

```
In [1]: import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

import warnings
warnings.filterwarnings('ignore')
```

```
In [2]: print(("*")*120)
print(" \t\t\t\t*****DATASET:- Computer Science / Programing Books*****\n\nHigh rated book")
print(("*")*120)
```

```
*****
*****DATASET:- Computer Science / Programing Books*****

High rated book information in the field of computer science and programing
*****
```

```
In [3]: df=pd.read_csv("prog_book.csv")
df
```

Out [3]:

	Rating	Reviews	Book_title	Description	Number_Of_Pages	Type	Price
0	4.17	3,829	The Elements of Style	This style manual offers practical advice on i...	105	Hardcover	9.323529
1	4.01	1,406	The Information: A History, a Theory, a Flood	James Gleick, the author of the best sellers C...	527	Hardcover	11.000000
2	3.33	0	Responsive Web Design Overview For Beginners	In Responsive Web Design Overview For Beginner...	50	Kindle Edition	11.267647
3	3.97	1,658	Ghost in the Wires: My Adventures as the World...	If they were a hall of fame or shame for compu...	393	Hardcover	12.873529
4	4.06	1,325	How Google Works	Both Eric Schmidt and Jonathan Rosenberg came ...	305	Kindle Edition	13.164706
...
266	3.76	0	3D Game Engine Architecture: Engineering Real...	Dave Eberly's 3D Game Engine Design was the fi...	752	Hardcover	203.108823
267	3.94	22	An Introduction to Database Systems	Continuing in the eighth edition, An Introduct...	1040	Paperback	212.097059
268	4.49	36	The Art of Computer Programming, Volumes 1-3 B...	Knuth's classic work has been widely acclaimed...	896	Boxed Set - Hardcover	220.385294
269	4.77	4	The Art of Computer Programming, Volumes 1-4a ...	"The bible of all fundamental algorithms and t...	3168	Hardcover	220.385294
270	3.84	5	A Discipline for Software Engineering	Designed to help individual programmers develo...	789	Hardcover	235.650000

271 rows × 7 columns

```
In [4]: print(("*")*120)
print(" \t\t\t\t*****UNDERSTANDING OF DATA GIVEN*****\n\n(a) Rating ----- > The user rating")
print(("*")*120)
```

```
*****
*****UNDERSTANDING OF DATA GIVEN*****

(a) Rating ----- > The user rating for the book. the rating score ranges between 0 and 5.
(b) Reviews ----- > The number of reviews found on this book.
(c) Book_title ----- > The name of the book.
(d) Description ----- > A short description of the book.
(e) Number_Of_Pages ----- > Number of pages in the book.
(f) Type ----- >The type of the book meaning is it a hardcover book or an ebook or a kindle book etc.
(g) Price ----- > The average price of the book in USD where the average is calculated according to the 5 web sources.
*****
```

BASIC INFORMATION OF THE DATASET

```
In [5]: print(("*")*120)
print(" \t\t\t\t*****ANALYSIS*****\n\nMY DATASET HAS 7 COLUMNS AND 48895 ROWS.\n\n")
print(("*")*120)
df.info()
```

```
*****
*****ANALYSIS*****
MY DATASET HAS 7 COLUMNS AND 48895 ROWS.
IT HAS 10 NUMERICAL COLUMNS AND 6 CATEGORICAL COLUMNS.
4 COLUMNS HAS NULL VALUES.
*****
<class 'pandas.core.frame.DataFrame'>
```

```

RangeIndex: 271 entries, 0 to 270
Data columns (total 7 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Rating                 271 non-null   float64
1   Reviews                271 non-null   object
2   Book_title             271 non-null   object
3   Description             271 non-null   object
4   Number_Of_Pages        271 non-null   int64
5   Type                   271 non-null   object
6   Price                  271 non-null   float64
dtypes: float64(2), int64(1), object(4)
memory usage: 14.9+ KB

```

FEATURES OF DATASET

```

In [6]: print(("*")*120)
print("\t\t\t\t*****ANALYSIS*****\n~ PRICE:-Price of airbnb ranges between 0$ to 10000$, me
print(("*")*120)
df.describe()

*****
*****ANALYSIS*****
~ PRICE:-Price of airbnb ranges between 0$ to 10000$, mean cost of an airbnb is 152$,but median cost is 106$ this implies there are large number of
~ ALL FEATURES OF DATASET WILL BE ANALYSED IN DETAIL AS WE GO AHEAD IN PROJECT.
*****

```

```

Out [6]:

```

	Rating	Number_Of_Pages	Price
count	271.000000	271.000000	271.000000
mean	4.067417	475.077491	54.541860
std	0.290749	306.147492	35.711485
min	3.000000	50.000000	9.323529
25%	3.915000	289.000000	30.751471
50%	4.100000	384.000000	46.317647
75%	4.250000	572.500000	67.854412
max	5.000000	3168.000000	235.650000

```

In [7]: print(("*")*120)
print("\t\t\t\t*****ANALYSIS*****\n\t\t\t\t~ WE HAVE 47906 UNIQUE RECORDS IN NAME.\n\t\t\t\t~
print(("*")*120)
df.describe(include="O")

*****
*****ANALYSIS*****
~ WE HAVE 47906 UNIQUE RECORDS IN NAME.
~ WE HAVE 11452 UBNIQUE ROWS IN HOST_NAME
~ Neighbourhood_group has 5 unique columns with highest frequency of Manhattan
~ neighbourhood has 221 unique columns
~~ ALL THIS FEATIRES WE WILL ANALYSED IN DETAIL AS WE GO AHEAD IN THIS PROJECT
*****

```

```

Out [7]:

```

	Reviews	Book_title	Description	Type
count	271	271	271	271
unique	133	271	270	6
top	0	The Elements of Style	Learn You a Haskell for Great Good! is a fun, ...	Paperback
freq	22	1	2	156

CLEANING DATA

Finding null/missing values in the data if any.

```

In [8]: print(("*")*120)
print("\t\t\t\t*****ANALYSIS*****\n\t\t\t\t~We can clearly see True in output of some columns
print(("*")*120)
df.isnull().any()

*****
*****ANALYSIS*****
~We can clearly see True in output of some columns which contains null values
*****

