

# Air Travel Report - Will my baggage make it to my final destination?

By [Hewei Shen](#)

Created on Oct 22, 2023

## BUSINESS PROBLEM

Baggage loss or mishandling can significantly disrupt travel plans and lead to increased client dissatisfaction, often reflecting poorly on travel agents. This report provides a detailed analysis of the **factors that most influence baggage claims**—specifically, the number of passenger complaints related to lost, damaged, or misrouted baggage—for two major airlines, **United Airlines and American Eagle**.

By reviewing this report, you will gain valuable insights that enable you to offer more informed advice to your clients for their travel plans over the next six months of 2011. Additionally, the report includes a **6-month forecast for baggage claims and canceled flights** for these two airlines, helping you anticipate potential issues and better serve your clients' needs.

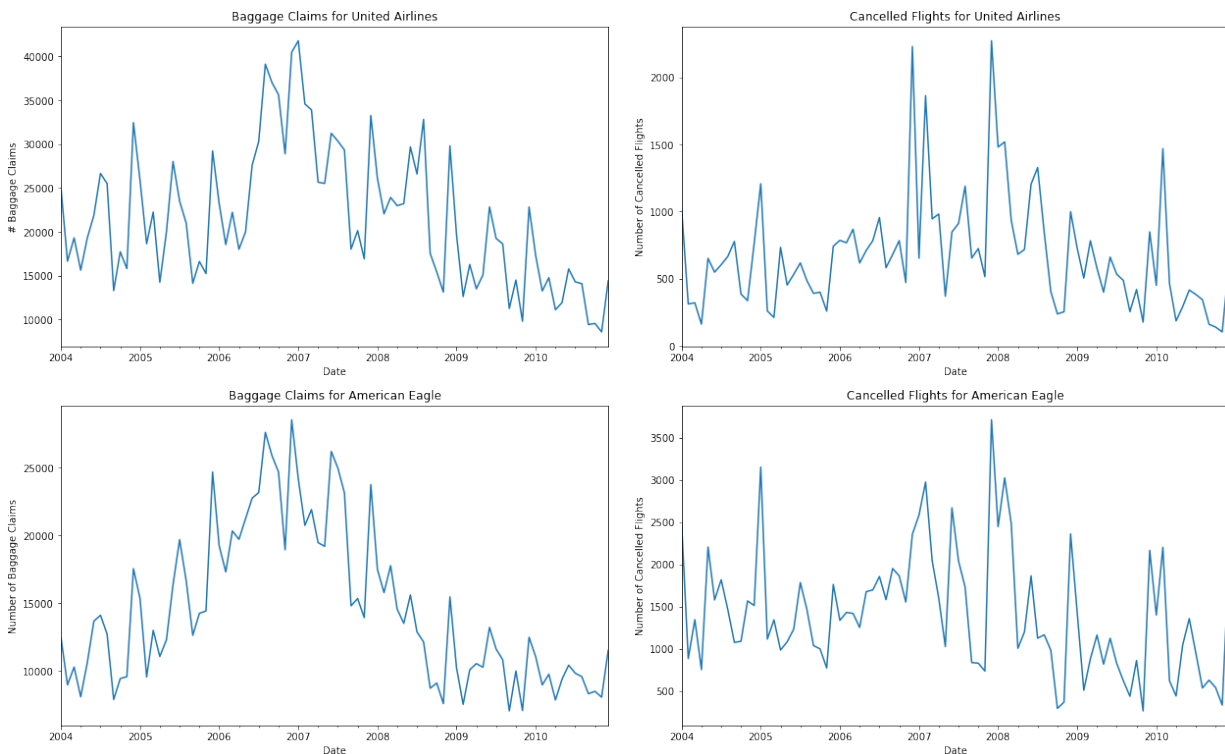
## CONCLUSIONS

- For passengers planning to fly in the next six months, **February and April** are the best months to minimize the risk of baggage loss or mishandling, while **June** is projected to be the worst month.
- For those looking to avoid canceled flights, **April and May** are the optimal months, whereas **February** is expected to be the worst month for flight cancellations.
- The forecast models for baggage claims are generally more accurate than those for canceled flights. This is largely because canceled flights are highly influenced by weather conditions, which are unpredictable and can lead to greater variability.
- Passengers seeking to avoid baggage loss or mishandling should consider flying with **United Airlines** rather than American Eagle, as United has consistently shown a lower average baggage claims rate over the past year.

