```
#ENGINEERING FEATURE-1
import pandas as pd
import numpy as np
# Creating a dummy DataFrame of 15 numbers randomly
# ranging from 1-100 for age
df = pd.DataFrame({'Age': [42, 15, 67, 55, 1, 29, 75, 89, 4,
10, 15, 38, 22, 77]})
# Printing DataFrame Before sorting Continuous
# to Categories
print("Before: \n")
print(df)
# A column of name 'Label' is created in DataFrame
# Categorizing Age into 4 Categories
# Baby/Toddler: (0,3], 0 is excluded & 3 is included
# Child: (3,17], 3 is excluded & 17 is included
# Adult: (17,63], 17 is excluded & 63 is included
# Elderly: (63,99], 63 is excluded & 99 is included
df['Label'] = pd.cut(x=df['Age'], bins=[0, 3, 17, 63, 99],
labels=['Baby/Toddler', 'Child', 'Adult',
'Elderly'])
# Printing DataFrame after sorting Continuous to
# Categories
print("After: \n")
print(df)
# Check the number of values in each bin
print("Categories: \n")
print(df['Label'].value_counts())
```

```
Before:

    Age
0    42
1    15
2    67
3    55
4     1
5    29
6    75
7    89
8     4
9    10
10   15
11   38
12   22
13   77
After:

    Age         Label
0    42         Adult
1    15         Child
2    67       Elderly
3    55         Adult
4     1  Baby/Toddler
5    29         Adult
6    75       Elderly
7    89       Elderly
8     4         Child
9    10         Child
10   15         Child
11   38         Adult
12   22         Adult
13   77       Elderly
Categories:

Label
Adult          5
Child          4
Elderly        4
Baby/Toddler   1
Name: count, dtype: int64
```

```
#ENGINEERING FEATURE-2
# Importing pandas and numpy libraries
import pandas as pd
import numpy as np
# Creating a dummy DataFrame of 12 numbers randomly
# ranging from 150-180 for height
df = pd.DataFrame({'Height': [150.4, 157.6, 170, 176, 164.2, 155,
159.2, 175, 162.4, 176, 153, 170.9]})
# Printing DataFrame Before Sorting Continuous to Categories
print("Before: ")
print(df)
# A column of name 'Label' is created in DataFrame
```

```
# Categorizing Height into 3 Categories
# Short: (150,157], 150 is excluded & 157 is included
# Average: (157,169], 157 is excluded & 169 is included
# Tall: (169,180], 169 is excluded & 180 is included.
df['Label'] = pd.cut(x=df['Height'],
bins=[150, 157, 169, 180],
labels=['Short', 'Average', 'Tall'])
# Printing Data Frame After Sorting Continuous to Categories
print("After: ")
print(df)
# Check the number of values in each bin
print("Categories: ")
print(df['Label'].value_counts())
```

```
Before:
    Height
0    150.4
1    157.6
2    170.0
3    176.0
4    164.2
5    155.0
6    159.2
7    175.0
8    162.4
9    176.0
10   153.0
11   170.9
After:
    Height    Label
0    150.4    Short
1    157.6  Average
2    170.0     Tall
3    176.0     Tall
4    164.2  Average
5    155.0    Short
6    159.2  Average
7    175.0     Tall
8    162.4  Average
9    176.0     Tall
10   153.0    Short
11   170.9     Tall
Categories:
Label
Tall       5
Average    4
Short      3
Name: count, dtype: int64
```