

In [1]: !pip3 install seaborn
Requirement already satisfied: seaborn in c:\users\aboya\anaconda3\lib\site-packages (0.11.2)
Requirement already satisfied: scipy>=1.0 in c:\users\aboya\anaconda3\lib\site-packages (from seaborn) (1.7.3)
Requirement already satisfied: pandas>=0.23 in c:\users\aboya\anaconda3\lib\site-packages (from seaborn) (1.4.2)
Requirement already satisfied: matplotlib>=2.2 in c:\users\aboya\anaconda3\lib\site-packages (from seaborn) (3.5.1)
Requirement already satisfied: numpy>=1.15 in c:\users\aboya\anaconda3\lib\site-packages (from seaborn) (1.21.5)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\aboya\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (1.3.2)
Requirement already satisfied: packaging>=20.0 in c:\users\aboya\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (21.3)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\aboya\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (4.25.0)
Requirement already satisfied: pillow>=6.2.0 in c:\users\aboya\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (9.0.1)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\aboya\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (2.8.2)
Requirement already satisfied: pyparsing>=2.2.1 in c:\users\aboya\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (3.0.4)
Requirement already satisfied: cycler>=0.10 in c:\users\aboya\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (0.11.0)
Requirement already satisfied: pytz>=2020.1 in c:\users\aboya\anaconda3\lib\site-packages (from pandas>=0.23->seaborn) (2021.3)
Requirement already satisfied: six>=1.5 in c:\users\aboya\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib>=2.2->seaborn) (1.16.0)

In [2]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
#import seaborn.objects as so

In [3]: # Load Dataset
df = pd.read_csv("fashion dataset.csv")
df

Out[3]:

| | p_id | name | price | colour | brand | ratingCount | avg_rating | | description | p_attributes |
|-------|------------|---|--------|-----------|-------------------|-------------|------------|--|----------------------------|---|
| 0 | 1518329.0 | Dupatta Bazaar White Embroidered Chiffon Dupatta | 899.0 | White | Dupatta Bazaar | 1321.0 | 4.548827 | embroidered | White Chiffon Hand- ... | {'Occasion': 'Daily', 'Pattern': 'Embroidered'... |
| 1 | 5829334.0 | Roadster Women Mustard Yellow Solid Hooded Swe... | 1199.0 | Mustard | Roadster | 5462.0 | 4.313255 | Mustard yellow solid sweatshirt, has a hood, t... | | {'Body Shape ID': '443,424,324', 'Body or Garm... |
| 2 | 10340119.0 | Inddus Peach-Coloured & Beige Unstitched Dress... | 5799.0 | Peach | Inddus | 145.0 | 4.068966 | Peach-Coloured and beige woven design unstitch... | | {'Bottom Fabric': 'Cotton Blend', 'Bottom Patt... |
| 3 | 10856380.0 | SASSAFRAS Women Black Parallel Trousers | 1499.0 | Black | SASSAFRAS | 9124.0 | 4.147523 | Black solid woven high-rise parallel trousers,... | | {'Add-Ons': 'NA', 'Body Shape ID': '424', 'Bod... |
| 4 | 12384822.0 | Kotty Women Black Wide Leg High-Rise Clean Loo... | 1999.0 | Black | Kotty | 12260.0 | 4.078467 | Black dark wash 4-pocket high-rise jeans, clea... | | {'Add-Ons': 'NA', 'Brand Fit Name': 'NA', 'Clo... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 14324 | 17029604.0 | The Chennai Silks Pink & Silver-Toned Floral Z... | 3999.0 | Pink | The Chennai Silks | NaN | NaN | Design Details Pink and silver-... | | {'Better Cotton Initiative': 'Regular', 'Blous... |
| 14325 | 17600212.0 | Kinder Kids Girls Blue & Green Printed Foil Pr... | 2050.0 | Blue | Kinder Kids | NaN | NaN | Blue and green printed lehenga choli, has foi... | | {'Blouse Closure': 'NA', 'Blouse Fabric': 'Cot... |
| 14326 | 18159266.0 | KLOTTHE Women Green & Black Floral Printed Pal... | 1659.0 | Green | KLOTTHE | NaN | NaN | Green and black woven palazzos ... | | {'Body or Garment Size': 'To-Fit Denotes Body ... |
| 14327 | 18921114.0 | InWeave Women Red Printed A-Line Skirt | 2399.0 | Red | InWeave | NaN | NaN | <p>Red printed A-line skirt, has drawstring cl... | | {'Add-Ons': 'NA', 'Body Shape ID': '324,333,42... |
| 14328 | 19361058.0 | BoStreet Women Navy Blue Tapered Fit Trousers | 2599.0 | Navy Blue | BoStreet | NaN | NaN | Navy blue knitted trousers ... | | {'Add-Ons': 'NA', 'Body Shape ID': '443,333,42... |

14329 rows × 9 columns

In [4]: df.info()

