



TM

# CMPUT 191: Starbucks Latte Index

By Miguel Gil

# Introduction

The index I have created is a Starbucks Latte index

Reason: Because I was curious about the different prices around the world and I was wanted Starbucks at that moment

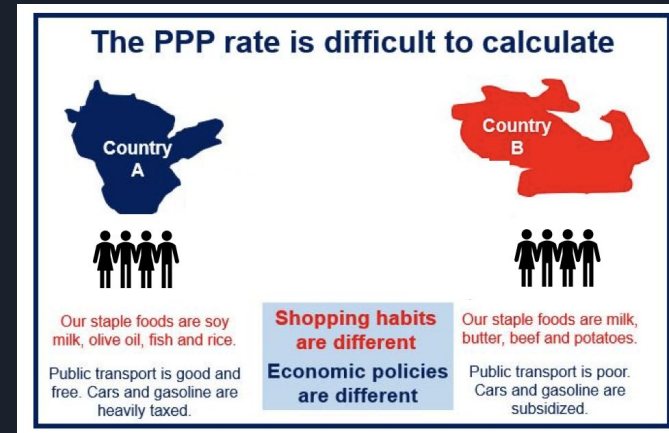
All latte prices are in Tall sizes



# Purchasing Power Parity

It is the concept that there is an exchange rate which makes it so you can buy the same amount of goods or services in any country

Problem - Differences in taxing, importing and exporting, world events, availability, etc



# Scraping Tables

Price	Country
\$4.09	Canada
HK \$43.00	Hong Kong
\$4.62	United States
A\$5.50	Australia
₹235	India
£2.25	United Kingdom
S\$6.90	Singapore
Rp44.000	Indonesia
₱145.00	Philippines
EGP 40.000	Egypt

Currency	Country	Currency code
Euro	Euro area countries	EUR
Albanian lek	Albania	ALL
Belarusian ruble	Belarus	BYN
Bosnia and Herzegovina convertible mark	Bosnia and Herzegovina	BAM
Bulgarian lev	Bulgaria	BGN
Croatian kuna	Croatia	HRK
Czech koruna	Czechia	CZK
Danish krone	Denmark	DKK
Croatian Kuna	Croatia	HRK
Georgian lari	Georgia	GEL
... (170 rows omitted)		

Currency Name	CAD to Currency	Conversion To CAD	Currency
Canadian Dollar	1.00	1.00	CAD
Australian Dollar	1.09023	0.917238	AUD
Hong Kong Dollar	5.700558	0.175421	HKD
Indian Rupee	60.199502	0.016611	INR
Indonesian Rupiah	11426.879938	8.8e-05	IDR
Philippine Peso	40.57175	0.024648	PHP
Singapore Dollar	0.992428	1.00763	SGD
British Pound	0.599883	1.666992	GBP
US Dollar	0.731776	1.366538	USD
Egypt Pound	18.28	0.055	EGP

All tables I used came from the web

Even made a function specifically for scraping

Tables:

