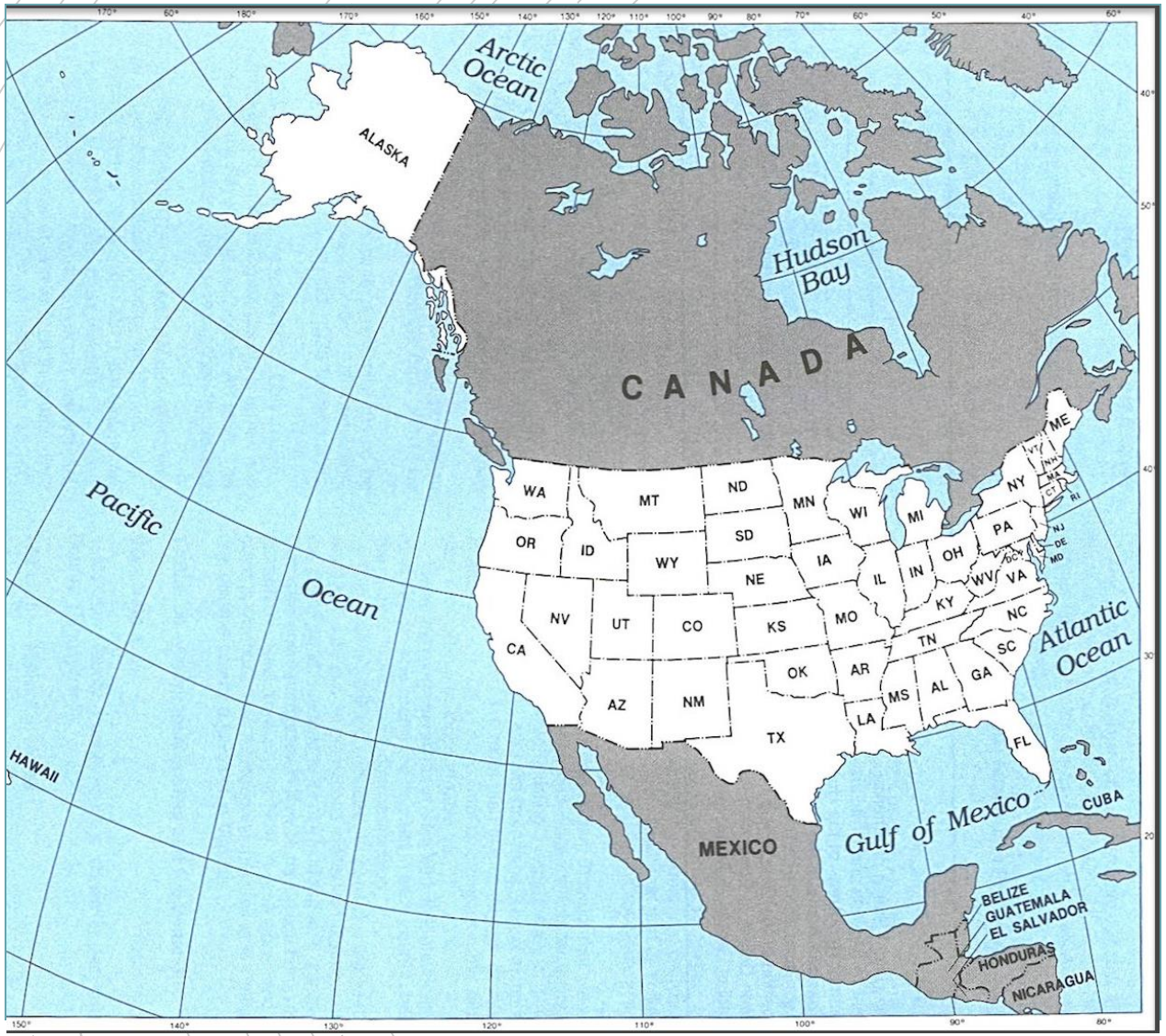


USA
Crossing Border EDA and Time Series.

- The U.S. – Mexico Border: identify total number of legal immigration.

- The U.S.-Canada Border: recognize total number of inbound activities.

- **Understanding:**

Data of inbound number of people crossing the USA Canada/Mexico borders beginning of January 1996 through December 2021 by different type of transportation and incoming through all 121 Ports and 15 States.

- **Motivation:**

Using this dataset to learn about the incoming value into the USA. It can also prove to useful to predict the monthly and annually traffic that's going to accumulate on the borders so that the authorities can be aware of the number of people beforehand.

- **Goals:**

Using EDA and Visualization techniques to define the Top number of Value in State, Border, Country, by Measure , Monthly and Annually. Apply Time Series Model and Predict 2-3 Years of inbound Value. *'Border_Crossing_Entry_Data.csv'*