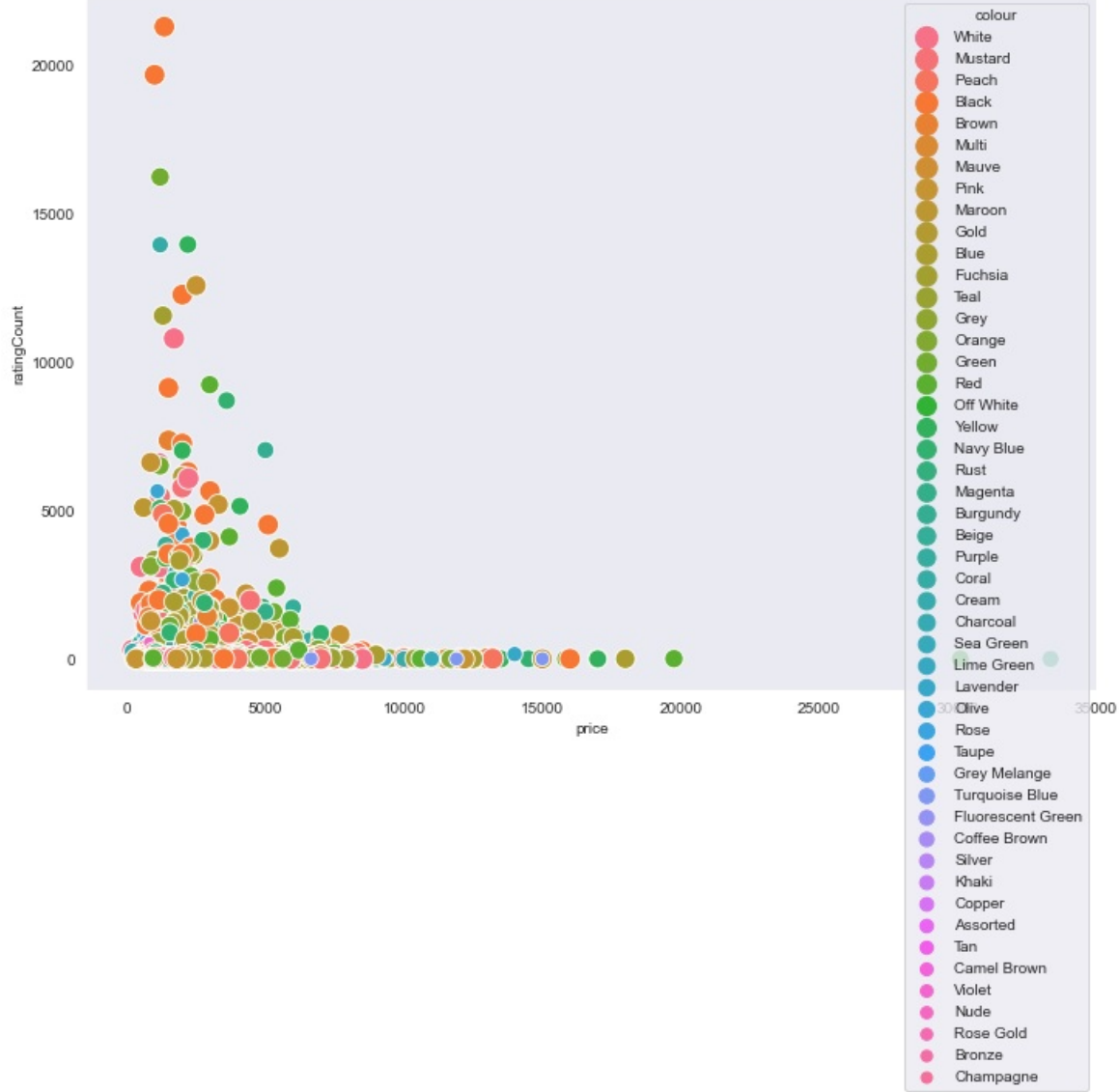
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14329 entries, 0 to 14328
Data columns (total 9 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   p_id        14311 non-null  float64
1   name        14310 non-null  object
2   price       14310 non-null  float64
3   colour      14307 non-null  object
4   brand       14305 non-null  object
5   ratingCount  6581 non-null   float64
6   avg_rating  6581 non-null   float64
7   description 14310 non-null  object
8   p_attributes 14310 non-null  object
dtypes: float64(4), object(5)
memory usage: 1007.6+ KB
In [5]: df.describe()
```

Out[5]:

| | p_id | price | ratingCount | avg_rating |
|-------|--------------|--------------|--------------|-------------|
| count | 1.431100e+04 | 14310.000000 | 6581.000000 | 6581.000000 |
| mean | 1.569129e+07 | 2964.168484 | 184.479410 | 4.101226 |
| std | 3.153525e+06 | 2564.014851 | 782.501137 | 0.475633 |
| min | 7.016600e+04 | 169.000000 | 1.000000 | 1.000000 |
| 25% | 1.413618e+07 | 1599.000000 | 9.000000 | 3.888889 |
| 50% | 1.638217e+07 | 2200.000000 | 23.000000 | 4.180822 |
| 75% | 1.808452e+07 | 3495.000000 | 80.000000 | 4.392857 |
| max | 1.941576e+07 | 47999.000000 | 21274.000000 | 5.000000 |

EDA

