**Air Travel Report** - Hewei Shen

# Business Problem

Baggage loss or mishandling can always disrupt a perfect travel plan and can lead to an increase in the dissatisfaction rate from clients towards their travel agents. The report will help you understand which factor has the biggest impact on baggage claims, defined as the number of passenger complaints for lost, damaged, or misrouted baggage, for the two main airlines, United Airlines and American Eagle. After reading through the report, you will be able to give wiser advice to your clients regarding their travel plans for the next six months of 2011. The report also gives a 6-month forecast on baggage claims and canceled flights for the two airlines.

# overview of the Data

The data that the report used is the operational data from three airlines for the years 2004 - 2010, recorded at a monthly level. These data are available from the U.S. Department of Transportation, “Air Travel Consumer Report,” the Office of Aviation Enforcement and Proceedings, Aviation Consumer Protection Division. The data for baggage complaints and enplaned passengers cover domestic travel only. For the original data see appendix [1].

The following graphs are historical records of baggage claims and canceled flights for United Airlines and American Eagle from 2004 - 2010. As is shown in the graph, all four metrics showed a seasonal pattern. The number of baggage claims and canceled flights peaked in the summer and winter seasons, while bottomed in the spring and fall seasons.

|  |  |
| --- | --- |
|  |  |
|  |  |

Considering that different airlines may have different scales, it’s hard for us to compare the number of baggage claims between the two airlines. Therefore, we calculated the baggage claims rate for each airline to eliminate the impact of scale. Baggage claims rate = baggage claims / enplaned passengers. We also calculated the average baggage claims rates for the two airlines.

|  |  |
| --- | --- |
|  | Average baggage claims rate for United Airlines: **0.342** |
|  | Average baggage claims rate for American Eagle: **0.722** |

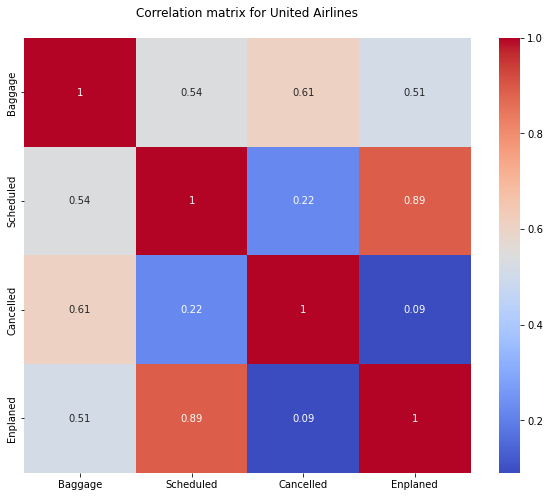
# What influenced baggage claims the most

To understand which factor might influence baggage claims for each airline, we analyzed the correlation among baggage complaints, scheduled flights, canceled flights, and enplaned passengers for United Airlines and American Eagle respectively.

Heatmap is a simple and intuitive way to show the relationship between metrics. In the following heatmaps, blocks with a value close to 1 (or a color close to red) have the highest correlation with each other, while blocks with a value close to 0 (or a color close to dark blue) have the lowest.

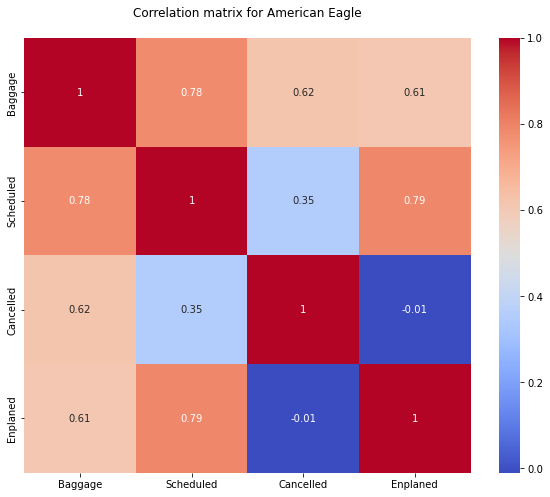
## Key Findings for United Airlines

For United Airways, baggage complaints have the highest correlation with canceled flights, with a positive correlation coefficient of 0.61. That means the number of canceled flights influences the baggage complaints the most among the three factors.



## Key Findings for American Eagle

For American Eagle, baggage complaints have the highest correlation with scheduled flights, with a positive correlation coefficient of 0.78. That means the number of scheduled flights influences the baggage complaints the most among the three factors. Canceled flights and enplaned passengers also have a relatively high impact, with a positive correlation coefficient of over 0.6.



# Forecast for the next 6 months

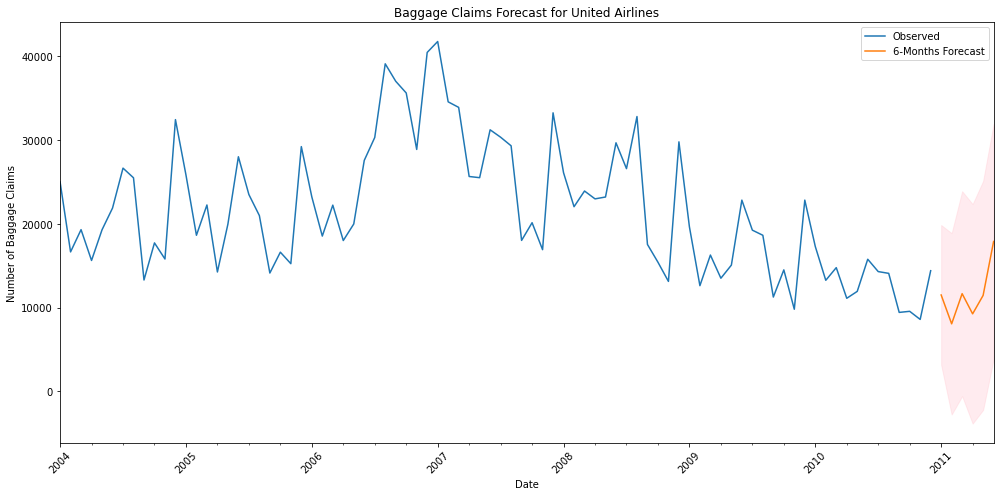
This part shows the 6-month forecast for baggage claims and canceled flights for United Airlines and American Eagle. For detailed summary of models see appendix [2].

In each graph, the blue line is the historical values, and the orange line is the extended forecast part starting from Jan 2011 to Jun 2011. The shaded area in faint red is the 95% confidence interval for the forecasting, which means we have a 95% confidence that the actual value in the following 6 months will be in this shaded area.

In each table, we provided three forecast values for each metric. The column Forecast is the expected mean value. The Lower\_Bound and Upper\_Bound are the lower and upper bound for the confidence interval.

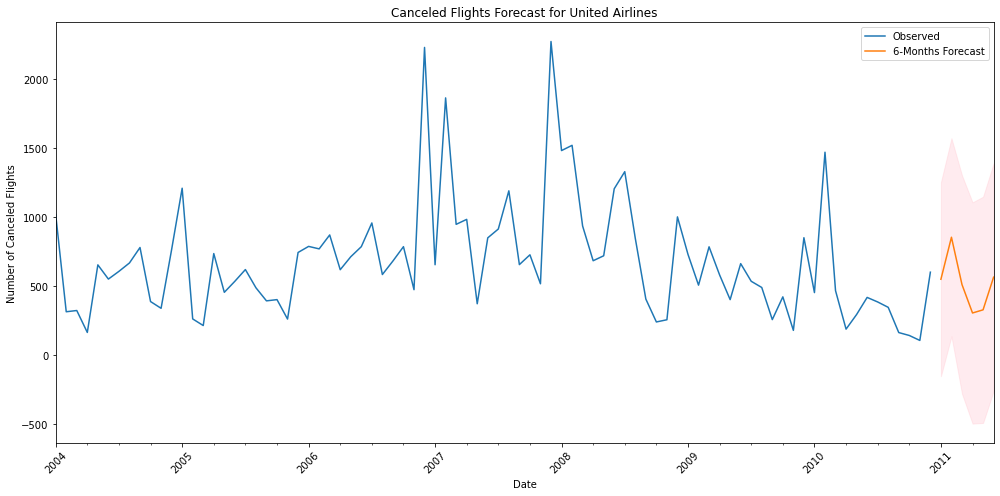
MAPE and RMSE are the two indexes to measure the forecast model's accuracy. The lower these two values are, the more accurate the model is.

