



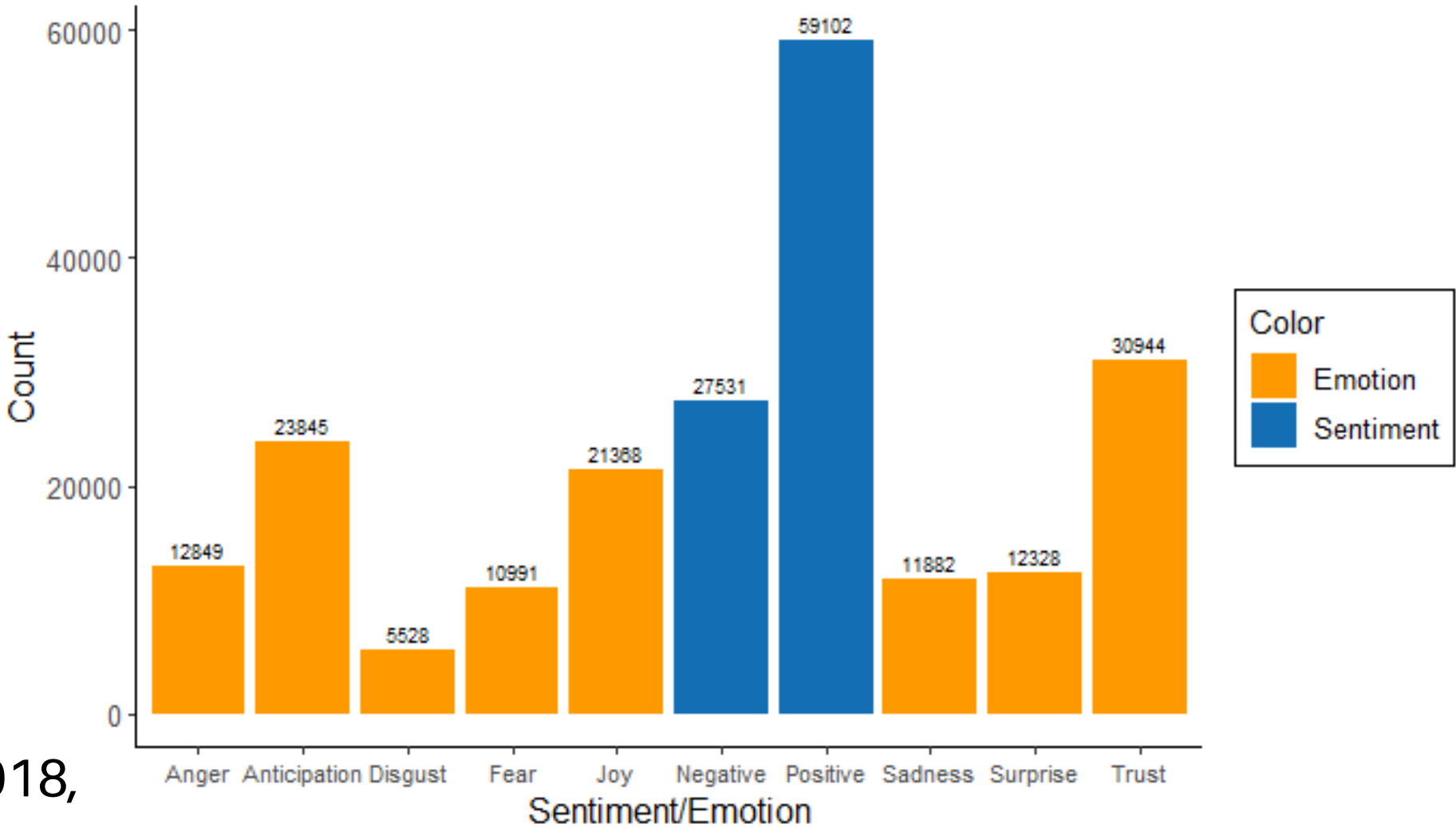
A Deep Dive Into Spotting Fake Reviews On amazon

