

Data Quality Assurance Check

Hongting Chen¹

¹New York University (NYU)

February 14, 2018

To estimate Groupon's 4Q13 North America gross billings for the investment research, we face a few challenges due to data limitations. This article is the

Examining the Proprietary Data

The dataset is YipitData's proprietary estimate of billings and units for each deal that was active in Groupon's North America segment in Q4 2013. Each row represents a Groupon deal and we provide features including

- Deal ID
- Units sold during the period
- Billings during the period
- Date that the deal started
- URL of the deal page
- Segment of the deal(Local, Travel, Goods)
- Inventory type

Data Challenge

It is informed that due to database outage, the dataset includes zero Local deals that started from October 20 to October 30, 2013 inclusively.

To address the missing Local data and better estimate the billings, an exploratory analysis is conducted.

