UConn top 25 Public Research Institution

92% secured post – graduate employment or service

2,249 undergrads, 329 grads, 2,578 total

70% are of diverse backgrounds, 52% are first gen, 47% are Pell grant eligible

UCONN Stamford Computer Science program started in 2016

Focus on Internships and Real-World Challenges

Synchrony Financial build a co-location lab on Campus hired 20+ students

Financial companies -Morgan Stanley, UBS, RBS, Mckinsey, AQR and Point 72.

State Street Project Team

Dr. Phil Bradford
20+ Years Financial Services
Blackrock, GE. State Street
22+ Teaching
Univ Alabama, Rutgers, UCONN

John Iacovacci, MBA
20+ Years Financial Services
Citigroup, GE, State Street
20+ Teaching
Fordham , UCONN

Haris Mahmood
Google DSC Lead
UCONN Senior
The Hartford

Shalini Kuzhippat
UCONN Senior
Computer Science

Charlton Toussaint
UCONN Senior
Computer Science

Charlie Kuryluk
UConn Student
Cyber Security

Dedeep Singu
UCONN CS Major
Math Minor
intern Synchrony

Rauni Machado-Filho
UCONN Senior

David Keenan
UCONN Junior

Google   UCONN Stamford Google Club

Google Developer Student Clubs for students interested in Google technologies.

Stamford Google Club - only club in CT - 170 students

Speakers
GE, Bloomberg, Lockheed Martin, Google, Twitter, Deutsche Bank, State Street Bank.
Cloud development platform creating real-world applications for financial sector.

Google Developer Student Clubs
University of Connecticut

Susan Lasota
Chief Technology Officer
STATE STREET GLOBAL ADVISORS

Thu. Oct. 13 | 5pm EST | Auditorium 1

- Learn how to advance your career in tech
- State Street manages over 3.4 trillion dollars in assets
- Potential job/internship opportunities at State Street!
- Free Pizza and Drinks

RSVP Today!

Public Equities – Sentiment Project

- Build a library of key words and topics for sentiment model
- Build and develop sentiment model to analyze company conference call transcripts to identify shift in company fundamentals
- Develop sentiment score using sentiment model analysis
- Train sentiment model to avoid « hallucination"
- Incorporate sentiment score into Fundamental Equity team CQ Calibration dashboard (in Spotfire)

Ability to monitor conference calls for covered companies and identify sentiment of categories:

- Financial metric   - Macro
- Sector trend   - Regulation

Keywords / Topics

Inflation
Interest rate
Raw Materials inflation
Volume
Revenue
Earnings
Earnings per Share / EPS
Share Repurchase
Buy back
Dividends
Economy
Operating Expenses
Operating margins
Pricing
Gross margin
Market share
Employee attrition
Inventory
Fuel cost
Wage increase
Price increases
Guidance
Sales drivers
Margin improvement

Revenue Growth
Competitive Environment
EBIT Margins
Organic growth
Book-to-bill
Contribution margin
Orders
Restructuring
Inventory destocking
Backlogs
Order rates
Incremental margins
Recurring revenue
Aftermarket growth
Labor shortage
Capacity utilization
Economic stimulus
Renewables
Megaprojects
Reshoring
Regulation
Capex
Lay offs
Unemployment

Net Interest Income
Mortgage Fees
Reserves
Investment Banking Fees
Trading revenues
Credit cost
Delinquency
Capital Requirements
Loan-to-Value
Interest coverage
Defaults
Commercial Real Estate
Collateral value
Net Interest Margin (NIM)
Loan growth
Cost/income ratio
Provisions
Non-performing loans
Credit card balance
Delinquencies
Mortgage Refinancing
Debt issuance
IPO
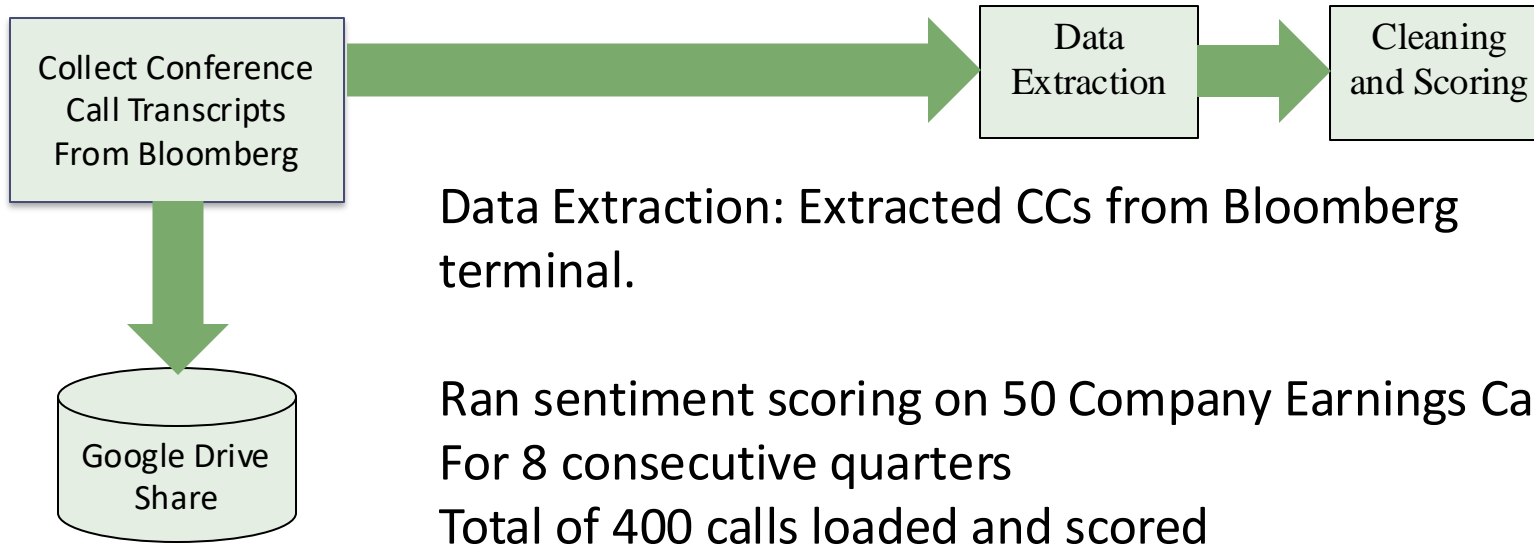Deposit

