DSC 425 FINAL PROJECT - STORE ITEM SALES PREDICTION

Team: Andrew Relstab, Akbar Aidarov, Yizi (Winnie) Huang, Patrick Maguire

Table of Contents

Store Item Demand Forecasting - Non-Technical Summary

Our group analyzed a set item sales data from an unnamed retail location. This data set provided five years (2013-2015) of daily sales totals for 50 items at 10 different stores. Before we began our work, we aggregated the data across the 10 stores, which gave us a measure of the "corporate" sales for each item. We then derived two additional, related sets—sales by week and sales by day. When we began our analysis, we isolated single items to allow us to compare individual item behavior. This approach reduced our working data set and focused our analysis.

Our best-performing model takes several important characteristics of the data into account in order to make accurate predictions. First, our model looks at the actual previous values of the time series. For instance, if the last three values of our data were [2, 1, 0], then our model would probably predict a low number around 1; if the last three values of our data were [31, 50, 44], our model would predict a value between 30 and 50. Our model also takes into account the behavior of the "shocks" in the data—that is, the value of the difference between two consecutive values, as well as the likelihood that a given shock is small or large.

A key component of our data is strong seasonal trends, at different time scales. Interestingly, the data shows a "trend within a trend," with weekly sales patterns mirroring, on a smaller scale, the overall yearly trend. Due to this fact, we incorporated techniques into our analysis that allows us to model seasonality: cycles of high and low sales that are very evident in the data. We also ended up building two primary models that capture distinct seasonal sales behavior apparent in our data. The first model is built on data aggregated by the week, which allows us to model month- and year-long trends. However, the model smooths out oscillations in daily sales. The second model has daily-aggregated data, which is useful for short-term forecasting, but here we lose the ability to track the clear 12-month, wave-like seasonal pattern.

We produced forecasts on both models, choosing the weekly-aggregated model to make predictions months into the future, and the daily-aggregated model to forecast the sales for only the next few days. Our models performed well: the weekly model predicts with accuracy of +/- 2.83% and the daily model predicts with accuracy of +/- 10.3%.

We can imagine a situation where these models prove useful to an organization desiring more information about item-by-item sales, in both the short and long term. Ultimately, each approach has its own advantages and disadvantages, but taken together, these models represent a potent business asset for any entity who wish to model and predict the sales of a given item.

Store Item Demand Forecasting - Technical Summary

INTRODUCTION - Statement of the Problem:

As a retailer it is important to align inventories and work with suppliers to ensure that the correct items are in our stores to serve customers. The problem, stated generally, is to forecast sales for an item with large seasonal swings in sales, as a means to provide seamless availability for our customers while simultaneously minimizing on-hand inventory. Organizations rely on analytics and time series forecasting techniques to predict sales behavior, serve customers efficiently, and produce positive effects on the bottom line. We will try to replicate such work in this project.

RESEARCH STRATEGIES AND APPROACH

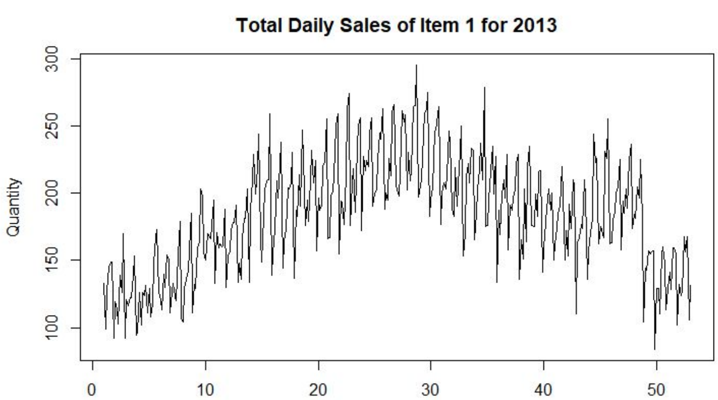
As a means to solve this problem, our team employed several techniques. One involved working to understand the seasonal pattern of sales for the item we are trying to forecast. Another examined how the differences in sales could be modeled to better model the effects of the volatility of the item to make a better prediction. We also asked ourselves if items with highly-correlated sales could in fact be used to predict the sales of one another.

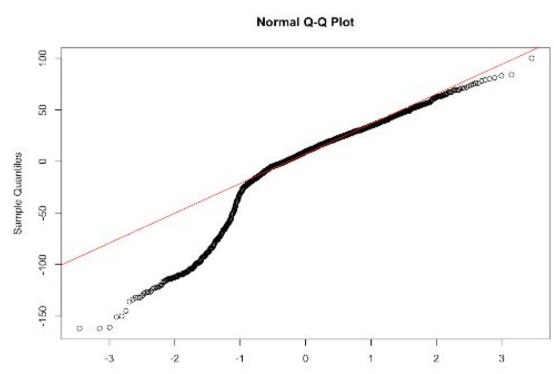
This data set used in this project describes the number of daily sales for 50 different common grocery store items across 10 different stores for 5 years. The whole dataset contains more than 900,000 individual records. This data is already very clean, with four components (date, item number, store number, number of sales) that are fully populated with no null values. While the cleanliness of the data is a plus, one disadvantage of this data set is that the items themselves are not described at all, so we can't make "real world" inferences about product category popularity, sale patterns, etc.

Due to the size of the data set, we separated the data into smaller sets representing single items. This allowed us to focus our attention to modeling sales at the item-level (rather than at the store level). Then, using seasonal ARIMA models as our primary technique, we decided to build two variations of a model for single item: one to model short-term behavior (no longer than 1 week out), and one to model the long term trend and cycle of the item sales over several years.

