From Zero to Hero

Identifying Vendor Characteristics that Impact Vendor Performance on Darknet Markets

From Zero To Hero
Identifying Vendor Characteristics that Impact Vendor Performance on Darknet Markets

Master thesis submitted to Delft University of Technology in partial fulfilment of the requirements for the degree of

Master of Science
in
Complex Systems Engineering and Management

by

Valentijn Frederik Grapperhaus
Student Number: 4250338

To be defended publicly on Wednesday September 25, 2019 at 09:00 AM.

Project duration:	February 25, 2019 – September 25, 2019	
Thesis committee:	Prof. dr. M. J. G. van Eeten,	TU Delft, Chair
	R. S. van Wegberg,	TU Delft, First Supervisor
	Prof. dr. P. H. Hartel,	TU Delft, Second Supervisor
	dr. G. J. van Hardeveld,	FIOD, External Supervisor

An electronic version of this thesis is available at http://repository.tudelft.nl/

TU Delft

Preface

Dear reader,

During the past 7 months I experienced the process of doing research for a master's degree in Complex Systems Engineering and Management at the TU Delft. At times, the process of doing research and writing a master thesis, felt like doing two steps forward and one step back, on repeat.

This thesis also marks the end of an academic career that I commenced with the BSc programme of Systems Engineering Policy Analysis and Management at the TPM faculty, 7 years ago. Therefore, I would like to thank all the people from the TPM faculty that I met during my years in Delft. In particular, I would like to thank Dr.ir. Tineke Ruijgh-van der Ploeg for supporting me in the year after my mother passed away. Tineke provided me with guidance and inspiration for completing the final year of my Bachelor.

With regards to this research, I would like to begin with thanking my entire graduation committee. I would like to thank Prof.dr. Michel van Eeten for being the chair of my graduation comittee, his patience, his straightforward and valuable feedback, which helped lay the foundations for this research.
Also, I would like to thank Prof.dr. Pieter Hartel, my second supervisor, for making time and providing me with new perspectives in the light of this research.
Furthermore, I would like to thank Rolf van Wegberg, my daily and first supervisor, for guiding me through this process. Rolf was always around, and his enthusiasm and optimism gave me the strength and willingness to persevere in this research, even during its darkest moments.
Especially, I'm thankful to the the final member of the graduation committee, dr. Gert Jan van Hardeveld as external supervisor at the FIOD, for guiding, teaching, and supporting me through this entire process.

This research was not possible without the FIOD, and the Financial Advanced Cyber Team (FACT) in particular. Everybody from FACT immediately gave me the feeling of being a full member of the team.

Finally, I would like to thank all my friends and family for their love, support, and patience over these past years.

V.F. Grapperhaus
Delft, September 2019

Table of Contents

List of Figures

List of Tables

1 Introduction

The first online drug trade was made around 1970 between students from of two American universities, using Arpanet, the precursor to Internet (Markoff, 2005; Buxton and Bingham, 2015). While drugs were available online in small quantities ever since the beginning of the internet, only in recent years drugs and other illegal goods became easy to order from the web. This is mainly due to increasing availability of internet, new communication technologies and state-of-the-art encryption techniques.

The phenomenon in the cyberspace of online markets facilitating the trade of illegal goods are often referred to as darknet markets. The first-ever darknet market to gain worldwide attention was Silk Road 1.0, which opened in the beginning of 2011 (Commission and Europol, 2017). After the take-down of Silk Road 1.0 at the end of 2013, numerous of other darknet markets emerged. To this day, more than 100 darknet markets have existed (Aldridge and Décary-Hétu, 2016). These darknet markets have great similarities with earlier versions of ordinary online markets for licit goods such as eBay and Amazon: the interface is user-friendly and straight to the point, as the example in Figure 1 shows (Aldridge and Décary-Hétu, 2014;Paquet-Clouston, Décary-Hétu and Morselli, 2018).

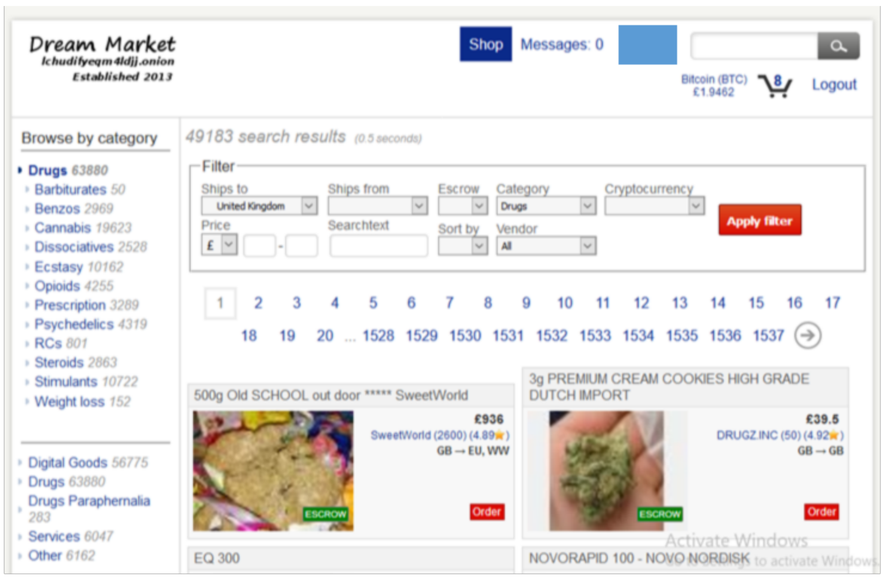


Figure 1: Interface of Dream Market

On darknet markets, users are either active as buyers of illegal goods, or as vendors that engage in selling these goods. Vendors are not much different from drug dealers in an offline environment, in the sense that they want to avoid detection, and minimize the risk of getting caught by law enforcement (Décary-Hétu and Leppänen, 2016). A big difference however, is that ordinary criminals need to be physically present in an offline environment in order to trade illegal goods: they are bound to a physical location and in most cases one transaction a time (Commission and Europol, 2017). While for online markets a physical infrastructure is not needed, and therefore vendors can easily migrate their businesses to other markets, together with their reputation. Also, this limitless digital infrastructure enables the possibility of engaging in doing transactions anywhere, anytime and even doing them simultaneous. For these reasons, darknet markets are a welcome addition, alternative, or both, for criminals in comparison to the physical market.

These features of darknet markets have made it extremely difficult for law enforcement to successfully intervene. And even when an intervention was successful, the impact on the online trade of illegal goods by taking down a darknet market is often of limited duration: revealing that the demand for these illegal goods through these platforms is significant (Décary-Hétu and Giommoni, 2017; Soska and Christin, 2015).
Because of these difficulties, law enforcement agencies are in need of novel strategies for interventions of darknet markets with a more sophisticated approach. An alternative way of inflicting damage to the ecosystem could be to target specific types of vendors on these darknet markets, i.e. vendors with the biggest market share. However, law enforcement agencies are better of with a more thorough understanding on how vendors operate, so that they can optimize their strategies. Through which characteristics can vendors be distinguished and how do vendors use these characteristics to make money or become important in the ecosystem of darknet markets? These are the questions that should be answered when the aim is to gain a better understanding in the modus operandi of vnedors. Moreover these insights can turn out to be valuable for law enforcement agencies, using them to improve their intervention strategies.

1.1 Research Objective

This study aims to assess the entire market activity over time and analyze the vendors that come and go during the lifetime of a darknet market. This will enable the identification of different vendor characteristics, and types of vendors, and how they are related to vendor performance on darknet markets. Moreover, it is examined whether the importance of certain characteristics for performance changes over time, i.e. due to a maturing market or changing market environment. The ultimate goal is to provide insights in what characteristics are important for vendors to generate revenue and sales. A secondary objective is to translate these findings to meaningful insights for law enforcement, that can be used to improve vendor targeted intervention strategies.

1.2 Contributions to Research

This research contributes to literature on competition on darknet markets and on different selling strategies that are adopted by vendors. Moreover, it is the first time that a darknet market analyzed from beginning to end. Another contribution is the fact that the back end data of this market is available for analysis: this enables the inclusion of more meaningful metrics to determine what impacts revenue and sales generated by vendors. Altogether, this research offers the unique opportunity to analyze a darknet market's activity over time and assess what indicators are for vendor performance.

1.3 Research Scope

Darknet markets facilitate the trade of a wide range of illegal goods. While there are darknet markts that only focus on certain segments, the darknet market of which the data for this research was collected offered trade opportunities in more than 10 different segments. Ranging from Drugs to Erotica and Jewellery. However, as found by earlier research, the market share of Drugs is by far the biggest on most darknet market. This also goes for our data. Therefore, the decision is made to focus on the trade in the 'Drugs' segment.

1.4 Research Question and Approach

To realize the above mentioned aims and contributions the following research question is defined:

To what extent do vendor characteristics impact vendor performance on darknet markets?

In order to answer the main research question a set of five sub-research questions are formulated:

1. What metrics are considered to indicate vendor performance on darknet markets?
2. What vendor characteristics are considered to have an effect on performance and to what degree are they present on a darknet market?
3. What profiles of vendors can be identified, and how do these profiles perform relative to each other?
4. What are the effects of vendor characteristics and vendor profiles on vendor performance?
5. Do these effects change when the market grows and matures, and how do they change?

The remainder of this thesis consists of literature reviews in Chapters 2 and 3. Chapter 4 outlines the methodology for performing this research. In chapter 5, the data is prepared for analysis, step by step. in chapter 6, 7 and 8 the different analyses are performed. The results, limitations are discussed in chapter 9, together with the conclusion, recommendations and perspectives for law enforcement agencies.

1.5 Linkage with the CoSEM programme

The research objective of determining to what degree vendor performance on darknet markets is influenced by vendor characteristics, follows the principles of the CoSEM program. The objective not only examines how actor behaviour within a socio-technical system can result in different outcomes. But also, how the system itself changes. Moreover this research includes a big technical component, namely understanding the driving forces of a darknet market. Together with examining how certain characteristics play a role in the socio-technical environment of a darknet market, it offers law enforcement agencies and science a new view on the socio-economic relations that are formed in the system. Altogether, the objective of this research is about getting a better understanding of the market dynamics of darknet markets, and on how the modus operandi of vendors influences their performance.

2 The Phenomenon of Darknet Markets

Criminal networks are typically characterized by the fact that they are organized underground and operate outside of the law (Morselli, 2009). Most criminal networks aim to pursue relationships with the outside and legal world, such as companies and governments. In doing so they undermine the goals of a just society in different ways (Duijn and Klerks, 2014; United Nations Office on Drugs and Crime, 2010). For this reason, law enforcement agencies desire intervention strategies that target criminal networks more effectively (Duijn and Klerks, 2014). This chapter discusses characteristics and the processes of *darknet markets*, that are in place to enable users to transact and interact with each other. Also, the underlying mechanisms that make the concept of a darknet market work, such as trust, anonymity, and security are discussed.

2.1 Darknet Markets in General

A darknet market can be viewed as an online criminal network in such a way that it is organized outside the scope of the ordinary internet and thereby escapes the need of compliance. Together with the lack of a judicial system targeting these illegal activities, a darknet market functions as a trading platform for illicit goods. Their users can either be identified as vendors who engage in selling goods, or as buyers who engage in buying goods. As users interact with each other by performing transactions, a network is formed. A dark market makes it possible for vendors to sell their goods anytime and anywhere: there are no geographical boundaries and time zones that limit trade opportunities (Christin, 2012; Décary-Hétu and Leppänen, 2016).

2.2 Trust, Anonimity and Security

The reason for darknet markets being a popular online platform for vendors and buyers that want to trade illegal goods, is mostly due to its three characteristics. Darknet markets facilitate and ensure anonymity, security, and trust at the same time (Verburgh, Smits and van Wegberg, 2018; Aldridge and Décary-Hétu, 2016).
For one, darknet markets are hosted on, and only accessible through, the Tor network: making the activity of its users untraceable, providing security and anonymity (McCoy, Bauer, Grunwald, Kohno and Sicker, 2008; Hout and Bingham, 2013).
Secondly, darknet markets often provide means of encrypted messaging. An example of this is he pretty-good-privacy (PGP) encryption protocol for person to person communication (Tzanetakis, Kamphausen, Werse and von Laufenberg, 2016a,Kruithof, Aldridge, Hétu et al., 2016), facilitating a completely secure communication between counter-parties. Everyone can encrypt a message, so that only the receiver can decrypt the message. Users of this encrypytion protocol often share their public PGP keys, which in turn also function as a valuable alternative for verifying identities (Broséus et al., 2016). This comes at hand, for example, when a buyer needs to leave his home address for the illegal goods he bought on the darknet market.
Third, darknet markets require its users to pay with cryptocurrencies. Cryptocurrenies enable direct quasi-anonymous transactions: information on the transactions are publicly accessible and traceable through the bitcoin blockchain (Tzanetakis et al., 2016a). As long as the addresses are used anonymous, these transactions are not traceable to real-world identities (Aldridge and Décary-Hétu, 2014). Despite this quasi-anonymous characteristic, Bitcoin remains the number one used cryptocurrency by darknet markets and vendors. An additional feature of bitcoin is the escrow feature, which is also implemented by most darknet markets. The escrow feature counters fraudsters by involving the darknet market in the transaction as a third party: at least two of the three parties need to agree on the transaction to be finalized (Verburgh et al., 2018).
On darknet markets, trust is further facilitated by the integration of a refined feedback systems. Buyers can rate vendors based on the satisfaction over the purchased goods and the delivery method. Moreover, most markets facilitate the rating of buyers as well: vendors can rate buyers for the same

reasons as in the opposite case. This feedback is visible to all users on the market. A vendor's rating therefore functions as guidance for buyers to decide on which vendor is to be trusted(Kruithof, Aldridge, Décary-Hétu et al., 2016Verburgh et al., 2018). Figure 2 gives an example of an vendor profile, where a vendors rating is clearly visible. In turn, this encourages vendors to up their game, and keep delivering, in order to improve their reputation.

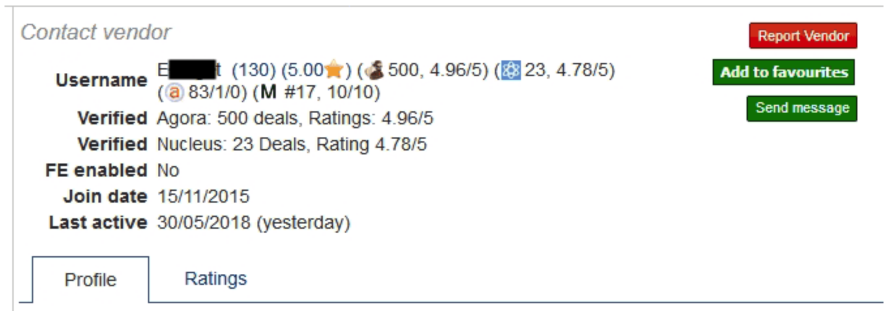


Figure 2: A Vendor Profile on Dream Market

2.3 Intervening in Darknet Markets

Darknet markets have proved themselves to be resilient against law enforcement intervention strategies (Christin, 2012; Soska and Christin, 2015; Décary-Hétu and Giommoni, 2017). Christin (2012) names four kind of strategies practiced by law enforcement agencies to counter dark market networks: attacking the network; targeting the financial infrastructure; targeting the delivery model; or leave the market be. However, in a later paper Soska & Christin (2015) state that interventions that target key vendors might have more impact altogether and suggest future research to look into it. This is reinforced by Duxbury & Haynie (2018), who also propose research that evaluates the darknet market structure after important vendors are removed.
Nevertheless, the activity on dark markets seem to have grown continuously over the recent years at a constant rate, diminishing the different intervention attempts from law enforcement in the past (Kruithof, Aldridge, Décary-Hétu et al., 2016; Décary-Hétu and Giommoni, 2017). According to Décary-Hétu & Quessy-Doré (2017) the trading volume on dark markets is likely to grow bigger and bigger, which must be considered by governments and law enforcement when developing new intervention strategies. Future investigations by law enforcement are likely to become more difficult, due to the platforms adapting more and evolving rapidly (Paquet-Clouston et al., 2018).

3 Discerning Darknet Vendors

3.1 Vendor Performance

The question whether a vendor performs well depends on the unit of measurement. Most commonly, market share is taken as performance indicator as it reflects the relative performance of a vendor. Market share can be defined by different metrics. For example, in some markets it is more valuable to measure market share in terms of the number of clients, while in other markets the biggest share in terms of sales is more important.
In research on darknet markets, the performance of vendors is most commonly measured by sales volume. However, by only looking at sales, it is possible that vendors with less sales but a lot revenue are not considered to be performing well.The reason that vendor performance on darknet markets is measured as a total amount of sales in most literature, is due to the fact that sales result in feedback, which again results in reputation and trust. Therefore, a vendor that handles more transactions, achieves a higher reputation and gains a trustworthy position within the community.

3.2 Darknet Vendor Characteristics

The large amount of vendors active on darknet markets compete with each other for a bigger market share. This section gives a brief summary on the characteristics of vendors that are known to be related with the performance of these vendors. These characteristics were also used in other studies to examine and analyze market performance and buyer-vendor relationships. The available data from darknet market X makes it possible to construct characteristics that were not used before or only in other fields of research.

Experience
The lifetime of a vendor on a dark market can be defined as experience (Décary-Hétu and Quessy-Doré, 2017). Moreover, research found that vendors that were active less frequently for longer periods of time on dark markets experienced more performance in business than vendors that were frequently active but for a shorter period of time (Décary-Hétu and Leppänen, 2016). More general research in online crimnial networks confirmed this, stating that experience is an important factor for performance (Morselli, 2009; Bouchard and Nguyen, 2011).

Reputation
Another valuable factor for performance is trust. Trust is obtained through reputation, which is embedded in the process of making purchases by vendors. Buyers are always asked to give feedback on the vendor after they received their purchased goods. This functions as a sort of quality control (Soska and Christin, 2015; Décary-Hétu and Quessy-Doré, 2017). Because reputation grows bigger overtime, it can become more difficult for new vendors to gain market share. Vendors that are new to the market are always 'two steps' behind the already established vendors. According to Duxbury and Haynie (2018), this indicates that the structure of dark markets becomes more centralized over time, which results in fewer vendors being responsible for most transactions. Moreover, when a vendor establishes the reputation of being reliable, buyers have found to be less precarious doing transactions with them, resulting in transaction being finalized earlier than than would be the case with a vendor lacking this reputation (Tzanetakis, Kamphausen, Werse and von Laufenberg, 2016b).

Loyalty
This resulting behaviour of buyers is translated to loyalty of buyers to vendors. Loyalty was first used to analyze vendor-buyer relations on darknet markets by Décary-Hétu and Quessy-Doré (2017) and is based on the idea that making purchases from the same vendor on different moments in time results into loyalty. Loyalty is thus strongly dependent on trust. While vendors had gained more trust from returning buyers, or repeat buyers as they are called in literature, potential new buyers are more

inclined to buy from the trusted vendor (Tzanetakis et al., 2016b). Moreover, the more transactions are completed and the more pleasant the experience around the transaction is, the greater the chance that the buyer returns to the vendor for another purchase (Hout and Bingham, 2013).
Loyalty is not to be confused with exclusivity: a buyer can still be loyal to a vendor when he, for a smaller share, also trades with other vendors (Neal, 1999), where exclusivity indicates that a buyer trades with just a single vendor (Décary-Hétu and Quessy-Doré, 2017).

Exposure

The term 'exposure' is used to describe the visibility of a vendor on a darknet market. It almost goes without saying that for an vendor to improve his visibility on the darknet market, the first thing he can do is offer more goods on the platform through listings. By having listings on the platform, the vendor becomes visible to other actors on the market. Exposure is not to be confused with diversity: it doesn't necessarily mean that the vendor has a wide range of different products in his inventory. It even is common that a vendor posts multiple listings with different quantities for the same product (Paquet-Clouston et al., 2018), these are also referred to as clone listings and are created with the purpose of offering different quantities of the same product. This is likely to target different kind of customers, i.e. buyers that are only interested in goods for self-consumption, or buyers that are actually sellers themselves and want to buy in bulk quantities (Broséus et al., 2016; Aldridge and Décary-Hétu, 2016). Previous research on darknet markets use the amount of listings of a vendor as a proxy for the marketing of their goods (Paquet-Clouston et al., 2018; Décary-Hétu and Quessy-Doré, 2017). However, a vendors exposure on darknet markets is more important than on traditional on-line markets. Where communication outside 'advertisements' on traditional markets is possible through numerous possibilities, communication between vendors and buyers is drastically limited on most darknet markets.

