

eBay Reputation Extractor and Analysis Tools

A Major Qualifying Project Report:
submitted to the faculty
of the
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfillment of the requirements for the
Degree of Bachelor of Science

by

Joseph Beynon

Marie Russo

Date: March 1, 2007

Approved:

Professor David Finkel, Major Advisor

Abstract

For this project, we developed statistics and analysis tools associated with reputation on eBay to make feedback more useful for buyers and sellers. We created feedback scores base on the volume of each transaction. We then created reputation visulizations, namely volume sparklines and logarithmic stars. We also developed applications to compare proxy bids with the prices the items sold for in order to study causes of negative transactions. Finally, we developed a new system for sampling listings, weighted by volume.

Executive Summary

An important part of the success of eBay is the trust that is developed between users. Without trust, users would not be as likely to make transactions with others. The current reputation system in place on eBay allows users to rate others with either a positive, negative, or neutral comment. There are many types of fraud that are not prevented by this reputation system. Making the reputation system more transparent to users will allow them to avoid fraudulent users without decreasing the number of positive transactions or increasing costs to eBay.

We were asked to create and implement new methods for calculating feedback and reputation. These methods were based on the end price of items as opposed to the current system which has no monetary basis at all. This project stands as a proof-of-concept for the idea of basing reputation off of volume. In a related area, graphical representations of reputation, both volume-based and not, were implemented. We also developed two new tools for use in the analysis of negative transactions and other applications.

In volume reputation, a positive comment is worth the selling price of the item instead of one point. Negative comments subtract the price of the item from the user's score. This system helps to prevent certain types of fraud. The visualizations we developed were volume sparklines and logarithmic stars. Currently a user's score is given using a number and a percentage. Sparklines show a recent list of feedbacks and their volume to give others a brief history of the user. Logarithmic stars use a new method of scaling to give each user a rating from zero to five and then display the result as a number of stars.

Overall our volume feedback reputation application meets all of the requirements set by our mentor and fits the design we created originally. There is some room for improvement within the application, mostly in reference to the API. If the API calls were replaced with internal data calls, the program would run much quicker. As for the data collected, studies should be conducted to determine the usefulness of volume-based feedback systems.

Reputation is not presented in a way that makes it intuitive. Sparklines and logarithmic stars attempt to solve this problem. Sparklines allows advanced users to see large amounts of past data in a small space and without having to jump between many different pages on the site. Adding volume to normal sparklines allows them to show much more information in the same amount of space as they use without it. We recommend that more tests be done using sparklines, including user tests. Also, a good scale for the maximum sparkline value should be determined. One possibility is to have a varying scale based on the category of the listing.

Logarithmic stars can be used for comparing many users at once. They separate users into distinct level that allow for buyers to easily eliminate sellers they are not comfortable purchasing from. Using volume to determine the level of stars for a user helps to make users with many small positive transactions and a few large negative transactions have lower ratings. Future work on logarithmic stars would be to determine whether or not half star increments are necessary and also to develop new graphics for the star. Perhaps always showing 5 stars but having some grayed out or hollow when a user has fewer would make the presentation look cleaner. Like sparklines. User tests to see whether or not users like the stars for rating would be useful.

The other tools that we developed were a proxy bid vs. final value analysis tool and a method of sampling listings by volume. Proxy bid vs. final value compares the amount that a user was willing to pay for an item with the price that the auction ended at. This difference might be a linked to negative transactions. Sampling by volume is a method of creating a sample set of listings using the price of each listing to weight its probability of being selected.

There is still some room for improvement in the flow and speed of our proxy bid vs. final value program. The program would run a lot faster with internal data access and the ability to interface multiple data sources together without user help. If this concept is to be used frequently, it might be worth making these changes. As for the data returned by the limited test sampling that was done, studies should be conducted to determine the usefulness of proxy bid vs. final value comparisons.

Sampling by volume is an interesting concept but needs to be worked on further. Currency differences are an inherent problem with this concept. Also, using regular

volume may not be the best method to sample by for eBay. Since eBay's profits are not based directly on volume, using the volume that is made by eBay for each listing might be a better property to use as a weight. Sampling by volume needs to be tested in different situations to determine whether the difference in sampling has a significant impact on the measurements taken from the data. We hope that it will be very useful for other projects.

Table of Contents

ABSTRACT	II
EXECUTIVE SUMMARY	III
TABLE OF CONTENTS	VI
LIST OF FIGURES & TABLES	VIII
1 INTRODUCTION	1
2 BACKGROUND	2
2.1 About eBay.....	2
2.2 eBay’s Feedback System.....	3
2.3 Other Feedback Systems	4
2.4 Studies on eBay’s Feedback	5
2.5 Reputation Systems.....	7
2.6 eBay Data Access.....	8
2.6.1 eBay API	8
2.6.2 Sojourner	9
2.6.3 Voyager	10
2.7 The Long Tail	10
2.8 Feedback Abuse.....	11
2.9 Uses of Feedback	13
2.10 Identity	14
2.11 Shannon’s Information Theory.....	17
3 DESIGN & IMPLEMENTATION	17
3.1 Maximum Proxy Bid vs. Final Value.....	18
3.2 Volume Reputation	23
3.3 Reputation Images	26
3.3.1 Volume Sparklines	27

3.3.2	Logarithmic Stars	28
3.4	Sampling by Volume	29
4	RESULTS & ANALYSIS	30
4.1	Proxy Bid vs. Final Value	30
4.2	Volume Reputation	31
4.3	Reputation Images	31
4.3.1	Volume Sparklines	32
4.3.2	Logarithmic Stars	33
4.4	Sampling by Volume	33
5	RECOMMENDATIONS & CONCLUSION.....	34
	APPENDIX A: SAMPLE SOJOURNER OUTPUT FILE.....	38
	APPENDIX B: SAMPLE PROXY BID VS. FINAL VALUE OUTPUT	39
	APPENDIX C: SAMPLE VOLUME REPUTATION OUTPUT.....	40
	APPENDIX D: SAMPLE LONG TAIL IMAGE.....	42
	APPENDIX E: SAMPLE GREASEMONKEY OUTPUT.....	43
	Get Results Page	43
	Compare Items Page	44
	View Item Page	44

List of Figures & Tables

Figure 2.1 – Graph of the Long Tail	11
Figure 3.1 – Volume Sparkline	27
Figure 4.1 – Volume Sparkline of a Padder.....	32
Figure 4.2 – Volume Sparkline of a Misleading Advertiser.....	32
Table 3.1 – Table of Bid Increments	19
Table 3.2 – Percent Positive Feedback and Corresponding Star Rating.....	29

1 Introduction

An important part of the success of eBay is the trust that is developed between users by the reputation system. Without trust, users would not be as likely to make transactions with others. The current reputation system in place on eBay allows users to rate others with either a positive, negative, or neutral comment. There are many types of fraud that are not prevented by this reputation system. Making the reputation system more transparent to users will allow them to avoid fraudulent users without decreasing the number of positive transactions or increasing costs to eBay.

The inspiration for the majority of this project arose from our mentor, Mark Boyd, and his patent for take volume reputation [1]. Mark's idea is to create a reputation system that is weighted by how much money eBay itself makes. As a possible precursor to that system, we were asked to create and implement new methods for calculating feedback and reputation. These methods were based on the end price of items as opposed to the earnings eBay made off of that transaction or the current system that has no monetary basis at all. This project was a logical step because the data on item value is exponentially easier to gain access to than the revenue information. This project also stands as a proof-of-concept for the idea of basing reputation off of volume of any kind. In a related area, graphical representations of reputation, both volume-based and not, were implemented. We also developed two new tools for use in the analysis of negative transactions and other applications.

The new method for determining a user's score is by using volume reputation. In this system, a positive comment is worth the selling price of the item instead of one point. Negative comments subtract the price of the item from the user's score. This system helps to prevent certain types of fraud. The visualizations we developed were volume sparklines and logarithmic stars. Currently a user's score is given using a number and a percentage. Sparklines show a recent list of feedbacks and their volume to give others a brief history of the user. Logarithmic stars use a new method of scaling to give each user a rating from zero to five and then display the result as a number of stars.

The other tools that we developed were a proxy bid vs. final value analysis tool and a method of sampling listings by volume. Proxy bid vs. final value compares the

amount that a user was willing to pay for an item with the price that the auction ended at. This difference might be linked to negative transactions. Sampling by volume is a method of creating a sample set of listings using the price of each listing to weight its probability of being selected.

2 Background

Prior to arriving at eBay, we conducted research on some general topics pertaining to this project. Subjects covered include background on eBay as a company as well as a description of eBay's feedback system and comparisons to other feedback systems. A brisk analysis of eBay's feedback system referencing studies on the subject was conducted as well as research into other types of reputation systems.

While working at eBay, we had many conversations with Mark J. Boyd about topics pertinent to eBay and eBay's feedback system. The long tail distribution of the products sold on eBay was discussed. Most of the topics in question had a more direct correlation to feedback. These included feedback abuse, uses of feedback, identity, and Shannon's Law of Self-Information. All of these subjects are detailed below.

2.1 About eBay

Today eBay is known as the king of online commerce with three times as much profit as Amazon.com yet only half the revenue. eBay has reported a 70% increase in profits every year and a 21-fold increase in share price since going public [2]. Despite increase in costs, eBay is holding steady in the current market.

eBay was founded in 1995 by Pierre Omidyar, who ran it out of his living room. The original name of the company was Echo Bay Technology; however that name was already taken. Omidyar shortened the name to eBay and the successful online auction house known today was born. eBay went public in 1998 competing against Amazon.com and Yahoo. eBay bought Half.com in 2000 and subsequently integrated into eBay by 2001 [3]. In 2002 PayPal was bought by eBay, replacing their old payment method, Billpoint [4].

eBay allows people to buy and sell a large variety of items across the globe. eBay stores are used by more extensive sellers to organize their business [5]. There are many

subdivisions of eBay that serve different purposes. Half.com sells items in a similar way to eBay, but in less of the auction style which allows for lower costs per transaction [6]. Half.com allows sellers to post items at a set price for other users to buy. Rent.com allows users to search and rent properties around the world [7]. eBay Motors is a forum for selling cars, motorcycles, boats, and parts for all of those items [8]. eBay has sold over two million vehicles since its creation [9]. My eBay allows users to track the items they are buying, selling, or interested in [10]. My eBay also aids the user in keeping their account information current.

PayPal is an eBay service available to anyone with an email address [11]. PayPal allows users to make financial transactions using their bank account or credit card. Transactions are made with others, without using their real financial information. PayPal offers financial protection as well. There are currently over one hundred million PayPal accounts across the world. PayPal was founded in 1998 and acquired by eBay in 2002 [4].

From the beginning eBay's biggest concern was the customer, mainly the buyer. Currently eBay offers many resources to both the buyer and the seller. These resources vary from simple online tutorials on how to buy and sell items to information on what to do in the case of a fraudulent transaction. eBay supplies documents on how to prevent fraud [12] as well as their policies concerning fraud and many other topics [13]. Real-life courses are offered on how to buy and sell on eBay [14]. eBay's glossary [15] and acronym dictionary [16] help new users understand some of the lingo used by eBay and their user. Details of eBay's financial status are available [17] as well as information on eBay Foundation, eBay's charitable fund that has donated over eight million dollars to non-profit organizations [18]. Available careers within eBay are listed directly on eBay's website [19]. All of these resources and more are available directly at eBay's website.

2.2 eBay's Feedback System

The current feedback system used by eBay was created in February 1996 by eBay founder Pierre Omidyar. The Feedback Forum is an open forum for both complaint and praise of members that others have dealt with. When someone purchases or sells an item they can leave feedback for the member they traded with [20].

The three ratings that can be left are positive, neutral, and negative. These ratings are then translated into a score for each member. A member's feedback score is equal to the number of positive ratings they've been given minus the number of negative ratings they've been given [21]. In addition to a number rating, members receive stars based on their overall feedback score. The feedback stars range from a plain yellow star for a feedback score between 10 and 49 inclusive, to a red shooting star for a feedback score of 100,000 or greater [22].

If a member leaves multiple positive ratings or multiple negative ratings for a given member only one of them is counted towards the rated member's feedback score. If a member leaves feedback with different ratings for the same user the rated member's feedback is affected by at most one point, either positive or negative, dependent on the sum of the ratings given by the user [23].

Each feedback rating is attached to a comment about the transaction with the member. eBay suggests not relying solely on feedback score, but also on the comments left for someone when deciding whether or not to deal with a member. Comments do not affect a user's feedback score, but can affect others' decisions about making bids [21].

eBay has a strict feedback removal policy. Feedback cannot be removed or edited normally. Feedback can be mutually withdrawn if a disagreement has been resolved. A request for removal must be submitted soon after the feedback was left, or the transaction ended [24]. Other than a mutual request, feedback will only be removed if the system is abused [25].

2.3 Other Feedback Systems

Many other online commerce sites use feedback systems to rate items and sellers. These sites share some aspects of the systems in common with each other and eBay but also have many differences. Some sites that use feedback systems are Amazon.com, Buy.com, and Google Checkout.

Amazon.com has a feedback system for their merchant shopping section that is similar to eBay's feedback system. Feedback includes a rating and a comment [26]. Both buyers and sellers may leave feedback for each other. Seller's feedback and scores are shown on their site. Buyer's feedback and scores are shown on their profile page.

A seller's feedback score is calculated based on only sellers' ratings and use a five star rating system. One or two stars are considered negative feedback while four or five stars are considered positive feedback. Three stars are considered neutral feedback. These scores are used to calculate the feedback scores which include average star rating and the percentage of positive, neutral, and negative feedbacks.

On Buy.com, items can be purchased through other sellers in the Marketplace. Marketplace buyers can leave feedback for the sellers using a star rating. A seller's rating is the average of the number of stars they've been rated. Ratings of one or two stars are considered negative, three stars is neutral, and four or five stars are positive. The average rating is shown to the nearest whole number of stars but is also offered as a numeric value. With each rating, buyers can leave a comment about their transaction as well. All of a seller's previous ratings are available to potential buyers. Ratings cannot be altered in the Buy.com system [27].

Google Checkout is another service with a seller feedback system. Checkout's reviews are a star rating combined with comments. This is similar to eBay's system but varies in the calculation of a seller's rating. The rating is calculated as the average number of stars awarded to a seller in all ratings left for them. The ratings go from one to five stars, and there is no way to leave "negative" ratings. Unlike eBay's current feedback system, there is no way for sellers to rate buyers [28]. They are able to comment on reviews left for them though. After a comment has been left on a review, the reviewer has a chance to make a follow-up comment [29].

The reviews on Google Checkout are only visible by the buyer who wrote it and the seller it was left for. According to Google's help site, there are future plans for reviews to be visible by anyone. Although shoppers cannot see the reviews, they are still able to see the average rating of merchants that they might want to buy from [30].

2.4 Studies on eBay's Feedback

The purpose of eBay's feedback system is to give the users information about the people they are making transactions with. This information is meant to give the user a sense of trust and respect, and promote honest and fair transactions [12]. This information affects with whom users are willing to make transactions, but also can affect the amount

of money a buyer is willing to spend on a seller's item [31]. Different representations of this information will yield different effects.

Several studies have been performed that show that buyers are willing to spend more for an item from a seller with a positive reputation. Other studies show that buyers are willing to spend more for an item from a seller that has had more experience shown by more overall feedbacks. It has been shown that long-term users are less likely to be dishonest in a transaction [32]. Yet other studies show that a newer seller with a few negative feedbacks will be treated the same as a newer seller with no negative feedbacks. If a user values an item enough, they may disregard the feedback scores simply to obtain the item. These are a few of the ways the feedback score system can affect prices of eBay items and the profit of the sellers. Ba and Pavlou [32] detail other effects of feedback scores on prices of items in their studies.

The representation of a feedback score is vague in some situations [31]. If a user has 195 positive feedbacks from discrete users and no negative feedbacks, their feedback score would be 195. However, a user with 223 positive feedbacks and 28 negative feedbacks, their feedback score would also be 195. While the feedback score shows similar caliber in the two users, the value of the individual feedbacks is more important. In this case, the percent of positive feedbacks is a better indicator of the caliber of the user and how seasoned he or she is.

Currently, the only criteria by which the feedback scores are broken down is time. eBay displays the number of positive, negative, and neutral feedbacks over the time periods of one month, six months, twelve months, and the total. There are several other possibilities for the manner in which to show this information.

Buyer and seller feedbacks are not separated except in reference to comments, not ratings. This information could be useful for buyers. Imagine a user with a very good feedback score, but most of the ratings that make up that score are from buying items. This user may not be a good choice to buy from as opposed to someone who has more selling experience.

Another concept that is not represented is a user's knowledge of different categories of items. A buyer might not wish to purchase an item from a seller who has never sold items of that category before. An example would be a seller who specializes in

collectable candy dispensers posting a sale for a high-powered sports car. In the aforementioned example the price of the items may also be valuable information.

A different circumstance could involve the caliber of the user leaving the feedback. The opinion of a user who has more negative feedbacks than positive may be less beneficial than that of a user with an average feedback score. Similarly, the opinion of someone who leaves more negative feedback for others than good might not be very valuable. These are several aspects of feedback that could be useful to users in choosing where they buy their items and how much they are willing to pay for said item.

2.5 Reputation Systems

A reputation system is an important part of any electronic commerce system. In the digital world, people do not know the reputations of sellers like they would in their own community. It is important for buyers to have a level of trust in a seller before buying items from them. Reputation systems allow buyers to determine a level of trust that a seller has before engaging in a transaction with them. According to Resnick, Zeckhauser, Friedman, and Kuwabara [33], “a reputation system collects, distributes, and aggregates feedback about participants’ past behavior.” The system helps people know who they can trust to make deals with, and who they should avoid. Reputation systems also “deter participation by those who are unskilled or dishonest.” [33]

Reputation systems come in many varieties. There are different methods of rating users, different methods of calculating rating scores, and there are other design factors of reputation systems that can be changed. One thing that needs to be handled in reputation systems is the problem of unfair ratings. High ratings may be given unfairly to increase a seller’s reputation. On the other hand, unfairly low ratings may be given to drive a seller out of the market [34]. Currently there are many proposed reputation systems that handle unfair ratings. These systems use a range of methods from collaborative filtering to probabilistic modeling and graph theory [34, 35]. Some of these systems handle different situations such as distributed or centralized systems, handling of unfairly low, high, or both ratings, etc.

TrustDavis [35], is a reputation system developed by the University of California, Davis which gives incentives for accurate rating and resists exploitation of various

means. This is achieved through references. TrustDavis' reputation system is a graph with weighted, directed edges that represent references to other users. The weight of the edge is the amount that a user is willing to be liable for if the person they refer is bad. With this system, there is no risk in dealing with another person if you can find connections to the dealer with enough liability to make up for bad trades. In order for someone to get a reference they would generally buy a reference from someone they trust. This reference then has a weight based on how much they paid so that the referrer has the money to accept limited liability for the other user's actions.

Another problem with some reputation systems is the style in which buyers are asked to rate sellers. [36] is a study of how information available to the rater at the time of rating affects their decision. The study found that users rate consistently with different rating scales. They also found that if the current rating is shown while the user is making their own decision, their choice can be manipulated [36]. This study performed its experiments using a recommendation system. Many of the concepts can still be applied to reputation, especially interface considerations. Altered ratings can have a major effect on the performance and quality of the system [36].

2.6 eBay Data Access

In order to study eBay's feedback system and develop tools to work with it, we needed to be able to access data about users and item listings from eBay. For this project we used three methods of accessing the data that we needed. We were given access to one public and two private eBay data retrieval systems. The public method we used was the eBay API for developers. The private systems were Sojourner, a user click-tracking system, and Voyager, the eBay search engine.

2.6.1 eBay API

The eBay Developers Program [39] was started in the year 2000 and provides resources for developers of 3rd party applications on eBay. The Developers Program provides documentation for the XML and SOAP APIs as well as software development kits for various programming languages. The eBay Developers Program also provides a Solutions Directory [37] on which developers can list their applications using the various

APIs provided. Other users can also use the Solutions Directory to find applications that fit their particular needs. Everything from complete selling solutions to category-specific applications can be found in the directory. An example of an application that uses the eBay API is Configurator [38]. Configurator is an application that lets users customize items that they want to purchase and have it automatically added to the seller's eBay store. This could be used to customize a computer and have it automatically set up as a Buy It Now listing through eBay for the customer to purchase it.

The eBay API is a web service which allows developers to create applications using the information available through eBay's website in a more computer-readable manner. The requests and responses are made up of XML code and can be easily translated into data structures for many languages. Libraries are available publicly from eBay's developer website for Java, Microsoft .NET, and PHP [39]. The eBay API allows for users to rapidly expand their eBay businesses by automating repetitive tasks. Many programs have been written using the API, creating a full workflow for managing eBay businesses. Currently, 23% of eBay listings are created through 3rd party applications [40].