```
In [6]: sns.set_style("dark")
fig, ax = plt.subplots()
fig.set_size_inches(11.7, 8.27)
sns.scatterplot(x="price", y="ratingCount", data=df, hue='colour',size="colour",
                sizes=(50, 200), legend="full");
```

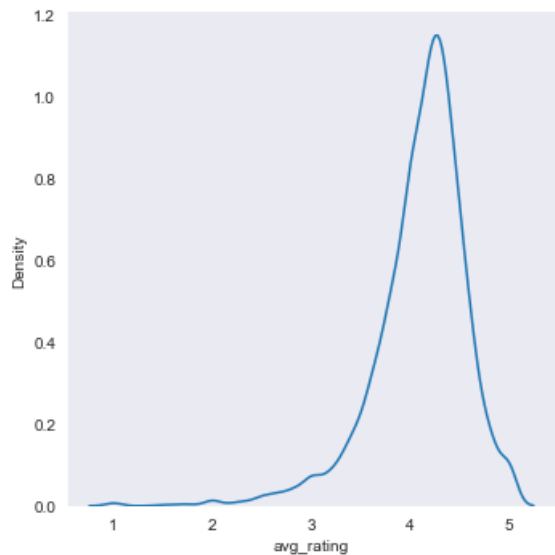


In [7]: x= df.name.unique()

In [8]: sns.set_style("dark")
fig, ax = plt.subplots()
fig.set_size_inches(11.7, 8.27)
sns.scatterplot(x="price", y="ratingCount", data=df, hue='colour',size="ratingCount",
sizes=(50, 200), legend="full");



```
In [9]: sns.displot(df, x="avg_rating", kind="kde");
```



```
In [ ]:
```

```
In [10]: for label, content in df.items():
          if not pd.api.types.is_numeric_dtype(content):
              print(label)
```

```
name
colour
brand
description
p_attributes
```

```
In [11]: for label, content in df.items():
          if not pd.api.types.is_numeric_dtype(content):
              df[label]= content.astype(str)
```

```
In [12]: df.isnull().sum()
```

```

Out[12]:p_id      18
        name       0
        price     19
        colour     0
        brand      0
        ratingCount 7748
        avg_rating 7748
        description 0
        p_attributes 0
        dtype: int64

In [13]: df["price"] = df.price.fillna(df.price.median())
        df["name"] = df.name.fillna(df.name.mode())
        #df["description"] = df.description.fillna(df.description.mode())
        #df["brand"] = df.brand.fillna(df.brand.mode())
        df["p_id"] = df.p_id.fillna(df.p_id.median())
        df["colour"] = df.colour.fillna(df.colour.mode())
        df["ratingCount"] = df.ratingCount.fillna(df.ratingCount.median())
        df["avg_rating"] = df.avg_rating.fillna(df.avg_rating.median())

In [14]: df.isnull().sum()

Out[14]:p_id      0
        name      0
        price     0
        colour     0
        brand      0
        ratingCount 0
        avg_rating 0
        description 0
        p_attributes 0
        dtype: int64

In [15]: df.groupby("price")["colour", "brand"].agg(sum).sort_values(by=["price"], ascending=False)

C:\Users\Aboya\AppData\Local\Temp\ipykernel_19312\197361408.py:1: FutureWarning: Indexing with multiple keys (implicitly converted to a tuple of keys) will be deprecated, use a list instead.
  df.groupby("price")["colour", "brand"].agg(sum).sort_values(by=["price"], ascending=False)
Out[15]:
           colour      brand
price
47999.0      Maroon  MOKSHA DESIGNS
33350.0    Magenta  Readiprint Fashions
30500.0       Blue  Readiprint Fashions
30090.0  Off White           Shaily
30000.0 BrownWhiteBlack  Readiprint FashionsMasabaMasaba
...
298.0      Navy Blue           Soch
295.0       Black           Jockey
249.0       White    Juniors by Lifestyle
199.0  WhiteCream           maxmax
169.0      Peach           max

1209 rows x 2 columns

In [16]: def remove_char(x):
        transformed_col = x.str.replace('\W', '', regex=True)
        return transformed_col

In [17]: df.description = remove_char(df.description)
        df.p_attributes = remove_char(df.p_attributes)

In [18]: # Creating a function which will remove extra leading
        # and trailing whitespace from the data.
        # pass dataframe as a parameter here
        def whitespace_removal(dataframe):

            # iterating over the columns
            for i in dataframe.columns:

                # checking datatype of each columns
                if dataframe[i].dtype == 'object':

                    # applying strip function on column
                    dataframe[i] = dataframe[i].map(str.strip)
                else:

                    # if condn. is False then it will do nothing.
                    pass

```

```
# applying whitespace_remover function on dataframe
whitespace_remover(df)

