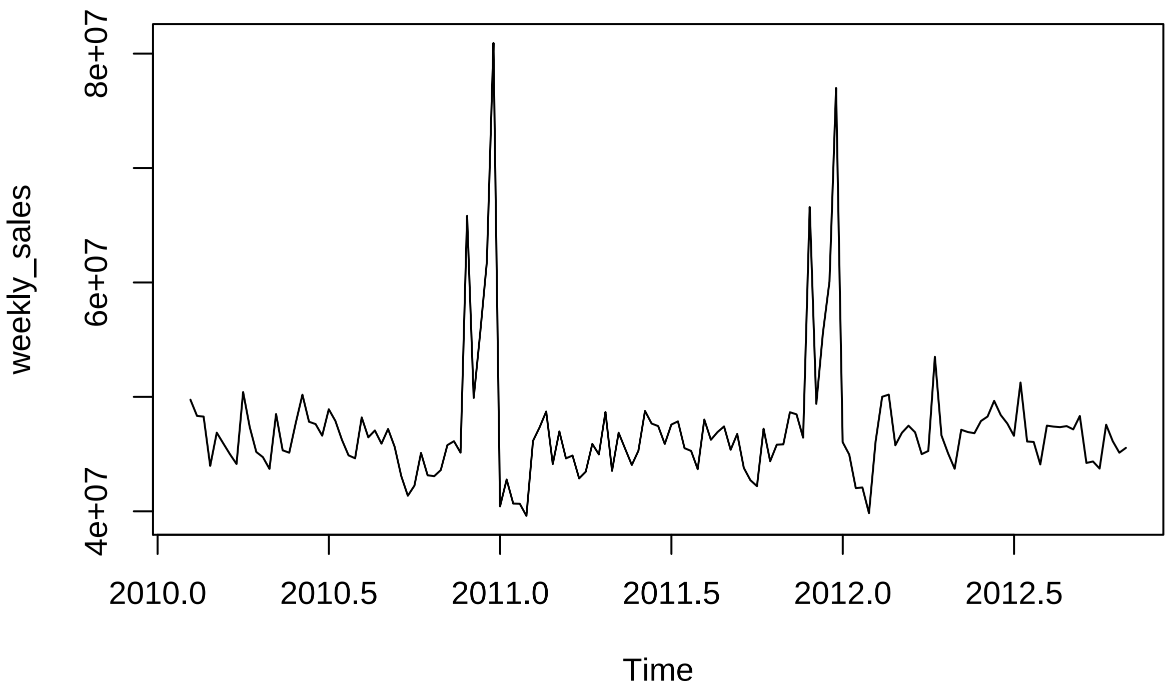
Sales

Oscar

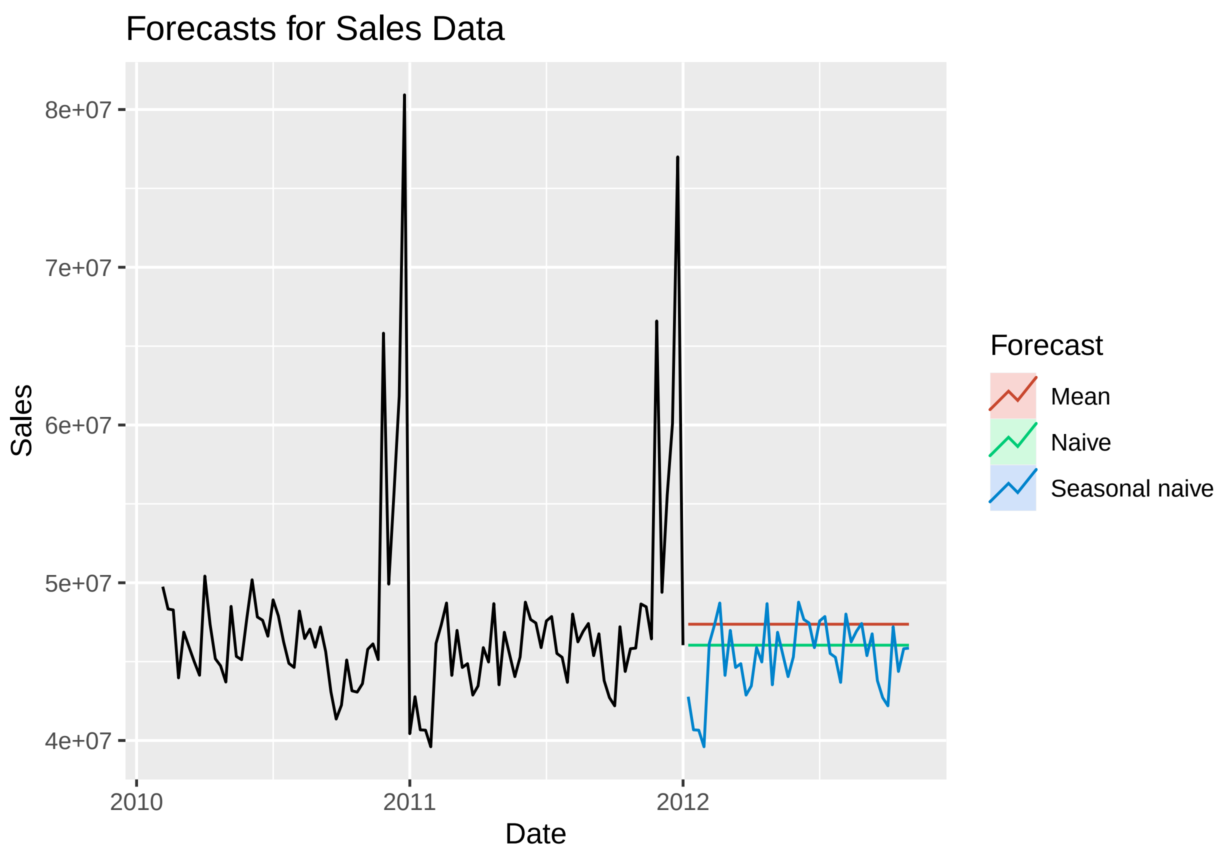
2/6/2021

Time Series Analysis We will forecast our sales by using weekly data from the data set that we used in our python project. This will give us about 143 observations with which to forecast. First here is a plot of our weekly sales from 2010-02-05 to 2012-10-26.



As you can see our sales seem to spike around the end of the year, most likely because of the holiday sales boom that retail stores typically get.

Before we do any forecasting we need to decide on some baseline models to use as a benchmark. This will allow us to compare seemingly simpler forecasting models with more complicated models. We will use the mean, naive, and seasonal naive as our baseline models. Down below we plot them.

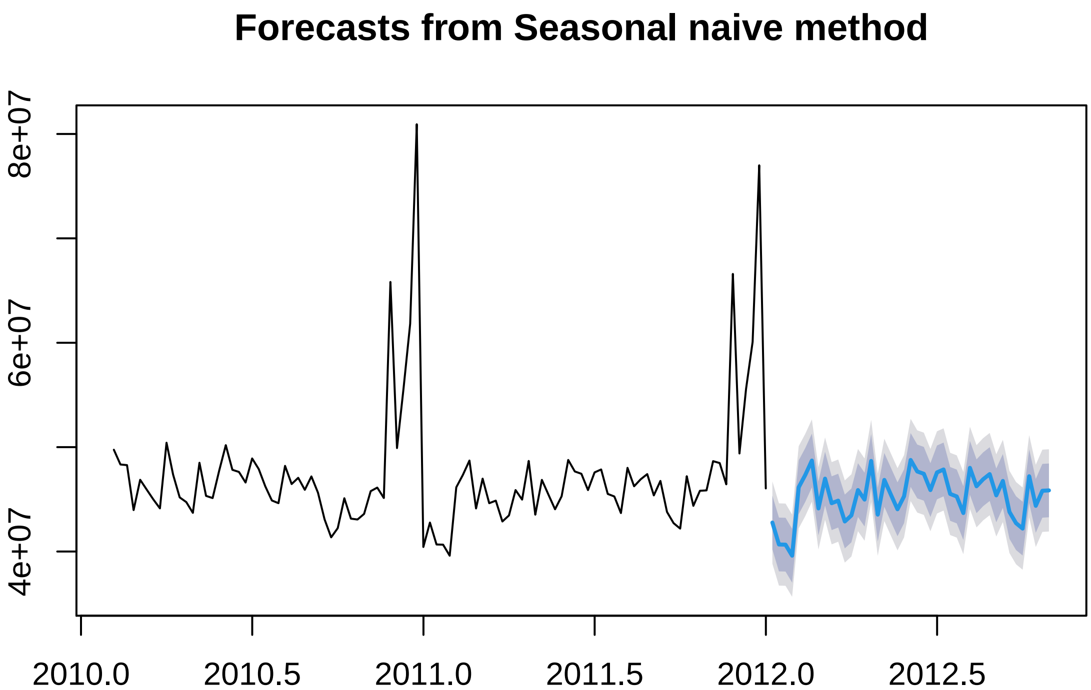


We first partition our data into train and test splits. The training data is what we use to train our model on. While the testing data is how we evaluate our models performance against data that it hasn't been trained on. Up above the *mean* forecast is merely the historical average of the data. The *naive* forecast is simply the value of the last observation. While the *seasonal naive* forecast is equal to the last observed value from the same period in the previous year.

Since our data is highly seasonal we are most concerned with the seasonal naive method. Below we show the accuracy metrics for the seasonal naive method.

```
##                    ME    RMSE     MAE       MPE     MAPE     MASE       ACF1
## Training set  -91460 2009836 1387984 -0.160249 2.917466 1.000000 0.19892562
## Test set     1166821 1985905 1483137  2.437898 3.129202 1.068555 0.03856437
##              Theil's U
## Training set        NA
## Test set     0.7047942
```

As you can see our mean absolute error (MAE) is 1,387,984 for the training set and 1,483,137 for the test set. This means that on average with our testing data we are off by about 1,483,137 from the true value. There's a couple of things to keep in mind when hearing this forecast. First, we only have about 2 years or 104 weeks of weekly data to forecast off of. Ideally we would want several years worth of data to give effective accurate forecasts. This is to ensure that we have enough data to capture and deal with periodic effects such as seasonality. Second, we are forecasting quite far out into the future (about 43 weeks). If we were forecasting not so far out say only a couple of weeks our forecasts and thus MAE would be more accurate.



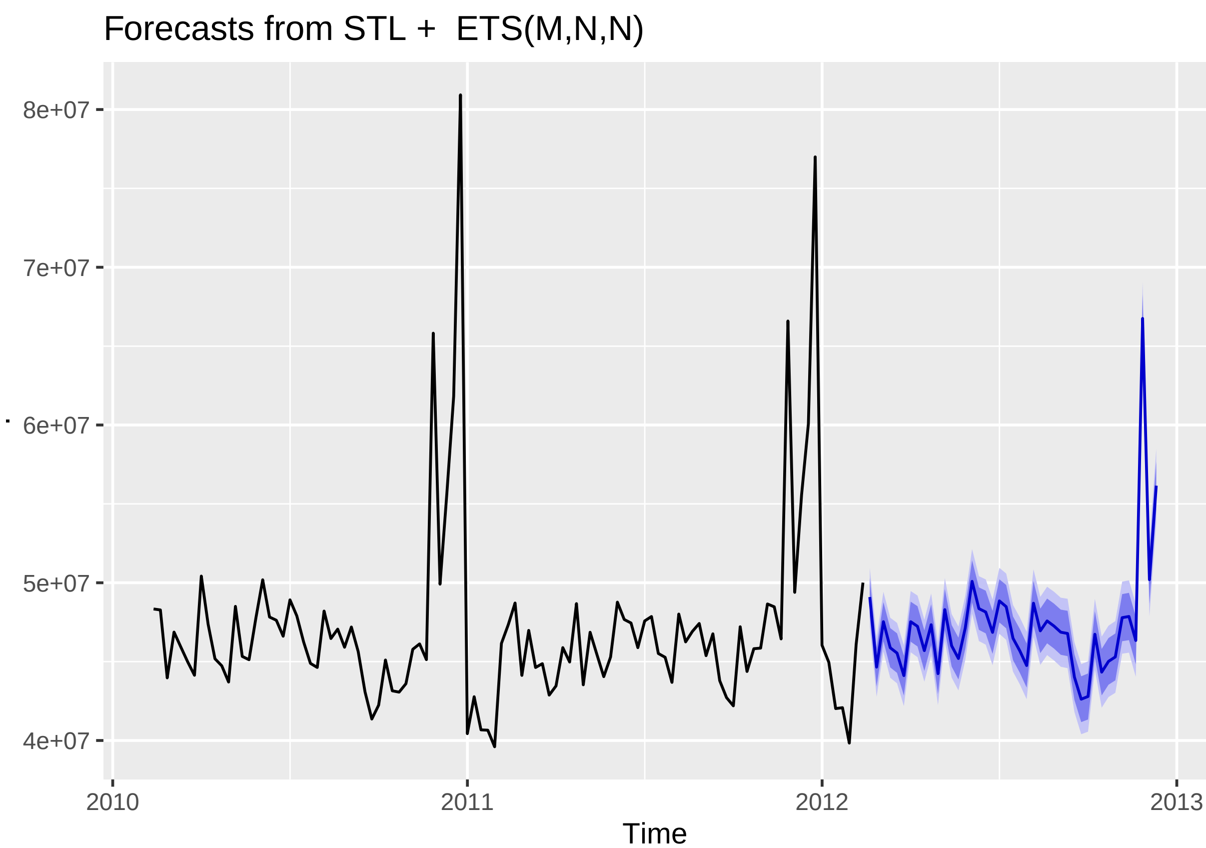
The forecast you see above shows the prediction intervals. The 95% prediction interval is shown in light gray while the 80% interval is shown in dark gray. Below I show our forecast numbers which include the date, point estimate, 80% and 95% prediction interval for all 43 weeks of our forecast.

```
##          Date Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 1  2012-01-07       42775788 40200079 45351497 38836581 46714995
## 2  2012-01-14       40673678 38097969 43249387 36734471 44612885
## 3  2012-01-21       40654648 38078939 43230357 36715441 44593855
## 4  2012-01-29       39599853 37024144 42175562 35660646 43539060
## 5  2012-02-05       46153111 43577402 48728820 42213904 50092318
## 6  2012-02-12       47336193 44760484 49911902 43396986 51275400
## 7  2012-02-19       48716164 46140455 51291873 44776957 52655371
## 8  2012-02-26       44125860 41550151 46701569 40186653 48065067
## 9  2012-03-04       46980604 44404895 49556313 43041397 50919811
## 10 2012-03-11       44627319 42051610 47203028 