

MGMT 190: Data and Programming for Analytics

Group 4: Anna Zhang, Ying Jin, Sang Do, Anuj Shah, Milisha Merchant, Egemen Can Gok

Project Plan

Topic: Analyze Online Retailers (Amazon vs. eBay vs. Alibaba)

Every year, more than 100 million Americans purchase goods from the online retail marketplace, which has become one of the fastest-growing sales channels in the United States. With the online retail industry on a rise, Amazon, eBay, and Alibaba are three strong competitors each with different user pools and as a result different product reviews.

Amazon is primarily a retailer meaning that it exists to sell products to a wide customer pool at reasonable prices. It targets the middle and upper class individuals who value convenience when it comes to online shopping. On the other hand, eBay primarily plays the role of a wholesaler and sells merchandise at the lowest price possible mostly to retailers who will most likely resell the goods to consumers. Lastly, Alibaba serves as a B2B website as it tries to serve the needs of other businesses and its employees. It also offers a platform for individuals seeking fast shipment and small quantity orders. Unlike Amazon and eBay, Alibaba is primarily based and founded in China.

Although all three companies are essentially online retailers, they serve different purposes and as a result attract different consumers. Through this project, we hope to gain a better understanding of the difference in consumers between each of the platforms and essentially why reviews for similar products vary so greatly.

In this project, our group plans on analyzing the differences in users and products sold on Amazon, eBay, and Alibaba. It is typical that one product receives drastically different ratings on different sites. However, if there are a large group of people reviewing the products and if the sample represents the population, the ratings should be approximately the same. Therefore, we are interested in discovering the reason why the ratings are so different across

the various websites. We plan to do this by discovering popular products that are sold on all three sites, acquiring the profiles of the users who gave reviews to those products, and concluding on a trend about the sites' user pools.

Here are some of the questions we hope to answer after completing this project:

- What are the most popular products sold on Amazon, eBay, and Alibaba?
- What are the profiles of the people who purchased and reviewed these products?
- What do the ratings tell us about the product(s) and the websites?
- Why do the ratings differ across the three websites?
- What do the user profiles tell us about the product reviews?

We plan to collect user profile data by web scraping the product review page and finding links to each user's profile. By scraping the profile page of each user, we can collect data about their region, occupation, age, and ratings.

We are going to present the user profiles via pie charts, bar graphs, and scatterplots.

Additionally, we plan on calculating the average ratings of each occupation, region, and age group in order to find relationships between ratings and user groups.

If possible, we will use these APIs to gather additional data:

- [Amazon](#)
- [eBay](#)
- [Alibaba](#)

In the end, we hope that we will be able to gain valuable information about the consumers, which will help strengthen our analysis and knowledge of the three platforms.

Project Plan Guidelines:

- Topic — What topic did the team choose to focus on? This should include a brief overview of the background of the topic.
- Justification — Brief description of why you are proposing this topic, what questions you attempt to answer from the data, and how you plan to collect and analyze the data.