Diversity

The definition of the diversity characteristic is in itself clear: the degree of different product categories a vendor is offering goods for sale. Often, vendors on darknet markets are either specialized in a certain product category, or they sell a wide range of goods Broséus et al., 2016.
Diversity is defined in different ways in literature. Pacquet-Clouston et al. (2018) measure diversity according the Diversity Index (Agresti and Agresti, 1978). Diversity in this sense, is seen as the distribution of the amount of listings in different categories, i.e. when a vendor has 2 listings in category 'y' and 2 listings in category 'z' the *Diversity Index* is 1.0: the vendor is perfectly diversified. While when a vendor only has listings in a single product category, his diversity score is 0.
Rolf et al. (2019) measure diversity in a different way: by constructing a binary variable that is 1 if a vendor is active in more than one category, and 0 if a vendor is active in a single category.

Forum Activity

Another important characteristic is overlooked by many researchers: a vendor's forum activity. Forums enable vendors to interact with potential buyers in a different way than on the marketplaces; e.g. sharing their experiences, promoting their goods, or engaging with the community. The forum can be seen as an additional opportunity for vendors to expand their exposure and reputation: by interacting with the community, more exposure is achieved which in turn influences reputation (Tzanetakis et al., 2016a). Moreover, it is often the case that the aliases of vendors on forums are the same as on the darknet markets.

Multihoming

The last vendor characteristic that this research includes is Multihoming. In the context of darknet markets, multihoming describes the phenomenon of vendors being active on various darknet markets at the same time (T. Calis, 2018). Soska and Christin (2015) state that it is actually preferable for a vendor to run business on multiple markets at the same time, so that when one market goes down, the business on the other market can seamlessly continue. The benefits of multi-homing were confirmed

by research that found that multi-homing vendors generated more sales relative to other vendors after the takedown of a darknet market (T. Calis, 2018).
Due to the ease of copy-pasting listings to other markets and the possibility of vendors to import their rating from one darknet market to another, vendors' reputation is preserved. It is estimated that around 25% of the vendors are active on multiple darknet markets Soska and Christin, 2015.

4 Methodology

To what extent do vendor characteristics impact vendor performance on darknet markets?

As the objective of this research is to provide an answer to the question whether vendor characteristics impact vendor performance on darknet markets, this chapter outlines the different steps necessary. First, this chapter elaborates on the research approach and the design of the conceptual model, which both form the basis for this research. Secondly, the process for preparing the data for analyses is described. Third, three methods for analyses are proposed, substantiated and justified. The last section of this chapter elaborates on how these methods provide a solid approach to analyse the concepts defined in the conceptual model, and how a conclusion can be drawn, taking all aforementioned into account.

4.1 Research Approach

4.1.1 Conceptual Model

Vendor characteristics have an effect on the market performance of vendors, as discussed in section 3. Taking these vendor characteristics into consideration, it becomes clear that there is a clear distinction to be made. Some characteristics are likely to have a direct impact on vendor performance, while others are affected by vendor performance and at the same time influence vendor performance: a lagged relation exists between these particular characteristics and vendor performance. The vendor characteristics that are assumed to have a 'direct' effect on vendor performance can be referred to as *active* characteristics. This is due to the fact that the characteristics are modifiable by vendors themselves. In Figure 3 the causality between vendor characteristics and vendor performance is illustrated. The outgoing arrow from the passive characteristics to performance, represents the lagged effect, or feedback loop.

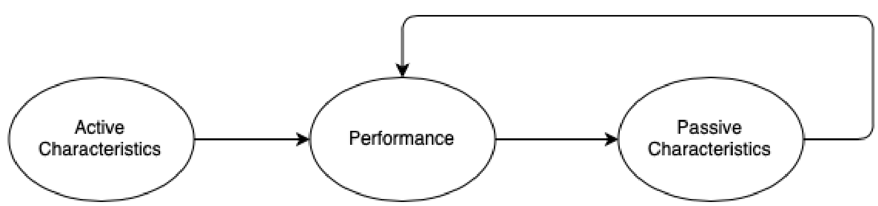


Figure 3: Conceptual Model of Vendor Performance I

However, due to the market growing over time and the possibility of changing market dynamics, it is plausible to assume that there could be changes in the effects of *active* characteristics, and *passive* characteristics, on vendor performance. For example, a certain *active* characteristic could have a strong impact on performance in a early stage of the market, while this effect weakens in later stages. As for the *passive* characteristics it could be possible that the lagged effects on performance are stronger in a mature market, than in a market that still has to *lift off.* Because, in the former case, there are established vendors, likely having a trusted reputation build on many feedbacks, which makes the feedback loop even stronger, making it for new or upcoming vendors impossible to break the wheel. However, vendors exhibit differences in active characteristics, resulting into numerous different compositions of active characteristics per vendor. These different compositions of active characteristics will be conceptualized as vendor profiles. Figure 4 shows transformation of active characteristics into vendor profiles. In the causal model of Figure 3, the active characteristics are individually related to vendor performance. While in Figure 4 the the different compositions, thus vendor profiles, are related to performance.

The feedback loop between *passive* characteristics and performance is depicted by two relations. First, it's important to understand that reputation and loyalty only comes to existence after a first transaction is completed: When an order is finalized a sell is made and revenue is generated. Then, the second relationship comes into play, from *passive* characteristics to performance: a higher rating or lower loyalty level is expected to impact performance.

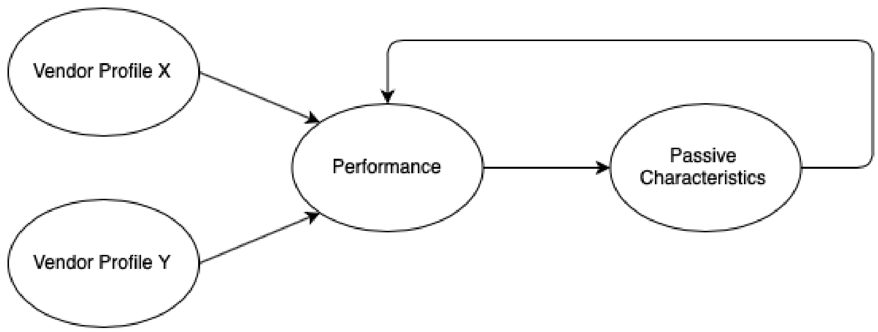


Figure 4: Conceptual Model of Vendor Performance II

This research will examine and analyze the different causal relationsips. First, the different compositions of active characteristics present among vendors are examined and interpreted with a clustering analyses. The clustering analysis aims to find what profiles follow from the characteristics that are extracted from the data. The relations between vendor performance and the active characteristics and vendor profiles is analysed through several multiple regressions. With these analysis the strength of *active* characteristics individually, or composed as a profile, and *passive* characteristics on performance are examined. Finally, the arrow from performance to passive characteristics is taken into account during the last series of regression analysis: a multiple regression is performed for every month. These relationships focus on the relative impact of *active* and *passive* characteristics and how the strength of these impacts change over time, inherent to the changing or evolving market dynamics. The feedback loop in Figure 3 is included in this analyses by altering the *passive* characteristics into *lagged variables.*

4.2 Data

4.2.1 Collecting Darknet Market Data

Previous research on darknet markets made use of different methods to collect data, in order to perform analyses. Such methods are known as 'scraping' or 'crawling'. The biggest disadvantage of such methods is that it only extracts the information that is visible at the front end: likely only revealing the tip of the ice berg.

To get a good idea of the finalized transactions between vendors and buyers, crawlers scrape the feedback section of vendor's listings. Researchers have based their research on market performance of vendors of dark market on the 'publicly' available feedbacks from buyers on vendors. This is possible due to each feedback containing general but relevant information about the transaction: listing title; price of the listing; amount of bitcoin transacted; and the time of feedback. However, there are at least three caveats in analyzing a vendor's market performance based on the feedbacks.

The first caveat is that in general, feedbacks on a darknet market are publicly visible only for a limited time. Every dark market has its own set of rules for this: Some markets show the feedback from the moment it was posted with a maximum of one month back. However, most markets keep feedbacks visible for between three to six months. This limited visibility of feedbacks makes it difficult for researchers to reconstruct the complete network of transactions. Further, this complicates the process of doing longitudinal analysis which captures the growth of the market and changing market conditions.

Another caveat with deriving vendor performance from feedbacks, is the fact that a certain amount bitcoins was exchanged during the transaction. The exact time of transaction is not visible from the feedback and therefore the transaction value in USD is estimated, potentially with great inaccuracy due to the volatile value of BTC. Together with varying listing prices, making it more difficult for researchers or law enforcement to establish an accurate estimation of the US Dollars that were involved in a transaction. Or, even more important specifically for law enforcement: to predict the total revenue a vendor made.

The third caveat that concerns the assessment of vendor-buyer relations, is that on most darknet markets the buyers identity is made anonymous. At some point a certain darknet market[1] was considered giving the most complete information on vendor-buyer relationships(Décary-Hétu and Quessy-Doré, 2017). Still, the information that this market unveiled is fairly limited: Buyers were pseudo-anonymized by only showing the first and last character of their username, making it hard to determine the significance of the buyer-vendor relationship, e.g. determining whether they were returning buyers.

Leaving the notion of front end data behind, this research has the back end database of a particular darknet market at its disposal, which from now on is referred to as *market X*. The back end data offers an unique opportunity to visualize and analyze the market activity on *market X* over time and to provide more accurate insights in the why and how vendors generate sales and revenue.
The back end data of *market X* is accessible through a SQL-database structure. Numerous tables are present that structure user information, listing information, product categories, orders, transactions, feedbacks and many more. In order to perform meaningful analysis the data needs to be cleaned and prepared so that only the relevant data needed for this research is extracted and constructed.

[1] The name of the market was not given in this study

4.2.2 Preparing the Data for Analysis

To analyze the effects of vendor characteristics on performance, the previously introduced metrics in section 3 and section 4.1.1 need to be constructed from the back end data. Preparing the darknet market data in this way demands a process that deals with superfluous and missing data. On top of that, it needs to reconstruct the required data for analyses. Below, Figure 5 illustrates the entire process on how the data is prepared.

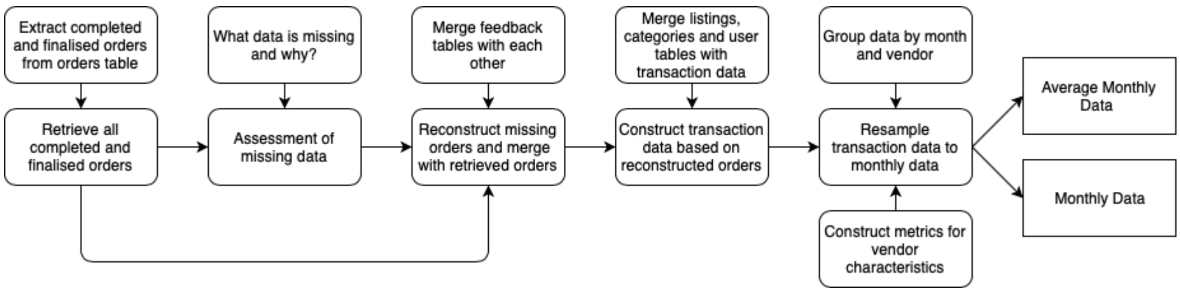


Figure 5: data preparation process

4.3 Cluster Analysis

Cluster analysis is seen as an important technique for exploratory and descriptive analysis of large data samples (Anil, 2008). Clustering is an unsupervised learning method: it identifies patterns in data, where these are not known on beforehand. By clustering data, samples are grouped together based on their similarity in measurement.

4.3.1 Aim

A cluster analysis is performed due to one of the aims of this research being to find groups of vendors with similar selling behaviour, based on their active characteristics. A segmentation of vendors provides insights in what is believed to be important for a vendor to gain performance or ensure business continuity. Vendors may vary in terms of how they present their products, or how much experience they need, to perform in terms of sales. Moreover, it is interesting to compare different vendor profiles and see if some profiles perform better or worse. The expectation is that there are a handful of distinct vendor profiles to be found, given the limited amount of active characteristics that are taken into consideration.

4.3.2 Choice of Clustering Method

In general, various clustering methods exist for the purpose of segmenting individuals based on their behaviour or features. Two of the most popular methods in this field are *hierarchical clustering* and *K-means clustering*. Hierarchical clustering methods are considered to form the basis of cluster analysis in general (Everitt, 2011), and are widely adopted in studies on customer segmentation, which identify distinct groups of customers (Chicco, Napoli and Piglione, 2006; Tripathi, Bhardwaj and E, 2018). Due to the relative equal nature of this analysis, the hierarchical clustering method is adopted. Another reason for adopting hierarchical clustering instead of k-means, is the fact that the number of k clusters are not chosen on before hand. Moreover, with hierarchical clustering the steps of the algorithm are visualized with a dendogram. In this way, the researcher can visually inspect where to cut the 'tree' and decide on the appropriate amount of clusters. The process is hierarchical because the clustered samples at level k remain also clustered at level k + 1 (Varghese, 2013). Figure 6 shows an hypothetical example of the hierarchical structure of vendor profiles.

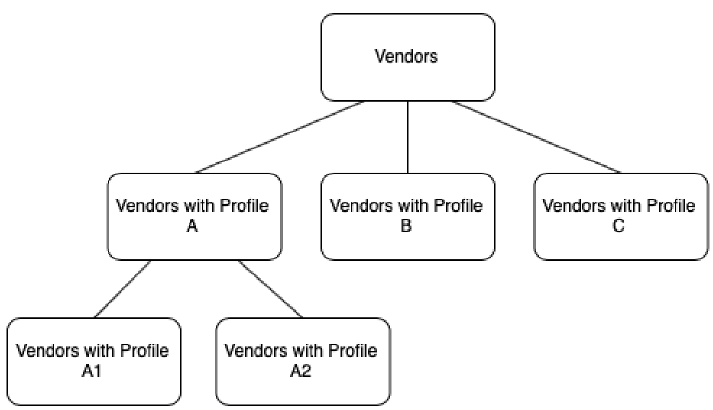


Figure 6: conceptual model of hierarchically structured vendor profiles

In general, there are two approaches of hierarchical clustering: First, there is the bottom-up approach, known as *agglomerative clustering*, where all samples begin in their own cluster, and ultimately end up together in a single cluster. Secondly, the top-down approach, known as *divisive clustering*. The divisive approach starts with a single cluster containing all individuals, and works down until all individuals reside in their own cluster. The crux of both approaches is that the researcher has to decide, depending on the preferred number of clusters, where the process of merging or separating the clusters should stop. The agglomerative approach is the most widely used of the hierarchical methods (Everitt, 2011), whereas the divisive approach is used less due to its complexity (Hexmoor, 2015; Wittek, 2014). Moreover, van Wegberg et al. (2019), performed an agglomerative hierarchical cluster analysis on this topic. For these reasons, the agglomerative approach is adopted.

4.3.3 Procedure

When performing a hierarchical agglomerative cluster analysis, the distance measure needs to be chosen first, which assesses the similarity between individuals on variables. The Euclidean distance is chosen. it defines the distance between subjects, based on the similarity in magnitude of the characteristics. However, due to different measurement scales, one variable could have more weight in the distance calculation than other variables. Therefore, the variables are min-max normalized so that the measurements are equally weighted. Then, the euclidean distance matrix is computed. In the context this research, it means that vendors are likely to be clustered, if they both score high on exposure and low on experience.
The following step is to define the linkage method that groups the instances based on their similarities and differences. These groups of paired individuals are linked to other groups of matched individuals to create bigger groups. Ward's method is chosen for the the linkage of these groups, because its objective is to minimize the growth of the sum of squares, when merging the different groups into bigger groups. This method of linkage also results in relative equally sized clusters, which is preferable for the aim of this analysis (Everitt, 2011).
The final step of the analysis is described as the following: The decision on the amount of clusters. Due to the fact that clustering is used for exploratory data analyses, there is not a single *right* amount of clusters. So it's up to the researcher to decide on the number of clusters.
After the clustering and interpretation of clusters is complete, the clusters will be compared with each other regarding their difference in performance and revenue. It is examined if particular clusters achieve higher performance than other clusters. The decision on what statistical test is appropriate for comparing the differences is made after the clusters are formed. Subsequently, a post-hoc analysis is performed to examine if the difference in performance between each pair of clusters is statistically significant.

4.3.4 Limitations

There are limitations associated with hierarchical clustering. One limitation is that due to the hierarchical nature, once samples are assigned to a cluster, they can not switch to other clusters further down the process: 'mistakes' can't be undone . Another limitation is that hierarchical clustering has a high time complexity. A last limitation is often mentioned in literature: there exists no statistic for determining the right amount of clusters: it is up to the performer, and his knowledge on the subject and data, to decide on the amount of clusters. Therefore, no single solution is correct.

4.3.5 Tools

The hierarchical agglomerative cluster analysis is performed with the *scipy* and *sklearn* packages, in *Python*. A variety of Python scripts that run agglomerative clustering algorithms can be found on the internet. However, due to multiple trial by error testing, customized scripts were written.

4.4 Regression Analysis

In a simple linear regression, the relationship between one independent variable and a dependent variable is inspected. However, not only the impact of the effect from the independent variable on the dependent variable is assessed, but also the strength and direction of the effect are examined. When more than one independent variables are included in the regression, the regression becomes a multiple regression analysis. In this research, several multiple regression analyses are performed.

4.4.1 Aim

The aim of this analysis is to determine the impact of vendor characteristics on performance. Also, this research examines whether vendor profiles have an effect on performance, and if profiles can be used to predict performance. However, the objective of this analysis is not to accurately predict or forecast vendor sales and revenue. Rather, it is to compare the relative impact of vendor characteristics on performance, and how that impact changes as the market evolves. The intended outcome is to describe what vendor characteristics are important for vendors to influence their performance.

4.4.2 Procedure

To achieve this objective, the analysis is divided into two parts. The first part, focuses on the relative impact strength of the vendor characteristics and vendor profiles on performance: to determine what active characteristic is most important for vendors to increase performance. This part of the analysis uses the 'vendor average' dataset as input. While the second part, examines how the strength of these impacts change over time. This second part uses the 'vendor monthly' dataset as input.

Measuring the Impact on Performance
The objective of the first part is to determine what characteristics contribute to performance and in what direction. This can be achieved by standardizing the coefficients of the predictors. In that way, the weights or impacts of the characteristics can be compared to each other in terms of influencing performance: sales or revenue in this case. These standardized variables are referred to as β coefficients. The β coefficients are the coefficients that are obtained in a regression when both the independent outcome, as well as the dependent variables, are transformed to standard scores, also known as Z-scores. In Figure 7, the first regression model is depicted. This model inspects the relations between vendor characteristics and performance. The second model (Figure 8) inspects relations of vendor profiles, and passive characteristics with performance. The reason for having two separate models is that the vendor profiles are based on the active characteristics.