Data

Understanding

Motivation

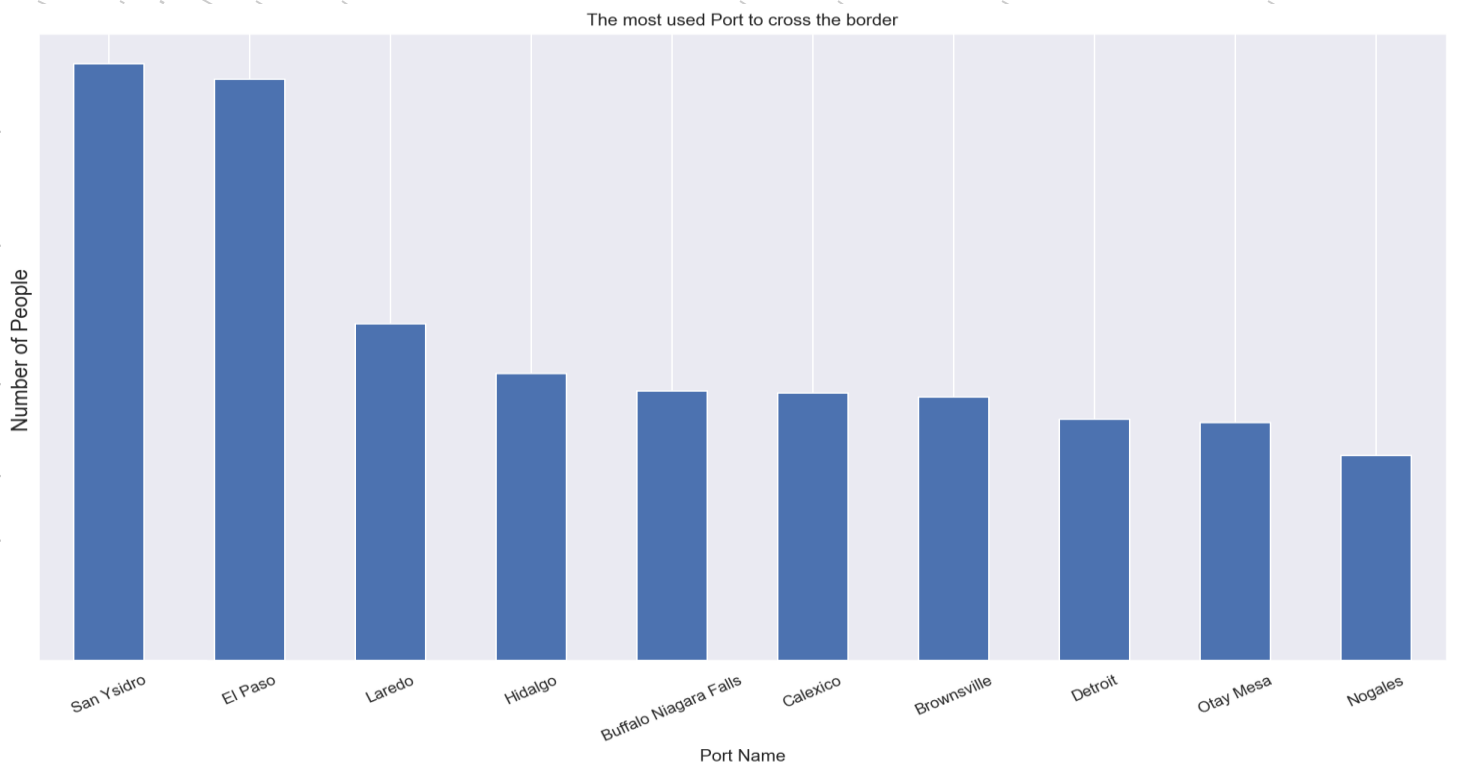
Goals

Determine the incoming number of people to the USA by following method

- Finding if there is a difference and/or any consequence between Ports by number of inbound traffic and people.
- Visualizing and displaying: The most frequent Port Name and the top 10 Port Names with the highest number of people crossing it.
- Locating the most frequent and the most crossed of top 10 States by number of people.
- Calculate the leading neighboring country from which the highest number of people crossed the border.
- Evaluating the most commonly used Type of transportation (Measure) in order to cross the border and number of People through US-Mexico and US-Canada by Value .
- Investigating if there's any specific relation with the Value of border crossing with the Date.
- Forecast and create a model with SARIMAX ,search for best parameters with ARIMA and Predict 2-3 years ahead of 2021 using Time Series of the value growth of incoming traffic and people to USA.

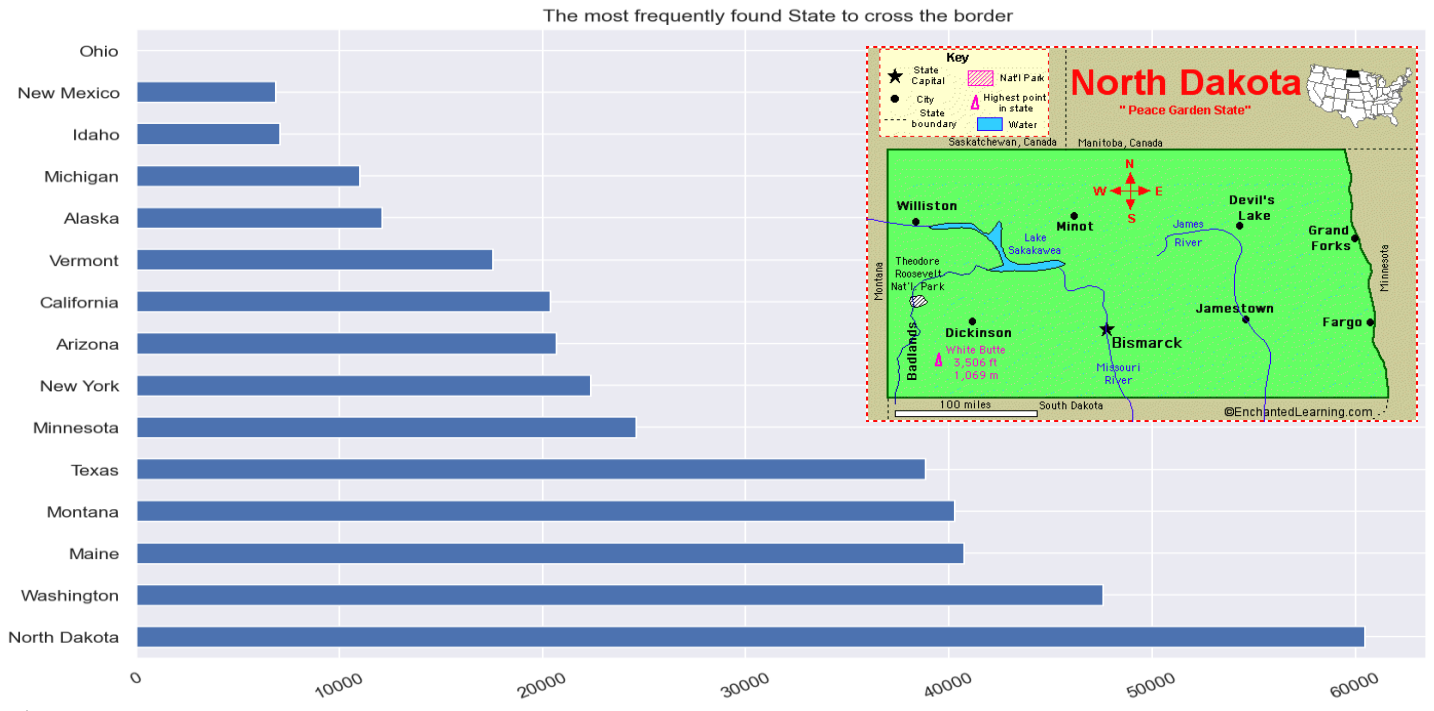
Port Names crossing

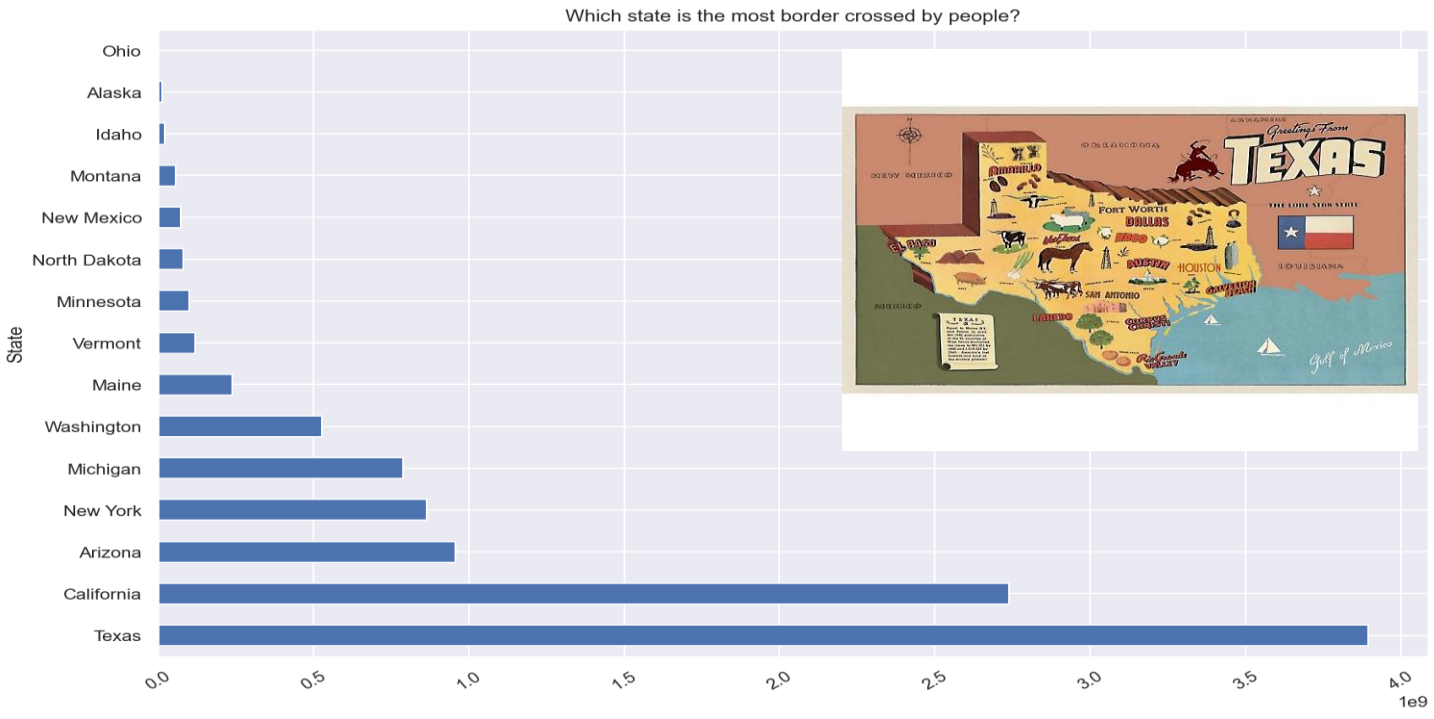
- **East Port and Buffalo Niagara Falls are the most frequently crossed Ports.**
- **San Ysidro and El Passo are two highest out of top ten Port Names by number of people.**

