

# Data Challenge Answers

1a)

Most Trafficked Route: Sydney - Auckland

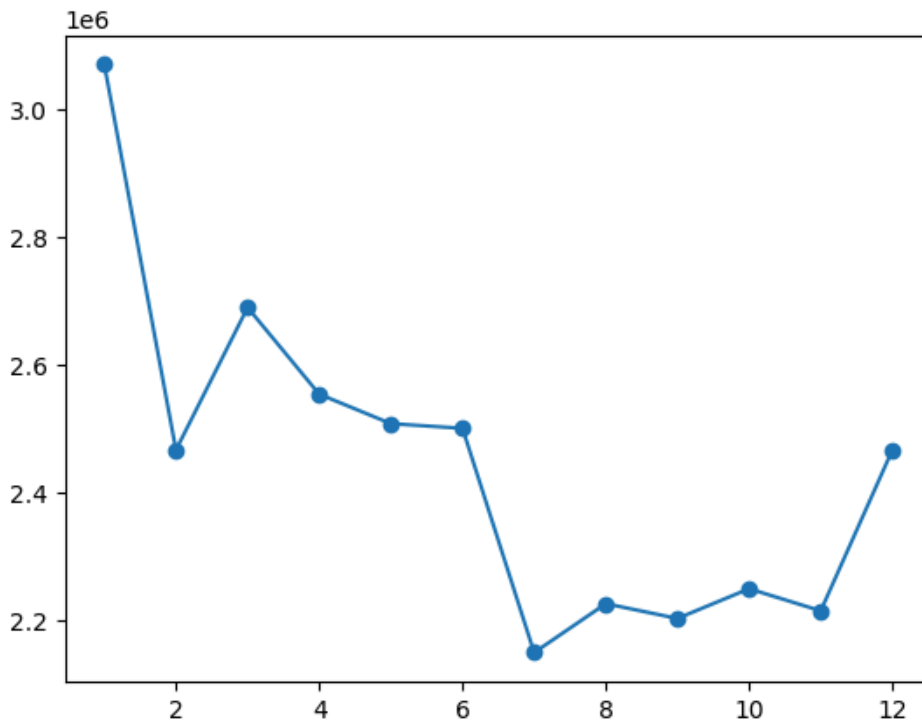
Least Trafficked Routes: Melbourne-Denver, Brisbane-Colombo, Brisbane-Chicago, Perth-Bandar Seri Begawan, Adelaide-Harare. They all have 0 passengers

1b)

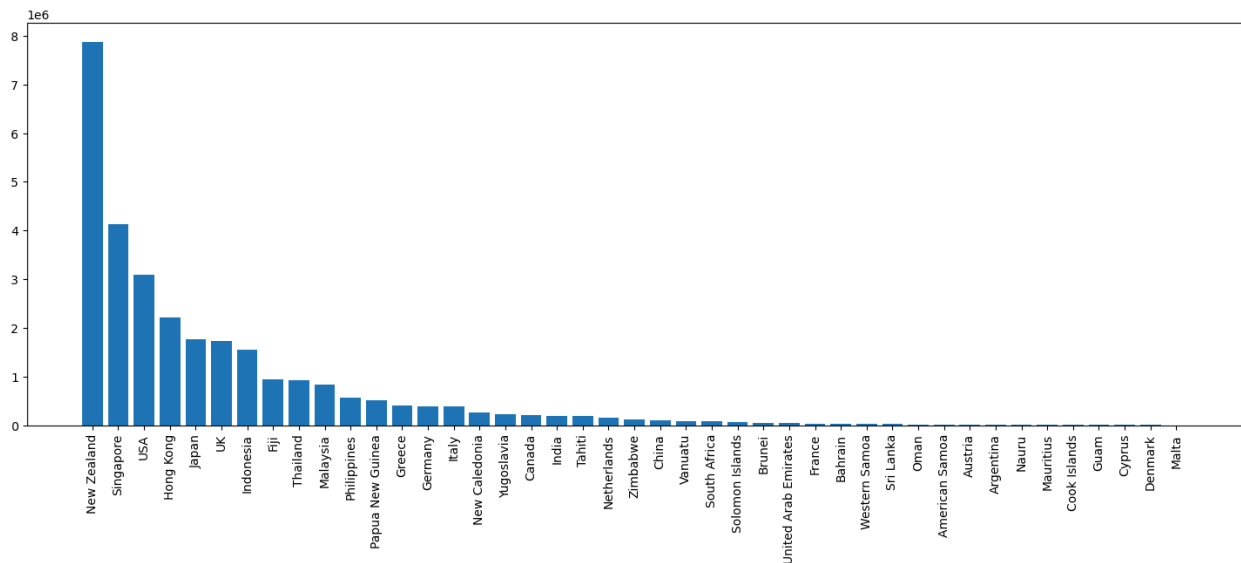
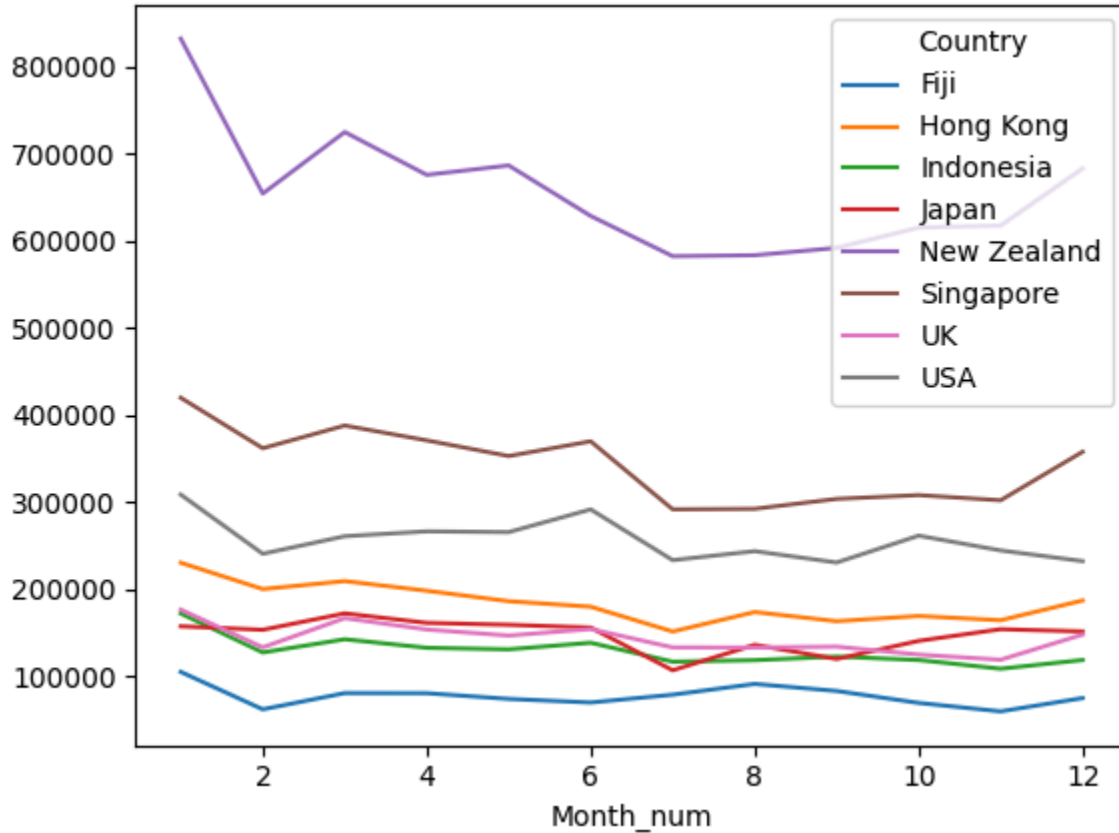
There is a clear overall growth across the years. The total number is going up every year, which makes sense with international travelling becoming more popular.

There is also a strong seasonal pattern. The traffic is usually highest in December and January likely due to the holiday season (Christmas and New Year). After that, there's a sharp drop in February, followed by more stable levels during spring.

Traffic then rises again in the summer, slightly decreases in the Fall, and steadily goes up again by December where it reaches its peak. Geographically, the busiest routes tend to be mostly countries around Australia (i.e APAC).



Passengers by month



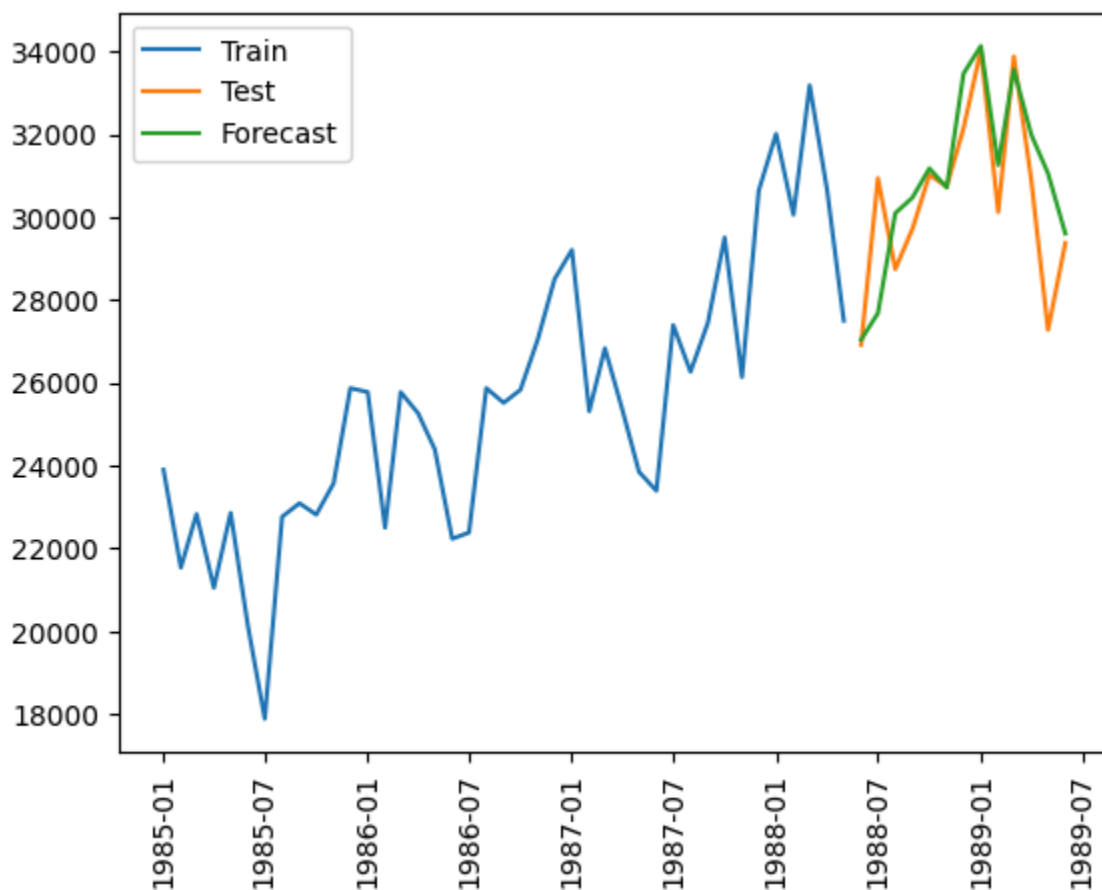
3a)

I tried several different approaches, including linear regression and other regressors, but Exponential Smoothing gave the best results when I compared

them visually using plots. The dataset clearly shows a seasonal pattern (big spike in December/January, sharp drop in February, and smaller fluctuations through the year), so I felt ETS was the most appropriate choice.

3b)

For evaluation, I created plots that compare the model's predictions with the actual data. Visually, the ETS model captured both the repeated seasonal patterns and the overall upward growth trend better than the other models I tested. In addition, I used quantitative metrics like MAPE and RMSE to measure accuracy. The results showed relatively low error values, which confirmed that ETS performed reliably.



Test RMSE: 1563 - about 5% error

Test  $R^2$ : 0.429 - not the best result, but i believe that the problem is in the data. 1989 exhibits a sudden drop in passenger counts, deviating largely from previous upward trend across all other years. I'm not sure what the reason is, but it's likely because of an external event? (the model cannot account for it)

4a)

Based on the model results and historical patterns, AeroConnect should consider investing more in the Sydney–Singapore route. Unlike other popular routes such as Sydney–Auckland or Sydney–Tokyo, which experienced a sharp and unusual decline in 1989, the Sydney–Singapore route showed greater stability over time. Although there was an overall dip in traffic during that year, this route consistently performed better compared to others, making it a more reliable investment.

4b)

This model can be used to estimate passenger traffic on different routes, which helps AeroConnect make better decisions about resource allocation (aircraft, crew, fuel, and scheduling). By forecasting demand, the airline can plan fleet usage more efficiently.