Raul G. Garcia III
Dr. Jeffry Babb, Ph.D.
Business Intelligence & Decision Support Systems – 4310
28 July 2024

FINAL PROJECT - FLIGHT DELAY PREDICTIONS

Problem Domain: Improving Flight Efficiency

Project Name: *AUS Flight Delays*

Project Concept: Develop a dashboard to identify carrier flights at high risk of delays based on origin, delay time historical data, allowing carriers to implement operational improvement strategies to improve flight delay issues.

Supplemental Materials Enclosed:

1. PBIX data models
2. Exported PDF
3. Project Storyboard
3. 4V Model analysis of these sources
4. CRISP-DM diagram modeling how the project will work

PROJECT STORYBOARD: Identify flights at risk of delay for resolution implementation

	Setup	Actions	Outcomes	Results
Current implementation	o Airlines use flight delay data in excel to understand problem areas o The data is manually compiled and analyzed to pull insights where flight delays occur	o Carrier analysts review flight performance periodically o Recommend points of improvement to improve delay times	o Flights at risk are identified based on manual reviews and anecdotal evidence o Carrier efforts to improve customer satisfaction and operational flows based on manually compiled information	o Large data sets are difficult to compile and make sense of the data o Areas of improvement are difficult to identify as the problem areas are not always easily found
Future implementation	o Integrate data from various sources and catalog into a centralized platform o Implement real-time data collection on flight tracking performance o Coordinate flight analyst outputs with operations for improvement reviews o Collect and analyze flight prediction data for continuous improvement	o Analyze integrated data to identify risk factors for carrier delays o Develop predictive models to calculate delay time seasonality, location, and frequency o Create an early warning system to flag risky flights o Expand communication efforts between operations and flight analysts	o Identify improvement domain; geography, flight times, operations o Early flags of high-risk flights o Draft operational improvements plans where identifiable o Timely intervention by analysts o Timely operational improvements	o Improved operational work-flows which reduce flight delays o Improved customer satisfaction o Reduced fees for regulatory infractions caused by delays o Improved reputation of the carriers and their operations

4V Model Outcomes:
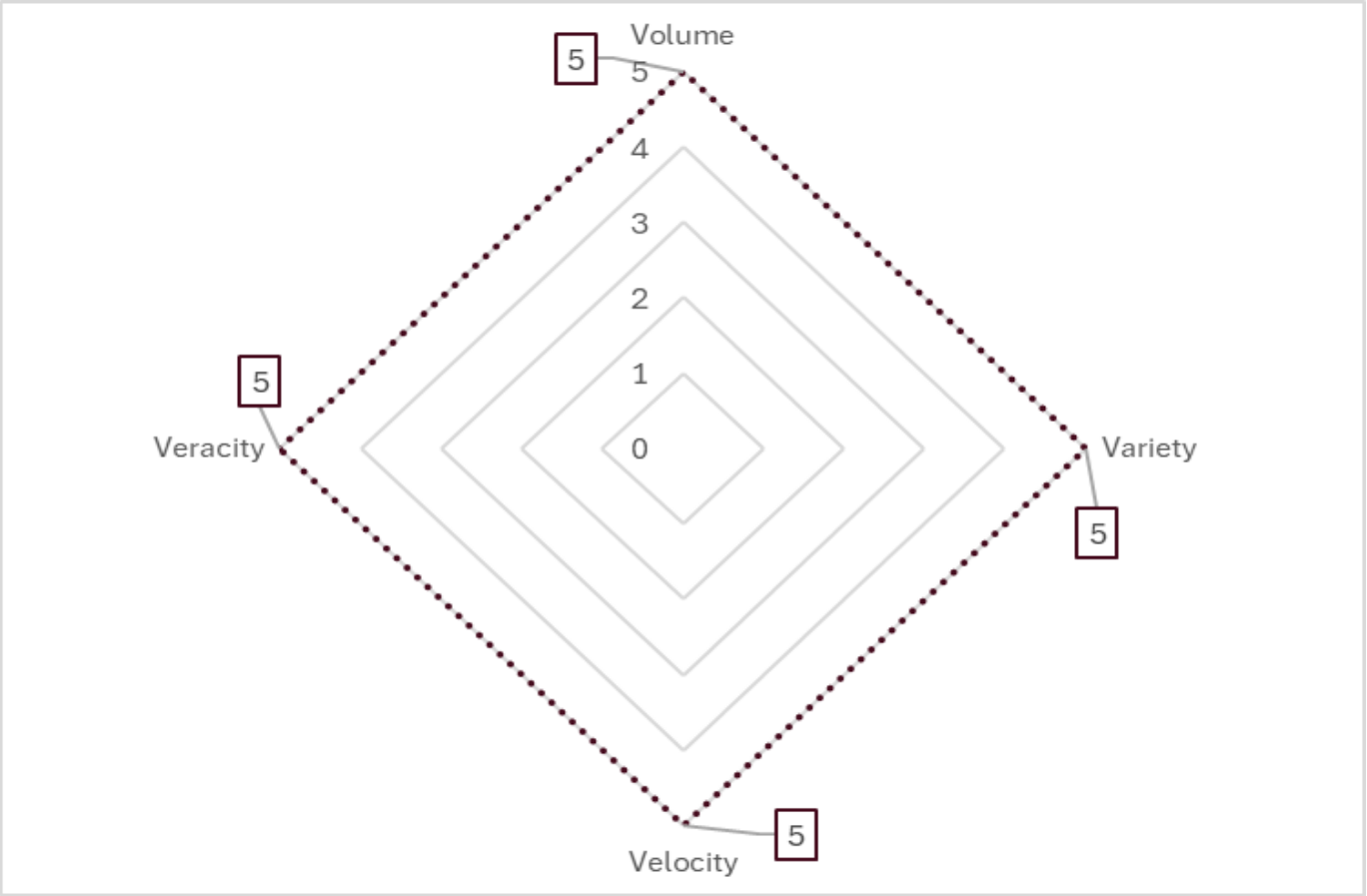
Consolidated and *Averaged Data Readiness Scores

Dimension	Considerations
Volume	How much data do you have available?
	How much data will be produced?
Variety	Does the data contain enough variety to capture even rare events?
	Does the data contain so much variety that it holds too much noise and requires heavy data cleaning?
Velocity	How often are relevant data sources produced or updated?
	Are data sources updated often enough to retrain the model soon enough to mitigate the risk of data drift?
Veracity	How accurate is the data, How complete is the data?
	How consistent is the data? Labels present?

CRISP-DM

Business Understanding

Improve flight delays by identifying at-risk flights by carrier, geography and time
Early warning systems & operational improvements

Data Understanding

Collect historical flight data
Explore and analyze data quality and relevance, identify key variables and metrics for the objectives

Data Preparation

Clean data and handle missing
Values; Integrate and merge datasets
Select and create relevant features, geographical data, time data, and carrier data

Modeling

Select modeling techniques
Train and evaluate models
Choose the best-performing model

Deploy the predictive model for use by flight analysts
Implement the model and early warning system
Coordinate with operational staff to implement improvement points to improve flight delays
Monitor and maintain the system

Evaluate the validity of the model
Assess accuracy and business relevance
Reiterate business understanding and alignment

| Business Understanding | Data Understanding |
| Data Preparation |
| Modeling |
| Deployment | Data | Evaluation |