2.6.2 Sojourner

Although the eBay is available from outside of eBay, Sojourner is only available to eBay employees. Sojourner is a system which tracks individual users' clicks on the eBay website and provides an easy method to search for information about user sessions. The Sojourner database is accessible through the eBay Research Labs and records the sessions of 25% of the site's users. Third Eye Seer is a Web interface for Sojourner that allows its users to create C++ graphs which are run against the data stored in the Sojourner database [41]. A graph is a query which processes and outputs data from the Sojourner system in a tree format. Graphs are used to filter data by which site the data was inputted and what types of data to output. This provides a powerful way to interface with the massive amounts of data available in the system. Sojourner provides access to data that is not available through either of the other methods such as proxy bids.

When a user accesses a page on eBay, a message is sent containing all of the information submitted by the user plus extra data generated by the Web servers. This

information is recorded in Sojourner. The first time a browser accesses an eBay page, a GUID (Globally Unique Identifier) is generated for it and saved in the browser's cookies. This is a very important piece of information in the Sojourner system. If the cookie is not deleted, then any page on eBay that the browser goes to is recorded in Sojourner with the same GUID and sessions can be tracked for as long as the cookie remains. The GUID is not tied to a specific user or IP address but only the browser used to view the Webpage [42].

2.6.3 Voyager

Unlike Sojourner, Voyager is partially available outside of eBay but most of the functionality is only available to eBay employees. Voyager is the current eBay search platform. Voyager is comprised of many nodes which are accessed through aggregators and load balancers. When a search is performed through the eBay website, Voyager is the system that handles the request and finds the search results. Voyager is a very powerful system handling millions of searches each day. Voyager is accessible to all users of the eBay website in a limited capacity, but can be used to a much greater extent from inside of eBay.

When users perform a search through the site, they are only able to access some of the data that is available through Voyager. Querying Voyager nodes directly gives access to more information than is accessible through the eBay website and provides it at greater speeds, in a tab-delimited format. The data returned by a query on the nodes can be changed to suit the needs of the user. Accessing the search nodes of Voyager directly allows for much quicker access to data than using the eBay API. The information provided by Voyager allows for the creation of fast internal applications.

2.7 *The Long Tail*

A very distinctive set of items are sold on eBay; ones that are impossible to find anywhere other than your neighborhood yard sale. These items are generally one-of-a-kind, difficult to find, or generally unique. If one were to plot the volume per item against the range of every item available including brick-and-mortar stores, hybrid stores, online stores, yard sales, and more unique sites like eBay, one would see an inverse exponential

relationship as shown in Figure 2.1. The head, or body, consists of the mainstream retail items such as popular console systems, DVDs of recent box office movies, and the like. The tail consists of the scarcer or singular items that can include antiques, original artwork or crafts, and other less popular items. When analyzing the monetary volume of each of these two sections, there is a distinct, but unapparent difference. Focusing on the height of the head items, there is a cap to how much most people will pay for any given item. This limits the amount of money possible to make on these types of items. However, in the tail section, there are an infinite number of unique items that people can buy and sell. Tapping this area of the marketplace leaves no limit on the amount of money available to be made since there is no limit on the number of items to be sold. The major goal of eBay is to harness the income of the long tail while allowing users to buy and sell items that would not be available anywhere else.



Figure 2.1 – Graph of the Long Tail

2.8 Feedback Abuse

eBay's feedback system, like most other feedback systems, is not perfect and can allow a user's reputation to be presented in ways that are not truthful. This can result from purposeful action, lack of awareness, or it could be completely accidental. There are several different types of actual feedback abuse. A few of these include shilling, padding, extortion, and general anticompetitive behavior. Feedback abuse not only affects the user committing the abuse, but can mar other users' reputations. In general, it is a good idea to

minimize any and all of these issues to maintain a functional feedback system. Each type of feedback abuse has different ways to prevent it.

Shilling occurs when transactions are made between two accounts that are either owned by the same person or two people that are conspiring with each other. This can be done to achieve various outcomes. In relation to the item being sold, shilling can raise its demand or the price at which an unknowing person would have to pay to obtain that item. Shilling can also be used to inflate feedback scores. This would be done by appearing to sell items between two accounts, but never exchanging money or merchandise. This would allow the accounts to leave positive feedback for each other with no cost. In an attempt to thwart this, eBay uses unique identifiers that prevent people from having more than one account. This prevents self to self shilling, but doesn't succeed in deterring cooperative shilling.

A user may partake in padding, which consists of participating in many small transactions to boost their feedback score. A single feedback has the same affect on a user's feedback score, regardless of whether that feedback is from a buyer or a seller and whether the item sold was expensive or lesser in value. Padding can be conducted without the other party knowing because the transaction will seem, and is, completely legitimate. This could allow a user to purchase many small items then attempt to sell a much larger item, possibly with fraudulent intentions. An unknowing buyer may see the higher feedback score and think the seller is genuine in their attempt to sell an item. As an example of fraudulent behavior, the seller may then take the buyer's money and never deliver the item. Currently, eBay has no policies directly against padding.

Using the feedback system as leverage, users can extort other users into doing deeds they would not do otherwise. These actions can include lowering the price of an item, cancelling a shipping fee, including an extra item at no charge, amongst many others. A user could purchase an item and then threaten the seller with a negative feedback if they do not complete the task stated by the buyer. eBay has strict policies against extortion, but no action can be taken if the situation is not reported to eBay.

Punitive feedback abuse takes place when users employ the feedback system to attack someone whom they don't like or feel has wronged them. This tends to happen mostly when the targeted user sells items on a regular basis. The assailant or assailants

would buy several items from the victim, generally low in price, and leave negative feedbacks. The aggressor could involve other eBay users they know by asking them to do the same. This type of feedback abuse is also hard to spot because there are legitimate transactions taking place, but illegitimate feedback being left for those transactions. A user who has fallen victim to punitive feedback abuse can report the offending users to eBay and petition to have the feedback removed. The offending users may be forced to leave by eBay as well.

These are some of the behaviors that negatively affect eBay and its feedback system. While there are other types anti-competitive behavior, the ones described above are the most prominent and most detrimental to eBay's community. Some of the concepts researched and implemented in this project help to deter several of these types of behaviors.

2.9 Uses of Feedback

The feedback system and resulting user reputation is valuable to several different groups of people, each for different reasons. eBay users can include this information in their decision of with whom to do business. eBay employs this data internally for customer service and program eligibility. In each situation, users with different reputations receive different treatment.

Buyers and sellers on eBay use feedback and reputation to determine how willing they would be to conduct transactions with another user. A buyer can look back indefinitely into a seller's past feedbacks and ninety days into the details of the related transactions. These details can show how consistent the seller is, whether they've sold the caliber and type of item the buyer is looking for, and other information that is pertinent to a buyers' choices. Sellers can view similar data about users that may be interested or participating in their listings. The seller can then determine if the buyer seems legitimate enough to be allowed to bid on or purchase an item. A seller may also make allowances for pristine buyers. The seller may be willing to ship an item before receipt of payment if the buyer's reputation is clean enough. If these decisions are made correctly, with accurate information, there is a higher probability that the transaction will be completed with no problems. Issues like significantly not as described items, items not being

received, and other PayPal claims can be prevented if the reputation system is accurate and robust, and eBay users make use of it correctly.

A user's feedbacks and overall reputation are also considered in deciding whether or not a user should be given extra monetary protection. PayPal's Buyer Protection covers buyers for items that were never sent and items that arrive but are not as described by the seller [43]. For a seller's listings to be eligible for buyer protection the seller must have a feedback score of at least 50 and 98% positive feedback [44]. This helps eBay and PayPal save money by not insuring users who are prone to being involved in claims, however negative feedback is not required if a user is involved with a claim.

Internal eBay customer service personnel can use the feedback system to aid in decisions pertaining to claims against eBay users. Generally one user makes a claim against another user. The customer service agent can review the feedbacks and overall reputation of each of the users in question. With this information, the customer service agent can begin to form opinions about the users. In the case that a user makes a false claim against the other user, the customer service agent will see evidence that the accusing user does not have a very good reputation or that the user being accused has a pristine reputation. In either of these cases, the customer service agent will have an idea as to whether the claim is plausible or not. This can reduce the time and effort spent on each case and allow for more pertinent actions on the part of the customer service representative.

2.10 Identity

Identity is a very important concept in reference to a feedback system like eBay's. If there is a weak identity system, users who have negative feedbacks and an overall poor reputation could simply shed the skin of their old account and sign up for another, unconnected account and start from scratch. eBay attempts to prevent this by using unique identifiers that connect any accounts owned by the same person. There are several different ways to identify a person that are in use currently. These vary in strength from temporary email address to government issued identification numbers. The strength of the identify characteristic determines how robust the verification procedure is required to be

to validate the given information. This also affects how reliable the characteristic is for determining unique individuals.

A few of the identity attributes that eBay requires when creating an account with them include name, street address, phone number, email address, birthday, and a user ID [45]. PayPal requires the same information but uses your email address as your user ID. These pieces of data combine into a user's identity. The larger quantity of more distinctive pieces of data that are used, the more unique and reliable that identity is. If one were to only use an email address as an identifier, it would not be very dependable because any one person can sign up for multiple free email address from any number of sites online. However, if one were to have the ability to verify a social security number or a credit card number, that would be more trustworthy than an email address or many other personal identifiers. Each bit of information about a person has verification challenges that must be met to validate the fact.

There are situational difficulties to verifying different types of identifying data. For confirming an email address, one could send an email to that address and require the user to either reply to that email or input information that can only be found in that email, like a confirmation number. This would at least verify that the user has access to the email account given.

PayPal has a similar system that is used to verify ownership of a bank account [46]. After a user has added a bank account to their PayPal account, they have the option to confirm that account. This gives the user added benefits of higher spending power, increased security, and more buyer and seller confidence [47]. When a user wishes to confirm a bank account, they notify PayPal. PayPal then makes two deposits to the bank account each with amounts between \$0.01 and \$0.99. Once the user receives these deposits, they return to their PayPal account and enter the two deposit amounts. If the two amounts are correct, then the account becomes a PayPal confirmed bank account.

