PROJECT TITLE

Name : KUBERAN T

Register No : 111421104303

Department :cse

Year : 3RD THIRD

College Name::PRATHYUSHA ENGINEERING COLLEGE

ABSTRACT

1. Problem Statement: The flight dataset comprises comprehensive aviation data, departure and arrival times, aircraft type and passenger counts.

2. Proposed Solution: The proposed models and algorithms and the user satisfaction achieved through the passenger service recommendation system.

INTRODUCTION - 2 SLIDES

The flight dataset is a comprehensive collection of aviation-related data that captures various aspects of flights, airlines, airports, and related information. This dataset is valuable for analyzing trends in air travel, understanding flight patterns, assessing airline performance, and conducting research on aviation operations. It contains data points such as departure and arrival times, flight durations, airline and aircraft details, origin and destination airports, and more. Researchers, analysts, and aviation enthusiasts can leverage this dataset to gain insights into the dynamics of air travel and make informed decisions within the aviation industry.

LITERATURE SURVEY - 10 PAPERS

S NO	FLIGHT DATASET	PROPOSED	LIMITATION	JOURNAL/ CONFERENCE NAME	YEAR OF PUBLICATION
1	Quallity assessment for liked data a survey	Where **flights** can periodically broadcast their current sequences (the training **data set** is half the size of full **data set**).	Given the test **data set** of n observations, of the model refers to the mean of the absolute values of each prediction error (residual) on all instances of the test **data set**.	D**ataset** on the ethnicity of refugees, covering the years this study analyzes refugee **flight** such as ethnicity, are essential to understanding the **flight** direction of refugees.	2021
2	Flight abroad and international credit cycle	Method **proposed** in this paper consists of the following. First, to cope with the high heterogeneity of parameters.	**A set** of **flights** where the trajectories executed by the ornithopter were imitated multirotor.	F**light** operations quality assurance program, also called **flight data** monitoring in Europe, aims to use **data.** current programs, the analysis of **flight data** is conducted	2021

3	Predicting flight delay based on multiple linear regression	In this work, we **propose** a crowd detection method for drone **flight** safety..	Besides the French **data set**, we a **data set** with three days of **flights** over second **limitation**.l	**Flight data** recorders generate **data** that is collected on an embedde **data** to a remote cloud storage system, so the **data** memory device will be unbounded.	2021
4	Flight departure time prediction based on deep learning	**Flight data sets** illustrate that the **proposed** algorithm can accurately and quickly detect and identify anomalous sources in **flight data.**	For the mica and **data sets**, a selection of peaks variance in the detection **limit** for variation in the **data set** detection.	We begin by outlining our required **data** products and methodology. We also consider fits to the entire B03 **data set data** alone, and a combined	2022
5	Impact of operational flight predictability on airline fuel cost	The purpose of this paper is to **propose** and change detection.	Coefficients from free-**flight data** introduce assumptions and **limitations** into the equations Experimental free-**flight data** are then "fit" by adjusting coefficients and initial condition.	The results of the **flight data set** contain information about the **flight** code, aircraft code, airline name, departure airport, departure city, arrival airport, arrival city, date/time of departure.	2021

LITERATURE SURVEY - 10 PAPERS

S NO	FLIGHT DATASET	PROPOSED	LIMITATION	JOURNAL/ CONFERENCE NAME	YEAR OF PUBLICATION
6	Agial object trajectory classification by training of fulight controller	To address this problem, we **propose** a novel incremental anomaly detection method digital **flight data**.	Traditionally, this is performed by checking exceedances in pre-**set limits** to the **flight data Flight data** monitoring is the analysis of **flight data** recorded by the digital **flight data**.	Spectrum of the Boomerang. A test **flight** to constrain the curvature of the universe. Given that we have based our results on this **data set** alone, our results are completely independent.	2021
7	Design and research of intelligent as system for flight rew operating manual	P**roposed** model on **flight dataset dataset**. In order to create balance Due of imbalanced **data sets**.	One **flight** is eliminated from the **data set** because there was no clear decision. unstable when this **limit** is exceeded	This article offers a linear regression model to forecast **flight** searches. Experiments on realistic **data sets** of domestic suggested model.	2022

