Billing, Units Sold and New Deals 4Q13 estimation

File: 4Q2013 Estimation.xlsx

	Q4 2013			Q4 total
	Oct-13	**Nov-13**	**Dec-13**	
Billings ($ million)				
Local	$140.6	$148.9	$165.3	$454.9
Goods	$84.9	$104.8	$92.6	$282.2
Travel	$23.9	$21.4	$25.2	$70.6
Total	**$249.4**	**$275.1**	**$283.2**	**$807.7**

	Q4 2013			
	Oct-13	**Nov-13**	**Dec-13**	
Units Sold				
Local	4,957,175	4,804,672	5,315,744	15,077,591
Goods	3,096,614	3,914,936	3,408,196	10,419,746
Travel	113,996	129,649	135,265	378,910
Total	**8,167,785**	**8,849,256**	**8,859,206**	**25,876,247**

	Q4 2013			
	Oct-13	**Nov-13**	**Dec-13**	
New Deals Started				
Local	19,207	19,322	19,888	58,417
Goods	3,649	4,520	4,581	12,750
Travel	466	855	856	2,177
Total	**23,322**	**24,697**	**25,325**	**73,344**

Totals by segment is computed in python using 4Q2013 raw data. Simple filtering by segments and adding all rows for each category.

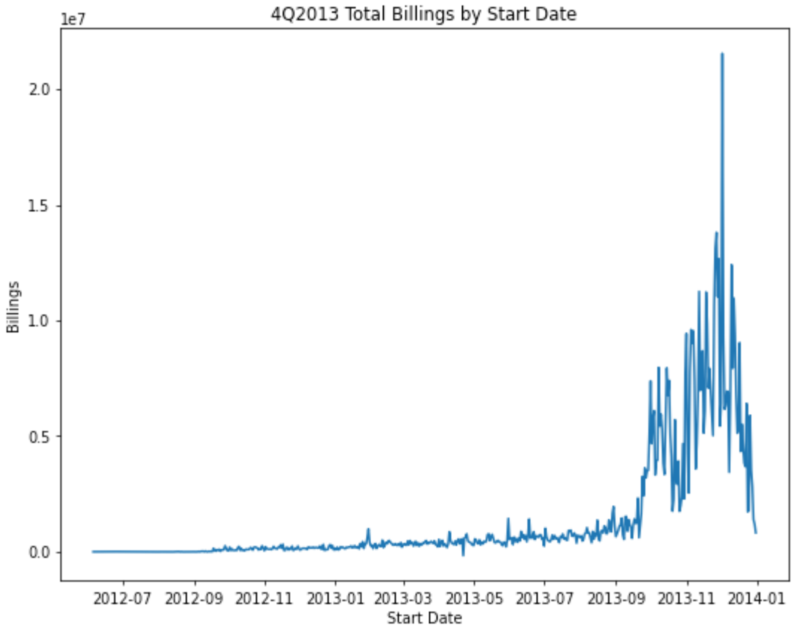
New deals is computed in python using 4Q2013 raw data. Filter rows by segments and start date month, then sum rows by filter.

