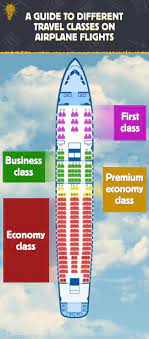
**PROJECT ON**

**EXPLORATORY DATA ANALYSIS ON FLIGHT**



**SUBMITTED BY:**

**NAME:** AKSHAYKUMAR.C

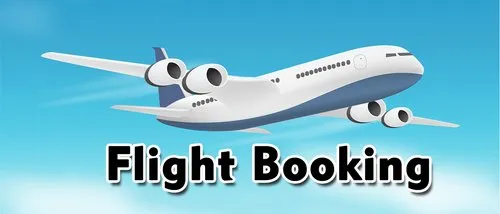
**EMP.ID:** EZ-1041

**DESIGNATION:** ASSOCIATE SOFTWARE ENGINEER

**MOB:**9611653780

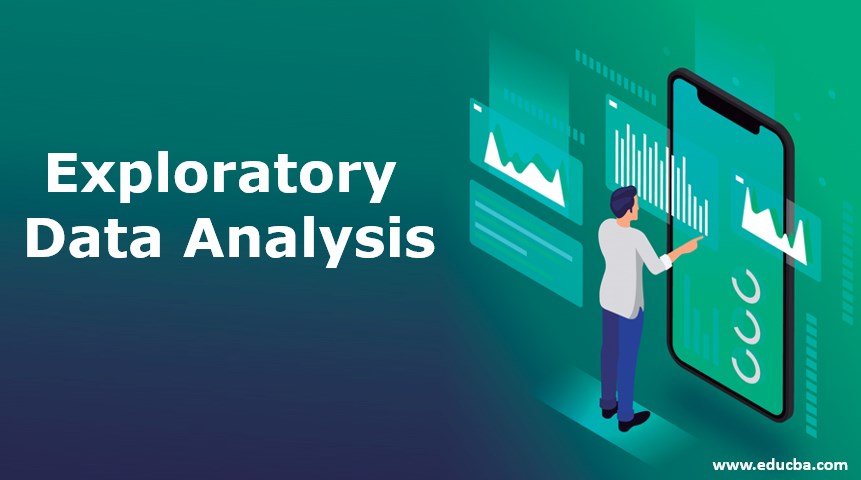
**CLIENT REQUIREMENT**





**NEED TO FIND THE PRICES OF FLIGHTS AND WHICH FLIGHT HAS TOP NUMBER OF BOOKINGS.**

**EXPLORATORY DATA ANALYSIS (EDA)**



**EDA:** IT IS A DATA ANALYTICS PROCESS TO UNDERSTAND THE DATA IN DEPTH AND LEARN THE DIFFERENT DATA CHARACTERISTICS ,OFTEN WITH VISUAL MEANS.

**THE FOUR MAJOR STEPS IN EDA :**

1. CLIENT REQUIREMENT CHECKING
2. DATA PRE-PROCESSING /DATA CLEANING
3. DATA VISUALIZATION
4. CONCLUSION FOR CLIENT

**DATA PRE-PROCESSING /DATA CLEANING**

**STEP 1:** IMPORTING LIBRARIES

**PANDAS:** IT IS PYTHON LIBRARY USED FOR WORKING WITH DATA SETS.

IT HAS FUNCTIONS FOR ANALYZING ,CLEANING ,EXPLORING,MANIPULATING DATA.

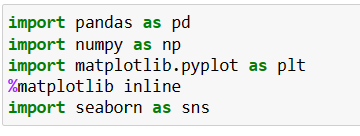
**NUMPY:** IT IS PYTHON LIBRARY USED FOR WORKING WITH ARRAYS.

NUMPY STANDS FOR NUMERICAL PYTHON.

