Name: Tejas Chincholkar

Roll no:CS5-71

PRN: 202401100147

## **EDS ACTIVITY**

1. How many rows and columns are there?

2.Get a summary (info) of the dataset.

```
Welcome
                 newpanda.py X
                                    books.csv
 newpanda.py > ...
       import pandas as pd
       import numpy as np
       df = pd.read_csv('E:\\Pandas\\books.csv')
       print(df.info())
 PROBLEMS
           OUTPUT
                    DEBUG CONSOLE
                                    TERMINAL
                                              PORTS
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
 <class 'pandas.core.frame.DataFrame'>
 RangeIndex: 11127 entries, 0 to 11126
 Data columns (total 13 columns):
     Column
                         Non-Null Count Dtype
     bookID
  0
                         11127 non-null int64
  1
     title
                         11127 non-null object
     authors
                         11127 non-null object
  2
                         11127 non-null object
  3
     average_rating
  4
     isbn
                         11127 non-null object
     isbn13
  5
                         11127 non-null object
                         11127 non-null
  6
     language_code
                                        object
  7
      num_pages
                         11127 non-null object
                         11127 non-null int64
  8
    ratings_count
  9
    text reviews count 11127 non-null int64
  10 publication_date
                        11127 non-null object
  11 publisher
                         11127 non-null object
  12 Unnamed: 12
                         4 non-null
                                         object
 dtypes: int64(3), object(10)
 memory usage: 1.1+ MB
 None
 PS E:\Pandas>
```

## 3. Find the book with the highest rating.

```
Welcome
               newpanda.py X Books.csv
 newpanda.py > ...
       import pandas as pd
       import numpy as np
      df = pd.read csv('E:\\Pandas\\books.csv')
       print(df.sort_values(by='average_rating', ascending=False).head(1))
                                                                                                                        Pyth
                                  TERMINAL
PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
                                                       title
                                                                   authors ... publication_date
                                                                                                       publisher Unnamed: 12
     20957 Colossians and Philemon: A Critical and Exeget... R. McL. Wilson ...
                                                                                     12-07-2005 T&T Clark Int'l
 [1 rows x 13 columns]
PS E:\Pandas>
```

4. Find the book with the lowest rating.

```
★ Welcome

               newpanda.py > ...
       import numpy as np
      df = pd.read_csv('E:\\Pandas\\books.csv')
       print(df.sort_values(by='average_rating', ascending=True).head(1))
 PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
                                                                                                                  Python + √ [
PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
                                                    title
                                                                authors ... publication_date publisher
                                                                                                                   Unnamed: 12
      12224 Streetcar Suburbs: The Process of Growth in Bo... Sam Bass Warner ...
                                                                                           6 4/20/2004 Harvard University Press
 [1 rows x 13 columns]
 PS E:\Pandas>
```

5. Find all books written by 'J.K. Rowling'.

```
₩elcome
                   newpanda.py X Books.csv
 newpanda.py > ...
        import numpy as np
        df = pd.read_csv('E:\\Pandas\\books.csv')
        print(df[df['authors'] == 'J.K. Rowling'])
 PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

    Python + ∨ |

 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
                                                                    title
                                                                                                                                      publisher Unnamed: 12
                                                                                 authors ... publication_date
              4 Harry Potter and the Chamber of Secrets (Harry... J.K. Rowling \,\ldots\,
                                                                                                       11-01-2003
                                                                                                                                     Scholastic
                                                                                                                                                           NaN
                       Harry Potter Collection (Harry Potter #1-6) J.K. Rowling
                                                                                                       09-12-2005
                                                                                                                                     Scholastic
                                                                                                                                                           NaN
             10
 613
           2002 Harry Potter Schoolbooks Box Set: Two Classic ... J.K. Rowling
                                                                                                       11-01-2001
                                                                                                                              Arthur A. Levine
                                                                                                                                                           NaN
           2005 Harry Potter and the Half-Blood Prince (Harry ... J.K. Rowling 3357 Harry Potter Y La Piedra Filosofal (Harry Pott... J.K. Rowling
                                                                                                                         Bloomsbury Publishing
Turtleback Books
                                                                                                        6/23/2006
 988
                                                                                                       03-06-2001
                                                                                                                                                           NaN
          4256 Harry Potter and the Prisoner of Azkaban (Harr... J.K. Rowling
15876 Harry Potter y la Orden del Fénix (Harry Potte... J.K. Rowling
 1233
                                                                                                       07-01-2008
                                                                                                                                 Bloomsbury UK
                                                                                                                                                           NaN
                                                                                                                                 Emece Editores
 4413
                                                                                                        2/21/2004
                                                                                                                                                           NaN
 10347
          41907 Harry Potter und die Kammer des Schreckens (Ha... J.K. Rowling
                                                                                                        9/38/2888
                                                                                                                                         Carlsen
                                                                                                                                                           NaN
          41911 Harry Potter und der Gefangene von Askaban (Ha... J.K. Rowling
43504 Harry Potter and the Philosopher's Stone (Harr... J.K. Rowling
                                                                                                       84-81-2881
                                                                                                                                                           NaN
                                                                                                       07-01-2010 Bloomsbury USA Childrens
 10678
                                                                                                                                                           NaN
         43509 Harry Potter and the Goblet of Fire (Harry Pot... J.K. Rowling
                                                                                                       07-08-2000
                                                                                                                                     Bloomsbury
                                                                                                                                                           NaN
 [11 rows x 13 columns]
 PS E:\Pandas>
```

