# Report

Group B

21/10/2020

## Data collection

The data collection process was more difficult than expected. There were many restaurants that we had to exlude because they did not have a 0.5L beverage or a margarita pizza. For the beverage we always tried to pick Coca Cola or Pepsi when available or otherwise water. If those were not available, we took lemonade even though it is usually more expensive. For the extra measures we recorded, numbers of pizzas they offered, opening hours (which we converted into minutes opened in this analysis), as well as rating on Google and the number of ratings they received. The number of ratings is important since it gives the rating more credibility. We also recorded the longitude and latitude to calculate the distance from CEU and give students another variable if they want to pick one of these places for lunch or dinner.

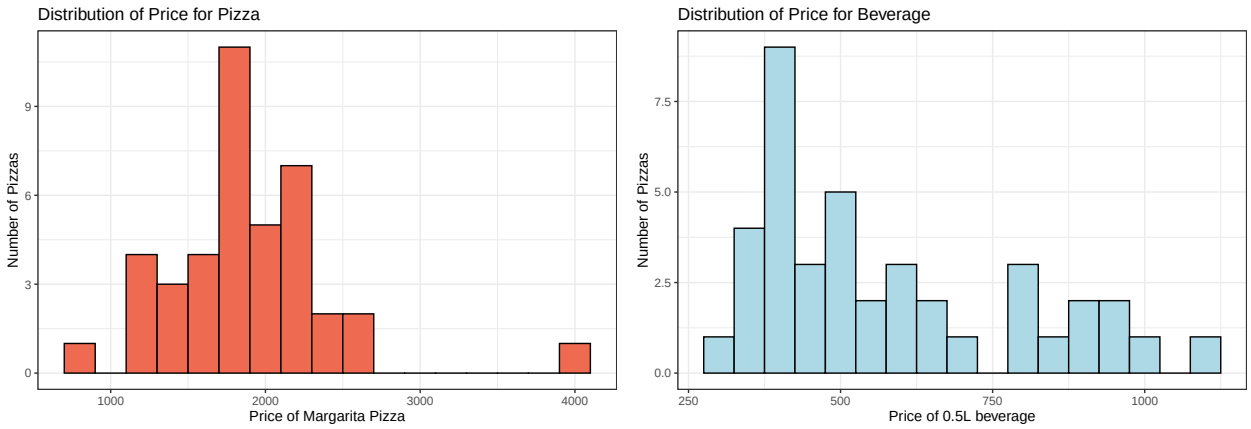
## Calculations for opening hours and distance from CEU

## Descriptive Statistics

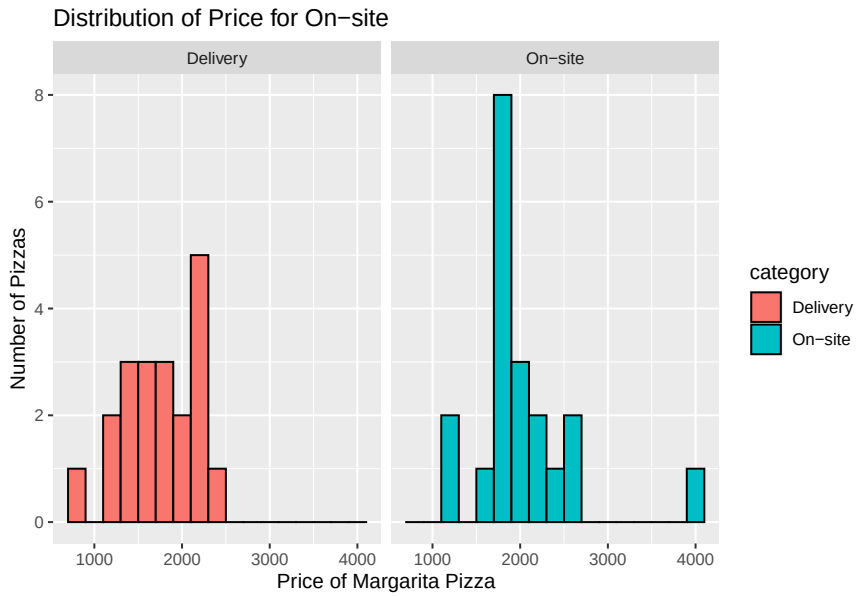
| n | mean | median | min | max | sd | skew |
|---|---|---|---|---|---|---|
| 40 | 1916.525 | 1885 | 890 | 3920 | 510.6978 | 1.3210519 |
| 40 | 573.500 | 495 | 280 | 1090 | 215.2464 | 0.8007159 |

# Distributions

## Price of pizza margarite and beverage




## Price distribution on-site vs delivery




# T-test

## Hypothesis

Null mean(on-site)=mean(delivery) Alternative mean(on-site) not = mean(delivery)

```
## Warning: package 'pander' was built under R version 4.0.3
```

Table 2: Welch Two Sample t-test: `price_offline` and `price_online`

| Test statistic | df | P value | Alternative hypothesis | mean of x | mean of y |
|---|---|---|---|---|---|
| 1.457 | 33.89 | 0.1542 | two.sided | 2033 | 1800 |

## Summary and Conclusion

Overall the distribution of price of the beverages has a larger range than the prices of the pizza. And overall the prices of the pizzas on-site were more expensive then the ones for delivery, although with the delivery fee this may be evened out. The t-test also comes to the same conclusion where