Report on Snickers and Coca-Cola Prices

Lagos Team

1 November 2021

Contents

Introduction

HERE COMES THE MOTIVATION WHY THIS IS A MEANINGFUL PROJECT AND WHAT IS THE MAIN GOAL!

Data

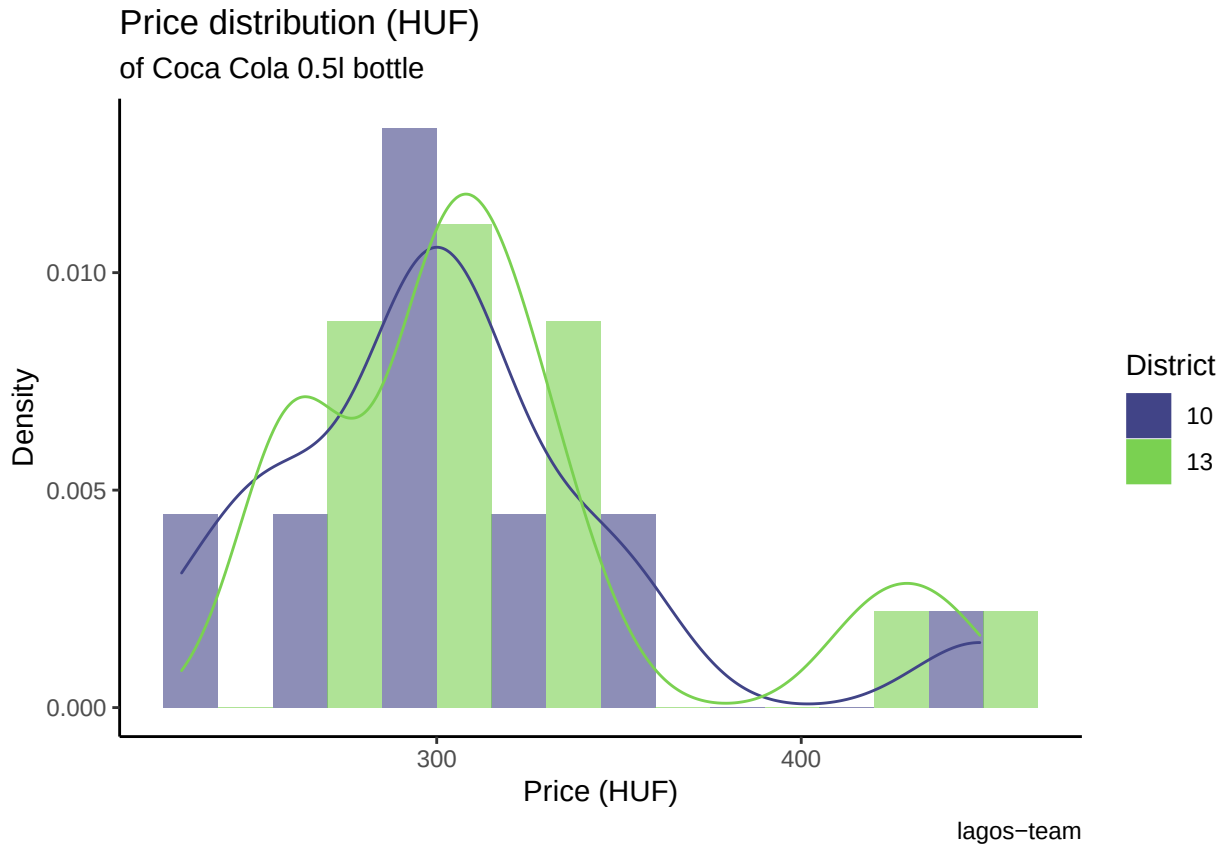
HERE COMES A DETAILLED EXPLANATION ABOUT WHERE THE DATA COMES FROM AND IF IT IS REPRESENTATIVE OR NOT.

