BUS 139 Data-driven Marketing

Stanford | Continuing Studies

Prediction

1. GOAL SETTING
2. COLLECTION & PREPARATION
3. ANALYSIS
4. DECISION
5. VISUALIZATION
6. PRESENTATION
7. PREDICTION

How prediction fits into marketing

Three Examples:

1. Eliminating risk. For example, predicting which campaigns are likely to perform best.

2. Budget Allocation. For example, if you know sales are highest during the holiday season, you may consider allocating more marketing spend to that time period.

3. Resource planning. For example, predicting which types of social media content are likely to perform best.

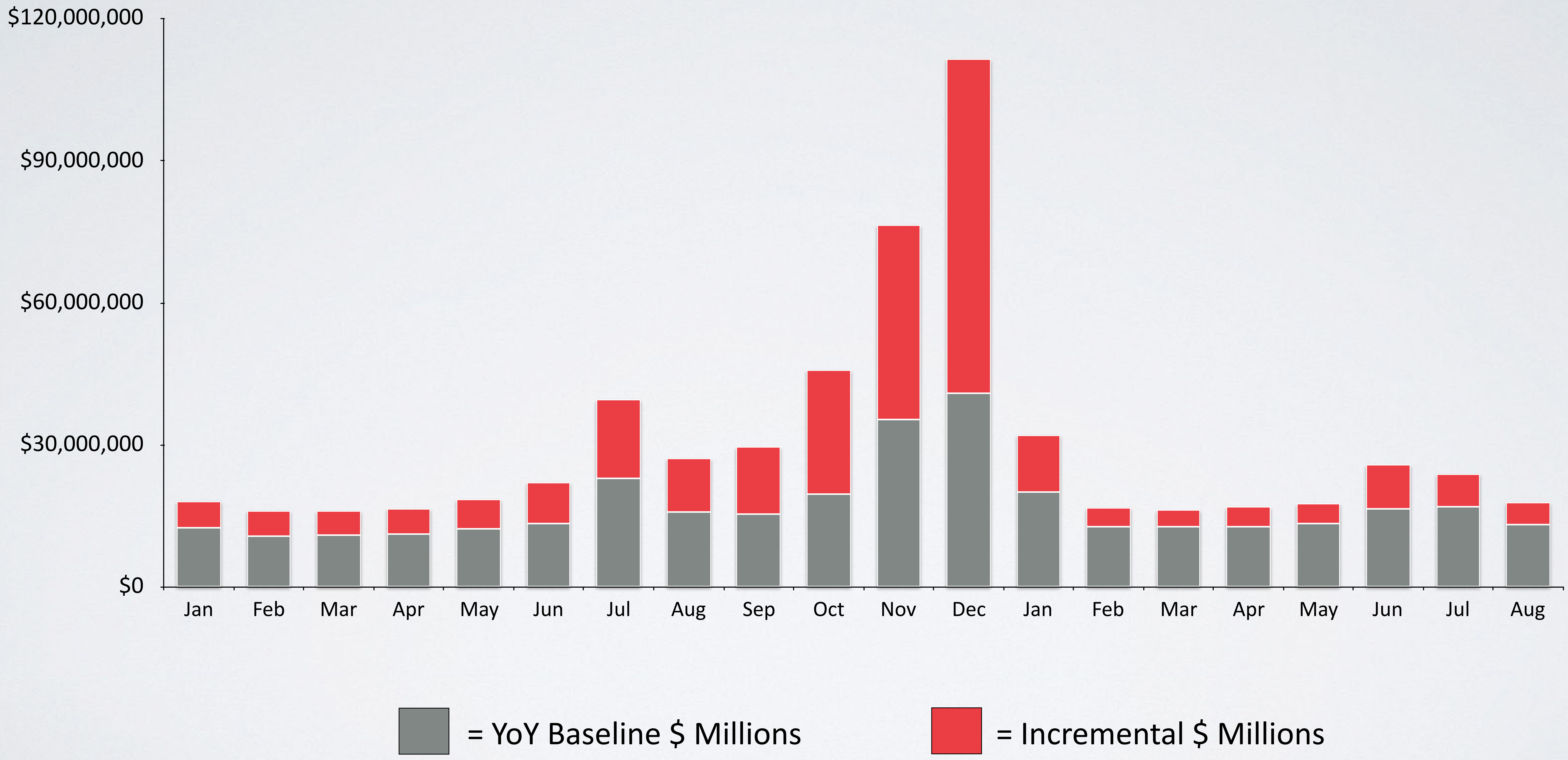
Forecasting vs. Predicting

Key differences and uses

1. Forecasting is associated with a time series relationship and is used for defining some number in the future, e.g., weather forecasts, stock prices, etc.

2. Predictions are most often outcome focused, e.g., who wins an election, whether a customer is likely to churn, etc.

Forecasting Wine Sales



Predicting Customer Lifetime Value

CLV = Customer Lifetime Value

P = Average number of bottles purchased monthly

AOV = Average order value

ACL = Average Customer Lifespan (in months)

AGM = Average gross margin

$$CLV = ((P \times AOV)\ AGM)\ ACL$$

P	AOV	AGM	ACL
2	$19.99	.50	12

Predicting Customer Lifetime Value

CLV = Customer Lifetime Value

P = Average number of bottles purchased monthly

AOV = Average order value

ACL = Average Customer Lifespan (in months)

AGM = Average gross margin

$$CLV = ((P \times AOV)\ AGM)\ ACL$$

\$239.88	2	\$19.99	.50	12

Five examples of prediction

1. CAGR

Definition?

The year-over-year growth of some number over a specified period of time, calculated by taking the nth root of the total percentage growth rate, where n is the number of years in the period being considered.

In plain english:

If you were to draw a line showing the year over year trend in growth, it would be a curvy line due to the variance in the data. CAGR flattens out the curves by looking at the average of the data.

YEAR	2011	2012	2013	2014	2015	2016
eCommerce Revenue	$1,325,000	$1,350,000	$1,450,000	$1,500,000	$1,650,000	$1,800,000
Growth		1.89%	7.41%	3.45%	10.00%	9.09%

1. CAGR

5-year CAGR = 6.32%

2. Bayes Rule

What is it?

Bayes' rule says that if you have a hypothesis H and evidence E that bears on that hypothesis, then we use the notation Pr[A] denotes the probability of an event A and Pr[A | B] denote the probability of A conditional on another event B. The hypothesis H is that play will be say, yes, and Pr{H | E] is going to turn out to be 20.5%, just as determined previously.

In plain english:

Bayes' rule is used to predict probability of outcome. It is calculated like this: probability = desired event / all possible events.

RED

BLACK

The chance of RED coming up four times in a row in a roulette spin is:

1/2 x 1/2 x 1/2 x 1/2

1/16 or 6.25%

3. Linear and logistic regression

What is the difference?

The main difference between linear and logistic regression is that linear regression outcomes are continuous and therefore have an infinite number of possible values. Logistic regression has only a limited number of possible values.

Linear Regression

What is it?

Simple regression analysis is a statistical tool that proves the mathematical relationship between a dependent variable (usually called y) and an independent variable (usually called x).

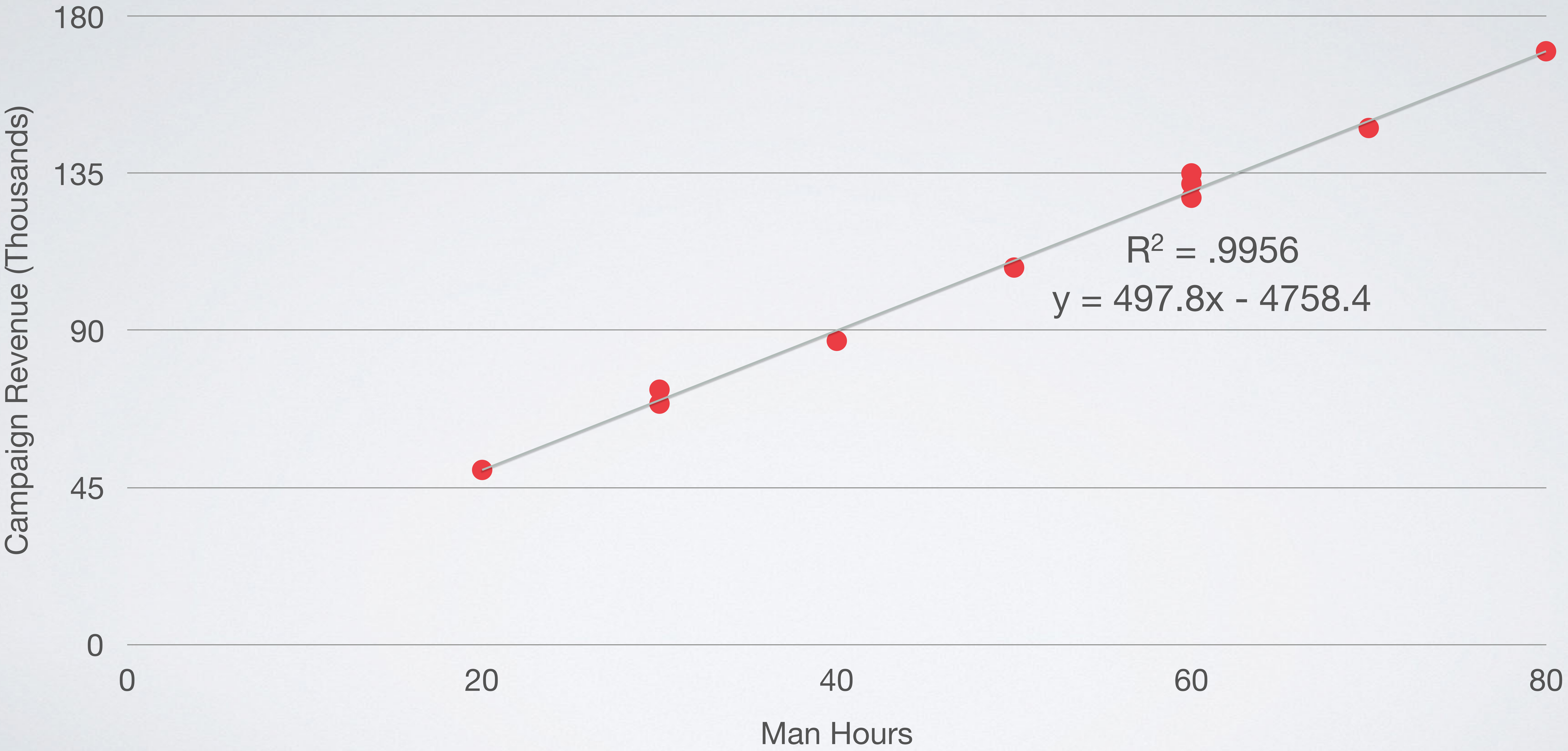
In plain english:

Regression gives us the ability to understand the relationship between two variables and thereby predict outcomes by creating a linear formula.

x=independent y=dependent

Man Hours	Campaign Revenue in Thousands
73	30
50	20
128	60
170	80
87	40
108	50
135	60
69	30
148	70
132	60

(y = 497.8x - 4758.4) means that for every man hour added you will create $497.8 in additional revenue.



Logistic Regression

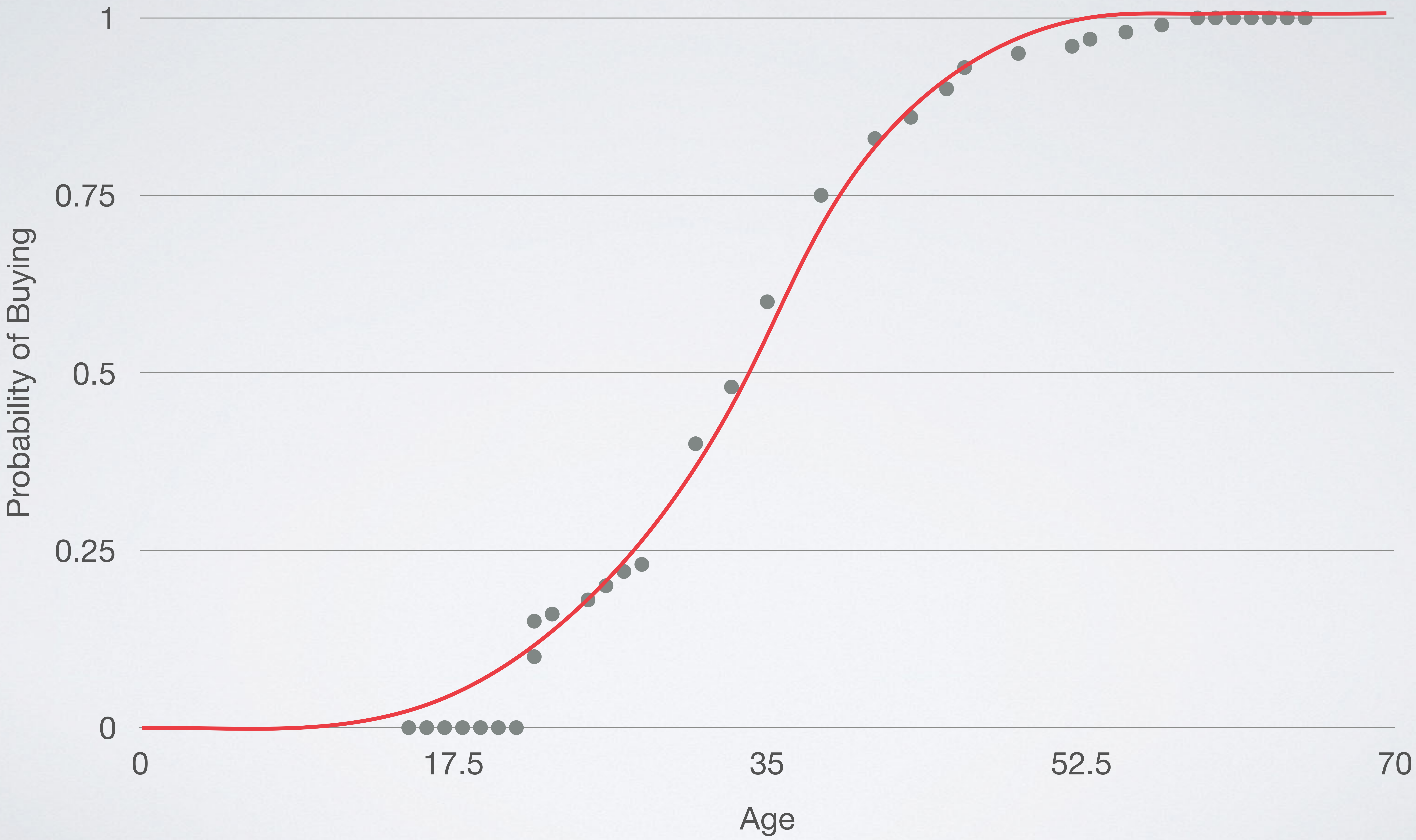
What is it?

Logistic regression, or logit regression, is used to predict a binary response based on one or more predictor variables (features). That is, it is used in estimating the parameters of a qualitative response model.

In plain english:

Predicting a binary outcome represented between 0 and 1. For example: a customer who is likely to buy (1) vs. unlikely to buy (0).

Probability of buying a product



4. Algorithms

Data Mining

There is information that can be obtained (and retained) from single variables and there is also information that can be obtained between the relationships of variables. This is the value of data mining and why marketers use algorithms.

Decision Tree Algorithm

What is it?

The Decision Tree algorithm is based on conditional probabilities generated by rules that map observations about an item to conclusions about the item's target value.

In plain english:

A Decision Tree predicts (or classifies) the outcome of individual instances based on some combination of input variables.

Predicting if we will play golf

outlook = sunny
- humidity = high: yes (2.0)
- humidity = low: no (3.0)

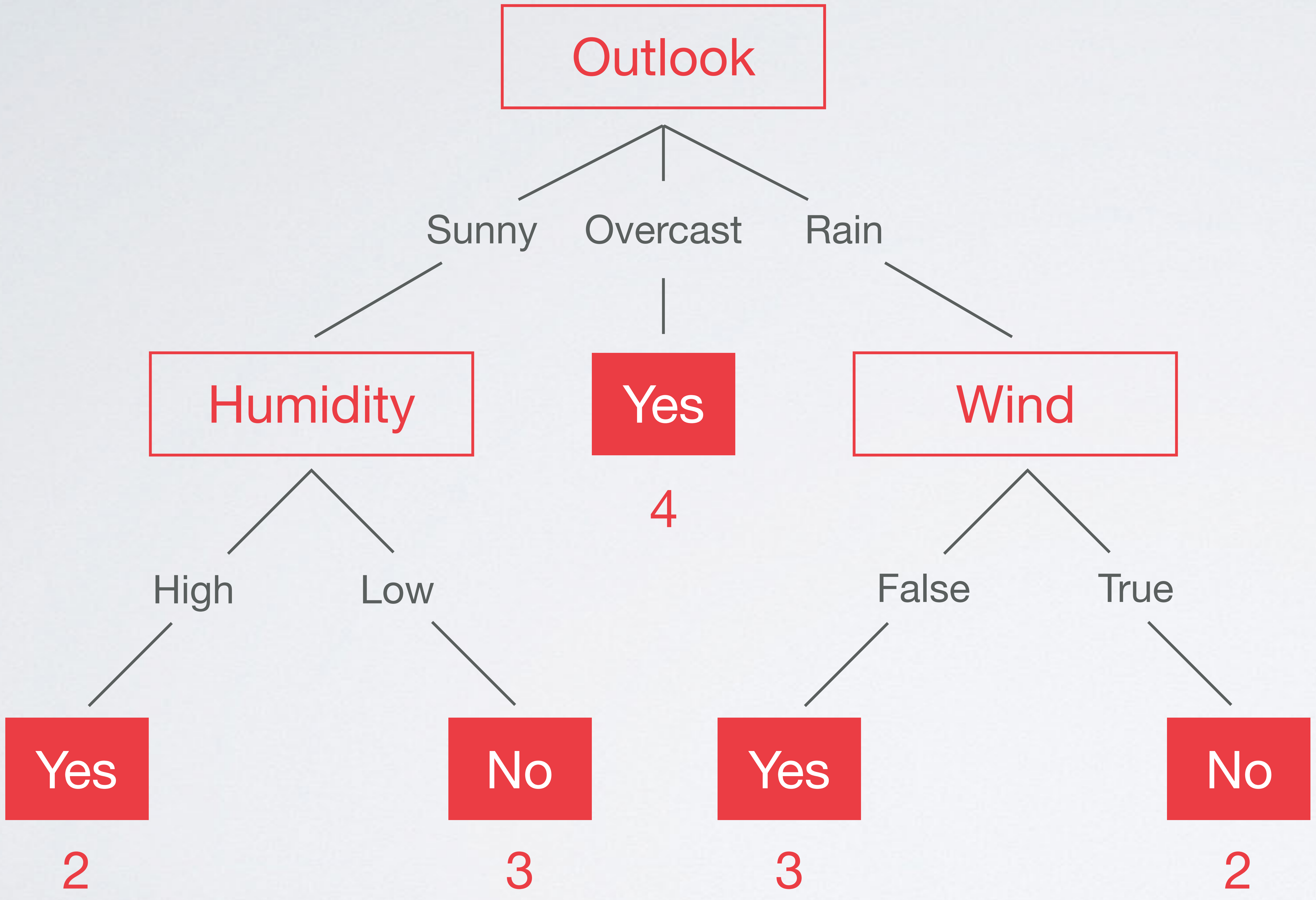
outlook = overcast: yes (4.0)

outlook = rain
- windy = true: no (2.0)
- windy = false: yes (3.0)

