BOOSTING

- AdaBoost
- Gradient Boost
- XGBoost

GRADIENT BOOST

Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88
1.6	Green	Female	76
etc…	etc…	etc…	etc…

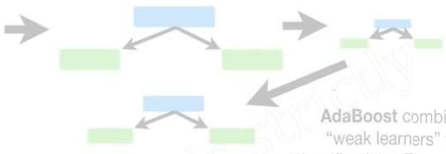
we want to use these measurements to **Predict Weight**…

Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88
1.6	Green	Female	76
1.5	Blue	Female	56
1.8	Red	Male	73
1.5	Green	Male	77
1.4	Blue	Female	57

NOTE: When **Gradient Boost** is used to **Predict** a continuous value, like **Weight**, we say that we are using **Gradient Boost** for **Regression**.

Using **Gradient Boost** for **Regression** is different from doing **Linear Regression**, so while the two methods are related, don't get them confused with each other.

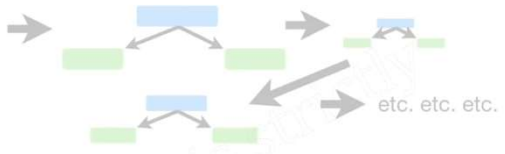
Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88
1.6	Green	Female	76
etc…	etc…	etc…	etc…



AdaBoost combines a lot of "weak learners" to make classifications. The weak learners are almost aways **stumps**.

In contrast, **Gradient Boost** starts by making a single leaf, instead of a tree or stump.

Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88
1.6	Green	Female	76
etc…	etc…	etc…	etc…



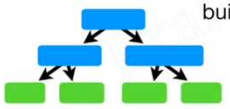
This leaf represents an initial guess for the **Weights** of all of the samples.

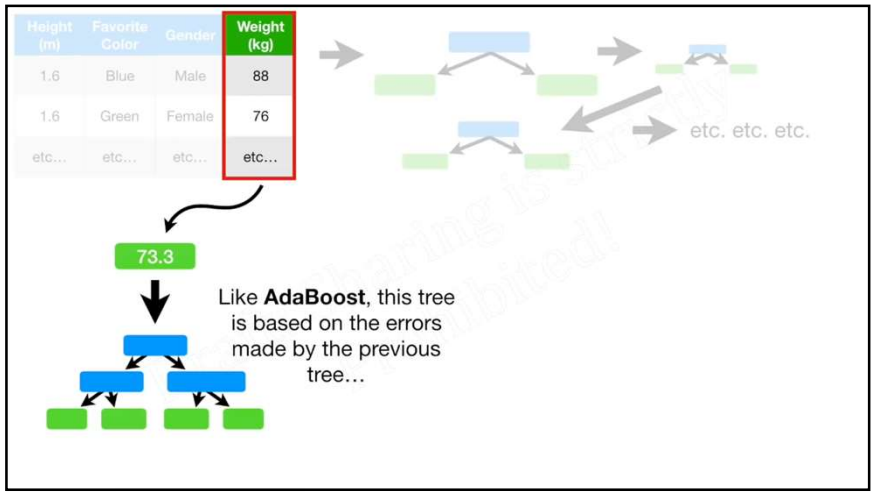
Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88
1.6	Green	Female	76
etc…	etc…	etc…	etc…

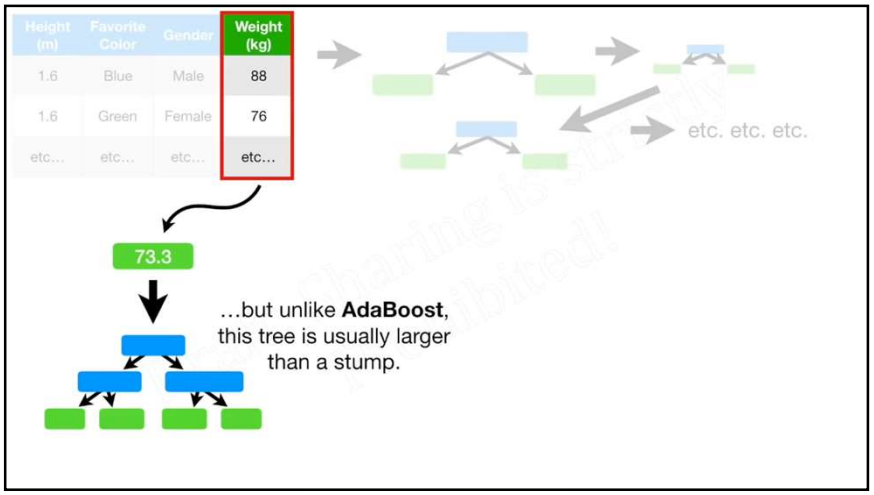
When trying to **Predict** a continuous value like **Weight**, the first guess is the the average value.

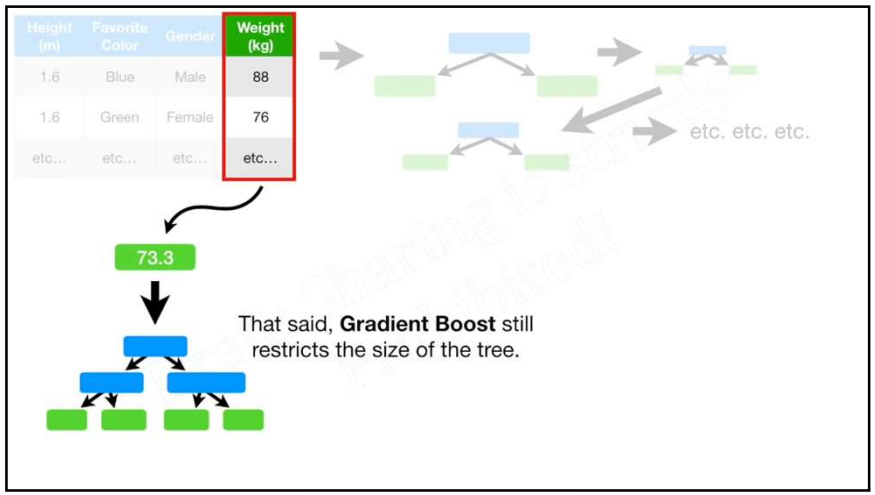
Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88
1.6	Green	Female	76
etc…	etc…	etc…	etc…

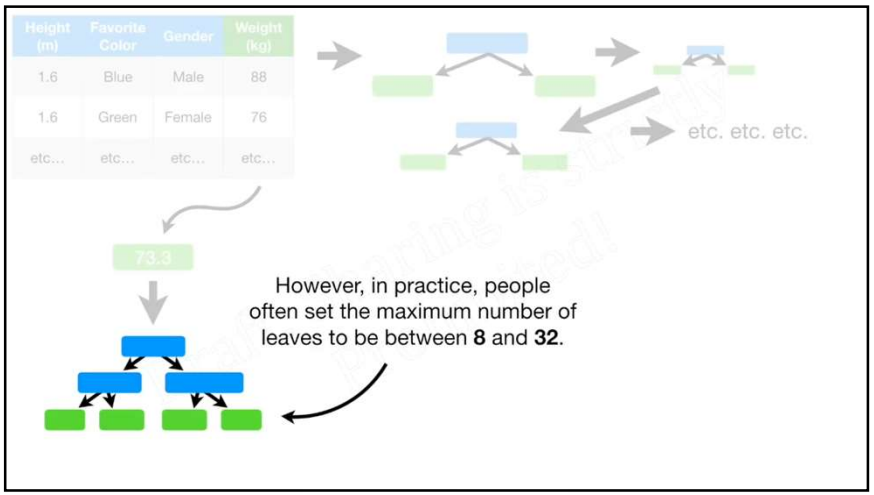
Then **Gradient Boost** builds a tree.

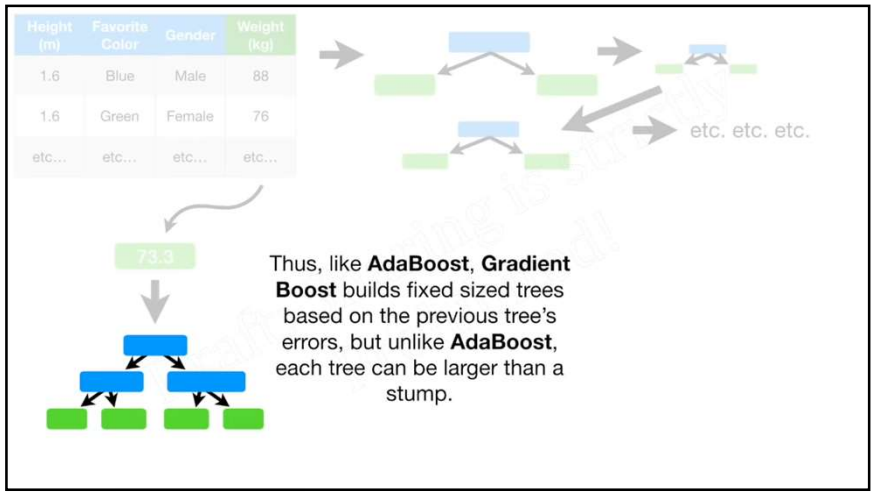


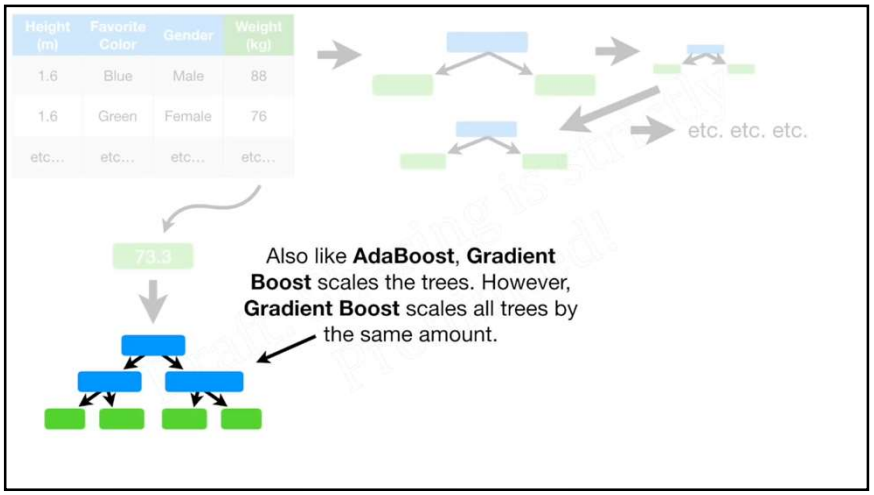

Height (m)
Favorite Color
Gender
Weight (kg)
1.6 Blue Male 88
1.6 Green Female 76
etc... etc... etc... etc...
etc. etc. etc.
73.3
Like AdaBoost, this tree is based on the errors made by the previous tree…

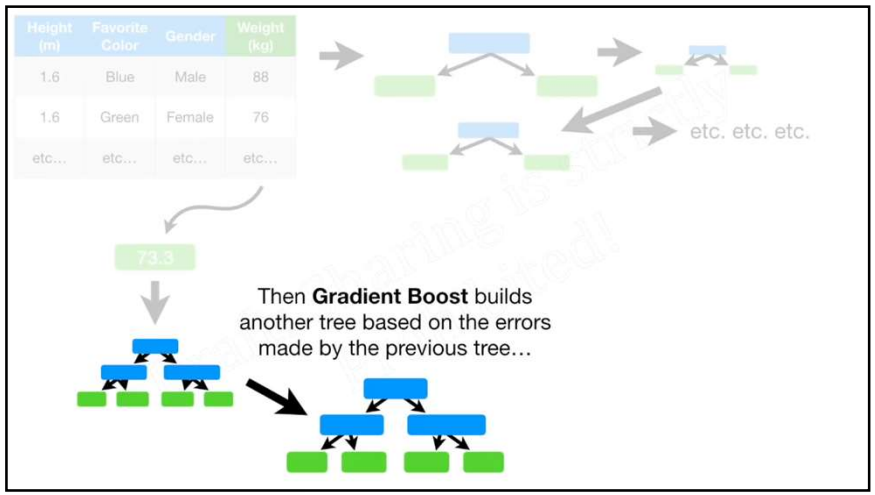

Height (m)
Favorite Color
Gender
Weight (kg)
1.6 Blue Male 88
1.6 Green Female 76
etc... etc... etc... etc...
etc. etc. etc.
73.3
…but unlike AdaBoost, this tree is usually larger than a stump.

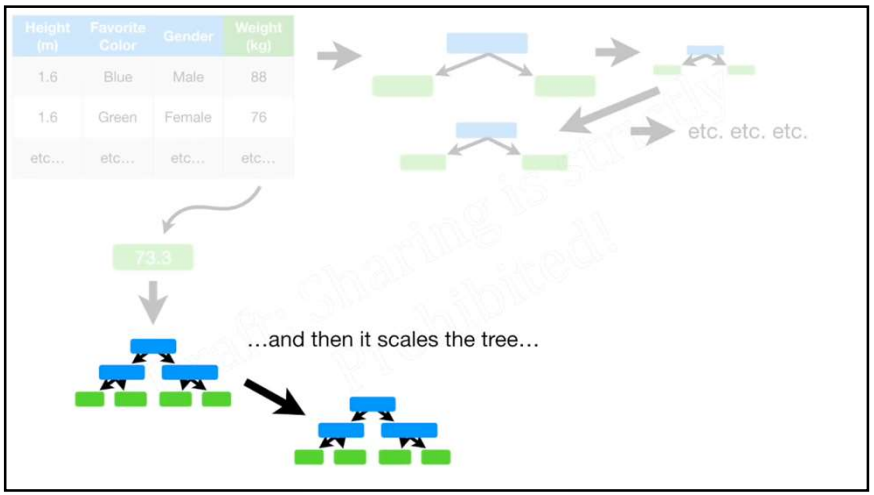

Height (m)
Favorite Color
Gender
Weight (kg)
1.6 Blue Male 88
1.6 Green Female 76
etc... etc... etc... etc...
etc. etc. etc.
73.3
That said, Gradient Boost still restricts the size of the tree.

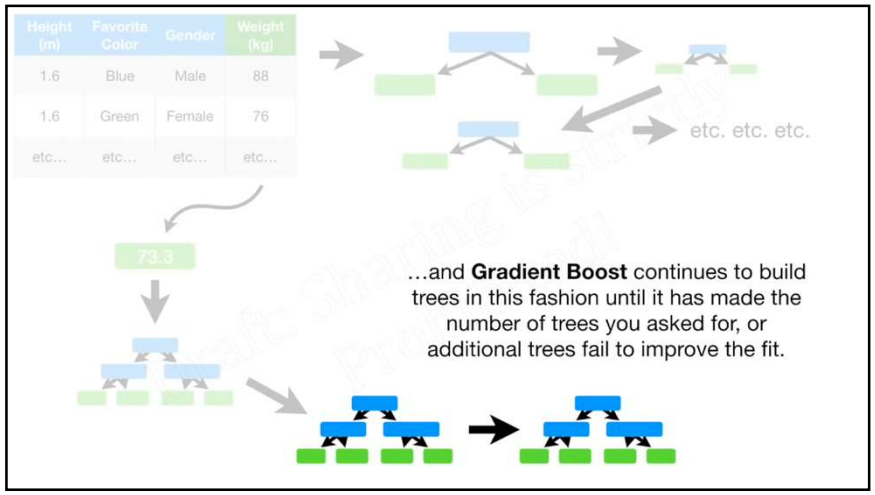

Height (m)
Favorite Color
Gender
Weight (kg)
1.6 Blue Male 88
1.6 Green Female 76
etc... etc... etc... etc...
etc. etc. etc.
73.3
However, in practice, people often set the maximum number of leaves to be between 8 and 32.

