An Investigation on Mitigating Airline Pricing Model's Bias

DSC 180A,B Capstone Group A-05-2

Edwin Tse, Garrick Su, Joseph Perez, Maricela Vasquez, Qixi Huang
University of California, San Diego

Abstract

When buying an airplane ticket there are multiple factors that can result in largely different prices in flights. As a result, there exist multiple different avenues for price discrimination to occur. Here, we investigate the effects of local demographics on the pricing of a given flight. Specifically, we look into potential biases with regard to race and income in terms of the metropolitan area. This is done through bias analysis and mitigation via the AIF360 toolkit. In turn, we aim to develop models that balance both accuracy as well as fairness between our classes to provide insights for a better pricing strategy for airlines and pricing information to consumers. In terms of pricing bias based upon metropolitan demographics, we found that there was little bias present in the data and in the models we developed. While the bias was not significant, we found a correlation between our set demographics and pricing as well as created models that balance both fairness and accuracy which demonstrates practical applicability of bias mitigation techniques.

Hypothesis

We anticipated that there would be a significant difference between flight fares when comparing our privileged and unprivileged classes. Here, our classes will be determined by race (white and non-white), and income (high and low) of the local airport metro area population. Low income here is defined to be whether the median of the local income is less than or equal to the 25th quantile threshold found from a distribution of median incomes across all metro areas.

Introduction

Price discrimination is the practice of setting significantly different prices for different groups of people for the same or similar commodity. This is specifically known as third-degree price discrimination. As a result, this can often lead to optimized profits for the seller, with a subset of buyers left paying higher prices. In the airline industry, there are several features such as distance and airport that can affect the fare price as well as external forces such as market concentration and competitors. In turn, the local demographics of airport metro areas are features we hypothesize may distinguish significant pricing discrimination between majority and minority populations in airfare prices when comparing similarly comparable flights. The goal here is to identify potential factors in the dataset that may be causing biases in model implementations, mitigating these biases, and ultimately developing fairer machine learning models for fairer airfare pricing.

Related Literature

Price discrimination is an issue that is being observed by many through various similar projects. The first project we will be evaluating is a project that focuses on investigation on price discrimination within the airline industry by comparing prices to similar users. The project was conducted by Stefano Azzolina from the Department of Economics at the University of Bologna, Manuel Razza from the Italian Competition Authority, Kevin Sartiano from the Department of Engineering at Uninettuno University, and Emanuel Weitschek from the Department of Engineering at Uninettuno University and the Italian Competition Authority. Data for study was gathered through a software that obtained user information from information provided as they purchase airline tickets online. The Flight Data Acquisition Software collected data under two workflows which are flight search similarities and flight search differences. The results of this study revealed that price discrimination was prevalent as we see differences in prices amongst different types of users. The study goals are similar to the investigation we analyzed through this project. However, instead of using data from users, we collected data from survey data and utilized the AI Fairness 360 toolkit. Another difference is that our project observes comparison among flight paths, which entails ticket purchasers within the same origin and destination locations.

The next study was conducted by Kevin R. Williams used airline dynamics to create a model to predict pricing. Airlines dynamics are factors such as demand and scarcity that make airlines adjust their prices daily based on potential customers' willingness to purchase tickets for airlines set prices. The reality is that airline industries use these factors to make a profit. The proposed model utilizes stochastic demand features and revenue management models to predict airline ticket pricing. The model reveals the competitiveness that airlines pricing can lead to price discrimination. Employing similar models and analysis on airline data can help prevent price discrimination that are created out of the competitiveness of purchasing airline tickets. Compared to our projects, both create models to predict ticket pricing. We believe that the model from Kevin R. Williams focuses on the domain knowledge of economics that reveals the reasoning behind pricing. Our model focuses on the fairness of the model, in fact attempting to mitigate the bias within our model.

Datasets

The airline ticket and pricing data are provided by the Airline Origin and Destination Survey (DB1B). The DB1B database is a dataset that is maintained by the United States Department of Transportation Bureau of Transportation Statistics. Origin and Destination Survey (DB1B) is a 10% sample of airline tickets from reporting carriers in the United States. Data includes origin, destination and other itinerary details of passengers transported. The DB1B database has data from 1993 to the 2nd Quarter of 2022, however, due to the constraints of our computing environment capabilities, we are only using the data from 2018 to the most recent available record in our project.

While the DB1B database does not include demographics such as race, age, or income, for our purposes, we are instead using U.S. Census data in order to get feature variables that describe the local populations of the origin airport metropolitan area and the destination airport metropolitan area.

By merging these two datasets, we are able to investigate the relationship between local population demographics and airline ticket prices. Before beginning our analysis, it is important to establish that our target variable, airline ticket fare per mile (FarePerMile), will be categorized into two classes under a new variable called fare_class. For the preferred outcome or group, we will have fares that are below the third quartile or 75th percentile of the fare-per-mile distribution. For the unpreferred outcome or group, we will have fares that are above the third quartile which essentially represents a high-cost fare. This then turns our predictive task into a classification problem instead of a regression which is necessary as we are concerned with a range of values being fair or belonging to the same group instead of how precisely the exact cost of two tickets is.

The database also does not include the general details (such as total passenger count, amount of flights on a given route) from a given set of flight details. To further enhance our ability to capture a well-behaved model for our analysis, we also incorporated data from the Air Carrier Statistics (Form 41 Traffic)- U.S. Carriers (T-100) from the United States Department of Transportation Bureau of Transportation Statistics. This dataset contains domestic market data reported by U.S. air carriers, including carrier, origin, destination, and service class for enplaned passengers, freight and mail when both origin and destination airports are located within the boundaries of the United States and its territories.