## Baggage Claims Forecast for United Airlines



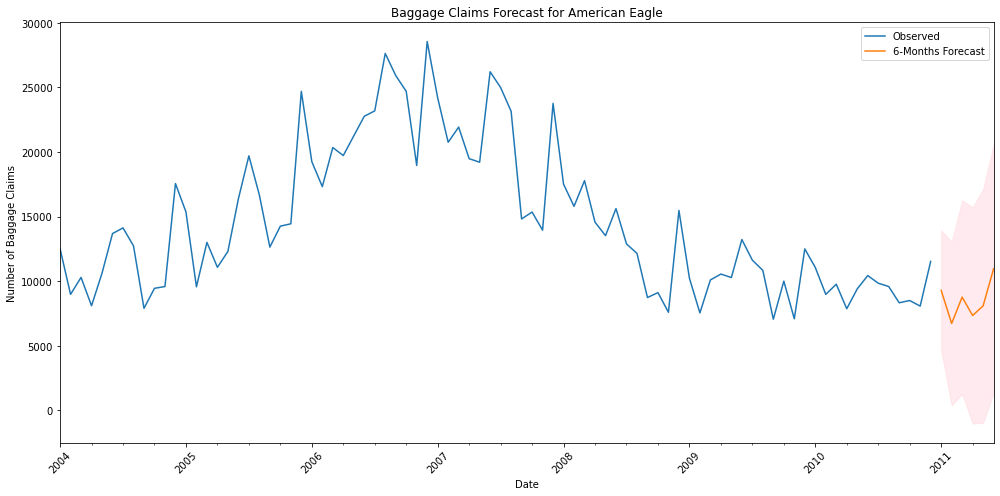
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | **Forecast** | **Lower\_Bound** | **Upper\_Bound** | | --- | --- | --- | --- | | **2011-01-01** | 11517.990008 | 3166.755594 | 19869.224422 | | **2011-02-01** | 8070.936396 | -2757.702052 | 18899.574844 | | **2011-03-01** | 11672.924019 | -560.511300 | 23906.359338 | | **2011-04-01** | 9262.037183 | -3843.209153 | 22367.283519 | | **2011-05-01** | 11463.779640 | -2204.058935 | 25131.618214 | | **2011-06-01** | 17903.990421 | 3865.490765 | 31942.490076 | | MAPE = 17.942%  RMSE = 5,462.825 |

