

Competitive Analysis of Value Fashion Brands in India with Text Analytics



**Tharuka
Gallekankanamge**

SRN: R19MBA82

Date: 20/08/2022

MBA in Business Analytics

Capstone Project Presentation
Year: I

- 01. Introduction**
- 02. Literature Review**
- 03. Problem Statement**
- 04. Objectives of Study**
- 05. Project Methodology**
- 06. Business Understanding**
- 07. Data Understanding**
- 08. Data Preparation**
- 09. Modeling**
- 10. Evaluation**
- 11. Analysis and Results**
- 12. Conclusions and Future Scope**
- Bibliography**
- Plagiarism Report**



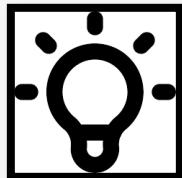
01. Introduction



Ever Changing | Ever Evolving | Ever Challenging



Competitive Landscape



What triggered this study?

Background

- Indian value fashion is an **increasingly fickle** environment
- **Competition** and fashion **heterogeneities**
- **Fast-paced** consumer evolution



Current status

Why this study?

- Provide **Business Insights** to an International value fashion Brand, “R&B” to **formulate a strategic marketing plan**”

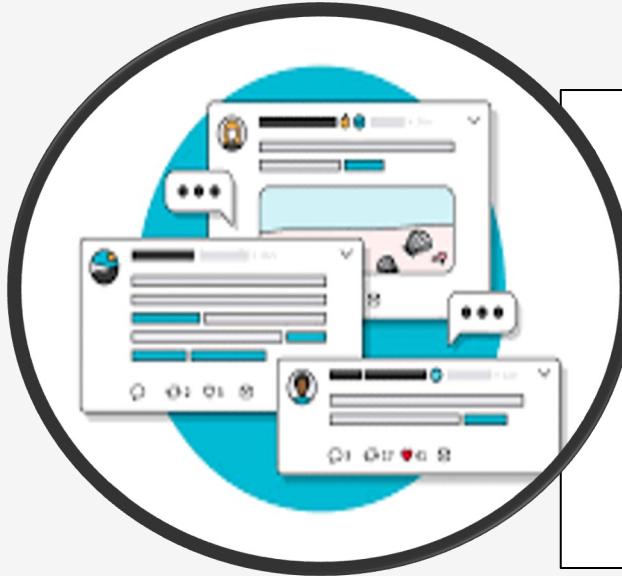


- It's a shark tank!
- Crowded with global and local brands.
- Content generated by these Brands on Social media is massive!

Competitive Landscape

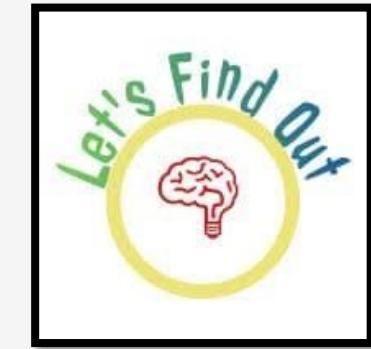


What triggered the study?



Is there a strategy and a thought process behind brand-generated content?

How can new brands leverage on this content to formulate a solid marketing strategy?



02. Literature Review

Let's gather some research around the study!

Business and Industry Related



- **Bigger organizations** have an upper hand due to **more workforce and investment**.
- New-gen companies leverage **emerging tech**.
- India will be the **world's most tech-savvy e-commerce market**
- **Exponential growth** due to the rapid growth of **internet users** in the country



Text mining & modeling techniques



- Derive **insights**, identify **patterns and opportunities**, and define **new algorithms** to tackle social media content.
- **Text mining** and **competitive intelligence**.
- Research on the Application of **topic modeling techniques (LDA & NMF)**



03. Problem Statement

“If you define the problem correctly, you almost have the solution”.

