



Zomato - Best of Melbourne Analysis

Dea Avega Editya

Master of BA

Muhammad Soban Qasim

Master of BA

Abhishek Sinha

Master of BA

Vinny Vu

Master of BA

Report for

ETC5513 Collaborative and Reproducible Practices

10 June 2020

**Our consultancy
add names &
add names**

📞 (03) 9905 2478
✉️ questions@company.com

ABN: 12 377 614 630

Contents

Introduction

For this assignment we will be looking at scrapping data from Zomato, one of the largest food aggregators in the world Zomato.com (2020b). Zomato provides information on restaurants including, price, location, reviews and more. For the purposes on our analysis we will be looking at the best of Melbourne list Zomato.com (2020a) to gain insight into some of the most popular restaurants in Melbourne.

This report will include four sections looking at:

- Is expensive food better than cheap food?
- Popularity of different cuisine in Melbourne
- Exploring gluten free food/restaurants in Melbourne
- Distribution of the top 10 restaurants in Melbourne

In this report we have used the following packages Wickham et al. (2019), Wang (2020), Zhu (2019), Xie (2020), Wickham et al. (2020), Wickham (2016), Cheng, Karambelkar, and Xie (2019), Bivand, Keitt, and Rowlingson (2019), Wickham, Hester, and Francois (2018), Appelhans et al. (2020), Garnier (2018).

Is expensive food better than cheap food?

In this section we will be using the Zomato Best of Melbourne to explore whether expensive food is “better” than cheap food. To conduct this analysis, we will be looking at:

- What is considered cheap and expensive food
- What the breakdown of price ranges of the restaurants is
- Comparing prices and aggregate ratings
- Determining whether there is a relationship between prices and ratings

0.0.1 What is considered cheap and expensive food?

From Figure 1 we can see Zomato breaks down price ranges into 4 categories with 4 being the most expensive to 1 being the least expensive.

From Table 1 we can see the range of prices that make up each price category.

