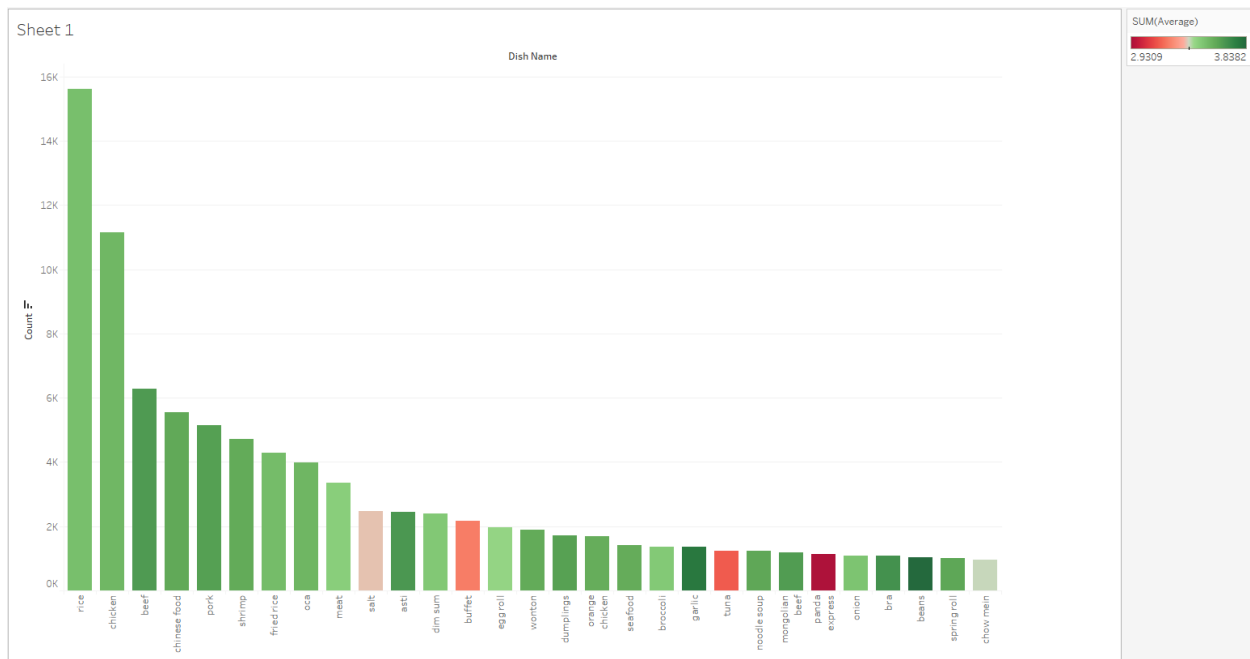## Note

The cuisine used in this task is 'Chinese'
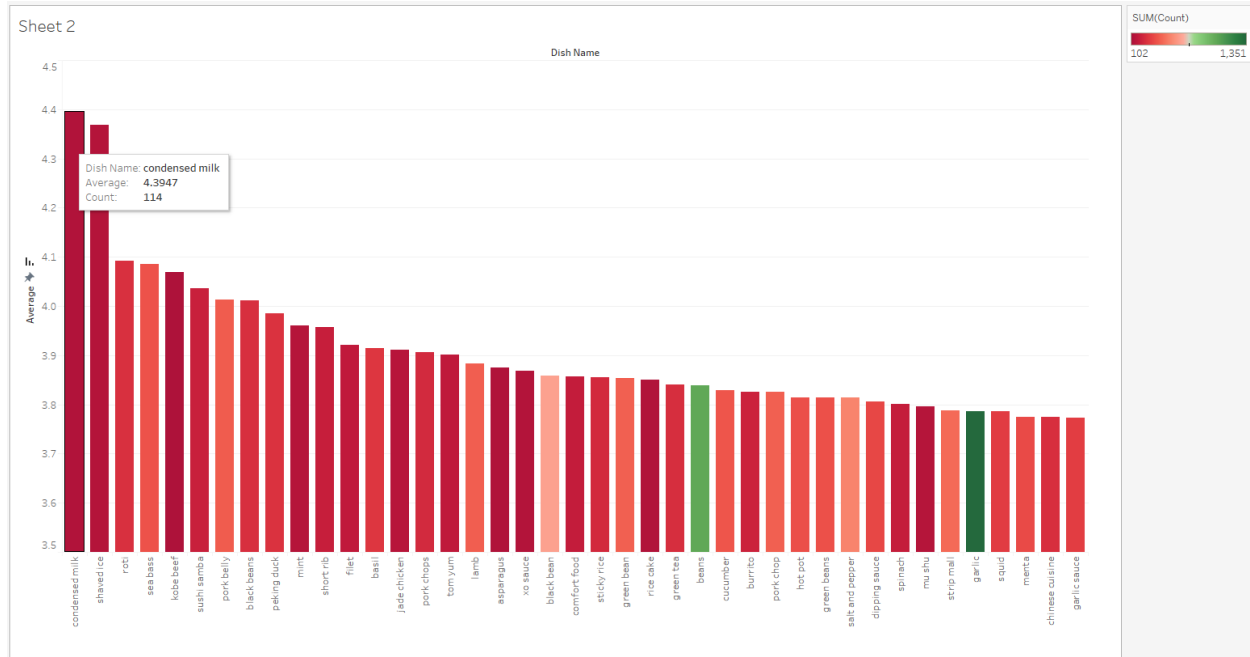
Task 4:

In task 4, dishes are ranked based on their amount of time mentioned in all reviews in category 'Chinese'.

Colors in this chart represents average ranking of corresponding dishes. With higher ranking toward green and lower ranking towards red.

The highest ranked dishes are 'rice', 'chicken', 'beef' and so on. This is obviously not very meaningful. The reason is that the single words are more likely to be mentioned in reviews, like 'fried rice' is also contributing to dish 'rice'.

Another approach is to rank by average ratings. See visualization below:

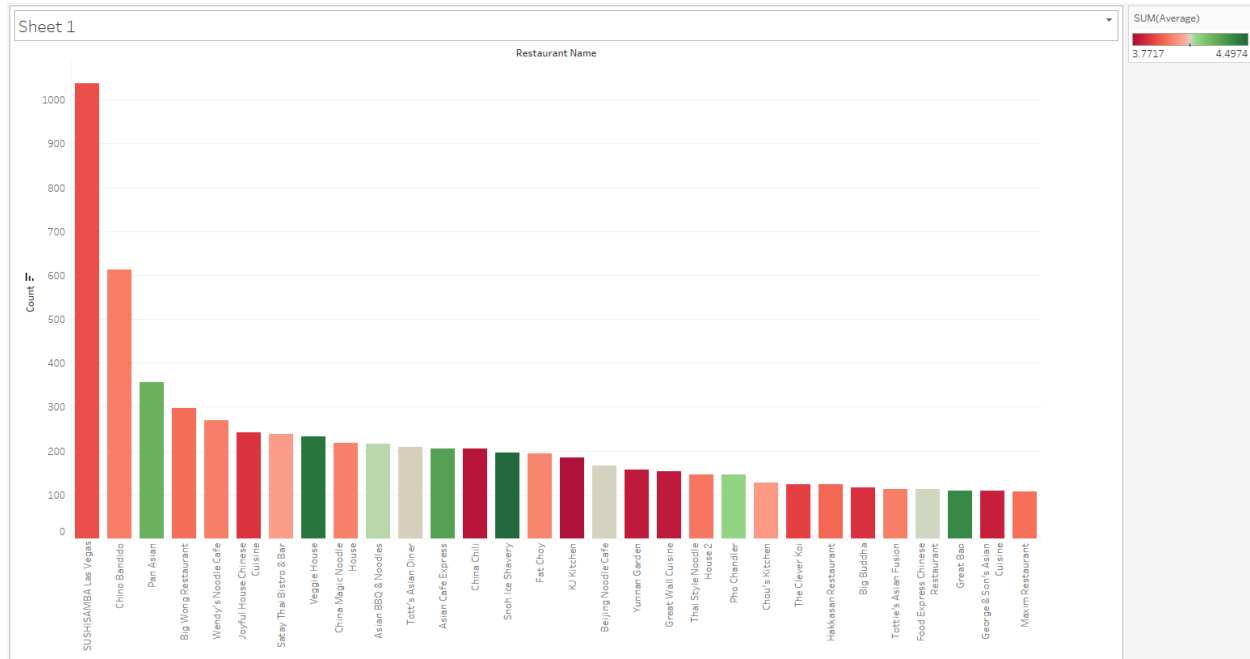Please notice that the y-axis's range is from 3.5 to 4.5

Colors in this chart represents amount of time mentioned in all reviews. To remove outliers, I chose only dishes with more than 100 reviews.

