

Bar Plot in Python

by [Venmani A D](https://www.machinelearningplus.com/author/venmani/) (<https://www.machinelearningplus.com/author/venmani/>)

A bar plot shows categorical data as rectangular bars with the height of bars proportional to the value they represent. It is often used to compare between values of different categories in the data.

Content

1. What is a barplot?
2. Simple bar plot using matplotlib
3. Horizontal barplot
4. Changing color of a barplot
5. Grouped and Stacked Barplots
6. Grouped barplot
7. Stacked barplot
8. List of available palettes

What is a barplot?

A bar plot shows categorical data as rectangular bars with heights proportional to the value they represent. It is often used to compare between values of different categories in the data.

What is categorical data?

A categorical data is nothing but a grouping of data into different logical groups, for example, data on the height of persons being grouped as 'Tall', 'Medium', 'Short' etc.

To make a bar plot, you need to calculate the count of each category.

First you need to install all the required libraries which we will be using. I also changed the default `figsize` and `dpi` [dots per inch] parameters by using `plt.rcParams.update()` function.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
plt.rcParams.update({'figure.figsize':(7,5), 'figure.dpi':100})
```

Lets create a dataset containing 10 discrete categories and assign values to each category.

To create a random array, use `np.random.randn()` function with lower limit, upper limit and size of the array as arguments.

Recent Posts

[Requests in Python \(Guide\)](https://www.machinelearningplus.com/python/requests-in-python/)
(<https://www.machinelearningplus.com/python/requests-in-python/>)

[Matplotlib Pyplot](https://www.machinelearningplus.com/plots/matplotlib/matplotlib-pyplot/)
([https://www.machinelearningplus.com/plots/matplotlib-pyplot/](https://www.machinelearningplus.com/plots/matplotlib/matplotlib-pyplot/))

[Python Boxplot](https://www.machinelearningplus.com/plots/python-boxplot/)
(<https://www.machinelearningplus.com/plots/python-boxplot/>)

[Bar Plot in Python](https://www.machinelearningplus.com/plots/bar-plot-in-python/)
(<https://www.machinelearningplus.com/plots/bar-plot-in-python/>)

[datatable in R – The Complete Beginners Guide](https://www.machinelearningplus.com/data-manipulation/datatable-in-r-complete-guide/)
(<https://www.machinelearningplus.com/data-manipulation/datatable-in-r-complete-guide/>)

[Augmented Dickey Fuller Test \(ADF Test\) – Must Read Guide](https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/)
(<https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/>)

[KPSS Test for Stationarity](https://www.machinelearningplus.com/time-series/kpss-test-for-stationarity/)
(<https://www.machinelearningplus.com/time-series/kpss-test-for-stationarity/>)

[101 R datatable Exercises](https://www.machinelearningplus.com/data-manipulation/101-r-data-table-exercises/)
(<https://www.machinelearningplus.com/data-manipulation/101-r-data-table-exercises/>)

[P-Value – Understanding from Scratch](https://www.machinelearningplus.com/statistics/p-value/)
(<https://www.machinelearningplus.com/statistics/p-value/>)

[101 Python datatable Exercises \(pydatatable\)](https://www.machinelearningplus.com/data-manipulation/101-python-datatable-exercises-pydatatable/)
(<https://www.machinelearningplus.com/data-manipulation/101-python-datatable-exercises-pydatatable/>)

[Vector Autoregression \(VAR\) – Comprehensive Guide with Examples in Python](https://www.machinelearningplus.com/time-series/vector-autoregression-examples-python/)
(<https://www.machinelearningplus.com/time-series/vector-autoregression-examples-python/>)

[Mahalanobis Distance – Understanding the math with examples \(python\)](https://www.machinelearningplus.com/statistics/mahalanobis-distance/)
(<https://www.machinelearningplus.com/statistics/mahalanobis-distance/>)

[datetime in Python – Simplified Guide with Clear Examples](https://www.machinelearningplus.com/python/datetime-in-python/)
(<https://www.machinelearningplus.com/python/datetime-in-python/>)

