

Product Development Specialist

Groupon Exercise

Part 1

At YipitData, we collect and analyze unique data sets that allow us to track various metrics on many publicly traded companies. This helps our clients, who are investors in the stock market, better understand the companies they invest in.

One of the companies we cover is Groupon. For Groupon, the main metric we track is called gross billings. Every quarter, Groupon reports gross billings in their financial statements. We use our proprietary data to estimate gross billings before Groupon reports it in their financial statements, giving an edge to investors who purchase our reports.

Your goal is to use the attached data file to estimate Groupon's 4Q13 North America gross billings by segment (Local, Travel, and Goods). Then, using this estimate, make a buy or sell recommendation for Groupon stock, pretending it is January 2014, before Groupon reports 4Q13 earnings in February 2014. The data is based on real data we collected for Groupon in 2013, and you will need to overcome real challenges we faced back then in order to arrive at your estimate. Remember, we are pretending it is January 2014, so you can't use any information that is after January 2014.

Your response should include:

- Your billings estimate by segment.
- A concise summary of how you arrived at that estimate including any adjustments to the data and the rationale for those adjustments.
- A summary of any patterns you notice in the data, what is causing those patterns, and any relevant adjustments you made for them, if necessary.
- Key calculations you made so we can check your work.
- Any other relevant irregularities in the data and how you dealt with them.
- A buy or sell recommendation for Groupon stock from the perspective of January 2014. Your recommendation should be based on comparing your estimate to Wall Street consensus (you can estimate this using the attached equity research reports).

You will be graded on:

- Accuracy of your billings estimates.

- Attention to detail. This is very important. We are looking for candidates who can produce data outputs with zero errors and **who are paranoid about finding every data inaccuracy.**
- Logic of your data adjustments and demonstration of understanding the data set and patterns in the data set.
- Clearness of your communication. We also look to make sure that your writing is not only clear, but also free of errors. Strong written communication is an important part of this job.
- Soundness of your logic for your buy or sell recommendation.
- Many candidates choose to use an advanced model to help deal with data adjustments. This is not necessary – simple solutions often work best. What's important is that you understand exactly how you arrive at your solution and why it is the optimal solution given the nature and patterns of the data, and that the solution is truly accurate.

Methods & Deliverable:

- **Please read all directions about the data set carefully. The data set is challenging – make sure you fully understand what the data is!**
- Please include a write up that explains your answer and your methods. This should be in a .pdf file. It should read like a written report, not a notebook format.
- Please include your calculations. This can be done in excel, python, or any other software or language that you prefer.
- **You should not receive help from any person on this assignment. If we find out that you received help, we will disqualify you from the process.**
- We recommend that you review Groupon's 2012 annual report [here](#) to learn more about Groupon's business and strategy. This is a document that would have been available to you had you been doing this analysis back in early 2014, which is when we did this analysis ourselves.
- The whole annual report is over 100 pages, so please do not read the entire thing. We recommend reading Item 1: (pages 1 through 11) and item 7 (pages 30-57). Skimming these sections is sufficient – you do not need to know all of the details, you just need to understand a bit about Groupon's business so you can analyze the data provided.
- Please also read the attached equity research reports. All of the reading of the equity research reports and the annual report combined probably shouldn't take more than 2 hours, and could take a lot less. You don't need to know all of the details, you should just read enough to understand how Groupon's business works and what the different segments are.

Background on the Groupon Q4 2013 North America dataset:

- This dataset is YipitData's proprietary estimate of gross billings and units sold for each deal that was active in Groupon's North America segment in Q4 2013.