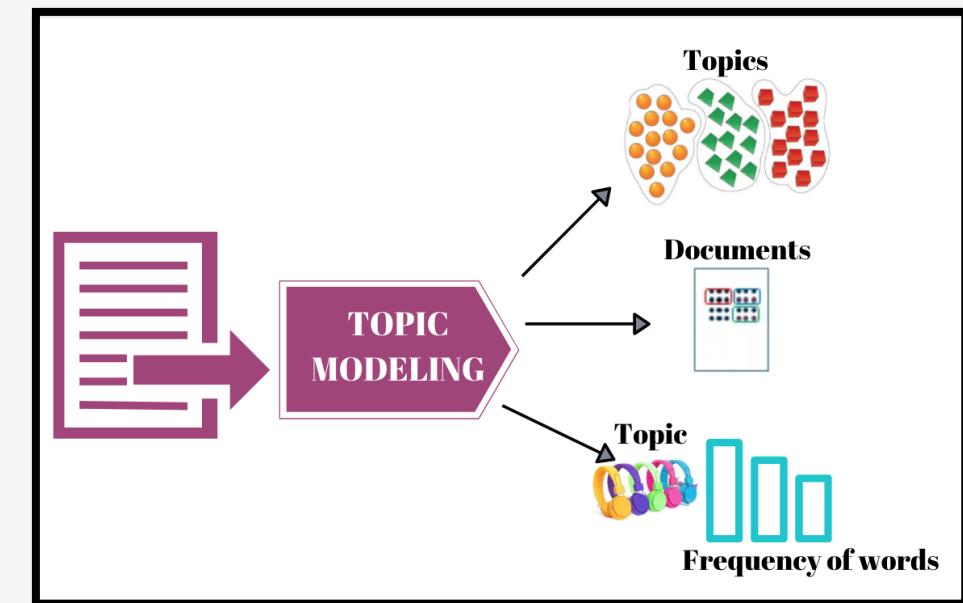
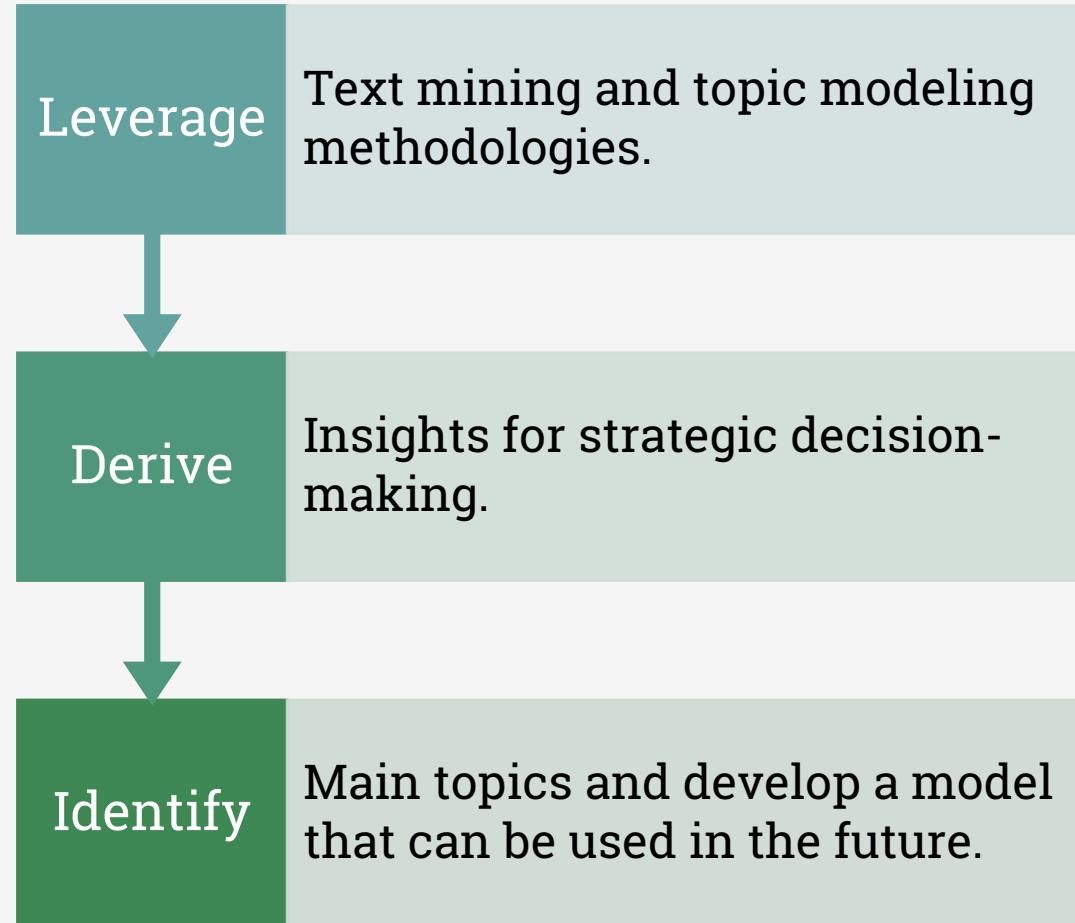
- Steve Jobs



04. Objectives of the study

“No project is completed until its objectives are achieved”.

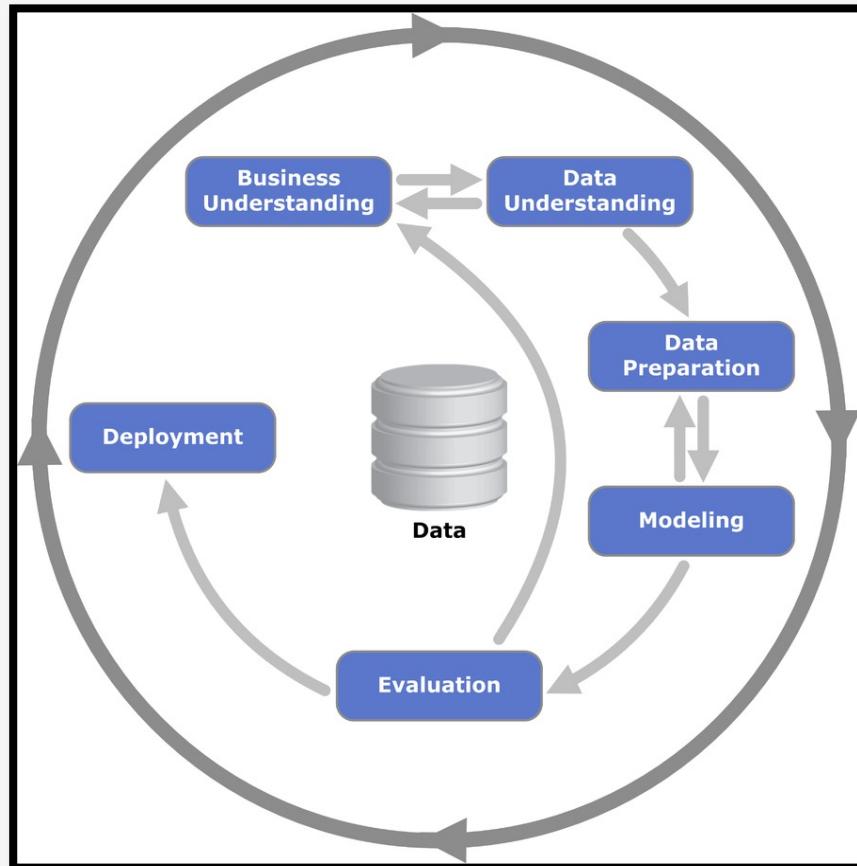
- Paulo Coelho



05. Project Methodology

A framework for data mining and analysis

CRISP - DM



- This study has been conducted as per Cross Industry Standard Process for Data Mining (CRISP-DM) methodology.
- This method has been adopted to derive insights from brand-generated content and used for topic modeling.
- The CRISP-DM methodology is a well-adapted process model across various domains, and it has six phases which is a step-by-step approach to any data science problem.

Figure: Cross-industry standard process for data mining Lifecycle source IBM.

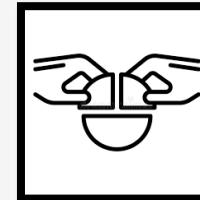
05. Business Understanding

“Understanding the needs of a business is the starting point of any project”.

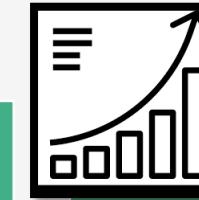
- John Williams



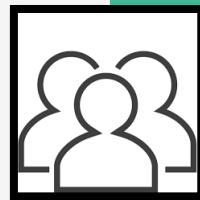
Revenue share of fashion segment: **US\$19.69 bn** in 2022



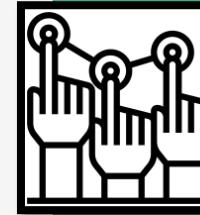
The projected revenue share: **US\$33.11 bn** by 2025



Expected CAGR (2022-2025): **18.92%**



Expected no. of users: **446.2m** users by 2025



User penetration in 2022: 22.8% & hit **30.9%** by 2025



ARPU (The average revenue per user): **US\$61.46**

Source: Fashion E-commerce Report (Statista Digital Market Outlook, 2021)

“If you don’t understand the details of your business you are going to fail”.

– **Jeff Bezos**

Value market segments under Fashion:



Women Apparel



Men Apparel



Kids Apparel



Footwear



Accessories and bags

Source: Fashion E-commerce Report (Statista Digital Market Outlook, 2021)

Dressing The Indian Masses & Unlocking the Value Fashion Retail

- **per capita expenditure** on fashion is expected to reach **\$ 95** by 2023!
- Value-Retailers Are Capturing Consumers Across Bharat!
 - H&M 50+ stores
 - Max Fashions 400+ stores
 - Trends: 500+ stores
- Online fashion industry continues to grow upwards!
 - Order volume growth of **51%** (FY21)
 - GMV (Gross Merchant Value) growth of **45%** (FY21).



Source: Official websites of above organizations, Fashion E-commerce Report, e-commerce focused supply-chain SaaS technology platform Unicommerce)

07. Data Understanding

“In God we trust, All others must bring data ”.

- W. Edward Demming



You only need a cup of data to make magic.

 H&M 
April 20, 2021 at 4:30 PM · 

Inspired by sunshine, created for spring. Our new kids' styles for the season are here! #HMKids

 Max Fashion
January 17, 2021 · 

Pair up this casual sweatshirt alongwith your favourite denim for a perfect sunday evening hangout. 😍

With upto 50% OFF on our #EndOfSeasonSale you can now avail the collection from our #MaxFashionStores near you. 🛍

#MaxFashion #EOSS #Sale #Discount
#WinterCollection #Fashion #Shopping #Deals
#MaxFashionStores #shopping #fashion #style
#ootd #outfitoftheday #lookoftheday
#fashiongram #style #currentlywearing
#lookbook #whatiworetoday #outfit #clothes
#todayimwearing #ou... See more

 Reliance Trends - Official
February 13, 2021 · 

Get some hearts fluttering in this pretty off-shoulder top with delicate embroidery. Shop in-store or online at trends.ajio.com for the perfect date-night look today!

