Nike Data Analysis

In this project I will analyze Nike Store Data. The data has 17 columns and 112 Rows. Here are few things I would analyze:

- Gender Wise Product Distribution
- Top 5 Nike Colors
- Top 5 Nike Sizes
- Price Bar Chart
- Nike Sub_Title Bar Chart
- Nike Price Bar Chart
- Stock Availability

Step 1:

Importing necessary Python Libraries

```
In [1]:  import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import seaborn as sns
         import plotly.express as px
         from plotly.offline import init_notebook_mode, iplot
```

Step 2:

Import Nike dataset and display its head

```
In [2]:  df = pd.read_csv('nike_data.csv')
         df.head(10)
```

Out[2]:

	url	name	sub_title	brand	model	color	price	currency	availability	description	raw_description	avg_rating	review_count	images	available_sizes	uniq_id	scraped_at
0	https://www.nike.com/t/dri-fit-team-minnesota-...	Nike Dri-FIT Team (MLB Minnesota Twins)	Men's Long-Sleeve T-Shirt	Nike	14226571	Navy	40.00	USD	InStock	SWEAT-WICKING COMFORT.The Nike Dri-FIT Team (M...	<div class="pi-pdpmainbody"><p><b class="headl...	NaN	NaN	https://static.nike.com/a/images/t_PDP_1280_v1...	S | M | L | XL | 2XL	c3229e54-aa58-5fdd-9f71-fbe66366b2b2	20/09/2022 23:32:28
1	https://www.nike.com/t/club-américa-womens-dri...	Club América	Women's Nike Dri-FIT Soccer Jersey Dress	Nike	13814665	Black/Black	90.00	USD	InStock	Inspired by traditional soccer jerseys, the Cl...	<div class="pi-pdpmainbody"> <p>Inspired b...	5.0	1.0	https://static.nike.com/a/images/t_PDP_1280_v1...	L (12–14)	f8ebb2ed-17ae-5719-b750-5ea3ec69b75c	20/09/2022 23:32:40
2	https://www.nike.com/t/sportswear-swoosh-mens-...	Nike Sportswear Swoosh	Men's Overalls	Nike	13015648	Black/White	140.00	USD	OutOfStock	WORKING HARD TO KEEP YOU COMFORTABLE.The Nike ...	<div class="pi-pdpmainbody"><p><b class="headl...	4.9	11.0	https://static.nike.com/a/images/t_PDP_1280_v1...	NaN	88120081-e6cb-5399-b9dc-a2d3d5dd5206	20/09/2022 23:33:16
3	https://www.nike.com/t/dri-fit-one-luxe-big-ki...	Nike Dri-FIT One Luxe	Big Kids' (Girls') Printed Tights (Extended Size)	Nike	13809796	Black/Rush Pink	22.97	USD	OutOfStock	ELEVATED COMFORT GOES FULL BLOOM.The Nike Dri-...	<div class="pi-pdpmainbody"><p><b class="headl...	NaN	NaN	https://static.nike.com/a/images/t_PDP_1280_v1...	NaN	98348cc5-1520-5b6e-a5f6-c42547b6a092	20/09/2022 23:33:17
4	https://www.nike.com/t/paris-saint-germain-rep...	Paris Saint-Germain Repel Academy AWF	Big Kids' Soccer Jacket	Nike	13327415	Dark Grey/Black/Siren Red/Siren Red	70.00	USD	InStock	WATER-REPELLENT COVERAGE GETS PSG DETAILS.The ...	<div class="pi-pdpmainbody"><p><b class="headl...	NaN	NaN	https://static.nike.com/a/images/t_PDP_1280_v1...	XS | S | M | L | XL	f15981a5-d8c9-53fa-880d-80606be188fe	20/09/2022 23:33:22
5	https://www.nike.com/t/miami-dolphins-mike-ges...	NFL Miami Dolphins (Mike Gesicki)	Men's Game Football Jersey	Nike	14057953	White	130.00	USD	InStock	HIGHLIGHT YOUR LOYALTY TO YOUR TEAM.Rep your t...	<div class="pi-pdpmainbody"><p><b class="headl...	NaN	NaN	https://static.nike.com/a/images/t_PDP_1280_v1...	S | M | L | XL | 2XL	32f47015-e055-5757-beb4-6b57604e3141	20/09/2022 23:33:32
6	https://www.nike.com/t/knicks-icon-edition-202...	Kristaps Porziņģis Knicks Icon Edition	Nike NBA Swingman Jersey	Nike	11815891	Rush Blue	104.99	USD	OutOfStock	THE HEART OF YOUR TEAM'S IDENTITY.The Icon Edi...	<div class="pi-pdpmainbody"><p><b class="headl...	5.0	3.0	https://static.nike.com/a/images/t_PDP_1280_v1...	NaN	4b8de2f1-af57-5e5d-a208-371a2dfaa3ff	20/09/2022 23:33:38
7	https://www.nike.com/t/oregon-mens-max-90-t-sh...	Nike College (Oregon)	Men's Max 90 T-Shirt	Nike	13817332	White	18.97	USD	InStock	With a roomy fit and '90s-style detailing, th...	<div class="pi-pdpmainbody"><p><b class="headl...	NaN	NaN	https://static.nike.com/a/images/t_PDP_1280_v1...	S | M | L | XL | 2XL	de86e49a-23df-5149-b801-407fbc6b8ef9	20/09/2022 23:33:54
8	https://www.nike.com/t/sportswear-mens-logo-t-...	Nike Sportswear	Men's Logo T-Shirt	Nike	13918594	Redstone	28.97	USD	InStock	SWOOSH DETAILS.Keep it classic in the Nike Spo...	<div class="pi-pdpmainbody"><p><b class="headl...	5.0	1.0	https://static.nike.com/a/images/t_PDP_1280_v1...	S	b24f9f9b-c408-5d83-b7a6-4980af41e9d7	20/09/2022 23:34:30
9	https://www.nike.com/t/therma-crucial-catch-mi...	Nike Therma Crucial Catch (NFL Miami Dolphins)	Men's Pullover Hoodie	Nike	14233819	Black	85.00	USD	InStock	WARM, BREATHABLE FLEECE.The Nike Therma Crucia...	<div class="pi-pdpmainbody"><p><b class="headl...	NaN	NaN	https://static.nike.com/a/images/t_PDP_1280_v1...	S | M | L | XL | 2XL	41b961cc-50f1-5ed7-85a2-836a7cffceae	20/09/2022 23:34:35

For EDA, I dont need following columns:

- url
- brand
- description
- raw_description
- images
- uniq_id
- scraped_at
- currency

```
In [3]: df = df.drop(['url','brand','description','raw_description','images','uniq_id','scraped_at', 'currency'], axis = 1)
        df.head(5)
```

Out[3]:

	name	sub_title	model	color	price	availability	avg_rating	review_count	available_sizes
0	Nike Dri-FIT Team (MLB Minnesota Twins)	Men's Long-Sleeve T-Shirt	14226571	Navy	40.00	InStock	NaN	NaN	S | M | L | XL | 2XL
1	Club América	Women's Nike Dri-FIT Soccer Jersey Dress	13814665	Black/Black	90.00	InStock	5.0	1.0	L (12–14)
2	Nike Sportswear Swoosh	Men's Overalls	13015648	Black/White	140.00	OutOfStock	4.9	11.0	NaN
3	Nike Dri-FIT One Luxe	Big Kids' (Girls') Printed Tights (Extended Size)	13809796	Black/Rush Pink	22.97	OutOfStock	NaN	NaN	NaN
4	Paris Saint-Germain Repel Academy AWF	Big Kids' Soccer Jacket	13327415	Dark Grey/Black/Siren Red/Siren Red	70.00	InStock	NaN	NaN	XS | S | M | L | XL

Now I will analyze this available data. First I will check how many rows and columns I have in this dataset.

```
In [4]: df.shape
```

Out[4]: (112, 9)

There are 112 rows and 9 columns. Now I will check how many rows have null or NaN values

```
In [5]: df.isnull().sum()
```

Out[5]:
```
name               0
sub_title          0
model              0
color              2
price              0
availability       4
avg_rating        89
review_count      89
available_sizes   56
dtype: int64
```

There are 5 columns who have null values. Let's drop them

```
In [6]: df.dropna()
```

Out[6]:

	name	sub_title	model	color	price	availability	avg_rating	review_count	available_sizes
1	Club América	Women's Nike Dri-FIT Soccer Jersey Dress	13814665	Black/Black	90.00	InStock	5.0	1.0	L (12–14)
8	Nike Sportswear	Men's Logo T-Shirt	13918594	Redstone	28.97	InStock	5.0	1.0	S
23	Air Jordan 1 Zoom CMFT	Shoes	13941353	Fossil Stone/Fossil Rose/Psychic Purple/White	150.00	InStock	4.8	120.0	M 9 / W 10.5 | M 9.5 / W 11 | M 10 / W 11.5 | ...
37	Zion 2	Men's Basketball Shoes	14027597	Flax/Sesame/Fauna Brown/Muslin	119.97	InStock	4.7	3.0	8
41	Nike Air Max 90 SE	Men's Shoes	13823256	Black/Scream Green/Summit White/Obsidian	119.97	InStock	4.8	31.0	6 | 6.5 | 7 | 7.5 | 8 | 8.5 | 9 | 10
49	Nike SFB Gen 2 8"	Tactical Boot	11825513	Black/Black/Black	140.00	InStock	3.8	28.0	M 6 / W 7.5 | M 6.5 / W 8 | M 7 / W 8.5 | M 7....
70	Nike Offline Pack	Men's Shoes	13824254	Light Smoke Grey/Flat Pewter/Grey Fog/Iron Grey	92.97	InStock	4.8	5.0	8 | 8.5 | 9 | 9.5 | 10 | 10.5 | 11 | 11.5
96	MLB Houston Astros	Women's Replica Baseball Jersey	13731002	Midnight Navy	115.00	InStock	4.8	5.0	S (4–6) | M (8–10) | L (12–14) | XL (16–18) | ...
103	Nike Zoned Support	Calf Sleeves	12090072	Black	45.00	InStock	4.8	11.0	S | M | L | XL

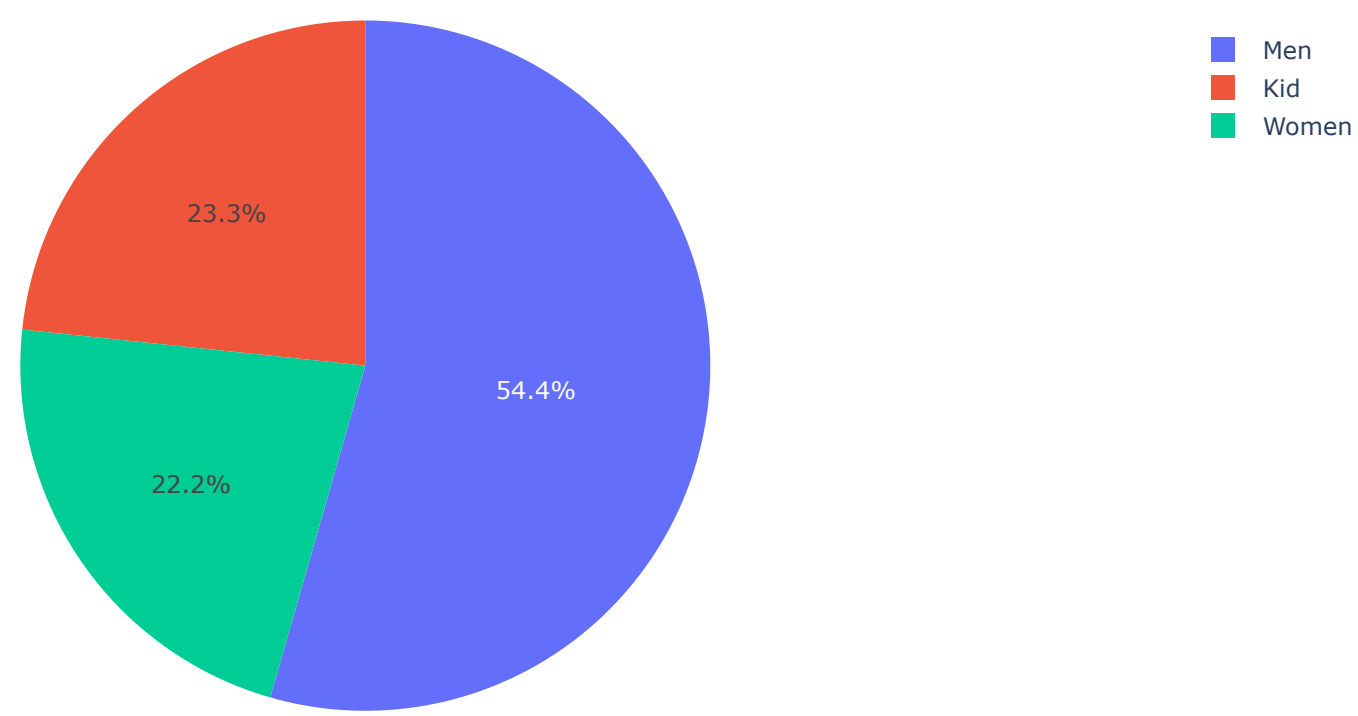## Gender wise product distribution

```
In [7]: nike_men = df[df['sub_title'].str.contains('Men')]
        nike_women = df[df['sub_title'].str.contains('Women')]
        nike_kid = df[df['sub_title'].str.contains('Kid')]

        nike_total = [nike_men.sub_title, nike_women.sub_title,nike_kid.sub_title]
        nike_total_count = [nike_men.sub_title.count(),nike_women.sub_title.count(),nike_kid.sub_title.count()]
        nike_total_count
```

Out[7]: [49, 20, 21]

```
#Pie chart showing gender wise product distribution
fig = px.pie(df, values = nike_total_count, names = ['Men',"Women",'Kid'], title = 'Gender Wise Product Distribution')
fig.show()
```
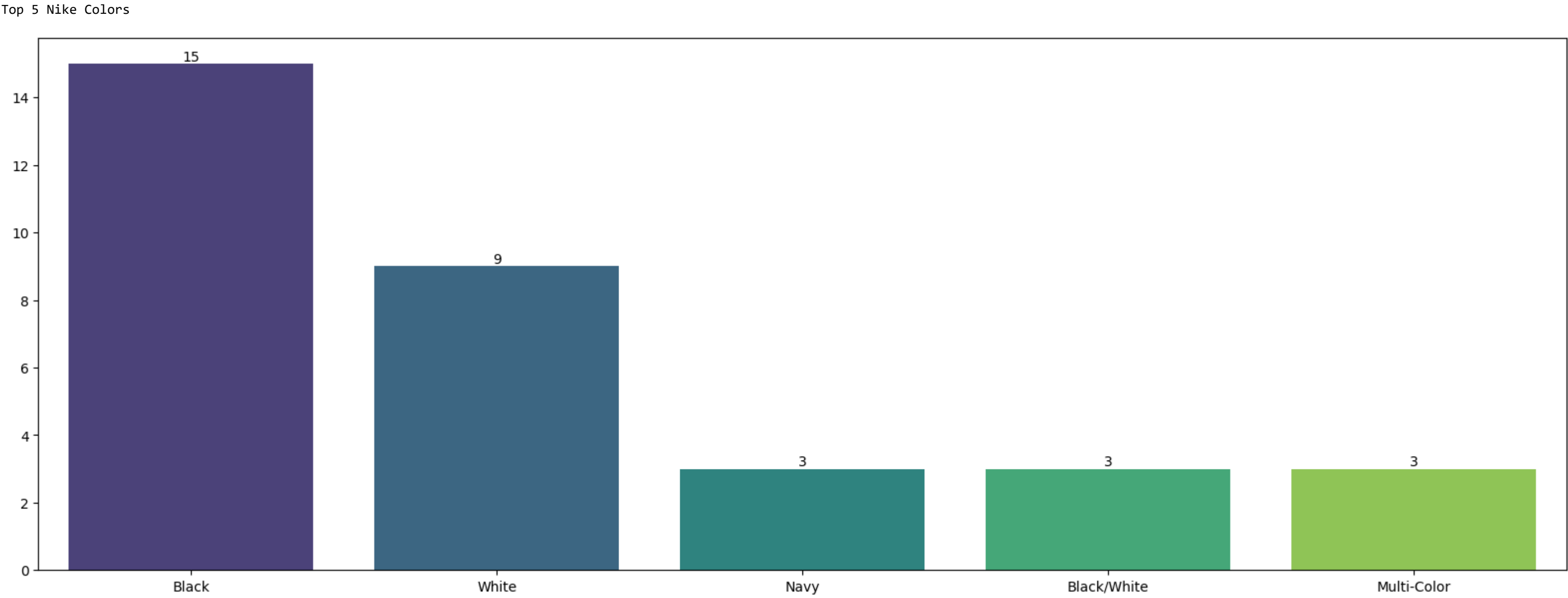
Gender Wise Product Distribution

Men
Kid
Women

Top 5 Nike Colors

```python
try:
    color = df['color'].value_counts()[:5]
    plt.figure(figsize = (20,7))
    ax = sns.barplot(x=list(color.keys()),
                     y=list(color.values),
                     palette='viridis')

    #annotate bars
    ax.bar_label(ax.containers[0])

except AttributeError:
    print('AttributeError Spotted')

finally:
    print('Top 5 Nike Colors')
```

Top 5 Nike Colors

Top 5 Nike Sizes
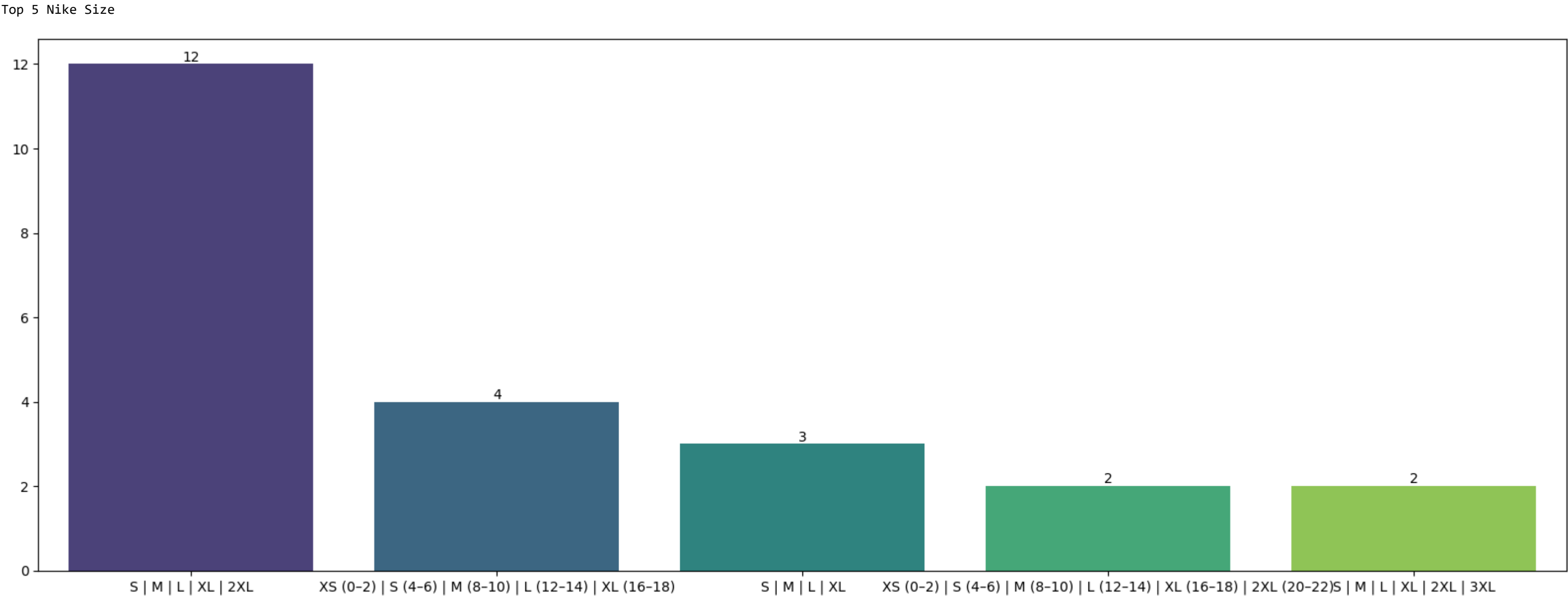
```
In [10]: try:
             size = df['available_sizes'].value_counts()[:5]
             plt.figure(figsize = (20,7))
             ax = sns.barplot(x=list(size.keys()),
                              y=list(size.values),
                              palette='viridis')

             #annotate bars
             ax.bar_label(ax.containers[0])

         except AttributeError:
             print('AttributeError Spotted')

         finally:
             print('Top 5 Nike Size')
```
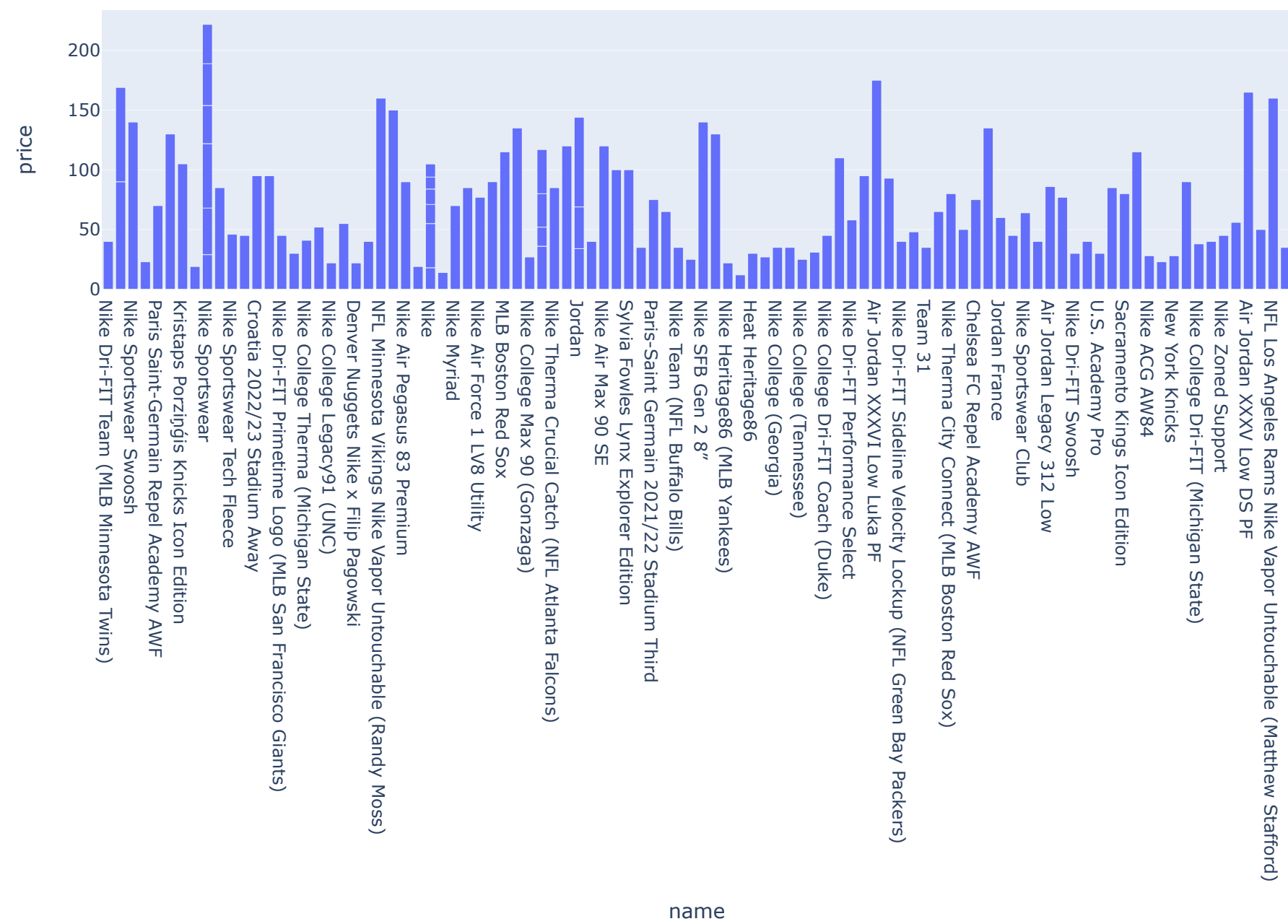
Top 5 Nike Size

Price Bar Chart

In [11]:
```
#Price Bar Chart
fig = px.bar(df, x=df['name'], y=df['price'],height=700)
fig.show()
```
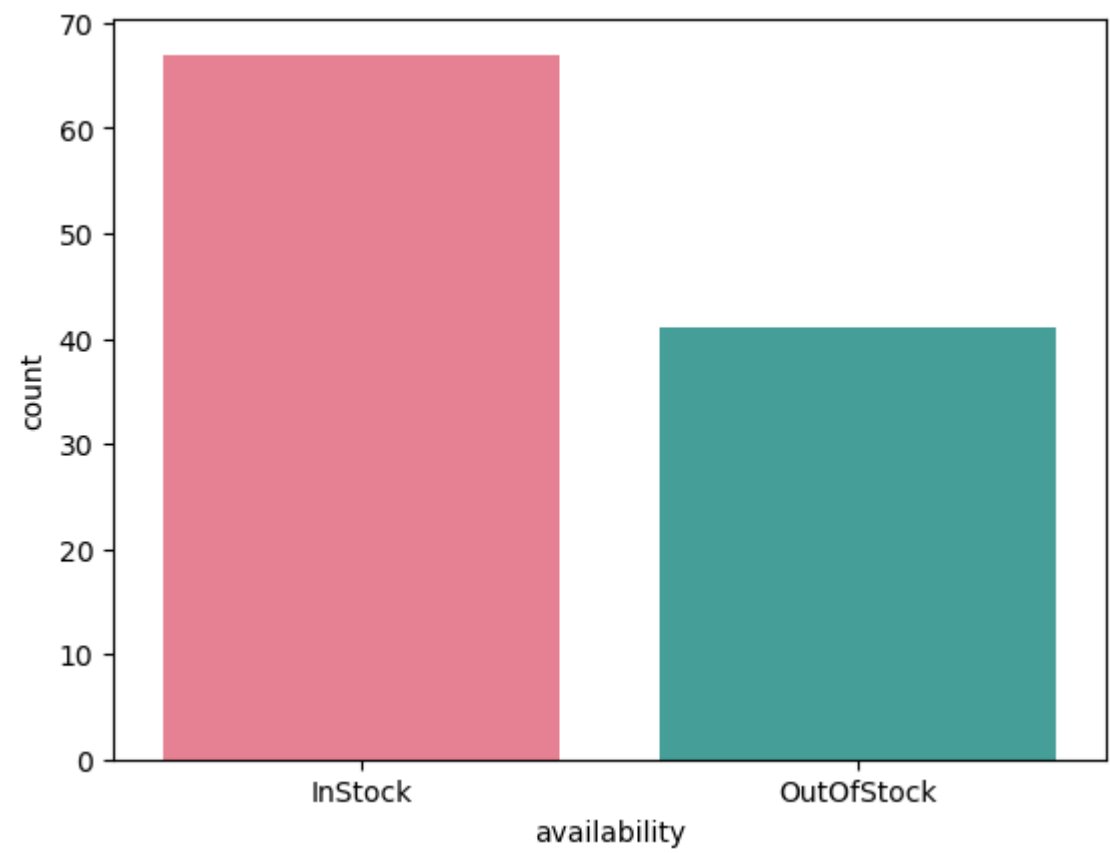

InStock/OutStock Products

In [12]:
```
#Count of Product Available
count_availability = df['availability'].value_counts()
```

```
In [13]: sns.countplot(df['availability'],palette='husl')
```

C:\Users\muhmdhas\Anaconda3\lib\site-packages\seaborn_decorators.py:36: FutureWarning:

Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
Out[13]: <AxesSubplot:xlabel='availability', ylabel='count'>
```


Average Rating

```
In [14]: import statsmodels as sm
         import plotly.express as px
         %matplotlib inline

         df['avg_rating'] = df['avg_rating'].fillna(df['avg_rating'].median(),axis=0)
         df['review_count'] = df['review_count'].fillna(df['review_count'].median(),axis=0)

         px.scatter(df,x="avg_rating", y="price", size="review_count", color="review_count",hover_name="name",log_x=True,
                    size_max=100)
```