RESEARCH DESIGN AND ANALYSIS IN DATA SCIENCE

UNIVERSITI TEKNOLOGI MALAYSIA

PROJECT PROPOSAL PRESENTATION

Forecasting AirAsia’s Profitability Based on Fuel Price Trends Using ARIMA and XGBoost

Video Link: https://www.youtube.com/watch?v=sfuBBFJsr8Y

UNIVERSITI TEKNOLOGI MALAYSIA

CANDIDATE : MIRA EDORA BINTI YUNOS
LECTURER : ASSOC. PROF. DR MOHD SHAHIZAN BIN OTHMAN

DATE : 30 JUNE 2025

PRESENTATION CONTENTS

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

Forecasting AirAsia’s Profitability Based on Fuel Price Trends Using ARIMA and XGBoost

1 RESEARCH INTRODUCTION
2 LITERATURE REVIEW
3 RESEARCH METHODOLOGY
4 INITIAL FINDINGS

RESEARCH INTRODUCTION

INTRODUCTION

PROBLEM BACKGROUND

Problem Background

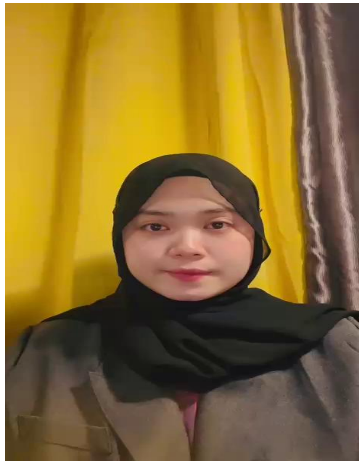
- The airline industry, especially low-cost carriers like AirAsia, is highly sensitive to fuel price volatility, with fuel accounting for up to 30%–50% of operating costs
- Despite digital transformation efforts, AirAsia remains vulnerable to external economic shocks, such as fuel market instability.
- Fuel price increases (e.g., a 10% rise) can lead to as much as a 15% drop in profits (Cai et al., 2025)

INTRODUCTION

PROBLEM STATEMENT

Problem Statement

- There is no integrated forecasting framework tailored to assessing the financial impact of fuel price volatility on low-cost carriers like AirAsia.
- Traditional models iike ARIMA handle linear trends but fail to capture nonlinear patterns caused by complex external variables.
- This study proposes a hybrid forecasting model that integrates ARIMA and XGBoost to overcomee these limitations and provide more robust, adaptive predictions



INTRODUCTION

THE NEED OF STUDY

KEY ISSUE

High Sensitivity to Fuel Price Volatility

Low-cost carriers like AirAsia are highly sensitive to fluctuations in fuel prices due to their significant contribution to operating costs up to 30% of total expenditures (Wu, 2024).

Limited Hedging Capacity

Airlines like AirAsia have limited options to protect themselves against rising fuel prices, leaving them vulnerable to sudden increases (Cai et al., 2025).

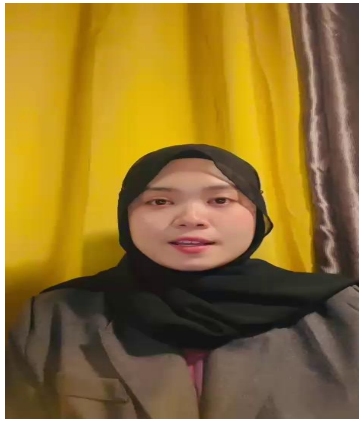
CONSEQUENCES

Financial Instability

Fluctuations in fuel prices lead to significant volatility in AirAsia's profitability, as seen in sharp increases and decreases in net profit/loss over time.

Operational Disruptions

Rising fuel costs compel airlines to optimize operational efficiency, potentially leading to service quality reductions or route adjustments. These measures can negatively affect passenger satisfaction and revenue generation (Sokkalingam et al., 2021).



INTRODUCTION

RESEARCH QUESTIONS

Research Questions

- How accurately can ARIMA and XGBoost forecast AirAsia's profitability based on historical fuel price trends?
- What are the key patterns and influencing factors behind fluctuations in fuel prices that affect airline profitability?
- How can a hybrid forecasting model improve financial planning and risk mitigation for low-cost carriers like AirAsia?