Sentiment Scoring

Conference calls are loading into the process and searched by keyword depending on sector. Paragraphs containing keywords are scored via sentiment model and weighted based on the importance of the word.

Sector

Industrials
Banks
All

Categories

Financial metric
Macro
Sector trend
Regulation

Weights (adjustable)

Very important = 1.5
Important = 1
Less so important = 0.5

Keywords/Topics	Sector	Key Word Category	Weights
Revenue	All	Financial metric - All	Very Important
Earnings per Share / EPS	All	Financial metric - All	Very Important
Share Repurchase	All	Financial metric - All	Important
Buy back	All	Financial metric - All	Very Important
Dividends	All	Financial metric - All	Important
Net Interest Income	Banks	Financial metric - Bank	Very Important
Economy	All	Macro	Less so important
Mortgage Fees	Banks	Financial metric - Bank	Important
Reserves	Banks	Financial metric - Bank	Very Important
Operating Expenses	All	Financial metric - All	Important
Credit cost	Banks	Financial metric - Bank	Important
Delinquency	Banks	Sector trend	Important

Sentiment Loading Process

Conference call transcripts for a two-year period for each company

Collect Conference Call Transcripts From Bloomberg → Data Extraction → Cleaning and Scoring

Google Drive Share

Data Extraction: Extracted CCs from Bloomberg terminal.

Ran sentiment scoring on 50 Company Earnings Calls
For 8 consecutive quarters
Total of 400 calls loaded and scored

Documents stored by ticker, period, date of call.

e.g. WSP.TO_Q123_5-11-23

STATE STREET
GLOBAL ADVISORS

Conference Call Sentiment score validation

Team performed manual score review which validated our automatic testing.

Individually scored Conference calls using generic Google sentiment scoring process results were in line with the BERT financial trained model we used.

Paragraph	Cloud NLP API Score	Model Score	Model Score weighted
Let me now address some addit	0.558	0.613	0.829
ber two in leverage finance with	0.837	0.945	0.999
And so those investments acros	0.895	0.786	0.915
I'd also like to caution you that w	-0.023	-0.044	0.483

Positive   Negative   Neutral

Sentiment Scoring Process

Remove Stop Words	→	Tokenize the Text	→	Score Paragraph Sentences	→	Store Details