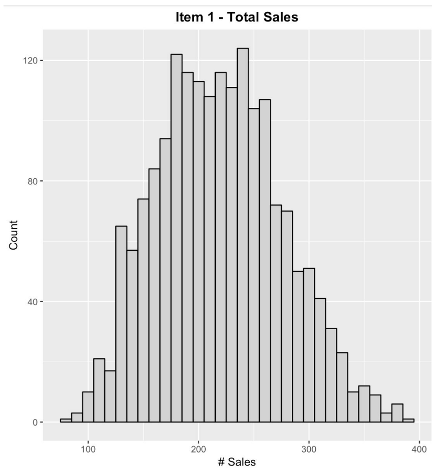
EXPLORATORY ANALYSIS OF DATA

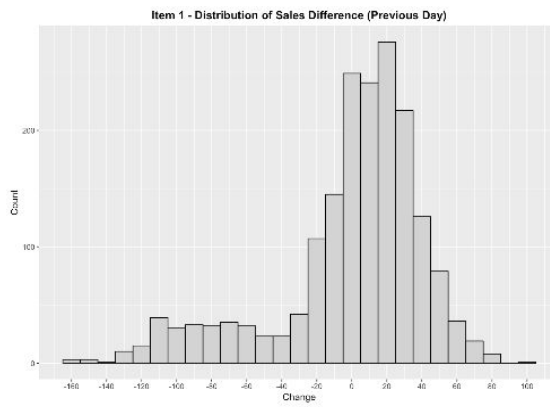
Our dataset consists of daily sales of 50 items across all stores since January 1, 2013. We aggregated Item 1 sales data in two ways - total daily sales of all stores for the first year and total weekly sales of all stores for all years.





As part of continuing the exploration phase of the analysis our group also examined the differences of sales to gain a better understanding on the net effects of the previous period on future sales. Ais the distribution of the differences of sales change of item 1. Below is the histogram of t he sales difference of item 1 and the Q-Q plot it would appear that the distribution is skewed to the right.

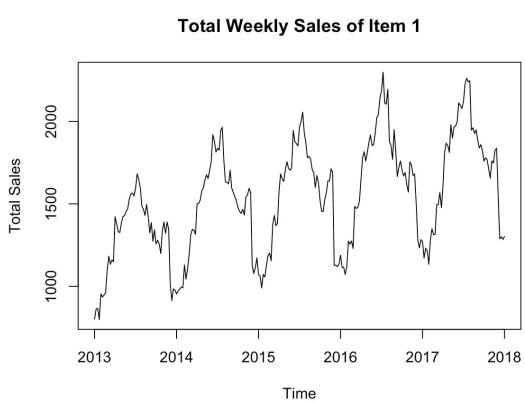


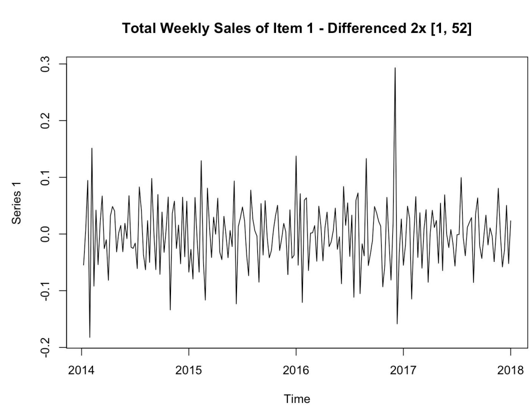


SEASONALITY

Initial visualization and exploration of the item sales data provided strong evidence that we would need to incorporate some techniques for dealing with the heavily seasonal component. All the items (item 1, item 2, etc.) have sales profiles that are clearly seasonal, with spikes and lulls in sales occurring at distinct points throughout the calendar year. The data shows very clear seasonal trends, both on a weekly and yearly scale.

We used residual analysis to identify the appropriate lags at which to difference the data. On the weekly-aggregated item 1 data, we differenced twice: at 1 and 52. The resultant set was much closer to stationarity than the untreated data.





SARIMA

As mentioned in the Research Strategies and Approach section, our approach included building two related SARIMA models: one that is focused on short-term predictions, and one that is meant for long-term predictions. Such models would be applicable at different levels of the corporate structure and would be easily reproduced and customized for different items in the store. Because the model for short-term predictions was based on aggregations at the daily level, we will call this the daily model. The model for long-term predictions was based on aggregations at the weekly level; we will call this the weekly model.

Despite its obvious seasonality, our team began our analysis by running the auto.arima function in R on our item 1 data. This provided a performance benchmark for our future models and gave us an idea of what the final order of the model would be. Once we had performed this work, we moved into fine-tuning our SARIMA in an attempt to match or exceed the performance mark set by auto.arima.

Daily Model - forecast several days ahead

We experimented with different combinations of AR and MA coefficients in the model-building iteration for daily-aggregated sales data. We were able to produce SARIMA models that rivaled auto.arima with regard to the information criteria scores, but not in the simplicity of the model order. For that reason, we decided that the following SARIMA was our best model for near-term forecasting. The Daily Model is built on item 1 sales aggregated on a daily level for 2013 (the details of deriving this model and its performance comparison can be found in Appendix 4 as well as in R code file 'Akbar_Daily.R'). The coefficients and the intercept of this model are all significant:

ARIMA: order (1,0,1), seasonal order (2,0,0), period = 7, include mean = True

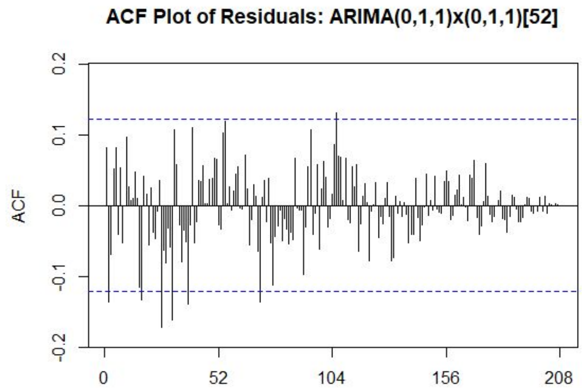
Weekly Model - forecasting long-term sales

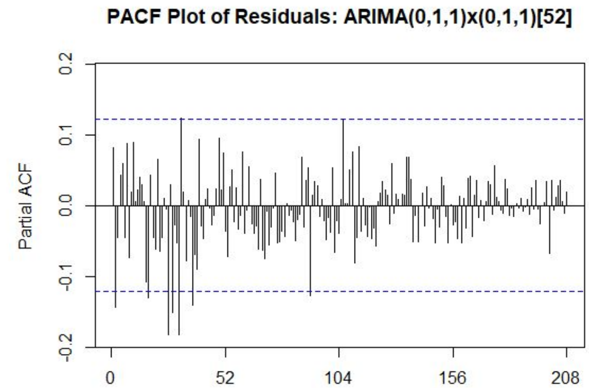
For a different look at the data, we aggregated on a weekly level in an attempt to recreate the cyclical behavior of item sales throughout the year. This Weekly Model is less accurate in predicting short-term variation in sales, but it does quite accurately model the shape and slight upward trend of the data at the scale of months and years. The Weekly Model is built as following with all of its coefficients tested as significant (see details in Appendix 4 as well as in R code file 'Akbar - Weekly Models.R'):

