**Amazon Books**

Christine Gardner

Colorado State University Global

Course Code: MIS 480

Dr. Orenthio Goodwin

March 10, 2024

**Amazon Books**

There are a lot of e commerce businesses that are successful, and those companies’ collect lots of data daily about customers that help them to make informed business decisions. There are multiple ways that an ecommerce business can benefit from business intelligence tools. These tools will need to be set up to handle big data so that the company can maintain a competitive edge. BI tools help companies transform massive amounts of data into valuable insights that can help increase sales by tailoring recommendations, improving marketing capabilities, and mitigating customer complaints. Amazon is a major e commerce company that uses data to grow its business. Amazon has multiple facets to its ecommerce business, and it currently dominates the ecommerce market. This company has lots of theories that can be solved using a BI tool such as Tableau to gather useful insights on the large amounts of data available to this company.

**Amazon Story**

Amazon started out as a book retailer and evolved into the major ecommerce business that many consumers know today. At first Amazon “was founded on July 5, 1994, as a website that only sold books, founder Jeff Bezos had a vision for the company's explosive growth and ecommerce domination. He knew from the very beginning that he wanted Amazon to be "an everything store" (Hartmans, 2015, para. 1). Amazon has grown and expanded its services to include markets such as online services, and subscriptions. Amazon has “grown over the decades to become the largest e-commerce company by revenue in the United States” with many competitors Amazon needs to have fast insights to stay ahead of the competition (Segal, para. 1). One main part of the business still includes book sales. Gaining insights into the book market can be very useful to the company while helping to maintain its large portion of the retail industry. In order to gain useful insights Amazon can implement Tableau for its business intelligence needs. By using Tableau Amazon can answer impactful business questions about the book industry as well as financials, and company metrics.

**Theories to Solve**

To answer questions using big data Amazon must take the time to plan and decide what questions to answer about the business. These questions should lead to a course of action that helps the company to improve and reach essentail goals. The retail world is extremely fast paced, and BI helps companies like Amazon to make sense of large volumes of data and turn it into actionable insights. Amazon’s business is comprised of selling books and a BI tool could help Amazon analyze books that sold well in the past. Amazon could use this information to compare different attributes of books to see if they can help predict the impact that customer reviews, books pricing, genre, and authors have on sales. Amazon could review book sales on the company’s website and gather data about each book and analyze the score that consumers use to review the books. Amazon could then use that data to see what attributes are most common for the highest rated books in a specific time-period. For example, in Figure 1 the dashboard depicts that books in the genre fiction are the highest rated in the years 2009 – 2013. Amazon can drill down into the data for the books with the rating of five for each year and see if there are any attributes that the books have in common. For example, the majority of the books may have all been published by the same publisher, or have a publish date of the same year that it was rated a five. These insights can help Amazon, because in response to this data the company could acquire more books from the publishers that have the highest rated books, which could increase sales. Amazon could also create a market strategy based around new books that are published in the current year to increase sales. There are many other insights that Amazon can uncover using Tableau to help increase sales or improve the company’s marketing strategy. The other two data sets included data that was combined to create the insights found in Figure 1. This dashboard shows units sold compared to genre and unit price. It also shows gross sales by genre. According to blank “the purpose of the Dashboard is to display information on a single screen in a clear manner, in order to be understood by everyone”, and Tableau makes it extremely easy to combine multiple visualization on screen in an organized visually pleasing manner (Ioana, 2014, pg. 852). Tableau allows for the easy combination of many visuals from 3 different data sets. This dashboard can be used to help increase sales and improve marketing strategies. BI tools don’t just help with marketing, these tools also can help track company metrics. The company can create a dashboard centered around book sales from year to year, shipping lead times for book orders, and cost versus revenue for the book inventory.

**Dashboard Insights**

To display these theories Amazon can create insights similar to the Dashboard in Figure 1 by collecting and using a data set that records book price, review score, title, and units sold. This data set can be created multiple ways, Amazon could pull this information from an internal information management system or create an API to scrap the web for information on reviews about specific book titles that Amazon has the sales data for. The data used for the dashboard in **Figure 1** includes three different sets with thousands of entries. Even though all these data sets have different origins the company is still able to gain insights that help to improve the company. Without a BI tool like Tableau Amazon would have to dedicate lots of time and team members to gain information that can lead to helpful/ profitable conclusion for the company.

