MM.DD.YY

Brewing Insights:

How Economic Factors Shape Coffee Shop Ratings

Group 7: Dierre Johnson, Cristian Gaggi, Noah Pelberg. Guanchen Wu, Jinning Zhang, Aoqi Zeng

Agenda

- Project Background
- Data Analysis
- Key Findings

Project Background

Business Question

How do macroeconomic factors like US inflation and stock market performance influence average star ratings for coffee shops?

Hypotheses

1)

As stock market performance goes up, coffee shop ratings will increase

2)

As US inflation goes up, coffee shop ratings will decrease

Data Sources

Yelp Dataset

Tables: business, user, review_notext, category

Columns: stars, date, average_stars, category_name

Yahoo Finance

S&P 500 - (GCSP)

Federal Reserve Economic Data (FRED)

Consumer Price Index (CPI)

API

FRED API

```python
#get the yahoo data
import yfinance as yf
from datetime import datetime, timedelta

ticker_symbol = "^GSPC"

ticker = yf.Ticker(ticker_symbol)

start_date = "2005-01-01"
end_date = (datetime.now() - timedelta(days=datetime.now().weekday() + 2)).strftime("%Y-%m-%d")
historical_data = ticker.history(start=start_date, end=end_date)
sp500_data = historical_data[["Open", "Close"]]

print(sp500_data.head())
sp500_data.to_csv("sp500_open_close.csv", index=True)
print("data saved sp500_open_close.csv")
```

Yfinance library

```python
#get cpi
import requests
import pandas as pd
from datetime import datetime, timedelta


API_KEY = "e4d0cd3909f4ad1a5a9662303565ea43"


series_id = "CPIAUCSL"
start_date = "2005-01-01"
end_date = (datetime.now() - timedelta(days=datetime.now().weekday() + 2)).strftime("%Y-%m-%d")
url = f"https://api.stlouisfed.org/fred/series/observations"

params = {
    "series_id": series_id,
    "api_key": API_KEY,
    "file_type": "json",
    "observation_start": start_date,
    "observation_end": end_date
}
```

Joining Data Sets

5 columns from 3 tables in the yelp data

```
-- project code
select c.business_id, category_name, date, r.stars, b.city
from yelp.category as c
left join yelp.review_notext as r
on c.business_id = r.business_id
left join yelp.business as b
on r.business_id = b.business_id
where category_name like '%Coffee%';
```

Challenges

Reliable API Access

Creating Effective Visuals

Data Merging

Data Cleaning & Troubleshooting

Data Analysis

Checking for balance

- Remove any city with less than 10 reviews to ensure balanced data
- Check for balance on average stars

```
#checking for balance
balance_table = yelp_data.groupby('city').count()
balance_table['observations'] = yelp_data['city'].value_counts()
balance_table
```

```
#remove anything below 10 observations
yelp_data = yelp_data.groupby('city').filter(lambda x: len(x) > 10)
yelp_data['city'].unique()
```

city	stars
Boise	4.5
Clearwater	4.764706
Dunedin	4.388889
Edmonton	4.0
Indianapolis	4.514286
Maplewood	3.619048
Moorestown	4.533333
Nashville	4.0
New Orleans	4.402439
Newtown	4.264706
Philadelphia	4.12963
Reno	4.488889
Saint Louis	3.4375
Santa Barbara	3.75
Tampa	3.865385