8	Crow sourced Air traffic data from the opensky	In the **data** acquisition system. The novelty of the work presented in this paper lies in **proposing.**	**Flight** studies are performed in still air or artificially smooth flow. Here we show that variability in external airflow.	Effect capital **flight** flows this question using a new panel **dataset** test the hypothesis that inflation has a positive differential effect on capital **flight** after war.	2021
9	A transferable deep network of aircraft engine damange	The **flight** delay is difficult to predict, this study **proposes** a method to model the arriving. This reduces the size of our **data set** to around, records. For each record.	An obvious, albeit self-imposed, **limitation** is that only the period of **flight** from to the same feature **set** could identify unusual occurrences in other **flight** phases eg departures.	In the present paper, new field **data** are added to the original **data set**, bringing the number of species to, on all of which wing measurements and wingbeat frequencies have been .	2021
10	Flight a case study of harts process connect fileld	The **proposed** algorithm on experiment that combines estimation, with control and planning.	Since, both in PPM and SM, PICsIT single and multiple events are contained in two different and independent **data sets**.	Such system generates a common model based on the learning **data** so that it can predict the results of new **data sets**.	2022

Problem definition

A flight dataset typically refers to a collection of structured data containing information about various aspects of flights. This dataset can be used for analysis, prediction, optimization, or other data-driven tasks related to the aviation industry. The problem definition can vary based on the specific objectives of your analysis. Here are a few potential problem definitions that can be tackled using a flight dataset:

Dataset description

About the Dataset

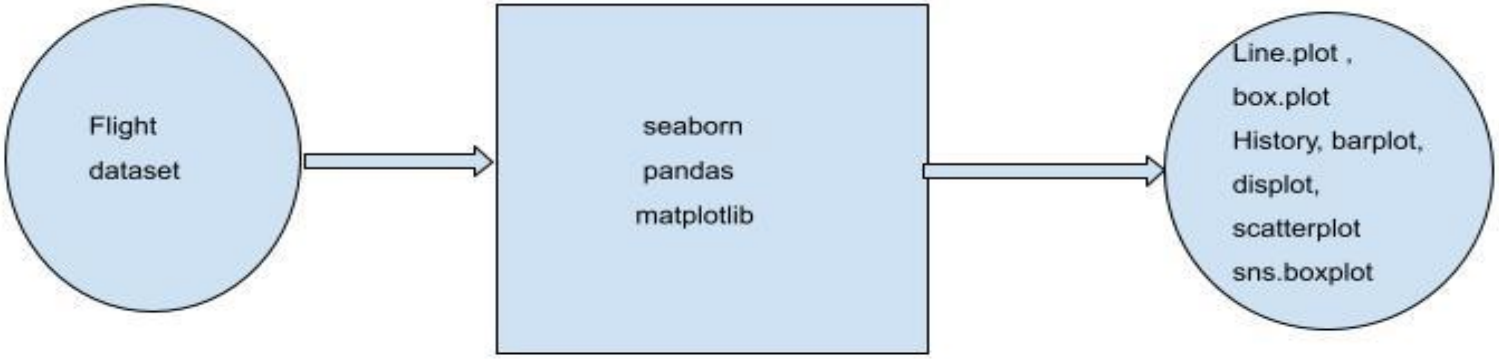
The "Flight Data Set" is a collection of aviation-related data containing information about flights, airlines, airports, and associated details. It likely includes data such as flight numbers, departure and arrival times, origin and destination airports, airline information, aircraft types. Researchers, analysts, and aviation professionals can use this dataset to gain insights into flight scheduling, delays, connections, and other aspects of air travel.

1. No.of. Rows - 12 rows
2. No.of. Columns - 3 columns

Columns description

Column Name	Data Type	Values(eg:0/1 , Multiple values, range)
year	integer	discrete
month	category	categorical
passengers	integer	discrete

SYSTEM ARCHITECTURE



Modules

PANDAS

Pandas is a popular python library used for data manipulation and analysis. It provides powerful and flexible data structures. The main data structure in pandas is the DataFrame, which is a two-dimensional table of data with labeled rows and columns.

MATPLOTLIB

Matplotlib is a widely-used data visualization library in python. It provides and python. With Matplotlib, you can create a wide range of plots, charts, histograms, scatterplots, and so on.

SEABORN

Seaborn is a python data visualization library built on top of Matplotlib. It provides a high-level interface for creating attractive and informative statistical graphics.

