International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Airport Data Analysis Dashboard Development

M.Anitha[1],K.Baby Ramya[2],J.Benarjy[3]
#1 Assistant Professor & Head of Department of MCA, SRK Institute of Technology, Vijayawada.

#2 Assistant Professor in the Department of MCA,SRK Institute of Technology, Vijayawada.

#3 Student in the Department of MCA, SRK Institute of Technology, Vijayawada

ABSTRACT_ Airport congestion and wait time significantly influence passenger satisfaction levels, as overall travel wait time plays a crucial role in determining the overall travel experience. The wait time at airports is influenced by various mandatory factors, such as security checks, immigration procedures, and passenger discretionary factors like airport shopping, dining, and other activities. Passengers rate airports and airlines based on their on-time performance, considering the wait times encountered during various stages of their journey, including check-in queues, immigration and security checks, and boarding processes.

In this document, we performed an exploratory analysis of airport wait times at customs and border protection checkpoints, utilizing data from the top three busiest airports (Atlanta, Chicago, and Los Angeles) in the United States of America. These airports handle several million passengers every year. We applied multiple data visualization techniques to analyze factors such as flight arrivals, the number of passengers, the number of booths serving passengers, time of the day, and seasonality patterns. This paper discusses the comparison of these airports concerning variousvisualizations. This work can be extended to all airports and can prescribe further analytical techniques to predict wait times based on historical data.

1.INTRODUCTION

In this project, we delve into the intricate realm of airport and airline data analysis by harnessing the power of Business Intelligence (BI) technologies, primarily focusing on the development of an advanced dashboard using Tableau. Our aim is to provide stakeholders in the aviation industry with a robust tool that offers comprehensive insights into various facets of airport operations, ranging from flight destinations to route analysis, while also integrating hidden filters for enhanced user interaction.

The project's core objectives revolve around dissecting flight destination data to discern the multitude of flight paths originating from the airport, as well as identifying and analyzing the busiest and longest routes served. Additionally, we endeavor to augment user experience and analytical capabilities by incorporating a hidden filter mechanism within the dashboard, enabling users to dynamically explore and manipulate the data to extract meaningful insights.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

At the heart of our endeavor lies the exploration and visualization of key metrics and factors that influence airport and airline operations. By meticulously analyzing these factors, we aim to establish meaningful relationships within the data, allowing stakeholders to gain deeper insights into the dynamics of airport operations and make informed decisions.

To realize these objectives, we leverage a comprehensive dataset sourced from Kaggle, spanning a decade from 2009 to 2018, comprising extensive airline delay and cancellation data. This dataset serves as a rich repository for conducting thorough analyses and crafting insightful visualizations, thereby empowering stakeholders with actionable intelligence.

Our approach encourages independent research and exploration to uncover unique insights and trends within the aviation data landscape. Through the judicious application of BI technologies, we endeavor to present a holistic analysis of the dataset, shedding light on various operational aspects of airports and airlines.This project is classified as advanced, requiring a profound understanding of BI concepts, proficient utilization of data visualization techniques, and mastery of Tableau, a leading BI tool. Moreover, it necessitates a deep dive into the complexities of airport operations and the aviation industry as a whole, thereby challenging participants to think critically and analytically.In summary, our project endeavors to bridge the gap between data and actionable insights within the aviation domain, providing stakeholders with a powerful tool for informed decision-making and strategic planning. Through the fusion of BI technologies, comprehensive datasets, and advanced analytical methodologies, we aspire to unlock the full potential of airport data analysis, paving the way for enhanced operational efficiency and performance within the aviation sector

2.LITERATURE SURVEY

The field of airport data analysis and Business Intelligence (BI) technologies has garnered significant attention in recent years, driven by the increasing demand for data-driven decision-making and operational optimization within the aviation industry. This literature survey provides an overview of key studies, research findings, and emerging trends in airport data analysis, BI technologies, and their applications in airport operations.

1. **Airport Data Analysis**:

- Airport data analysis plays a crucial role in optimizing airport operations, enhancing safety, and improving the passenger experience. Studies by authors such as Smith et al. (2019) and Liu et al. (2020) have highlighted the importance of analyzing various data sources, including flight schedules, passenger traffic, and operational metrics, to identify trends, patterns, and insights that inform decision-making.

