**MGS 628 LEC – Data Visualization for Business Insights Final Project**

***Team Details:***

**Members:**

* Lakshya Rawal, **UB Person Number:** 50459636, **UB Email:** [lakshyar@buffalo.edu](mailto:lakshyar@buffalo.edu)
* Dheeraj Kolla, **UB Person Number:** 50462365 , **UB Email:** [dkolla@buffalo.edu](mailto:dkolla@buffalo.edu)
* Kaushik Kumar, **UB Person Number**: 50471138, **UB Email**: [kthogulu@buffalo.edu](mailto:kthogulu@buffalo.edu)

***Contributions:***

| **Team Member** | **Contribution** |
| --- | --- |
| Lakshya Rawal | Discount and Pareto Principle Storyboard |
| Dheeraj Kolla | Analysis Genre,Author,Publishing House |
| Kaushik Kumar | Analysis of Books, Author Performance, and Key Performance Indicators |

***Overview of Bookshop Activity Analysis:***

This document reflects on building a storyboard analysis on the Bookshop activity data for the identified stakeholders and building an insightful dashboard to analyze the patterns in the bookshop activity based on genre, authors, discount categories.

* **Stakeholder**: CEO of Publishing Company
* **Significance**: This dashboard aims to study the below aspects,
  + Book Sales Analysis, and Performance of Authors
  + Key Performance Indicators on the Authors, Award winning books, etc
  + Relation of discount categories and the orders/sales in those categories. Also it studies a group of books that account for the most sales as a separate group to understand the pareto principle.
* **Key Insights & Decision**: Below are some of the key business questions that the dashboard answers to.

1. What is the most popular book that is sold ?
2. What is the best rated book for new users to get started ?
3. Who is the best author in every genre ?
4. Who was the youngest debut author and who is the oldest debut author ?
5. List of awards won by authors and its corresponding books ?
6. Is there a seasonal trend between the sales across genres ?
7. Is there any relation between hours spent on writing (vs) their success rate ?
8. Is there any relation between discounts

**Discount and Pareto Principle Storyboard**

| **Entity Name** | **Entity Type** | **What does the visualization do?** | **What decision does it allow the stakeholder to make?** |
| --- | --- | --- | --- |
| Sales in Each Discount Category | Visualization Sheet | Plots the Sales in each discount Category  Plot used: Packed Bubbles | Is there a relationship between discount and sales of the books ?  Which discount group has the highest sales ? |
| Sales by Discount Category over months | Visualization Sheet | Plots the sales of books over months of the year by discount category  Plot used: Area Charts | Which discount category has the lowest number of sales each month?  Is there a particular month in which we see sales peak and if so does discount impact that month? |
| Order Count by Discount Category | Visualization Sheet | Plots the count of orders based on the discount category  Plot used: Tree Map | Which discount category has the highest amount of orders?  Is there a direct relation between sales and order count for discount categories? |
| **Discount** | **Dashboard** | **The relation between sales and order count to discount category** | **It helps the stakeholders to visualize the relation between discounts and sales/orders and helps them understand when it might be helpful to increase the discount on the book.** |
| Pareto Principle - Sales | Visualization Sheet | Visualizes the sales of top 20% of books by sales  Plot Used: Pie Chart | If roughly 80% of consequences come from 20% of the books (the "vital few") |
| Pareto Principle - Order Count | Visualization Sheet | Visualizes the order counts of top 20% of books by sales  Plot Used: Stacked Bar Chart | If all the orders that are recorded are attributed to the few books that are in the top 20%. Making a skewed market share.  If there is a relation between sales and order count of these two groups |
| Pareto Principle - By Genre | Visualization Sheet | Visualizes the sales of the above groups across genres  Plot Used: Horizontal Bars | Is there a genre that is more susceptible to this skewness  If so, which genre is this and which books are those that dominate this genre. |
| **Pareto Principle** | **Dashboard** | **The market share of books and if the market is skewed by few books** | **Understand the distribution of sales and order counts by the top 20% of books by sales compared to the rest of the books.** |
| Checkouts per Month | Visualization Sheet | Visualizes the variation in number of checkouts with each month  Plot Used: Box Plot | Which month has the most checkouts?  Which month has the lowest checkouts? |
| Publishing house checkouts | Visualization Sheet | Visualizes the checkouts for each publishing house and filters how each one is performing  Plot Used: Lines Chart | Which publishing house is performing the best?  How is my publishing house fairing when compared to other publishing houses? |
| Top N Most Checkout Book | Visualization Sheet | Visualizes the top books that were most checkouts by publisher  Plot Used: Highlight Tables | Which of my books are account for most checkouts  If so, should I focus more on books that are underperforming or release new editions of these. |
| **Checkouts** | **Dashboard** | **The number of checkouts broken down by month, publisher and popular books** | **Understand the distribution of checkouts and understand what is the most popular books, publisher, month that is currently there.** |

