Yelp Review Frequency

Maxwell Peterson

When do people post Yelp reviews?

The number of reviews posted to Yelp per day is a proxy to a number of interesting questions. Is Yelp usage still growing? What days of the week do people visit restaurants? Do people go out on holidays (or: are restaurants open on holidays?)?

Input Data

Input data is a subset of Yelp data. Data was provided for 11 metropolitan areas. Dates range from July 22, 2004 to December 11, 2017. We examine trends at the state and province level (from here-on just "states"). Input data contains reviews from 68 states, but not all states saw enough Yelp activity to say interesting things about; just 22 had more than 100 reviews all-time. If this seems low, note that the publicly-available data is only a small subset of Yelp's data, and than some very small "states" are likely just incorrectly marked (e.g. state "01")

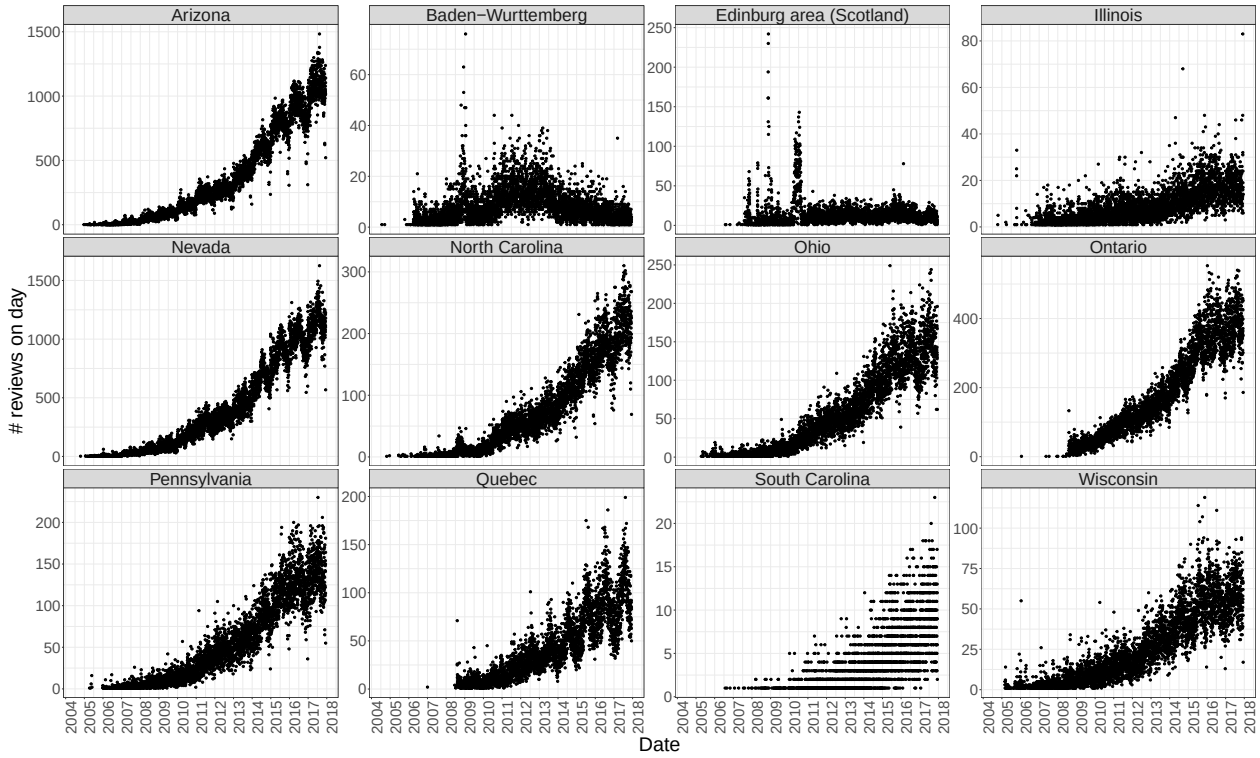
Part of our objective will be to forecast future review counts. So we'll keep only those states that have a decent amount of activity in the recent past. Specifically, we only analyze states with 300 or more reviews in the year 2017. These are:

State	Reviews in 2017
Nevada	392,307
Arizona	368,147
Ontario	131,353
North Carolina	72,092
Ohio	50,071
Pennsylvania	48,060
Quebec	31,733
Wisconsin	20,043
Illinois	6,286
Edinburg area (Scotland)	3,338
South Carolina	3,088
Baden-Wurttemberg	1,384

Going forward, we only use these states.

Reviews Per Day

Reviews per day per state all-time activity. Note that each y axis is different.



Observe:

- There is an upward trend in most states.

  This makes sense – Yelp was founded in 2004. More people learned of the site over time and so reviews grew over time. What's interesting is the states that *didn't* trend upward over the past 12 years: Baden-Wurttemburg in Germany, and the Edinburg region in Scotland. A few possible explanations:

  – Competition

    Yelp did have a serious European competitor: Qype. In May 2012, Qype may have had as much as 5 times Yelp's traffic (Lunden, Ingrid 2012a). Yelp bought Qype 5 months later (Lunden, Ingrid 2012b), and the beginning of an upward trend in Baden-Wurttemburg happens around the same time! But it doesn't continue.

- Large, infrequent spikes downward in reviews posted

  This is most obvious in Arizona, but is the case in many states, including the two Canadian provinces. These are probably the big winter holidays.

- Infrequent high outliers

  At least Illinois, Wisconsin, and Quebec have a few days with a very high number of reviews.

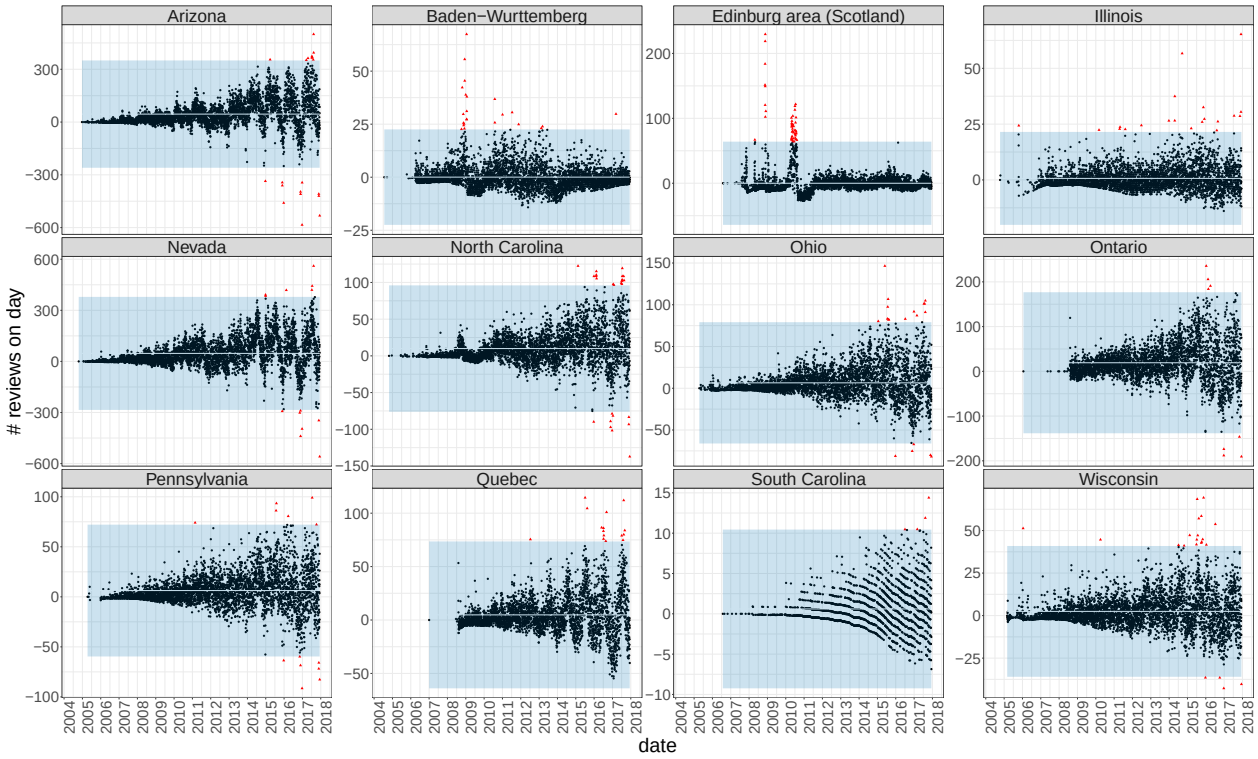
- South Carolina South Carolina's utilization looks odd, but there are very few reviews in the state.

Outliers

There are outliers in the dataset. We'd like to know:

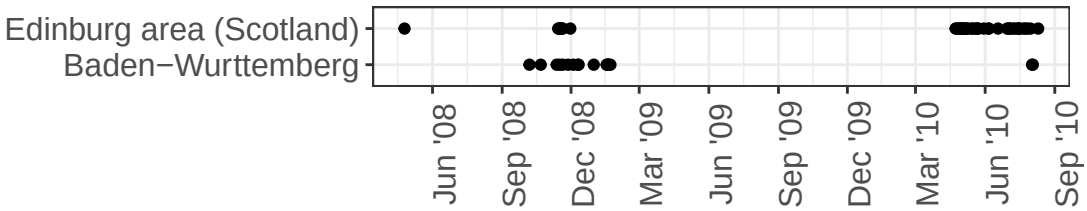
- Did the single dramatic spikes upward in the two European countries occur on the same day?
- Do the yearly dips in AZ, PA OH, WI, and more fall on Christmas, or Thanksgiving?

To answer these questions, any rough definition of outliers will do. We just need a scheme that picks up the outliers obvious to the eye, and that doesn't pick up too many days. We landed on taking those days that were 4 or more standard deviations from the yearly-detrended mean. In other words: we removed the upward trend from each state, then took the points that were unusually far away from the average. Then graphically, the average is the horizontal white line and the outliers are the red triangles:



Did a spike in usage occur in two European states at the same time?

The far-away view suggests that the Baden-Wurttemburg outlier spike and the Edinburg spike may fall on the same days. But closer examination is less exciting; below, we mark each outlier day with a black dot, and leave non-outlier days blank. The spikes we are interested occur between 2008 and 2010, so only those years are included.



Though the outlier days sometimes overlap for the two regions, the patterns are less similar than the red-triangled outlier plot suggested. So there's not much here. Let's move on.

Are holidays causing the infrequent downward jumps in some states?

Given a holiday H, we consider H, the day before H, and the day after to be the same holiday, and name them all H. Then the counts of holidays that are also outlier days are as follows:

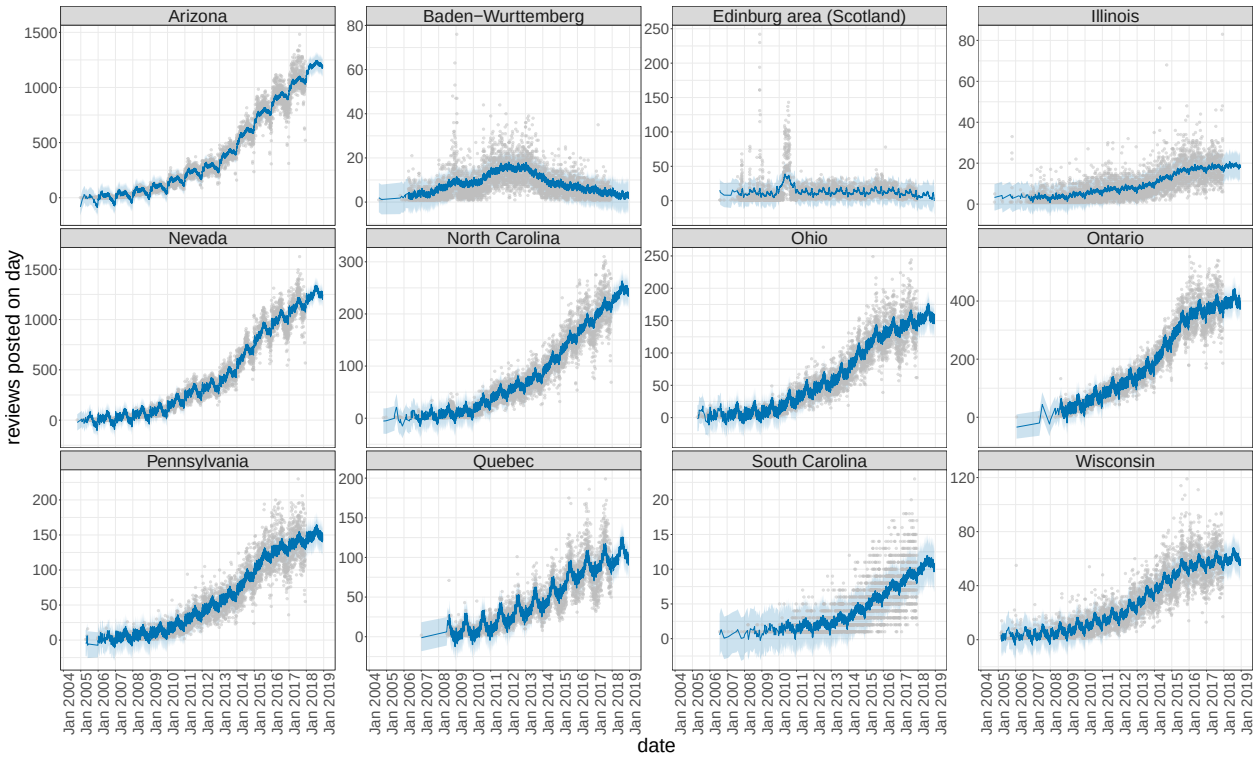
State	Holiday	Outliers (all-time)
Arizona	Thanksgiving	6
Nevada	Thanksgiving	4
North Carolina	Thanksgiving	4
Pennsylvania	Thanksgiving	3
Ohio	Thanksgiving	2
Arizona	Christmas	4
North Carolina	Christmas	2
Ohio	Christmas	2
Ontario	Christmas	2
Pennsylvania	Christmas	2
Wisconsin	Christmas	2
Nevada	Christmas	1
North Carolina	GWBirthday	3
Nevada	GWBirthday	1
Ontario	GWBirthday	1
Wisconsin	GWBirthday	1
Quebec	Independence	2
South Carolina	Independence	1
Quebec	Labor	2
North Carolina	MLKing	1
Ohio	MLKing	1
Pennsylvania	Columbus	1
Wisconsin	Columbus	1
Pennsylvania	GoodFriday	1
Quebec	Memorial	1

So Thanksgiving and Christmas do often correspond to the low days. This is reasonable; people probably go to restaurants less on these big family-oriented holidays. This also suggests we should include these holidays in any forecasting we do.

There are some good strange cases here, too; Quebec's apparent celebration of Independence Day is especially charming.

Forecasting future reviews-per-day

Auto-fitting a timeseries model to each state using default settings (and denying holidays any special treatment) gives the following fits to the data. We also include a forecast for one year past the final data point:

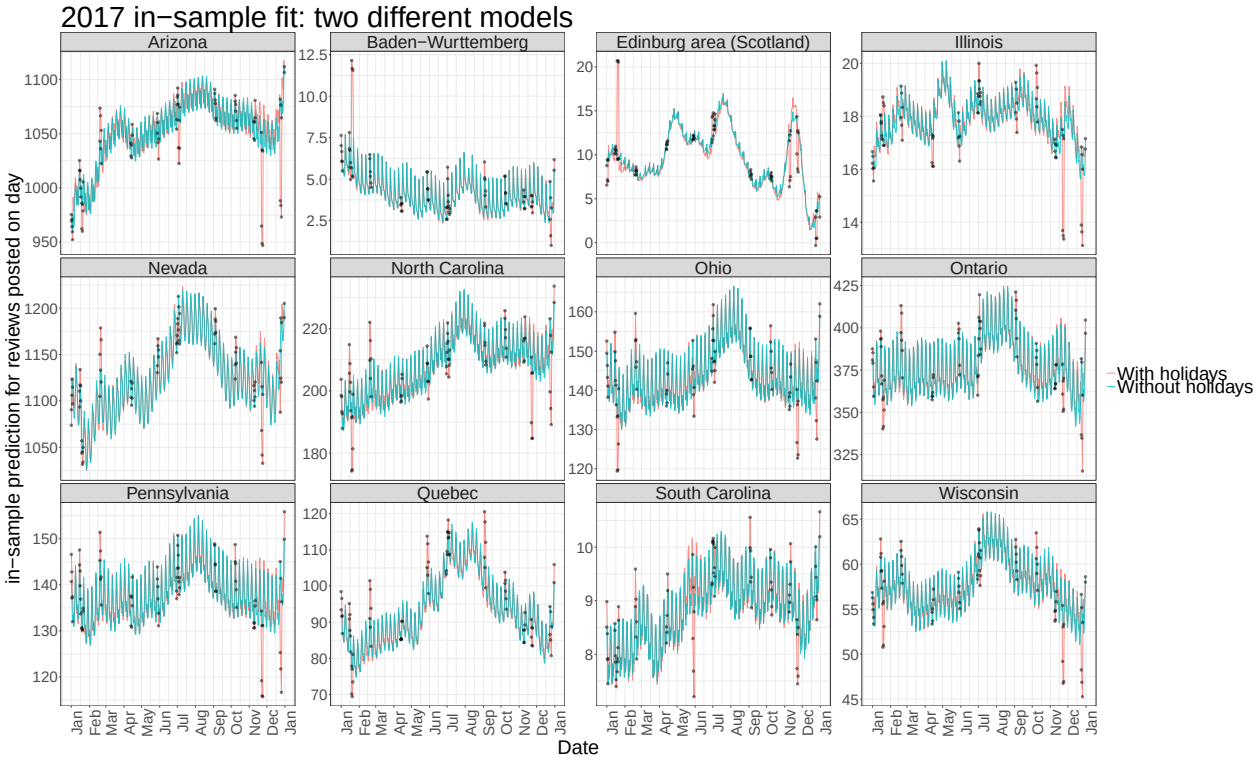


The upward trends are captured nicely, and the within-year variation also looks good. This fit can be used to predict, but also provides a clearer picture of the different trends over time. The fitted line allows us to observe:

- Almost all states have their highest yearly yelp review activity in the middle of the year.
- Arizona exhibits a strange step-like behavior in its upward trend.
- The year-of-year acceleration in reviews has leveled off in the past few years: Nevada, Pennsylvania, Ohio, and especially Wisconsin and Ontario appear to perhaps be going level (though the increasing spread in the data in recent years in some states [esp. Ontario] complicates this conclusion).

Marking holidays

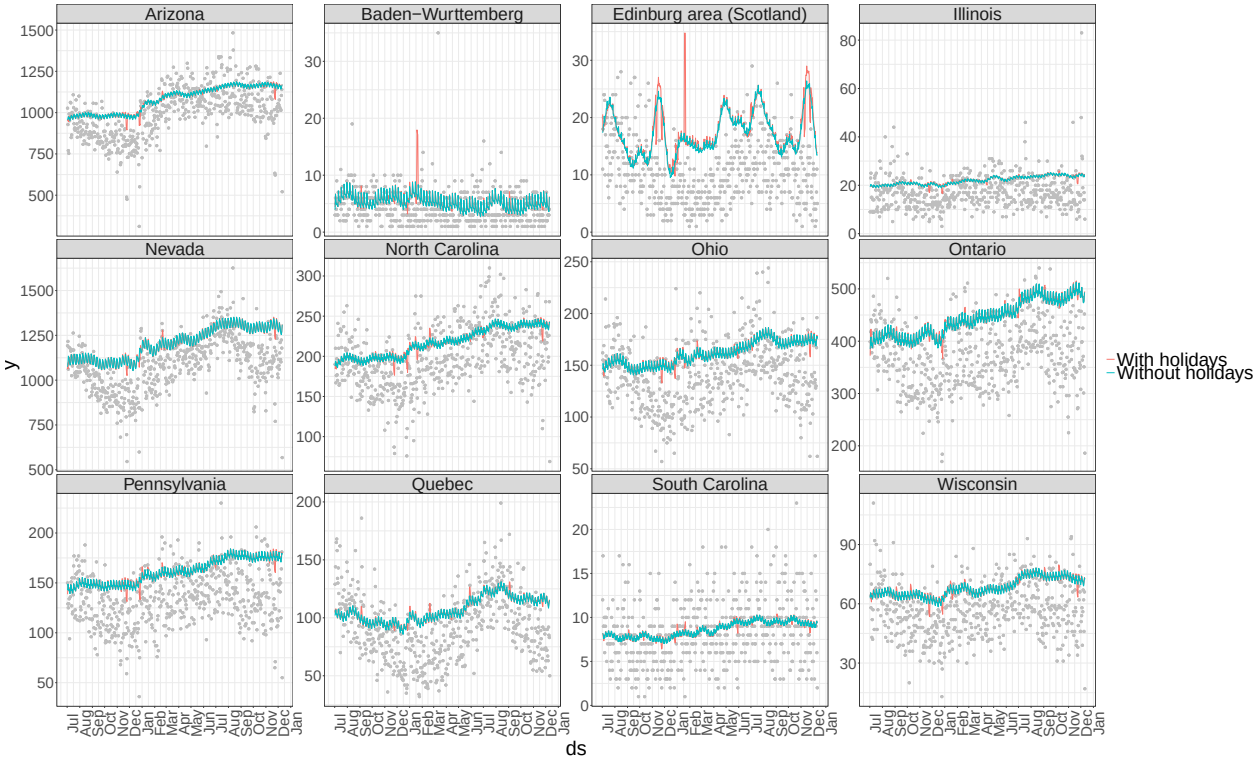
We discovered in a previous section that a lot of outliers coincide with holidays. Treating holidays specially may improve the forecast. Fitting the same model as above, but including the holidays from the previous section, results in the following in-sample fit (zoomed-in to 2017 for better resolution). Holidays are marked with points (recall that we consider the 3-day area around a holiday to all be holidays).



