In [ ]:
```
#Normal distribution
```

In [ ]:
```
dataset['column'] - #shows all members of the column
```

In [ ]:
```
dataset['column']/100 = # divides all the members with 100
```

In [ ]:
```
dataset['column'] > 5 = #shows the values (rows) with greater than 5 values
```

In [ ]:
```
dataset['column'].replace([value], np.nan, inplace=True) - # replaces the value with NaN, we can also put somet
```

In [ ]:
```
function = dataset['column'] == 1
~function = #everything besides the condition
```

In [2]:
```
pmf = Pmf(dataset['column'], normalize = True/False) # probability mass function
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
<ipython-input-2-134f4d357463> in <module>
----> 1 pmf = Pmf(dataset['column'], normalize = True/False) # probability mass function

NameError: name 'Pmf' is not defined
```

In [ ]:
```
cdf = Cdf(dataset['column']) # cummulative distribution functions
cdf.plot()
```

In [ ]:
```
q = 50
p = cdf(q)    #finds the percent, can be used with cdf.inverse(q) and q < 1
```

In [21]:
```
xs = np.linspace(-5,5)
ys = norm(0,2).pdf(xs)
plt.plot(xs, ys)
```

Out[21]:
```
[<matplotlib.lines.Line2D at 0x2291ace40a0>]
```