EDA

We conducted our EDA using the database listed above and on other associated datasets. In this section, we would present our methodology during our analysis and how the detailed findings help our future progress in project findings.

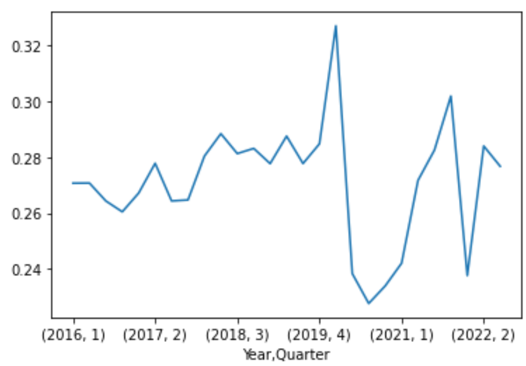
Processing the DB1B Dataset

The original Ticket in the DB1B dataset records information on the itinerary level. Since one itinerary might have multiple flights and destinations depending on whether it is a round trip or has multiple stops along the way. Therefore, we merge the Ticket dataset with the Market/Coupon dataset on itinerary ID, and it allows us to look closer into ticket information on an individual flight level. On average for each quarter the combined dataset has around 6 million rows and 26 useful features after we exclude other redundant columns, where each row represents a ticket and its associated information. Details of the columns and variables are available on the project website.

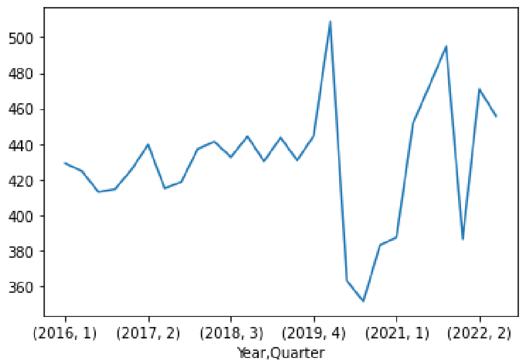
Since the DB1B dataset is build on the ticket level, meaning that the ticket could have segments that represents either be a one-way, a round-trip or a muilt-destination ticket. Therefore, there is not a clearly define destination. We decided to use observe the destination in dataset in based on various assumptions shown below, which are also implemented in our "`sparkmanager.py`":

Assumption	Method name in module	Resulting dataset size
The last destination is the real destination	"default"	98898392 (total amount of tickets)
The median point is the real destination	"midpoint"	98898392 (total amount of tickets)
Each segment should be separate	"segment"	247175573 (all avaiable rows in the dataset)

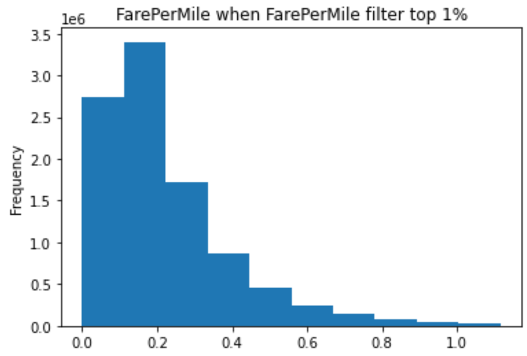
After our data wrangling work, we immediately found the default approach is problematic because over 60% of the tickets are considered roundtrips. As a result, over 60% of tickets using such an approach results in the same origin and destination. The median approach shows some promising results, however it also means it is losing a lot of information that could be able to be represented in the segments. Therefore, after our experiment on the three approaches, we decided to treat each segment as an individual ticket in our future analysis as it would keep the most information intact.



Fare Per Mile in 2022 dollars



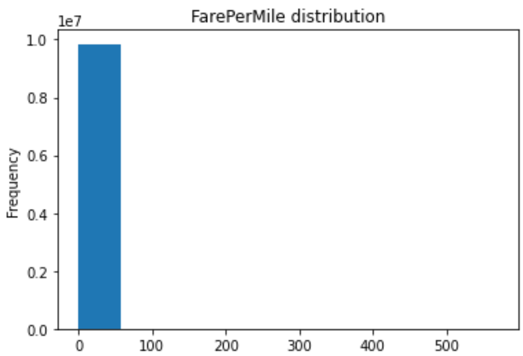
Ticket Prices in 2022 dollars



Distribution of FarePerMile in DB1B after to outlier filtering

To further perform a better and more accurate analysis, we also used the cpi module in Python to change all of our prices into 2022 prices from their respective years. The prices shown below the EDA section are the original number of prices, unless specifically stated of using the cpi adjusted 2022 prices.

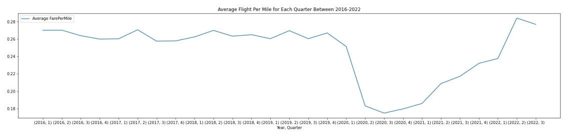
Trends and Findings of the DB1B Dataset

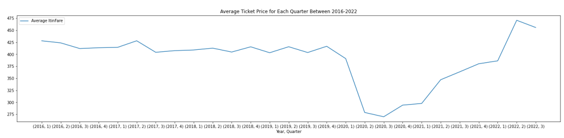


Distribution of FarePerMile in DB1B prior to outlier filtering

Our focus on pricing prompted us to focus most of our effort in it when it comes to finding trends in the dataset. Therefore, it becomes essential to analyze the FarePerMile (fare per mile) and ItinFare (ticket itinerary fare). We immediately look into the general distribution of the two variables and find a small but extremely strong set of outliers.

After a thorough investigation, we discovered that the bottom 1%-tile and the upper 99%-tile had huge outliers that swayed the variable. Therefore, we decided to drop those rows in our following analysis due to the belief that such may be a result of human error.