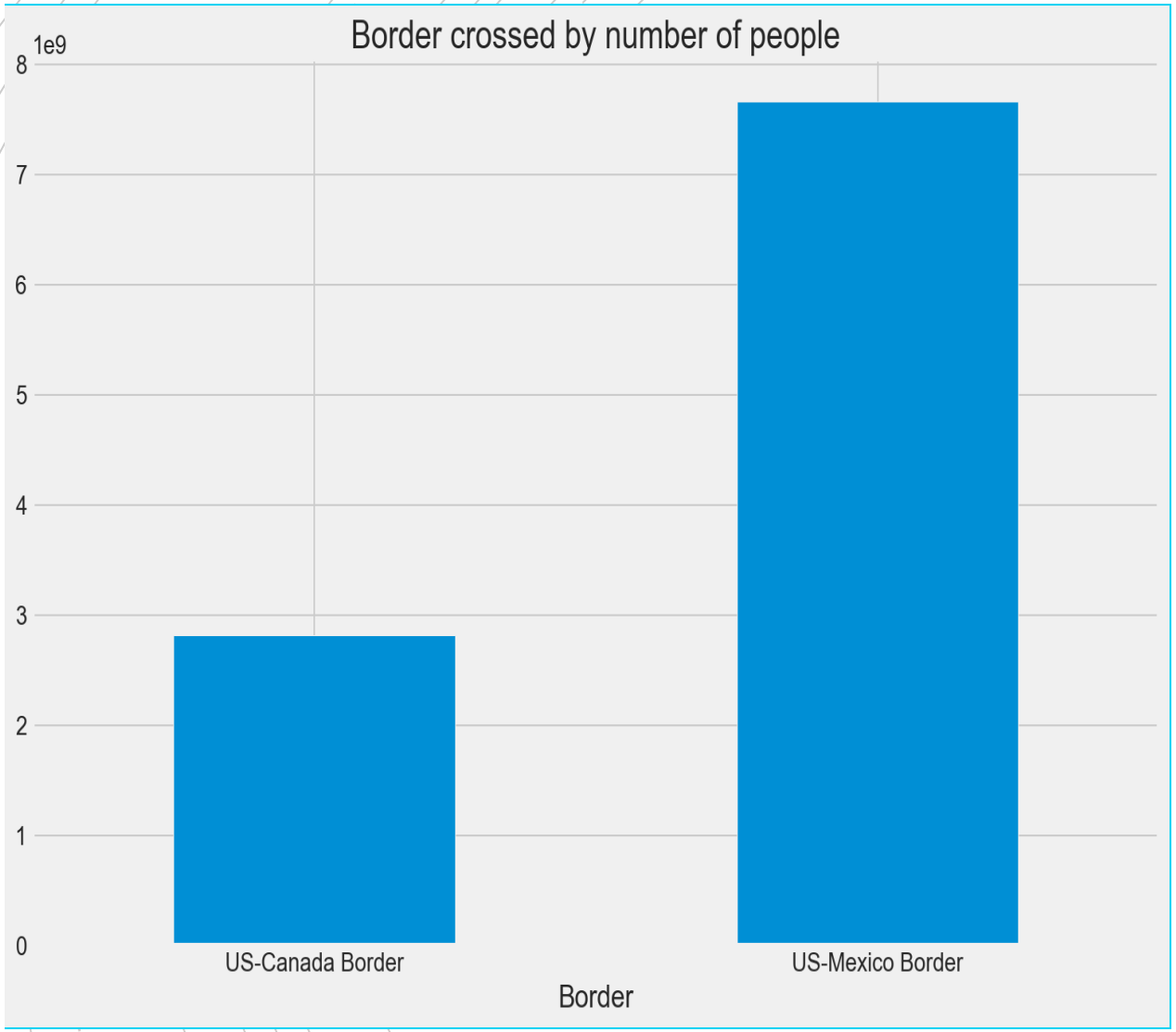


- Per the above plots , States North Dakota-ND and Washington-WA appear to be the most frequently found States to cross the border.

- The highest value of crossing border is through Texas-TX and California-CA.







- Per plot above , the US-Mexico Border has a drastically higher value of crossing.
- Number of inbound People (Value) in US-Mexico Border is 76,630,976,310
- Number of inbound People (Value) in US-Canada Border is 2,817,379,880.

Measure	Border		
Personal Vehicle Passengers	US-Mexico Border	19163986	4198297783
Personal Vehicles	US-Mexico Border	19163986	1987320131
Personal Vehicle Passengers	US-Canada Border	63809550	1596196106
Pedestrians	US-Mexico Border	19375565	1121819338
Personal Vehicles	US-Canada Border	63960401	767901810
Trucks	US-Canada Border	62792587	157374124
	US-Mexico Border	18371885	129495638
Truck Containers Loaded	US-Canada Border	61657890	118131725
	US-Mexico Border	18367996	83192986
Bus Passengers	US-Canada Border	55997608	75997886
	US-Mexico Border	17597170	72000097
Truck Containers Empty	US-Mexico Border	18307657	45081279
Rail Containers Loaded	US-Canada Border	53686121	35320198
Truck Containers Empty	US-Canada Border	62487461	28364276
Rail Containers Empty	US-Canada Border	53692145	14370761
Pedestrians	US-Canada Border	55132680	13531610
Rail Containers Empty	US-Mexico Border	16238944	10449400
Rail Containers Loaded	US-Mexico Border	16095892	9140909
Train Passengers	US-Canada Border	53780702	6263066
Buses	US-Mexico Border	17601994	5749108
	US-Canada Border	56006690	3171696
Trains	US-Canada Border	53788062	756622
Train Passengers	US-Mexico Border	15766224	319078
Trains	US-Mexico Border	16238734	231884

Number of Personal Vehicle Passenger through US-Mexico by Value is essentially higher.

- Number of people crossing the USA border by type of transportation/pedestrians:

Personal Vehicle Passengers 5794493889

Personal Vehicles 2755221941

Pedestrians 1135350948

Trucks 286869762

Truck Containers Loaded 201324711

Bus Passengers 147997983

Truck Containers Empty 73445555

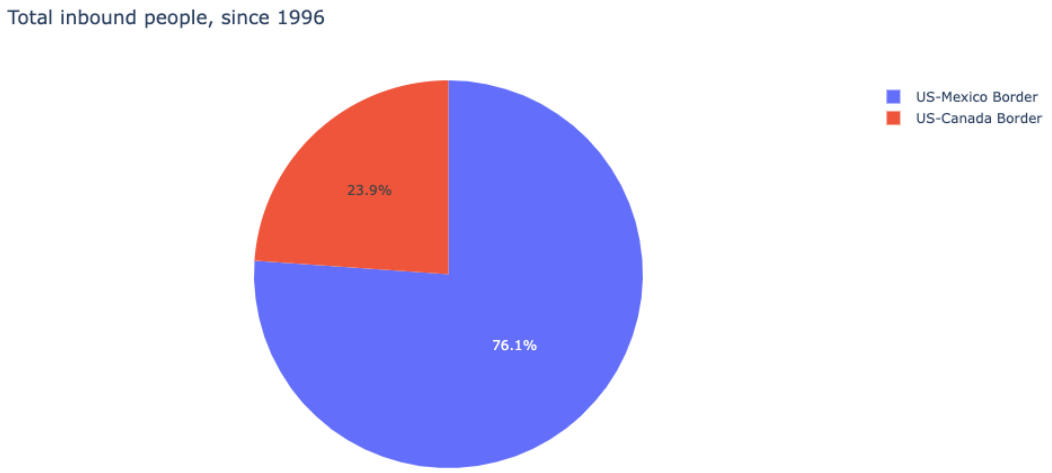
Rail Containers Loaded 44461107

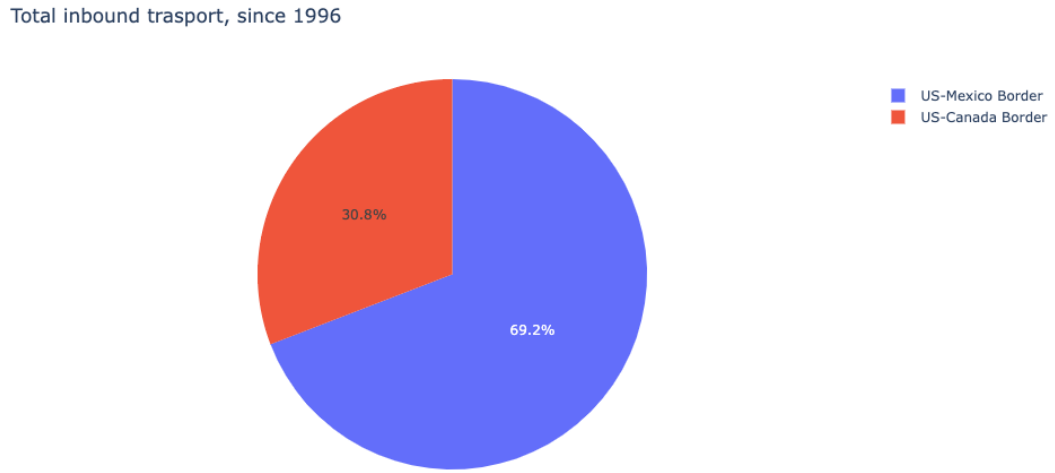
Rail Containers Empty 24820161

Buses 8920804

Train Passengers 6582144

Trains 988506





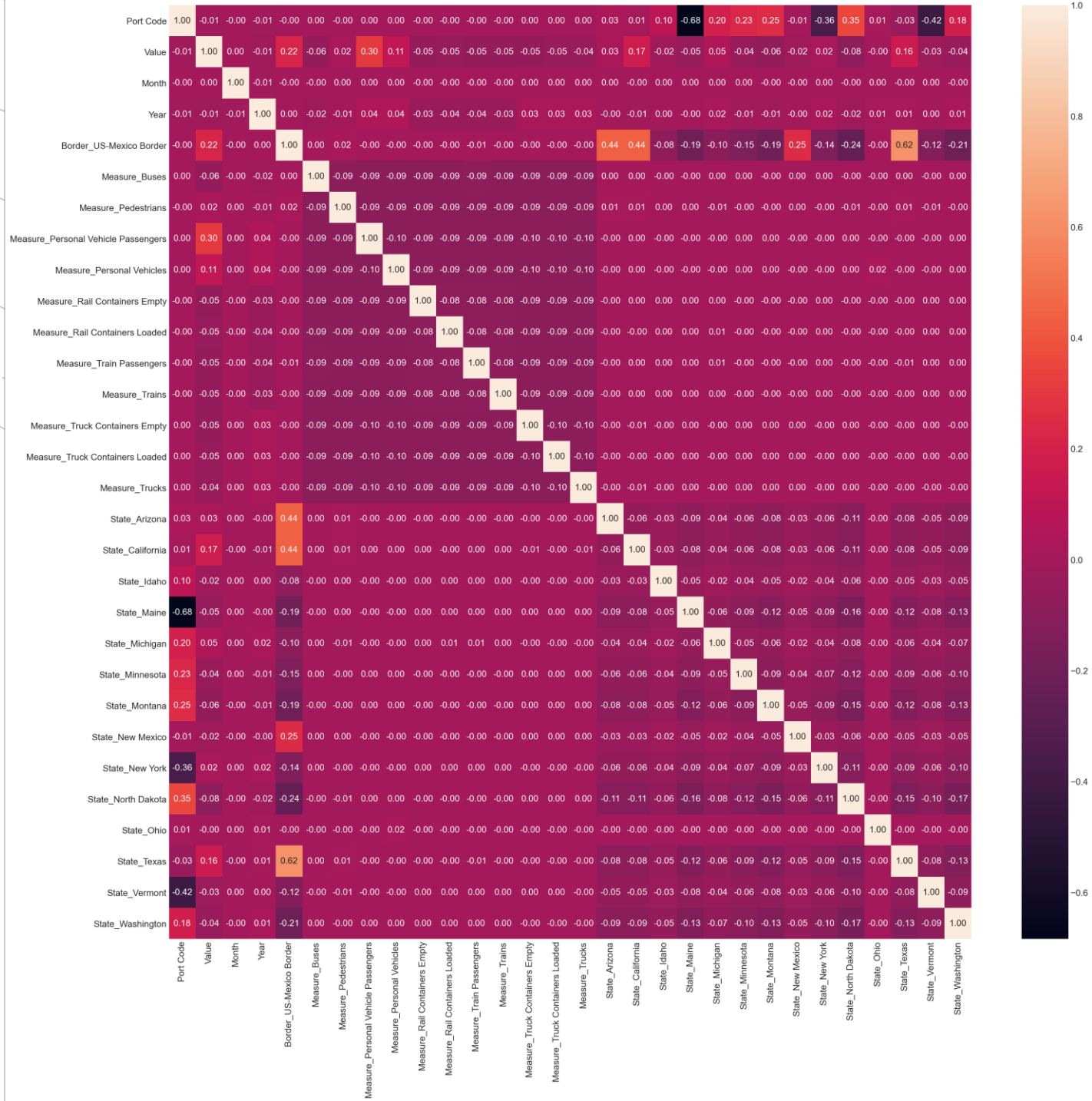
Per pie , out of total inbound people there is 23.9% of people US-Canada Border and the rest of 76.1% US-Mexico Border by all type of Transportation (Measures).

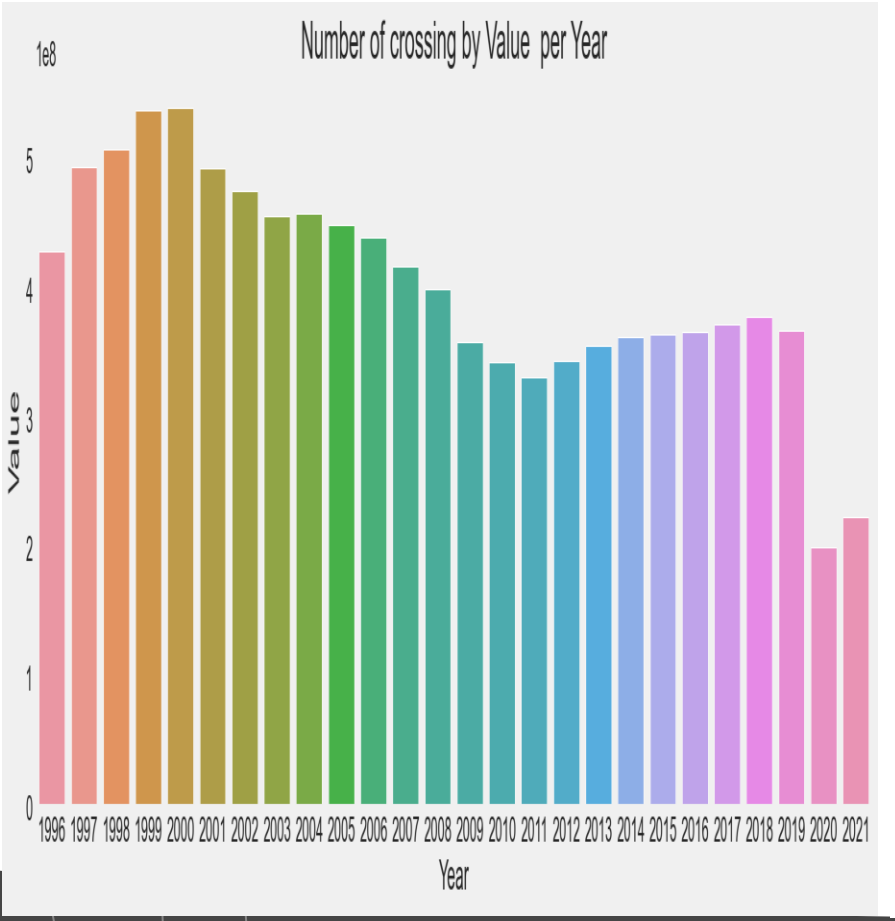
Per pie ,out of total inbound traffic there is 30.8% of all 12 type of transportation inbound US-Canada Border, 69.2%-US-Mexico Border.

