**Module Six Critical Thinking**

**Option One: Capstone Project Rough Draft: US Organization**

**Andrew Arendt**

**Colorado State University Global**

**MIS581: Capstone - Business Intelligence and Data Analytics**

**Dr. Kimberly Ford**

**Due 5/2/2021**

**Contents**

Title Page………………………………………………………………………………………….1

Table of Contents………………………………………………………………………………….2

Introduction and Objectives………………………………….…………………………………....3

Explanation of Hypotheses………………………………………………………………………..4

Methodology……………….……………………………………………………………………...5

Ethics Concerns and Considerations……………………………………………………………....6

Conclusions………………………………………………………………………………..…..…..8

References……………………………………………………………………………..……..…....9

Throughout the scope of this study, it is intended that a deeper understanding of the sales patterns of electronics through the Amazon sales platform. Amazon is a top US based online retailer specializing in home delivery. Its online presence is touted as among the world’s most advanced and it is number two on the Fortune 500 as of 2020 (Murray & Meyer, 2020). It therefore can be stated that understanding the electronics sales done through Amazon’s platform can be indicative of the greater market. Using the data from this study, conclusions can be drawn about the electronics market and current industry trends and customer opinions.

The objective of this study is to determine trends in the sales of electronics in the current year. Through the use of the metrics provided, commonly occurring data points which are high in the rankings or have high customer review averages will be singled out for analysis. The wealth of data in this data set will allow the study to have a comprehensive view of the patterns in sales data and highly ranked products for the purposes of marketing, production, and supplier relations. Electronics sales are a critical part of the retail market as a whole as they represent the cutting edge of modern consumer science. “As we enter the new year, several promising technologies are poised to lead the way by improving how businesses and consumers use and experience the digital world.” (Schelsinger, 2021)

The data set being used was set up by a Ukrainian data scientist by the name of Anna Pastushko who works for GlobalLogic. It summarizes the top 100 selling electronics items for each calendar day in the months of February and March of 2021. Along with the electronics items and their positions, the data also includes the metrics for the number of total reviews as well as the average review score among them. It also includes the retails price for the item on the day in question. The data set itself is quite robust, consisting of thousands of data points. This will help to provide a very comprehensive view of the electronics sales data within the Amazon platform over the months set forth within the scope of this report.

It is hypothesized that the study will be able to identify popular products through the use of analyses and predictive analytics. This would allow the company as a whole to identify which products will continue to be popular through the course of the next several months. In addition to this, patterns will be able to be tracked among the review scores and total number of reviews as well as price. Occurrences such as sales surges and discounts can play a part in determining the top selling items, and with the data set, these variables can be analyzed in detail. A number of other valuable insights can be obtained from this data, and through tinkering with it within the SAS platform it is hoped that more hidden relationships will be ascertained and exploited in the search for more data points.

To track the results of these concerns, some important questions need to be asked. First, “Are there any products or product types which consistently show up in the top of the charts?” The null hypothesis is that there are no consistent patterns in the top of the charts that are able to be tracked. The alternate hypothesis is that certain products or product types are consistently sold more than others. This is important because it identifies which products are the big money makers within the Amazon electronics platform. Another important question is, “Is there a correlation between review scores and position within the top of the charts?” The null hypothesis is that there is no correlation. The alternate hypothesis is that there is a measurable correlation between these metrics. This is important because it identifies customer opinions as a significant metric in determining sales volume. Another question is, “Does the sales prices affect the position within the charts?” The null hypothesis is that there is not a significant correlation between these two variables. The alternate hypothesis is that there is a measurable correlation that can be observed. This is important because it identifies whether sales prices is an important metric to track when selling certain items on the platform.

A number of sources will be used for the purposes of this study. These sources will serve to reinforce talking points throughout the study. The first type of source will be those which illuminate Amazon’s business model and objectives as a company. These will serve to provide context for the data as well as help illustrate the story behind the data as a whole. To this point in the study, these include an article from Fortune by Murray and Meyer in May of 2020 which give an idea of the total business that Amazon does within the US to give an idea of the retailer as a whole. Another source within this category is an article from Bloomberg by Matt Day which highlights electronics trends within Amazon, specifically a potential plan to open brick and mortar electronics discount stores. The second type of source will be those which help to identify market trends within the electronics market. These will provide context for the top performing products over the course of the data’s scope. To this point in the study, these include an article from analytics firm Deloitte by Managing Director and Chief Futurist Mike Bechtel from earlier this month which is about emerging trends within the electronics industry as a whole. This will help to really get into the meat of what the study is finding in terms of which products are performing well. Another source within this category is an article from Forbes by author Mark Schlesinger from January of this year which also helps to illuminate some of the current and emerging trends within the technology field. All told, the sources for this project will expand greatly and will serve to create a framework of context for the conclusions that will be drawn based on the data set and the analytics performed on it.

The bulk of the analytics for this project will be done using the SAS platform. Using this platform, a number of detailed analytics can be performed on the dataset as a whole which will serve to break the data down into digestible chunks. Predictive analytics will be performed on individual variables, and in addition, multi variable relationships will be detailed and analyzed. A number of actionable relationships between price, reviews, number of reviews, and position within the top 100 will serve to break down market trends as well as trends within individual products. Once these analyses have been performed it will then be time to create visualizations which will be digestible for the analytics layman. These will be created through the use of Tableau, as it provides a plethora of different types of graphs and charts which can help to create a vivid picture of the analytics.   
 Limitations within this study stem from the data’s limited scope. The data is a two month snapshot of the electronics market as a whole within the Amazon platform. Any more and the data would take a team to parse through and produce useful analytics. Any less and the time period could not provide any useful insight. It is speculated that the scope, although limiting, serves the purposes of this study perfectly.

The ethical concerns of this sort of project are extremely limited. The data does not deal with any personally identifiable information, nor does it deal with any particularly sensitive topics. It is therefore postulated that the ethics of this project will not be questionable in any way and therefore do not merit much consideration. This being stated, it is important to note the gravity of concerns within the industry that big data analysts need to keep in mind. Since the adoption of the internet en masse by the public at large, data security and privacy has become more and more of a concern over the years. There are a bevy of challenges that are presented by the ability to share data instantly across large distances and through many different servers and routers. Working so closely with data that is commonly acquired over the net, data analytics professionals share these challenges with the rest of the economy at large. When working with the private and personal information of so many people and transforming it in such a way as to be useful to decision makers, analysts should take great pains to make sure that the integrity of security processes is maintained to the highest standard possible.

Both from a professional and ethical standpoint, making sure data is secure is an absolute imperative. The biggest hurdle for making sure that this is achieved is during the transfer of data. Encryption of email attachments is a great step to make sure that data gets from one point to another without becoming compromised. The second step is making sure that security is maintained while working with the data. This can be done through a combination of firewalls, virus protection, and VPNs. Making sure that cybersecurity departments are regularly working with IT in order to keep every terminal up to date in its security allows analysts to be relatively worry free, but also allows the data to link up to the intranet without concerns. VPNs can also aid in this process as they are an ideal choice during remote work situations. This is especially important during a pandemic lockdown. Finally the terminal itself must be secured. This can be done through a number of methods, most notably tokens and/or strong passwords. In this regard having solid policy in place to assist in this effort is invaluable.

In regards to maintaining data integrity when disbursing it within an organization, there are many steps that can and must be taken to ensure that any sensitive data does not fall into the wrong hands. Within the finance industry, this occurs multiple times a day for each employee within a given organization. The main ways that banks avoid losing customer data is by encryption and layering. Encryption is a vital step because bank branches are spread throughout the market and therefore must sync up to the entire internet in order to access the company intranet. Great care is taken to ensure that each individual terminal is granted access, but bad actors such as hackers and thieves are not allowed to access the financial network. When receiving reports both visual and data oriented in nature these methods are vital in order to allow data to arrive at its proper destination without incident. The second method is layering, which is done to ensure that even though a breach may occur its effects are limited. This is done by setting up different portals for each individual employee. These portals are uniquely encrypted such that even when one may be compromised the rest are not. This ensures that any potential effects of a data security failure are mitigated.

Data security is the responsibility of every element of an organization. One weak link can cause a cascade of reputation loss and financial penalties as a result of a large scale data breach. By avoiding the large scale data security and privacy issues presented by large firms like Experian, Equifax and others, companies can stay ahead of the game and avoid serious complications. In addition, by maintaining high ethical standards, companies can use these policies to inspire loyalty within their market and allow the future of data privacy to be one where the bad elements of society are kept at bay and the free market is allowed to flourish in their absence.

This study should create some interesting conclusions. As Amazon is the biggest retailer within the United States, it stands to reason that any analytics done on sales data within the platform will have wider implications on the market as a whole. It is thought that through the descriptive and predictive analytics that will be performed on this data set, wider conclusions about the electronics market can be made. With these conclusions, any firm that would have access to the predictive analytics gleaned from the study would stand to benefit, as they would be able to alter their marketing and sales strategy in order to compete. For Amazon itself, having access to the results of this study would allow them to do the same and further cement their grip as the top retailer in the country.

References

Bechtel, M (2021, April 5). *Technology Futures Report 2021*. Deloitte.

https://www2.deloitte.com/global/en/pages/about-deloitte/articles/world-economic-forum

-technology-futures-report.html?id=us%3A2ps%3A3gl%3Acon%3Aweftech%3A040621%3Aad1%3Akwd-40898242%3Atechnology+report%3A%7Bcriterionid%7D&gclid=Cj0KCQjwppSEBhCGARIsANIs4p7b7K-HWCxOZ--m9sN59fj\_-v5Vhtz\_B7sVZtU908BDW-iCpYJV4vMaAgz8EALw\_wcB.

Davis, K., & Patterson, D. (2012). *Ethics of big data*. O'Reilly.

Day, M. (2021, April 1). *Bloomberg.com*. Amazon Explored Opening Home Goods, Electronics

Discount Stores. https://www.bloomberg.com/news/articles/2021-04-01/

amazon-explored-opening-home-goods-electronics-discount-store.

Murray, A., & Meyer, D. (2020, May 18). *Guess who's Number Two on this year's Fortune 500*

*list*. Fortune. https://fortune.com/2020/05/18/amazon-number-two-fortune-

500-ceo-daily/.

O'Leary, Z. (2017). *The essential guide to doing your research project*. Sage Publications Ltd.

Pastushko, A. (2021, April 3). *Amazon Top 100 Best Sellers in Electronics 2021*. Kaggle.

https://www.kaggle.com/annpastushko/amazon-top-100-best-sellers-in-electronics-2021.

Schlesinger, M. (2021, January 21). *Council Post: Technology Trends That Will Lead The Way In*

*2021*. Forbes. https://www.forbes.com/sites/forbestechcouncil/2021/01/22/

technology-trends-that-will-lead-the-way-in-2021/?sh=4ce679f73ca3.