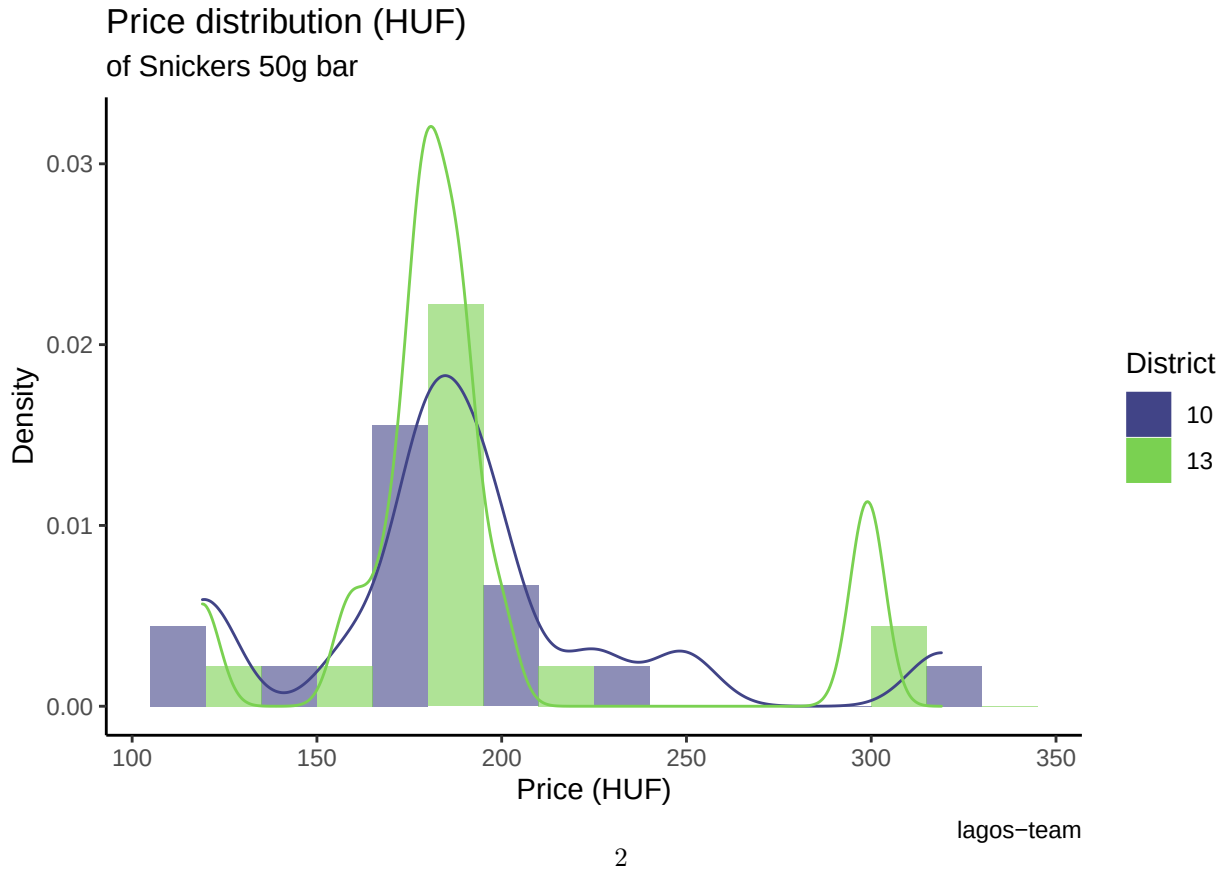
Table 1: Descriptive statistics of prices by district