#RelianceTrends #GetThemTalking
#TrendMyDate

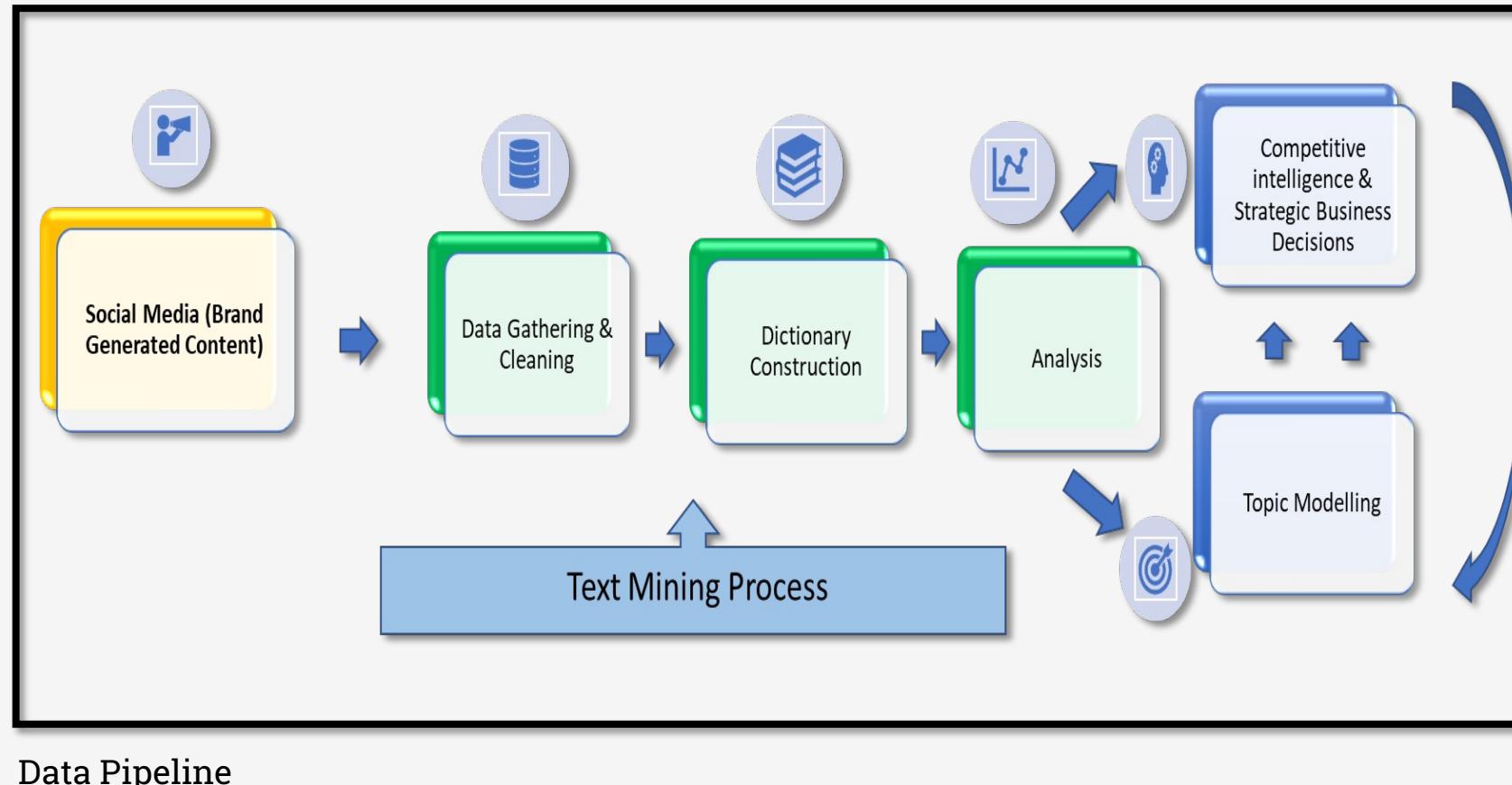
**Six Attributes/columns
(441 brand-generated content)**

- Brand
- Date
- Post
- No. of comments
- No. of Likes
- Topic (manually labelled)

Examples of brand generated posts

“process is more important than the results”.

- M. S. Dhoni



Data Pipeline

“Torture the data, and it will confess to anything”.

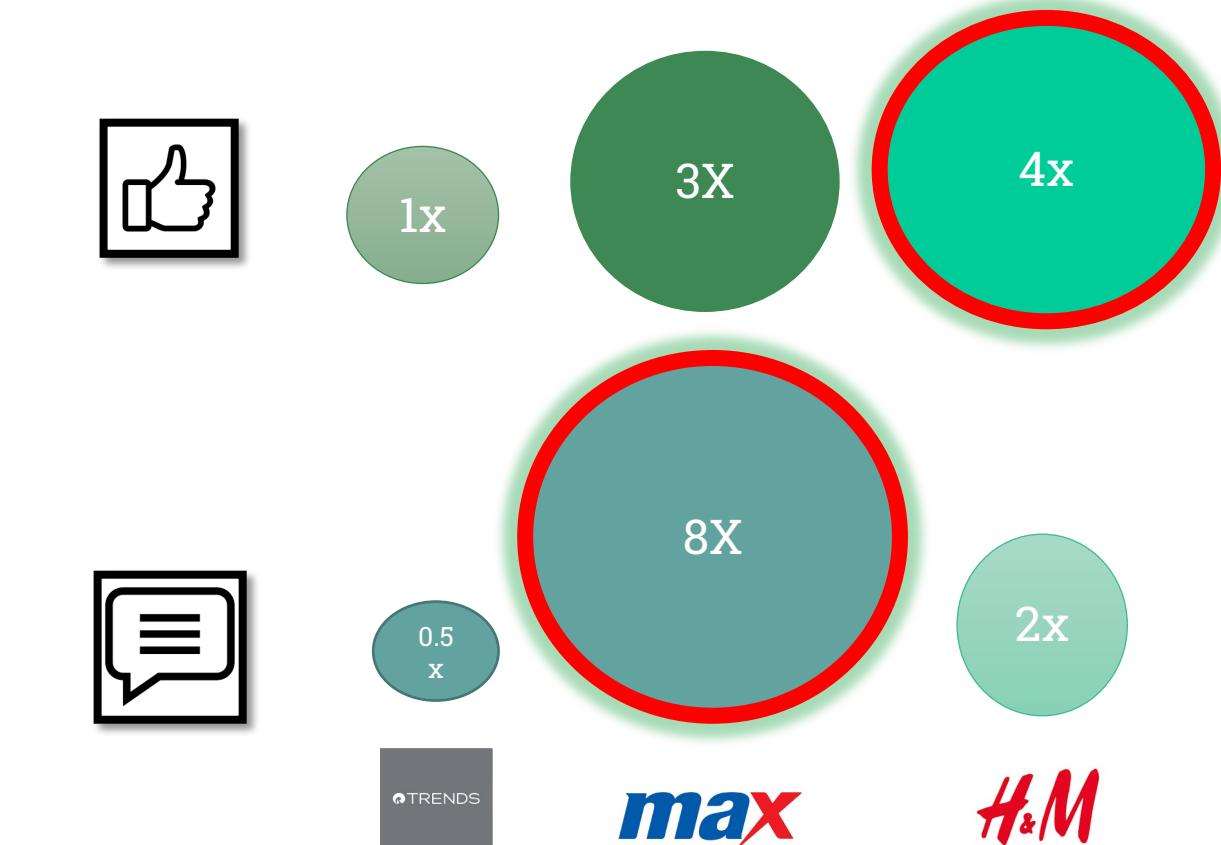
- Ronald Coase

No. of Posts generated by each brand:



- More the Merrier!
- Max Fashion India has cracked the code in engaging with their audience followed by H&M.

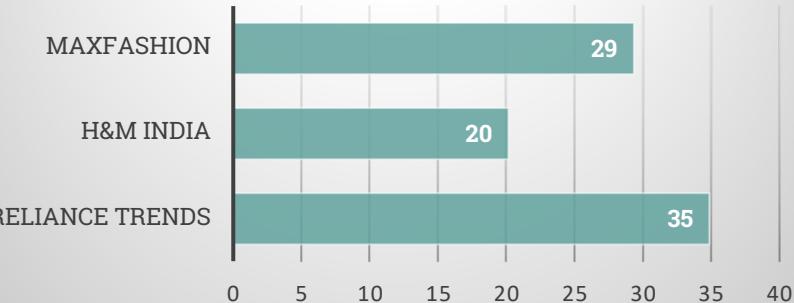
No. of Likes & Comments attracted by each brand:



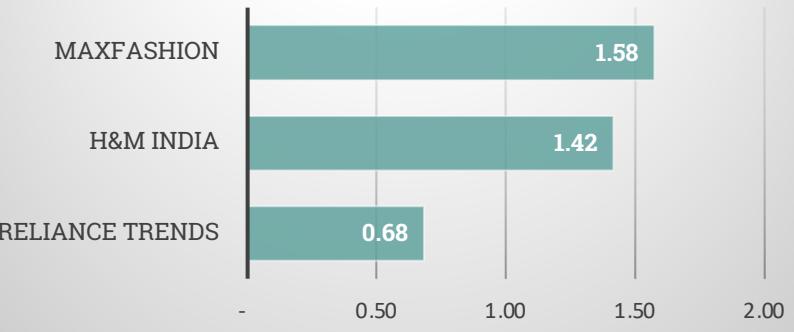
“Torture the data, and it will confess to anything”.

- Ronald Coase

Avg. No. of words per post



Avg. Posts per day

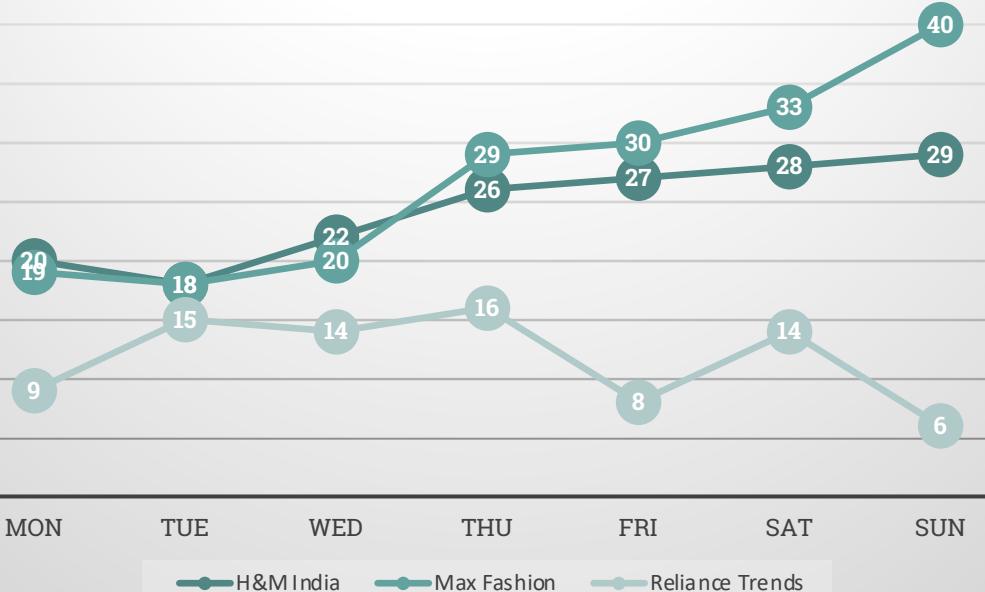


- Crisp, clear, to-the-point content grabs attention! Not the verbose ones!
- More Frequent you are, the better the engagement!