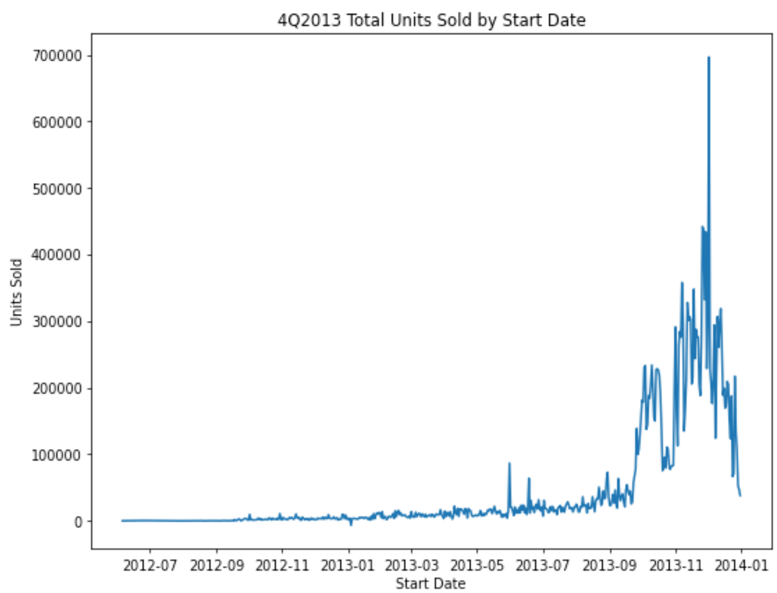
10/20/13 - 10/30/13 new deals in local segment were not captured during these dates. Calculated new deals that started in 10/1/13 - 10/19/13 and proportioned it to estimate missing deals of the last week of October. Billings and Unit Sold for these missing dates are calculated using ARIMA time series which is explained later.

After totals(blue) was calculated, I multiplied the totals by a percentile DataFrame to breakdown Q4 totals into months. Percentile DataFrame (explained below) was calculated in excel and imported into python to calculate breakdown.

10/20/13 – 10/30/13 local new deals, local billing and local units sold prediction (timeseries)

In order to estimate 2013Q4 totals, I needed to estimate the billing and local units sold for the dates missing. I turned the raw data into a timeseries dataset by grouping daily billings and units sold by start date.



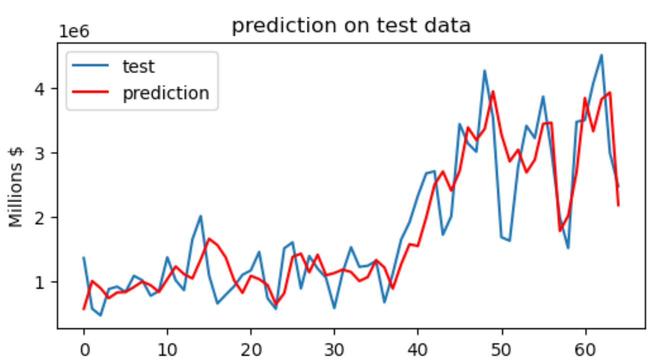


The drop of billings and units sold on the last week of October is attributed to the missing data. There is significantly lower billing and units sold for start dates before 9/2013 because typically Groupon deals will not last very long. It is not normal for a deal to start in 2012 and be claimed in 4Q2013. In addition, there is a spike in deals created in Q4 and a heavy drop towards the end of year. Typically November and December consumption increases dramatically due to the holiday spirits and the busy season in most industries. To capitalize on that, Groupon started many deals during this time knowing these deals would be claimed. As all the holidays are over, the consumption also drops moving businesses into slow season Q1.

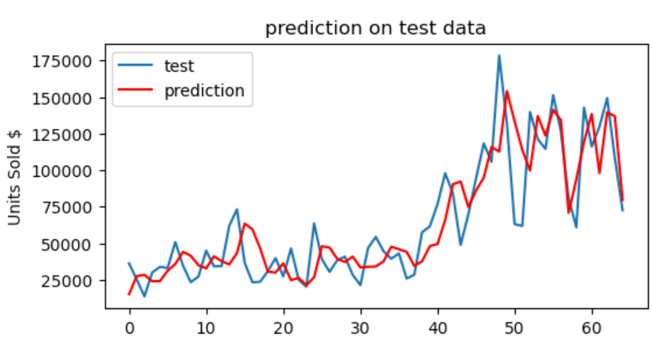
The dataset was turned into a timeseries dataset because the billing estimates and unit sold are heavily focused on the date of the year. ARIMA was the timeseries I chose because I believe it had the perfect characteristics to fit the data. ARIMA is simple enough to not overfit the data, but it is also complex enough to reflect big recent changes. The AR (autoregressive) reflects the changes, the MA (moving average) smoothen the trends and lastly I (integrated) helps make the dataset stationary through differentials. ARIMA forecast the future by using the correlation by some nearby points with some added lag. ARIMA's fails to predict long period of times because correlation weakens as the prediction is further in the future.

<u>ARIMA parameters:</u>
p : order of AR term
q: order of MA term
d: # of differencing to make time series stationary.

After finding optimal parameters, I trained data on daily billings and units sold from the earliest start date to 8/15/2013 that was billed in 4Q2013. Then I predicted daily billings and units sold from 8/16/2013 to 10/19/2013 start dates. Graph and MSE shows that my algorithm is performing better than I expected.

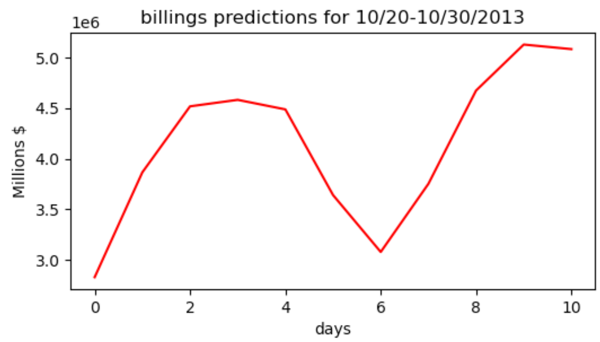


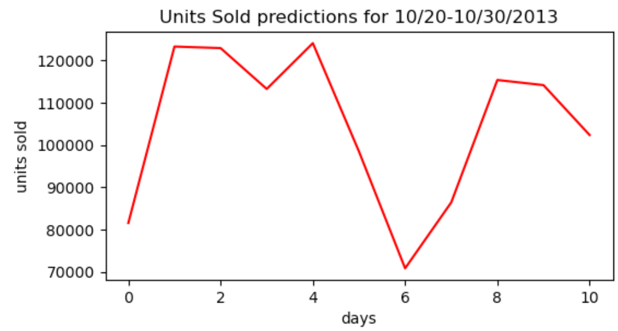
Billings



Units Sold

Lastly, predicted 10/20/2013 – 10/30/2013





Groupon Stock Recommendation and Thoughts

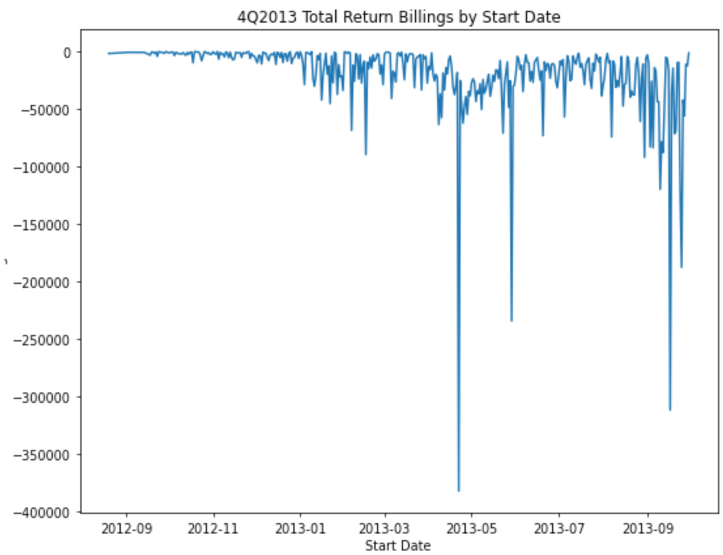
Groupon has established a well known name for itself and investors are keeping a keen eye on this promising company. Quarter after quarter Groupon have proven to be successful with steady growth. Multiple equity research reports have a consensus that the rewards outweigh the risks and categorized the stock as buy.