Most companies that accept credit cards as a payment form require the user to supply some information about the billing address of the credit cards. The zip code of the billing address is the most basic field to verify, but some companies require the full mailing address, contact phone number, and email address to be verified. This system may or may not work for prepaid credit cards that are now available. The information

connected with the credit card number most likely does not match the information of the person who possesses the card itself. The card may seem to belong to “Prepaid Credit Card Customer” and have a billing address of the retail store where the card was purchased. If the user attempts to input their actual information, the system would deny the charge. Also, if a user does know the information connected to the card, they could put that information in and of course it would not make any connections between that card and any of the user’s other cards on file. In either case, some legitimate users may be denied while some illegitimate users could be allowed to use others’ credit cards, especially pre-paid ones.

Verifying something as secure as a social security number poses several other issues. In this case, not everyone has a social security number so this would not be a good choice of verification attribute to use with a system that is not solely based in the United States. Even in this case, people who live in the United States who don’t have social security numbers would not be able to use this system. Another issue would be a person’s willingness to supply their social security number to a company, especially online. Not every web page is completely secure and users would be concerned with unauthorized access to their personal information. Lastly, even if a user is willing to give their social security number, one would most likely wish to verify that the user is giving their genuine social security number. To do this, one would require access to a government database containing social security numbers and detailed information about each of the entries. This would add the issue of keeping the connection to the database, and the database itself, secure. While validating a user’s social security number may be a very strong form of identity, it also has the most risks for both the user and the system itself.

Another way of identifying a user is by the computer that they are using to access eBay’s webpage. Each internet browser session can be tracked by placing a cookie on that computer. This can allow the site being viewed, in this case eBay, to track information about what the user is doing. This information can include where the user is traveling to on that site, what they are clicking on, what information is being typed into fields on pages, etc. This form of identity is used less for security reasons and more for information gathering on the part of the company whose page is being visited.

Each of these forms of identity verification has slightly different uses, very diverse procedures required to verify the given information, and associated risks and rewards. PayPal requires credit card billing address verification for credit cards added to users' accounts. PayPal's bank account verification procedure has previously been described. eBay uses a combination of the above-described identity datum for several different applications.

2.11 Shannon's Information Theory

Claude Shannon, the mathematician and engineer, created information theory in the 1940's. Originally, his research was directed towards communication theory but later became of significant value to mathematics and probability theory [48]. Information theory as created by Shannon relies on the ideas of information and bits. When a certain symbol is chosen from a set of possibilities and is transmitted, the knowledge that is gained from that symbol is its information. The basic unit of information in Shannon's theory is the bit and logarithmic measures are used in the determination of information gain. Therefore, a switch with two positions can store one bit of information. Shannon also created the idea of self-information. When many data symbols are being transmitted, the information gained by each one is equal to the logarithm of the inverse of its probability [49]. This concept is very useful in this project for determining the information gained by a negative feedback that has been left for a user. For the first negative feedback a user gets, the information gained is tremendous because the probability of a negative feedback was very small until that point. Because most of the users on eBay have between 85 and 100 percent positive feedback, linear scaling techniques would not show much difference in the range. Shannon's theories point towards the use of logarithmic scaling for the information that we are using.

3 Design & Implementation

During our nine weeks at eBay we created many new representations of reputation in eBay's feedback system. These representations range from new numerical calculations using the existing transactional information and using probabilistic models to predict the occurrence of negative transaction to making graphical representations of

feedback. The different methods are described in the following sections. Our mentor was also interested in testing a new method of sampling data in eBay's system to determine if it may be a better method than what is currently used.

3.1 Maximum Proxy Bid vs. Final Value

eBay's auction system uses what are called proxy bids to allow users to bid on an item, but not have to constantly watch the auction to see if another user outbid them. [50] When a user places a bid, they are required to bid at least a small percentage above the current highest bid. [51] This amount is calculated based on the current price of the item. Table 3.1 shows the price ranges and associated minimum bid increments. A user may choose to simply bid the bid increment above current highest bid or they can input a maximum proxy bid. In the case of the user inputting a proxy bid, the system bids the bid increment above the last highest bidder's bid until one of two things happens. The first is that other users are bidding or have placed proxy bids above this user's proxy bid. In this case, the user does not win the auction and the only way to win the auction would be to place another bid. The second situation is that the user's bid becomes the highest current bid if no one else has a higher bid in place. In this case, other users may place more bids in attempt to win the item. If this becomes the case, the system continues to place bids as before, again, until either the user becomes the highest bidder or the highest bid exceeds the user's maximum proxy bid. In all cases, the user's maximum proxy bid is known only to the bidder and never the seller or the competing bidders. These cycles repeat until the close of the auction.

Current Price	Bid Increment
\$ 0.01 - \$ 0.99	\$ 0.05
\$ 0.99 - \$ 4.99	\$ 0.25
\$ 5.00 - \$ 24.99	\$ 0.50
\$ 25.00 - \$ 99.99	\$ 1.00
\$ 100.00 - \$ 249.99	\$ 2.50

\$ 250.00 - \$ 499.99	\$ 5.00
\$ 500.00 - \$ 999.99	\$ 10.00
\$ 1000.00 - \$ 2499.99	\$ 25.00
\$ 2500.00 - \$ 4999.99	\$ 50.00
\$ 5000.00 and up	\$ 100.00

Table 3.1 – Table of Bid Increments

The concept behind a proxy bid is that a user puts in the maximum amount of money they would be willing to pay for a given item. As described above, if this user wins the auction, they only have to pay an increment above what the second highest bidder bid. This could allow for the winning user to have to pay much less than they were willing. Of course, this is a good situation for the buyer, they save money. However, if you imagine this same situation, but from the seller’s perspective, they are missing out on a lot of revenue. Since the seller never knows what the winner’s proxy bid is, no seller will realize the loss of money in that manner. However, if the auction ends at an unsatisfactorily low price for the seller, they may be inclined to not sell the item even though a legitimate auction has been completed for it. This would lead to seller-nonperformance and subsequent PayPal claims. Since eBay has access to the proxy bid amounts, it could be possible to flag transactions with large differences between the maximum proxy bid and the final price that the item sold for. Having flagged those transactions, eBay could take steps toward preventing this particular situation or they could even intervene and attempt to keep negative transactions from taking place.

In designing a tool to aid in comparing proxy bids to final values of a given auction, we came across several hurdles. The largest of these issues was that the two main pieces of data we were looking for, the proxy bid and the final value, were not collocated anywhere. The proxy bids were only available through Sojourner, the click-tracking system, and the final value was not accessible through Sojourner. The most logical place to get the final value was from the API. Sojourner was able to list an item ID, the user ID of a user who bid on that item, and the amount of the bid placed. This information was given for every bid made by a user within Sojourner’s limited scope.

Given an item ID, the API gave us the user ID of the winner of the auction and the final value of the item. Combining these pieces of information, it can be determined whether the given user was the winner of the item, a loser, or if there were odd circumstances and the bids and final value can be compared.

Several situations can arise when comparing all of these bits of information. Obviously, if the winner of the item is listed as the original user, then they won the item and their proxy bid must be above the final value. It can also happen that the original user is not the winning bidder, but shouldn't be because their highest proxy bid was not above the final selling amount. However, if the original user is not listed as the winner but has a proxy bid higher than the final value, several exceptions could have occurred.

We observed a small handful of the circumstances surrounding the situation where a user has a proxy bid higher than the final value, yet is not the winner of the auction. In one case, the seller had ended the auction early and cancelled all bids on it. In other cases, the individual bids had been retracted for any number of reasons. Some of these reasons include the bidder entering the incorrect bid amount, the bidder deciding not to buy the item because the seller changed the description, the seller denying a bid because of low or poor bidder feedback, and any other reason that are possible to use while making a bid retraction. We have also come across several situations that it seems prevent the winning user ID from being available through the API. The three reasons we've found that can cause this are: private listings where the user names are not disclosed, auctions with more than one item available, and if the reserve price was not met. There may be other issues similar to this, but within the narrow time window available, further testing was not possible.

Sojourner caused a small inconvenience by listing each day's worth of data in a separate file. Originally our program was designed to take in multiple Sojourner output files and output one single summary file. However, an issue arose, not with the design of our code, but with the robustness of the eBay API. It became a recurrent issue that the connection to the API would drop after any large amount of time. It just so happened that the amount of time to run a single Sojourner file, only one day's worth of data, was just short enough that there were few API disconnections when running only one file. As a solution to this problem, our software only runs one Sojourner file at a time.

The end result of this portion of our project was to gather the two major pieces of data into one place to then be analyzed by others. The program written to do this takes a file from Sojourner, parses then compares the data to information from the API, then outputs a file giving summary information. This output file is tab-delimited and therefore can be easily accessed through Microsoft Excel or most data-analysis tools. This should allow for easy examination of the data.

There is a two-step process in obtaining and analyzing the specified data. First, a graph is run through Sojourner which gathers the first section of data. This information is then compared to data accessed through the eBay API. The results from these comparisons are printed into an output file.

The first step is to run a Sojourner graph and retrieve the returned files. The graph that is run calls for each bid placed on a given eBay site that is a confirmed bid placed by a user within Sojourner's limited user base. For each of these bids, Sojourner creates a branch of a tree. The item ID is the first node with any users who bid on that item in the second level, as the child nodes of the item ID. The proxy bids of those users on that item are the third level of the tree, as children of the user ID. Sojourner outputs a tab-delimited text file of this tree structure. After the Sojourner files are gathered, they are placed in a folder in the working directory for Java within Eclipse and our Java application is run on these files to gather the final data. A sample Sojourner output file is included in Appendix A.

As previously mentioned, Sojourner outputs a data file per day's worth of data. Our code circumvents this using a directory object. The Sojourner files are placed in a predetermined directory in the current workspace. The directory object keeps track of the directory and file names within the given directory by keeping an array of the file names within the folder. To prevent accidental reading of the versioning files within the folder, each filename is tested to match the form of a Sojourner data file, before being processed. Each of the Sojourner data files are opened and processed consecutively. However, with the aforementioned issue of the API disconnections, not all files can be processed in one run of the software. Generally, about 5000 lines of Sojourner data can be processed before issues with the API arise. If the Sojourner input files are small enough, multiple files can be processed in one application run. If not, each file must be run separately. In

either case, for each Sojourner input file, there is one output file. Currently, to allow for both of these situations, there is a loop that controls how many and which files in the data folder are processed in one application run. This loop should be adjusted depending on the size of the Sojourner files being inputted.

Given the form of the Sojourner files, the parsing of this data was not entirely difficult. Each level of the aforementioned tree is indented a different amount. Each line of the Sojourner text file is read in and parsed into an array by tabs. There are tests along the way to determine if there are more items, more users who bid on a given item, or more bids by one user on the given item. As the Sojourner data file is traversed, data is stored and gathered, and then summary lines are appended to the output file.