```

```

Out [8]: Rating      False
Reviews    False
Book_title  False
Description False
Number_Of_Pages False
Type        False
Price       False
dtype: bool

```

finding % of null values

```
In [9]: print(("**")*120)
print("\t\t\t\t\t*****ANALYSIS*****\n~ reviews_per_month column has approximately 20% null va
print(("**")*120)
df.isnull().sum()/len(df)*100
```

```
*****
*****ANALYSIS*****
~ reviews_per_month column has approximately 20% null values which i will fill with mean value of this columns.
~ last_review column also has 20% null values, for my analysis i donot require this column so i will be dropping whole column
~ name and host_name has less than 0.05% null values, i will be dropping this null rows.
*****
```

```
Out [9]: Rating      0.0
Reviews    0.0
Book_title  0.0
Description 0.0
Number_Of_Pages 0.0
Type        0.0
Price       0.0
dtype: float64
```

```
In [10]: df['Reviews']=df['Reviews'].map(lambda x: x.replace(',',''))
df['Reviews']=pd.to_numeric(df['Reviews'])
```

Some books have no reviewers

```
In [11]: df_corr=df.copy()
df_Reviews=df_corr["Reviews"]
```

```
In [12]: point_0=(df_Reviews==0)
```

```
In [13]: df[point_0]
```

	Rating	Reviews	Book_title	Description	Number_Of_Pages	Type	Price
2	3.33	0	Responsive Web Design Overview For Beginners	In Responsive Web Design Overview For Beginner...	50	Kindle Edition	11.267647
9	4.62	0	ZX Spectrum Games Code Club: Twenty fun games ...	This book is ideal for ZX Spectrum fans and be...	128	Paperback	14.641176
24	3.20	0	Game Programming Golden Rules	'Game Programming Golden Rules' provides indis...	318	Paperback	19.152941
41	4.33	0	Beginners Guide On PHP Programming: Quick And ...	The Book - Beginners Guide On PHP Programming ...	410	Paperback	23.226471
44	4.00	0	Python Programming Books Set: Python Programmi...	Quick & Easy Guide to Python Programming For B...	446	Paperback	23.858824
54	5.00	0	Your First App: Node.js	A tutorial for real-world application developm...	317	ebook	25.855882
64	3.91	0	Python Programming For Beginners: Quick And Ea...	Quick & Easy Guide on Python Programming For B...	190	Paperback	29.020588
148	4.00	0	Building an FPS Game with Unity	Create a high-quality first person shooter gam...	326	Paperback	50.726471
177	3.00	0	Advanced Game Programming: A Gamedev.Net Colle...	Welcome to "Advanced Game Programming: A GameD...	464	Paperback	59.082353
178	3.83	0	Basic Proof Theory	This introduction to the basic ideas of struct...	355	Hardcover	59.526471
182	3.00	0	Cross-Platform Game Programming (Game Developm...	With many of today's games being released simu...	460	Paperback	60.391176
187	3.00	0	Lambda-Calculus, Combinators and Functional Pr...	Originally published in 1988, this book presen...	192	Paperback	61.167647
199	3.92	0	The Calculus of Computation: Decision Procedur...	This textbook introduces computational logic f...	366	Hardcover	66.820588
212	4.13	0	Communication Networks: Fundamental Concepts a...	This book is designed for introductory one-sem...	928	Hardcover	75.700000
226	3.59	0	Genetic Programming II: Automatic Discovery of...	Genetic Programming II extends the results of ...	768	Hardcover	82.367647
238	4.01	0	Game Programming Gems 2	Blazing through the excitement generated by th...	575	Hardcover	91.047059
239	4.03	0	Game Programming Gems 5	With every new volume, the Game Programming Ge...	791	Hardcover	91.047059
246	4.37	0	A First Course in Logic: An Introduction to Mo...	The ability to reason and think in a logical m...	431	Paperback	96.494118
261	3.82	0	Game Programming Gems 6	Welcome to the sixth volume of the Game Progra...	700	Hardcover	119.202941
263	3.77	0	Game Programming Gems 4	Uncover the secrets of the game industry's bes...	703	Hardcover	139.435294
264	3.82	0	Game Programming Gems 3	The journey continues with this ALL NEW volume...	663	Hardcover	143.138235

	Rating	Reviews	Book_title	Description	Number_Of_Pages	Type	Price
266	3.76	0	3D Game Engine Architecture: Engineering Real...	Dave Eberly's 3D Game Engine Design was the fi...	752	Hardcover	203.108823

```
In [14]: print(("*")*120)
print("\tSome books have no reviewers but have rating .We have inconsistent data.I will fill it with the average")
print(("*")*120)
```

```
*****
Some books have no reviewers but have rating .We have inconsistent data.I will fill it with the average value instead of inconsistent data
*****
```

```
In [15]: from math import *
df_Reviews[point_0]=ceil(df.Reviews.mean())
```

```
In [16]: df["Reviews"]=df_Reviews
df[point_0]
```

Out [16]:

	Rating	Reviews	Book_title	Description	Number_Of_Pages	Type	Price
2	3.33	186	Responsive Web Design Overview For Beginners	In Responsive Web Design Overview For Beginner...	50	Kindle Edition	11.267647
9	4.62	186	ZX Spectrum Games Code Club: Twenty fun games ...	This book is ideal for ZX Spectrum fans and be...	128	Paperback	14.641176
24	3.20	186	Game Programming Golden Rules	'Game Programming Golden Rules' provides indis...	318	Paperback	19.152941
41	4.33	186	Beginners Guide On PHP Programming: Quick And ...	The Book - Beginners Guide On PHP Programming ...	410	Paperback	23.226471
44	4.00	186	Python Programming Books Set: Python Programmi...	Quick & Easy Guide to Python Programming For B...	446	Paperback	23.858824
54	5.00	186	Your First App: Node.js	A tutorial for real-world application developm...	317	ebook	25.855882
64	3.91	186	Python Programming For Beginners: Quick And Ea...	Quick & Easy Guide on Python Programming For B...	190	Paperback	29.020588
148	4.00	186	Building an FPS Game with Unity	Create a high-quality first person shooter gam...	326	Paperback	50.726471
177	3.00	186	Advanced Game Programming: A Gamedev.Net Colle...	Welcome to "Advanced Game Programming: A GameD...	464	Paperback	59.082353
178	3.83	186	Basic Proof Theory	This introduction to the basic ideas of struct...	355	Hardcover	59.526471
182	3.00	186	Cross-Platform Game Programming (Game Developm...	With many of today's games being released simu...	460	Paperback	60.391176
187	3.00	186	Lambda-Calculus, Combinators and Functional Pr...	Originally published in 1988, this book presen...	192	Paperback	61.167647
199	3.92	186	The Calculus of Computation: Decision Procedur...	This textbook introduces computational logic f...	366	Hardcover	66.820588
212	4.13	186	Communication Networks: Fundamental Concepts a...	This book is designed for introductory one-sem...	928	Hardcover	75.700000
226	3.59	186	Genetic Programming II: Automatic Discovery of...	Genetic Programming II extends the results of ...	768	Hardcover	82.367647
238	4.01	186	Game Programming Gems 2	Blazing through the excitement generated by th...	575	Hardcover	91.047059
239	4.03	186	Game Programming Gems 5	With every new volume, the Game Programming Ge...	791	Hardcover	91.047059
246	4.37	186	A First Course in Logic: An Introduction to Mo...	The ability to reason and think in a logical m...	431	Paperback	96.494118
261	3.82	186	Game Programming Gems 6	Welcome to the sixth volume of the Game Progra...	700	Hardcover	119.202941
263	3.77	186	Game Programming Gems 4	Uncover the secrets of the game industry's bes...	703	Hardcover	139.435294
264	3.82	186	Game Programming Gems 3	The journey continues with this ALL NEW volume...	663	Hardcover	143.138235
266	3.76	186	3D Game Engine Architecture: Engineering Real...	Dave Eberly's 3D Game Engine Design was the fi...	752	Hardcover	203.108823

```
In [17]: df['Reviews']=df['Reviews'].convert_dtypes()
```

```
In [18]: df['Reviews']=df['Reviews'].astype(str)
```

```
In [19]: df['Reviews'] = df['Reviews'].astype(int)
```

```
In [20]: df['Reviews'] = df['Reviews'].replace(',', '')
df.head()
```

Out [20]:

	Rating	Reviews	Book_title	Description	Number_Of_Pages	Type	Price
0	4.17	3829	The Elements of Style	This style manual offers practical advice on i...	105	Hardcover	9.323529
1	4.01	1406	The Information: A History, a Theory, a Flood	James Gleick, the author of the best sellers C...	527	Hardcover	11.000000
2	3.33	186	Responsive Web Design Overview For Beginners	In Responsive Web Design Overview For Beginner...	50	Kindle Edition	11.267647
3	3.97	1658	Ghost in the Wires: My Adventures as the World...	If they were a hall of fame or shame for compu...	393	Hardcover	12.873529
4	4.06	1325	How Google Works	Both Eric Schmidt and Jonathan Rosenberg came ...	305	Kindle Edition	13.164706

In []:

In [21]:

```
print(("*")*120)
print("\t\t\t\tDropping null rows of columns name and host_name")
print(("*")*120)
df.dropna(inplace=True)
df
```

Dropping null rows of columns name and host_name

Out [21]:

	Rating	Reviews	Book_title	Description	Number_Of_Pages	Type	Price
0	4.17	3829	The Elements of Style	This style manual offers practical advice on i...	105	Hardcover	9.323529
1	4.01	1406	The Information: A History, a Theory, a Flood	James Gleick, the author of the best sellers C...	527	Hardcover	11.000000
2	3.33	186	Responsive Web Design Overview For Beginners	In Responsive Web Design Overview For Beginner...	50	Kindle Edition	11.267647
3	3.97	1658	Ghost in the Wires: My Adventures as the World...	If they were a hall of fame or shame for compu...	393	Hardcover	12.873529
4	4.06	1325	How Google Works	Both Eric Schmidt and Jonathan Rosenberg came ...	305	Kindle Edition	13.164706
...
266	3.76	186	3D Game Engine Architecture: Engineering Real...	Dave Eberly's 3D Game Engine Design was the fi...	752	Hardcover	203.108823
267	3.94	22	An Introduction to Database Systems	Continuing in the eighth edition, An Introduct...	1040	Paperback	212.097059
268	4.49	36	The Art of Computer Programming, Volumes 1-3 B...	Knuth's classic work has been widely acclaimed...	896	Boxed Set - Hardcover	220.385294
269	4.77	4	The Art of Computer Programming, Volumes 1-4a ...	"The bible of all fundamental algorithms and t...	3168	Hardcover	220.385294
270	3.84	5	A Discipline for Software Engineering	Designed to help individual programmers develo...	789	Hardcover	235.650000

271 rows x 7 columns

In [22]:

```
print(("*")*120)
print("\t\t\t\t\tNOW RESETTING INDEX TO GET DEFAULT INDEXES.")
print(("*")*120)
df.reset_index(drop=True,inplace=True)
df
```

NOW RESETTING INDEX TO GET DEFAULT INDEXES.

