We have collected the data, now what?

We have collected the data, now what?

Goal:

We have collected the data, now what?

Goal:

What we want to learn from the data

We have collected the data, now what?

Goal: We want to understand the effect of light on the height of the plants

We have collected the data, now what?

Goal: We want to understand the effect of light on the height of the plants

Hypothesis:

We have collected the data, now what?

Goal: We want to understand the effect of light on the height of the plants

Hypothesis:

What we think the outcome will be

We have collected the data, now what?

Goal: We want to understand the effect of light on the height of the plants

Hypothesis:

We think that seeds exposed to more light will be taller because plants need light

We have collected the data, now what?

Goal: We want to understand the effect of light on the length of the plants

Hypothesis:

We think that seeds exposed to more light will be taller because plants need light

We think that light has a positive effect on the height of the plant

We have collected the data, now what?

Goal: We want to understand the effect of light on the length of the plants

Hypothesis:

We think that seeds exposed to more light will be taller because plants need light

We think that light has a positive effect on the height of the plant

We think that plants not exposed to enough light will be short

Which are our variables of interest?

Which are our variables of interest?

The variables of interest depend on our hypothesis

Which are our variables of interest?

The variables of interest depend on our hypothesis

Hypothesis:

We think that seeds exposed to more light will be taller because plants need light

We think that light has a positive effect on the height of the plant

We think that plants not exposed to enough light will be short

Which are our variables of interest?

The variables of interest depend on our hypothesis

Hypothesis:

We think that seeds exposed to more light will be taller because plants need light

We think that light has a positive effect on the height of the plant

We think that plants not exposed to enough light will be short

Which are our variables of interest?

The variables of interest depend on our hypothesis

Light

Plant height

Which are our variables of interest?

The variables of interest depend on our hypothesis

Light ⬅ Treatment (ecosystem)

Plant height

Now, let's look at our data

Now, let's look at our data

Plant	Dark length	Plant	Light length
1	28	1	42
2	30	2	40
3	6	3	65
4	14	4	64
5	26	5	60
6	4	6	70
7	20	7	58

Now, let's look at our data

Better

Plant	Treatment	Root_length_mm	Shoot_length_mm	Root_to_Shoot_ratio	Total_length
1	D	18	10	2	28
2	D	18	12	2	30
3	D	2	4	1	6
4	D	3	11	0	14
1	L	10	32	0	42
2	L	10	30	0	40
3	L	30	35	1	65
4	L	32	32	1	64

Good practices storing data in spreadsheets

1. Save file as csv
2. Name file without spaces or special characters
3. Meaningful column names without special characters
4. No empty cells
5. One thing per cell
6. Create a data dictionary
7. No calculations in the raw data files
8. Do not use font/color to highlight data

Good practices storing data in spreadsheets

1. Save file as csv
2. Name file without spaces or special characters
3. Meaningful column names without special characters
4. No empty cells
5. One thing per cell
6. Create a data dictionary
7. No calculations in the raw data files
8. Do not use font/color to highlight data

Not meant to be human-readable, but computer-readable

Good practices storing data in spreadsheets

1. Save file as csv
2. Name file without spaces or special characters
3. Meaningful column names without special characters
4. No empty cells
5. One thing per cell
6. Create a data dictionary
7. No calculations in the raw data files
8. Do not use font/color to highlight data


THE AMERICAN STATISTICIAN
2018, VOL. 72, NO. 1, 2–10
https://doi.org/10.1080/00031305.2017.1375989

Data Organization in Spreadsheets

Karl W. Broman[a] and Kara H. Woo[b]

Now, let's look at our data

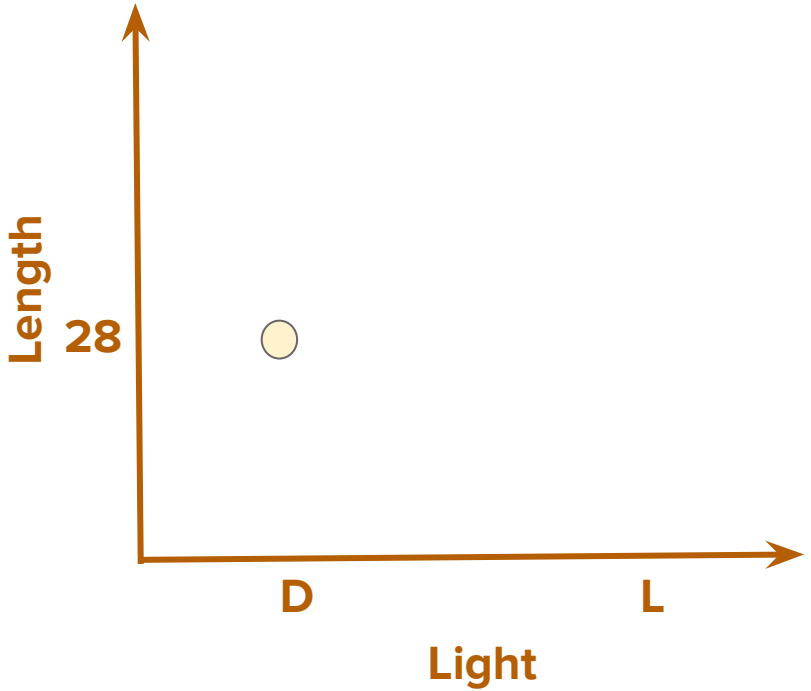
Plant	Treatment	Root_length_mm	Shoot_length_mm	Root_to_Shoot_ratio	Total_length
1	D	18	10	2	28
2	D	18	12	2	30
3	D	2	4	1	6
4	D	3	11	0	14
1	L	10	32	0	42
2	L	10	30	0	40
3	L	30	35	1	65
4	L	32	32	1	64