Authors: Uyen Ha, Duy Nguyen, Thao Duyen Tran

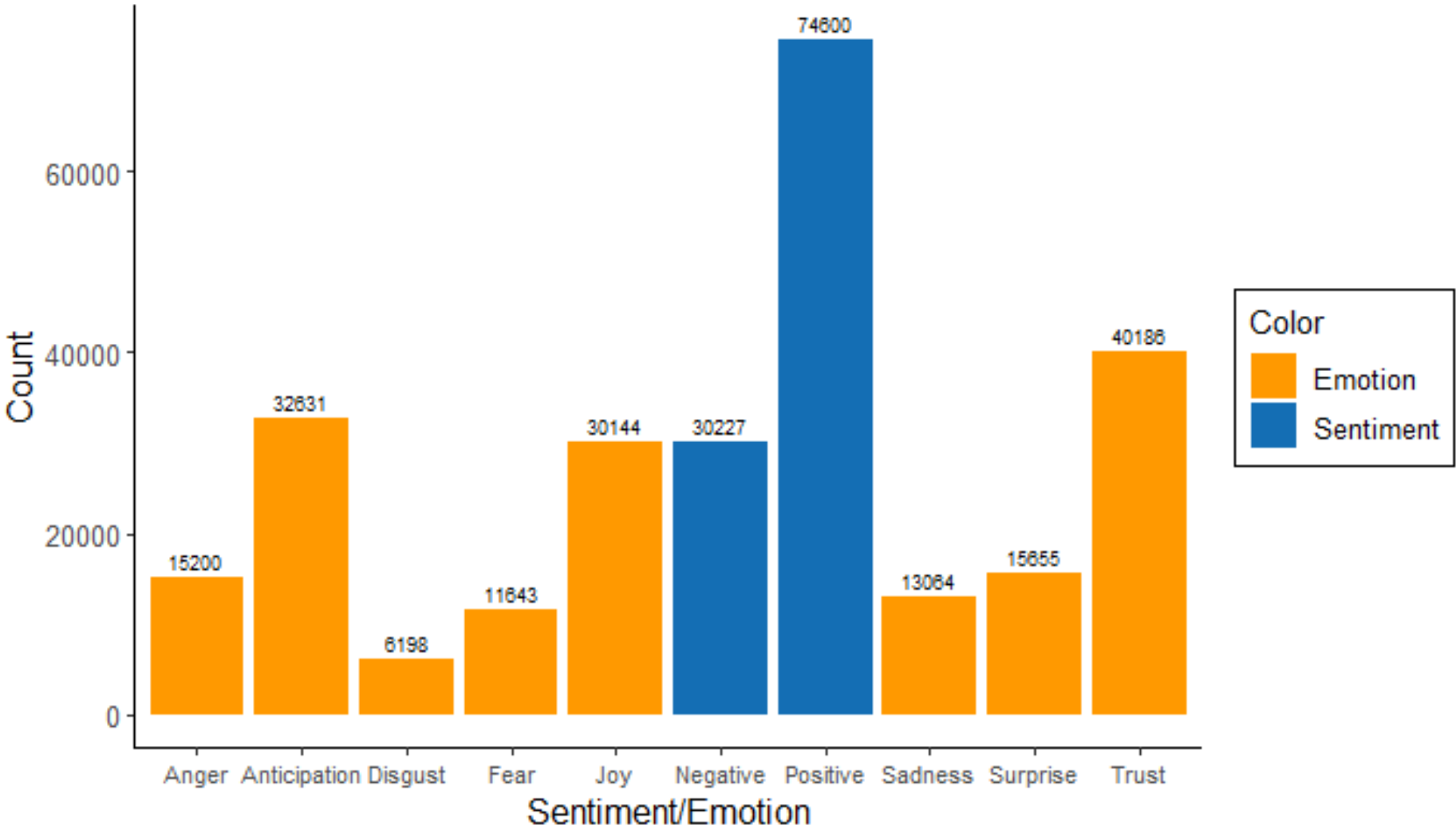
Sentiment Analysis

- Unpleasant/negative reviews increased from 2008 to 2018
- Positive sentiment and trust also increased.
- However, considering the vast expansion of Amazon from 2008 to 2018, it is expected that the number of positive sentiment and emotion should increase by a higher magnitude that is comparable with its scale.

Sentiment Analysis for 2008's 5-star Reviews (June-Sept)



Sentiment Analysis for 2018's 5-star Reviews (June-Sept)



