Navigating the Skies: A Study on Airline Delays in the Aviation Industry
By: Alex Pettis

I. Industry Overview

The aviation industry encompasses all activities related to mechanical air transportation, primarily using aircraft such as airplanes, helicopters, and drones. It includes various sectors and businesses, including airlines, aircraft manufacturers, researchers, air safety specialists, and military aviation. The industry is crucial for global connectivity, facilitating travel and trade between continents, countries, and cultures. It plays a significant role in global economic prosperity by boosting local economies through tourism and enhancing global trade efficiency.

One of the industry's key contributions is job creation, providing employment for millions of people worldwide in various roles, from pilots and cabin crew to aerospace engineers and air traffic controllers. Additionally, the aviation industry indirectly supports jobs in the travel and tourism sector. The aviation industry is divided into civil and military aviation. Civil aviation includes private and commercial air travel, while military aviation involves aircraft used in military settings for aerial warfare and surveillance operations.

Key challenges facing the aviation industry include workforce shortages, safety and security concerns, sustainability issues, complex regulatory environments, and the need to adapt to technological advancements. Major aircraft manufacturing companies, such as Airbus, Boeing, and Lockheed Martin, play a crucial role in the aviation industry, producing aircraft for civilian and military markets. These companies drive innovation and technology advancements in aviation.

The airline industry, a subset of the aviation industry, includes businesses offering air transportation services for passengers and cargo. It can be categorized into international, national, and regional airlines, each serving different markets and operating with varying fleet sizes and revenue streams. Overall, the aviation industry is essential for global connectivity, economic growth, and job creation, making it a vital sector in the modern world.

II. Sense the Problem Broadly

The aviation industry faces a myriad of challenges that impact its operations and performance. Foremost among these challenges it the persistent issue or workforce shortages, encompassing pilots, mechanics, air traffic controllers, and ground crews. Ensuring the safety and security of passengers, crews, and the public remains a top priority, necessitating ongoing efforts and investment. Additionally, airlines are increasingly focusing on sustainability, striving to reduce their carbon footprint through technological advancements and operational improvements.

Navigating the complex regulatory environment is another significant challenge for the industry, as it must balance stringent safety requirements with the imperative for innovation and efficiency. The emergence of new technologies, such as supersonic flight and electric air taxis, presents both opportunities and challenges regarding regulation, safety, and public acceptance.

On the operational front, the industry grapples with issues like flight delays and cancellations, driven in part by workforce shortages and other operational issues. Moreover, the infrastructure of airports and air traffic control systems may struggle to keep pace with the growing demand for air travel. Disruptions in the supply chain, including shortages or parts or materials, can further impact the industry’s ability to maintain and operate aircraft efficiently.

As the industry becomes increasingly reliant on digital technologies, cybersecurity threats pose a significant risk to operations and data security. Meeting customer expectations for convenience, comfort, and reliability is an ongoing challenge, particularly during times of disruption.

Addressing these challenges and operational problems is essential for the aviation industry to improve the overall travel experience, enhance safety and security, reduce environmental impact, and maintain a competitive edge, Closing the performance gap requires strategic planning, innovation, and collaboration among industry stakeholders, government agencies, and regulatory bodies.

III. Define a Specific Outcome You Want to Achieve and Define the Performance Gap

A flight is considered 'on time' according to the Bureau of Transportation Statistics if it arrives at the gate less than 15 minutes later than the scheduled time shown in the carriers' Computerized Reservations Systems (CRS). Departure performance is based on the time of departure from the gate. Based on this information from the Bureau of Transportation Statistics there is just over 20% of all flights from June 2003 to January 2024 that are delayed. Airlines categorize the causes of flight delays and cancellations into five main categories. The "Air Carrier" category includes issues within the airline's control, such as maintenance problems, crew scheduling issues, baggage handling errors, and fueling problems. "Extreme Weather" covers delays caused by severe meteorological conditions like tornadoes, blizzards, and hurricanes, which are beyond the airline's control. "National Aviation System (NAS)" delays are due to broader system issues like non-extreme weather conditions, airport operations, heavy traffic, and air traffic control problems. "Late-Arriving Aircraft" delays occur when a previous flight using the same aircraft arrives late, causing a delayed departure for the next flight. Finally, "Security" delays are caused by security-related incidents such as terminal evacuations, security breaches, and long screening lines. These categories help airlines and authorities analyze and address the causes of delays to improve overall flight performance and passenger satisfaction.

To enhance overall flight performance and elevate passenger satisfaction, our primary objective is to minimize the incidence of flight delays across the board. This entails implementing targeted strategies and operational enhancements aimed at achieving a substantial reduction in delays, thereby bringing us closer to the industry benchmark of consistently achieving on-time performance. By proactively addressing delays and optimizing operational efficiency, we aim to deliver a more reliable and enjoyable travel experience for our passengers, reinforcing our commitment to excellence in air travel.

IV. Collect Data to Analyze the Status Quo

We analyze the operational performance of Delta Airlines and Frontier Airlines by examining specific criteria such as Carrier Delays, National Aviation System (NAS) Delays, and Late Arrival Delays. The Factor Weighted Table (Figure 1) presented below provides a structured framework for evaluating these performance metrics, with each criterion assigned a weight reflecting its importance. By comparing the scores of Delta and Frontier Airlines on each criterion, derived from both subjective assessments and objective operational data, the aim is to identify areas for improvement and strategic priorities for enhancing operational efficiency, customer satisfaction, and overall competitiveness within the aviation industry.

Criteria	Criteria Weight	Criteria Score (1-10)	Total	Criteria Score (1-10)	Total
		Delta Airlines		Frontier Airlines	
Carrier Delays	0.20	9	1.80	7	1.40
Weather Delays	0.10	4	0.40	3	0.30
NAS Delays	0.15	7	1.05	8	1.20
Security Delays	0.05	1	0.05	1	0.05
Late Arrival Delays	0.40	6	2.40	9	3.60
Cancelled Flights	0.10	3	0.30	4	0.40
			6		6.95

Figure 1: Factor Weighted Table Comparing Operational Performance Criteria for Delta Airlines and Frontier Airlines

The comparison between the Factor Weighted Table and the % of Total Operations data provides a comprehensive understanding of airlines' operational performance, particularly within the airline industry. By analyzing specific criteria such as Carrier Delays, NAS Delays, and Late Arrival Delays, derived from both subjective assessments and objective operational data, airlines can assess their strengths and weaknesses in key operational areas.

For instance, looking at Carrier Delays, which have a weight of 0.20 in the Factor Weighted Table, Delta received a score of 9, resulting in a total weighted score of 1.80, while Frontier scored 7, with a total weighted score of 1.40. Correspondingly, Delta has an Air Carrier Delay percentage of 6.68%, while Frontier's percentage is higher at 11.81%. This alignment between the Factor Weighted Table scores and the objective operational data indicates that Frontier experiences more Carrier Delays than Delta, highlighting an area for potential improvement for Frontier.

Similarly, for NAS Delays, with a weight of 0.15, Delta scored 7, resulting in a total weighted score of 1.05, while Frontier scored 8, with a total weighted score of 1.20. Comparing this with the % of Total Operations data, Delta has a NAS Delay percentage of 5.98%, slightly lower than Frontier's 7.74%. Again, this correspondence suggests that Frontier experiences more NAS Delays than Delta, as reflected in both the Factor Weighted Table and the operational data.

Late Arrival Delays, with the highest weight of 0.40, show a similar trend. Delta scored 6, resulting in a total weighted score of 2.40, while Frontier scored 9, with a total weighted score of 3.60. In the % of Total Operations data, Delta has an Aircraft Arriving Late percentage of 4.94%, while Frontier's percentage is higher at 7.36%. This consistency between the Factor Weighted Table and the operational data indicates that Frontier experiences more Late Arrival Delays compared to Delta.

Operational performance is paramount in the aviation industry, directly impacting customer satisfaction, safety, and financial stability. By evaluating specific criteria such as Carrier Delays, NAS Delays, and Late Arrival Delays, airlines can gain insights into their performance and make informed decisions to improve efficiency and reliability.