Now, let's look at our data

Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64

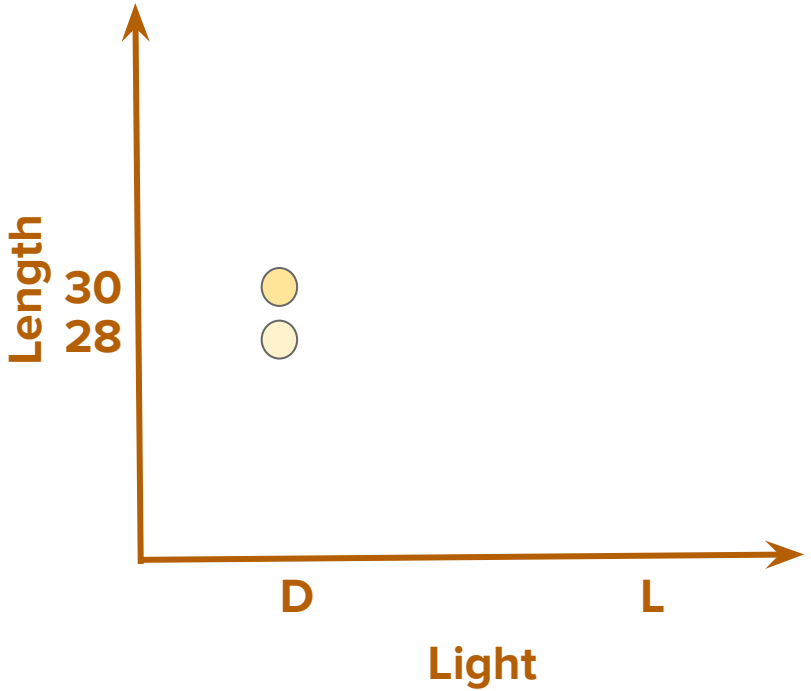
Now, let's look at our data

Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



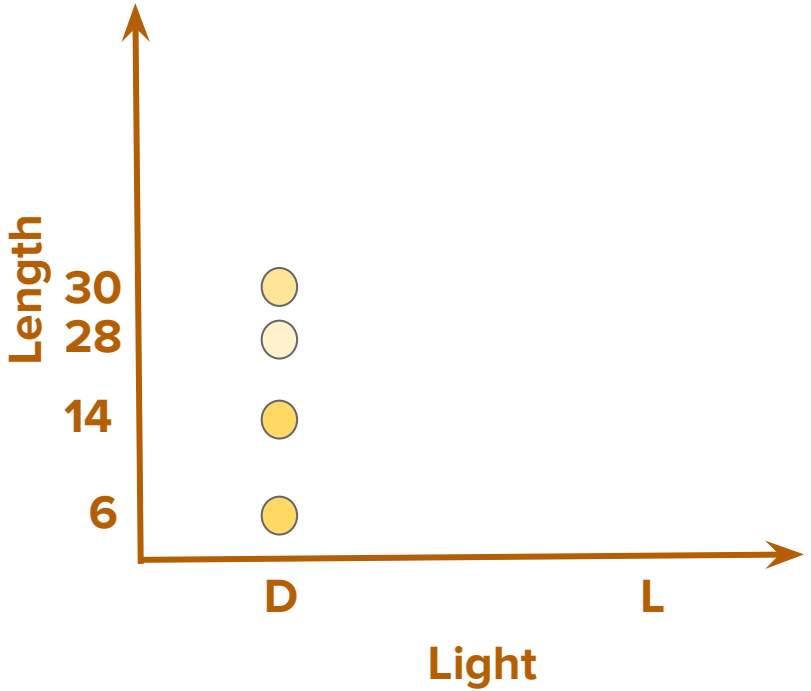
Now, let's look at our data

Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



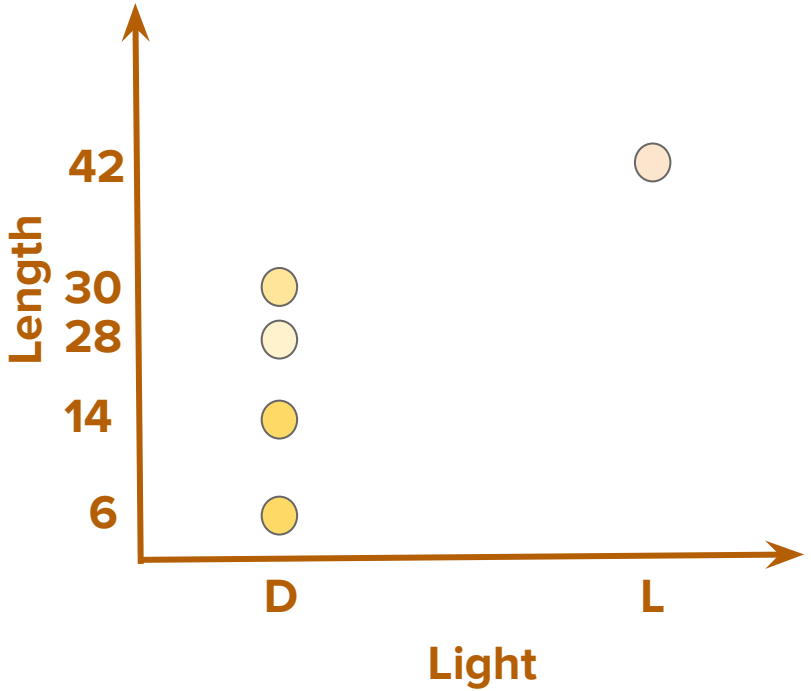
Now, let's look at our data

Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



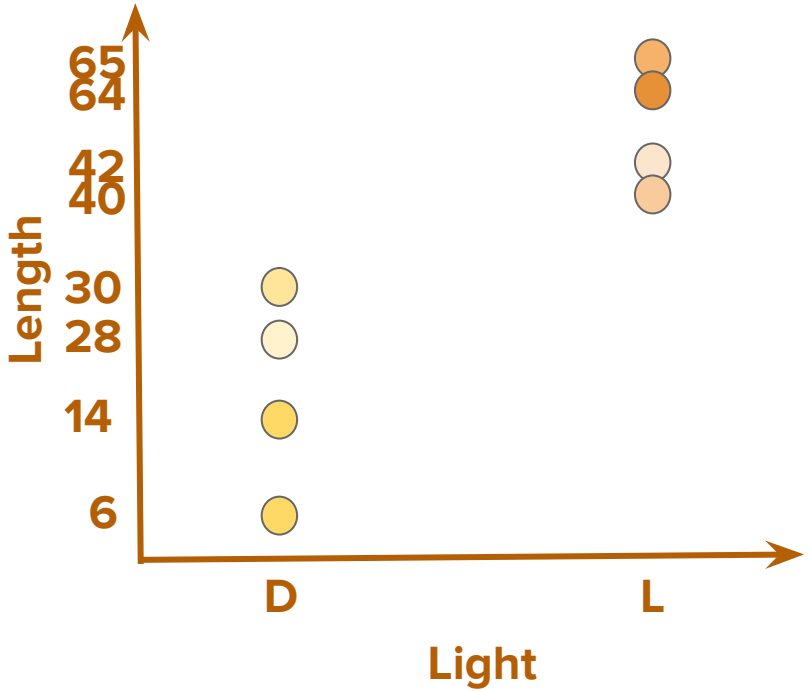
Now, let's look at our data

Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



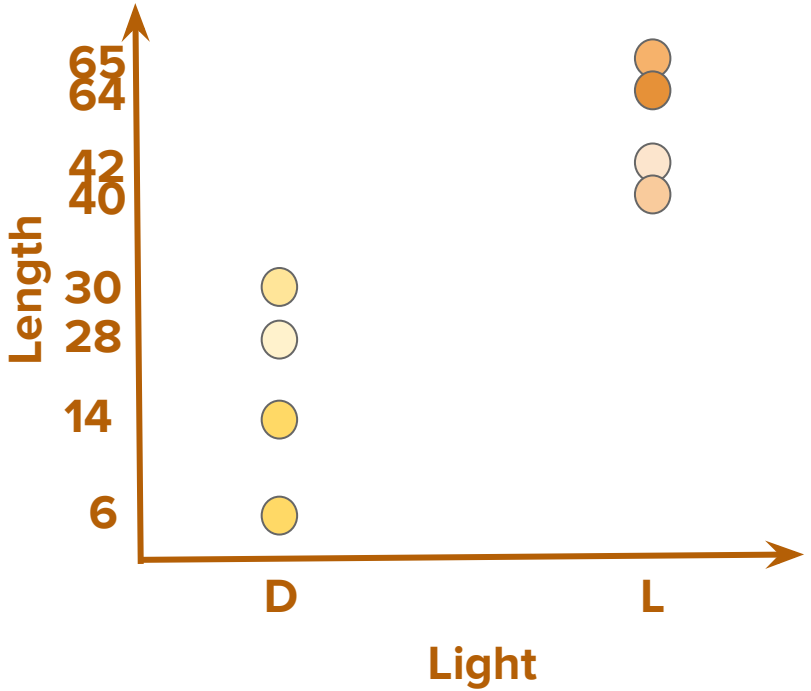
Now, let's look at our data

Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

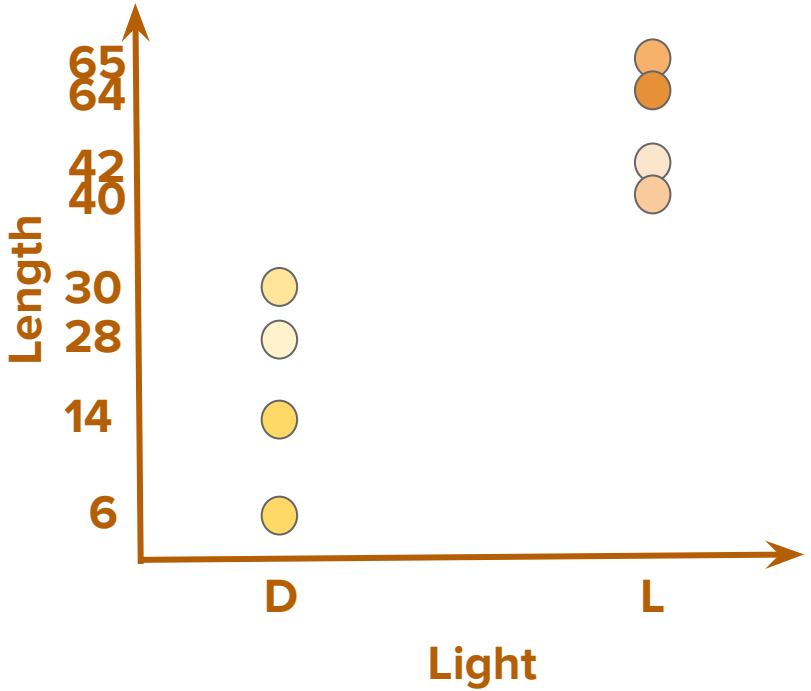
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that seeds exposed to more light will be taller