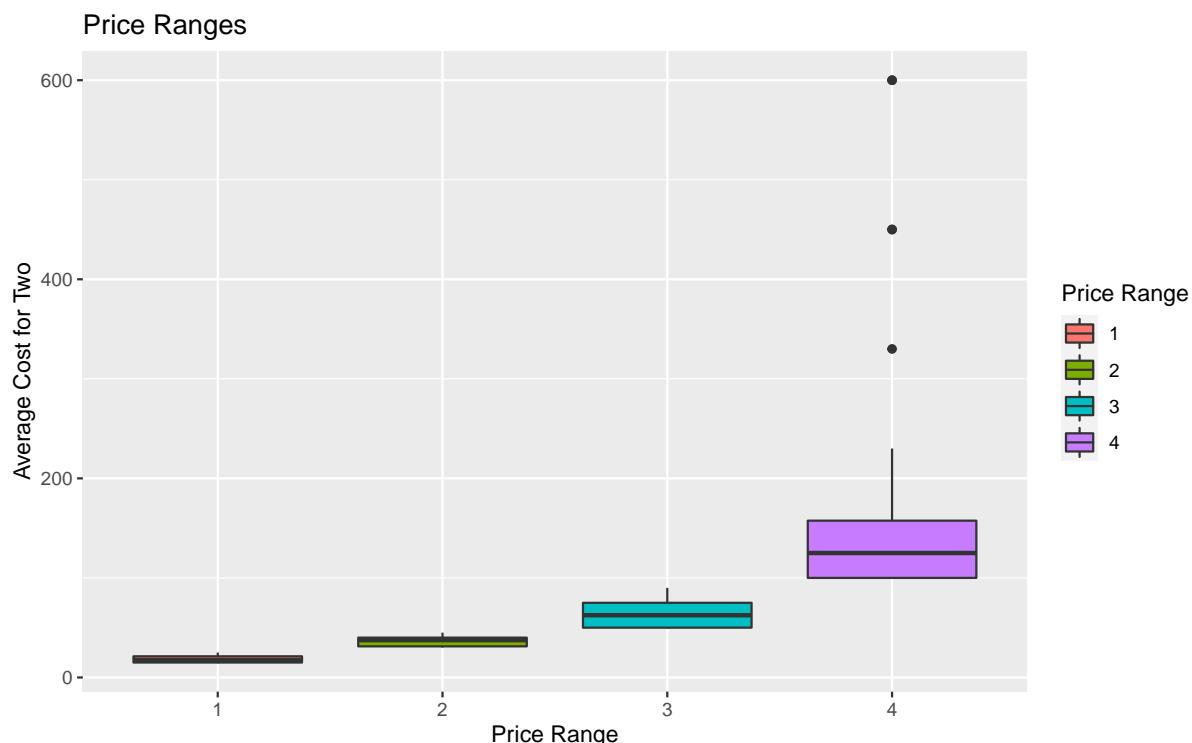


Figure 1: Price Ranges

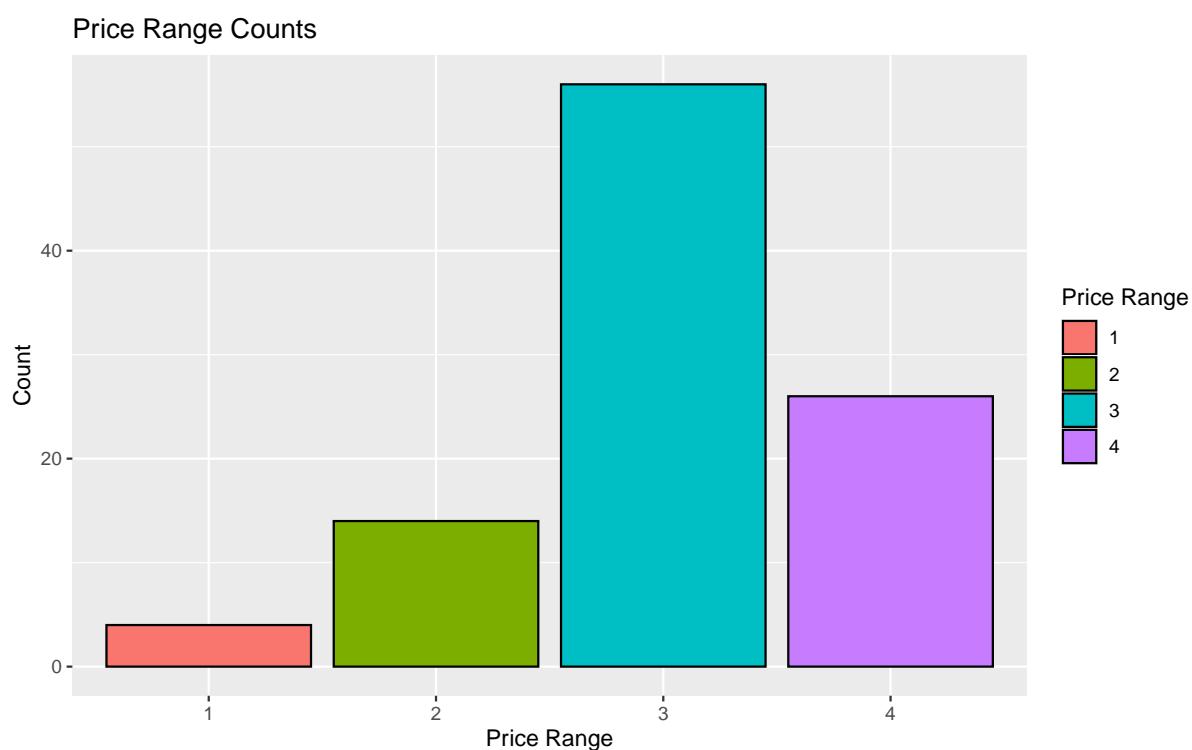


Figure 2: Price Range Counts

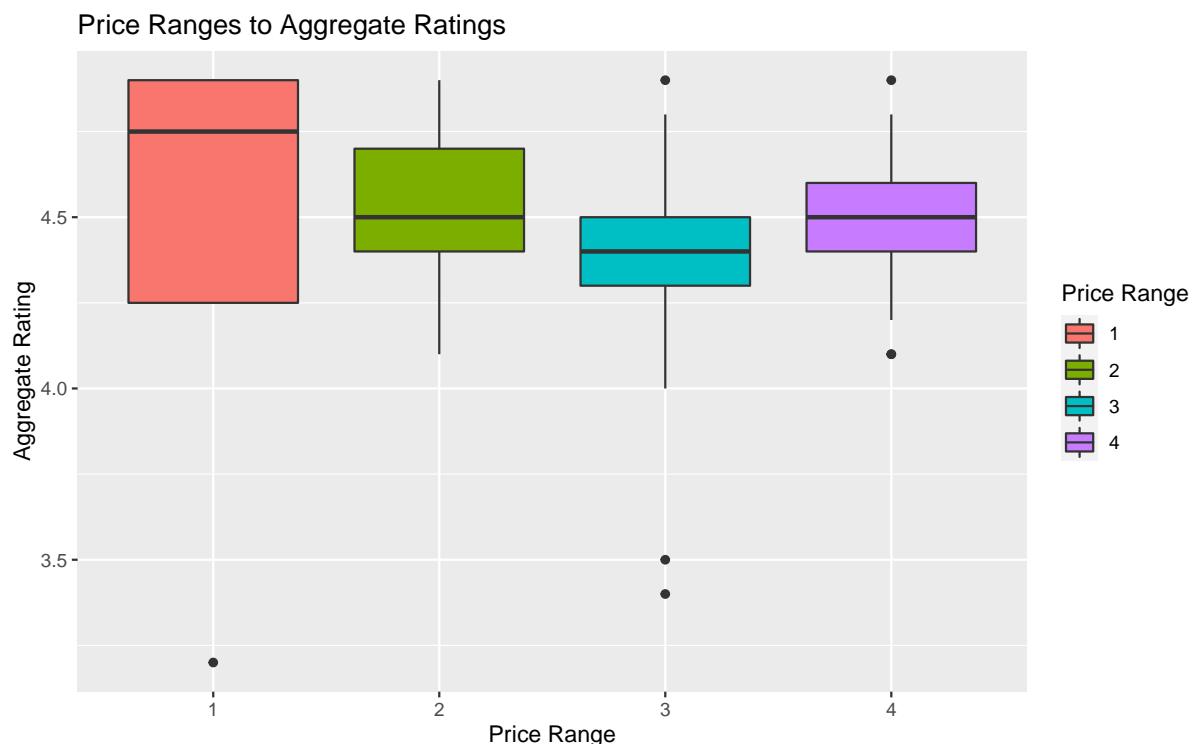
Table 1: Price Ranges

Price Range	Maximum	Median	Minimum	Range
1	25	17.5	15	10
2	45	37.5	30	15
3	90	62.5	50	40
4	600	125.0	100	500

0.0.2 What the breakdown of price ranges of the restaurants?

From Figure 2 We can see price range 3 has the most counts. One thing to note however is price range 3 has a range of \$40 compared to 1 and 2 which is \$10 and \$15, respectively. It may be unfair to conclude there is a higher count of more expensive food as the ranges included in the expensive ranges are larger than the lower ones.

0.0.3 Comparing prices and aggregate ratings

**Figure 3: Price Ranges to Aggregate Ratings**

From Figure 3 we can see the median aggregate rating for price range 1 is the highest. In fact, we can see a large amount of lower priced restaurants outranking higher priced restaurants in terms of aggregate ratings.

0.0.4 Is there a relationship between prices and ratings

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

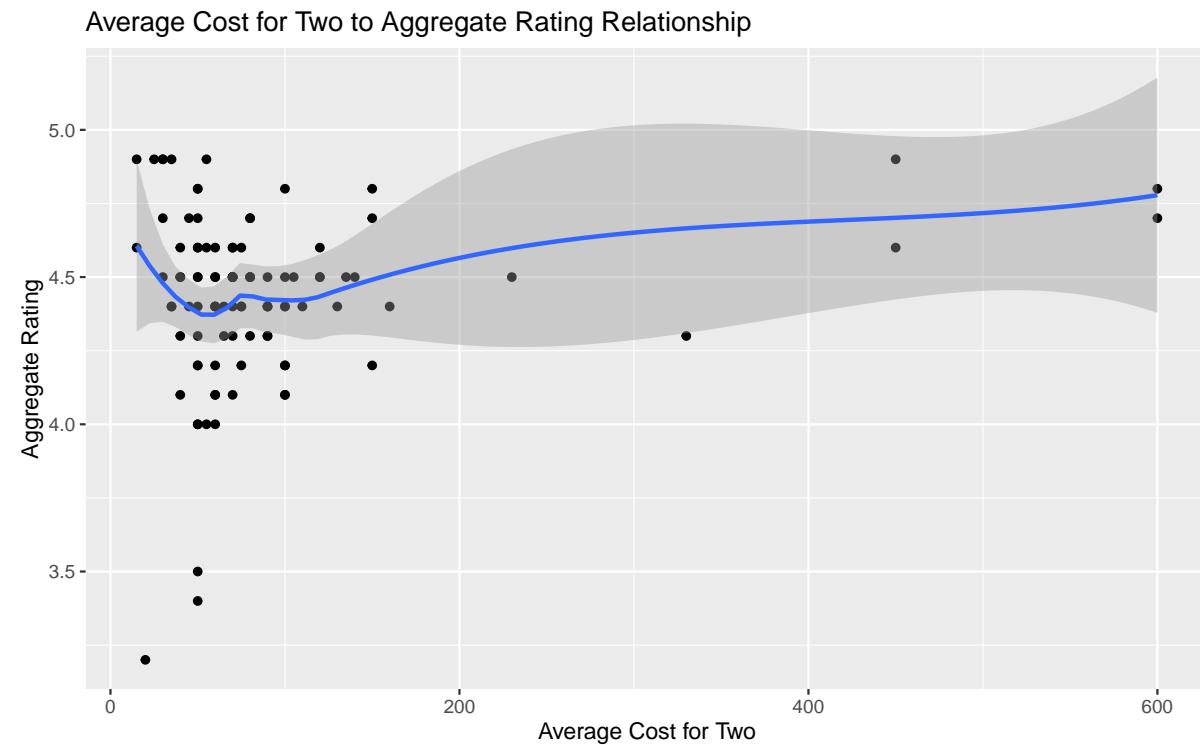


Figure 4: Average Cost for Two to Aggregate Rating Relationship

From Figure 4 we can see there is not a strong relationship between the average cost for two and the aggregate ratings suggesting there is no relationship between price and food quality. We see a slight upwards trend at the higher price levels. However, this is only based on 5 restaurants with drastically higher prices. We see most of the restaurants fall below the \$200 price point. This might even suggest lower priced restaurants out rank the higher priced restaurants due to the higher count.

Table 2: Rating Ranges

Rating Text	Maximum	Minimum
Average	3.4	3.2
Excellent	4.9	4.5
Good	3.5	3.5
Very Good	4.4	4.0

Table 2 shows range of aggregate ratings that make up each rating levels.

From Figure 5 we can see restaurants in the second and fourth price category all rank in either the very good or excellent category. 3 has the highest count in all categories, however, this may be due to the higher count. We see all price levels have restaurants with high ratings.