```
# Create Data
x=['A','B','C','D','E','F','G','H','I','J']
y = np.random.randint(low=0, high=100, size=10)
y
```

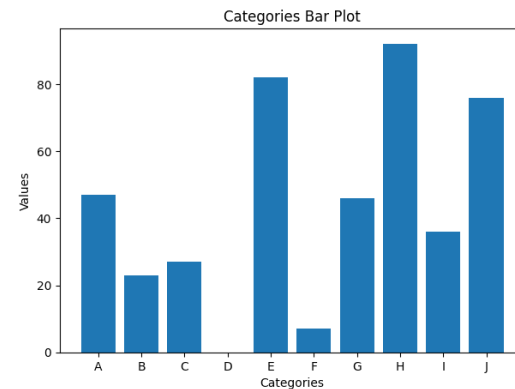
```
array([47, 23, 27,  0, 82,  7, 46, 92, 36, 76])
```

You can see that y contains an array of randomly assigned values

Simple bar plot using matplotlib

For plotting a barplot in matplotlib, use `plt.bar()` function passing 2 arguments – [`x_value` , `y_value`]

```
# Simple Bar Plot
plt.bar(x,y)
plt.xlabel('Categories')
plt.ylabel("Values")
plt.title('Categories Bar Plot')
plt.show()
```



(<https://www.machinelearningplus.com/wp-content/uploads/2020/04/simple-bar-plot.png>)

In the above barplot we can visualize the array we just created using `random()` function.

Horizontal barplot

You can also visualize the same graph horizontally using the `barh()` function with the same values as arguments.

[Principal Component Analysis \(PCA\) – Better Explained](https://www.machinelearningplus.com/machine-learning/principal-components-analysis-pca-better-explained/)
(<https://www.machinelearningplus.com/machine-learning/principal-components-analysis-pca-better-explained/>)

[Python Logging – Simplest Guide with Full Code and Examples](https://www.machinelearningplus.com/python/python-logging-guide/)
(<https://www.machinelearningplus.com/python/python-logging-guide/>)

[Matplotlib Histogram – How to Visualize Distributions in Python](https://www.machinelearningplus.com/plots/matplotlib/matplotlib-histogram/)
([https://www.machinelearningplus.com/plots/matplotlib-histogram-python-examples/](https://www.machinelearningplus.com/plots/matplotlib/matplotlib-histogram/))

[ARIMA Model – Complete Guide to Time Series Forecasting in Python](https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/)
(<https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/>)

[Time Series Analysis in Python – A Comprehensive Guide with Examples](https://www.machinelearningplus.com/time-series/time-series-analysis-python/)
(<https://www.machinelearningplus.com/time-series/time-series-analysis-python/>)

[Matplotlib Tutorial – A Complete Guide to Python Plot w/ Examples](https://www.machinelearningplus.com/plots/matplotlib/matplotlib-tutorial-complete-guide-python-plot-examples/)
([https://www.machinelearningplus.com/plots/matplotlib-tutorial-complete-guide-python-plot-examples/](https://www.machinelearningplus.com/plots/matplotlib/matplotlib-tutorial-complete-guide-python-plot-examples/))

[Topic modeling visualization – How to present the results of LDA models?](https://www.machinelearningplus.com/nlp/topic-modeling-visualization-how-to-present-results-lda-models/)
(<https://www.machinelearningplus.com/nlp/topic-modeling-visualization-how-to-present-results-lda-models/>)

Top Posts & Pages

[ARIMA Model - Complete Guide to Time Series Forecasting in Python](https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/)
(<https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/>)

[Time Series Analysis in Python - A Comprehensive Guide with Examples](https://www.machinelearningplus.com/time-series/time-series-analysis-python/)
(<https://www.machinelearningplus.com/time-series/time-series-analysis-python/>)

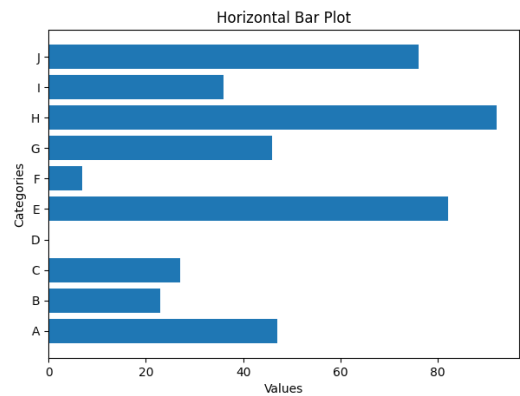
[Top 50 matplotlib Visualizations - The Master Plots \(with full python code\)](https://www.machinelearningplus.com/plots/top-50-matplotlib-visualizations-the-master-plots-python/)
(<https://www.machinelearningplus.com/plots/top-50-matplotlib-visualizations-the-master-plots-python/>)

[Cosine Similarity - Understanding the math and how it works \(with python codes\)](https://www.machinelearningplus.com/nlp/cosine-similarity/)
(<https://www.machinelearningplus.com/nlp/cosine-similarity/>)

[Parallel Processing in Python - A Practical Guide with Examples](https://www.machinelearningplus.com/python/parallel-processing-python/)
(<https://www.machinelearningplus.com/python/parallel-processing-python/>)

[Topic Modeling with Gensim \(Python\)](https://www.machinelearningplus.com/nlp/topic-modeling-gensim-python/)
(<https://www.machinelearningplus.com/nlp/topic-modeling-gensim-python/>)