## Canceled flights Forecast for United Airlines



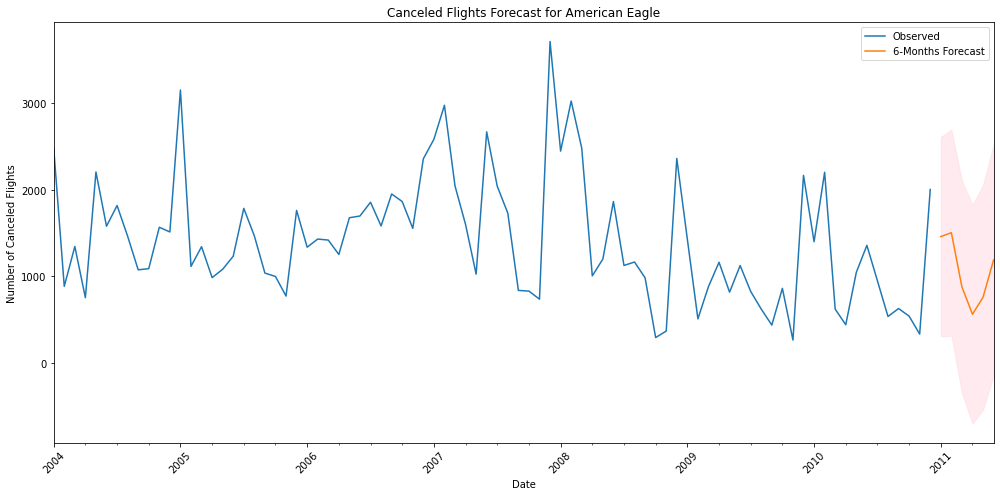
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | **Forecast** | **Lower\_Bound** | **Upper\_Bound** | | --- | --- | --- | --- | | **2011-01-01** | 547.621832 | -155.670041 | 1250.913705 | | **2011-02-01** | 852.413748 | 134.680391 | 1570.147104 | | **2011-03-01** | 509.236256 | -282.371499 | 1300.844010 | | **2011-04-01** | 303.110520 | -499.121728 | 1105.342768 | | **2011-05-01** | 325.908943 | -494.398470 | 1146.216356 | | **2011-06-01** | 562.942854 | -262.429258 | 1388.314966 | | MAPE = 48.698%  RMSE = 380.126 |

## Baggage Claims Forecast for American Eagle



|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | **Forecast** | **Lower\_Bound** | **Upper\_Bound** | | --- | --- | --- | --- | | **2011-01-01** | 9288.020443 | 4638.840878 | 13937.200008 | | **2011-02-01** | 6723.942849 | 388.343373 | 13059.542324 | | **2011-03-01** | 8761.914097 | 1267.439943 | 16256.388252 | | **2011-04-01** | 7338.663764 | -1031.274599 | 15708.602127 | | **2011-05-01** | 8084.123928 | -975.968680 | 17144.216535 | | **2011-06-01** | 10955.028183 | 1337.669734 | 20572.386632 | | MAPE = 14.710%  RMSE = 3,004.210 |

## Canceled flights Forecast for American Eagle



|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | **Forecast** | **Lower\_Bound** | **Upper\_Bound** | | --- | --- | --- | --- | | **2011-01-01** | 1460.406849 | 309.808730 | 2611.004968 | | **2011-02-01** | 1505.445611 | 316.510764 | 2694.380459 | | **2011-03-01** | 879.431004 | -346.615745 | 2105.477752 | | **2011-04-01** | 564.038164 | -698.025005 | 1826.101333 | | **2011-05-01** | 760.709863 | -536.369252 | 2057.788978 | | **2011-06-01** | 1191.057523 | -140.116648 | 2522.231695 | | MAPE = 40.851%  RMSE = 606.978 |

# Recommendations and conclusions

* For passengers who decide to travel by plane in the next six months, February and April will be the best months for those who don’t want to run into baggage loss or mishandling. April and May will be the best months for those who want to avoid canceled flights.
* Forecast models for baggage claims are generally more accurate than models for canceled flights. This is probably because the number of canceled flights is more subject to the impact of weather, which can be very random.
* Passengers who don’t want to run into baggage loss or mishandling may want to choose United Airlines rather than American Eagle because the former had an average smaller rate of baggage claims over the past year.

# Appendix

## [1] original data



|  |  |
| --- | --- |
| **Variable Name** | **Variable Description** |
| Airline | American Eagle, Hawaiian, United |
| Date | Month and year |
| Month | Month |
| Year | Year |
| Baggage | The total number of passenger complaints for lost, damaged, or misrouted baggage |
| Scheduled | The total number of flights scheduled for the airline that month |
| Cancelled | The total number of flights cancelled by that airline that month |
| Enplaned  Baggage\_rate | The total number of passengers who boarded a plane with the airline that month  Baggage / Enplaned |

## [2] summary of ARIMA model