As the data from the Sojourner file is collected, API calls are made and conclusions are made about the bid. The item ID is parsed out of the Sojourner file first and then a call is made to the API to retrieve information about who the actual winner was and how much the final value was. The bidding user ID is then extracted from the Sojourner output file and compared to the winning user ID. The proxy bid amount is then extracted and compared to the final value. After gathering these pieces of information, the status of the bid being examined is determined.

The possible statuses consist of “WINNER”, “LOSER”, and “ODDITY”. If the bidding user ID matches the winning user ID and the proxy bid is greater than or equal to the final value, then the bid is determined to be a winner and is marked as such in the output file. If the bidding ID does not match the winning user ID and the proxy bid is not higher than the final value, then the bid is marked as a loser. If anything other than these two situations occurs, then the bid is marked as an oddity. As mentioned before, these usually occur due to bid retractions and cancellations.

For each bid extracted from Sojourner, a line in the output file is created. This line includes the bid status, the item ID, the bidding user ID, the final value of the item, the bid placed by the user, the difference between those two amounts, and the category ID of the item. The category ID was requested by our mentor and was already returned by the API call required to gather the winning user ID and the final value. This output file is tab delimited to allow for easy import into many other data-analysis programs. An example of the output file is included in Appendix B.

3.2 Volume Reputation

Volume reputation includes several that could allow for more transparency of an eBay user's reputation. These statistics are based on the final selling value of the item that a feedback is left for. The most important of these statistics are: volume feedback score, volume percent positive, ratio of positive volume to negative volume, the average of the positive and total volume, the median of the positive and total volume, and long tail images.

Volume feedback score is essentially the same as eBay's current feedback score, except weighted by volume. eBay's feedback score consists of the number of positive feedbacks minus the number of negative feedbacks. Applying that to volume, a volume feedback score is calculated by subtracting the sum of the values of the negative transactions of a user from the sum of the value of their positive transactions. This shows not only the number of positive transactions compared to the negative transactions, but also shows the volume of positive transactions compared to the volume of negative transactions.

The volume percent positive is also very similar to eBay's current percent positive feedback, but with volume taken into account. This statistic is calculated by dividing the sum value of positive transactions of a user by the sum value of all of their transactions. This is taken as a percentage. Similarly, negative and neutral percentages can be calculated, but seemed less important than the volume percent positive.

Our mentor also suggested that the ratio of positive volume to negative volume be included. This is calculated simply by dividing the sum of the value of a user's positive transactions by the sum of the value of their negative transactions.

Two other metrics included were the median and average of a user's positive, negative, neutral, and total volume. Of these eight statistics, the four that seemed most important were the positive and total median and the positive and total average. The medians seemed to help single out both users who were padding their feedback scores and users that were breaking out of their current price bracket. The difference between the medians and averages helped point these distinctions out as well.

The last piece of volume reputation is the long tail images. These images are essentially sorted, colored bar graphs of the volume and feedback of a transaction. A

positive transaction was colored green, a negative transaction was colored red, and a neutral transaction was colored grey. The width of the bars was a fixed width of one pixel and the height of the bar was the volume of the transaction. The transactions were placed on the graph in decreasing volume from left to right. A sample long tail graphic is included with the example output in Appendix D.

In designing an application to calculate all of these statistics, we used only the eBay API. Unfortunately, the two pieces of information that were the most useful, the feedback rating and the value of the item, were not easily accessible from the same API call so we used two different API calls to gather the required data. With the first API call we were able to access pages of feedbacks. Each page contained 25, 50, 100, or 200 feedbacks and multiple pages could be accessed one at a time. For each feedback there were many fields of data but the ones we used were the feedback type (positive, negative, or neutral) and the item ID. From the item ID, we were able to make the second API call to retrieve the item information including the final selling price of the item. Using these pieces of information for multiple transactions, the aforementioned statistics could be calculated and stored. To allow for portability into other portions of the project, an object was used to store both the gathered data of the transactions as well as the summary statistics.

The basis of the implementation of this portion of the project consists of a volume feedback score object. This object stores both the transaction data gathered and the calculated summary statistics. The transaction data is stored in one of two types of arrays. One simply holds the value of the item while the other array holds much more detailed information. Each element in the arrays represents a transaction with feedback. The detailed data kept for each transaction includes: the role that the user played, the feedback type, the time the feedback was placed, the comment given, the commenting user, the feedback score of the commenting user, a boolean which is true if the data for the item is available, the price the item sold for, the item ID, the URL for the item, and details about the images for the item. There are also variables for each of the stored summary statistics. There are functions to draw each of the positive, negative, neutral, and total long tail images which access the volume feedback score object for its data.

The gathering of the individual transaction data was the most complex step in this process. In creating a volume feedback score object, the user specifies which eBay user's feedback information to use and how many feedbacks they would like to have collected to have the statistics calculated on. Using the number of feedbacks requested by the user, the size and number of feedback pages is calculated. Since the smallest number of feedbacks that can be requested is 25, the number of feedbacks requested by the user is rounded up to the next increment of 25 for simplicity. The software then retrieves that number of feedbacks, one page at a time, and compiles the data.

Two different API calls are required to gather all of the data. Each page of feedbacks requires one API call where each item requires its own API call. For each feedback returned from the feedback API call an item API call is made. This requires at least one API call per item that is included in the data set. Once the data for a transaction is gathered, it is stored into select arrays and sum and count variables are incremented.

Several arrays of feedback and item data are kept, each for different reasons. Four arrays of doubles are kept, one for each of the sets of positive, negative, neutral, and all item values. These arrays are sorted by the value of its entries. These are used to calculate the positive, negative, neutral, and total medians. Two special arrays are kept that contain very detailed information for all the feedback transactions collected. One of these arrays is not sorted, therefore in the order received. This array is used in the sparkline drawing software. The second array is sorted by the value of the item. This array is used in the drawing of the long tail images. Each of these arrays has a get-function so that the data may be accessed outside the scope of this individual application.

In addition to arrays, several variables are kept as running sums or counts of various pieces of data. There are four double variables that keep a running sum of the values of the positive, negative, neutral, and all transactions. There are five integer variables that keep a running count of the positive, negative, neutral, all, and bad items. These variables also have get-methods to allow access outside of this application.

Once the predetermined number of feedbacks is processed, the statistics are calculated and printed to a file. Each of the statistics is calculated as described above and stored in its own variable. Along with calculating the statistics, two more API calls are made to retrieve eBay's actual feedback score for the user and their PayPal protection

status. The volume feedback score, volume feedback percentages, and volume ratios are all calculated using the running value sum variables. A comparable eBay feedback score and feedback percentages are calculated using the running count variables. The minimum and maximum values are found for the double arrays containing only the transaction values. A function was written to calculate the median of these arrays as well. Another function was written to calculate the average of the same arrays using the running sum of the positive, negative, neutral, and total values and the running counts of the same. Other ratios between the positive and total medians and averages are also calculated. After the calculations are complete, each summary statistic is printed to an output file that is named after the user the program was run on and the time and date of the run. A sample output file is included in Appendix C.

After this is complete, the long tail graphics are drawn. These graphics were quite simple to draw since there was already a sorted array with all of the required information in it. A long tail graph is drawn for each of the positive, negative, neutral, and total value arrays. Starting on the left, a vertical line is drawn from the bottom of the image to the value of the largest item. The scale is one pixel wide per item and one pixel tall per dollar. If the feedback was positive the line is drawn in green, if the feedback was negative the line is drawn in red, and if the feedback was neutral the line is drawn in a dark grey. The second largest item is then drawn in the same manner and so on until there are no items left. If there are no feedbacks in the given array, the long tail is not drawn. Samples of these images are included in the aforementioned sample output in the appendix.

3.3 Reputation Images

For this project we created a couple of image types that represent reputation for users on the eBay website. The images that we are able to create are sparklines and stars. Both of the images were made to be easy to add new features to or update with different reputation metrics. The superclass for reputation images allows an implementation to use volume feedback score objects to gather the data required for creating the image.

With increasing adoption of broadband connections, images can be used increasingly more often. Many people can understand simple concepts in images such as

color, shape, and size in order to replace tables of data in a much smaller space. Volume sparklines and logarithmic stars attempt to achieve simplicity in design and should be easy for users to understand with little instruction.

3.3.1 Volume Sparklines

Volume sparklines are graphical views of past transactions a user has had a role in. The sparklines show vertical whiskers along a horizontal axis with each whisker representing the feedback for one transaction. The size and direction of each whisker represents the volume or price of the transaction and whether the feedback left was positive, negative, or neutral. An example of a volume sparkline can be seen in Figure 3.1. The green whiskers that go upward from the center are positive feedbacks. The red, downward whiskers are negative feedbacks, and the grey centered whiskers are neutral feedbacks. The full height of the image represents \$200. Because of this, the largest volume that can be represented for a positive or negative transaction is \$100 while neutral can show up to \$200 of volume.



Figure 3.1 – Volume Sparkline

The dimensions and scale of the sparkline can be different. For some applications, a long sparkline showing the last 200 transactions is useful but there may not always be space for it. In this situation a shorter sparkline showing only the last 50 transactions might be a better fit. For this project we made the sparklines very modular. The class that generates the volume sparklines can be customized to have a different height, width, scaling factor, or whisker thickness. For users who deal primarily on the eBay Motors site, a scale that allows users to see deals of more than \$1000 is very useful.

Another important option that we added to the volume sparkline class shows only the feedback from when the user was the seller. This option has a striking effect on some users' sparklines. Someone may look like they have many high value transactions but they've only sold items that cost one cent.

The volume sparkline generating class takes in a volume feedback score object and uses the data that it gathered. The information required for the sparkline are the role

of the user in each transaction, the final value of the transaction, and whether the feedback left was positive, negative, or neutral.

If the option for seller only is set, the program starts by removing all of the non-seller transactions from the list, otherwise the original list is left intact. The generator then determines the scaling factor for the whiskers based on the height of the image and the scaling mode set through the options. If it is set to automatic scaling, then the largest value of all of the transaction must be determined first. This value is used as the largest value that can be displayed in either the positive or negative direction. The generator traverses the list of feedbacks, drawing a whisker for each one starting from the right edge of the image. We made the most recent transaction on the right side to be more intuitive. The sparkline is similar to a timeline which goes from oldest of the left to newest on the right.

