1. Understanding the Data
    a. Identify the most and least trafficked routes

```
MOST TRAFFICKED ROUTES BY YEAR:
 Year      Total_Route Total_Value
 1985 Sydney-Auckland      556423
 1986 Sydney-Auckland      610302
 1987 Sydney-Auckland      717006
 1988 Sydney-Auckland      770231
 1989 Sydney-Auckland      307250

LEAST TRAFFICKED ROUTES BY YEAR:
 Year                                                          Total_Route Total_Value
 1985                            Adelaide-Bangkok, Perth-Bandar Seri Begawan           0
 1986                     Adelaide-Harare, Brisbane-Colombo, Melbourne-Denver           0
 1987 Cairns-Manila, Cairns-Wellington, Melbourne-Harare, Townsville-Honolulu           2
 1988                                           Hobart-Tokyo, Sydney-Seattle           0
 1989                                Brisbane-Chicago, Port Hedland-Jakarta           0
```

    b. Analyze trends and/or geographical patterns
        i. Asia-Pacific dominates in passenger volume
            1. New Zealand leads with 7.9M passengers, followed by Singapore (4.1M) and other Asia-Pacific nations. This suggests Australia's strongest aviation connections are regional, likely due to geographic proximity and established trade/tourism relationships.

```
TOP COUNTRIES BY TOTAL PASSENGER VOLUME:
    Country  Passengers_Total  Route
New Zealand           7881613     24
  Singapore           4130761      9
        USA           3084101     39
  Hong Kong           2217453      6
      Japan           1773146     13
         UK           1729405     16
  Indonesia           1552095     16
       Fiji            933701      7
   Thailand            930728      8
   Malaysia            835317      6
```

        ii. YoY saw emerging South American markets grow while several small Asia-Pacific countries saw severe declines.
            1. Argentina experienced 446% growth while Malaysia (-46%), Fiji (-45%), Brunei (-45%), and Thailand (-37%).

COUNTRY GROWTH RATES (1985-1989):

Country	growth_rate	Passengers_Total
Argentina	446.326014	12937
Austria	0.728036	8163
Sri Lanka	-29.809843	1255
Thailand	-36.542531	168814
Bahrain	-38.099454	4195
Western Samoa	-40.927714	4037
Zimbabwe	-43.631798	21102
Brunei	-44.767616	7368
Fiji	-44.886452	112899
Malaysia	-46.367297	110286

iii. Sydney is clearly the Hub
    1. Connected to 43 countries with 63 foreign ports, handling 15.5 million passengers. Sydney dominates both domestic and international flights.

PORT CONNECTIVITY ANALYSIS:

AustralianPort	Country	ForeignPort	Passengers_Total
Sydney	43	63	15499688
Melbourne	35	49	6180004
Brisbane	31	43	3153897
Perth	29	37	2811590
Adelaide	25	34	604937
Cairns	20	28	474393
Darwin	16	21	277604
Townsville	15	19	110539
Hobart	6	9	70312
Norfolk Island	1	1	91120

iv. Route concentration is a big indicator of efficiency and growth potential.
    1. The USA has 39 routes but only 3.1 million passengers, suggesting several low-volume routes, while Singapore has 4.1 million passengers with just 9 routes.

PORT CONNECTIVITY ANALYSIS:

AustralianPort	Country	ForeignPort	Passengers_Total
Sydney	43	63	15499688
Melbourne	35	49	6180004
Brisbane	31	43	3153897
Perth	29	37	2811590
Adelaide	25	34	604937
Cairns	20	28	474393
Darwin	16	21	277604
Townsville	15	19	110539
Hobart	6	9	70312
Norfolk Island	1	1	91120

2. Build a Model

a. Your model should predict passenger traffic for the next 6–12 months on at least 1 city pair
    i. I chose the Adelaide-London route due to it having the largest sample size. My SARIMA model forecasted 2,170 passengers/month with a range of 1,713-2,731.

3. Evaluate Your Model
    a. Explain your model choices — why did you choose the elements you did
        i. I chose a model that is built specifically for time series predictions, the SARIMA model. Initially I chose the ARIMA model but identified strong seasonal heterogeneity so I opted for the SARIMA model which takes seasonality into account. I validated the SARIMA(1,1,1)(1,1,1,12) specification assumptions pre model and validated the results. The ADF test (p-value: 0.9660) and KPSS test (p-value: 0.0100) confirmed the series was non-stationary, requiring first differencing (d=1). Strong seasonality was detected with a seasonal strength of 0.752, well above the 0.3 threshold, justifying seasonal differencing (D=1) within a 12-month period. I do not have extensive experience with time series forecasting so I used online resources like youtube to learn about SARIMA and how to implement it.
    b. Evaluate the model's performance & report the accuracy of the model
        i. The model demonstrates mixed but acceptable performance for aviation forecasting. The Ljung-Box test (p-value: 0.0640) successfully validates that residuals show no significant autocorrelation, confirming the model captures seasonality effectively. With a MAPE of 12.1% and MAE of 242.5 passengers, the forecast accuracy falls within industry-acceptable ranges for aviation traffic prediction. Here I was most concerned with MAPE as MAPE deals better with the scaling issues that come with time series better than MAE. The failed normality test (p-value: 0.0000) indicates non-normal residual distribution, which effects the confidence in the prediction. Despite this limitation, the model's ability to generate reasonable forecasts (average 2,170 passengers/month, range 1,713-2,731) makes it suitable for operational planning purposes on the Adelaide-London route.
4. Provide Recommendations
    a. Which routes should AeroConnect invest more in or scale back from?
        i. Based on the passenger volume and growth data, AeroConnect should prioritize investment in Asia-Pacific routes, particularly those connecting to Singapore, which demonstrates both high volume

(4.1M passengers) and efficient route density with only 9 routes. New Zealand represents the strongest market with 7.9M passengers across 24 routes, suggesting room for route optimization and frequency increases. Conversely, AeroConnect should scale back from routes showing severe decline during the 1985-1989 period, particularly Malaysia (-46% growth), Fiji (-45%), and Brunei (-45%), unless market conditions have fundamentally changed. The USA market presents a mixed opportunity, while it has high route diversity (39 routes), the relatively low passenger density suggests many underperforming routes that could be consolidated or eliminated. I would suggest improving marketing initiatives in the USA, understanding how consumers react, and make a decision from there. While the USA is an inefficient market, there is substantial route diversity, and the reasoning for the lack of volume is non-conclusive.

b. How can AeroConnect use this model going forward?
    i. AeroConnect can leverage the SARIMA model for connecting flight optimization by forecasting demand patterns across feeder routes that connect through major hubs like Sydney (15.5M passengers, 63 foreign ports). The model's ability to predict seasonal peaks and troughs enables optimal scheduling of connecting flights. For example, if the Adelaide-London route shows predictable monthly patterns, AeroConnect can coordinate domestic feeder flights from smaller cities to Adelaide to maximize passenger capture during high-demand periods. By applying SARIMA models to multiple route pairs feeding into hub airports, the airline can synchronize arrival and departure times to minimize connection times while maximizing load factors. This SARIMA model also enables AeroConnect to utilize dynamic pricing, maximising revenue and route efficiency. As for continuing my education on time series models, I would like to learn more about dynamic pricing.