



Correlation

When one trait or behavior accompanies another, we say the two correlate.

Correlation
Coefficient (r)

Indicates strength
of relationship
(0.00 to 1.00)

Correlation Coefficient is a statistical measure of relationship between two variables.

- ± 1

$r = +0.37$

Indicates direction
of relationship
(positive or negative)



A. Which correlation coefficient is the strongest?

Example 1

a. $+.95$

b. $+.75$

Example 2

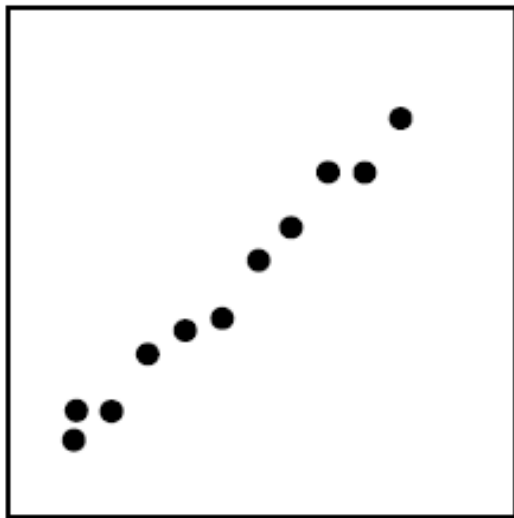
a. $-.45$

b. $-.15$

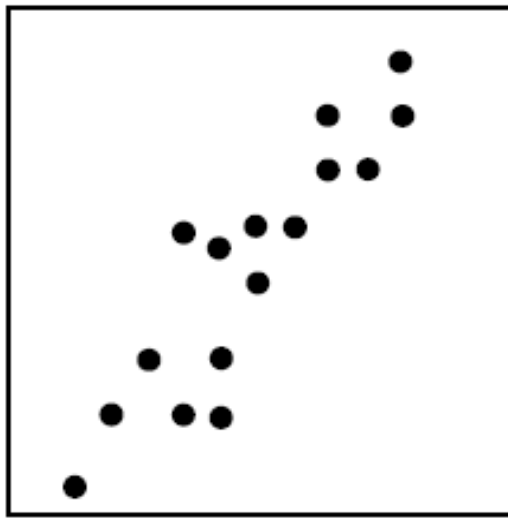


B. Match the correct scatterplots

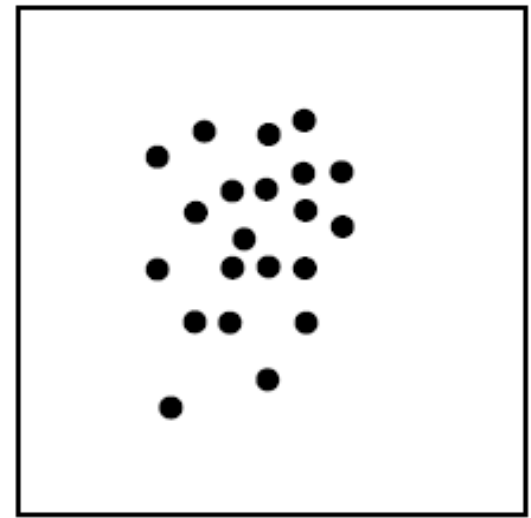
1. Which ones are positively correlated?
2. Which ones are negatively correlated?
3. Which one has strong negative correlation?
4. Which one has moderately positive correlation?
5. Which one has no correlation?



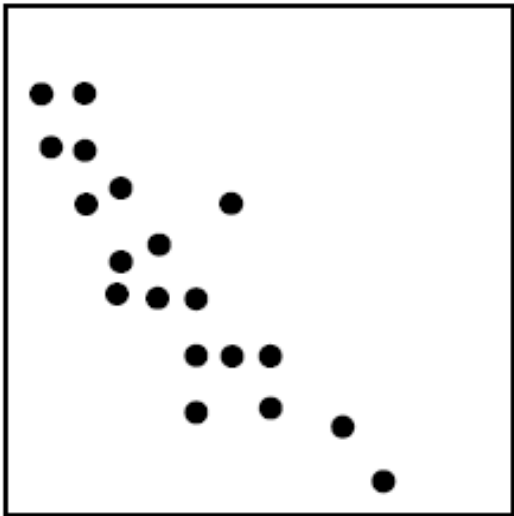
A



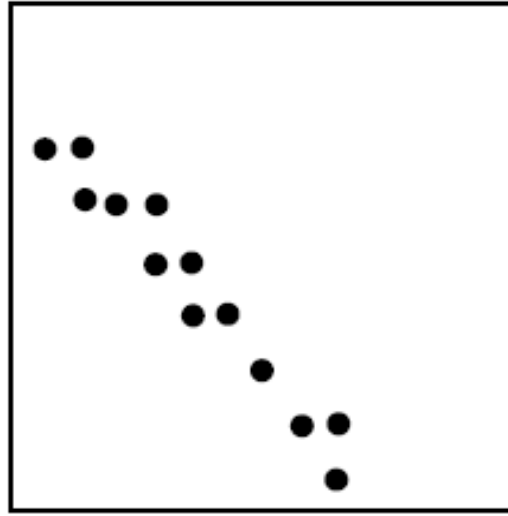
B



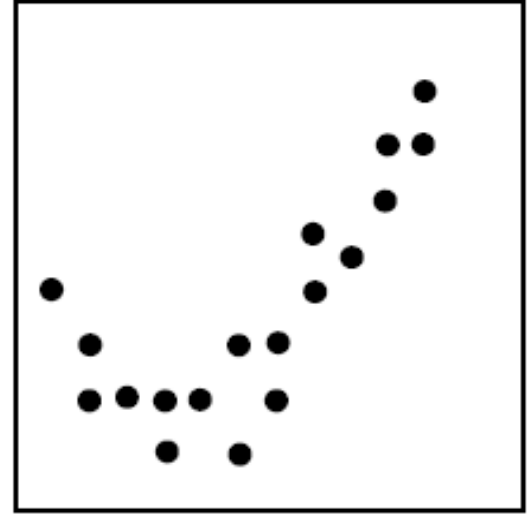
C



D



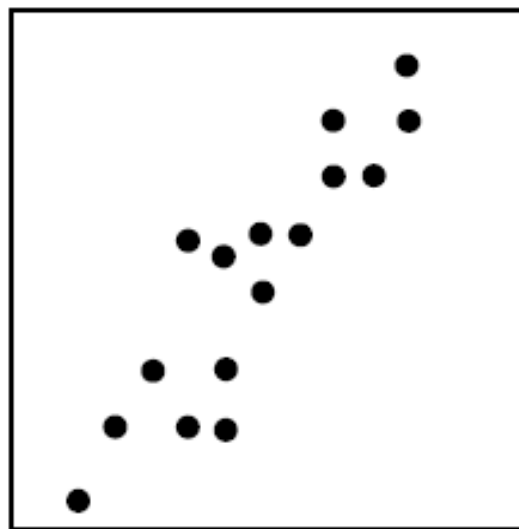
E



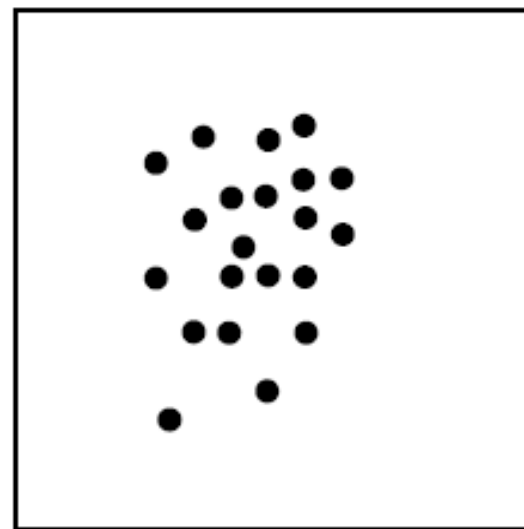
Curvilinear
relationship



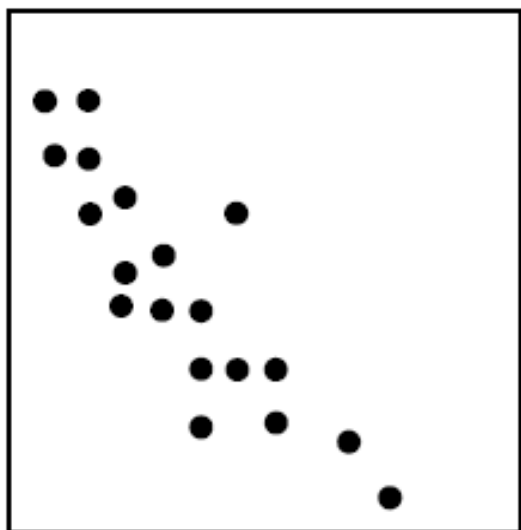
Strong positive correlation



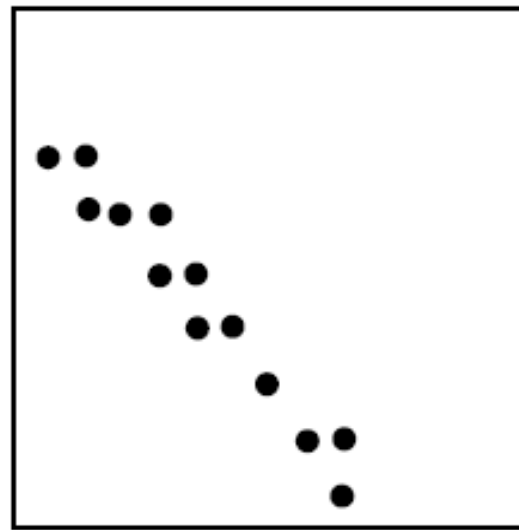
Moderate positive correlation



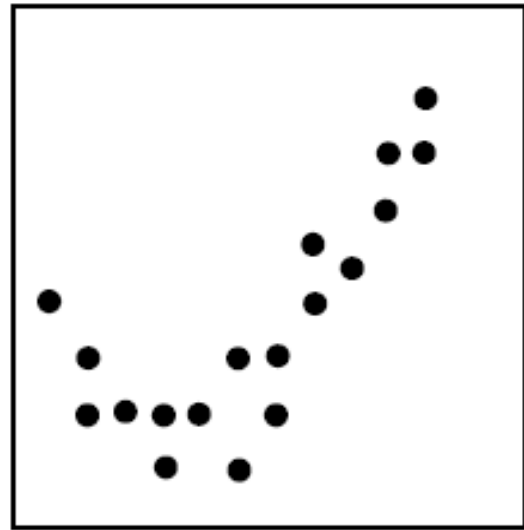
No correlation



Moderate negative correlation



Strong negative correlation

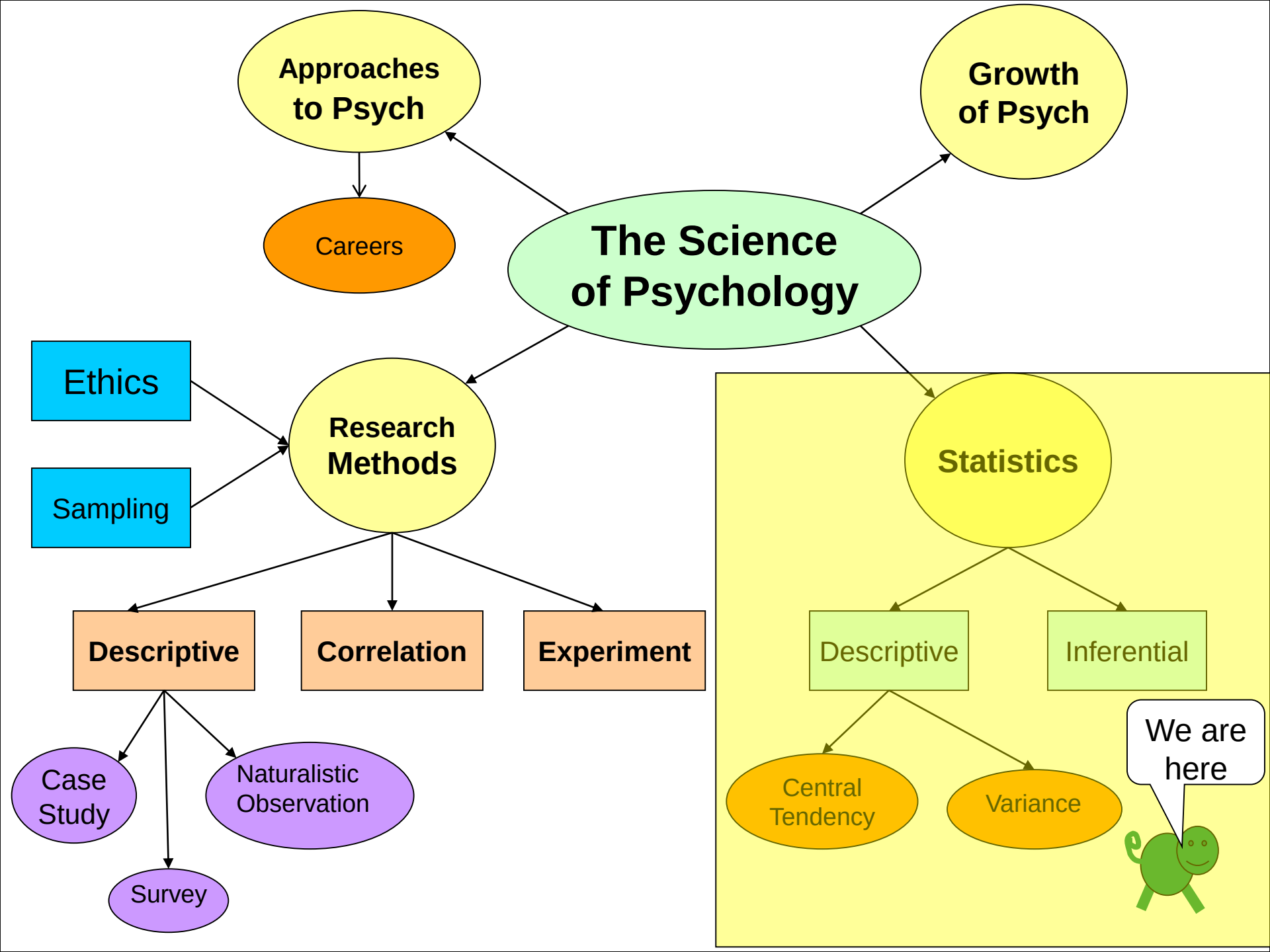


Curvilinear relationship

Unit 1: Science of Psychology

Essential Task 1-8: Apply basic statistical concepts to explain research findings:

- Descriptive Statistics: Central Tendency (*mean, median, mode, skewed distributions*) Variance (*range, standard deviation, and normal distributions*)
- Inferential Statistics: *Statistical significance (t-test and p-value)*





Essential Task 1-:

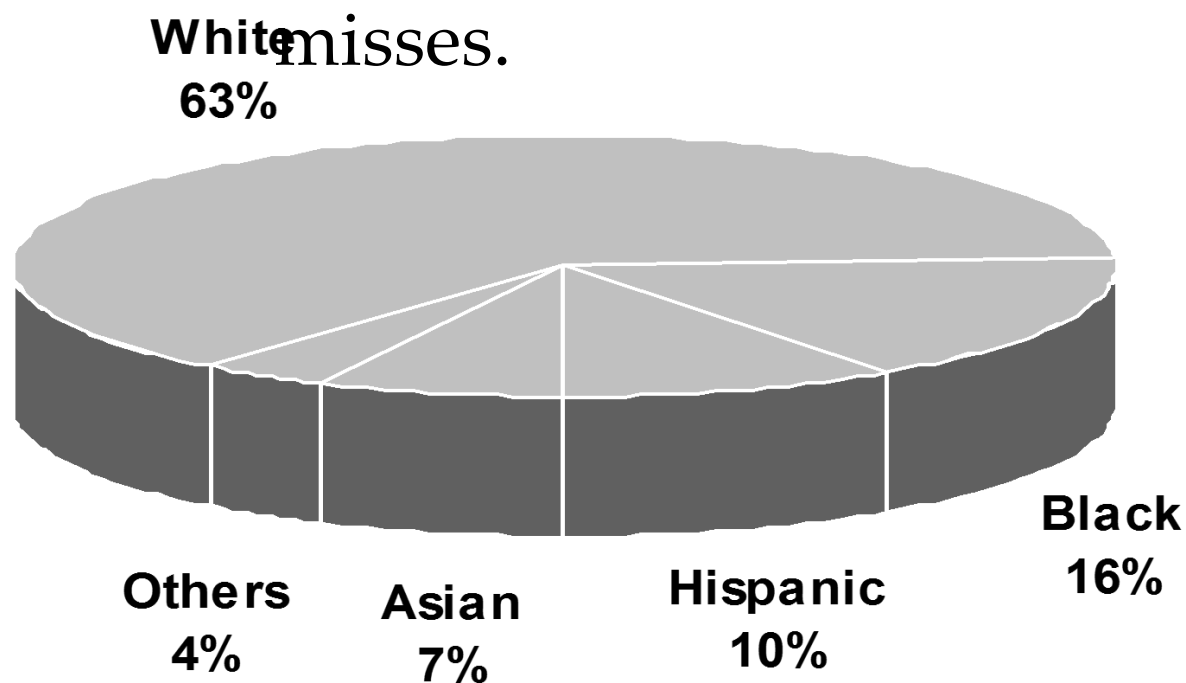
Outline

- Descriptive Statistics:
 - Central Tendency
 - *Mean, median, and mode*
 - *skewed distributions*
 - Variance
 - *Range*
 - *standard deviation*
 - *normal distributions*
- Inferential Statistics:
 - *Statistical significance*
 - *t-test* and the *p-value*
 - Confidence intervals



Statistical Reasoning

Statistical procedures analyze and interpret data and let us see what the unaided eye misses.



Composition of ethnicity in urban locales



Descriptive Statistics

- Summarizes numbers so they became more meaningful and easier to communicate to others
 - Describes data
 - Measures central tendency



Central Tendency

- Tendency of scores to congregate around some middle variable
- A measure of central tendency identifies what is average or typical in a data set



Measures of Central Tendency

Mode: The most frequently occurring score in a distribution.

- Easy to obtain, but often unreliable

Mean: The arithmetic average of scores in a distribution obtained by adding the scores and then dividing by their number.

- sensitive to extremely high and low numbers (skews the data)

Median: The middle score in a rank-ordered distribution.



C. Find the central tendencies

1, 2, 3, 3,3, 4, 5, 6, 7

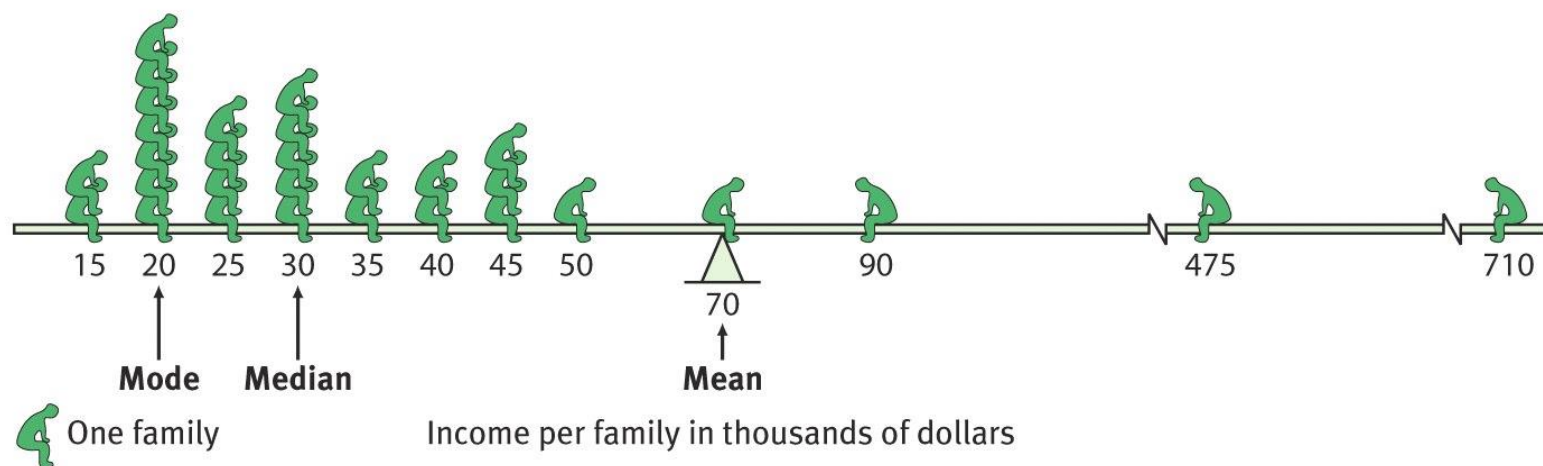
Mean:

Median:

Mode:



But the mean doesn't work in a skewed distribution



The Median is a much better measure of the center



Normal and Skewed Curves

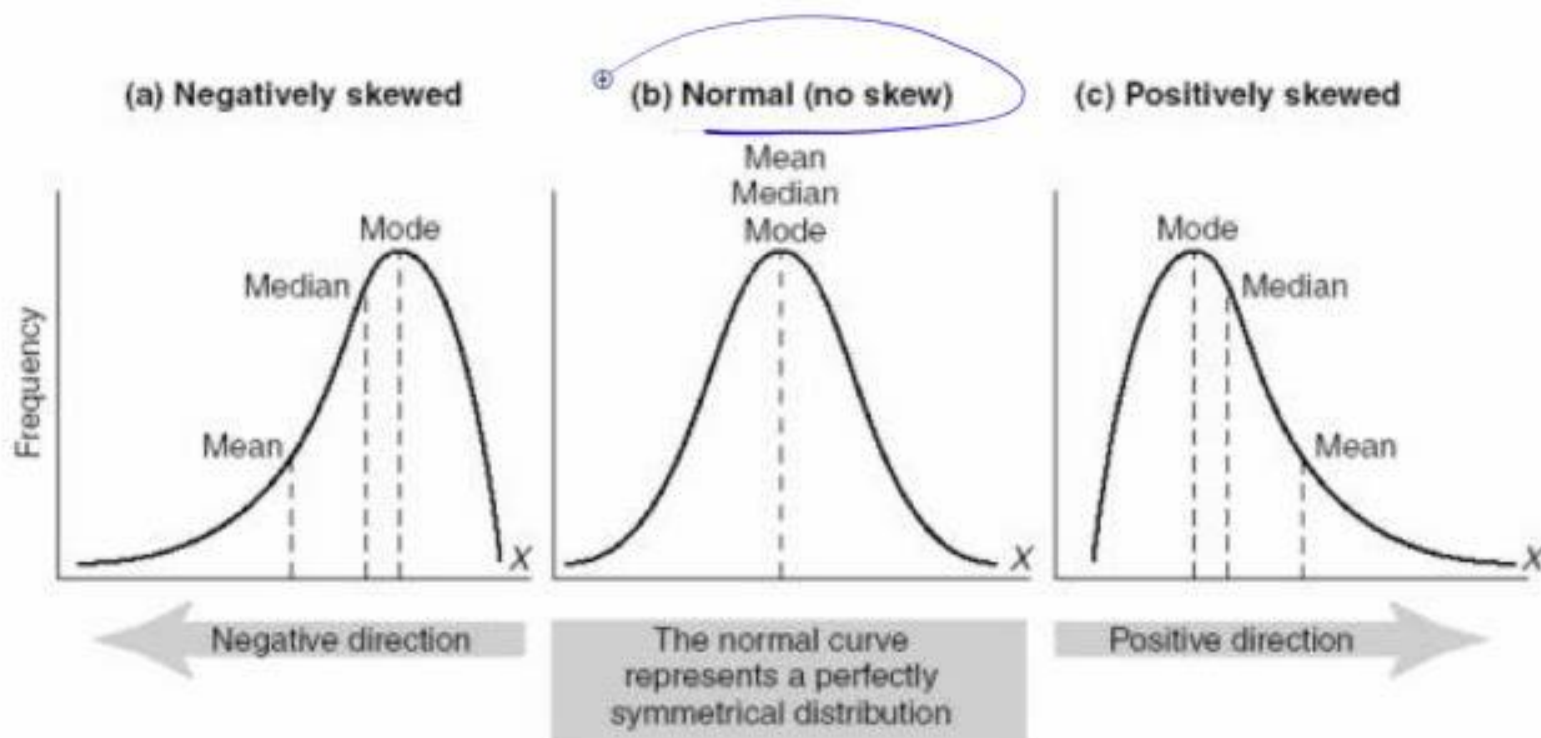


FIGURE 15.6 Examples of normal and skewed distributions

Image courtesy of: http://edn523.wikispaces.com/file/view/distribution_curves.jpg/77287639/distribution_curves.jpg



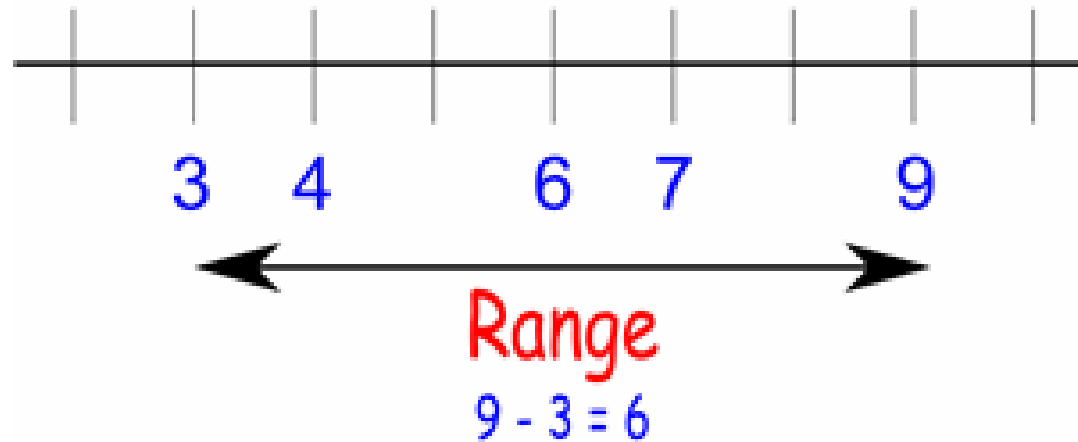
Measures of Variation

- Statistical dispersion (how distributed the data points are) is a key concept in statistics.
- Two key ways of measuring statistical dispersion
 - » Range
 - » Standard Deviation



1. Range

- The range simply gives the lowest and highest values of a data set.





D. Sample Questions

- What is the **range** between 4 to 25?
- What is the **range** between 5, 7, 8, 10?



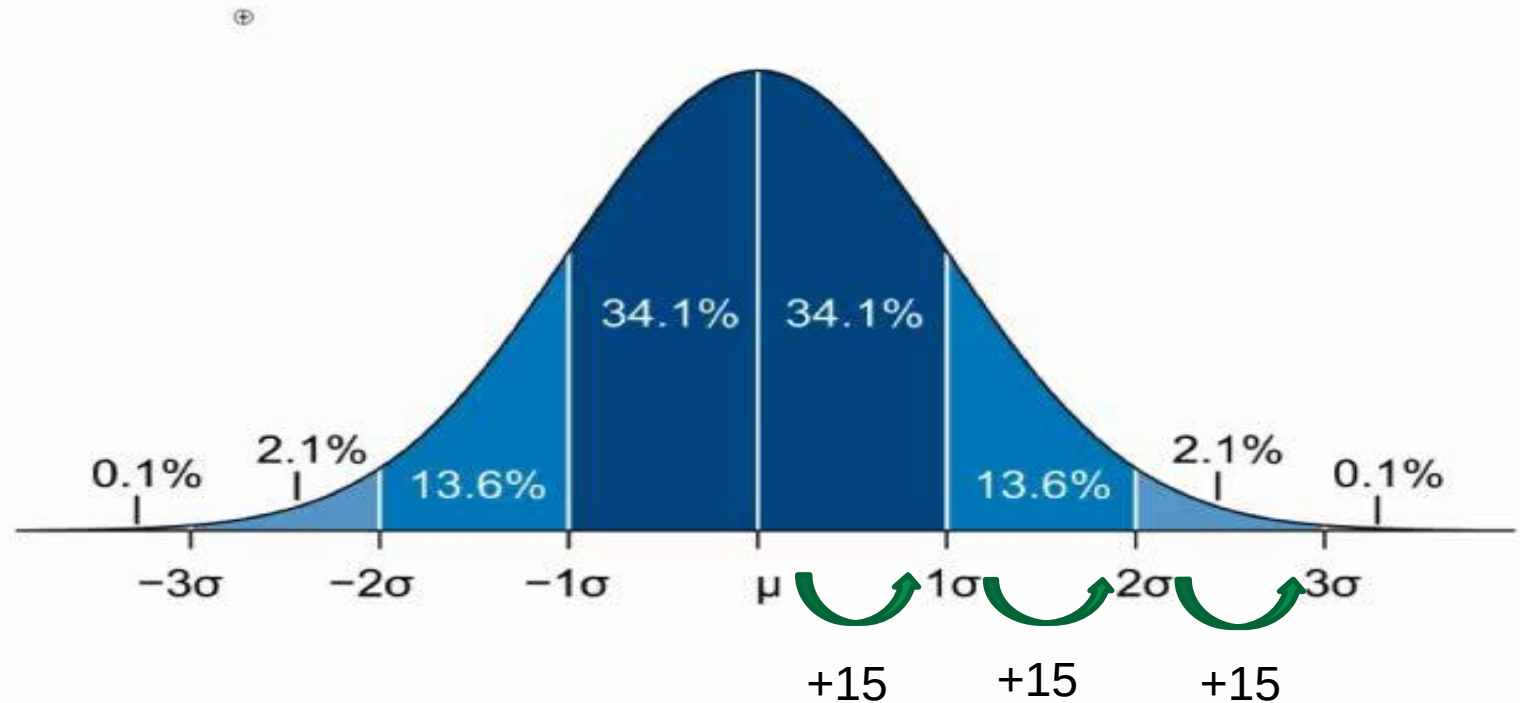
2. Standard Deviation

- Standard deviation gives a measure of dispersion.
- Essentially, they are measures of the average difference between the values.
 - Low deviation means scores are compressed
 - High deviation means scores are spread out
- Standard deviation gives a value that is directly comparable to your mean values.



Standard
Deviation: 15

Standard Deviation





E. Sample Questions

The mean in an IQ test is 100 . the Standard Deviation is 15.

1. What is Jake's IQ if he is 2 Standard Deviation above the mean?
2. What is Finn's IQ if he is 1 Standard Deviation below the mean?
3. If John Cena's IQ score is 85 and the mean is 100 and the S.D. is 5, which Standard deviation does he fall under?



E. Sample Questions

Example:

If Johnny Uppercut's IQ score is 85 and the mean is 100 and the S.D. is 10, which Standard deviation does he fall under?

3. If John Cena's IQ score is 85 and the mean is 100 and the S.D. is 5, which Standard deviation does he fall under?



Z score

Z score measure distance of a score from the mean in units of standard deviation.

Example 1

Clarence scored a 72 on a test with a mean of 80 and SD of 8. What would Clarence's Z score be?



Z score

Example 2

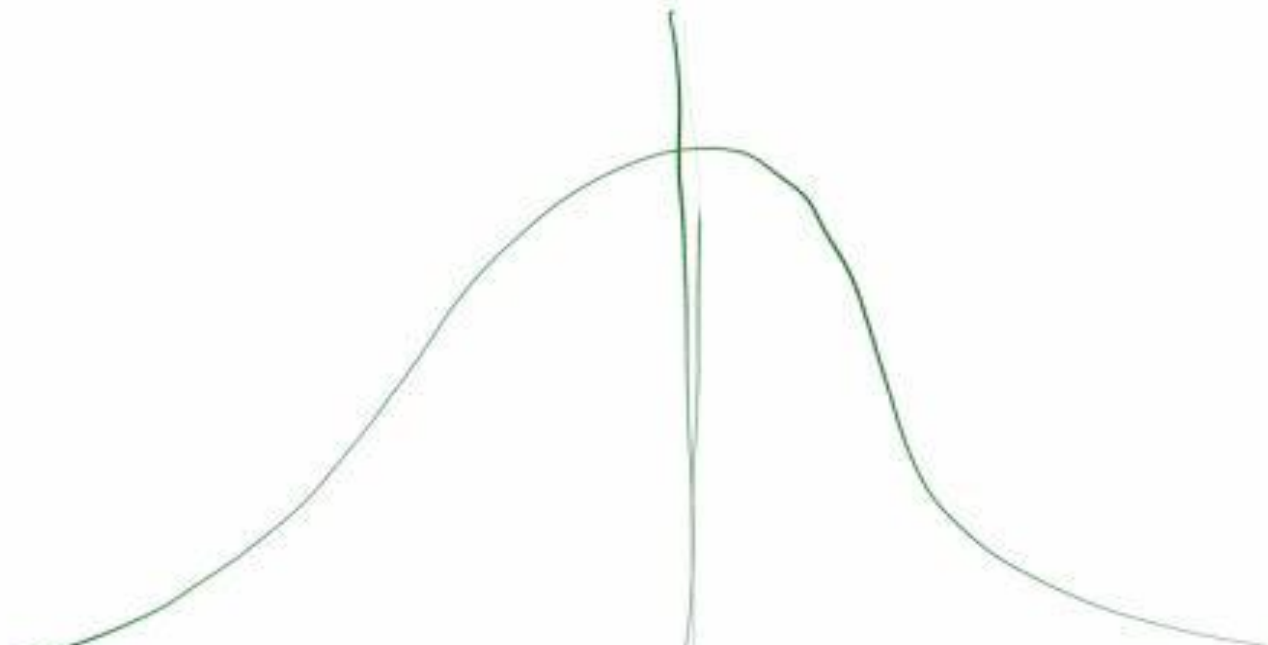
Sumo scored a 84 on a test with a mean of 80 and SD of 8. What would Sumo's Z score be?

Examples of Standard Deviation Problems

- What if the mean is not 100?
- E.g. mean is 85 and SD is 6
- E.g. mean is 45 and SD is 3

NORMAL
CURVE

-3	-2	-1	$\bar{x}$	+1	+2	+3
67	73	79	85	91	97	103



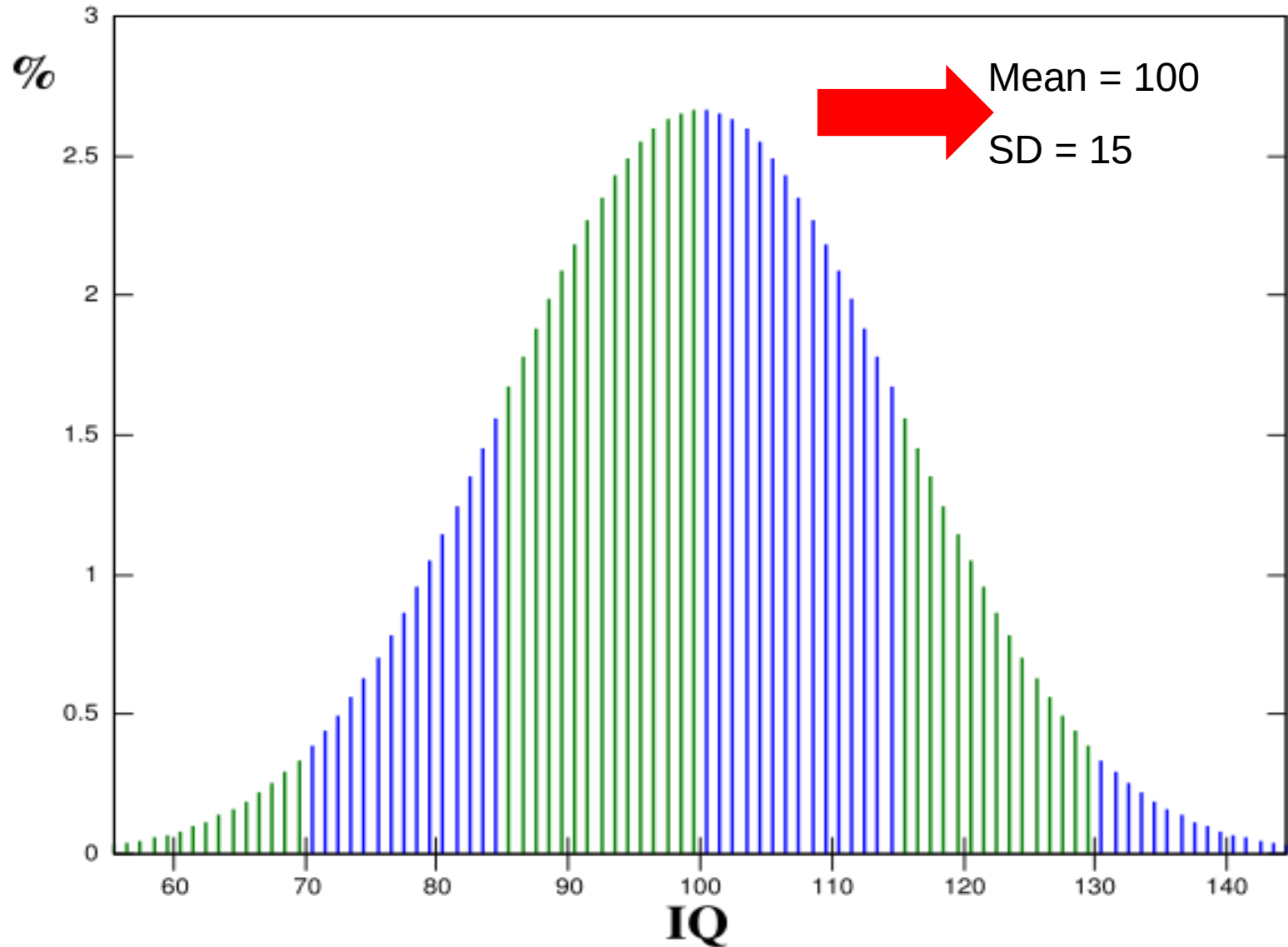


Normal Distribution (Bell Curve)

- If it is a Normal Distribution/Curve, always assume the Mean/Median/Mode is 100.
- IQ score will always have a normal distribution
- Standard Deviation is 15

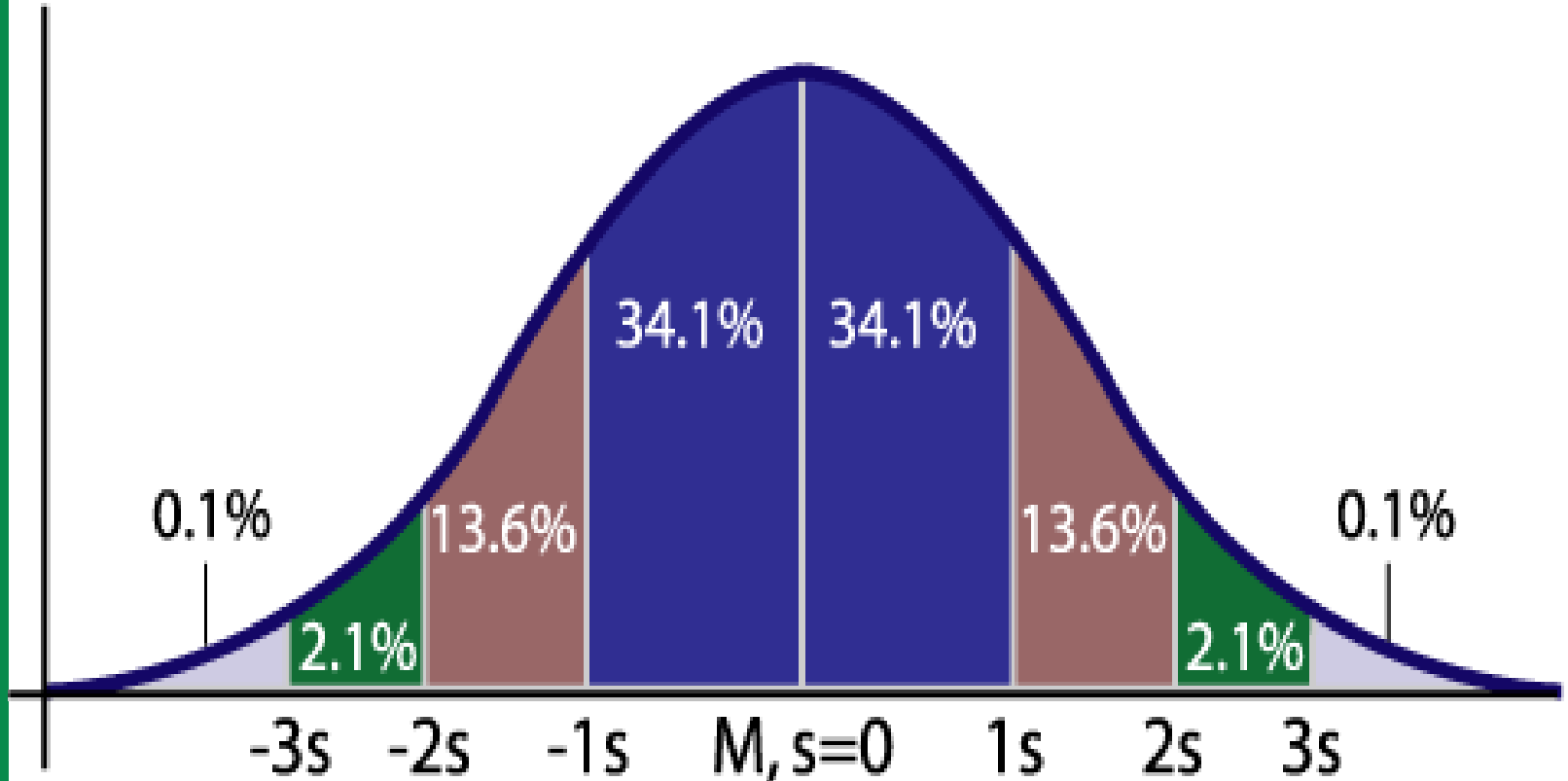


IQ follows a Normal Distribution



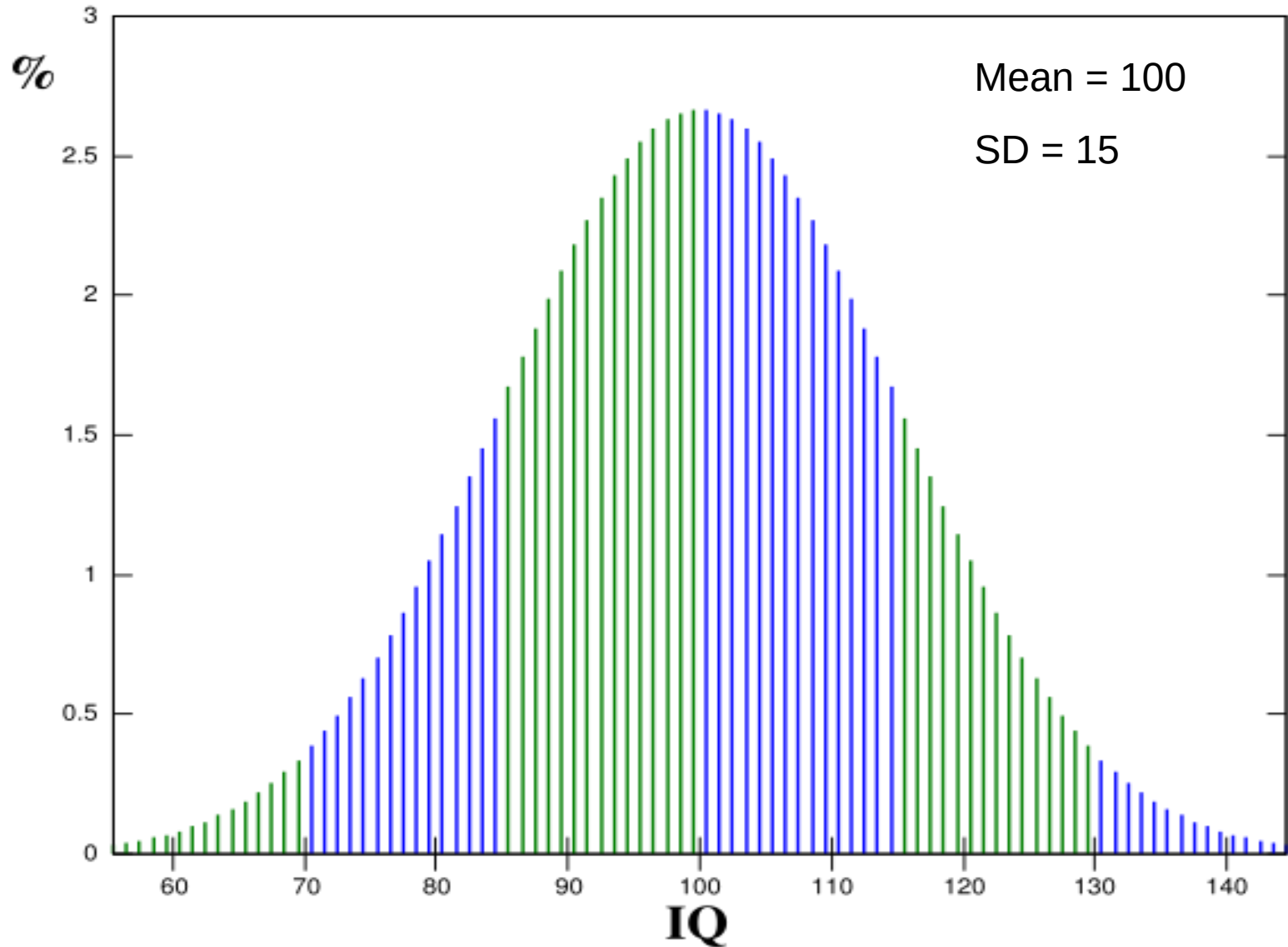


Normal Distribution (Bell Curve)



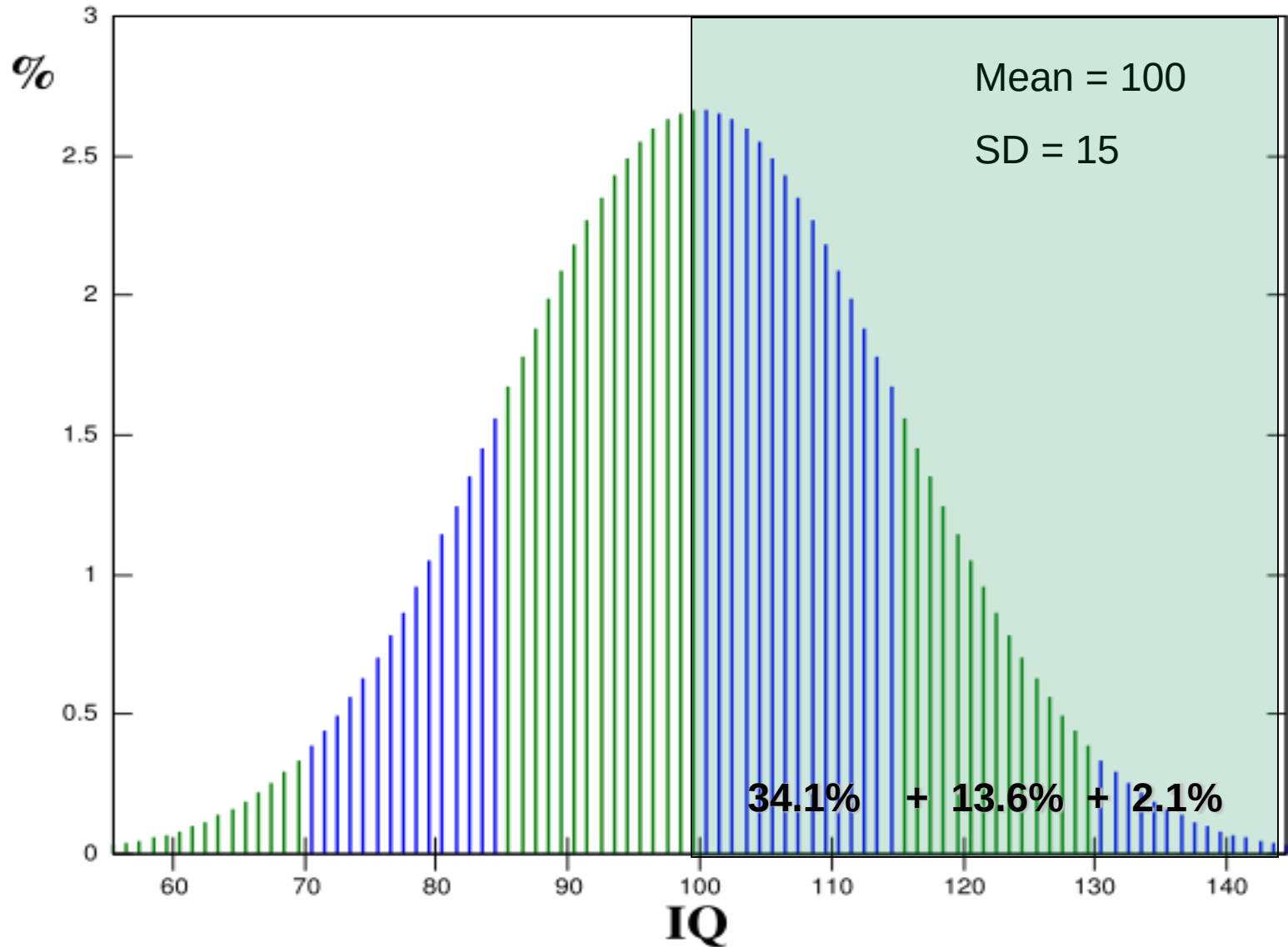


What percentage score below 100?



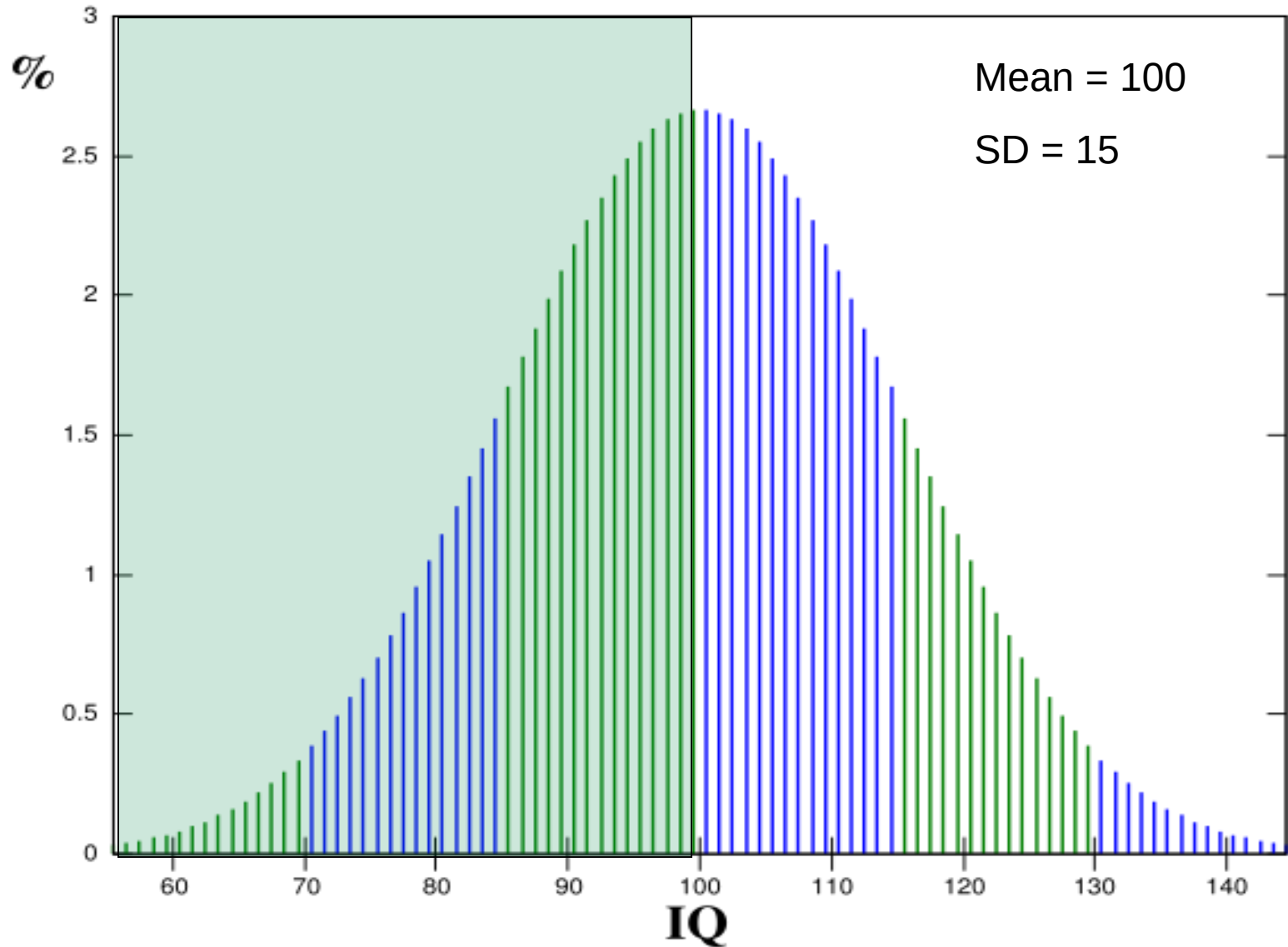


What percentage score above 100?





What percentage score below 100?





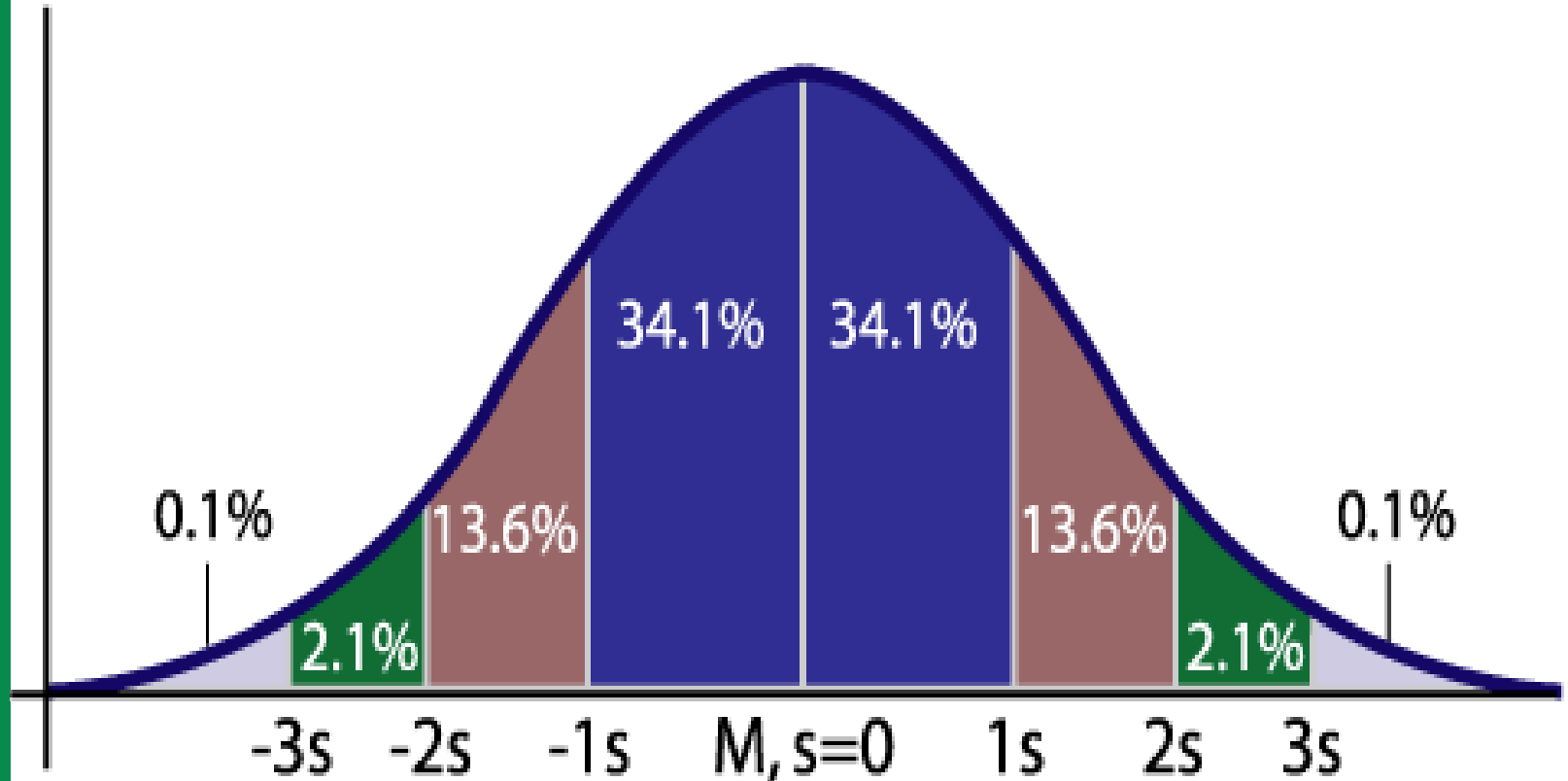
F. Sample Question

If the Mean is at 100 and SD is 15

1. What is the percentage of people scored between 85 and 100?
2. What percentage of people scored between 85 and 115?
3. What percentage of people scored between 70 and 130?
4. What percentage score below 70 and above 130?

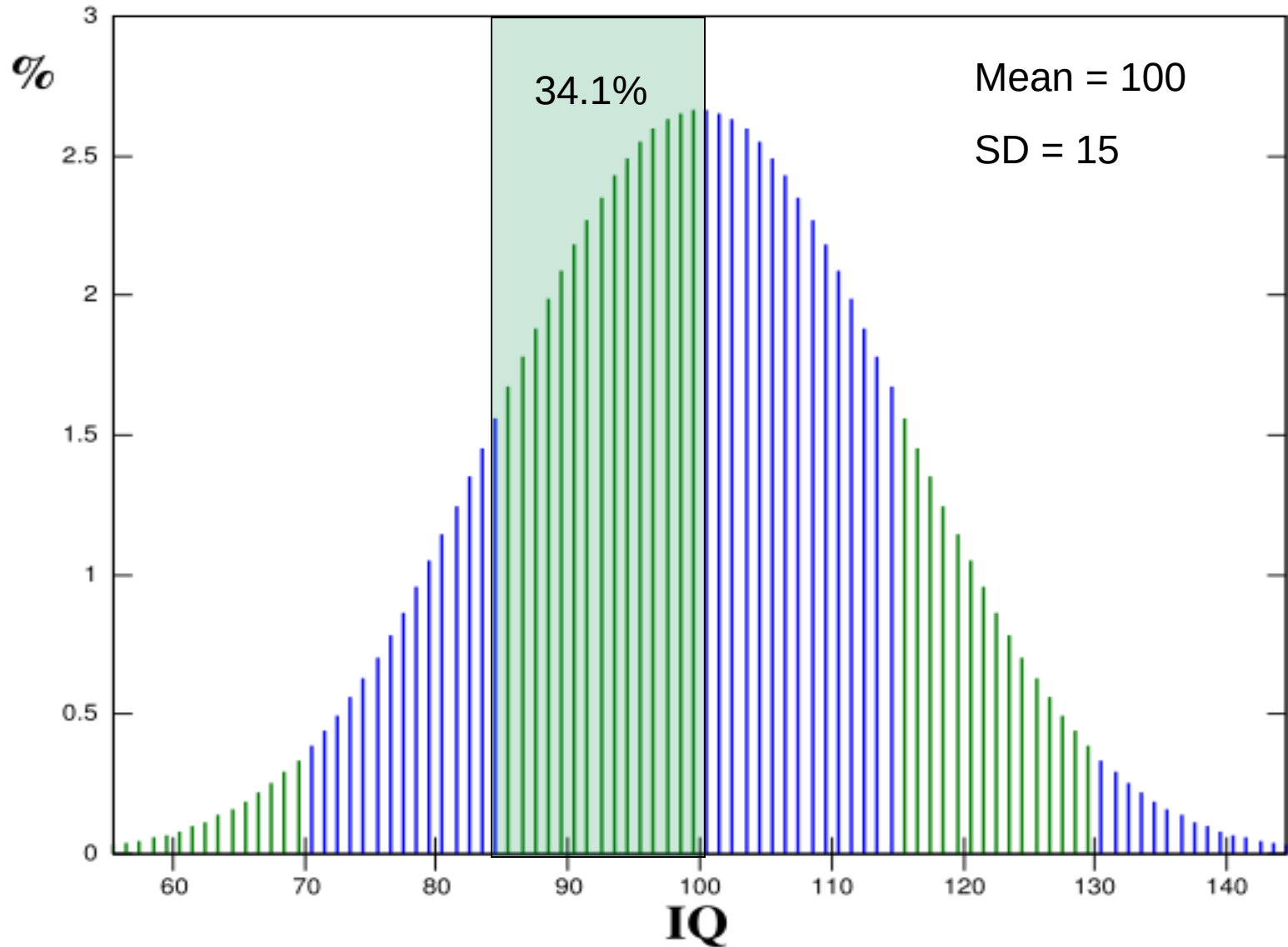


Normal Distribution



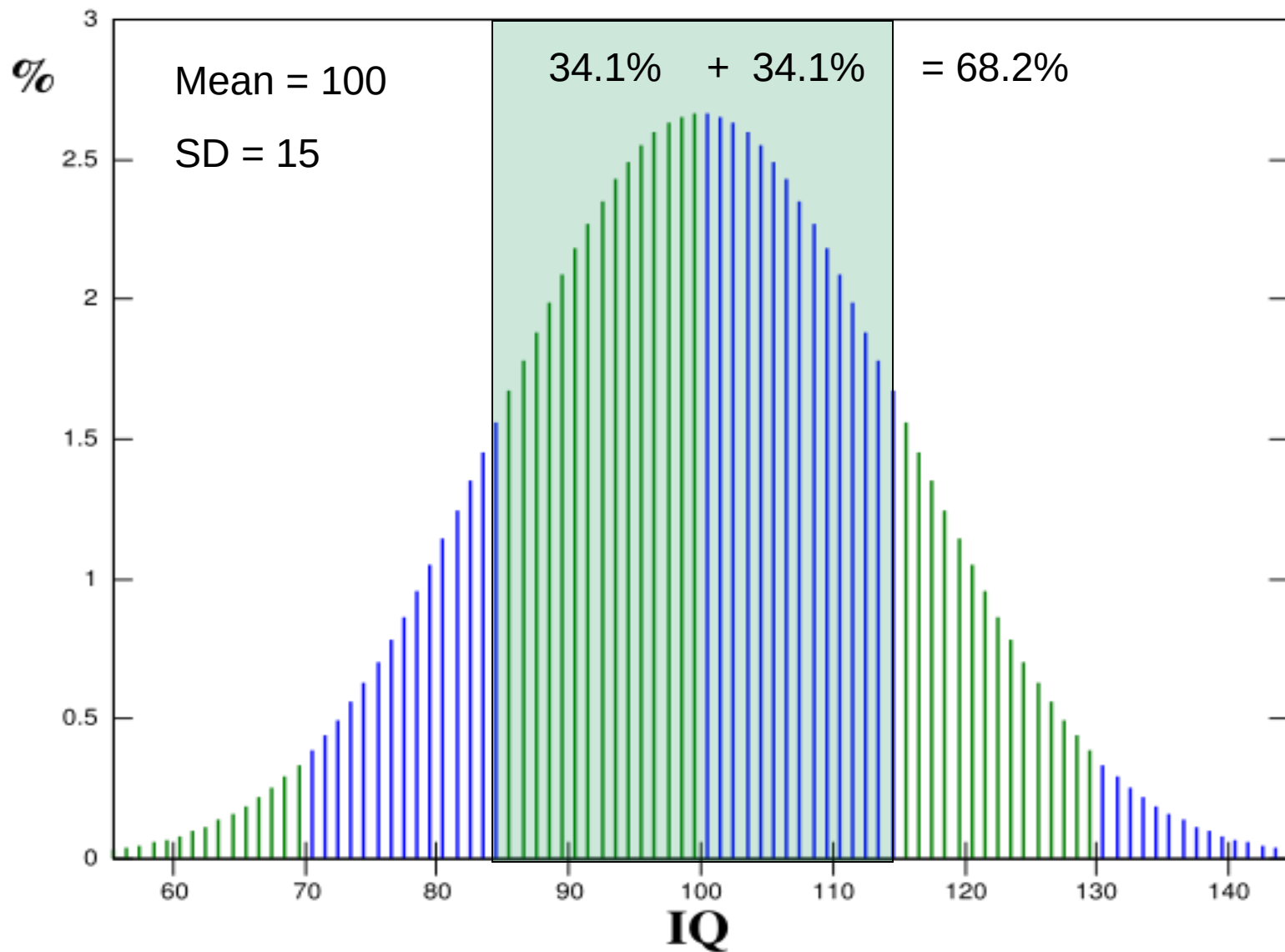


1. What percentage score between 85 and 100?



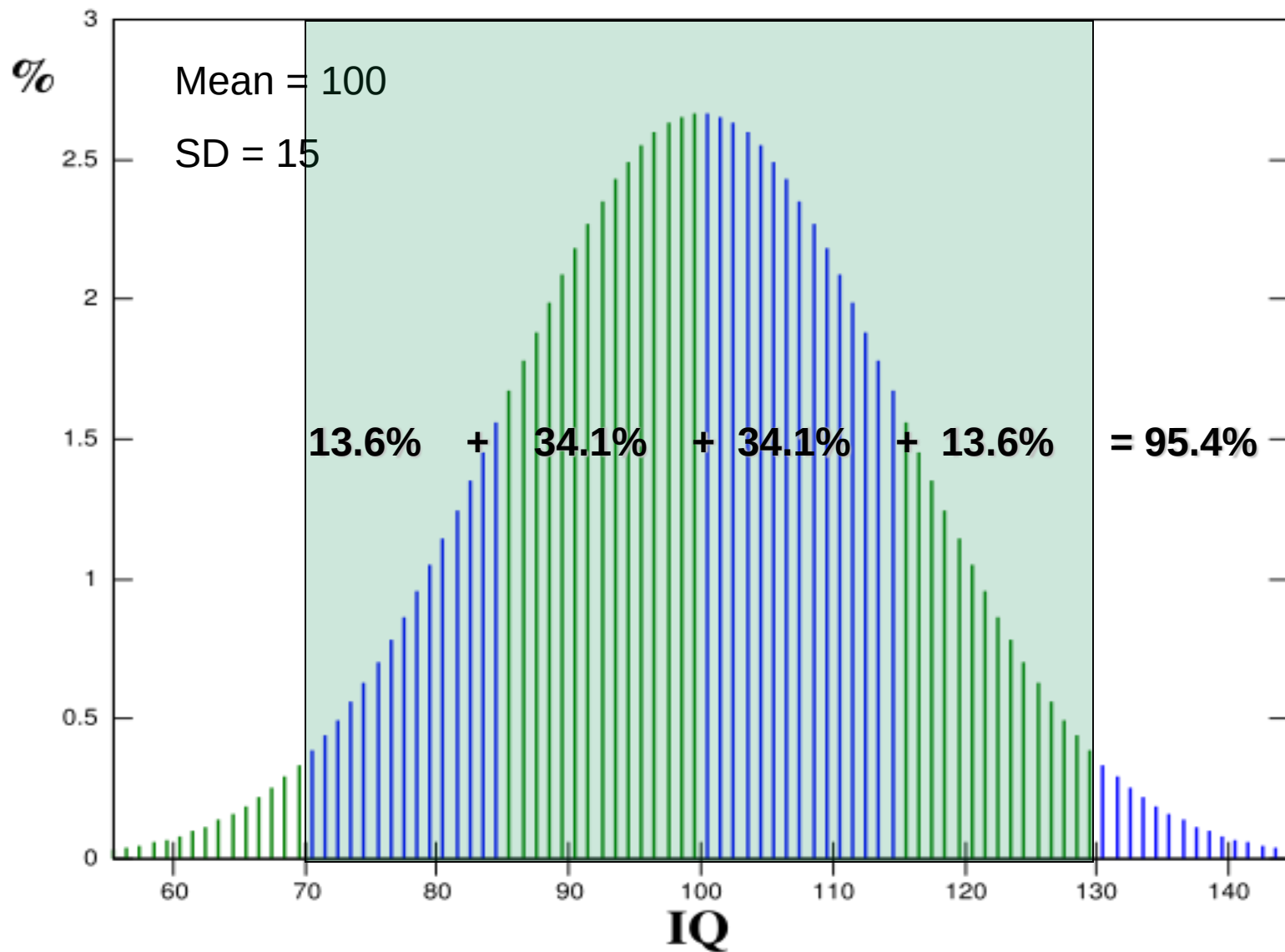


What percentage score between 85 and 115?



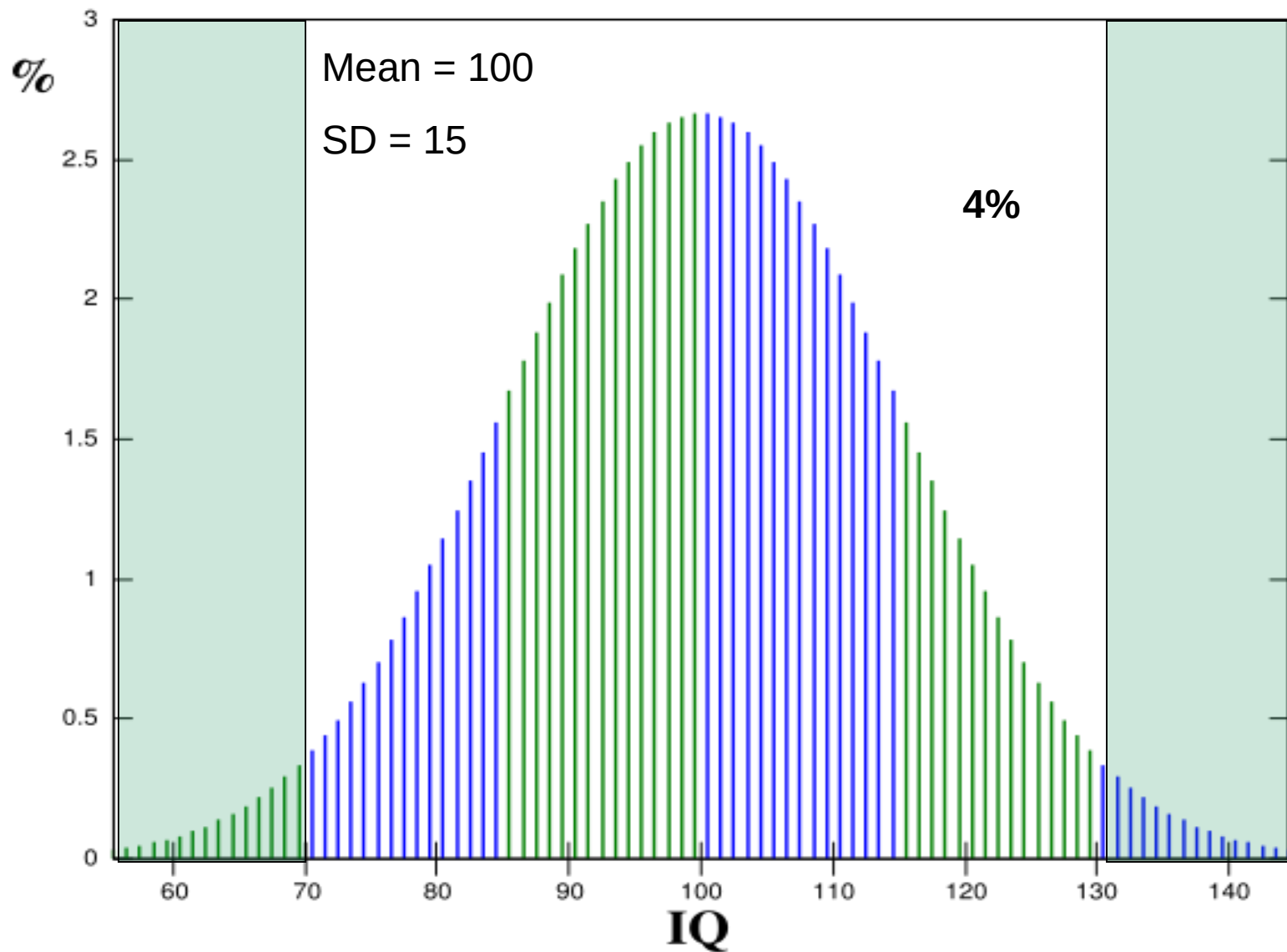


What percentage score between 70 and 130?





What percentage score below 70 and above 130?





F. Sample Questions

5. What percentage of people that will score 1 SD above the mean?

6. What percentage of people will score below 2 SD below the mean?



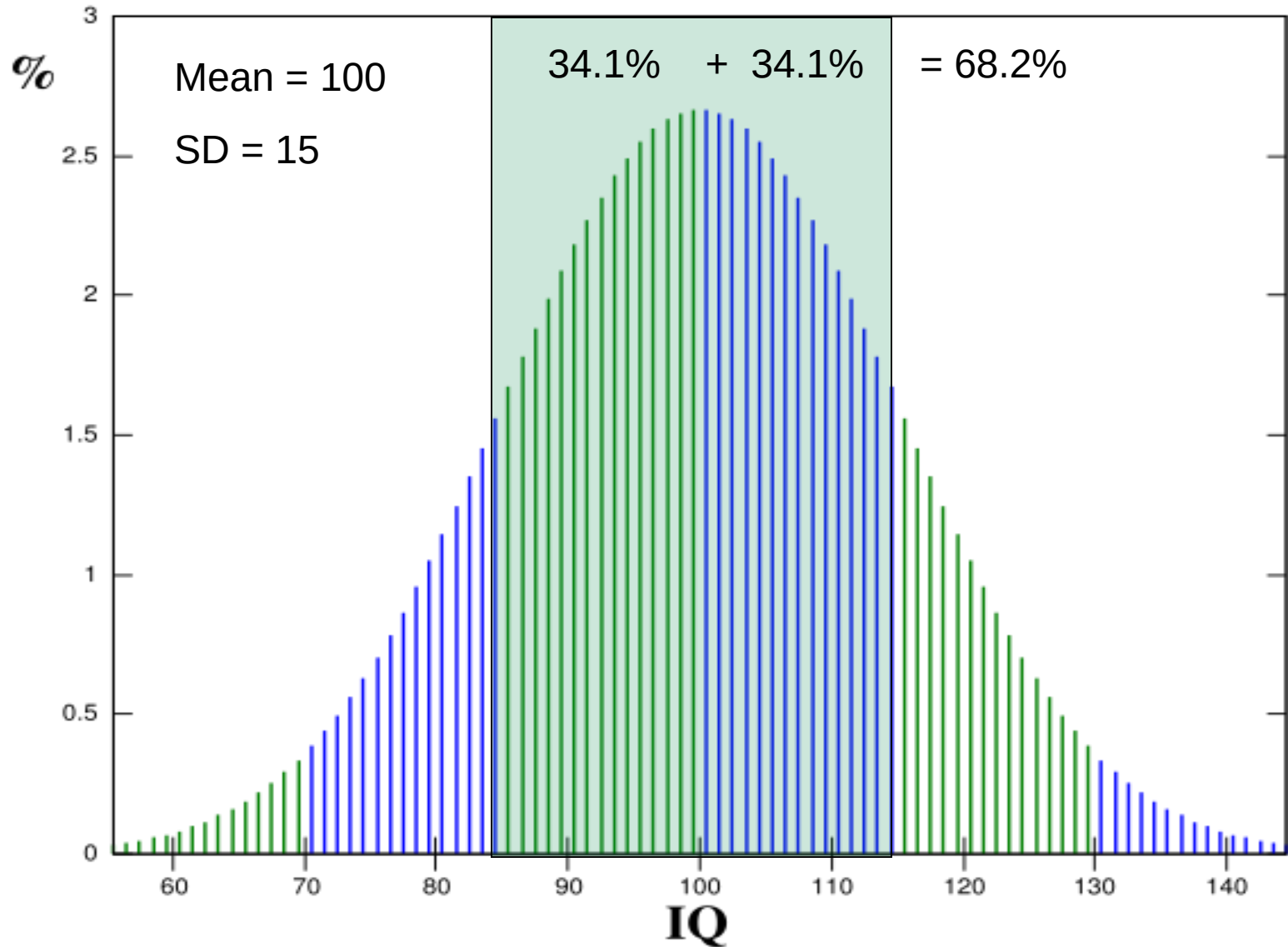
Curve Ball! Read carefully!!

1. Approximately what percentage of people will fall within one standard deviation of the mean?

- a) 34%
- b) 68%
- c) 95%
- d) 99%



Approximately what percentage of people will fall within one standard deviation of the mean? **68**





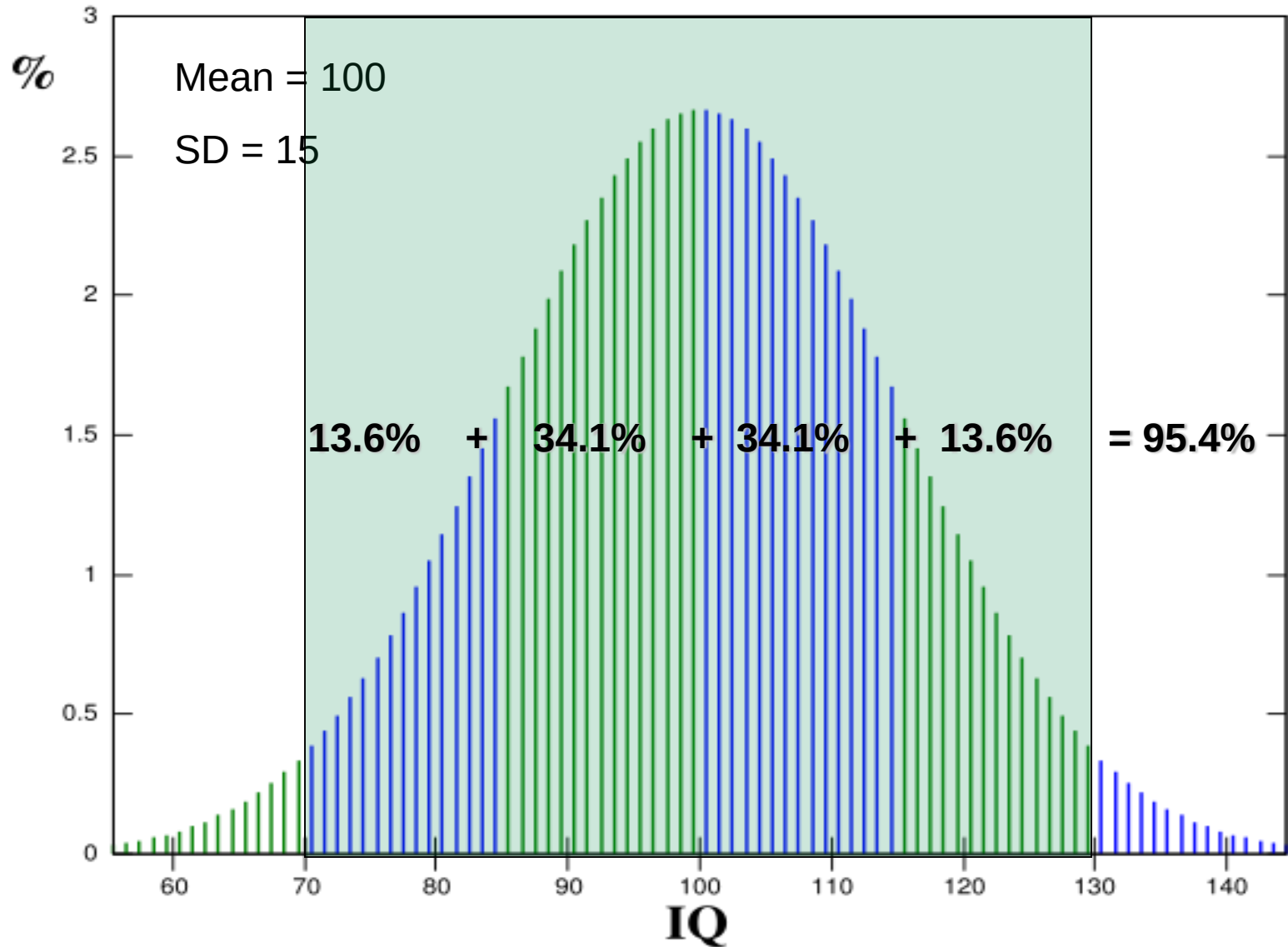
Read carefully!!

2. Approximately what percentage of people will fall within two standard deviation of the mean?

- a) 34%
- b) 68%
- c) 95%
- d) 99%



Approximately what percentage of people will fall within two standard deviation of the mean? **95**





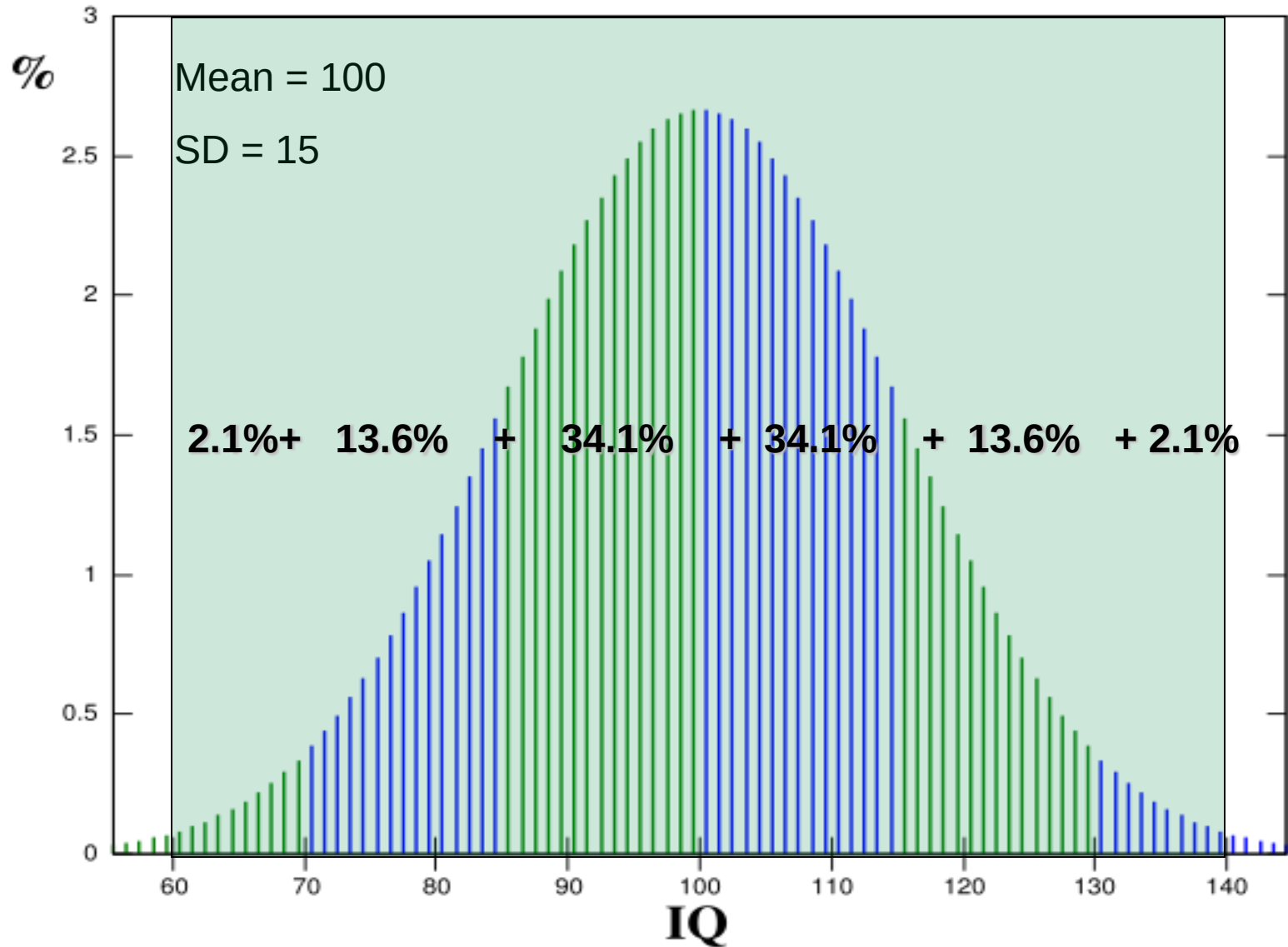
Read carefully!!

3. Approximately what percentage of people will fall within three standard deviation of the mean?

- a) 34%
- b) 68%
- c) 95%
- d) 99%

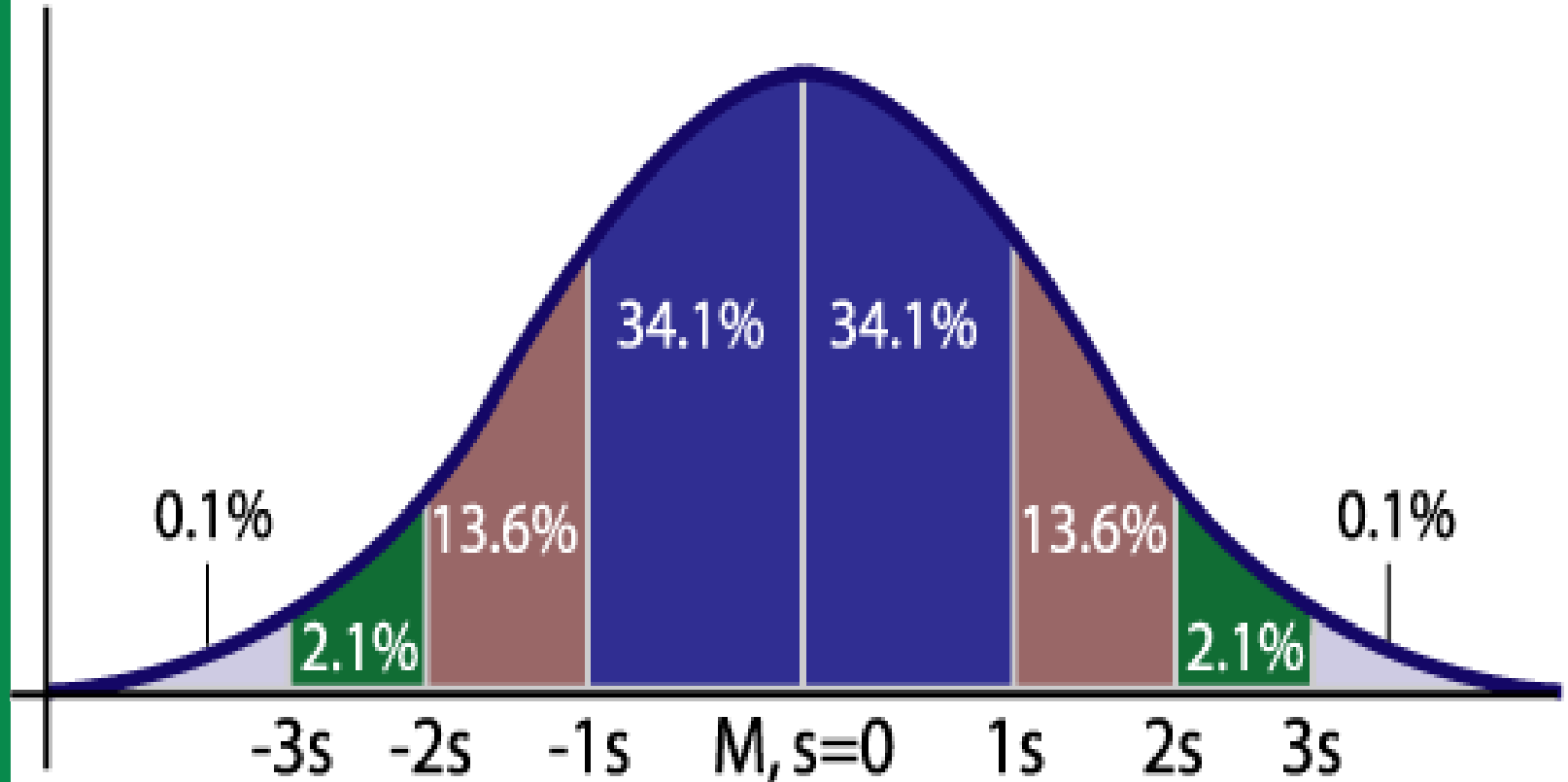


Approximately what percentage of people will fall within three standard deviation of the mean? **99**



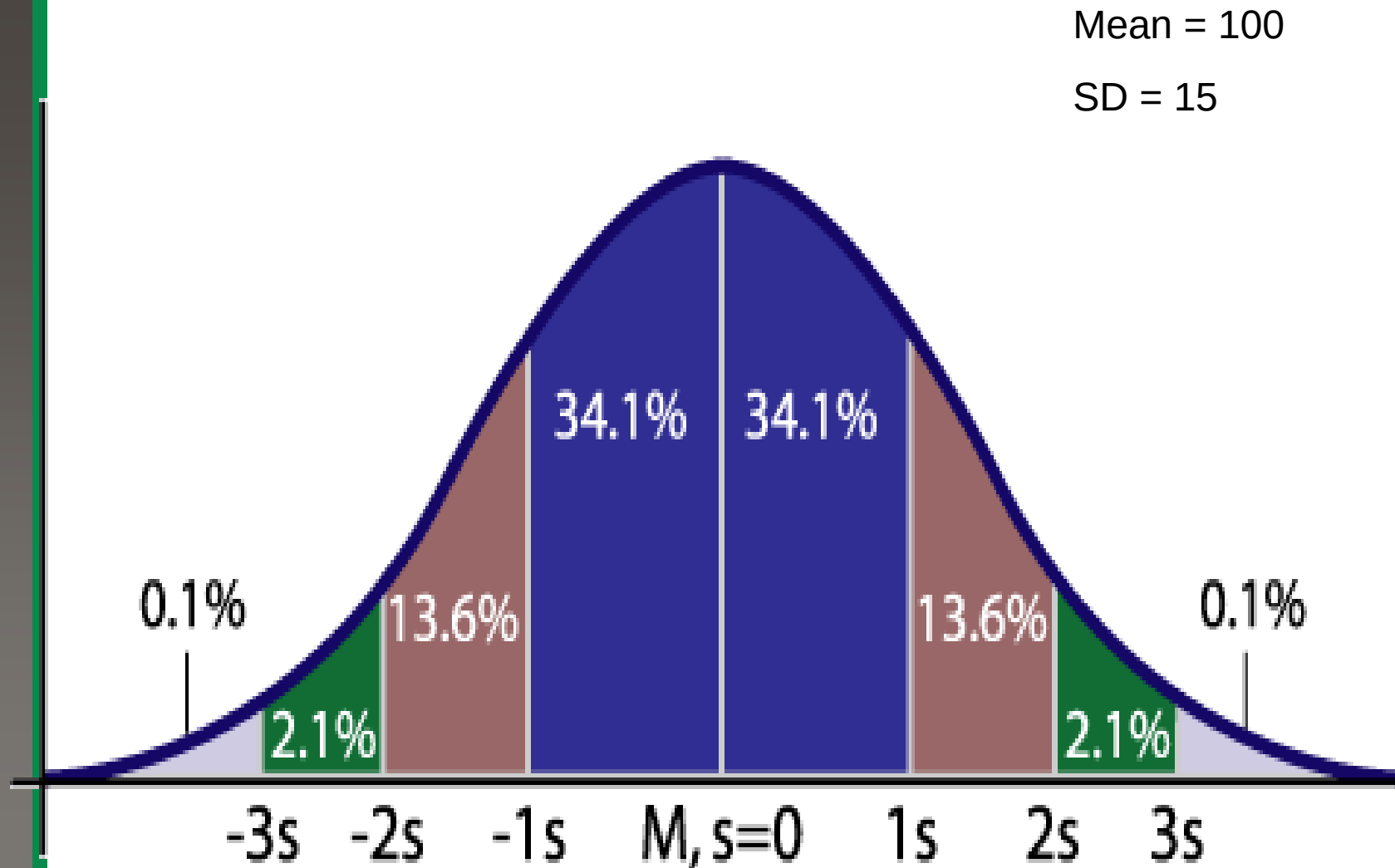


Normal Distribution





What percentage of people will score 1 SD above the mean?





Inferential Statistics

- Infer (draw-out / conclusion)
- You are trying to reach conclusions that extend beyond just describing the data.
- These are used to test hypothesis about samples.



T-test yields a p-value

- Generally, the t test gives a P value that allows us a measure of confidence in the observed difference.
- It allows us to say that the difference is real and not just by chance.
- A p value of less than 0.05 is a common criteria for significance.
 - 5% chance it was due to chance
- We call this **statistically significant** – results of the experiment did NOT happen by chance.

Graphical Representations

- Presenting numbers pictorially (usually in a graph) so they are easier to visualize
 - Subset of descriptive statistics
- Frequency Distribution: Table that divides an entire range of scores into a series of equal classes and then records the number of scores that fall into each class
- Histogram: Graph of a frequency distribution; scores are represented by vertical bars
- Frequency Polygon: Graph of a frequency distribution where the number of scores in each class is represented by points on a line



Frequency distribution

Pulse/min	Frequency	Cumulative frequency	Relative cumulative frequency (%)
60-64	2	2	3.17
65-69	7	9	14.29
70-74	11	20	31.75
75-79	15	35	55.56
80-84	10	45	71.43
85-89	9	54	85.71
90-94	6	60	95.24
95-99	3	63	100

frequency distribution table

A data table that lists a set of scores and their frequency.

score	tally	frequency (f)
1		4
2	 	9
3	 	6
4	 	7
5		3
6		2



Frequency Histogram

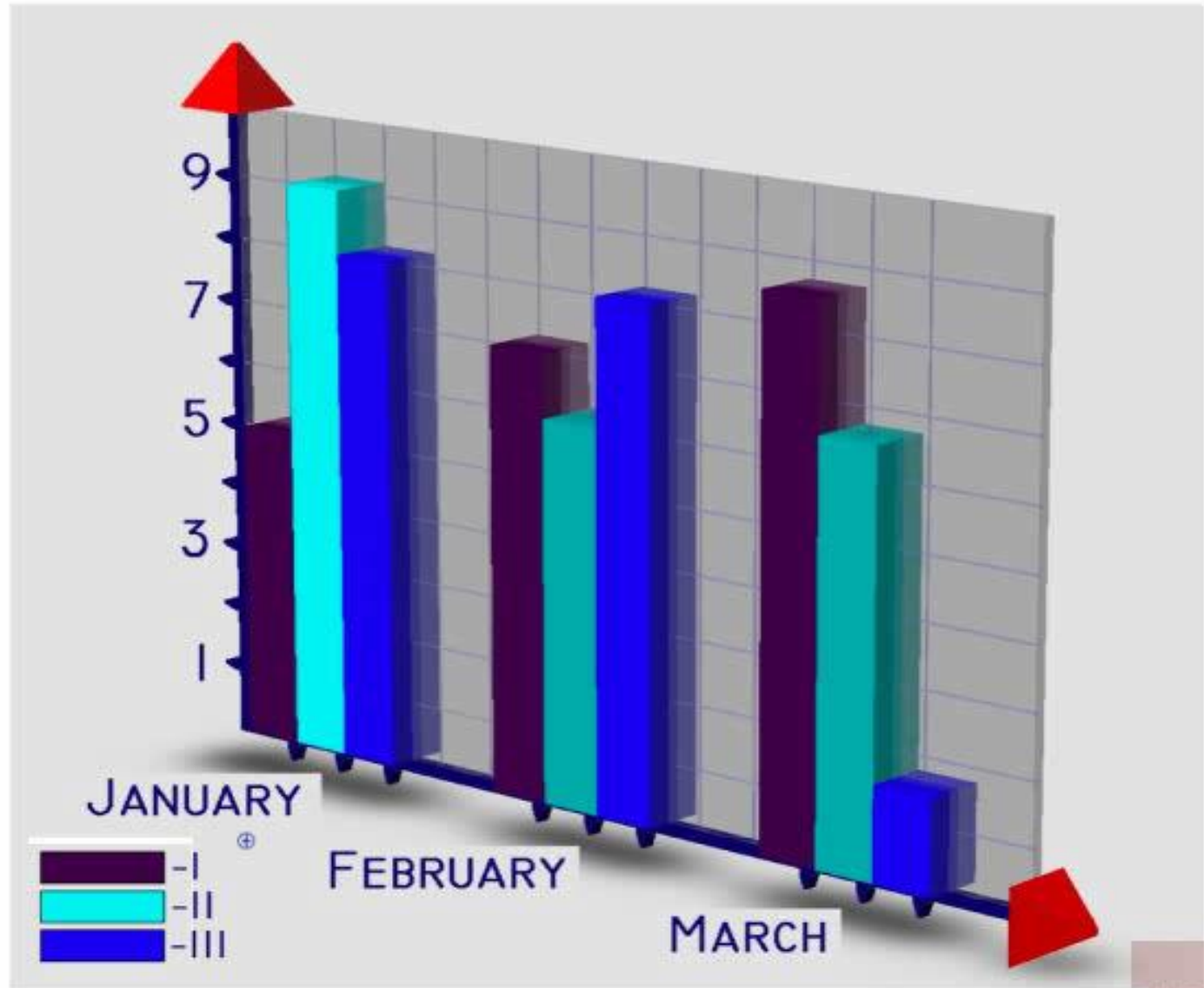
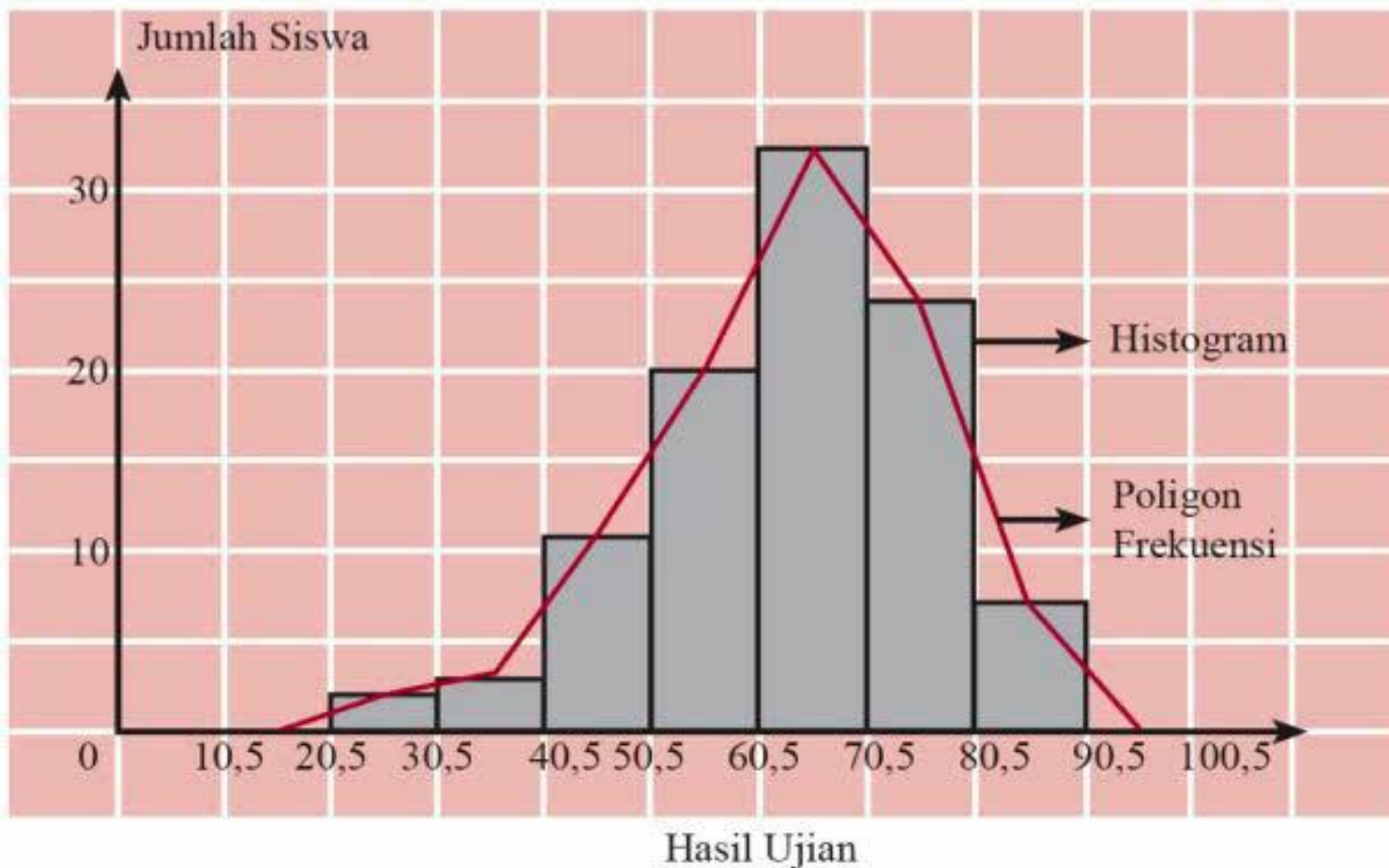


Image courtesy: https://upload.wikimedia.org/wikipedia/commons/c/c5/l8_3d_Histogram.PNG



Frequency Polygon



Dari histogram tersebut tampak bahwa kebanyakan siswa memperoleh nilai antara 60,5 dan 70,5. Coba Anda ceritakan hal lain dari histogram tersebut.