3.3.2 Logarithmic Stars

Logarithmic degradation stars is a method of mapping eBay’s feedback into a five star rating system. This method uses Shannon’s Law of Self-Information [49] to determine the effect of negative feedback on a user’s star rating. If a user has only one negative feedback, it contains more information than one of a user who has 100% negative feedback. With this idea a logarithmic scale is created. For a user with 1024 total feedbacks, the first negative feedback the user has takes away half of a star. In order to lose the next half star they need to be given two more negative feedbacks. The third half of a star is lost by being given four more negative feedbacks. Table 3.2 lists the percent of positive feedbacks the user has and the star rating they would have. The program we have written determines star rating of a given user and outputs a star rating image with 0 to 5 stars in half star increments.

% Positive Feedback	Star Rating
100 – 99.90234375	★★★★★
99.90234375 – 99.8046875	★★★★☆
99.8046875 – 99.609375	★★★★
99.609375 – 99.21875	★★★☆☆
99.21875 – 98.4375	★★★

98.4375 – 96.875	★★★★
96.875 – 93.75	★★★
93.75 – 87.5	★★
87.5 – 75	★
75 – 50	☆
50 – 0	

Table 3.2 – Percent Positive Feedback and Corresponding Star Rating

This progression is defined using a formula using logarithms and the percent of positive feedback. The class which generates the star rating for a user can be changed through options to use different rounding methods and different feedback metrics. The two methods of rounding implemented in the generating classes are floor and round. Using floor brings a users rating down to the first discrete cutoff point lower then their actual rating. For example, if someone should have a rating of 4.78 stars they would be brought down to 4.5 stars. Using round, the star rating is changed to the nearest discrete cutoff point from their original rating. The same user with 4.78 stars would go up to 5 stars with this method.

In addition to the different rounding methods, different feedback metrics can be used to determine the percent of feedback that is positive. The generating class can use binary feedback scores or volume feedback scores.

3.4 Sampling by Volume

For sampling by volume a cumulative distribution is created with each item listed on eBay represented by one unit on the x-axis and the selling price of the item as its contribution to the cumulative distribution. To obtain a sample of the objects represented random values are selected between \$0 and the total volume of all of the objects. The item on the x-axis at which the distribution is equal to the selected value is added to the sample set.

Because this operation can take a long time to go through the items to find a particular item, we’ve also developed a quicker method for sampling many items. The class can take in a list of positions to take samples from. This list is then sorted and the

samples are taken. Because the sample points are sorted, the program can go through the list one time and find the items that match each sample point.

In order to implement this, we created a basic sampler class. Subclasses of this base class implement functions that return the value of an item, the total value of all items, and items based on an index position. These values allow the base class to sample items based on many types of volume. We created two subclasses using this system. The first subclass demonstrates sampling that gives each item equal weight. In this implementation, the returned value of each item is 1 and the total value is simply the number of items.

We also implemented sampling by volume using this method. The volume of all items is determined when the object is first created. A loop goes through all of the items and totals the selling price of each one. When the value of a particular item is requested, the current selling price is returned. This allows the base class for sampling to perform sampling by volume on a list of items.

4 Results & Analysis

At the culmination of our project, we reassessed our original goals and reflected upon the progress made. Overall, each objective was met with few obstacles and even fewer known bugs. The remainder of this section details our thoughts and reflections upon our completed work.

4.1 Proxy Bid vs. Final Value

The goal of this portion of our project was to create an application which would gather and compare data about the maximum proxy bid and final value of auctions on eBay. This goal was met, although the resulting application is not the most elegant of solutions.

This application succeeded in gathering data about proxy bids, the auctions those bids were made on, and the final value of that auction. This data is compared and displayed in a format that is easy to understand and to analyze using external tools. The only negative aspects of our application deal with what is required to run data, especially large amounts of data.

The first undesirable aspect of our application is that for each set of data, two individual steps must be taken by an individual to complete our analysis. First, someone must run a Sojourner graph, which if there is a large amount of data, could take as long as several days. The resulting files must then be moved into a folder in the working directory of our application before our software can be run. Because our software uses API calls, which are not designed to be streamlined and quick, running it on large amounts of data also takes a large amount of time. These issues were unavoidable due to the lack of appropriate interface with Sojourner and the constraints of using the API.

The second undesirable aspect of our application is the inability to condense the data from a set of files from Sojourner into a single output file. This was also caused by constraints from using the API.

4.2 Volume Reputation

The goal of this portion of our project was to create an application which would extract feedback data and calculate volume-based feedback statistics. This goal was met without any major negative aspects.

This application succeeded in gathering data about feedbacks and the item for which the feedback was left. This data was used to calculate several feedback statistics and images that were organized in one output file and at most four images.

4.3 Reputation Images

The reputation images that we were able to generate based on the feedback and listing information available through the API ended up being very successful. We were able to accomplish all of the goals set forth for them and more. The people who have tested sparklines and logarithmic stars during the development process were happy with them and have even stated that they will no longer shop eBay without them. We feel that this means that these new representations of reputation have the utility that was expected and possibly more.

For demonstrating these new feedback visualizations, scripts were created for the Greasemonkey plug-in for Mozilla Firefox. Greasemonkey allows JavaScript scripts to be run when certain Web pages are loaded. These scripts allow for sparklines or logarithmic

stars to be inserted into live eBay pages at the user's computer. These scripts inserted the images on the Get Result, Compare Items, and View Item pages. The Get Result page is shown after the user performs a search. Each listing in the search results can be given a reputation image making the comparison of many items concurrently fairly simple at a glance. The Compare Items page is shown when the user selects multiple items on the Get Result page and presses the Compare button. On this page, a table is shown with different properties of the selected listings shown side by side. Using the script, an extra row is added to the table which houses the reputation images. Lastly, the View Item page is another place where reputation images fit very well into the current eBay site. The View Item page is shown when the user has clicked on an item to view its full description. The page has a box with information about the seller. This space fits either reputation image very well. Examples can be seen in Appendix E.

4.3.1 Volume Sparklines

Volume sparklines work very well to help show the volume of users' recent transactions. In a buying situation, it is important to be able to tell whether or not the seller has experience at the volume of the current listing. Volume sparklines allow users to spot sellers who are breaking out of their selling profile with a quick look.

Certain types of abuse and fraud are easy to identify visually using volume sparklines. Figure 4.1 is an example of what the sparkline of a feedback padder looks like. There are many small transactions that are for less than a dollar or two, with occasional high-value transactions. Figure 4.2 shows the sparkline of a user who writes misleading descriptions in their listings. This user was selling plasma television wholesale lists. Many users were tricked by the listings and left negative feedback for the user.



Figure 4.1 – Volume Sparkline of a Padder



Figure 4.2 – Volume Sparkline of a Misleading Advertiser

4.3.2 Logarithmic Stars

Logarithmic stars were also a success. Using a logarithmic scale to help differentiate the small range of typical percent positive scores makes the difference between a 99% positive user and a 90% positive user much easier to see. New users who are familiar with other star rating systems should be able to understand this system very well.

When given many users to compare, logarithmic stars allow users to easily distinguish different levels of reputation. By having distinct levels for sellers, buyers can easily select a level of seller that they feel comfortable buying from.

4.4 Sampling by Volume

Our sampling by volume system worked very well. Unfortunately we were not able to test sampling by volume against the normal methods of sampling by listing or user. Whether or not sampling by volume will make a significant difference in the results is still unclear at this time. At this time, other departments inside of eBay have inquired about sampling by volume and hopefully it will undergo further testing and use.

There are some problems with sampling by volume. Some items are very over-represented in the result set. If sampling over a wide variety of items, real estate and eBay Motors items are often selected very often. Having a set of only these very high value items can easily throw off any measurements done on the set if this selection is not noticed.

Another problem with sampling by volume that needs to be addressed is the difference between different currencies. A \$300 laptop might show up as have a volume of 20,000 in some other currency causing it to be highly weighted compared to most other items in the original set of listings.

Despite these problems, the sampling by volume prototype that we made works very well as a proof of concept. There is still work to be done with this concept but our preliminary tests show that it could be a very useful analysis tool.

5 Recommendations & Conclusion

eBay's reputation system works to give users a sense of trust and respect for the other users of the system. The system does this through comments that users can leave for others. The comments can be given a positive and negative score which contributes to an overall score. A user's score only provides numerical measures for reputation, which do not have a maximum value. New users can have a very difficult time deciding whom to trust because they do not fully understand what these numbers mean. Extra information about particular feedbacks could help to make reputation more transparent. Reputation systems need to be robust and have a high level of success in giving users accurate information about other users. The system must also prevent users from providing inaccurate information or trying to abuse the system.

Overall our proxy bid vs. final value application meets all of the requirements set by our mentor and fits the design we created originally. There is still some room for improvement in the flow and speed of the program. If internal non-API calls could be used, the program would run a lot faster. Also, if it were possible to interface Sojourner directly with our software, it could remove the need for manual transfer of data and streamline the entire process of running our software. If this concept is to be used frequently, it might be worth making these changes. As for the data returned by the limited test sampling that was done, studies should be conducted to determine the usefulness of proxy bid vs. final value comparisons.

Overall our volume feedback reputation application meets all of the requirements set by our mentor and fits the design we created originally. There is some room for improvement within the application, mostly in reference to the API. If the API calls were replaced with internal data calls, the program would run much quicker. As for the data collected, studies should be conducted to determine the usefulness of volume-based feedback systems.

Reputation is not presented in a way that makes it intuitive. Sparklines and logarithmic stars attempt to solve this problem. Sparklines allows advanced users to see large amounts of past data in a small space and without having to jump between many different pages on the site. Adding volume to normal sparklines allows them to show much more information in the same amount of space as they use without it. We

recommend that more tests be done using sparklines, including user tests. Also, a good scale for the maximum sparkline value should be determined. One possibility is to have a varying scale based on the category of the listing.

Logarithmic stars can be used for comparing many users at once. They separate users into distinct level that allow for buyers to easily eliminate sellers they are not comfortable purchasing from. Using volume to determine the level of stars for a user helps to make users with many small positive transactions and a few large negative transactions have lower ratings. Future work on logarithmic stars would be to determine whether or not half star increments are necessary and also to develop new graphics for the star. Perhaps always showing 5 stars but having some grayed out or hollow when a user has fewer would make the presentation look cleaner. Like sparklines. User tests to see whether or not users like the stars for rating would be useful.

Sampling by volume is an interesting concept but needs to be worked on further in order to be useful for any applications. It needs to be able to convert different currencies to be able to compare them correctly. Also, using regular volume may not be the best method to sample by for eBay. Since eBay's profits are not based directly on volume, using the volume that is made by eBay for each listing might be a better property to use as a weight. Sampling by volume needs to be tested in different situations to determine whether the difference is sampling has a significant impact on the measurements taken from the data. We hope that it will be very useful for other projects.

