

In [1]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
```

Reviews

In [2]:

```
reviews = pd.read_excel('reviews.xlsx')
```

In [3]:

```
reviews.head()
```

Out[3]:

	abuseCount	customerNickname	helpfulNo	helpfulYes	id	imagesCount	languageCode	postedDate	pro
0	0	iHerb Customer	0	6	05c2b17e-c28d-4792-930d-27e787d8d4ad	1	en-US	2021-01-27T09:04:10.569Z	1
1	0	iHerb Customer	0	0	80e44af8-2edf-4b81-a80a-7e7888d03cc0	0	ru-RU	2021-02-07T00:56:39.055Z	1
2	0	iHerb Customer	0	0	9a76e047-21e4-4da3-8b50-9d2396519b6b	0	en-US	2021-02-06T21:40:02.886Z	1
3	0	Innalgorevna	0	0	2890ac54-8707-418e-be3e-8d46231e3672	0	ru-RU	2021-02-05T16:29:28.906Z	1
4	0	iHerb Customer	0	0	9db33354-0457-4efa-bc9c-b5f7ee0eff31	0	ru-RU	2021-02-05T09:43:42.367Z	1



In [4]:

```
reviews.dtypes
```

Out[4]:

```
abuseCount          int64
customerNickname    object
helpfulNo           int64
helpfulYes          int64
id                  object
imagesCount         int64
languageCode        object
postedDate          object
productId           int64
profileInfo.ugcSummary.answerCount  float64
profileInfo.ugcSummary.reviewCount  float64
ratingValue         int64
reviewText          object
reviewTitle         object
reviewed            bool
score              int64
```

languageCode.1object

translation.reviewTextobject

translation.reviewTitleobject

dtype: object

In [5]:

```
reviews.describe()
```

Out[5]:

	abuseCount	helpfulNo	helpfulYes	imagesCount	productId	profileInfo.ugcSummary.answerCount	profileInfo.u
count	3849.000000	3849.000000	3849.000000	3849.000000	3849.000000		3843.000000
mean	0.045986	0.053520	0.609769	0.081320	101262.309951		3.910487
std	0.247051	0.348361	4.036665	0.429411	1168.525566		28.860762
min	0.000000	0.000000	0.000000	0.000000	99829.000000		0.000000
25%	0.000000	0.000000	0.000000	0.000000	100237.000000		0.000000
50%	0.000000	0.000000	0.000000	0.000000	101691.000000		0.000000
75%	0.000000	0.000000	0.000000	0.000000	101955.000000		1.000000
max	4.000000	9.000000	139.000000	5.000000	104023.000000		673.000000

Products

In [6]:

```
products = pd.read_excel('products.xlsx')
```

In [7]:

```
products.head()
```

Out[7]:

	product_id	product_name	product_price	price_currency	product_availability	product_url
0	103205	Hwipure, Disposable KF94 (N95 / KN95/ FFP2) ...	2.95	AUD	http://schema.org/InStock	https://au.iherb.com/pr/Hwipure-Disposable-KF9... http
1	101774	HIGUARD, Disposable KF94 (N95 / KN95/ FFP2) ...	2.95	AUD	http://schema.org/InStock	https://au.iherb.com/pr/HIGUARD-Disposable-KF9... http
2	101955	SunJoy, KN95, Professional Protective Disposab...	8.86	AUD	http://schema.org/InStock	https://au.iherb.com/pr/SunJoy-KN95-Profession... http
3	103838	Lozperi, Copper Mask, Adult, Black, 1 Mask	6.85	AUD	http://schema.org/InStock	https://au.iherb.com/pr/Lozperi-Copper-Mask-Ad... http
4	102734	Zidian, Disposable Protective Mask, 50 Pack	15.35	AUD	http://schema.org/InStock	https://au.iherb.com/pr/Zidian-Disposable-Prot... http

In [8]:

```
products.dtypes
```

products.dtypes

Out[8]:

```
product_id          int64
product_name        object
product_price       float64
price_currency      object
product_availability object
product_url         object
source_url          object
dtype: object
```

In [9]:

```
products.describe()
```

Out[9]:

	product_id	product_price
count	27.000000	27.000000
mean	102093.888889	16.969259
std	1499.395964	15.505124
min	99829.000000	2.950000
25%	100838.000000	6.260000
50%	101955.000000	8.860000
75%	103790.500000	22.440000
max	104023.000000	50.190000

Visualisation and Analysis

Tip: When plotting more than 10 categories use patterns of colors in bar plot

In [10]:

```
products['product_id'].nunique()
```

Out[10]:

27

In [11]:

```
reviews['productId'].nunique()
```

Out[11]:

27

In [12]:

```
products['productId'] = products['product_id']
products.drop('product_id', axis=1, inplace=True)
products.head()
```

Out[12]:

	product_name	product_price	price_currency	product_availability	product_url
0	Hwipure, Disposable KF94 (N95 / KN95/ FFP2) ...	2.95	AUD	http://schema.org/InStock	https://au.iherb.com/pr/Hwipure- Disposable-KF9... https://au.iherb.com/pr/Hwipure- Disposable-KF9...
	HIGUARD, Disposable				

	product_id	product_name	price_current	currency	product_url	product_url	product_url
1	195	KN95/ FFP2)					
2		SunJoy, KN95, Professional Protective Disposab...	8.86	AUD	http://schema.org/InStock	https://au.iherb.com/pr/SunJoy-KN95-Profession...	https://au.iherb.com/pr/SunJoy-KN95-Profession...
3		Lozperi, Copper Mask, Adult, Black, 1 Mask	6.85	AUD	http://schema.org/InStock	https://au.iherb.com/pr/Lozperi-Copper-Mask-Ad...	https://au.iherb.com/pr/Lozperi-Copper-Mask-Ad...
4		Zidian, Disposable Protective Mask, 50 Pack	15.35	AUD	http://schema.org/InStock	https://au.iherb.com/pr/Zidian-Disposable-Prot...	https://au.iherb.com/pr/Zidian-Disposable-Prot...

In [13]:

```
combined_prod_id = reviews.merge(products[['productId', 'product_name']], on='productId',
how='left')
combined_prod_id.head()
```

Out[13]:

	abuseCount	customerNickname	helpfulNo	helpfulYes	id	imagesCount	languageCode	postedDate	pro
0	0	iHerb Customer	0	6	05c2b17e-c28d-4792-930d-27e787d8d4ad	1	en-US	2021-01-27T09:04:10.569Z	1
1	0	iHerb Customer	0	0	80e44af8-2edf-4b81-a80a-7e7888d03cc0	0	ru-RU	2021-02-07T00:56:39.055Z	1
2	0	iHerb Customer	0	0	9a76e047-21e4-4da3-8b50-9d2396519b6b	0	en-US	2021-02-06T21:40:02.886Z	1
3	0	Innalgorevna	0	0	2890ac54-8707-418e-be3e-8d46231e3672	0	ru-RU	2021-02-05T16:29:28.906Z	1
4	0	iHerb Customer	0	0	9db33354-0457-4efa-bc9c-b5f7ee0eff31	0	ru-RU	2021-02-05T09:43:42.367Z	1

In [18]:

```
product_ratings_avg = combined_prod_id[['product_name', 'ratingValue']].groupby(['product_name']).mean().sort_values(by='ratingValue', ascending=False)
product_ratings_avg
```

Out[18]:

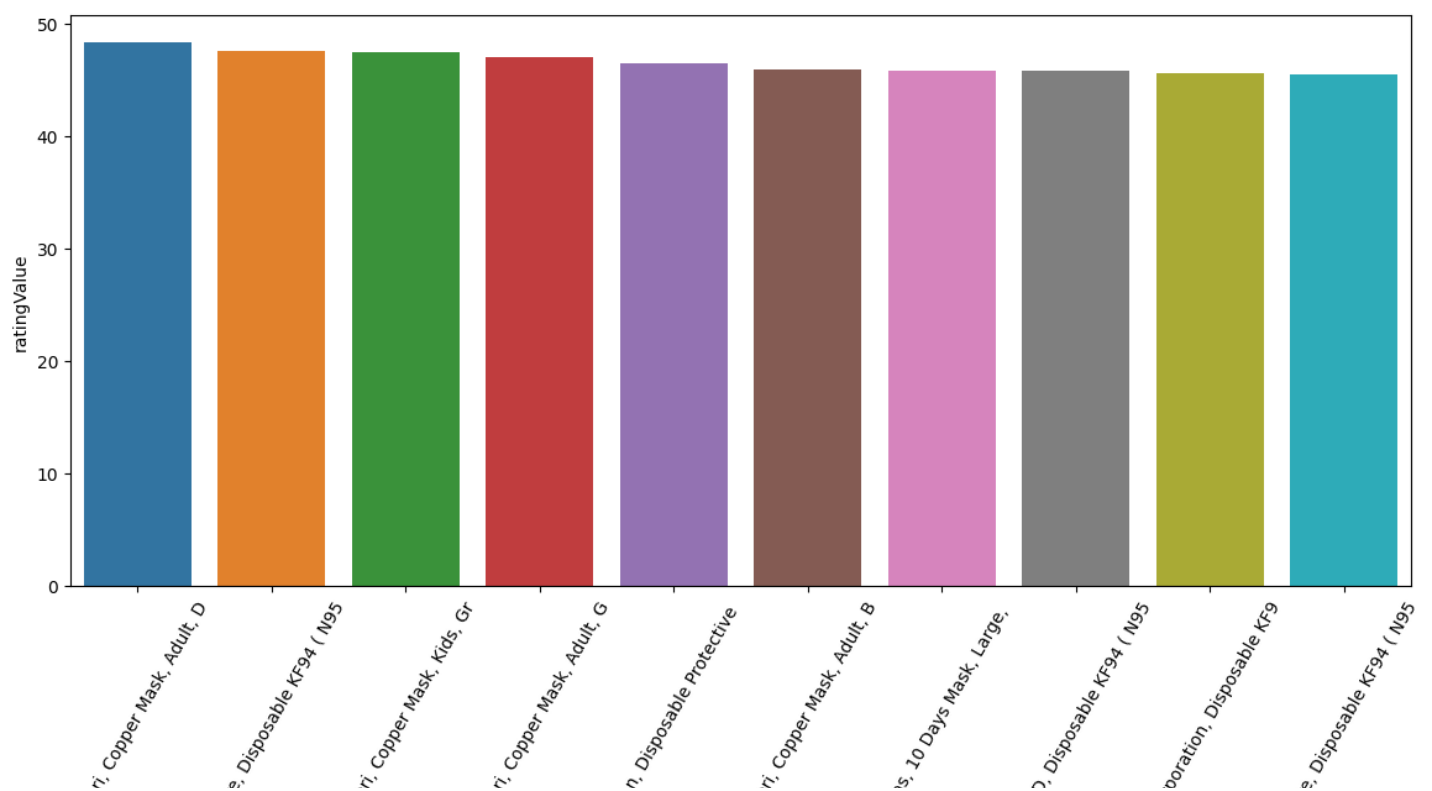
	ratingValue
product_name	
Lozperi, Copper Mask, Adult, Dot, 1 Count	48.333333
Hwipure, Disposable KF94 (N95 / KN95/ FFP2) Mask, 25 Masks	47.619048
Lozperi, Copper Mask, Kids, Gray, 1 Count	47.500000
Lozperi, Copper Mask, Adult, Gray, 1 Mask	47.000000
Zidian, Disposable Protective Mask, 50 Pack	46.498516

Lozperi, Copper Mask, Adult, Black, 1 Mask	45.30289
Puritas, 10 Days Mask, Large, 3 Pack	45.869565
HIGUARD, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	45.833333
YJ Corporation, Disposable KF94 (N95 / KN95/ FFP2) Mask, for Kids, 1 Mask	45.588235
Hwipure, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	45.555556
La Hauteur, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	45.511111
Kosette, Nano Reusable Face Protection Mask, Large, 1 Mask	45.361842
Kosette, Nano Reusable Face Protection Mask, Medium, 1 Mask	45.326316
Kitsch, 100% Cotton Reusable Face Masks, Neutral, 3 Pack	45.019305
Kosette, Fashion Mask with Filter Pocket, 3 Masks	45.000000
Kitsch, 100% Cotton Reuseable Face Masks, Blush, 3 Pack	44.579439
SunJoy, KN95, Professional Protective Disposable Face Mask, 10 Pack	44.285714
Tony Moly, CTT KN95 Respirator Mask, 5 Count	44.275862
Dr. Puri, Disposable KF94 (N95 / KN95/ FFP2) Mask, 20 Masks	44.250000
Kitsch, 100% Cotton Reuseable Face Masks, Leopard, 3 Pack	42.283019
One Fine Day, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	41.200000
Now Foods, Face Mask, Grey, S-M, 1 Mask	40.588235
Landsberg, 3 Ply Disposable Protective Face Mask, 50 Pack	40.465116
Kosette, PM 2.5 Replaceable Filter, 24 Filters	40.000000
Lozperi, Copper Mask, Kids, Black, 1 Mask	40.000000
Luseta Beauty, Disposable Protection Face Mask, 50 Pack	37.333333

In [15]:

```
#fig = px.bar(product_ratings_avg['ratingValue'])
#fig.show()

plt.figure(figsize=(14,6))
ax = sns.barplot(y=product_ratings_avg.head(10)['ratingValue'], x=product_ratings_avg.index[:10])
ax.set_xticklabels(product_ratings_avg.index[:10].map(lambda x:x[:30]), rotation=60)
pass
```



Product Ratings including quantity

In [17]:

```
combined_prod_id['product_name'].value_counts()
```

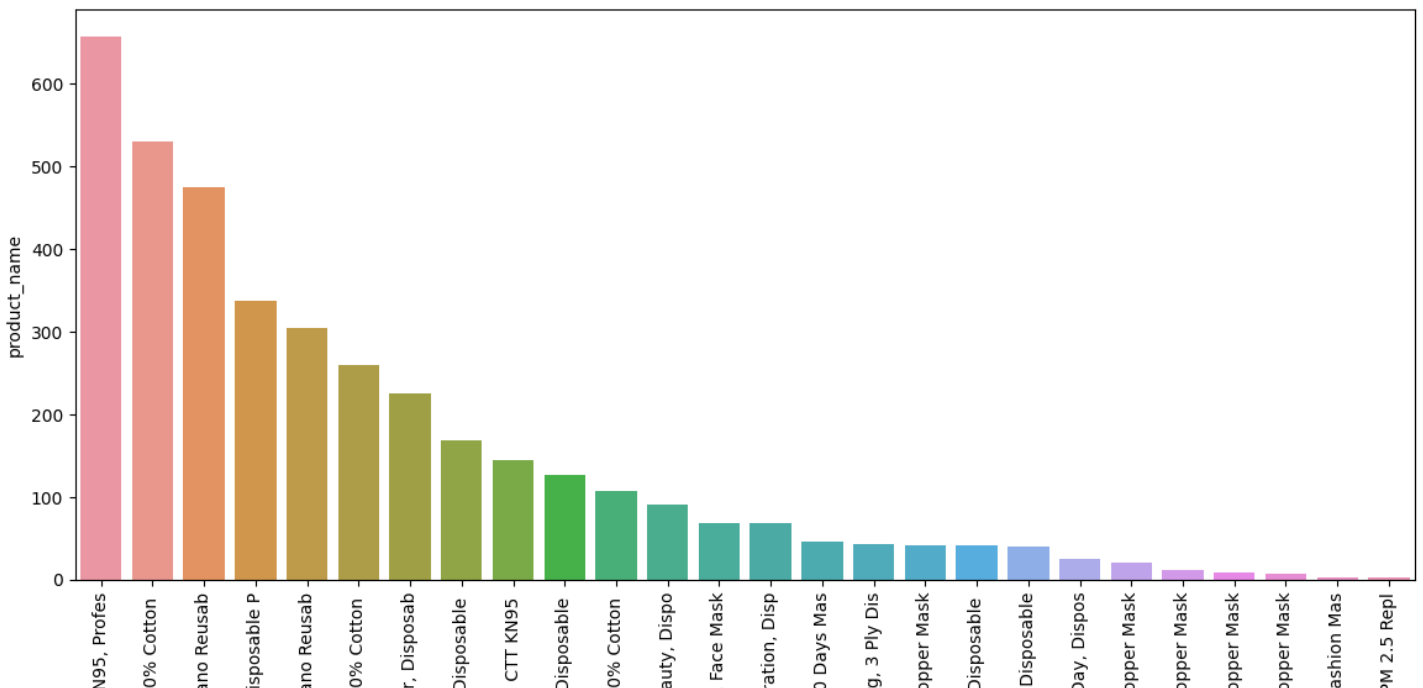
Out[17]:

SunJoy, KN95, Professional Protective Disposable Face Mask, 10 Pack	658
Kitsch, 100% Cotton Reuseable Face Masks, Leopard, 3 Pack	530
Kosette, Nano Reusable Face Protection Mask, Medium, 1 Mask	475
Zidian, Disposable Protective Mask, 50 Pack	337
Kosette, Nano Reusable Face Protection Mask, Large, 1 Mask	304
Kitsch, 100% Cotton Reusable Face Masks, Neutral, 3 Pack	259
La Hauteur, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	225
HIGUARD, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	168
Tony Moly, CTT KN95 Respirator Mask, 5 Count	145
Hwipure, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	126
Kitsch, 100% Cotton Reuseable Face Masks, Blush, 3 Pack	107
Luseta Beauty, Disposable Protection Face Mask, 50 Pack	90
Now Foods, Face Mask, Grey, S-M, 1 Mask	68
YJ Corporation, Disposable KF94 (N95 / KN95/ FFP2) Mask, for Kids, 1 Mask	68
Puritas, 10 Days Mask, Large, 3 Pack	46
Landsberg, 3 Ply Disposable Protective Face Mask, 50 Pack	43
Lozperi, Copper Mask, Adult, Black, 1 Mask	42
Hwipure, Disposable KF94 (N95 / KN95/ FFP2) Mask, 25 Masks	42
Dr. Puri, Disposable KF94 (N95 / KN95/ FFP2) Mask, 20 Masks	40
One Fine Day, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	25
Lozperi, Copper Mask, Adult, Gray, 1 Mask	20
Lozperi, Copper Mask, Adult, Dot, 1 Count	12
Lozperi, Copper Mask, Kids, Gray, 1 Count	8
Lozperi, Copper Mask, Kids, Black, 1 Mask	7
Kosette, Fashion Mask with Filter Pocket, 3 Masks	2
Kosette, PM 2.5 Replaceable Filter, 24 Filters	2

Name: product_name, dtype: int64

In [29]:

```
plt.figure(figsize=(14,6))
ax = sns.barplot(y=combined_prod_id['product_name'].value_counts(), x= combined_prod_id[
'product_name'].value_counts().index)
ax.set_xticklabels(pd.Series(combined_prod_id['product_name'].value_counts().index).map(
lambda x:x[:20]), rotation=90)
pass
```



Sunjoy, Ki
Kitsch, 10
Kosette, Ni
Zidian, D
Kosette, Na
Kitsch, 10
La Hauteu
HIGUARD, I
Tony Moly,
Hwipure, I
Kitsch, 10
Luseta Be
Now Foods,
YJ Corpor
Puritas, 1
Landsber
Lozperi, Cu
Hwipure, I
Dr. Puri,
One Fine L
Lozperi, Cu
Lozperi, Cu
Lozperi, Cu
Lozperi, Cu
Kosette, F
Kosette, F

Ratings for products with more than 50 units sold

In [35]:

```
product_counts = combined_prod_id[['product_name', 'ratingValue']].groupby(by='product_name').count()  
product_ratings_avg_50_plus = product_counts.merge(product_ratings_avg, on='product_name')
```

Out[35]:

	ratingValue_x	ratingValue_y
product_name		
Dr. Puri, Disposable KF94 (N95 / KN95/ FFP2) Mask, 20 Masks	40	44.250000
HIGUARD, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	168	45.833333
Hwipure, Disposable KF94 (N95 / KN95/ FFP2) Mask, 25 Masks	42	47.619048
Hwipure, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	126	45.555556
Kitsch, 100% Cotton Reusable Face Masks, Neutral, 3 Pack	259	45.019305
Kitsch, 100% Cotton Reuseable Face Masks, Blush, 3 Pack	107	44.579439
Kitsch, 100% Cotton Reuseable Face Masks, Leopard, 3 Pack	530	42.283019
Kosette, Fashion Mask with Filter Pocket, 3 Masks	2	45.000000
Kosette, Nano Reusable Face Protection Mask, Large, 1 Mask	304	45.361842
Kosette, Nano Reusable Face Protection Mask, Medium, 1 Mask	475	45.326316
Kosette, PM 2.5 Replaceable Filter, 24 Filters	2	40.000000
La Hauteur, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	225	45.511111
Landsberg, 3 Ply Disposable Protective Face Mask, 50 Pack	43	40.465116
Lozperi, Copper Mask, Adult, Black, 1 Mask	42	45.952381
Lozperi, Copper Mask, Adult, Dot, 1 Count	12	48.333333
Lozperi, Copper Mask, Adult, Gray, 1 Mask	20	47.000000
Lozperi, Copper Mask, Kids, Black, 1 Mask	7	40.000000
Lozperi, Copper Mask, Kids, Gray, 1 Count	8	47.500000
Luseta Beauty, Disposable Protection Face Mask, 50 Pack	90	37.333333
Now Foods, Face Mask, Grey, S-M, 1 Mask	68	40.588235
One Fine Day, Disposable KF94 (N95 / KN95/ FFP2) Mask, 1 Mask	25	41.200000
Puritas, 10 Days Mask, Large, 3 Pack	46	45.869565
SunJoy, KN95, Professional Protective Disposable Face Mask, 10 Pack	658	44.285714
Tony Moly, CTT KN95 Respirator Mask, 5 Count	145	44.275862
YJ Corporation, Disposable KF94 (N95 / KN95/ FFP2) Mask, for Kids, 1 Mask	68	45.588235
Zidian, Disposable Protective Mask, 50 Pack	337	46.498516

In []:

```
plt.figure(figsize=(14,6))  
ax = sns.barplot(y=product_ratings_avg.head(10)['ratingValue'], x=product_ratings_avg.index[:10])  
ax.set_xticklabels(product_ratings_avg.index[:10].map(lambda x:x[:30]), rotation=60)  
pass
```