CH. 4 – z-scores and the Standard Normal Distribution

PY 221 **Research Methods & Statistics I**

Dr. Valenti

Tomorrow (Wednesday 2/16) – Harbert 301

- We'll be in the lab to work on Practice Lab #2, using JAMOVI
- Arrive a few minutes before 2:00 and find a computer that is turned on and connected to the internet
  - Please try to sit near the middle aisle and/or close to the front of room; this allows me to get to you more easily, in the event that you need help!
- Log in to Moodle, and open the Chapter 3 lecture slides as a reference
- Consider where you are going to save your practice lab work so that you'll have access to it later
  - E.g., USB, desktop and then upload to *OneDrive*
- I'd recommend also having a pen/pencil and your notebook for taking notes on how to use JAMOVI

This YouTube channel has a lot of helpful tutorials about JAMOVI.
https://www.youtube.com/playlist?list=PLkk92zzyru5OAtc_ItUubaSSq6S_TGfRn

Outline for Ch. 4

1. Standard Normal Distribution
2. Defining and calculating z-scores (aka *standardized scores*)
3. z-scores and the area under the curve

Histogram of Heights of Grad Students

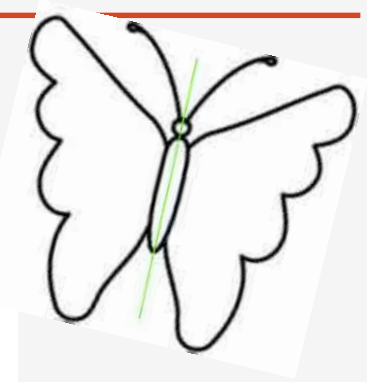
Frequency
(# Ps with each score)

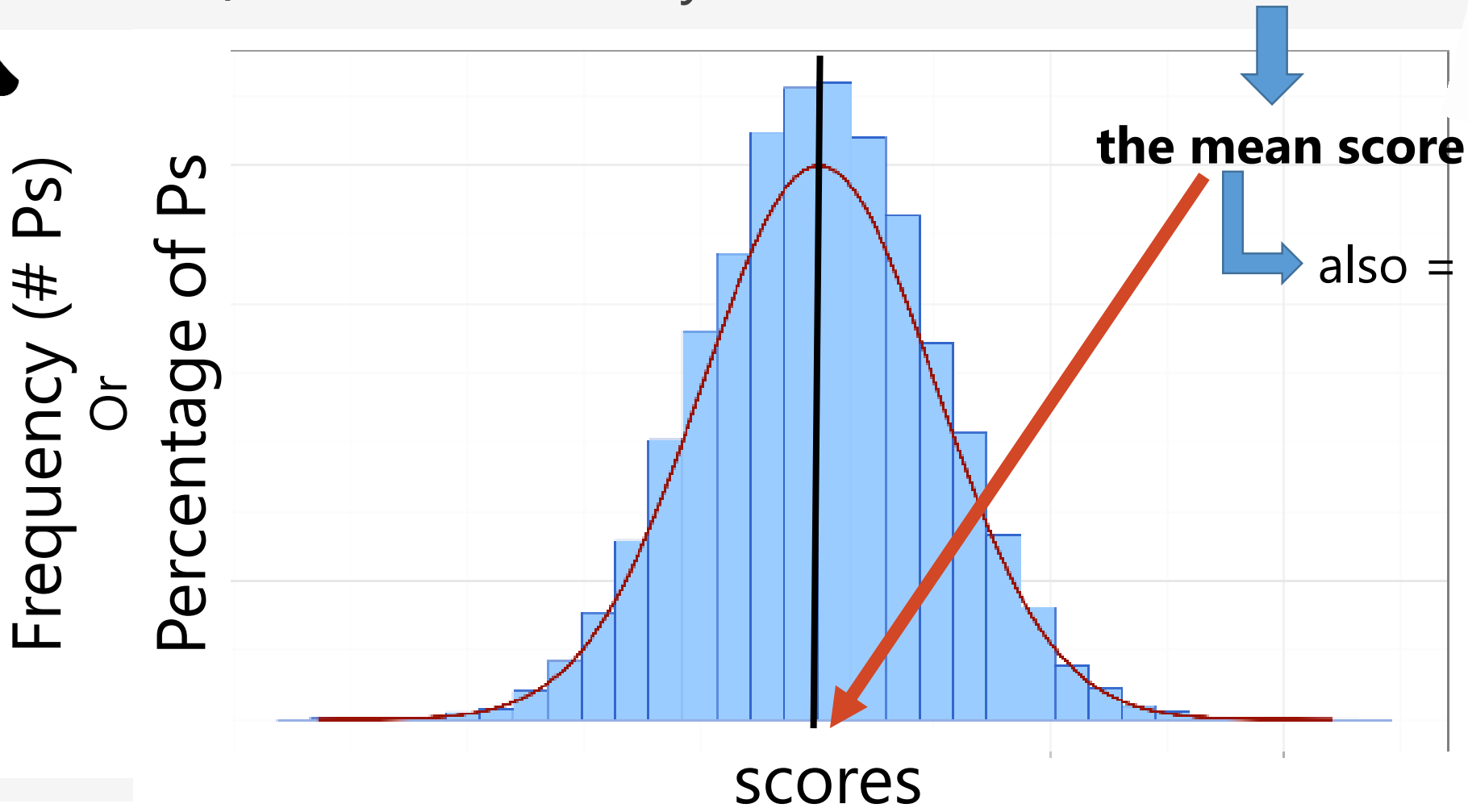
Score (which is height)

Standard Normal Distribution

What characteristics do data with this distribution have?