```
# Horizontal Bar plot
plt.barh(x,y)
plt.xlabel("Values")
plt.ylabel('Categories')
plt.title('Horizontal Bar Plot')
plt.show()
```



<https://www.machinelearningplus.com/wp-content/uploads/2020/04/horizontal-bar-plot.png>

Changing color of a barplot

You can also change the color of the bar by using the [color= ' '] command in the plt.bar() fuction.

```
# Change color of each bar
plt.bar(x,y, color=['firebrick', 'green', 'blue', 'black', 'red',
                    'purple', 'seagreen', 'skyblue', 'black', 'tomato'])
plt.xlabel('Categories')
plt.ylabel("Values")
plt.title('Barplot with colored bars')
plt.show()
```

101 Pandas Exercises for Data Analysis
(<https://www.machinelearningplus.com/python/101-pandas-exercises-python/>)

Complete Introduction to Linear Regression in R
(<https://www.machinelearningplus.com/machine-learning/complete-introduction-linear-regression-r/>)

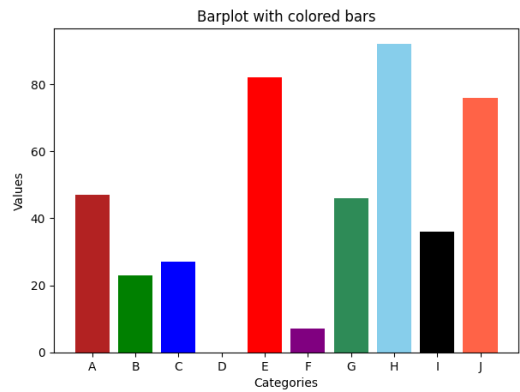
Matplotlib Histogram - How to Visualize Distributions in Python
(<https://www.machinelearningplus.com/plots/matplotlib-histogram-python-examples/>)

101 NumPy Exercises for Data Analysis (Python)
(<https://www.machinelearningplus.com/python/101-numpy-exercises-python/>)

Tags

- Bigrams (<https://www.machinelearningplus.com/tag/bigrams/>)
- Classification (<https://www.machinelearningplus.com/tag/classification/>)
- Corpus (<https://www.machinelearningplus.com/tag/corpus/>)
- Cosine Similarity (<https://www.machinelearningplus.com/tag/cosine-similarity/>)
- data.table (<https://www.machinelearningplus.com/tag/data-table/>)
- Data Manipulation (<https://www.machinelearningplus.com/tag/data-manipulation/>)
- debugging (<https://www.machinelearningplus.com/tag/debugging/>)
- Doc2Vec (<https://www.machinelearningplus.com/tag/doc2vec/>)
- Evaluation Metrics (<https://www.machinelearningplus.com/tag/evaluation-metrics/>)
- FastText (<https://www.machinelearningplus.com/tag/fasttext/>)
- Feature Selection (<https://www.machinelearningplus.com/tag/feature-selection/>)
- Gensim (<https://www.machinelearningplus.com/tag/gensim/>)
- LDA (<https://www.machinelearningplus.com/tag/lda/>)
- Lemmatization (<https://www.machinelearningplus.com/tag/lemmatization/>)
- Linear Regression (<https://www.machinelearningplus.com/tag/linear-regression/>)
- Logistic (<https://www.machinelearningplus.com/tag/logistic/>)
- LSI (<https://www.machinelearningplus.com/tag/lsi/>)
- Matplotlib (<https://www.machinelearningplus.com/tag/matplotlib/>)
- Multiprocessing (<https://www.machinelearningplus.com/tag/multiprocessing/>)
- NLP (<https://www.machinelearningplus.com/tag/nlp/>)
- NLTK (<https://www.machinelearningplus.com/tag/nltk/>)
- Numpy (<https://www.machinelearningplus.com/tag/numpy/>)
- P-Value (<https://www.machinelearningplus.com/tag/p-value/>)
- Pandas (<https://www.machinelearningplus.com/tag/pandas/>)
- Parallel Processing (<https://www.machinelearningplus.com/tag/parallel-processing/>)

Feedback



<https://www.machinelearningplus.com/wp-content/uploads/2020/04/coloured-bar-plot.png>

Grouped and Stacked Barplots

There are 2 types of barplots – Grouped and stacked barplots.

Let's look into this with an example of the famous titanic dataset.

This dataset contains the data of whether the person has survived or not during the sink of titanic and different details of the person.

You can download the below used dataset from the [link](https://github.com/ven-27/datasets/blob/master/titanic.csv) (<https://github.com/ven-27/datasets/blob/master/titanic.csv>).

Grouped barplot

In a grouped bar chart, for each categorical group there are two or more bars.