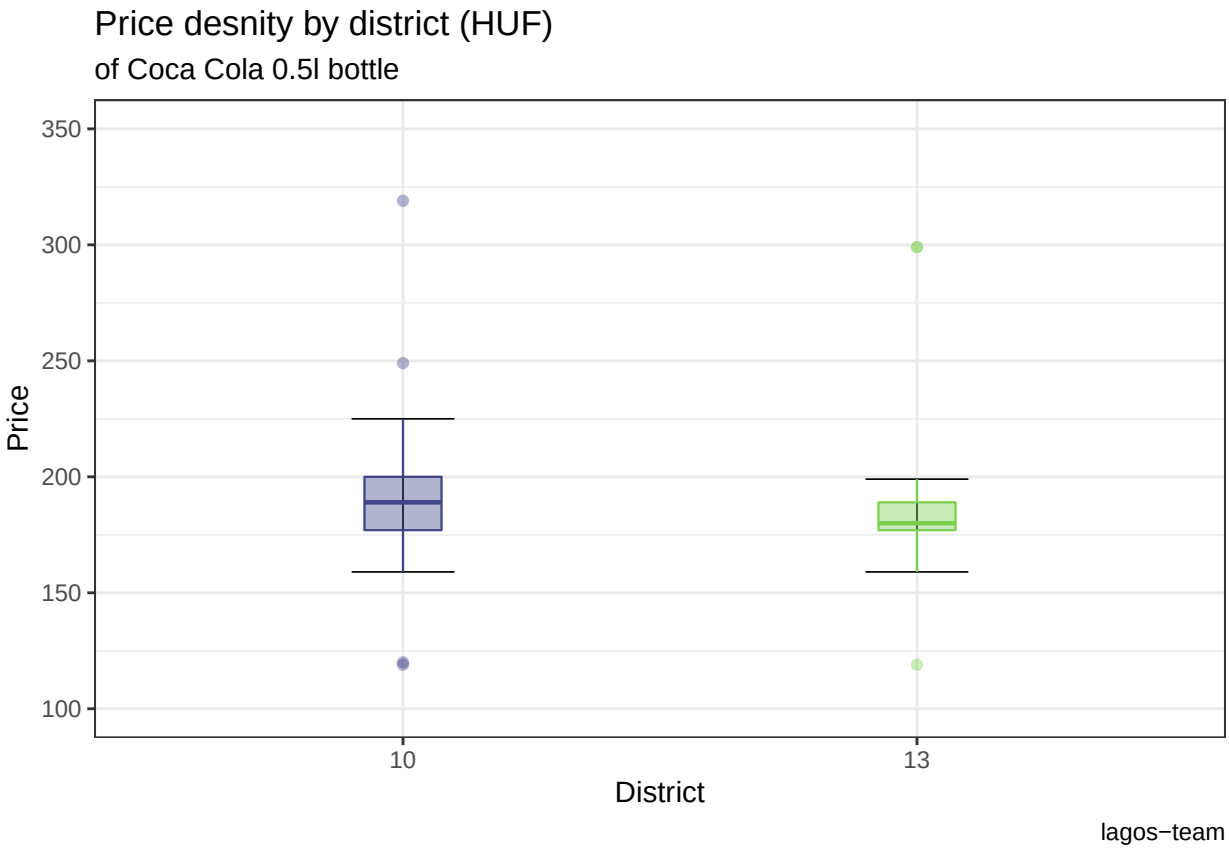
Product	District		NUnique	PercentMissing	Mean	SD	Min	Max	P05	Median
Coke 0.5l	10	Price	11	0	304.53	52.43	230	449	244.00	299.00
	13	Price	11	0	314.40	52.98	259	439	259.00	300.00
Snickers 50g	10	Price	12	0	191.47	48.54	119	319	119.70	189.00
	13	Price	10	0	192.60	46.96	119	299	147.00	180.00