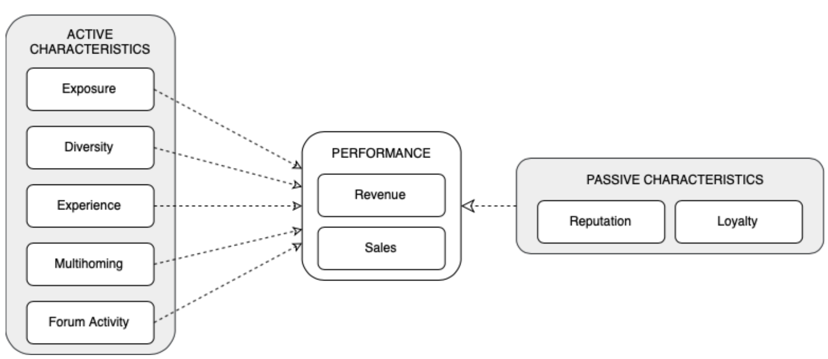


Figure 7: Regression Analysis Framework 1

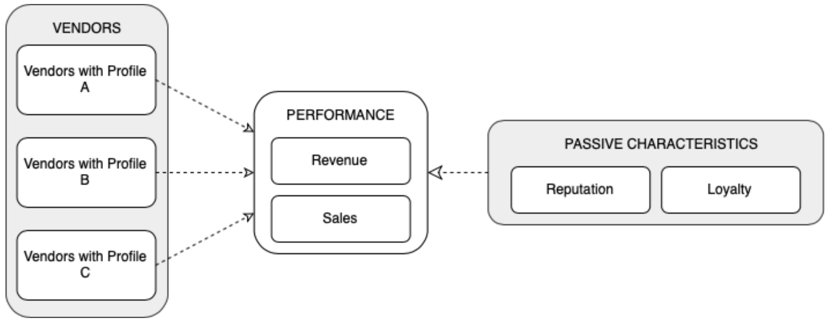


Figure 8: Regression Analysis Framework 2

Measuring the Change in Impact
In the second part, an important aspect of this analysis is the fact that time is taken into account. For every month a separate multiple regression is performed, with the focus of the analysis being the relative weights of the β coefficients. The standardized coefficients for the main effects are measured and plotted together for visualization and better interpretation. By looking at the relative impacts of *active* vendor characteristics at each time point, conclusions can be drawn without the risk of doing biased predictions through sub-optimally configured advanced regression models, e.g. multi-level regression analyses. In this way, the change in impact can be observed. Due to changing market conditions or a maturing darknet market, it is expected that the impact of some vendor characteristics also changes over time, either diminishing or strengthening. The framework from Figure 7) is adapted, but now the effect of time on the relations is the key point of interest. This is further emphasized by transforming the passive characteristics into lagged variables: Figure 3 already showed the causal relation directed from the passive characteristics to performance.

4.4.3 Limitations

The decision to perform a multiple regression for every month individually, was due to the many dependencies present in the monthly data, violating the assumption of independence of measurements. The measurements in the monthly data are correlated with each other on two levels. The first level of correlation is present at the subject level. The 'repeated' measurements of each subject are correlated with each other. And the second level of correlation is present between subjects measured at the same moment in time t. These subjects are correlated with each other in the fact that their measurements are partially characterized by time t. This form of correlation is known as auto-correlation, and is considered a problem when performing a regression.
Another limitation of the proposed approach is that by standardizing across individuals per point in time t, the mean levels of the outcome variable also changes per point in time t. This results in not knowing whether the raw-score is increased or decreased over time.
This is illustrated by the following example: the same impact is predicted for an in- or decrease of standard deviation for performance by a one unit increase of standard deviation of the mean of exposure in time t=1 and t=10. In t=1 a one unit increase of listings predicts a certain amount of performance, whereas in t=10 a ten unit increase of listings are needed for the same relative increase in performance. In this sense the relative impact on the dependent variable is equal, yet the raw scores indicate that more listings, and thus exposure, are needed to achieve this change in performance.

4.4.4 Tools

R statistical software was used to perform all the necessary tests and analyses. While *R* offers almost every possible statistical analysis through packages, the user interface is very basic. Therefore, *SPSS Statistics* was used as well to gain more insights in the results. However, *R* was used predominantly: the activities performed in *R* included testing the assumptions for regression, inspecting the relations between the independent- and outcome variables, testing whether the expansion of models was significantly 'better', preparing tables for regression analyses and finally structuring outcomes and preparing result tables.

5 Data Preparation

This chapter discusses what steps were taken to prepare the data for further analyses. The process as shown in Figure 5 is followed: first, available complete data is extracted from the database. Then, the necessary data for analyses that is missing is assessed. The missing orders are then reconstructed, based on the available feedbacks, and merged with the already available orders. Following this, the transaction level data is constructed by including information, such as the value of every transaciton in bitcoin, and if available, also in US Dollars. Finally, the data is resampled into monthly data, grouped by vendor ID's: as a result, a new dataset is formed that consists of vendors with monthly values that enable the construction of characteristics and calculation of performance values. When the characteristics are constructed per vendor the dataset is ready for analysis. This dataset functions as the main dataset for this research. All these steps are further elaborated on in the following subsections.

5.1 Retrieving Finalized Transactions

The back-end database consists of orders that have different *statuses* assigned. The lowest status concerns orders that are canceled. While the highest status are assigned to orders that are finalized. Between these statuses, other statuses exist that depict the current progress of an order. Altogether, there are only two statuses that depict whether an order is completed. Because the scope of this research only considers finalized transactions, or completed orders, only the orders with either one of these two statuses are extracted from the back end data.

5.2 Assessment of Missing Data

5.2.1 Self-Purging Database

When this research was initiated, the presumption was made that the data on *market X* was complete. However, after some time it became clear this was not the case. Not only was the database imported one or two times: this conclusion was drawn by the fact that at some point in time user IDs were re-assigned. But also, that a great amount of orders were missing from the database. In fact, it turned out that of all the orders that were *still* present in the data base, almost 50% consisted of either canceled orders or orders that did not even went through. The reason for this being that the database used a script for automatic data purging. Purging is different from archiving in the sense that when data is archived it is still retrievable but stored away, whereas purging means that the data is completely removed and not retrievable: It's gone for good. The motivation behind the purging is likely due to the data on transactions containing incriminating information. Order information together with the correspondence between the vendor and the buyer during the process of the order were purged 30 days after the order had an finalized or complete status. The purge script also deleted feedbacks after a minimum amount of time of 180 days, **only** if vendors enabled this option. Many vendors however did not use of this feature. Some vendors even decided to set the treshold to 360 days. A plausible explanation for this would be that vendors want to keep their feedbacks, as this is likely to have an reinforcing effect on their performance (see section 5.5.2.

5.2.2 Order Information

An important element of order information is the time of purchase and the time of finalization of the order. With this information accurate estimations for the USD value of orders can be made. However, for the missing orders that have been deleted by the database through the purge script, this information is not available anymore. Moreover, some orders only show the transacted value in BTC and leave out traditional currency values. This emphasizes the earlier mentioned problem caused by the purge script even more. In summary, there are two important elements of order information missing that are necessary to perform the analyses proposed in section 4.

1. Time of Purchase
2. USD Value of Transaction

5.2.3 Listing Information

In the case of listings, there are not necessarily any entries missing. However, for the characteristics of *exposure* and *diversity* it is important to accurately determine what the amount of active listings per month for every vendor is. The back end data imposes no clear boundary on whether listings are still active or not. Therefore, a method is required that only takes into account active listings.

5.3 Reconstructing Missing Data

5.3.1 Estimating the Time of Purchase

In order to reconstruct the transactions it is important that the time of purchase for orders/transactions where this information is missing is estimated. The estimation is based on examining the difference between time of feedback and time of purchase from orders where both information is present. However, there is a difference in the kind of feedback that could be present: the feedback from vendors on buyers follows a different distribution than the feedback from buyers on vendors. Evaluating these difference, the decision was made that for transactions where both the time of feedback by vendors and the time of purchase was missing, the time of purchase was estimated by subtracting the median time until feedback by buyers from the time of feedback by buyers value.

Feedback by Vendors on Buyers	
feedbacks	120542
mean	7 days 04:10:54
std	9 days 13:28:45
median	4 days 21:44:11

Feedback by Buyers on Vendors	
feedbacks	154906
mean	9 days 16:25:17
std	16 days 02:03:45
median	6 days 08:54:05

Table 1: Time between time of purchase and feedback

5.3.2 Estimating the USD Value of Transactions

Besides estimating the time of purchase, finding the right USD value for each transaction is a lot trickier. Not only due to having almost 50% of the time of purchase being estimated, but also because of the highly volatile value of BTC. Moreover, the market seemed to had implemented a BTC ticker that updated the USD/BTC per minute. In order to convert the value of BTC to USD, external data from https://bitcoincharts.com/ was extracted. This source provided price history data on BTC per minute, for the months the market was active. The price data from the ticker was adopted as it was considered accurately enough when comparing it with the transactions that showed the transacted value in USD, next to the time of transaction.

5.3.3 Determining whether an Listing is Active

When determining the levels of diversity and exposure for vendors, the problem arises that listings are present that are inactive, outdated, or deleted. Moreover, skewed comparisons are possible, when taking into consideration that some vendors might update their offered goods regularly, by creating new listings every month and deleting old ones. While other vendors keep their listing entries, and merely update their listings once in a while. The consequence being for example, that *vendor A* has more than 500 listings in total, and *vendor B* has 50 in total. While on average, both vendors have effectively 50 listings active every month. To counter this caveat listings are considered active in a month, if:

- The listing is created in OR before that month

Listings are not considered active in a month, if:

- The last purchase of that listing was before that month AND the listing is marked as deleted.

5.4 Reconstructing Transactions

The finalized orders, thus transactions, are reconstructed by merging both feedback tables and the table containing order information. The latter is filtered beforehand, so that only orders with a finalized or completed status are included. Following the steps described in section 5.2, a transaction-level dataset is constructed that can answer the following questions:

- When did the transaction take place?
- What item or product was transacted?
- What was the value of the transaction?
- Who participated in the transaction (vendor and buyer)?

In total, this reconstructed data consists of around 364.000 finalized transactions. However, according to the back end database, 362.000 transactions were finalized. This number is found by taking the sum of finalized orders that were saved for every vendor in a table of the database that contained user information. Altogether, this comes down to a little more than 0.5 percent of reconstructed transactions are overestimated.
Due to possible corrupted and incomplete data in the final stage of market X, transactions in this period are excluded from the dataset: as a consequence, around 328,000 transactions remain, and form the final dataset for this research. Figure 9 shows the remaining reconstructed transactions in blue, as compared to the transactions that were originally found in the table containing order information, in orange.

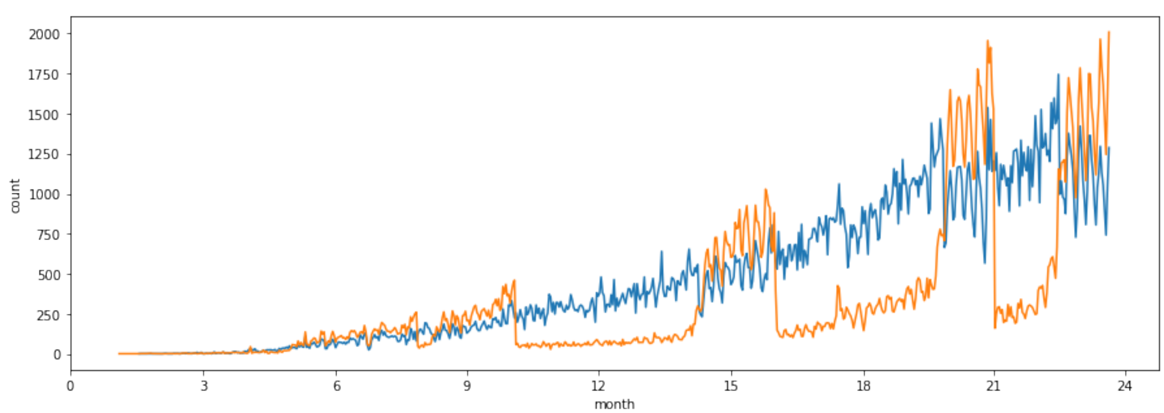


Figure 9: reconstructed transactions (blue) and finalized orders (orange)

5.5 Constructing Metrics for Characteristics and Performance

As reviewed in section 3.2, meaningful metrics must be constructed for measuring vendor characteristics, so that the relations discussed in section 4.1.1 can be properly assessed and examined in the different analyses. As the aim is also to perform analyses with a longitudinal nature, the characteristics are calculated on a monthly basis. The following paragraphs will briefly outline how for each characteristic a metric is constructed.

5.5.1 Active Characteristics

Exposure
The exposure metric is computed by taking the sum of listings a vendor has (Paquet-Clouston et al., 2018;Décary-Hétu and Quessy-Doré, 2017; van Wegberg et al., 2018). However, only the listings that are active in the specified month are taken into consideration for this research. Listings that are hidden or deleted, are excluded. This ensures that exposure does not become a cumulative sum of all the listings a vendor had in the past. Listings are considered active in a month if they are created in or before that month. Listings are not included if the last purchase of that listing was before that month and the listing is marked deleted.

Diversity
Concerning diversity, there are two metrics used. The first metric is computed following the same procedure as the Diversity Index (Agresti and Agresti, 1978) and is labeled as *degree of diversity*. The formula of degree of diversity is shown in Equation 1. stands for the amount of categories a vendor is active in, and p_i stands for the proportion of listings in the ith category.

$$DegreeDiversity_{vm} = 1 - \sum_{i=1}^{k} p_i^2 \tag{1}$$

The second metric for diversity is the *diversity in categories*: the amount of sub-categories in drugs a vendor is active in: defined as count variable from 1 to 7. However, for the analyses the variable is considered continous, due to the 'vendor average' data calculating the average amount of categories a vendor was active in a month.

Experience
Experience is calculated conform the method adopted by Pacquet-Clouston & Décary-Hétu (2018); and van Wegberg et al. (2019). A vendor's experience is measured in the amount of days between the most recent finalized transaction and the first-ever finalized transaction by that vendor. Another method for calculating the experience of vendors is to only count the months they were active. However, using this approach the effect of returning vendors is left out. On the contrary, the months-active approach has a lower risk of overrating the experience of vendors. This overestimation of the adopted metric is simply illustrated by the following example: imagine a vendor making a transaction at time t=2 days and the next transaction at t=200 days, the vendor's experience is then calculated to be 200 - 2 = 198 days of experience. The calculation of the experience metric is shown by the following formula.

$$Experience_{vm} = max(T_m) - min(T_1) \tag{2}$$

Multihoming
Multihoming is measured as a binary variable. The information whether a vendor is or was active on other markets was found in the database. The database of *market X* consists of a table with vendor aliases that are imported from other markets. The active characteristic of multihoming is measured 1 if a the vendor is imported: being active at least on one other market. Or 0, meaning that a vendor

is solely active on *market X*.

$$Multihoming = \begin{cases} 0, & \text{if } M = 0 \\ 1, & \text{otherwise} \end{cases} \tag{3}$$

Forum Activity

Next to the trading platform, *market X* facilitates a forum for communication between the users of the market. On this forum the same aliases are used on the market. Forum Activity is measured as a binary variable: 1 means that a vendor posted at least one message. 0 if a vendor did not post anything on the forum of *market X*. Considering the longitudinal nature of this research, this means that a vendor can be active on the forum in month 1, but not active on the forum in month 2, thus scoring a 1 and 0 respectively.

$$ForumActivity_m = \begin{cases} 0, & \text{if } F = 0 \\ 1, & \text{otherwise} \end{cases} \tag{4}$$

5.5.2 Passive Characteristics

Reputation

The reputation of vendors is measured in different ways in literature. Often, the way reputation is measured, depends on the feedback scoring mechanisms that applies on the darknet market under research. However, for *market X* the reputation of vendors, or rating, was depicted by the amount of positive and negative ratings. Therefore, these two variables are included as metrics for measuring reputation in the final data-set.

Loyalty

Customer loyalty is difficult to model. There are different approaches and definitions on what loyalty is and how it's measured. Using the same definition as discussed in section 3.2 the loyalty score is calculated according to the procedure as found in darknet market literature Décary-Hétu and Quessy-Doré, 2017. The loyalty score L_v for vendor v with buyer i in category C is calculated by dividing the amount of purchases P buyer i does with vendor v in category C with the total amount of purchases P from buyer i in category C. This is shown in Equation (5):

$$Loyalty_{vC^i} = \frac{P_{iC^v}}{P_{iC}} \tag{5}$$

However, this gives the loyalty score for vendor v based on only one buyer b: The loyalty score of vendor v for the total amount of buyers N for category C is calculated as displayed in Equation (6):

$$Loyalty_{vC} = \frac{\sum_{i=1}^{N} Loyalty_{vC^i}}{N} \tag{6}$$

Due to vendors being active in multiple categories the weighted average of the scores per category was calculated to determine the overall loyalty score for a vendor v.

5.6 Constructing the Datasets for Analysis

This research prepared the transaction-level dataset in such a way, that it consists of monthly records, for each vendor that was active in the drugs segment. This is achieved by taking a subset from the reconstructed transaction-level dataset: transactions are extracted if they took place in the drugs segment. The final dataset contains transaction information for 1435 unique vendors, each entering and leaving the drugs market at different points in time. A new dataframe is created, by constructing the vendor characteristic metrics based on the monthly transaction-level data, as discussed in section 5.5.1 and section 5.5.2.

The resulting dataframe with monthly values for vendors, referred to as the 'vendor monthly' dataset, is used as input data for the multiple regression analyses in section 8.3, and for the descriptive analysis in section 6. A preview of this data is shown in Table 2. For the cluster analysis in section 7, and the multiple regression analyses over time in section 8.2, an aggregated version of this dataframe is used as input data, which is referred to as the 'vendor average' dataset, is used as input data. A preview [2] of this dataframe is shown in Table 3. In the 'vendor average' dataset, for every vendor the means are calculated for performance, revenue, loyalty, exposure, degree of diversity, diversity in categories, forum activity, negative ratings and positive ratings. For experience the last known value is extracted.

ID	Month	Sales	Revenue	Rating Positive	Rating Negative	Loyalty	Experience	Exposure	Diversity Degree	Diversity Categories	Forum Activity	Multihoming
A	21	17.0	487.93	37.0	0.0	0.73	121.0	23.0	0.61	3.0	0.0	1.0
A	22	42.0	1379.68	70.0	2.0	0.59	152.0	27.0	0.62	4.0	0.0	1.0
A	23	15.0	409.94	79.0	2.0	0.52	171.0	26.0	0.64	4.0	0.0	1.0
B	18	4.0	936.37	4.0	0.0	0.00	7.0	18.0	0.00	1.0	0.0	0.0
B	19	15.0	7960.75	18.0	0.0	1.00	40.0	17.0	0.00	1.0	0.0	0.0
B	20	23.0	9284.14	40.0	0.0	0.60	68.0	17.0	0.00	1.0	0.0	0.0
B	21	23.0	12462.82	59.0	0.0	0.51	100.0	17.0	0.00	1.0	0.0	0.0
B	22	7.0	2916.62	64.0	0.0	0.48	130.0	17.0	0.00	1.0	0.0	0.0
C	18	19.0	582.45	16.0	0.0	0.42	12.0	10.0	0.00	1.0	1.0	1.0
C	19	42.0	1890.88	52.0	0.0	0.45	40.0	23.0	0.53	3.0	1.0	1.0
C	20	36.0	1545.64	80.0	1.0	0.45	70.0	27.0	0.47	3.0	1.0	1.0
C	21	57.0	3047.22	126.0	1.0	0.40	101.0	31.0	0.42	3.0	0.0	1.0
C	22	41.0	2340.85	155.0	1.0	0.34	129.0	47.0	0.29	3.0	0.0	1.0
C	23	5.0	313.33	159.0	1.0	0.33	137.0	44.0	0.31	3.0	0.0	1.0
D	18	9.0	29.12	9.0	0.0	0.00	4.0	1.0	0.00	1.0	1.0	1.0
D	21	115.0	4049.30	113.0	0.0	0.72	97.0	45.0	0.74	2.0	1.0	1.0
D	22	173.0	13842.56	259.0	1.0	0.58	128.0	53.0	0.66	2.0	0.0	1.0
D	23	143.0	10692.75	371.0	1.0	0.56	147.0	52.0	0.62	2.0	0.0	1.0
E	18	15.0	937.27	15.0	0.0	0.78	14.0	194.0	0.00	1.0	0.0	1.0
E	19	21.0	1200.48	35.0	0.0	0.49	42.0	186.0	0.00	1.0	0.0	1.0
E	20	22.0	2275.41	55.0	0.0	0.44	71.0	179.0	0.00	1.0	0.0	1.0
E	21	30.0	5361.62	80.0	0.0	0.49	104.0	170.0	0.00	1.0	0.0	1.0
E	22	13.0	3409.20	93.0	0.0	0.44	134.0	155.0	0.00	1.0	1.0	1.0
E	23	4.0	1882.24	96.0	0.0	0.40	143.0	146.0	0.00	1.0	0.0	1.0
F	18	12.0	29.87	11.0	0.0	0.30	4.0	15.0	0.00	1.0	1.0	1.0

Table 2: Preview of the 'Vendor Monthly' Dataset

ID	Sales	Revenue	Rating Positive	Rating Negative	Loyalty	Experience	Exposure	Diversity Degree	Diversity Categories	Forum Activity	Multihoming
A	24.67	759.18	62.00	1.33	0.61	171.0	25.33	0.62	3.67	0.00	1.0
B	14.40	6712.14	37.00	0.00	0.52	130.0	17.20	0.00	1.00	0.00	0.0
C	33.33	1620.06	98.00	0.67	0.40	137.0	30.33	0.34	2.67	0.50	1.0
D	110.00	7153.43	188.00	0.50	0.46	147.0	37.75	0.50	1.75	0.50	1.0
E	17.50	2511.04	62.33	0.00	0.51	143.0	171.67	0.00	1.00	0.17	1.0
F	12.00	29.87	11.00	0.00	0.30	4.0	15.00	0.00	1.00	1.00	1.0

Table 3: Preview of the 'Vendor Average' Dataset

[2] The values are based on the preview shown in Table 2

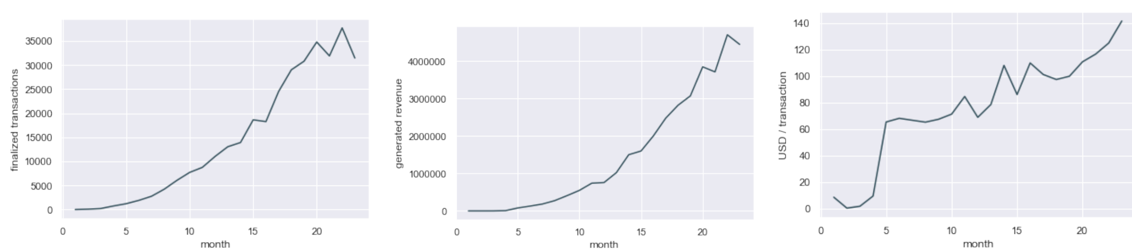
6 Descriptive Analysis

The first section provides general analysis on *market X* as a whole, whereas the second section focuses on the drugs segment. The last section performs a descriptive analyses on the drugs segment with vendors being the unit of analysis. Moreover, this sections examines how their characteristics, as defined in section 3.2, evolve over time.