A pricing table

A currency symbol table

A currency conversion table

# Cleaning Data

Removed the currency symbols off of the pricing with a function

Narrowed down conversion table to chosen countries

Added currency codes to conversion table

Dropped unneeded columns

Needed to convert countries in currency symbol tables

```
def clean_price(table): # removes non numeric characters
    column = table.column("Price")
    price_list = []

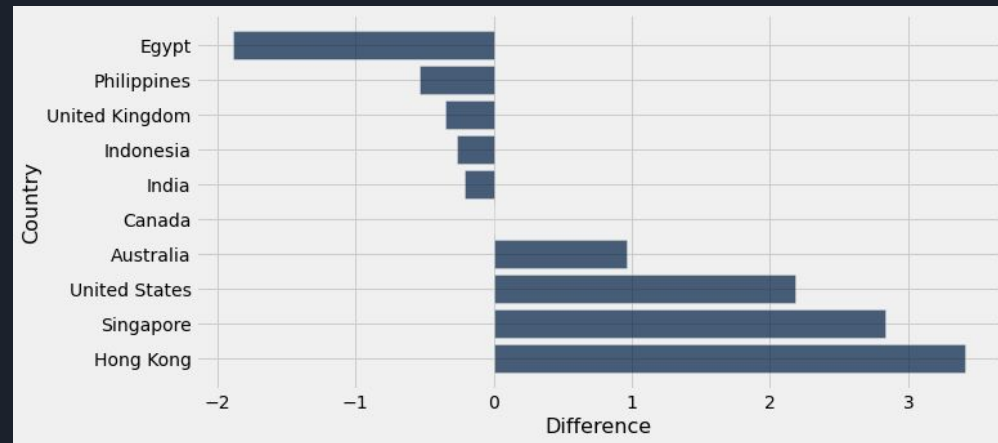
    for x in range(len(column)):
        split_string = column[x].split(".")
        split_string[0] = ''.join(filter(str.isdigit, split_string[0]))
        if table.column("Country")[x] == "Indonesia":
            rejoin = "".join(split_string)
        else:
            rejoin = ".".join(split_string)
        price_list.append(rejoin)
    return price_list
```

Currency	Price	Country	Conversion To CAD
AUD	5.50	Australia	0.917238
CAD	4.09	Canada	1.00
EGP	40.000	Egypt	0.055
GBP	2.25	United Kingdom	1.666992
HKD	43.00	Hong Kong	0.175421
IDR	44000	Indonesia	8.8e-05
INR	235	India	0.016611
PHP	145.00	Philippines	0.024648
SGD	6.90	Singapore	1.00763
USD	4.62	United States	1.366538

# Difference Visualization

Lowest is Egypt

Highest is Hong Kong





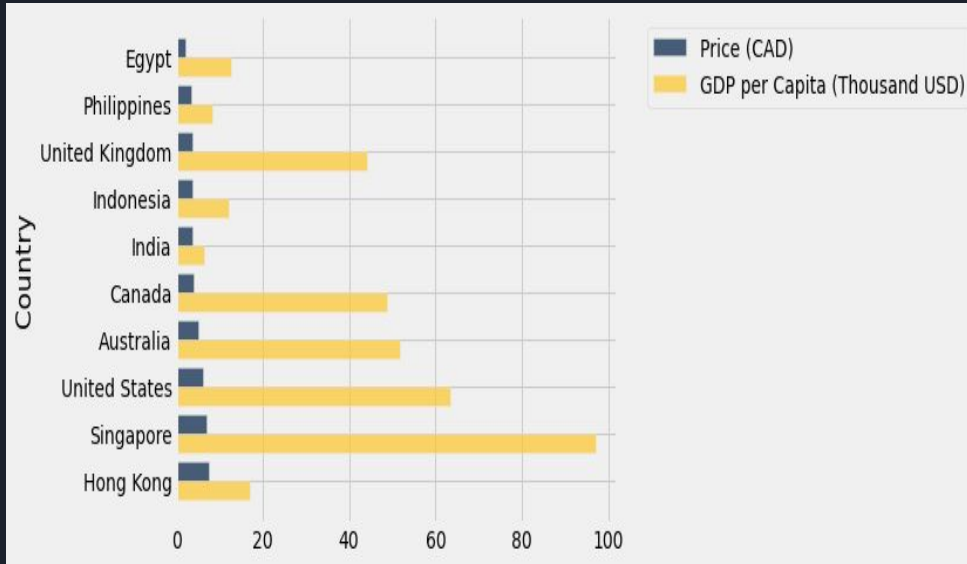
# External Factor (GDP Per Capita)

GDP per capita seemed like a factor that would likely affect prices

Consumer has more money -> more luxury item purchases

Country	GDP per Capita
Afghanistan	2390
Albania	14218
Algeria	11112
Angola	6932
Argentina	20751
Armenia	13261
Australia	51680
Austria	55218
Azerbaijan	14431
Bahrain	48766
... (174 rows omitted)	

