Identification of new air travel routes in emerging markets

2 Our team

Karthik Arunkumar
Software Engineer

Bhagyashree Gohiya
Analyst – Software Engineer

Udatha Venu
Senior Software Engineer

Agenda

TOOLS AND TECHNOLOGIES

5 ETL PIPELINE

Dataset is taken from Kaggle

Scrapped data is now extracted, transformed and made ready for visualization using Jupyter and Databricks

The Transformed Data is Visualized for the further analysis

6 Introduction

To identify new route opportunities in the emerging markets by tracking macro-economic activities and data from other means of transport. The focus is on tier II cities from India. Aviation sector is showcasing intriguing growth rate in the emerging economies like India and other APAC nations. Due to various government initiatives, new aviation infrastructure is being developed in India. There is a push to provide affordable air travel for wider population. In addition, the economic growth of India is leading to distribution of economic activities across tier-II cities, which wasn't necessarily the case in the recent past.

Problem Statement

8 Problem Statement

In this fast-growing market, we need to develop new air- travel connectivity for faster and more profits to the economy. By this we might connect with the across tier-II cities and between metro / tier-I cities. And there will be better connectivity among narrow and wider body segments.

Solution

10 Solution

We will analyze the Marco - economic data and the number of Travellers in Tier 1 cities and will be compared them with the Tier 2 cities. By this, we will conclude where tier 2 cities lag and what needs to be improved to increase the air travel routes.

Summary

Data Scrapping

Datasets have been downloaded from Kaggle

Dataset link --->> International Airtraffic to and from India

---->> List of companies

---->> Tourism

11 Data Transformation

- Here the datasets are being cleaned.
- Applying the required filter conditions on the datasets.
- Making the datasets ready for visualization.

12 Data Transformation

Here are the code snippets of transformation

```
#df9=df4.withColumnRenamed("Name of the Monument","Location")
Tourismdf=df4.select("Circle","Name of the Monument ", "Domestic-2019-20", "Domestic-2020-21")
Tourismdf1=Tourismdf.withColumnRenamed("Circle","City Name")
Tourismdf2=Tourismdf1.where((col("Domestic-2019-20") != 0) & (col("Domestic-2020-21") != 0))
Tourismdf3=Tourismdf2.filter(col("City Name") != "Total")
Tourismdf4=Tourismdf3.filter(col("City Name") != "Grand Total")
display(Tourismdf4)
#Tourismdf4.write.option("header",True).csv("/FileStore/tables/Torism_data_csv")
```

```
from pyspark.sql.functions import regexp_replace
Companydf1 = Companydf.withColumn("Headquater", regexp_replace("Headquater", "\+", ""))
#Companydf = Companydf.withColumn("location", F.split("location", "\+")[0])
```

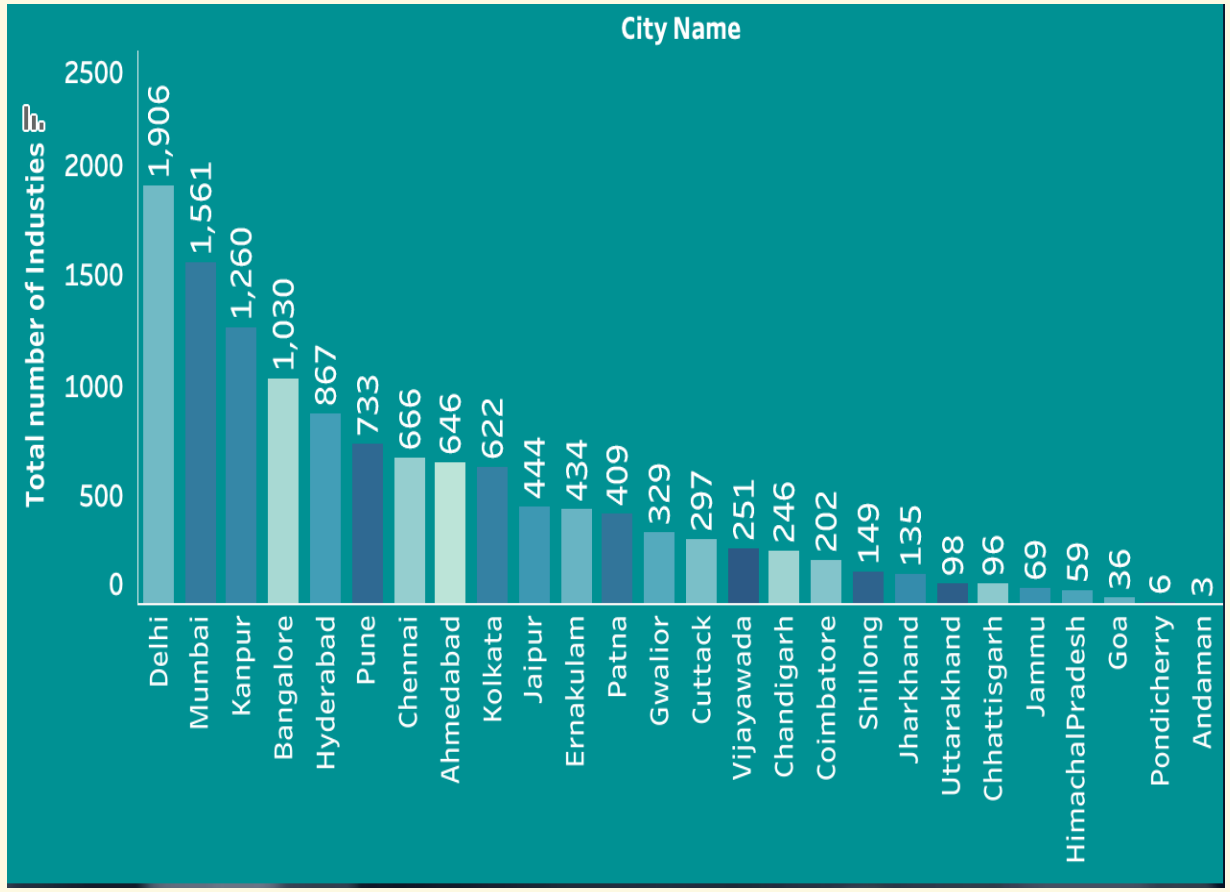
```
df5.show()
companydf=df5.select("Company Name", "Location", "Salary")
companydf1=companydf.withColumnRenamed("Location","City Name")
display(companydf1)
companydf1.count()
```

Data Visualization

Data Visualization

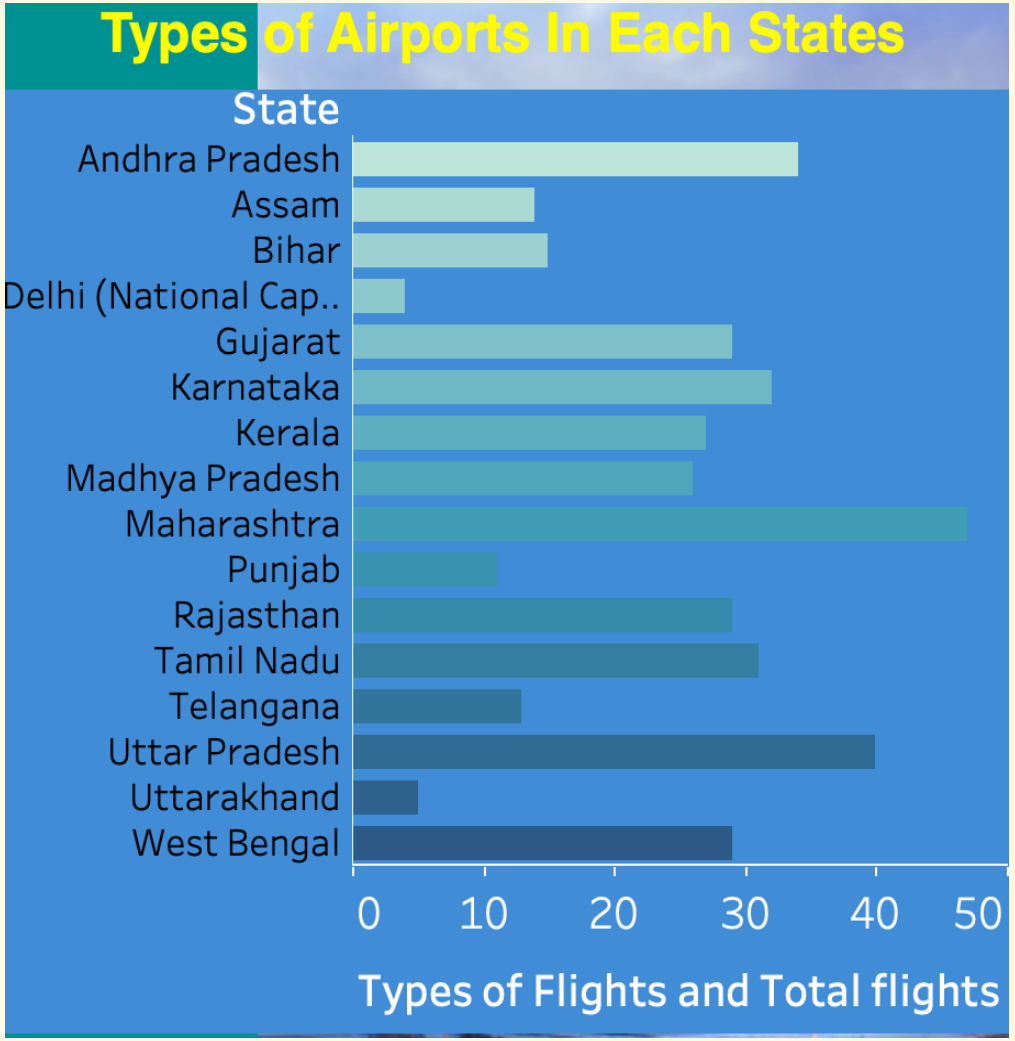
- **Tourism plays an important role in a city's economy.**
- **It is evident that there are some tourist places in tier-II cities like: Vadodara, Indore, Jaipur etc.**
- **But people prefer to visit Tier-I since the mode of transport is highly available and easy commuting.**

Industries in India



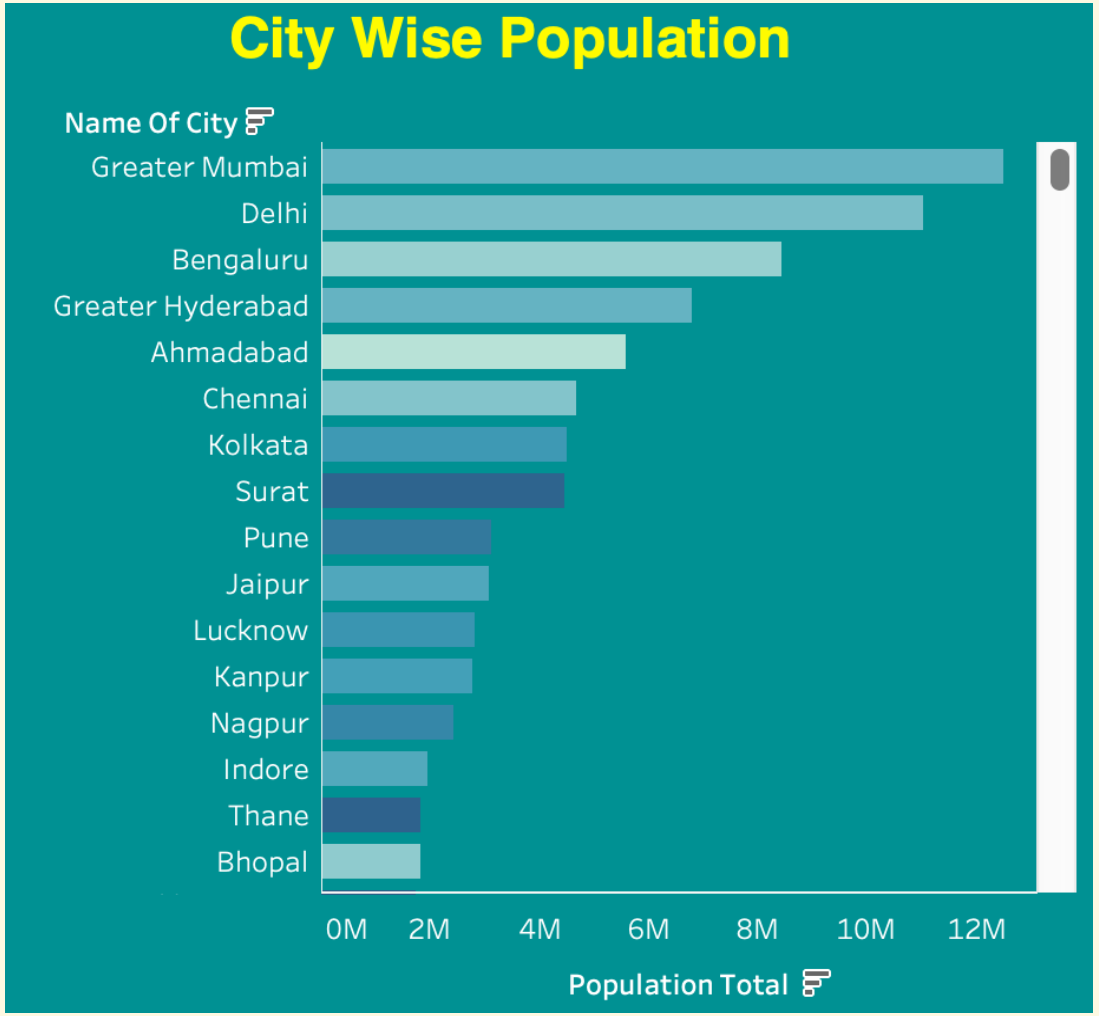
- **From the visualization we can say, Delhi has more industries and Andaman has the least, This might be the major factor why people's population is lesser in tier-II**

Airports in Individual States



- There are many airports in all major and minor cities where we can say, tier-II cities have many airstrips , heliport etc. But unfortunately, many of them are closed and it is kind of evident from our visualization.

City-wise Population



- **From this, we can conclude population is higher in tier-I rather than tier-II.**

Conclusion

- **After all the visualization , we can conclude that , tourism, GDP, Airport connectivity, Industries plays a major role in macro-economy.**
- **And People tend to be in comfort zone and look for quick commuting for daily life.**
- **It is very much evident that , people choose the cities which has more industry.**
- **Cities which have very good monuments are helping to attract the international tourist to visit and helps in economic growth.**

From all the research and analysis we did, and putting up the problem statement "To improve and identify the new air travel routes" we are concluding that to improve the air travel the government must invest half of its high GDP cites economy to the tier-II cities, this will make the people to travel to those cities and which may lead to more economic gain for each state. And it is also notable that all tier-II cities too have the tourist places to visit but the major drawback is the mode of transport. So, we must look to improve the transportation for easy communication to nearby urbanized cities. We also noted that there is many airports, airstrips, heliports which has been closed and not maintained well. If we are able to re-open it for small airports, use it for the transit purpose and it will automatically lead to new air travel routes.

Thank you

Karthik.arunkumar@agilisium.com

Bhagyashree.gohiya@agilisium.com

Udatha.venu@agilisium.com