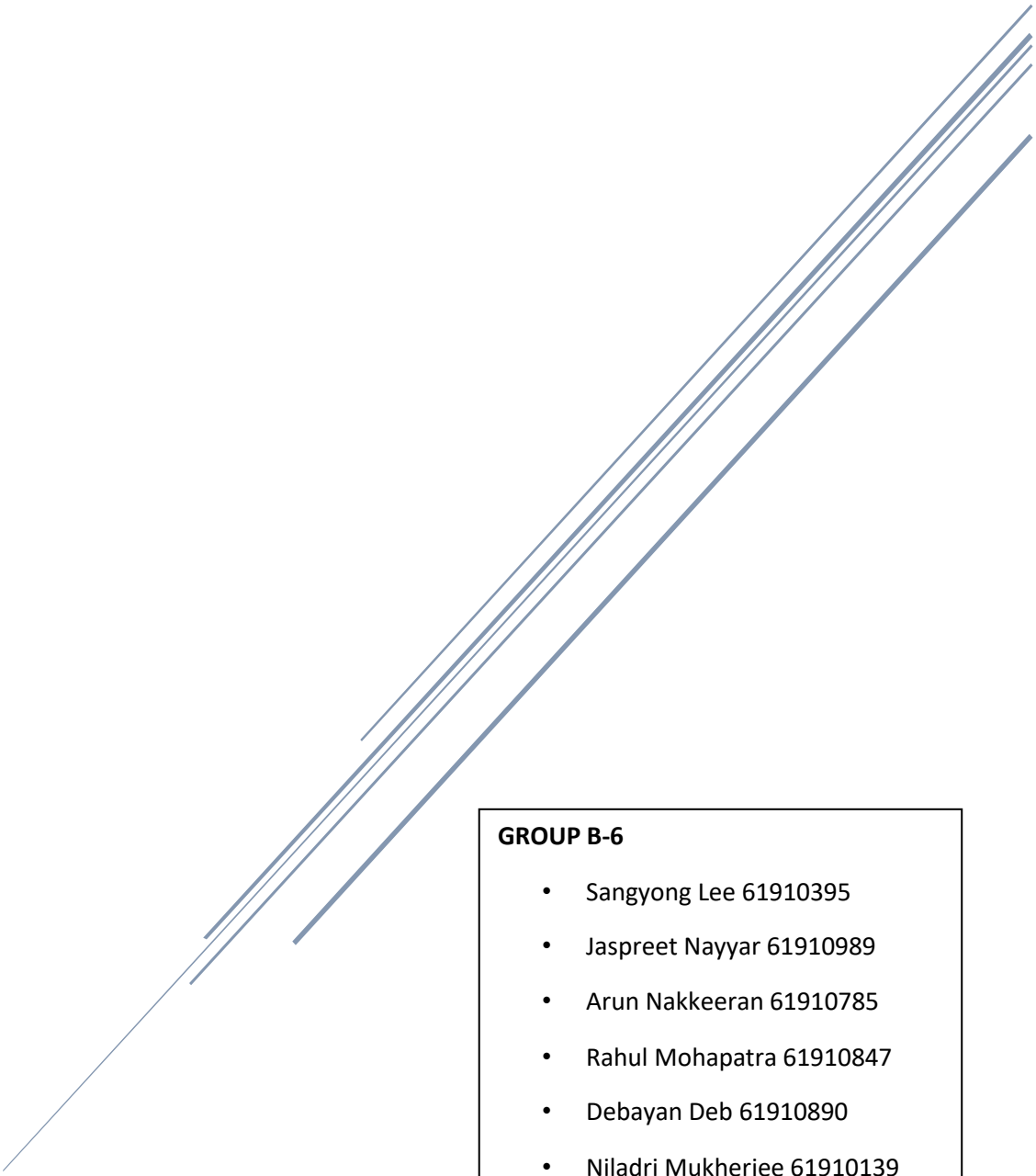
FORECASTING TRAFFIC AND FREIGHT DEMAND IN ORDER TO DECIDE ON EXPANSION

FCAS FINAL PROJECT REPORT



GROUP B-6

- Sangyong Lee 61910395
- Jaspreet Nayyar 61910989
- Arun Nakkeeran 61910785
- Rahul Mohapatra 61910847
- Debayan Deb 61910890
- Niladri Mukherjee 61910139