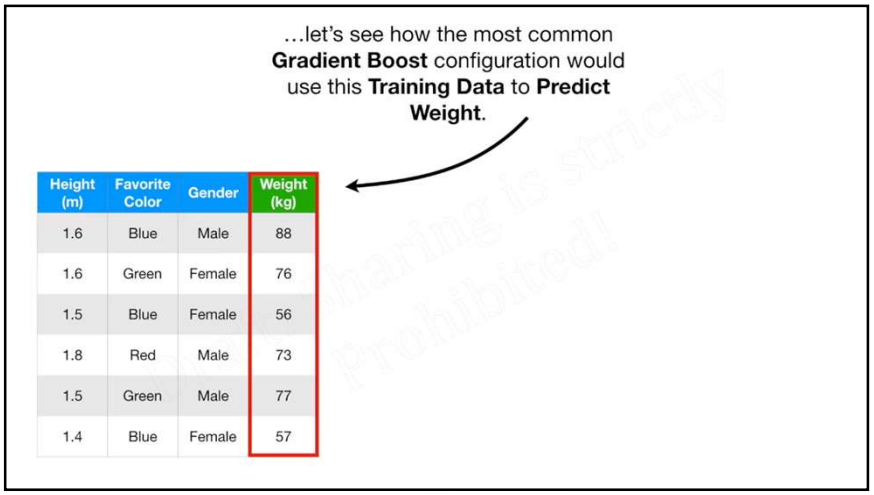

Height (m)
Favorite Color
Gender
Weight (kg)
1.6 Blue Male 88
1.6 Green Female 76
etc... etc... etc... etc...
etc. etc. etc.
73.3
Thus, like **AdaBoost**, **Gradient Boost** builds fixed sized trees based on the previous tree's errors, but unlike **AdaBoost**, each tree can be larger than a stump.

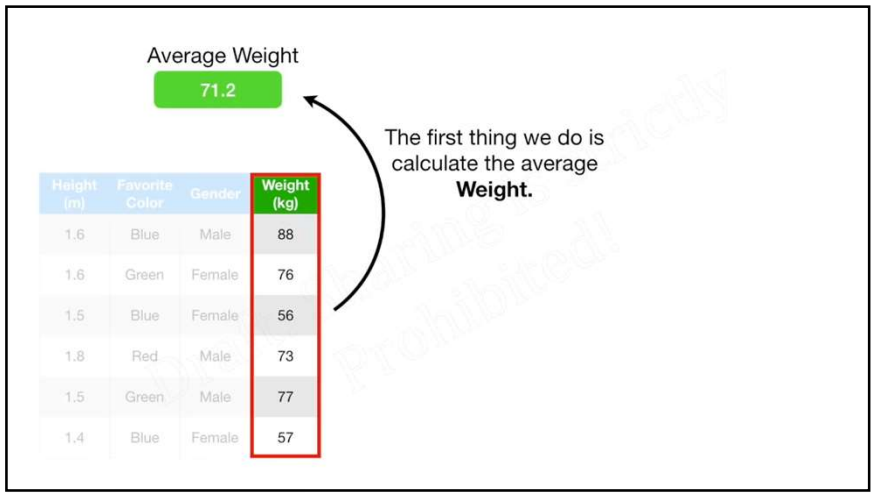

Height (m)
Favorite Color
Gender
Weight (kg)
1.6 Blue Male 88
1.6 Green Female 76
etc... etc... etc... etc...
etc. etc. etc.
73.3
Also like **AdaBoost**, **Gradient Boost** scales the trees. However, **Gradient Boost** scales all trees by the same amount.

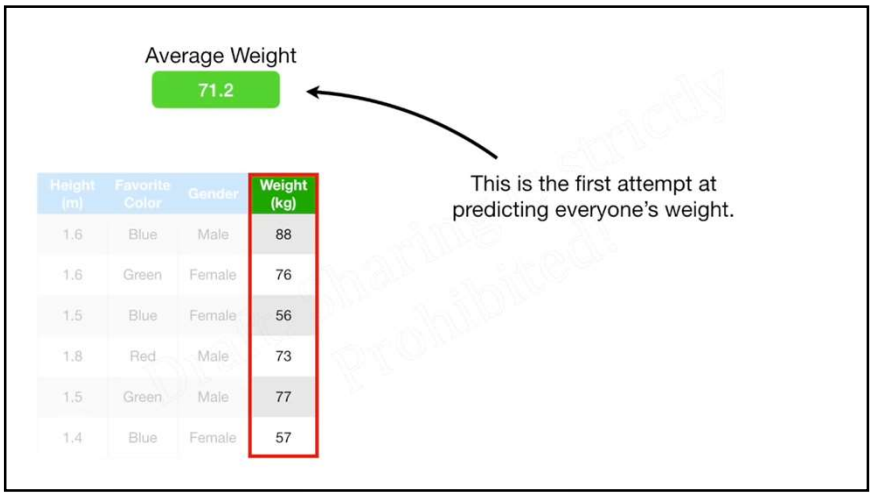

Height (m)
Favorite Color
Gender
Weight (kg)
1.6 Blue Male 88
1.6 Green Female 76
etc... etc... etc... etc...
etc. etc. etc.
73.3
Then **Gradient Boost** builds another tree based on the errors made by the previous tree…

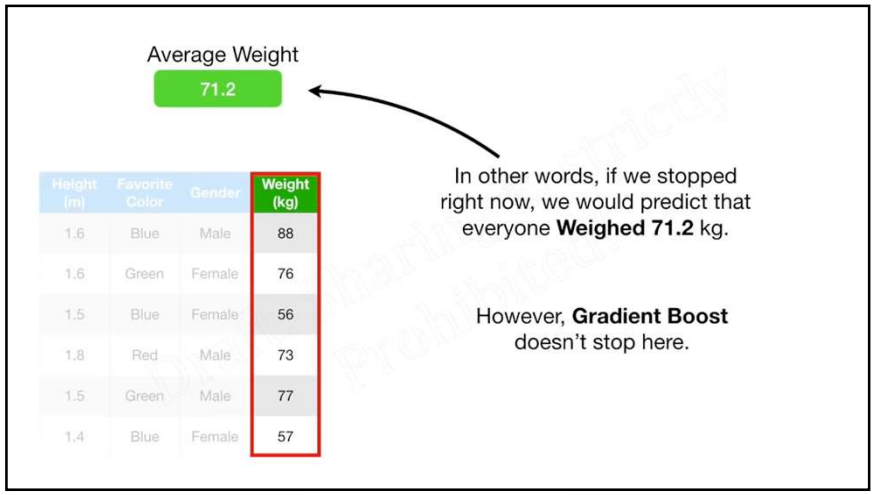

Height (m)
Favorite Color
Gender
Weight (kg)
1.6 Blue Male 88
1.6 Green Female 76
etc... etc... etc... etc...
etc. etc. etc.
73.3
…and then it scales the tree…


Height (m)
Favorite Color
Gender
Weight (kg)
1.6 Blue Male 88
1.6 Green Female 76
etc... etc... etc... etc...
etc. etc. etc.
73.3
...and **Gradient Boost** continues to build trees in this fashion until it has made the number of trees you asked for, or additional trees fail to improve the fit.

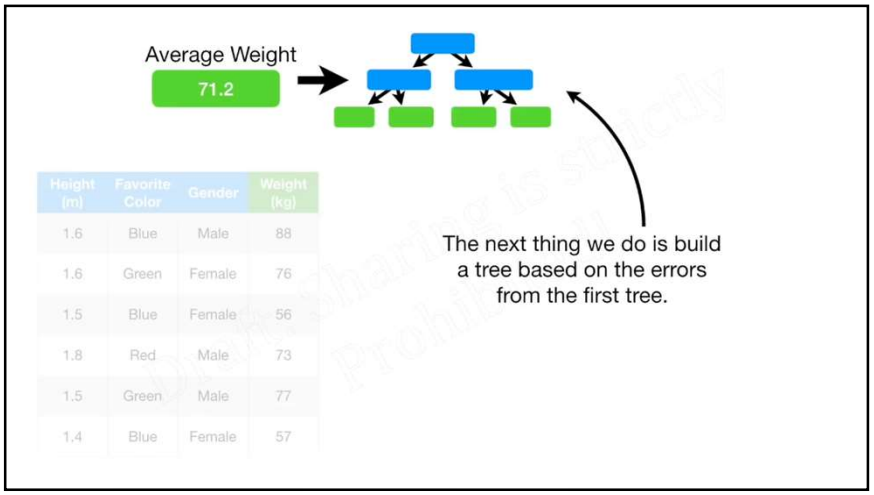

...let's see how the most common **Gradient Boost** configuration would use this **Training Data** to **Predict Weight**.
Height (m) | Favorite Color | Gender | Weight (kg)
1.6 | Blue | Male | 88
1.6 | Green | Female | 76
1.5 | Blue | Female | 56
1.8 | Red | Male | 73
1.5 | Green | Male | 77
1.4 | Blue | Female | 57


Average Weight
71.2
The first thing we do is calculate the average **Weight.**
Height (m) | Favorite Color | Gender | Weight (kg)
1.6 | Blue | Male | 88
1.6 | Green | Female | 76
1.5 | Blue | Female | 56
1.8 | Red | Male | 73
1.5 | Green | Male | 77
1.4 | Blue | Female | 57

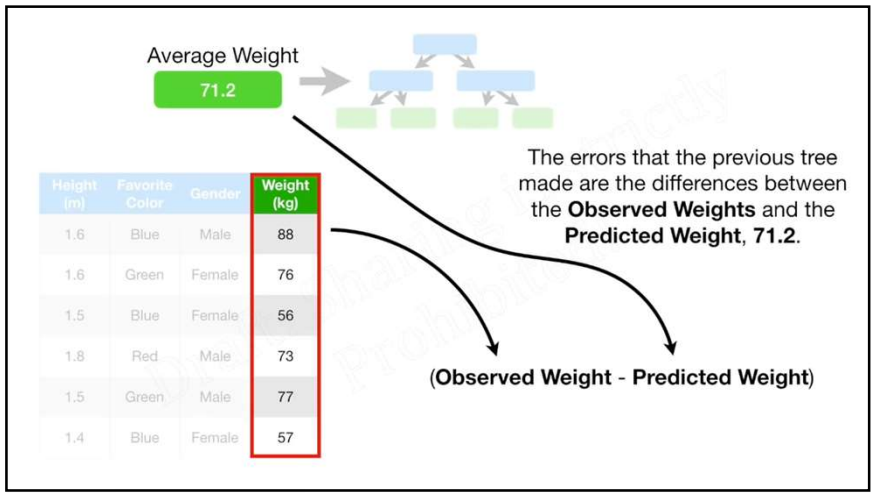

Average Weight
71.2
This is the first attempt at predicting everyone's weight.
Height (m) | Favorite Color | Gender | Weight (kg)
1.6 | Blue | Male | 88
1.6 | Green | Female | 76
1.5 | Blue | Female | 56
1.8 | Red | Male | 73
1.5 | Green | Male | 77
1.4 | Blue | Female | 57



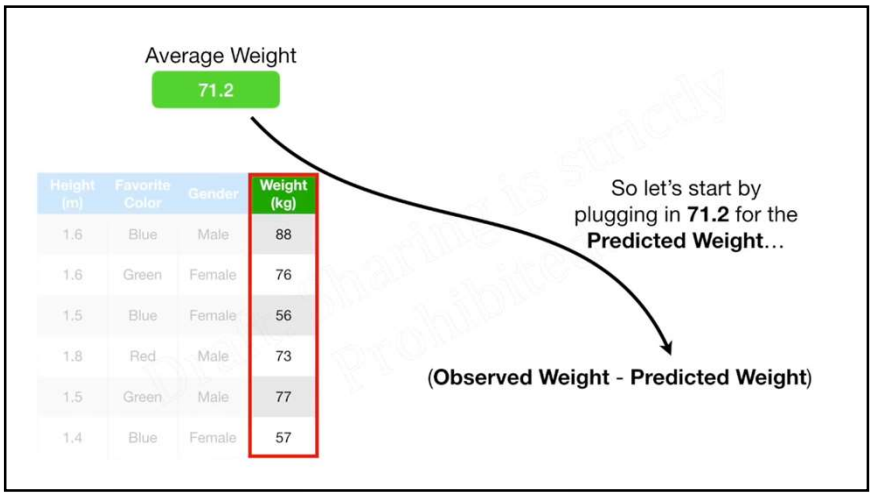
Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88
1.6	Green	Female	76
1.5	Blue	Female	56
1.8	Red	Male	73
1.5	Green	Male	77
1.4	Blue	Female	57



Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88
1.6	Green	Female	76
1.5	Blue	Female	56
1.8	Red	Male	73
1.5	Green	Male	77
1.4	Blue	Female	57



Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88
1.6	Green	Female	76
1.5	Blue	Female	56
1.8	Red	Male	73
1.5	Green	Male	77
1.4	Blue	Female	57



Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88
1.6	Green	Female	76
1.5	Blue	Female	56
1.8	Red	Male	73
1.5	Green	Male	77
1.4	Blue	Female	57

Average Weight

71.2

Height (m)	Favorite Color	Gender	Weight (kg)	Residual
1.6	Blue	Male	88	16.8
1.6	Green	Female	76	
1.5	Blue	Female	56	
1.8	Red	Male	73	
1.5	Green	Male	77	
1.4	Blue	Female	57	

…and save the difference, which is called a **Pseudo Residual**, in a new column.

(**88** - **71.2**) = 16.8

Average Weight

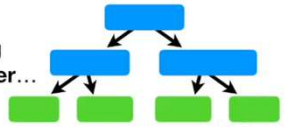
71.2

Height (m)	Favorite Color	Gender	Weight (kg)	Residual
1.6	Blue	Male	88	16.8
1.6	Green	Female	76	4.8
1.5	Blue	Female	56	-15.2
1.8	Red	Male	73	1.8
1.5	Green	Male	77	5.8
1.4	Blue	Female	57	-14.2

Now do the same thing for the remaining **Weights**…

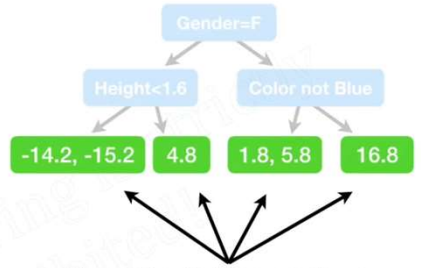
(**57** - **71.2**) = -14.2

Now we will build a **Tree**, using **Height**, **Favorite Color** and **Gender**…



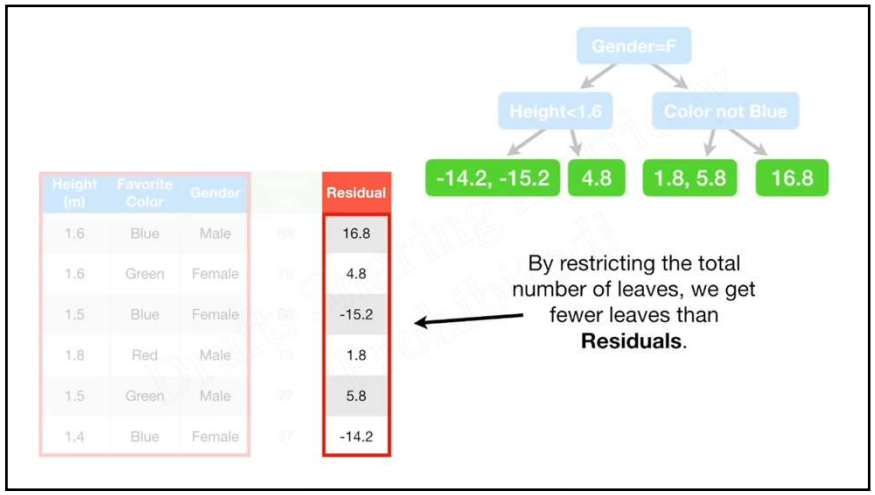
Height (m)	Favorite Color	Gender	Weight (kg)	Residual
1.6	Blue	Male	88	16.8
1.6	Green	Female	76	4.8
1.5	Blue	Female	56	-15.2
1.8	Red	Male	73	1.8
1.5	Green	Male	77	5.8
1.4	Blue	Female	57	-14.2

…to **Predict** the **Residuals**.

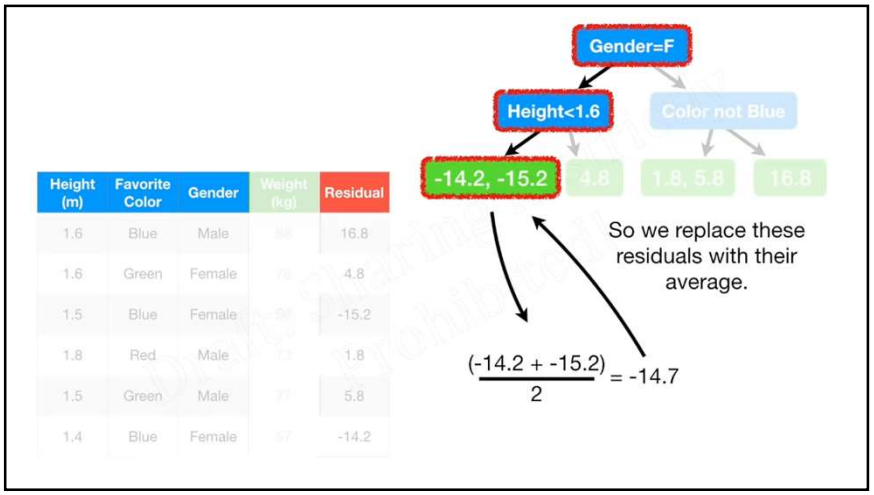


Remember, in this example we are only allowing up to four leaves…

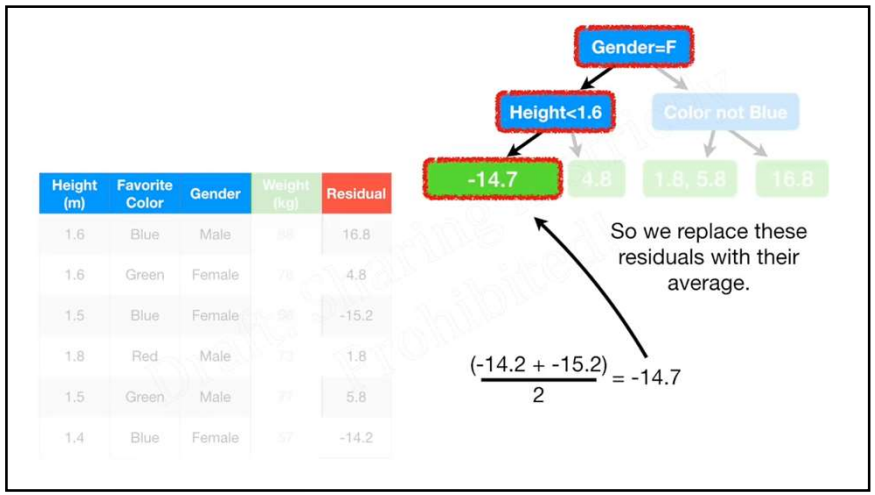
…but when using a larger dataset, it is common to allow anywhere from **8** to **32**.



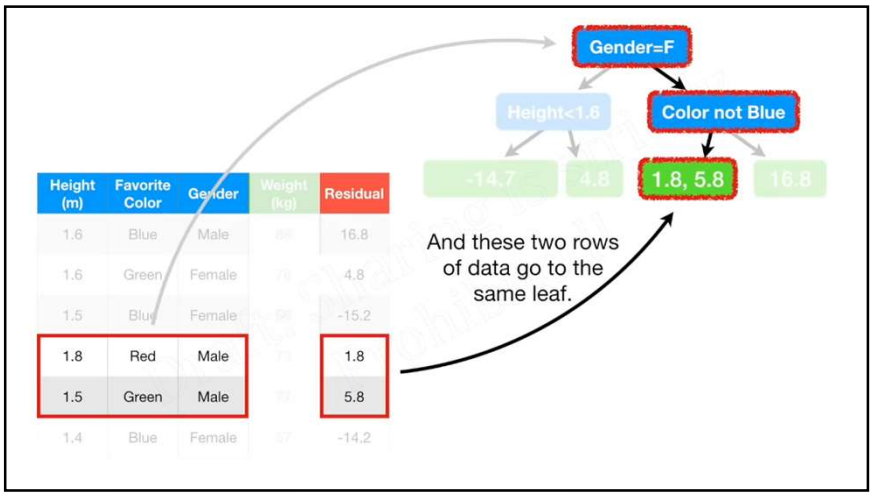
Height (m)	Favorite Color	Gender	Residual
1.6	Blue	Male	16.8
1.6	Green	Female	4.8
1.5	Blue	Female	-15.2
1.8	Red	Male	1.8
1.5	Green	Male	5.8
1.4	Blue	Female	-14.2



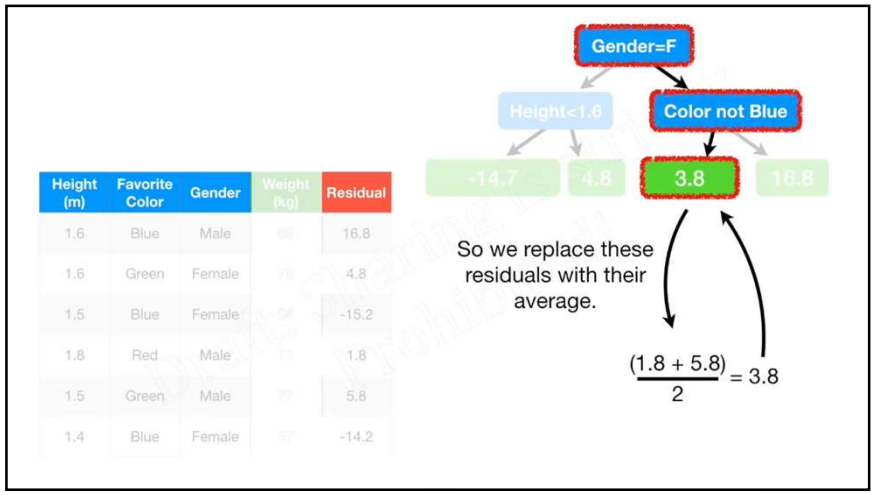
Height (m)	Favorite Color	Gender	Residual
1.6	Blue	Male	16.8
1.6	Green	Female	4.8
1.5	Blue	Female	-15.2
1.8	Red	Male	1.8
1.5	Green	Male	5.8
1.4	Blue	Female	-14.2



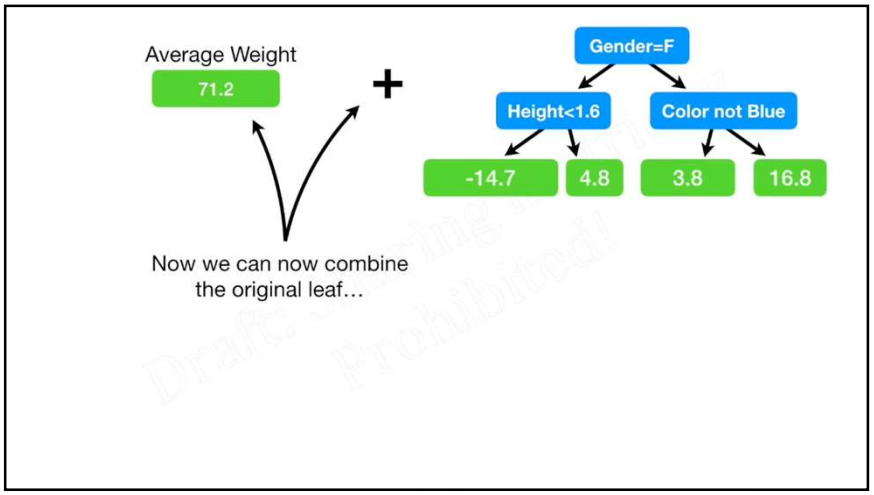
Height (m)	Favorite Color	Gender	Residual
1.6	Blue	Male	16.8
1.6	Green	Female	4.8
1.5	Blue	Female	-15.2
1.8	Red	Male	1.8
1.5	Green	Male	5.8
1.4	Blue	Female	-14.2

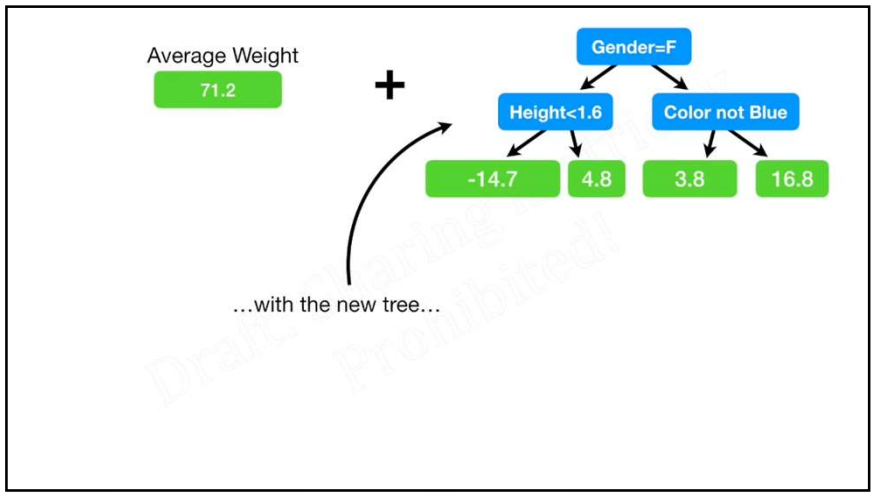


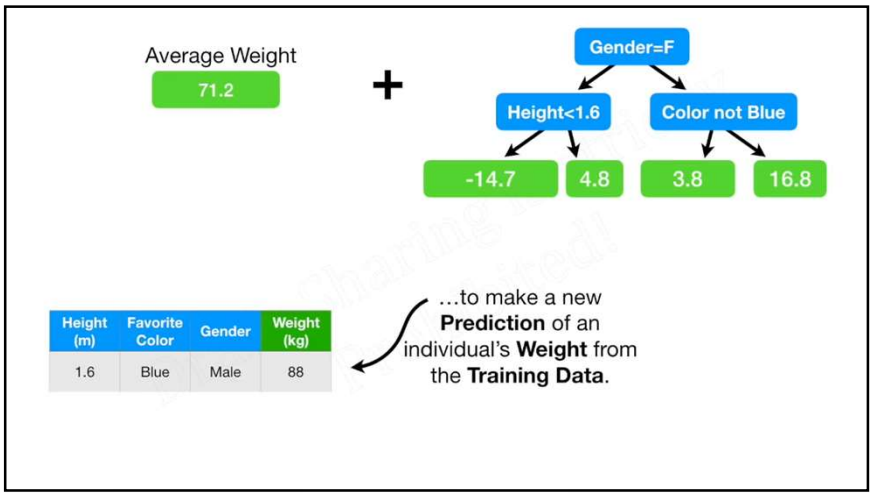
Height (m)	Favorite Color	Gender	Residual
1.6	Blue	Male	16.8
1.6	Green	Female	4.8
1.5	Blue	Female	-15.2
1.8	Red	Male	1.8
1.5	Green	Male	5.8
1.4	Blue	Female	-14.2



Height (m)	Favorite Color	Gender	Weight (kg)	Residual
1.6	Blue	Male	88	16.8
1.6	Green	Female	76	4.8
1.5	Blue	Female	56	-15.2
1.8	Red	Male	73	1.8
1.5	Green	Male	77	5.8
1.4	Blue	Female	57	-14.2

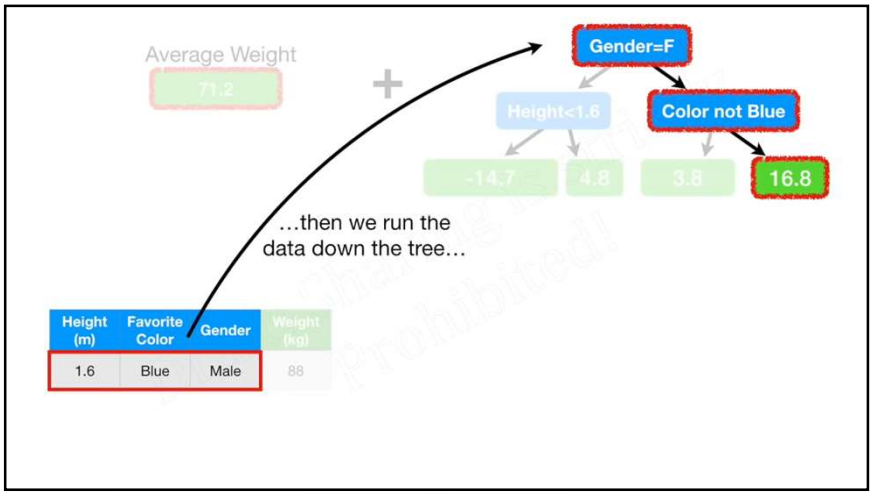


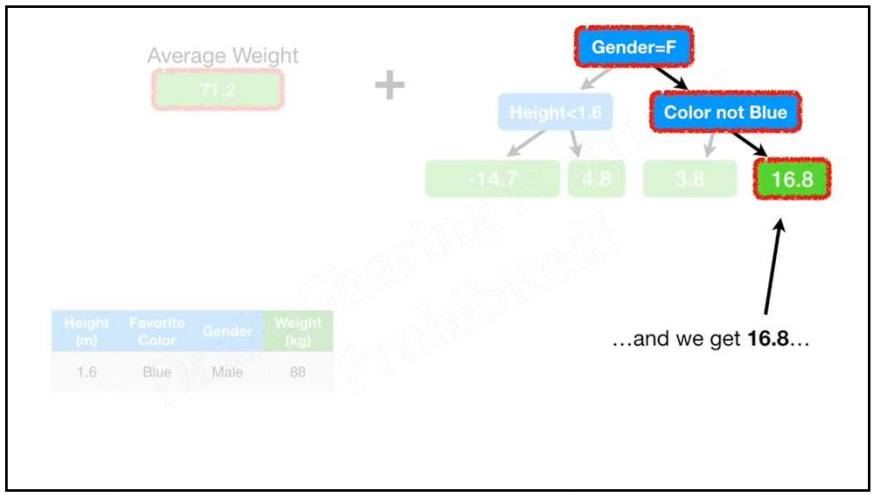


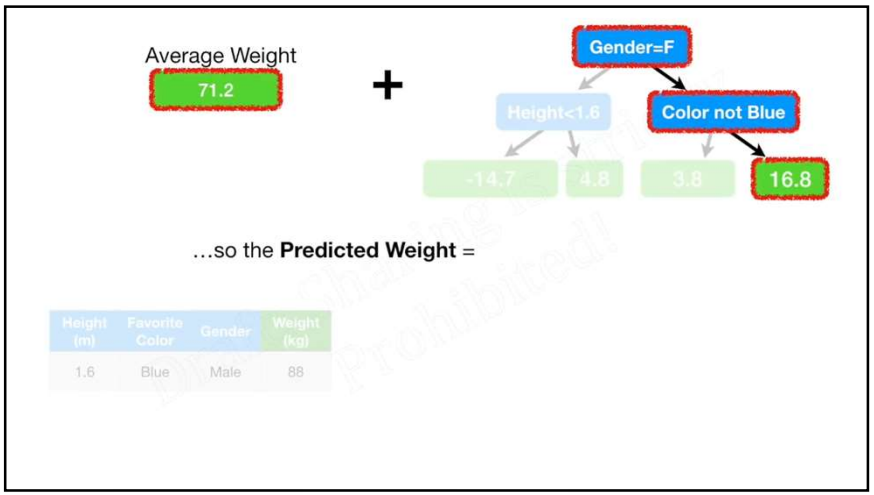


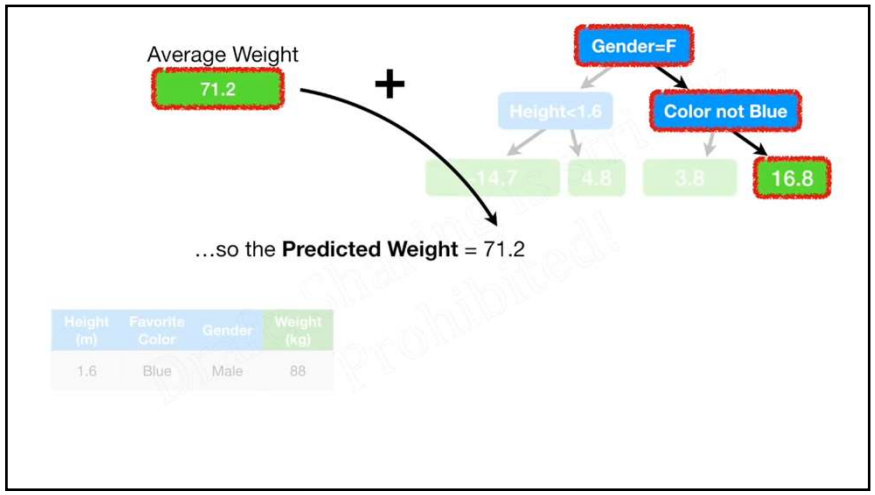
Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88

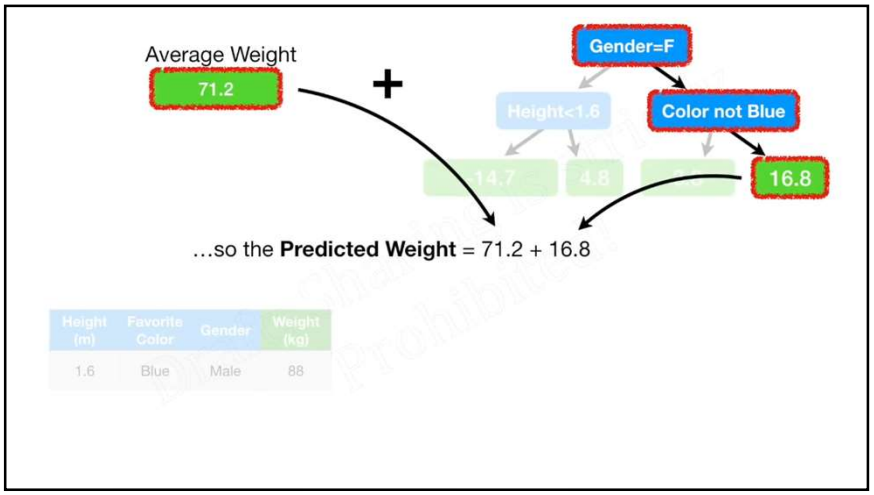

Average Weight
71.2
We start with the initial
Prediction, **71.2**…
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
Height (m)
Favorite Color
Gender
Weight (kg)
1.6
Blue
Male
88

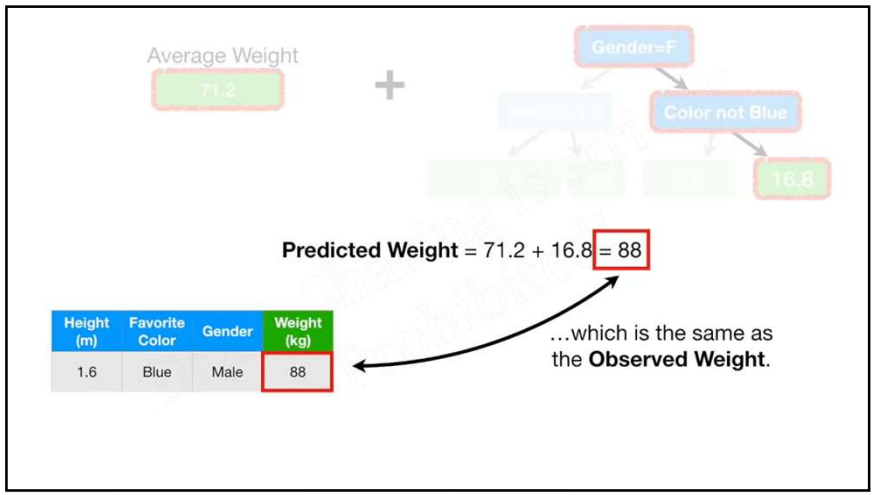

Average Weight
71.2
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
…then we run the
data down the tree…
Height (m)
Favorite Color
Gender
Weight (kg)
1.6
Blue
Male
88

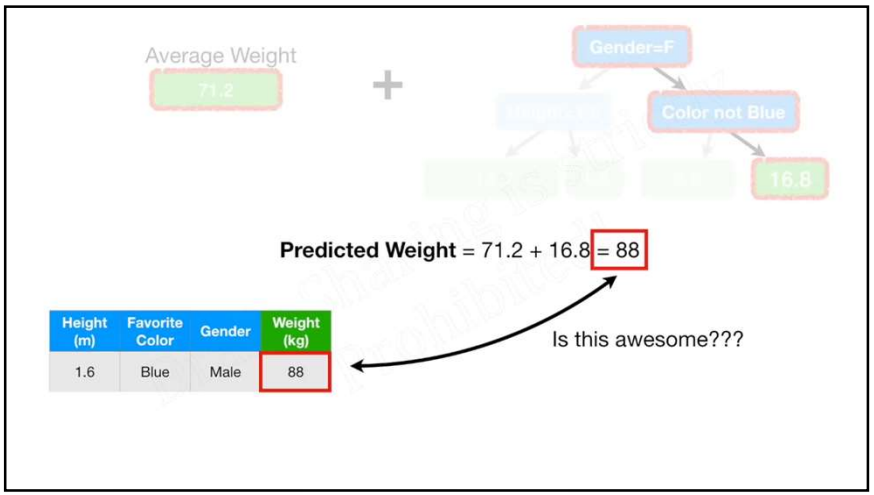

Average Weight
71.2
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
…and we get **16.8**…
Height (m)
Favorite Color
Gender
Weight (kg)
1.6
Blue
Male
88

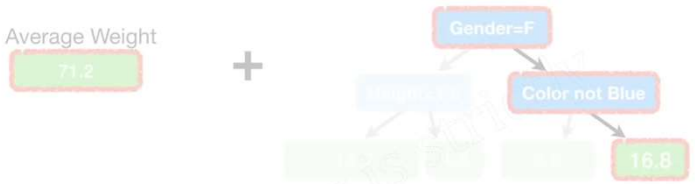

Average Weight
71.2
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
…so the **Predicted Weight** =
Height (m)
Favorite Color
Gender
Weight (kg)
1.6
Blue
Male
88


Average Weight
71.2
+
Gender=F
Height<1.6
Color not Blue
14.7
4.8
3.8
16.8
...so the **Predicted Weight** = 71.2
Height (m) | Favorite Color | Gender | Weight (kg)
1.6 | Blue | Male | 88


Average Weight
71.2
+
Gender=F
Height<1.6
Color not Blue
14.7
4.8
16.8
...so the **Predicted Weight** = 71.2 + 16.8
Height (m) | Favorite Color | Gender | Weight (kg)
1.6 | Blue | Male | 88


Average Weight
71.2
+
Gender=F
Color not Blue
16.8
Predicted Weight = 71.2 + 16.8 = 88
Height (m) | Favorite Color | Gender | Weight (kg)
1.6 | Blue | Male | 88
...which is the same as the **Observed Weight**.

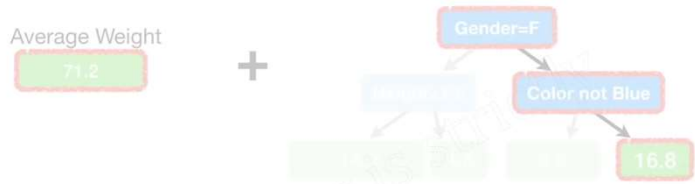

Average Weight
71.2
+
Gender=F
Color not Blue
16.8
Predicted Weight = 71.2 + 16.8 = 88
Height (m) | Favorite Color | Gender | Weight (kg)
1.6 | Blue | Male | 88
Is this awesome???



Predicted Weight = 71.2 + 16.8 = 88

Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88

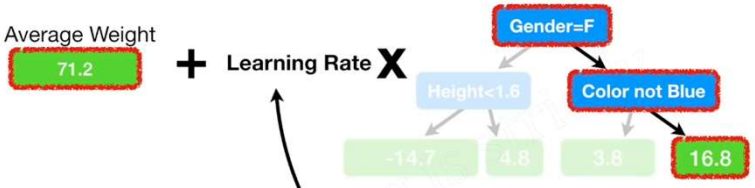
No. The model fits the **Training Data** too well.



Predicted Weight = 71.2 + 16.8 = 88

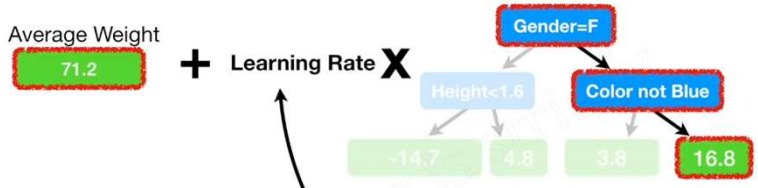
Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88

In other words, we have low **Bias**, but probably very high **Variance**.



Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88

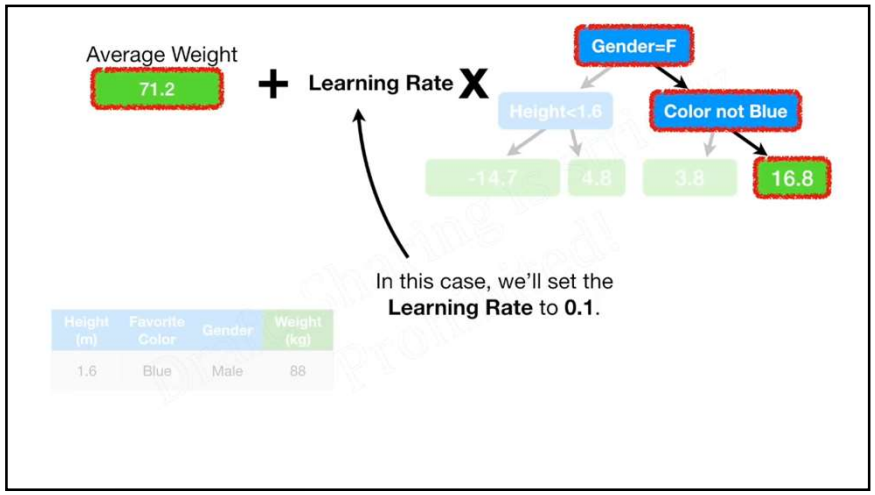
Gradient Boost deals with this problem by using a **Learning Rate** to scale the contribution from the new tree.

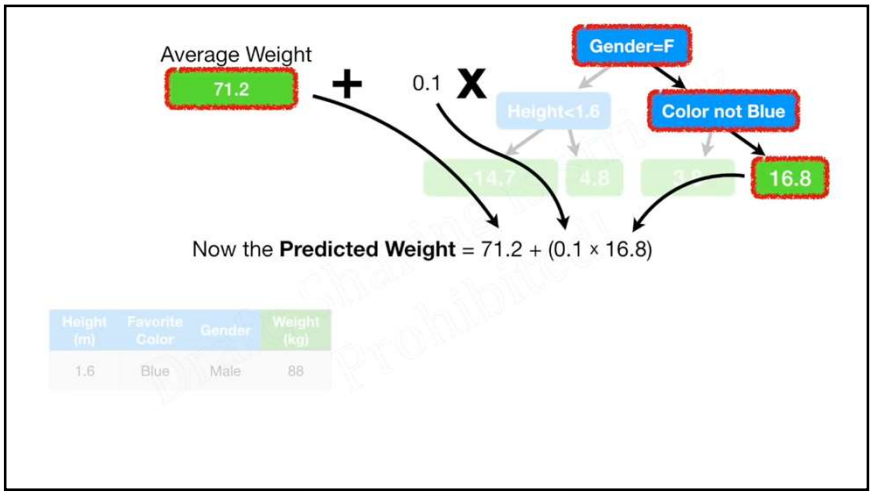


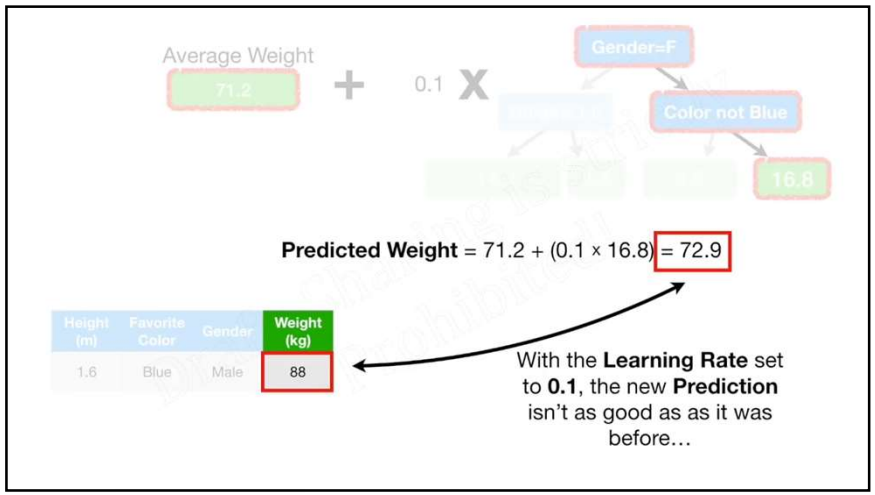
Height (m)	Favorite Color	Gender	Weight (kg)
1.6	Blue	Male	88

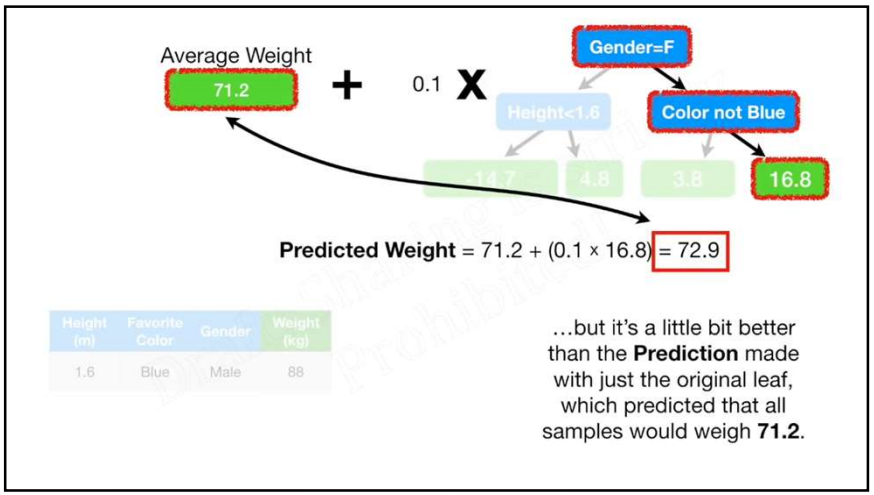
Gradient Boost deals with this problem by using a **Learning Rate** to scale the contribution from the new tree.

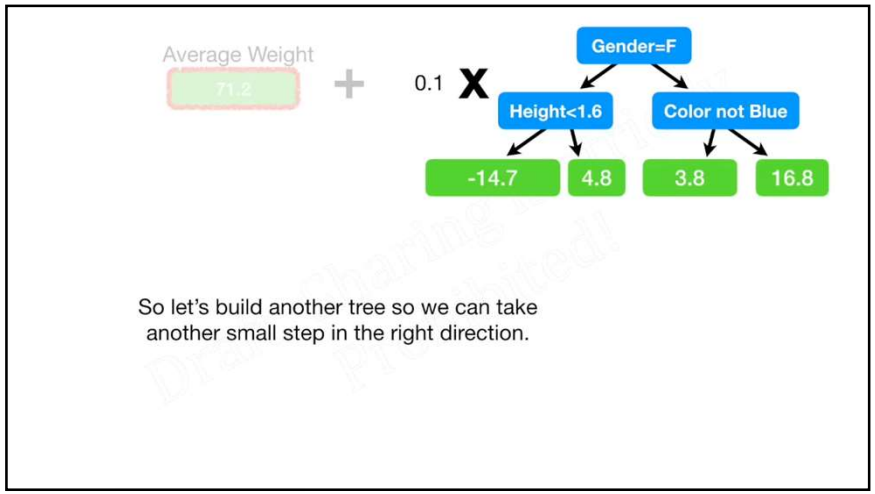
The **Learning Rate** is a value between **0** and **1**.


Average Weight
71.2
+
Learning Rate
X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
In this case, we'll set the Learning Rate to 0.1.
Height (m)
Favorite Color
Gender
Weight (kg)
1.6
Blue
Male
88

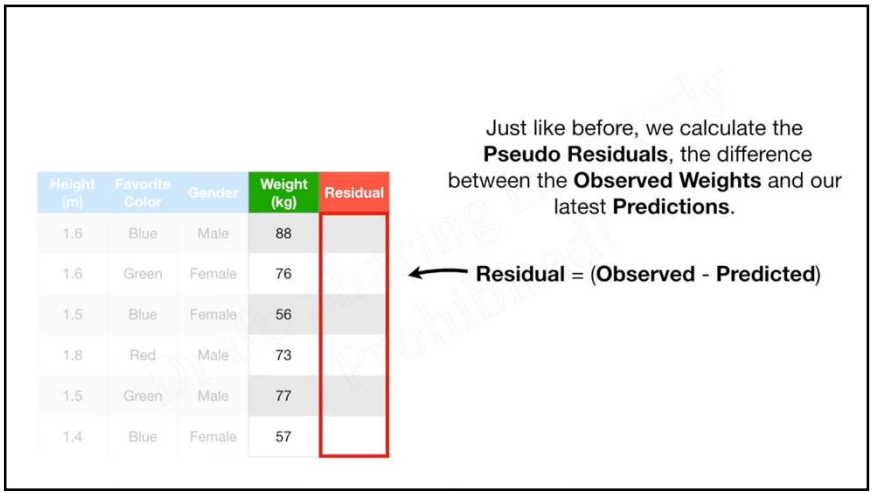

Average Weight
71.2
+
0.1
X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
16.8
Now the Predicted Weight = 71.2 + (0.1 × 16.8)
Height (m)
Favorite Color
Gender
Weight (kg)
1.6
Blue
Male
88


Average Weight
71.2
+
0.1
X
Gender=F
Color not Blue
16.8
Predicted Weight = 71.2 + (0.1 × 16.8) = 72.9
Height (m)
Favorite Color
Gender
Weight (kg)
1.6
Blue
Male
88
With the Learning Rate set to 0.1, the new Prediction isn't as good as as it was before…


Average Weight
71.2
+
0.1
X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
Predicted Weight = 71.2 + (0.1 × 16.8) = 72.9
Height (m)
Favorite Color
Gender
Weight (kg)
1.6
Blue
Male
88
…but it's a little bit better than the Prediction made with just the original leaf, which predicted that all samples would weigh 71.2.

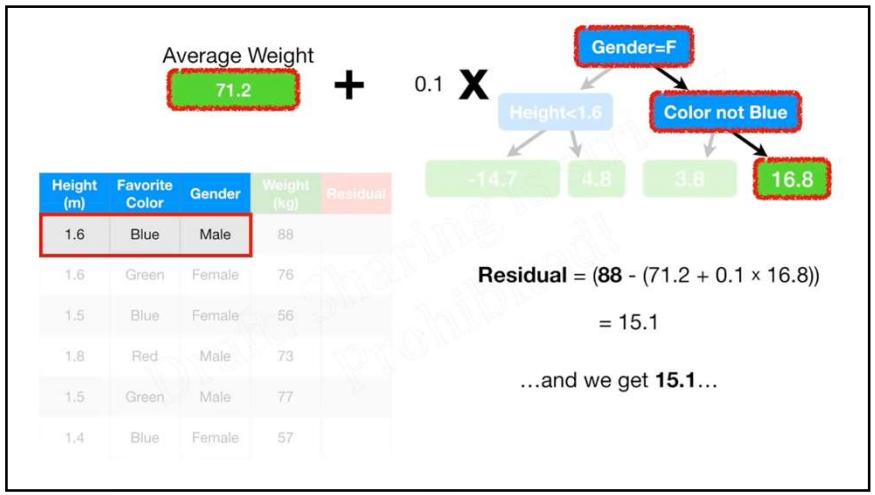


So let's build another tree so we can take another small step in the right direction.



Just like before, we calculate the **Pseudo Residuals**, the difference between the **Observed Weights** and our latest **Predictions**.

Residual = (**Observed** - **Predicted**)



Residual = (**88** - (71.2 + 0.1 × 16.8))

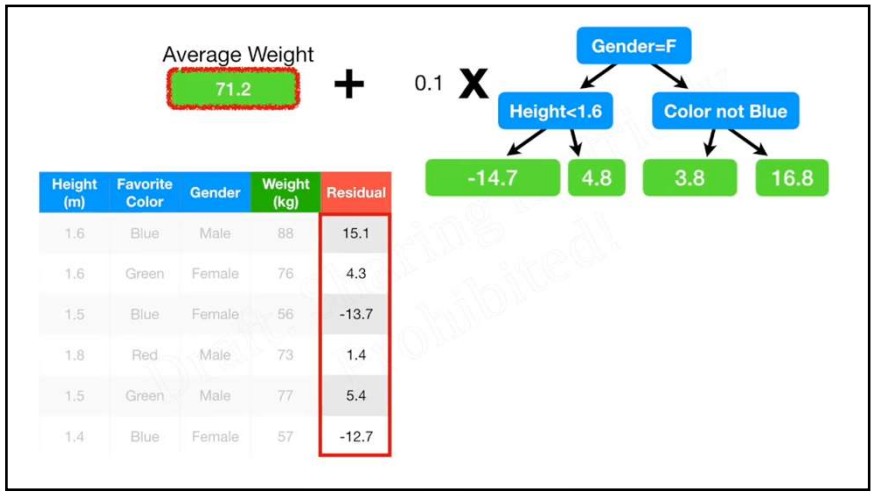
= 15.1

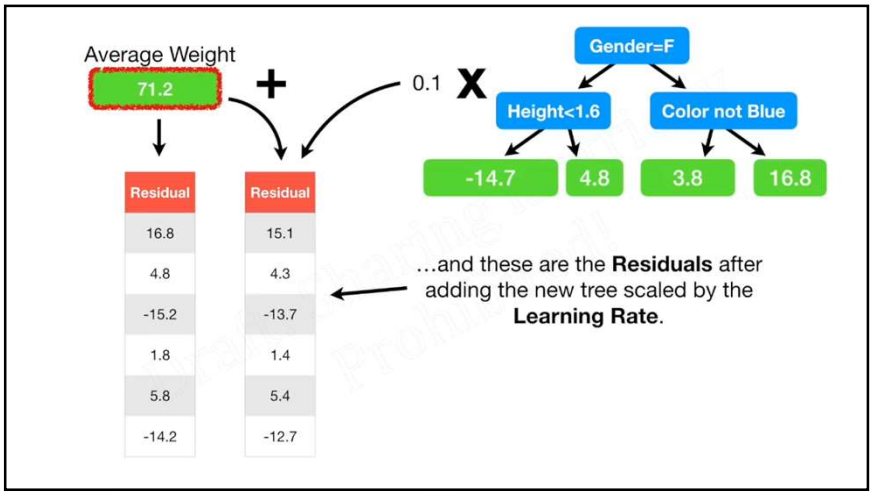
…and we get **15.1**…

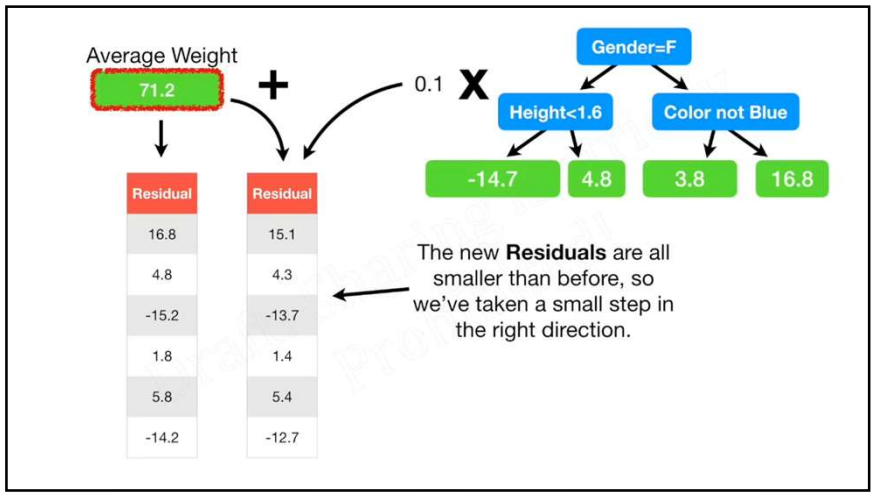
Residual = (**88** - (71.2 + 0.1 × 16.8))

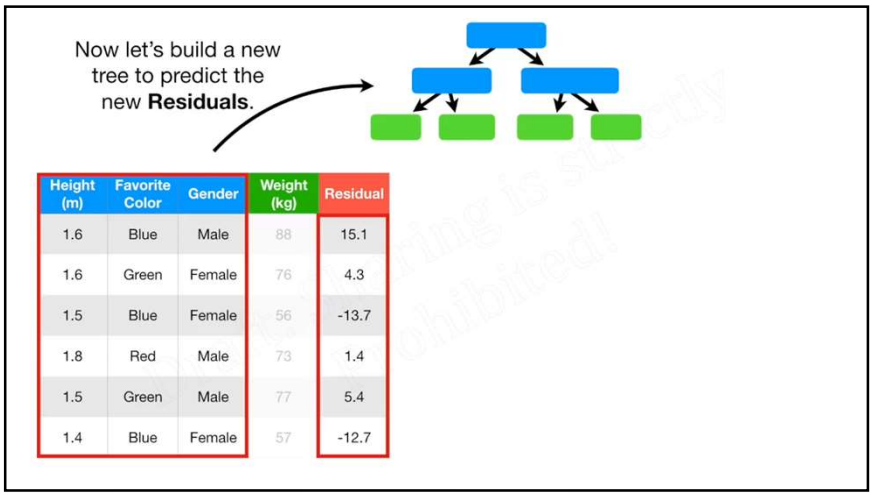
= 15.1

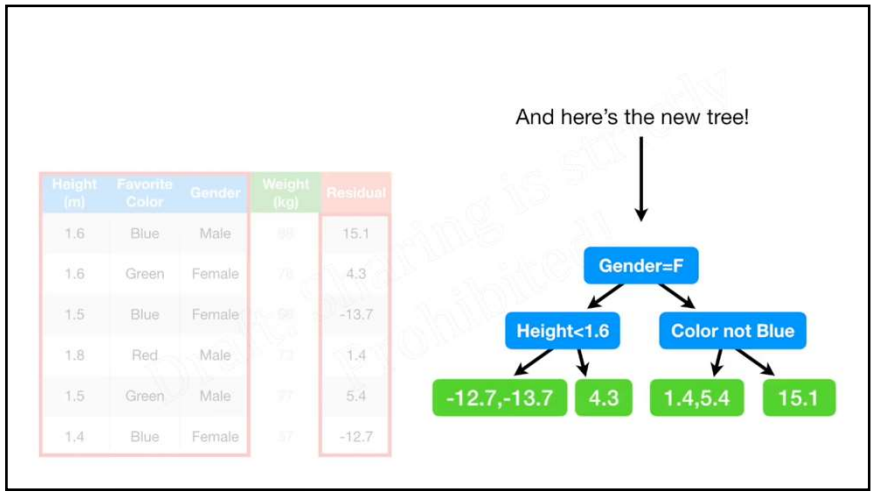
…and we save that in the column for **Pseudo Residuals**.

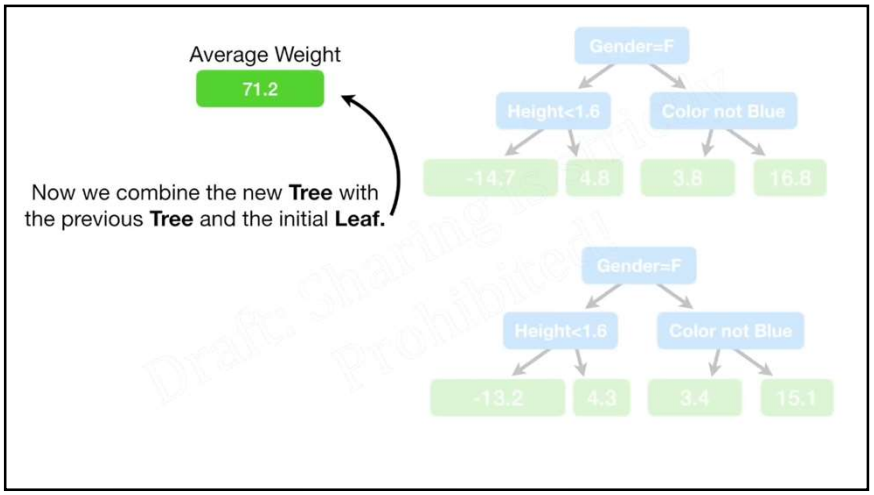

Average Weight
71.2
+
0.1
X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
Height (m)
Favorite Color
Gender
Weight (kg)
Residual
1.6 Blue Male 88 15.1
1.6 Green Female 76 4.3
1.5 Blue Female 56 -13.7
1.8 Red Male 73 1.4
1.5 Green Male 77 5.4
1.4 Blue Female 57 -12.7

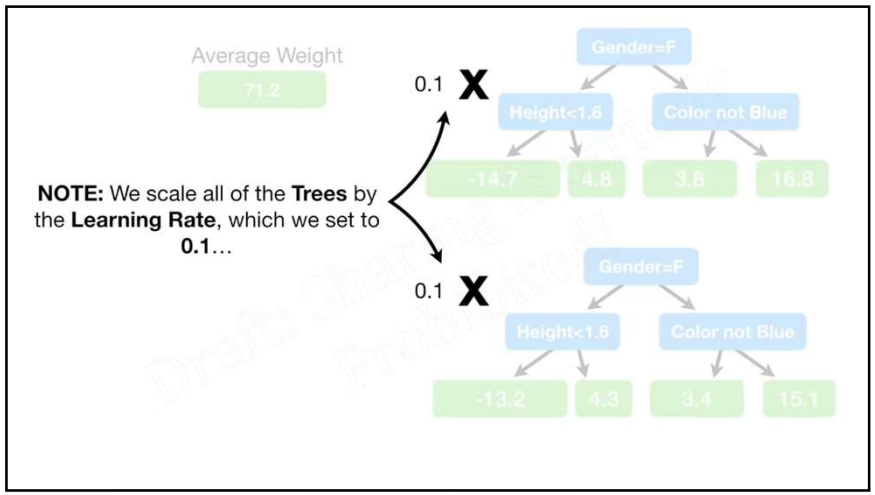

Average Weight
71.2
+
0.1
X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
Residual
16.8
4.8
-15.2
1.8
5.8
-14.2
Residual
15.1
4.3
-13.7
1.4
5.4
-12.7
…and these are the **Residuals** after adding the new tree scaled by the **Learning Rate**.

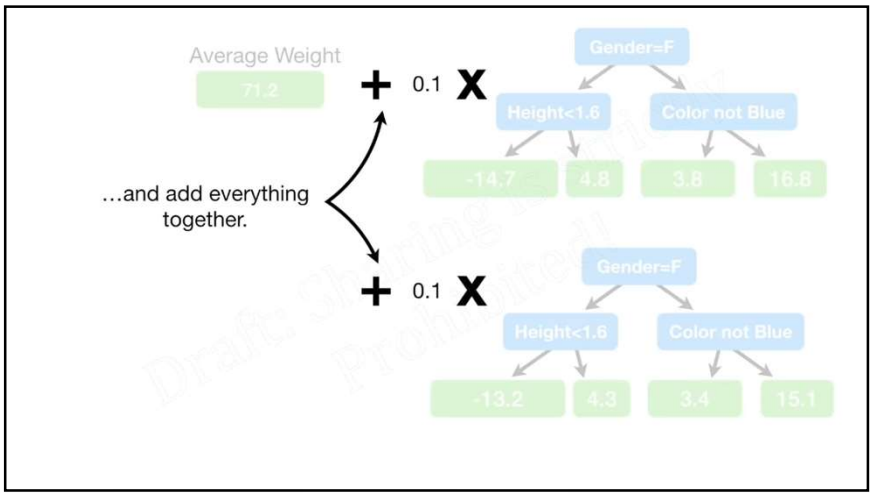

Average Weight
71.2
+
0.1
X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
Residual
16.8
4.8
-15.2
1.8
5.8
-14.2
Residual
15.1
4.3
-13.7
1.4
5.4
-12.7
The new **Residuals** are all smaller than before, so we've taken a small step in the right direction.

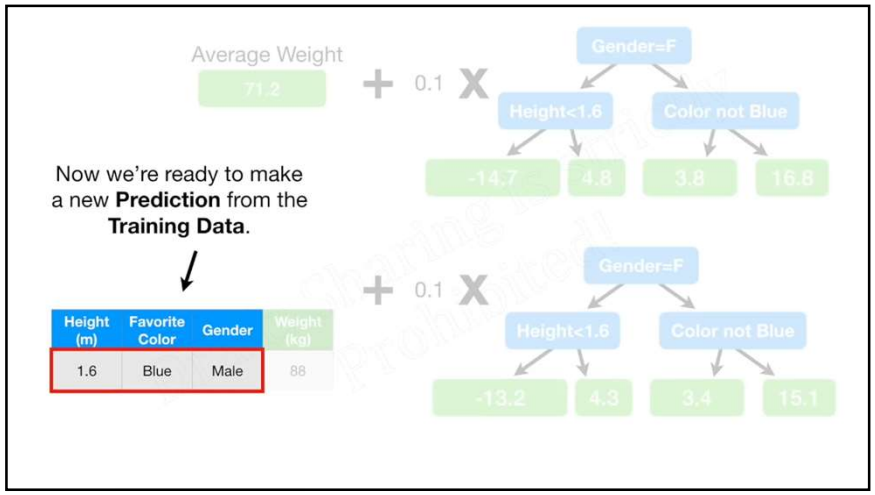

Now let's build a new tree to predict the new **Residuals**.
Height (m)
Favorite Color
Gender
Weight (kg)
Residual
1.6 Blue Male 88 15.1
1.6 Green Female 76 4.3
1.5 Blue Female 56 -13.7
1.8 Red Male 73 1.4
1.5 Green Male 77 5.4
1.4 Blue Female 57 -12.7

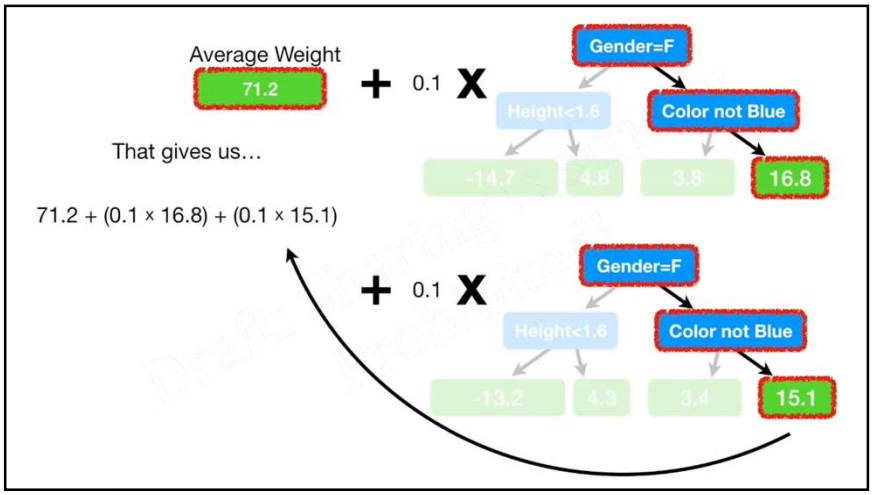

And here's the new tree!
Gender=F
Height<1.6
Color not Blue
-12.7,-13.7
4.3
1.4,5.4
15.1


Average Weight
71.2
Now we combine the new **Tree** with the previous **Tree** and the initial **Leaf.**
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
Gender=F
Height<1.6
Color not Blue
-13.2
4.3
3.4
15.1

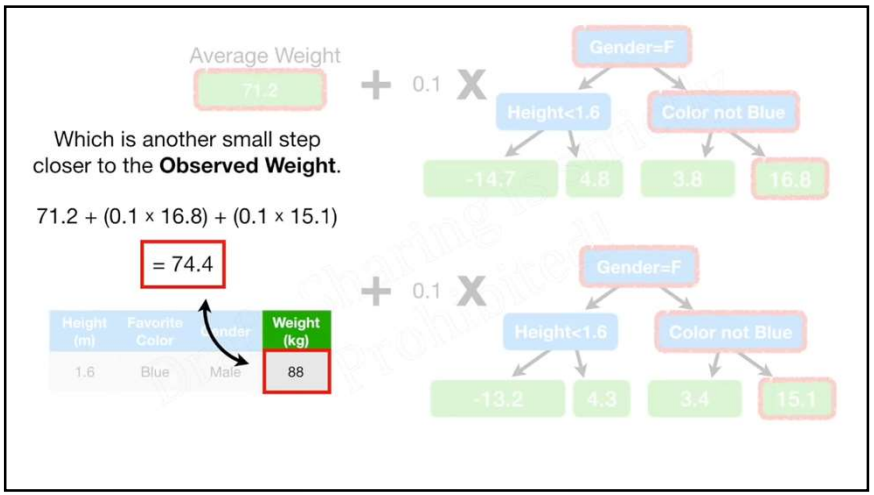

Average Weight
71.2
NOTE: We scale all of the **Trees** by the **Learning Rate**, which we set to **0.1**…
0.1 X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
0.1 X
Gender=F
Height<1.6
Color not Blue
-13.2
4.3
3.4
15.1

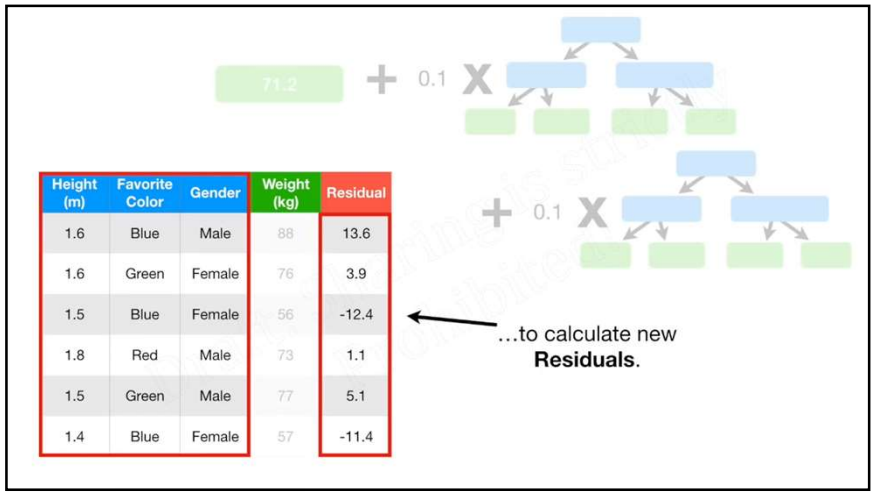

Average Weight
71.2
…and add everything together.
+ 0.1 X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
+ 0.1 X
Gender=F
Height<1.6
Color not Blue
-13.2
4.3
3.4
15.1

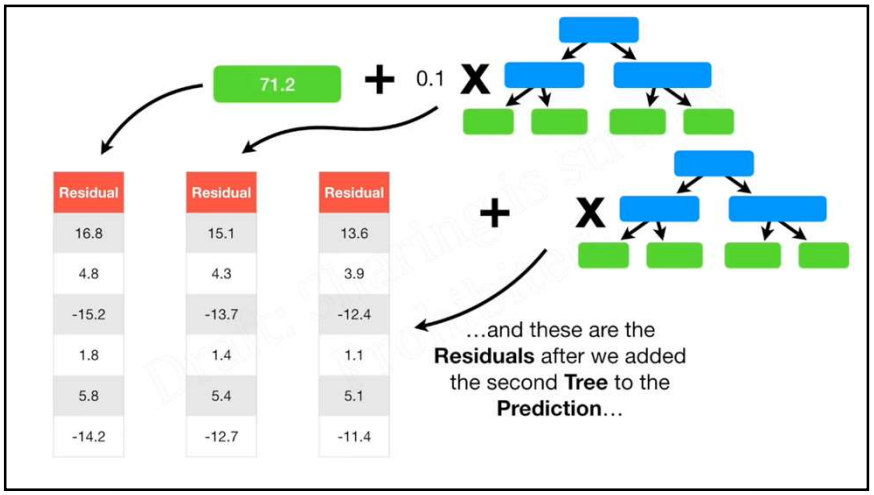

Average Weight
71.2
+ 0.1 X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
Now we're ready to make a new **Prediction** from the **Training Data**.
Height (m)
Favorite Color
Gender
Weight (kg)
1.6
Blue
Male
88
+ 0.1 X
Gender=F
Height<1.6
Color not Blue
-13.2
4.3
3.4
15.1

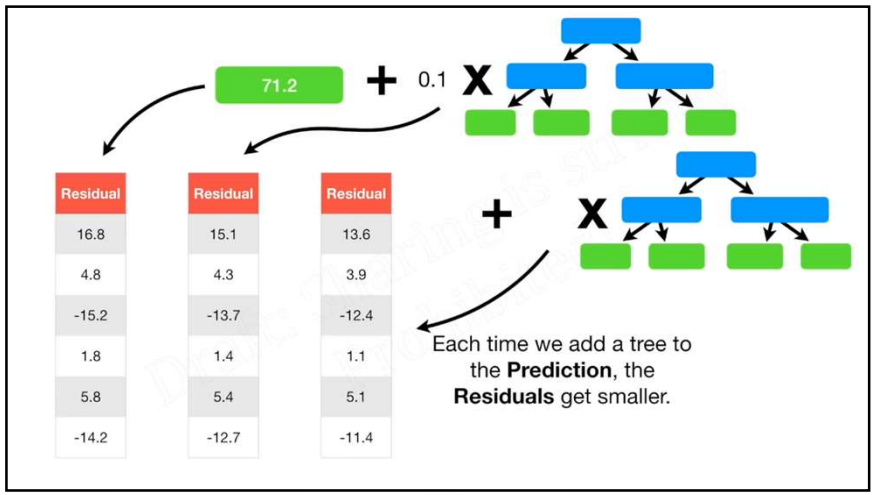

Average Weight
71.2
+ 0.1 X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
That gives us…
71.2 + (0.1 × 16.8) + (0.1 × 15.1)
+ 0.1 X
Gender=F
Height<1.6
Color not Blue
-13.2
4.3
3.4
15.1

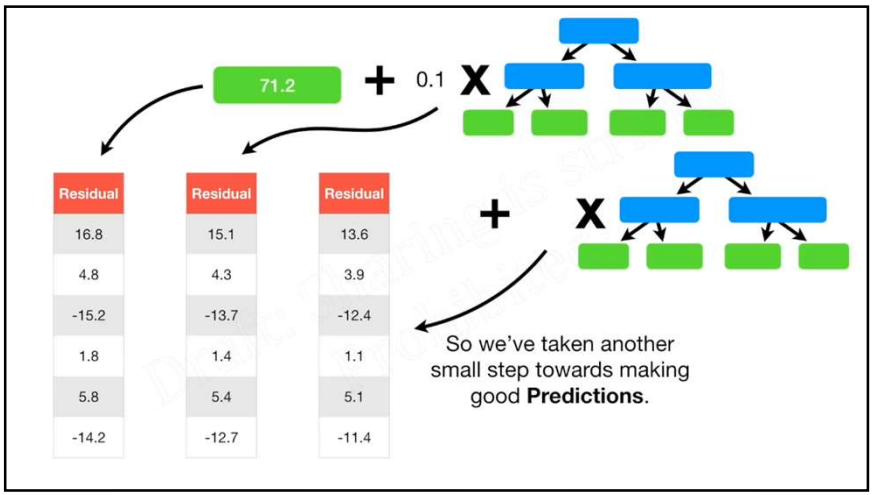

Average Weight
71.2
+ 0.1 X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
That gives us…
71.2 + (0.1 × 16.8) + (0.1 × 15.1)
= 74.4
+ 0.1 X
Gender=F
Height<1.6
Color not Blue
-13.2
4.3
3.4
15.1

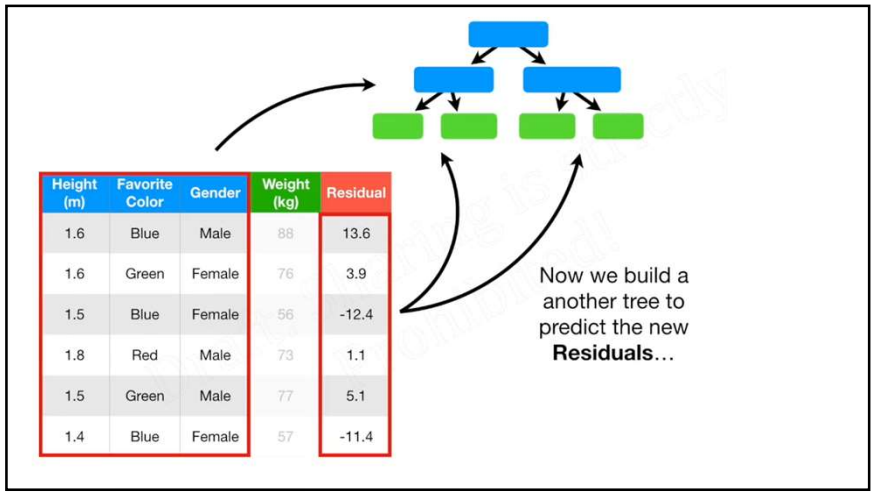

Average Weight
71.2
+ 0.1 X
Gender=F
Height<1.6
Color not Blue
-14.7
4.8
3.8
16.8
Which is another small step closer to the **Observed Weight**.
71.2 + (0.1 × 16.8) + (0.1 × 15.1)
= 74.4
Height (m)
Favorite Color
Gender
Weight (kg)
1.6
Blue
Male
88
+ 0.1 X
Gender=F
Height<1.6
Color not Blue
-13.2
4.3
3.4
15.1

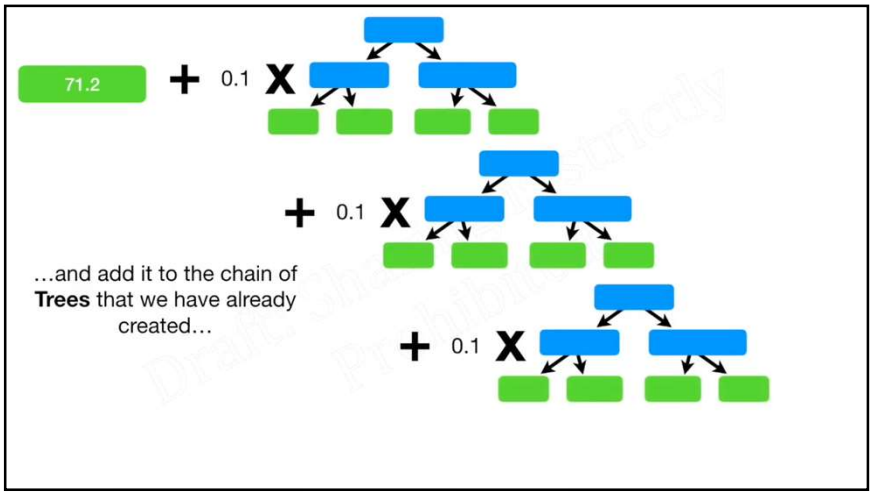

71.2
+ 0.1 X
+ 0.1 X
Height (m)
Favorite Color
Gender
Weight (kg)
Residual
1.6 Blue Male 88 13.6
1.6 Green Female 76 3.9
1.5 Blue Female 56 -12.4
1.8 Red Male 73 1.1
1.5 Green Male 77 5.1
1.4 Blue Female 57 -11.4
...to calculate new Residuals.

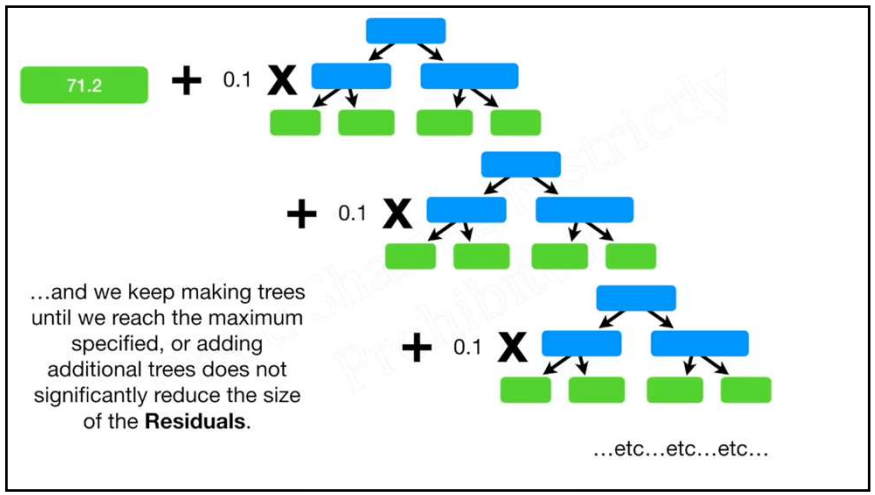

71.2
+ 0.1 X
Residual
16.8
4.8
-15.2
1.8
5.8
-14.2
Residual
15.1
4.3
-13.7
1.4
5.4
-12.7
Residual
13.6
3.9
-12.4
1.1
5.1
-11.4
+ X
...and these are the Residuals after we added the second Tree to the Prediction...

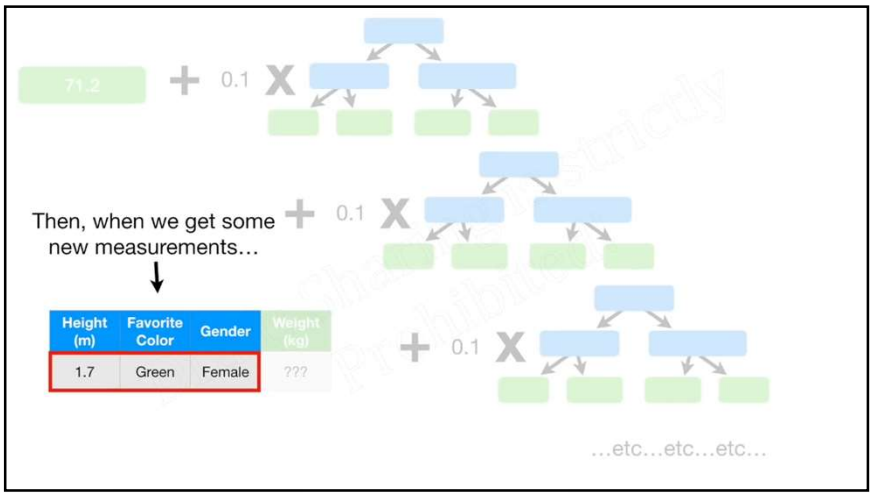

71.2
+ 0.1 X
Residual
16.8
4.8
-15.2
1.8
5.8
-14.2
Residual
15.1
4.3
-13.7
1.4
5.4
-12.7
Residual
13.6
3.9
-12.4
1.1
5.1
-11.4
+ X
Each time we add a tree to the Prediction, the Residuals get smaller.

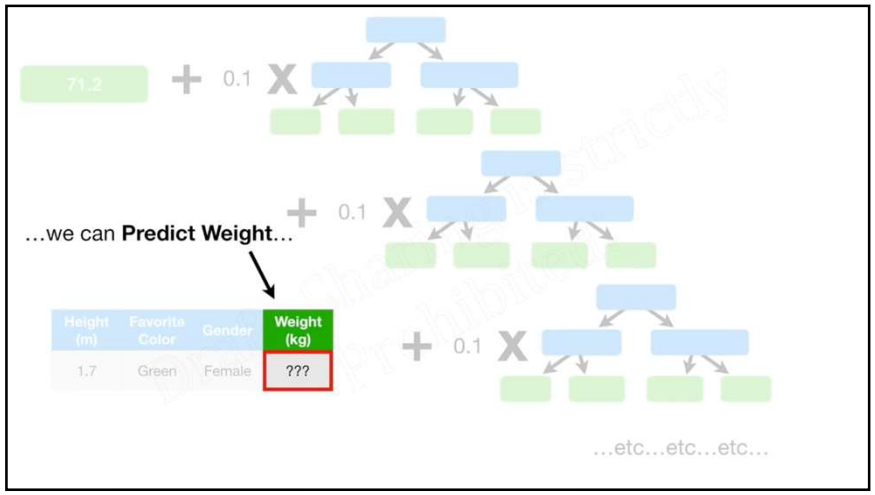

71.2
+ 0.1 X
Residual
16.8
4.8
-15.2
1.8
5.8
-14.2
Residual
15.1
4.3
-13.7
1.4
5.4
-12.7
Residual
13.6
3.9
-12.4
1.1
5.1
-11.4
+ X
So we've taken another small step towards making good Predictions.

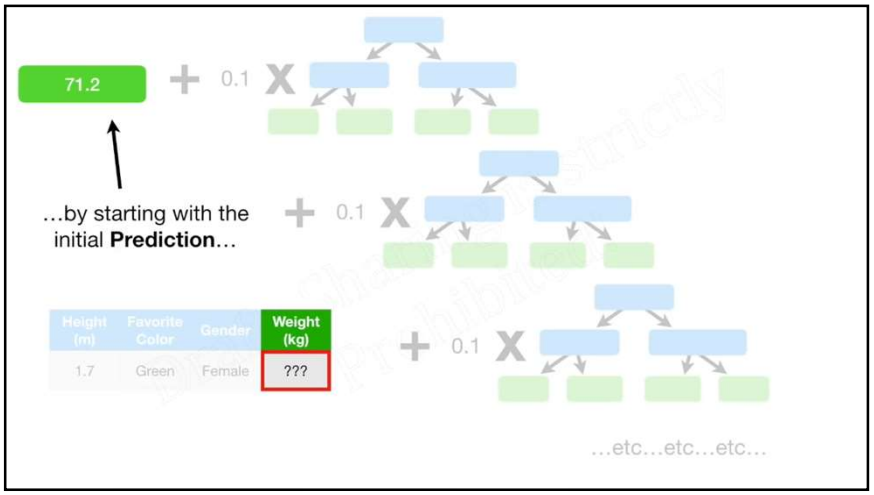

Height (m) | Favorite Color | Gender | Weight (kg) | Residual
1.6 | Blue | Male | 88 | 13.6
1.6 | Green | Female | 76 | 3.9
1.5 | Blue | Female | 56 | -12.4
1.8 | Red | Male | 73 | 1.1
1.5 | Green | Male | 77 | 5.1
1.4 | Blue | Female | 57 | -11.4
Now we build a another tree to predict the new **Residuals**…

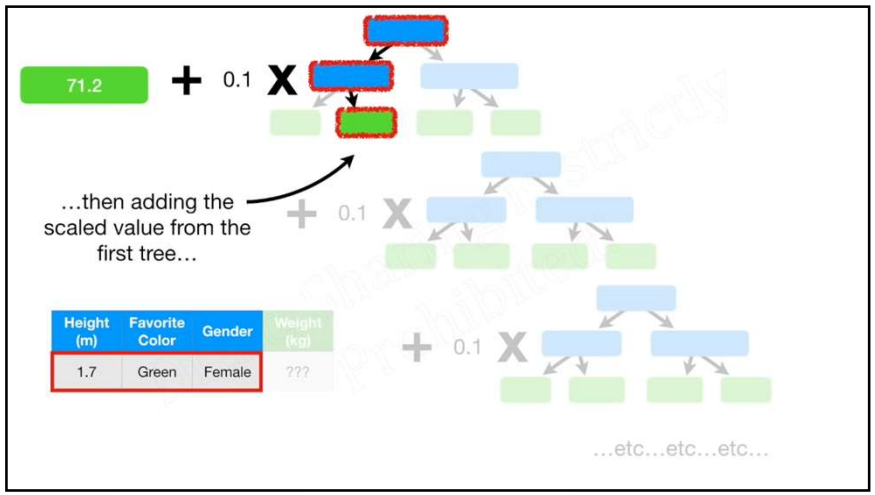

71.2
+ 0.1 X
+ 0.1 X
+ 0.1 X
…and add it to the chain of **Trees** that we have already created…

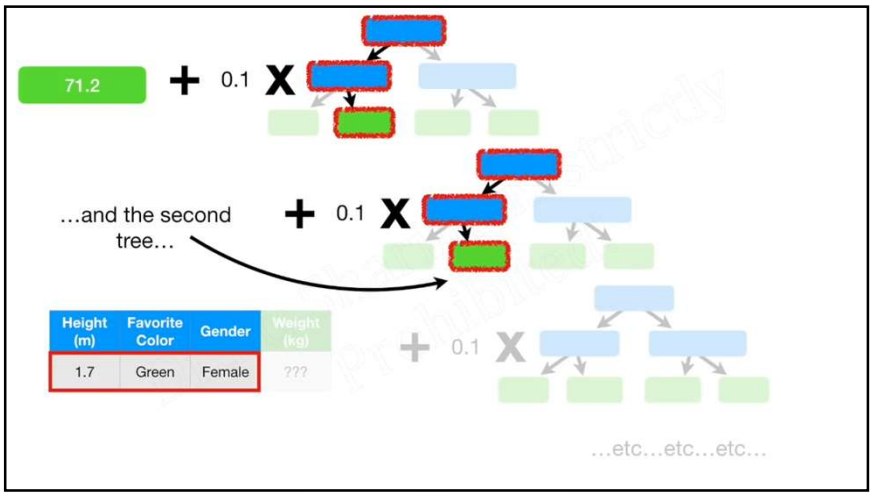

71.2
+ 0.1 X
+ 0.1 X
+ 0.1 X
…and we keep making trees until we reach the maximum specified, or adding additional trees does not significantly reduce the size of the **Residuals**.
…etc…etc…etc…

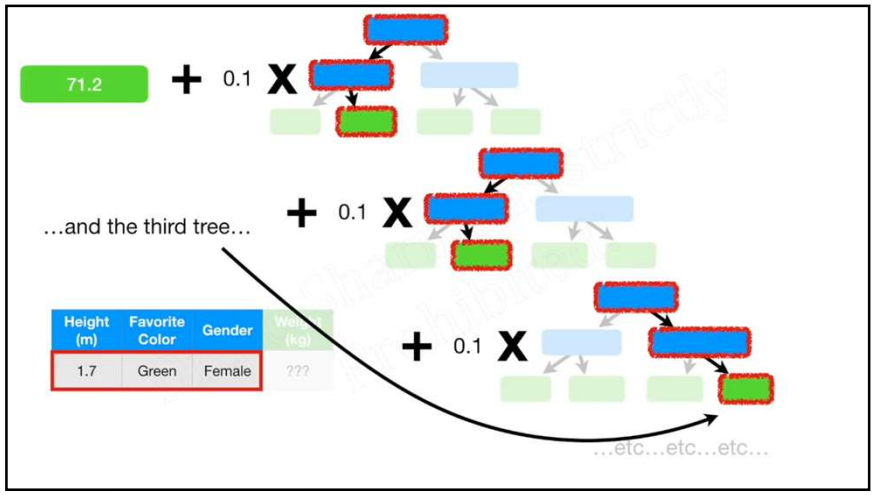

71.2
+ 0.1 X
+ 0.1 X
+ 0.1 X
Then, when we get some new measurements…
Height (m) | Favorite Color | Gender | Weight (kg)
1.7 | Green | Female | ???
…etc…etc…etc…

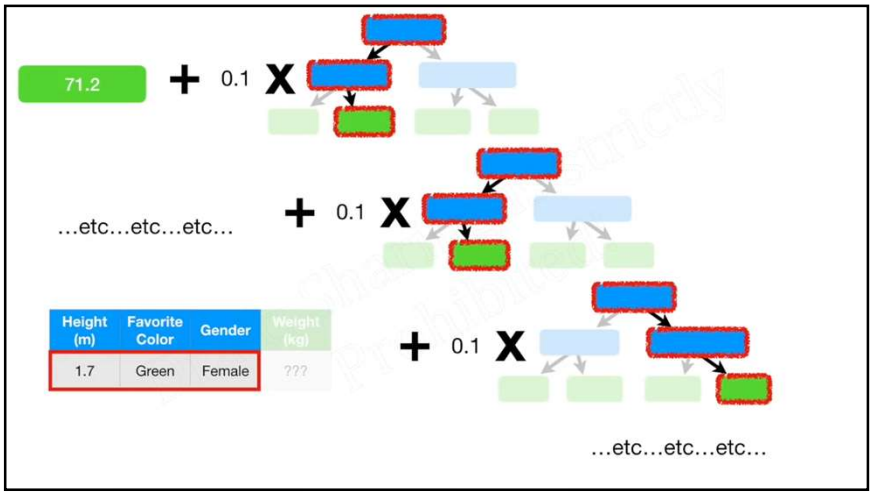

71.2
+ 0.1 X
+ 0.1 X
+ 0.1 X
…we can **Predict Weight**…
Height (m) | Favorite Color | Gender | Weight (kg)
1.7 | Green | Female | ???
…etc…etc…etc…

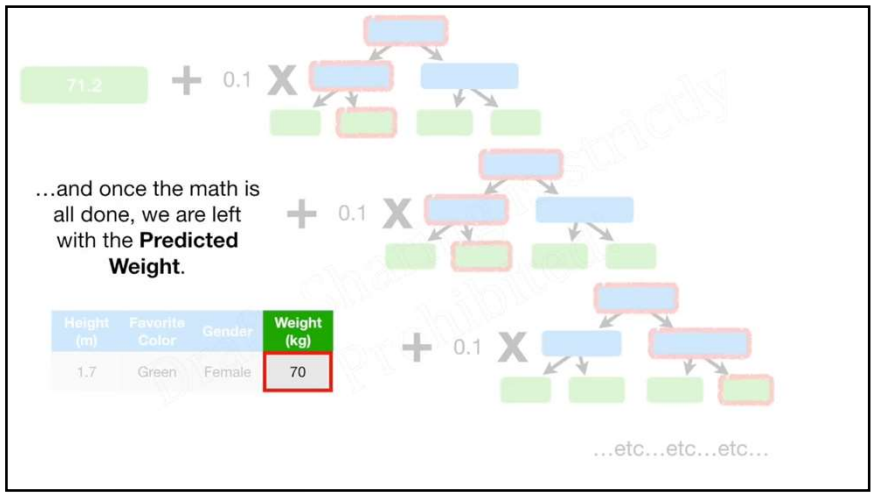

71.2
…by starting with the initial **Prediction**…
+ 0.1 X
+ 0.1 X
Height (m) | Favorite Color | Gender | Weight (kg)
1.7 | Green | Female | ???
…etc…etc…etc…

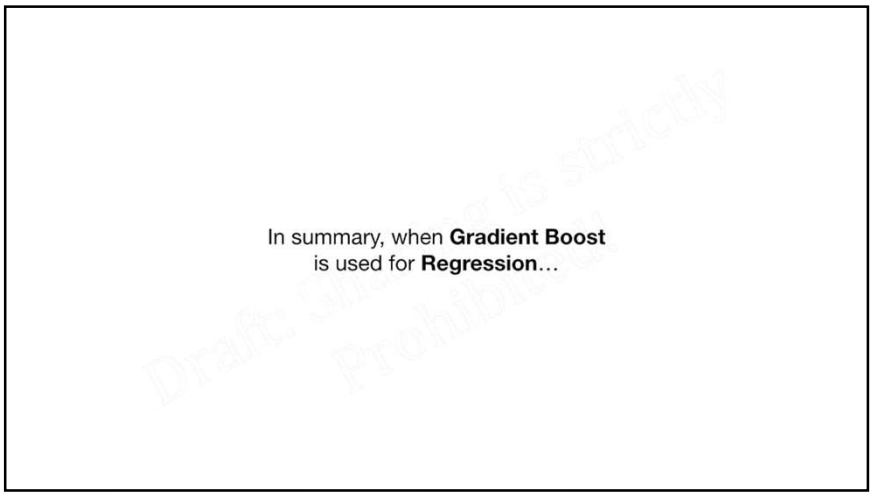

71.2
+ 0.1 X
…then adding the scaled value from the first tree…
+ 0.1 X
+ 0.1 X
Height (m) | Favorite Color | Gender | Weight (kg)
1.7 | Green | Female | ???
…etc…etc…etc…

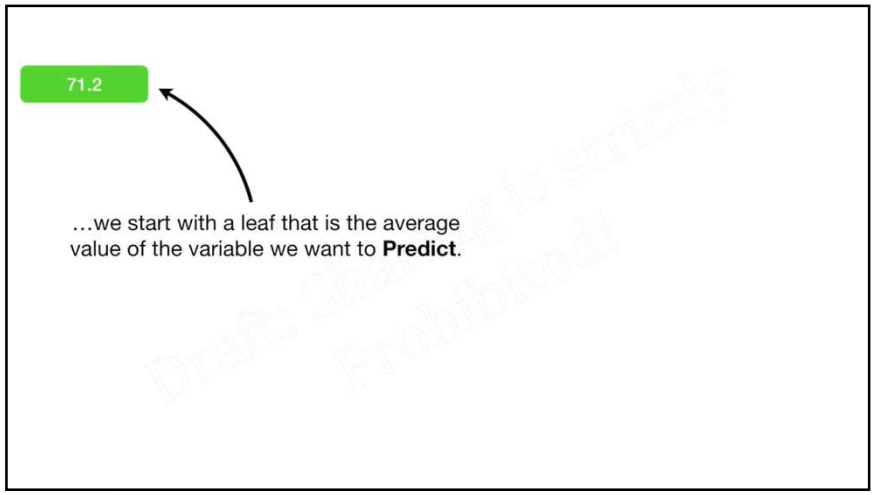

71.2
+ 0.1 X
…and the second tree…
+ 0.1 X
+ 0.1 X
Height (m) | Favorite Color | Gender | Weight (kg)
1.7 | Green | Female | ???
…etc…etc…etc…

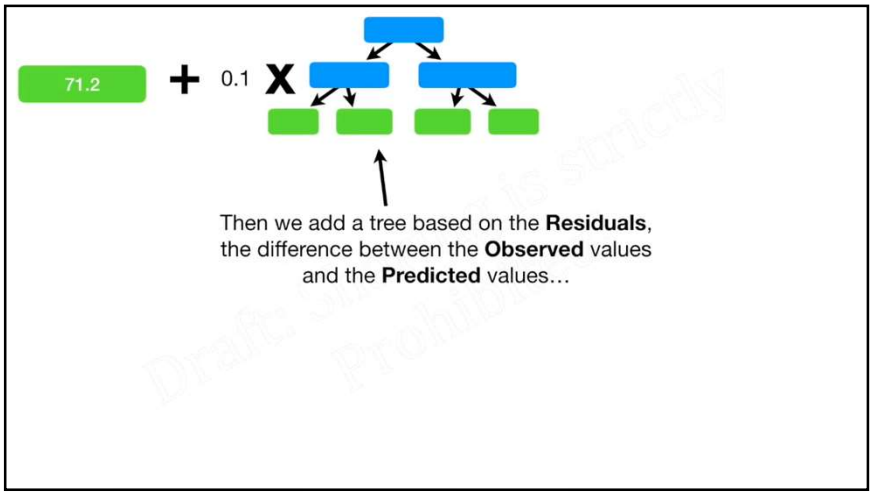

71.2
+ 0.1 X
+ 0.1 X
+ 0.1 X
...and the third tree...
Height (m) | Favorite Color | Gender | Weight (kg)
1.7 | Green | Female | ???
...etc...etc...etc...

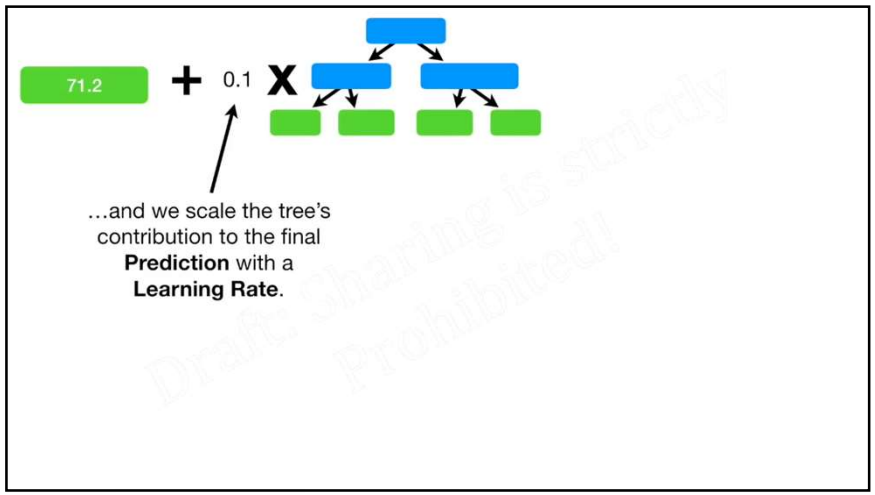

71.2
+ 0.1 X
+ 0.1 X
+ 0.1 X
...etc...etc...etc...
Height (m) | Favorite Color | Gender | Weight (kg)
1.7 | Green | Female | ???
...etc...etc...etc...

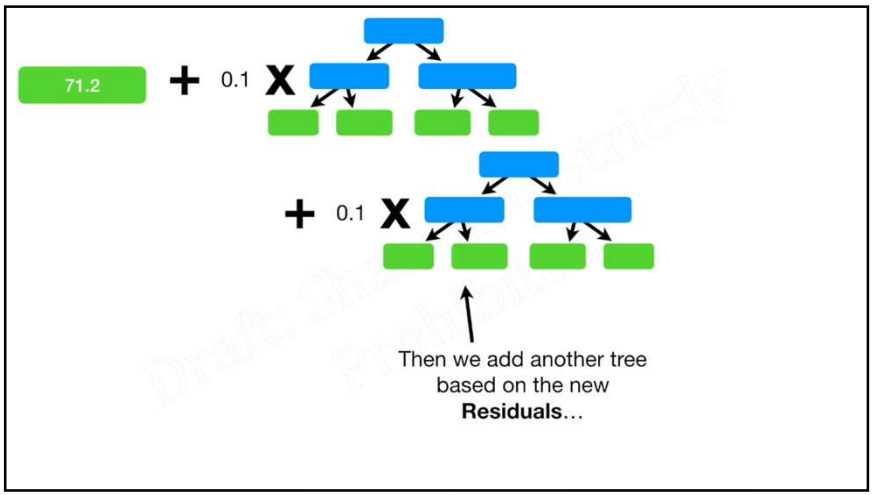

71.2
+ 0.1 X
+ 0.1 X
+ 0.1 X
...and once the math is all done, we are left with the **Predicted Weight**.
Height (m) | Favorite Color | Gender | Weight (kg)
1.7 | Green | Female | 70
...etc...etc...etc...

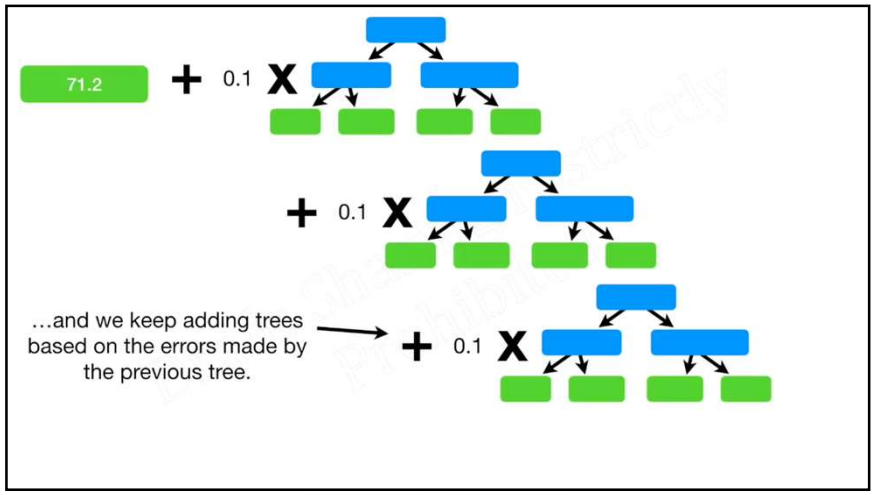

In summary, when **Gradient Boost** is used for **Regression**...

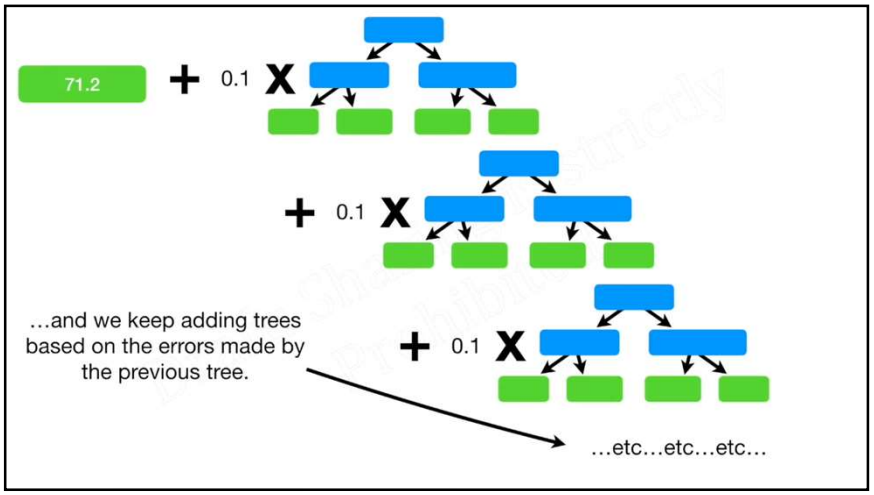

71.2
…we start with a leaf that is the average value of the variable we want to **Predict**.

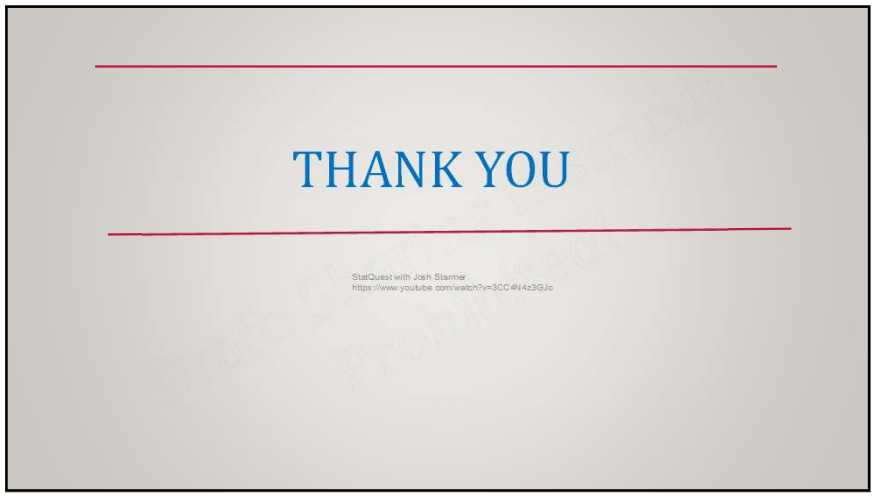

71.2
+
0.1
X
Then we add a tree based on the **Residuals**, the difference between the **Observed** values and the **Predicted** values…


71.2
+
0.1
X
…and we scale the tree's contribution to the final **Prediction** with a **Learning Rate**.


71.2
+
0.1
X
+
0.1
X
Then we add another tree based on the new **Residuals**…


71.2
+ 0.1 X
+ 0.1 X
…and we keep adding trees based on the errors made by the previous tree.
+ 0.1 X


71.2
+ 0.1 X
+ 0.1 X
…and we keep adding trees based on the errors made by the previous tree.
+ 0.1 X
…etc…etc…etc…


THANK YOU
StatQuest with Josh Starmer
https://www.youtube.com/watch?v=3CC4N4z3GJc