

DSCI 6840 Homework 6

Forecasting

Task 1: Using the Whole Data Available to Forecast (40%)

1. Conduct the forecast for the next 8 quarters in Excel. Record the RMSE value. **Excel Skip**
2. Conduct the forecast for the next 8 quarters in R using Seasonal Naïve, ETS, and ARIMA methods. Record the RMSE value.

```
> accuracy (snaive (Train, h=8), Test)
```

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	18750.43	26058.45	22806.75	3.33619	4.096587	1.00000
Test set	122578.25	125089.71	122578.25	14.05121	14.051212	5.374648

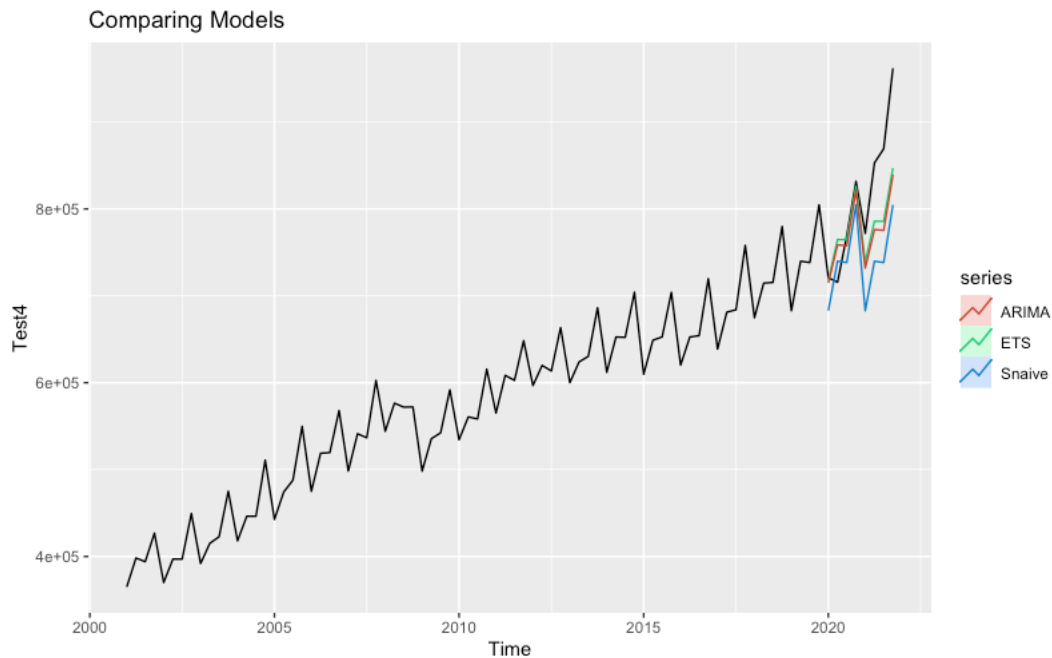
```
> accuracy (forecast (fit2, h=8), Test)
```

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	-743.0113	10760.20	7484.503	-0.1756672	1.360322	0.3281705
Test set	74679.7177	80285.57	74679.718	8.4302608	8.430261	3.2744568

```
> accuracy (forecast (fit3, h=8), Test)
```

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	-35.91559	10539.51	7197.935	-0.04336054	1.287306	0.3156055
Test set	83269.33105	88471.07	83269.331	9.42713421	9.427134	3.6510827

- Plot the forecasts for the next 8 quarters from Seasonal Naïve, ETS, and ARIMA on the same graph.