6. Find all books written in en-UK language.

```
× Welcome
                newpanda.py X
                                   books.csv
 newpanda.py > ...
       import pandas as pd
       import numpy as np
       df = pd.read_csv('E:\\Pandas\\books.csv')
       print(df[df['language_code'] == 'en-US'])
 PROBLEMS
           OUTPUT
                    DEBUG CONSOLE
                                   TERMINAL
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
                                                          title ...
       bookID
                                                                                     publisher Unnamed: 12
            9 Unauthorized Harry Potter Book Seven News: "Ha... ...
                                                                                  Nimble Books
                                                                ... Houghton Mifflin Harcourt
               The Lord of the Rings (The Lord of the Rings ...
 24
                  Hatchet Jobs: Writings on Contemporary Fiction
                                                                                 The New Press
 32
                                                                                                      NaN
                                                The Known World ...
                                                                                   HarperAudio
 40
           68
                                                                                                      NaN
 45
           75
                                              Uncommon Carriers
                                                                          Farrar Straus Giroux
                                                                                                      NaN
                                           Live Rich Die Broke
 11094
        45515
                                                                                  Rodale Books
                                                                                                      NaN
 11095
        45516
                         The Die Broke Financial Problem Solver
                                                                     William Morrow Paperbacks
                                                                                                      NaN
                                                    On the Road ...
                                                                       Penguin Classics
 11100
        45536
                                                                                                      NaN
                                                 Ajax in Action ...
                                                                          Manning Publications
 11108
        45574
                                                                                                      NaN
                                              Whores for Gloria ...
                                                                                 Penguin Books
 11121
        45630
                                                                                                      NaN
 [1408 rows x 13 columns]
 PS E:\Pandas>
```

7. Find the average rating count for each author.

```
★ Welcome
                newpanda.py > ...
       import pandas as pd
       import numpy as np
      df = pd.read_csv('E:\\Pandas\\books.csv')
      print(df.groupby('authors')['ratings_count'].mean().sort_values(ascending=False))
 PROBLEMS OUTPUT DEBUG CONSOLE
                                  TERMINAL
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
 authors
 Stephenie Meyer
                                                                                                         4597666.0
 George Orwell/Boris Grabnar/Peter Škerl
                                                                                                         2111750.0
 Paulo Coelho/Alan R. Clarke/Özdemir İnce
                                                                                                         1631221.0
 Markus Zusak/Cao Xuân Việt Khương
                                                                                                         1516367.0
 J.K. Rowling/Mary GrandPré
                                                                                                         1487330.0
                                                                                                              0.0
 Adrian Guelke/Michael Cox/Fiona Stephen
 Graham Handley
                                                                                                              0.0
 Herman Melville/Hershel Parker/G. Thomas Tanselle/Harrison Hayford/Robert Sandberg/Alma MacDougall Reising
                                                                                                              0.0
 Eleanor J. Gibson
                                                                                                              0.0
 Robert L. Geneve
                                                                                                              0.0
 Name: ratings_count, Length: 6643, dtype: float64
 PS E:\Pandas>
```