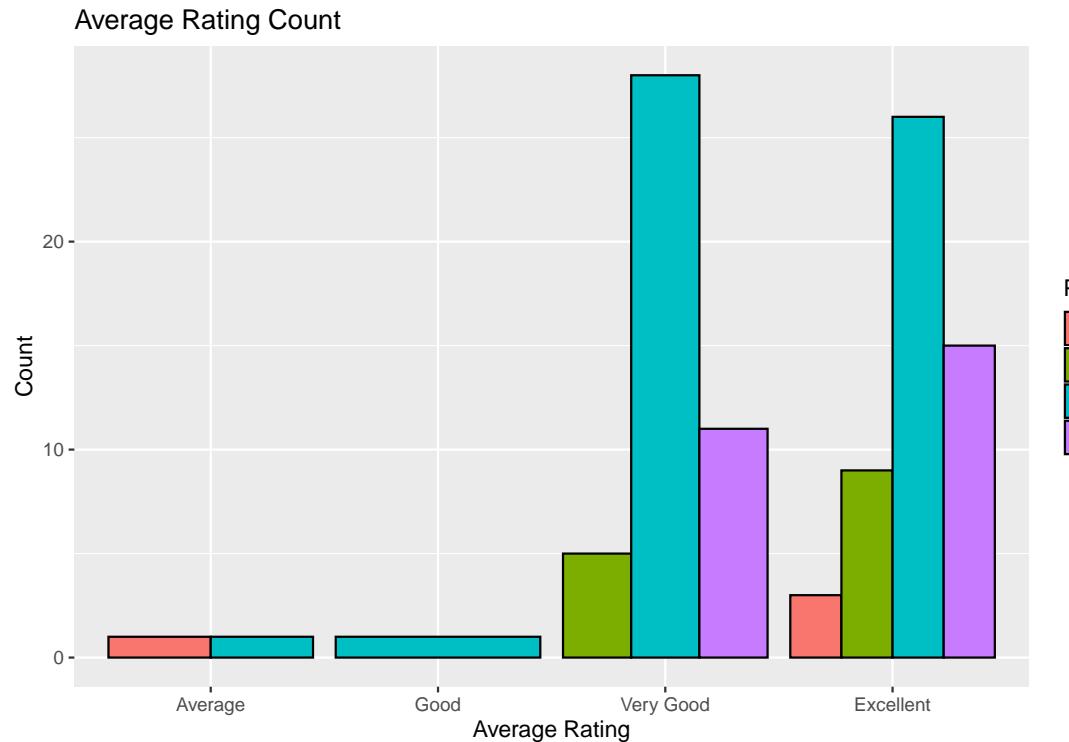


Figure 5: Average Rating Count

0.0.5 Conclusion

From this analysis we find there is no real relationship between restaurant prices and quality suggesting expensive food is not better than cheap.

References

- Appelhans, T, F Detsch, C Reudenbach, and S Woellauer (2020). *mapview: Interactive Viewing of Spatial Data in R*. R package version 2.7.8. <https://CRAN.R-project.org/package=mapview>.
- Bivand, R, T Keitt, and B Rowlingson (2019). *rgdal: Bindings for the 'Geospatial' Data Abstraction Library*. R package version 1.4-8. <https://CRAN.R-project.org/package=rgdal>.
- Cheng, J, B Karambelkar, and Y Xie (2019). *leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library*. R package version 2.0.3. <https://CRAN.R-project.org/package=leaflet>.
- Garnier, S (2018). *viridis: Default Color Maps from 'matplotlib'*. R package version 0.5.1. <https://CRAN.R-project.org/package=viridis>.
- Wang, E (2020). *zomato: API to Zomato*. R package version 0.1.0.
- Wickham, H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.

- Wickham, H, M Averick, J Bryan, W Chang, LD McGowan, R François, G Grolemund, A Hayes, L Henry, J Hester, M Kuhn, TL Pedersen, E Miller, SM Bache, K Müller, J Ooms, D Robinson, DP Seidel, V Spinu, K Takahashi, D Vaughan, C Wilke, K Woo, and H Yutani (2019). Welcome to the tidyverse. *Journal of Open Source Software* 4(43), 1686.
- Wickham, H, R François, L Henry, and K Müller (2020). *dplyr: A Grammar of Data Manipulation*. R package version 1.0.0. <https://CRAN.R-project.org/package=dplyr>.
- Wickham, H, J Hester, and R Francois (2018). *readr: Read Rectangular Text Data*. R package version 1.3.1. <https://CRAN.R-project.org/package=readr>.
- Xie, Y (2020). *knitr: A General-Purpose Package for Dynamic Report Generation in R*. R package version 1.28. <https://yihui.org/knitr/>.
- Zhu, H (2019). *kableExtra: Construct Complex Table with 'kable' and Pipe Syntax*. R package version 1.1.0. <https://CRAN.R-project.org/package=kableExtra>.
- Zomato.com (2020a). *Best of Melbourne*. <https://www.zomato.com/melbourne/great-food-no-bull>.
- Zomato.com (2020b). *Shaping the future of food*. <https://www.zomato.com/about>.