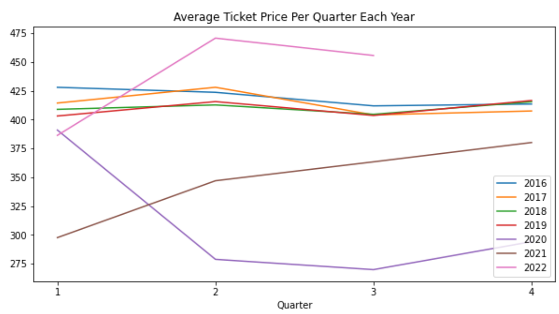


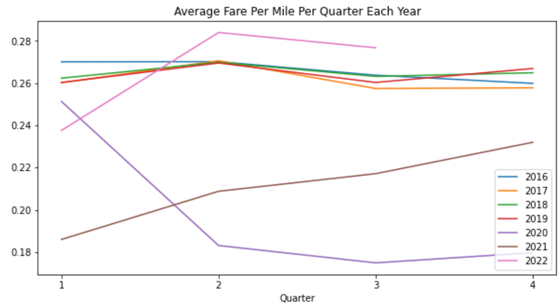


By plotting them against their respected year and quarters, we found that they in general follow each other quite well.

In the case of FarePerMile, the variable is relatively stable, prior to the 2020 Q1, hovering between 26-27 cents in each quarter. There is a significant drop during the first 3 quarters of 2020 from the average of 25 cents to the lowest at 18 cents. The variable since then steadily rebounded to 24 cents at the last quarter of 2021, and spiked to the highest point in 2022 Q2 and Q3 of 28 cents.

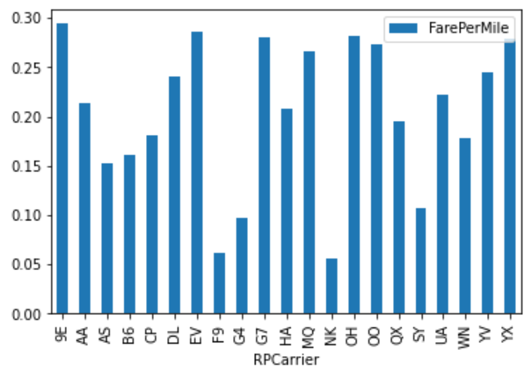
In the case of ItinFare, the variable is relatively stable, prior to the 2020 Q1, hovering between \$400-\$430 in each quarter. There is a significant drop during the first 3 quarters of 2020 from the average of \$390 to the lowest at \$270. The variable since then steadily rebounded to \$390 at the last quarter of 2021, and spiked to the highest point in 2022 Q2 and Q3 of \$470. It is important to note that this variable is highly influenced by the rate of round-trip travel, as a round-trip ticket fare is counted as the same as a one-way trip ticket in this variable. Our analysis found (also shown below), on average 60% of the tickets are classified as round-trip travels.





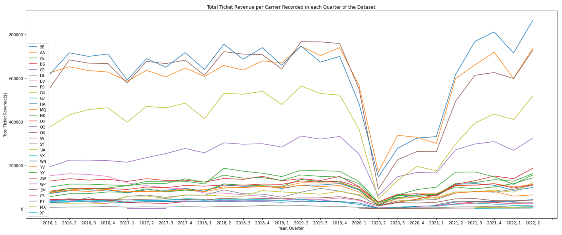
We then focused on finding whether there is a noticeable difference between the quarters of the ticket used on travel.

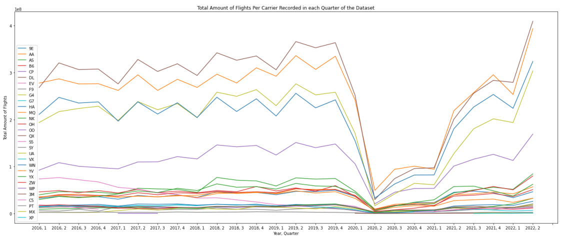
Our analysis shows that in the years prior to the pandemic (2016-2019), both FarePerMile and ItinFare follows a general trend with both Q2 and Q4 peaking in each respective years. However, the same analysis provided insight on the significance of the pandemic effects during the later years of the dataset.



Average FarePerMile of each reporting carrier in DB1B

Airline-wise, low-cost airlines like Spirit have an average fare-per-mile of 0.1\$, while Regional Airline like Silver Airways has the highest average fare-per-mile of 0.6\$. On the other hand, Legacy carriers like United, Delta, and American Airlines have relatively moderate and similar average fare-per-mile (~ 0.3$), which we think can be a scenario of market competition. Nevertheless, Legacy carriers have a dominant market share, over 65% of the tickets in this combined dataset belong to these airlines.

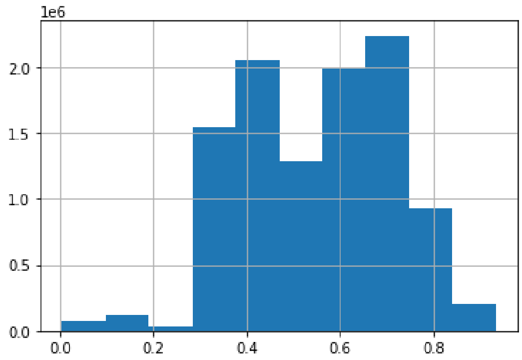




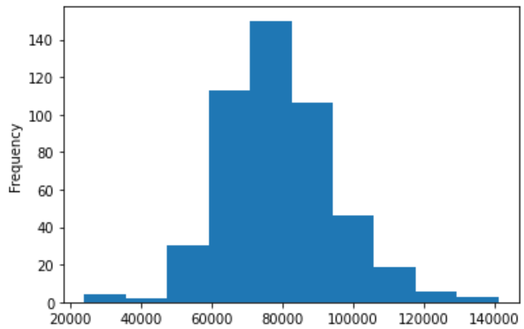
By visualizing the dollar amount presented as revenue, we allow us to investigate whether there are airlines that have a sharp competitive edge, where they earn more money yet by flying fewer passengers. Our analysis shows that this is likely not the case.