In conclusion, eBay's reputation system is a very useful tool for users of eBay's site. Although there are many legitimate users of the site, there are still ways in which fraudsters can try to manipulate the system and rip off other users. The tools we've created during the time spent on our MQP will hopefully make buyer and seller reputation more transparent to others so that fraud will become less common. Also, we hope that they can be used internally to help police the system.

[1] Boyd, M. J. "System and Method for Measuring Reputation Using Take Volume." U.S. Patent Pending. December 2006.

[2] Rivlin, G. "eBay's Joy Ride: Going Once..."
<http://www.nytimes.com/2005/03/06/business/yourmoney/06ebay.html?pagewanted=1&%2338;ei=5088&%2338;en=dbe72934a0e4f364&ex=1267851600&adxnnl=1&adxnnlx=1163970469-k0qwAwjEbbb3Rv6qYvprdA>. Page 2, March 6, 2005.

[3] Mullen, A. "The History of eBay." <http://www.happynews.com/living/online/history-ebay.htm>. November 12, 2006.

-
- [4] Wolverton, T. "Billpoint Failure A Lesson for eBay?" http://news.com.com/Billpoint+failure+a+lesson+for+eBay/2100-1017_3-942231.html. July 8, 2002.
- [5] eBay, Inc. "eBay Stores Toolkit." <http://pages.ebay.com/storefronts/seller-landing.html>. November 12, 2006.
- [6] Half.com, Inc. "Half.com." <http://www.half.ebay.com/>. November 12, 2006.
- [7] Viva Group, Inc. "Rent.com." <http://www.rent.com/>. November 12, 2006.
- [8] eBay, Inc. "eBay Motors." <http://www.motors.ebay.com/>. November 12, 2006.
- [9] eBay, Inc. "Two Millionth Passenger Vehicle Sold on eBay Motors." <http://investor.ebay.com/releasedetail.cfm?ReleaseID=206868>. August 8, 2006.
- [10] eBay, Inc. "Using My eBay – Overview." <http://pages.ebay.com/help/myebay/myebay-overview.html>. November 12, 2006.
- [11] PayPal, Inc. "About Us – PayPal." <http://www.paypal.com/cgi-bin/webscr?cmd=p/gen/about-outside>. November 12, 2006.
- [12] eBay, Inc. "Trust and Safety." <http://pages.ebay.com/aboutebay/trustandsafety.html>. November 12, 2006.
- [13] eBay, Inc. "eBay Help: Rules and Policies." <http://pages.ebay.com/help/policies/index.html>. November 12, 2006.
- [14] eBay, Inc. "eBay University." <http://pages.ebay.com/university/>. November 12, 2006.
- [15] eBay, Inc. "eBay Glossary." <http://pages.ebay.com/help/newtoebay/glossary.html>. November 12, 2006.
- [16] eBay, Inc. "eBay Acronyms." <http://pages.ebay.com/help/newtoebay/acronyms.html>. November 12, 2006.
- [17] eBay, Inc. "eBay – Investor Relations." <http://investor.ebay.com/index.cfm>. November 12, 2006.
- [18] eBay, Inc. "eBay Foundation." <http://pages.ebay.com/aboutebay/foundation.html>. November 12, 2006.
- [19] eBay, Inc. "eBay Careers." <http://www.ebaycareers.com/>. November 12, 2006.
- [20] Omidyar, P. "Founder's Note." <http://pages.ebay.com/services/forum/feedback-foundersnote.html>. November 11, 2006.
- [21] eBay, Inc. "About Feedback." <http://pages.ebay.com/help/feedback/feedback.html>. November 11, 2006.
- [22] eBay, Inc. "eBay Star Glossary." <http://pages.ebay.com/help/feedback/reputation-stars.html>. November 11, 2006.
- [23] eBay, Inc. "Understanding Feedback Scores." <http://pages.ebay.com/help/feedback/feedback-scores.html>. November 11, 2006.
- [24] eBay, Inc. "What is Mutual Feedback Withdrawal?" <http://pages.ebay.com/help/feedback/questions/mutual-withdrawal.html>. November 11, 2006.
- [25] eBay, Inc. "Feedback Abuse, Withdrawal and Removal." <http://pages.ebay.com/help/policies/feedback-abuse-withdrawal.html>. November 11, 2006.
- [26] Amazon.com, Inc. "Seller Feedback." <http://www.amazon.com/gp/help/seller/feedback-popup.html/102-9802966-8697718?ie=UTF8&seller=A23R4GCC2FFZBD>. November 19, 2006.
- [27] Buy.com, Inc. "Marketplace Buyer's Guide." https://secure.buy.com/corp/support/marketplace_help.asp. November 15, 2006.
- [28] Google, Inc. "What Are Reviews?" <http://checkout.google.com/support/sell/bin/answer.py?answer=31685&topic=8963>. November 15, 2006.
- [29] Google, Inc. "What If I Disagree With a Buyer's Comments?" <http://checkout.google.com/support/sell/bin/answer.py?answer=31689&topic=8963>. November 15, 2006.
- [30] Google, Inc. "Where Do These Reviews Appear? How Are They Used?" <http://checkout.google.com/support/sell/bin/answer.py?answer=31688&ctx=sibling>. November 15, 2006.
- [31] Keser, C. "Experimental Games for the Design of Reputation Management Systems." *IBM Systems Journal* 42.3 (Sept 2003): p498(9).
- [32] Ba, S. and Pavlou, P. A. "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior." *MIS Quarterly* 26.3 (Sept 2002): p243(26).

-
- [33] Resnick, P., Zeckhauser, R., Friedman, E., and Kuwabara, K. "Reputation Systems." *Communications of the ACM* 43.12 (Dec 2000): p45(4).
- [34] Zhang, J. and Cohen, R. "Trusting Advice from Other Buyers in E-Marketplaces: The Problem of Unfair Ratings." *Proceedings of the 8th International Conference on Electronic Commerce: The New E-Commerce: Innovations for Conquering Current Barriers, Obstacles and Limitations to Conducting Successful Business on the Internet ICEC '06* (Aug 2006): p225(10).
- [35] DeFigueiredo, D. B. and Barr, E. T. "TrustDavis: A Non-Exploitable Online Reputation System." *Proceedings of the Seventh IEEE International Conference on E-Commerce Technology CEC '05* (2005).
- [36] Cosley, D., Lam, S. K., Albert, I., Konstan, J. A., and Riedl, J. "Is Seeing Believing? How Recommender Interfaces Affect Users' Opinions." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Apr 2003): p585(8).
- [37] eBay, Inc. "eBay Solutions Directory – Tools & Services That Make Buying and Selling on eBay Even Easier." <http://cgi6.ebay.com/ws/eBayISAPI.dll?SolutionsDirectory>. February 14, 2007.
- [38] eBay, Inc. "Success Story – Infopia." http://developer.ebay.com/businessbenefits/success/InfopiaSS.html/pdf_file. February 23, 2007.
- [39] eBay, Inc. "Tap Into the eBay Platform – eBay Developer's Program." <http://developer.ebay.com/>. February 12, 2007.
- [40] eBay, Inc. "About Us – eBay Developers Program." <http://developer.ebay.com/businessbenefits/aboutus>. February 14, 2007.
- [41] Runes, J., Schweers, J., and Sreenath, N. "ThirdEye Seer - Sojourner Data Mining Platform." WPI Major Qualifying Project CS-DXF-EB62. March 3, 2006.
- [42] Gupta, R. (eBay, Inc.) "Sojourner Analytics Platform User Guide." December 2005.
- [43] PayPal, Inc. "Help Center – PayPal: What does PayPal Buyer Protection cover?" https://www.paypal.com/cgi-bin/webscr?cmd=_help-ext&nodeid=56425&leafid=56468&prior_transaction_id=402470&answer_id=16777217. February 14, 2007.
- [44] PayPal, Inc. "PayPal Expands Buyer Protection on eBay." <http://www.shareholder.com/paypal/releaseDetail.cfm?ReleaseID=145650&Category=US>. February 14, 2007.
- [45] eBay, Inc. "eBay.com – Registration: Enter Information." <https://scgi.ebay.com/ws/eBayISAPI.dll?RegisterEnterInfo>. February 15, 2007.
- [46] PayPal, Inc. "Help Center – PayPal: How do I confirm a U.S. bank account already added to my PayPal account?" https://www.paypal.com/cgi-bin/webscr?cmd=_help-ext&nodeid=25679&leafid=1199945&prior_transaction_id=141386&answer_id=16777220. February 16, 2007.
- [47] PayPal, Inc. "Help Center – PayPal: What is the advantage of confirming my bank account?" http://www.paypal.com/cgi-bin/webscr?cmd=_help-ext&eloc=4478&loc=4434&unique_id=1696&source_page=_home&flow=. February 16, 2007.
- [48] Wyner, A. D. "The Significance of Shannon's Work." <http://cm.bell-labs.com/cm/ms/what/shannonday/work.html>. June 28, 1997.
- [49] Shannon, C. E. "A Mathematical Theory of Communication." *The Bell System Technical Journal* 27 (July, October 1948): p379(45), p623(34).
- [50] eBay, Inc. "eBay's Automatic Bidding System (Proxy Bidding)." <http://pages.ebay.com/help/buy/proxy-bidding.html>. February 22, 2007.
- [51] eBay, Inc. "Bid Increments." <http://pages.ebay.com/help/buy/bid-increments.html>. February 22, 2007.