Let the weekend begin...



Weekly pattern of posting



- Planned & structured approach wins the race!
- It is all about the weekend!
- No vision makes it impossible!

Data Exploration:



08. Data Preparation

“Give me six hours to chop down a tree and I will spend the first four sharpening the axe”.

- Abraham Lincoln

Facebook Features



Word Features



Tokenization



Stemming

- Replace @ user tag
- Replace emojis with a meaningful text
- Replace occurrences of urls.
- Replace # hashtags

- Remove upper capitalization
- Word repetition
- Punctuation repetition
- Word contraction

- NLTK
- word_tokenize module from NLTK

- PorterStemmer

Original Post (input): Lookback for a moment, thank yourself for getting through this year and get ready for new 365 days, only and only filled with happiness!!

Tokenized Post (output): ['lookback', 'moment', 'thank', 'get', 'year', 'get', 'ready', 'new', 'day', 'fill', 'happiness']

“No data is clean, but most is useful”.

- Dean Abbott

Sale

Posts that are about offers, and discounted merchandise communication.

- Example: It is never too late to grab amazing deals.

With #NewBeginnings revamp your wardrobe with #MaxEndOfSeasonSale only at our #MaxFashionStore near you.

Hurry! Get up to 50%

Season/Campaign

Posts dedicated to seasonality and marketing communication.

- Example: With #NewBeginnings it is time to give your closet a fresh makeover. 🛍️ Shop at the nearest #MaxFashionStore now and avail the amazing collection.

Product

Posts that are about the product and the silhouettes.

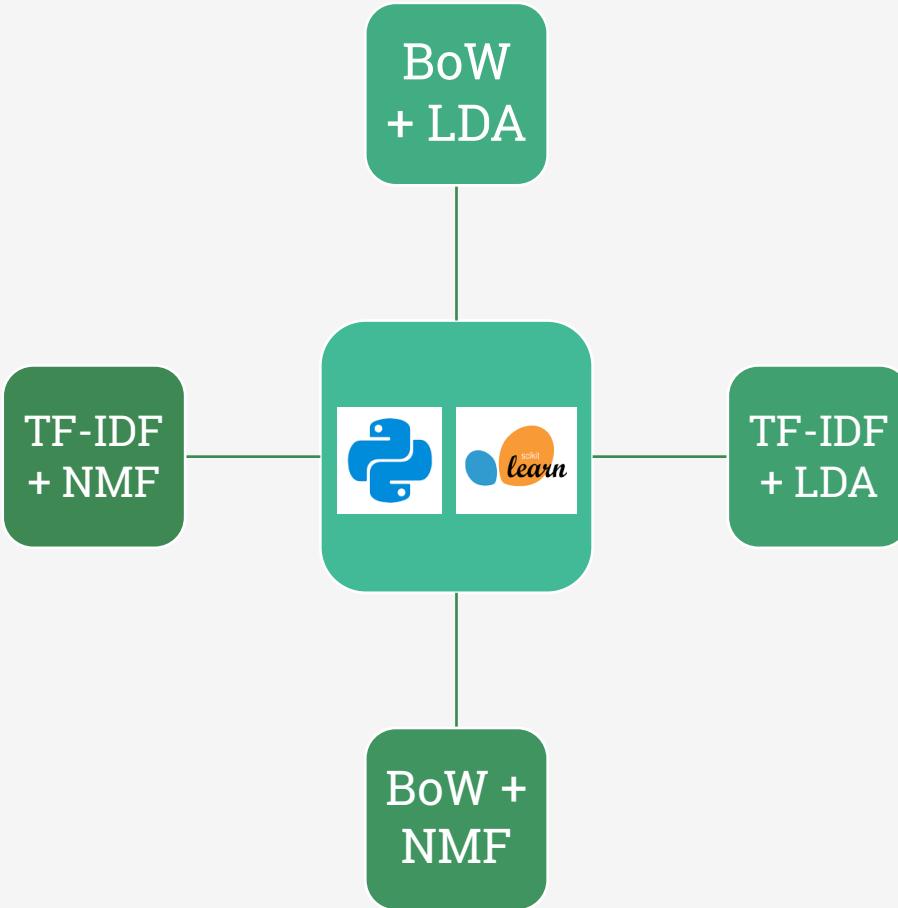
- Example: Pair up a Hooded Sweatshirt with your favourite Blue Denims, and you are set for the #NewBeginnings. Check out the collection at our #MaxFashionStore near you. 🛍️ #HappyWeekendShopping

Local language

Posts that are in local languages such as Hindi, Kannada, etc.

- Example: Logo ka kaam hai kehna, log toh kuch na kuch kahnege hi. But all I want to say to all the women out there is #YouAreSoBeautifull. Check out @maxfashionindia to celebrate and appreciate the smart, kind, funny, and lovable woman that is you.