Operational delays, such as those caused by carrier or NAS issues, can lead to passenger frustration and dissatisfaction. By identifying areas of weakness through analyses like the Factor Weighted Table, airlines can implement strategies to enhance reliability and customer experience. This, in turn, fosters loyalty and positive word-of-mouth, crucial for long-term success in the competitive airline industry.

Operational performance is closely tied to safety in aviation. Delays and operational issues can compromise safety protocols or increase risks. By addressing areas of concern highlighted by performance analyses, airlines can enhance safety standards, ensuring compliance with regulatory requirements and maintaining a strong safety record, which is paramount for maintaining public trust and confidence.

Efficient operations are essential for airlines to remain competitive and profitable. High rates of delays and cancellations disrupt schedules, increase costs, and diminish overall efficiency. Analyzing performance data enables airlines to identify inefficiencies and implement strategies to streamline operations, reduce delays, and optimize resource utilization, ultimately improving financial performance and sustainability.

Comparing performance metrics with industry benchmarks and competitors is vital for airlines to gauge their standing within the market. The Factor Weighted Table provides a structured framework for benchmarking and allows airlines to identify best practices and areas for improvement relative to their competitors. This competitive analysis informs strategic decision-making and drives continuous improvement efforts, crucial for maintaining competitiveness and operational excellence in the dynamic aviation industry.

Overall, this integrated analysis enables airlines to identify areas for improvement and prioritize strategic initiatives to enhance operational efficiency, customer satisfaction, and overall competitiveness within the aviation industry. By combining subjective assessments with objective operational data, airlines can gain a more accurate understanding of their performance and take targeted actions to drive continuous improvement. This analysis is highly relevant to the aviation industry, enabling airlines to evaluate their performance comprehensively, identify areas for improvement, and implement strategies to enhance operational efficiency, safety, and customer satisfaction. By leveraging both subjective assessments and objective operational data, airlines can make informed decisions to drive continuous improvement and maintain a competitive edge in the ever-evolving airline sector.

V. Consider Opportunities for Improvements

Figure 2 (Below) shows the delay caused by year as a percent of total delay minutes from 2003 to 2023. To help improve operations we could forecast for each of the factors that are causing delays and then determine which factor should have the most resources allocated towards it.

Year	Air Carrier Delay	Aircraft Arriving	National Aviation	Security Delay	Extreme Weather
2003	26.33	30.86	36.45	0.25	6.11
2004	25.77	33.61	33.48	0.25	6.89
2005	28.03	34.19	31.43	0.18	6.16
2006	27.82	36.98	29.37	0.25	5.57
2007	28.54	37.65	27.94	0.18	5.69
2008	27.76	36.55	30.21	0.13	5.35
2009	28.04	36.22	30.63	0.12	4.98
2010	30.38	39.39	25.66	0.17	4.40
2011	30.08	40.83	24.81	0.13	4.15
2012	31.92	41.41	22.54	0.13	4.01
2013	29.38	42.11	24.22	0.14	4.13
2014	30.23	41.93	23.49	0.09	4.25
2015	32.20	39.84	22.88	0.13	4.95
2016	32.64	39.20	23.68	0.14	4.35
2017	31.17	39.36	25.07	0.14	4.25
2018	30.06	39.63	24.55	0.14	5.62
2019	30.61	39.71	24.03	0.14	5.51
2020	42.00	29.20	21.70	0.22	7.00
2021	40.80	35.30	16.70	0.30	6.90
2022	39.80	37.70	16.80	0.20	5.60
2023	36.40	40.00	18.10	0.20	5.20

Figure 2. Current data on the Delay Cause by Year, as a Percent of Total Delay Minutes from 2003 to 2023

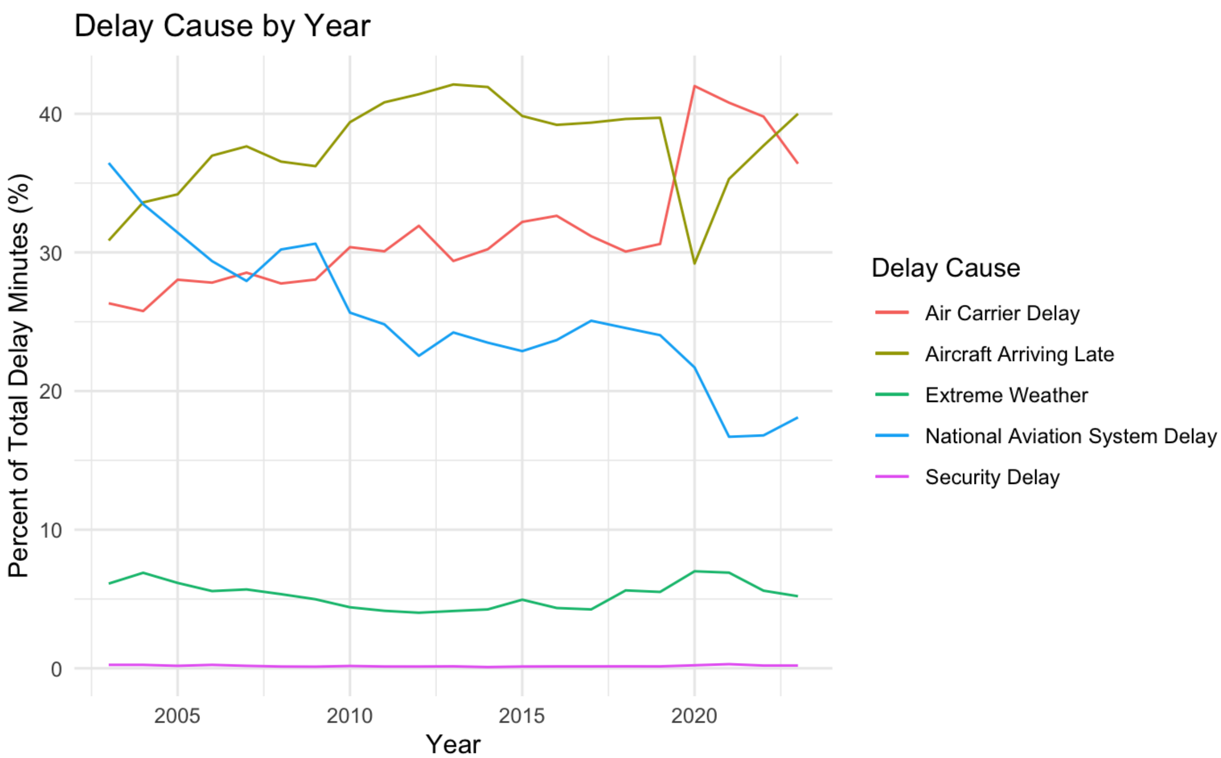


Figure 3. Graph of the current data of Delay Cause by Year, as a Percent of Total Delay Minutes from 2003 to 2023

This graph shows the current outlook of the data we have from 2003 to 2023 of factors affecting flight delays. To forecast for each of the factors contributing to the delays we would want to use either regression or exponential smoothing methods. To determine which method is better we would want to consider the forecast accuracy using both the MASE and MAPE. By analyzing forecast accuracy metrics such as MASE and MAPE for both regression and exponential smoothing methods, we can identify the most effective approach for predicting future trends in flight delays, thereby enabling proactive measures to mitigate potential disruptions and improve overall operational efficiency.

Model	MASE	MAPE
Regression	3.855774	11.94303
Exponential Smoothing	4.967502	14.93985

Table 1. Forecast Accuracy for Air Carrier Delays

Based on the forecast accuracy metrics presented in Table 1, regression emerges as the preferred method for forecasting Air Carrier Delays, as indicated by its lower values of Mean Absolute Scaled Error (MASE) and Mean Absolute Percentage Error (MAPE) compared to

Exponential Smoothing. This choice ensures more precise predictions, allowing for better-informed decision-making and proactive management of air carrier delays.