To further investigate the significant impact during the pandemic, we decided to look into the flight volumes associated with the dataset more. We found it matches our findings and expectations.

Processing the US Census and Defining the Protected Groups



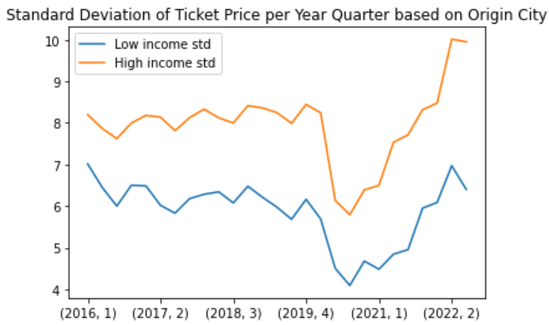
Distribution of White Population Ratio in the Dataset

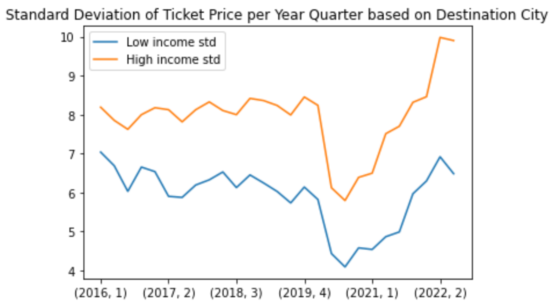


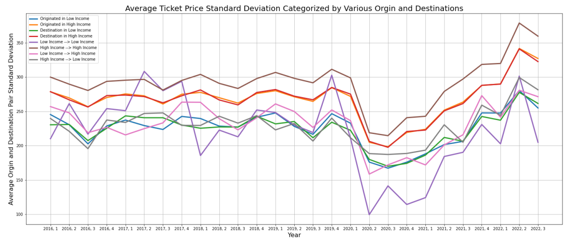
Distribution of Household Median Income in the Dataset

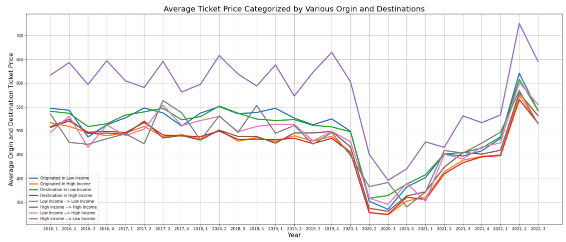
The racial and income groups for each airport's metropolitan area information taken from the US Census allows us to gather information of our protected groups. After conducting data cleaning with the Census data, we produce two separate datasets, Race and Income. Both dataset has 754 rows and 8 columns, where each row represents a Micro/Metro Area that is related to an airport code that is listed in the DB1B dataset and has corresponding demographic details (e.g. ratio of White and Non-White population).

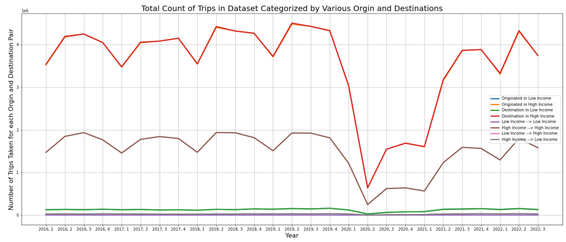
The Income dataset has 513 rows and 5 columns, where each row represents a Micro/Metro Area and has corresponding income statistics. Based on the shared Area code, we will merge these two datasets with the combined ticket dataset above. Rows that don't have matching area codes are omitted during the merging process. In other words, it is an inner merge as we want to ensure each ticket purchased comes with detailed information about its origin/destination city.











Bias Discoveries on the DB1B Dataset

To ensure our bias discovery process would be as fair as possible, we used the 2022 dollars for this part of the analysis. We first started with uncovering the difference of means in fare per mile between different protected groups and privileged groups based on their destinations.

Based on origination of flight:

- Privileged Group:
  - High income
  - Dominant Race
- Protected Group
  - Low income
  - Minority Race

Protected Attribute	Privileged group mean	Protected group mean
Income	0.2373	0.3177

Ethnicity	0.2978	0.2859

The privileged group for the attribute of income pays the least of 0.2373 while the protected group of income pays the most at 0.3177 fare per mile. For ethnicity, the privileged group pays more than the protected group, and there is a much closer margin between them which implies that pricing is more balanced than income.

Based on destination of flight:

- Privileged Group:
    - High income
    - Dominant Race
- Protected Group
    - Low income
    - Minority Race

Protected Attribute	Privileged group mean	Protected group mean
Income	0.2385	0.3242
Ethnicity	0.3106	0.2894

Similar to above, when grouped on destination of flight, the privileged group of income pays the least at 0.2385 while the protected group pays the most at 0.3242. With ethnicity, the privileged group pays more than the protected group with a smaller margin.

Flight with different groups in origin and destination

- Privileged Group:
    - High income
    - Dominant Race
- Protected Group
    - Low income
    - Minority Race

Groups	Fare Per mile mean
Income Privileged → Protected	0.2734
Income Protected → Privileged	0.2648
Ethnicity Protected → Privileged	0.2648
Ethnicity Privileged → Protected	0.2502

When looking at FarePerMile across income and ethnicity at the same time, we can see that going from an income-privileged area to a protected area has the highest fare per mile, but the margins between each are generally small.

Flight with same groups in origin and destination:

- Privileged Group:
    - High income
    - Dominant Race
- Protected Group
    - Low income
    - Minority Race

Protected Attribute	Privileged group mean	Protected group mean
Income	0.2077	0.4519
Ethnicity	0.2416	0.3067

When we make both the origin and destination city the same type (privileged to privileged), we see that margins increase even further. With the protected income group having a much higher cost of 0.4519 than the privileged group income class of 0.2077, and the protected ethnicity class also pays more at 0.3067 compared to 0.2416.

Since the small amount of representation in the dataset of the protected groups shown above, compared to the privilege groups. Although we find some great differences between groups, we did not find such a difference statistically significant enough. Therefore, we may not conclude that the dataset itself is inherently biased.

Another discovery that we take to further investigate the relative larger differences between the high income and low income group is to see whether there is a strong difference in their standard deviation, which if it is lower shows a low price variability and vice-versa.

We found out that the low-income groups consistently has a lower standard-deviation compared to the high income group. However, this may also be the result of there is far less samples of low income related flights included.

Further Discoveries and Engineered Variables

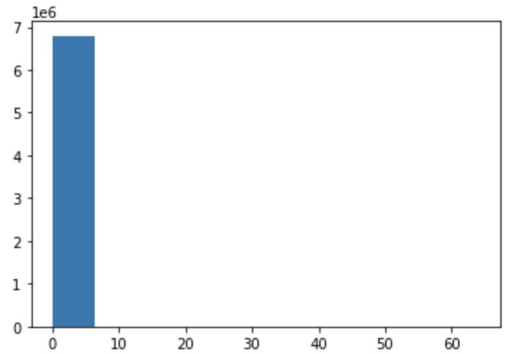
After recognizing the general trend of the pricing variables and the relationship with other variables in the dataset, we looked into other deeper trends to determine the predictors and predicted variables.

<u>Fare Class Model</u>

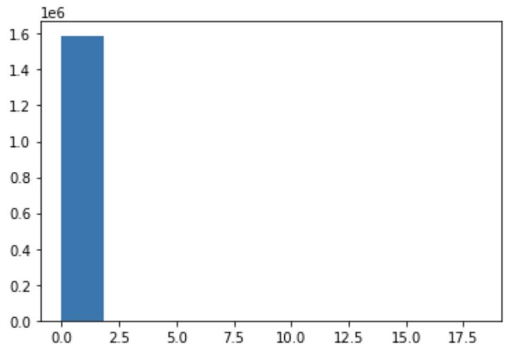
Distribution of FarePerMile Based on FareClass

For the Fare Class Model, further analysis of the relationship between FareClass and FarePerMile features needs to be evaluated to help classify the various fare classes as favorable and unfavorable fare class for bias analysis which will be described in more detail later on in the report. The following histograms show the frequency of FarePerMile for the different Fare Classes.