```
# Import data
# df=pd.read_csv("titanic.csv")
df=pd.read_csv("https://raw.githubusercontent.com/ven-27/datasets/master/titanic.csv")
df.head()
```

PassengerId	Survived	Price	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A5 21171	7.2500	NaN	S
1	2	1	1	Cummings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Hekkinen, Miss. Laina	female	26.0	0	0	STON/O2 3101282	7.9200	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.5500	NaN	S

https://www.machinelearningplus.com/wp-content/uploads/2020/04/titanic_data-min.png

If you want to plot Grouped barplots then use seaborn package using hue='groupColumnName' which will contain the category which you are grouping into x .

So basically you are splitting the x[Sex] category further into categories 'Pclass' .

This is called as grouped barplot.

(<https://www.machinelearningplus.com/tag/parallel-processing/>)
Phrase (<https://www.machinelearningplus.com/tag/phrase/>)
Practice Exercise (<https://www.machinelearningplus.com/tag/practice-exercise/>)

Python

(<https://www.machinelearningplus.com/tag/python/>)

R (<https://www.machinelearningplus.com/tag/r/>)
Regression (<https://www.machinelearningplus.com/tag/regression/>)

(<https://www.machinelearningplus.com/tag/residual-analysis/>)
Residual Analysis (<https://www.machinelearningplus.com/tag/residual-analysis/>)

Scikit Learn (<https://www.machinelearningplus.com/tag/scikit-learn/>)

Significance Tests (<https://www.machinelearningplus.com/tag/significance-tests/>)

Soft Cosine Similarity (<https://www.machinelearningplus.com/tag/soft-cosine-similarity/>)

spaCy (<https://www.machinelearningplus.com/tag/spacy/>)

Stationarity (<https://www.machinelearningplus.com/tag/stationarity/>)

Summarization (<https://www.machinelearningplus.com/tag/summarization/>)

TaggedDocument (<https://www.machinelearningplus.com/tag/taggeddocument/>)

TextBlob (<https://www.machinelearningplus.com/tag/textblob/>)

TFIDF (<https://www.machinelearningplus.com/tag/tfidf/>)

Time Series (<https://www.machinelearningplus.com/tag/time-series/>)

Topic Modeling (<https://www.machinelearningplus.com/tag/topic-modeling/>)

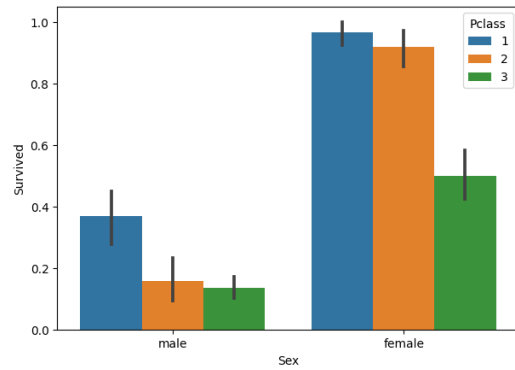
Visualization (<https://www.machinelearningplus.com/tag/visualization/>)

Word2Vec (<https://www.machinelearningplus.com/tag/word2vec/>)

(<https://www.machinelearningplus.com/tag/word2vec/>)

Feedback

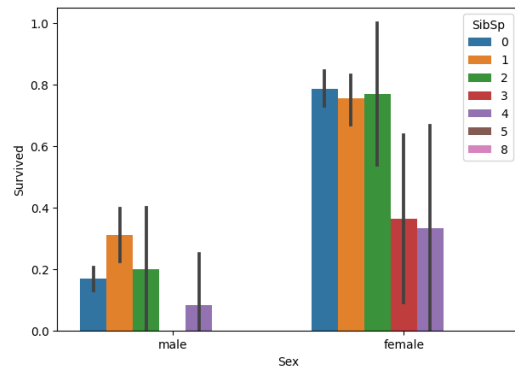
```
# Grouped bar plot with seaborn
import seaborn as sns
sns.barplot(y='Survived',x='Sex',hue='Pclass',data=df);
```



(<https://www.machinelearningplus.com/wp-content/uploads/2020/04/titanic-barplot1.png>)

Let's see another example.

```
# Another example
sns.barplot(y='Survived',x='Sex',hue='SibSp',data=df);
```



(<https://www.machinelearningplus.com/wp-content/uploads/2020/04/grouped-bar-plot1.png>)

You can also see the error bar for each category.

Error bars are graphical representations of the variability of data and used on graphs to indicate the error or uncertainty in a reported measurement.

Error bars often represent one standard deviation of uncertainty, one standard error, or a particular confidence interval (e.g., a 95% interval).

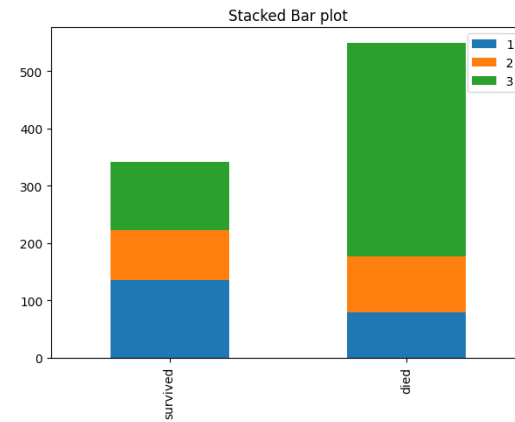
Stacked barplot

The stacked bar chart stacks bars that represent different groups on top of each other.

This can be done in pandas library by using `stacked=True` command in `df.plot()` function.

```
# Stacked barplot with pandas
survived = df.loc[df['Survived']==1, :]['Pclass'].value_counts()
died = df.loc[df['Survived']==0, :]['Pclass'].value_counts()
df_plot = pd.DataFrame([survived,died])
df_plot.index=['survived','died']

# Plot
df_plot.plot(kind='bar',stacked=True, title='Stacked Bar plot');
```



(<https://www.machinelearningplus.com/wp-content/uploads/2020/04/stacked-barplot-1.png>)

The above graph is categorized based on whether the passenger survived or not and also stacked based on the class in which the passenger is traveling.

Let's draw with different palette.

```
# Stacked barplot with pandas with differnt palette
survived = df.loc[df['Survived']==1, :]['Sex'].value_counts()
died = df.loc[df['Survived']==0, :]['Sex'].value_counts()
df_plot = pd.DataFrame([survived,died])
df_plot.index=['survived','died']

# Bar pLot
df_plot.plot(kind='bar',stacked=True, colormap='Spectral', title='Stacked Bar p
```