Word use in unverified and verified reviews are identical

Text Mining

Most frequent-used words in Electronics product reviews on Amazon

2008



2018



verified

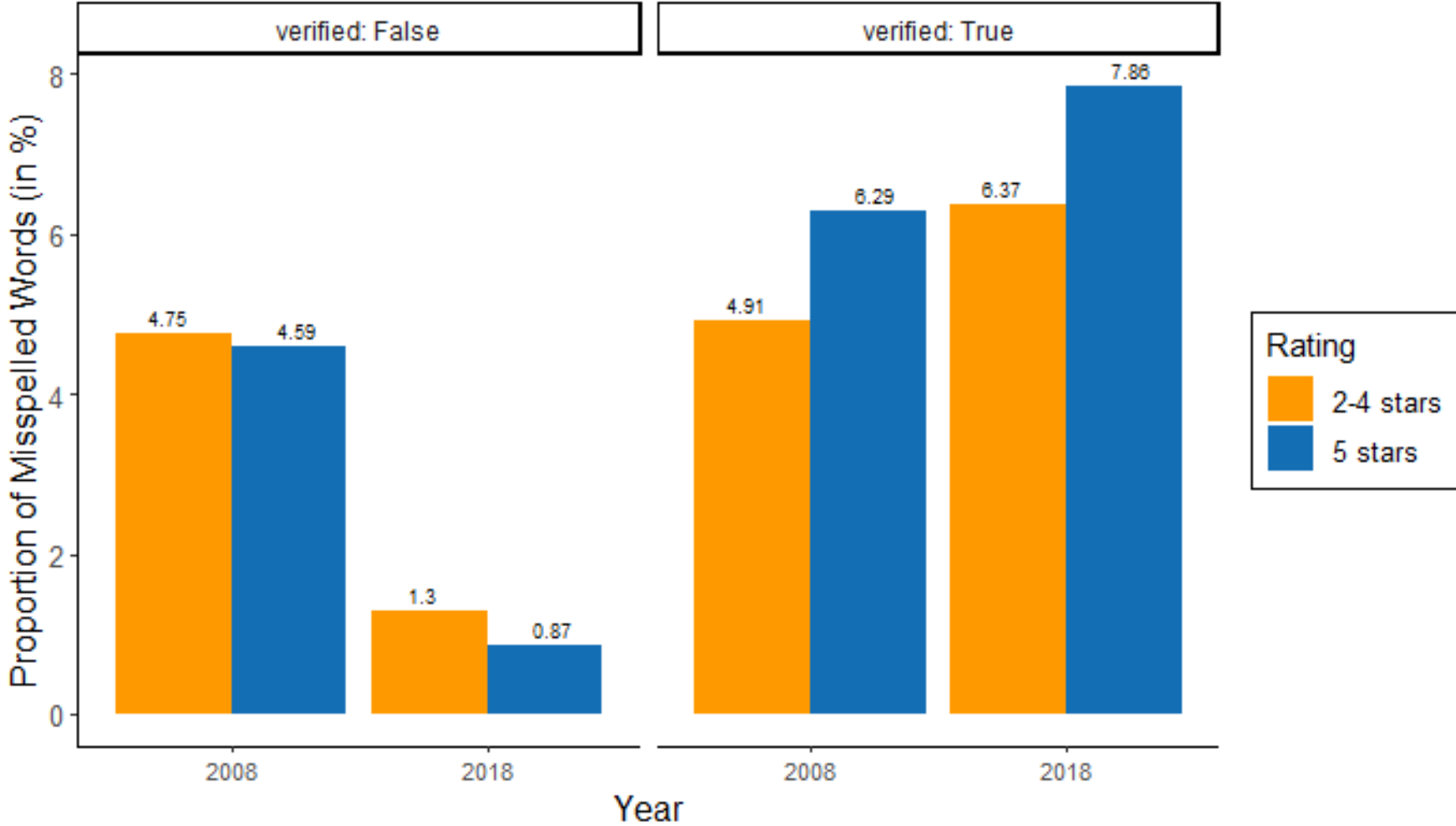
unverified

- NLP helps provide an overall summary of the content in the reviews, in other words, how customers think about the products.
- However, as unverified and verified reviews are very much similar. It has become trickier to detect fake or bot reviews.

Possible methods to detect fake reviews

- Misspelling check
- Similar or very short reviews detection
- A large number of reviews of one product is generated within a short amount of time
- Unusual proportion in the number of 4-5-star vs 1-star reviews. (eg: 70% 5-star vs 30% 1-star)

Proportion of Misspelled Words in 2008 and 2018 (Jun to Sept)



- Unverified reviews that give 2-4-star rating have higher proportion of misspelling words
- Fake reviews for 5-star rating are more generated more sophisticatedly with less misspelling words

Similar Reviews	Freq
works great	1332
good	834
great	773
ok	396
excellent	193
great product	92
works well	89
perfect	82
works	81
good product	76
love it	74
nice	63
very good	59

Investigate similar/duplicated reviews

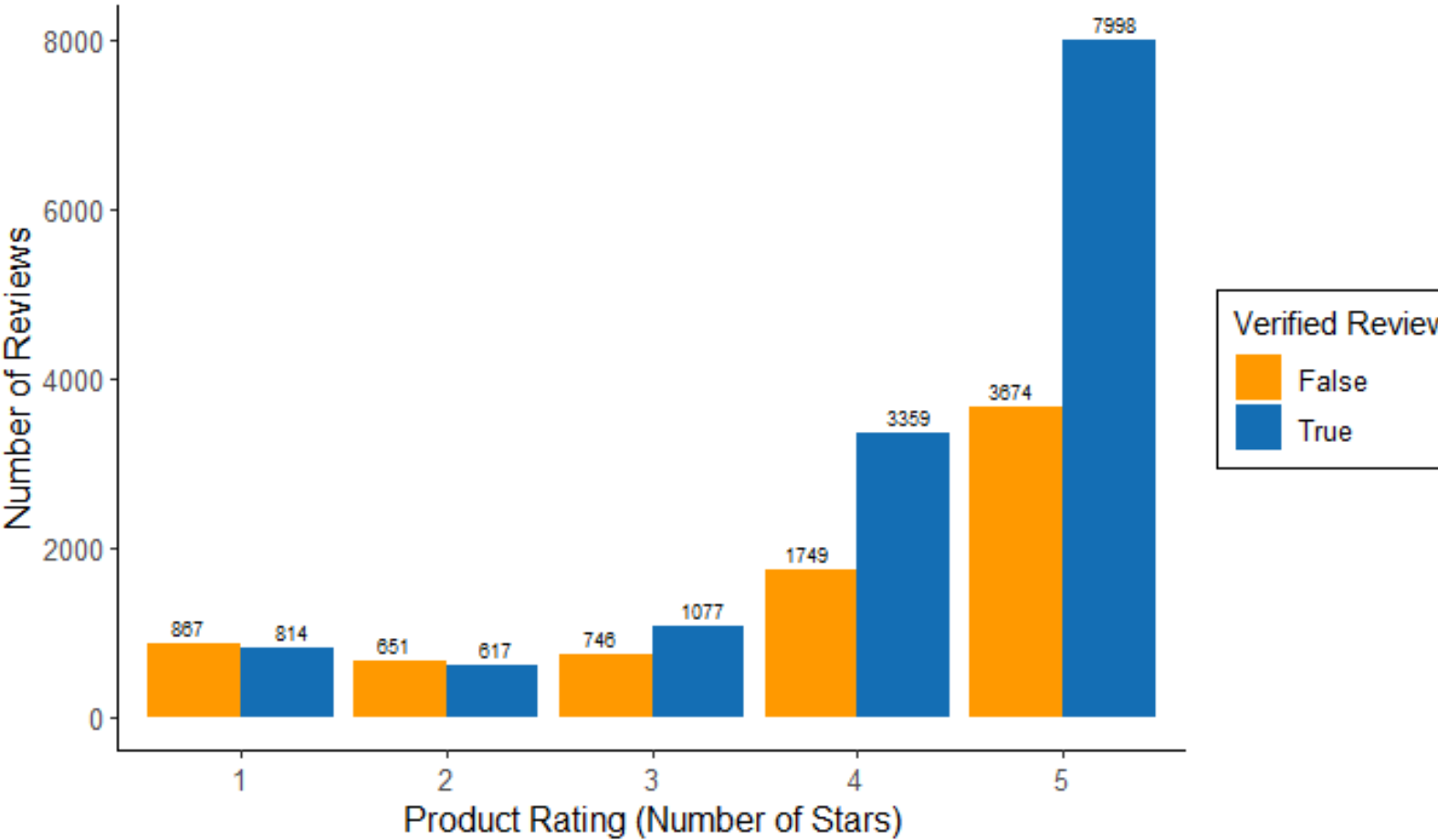
- Thousands similar/duplicated reviews from different users and different products are detected.
- Most of them contain generic words such as: great, good, perfect, nice... and either are extremely positive (5-star) or negative (1-star) but don't specify the reasons. These could be the red flags that the review is fabricated or bot-generated.
- Word cloud of similar reviews particularly contains higher frequency of phony/ extreme positive/ negative words