Pandas Methods & Attributes

1. LINE.PLOT

2. BOX.PLOT

3.HISTOGRAM.PLOT

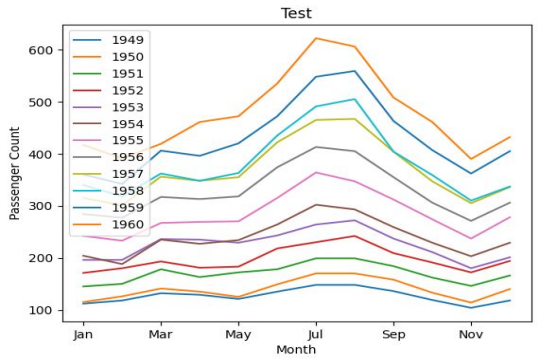
4.BAR.PLOT

5.BIS.PLOT

6.SNS.BOX.PLOT

7.SCATTER.PLOT

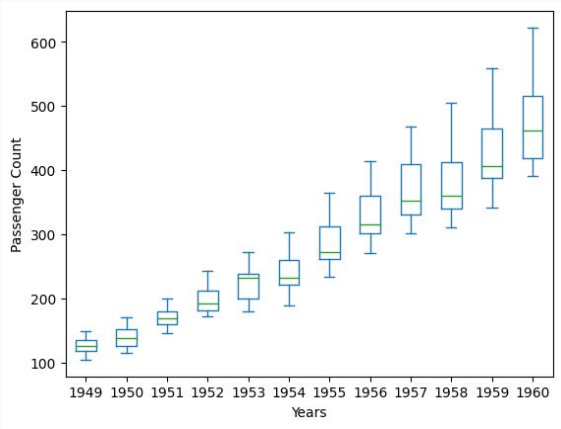
Matplotlib Plots



Inference: **PLOT LINE**

From this graph x is a months next y is a passenger counts list graph name is line plot test

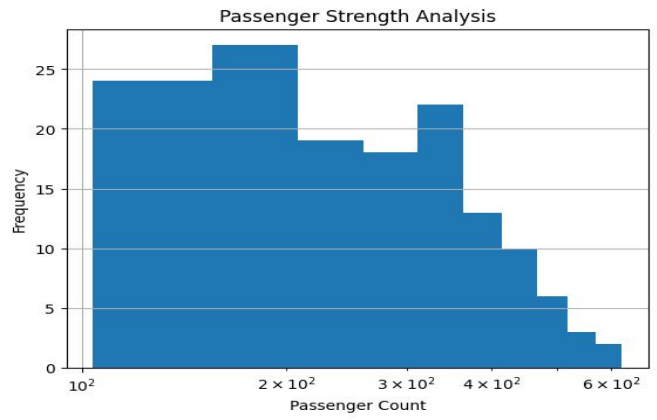
Matplotlib Plots



INFERENCE: BOX PLOT

From this graph x is a year next y is a passenger counts list graph name is box plot

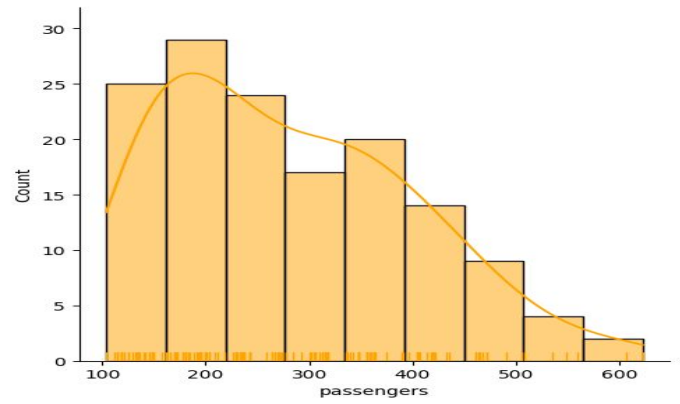
Matplotlib Plots



INFERENCE: HISTOGRAM PLOT

From this plot x is a passenger count flight dataset y is a frequency line it's used to passenger strength analysis