There are a lot of BI tools in the world today that are built to help companies gain insights. Tableau is a user-friendly tool, with an easy-to-understand intuitive design. Tableau is the Business intelligence tool that would make it easy for Amazon to combine large amounts of data from data sets to gain insights. Tableau is a user-friendly visualization tool that helps users to uncover insights in a drag and drop enviroment. Amazon can use Tableau data prep clean and import data into Tableau desktop to get easy to understand visualizations. These visualizations can be combined into dashboards that summarize the data sets.

**Other Insights Outside of Book Data**

Tableau not only provides a platform for data visualization, but Tableau Prep takes data sources and helps prepare them to be analyzed. All these resources are available within the same major product Tableau unlike other BI solutions that require companies to use 3rd party data prep tools, or even excel. Choosing Tableau ensures that the staff won’t have a steep learning curve when trying to use and implement Tableau. Tableau has lots of tutorials available and lots of online resources to help users understand how to get the most out of this business intelligence platform. Using Tableau and Tableau Prep allows for a seamless transition for data because of the similar interfaces and internal connectors that make it easy for information to travel from application to the other. Tableau prep has special features to make visualizing data flows easier as well as repeatable independent of the data set. Figure 1 depicts a data flow that uses essentail steps such as clean and join to prepare a large data set for visualization. Tableau Prep also has the ability to work on a small sample of the data while a user creates the flow. This ensures that large volume data sets don’t slow down the application or crash the system. Trying to open the same amount of data in excel is most likely to crash the application, or make the application run at a very slow speed, which creates a barrier between Amazon and gaining insights from the data. All of these features make Tableau and Tableau prep a great candidate for a BI tool solution for Amazon. Since an “investment in big data analytics has almost become a necessity in large-sized firms, particularly multinational companies, for its unique benefits, particularly in prediction and identification of various trends” it makes a lot of sense for Amazon to adopt Tableau as the BI tool for the company (Biju, 2017, pg. 2). Having this BI tool in place will help ensure the company’s success in a very competitive market.

**Data Set 1 – Book Sales Ratings**

This data set was downloaded from Kaggle and it includes a lot of useful fields for analysis. This data set includes book name, book average rating, genre, gross sales, sales price, and units sold. There are other fields in this data that can be used for futher insights such as author rating, and author. For the visualization pictured in Figure 2 the first listed set of fields is what was need to create the insights displayed. This data set contributes to the bar chat that compares genre type to the price of the book sold with the total units sold. This graph is separated by minimum and maximum price. Amazon users can use these visualizations to try and predict how to price books in a specific genre. For example, if Amazon wanted to choose between two fiction books to build a marketing campagin around and they only had the operating income to fund marketing for one they can use this visualization to help them choose. If one book is priced at $45 and the other is priced at $32, the $32 book falls into the range of books that sold the most units between the years 2009 to 2013. Amazon could select this book because the graphic also shows that as the highest price for books with a large number of units sold. This data could be updated to include different years, which would just help Amazon base decisions off of something more than just guessing which to choose.

**Data Set 2 and 3 – Book Ratings and Book Details**

The last two data sets are also from Kaggle and the data sets work together to provide some useful insights that Amazon can Use. Each data set includes book titles, but they separate because one data set has fields that relate to the descriptors for the book while the other data set is focused around review results. The book review data set includes fields for review/ score for the book, review/time, book price, and book id. The book detail data set includes fields for publisher, genre, and published date. Each data set has more fields listed that can be used in other visualizations, but the ones listed were used to create the visualizations in Figure 3. This visualization shows the top total review score for a book by year. This visualization only presents this data for the years 2009 – 2013. Amazon can use this type of visualization to try and understand what books have the highest combined review scores. The visualization depicts pride and prejudice with the highest score for 2012, 2011, 2010, and 2009. Amazon can use the attributes for the top highest rated books and analyze their attributes and compare them to sales for that specific year to see if the rating score influences sales. If Pride and Prejudice has lots of sales for 2012 and 2011 Amazon users can try to understand what makes that book so highly rated and stock similar books that will increase review scores. This data set can be combined to answer future questions and build on this data set with a comparison with top rated books and best-selling books to see if review scores have any correlation to units sold of specific books. Tableau gives Amazon the foundation to combine various data sets that can impact business decisions to increase sales and positively impact marketing.

Please see the figure below for the Tableau Book Insights Dashboard.

**Figure 1**

*Tableau Book Insights Dashboard*A screenshot of a computer

Description automatically generated

*Note*. This Tableau dashboard is an example of how Amazon can present multiple different visualizations that can help the company answer questions and research theories.

Please look at the Figure 2 below for the Gross Sales by Genre.

**Figure 2**

*Gross Sales by Genre Visualization* A screenshot of a computer

Description automatically generated

*Note*. This visualization makes it easier for the user to conclude that Fiction is the highest selling genre with the included data. This visualization lets the user have a perspective of the scale of 1,684,357 units looks like compared to the second largest genre sales at 220,958 units.

Please look at the Figure 3 below.

**Figure 3**

*Book Detail and Book Ratings in Tableau Prep*A screenshot of a computer

Description automatically generated

*Note*. Ada Tableau Prep makes it easy to combine data sets and create flows. The layout is very user friendly, and the application allows you to drag and drop different aspects to edit and clean the data.

**Benefits of Using Tableau on Answering Theories**

When considering the benefits of BI user may find themselves asking “what if data analysis answered the original question but generated many more? Rapid analytic and data visualization software offers” many solutions but it also needs to be able to answer questions the constant stream of questions that come with handeling big data (Murphy, 2013, pg. 465). Some BI tools are so complex and require a large learning curve that some of the value of the system is lots in the amount of time that it takes to answer questions and gain insights. Since Tableau is so easy to use Amazon can have employees spend more time focusing on what questions to ask that they can quickly and efficiently answer using Tableau. Some questions that may be beneficail for a company like Amazon to find answers to and gain insights from in the future are as follows. Does the sentiment recorded in the book reviews support the best sellers for that year? Can Amazon use customer reviews to build profiles that build a base for customer segmentation based on reviews and recommend books from certain profile to other users that have one book review in common? These are just a few of the questions that Amazon can build a theory to try and answer about book reviews and top selling books. Amazon would just need to gather the top selling book list and set the data set up to present insights inside Tableau. This list could also be broken down by which genres sell the most books during certain timeframes of the year. Do romance novel sales increase from January to February. What price range do the bestselling books fall into? Or does the bestselling book price vary greatly? Amazon has all the data to answer these questions, but they need to ensure that there is a tool to make those insights easy to unpack and visualize.

**Conclusion**

Amazon has lots of data that they can collect but I choose data sets on books because this is the origin of the company. I chose Amazon because I thought they would have a lot of data sets available to analyze which would lead to a more interesting comparison and analysis. Plus since Amazon is a large, succesful company it was a good candidate to explore ways to improve using big data and business intelligence. I chose Tableau because of its user-friendly structure and easy to build visualizations, which helps users focus more on insights instead of struggle to process the raw data. Tableau is a great BI tool for Amazon to use this application can help answer questions that promote sales, record metrics, and improve marketing strategies. Overall Tableau has a lot to offer Amazon and if the questions and theories are well aligned with the company’s goals it can help push Amazon further into the lead as a retailer.

**References**

Hartmans, A. (n.d.). Jeff Bezos originally wanted to name Amazon “Cadabra,” and 14 other little-known facts about the early days of the e-commerce giant. Business Insider. https://www.businessinsider.com/jeff-bezos-amazon-history-facts-2017-4#amazon-wasnt-the-companys-original-name-1 Devereaux, A. (2015). Pandemic influenza: An evolutionary concept analysis. *Journal of Advanced Nursing, 71*(8), 1787–1796. <https://doi.org/10.1111/jan.12654>

Biju, S. M., & Mathew, A. (2017). Comparative Analysis of Selected Big Data Analytics Tools. Journal of International Technology & Information Management, 26(2), 2–22. https://doi.org/10.58729/1941-6679.1301

Ioana, B., Diana Claudia, S.-P., & Ioan, B. M. (2014). Using Dashboards in Business Analysis. Annals of the University of Oradea, Economic Science Series, 23(1), 851–856.

Murphy, S. (2013). Data Visualization and Rapid Analytics: Applying Tableau Desktop to Support Library Decision-Making. Journal of Web Librarianship, 7(4), 465–476. https://doi.org/10.1080/19322909.2013.825148

Segal, T. (n.d.). Who are Amazon’s (AMZN) main competitors? Investopedia. https://www.investopedia.com/ask/answers/120314/who-are-amazons-amzn-main-competitors.asp