

GOVERNMENT OF INDIA

**MINISTRY OF SKILL DEVELOPMENT & ENTREPRENEURSHIP**

DIRECTORATE GENERAL OF TRAINI

**NATIONAL SKILL TRAINING INSTITUTE**

NSTI (W), Salt Lake, Kolkata-700091

**CERTIFICATE**

This is to certify that following trainees have completed their project titled

**“NSTI Smart Cloud Campus”**

**For IBM Program – IT, Networking and Cloud (Technical Diploma)**

|  |  |
| --- | --- |
| **Roll Number** | **Name** |
| ADIT19AU01323 | Sakshi Kumari |
| ADIT19AU00030 | Shreya Das |
| ADIT19AU01056 | Aditi Sau |
| ADIT19AU04163 | Shweta Roy |

|  |  |  |  |
| --- | --- | --- | --- |
| Mr. Abhishek Raman | Mr. K. L. Kuli | Mr. G. C. Ramamurthy | Mr. Sarbojit Neogi |
| IBM Faculty | ADIT Director | Principal | Section-In-Charge |

##### ACKNOWLEDGEMENT

We would like to express my sincere gratitude to several individuals and organizations for supporting me throughout my diploma study. First, I wish to express my sincere gratitude to my trainer, Mr. Sarbojit Neogi, for her enthusiasm, patience, insightful comments, helpful information, practical advice and unceasing ideas that have helped me tremendously at all times in my study and writing of the project report. Her immense knowledge, profound experience and professional expertise in computer science has enabled me to complete the project successfully.

I also wish to express my sincere thanks to the National Skill Training Institute, Salt Lake for accepting me into the diploma program. In addition, I am deeply indebted to the Ministry of Skill Development & Entrepreneurship and IBM for granting me the diploma course. The technical and financial support has enabled me to complete my diploma course studies successfully. Also, I am grateful the Mr. Sarbojit Neogi faculty of NSTI Salt Lake for supporting me for course completion in the specific subject.

I am also grateful to the following NSTI Regional Director Sir Mr. K. L. Kuli and Principal Sir Mr. G. C. Ramamurthy for their consistent support and assistance.

## Project on Data Visualization

REPORT OF MAJOR PROJECT SUBMITTED FOR PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DIPLOMA IN IT, NETWORKING AND CLOUD COMPUTING

*…..Submitted by*

***Sakshi Kumari***

### TABLE OF CONTENTS

1. Introduction
2. Technology used
3. Topic
4. Dashboard of Data Visualization
5. Data Visualization and Excel
6. Types of Data Visualization
7. Data Visualization Tools
8. Reports created with description
9. Conclusion

#### Introduction

Data visualization is the representation of data or information in a graph, chart, or other visual format. It communicates relationships of the data with images. It helps users in analyzing a large amount of data in a simpler way. It makes complex data more accessible, understandable and usable.

Data visualization has its roots in the field of [Statistics](https://en.wikipedia.org/wiki/Statistics) and is therefore generally considered a branch of [Descriptive Statistics.](https://en.wikipedia.org/wiki/Descriptive_Statistics) However, because both design skills and statistical and computing skills are required to visualize effectively.

#### Technology used

Excel 2013 is used on Windows 10 platform to create this real time report. Microsoft Excel is a spreadsheet developed by Microsoft for Windows, macOS, Android and iOS. It features calculation, graphing tools, charts, pivot tables, and a macro programming language called Visual Basic for Applications (VBA). It is typically used to record and analyze numerical and statistical data.

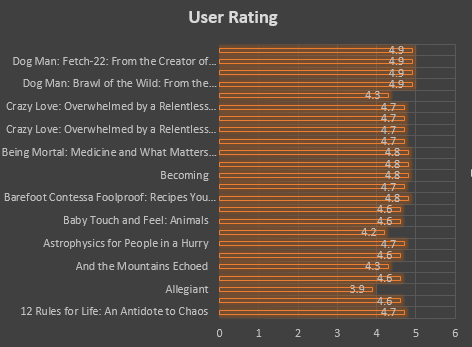
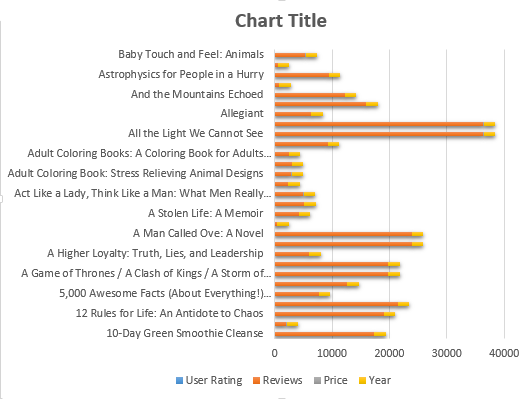
**Topic**

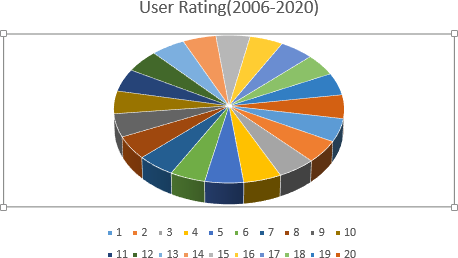
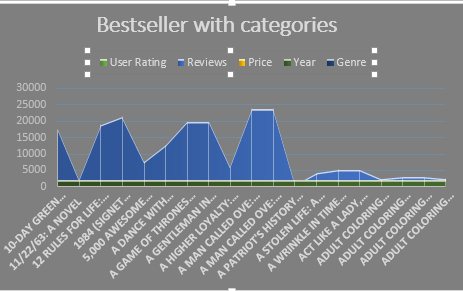
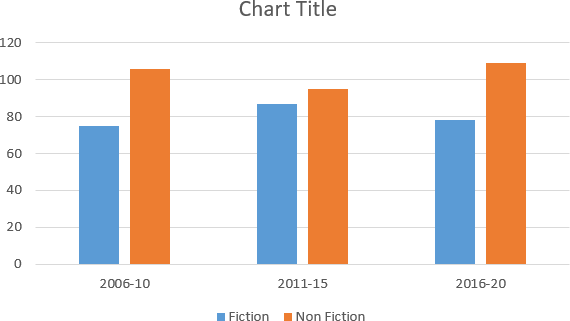
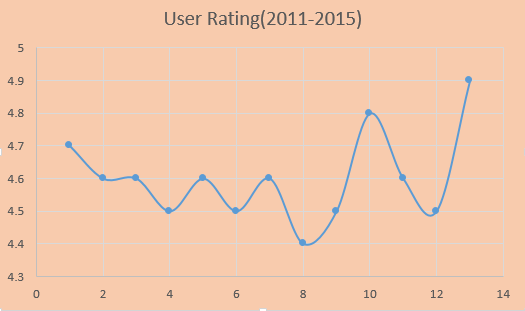
# Data Visualization on Amazon Best seller books

2006-2020



Dashboard of Data Visualization



#### Data Visualization and Excel

Excel can perform basic data analysis functions, it has a surprising number of data visualization tools under the hood. “Excel isn't explicitly a data visualization tool, it's a spreadsheet,” says Excel developer and consultant Jon Peltier. But we can display our data analysis reports in a number of ways in Excel, because- in Excel, charts are used to make a graphical representation of any set of data. A chart is a visual representation of the data and Excel Recommended Charts option to customized data.

Excel covers a vast range of data sources to connect with for data visualization. Excel provides visualization tools for free.

The most popular Data visualization software Power Bi, Data pine are provides a strong backend data manipulation feature with access to simple visualizations and connect much larger datasets. But the software like power bi are heavy software to load data and made visualization in computer. Sometimes it take much more time but not work properly. For some queries, Power BI was only able to process 20000-30000 rows without running into issues. There's no clear reason why this is happening, it would just sometimes timeout during the processing if the rows amount was over 20000.