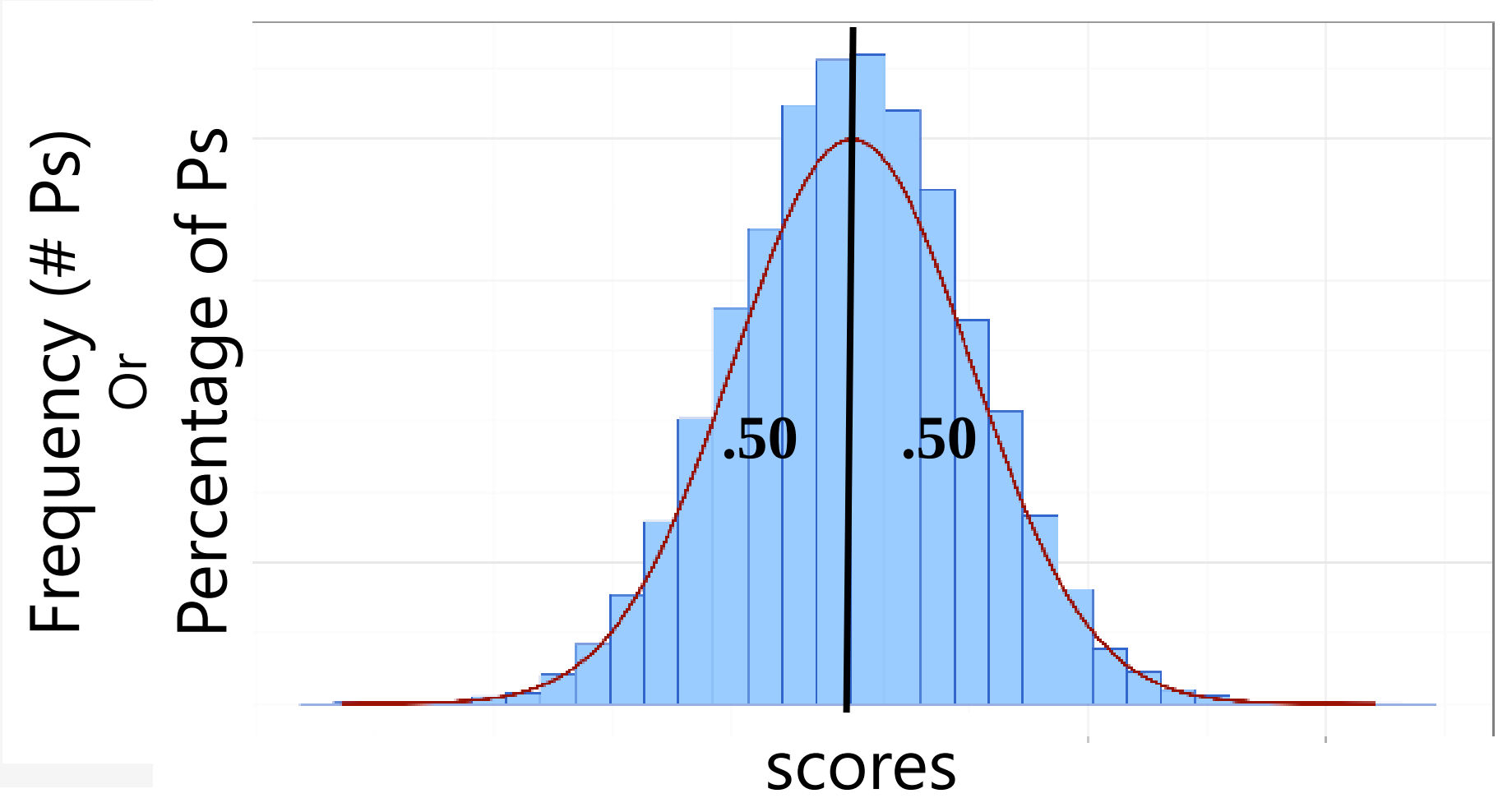
- Bell shaped-*ish*
- Symmetrical around center





Standard Normal Distribution

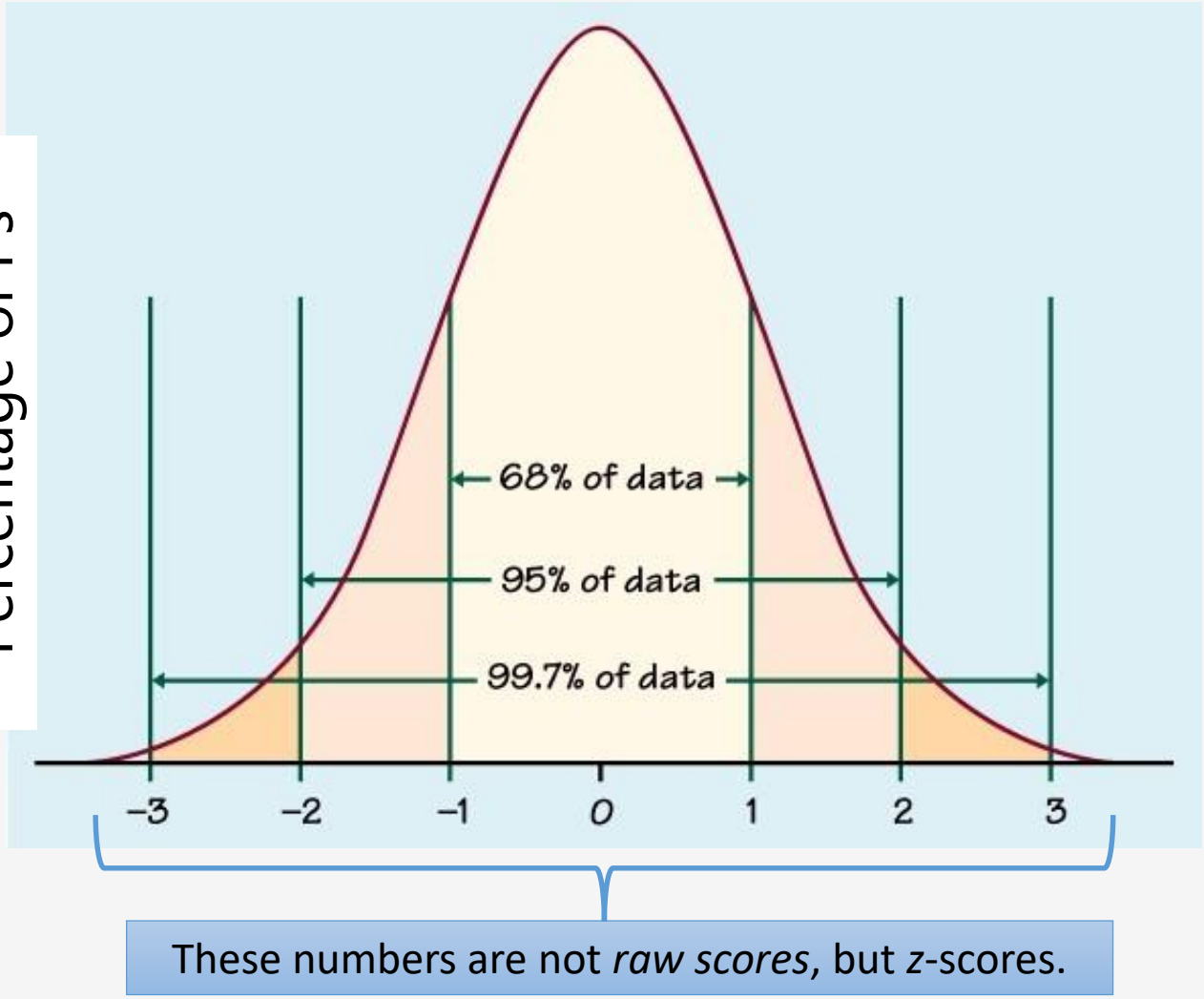
What characteristics do data with this distribution have?



Standard Normal Distribution

What characteristics do data with this distribution have?

- "area under the curve"
- **68%** of the area is (68% of people's scores are) w/in **1 standard deviation** of the mean



Outline for Ch. 4

1. Standard Normal Distribution
2. **Defining and calculating *z*-scores (aka *standardized scores*)**
3. z-scores and the area under the curve

Suppose Lisa receives scores of 76% on both her PY 101 exam and her BI 110 exam . . .

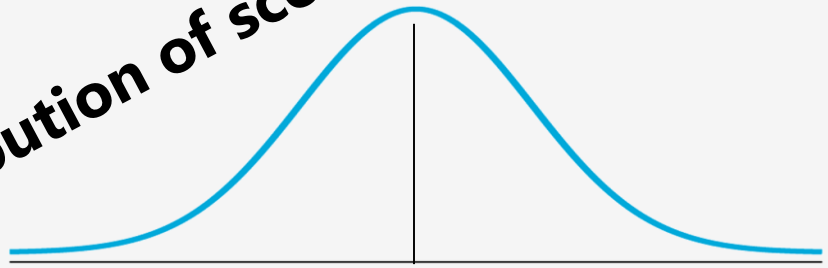
- how did Lisa do in each class?
- *and*, is she equally good at psychology and biology?

- On its own, "76%" doesn't provide all the information we need.
- **What additional information do we need?**

What additional information do we need?

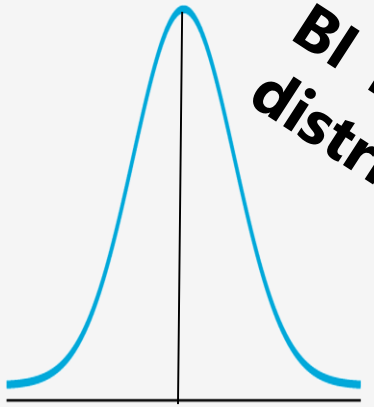
- Mean for each class → tells you whether she did better or worse than the average grade in that class
  - But what if mean for each class is the same? (70%) Is she equally good in PY 101 and BI 110?
- Standard deviation for each class →
  - tells you whether scores are really spread out from mean or close to mean . . .

PY 101
class distribution of scores



Mean = 70

BI 110 class
distribution of scores

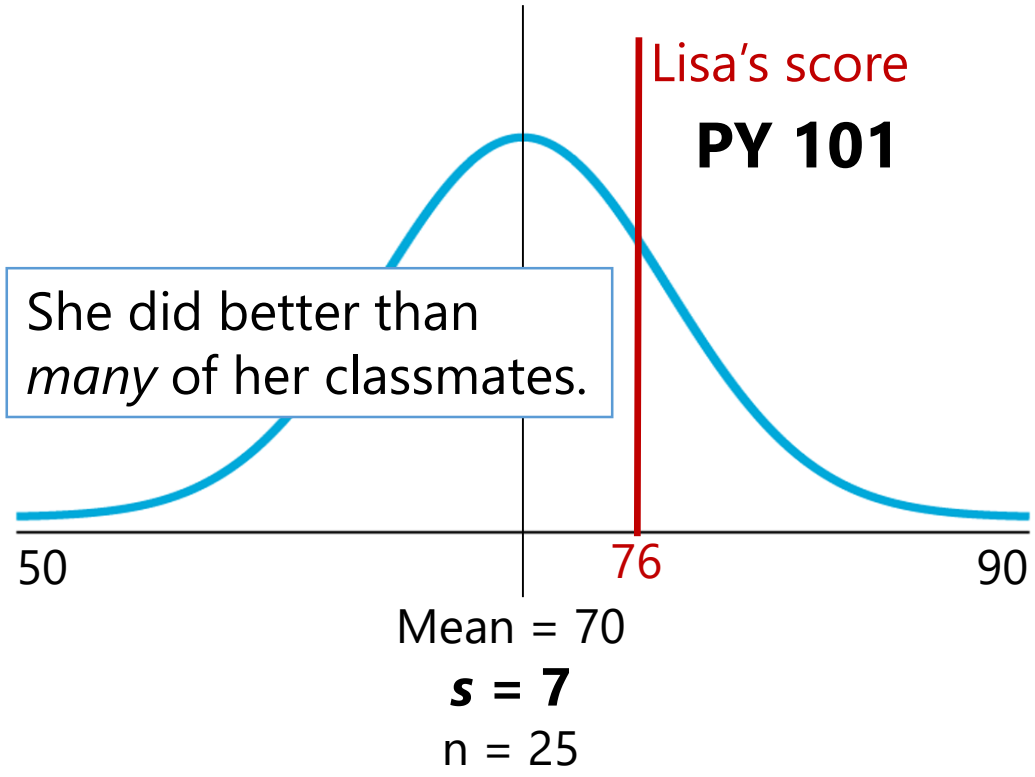


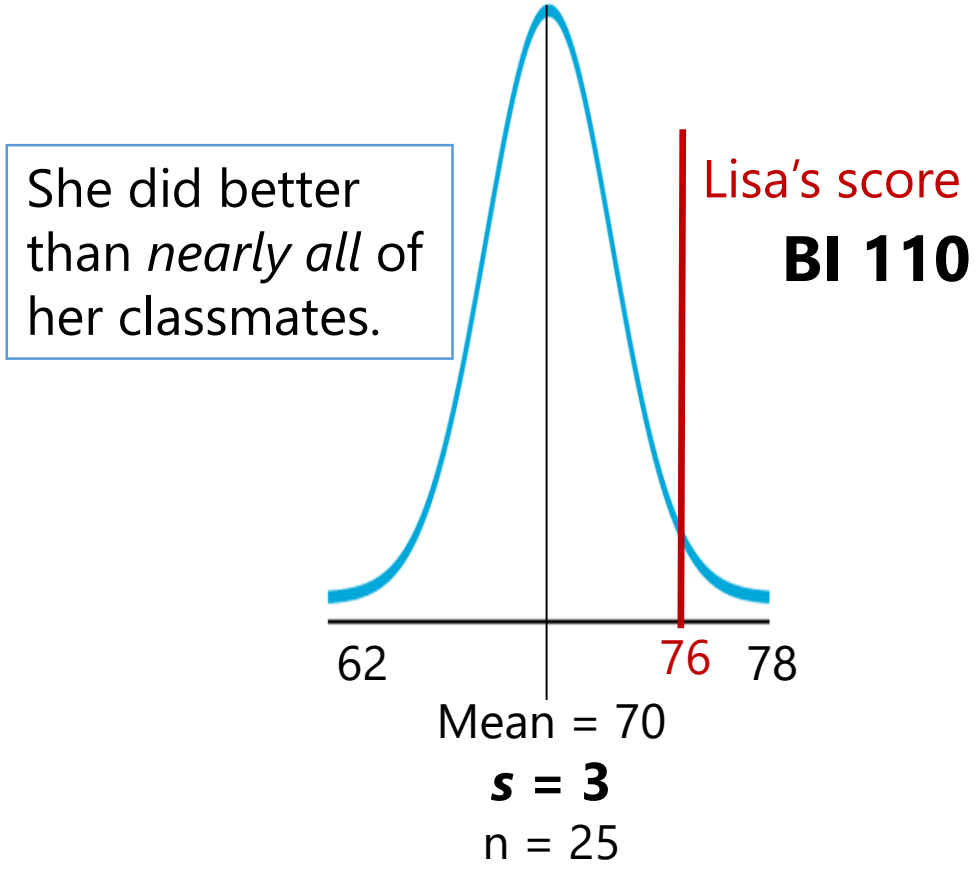
Mean = 70

It'd be helpful to know, for each class, the *location* of Lisa's score *within the distribution of the entire class' scores*

76 is Lisa's "raw score"

In which class is Lisa one of the best performers?
In which class is Lisa simply a good performer?



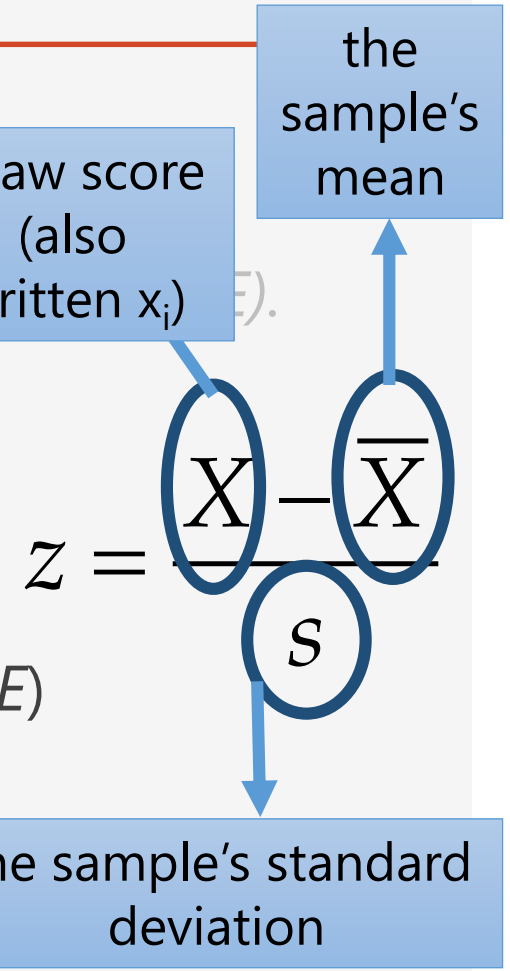


What is a *z*-score?

- Each raw score can be converted into a *z*-score, as long as you know the mean (x-bar) and standard deviation (*s*) for the sample.
- Each *z*-score tells the *exact location* of the raw score *within the distribution*.
- In particular, on its own, each *z*-score tells us…
  - whether the raw score is *above* or *below* the mean (*SIGN, +/-*)
  - how many *standard deviations from the mean* the raw score is (*MAGNITUDE*)

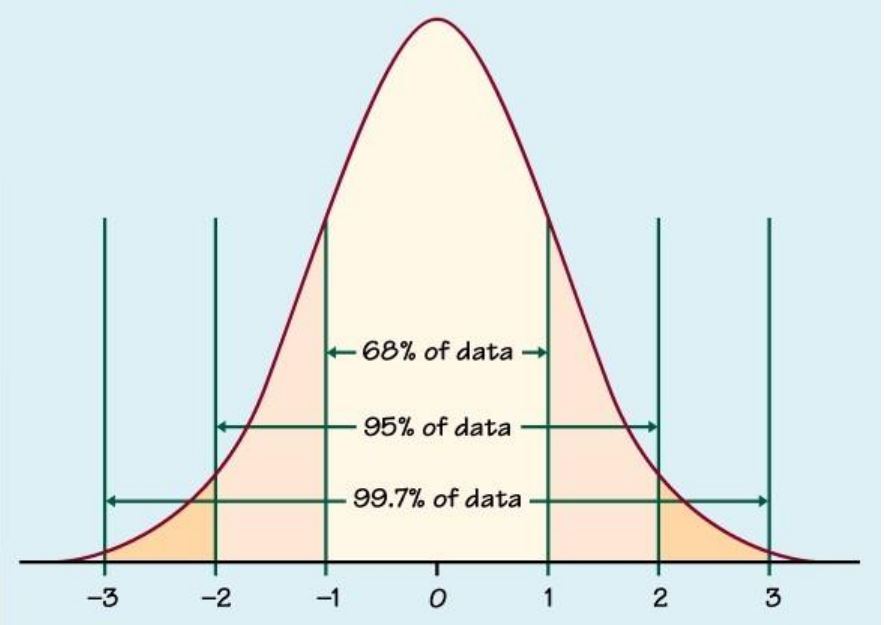
What is a *z*-score?

- In particular, each *z*-score tells us…
  - whether raw score is *above* or *below* mean (*SIGN*)
  - how many *standard deviations from mean* the raw score is (*MAGNITUDE*).
- Raw scores are converted into *z*-scores using a formula
  - *z*-scores can be positive or negative (*SIGN*)
    - *z* is *positive* when the raw score is *above (greater than)* the mean
    - *z* is *negative* when the raw score is *below (less than)* the mean
  - *z*-scores can be larger or smaller in absolute value (*MAGNITUDE*)
    - *z*-scores of *larger* magnitude indicate a raw score *far from* mean/center (i.e., closer to tails)
    - *z*-scores of *smaller* magnitude indicate a raw score *close to* mean/center (i.e., farther from tails)
  - a *z*-score of 0 indicates the raw score is *equal to* mean



What is a *z*-score?

- In particular, each *z*-score tells us…
  - whether raw score is *above* or *below* mean (*SIGN*)
  - how many *standard deviations from mean* the raw sco
- Raw scores are converted into *z*-scores using a formula
  - *z*-scores can be positive or negative (*SIGN*)
    - *z* is *positive* when the raw score is *above* the mean
    - *z* is *negative* when the raw score is *below* the mean
  - *z*-scores can be larger or smaller in absolute value (*MAGNITUDE*)
    - *z*-scores of *larger* magnitude indicate raw score *far from* mean/center (closer to tails)
    - *z*-scores of *smaller* magnitude indicate raw score *close to* mean/center (farther from tails)
  - a *z*-score of zero indicates the raw score is *equal to* mean



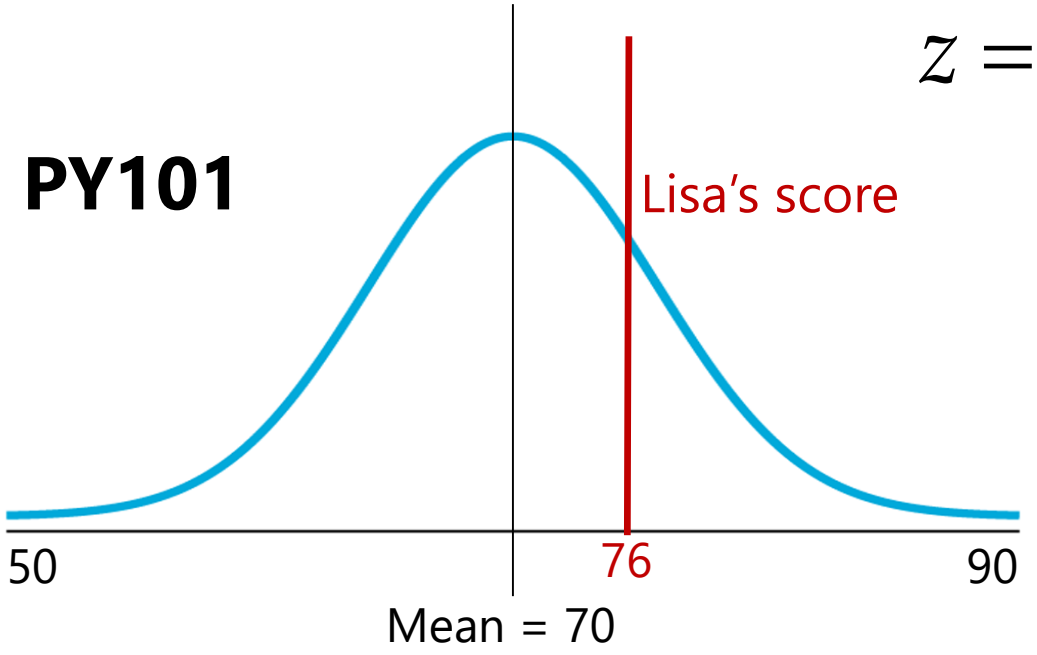
$$z = \frac{X - \overline{X}}{s}$$

Let's convert Lisa's *raw score* into a *z*-score for each class

Lisa's z-score for PY 101

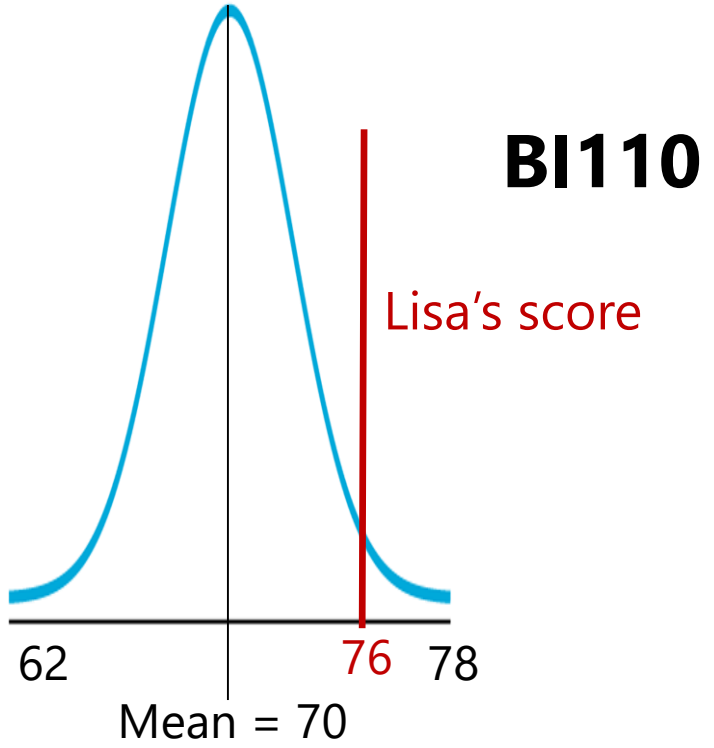
z = (76-70) / 7 = 6/7 = **+0.86**

$$z = \frac{X - \overline{X}}{s}$$



Lisa's z-score for BI 110

z = (76-70) / 3 = 6/3 = **+2.00**



Which class is Lisa doing better in, relative to the class' performance?

BI 110

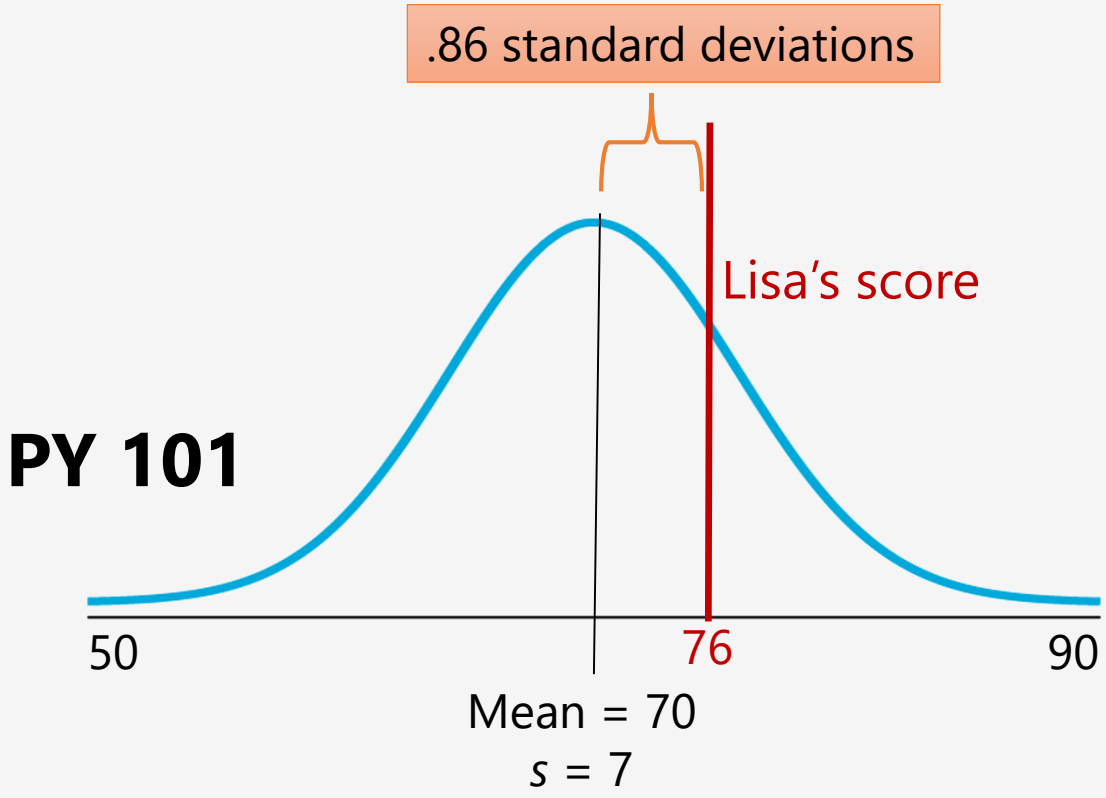
How do you interpret *z*-scores? What do they mean, in words?

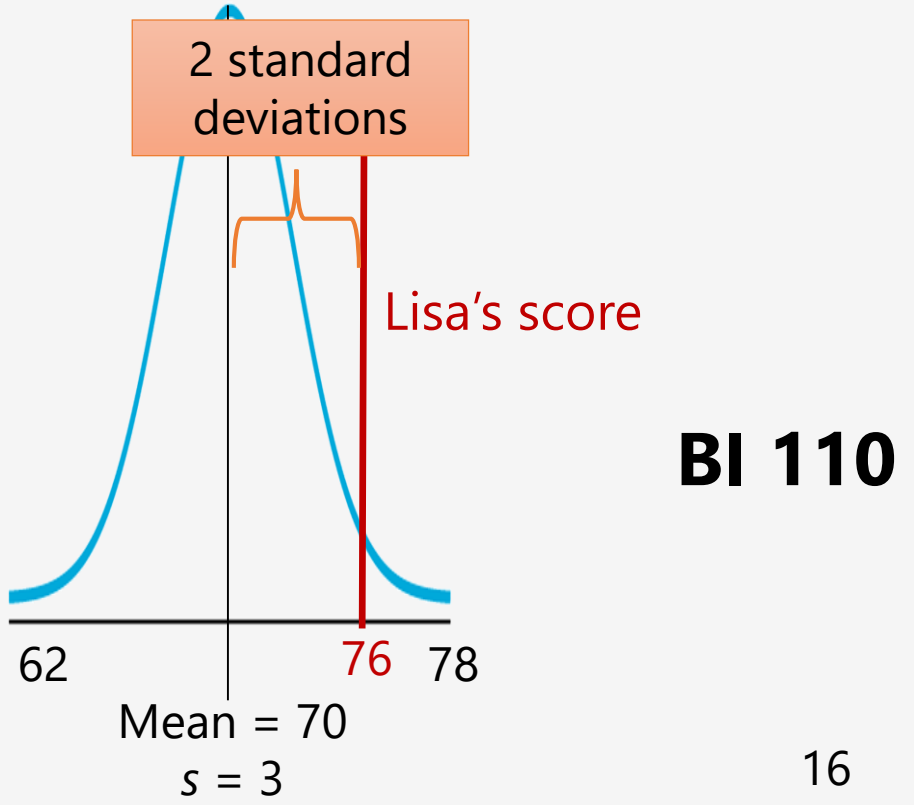
From prior slide…for PY 101, ***z* = +0.86**

- Lisa's raw score of 76 is **0.86 standard deviations above** the class mean in PY 101.

From prior slide…for BI 110, ***z* = +2.00**

- Lisa's raw score of 76 is **2.00 standard deviations above** the class mean in BI 110.





Bart earned a 65% on both exams. What are his z-scores? *Calculate for each class distribution.*

Reminder, for PY 101

Mean = 70

Standard deviation = 7

Reminder, for BI 110

Mean = 70

Standard deviation = 3

$$z = \frac{X - \overline{X}}{s}$$

Interpret Bart's z-score in words.

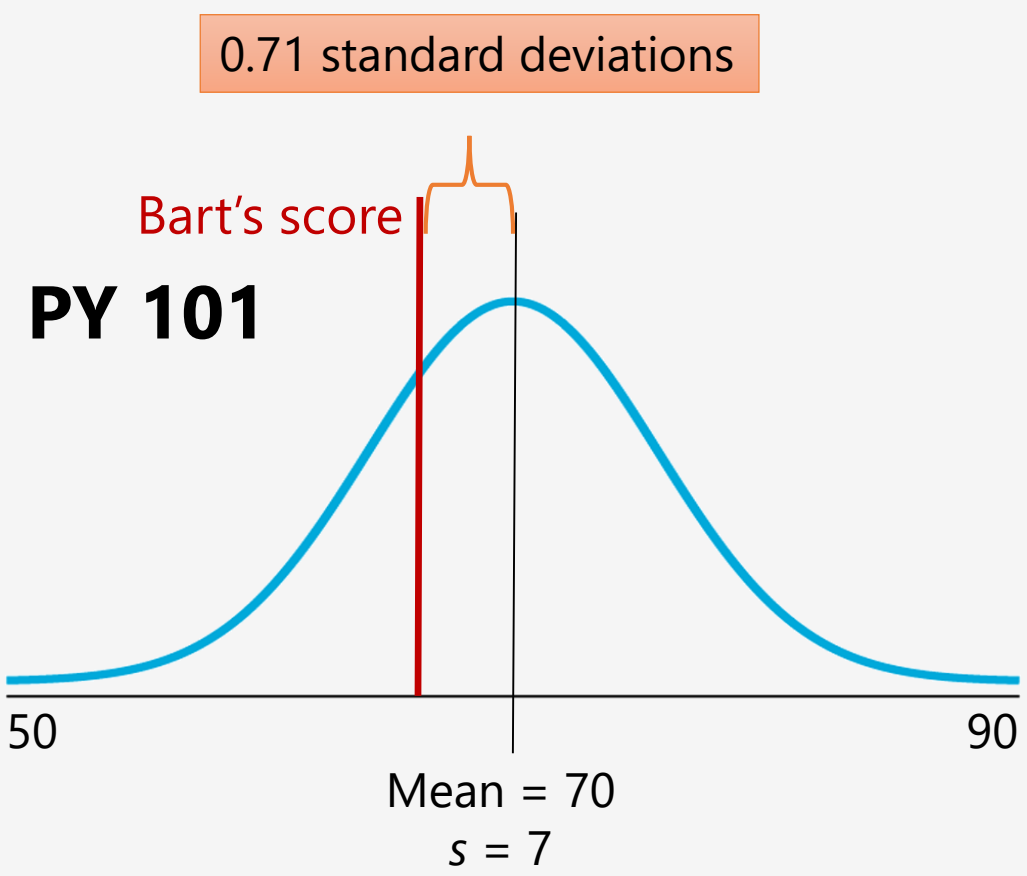
- ***z* = -0.71**
- Bart's score of 65 is **.71 standard deviations <u>below</u>** the mean PY 101 score.

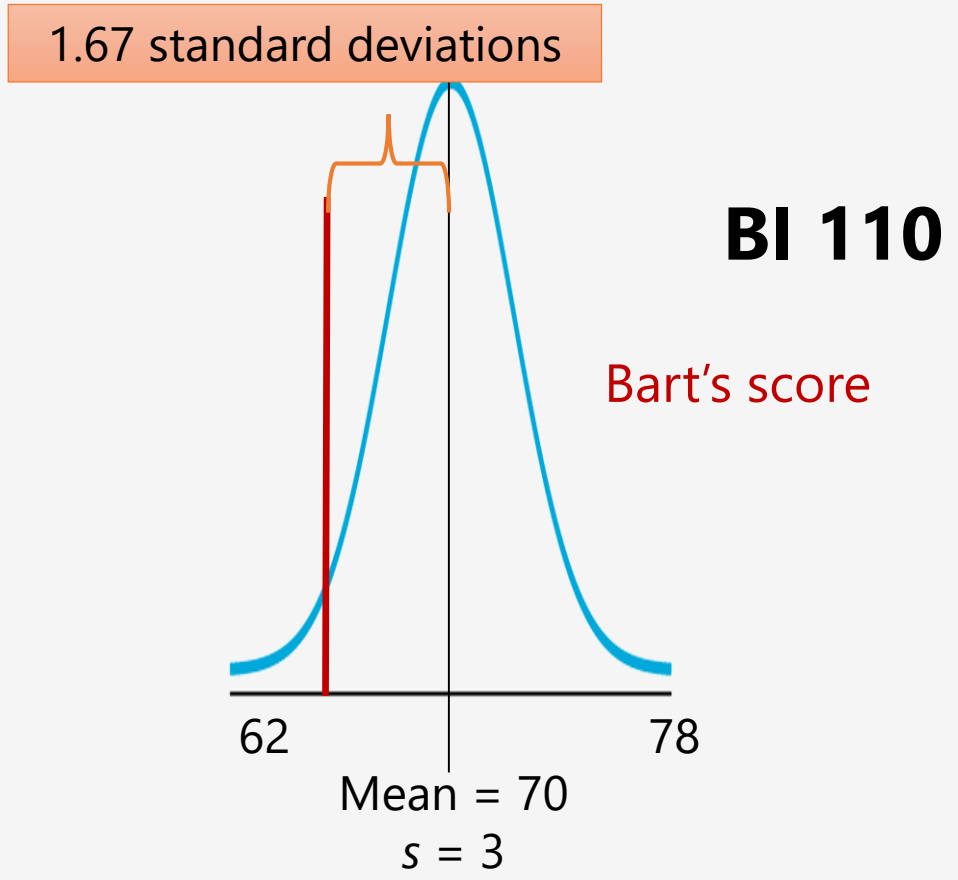
- ***z* = -1.67**
- Bart's score of 65 is **1.67 standard deviations <u>below</u>** the mean BI 110 score.

Which class is Bart doing better in, relative to the class' performance?

PY 101

Bart earned a 65% on both exams.





What are the advantages of *z*-scores, over and above *raw scores*?

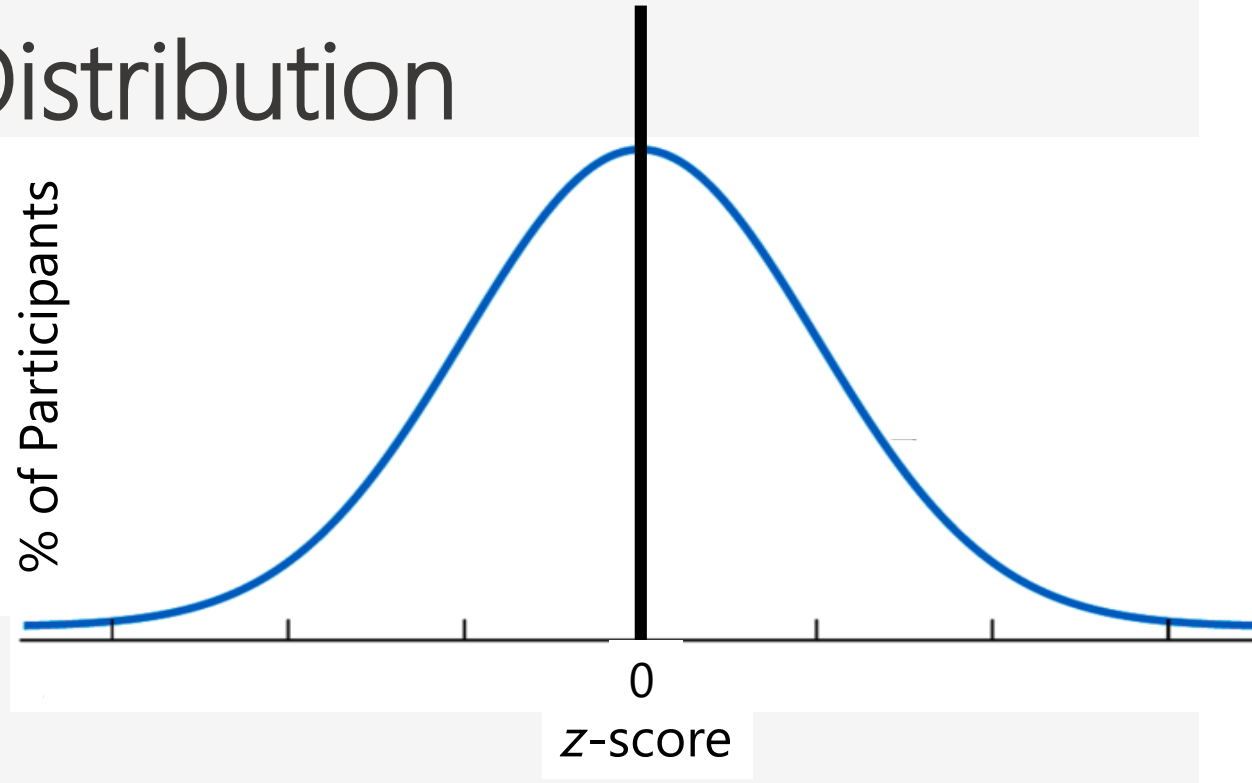
1. *Z*-scores make it possible to compare scores that come from different distributions.
   - Our original example: Lisa's *z*-score in **PY 101** & her *z*-score in **BI 110**
   - Another example: Lisa's *z*-score in **MU 100 Section *A*** & Allison's *z*-score in **MU 100 *B***

2. *Z*-scores allow us to determine the **percentage** of raw scores that fall above or below a given raw score.
   - EX: what % of PY 101 classmates performed better than and worse than Lisa, who scored a 76 on the exam?

… the topic of Chapter 5

The Standard Normal (*z*) Distribution

- If you converted ALL raw scores in your normally-distributed sample to z-scores, & created a frequency distribution (histogram), it would be called the *z* distribution, which always has a mean of zero.



If your z-score = 0, what % of scores in the sample are *above* your score?

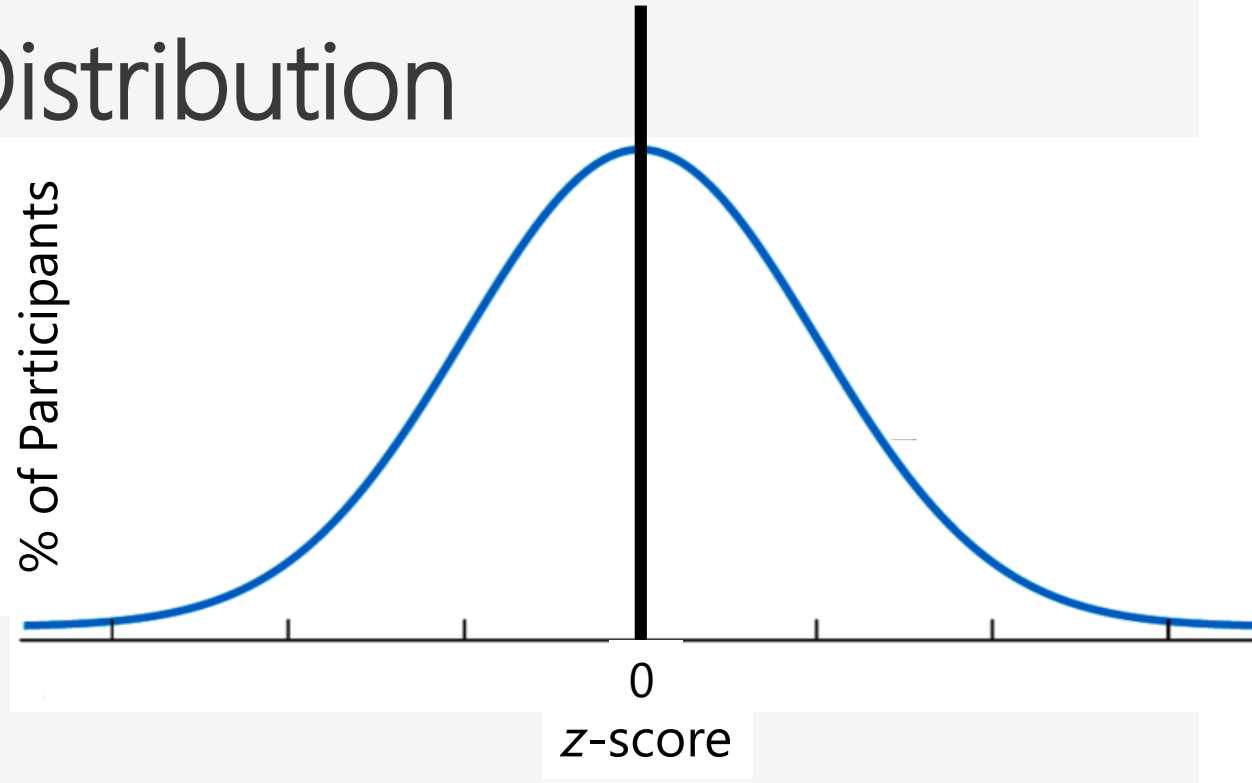
- **~50%**

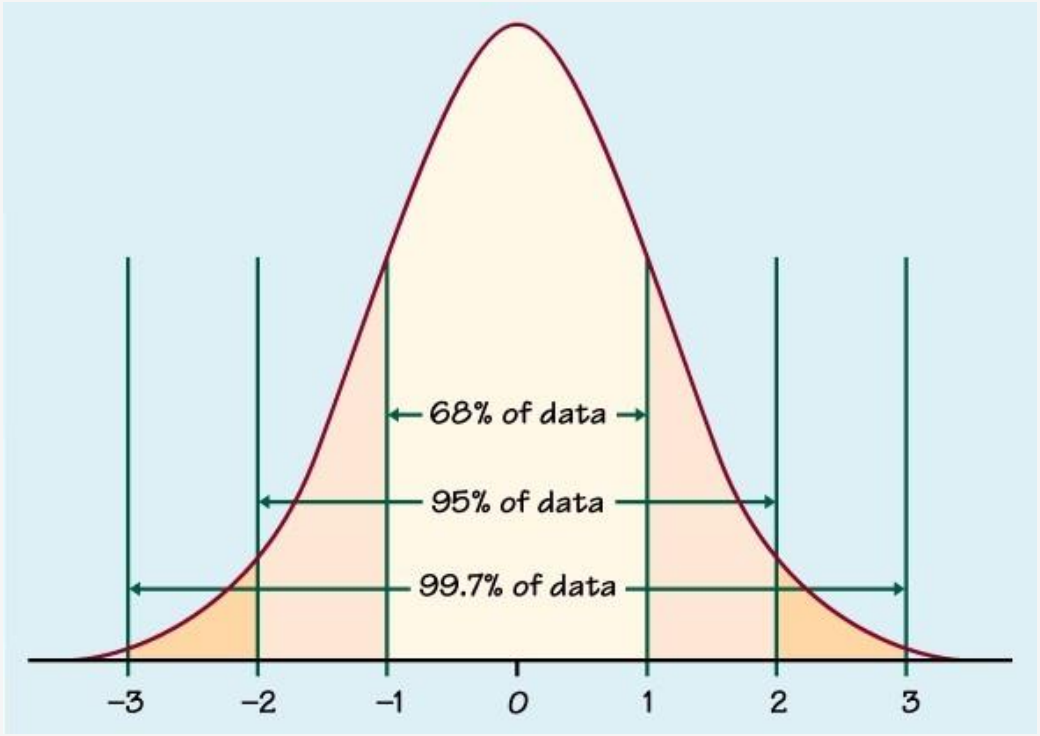
If your z-score = 0, what % of scores in the sample fall *below* your score?

- **~50%**

The Standard Normal (*z*) Distribution

- If you converted ALL raw scores in your normally-distributed sample to z-scores, & created a frequency distribution, it would be called the *z* distribution, which always has a mean of zero.





Check your understanding of the z distribution

1. If your z-score = 0, what % of scores in the sample fall *above* your raw score?
2. In any z-distribution, what % of all scores have a z-score of >+2 or < –2?
3. If your z-score for a class exam is +.75, does your raw exam score fall above or below the class mean?
4. If your raw exam score equals the class mean, what is your z-score?
5. If your z-score is –2.98, are you closer to the center of the distribution or tails?
6. If you randomly select a z-score from all possible z-scores in our sample, what's the probability that the chosen score will be between -2 and +2?
7. If the mean height of all 5^{th} graders is 50 inches, with a standard deviation of 5 inches and assuming the distribution is normal, what % of 5^{th} graders have a height between 45 inches and 55 inches?
8. In a given sample, Person A has a z-score of –1.75 and Person B has a z-score of 2.30. Whose raw score is closer to the mean of the sample?
9. In a given sample, Person A has a z-score of +1 and Person B has a z-score of -1. What % of participants in the sample have scores between Person A and B?

Check your understanding of the z distribution

50% 1. If your z-score = 0, what % of scores in the sample are *above* your raw score?

5% 2. In any z-distribution, what % of all scores have a z-score of >+2 or < –2?

above 3. If your z-score for a class exam is +.75, does your raw exam score fall above or below the class mean?

z = 0 4. If your raw exam score equals the class mean, what is your z-score?

tail 5. If your z-score is –2.98, are you closer to the center of the distribution or tails?

95% 6. If you randomly select a z-score from all possible z-scores in our sample, what's the probability that the chosen score will be between -2 and +2?

68% 7. If the mean height of all 5th graders is 50 inches, with a standard deviation of 5 inches and assuming the distribution is normal, what % of 5th graders have a height between 45 inches and 55 inches?

A 8. In a given sample, Person A has a z-score of –1.75 and Person B has a z-score of 2.30. Whose raw score is closer to the mean of the sample?

68% 9. In a given sample, Person A has a z-score of +1 and Person B has a z-score of -1. What % of participants in the sample have scores between Person A and B?

“Simple Learning Pro” YouTube channel

Playlist for ***statistics***

https://www.youtube.com/playlist?list=PL0KQuRyPJoe6KjlUM6iNYgt8d0DwI-IGR

I also put this link on Moodle. Check it out!