
Books Analysis

BY SHYNAR BOLATOVA

id: 23884

IITU, Almaty

Cross-mobile Development | December 2020

Description

The system picking a preferable books and making analysis and book selections from database.

Project Requirement:

Using Python language, for data analysis

Software Requirement:


Jupyter Notebook Application (Anaconda), or other text editors that display the input and output of a Python or language script.

Development stages:

- Gathering data

Data in .CSV File

Link: <https://data.world/josh-nbu/books>

 Books_Data_Clean.csv

[Query](#) [Download](#) [Open in app](#)

	year	book_name	author	language_code	author_rating	#	book_average_rating
1	1975	Beowulf	Unknown, Seamus Heaney	en-US	Novice		3,42
2	1987	Batman: Year One	Frank Miller, David Mazzucchelli, Ric	eng	Intermediate		4,23
3	2015	Go Set a Watchman	Harper Lee	eng	Novice		3,31
4	2008	When You Are Engulfed in Flames	David Sedaris	en-US	Intermediate		4,04
5	2011	Daughter of Smoke & Bone	Laini Taylor	eng	Intermediate		4,04
6	2015	Red Queen	Victoria Aveyard	eng	Intermediate		4,08

Development stages:

- Exploring the data
 - Table creation
 - Table Shape
 - Data types
 - Table properties
 - Column properties

Development stages:

- Preproceccing

Data preprocessing includes cleaning, Instance selection, normalization, transformation, feature extraction and selection, etc.

Development stages:

- Data Analysis
 - Summarizing the Data
 - Finding relevant information from this data

Development stages:

- Visualization

Python data visualization libraries I used:

- Matplotlib
- Seaborn

Development stages:

- Documentation
 - headings
 - markdowns
 - comments

Goal result:

I will make an analysis and find answers to these questions:

Which are the most famous books today?

Which are the best book authors?

Which books may fit your preferences?

Who are the publishers in this dataset?

How many books was written by each athor? Who had wrote the greatest amount of book?

What is the best seller book?