

File

Edit

Selection

View

Go

Run

...

←

→

Assignment 4

□

□

□

□

—

□

×

assignment.ipynb ×

assignment.ipynb > MData analysis > MMost selling harry potter books > Mdisplay most selling books of harry potter (in order) > print("Most selling harry potter books:")

+ Code + Markdown | ▶ Run All ↺ Restart ≡ Clear All Outputs | [V] Variables ≡ Outline ...

base (Python 3.11.7)

▶

```
# import data
data = pd.read_csv("books.csv")
# display first 5 records
data.head()
```

[3] ✓ 0.0s Python

...

	book_id	goodreads_book_id	best_book_id	work_id	books_count	isbn	isbn13	authors	original_publication_year	original_title	...	ratings_count
0	1	2767052	2767052	2792775	272	439023483	9.780000e+12	Suzanne Collins	2008.0	The Hunger Games	...	47806
1	2	3	3	4640799	491	439554934	9.780000e+12	J.K. Rowling, Mary GrandPré	1997.0	Harry Potter and the Philosopher's Stone	...	46024
2	3	41865	41865	3212258	226	316015849	9.780000e+12	Stephenie Meyer	2005.0	Twilight	...	38668
3	6	11870085	11870085	16827462	226	525478817	9.780000e+12	John Green	2012.0	The Fault in Our Stars	...	23464
4	12	13335037	13335037	13155899	210	62024035	9.780000e+12	Veronica Roth	2011.0	Divergent	...	19035

5 rows × 23 columns

1

✕ 0 ⚠ 0 🔊 0

CRLF 🔔 {}

79°F Sunny

Search

4:35 PM 4/22/2024

ENG 🔊 🔌 🔋

PRE

File

Edit

Selection

View

Go

Run

...

←

→

Assignment 4

□

□

□

□

—

□

×

assignment.ipynb ×

assignment.ipynb > MData analysis > MMost selling harry potter books > Mdisplay most selling books of harry potter (in order) > print("Most selling harry potter books:")

+ Code + Markdown | ▶ Run All ↺ Restart ≡ Clear All Outputs | [x] Variables ≡ Outline ...

base (Python 3.11.7)

▶

data.isnull().sum()

[4] ✓ 0.0s

Python

...

book_id	0
goodreads_book_id	0
best_book_id	0
work_id	0
books_count	0
isbn	52
isbn13	44
authors	0
original_publication_year	3
original_title	52
title	0
language_code	109
average_rating	0
ratings_count	0
work_ratings_count	0
work_text_reviews_count	0
ratings_1	0
ratings_2	0
ratings_3	0
ratings_4	0
ratings_5	0
image_url	0
small_image_url	0
dtype:	int64

<

0

0

0

0

CRLF

🔔

{ }

79°F

Sunny

🌐

🔍 Search

🌐

📄

🗨

📊

📺

🎮

🔧

🌐

🔍

📁

🌐

^

ENG

📶

🔊

🔌

4:36 PM

4/22/2024

🔔

PRE

File

Edit

Selection

View

Go

Run

...

←

→

Assignment 4

🏠

📄

🔍

🔗

🔧

📦

🧪

assignment.ipynb

✕

assignment.ipynb > MData analysis > MMost selling harry potter books > Mdisplay most selling books of harry potter (in order) > print("Most selling harry potter books:")

+ Code + Markdown | ▶ Run All ↺ Restart ≡ Clear All Outputs | 📄 Variables ≡ Outline ...

base (Python 3.11.7)

Handle duplicates

+ Code + Markdown

```
# Check for duplicated books count
duplicated_books_count = data.duplicated().sum()
duplicated_books_count
```

[32] ✓ 0.0s Python

... 0

Data analysis

Select Harry Potter books

```
# select books that contains the term "harry potter" in ["title","original_title"] features
harry_potter_books = data.loc[ data["original_title"].str.contains("harry potter" , case=False) |
                               data["title"].str.contains("harry potter" , case=False)]
# Reset index
```

🔍 0 ⚠️ 0 🗣️ 0

Ln 9, Col 19 CRLF

🔔 {}

79°F Sunny

🪟 🔍 Search 🌐 📄 🗨️ 📊 🖨️ 🗑️ 🔄

ENG 📶 🔊 🔌 4:36 PM 4/22/2024 🌤️ PRE

FileEditSelectionViewGoRun

Assignment 4

assignment.ipynb

assignment.ipynb > MData analysis > MMost selling harry potter books > Mdisplay most selling books of harry potter (in order) > print("Most selling harry potter books:")

CodeMarkdownRun AllRestartClear All OutputsVariablesOutline

base (Python 3.11.7)

book_idgoodreads_book_idbest_book_idwork_idbooks_countisbnisbn13authorsoriginal_publication_yearoriginal_title...ratings_co

023346407994914395549349.780000e+12J.K. Rowling, Mary GrandPré1997.0Harry Potter and the Philosopher's Stone...4602

118552402163376043965548X9.780000e+12J.K. Rowling, Mary GrandPré, Rufus Beck1999.0Harry Potter and the Prisoner of Azkaban...1832

221228092033074393580789.780000e+12J.K. Rowling, Mary GrandPré2003.0Harry Potter and the Order of the Phoenix...173!

323158811588162311713984390648649.780000e+12J.K. Rowling, Mary GrandPré1998.0Harry Potter and the Chamber of Secrets...1779

424630465723324391396009.780000e+12J.K. Rowling, Mary GrandPré2000.0Harry Potter and the Goblet of Fire...1753

5251262511262512622102625150102259.780000e+12J.K. Rowling,2007.0Harry Potter and the...174

Ln 9, Col 19CRLF

79°F Sunny

Search

4:36 PM4/22/2024

File

Edit

Selection

View

Go

Run

...

←

→

Assignment 4

□

□

□

□

—

□

×

assignment.ipynb ×

assignment.ipynb > MData analysis > MMost selling harry potter books > Mdisplay most selling books of harry potter (in order) > print("Most selling harry potter books:")

+ Code + Markdown | ▶ Run All ↺ Restart ≡ Clear All Outputs | [x] Variables ≡ Outline ...

base (Python 3.11.7)

Average rating of harry potter books

▶

avg_rating_of_harry_potter_books = (harry_potter_books["average_rating"]).sum()/harry_potter_books["average_rating"].count()
print(f"Average rating of harry potter books: {avg_rating_of_harry_potter_books}")

[9] ✓ 0.0s Python

... Average rating of harry potter books: 4.4910000000000005

Most selling harry potter books

display data related to sales

[16] ✓ 0.0s Python

	books_count	work_ratings_count	work_text_reviews_count
0	491	4800065	75867
1	376	1969375	36099

79°F Sunny

Search

4:36 PM 4/22/2024

