

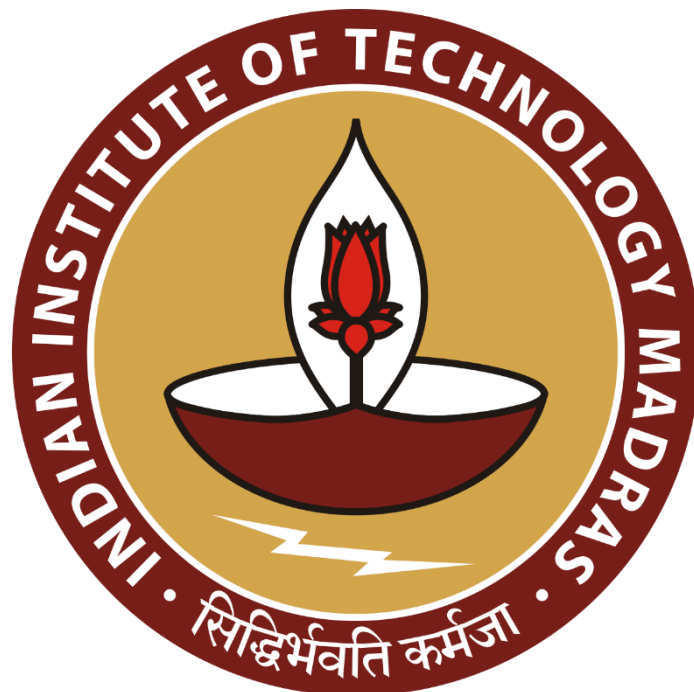
Digital Disruption: The Decline of Book Purchases in the Digital Era

A Mid-Term report for the BDM capstone Project

Submitted by

Name: Sudiksha Singh

Roll number: 22F3001123



IITM Online BS Degree Program,
Indian Institute of Technology, Madras, Chennai
Tamil Nadu, India, 600036

Contents

0	Declaration Statement	2
1	Executive Summary and Title	3
2	Proof of Originality of Data	4
3	Meta Data	5
4	Descriptive Statistics	7
5	Detailed Explanation of Analytical Process	8
6	Results and Findings	9

Declaration Statement

I am working on a Project titled “**Digital Disruption: The Decline of Book Purchases in the Digital Era**”. I extend my appreciation to **NIMBUS BOOK RETAIL OUTLET**, for providing the necessary resources that enabled me to conduct my project.

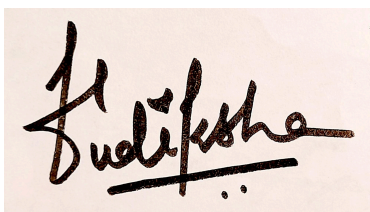
I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered from primary sources and carefully analyzed to assure its reliability.

Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the principles of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I understand that all recommendations made in this project report are within the context of the academic project taken up towards course fulfillment in the BS Degree Program offered by IIT Madras. The institution does not endorse any of the claims or comments.

A handwritten signature in black ink on a light pink background. The signature is stylized and appears to read 'Sudiksha Singh'.

Signature of Candidate

Name: Sudiksha Singh

Date: 01-08-2024

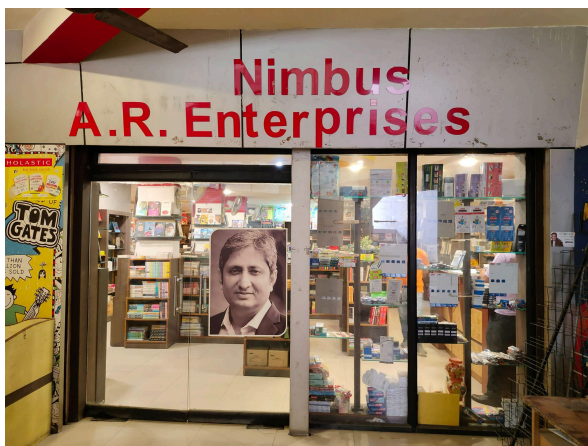
01. Executive Summary

Nimbus Book Retail Outlet located at F 25, CMC, R G College Road, Meerut, Uttar Pradesh, Nimbus Book Retail Outlet is a B2C bookstore founded in 2003 by Mr. Ram Kumar Gupta. Over the years, it has grown to become the largest bookstore in Meerut. The store offers a broad selection of books, including school textbooks, competitive exam preparation materials, fiction, non-fiction, and religious texts, catering to a diverse clientele that includes students, professionals, and avid readers.