- Each row represents a Groupon deal that was active for some or all of Q4 2013. For each row we've provided data on units sold during Q4 2013, gross billings during Q4 2013, the date that the deal started, the URL of the deal page, the product segment of the deal (Local, Travel, or Goods), and the inventory type of the good (first-party means Groupon owns the inventory).
- Groupon offers thousands of deals on its platform, with new ones starting each day. Some deals stay active for weeks, months, or even years, while other deals are only active for a day or a few days. In the data file, we include the start date of each deal that was active in Q4 2013. Remember, these dates are not necessarily the date the billings occurred, but the date that the deal launched. Since deals can be active for many days, or even weeks or months, these deals can have billings on many days, not just the date that the deal started.
- We collect this data by finding all deals on Groupon's platform, and then tracking the quantity sold and price information that is available on each deal webpage. Since we have the price and quantity for each deal, we can generate companywide estimates of gross billings.
- However, our system that finds the deals on the Groupon North America website and adds them to our database broke from October 20 to October 30 2013 (inclusive) for the Local segment, so we did not add any Local segment deals that started between those dates to our database (you'll notice this in the dataset). This means we missed all billings from deals that started during this period, even if they remained active after the period. In other words, our dataset includes zero Local deals that started from October 20 to October 30, 2013 inclusive, and 100% of all other Local deals that were active in Q4 2013. You'll need to adjust the data to deal with this outage as you make an estimate for North America Q4 2013 billings by segment.
- Note that purchasers of Groupon deals can also return deals if they decide they don't want to use them. The data we present in the spreadsheet is net of returns. For example, let's say a row of data has \$100 in billings. This means there were \$100 gross billings for this deal in Q4 2013, which could mean there was \$100 of sales and \$0 of returns, it could mean \$110 of sales and \$10 of returns, it could mean \$500 of sales and \$400 of returns. We don't breakout the data by purchases and returns, we just provide the net number, which is called "gross billings". This is in line with how Groupon reports Gross billings in their financial statements. Note that some rows in our data have negative gross billings. This means that these deals had more returns than purchases in Q4 2013.
- You'll notice that units sold are often in decimals. This is because we are employing estimation techniques behind the scenes - you can ignore the methodology behind these estimations and just take the data as given.
- There is a tab in the Q4 2013 North America Groupon Data XLSX file called "Historical Data," which shows you YipitData estimates from previous quarters. This tab may be helpful in understanding to what extent YipitData estimates can accurately estimate Groupon reported gross billings, which is a number that investors care about. Although we broke out the historical figures on a monthly basis, the raw data we provided for Q4

2013 does not break gross billings out by month. Additionally, we do not provide the raw data that was used to generate the historical estimates, only the Q4 2013 raw data.

- Data is not always as clean as you would hope. As such, there may be inaccuracies in the data in addition to the ones described above. If you find any, you should determine their impact and adjust accordingly if you think an adjustment is warranted. As a reminder, please describe any adjustments you made and the rationale for those adjustments.

Specifically, here is the data dictionary of the data:

- deal_id - Unique identifier for each deal
- units - Total number of units sold of each deal during 4Q 2013
- billings - Total billings from each deal during 4Q 2013
- start_date - The date the deal started
- segment - The product segment of the deal (Local, Travel, Goods)
- inventory - The type of inventory of the deal (First party means Groupon owns the inventory)

Part 2

At YipitData, one of the ways we think about our job is, we try to identify the key questions that investors are asking about the companies that we cover, and then we use data to try to answer those questions. For example, for Groupon, one of the key questions that investors ask is, what will Groupon report for gross billings this quarter?

Given this, please answer the following questions:

- 1) **What do you think are the most important questions that investors in Groupon would want to know in order to form an opinion on the stock?** Please come up with 5-10 questions (more isn't necessarily better, we are looking for quality).
- 2) The Part 1 exercise above is based on what we call "web data" – data that is publicly available that we collect from the internet. Another data set we have access to is called email receipt data. With email receipt data, we have access to the email receipts of about 1 million people in North America. These 1 million people are the users of a particular app that opt into data sharing – that is how we have access to the data. Each of these 1 million people have an anonymized ID number so we can track their purchases across the data. For each receipt we see:
 - the anonymized user ID number
 - the merchant the person bought from
 - the date of the purchase

- the items that were purchased
- the price of the items (also can see things applied to the price such as discounts, shipping fees, taxes, etc.)

What additional investor questions about Groupon could email receipt data help us answer that web data cannot answer?

- 3) Which data set, web data or email receipt data, would you expect to be more accurate for tracking Groupon's North America gross billings and why? What do you think are the strengths and weaknesses of each data set (web data vs email receipt data)?**