6.1 Overall Market Analysis

The analyzed market activity on *Market X* consists of 328.819 transactions finalized over a 23 month time-frame, 665 days in particular. The total amount of revenue generated by market X is estimated at 34,36 million US Dollars. On average that would be 51.669 US dollars per day and 104,50 US dollars per transaction.
In total, there were 59.201 unique buyers and 1.755 vendors active on market X. The monthly growth that explains these numbers is displayed in Figure 10.

Market Performance
In section 5.2, Figure 9 provided an overview of market X' activity over time, in terms of finalized transactions. The plots in Figure 10 however, display the aggregated market activity per month in terms of sales, revenue and value per transaction respectively. In Figure 10a and Figure 10b the maximum is reached in month 22 with 37,7 thousand transactions and 4,71 million revenue in US Dollars. The explanation for the fact that month 23 shows declining sales and revenue, is that the last 10 days were excluded from the data [3]. More interesting is the increasing average value per transaction, shown in Figure 10c. This indicates that either vendors raise their prices over time, or that buyers buy higher quantities per transaction.

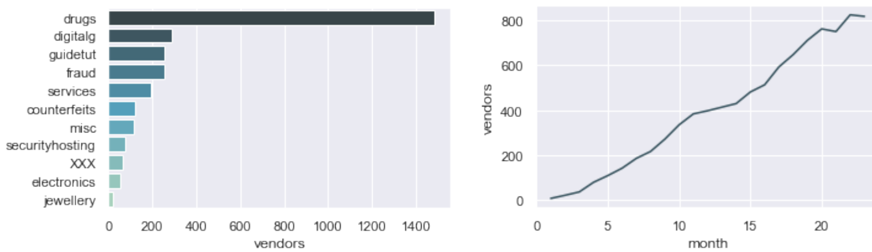


(a) total transactions (b) total revenue (c) average value of a transaction

Figure 10: Global market statistics

Vendors
Vendors are considered active in a month, if they finalize at least one transaction in the corresponding month. Keeping in mind that in total 1,755 vendors were active on *market X* over a period of 23 months: Figure 11 shows how many vendors were active each month and in what segments they were selling goods. In month 23, 826 vendors were actively selling goods. While on average, a vendor was active for 5 months and 6 days on *market X*. Figure 11a shows that 1435 vendors were active in the drugs segment, which is 81.7% of total vendors. Other segments like *digital items*; *tutorials*; *fraud*; are represented by 16.4%, 14,7% and 14,7% of total vendors respectively. While in the least populated segment *jewellery*, only 1,6% of total vendors were active.'

[3] This data was excluded because it contained corrupt or inaccurate data.



(a) vendors per segment (b) active vendors per month

Figure 11: Segments and Vendor Activity

6.2 The Drugs Segment

In total, there were 214,154 transactions finalized in the drugs segment. This comes down to 65.1% of total transactions on *market X*. Around 31,7 million US Dollars in revenue was generated, which is 92.4% of the market's total revenue. The difference in percentages between revenue and transactions is due to drugs having on average a higher unit price, as compared to most products in other segments. Figure 11a already showed that there were 1435 vendors active in the drugs segment.

Drug Categories

In the drugs segment on market X, drugs are categorized into different categories. This research used the same categorization as Pacquet-Clouston et al. (2018), which resulted in a total amount of seven categories. Figure 12 lists the seven categories, together with the amount of vendors active in the respective category. Cannabis is by far the most popular category among vendors, whereas the opioids is the least popular category: merely 200 vendors were selling opioids.

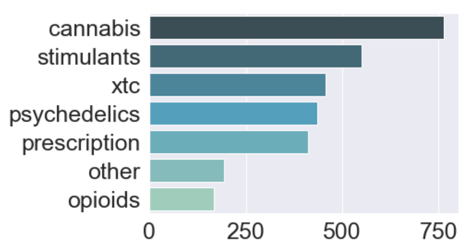


Figure 12: vendors per drug category

Vendor Performance

In Table 4[4] the monthly sales and revenue generated for the corresponding months are shown. Equivalent to the growth of *market X*, the drugs segment grows in both revenue and sales over time as well. When comparing the rates of increase for both the median and mean for sales, it becomes clear that the increased market activity has a bigger effect on the upper 50% of the vendors, and less on the lower 50% of vendors. The same holds for revenue. This is explainable due to new vendors registering and performing worse, bringing the median down. But also, due to established vendors that attract new or existing buyers more easily than beginning vendors: with a more increasing mean as a consequence. On average, in the drugs segment a vendor finalizes 147 transactions in total. While the average amount per month for a vendor comes down to 18 transactions.

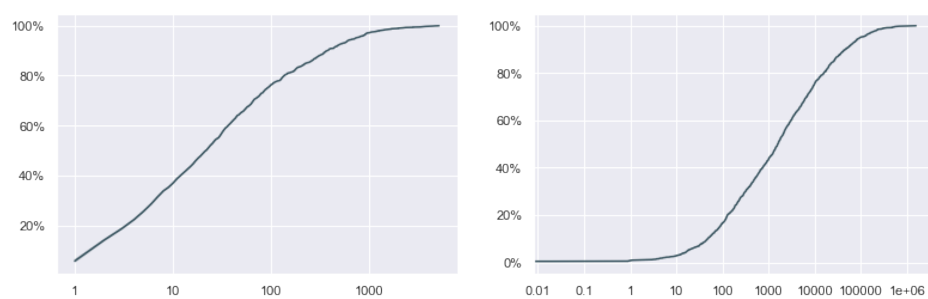
To get a better view on how vendors in the drug segment are performing, the total sales and revenue

[4]The table shows values per month, only for the months displayed

month	vendors	sales sum	sales mean	sales median	revenue sum	revenue mean	revenue median
1	4	5.00e+00	1.25e+00	1.00e+00	5.12e+01	1.28e+01	0.00e+00
4	55	2.73e+02	4.96e+00	2.00e+00	6.74e+03	1.22e+02	2.12e+01
7	145	1.50e+03	1.03e+01	4.00e+00	1.72e+05	1.18e+03	4.12e+02
11	304	5.28e+03	1.74e+01	6.00e+00	7.02e+05	2.31e+03	6.08e+02
14	339	8.58e+03	2.53e+01	8.00e+00	1.46e+06	4.29e+03	6.25e+02
17	471	1.49e+04	3.16e+01	9.00e+00	2.37e+06	5.04e+03	8.49e+02
20	606	2.35e+04	3.88e+01	1.05e+01	3.51e+06	5.78e+03	8.23e+02
23	682	2.50e+04	3.67e+01	1.00e+01	4.07e+06	5.96e+03	9.97e+02

Table 4: Market Performance over time

generated are summed per vendor. Given this data, two cumulative distribution plots are shown in Figure 13 that show how vendors perform relative to other vendors.
Regarding the sales performance, around 80% of the vendors perform no more than 100 transactions during their lifetime, while only 2% of the vendors finalizing 1000 transactions or more.
The line in Figure 13b is rather distorted, this is due to the database not registering transaction values (see chapter data for explanation). Ignoring this distortion, an equivalent can be drawn as for sales performance: 80% of the vendors did not generate more than 10.000 US Dollar. Soska and Christin 2015 argue that only vendors that generate more than 10.000 US Dollars should be considered as being successful in the ecosystem. However, as this only concerns the transaction data of *market X*, it could be that some of the vendors in the 80% are also active on other markets.



(a) CDF of sales (b) CDF of revenue

Figure 13: Cumulative Distribution Plots of Performance

6.3 Vendor Characteristics

The descriptive statistics on vendor characteristics are shown in Table 5, which is based on the monthly aggregated data of vendors. On average in a month, vendors have 33 listings and are active in two drugs categories. The maximum of 718 listings in a month is extremely high and could have different causes: either a vendor has many clone listings, listings that are exact copies of existing listings but differ in the quantity offered. Or, vendors keep their old listings active for referencing due to the feedbacks attached to those listings.
On average, vendors in the drugs segment are intermediate experienced with 147 days. This seems to be in line with previous research, although due to the different time frames that were analyzed, it remains difficult to make a one-on-one comparison (Soska and Christin, 2015; Paquet-Clouston et al., 2018).The newly introduced characteristic for forum activity, shows that on average a vendor is 21% of the time active on the forum. Whereas only 36% of all vendors were at least one month active on

the forum. The monthly average for multihoming states that 56% of vendors come from, or are active on, other darknet markets.

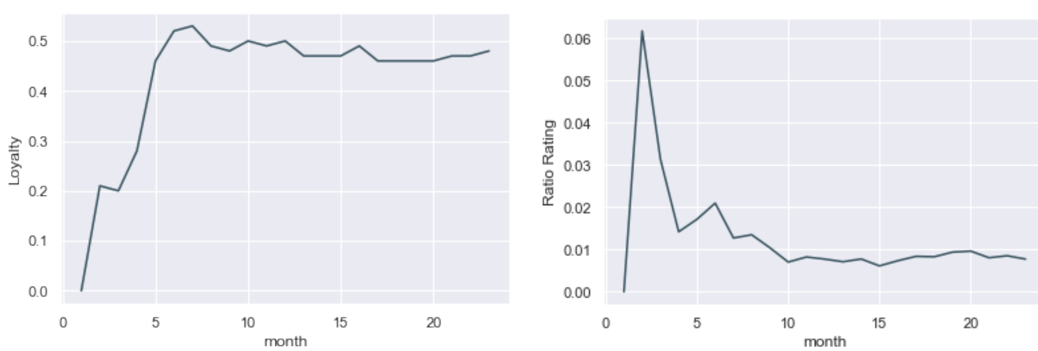
	vendors	min	max	mean	std
Exposure	1435	1	718	33	47
Experience	1435	0	731	147	156
Degree Diversity	1435	0	1.00	0.36	0.40
Category Diversity	1435	1	6	2	1
Forum Activity	1435	0	1.00	0.21	0.34
Multihoming	1435	0	1.00	0.56	0.50
Loyalty	1435	0	1.00	0.36	0.30
Positive Ratings	1435	0	3800.00	113.90	311.90
Negative Ratings	1435	0	84.00	1.09	4.13

Table 5: Average Vendor Characteristics

The values for passive characteristic plots show that vendors, on average, score 36% on loyalty. While, vendors gain almost 114 positive and only 1 negative feedback during their lifetime on a *market X*. Considering the average amount of finalized transactions per vendor in section6.2, 77% of those transactions are positively rated and 0,7% are negatively rated.

Passive Characteristics over Time

The change in means for passive characteristics are shown in **??**. On avarage, the loyalty score of vendors stabilizes from month 6 on at around 0,48 and stays at this level for the following months. Figure 14b shows the ratio of negative to positive ratings. In the early phases of the market, a lot more negative ratings are given. This is likely due to the small amount of vendors that are active on the market: in the fourth active month of the market, ther Table 5

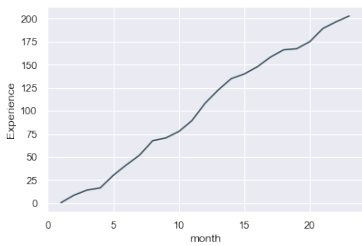


(a) Loyalty

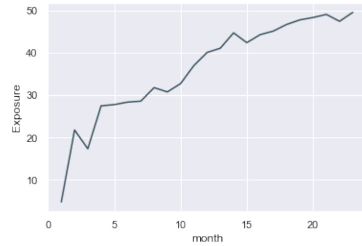
(b) Ratio of Ratings

Active Characteristics over Time

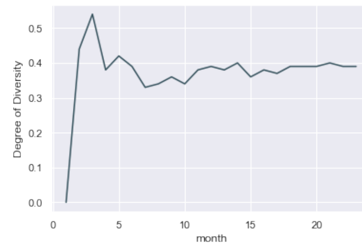
In Figure 15, changes in the means for active characteristics over time are shown. Exposure in particular is interesting, as Figure 15b implies that the average amount of active listings per vendors increases as the market activity increases as well: This is in line with previous research on vendor' exposure on darknet markets (Paquet-Clouston et al., 2018). The numbers in Figure 15e and Figure 15f state that starting from month 6, on average, between 15% and 20% of vendors were active on the forum, and around 70% of the active vendors on *market X* were also active on other markets. Figure 15c and Figure 15d show that on average vendors are active in 2 categories and score just below 0.4 on average in degree in diversity. The latter is higher than the 0,25 found in the study of Pacquet-Clouston et al. (2018). This is likely due to the fact that this research only takes into account vendors that were active, i.e. made at least one transaction during their lifetime, whereas in the other study also experimenting vendors were considered.



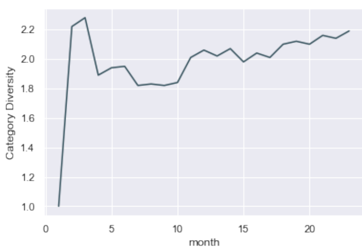
(a) Experience



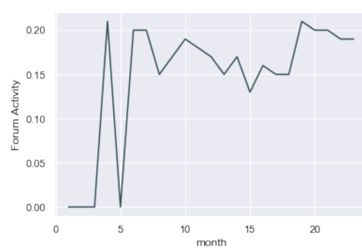
(b) Exposure



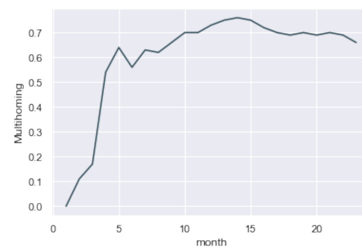
(c) Degree in Diversity



(d) Category Diversity



(e) Forum Activity



(f) Multihoming

Figure 15: Active characteristics over time

6.4 Top 25 Performing Vendors

In this final section, the vendor characteristics of the top 25 performing vendors are examined and evaluated. Due to the difference between performance and revenue, two sets of 25 vendors are discussed.

characteristic	Sales mean	min	max
Loyalty	0.65	0.49	0.84
Experience	384.16	112.00	604.00
Exposure	81.37	15.44	400.14
Degree Diversity	0.33	0.00	0.96
Category Diversity	2.17	1.00	5.00
Forum Activity	0.27	0.00	1.00
Multihoming	0.80	0.00	1.00
Sales	221.97	79.00	503.33
Revenue	32,395.49	6500.89	148,934.34

characteristic	Revenue mean	min	max
Loyalty	0.64	0.40	0.85
Experience	338.28	112.00	604.00
Exposure	73.20	15.44	193.40
Degree Diversity	0.39	0.00	0.96
Category Diversity	2.49	1.00	5.00
Forum Activity	0.23	0.00	1.00
Multihoming	0.76	0.00	1.00
Sales	161.21	7.00	503.33
Revenue	50,682.17	17,916.18	148,934.34

Table 6: top 25 performing vendors in sales and revenue

The left table in Table 6 shows the mean values for the the top 25 vendors in sales, while the right table shows the same for the top 25 vendors in revenue. Both sets of vendors score relatively high on loyalty: buyers that perform multiple transactions in the same category, often remain loyal to these vendors. Furthermore, the minimum value for experience stands out in both tables: which indicates that one vendor earned between 50.000$ and 150.000$ in just 112 days. The other characteristics show a broader distribution. From these tables, it becomes clear that top performing vendors in sales and revenue have different values for their vendor characteristics, and a pattern is yet to be found.

7 Clustering Vendor Profiles

This chapter performs hierarchical agglomerative clustering on the dataset with average values, as described in section 5.6. The aim is to determine what distinct vendor profiles, through different combinations of *active* characteristics, are observable on *market X*.

7.1 Running the Agglomerative Clustering Algorithm

Following the procedure outlined in section 4.3, the data was first scaled in order to run the algorithm correctly. The algorithm clusters similar vendors into the same profiles, based on their active characteristics. The reason for only including active characteristics is that these are directly linked to a vendor's behaviour on the market. In other words, the active characteristics are adjustable by vendors at any time.

The elbow method is used to assist in chosing the right amount of clusters. However, in the end the interpretation of the researcher determines the right amount of clusters. The upper plot in Figure 23 shows that 'elbow' in the sum of squared distances is located in the range between 2 and 6 clusters. After arriving at 6 clusters the slope is less steep and continuous the same direction for the rest of the graph. For determining the exact amount of clusters, the dendogram in the lower plot in Figure 23 is inspected. Evaluating the dendogram top-down, the 2-cluster solution consists of either multihoming vendors in one cluster, and non-mulithoming vendors in the other. At distance 17.1, the non-multihoming vendor cluster is split into two new clusters: vendors that are mostly active on the forum, and vendors that are not. At distance 13.3, the multihoming vendor cluster is split up in the same way. At this moment, there are 4 clusters. Continuing this process and separating the clusters at distance 10.2 and distance 10, there are 6 six clusters formed. These last two 'cuts' resulted in a clear distinction in experience and diversity within all clusters. Higher cluster solutions separate the vendors predominantly on experience, which provides little additional information, Moreover, it results in a less parsimonious solution. For these reasons, together with the result of the elbow heuristic, the 6-cluster solution is chosen.