However, Nimbus Bookstore faces several challenges. The yearly turnover, which once exceeded ₹1 crore, has fallen to approximately ₹70 lakhs due to the pandemic and lockdowns. Additionally, the rise of digital media and online book purchases has further reduced physical book sales, particularly among younger readers. The store is also contending with old inventory that is difficult to clear, impacting profitability.

This BDM capstone project aims to address these challenges. Data analysis will identify key patterns in sales trends, inventory management, and customer behavior. Tools such as Google Sheets, Google Collab and other analytical software will be utilized to process and interpret this data. The insights gained will guide future decision-making and strategic planning, with the goal of enhancing profitability and ensuring the long-term sustainability of Nimbus Book Retail Outlet.

02. Proof of Originality of Data



NIMBUS BOOKS RETAIL OUTLET

Date- 24/06/2024

Subject:- Proof of originality of Data for Project Submission

I, **Ram Kumar Sharma**, confirm that the data provided to Sudiksha Singh for her project submission of the Business data Management Capstone project is original and authentic . The data has been collected from our business, Nimbus Books Retail Outlet, F-25 to 29 Chetan Medical Complex, R.G College Road ,Meerut .

I have personally overseen the data collection process and can vouch for its accuracy and confidentiality. The data reflects the actual state of our business operations.

Thank You

Regards,



(Ram Kumar Sharma)

Nimbus Books Retail Outlet,
F-25 to 29 Chetan Medical Complex,
R.G College Road ,
Meerut,
Uttar Pradesh

[**Click here for Interview and Photos of Book Store**](#)

03. METADATA

The collected data for Nimbus Book Retail Outlet covers the period from April 1, 2018, to March 31, 2024. It includes a raw dataset consisting of both Sales Data and Inventory Analysis.

The raw dataset is organized into 12 sheets, with six dedicated to purchase data and six to sales data, spanning six years from 2018 to 2023.

Metadata provides descriptive details about the dataset, offering context and clarity on the data fields.

Data Sheets Columns:

- **Particular:** Indicates the name of the publisher or bookstore from which books are purchased.
- **Inwards Qty:** The quantity of stock items received from each publisher.
- **Rate:** The per-unit cost price of the incoming stock from each publisher.
- **Inwards Amount:** The total cost value of the stock received from each publisher.

- **Outwards Qty:** The quantity of stock items sold or transferred out.
- **Selling Rate:** The per-unit selling price of the books.
- **Outwards Amount:** The total sales amount for each publisher's books.

Purchase data-

Particulars	Inwards Qty	Rate	Amount
Cash	16 pcs	313.80	5020.74
<i>Analysis of Financial Statements</i>	2 pcs	449.63	899.25
<i>Business Studies</i>	2 pcs	604.80	1209.60
<i>Chemistry XII Part-II</i>	5 pcs	119.00	595.00

Sales data-

Particulars	Outwards Qty	Selling Rate	Amount
Asha Book Depot	6 pcs	277.23	1663.35
<i>101 Bedtime Stories</i>	1 pcs	200.00	200.00
<i>Political Science(Mishra)</i>	1 pcs	688.00	688.00
<i>Railway Samanya Vigyan</i>	1 pcs	201.60	201.60

[Click here for Raw Data of NIMBUS BOOK STORE](#)

Data cleaning

To prepare the data for analysis in the Nimbus Book Retail Outlet project, all raw data is first gathered and imported into Google Sheets. The data is then carefully examined for any missing values, duplicates, or outliers. Missing values are either imputed based on logical estimates or removed if deemed non-critical. Duplicate records are identified and deleted to prevent skewed results. Date formats, product names, and categories are standardized to ensure consistency. Anomalies are investigated, and corrections or removals are made as necessary. Data across different datasets, such as sales and inventory, is validated to ensure alignment. The data is then normalized and aggregated as required for analysis. Totals are cross-checked, and an initial analysis is performed to confirm data quality. Finally, the entire cleaning process is documented, with any significant changes noted.

The cleaned data includes Sales Data and Inventory Analysis for **Nimbus Book Retail Outlet**.

Inwards and Outwards Stock Data-

This dataset provides a detailed account of stock movements, categorized by supplier, at Nimbus Book Retail Outlet. It includes data on the quantities and financials associated with both incoming and outgoing stock. There are 5 Data Columns:

1. **Particulars:** Lists the publishers from whom the stock has been purchased.
2. **Inwards Qty:** The quantity of stock items received from each publisher.
3. **Rate:** The per-unit cost price of the inwards stock for each publisher.

4. **Inwards Amount:** The total cost value of the stock received from each publisher.
5. **Outwards Qty:** The quantity of stock items sold or transferred out.
6. **Outwards Amount:** The total sales amount for each publisher.

Particulars	Inwards Qty	Rate	Amount	Outwards Qty	Amount
Asha Book Depot	8 pcs	40.00	320.00	7 pcs	310.00
Bharatiya Jnanpith	40 pcs	60.00	2400.00	15 pcs	1289.00
Cash	283 pcs	150.00	42450.00	180 pcs	38595.00
Children Choice Publication	4055 pcs	500.00	2027500.00	----	0.00
Chitra Exports	920 pcs	117.00	107640.00	1 pcs	135.00

Annual Financial Data for Nimbus Book Retail Outlet (2018-2023)

This dataset gives a year-by-year look at the financial performance of Nimbus Book Retail Outlet from 2018 to 2023. It shows how much was spent on purchasing stock, how much money was made from sales, and whether the store made a profit or suffered a loss each year. The dataset includes 5 columns:

1. **Year:** The fiscal year for the data, showing the time period for each set of figures.
2. **Duration:** The specific dates for the fiscal year, from April 1st to March 31st of the next year.
3. **Purchase Amount:** The total money spent on buying books and other stock during the year.
4. **Sales Amount (Revenue):** The total income earned from selling those books throughout the year.
5. **Profit/Loss Amount:** The difference between what was earned and what was spent. If this number is positive, the store made a profit; if it's negative, the store experienced a loss.

Year	Duration	Purchase Amount	Sales Amount	Profit/Loss Amount
2018	01/04/2018 to 31/03/2019	8566900	10748610	2181710
2019	01/04/2019 to 31/03/2020	9232385	13675753	4443368
2020	01/04/2020 to 31/03/2021	6546721	1489205	-5057516
2021	01/04/2021 to 31/03/2022	1277508	1318318	40810
2022	01/04/2022 to 31/03/2023	2472210	3992391	1520181
2023	01/04/2023 to 31/03/2024	4017675	7820800	3803125

[**Click here for Cleaned Data of Nimbus Bookstore**](#)

04. Descriptive Statistics

The cleaned data was imported into Google Colab and transformed into a DataFrame using pandas 'read_excel()' function. Subsequently, the 'df.describe()' function was applied to the DataFrame to generate summary statistics, providing an overview of the key metrics such as mean, standard deviation, minimum, and maximum values for the sales data.

Descriptive Statistics of sales of books of each year

```
[3] import numpy as np
import pandas as pd
df = pd.read_excel('/content/Data (2).xlsx', sheet_name= 'sales')
df.describe()
```

index		Unnamed: 0	2018	2019	2020	2021	2022	2023
count		0.0	21.0	42.0	38.0	48.0	44.0	49.0
mean		NaN	511838.5714285714	325613.1666666667	39189.60526315789	27464.958333333332	90736.15909090909	159608.16326530612
std		NaN	562743.7254762216	384929.2571232529	94040.39411281743	99691.85881909684	181311.41897465257	199724.01719084207
min		NaN	21800.0	1290.0	0.0	80.0	468.0	9950.0
25%		NaN	185420.0	76448.75	607.5	1107.5	14296.25	53600.0
50%		NaN	255000.0	222384.0	2550.0	4840.0	30585.0	92400.0
75%		NaN	755800.0	483875.0	18662.5	16225.0	85402.5	172750.0
max		NaN	1950000.0	2218900.0	382500.0	680500.0	1085000.0	1254110.0

