Amazon product dataset

Name : Shrutika gaikwad
Batch : ET2
Roll no : ET2-17
Prn : 2024010701120

shrutika eds > prob1 copy 8.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#9. Find the number of air conditioners under the "Portable" sub-category.

portable_ac_count = (df['sub_category'] == 'Portable').sum()
print('9. Number of Portable air conditioners:', portable_ac_count)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika eds
prob1 copy 8.py"
9. Number of Portable air conditioners: 0
```

shrutika eds > prob1 copy 7.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#7. Find the air conditioner with the highest rating.
df['ratings'] = pd.to_numeric(df['ratings'], errors='coerce')
highest_rated_ac = df.loc[df['ratings'].idxmax()]
print('7. Highest rated air conditioner:', highest_rated_ac['name'])

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika eds/
prob1 copy 7.py"
7. Highest rated air conditioner: Hitachi Split Ac - 1.5 Ton Kiyora 5200Fx I Fresh Inverter - R32 - RSRG518FFEO (
Gold)
```

shrutika eds > prob1 copy 20.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#8. Find the air conditioner with the lowest rating.
df['ratings'] = pd.to_numeric(df['ratings'], errors='coerce')
lowest_rated_ac = df.loc[df['ratings'].idxmin()]
print('8. Lowest rated air conditioner:', lowest_rated_ac['name'])
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

Python: prob1 copy 20

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika eds/
y"
8. Lowest rated air conditioner: CARRIER SPLIT AC 1 TON 3 STAR FIX SPEED (12K 3 STAR DURAFRESH NEO,COPPER)
```

shrutika eds > prob1 copy 13.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#14. Find the total number of air conditioners under the "Window" sub-category.
window_ac_count = (df['sub_category'] == 'Window').sum()
print('14. Number of Window air conditioners:', window_ac_count)

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika eds
prob1 copy 13.py"
14. Number of Window air conditioners: 0
```

shrutika eds > prob1 copy 12.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#13. Find the most common main category for air conditioners.
most_common_category = df['main_category'].mode()[0]
print('13. Most common main category:', most_common_category)

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutik
prob1 copy 12.py"
13. Most common main category: appliances
```

shrutika eds > prob1 copy 19.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#20. Find the average rating for each main category.
df['ratings'] = pd.to_numeric(df['ratings'], errors='coerce')
average_rating_by_main_category = df.groupby('main_category')['ratings'].mean()
print('20. Average rating by main category:\n', average_rating_by_main_category)

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Python: prob1 copy 19

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika eds/prob1 copy 19
y"
20. Average rating by main category:
 main_category
appliances    3.809693
Name: ratings, dtype: float64
```

shrutika eds > prob1 copy 18.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#19. Find the top 3 air conditioners with the highest discount price.
df['discount_price'] = pd.to_numeric(df['discount_price'], errors='coerce')
top_3_discount = df.nlargest(3, 'discount_price')
print('19. Top 3 air conditioners with highest discount price:\n', top_3_discount[['name', 'discount_price']]
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Python: prob1 copy 18

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika eds/py"
19. Top 3 air conditioners with highest discount price:
                                                name  discount_price
0  Lloyd 1.5 Ton 3 Star Inverter Split Ac (5 In 1...             NaN
1  LG 1.5 Ton 5 Star AI DUAL Inverter Split AC (C...             NaN
2  LG 1 Ton 4 Star Ai Dual Inverter Split Ac (Cop...             NaN
```

shrutika eds > prob1 copy 16.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")
#17. Find the air conditioner with the maximum discount (actual price - discount price).
df['actual_price'] = pd.to_numeric(df['actual_price'], errors='coerce')
df['discount_price'] = pd.to_numeric(df['discount_price'], errors='coerce')
df['discount_amount'] = df['actual_price'] - df['discount_price']
print(df[['actual_price', 'discount_price', 'discount_amount']].head())

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

Python: prob1 copy 16

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika eds/
y"
   actual_price  discount_price  discount_amount
0           NaN             NaN              NaN
1           NaN             NaN              NaN
2           NaN             NaN              NaN
3           NaN             NaN              NaN
4           NaN             NaN              NaN
```

shrutika eds > prob1 copy 17.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#18. Find the air conditioners with a rating of 5 stars.
five_star_ac = df[df['ratings'] == 5]
print('18. Air conditioners with 5-star rating:\n', five_star_ac[['name', 'ratings']])

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika eds
prob1 copy 17.py"
18. Air conditioners with 5-star rating:
 Empty DataFrame
Columns: [name, ratings]
Index: []
```

shrutika eds > prob1 copy 11.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#12. Find the top 5 air conditioners with the highest number of ratings.
df['no_of_ratings'] = pd.to_numeric(df['no_of_ratings'], errors='coerce')
top_5_rated = df.nlargest(5, 'no_of_ratings')
print('12. Top 5 air conditioners with highest ratings:\n', top_5_rated[['name', 'no_of_ratings']])

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Python: prob1 copy 11

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika
y"
12. Top 5 air conditioners with highest ratings:
                                                  name  no_of_ratings
101  LG 1.5 Ton 3 Star DUAL Inverter Split AC (Copp...          926.0
32   Whirlpool 1.0 Ton 3 Star, Flexicool Inverter S...          925.0
257  Symphony Diet 22i 22 Litre Air Cooler (White) ...          913.0
165  LG 1.5 Ton 3 Star Inverter Window AC (Copper, ...          886.0
9    Voltas 1.5 Ton, 5 Star, Inverter Split AC(Copp...          801.0
```

shrutika eds > prob1 copy 14.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")
#15. Calculate the proportion of air conditioners that have a discount.
ac_with_discount = (df['discount_price'] < df['actual_price']).sum()
total_ac = df.shape[0]
discount_proportion = ac_with_discount / total_ac
print('15. Proportion of air conditioners with discount:', discount_proportion)

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika
prob1 copy 14.py"
15. Proportion of air conditioners with discount: 0.6138888888888889
```

shrutika eds > prob1 copy 15.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#16. Find the average rating for each sub-category.
df['ratings'] = pd.to_numeric(df['ratings'], errors='coerce')
average_rating_by_sub_category = df.groupby('sub_category')['ratings'].mean()
print('16. Average rating by sub-category:\n', average_rating_by_sub_category)

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Python: prob1 copy 15

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika e
y"
16. Average rating by sub-category:
 sub_category
Air Conditioners    3.809693
Name: ratings, dtype: float64
```

shrutika eds > prob1 copy 20.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#8. Find the air conditioner with the lowest rating.
df['ratings'] = pd.to_numeric(df['ratings'], errors='coerce')
lowest_rated_ac = df.loc[df['ratings'].idxmin()]
print('8. Lowest rated air conditioner:', lowest_rated_ac['name'])
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika ed
prob1 copy 20.py"
8. Lowest rated air conditioner: CARRIER SPLIT AC 1 TON 3 STAR FIX SPEED (12K 3 STAR DURAFRESH NEO,COPPER)
```

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#3. Find the number of air conditioners with ratings above 4.
df['ratings'] = pd.to_numeric(df['ratings'], errors='coerce')
high_rated_ac = (df['ratings'] > 4).sum()
print('3. Number of air conditioners with ratings above 4:', high_rated_ac)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika
s/prob1 copy 3.py"
3. Number of air conditioners with ratings above 4: 170
```

shrutika eds > prob1 copy 5.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#5. Find the average price of all air conditioners (actual price).
df['actual_price'] = pd.to_numeric(df['actual_price'], errors='coerce')
average_price = df['actual_price'].mean()
print('5. Average actual price:', average_price)

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Python: prob1 copy 5

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika
py"
5. Average actual price: nan
```

shrutika eds > prob1 copy 4.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")
#4. Find the number of air conditioners with ratings below 2.
df['ratings'] = pd.to_numeric(df['ratings'], errors='coerce')
low_rated_ac = (df['ratings'] < 2).sum()
print('4. Number of air conditioners with ratings below 2:', low_rated_ac)
low_rated_ac = (df['ratings'] < 2).sum()
print('4. Number of air conditioners with ratings below 2:', low_rated_ac)

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Python: prob1 copy 4

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika e
py"
4. Number of air conditioners with ratings below 2: 15
4. Number of air conditioners with ratings below 2: 15
```

shrutika eds > prob1 copy 6.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#6. Find the highest and lowest priced air conditioners.
df['actual_price'] = pd.to_numeric(df['actual_price'], errors='coerce')
max_price = df['actual_price'].max()
min_price = df['actual_price'].min()
print('6. Highest price:', max_price)
print('Lowest price:', min_price)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika
s/prob1 copy 6.py"
6. Highest price: nan
Lowest price: nan
```

shrutika eds > prob1 copy 9.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#Find the average discount price of air conditioners.
df['discount_price'] = pd.to_numeric(df['discount_price'], errors='coerce')
average_discount_price = df['discount_price'].mean()
print('10. Average discount price:', average_discount_price)

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Python: prob1 copy 9

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika e
"
10. Average discount price: nan
```

shrutika eds > prob1 copy 2.py > ...

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#2. Find the average rating of all air conditioners.
df['ratings'] = pd.to_numeric(df['ratings'], errors='coerce')
average_rating = df['ratings'].mean()
print('2. Average rating:', average_rating)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Python

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutik
py"
2. Average rating: 3.809692671394799
```

```python
import pandas as pd
import numpy as np

# Load dataset
df = pd.read_csv(r"C:\Users\yrugw\OneDrive\Desktop\shrutika eds\archive (2)\Air Conditioners.csv")

#11. Find the number of air conditioners with a discount price below $200.
df['discount_price'] = pd.to_numeric(df['discount_price'], errors='coerce')
discount_below_200 = (df['discount_price'] < 200).sum()
print('11. Air conditioners with discount below $200:', discount_below_200)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Python: prob1 copy 10

```
yrugw@Deadlux MINGW64 ~/OneDrive/Desktop/shrutika eds
$ C:/Users/yrugw/AppData/Local/Programs/Python/Python39/python.exe "c:/Users/yrugw/OneDrive/Desktop/shrutika
y"
11. Air conditioners with discount below $200: 0
```