Apparel Brand Sentiment Analysis Using Big Data

Course: CS 226 Big Data Management Group Members: Cristina Lawson, Sridevi Raju, Varshini Sampath, Hamsa Gouda Veerendra





Outline

Motivation

Problem Statement

Relevance to Big Data

Related Work

Steps Involved

Main Proposed Work

Project Progress

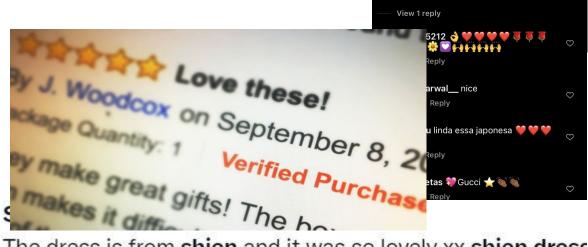
Evaluation

Conclusion

Motivation

Shopping???





6 likes Reply

5d Reply

debbie_wingham . Love it

The dress is from shien and it was so lovely xx shien dress great for plus size too

my chanel ss2020 fashion show review (1/3)





**** Very disappointed

Reviewed in the United States on October 20, 2020

Size: XX-Large Color: Black Verified Purchase

I was so excited for this product because I've been wanting a good fleece but I just didn't want to pay all that money. I bought a size 2xl. It came a day ea supposed to. As soon as I opened it and took it out, I knew it wasn't going to fit. Normally, I fit into Amazon's clothing pretty well and buy in a size 2xl. The especially around the chest area. I'm a 44D. And it fell just a few centimeters below my belly button. I like my clothes fall pretty far below my abdominal super curvy girl with love handles and all and this product did not do a good job at covering and it was not flattering. The material felt like those sheets would use in art class in first grade. Therefore, not great quality. Moral of my story, if you're curvy and busty, this product will not be your friend. However,

Report abuse

GitmoGecko

***** Very disappointed.

Reviewed in the United States on November 3, 2020

Order way up if you are ordering them. I ordered the white buffalo plaid and the red buffalo plaid ones. I opened them up and they are really small. I ord expecting them to run small, but I didn't expect them to be tiny! They will be going to my granddaughters that are in sizes 10/12 and the older one wear seriously do not think they are for women, but more for little girls. I was very disappointed.



Report abuse



******** Would not buy again

Reviewed in the United States on February 6, 2021

Verified Purchase

I purchased the fushia color in size small. I have a small frame but still the arms are snug, which forces me to wear a short-sleeved shirt rather than the lo preferred. Otherwise the fit is good. On the plus side this top is soft and warm. I wish the collar was not quite so high. The thing I really don't like is the fi would guess it will last maybe 10 wearings before it gets tossed.

Helpful

Report abuse

Problem Statement

• The aim of the project is to **analyze the sentiments** of the people on various clothing brands.

Adidas, Asos, Boohoo, Chanel, Gucci, H&M, Nike, Shein, Victoria's secret, Zara

- The reviews of people in twitter form the data for this project.
- We are going to evaluate how these brands are perceived across the world by capturing negative, positive, and neutral sentiments and and compare them.
- We also visualize them on graphs.

Big Data Relevance

VOLUME

Data from social media - Around 175k tweets from Twitter

VARIETY

Actions of customers, suppliers etc., from various sources

Text, images etc..

VERACITY

Quality of the data

Social media nowadays has become the basis when it comes to shopping.

VELOCITY

High data generation speed - online reviews

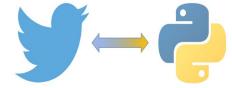
Project Steps

- → Data Collection
- → Data Preprocessing
- → Data Processing
- → Sentiment Analysis
- → Visualization

Main Proposed Work

Data Collection

- Tweets retrieval using brand names
- Collected tweet text, tweet date, location
- Tweets collected for this project = 175000



text	geo	created_at
if u kept up with my shein haul on my snap u a real one	{'place_id': '00a4bbcc0dcd7572'}	2021-12-06 2
Day 2 of Lana shein outfit discourse. No signs of stopping	{'place_id': 'e4a0d228eb6be76b'}	2021-12-06 2
idk why shein sweats are so high waisted	{'place_id': 'de599025180e2ee7'}	2021-12-06 2
@JuniorAyress @Kellychrystini rainha da Shein	{'place_id': '97bcdfca1a2dca59'}	2021-12-06 2
B @breeeeezuh LMFAO, getting they're evil eye from shein and think they protected from	{'place_id': '70707114a9202b98'}	2021-12-06 2
Casual and comfortable look by @sheinofficial use my	{'coordinates': {'type': 'Point', 'coordi	n 2021-12-06 2
7 Especially the oxtail from SHEIN https://t.co/hLeTiyYMbC	{'place_id': 'dd3b100831dd1763'}	2021-12-06 2
7 Shein is going to take all my money I can't stay off that app	{'place_id': 'c1e1a49e92ce1ba0'}	2021-12-06 2
B Let's start with my shein cart pls. https://t.co/WRJMg97Jr1	{'place_id': 'e564d30dc173d2a8'}	2021-12-06 2
4 S/o to Mel aka the bestest friend in the whole mf world she just bought my whole cart from S	{'place_id': '04cb31bae3b3af93'}	2021-12-06 1
Got a full SHEIN cart, who paying?!	{'place_id': '76a51908fb719a47'}	2021-12-06 1
0 @migurtcita Shein	{'place_id': '019f2f4e187ace41'}	2021-12-06 1
B the amount of shein orders i've done is ridiculous	{'place_id': '1c69a67ad480e1b1'}	2021-12-06 1
shein gotta chill w these notifications	{'place_id': 'c807df9bb77ea0b3'}	2021-12-06 1
7 how did you guys work when you missed your SHEIN delivery? asking for a dumb friend of mir	{'place_id': 'd98aef293ae7311d'}	2021-12-06 1
Uninstall shein	{'place_id': '002c292a55e071e0'}	2021-12-06 1
B Honestly after my last package from Shein I think I'm done asf with them!!	{'place_id': '609b31e0a50b4474'}	2021-12-06 1
@carrymeOut The dress is literally from Shein	{'place_id': '0138153149b79c7f'}	2021-12-06 1