**MATPLOTLIB:** IT IS A COMPREHENSIVE LIBRARY FOR CREATING STATIC ,ANIMATED ,INTERACTIVEVISUALIZATIONS IN PYTHON.

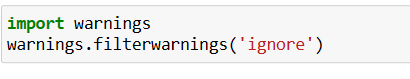
(**INLINE**: IT IS USED TO UPDATE IN THE SAME GRAPH /DATAFRAME)

**SEABORN**: IT IS A LIBRARY FOR MAKING STATISTICAL GRAPHICS IN PYTHON.



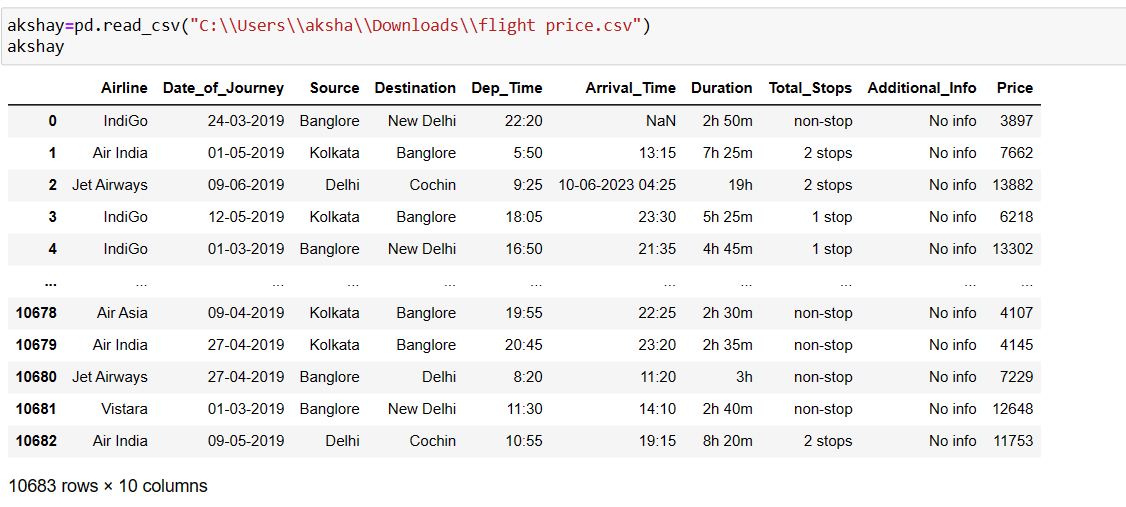
**STEP 2:** IMPORTING WARNINGS

WARNING MESSAGES ARE NORMALLY WRITTEN TO SYS.STDERR,BUT THEIR DISPOSITION CAN BE CHANGED FLEXIBLY, FROM IGNORING ALL WARNINGS TO TURNING THEM INTO EXCEPTIONS.



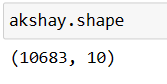
**STEP 3**: READING /IMPORTING DATA

USING PANDAS LIBRARY WE CAN IMPORT THE DATA , COPYING ITS PATH , BY CREATING / TAKING ANY VARIABLE NAME , STORING TO IT , AND THEN CALL THAT VARIABLE NAME .HERE WE MENTIONED PANDAS AS PD .



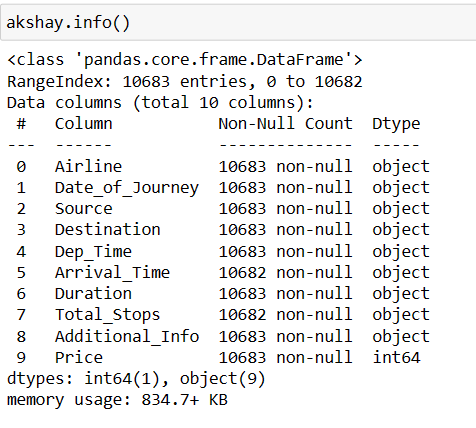
**STEP 4:** SHAPE OF THE DATAFRAME .

INTINAL SHAPE OF THE DATAFRAME CAN BE OBTAINED USING SHAPE , WE NEED TO DO THIS STEP BECAUSE AFTER REMOVING NULL /DUPLICATED VALUES …FOR CONFIRMATION PURPOSE ..WHELTHER THE NULL AND DUPLICATED VALUES ARE REMOVED .

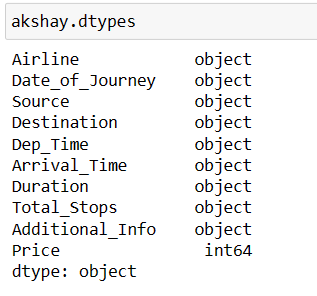


**STEP 5:** CHECKING THE INFORMATION OF THE DATAFRAME .

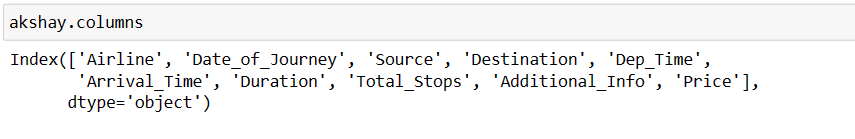
BASICALLY WE SHOULD CHECK THE INFORMATION OF THE CLIENT DATA … BY USING DATAFRAME NAME . INFO()…… WE CAN GET /FETCH THE RANGE INDEX



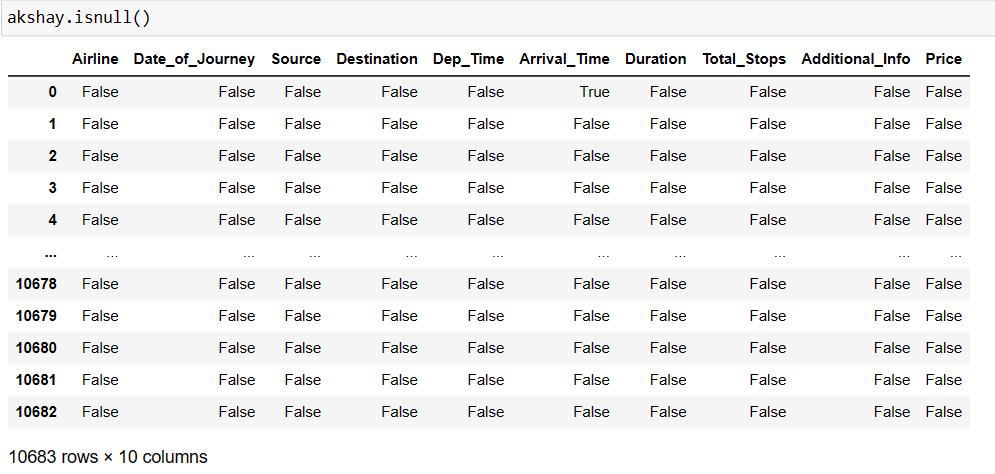
**STEP 6:** CHECKING THE DATA-TYPES OF THE DATA FRAME



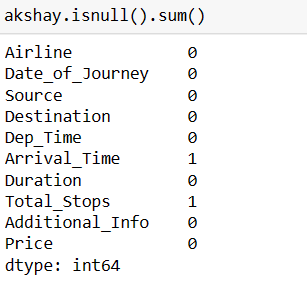
**STEP 7**: CHECKING THE COLUMNS NAME OF THE DATA-FRAME



**STEP 8:** CHECKING THE NULLVALUES IN THE DATA FRAME



**STEP 9:** FETCHING THE SUM OF NULL VALUES



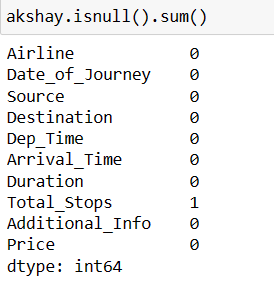
**STEP 10:** FILLNA PROCESS

FOR NAN VALUES WE ARE FILLING WITH “13:10” VALUE USING FILLNA FUNCTION ,FOR THE COLUMN “ARRIVAL\_TIME”.



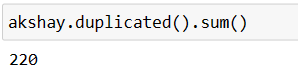
**STEP 11:** AFTER PROCESS , FETCHING SUM OF NULL VALUES.

AFTER THE REMOVAL OF NAN VALUES OR REPLACE NAN VALUES WITH UR SPECIFIED VALUES ..WE ARE FETCHING SUM OF NULL VALUES FOR DIFFERENT COLUMNS .