Ex: the, is, with, and, on, of

Parse through keywords and extract paragraphs

Run paragraph thru sentiment model to get score (-1 to 1)
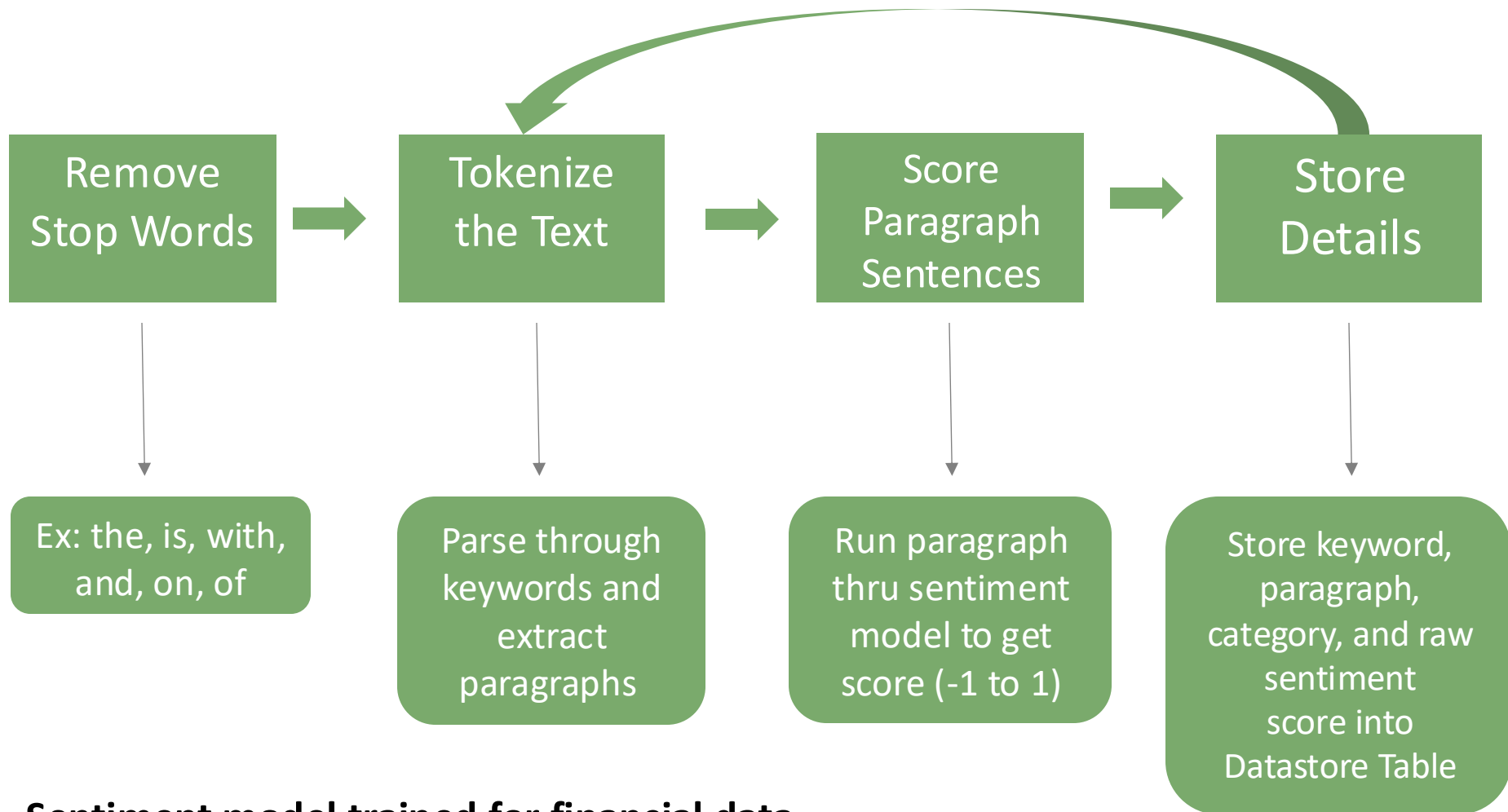
Store keyword, paragraph, category, and raw sentiment score into Datastore Table

Sentiment model trained for financial data

Sample of Stored Details

Call Date: 10/14/2021

Ticker: C US

Category: Financial metric

Period: Q42021

Document Type: CC

Score: 0.265

Keyword: Dividends

Sector: All

Paragraph: Let me briefly touch on our US retail strategy. Now, our digital strategy and investments are starting to pay **dividends**. Digital deposits stand at $19 billion, having grown 26% in the last year. And we see more than two-thirds of our digital deposit balances coming from customers outside of our branch footprint, with about half of those deposit from our card holders who did not previously have a retail relationship with us.

Weighted Sentiment: 0.397

Sample of Sentiment Summary:

Ticker: C US   **Period:** Q42021

Financial metric - Average: 0.014
Financial metric - Count: 32
Financial metric - Weighted Average: 0.035

Macro Average: 0.402
Macro Count: 3
Macro Weighted Average: 0.361

Regulation Average: 0
Regulation Count: 0
Regulation Weighted Average: 0

Sector trend Average: 0.383
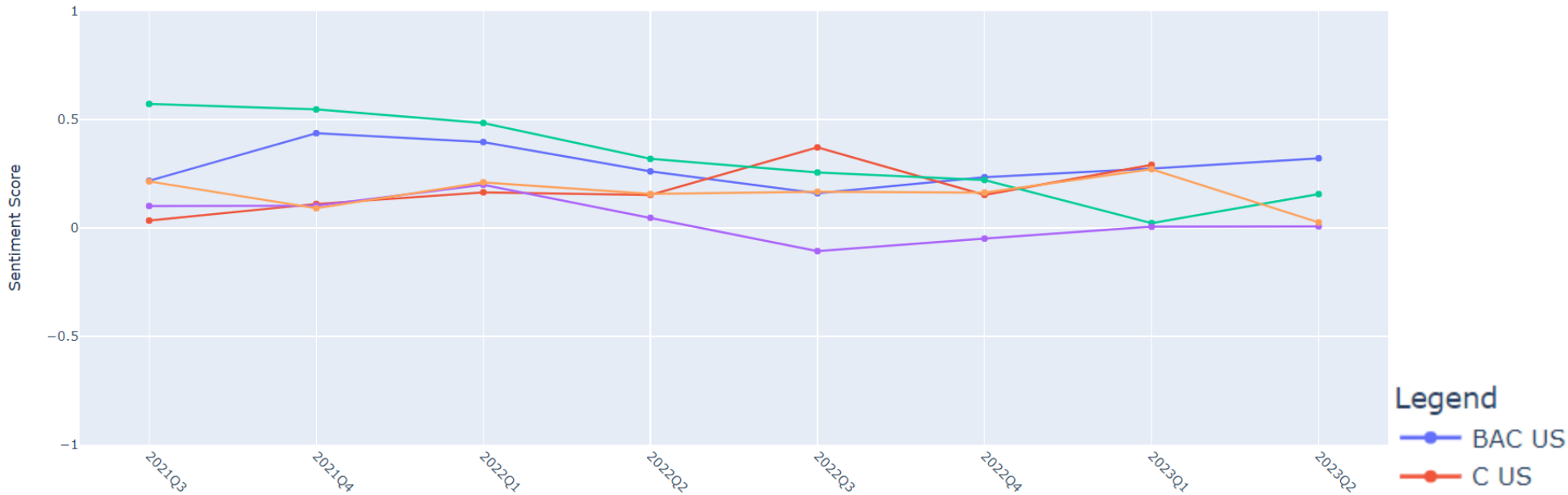**Sector trend Count:** 8
Sector trend Weighted Average: 0.321