Here condensed milk is the highest ranked dish.
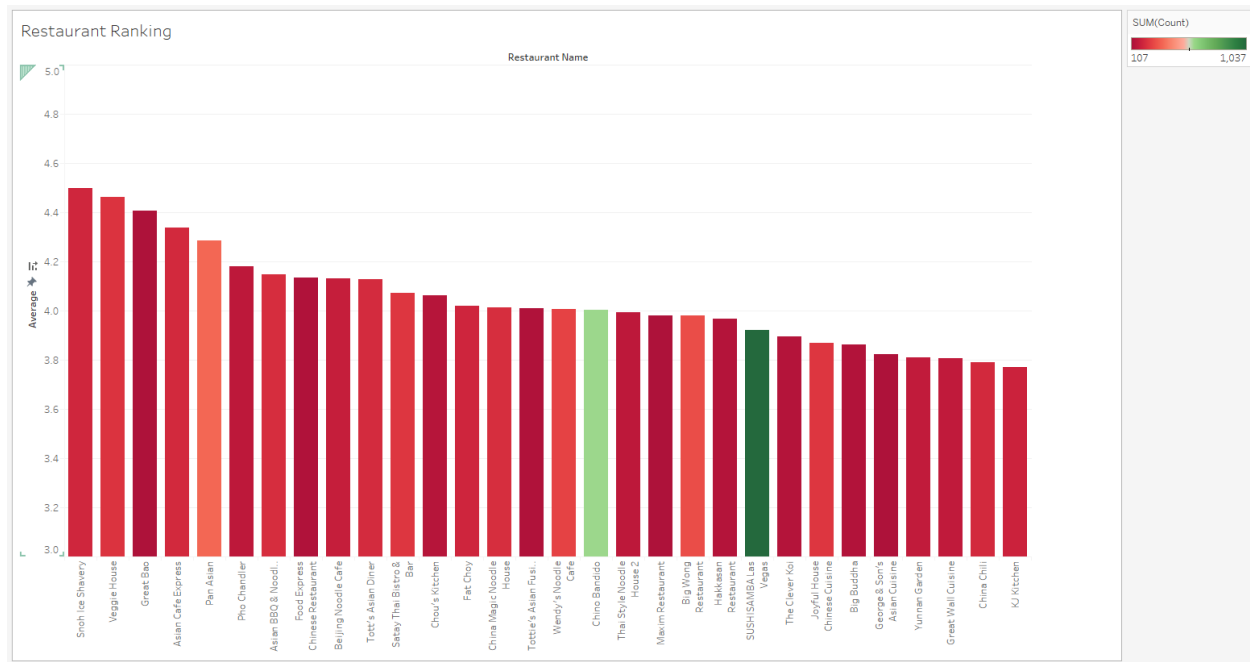
Task 5:

To find the highest ranked restaurant for a dish, I choose to rank by the amount of time this dish mentioned in all reviews, providing the average rating. See visualization below:

The dish I queried is 'Chicken'

Colors in this chart represents the average rating of restaurants. With higher ratings goes toward green and lower ratings goes toward red.

The highest rated restaurant is a Sushi house. I personally cannot judge this result, but according to previous task, I also created a visualization based on average rating:

Please notice that the y-axis's range is from 3.5 to 5.0. This chart provides another way of ranking all restaurants.

Notes:

One thing I would like to introduce for your replication is that I used BOW to decide whether a dish is really mentioned in a review. The initiation of this concern is that there are two dishes in the provided dish list:

Cucumber yogurt

Yogurt cucumber

These two dishes are the same. The way I used to check whether a dish is mentioned in a review is to directly look it up in the string. This will lead them to different dishes. Thus, I decided to use BOW to query dish names, which is to check whether every word in this dish name appeared in this review, instead of directly check the entire phrase.