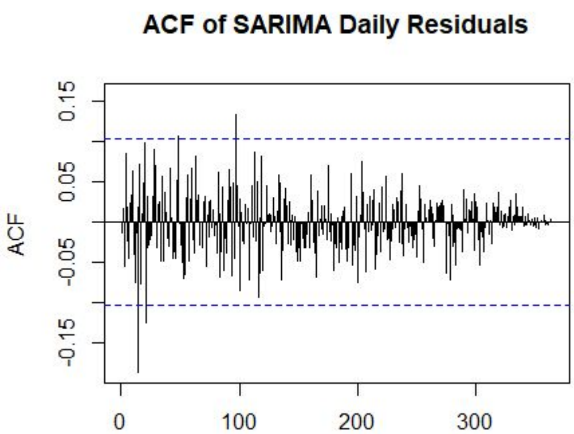
ARIMA: order (0,1,1), seasonal order (0,1,1), period = 52

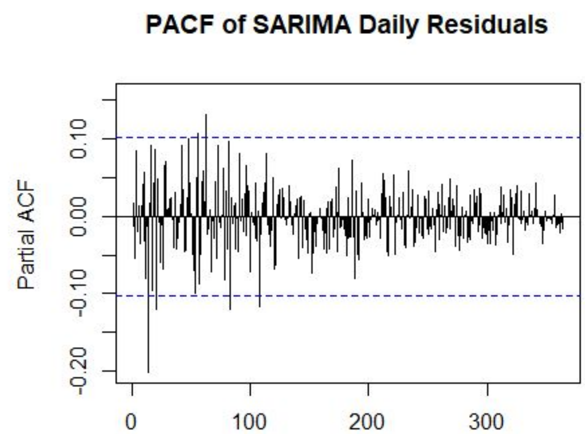
Residual Analysis and Model Diagnostic

The ACF and PACF plots of both Weekly and Daily models show predominantly values that are not significantly different from zero, while the rest of them show small ACF and come from large random lags. In other words, the residuals appear stationary. The Ljung-Box tests for both model residuals showed p-values > 0.05 meaning that the residuals are uncorrelated. The Jarque-Bera Normality tests for both model residuals show p-values < 0.05 suggesting that the residuals are not normally distributed. Thus, we can't claim that the residuals are white noise.



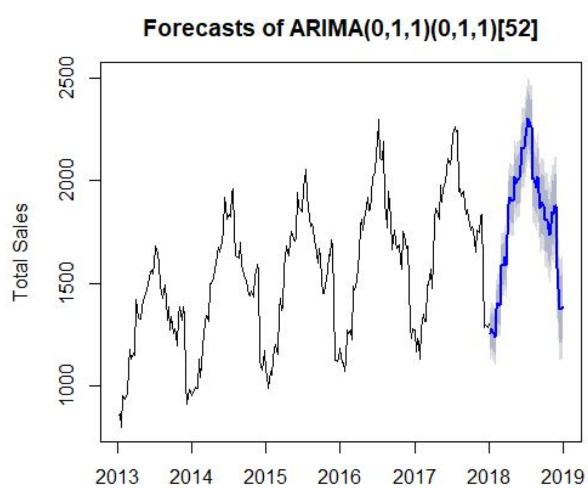


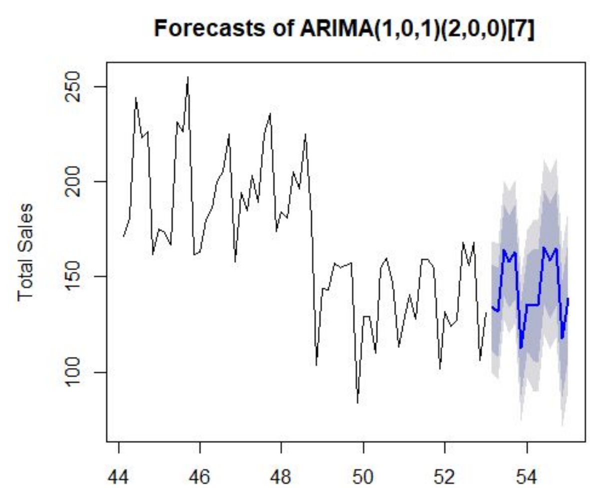




Forecast Analysis

In the plots below we can see that our Daily Model forecasts follow the same basic pattern as shown in the training set (zoomed in on the right); we can make predictions up to 2 weeks out (as shown here) before the confidence interval expands too far.



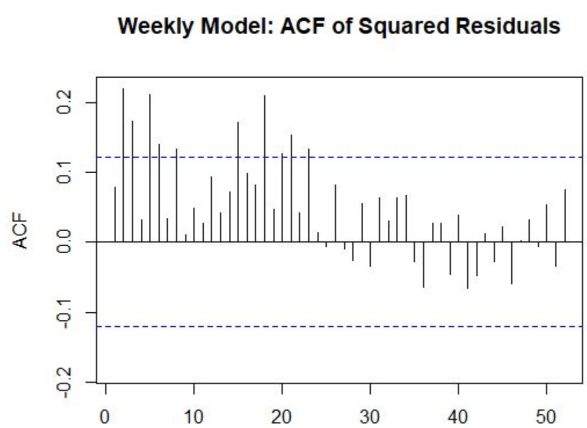


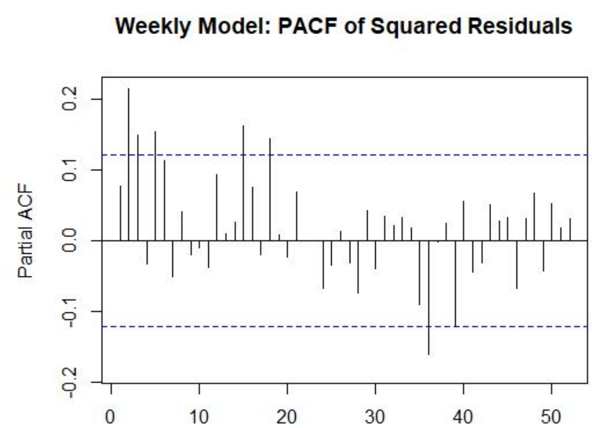
On the left, we see the forecasts of our Weekly Model providing a "birds-eye" view of the sales trends of Item 1. We can reasonably predict the rise and fall of sales through the course of the year, as indicated by the tight confidence intervals. This would be useful for a corporate analysis of all stores' yearly sales and how they help define sales numbers in the year to come.