## DATA OVERVIEW

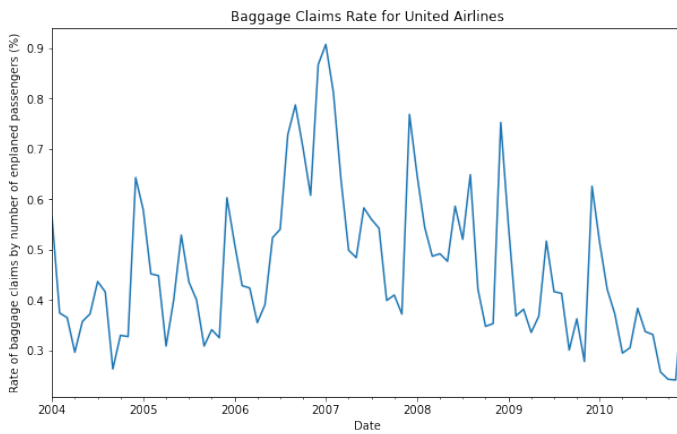
The data used in this report comprises operational records from three airlines, covering the period **from 2004 to 2010**, with **monthly-level** details. This data is sourced from the "Air Travel Consumer Report," specifically from the U.S. Department of Transportation's, Office of Aviation Enforcement and Proceedings. Notably, the data on baggage complaints and enplaned passengers cover domestic travel only. For the original dataset, please refer to Appendix [1].

The following graphs present historical trends for baggage claims and canceled flights for United Airlines and American Eagle from 2004 to 2010. All four metrics exhibit **clear seasonal patterns**, with the number of baggage claims and canceled flights peaking during the summer and winter seasons, and reaching their lowest during the spring and fall seasons.

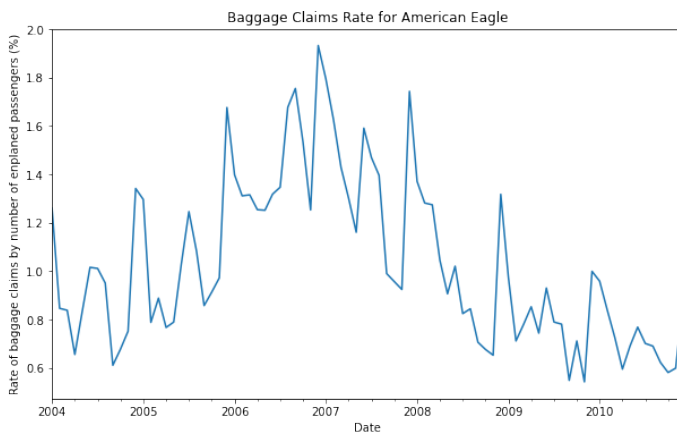


Since airlines operate at different scales, directly comparing the number of baggage claims between United Airlines and American Eagle can be challenging. To address this, we calculated the **baggage claims rate** for each airline, which allows for a fair comparison by eliminating the impact of varying passenger volumes. The baggage claims rate is calculated as follows:

$$\text{Baggage claims rate} = (\text{Baggage claims}) / (\text{Enplaned passengers})$$



Average baggage claims rate for United Airlines airline: **0.34%**



Average baggage claims rate for American Eagle airline: **0.72%**

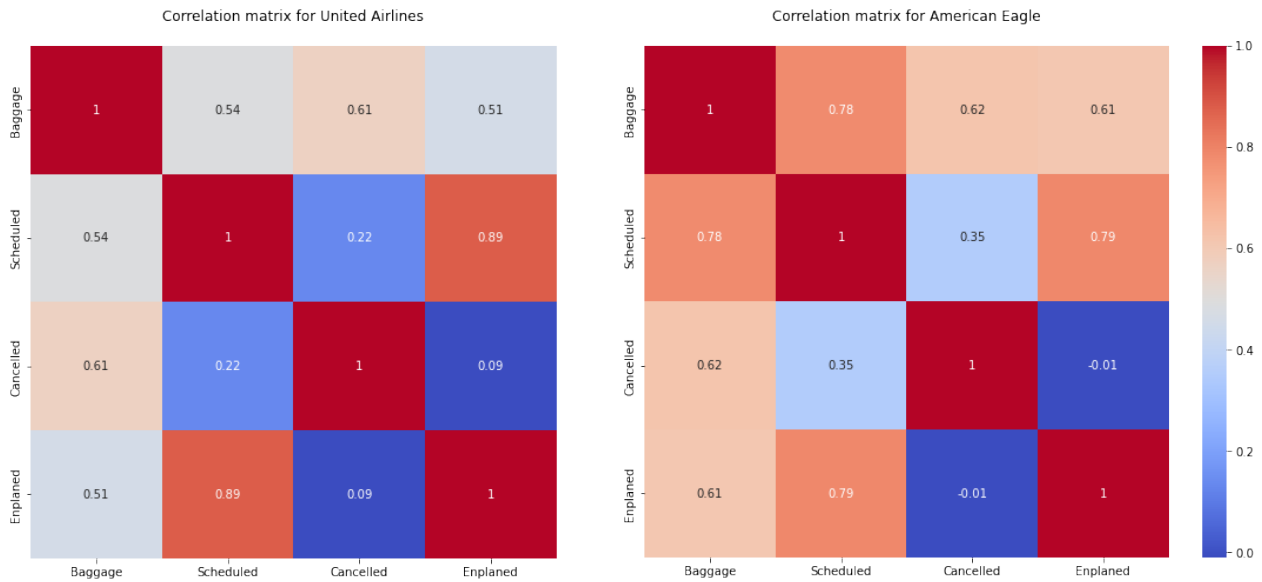
Additionally, we computed the **average baggage claims rate** for both airlines from 2004 to 2010. Overall, American Eagle has a significantly higher rate of baggage complaints per customer, more than double that of United Airlines. This highlights a greater likelihood of baggage issues for American Eagle passengers compared to those flying with United.

## WHAT INFLUENCED BAGGAGE CLAIMS

To identify the factors that may influence baggage claims for each airline, we conducted a **correlation analysis** between several key metrics: baggage complaints, scheduled flights, canceled flights, and enplaned passengers for both United Airlines and American Eagle.

A **heatmap** provides a simple and intuitive visualization of these relationships. In the following heatmaps, blocks with values close to 1 (or colors approaching red) indicate a strong positive correlation between baggage claims and the specific metric, while blocks with values close to 0 (or colors approaching dark blue) reflect little to no correlation.

For United Airlines (left graph), baggage complaints show the strongest correlation with **canceled flights**, with a positive coefficient of 0.61. This indicates that among the three factors analyzed, the number of canceled flights has the greatest influence on baggage complaints.



In contrast, for American Eagle (right graph), baggage complaints are most strongly correlated with **scheduled flights**, displaying a positive coefficient of 0.78. Additionally, both canceled flights and enplaned passengers show a significant influence, each with a coefficient above 0.6.

When comparing United Airlines and American Eagle, it's evident that United's baggage complaints are primarily driven by the inconvenience of canceled flights. In contrast, American Eagle shows a higher baggage complaint rate in its regular daily operations. This suggests that even when passengers are successfully onboarded, they still face a significant risk of baggage loss, damage, or mishandling.

## 6 MONTHS FORECAST

This section presents a 6-month forecast for baggage claims and canceled flights for United Airlines and American Eagle. For a detailed model summary, please refer to Appendix [2].

In each graph, the blue line represents historical values, while the orange line shows the extended forecast from January 2011 to June 2011. The shaded area in faint red indicates the 95% confidence interval, meaning we are 95% confident that the actual values over the next six months will fall within this range.

Additionally, each table provides three forecast values for each metric:

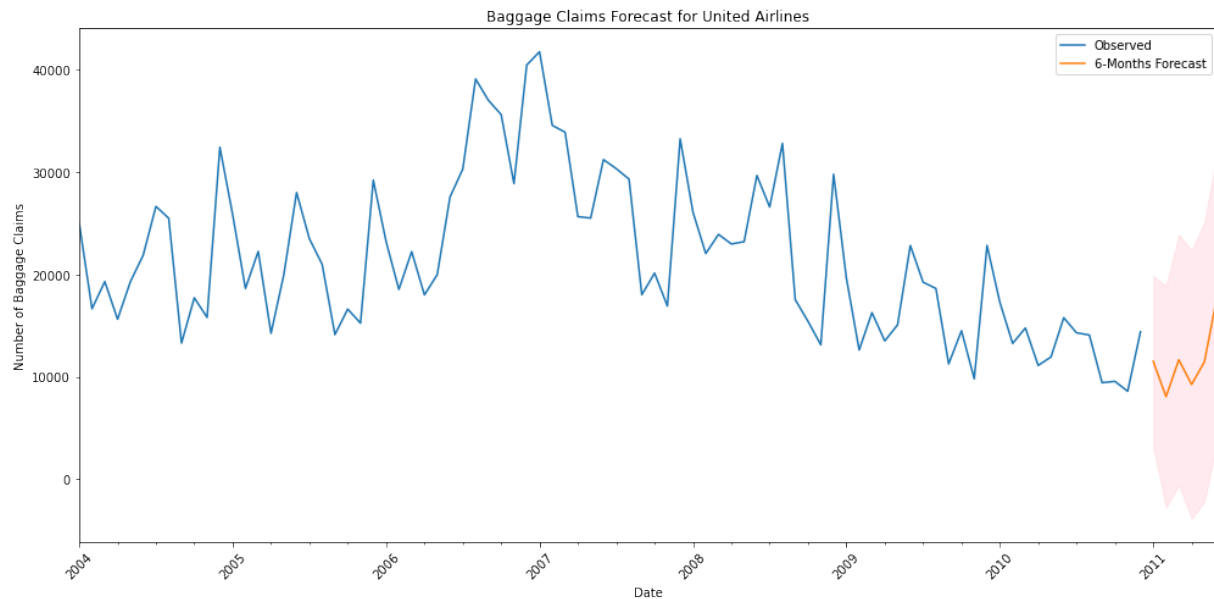
**Forecast:** The expected mean value

**Lower Bound:** The lower limits of the confidence interval

**Upper Bound:** The upper limits of the confidence interval

To assess the accuracy of the forecast models, we use two key indicators: MAPE (Mean Absolute Percentage Error) and RMSE (Root Mean Square Error). Lower values for both MAPE and RMSE indicate higher model accuracy.

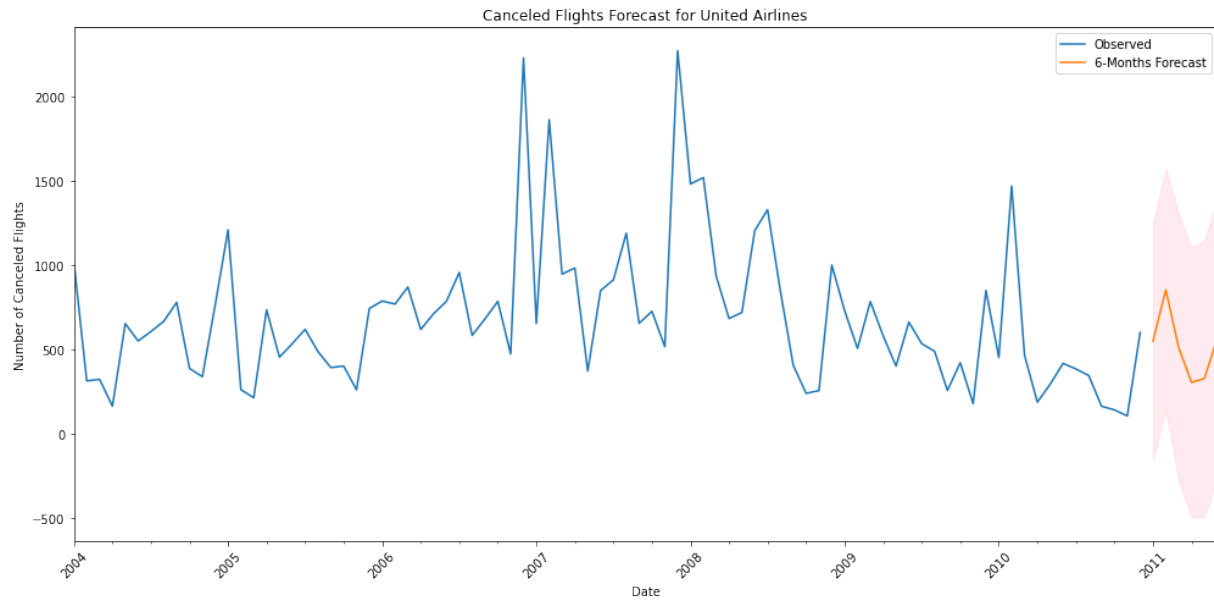
## BAGGAGE CLAIMS - UNITED AIRLINES



	Forecast	Lower Bound	Upper Bound
<b>2011-01-01</b>	11517.99	3166.76	19869.22
<b>2011-02-01</b>	8070.94	-2757.70	18899.57
<b>2011-03-01</b>	11672.92	-560.51	23906.36
<b>2011-04-01</b>	9262.04	-3843.21	22367.28
<b>2011-05-01</b>	11463.78	-2204.06	25131.62
<b>2011-06-01</b>	17903.99	3865.49	31942.49

MAPE = 17.94%  
RMSE = 5,462.83

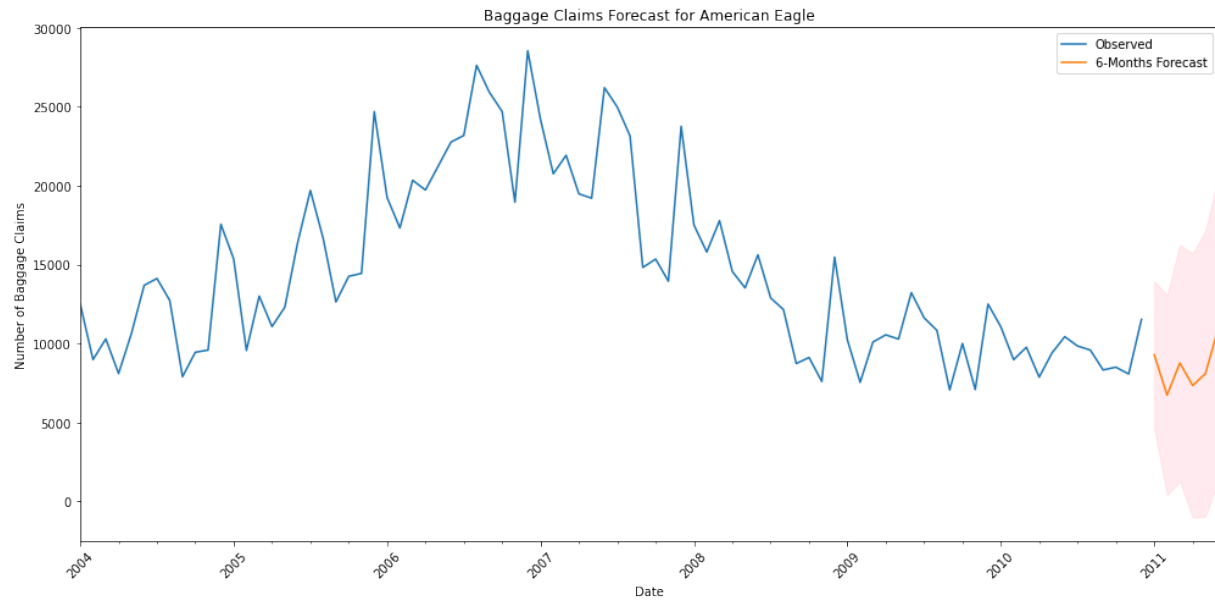
## CANCELED FLIGHTS - UNITED AIRLINES



	Forecast	Lower Bound	Upper Bound
<b>2011-01-01</b>	547.62	-155.67	1250.91
<b>2011-02-01</b>	852.41	134.68	1570.15
<b>2011-03-01</b>	509.24	-282.37	1300.84
<b>2011-04-01</b>	303.11	-499.12	1105.34
<b>2011-05-01</b>	325.91	-494.40	1146.22
<b>2011-06-01</b>	562.94	-262.43	1388.31

