# Comparing, Contrasting, and Predicting Commercial Flight Characteristics, Pre-, Mid-, and Post-COVID-19 Pandemic

Zhexuan Tang, Ruixin Gan, Zheyan Chen, Yin Wang, Weihao Xia

MLC
Spring 2024

NYU TANDON SCHOOL OF ENGINEERING

## Introduction

Before the COVID-19 pandemic, the commercial aviation industry had reached new heights.
2016-2019 marked robust expansion in air travel (Mutzabaugh, 2019 & Morris, 2018).
Airlines expanded fleets and added routes to meet growing demand.
Post-COVID, the industry evolved to meet new realities.
Strategic realignments by airlines aimed to optimize operations and adapt to changed market.

## Research Question

1. How did ridership trends shift pre-, during, and post-pandemic?
2. What changes occurred in aircraft types used, and how do these reflect airline strategies and passenger preferences?
3. What are the projections for post-pandemic ridership recovery, and how do they compare to pre-pandemic growth projections?

## Hypothesis

1. Post-pandemic, narrowbody aircraft have become more prevalent in flight clusters than widebody/mid-sized aircraft.
2. Based on Gaussian Prediction (GP) Post-Pandemic ridership recovery can be predicted to continue to rise, but would most likely struggle to reach a pre-pandemic growth levels had the pandemic never happened.

## Methodology

## Time Series -Passanger Amount

## Analysis

### GP Prediction-All Flights

```
Log Marginal Likelihood (optimized): -3749.474479552617
Optimized Kernel:
 316**2 * Matern(length_scale=7.27e-12, nu=1.5) + DotProduct(sigma_0=7.32e+04) + 316**2 * RationalQuadratic(alpha=7.24, length_scale=5.79e-15) + 316**2 *
RBF(length_scale=0.625) + WhiteKernel(noise_level=1e+05) + 316**2 * ExpSineSquared(length_scale=0.01, periodicity=31) + 316**2 * ExpSineSquared(length_sc
ale=0.0667, periodicity=365)
```

```
Log Marginal Likelihood (optimized): -6823.333660898841
Optimized Kernel:
 316**2 * Matern(length_scale=0.0848, nu=1.5) + DotProduct(sigma_0=1.15e+04) + 316**2 * RationalQuadratic(alpha=1e+05, length_scale=1.17) + 316**2 * RBF
(length_scale=35.6) + WhiteKernel(noise_level=1e+05) + 316**2 * ExpSineSquared(length_scale=0.01, periodicity=30.1) + 316**2 * ExpSineSquared(length_scal
e=0.0103, periodicity=363)
```

### GP Prediction-Mid Body

```
Log Marginal Likelihood (optimized): -2285.920166757763
Optimized Kernel:
 5.29**2 * Matern(length_scale=20, nu=1.5) + DotProduct(sigma_0=1.52e+03) + 113**2 * RationalQuadratic(alpha=1.19, length_scale=20) + 7.61**2 * RBF(lengt
h_scale=20) + WhiteKernel(noise_level=9.09e+03) + 151**2 * ExpSineSquared(length_scale=0.108, periodicity=7) + 118**2 * ExpSineSquared(length_scale=0.043
9, periodicity=360)
```




```
Log Marginal Likelihood (optimized): -4798.536994459281
Optimized Kernel:
 130**2 * Matern(length_scale=1, nu=1.5) + DotProduct(sigma_0=1.29e+03) + 140**2 * RBF(length_scale=1) + WhiteKernel(noise_level=5.25e+03) + 160**2 * Exp
SineSquared(length_scale=1.49, periodicity=7) + 148**2 * ExpSineSquared(length_scale=360, periodicity=30.2) + 125**2 * ExpSineSquared(length_scale=0.824,
periodicity=363)
```

### GP Prediction-Wide Body

```
Log Marginal Likelihood (optimized): -2171.689323737707
Optimized Kernel:
 10.1**2 * Matern(length_scale=20, nu=1.5) + DotProduct(sigma_0=1.29e+03) + 103**2 * RationalQuadratic(alpha=1.65, length_scale=7.1) + 17.6**2 * RBF(leng
th_scale=20) + WhiteKernel(noise_level=3.81e+03) + 80.8**2 * ExpSineSquared(length_scale=0.0177, periodicity=361)
```

```
Log Marginal Likelihood (optimized): -4322.678825251027
Optimized Kernel:
 17.4**2 * Matern(length_scale=1, nu=1.5) + DotProduct(sigma_0=1.66e+03) + 38**2 * RBF(length_scale=1) + WhiteKernel(noise_level=8.51e+03) + 60.4**2 * Ex
pSineSquared(length_scale=1e+03, periodicity=30) + 70.5**2 * ExpSineSquared(length_scale=0.209, periodicity=360)
```

### Cluster-Before

| cluster | type | flight_count |
|---|---|---|
| 0 | B77 | 331943 |
| 1 | A320 | 1374012 |
| 2 | B73 | 2440608 |

### Cluster-After

| cluster | type | flight_count |
|---|---|---|
| 0 | A320 | 1784200 |
| 1 | B73 | 4194800 |
| 2 | CRJ9 | 526490 |

## References


Bauranov, Aleksandar, et al. "Quantifying the Resilience of the U.S. Domestic Aviation Network During the COVID-19 Pandemic." *Frontiers in Built Environment*, vol. 7, 12 Apr. 2021, doi:10.3389/fbuil.2021.642295.

Barczak, Agnieszka, et al. "The Impact of COVID-19 Pandemic on Air Transport Passenger Markets: Implications for Selected EU Airports Based on Time Series Models Analysis." *Sustainability*, vol. 14, no. 4345, 2022, doi:10.3390/su14074345.

Ehsani, Sina, et al. "Predicting the Skies: A Novel Model for Flight-Level Passenger Traffic Forecasting." *arXiv preprint arXiv*: 2401.03397v2, 9 Jan. 2024, arxiv.org/abs/2401.03397v2.

Firat, Murat, Yiltas-Kaplan, Derya, and Samli, Ruya. "Forecasting Air Travel Demand for Selected Destinations Using Machine Learning Methods." *Journal of Universal Computer Science*, vol. 27, no. 6, 28 Jun. 2021, doi:10.3897/jucs.68185.

Gao, Yi. "Benchmarking the Recovery of Air Travel Demands for US Airports During the COVID-19 Pandemic." *Transportation Research Interdisciplinary Perspectives*, vol. 13, 2022, doi:10.1016/j.trip.2021.100570.

IP2Location. "ip2location-iata-icao." *GitHub*, updated 28 Apr. 2024, github.com/ip2location/ip2location-iata-icao.
Release note: Add Milan-Linate Airport.

Tolcha, Tassew Dufera. "The State of Africa's Air Transport Market Amid COVID-19, and Forecasts for Recovery." Journal of Air Transport Management, vol. 108, 2023, doi:10.1016/j.jairtraman.2023.102380.


## Limitations

**Airline Load Factor Accuracy:**
- Lack of accruate open data on flight passenger load (guarded commercial secret) to gauge ridership count
- Predetermined arbitrary airline load pre-, mid-, and post-pandemic at 90%, 65%, and 90% respectively

**GPR Model Precision and Data Availability:**
- the Gaussian Prediction Regression (GPR) model's precision is limited due to the lack of available, free opendata between 2017-2018.
- constrained to make predictions based on single 2019 year for pre-pandemic training data and negatively impacts accuracy

## Future Research

**Extending the Scope of Data**
- ICAO conforming Flight data from 2016-2018 will help in developing a more precise model and will enable more accurate predictions
- Using 2024 data to further extrapolar recovery trends moving forward

**Extending the Granularity of Data**
- segmenting by region according to the OD pairs to analyze the specific impact of the pandemic on airline load factors
- integrate a severity index to visualize differentiated impacts

## Conclusions

- **Ridership recovery prediction based on 2021-2022 data shows a slow pace to 75% of projected 2018-2019 data growth**
- **Projected Wide body ridership (B777s, A340s, A380s) are recovering at a slightly faster rate than Mid-body ridership (A330s, B767s, B787s)**
- **Narrowbody jets comprised of highest flight count amongst all three post-Pandmic, compared to pre-pandemic widebody prevalence**

## Data Sources

ICAO
OPENSKY NETWORK
GitHub