

Detection of Emerging Food Trends

-Based on Facebook posts text data

Yunqing Yu
MS Marketing Analytics, Simon Business School



Preview

Since popularity of certain food is positively correlated with term frequency, this Facebook text analysis is mainly using unigram model to generate document term matrix to detect trends. I used cauliflower rice and vegetable noodle, which became popular in 2016 and 2015 respectively, to test the effectiveness of this model.

Data Preprocess

Create word corpus and time index. Remove punctuations, numbers, whitespaces and stopwords (most commonly used words in English i.e. 'I', 'are', 'and') in Facebook data.

Key Term

Examined the most representative term to analyze the food trend.

Term Frequency & Moving Average Trend

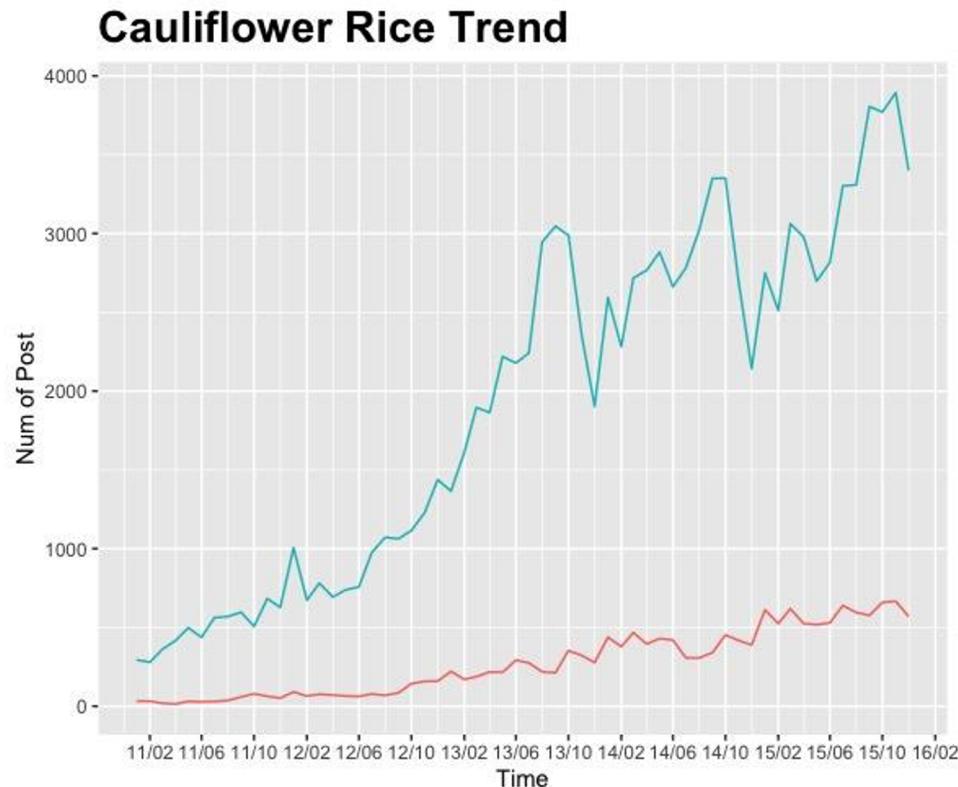
Explored term frequency in time range and compared with moving average line to see whether it has formed a trend.



Use 'Cauliflower' to Detect Trend

After some data cleaning, the frequency lines of two ingredients are plotted.

"**Rice**" is a much frequently used word in dataset, the trend of which cannot represent 'Cauliflower rice'. Thus I decided to use the term '**Cauliflower**'.

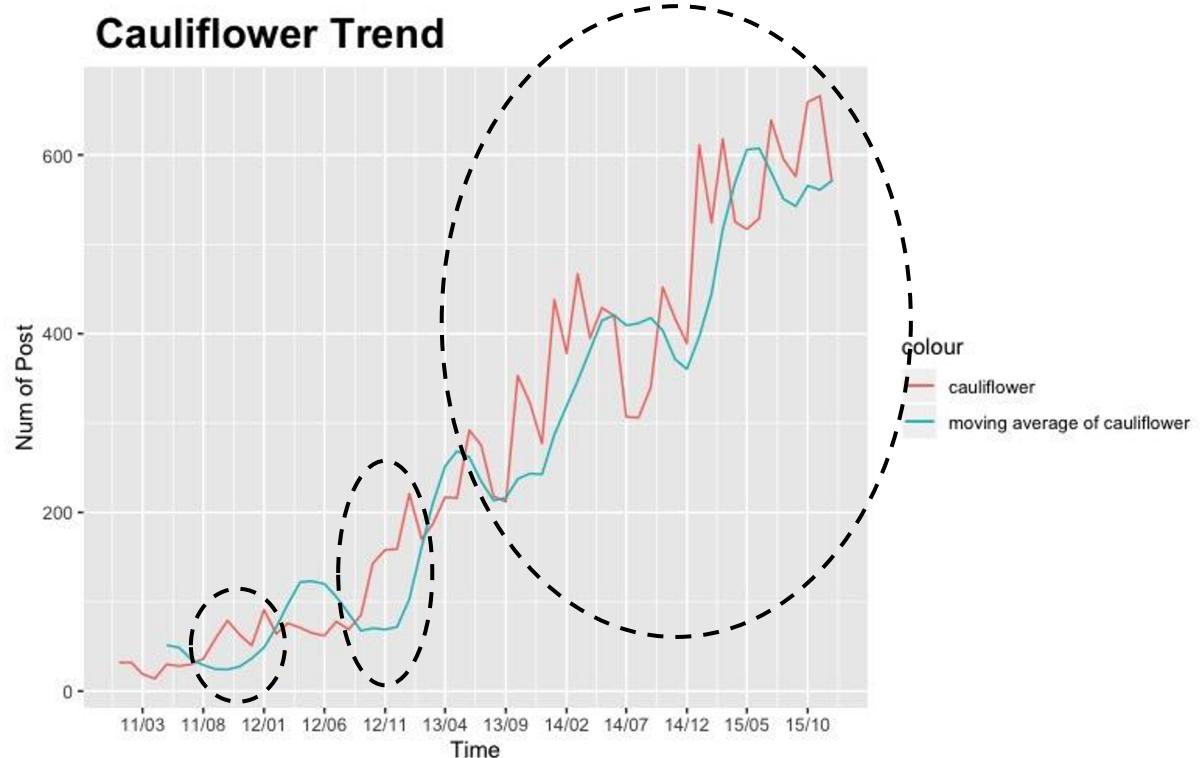




'Cauliflower' became a trending topic

Based on moving average in 5 months, '**Cauliflower**' first became popular between **Aug. 2011-Jan.2012**; Around **Oct. 2012- Mar.2013**, it grasp our attention again. Since **mid 2013**, there has formed several short-term hot topics.

This indicates the fact that cauliflower rice became a trend in 2016.

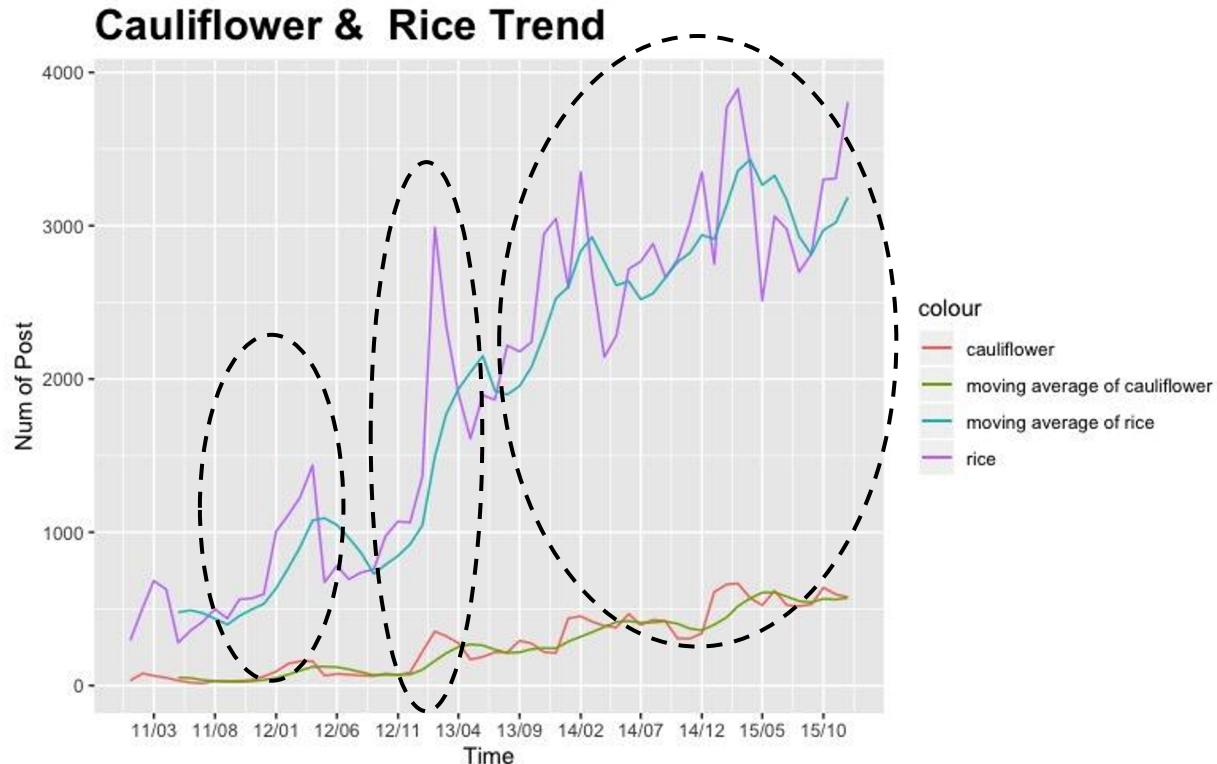




'Cauliflower rice' became a trending topic

Although 'Rice' was not a good indicator of 'Cauliflower rice' trend, it shares a similar pattern in popularity with 'cauliflower'.

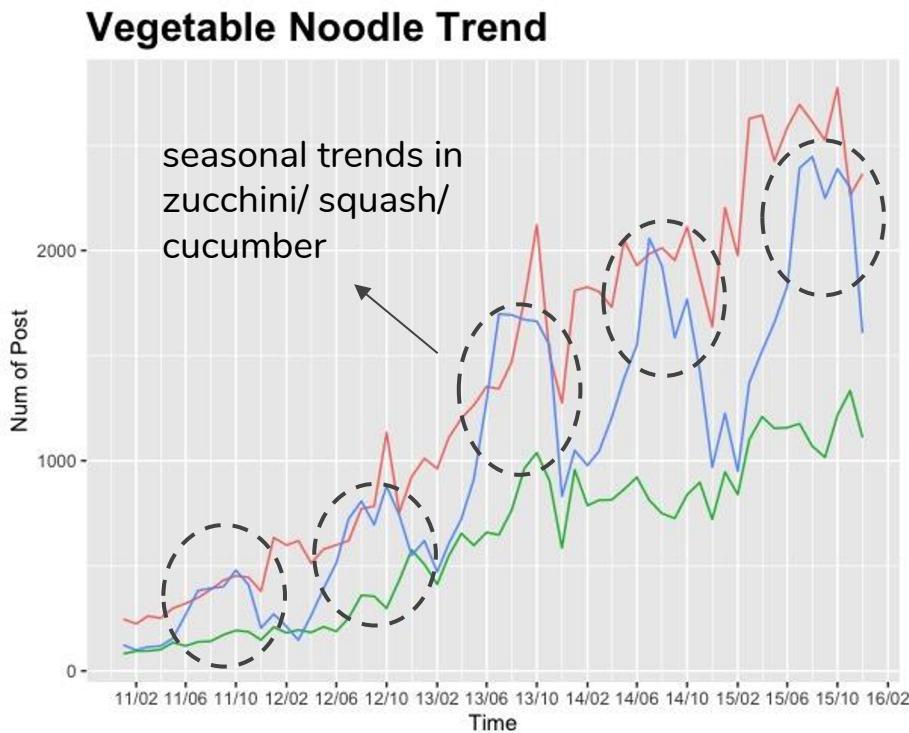
Takeaway: Using the model, we can extract the signals as early as **late 2011**, and we can see the rise of cauliflower rice since **mid 2013**.