1. **Executive Summary**

a. Problem Description

Our client, The Australian Airports Association, appointed us for forecasting the passenger and freight traffic for the month of September 2019. Based on our forecast they intend to take the decision of whether to go ahead with investing a budgeted $ 1.2 Billion on building new infrastructure (including runways, facilities) and recruiting staff. As per their own business estimates, they have set a threshold of a 40% increase in their YoY traffic for both passenger and freight data to decide for a go/ no go on the Investment.

To undertake the project of infrastructure expansion, the Airport Association needs one year for implementation and needs the forecasted traffic for the next year every month, so that the project can be started at an implementable pace. Hence the Business model is based on monthly forecasted Traffic for one year ahead to facilitate the authorities in the decision making.

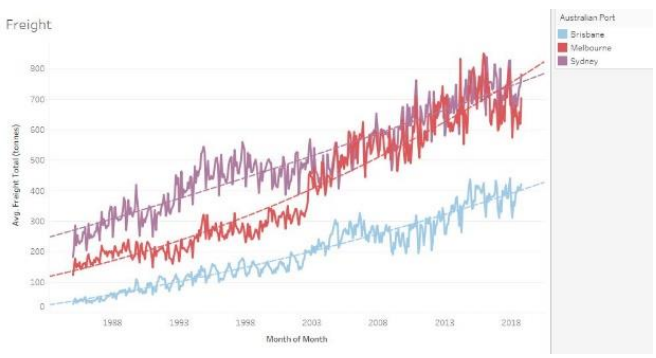
b. The dataset

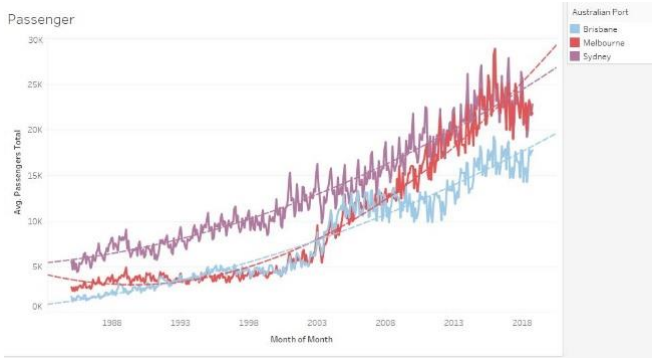
To accurately forecast the passenger and freight traffic our team has used government data from the Australian Government; Data source: https://data.gov.au/dataset/ds-dga-d9fbffaa-836f-4f52-80e8-324249ff269f/details

A brief snapshot of the dataset gives us the following outlook:

Month	Australian Port	Foreign Port	Country	Passengers In	Freight In (tonnes)	Mail In (tonnes)	Passengers Out	Freight Out (tonnes)	Mail_Out_ (tonnes)	Passengers Total	Freight Total (tonnes)	Mail Total (tonnes)	Year
01 January 1985	Adelaide	Auckland	New Zealar	1513	42.167	0.311	985	18.704	0.924	2498	60.871	1.235	1985
01 January 1985	Adelaide	Bahrain	Bahrain	12	0	0	5	0.033	0	17	0.033	0	1985
01 January 1985	Adelaide	Bombay	India	7	0	0	5	0	0	12	0	0	1985
01 January 1985	Adelaide	Frankfurt	Germany	115	0.009	0	171	0	0.248	286	0.009	0.248	1985
01 January 1985	Adelaide	London	UK	1567	2.8	0	1472	10.618	2.487	3039	13.418	2.487	1985
01 January 1985	Adelaide	Muscat	Oman	17	0	0	14	0.1	0	31	0.1	0	1985
01 January 1985	Adelaide	Rome	Italy	79	0.005	0	44	0	0	123	0.005	0	1985
01 January 1985	Adelaide	Singapore	Singapore	2496	37.345	0	2037	133.203	0.112	4533	170.548	0.112	1985
01 January 1985	Brisbane	Abu Dhabi	United Aral	0	0	0	3	0	0	3	0	0	1985
01 January 1985	Brisbane	Auckland	New Zealar	7157	223.258	0.671	5652	33.032	3.218	12809	256.29	3.889	1985
01 January 1985	Brisbane	Bahrain	Bahrain	12	0	0	15	0.149	0	27	0.149	0	1985

We have plotted the dataset against time (as depicted below) and the brief analysis of the plot shows a shift in traffic trend post Sep 2001. To make the data more relevant to the current demand forecast,





we have truncated the data from Oct 2001 to Sep 2018 (This truncation was done based on changes in world-wide air transport regulations post 9/11 attacks in the USA).

c. Metrics

Since the business decision is based on the incremental change in the Passenger & Freight traffic, we have used MAPE (which acts as a proxy for the incremental change) to benchmark different models as it provides the percentage change in forecasted traffic vs actual Traffic.

To decide the best model for predicting each data set, The MAPE for both training and validation were compared to look for overfitting and then the model with least values of MAPE have been decided as the best predictor (since the data differences were very small, the decision to go ahead with the model with the minimum MAPE was used to forecast the data).

d. Analysis and Final forecasting method

We have analysed each of 6 datasets- passenger and freight data for cities of Brisbane, Sydney and Melbourne- with multiple methods of prediction (as indicated in technical summary) and each model used a different final method for forecasting based on MAPE data.

After a detailed analysis we have concluded that the decision is a no go for all 3 airports as the increment of traffic in each of the 6 cases is less than the stipulated 40%

Technical Summary

The section presents a summary of the technical details covering, among other, aspects like data preparations and related issues, performance of the models employed against adopted benchmarks and the details of the forecasting methods themselves.

a. Data Preparation

The Raw data (containing traffic across various international cities from 19 Airports of Australia) obtained could not be used directly for forecasting the traffic and hence the data had to be cleaned and prepared accordingly. Due to the nature of assignment, the dataset was partitioned into 6 sections 3 sets of Freight and Passenger data each for the 3 busiest cities Melbourne, Sydney and Brisbane.

The imputed dataset was extracted using pivot tables (after computing monthly total traffic for each airport) following which the data was truncated from October 2001 to September 2018 (both months included). Since the regulatory norms regarding air security had changed significantly since sep-2001, it made sense to use only the above period as the relevant data for our forecast. Further

to make the data parsimonious, the variables not of business concern were removed from the data set.

To facilitate understanding of the forecasting, a continuous time index variable was computed (based on the number of months since October 2001) in the time series data, the months were considered as dummy variables (to understand the monthly seasonality in the data including the seasonal impact of holiday traffic like Easter & Christmas),with January as 1, followed by other months in sequence. The Time series data also had another continuous time variable namely, the year index with Year 2001 as value 1 followed by further years till 18 for Year 2018. The motive behind adding the year index as another continuous variable along with time index variable was to capture the yearly trend in the data and to capture external factors like oil prices and economic policies in the country. Using ***Monthly data to capture trends in the traffic had salient benefits such as, it could be used to capture/remove the effects of holidays and seasonal spikes on traffic when using a rolling over method to forecast the traffic in future and in absence of which the model could have overestimated the traffic leading to excess investment from the Airport Association***.

The imputed data was partitioned into training and validation sets based on time series partitioning, with the data from the last twelve months (Oct 2017 to Sep 2018) into the validation set and the rest (192 data points Oct 2001 to Sep 2018) into the training data set.

b. **Model Performance**

i. Brisbane Passenger Traffic

Model	MAPE value Training	MAPE value validation
Naïve		6.584
Seasonal Naïve		6.067
Linear Regression	5.027	3.8
Linear regression with AR Model		2.316
Holts-Winter Additive (Default)	8.133	6.176
Holt-Winter Additive (optimised) α = 0.008 β = 0.204 γ = 0.099	7.917	6.199
Holt-Winter Multiplicative (optimised) α = 0.614 β = 0.013 γ = 0.446	2.649	3.148
Holt-Winter Multiplicative	3.542	2.824

ii. Brisbane Freight Traffic

Model	MAPE(%)	Alpha	Beta	Gamma
Naïve	8.29			
Naïve Seasonal	8.63			
Linear Reg	10.15			
Linear with Arima	4.73			
HWA	4.10	0.2	0.15	0.05
HWA optimized	4.11	0.572954	0.005921	0.499619
HWM	4.67	0.2	0.15	0.05
HWM optimized	4.32	0.572954	0.005921	0.499619

i. Melbourne Passenger Traffic

Model	MAPE (Validation Data)
Naïve	6.93%
Seasonal Naïve	7.66%
Linear Regression (Additive Seasonality, Linear Trend) with ARIMA Model	3.22%
Linear Regression (Additive Seasonality, Quadratic Trend) with ARIMA Model	1.82%
Linear Regression (Multiplicative Seasonality, Linear Trend) with ARIMA Mode	1.96%
Holts-Winter Additive Smoothing (Optimized)	2.84%
α = 0.34 β = 0.02 γ = 0.64	
Holts-Winter Multiplicative Smoothing (Optimised)	2.35%
α = 0.28 β = 0.03 γ = 0.52	

iv. Melbourne Freight Traffic

Method	Training	Validation
Naïve		7.4
Moving A verage	6.8	9.7
Additive Holt Winters	4.9	5.2
Multiplicative Holt Winters	5.0	4.9
Multiplicative Holt Winters Optimized	3.8	3.9

v. Sydney Passenger Traffic

Sl. No.	Method name	MAPE Training	MAPE Validation
1	Naïve		5.58%
2	Naïve Seasonal		4.78%
3	Linear Regression	3.46%	6.69%
4	Linear Regression with ARIMA Model	2.60%	2.28%
5	Holts Winter Additive	2.99%	2.53%
6	Holts Winter Additive (optimized)	2.46%	1.50%
7	Holts Winter Multiplicative	3.09%	3.07%
8	Holts Winter Multiplicative (optimized)	2.39%	1.18%

vi. Sydney Freight Traffic

Model	MAPE(%)	α	β	γ
Naïve Seasonal	7.86%			
Linear Reg	10.51%			
Linear with Arima	9.32%			
HWA	3.21%	0.20	0.15	0.05
HWA Optimized	2.83%	0.61	0.01	0.45
HWM	3.45%	0.20	0.15	0.05
HWM Optimized	2.67%	0.55	0.03	0.33

c. Forecasting Methods

The Best model for each Time series, with respect to the mentioned criteria is as depicted below;

Sl no	Airport City	Type of Traffic	Best Model
1	Melbourne	Passenger Data	Linear Regression with Arima lag 1
2	Melbourne	Freight Data	Holt Winter's Multiplicative Optimized
3	Sydney	Passenger Data	Holt Winter's Multiplicative (Non-optimized)
4	Sydney	Freight Data	Linear regression with t and t^2
5	Brisbane	Passenger Data	Linear Regression with AR lag1
6	Brisbane	Freight Data	Holt Winter's Additive (Non-optimized)

d. Recommendation

City	Freight		Passenger	
	Sep-18	Sep-19	Sep-18	Sep-19
Melbourne	27203.94	28267.78	951523	1017328
% change	3.91%		6.92%	
Go/No Go	No Go		No Go	
Sydney	45998.26	48699.72	1424726	1416220
% change	5.87%		-0.597%	
Go/No Go	No Go		No Go	
Brisbane	11269.6	12430.91	528778	642764
% change	10.30%		21.56%	
Go/No Go	No Go		No Go	

Based on our model estimate as depicted above the decision is a No Go for all the 3 airports considering that all the increments are less than the mentioned 40% by our client. However, there is a considerable increase in passenger traffic in the year on year data for September. Thus, to account for this change we recommend subsequent forecast of monthly data to get a better estimate.