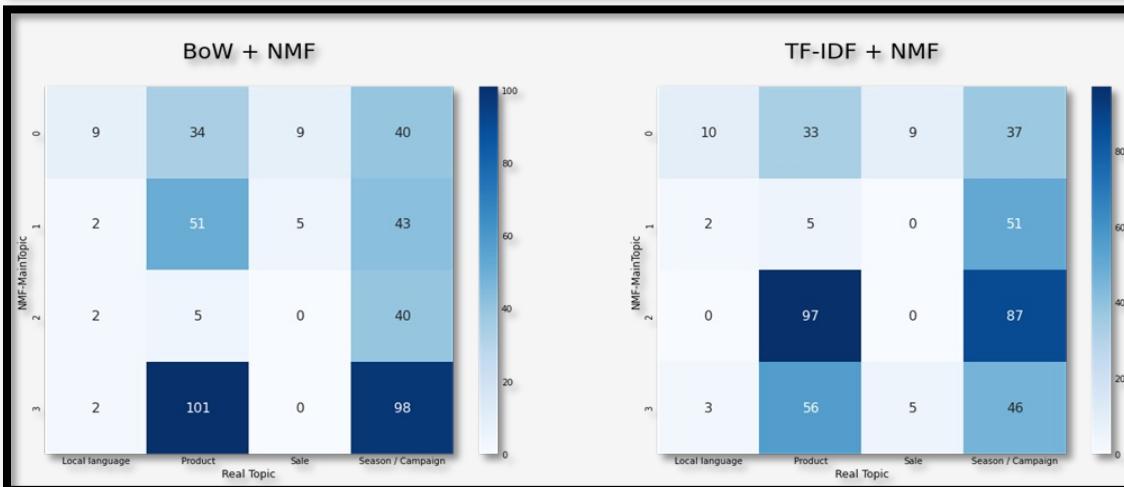
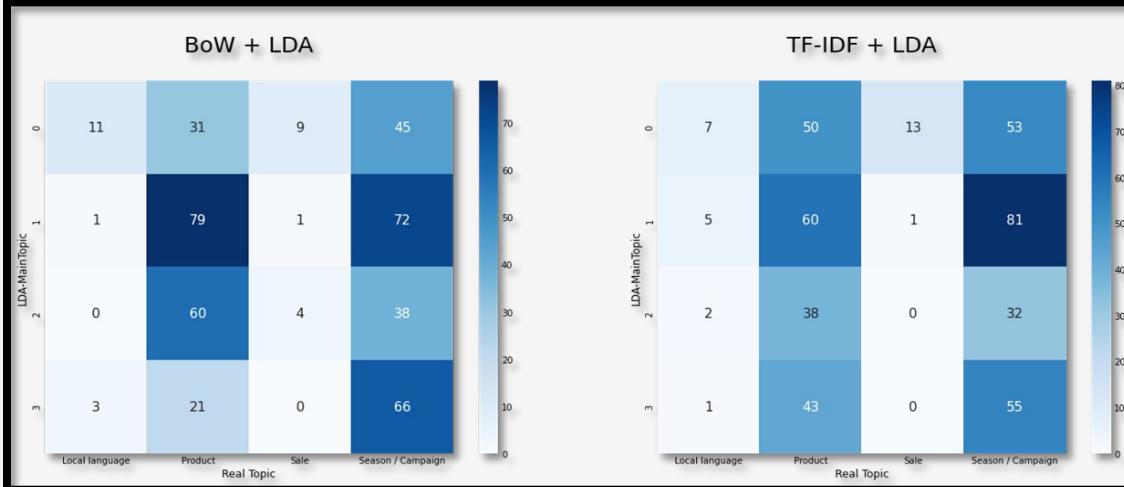
09. Modelling



- Can these unsupervised models detect the topics and is it as per the manually labeled topics?

10. Model Evaluation

Confusion Matrices:



Let's assess the models!

Model Results:

Topic	TF-IDF + LDA	BoW + LDA	TF-IDF + NMF	BoW+ NMF
Local Language	71.88%	79.82%	80.95%	79.82%
Product	50.57%	57.82%	45.80%	56.92%
Sale	80.50%	75.51%	55.10%	86.17%
Season/Campaign	52.385	59.41%	45.80%	48.75%



BoW + LDA is the best approach!

11. Analysis and Results

“We are surrounded by data but starved for insights”.

- Jay Baer

Insights:

- Crisp, clear, to-the-point content grabs attention! Not the verbose ones!
- More Frequent you are, the better the engagement!
- Planned & structured approach wins the race!
- It is all about the weekend!
- No vision makes it impossible!

Business Direction:

- A new entrant must focus on brand building via Facebook.
- The stronger category of the new entrant needs to be the hero of the content.
- Add newness and freshness!
- Product-specific attributes are highlighted in the content such as colors, prints, designs, etc.
- Promote other channels such as offline stores, Ecommerce, etc.

12. Conclusions and Future Scope

“All good things come to an end”.

- Geoffrey Chaucer

Conclusion:

- BoW + LDA is the best approach.
- To enhance the performance, the corpus needs to be expanded.
- The scope of this study was limited to four months brand generated content.

Future scope and points of discussion:

- Acquire data for at least an year to ensure all seasonality is covered.
- Apart from Facebook posts, collect data from Instagram & Twitter.
- Expand the competitive intelligence analysis to other value players such as easy buy, Zudio, westside, etc.

