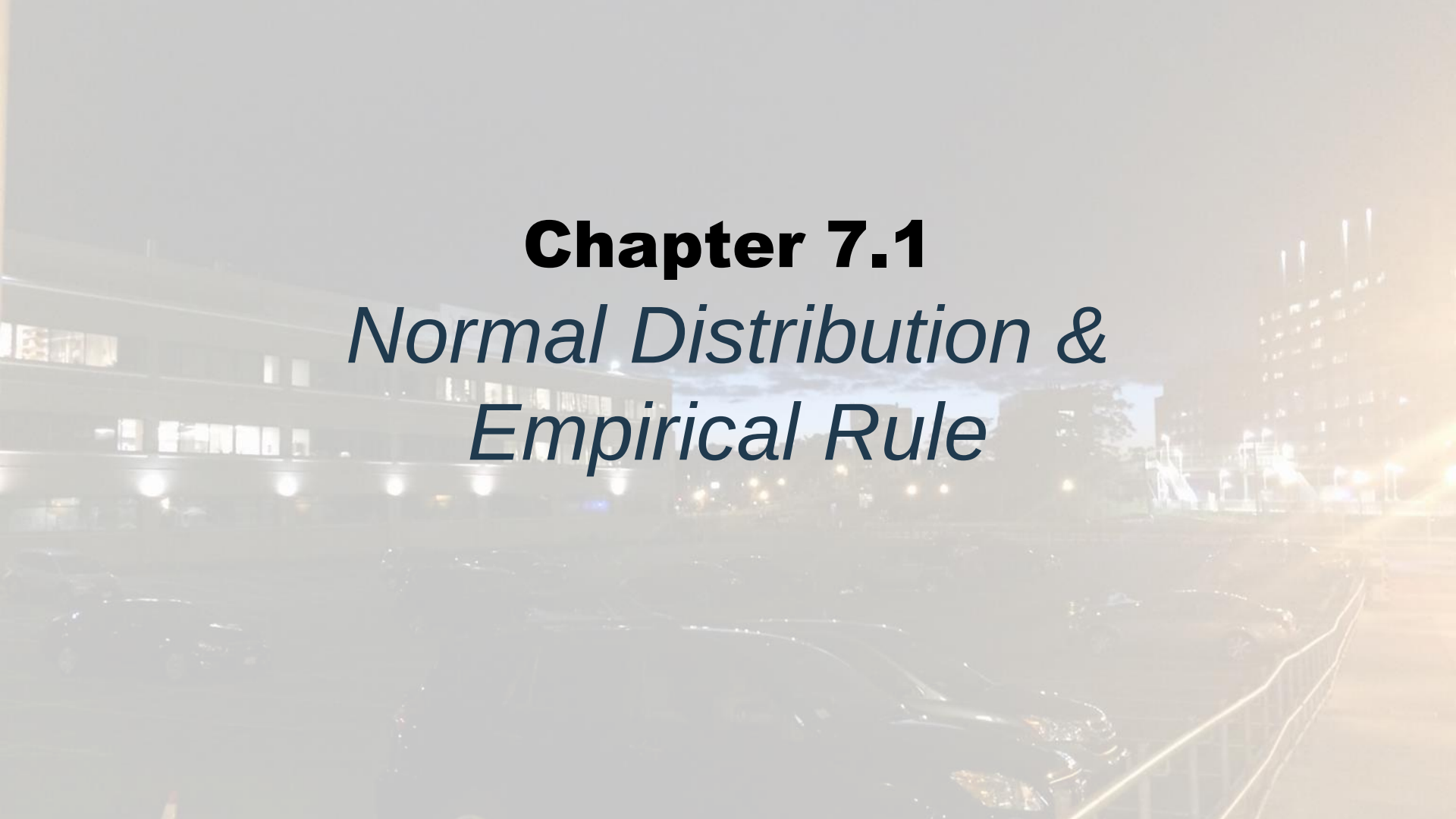


Chapter 7.1

Normal Distribution & Empirical Rule

Learning Objectives

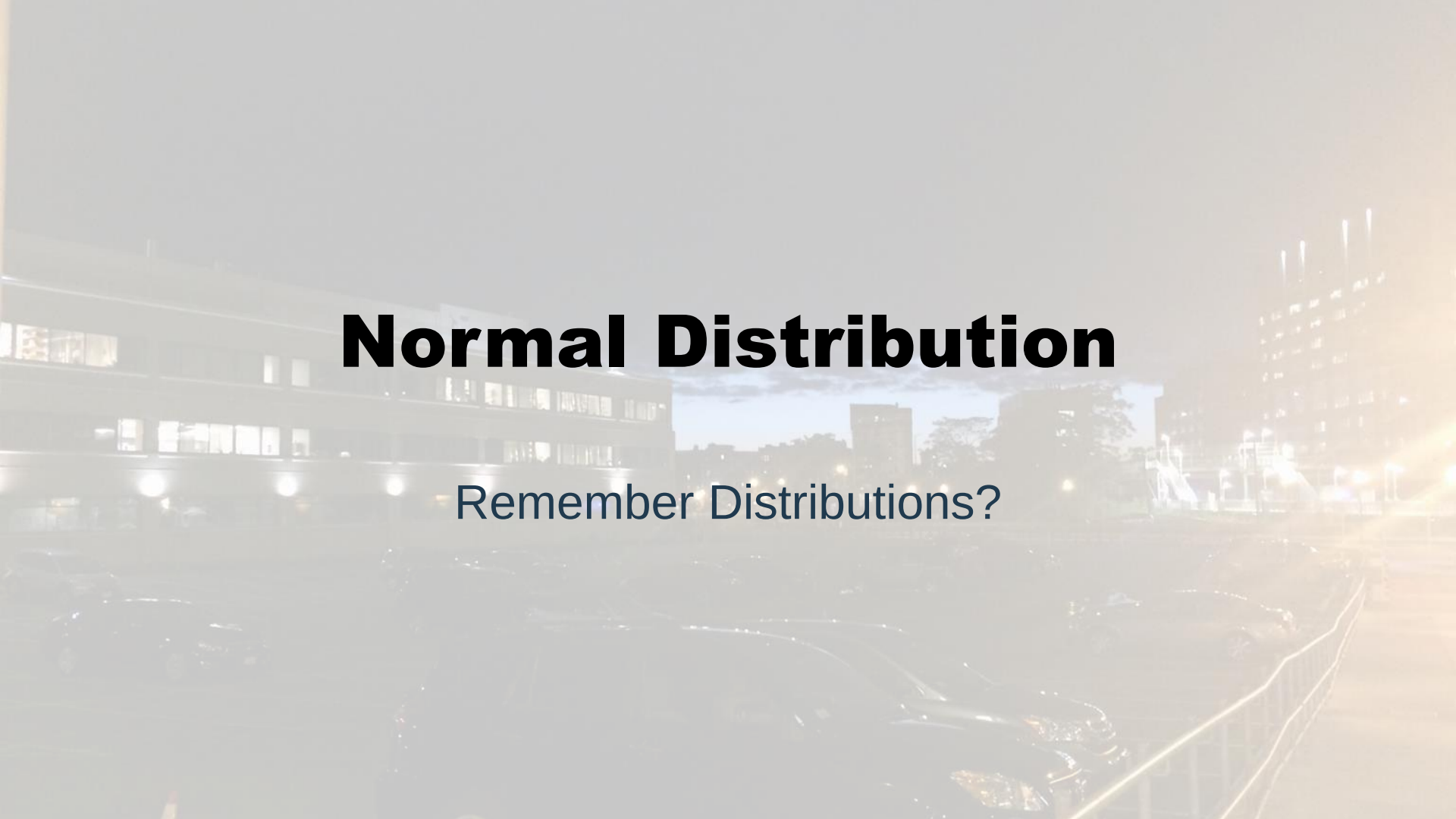
At the end of this lecture, the student should be able to:

- State two properties of the normal curve.
- State two differences between Chebyshev Intervals and the Empirical Rule
- Explain how to apply the Empirical Rule to a normal distribution

Introduction

- Remembering distributions
- Properties of the normal curve
- Remembering Chebyshev Intervals
- Empirical Rule
- Applying the Empirical Rule

Photograph by Mightyhansa



Normal Distribution

Remember Distributions?

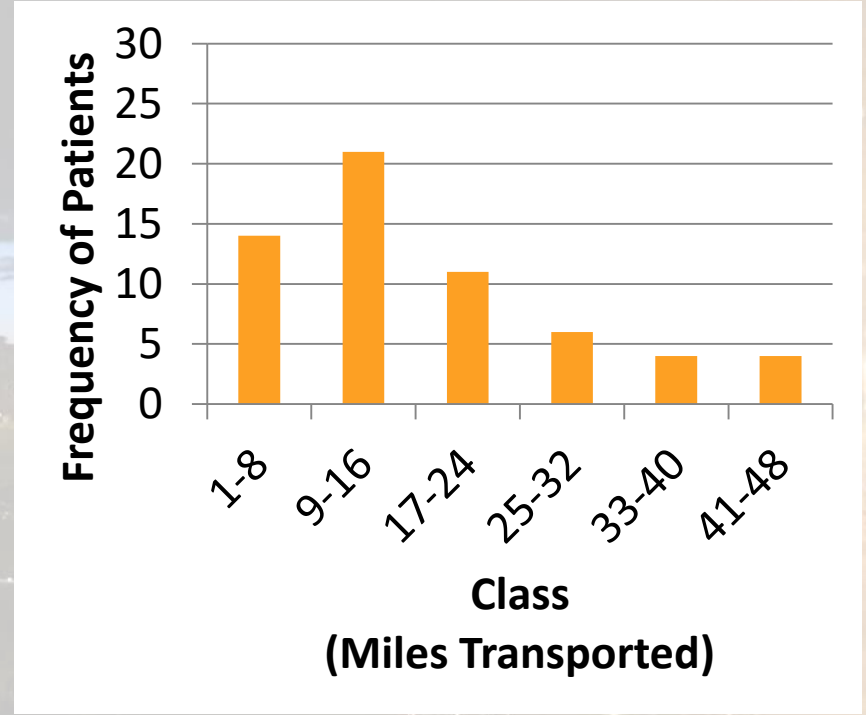
Remember Distributions?