Total Average: 0.110
Weighted Average: 0.111

Summarize Sentiment

Aggregate Sentiment Details into Summary Table

STATE STREET GLOBAL ADVISORS

Comparison Graphs (Banks)

Raw score

Weighted score
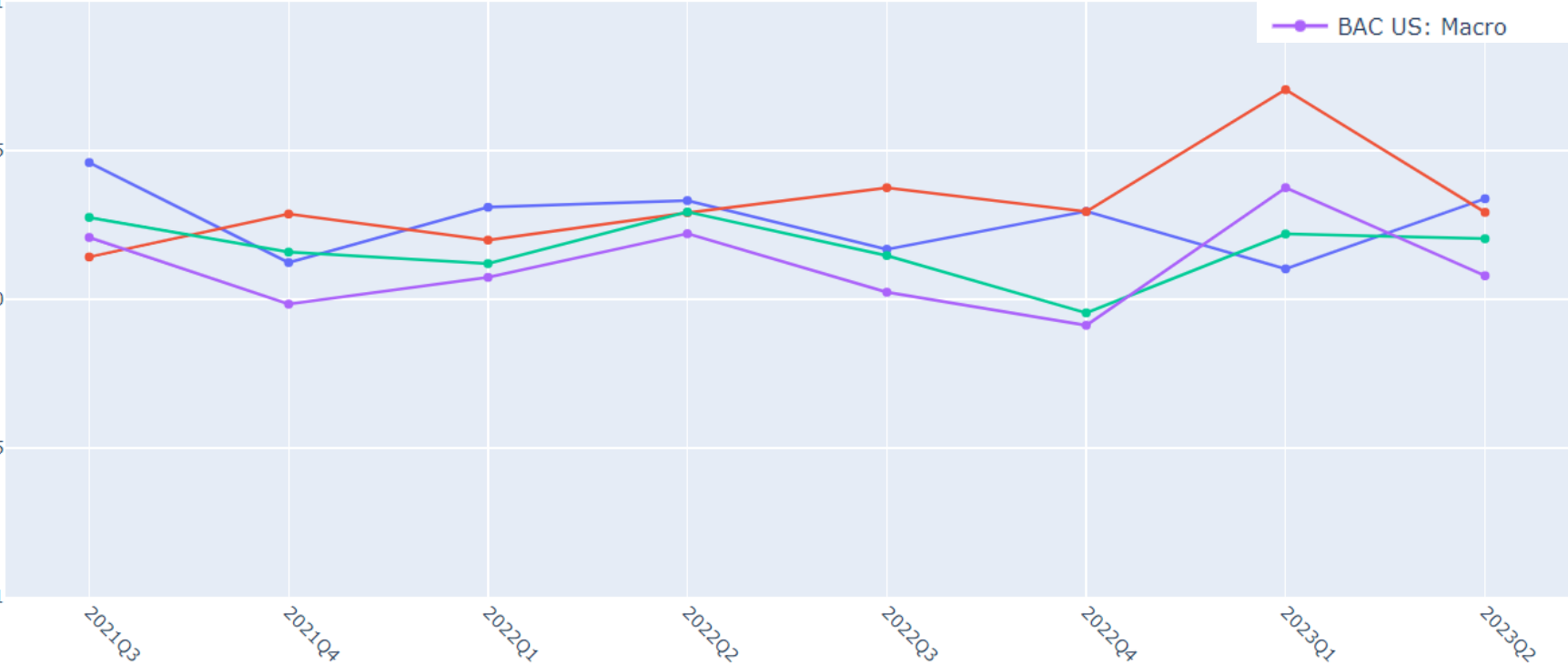
Legend
- BAC US
- C US
- GS US
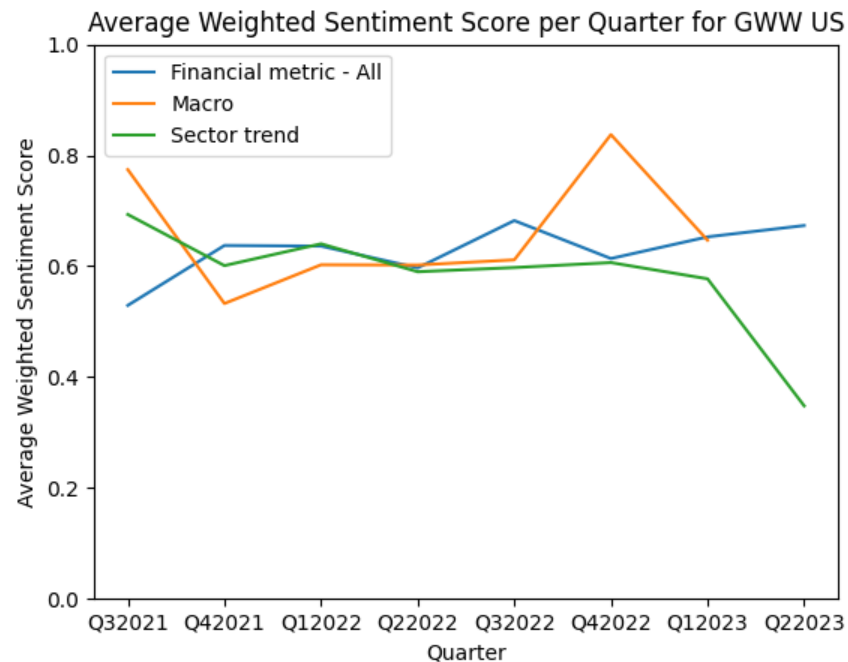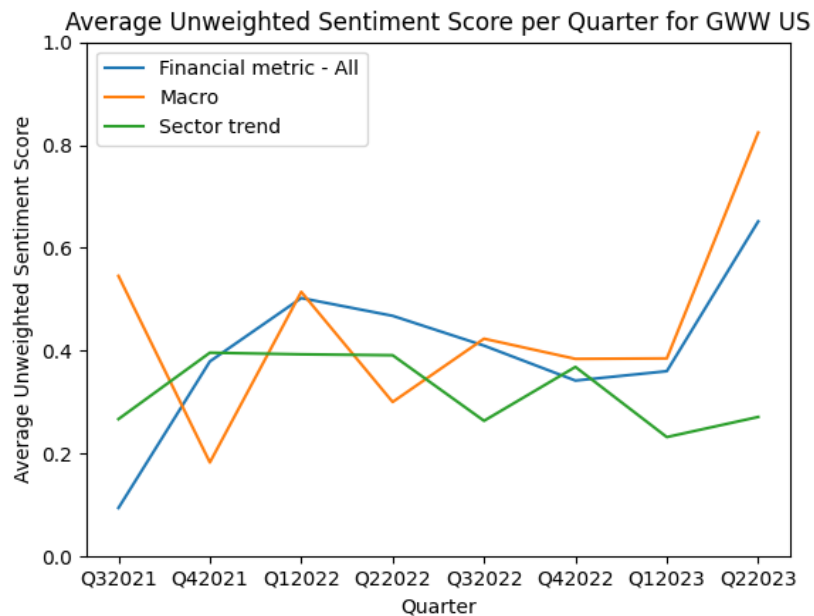- JPM US
- WFC US

