

# **RECOMMENDATION SYSTEM PROJECT REPORT**

**SAKSHAM CHAUDHARY**

**15th MAY**

## **ABSTRACT**

For making this project I have taken around 2-3 days for studying about all different library used in this project like numpy , pandas etc. In this project I have taken dataset from kaggle website. So in this project I have used basic math , correlation to pearson correlation for better recommendation of dataset.

## **1 INTRODUCTION**

### **1.1 Problem Statement**

Make a book recommendation system that is filtration program whose prime goal is to predict the rating or preference of user towards a domain specific item.

### **1.2 Engineering Design Task**

1) In this project first I have use pivot table for help to represent the user who have not voted or have not giving rating to any book denoted as NaN.

2) Pivot converts rating table to 2D matrix.

3) I have used 3 datasets in my project for better comparison and recommendation.

4) I have convert the rating table into 2d matrix for better recommendation.

5) I have use pearson correlation for better recommendation.

## **2 LIBRARIES & FUNCTIONS**

### **2.1 Pandas Library**

Open a local file using Pandas, usually a CSV file, but could also be a delimited text file (like TSV), Excel, etc.

### **2.2 Numpy Library**

NumPy is a Python library used for working with arrays. It also has functions for working in linear algebra, Fourier transform, and matrices and performs other mathematical operations.

### **2.3 Matplotlib Library**

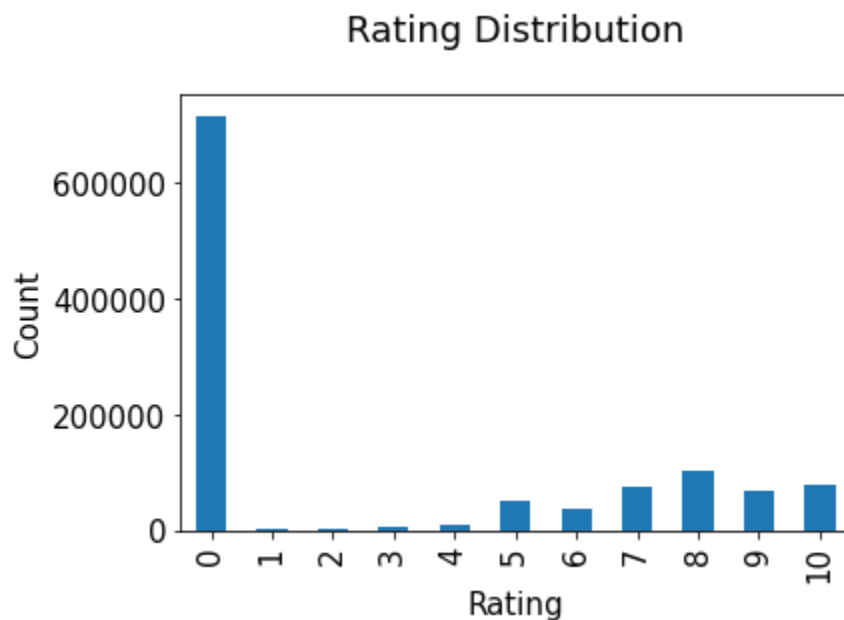
Matplotlib is a library for creating static, animated, and interactive visualizations in Python. With the help of Matplotlib library we can plot bar graphs, pie charts, histograms etc.

## 2.4 Pivot Table

Pivot converts rating table to 2D matrix. It is also use for eg- for help to represent the user who have not voted or have not giving rating to any book denoted as NaN.

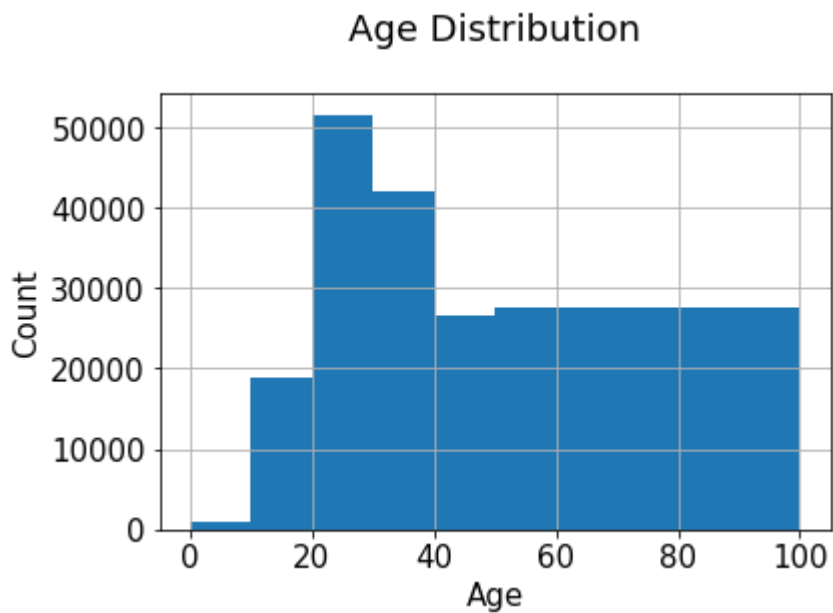
## 3 Visual Representations Used in Project

### 3.1



The above graph shows rating distribution in the form of a bar graph. It shows many how people give book ratings based on age distribution. Comparison between count and rating as factor in the graph.