Show 25 per page

Descriptive Statistics of purchases of books of each year

```
import numpy as np
import pandas as pd
df = pd.read_excel('/content/Data (1).xlsx', sheet_name= 'purchase')
df.describe()
```

index		Unnamed: 0	2018	2019	2020	2021	2022	2023
count		0.0	21.0	42.0	38.0	48.0	44.0	49.0
mean		NaN	407947.61904761905	219818.69047619047	172282.13157894736	26614.75	56186.59090909091	81993.36734693877
std		NaN	418146.1402656344	193040.8569197632	451337.6303625774	86852.26597406533	66322.8324690108	57695.64246785811
min		NaN	17100.0	1200.0	220.0	175.0	1080.0	6760.0
25%		NaN	135000.0	51250.0	2193.75	2385.0	16407.5	41600.0
50%		NaN	200000.0	177410.0	7140.0	5390.0	33200.0	63000.0
75%		NaN	600000.0	312375.0	42400.0	12762.5	73800.0	110000.0
max		NaN	1400000.0	688750.0	2027500.0	582750.0	297500.0	234000.0

Show 25 per page

[Click here for Test Run results of Google Colab](#)

Upon analyzing the statistics, it was observed that:

2018-2019: The data was quite high and varied a lot, showing strong performance but with big swings.

2020: There was a sharp drop in values, likely due to external factors like the pandemic, leading to more consistent but lower numbers.

2021-2022: Performance started to recover, with values rising again, but there were still some fluctuations.

2023: Performance is strong and improving, but there's still significant variability.

Percentage of book sales distribution by genre

GENRE	Book sales distribution by genre
Fiction	18.30%
Non-fiction	15.50%
Science Fiction	10.20%
Mystery	7%
Fantasy	9.10%
Romance	5.60%
Biography	4%
Children's Books	4.20%
Competitive Exam Books	14.70%
School Books	9.45%
Others	2%

The table shows that **Fiction** and **Competitive Exam Books** lead in sales, while **Non-fiction**, **Science Fiction**, and **School Books** also perform well. **Fantasy**, **Mystery**, and **Children's Books** have smaller but notable shares. **Romance** and **Biography** have even smaller shares, suggesting potential for growth. The **Others** category has the smallest share, indicating niche genres with limited demand. This highlights key areas for focus and potential opportunities for expansion.

By analyzing the data, it was noted that the top-selling book is by **Premchand**, highlighting its strong popularity. In terms of pricing, the most expensive books are usually high-end editions such as rare collectibles, limited editions, or comprehensive academic texts.

05. Detailed Explanation of Analytical Process

1. Sales and Inventory Overview: Start by cleaning up and organizing the historical sales and inventory data in Google Sheets. Then, import this cleaned data into Google Colab. Use Pandas to generate summary statistics like average sales, highs and lows, and variations. This will give us a clear snapshot of how the bookstore has been performing and how its inventory is turning over.