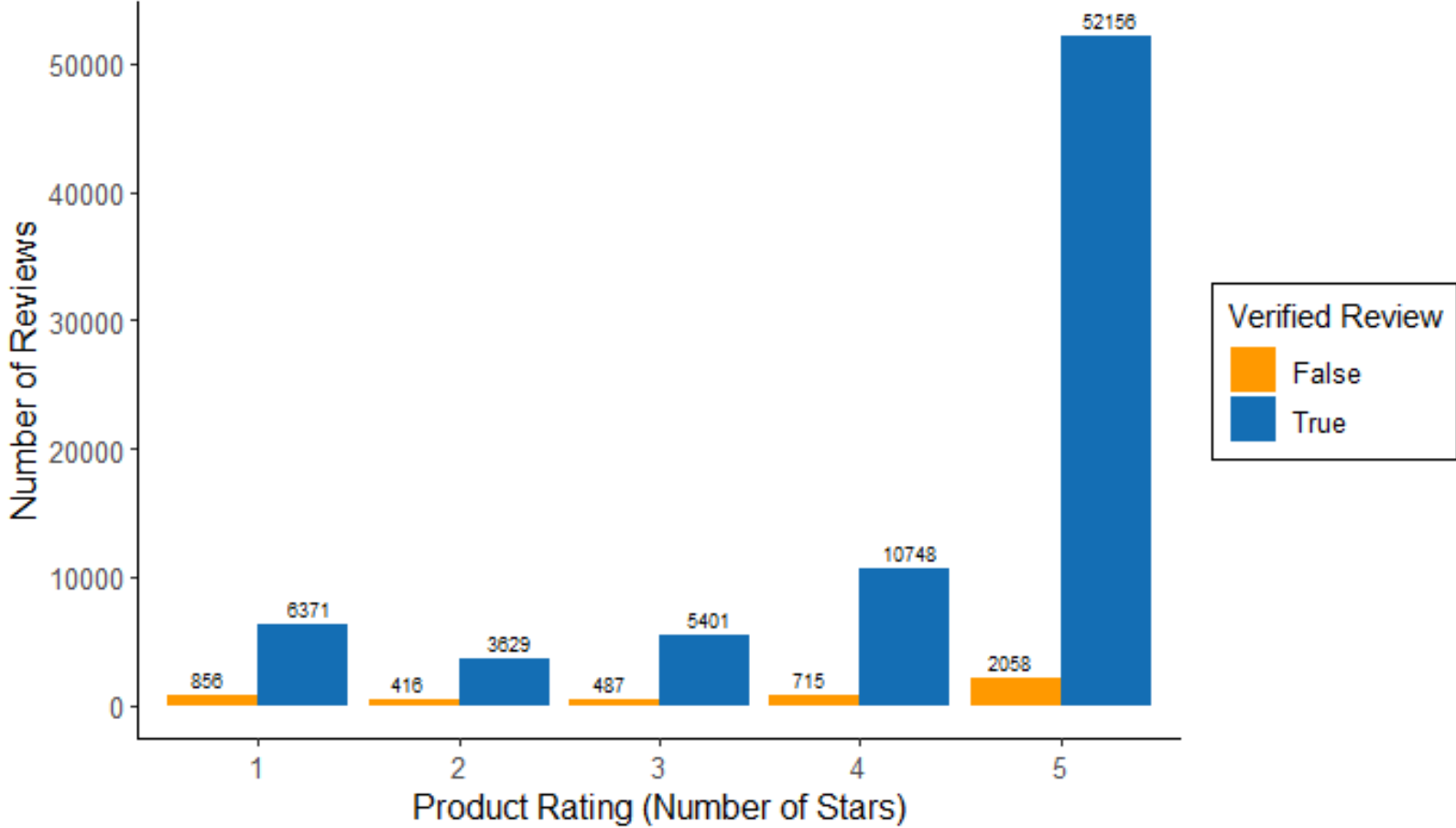
Amazon took action

- Invested more than \$400M
- Blocked more than a million bad actor accounts before they published a single listing
- Blocked more than 3B suspected bad listings before they were published to our stores
- Prevented more than 13M attempts to leave an inauthentic review
- Recommendations
 - Build bot chat detecting review trading conversations on Amazon web chat
 - Set stricter standards for verified purchases
 - Increase account privacy, security and protection

Number of Verified Reviews in 2008 (June-Sept)



Number of Verified Reviews in 2018 (June-Sept)



Unverified Reviews on Amazon decreased significantly from 2008 - 2018 thanks to serious effort that Amazon made to address the issue

A Deep Dive Into Spotting Fake Reviews on Amazon

Background

Amazon is one of the world's most successful technology companies, known worldwide for its e-commerce, digital streaming, and artificial intelligence capabilities. What started as an online bookstore in 1994, founded by Jeff Bezos, quickly rose above its competitors as Bezos realized the value of the e-commerce industry and expanded the company's offerings. As of now, nearly three decades after its conception, Amazon is one of the world's largest online markets, with over 12 million Amazon-branded products, as well as over 300 million marketplace products for sale¹. In the U.S, Amazon accounts for 49% of the e-commerce market, and 5% of all retail spending.²

To better understand how Amazon operates and generates value, we can assess its value chain model. A company's value chain can be studied to see gaps and opportunities for expansion. Evaluating each phase in a company's value chain allows us to find opportunities for development, specific to said phase. The design, implementation, and, subsequently, optimization of Amazon's value chain analysis are vital to the company's success.

Amazon's Value Chain

Inbound Logistics

Inbound logistics refers to a company's actions that involve acquiring raw materials. Amazon rarely has its products. It operates as a marketplace for other companies' goods and services. Amazon has few long-term partnerships with these companies or suppliers. Fulfillment by Amazon (FBA) is offered to inbound logistics companies.

