

Flight Prices Analysis

Data Management

Project in Data Management with a specialisation in Data collection, Data cleaning, Data visualisation and Modelling.

Student 1: Adelia Rahmawati Safitri

Student 2: Céline Lam

Student code 1: s175873

Student code 2: s191578

Course: ECON2306-1

Lecturer: Malka Guillot

June 2023

Abstract

The goal of this project is to develop a model that can accurately predict the prices of round-trip airline tickets from Lufthansa and Swiss airlines to São Paulo and New York City. By determining the optimal time to purchase tickets for each destination, it is hoped that significant savings can be achieved, allowing for the maximisation of the travel budget. To accomplish this, historical data on ticket prices and other relevant factors such as holidays will be analysed in order to identify patterns and trends that can be used to forecast future ticket prices. This project is expected to provide valuable experience in working with time series data and developing predictive models, while also uncovering practical ways to save money on air travel during peak travel times.

Keywords: Destinations, Historical data, Ticket prices, Patterns, Trends, Forecasting, Time series data, Predictive models

Contents

1	Introduction	1
1.1	Research questions/hypotheses	1
2	Background	1
3	Data	1
3.1	Data preparation	2
4	Approach	2
5	Results	2
6	Discussion	2
6.1	Limitations and Challenges	2
7	Conclusion	2
8	Reflections on own work	3

1. Introduction

As students, we are constantly seeking opportunities to maximise our limited budgets. One area where we identified potential for cost savings is air travel, particularly during peak travel times such as holidays. During these periods, ticket prices can increase significantly, making it challenging for us to afford travel for visiting family or taking vacations.

With this in mind, we decided to undertake a project to develop a model capable of predicting the prices of round-trip tickets from Lufthansa and Swiss airlines to São Paulo and New York City. By analysing historical data on ticket prices and other relevant factors such as holidays, we aim to identify patterns and trends that can assist us in forecasting future ticket prices. Utilising this information, our goal is to determine the optimal time to purchase a ticket for each destination, allowing us to save money and make the most of our travel budget.

Through this project, we hope to gain valuable experience in working with time series data and developing predictive models while also discovering practical ways to save money on air travel during peak travel times. By focusing on two specific airlines and destinations, we aim to develop a model that is both accurate and applicable to our travel needs.

1.1 Research questions/hypotheses

Based on the problem domain described above, list and describe your research questions/hypotheses to be answered in your report. These questions can also be used during your user evaluation, i.e., will the users be able to answer these questions with the help of your tool?

2. Background

How do you plan to address the problem and answer your hypotheses? Review and describe the related literature. It is encouraged to make use of the seminar papers already provided, if applicable. What are the previous researches' questions, results, and solutions and how do they relate to your problem? For example, if your project is about analyzing predictions of the effects of global warming, how have others addressed this problem? Which analysis techniques have they used for their particular data and why? That is, can the related research found help you in your work on your data, analysis, visualization, and evaluation?

3. Data

Briefly describe the dataset and the parameters that will be important for you in order to answer your questions/hypotheses. Indicate the source. If available online, providing a URL link helps the readers.

3.1 Data preparation

Describe what you did in order to use the data for analysis.

4. Approach

Describe the techniques that you will use for managing and analyzing your data and the tool(s) that you will use for data management/data analysis and evaluation. Motivate why you have chosen a certain method over others and describe how you plan to evaluate your solution. This section should give a detailed account of what you did in order to answer your questions.

5. Results

Describe your final data analysis results and everything that you discovered on the way the put you forward and you think is worth describing here. Elaborate on the significance of your findings/results and convey your thoughts.

6. Discussion

What do the results mean (their significance and to whom) and how do they answer your research questions/hypotheses? From the evaluation and validation of the tool, what can you conclude?

6.1 Limitations and Challenges

What could have been investigated if given more time? What have been difficult when solving the problem and getting answers for your research questions/hypotheses?

7. Conclusion

Write the conclusion. What did you gain from the project assignment? Briefly explain your questions/hypotheses, findings, and meaningful discussion points in relation to the data collection, data management, data analysis, visualization and interaction concepts, and evaluation of your tool. What additional investigations need to be performed (or what is the limitation) in order to say that your solution is a good one for the problem?

8. Reflections on own work

- Describe how you decided to scope (and/or re-scope) your problem formulation during the work, and given the data you had access to.
- Describe how you searched for knowledge on how to scope the DS question you wanted to answer
- understand how to implement, test, and validate your results.
- Which sources helped you to get progress and how?
- What would you have done differently if you were to start over again, for understanding the problem faster and for understanding what could be done with the data you had access to?
- What else is there that you would have changed about this assignment?