

PROJECT NAME:

Exploratory Data Analysis on Indian Tourism

Introduction:

India being one of the oldest civilizations in the world, is famous for rich heritage and myriad attractions. As the 7th largest country in the world, India stands apart from the rest of Asia, marked off as it is by mountains and the sea, which give the country a distinct geographical entity and one of the most popular tourist destinations in the world.

Objective:

Purpose of this project is to visualize Tourism in India as a major tool to bring socio-economic development to the people of the country. This project highlights on following facts which are interrelated to Indian Tourism:

- Visualization & Statistical Analysis of Foreign Tourist Arrival (FTA) in India over 2009-2019
- Visualization & Statistical Analysis of Foreign Exchange Income of India over 2009-2019
- Visualization & Statistical Analysis of GDP of India wrt Indian Tourism
- Visualization & Statistical Analysis of Employment Generation in India w.r.t Indian Tourism

Data Source:

<https://tourism.gov.in/>
<https://tourism.gov.in/sites/default/files/2022-09/India%20Tourism%20Statistics%202021%20%281%29.pdf>

Collected Data & Source Code:

From the above data source; data have been captured in "Stats_Tourism.xlsx" spreadsheet.



Stats_Tourism.xlsx

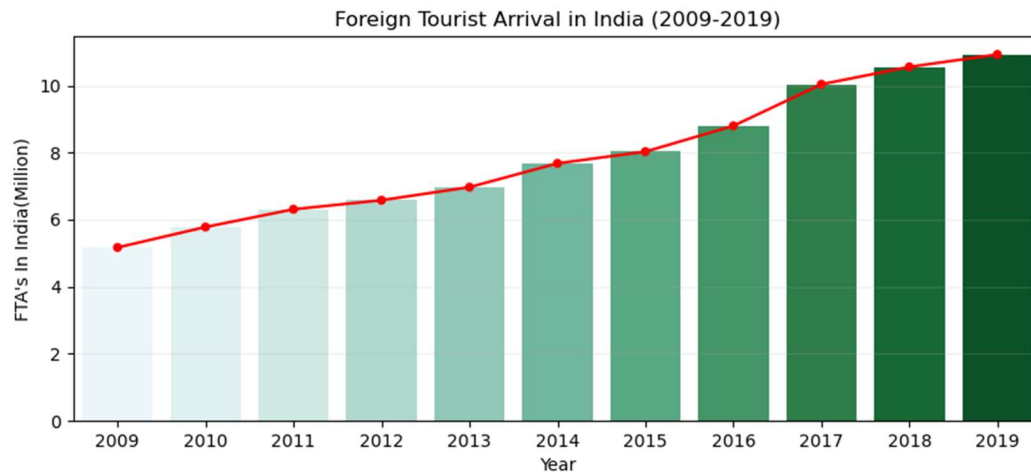
Data Description:

From the above data source, we have jotted down the following data into a spreadsheet named "Stats_Tourism.xlsx".

- Tab1: "YoY Growth Of FTA" contains Year & Foreign Tourist Arrival (FTA) Values
- Tab2: "Region wise FTA" contains Year, Continents & Region wise FTA Values
- Tab3: "Forex Income" contains Year & Forex Income Values
- Tab4: "GDP" contains % contribution of tourism on GDP of India over years
- Tab5: "Employment_Generation" contains Total Employment & Contribution of Tourism towards Employment

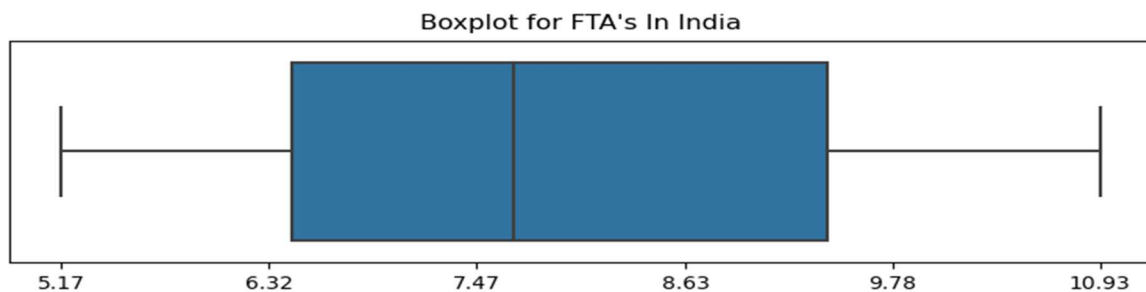
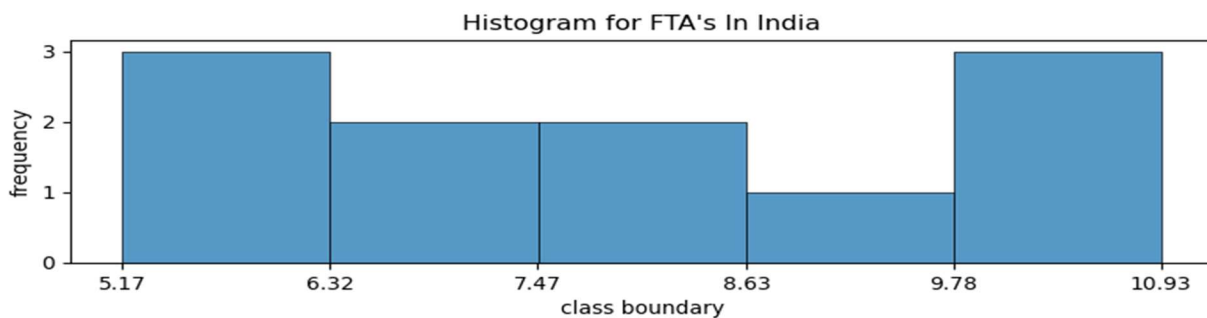
Univariate Analysis:

i. Foreign Tourist Arrival (FTA):



Interpretation:

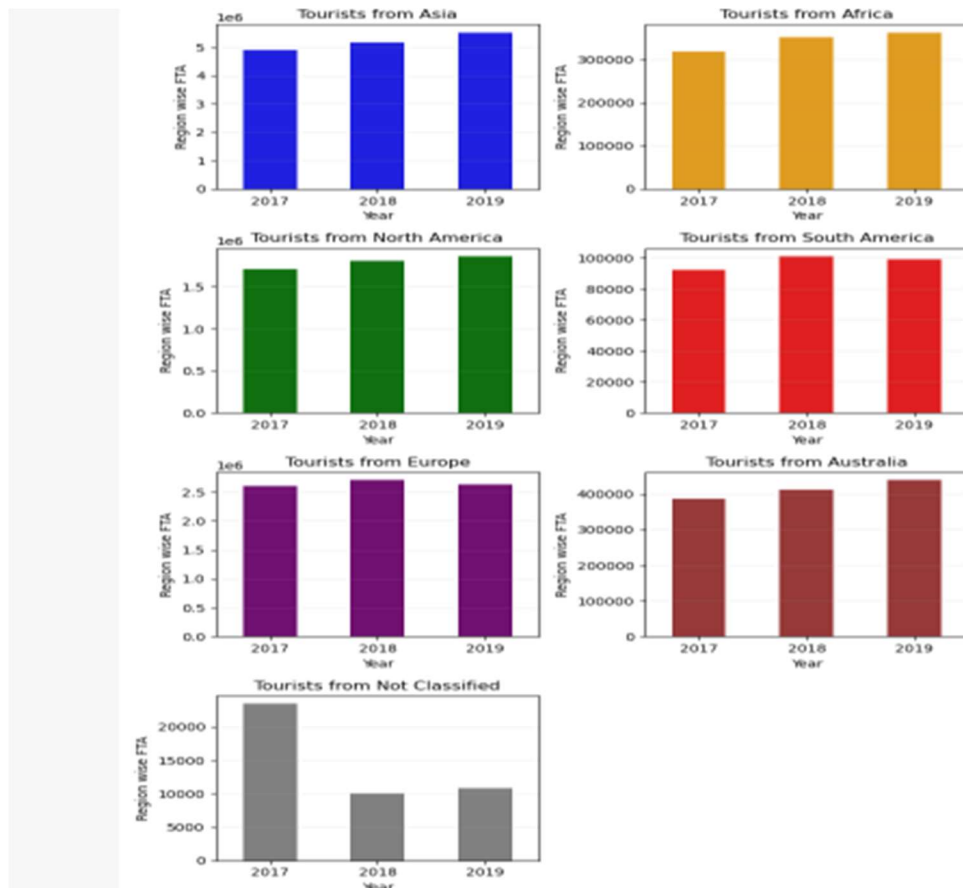
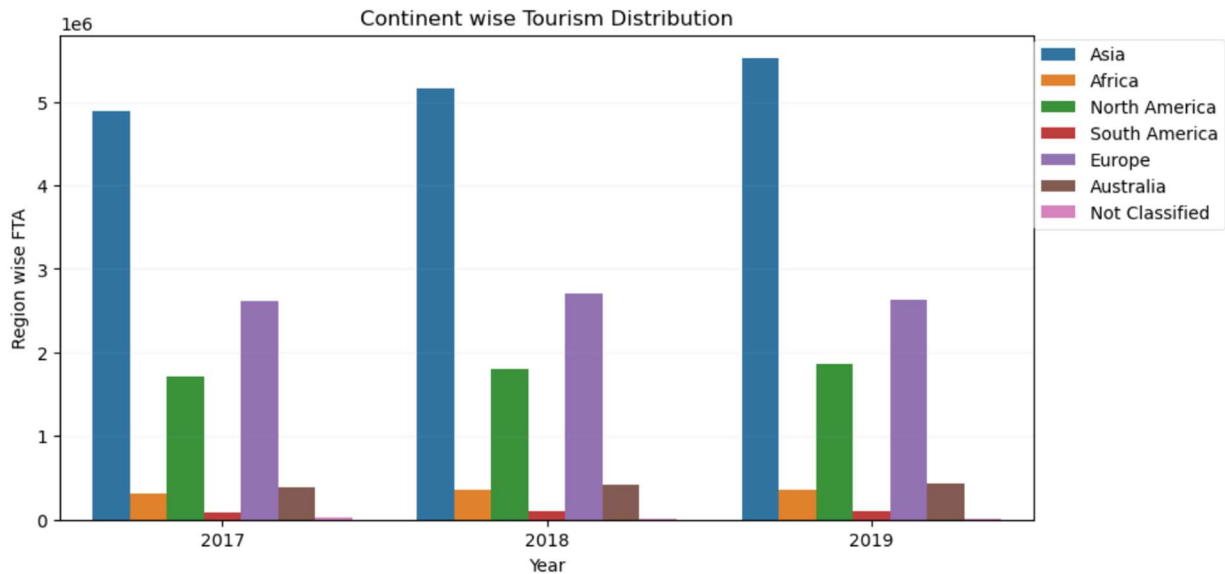
- From above graph, we can see number of foreign tourists arriving in India is increasing gradually over the years (2009-2019).



Interpretation:

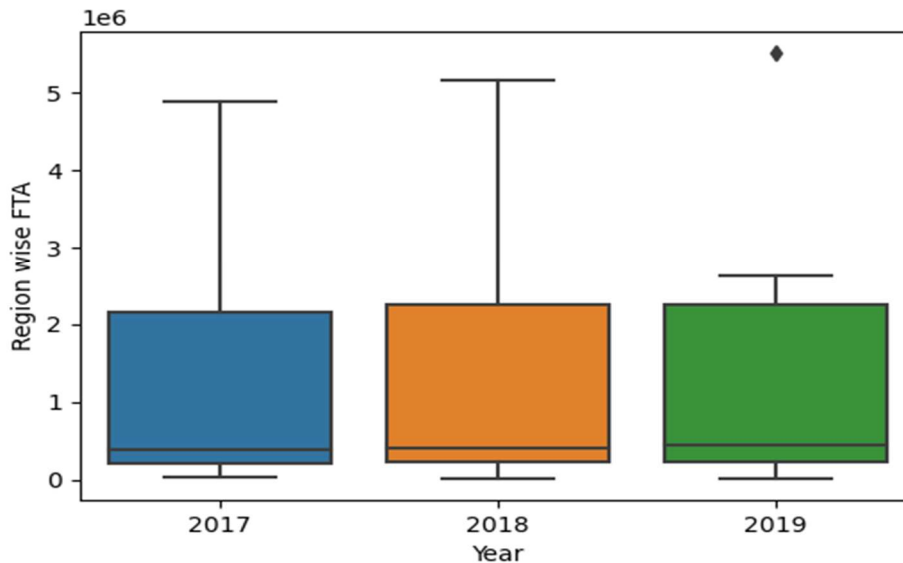
- For the given data for the span 2009 to 2019 the median value of number of tourist is 7.68 million.
- There is no outlier value for the given data
- According to Box plot 25% of number of tourists between 2009 to 2019 is less than 6.44M
- 50% of number of tourist is less than 7.68M
- 75% of number of tourist is less than 9.42M
- Middle 50% of data is lying in between 6.44M and 9.42M. IQR is $(9.42 - 6.44) = 2.98\text{M}$
- Since $Q2 < Q3$ is greater than $Q1 < Q2$; above data is a right skewed / positively skewed.

ii. Continent wise Tourism Distribution:



Interpretation:

- Over the years from 2017 to 2019; in Europe and South America, Number of tourist slightly decreases in 2019 as compared to 2018.
- Possible reasons can be:
 - lack of overseas promotion by the government,
 - outdated brand image,
 - no heads for India Tourism offices overseas,
 - poor PR efforts to counter negative publicity.



Interpretation:

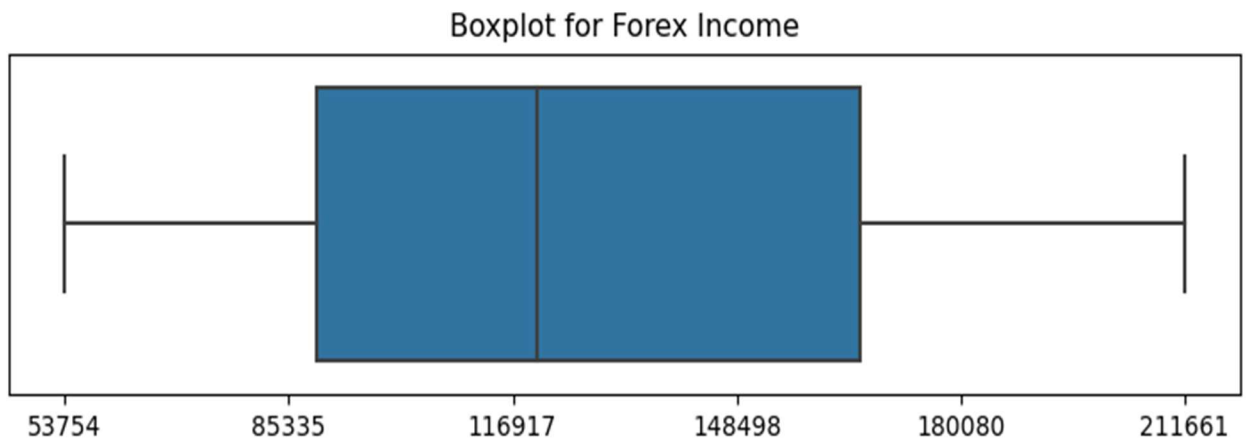
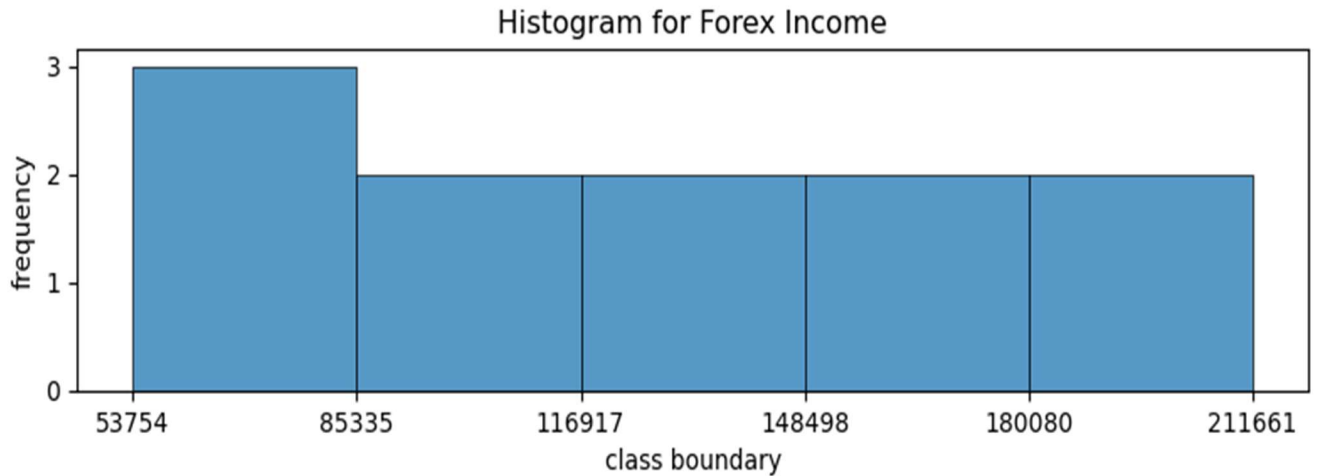
- All the above three years data (for Region wise FTA); All three are positively skewed data.
- There is an outlier value for the year 2019.
- Reason for outlier is: in 2019 a significantly high number of tourists arrived in India from Asia.

iii. Foreign Exchange Income depending on Tourism:



Interpretation:

- Since number of tourist arrived india is gradually increasing over the years, the amount of Foreign exchange Income is also increasing.
- This is a good sign for the economy Of India as its foreign reserve increases which in turn strengthens our economy.

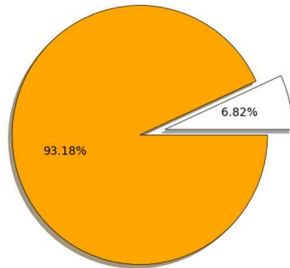


Interpretation:

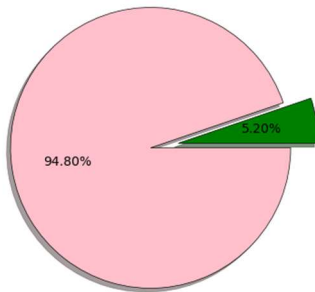
- For the given years the range of foreign exchange income lies between 53754 crore to 211661 crore
- Median of Foreign Exchange income for the given years is 120367 crore Rs
- According to Box plot 25% of Foreign Exchange income amount is less than 89321 crore
- 50% of Foreign Exchange income amount is less than 120367
- 75% of Foreign Exchange income amount is less than 166010
- Middle 50% of data is lying in between 89321 and 166010.
- IQR is $(89321 - 166010) = 76689$ crore Rs
- here $Q2Q3$ is greater than $Q1Q2$. So above data is right skewed.

iv. GDP of India depending on Tourism:

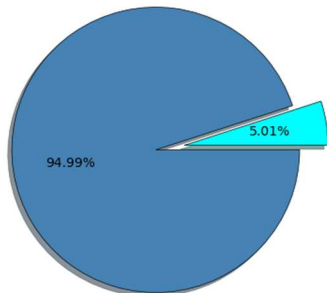
% contribution(2009-2010) of Tourism on GDP of India



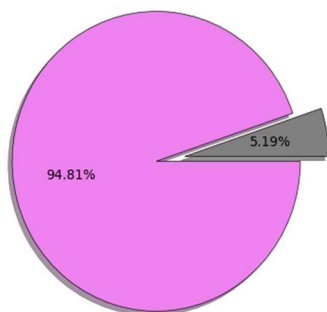
% contribution(2015-2016) of Tourism on GDP of India



% contribution(2018-2019) of Tourism on GDP of India



% contribution(2019-2020) of Tourism on GDP of India

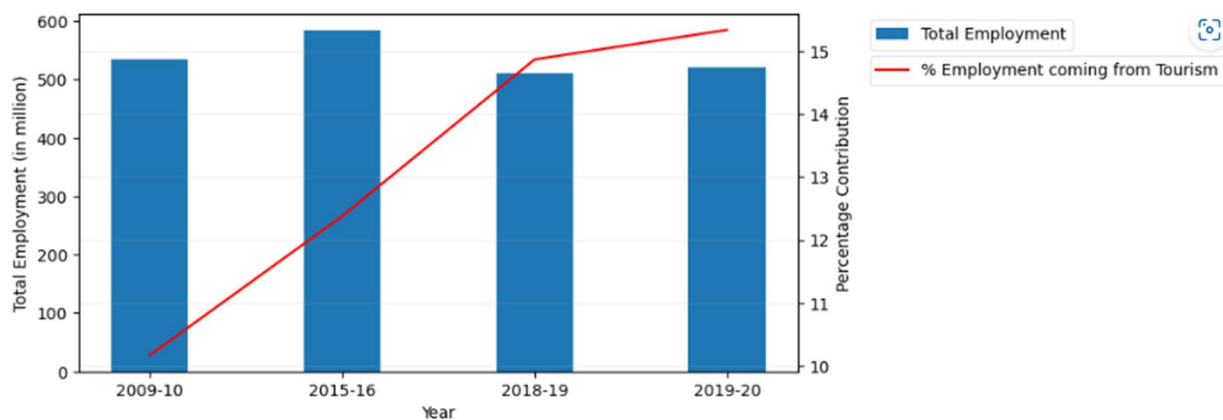
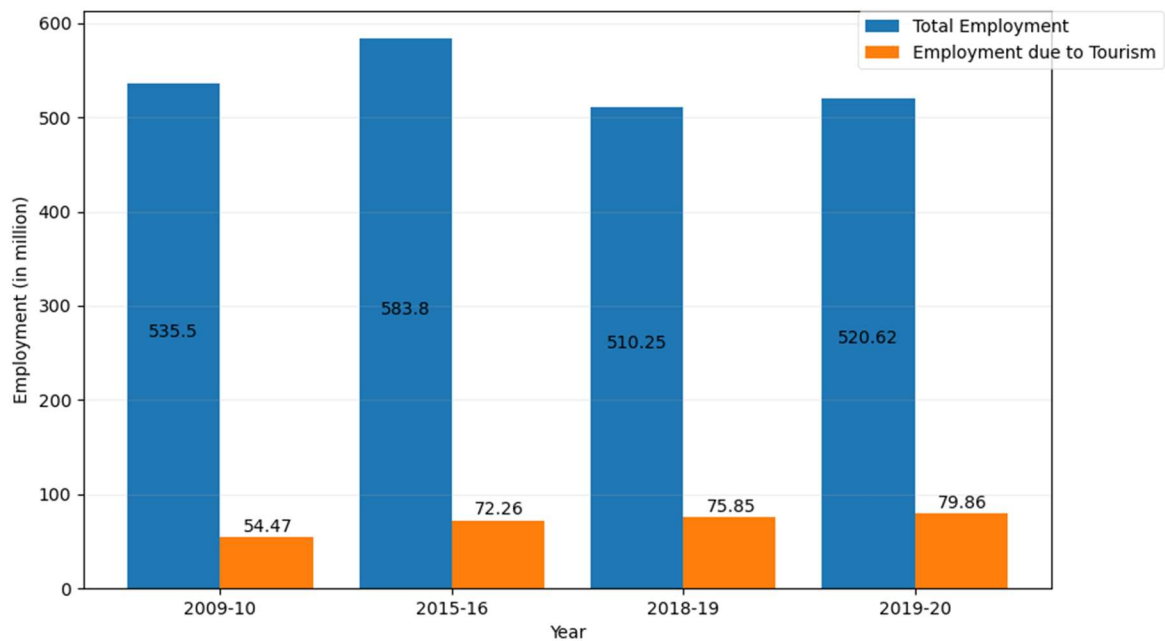


Interpretation:

- Contribution of Tourism Industry towards India's GDP was 3.68% in the year 2009-10
- But in the subsequent year (2015-16) the percentage contribution decreased and it remained constant for the next years.
- Over the years India's GDP has grown rapidly. But compared to that, tourism industry has not grown that much. Due to that reason, overall % contribution from Tourism has been decreased.

v. Employment Generation in India depending on Tourism:

Total Employment vs Contribution of Tourism towards Employment

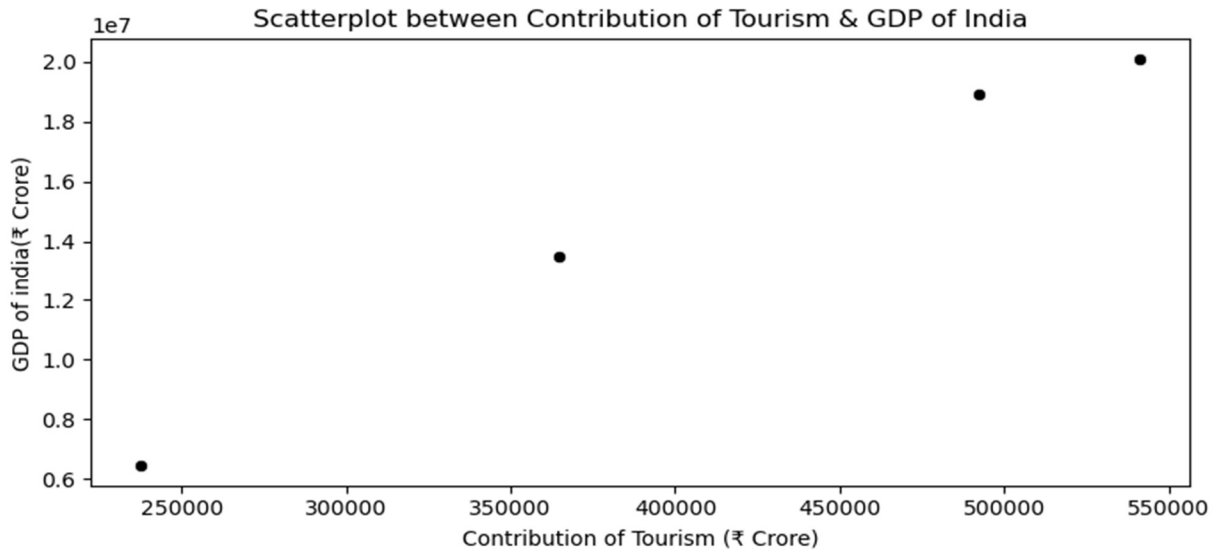


Interpretation:

- Overall employment in India around 2015-16 increased as compared to 2009-10.
- But after that the total employment generated in India is gradually decreasing over years.
- On the contrary; for Tourism, number of employment generated is gradually increasing.
- So we can assume over the years the tourism industry is flourishing.

Bivariate Analysis:

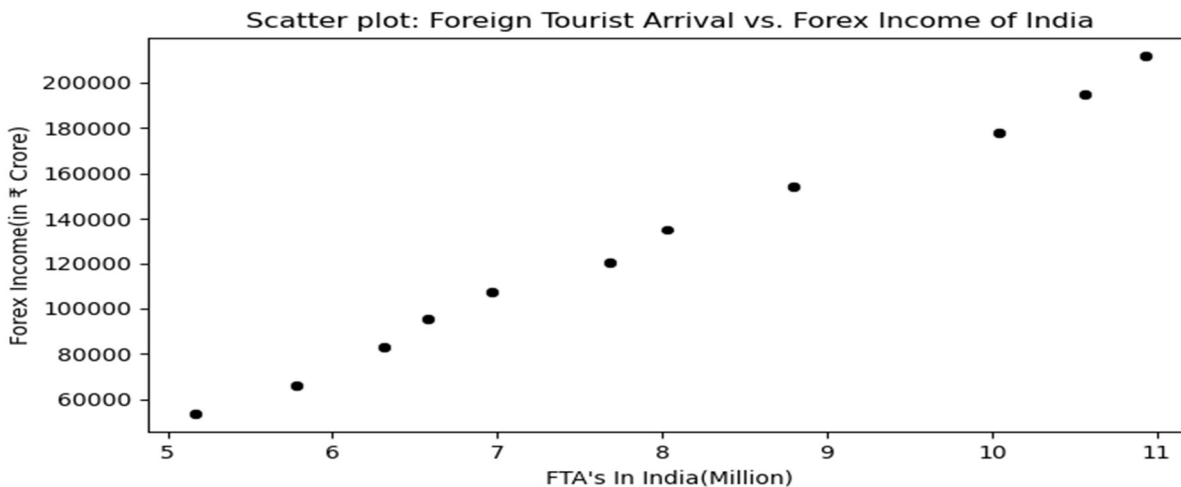
i) Bivariate Analysis for "Contribution of Tourism" & "GDP of India" :



Interpretation:

- "GDP of India" and "Contribution of Tourism" is linearly dependent to each other.
- Correlation Coefficient between these two variables is: 0.994

ii) Bivariate Analysis for "Foreign Tourist Arival" & "Forex income of India":



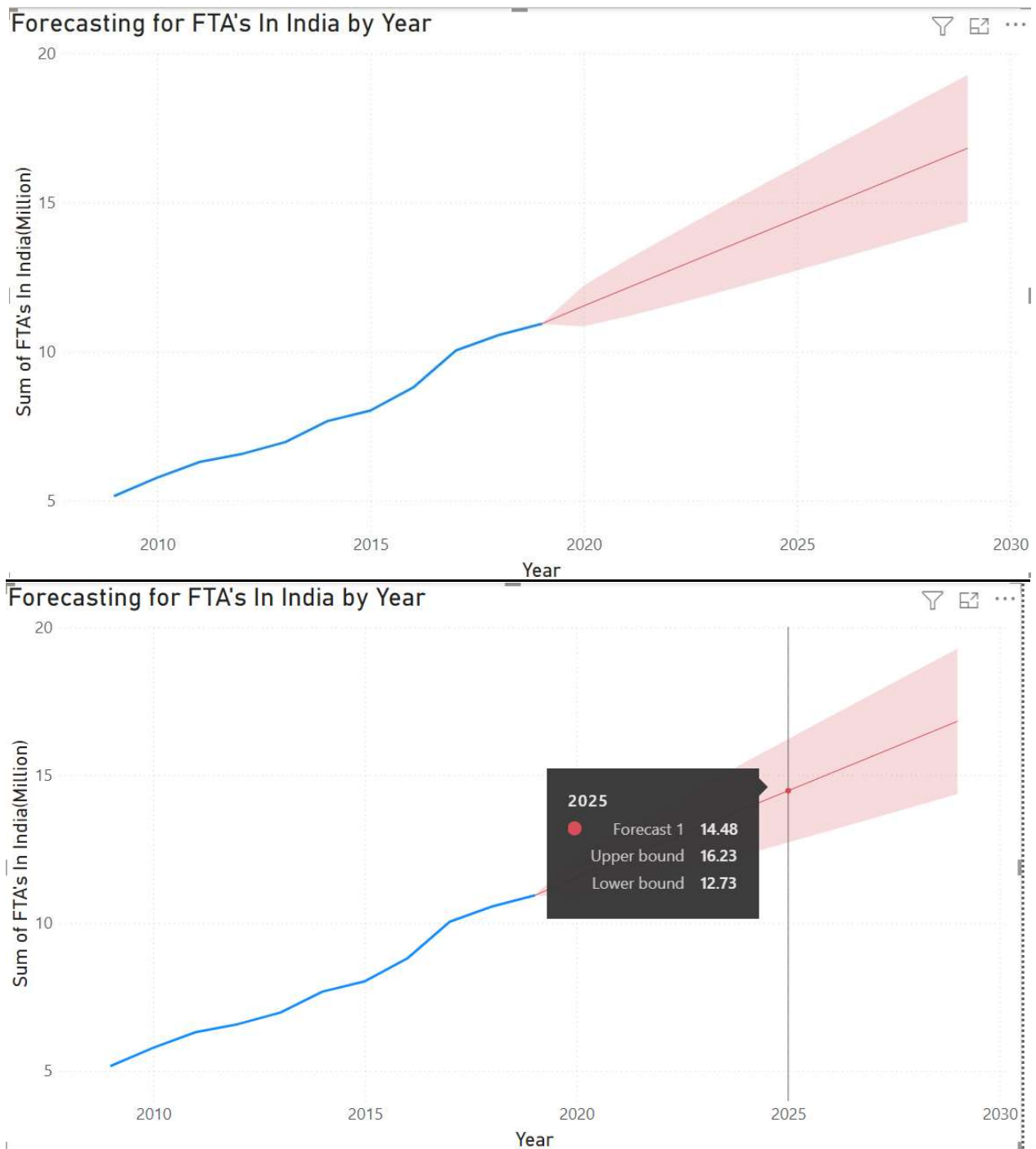
Interpretation:

- "Foreign Tourist Arrival in India" and "Forex Income of India" are highly correlated with each other.
- Correlation Coefficient between these two variables is: 0.997

Regression Analysis:

(The following graphs have been created by the help of Power BI tool)

I) Forecasting of foreign tourist arrival In India By Year

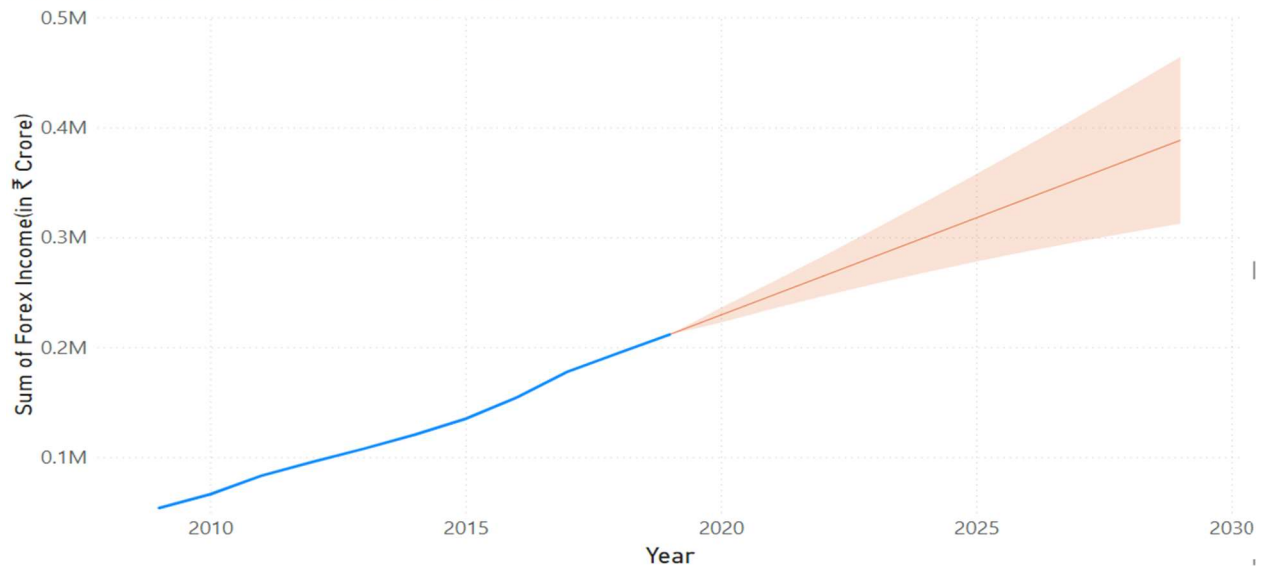


Interpretation:

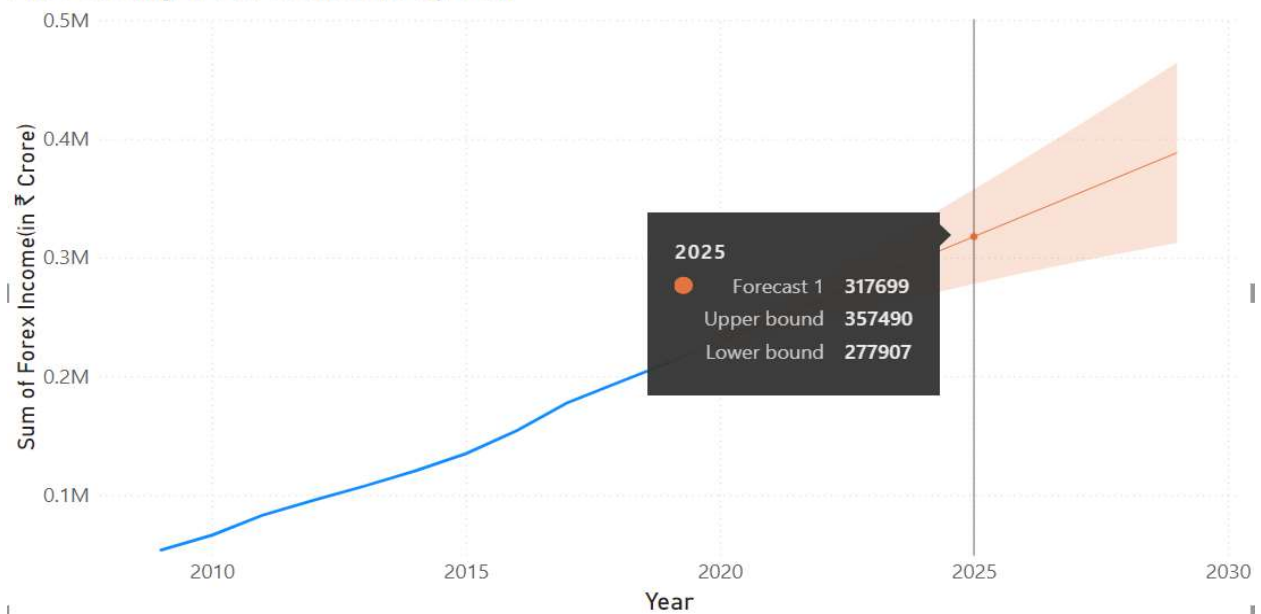
- We have performed above Regression plot at 95% confidence level & with 5 % permissible error.
- We can also forecast that in year 2025, total number of foreign tourist arrival in India will be 14.48 Million.

II) Forecasting Of Forex Income By Year Depending On Tourism

Forecasting of Forex Income by Year



Forecasting of Forex Income by Year



Interpretation:

- We have performed above Regression plot at 95% confidence level & with 5 % permissible error.
- Above graph is also forecasting that in year 2025, total Forex income of India will be 317699 crore.

Conclusion

From all of the above exploratory data analysis, we can conclude the following facts-

- Foreign Tourist Arrival (FTA) in India will be increasing over the years.
- Since Foreign Exchange income is highly correlated with FTA, we can conclude that, over the years Forex income of India will be showing a sharp rise pattern.
- For the same reason; %contribution of Tourism towards India's GDP will also be increased.
- As a result, Indian Tourism will generate more and more employment opportunities in near future.