1. Understanding Data

a. Most and Least Trafficked Routes

Based on total passenger numbers, here are the top 10 most and least trafficked routes:

Top 10 Most Trafficked Routes

1. Sydney - Auckland | 2,961,212
2. Sydney - Singapore | 1,440,018
3. Sydney - Tokyo | 1,292,116
4. Sydney - Hong Kong | 1,151,900
5. Perth - Singapore | 952,926
6. Brisbane - Auckland | 893,246
7. Sydney - Christchurch | 882,357
8. Melbourne - Singapore | 865,251
9. Sydney - Los Angeles | 862,964
10. Sydney - Honolulu | 861,814

Top 10 Least Trafficked Routes

1. Hobart - Los Angeles | 2
2. Hobart - Tokyo | 1
3. Cairns - Honiara | 1
4. Townsville - San Francisco | 1
5. Darwin - Zagreb | 1
6. Melbourne - Denver | 0
7. Perth - Bandar Seri Begawan | 0
8. Brisbane - Colombo | 0
9. Adelaide - Harare | 0
10. Brisbane - Chicago | 0

b. Trends and Geographical Patterns

This section highlights the regions and routes that have shown the fastest growth and decline over time based on CMGR.

Top 5 Countries by Monthly CMGR (%)

1. **Guam:** 41.67
2. **Austria:** 9.93
3. **Cyprus:** 8.67
4. **Brunei:** 0.96
5. **Cook Islands:** 0.74

Top 5 Routes by Monthly CMGR (%)

1. **Perth - Zagreb:** 192.40
2. **Sydney - Honiara:** 157.13
3. **Adelaide - Zagreb:** 138.05
4. **Darwin - Christchurch:** 106.06
5. **Adelaide - Kuala Lumpur:** 98.09

To understand growth patterns, I analyzed Compounded Monthly Growth Rates (CMGR) which shows the average monthly growth of a metric over a given period. The results show that Guam leads all countries with a 41.67% CMGR, while routes like Perth → Zagreb (192.40%) and Sydney → Honiara (157.13%) are among the fastest growing, indicating rising traveler demand.

I also examined the average monthly passenger volumes for the top 10 highest trafficked routes, which revealed clear seasonal peaks. For example, Sydney → Auckland consistently surges in December–January during summer holidays, while Sydney → Singapore sees spikes in November–January and July, aligning with holiday travel and school breaks. These patterns help anticipate how emerging high growth routes like Perth → Zagreb may evolve despite historically lower traffic. Overall, this analysis distinguishes mature, stable routes from rapidly expanding opportunities, providing actionable insights for capacity planning and network expansion strategies.

2. Build a Model

a. Explain your model choices

For predicting passenger traffic, I chose to build a Seasonal AutoRegressive Integrated Moving Average (SARIMA) model because the dataset exhibits strong seasonality and trend patterns due to monthly travel cycles. SARIMA is suited for time series forecasting as it extends the ARIMA framework while incorporating the seasonal component allowing the model to capture fluctuations in passenger demand. To optimize performance instead of inputting my own parameters, I used the auto arima library to auto select the best combination of parameters depending on the route. Once the optimal parameters were identified, I fit a SARIMA model for the route and selected a 12 month forecasting period to provide insights into understanding potential traffic growths and declines for the route.

b. Evaluate the model's performance

To evaluate the SARIMA models, I trained each route's model using all historical data except the most recent 12 months and then compared the forecasts against the actual passenger counts for that period. Model accuracy was assessed using Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Symmetric Mean Absolute Percentage Error (sMAPE). MAE measures the average absolute difference between predicted and actual passenger counts, MAPE expresses this difference as a percentage of the actual values, and sMAPE adjusts for scaling to provide a balanced percentage error even when volumes vary across routes. Overall, the SARIMA models produced stable and accurate forecasts across most city pairs, performing especially well on routes with clear seasonal trends and sufficient historical data.

For the **Sydney-Singapore** route the model evaluation metrics were:

MAE = 862.0

MAPE = 2.85%

sMAPE = 2.83%

These values indicate that the SARIMA model performs very well for this route, with an average absolute forecast error of only 862 passengers per month and percentage errors under 3%, which is highly accurate given the overall passenger volumes. The low MAPE and sMAPE values demonstrate that the model successfully captures the seasonal trends and short term fluctuations in traffic for this route, making the forecasts both reliable and actionable for capacity planning and strategic decision making.

3. Provide Recommendations

a. Which routes should AeroConnect invest more in or scale back from?

To identify where AeroConnect should focus resources, I combined historical growth (CMGR) with forecasted growth from the SARIMA based forecasts. Routes showing both strong historical growth and strong future demand were classified as Invest, while routes with consistent declines in both were classified as Scale Back. Forecasted growth is calculated by comparing the total passenger traffic predicted for the next 12 months with the total passenger traffic from the previous 12 months of historical data. The SARIMA time series model is used to generate the monthly passenger forecasts for the upcoming year. The growth is then calculated as the percentage change between these two totals.

Based on this analysis, I recommend investing in the following top three routes:

Darwin → Bangkok: CMGR 11.3%, forecasted growth 799%

Adelaide → Los Angeles: CMGR 5.6%, forecasted growth 84.7%

Perth → Christchurch: CMGR 8.7%, forecasted growth 62.2%

These routes demonstrate strong past performance and significant expected demand growth.

Conversely, the top three routes to scale back are:

Adelaide → Muscat: CMGR -6.9%, forecasted growth -100%.

Sydney → Johannesburg: CMGR -0.7%, forecasted growth -100%.

Brisbane → Manila: CMGR -0.2%, forecasted growth -97.2%.

These routes show sustained declines both historically and in future projections.

b. How can AeroConnect use this model going forward?

AeroConnect can use this SARIMA based forecasting model to continuously monitor route performance and anticipate future demand. By combining historical growth trends (CMGR) with forecasted passenger traffic, the model helps identify which routes are likely to deliver strong returns and which may require scaling back. As new monthly passenger data becomes available, AeroConnect can update the model to refine forecasts and adjust investment decisions proactively. Over time, this approach enables data driven route optimization, ensuring resources are allocated to the most profitable and high growth opportunities.