

# Foundations of Graphical Models

## Problem Set #1

Si Kai Lee `sl3950@columbia.edu`

October 6, 2016

### Problem 1

**a**

$$2^6 = 64$$

**b**

According to the readings, size of table of conditional probability associated with node  $X_i = r^{m_i+1}$  where  $m_i$  are the number of variables conditioned on.  $P(X_1|X_2, X_3, X_4) = 2^{3+1} = 16$ .

**c**

$$P(X_1, X_2|X_3, X_4, X_5) = 2^{4+1} = 32$$

**d**

$$p(x_{1:6}) = p(x_6) \cdot p(x_5|x_6) p(x_4|x_6) p(x_3) p(x_2|x_3, x_4, x_5) p(x_1|x_2, x_3)$$

**e**

The number of entries are (w.r.t. to the above factorisation) is  $2 + 2^2 + 2^2 + 2 + 2^4 + 2^3 = 36$

**f**

$P(X_1|X_2, X_3, X_4) = P(X_1|X_2, X_3)$  as  $X_1 \perp\!\!\!\perp X_4 | \{X_2, X_3\}$ . Hence we will touch  $2^3 = 8$  entries.

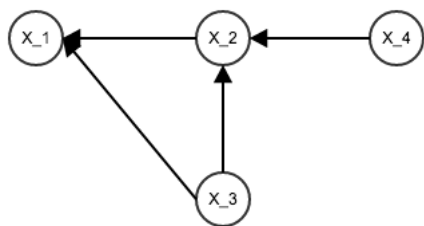
**g**

$P(X_1, X_2|X_3, X_4, X_5) = \sum_{X_3} P(X_1|X_2, X_3) P(X_2|X_3, X_4, X_5)$ . Hence we will touch  $2^3 + 2^4 = 24$  entries.

## Problem 2

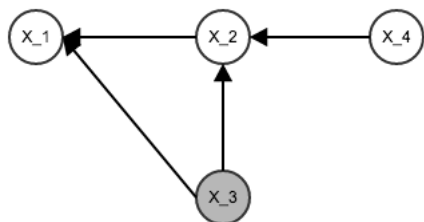
**a**

We redraw the affected nodes in the following manner:



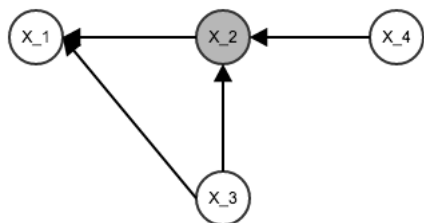
Then we can see that  $X_1 \not\perp\!\!\!\perp X_4$

**b**



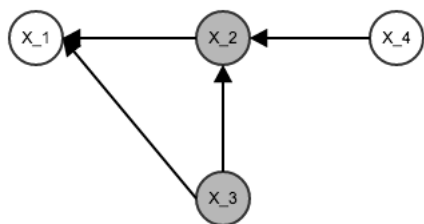
From the diagram, it is obvious that that  $X_1 \not\perp\!\!\!\perp X_4 | X_3$

**c**



We can have a ball going from  $X_1$  to  $X_3$  to  $X_2$  and then  $X_1$ , hence  $X_1 \not\perp\!\!\!\perp X_4 | X_2$

**d**

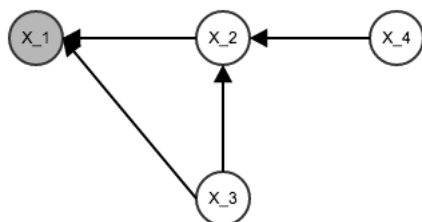


Balls are not permitted to go from  $X_1$  to either  $X_2$  and  $X_3$  or from  $X_4$  to  $X_3$  to  $X_1$  so  $X_1 \perp\!\!\!\perp X_4 | X_2, X_3$

e

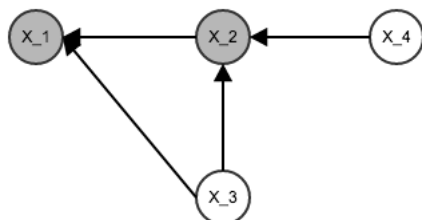
Although we have a v-structure between  $X_3, X_2, X_4$ , we have an inverted v-structure between  $X_3, X_1, X_2$ . We then go from  $X_3$  to  $X_2$  via  $X_1$  then to  $X_4$  which gives  $X_3 \not\perp\!\!\!\perp X_4$

f



We cannot have a ball going from  $X_3$  to  $X_2$  via  $X_1$  as  $X_1$  is now shaded thus  $X_3 \perp\!\!\!\perp X_4 | X_1$

