Project 2 Report

Anthony Albanese
Binghamton University
Binghamton, New York, USA
aalbanese6@binghamton.edu

Kyle Enriquez
Binghamton University
Binghamton, New York, USA
kenriqu1@binghamton.edu

Ameen Kunnathu
Binghamton University
Binghamton, New York, USA
akunnat1@binghamton.edu

Andy Zheng
Binghamton University
Binghamton, New York, USA
azheng74@binghamton.edu

Abstract

In this study, we explore the landscape of fashion trends, and consumer behavior through an analysis of data from Reddit and Ebay. The primary objective was to identify correlations between user discussions on Reddit's fashion-related subreddits and fashion apparel sales on eBay. This exploration aimed to uncover the underlying trends in fashion preferences, and price sensitivity within these platforms. Our analysis employed data from subreddits such as r/findfashion and r/FashionReps alongside eBay apparel sales,to gain insights into real-time fashion trends, user discussions, and pricing strategies. We examined the frequency of discussions based on brands, toxicity levels, and diversity of products and brands within these subreddits. Additionally, we analyzed eBay's apparel sales data to determine pricing trends across various fashion categories. The findings revealed contrasts between the userbases of Reddit and Ebay. Contrary to our initial hypothesis, there was no strong correlation between the user demographics of the two platforms. Our research also highlighted a trend toward individualized fashion choices over popular brands, along with a discernible preference for cost-effective, thrifty fashion purchases. This study contributes to the understanding of online consumer behavior in the fashion industry, emphasizing the diverse and dynamic nature of fashion trends as reflected in social media discussions and e-commerce activities. It provides a foundational framework for future research, pointing to the need for more extended and diverse data collection and analysis to fully grasp the intricacies of fashion trends and consumer preferences in the digital age.

1 Introduction

For our project, we analyzed data from different sources to determine general trends and correlations in fashion apparel. To do so, we closely analyzed Reddit and eBay as sources of data. For Reddit, we analyzed the subreddits r/findfashion and r/FashionReps to analyze discussion and trending topics in fashion. As for eBay, we focused on sales of apparel and clothing being auctioned recently to view their prices and categories. Through these sources, we are motivated to determine notable parallels and contrasts in the data to give us a stronger understanding of fashion trends. In particular, this would give us crucial information on the user base demographic of the sources, as well as overall fashion trends over time and as a whole. Through this, we can make accurate and educated decisions and conclusions on the matter by understanding the corresponding patterns and opinions. As a result, being able to continuously gather data and display our findings in a visually useful and understandable manner is crucial for making accurate and concrete claims.

2 Background and Related Work

For our analysis, the primary sources we studied included Reddit and eBay as domains. As a whole, both Reddit and eBay are consistently active sites which continue to rise in popularity. Through the rapid growth of technology in recent years, modern means of communication, such as smartphones, computers, and the internet, become the norm of society as they are widely made available worldwide. Through these devices, the development and distribution of social media platforms were effectively facilitated, growing in parallel to accompanying technology advancements [4]. As a result, social media evolved to be a central part of an increasingly digital society, being utilized as a primary source to accomplish multiple tasks. In particular, these sites established a role as being a means of evaluating the quality and trends

in product choices. To begin, social media allows for the ubiquitous and easily accessible generation and exchange of information. Through this, an increasing number of people use social media as a means of acquiring feedback of specific products which spark their interest [3]. This trend directly correlates to the realms of Reddit and eBay. For example, Reddit was initially distrusted as a reliable source of branding due to its notoriety of harassment and internet trolling. However, its reputation shifted over time as people realized the benefits of its varied community platform, as well as its comprehensive and straightforward means of accumulating and spreading information. This transcended to brands utilzing Reddit as a means of advertisement, including eBay investing in Reddit as a way to tailor specific products to a desired audience. [2] Through this, Reddit has proven to be a primary source to determine the general opinion of a product. This has been proven in past studies of internet phenomenons, such as being used to gauge user feedback on Etsy activity [1]. As a result, the development of social media trends and uses acts as the foundation for our experiment, being used to gauge user activity and general consensus in a specific industry, fashion.