INTRODUCTION

RESEARCH OBJECTIVES

Research Objective

- Develop and apply an ARIMA model to forecast fuel price trends over time
- Utilize XGBoost to model the nonlinear relationships between fuel prices and AirAsia's profitability
- Integrate both models (ARIMA and XGBoost)) into a hybrid framework to improve the accuracy of financial forecasting

INTRODUCTION

RESEARCH MAPPING

	Problem Statements	Research Questions	Research Objectives
1	Lack of an integrated forecasting framework to assess the financial impact of fuel price volatility on AirAsia's profitability, especially under nonlinear and dynamic fuel price behavior.	How accurately can ARIMA and XGBoost forecast AirAsia's profitability based on historical fuel price trends?	To develop and apply an ARIMA model to forecast fuel price trends over time.
2	Inability of traditional models alone to capture complex external influences on aviation fuel prices and profitability.	What are the key patterns and influencing factors behind fluctuations in fuel prices that affect airline profitability?	To utilize XGBoost to model the nonlinear relationships between fuel prices and AirAsia's profitability.
3	Limited use of hybrid models in Malaysia's aviation financial forecasting context, reducing model adaptability and strategic planning accuracy.	How can a hybrid forecasting model improve financial planning and risk mitigation for low-cost carriers like AirAsia?	To integrate both models (ARIMA and XGBoost) into a hybrid framework to improve the accuracy of financial forecasting.

INTRODUCTION

RESEARCH SCOPE

RESEARCH SCOPE

The study specifically focuses on forecasting the financial performance of AirAsia

Historical data on fuel prices and AirAsia's financial performance will be utilized

The study applies:

- ARIMA for time series analysis of fuel price trends
- XGBoost for structured data modeling to capture nonlinear relationships between fuel prices and profitability

LITERATURE REVIEW

UTM UNIVERSITI TEKNOLOGI MALAYSIA

LITERATURE REVIEW

KEY COMPONENTS

LITERATURE REVIEW

Identification of Research Gaps
- Summary of existing research
- Identification of Research Gaps

Overview of AirAsia and the Volatile Airline Sector

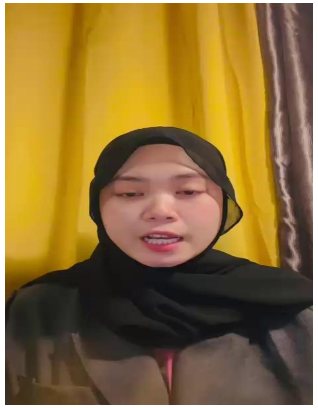
Global and Regional Airline Industry Landscape

Forecasting Models; ARIMA & XGBoost
- ARIMA in fuel price forecasting.
- Limitations of ARIMA for Nonlinear Events.
- Applications of XGBoost in Energy and Aviation Forecasting

Fuel Price Volatility and Its Impact on Airline Profitability
- Fuel Pricing Trends in Malaysia and Global Drivers.
- Fuel as a Primary Operational Cost.

Overview of Hybrid Models

Summary of Existing Research
- ARIMA in Fuel Price Forecasting.
- Limitations of ARIMA for Nonlinear Events.



LITERATURE REVIEW

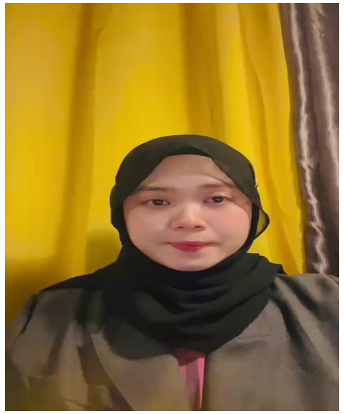
KEY THEMES

THEME	PREVIOUS STUDIES
Airline Profitability & Fuel Price Sensitivity	Airlines are highly sensitive to fuel price fluctuations, with fuel costs accounting for up to 30% of total operating expenses (Chen et al., 2024)
Time Series Forecasting Models (ARIMA)	ARIMA is widely used for linear trend analysis in fuel price forecasting (Baumann et al., 2021).
Machine Learning in Aviation Forecasting (XGBoost)	XGBoost is effective in modeling complex, nonlinear relationships in structured datasets (Yunos et al., 2024).
Hybrid Forecasting Models	Hybrid models combining ARIMA and machine learning algorithms are gaining traction (Baumann et al., 2021).
Data and Feature Engineering	Data preprocessing steps such as normalization, outlier detection, and feature engineering are critical for model performance (Yunos et al., 2024).
Model Evaluation and Interpretability	The study evaluates both ARIMA and XGBoost using MAE, RMSE, and R^2.