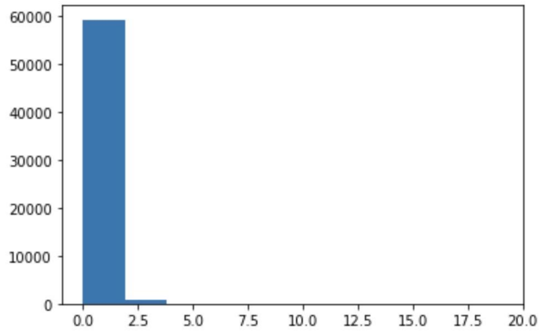
FarePerMile Distribution for Class X



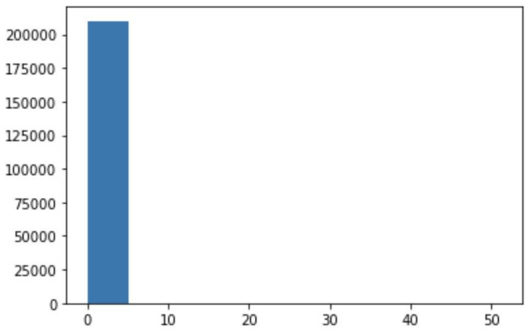
FarePerMile Distribution for Class Y



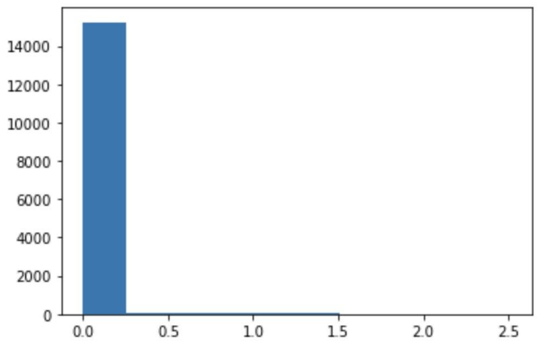
FarePerMile Distribution for Class C



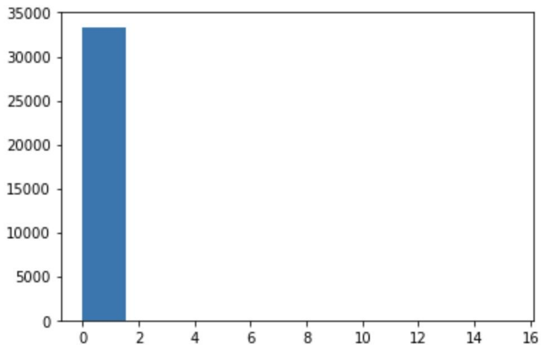
FarePerMile Distribution for Class D



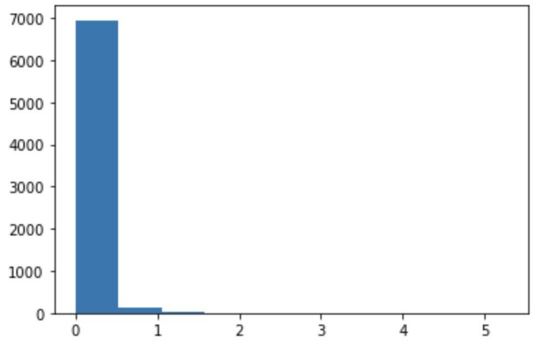
FarePerMile Distribution for Class F



FarePerMile Distribution for Class G



FarePerMile Distribution for Class U

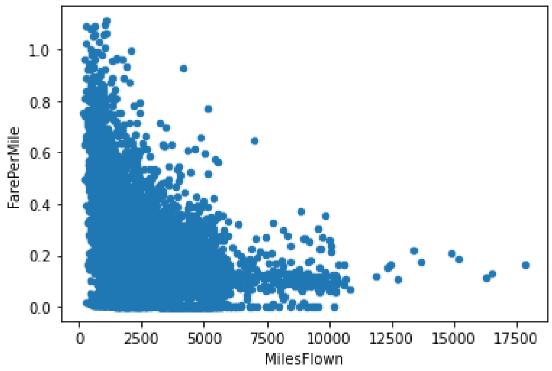


As noted in previous analysis, outliers had to be removed, specifically for FareClass X analysis. Based on the distributions, we can categorize classes Y, C, and D as favorable classes, with their distributions being close to the mean or even lower to the mean of FarePerMile.

Transforming Distance to Categorical Variable

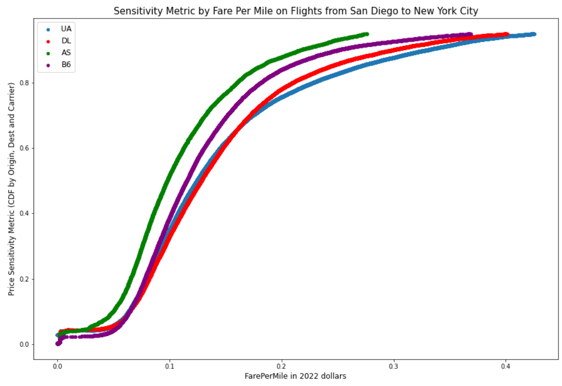
To further discover how flights are different based on the distance traveled, we create a categorical variable, “Flight Length,” and segment our ticket into three categories: “Short-haul” flights with Miles flown less than 1725 Miles, “Medium-haul” flights with Miles flown between 1725 and 3450 Miles, and “Long-haul” flights with Miles flown bigger than 3450 Miles. We group by the variable “Flight Length” and discover that “Short-haul” and “Medium-haul” take up around 80% of all the recorded flights (around 40% each).

Although our EDA focuses on the first quarter of 2022, this ratio varies across years and quarters. The average fare per mile is the highest for “Short-haul” flights and the lowest for “Long-haul” flights. In other words, passengers on short-haul flights tend to pay more for each mile they are flying. If most of the flights that come out of an airport are short-haul, then residents that live around this airport might have to bear a higher fare per mile when purchasing flight tickets.



The correlation values also confirm the above finding; we saw a negative correlation value (-0.16) between the target variable “fare-per-mile” and the “MilesFlown.” In addition, we also saw other features like “Coupons”, “Passengers”, “RoundTrip”, “median_income”, “None-white alone proportion_dest”, and “white alone proportion_dest” having a moderate correlation (absolute value of correlation > 0.01, choosing this threshold given that there are so many features in the dataset) with “fare-per-mile.” The correlation values suggest there are connections between flight fare and demographic information, implying potential biases with regard to race and income in price-setting models.

<u>High Price Indicator Model</u>



As shown in below section, the model uses the CDF of the price variables as the sensitivity metric, based on the origin, destination and the reporting carrier.

The sensitivity metric is calculated by taking all the ticket prices and the related fare per mile in 2022 dollars from the dataset (2016-2022) and ranking them in ascending order between the interval of 0 to 1 in groups of origin, destination, and the airline. On the training data, 0 represents the lowest price available in the dataset on a given city-pair and airline, and 1 represents the highest price available in the dataset on a given city-pair and airline. Essentially, the values are the CDF of the distribution. All price of the data is adjusted by the CPI index (Airfare category) to 2022 dollars. The model was trained based on data from 2016-2022 DB1B dataset.

We interpret if the model predicts 1, means that the passenger would be least likely to make a purchase of a fare and the airline would most likely offer such a fare. If the model predict 0, means the passenger would be most likely to make a purchase of a fare and the airline would least likely offer such a fare. If the model predicts anything that is > 1 or < 0, it means that the price is never going to happen.

Model Development

High Price Indicator Model

For our models, the data is grouped by flight path in order to determine if our privileged class would be predicted to pay lower FarePerMiles in contrast to our unprivileged groups. Here, our privileged group would be white majority populations when summing up flight origin and flight destination metropolitan areas. Our unprivileged group is non-white majority when summing up flight origin and flight destination areas. This variable that we created is defined as our "Race" variable and it contains the group information we will be using as previously described. The variables used in our models for predicting high or low FarePerMile are "RoundTrip", "OnLine", "DistanceGroup", "OriginCityMarketID", "LastCityMarketID", "RPCarrier", and "ItinGeoType." Our group has created various models that aim to predict the FarePerMile model while mitigating bias. We will be listing the models that are the most accurate or insightful in understanding and mitigating the existing bias to demonstrate the trade-off between debiasing and accuracy.

Our baseline models are simple logistic regression and random forest regression models in order to gauge predictive bias and performance. Our target variable is created by applying a threshold to FarePerMile where any flight path that has a median income greater than the 75th percentile in terms of median income In testing, we are using 2019 data in order to avoid COVID-related shifts in typical air trends as we earlier found in exploratory data analysis that the trend of average FarePerMile drastically changed during 2020, 2021, and even still in the currently available 2022 data.