13
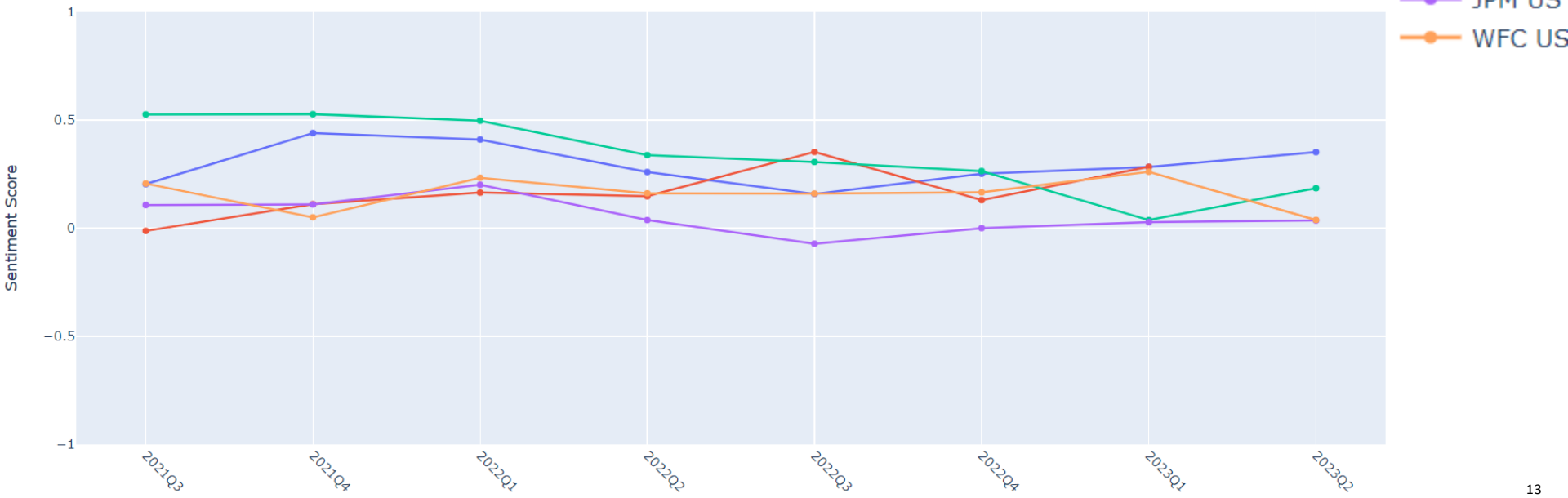
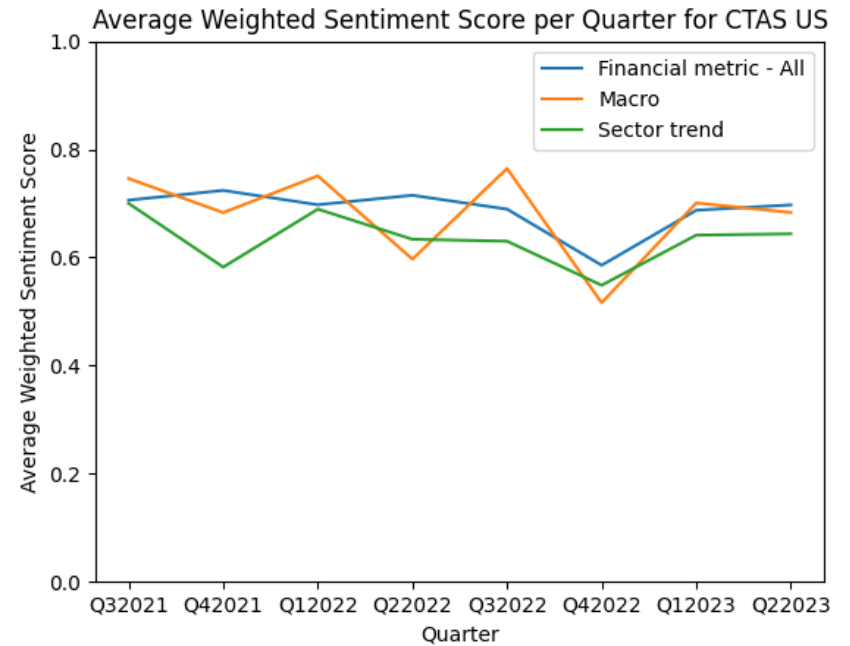# Keyword Occurrences – BAC US
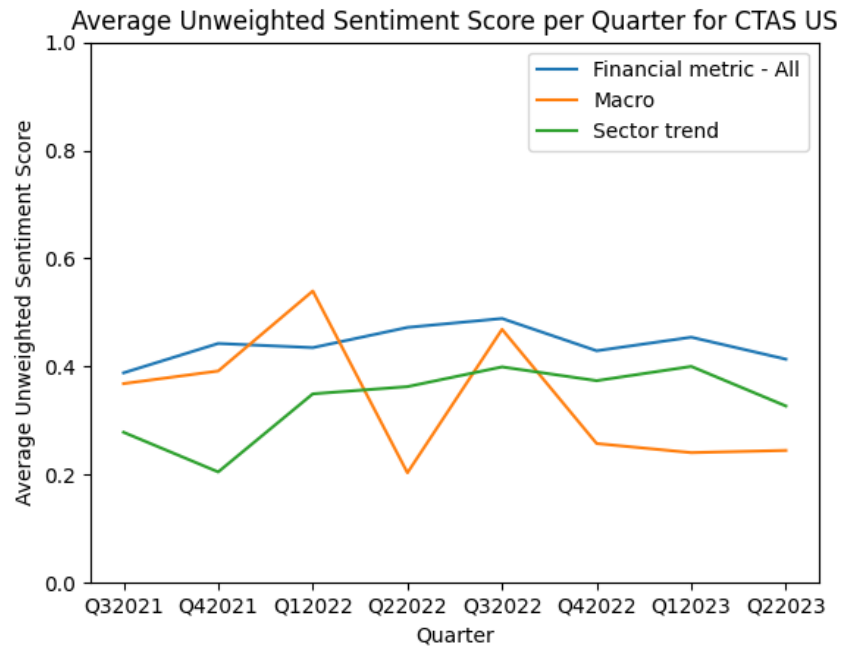

BAC US Sector Occurrences
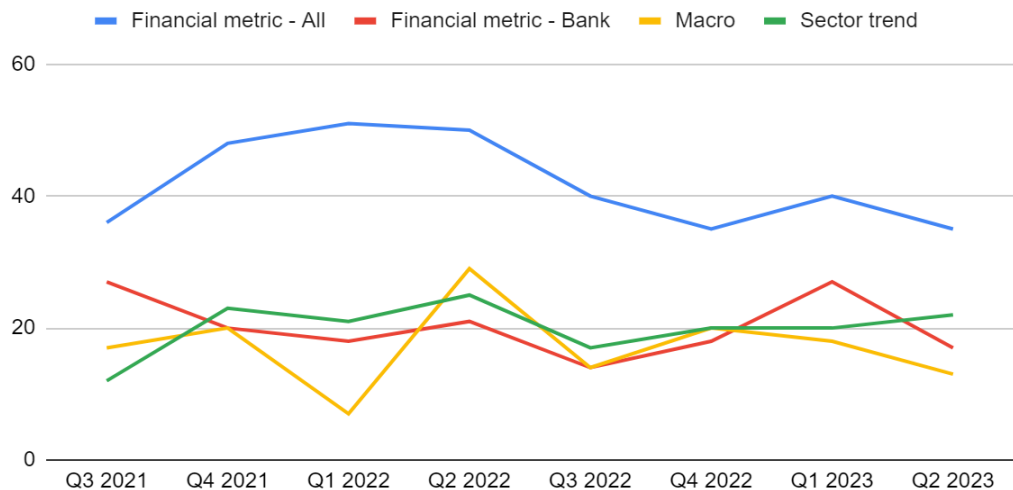
Raw score

Graphs- WW Grainger Inc. (GWW US)