Electric and the contract of t

Data Preprocessing

#vzv #veezy700 #overkillshop https://t.co/gBOiVy7Hhz",,2021-12-08 11:05:43+00:00,0,0,0,0

Emojis removal to retain only the text to facilitate the processing of data.

1461394273820160009, "!!CAN YOU BELIEVE IT! Now selling at \$131.00Adidas - Nmd_r1 by Adidas Shop here https://t.co/d1pIeOWo33 ☆☆☆☆ Rating: 4.9 https://t.co/oEmgalWlazS",,2021-12-08 11:05:21+00:00,0,0,0,0 1890473725,@mmxxiiimm Pharell x Adidas human race,,2021-12-08 11:04:28+00:00,0,0,0,0 1069306707644432385, "Precision to blur 4 Dope winter c'way !! adidas X Ghosted.1 SG 🁌 Multiple sizes available, tap to shop # #WearItOrChaseIt 🔒 @ UEFA Champions League https://t.co/2Kityg7Ghc",,2021-12-08 11:03:05+00:00,0,0,0 1392519908446117888, "@AltcoinDailyio https://t.co/HImtqxqwZP If you're in NFT, then your team Adidas (a)",,2021-12-08 11:03:01+00:00,0,1,1,0 4851402556, "[NEW MID LEVEL ROLE] @trippin world are looking for an ACCOUNT MANAGER to join! Trippin has partnered with some of the world's most progressive brands incl. Nike, Versace, adidas, Converse & Dry Farfe APPLY via https://t.co/i6U8ioJKDu #SocialFIXT #HireBlack https://t.co/0V9CZSxsAc",,2021-12-08 11:01:45+00:00,1,0,0,0 1141678518411124742, "RCSA Stransbourg | Adidas | Concept Kit #rcstrasbourg #rcstrasbourgalsace #RCSA #LiguelUberEats #Strasbourg #Alsace #allezracing

#yzy #yeezy700 #overkillshop https://t.co/g80iVy7Hhz",,2021-12-08 11:05:43+00:00,0,0,0,0,0,0,aidas
1461394273820160009,"PCAN YOU BELIEVE TTP Now selling at \$131.00Adidas - Nmd_r1 by Adidas
Shop here https://t.co/oEmgaNWazS",,2021-12-08 11:05:21+00:00,0,0,0,0,0,aidas
1890473725,@mmxxiiimm Pharell x Adidas human race,,2021-12-08 11:04:28+00:00,0,0,0,0,adidas
1069306707644432385,"Precision to blur Dope winter c'way !!
adidas X Ghosted.1 SG
Multiple sizes available, tap to shop
#WearItOrChaseIt @ UEFA Champions League https://t.co/ZKityg7Ghc",,2021-12-08 11:03:05+00:00,0,0,0,0,adidas
1392519908446117888,"@AltcoinDailyio https://t.co/HImtqxqwZP

If you're in NFT, then your team Adidas ",,2021-12-08 11:03:01+00:00,0,1,1,0,adidas
4851402556,"[NEW MID LEVEL ROLE] @trippin_world are looking for an ACCOUNT MANAGER to join!

Trippin has partnered with some of the world's most progressive brands incl. Nike, Versace, adidas, Converse & Damp; Farfetch.

#SocialFIXT #HireBlack https://t.co/0V9CZSxsAc",,2021-12-08 11:01:45+00:00,1,0,0,0,adidas

#rcstrasbourg #rcstrasbourgalsace #RCSA #Ligue1UberEats #Strasbourg #Alsace #allezracing

1141678518411124742, "RCSA Stransbourg | Adidas | Concept Kit

APPLY via https://t.co/j6U8ioJKDu

Data Processing

Used Spark for data processing

Why???