Investigating GARCH effects

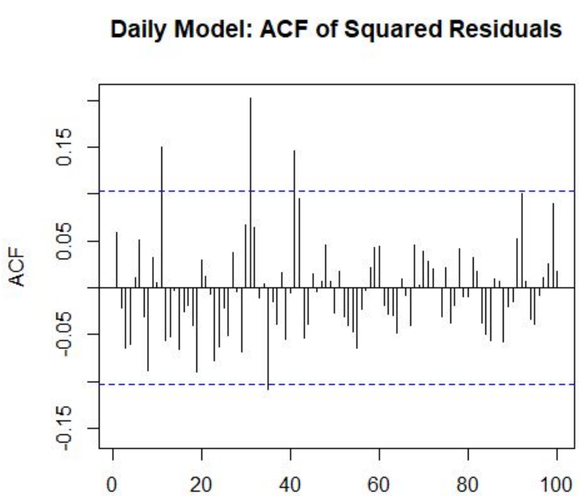
The group investigated GARCH effects in both of our final SARIMA models in an effort to find effects of volatility in the data that we could usefully model. We saw in the ACF and PACF plots of the squared residuals of our Weekly Model below that there was still an evident visual pattern, as well as several values that spiked above the confidence line. Ljung-box tests on the squared residuals yielded significant values lag-4, suggesting that autocorrelation exists for in the residuals.

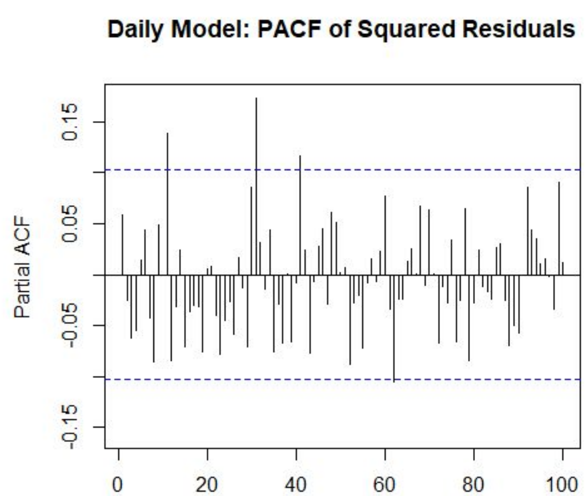
Weekly data:





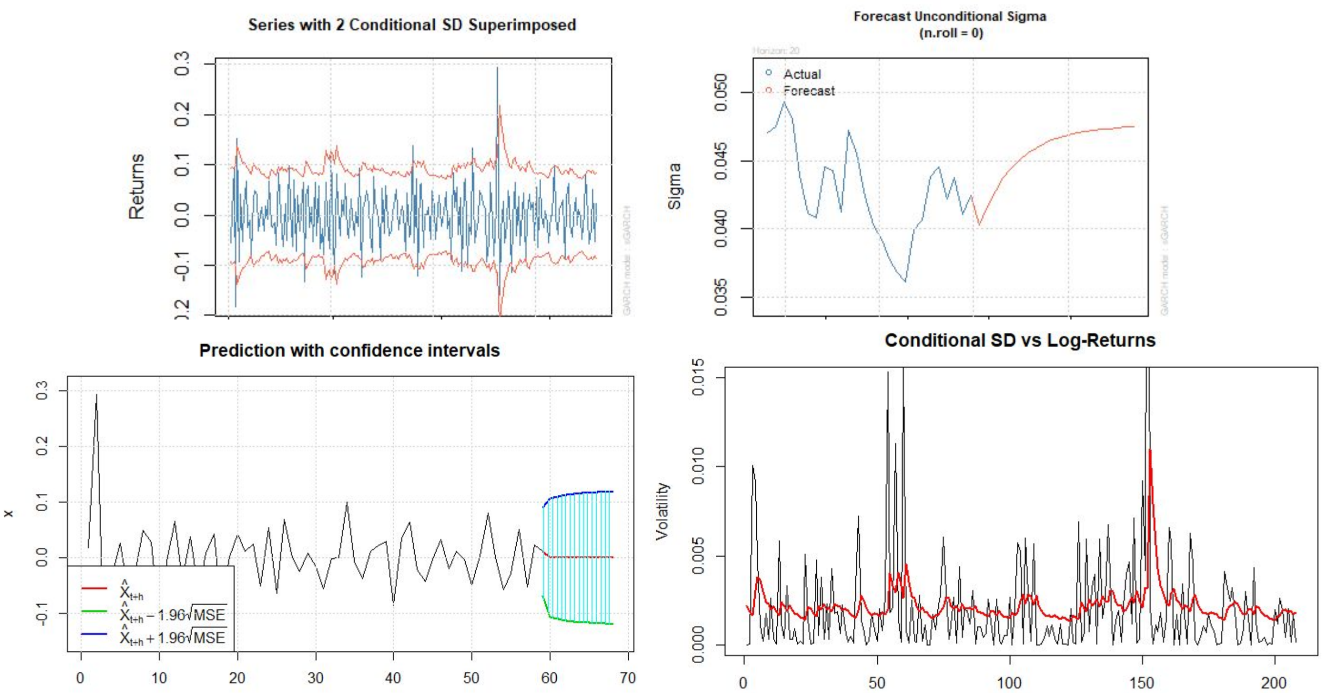
As for the Daily Model, Ljung-Box test suggested the the squared residuals are not autocorrelated.