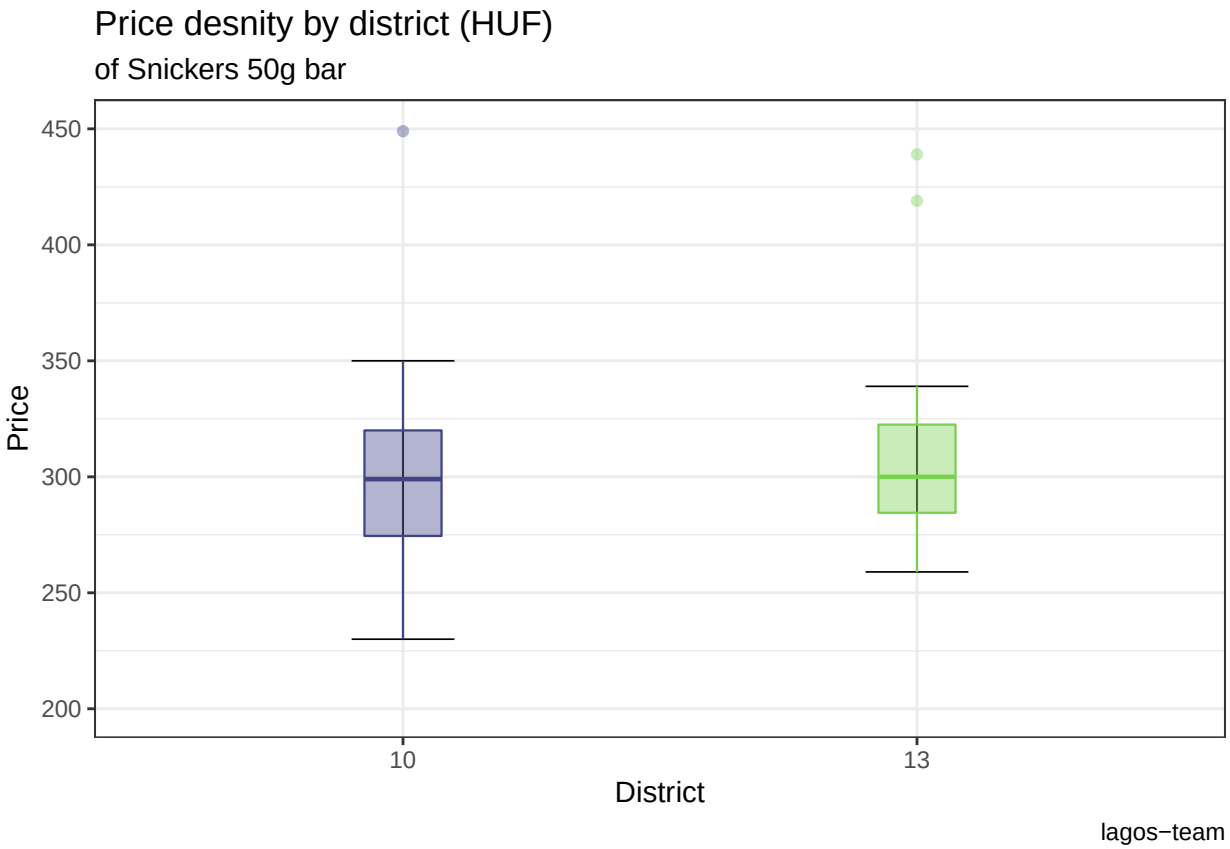
Intuitively the mean of both products are slightly higher in the inner 13th district. Also there is less variation in price of S

Viz


Price distribution (HUF)
of Coca Cola 0.5l bottle
Density
0.010
0.005
0.000
300
400
Price (HUF)
District
10
13
lagos−team