The model has not factored in the crude oil price change data to forecast. We recommend including that in future regression as a numerical or as a dummy variable to better forecast passenger traffic.

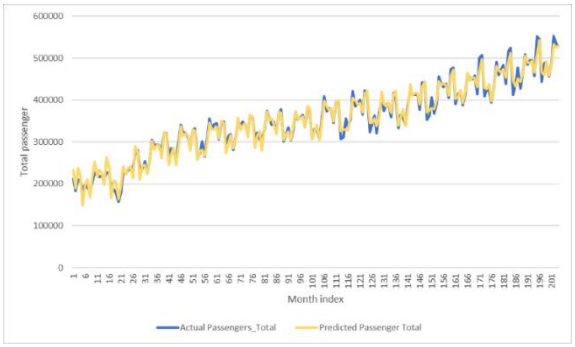
It is also recommended to factor in economic depression as a dummy for regression to get a better forecast.

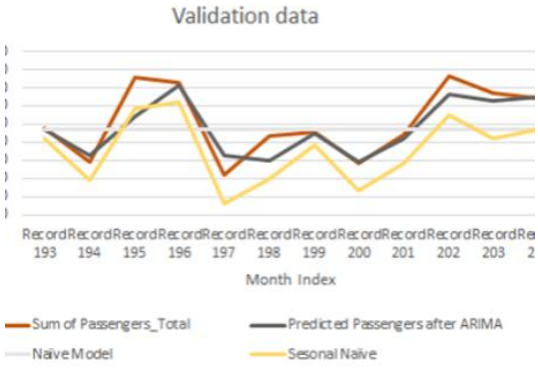
Appendix

Brisbane Passenger Data- Linear Regression with AR lag1:

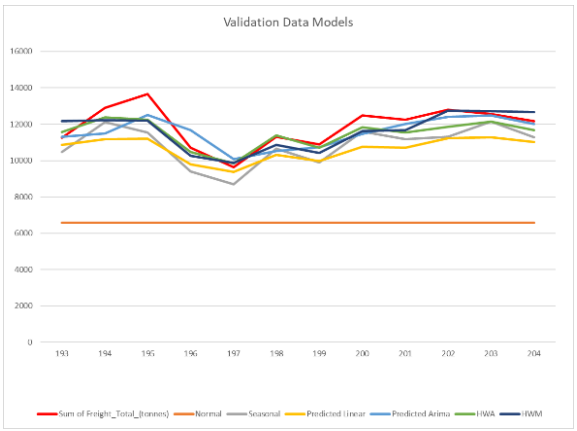
Arun Nakkeeran

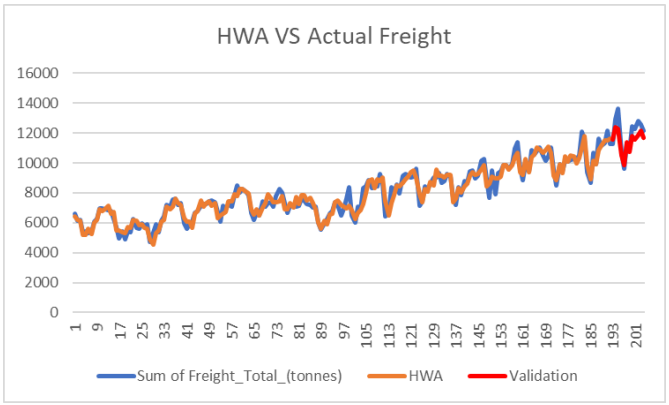
Year	Row Lab	Year_no	Month	Sum of Passengers_Total	Residual	Fitted Residual	New Predicted Sum of Passengers
2018	204	18	9	528778	24897.69063		528778
2018	205	18	10	504838.32	18970.01365	18970.01365	316.2471947
2018	205	18	11	488191.2134	14528.89887	14528.89887	502720.1123
2018	205	18	12	548889.6693	11201.54136	11201.54136	560091.2106
2019	205	19	1	561029.0627	8708.629431	8708.629431	569737.6921
2019	205	19	2	500932.331	6840.898096	6840.898096	507773.2291
2019	205	19	3	528644.1619	5441.562526	5441.562526	534085.7244
2019	205	19	4	549855.8678	4393.156899	4393.156899	554249.0247
2019	205	19	5	534181.5111	3607.67386	3607.67386	537789.185
2019	205	19	6	576825.6545	3019.176822	3019.176822	579844.8313
2019	205	19	7	636139.9229	2578.264993	2578.264993	638718.1879
2019	205	19	8	627638.3163	2247.926473	2247.926473	629886.2427
2019	205	19	9	641985.8346	2000.431326	2000.431326	643986.266