Forecast Analysis:

We fit GARCH(1,1) on the Weekly Model using ARMA(0,1) and de-seasonalized weekly log-returns series. The GARCH model produced uncorrelated residuals (L-B test p-value > 0.05) that are not distributed normally (J-B Normality test p-value < 0.001). We ran forecasts and backtesting (see details in Appendix 4 and in R code file ‘Akbar - Weekly Models.R’), which yielded 0.03754 RMSE and 1.212243 Symmetric Mean Absolute Error. You can see from the plots below that the GARCH is capturing and forecasting volatility pretty well.



<u>**TIME SERIES REGRESSION**</u>

We made use of the R’s cross-correlation and prewhiten functions to determine if there was any useful correlation between the sales pattern of different items.

We tested the first 10 items against are selected to test correlations and prewhitting. On weekly basis, the item 3,item 6, and item 8 are included in final regression model with ARMA errors because their values at lag 0 are the largest which are equal or larger than 0.8 among lags after prewhitening and their coefficients are significant. On daily basis, the item 2,item 7, and item 8 are included in final regression model with ARMA errors because their values at lag 0 are the largest which are equal or larger than 0.3 among lags after prewhitening and their coefficients are significant. The first 356 data on daily basis are used to model.

Regression model with ARMA error on weekly basis

Before building the regression model with ARMA error, items 2 through 10 are tested against item 1 using CCF and prewhiten functions, to see whether there is real correlation among any pairs. The item

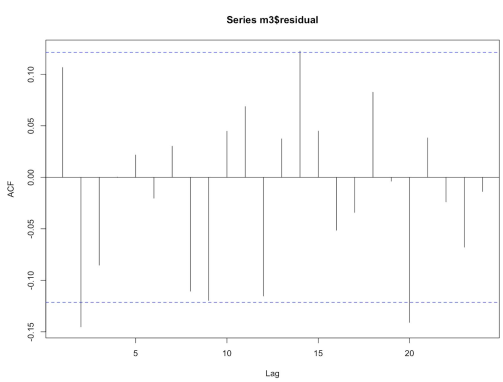
3,item 6, and item 8 are included in final regression model with ARMA errors because their values at lag 0 are the largest which are equal or larger than 0.8 among lags after prewhitening and their coefficients are significant.