Appendix A: Sample Sojourner Output File

#	Count	%age	Name	SumOfChildren
0	100.00%		GetBids	0
0	100.00%		200068550402	0
0	100.00%		pauldome	1
		1	100.00%	9.99
0	100.00%		170070289884	0
0	100.00%		mr.murray	1
		1	100.00%	16.83
0	100.00%		160075333229	0
0	100.00%		drpezy154	1
		1	100.00%	15.0
0	100.00%		160073319484	0
0	100.00%		weisigns	1
		1	100.00%	2.0
0	100.00%		290073385211	0
0	100.00%		wmf723	1
		1	100.00%	75.0
0	100.00%		220072158175	0
0	100.00%		yellowfish	1
		1	100.00%	5.5
0	100.00%		130068754006	0
0	100.00%		kle8763	2
		1	50.00%	10.99
		1	50.00%	11.99
0	100.00%		170069925928	0
0	100.00%		cloudlessky	1
		1	100.00%	41.66
0	100.00%		260075177229	0
0	100.00%		liteflyer06	1
		1	100.00%	5.0
0	100.00%		bob315751	1
		1	100.00%	11.0
0	100.00%		120075084111	0
0	100.00%		evu348	3
		1	33.33%	270.0
		1	33.33%	295.0
		1	33.33%	310.0
0	100.00%		180076096503	0
0	100.00%		lucky13een	1
		1	100.00%	0.99
0	100.00%		110078400591	0
0	100.00%		redskinlvr	1
		1	100.00%	35.99
0	100.00%		110079079695	0
0	100.00%		singer89	1
		1	100.00%	37.0
0	100.00%		130068726011	0
0	100.00%		cheesewhiz	1
		1	100.00%	5.25
0	100.00%		190072550491	0
0	100.00%		shorty07	1
		1	100.00%	57.0

Appendix B: Sample Proxy Bid vs. Final Value Output

Status	userID	itemID	finalValue	bidPlaced	diff	catID
WINNER	pauldome	200068550402	9.99	9.99	0.00	3225
WINNER	mr.murray	170070289884	10.38	16.83	6.45	86958
ODDITY	drpezy154	160075333229	9.95	15.00	5.05	16285
WINNER	weisigns	160073319484	0.99	2.00	1.01	146390
LOSER	wmf723	290073385211	151.09	75.00	-76.09	85805
WINNER	yellowfish	220072158175	4.99	5.50	0.51	16476
LOSER	kle8763	130068754060	107.51	10.99	-96.52	134653
LOSER	kle8763	130068754060	107.51	11.99	-95.52	134653
LOSER	cloudlessky	170069925982	43.99	41.66	-2.33	1288
LOSER	liteflyer06	260075177229	27.55	5.00	-22.55	20756
LOSER	bob315751	260075177229	27.55	11.00	-16.55	20756
LOSER	evu348	120075084111	380.00	270.00	-110.00	62054
LOSER	evu348	120075084111	380.00	295.00	-85.00	62054
LOSER	evu348	120075084111	380.00	310.00	-70.00	62054
LOSER	lucky13een	180076096503	57.00	0.99	-56.01	289
LOSER	redskinlvr	110078400591	42.00	35.99	-6.01	64462
LOSER	singer89	110079079695	42.95	37.00	-5.95	16489
LOSER	cheesewhiz	130068726011	40.00	5.25	-34.75	39462
LOSER	shorty07	190072550491	150.00	57.00	-93.00	3697

Appendix C: Sample Volume Reputation Output

bargainland - 02-07-2007 16.51

Official eBay
Feedback Score: 191520
Percent Positive: 90.00%
User does not have protection from PayPal

Repeat User
Positive Count: 197
Neutral Count: 1
Negative Count: 2
Total Count: 200
"My" Feedback Score: 195

Positive Percentage: 98.5000%
Neutral Percentage: 0.5000%
Negative Percentage: 1.0000%

Volume With Repeat User
Positive Value: 3806.69
Neutral Value: 192.49
Negative Value: 26.99
Total Value: 4026.17
Feedback Value: 3779.70

Positive percentage: 94.5487%
Neutral percentage: 4.7810%
Negative percentage: 0.6704%
Ratio of Pos Value to Neg Value: 141.0407558354946

Min of Positive Values: 0.99
Min of Neutral Values: 192.49
Min of Negative Values: 0.99
Min of Total Values: 0.99

Max of Positive Values: 401.00
Max of Neutral Values: 192.49
Max of Negative Values: 26.00
Max of Total Values: 401.00

Median of Positive Values: 6.50
Median of Neutral Values: 192.49
Median of Negative Values: 13.50
Median of Total Values: 6.50

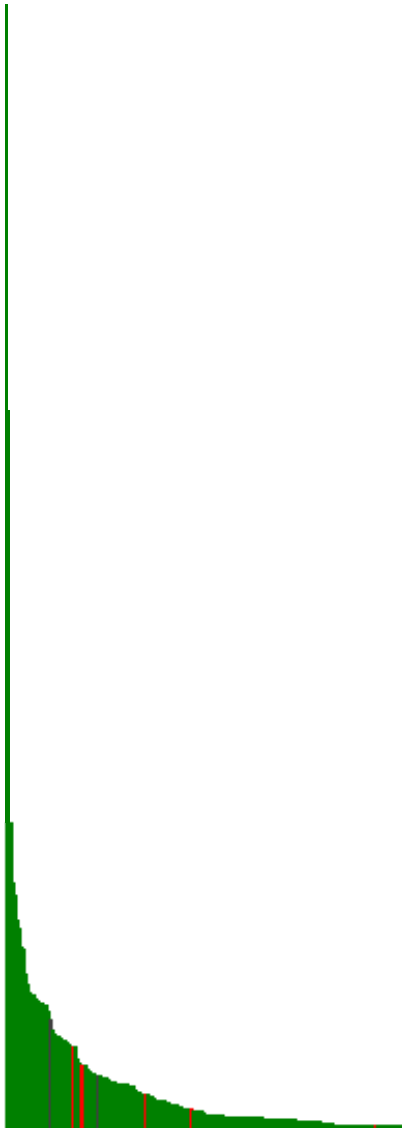
Average of Positive Values: 19.32
Average of Neutral Values: 192.49
Average of Negative Values: 13.50
Average of Total Values: 20.13

Ratio of Pos Median to Pos Average: 0.34
Ratio of Total Median to Total Average: 0.32

Ratio of Pos Median to Total Median: 1.00
Ratio of Pos Average to Total Average: 0.96

Summary
Positive Count Percentage (repeat): 98.5000%
Positive Volume Percentage (repeat): 94.5487%
Bad Count: 0

Appendix D: Sample Long Tail Image



Appendix E: Sample Greasemonkey Output

Get Results Page

[All Items](#) | [Auctions](#) | [Buy It Now](#)

All Categories [Advanced Search](#)






Search title and description
 Related Searches: [ipod nano](#), [apple ipod](#), [ipod video](#), [ipod 30gb](#), [ipod mini](#)

Matching Categories
[Consumer](#)
[Electronics](#) (99445)
 MP3 Accessories (81110)
 Apple iPod, MP3 Players (10049)
 Car Electronics (1480)
 more...
[Computers & Networking](#) (1183)
 Desktop & Laptop Accessories (283)
 Software (173)
 Apple, Macintosh Computers (109)
 more...
[Cell Phones & PDAs](#) (1148)
 Accessories, Parts (1402)
 PDAs & Pocket PCs (67)
 Bluetooth Wireless Accessories (23)
 more...
[See all categories ...](#)

Related Guides
[In-Dash Car Receivers](#)
[Car Amplifiers](#)
[Car Audio and Car Video](#)
[Car Speakers & Car Subwoofers](#)
[Car Video](#)
[Home Audio](#)













106384 items found for ipod
 Get it new. Buy it now: 19114 [ipod](#) items on eBay Express!

List View | [Picture Gallery](#)

	Item Title	PayPal	Bids	Price	Shipping to 01609, USA
<input type="checkbox"/>	 iPOD PROFESSIONAL CAR RADIO FM TRANSMITTER & CHARGER UK Feedback: (3550) 100%	<input type="checkbox"/>	-	GBP 4.99	GBP 8.89
<input type="checkbox"/>	 iPod Video 30 60 Photo TV AV RCA Cable+Car+Wall Charger Feedback: (70546) 98.5%	<input type="checkbox"/>	-	\$6.99 =Buy It Now \$7.99	\$5.95
<input type="checkbox"/>	 Portable Dock Speaker for iPod Mini Video MP3 MP4 White Feedback: (20399) 97.6%	<input type="checkbox"/>	-	AU \$9.99 =Buy It Now AU \$13.99	AU \$29.99
<input type="checkbox"/>	 Apple iPod Wireless Remote Control Photo Video Nano Feedback: (8310) 98.6%	<input type="checkbox"/>	-	\$8.99 =Buy It Now \$10.99	\$5.99
<input type="checkbox"/>	 Griffin iClear Case for iPod nano 1GB 2GB 4GB 1 2 4 GB Feedback: (42240) 99.6%	<input type="checkbox"/>	-	\$4.99 =Buy It Now	\$4.99

Compare Items Page

Showing 1-4 of 4 items Previous 4 | Next 4 [Show All Items](#)

Compare Items				
<input type="button" value="Watch All"/> <input type="button" value="Remove All"/>		Sort by: Order Selected: recent first		
Item	Remove Item	Remove Item	Remove Item	Remove Item
	 iPOD PROFESSIONAL CAR RADIO FM TRANSMITTER & CHARGER UK	 3.5mm Stereo Earphones Headphones for iPod MP3 Player	 Apple iPod Mini Portable Speaker Nano Video Photo Color	 Portable Dock Speaker for iPod Mini Video MP3 MP4 White
	Watch this Item	Watch this Item	Watch this Item	Watch this Item
Time Left	Ended	Ended	Ended	Ended
Bids	0 bids	0 bids	1 bid	0 bids
Seller	kmsdirect (3552 ★) 100% Positive 	discounttown (61000 ★) 97.8% Positive 	mlaccessories (1877 ★) 97.9% Positive 	buonshopping (20403 ★) 97.6% Positive 
Repech (last 50)				
Price	GBP 4.99 (Approximately US \$9.79)	US \$1.99	US \$4.99	AU \$9.99 (Approximately US \$7.83) AU \$13.99 <i>=Buy It Now</i> (Approximately US \$10.96)

View Item Page

Apple 2 GB iPod Nano Silver 2nd Generation mp3 player		Item number: 290083407917
Bidder or seller of this item? Sign in for your status		Watch this item in My eBay Email to a friend
 View larger picture	Current bid: US \$31.00 <input type="button" value="Place Bid >"/> End time: 43 mins 5 secs (Feb-15-07 15:40:44 PST) Shipping costs: US \$18.49 UPS Ground Service to United States (more services) Ships to: Worldwide Item location: Old Greenwich, United States History: 9 bids High bidder: jay51669 (11 ★)	Meet the seller Seller: auctionalliesgreenwich (2129 ★)  Feedback:  95.8% Positive Member: since Jun-15-02 in United States <ul style="list-style-type: none"> ■ Read feedback comments ■ Ask seller a question ■ Add to Favorite Sellers ■ View seller's other items: Store List ■ Visit seller's Store: Auction Allies in Greenwich
Listing and payment details: Show/Show		Buy safely <ol style="list-style-type: none"> 1. Check the seller's reputation Score: 2129 95.8% Positive Read feedback comments 2. Check how you're protected If paying with a credit card, see your card issuer's terms Returns: Seller accepts returns. 3 Days Money Back