So marking holidays has induced the fit to dive and arc aggressively for holidays on which there were extreme values; but at the same time, no dives or arcs are made for days marked holidays that had values in normal ranges. So our strategy of marking a slew of holidays, even ones that do not commonly coincide with extreme values, did not hurt the model's ability to discriminate between (in-sample) true extreme holidays and technically-holidays-but-pretty-normal days.

Choosing a model

The in-sample twelve-year view suggests that our fits are smart and reasonable. But so far we've only seen fits on in-sample data. We now examine out-of-sample error. As before, we fit models with and without holidays, but this time the models are trained only up to the end of 2016. Then we forecast 2017.



So although the fits are reasonable, we also get a decent amount of error day-to-day. We also see that, for all our holiday considerations in previous sections, the model actually declines to dive and arc as it did on the in-sample holidays, and the two models are almost identical! So we go on to report accuracy measures on the with-marked-holidays model, but we would do just as well with the holiday-agnostic model. Finally, we notice that though the direction and acceleration of the year-to-year upward trends are nicely captured, forecasts often fail to cut through the middle of the points, instead tending to over-forecast most days.

Accuracy

The average *daily* percentage error per state for this 365-day forecast on 2017 is as follows:

State	Avg daily % error	Sum-of-year % error
Arizona	8%	3.5%
Baden-Wurttemberg	93%	-7.1%
Edinburg area (Scotland)	111%	-47.9%
Illinois	33%	-5.2%
Nevada	8%	-0.4%
North Carolina	14%	-0.7%
Ohio	16%	-2.4%
Ontario	12%	-3.9%
Pennsylvania	16%	-4.7%
Quebec	29%	-12.7%
South Carolina	41%	3.2%
Wisconsin	21%	-6.9%

- The interpretation of Avg daily % error is: If we forecasted the reviews for every day in 2017 for, e.g., North Carolia, the forecasted value for a given day would be, on average, 14% off the mark.
- The interpretation of the sum-of-year percentage error is: If, at the end of 2016, we had forecasted the total-reviews-in-year for 2017 for, e.g., North Carolia, the forecasted year total would have been 0.7% below the actual total.

Some states see impressive sum-of-year accuracies; others see impressive average daily accuracy. Whether any state model is sufficiently accurate must be considered in the context of what the important accuracy measure is.

Depending on the application, forecasts can be made on horizons much shorter than 365 days, and will be more accurate the shorter the horizon. But yearly forecasts have their place too. For example, if Yelp were considering how many more servers to buy in a state, a yearly forecast of total reviews can inform the decision (though modeling traffic per day would be better).

Lunden, Ingrid. 2012a. "Qype, the Yelp of Europe, Claims Top Dog Status with 860,000 Places Reviewed, Expands Daily Deals." Tech Crunch. May. https://techcrunch.com/2012/05/11/qype-the-yelp-of-europe-reaches-860000-places-reviewed-and-expands-its-daily-deals-service.

———. 2012b. "Yelp Pays $50M to Acquire Its Big European Rival, Qype, to Beef up Its Recommendations and Listings Business." Tech Crunch. October. https://techcrunch.com/2012/10/24/yelp-pays-50m-to-acquire-its-big-european-rival-qype/.