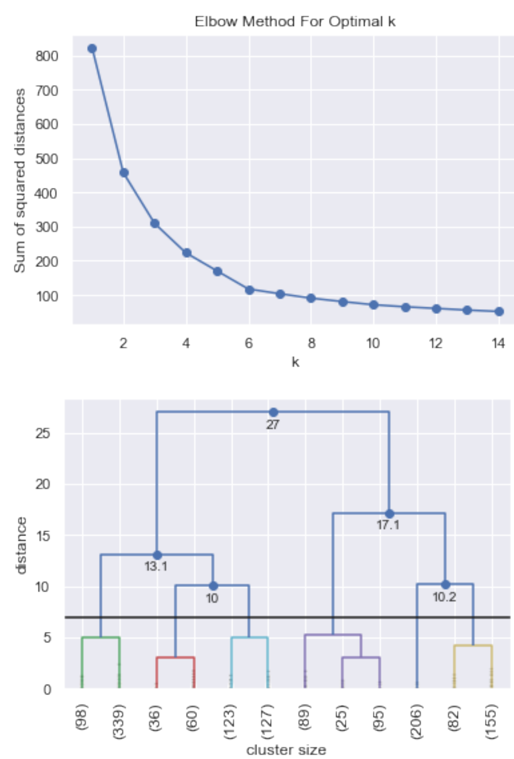
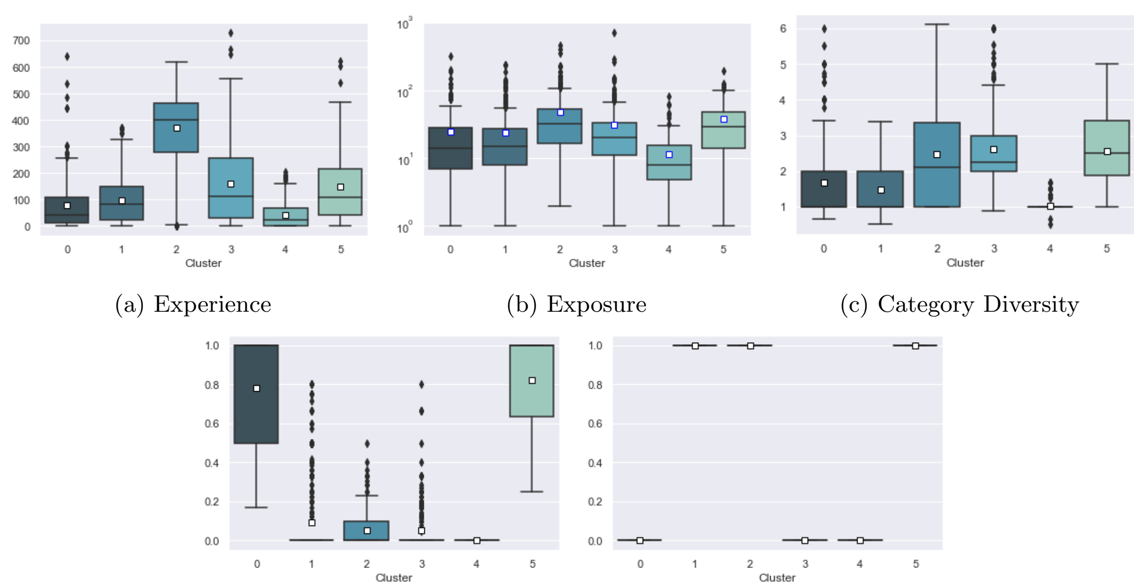


Figure 16: chosing the amount of clusters

In Figure 23, the six clusters are depicted by different colors in the dendogram, together with their respective size, which are shown on the x-axis. The different cluster sizes under the red tree, 36 and 60, are summed together, because they fall below the cut-off point, thus forming one cluster of 96 vendors in total.

7.2 Interpreting the Clusters

To facilitate the interpretation of the 6-cluster solution, the different distributions of the active characteristic per cluster are shown in Figure 17.
There are three distributions found of the active characteristic experience, among the six clusters in Figure 17a: a vendor is either active for a relatively short time, or for a moderate amount of time, and there are vendors that active for an extensive amount of time. The hypothesis here is that the short-stay vendors come and go during the whole lifetime of the market, and the long-stay vendors are active from the early beginnings of the dark market. However, the distributions in Figure 17a show skewness, and in the case of cluster 2 a heavy tailed distribution.
The boxplot for the exposure characteristic is log-transformed, to make the differences between the clusters more clear. The clusters follow the same distributions: skewed to the right. From Figure 17b it seems that, on average, clusters 2 and 5 have more listings than other clusters, where cluster 4 stands out with the lowest median.
In Figure 17c, the category diversity characteristic shows very different distributions among the clusters. Approximately three distributions seen. Vendors within the Drugs segment are on average either active in a lower amount of product categories, with the mean at at around 1.5 product categories, or vendors are active more categories with the mean at around 2.5 product categories. Cluster 4 shows a complete different distribution: almost every vendor is active in only one category.
Moving on to the forum activity of vendors, vendors are either active, not active, or seldomly active on the forum. Clusters 1, 2, 3 are predominantly not active on the forum. While clusters 0 and 5 consists of vendors that are active for most of their lifetime on *market X*. Cluster 4 consists of vendors that were never active on the forum.
Interpreting the distribution of multihoming among the different clusters is relatively easy, due to the characteristic is measured dichotomously. Clusters 1, 2 and 5 consists of vendors that are, or were, also active on other markets during the lifetime of market.



(a) Experience (b) Exposure (c) Category Diversity

(d) Forum Activity (e) Multihoming

Figure 17: active characteristics in clusters

Cluster 0: Single-Market Specialist I
These vendor have their focus on one or two product categories, and sell their goods only on *market X*. On average, they have relatively little experience, and are active on the forum for most of their lifetime on *market X*.

$n = 209$	mean	std	min	median	max
Experience	79.06	101.67	0.00	41.0	641.0
Exposure	24.78	35.57	1.00	14.0	324.0
Category Diversity	1.68	1.05	0.67	1.0	6.0
Forum Activity	0.78	0.28	0.17	1.0	1.0
Multihoming	0.00	0.00	0.00	0.0	0.0

Figure 18: Cluster 0 Descriptives

Cluster 1: The Multi-Market Specialist
Cluster 1, the biggest cluster, consists vendors that are also specialized in either one or two drugs categories. However, these vendors also sell their goods on other markets, and are rarely seen on the forum of *market X*. They have a little more experience than the single market specialists.

$n = 437$	mean	std	min	median	max
Experience	97.25	83.85	0.0	82.0	373.0
Exposure	24.07	30.48	1.0	15.0	239.0
Category Diversity	1.48	0.69	0.5	1.0	3.4
Forum Activity	0.09	0.18	0.0	0.0	0.8
Multihoming	1.00	0.00	1.0	1.0	1.0

Figure 19: Cluster 1 Descriptives

Cluster 2: Multi-Market Generalist
Vendors with the multi-market generalist profile have a lot of experience. On average, they are active in two product categories. Vendors with this profile are active in selling their goods on other markets. Just a few of these vendors have ever been active on the forum.

$n = 250$	mean	std	min	median	max
Experience	371.40	143.70	0.0	400.50	619.00
Exposure	48.24	57.06	2.0	32.25	461.00
Category Diversity	2.47	1.34	1.0	2.11	6.12
Forum Activity	0.05	0.09	0.0	0.00	0.50
Multihoming	1.00	0.00	1.0	1.00	1.00

Figure 20: Cluster 2 Descriptives

Cluster 3: Single-Market Generalist
These vendors only trade their goods on *market X*. They are relatively experienced, compared to the other single-market vendor profiles. Despite their experience, these vendors do not make use of the forum. On average, these vendors are active in almost three categories.

$n = 237$	mean	std	min	median	max
Experience	159.05	157.15	0.00	113.00	731.00
Exposure	31.40	56.23	1.00	20.00	717.25
Category Diversity	2.63	1.17	0.88	2.25	6.00
Forum Activity	0.05	0.12	0.00	0.00	0.80
Multihoming	0.00	0.00	0.00	0.00	0.00

Figure 21: Cluster 3 Descriptives

Cluster 4: Single-Market Specialist II
Cluster 4 consists of vendors that are almost equal to the single market specialists of cluster 0. However, vendors have never been active on the forum and are solely active in just one category. Moreover, this profile consists of vendors with very little experience: on average, they have the least experience.

$n = 206$	mean	std	min	median	max
Experience	79.06	101.67	0.00	41.0	641.0
Exposure	24.78	35.57	1.00	14.0	324.0
Category Diversity	1.68	1.05	0.67	1.0	6.0
Forum Activity	0.78	0.28	0.17	1.0	1.0
Multihoming	0.00	0.00	0.00	0.0	0.0

Figure 22: Cluster 4 Descriptives

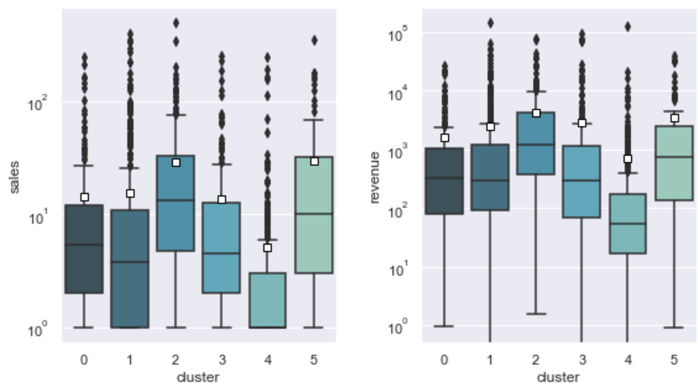
Cluster 5: Multi-Market Generalist II
Cluster 5 stands out because its relative small size: only 96 vendors share this profile. These generalist vendors trade in multiple product categories, are relatively experienced, and are active on multiple markets. More notable about this profile is that on average, vendors were active on the forum of *market X* for 82% of their time.

$n = 96$	mean	std	min	median	max
Experience	148.50	145.23	0.00	110.50	624.0
Exposure	38.67	33.91	1.00	29.08	193.4
Category Diversity	2.55	1.17	1.00	2.50	5.0
Forum Activity	0.82	0.21	0.25	1.00	1.0
Multihoming	1.00	0.00	1.00	1.00	1.0

Figure 23: Cluster 6 Descriptives

7.3 Comparing Performance Between Clusters

This section will compare the clusters and determine if there are differences between the clusters in the degree of vendor performance. As performance is measured by two metrics, sales and revenue, the necessary tests are performed simultaneously. Figure 24 shows the distributions among the clusters for vendor performance. All clusters show a similar distribution and outliers. The white squares inside the boxes in Figure 24 respresent the cluster means.



(a) Sales (b) Revenue

Figure 24: performance in clusters

To determine whether multiple groups score differently on a dependent variable, a one-way analysis of variances (ANOVA) is usually adopted. However, for ANOVA it is important that the assumptions for performing the procedure are met. The first being that the residuals follow a normal distribution. The second is that there should be equal variances between groups, also referred to as homogeneity of variances. To test the whether the assumption of normality is passed, the Shapiro-Wilkinson test is performed. The null hypothesis is that the population is normally distributed. Both sales and revenue are not normal distributed, as the Shapiro-Wilkinson test is significant ($p < .001$), which translates to the first assumption being violated.

To test whether the second assumption holds, the Levene test is performed. For both sales and revenue, the Levene's statistic is tested as significant, $\rho < 0.001$. As a consequence, the second assumption is violated as well. With both assumptions violated, a non-parametric test is performed. Subsequently, a post-hoc test is performed that relaxes these assumptions. The non-parametric test that is selected is the Kruskal-Wallis test: It assesses whether groups are significantly different on a dependent variable. An examination of the boxplots in Figure 24 shows that all distributions have the same shape for each group in revenue.

Now, performing the Kruskal-Wallis test for sales, the difference in median sales values between clusters is statistically significant, with $\chi^2(5) = 135.8$, $\rho < 0.001$. The difference in the median of revenue between clusters, is also proven to be statistically significant with $\chi^2(5) = 130.65$ and $\rho < 0.001$.

Finally, a post-hoc test is performed to do a pair-wise comparison of the clusters. The pairwise comparison follows Dunn's method. Furthermore, it corrects for the increasing chance of making a Type 1 error by adjusting the ρ −values according to the Bonferroni method.

For sales, the Dunn-test reveals that there are statistically significant differences between most of the clusters. In total there are 4 pairs of clusters where the difference is not significantly different. First, between clusters 0 (Mdn = 5.33) and 1 (Mdn = 3.75) the difference is not significantly different (p = 1.00). Second, the difference between clusters 0 (Mdn = 5.33) and 3 (Mdn = 4.50) is also not statistically significant (p = 1.00). Third, clusters 1 (Mdn = 3.75) and 3 (Mdn = 4.50) the difference

Cluster	Sales	Revenue
0	5.33	324.79
1	3.75	295.18
2	13.21	1203.45
3	4.50	294.27
4	1.00	53.75
5	10.00	752.18

Table 7: median values for performance per cluster

is also not statistically significant (p = 1.000). Finally, the comparison of clusters 2 (Mdn = 13.21) and 5 (10.00) also yields a difference that is not significant(p = 1.000).
When assessing the results from the Dunn-test for revenue, the same results are achieved in the difference between clusters aside from one pair of clusters. Namely, the difference between clusters 1 (Mdn = 295.18) and 5 (Mdn = 752.18) is found to be not statistically significant (p ¡ 0.001).

Sales			Revenue		
Clusters	Z	p	Clusters	Z	p
0 - 1	0.233	1.000	0 - 1	-1.757	0.591
0 - 2	-5.897	0.000	0 - 2	-6.578	0.000
1 - 2	-7.217	0.000	1 - 2	-5.911	0.000
0 - 3	0.630	1.000	0 - 3	0.013	1.000
1 - 3	0.498	1.000	1 - 3	1.848	0.485
2 - 3	6.756	0.000	2 - 3	6.814	0.000
0 - 4	4.749	0.000	0 - 4	4.067	0.000
1 - 4	5.285	0.000	1 - 4	6.473	0.000
2 - 4	10.829	0.000	2 - 4	10.795	0.000
3 - 4	4.267	0.000	3 - 4	4.178	0.000
0 - 5	-3.314	0.007	0 - 5	-3.107	0.014
1 - 5	-3.799	0.001	1 - 5	-2.087	0.277
2 - 5	1.200	1.000	2 - 5	1.945	0.389
3 - 5	-3.872	0.001	3 - 5	-3.176	0.011
4 - 5	-7.080	0.000	4 - 5	-6.330	0.000

Table 8: Dunn-test on Sales and Revenue

The full output of the post-hoc analysis is shown in Table 8. The results confirm the expectations on the difference in medians between the clusters, when the box plots in Figure 24 are considered. Altogether, these results indicate that clusters 0, 1 and 3 achieve performance at comparable levels, and the same goes for clusters 2 and 5. Cluster 4, having a statistically significant lower median as compared to all other clusters in revenue *and* sales, was already expected to perform less.

Sales			Revenue	
Clusters	Vendors		Clusters	Vendors
0	2		1	5
1	5		2	10
2	11		3	5
3	3		4	1
5	4		5	4

Table 9: Top 25 Performing Vendors with Respective Clusters

Now, it is interesting to assess how the clusters are distributed among the top 25 performing vendors, in sales and revenue respectively. Table 9 shows the vendor counts for every cluster. In both sales and revenue, the top 25 consists mainly vendors that belong to cluster 2: around 40 percent. Cluster 5 is represented by merely 16 percent in both tables. However, clusters 1 and 3 are represented by the same amount of vendors. It stands out, that a vendor from cluster 4 belongs to the top 25 performing vendors in revenue. However, when the outliers are inspected in Figure 24, cluster 4 shows a isolated outlier, outperforming most outliers from the other clusters. This brief examination of the top 25 vendors, indicates that there is no profile that is predominantly present in the top 25 performing vendors, Although cluster 2 scores the best. This also confirms the findings from the descriptive analysis in section 6.4.

7.4 Summary of Results

This analysis clustered vendors into different profiles based on their active characteristics. The analysis shows that while cluster 5 is active on the forum, they have relatively little experience. These characteristics distinguish this vendor profile clearly from cluster 2. Nevertheless, the difference in performance between these clusters is not statistically significant. Both clusters consists of vendors that are multihoming. But what distinguishes these two clusters from the other clusters, enabling them to perform better? The one thing that comes forward from the boxplots in Figure 24, is that clusters 2 and 5 have more exposure. These differences in characteristics among clusters, together with the differences in performance, indicate that the characteristics exposure and multihoming have a positive relation with vendor performance.

8 The Impact of Characteristics on Performance

This chapter entails the regression analyses that are performed on the data. First, the input variables are discussed together with how they are included on the models. In the second section the OLS regressions are performed on the '*averaged vendor*' dataset. These models aim to provide information on the individual effects of characteristics and vendor profiles on performance. In the third section, OLS regression models are performed for every subsequent month, using the '*monthly vendor*' dataset. The OLS regression results for every month are then taken together, and plotted over time. This procedure makes it possible to determine whether the impact of vendor characteristics on performance changes over time. The final section summarizes the results into a brief conclusion.

8.1 Variables

8.1.1 Dependent Variables

Sales
The dependent variable *sales*, measured in finalized transactions. However, the sales variable is not normally distributed: It is strongly positive skewed. Therefore, a log transformation is applied. The logistic transformation of sales is shown in Figure 25. What stands out from the log-transformation is that there are two peaks observed. The smaller peak is explained by the fact that each month there is a large amount of vendors almost selling almost nothing. When the monthly data is filtered in such a way that only samples with at least on month experience (30 days) are included: A large part of this peak is explained away. When only including monthly data where vendors have experience less than one month: The cause for the peak is clearly visible. However, a smaller peak is still visible in the filtered data. This peak could be due to experienced vendors either performing worse or having vacation.

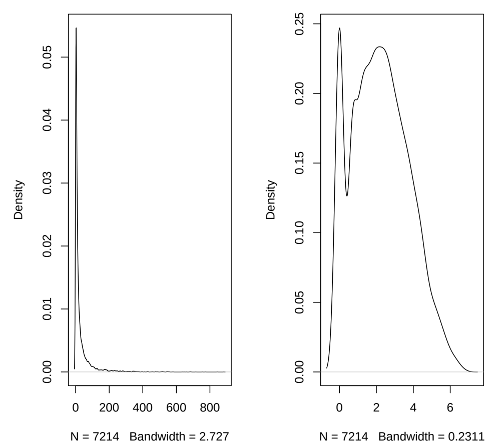


Figure 25: log of sales

Revenue
For the second leg of analyses, revenue is added as the dependent variable and is measured on a continuous scale: the estimated amount of revenue in USD per month. Revenue is also not normally distributed, a strong positive skewness is present. Therefore, the same steps are performed as in the previous section with performance as dependent variable and a log transformation is applied Figure 26.

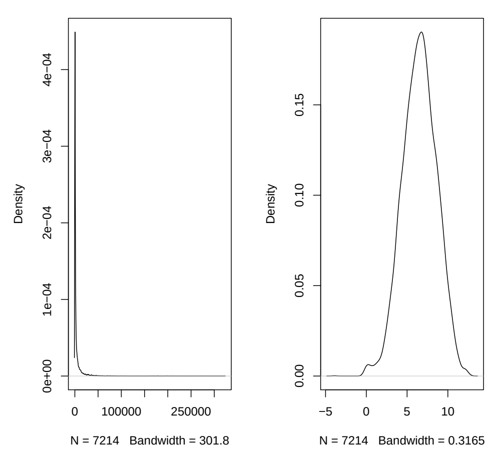


Figure 26: log of revenue

8.1.2 Independent Variables

For each independent variable a brief description is given and whether or not the relation with the dependent variable is linear. Due to some independent variables violating the assumptions of having a linear relation with the dependent variable and violating the assumption of homoscedacity, transformations of these variables were performed. The defined strategies in section 7 are also included as categorical dummy variables. An overview of the included independent variables for the regression model is given by Table 10.