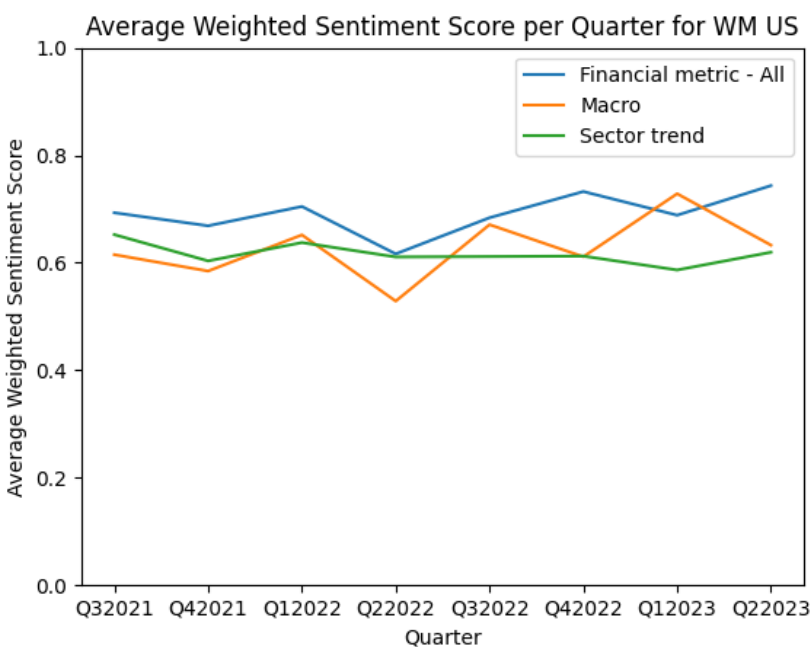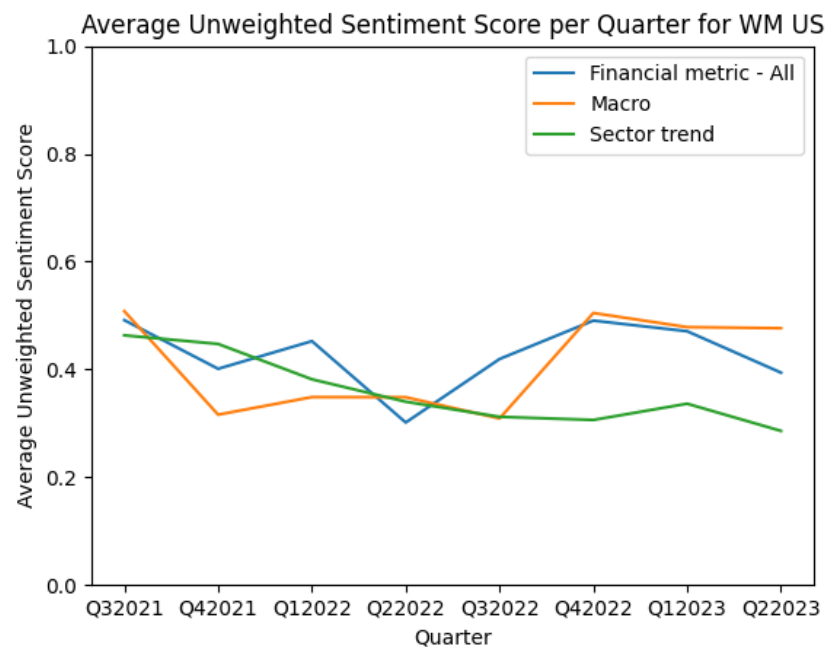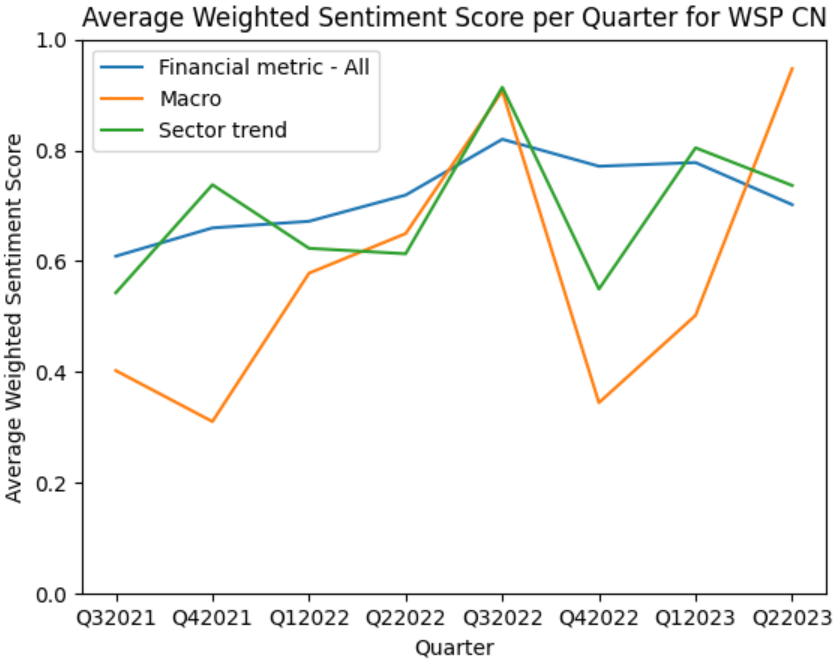# Graphs- Cintas Corp (CTAS US)

Average Unweighted Sentiment Score per Quarter for CTAS US

Average Weighted Sentiment Score per Quarter for CTAS US

Graphs- Waste Management (WM US)

STATE STREET
GLOBAL ADVISORS

Graphs- From WSP Global Inc. (WSP CN)

Average Unweighted Sentiment Score per Quarter for WSP CN

Average Weighted Sentiment Score per Quarter for WSP CN

Drill Down

Input Ticker, Period, and Category.
Retrieves paragraph, weighted and unweighted scores, and keyword
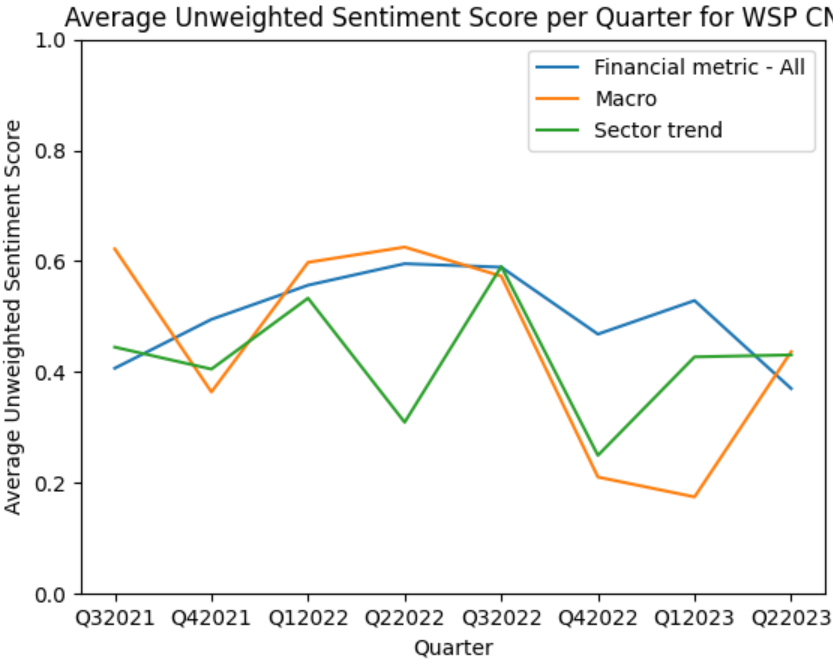
UConn Sentiment Data Manager

Info Drill Down

Yahoo Ticker: C US

Period: Q22022

Category: Financial metric - All

Submit

Paragraph	Weighted Sentiment	Score	Keyword	Category
So those are a couple of things. Again, we're consistent with guidance. We are -- we believe we're on track for the guidance we gave for the full-year of mid-single digits and we are looking forward to generating the efficiencies that come out of this spend and put a dent in our structural expense base over time.	0.828	0.923	Guidance	Financial metric - All

Model

o BERT (Bidirectional Encoder Representations from Transformers)

   o Developed by Google AI, Industry standard

   o Used by major firms like Bloomberg for text analysis.

   o Pretrained on large-scale unlabeled data: Book Corpus & English Wikipedia

   o Reads context in both directions for deeper language comprehension

   o Finetune for specific tasks --> Sentiment Analysis

Training

o Custom SSGA Training

    o Model can be trained on specific term

    o Industry Specific customizable weights

o Example: the word "reserves" can be trained as negative in the banking sector, despite being neutral in general sentiment analysis

sentence string · lengths	label class label
According to Gran , the company has no plans to move all production to Russia , although that is where the company is growing .	1 neutral
Technopolis plans to develop in stages an area of no less than 100,000 square meters in order to host companies working in computer technologies and telecommunications , the statement sai…	1 neutral
The international electronic industry company Elcoteq has laid off tens of employees from its Tallinn facility ; contrary to earlier layoffs the company contracted the ranks of its office…	0 negative
With the new production plant the company would increase its capacity to meet the expected increase in demand and would improve the use of raw materials and therefore increase the…	2 positive

Roadmap

o **Data Acquisition**

    o Collect audio & text from financial conference calls

o **Data Preparation**

    o Clean data

o **Model Training**

    o Train using cleaned & segmented data

o **Model Evaluation**

    o Access the model's performance using validation sets in additional to manual testing.

o **Adversarial Training**

    o Strengthen model robustness by training with Adversarial examples

o **Continuous Learning**

    o Regularly update the model with fresh data to adapt to evolving