Amazon requires them to supply their goods in advance to sell them. Amazon handles shipping, customer service, and returns after receiving the products. Sellers can choose not to use FBA services at all. So retailers can handle their shipping and other services. But most vendors use Amazon's services. Logistics serve Amazon's marketplaces. The firm also provides logistical support to other firms. Various subsidiaries offer business shipping services. Amazon has a considerable logistical advantage over many of its competitors.

Additionally, when Amazon expanded, it began offering its own "basic product" lines to customers. In comparison to third-party products, Amazon-branded products have significantly more simplified logistics as they are able to ship products immediately from their storage to the client, without having to wait for a third-party seller to deliver on the product beforehand. As a result, Amazon can secure a competitive advantage over third-party sellers by creating their own lines of popular products such as housewares, clothing, basic electronics.

Operations

In a value chain, operations refers to the actions that transform raw resources into finished commodities. The operations of Amazon are spread across a variety of channels, the

¹ <https://www.retailtouchpoints.com/resources/how-many-products-does-amazon-carry>

² <https://techcrunch.com/2018/07/13/amazons-share-of-the-us-e-commerce-market-is-now-49-or-5-of-all-retail-spend/>

North American and the international sectors are two examples of said geographically diverse sections. Amazon Web Services (AWS) also fits into this category. AWS offers a wide range of services, including cloud computing, storage, databases, and other types of data infrastructures; much of Amazon's competitive edge is derived from its services and operations.

By using cutting-edge and dependable solutions, Amazon was able to rise above its competitors and secure its high standing in the current market. Today's popular cloud-computing services have their roots in in-house implementations at a single organization. However, Amazon could commercialize these solutions because of their sophistication and robustness. As a result, the corporation benefits greatly from its successful operations procedures.

Additionally, Amazon is responsible for managing the storing and tracking processes of both inbound and outbound shipments from sellers that choose to ship with Amazon, as opposed to shipping independently, as well as delivering them to the client. Additionally, they maintain their own delivery networks for addresses located near their storage facilities.

Outbound Logistics

Outbound logistics are the activities that deliver Amazon services and products to its customers - this is where finished products are stored, moved, and distributed. Amazon's outbound logistics include a wide range of activities. Automated inventories integrated into Amazon's fulfillment centers and assist in managing, storing, picking out, and transporting products. Worldwide, the company has a total of 175 fulfillment centers.

Aside from utilizing fulfillment facilities, Amazon also operates other outbound logistical tasks, primarily co-sourcing and outsourcing. As with online products like books and digital media, the corporation uses digital distribution systems. Additionally, the corporation has many physical stores worldwide, which provides a significant amount of money for the business. Amazon uses DHL, FedEx, UPS, and other delivery providers globally. Not only that, the firm has been establishing and enhancing its own dedicated logistics system, which consists of even more planes, trains, ships, trucks, and vans. More recently, Amazon has made headlines for investing in, and developing, drone technology to further innovate its delivery processes.

Marketing and Sales

Amazon spends a lot on marketing globally. Marketing and sales operations in the value chain include the promotion and sales tactics (e.g. pricing, discounting) of products. Similarly, the corporation has continuously increased its capital inflow. In 2018, the corporation spent nearly \$13 billion on advertising. According to the market and consumer database company Statista, in 2019, Amazon spent almost \$18 billion on advertising, up to \$5 billion from 2018, almost a 37% increase in advertising spending.

The firm has always believed in marketing. As Amazon sells various products and services in multiple markets, it must constantly innovate itself and its image. The company's focus on marketing and sales have been a significant source of value as successful targeted advertisements can promote sales and drive interest for specific items or for the brand itself. Customers often are satisfied with Amazon's vast assortment of products, competitive prices, seamless shopping experiences, almost-instant delivery, and excellent customer service. Amazon

Prime is positioned as a higher standard of these services. The organization uses a variety of marketing mix tactics to reach its target market.

Services

TA value chain study of Amazon reveals that the organization's services post-sales are critical to gaining a competitive advantage. It offers premium services to customers and vendors, resulting in a smooth and stable business strategy. It provides comprehensive vendor support, providing training, literature, and facilities. Moreover, customers can benefit from Amazon Marketplace and Prime services. Amazon's after-sale service is excellent and is often the factor that maximizes its customer retention.

Ethical Issues Arising from Rampant Fake Reviews

Amazon's popularity and usage has risen at an exponential rate over the last few years. According to Statista, Amazon's brand value has surpassed 68 billion dollars with 200 million Prime members.³ Amazon has been exposed to numerous challenges as a result of its large user base, as new sellers struggle to generate enough traffic to their products. As a result, one of the most frequently encountered issues on Amazon is the use of phony positive reviews to increase sales of certain products. According to data gathered from Fakespot, a company that specializes in the identification of fake reviews and counterfeits for online retailers, the number of phony Amazon reviews is predicted to be between 40% and 45% of all global reviews as of June/July 2020.⁴ To demonstrate the presence of phony reviews on Amazon, we evaluated the amount of false verified accounts. The following graph illustrates the outcome. Appendix 1 contains the result demonstrating the existence of False Verify accounts.