Characteristic	Transformation?	Input Variable
Exposure	log	log_real_exposure
Category Diversity	none	active_categories
Experience	none	experience
Forum Activity	none	forum_act
Multihoming	none	is_imported
Reputation (positive)	log	log_rpos
Reputation (negative)	log	log_rneg
Loyalty	none	loyalty

Table 10: Independent Variables

8.1.3 Interpretation of log-linear and log-log relations

Due to the transformation of dependent and some independent variables, the interpretation of their regression coefficients changes. Due to the standardized regression coefficients being the focus of this research, the unstandardized regression coefficients are not interpreted in this research. Still, it remains important to understand the difference in interpretation.
For log-linear relations, meaning the dependent variable is log transformed, the interpretation of regression coefficients is shown in Equation 7. Example: if β_1 is 0.07, a one-unit increase in independent variable x translates to roughly a 7% increase in y.

$$\%\Delta y = 100 \times (e^{\beta_1 x} - 1) \tag{7}$$

For log-log relations, meaning both dependent and independent variables are log transformed, the interpretation of regression coefficients is shown in Equation 8. Example: if β_1 is 0.07, a one percentage increase in independent variable x would yield a 0.07% increase in dependent variable y.

$$\%\Delta y = \beta_1 \times \%\Delta x \tag{8}$$

8.2 Multiple Regression

As the goal is to compare the relative impact of the independent variables on the outcome variables, all variables are standardized so that the effects of different scales are eliminated, and the regression coefficients can be compared with each other.
To measure what the relative impacts of vendor characteristics and vendor profiles are on performance, multiple regression analyses are performed. Due to the profiles being derived from active characteristics, these groups are not combined in a regression model. For this reason, there are two main regression models constructed. The first regression analysis examines the relations depicted by the framework in Figure 7, while the second analysis inspects the relations as shown in Figure 8. First, these two regression analyses are performed with Sales as the dependent variable. Then, the analyses are performed with Revenue as the dependent variable. In both models, the reputation characteristics function as control variables. Due to to a near one-on-one relation between feedbacks and sales, the association found in the regression is rather trivial, and functions as control variable.

8.2.1 Predicting Impacts on Sales

Table 11 and Table 12 show the results for the two regression models with sales as dependent variable. For both models, groups of independent variables are included individually to the regression model, to examine their isolated impacts on sales. By adding other independent variables step-wise, the change in strength or direction of the regression coefficients can be observed. This enables to draw conclusions on the interplay of the independent variables. Finally, *Model 6* represents the complete model, where all dependent variables are included.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0110)	(0.0109)	(0.0216)	(0.0211)	(0.0092)	(0.0090)
log_rpos	0.9100***	0.9348***			1.0580***	1.0069***
	(0.0110)	(0.0122)			(0.0136)	(0.0151)
log_rneg		-0.0551***			-0.0521***	-0.0440***
		(0.0122)			(0.0105)	(0.0104)
log_real_exposure			0.4017***		0.1120***	0.1098***
			(0.0255)		(0.0114)	(0.0112)
active_categories			-0.1237***		-0.0315***	-0.0324***
			(0.0242)		(0.0105)	(0.0104)
experience			0.3312***		-0.2621***	-0.2621***
			(0.0236)		(0.0124)	(0.0122)
forum_act			0.1179***		0.0157*	0.0167*
			(0.0222)		(0.0095)	(0.0093)
is_imported			0.0373		-0.0007	-0.0007
			(0.0230)		(0.0098)	(0.0096)
loyalty				0.6023***		0.0809***
				(0.0211)		(0.0113)
R-squared	0.83	0.83	0.33	0.36	0.88	0.88
Adj. R-squared	0.83	0.83	0.33	0.36	0.88	0.88
Log-Likelihood:	-772.58	-762.38	-1747.05	-1712.82	-513.17	-487.94
F-statistic:	6905.51	3509.97	141.83	815.91	1499.03	1364.02
No. observations	1435	1435	1435	1435	1435	1435

Standard errors in parentheses
* p<.1, ** p<.05, ***p<.01

Table 11: Multiple Regression Vendor Characteristics, Sales

Impact of Vendor Characteristics on Sales
In *Model 1* and *Model 2* only the control variables are included. The results are not surprising: negative ratings have a negative but very low impact, in comparison with the positive ratings. This is in line with the findings from section 6.3 that on average vendors receive on less than 1% of their

transactions, a negative feedback. The high R-squared value that is observed, is purely explainable by the feedbacks. The reason for a lot of the variance being explained, is that positive feedbacks are highly correlated with sales. Going from *Model 1* to *Model 2*, adding negative feedbacks to the model, there is no additional variance in sales explained.

Model 3 shows the standardized regression coefficients, when only the active characteristics are included as independent variables. Only *Multihoming* is not significant. As all variables are standardized we can determine the relative impact on sales. Exposure has the biggest impact, followed by experience. Then, the amount of categories has the third biggest impact, although negative. This suggests that vendors active in less categories perform better in sales, than vendors that are active in more categories. Forum activity is significant, and has a weak but positive impact. The impact of Multihoming on sales is not statistically differet from zero.

Model 4 includes only loyalty as independent variable. Loyalty has a positive impact of loyalty on sales.

Model 5 and *Model 6* show the final models. Including loyalty to the *Model 6* does not explain more variance: the R-squared remains at 88%. However, compared to the model with the control variables, there is 5% of the variance is purely explained by including the active characteristics. Even more interesting is the fact that the direction and impact strength of some active characteristics change, compared to *Model 3*. What stands out from these results, is that experience changes sign and that the strength of impact remains. This indicates that when (positive) ratings are included as a independent variable, the effect of experience on sales is negative: meaning that if experience increases, while the amount of (positive) ratings remains the same, this has a negative effect on sales. Another active characteristic, that has a relatively strong impact on sales is exposure. Loyalty has a relatively low impact, due to the high association between positive ratings and transactions. However, the effect of loyalty is stronger than the effect of negative ratings. The multihoming characteristic remains insignificant in the final model.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	-0.0000	-0.0000	-0.0578	-0.0000	0.1171***	0.1146***
	(0.0110)	(0.0109)	(0.0661)	(0.0211)	(0.0276)	(0.0272)
log_rpos	0.9100***	0.9348***			0.9730***	0.9192***
	(0.0110)	(0.0122)			(0.0125)	(0.0148)
log_rneg		-0.0551***			-0.0420***	-0.0339***
		(0.0122)			(0.0119)	(0.0118)
C(cluster)[T.1]			0.0020		-0.0620*	-0.0601*
			(0.0804)		(0.0334)	(0.0329)
C(cluster)[T.2]			0.5605***		-0.3595***	-0.3580***
			(0.0895)		(0.0391)	(0.0385)
C(cluster)[T.3]			-0.0534		-0.1736***	-0.1740***
			(0.0907)		(0.0378)	(0.0372)
C(cluster)[T.4]			-0.4261***		-0.0277	-0.0178
			(0.0938)		(0.0392)	(0.0387)
C(cluster)[T.5]			0.4416***		-0.0432	-0.0395
			(0.1178)		(0.0497)	(0.0490)
loyalty				0.6023***		0.0847***
				(0.0211)		(0.0130)
R-squared	0.83	0.83	0.09	0.36	0.84	0.85
Adj. R-squared	0.83	0.83	0.09	0.36	0.84	0.85
Log-Likelihood:	-772.58	-762.38	-1967.74	-1712.82	-704.57	-683.43
F-statistic:	6905.51	3509.97	28.60	815.91	1100.31	996.19
No. observations	1435	1435	1435	1435	1435	1435

Standard errors in parentheses
* p<.1, ** p<.05, ***p<.01

Table 12: Multiple Regression Vendor Profiles, Sales

Impact of Vendor Profiles on Sales

Looking at the results of the regression in Table 12, *Models 1, and 4*, are equal to the equally numbered models in Table 11. The reason to include these once more, is to examine the difference in additional variance that is explained when the vendor profiles are included in the final model.

Model 3 only includes the clusters, or vendor profiles, as independent variables. Due to the clusters being added to the regression model as dummy variables, cluster 0 is left out and represented by the intercept. This means that the regression coefficients for the other clusters are relative to the intercept. The intercept has a negative impact of -0.99, thus if a vendor belongs to cluster 2 his sales is positively impacted with strength $-0.06 + 0.56 = .50$. Only clusters 2, 4 and 5 are significant. This is in line with the results from the cluster analysis in section 7, where the difference in sales between clusters 0, 1 and 3 was not significantly significant. *Model 3* explains 9% of the variance, indicating that clusters, or vendor strategies, can be used to predict performance. The sequence of clusters from biggest to weakest impact is: cluster 5; cluster 3; cluster 0; cluster 2; cluster 4; and cluster 1. With cluster 0 having a negative impact.

The final results in *Model 6* show that compared to the baseline model, 1% of the variance is purely explained by vendor profiles. More interestingly, is the effect of ratings on the profiles, with relation to sales. It stands out that the regression coefficient for cluster 2 switches sign. This is likely due to the fact that there are no vendors in cluster 5 with zero (positive) ratings. So when all other effects are held constant, the sole effect of belonging to cluster 2 has a negative correlation with sales. By which is meant, that the positive effect of cluster 2 is explained by positive ratings. The change of sign of the intercept is also interesting. It indicates that belonging to cluster 0, has a slight positive impact on sales, coexisting with to the effect of (positive) ratings.

8.2.2 Predicting Impacts on Revenue

Table 13 and Table 14 show the results for the two regression models with revenue as dependent variable. The same procedure is followed as in section 8.2.1.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0179)	(0.0176)	(0.0222)	(0.0219)	(0.0170)	(0.0166)
log_rpos	0.7347***	0.8002***			0.7972***	0.6929***
	(0.0179)	(0.0197)			(0.0251)	(0.0279)
log_rneg		-0.1457***			-0.1484***	-0.1319***
		(0.0197)			(0.0195)	(0.0192)
log_real_exposure			0.3892***		0.1818***	0.1772***
			(0.0261)		(0.0211)	(0.0206)
active_categories			-0.1436***		-0.0552***	-0.0569***
			(0.0248)		(0.0195)	(0.0191)
experience			0.2969***		-0.1235***	-0.1235***
			(0.0242)		(0.0230)	(0.0225)
forum_act			0.0587***		-0.0127	-0.0107
			(0.0227)		(0.0175)	(0.0172)
is_imported			0.0753***		0.0531***	0.0533***
			(0.0236)		(0.0181)	(0.0177)
loyalty				0.5596***		0.1652***
				(0.0219)		(0.0209)
R-squared	0.54	0.56	0.30	0.31	0.59	0.61
Adj. R-squared	0.54	0.56	0.30	0.31	0.59	0.60
Log-Likelihood:	-1479.33	-1452.41	-1782.84	-1766.66	-1397.24	-1366.43
F-statistic:	1680.86	899.33	121.02	653.32	292.81	275.09
No. observations	1435	1435	1435	1435	1435	1435

Standard errors in parentheses
* p<.1, ** p<.05, ***p<.01

Table 13: Multiple Regression Active Characteristics Revenue

Impact of vendor characteristics on revenue

The regression model with revenue as a dependent variable shows a big difference in the explained variance by the control variables in *Model 1* and *Model 2*, as compared to the regression with sales as dependent variable. This difference is attributable to the fact that a feedback, does not necessarily translate to a high value transaction. For example, vendor A receives many feedbacks and makes little revenue, while vendor B receives few feedbacks and makes a lot of revenue. The difference could be in the kind of product sold, or the price of the listing. In addition to *Model 1*, in *Model2* 2% of the variance is explained purely by negative feedbacks. *Model 3* only includes the active characteristics as independent variables. However, as opposed to the regression model with sales as dependent variable, the regression coefficient for multihoming is significant in predicting revenue. Apart from this difference, the rest of the regression coefficients are similar to those in Table 11

Model 6 includes the effects of all passive characteristics and active characteristics. Similar to the sales regression, the regression coefficient for experience changes sign, but in this case its impact strength decreases. This indicates that when (positive) ratings are included as a independent variable, the effect of experience on sales is negative: meaning that if experience increases, while the amount of (positive) ratings remains the same, this has a negative effect on revenue. When including the active characteristics into the regression, an additional 3% of the variance is explained.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	-0.0000	-0.0000	-0.1195*	-0.0000	0.0012	-0.0037
	(0.0179)	(0.0176)	(0.0662)	(0.0219)	(0.0460)	(0.0451)
log_rpos	0.7347***	0.8002***			0.8044***	0.6966***
	(0.0179)	(0.0197)			(0.0208)	(0.0245)
log_rneg		-0.1457***			-0.1402***	-0.1239***
		(0.0197)			(0.0199)	(0.0196)
C(cluster)[T.1]			0.1626**		0.1166**	0.1206**
			(0.0805)		(0.0557)	(0.0545)
C(cluster)[T.2]			0.5958***		-0.0879	-0.0850
			(0.0897)		(0.0652)	(0.0638)
C(cluster)[T.3]			-0.0001		-0.0699	-0.0706
			(0.0908)		(0.0630)	(0.0617)
C(cluster)[T.4]			-0.4060***		-0.0913	-0.0715
			(0.0939)		(0.0655)	(0.0642)
C(cluster)[T.5]			0.3665***		0.0484	0.0558
			(0.1180)		(0.0830)	(0.0813)
loyalty				0.5596***		0.1697***
				(0.0219)		(0.0215)
R-squared	0.54	0.56	0.09	0.31	0.56	0.58
Adj. R-squared	0.54	0.56	0.09	0.31	0.56	0.58
Log-Likelihood:	-1479.33	-1452.41	-1969.90	-1766.66	-1440.04	-1409.45
F-statistic:	1680.86	899.33	27.66	653.32	264.05	248.71
No. observations	1435	1435	1435	1435	1435	1435

Standard errors in parentheses
* p<.1, ** p<.05, ***p<.01

Table 14: Multiple Regression Vendor Profiles, Revenue

Impact of vendor profiles on revenue

Model 3 consists of the clusters, the vendor profiles, and the Intercept represents the standardized mean for the cluster 0. The coefficients for the other clusters are relative to cluster 0. The sequence of clusters from strongest to weakest impact is: cluster 2; cluster 4; cluster 5; cluster 0; cluster 2; and cluster 1. With cluster 0 having a negative impact. Cluster 3 is not significantly different from the intercept. This is slightly different from the sales regression model.

The final model, *Model 6*, shows that there is no variance uniquely explained by including the clusters as independent variables. In this regression model, the regression coefficients for the clusters are insignificant, except for cluster 1. It is likely due to (positive) feedback being the confounding variable, explaining a lot of the relation between the other clusters and revenue. This leads to the same conclusion as the sales regression model: the positive effect of clusters on revenue is mainly due to the presence of ratings. Cluster 1 has a positive and significant impact on revenue: this indicates that this vendor profile has a positive relation with revenue

8.3 Monthly OLS Regression to Assess Change in Impact

The aim of the regression models in this subsection is to assess the change in relative impact of vendor characteristics over time. Vendor strategies are not included in these analyses, as the clustering procedure in section 7 did not take time into account. However, because these models do take time into consideration, the passive characteristics are lagged to simulate the feedback loop from the conceptual model discussed in section 4.1.1 and displayed in Figure 3. The first and second paragraph assess the change in impact on sales, and revenue, respectively. This is achieved by performing an OLS regression for every month and plotting the resulting β coefficients. Although both paragraphs aim to do the same, a difference between between the sales and revenue models exist. Due to missing revenue data, the regression output for predicting revenue from month 1 to 7 are unreliable.

8.3.1 Impact on Performance over Time

For every month a OLS regression model is estimated. Due to limited sample size in first 3 months, the results of these models are not significant. The regression formula is given below.

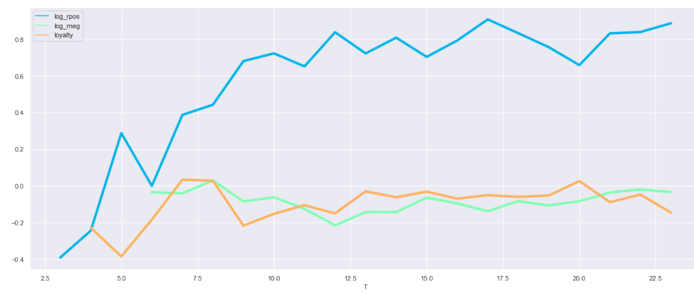
$$
\begin{aligned}
\log(\text{Sales})|\log(\text{Revenue}) = \alpha &+ \beta_1 \times \log(\text{Exposure}) \\
&+ \beta_2 \times \text{CategoryDiversity} \\
&+ \beta_3 \times \text{Experience} \\
&+ \beta_4 \times \text{ForumActivity} \\
&+ \beta_6 \times \text{Multihoming} \\
&+ \beta_7 \times \text{RatingPositive} \\
&+ \beta_8 \times \text{RatingNegative} \\
&+ \beta_9 \times \text{Loyalty} + \epsilon
\end{aligned}
$$

Results
Figure 27a and Figure 28a show the standardized β's of the active characteristics: The passive characteristics are left out of these plots to enable a better assessment of the relative impact change of the active characteristics. The plot clearly shows that exposure and forum activity have the strongest positive impact, and experience has the only negative impact. Experience starts in the first months having a positive impact on sales and revenue, while as time goes by the impact becomes negative. This likely due to vendors that remain active for a long time but fail to up their sale game and generate revenue.

Looking at Figure 27b and Figure 28b, it is evident that the relative weight of the positive ratings are really strong. This confirms the expectation and that ratings are most important for vendor performance. What also comes forward is the relative low impact of loyalty as compared to the positive ratings. Moreover, the positive rating line seems to stabilize during the second half of the markets lifetime. The standardized β for negative ratings seem to have a greater negative impact in the beginning than in the later stages in the market. This is possibly explained by the fact that vendors gain more ratings over time, so the effect of the negative ratings diminishes against the increasing rate of positive ratings: The fifth month shows a strong negative impact as opposed to very weak impact during the last couple of months.

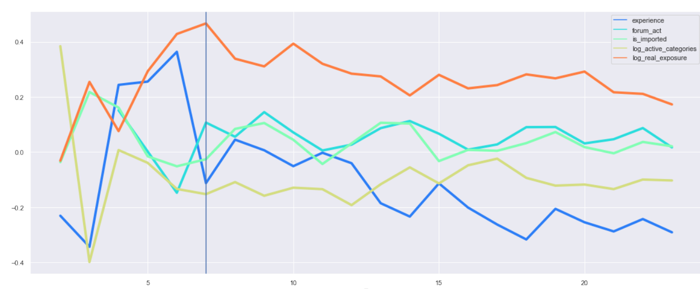
Interpreting the results this way offers mostly qualitative insights (see **??**). The full results of all regression models are found in the Appendix (section 11). The only independent variable that was insignificant almost every month was is_imported ($p > 0.10$).

(a) Active characteristics

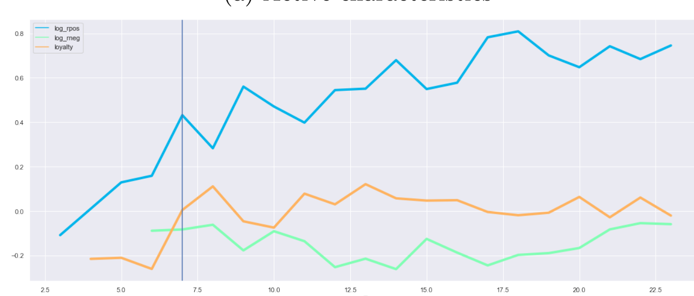


(b) Passive characteristics

Figure 27: Standardized Beta's Plotted over Time - Sales



(a) Active characteristics



(b) Passive characteristics

Figure 28: Standardized Beta's Plotted over Time - Revenue

8.4 Summary of Results

Impact of Characteristics and Strategies on Performance
The active characteristics exposure and experience have the biggest impact on sales and revenue. In the both final models, experience is negatively correlated with sales and revenue. This is interesting as it indicates that the positive relation between experience and performance and revenue is explained by positive ratings. Diversity is not rewarded: there is a slight advantage for vendors to be active in less categories, thus being specialized. Being active on the forum also slightly positively impacts sales, while having a negative impact on revenue. There was no effect measured of multihoming on sales. While multihoming tested to have a postive impact on revenue.
Moreover, it was found that vendors strategies can be used to predict performance and revenue. While the aim of these analyses was not to do exact predictions, the results from section 7 were confirmed.

Change of Impact over Time
Plotting the standardized betas for over time revealed that the relative impacts of active and passive characteristics did not change, except for Experience. The conclusion can be drawn that that a higher amount experience during later stages of the market have a negative impact on performance and revenue. The fact that the rest of the relative impacts did not change is somewhat surprising. This can also be due to the fact that only a small fraction of the vendors is responsible for a big fraction of the sales and revenue, which was already highlighted in Figure 13.