MAPE = 48.70%  
RMSE = 380.13

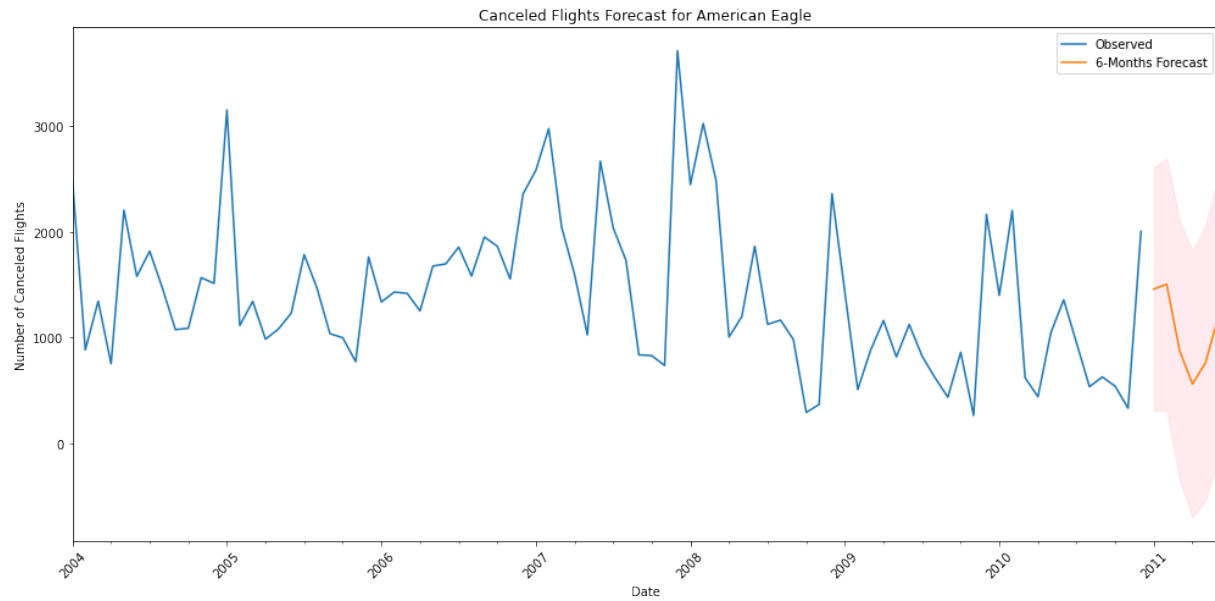
## BAGGAGE CLAIMS - AMERICAN EAGLE



	Forecast	Lower Bound	Upper Bound
<b>2011-01-01</b>	9288.02	4638.84	13937.20
<b>2011-02-01</b>	6723.94	388.34	13059.54
<b>2011-03-01</b>	8761.91	1267.44	16256.39
<b>2011-04-01</b>	7338.66	-1031.27	15708.60
<b>2011-05-01</b>	8084.12	-975.97	17144.22
<b>2011-06-01</b>	10955.03	1337.67	20572.39

MAPE = 14.71%  
RMSE = 3,004.21

## CANCELED FLIGHTS - AMERICAN EAGLE



	Forecast	Lower Bound	Upper Bound
<b>2011-01-01</b>	1460.406849	309.808730	2611.004968
<b>2011-02-01</b>	1505.445611	316.510764	2694.380459
<b>2011-03-01</b>	879.431004	-346.615745	2105.477752
<b>2011-04-01</b>	564.038164	-698.025005	1826.101333
<b>2011-05-01</b>	760.709863	-536.369252	2057.788978
<b>2011-06-01</b>	1191.057523	-140.116648	2522.231695

MAPE = 40.85%  
RMSE = 606.98



## APPENDIX

### [1] ORIGINAL DATA

	Airline	Date	Month	Year	Baggage	Scheduled	Cancelled	Enplaned	Baggage_rate
0	American Eagle	01/2004	1	2004	12502	38276	2481	992360	1.259825
1	American Eagle	02/2004	2	2004	8977	35762	886	1060618	0.846393
2	American Eagle	03/2004	3	2004	10289	39445	1346	1227469	0.838229
3	American Eagle	04/2004	4	2004	8095	38982	755	1234451	0.655757
4	American Eagle	05/2004	5	2004	10618	40422	2206	1267581	0.837659
...	...	...	...	...	...	...	...	...	...
247	United	08/2010	8	2010	14099	30637	344	4263211	0.330713
248	United	09/2010	9	2010	9435	28072	161	3679517	0.256420
249	United	10/2010	10	2010	9565	29144	140	3952549	0.241996
250	United	11/2010	11	2010	8597	27318	104	3573268	0.240592
251	United	12/2010	12	2010	14415	27619	599	3493643	0.412607

252 rows x 9 columns

#### Variable Name      Variable Description

<b>Airline</b>	American Eagle, Hawaiian, United
<b>Date</b>	Month and year
<b>Month</b>	Month
<b>Year</b>	Year
<b>Baggage</b>	The total number of passenger complaints for lost, damaged, or misrouted baggage
<b>Scheduled</b>	The total number of flights scheduled for the airline that month
<b>Cancelled</b>	The total number of flights cancelled by that airline that month
<b>Enplaned</b>	The total number of passengers who boarded a plane with the airline that month
<b>Baggage_rate</b>	Baggage / Enplaned

## [2] MODEL SUMMARY

### Summuary of ARIMA model for United Airlines, Baggage Claims

SARIMAX Results						
Dep. Variable:		y	No. Observations:		84	
Model:		ARIMA(1, 0, 0)x(0, 1, [1], 12)	Log Likelihood		-706.585	
Date:		Sun, 22 Oct 2023	AIC		1419.170	
Time:		10:52:27	BIC		1426.000	
Sample:		01-01-2004	HQIC		1421.890	
		- 12-01-2010				
Covariance Type:		opg				
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.8258	0.035	23.635	0.000	0.757	0.894
ma.S.L12	-0.6496	0.116	-5.580	0.000	-0.878	-0.421
sigma2	1.813e+07	1.72e-09	1.05e+16	0.000	1.81e+07	1.81e+07
Ljung-Box (L1) (Q):			0.48	Jarque-Bera (JB):		17.36
Prob(Q):			0.49	Prob(JB):		0.00
Heteroskedasticity (H):			0.30	Skew:		-0.47
Prob(H) (two-sided):			0.00	Kurtosis:		5.21

### Summuary of ARIMA model for United Airlines, Cancelled Flights

SARIMAX Results						
Dep. Variable:	y			No. Observations:	84	
Model:	ARIMA(2, 0, 0)x(0, 1, [1], 12)			Log Likelihood	-528.251	
Date:	Sun, 22 Oct 2023			AIC	1064.501	
Time:	10:48:39			BIC	1073.608	
Sample:	01-01-2004			HQIC	1068.127	
	- 12-01-2010					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.2038	0.091	2.246	0.025	0.026	0.382
ar.L2	0.4335	0.086	5.030	0.000	0.265	0.602
ma.S.L12	-0.6021	0.110	-5.452	0.000	-0.819	-0.386
sigma2	1.287e+05	1.78e+04	7.243	0.000	9.39e+04	1.63e+05
Ljung-Box (L1) (Q):			0.05	Jarque-Bera (JB):		14.42
Prob(Q):			0.82	Prob(JB):		0.00
Heteroskedasticity (H):			0.74	Skew:		0.90
Prob(H) (two-sided):			0.47	Kurtosis:		4.25

#### Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

# Summary of ARIMA model for American Eagle, Baggage Claims

## SARIMAX Results

Dep. Variable:	y	No. Observations:	84			
Model:	ARIMA(1, 0, 0)x(0, 1, [1], 12)	Log Likelihood	-669.273			
Date:	Sun, 22 Oct 2023	AIC	1344.546			
Time:	11:09:38	BIC	1351.376			
Sample:	01-01-2004	HQIC	1347.265			
	- 12-01-2010					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]
-----						
ar.L1	0.9322	0.025	37.466	0.000	0.883	0.981
ma.S.L12	-0.8991	0.167	-5.397	0.000	-1.226	-0.573
sigma2	5.34e+06	3.34e-08	1.6e+14	0.000	5.34e+06	5.34e+06
=====						
Ljung-Box (L1) (Q):		0.40	Jarque-Bera (JB):		5.16	
Prob(Q):		0.53	Prob(JB):		0.08	
Heteroskedasticity (H):		0.24	Skew:		-0.21	
Prob(H) (two-sided):		0.00	Kurtosis:		4.24	
=====						

# Summary of ARIMA model for American Eagle, Cancelled Flights

## SARIMAX Results

Dep. Variable:	y	No. Observations:	84			
Model:	ARIMA(0, 1, 1)x(0, 1, 1, 12)	Log Likelihood	-555.732			
Date:	Sun, 22 Oct 2023	AIC	1117.464			
Time:	11:11:26	BIC	1124.252			
Sample:	01-01-2004	HQIC	1120.163			
	- 12-01-2010					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]
-----						
ma.L1	-0.7398	0.080	-9.236	0.000	-0.897	-0.583
ma.S.L12	-0.6459	0.131	-4.924	0.000	-0.903	-0.389
sigma2	3.44e+05	5.81e+04	5.920	0.000	2.3e+05	4.58e+05
=====						
Ljung-Box (L1) (Q):		0.04	Jarque-Bera (JB):		5.88	
Prob(Q):		0.83	Prob(JB):		0.05	
Heteroskedasticity (H):		0.97	Skew:		0.41	
Prob(H) (two-sided):		0.94	Kurtosis:		4.14	
=====						