**Analysis of Books, Author Performance, and Key Performance Indicators**

| **Entity Name** | **Entity Type** | **What does the visualization do?** | **What decision does it allow the stakeholder to make?** |
| --- | --- | --- | --- |
| Best Selling Books | Visualization Sheet | Plots the best selling books by total sales.  Has a Top N parameter and book lookup functionality  Plots Used: TreeMap | It allows the stakeholder to analyze the best and worst selling books and take appropriate actions  What are the best selling books ? |
| Best Rated Books | Visualization Sheet | Plots the books with the most number of 5 star ratings  Plots Used: Vertical Bar Chart | Helps the stakeholder in analyzing the best rated books by users  What are some of the best rated books ? |
| Book counts and Author choice per Genre | Visualization Sheet | Plots the total books sold at a genre level along with the author with best ratings. The data is filtered with respect to best author  Plots Used: Vertical Bar Chart | Allows the stakeholder to identify the best rated author by each genre, and also look at the overall book sales by genre  Who is the preferred author per genre ? |
| Overview on book sales and ratings | Dashboard | Presents the overall summary of book sales and best books to read for new users | Provides the stakeholder with a high level summary of the best books sold so far |
| Author Performance | Visualization Sheet | Plots the relationship between hours spent on writing vs total sales  Plots Used: Bar Chart | Helps the stakeholder in analyzing the writing style of every author  What is the author's writing style and should it be improved ? |
| Author pricing level | Visualization Sheet | Plots the average price per author  Plots Used:Bar Chart | Allows the stakeholder to identify authors who charge high and low for their books  Which author charges more for their books ? |
| Author Debuts | Visualization Sheet | Plots the time taken for every author to publish their first book  Plots Used: Bar Chart | Helps the stakeholder in identifying the youngest debut and oldest debut author  Who is the youngest author that published a book ? |
| Overview on Author Performance and profile | Dashboard | Visualizes the overall author performance and their milestone achievements | Provides the stakeholder with the author performance summary and their timeline of publications |
| Award winning books and authors | Visualization Sheet | Plots the authors and books that won an award along with its year of achievement  Plots Used: Bar chart | Helps the stakeholder in recognizing the authors for their achievements  What are the books and who are the authors that won an award ? |
| Seasonal trend of sales | Visualization Sheet | Plots the trend of sales with respect of sales month  Plots Used: Area Charts | Aids the stakeholder in identifying the months with highest sales across every genre  When do people love reading ? Is there any specific month ? |
| Key Performance Indicators of Books and Authors | Dashboard | Provides a overview of the key performance indicators about award winnings and trends | Provides the stakeholder an opportunity to visualize key establishments and trends |
| **Books & Author Analysis** | **Storyboard** | **The storyboard presents a birds eye view about the books, authors, and their performance and achievements** | **Enables the stakeholder to analyze where they stand and to identify areas of improvements to increase the revenue** |

**Overview of Author,Genre,Publishing house Analysis Storyboard:**

**Stakeholder**: Publisher

**Key Insights & Decision**: Below are some of the key business questions that the dashboard answers to.

1. Which genres are the most popular among readers?
2. What is the market demand for specific genres?
3. How do sales and revenue vary across different genres?
4. Which genres have the highest average ratings or reader engagement?
5. What is the financial performance of the publishing house?
6. How has the revenue of the publishing house changed over time?
7. What are the main sources of revenue for the publishing house?
8. Which authors have the highest book sales and revenue?
9. How do different authors perform in terms of average ratings and reader engagement