9 Discussion, Conclusion, and Recommendations

In this chapter, first the discussion of results and limitations takes place. Then, in the conclusion section the answers to the research questions are summarized, and final statements are made regarding this research. Finally, this chapter concludes with recommendations for further research.

9.1 Discussion

This section discusses the results of this research, together with its limitations and weaknesses. Also, suggestions are made to counter these limitations in further research on this topic.

9.1.1 Preparing the Data

Reconstructing Transactions
The back-end database provided rich information on the operation of market X and the interactions between its users. Due to the fact that the data was found to be incomplete and partly corrupted, a reconstruction of transactions was necessary in order to provide the data needed for analysis. On top of that, the reconstruction required missing data to be estimated.
For one, this research based the time of transaction on the time of purchase. However, the time of purchase was not known for almost half of the reconstructed transactions. This means that for more than 50 percent o of all transactions, the time of purchase is estimated, followed by estimations on the value in Bitcoin and dollars. Due to the volatile value of Bitcoin, wrongly estimated purchase times can impact the accuracy of revenue estimations.
Secondly, Market X did not registered the btc value of the transactions in the feedback tables, during the first four months of market X. A possible workaround was to estimate the price of these transactions by mapping the prices of listings. However, due to the listing information very limited: it was impossible to determine if listings changed or updated their prices, and therefore this approach was discontinued in this research. As a consequence, the revenue numbers in the first 5 months are highly underestimated. The effect of underestimation on market X's is diminishable however: in month 4 merely 273 transactions took place.

9.1.2 Darknet Market X

The total revenue of market X was estimated at $34.4 million, with the drugs segment around $30.3 million. With a monthly revenue in drugs of around $2.5 million, market X is a small market compared to the eight biggest darknet markets, where the monthly revenue was estimated at $14.2 million on average (Kruithof, Aldridge, Hétu et al., 2016. While the difference in generated revenue between markets is big, there are not substantially less active vendors on market X. (Paquet-Clouston et al., 2018;Aldridge and Décary-Hétu, 2016). The reconstructed transaction level data showed that 'drugs' is by far the most popular segment on market X, which is in line with previous research: showing drugs outperforming other segments on the dark web. Market X showed an almost constant growth rate over the course of 22 months. Moreover, the average value of a transaction increased from month 5 to month 22 with 52%. Also, descriptives of the transaction data showed that around 80% of the vendors never achieved to generate more than $10.000 in revenues on market X alone. Around 2% to 3% of the vendors finalized more than 1000 transactions and generated more than $100.000 in revenues. However, it was also shown that the top revenue generating vendors do not necessarily belong to the top transaction finalizing vendors. A vendor with 'only' 288 finalized transactions generating more than $1.1 million is exceptional and indicates that this vendor predominantly sold drugs in bulk.

9.1.3 Active Characteristics

Experience
While literature provided different metrics for measuring vendor experience on darknet markets (Soska

and Christin, 2015;Paquet-Clouston et al., 2018), this research used the same method as Christin (2012): only vendors were considered active in a month, when they finalized a transaction in that month. On average, the experience of vendors on market X was 147 days. This is almost half as many days of experience found in literature (Kruithof, Aldridge, Hétu et al., 2016; Soska and Christin, 2015). However, previous research worked with scrapes of front-end data, where the experience of vendors was based on feedbacks. The timeframes of these scrapes were often much smaller, than the timeframe used in this analysis, making the direct comparison of average vendor experience between different markets difficult to realize.
This research showed that the vendor characteristic experience, in the way it was defined in this report, does not have a positive relation with vendor performance. However, maybe academics that research darknet markets are lacking in constructing the right metric for experience. For example, experience could also be measured by the amount of days the alias of the vendor exists on the entire darkweb, i.e. the time between the first appearance of the alias and his latest feedback, or finalized transaction in the case of this research. Or by a more detailed version of the multi-homing characteristic: the total amount of darknet markets where the vendor has been active. In section 5.5 an alternative method for calculating experience is indicated. It is proposed to further research the meaning of vendor experience on darknet markets.

Exposure
Almost every paper that measures exposure, follows the same approach: exposure is measured by counting the number of *active* listings. These papers used scraping methods, which are accurate in detecting whether listings are active or not. During the data collection and construction of characteristics, there were some difficulties in estimating the measure for exposure: the back-end data does not provide information on the timeframe of listings being active. Through trial and error, ultimately a workaround was used that gave the most probable amount of active listings. Still, the high amount of active listings per month of some vendors, are suspected to be inaccurate: this research found that, on average, vendors had around 33 listings active per month. This is more than twice the amount as found in previous research (Paquet-Clouston et al., 2018; Kruithof, Aldridge, Hétu et al., 2016; Aldridge and Décary-Hétu, 2014). Moreover, the distribution of exposure was heavy tailed: a lot of outliers were present. As an example: one vendor had over 700 active listings active per month, on average.
Also, the difference of vendors using clone listings- and vendors not using them, was not taken into account. Clone listings are copies of listings, and are essentially the same, however differ in the quantity offered. The existence of this option might indicate that vendors want to expand their exposure, and at the same time keeping the rating of the original listing, to attract buyers.

Diversity
Initially, the characteristic of diversity was measured by two metrics: the diversity degree and the diversity in categories. On average, vendors have listings in 2 different categories in the drugs segment. However, the diversity degree shows that vendors are often more specialized in one category, meaning they have more listings in a single category. This is illustrated by a 0.4 score in diversity degree. The decision to include both, measuring essentially the same characteristic, was made due to both metrics are adopted by academics in different instances (Paquet-Clouston et al., 2018; Soska and Christin, 2015; Wegberg et al., 2019). Because the high correlation between the variables, only category diversity was included in the cluster and regression analyses. The metric for category only measured in the quantity of drugs categories a vendor was active in: it provides a high-level insight when measuring the impact on performance. A suggestion for a more sophisticated metric for diversity, that reveals more about the relation with vendor performance: one that also takes into account in what kind of categories the vendor is active in.

Forum Activity
This research was one of the first of its kind, that considered forum activity as a vendor characteristic on darknet markets. On average, it was found that vendors were active on the forum for around 21% of their lifetime on *market X*, while only 36% of all vendors were at least once active on the forum. Moreover, this research counted vendors as active on the forum, leaving out the content or quantity of posts. The initial idea by taking forum activity into consideration, was the hypothesis that it functions as an extension to exposure and reputation, and influences vendor performance. This is confirmed by the results, which show that forum activity positively impact vendor performance. However, not having examined the content or the number of forum posts, the strength of the effect is likely lower, than when the metric is more detailed: Therefore, for further research, more detailed characteristics considering forum activity are proposed. For example, a metric that only considers forum messages that resemble listings.

Multihoming
Interesting is the fact that on average 56% of the active vendors in a month, are or were also active on other markets. When considering active vendors from month 10 and onward, this percentage lies around 70%. The biggest increase in the presence of multihoming happens from month 3 to 4. This indicates that experienced vendors from other markets decided to enter market X, after it had gained a reputable and established status. A further interesting finding of this research, is that 80% of the 25 top performing vendors were active on other markets. This further indicates, that the same vendors are active on different darknet markets, and together are responsible for a substantial part of the drug trade. Interesting would be for future research to examine when and whether markets are seen as attractive for vendors to migrate or start their businesses.

9.1.4 Passive Characteristics

Loyalty
Customer loyalty is an important phenomenon on darknet markets. Vendors want to maintain their existing buyers, and can do so by remaining attractive by offering high quality of goods, or having a good reputation. In this research, a vendor's loyalty score is based on the average buyer loyalty per category. The descriptive statistics showed that the 25 top performing vendors, on average, retained a loyalty score of 65%. This research indicates that loyalty has a higher relative impact on revenue compared to other characteristics, than is the case of the performance model. This translates to customer loyalty being more important for generating revenue, than increasing sales numbers. In other words: loyalty is likely more associated with higher value transactions.
However, this way of measuring loyalty, has the risk of doing not any justice to the real meaning of loyalty. When a vendor is active in multiple categories, the weighted average of the different loyalty scores is taken as the final score. The limitation is that if a vendor has a lot of highly loyal buyers in category A, and a lot of less loyal buyers in category B, the overall score is brought down greatly. While in fact, a vendor achieves can achieve a high loyalty score in one category. Alternative ways of measuring loyalty could also be considered, by for example examining the customer retention rate.

Reputation
The metrics for reputation in this research were constructed by taking the positive and negative ratings. On market X, the cumulative sum of all postive ratings were shown on the vendor's profile page, or listing page. Therefore, the comparison between markets should be done very carefully, taking into account these possible differences between the appearance of ratings. On average, it was found that very few negative feedbacks are given, only 0.7% of the transactions. This is likely due to finalized orders, being completed: the goods are delivered, and any potential disputes are already resolved. Moreover, as vendors regard their reputation as the most valuable asset, it is unlikely that when they 'deliver' they will disappoint. The ratings were used as control variables in the regression analyses, because they are strongly correlated with sales numbers: a feedback has a near one-on-one relation

with a transaction. Also, it is important for determining the effects of, or changes in the direction and strength, of the active characteristics. In the end, the obvious needs to be controlled for: ratings, remain the most important asset for vendors to increase their sales numbers and total revenues.

9.1.5 Vendor Profiles

The clustering of vendors based on their active characteristics resulted in six different vendor profiles. Some of the profiles showed similarity with the vendor profiles found in previous research (van Wegberg et al., 2019). The main takeaway of this analysis is that vendors with strategy 2 or 5, on average, have more listings and are active in more categories, as compared to other clusters. However, compared to profile 2, profile 5 consists of vendors that have more experience and are less active on the forum. Both of profiles 2 and 5 also performed better relative to the other strategies. Strategies 0, 1 and 3 seem to perform equally well.
Agglomerative hierarchical clustering is but one of the many possible methods for clustering data. It should be researched what kind of clustering method is best applied on darknet vendor data, and on these metrics specifically. The k-means algorithm was trialed on the data, but resulted in sparsely populated clusters (N ¡ 30). On top of that, the formed clusters were not much different from each other. Future research should look into the use of latent class analysis (LCA) as an alternative clustering method. LCA is a model-based clustering method that arguably deals with all the disadvantages associated with traditional clustering algorithms, like hierarchical and k-means.

9.1.6 Impacts on Performance

The results from the different regression models revealed that exposure and experience have the biggest impact on performance, relative to other active characteristics. However, when reputation is introduced in the model, the effect of experience becomes negative. This confirms the idea that experience alone is no guarantee for performance on darknet markets. Moreover, a higher diversity in categories was found to negatively impact performance. This invalidates the assumption that a higher diversity in categories has a positive effect on performance, that the results from the cluster analysis indicated. An interesting result from the regression analyses is that forum activity does have an effect on sales, albeit very weak. Therefore, it would be interesting for future research to focus on the role of darknet market forums. This is also interesting for law enforcement agencies, as in most cases the same aliases are used as on the market platform.

Change in Impact over Time

The monthly regressions models delivered somewhat surprising results. It was expected that some characteristics would become more important as the market matured. However, only the impact of experience changed, from having a positive impact on performance in the first months to a negative impact in later months. At first, performing an OLS regression for every month and plotting the standardized beta's seemed as a good idea. Yet, during this research other regression methods came forward, that possible are better alternatives for inspecting the change over time. This is mainly due to the 'monthly vendor' data having a panel structure. Therefore, future research that aims to assess whether, and how, time influences the effects of characteristics on performance: it should look into adopting either Fixed effects (FE) or mixed effects (ME) regression, as these are more suitable methods for performing regression analysis on panel-data.

9.2 Conclusion

In the previous section, the discussion treated the results extensively, and posed limitations of this research. Therefore, this section provides a high-level overview on the results and insights gained in this study in the form of conclusive answers on the sub research questions. The section concludes with answering the main research question and summarizing the scientific and societal relevance of this research.

9.2.1 Answering the Sub Research Questions

In this section, all sub research questions of this research are discussed. The sub questions were structured in such a way that every question is answered by an analysis of some kind.

What metrics are considered to indicate vendor performance on darknet markets?
The answer to this question provides insights in what metrics can be used to measure vendor performance on darknet markets. Through desk-research and literature review, it was found that the total amount of sales or revenue, is the predominantly used in literature on darknet markets as metrics for meausuring vendor performance. This is mainly due to sales resulting in feedbacks. In turn, feedbacks result in reputation and trust on the marketplace. Trust and reputation are, as found by the literature review, highly valuable assets an vendor has on darknet markets. This due to darknet markets being completely anonymous and build around the feedback mechanisms, which in turn safeguards trust.

What vendor characteristics are considered to have an effect on performance and to what degree are they present on a darknet market?
The literature review treated seven different vendor characteristics, most of which are predominantly researched in the field of darknet markets. Some of the characteristics are but recently introduced in the research field, offering room for improvement. The elaboration on these characteristics, together with the methods for constructing meaningful metrics, formed the basis for the answer to the second part of this question. A descriptive analysis was used to assess the distribution of characteristics among the vendors on market X. Additionally, this research examined the features of the top 25 performing vendors. The values for characteristics are evenly distributed among vendors, and show mixed values, which even holds for the top 25 performing vendors.

What profiles of vendors can be identified, and how do these types perform relative to each other?
The answers from the previous sub questions form the basis for answering this research question. The cluster analysis uses the active characteristics as features for assessing similarity with vendors. In this way, vendor profiles submerge from the data. The vendor profiles are important for identifying different types of vendors and offer insights in what the dominant distributions and compositions of characteristics are, among vendors. Comparing the vendor profiles in terms of performance through a pair-wise comparison, it was possible to determine that some vendor profiles performed better than others. The most important part of the answer to this research question is that although some profiles appear to perform better, there is no clear-cut in what the superior vendor profile would be.

What are the effects of vendor characteristics and vendor profiles on vendor performance?
The answer to this question was found by performing several multiple regressions. By using separate regression models for vendor profiles and vendor characteristics respectively, the effects on performance were measured. The regression coefficients were standardized with the aim of comparing the strength of the regression coefficients. Ratings functioned as control variables to determine whether the profiles had an additional effect on performance. For three of the six clusters it was found that belonging to those clusters, has an additional and positive impact on performance in terms of sales,

holding all other effects constant. While only one profile had an additional positive effect on revenue. The effects of vendor characteristics on performance indicated that multihoming is merely of importance, despite the fact that most of the top performing vendors possess this characteristic. For the other characteristics, it became clear what their relative importance for vendor performance is, when rating is included into the model as control variable.

Do these effects change when the market grows and matures, and how do they change?
A rather extensive approach was chosen to answer this sub research question. This research performed a multiple regression for month, standardizing the regression coefficients and plotting the results together over time. This method enabled the assessment whether the relative importance of vendor characteristics for performance changed. The results indicated that the strength of impact changed for some characteristics. Only in the case of one characteristic, experience, it was found that effect turned from positive to negative, as the market grew.

9.2.2 Answering the main research question

To what extent do vendor characteristics impact vendor performance on darknet markets?
In order to gain a better understanding on the modus-operandi of vendors on darknet market, this research examined and assessed the degree to which particular vendor characteristics have an impact on performance. By reconstructing transactions, not only the interactions between vendors and buyers were revealed, but also the different vendor profiles, that emerged from the data after a cluster analysis. Every analysis offered new insights in the mechanisms of darknet markets, and in particular revealed how vendors operate in the drugs segment. Exposure and experience came forward as being the most important indicators for performance. With experience negatively impacting performance as the market ages. Other characteristics also showed impact, but their relative importance was rather low. Taken together, this research provided an peek into the world of darknet vendors, revealing how they operate and what factors are proven to be important for improving vendor performance.

9.2.3 Scientific Contribution

This research was the first in its kind to reconstruct and assess the transaction level data of the entire lifetime of a darknet market. With this data, a longitudinal analysis was performed to assess the distribution and development of characteristics among vendors over time. Moreover, this research confirmed that experience is no guarantee for improving vendor performance. Additionally, the introduction of forum activity as a vendor characteristic was found to provide new insights. However, further refinement of the metric could lead to even more substantial results.
Another contribution to the literature is the approach of clustering vendors into profiles. Where previous research included feedbacks or performance as a feature, to group vendors on darknet markets, this research only included the so-called active characteristics as features. In this sense, the profiles describe the different behaviours, or strategies, of vendors on darknet markets. This helps to understand the social-economic motivations of darknet vendors even more.

9.3 Recommendations

Given the discussion of results and the posed limitations of this research in section 9.1, future research is proposed both of the following directions:

- Further research using the same (reconstructed) data
- Further research on Darknet Markets in General

9.3.1 Further Research Using the Same (Reconstructed) Data

This research predominantly focused on vendors and their behaviour on darknet markets. In literature, this is equally the case. While the assessment and examination of buyers and their behaviour on darknet markets is underrepresented in research. For one, the same clustering method could be adopted to identify different types of buyers. But first, metrics for buyer-specific characteristics should be designed and constructed, by answering questions like:

1. What kind of distinctive buying behaviour is present among buyers?
2. Are there occasional buyers, frequent buyers or bulk buyers present?
3. How much time is there between a first and second purchase, for repeat buyers?
4. Why do repeat buyers switch vendors?

Another suggestion for further research on this dataset are vendor-buyer relations. Having information the counter-parties of almost every transaction, different directions of research are possible.

A third suggestion for further research would be to test different clustering methods on the dataset and evaluate what method gives the most fitting results. This paper used the hierarchical clustering method. However the clustering resulted into clusters with many outliers and heavy tailed distributions. Model-based clustering techniques should be considered due to its ability to provide estimates on the likelihood an individual belongs to a cluster. Moreover, it enables the comparison in goodness of fit of different cluster solutions. One such method is Latent Class Analysis (LCA).

9.3.2 Further Research on Darknet Markets in General

Darknet Markets in General

One suggestion for further research, is to determine the role and functions of darknet market forums for vendors. How does the communication on these boards relate to the market activity on the platform? Social Network Analysis could be applied to determine which users fulfill important positions on the forum. These results could then be compared and linked to the behaviour of those users on the darknet markets. Doing this, likely provides more insights on the link between the forum and the market.

9.4 Perspectives for Law Enforcement

9.4.1 Vendor-Targeted Interventions

One of the aims of this research was to provide law enforcement agencies a better understanding of how vendors operate on darknet markets. Law enforcement agencies continuously seek new approaches and strategies for effectively intervening on darknet markets. Currently, the strategies that target the vendor instead of the platform are the topic of discussion. By focusing on vendors, the impact of interventions has a potentially longer lasting effect. Below, some key findings of this research are translated to takeaways for law enforcement agencies.

Broaden the Scope on the Dark Web
One outcome of this study found that nearly 80 per cent of the actively selling vendors are or were also active in other markets. This finding indicates that targeting markets is not the (only) key to bringing damage to the ecosystem. Law Enforcement Agencies should broaden their scope and not limit their scope to one market at a time. Instead, vendors should be targeted based on their activity on the entire dark web. The fact that a significant share of active vendors, even 20 of the top 25 performing vendors, consist of multihoming vendors, implies that these vendors are well experienced and able to survive the attempts of law enforcement agencies, to detect and track their activity.
Currently, some firms provide services that monitor the dark web, and pick up on specific (predefined) information. One example is a company that provides a proactive monitoring service for firms with stolen or compromised credentials. The service alerts the clients whenever it discovers the stolen data on the dark web[5].
Another relevant insight is the fact that vendors can use the forum to improve their performance on darknet markets. This suggests that law enforcement can use the information on publicly available forums to gather more information about potential targets.
Law enforcement agencies should focus on developing services or tools, that continuously scrapes the dark web (only) on popular vendor aliases. In this way, law enforcement agencies build up records of vendors, registering upcoming and important vendors, and at the same time, enlarge the possibility of detecting some activity that could eventually lead to their personal identity. Simply said: a dark web version of Google Trends[6].

Using Profiles to Target Vendors
This study found that vendor segmentation resulted in six different vendor profiles. Two of those profiles performed, on average, better than the other profiles. For starters, Law enforcement agencies can use these profiles to narrow their focus, by targeting only vendors that exhibit these characteristics. Because, by only targeting the already established vendors, through focusing on their reputation and feedbacks, small and upcoming vendors remain undetected.

Reconstruction of Data
A final contribution of this research for law enforcement is the way the transaction-level data is reconstructed. Whereas the original database consisted of mostly incomplete or canceled orders, there were around 150.000 finalized orders present. The reconstruction of the database reconstructed around 179.000 finalized orders. While scraping is still predominantly used, the process for reconstructing transaction-level data can be used by law enforcement agencies for similar cases to get a better picture of the activity on markets.

[5] https://www.momentumit.com/dark-web-monitoring
[6] https://trends.google.com/trends/

References

Aldridge, J. & Décary-Hétu, D. (2014). Not an 'Ebay for Drugs': The Cryptomarket 'Silk Road' as a Paradigm Shifting Criminal Innovation. SSRN Electronic Journal. doi:10.2139/ssrn.2436643

Aldridge, J. & Décary-Hétu, D. (2016). Hidden wholesale: The drug diffusing capacity of online drug cryptomarkets. International Journal of Drug Policy, 35, 7–15. doi:10.1016/j.drugpo.2016.04.020

Anil, J. K. (2008). Data Clustering: 50 Years Beyond K-means, 26–27.

Bouchard, M. & Nguyen, H. (2011). Professionals or amateurs? revisiting the notion of professional crime in the context of cannabis cultivation. World wide weed: global trends in cannabis cultivation and its control, Ashgate, Farnham, 109–126.

Broséus, J., Rhumorbarbe, D., Mireault, C., Ouellette, V., Crispino, F. & Décary-Hétu, D. (2016). Studying illicit drug trafficking on Darknet markets: Structure and organisation from a Canadian perspective. Forensic Science International, 264, 7–14. doi:10.1016/j.forsciint.2016.02.045

Buxton, J. & Bingham, T. (2015). The Rise and Challenge of Dark Net Drug Markets, 24.

Chicco, G., Napoli, R. & Piglione, F. (2006). Comparisons among clustering techniques for electricity customer classification. IEEE Transactions on Power Systems, 21(2), 933–940. doi:10.1109/TPWRS.2006.873122

Christin, N. (2012). Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. arXiv:1207.7139 [cs]. arXiv: 1207.7139. Retrieved February 28, 2019, from http://arxiv.org/abs/1207.7139

Commission, E. & Europol (Eds.). (2017). Drugs and the darknet: Perspectives for enforcement, research and policy. OCLC: 1027963810. Luxembourg: Publications Office of the European Union.

Décary-Hétu, D. & Giommoni, L. (2017). Do police crackdowns disrupt drug cryptomarkets? A longitudinal analysis of the effects of Operation Onymous. Crime, Law and Social Change, 67(1), 55–75. doi:10.1007/s10611-016-9644-4

Décary-Hétu, D. & Leppänen, A. (2016). Criminals and signals: An assessment of criminal performance in the carding underworld. Security Journal, 29(3), 442–460. doi:10.1057/sj.2013.39

Décary-Hétu, D. & Quessy-Doré, O. (2017). Are Repeat Buyers in Cryptomarkets Loyal Customers? Repeat Business Between Dyads of Cryptomarket Vendors and Users. American Behavioral Scientist, 61(11), 1341–1357. doi:10.1177/0002764217734265

Duijn, P. & Klerks, P. (2014). De brug tussen wetenschap en opsporingspraktijk: Onderzoek naar de toepassing van sociale netwerkanalyse in de opsporing. Tijdschrift voor Criminologie, 56(4). doi:10.5553/TvC/0165182X2014056004003

Duxbury, S. W. & Haynie, D. L. (2018). The Network Structure of Opioid Distribution on a Darknet Cryptomarket. Journal of Quantitative Criminology, 34(4), 921–941. doi:10.1007/s10940-017-9359-4

Everitt, B. (2011). Cluster Analysis, 348.

Hexmoor, H. (2015). Chapter 6 - diffusion and contagion. In H. Hexmoor (Ed.), Computational network science (pp. 45–64). Emerging Trends in Computer Science and Applied Computing. Boston: Morgan Kaufmann. doi:https://doi.org/10.1016/B978-0-12-800891-1.00006-8

Hout, M. C. V. & Bingham, T. (2013). 'Surfing the Silk Road': A study of users' experiences. International Journal of Drug Policy, 24(6), 524–529. doi:10.1016/j.drugpo.2013.08.011

Kruithof, K., Aldridge, J., Décary-Hétu, D., Sim, M., Dujso, E. & Hoorens, S. (2016). Internet-facilitated drugs trade. An analysis of the size, the scope and the role of the Netherlands. Santa Monica/Cambridge: RAND.

Kruithof, K., Aldridge, J., Hétu, D., Sim, M., Dujso, E. & Hoorens, S. (2016). Internet-facilitated drugs trade: An analysis of the size, scope and the role of the Netherlands. RAND Corporation. doi:10.7249/RR1607

Markoff, J. (2005). What the dormouse said: How the 60s counterculture shaped the personal computer. Viking Adult.

McCoy, D., Bauer, K., Grunwald, D., Kohno, T. & Sicker, D. (2008). Shining Light in Dark Places: Understanding the Tor Network. In N. Borisov & I. Goldberg (Eds.), Privacy Enhancing Technologies (Vol. 5134, pp. 63–76). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-70630-4_5

Morselli, C. (2009). Inside Criminal Networks. Studies of Organized Crime. New York, NY: Springer New York. doi:10.1007/978-0-387-09526-4

Neal, W. D. (1999). Satisfaction is nice, but value drives loyalty. Marketing Research, 11(1), 20–23. Copyright - Copyright American Marketing Association Spring 1999; Last updated - 2014-05-24; SubjectsTermNotLitGenreText - US. Retrieved from https://search.proquest.com/docview/202676890?accountid=27026

Paquet-Clouston, M., Décary-Hétu, D. & Morselli, C. (2018). Assessing market competition and vendors' size and scope on AlphaBay. International Journal of Drug Policy, 54, 87–98. doi:10.1016/j.drugpo.2018.01.003

Soska, K. & Christin, N. (2015). Measuring the Longitudinal Evolution of the Online Anonymous Marketplace Ecosystem, 17.

Tripathi, S., Bhardwaj, A. & E, P. (2018). Approaches to Clustering in Customer Segmentation. International Journal of Engineering & Technology, 7(3.12), 802. doi:10.14419/ijet.v7i3.12.16505

Tzanetakis, M., Kamphausen, G., Werse, B. & von Laufenberg, R. (2016a). The transparency paradox. Building trust, resolving disputes and optimising logistics on conventional and online drugs markets. International Journal of Drug Policy, 35, 58–68. doi:10.1016/j.drugpo.2015.12.010

Tzanetakis, M., Kamphausen, G., Werse, B. & von Laufenberg, R. (2016b). The transparency paradox. building trust, resolving disputes and optimising logistics on conventional and online drugs markets. International Journal of Drug Policy, 35, 58–68. Drug Cryptomarkets. doi:https://doi.org/10.1016/j.drugpo.2015.12.010

United Nations Office on Drugs and Crime. (2010). The globalization of crime: A transnational organized crime threat assessment. OCLC: 654315644. Vienna: United Nations Office on Drugs and Crime.

Varghese, B. M. (2013). SPATIAL CLUSTERING ALGORITHMS - AN OVERVIEW, 8.

Verburgh, T., Smits, E. & van Wegberg, R. (2018). Uit de schaduw: Perspectieven voor wetenschappelijk onderzoek naar dark markets. Justitiële verkenningen, 44(5), 68–82. doi:10.5553/JV/016758502018044005006

Wittek, P. (2014). 5 - unsupervised learning. In P. Wittek (Ed.), Quantum machine learning (pp. 57–62). Boston: Academic Press. doi:https://doi.org/10.1016/B978-0-12-800953-6.00005-0

10 APPENDIX A: Academic Article

includepdf[pages=-]Paper$_{D}arknet_{V}endors.pdf$

11 Appendix B: Regression Results

11.1 Multiple Regression Analyses on 'average vendor' data

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	-0.3126***	-0.3348***	0.1012	0.8584***	-0.6395***	-0.6484***
	(0.0302)	(0.0304)	(0.0889)	(0.0456)	(0.0396)	(0.0389)
log_rpos	0.7901***	0.8116***			0.9185***	0.8742***
	(0.0095)	(0.0106)			(0.0118)	(0.0131)
log_rneg		-0.1505***			-0.1422***	-0.1202***
		(0.0333)			(0.0288)	(0.0284)
log_real_exposure			0.5213***		0.1453***	0.1424***
			(0.0331)		(0.0148)	(0.0145)
active_categories			-0.1509***		-0.0385***	-0.0395***
			(0.0296)		(0.0129)	(0.0126)
experience			0.0029***		-0.0023***	-0.0023***
			(0.0002)		(0.0001)	(0.0001)
forum_act			0.4684***		0.0622*	0.0662*
			(0.0880)		(0.0376)	(0.0370)
is_imported			0.1031		-0.0020	-0.0018
			(0.0637)		(0.0271)	(0.0266)
loyalty				2.7582***		0.3704***
				(0.0966)		(0.0518)
R-squared	0.83	0.83	0.33	0.36	0.88	0.88
Adj. R-squared	0.83	0.83	0.33	0.36	0.88	0.88
Log-Likelihood:	-1232.48	-1222.28	-2206.95	-2172.72	-973.06	-947.84
F-statistic:	6905.51	3509.97	141.83	815.91	1499.03	1364.02
No. observations	1435	1435	1435	1435	1435	1435

Standard errors in parentheses
* p<.1, ** p<.05, ***p<.01

Table 15: Multiple Regression Active Characteristics, Sales, Unstandardized

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	-0.3126***	-0.3348***	1.7828***	0.8584***	-0.2736***	-0.2951***
	(0.0302)	(0.0304)	(0.0911)	(0.0456)	(0.0454)	(0.0449)
log_rpos	0.7901***	0.8116***			0.8447***	0.7980***
	(0.0095)	(0.0106)			(0.0108)	(0.0128)
log_rneg		-0.1505***			-0.1148***	-0.0926***
		(0.0333)			(0.0325)	(0.0322)
C(cluster)[T.1]			0.0027		-0.0855*	-0.0828*
			(0.1107)		(0.0460)	(0.0453)
C(cluster)[T.2]			0.7723***		-0.4953***	-0.4933***
			(0.1234)		(0.0538)	(0.0530)
C(cluster)[T.3]			-0.0735		-0.2392***	-0.2397***
			(0.1249)		(0.0520)	(0.0513)
C(cluster)[T.4]			-0.5870***		-0.0382	-0.0246
			(0.1292)		(0.0540)	(0.0533)
C(cluster)[T.5]			0.6084***		-0.0595	-0.0545
			(0.1623)		(0.0685)	(0.0676)
loyalty				2.7582***		0.3880***
				(0.0966)		(0.0594)
R-squared	0.83	0.83	0.09	0.36	0.84	0.85
Adj. R-squared	0.83	0.83	0.09	0.36	0.84	0.85
Log-Likelihood:	-1232.48	-1222.28	-2427.64	-2172.72	-1164.47	-1143.33
F-statistic:	6905.51	3509.97	28.60	815.91	1100.31	996.19
No. observations	1435	1435	1435	1435	1435	1435

Standard errors in parentheses
* p<.1, ** p<.05, ***p<.01

Table 16: Multiple Regression Vendor Profiles, Sales, Unstandardized

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	3.4207***	3.3329***	3.6043***	4.6494***	2.6837***	2.6565***
	(0.0738)	(0.0734)	(0.1360)	(0.0707)	(0.1094)	(0.1071)
log_rpos	0.9520***	1.0368***			1.0330***	0.8978***
	(0.0232)	(0.0255)			(0.0325)	(0.0362)
log_rneg		-0.5939***			-0.6051***	-0.5378***
		(0.0803)			(0.0795)	(0.0783)
log_real_exposure			0.7539***		0.3521***	0.3432***
			(0.0506)		(0.0408)	(0.0399)
active_categories			-0.2615***		-0.1005***	-0.1037***
			(0.0452)		(0.0355)	(0.0348)
experience			0.0039***		-0.0016***	-0.0016***
			(0.0003)		(0.0003)	(0.0003)
forum_act			0.3482***		-0.0755	-0.0633
			(0.1347)		(0.1039)	(0.1018)
is_imported			0.3110***		0.2195***	0.2199***
			(0.0975)		(0.0748)	(0.0732)
loyalty				3.8244***		1.1289***
				(0.1496)		(0.1427)
R-squared	0.54	0.56	0.30	0.31	0.59	0.61
Adj. R-squared	0.54	0.56	0.30	0.31	0.59	0.60
Log-Likelihood:	-2513.84	-2486.92	-2817.35	-2801.17	-2431.75	-2400.94
F-statistic:	1680.86	899.33	121.02	653.32	292.81	275.09
No. observations	1435	1435	1435	1435	1435	1435

Standard errors in parentheses
* p<.1, ** p<.05, ***p<.01

Table 17: Multiple Regression Active Characteristics, Revenue, Unstandardized

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	3.4207***	3.3329***	5.7958***	4.6494***	3.3148***	3.2508***
	(0.0738)	(0.0734)	(0.1361)	(0.0707)	(0.1132)	(0.1112)
log_rpos	0.9520***	1.0368***			1.0422***	0.9026***
	(0.0232)	(0.0255)			(0.0270)	(0.0318)
log_rneg		-0.5939***			-0.5714***	-0.5050***
		(0.0803)			(0.0810)	(0.0798)
C(cluster)[T.1]			0.3343**		0.2398**	0.2479**
			(0.1655)		(0.1145)	(0.1121)
C(cluster)[T.2]			1.2251***		-0.1807	-0.1748
			(0.1844)		(0.1341)	(0.1313)
C(cluster)[T.3]			-0.0001		-0.1437	-0.1451
			(0.1867)		(0.1296)	(0.1269)
C(cluster)[T.4]			-0.8348***		-0.1877	-0.1471
			(0.1932)		(0.1346)	(0.1320)
C(cluster)[T.5]			0.7536***		0.0996	0.1147
			(0.2426)		(0.1708)	(0.1672)
loyalty				3.8244***		1.1594***
				(0.1496)		(0.1471)
R-squared	0.54	0.56	0.09	0.31	0.56	0.58
Adj. R-squared	0.54	0.56	0.09	0.31	0.56	0.58
Log-Likelihood:	-2513.84	-2486.92	-3004.41	-2801.17	-2474.55	-2443.96
F-statistic:	1680.86	899.33	27.66	653.32	264.05	248.71
No. observations	1435	1435	1435	1435	1435	1435

Standard errors in parentheses
* p<.1, ** p<.05, ***p<.01

Table 18: Multiple Regression Vendor Profiles, Revenue, Unstandardized

11.2 Multiple Regression Analyses on 'monthly vendor' data

	T	r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual
1	1	1.00	1.00	0.00	Inf	0.00	3	Inf	-Inf	-Inf	0.00	1
2	2	0.99	0.90	0.26	11.34	0.22	8	9.34	-0.69	1.09	0.07	1
3	3	0.73	0.54	0.49	3.81	0.03	8	-7.45	32.91	40.92	2.41	10
4	4	0.87	0.84	0.39	37.78	0.00	9	-22.01	64.01	84.08	7.17	46
5	5	0.83	0.81	0.45	46.97	0.00	8	-44.07	106.15	127.24	14.16	69
6	6	0.85	0.84	0.47	69.17	0.00	9	-67.96	155.92	182.75	22.26	99
7	7	0.80	0.79	0.54	67.66	0.00	9	-110.60	241.20	270.97	39.03	136
8	8	0.74	0.72	0.65	55.08	0.00	9	-161.40	342.79	373.97	67.56	158
9	9	0.83	0.83	0.55	126.88	0.00	9	-171.92	363.83	397.44	62.65	204
10	10	0.78	0.77	0.61	112.29	0.00	9	-240.91	501.82	537.66	95.30	257
11	11	0.79	0.78	0.62	136.80	0.00	9	-283.03	586.06	623.23	114.57	295
12	12	0.76	0.75	0.67	118.20	0.00	9	-311.16	642.32	679.75	134.26	303
13	13	0.70	0.69	0.80	90.46	0.00	9	-384.84	789.68	827.49	204.06	315
14	14	0.75	0.74	0.78	123.74	0.00	9	-390.78	801.56	839.82	199.06	330
15	15	0.75	0.75	0.77	136.43	0.00	9	-419.78	859.56	898.56	213.19	356
16	16	0.71	0.71	0.82	121.79	0.00	9	-482.12	984.23	1024.12	261.83	390
17	17	0.75	0.74	0.78	170.79	0.00	9	-548.61	1117.22	1158.77	283.31	462
18	18	0.72	0.72	0.83	166.34	0.00	9	-627.61	1275.22	1317.66	345.02	506
19	19	0.72	0.72	0.81	181.11	0.00	9	-691.30	1402.60	1446.10	374.62	564
20	20	0.74	0.73	0.82	209.06	0.00	9	-737.30	1494.60	1538.67	404.37	597
21	21	0.71	0.71	0.83	183.25	0.00	9	-730.79	1481.57	1525.48	405.32	587
22	22	0.76	0.76	0.79	265.28	0.00	9	-780.71	1581.41	1626.42	406.62	657
23	23	0.73	0.72	0.84	222.41	0.00	9	-848.10	1716.21	1761.46	480.23	673

Table 19: Multiple Regression per Month, Sales

	T	r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual
1	1	0.84	0.51	1.39	2.55	0.40	3	-4.21	16.42	13.97	1.92	1
2	2	0.44	-3.51	1.90	0.11	0.98	8	-8.64	35.28	37.05	3.59	1
3	3	0.45	0.07	1.69	1.17	0.40	8	-29.72	77.44	85.46	28.64	10
4	4	0.14	-0.02	2.59	0.90	0.53	9	-125.55	271.11	291.18	309.52	46
5	5	0.49	0.44	1.34	9.52	0.00	8	-127.70	273.41	294.50	124.33	69
6	6	0.60	0.57	1.20	18.89	0.00	9	-167.83	355.67	382.49	141.50	99
7	7	0.59	0.56	1.12	24.03	0.00	9	-217.78	455.57	485.33	171.19	136
8	8	0.58	0.56	1.03	27.32	0.00	9	-237.51	495.02	526.20	168.10	158
9	9	0.65	0.64	1.12	47.89	0.00	9	-321.35	662.70	696.31	254.87	204
10	10	0.60	0.58	1.25	47.57	0.00	9	-432.76	885.52	921.36	403.21	257
11	11	0.58	0.57	1.23	51.18	0.00	9	-488.51	997.01	1034.18	442.75	295
12	12	0.59	0.58	1.21	54.32	0.00	9	-498.73	1017.45	1054.88	446.79	303
13	13	0.53	0.51	1.43	43.73	0.00	9	-571.86	1163.73	1201.53	647.35	315
14	14	0.58	0.57	1.37	57.07	0.00	9	-582.81	1185.61	1223.87	618.02	330
15	15	0.57	0.56	1.32	59.59	0.00	9	-616.07	1252.14	1291.14	625.00	356
16	16	0.52	0.51	1.50	53.59	0.00	9	-724.56	1469.12	1509.01	882.68	390
17	17	0.60	0.59	1.33	87.26	0.00	9	-798.43	1616.85	1658.40	818.38	462
18	18	0.61	0.60	1.33	98.08	0.00	9	-874.49	1768.98	1811.43	899.98	506
19	19	0.56	0.55	1.40	90.17	0.00	9	-1000.80	2021.61	2065.12	1103.47	564
20	20	0.59	0.58	1.36	105.41	0.00	9	-1043.43	2106.87	2150.93	1110.62	597
21	21	0.59	0.58	1.34	105.66	0.00	9	-1016.85	2053.70	2097.60	1058.51	587
22	22	0.62	0.62	1.28	134.69	0.00	9	-1107.09	2234.18	2279.19	1083.56	657
23	23	0.58	0.58	1.33	117.57	0.00	9	-1157.90	2335.79	2381.04	1191.23	673

Table 20: Multiple Regression per Month, Revenue

12 APPENDIX C: Python Scripts

12.1 Reconstruction of Transactions

12.2 Construction of Characteristics

12.3 Hierarchical Clustering Script