Large amount of data - processing made easy

Made the integration with further processing easier

Sentiment Analysis

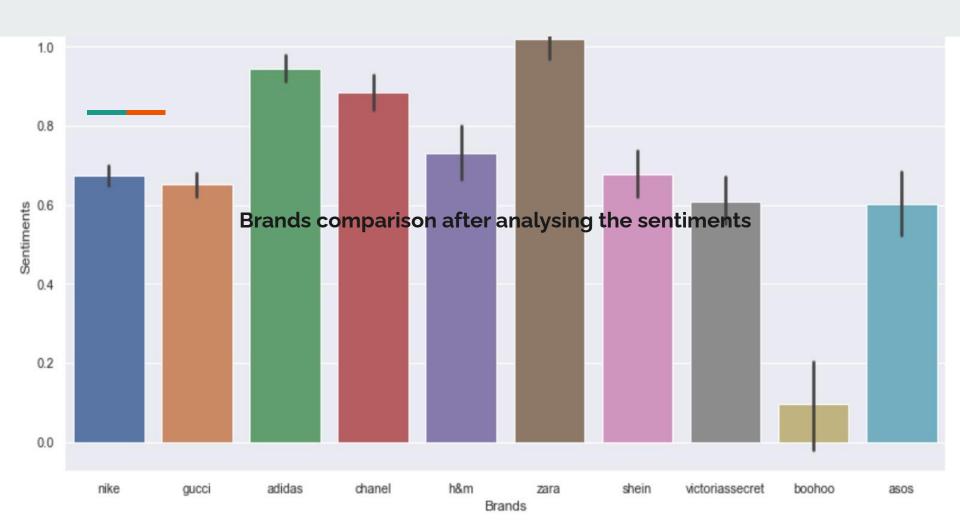
Technologies used

Afinn

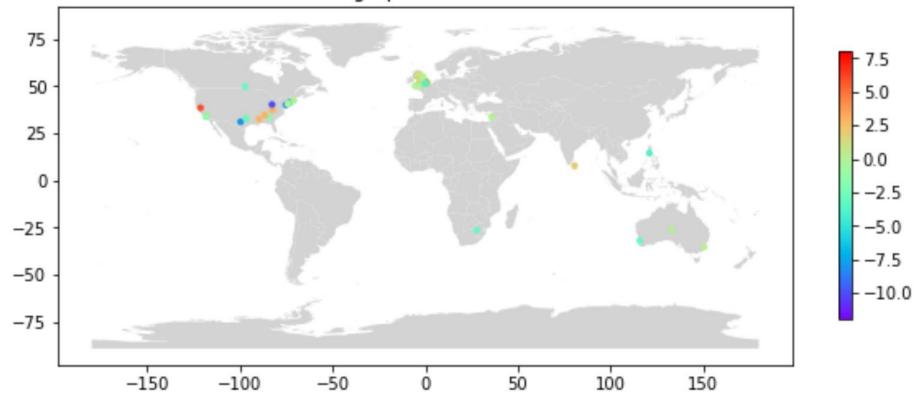
gives the sentiment of the tweet in the range of -50(negative sentiment) to +50(positive sentiment)

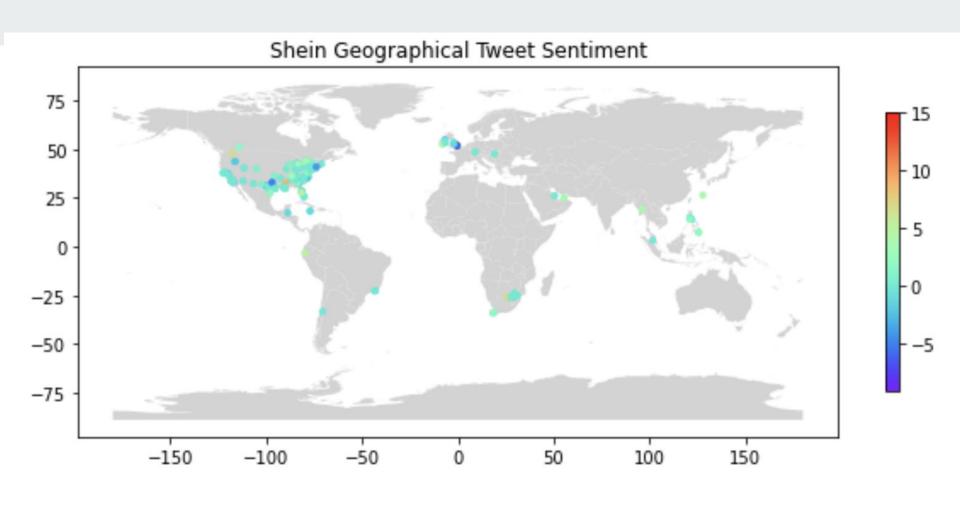
NLTK

NLTK gives positive, negative and neutral sentiment results



Boohoo Geographical Tweet Sentiment





Evaluation

- This project can accommodate the growing data.
- It is scalable to multiple bands across the world.
- Since spark is being used, the processing speed should not be affected to a large extent.

Conclusion

Started the project with comparison of two brands, Shein and H&M

To widen the scope and make it more meaningful added multiple brands

Now the project compares ten popular brands across the world

On the other side,

We can improve this project by adding ways to remove the unrelated tweets during preprocessing

We didn't have enough geospatial information to do the plotting