Term Selection

It's clear that '**vegetable**', '**noodle/pasta/spiralizer**' and '**zucchini/squash/cucumber**', the ingredients related with 'vegetable noodle' share similar pattern. But for '**zucchini, squash and cucumber**', it has a seasonal variance (popular in summer), thus the terms might not be accurate to detect the trend.

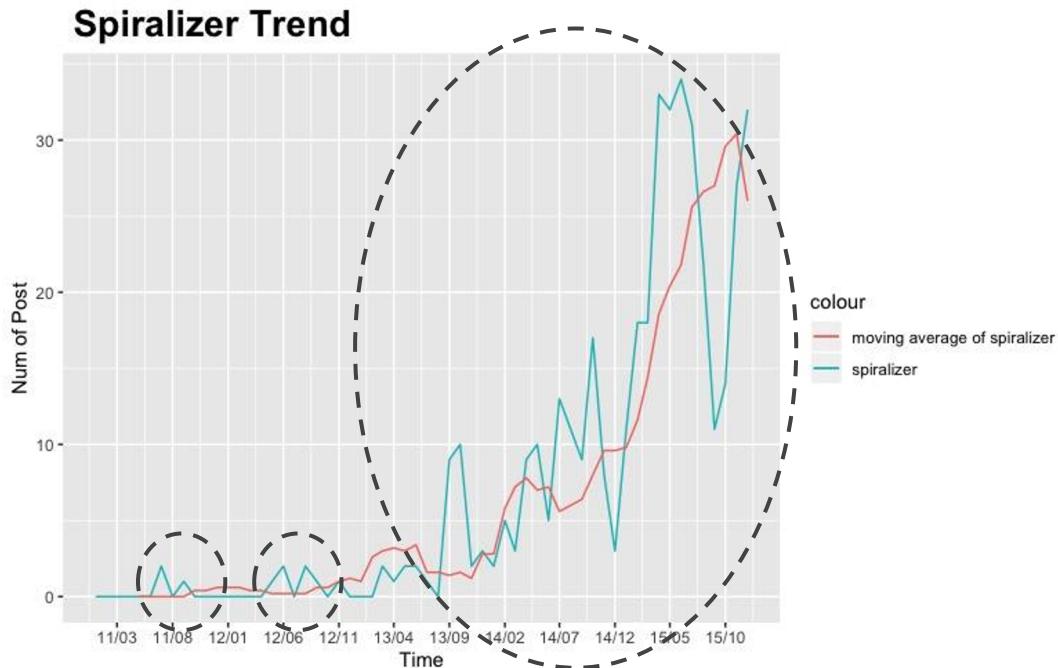




Examine the trend in 'Vegetable Noodle'

'Spiralizer' could better capture the trend. It was continuously frequently discussed from **mid 2013**.

Takeaway: Based on the model, we can detect the rising popularity as early as **mid 2011**, and we can pay attention to the new food combination from **mid 2013**.



Detection of Emerging Food Trends

-Based on Facebook posts text data

Yunqing Yu

MS Marketing Analytics, Simon Business School

Yunqing.Yu@simon.rochester.edu