Types of Data Visualization

1. **Common general types of data visualization:**

* Charts
* Tables
* Graphs
* Maps
* Info graphic
* Dashboards

1. **More specific examples of methods to visualize data:**

* Area Chart
* Bar Chart
* Highlight Table
* Histogram
* Scatter Plot (2D or 3D)
* Streamgraph
* Text Tables
* Timeline
* Tree map
* Box-and-whisker Plots
* Bubble Cloud
* Bullet Graph

**Data Visualization Tools**

Data visualization tools can be used in a variety of ways. The most common use today is as a business intelligence ([BI](https://searchbusinessanalytics.techtarget.com/definition/business-intelligence-BI)) reporting tool. Users can set up visualization tools to generate automatic dashboards that track company performance across key performance indicators ([KPIs](https://searchbusinessanalytics.techtarget.com/definition/key-performance-indicators-KPIs)) and visually interpret the results.

The generated images may also include interactive capabilities, enabling users to manipulate them or look more closely into the data for questioning and analysis. Indicators designed to alert users when data has been updated or when predefined conditions occur can also be integrated.

**Popular tool for data visualization**

* **Microsoft Excel**
* **Microsoft Power BI**
* **IBM Cogons Analytics**
* **Qlik Sense and QlikView**
* **Oracle Visual Analyzer**
* **SAP Lumira**
* **SAS Visual Analytics**
* **Jupyter**
* **MicroStrategy**
* **Google Charts**

**Feasibility Study**

Excel was developed and released by Microsoft Corporation in 1985 for Microsoft Windows OS. It had undergone many modifications during the last 30 years and many advanced versions are now available for other operating systems like MacOS X, iOS etc. Microsoft Excel is one of the most powerful tools used for data management.

It is a software which has a battery of supplied functions to answer statistical, engineering, and financial needs. In addition, it can display data as line graphs, histograms and charts, and with a very limited three-dimensional graphical display. Recently Excel supports programming through Microsoft's Visual Basic for Applications (VBA), which includes a window for writing code, debugging code, and code module organization environment.

Benefits:

* + - In Excel, charts are used to make a graphical representation of any set of data. A chart is a visual representation of the data and Excel Recommended Charts option to customized data.
* Excel covers a vast range of data sources to connect with for data visualization.
  + - Excel provides visualization tools for free.

Reports created with description

**Important KPIs on which reports are made**

* Books price are given in hundreds
* User rating are given in decimal
* Genre of books
* Time span of 15 years(2006- 2020)

**Graph Of the above KPIs are given below with description**

**1) Book reviews v/s prices during 2006-2010**



**2006-2010 best books reviews & price**

In the year of 2006-2015 best books reviews & prices chart we used Combo chart where Reviews are shown with the chart type Clustered bar in light orange color and Price are shown with the chart type Stacked area in light green color.

Here we found the top 10 highest reviews get most of the books on the genre of Fiction. So we understood that public love fiction books too much. But also there is one non -fiction book which is getting the highest review but it is only one out of ten. Actually public can give the rating and review after buying this book. So we can see here how many of people are involved with this book and they have good reading habit in their busy life. And here the graph of the prices of books, where we can see most people involved with the cheap books. The top 10 costly books are get reviews between 900-8600 and most of them are released in 2009.

##### 2) Book reviews v/s prices during 2011-2015



In the year 2011-2015 best books reviews & prices chart we used Combo chart where Reviews are shown with the chart type Area in yellow color and Price are shown with the chart type Line in black color. Here we can see most people involved with the cheap books which is shown in the graph. And also we can see that 30000 people gives reviews in the previous graph and here 80000 people gives reviews so day by day more people are involved with books are increasing, they increased their good habit- reading habit. And here also public loved fiction books because top 10 books by reviews are from the genre fiction.

##### 3) Book reviews v/s prices during 2016-2020

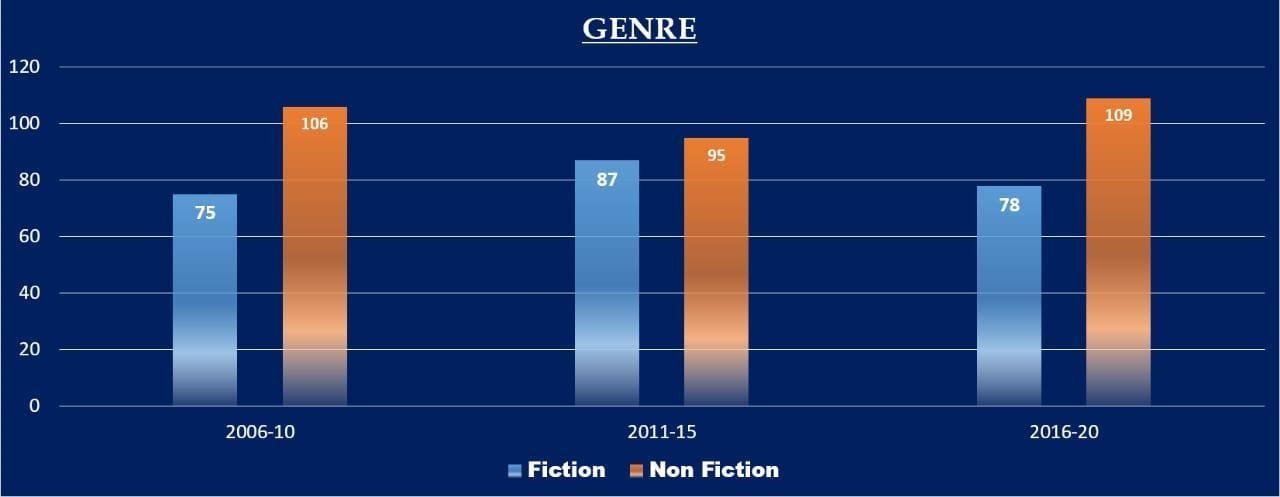


In the year 2016-2020 best books reviews & prices chart we used Combo chart where Reviews are shown with the chart type Clustered bar in pink color and Price are shown with the chart type Clustered column in green color. Here we can see that 80000 people gives reviews in the previous graph and here 88000 people gives reviews so day by day more people are involved with books are increasing, they increased their good reading habit.

As we can see through this chart the reviews each book has increased by the time. It means more people have read these books. We know by increasing demand the price of the things also increased. And the increased rate of books also shown in the graph.

##### 4) Demanding Genre of 2006-2010/2011-2015/2016-2020

In this three section we found year wise public demanding books in those particular genre Fiction and Non-fiction.