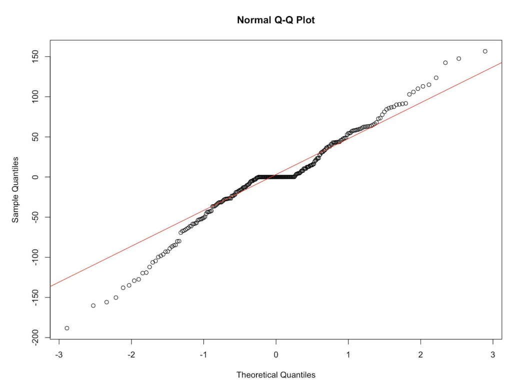
Model

The selected item 3, item 6, item 8 and ma1 are significant with aic = 2319.1

```
m3 = arima(ts(week_item1), order = c(0,1,1), seasonal = list(order=c(0,1,1), period=52),xreg=data.frame(week_item3,week_item6,week_item8),fixed = c(NA,0,NA,NA,NA))

          Estimate Std. Error  z value  Pr(>|z|)

ma1       -0.999937   0.063805 -15.6718 < 2.2e-16 ***

week_item3  0.192092   0.057119   3.3630  0.000771 ***

week_item6  0.110873   0.038228   2.9003  0.003728 **

week_item8  0.105275   0.032194   3.2700  0.001075 **

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual analysis



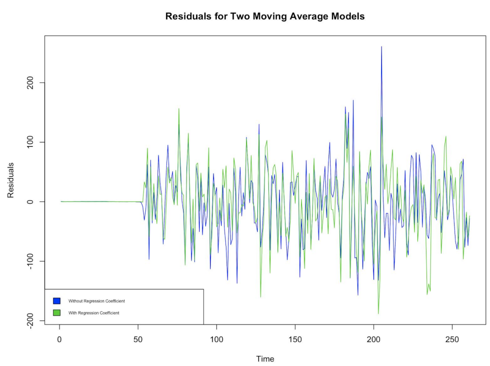


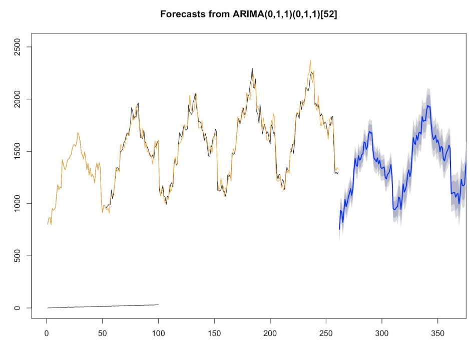
The residuals of Regression model with ARMA errors on weekly basis is not white noise. From ACF, there are three lags not significant showing some autocorrelations. Even though the p-value obtained by Ljung box test is 0.08313 which can not reject null hypothesis so that residual may be white noise, the distribution of residuals is not normally distributed supported by p-value is 0.001249 rejecting null hypothesis obtained by Jarque Bera test and QQ plot.

Back Test & Forecast

The back test with 85% training dataset is apply to test Regression model with ARMA errors on weekly basis. The performance of this model is exactly the same with the model with ARMA model. Both their

RMSE of out-of-sample forecasts are 53.98 and mean absolute error of out-of-sample forecasts of are 47.22. As you can see the residuals plot as below, even though the Regression model with ARMA errors smooth the residual curves that seems to capture more information from residuals, while it does not improve the performance at all. The model is fitted to to predict item1. The yellow line is fitted values by model, it looks to catch most information of original time series and the blue line is forecast which follows the similar pattern. The model is fitted to to predict item 6 as well after that. However, the performance is worse than that of predicting item 1 because both RMSE of out-of-sample forecasts of models with and without autocorrelation errors are 116.0017 and mean absolute error of out-of-sample forecasts of are 91.80834.





CONCLUSION

The goal of our project was to come up with robust ways of predicting short-term and long-term sales. It was quickly discovered that building a single model to account for both levels of seasonality might overcomplicate the model and be mathematically and computationally cumbersome. To this end, we divided our work into two main streamlines with the focus on daily sales and weekly sales of Item 1. The daily sales were forecasted using the daily sales data of the first year, whereas the weekly sales were forecasted using weekly aggregated data for all years. This way, we were able to capture the fluctuations on different scales and trends that are susceptible to weekly and yearly seasonality. Our Weekly and Daily models performed well and proved to be a potential business asset with MAE of 46.62 and 13.5 correspondingly. Throughout the project, many different attempts were made to improve the models, including volatility anticipation via GARCH and external time series incorporation via Time Series Regression. While GARCH was implemented successfully and it helped tame down some volatility, it couldn't address the fact that the residuals were not normally distributed hinting on a more complex behavior of the residuals. Even though there were many other items that were in correlation with Item 1, time series regression did not improve our model significantly and was ruled out as unnecessary complication of the existing SARIMA models.

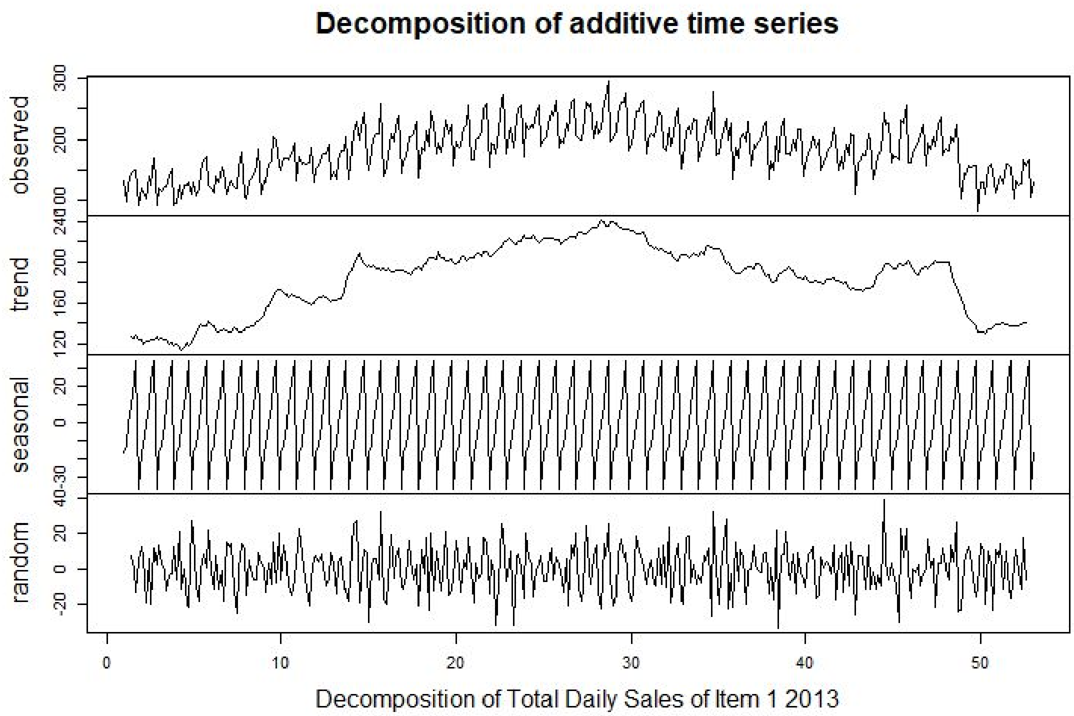
FUTURE WORK

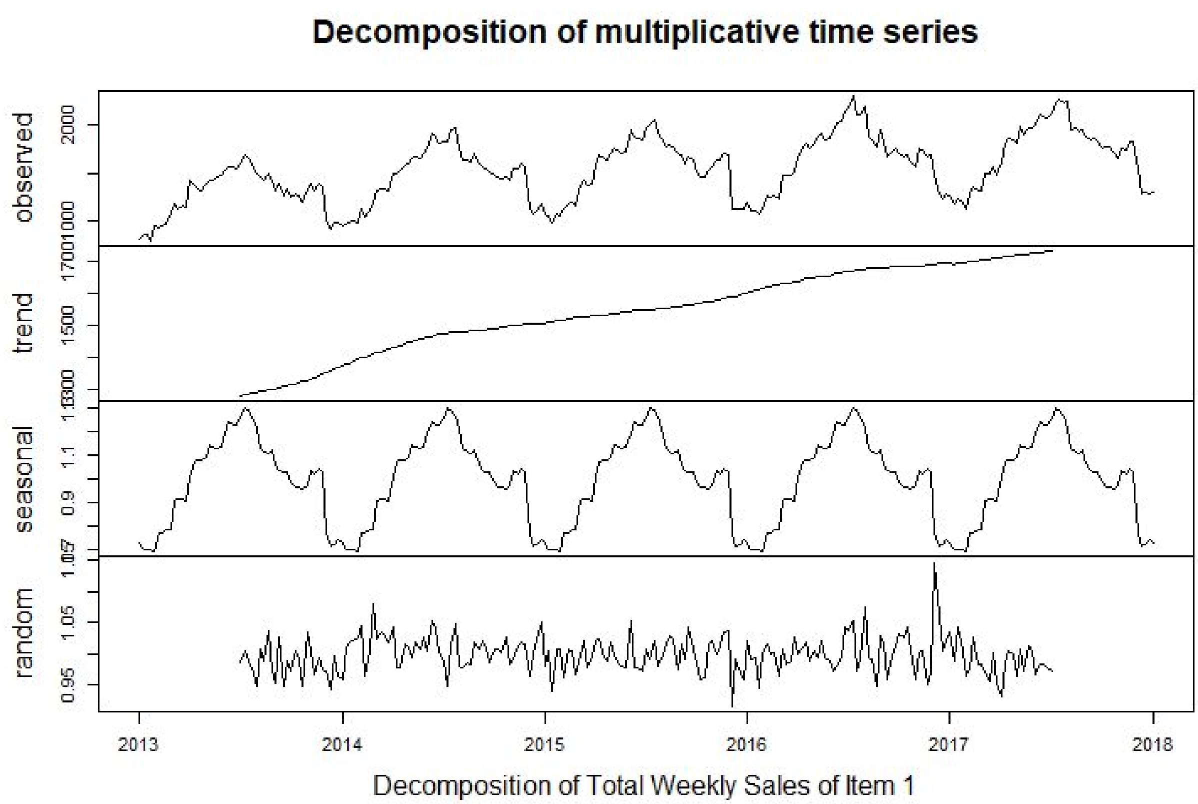
In an effort to continue strengthening the analytical power of our models, we suggest to test the regression methods on the rest of the items in the set as a means to discover value and insight from other items. An alternative approach would include the lag influences of the selected items in Regression model with ARMA error. Besides regression models, we suggest testing intervention models as means of accounting for major inflection points in the trend such as the one around November.

Appendix

Results of the Regression models with the ARMA function

Data	Model name	Significant Parameters	AIC	Ljung box	JB Test	RMSE
Weekly	SARIMA	ma1 sma1	2342.95	0.1895	5.644e-07	53.98
	Regression model with ARMA error	ma1, week_item3, week_item6, week_item8	2319.11	0.08313	0.001249	53.98
Daily	SARIMA	ar1, ma1, sar1, sar2	3143.61	0.7984	0.01796	19.04367
	Regression model with ARMA error	sar1, daily_item2, daily_item7, daily_item8	2935.57	0.7433	0.645	19.04367


Decomposition of additive time series
observed
trend
seasonal
random
Decomposition of Total Daily Sales of Item 1 2013


Decomposition of multiplicative time series
observed
trend
seasonal
random
Decomposition of Total Weekly Sales of Item 1

Individual Report - Yizi(Winnie) Huang

1.Initial data exploratory

At the beginning of the project, the dataset was aggregated on a daily, weekly, monthly basis from aspects of items and stores by me. My data exploratory covers line graphs, box plots, and basic time series plots.

2.Model building

2.1 Try SARIMA model

The SARIMA model is built for daily sales of item 1 across 10 stores. My contribution covers related time series plots, identifying p and q in the SARIMA model, residual analysis, model evaluation, and forecasting.

Model:

sarima(diff(ts_store2),2,0,2,1,0,1,7)) model

The sales of items are non-stationary. It has very strong seasonality pattern in the series and has a yearly pattern and the spikes occur in the middle of each year. The difference between sales and log price also show strong seasonality.

2.2 Try GARCH model

The residuals are tested based on ARIMA model with seasonality and there are few lags showing some autocorrelations. There is a very weak GARCH effect in our dataset. However, in order to improve the performance of the model, three GARCH models are built because the absolute residuals and squared residuals plots show that there is a weak GARCH effect. My contribution covers GARCH model building, residual analysis, model evaluation, and forecasting based on sales of item1, sales of item1 after twice differences, sales of item1 after twice difference with arma model.

Model:

gfit1.ts = garchFit(~ garch(1, 1), data=item1, trace=F)

gfit2.ts = garchFit(~ garch(1, 1), data=diff(diff(item1),7), trace=F)

gfit3.ts =garchFit(~ arma(1, 1) + garch(1, 1), data=diff(diff(item1),7), trace=F)

There are more parameters included in the GARCH model while the performance is getting worse than that of the ARIMA model with seasonality. The GARCH models have larger AIC and larger RMSE than those of ARIMA model with seasonality since they do not take account of the effects of q in seasonality part. To sum up, there are limited GARCH effects on our sales data.

2.3 Regression model with ARMA errors

The basic problem we're considering is the description and modeling of the relationship between two or more than time series. First 10 items are selected to using CCF and prewhiten() to test their real correlation with item1. If their values after prewhiten() are very high which is higher than our threshold, and then those items will be included in my regression model with ARMA errors. Two Regression models with ARMA errors were built on a daily and weekly basis to predict item 1 and their performances are compared with those of ARIMA models correspondingly. What more, two regression models are applied to predict other items. My contribution part covers Regression model with ARMA errors building, residual analysis, model evaluation, and forecasting.

Model:

Weekly basis

m3 = arima(ts(week_item1), order = c(0,1,1), seasonal = list(order=c(0,1,1), period=52),xreg=data.frame(week_item3,week_item6,week_item8),fixed = c(NA,0,NA,NA,NA))

Daily basis

m6 = arima(daily_item1, order = c(1,0,1), seasonal = list(order=c(2,0,0), period=7),xreg=data.frame(daily_item2,daily_item7,daily_item8),fixed=c(0,0,NA,0,0,NA,NA,NA))

The Regression model with ARMA error does not improve the performance based on RMSE and mean absolute error. The models based on weekly data have lower AIC than models based on daily data. The model built for predicting item 1 has a worse performance (RMSE) to predict other items.

3. Other contribution

I also contributed to providing the template, combining parts of ppt slides, and providing the draft of the technical report of my models.

Learn from the project

This project makes me clearer to choose a good model. At first, I think the GARCH model or Regression model with ARMA errors can work better compared with another basic model. However, the residuals of those two kinds of models are more like to be the random walk because fewer information will be captured after that. More information obtained from residual may not make the performance of the model get better. In the real case, the simple model with few parameters with great prediction/forecast is the best.

Individual Report - Patrick Maguire

I contributed to the group final project by working with the other group members to plan our approach to working our data set; perform model-building, analysis, and testing in R; and articulate our findings in the project report.

After we decided on the data set that we wanted to use for our project, the group met over several sessions outside of class to perform initial analysis and model building. During this time, I helped develop the aggregation logic for single item sales in our data and produced initial exploratory visualizations of the distribution of our data and a side-by-side comparison graph of original and differenced data behavior.

At this point we began working to develop the best model possible for item 1 sales behavior aggregated at daily and monthly scales. I produced a SARIMA model and investigated the presence of GARCH effects in the model. I back tested my model and got middling results; it was, however, useful to have a benchmark of what did and did not seem to work when our group met to finalize the models. Working from item 1 data differenced at 1 and 7, the best model.