3 Data Sets

Through our experiment and data collection, various data sets can be shown exhibiting relations in our dataset. To begin, below shows the data collected from r/politics:
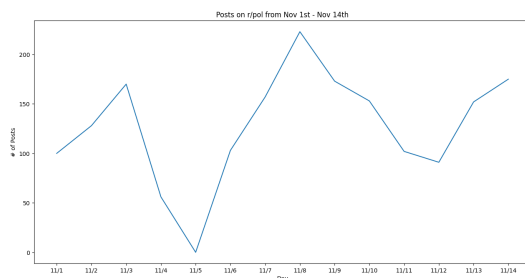

Figure 1. Number of Submissions in r/politics

The graph demonstrates the number of submission made in r/politics daily, with the x-axis being time in days while y is the number of submissions made, with early discrepancies in the data deriving from the virtual machine shutting down. In addition, we also analyzed data from our studied subreddits using ModerateHateSpeech, which is a 3rd party API that adds real-time measurements of toxicity to our data collection system:
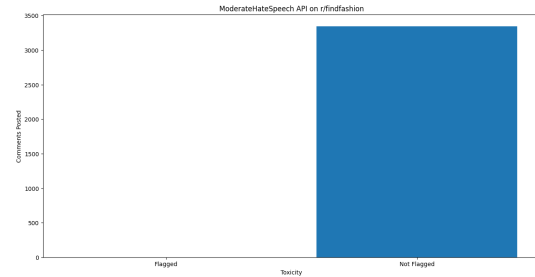

Figure 2. Toxicity in r/findfashion
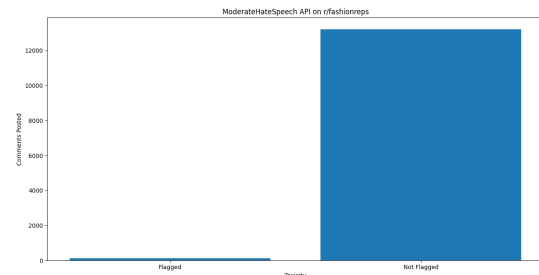
Figure 3. Toxicity in r/fashionreps

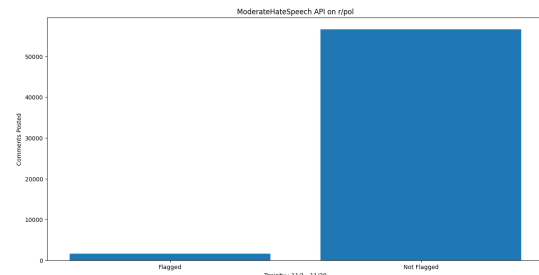**Figure 4.** Toxicity in r/politics

Furthermore, various datasets can be derived from the main sources we studied, which were r/FashionReps, r/findfashion, and the apparel category on eBay. To begin, as a means of comparing all of our datasets in one figure, we created a graph which demonstrates the compared variety of product types found between r/FashionReps, r/findfashion, and eBay listings seen here:
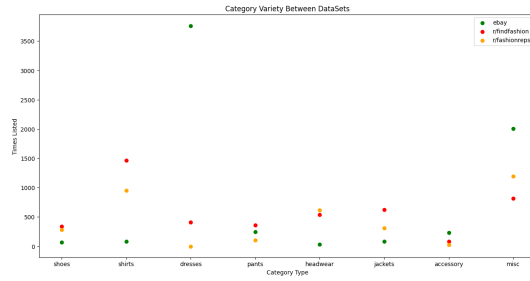
Figure 5. Category Variety in Each Source

In addition, we examined elements of Reddit and eBay individually to gather further understanding of our datasets. To begin, below is the average prices of eBay listings, divided based on their associated category:
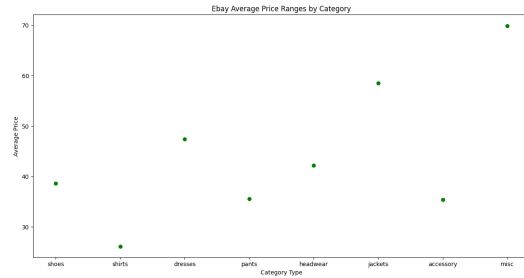