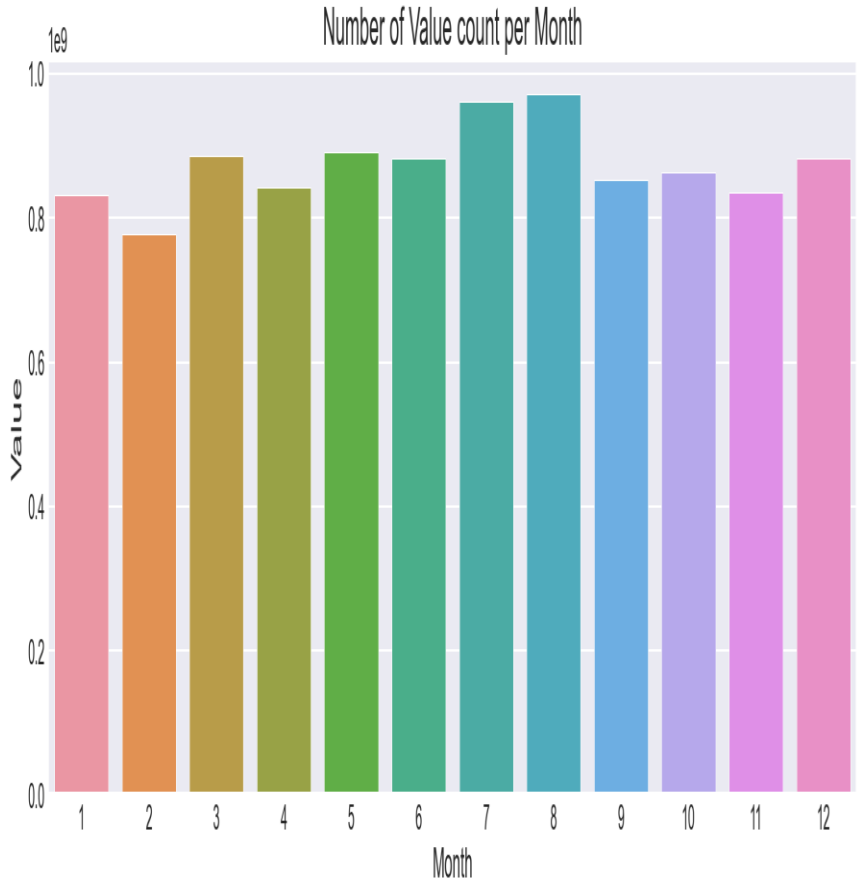
Correlation heatmap

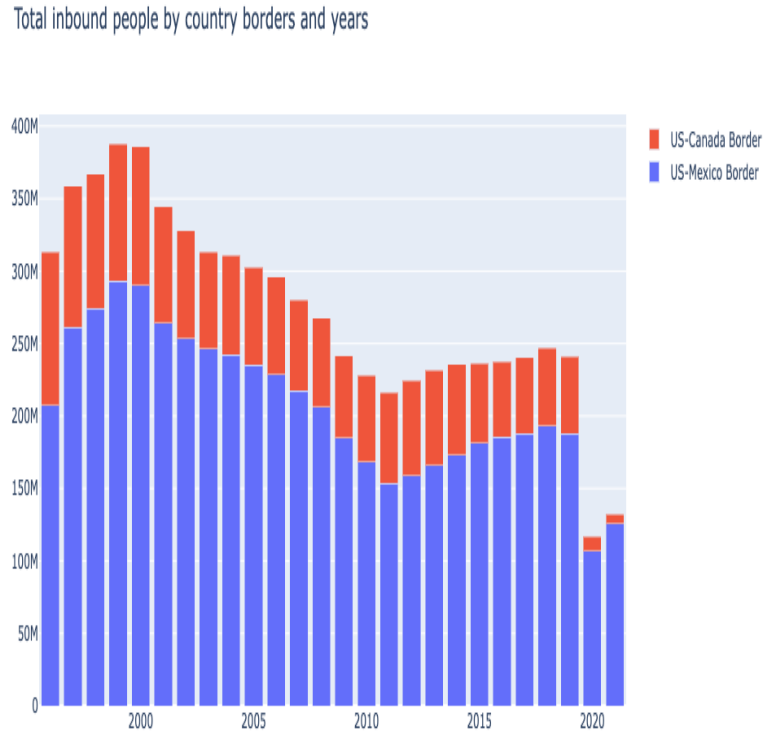
The heatmap suggests a slightly strong correlation between US/Mexico/Border in State-TX ,

Between Measure Personal/vehicle/Passenger and Value.





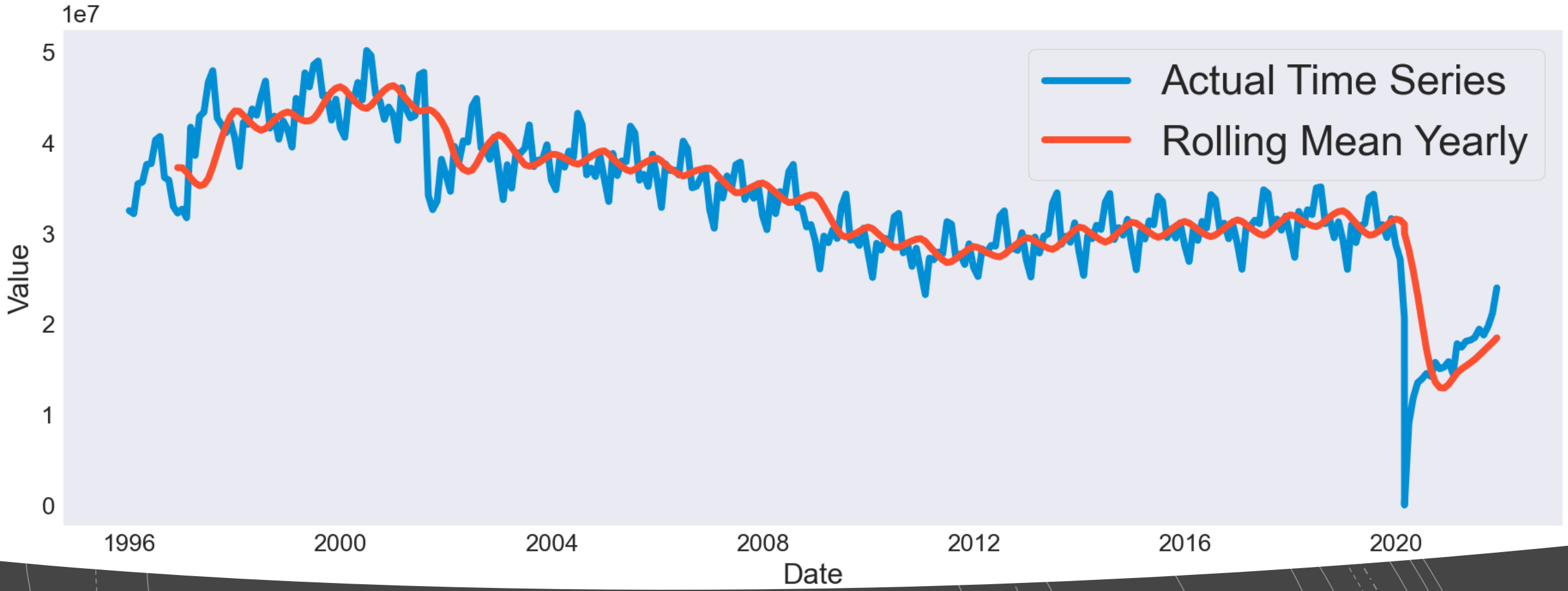




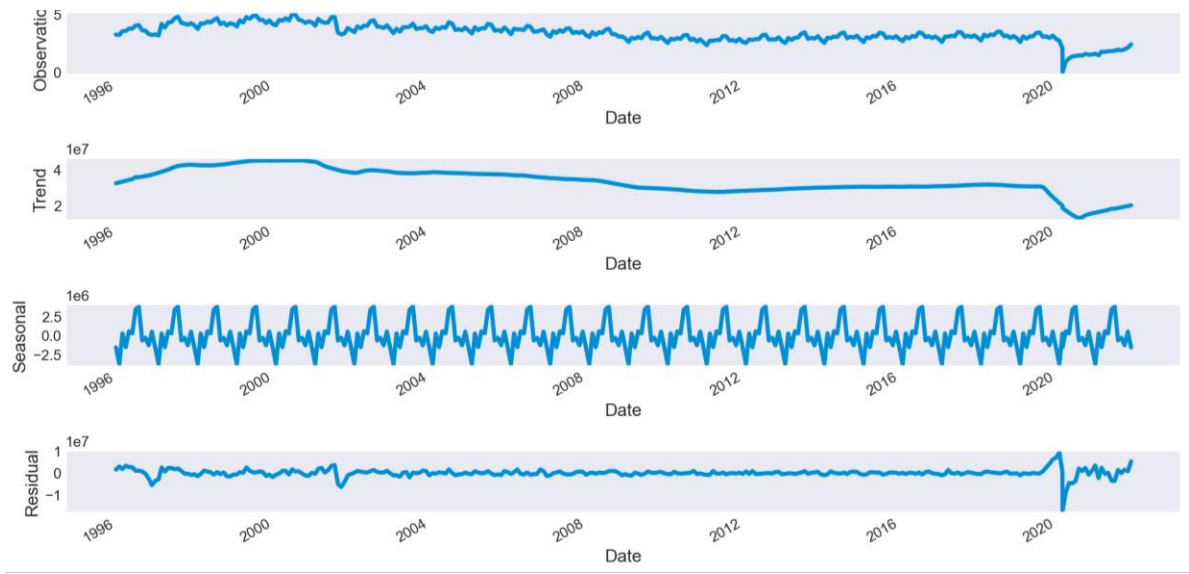
Annual growth of inbound people by borders: US-Mexico and US-Canada are displaying

- The number of inbound people by Year is significantly lower in 2011 and the lowest number is In 2008. The highest number of inbound people is withing 1999 and 2000.
- The number of inbound people by Month is highest in July and August and the lowest valuable Month is in February.
- Annual growth of inbound number of people by borders:

US-Mexico/US-Canada.



As it displayed on the plot below mid 2001 - 2012 are in a downward trend at the time of 9/11 attacks/financial crisis of 2007-2008, and slightly stable up till 2013 and even slightly increases up until 2019 right to the beginning of the Pandemic, 2020 is the Covid-19 shutdown then it tends to sharp decline. However in 2021 and further the number of people tends to increase .

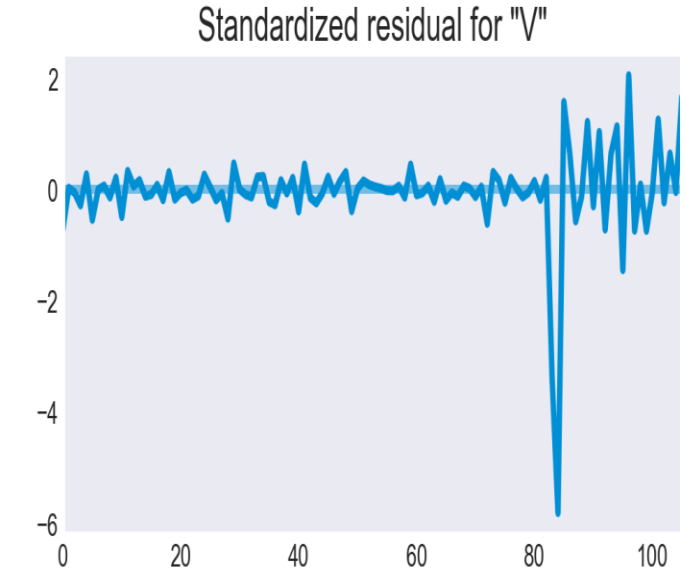


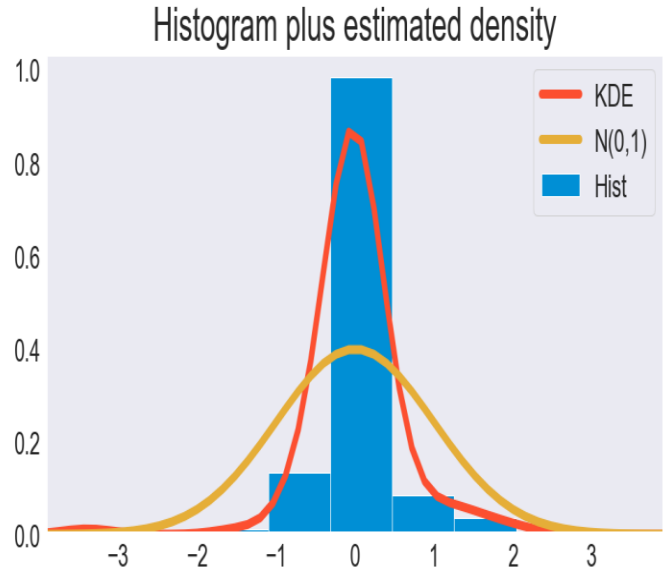
Decomposed Time Series

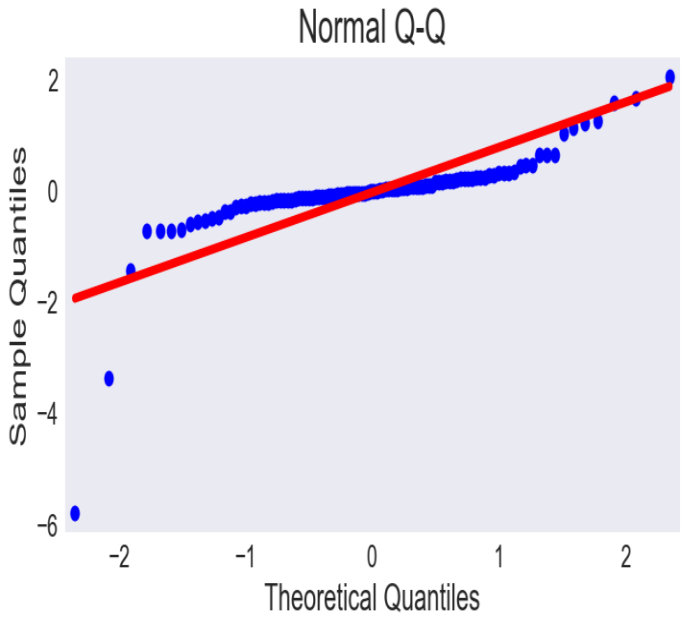
- Plotted Decomposed Data displays normal .
- Seasonality is clearly observed. It's highly related to holidays, tourism, weather condition, etc. It is clearly displayed how the mid summer season arises and a slightly jump up before Christmas and finally there is a sharp decline right after New Year events .