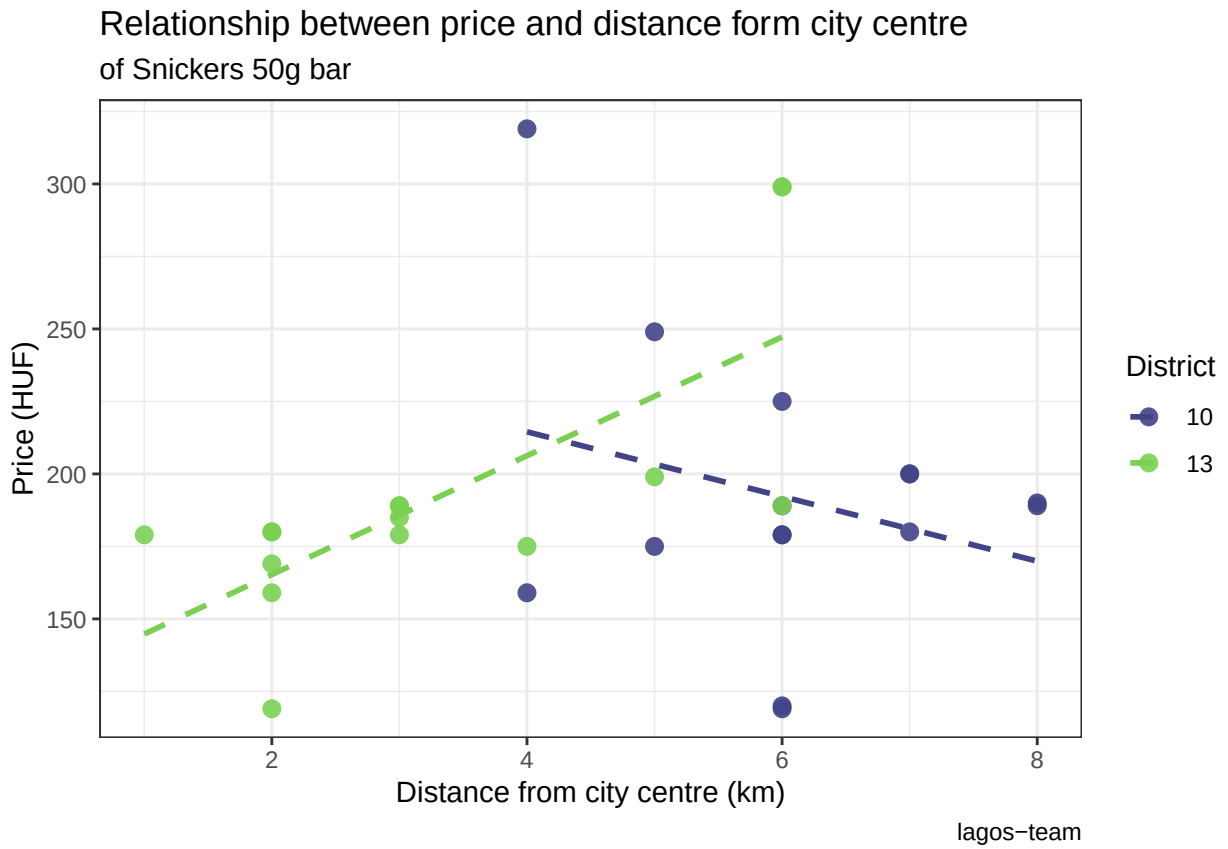

Price distribution (HUF)
of Snickers 50g bar
Density
0.03
0.02
0.01
0.00
100
150
200
250
300
350
Price (HUF)
District
10
13
lagos−team


Price desnity by district (HUF)
of Coca Cola 0.5l bottle
350
300
250
200
150
100
Price
10
13
District
lagos−team