Model	MASE	MAPE
Exponential Smoothing	2.326252	8.975497
Regression	6.040573	21.561436

Table 2. Forecast Accuracy for Aircrafts Arriving Late

Exponential Smoothing outperforms Regression for forecasting Aircraft Arriving Late. With a lower Mean Absolute Scaled Error (MASE) of 2.33 and Mean Absolute Percentage Error (MAPE) of 8.98%, Exponential Smoothing demonstrates superior accuracy compared to Regression, which has a higher MASE of 6.04 and MAPE of 21.56%. This indicates that Exponential Smoothing produces smaller forecast errors and less deviation from the actual values, making it the preferred method for this forecasting task.

Model	MASE	MAPE
Regression	1.366412	11.06366
Exponential Smoothing	2.717201	26.63977

Table 3. Forecast Accuracy for NAS Delays

In comparing Regression and Exponential Smoothing for forecasting National Aviation System Delays, Regression shows a lower Mean Absolute Scaled Error (MASE) of 1.37 and a Mean Absolute Percentage Error (MAPE) of 11.06%. In contrast, Exponential Smoothing exhibits a higher MASE of 2.72 and a significantly higher MAPE of 26.64%. These metrics suggest that Regression outperforms Exponential Smoothing in accuracy, making it the preferred method for this forecasting task.

Model	MASE	MAPE
Exponential Smoothing	1.873601	27.18698
Regression	3.776717	60.53097

Table 4. Forecast Accuracy for Security Delays

In comparing Exponential Smoothing and Regression for forecasting Security Delays, Exponential Smoothing shows superior performance with a lower Mean Absolute Scaled Error (MASE) of 1.87 and a Mean Absolute Percentage Error (MAPE) of 27.19%, compared to Regression's higher MASE of 3.78 and MAPE of 60.53%. These results indicate that Exponential Smoothing generates more accurate forecasts with smaller errors and lower percentage deviations from actual values, making it the preferred method for forecasting Security Delays.

Model	MASE	MAPE
Exponential Smoothing	4.349201	27.87527
Regression	6.940639	45.36848

Table 5. Forecast Accuracy for Extreme Weather

In forecasting Extreme Weather delays, Exponential Smoothing and Regression differ in accuracy. Exponential Smoothing has a MASE of 4.35 and a MAPE of 27.88%, while Regression shows higher values with a MASE of 6.94 and a MAPE of 45.37%. These figures suggest that Exponential Smoothing offers more precise forecasts with smaller errors and lower percentage deviations from actual values compared to Regression. Thus, Exponential Smoothing is likely more suitable for forecasting Extreme Weather delays due to its superior accuracy.

Year	Air Carrier Delay	Aircraft Arriving Late	National Aviation System Delay	Security Delay	Extreme Weather
2024	38.20	39.11	17.02	0.21	5.20
2025	38.82	39.11	16.25	0.21	5.20
2026	39.43	39.11	15.49	0.21	5.20
2027	40.05	39.11	14.73	0.21	5.20
2028	40.66	39.11	13.96	0.21	5.20
2029	41.28	39.11	13.20	0.21	5.20

Figure 4. Forecasted Data for each factor causing a delay from 2024 to 2029

To reduce airline delays, it's essential to focus on the types of delays that have the greatest impact and are most amenable to improvement. Looking at the forecasted data from 2024 to 2029, we can see that the Aircraft Arriving Late delay is projected to remain relatively stable around 39.11. In contrast, the Air Carrier Delay is forecasted to increase slightly from 38.20 in 2024 to 41.28 in 2029. This suggests that the Air Carrier Delay may become an increasingly significant contributor to overall delays and should be a key focus area for reducing delays in the airline industry. Addressing factors contributing to Air Carrier Delays could lead to more effective strategies for improving overall on-time performance.

The forecasted data from 2024 to 2029 suggests stable Aircraft Arriving Late delays but a slight increase in Air Carrier Delays. Effective cost management, particularly in crew expenses and maintenance, is crucial to mitigate delays' impact on operations. Aligning cost strategies with impactful delays can help airlines improve on-time performance and reduce financial burdens.

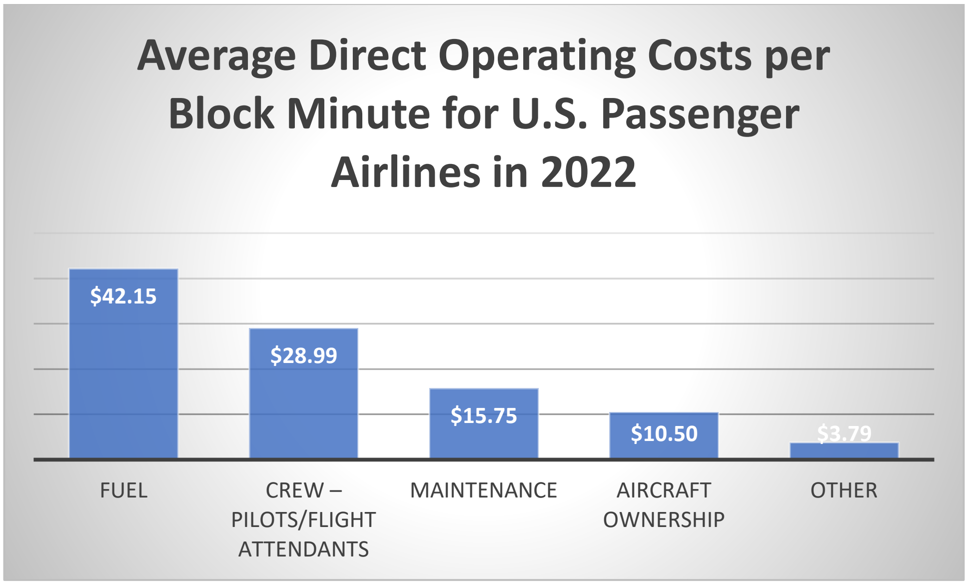


Figure 5. Average Direct Operating Costs per Block Minute for U.S. Passenger Airlines in 2022

In 2022, the direct aircraft operating costs per block minute saw various changes compared to the previous year. Figure 5 (above) shows the breakdown of these costs. The cost of fuel per block minute increased slightly to $42.15, with a 3.0% year-over-year (YOY) change. Crew costs, including pilots and flight attendants, significantly rose to $28.99 per block minute, showing an 87.3% YOY increase. Maintenance costs also increased moderately to $15.75 per block minute, with a 6.1% YOY change. However, aircraft ownership costs decreased to $10.50 per block minute, with a negative YOY change of -12.9%. Other direct operating costs per block minute increased to $3.79, showing a 26.8% YOY change. The total direct operating cost per block minute, encompassing all these factors, was $101.18, representing a 25.7% YOY increase.

Understanding the direct aircraft operating costs per block minute for 2022 can provide valuable insights for improving operational efficiency and reducing flight delays. By analyzing the cost breakdown, airlines can identify areas where cost-saving measures can be implemented. For example, the significant increase in crew costs may indicate the need for more efficient crew scheduling or training programs to optimize staffing levels. Similarly, the moderate increase in maintenance costs could prompt airlines to enhance aircraft maintenance practices to minimize downtime and delays caused by maintenance issues. Moreover, the decrease in aircraft ownership costs presents an opportunity for airlines to reassess their fleet management strategies and potentially invest in newer, more fuel-efficient aircraft to reduce operating costs in the long term. Overall, a detailed analysis of these cost factors can help airlines make informed decisions to improve operational efficiency and reduce the likelihood of flight delays.

VI. Understand the benefits of improvements; Pick one opportunity, create a small experiment, and create an implementation plan that management would sign-off on

Our analysis of the factor-weighted table and forecasts highlights late arrivals and air carriers as the area's most prone to delays. In this discussion, we will focus on addressing air carrier delays to improve the overall efficiency of the aircraft turnaround process.

The aircraft turnaround process is a critical operation that involves several key activities aimed at minimizing the turnaround time (TAT) and ensuring the timely departure of aircraft for their next flights. One of the primary tasks in this process is passenger and baggage handling, which includes the efficient offloading of baggage from arriving flights, processing baggage for departing flights, and managing passenger boarding and deboarding. Additionally, cargo and mail handling play a crucial role, as they involve the unloading and loading of cargo and mail onto the aircraft for both arriving and departing flights.

Another essential aspect of the turnaround process is load control, which focuses on determining the optimal distribution of weight on the aircraft for balance and safety. This includes managing fuel, baggage, cargo, and passenger loads to ensure safe and efficient flight operations. Additionally, ramp handling activities are conducted on the ground to prepare the aircraft for departure. These activities include refueling, catering, cleaning, and other maintenance tasks necessary for the aircraft's readiness for the next flight.

Efficient management of these activities is vital for airlines and airports, as it directly impacts aircraft utilization, operational costs, and overall profitability. Airlines strive to minimize TAT to increase the number of flights an aircraft can perform per day, thereby reducing the cost per available seat kilometre (CASK) and enhancing profitability. Similarly, airports benefit from shorter TATs by increasing the number of aircraft served and improving overall operational performance.

Efforts to reduce TAT often focus on optimizing ground operations, improving coordination between airline and airport staff, and leveraging technology to streamline processes. Advanced planning, efficient resource utilization, and effective communication are key elements in achieving a smooth and efficient aircraft turnaround process.

To improve the efficiency of the aircraft turnaround process, focusing on addressing air carrier delays is crucial, as they are identified as a significant area prone to delays. By streamlining air carrier operations, airlines can reduce the turnaround time (TAT) and enhance overall operational efficiency. One potential opportunity for improvement is to implement a more efficient communication and coordination system between ground staff and air carrier personnel. This can help reduce delays caused by miscommunication or inefficiencies in the handover process between these two groups. For the small experiment, the airline can pilot a new communication protocol or technology, such as a digital communication platform, to facilitate real-time updates and instructions. The implementation plan would involve training staff on the new system, testing it in a controlled environment, and gradually integrating it into regular operations.

The value of this improvement can be measured by tracking the reduction in air carrier delays and overall TAT. By reducing delays, the airline can increase aircraft utilization, leading to potential cost savings and increased profitability. In just one day with little cash, the airline can learn valuable insights about the effectiveness of the new communication system. By collecting feedback from ground staff and air carrier personnel, as well as monitoring key performance indicators such as TAT and delay times, the airline can quickly assess the impact of the improvement and make necessary adjustments. Overall, focusing on improving air carrier operations through better communication and coordination can lead to significant benefits for airlines, including reduced TAT, increased aircraft utilization, and improved operational efficiency.

VII. Conclusion

In conclusion, our extensive study on airline delays in the aviation industry has provided valuable insights into the challenges and opportunities for improvement in air travel operations. Through rigorous analysis of industry trends, operational data, and forecasting future trends, we have identified key areas for enhancing operational efficiency and reducing flight delays. The aviation industry confronts various challenges, including workforce shortages, safety and security concerns, sustainability issues, and complex regulatory environments, necessitating strategic planning, innovation, and collaboration among industry stakeholders. One of the primary objectives of our study was to minimize the incidence of flight delays across the board by employing targeted strategies and operational enhancements, such as implementing a more efficient communication and coordination system between ground staff and air carrier personnel, which can reduce turnaround times and enhance overall operational efficiency.

Additionally, our analysis encompassed the development of a factor weighted table and a Pareto chart, instrumental in identifying and prioritizing the most significant factors contributing to delays, enabling airlines to focus their efforts and resources on areas with the greatest impact. Our study not only sheds light on the current challenges faced by the aviation industry but also highlights the potential for significant improvements in operational efficiency and customer satisfaction. By addressing workforce shortages, enhancing safety and security measures, and navigating complex regulatory landscapes, airlines can create a more sustainable and competitive industry. Implementing targeted strategies, informed by data analysis and forecasting, will be crucial in achieving these goals. Through continuous improvement and collaboration, the aviation industry can provide a more reliable and enjoyable travel experience for passengers worldwide.

Citations

(14) Flight Delays: What can be improved? | LinkedIn. (n.d.). Retrieved April 30, 2024, from https://www.linkedin.com/pulse/flight-delays-what-can-improved-vinod-k-vijayan/

Airline On-Time Performance and Causes of Flight Delays | Bureau of Transportation Statistics. (n.d.). Retrieved April 30, 2024, from https://www.bts.dot.gov/explore-topics-and-geography/topics/airline-time-performance-and-causes-flight-delays

Barten, M. (2024, January 5). *Aviation Industry: All You Need to Know About the Aviation Sector*. Revfine.Com. https://www.revfine.com/aviation-industry/

Delay Cause by Year, as a Percent of Total Delay Minutes | Bureau of Transportation Statistics. (n.d.). Retrieved April 30, 2024, from https://www.bts.gov/content/delay-cause-year-percent-total-delay-minutes

OST_R | BTS | Title from h2. (n.d.). Retrieved April 30, 2024, from https://www.transtats.bts.gov/OT_Delay/OT_DelayCause1.asp?20=E

Picchi Scardaoni, M., Magnacca, F., Massai, A., & Cipolla, V. (2021). Aircraft turnaround time estimation in early design phases: Simulation tools development and application to the case of box-wing architecture. *Journal of Air Transport Management*, *96*, 102122. https://doi.org/10.1016/j.jairtraman.2021.102122

U.S. Passenger Carrier Delay Costs. (n.d.). Airlines For America. Retrieved April 30, 2024, from https://www.airlines.org/dataset/u-s-passenger-carrier-delay-costs/

Delay Forecast

Alex Pettis

05-01-2024

```r
library(fpp3)
library(readxl)
library(readr)
library(knitr)
library(DT)
```

Original Data

```r
my_data <- read_xlsx('/Users/alex/Desktop/Spring 2024/Operations and Supply Chain Management/Delay Data

kable(my_data)
```