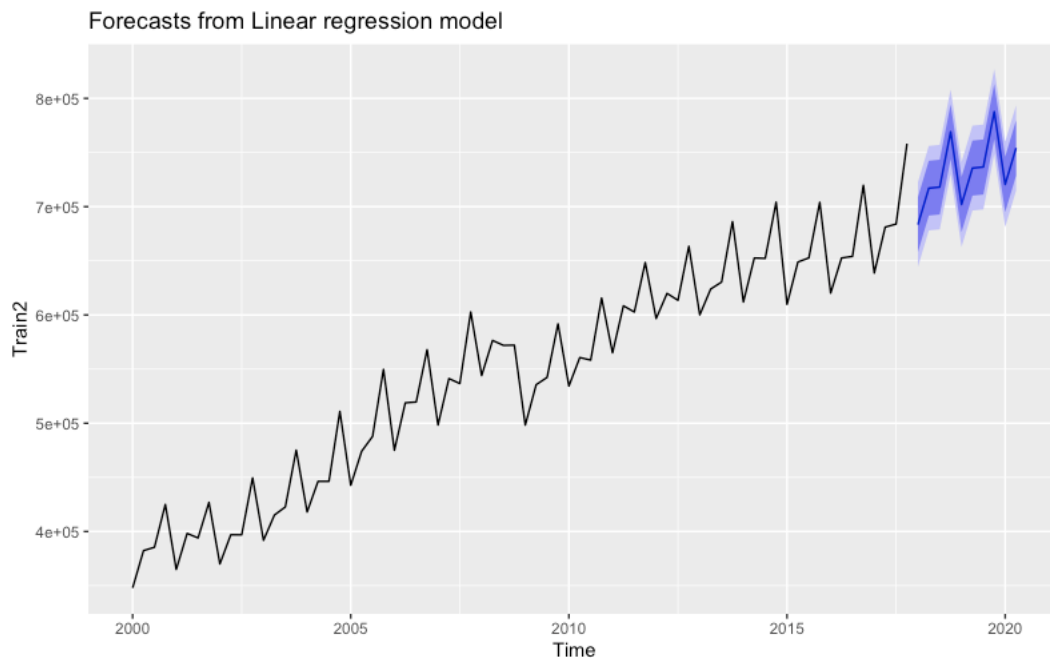
- Which forecast result will you report? Why? **ETS model would be better fit because RMSE and MAPE results are smaller than other model's values.**
- Please report your point forecasts of your selected method for the next 8 quarters.

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2020 Q1	717442.9	699458.7	735427.1	689938.4	744947.3
2020 Q2	764726.6	740802.5	788650.7	728137.9	801315.3
2020 Q3	764610.8	736235.8	792985.8	721214.9	808006.7
2020 Q4	825841.4	792686.1	858996.6	775134.8	876548.0
2021 Q1	738594.2	701123.1	776065.4	681287.1	795901.4
2021 Q2	785878.0	744850.8	826905.1	723132.4	848623.6
2021 Q3	785762.2	741835.5	829688.9	718582.1	852942.3
2021 Q4	846992.8	799438.8	894546.7	774265.2	919720.3

Task 2: Selecting Forecasting Models (60%)

Let's split the original data into two sets: **Training data** from 2000Q1 to 2017Q4 and **Test data** from 2018Q1 to 2019Q4 (to avoid the pandemic impact).

- Using a regression model (either in Excel or in R) to forecast the test period (2018Q1 to 2019Q4). Calculate/Get RMSE and MAPE for the test period and record these two values
Point Forecast Lo 80 Hi 80 Lo 95 Hi 95



2018 Q1	683413.6	658123.3	708703.9	644412.6	722414.6
2018 Q2	717066.4	691776.1	742356.7	678065.4	756067.5
2018 Q3	718062.8	692772.5	743353.1	679061.8	757063.8
2018 Q4	769175.6	743885.4	794465.9	730174.6	808176.7
2019 Q1	701973.8	676565.4	727382.2	662790.6	741157.0
2019 Q2	735626.6	710218.2	761035.0	696443.4	774809.8
2019 Q3	736623.0	711214.6	762031.4	697439.8	775806.2
2019 Q4	787735.9	762327.4	813144.3	748552.7	826919.1

	ME	RMSE	MAE	MPE	MAPE
Training set	-3.877984e-12	17954.35	14703.898	-0.15957776	2.744966
Test set	6.302649e+01	10466.69	8303.434	-0.06788277	1.134272

- Conduct the forecast for the test period (2018Q1 to 2019Q4) in Excel. Calculate RMSE and MAPE and record these two values. **Excel Skip**
- Fit the models on the Training Data using Seasonal Naïve, ETS, and ARIMA methods in R.
- Using model fit from Step 3 to forecast the test period (2018Q1 to 2019Q4) and check forecasting accuracy. Record RMSE and MAPE.