Number of leaves: 5

Size of tree: 8

Decision Tree Algorithm



Predicting if we will play golf

outlook = sunny

humidity = high: yes (2.0)

humidity = low: no (3.0)

outlook = overcast: yes (4.0)

outlook = rain

windy = true: no (2.0)

windy = false: yes (3.0)

Number of leaves: 5

Size of tree: 8

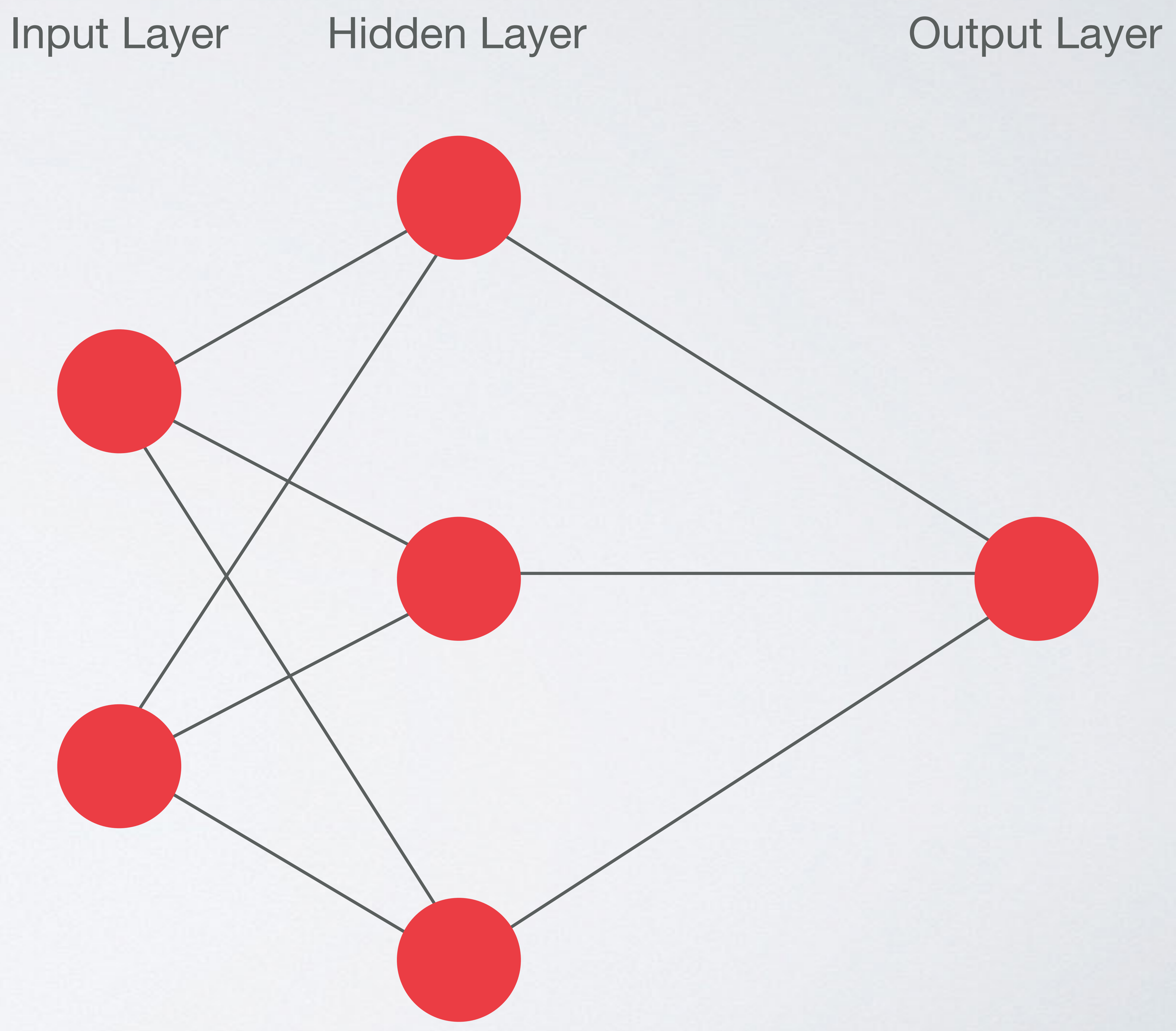
5. Artificial Neural Networks

What is it?