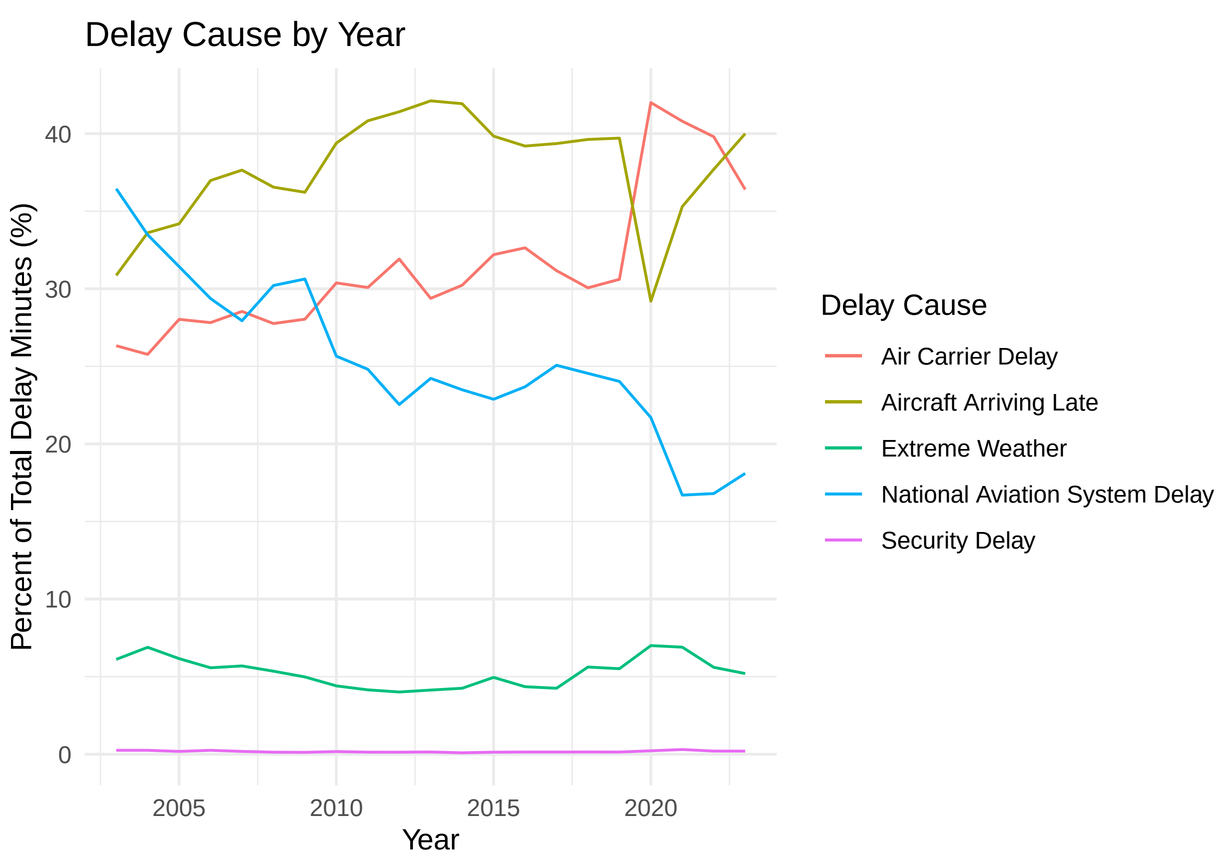
Year	Air Carrier Delay	Aircraft Arriving Late	National Aviation System Delay	Security Delay	Extreme Weather
2003	26.33000	30.86000	36.45000	0.2500000	6.110000
2004	25.77000	33.61000	33.48000	0.2500000	6.890000
2005	28.03000	34.19000	31.43000	0.1800000	6.160000
2006	27.82000	36.98000	29.37000	0.2500000	5.570000
2007	28.54000	37.65000	27.94000	0.1800000	5.690000
2008	27.76000	36.55000	30.21000	0.1300000	5.350000
2009	28.04000	36.22000	30.63000	0.1200000	4.980000
2010	30.37966	39.39165	25.65594	0.1697973	4.402955
2011	30.08181	40.82591	24.81228	0.1319039	4.148089
2012	31.92000	41.41000	22.54000	0.1300000	4.010000
2013	29.38227	42.11432	24.22497	0.1440846	4.134351
2014	30.23000	41.93000	23.49000	0.0900000	4.250000
2015	32.20000	39.84000	22.88000	0.1300000	4.950000
2016	32.64000	39.20000	23.68000	0.1400000	4.350000
2017	31.17000	39.36000	25.07000	0.1400000	4.250000
2018	30.06278	39.62649	24.54697	0.1445406	5.619226
2019	30.61000	39.71000	24.03000	0.1400000	5.510000
2020	42.00000	29.20000	21.70000	0.2195141	7.000000
2021	40.80000	35.30000	16.70000	0.3000000	6.900000
2022	39.80000	37.70000	16.80000	0.2000000	5.600000
2023	36.40000	40.00000	18.10000	0.2000000	5.200000

Graph of the Original Data

```r
# Convert the "Year" column to Date format
my_data <- as.data.frame(my_data)
my_data$Year <- as.Date(paste0(my_data$Year, "-01-01"), format = "%Y-%m-%d")

view(my_data)
# Extract the year from the "Year" column and set up data as a time series object
my_data <- my_data |>
  mutate(Year = year(Year)) |>
  as_tsibble(index = Year)


 ggplot(my_data, aes(x = Year)) +
  geom_line(aes(y = `Air Carrier Delay`, color = "Air Carrier Delay")) +
  geom_line(aes(y = `Aircraft Arriving Late`, color = "Aircraft Arriving Late")) +
  geom_line(aes(y = `National Aviation System Delay`,
                color = "National Aviation System Delay")) +
  geom_line(aes(y = `Security Delay`, color = "Security Delay")) +
  geom_line(aes(y = `Extreme Weather`, color = "Extreme Weather")) +
  labs(title = "Delay Cause by Year",
       x = "Year",
       y = "Percent of Total Delay Minutes (%)",
       color = "Delay Cause") +
  theme_minimal()
```



```r
 # extract the training set to fit the data
training_set <- my_data |>
  filter_index("2003" ~ "2017")

# extract the test set
test_set <- my_data |>
  filter_index("2018" ~ "2023")

# set the forecast horizon equals to the test set
h <- nrow(test_set)

my_fit <- training_set |>
  model(
    Drift = NAIVE(my_data ~ drift()))

my_fit <- training_set |>
  model(
    Regression = TSLM(`Air Carrier Delay` ~ Year),
    Exponential_Smoothing = ETS(`Air Carrier Delay` ~ error("A") +
                                trend("N") + season("N")))

my_fit2 <- training_set |>
  model(
    Regression = TSLM(`Aircraft Arriving Late` ~ Year),
    Exponential_Smoothing = ETS(`Aircraft Arriving Late` ~ error("A") +
                                trend("N") + season("N")))

my_fit3 <- training_set |>
  model(
    Regression = TSLM(`National Aviation System Delay` ~ Year),
    Exponential_Smoothing = ETS(`National Aviation System Delay` ~ error("A") +
                                trend("N") + season("N")))

my_fit4 <- training_set |>
  model(
    Regression = TSLM(`Security Delay` ~ Year),
    Exponential_Smoothing = ETS(`Security Delay` ~ error("A") + trend("N") +
                                season("N")))

my_fit5 <- training_set |>
  model(
    Regression = TSLM(`Extreme Weather` ~ Year),
    Exponential_Smoothing = ETS(`Extreme Weather` ~ error("A") + trend("N") +
                                season("N")))
```

Forecast Accuracy for Air Carrier Delays

```r
my_fc <- my_fit |>
  forecast(h = h)


kable(accuracy(my_fc, my_data) |>
```

```
         select(Model = .model, MASE, MAPE)|>
        arrange(MASE, MAPE))
```

Model	MASE	MAPE
Regression	3.855774	11.94303
Exponential_Smoothing	4.967502	14.93985

```
accuracy_output <- accuracy(my_fc, my_data) |>
  select(.model, MASE, MAPE) |>
  arrange(MASE, MAPE)
```

Forecast Accuracy for Aircraft Arriving Late

```
my_fc2 <- my_fit2 |>
  forecast(h = h)


kable(accuracy(my_fc2, my_data) |>
         select(Model = .model, MASE, MAPE)|>
        arrange(MASE, MAPE))
```

Model	MASE	MAPE
Exponential_Smoothing	2.326252	8.975497
Regression	6.040573	21.561436

```
accuracy_output <- accuracy(my_fc2, my_data) |>
  select(.model, MASE, MAPE) |>
  arrange(MASE, MAPE)
```

Forecast Accuracy for National Aviation System Delays

```
my_fc3 <- my_fit3 |>
  forecast(h = h)


kable(accuracy(my_fc3, my_data) |>
         select(Model = .model, MASE, MAPE)|>
        arrange(MASE, MAPE))
```

Model	MASE	MAPE
Regression	1.366412	11.06366
Exponential_Smoothing	2.717201	26.63977

```
accuracy_output <- accuracy(my_fc3, my_data) |>
  select(.model, MASE, MAPE) |>
  arrange(MASE, MAPE)
```

Forecast Accuracy for Security Delays

```
my_fc4 <- my_fit4 |>
  forecast(h = h)

kable(accuracy(my_fc4, my_data) |>
         select(Model = .model, MASE, MAPE)|>
        arrange(MASE, MAPE))
```

Model	MASE	MAPE
Exponential_Smoothing	1.873601	27.18698
Regression	3.776717	60.53097

```
accuracy_output <- accuracy(my_fc4, my_data) |>
  select(.model, MASE, MAPE) |>
  arrange(MASE, MAPE)
```

Forecast Accuracy for Extreme Weather

```
my_fc5 <- my_fit5 |>
  forecast(h = h)


kable(accuracy(my_fc5, my_data) |>
         select(Model = .model, MASE, MAPE)|>
        arrange(MASE, MAPE))
```

Model	MASE	MAPE
Exponential_Smoothing	4.349201	27.87527
Regression	6.940639	45.36848

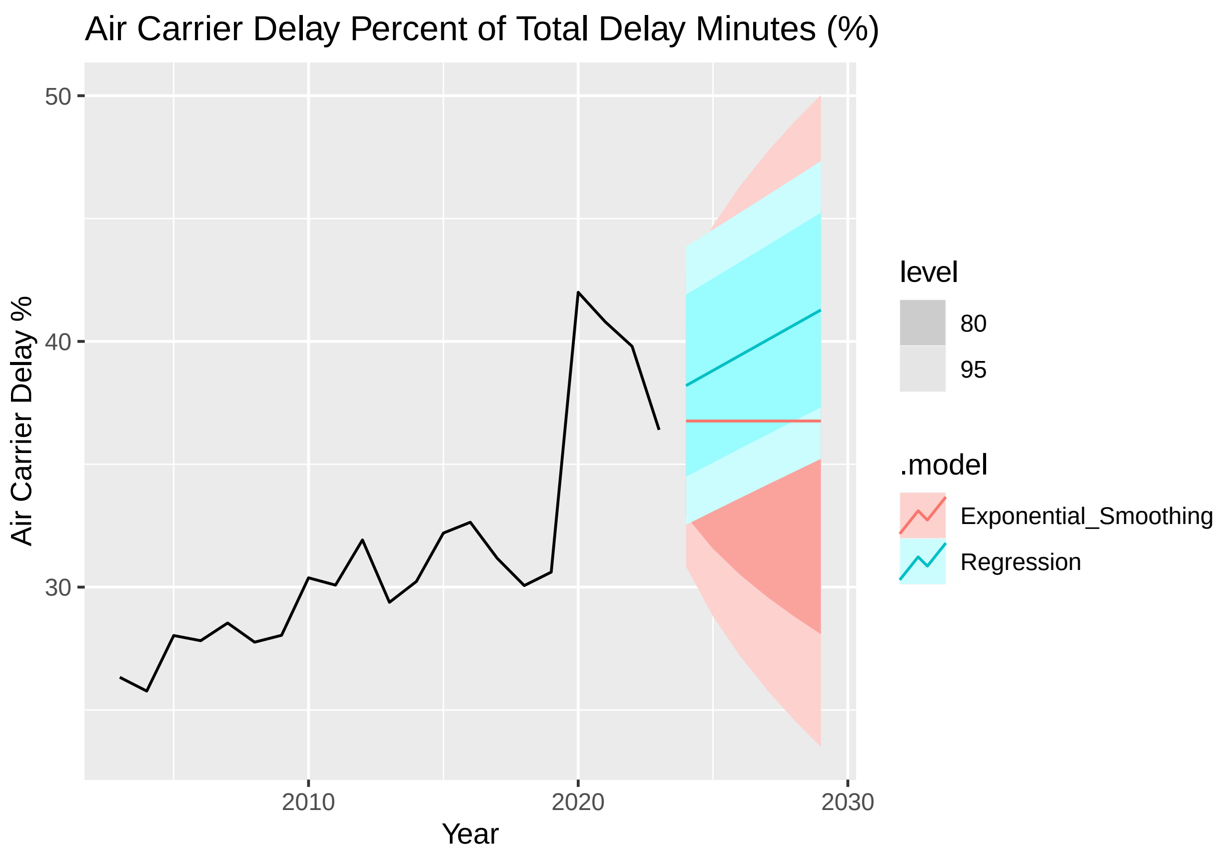
```
accuracy_output <- accuracy(my_fc5, my_data) |>
  select(.model, MASE, MAPE) |>
  arrange(MASE, MAPE)
```

Forecasted Graph

This graphs shows the forecast for years 2024-2029 for both Regression as well as Exponential Smoothing.

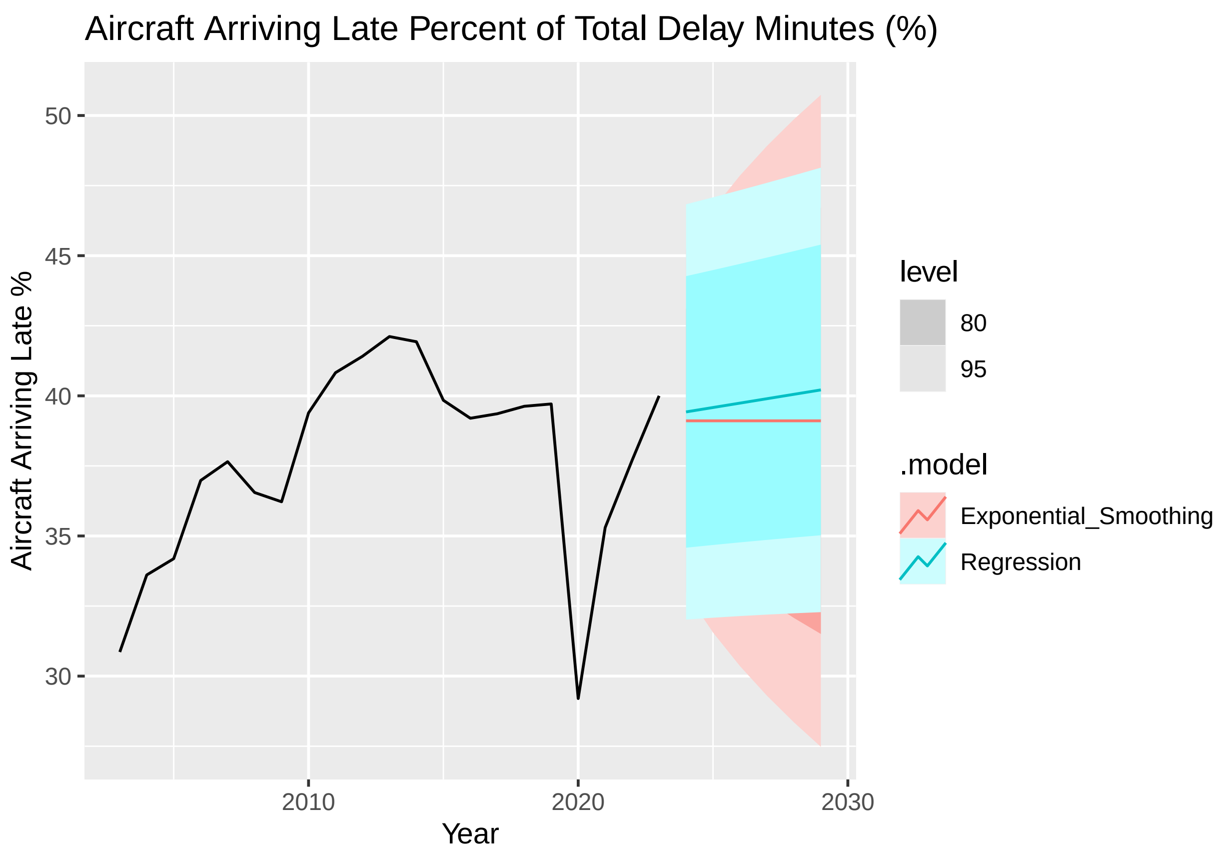
```r
my_fit <- my_data |> # notice you are using the complete data here, not the training set
  model(
    Regression = TSLM(`Air Carrier Delay` ~ Year),
    Exponential_Smoothing = ETS(`Air Carrier Delay` ~ error("A") + trend("N") +
                                  season("N")))

my_fit |>
  forecast(h = h) |>
  autoplot(my_data) +
  labs(title = "Air Carrier Delay Percent of Total Delay Minutes (%)",
       x = "Year",
       y = "Air Carrier Delay %")
```



