

Ellie Meltzer

Generate Tech Challenge

1. Understanding the Data

a) Identify the most and least trafficked routes

Least:

Route	Passengers_Total
Townsville-San Francisco	1
Cairns-Honiara	1
Darwin-Zagreb	1
Hobart-Tokyo	1
Hobart-Los Angeles	2
Perth-Port Moresby	2
Cairns-Belgrade	3
Hobart-Denpasar	3
Adelaide-Jakarta	3
Townsville-Zagreb	4

Most:

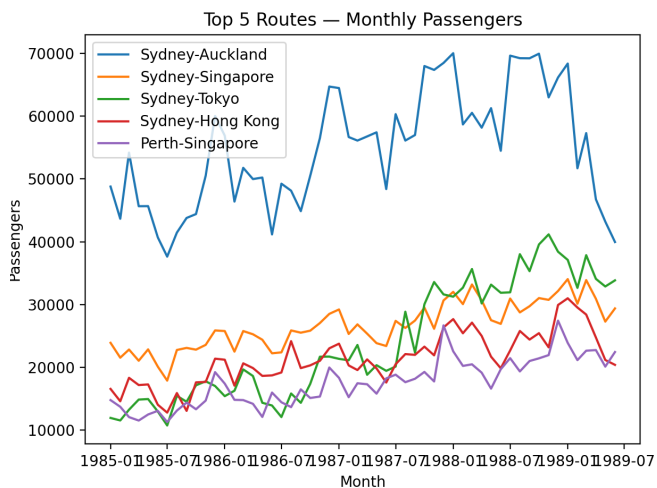
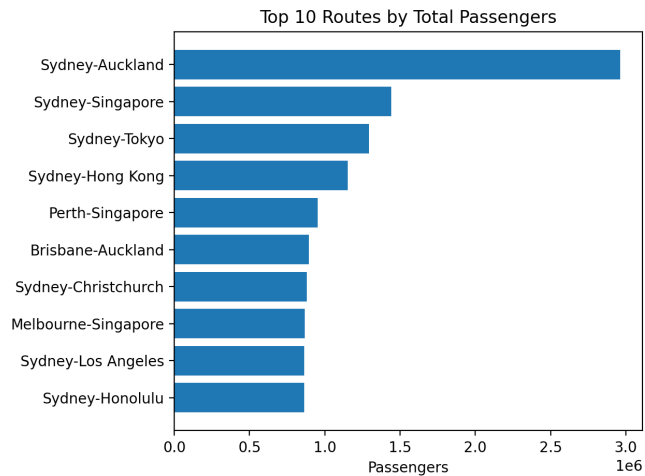
Route	Passengers_Total
Sydney-Auckland	2961212
Sydney-Singapore	1440018
Sydney-Tokyo	1292116
Sydney-Hong Kong	1151900
Perth-Singapore	952926
Brisbane-Auckland	893246
Sydney-Christchurch	882357
Melbourne-Singapore	865251
Sydney-Los Angeles	862964
Sydney-Honolulu	861814

b) Analyze trends and/or geographical patterns

Country	YoY_pct
Thailand	34.362785435258600
Fiji	25.488811597992300
Japan	23.610533709414600
New Caledonia	12.785025955405900
Philippines	9.30876251584718
Canada	9.019753037419800
Papua New Guinea	8.807047207828300
Germany	5.040285280232010
USA	3.619259429109660
Singapore	3.333870468720520
Indonesia	0.9701514103675870
Hong Kong	-1.410759380145960
Malaysia	-3.2805158966272600
Greece	-3.567919807764010
UK	-7.207684547861020
Italy	-9.376148781608660
New Zealand	-9.655780990822480

To evaluate trends, I calculated YoY (year-over-year) passenger growth by country and averaged the last twelve months. The results shown above show strong growth in Thailand (+34.4%), Fiji(+25.5%), and Japan(+23.6%), while notable declines occurred in New Zealand (-9.7%), Italy (-9.4%), and the UK (-7.2%). These findings suggest AeroConnect should prioritize expanding service in rapidly growing Asian and Pacific markets while carefully managing capacity in established but declining regions.

c) Create visualizations to demonstrate trends & patterns determined in part b



2. Build a Model

a) Your model should predict passenger traffic for the next 6–12 months on at least 1 city pair

model	rmse	mae	mape
seasonal model	9855.925527316040	8263.333333333330	15.613933386630500
holt_winters	10732.270168894300	9600.818355329000	17.694704470801900

3. Evaluate your model

a) Explain your model choices

I implemented a seasonal-naive baseline where each forecasted month is set equal to the observed value from the same month one year prior. This choice provides a

transparent and acceptable benchmark in time series forecasting, especially for strongly seasonal data like air travel. I also tested a Holt-Winters model, which can capture trend and seasonality to compare performance against the baseline.

b) Evaluate the model's performance & report the accuracy of the model

To evaluate model accuracy, I used a 12-month period at the end of the time series for the busiest route. On the busiest route (Sydney → Auckland), the baseline model achieved lower error (RMSE = 9,856; MAPE = 15.6%) compared to the Holt-Winters model (RMSE = 10,732; MAPE = 17.7%). This indicated that with the relatively short historical dataset, the simpler baseline approach was more accurate. Holt-Winters will most likely become more effective with longer data histories or more complex seasonal structures.