Out [22]:

	Rating	Reviews	Book_title	Description	Number_Of_Pages	Type	Price
0	4.17	3829	The Elements of Style	This style manual offers practical advice on i...	105	Hardcover	9.323529
1	4.01	1406	The Information: A History, a Theory, a Flood	James Gleick, the author of the best sellers C...	527	Hardcover	11.000000
2	3.33	186	Responsive Web Design Overview For Beginners	In Responsive Web Design Overview For Beginner...	50	Kindle Edition	11.267647
3	3.97	1658	Ghost in the Wires: My Adventures as the World...	If they were a hall of fame or shame for compu...	393	Hardcover	12.873529
4	4.06	1325	How Google Works	Both Eric Schmidt and Jonathan Rosenberg came ...	305	Kindle Edition	13.164706
...
266	3.76	186	3D Game Engine Architecture: Engineering Real...	Dave Eberly's 3D Game Engine Design was the fi...	752	Hardcover	203.108823
267	3.94	22	An Introduction to Database Systems	Continuing in the eighth edition, An Introduct...	1040	Paperback	212.097059
268	4.49	36	The Art of Computer Programming, Volumes 1-3 B...	Knuth's classic work has been widely acclaimed...	896	Boxed Set - Hardcover	220.385294

	Rating	Reviews	Book_title	Description	Number_Of_Pages	Type	Price
269	4.77	4	The Art of Computer Programming, Volumes 1-4a ...	"The bible of all fundamental algorithms and t...	3168	Hardcover	220.385294
270	3.84	5	A Discipline for Software Engineering	Designed to help individual programmers develo...	789	Hardcover	235.650000

271 rows × 7 columns

```
In [23]: print(("*")*120)
print("\t\t\t\tI HAVE CLEANED ALL NULL VALUES AND UNWANTED COLUMNS FROM MY DATA SET.")
print(("*")*120)
df.isnull().sum()/len(df)*100
```

```
*****
I HAVE CLEANED ALL NULL VALUES AND UNWANTED COLUMNS FROM MY DATA SET.
*****
```

```
Out [23]: Rating      0.0
Reviews    0.0
Book_title  0.0
Description 0.0
Number_Of_Pages 0.0
Type        0.0
Price       0.0
dtype: float64
```

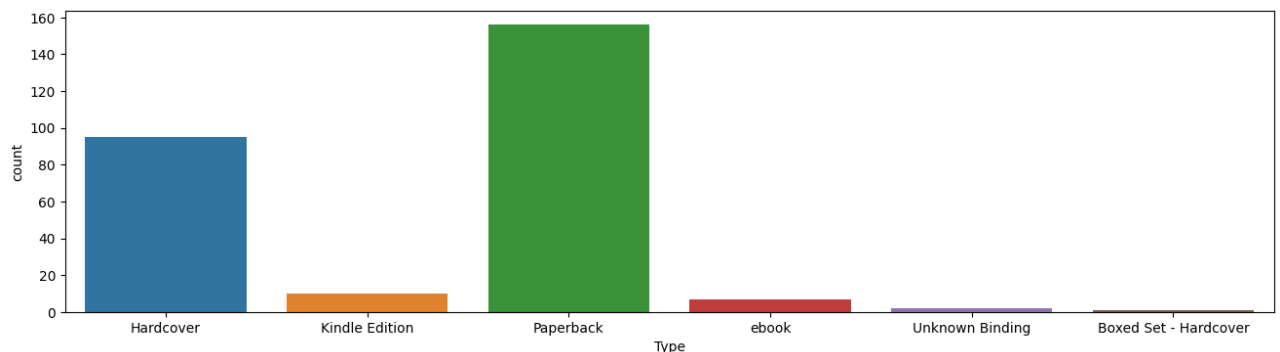
Q1) Counts of different types of Books

```
In [24]: df["Type"].value_counts()
```

```
Out [24]: Type
Paperback      156
Hardcover       95
Kindle Edition  10
ebook           7
Unknown Binding 2
Boxed Set - Hardcover 1
Name: count, dtype: int64
```

```
In [25]: print(("*")*120)
print("\t\t\t\t*****ANALYSIS*****\n\t\t\t(1)Highest types of book is Paperback \n\t\t\t(2)Lowest types of books is Boxed Set - Hardcover")
print(("*")*120)
plt.figure(figsize=(16,4))
ax = sns.countplot(x="Type", data=df);
```

```
*****
*****ANALYSIS*****
(1)Highest types of book is Paperback
(2)Lowest types of books is Boxed Set - Hardcover
*****
```



Q2) Distribution of Book Prices based on Rating

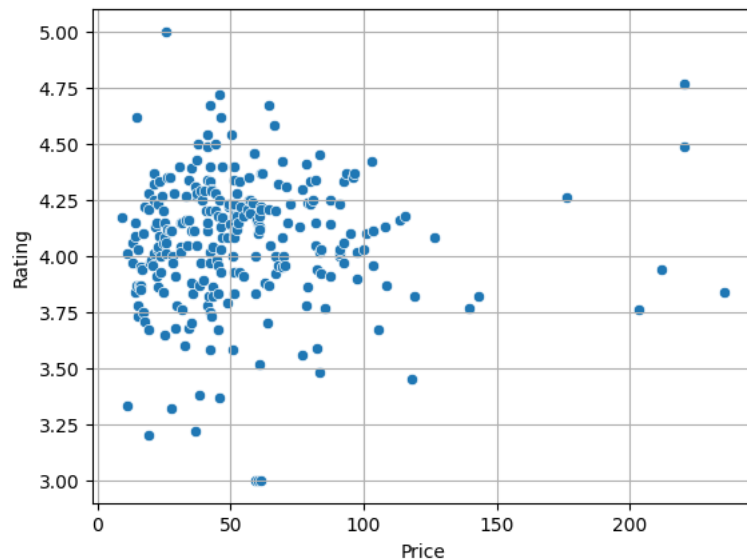
```
In [26]: p1=df.groupby("Price")["Rating"].count().sort_values(ascending=False)
p1
```

```
Out [26]: Price
51.473529      5
60.682353      4
91.047059      3
69.379412      2
59.082353      2
..
37.323529      1
37.432353      1
37.661765      1
38.235294      1
235.650000      1
Name: Rating, Length: 241, dtype: int64
```

```
In [27]: print(("*")*120)
print("\t\t\t\t*****ANALYSIS*****\n\t\t\tMost books are between price range 30 to 100 with av")
print(("*")*120)
sns.scatterplot(data=df,x='Price',y='Rating')
```

```
plt.grid()
plt.show()
```

```
*****
*****ANALYSIS*****
Most books are between price range 30 to 100 with average rating
*****
```