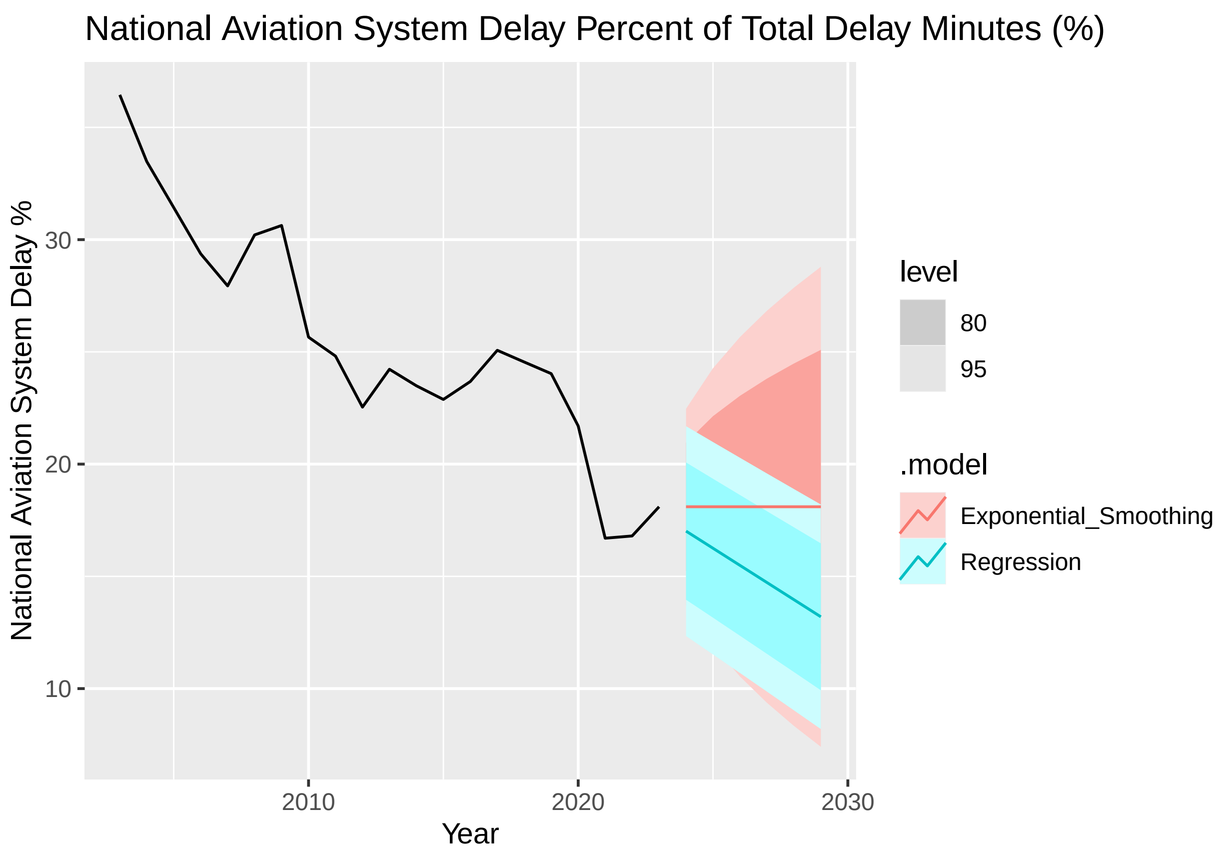

```r
my_fit2 <- my_data |> # notice you are using the complete data here, not the training set
  model(
    Regression = TSLM(`Aircraft Arriving Late` ~ Year),
    Exponential_Smoothing = ETS(`Aircraft Arriving Late` ~ error("A") + trend("N") +
                                  season("N")))

my_fit2 |>
  forecast(h = h) |>
  autoplot(my_data) +
  labs(title = "Aircraft Arriving Late Percent of Total Delay Minutes (%)", x = "Year",
       y = "Aircraft Arriving Late %")
```



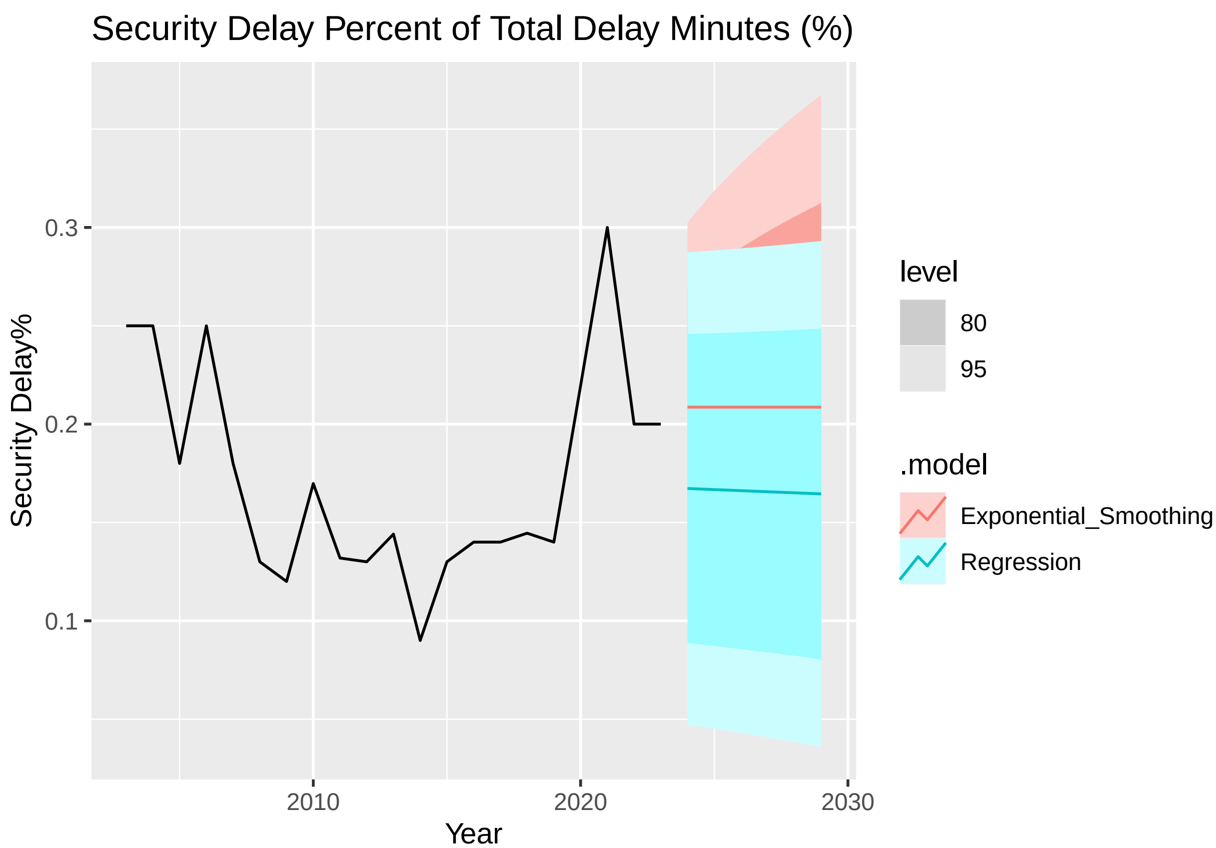

```
my_fit3 <- my_data |> # notice you are using the complete data here, not the training set
  model(
    Regression = TSLM(`National Aviation System Delay` ~ Year),
    Exponential_Smoothing = ETS(`National Aviation System Delay` ~ error("A") +
                                  trend("N") + season("N")))

my_fit3 |>
  forecast(h = h) |>
  autoplot(my_data) +
  labs(title = "National Aviation System Delay Percent of Total Delay Minutes (%)",
       x = "Year",
       y = "National Aviation System Delay %")
```