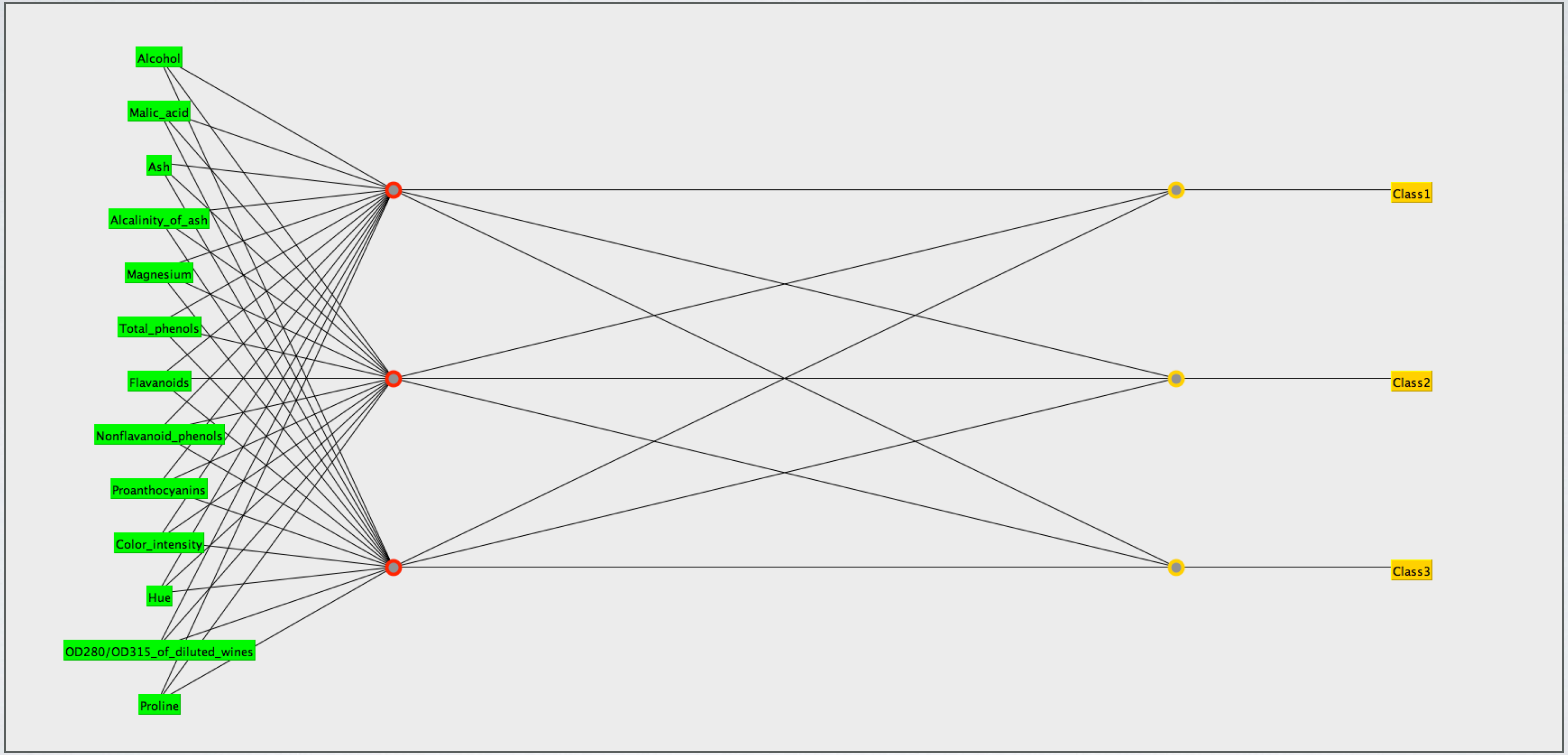
A computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs.