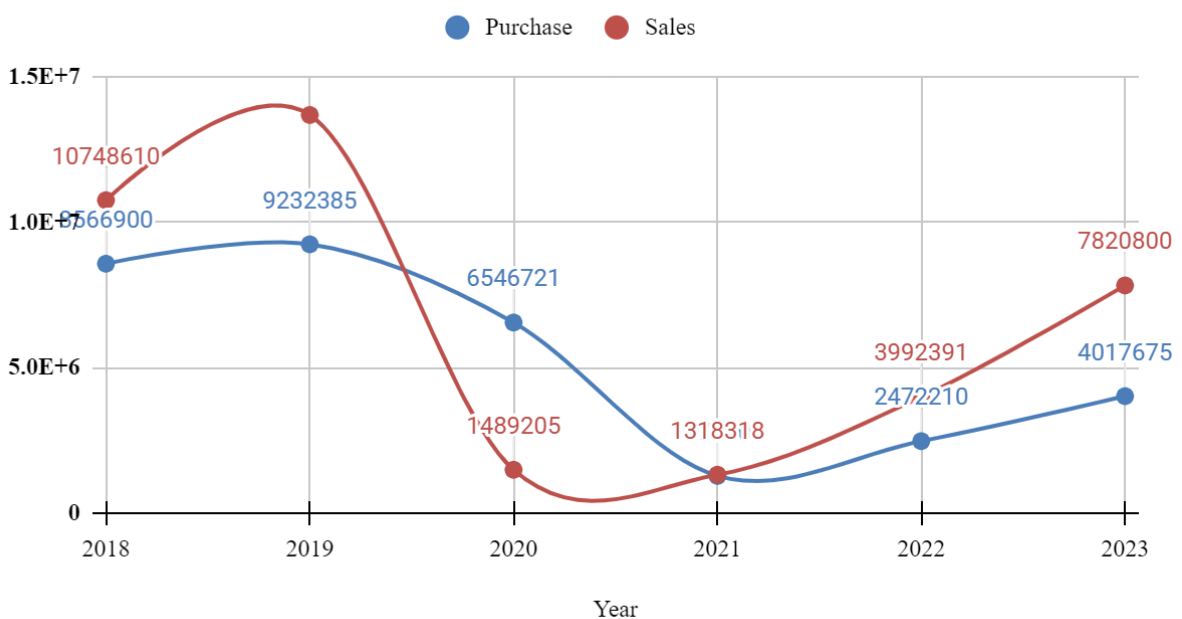
2. Trend Analysis: Look at the sales data over the years to spot trends and patterns. Create a Line Chart to track how sales have changed over time, including any seasonal ups and downs or the impact of events like the pandemic. This helps us see how the bookstore has evolved and how it's recovering or adapting to changes.

3. Comparing Sales and Inventory: Compare sales numbers with inventory levels to spot any mismatches. Use a Column Chart to show how sales stack up against inventory, helping to identify if there are too many books sitting around or if sales are lagging. We'll also compare our performance to industry standards to see how we stack up against competitors.

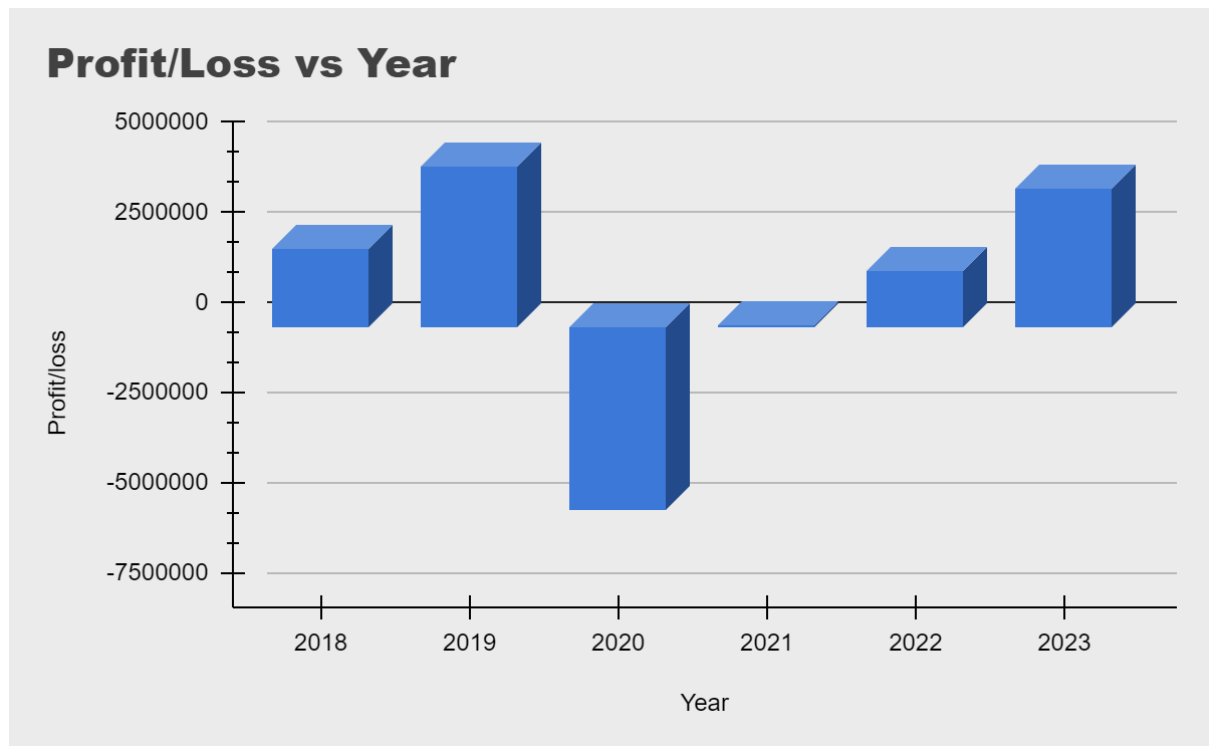
4. Performance Metrics: Summarize important figures like total sales and average sales per month, and look at how fast inventory is turning over. Use visual tools like Bar Charts or Pie Charts to make these numbers easy to understand. This helps us see which books are selling well and gives a clear picture of how the business is doing overall.

06. Results and Findings

Sales vs Purchase

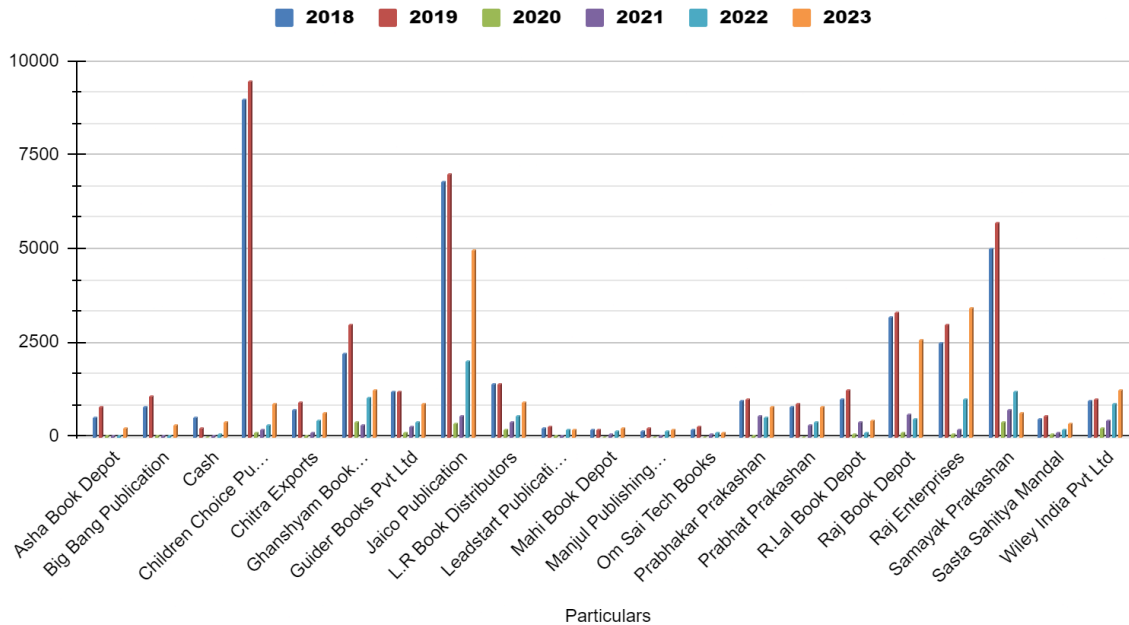


Looking at the line chart for purchases and sales over time, it's clear that in 2018 and 2019, sales were higher than purchases, leading to good profits. However, 2020 brought a sharp drop in both areas, likely due to the pandemic's impact. There was a bit of a rebound in 2021, but profits didn't reach the pre-pandemic levels. In 2022 and 2023, we see growth in both purchases and sales, but profits are still catching up to where they were before the pandemic. This shows that while the business is recovering, it's still working to return to its previous level of profitability.



The bar chart clearly highlights the store's financial fluctuations over the years. It reflects a period of strong profitability before the pandemic, a significant downturn in 2020 due to external disruptions, and a gradual recovery in subsequent years. The increasing profits in 2022 and 2023 suggest that the store is on a positive trajectory, recovering from the pandemic's impact and showing signs of financial stability and growth.

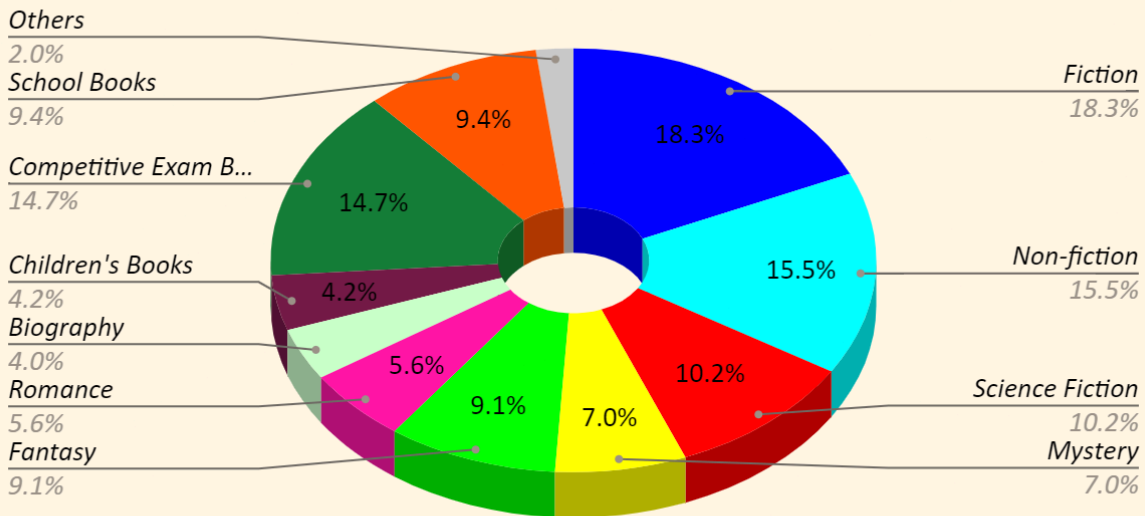
Trends in Publisher Sales Over Time



The chart reveals that while some publishers experienced a significant dip in sales during 2020, likely due to the pandemic, most have shown a recovery or growth trend in the subsequent years, reflecting their ability to adapt and rebound from market disruptions.

In summary, the bar chart provides a clear view of how different publications fared over the years, highlighting trends, recoveries, and the overall dynamics of the book sales market

Book sales distribution by genre



The pie chart of book sales distribution reveals clear patterns in reader interests. **Fiction** leads with 18.30% of total sales, reflecting its broad appeal and widespread popularity. **Non-fiction** follows closely at 15.50%, indicating a strong interest in factual and educational content. The notable share of **Competitive Exam Books** at 14.70% suggests a significant demand for study materials among readers preparing for exams. **Science Fiction** and **Fantasy** capture 10.20% and 9.10%, respectively, showing a robust interest in imaginative and speculative genres. **School Books** hold 9.45%, underscoring a steady need for educational resources. Genres like **Mystery** and **Romance** have smaller shares at 7% and 5.60%, indicating more niche interests. **Biography** and **Children's Books** attract modest attention with 4% and 4.20%, respectively, while the **Others** category at 2% represents more specialized genres. Overall, the data reveals a diverse range of reading interests, with a strong preference for Fiction and educational materials.