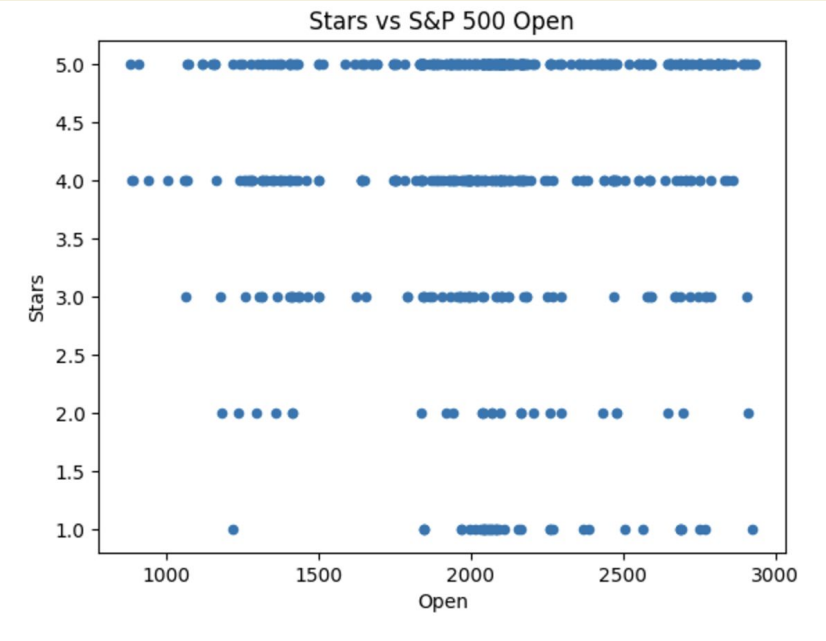
Yelp_Sp500 Table

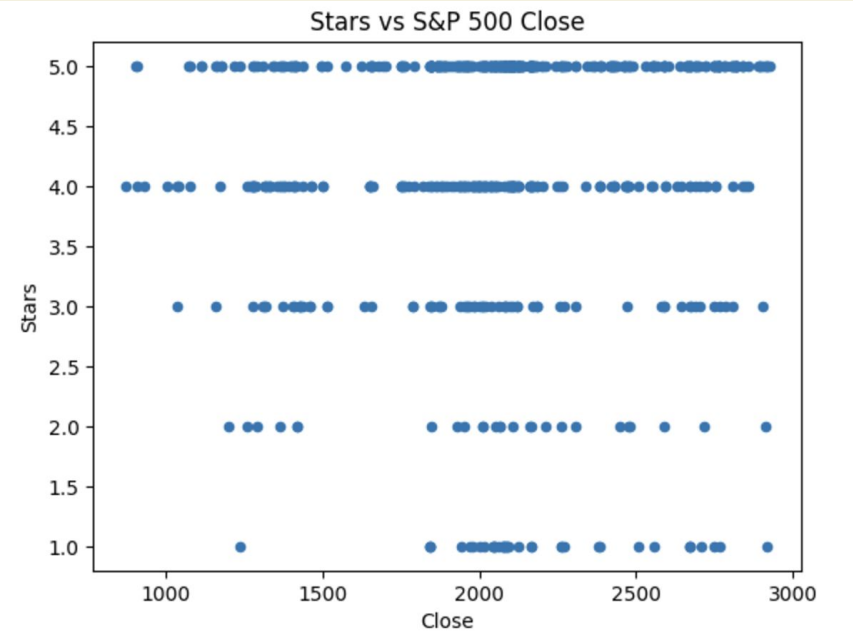
	business_id	category_name	date	stars	city	Open	Close
59	0zH0l4Jbf-oove3cLvrFOg	Coffee & Tea	2012-01-03	3	Newtown	1258.859985	1277.060059
64	ZyyExneJPXjci7yc2CGAPQ	Coffee & Tea	2012-01-03	4	Philadelphia	1258.859985	1277.060059
488	oaboaRBUgGjbo2kfUlKDLQ	Coffee & Tea	2012-01-04	5	Nashville	1277.030029	1277.300049
288	9sdu-m075T6qdB46wXP46w	Coffee & Tea	2012-01-04	4	Nashville	1277.030029	1277.300049
254	oaboaRBUgGjbo2kfUlKDLQ	Coffee & Tea	2012-03-01	5	Nashville	1365.900024	1374.089966
...	...	...	...	...	...	...	...
230	GOSVLhjUT6TnrPhkkmDOpw	Coffee & Tea	2018-09-25	5	Philadelphia	2921.750000	2915.560059
443	DZdE09qqCcX7Atf3Df7Yqg	Coffee & Tea	2018-09-25	1	Nashville	2921.750000	2915.560059
833	9_m_iA3VAqGBJEW5XLc-5w	Coffee & Tea	2018-09-27	5	Indianapolis	2911.649902	2914.000000
720	EQ-TZ2eeD_E0BHuvoaeG5Q	Coffee & Tea	2018-09-27	2	Indianapolis	2911.649902	2914.000000
719	EQ-TZ2eeD_E0BHuvoaeG5Q	Coffee & Tea	2018-10-03	5	Indianapolis	2931.689941	2925.510010

Yelp_CPI Table

	business_id	category_name	date	stars	city	CPI
6	lDtLPgUrqorrpqSLdfMhZQ	Coffee & Tea	2018-01-09	5	Santa Barbara	252.182
59	0zH0l4Jbf-oove3cLvrFOg	Coffee & Tea	2012-01-03	3	Newtown	228.807
64	ZyyExneJPXjci7yc2CGAPQ	Coffee & Tea	2012-01-03	4	Philadelphia	228.807
93	CPFKi2lZJazP6ldtCdDDyg	Coffee & Tea	2016-01-04	5	Reno	238.992
116	EQ-TZ2eeD_E0BHuvoaeG5Q	Coffee & Tea	2018-01-10	5	Indianapolis	252.772

Average star rating vs S&P 500 Open and Close Price





Correlation Matrix: Average Stars vs S&P 500 Open/Close

	stars	Open	Close
stars	1.000000	0.040307	0.040331
Open	0.040307	1.000000	0.999480
Close	0.040331	0.999480	1.000000

Average star rating vs CPI

Correlation Matrix: Average Stars vs CPI

	stars	CPI
stars	1.000000	0.040501
CPI	0.040501	1.000000

Key Findings

Summary

1. **No Correlation Observed:** Neither US stock market performance (S&P 500) nor inflation (CPI) showed any significant impact on the average star ratings of coffee shops.

2. **Consistent Ratings:** Average star ratings remained unchanged regardless of fluctuations in daily S&P opening/closing prices or varying CPI levels.

Key Insights

1. **Customer Loyalty**: Regular customers may continue to support their favorite coffee shops regardless of economic fluctuations.

2. **Miniscule Price Changes**: Small price adjustments may not significantly affect customer perceptions or ratings.

3. **Service-Based Industry**: Ratings are primarily influenced by service quality, not external economic factors.

Possible New Hypotheses

1. **Operating Hours**: People care a lot about when they are able to get coffee (such as before work). If a coffee shop is not open during convenient hours or days, it may receive lower ratings.

2. **Taste of Product**: Where they source their coffee beans from, the types of drinks and food they offer, and the freshness of their products are all factors that could impact ratings.

3. **Ambiance of Shop**: With the growing number of work-from-home jobs, coffee shops have become popular places to work. If the ambiance is not conducive to work, it could lead to lower reviews