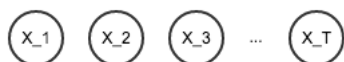
g



We can have a ball going from  $X_3$  to  $X_4$  via  $X_2$  thus  $X_3 \not\perp\!\!\!\perp X_4 | X_1, X_2$

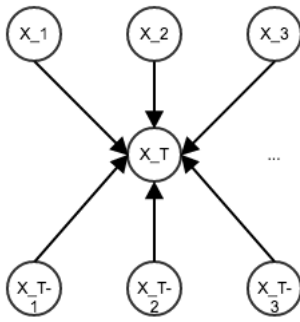
### Problem 3

a



b

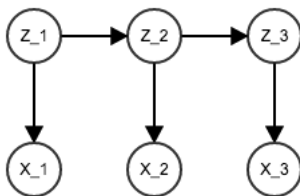


**c****d**

All nodes in a) are independent on each other, the  $i^{th}$  node in b) is only dependent on the  $i - 1^{th}$  node and the  $T^{th}$  node is dependent on all nodes in c).

We could model coin flips with a), speech with b) and the indicator function that all items on this week's shopping list have been bought ( $X_T$ ) with c) and  $X_1 \dots X_T$  being all ingredients needed.

## Problem 4

**a****b**

We begin by converting the HMM to a tree by hanging the graph on the node representing the hidden state we want to marginalise and pass messages from the leaves to the root to obtain the marginal probability of the node. A good ordering can be obtained by performing a depth-first search on the node we want to marginalise which would ensure that we do not do any unnecessary calculations. Since we are marginalising on  $z_2$ , have  $I = \{x_1, z_1, x_2, x_3, z_3, z_2\}$ .

**c**

The outcome observed is  $\{H, T, H\}$  and we want to know  $p(Z_2 = F)$ .

$$\begin{aligned}
 m_{X_1 \rightarrow Z_1}(Z_1 = F) &= m_{X_3 \rightarrow Z_3}(Z_3 = F) = 0.5 \\
 m_{X_1 \rightarrow Z_1}(Z_1 = L) &= m_{X_3 \rightarrow Z_3}(Z_3 = L) = 0.8 \\
 m_{X_2 \rightarrow Z_2}(Z_2 = F) &= 0.5 \\
 m_{X_2 \rightarrow Z_2}(Z_2 = L) &= 0.2 \\
 m_{Z_1 \rightarrow Z_2}(Z_2 = F) &= 0.5 * 0.75 * 0.8 + 0.8 * 0.25 * 0.1 = 0.32 \\
 m_{Z_1 \rightarrow Z_2}(Z_2 = L) &= 0.5 * 0.75 * 0.2 + 0.8 * 0.25 * 0.9 = 0.255 \\
 m_{Z_3 \rightarrow Z_2}(Z_2 = F) &= 0.5 * 0.8 + 0.8 * 0.2 = 0.56 \\
 m_{Z_3 \rightarrow Z_2}(Z_2 = L) &= 0.5 * 0.1 + 0.8 * 0.9 = 0.77 \\
 m_{z_2=F} &= 0.32 * 0.5 * 0.56 \\
 m_{z_2=L} &= 0.255 * 0.2 * 0.77 \\
 p(z_2 = F) &= \frac{m_{z_2=F}}{m_{z_2=F} + m_{z_2=L}} = 0.695
 \end{aligned}$$

## Problem 5

**a**

I am a massive foodie as the national pastime of Singapore (where I am from) is to search for good and cheap food. This summer in San Francisco, I embarked on a quest to find the best burrito in the city as it meets both criteria of being good and cheap. However, my favourite burrito place was Papalote while FiveThirtyEight and my housemates' favourite joint was La Taqueria. I initially tried to play with data from Citibikes but the dataset was too large for my computer to visualise so I went on Kaggle to obtain a data interesting to me and found this.

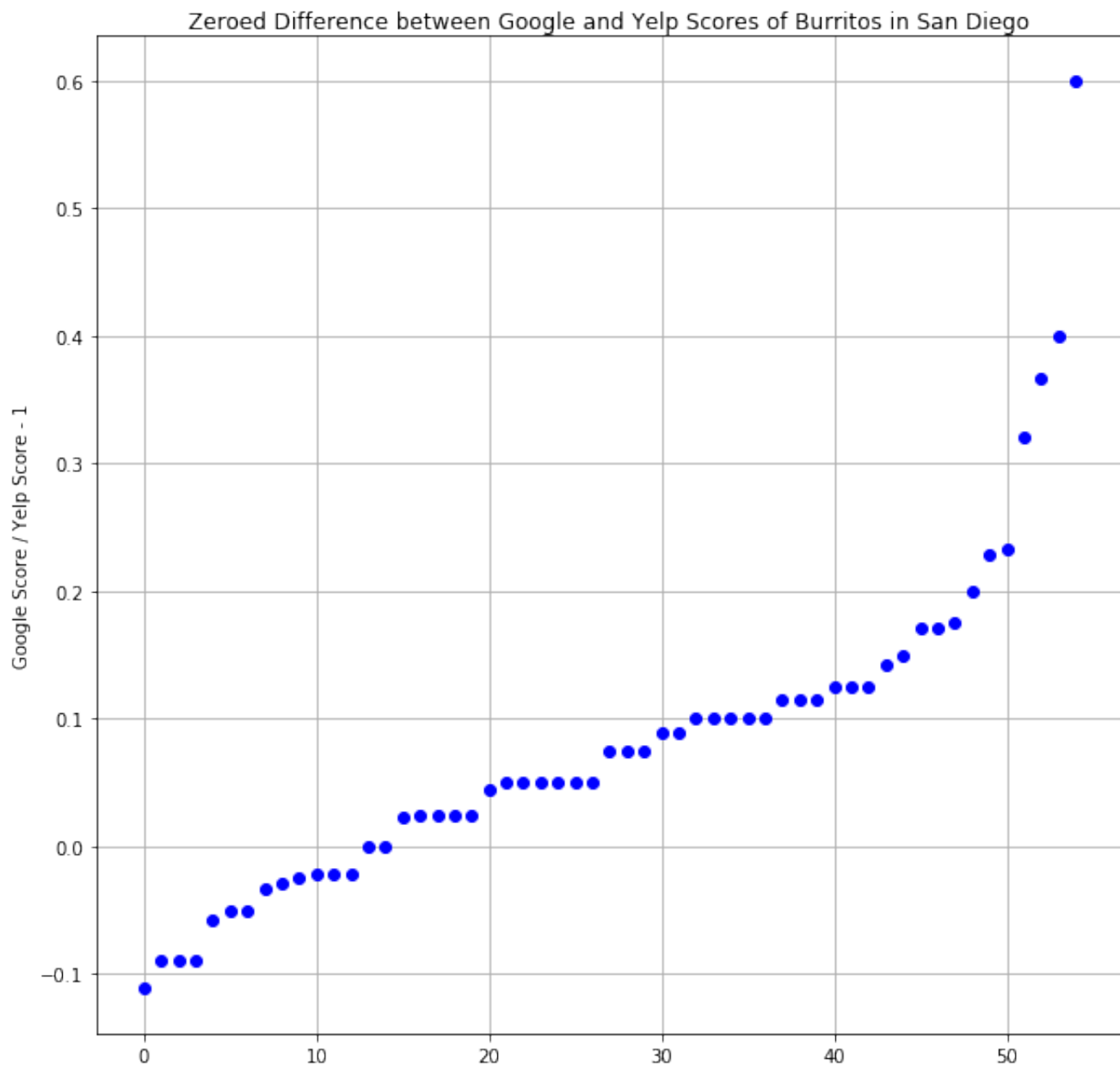
This dataset<sup>1</sup> is a collection of reviews of burrito spots in San Diego. It is actively added to by reviewers from the area. As the reviewers are from the UCSD Graduate Neuroscience program and their friends, I believe that the data is relatively clean and well-collected. The aim is to find burritos consistently rated the best by multiple reviewers due to high variance of the burritos and subjectively of each person's tastebuds and also figure interesting trends relating features of burritos to their overall ratings. I love burritos so I would want to know where and which burrito to get if I am ever in San Diego.

The main problem with the dataset is the sparsity in terms of certain columns such as neighbourhoods since the list of burritos and burrito joints have expanded since the conception of the website and the subjectively of the scores for the features given by reviewers.

