There is a huge library for Probability and Statistics. We will explore some simple subjects.
To truly understand this subject requires fairly advanced Mathematics. It is not difficult, it just requires patience and time.

Factorials come into use when evaluating the number of ways to arrange objects.
Factorial[5]= 5 ! = 5*4*3*2*1

```
In[]:= 5 !
Out[]= 120
```

Binomial[n, m] gives the binomial coefficient n !/(m ! (n - m)!) .

Suppose you have 4 boys in the class, and you must pick 2 to run across the school to get you some colored pencils and chaulk.

Let’s call the boys “Ketevi”, “Bob”, “Yash”, and “Johannes”
How many ways can you get 2 boys out of 4.

```
In[7]:= listofboys = {"Ketevi", "Bob", "Yash", "Johannes"};
        chooseboys = Subsets[listofboys, {2}]
Out[8]= {{Ketevi, Bob}, {Ketevi, Yash}, {Ketevi, Johannes},
         {Bob, Yash}, {Bob, Johannes}, {Yash, Johannes}}
```

This should be 6 ways or Binomial [4,2]

```
In[10]:= Binomial[4, 2]
Out[10]= 6
```

To randomly pick from these 6 choices we can use RandomChoice which picks an element from a given list.

```
In[13]:= RandomChoice[chooseboys]
Out[13]= {Yash, Johannes}
```

Some probability density functions.

Distributions can be specified by a name and its parameters.

BinomialDistribution[n, p] where n is the number of elements and p is the probability of choosing an element. The random variable k is the number of chosen elements. PDF[BinomialDistribution[n, p], k] is the function of k. We can make a plot of this function

```
In[14]:= PDF[BinomialDistribution[n, p], k]
```

Out[14]=

$$\begin{cases} (1-p)^{-k+n} p^k \text{Binomial}[n, k] & 0 \le k \le n \\ 0 & \text{True} \end{cases}$$

```
In[26]:= Table[{k, PDF[BinomialDistribution[5, 1/2], k]}, {k, 0, 5}] // TableForm
```

Out[26]//TableForm=

0	$\frac{1}{32}$
1	$\frac{5}{32}$
2	$\frac{5}{16}$
3	$\frac{5}{16}$
4	$\frac{5}{32}$
5	$\frac{1}{32}$

In[32]:=
```
? ListPlot
```

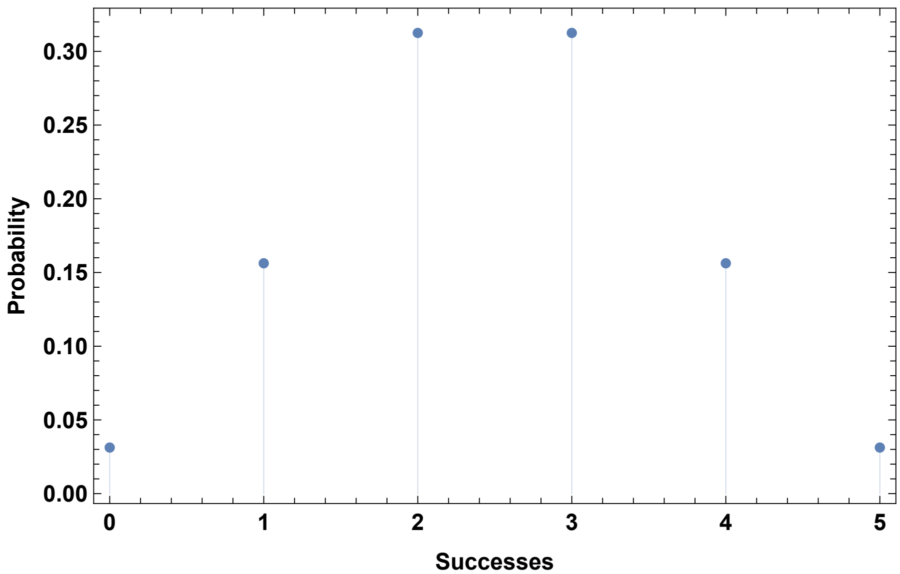
ListPlot[{y_1, y_2, …}] plots points {1, y_1}, {2, y_2}, ….

ListPlot[{{x_1, y_1}, {x_2, y_2}, …}] plots a list of points with specified x and y coordinates.

ListPlot[{$data_1$, $data_2$, …}] plots data from all the $data_i$.

ListPlot[{…, w[$data_i$, …], …}] plots $data_i$ with features defined by the symbolic wrapper w. >>

In[37]:=
```
ListPlot[Table[{k, PDF[BinomialDistribution[5, 1/2], k]}, {k, 0, 5}],
 Frame → True, FrameLabel → {"Successes", "Probability"},
 Filling → Axis, FrameStyle → Bold, LabelStyle → Directive[Bold, Medium]]
```

Out[37]=


Some Continuous probability density functions.
Normal of Gaussian distribution is

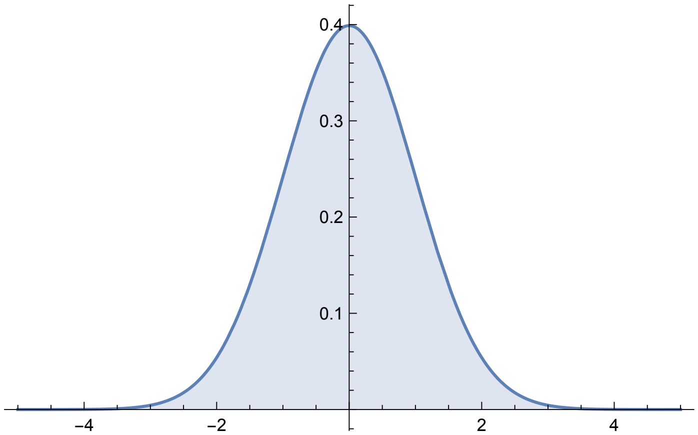
NormalDistribution[mean, sigma]

In[28]:=
```
PDF[NormalDistribution[0, 1], x]
```

Out[28]=
$$\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

In[29]:=
```
Plot[PDF[NormalDistribution[0, 1], x], {x, -5, 5}, Filling → Axis]
```

Out[29]=



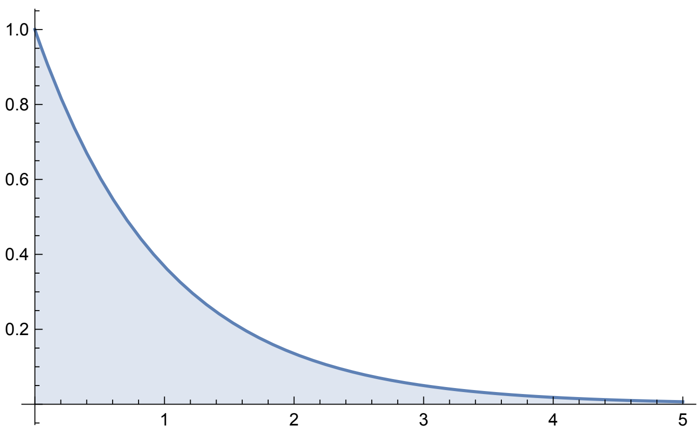
In[41]:=
```
PDF[ExponentialDistribution[1], x]
```

Out[41]=

$$\begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{True} \end{cases}$$

In[44]:=
```
Plot[PDF[ExponentialDistribution[1], x], {x, 0, 5}, Filling → Axis]
```

Out[44]=



What is the chance that none of you 70 people have the same birthday ?
First person has some birthday out of d=365 days
Chance that second person does not share the birth of first is: (365-1)/365
(d-1)/d
Chance that third person does not share with first two is: (d-2)/d
etc.

In[64]:=
```
q1 = Product[(d - i) / d, {i, 1, 69}]
```

Out[64]=
$$\frac{1}{d^{69}} (-69+d)(-68+d)(-67+d)(-66+d)(-65+d)(-64+d)(-63+d)(-62+d)(-61+d)(-60+d)(-59+d)(-58+d)(-57+d)(-56+d)(-55+d)(-54+d)(-53+d)(-52+d)(-51+d)(-50+d)(-49+d)(-48+d)(-47+d)(-46+d)(-45+d)(-44+d)(-43+d)(-42+d)(-41+d)(-40+d)(-39+d)(-38+d)(-37+d)(-36+d)(-35+d)(-34+d)(-33+d)(-32+d)(-31+d)(-30+d)(-29+d)(-28+d)(-27+d)(-26+d)(-25+d)(-24+d)(-23+d)(-22+d)(-21+d)(-20+d)(-19+d)(-18+d)(-17+d)(-16+d)(-15+d)(-14+d)(-13+d)(-12+d)(-11+d)(-10+d)(-9+d)(-8+d)(-7+d)(-6+d)(-5+d)(-4+d)(-3+d)(-2+d)(-1+d)$$

In[65]:=
```
q1 /. d → 365 // N // ScientificForm
```

Out[65]//ScientificForm=
8.40424×10^{-4}

This should be the same as (which is very small)

In[66]:=
```
Q1[n_, d_] := d ! / ((d - n) ! * d^n)
```

In[67]:=
```
Q1[70, 365] // N
```

Out[67]= 0.000840424

What is the chance that out of 10 people 2 or more share a birthday. It is amazingly large. Nearly 12 %.

In[70]:=
```
p1 = 1 - Q1[5, 365] // N
```

Out[70]= 0.0271356

Statistics is the subject of extracting information about the parameters of the underlying probability distribution from data.

We can put data into Mathematica and do simple and very advanced analysis on it.

Let’s first create some data.

In[79]:=
```
mydata = {2, 1, 3, 5, 3.2, 1.1, 5.6, 2, 1., 1.1, 3.6, 9.4, 5., 5.1, 5.4, 4.9};
```

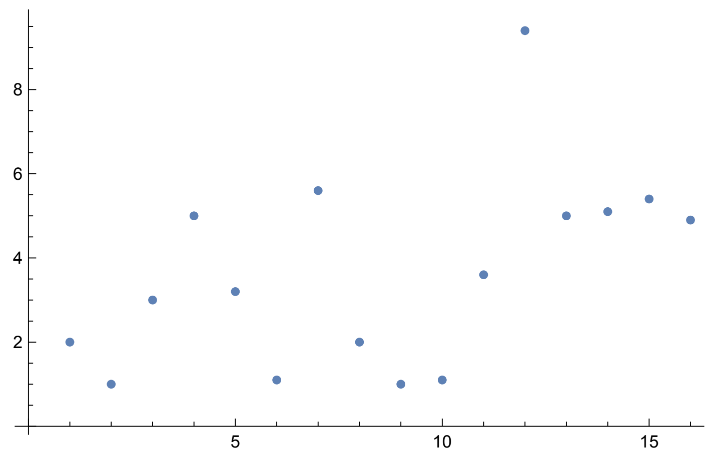
In[80]:=
```
Mean[mydata]
```

Out[80]= 3.65

In[81]:=
```
StandardDeviation[mydata]
```

Out[81]= 2.31459

In[82]:=
```
ListPlot[mydata] (* makes a plot of the data versus its index number *)
```

Out[82]=


We can use Histogram function to plot the frequency of data. How many in each bin.
Histogram[data, bins]

In[78]:=
```
? Histogram
```

Histogram[{$x_1, x_2, \ldots$}] plots a histogram of the values x_i.

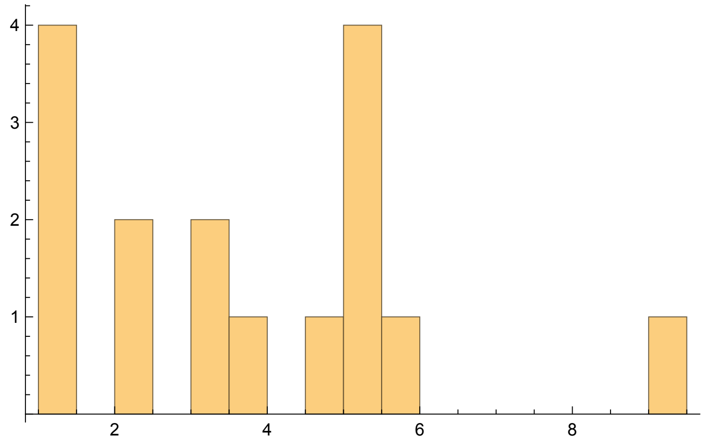
Histogram[{$x_1, x_2, \ldots$}, *bspec*] plots a histogram with bin width specification *bspec*.

Histogram[{$x_1, x_2, \ldots$}, *bspec*, *hspec*] plots a histogram with bin heights computed according to the specification *hspec*.

Histogram[{$data_1, data_2, \ldots$}, …] plots histograms for multiple datasets $data_i$. >>

In[87]:=
```
Histogram[mydata, 20]
```

Out[87]=



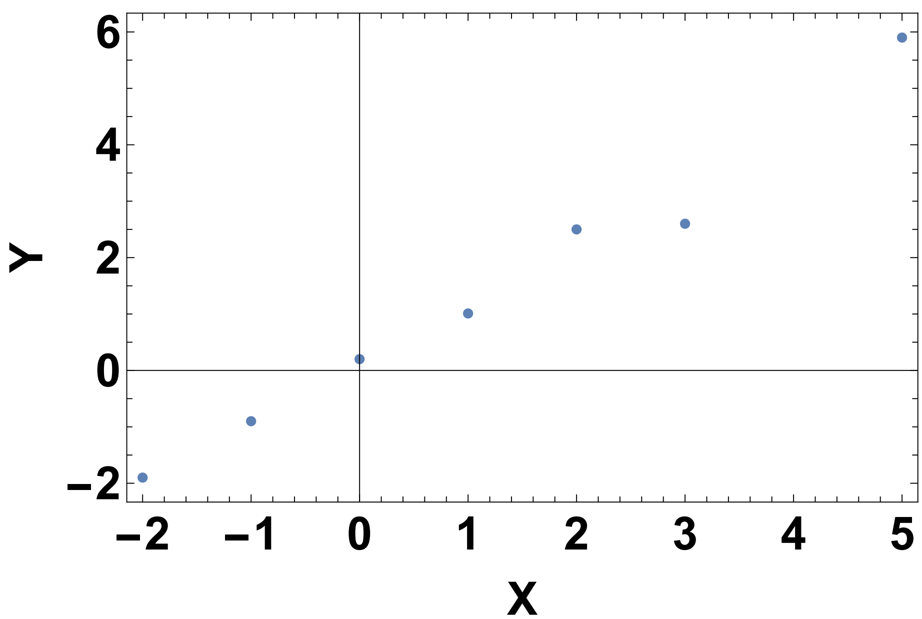
How about pairs of data points {x,y}

In[88]:=
```
pairdata =
  {{-2., -1.9}, {-1., -0.9}, {0., 0.2}, {1, 1.01}, {2., 2.5}, {3, 2.6}, {5, 5.9}}
```

Out[88]=
```
{{-2., -1.9}, {-1., -0.9}, {0., 0.2}, {1, 1.01}, {2., 2.5}, {3, 2.6}, {5, 5.9}}
```

In[99]:=
```
dataplot = ListPlot[pairdata, Frame → True,
   FrameLabel -> {"X", "Y"}, FrameStyle → Bold, LabelStyle → Large]
```

Out[99]=



How about making a linear fit to this data ?

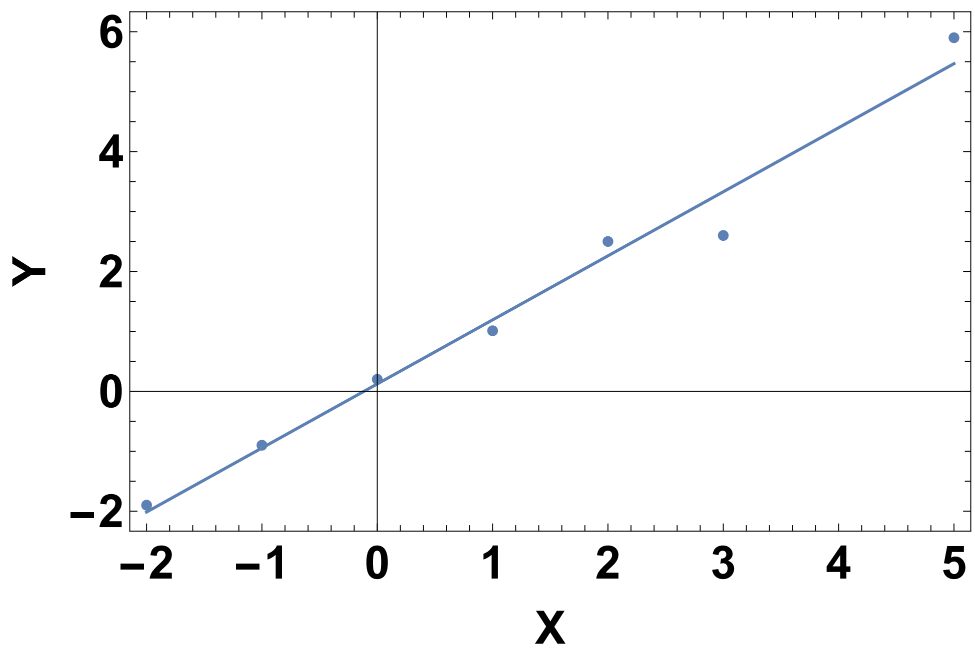
In[95]:=
```
? Fit
```

Fit[*data*, *funs*, *vars*] finds a least-squares fit to a list
of data as a linear combination of the functions *funs* of variables *vars*. >>

In[98]:=
```
myfit = Fit[pairdata, {1, x}, x]
```

Out[98]=
```
0.122787 + 1.06881 x
```

In[100]:=
```
Show[dataplot, Plot[myfit, {x, -2, 5}]]
```

Out[100]=


This is the end of the tutorial on Mathematica. It was meant to get you comfortable with the system.

The best way to use this tool is by trying simple things first. Use the
website to learn about functions. Most simple problems have been solved by someone and the solution is

available on the web.

Complicated problems can be pieced together by breaking them up in simple ones. Setup each simple one and make sure it works, and then move on to put it all together.