Our Analyses

To have a better look at this issue, we examined Amazon's historical statistics for the years 2008 and 2018 (from June 1st to September 30th). We chose these years as they can give us a good understand of the decade-long differences/similarities between Amazon usage and reviews, and these months do not include any major holidays, which can often overinflate reviews and purchases due to spikes in sale deals and offerings. Additionally, we use comments from Amazon's "Electronics" category, as it is one of the most popular product categories available, which means it also has the highest number of reviews and review diversity. According to the Amazon website, to become a verified customer, an account must spend at least \$50 on Amazon in the preceding 12 months or purchase a product without receiving a significant discount; this will be important for our analysis, which will sometimes include splits between verified and unverified reviews.⁵

According to Appendix 1, half of the reviews in the Electronics department between June and September 2008 are False Negative. Buyers are experiencing actual difficulties as a result of fake reviews. Additionally, it leads the consumer to believe in the product's quality only to receive one that is defective or does not match the image. However, consumers are not the only ones that suffer when sellers fabricate reviews. As phony reviews erode consumer confidence in

³ <https://www.statista.com/statistics/546894/number-of-amazon-prime-paying-members/>

⁴ <https://www.techradar.com/news/how-to-spot-fake-amazon-reviews-during-prime-day-2021>

⁵ <https://www.aboutamazon.com/news/innovation-at-amazon/update-on-customer-reviews>

the ratings system, the platform as a whole suffers. Additionally, Amazon could have done nothing because increased false reviews result in increased sales on their site. However, if these incidents continue, they will eventually lose all credibility as customers lose trust in their platform. Amazon must address this issue if it wishes to remain in business.

Amazon has invested a significant amount of money on this issue in order to preserve their reputation over time. According to Amazon, they spent more than \$400 million in 2018 to ensure that customers have a positive shopping experience and that selling partners have a positive selling experience in the store.⁶ Additionally, they blocked over a million bot accounts prior to posting even a single listing. Additionally, it halted the publication of almost three billion potentially dangerous listings. Additionally, they thwarted almost 13 million attempts to post a bogus review.

A Harvard Business Review reports that Amazon spent more than \$500 million and employed more than 8,000 people in 2019 to reduce fraud and abuse on the marketplace.⁷ They're also discovered that Amazon was deleting about 40% of these phony reviews in their investigation.

We're also conducting our own test by comparing the quantity of the Unverified accounts in 2018's data to the amount in 2008's data, which is shown in our 2nd Appendix. As a result, we can see that the number of "False Verify" accounts has decreased dramatically since 2008, indicating that Amazon is serious about resolving this issue. As a result, all of the sources from Amazon, HBR, and our own analysis give the same result and corroborate one another.

Sentiment Analysis

To further investigate the bot account, we performed a sentiment analysis on all of the reviews in our 2008 and 2018 data sets. Appendix 3 (for 2008) and Appendix 4 include the findings of the investigation (for 2009). (As of the year 2018). In general, statistics show that all negative emotions were higher in 2008 than in 2018. However, negative attitude feelings were lower in 2008 than they are now. Positive feelings should have increased by a magnitude comparable to its size over the course of the year, while the quantity of good sentiments increases over the course of the year. This demonstrates that, despite Amazon's efforts to reduce the number of fraudulent reviews, some still fall through the cracks, resulting in these results.

As a result, we may assume that Amazon has attempted and is resolving these issues. However, as sellers get more creative in their use of false reviews, online marketplaces like Amazon find themselves in an endless arms race to develop effective methods for identifying and removing them while maintaining consumer trust.

Text Mining

⁶ <https://www.aboutamazon.com/news/company-news/ensuring-customers-receive-the-products-they-expect>

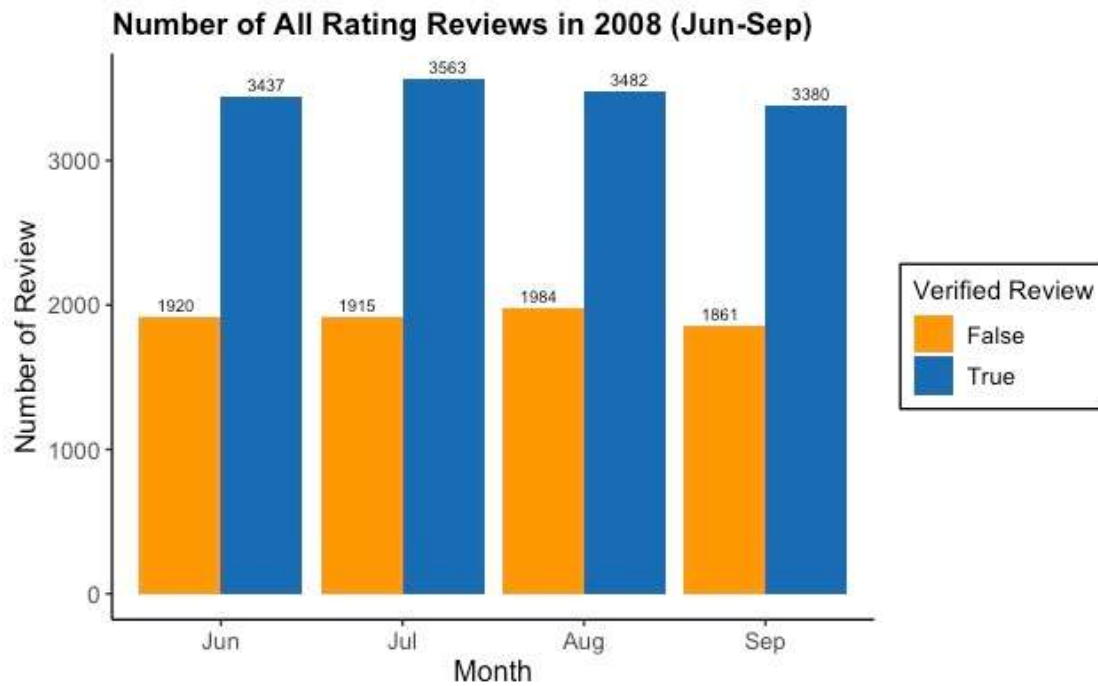
⁷ <https://hbr.org/2020/11/how-fake-customer-reviews-do-and-dont-work>

Text Mining is another data analysis method that we used. Natural Language Processing allows us to have a more comprehensive overview of what terms are popular when it comes to reviews for electronic products on Amazon – which allows us to understand what consumers often think about and refer to when assessing these products. We compared the popularity of words in five-star reviews for both unverified and verified reviews for the years 2008 and 2018. We can actually see some very interesting similarities between the two years, as well as the two review types. For instance, the word “easy” is, by far, the most often used word in our analysis, followed by “excellent” and “recommend”. This is valuable information to note as we can almost exclude this from the list of notable words when trying to decide what might be a fake review.

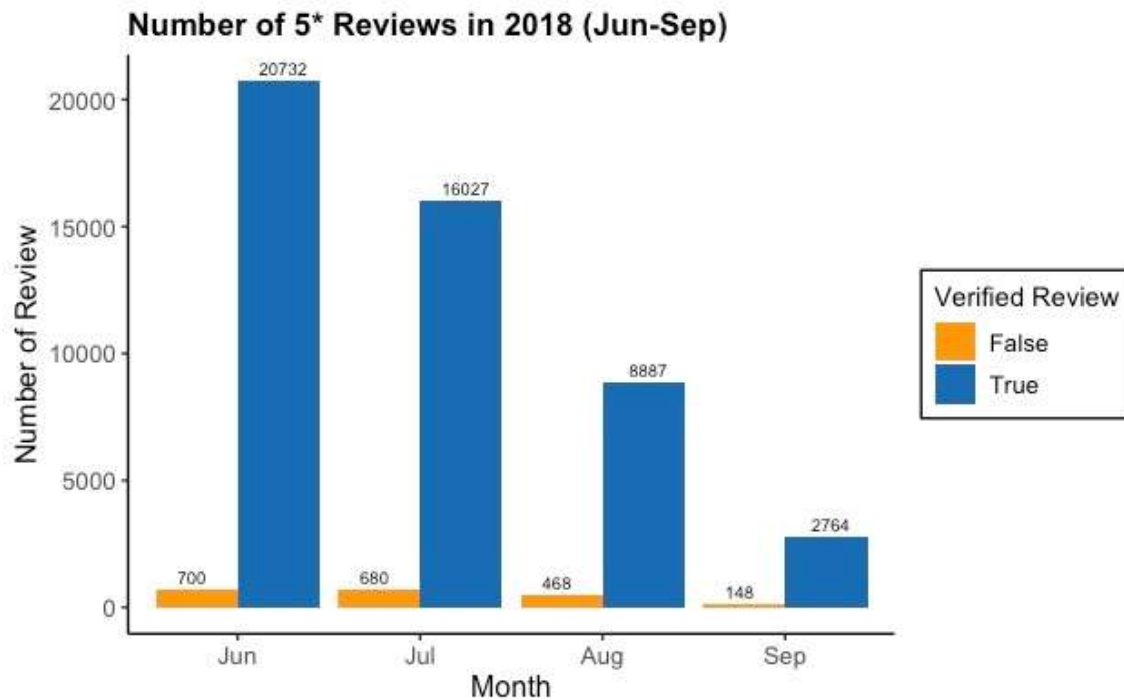
Managerial Conclusion

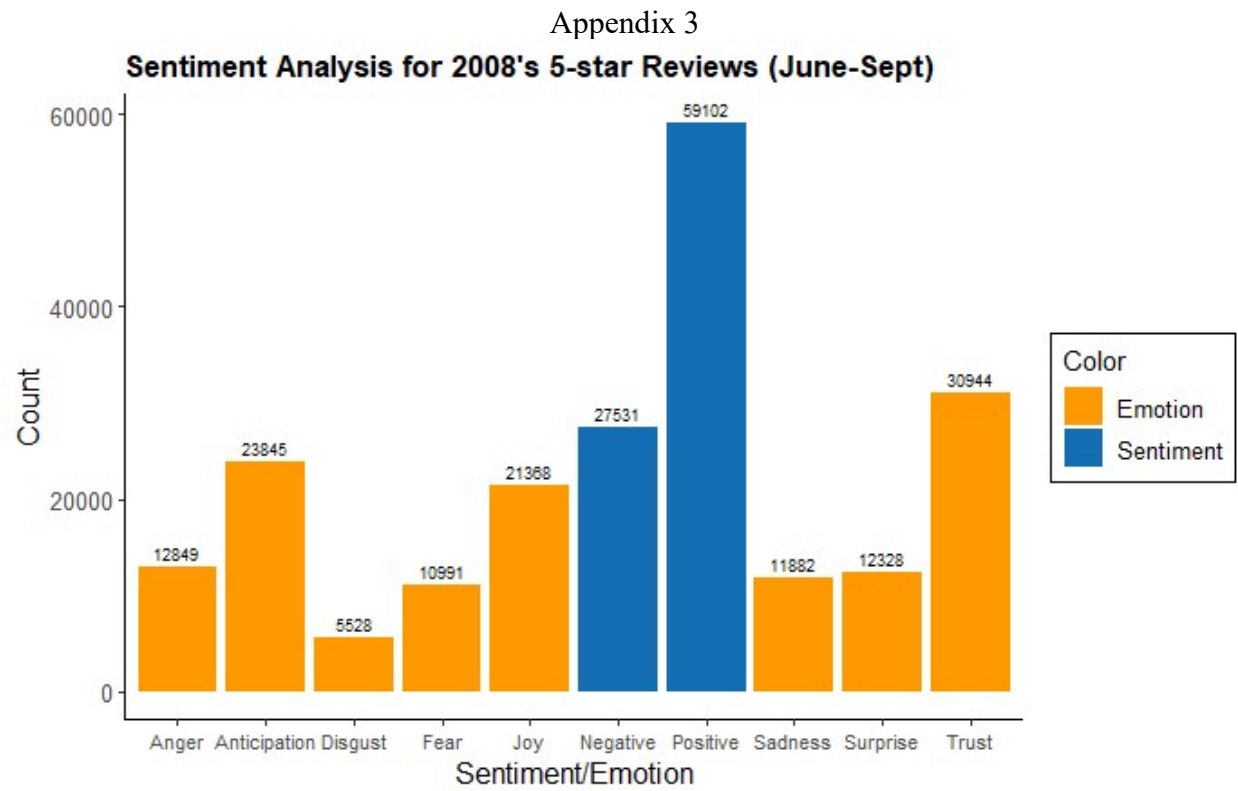
For a massive e-commerce platform like Amazon, fake/bot reviews are an inherent concern. Although they have invested the money and made every effort to remove it, people continue to devise new ways to circumvent it. From our analyses, it is actually much more difficult to identify fake reviews from authentic ones, especially given the new methods that phony sellers are using to circumvent Amazon’s automated flagging systems. Our recommendations for Amazon's future actions, beyond what Amazon is already implementing in its own business, is to include the development of bot chats that are capable of identifying review trade conversations between sellers and buyers on Amazon web chat, more comprehensive purchase verification standards, and increased account privacy, security, and protection.