8. Average number of reviews per rating category.

```
★ Welcome

                 newpanda.py X
                                    books.csv
 💠 newpanda.py > ...
       import pandas as pd
       import numpy as np
       df = pd.read_csv('E:\\Pandas\\books.csv')
   4
        print(df.groupby('ratings_count')['text_reviews_count'].mean())
 PROBLEMS
            OUTPUT
                     DEBUG CONSOLE
                                    TERMINAL
                                               PORTS
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
 ratings_count
 0
                0.025000
 1
                0.144737
 2
                0.281690
 3
                0.390244
 4
                0.535211
 2339585
            36325.000000
 2418736
            21303.000000
 2457092
           43499.000000
 2530894
            32871.000000
            94265.000000
 Name: text_reviews_count, Length: 5294, dtype: float64
 PS E:\Pandas>
```

9. Correlation between number of reviews and rating count.

```
★ Welcome

                newpanda.py X
                                   books.csv
 newpanda.py > ...
       import pandas as pd
       import numpy as np
       df = pd.read_csv('E:\\Pandas\\books.csv')
       print(df[['text_reviews_count', 'ratings_count']].corr())
   4
 PROBLEMS
           OUTPUT
                    DEBUG CONSOLE
                                   TERMINAL
                                              PORTS
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
                    text reviews count ratings count
                              1.000000
                                            0.865979
 text reviews count
 ratings_count
                              0.865979
                                            1.000000
 PS E:\Pandas>
```

10. Find the oldest book in the list.

12. Correlation between number of reviews and ratings.

```
★ Welcome

                 newpanda.py X
                                    books.csv
 newpanda.py > ...
       import pandas as pd
       import numpy as np
       df = pd.read_csv('E:\\Pandas\\books.csv')
       print(df[['text_reviews_count', 'ratings_count']].corr())
   4
 PROBLEMS
           OUTPUT
                    DEBUG CONSOLE
                                              PORTS
                                    TERMINAL
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
                    text_reviews_count ratings_count
 text_reviews_count
                              1.000000
                                             0.865979
 ratings count
                              0.865979
                                             1.000000
 PS E:\Pandas>
```

13. Find the books published by 'Broadway Books'.

```
★ Welcome

                 newpanda.py X
                                     ■ books.csv
 newpanda.py > ...
        import pandas as pd
        import numpy as np
        df = pd.read csv('E:\\Pandas\\books.csv')
   4
        print(df[df['publisher']=='Broadway Books'])
                     DEBUG CONSOLE
                                     TERMINAL
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
                                                            title
        bookID
                                                                             publisher Unnamed: 12
 12
            21
                             A Short History of Nearly Everything
                                                                        Broadway Books
 13
                                      Bill Bryson's African Diary ... Broadway Books
            22
                                                                                               NaN
            23
                Bryson's Dictionary of Troublesome Words: A Wr... ...
 14
                                                                        Broadway Books
                                                                                               NaN
 15
            24
                                           In a Sunburned Country ...
                                                                        Broadway Books
                                                                                               NaN
            25
                I'm a Stranger Here Myself: Notes on Returning...
 16
                                                                        Broadway Books
                                                                                               NaN
                                                The Da Vinci Code
                                                                        Broadway Books
 311
           969
                                                                                               NaN
          1059
 325
                                                          Shibumi
                                                                        Broadway Books
                                                                                               NaN
                The Last Hours of Ancient Sunlight: The Fate o...
 661
          2127
                                                                        Broadway Books
                                                                                               NaN
 766
          2521
                            Blood Done Sign My Name: A True Story
                                                                        Broadway Books
                                                                                               NaN
 1012
          3444
                                                  Northern Lights ...
                                                                        Broadway Books
                                                                                               NaN
 1013
          3445
                If I Die in a Combat Zone Box Me Up and Ship ... ...
                                                                        Broadway Books
                                                                                               NaN
 1044
          3536
                                                Love @ First Site
                                                                        Broadway Books
                                                                                               NaN
 1157
          3858
                                    Holy Cow: An Indian Adventure ...
                                                                        Broadway Books
                                                                                               NaN
                             A Short History of Nearly Everything
          3870
 1161
                                                                        Broadway Books
                                                                                               NaN
 1534
          5319
                                                A Christmas Story
                                                                        Broadway Books
                                                                                               NaN
 2006
          7268
                          Under The Tuscan Sun - At Home In Italy
                                                                        Broadway Books
                                                                                               NaN
                     The Metamorphosis (Graphic Novel Adaptation) ... Broadway Books
 2144
          7718
                                                                                               NaN
 2485
          9073
                See No Evil: The True Story of a Ground Soldie... ...
                                                                                               NaN
                                                                        Broadway Books
 2661
          9789
                A Walk in the Woods: Rediscovering America on ... ...
                                                                        Broadway Books
                                                                                               NaN
                      Youth in Revolt: The Journals of Nick Twisp ...
 2726
                                                                        Broadway Books
                                                                                               NaN
         11077
                The Sex Lives of Cannibals: Adrift in the Equa... ...
                                                                        Broadway Books
 3010
                                                                                               NaN
         16007
 4451
                Last Train to Paradise: Henry Flagler and the ...
                                                                        Broadway Books
                                                                                               NaN
                The Corner: A Year in the Life of an Inner-Cit... ...
 5253
                                                                        Broadway Books
                                                                                               NaN
                The Hunt for Zero Point: Inside the Classified... ...
 5896
         22198
                                                                        Broadway Books
                                                                                               NaN
                For the Sins of My Father: A Mafia Killer His... ...
 5916
         22239
                                                                        Broadway Books
                                                                                               NaN
         22649
 6844
                                                         Bookends ...
                                                                        Broadway Books
                                                                                               NaN
 6231
         23503
                                         The Alchemist's Daughter ...
                                                                                               NaN
                                                                        Broadway Books
 6244
         23526
               Classic Myths to Read Aloud: The Great Stories... ...
                                                                                               NaN
                                                                        Broadway Books
 6544
                Isolde Queen of the Western Isle (Tristan and...
         24674
                                                                        Broadway Books
                                                                                               NaN
                The Best Recipes in the World: More Than 1 000... ...
 6810
                                                                        Broadway Books
                                                                                               NaN
                The Minimalist Cooks at Home: Recipes That Giv... ...
 6811
                                                                        Broadway Books
 6857
         25812
                                                    The Ruby Ring
```

14Find the number of books published per day.

```
⋈ Welcome
                newpanda.py X | Books.csv
 newpanda.py > ...
   1 import pandas as pd
   2 import numpy as np
   3 df = pd.read_csv('E:\\Pandas\\books.csv')
     books_per_date = df['publication_date'].value_counts()
     print(books_per_date[books_per_date > 1])
   6
 PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
 publication_date
 10-01-2005
 11-01-2005
             53
 09-01-2006
             51
 10-01-2006
             48
 11-01-2006
             40
 1/26/2005
             2
 12/31/2003
 02-01-1984
             2
 7/17/2000
 5/15/1992
 Name: count, Length: 1658, dtype: int64
 PS E:\Pandas>
```

15. Find the list of books of 'Harry Potter'.

```
★ Welcome

                                     ■ books.csv
                 newpanda.py X
 newpanda.py > ...
        import pandas as pd
        import numpy as np
       df = pd.read_csv('E:\\Pandas\\books.csv')
       harry_potter_books = df[df['title'].str.contains('Harry Potter', case=False, na=False)]
        print(harry_potter_books[['title', 'authors']])
                                     TERMINAL
                                               PORTS
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
                                                    title
 θ
        Harry Potter and the Half-Blood Prince (Harry ...
                                                                                  J.K. Rowling/Mary GrandPré
        Harry Potter and the Order of the Phoenix (Har...
                                                                                  J.K. Rowling/Mary GrandPré
 1
        Harry Potter and the Chamber of Secrets (Harry...
                                                                                                J.K. Rowling
 2
 3
        Harry Potter and the Prisoner of Azkaban (Harr...
                                                                                  J.K. Rowling/Mary GrandPré
        Harry Potter Boxed Set Books 1-5 (Harry Potte...
                                                                                  J.K. Rowling/Mary GrandPré
 4
 5
        Unauthorized Harry Potter Book Seven News: "Ha...
                                                                                      W. Frederick Zimmerman
            Harry Potter Collection (Harry Potter #1-6)
                                                                                                J.K. Rowling
 6
        Harry Potter Schoolbooks Box Set: Two Classic ...
                                                                                                J.K. Rowling
 613
        J.K. Rowling's Harry Potter Novels: A Reader's...
                                                                                                  Philip Nel
 614
        Harry Potter and the Half-Blood Prince (Harry ...
 615
                                                                                                J.K. Rowling
        Harry Potter Y La Piedra Filosofal (Harry Pott...
 988
                                                                                                J.K. Rowling
 1233
        Harry Potter and the Prisoner of Azkaban (Harr...
                                                                                                J.K. Rowling
 4243
                          Looking for God in Harry Potter
                                                                                                John Granger
 4411
        Mugglenet.Com's What Will Happen in Harry Pott... Ben Schoen/Andy Gordon/Gretchen Stull/Emerson ...
 4412
        Harry Potter y el misterio del príncipe (Harry...
                                                                            J.K. Rowling/Gemma Rovira Ortega
                                                                                                J.K. Rowling
 4413
       Harry Potter y la Orden del Fénix (Harry Potte...
       Ultimate Unofficial Guide to the Mysteries of ...
                                                                           Galadriel Waters/Astre Mithrandir
 4414
 4416
       Harry Potter and the Chamber of Secrets (Harry...
                                                                                  J.K. Rowling/Mary GrandPré
                                                                               David Baggett/Shawn E. Klein
 8266
        Harry Potter and Philosophy: If Aristotle Ran ...
 8876
        Harry Potter and the Sorcerer's Stone (Harry P...
                                                                                  J.K. Rowling/Mary GrandPré
 10347 Harry Potter und die Kammer des Schreckens (Ha...
                                                                                                J.K. Rowling
 10348 Harry Potter und der Gefangene von Askaban (Ha...
                                                                                     J.K. Rowling/Rufus Beck
 10349
         Harry Potter ve Sırlar Odası (Harry Potter #2)
                                                                                    J.K. Rowling/Sevin Okyay
 10350 Harry Potter und der Gefangene von Askaban (Ha...
                                                                                                J.K. Rowling
       Harry Potter and the Philosopher's Stone (Harr...
                                                                                                J.K. Rowling
 10679 Harry Potter and the Goblet of Fire (Harry Pot...
                                                                                                J.K. Rowling
 PS E:\Pandas>
```

16. Find the number of books available in english.

```
💢 Welcome
                newpanda.py X
                                   books.csv
 newpanda.py > ...
       import pandas as pd
       import numpy as np
       df = pd.read_csv('E:\\Pandas\\books.csv')
       english_books = df[df['language_code'].str.lower() == 'eng']
   4
       print(f"Number of English books: {len(english_books)}")
PROBLEMS
           OUTPUT
                    DEBUG CONSOLE
                                   TERMINAL
                                             PORTS
PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
Number of English books: 8908
PS E:\Pandas>
```

17. Find book having maximum number of reviews.



18. Find authors who have written more than 5 books.

```
★ Welcome

                newpanda.py X
                                   books.csv
 newpanda.py > ...
       import numpy as np
       import pandas as pd
       df=pd.read_csv("E:\\Pandas\\books.csv")
       popular_authors = df['authors'].value_counts()
   4
       print(popular_authors[popular_authors > 5])
 PROBLEMS
           OUTPUT
                    DEBUG CONSOLE
                                   TERMINAL
                                              PORTS
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
 authors
 Stephen King
                        40
 P.G. Wodehouse
                        40
 Rumiko Takahashi
                        39
 Orson Scott Card
                        35
 Agatha Christie
                        33
 Oscar Wilde
                         6
 Junji Ito/伊藤潤二
                             6
 Barbara Kingsolver
                         6
 Catherine Coulter
                         6
 Patricia A. McKillip
                         6
 Name: count, Length: 296, dtype: int64
 PS E:\Pandas>
```

19. Find the most common first letter of all book titles.

```
books.csv

★ Welcome

                 newpanda.py X
 newpanda.py > ...
       import numpy as np
       import pandas as pd
       df=pd.read_csv("E:\\Pandas\\books.csv")
       first_letters = df['title'].str[0]
       print(first_letters.value_counts().head(1))
   6
 PROBLEMS
           OUTPUT
                    DEBUG CONSOLE
                                    TERMINAL
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
 title
      3502
 Name: count, dtype: int64
 PS E:\Pandas>
```

20. Find the book with the longest title.

```
✓ Welcome

                                  ■ books.csv
                newpanda.py X
 newpanda.py > ...
      import numpy as np
   2 import pandas as pd
   3 df=pd.read_csv("E:\\Pandas\\books.csv")
       df['title_length'] = df['title'].apply(len)
       print(df.sort_values(by='title_length', ascending=False).head(1))
   6
 PROBLEMS OUTPUT DEBUG CONSOLE
                                  TERMINAL
                                            PORTS
 PS E:\Pandas> & C:/Users/HP/AppData/Local/Programs/Python/Python312/python.exe e:/Pandas/newpanda.py
                                                       title ... Unnamed: 12 title_length
      bookID
 7455 28639 Yours in Food John Baldessari: with meditatio... ...
 [1 rows x 14 columns]
 PS E:\Pandas>
```