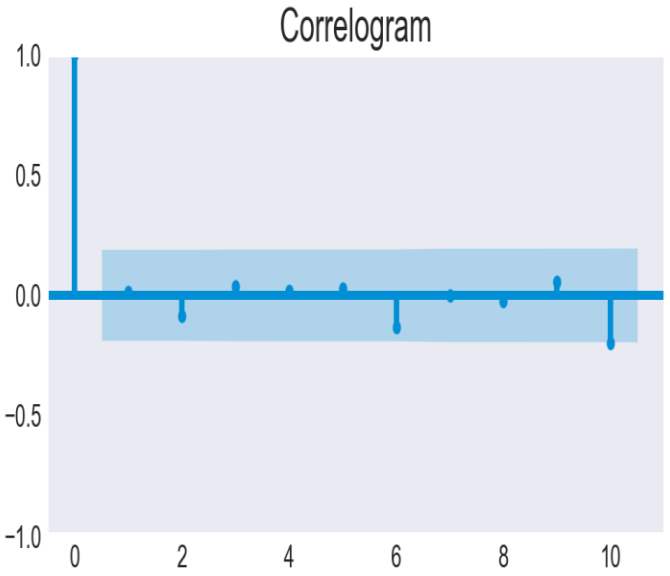
Validated Time Series Model

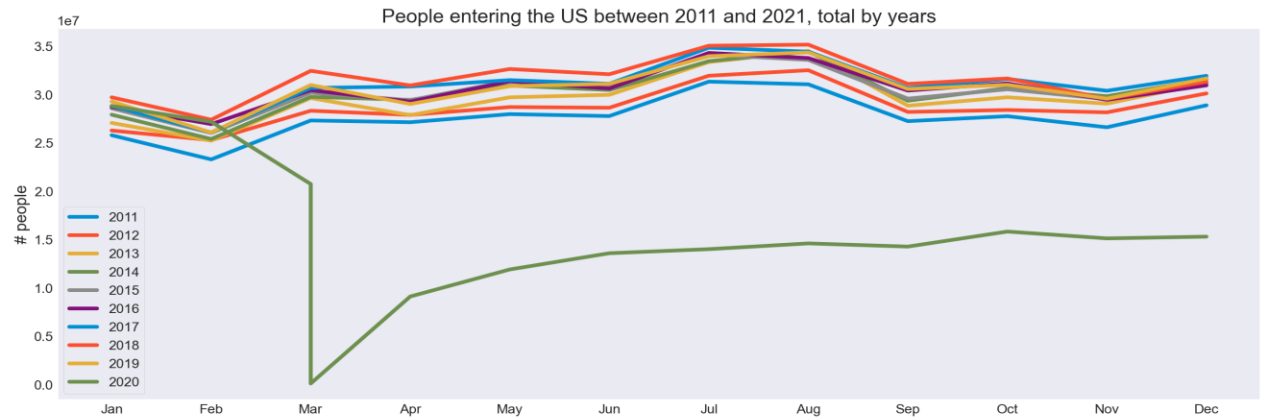
There is a slight alignment with the observed value in the forecast and only the few past years are better aligned. The histogram looks pretty much normally distributed with a very little skewness to the left.

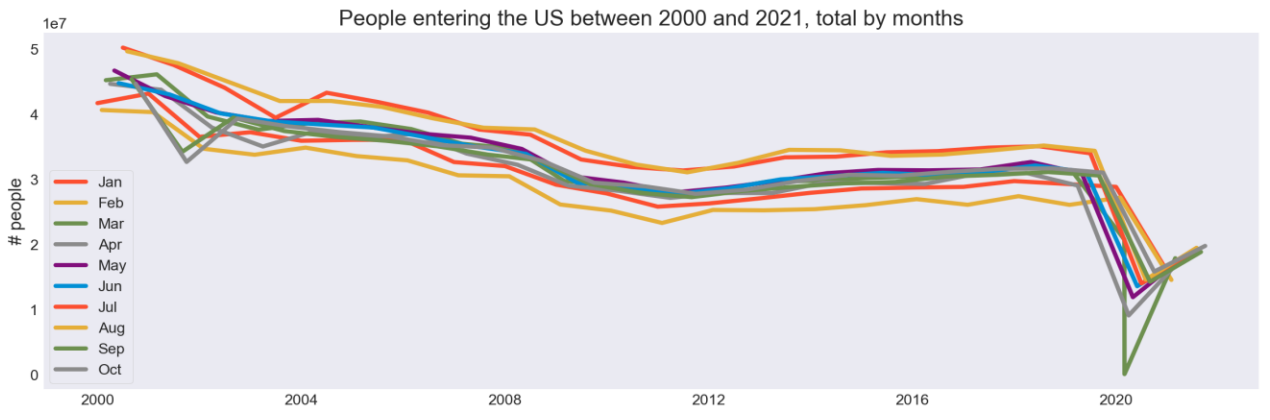


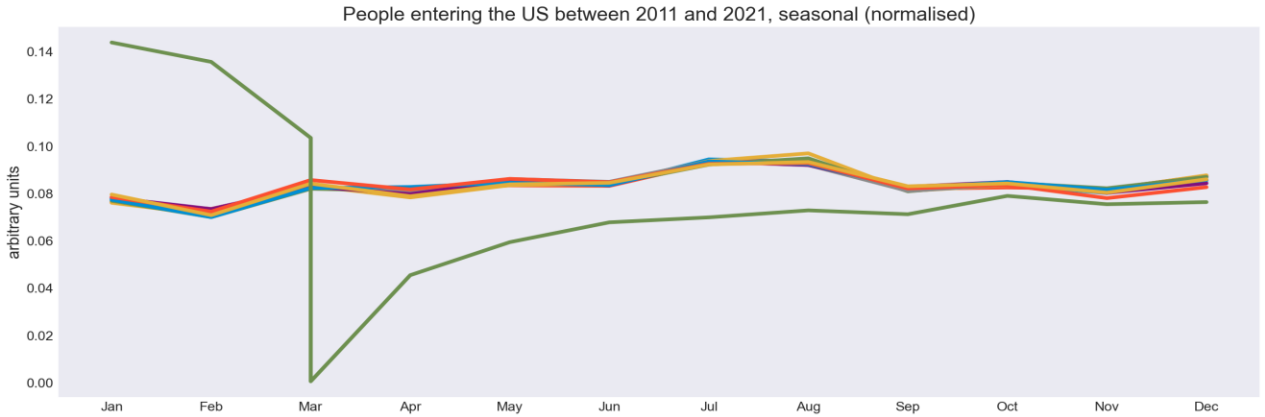


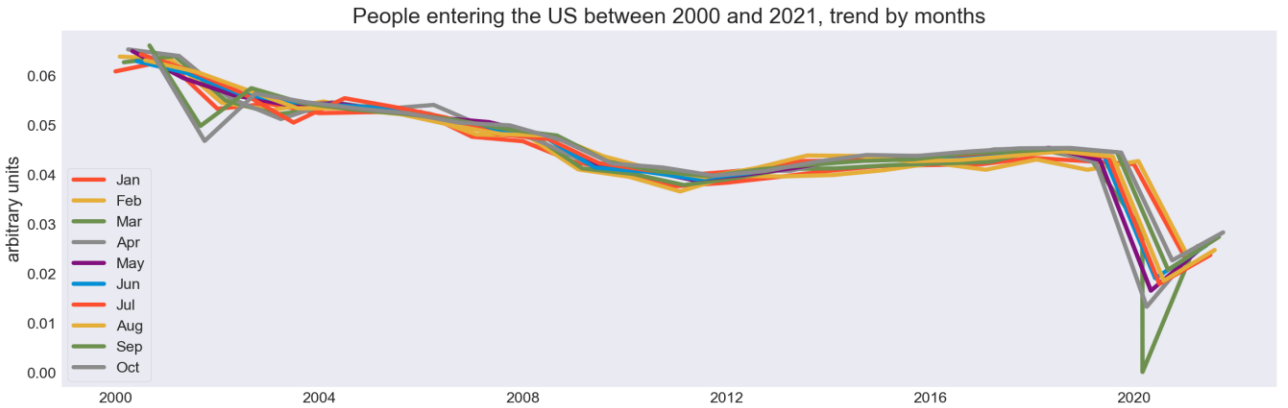






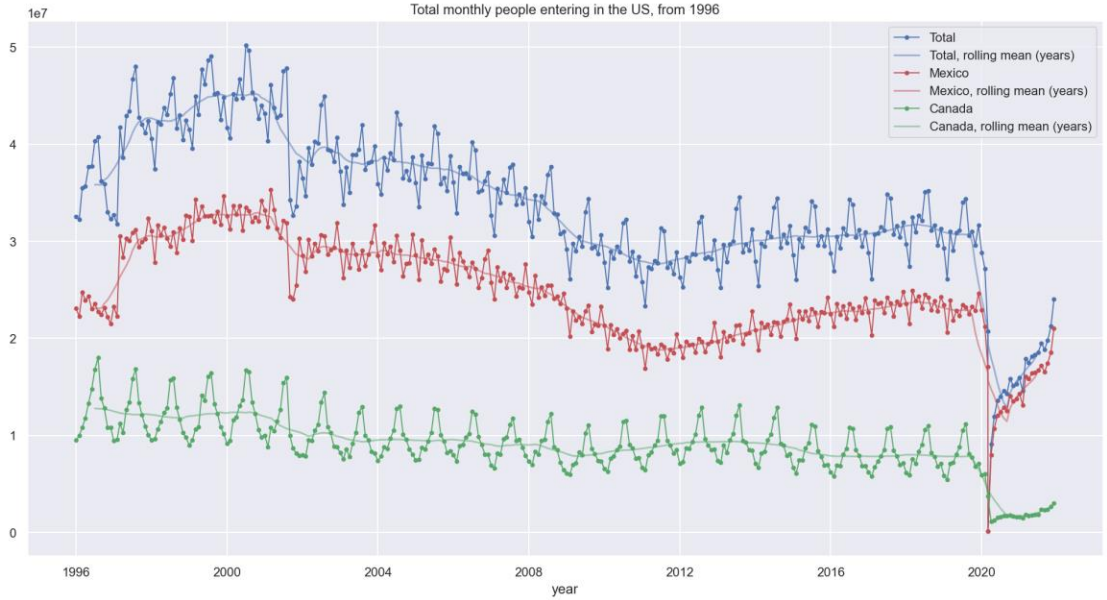






Observing and comparing seasonal and annual incoming traffic and number of inbound people to USA.

- The monthly forecast displays a slump in September & October at the end of 2002 and a way more noticeable slump on March of 2020. For the period of 2007-2008 during the financial crisis trend slides down and it's stabilizes up until end of 2019.
- Covid-19 is a fair explanation for the drastic and sharp tend down in March 2020

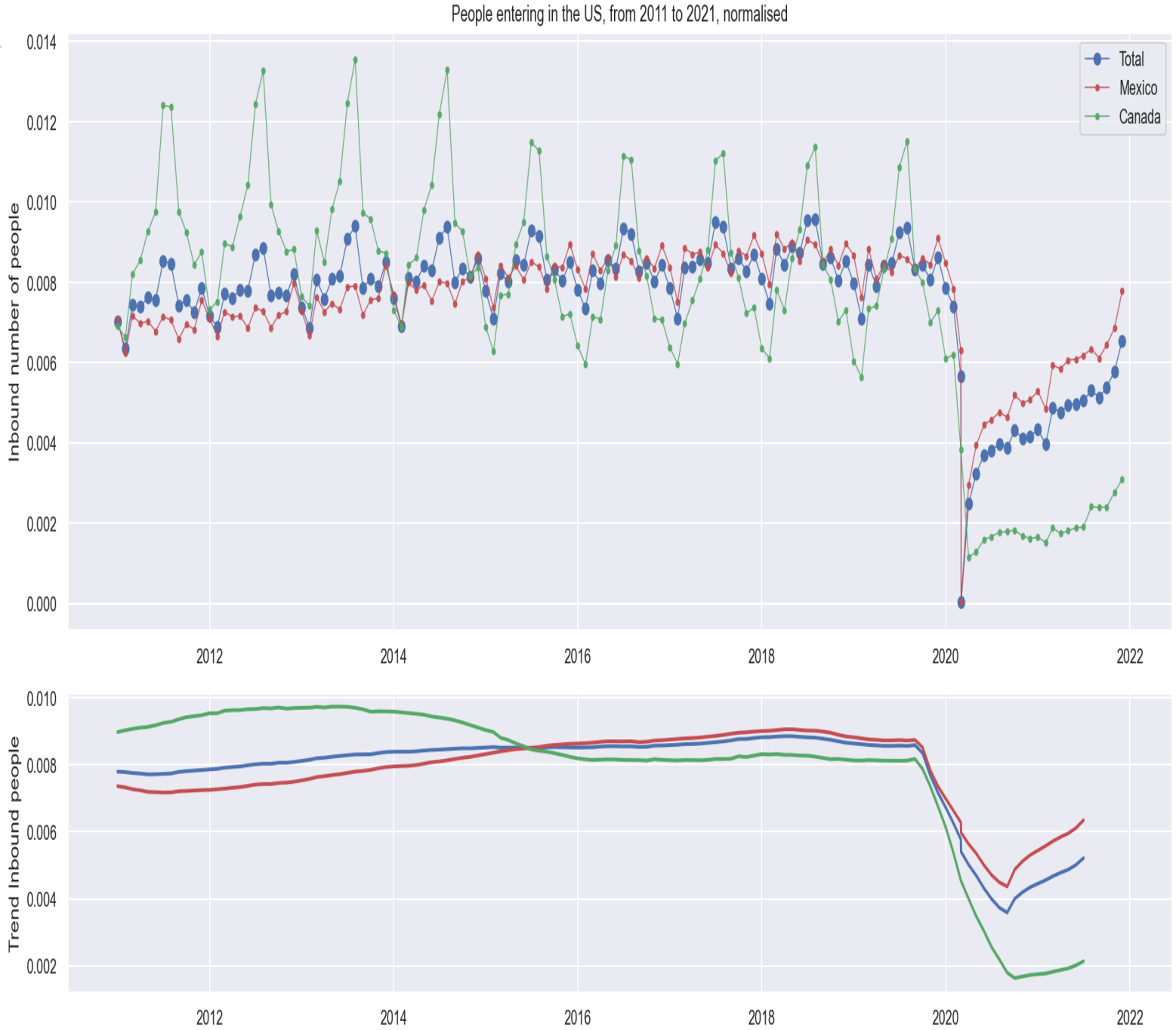


Exploring timeseries for both US-Mexico and US_Canada borders.

There's seasonal component, with a period of one year between 2020-2021.It's minimums are during the spring, especially in March, in the meantime the maximums are in summer, during July & August. The drastic tend to drop in March of 2020 is explained by Covid-19 shutdown.

Prediction of 2022

In a prediction of 2022 there is an increase of the inbound number of people from both Mexico and Canada countries, it is definitely delivering a good impact to economy and growing number of population in general after a very big distract of Covid-19 shutdown.



Conclusion

On analyzing the Border parameter, about 76.1% of the inbound crossing in the dataset consisted of US-Mexico border meaning people from Mexico have tended to come in more frequently into the US as compared to the people from Canada beginning January,1996 to December,2021 , State Texas is having the maximum share throughout the Country.

About 5,794,493,889 people have crossed the US-Mexico and US-Canada border by Personal Vehicles as Passengers which shows people avoid the public transport.