- Plot the forecasts VS actual values for the test period from Seasonal Naïve, ETS, and ARIMA on the same graph.

Answer for 3,4 and 5

```
> accuracy (snaive(Train2, h=8), Test2)
```

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	17960.6	25976.08	22494.13	3.317122	4.166977	1.000000
Test set	40792.5	42413.11	40792.50	5.592645	5.592645	1.813473

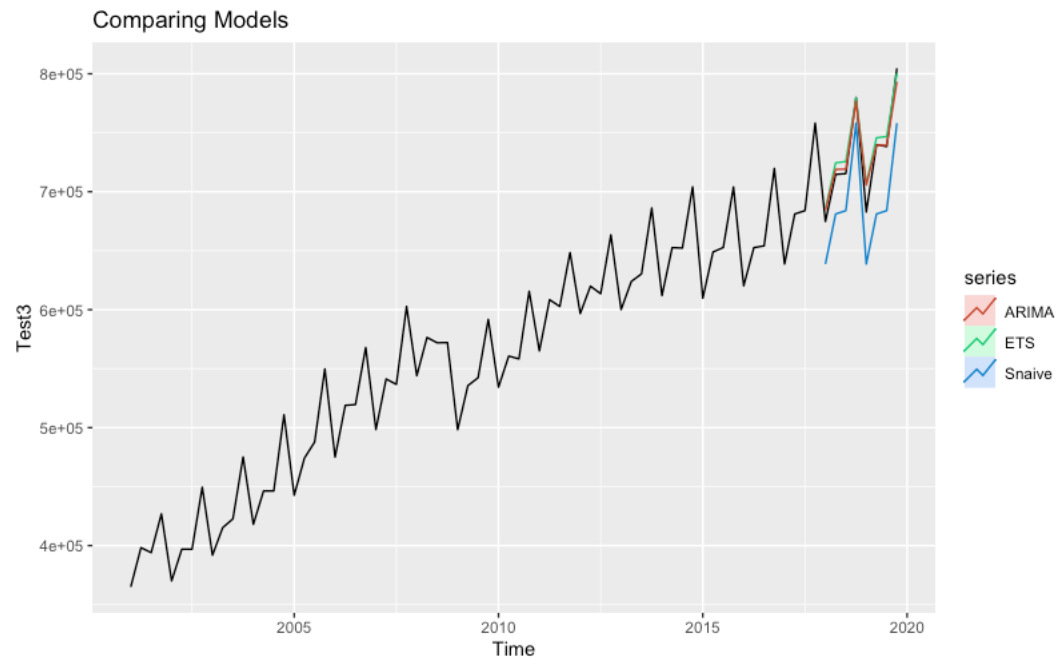
```
> accuracy (forecast (fit5, h=8), Test2)
```

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	-721.0474	10893.24	7669.680	-0.1716944	1.417219	0.3409636
Test set	-8017.2517	11175.05	9162.315	-1.1524165	1.295161	0.4073202

```
> accuracy (forecast (fit6, h=8), Test2)
```

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	218.5748	10279.827	7107.849	0.007238856	1.283993	0.3159868
Test set	-3300.2184	9891.629	7058.753	-0.516517506	0.988055	0.3138042

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2018 Q1	684092.4	670059.7	698125.0	662631.3	705553.5
2018 Q2	718919.1	701006.6	736831.5	691524.4	746313.7
2018 Q3	719121.2	699148.9	739093.4	688576.3	749666.1
2018 Q4	776720.5	755557.2	797883.7	744354.1	809086.9
2019 Q1	705635.0	682035.5	729234.4	669542.7	741727.2
2019 Q2	739319.7	714306.4	764332.9	701065.2	777574.2
2019 Q3	739521.3	713660.4	765382.2	699970.5	779072.1
2019 Q4	793254.7	766880.2	819629.2	752918.4	833591.0



6. Based on RMSE and MAPE, among the 5 models (Regression, Excel Built-in Model, Seasonal Naïve, ETS, and ARIMA), which model performs the best?

Arima model is the best according to RMSE and MAPE.