In terms of performance, using the aforementioned variables alongside a logistic regression classifier resulted in the highest best balanced accuracies of about 87% when using data from 2019. The random forest classifier has worse performance with about 60% accuracy.

Price Sensitivity Model

As we focuses on how to reduce biases in model in pricing models, we believe it is essential to develop a tool kit that has the ability to measure the chances of whether a flight ticket would be a valid and accepted by a given stakeholder (consumer and the airlines), given all the attributes of such a ticket. Not only because such a tool would be able to measure whether a difference between different protected groups exist in terms of price acceptance and the price of that they are being offered, which would allow a thorough investigation in pricing. But also we believe it would be useful for all stakeholders to see whether a given price and a given set of ticket attributes would be realistic or not. Since all the data in the datasets are the prices that are already being accepted by both sides. Therefore, we would be able to use the data from the dataset to measuring the chances of a hypothetical ticket would be accepted, by developing a model to compare the hypothetical ticket with the dataset itself, given we have an ability to determine a variable that would represents the probability of a given entry in the dataset.

The price sensitivity model is a model that predicts on a scale from 0-1 to showcase whether the price would be offered. This is a model based on using the distribution (CDF) of the airfare of a given city-pair origin and destination and the carrier from the DB1B dataset. The CDF serves as an indicator of whether a given airfare would appear to indicate how sensitive a given fare is to the stakeholders (Airline and Passengers), as it is able to capture the probability of whether a given ticket would be valid.

The initial model of such a prediction runs extremely well, our first initial model uses sci-kit learn's linear regression (sklearn.linear_model.LinearRegression), and LightGBM's (lightgbm.LGBMRegressor). Since the large scale of the dataset, we decided the initial model would only train a small subset that is randomly selected from 2018 Q2, and test on another subset of the same quarter. We believe such an approach to test and train our model would appropiate during our initial stage, as we are only finding ways to see what model approach and variable would be a best predictor. Details of the best initial model are shown on the website.

FareClass Model

Another model utilized within model development was the FareClass Model. The FareClass Model is a random forest classifier that predicts if the given ticket is either a favorable or unfavorable FareClass. The classification for the FareClass feature was determined through the distribution of the FarePerMile attribute for each of the FareClass. Observing the distributions, favorable classes had distributions that were centered either below or right at the average FarePerMile for all FarePerMiles values together. Other features incorporated into the model included RoundTrip, DollarCred, Passengers, ItinFare, BulkFare, Distance, MilesFlown, White alone proportion and Non-White alone proportion. The results from this model revealed the limited bias within the dataset.

Here's the results of the following model:

Bias Mitigator	Classifier	Balanced Accuracy	Average Odds Difference	Disparate Impact	Statistical Parity Difference	Equal Odds Difference	Theil
None	Random Forest	0.6088	0.1740	1.2261	0.1707	0.1957	0.154
Reweighing	Random Forest	0.6118	-0.0121	1.1845	-0.0225	0.0059	0.144

The following results were gathered based on the data from the first Q1 of 2022. The Disparate Impact of 1.2261 on the model prior to any mitigation methods suggests a fair model from the very beginning. Now the focus would be to see how Reweighing will improve the fairness of the model. After applying the preprocessing technique of Reweighing, we see the model fairness improves slightly including the balanced accuracy of the model. We see disparate impact reducing to get closer to the ideal value of 1 and equal opportunity difference and theil index reducing from 0.1957 to 0.059 and 0.1542 to 0.1440 respectively. Further details about the fairness metrics are detailed in the next section.

Fairness Metric Introduction

Bias Mitigator	Classifier	BC	AOD	DI	SPD	EOD	TI
None	Logistic Regression	0.876	0.059	1.243	0.166	0.067	0.065

In the next section, we will evaluate our bias-mitigated models using various fairness metrics provided by the AIF360 package. We will give a thorough touch on each metric used to better help comprehend how our models perform. And we will interpret metrics from our baseline model (no bias mitigator, logistic regression) as an example.

Balanced Accuracy (BC): 0.876. Balanced accuracy is the mean between the true positive rate and the true negative rate. It measures the average accuracy obtained from both the minority and the majority class. A score of 0.876 signals a good model performance in identifying negative and positive classes. However, the metric itself offers trivial indications regarding fairness.

Average Odds Difference (AOD): 0.059. The average odds difference value is the average difference in False Positive Rate (FPR) and True Positive Rate (TPR) for unprivileged and privileged groups. A value of 0 would indicate equality of chance of odds. Therefore, the metric with value close to 0 would suggest fairness: The FPR and TPR are relatively similar between privileged groups and unprivileged groups.

Disparate Impact (DI): 1.243. It is the probability of positive classification in the unprivileged group divided by the probability of positive classification in the privileged group. In a fair situation, we expected DI to be close to 1. Therefore, the metric with value of 1.243 signals unfairness even though the results favor the under-privileged groups: the probability of positive classification is significantly higher in the under-privileged group than the privileged group.

Statistical Parity Difference (SPD): 0.166. The metric measures the differences between the probability of positive classification in the unprivileged group and the probability of positive classification in the privileged group. A value that is different from 0 will indicate unfairness as the probability of positive classification is different between privileged and unprivileged groups. A positive value here would agree with the DI value found above, where the probability of positive classification is significantly higher in the under-privileged group than the privileged group.

Equal Opportunity Difference (EOD): 0.067. This metric measures the difference between the TPR of the unprivileged group and the true positive rate of the privileged group. A value that is significantly different than 0 will indicate unfairness as the TPR is different between privileged and unprivileged groups. Therefore, the metric with value close to 0 would suggest fairness: The TPR is relatively similar between privileged group and unprivileged group.

Theil Index (TI): 0.065. Theil index is the generalized entropy index with alpha = 1. It measures an entropic "distance" the population is away from the "ideal" egalitarian state. 0 will indicate perfect equality, and 1 will indicate maximum inequality. Since the value is not exactly 0, the metric reveals at least some level of unfairness.

Bias Analysis & Mitigation

High Price Indicator Model

In order to measure for potential biases, we are using the AI Fairness 360 package. However, throughout the model development process, we have found that there was often little to no bias in the models to mitigate in the first place. The baseline logistic regression model had a disparate impact value of 1.24 and a statistical parity difference of 0.16, which are both signs that our unprivileged class is receiving more favorable outcomes than our privileged class. Using a RandomForestClassificaiton model instead results in less bias, but the performance of the model suffers down to about 60% accuracy. As mentioned prior, the magnitude of these biases is already small but we can still mitigate them in order to see how they would affect the performance of the model.

For example, with a preprocessing technique known as reweighing, where weights are assigned to each group combination prior to classification, our logistic regression model has a disparate impact value of 1 and a statistical parity difference of almost 0 which means that the bias was mitigated. Additionally, the performance only drops down to 86%, which can be considered negligible. This similarly improved the bias mitigation for our Random Forest model too and even resulted in a small boost in accuracy.

Another technique we can apply called prejudice remover is an in-processing technique where a discrimination-aware regularization term is added. This managed to mitigate the bias by a small amount but was largely ineffective in significantly dealing with bias. It did, however, manage to maintain a high accuracy of about 87% which is as good as the accuracy seen in the reweighed logistic regression classifier.

Bias Mitigator	Classifier	BC	AOD	DI	SPD	EOD	TI
None	Logistic Regression	0.8759	0.0598	1.2428	0.1661	0.0678	0.0649
None	Random Forest	0.5945	0.1017	1.0759	0.0698	0.0064	0.0512
Reweighting	Logistic Regression	0.8644	-0.0405	1.0031	0.0023	0.0183	0.0756
Reweighting	Random Forest	0.6106	0.0033	1.0082	0.0077	0.0052	0.0511
Prejudice Remover	Logistic Regression	0.8693	0.03492	1.2126	0.1444	0.05648	0.0720

BC: Best Balanced Accuracy
AOD: Average Odds Difference
DI: Disparate Impact
SPD: Statistical Parity Difference
EOD: Equal Opportunity Difference
TI: Theil Index

With regards to the actual application of a bias mitigator in the model pipeline and the fairness implications, this would add weights to our data with the goal of achieving a disparate impact of 1 at the preprocessing stage or before any actual decision making is done. Here, group affiliation and fairness is taken into consideration for each individual prior to any decision. The cost of a false negative or a false positive in this case would most likely be more severe for our unprivileged class but we have found that there did not seem to be much quantifiable bias in our model results against the unprivileged class to begin with. However, as we are determining our classes by the majority population in the flight origin and destination, the bias we mitigated through reweighing could prove to still be beneficial to the minority populations traveling in the white majority cities as the FarePerMile in general would be lower.

Price Sensitivity Model

To improve accuracy on the protected group compared to the privilege group, reweighting was applied to the protected group to address the issue of bias. As a result, the accuracy of the protected group increased from .85 to .9, indicating some degree of bias mitigation. However, the accuracy of the privilege group decreased from .97 to .94. It is important to note that the accuracy of the original test set remained unchanged, indicating that oversampling did not negatively affect the generalizability of the model. The oversampling technique represents an effort to promote fairness in the model's predictions for both the privilege and protected groups.

Since the oversampling technique work reasonably well on the initial model, therefore, we decided to build our final model based on such an assumption. Details of performance of the final model could be found on the website.

Conclusion

By investigating the potential bias within flight ticket pricing models, we obtain a thorough understanding of current problems within this complicated system. We do discover correlations between pricing differences and demographics. Yet, we did not observe evident bias with respect to race and income through our pricing model development. Regardless, we utilize fairness metrics and bias mitigation methods to create and evaluate a model that is both accurate and unbiased. And we believe fairness metrics, like Disparate Impact, would serve as a great indicator that allows consumers to understand whether they are price discriminated against and whether they are paying the fair price. We believe our findings provide inspiration and a solid foundation for the next step of discoveries, where the flight ticket pricing models can be held under even more stringent inspection when future researchers have access to ticket data with finer details or more advanced computation methods. We think that the variation of price sensitivity between privileged and underprivileged groups would be an interesting research topic that builds on top of our project. We expect that fair pricing models should result in similar price sensitivity between the privileged and underprivileged groups. “Under the current pricing model, do underprivileged groups have to accept limited ranges of high prices, thus having lower price sensitivity compared to the privileged groups?” would be a potential question to ask in that case.

References

Azzolina, Stefano, et al. “Price Discrimination in the Online Airline Market: An Empirical Study.” Journal of Theoretical and Applied Electronic Commerce Research, vol. 16, no. 6, Sept. 2021, pp. 2282–303. Crossref, https://doi.org/10.3390/jtaer16060126.

Lewis, Matthew S. "Identifying airline price discrimination and the effect of competition." International Journal of Industrial Organization 78 (2021): 102761. https://www.sciencedirect.com/science/article/pii/S0167718721000540

Stavins, Joanna. “Price Discrimination in the Airline Market: The Effect of Market Concentration.” Review of Economics and Statistics, vol. 83, no. 1, 2001,pp.200–202., https://doi.org/10.1162/rest.2001.83.1.200.

Williams, Kevin, Dynamic Airline Pricing and Seat Availability (May 26, 2020). Cowles Foundation DiscussionPaperNo.2103R, http://dx.doi.org/10.2139/ssrn.3611696