40688113 48566526
## 11 2012-03-18       44872326 42296617 47448035 40933119 48811533
## 12 2012-03-25       42876199 40300490 45451908 38936992 46815406
## 13 2012-04-01       43458991 40883282 46034700 39519784 47398198
## 14 2012-04-08       45887467 43311758 48463176 41948260 49826674
## 15 2012-04-15       44973328 42397619 47549037 41034121 48912535
## 16 2012-04-22       48676692 46100983 51252401 44737485 52615899
## 17 2012-04-29       43530033 40954324 46105742 39590826 47469240
## 18 2012-05-06       46861958 44286249 49437667 42922751 50801165
## 19 2012-05-13       45446145 42870436 48021854 41506938 49385352
## 20 2012-05-20       44046598 41470889 46622307 40107391 47985805
## 21 2012-05-27       45293457 42717748 47869166 41354250 49232664
## 22 2012-06-03       48771994 46196285 51347703 44832787 52711201
```

```
## 23 2012-06-10        47669735 45094026 50245444 43730528 51608941
## 24 2012-06-17        47447562 44871853 50023271 43508355 51386769
## 25 2012-06-25        45884095 43308386 48459803 41944888 49823301
## 26 2012-07-02        47578520 45002811 50154228 43639313 51517726
## 27 2012-07-08        47859264 45283555 50434973 43920057 51798471
## 28 2012-07-15        45515930 42940221 48091639 41576723 49455137
## 29 2012-07-22        45274411 42698702 47850120 41335205 49213618
## 30 2012-07-30        43683274 41107565 46258983 39744067 47622481
## 31 2012-08-06        48015467 45439758 50591176 44076260 51954674
## 32 2012-08-13        46249569 43673860 48825278 42310362 50188776
## 33 2012-08-20        46917348 44341639 49493057 42978141 50856554
## 34 2012-08-27        47416948 44841240 49992657 43477742 51356155
## 35 2012-09-03        45376623 42800914 47952332 41437416 49315830
## 36 2012-09-10        46763228 44187519 49338936 42824021 50702434
## 37 2012-09-17        43793960 41218251 46369669 39854753 47733167
## 38 2012-09-24        42718097 40142388 45293806 38778890 46657304
## 39 2012-10-01        42195831 39620122 44771540 38256624 46135038
## 40 2012-10-08        47211688 44635979 49787397 43272482 51150895
## 41 2012-10-15        44374820 41799111 46950529 40435613 48314027
## 42 2012-10-22        45818953 43243245 48394662 41879747 49758160
## 43 2012-10-29        45855821 43280112 48431530 41916614 49795028
```