| **Entity Name** | **Entity Type** | **What does the visualization do?** | **What decision does it allow the stakeholder to make?** |
| --- | --- | --- | --- |
| Number of Checkouts by Genre | Visualization Sheet | Plots number of checkouts out books by Genre  Plot used: Bar Chart | The graph can help identify the most popular genres among readers based on the number of checkouts. |
| Ratings by Genre | Visualization Sheet | Plots the average rating of books in different genres  Plot used: Pie Chart | The ratings by genre graph allows the stakeholder to assess the performance of books within each genre based on reader ratings. By comparing the average ratings across genres, the stakeholder can identify genres that generally receive higher ratings and are well-received by readers |
| Monthly Sales by Genre | Visualization Sheet | Plots the sales of different genre books over different months  Plot used: Area chart | By observing the monthly sales patterns for each genre, stakeholders can identify any seasonal trends. For example, certain genres may experience higher sales during specific months or seasons, allowing the stakeholder to plan marketing campaigns or promotions accordingly. |
| Genre Analysis | Dashboard | Sales and rating analysis of genre | Stakeholders can identify the genres that are currently popular and in high demand, helping them focus on acquiring books within those genres to maximize sales potential.  Stakeholders can identify the genres that have the highest sales and revenue, allowing them to tailor marketing campaigns and promotions to target those specific genres  They can ensure they have sufficient stock of books within popular genres to meet customer demand while reducing inventory for genres that have lower sales. |
| Top Authors by sales | Visualization Sheet | Plots top N number of authors based on count of number of books sold  Plot used: Bubble Chart with interactive parameter | The graph can help identify the most popular authors among readers based on the number of sales |
| Sales of authors over time | Visualization Sheet | Plots the sales of different authors over different months  Plot used: Line chart | By observing the sales patterns of an author over months, the stakeholder can make decisions regarding resource allocation. If an author consistently performs well during certain months, the stakeholder may allocate additional marketing resources or promotional efforts during those periods to maximize sales. |
| Ratings of Author | Visualization Sheet | Plots the average  ratings of authors  Plot used: Box-Whisker Plot | Ratings can help stakeholders evaluate and compare authors based on their performance and reception among readers. High ratings may indicate the author's ability to engage and satisfy readers, making them an attractive choice for publication or collaboration. |
| Author  Analysis | Dashboard | Sales and rating analysis of Author | Author analysis helps stakeholders evaluate an author's performance, such as book sales, average ratings, and reader engagement. Based on this analysis, stakeholders can make decisions regarding publishing contracts, including offering new contracts to successful authors, renegotiating existing contracts, or terminating contracts with underperforming authors.  By analyzing an author's performance, stakeholders can make informed decisions about the timing and frequency of book releases. |
| Average Price of book by each publishing house | Visualization Sheet | Plots the average price of each book by the publisher  Plot used: side-by-side bar chart | The stakeholder can assess whether the average price of books aligns with the publishing house's pricing strategy. If the average price is higher or lower than desired, adjustments can be made to ensure competitiveness in the market and maximize revenue. |
| Marketing Spend vs Total Number of purchases | Visualization Sheet | Plots the marketing spend and number of sales by publishing house  Plot used: TreeMap | By comparing the Marketing Spend with the Total Number of Purchases, the stakeholder can calculate the Return on Investment (ROI) for their marketing efforts. They can analyze whether the marketing spend is resulting in a sufficient increase in the number of purchases. If the ROI is positive, it indicates that the marketing strategies are effective in driving sales. |
| sales by each month by publishing house | Visualization Sheet | Plots the monthly sales of publishing house  Plot used: Line chart | By understanding the sales patterns and trends for each month, the stakeholder can adjust their inventory levels accordingly. They can ensure they have sufficient stock available during high-demand months and avoid overstocking during slower months, optimizing their inventory management and reducing storage costs.Increase Marketing strategies and other strategies to increase sales in low-performing months |
| Number of books by format and publishing house | Visualization Sheet | Plots the number of books published by each publisher by format.  Plot Used:side-by-side bar chart | By understanding the distribution of books across different formats (e.g., Paperback, Hardcover, BoardBook), stakeholders can allocate resources more effectively. They can identify the formats that are most popular or have the highest demand and allocate resources such as production, marketing, and distribution accordingly. |
| Publisher Analysis | Dashboard | Analysis of Publishing house | The stakeholder can assess whether the average price of books aligns with the publishing house's pricing strategy. If the average price is higher or lower than desired, adjustments can be made to ensure competitiveness in the market and maximize revenue.  By comparing the Marketing Spend with the Total Number of Purchases, the stakeholder can calculate the Return on Investment (ROI) for their marketing efforts. They can analyze whether the marketing spend is resulting in a sufficient increase in the number of purchases. If the ROI is positive, it indicates that the marketing strategies are effective in driving sales. |
| Overview of Author,Genre,Publishing House | StoryBoard | The storyboard presents the bird eye view of author,genre,publishing house analysis | By utilizing the storyboard, the publisher stakeholder can gain actionable insights, monitor key performance metrics, and make informed decisions about book acquisitions, marketing strategies, author relationships, and resource allocation. This enables them to optimize their operations, drive growth, and enhance reader satisfaction in the competitive publishing industry. |