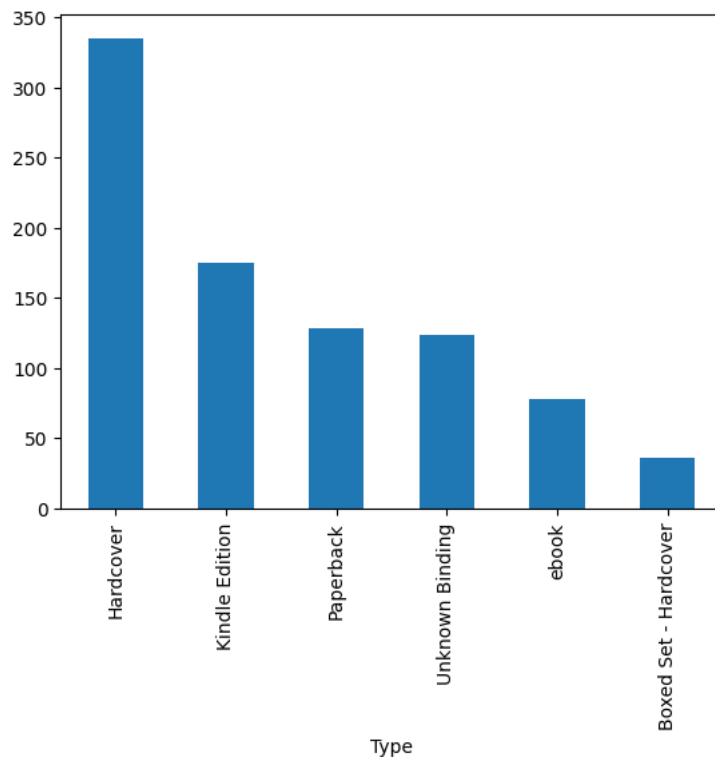


Q3) Reviews according to different types of books.

```
In [28]: p2=df.groupby("Type")["Reviews"].mean().sort_values(ascending=False)
p2
```

```
Out [28]: Type
Hardcover      334.757895
Kindle Edition  175.100000
Paperback      128.179487
Unknown Binding 123.500000
ebook          78.000000
Boxed Set - Hardcover 36.000000
Name: Reviews, dtype: float64
```

```
In [29]: df.groupby("Type")["Reviews"].mean().sort_values(ascending=False).plot(kind="bar")
plt.show()
```



Q4) FIND 5 BOOK TITLE WITH LOWEST RATINGS

```
In [30]: df.groupby("Book_title")["Rating"].value_counts().sort_values(ascending=False).tail(5)
```

```
Out [30]: Book_title
Expert C Programming: Deep C Secrets    4.32    1
```

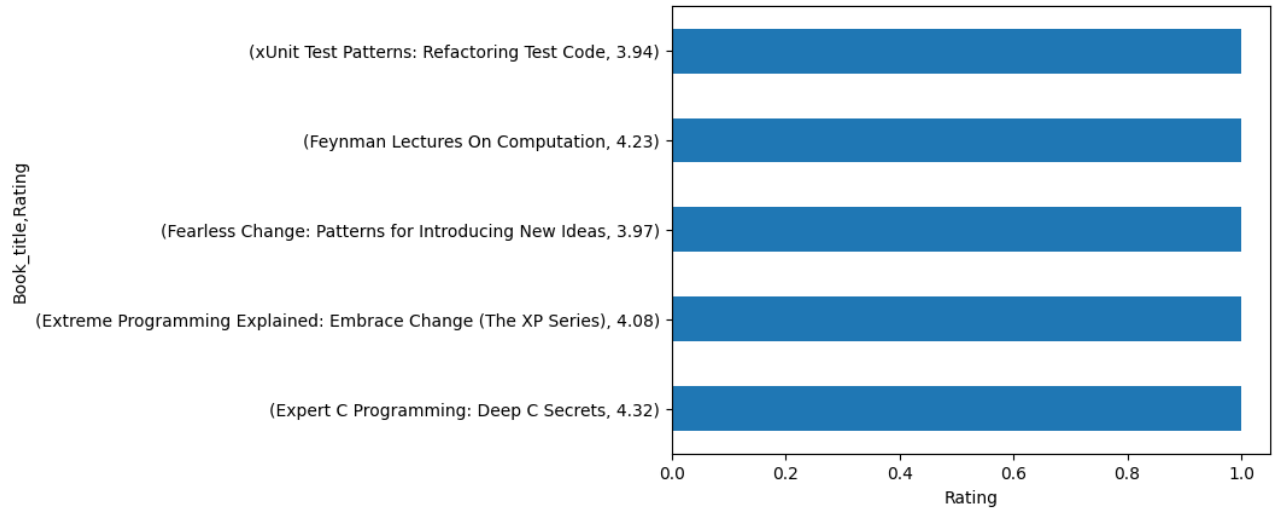
Extreme Programming Explained: Embrace Change (The XP Series)	4.08	1
Fearless Change: Patterns for Introducing New Ideas	3.97	1
Feynman Lectures On Computation	4.23	1
xUnit Test Patterns: Refactoring Test Code	3.94	1

Name: count, dtype: int64

```
In [31]: print(("*")*120)
print("\t\t\t\t\t*****ANALYSIS*****\n\n~~(1) We can see that there are multiple neighbourhoods")
print(("*")*120)
df.groupby("Book_title")["Rating"].value_counts().sort_values(ascending=False).tail(5).plot(kind="barh")
plt.xlabel("Rating")
plt.show()
```

```
*****
*****ANALYSIS*****

~~(1) We can see that there are multiple neighbourhoods with only 1 AIRBNB, tourists are least intersted in this places.
*****
```



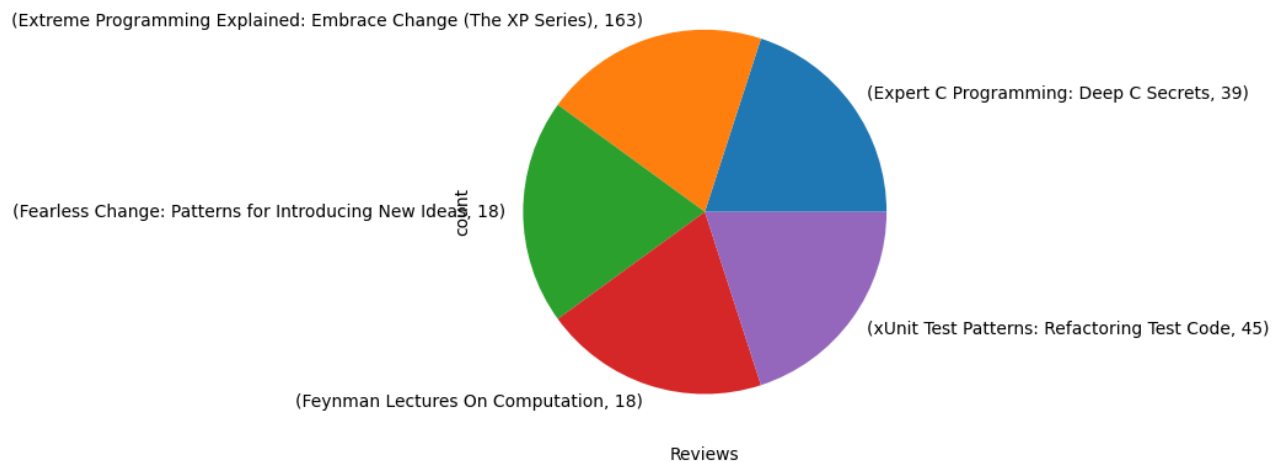
```
In [39]: df.groupby("Book_title")["Reviews"].value_counts().sort_values(ascending=False).tail(5)
```

```
Out [39]: Book_title
Expert C Programming: Deep C Secrets      39      1
Extreme Programming Explained: Embrace Change (The XP Series)  163      1
Fearless Change: Patterns for Introducing New Ideas          18      1
Feynman Lectures On Computation             18      1
xUnit Test Patterns: Refactoring Test Code    45      1
Name: count, dtype: int64
```

```
In [33]: df.groupby("Book_title")["Reviews"].value_counts().sort_values(ascending=False).tail(5).plot(kind="pie")
plt.xlabel("Reviews")
plt.show()
```

```
*****
*****ANALYSIS*****

~~(1) We can see that there are multiple neighbourhoods with only 1 AIRBNB, tourists are least intersted in this places.
*****
```



```
In [40]: # Filter for only hardcover books
hardcover_df = df[df['Type'] == 'Hardcover']

# Sort by price descending and take the top 5 rows
top5 = hardcover_df.sort_values('Price', ascending=False).head(5)

# Create the plot
plt.figure(figsize=(13,9))
ax = plt.subplot(111)
ax.barh(top5['Book_title'], top5['Price'])
```



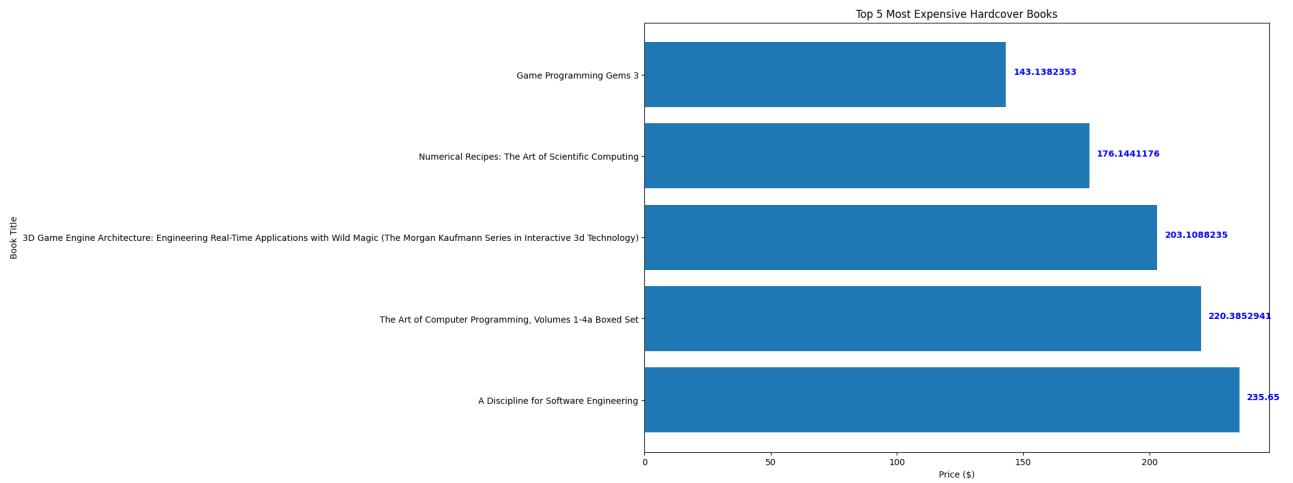
```

ax.set_ylabel('Book Title')
ax.set_xlabel('Price ($)')
plt.title('Top 5 Most Expensive Hardcover Books')

# Add data labels
for i, v in enumerate(top5['Price']):
    ax.text(v + 3, i, str(v), color='blue', fontweight='bold')

plt.show()

```

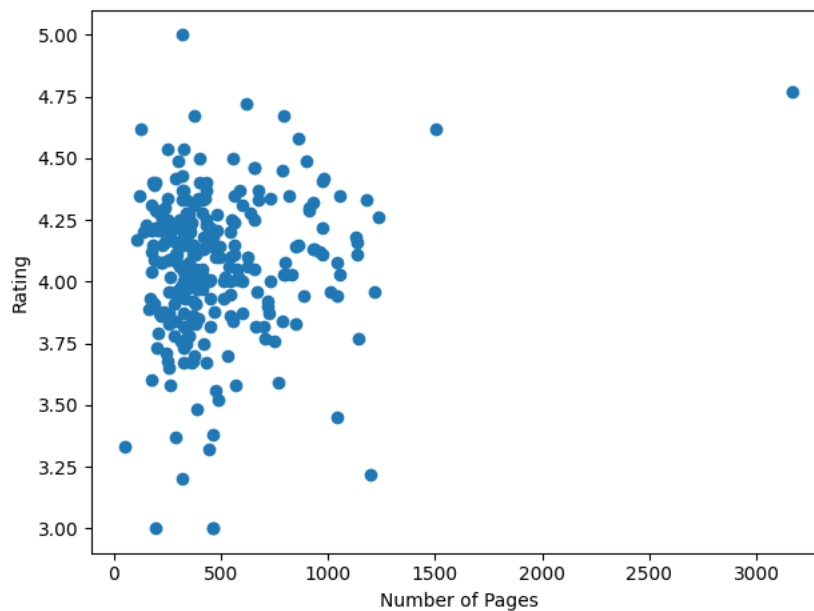


```

In [46]: # Scatter plot of rating vs pages
plt.scatter(df['Number_Of_Pages'], df['Rating'])
plt.xlabel('Number of Pages')
plt.ylabel('Rating')
plt.tight_layout()
plt.show

```

Out [46]: <function matplotlib.pyplot.show(close=None, block=None)>

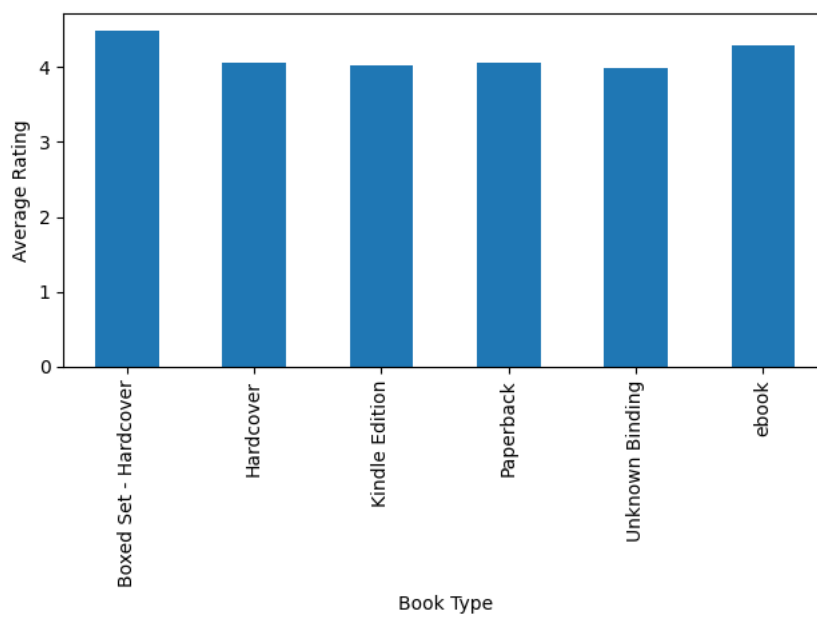


```

In [47]: # Bar chart of average rating by type
type_group = df.groupby('Type')['Rating'].mean()
type_group.plot.bar()

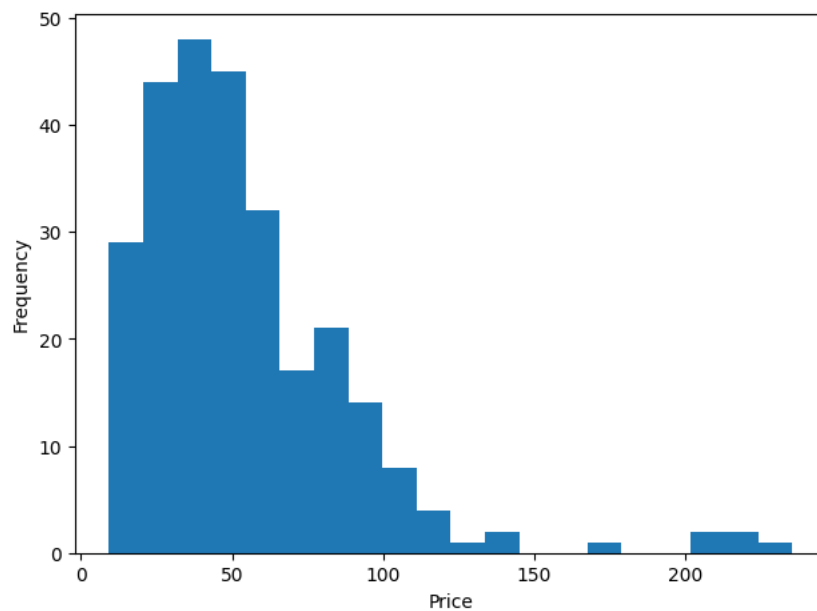
plt.xlabel('Book Type')
plt.ylabel('Average Rating')
plt.tight_layout()
plt.show()

```



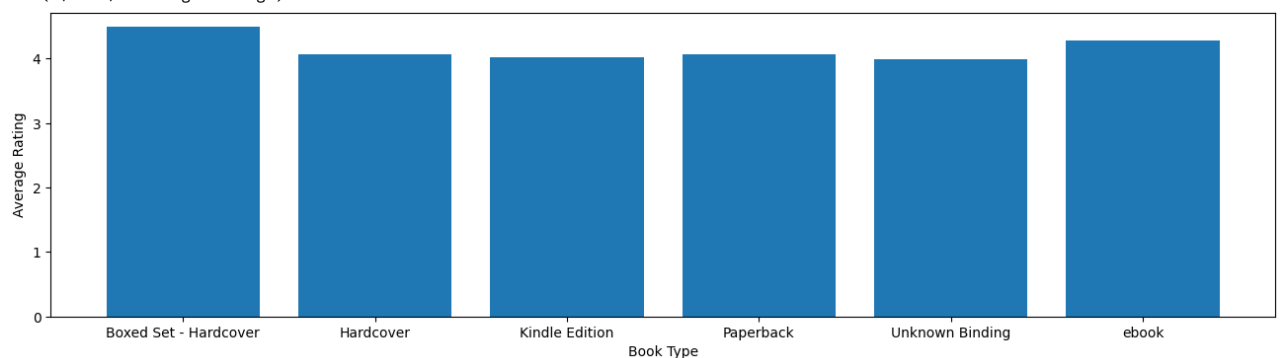
```
In [48]: # Histogram of book prices
plt.hist(df['Price'], bins=20)

plt.xlabel('Price')
plt.ylabel('Frequency')
plt.tight_layout()
plt.show()
```