**STEP 12:** DUPLICATED VALUES

FETCHING DUPLICATED VALUES SUM FOR REMOVING DUPLICATED VALUES .



**STEP 13:** CLEARING THE DUPLICATED VALUES

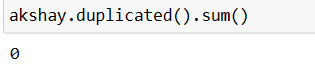
AFTER FETCHING DUPLICATED VALUES SUM, WE WILL DROP THE DUPLICATED VALUES USING DROPING METHOD …



**STEP 14:** AFTER THE PROCESS, CHECKING DUPLICATED VALUES SUM

FETCHING DUPLICATED VALUES SUM AND REMOVING DUPLICATED VALUES USING DROP METHOD .

FOR OUR CLARIFICATION WE AGAIN CHECK THE DUPLICATED VALUES SUM .



**STEP 15**: GRAPHICAL REPRESENTATION .

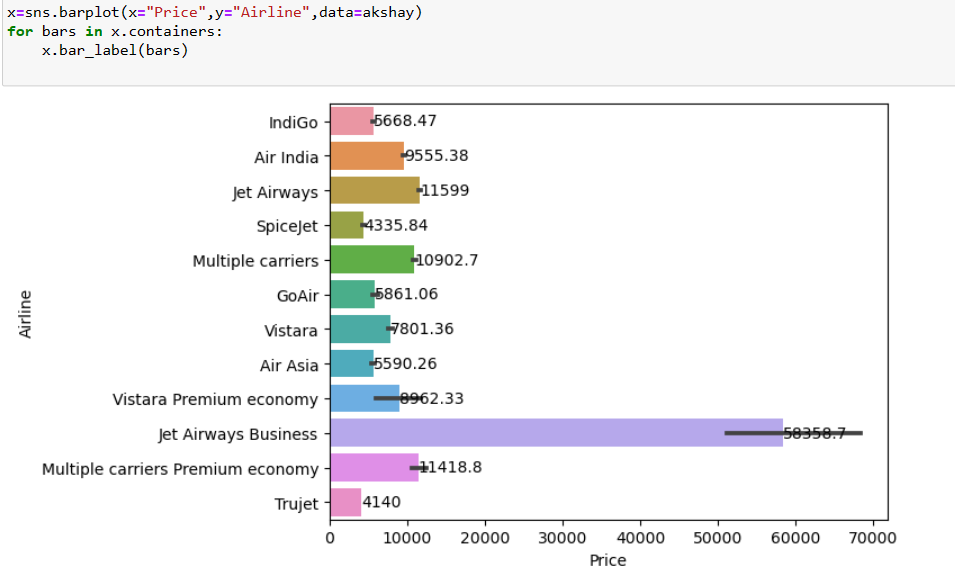
**DATA VISUALIZATION**

DATA VISUALIZATION IS THE REPRESENTATION OF DATA THROUGH USE OF COMMON GRAPHICS, SUCH AS CHARTS, PLOTS, INFOGRAPHICS, AND EVEN ANIMATIONS. THESE VISUAL DISPLAYS OF INFORMATION COMMUNICATE COMPLEX DATA RELATIONSHIPS AND DATA-DRIVEN INSIGHTS IN A WAY THAT IS EASY TO UNDERSTAND.

**SOLUTION FOR CLIENT REQUIREMNET**

**CLIENT REQUIREMENT : NEED TO FIND THE PRICES OF FLIGHTS**

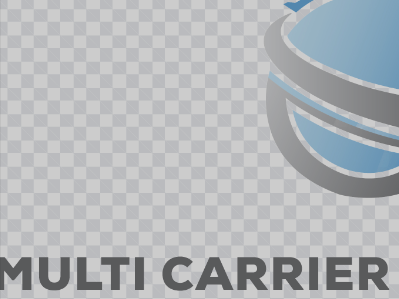
IM REPRESENTING THE PRICE OF EACH FLIGHTS , BY USING DATA VISUALIZATION PROCESS ,IN THAT IM USING BARPLOT , BY GIVING X-AXIS NAME AS: “PRICE”, Y-AXIS NAME AS: “AIRLINE” , FROM THE DATA : “AKSHAY”.