# External Factor Visualization



Singapore has a very high GDP per capita

Hong Kong has the highest price but low GDP per capita

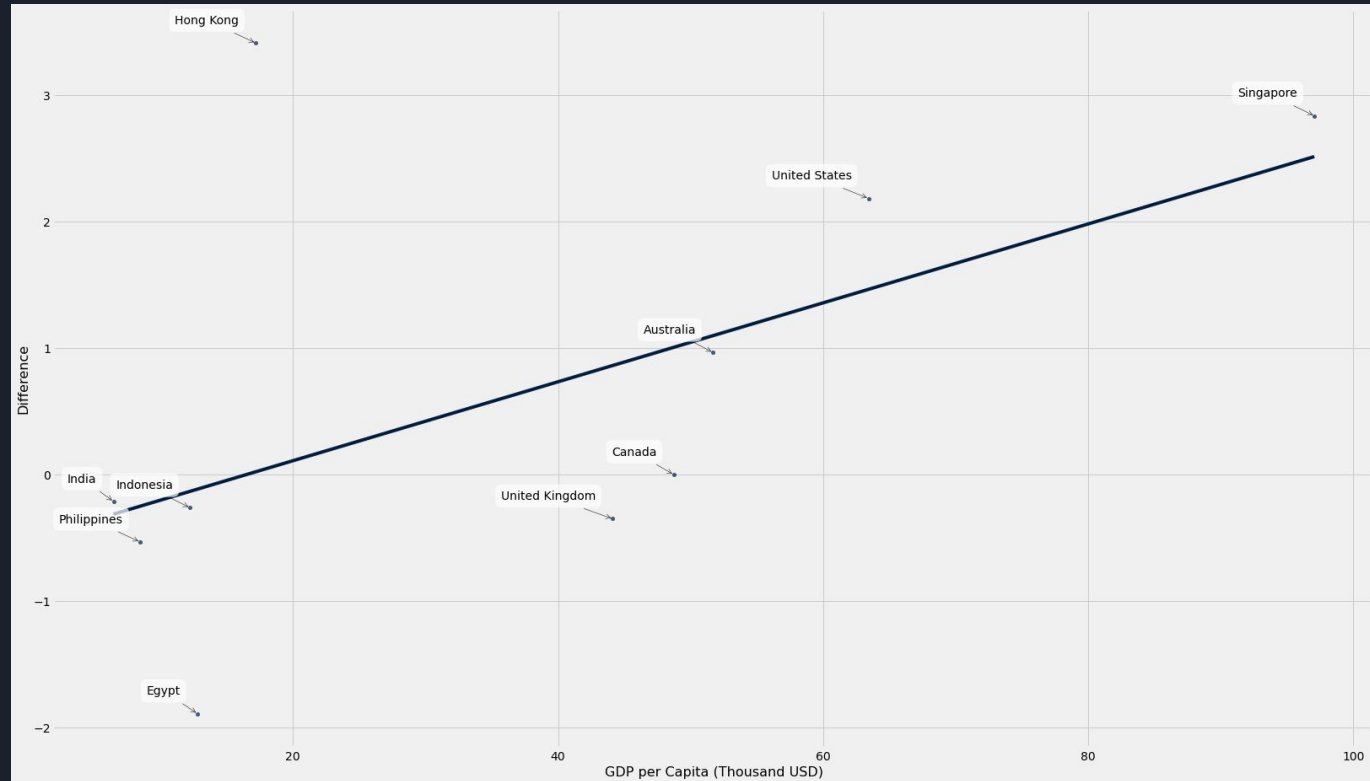
Weak trend



# Scatter Plot and Correlation Coefficient

Very weak correlation

$r$  coefficient = 0.552



# Conclusion

GDP per capita seems to have a very weak correlation to difference in pricing

Equator countries produce coffee beans -> Lower latte prices

Difficulties/Challenges:

- Finding price data in local currency
- Matching country names with currency codes (same currency countries)