- Using quantitative variable
- Classes determined
- Frequency table made

Class Limits	Freq-uency	Relative Freq-uency
1-8 miles	14	0.23
9-16 miles	21	0.35
17-24 miles	11	0.18
25-32 miles	6	0.10
33-40 miles	4	0.07
41-48 miles	4	0.07
Total	60	1.00

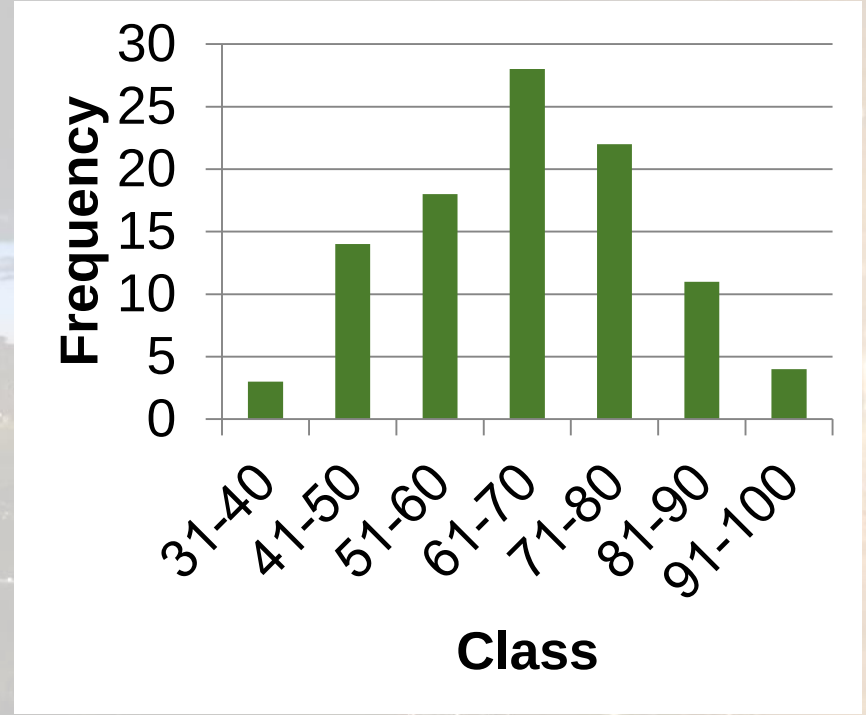
Remember Distributions?

- Using quantitative variable
- Classes determined
- Frequency table made
- Frequency histogram
- Then we could see the distribution



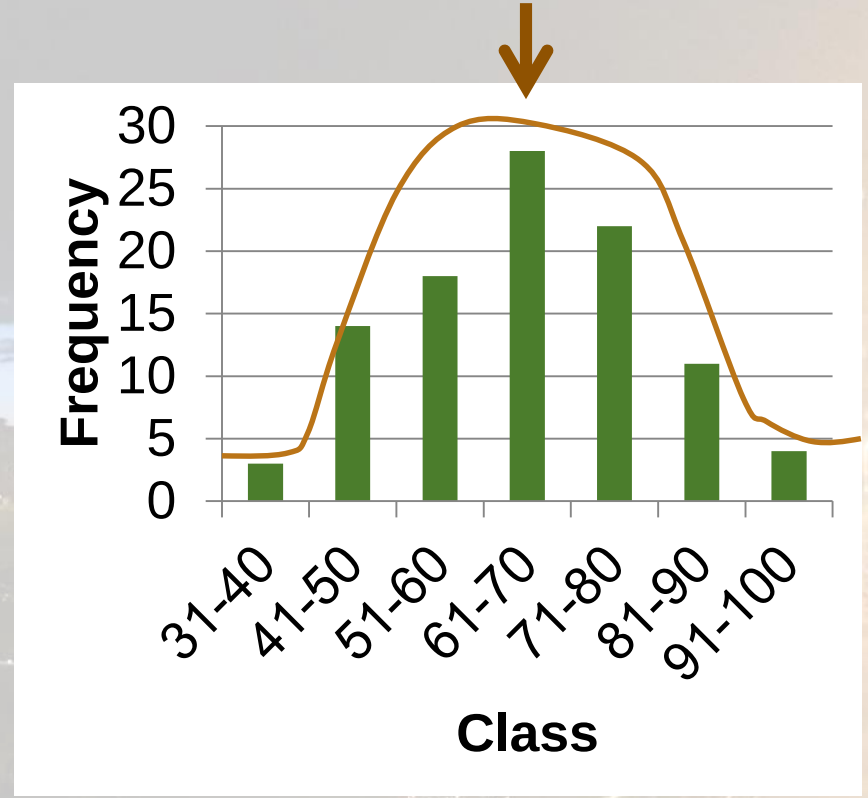
Remember the Normal Distribution?

- Imagine a large class (n=100) takes a very difficult test
- The test is worth 100 points, but it's so hard, no one actually gets 100 points
- Instead, the mode is near a C grade



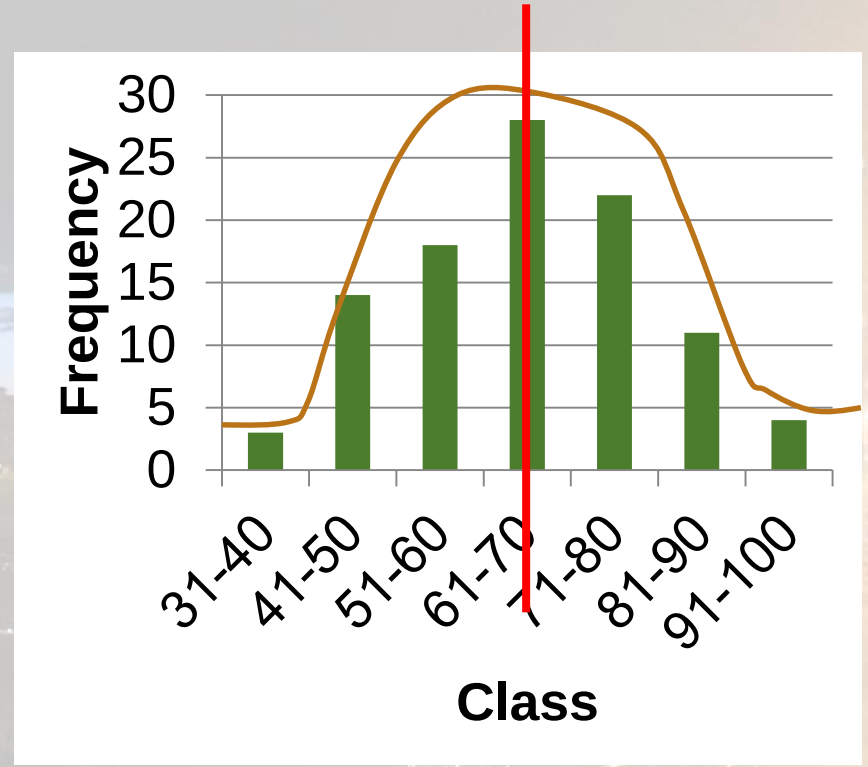
Properties of the Normal Curve

1. The curve is bell-shaped, with the highest point over the mean.



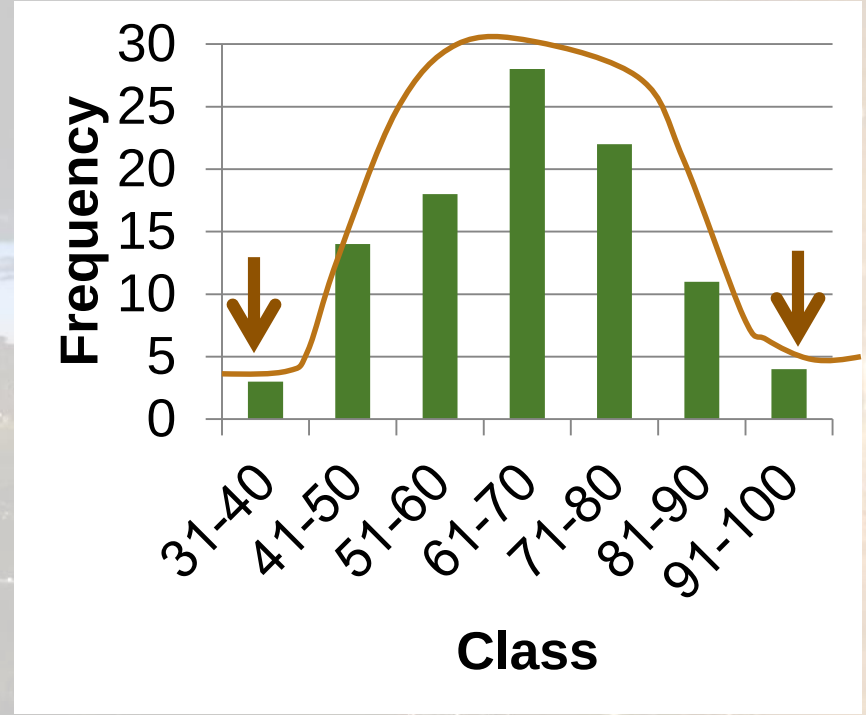
Properties of the Normal Curve

1. The curve is bell-shaped, with the highest point over the mean.
2. The curve is symmetrical around a vertical line through the mean.



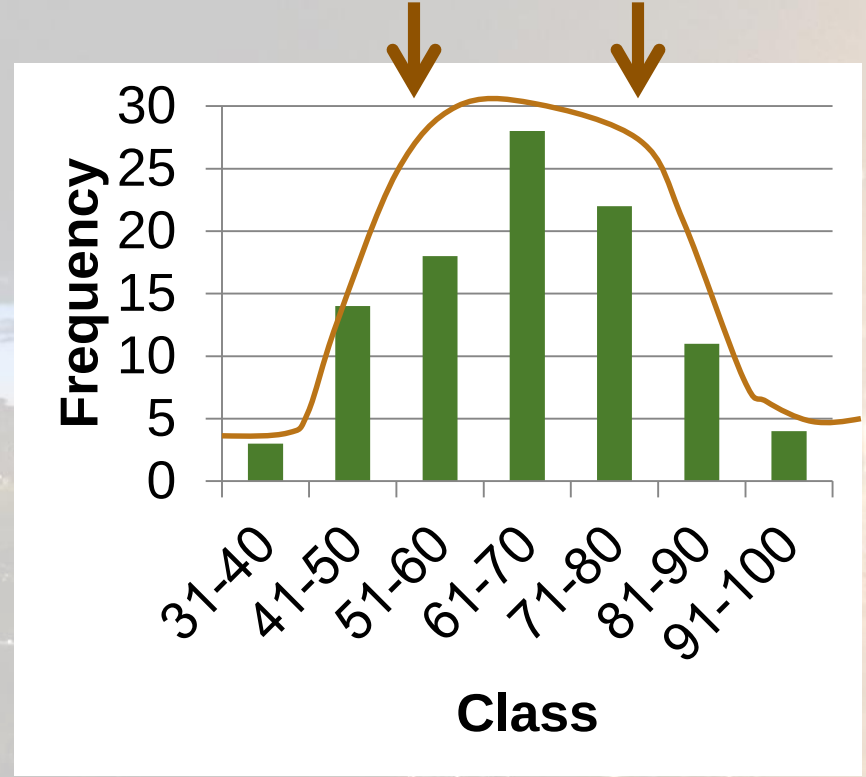
Properties of the Normal Curve

1. The curve is bell-shaped, with the highest point over the mean.
2. The curve is symmetrical around a vertical line through the mean.
3. The curve approaches the horizontal axis but never touches or crosses it.



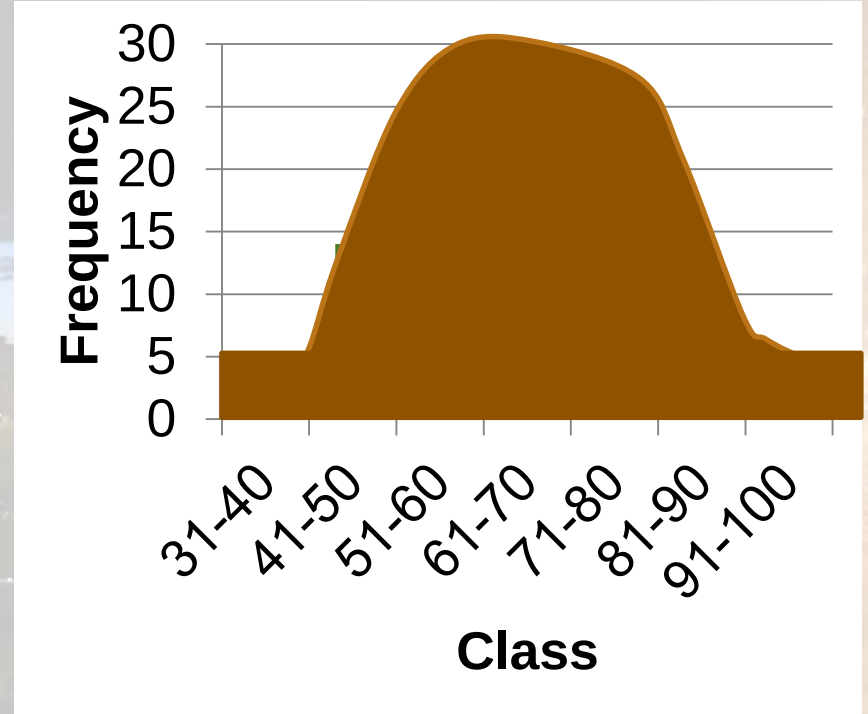
Properties of the Normal Curve

1. The curve is bell-shaped, with the highest point over the mean.
2. The curve is symmetrical around a vertical line through the mean.
3. The curve approaches the horizontal axis but never touches or crosses it.
4. The inflection (transition) points between cupping upward and downward occur at about mean +/- 1 sd



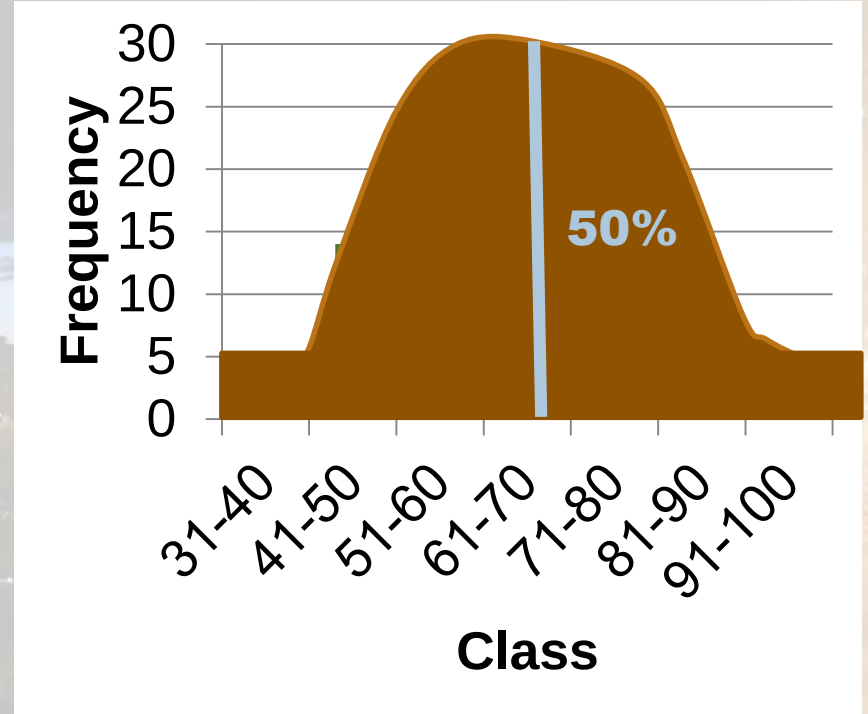
Properties of the Normal Curve

1. The curve is bell-shaped, with the highest point over the mean.
2. The curve is symmetrical around a vertical line through the mean.
3. The curve approaches the horizontal axis but never touches or crosses it.
4. The inflection (transition) points between cupping upward and downward occur at about mean +/- 1 sd
5. The area under the entire curve is 1 (think: 100%).



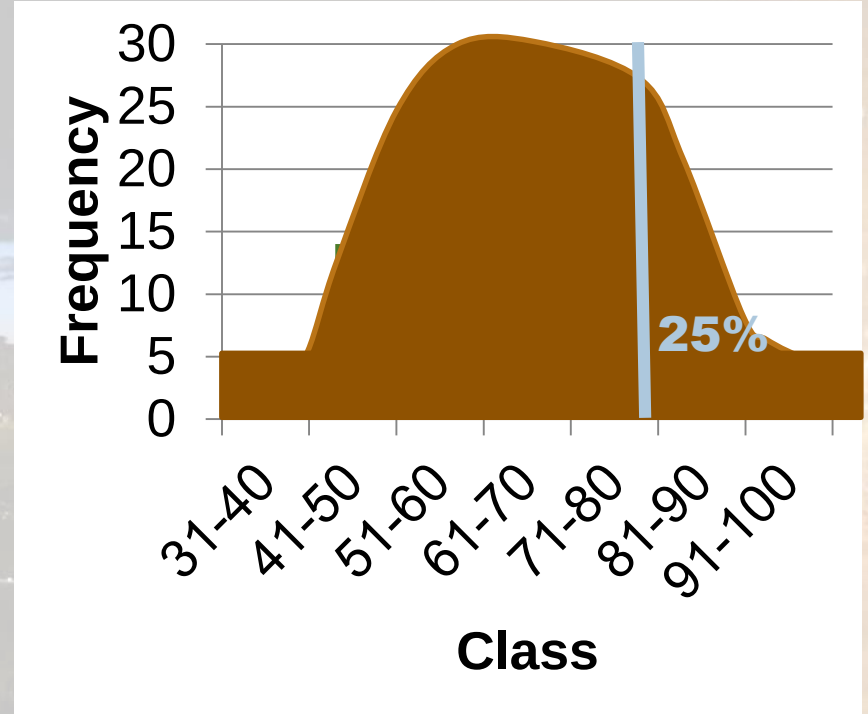
Properties of the Normal Curve

1. The curve is bell-shaped, with the highest point over the mean.
2. The curve is symmetrical around a vertical line through the mean.
3. The curve approaches the horizontal axis but never touches or crosses it.
4. The inflection (transition) points between cupping upward and downward occur at about mean +/- 1 sd
5. The area under the entire curve is 1 (think: 100%).



Properties of the Normal Curve

1. The curve is bell-shaped, with the highest point over the mean.
2. The curve is symmetrical around a vertical line through the mean.
3. The curve approaches the horizontal axis but never touches or crosses it.
4. The inflection (transition) points between cupping upward and downward occur at about mean +/- 1 sd
5. The area under the entire curve is 1 (think: 100%).



Empirical Rule

Remember Chebyshev?

Remember Chebyshev?

- Intervals have boundaries, or limits: lower limit and upper limit.
- Remember Chebyshev Intervals?
  - They say, “At least ____% of the data fall in the interval.”
  - When lower limit was μ-2σ, and upper limit was μ+2σ, at least 75% of the data were in the interval.
- Imagine n=100 students, μ score on test 65.5, σ = 14.5
  - Lower limit: 65.5 – (2*14.5) = 36.5
  - Upper limit: 65.5 + (2*14.5) = 94.5
  - So if you had 100 data points, *at least* 75 would be between 36.5 and 94.5.

Chebyshev vs. Empirical Rule

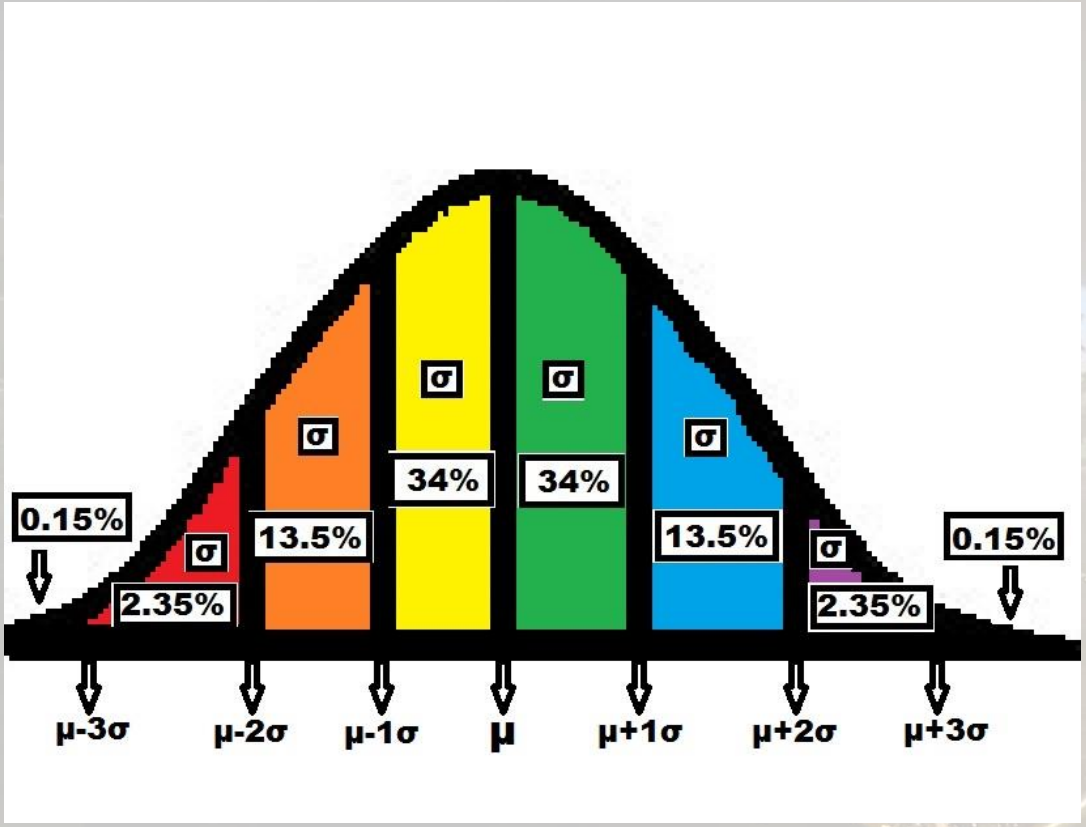
Chebyshev's Theorem

1. Applies to any distribution
2. Says "at least"
   - between μ +/- 2σ, there are *AT LEAST* 75% of the data
   - between μ +/- 3σ is at least 88.9%
   - between μ +/- 4σ is at least 93.8%

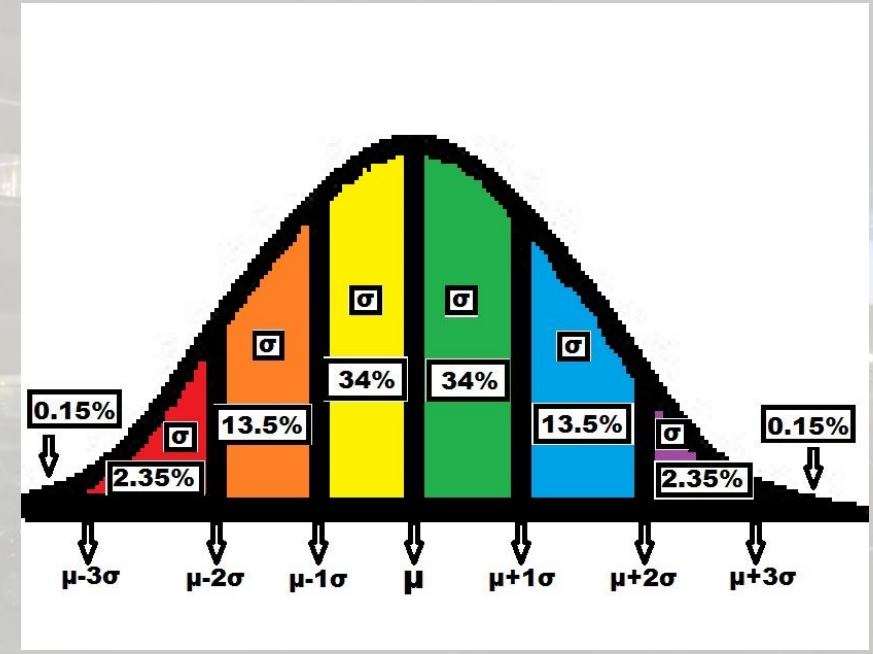
Empirical Rule

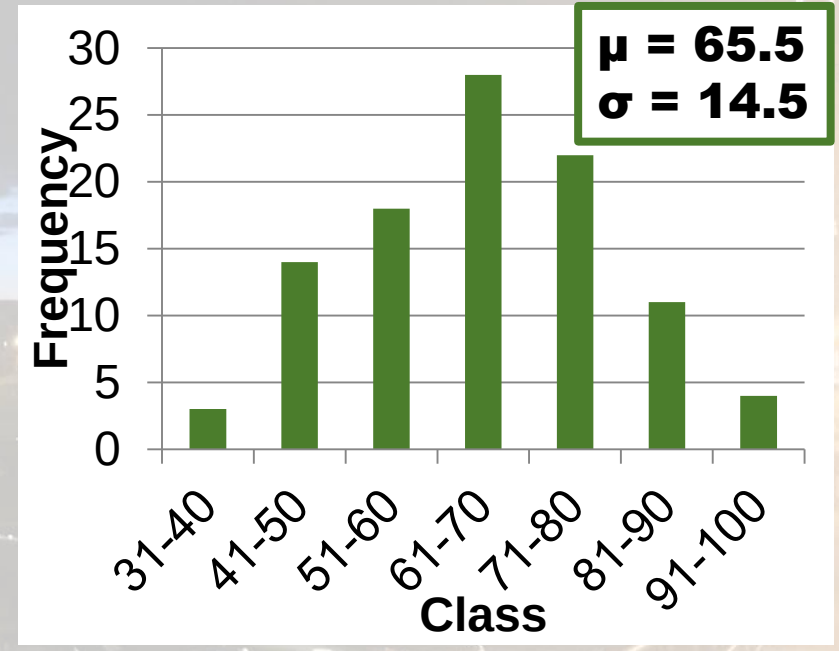
1. Applies to *ONLY* the normal distribution
2. Says "approximately"
   - 68% of the data are in interval μ +/- 1σ
   - 95% in interval μ +/- 2σ
   - 99.7% (almost all) in interval μ +/- 3σ

Empirical Rule Diagram

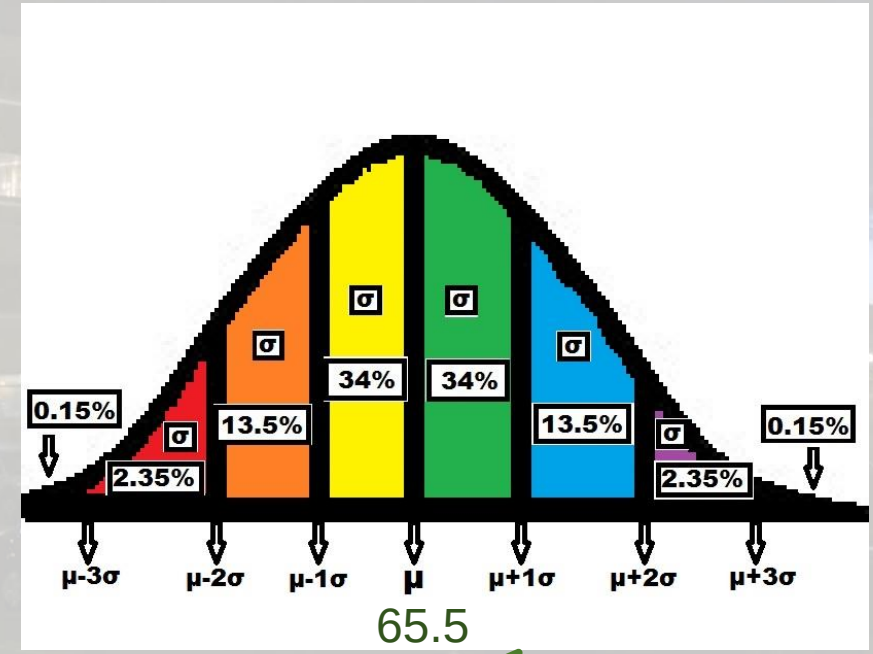


Applying Empirical Rule

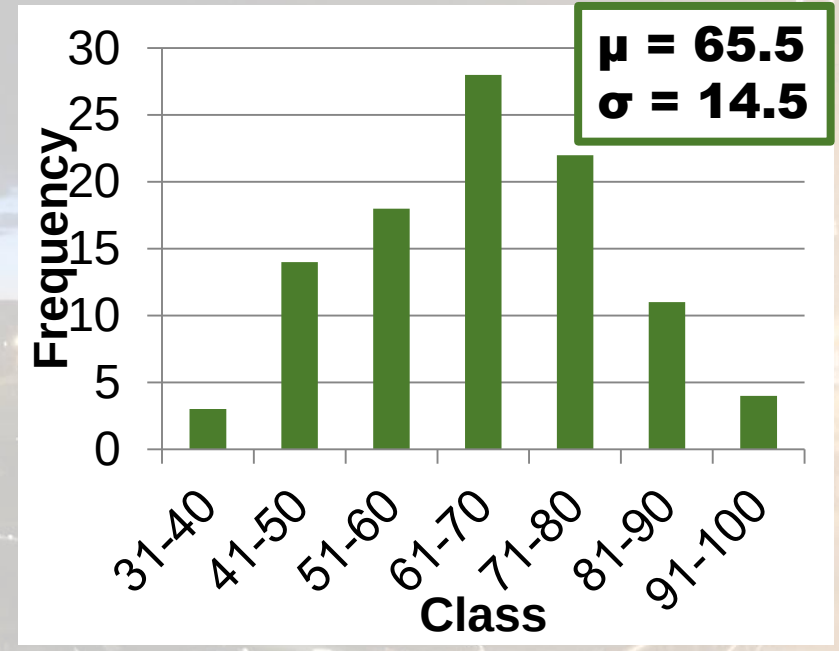




Applying Empirical Rule

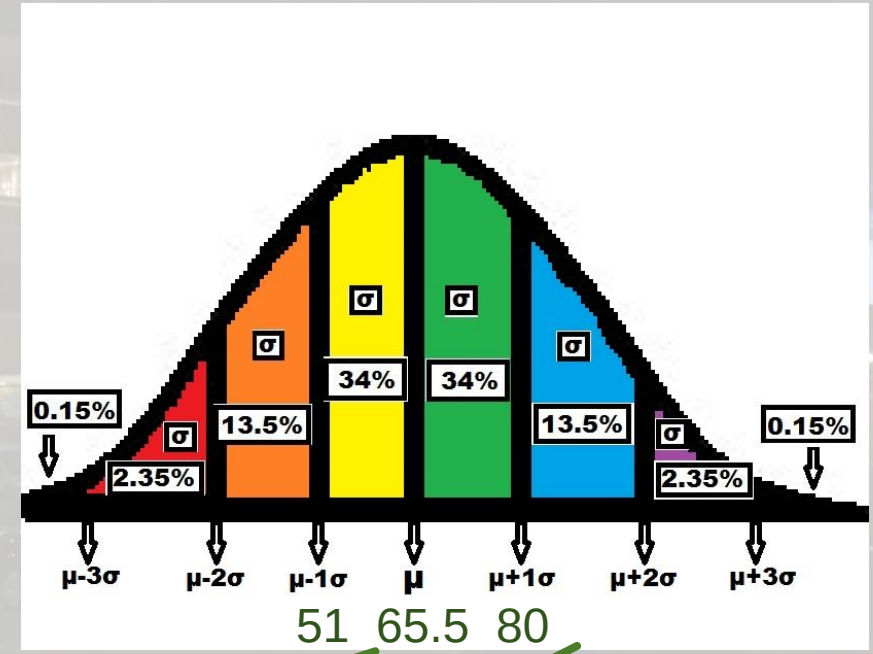


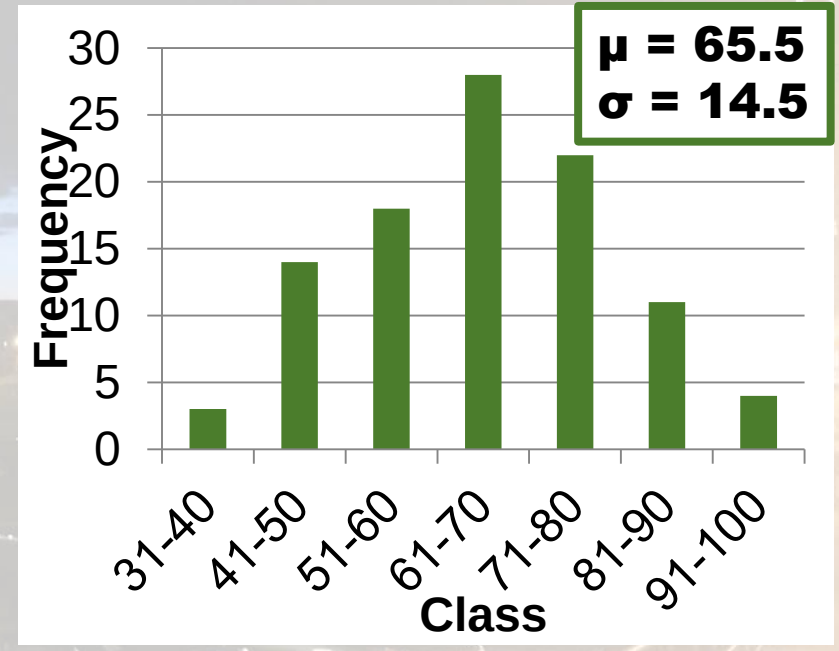
65.5



We can add the μ.

Applying Empirical Rule

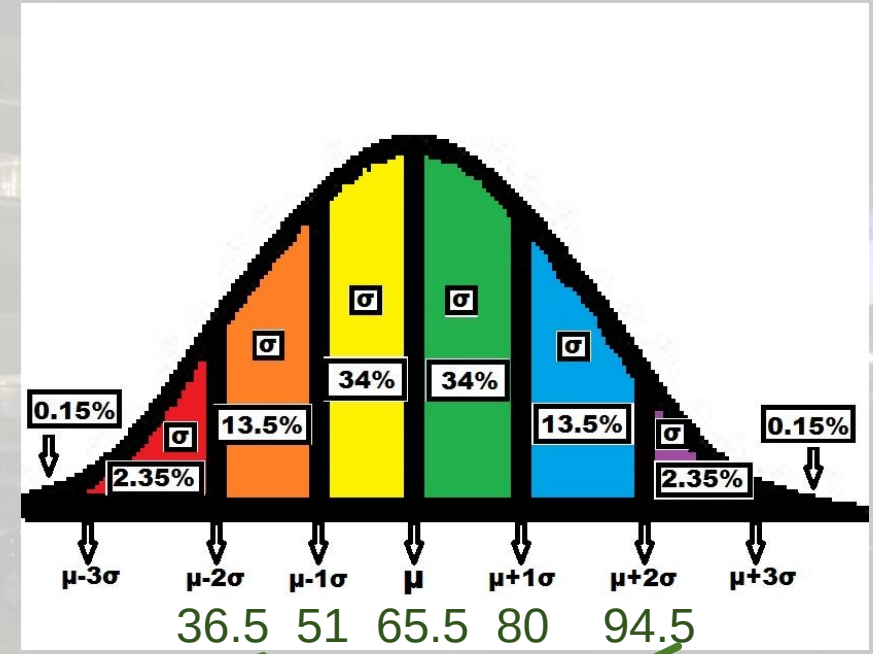


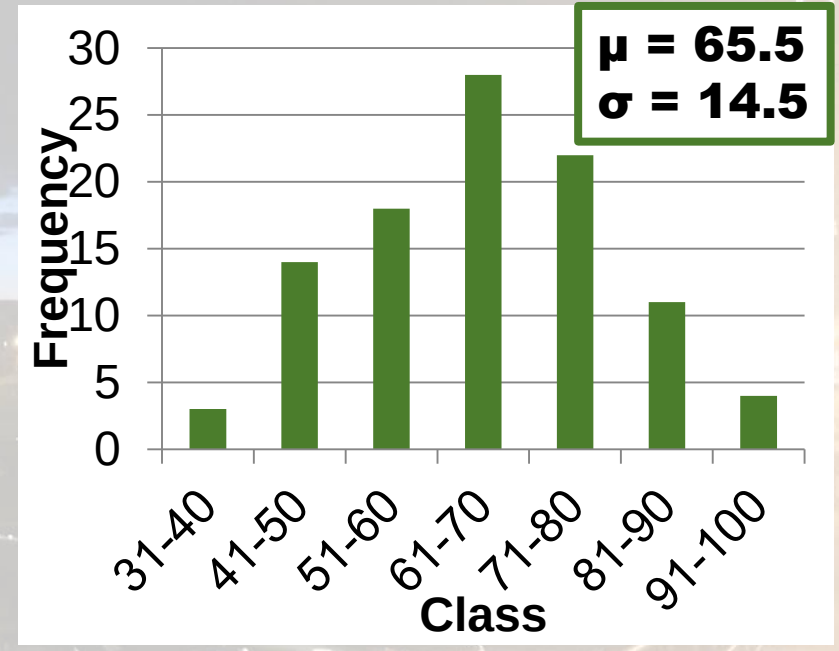


65.5 – (1*14.5) = 51

65.5 + (1*14.5) = 80

Applying Empirical Rule

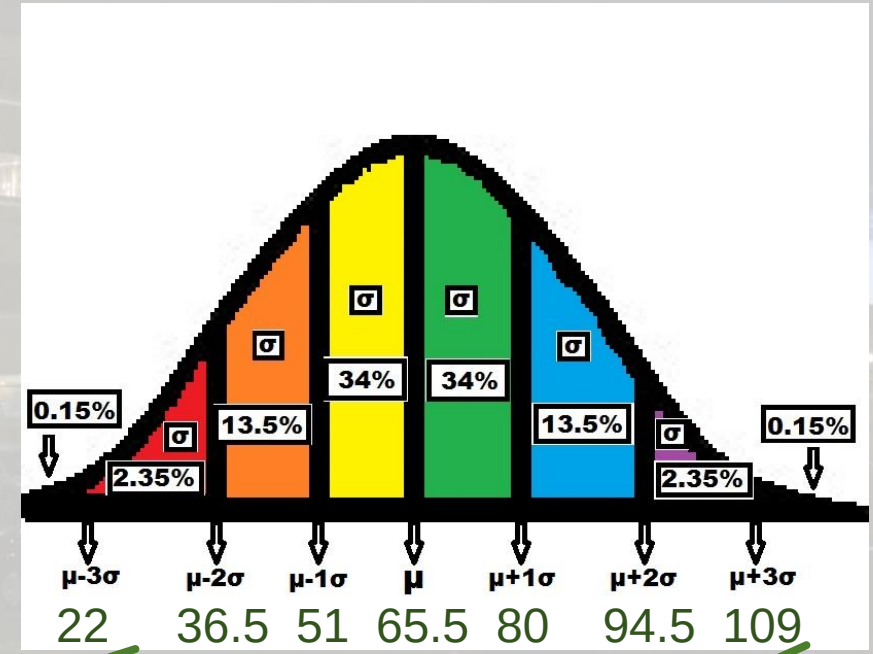


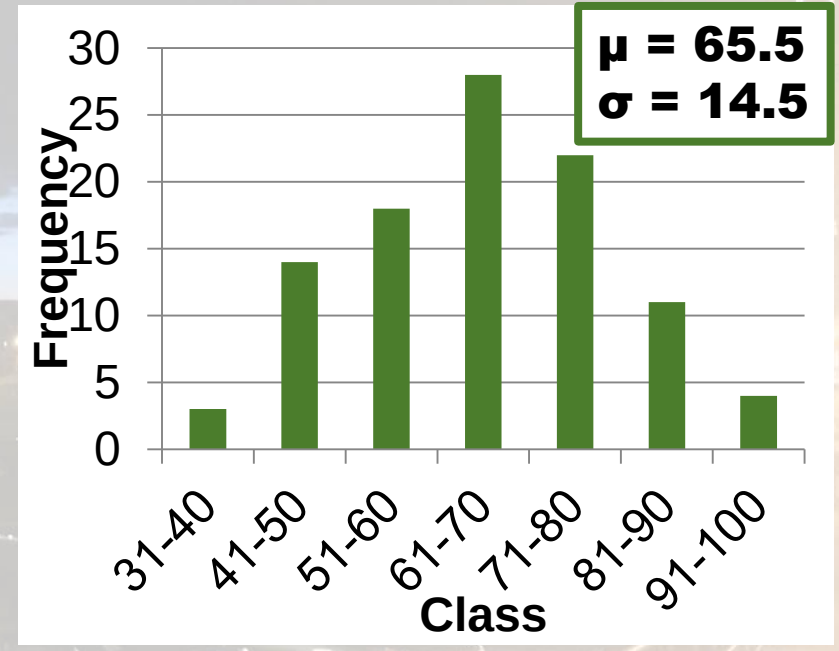


65.5 – (2*14.5) = 36.5

65.5 + (2*14.5) = 94.5

Applying Empirical Rule

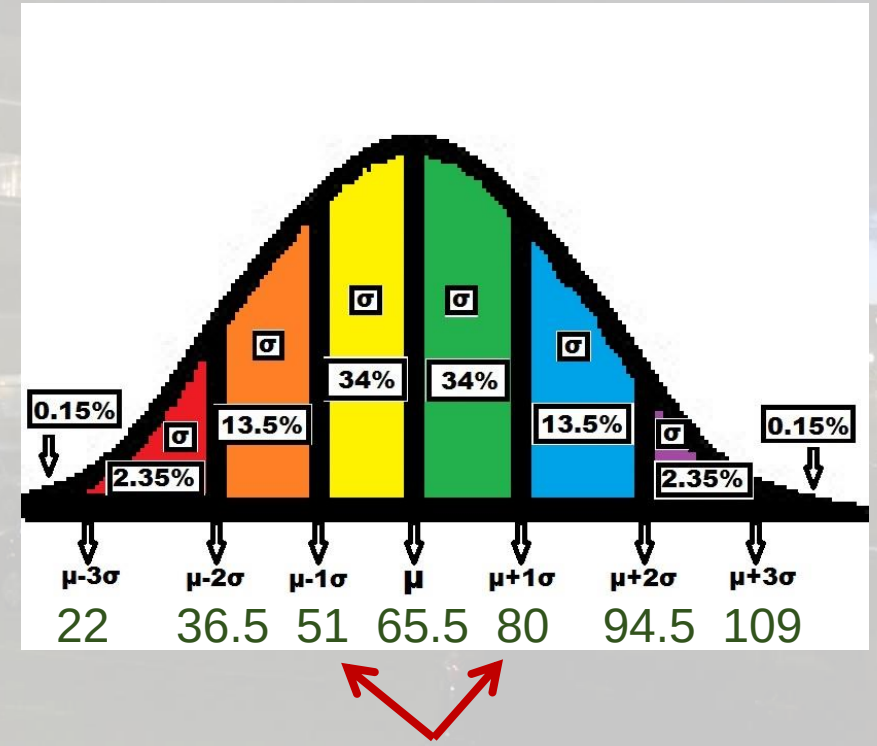




65.5 – (3*14.5) = 22

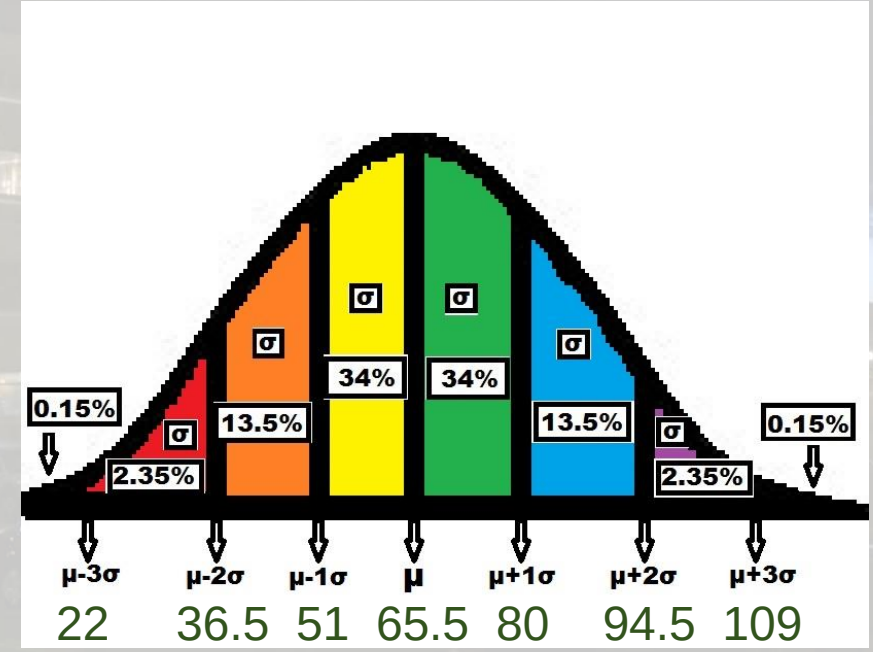
65.5 + (3*14.5) = 109

Applying Empirical Rule



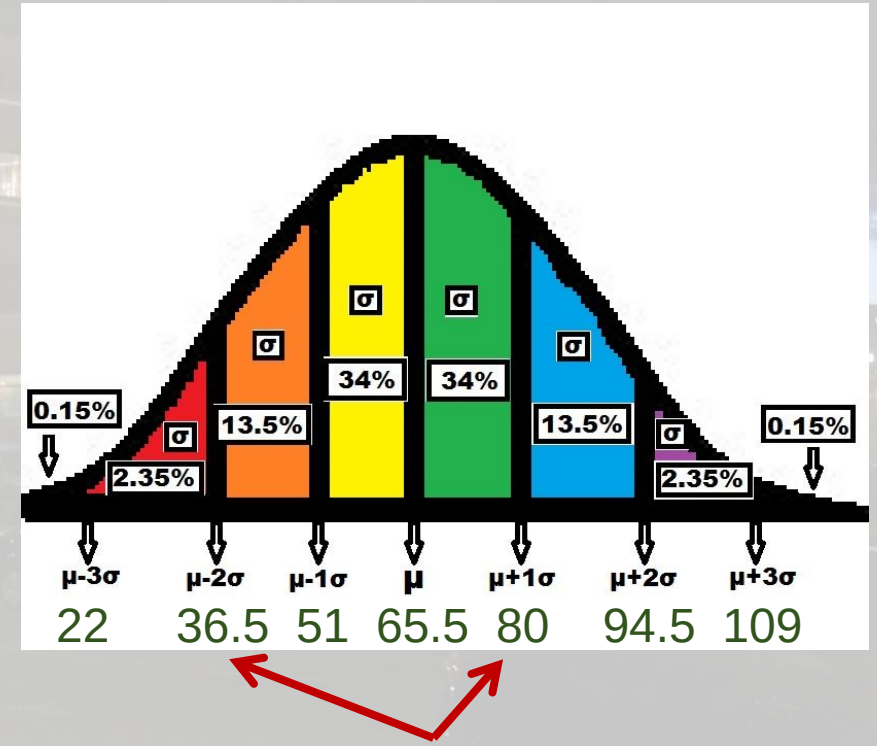
- Remember n=100?
- 34% of scores are between 51 and 65.5, meaning 34 scores.
- Another 34% (n=34) are between 65.5 and 80.
- This means 34%+34%=68% (n=68) are between 51 and 80.

Applying Empirical Rule



Question: What % of the data (student scores) are between 36.5 and 80?

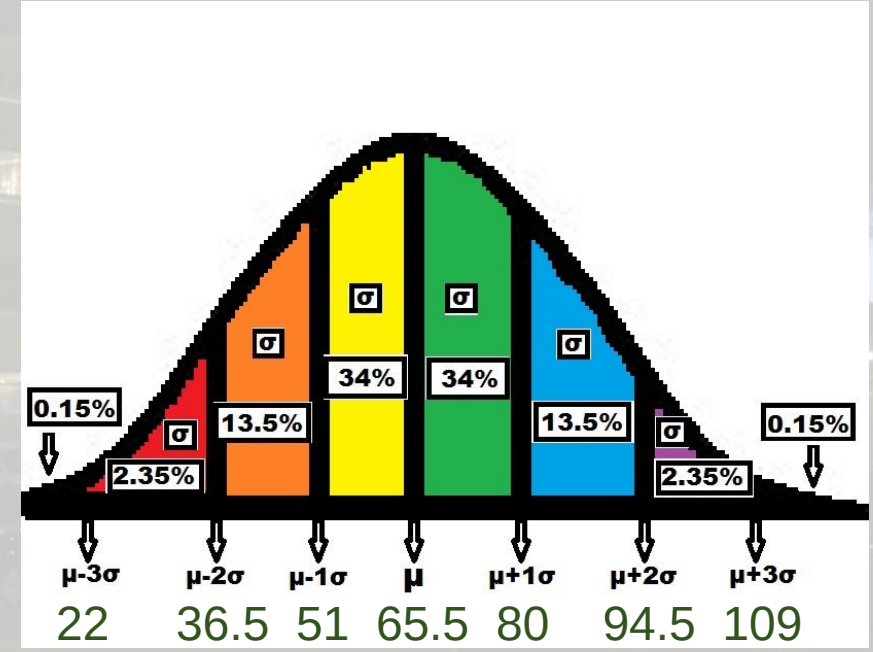
Applying Empirical Rule



Question: What % of the data (student scores) are between 36.5 and 80?

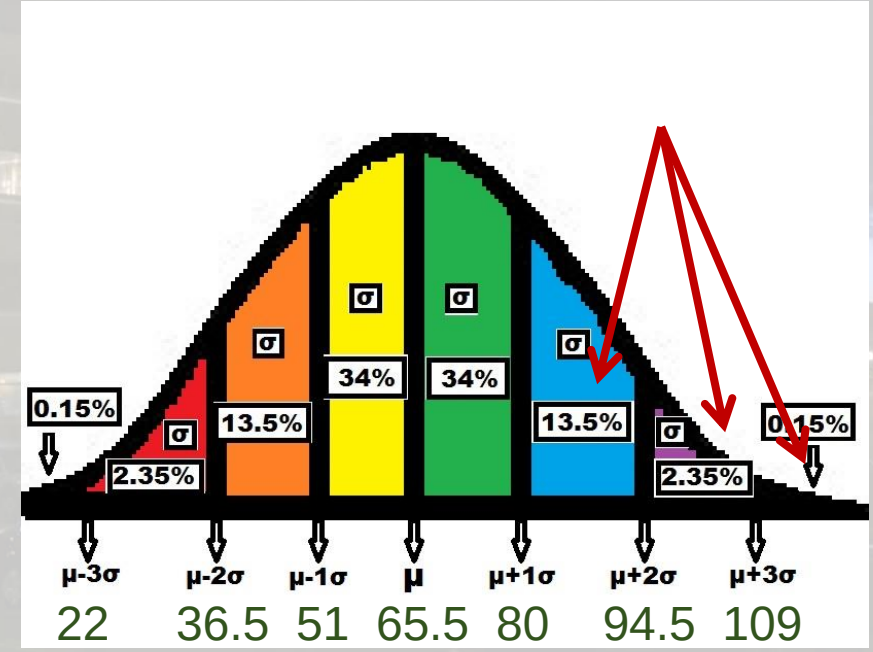
Answer: 13.5% + 34% + 34% = 81.5%

Applying Empirical Rule



Question: What cutpoint marks the top 16% of scores?

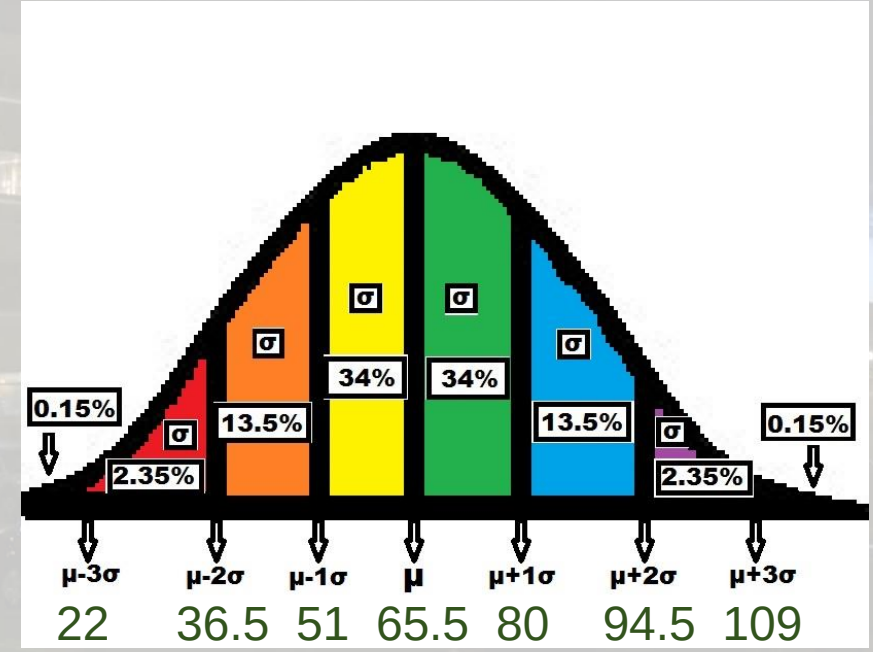
Applying Empirical Rule



Question: What cutpoint marks the top 16% of scores?

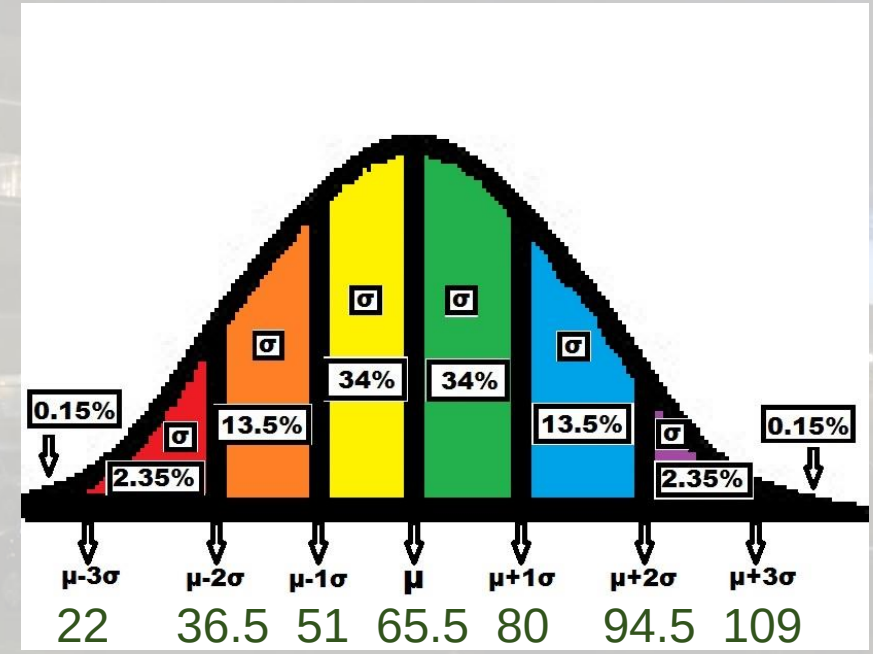
Answer: 0.15% + 2.35% + 13.5% = 16%. Top 16% cutpoint is 80.

Applying Empirical Rule



Question: What % of scores are below 94.5?

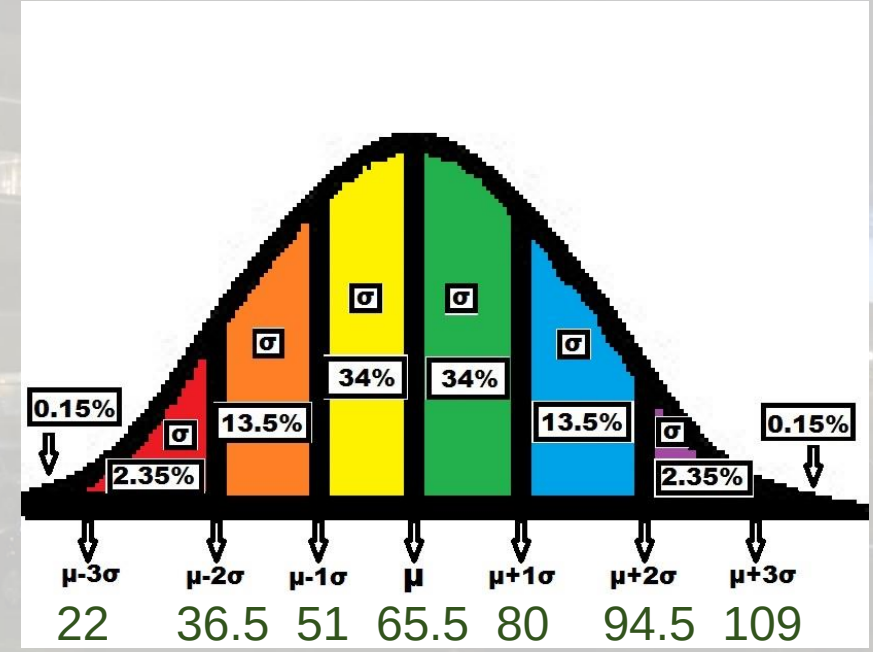
Applying Empirical Rule



Question: What % of scores are below 94.5?

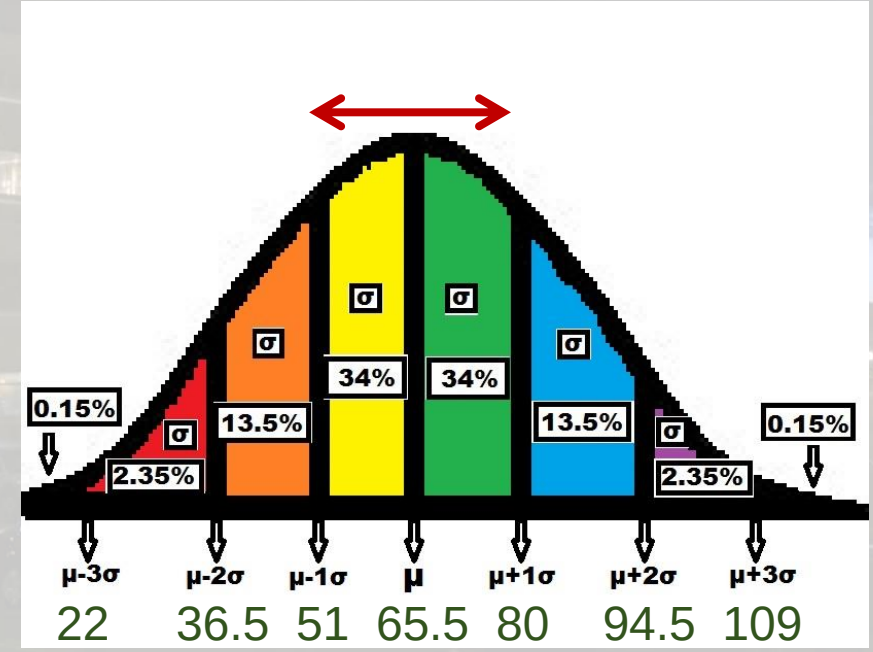
Answer: Add up % below 94.5: 13.5% + 34% + 34% + 13.5% + 2.35% + 0.15% = 97.5%

Applying Empirical Rule



Question: What are the cutpoints for the middle 68% of the data?

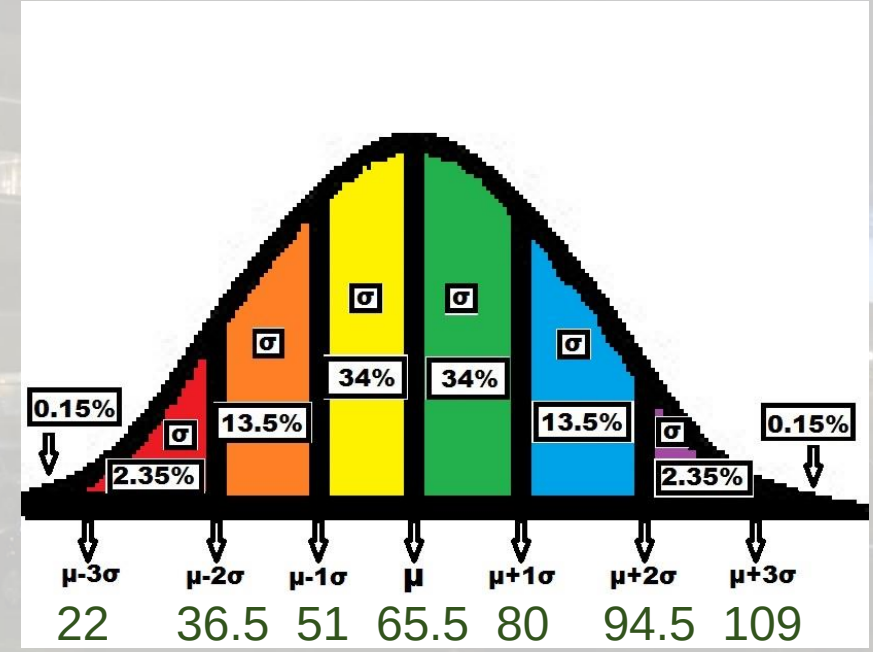
Applying Empirical Rule



Question: What are the cutpoints for the middle 68% of the data?

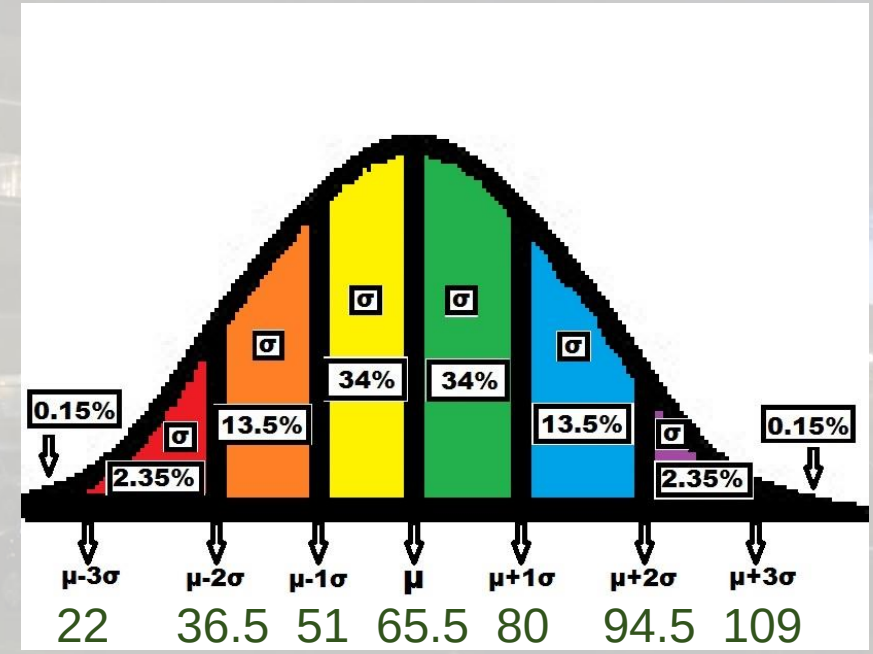
Answer: Middle 68% means 34% above mean (80) and 34% below mean (51).

Applying Empirical Rule



Question: What is the *probability* that if I select one student, that student will have a score less than 80?

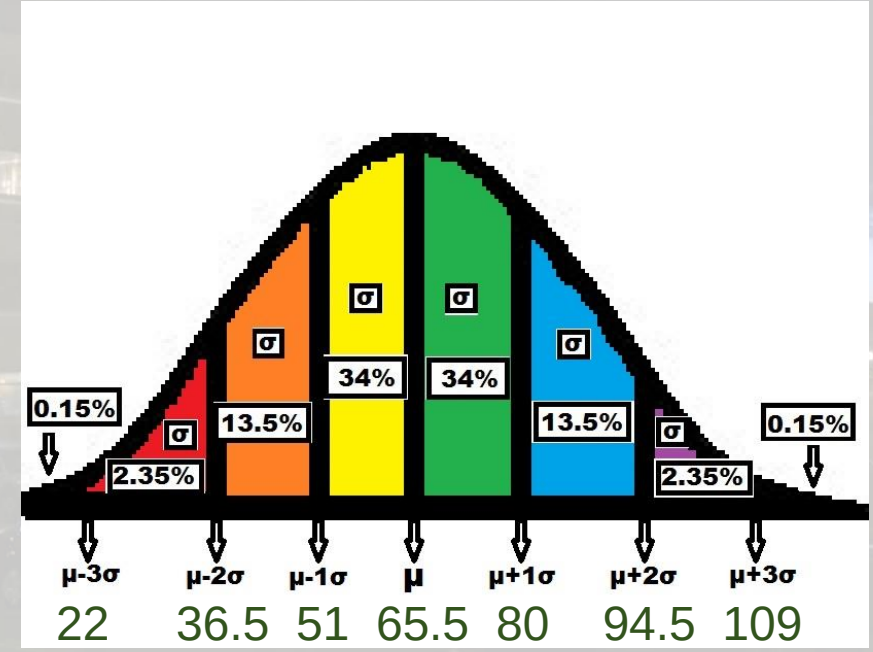
Applying Empirical Rule



Question: What is the *probability* that if I select one student, that student will have a score less than 80?

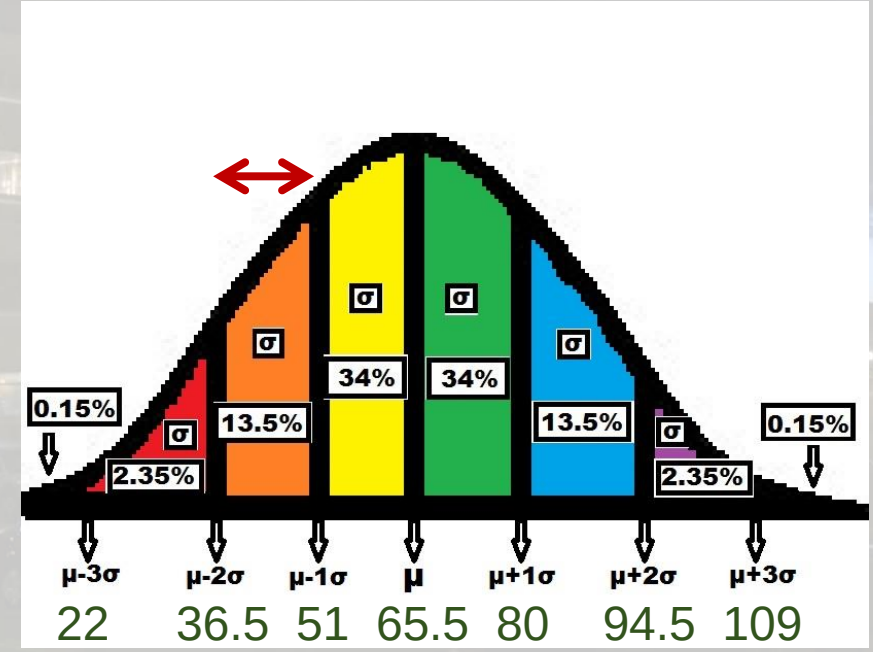
Answer: 50% + 34% = 84%

Applying Empirical Rule



Question: What is the *probability* I will select a student with a score between 36.5 and 51?

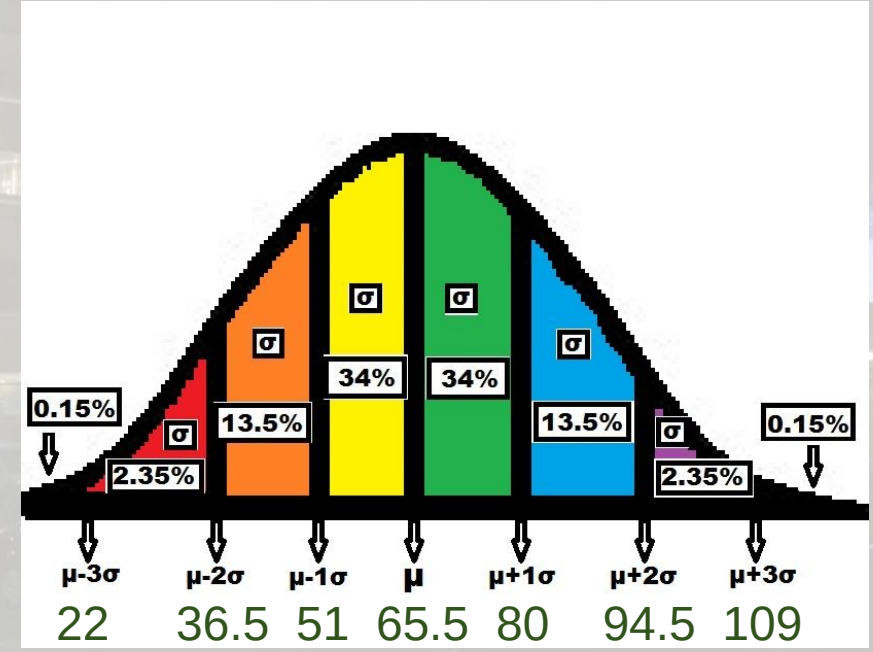
Applying Empirical Rule



Question: What is the *probability* I will select a student with a score between 36.5 and 51?

Answer: 13.5%

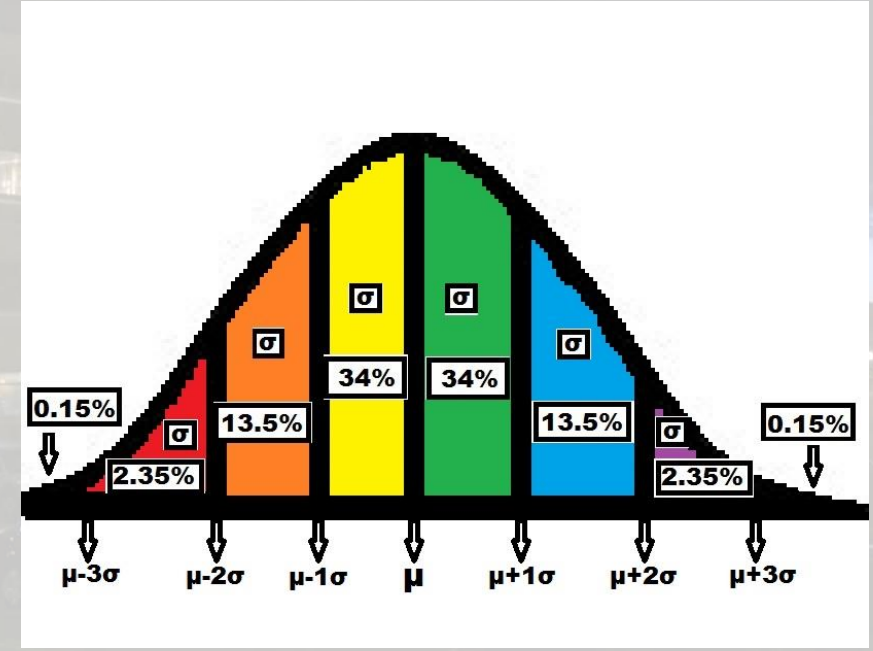
Applying Empirical Rule



Think about what would happen in a different class taking the same hard test.

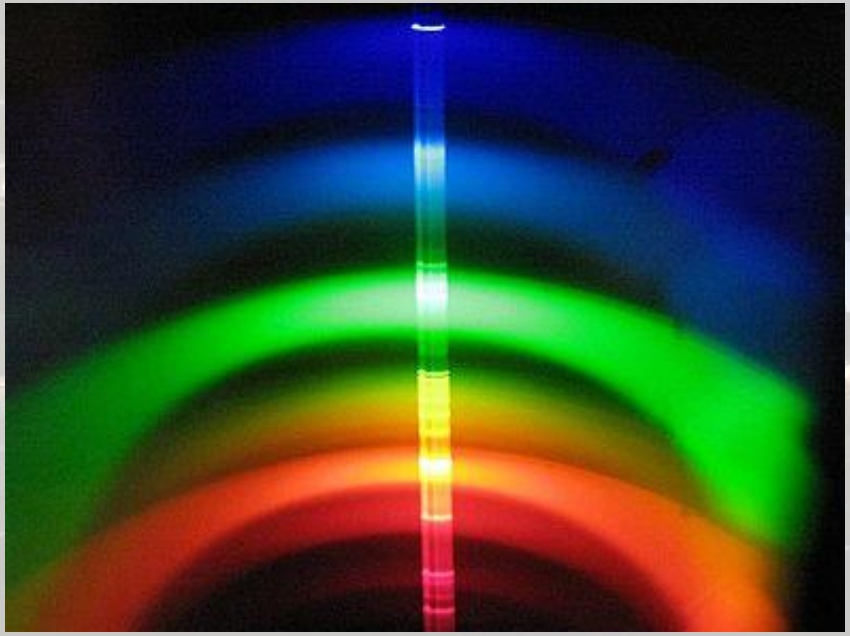
- What if the μ was the same, but the σ was larger than 14.5? What would that do to the intervals?
- What if σ was smaller than 14.5?

%, Area, and Probability: Did you Know?



- The %'s literally refer to the % of the area of the shape.
- Imagine the whole thing as 100%.
- Example:
  - The orange part is 13.5% of the area
  - It is also the probability that an "x" between μ-1σ and μ-2σ will be selected

Conclusion



- The Empirical Rule helps establish intervals that apply to normally distributed data
- These intervals have a certain percentage of the data points in them
- These intervals depend on the mean and standard deviation of the data distribution

Image by Zaereth