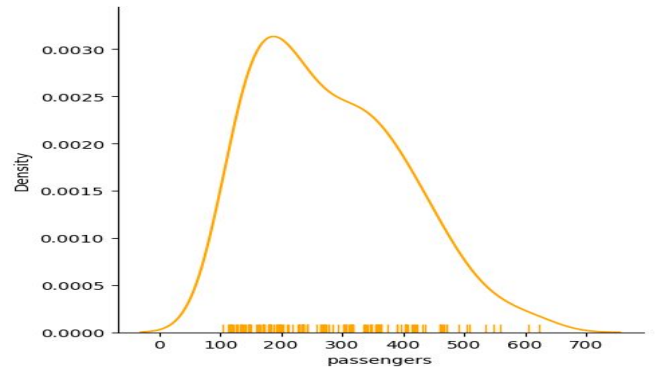
Seaborn



INFERENCE:BIS PLOT

From is x is a passenger list of bis plot and y is a density plot is used to seaborn plot

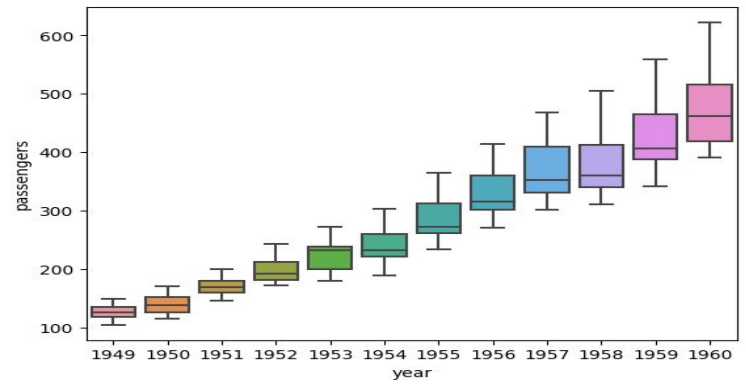
Seaborn



INFERENCE: BAR PLOT

In this plot x is a year and y passengers of flight data set it is a bar plot seaborn connect to the searborn bar

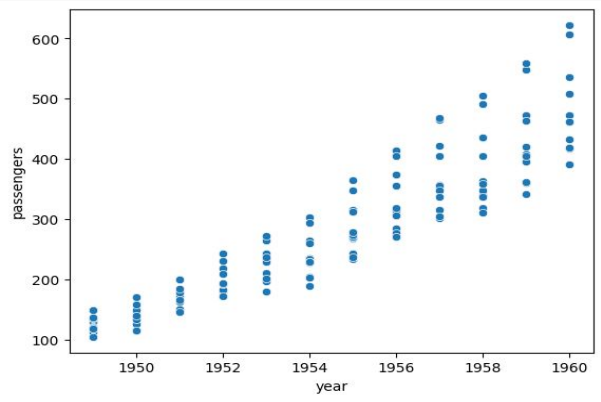
Seaborn



INFERENCE: BOX PLOT

From this graph x is a year and denotes of y is a passenger list is used to seaborn list connected on box plot

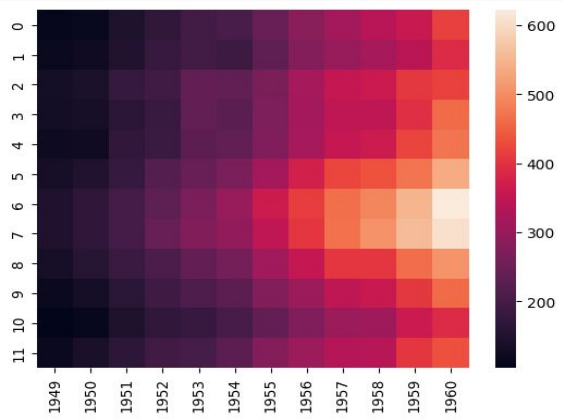
Seaborn



INFERENCE: SCATTER PLOT

From this graph x is a year and y is a passenger graph is used to increment of the list seaborn connected to scatter plot

Heatmap



INFERENCE: HEATMAP

Heatmap is a graphical representation of data where values of represented as x is a year passengers and y is a first passenger and second is a year connected is used to flight data set connect different of heatmap points connected

Conclusion

In conclusion, the flight dataset analysis has provided actionable insights that can drive strategic decisions across various aspects of the airline industry. By leveraging these insights, airlines and airports can enhance operational efficiency, optimize pricing experience for passengers. Further research and analysis could delve deeper into specific aspects highlighted by the dataset, contributing to ongoing improvements within the aviation sector.

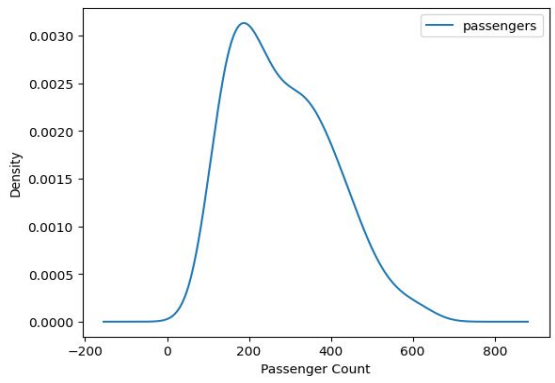
REFERENCE (minimum 7 papers)

1. Davis III, Roy B., et al. "A gait analysis data collection and reduction technique." *Human movement science* 10.5 (1991): 575-587.
2. Cox, Robert W. "AFNI: software for analysis and visualization of functional magnetic resonance neuroimages." *Computers and Biomedical research* 29.3 (1996): 162-173.
3. Kirschvink, JLhttps. "The least-squares line and plane and the analysis of palaeomagnetics data." *Geophysical Journal International* 62.3 (1980): 699-718.
4. Ionescu, Catalin, et al. "Human3. 6m: Large scale datasets and predictive methods for 3d human sensing in natural environments." *IEEE transactions on pattern analysis and machine intelligence* 36.7 (2013): 1325-1339.
5. Bandettini, Peter A., et al. "Processing strategies for time-course data sets in functional MRI of the human brain." *Magnetic resonance in medicine* 30.2 (1993): 161-173.
6. Ravel, Bruce, and M. A. T. H. E. N. A. Newville. "ATHENA, ARTEMIS, HEPHAESTUS: data analysis for X-ray absorption spectroscopy using IFEFFIT." *Journal of synchrotron radiation* 12.4 (2005): 537-541.
7. Fixsen, D. J., et al. "The cosmic microwave background spectrum from the full cobe* firas data set." *The Astrophysical Journal* 473.2 (1996): 576.

CODE

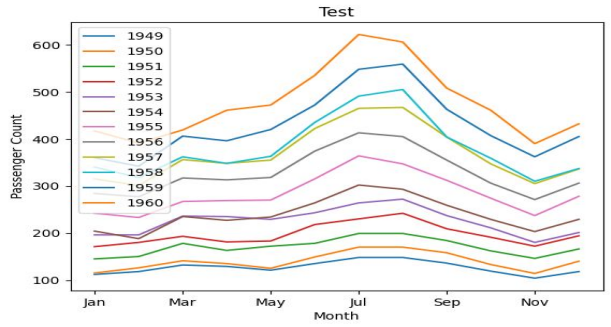
```
data = {}
year = []
for i in data_frame.values:
  if i[0] not in year:
    data[i[0]]=[i[2]]
    year.append(i[0])
  else:
    data[i[0]]+=[i[2]]
df = pd.DataFrame(data)
data
```

```
data_frame['passengers'].plot(kind="density")
plt.legend()
plt.xlabel("Passenger Count")
```