STL Decomposition and Exponential Smoothing

Time series decomposition simply means too break up time series date into different components. One such method is called "Seasonal and Trend decomposition using Loess" or STL for short. We can then use this decomposition to help us forecast values by combining it with another forecasting method. We will now forecast our sales data using STL decomposition along with exponential smoothing (ETS).




```
##                       ME      RMSE       MAE         MPE     MAPE      MASE
## Training set    6214.256  931633.6  653161.3 -0.02322703 1.369993 0.4878678
## Test set      459979.936 1606675.9 1185812.3  0.91312824 2.483784 0.8857225
##                     ACF1 Theil's U
## Training set   0.1389597        NA
## Test set      -0.2374639 0.6196751
```

As you can see we improved upon our previous seasonal naive model. Our mean absolute error using STL and ETS for the test set is 1,185,812.3. We beat our previous error for the seasonal naive model by about 297,324.7. The MASE error on the far right also confirms this. A MASE below one means that the current model is better than the average naive model.

It appears that our Exponential Smoothing model beat out our previous baseline model. We can now use this model for forecasting any future sales. Below we list our forecasting for the future 43 weeks using our new model.

```
##          Date Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 1  2012-02-19       49092492 47872280 50312705 47226338 50958647
## 2  2012-02-26       44651828 43423466 45880190 42773211 46530445
## 3  2012-03-04       47532516 46296059 48768973 45641518 49423514
## 4  2012-03-11       45879974 44635474 47124474 43976675 47783272
## 5  2012-03-18       45540412 44287921 46792903 43624892 47455932
## 6  2012-03-25       44112905 42852473 45373336 42185241 46040569
## 7  2012-04-01       47531408 46263086 48799731 45591676 49471140
## 8  2012-04-08       47237506 45961341 48513671 45285780 49189231
## 9  2012-04-15       45696073 44412114 46980032 43732427 47659718
```

```
## 10 2012-04-22        47339126 46047420 48630832 45363632 49314620
## 11 2012-04-29        44238330 42938923 45537738 42251058 46225602
## 12 2012-05-06        48296739 46989676 49603802 46297758 50295719
## 13 2012-05-13        46006856 44692182 47321530 43996235 48017477
## 14 2012-05-20        45196446 43874204 46518688 43174252 47218641
## 15 2012-05-27        47132226 45802460 48461993 45098524 49165928
## 16 2012-06-03        50089153 48751905 51426402 48044008 52134298
## 17 2012-06-10        48359902 47015213 49704591 46303377 50416426
## 18 2012-06-17        48144996 46792907 49497085 46077155 50212838
## 19 2012-06-25        46852653 45493204 48212101 44773556 48931749
## 20 2012-07-02        48849763 47482995 50216531 46759471 50940054
## 21 2012-07-08        48483426 47109377 49857475 46381999 50584852
## 22 2012-07-15        46479299 45098008 47860591 44366796 48591802
## 23 2012-07-22        45682639 44294143 47071135 43559117 47806161
## 24 2012-07-30        44750938 43355275 46146602 42616455 46885422
## 25 2012-08-06        48704738 47301943 50107533 46559348 50850128
## 26 2012-08-13        46949159 45539269 48359049 44792919 49105400
## 27 2012-08-20        47579311 46162362 48996261 45412274 49746349
## 28 2012-08-27        47256614 45832640 48680588 45078833 49434395
## 29 2012-09-03        46864502 45433538 48295466 44676031 49052973
## 30 2012-09-10        46784927 45347006 48222847 44585817 48984036
## 31 2012-09-17        44018809 42573966 45463652 41809112 46228506
## 32 2012-09-24        42618819 41167086 44070552 40398585 44839053
## 33 2012-10-01        42790280 41331689 44248870 40559558 45021001
## 34 2012-10-08        46734047 45268632 48199463 44492888 48975207
## 35 2012-10-15        44332587 42860377 45804796 42081037 46584136
## 36 2012-10-22        45015363 43536392 46494335 42753471 47277255
## 37 2012-10-29        45297351 43811648 46783055 43025164 47569538
## 38 2012-11-05        47786159 46293755 49278564 45503724 50068595
## 39 2012-11-12        47861767 46362691 49360843 45569128 50154406
## 40 2012-11-19        46337613 44831895 47843331 44034816 48640410
## 41 2012-11-26        66757434 65245103 68269765 64444524 69070344
## 42 2012-12-03        50192111 48673196 51711025 47869131 52515090
## 43 2012-12-10        56154396 54628926 57679866 53821390 58487402
```