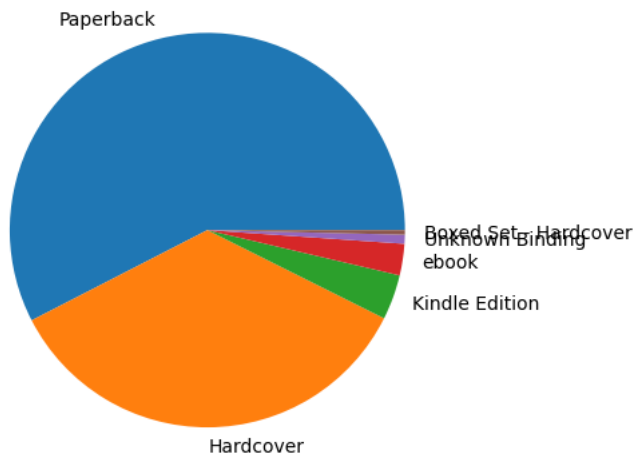
```
In [49]: avg_rating = df.groupby('Type')['Rating'].mean()
plt.figure(figsize=(16,4))
plt.bar(avg_rating.index, avg_rating)
plt.xlabel('Book Type')
plt.ylabel('Average Rating')
```

Out [49]: Text(0, 0.5, 'Average Rating')



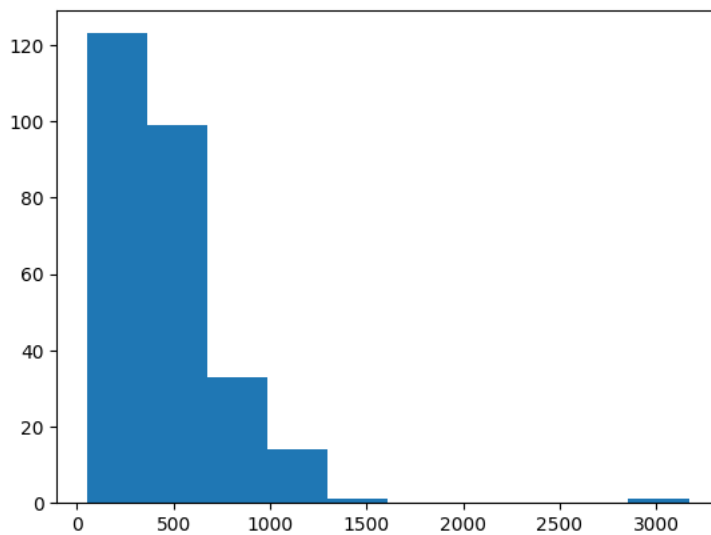
```
In [50]: formats = df['Type'].value_counts()
plt.pie(formats, labels=formats.index)
```

```
Out [50]: ([<matplotlib.patches.Wedge at 0x21808620d90>,
<matplotlib.patches.Wedge at 0x2180864bf50>,
<matplotlib.patches.Wedge at 0x2180865ced0>,
<matplotlib.patches.Wedge at 0x2180865df50>,
<matplotlib.patches.Wedge at 0x2180865f050>,
<matplotlib.patches.Wedge at 0x21808654abd0>],
[Text(-0.258959246721408, 1.0690837705893217, 'Paperback'),
Text(0.006375834103521128, -1.0999815219991127, 'Hardcover'),
Text(1.0341456116224346, -0.3748904559495493, 'Kindle Edition'),
Text(1.087532233188004, -0.1651473335452717, 'ebook'),
Text(1.0988175935413387, -0.05098917653798077, 'Unknown Binding'),
Text(1.0999260867891085, -0.0127516116941638, 'Boxed Set - Hardcover')])
```



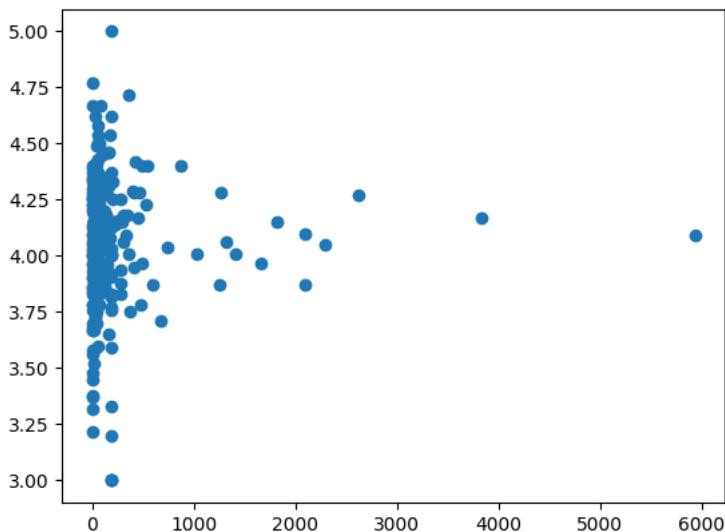
```
In [52]: plt.hist(df['Number_Of_Pages'])
```

```
Out [52]: (array([123., 99., 33., 14., 1., 0., 0., 0., 0., 1.]),
array([ 50., 361.8, 673.6, 985.4, 1297.2, 1609., 1920.8, 2232.6,
2544.4, 2856.2, 3168. ]),
<BarContainer object of 10 artists>)
```



```
In [53]: plt.scatter(x=df['Reviews'], y=df['Rating'])
```

```
Out [53]: <matplotlib.collections.PathCollection at 0x218086d7e90>
```



```
In [ ]:
```