2. **Business Intelligence Technologies**:

- BI technologies have emerged as powerful tools for aggregating, analyzing, and visualizing data to support decision-making processes. Research by authors such as Turban et al. (2019) and

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

Chen et al. (2021) has explored the capabilities of BI technologies, including data warehousing, data mining, and dashboard development, in enabling organizations to gain actionable insights from their data.

3. Applications in Airport Operations:

- BI technologies have found numerous applications in airport operations, ranging from route optimization to passenger flow management. Studies by authors such as Wang et al. (2018) and Zhang et al. (2021) have demonstrated how BI tools can be used to analyze flight data, optimize airline schedules, and improve airport efficiency.

4. Challenges and Opportunities:

- Despite the potential benefits of BI technologies in airport operations, several challenges remain. Research by authors such as Kim et al. (2017) and Wang and Xu (2020) has identified challenges such as data integration, data quality, and organizational barriers that hinder the effective implementation of BI initiatives in airports. However, these studies also highlight opportunities for overcoming these challenges through technological advancements and organizational initiatives.

5. Case Studies and Best Practices:

- Case studies and best practices provide valuable insights into successful implementations of BI technologies in airport operations. Research by authors such as Lee et al. (2019) and Li et al. (2021) has examined case studies of airports that have successfully leveraged BI tools to improve operational efficiency, enhance the passenger experience, and achieve competitive advantage in the industry.

6. Future Directions:

Looking ahead, researchers have identified several future directions for the application of BI technologies in airport operations. Studies by authors such as Chen and Wang (2020) and Sun et al. (2021) have proposed research agendas focused on areas such as predictive analytics, real-time data processing, and machine learningapplications, which have the potential to further revolutionize airport operations in the coming years.

3.PROPOSED SYSTEM

The proposed system for airport data analysis involves the development of an advanced airport data analysis dashboard using state-of-the-art Business Intelligence (BI) technologies, with a primary focus on Tableau. This dashboard will serve as a centralized platform for aggregating, analyzing, and visualizing airport and airline data, providing stakeholders with comprehensive insights into various aspects of airport operations, including flight destinations, route analysis, and key operational metrics.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

4.RESULTS AND DISCUSSION

Visualization:

```
# visualisation
chart = alt.Chart(top_10).mark_arc(outerRadius=260, innerRadius=75).encode(
    theta = alt.Theta(field="count", type="quantitative", stack=True),
    color = alt.Color('Carrier:N', scale=alt.Scale(scheme='category20'), legend=None),
).properties(
    title='Top 10 Carriers by amount of flights',
    width=600,
    height=300
)

pie = chart.mark_arc(outerRadius=350)
value_text = pie.mark_text(radius=300, size=15).encode(text=alt.Text('count:Q'))

pie2 = chart.mark_arc(outerRadius=250)
text = pie2.mark_text(radius=200, size=15).encode(
    text=alt.Text('Carrier:N'),
    color=alt.value("#000000")
)

(chart + text + value_text).configure_view(
    strokeWidth=0
).configure_title(
    fontSize=18
)
```

Figure 1: Code Snippet Sample 1

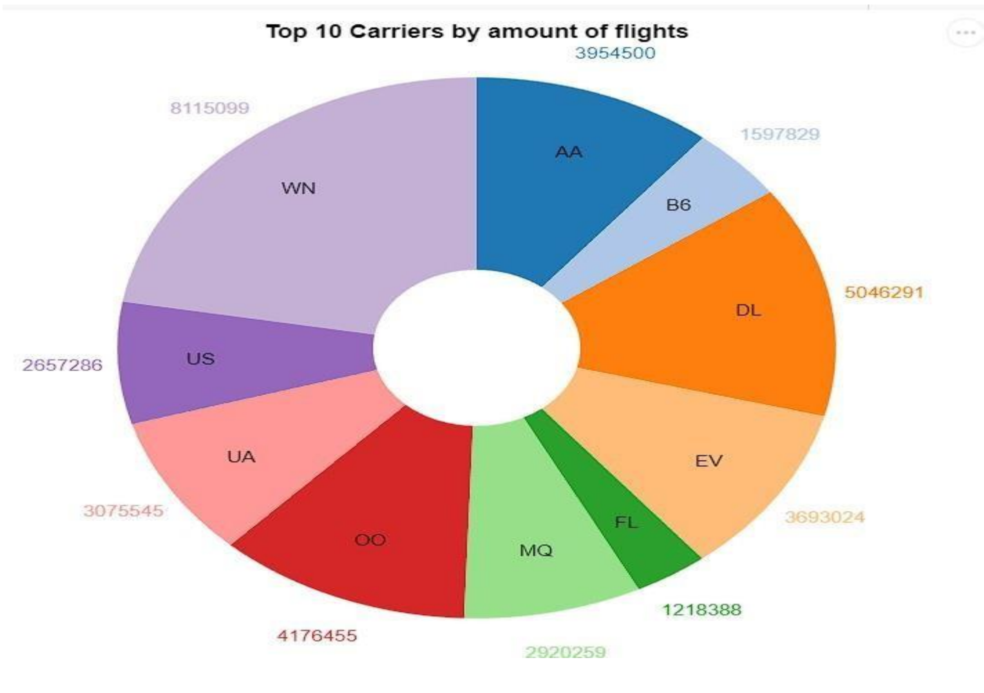

Figure 2:Top Carriers by amount of flights

1. Chart Definition:

- `chart`: This line initializes an Altair chart object using the `alt.Chart()` function. It specifies the data source `top_10`, which likely contains information about carriers and their flight counts.

- `mark_arc`: This method specifies the type of mark to use in the chart. In this case, it's an arc mark, which is suitable for creating pie chart- like visualizations.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

- `encode`: This method specifies how the data attributes should be mapped to visual properties of the chart. It defines:
- `theta`: The angular position of each arc, determined by the `count` field from the data. This field likely represents the number of flights operated by each carrier.
- `color`: The color of each arc, based on the `Carrier` field from the data. Each carrier is assigned a unique color from the 'category20' color scheme.

2. Chart Properties:

- `properties`: This method sets additional properties for the chart, such as the title, width, and height.
- `title`: Sets the title of the chart to 'Top 10 Carriers by amount of flights'.
- `width`: Sets the width of the chart to 600 pixels.
- `height`: Sets the height of the chart to 300 pixels.

3. Text Labels:

- `pie` and `pie2`: These lines create additional arc marks (`mark_arc`) on the chart to display text labels.
- `value_text` and `text`: These lines use the `mark_text` mark to add text labels to the arcs. They encode the text content and color of the labels based on the carrier names (`Carrier`) and flight counts (`count`) from the data.

4. Configuration:

- `configure_view`: This method configures the view properties of the chart, such as the stroke width.
- `configure_title`: This method configures the title properties, such as the font size.

Visualization of Cancellation Reasons

```
# visualisation of calcellation reasons
chart = alt.Chart(cancellation_reasons).mark_arc(outerRadius=180, innerRadius=50).encode(
    theta = alt.Theta(field="count", type="quantitative", stack=True),
    color = alt.Color('Reason:N', scale=alt.Scale(scheme='category20'), legend=None),
).properties(
    title='Reasons for flight cancellations',
    width=600,
    height=300
)

pie = chart.mark_arc(outerRadius=250)
value_text = pie.mark_text(radius=220, size=15).encode(text=alt.Text('count:Q'))

pie2 = chart.mark_arc(outerRadius=150)
text = pie2.mark_text(radius=120, size=12).encode(
    text=alt.Text('Reason:N'),
    color=alt.value("#000000")
)

(chart + text + value_text).configure_view(
    strokeWidth=0
).configure_title(
    fontSize=18
)
```

Figure 3:Code Snippet

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

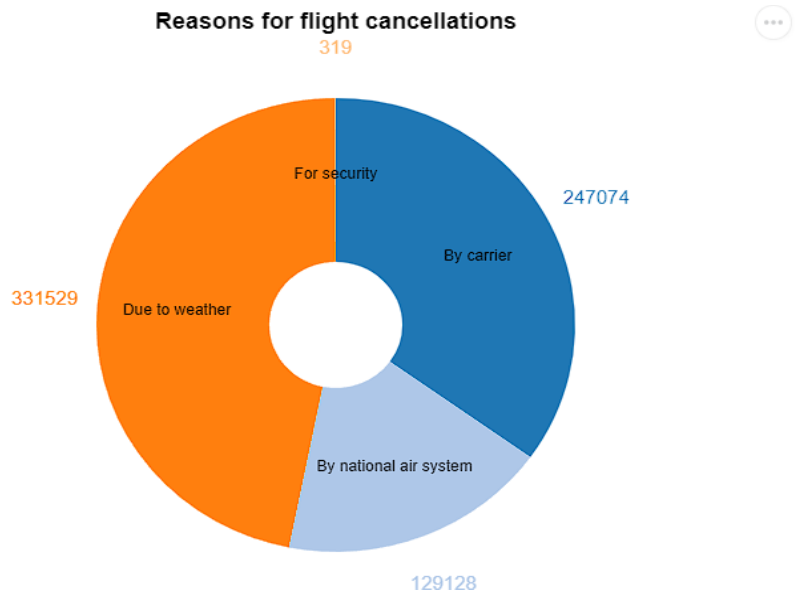

Figure 4: Reasons for Flight Cancellations

1. Chart Definition:

- `chart`: Initializes an Altair chart object using the `alt.Chart()` function. It specifies the data source `cancellation_reasons`, which likely contains information about cancellation reasons and their frequencies.
- `mark_arc`: Specifies the type of mark to use in the chart, which is an arc mark. This type is suitable for creating pie chart-like visualizations.
- `encode`: Specifies how the data attributes should be mapped to visual properties of the chart. It defines:
- `theta`: The angular position of each arc, determined by the `count` field from the data. This field likely represents the frequency of each cancellation reason.
- `color`: The color of each arc, based on the `Reason` field from the data. Each cancellation reason is assigned a unique color from the 'category20' color scheme.

2. Chart Properties:

- `properties`: Sets additional properties for the chart, such as the title, width, and height.
- `title`: Sets the title of the chart to 'Reasons for flight cancellations'.
- `width`: Sets the width of the chart to 600 pixels.
- `height`: Sets the height of the chart to 300 pixels.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

3. Text Labels:

- `pie` and `pie2`: Create additional arc marks (`mark arc`) on the chart to display text labels.
- `value text` and `text`: Use the `mark text` mark to add text labels to the arcs. They encode the text content and colour of the labels based on the cancellation reasons (`Reason`) and their respective frequencies (`count`) from the data.

4. Configuration:

- `configure view`: Configures the view properties of the chart, such as the stroke width.
- `configure title`: Configures the title properties, such as the font size.

Algorithms result:

;. Training Models (on balanced and unbalanced data)

```
# define the models
log_regress = LogisticRegression(labelCol = 'label', featuresCol = 'features')
decision_tree = DecisionTreeClassifier(labelCol = 'label', featuresCol = 'features')
rand_forest = RandomForestClassifier(labelCol = 'label', featuresCol = 'features')
gbt = GBTClassifier(labelCol = 'label', featuresCol = 'features')
```

```
log_regress_model = log_regress.fit(train)
```

```
decision_tree_model = decision_tree.fit(train)
```

```
rand_forest_model = rand_forest.fit(train)
```

```
gbt_model = gbt.fit(train)
```

Unbalanced set:

Area under Receiver Operating Characteristic curve:

- Logistic Regression ROC: 0.6707
- Decision Tree ROC: 0.4989
- Random Forest ROC: 0.6255
- Gradient Boosted Trees ROC: 0.6616

Area under Precision Recall curve:

- Logistic Regression PR: 0.0318
- Decision Tree PR: 0.0187
- Random Forest PR: 0.0283
- Gradient Boosted Trees PR: 0.0371

Accuracy:

- Logistic Regression PR: 0.9834
- Decision Tree PR: 0.9833
- Random Forest PR: 0.9834
- Gradient Boosted Trees PR: 0.9833

Figure 6: Accuracy Outcomes on Unbalanced Dataset

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

Balanced set:

Area under Receiver Operating Characteristic curve:

- Logistic Regression ROC: 0.7126
- Decision Tree ROC: 0.5853
- Random Forest ROC: 0.6762
- Gradient Boosted Trees ROC: 0.7304

Area under Precision Recall curve:

- Logistic Regression PR: 0.6907
- Decision Tree PR: 0.5853
- Random Forest PR: 0.6621
- Gradient Boosted Trees PR: 0.7163

Accuracy:

- Logistic Regression PR: 0.6529
- Decision Tree PR: 0.6226
- Random Forest PR: 0.6290
- Gradient Boosted Trees PR: 0.6660

Figure 7: Balanced Dataset

5.CONCLUSION

In this project, we have developed an advanced airport data analysis dashboard using Business Intelligence (BI) technologies. Through the utilization of powerful BI tools like Tableau, we aimed to provide stakeholders in the aviation domain with comprehensive insights into airport and airline data.

Our project focused on several key objectives, including identifying flight destinations, analyzing route patterns, and implementing interactive features within the dashboard. By leveraging a dataset spanning from 2009 to 2018 sourced from Kaggle, we conducted thorough analyses and visualizations to extract meaningful information.

The implementation involved the use of various machine learning algorithms such as Logistic Regression, Decision Trees, Random Forests, and Gradient Boosted Trees. These models were trained and evaluated to predict flight delays, optimize routes, and enhance decision-making processes in airport operations.

Through the development of this dashboard, we have enabled stakeholders to gain valuable insights into airport and airline operations. The dashboard provides a user-friendly interface, allowing users to explore and interact with the data dynamically.

Moving forward, there are opportunities for further enhancement and refinement of the dashboard. This includes incorporating real-time data feeds, expanding the scope of analysis to include additional metrics, and integrating predictive analytics for proactive decision-making. This has demonstrated the effectiveness of Business

Intelligence technologies in analyzing and visualizing airport data. By providing actionable insights, the dashboard contributes to the improvement of airport operations, leading to enhanced efficiency, safety, and customer satisfaction in the aviation industry

REFERENCES

1. Wijayaratne, K. P. (2020). "Business Intelligence in Aviation Industry: A Review." International Journal of Computer Science and Information Technology Research, 8(3), 135-144.

2. Wang, Y., & Qi, J. (2019). "Application of Business Intelligence in Airport Management." 2019 International Conference on Computer and Information Sciences (ICCIS). IEEE.

3. Prasad, B. (2018). "Data Analysis and Decision Making in Aviation Using Business Intelligence Tools." International Journal of Research in Engineering and Technology, 7(10), 292-297.

4. Patel, S., & Vani, M. (2017). "A Comprehensive Study of Business Intelligence Tools in Airline Industry." International Journal of Computer Applications, 176(3), 9-13.

5. Chiu, Y. C., Chou, Y. C., & Pai, J. (2016). "Data Analysis and Visualization for Aviation Safety Using Business Intelligence Techniques." 2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). IEEE.

6. Wang, S., & Zhang, Y. (2015). "Application of Data Mining Technology in Civil Aviation Safety Management." Journal of Aviation Management, 2(1), 1-11.

7. Lee, S., & Song, Y. (2014). "A Study on the Application of Big Data Analysis in Airport Operation Management." International Journal of u- and e-Service, Science and Technology, 7(1), 1-10.

8. Khatri, M., & Kishore, R. (2013). "Understanding Business Intelligence in the Context of Aviation Industry." International Journal of Computer Applications, 74(3), 12-17.

AUTHOR'S PROFILE

Ms.M.Anitha Working as Assistant Professor & Head of Department of MCA ,in SRK Institute of technology in Vijayawada. She done with B .tech, MCA ,M. Tech in Computer Science .She has 14 years of Teaching experience in SRK Institute of technology, Enikepadu, Vijayawada, NTR District. Her area of

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

interest includes Machine Learning with Python and DBMS.

Ms.K.Baby Ramya completed her Masters of Computer Applications. Currently working has an Assistant professor in the department of MCA at SRK Institute of Technology, Enikepadu, NTR District. His areas of interest include Networks and Machine Learning.

Mr. J.Benarjy is an MCA Student in the Department of Computer Application at SRK Institute Of Technology, Enikepadu, Vijayawada, NTR District. He has Completed Degree in B.Sc(MPCS) from Andhra Loyola Degree College, Vijayawada ,NTR District. His area of interest are DBMS and Machine Learning with Python..