## 3.2

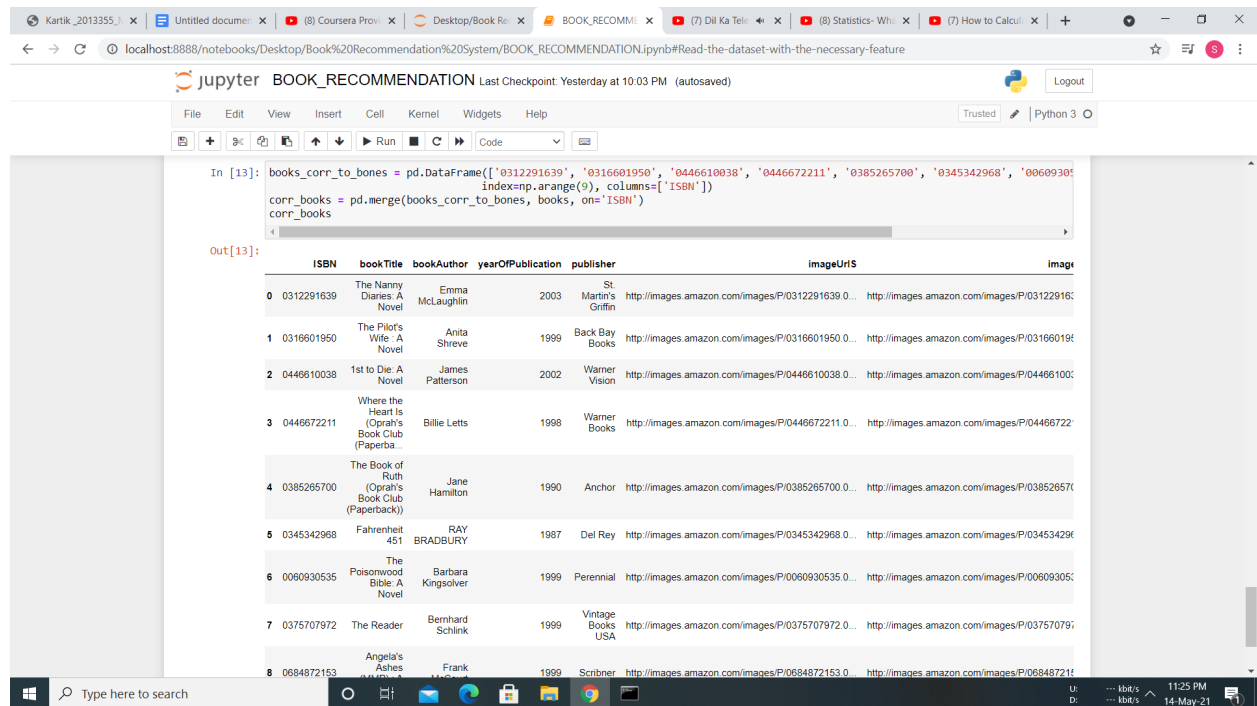


The above graph show age distribution by histogram graph. It show many how people give book rating based on age distribution. Comparison between count and age as factor.

## 4 Conclusions

After implementing my project my conclusions are I have used various data science technique for comparison and recommendation for dataset. I have basic math , rating counts to recommendation based on correlation techniques for enhancement for recommendation.

Here I am attaching the Screenshot of my final result.



```
In [13]: books_corr_to_bones = pd.DataFrame(['0312291639', '0316601950', '0446610038', '0446672211', '0385265700', '0345342968', '0060930535', '0375707972', '0684872153'],
      index=np.arange(9), columns=['ISBN'])
corr_books = pd.merge(books_corr_to_bones, books, on='ISBN')
corr_books
```

```
Out[13]:
```

	ISBN	bookTitle	bookAuthor	yearOfPublication	publisher	imageUriS	imageUriL
0	0312291639	The Nanny Diaries: A Novel	Emma McLaughlin	2003	St. Martin's Griffin	http://images.amazon.com/images/P/0312291639.0...	http://images.amazon.com/images/P/0312291639.0...
1	0316601950	The Pilot's Wife: A Novel	Anita Shreve	1999	Back Bay Books	http://images.amazon.com/images/P/0316601950.0...	http://images.amazon.com/images/P/0316601950.0...
2	0446610038	1st to Die: A Novel	James Patterson	2002	Warner Vision	http://images.amazon.com/images/P/0446610038.0...	http://images.amazon.com/images/P/0446610038.0...
3	0446672211	Where the Heart Is (Oprah's Book Club (Paperback))	Billie Letts	1998	Warner Books	http://images.amazon.com/images/P/0446672211.0...	http://images.amazon.com/images/P/0446672211.0...
4	0385265700	The Book of Ruth (Oprah's Book Club (Paperback))	Jane Hamilton	1990	Anchor	http://images.amazon.com/images/P/0385265700.0...	http://images.amazon.com/images/P/0385265700.0...
5	0345342968	Fahrenheit 451	RAY BRADBURY	1987	Del Rey	http://images.amazon.com/images/P/0345342968.0...	http://images.amazon.com/images/P/0345342968.0...
6	0060930535	The Poisonwood Bible: A Novel	Barbara Kingsolver	1999	Perennial	http://images.amazon.com/images/P/0060930535.0...	http://images.amazon.com/images/P/0060930535.0...
7	0375707972	The Reader	Bernhard Schlink	1999	Vintage Books USA	http://images.amazon.com/images/P/0375707972.0...	http://images.amazon.com/images/P/0375707972.0...
8	0684872153	Angela's Ashes	Frank McCourt	1999	Scribner	http://images.amazon.com/images/P/0684872153.0...	http://images.amazon.com/images/P/0684872153.0...

## 5 References :-

### 5.1 Numpy

<https://numpy.org/devdocs/user/quickstart.html>

### 5.2 Pandas