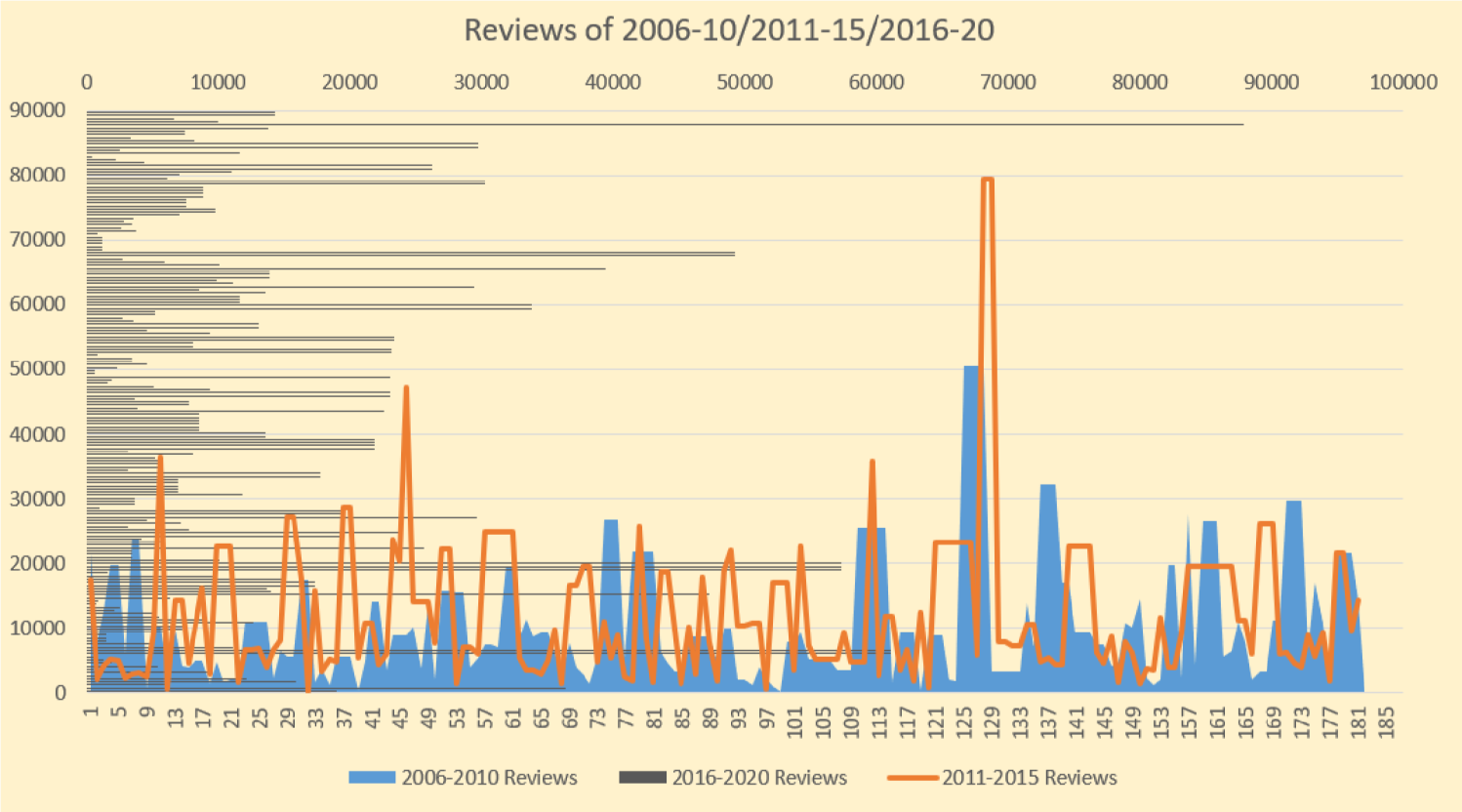
In this chart we used simple bar chart. Here we want to show how much people involved and buy the books type of genre fiction and non- fiction.

In 2006-2010 people choose to buy non-fiction genre book more than fiction genre, but in between 2011-2015 public changed their taste, so we can see the reviews of fiction & non-fiction keep their position quite equal, but non-fiction genre still high than fiction genre.

In between 2016-2020 the position of public reviews of fiction and nonfiction looks like as previous year 2006-2010. Here people choose to buy nonfiction genre book more than fiction genre.

##### 

##### 5) Analyzing Reviews Year wise



In this Reviews of analyzing 2006-2010/2011-2015/2020 chart we used

Combo chart where 2006-2010 Year Reviews are shown with the chart type Stacked Area in sky blue color, 2011-2015 Year Reviews are shown with the chart type Line in orange color and 2016-2020 Year Reviews are shown with the chart type Stacked bar in grey color. Where we can see people give the highest review in the year between 2006 - 2010 than 2016 - 2020.

As we can see through this chart the reviews of each book has increased by the time. It means more people have read these books. We know by increasing demand the price of the things also increased. And the increased rate of books also shown in the graph

**Conclusion**

**Reading helps a great deal in building your confidence, reduces the stress and puts you in a better mood. Once you start reading books and make it your habit, you will eventually get addicted to it. But at this time we all are dependable in many types of gadgets like mobile, laptop etc. And as well as public in their busy life really lost their good habit like reading books.**

**Amazon is an online market place is come to see by a little survey on their 2006-2020 bestseller books with buying and all information. But here we found some good things that people are really keep their good habit in their life. And also we saw that in 2006-2010 the involvement of people is very less and there have some issues like many people are not aware about amazon, gadget using. But day by day they involve them with books, amazon and improve their knowledge. Here amazon also can noticed about the public choice of books and it can help them to increase the business and its interests, like here amazon should keep more non- fiction genre books in their store and online site because after 5 year it can give them more success and profits.**

**So, we can understood how data visualization can help us to realize our profit, loss, do’s and don’ts. In every visualization have a good impact like here we found the reading habit of people as well as amazon can found which type of genre should keep them success in their business.**