Morgan Stanley reported gross billings vs YipitData estimates

MORGAN STANLEY								
		2012				**2013**		
USD millions	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
NA Deals/ Travel/ Goods	553.6	548.2	552.4	718.9	681.4	712.2	665	870.1
YipitData								
		2012				**2013**		
USD millions	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
NA Deals/ Travel/ Goods			566	694.5	669.9	725.4	668.8	807.7

YipitData's accuracy in estimating gross billings have increased over time. However, there is a big discrepancy between the two estimates in 4Q2013, (70M difference). This discrepancy is definitely raising some red flags. The only conclusion is Groupon 4Q2013 billings will underperform wealth management estimates OR my timeseries billings prediction for the 10 missing days is way off.

I graphed 4Q2013 returns by start date and noticed some large spikes in returns. Maybe this contributed to lower estimation in billings. However I would need to compare other quarter returns to see if there is a significance change in this recent quarter



Realistically I would have still recommended investors to buy Groupon stock despite the billing estimation discrepancy. During Jan 2014, endless articles and equity research reports have valued Groupon as a unicorn company. I would have been influenced by analysis of these well-known hedge funds and conclude that my estimation is inaccurate due to the missing data 10/20/13 – 10/30/13. In addition, I could be biased due the fact that I have used Groupon before in my personal life. Maybe if I was a more experienced data scientist, I would have more confidence in my analysis.

Part 2

1. Important questions that investors in Groupon would want to know in order to form an opinion on the stock

- What is Groupon's projections on their future sales and what actions are being taken to ensure such projections?
- What are the financial effects on merchants by partnering with Groupon and what is Groupon doing to measure merchants satisfaction
- What are some new products in the works that have yet to hit the market
- Groupon is growing extremely fast in the past two years, how is Groupon's infrastructure adapting to this growth and what are some scalability issues they have encountered?
- What is Groupon doing to help customer adoption on mobile consumerism
- The unit sold growth have slowed down significantly in 2013, what does this mean for Groupon and what attributed to this slow down?

2. Additional investor questions about Groupon that email receipt data can help us answer in which web data cannot.

Email receipt can provide a different layer of detail that is more related to customer behavior. Since each purchase is dated, we can study high and low seasons for certain category of good and services. In addition since we get details of all receipts, we can calculate what percentage of customer purchases utilizes Groupon. Since each customer have a unique ID, we can analyze what type of merchants have loyal customers and what percentage of the merchants' sales are repeating customers.

3. I would expect web data to have more accurate tracking of Groupon's North America gross billings than email receipt data. Email receipt data only entails around 1 million users which is not a good representation of Groupon's users as a whole. Web data does not have high level of details; however, it captures all the deals on Groupon. The billing estimates are calculated accurately by showing the daily unit change on the deals and price of each unit. Web data can capture billing, but it is missing level of details. Calculating Q4 total gross was quite simple, breaking down the total gross into months was a bit difficult. Email receipt data have high customer details of purchases with and without the usage of Groupon. Since each purchase is dated, it is easier to determine how well each month should perform in proportion to other months. Another benefit of receipt data is that we can analyze Groupon effects on shopping behavior by comparing users and their usage of Groupon.