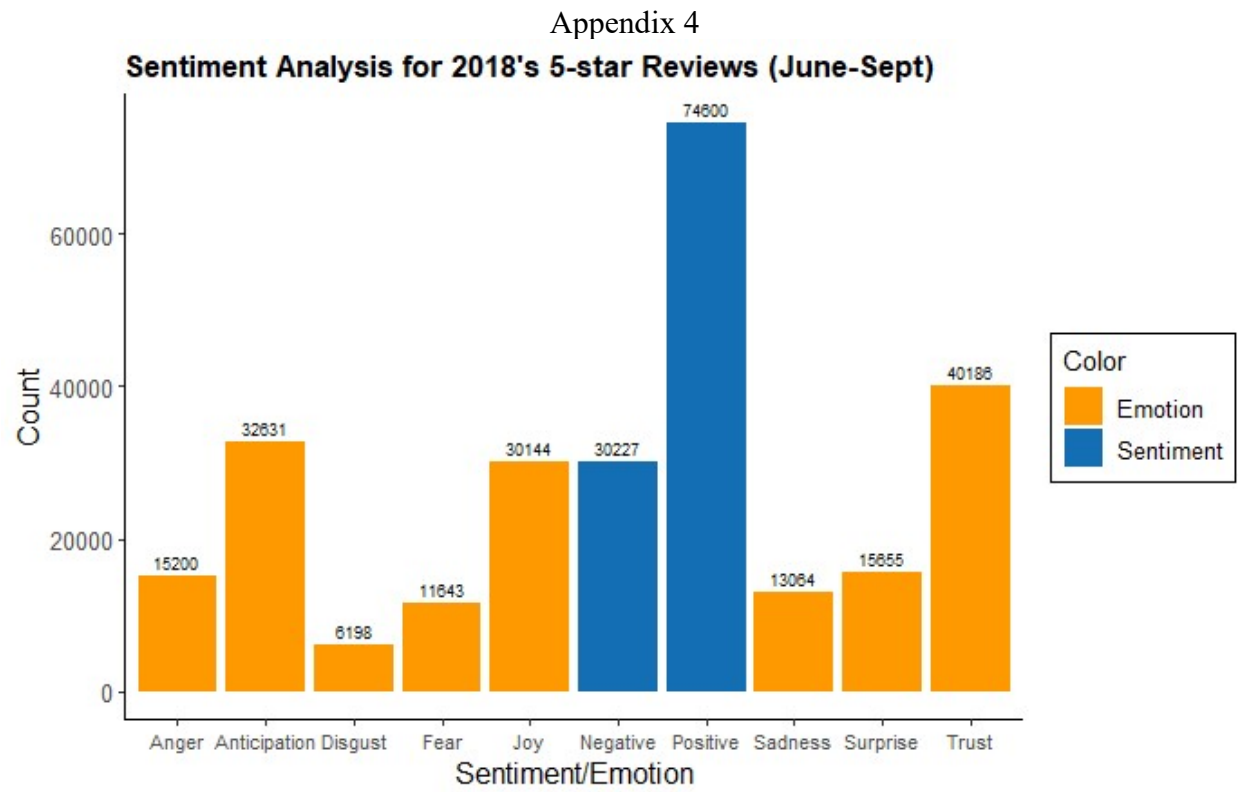
Appendix 1



Appendix 2







Appendix 5

2008



2018



verified



unverified

References

- “Ensuring Customers Receive the Products They Expect.” US About Amazon, Wwww.aboutamazon.com, 11 November. 2019, <https://www.aboutamazon.com/news/company-news/ensuring-customers-receive-the-products-they-expect>.*
- “Update On Customer Reviews.” US About Amazon, Wwww.aboutamazon.com, 3 October. 2016, <https://www.aboutamazon.com/news/innovation-at-amazon/update-on-customer-reviews>.*
- Watson, Josephine, and TechRadar. “How To Spot Fake Amazon Reviews During Prime Day 2021 | TechRadar.” *TechRadar*, Wwww.techradar.com, 20 June. 2021, <https://www.techradar.com/news/how-to-spot-fake-amazon-reviews-during-prime-day-2021>.
- “U.S. Amazon Prime Subscribers 2019 | Statista.” Statista, Wwww.statista.com, <https://www.statista.com/statistics/546894/number-of-amazon-prime-paying-members/>.*
- “Amazon’s Share Of the US E-commerce Market Is Now 49%, Or 5% Of All Retail Spend – TechCrunch.” TechCrunch, Techcrunch.com, <https://techcrunch.com/2018/07/13/amazons-share-of-the-us-e-commerce-market-is-now-49-or-5-of-all-retail-spend/>.*
- “How Many Products Does Amazon Carry? - Retail TouchPoints.” Retail TouchPoints, Wwww.retailtouchpoints.com, 8 January. 2020, <https://www.retailtouchpoints.com/resources/how-many-products-does-amazon-carry>.*