Figure 6. eBay Prices by Category

In addition, we also compared our two studied subreddits, r/findFashion and r/fashionReps, and compared the brand variety of each source seen here:
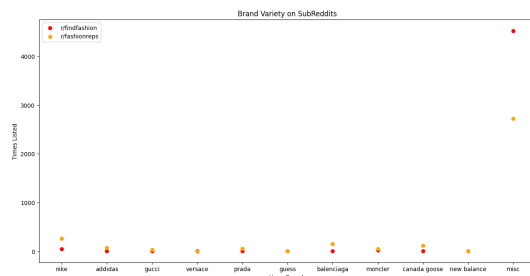

Figure 7. Brand Variety in Subreddits

Finally, we also created a table view of the average number of top brands found in r/findfashion, importing our data into a Latex table generator:

	Nike	Adidas	Gucci	Versace	Prada	Guess	Balenciaga	Moncler	Canada Goose	New Balance	Misc
Shoes	258	9	23	6	23	20	21	73	30	20	0
Shirts	940	53	115	21	126	105	125	352	185	90	12
Dresses	263	16	25	5	46	37	16	85	63	19	1
Pants	233	5	20	1	36	27	35	91	45	23	3
Headwear	0	0	0	0	0	0	0	0	0	0	0
Jackets	0	0	0	0	0	0	0	0	0	0	0
Accessory	0	0	0	0	0	0	0	0	0	0	0
Misc	752	66	75	22	94	114	75	311	162	52	19

Table 1. Average Brands in r/findfashion

4 Discussion

4.1 Rough Summary

Through our datasets, we were able to extract a vast amount of information from the sources we analyzed. To begin, from r/politics, there appears to be a sporadic level of activity on a day-to-day basis, as seen in figure 1. This is representative of subreddit activity as a whole, as it often derives from unpredictable roots and unforeseen causes on an exterior basis. In addition, when we analyzed the levels of hate speech in our analyzed subreddits, the presence of hate speech in r/politics shown in figure 4 eclipses that of r/findfashion, which has a negligible amount of hate speech, but contrarily is surpassed by that of r/fashionreps, as seen in figure 2 and figure 3, respectively. From this finding, it demonstrates the varied dynamics of the communities of individual subreddits, while also proving that they can be quite varied and dependent on real-time and outside events. Along with this, various interpretations can be made with the figures pertaining to our main data sources: r/findfashion, r/fashionreps. and the apparel section of eBay. We examined the clothing category variety between each source and determined that there is a notably varied mixture in each source, shown in figure 5. However, a distinct disparity exists in eBay, where there is a strong emphasis and outlier of women's fashion clothing, namely dresses, indicating differences in viewership between eBay and the subreddits. In addition, we also compared the two individual subreddits we studied and examined for trends in brand variety. As a whole, both subreddits portray revolve around the same trends in their clothing items, consisting of small subsets of large brand names in stark contrast with the predominant presence of miscellaneous brand items. Finally, we examined individual sources to more closely examine specific elements of each. Firstly, we determined from figure 6 that the largest pricing items originate from the dresses, jackets, and miscellaneous categories. These prices all reside under the 70 USD range, which indicates more frugal spending in online shopping as opposed to traditional retail shopping. Furthermore, when we viewed the average number of clothing category items in specific brands from a given day, as seen in table form in figure 7, we can determine that discourse predominantly revolves around traditional clothing items, such as shirts and pants, while there was seldom activity in niche items, such as headwear and accessories. This data can prove indicative of the typical activity in mainstream clothing companies.