LITERATURE REVIEW

RESEARCH GAP

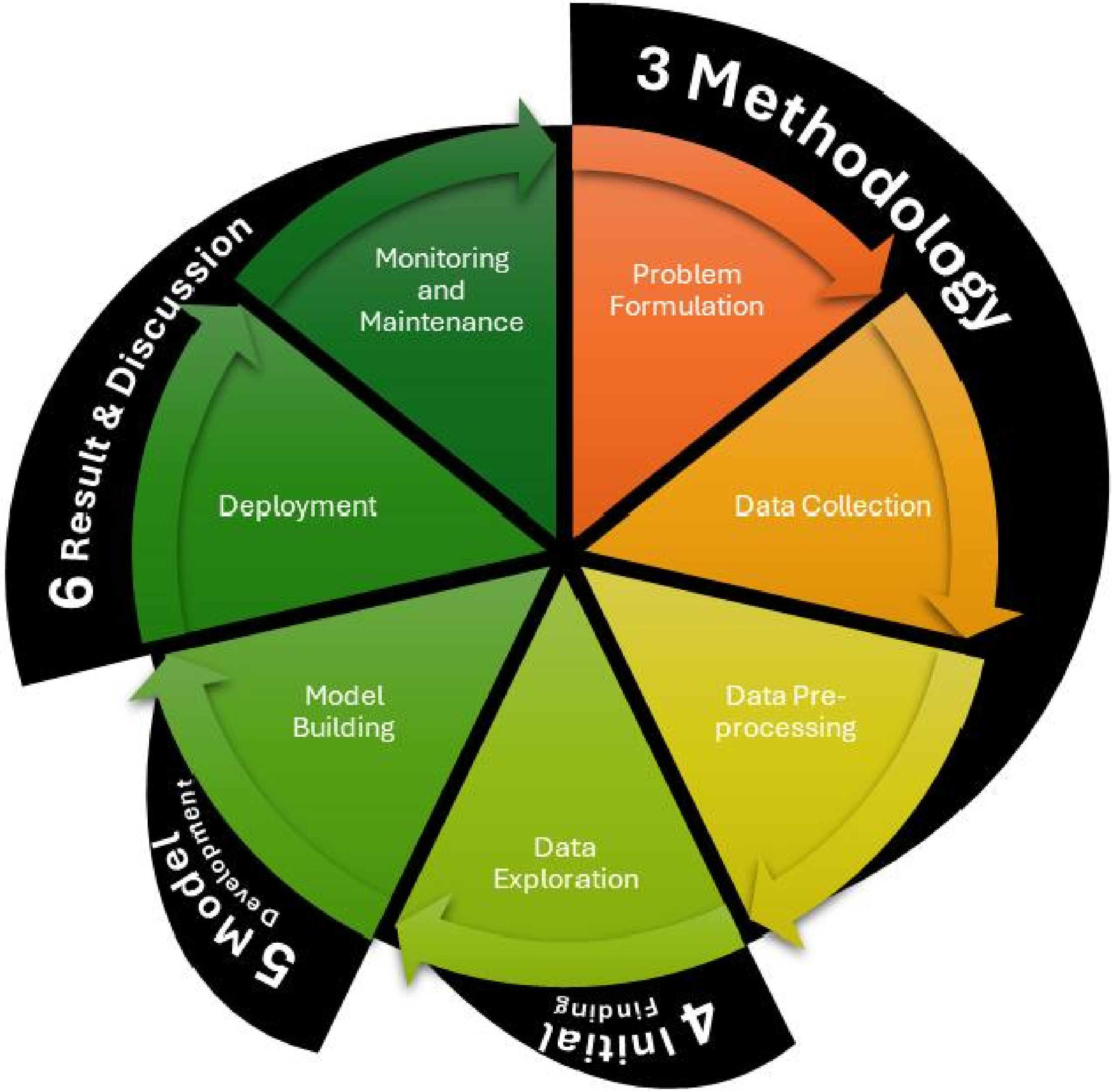
Aspect	Previous Studies	Current Research	Research Gap
Data Sources	• Use of historical fuel price data and financial records. • Limited focus on integrating multiple datasets	• Comprehensive dataset covering at least 3-5 years of data. • Inclusion of key variables such as fuel prices, revenue, and passenger load factor.	Need for integration of additional external factors to enhance predictive accuracy.
Methodology	• Primarily reliance on traditional statistical models like ARIMA. • Limited use of advanced machine learning techniques for forecasting profitability.	• Hybrid approach combining ARIMA for time series analysis and XGBoost for nonlinear modeling. • Implementation of feature engineering techniques. Such as lagged features, rolling averages.	Limited adoption of hybrid models in aviation financial forecasting.
Focus Area	Broad studies on airline profitability and fuel price volatility.	Specific focus on AirAsia, a leading low-cost carrier in Asia.	Lack of region-specific studies addressing the unique challenges faced by low-cost carriers in volatile markets.



RESEARCH METHODOLOGY

RESEARCH METHODOLOGY

RESEARCH FRAMEWORK



RESEARCH METHODOLOGY

RESEARCH FRAMEWORK

Phase 1: Problem Formulation
- Define research problem
- Investigate current data source
- Set Objective
- Data format and structure

Phase 2: Data Collection and preparation
- Data sources identification
- Data collection methods
- Data format and structure

Phase 3: Construction and Implementation
- Pre-processing and EDA
- Time series decomposition
- Feature seletion
- Feature engineering

Phase 4: Model Development and Training
- ARIMA model
- Random Forest model
- Model comparison

Phase 5: Validation &Conclusion

RESEARCH METHODOLOGY

PHASE 2 : DATA FORMAT

PARAMETER	VALUE
Time Frame	Minimum of 3-5 years data
Data Size	Dataset with up to 10,000 until 50,000
Variables	Include at least 5 to 7 key variables

RESEARCH METHODOLOGY

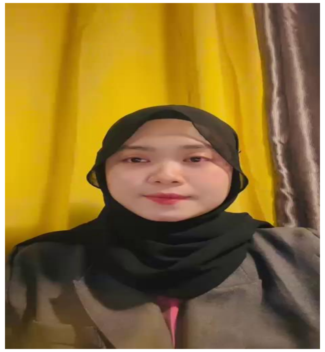
PHASE 2 : DATASET

Dataset	Source	Key Variable	Purpose
AirAsia Financial Data	AirAsia Investor Relations (Financial Reports from AirAsia's website)	• Date • Revenue • Net Income • Operating Cost • Fuel Cost • Passenger Load Factor	Forecast profitability using historical financial metrics
Fuel Price Data	Ministry of Energy and Natural Resources (KPKT)/ API(AviationStack)	• Jet Fuel Price • Crude Oil Price	Model fuel price impact on profitability
Passenger demand data	Malaysia Airports Authority (MAA)	• Number of passengers • Flight Frequency • Load Factor	Understad demand trends and their effect on revenue

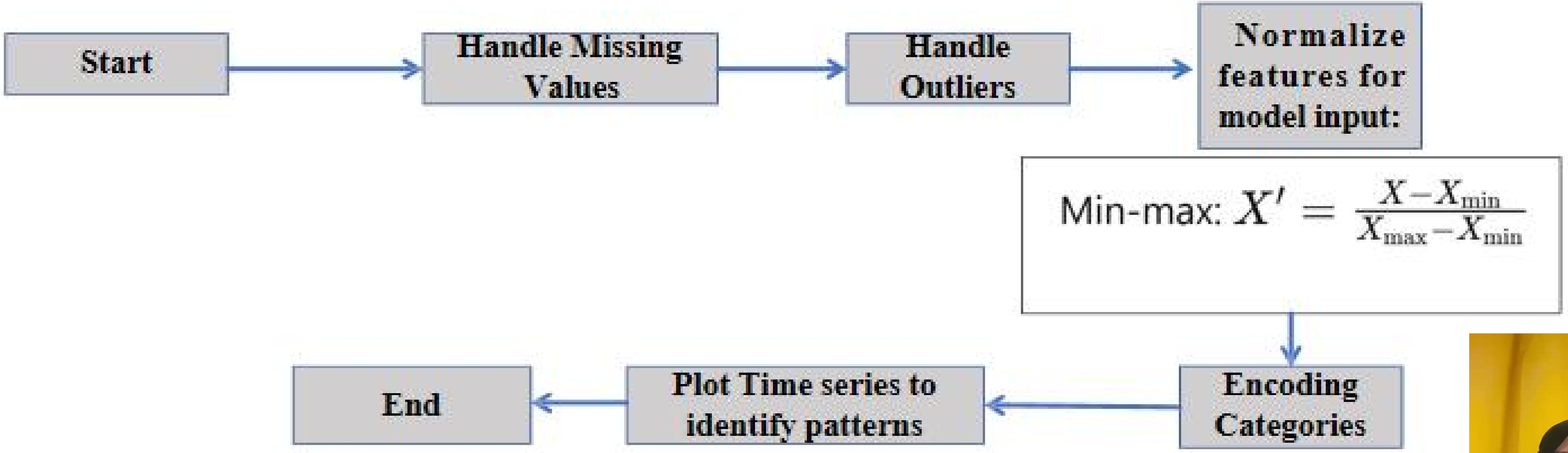
RESEARCH METHODOLOGY

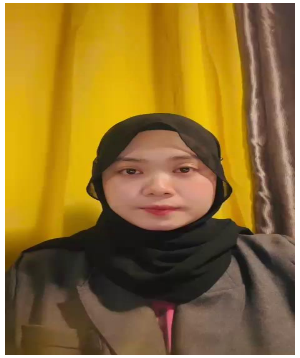
PHASE 3 : FEATURE ENGINEERING

Feature	Description	Formula used	Purpose
Lagged Feature	Previous values of fuel price and cost	Fuel_Price t−1, Fuel_Cost t−1	Forecast profitability using historical financial metrics (Comparing between predicted and actual profitability)
Moving Averages	Rolling averages of fuel price and cost	$\text{MA\_Fuel\_Price}_t = \frac{\text{Fuel\_Price}_t + \text{Fuel\_Price}_{t-1} + \text{Fuel\_Price}_{t-2}}{3}$	Model fuel price impact on profitability
Seasonal Features	Month and quarter of the year	$\text{Month}_t = \text{Month}(\text{Date}_t)$	Understad demand trends and their effect on revenue
Derived Metrics	Fuel cost and revenue per passenger	$\frac{\text{Fuel\_Cost}_t}{\text{Passengers}_t}, \frac{\text{Revenue}_t}{\text{Passengers}_t}$	Measure cost efficiency and revenue generation per passenger



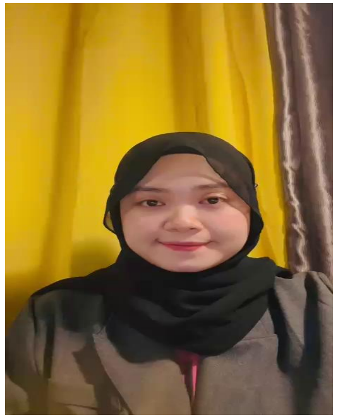
DATA CLEANING:
Flow of data cleaning





TIME SERIES DECOMPOSITION

STEP	DESCRIPTION
Formula	$$Y_t = T_t + S_t + R_t$$
Input time series Y_t	AirAsia's monthly profitability or fuel price over time.
Apply Decomposition	The goal is to split Y_t into three components using above formula: T_t Trend component (long-term movement) S_t Seasonal component (repeating patterns like monthly or quarterly) R_t Residual component (noise or irregularities).



RESEARCH METHODOLOGY

PHASE 4: Model Development

MODEL	DESCRIPTION
ARIMA	• A time series forecasting model for univariety data. • Captures trends, seasonality and autocorrelation in fuel price data without additional feature (Yunos et al., 2024).
XGBoost	• XGBoost is a powerful gradient-boosting algorithm that excels at handling complex, non-linear relationships in structured data (Yunos et al., 2024). • By effectively capturing nonlinear patterns and reducing overfitting through regularization techniques (Li & Zhang, 2018).
COMPARING MODEL	
Training and Validation	Split dataset into training and testing sets: $\frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$
Evaluation Metrics	Mean Absolute Error (MAE): MAE = Root Mea Squared Error (RMSE): RMSE= $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$

INITIAL FINDINGS

INITIAL FINDINGS

Exploratory Data Analysis (EDA)

Data Cleaning

- Shaping the dataset
- Detecting errors
- Correcting errors
- Processing values
- Transforming variables

INITIAL FINDINGS

PREVIEW DATASET

	count	mean	min	25%	50%	75%	max
Revenue (RM)	500000.0	799650142.630632	-648830802.0	596884669.25	799668436.0	1002156156.0	2168634417.0
Net_Profit_Loss (RM)	500000.0	-300813329.057666	-2138779273.0	-570849417.0	-300368642.0	-30925195.75	1571579640.0
Operating_Cost (RM)	500000.0	899735049.520496	-28883754.0	764684816.0	899727469.5	1034950700.0	1819165659.0
Fuel_Cost (RM)	500000.0	300083823.857914	-151346921.0	232731935.0	300164839.5	367368771.0	782762272.0
Fuel_Swap_Loss (RM)	500000.0	50037681.502318	-88757754.0	29770550.75	50057552.0	70287307.0	188316026.0
EBITDA (RM)	500000.0	199779974.427694	-1045036482.0	31508363.5	200015346.5	368567664.25	1458093548.0
Earnings_Per_Share (sen)	500000.0	-14.996068	-61.48	-21.73	-15.01	-8.26	35.34
Cash_Equivalents (RM)	500000.0	600124067.85264	-56762570.0	498748744.0	599752575.0	701130764.0	1338047291.0
Borrowings (RM)	500000.0	2000189541.3157	-277162843.0	1662263012.0	2001131209.5	2337517026.25	4180908742.0
Lease_Liabilities (RM)	500000.0	999968669.073888	-358544268.0	797254658.25	1000007258.5	1202278162.5	2357004771.0
Passengers	500000.0	5256116.942786	500021.0	2878045.5	5265464.5	7636628.25	9999997.0
Seat_Load_Factor (%)	500000.0	74.996177	60.0	67.52	74.98	82.48	90.0
ASK (mil)	500000.0	999.725053	-415.24	797.32	999.875	1201.86	2317.83
Fuel_Price_USD_per_Barrel	500000.0	90.04084	60.0	75.05	90.06	105.0525	120.0
Quarter_Date	500000	2022-11-15 16:29:59.999997696	2021-01-01 00:00:00	2021-12-09 00:00:00	2022-11-16 00:00:00	2023-10-24 00:00:00	2024-10-01 00:00:00
Year	500000.0	2022.5	2021.0	2021.75	2022.5	2023.25	2024.0

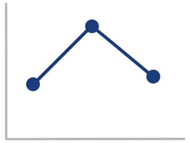
INITIAL FINDINGS

VISUALIZATIONS

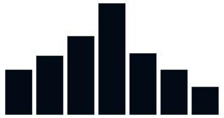
Time Series Plots: A plot where the x-axis represents time (in this case, quarters), and the y-axis represents a continuous variable

Correlation Heatmaps: Showing the Feature relationships & multicollinearity

Line Plots: Illustrating the trend analysis of quarterly performance



Derived Features: For Lag and Rolling Variables. Temporal modeling readiness



Bar Plot: Showing the Fuel Cost per Passenger, cost effieciency across quarters. Profit Margin per Passenger.

INITIAL FINDINGS

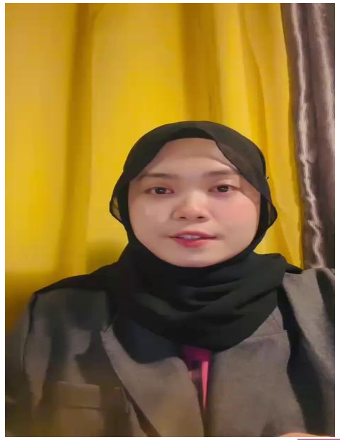
INITIAL INSIGHTS

High Volatility in Profitability AirAsia experiences significant fluctuations in net profit/loss over time, with notable periods of severe losses.

Inverse Relationship Between Fuel Prices and Profitability Higher fuel prices are strongly correlated with lower net profits or higher losses.

Negative Profit Margins Per Passenger Consistintently negative profit margins indicate ongoing operational challenges.

Strong Correlations Among Fuel Variables Fuel price, fuel cost, and rolling fuel price show strong positive correlations, highlighting the importance of fuel management.



INITIAL FINDINGS

EXPECTED OUTCOME

EXPECTED OUTCOME

- Improved Forecasting Accuracy
- Identification of Key Drivers of Profitability
- Contribution to Aviation Financial Modeling
- Enhanced Decision-Making Support

UTM UNIVERSITI TEKNOLOGI MALAYSIA

INITIAL FINDINGS

FUTURE WORK

FUTURE WORK

- Incorporate Forecasting with External Shocks
- Improve Model Interpretability
- Apply the Model to Other Regional Carriers
- Expand Data Sources

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

univteknologimalaysia

utm_my

utmofficial

Thank You

www.utm.my

innovative • entrepreneurial • global