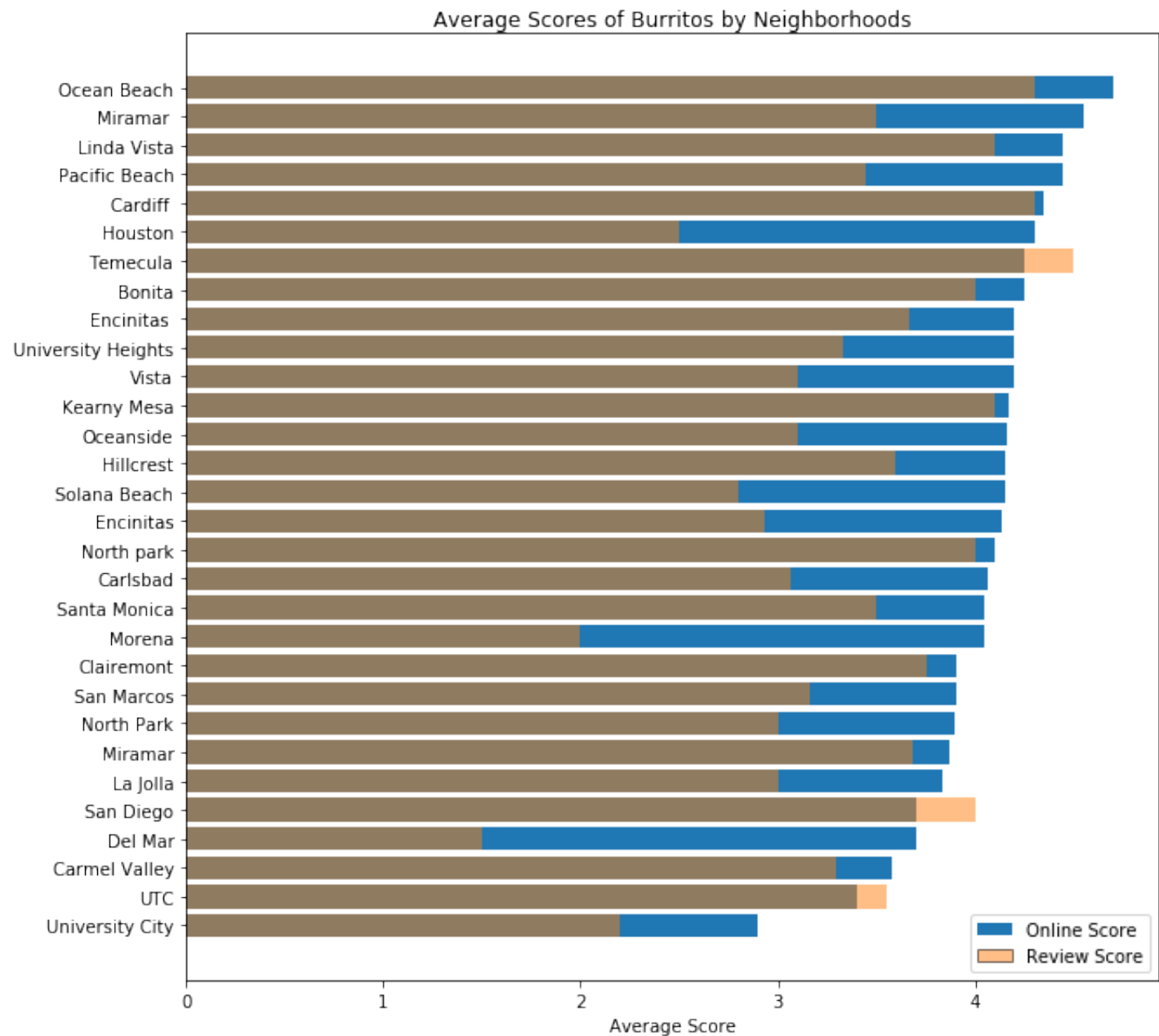
---

<sup>1</sup><https://srcole.github.io/100burritos/>

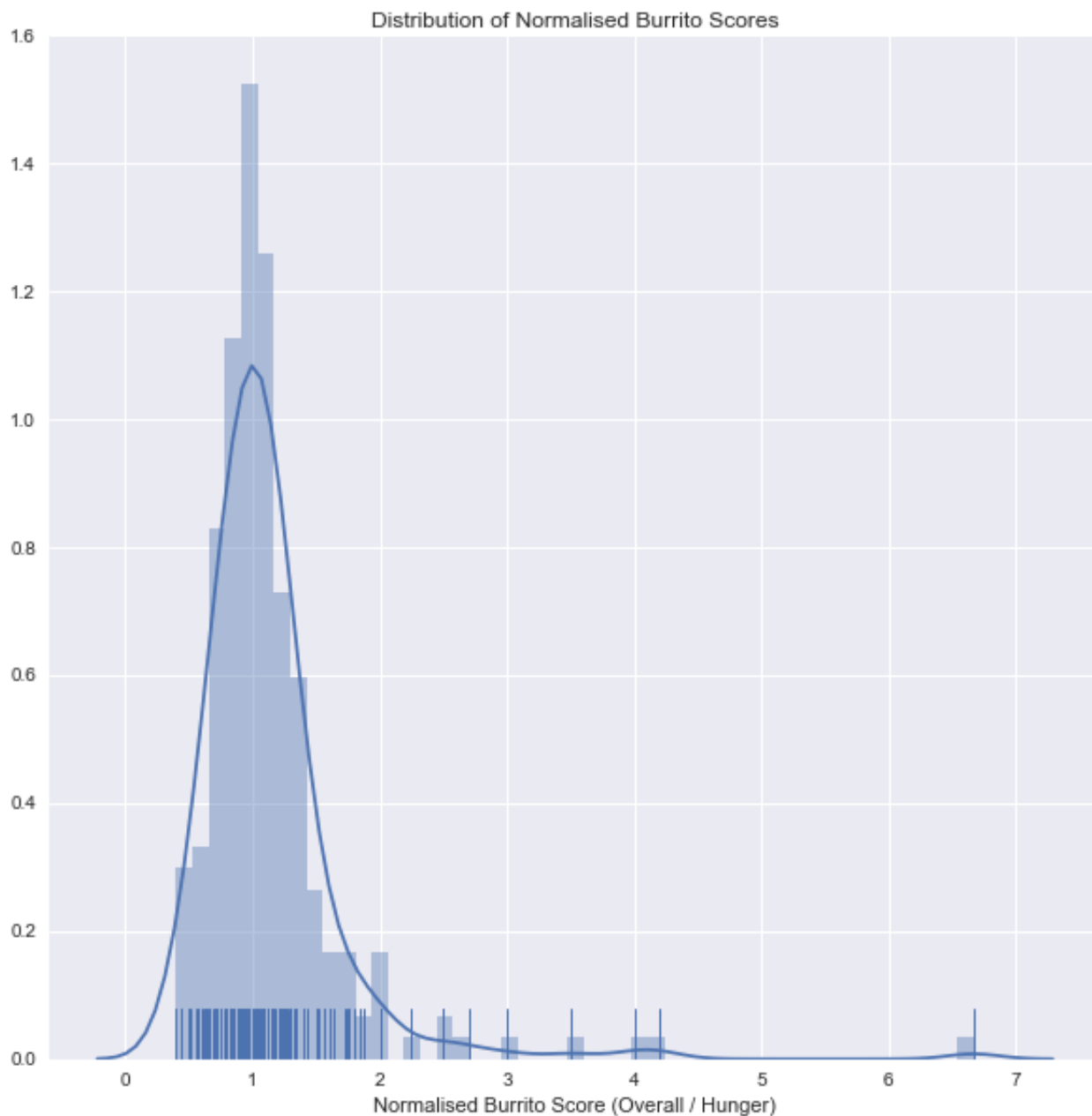
b



I have noticed a trend with Google reviews that the mean score of most food places, especially in San Francisco, tend to be around 4. I have also seen the Yelp reviewers are also more parsimonious when reviewing food. Therefore, I was curious about how much of a bonus (if it actually is a bonus) do Google reviews have over Yelp reviews.



From the above question, a logical step would be to ask how good are the tastebuds of people who review food on Google and Yelp compared to the people participating in this study. It was interesting to see that there is a large difference between online reviews scores and the scores from this study which could be due to the people participating in study being seasoned burrito eaters who have much higher standards than the general public.



I feel

that I am an impartial judge when it comes to food regardless of my state of hunger as long as I am not too full but I know intuitively that I am likely to be quite subjective. Since the reviewers stated their hunger levels, it would be good to figure how often hunger plays a substantial role in improving the taste of food among the reviewers so I can judge the quality of their recommendations.

(P.S. Working on this has me extremely hungry > . <)