# printing dataframe
#print(df)

In [19]: df

Out[19]:
```

| | p_id | name | price | colour | brand | ratingCount | avg_rating | description | p_attributes |
|-------|------------|---|--------|-----------|-------------------|-------------|------------|---|--|
| 0 | 1518329.0 | Dupatta Bazaar White Embroidered Chiffon Dupatta | 899.0 | White | Dupatta Bazaar | 1321.0 | 4.548827 | White embroidered nbsp dupattaChiffon br Hand ... | Occasion Daily Pattern Embroidered ... |
| 1 | 5829334.0 | Roadster Women Mustard Yellow Solid Hooded Swe... | 1199.0 | Mustard | Roadster | 5462.0 | 4.313255 | Mustard yellow solid sweatshirt has a hood t... | Body Shape ID 443 424 324 Body or Garmen... |
| 2 | 10340119.0 | Iddus Peach-Coloured & Beige Unstitched Dress... | 5799.0 | Peach | Iddus | 145.0 | 4.068966 | Peach Coloured and beige woven design unstitch... | Bottom Fabric Cotton Blend Bottom Patter... |
| 3 | 10856380.0 | SASSAFRAS Women Black Parallel Trousers | 1499.0 | Black | SASSAFRAS | 9124.0 | 4.147523 | Black solid woven high rise parallel trousers ... | Add Ons NA Body Shape ID 424 Body ... |
| 4 | 12384822.0 | Kotty Women Black Wide Leg High-Rise Clean Loo... | 1999.0 | Black | Kotty | 12260.0 | 4.078467 | Black dark wash 4 pocket high rise jeans clea... | Add Ons NA Brand Fit Name NA Closu... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 14324 | 17029604.0 | The Chennai Silks Pink & Silver-Toned Floral Z... | 3999.0 | Pink | The Chennai Silks | 23.0 | 4.180822 | b Design Details b ul li Pink and silver t... | Better Cotton Initiative Regular Blouse ... |
| 14325 | 17600212.0 | Kinder Kids Girls Blue & Green Printed Foil Pr... | 2050.0 | Blue | Kinder Kids | 23.0 | 4.180822 | Blue and green printed lehenga choli has foi... | Blouse Closure NA Blouse Fabric Cotto... |
| 14326 | 18159266.0 | KLOTTHE Women Green & Black Floral Printed Pal... | 1659.0 | Green | KLOTTHE | 23.0 | 4.180822 | ul li Green and black woven palazzos li ... | Body or Garment Size To Fit Denotes Body Me... |
| 14327 | 18921114.0 | InWeave Women Red Printed A-Line Skirt | 2399.0 | Red | InWeave | 23.0 | 4.180822 | p Red printed A line skirt has drawstring clo... | Add Ons NA Body Shape ID 324 333 424 ... |
| 14328 | 19361058.0 | BoStreet Women Navy Blue Tapered Fit Trousers | 2599.0 | Navy Blue | BoStreet | 23.0 | 4.180822 | ul li Navy blue knitted trousers li li ... | Add Ons NA Body Shape ID 443 333 424 ... |

```
14329 rows x 9 columns

In [20]: from collections import Counter
import matplotlib.pyplot as plt
from wordcloud import WordCloud, STOPWORDS, ImageColorGenerator

In [21]: cc = Counter(" ".join(df.p_attributes).split()).most_common(10)
cc

Out[21]:[('NA', 62167),
('Type', 31153),
('Fabric', 31148),
('Pattern', 30053),
('or', 25743),
('Wash', 24542),
('Regular', 24287),
('Body', 23665),
('Length', 17753),
('Care', 15859)]

Most Used Word in Attributes

In [22]: text_attribute = " ".join(review for review in df.p_attributes)

In [23]: wordcloud = WordCloud(background_color="white",
max_words=50,
max_font_size=300,
width=1024,
height=500,
colormap="Dark2_r"
).generate(text_attribute)

plt.figure(figsize=(10,10))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.margins(x=0, y=0)
plt.show()
```



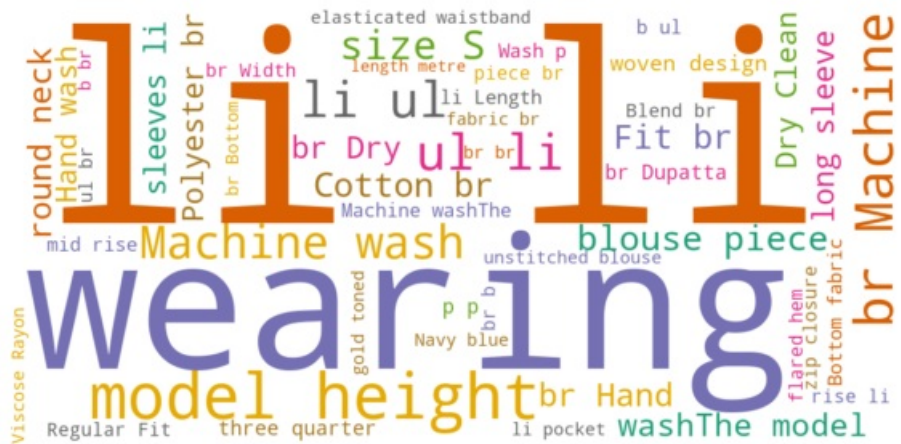

Most Used Word In Description

```
In [24]: text = Counter(" ".join(df.description).split()).most_common(10)

