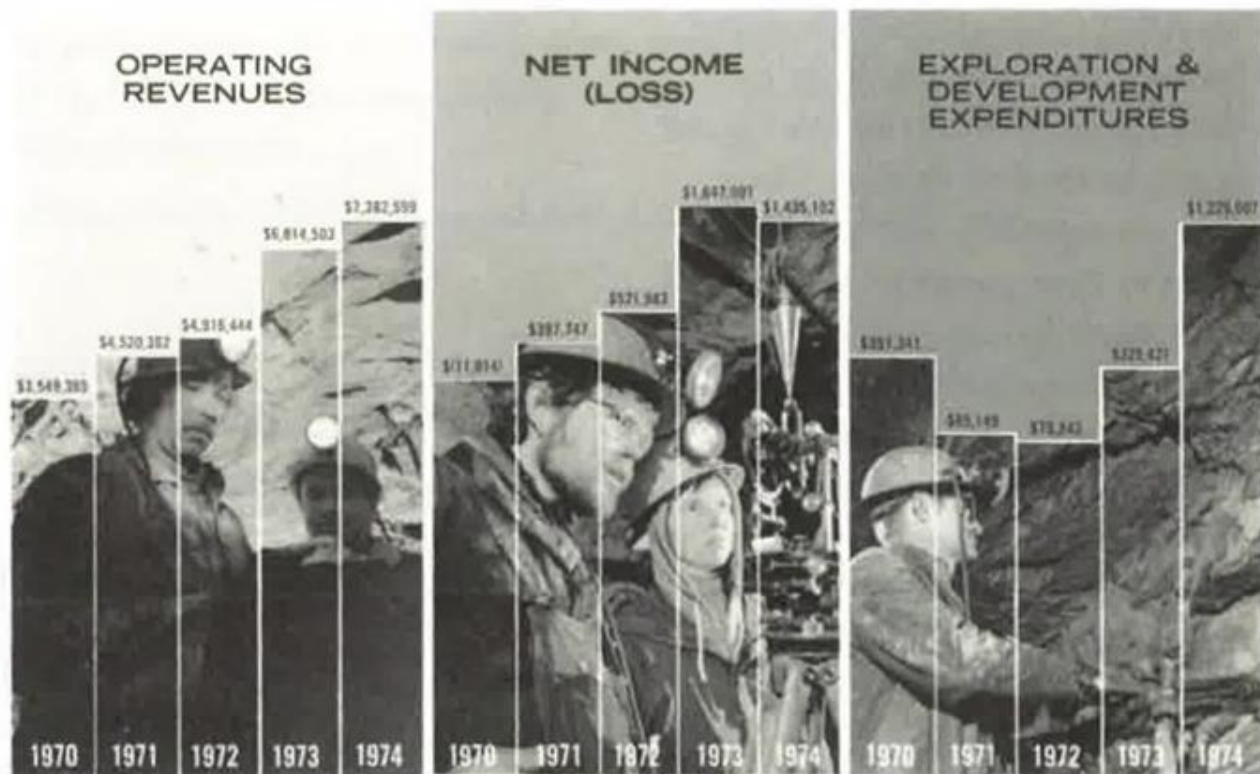
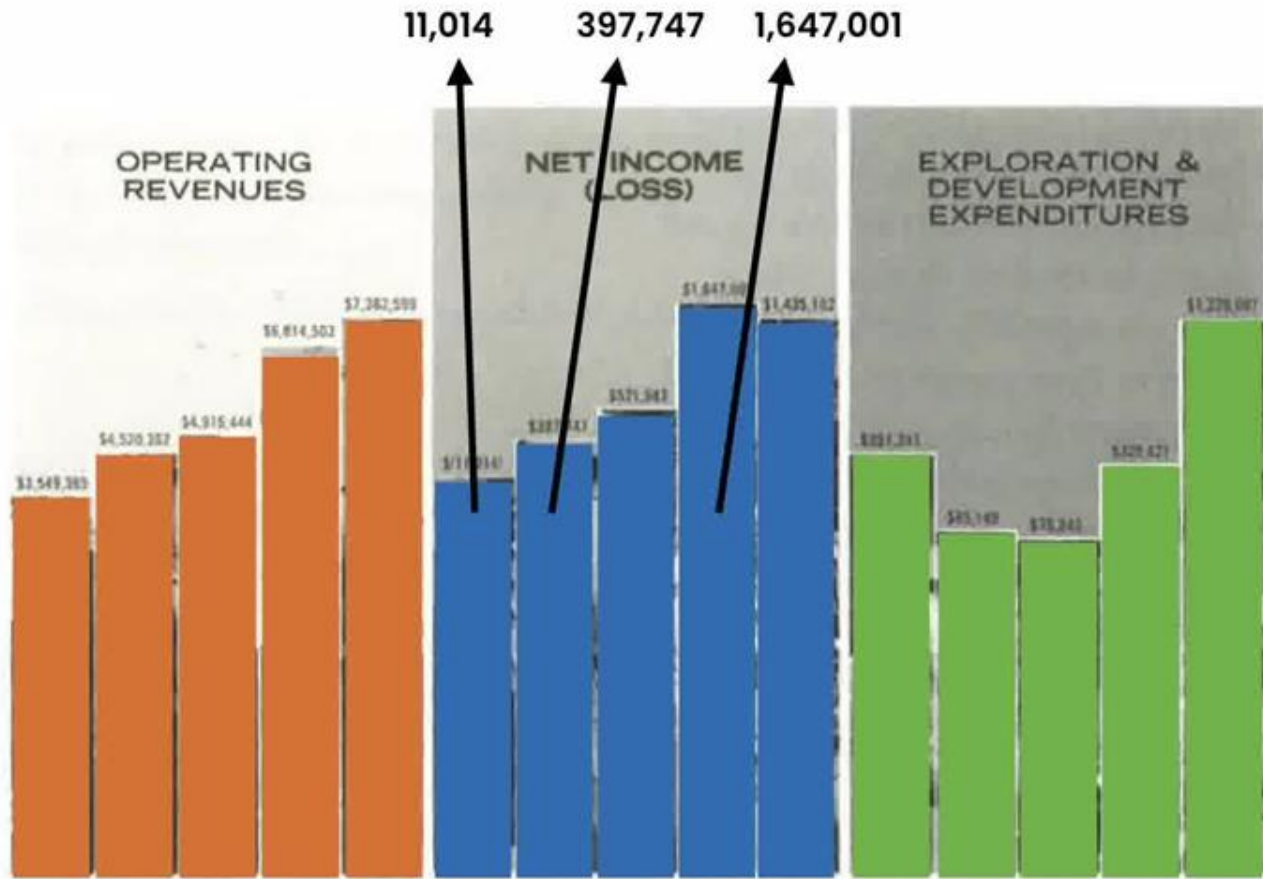


15. Graphical Integrity The Lie Factor

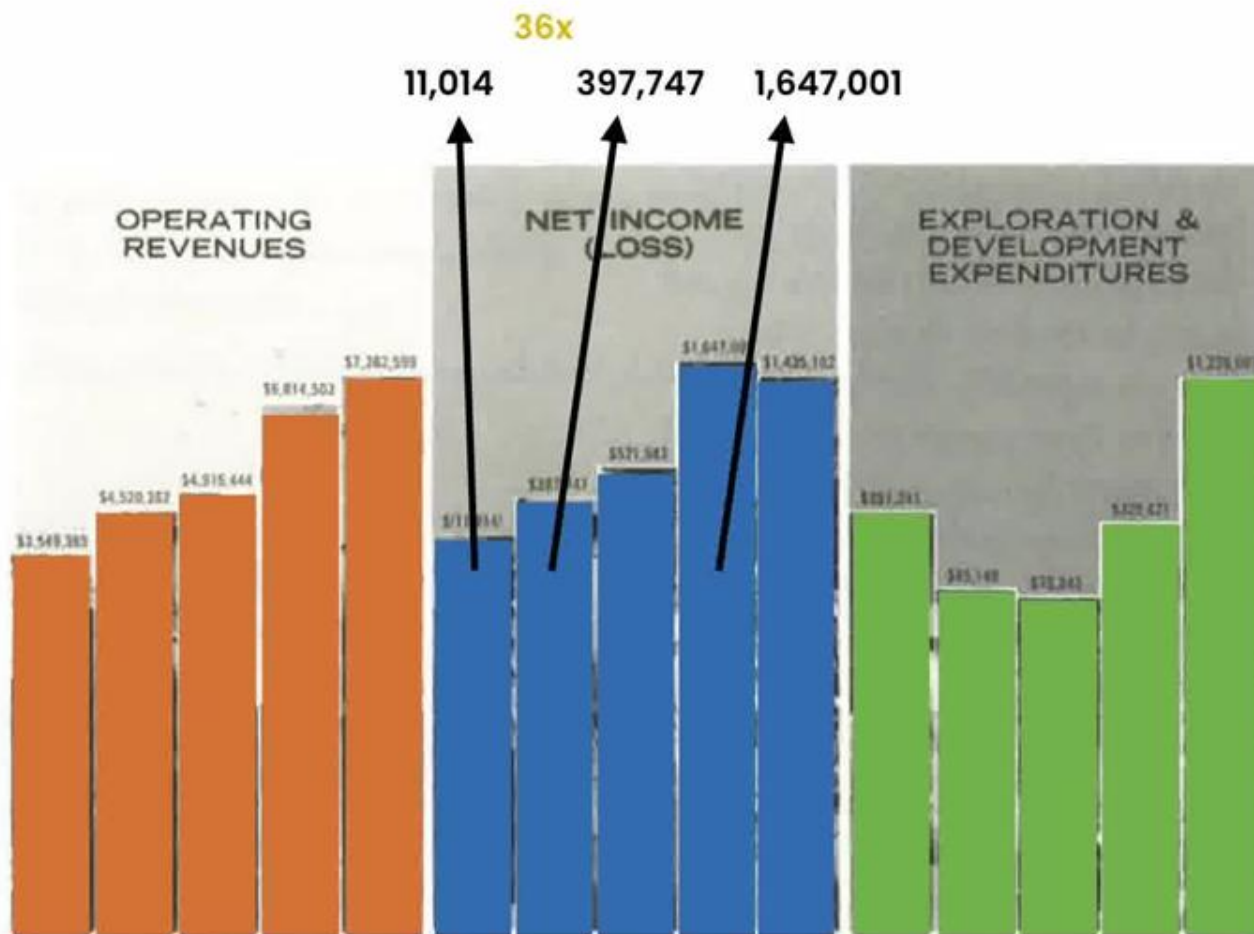
Graphical Integrity begins with
telling **the truth** about the data



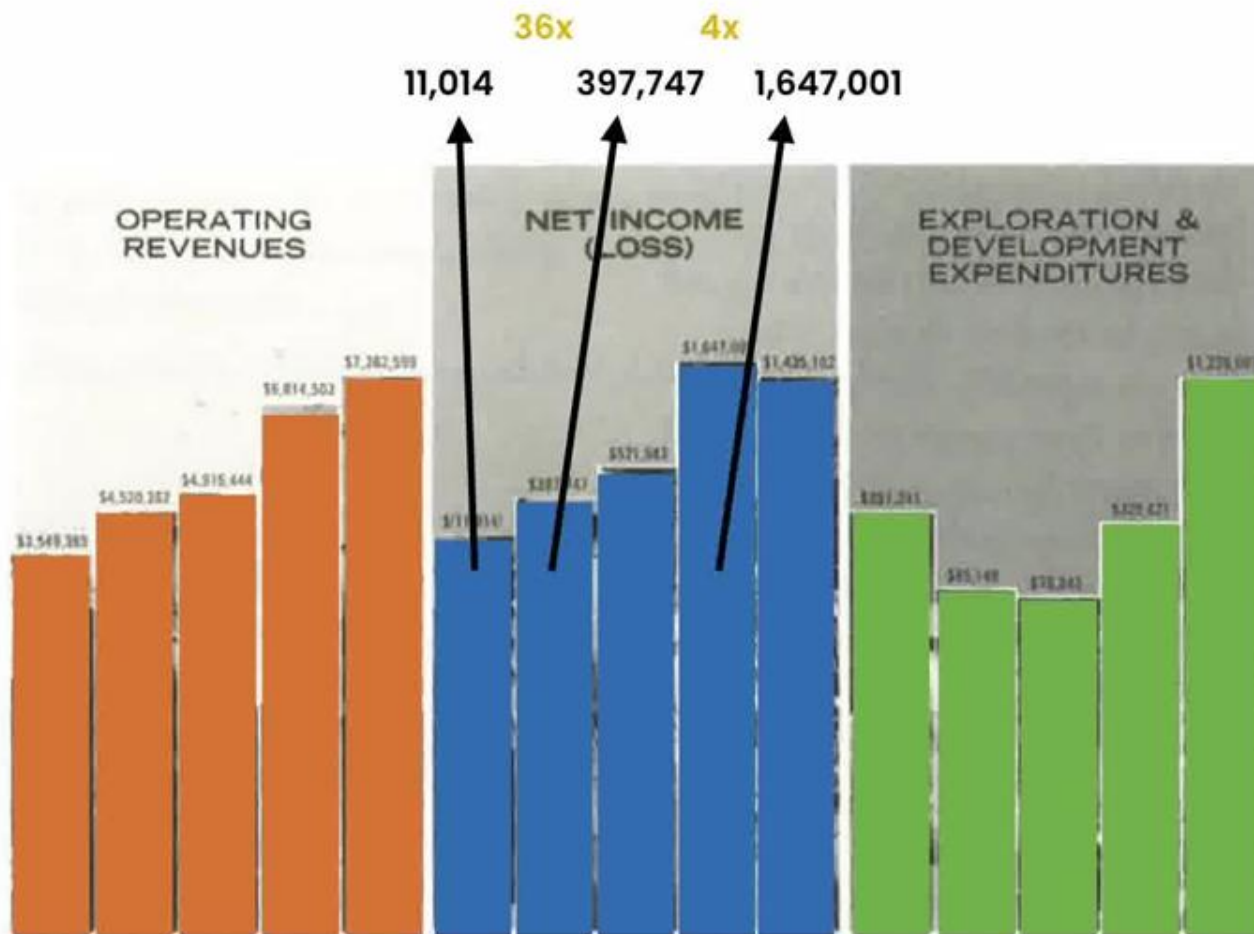
SOURCE: DAY MINES, INC., 1974 ANNUAL REPORT, P.1



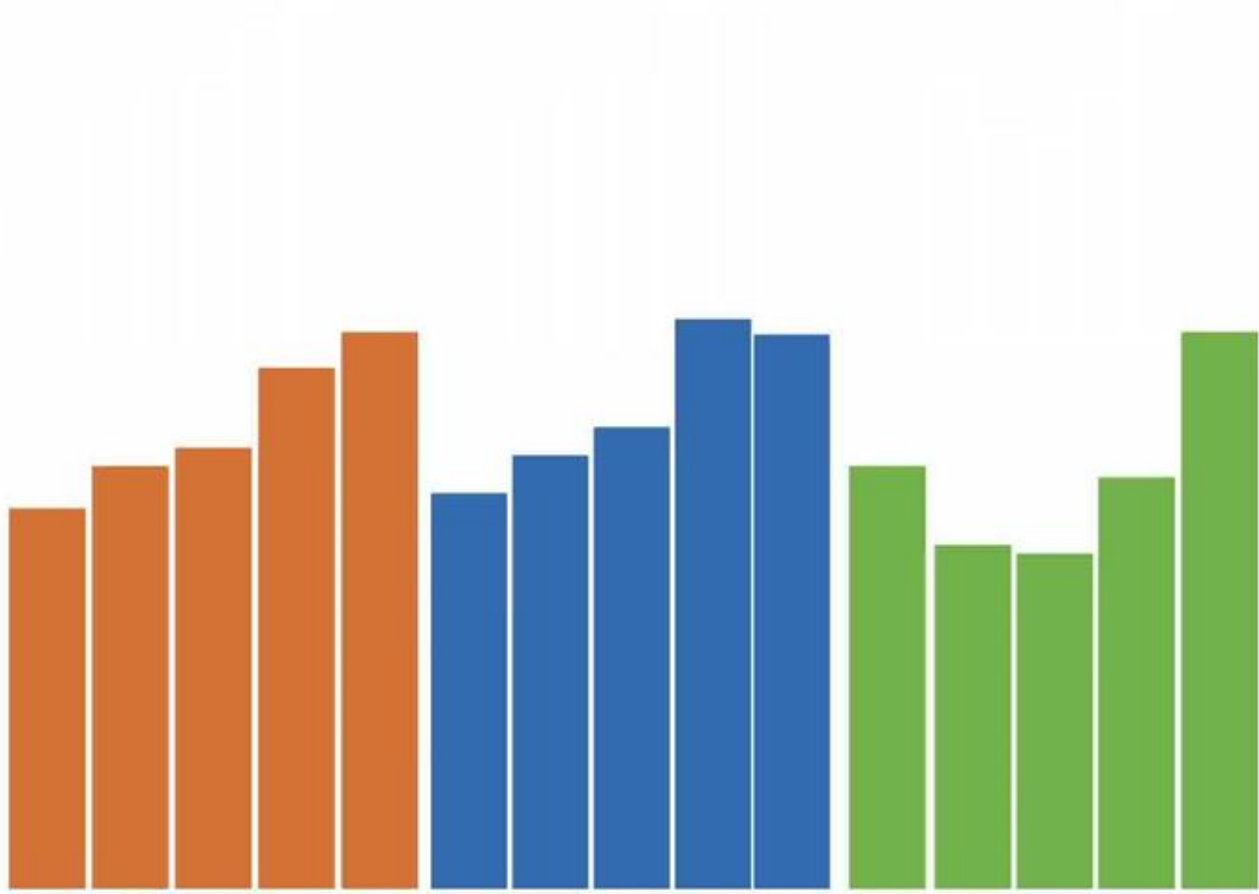
SOURCE: DAY MINES, INC., 1974 ANNUAL REPORT, P.1

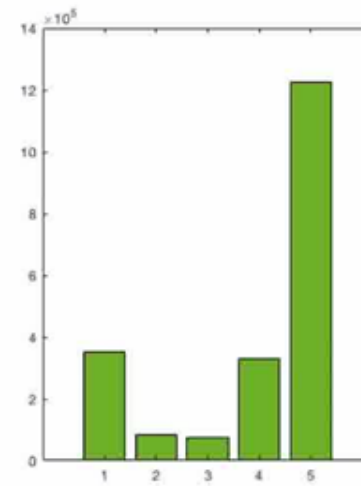
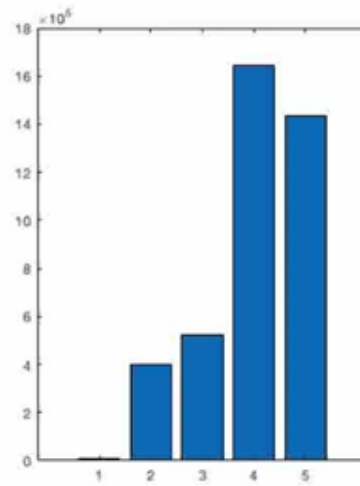
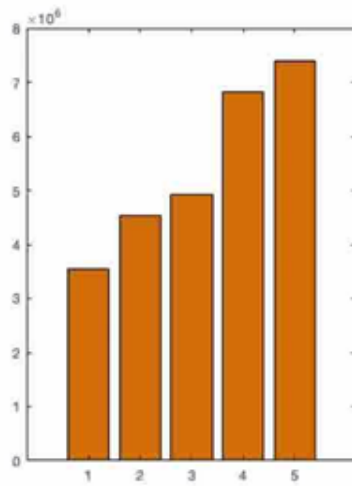
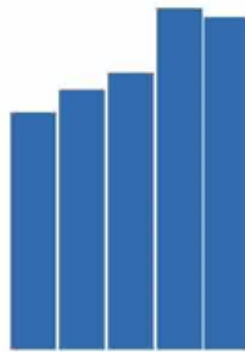
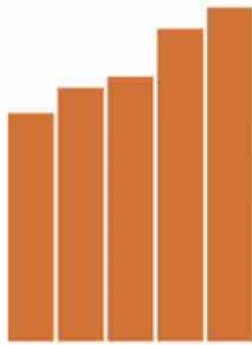


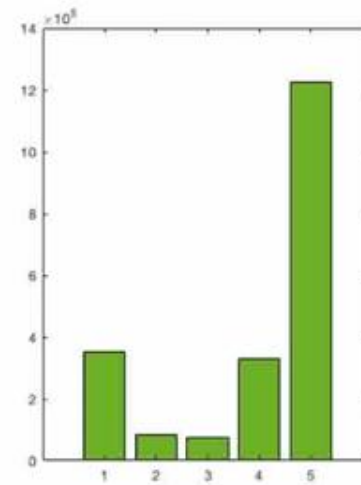
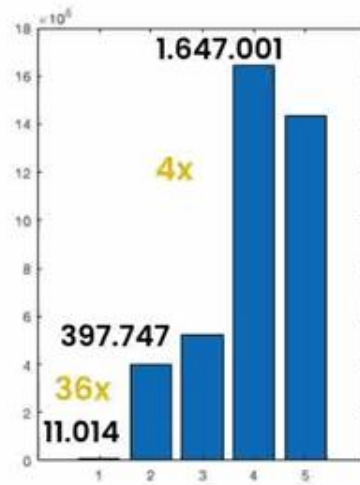
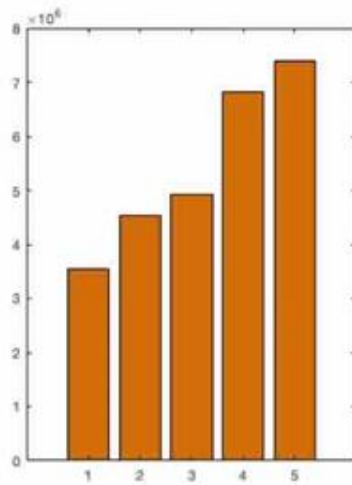
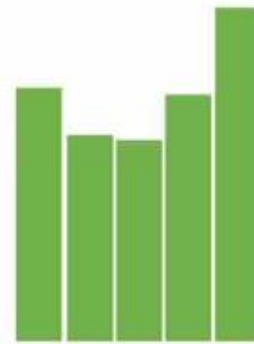
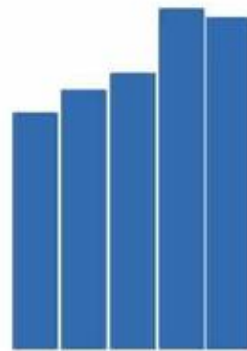
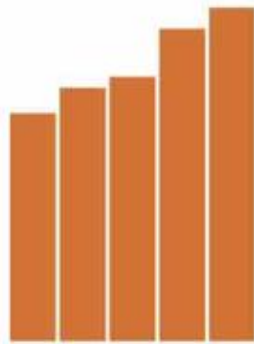
SOURCE: DAY MINES, INC., 1974 ANNUAL REPORT, P.1

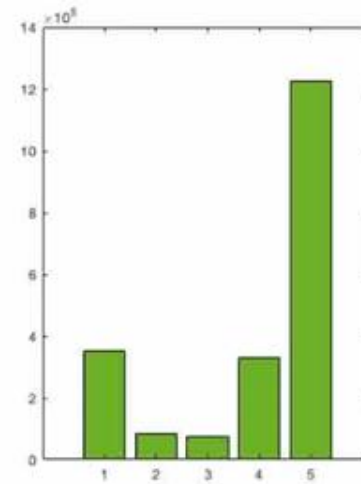
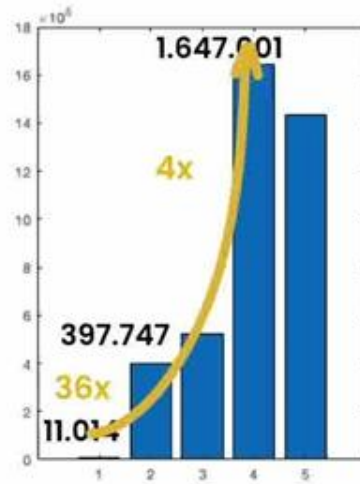
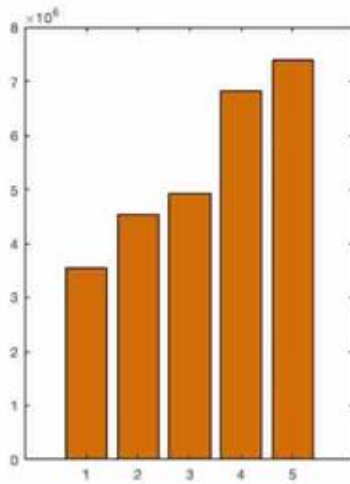
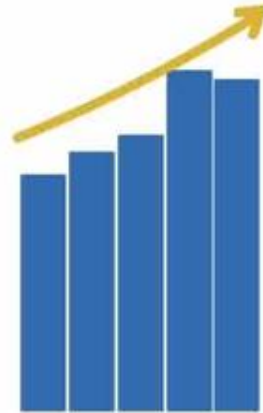
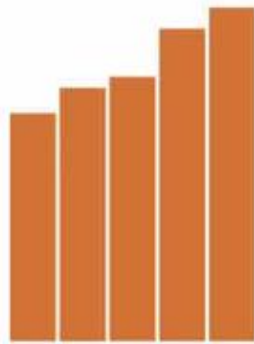


SOURCE: DAY MINES, INC., 1974 ANNUAL REPORT, P.1









Graphical Integrity 1st principle

To ensure that the visual representation of the data is perceived as accurately as possible by the user:

Graphical Integrity 1st principle

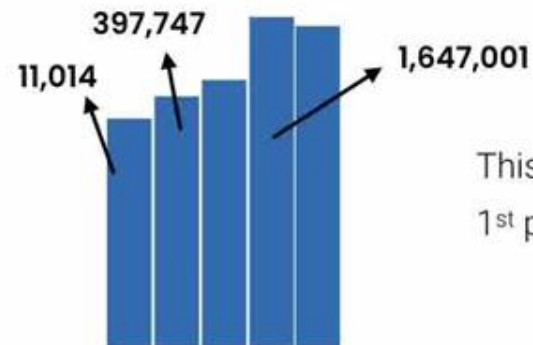
To ensure that the visual representation of the data is perceived as accurately as possible by the user:

THE REPRESENTATION OF NUMBERS, AS PHYSICALLY MEASURED ON THE SURFACE OF THE GRAPHIC ITSELF, SHOULD BE DIRECTLY PROPORTIONAL TO THE NUMERICAL QUANTITIES REPRESENTED.

Graphical Integrity 1st principle

To ensure that the visual representation of the data is perceived as accurately as possible by the user:

THE REPRESENTATION OF NUMBERS, AS PHYSICALLY MEASURED ON THE SURFACE OF THE GRAPHIC ITSELF, SHOULD BE DIRECTLY PROPORTIONAL TO THE NUMERICAL QUANTITIES REPRESENTED.



This previous example violates the 1st principle and it's totally **WRONG**.

Graphical Integrity 1st principle

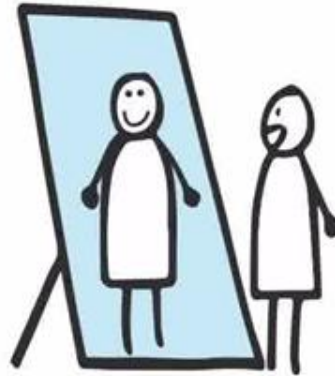
To ensure the visual representation of the data is perceived as accurately as possible by the user:

THE REPRESENTATION OF NUMBERS, AS PHYSICALLY MEASURED ON THE SURFACE OF THE GRAPHIC ITSELF, SHOULD BE DIRECTLY PROPORTIONAL TO THE NUMERICAL QUANTITIES REPRESENTED.

$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$

The lie factor

$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$

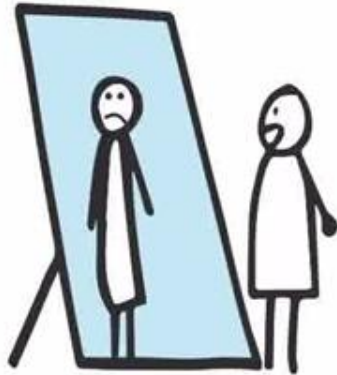


Lie Factor = 1

The representation of the data is true to the real quantitative data

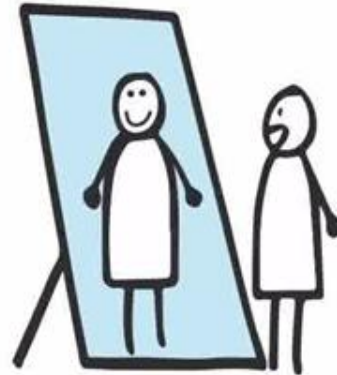
The lie factor

$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$



Lie Factor < 1

The representation of the data is under-representing the real effect

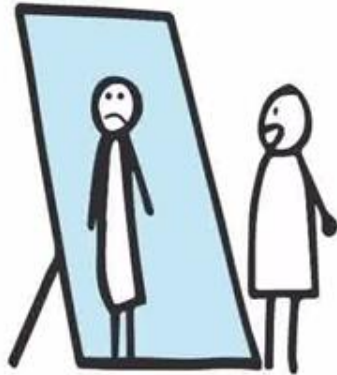


Lie Factor $= 1$

The representation of the data is true to the real quantitative data

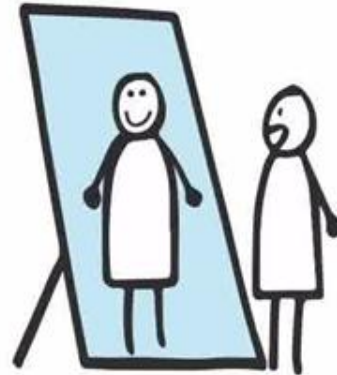
The lie factor

$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$



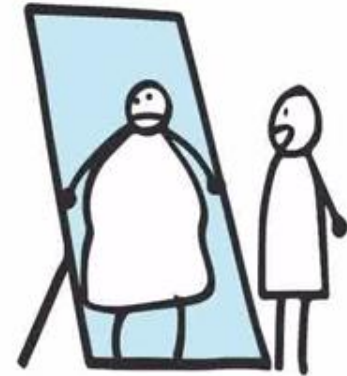
Lie Factor < 1

The representation of the data is under-representing the real effect



Lie Factor $= 1$

The representation of the data is true to the real quantitative data



Lie Factor > 1

The representation of the data is over-representing the real effect of the data

Calculation of the lie factor: example

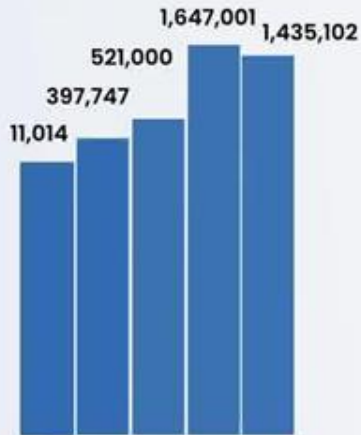
Original plot



$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$

Calculation of the lie factor: example

Original plot

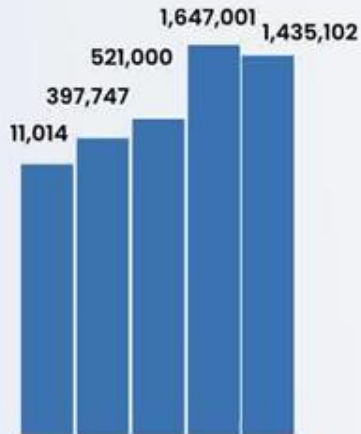


$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$

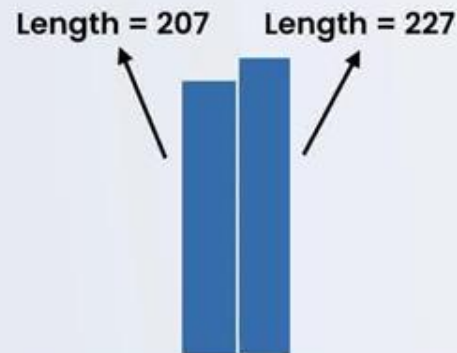
Calculation of the lie factor: example

$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$

Original plot



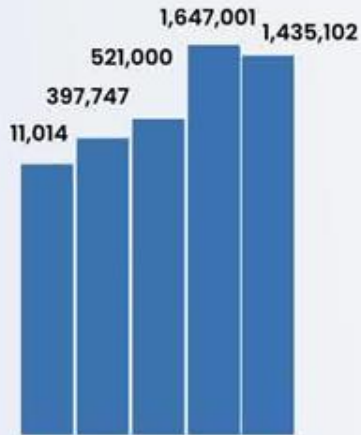
Effect shown in the graphic



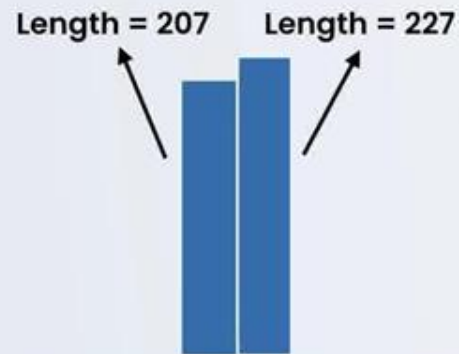
Calculation of the lie factor: example

$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$

Original plot



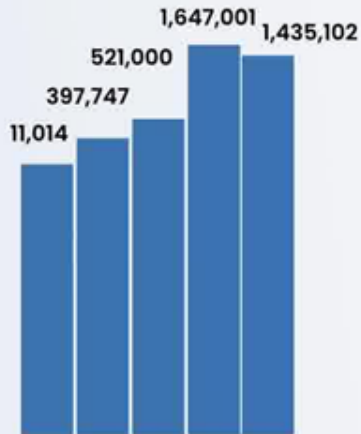
Effect shown in the graphic



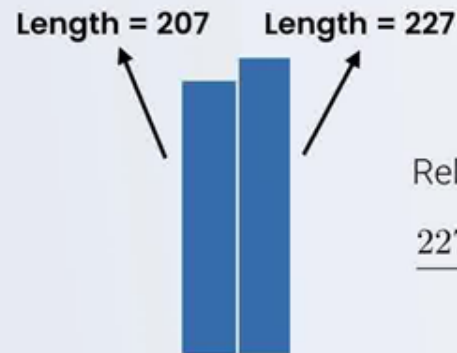
Calculation of the lie factor: example

$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$

Original plot



Effect shown in the graphic



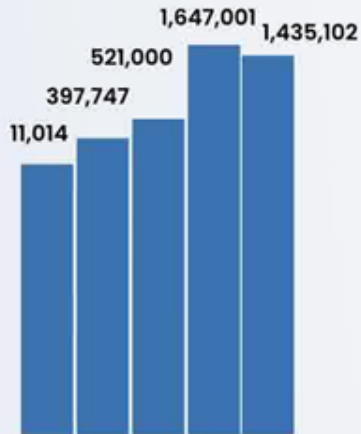
Relative change =

$$\frac{227 - 207}{207} = 0.0966$$

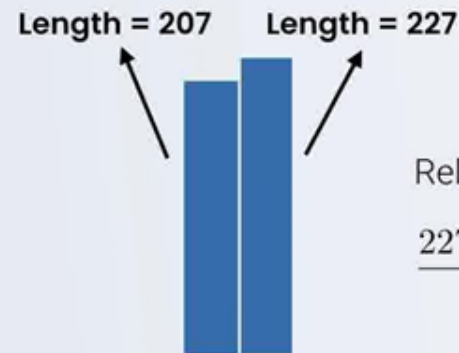
Calculation of the lie factor: example

$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$

Original plot



Effect shown in the graphic



Relative change =

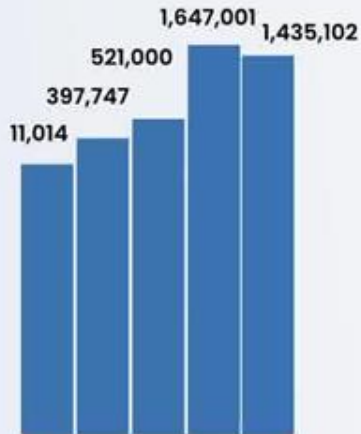
$$\frac{227 - 207}{207} = 0.0966$$

Effect in the data

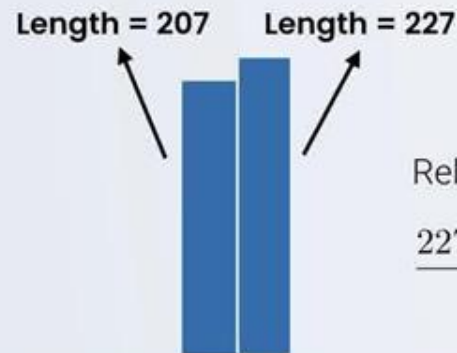
Calculation of the lie factor: example

$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$

Original plot



Effect shown in the graphic



Relative change =

$$\frac{227 - 207}{207} = 0.0966$$

Effect in the data

Data of bar 1 = 11,014

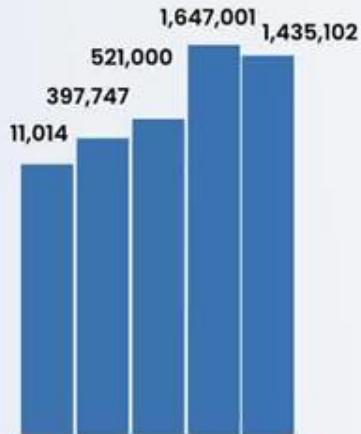
Data of bar 2 = 397,747

$$\text{Relative change} = \frac{397747 - 11014}{11014} = 35.113$$

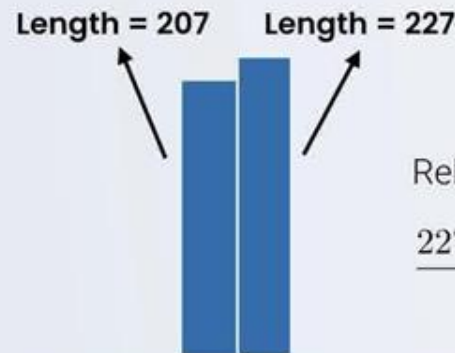
Calculation of the lie factor: example

$$\text{Lie Factor} = \frac{\text{size of effect shown in the graphic}}{\text{real size of effect in the data}}$$

Original plot



Effect shown in the graphic



Relative change =

$$\frac{227 - 207}{207} = 0.0966$$

Effect in the data

Data of bar 1 = 11,014

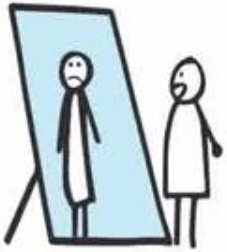
Data of bar 2 = 397,747

$$\text{Relative change} = \frac{397747 - 11014}{11014} = 35.113$$

$$\text{Lie Factor} = \frac{\text{graphic effect}}{\text{data effect}} = \frac{0.0966}{35.113} = 0.00275$$

Calculation of the lie factor: example

$$\text{Lie Factor} = \frac{\text{graphic effect}}{\text{data effect}} = \frac{0.0966}{35.113} = 0.00275$$



Lie Factor < 1

The representation of the data is under-representing the real effect



Lie Factor = 1

The representation of the data is true to the real quantitative data



Lie Factor > 1

The representation of the data is over-representing the real effect of the data

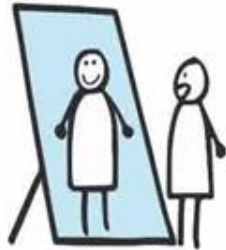
Calculation of the lie factor: example

$$\text{Lie Factor} = \frac{\text{graphic effect}}{\text{data effect}} = \frac{0.0966}{35.113} = 0.00275$$



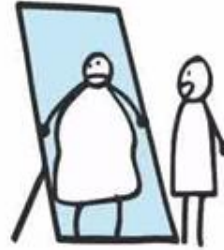
Lie Factor < 1

The representation of the data is under-representing the real effect



Lie Factor = 1

The representation of the data is true to the real quantitative data

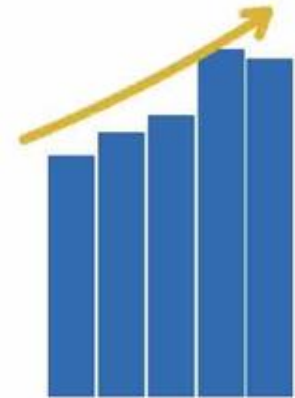


Lie Factor > 1

The representation of the data is over-representing the real effect of the data

Calculation of the lie factor: example

$$\text{Lie Factor} = \frac{\text{graphic effect}}{\text{data effect}} = \frac{0.0966}{35.113} = 0.00275$$



Lie Factor < 1

The representation of the data is under-representing the real effect



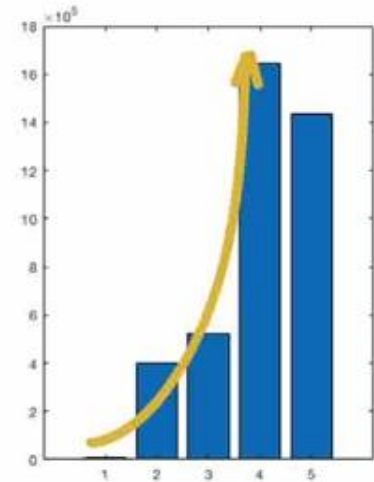
Lie Factor = 1

The representation of the data is true to the real quantitative data



Lie Factor > 1

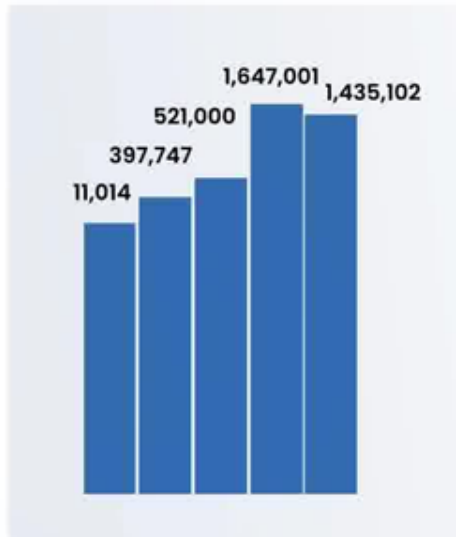
The representation of the data is over-representing the real effect of the data



A few notes

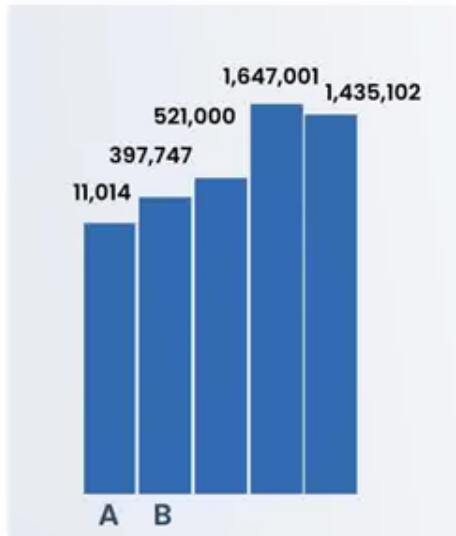
When calculating the lie factor...

The lie factor **might not be constant** along the plot.



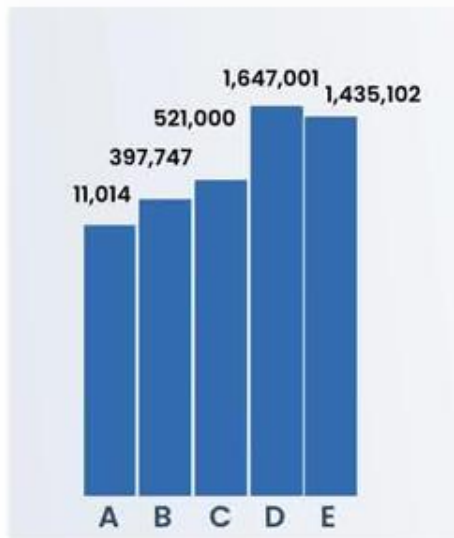
When calculating the lie factor...

The lie factor **might not be constant** along the plot.



When calculating the lie factor...

The lie factor **might not be constant** along the plot.



Lie factor (A-B) = 0.00275

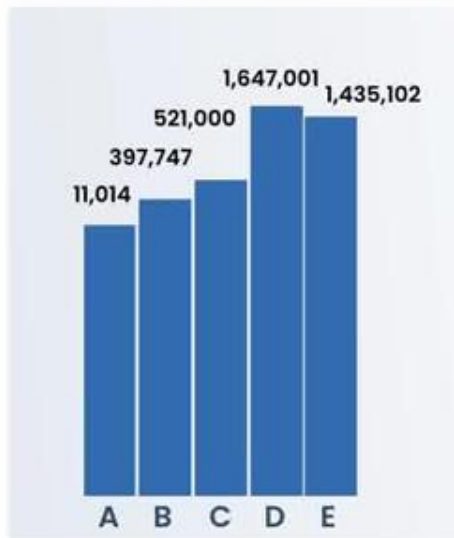
Lie factor (A-C) = 0.0092

Lie factor (A-D) = 0.0029

...

When calculating the lie factor...

The lie factor **might not be constant** along the plot.



Lie factor (A-B) = 0.00275

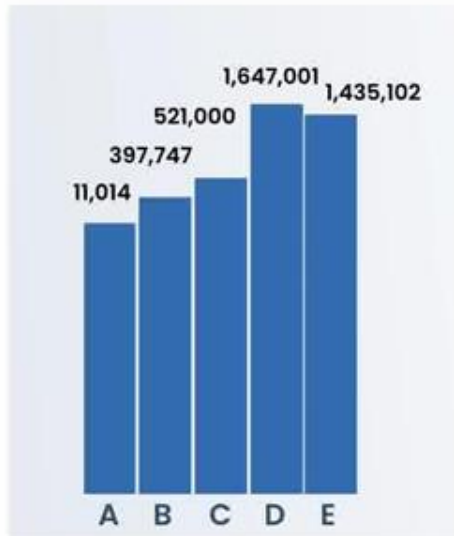
Lie factor (A-C) = 0.0092

Lie factor (A-D) = 0.0029

...

When calculating the lie factor...

The lie factor **might not be constant** along the plot.



$$\text{Lie factor (A-B)} = 0.00275$$

$$\text{Lie factor (A-C)} = 0.0092$$

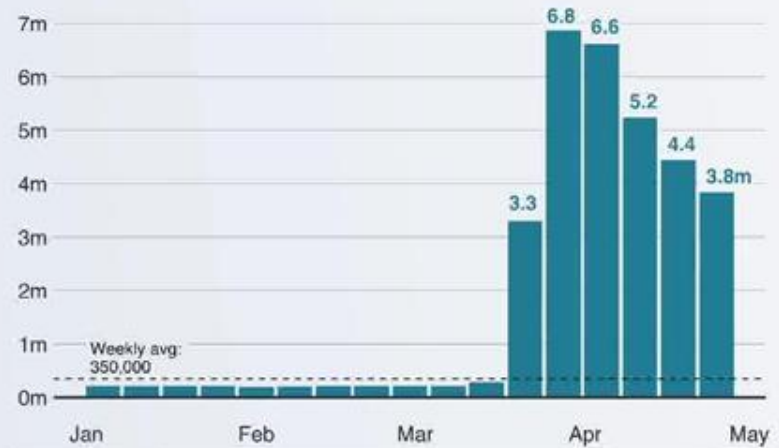
$$\text{Lie factor (A-D)} = 0.0029$$

...

In order for a graph not to lie, the lie factor should be 1 if we compare all pairs of data

More than 30m people in the US have filed for unemployment benefits in last six weeks

Weekly total of new unemployment claims in 2020



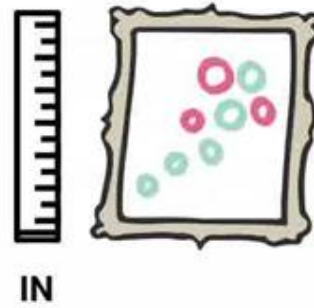
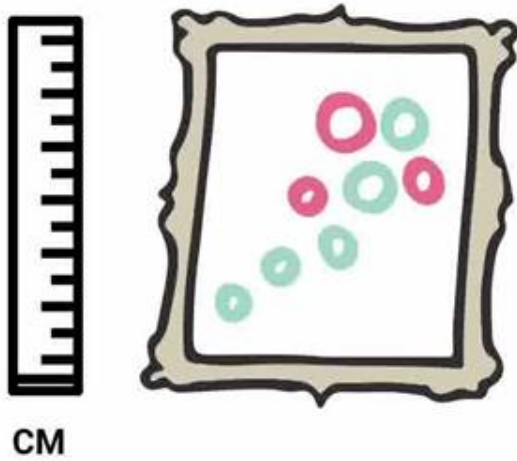
Source: US Bureau of Labor Statistics, 30 April 2020, 14:00 GMT

BBC

SOURCE: BBC GRAPHICS

When calculating the lie factor...

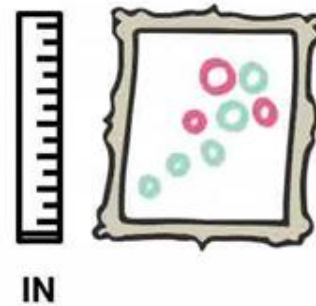
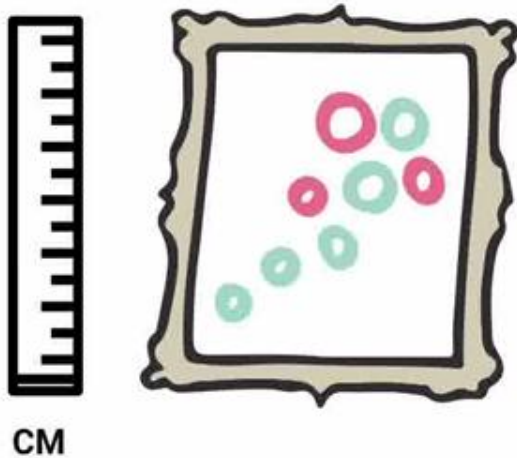
It doesn't matter if you take measurements in centimeters or inches or pixels.



When calculating the lie factor...

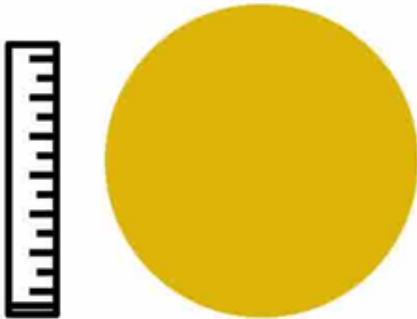
It doesn't matter if you take measurements in centimeters or inches or pixels.

Also it does not matter if you zoom in or out the image you're measuring, **only proportions matter.**



When calculating the lie factor...

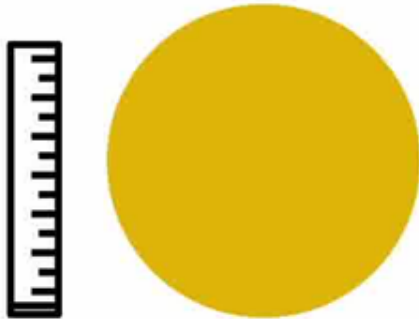
You should measure the magnitude that it's supposed to be decoded in each graph.



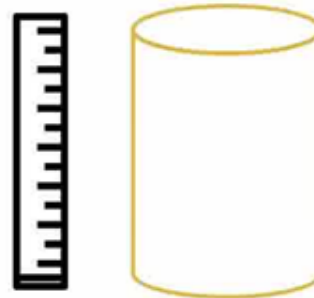
Area, not diameter or radius

When calculating the lie factor...

You should measure the magnitude that it's supposed to be decoded in each graph.



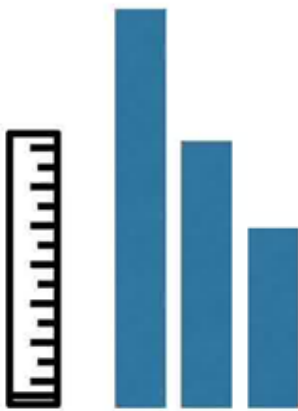
Area, not diameter or radius



Volume, not surface area or edge length or anything else

When calculating the lie factor...

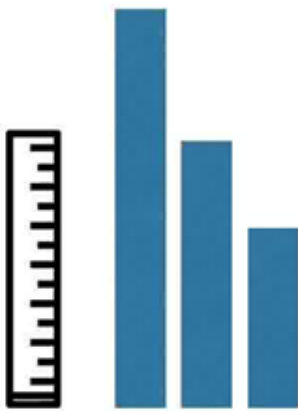
You should measure the magnitude that it's supposed to be decoded in each graph.



Length, not area

When calculating the lie factor...

You should measure the magnitude that it's supposed to be decoded in each graph.



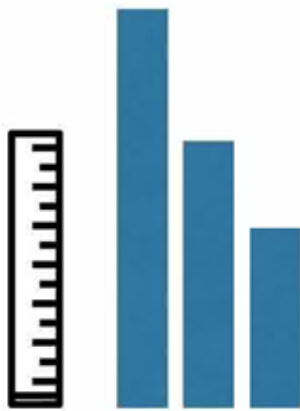
Length, not area



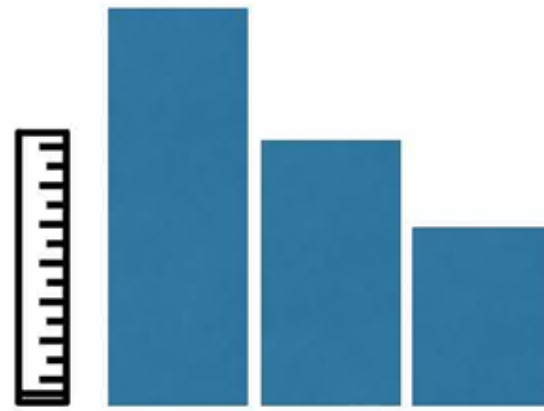
???

When calculating the lie factor...

You should measure the magnitude that it's supposed to be decoded in each graph.



Length, not area



???

Although if all bars are the same width, the result will be the same when taking length or area. Does not matter.

Your turn