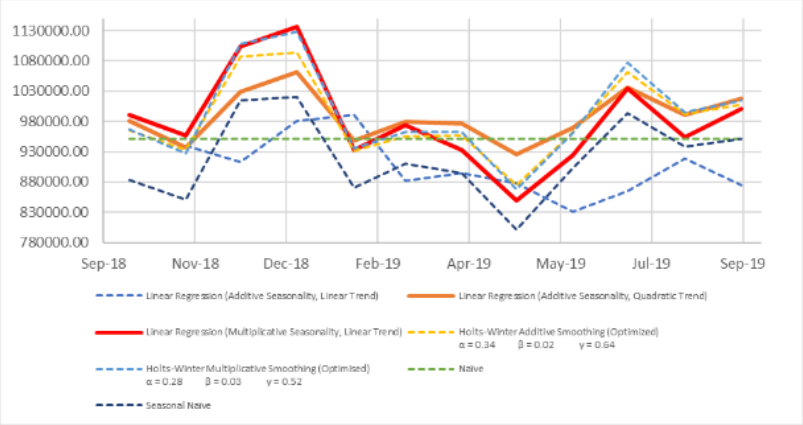
Brisbane Freight Data - Holt Winter's Additive (Non-optimized): **Niladri Mukherjee**

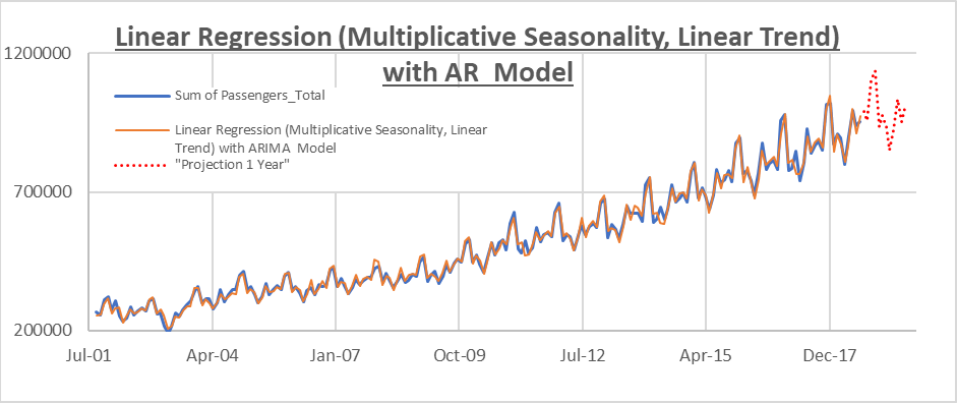




	HOLT Winter's Additive Forecast		
	Time Index	Forecast_Freight	
Oct-18	205	12871.2182	
Nov-18	206	12827.1878	
Dec-18	207	12619.1171	
Jan-19	208	11097.4808	
Feb-19	209	10904.5305	
Mar-19	210	11837.0577	
Apr-19	211	11568.7093	
May-19	212	12463.3161	
Jun-19	213	12399.5432	
Jul-19	214	12926.6878	
Aug-19	215	12717.6809	Sep-18
Sep-19	216	**12430.911**	11269.6
	% Increase	10.30%	

Melbourne Passenger Data : Jaspreet Nayyar

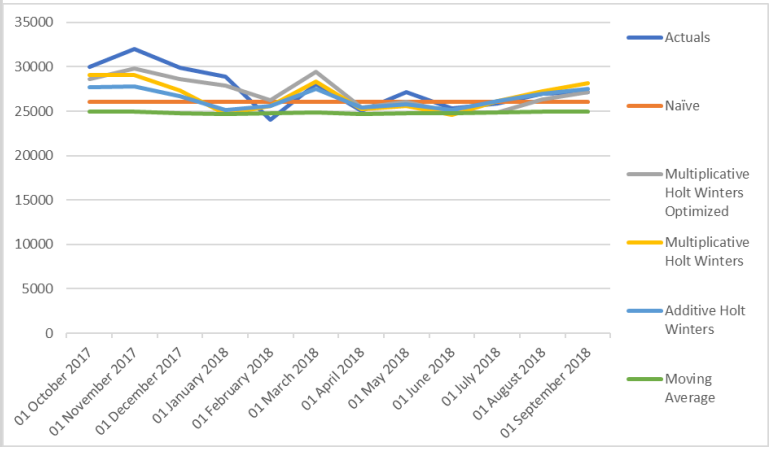




Melbourne Freight Data: Debayan Deb

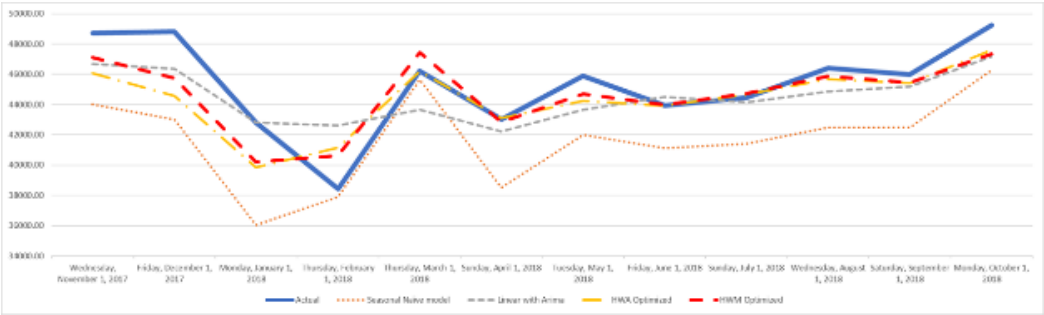
12 Month Forecast

Month	Forecast (tonnes)
01-Oct-18	30604.22457
01-Nov-18	31826.80885
01-Dec-18	30108.08425
01-Jan-19	29238.01083
01-Feb-19	26554.30127
01-Mar-19	30622.7118
01-Apr-19	26814.9004
01-May-19	27917.63436
01-Jun-19	26186.58313
01-Jul-19	26262.21533
01-Aug-19	27561.7153
01-Sep-19	28267.78025

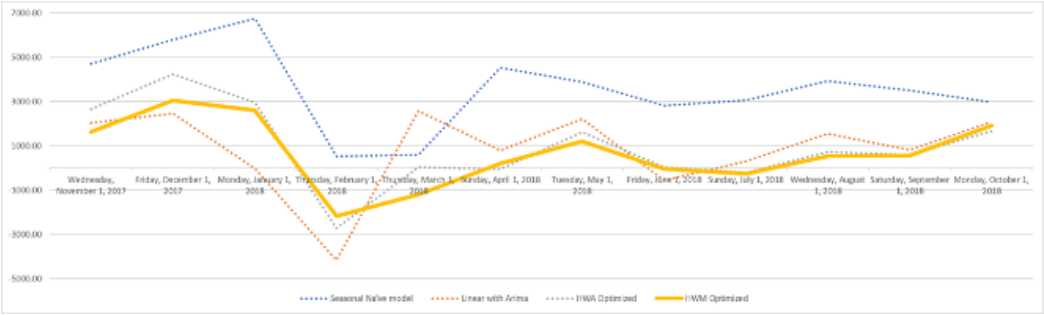


Sydney Freight Data – Linear regression with t and t^2: **Sangyong Lee**

Forecast plot by model

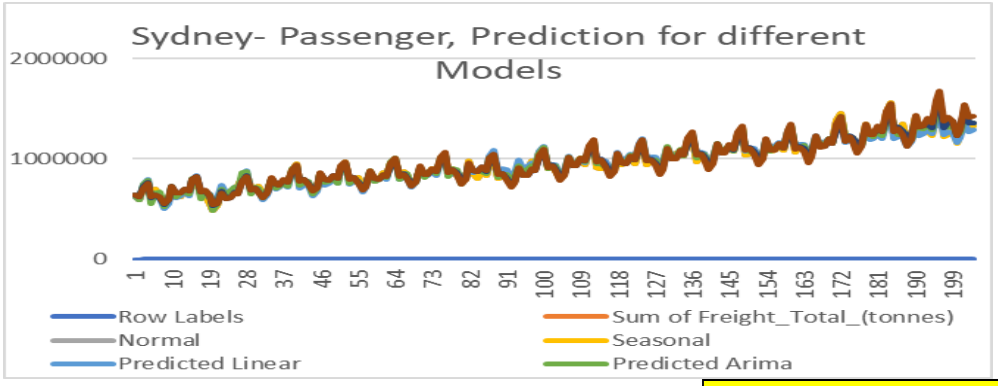


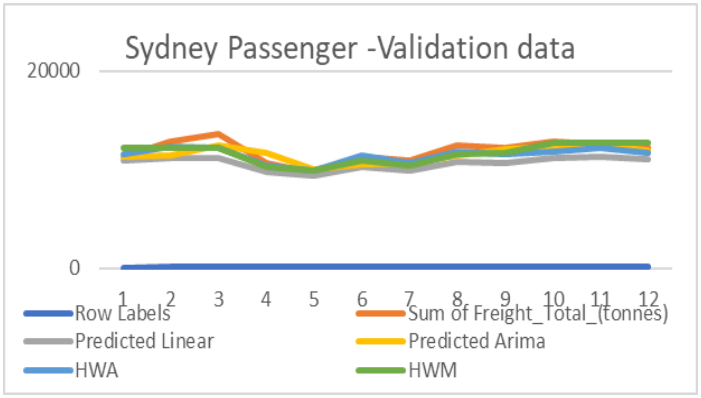
Residual plot by model



Month	HWM Optimized Forecast	vs. last year
Nov.2018	51325.31755	5.3%
Dec.2018	49766.66514	2.0%
Jan.2019	42287.85512	-1.2%
Feb.2019	41631.6554	8.3%
Mar.2019	50554.59122	9.4%
Apr.2019	45489.51567	5.7%
May.2019	47711.08633	4.0%
Jun.2019	46558.79428	6.0%
Jul.2019	47586.26224	7.0%
Aug.2019	49281.97101	6.2%
Sep.2019	48699.72028	5.9%
Oct.2019	51494.01331	4.6%

Sydney – Passenger Traffic Analysis- Holt Winter's Multiplicative (Non-optimized): **Rahul Mohapatra**





Forecast using HWM (non-optimised)			
Oct-18	205	1519074	
Nov-18	206	1449648	
Dec-18	207	1683038	
Jan-19	208	1710267	
Feb-19	209	1342965	
Mar-19	210	1374980	
Apr-19	211	1292755	
May-19	212	1171960	
Jun-19	213	1282703	
Jul-19	214	1593641	
Aug-19	215	1422622	Sep-18
Sep-19	216	1416220	1424726
	% change	-0.00597	