4.2 Implications

From our data, various implications can be made, which directly correlate to our research questions. To begin, the first question we aimed to answer is if there is a correlation between the user bases of Reddit and eBay. We initially predicted that the trends in both our studied subreddits and the apparel section on eBay would show notable parallels, as we believed that users who delve into these areas would be of a similar demographic. However, our experiment proved the opposite, with women's fashion and dresses being a predominant subset of the eBay listings collecting, while the subreddit categories were very mixed in contrast, as seen in figure 5. As a result, this demonstrates that there is no clear connection between the activity on the subreddits and the listings on eBay, indicating a clear disparity in their respective user bases. In addition, we also aimed to analyze the presence of top brand items in media and determine if they are the leading cause of discourse in fashion discussion. Through our findings, we found that this is predominantly not the case. As seen in figure 7, while certain top main name brands may occupy a significant subset of posts, it pales significantly to the percentage of posts made by non-leading brands. As a result, this implies a trend towards individualized styles and distinguishing oneself away from the norm, as opposed to conforming to popular clothing brands. Finally, we aimed to determine if price is a factor when accounting for common fashion trends. When analyzing the prices of average sales of items, this proves to be the case. In particular, as seen in figure 6, the average price of sales for all items prove to be below 70 USD, with each category's average prices being below retail pricing, especially designer items. This is further proven by the Bullseye deals section in eBay acting as a major source of sales in our data set. As a result, this is indicative of a pattern towards more thrifty and cost-effective spending in regard to fashion sales, proving price to be an important factor in fashion trends.

4.3 Limitations

While our exploration of the data proved to be extremely thorough, limitations exist within our study which should be acknowledged for future endeavors. For example, our study only took place over the course of the semester, meaning that data was only collected in the span of about 3 months and in around one seasonal period. To acquire a more well-rounded understanding of our target area, data can be collected throughout the year to remove the restriction of trends being confined to a specific season. This study can be expanded further by gathering data over the span of several years to analyze changes in fashion trends over time. In addition, we are also restricted by the number of data sources we utilize. While Reddit and eBay prove effective for determining user opinion and concrete purchases on items, being limited to only these sources may lead to discrepancies. This

is predominantly due to the user bases of these sites developing their own respective culture and viewpoints and not being directly indicative of the general public. To remedy this, we can expand this experiment to incorporate data from alternative platforms to vary the demographics in which we receive our information. This can include analyzing selling platforms such as Amazon, Etsy, and Facebook Marketplace along with eBay, as well as social media sites such as Pinterest, Twitter, and Instagram in conjunction with Reddit. Furthermore, we are confined by the technical limitations we have to work around. To begin, the virtual server has a finite storage capacity and inconsistent performance capabilities, which results in us being able to gather data only a limited number of times and potentially miss crucial data from being collected. Along with this, we are also limited by the APIs we need to use to access Reddit and eBay, as their advanced functionalities and prolonged usages are restricted by paywalls and specialized memberships. As a result, if given the proper authorization, resources, and funding, we can potentially expand the scope of our experiment from its current state. In the end, despite the limitations, the experiment proved to be very comprehensive and effective in accomplishing our main objective of determining fashion trends from Reddit and eBay.

5 Conclusion

In conclusion, as social media platforms developed and evolved, their current state proved to be an effective means of analyzing trends and patterns in fashion. For our experiment, we utilized Reddit and eBay as our main platforms for analysis, specifically the subreddits r/findfashion and r/fashionreps and the apparel section on eBay. Through the analysis of these sources, along with r/politics and the ModerateHateSpeech API, various figures were developed, allowing us to make corresponding interpretations and conclusions. These conclusions include a notable disparity in the user bases and typical activity of these sources, a shift in popularity towards non-traditional brands of clothing items, and the notable influence that price has on modern clothing trends. In the end, through our comprehensive analysis of the data we gathered, we can develop a firm understanding of this topic and create concrete and detailed claims to our primary research questions, which can continue to be expanded upon by defying our encountered limitations.

References

[1] Ryan Deffenbaugh. 2022. Etsy reports record sales: Sellers respond with 300 comments on Reddit in 24 hours. *Crain's New York Business* 38, 9 (2022), 10–. https://doi.org/10.1057/9781137399687_2

[2] Lauren Johnson. 2016. Big Brands Enter Reddit's Viral Fray. *Adweek* 57, 22 (2016), 8–. https://doi.org/10.4324/9781003150800-3

[3] Qingqi Pei Lichuan Ma, Xuefeng Liu and Yong Xiang. 2019. Privacy-Preserving Reputation Management for Edge Computing Enhanced Mobile Crowdsensing. *IEEE Transactions on Services Computing* 12, 5 (2019), 786–799. https://doi.org/10.1109/MSP.2013.2282414

[4] Esteban Ortiz-Ospina. 2019. The rise of social media. *Our World in Data* 1, 1 (2019), 1–12. https://doi.org/10.4324/9780203794807-7