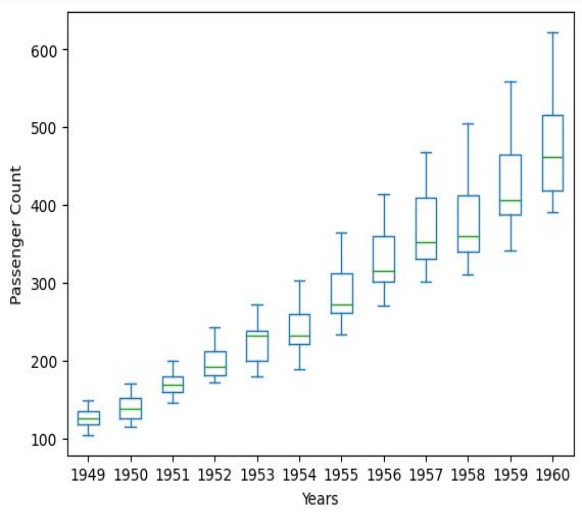


SCREEN SHOTS - OUTPUT

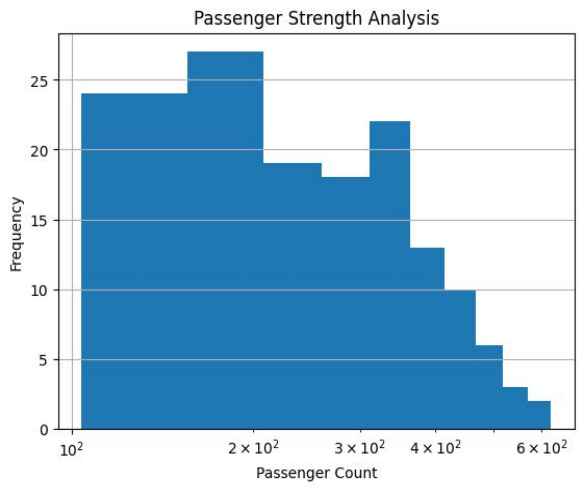
```
   df.plot.line()
plt.xlabel("Month")
plt.ylabel("Passenger Count")
plt.xticks([0,2,4,6,8,10],['Jan','Mar','May','Jul','Sep','Nov'])
plt.title("Test")
```



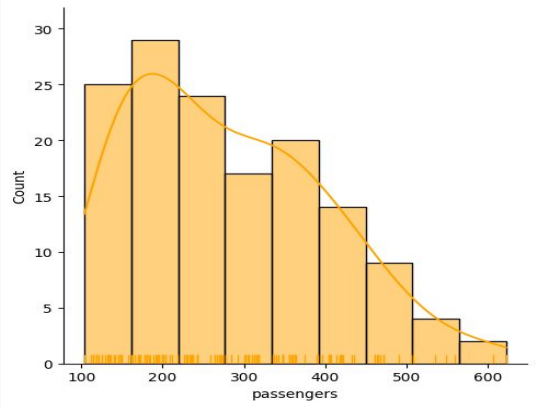
```
df.plot.box()
plt.xlabel("Years")
plt.ylabel("Passenger Count")
```



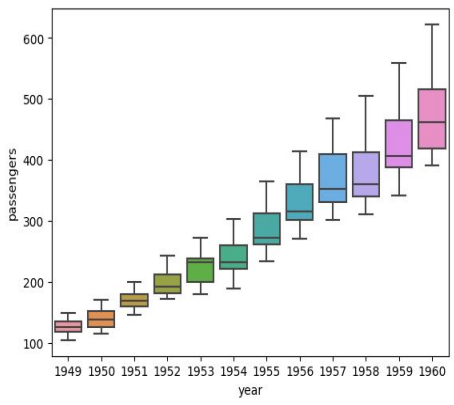
```
plt.hist(data_frame['passengers'])
plt.xlabel("Passenger Count")
plt.ylabel("Frequency")
plt.title("Passenger Strength Analysis")
plt.grid()
plt.xscale('log')
```



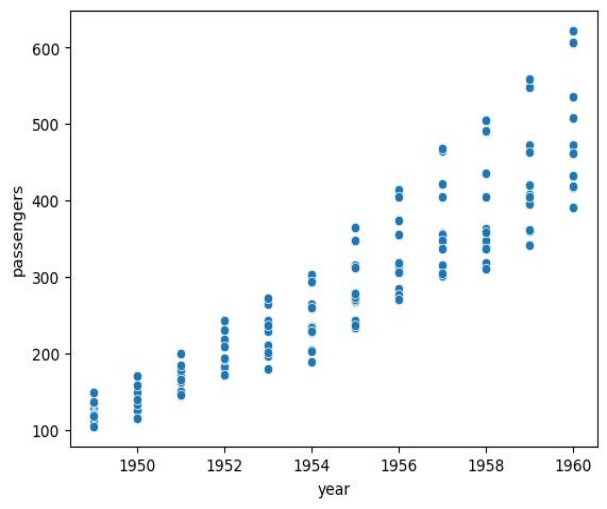
```
sns.displot(data=data_frame,x="passengers",kde=True,rug=True,color="orange"
```



```
sns.boxplot(data=data_frame,x="year",y="passengers")
```



```
sns.scatterplot(data=data_frame,x="year",y="passengers")
```



```
sns.heatmap(df)
```