```
garch_i1 = garchFit(~ arma(1, 1) + garch(1, 0), data=sd, trace=F, cond.dist = c("std"))
```

My GARCH model has a BIC of 9.2 but was ultimately outperformed by a SARIMA model built by a different group member.

I participated in a group brainstorming session after the midterm where our group focused on developing a plan for our analysis with regard to the final project. For the in-class presentation, I contributed my model findings and forecast visualizations to provide some sense of where the analysis of the monthly-aggregated data was heading. I helped produce slides and worked on the formatting the slide deck to increase the professionalism of our presentation.

The group's last in-person meeting occurred soon after we gave our in-class presentation. During this productive session decided on the best model for forecasting short-term behavior in item 1 sales and spent a good deal of fine-tuning the model. We ended up talking through each step in the SARIMA and GARCH analysis to make sure we understood the steps we were taking and why. I contributed to the discussion and helped with the final model order selection and iteration process. We looked for GARCH effects and discussed if our residuals results warranted such an approach.

Lastly, I wrote our project's non-technical summary and contributed heavily to the technical report. I edited the final report along with my teammates and helped standardize formatting and visualizations.

Project Takeaways - Patrick Maguire

I learned a lot about time series analysis from working on this project, both from my own analysis and from discussion with my teammates. To me, the thing that stands out is the recognition and treatment of non-stationary data sets and seasonal behavior. Our data happened to be very strongly seasonal and non-stationary, with both weekly and yearly trends. At first it was difficult for me to keep track of the "order of operations" for testing for and modeling seasonality and stationarity, but this project gave me ample practice in performing both types of work. We had to incorporate differencing and SARIMA techniques in order to accurately model the behavior of the item sales. I realized how important it is to explore various "views" of the data, whether it be by aggregating at different time scales (daily, weekly, etc.), differencing at several lags and choosing the best result, or even simply analyzing 1-year subsets of a multi-year data set. With the regularity of the patterns in our item sales data, it was only through this type of experimentation and exploration that we were able to tease out models that can accurately model the behavior of the item sales.

Individual Report – Akbar Aidarov

1. **Research Design**
   I actively participated in the discussion of the problem statement and different approaches of solution. Detected the two levels of seasonality (yearly and weekly) and pursued closer research of possibly building a single model that incorporates both. After finding it ‘too experimental’, proposed breaking it down to two different levels of aggregation: a) daily sales of the first year across the stores and b) weekly sales for all the years. This way we could analyze the weekly and yearly seasonal effects independently from each other and use such models for short- and long-term predictions respectively. One of the notes for future work made after the models were built is that intervention models might help improve the performance of the short term predictions (daily sales) by accounting for the major inflection points of the general trend. Another approach that was suggested by me was building models for two levels of forecasting: a) ‘managerial’, or store level, forecasting and b) ‘corporate’, or total sales, forecasting. It was decided to pursue only one of these routes (corporate).
2. **Models**
   I experimented extensively with total monthly, weekly, and daily sales data for Item 1. After trying many different models suggested by *ACF, PACF, and EACF plots*, I derived the best SARIMA models with significant coefficients supported by *backtesting at 80%* training size and ‘auto.arima()’ function’s *stepwise* and *brute force selections*:

	Monthly	Weekly	Daily
Model	ARIMA(1,1,0)(1,1,0)[12]	ARIMA(0,1,1)(0,1,1)[52]	ARIMA(1,0,1)(2,0,0)[7]
RMSE	98.18341	55.29384	19.600831
MAE	85.36758	46.61828	13.49969
MAPE	0.010847	0.028252	0.103051
SMAPE	0.010841	0.028065	0.082888
Sigma^2	39022	4230	309.4
AIC	636.04	2344.95	3145.61
BIC	641.59	2354.96	3169
Residuals			
L-B Test p-value	0.1524	0.1895	0.7984
J-B Test p-value	0.6206	< 0.001	0.01796
ADF Test p-value	< 0.01	< 0.01	< 0.01

As there were multiple lags at which the squared residuals of these models were showing autocorrelation, I investigated the GARCH effects on weekly and daily models using de-seasonalized log-returns. I ran residual analysis, forecasting and prediction for these GARCH models. Omegas of both GARCH models were not significant. Alpha-1, Phi-1, and intercept of Daily GARCH model were insignificant as well. Overall, these models have marginally improved our predictions of volatility; however, there is still room for investigating the residuals further:

	Weekly	Daily
GARCH Model	GARCH(1,1) + ARMA(0,1)	GARCH(1,1) + ARMA(1,1)
RMSE	0.03753947	0.1321138
MAE	0.03100983	0.09506996
MAPE	18.48622	Inf
SMAPE	1.212243	1.121034
Log-Likelihood (norm.)	1.663431	0.719713
AIC	-3.288401	-1.405812
BIC	-3.224217	-1.340640
Residuals^2		
L-B Test p-value	0.5252	0.07466
J-B Test p-value	< 0.001	< 0.001

3. **Contribution to the Reports**
   I helped defining the business problem and the value proposition of the models. I contributed extensively to visualizations of forecasts and residual analyses in SARIMA and GARCH models. I formatted 25% of the PowerPoint slides and participated heavily in the writing and formatting of the technical report.
4. **Personal Project Takeaways**
   One of the major takeaways for me was the fact that I learned how to approach the same dataset from different perspectives and levels. Working with a real world data with no labels had its pros and cons – while it'd be easier to have a contextual idea about the store items, it was a great exercise of abstract thinking and pattern finding. Aggregating sales on different levels showed me how 'corporate' level of decision making can be different from 'managerial' level decision making. The former is more strategic and is aimed at making longer term forecasts, whereas the latter is more short-term oriented and is more susceptible to shocks and innovations (and same goes to forecasting total daily sales on the corporate level). Breaking the data down in these ways helped us deal with two levels of seasonality, and that was another major lesson for me. It would be very interesting to see if it is mathematically possible and computationally feasible to combine multiple levels of seasonality in one model. Another lesson for me was to be cautious about auto.arima() function and compare its model performance with your model. This will give you a better idea of why auto.arima produce such model. Backtesting and checking the results might reveal that your model performs better, given that you are ready to add complexity for the marginal difference and risk overfitting.

Andrew Relstab individual Section

My contributions to the team was in two main areas, firstly organization of the team and the route we as a group are going as well as the completion of the written section of the report. With regards to the organization portion of the project I worked very closely with the team to make sure that we all were on the same page on the key deliverables. In terms of the contribution of the analysis, the other team members contributed the code my purpose in the project was to combine everything into the final paper. I also help contribute to the seasonality modeling as well as provided insight in the GARCH model creation. Also the editing of the paper was a large section of my contribution.

The key takeaway for me is that time series analysis is such a broad topic it is much more than just fitting a line through some data. This class also taught me to model the shocks and understand the variability of the residuals and a attempt to reduce the volatility.