Bibliography | Webliography

Bibliography:

- Capitalone. (n.d.). *Definition of TF-IDF - Capitalone*.
- Deloitte Digital. (2022). *Apparel Trends 2025*. <https://www.deloittedigital.com/content/dam/deloittedigital/us/documents/blog/blog-20200610-apparel-trends.pdf>
- Hassani, A., & Mosconi, E. (2022). Social media analytics, competitive intelligence, and dynamic capabilities in manufacturing SMEs. *Technological Forecasting and Social Change*, 175, 121416. <https://doi.org/10.1016/j.techfore.2021.121416>
- He, W., Shen, J., Tian, X., Li, Y., Akula, V., Yan, G., & Tao, R. (2015). Gaining competitive intelligence from social media data Evidence from the two largest retail chains in the world. In *Industrial Management and Data Systems* (Vol. 115, Issue 9, pp. 1622–1636).
- Emerald Group Holdings Ltd. <https://doi.org/10.1108/IMDS-03-2015-0098>
- He, W., Tian, X., Tao, R., Zhang, W., Yan, G., & Akula, V. (2017). Application of social media analytics: A case of analyzing online hotel reviews. *Online Information Review*, 41(7), 921–935. <https://doi.org/10.1108/OIR-07-2016-0201>
- He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, 33(3), 464–472. <https://doi.org/10.1016/j.ijinfomgt.2013.01.001>
- H&M. (n.d.). *H&M official facebook page*. Retrieved August 18, 2022, from <https://www.facebook.com/hm/>
- H&M website. (n.d.). *H&M official website*.
- IBM. (n.d.). *Cross-industry standard process for data mining Lifecycle*.
- Max Fashion India. (n.d.). *Max fashion official facebook page*. Retrieved August 18, 2022, from <https://www.facebook.com/maxfashions/>
- Maxfashion India official website. (n.d.). *Maxfashion India official website*.
- McKinsey & Company. (2022). *The State of Fashion 2022*. <https://www.mckinsey.com/~media/mckinsey/industries/retail/our%20insights/state%20of%20fashion/2022/the-state-of-fashion-2022.pdf>
- Mohammad, R. Z., Abbasi, A., & Liu, H. (2014). *Social Media Mining: An Introduction*. <http://dmml.asu.edu/smm>,
- Reliance Trends. (n.d.). *Reliance Trends official facebook page*.
- relianceretail. (n.d.). *Relianceretail official website*.
- Statista Digital Market Outlook. (2021). *Fashion eCommerce report 2021*.
- Techopedia. (n.d.). *Bag of Words (BoW)*.
- Tong, Z., & Zhang, H. (2016). *A Text Mining Research Based on LDA Topic Modelling*. 201–210. <https://doi.org/10.5121/csit.2016.60616>
- Wang, J., & Zhang, X.-L. (2021). *Deep NMF Topic Modeling*. <http://arxiv.org/abs/2102.12998>
- Wazir Advisors. (2022). *Wazir Report - The Road to 2025*. 1–32. <https://wazir.in/pdf/Wazir%20Report%20-%20The%20Road%20to%202025.pdf>
- Wikipedia. (n.d.-a). *Latent Dirichlet Allocation (LDA)*. Retrieved August 18, 2022, from https://en.wikipedia.org/wiki/Latent_Dirichlet_allocation#:~:text=In%20natural%20language%20processing%2C%20Latent,example%20of%20a%20topic%20model.
- Wikipedia. (n.d.-b). *Non-negative matrix factorization*.

Additional Information | Plagiarism score

Competitive Analysis of Value Fashion Brands in India with Social Media Analytics			
ORIGINALITY REPORT			
SIMILARITY INDEX	10% INTERNET SOURCES	4% PUBLICATIONS	7% STUDENT PAPERS
PRIMARY SOURCES			
1 www.researchgate.net Internet Source	3%		
2 www.capitalone.com Internet Source	1%		
3 ukcatalogue.oup.com Internet Source	1%		
4 github.com Internet Source	1%		
5 en.wikipedia.org Internet Source	1%		
6 Submitted to University of South Africa Student Paper	<1%		
7 export.arxiv.org Internet Source	<1%		
8 www.statista.com Internet Source	<1%		
9 Submitted to University of College Cork Student Paper	<1%		

10 www.ilsul6ana.com Internet Source	<1 %
11 repository.sharif.edu Internet Source	<1 %
12 www.alianzo.com Internet Source	<1 %
13 Submitted to Tampere University Student Paper	<1 %
14 Submitted to Massey University Student Paper	<1 %
15 mdpi-res.com Internet Source	<1 %
16 Submitted to University of Huddersfield Student Paper	<1 %
17 He, Wu, Shenghua Zha, and Ling Li. "Social media competitive analysis and text mining: A case study in the pizza industry", International Journal of Information Management, 2013. Publication	<1 %
18 core.ac.uk Internet Source	<1 %
19 Submitted to Army Institute of Technology Student Paper	<1 %
20 Submitted to SP Jain School of Global Management	<1 %

Student Paper			
21 www.preprints.org Internet Source	<1 %		
22 www.science.gov Internet Source	<1 %		

Exclude quotes On Exclude matches < 10 words
 Exclude bibliography On



*Thank
you!*