In [25]: text_description = " ".join(review for review in df.description)
```

```
wordcloud = WordCloud(background_color="white",
                        max_words=50,
                        max_font_size=300,
                        width=1024,
                        height=500,
                        colormap="Dark2"
                        ).generate(text_description)

plt.figure(figsize=(10,10))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.margins(x=0, y=0)
plt.show()
```



Most Rated Brand

```
In [28]: brand_rating = pd.DataFrame(df, columns=['brand','ratingCount'])
brand_rating=brand_rating.sort_values(by=['ratingCount'],ascending=False)
brand_rating.reset_index(drop=True,inplace=True)
brand_rating
```

| | brand | ratingCount |
|-------|----------------------|-------------|
| 0 | AHIKA | 21274.0 |
| 1 | SASSAFRAS | 19656.0 |
| 2 | AHIKA | 16219.0 |
| 3 | Varanga | 13947.0 |
| 4 | Roadster | 13938.0 |
| ... | ... | ... |
| 14324 | GRACIT | 1.0 |
| 14325 | Dupatta Bazaar | 1.0 |
| 14326 | Melange by Lifestyle | 1.0 |
| 14327 | Belle Fille | 1.0 |
| 14328 | Baby Lakshmi | 1.0 |

14329 rows × 2 columns

```
# Sum of Rating Count
brand_rating_sum = brand_rating.groupby("brand")["ratingCount"].agg(sum).sort_values( ascending=False)
pd.DataFrame(brand_rating_sum,columns=["ratingCount"])
```

Out[41]:

| | ratingCount |
|---------------|-------------|
| brand | |
| SASSAFRAS | 134826.0 |
| Roadster | 105342.0 |
| Tokyo Talkies | 63049.0 |
| AHIKA | 62876.0 |
| Anouk | 44643.0 |
| ... | ... |
| Seerat | 3.0 |
| THREAD MUSTER | 3.0 |
| DesiNoor.com | 3.0 |
| toothless | 2.0 |
| Mystere Paris | 2.0 |

1021 rows × 1 columns

Most Rated Product

```
In [42]: product_rating = pd.DataFrame(df, columns=['name','ratingCount'])
product_rating=product_rating.sort_values(by=['ratingCount'],ascending=False)
product_rating.reset_index(drop=True, inplace=True)
product_rating
```

Out[42]:

| | name | ratingCount |
|-------|---|-------------|
| 0 | AHIKA Women Black & Green Printed Straight Kurta | 21274.0 |
| 1 | SASSAFRAS Black High Neck Cropped Top | 19656.0 |
| 2 | AHIKA Floral Print Straight Cotton Kurta With ... | 16219.0 |
| 3 | Varanga Mustard Marigold Cotton Straight Kurta | 13947.0 |
| 4 | Roadster Women Coral Pink Solid Hooded Sweatshirt | 13938.0 |
| ... | ... | ... |
| 14324 | GRACIT Women Pack of 2 Red & Navy Blue Printed... | 1.0 |
| 14325 | Dupatta Bazaar Rust & Red Embroidered Dupatta | 1.0 |
| 14326 | Melange by Lifestyle Women Green Ethnic Motifs... | 1.0 |
| 14327 | Belle Fille Women Blue Solid Quilted Jacket | 1.0 |
| 14328 | Baby Lakshmi Girls Orange & Off White Ready to... | 1.0 |

14329 rows × 2 columns

Most Rated By Colour

```
In [43]: color_rating = pd.DataFrame(df, columns=['colour','ratingCount'])
color_rating=color_rating.sort_values(by=['ratingCount'],ascending=False)
color_rating.reset_index(drop=True, inplace=True)
color_rating
```


Out[43]:

| | colour | ratingCount |
|-------|--------|-------------|
| 0 | Black | 21274.0 |
| 1 | Black | 19656.0 |
| 2 | Green | 16219.0 |
| 3 | Yellow | 13947.0 |
| 4 | Coral | 13938.0 |
| ... | ... | ... |
| 14324 | Red | 1.0 |
| 14325 | Rust | 1.0 |
| 14326 | Green | 1.0 |
| 14327 | Blue | 1.0 |
| 14328 | Orange | 1.0 |

14329 rows × 2 columns

Fashion_Style with Rating Between 5000-10000

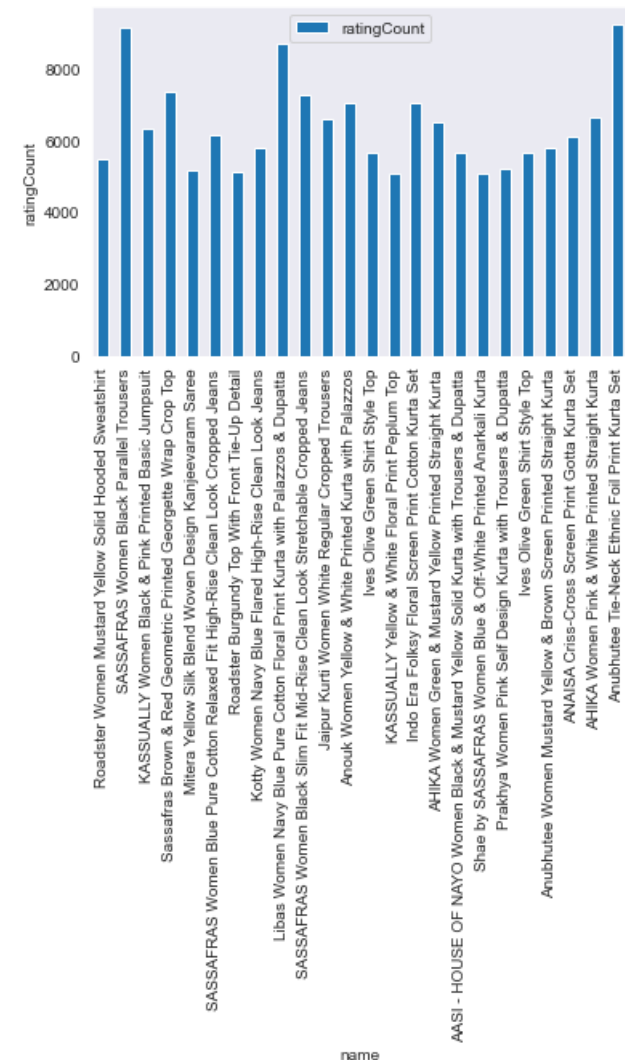
```
In [50]: fashion_rating = pd.DataFrame(df, columns=["name", "ratingCount"])
fashion_rating = fashion_rating[fashion_rating["ratingCount"].isin(range(5000,10000))]
fashion_rating.sort_values(by=["ratingCount"], ascending=False).reset_index(drop=True)
```

Out[50]:

| | name | ratingCount |
|----|---|-------------|
| 0 | Anubhutee Tie-Neck Ethnic Foil Print Kurta Set | 9229.0 |
| 1 | SASSAFRAS Women Black Parallel Trousers | 9124.0 |
| 2 | Libas Women Navy Blue Pure Cotton Floral Print... | 8695.0 |
| 3 | Sassafras Brown & Red Geometric Printed George... | 7358.0 |
| 4 | SASSAFRAS Women Black Slim Fit Mid-Rise Clean ... | 7252.0 |
| 5 | Indo Era Folksy Floral Screen Print Cotton Kur... | 7031.0 |
| 6 | Anouk Women Yellow & White Printed Kurta with ... | 7012.0 |
| 7 | AHIKA Women Pink & White Printed Straight Kurta | 6616.0 |
| 8 | Jaipur Kurti Women White Regular Cropped Trousers | 6590.0 |
| 9 | AHIKA Women Green & Mustard Yellow Printed Str... | 6504.0 |
| 10 | KASSUALLY Women Black & Pink Printed Basic Jum... | 6297.0 |
| 11 | SASSAFRAS Women Blue Pure Cotton Relaxed Fit H... | 6150.0 |
| 12 | ANAISA Criss-Cross Screen Print Gotta Kurta Set | 6076.0 |
| 13 | Anubhutee Women Mustard Yellow & Brown Screen ... | 5772.0 |
| 14 | Kotty Women Navy Blue Flared High-Rise Clean L... | 5771.0 |
| 15 | AASI - HOUSE OF NAYO Women Black & Mustard Yel... | 5652.0 |
| 16 | Ives Olive Green Shirt Style Top | 5648.0 |
| 17 | Ives Olive Green Shirt Style Top | 5648.0 |
| 18 | Roadster Women Mustard Yellow Solid Hooded Swe... | 5462.0 |
| 19 | Prakhya Women Pink Self Design Kurta with Trou... | 5199.0 |
| 20 | Mitera Yellow Silk Blend Woven Design Kanjeeva... | 5144.0 |
| 21 | Roadster Burgundy Top With Front Tie-Up Detail | 5097.0 |
| 22 | KASSUALLY Yellow & White Floral Print Peplum Top | 5069.0 |
| 23 | Shae by SASSAFRAS Women Blue & Off-White Print... | 5046.0 |

```
In [53]: fashion_rating.plot(x="name", y="ratingCount", kind="bar")
plt.xlabel("name")
plt.ylabel("ratingCount")
#plt.rcParams['figure.figsize']=[15,15];
```

Out[53]:Text(0, 0.5, 'ratingCount')



Fashion With Rating Above 20000

```
In [57]: above_20 = pd.DataFrame(df, columns=['name','ratingCount'])
         above_20=above_20.sort_values(by=['ratingCount'],ascending=False)
         above_20.reset_index(drop=True, inplace=True)
         above_20 = above_20[above_20["ratingCount"]>20000]
         above_20
```

```
Out[57]:
```

| | name | ratingCount |
|---|--|-------------|
| 0 | AHIKA Women Black & Green Printed Straight Kurta | 21274.0 |

Fashion Avg Rating Above 2

```
In [58]: avg_above_20 = pd.DataFrame(df, columns=['name','avg_rating'])
         avg_above_20=avg_above_20.sort_values(by=['avg_rating'],ascending=False)
         avg_above_20.reset_index(drop=True, inplace=True)
         avg_above_20 = avg_above_20[avg_above_20["avg_rating"]>2]
         avg_above_20
```

Out[58]:

| | | name | avg_rating |
|--|-------|---|------------|
| | 0 | MISH Blue Culottes & Fitted Crop Shirt Co-Ord Set | 5.000000 |
| | 1 | Amydus Women Plus Size Pink Printed Bomber | 5.000000 |
| | 2 | Silk Land White & Blue Zari Pure Silk Kanjeeva... | 5.000000 |
| | 3 | Readiprint Fashions Blue Embroidered Unstitch... | 5.000000 |
| | 4 | The Vanca Women Olive Green Hooded Sweatshirt | 5.000000 |
| | ... | ... | ... |
| | 14288 | SAADHVI Orange Printed Semi-Stitched Lehenga w... | 2.200000 |
| | 14289 | Zink London Women Blue Solid Tailored Jacket | 2.200000 |
| | 14290 | DeFacto Women Black Stretchable Jeans | 2.181818 |
| | 14291 | Inddus Women Burgundy Printed Co-Ords With Shrug | 2.166667 |
| | 14292 | Martini Women Black Self Design Basic Jumpsuit | 2.166667 |

14293 rows × 2 columns

Fashion Name With Black Colour

```
In [60]: black_name = pd.DataFrame(df, columns=['name','colour'])
black_name = black_name[black_name["colour"]=="Black"]
black_name
```

Out[60]:

| | | name | colour |
|--|-------|---|--------|
| | 3 | SASSAFRAS Women Black Parallel Trousers | Black |
| | 4 | Kotty Women Black Wide Leg High-Rise Clean Loo... | Black |
| | 5 | KASSUALLY Women Black & Pink Printed Basic Jum... | Black |
| | 8 | Tokyo Talkies Women Black Solid Regular Shorts | Black |
| | 9 | Anouk Stylish Black Solid Ready to Wear Leheng... | Black |
| | ... | ... | ... |
| | 14290 | ether Women Black Boyfriend Fit High-Rise Clea... | Black |
| | 14297 | Chkokko Women Black & Red Gym Sporty Zipper Ja... | Black |
| | 14309 | STREET 9 Black Puff Sleeves Blouson Crop Pleat... | Black |
| | 14315 | tantkatha Black Front Closure Saree Blouse Wit... | Black |
| | 14320 | Indo Era Women Black Floral Printed Pure Cotto... | Black |

1917 rows × 2 columns

Fashion Name With Black and White Colour

```
In [62]: black_w_name = pd.DataFrame(df, columns=['name','colour'])
black_w_name=black_w_name.sort_values(by=['colour'],ascending=False)
black_w_name.reset_index(drop=True, inplace=True)
black_w_name = black_w_name[black_w_name["colour"].isin(["Black","White"])]
black_w_name.reset_index(drop=True)
```

Out[62]:

| | | name | colour |
|--|------|---|--------|
| | 0 | KALINI Women White Ethnic Motifs Printed Empir... | White |
| | 1 | INDYA Women White & Green Printed Open Front S... | White |
| | 2 | Vero Moda Women White High-Rise Trousers | White |
| | 3 | Albion Girls White Solid Flared Knee Length Sk... | White |
| | 4 | URBANIC Women White & Blue Printed Crop Top Wi... | White |
| | ... | ... | ... |
| | 2723 | Sangria Women Black Pure Cotton Kurta with Tro... | Black |
| | 2724 | Puma Women Black Lightweight Padded Jacket | Black |
| | 2725 | URBANIC Black Ribbed Slim Fit Basic Jumpsuit | Black |
| | 2726 | Yuris Women Black & White Pure Cotton Printed ... | Black |
| | 2727 | STREET 9 Women Black Checked Lightweight Tailo... | Black |

2728 rows × 2 columns

```
In [103]: sns.distplot(df, x= "avg_rating", y="ratingCount", kind="kde");
```

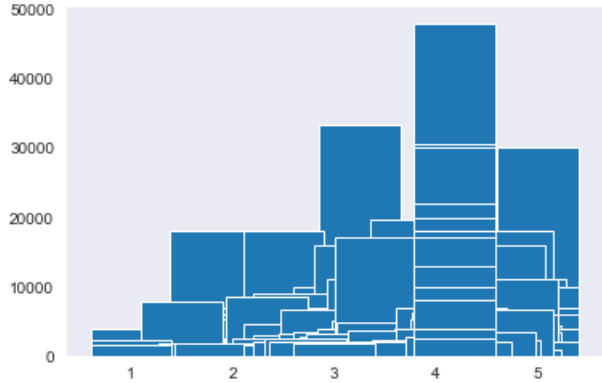
TypeError

Traceback (most recent call last)

Input In [103], in <cell line: 1>()
----> 1 sns distplot(df, x= "avg_rating", y="ratingCount", kind="kde")

TypeError: distplot() got an unexpected keyword argument 'y'

In [79]: y= df["price"]
x= df["avg_rating"]
plt.bar(x,y);



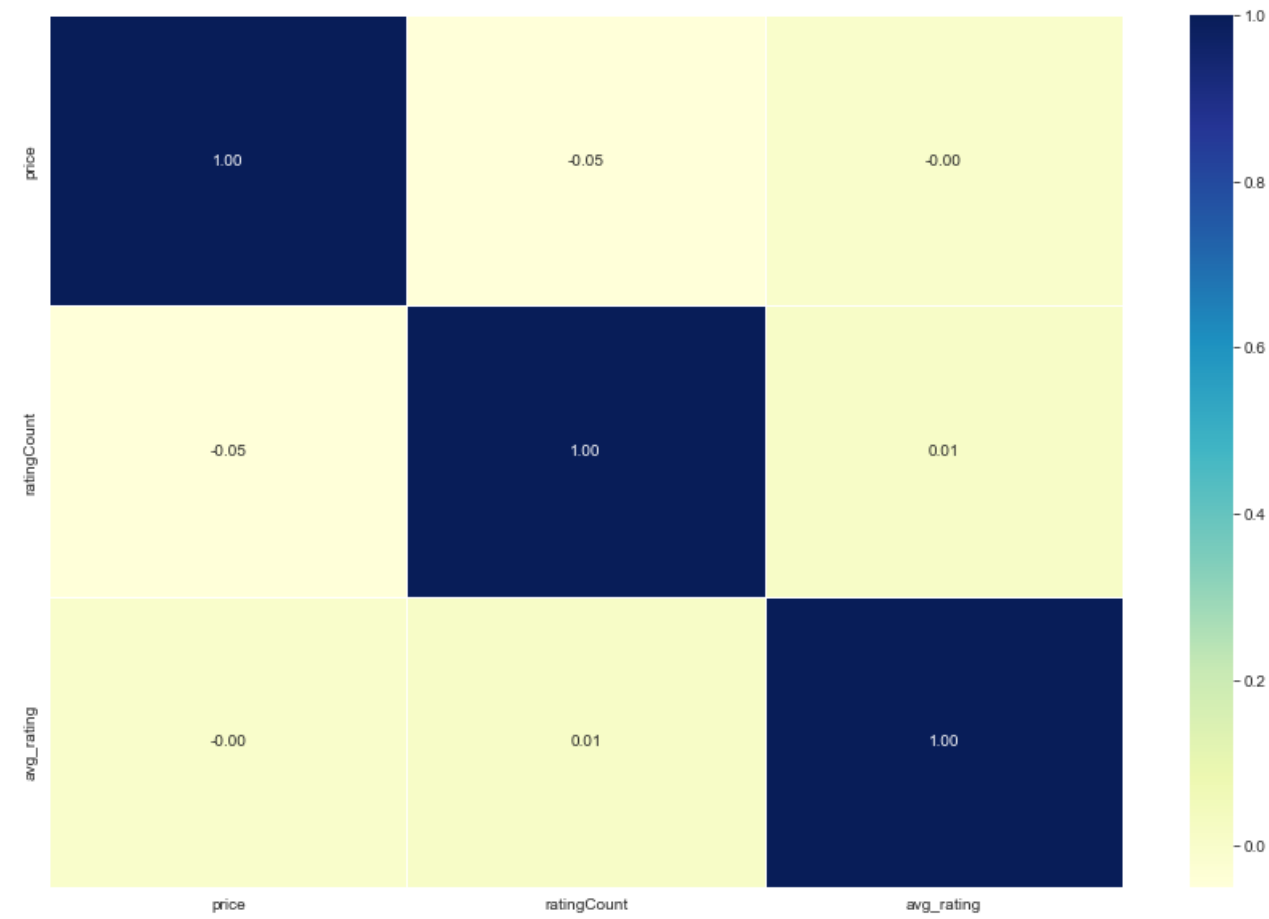
In [101]: price_rating_avg = pd.DataFrame(df, columns=["price", "ratingCount","avg_rating"])
price_rating_avg.sort_values(by=["ratingCount"], ascending=False).reset_index(drop=True)

Out[101]:

| | price | ratingCount | avg_rating |
|-------|--------|-------------|------------|
| 0 | 1350.0 | 21274.0 | 3.978377 |
| 1 | 999.0 | 19656.0 | 4.488146 |
| 2 | 1198.0 | 16219.0 | 4.093286 |
| 3 | 2199.0 | 13947.0 | 3.839607 |
| 4 | 1199.0 | 13938.0 | 4.226288 |
| ... | ... | ... | ... |
| 14324 | 2738.0 | 1.0 | 5.000000 |
| 14325 | 999.0 | 1.0 | 4.000000 |
| 14326 | 899.0 | 1.0 | 4.000000 |
| 14327 | 3999.0 | 1.0 | 1.000000 |
| 14328 | 698.0 | 1.0 | 2.000000 |

14329 rows × 3 columns

In [105]: ddf=price_rating_avg.corr()
ax,fig = plt.subplots(figsize=(15,10))
ax = sns.heatmap(ddf,
annot=True,
linewidths=0.5,
fmt=".2f",
cmap="YlGnBu")
ax;



In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []: