


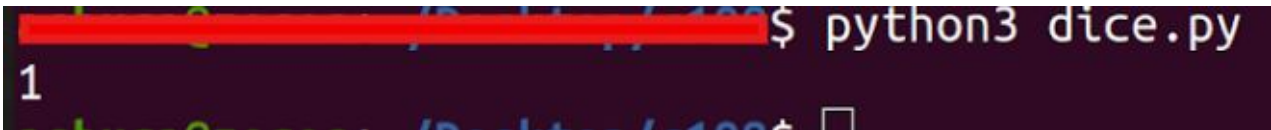


Topic	Bell Curve- the Normal Distribution	
Class Description	Student creates and plots different kinds of data to identify patterns in how data is distributed. Student is introduced to the normal distribution and its omni-presence in the nature.	
Class	C108	
Class time	45 mins	
Goal	<ul style="list-style-type: none"> • Student plots data from random dice throws • Student identifies the bell curve pattern in the data • Student plots other kinds of data and identifies similar bell-curve patterns 	
Resources Required	<ul style="list-style-type: none"> • Teacher Resources <ul style="list-style-type: none"> ◦ VCS Code ◦ Laptop with internet connectivity ◦ Earphones with mic ◦ Notebook and pen • Student Resources <ul style="list-style-type: none"> ◦ VCS Code ◦ Laptop with internet connectivity ◦ Earphones with mic ◦ Notebook and pen 	
Class structure	Warm Up Teacher-led Activity Student-led Activity Wrap up	5 mins 15 min 15 min 5 min
<u>CONTEXT</u> <ul style="list-style-type: none"> • Can random data have patterns in them? 		
Class Steps	Teacher Action	Student Action

Step 1: Warm Up (5 mins)	Hi <student name> We have worked with different kinds of data. Each piece of data could be anything, it could hold any value. But do you think all these data could have a pattern where they fit? Allow the student to think.	Student thinks about the question
	We'll be looking at a pattern today which holds universally true for almost all kinds of data in some way.	-
Teacher Initiates Screen Share		
<p style="text-align: center;"><u>CHALLENGE</u></p> <ul style="list-style-type: none"> • Collect data from throwing two dices • Plot the data as a bar graph and identify the bell-curve in the distribution • Plot the data of heights of individuals and identify the bell curve in the distribution 		
Step 2: Teacher-led Activity (15 min)	Let's go to Teacher Activity 1 (https://www.random.org/dice/?num=2). What do you see?	There are two dices and we can roll them up.

<div> Home Games Numbers Lists & More Drawings Web Tools Statistics Testimonials Learn More Login </div> <div> <h1>RANDOM.ORG</h1> <div> <input type="text"/> <input type="button" value="Search"/> </div> <div> <input type="button" value="Search"/> True Random Number Service </div> </div> <div> <h2>Dice Roller</h2> <p>You rolled 2 dice:</p> <div>   </div> <p>Timestamp: 2020-07-02 06:59:29 UTC</p> <div> <input type="button" value="Roll Again"/> <input type="button" value="Go Back"/> </div> <p>Does this not look random to you? Check out the FAQ.</p> <div> <p>© 1998-2020 RANDOM.ORG</p> <p>Follow us: Twitter Facebook</p> <p>Terms and Conditions</p> <p>About Us</p> </div> </div>		
	<p>Let us roll these dices and add the numbers on the two dice. Let's record the different sums we get on a file.</p> <p>8 3 2 ... 9</p> <p>Teacher and student rolls the dice for sometime and they record the data on a separate file</p>	<p>Student helps the teacher add the numbers on the dice and record them.</p>
	<p>If we do this, let's say 100 times, do you think there will be some pattern in the numbers we get?</p>	<p>ESR: Student thinks about it.</p>
	<p>Let's see... But wait...we are programmers. We don't need to roll the dice 100 times. We can write a python program which randomly generates two numbers between 1 to 6 for two different dice</p>	<p>ESR: Yes</p>

	<p>and adds them up.</p> <p>Can you help me write such a program?</p>	
	<p>Python has a random module which helps generate random number. Let's generate one random number between 1 to 6 and store it in a variable called dice1.</p> <p>random.randint() can be used to generate random number....you can pass the minimum number and maximum number between which you want to generate random values</p> <p>We can print the value.</p> <p>Teacher writes the code and runs it to show output to the student.</p>	<p>Student helps the teacher write the code to generate a random dice number for dice 1</p>
 		
	<p>We can easily write code to generate another random number between 1 and 6 for dice 2</p>	<p>Student helps the teacher write the code to generate a random dice number for dice 2</p>

```
import random
dice1 = random.randint(1, 6)
dice2 = random.randint(1, 6)
print(dice1 ,dice2)
```

Let's say we want to do this 100 times. What would we do?

ESR:
We can use a loop - for loop

Help me write code for that.
Teacher writes code to generate 100 random dice numbers for dice 1 and dice 2.

Teacher can re-iterate the of
range(0,100)
range(0,100) generates a list of numbers from 0 to 100 and iterates over the list.

Teacher can also print the numbers each time to show the results

Student can help the teacher write the code.

```
import random

dice_result = []
for i in range(0, 100):
    dice1 = random.randint(1, 6)
    dice2 = random.randint(1, 6)
    dice_result.append(dice1 + dice2)
```

	<p>We want to store the sum of numbers on the dice each time the dice is rolled.</p> <p>Let us create a list which would store the sum of numbers on the dice.</p> <p>We will append the sum to this list, each time the dice is rolled.</p> <p>Teacher writes code for this.</p>	<p>Student helps the teacher write the code while verbally articulating it.</p>
<pre>import random dice_result = [] for i in range(0, 100): dice1 = random.randint(1, 6) dice2 = random.randint(1, 6) dice_result.append(dice1 + dice2)</pre>		
	<p>Let us use plotly express to draw a histogram or bar graph for the data we have and see how the data is distributed.</p> <p>Teacher writes code to draw a histogram for the data with x axis containing the dice rolls and y axis containing the count.</p>	<p>Student helps the teacher with the code.</p>

```
import random
import plotly.express as px
count = []
dice_result = []
for i in range(0, 100):
    dice1 = random.randint(1, 6)
    dice2 = random.randint(1, 6)
    dice_result.append(dice1 + dice2)
    count.append(i)

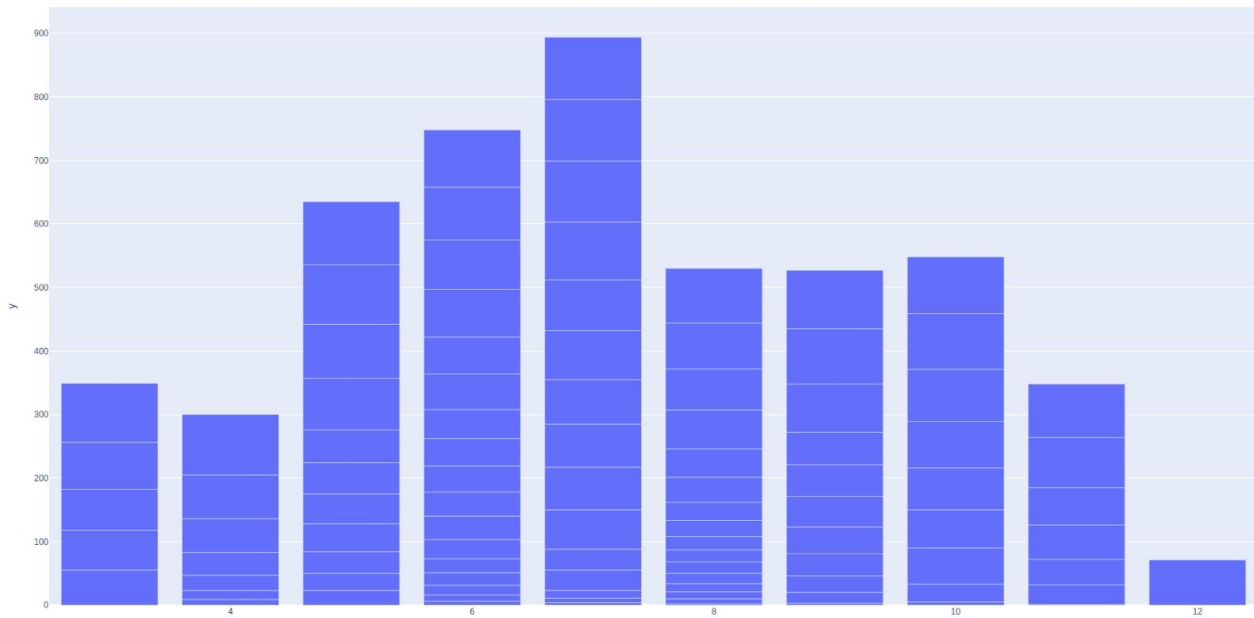
fig = px.bar(x=dice_result, y=count)
fig.show()
```

Let us run the code and look at how the distribution looks like.
Teacher runs the code.

What do you see?

Student observes the output.

ESR:
There are more values in the centre. The higher values and lower values are less.



If you could join all the left edges of the bar graph, you would see that the distribution of numbers is almost like a bell-shape.

This bell shape distribution is very common in nature. This is why it is called "normal distribution". Most of the different kinds of data in the universe follow this pattern. In fact, we expect different kinds of data to follow this pattern.

Student observes the bell curve pattern which the data follows.

We can also draw a distribution plot directly using plotly's figure_factory module.
We will need to additionally install scipy package on our system.

Student observes how to create distribution plots.

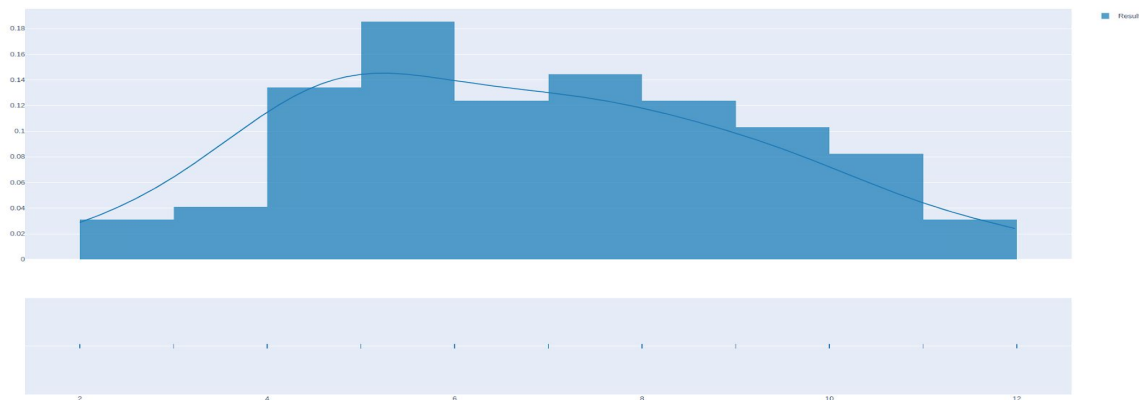
Teacher installs scipy using pip install scipy

Teacher imports plotly's figure_factory and uses distplot() function to draw the distribution graph for the data. It takes two arguments - the data and label for the data.

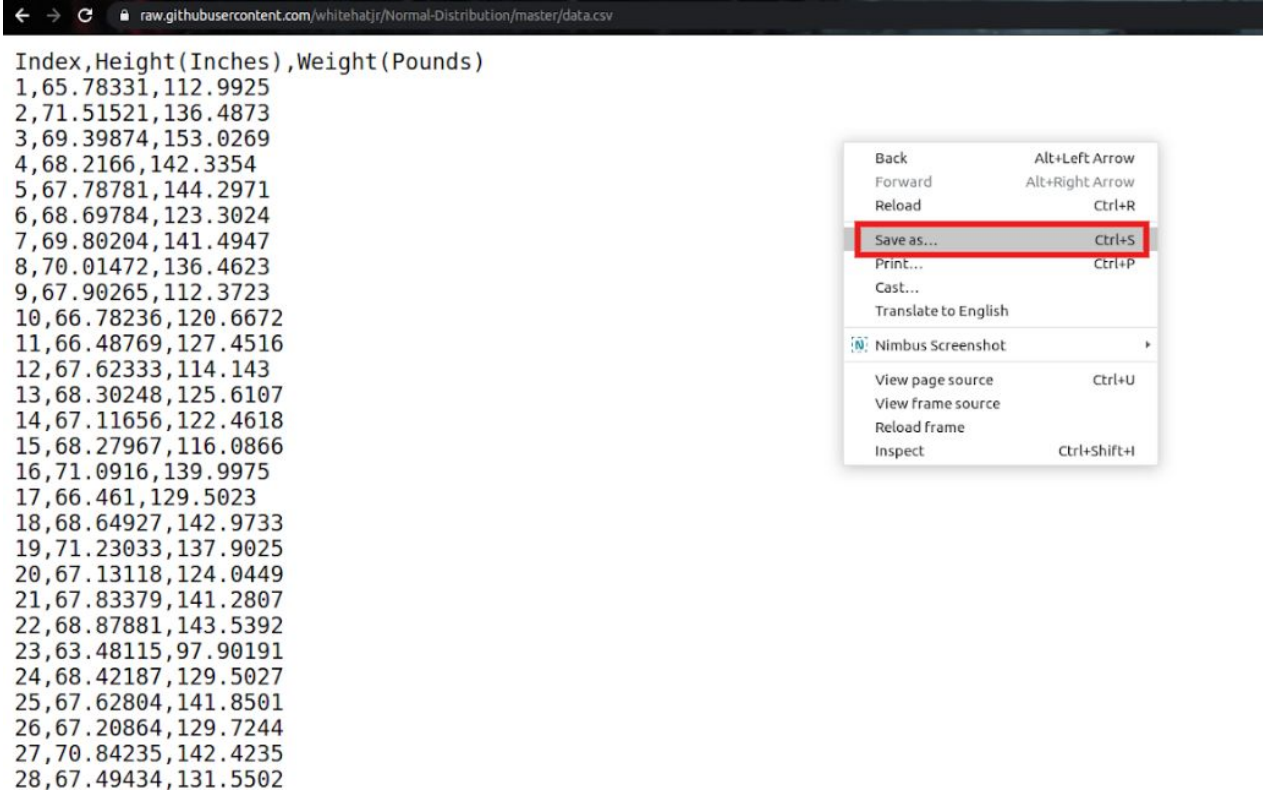
Teacher writes code to make the distribution plot.

Teacher also shows the bell curve line in the data.

```
1  import plotly.figure_factory as ff
2  import random
3
4  dice_result = []
5  for i in range(0, 100):
6      dice1 = random.randint(1, 6)
7      dice2 = random.randint(1, 6)
8      dice_result.append(dice1 + dice2)
9
10 fig = ff.create_distplot([dice_result], ["Result"])
11 fig.show()
12
13
```



	<p>Remember, the heights and weights data for 18 year olds which we used in our earlier class.</p> <p>What would be the distribution of heights and weights in these 18 year olds ? Do you think they will follow a pattern?</p> <p>Do you think more 18 year olds will be at the centre in the distribution and less people having very low heights/weights and very large heights / weights?</p> <p>What do you think will be the x-axis of the distribution?</p> <p>What do you think will be the y-axis of the distribution?</p>	<p>ESR:</p> <p>The distribution might be a bell curve</p> <p>There will be more 18 year olds having average height and less number of 18 year olds too short or too tall.</p> <p>x axis would be the weight / height.</p> <p>y axis would be the number of people having specific weights and heights</p>
	Let us try to plot the data distribution and check if this is actually the case.	-
Teacher Stops Screen Share		
	Now it's your turn. Please share your screen with me.	
<ul style="list-style-type: none"> • Ask Student to press ESC key to come back to panel • Guide Student to start Screen Share • Teacher gets into Fullscreen 		
<p style="text-align: center;"><u>ACTIVITY</u></p> <ul style="list-style-type: none"> • Plot the data of weights of individuals and identify the bell curve in the distribution. 		

Step 3: Student-Led Activity (15 min)	Guide the student to download the weights/ heights data for 18 year olds and move it to their working directory. Guide the student to start a new python project/ file.	Student downloads the data file and starts a new project. <Student downloads the data from student activity 1>
 <pre> Index,Height (Inches),Weight (Pounds) 1,65.78331,112.9925 2,71.51521,136.4873 3,69.39874,153.0269 4,68.2166,142.3354 5,67.78781,144.2971 6,68.69784,123.3024 7,69.80204,141.4947 8,70.01472,136.4623 9,67.90265,112.3723 10,66.78236,120.6672 11,66.48769,127.4516 12,67.62333,114.143 13,68.30248,125.6107 14,67.11656,122.4618 15,68.27967,116.0866 16,71.0916,139.9975 17,66.461,129.5023 18,68.64927,142.9733 19,71.23033,137.9025 20,67.13118,124.0449 21,67.83379,141.2807 22,68.87881,143.5392 23,63.48115,97.90191 24,68.42187,129.5027 25,67.62804,141.8501 26,67.20864,129.7244 27,70.84235,142.4235 28,67.49434,131.5502 </pre>		
	Guide the student to read the data from the csv file.	Student reads the csv data using pandas read_csv function

```
import pandas as pd
import csv

df = pd.read_csv("data.csv")
```

You can use the create_distplot function which we just discovered to plot how data is distributed in heights.

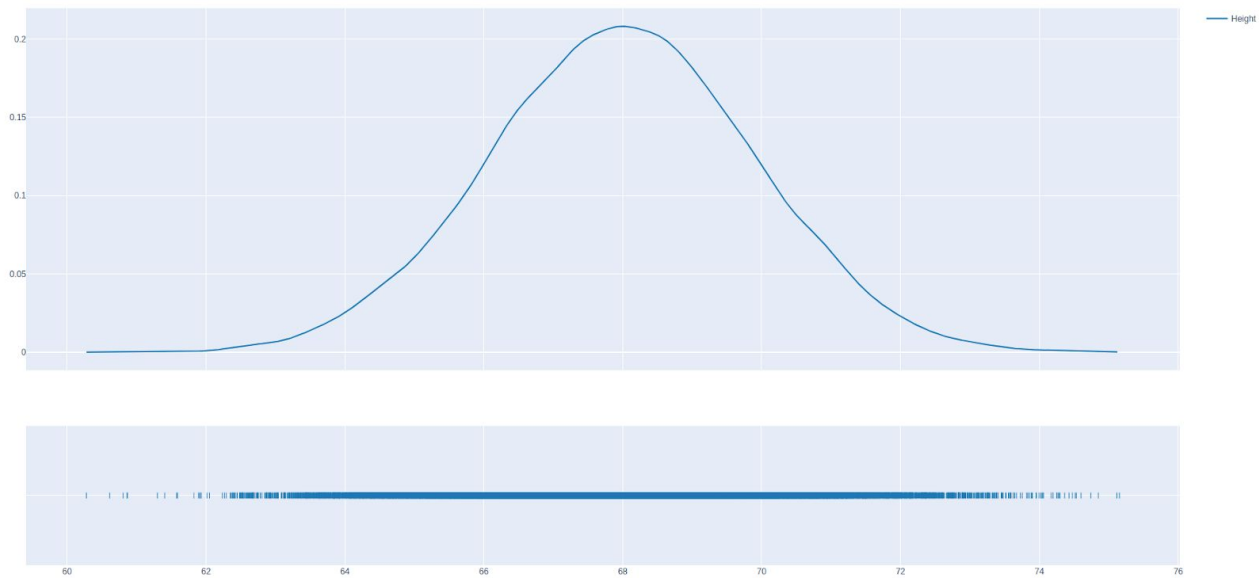
Student uses the create_distplot function to create a distribution graph for how height is distributed.
Student runs the code to see the output.

```
import plotly.figure_factory as ff
import pandas as pd
import csv

df = pd.read_csv("data.csv")
fig = ff.create_distplot([df["Height(Inches)"].tolist()], ["Height"], show_hist=False)
fig.show()
```

What do you observe?

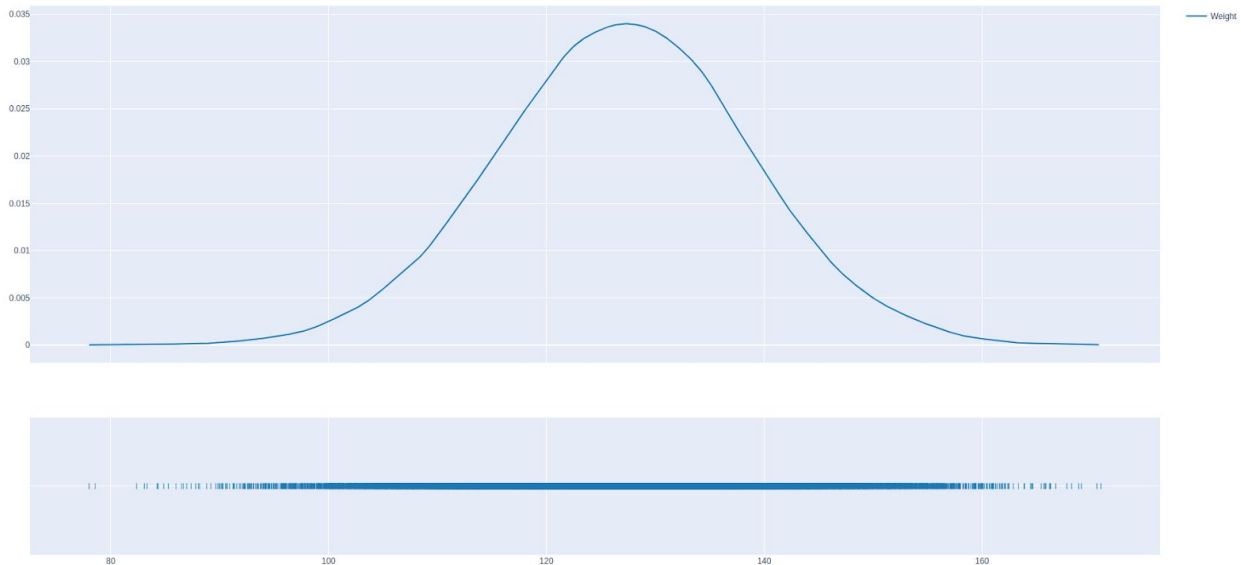
ESR:
We observe that the height is distributed in a bell-shaped curve.



Do you want to draw the distribution for weights in a similar way?

ESR:
Student draws the distribution for weights of 18 year olds.
Student writes the code and runs it to see the output.

```
1 import plotly.figure_factory as ff
2 import pandas as pd
3 import csv
4
5
6 df = pd.read_csv("data.csv")
7 fig = ff.create_distplot([df["Weight(Pounds)"].tolist()], ["Weight"], show_hist=False)
8 fig.show()
9
```



Ask the student to make observations

Student observes the bell-curve pattern for the data too

Teacher Guides Student to Stop Screen Share

FEEDBACK

- Appreciate the student for their efforts
- Identify 2 strengths and 1 area of progress for the student

Step 4:
Wrap-Up
(5 min)

What did you learn today?

ESR:
We learned how to draw distribution plots.
We also learned how data is normal distributed and follows a bell curve.

	<p>The normal distribution or bell curve pattern is universally present in nature. Most data distributions follow this pattern.</p> <p>In fact, we expect data to follow the normal distribution.</p> <p>We will be learning how this is a very important in statistical analysis. Knowing how data will be distributed also gives us a chance to predict what the new data could be.</p> <p>This will be very important in machine learning and AI algorithms.</p> <p>We will be learning amore about using normal data distributions in the coming classes.</p>	-
<div> <div>Teacher Clicks</div> <div>✕ End Class</div> </div>		
Additional Activities	<p>Encourage the student to write reflection notes in their reflection journal using markdown.</p> <p>Use these as guiding questions:</p> <ul style="list-style-type: none"> • What happened today? <ul style="list-style-type: none"> - Describe what happened - Code I wrote • How did I feel after the class? • What have I learned about programming and developing games? 	<p>The student uses the markdown editor to write her/his reflection in a reflection journal.</p>

	<ul style="list-style-type: none"> What aspects of the class helped me? What did I find difficult? 	
--	---	--

Activity	Activity Name	Links
Teacher Activity 1	Dice roller site	https://www.random.org/dice/?num=2
Teacher Activity 2	Solution	https://github.com/whitehatjr/Normal-Distribution
Student Activity 1	raw data	https://raw.githubusercontent.com/whitehatjr/Normal-Distribution/master/data.csv