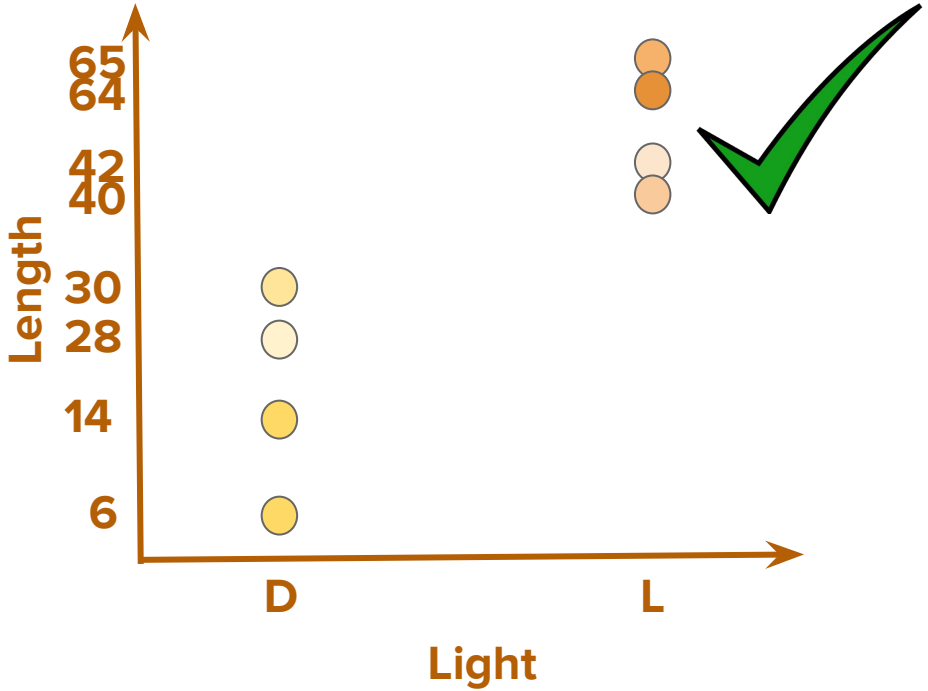
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that seeds exposed to more light will be taller

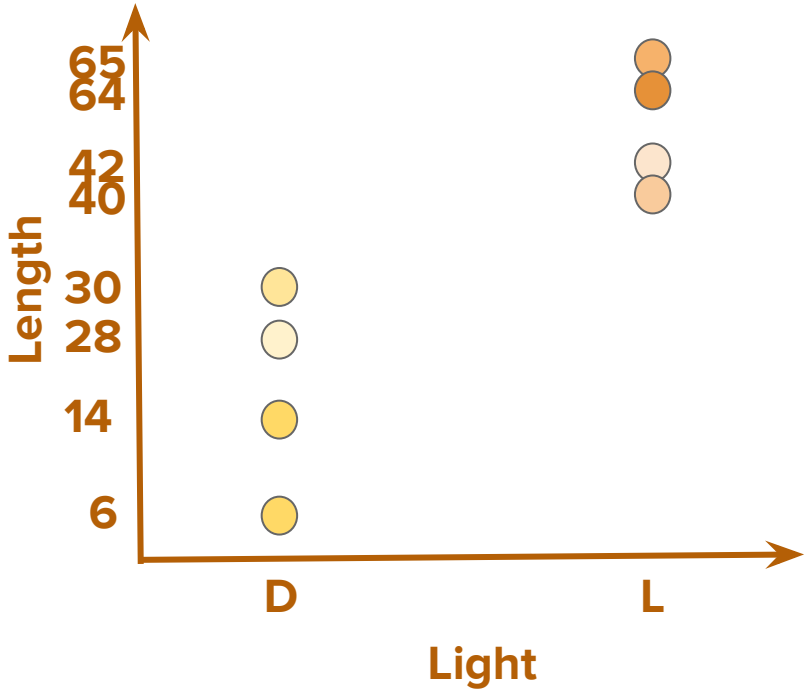
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that light has a positive effect on the height of the plant

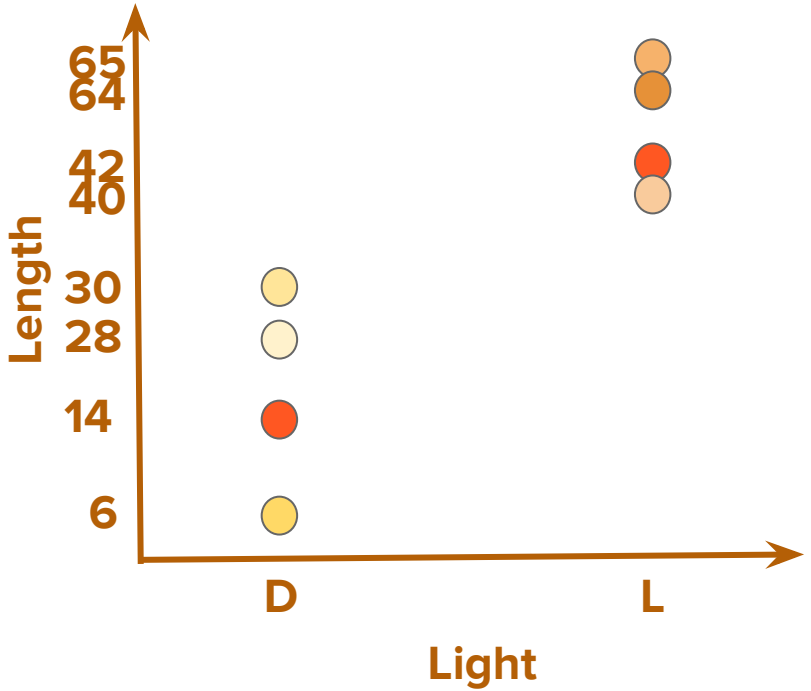
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that light has a positive effect on the height of the plant

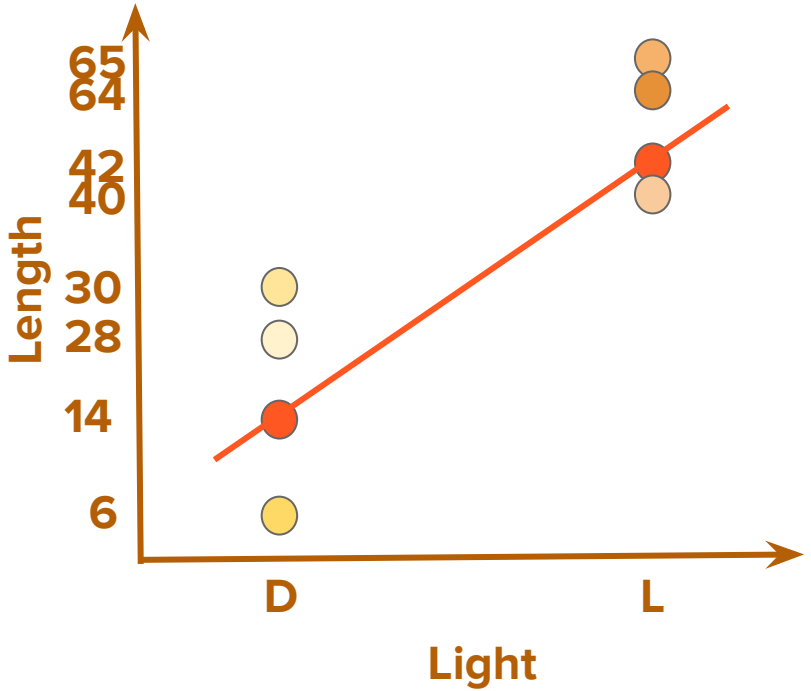
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that light has a positive effect on the height of the plant

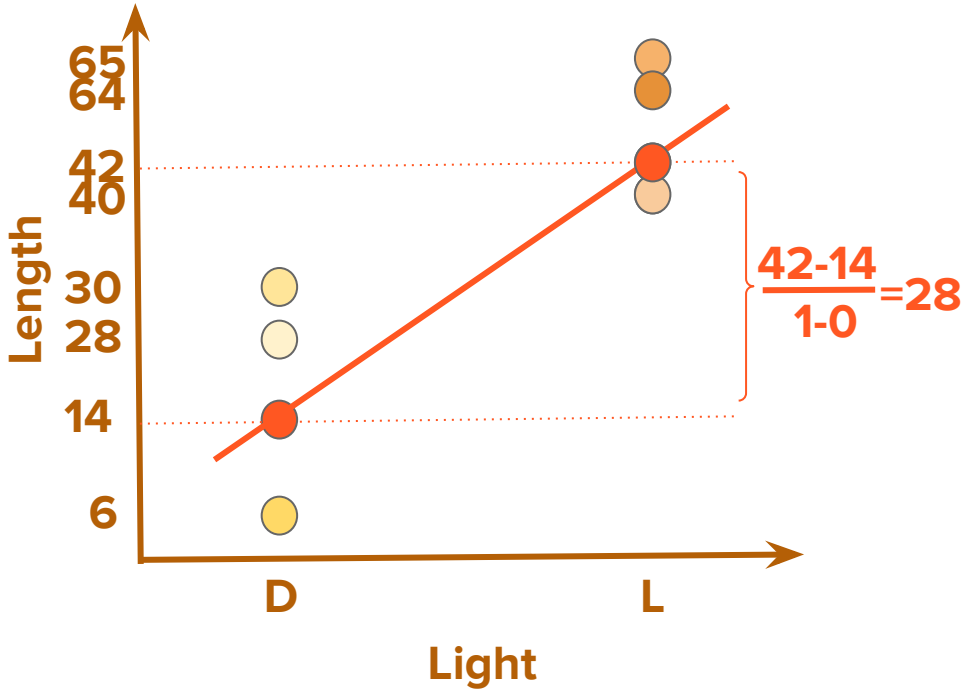
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that light has a positive effect on the height of the plant

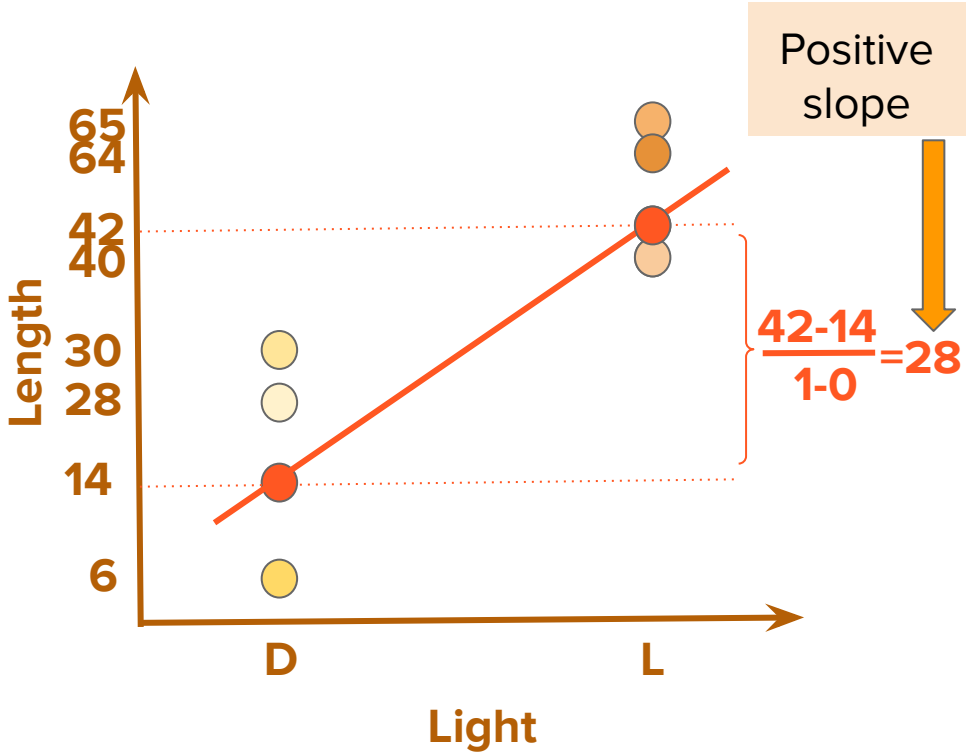
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that light has a positive effect on the height of the plant

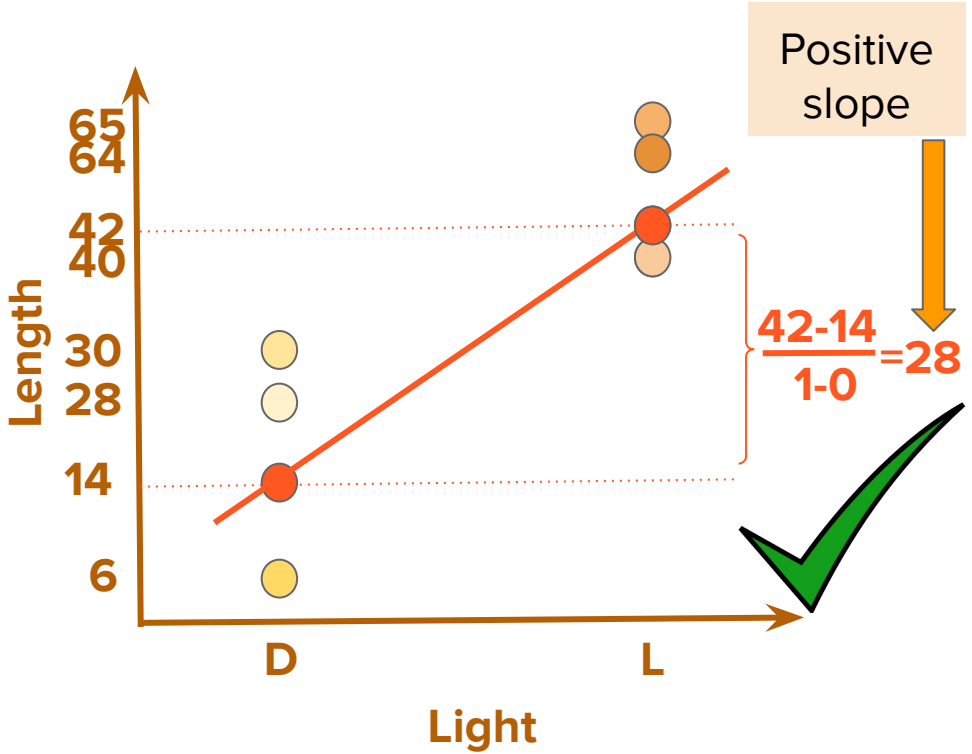
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that light has a positive effect on the height of the plant

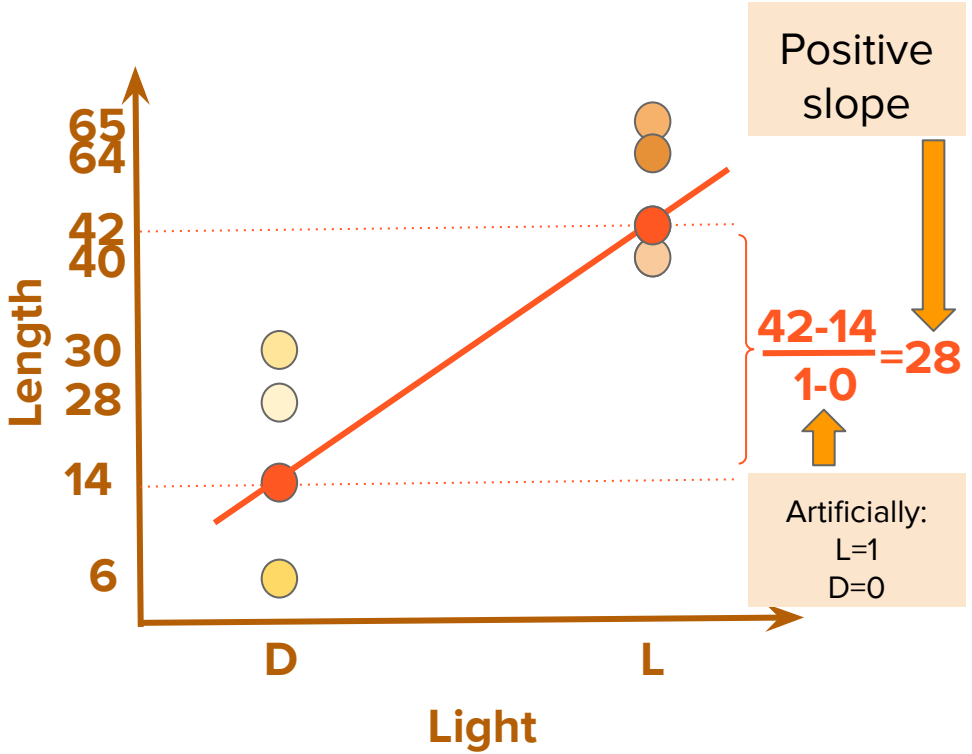
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that light has a positive effect on the height of the plant

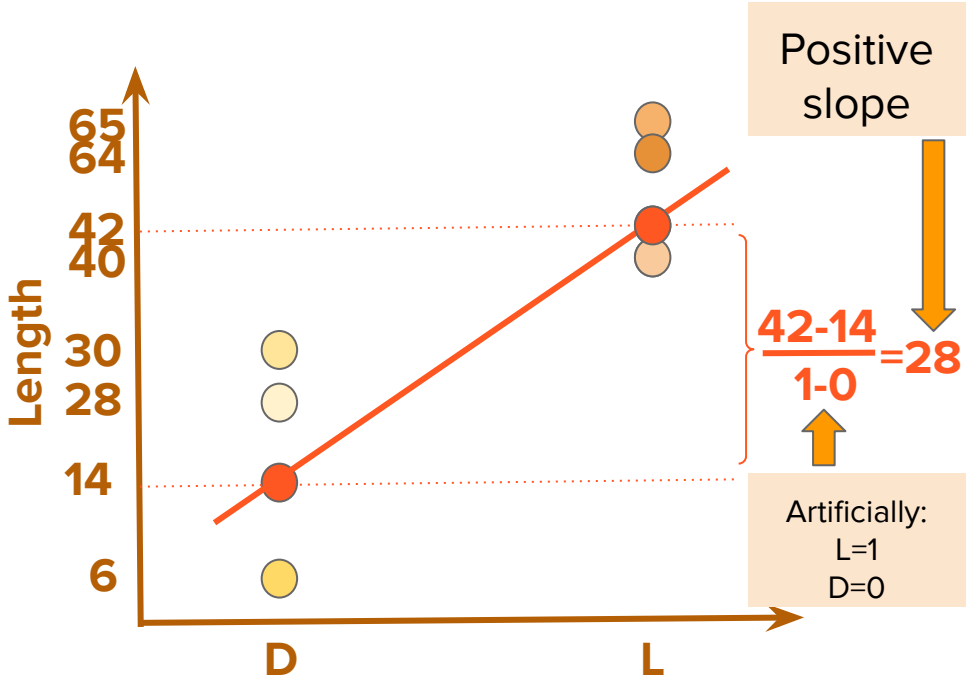
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that light has a positive effect on the height of the plant

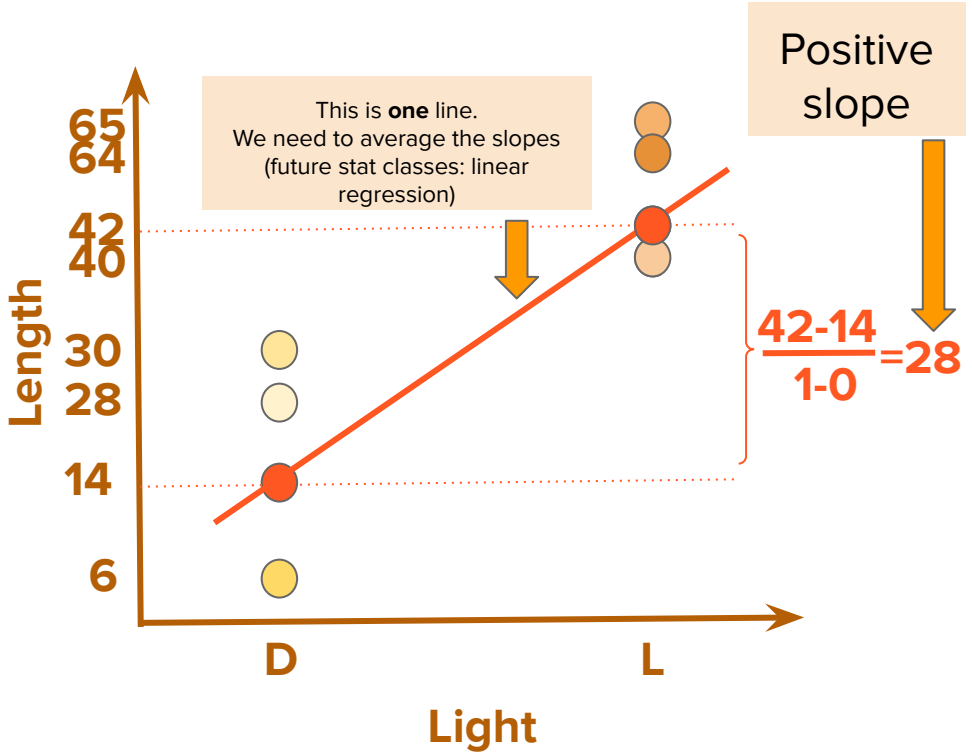
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that light has a positive effect on the height of the plant

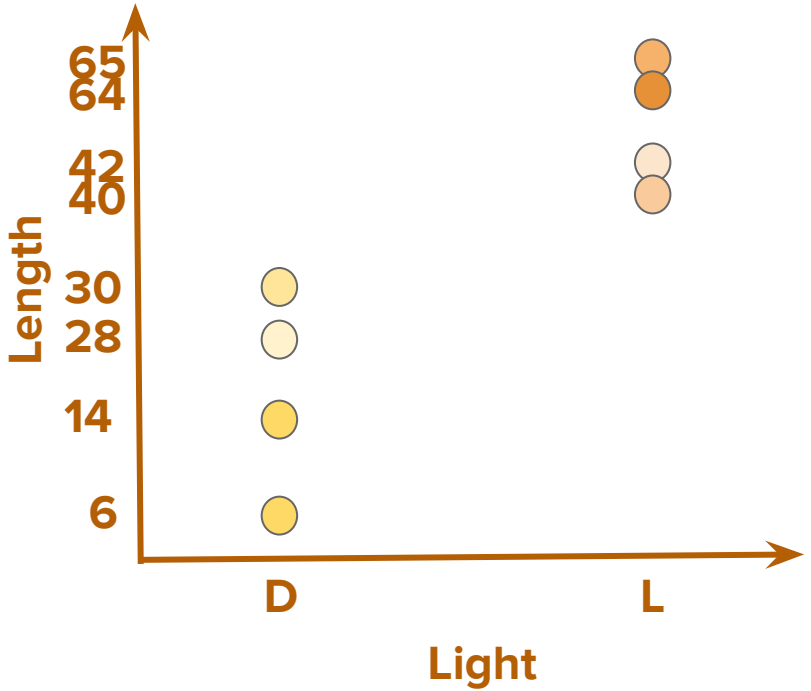
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that plants not exposed to enough light will be short

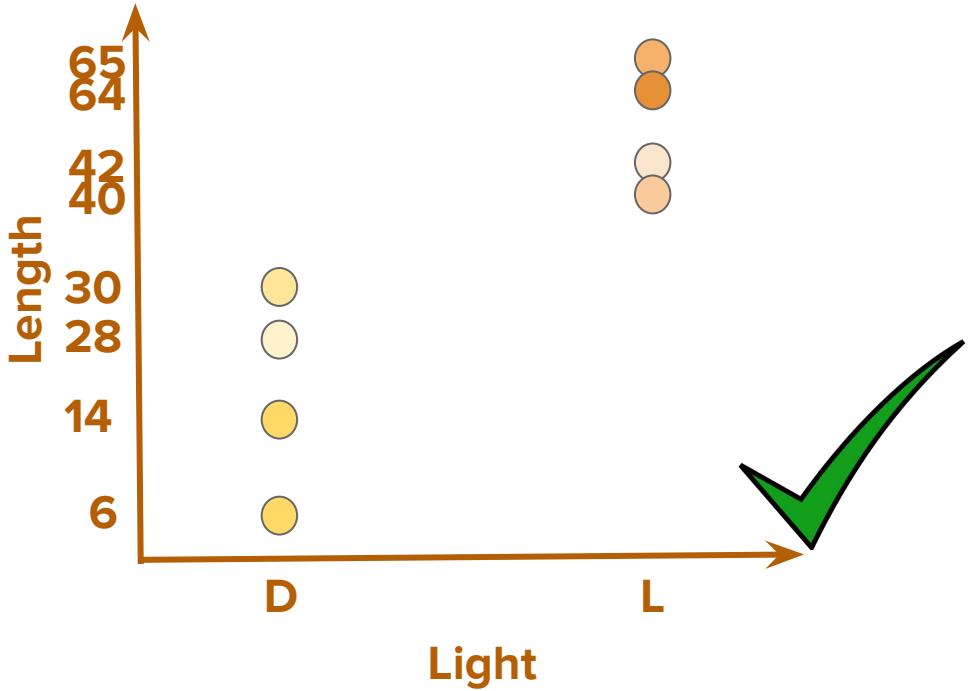
Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



Do the data agree with our hypothesis?

Hypothesis: We think that plants not exposed to enough light will be short

Plant	Treatment	Total_length
1	D	28
2	D	30
3	D	6
4	D	14
1	L	42
2	L	40
3	L	65
4	L	64



How do I make this plot with my own data?

New WI Fast Stats web app!

Publicly available and easy to use