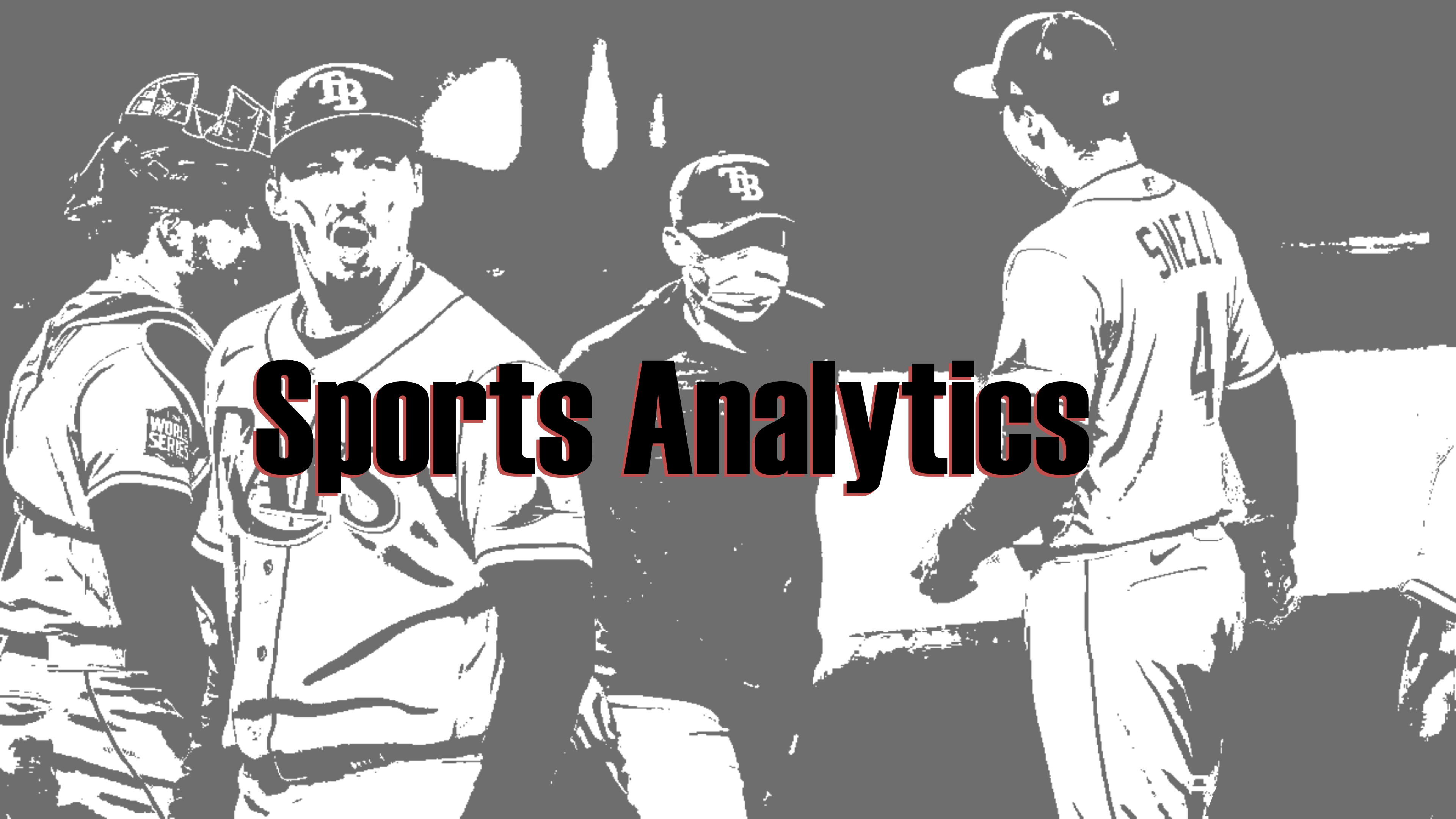
Sports Analytics



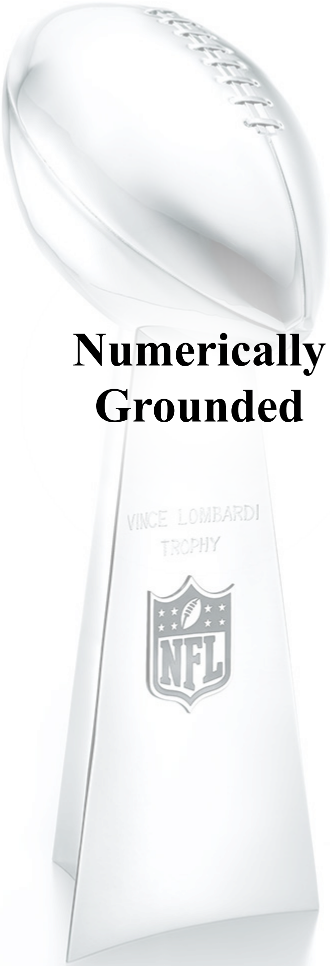

The GRANDSON of a Heisman Trophy winner playing against...A HEISMAN TROPHY WINNER!!! 😱😱😱
ESPN
NFL
ESTA NOCHE

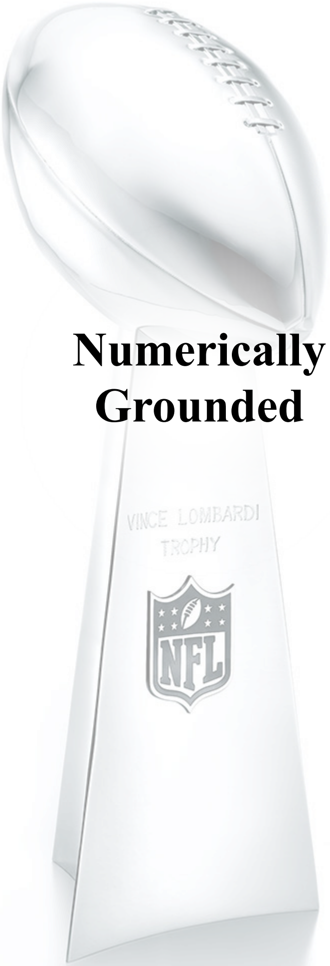

VINCE LOMBARDI
TROPHY
NFL


NBA

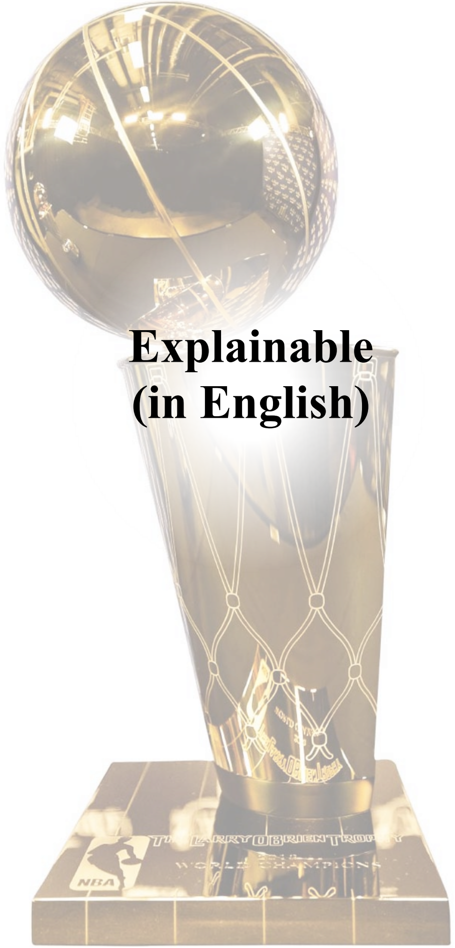

WORLD SERIES CHAMPIONS
NEW YORK YANKEES
1977
SEATTLE MARINERS
OAKLAND ATHLETICS
NEW YORK YANKEES
Minnesota TWINS
MILWAUKEE BREWERS
Kansas City ROYALS
DETROIT TIGERS
Cleveland INDIANS
San Diego PADRES
NEW YORK METS
Los Angeles DODGERS
HOUSTON ASTROS
Cincinnati REDS
CHICAGO CUBS
ATLANTA BRAVES
BALTIMORE ORIOLES

Numerically Grounded





Numerically Grounded



Explainable (in English)

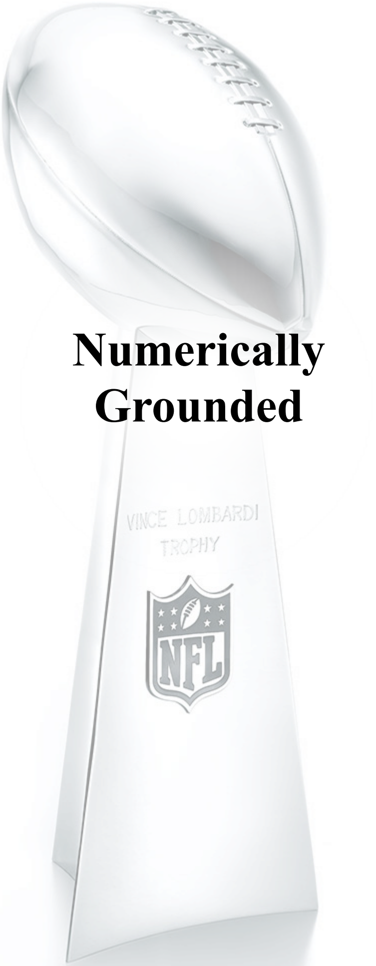

SPALDING
NBA

Jan Hatzius, chief economist at Goldman Sachs, November 15, 2007:

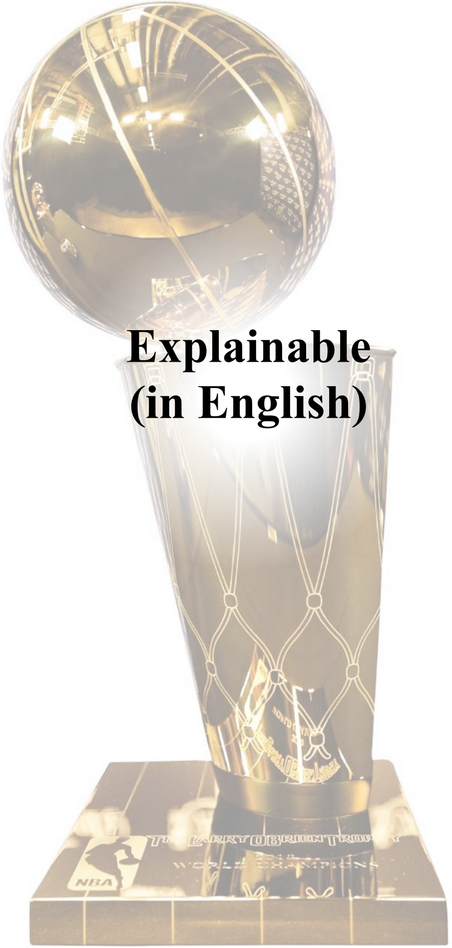
The likely mortgage credit losses pose a significantly bigger macroeconomic risk than generally recognized....The macroeconomic consequence could be quite dramatic. If leveraged investors see $200 billion in aggregate credit loss, they might need to scale back their lending by $2 trillion. This is a large shock....Is is easy to see how such a shock could produce a substantial recession or a long period of very sluggish growth.

ECRI, economic forecasting firm, September, 2011:

ECRI’s recession call isn’t based on just one or two leading indexes, but on dozens of specialized leading indexes, including the U.S. Long Leading Index....to be followed by downturns in the Weekly Leading Index and other shorter-leading indexes. In fact, the most reliable forward-looking indicators are now collectively behaving as they did on the cusp of full-blown recessions.



Numerically Grounded



Explainable (in English)

Actionable

More information ***should*** improve predictive accuracy.

More information tends to increase predictive confidence faster than it increases predictive accuracy.

People tend to be biased towards their most recent observations.

Rare events can be difficult to quantify with sampling data.

Human bias can both influence and outstrip sampling errors.

Crazy
Rich
Bayesians

Bayes Theorem:

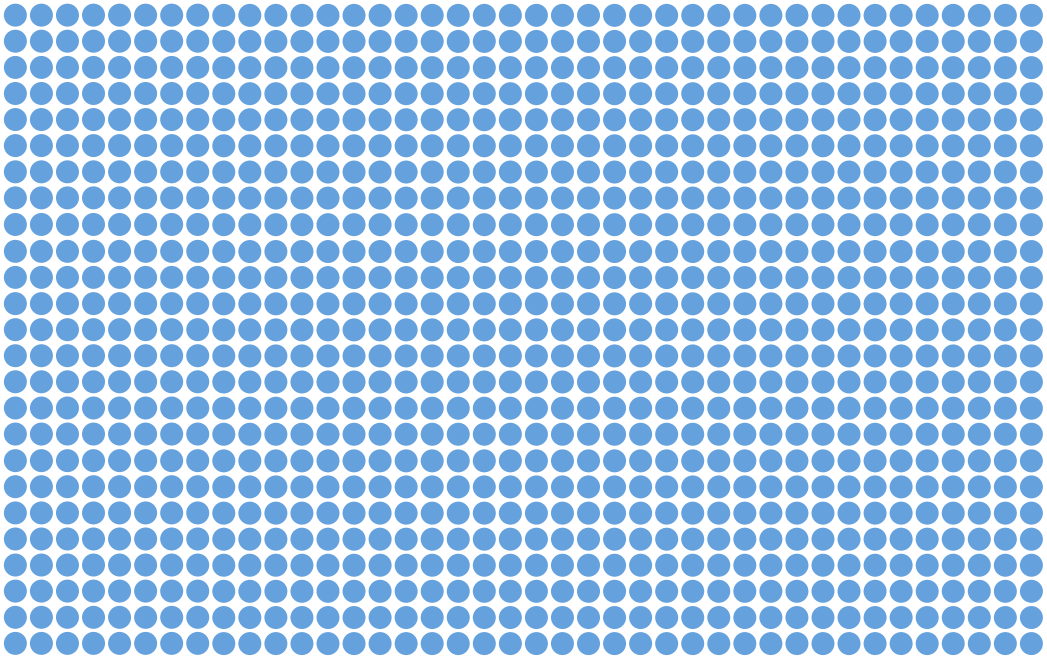
$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})\, P(\mathrm{A})}{P(\mathrm{B})}$$

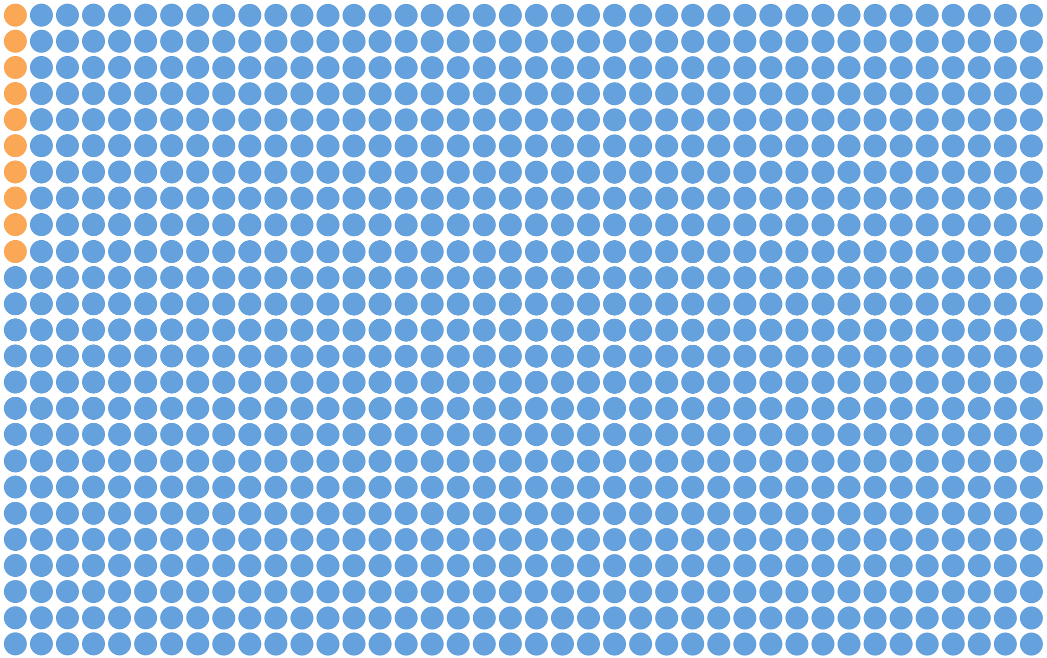
The probability that a woman at age 40 has breast cancer is 1%. According to the literature, the probability that the disease is detected by a mammography is 80%. The probability that the test misdetects the disease although the patient does not have it is 9.6%.

If a woman at age 40 is tested as positive, what is the probability that she indeed has breast cancer?

The probability that a woman at age 40 has breast cancer is 1%. According to the literature, the probability that the disease is detected by a mammography is 80%. The probability that the test misdetects the disease although the patient does not have it is 9.6%.

If a woman at age 40 is tested as positive, what is the probability that she indeed has breast cancer?



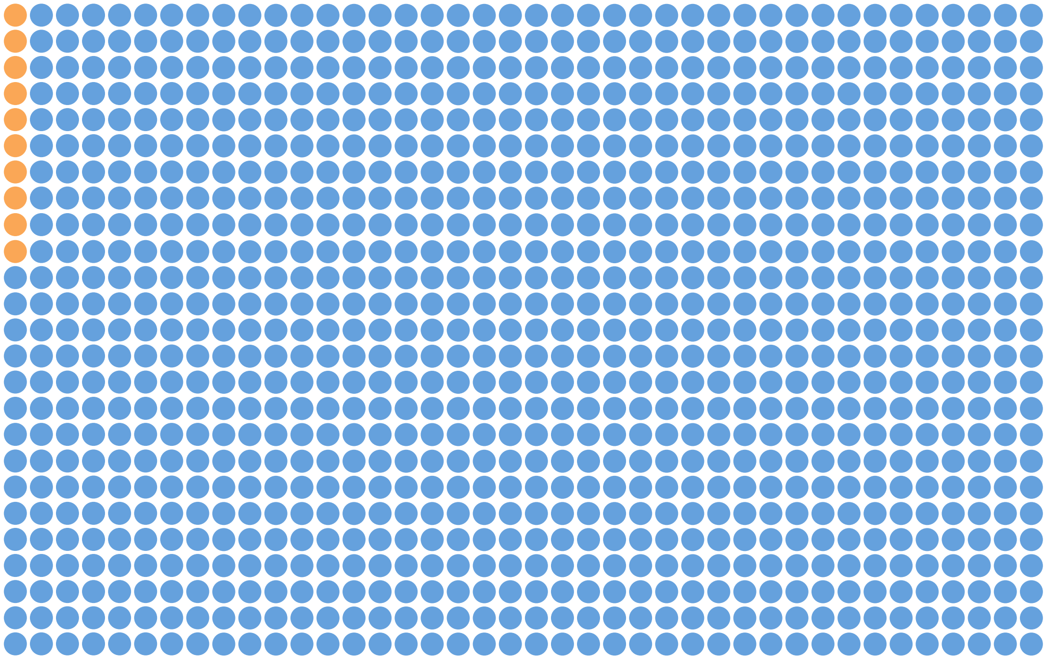


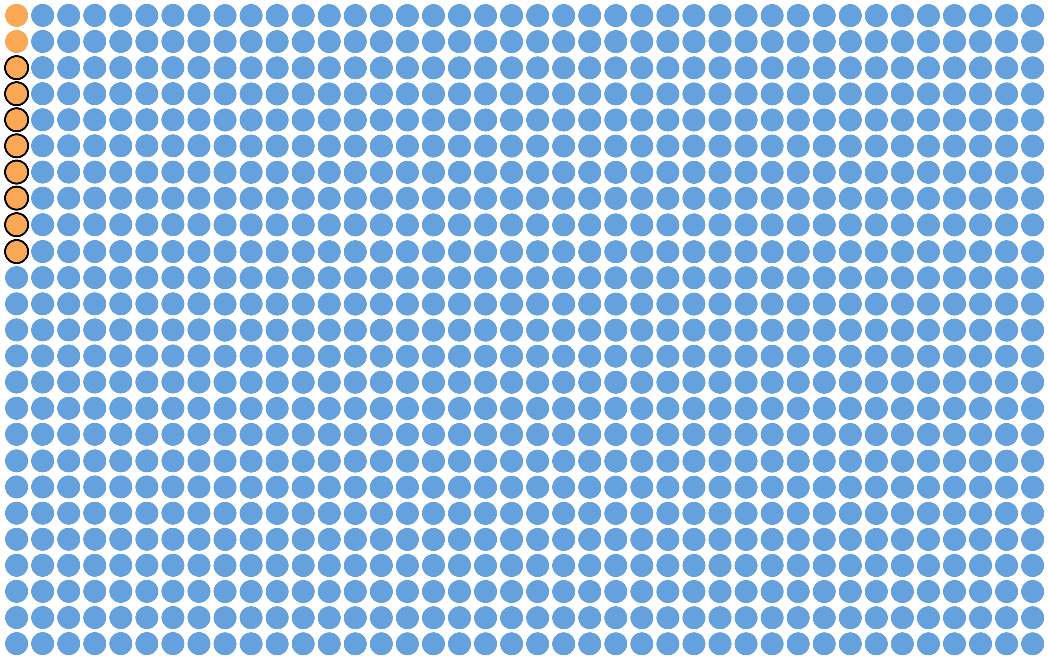
The probability that a woman at age 40 has breast cancer is 1%. According to the literature, the probability that the disease is detected by a mammography is 80%. The probability that the test misdetects the disease although the patient does not have it is 9.6%.

If a woman at age 40 is tested as positive, what is the probability that she indeed has breast cancer?

The probability that a woman at age 40 has breast cancer is 1%. According to the literature, the probability that the disease is detected by a mammography is 80%. The probability that the test misdetects the disease although the patient does not have it is 9.6%.

If a woman at age 40 is tested as positive, what is the probability that she indeed has breast cancer?



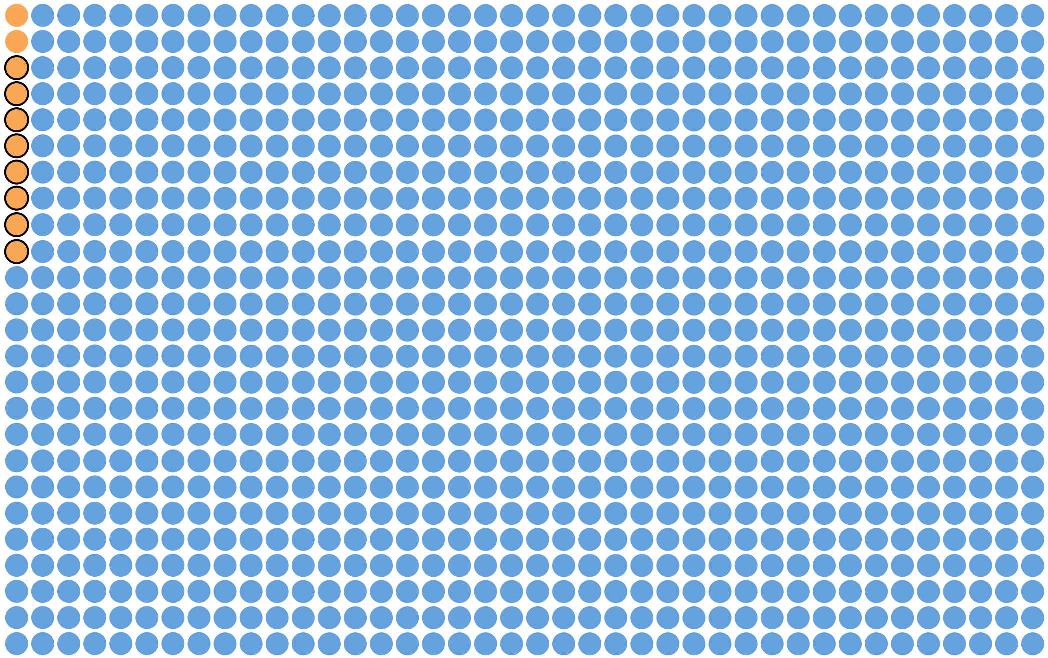


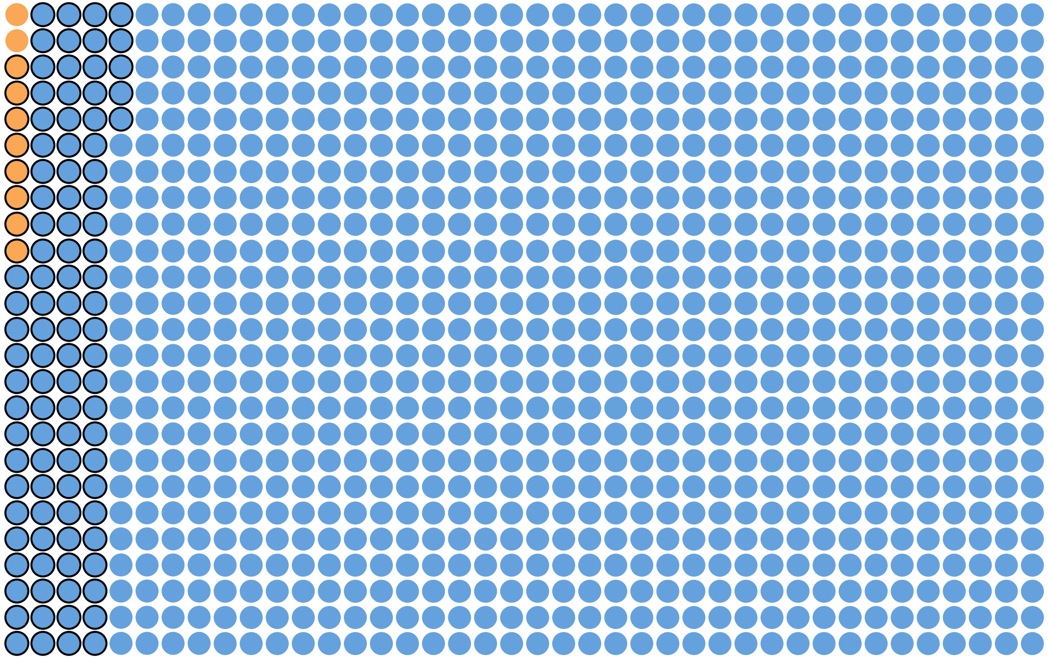
The probability that a woman at age 40 has breast cancer is 1%. According to the literature, the probability that the disease is detected by a mammography is 80%. The probability that the test misdetects the disease although the patient does not have it is 9.6%.

If a woman at age 40 is tested as positive, what is the probability that she indeed has breast cancer?

The probability that a woman at age 40 has breast cancer is 1%. According to the literature, the probability that the disease is detected by a mammography is 80%. The probability that the test misdetects the disease although the patient does not have it is 9.6%.

If a woman at age 40 is tested as positive, what is the probability that she indeed has breast cancer?



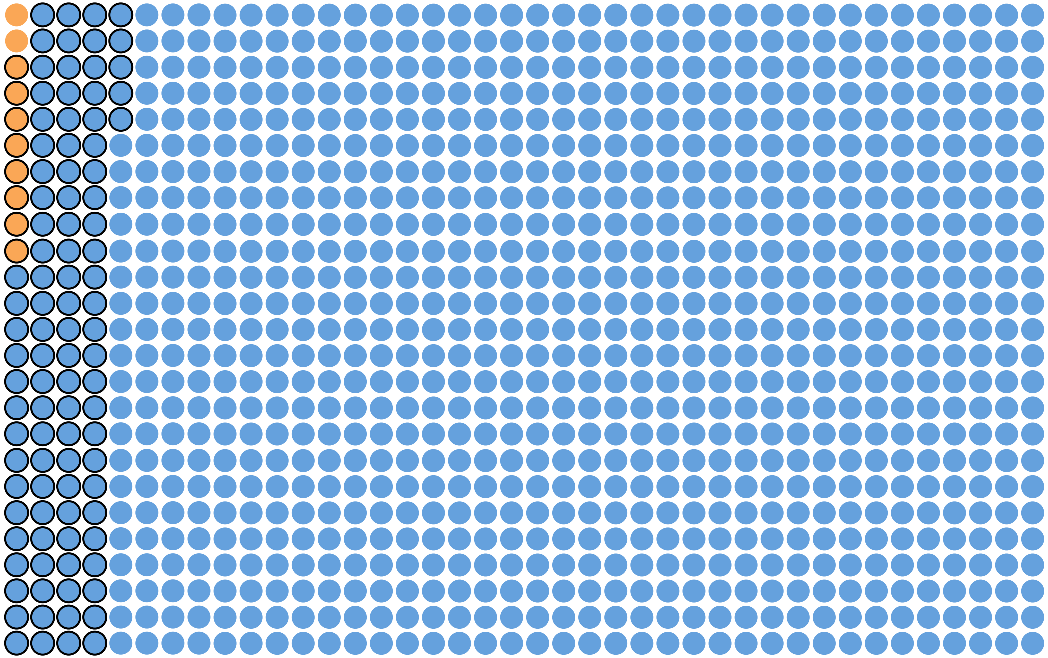


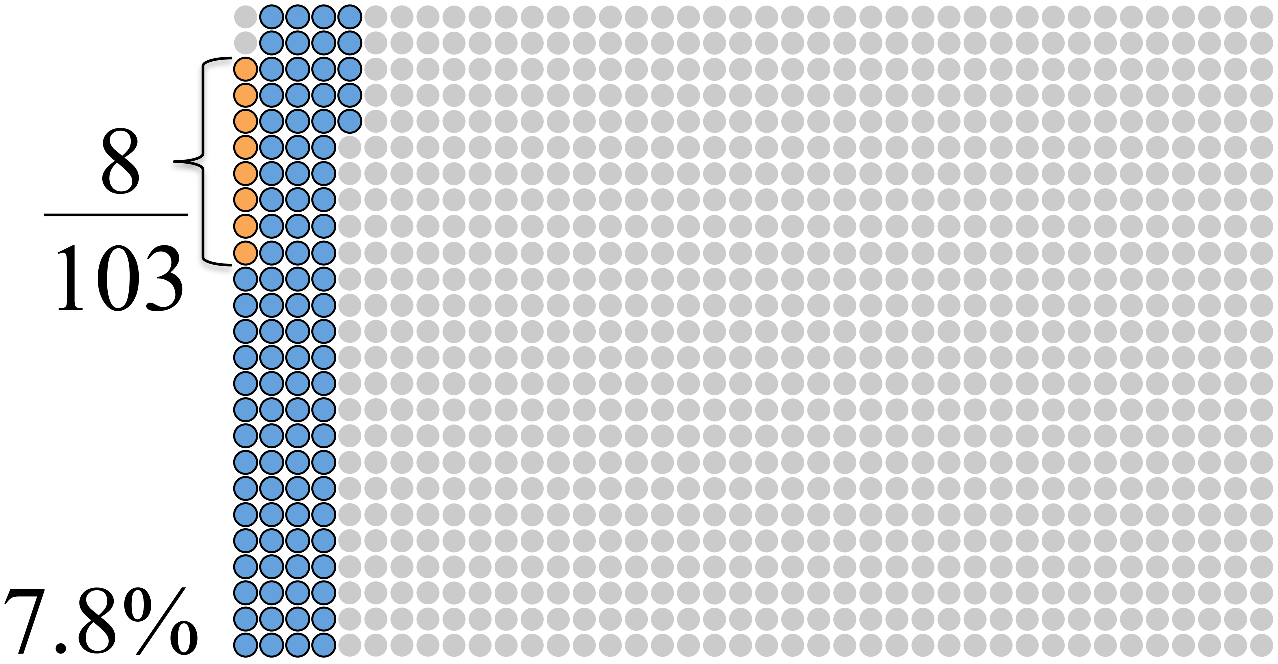
The probability that a woman at age 40 has breast cancer is 1%. According to the literature, the probability that the disease is detected by a mammography is 80%. The probability that the test misdetects the disease although the patient does not have it is 9.6%.

If a woman at age 40 is tested as positive, what is the probability that she indeed has breast cancer?

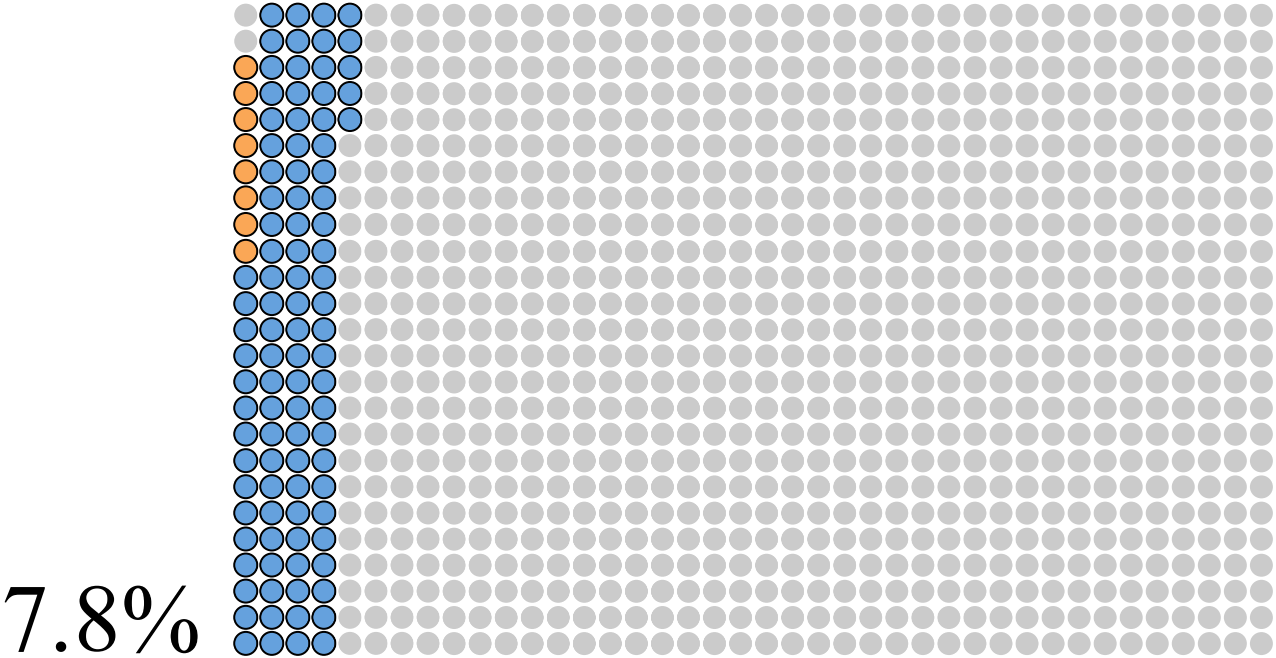
The probability that a woman at age 40 has breast cancer is 1%. According to the literature, the probability that the disease is detected by a mammography is 80%. The probability that the test misdetects the disease although the patient does not have it is 9.6%.

If a woman at age 40 is tested as positive, what is the probability that she indeed has breast cancer?




8
103
7.8%

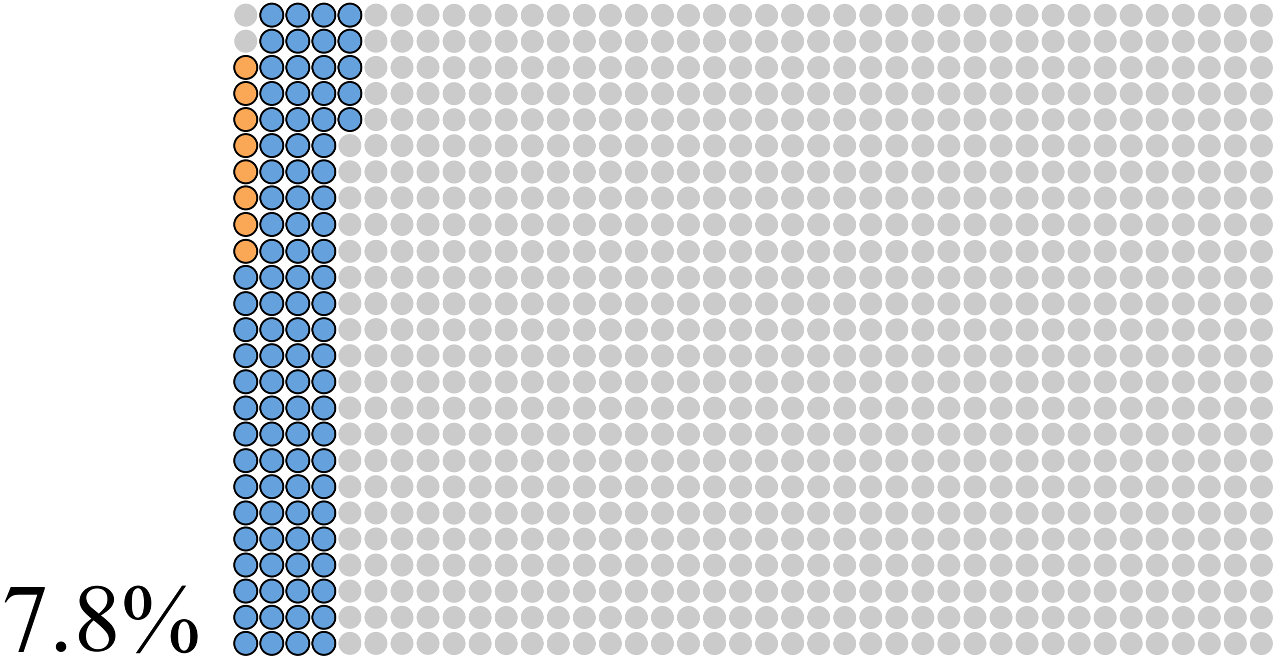
$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})\, P(\mathrm{A})}{P(\mathrm{B})}$$



A = Patient has breast cancer

B = Patient has positive mammography

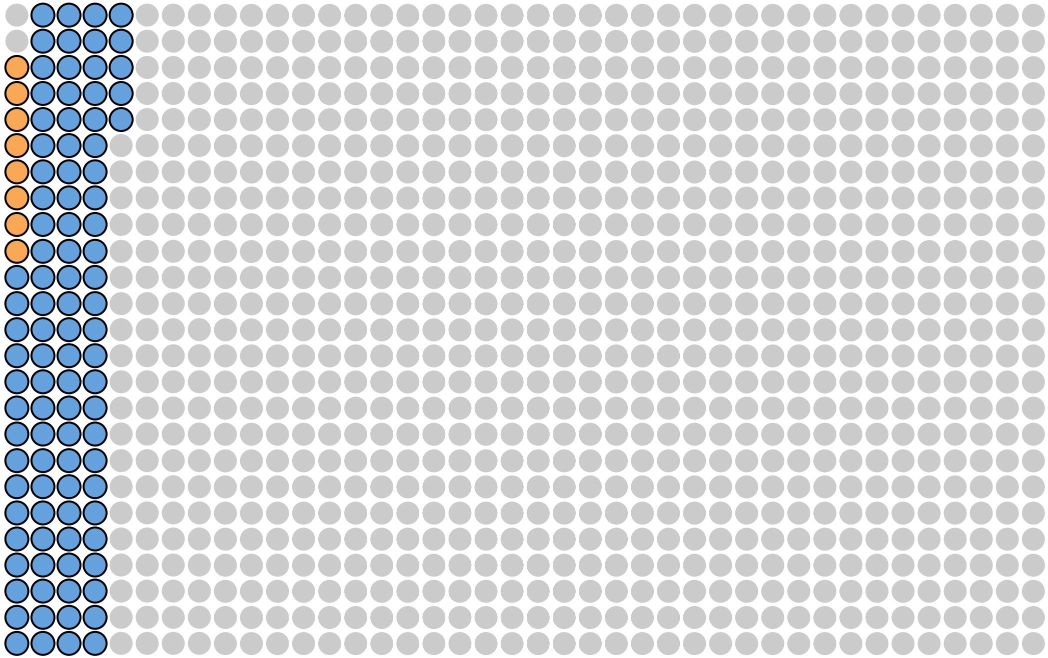
$$= \frac{P(\mathrm{B} \mid \mathrm{A})\, P(\mathrm{A})}{P(\mathrm{B})}$$



A = Patient has breast cancer

B = Patient has positive mammography

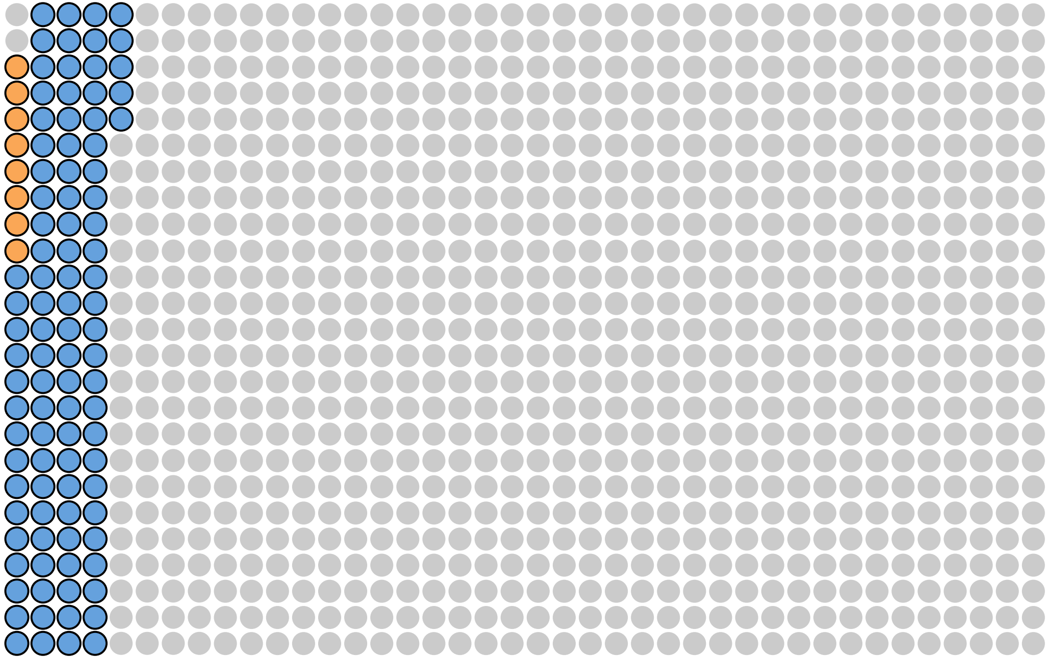
$$7.8\% = \frac{P(\mathrm{B} \mid \mathrm{A})\, P(\mathrm{A})}{P(\mathrm{B})} = \frac{}{\frac{103}{1000}}$$



A = Patient has breast cancer

B = Patient has positive mammography

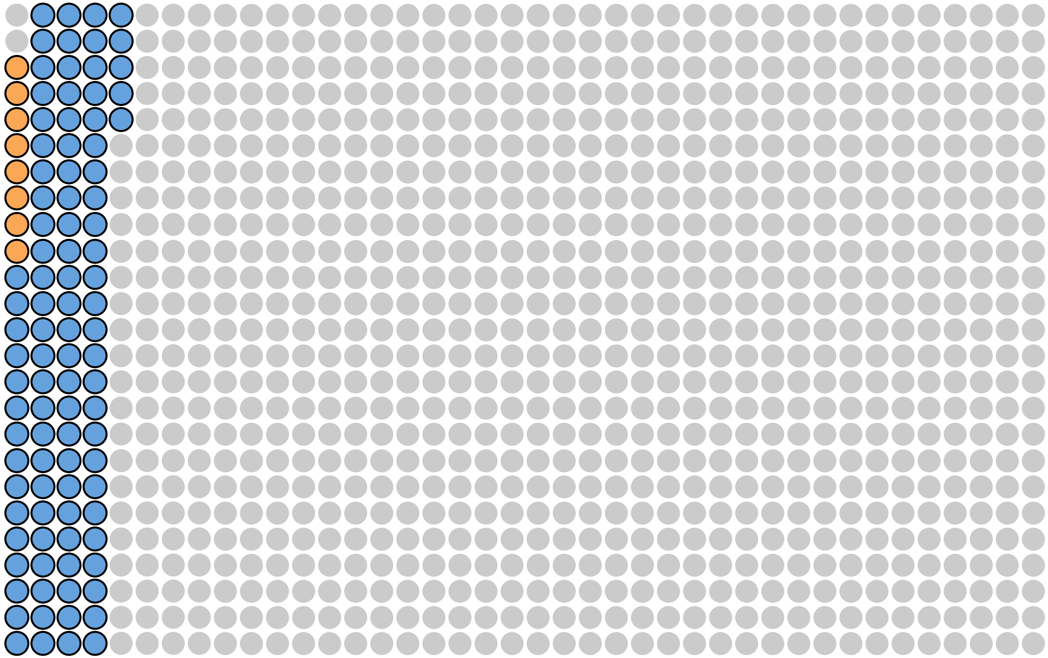
$$7.8\% = \frac{P(\mathrm{B} \mid \mathrm{A})\, P(\mathrm{A})}{P(\mathrm{B})} = \frac{}{\frac{103}{1000}}$$



A = Patient has breast cancer

B = Patient has positive mammography

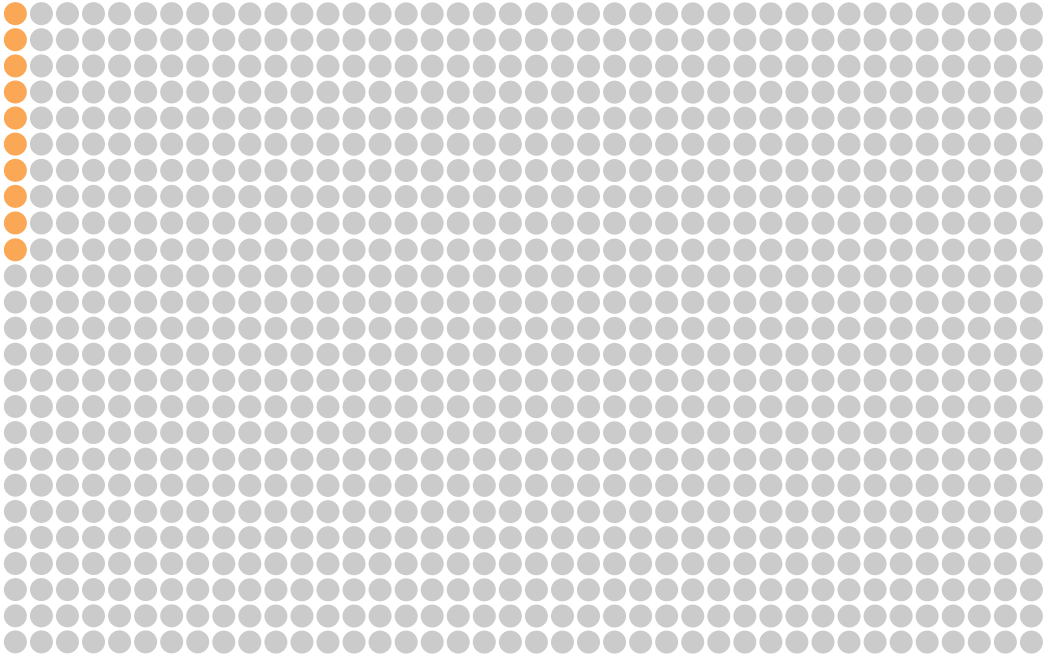
$$7.8\% = \frac{P(\mathrm{B} \mid \mathrm{A})\, P(\mathrm{A})}{P(\mathrm{B})} = \frac{\quad}{\frac{103}{1000}}$$



A = Patient has breast cancer

B = Patient has positive mammography

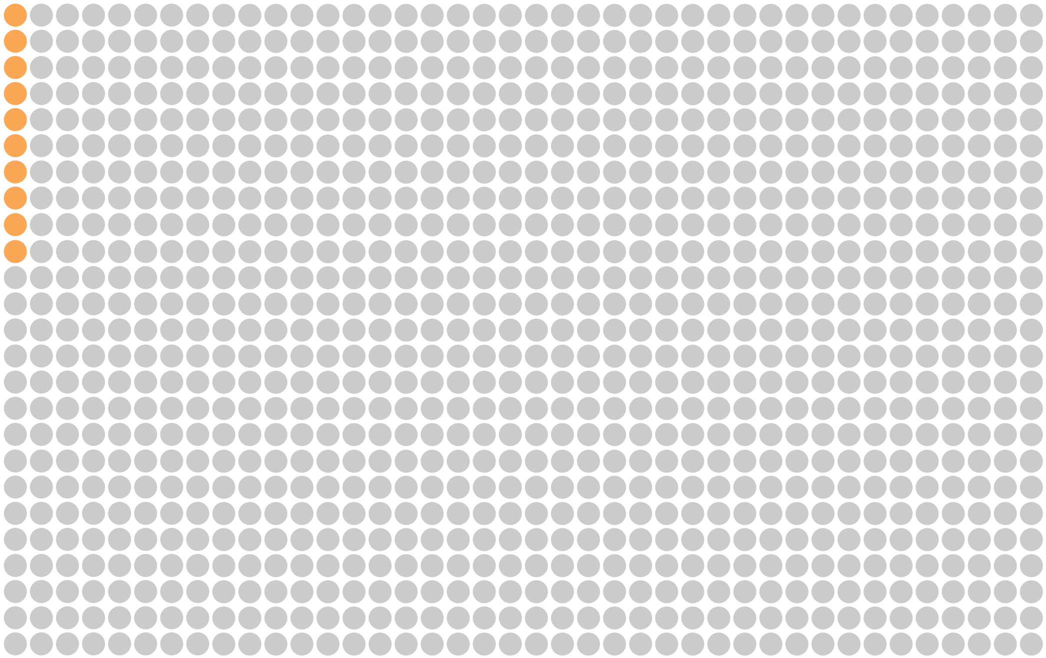
$$7.8\% = \frac{P(B \mid A)\, P(A)}{P(B)} = \frac{\frac{10}{\cancel{1000}}}{\frac{103}{\cancel{1000}}}$$



A = Patient has breast cancer

B = Patient has positive mammography

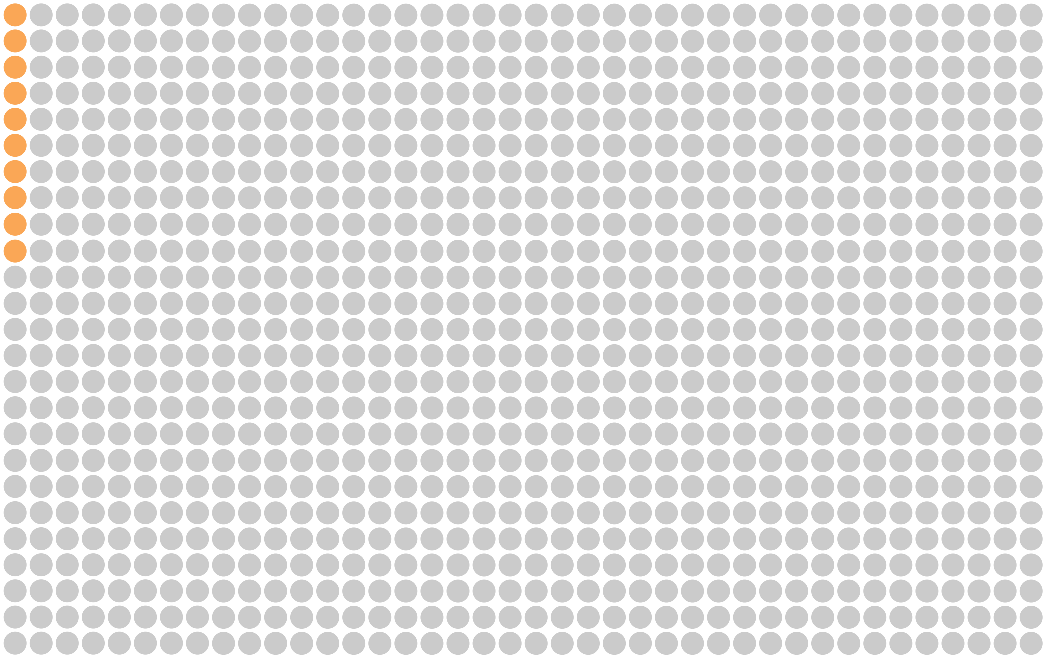
$$7.8\% = \frac{P(\mathrm{B} \mid \mathrm{A})\, P(\mathrm{A})}{P(\mathrm{B})} = \frac{10}{103}$$



A = Patient has breast cancer

B = Patient has positive mammography

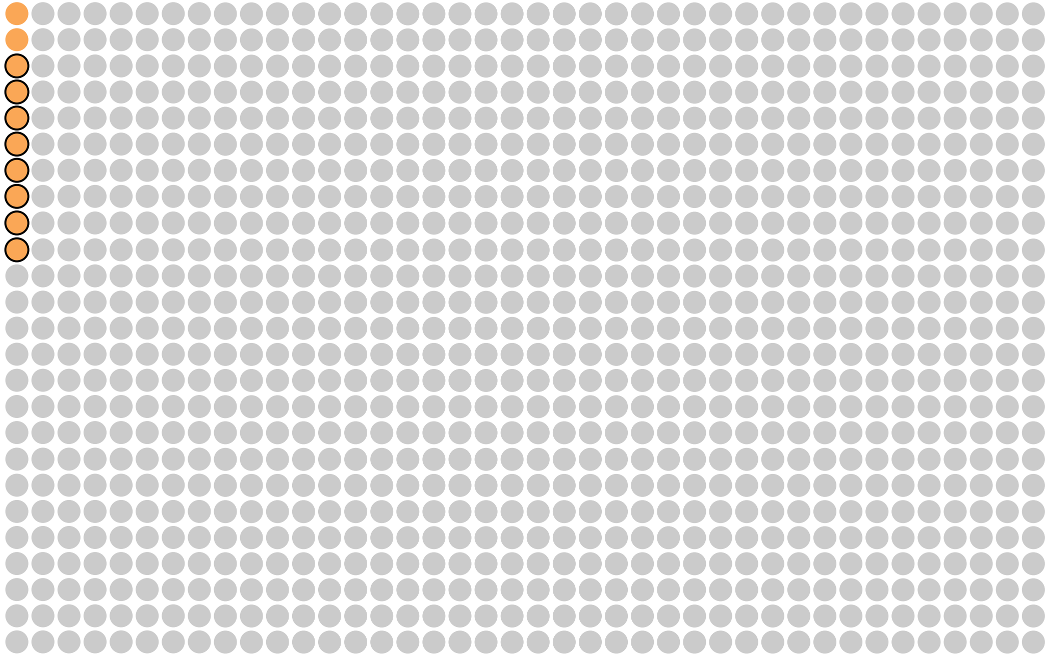
$$7.8\% = \frac{P(B \mid A)\, P(A)}{P(B)} = \frac{10}{103}$$



A = Patient has breast cancer

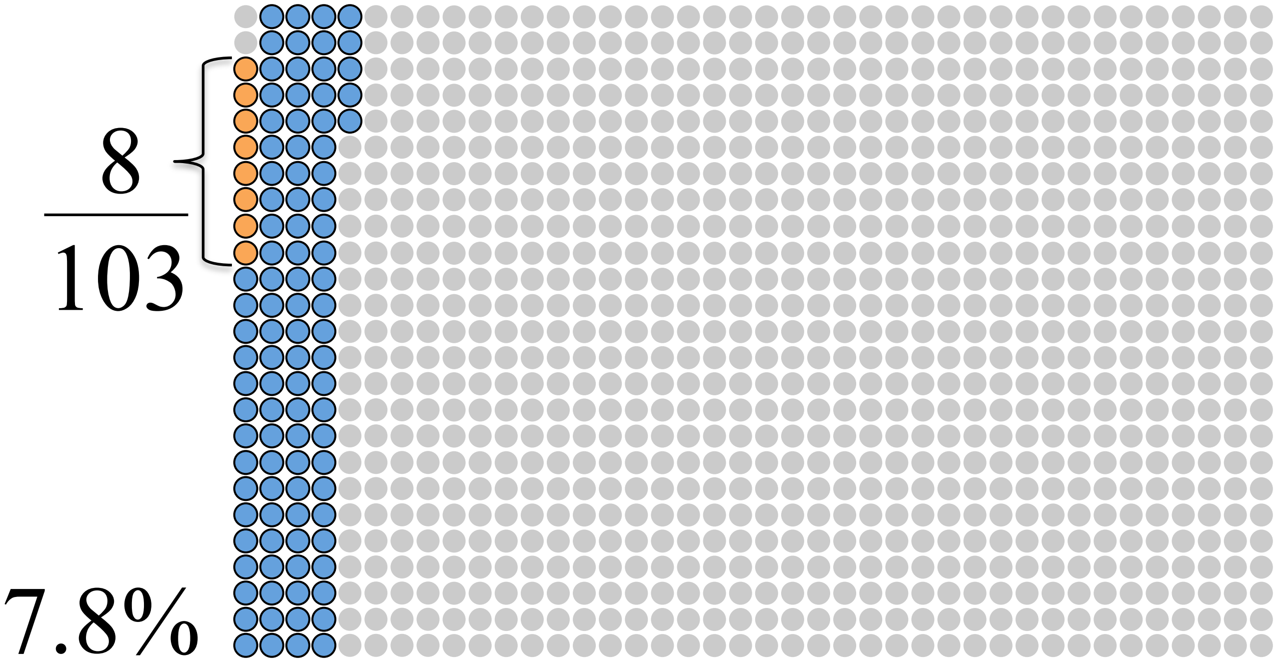
B = Patient has positive mammography

$$7.8\% = \frac{P(B \mid A)\, P(A)}{P(B)} = \frac{8}{\cancel{10}} \times \frac{\cancel{10}}{103}$$



A = Patient has breast cancer

B = Patient has positive mammography

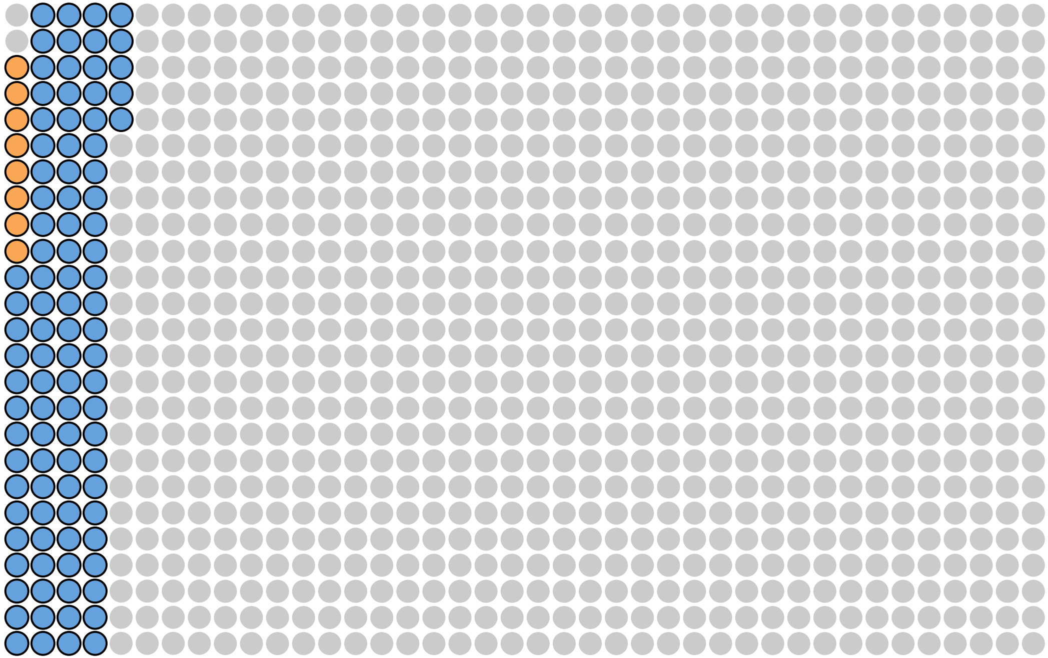


A = Patient has breast cancer

B = Patient has positive mammography

$$\frac{8}{103} = \frac{\text{true positives}}{\text{all positives}}$$

7.8%

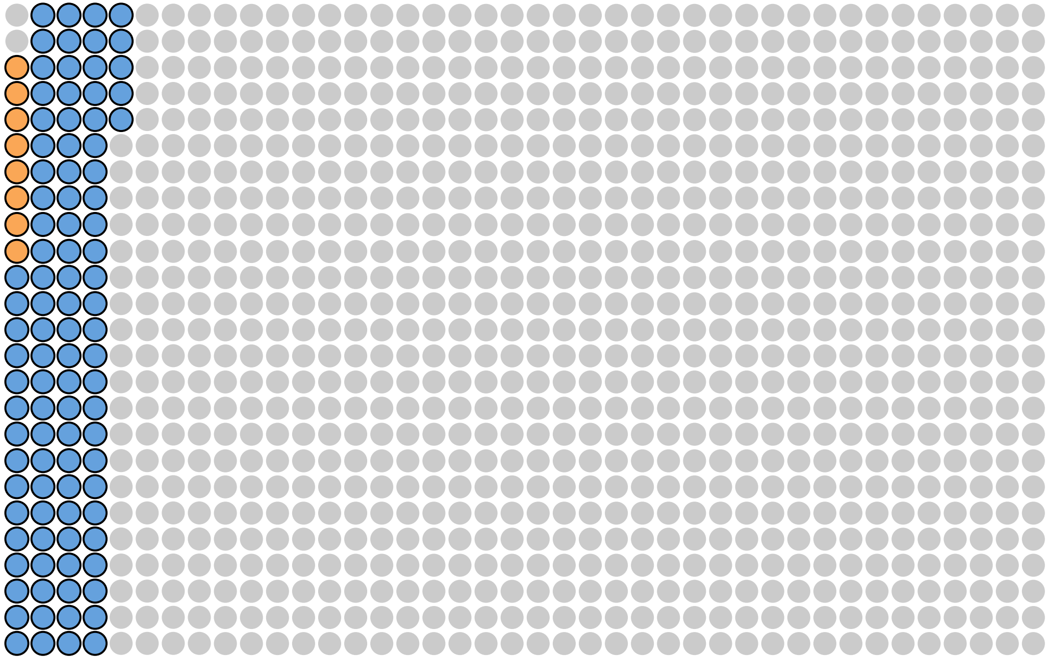


A = Patient has breast cancer

B = Patient has positive mammography

$$\frac{8}{103} = \frac{\text{total} \times P\,(\text{A}) \times P\,(\text{B} \mid \text{A})}{\text{all positives}}$$

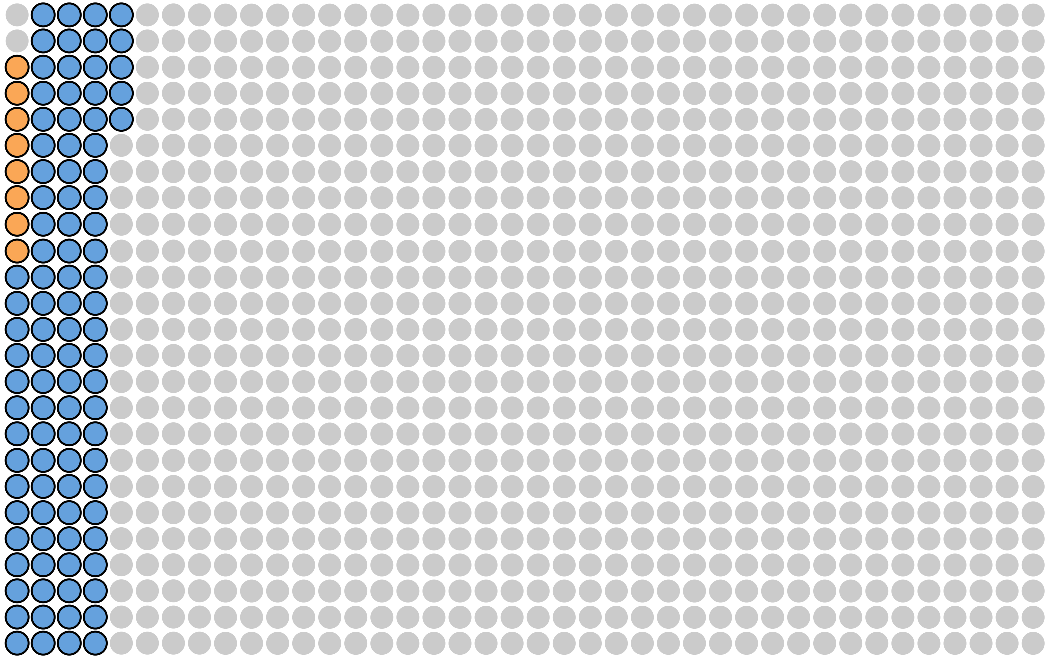
7.8%



A = Patient has breast cancer

B = Patient has positive mammography

$$\frac{8}{103} = \frac{\text{total} \times P(\text{A}) \times P(\text{B} \mid \text{A})}{\text{total} \times P(\text{A}) \times P(\text{B} \mid \text{A}) \quad + \quad \text{total} \times P(\neg\text{A}) \times P(\text{B} \mid \neg\text{A})}$$

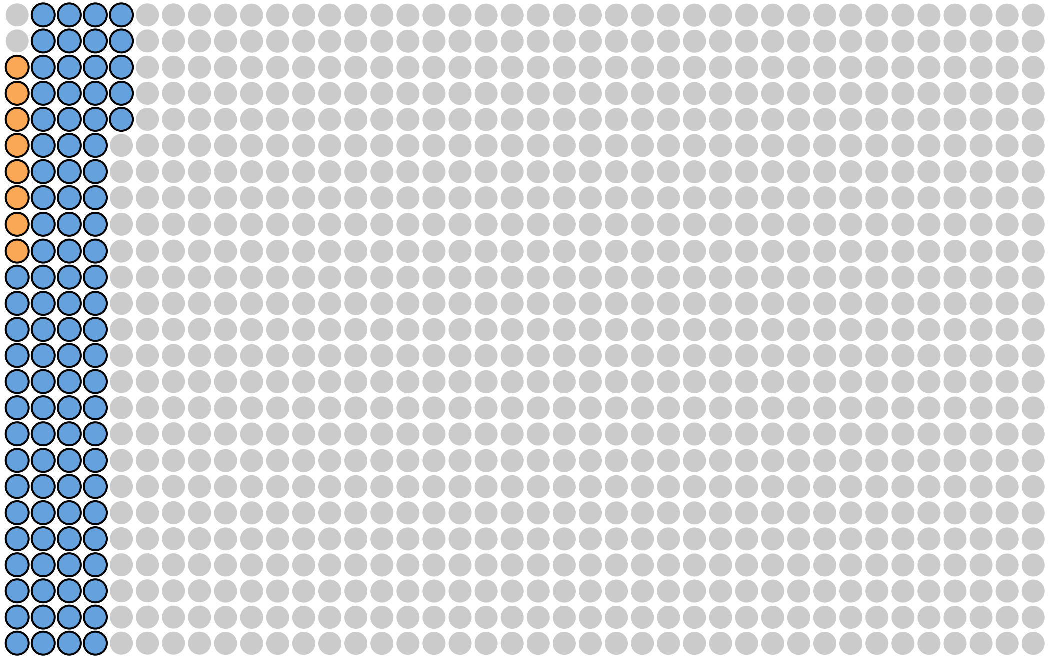


7.8%

A = Patient has breast cancer

B = Patient has positive mammography

$$\frac{8}{103} = \frac{\cancel{total} \times P(A) \times P(B \mid A)}{\cancel{total} \times P(A) \times P(B \mid A) \quad + \quad \cancel{total} \times P(\neg A) \times P(B \mid \neg A)}$$



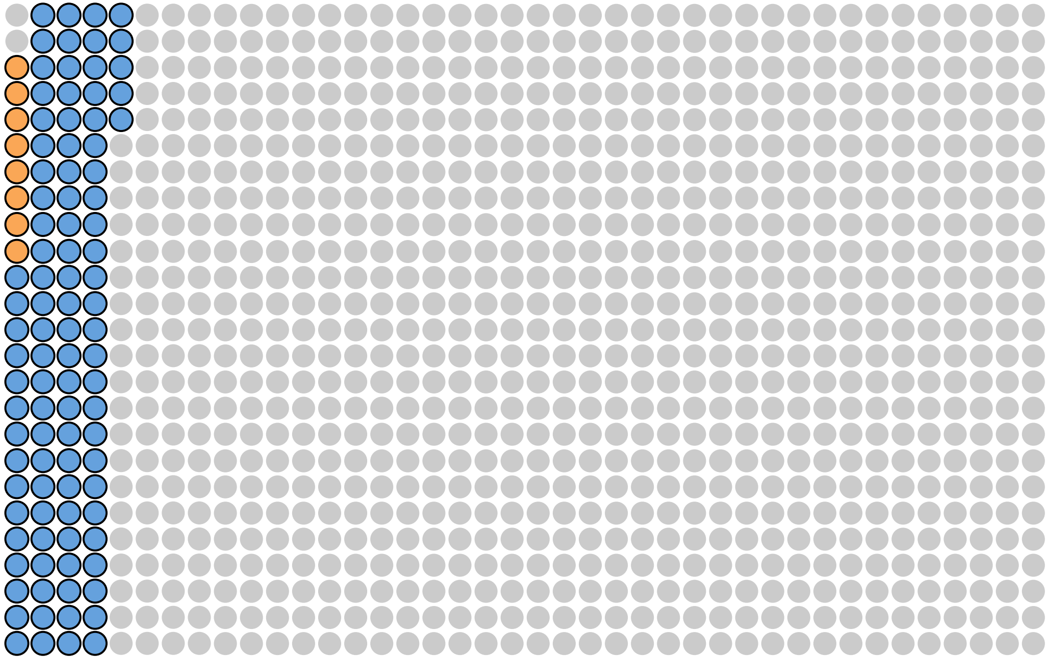
7.8%

A = Patient has breast cancer

B = Patient has positive mammography

$$\frac{8}{103} = \frac{P(\mathrm{A}) \times P(\mathrm{B} \mid \mathrm{A})}{P(\mathrm{A}) \times P(\mathrm{B} \mid \mathrm{A}) \quad + \quad P(\neg \mathrm{A}) \times P(\mathrm{B} \mid \neg \mathrm{A})}$$

7.8%

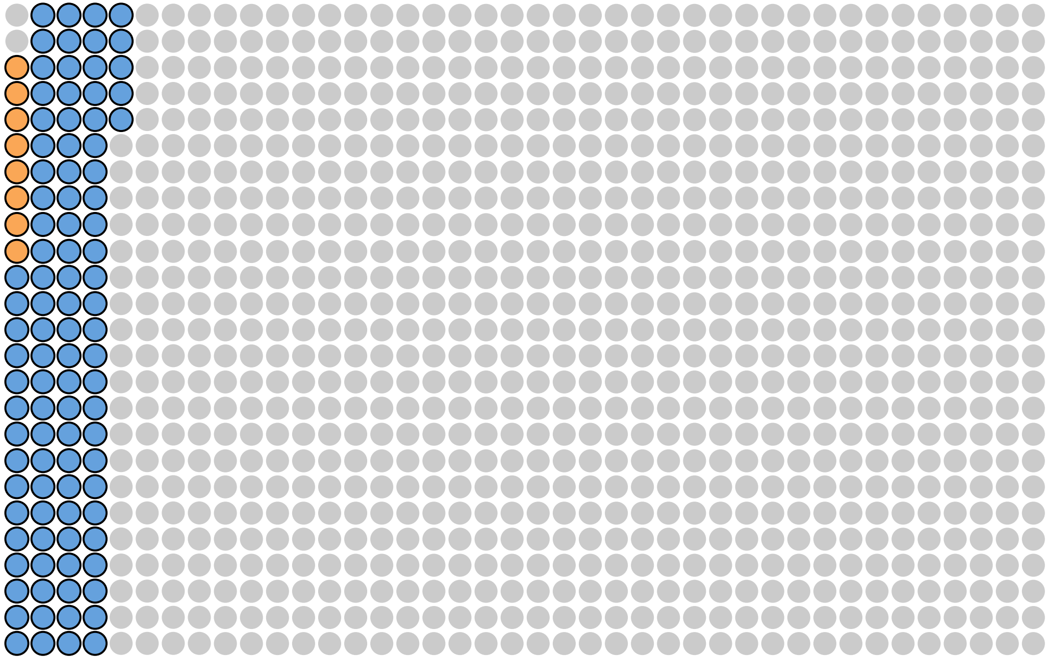


A = Patient has breast cancer

B = Patient has positive mammography

$$\frac{8}{103} = \frac{P(\mathrm{A}) \times P(\mathrm{B} \mid \mathrm{A})}{P(\mathrm{A}) \times P(\mathrm{B} \mid \mathrm{A}) + P(\neg\mathrm{A}) \times P(\mathrm{B} \mid \neg\mathrm{A})}$$

7.8%

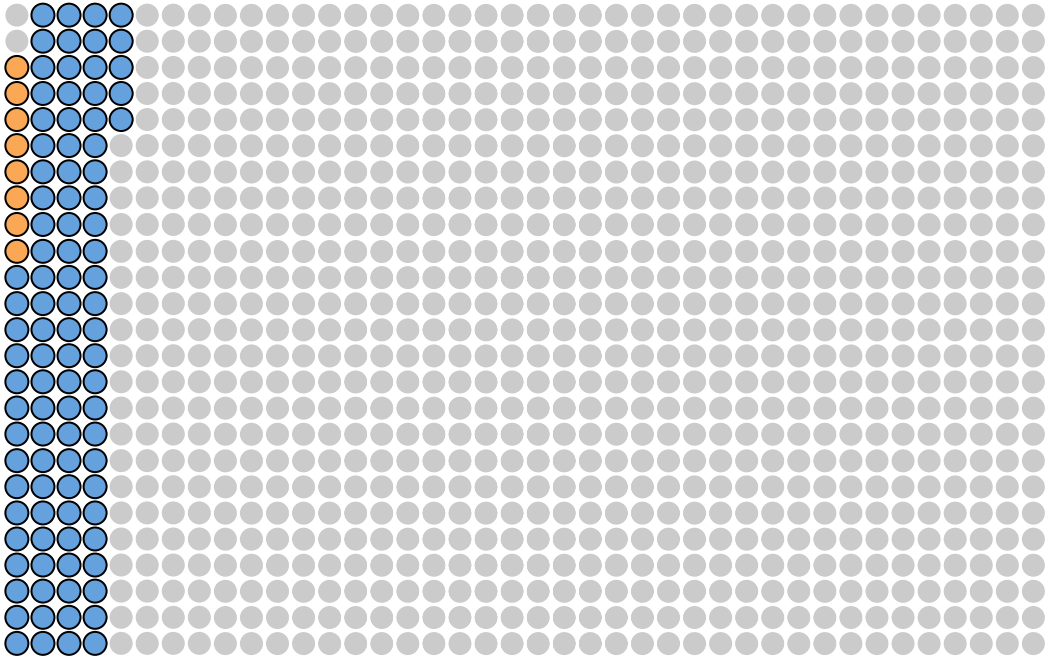


A = Patient has breast cancer

B = Patient has positive mammography

$$\frac{8}{103} = \frac{P(\mathrm{A}) \times P(\mathrm{B} \mid \mathrm{A})}{P(\mathrm{B})}$$

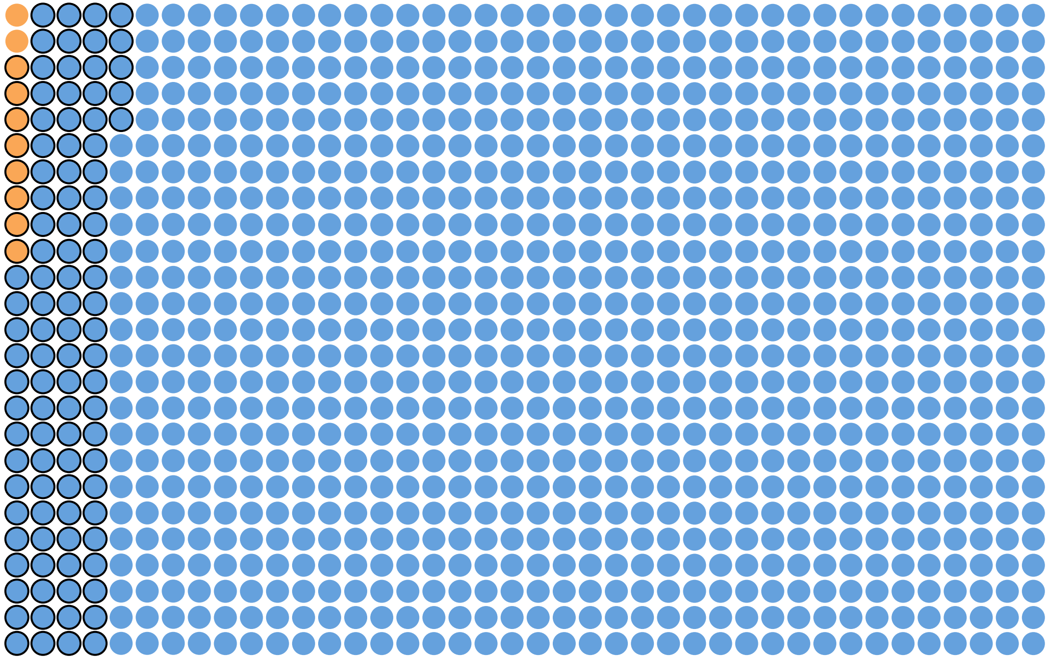
7.8%



A = Patient has breast cancer

B = Patient has positive mammography

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})\, P(\mathrm{A})}{P(\mathrm{B})} = 7.8\%$$



Bayes Theorem:

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})\, P(\mathrm{A})}{P(\mathrm{B})}$$

Bayes Theorem:

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{P(\mathrm{B})} \times P(\mathrm{A})$$

Bayes Theorem:

$$P(A \mid B) = \frac{P(B \mid A)}{\underbrace{P(B)}_{P(A) \times P(B \mid A) \; + \; P(\neg A) \times P(B \mid \neg A)}} \times \overbrace{P(A)}^{\text{prior}}$$

Bayes Theorem:

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{P(\underbrace{\mathrm{B}}_{\text{new evidence}})} \times P(\mathrm{A})$$

Bayes Theorem:

$$P(\mathrm{A} \mid \mathrm{B}) = \underbrace{\frac{P(\mathrm{B} \mid \mathrm{A})}{P(\mathrm{B})}}_{\text{influence of new evidence}} \times P(\mathrm{A})$$

Bayes Theorem:

$$\overbrace{P(\mathrm{A} \mid \mathrm{B})}^{\text{posterior}} = \frac{P(\mathrm{B} \mid \mathrm{A})}{P(\mathrm{B})} \times P(\mathrm{A})$$

Bayes Theorem: Test #2 is also positive

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{P(\mathrm{B})} \times P(\mathrm{A})$$

Bayes Theorem: Test #2 is also positive

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{P(\mathrm{B})} \times P(\mathrm{A})$$

Bayes Theorem: Test #2 is also positive

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{P(\mathrm{B})} \times 7.8\%$$

Bayes Theorem: Test #2 is also positive

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{\underbrace{P(\mathrm{B})}_{P(\mathrm{A}) \times P(\mathrm{B} \mid \mathrm{A}) \;+\; P(\neg \mathrm{A}) \times P(\mathrm{B} \mid \neg \mathrm{A})}} \times .078$$

Bayes Theorem: Test #2 is also positive

$$P(A \mid B) = \frac{P(B \mid A)}{\underbrace{P(B)}_{.078 \times P(B \mid A) + P(\neg A) \times P(B \mid \neg A)}} \times .078$$

Bayes Theorem: Test #2 is also positive

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{\underbrace{P(\mathrm{B})}_{.078 \times P(\mathrm{B} \mid \mathrm{A}) \; + \; .922 \times P(\mathrm{B} \mid \neg \mathrm{A})}} \times .078$$

Bayes Theorem: Test #2 is also positive

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{\underbrace{P(\mathrm{B})}_{.078 \times P(\mathrm{B} \mid \mathrm{A}) + .922 \times .096}} \times .078$$

Bayes Theorem: Test #2 is also positive

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{.80}{\underbrace{P(\mathrm{B})}_{.078 \times .80 \; + \; .922 \times .096}} \times .078$$

Bayes Theorem: Test #2 is also positive

$$41.3\% = \frac{.80}{\underbrace{P(\mathrm{B})}_{.078 \times .80 + .922 \times .096}} \times .078$$

Assumption: Test #1 and Test #2 are independent, given the patient's cancer status

Bayes Theorem:

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{P(\mathrm{B})} \times P(\mathrm{A})$$

Bayes Theorem: IS HE CHEATING?!

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{P(\mathrm{B})} \times P(\mathrm{A})$$

Bayes Theorem: IS HE CHEATING?!

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{\underbrace{P(\mathrm{B})}_{P(\mathrm{A}) \times P(\mathrm{B} \mid \mathrm{A}) \;+\; P(\neg \mathrm{A}) \times P(\mathrm{B} \mid \neg \mathrm{A})}} \times .04$$

Bayes Theorem: IS HE CHEATING?!

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{\underbrace{P(\mathrm{B})}_{.04 \times P(\mathrm{B} \mid \mathrm{A}) \; + \; .86 \times P(\mathrm{B} \mid \neg \mathrm{A})}} \times .04$$

Bayes Theorem: IS HE CHEATING?!

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{\underbrace{P(\mathrm{B})}_{.04 \times P(\mathrm{B} \mid \mathrm{A}) \; + \; .86 \times .05}} \times .04$$

Bayes Theorem: IS HE CHEATING?!

$$P(A \mid B) = \frac{.50}{\underbrace{P(B)}_{.04 \times .50 \;+\; .86 \times .05}} \times .04$$

Bayes Theorem: IS HE CHEATING?!

$$29\% = \frac{.50}{\underbrace{P(\mathrm{B})}_{.04 \ \times \ .50 \ + \ .86 \ \times \ .05}} \times .04$$

Bayes Theorem:

Mammogram #1

$$7.8\% = \frac{.80}{.01 \times .80 + .99 \times .096} \times .01$$

IS HE CHEATING?!

$$29\% = \frac{.50}{.04 \times .50 + .86 \times .05} \times .04$$

Bayes Theorem: 9-11

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{\underbrace{P(\mathrm{B})}_{P(\mathrm{A}) \times P(\mathrm{B} \mid \mathrm{A}) \;+\; P(\neg\mathrm{A}) \times P(\mathrm{B} \mid \neg\mathrm{A})}} \times P(\mathrm{A})$$

Bayes Theorem: 9-11

$$P(\mathrm{A}\,|\,\mathrm{B}) = \frac{P(\mathrm{B}\,|\,\mathrm{A})}{\underbrace{P(\mathrm{B})}_{P(\mathrm{A}) \times P(\mathrm{B}\,|\,\mathrm{A}) \;+\; P(\neg\mathrm{A}) \times P(\mathrm{B}\,|\,\neg\mathrm{A})}} \times .00005$$

Bayes Theorem: 9-11

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{\underbrace{P(\mathrm{B})}_{.00005 \times P(\mathrm{B} \mid \mathrm{A}) \; + \; .99995 \times P(\mathrm{B} \mid \neg \mathrm{A})}} \times .00005$$

Bayes Theorem: 9-11

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{B} \mid \mathrm{A})}{\underbrace{P(\mathrm{B})}_{.00005 \times P(\mathrm{B} \mid \mathrm{A}) \;+\; .99995 \times .00008}} \times .00005$$

Bayes Theorem: 9-11

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{.95}{\underbrace{P(\mathrm{B})}_{.00005 \times .95 \quad + \quad .99995 \times .00008}} \times .00005$$

Bayes Theorem: 9-11 Plane #1

$$37\% = \frac{.95}{\underbrace{P(\mathrm{B})}_{.00005 \times .95 \quad + \quad .99995 \times .00008}} \times .00005$$

Bayes Theorem: 9-11 Plane #2

$$37\% = \frac{.95}{\underbrace{P\,(\mathrm{B})}_{.00005 \times .95 \; + \; .99995 \times .00008}} \times .00005$$

Bayes Theorem: 9-11 Plane #2

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{.95}{\underbrace{P(\mathrm{B})}_{.37 \times .95 \quad + \quad .63 \times .00008}} \times .37$$

Bayes Theorem: 9-11 Plane #2

$$99.9\% = \frac{.95}{\underbrace{P\,(\mathrm{B})}_{.37 \times .95 \quad + \quad .63 \times .00008}} \times .37$$