

After conducting some background study, it was decided to reformat original information as Category-Type of export-Subcategory. For instance, the above-mentioned information is reformatted as “*Salmon-Cooked-Atlantic Salmon*”. So now, from this coloum we get to know the general category of the fish we are trading, whether it is cooked/fresh/frozen and if there is a specific category of fish. The data is then exported to RStudio, and the column COMMODITY\_DESC is duplicated, and the duplicated column is split into 3 columns: Category, Type of export and Subcategory, using stringer package. Other small inconsistencies are corrected in R

Now, we have 9 broad categories of the seafood export namely:

* "TROUT”
* “SALMON"
* "SHRIMPS AND PRAWNS"
* "OYSTERS"
* "MUSSELS"
* "CLAMS"
* "SCALLOPS"
* "ORNAMENTAL FISH"
* "SHRIMPS

The sub-categories are:

* NA
* "RAINBOW TROUT"
* "ATLANTIC SALMON"
* "DANUBE SALMON"
* "CHINOOK/KING SALMON"
* "CHUM/DOG SALMON"
* "PINK HUMPIE SALMON"
* "SOCKEYE/RED SALMON"
* "COHO/SILVER SALMON"
* "ROE"
* "PACIFIC SALMON"
* "OTHER"
* "SHELL ON"
* "PEELED"
* "SEED OYSTERS"
* "GEODUCK CLAMS"
* "GENERAL"
* "FRESHWATER"

The NA values are changed to “Not given” for the analysis.

The 6 different types of export are

* FRESH
* FROZEN
* COOKED
* CANNED
* SALTED
* LIVE

### Preprocessing of Import data

The Import data is preprocessed in the same way as the Export data. The unprocessed Import data had 243 unique values which are reformatted in the similar fashion as the Export data. After preprocessing, we have 10 categories, 6 Types of export and 22 Subcategories.

The 6 different types of export are

* FRESH
* FROZEN
* COOKED
* CANNED
* SALTED
* LIVE

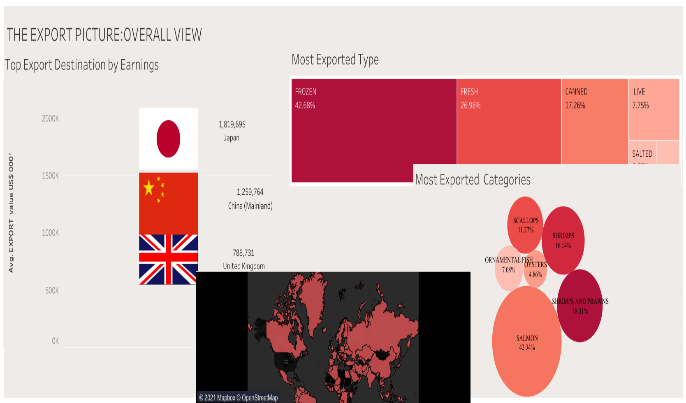
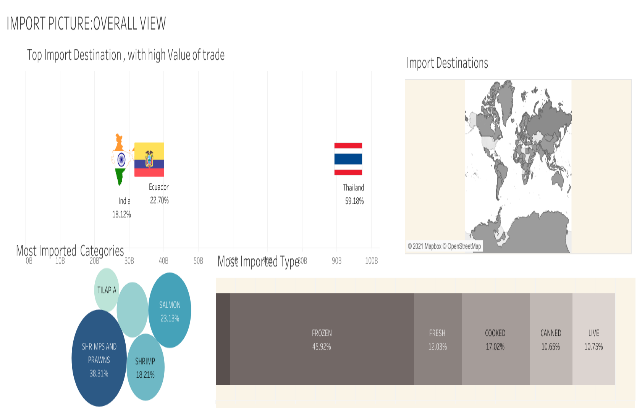
The categories are:

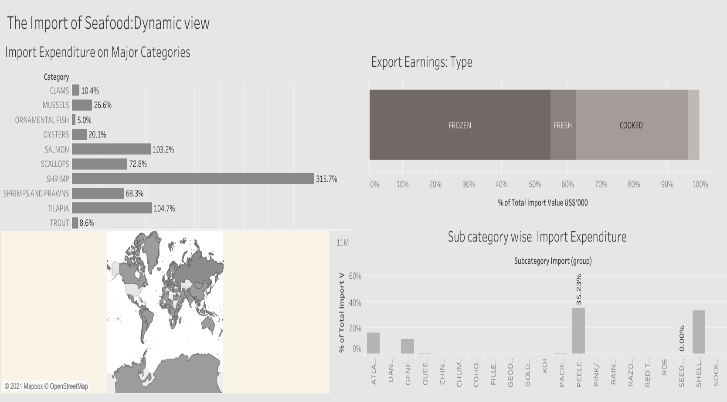
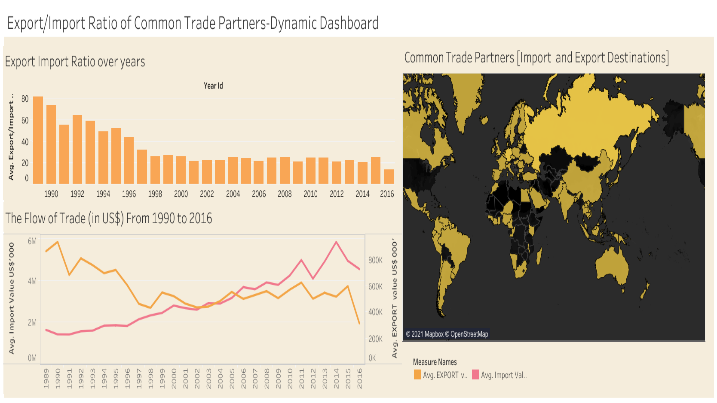
|  |
| --- |
| * TROUT |
| * SALMON |
| * TILAPIA |
| * SHRIMPS AND PRAWNS |
| * OYSTERS |
| * MUSSELS |
| * CLAMS |
| * SCALLOPS |
| * ORNAMENTAL FISH |
| * SHRIMP |
|  |

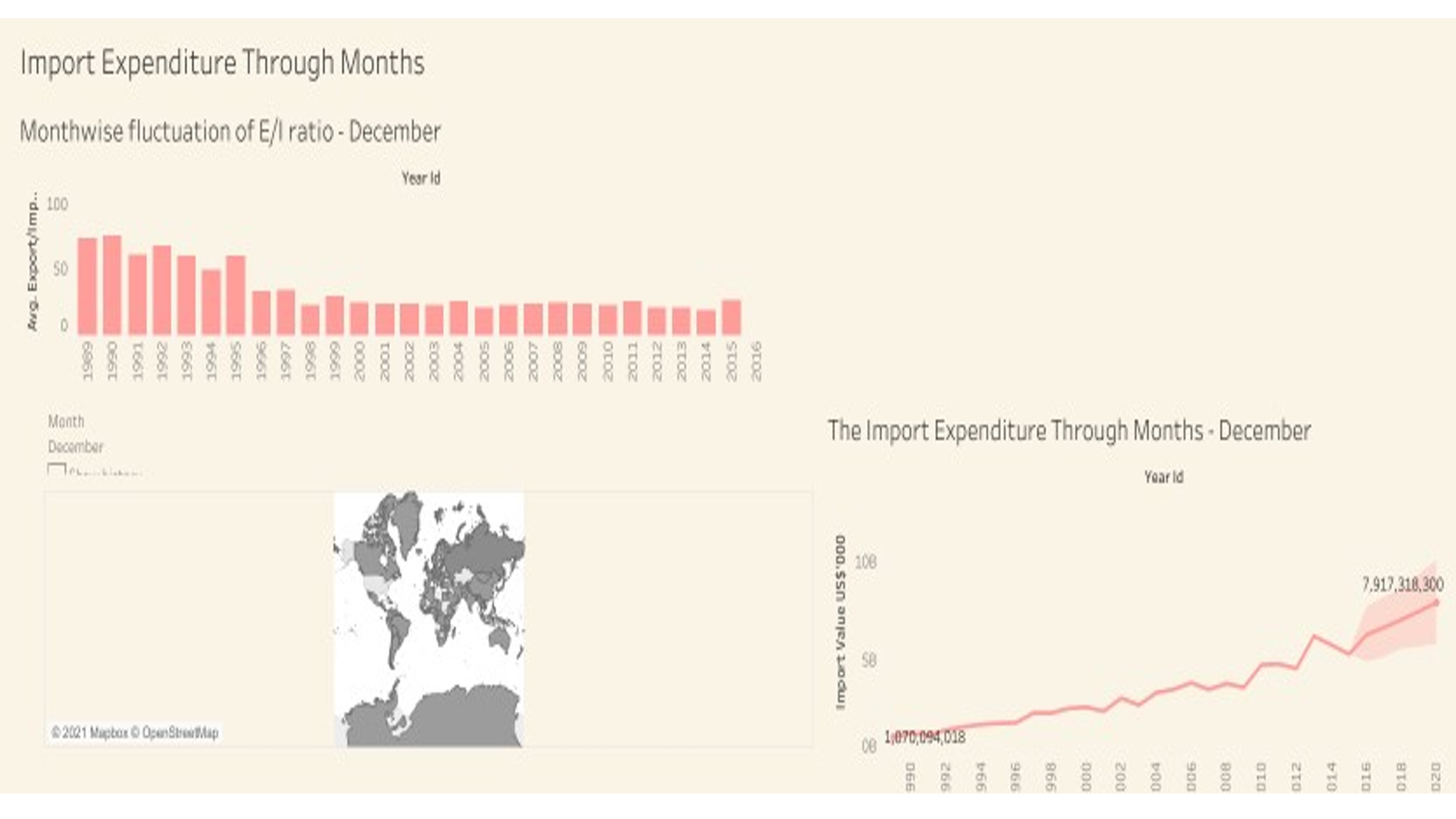
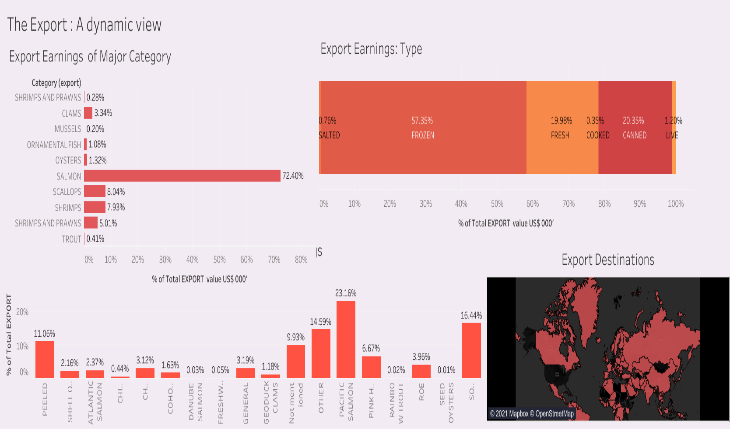
The Subcategories are:

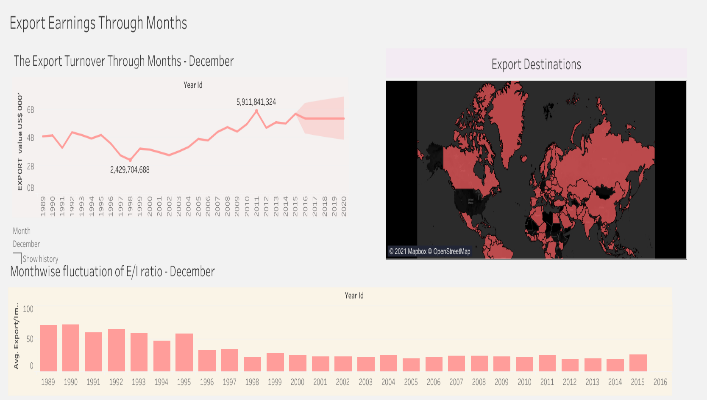
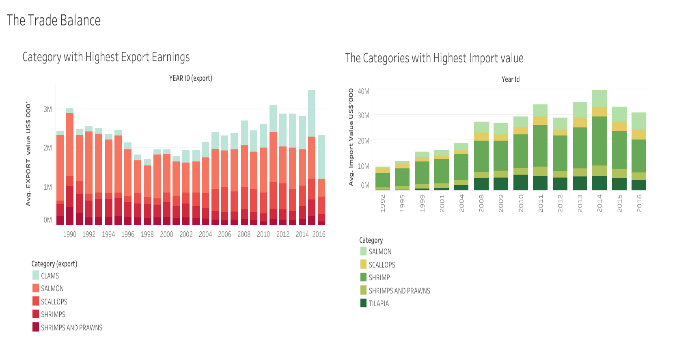
|  |
| --- |
| * GENERAL |
| * RAINBOW TROUT |
| * ATLANTIC SALMON |
| * DANUBE SALMON |
| * ROE |
| * CHINOOK /KING SALMON |
| * CHUM/DOG SALMON |
| * PINK/HUMPIE SALMON |
| * SOCKEYE/RED SALMON |
| * COHO /SILVER SALMON |
| * PACIFIC SALMON |
| * SHELL ON |
| * PEELED |
| * SEED OYSTERS |
| * GEODUCK CLAMS |
| * STIMSONS CLAMS |
| * RAZOR CLAMS |
| * QUEEN SCALLOPS |
| * KOI |
| * GOLDFISH |
| * FILLETS |
| * RED TILAPIA |

# DASHBOARD





*Please click on this link to view the dynamic dashboards:* <https://public.tableau.com/profile/neena.shereen.prathap>

# INSIGHTS FROM DATA

* The most prominent export of US is Salmon, within which the major item is Frozen Pacific Salmon.
* Majority of the export earnings were coming from China, Japan, and UK.
* The main items of Import export are Shrimp and Prawns.
* Frozen-peeled Shrimps are a major import item.
* US have bilateral trade relations with almost all countries.
* Over the years, Clams is slowly gaining importance as a highest grossing export of US; same goes to Tilapia as an item of Import trade.
* India, Ecuador, and Thailand are the destinations with high value Import trade.
* The Export Import gap is more pronounced before 2004.