**THE PRICES OF FLIGHTS**

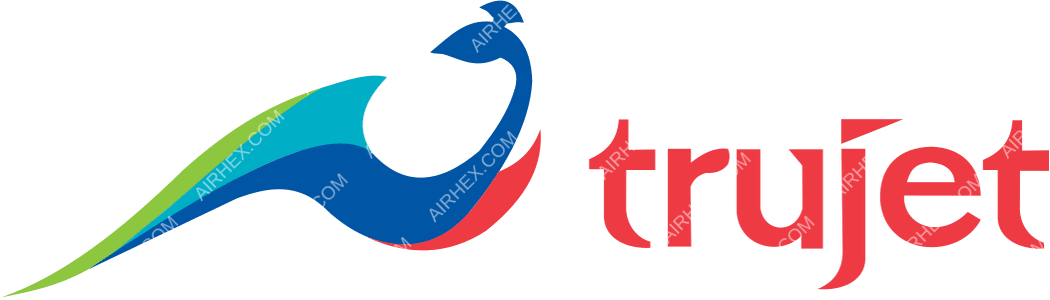
**6668.47 9555.38 11599**

**4335.84 10902.7 6861.06**

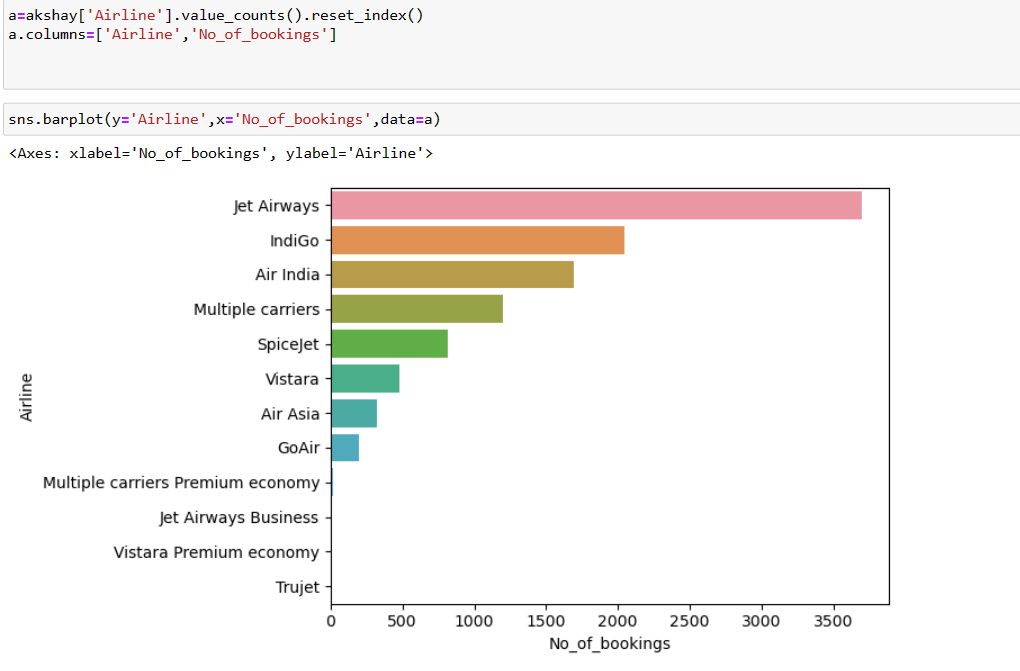
  

**7801.36 6590.26 8962.33**

**58358.7 11418.8 4140**

**CLIENT REQUIREMENT : NEED TO FIND WHICH FLIGHT HAS TOP NUMBER OF BOOKINGS.**



**CONCLUSION :**



**THE JET-AIRWAYS AS MORE NUMBER OF BOOKINGS**



**UNDER THE GUIDANCE OF :**

**SAI CHARAN**

TECHINCAL HEAD

EINFACHZAUBERER TECHNOLOGIES

PVT LTD, BANGALORE.