6.4.23 Phyton

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

In [2]:
```
df = pd.read_csv('../datacsv/M6_L4 Dataset_sheet1.csv')
```

Drop semua baris yang memuat value null values dan abaikan (exclude) value "#REF" di dalam kolom SKU Code.

In [3]:
```
df.head()
# drop SKU Code = #REF -------------------------
df2 = df.drop(df[df["SKU Code"] == "#REF"].index)
df2
# drop NaN -------
df3 = df2.dropna()
# df3.count()
```

In [4]:
```
df1 = df['SKU Code'].count()
print(f'Jumlah Row Sebelum Data Clean',df.count())
df3
print(f' ')
print(f'Jumlah Row Setelah Data Clean',df3.count())
```

```
Jumlah Row Sebelum Data Clean index         9271
SKU Code      9188
Design No.    9235
Stock         9235
Category      9226
Size          9235
Color         9226
dtype: int64

Jumlah Row Setelah Data Clean index         9188
SKU Code      9188
Design No.    9188
Stock         9188
Category      9188
Size          9188
Color         9188
dtype: int64
```

Pivot DataFrame agar barisnya dikelompokkan berdasarkan, Category dan Color, gunakan kolom Design No. sebagai kolom, dan kolom Stock sebagai value

In [5]:
```
pivot_df  = pd.pivot_table(df3, values='Stock', index=[ 'Category'],column
```

In [6]:
```
pivot_df
```

Out[6]:

Color Category	AQUA GREEN	BURGUNDY	Beige	Black	Blue	Brown	CORAL	CORAL ORANGE	CORAL PINK	Cha
AN : LEGGINGS	0	0	37	30	0	49	0	0	0	
BLOUSE	0	0	492	3440	279	187	0	0	0	
BOTTOM	0	0	12	0	0	0	0	0	0	
CARDIGAN	0	0	0	0	14	0	0	0	0	
CROP TOP	0	0	0	26	42	0	0	0	0	
CROP TOP WITH PLAZZO	0	0	0	0	0	0	0	0	0	
DRESS	0	0	39	248	2762	3	0	6	18	
JUMPSUIT	0	0	0	0	0	0	0	0	0	
KURTA	0	0	4010	12648	10220	1310	316	0	0	
KURTA SET	5	0	241	1278	2066	3895	0	0	0	
KURTI	0	0	0	0	0	250	0	0	0	
LEHENGA CHOLI	0	0	0	0	0	0	0	0	0	
NIGHT WEAR	0	0	167	0	528	0	0	0	0	
PALAZZO	0	0	0	47	0	86	0	0	0	
PANT	0	0	12	1238	0	244	0	0	0	
SAREE	0	0	75	129	171	100	0	0	0	
SET	0	33	606	1550	981	535	163	0	0	
SHARARA	0	0	0	305	0	0	0	0	0	
SKIRT	0	0	0	340	0	0	0	0	0	
TOP	0	45	439	740	2053	20	0	0	0	
TUNIC	0	0	0	174	182	0	0	0	0	

21 rows × 60 columns

In [7]:
```
avg_stock_df = pd.pivot_table(df3, values='Stock', index=[ 'Category'], agg
```

In [8]:
```
avg_stock_df.head()
```

Out[8]:

Category	Stock
AN : LEGGINGS	5.500000
BLOUSE	33.867521
BOTTOM	2.157895
CARDIGAN	3.625000
CROP TOP	9.642857

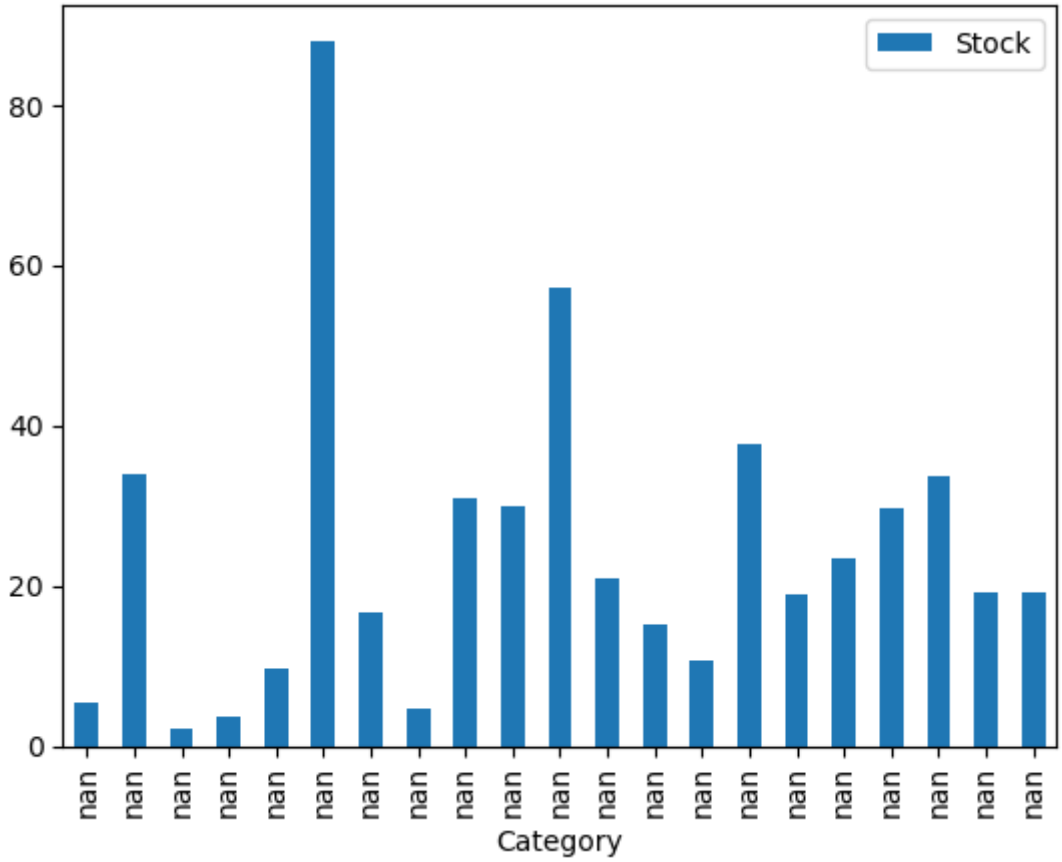
In [9]:
```
avg_csv = avg_stock_df.groupby(['Stock','Category']).mean()
avg_csv
```

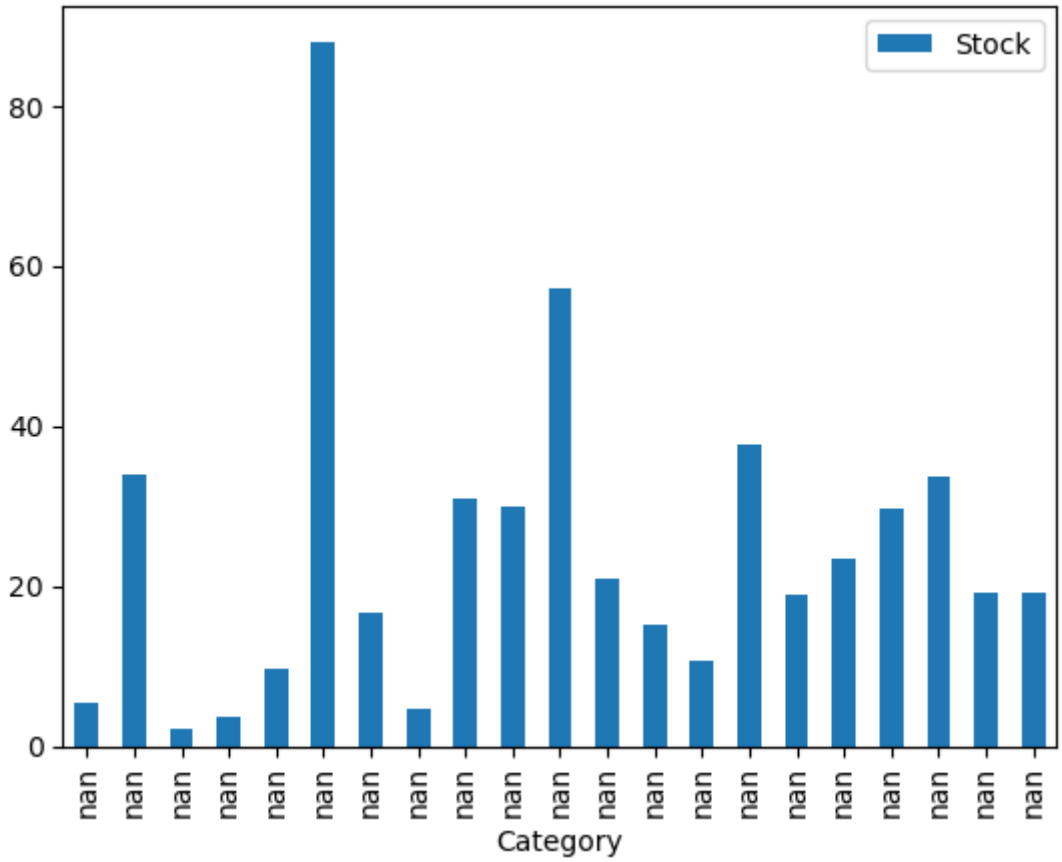
Out[9]:

Stock	Category
2.157895	BOTTOM
3.625000	CARDIGAN
4.714286	JUMPSUIT
5.500000	AN : LEGGINGS
9.642857	CROP TOP
10.582418	PALAZZO
15.152074	NIGHT WEAR
16.678571	DRESS
18.891156	SAREE
19.266234	TUNIC
19.289199	TOP
21.028571	LEHENGA CHOLI
23.491897	SET
29.750000	SHARARA
29.875940	KURTA SET
30.867441	KURTA
33.650000	SKIRT
33.867521	BLOUSE
37.758242	PANT
57.321429	KURTI
88.142857	CROP TOP WITH PLAZZO

In [10]:
```
## to csv file
avg_csv.to_csv('../datacsv/avg_stock.csv')
```

In [14]:
```
#print (dfg)
dfg = pd.DataFrame(avg_stock_df, columns = ['Category','Stock'])
dfg.plot(x='Category', y='Stock', kind='bar')
dfg.plot(x='Category', y='Stock', kind='bar')
plt.show()
```


Stock
80
60
40
20
0
nan
Category


Stock
80
60
40
20
0
nan
Category

In [ ]:

In [ ]: