Confidence Interval for One Mean

- Lesson 8.1 -

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

“I’m gonna drop out and become a…”

Musician

Ariana Grande
2020: **$72 Million**

Athlete

LeBron James
2020: **$92 Million**

Actor

Ryan Reynolds
2020: **$71 Million**

Source: Forbes

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Recently, it’s been more: “I’m gonna drop out and become a…”

+

YouTube Creator

Mr. Beast

2020: **$24 Million**

Insta Influencer

Kendall Jenner

2019: **$16 Million**

“TikToker”

Josh Richards

2020: **$1.5 Million**

Source: Forbes

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

YouTube Creator

Insta Influencer

“TikToker”

Today’s Key Analysis

Do social media creators, on average, make a livable wage?

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Lesson 8.1

Guided Notes

Handout: *skewthescript.org/8-1*

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Topics

1. Recall: sampling distribution for $\bar{x}$
2. The t-distribution and interval for $\bar{x}$
3. Four step process

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Topics

1. **Recall: sampling distribution for $\bar{x}$**
2. The t-distribution and interval for $\bar{x}$
3. Four step process

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

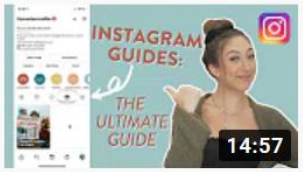
+

Skew The Script

How Much Money Do I Make on YouTube with 60K Subscribers...
Hustle at Home Mom
12K views • 1 month ago

MY FIRST *THREE* Youtube Paychecks | How Much Does ...
Madison Payne
16K views • 5 months ago

How Much YouTube Paid Me For My 2 MILLION Views Video
Liam James Kay
9.8K views • 1 month ago



NEW INSTAGRAM FEATURE GUIDES | What are Instagram...
Modern Millie
9.5K views • 1 week ago

My First YouTube Paycheck: How Much YouTube Pays +...
Journey With The Hintons
202K views • 5 months ago

This is Exactly What I Needed: Robinhood Challenge Week 8
Matthew Michael
1.2K views • 1 month ago

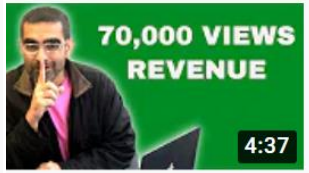
How much money did I make on YouTube??
Dr. Know-it-all Knows it all
5.7K views • 2 weeks ago

HOW LONG IT TAKES TO GET MONETIZED ON YOUTUBE: Th...
Cathrin Manning
1M views • 1 year ago

HOW MUCH MONEY $$$ I MAKE ON YOUTUBE WITH 380...
Kevin Eassa
706 views • 2 weeks ago

10 INSTAGRAM HACKS YOU DIDN'T KNOW EXISTED | How t...
Modern Millie
221K views • 2 months ago

How Much YouTube Paid Me My FIRST YEAR Being...
Cathrin Manning
248K views • 5 months ago



How Much YouTube Paid Me For 70000 Views | YouTube Ad...
Ali Mirza
292 views • 2 weeks ago

this is how much money YouTube paid me TOTAL in...
Annie Dubé
7.8K views • 1 week ago

This Is How Much YouTube Paid Me For My 1,000,000...
Shelby Church
8.8M views • 1 year ago

HOW MUCH I MADE ON YOUTUBE IN 1 MONTH WITH...
Modern Millie
4.1K views • 1 month ago

MY FIRST YOUTUBE PAYCHECK | HOW MUCH MONEY DO...
Sonja Lauren
131K views • 10 months ago

💰 My $2,000/Day Passive Income Online Store 🌞 (Case...
Wholesale Ted
256K views • 9 months ago

MY FIRST YOUTUBE PAYCHECK! How Much Money...
Alexandra Budgets
8.3K views • 8 months ago

How Much Youtube Paid me for 31 Days (30 million views)
Meet Kevin
222K views • 4 months ago

How Much Money YouTube Pays Us (125k Subscribers)
Colin and Samir
223K views • 5 months ago

MY FIRST YOUTUBE PAYCHECK / 1,000 subscribers
Samantha Piedra
55K views • 1 year ago

How Much $$ I Really Make as a Small YouTuber | Emma...
Emma Stevens
34K views • 11 months ago

how much youtube paid me for a 1 million viewed video
Reyes The Entrepreneur
3M views • 1 year ago

MY FIRST YOUTUBE PAYCHECK: How Much Money...
Morgan Stradling
12K views • 5 months ago

The Data

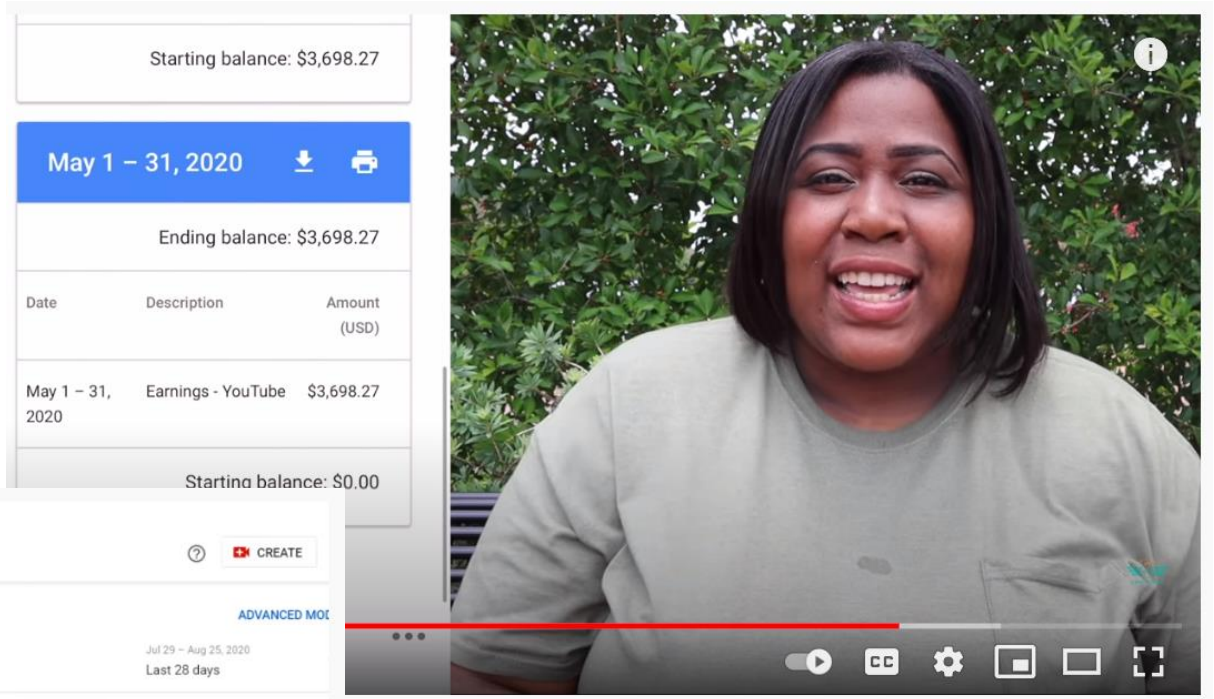
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

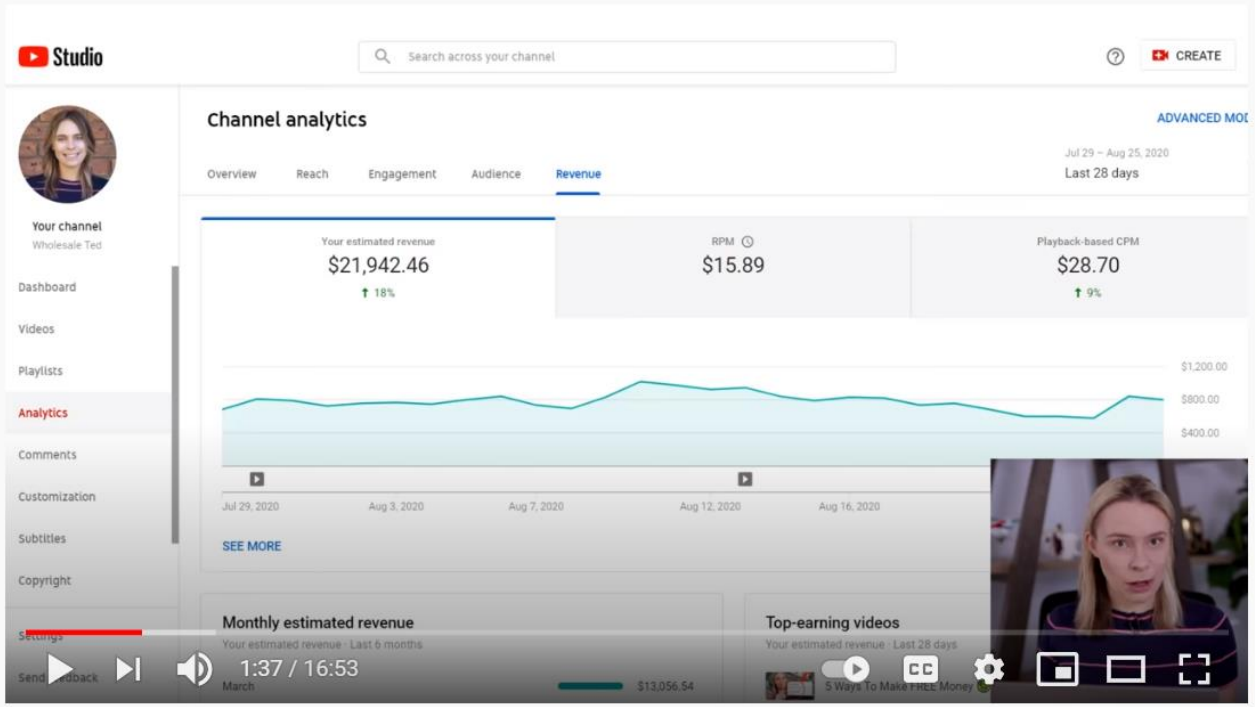
Searched “How much I make on YouTube.”

Randomly sampled 35 from hundreds of results



d #youtubepaymentprocess

k: How Much YouTube Pays + How To Get Monetized 2020 (Step



How Much Did YouTube Pay Me For 1 Million Views?! (How Much Do YouTubers REALLY Earn!)

49,475 views • Oct 1, 2020

3K 44 SHARE SAVE

Reliable data: They show their private channel revenue pages in the videos

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ vs. $\bar{x}$

μ = **population** mean

- Parameter

$\bar{x}$ = **sample** mean

- Statistic used to estimate μ

μ vs. $\bar{x}$

μ = **population** mean

- Parameter
- Ex: mean salary **among *all* YouTubers**

$\bar{x}$ = **sample** mean

- Statistic used to estimate μ

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ vs. $\bar{x}$

μ = **population** mean

- Parameter
- Ex: mean salary **among *all* YouTubers**

$\bar{x}$ = **sample** mean

- Statistic used to estimate μ
- Ex: mean salary **in our sample**

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

***In a world** where…*

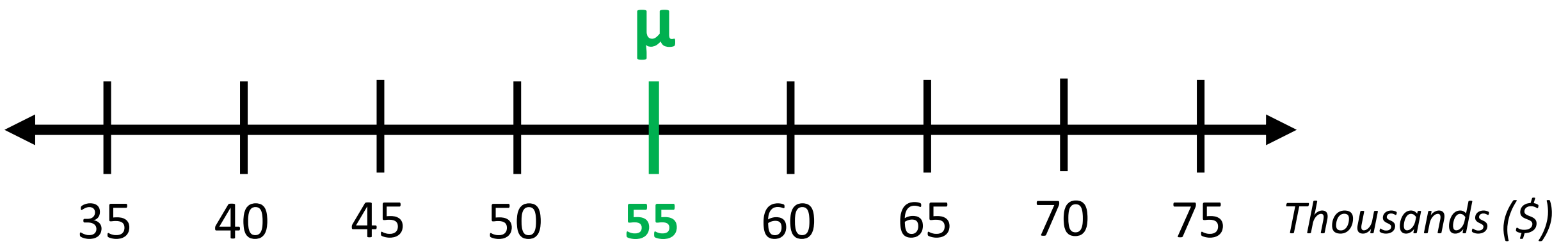
1. The true mean yearly salary among YouTubers is $55,000

2. The true standard deviation of salaries is $29,500

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

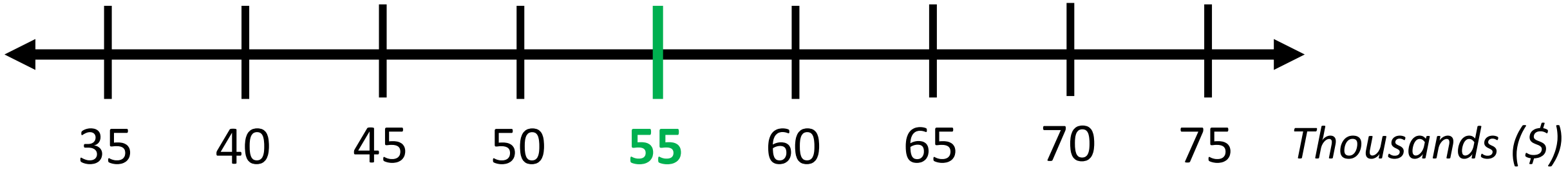
Skew The Script



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script



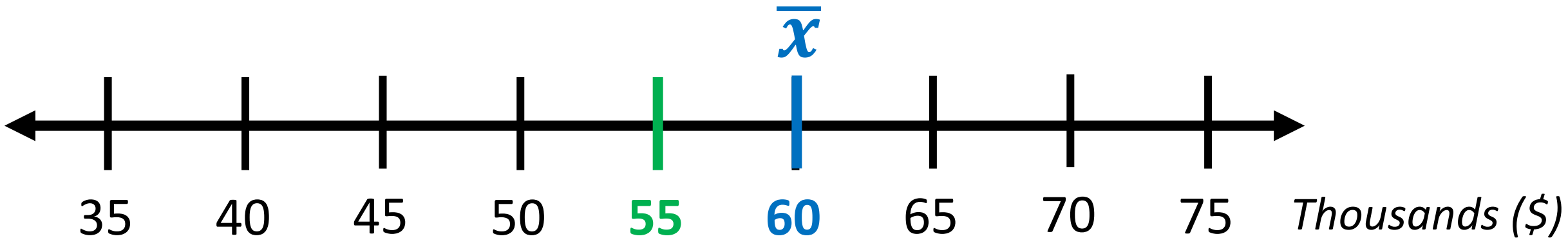
In most real-life scenarios:

-You **don't know** the true mean.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script



In most real-life scenarios:

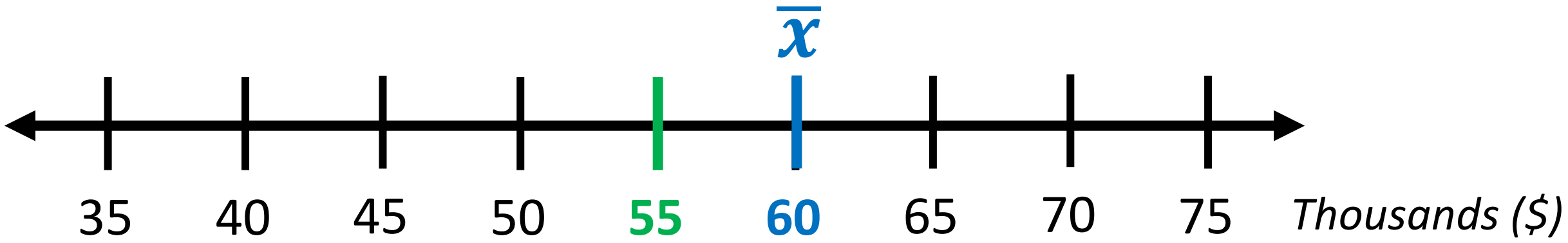
-You **don't know** the true mean.

-You survey a random sample of 35 YouTubers. Among them, the average yearly earnings was $\bar{x} = \$60{,}000$.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script



In most real-life scenarios:

-You **don't know** the true mean.

-You survey a random sample of 35 YouTubers. Among them, the average yearly earnings was $\bar{x} = \$60{,}000$.

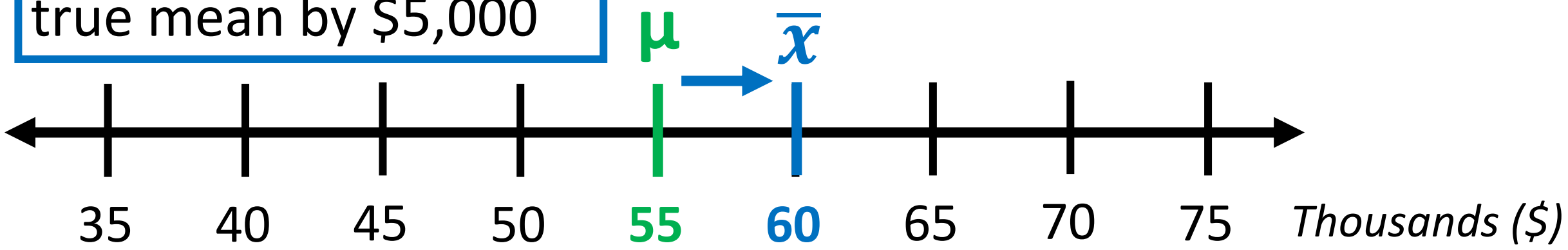
-Your estimate for the true mean (μ) is the sample mean ($\bar{x}$)

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

We **overestimated** the true mean by $5,000



In most real-life scenarios:

-You **don't know** the true mean.

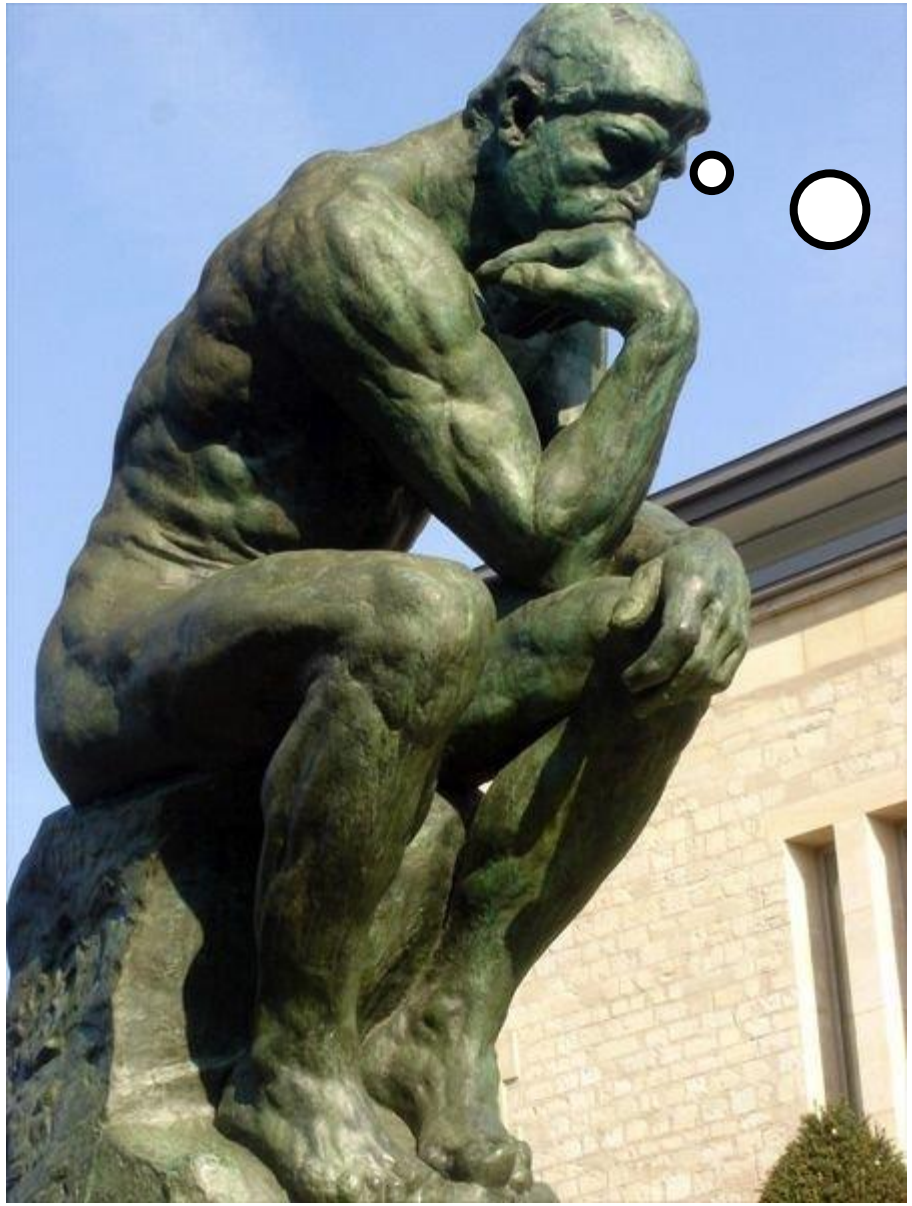
-You survey a random sample of 35 YouTubers. Among them, the average yearly earnings was $\bar{x} = \$60{,}000$.

-Your estimate for the true mean (μ) is the sample mean ($\bar{x}$)

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script



When I estimate something, how far off will I usually be?

Let's take a trip to ***Theory Land*** (the not real place)

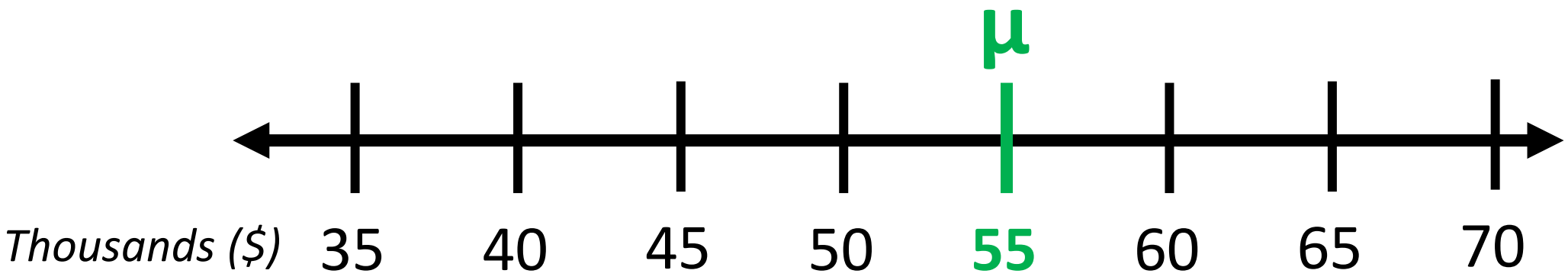
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

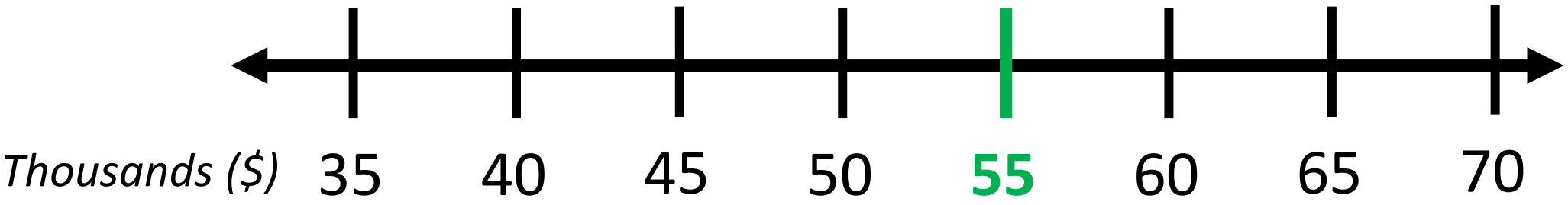
+

Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

In this world, get a random sample of 35 YouTubers.
Find their sample mean.



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

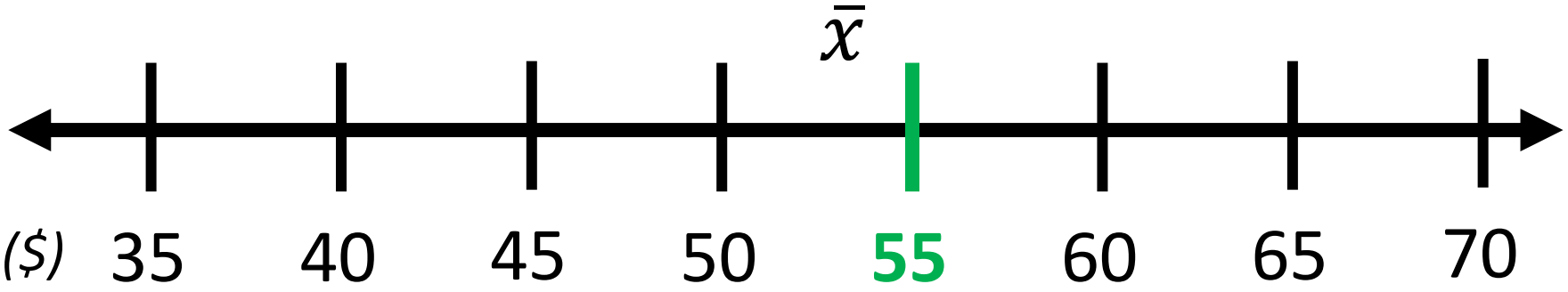
+

Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

In this world, get a random sample of 35 YouTubers.
Find their sample mean.



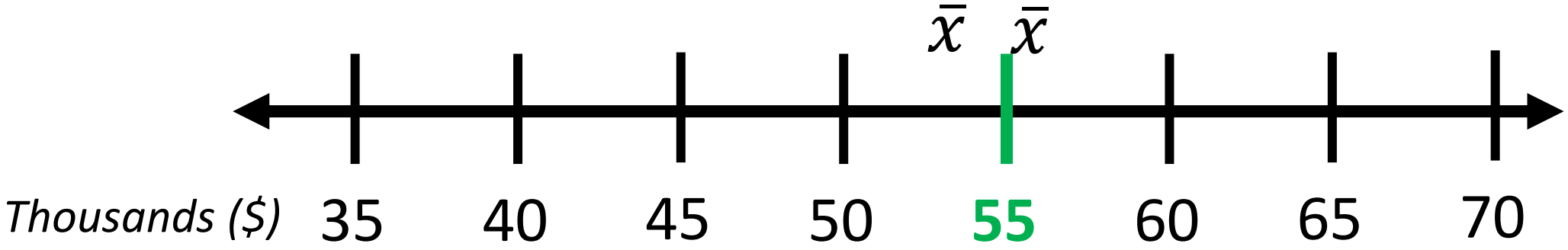
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO
+
Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

Get another random sample of 35 YouTubers. Find their sample mean.

Repeat a bunch of times



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

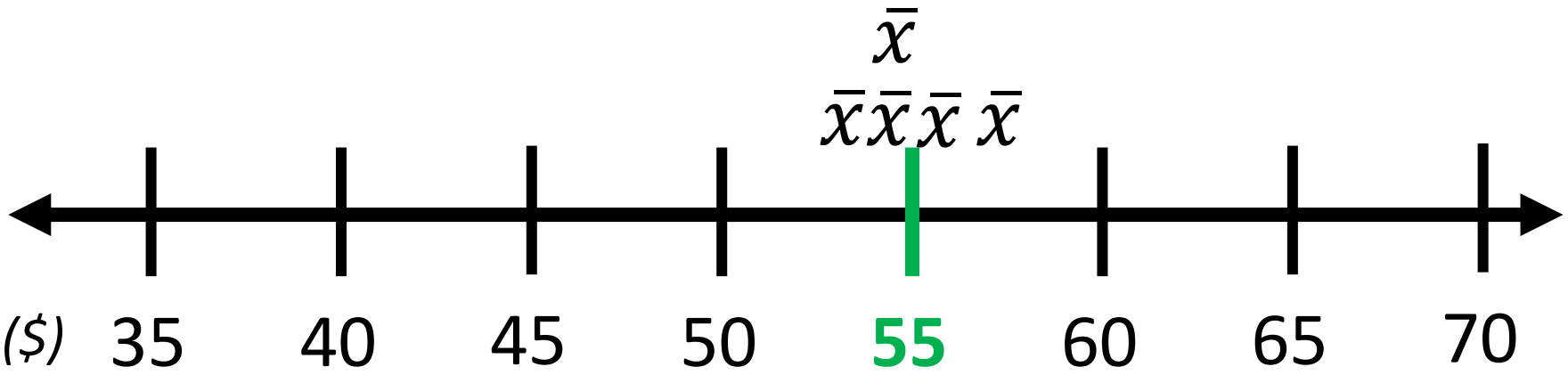
Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

Get another random sample of 35 YouTubers. Find their sample mean.

Repeat a bunch of times



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

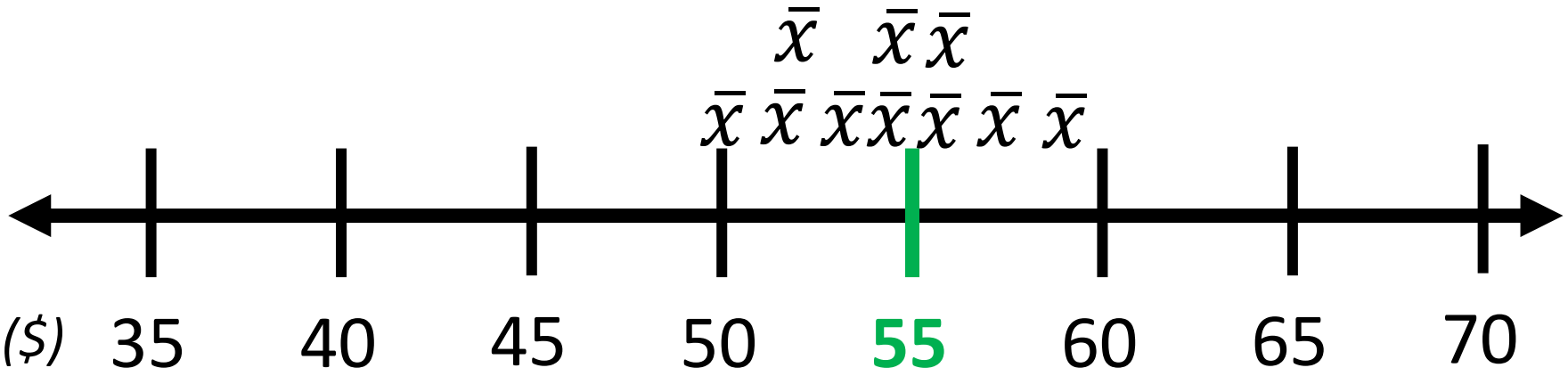
Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

Get another random sample of 35 YouTubers. Find their sample mean.

Repeat a bunch of times



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

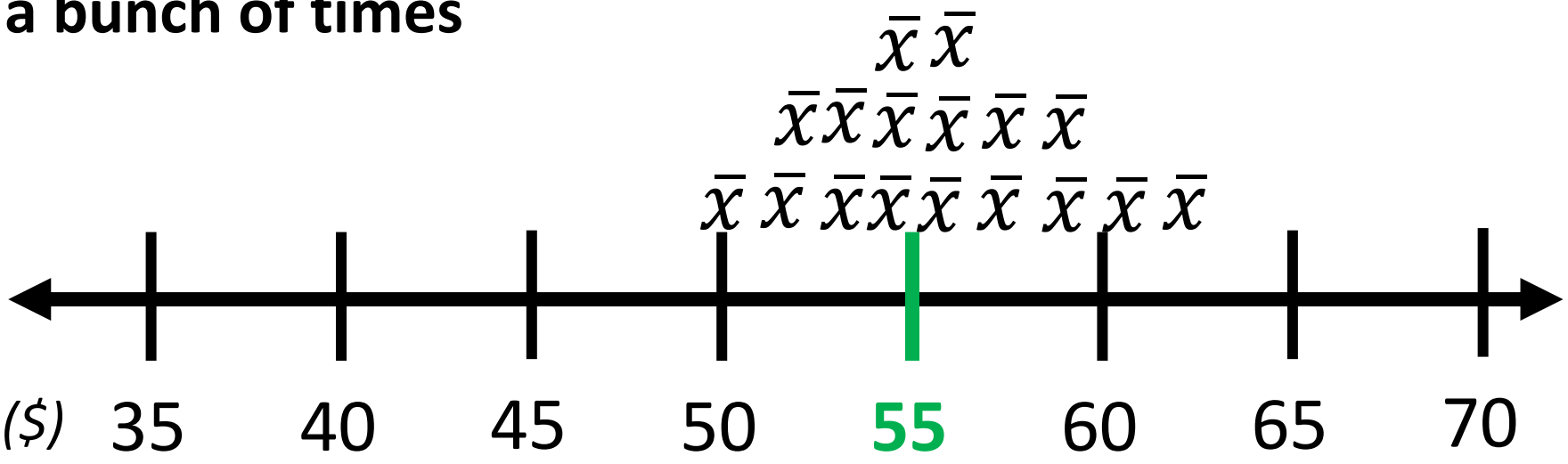
Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

Get another random sample of 35 YouTubers. Find their sample mean.

Repeat a bunch of times



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

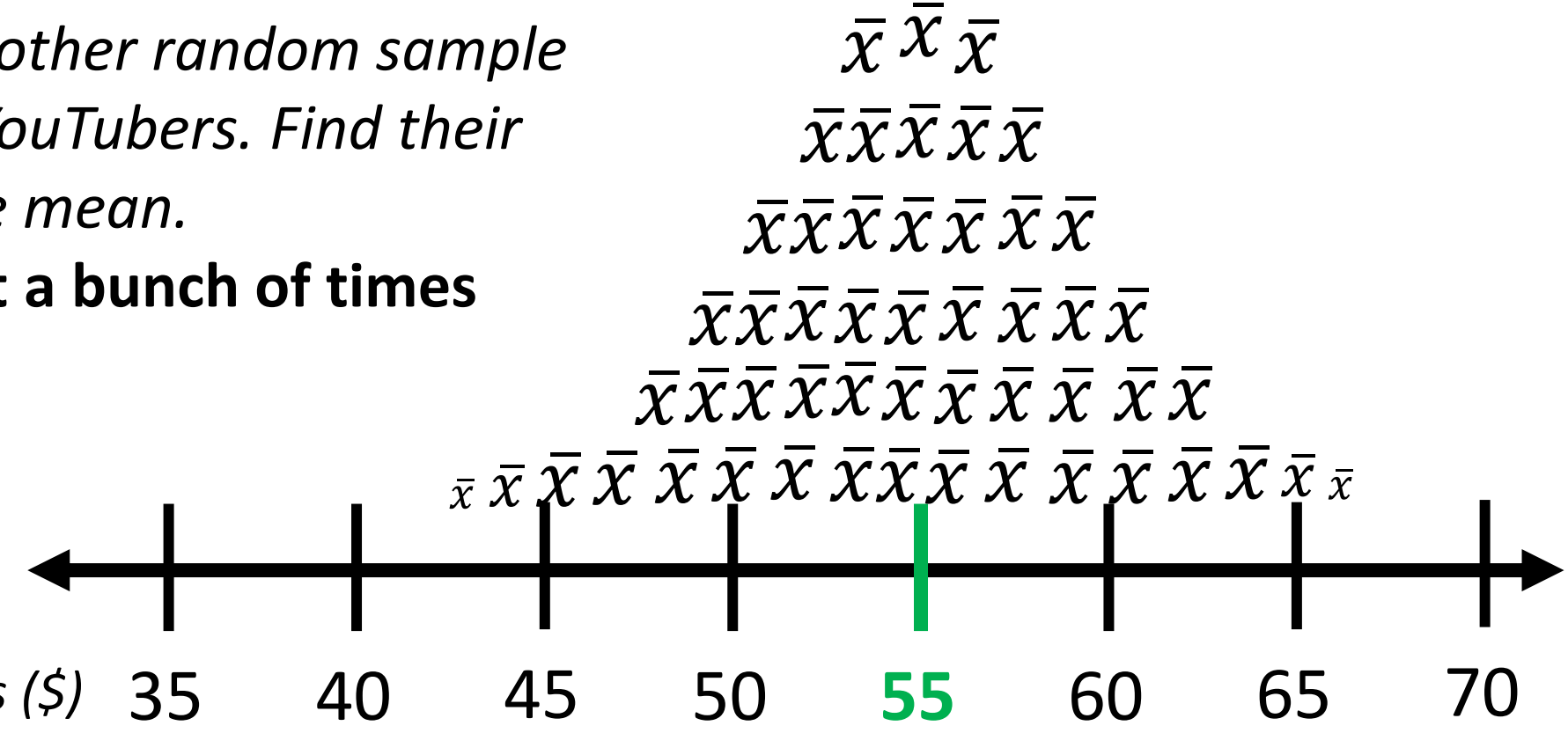
Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

Get another random sample of 35 YouTubers. Find their sample mean.

Repeat a bunch of times



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

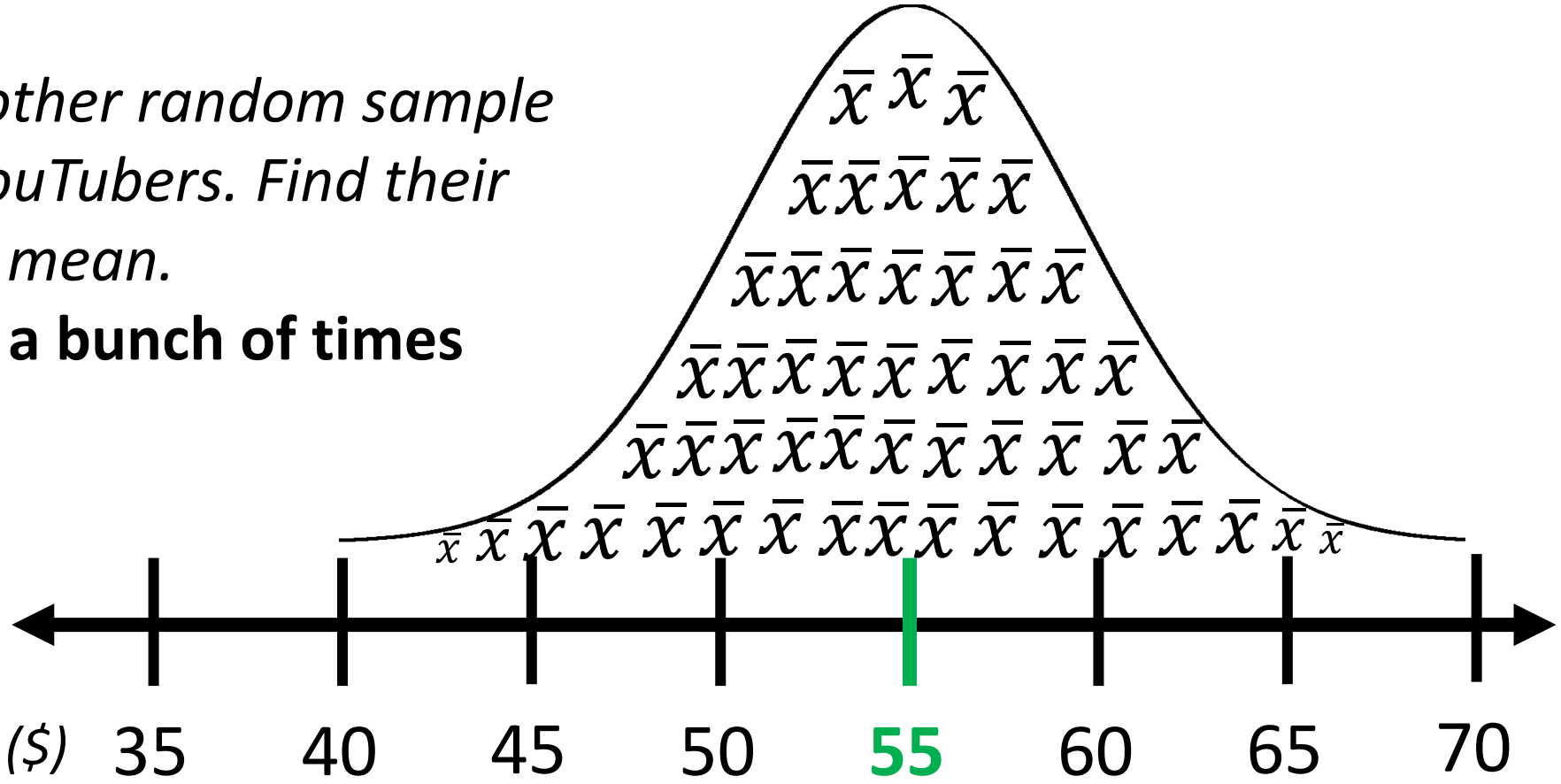
Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

Get another random sample of 35 YouTubers. Find their sample mean.

Repeat a bunch of times



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

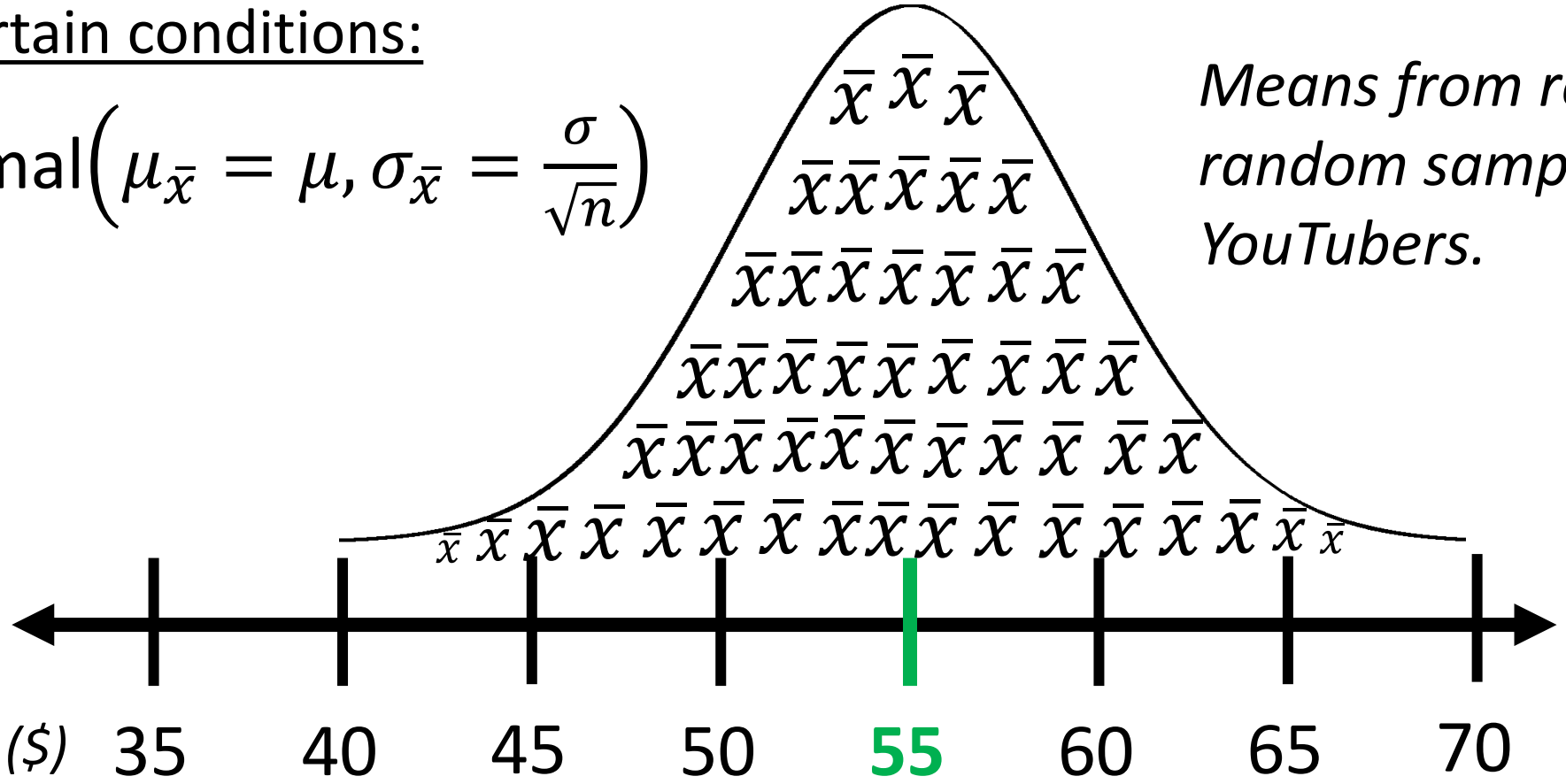
Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

Under certain conditions:

$$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \mu, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\right)$$



Means from repeated random samples of 35 YouTubers.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

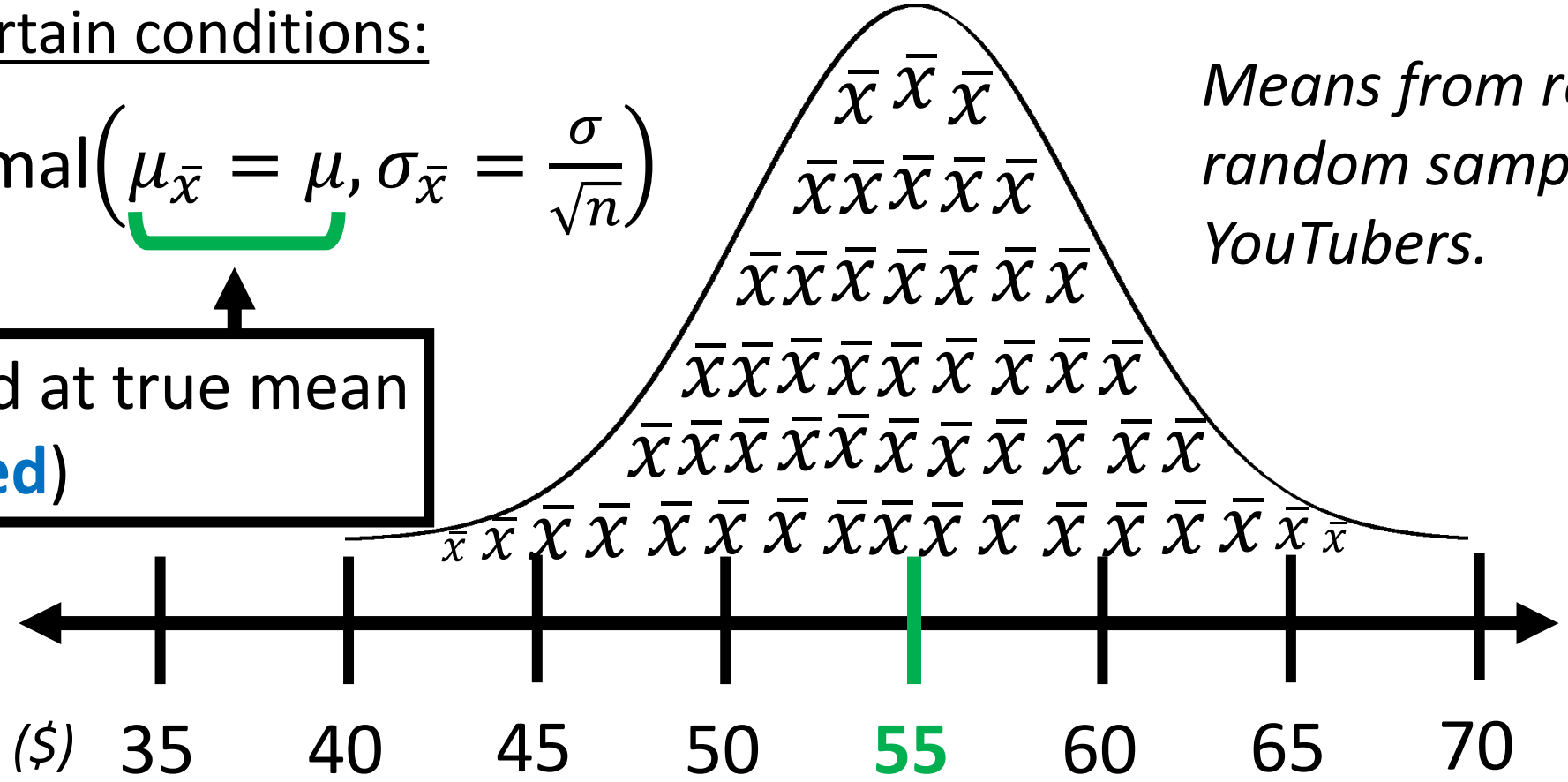
Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

Under certain conditions:

$$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \mu, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\right)$$

Centered at true mean (**unbiased**)



Means from repeated random samples of 35 YouTubers.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

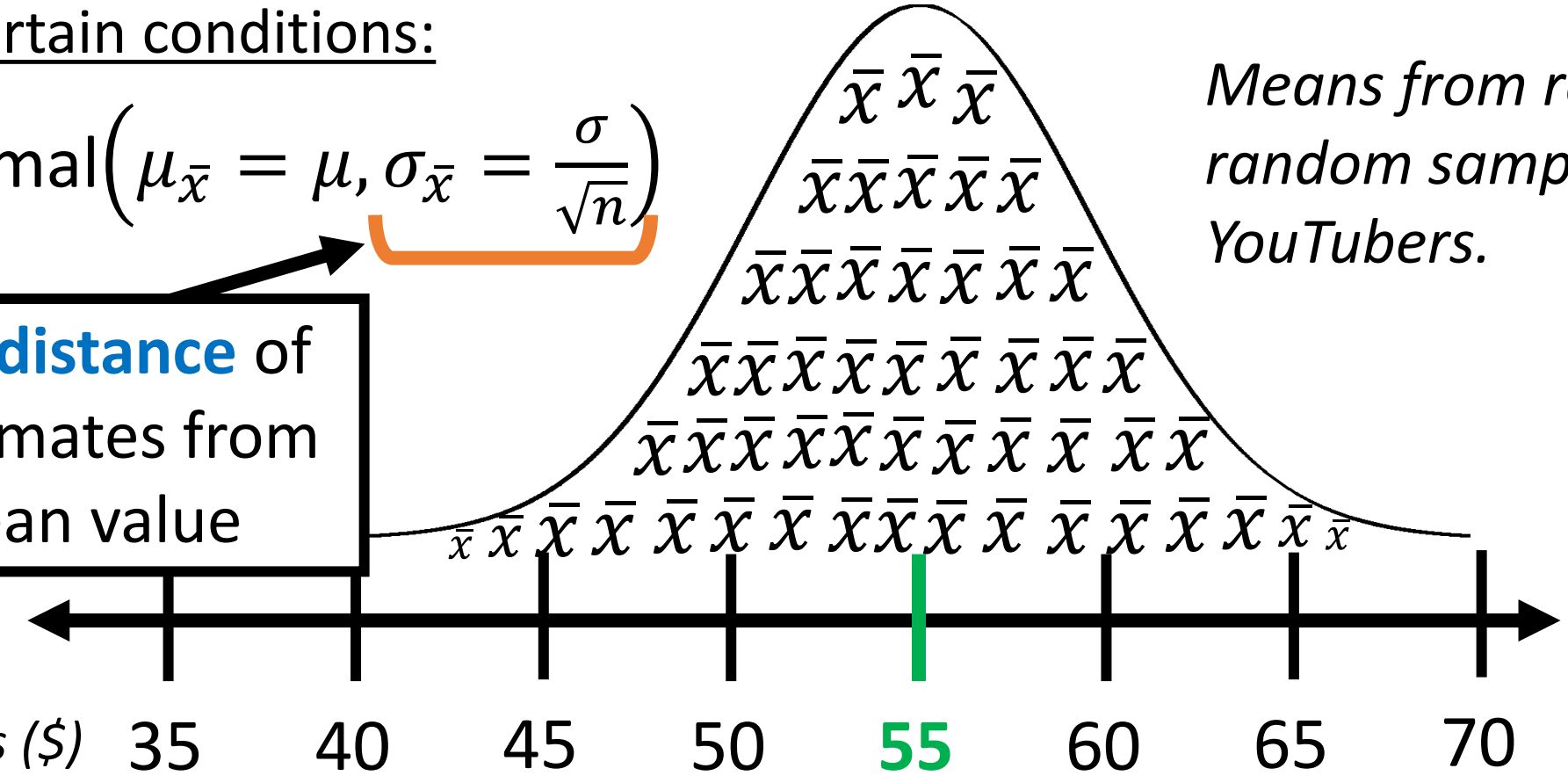
Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

Under certain conditions:

$$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \mu, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\right)$$

Typical distance of our estimates from true mean value



Means from repeated random samples of 35 YouTubers.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

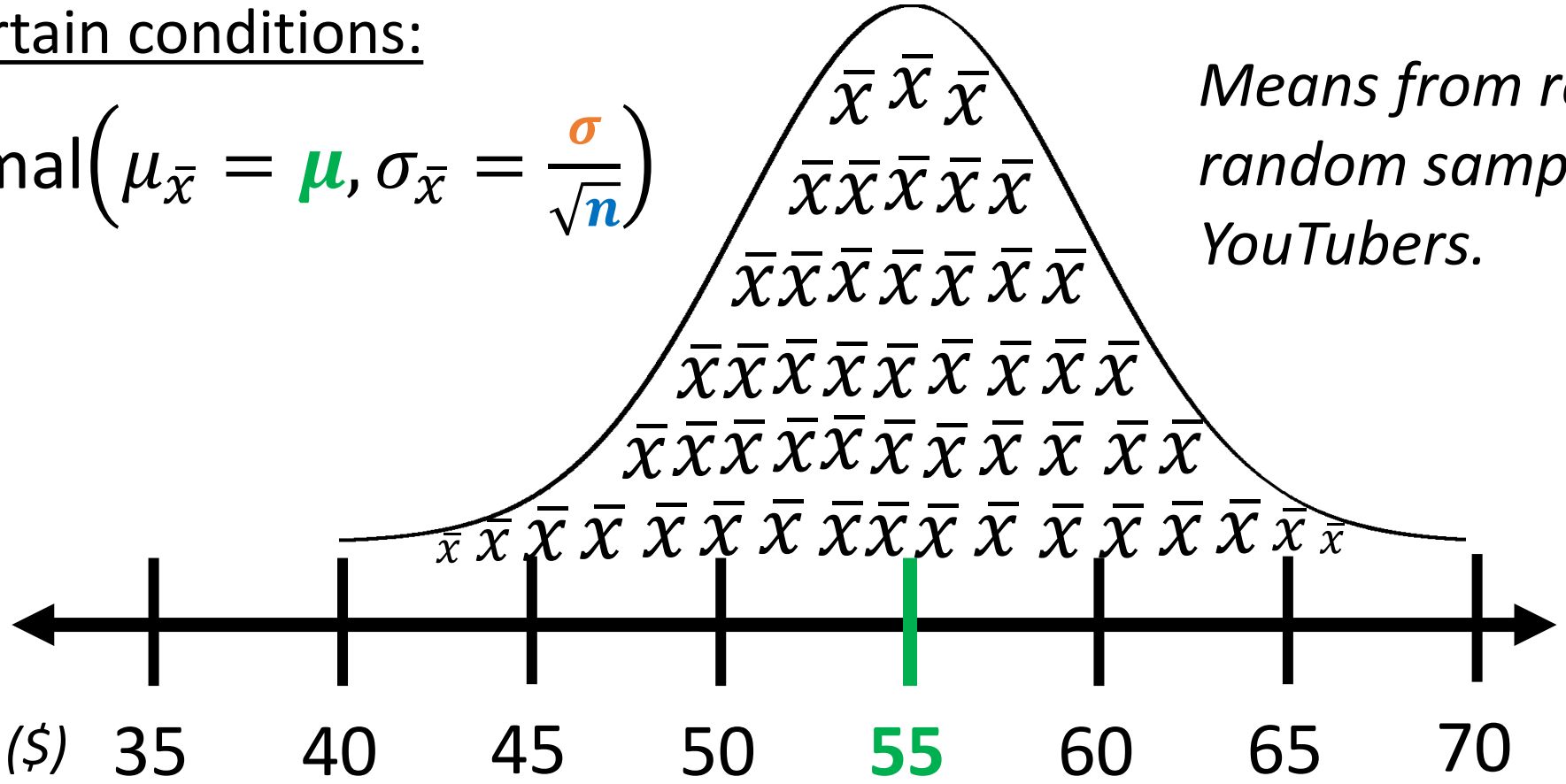
Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is **$55,000** and the standard deviation of salaries is **$29,500**:

Under certain conditions:

$$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \boldsymbol{\mu}, \sigma_{\bar{x}} = \frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}\right)$$



*Means from repeated random samples of **35** YouTubers.*

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

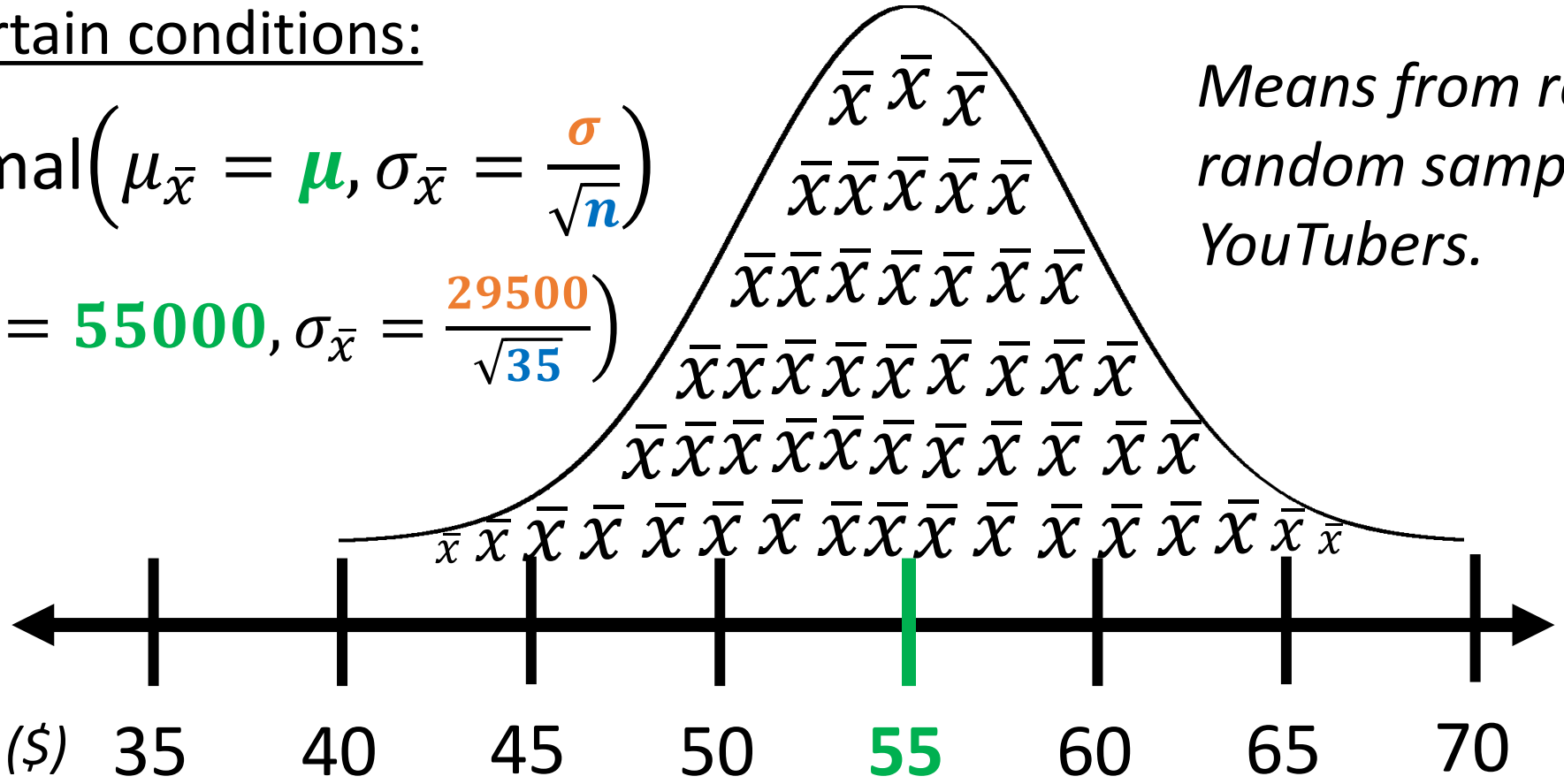
Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is **$55,000** and the standard deviation of salaries is **$29,500**:

Under certain conditions:

$$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \boldsymbol{\mu}, \sigma_{\bar{x}} = \frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}\right)$$

$$\bar{x} \sim \text{N}\left(\mu_{\bar{x}} = \mathbf{55000}, \sigma_{\bar{x}} = \frac{\mathbf{29500}}{\sqrt{\mathbf{35}}}\right)$$



*Means from repeated random samples of **35** YouTubers.*

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Sampling Distribution (Mean)

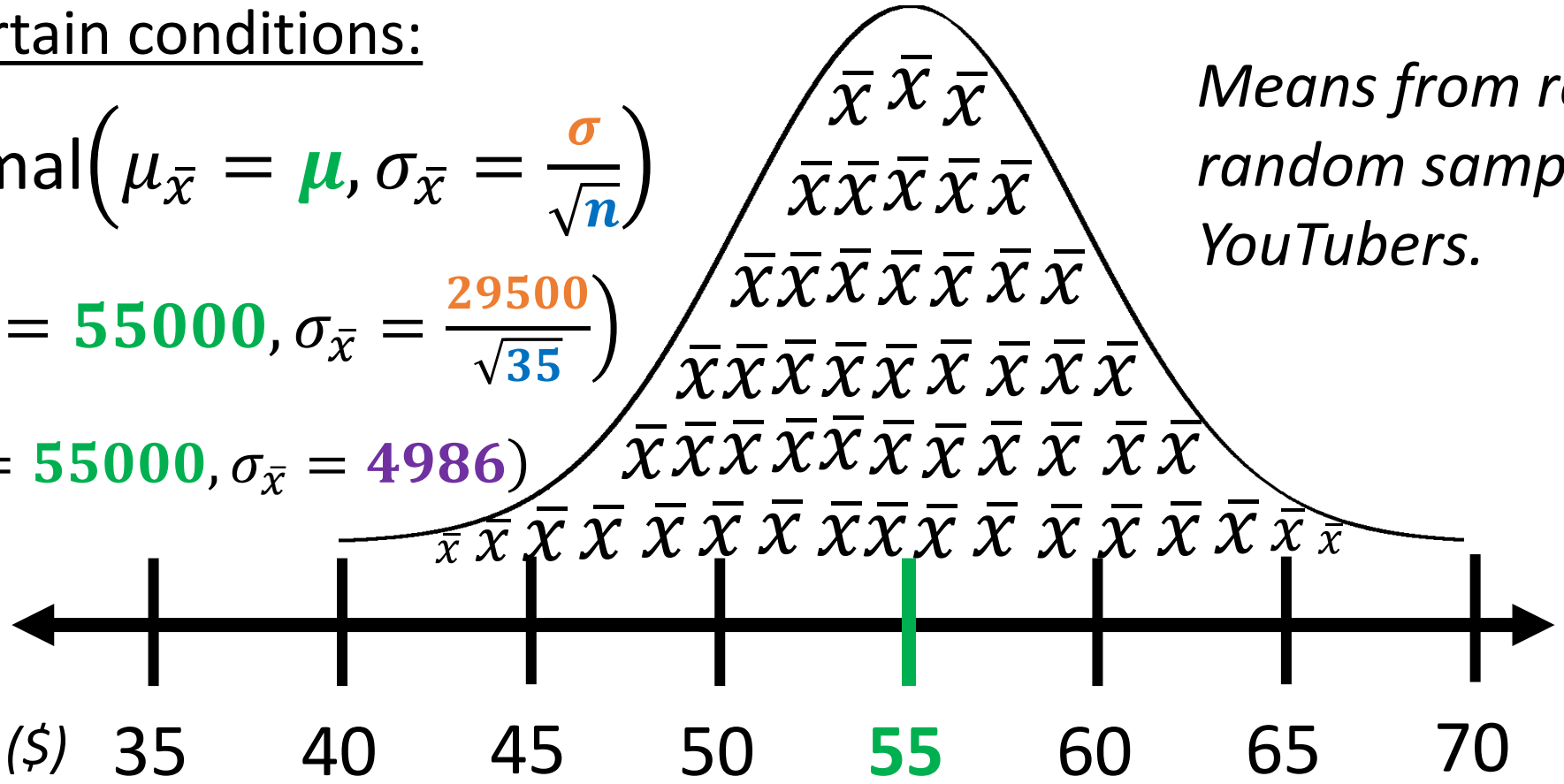
In a world where the true mean salary for *all* YouTubers is **$55,000** and the standard deviation of salaries is **$29,500**:

Under certain conditions:

$$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \boldsymbol{\mu}, \sigma_{\bar{x}} = \frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}\right)$$

$$\bar{x} \sim \text{N}\left(\mu_{\bar{x}} = \mathbf{55000}, \sigma_{\bar{x}} = \frac{\mathbf{29500}}{\sqrt{\mathbf{35}}}\right)$$

$$\bar{x} \sim \text{N}(\mu_{\bar{x}} = \mathbf{55000}, \sigma_{\bar{x}} = \mathbf{4986})$$



*Means from repeated random samples of **35** YouTubers.*

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Sampling Distribution (Mean)

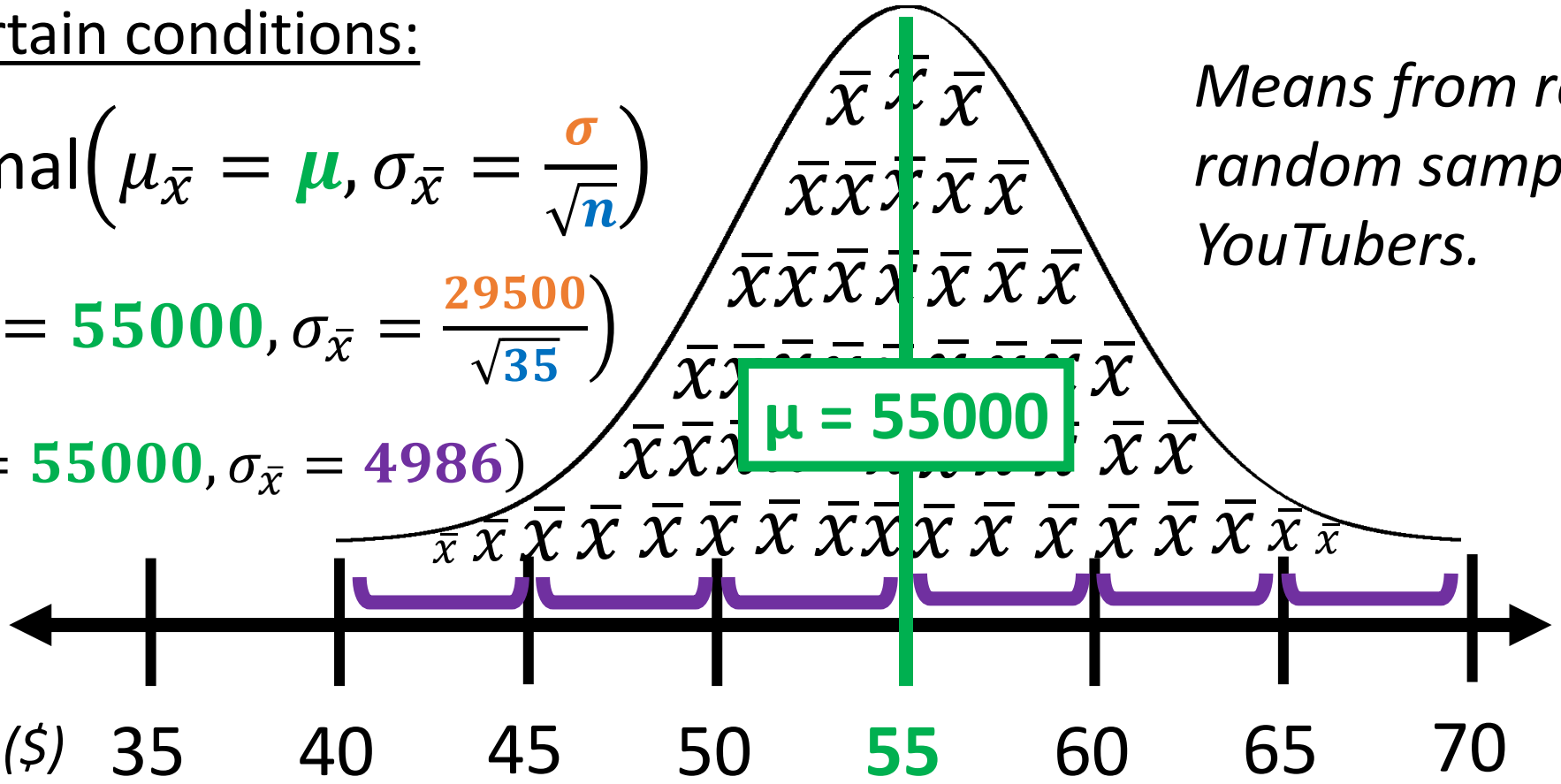
In a world where the true mean salary for *all* YouTubers is **$55,000** and the standard deviation of salaries is **$29,500**:

Under certain conditions:

$$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \boldsymbol{\mu}, \sigma_{\bar{x}} = \frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}\right)$$

$$\bar{x} \sim \text{N}\left(\mu_{\bar{x}} = \mathbf{55000}, \sigma_{\bar{x}} = \frac{\mathbf{29500}}{\sqrt{\mathbf{35}}}\right)$$

$$\bar{x} \sim \text{N}(\mu_{\bar{x}} = \mathbf{55000}, \sigma_{\bar{x}} = \mathbf{4986})$$



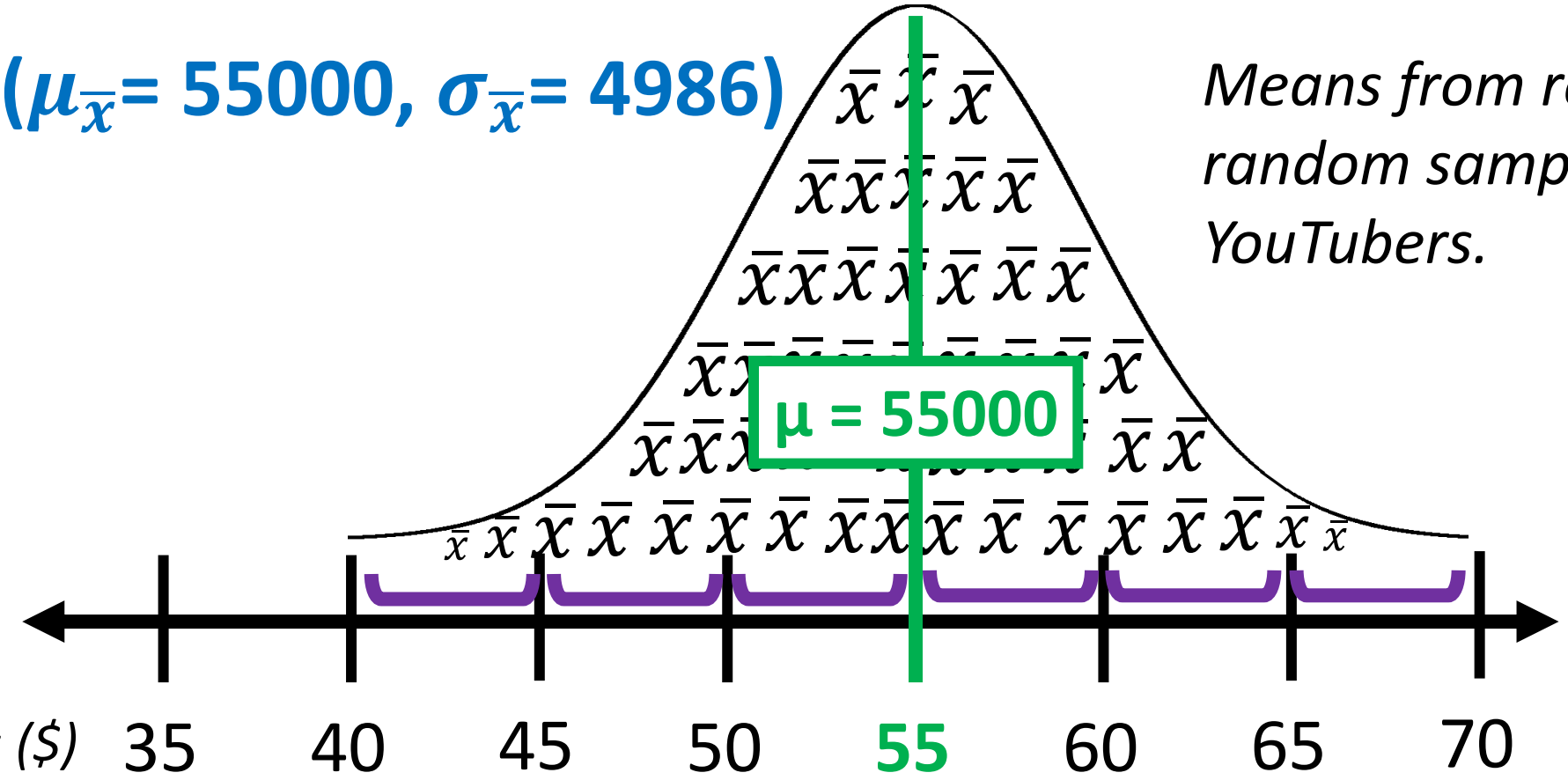
*Means from repeated random samples of **35** YouTubers.*

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO
+
Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

Normal($\mu_{\bar{x}}$= 55000, $\sigma_{\bar{x}}$= 4986)



Means from repeated random samples of 35 YouTubers.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

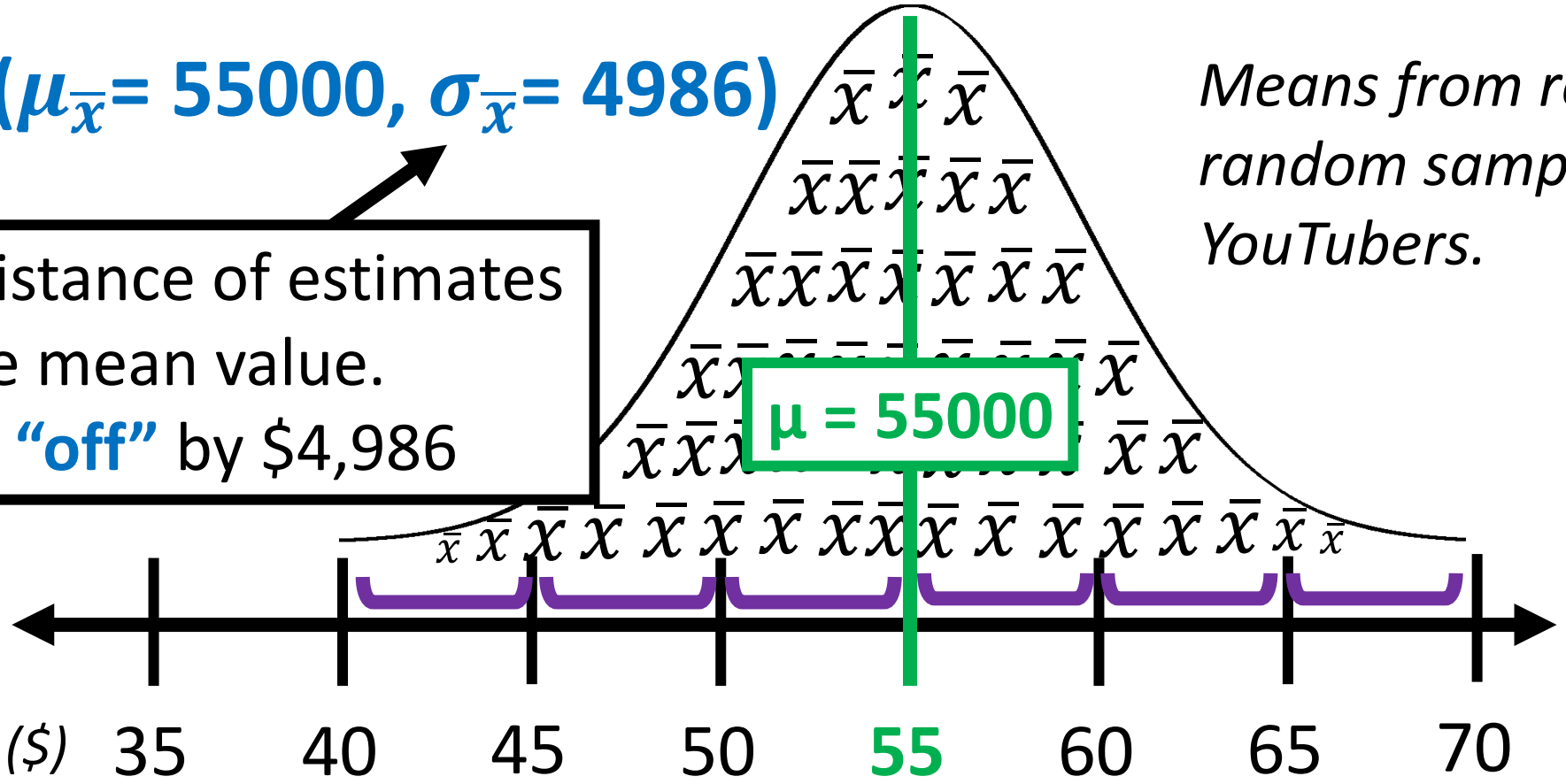
Skew The Script

Sampling Distribution (Mean)

In a world where the true mean salary for *all* YouTubers is $55,000 and the standard deviation of salaries is $29,500:

Normal($\mu_{\bar{x}}$= 55000, $\sigma_{\bar{x}}$= 4986)

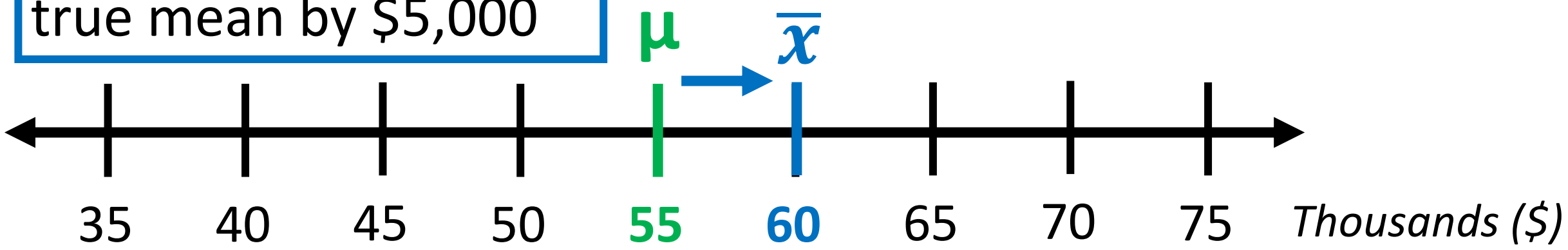
Typical distance of estimates from true mean value.
Typically "off" by $4,986



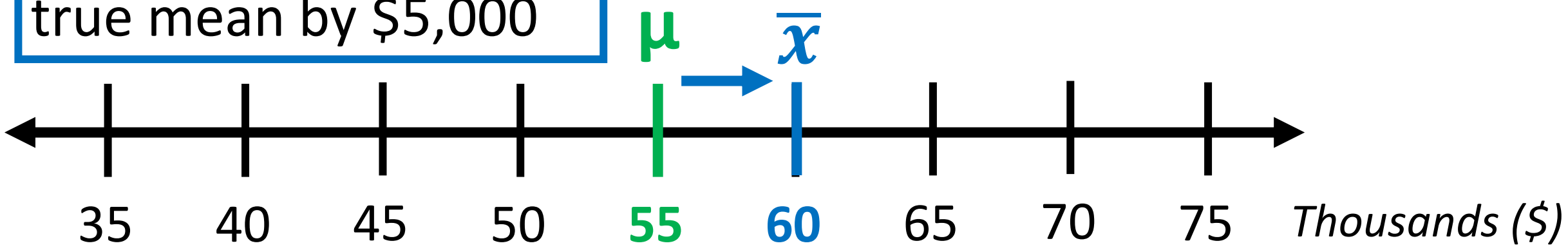
Means from repeated random samples of 35 YouTubers.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO
+
Skew The Script

We **overestimated** the true mean by $5,000



We **overestimated** the true mean by $5,000



This is a pretty typical estimation error in this situation.

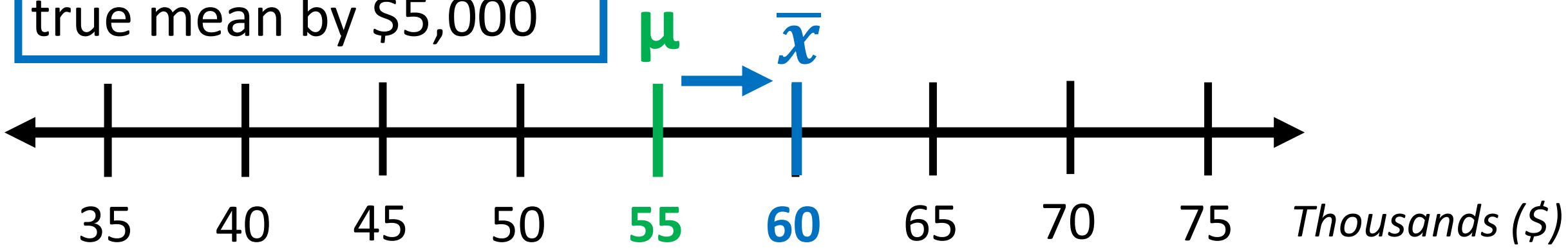
(**actual error = $5,000**, **typical error ($\sigma_{\overline{x}}$) = $4,986**)

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

We **overestimated** the true mean by $5,000



This is a pretty typical estimation error in this situation.
(**actual error = $5,000**, **typical error ($\sigma_{\overline{x}}$) = $4,986**)

-There should be a way to report both:

a) Our estimate

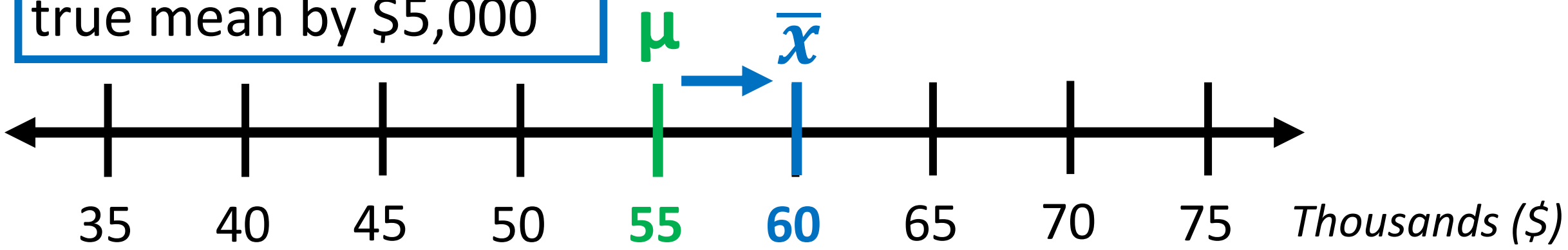
b) How far off this estimate might be

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

We **overestimated** the true mean by $5,000



This is a pretty typical estimation error in this situation.
(**actual error = $5,000**, **typical error ($\sigma_{\overline{x}}$) = $4,986**)

-There should be a way to report both:

a) Our estimate

b) How far off this estimate might be

One way to do this is with a **confidence interval**!

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

Topics

1. Recall: sampling distribution for $\bar{x}$
2. **The t-distribution and interval for $\bar{x}$**
3. Four step process

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

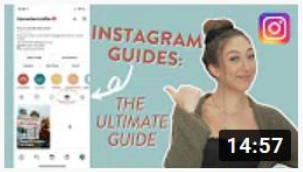
+

Skew The Script

How Much Money Do I Make on YouTube with 60K Subscribers...
Hustle at Home Mom
12K views • 1 month ago

MY FIRST *THREE* Youtube Paychecks | How Much Does ...
Madison Payne
16K views • 5 months ago

How Much YouTube Paid Me For My 2 MILLION Views Video
Liam James Kay
9.8K views • 1 month ago



NEW INSTAGRAM FEATURE GUIDES | What are Instagram...
Modern Millie
9.5K views • 1 week ago

My First YouTube Paycheck: How Much YouTube Pays +...
Journey With The Hintons
202K views • 5 months ago

This is Exactly What I Needed: Robinhood Challenge Week 8
Matthew Michael
1.2K views • 1 month ago

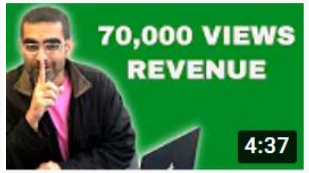
How much money did I make on YouTube??
Dr. Know-it-all Knows it all
5.7K views • 2 weeks ago

HOW LONG IT TAKES TO GET MONETIZED ON YOUTUBE: Th...
Cathrin Manning
1M views • 1 year ago

HOW MUCH MONEY $$$ I MAKE ON YOUTUBE WITH 380...
Kevin Eassa
706 views • 2 weeks ago

10 INSTAGRAM HACKS YOU DIDN'T KNOW EXISTED | How t...
Modern Millie
221K views • 2 months ago

How Much YouTube Paid Me My FIRST YEAR Being...
Cathrin Manning
248K views • 5 months ago



How Much YouTube Paid Me For 70000 Views | YouTube Ad...
Ali Mirza
292 views • 2 weeks ago

this is how much money YouTube paid me TOTAL in...
Annie Dubé
7.8K views • 1 week ago

This Is How Much YouTube Paid Me For My 1,000,000...
Shelby Church
8.8M views • 1 year ago

HOW MUCH I MADE ON YOUTUBE IN 1 MONTH WITH...
Modern Millie
4.1K views • 1 month ago

MY FIRST YOUTUBE PAYCHECK | HOW MUCH MONEY DO...
Sonja Lauren
131K views • 10 months ago

💰 My $2,000/Day Passive Income Online Store 🌞 (Case...
Wholesale Ted
256K views • 9 months ago

MY FIRST YOUTUBE PAYCHECK! How Much Money...
Alexandra Budgets
8.3K views • 8 months ago

How Much Youtube Paid me for 31 Days (30 million views)
Meet Kevin
222K views • 4 months ago

How Much Money YouTube Pays Us (125k Subscribers)
Colin and Samir
223K views • 5 months ago

MY FIRST YOUTUBE PAYCHECK / 1,000 subscribers
Samantha Piedra
55K views • 1 year ago

How Much $$ I Really Make as a Small YouTuber | Emma...
Emma Stevens
34K views • 11 months ago

how much youtube paid me for a 1 million viewed video
Reyes The Entrepreneur
3M views • 1 year ago

MY FIRST YOUTUBE PAYCHECK: How Much Money...
Morgan Stradling
12K views • 5 months ago

Got The Data

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Sampled YouTuber Yearly Revenues ($)

Yearly Revenue
$706.74
$46,532.04
$9,565.32
$48,000.00
$23,801.40
$3,359.17
$12,649.08
$73,030.80
$170,945.60
$220,200.84
$8,743.92
$19,873.68
$31,118.19
$68,777.16
$186,819.69
$30,824.60
$32,405.16
$18,625.29
$8,030.69
$53,044.98
$44,379.24
$17,572.93
$263,309.52
$193,918.33
$308,000.00
$7,137.12
$25,487.88
$17,953.89
$28,948.20
$14,172.00
$121,239.03
$34,418.04
$86,893.74
$190,841.66
$32,398.80

n = 35 YouTubers

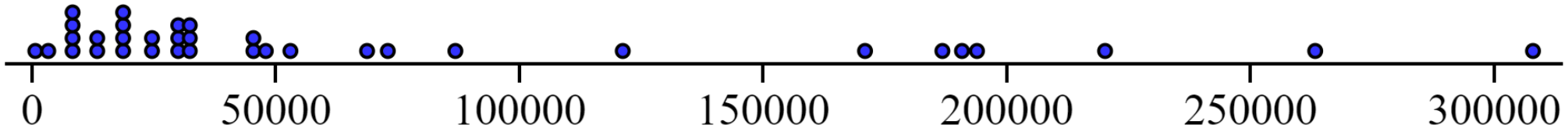
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Yearly Revenue
$706.74
$46,532.04
$9,565.32
$48,000.00
$23,801.40
$3,359.17
$12,649.08
$73,030.80
$170,945.60
$220,200.84
$8,743.92
$19,873.68
$31,118.19
$68,777.16
$186,819.69
$30,824.60
$32,405.16
$18,625.29
$8,030.69
$53,044.98
$44,379.24
$17,572.93
$263,309.52
$193,918.33
$308,000.00
$7,137.12
$25,487.88
$17,953.89
$28,948.20
$14,172.00
$121,239.03
$34,418.04
$86,893.74
$190,841.66
$32,398.80

Sampled YouTuber Yearly Revenues ($)



n = 35 YouTubers

Dotplot created with *stapplet.com*

2020 poverty guideline from US HHS, 2019 mean wage from US SSA

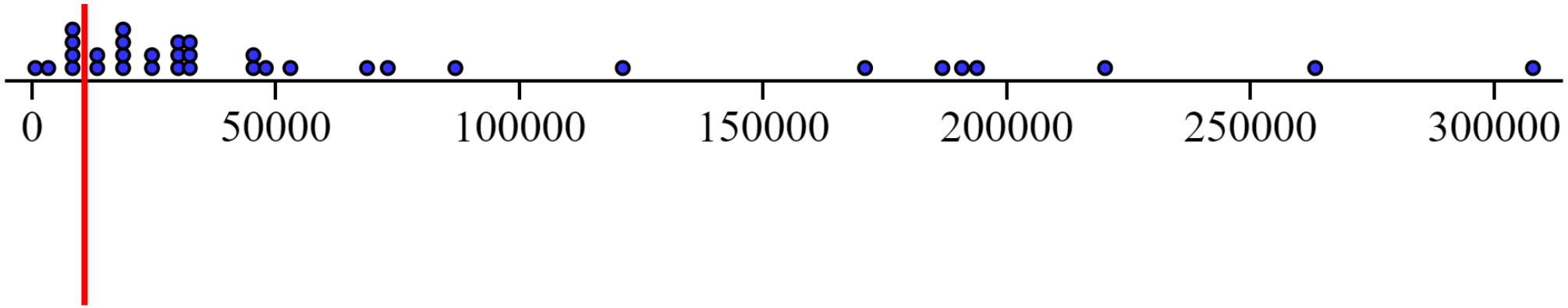
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Yearly Revenue
$706.74
$46,532.04
$9,565.32
$48,000.00
$23,801.40
$3,359.17
$12,649.08
$73,030.80
$170,945.60
$220,200.84
$8,743.92
$19,873.68
$31,118.19
$68,777.16
$186,819.69
$30,824.60
$32,405.16
$18,625.29
$8,030.69
$53,044.98
$44,379.24
$17,572.93
$263,309.52
$193,918.33
$308,000.00
$7,137.12
$25,487.88
$17,953.89
$28,948.20
$14,172.00
$121,239.03
$34,418.04
$86,893.74
$190,841.66
$32,398.80

Sampled YouTuber Yearly Revenues ($)



Individual Poverty Line: $12,760

n = 35 YouTubers

Dotplot created with *stapplet.com*

2020 poverty guideline from US HHS, 2019 mean wage from US SSA

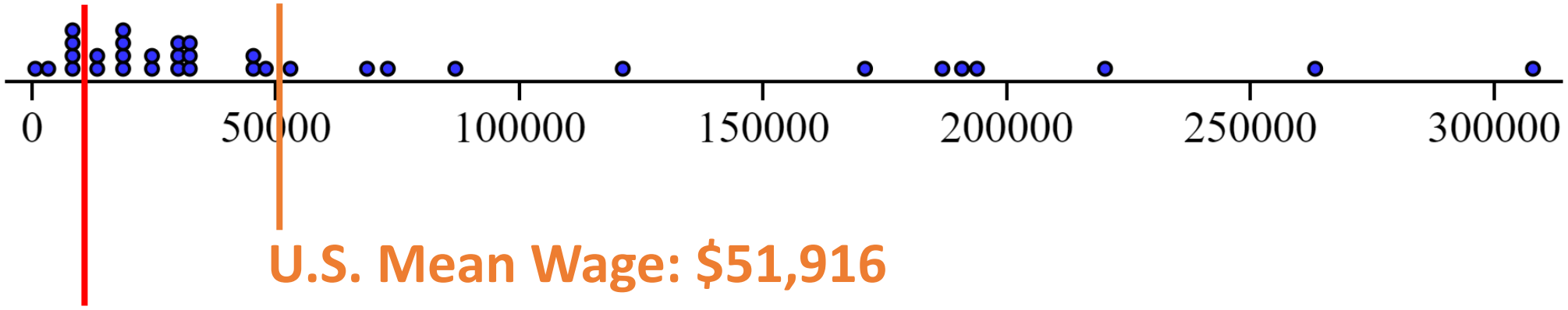
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Yearly Revenue
$706.74
$46,532.04
$9,565.32
$48,000.00
$23,801.40
$3,359.17
$12,649.08
$73,030.80
$170,945.60
$220,200.84
$8,743.92
$19,873.68
$31,118.19
$68,777.16
$186,819.69
$30,824.60
$32,405.16
$18,625.29
$8,030.69
$53,044.98
$44,379.24
$17,572.93
$263,309.52
$193,918.33
$308,000.00
$7,137.12
$25,487.88
$17,953.89
$28,948.20
$14,172.00
$121,239.03
$34,418.04
$86,893.74
$190,841.66
$32,398.80

Sampled YouTuber Yearly Revenues ($)



U.S. Mean Wage: $51,916

Individual Poverty Line: $12,760

n = 35 YouTubers

Dotplot created with *stapplet.com*

2020 poverty guideline from US HHS, 2019 mean wage from US SSA

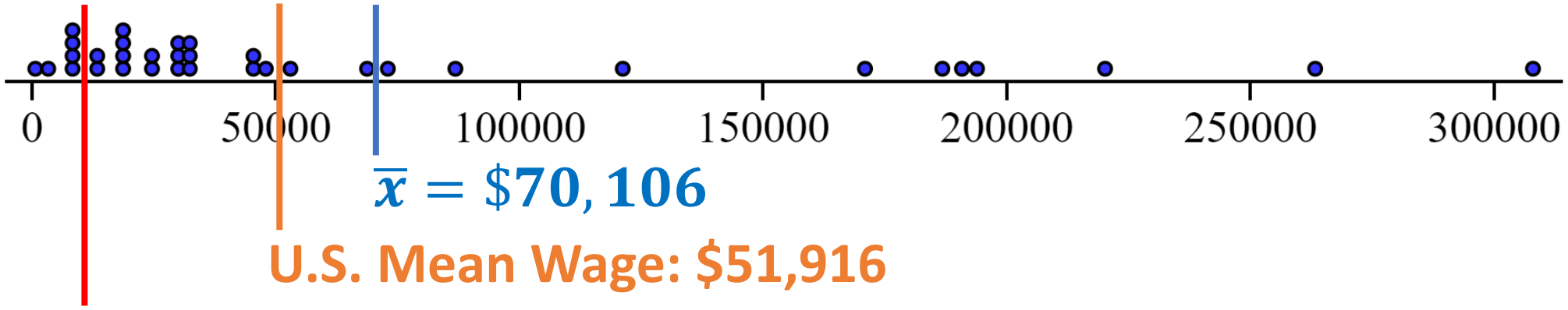
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Yearly Revenue
$706.74
$46,532.04
$9,565.32
$48,000.00
$23,801.40
$3,359.17
$12,649.08
$73,030.80
$170,945.60
$220,200.84
$8,743.92
$19,873.68
$31,118.19
$68,777.16
$186,819.69
$30,824.60
$32,405.16
$18,625.29
$8,030.69
$53,044.98
$44,379.24
$17,572.93
$263,309.52
$193,918.33
$308,000.00
$7,137.12
$25,487.88
$17,953.89
$28,948.20
$14,172.00
$121,239.03
$34,418.04
$86,893.74
$190,841.66
$32,398.80

Sampled YouTuber Yearly Revenues ($)



n = 35 YouTubers

Dotplot created with *stapplet.com*

2020 poverty guideline from US HHS, 2019 mean wage from US SSA

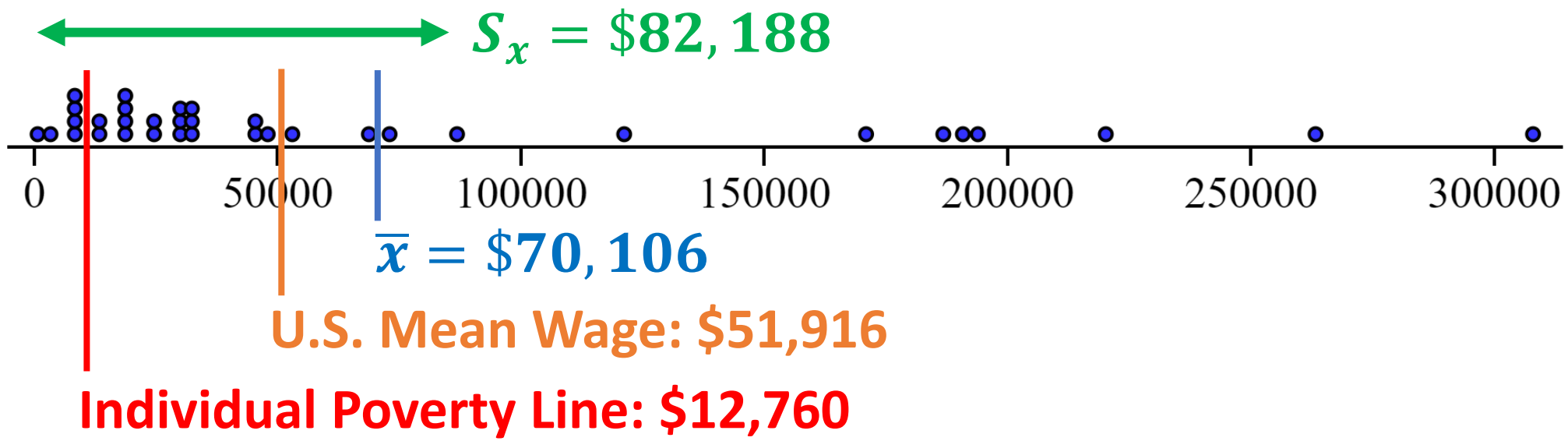
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Yearly Revenue
$706.74
$46,532.04
$9,565.32
$48,000.00
$23,801.40
$3,359.17
$12,649.08
$73,030.80
$170,945.60
$220,200.84
$8,743.92
$19,873.68
$31,118.19
$68,777.16
$186,819.69
$30,824.60
$32,405.16
$18,625.29
$8,030.69
$53,044.98
$44,379.24
$17,572.93
$263,309.52
$193,918.33
$308,000.00
$7,137.12
$25,487.88
$17,953.89
$28,948.20
$14,172.00
$121,239.03
$34,418.04
$86,893.74
$190,841.66
$32,398.80

Sampled YouTuber Yearly Revenues ($)



Impressive! But a lot of variation and skew. What if, by chance, we sampled unusually popular YouTubers and overestimated the true mean?

n = 35 YouTubers

Dotplot created with *stapplet.com*

2020 poverty guideline from US HHS, 2019 mean wage from US SSA

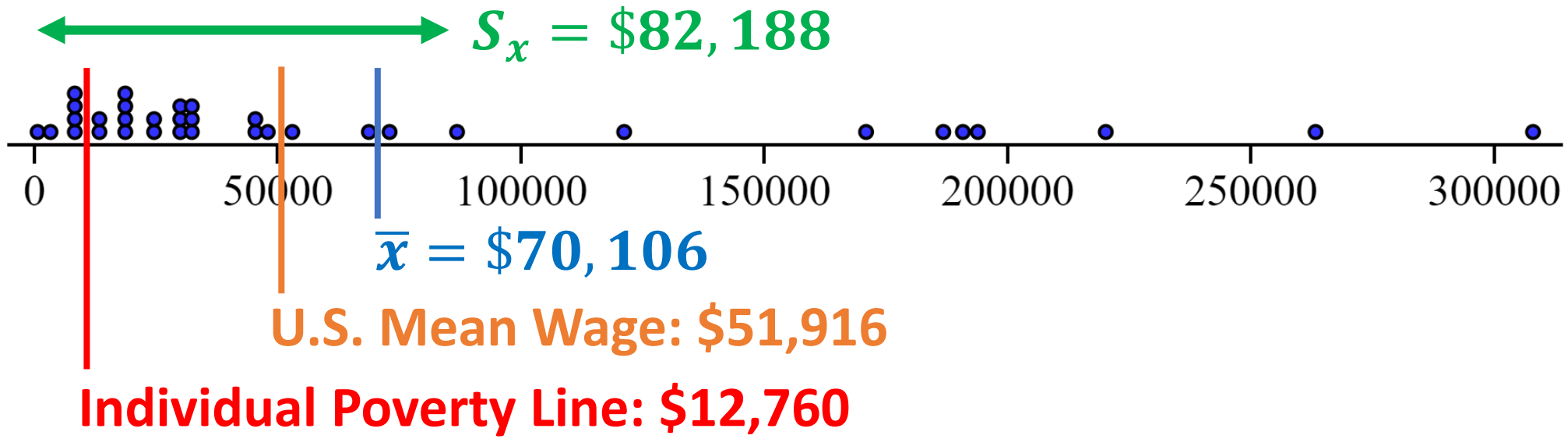
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Yearly Revenue
$706.74
$46,532.04
$9,565.32
$48,000.00
$23,801.40
$3,359.17
$12,649.08
$73,030.80
$170,945.60
$220,200.84
$8,743.92
$19,873.68
$31,118.19
$68,777.16
$186,819.69
$30,824.60
$32,405.16
$18,625.29
$8,030.69
$53,044.98
$44,379.24
$17,572.93
$263,309.52
$193,918.33
$308,000.00
$7,137.12
$25,487.88
$17,953.89
$28,948.20
$14,172.00
$121,239.03
$34,418.04
$86,893.74
$190,841.66
$32,398.80

Sampled YouTuber Yearly Revenues ($)



Impressive! But a lot of variation and skew. What if, by chance, we sampled unusually popular YouTubers and overestimated the true mean? → **confidence interval!**

n = 35 YouTubers

Dotplot created with *stapplet.com*

2020 poverty guideline from US HHS, 2019 mean wage from US SSA

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

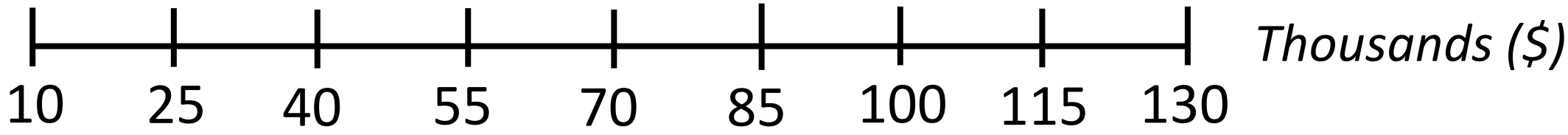
+

Skew The Script

μ = ? →

What is true mean salary of all YouTubers?

← μ? →

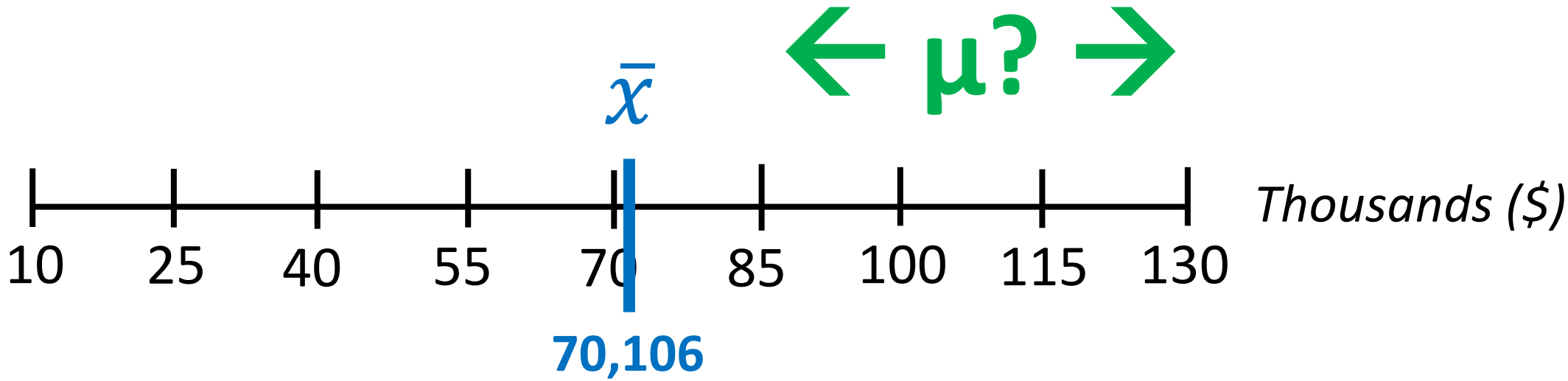


THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$

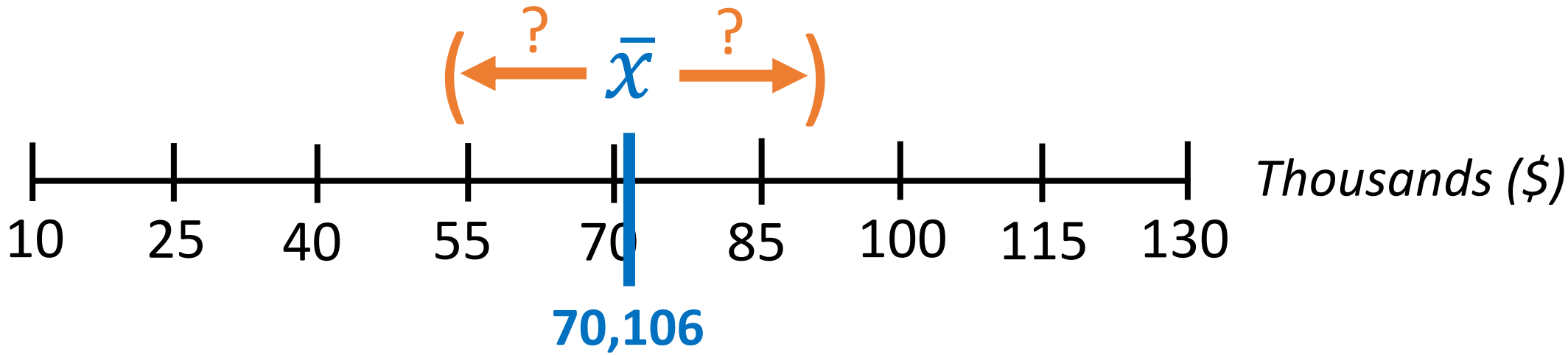


THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$

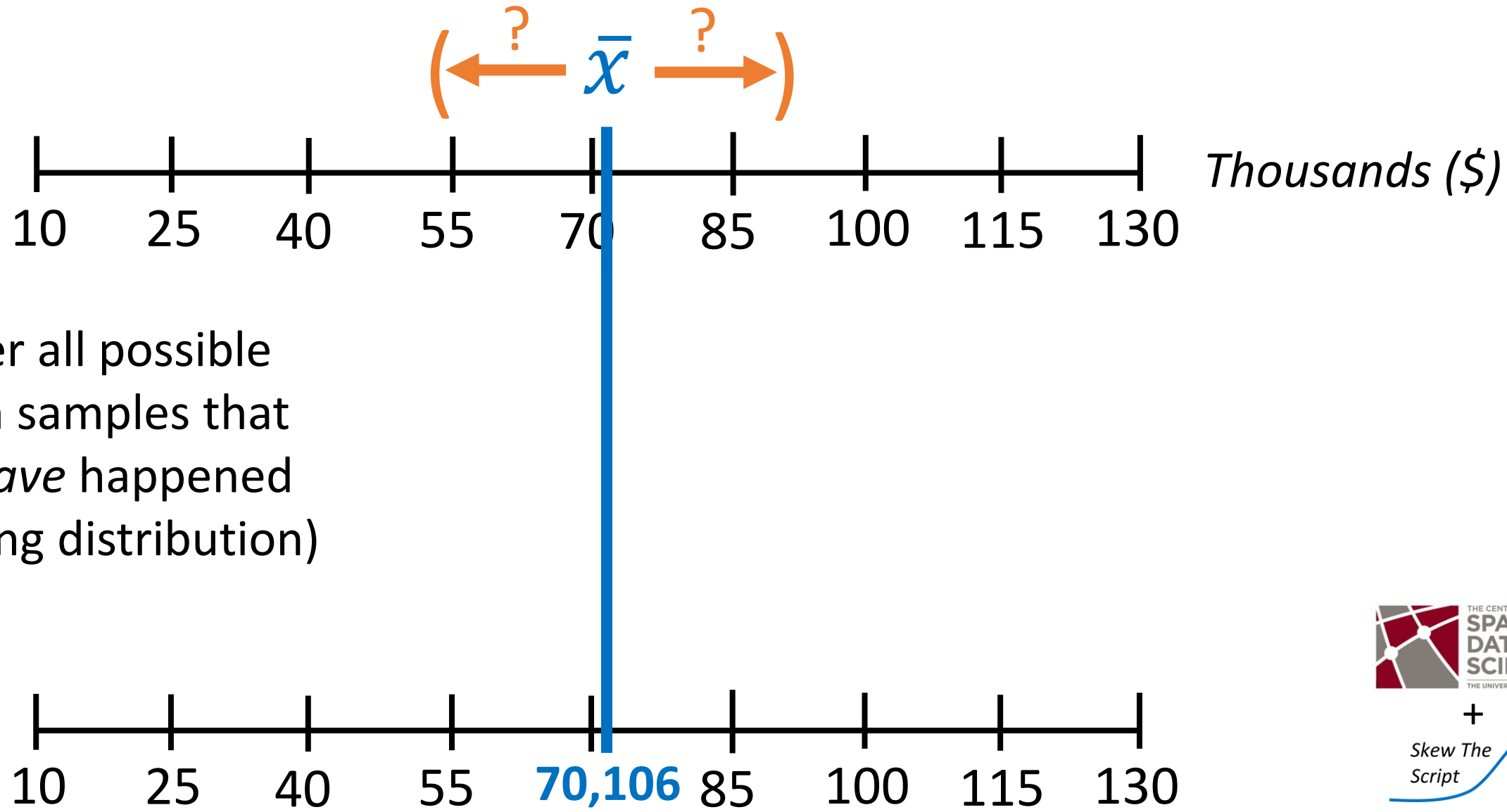


THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



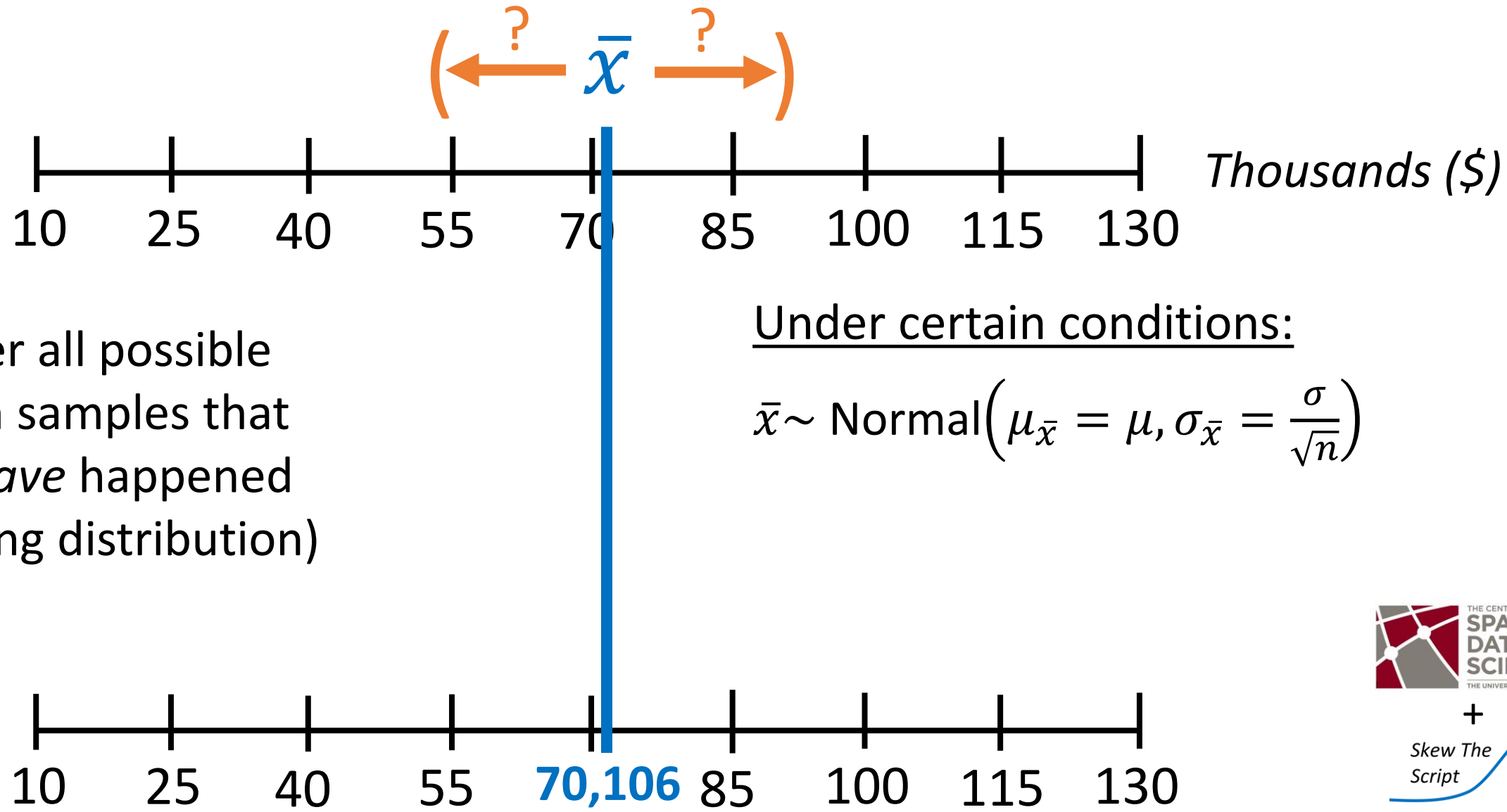
Consider all possible random samples that *could have* happened (sampling distribution)

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



Consider all possible random samples that *could have* happened (sampling distribution)

Under certain conditions:

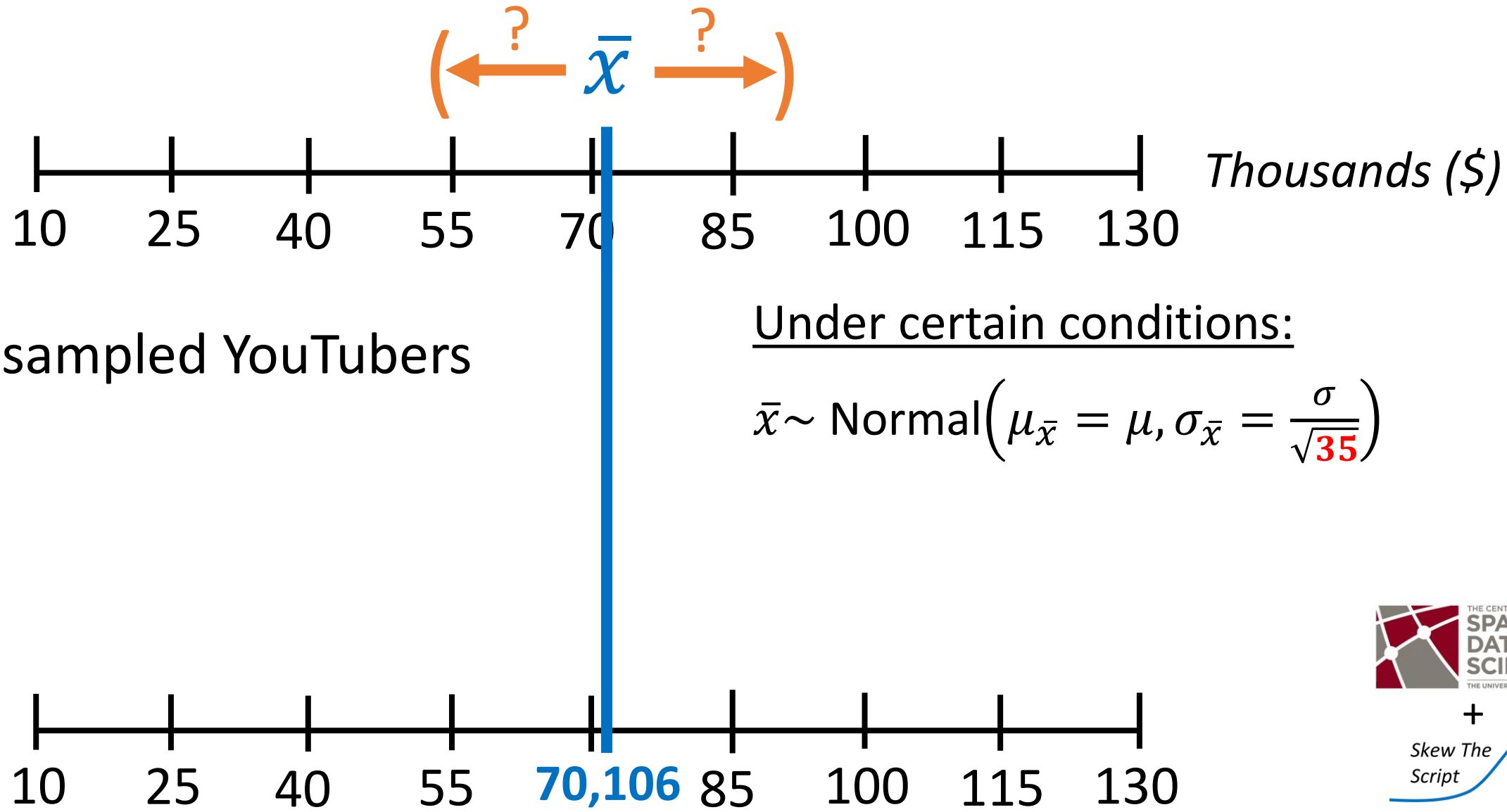
$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \mu, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\right)$

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



n = 35 sampled YouTubers

Under certain conditions:

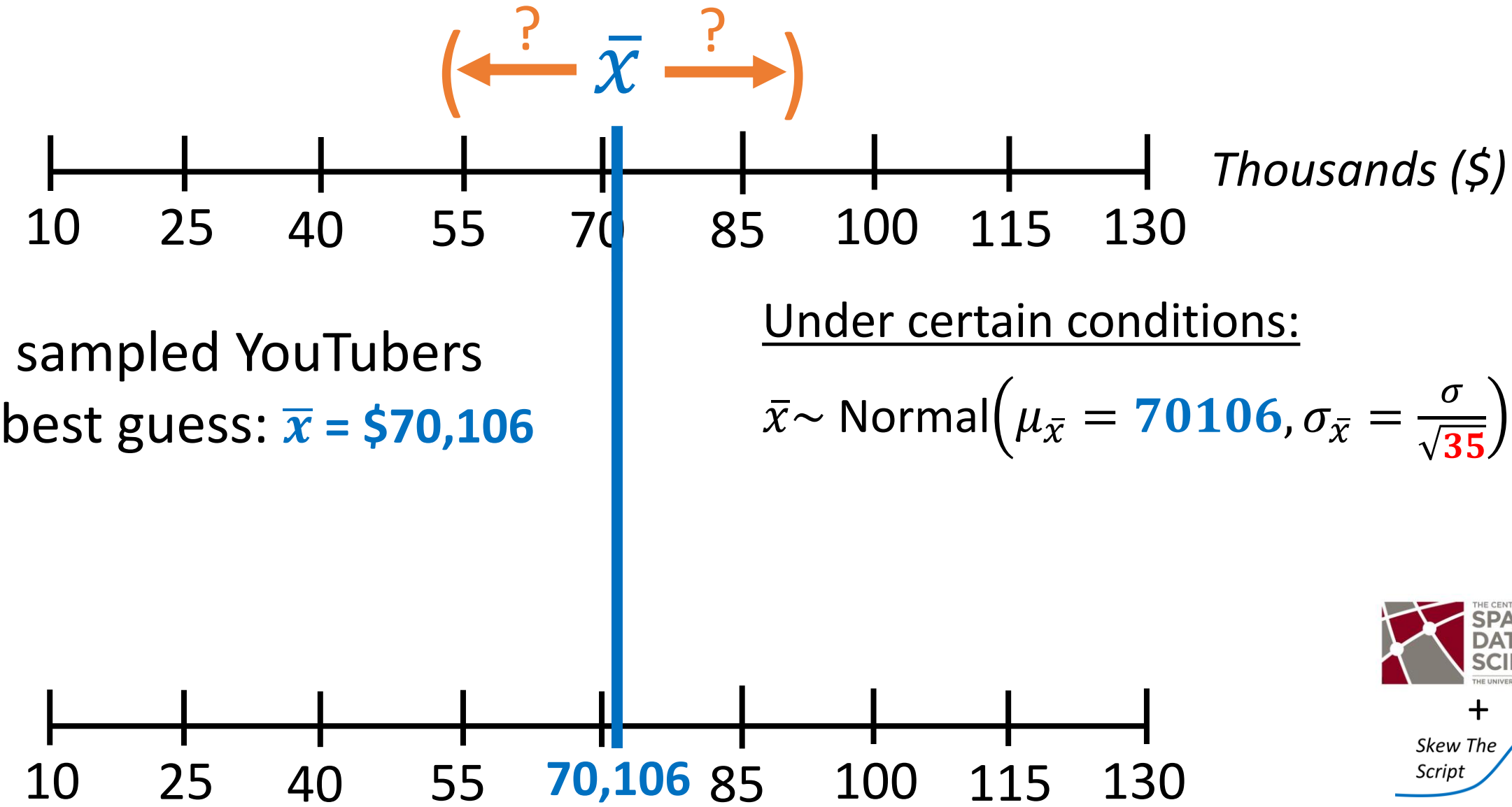
$$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \mu, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{35}}\right)$$

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



n = 35 sampled YouTubers

μ = ?, best guess: $\bar{x}$ **= $70,106**

Under certain conditions:

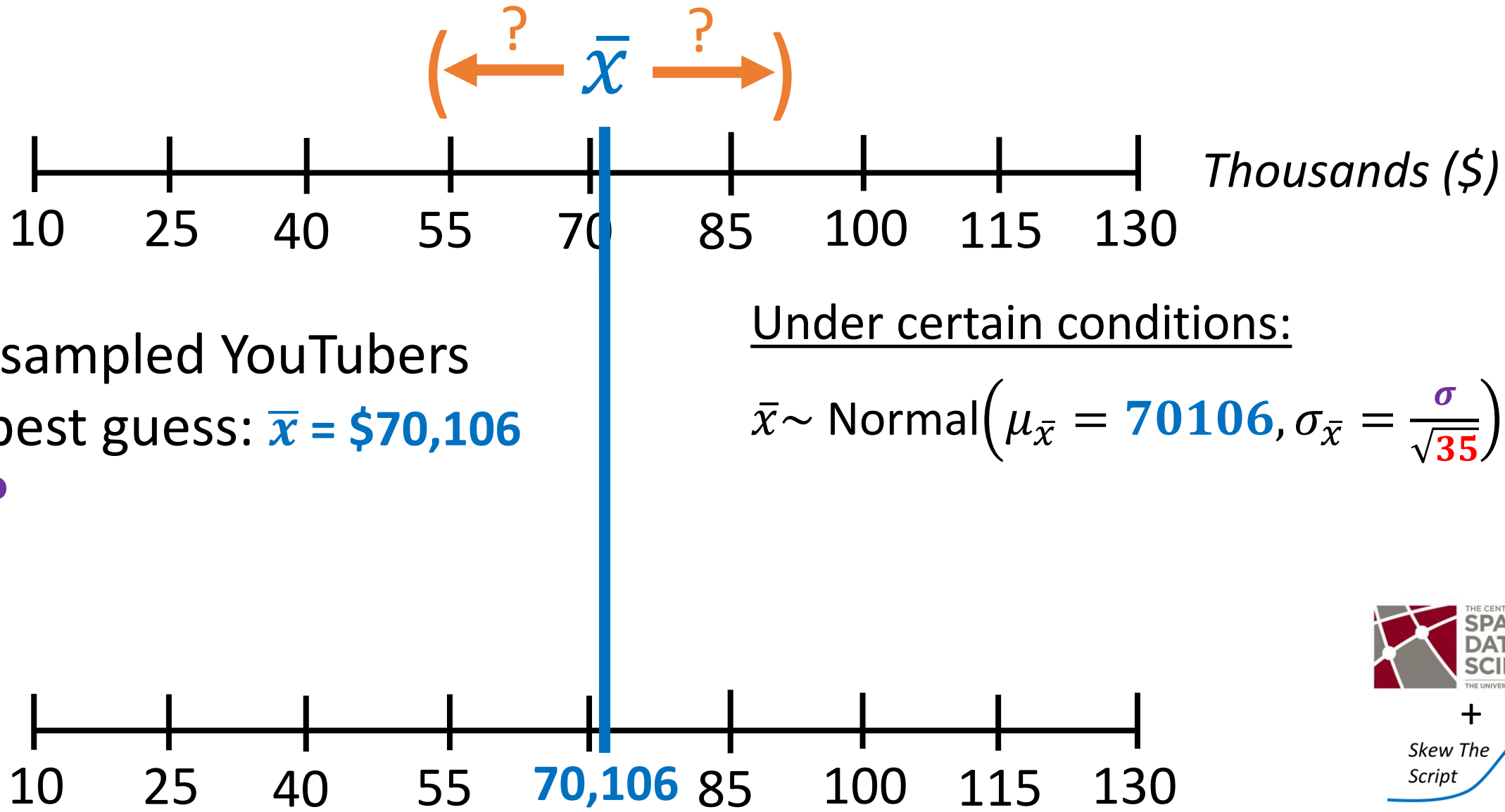
$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \mathbf{70106}, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{\mathbf{35}}}\right)$

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



n = 35 sampled YouTubers

μ = ?, best guess: $\bar{x}$ **= $70,106**

σ = ???

Under certain conditions:

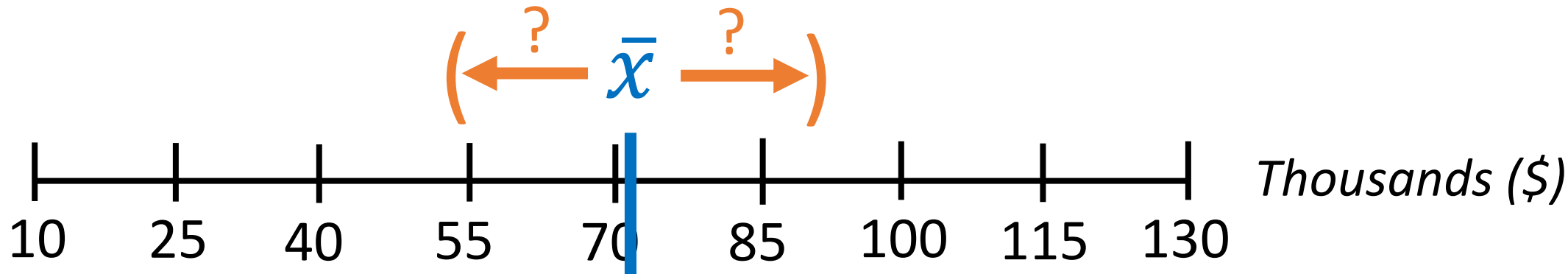
$$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \mathbf{70106}, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{\mathbf{35}}}\right)$$

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



n = 35 sampled YouTubers

μ = ?, best guess: $\bar{x}$ **= $70,106**

σ = ???

Under certain conditions:

$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \mathbf{70106}, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{\mathbf{35}}}\right)$

How do we estimate the true standard deviation (**σ**) of all YouTuber salaries?

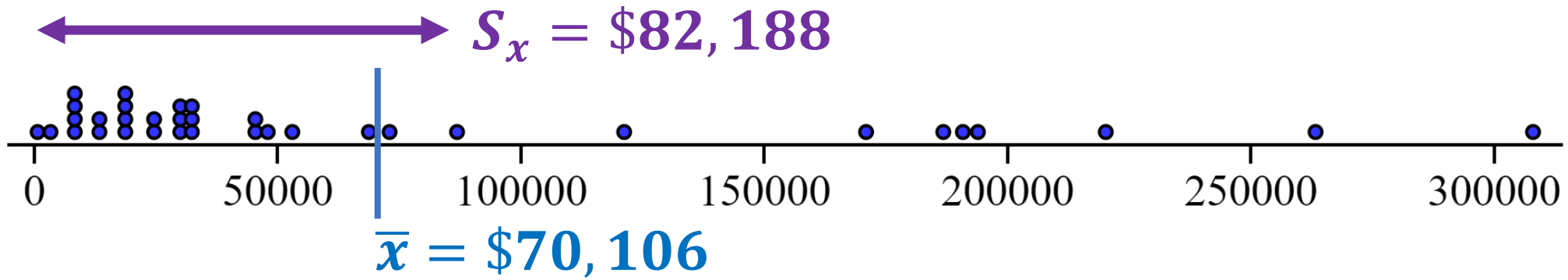
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Yearly Revenue
$706.74
$46,532.04
$9,565.32
$48,000.00
$23,801.40
$3,359.17
$12,649.08
$73,030.80
$170,945.60
$220,200.84
$8,743.92
$19,873.68
$31,118.19
$68,777.16
$186,819.69
$30,824.60
$32,405.16
$18,625.29
$8,030.69
$53,044.98
$44,379.24
$17,572.93
$263,309.52
$193,918.33
$308,000.00
$7,137.12
$25,487.88
$17,953.89
$28,948.20
$14,172.00
$121,239.03
$34,418.04
$86,893.74
$190,841.66
$32,398.80

Sampled YouTuber Yearly Revenues ($)



Dotplot created with *stapplet.com*

2020 poverty guideline from US HHS, 2019 mean wage from US SSA

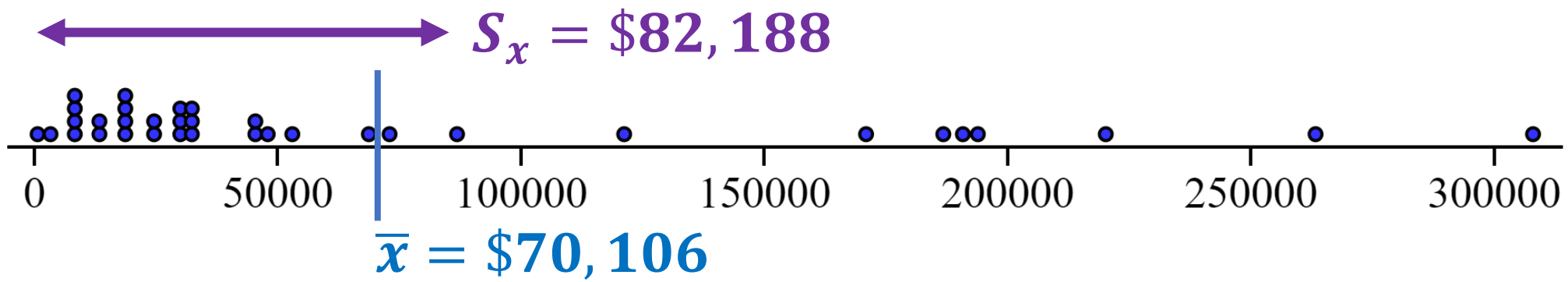
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Yearly Revenue
$706.74
$46,532.04
$9,565.32
$48,000.00
$23,801.40
$3,359.17
$12,649.08
$73,030.80
$170,945.60
$220,200.84
$8,743.92
$19,873.68
$31,118.19
$68,777.16
$186,819.69
$30,824.60
$32,405.16
$18,625.29
$8,030.69
$53,044.98
$44,379.24
$17,572.93
$263,309.52
$193,918.33
$308,000.00
$7,137.12
$25,487.88
$17,953.89
$28,948.20
$14,172.00
$121,239.03
$34,418.04
$86,893.74
$190,841.66
$32,398.80

Sampled YouTuber Yearly Revenues ($)



Use the **sample standard deviation (S_x)** from the data we already collected!

Dotplot created with *stapplet.com*

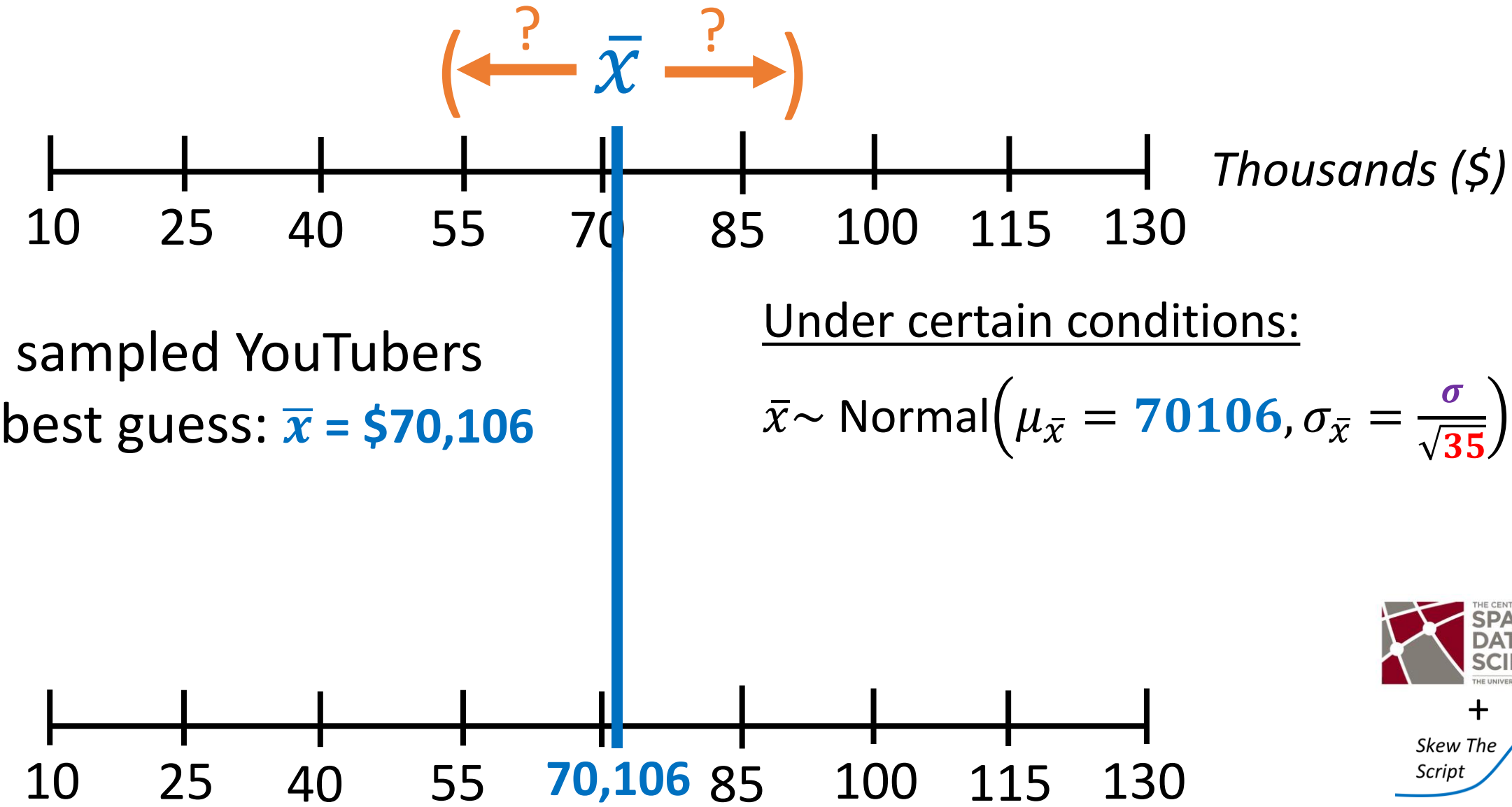
2020 poverty guideline from US HHS, 2019 mean wage from US SSA

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



n = 35 sampled YouTubers

μ = ?, best guess: $\bar{x}$ **= $70,106**

σ = ?

Under certain conditions:

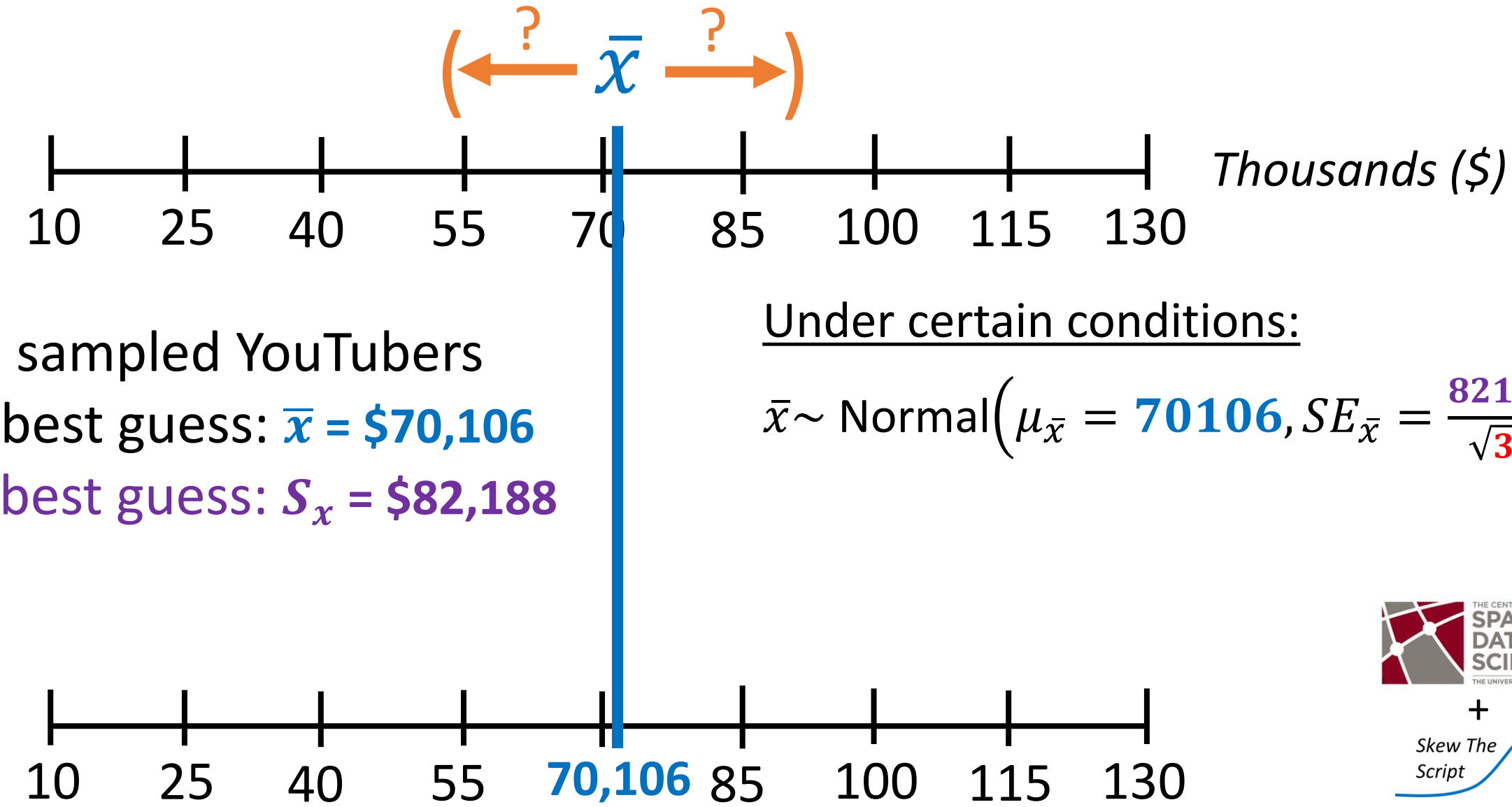
$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \mathbf{70106}, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{\mathbf{35}}}\right)$

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



n = 35 sampled YouTubers

μ = ?, best guess: $\bar{x}$ **= $70,106**

σ = ?, best guess: S_x **= $82,188**

Under certain conditions:

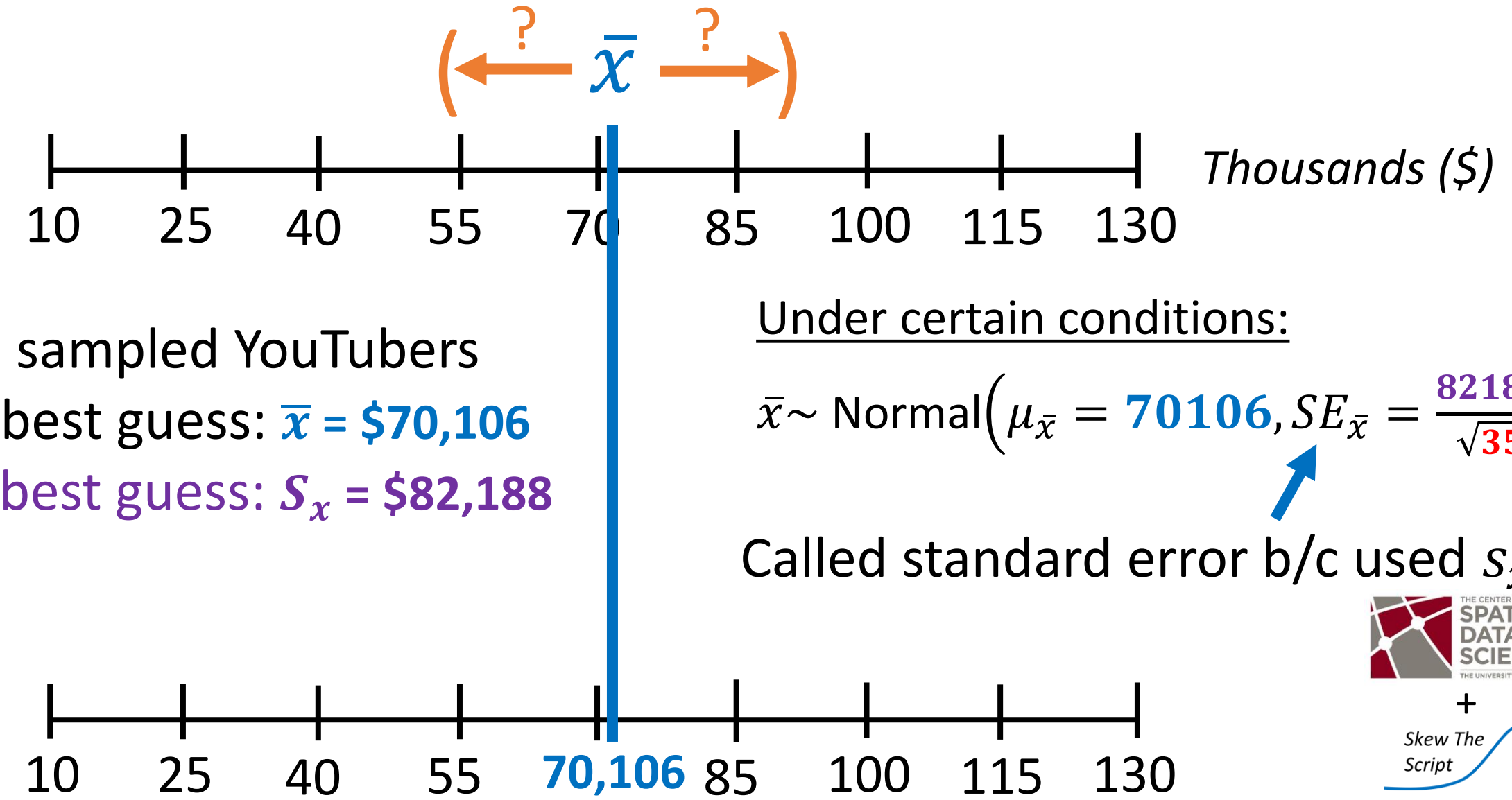
$$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \mathbf{70106}, SE_{\bar{x}} = \frac{\mathbf{82188}}{\sqrt{\mathbf{35}}}\right)$$

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



n = 35 sampled YouTubers

μ = ?, best guess: $\bar{x}$ **= $70,106**

σ = ?, best guess: S_x **= $82,188**

Under certain conditions:

$$\bar{x} \sim \text{Normal}\left(\mu_{\bar{x}} = \mathbf{70106}, SE_{\bar{x}} = \frac{\mathbf{82188}}{\sqrt{\mathbf{35}}}\right)$$

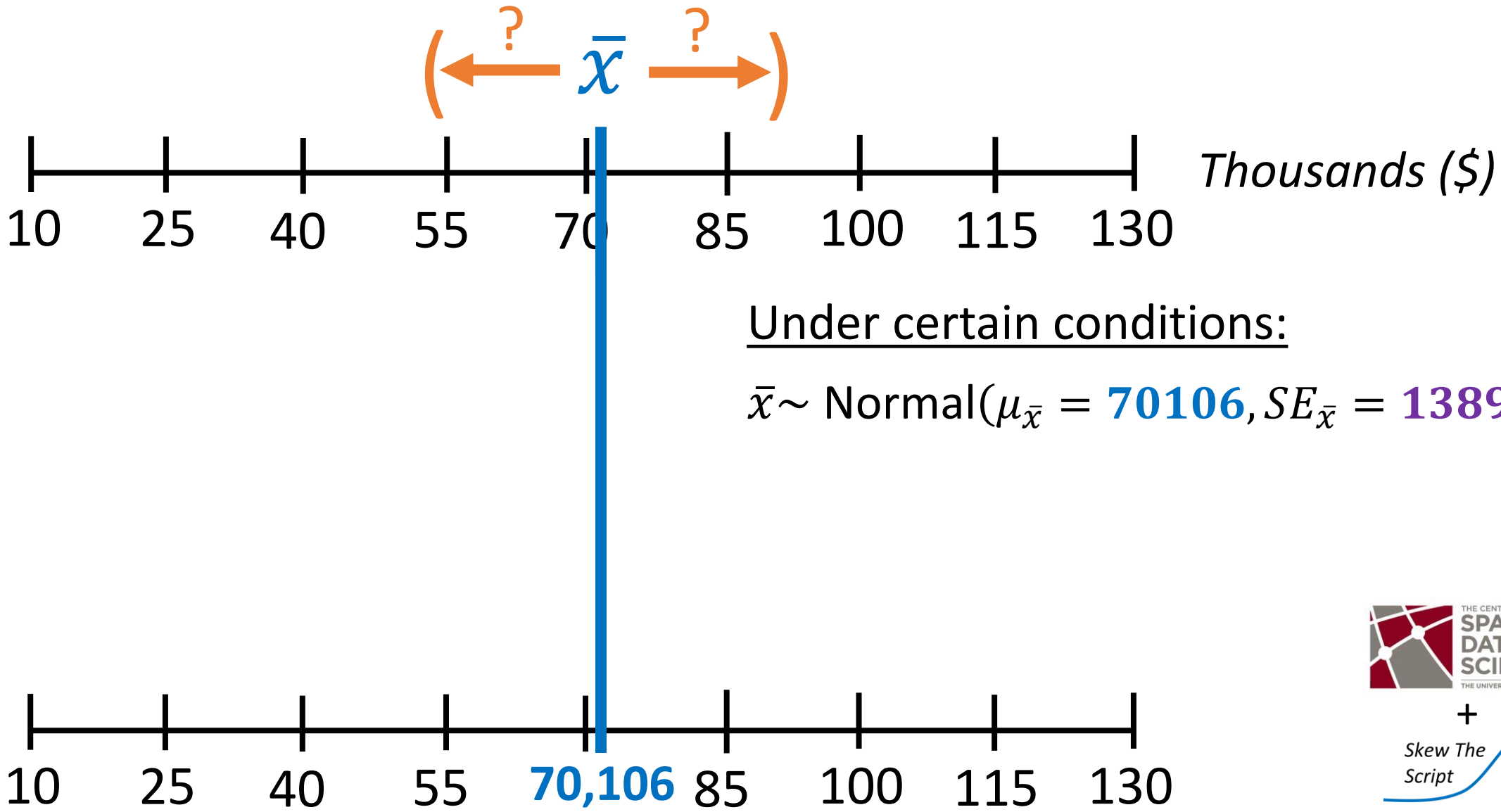
Called standard error b/c used s_x

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



Under certain conditions:

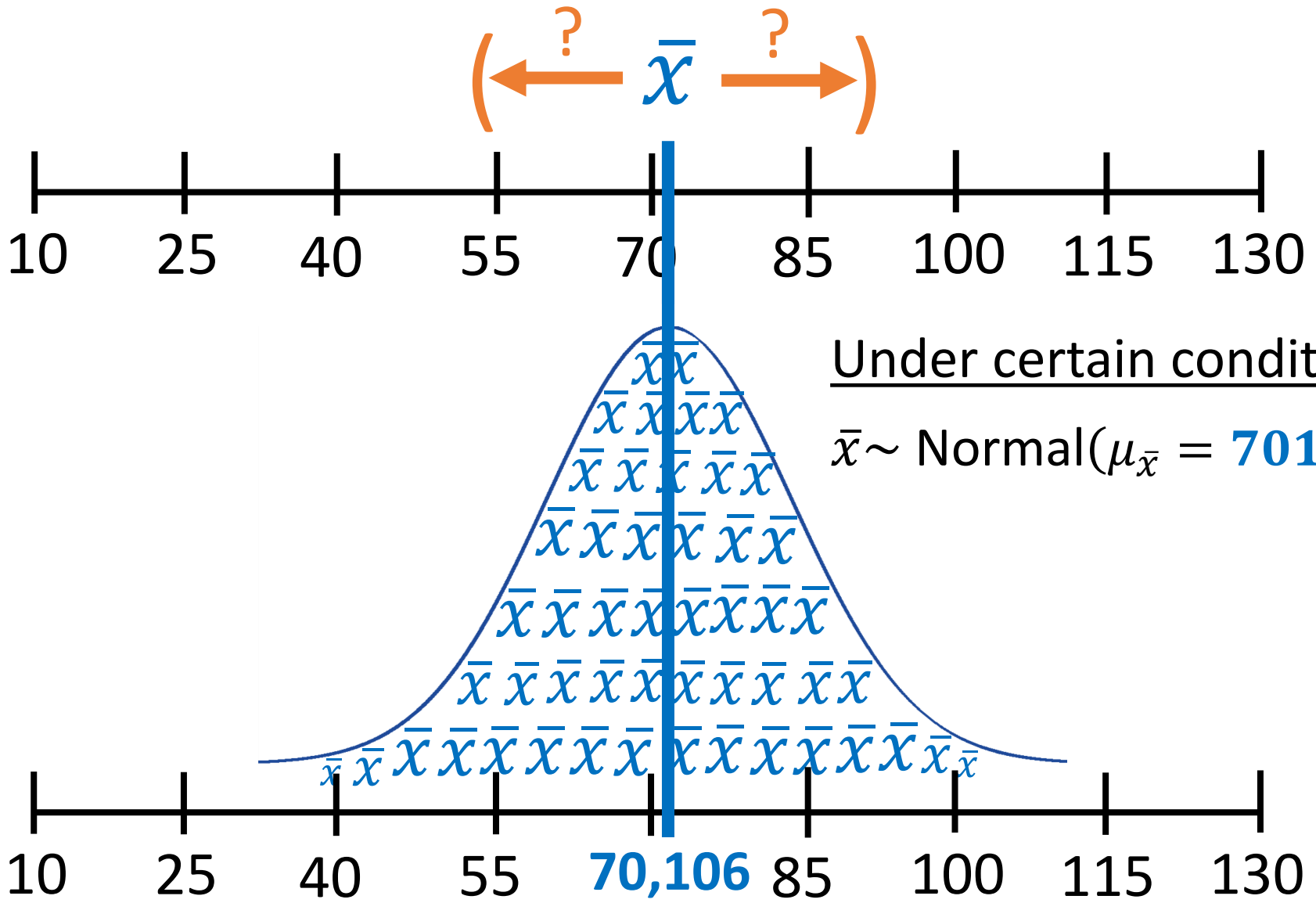
$\bar{x} \sim$ Normal($\mu_{\bar{x}} = \mathbf{70106}, SE_{\bar{x}} = \mathbf{13892}$)

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



Under certain conditions:

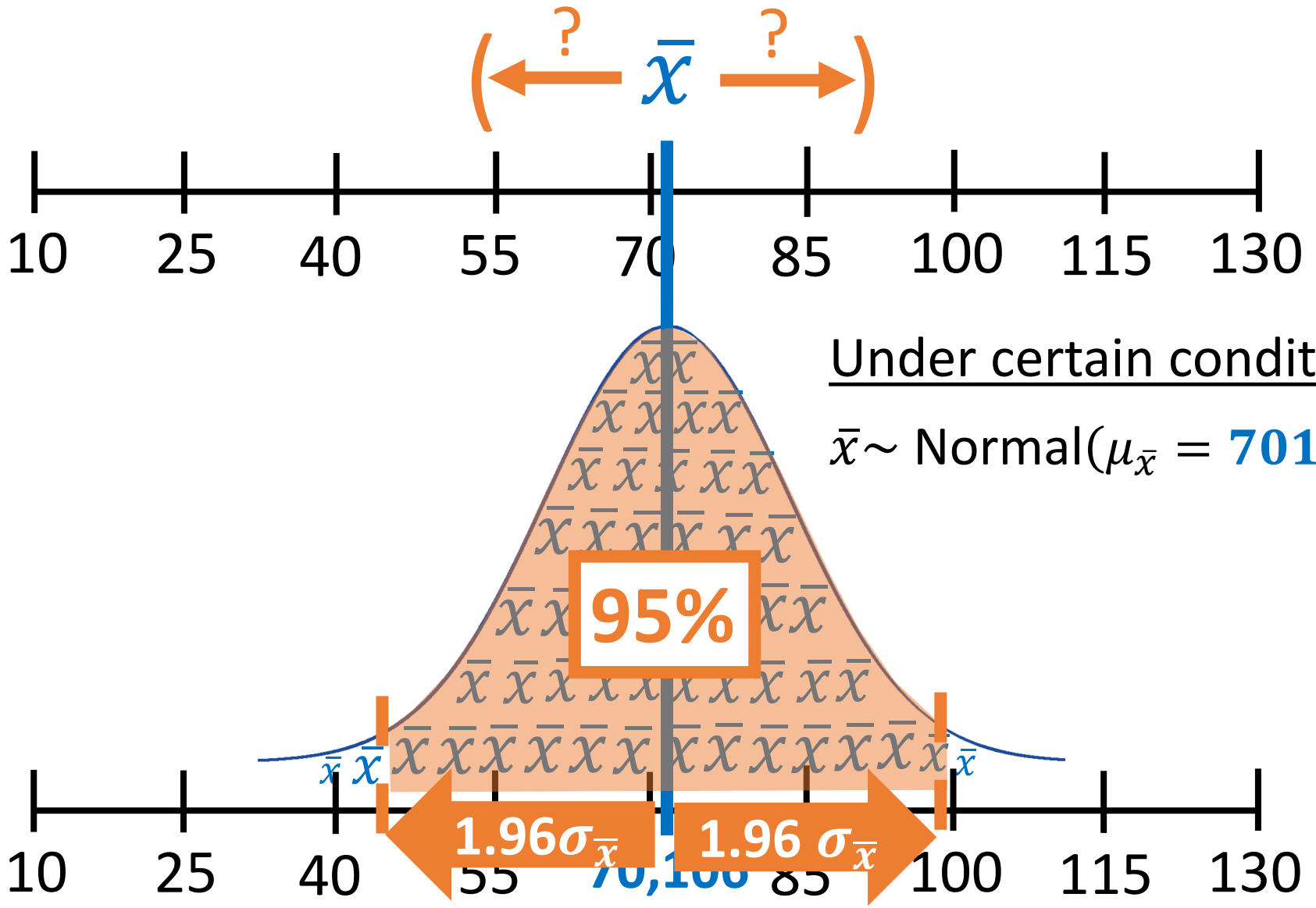
$\bar{x} \sim \text{Normal}(\mu_{\bar{x}} = \mathbf{70106}, SE_{\bar{x}} = \mathbf{13892})$

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



Under certain conditions:

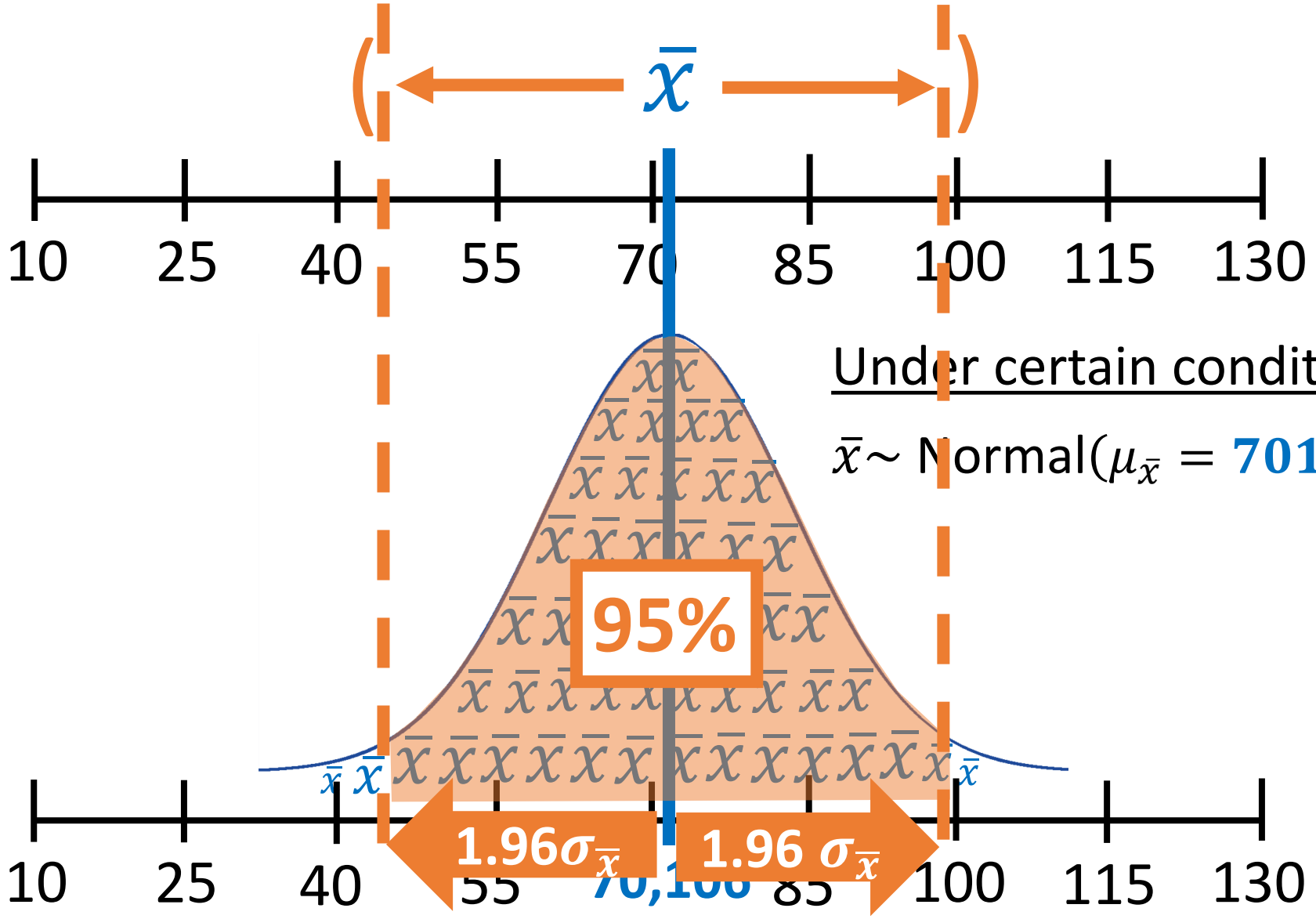
$\bar{x} \sim \text{Normal}(\mu_{\bar{x}} = \mathbf{70106}, SE_{\bar{x}} = \mathbf{13892})$

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



Under certain conditions:

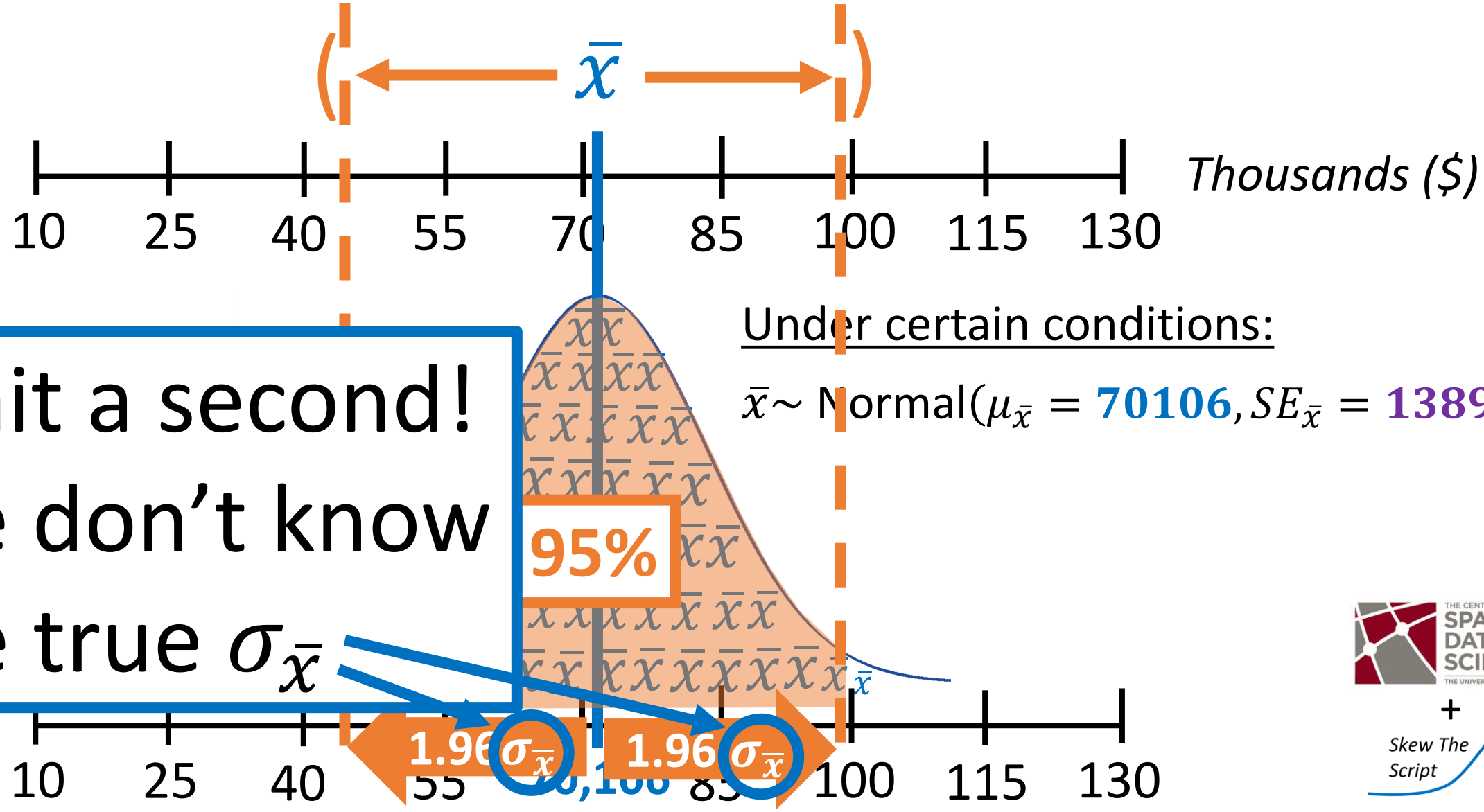
$\bar{x} \sim$ Normal($\mu_{\bar{x}} = \mathbf{70106}, SE_{\bar{x}} = \mathbf{13892}$)

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$



Wait a second!
We don't know the true $\sigma_{\bar{x}}$

Under certain conditions:

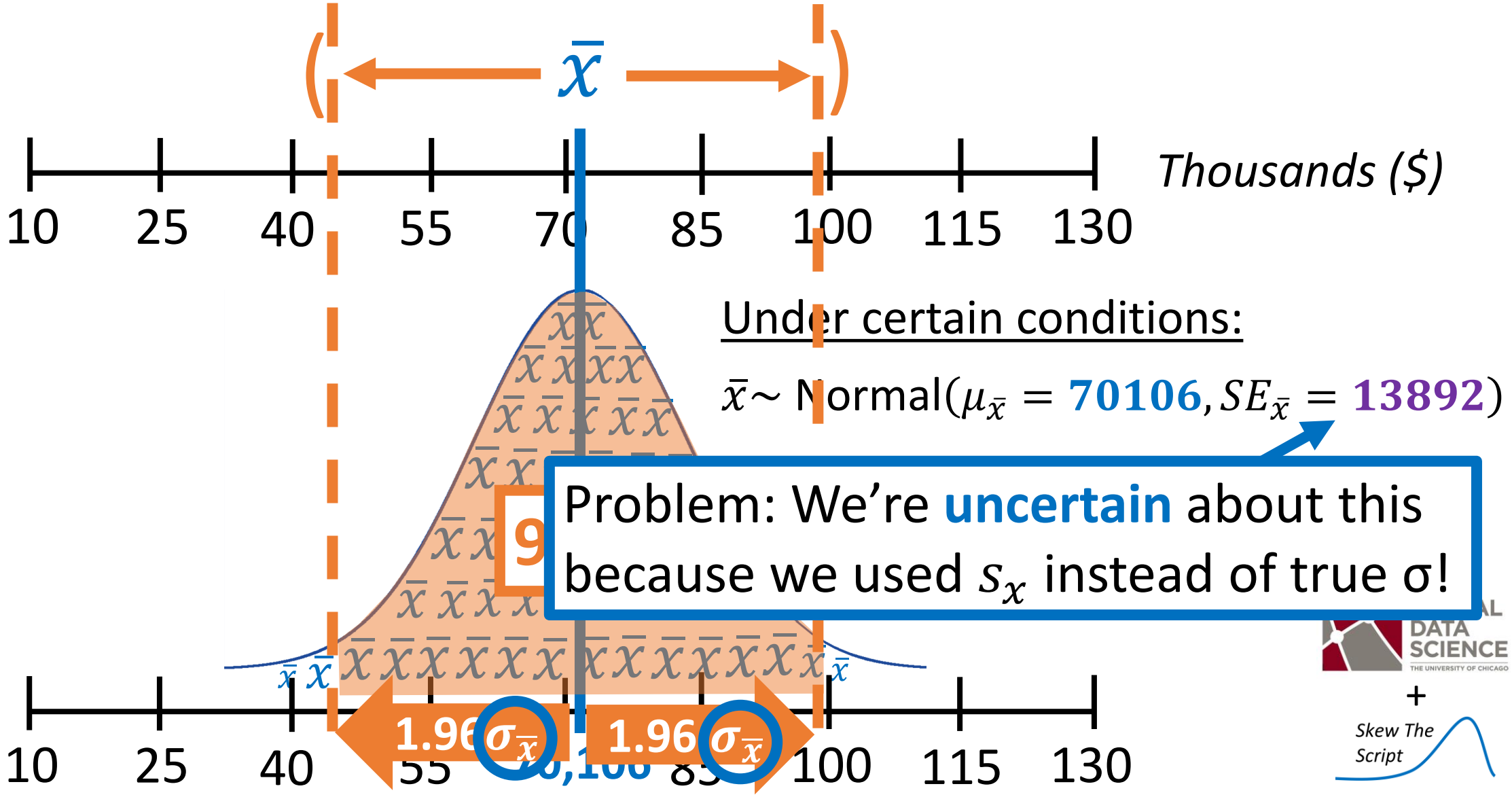
$\bar{x} \sim$ Normal($\mu_{\bar{x}} = \mathbf{70106}, SE_{\bar{x}} = \mathbf{13892}$)

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

μ = ?, best guess: $\bar{x} = \$70{,}106$

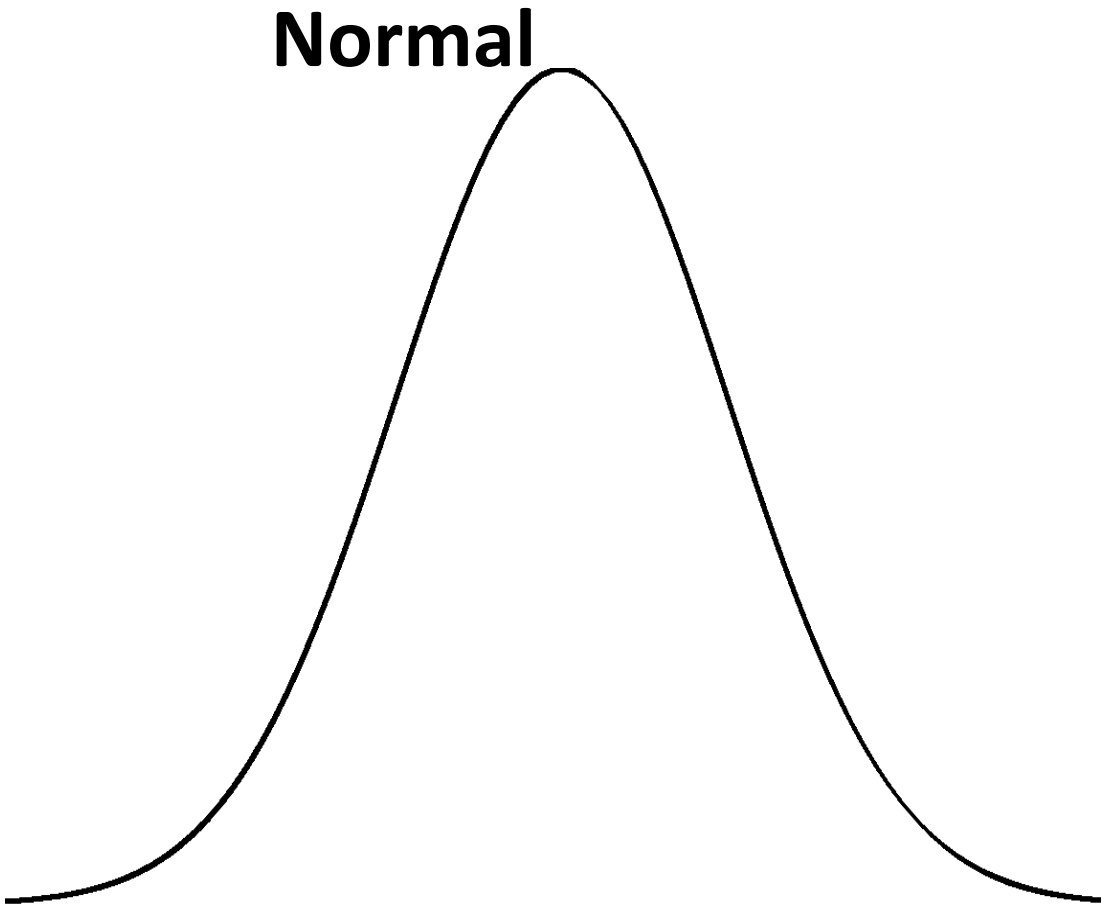


DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Solution: t-distribution

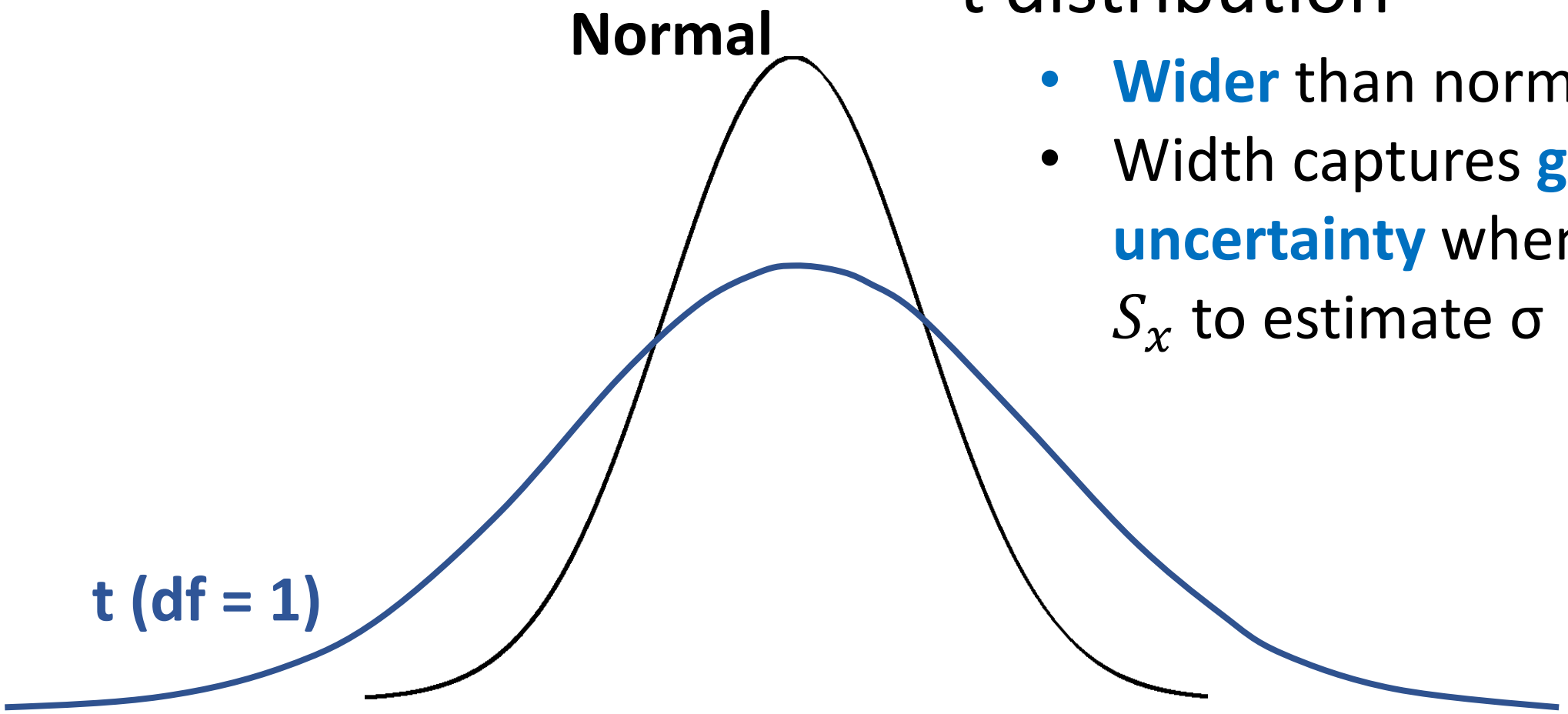


THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

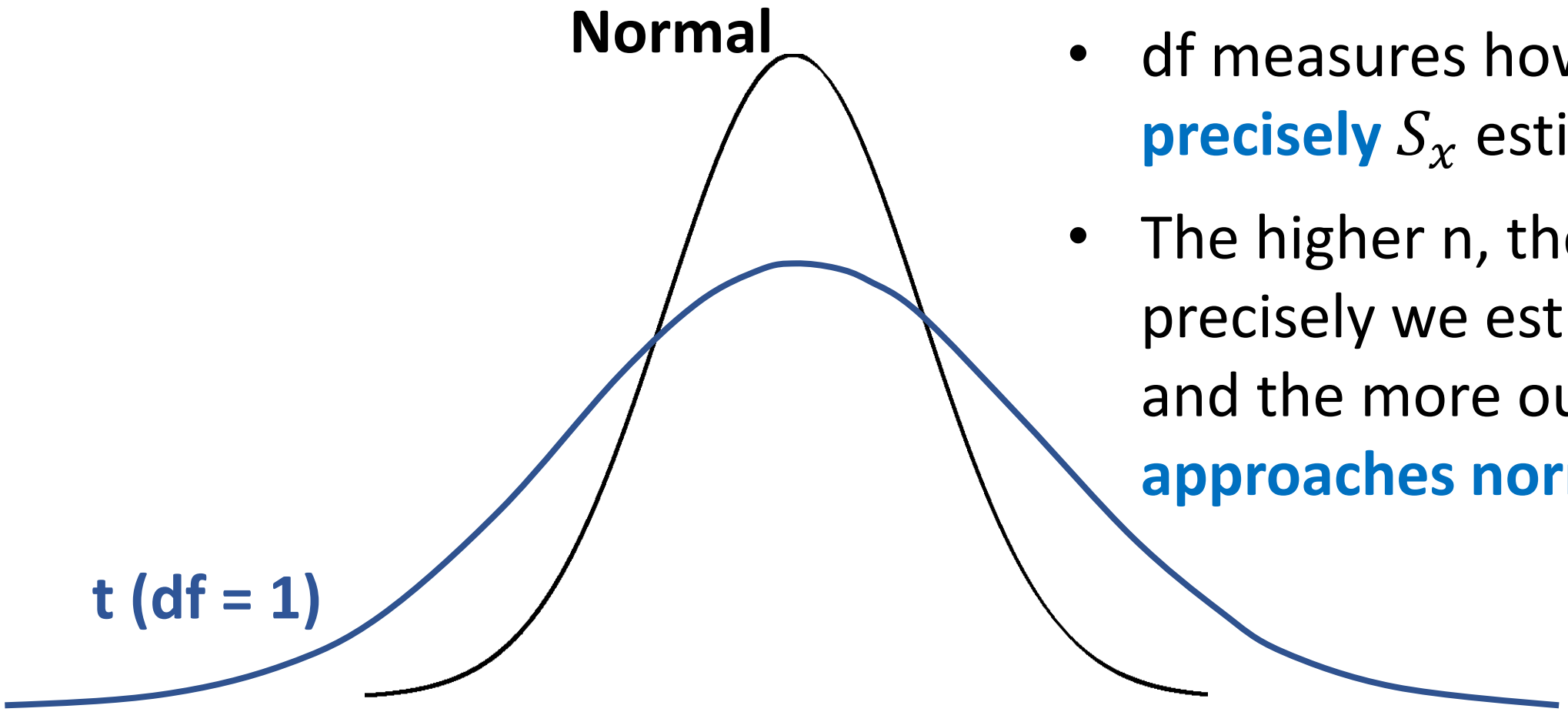
Solution: t-distribution



t distribution

- **Wider** than normal curve
- Width captures **greater uncertainty** when using S_x to estimate σ

Solution: t-distribution



degrees of freedom (df) = **n - 1**

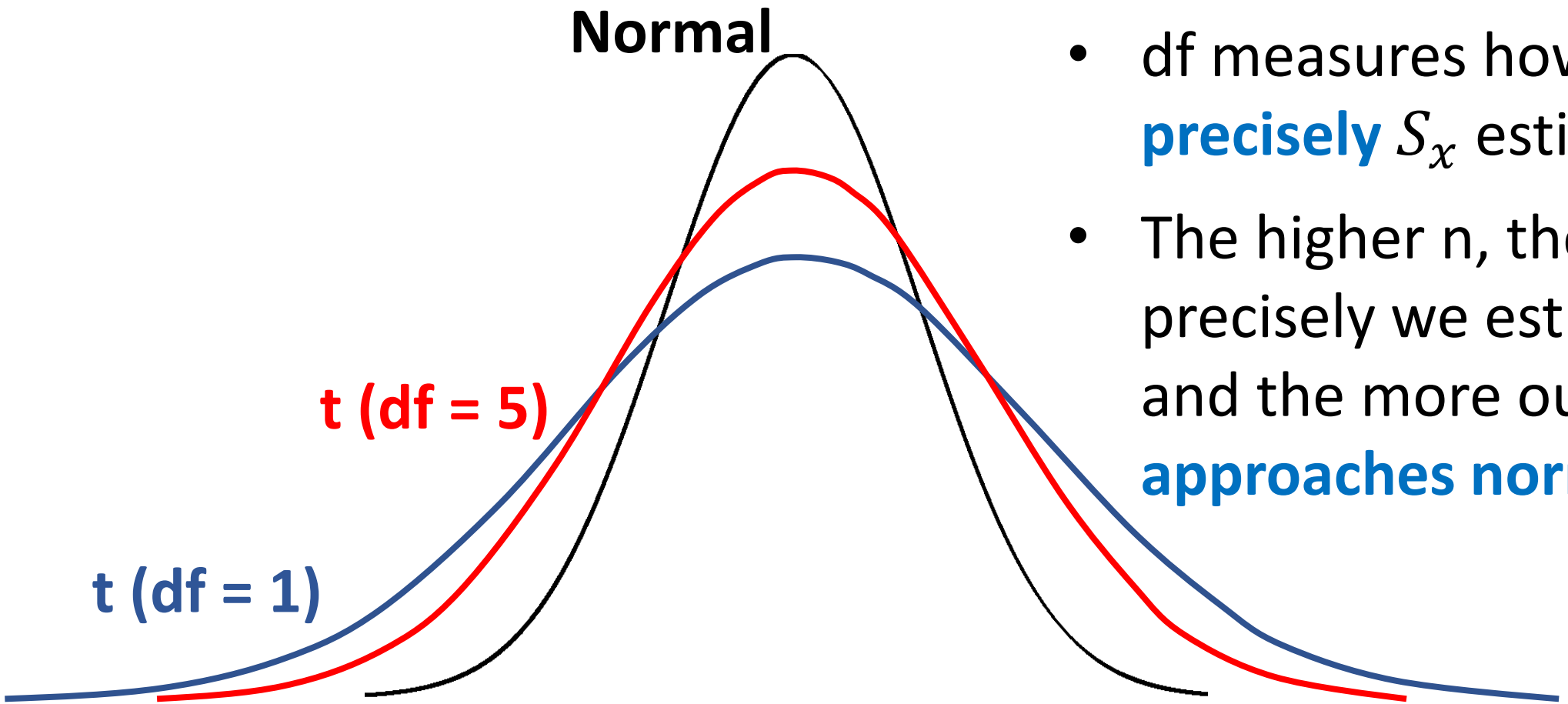
- df measures how **precisely** S_x estimates σ
- The higher n, the more precisely we estimate σ, and the more our t-curve **approaches normal!**

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Solution: t-distribution



degrees of freedom (df) = **n - 1**

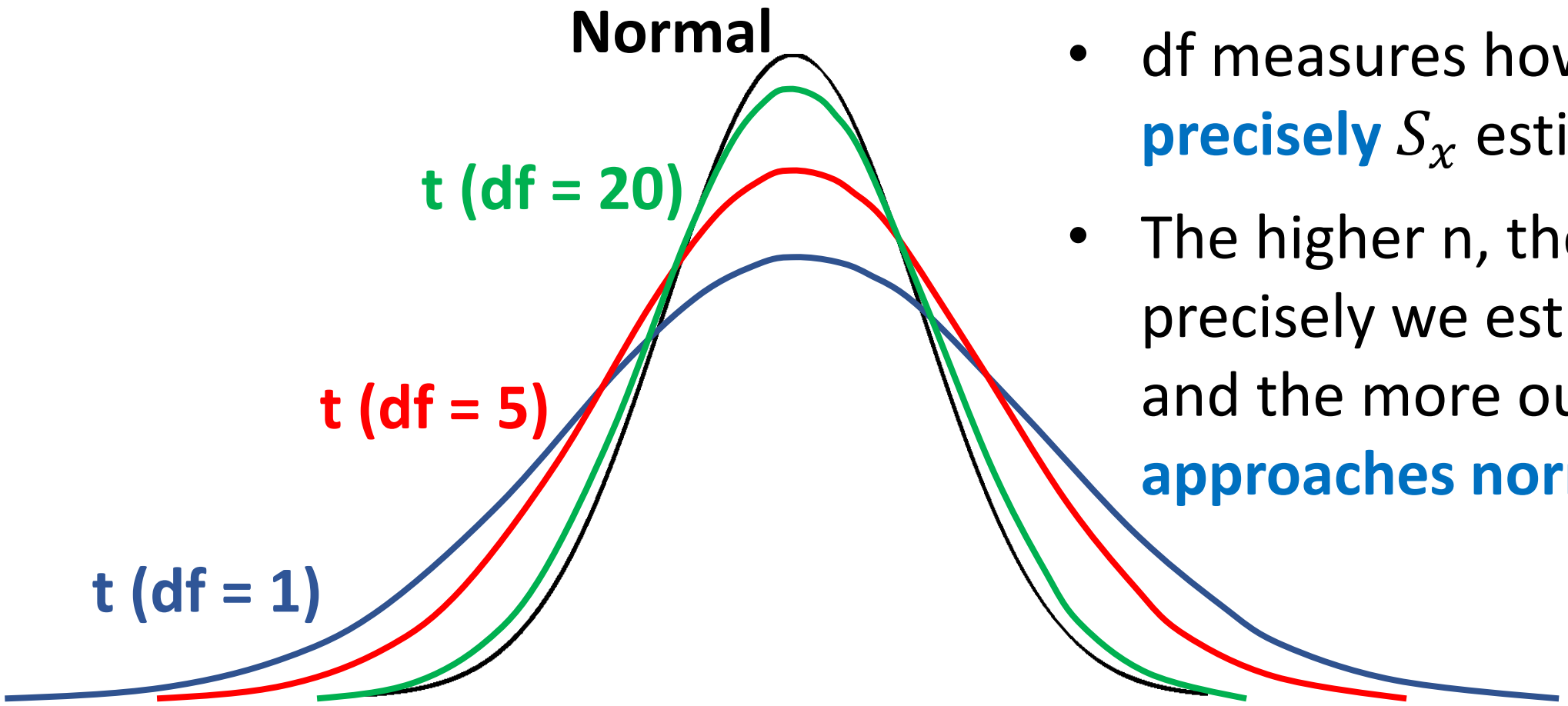
- df measures how **precisely** S_x estimates σ
- The higher n, the more precisely we estimate σ, and the more our t-curve **approaches normal!**

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Solution: t-distribution



degrees of freedom (df) = **n - 1**

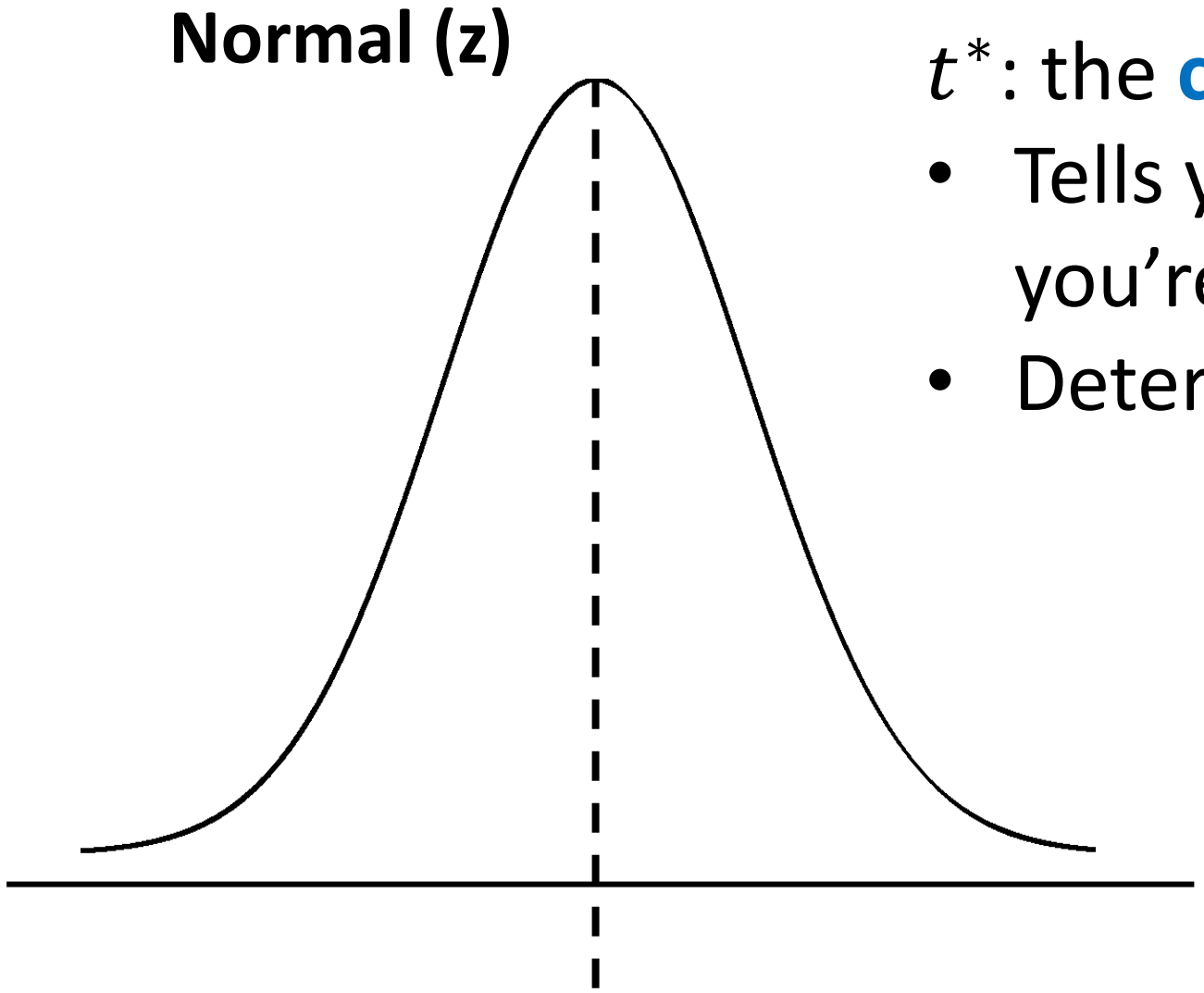
- df measures how **precisely** S_x estimates σ
- The higher n, the more precisely we estimate σ, and the more our t-curve **approaches normal!**

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

t critical values



t^*: the **critical value** of the t-interval

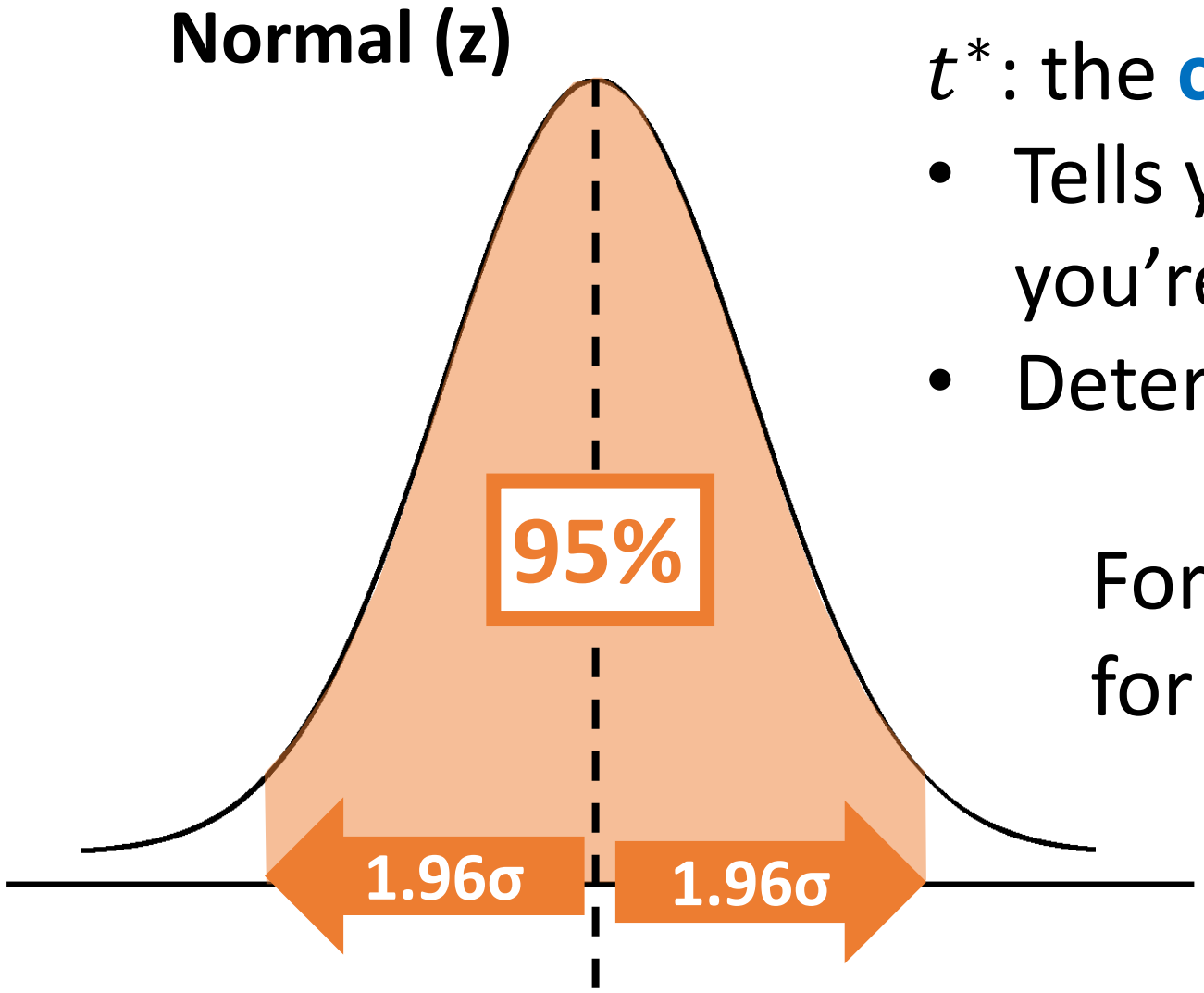
- Tells you how many **standard errors** you're including in your interval.
- Determines the **confidence level**.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

t critical values



t^*: the **critical value** of the t-interval

- Tells you how many **standard errors** you're including in your interval.
- Determines the **confidence level**.

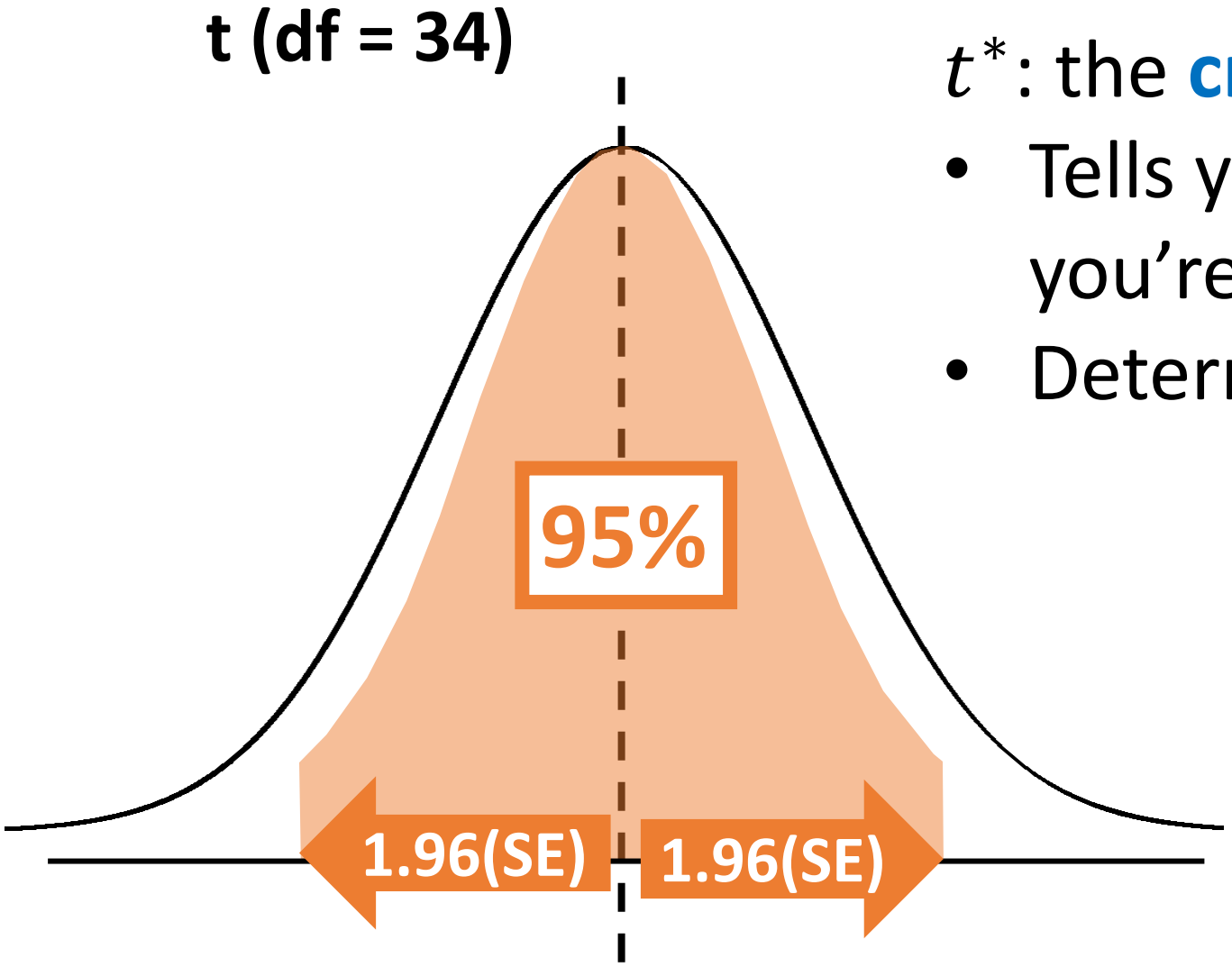
For normal curve, z* was 1.96 for 95% confidence

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

t critical values



t^*: the **critical value** of the t-interval

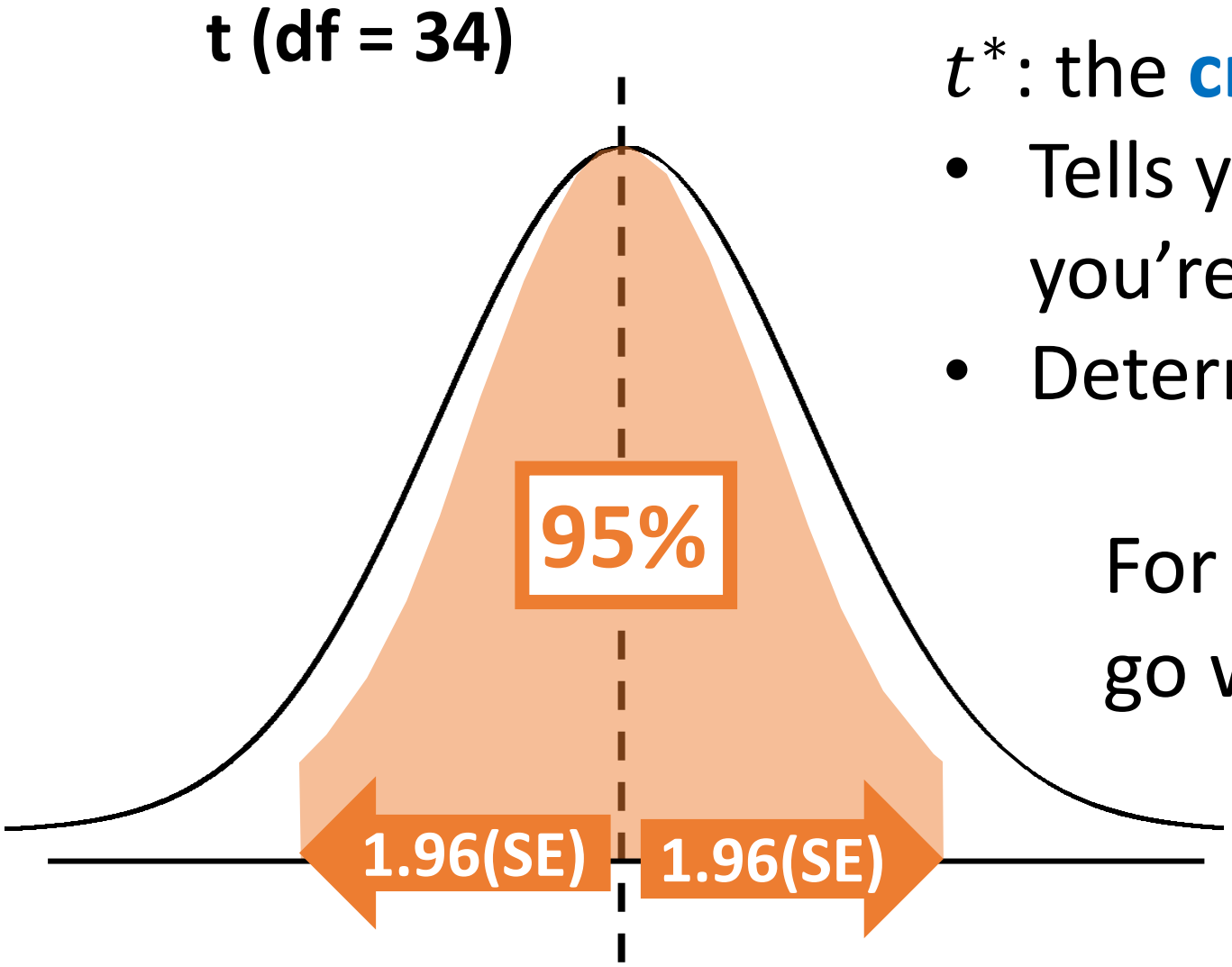
- Tells you how many **standard errors** you're including in your interval.
- Determines the **confidence level**.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

t critical values



t^*: the **critical value** of the t-interval

- Tells you how many **standard errors** you're including in your interval.
- Determines the **confidence level**.

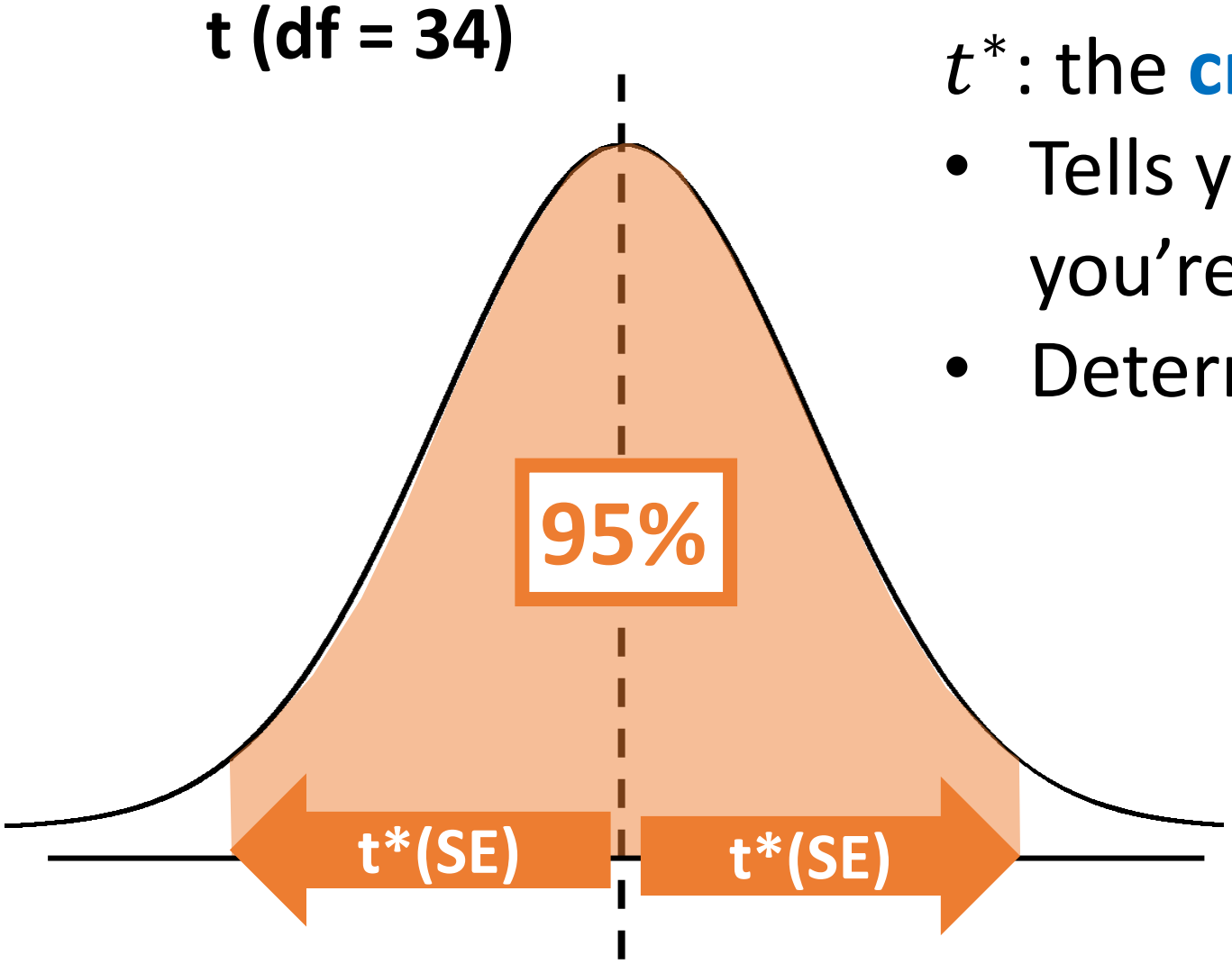
For t-distribution, we need to go wider to capture 95%

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

t critical values



t^*: the **critical value** of the t-interval

- Tells you how many **standard errors** you're including in your interval.
- Determines the **confidence level**.

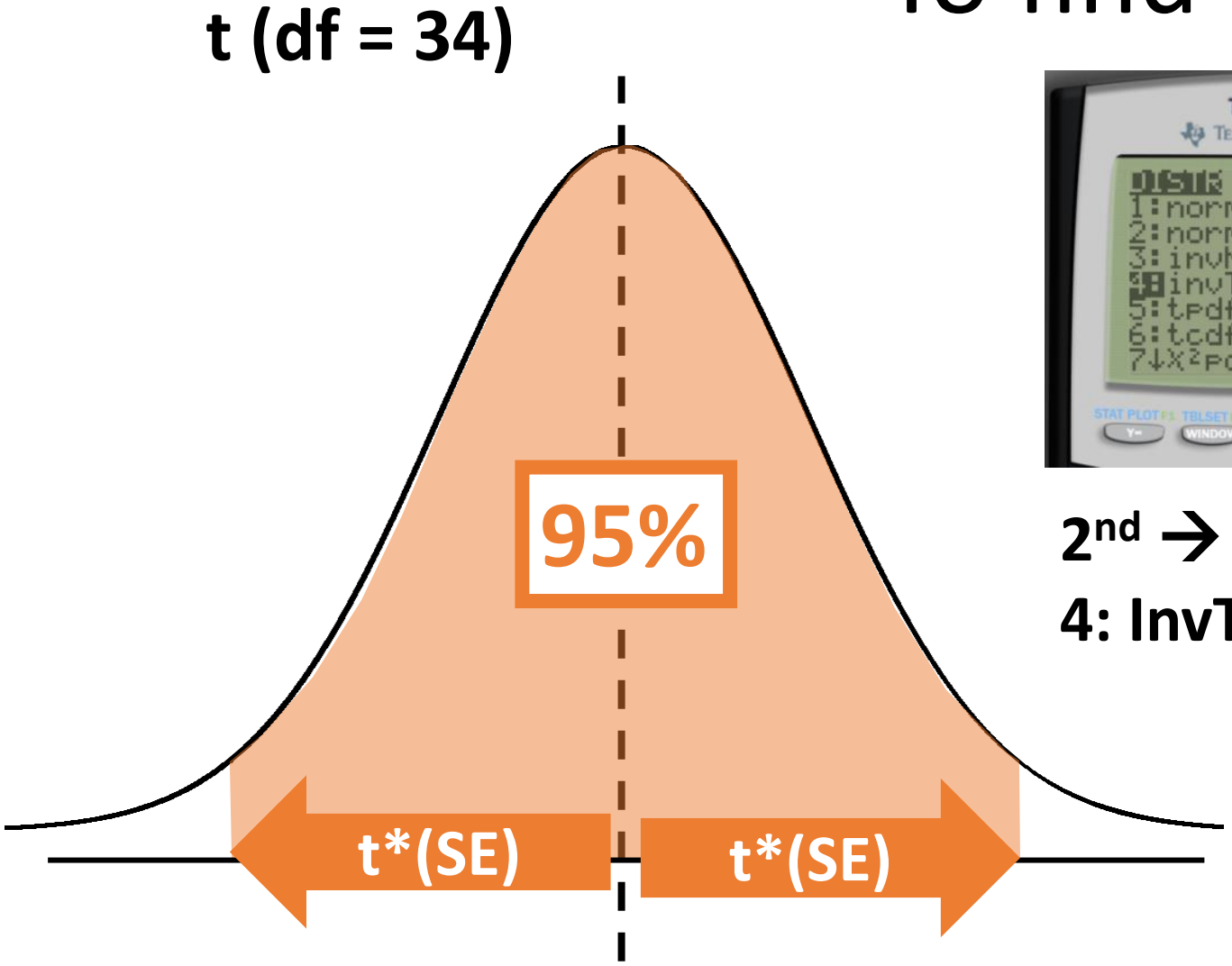
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

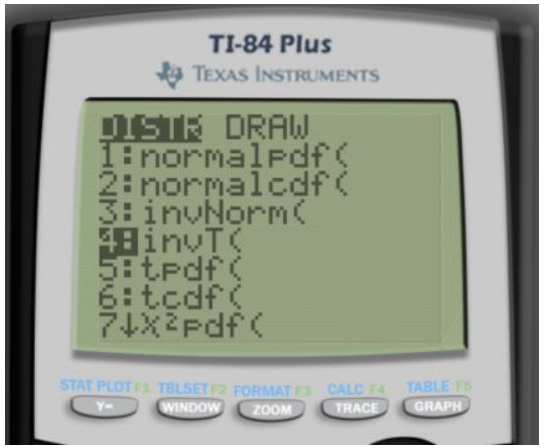
+

Skew The Script

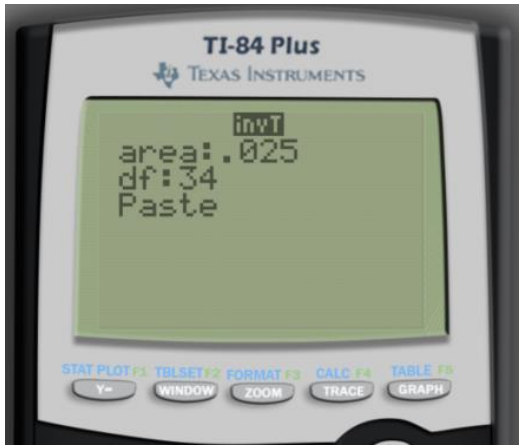
t critical values

To find t^*:





2nd → VARS → 4: InvT



-Area: percent below interval (2.5% = 0.025)

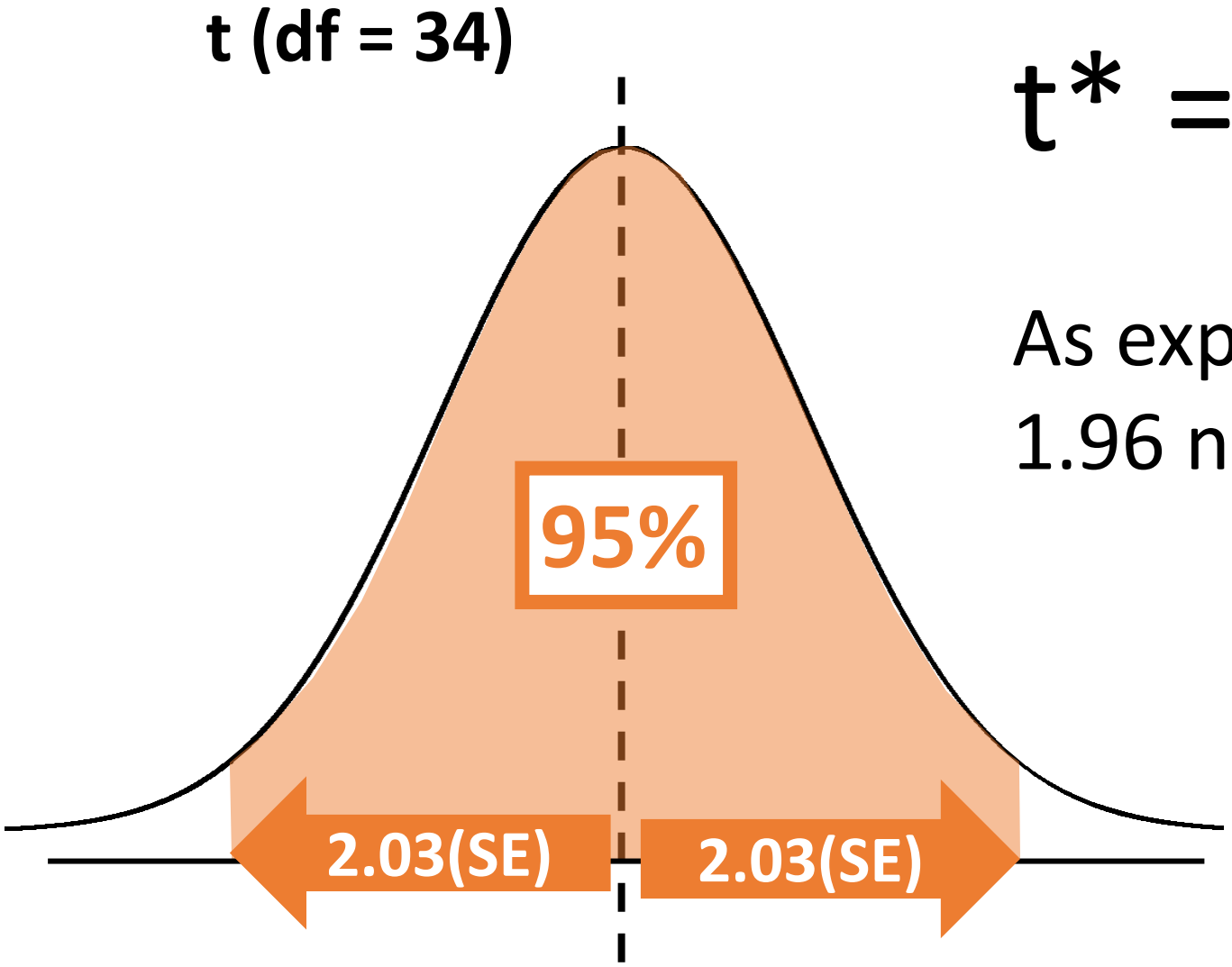
-df: n – 1 (35 – 1 = 34)

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

t critical values



$t^* = 2.03$

As expected, **slightly higher** than 1.96 needed for a normal curve.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

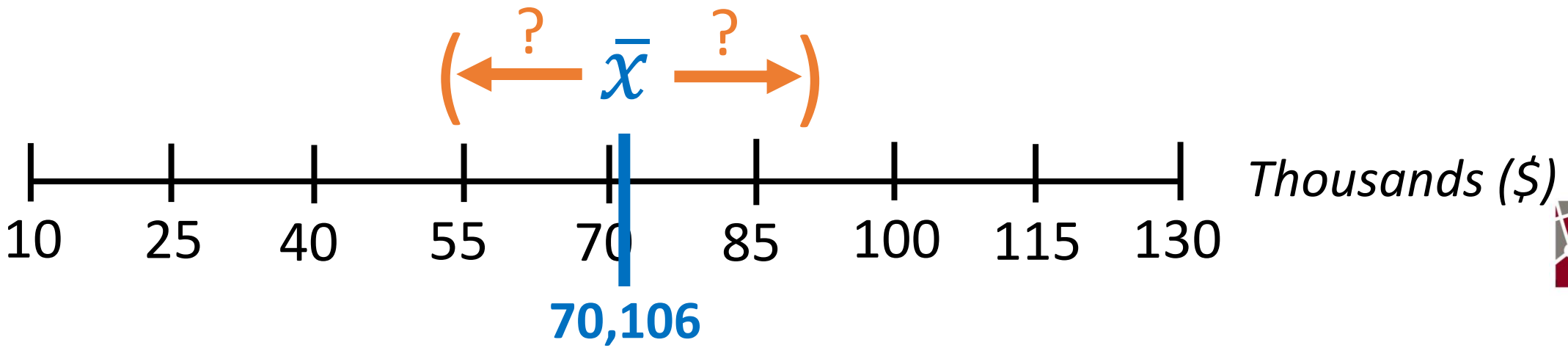
+

Skew The Script

Confidence interval for a mean

$$\text{point estimate} \pm \text{margin of error}$$

$$\bar{x} \pm t^*(SE_{\bar{x}})$$



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Confidence interval for a mean

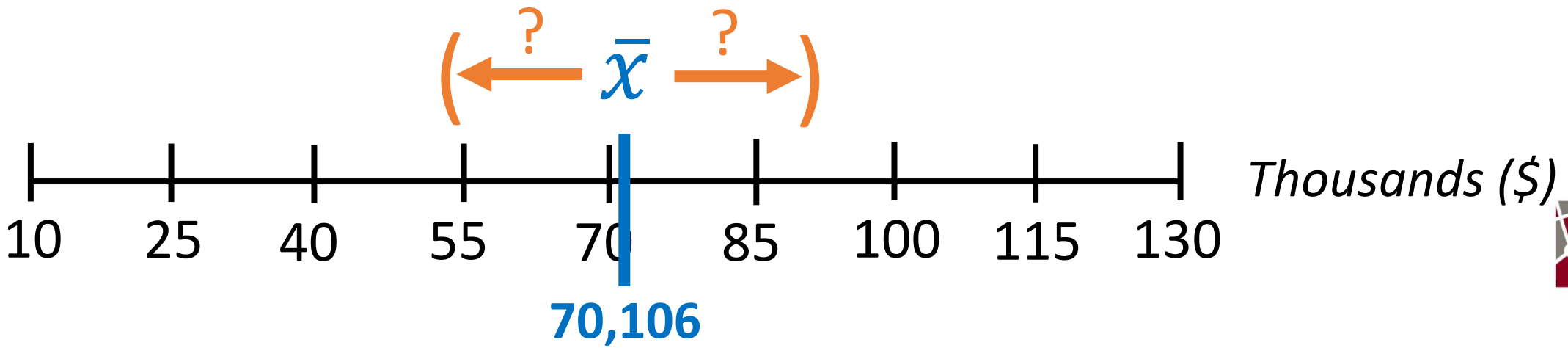
point estimate $\pm$ margin of error

$$\bar{x} \pm t^*(SE_{\bar{x}})$$

$$\bar{x} = \mathbf{70106}$$

$$SE_{\bar{x}} = \mathbf{13892}$$

t* = 2.03



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Confidence interval for a mean

point estimate $\pm$ margin of error

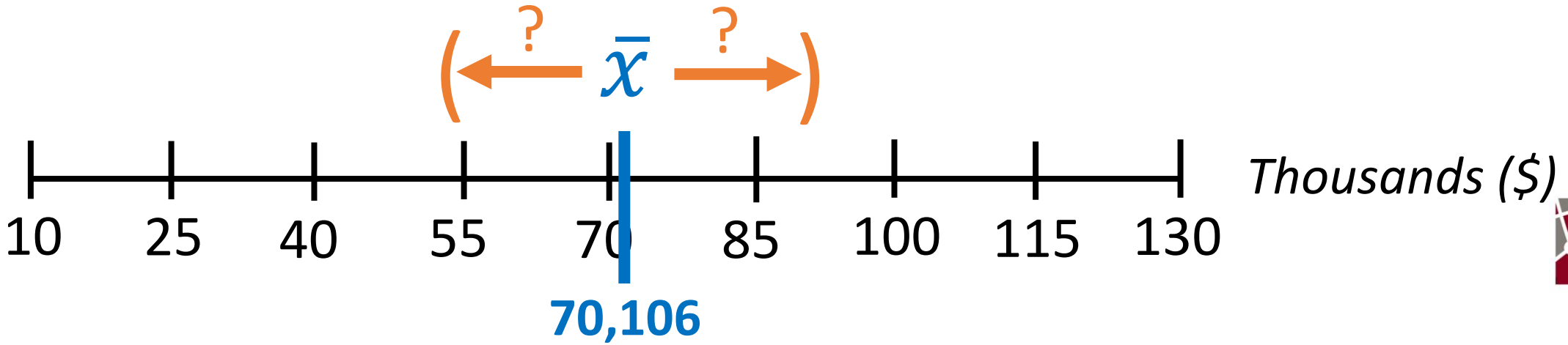
$\bar{x} \pm t^*(SE_{\bar{x}})$

$\mathbf{70106} \pm 2.03(\mathbf{13892})$

$\bar{x} = \mathbf{70106}$

$SE_{\bar{x}} = \mathbf{13892}$

t* = 2.03



Confidence interval for a mean

point estimate $\pm$ margin of error

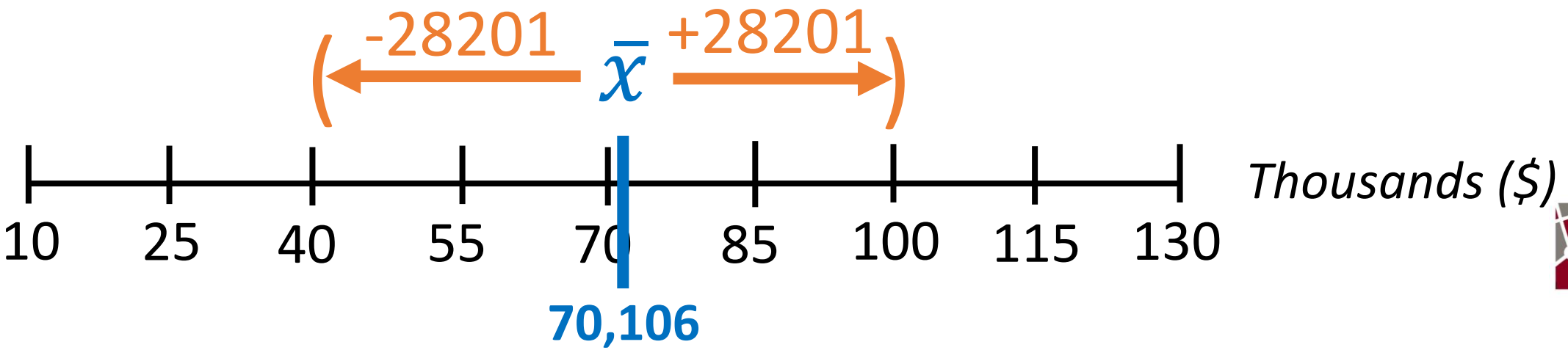
$\bar{x} \pm t^*(SE_{\bar{x}})$

$\mathbf{70106 \pm 28201}$

$\bar{x} = \mathbf{70106}$

$SE_{\bar{x}} = \mathbf{13892}$

t* = 2.03



THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

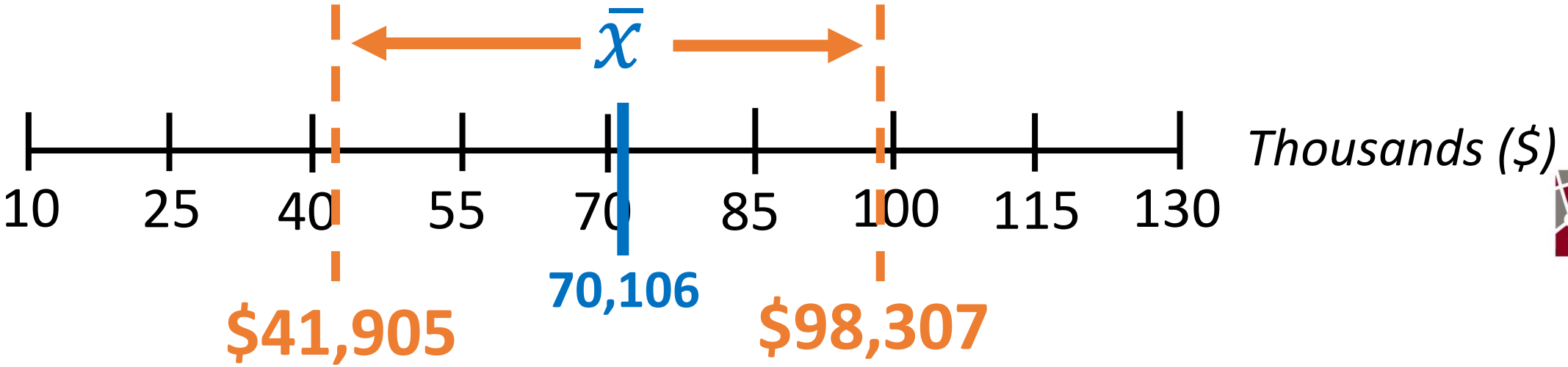
Skew The Script

Confidence interval for a mean

$$\text{point estimate} \pm \text{margin of error}$$

$$\bar{x} \pm t^*(SE_{\bar{x}})$$

($41905, $98307)



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

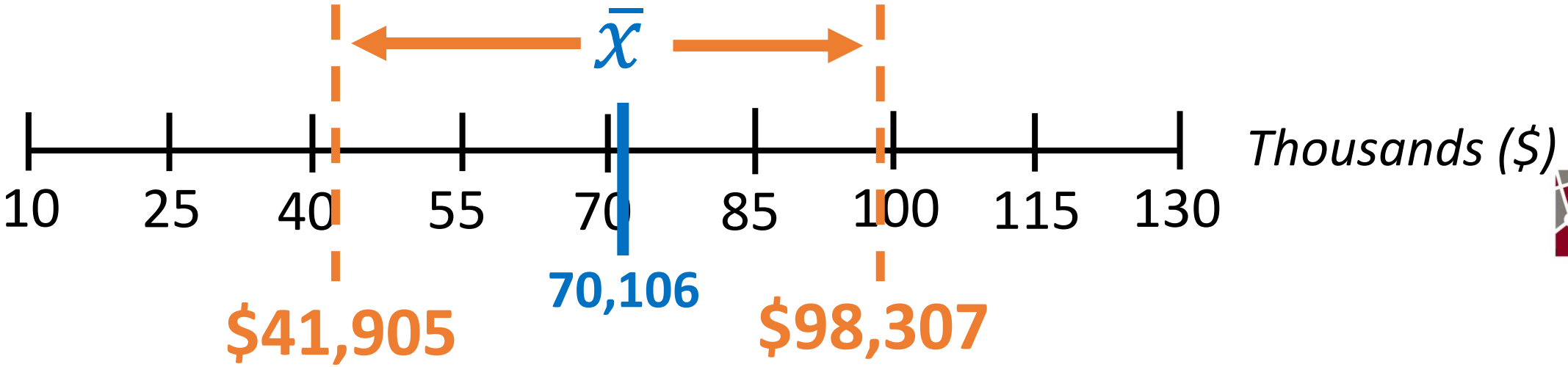
Confidence interval for a mean

$$\text{point estimate} \pm \text{margin of error}$$

$$\bar{x} \pm t^*(SE_{\bar{x}})$$

$$(\$41905, \$98307)$$

We are 95% confident the interval from \$41,905 to \$98,307 **captures** the **true mean** yearly income of YouTubers.



THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Topics

1. Recall: sampling distribution for $\bar{x}$
2. The t-distribution and interval for $\bar{x}$
3. **Four step process**

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

The Four Steps for Inference

*A suggested way to **organize** your work so that you get full credit on FRQ's!*

State: State the parameter you're estimating and the confidence level

Plan: Name your inference method and check conditions

Do: Perform calculations (if conditions met)

Conclude: Interpret your result with context

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Confidence Interval for $\hat{p}$

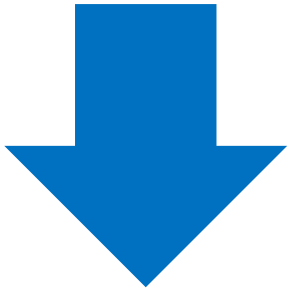
We obtained a random sample of 35 “How much I make” YouTube videos. Among the sample, the mean yearly income was $70,106 and the standard deviation was $82,188.

Construct and interpret a 95% confidence interval for the true mean salary of all YouTubers.

Confidence Interval for $\hat{p}$

We obtained a random sample of 35 “How much I make” YouTube videos. Among the sample, the mean yearly income was $70,106 and the standard deviation was $82,188.

Construct and interpret a 95% confidence interval for the true mean salary of all YouTubers.



Four Steps for Inference

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-Plan-Do-Conclude

State: State the parameter you're estimating and the confidence level

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-Plan-Do-Conclude

State: State the parameter you're estimating and the confidence level

We are estimating the **true mean yearly YouTuber salary** (μ), at **95% confidence.**

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-**Plan**-Do-Conclude

Plan: Name your inference method and check conditions

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-**Plan**-Do-Conclude

Plan: Name your inference method and check conditions

We will calculate a **one-sample t-interval** for μ, if all conditions are met.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-**Plan**-Do-Conclude

Plan: Name your inference method and check conditions

We will calculate a **one-sample t-interval** for μ, if all conditions are met.

Conditions

1. **Random:**

2. **10%:**

3. **Normal/Large Sample:**

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

Recall: Why we check conditions

$$\bar{x} \sim Norm\left(\mu_{\bar{x}} = \mu, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\right)$$

Note: We us a t-dist when substituting s_x for σ

3) Normal/Large Sample
→ approx. normal **shape** (by CLT)

2) 10% condition
→ calculable **spread**

1) Random condition
→ unbiased **center**

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-**Plan**-Do-Conclude

Plan: Name your inference method and check conditions

We will calculate a **one-sample t-interval** for μ, if all conditions are met.

Conditions

1. **Random:**

2. **10%:**

3. **Normal/Large Sample:**

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-**Plan**-Do-Conclude

Plan: Name your inference method and check conditions

We will calculate a **one-sample t-interval** for μ, if all conditions are met.

Conditions

1. **Random:** The sample of 35 YouTubers was collected randomly 

2. **10%:**

3. **Normal/Large Sample:**

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-**Plan**-Do-Conclude

Plan: Name your inference method and check conditions

We will calculate a **one-sample t-interval** for μ, if all conditions are met.

Conditions

1. **Random:** The sample of 35 YouTubers was collected randomly 

2. **10%:** $n \leq 0.10N$

3. **Normal/Large Sample:**

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-**Plan**-Do-Conclude

Plan: Name your inference method and check conditions

We will calculate a **one-sample t-interval** for μ, if all conditions are met.

Conditions

1. **Random:** The sample of 35 YouTubers was collected randomly 

2. **10%:** $n \leq 0.10N$

$35 \leq 0.10$ (all YouTubers)

3. **Normal/Large Sample:**

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-**Plan**-Do-Conclude

Plan: Name your inference method and check conditions

We will calculate a **one-sample t-interval** for μ, if all conditions are met.

Conditions

1. **Random:** The sample of 35 YouTubers was collected randomly

2. **10%:** $n \leq 0.10N$

$35 \leq 0.10$ (all YouTubers)

It's reasonable to assume 35 is less than 10% of all YouTubers

3. **Normal/Large Sample:**

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-**Plan**-Do-Conclude

Plan: Name your inference method and check conditions

We will calculate a **one-sample t-interval** for μ, if all conditions are met.

Conditions

1. **Random:** The sample of 35 YouTubers was collected randomly ✓

2. **10%:** $n \leq 0.10N$

$35 \leq 0.10$ (all YouTubers)

It's reasonable to assume 35 is less than 10% of all YouTubers ✓

3. **Normal/Large Sample:**

$n \geq 30$

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-**Plan**-Do-Conclude

Plan: Name your inference method and check conditions

We will calculate a **one-sample t-interval** for μ, if all conditions are met.

Conditions

1. **Random:** The sample of 35 YouTubers was collected randomly

2. **10%:** $n \leq 0.10N$

$35 \leq 0.10$ (all YouTubers)

It's reasonable to assume 35 is less than 10% of all YouTubers

3. **Normal/Large Sample:**

$$n \geq 30$$

$$35 \geq 30$$

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-**Plan**-Do-Conclude

Plan: Name your inference method and check conditions

We will calculate a **one-sample t-interval** for μ, if all conditions are met.

Conditions

1. **Random:** The sample of 35 YouTubers was collected randomly ✓

2. **10%:** $n \leq 0.10N$

$35 \leq 0.10$ (all YouTubers)

It's reasonable to assume 35 is less than 10% of all YouTubers ✓

3. **Normal/Large Sample:**

$n \geq 30$

$35 \geq 30$ ✓

The sample size is large

+

State-Plan-**Do**-Conclude

Do: Perform calculations

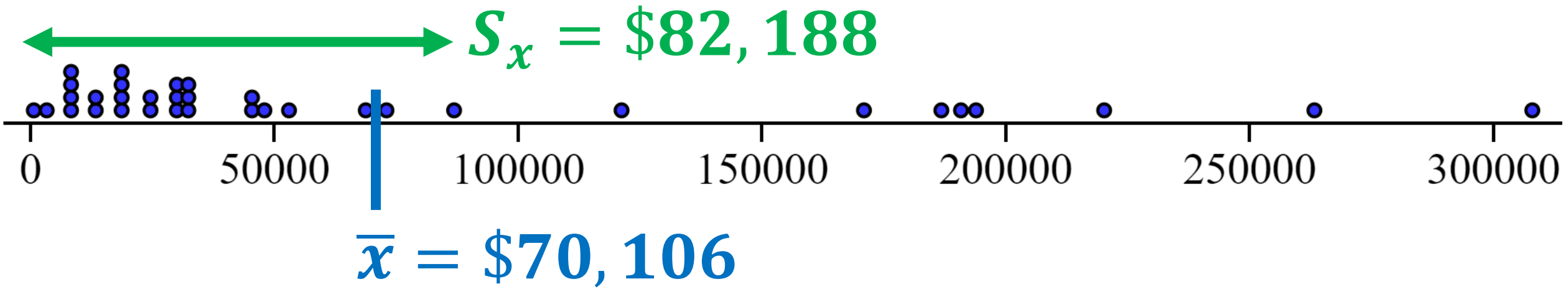
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-Plan-**Do**-Conclude

Do: Perform calculations



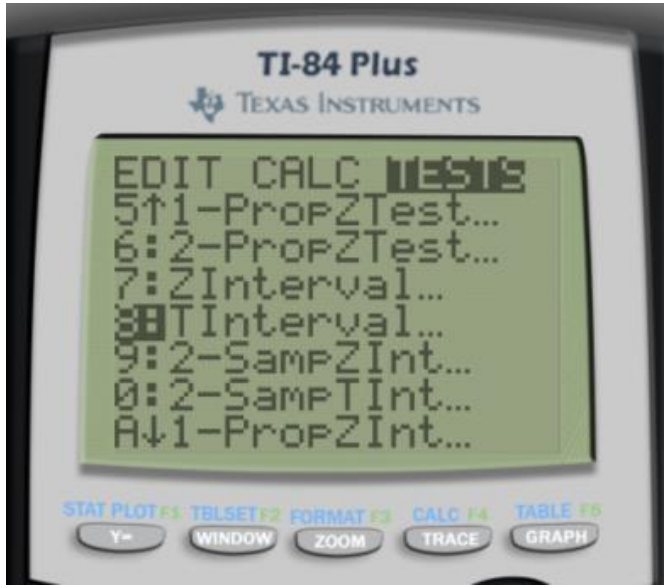
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

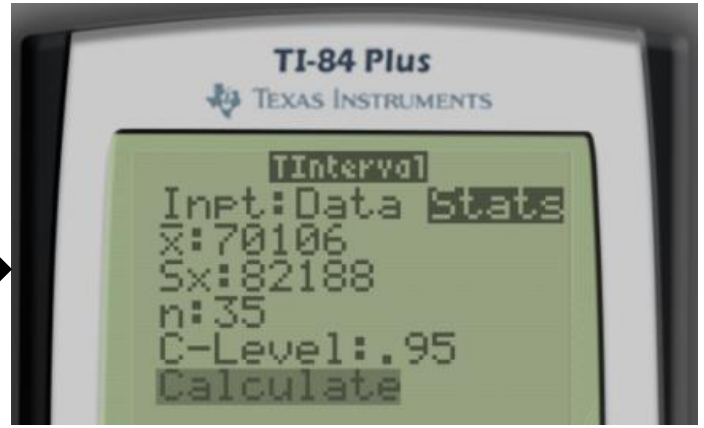
State-Plan-**Do**-Conclude

Do: Perform calculations



STAT → TESTS →

8: TInterval



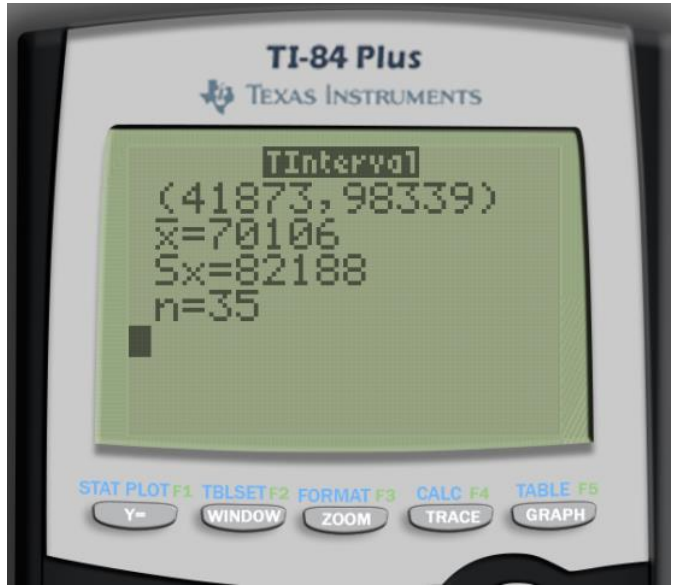
→Stats

$\bar{x}$: sample mean

S_x : sample stdev

n: sample size

Calculate



Output

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-Plan-**Do**-Conclude

Do: Perform calculations

T-interval($\bar{x}$ = 70106, s_x = 82188, n = 35, confidence = 0.95):

interval: $(41877, 98339)$

Note: This interval differs slightly from earlier interval due to rounding.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-Plan-Do-**Conclude**

Conclude: Interpret your result with context

$$(41877, 98339)$$

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

State-Plan-Do-**Conclude**

Conclude: Interpret your result with context

We are 95% confident the interval from \$41,877 to \$98,339 **captures** the **true mean** yearly income of YouTubers.

$$(41877, 98339)$$

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Lesson 8.1
Discussion

+

The **True** Mean

We can directly estimate the true mean income of YouTubers from Ad Revenue:

2019 YouTube Ad Revenue

$15,149,000,000

Sources:

- Alphabet 4th quarter earnings release: https://abc.xyz/investor/static/pdf/2019Q4_alphabet_earnings_release.pdf?cache=05bd9fe
- "How YouTube Ad Revenue Works," *Investopedia,* June 4, 2020: https://www.investopedia.com/articles/personal-finance/032615/how-youtube-ad-revenue-works.asp#:~:text=Enabling%20ads%20on%20your%20YouTube,get%20the%20remaining%2055%20percent.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

The **True** Mean

We can directly estimate the true mean income of YouTubers from Ad Revenue:

2019 YouTube Ad Revenue	**Creators Get**
$15,149,000,000	55%

Sources:

- Alphabet 4th quarter earnings release: https://abc.xyz/investor/static/pdf/2019Q4_alphabet_earnings_release.pdf?cache=05bd9fe
- "How YouTube Ad Revenue Works," *Investopedia,* June 4, 2020: https://www.investopedia.com/articles/personal-finance/032615/how-youtube-ad-revenue-works.asp#:~:text=Enabling%20ads%20on%20your%20YouTube,get%20the%20remaining%2055%20percent.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

The **True** Mean

We can directly estimate the true mean income of YouTubers from Ad Revenue:

Creator Share of Revenue

$$\$15{,}149{,}000{,}000 * 0.55 = \mathbf{\$8{,}331{,}950{,}000}$$

Sources:

- Alphabet 4th quarter earnings release: https://abc.xyz/investor/static/pdf/2019Q4_alphabet_earnings_release.pdf?cache=05bd9fe
- "How YouTube Ad Revenue Works," *Investopedia,* June 4, 2020: https://www.investopedia.com/articles/personal-finance/032615/how-youtube-ad-revenue-works.asp#:~:text=Enabling%20ads%20on%20your%20YouTube,get%20the%20remaining%2055%20percent.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

The **True** Mean

Creators Got:	**# of Creators**
$8,331,950,000	18,000,000

Sources:

- "How Many YouTube Channels Are There?", *Tubics,* https://www.tubics.com/blog/number-of-youtube-channels
- *SocialBlade*, data accessed 2019

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

The **True** Mean

Creators Got: $8,331,950,000 / **# of Creators** 18,000,000 =

Sources:

- "How Many YouTube Channels Are There?", *Tubics,* https://www.tubics.com/blog/number-of-youtube-channels
- *SocialBlade*, data accessed 2019

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

The **True** Mean

Creators Got:
$8,331,950,000

/

# of Creators
18,000,000

=

Average Yearly Pay Per Creator
$463

Sources:

- "How Many YouTube Channels Are There?", *Tubics,* https://www.tubics.com/blog/number-of-youtube-channels
- *SocialBlade*, data accessed 2019

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

The **True** Mean

Average Yearly Pay Per Creator: $463

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

The **True** Mean

Average Yearly Pay Per Creator: $463

Individual Poverty Line: $12,760

Source:

- 2020 poverty guideline from *US Dept of Health and Human Services*

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Discussion

Average Yearly Pay Per Creator: $463

Confidence Interval: $41,877 to $98,339

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Discussion

Average Yearly Pay Per Creator: $463

Confidence Interval: $41,877 to $98,339

Discussion Question: Why was our confidence interval so far off?

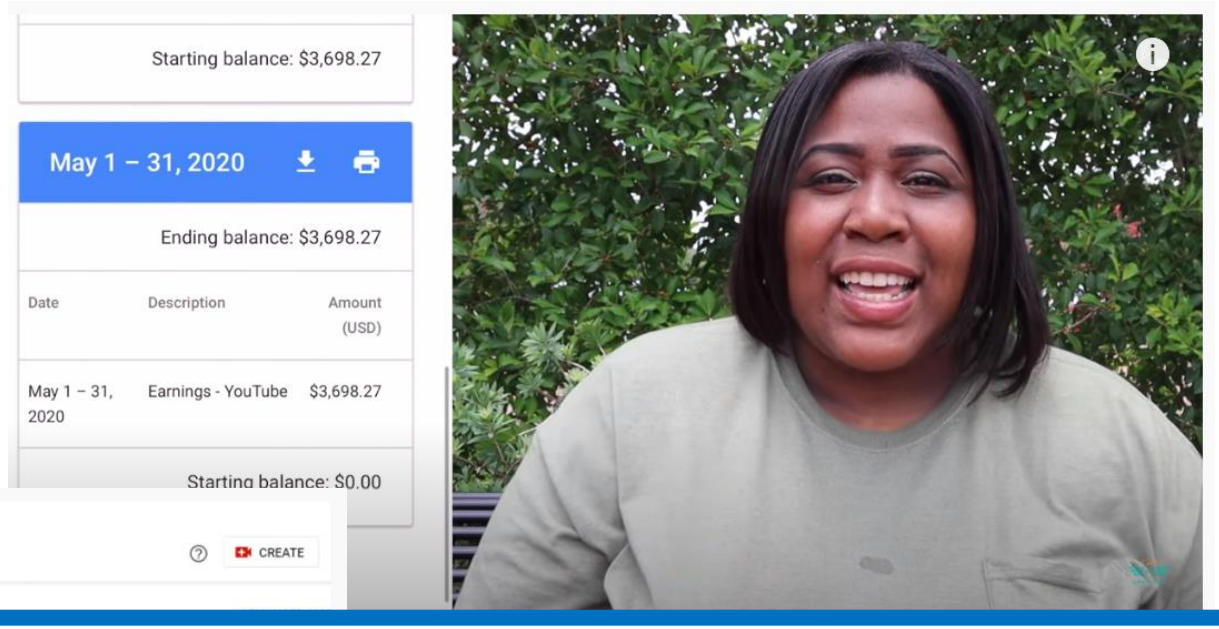
THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

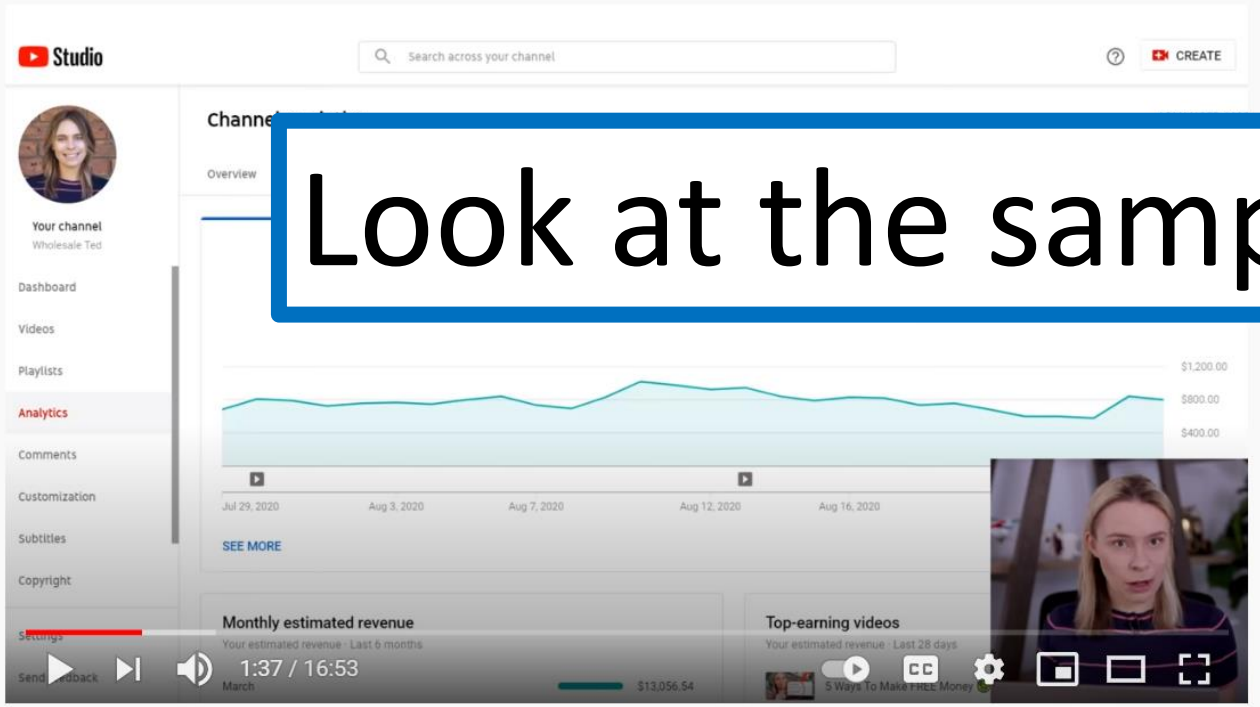
+

Skew The Script

Searched “How much I make on YouTube.”

Randomly sampled 35 from hundreds of results





How Much Did YouTube Pay Me For 1 Million Views?! (How Much Do YouTubers REALLY Earn!)

49,475 views • Oct 1, 2020 3K 44 SHARE SAVE

Look at the sampling method

Reliable data: They show their private channel revenue pages in the videos

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

+

Skew The Script

Discussion

We only **sampled among people who made videos** about their YouTube incomes.

- These people likely have **more to brag about** than those who don't make videos about their YouTube incomes.

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Discussion

Even though we sampled randomly ***among*** that group, the selection of **that group itself was biased**. So, we overestimated.

When checking the random condition, check that sampling from **whole population**.

Discussion

When you see these types YouTube videos on your “suggested” list, **don’t think** they are actually **representative** of the whole population!

Same with popular athletes and entertainers: they are not representative of all the less popular folks who wanted to be in their shoes

THE CENTER FOR SPATIAL DATA SCIENCE
THE UNIVERSITY OF CHICAGO

+

Skew The Script

Lesson 8.1
Practice

+