	A	B	C	D	E
1		Local Billings (million)	Goods Billings (million)	Travel Billings (million)	Total Billings (million)
2	7/1/12	133414773	26733909.13	16770765.93	176919448.1
3	8/1/12	137321693.4	33008914.69	15436645.08	185767253.2
4	9/1/12	138200000	50800000	14300000	203300000
5	10/1/12	126800000	64300000	16600000	207700000
6	11/1/12	139900000	85100000	14600000	239600000
7	12/1/12	164400000	64300000	18500000	247200000
8	1/1/13	162700000	47300000	22600000	232600000
9	2/1/13	143000000	46000000	13600000	202600000
10	3/1/13	163400000	51000000	20300000	234700000
11	4/1/13	143800000	62800000	18900000	225500000
12	5/1/13	162200000	71700000	22900000	256800000
13	6/1/13	153200000	67100000	22800000	243100000
14	7/1/13	139951253.5	61279983.68	22404491.6	223635728.8
15	8/1/13	136667939.6	67024982.15	22738887	226431808.8
16	9/1/13	133795040	63194983.17	21735700.81	218725724
17					
18					
19		Local Billings (million)	Goods Billings (million)	Travel Billings (million)	Total Billings (million)
20	7/1/12	0.326248168	0.241842104	0.360604161	
21	8/1/12	0.335802025	0.298607486	0.331917962	
22	9/1/12	0.337949807	0.45955041	0.307477877	
23					
24	10/1/12	0.294131292	0.300889097	0.334004024	
25	11/1/12	0.324518673	0.398221806	0.293762575	
26	12/1/12	0.381350035	0.300889097	0.3722334	
27					
28	1/1/13	0.346834364	0.327789328	0.4	
29	2/1/13	0.304839054	0.318780319	0.240707965	
30	3/1/13	0.348326583	0.353430353	0.359292035	
31					
32	4/1/13	0.31315331	0.311507937	0.292569659	
33	5/1/13	0.353222997	0.355654762	0.354489164	
34	6/1/13	0.333623693	0.332837302	0.352941176	
35					
36	7/1/13	0.341	0.32	0.335	
37	8/1/13	0.333	0.35	0.34	
38	9/1/13	0.326	0.33	0.325	
39					
40		0.309202359	0.300647395	0.339219797	
41		0.327397611	0.37123734	0.302969054	
42		0.363400029	0.328115265	0.357811149	
43					
44		Local Billings (million)	Goods Billings (million)	Travel Billings (million)	Total Billings (million)
45	10/1/13	0.309202359	0.300647395	0.339219797	
46	11/1/13	0.327397611	0.37123734	0.302969054	
47	12/1/13	0.363400029	0.328115265	0.357811149	

3Q12

July
Aug
Sept

1)
2)
3)

given historical data to find ratios.

copy of purple. These green cells were imported into python as a dataframe. (percentile DF)

Segment billing ratio of each month in their respective quarter.
Examples: (July2012/3Q12), (Aug2012/3Q12), (Sept2012/3Q12)
(April2013/2Q13), (May2013/2Q13), (June2013/2Q13)

Weighted average of each ratio

1) weighted average ratio of first month of each quarter=predicted billing ratio Oct4Q13
2) weighted average ratio of 2nd month of each quarter=predicted billing ratio Nov4Q13
3) weighted average ratio of 3rd month of each quarter=predicted billing ratio Dec4Q13

1)+2)+3) = 1.0 (4Q2013 Total Billing)

Weight is calculated by:
3Q2012 = 10%
4Q2012 = 60%
1Q2013 = 10%
2Q2013 = 10%
3Q2013 = 10%

4Q2012 have the highest weight because trying to predict 4Q2013 ratio. I believe there is seasonal factors that effects billing per quarter. I want previous 4Q ratio to play larger effect in prediction of future 4Q ratio

multiply 4Q2013 billing totals with ratios to calculate breakdown by month

Repeat in each Segment! (local goods travel)

1) $\frac{\text{July}}{3Q12}(0.1) + \frac{\text{Oct}}{4Q12}(0.6) + \frac{\text{Jan 13}}{1Q13}(0.1) + \frac{\text{April}}{2Q13}(0.1) + \frac{\text{July}}{3Q13}$

= predicted billing ratio of Oct 4Q2013

2) $\frac{\text{Aug}}{3Q12}(0.1) + \frac{\text{Nov}}{4Q12}(0.6) + \frac{\text{feb}}{1Q13}(0.1) + \frac{\text{may}}{2Q13}(0.1) + \frac{\text{Aug}}{3Q13}$

= predicted billing ratio of Nov 4Q2013

3) $\frac{\text{Sept}}{3Q12}(0.1) + \frac{\text{dec}}{4Q12}(0.6) + \frac{\text{mar}}{1Q13}(0.1) + \frac{\text{June}}{2Q13}(0.1) + \frac{\text{Sept}}{3Q13}$

= predicted billing ratio of Dec 4Q2013