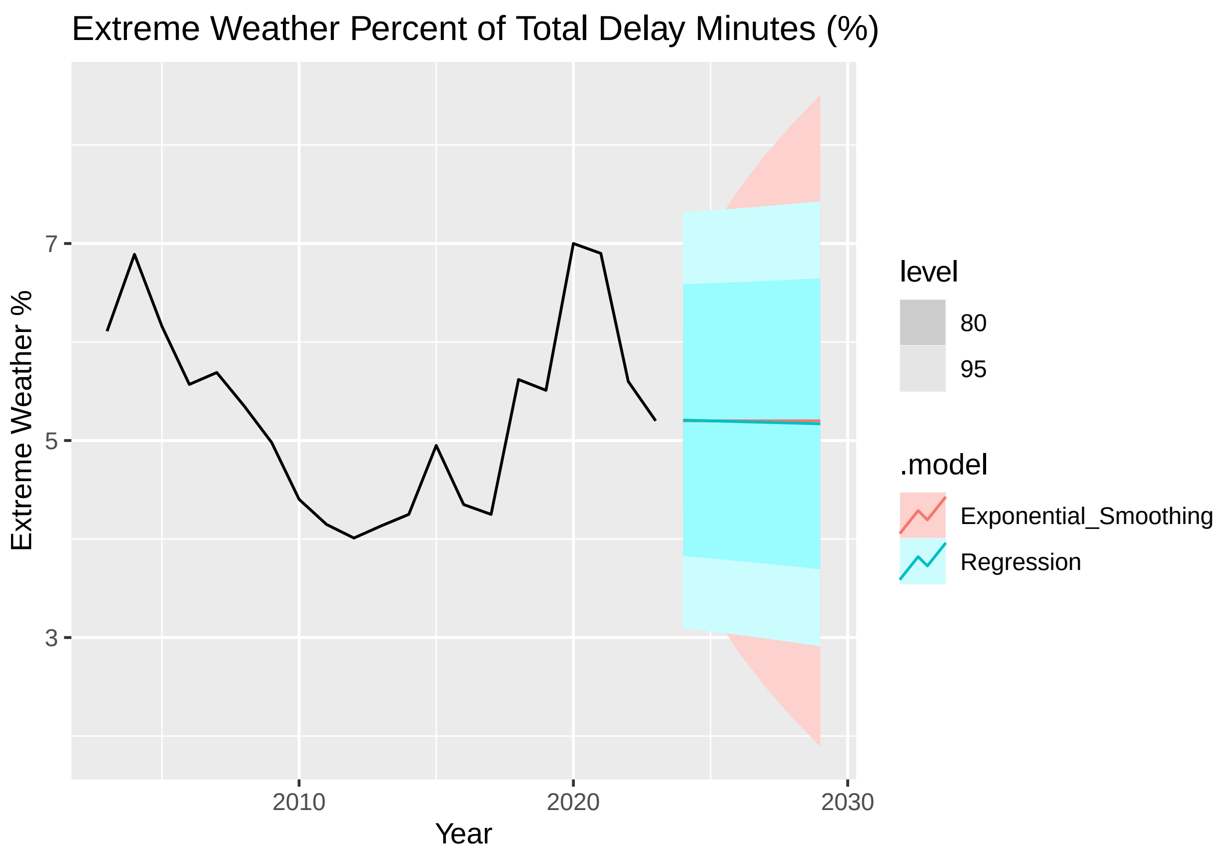

```r
my_fit4 <- my_data |> # notice you are using the complete data here, not the training set
  model(
    Regression = TSLM(`Security Delay` ~ Year),
    Exponential_Smoothing = ETS(`Security Delay` ~ error("A") + trend("N") +
                                  season("N")))

my_fit4 |>
  forecast(h = h) |>
  autoplot(my_data) +
  labs(title = "Security Delay Percent of Total Delay Minutes (%)", x = "Year",
       y = "Security Delay%")
```




```
my_fit5 <- my_data |> # notice you are using the complete data here, not the training set
  model(
    Regression = TSLM(`Extreme Weather` ~ Year),
    Exponential_Smoothing = ETS(`Extreme Weather` ~ error("A") + trend("N") +
                                  season("N")))

my_fit5 |>
  forecast(h = h) |>
  autoplot(my_data) +
  labs(title = "Extreme Weather Percent of Total Delay Minutes (%)", x = "Year",
       y = "Extreme Weather %")
```



Final Output

```
# Filter rows for Regression model
regression_data <- extracted_data |>
  filter(.model == "Regression") |>
  select(-.model)

# Filter rows for Exponential Smoothing model
exp_smoothing_data2 <- extracted_data2 |>
  filter(.model == "Exponential_Smoothing") |>
  select(-.model)

# Filter rows for Regression model
regression_data3 <- extracted_data3 |>
  filter(.model == "Regression") |>
  select(-.model)

# Filter rows for Exponential Smoothing model
exp_smoothing_data4 <- extracted_data4 |>
  filter(.model == "Exponential_Smoothing") |>
  select(-.model)

# Filter rows for Exponential Smoothing model
exp_smoothing_data5 <- extracted_data5 |>
```

```r
  filter(.model == "Exponential_Smoothing") |>
  select(-.model)

combine1 <- bind_rows(my_data, regression_data)

combine1 <- combine1 |>
  mutate(`Air Carrier Delay` = ifelse(is.na(`Air Carrier Delay`),
                                      .mean, `Air Carrier Delay`)) |>
  select(-.mean)


combine2 <- bind_rows(my_data, exp_smoothing_data2)

combine2 <- combine2 |>
  mutate(`Aircraft Arriving Late` = ifelse(is.na(`Aircraft Arriving Late`),
                                           .mean, `Aircraft Arriving Late`)) |>
  select(-.mean)

combine3 <- bind_rows(my_data, regression_data3)

combine3 <- combine3 |>
  mutate(`National Aviation System Delay` =
           ifelse(is.na(`National Aviation System Delay`),
                                                         .mean,
                                                         `National Aviation System Delay`)) |>
  select(-.mean)

combine4 <- bind_rows(my_data, exp_smoothing_data4)

combine4 <- combine4 |>
  mutate(`Security Delay` = ifelse(is.na(`Security Delay`),
                                   .mean, `Security Delay`)) |>
  select(-.mean)

combine5 <- bind_rows(my_data, exp_smoothing_data5)

combine5 <- combine5 |>
  mutate(`Extreme Weather` = ifelse(is.na(`Extreme Weather`),
                                    .mean, `Extreme Weather`)) |>
  select(-.mean)

forecasted_data <- bind_cols(combine1$Year, combine1$`Air Carrier Delay`,
                             combine2$`Aircraft Arriving Late`,
                             combine3$`National Aviation System Delay`,
                             combine4$`Security Delay`,
                             combine5$`Extreme Weather`)

forecasted_data <- forecasted_data |>
  rename(Year = ...1, `Air Carrier Delay` = ...2,
         `Aircraft Arriving Late` = ...3,
         `National Aviation System Delay` = ...4, `Security Delay` = ...5,
         `Extreme Weather` = ...6)
```

```r
kable(forecasted_data)
```

Year	Air Carrier Delay	Aircraft Arriving Late	National Aviation System Delay	Security Delay	Extreme Weather
2003	26.33000	30.86000	36.45000	0.2500000	6.110000
2004	25.77000	33.61000	33.48000	0.2500000	6.890000
2005	28.03000	34.19000	31.43000	0.1800000	6.160000
2006	27.82000	36.98000	29.37000	0.2500000	5.570000
2007	28.54000	37.65000	27.94000	0.1800000	5.690000
2008	27.76000	36.55000	30.21000	0.1300000	5.350000
2009	28.04000	36.22000	30.63000	0.1200000	4.980000
2010	30.37966	39.39165	25.65594	0.1697973	4.402955
2011	30.08181	40.82591	24.81228	0.1319039	4.148089
2012	31.92000	41.41000	22.54000	0.1300000	4.010000
2013	29.38227	42.11432	24.22497	0.1440846	4.134351
2014	30.23000	41.93000	23.49000	0.0900000	4.250000
2015	32.20000	39.84000	22.88000	0.1300000	4.950000
2016	32.64000	39.20000	23.68000	0.1400000	4.350000
2017	31.17000	39.36000	25.07000	0.1400000	4.250000
2018	30.06278	39.62649	24.54697	0.1445406	5.619226
2019	30.61000	39.71000	24.03000	0.1400000	5.510000
2020	42.00000	29.20000	21.70000	0.2195141	7.000000
2021	40.80000	35.30000	16.70000	0.3000000	6.900000
2022	39.80000	37.70000	16.80000	0.2000000	5.600000
2023	36.40000	40.00000	18.10000	0.2000000	5.200000
2024	38.20085	39.10824	17.01609	0.2086103	5.200040
2025	38.81666	39.10824	16.25244	0.2086103	5.200040
2026	39.43247	39.10824	15.48880	0.2086103	5.200040
2027	40.04827	39.10824	14.72515	0.2086103	5.200040
2028	40.66408	39.10824	13.96151	0.2086103	5.200040
2029	41.27989	39.10824	13.19786	0.2086103	5.200040