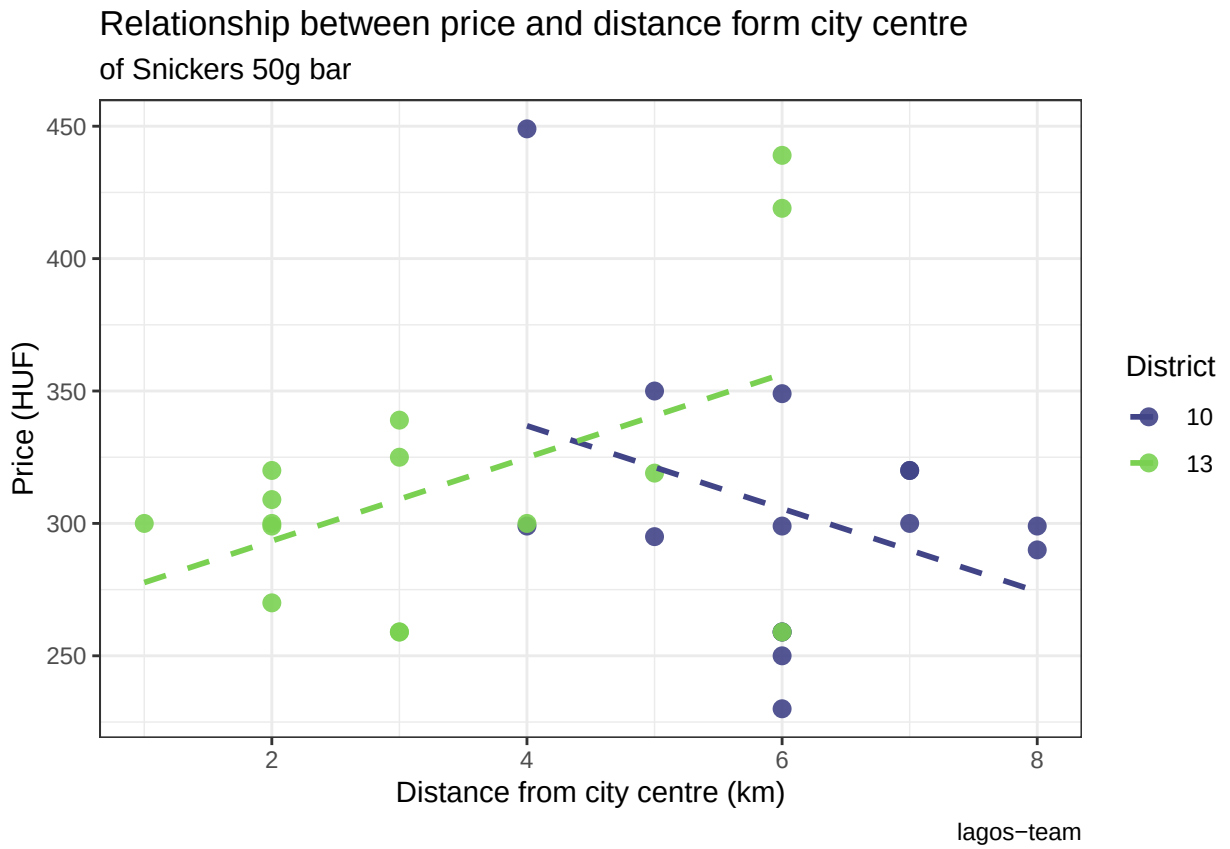


Price and distance

```
## `geom_smooth()` using formula 'y ~ x'
```




```
## `geom_smooth()` using formula 'y ~ x'
```



T-test by district

```
# 2 sample t-test

# Snickers
t.test(price_huf ~ district, data = filter(df_clean,item_name=='Snickers 50g')) # cannot reject H0 at 5
```

```
##
##  Welch Two Sample t-test
##
## data:  price_huf by district
## t = -0.064988, df = 27.97, p-value = 0.9486
## alternative hypothesis: true difference in means between group 10 and group 13 is not equal to 0
## 95 percent confidence interval:
##  -36.85737  34.59070
## sample estimates:
## mean in group 10 mean in group 13
##         191.4667         192.6000
```

```
# Cola
t.test(price_huf ~ district, data = filter(df_clean,item_name=='Coke 0.5l')) # cannot reject H0 at 5%
```

```
##
```

```
##  Welch Two Sample t-test
##
## data:  price_huf by district
## t = -0.51266, df = 27.997, p-value = 0.6122
## alternative hypothesis: true difference in means between group 10 and group 13 is not equal to 0
## 95 percent confidence interval:
##  -49.29090  29.55757
## sample estimates:
## mean in group 10 mean in group 13
##         304.5333         314.4000
```