

# Book Recommendation and Classification

Submitted by –

Jaya Krishna

Mithun Phadnis

Mukhar Gupta

Sumedh Khandeparkar

Varun Vohra

# Objective

- To build a model which will recommend books to users based on their interests and previous choices
- Aligns to other users with similar interests and makes appropriate suggestions to enhance their reading experience

# About the Dataset

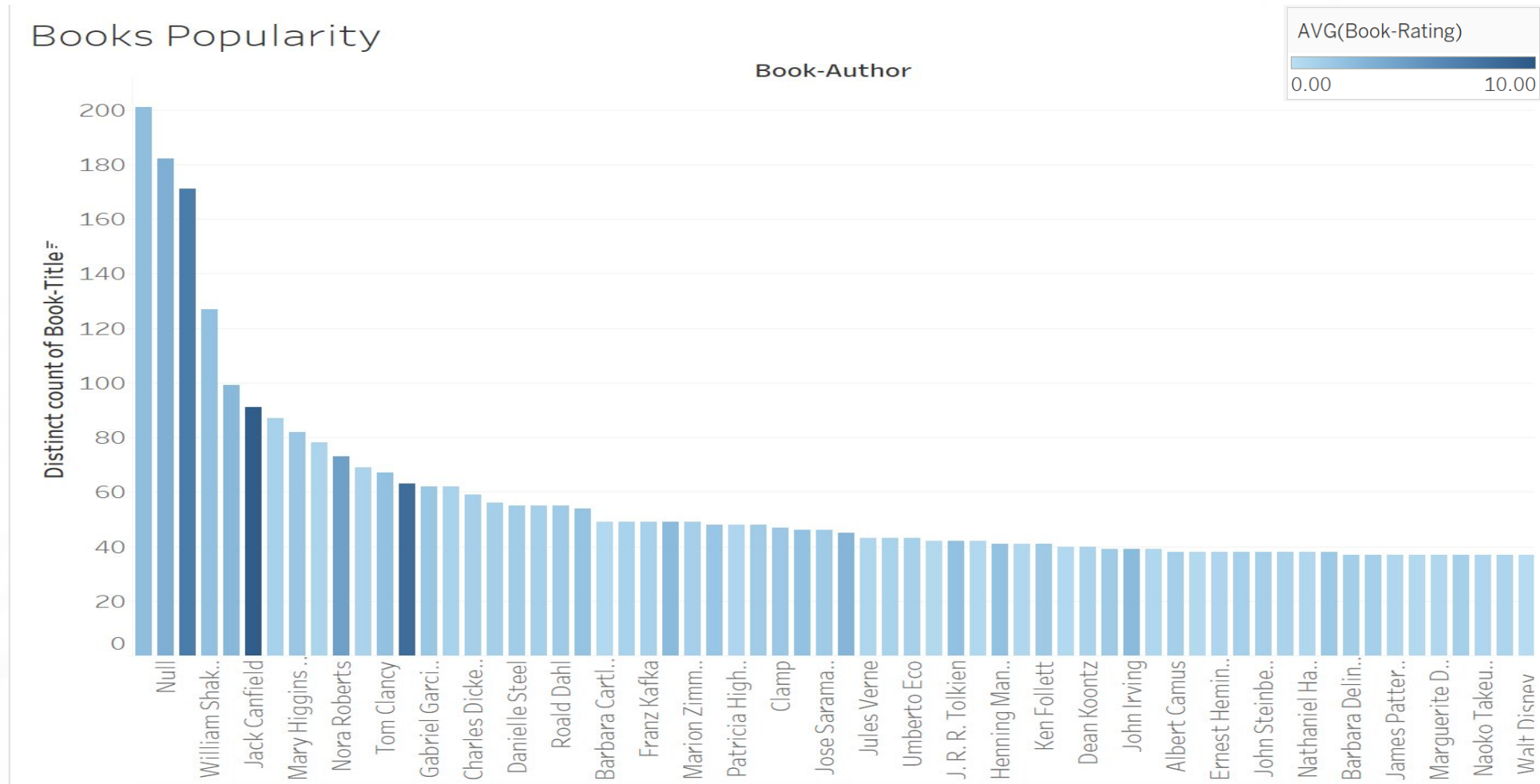
- **Book-Crossing dataset** -- <http://www.bookcrossing.com>
- **Source** -- <http://www2.informatik.uni-freiburg.de/~chiegler/BX/>
- Collected in a 4-week stretch (August / September 2004)
- Contains 278,858 users providing 1,149,780 ratings about 271,379 books

# About the Dataset (Conti.)

Dataset comprises 3 tables

- **Users:**
  - Contains the list of users and their demographic details
- **Books**
  - Content-based information about the books is provided
  - Books are identified by their respective International Standard Book Number (ISBN)
- **Book-Ratings**
  - Contains the book rating information
  - Ratings are expressed on a scale from 1-10

# Data Insight



# Algorithm : k-Nearest Neighbor (k-NN)

- Primarily used for Classification and Regression
- Instance-based and Lazy Learning Algorithm
- Determines its K nearest neighbors from the training data
- Distance is calculated using following measures:
  - Euclidean Distance
  - Minkowski Distance

## Goal

To achieve best possible accuracy for our recommendation system



[illegible]