In plain english:

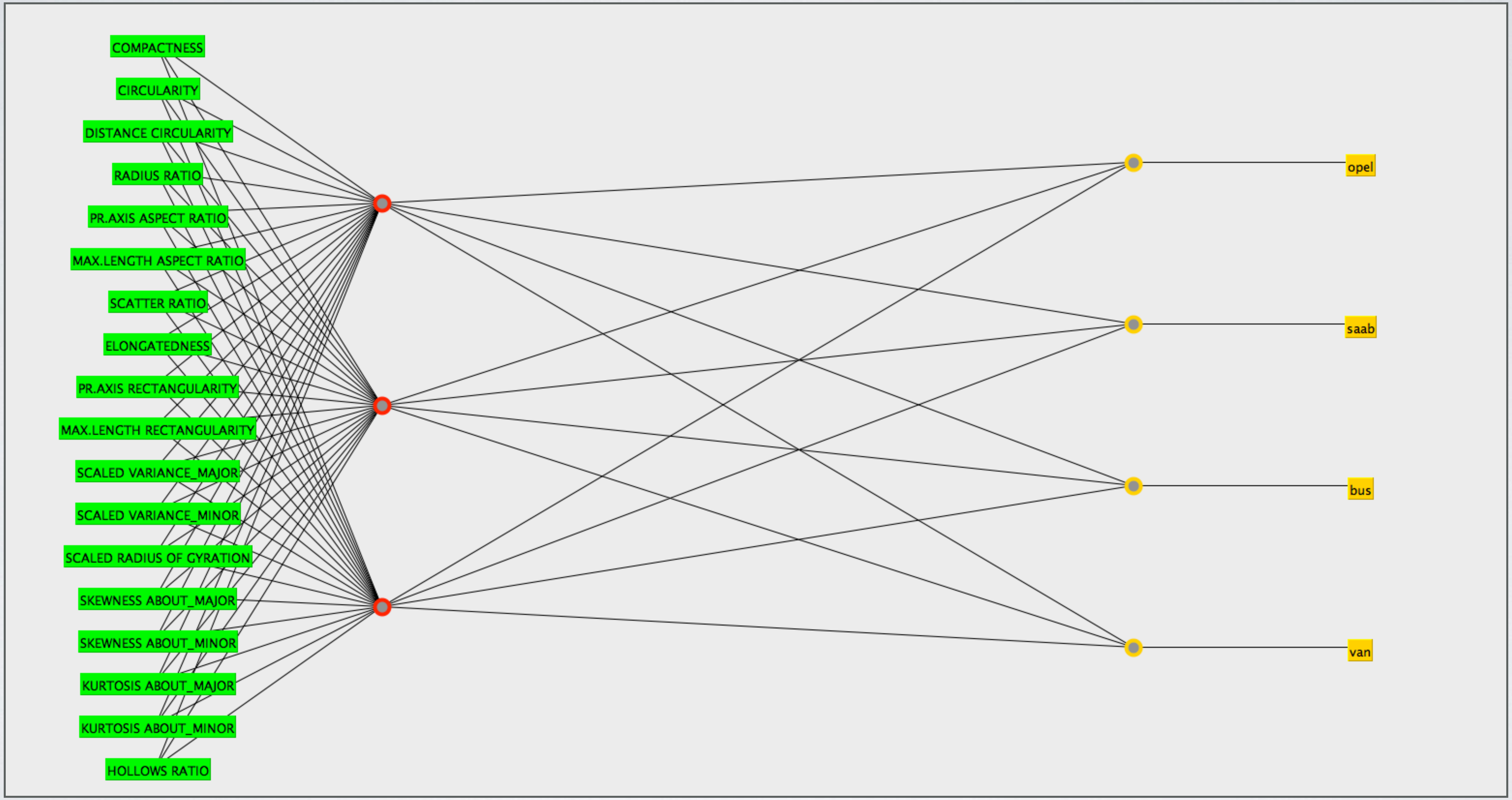
Neural Networks are “connectionist” networks capable of learning (and relearning) based on new examples that are fed into the model.



Predicting wine type based on chemical compounds.



Predicting car models based on measurement of parts



How to:

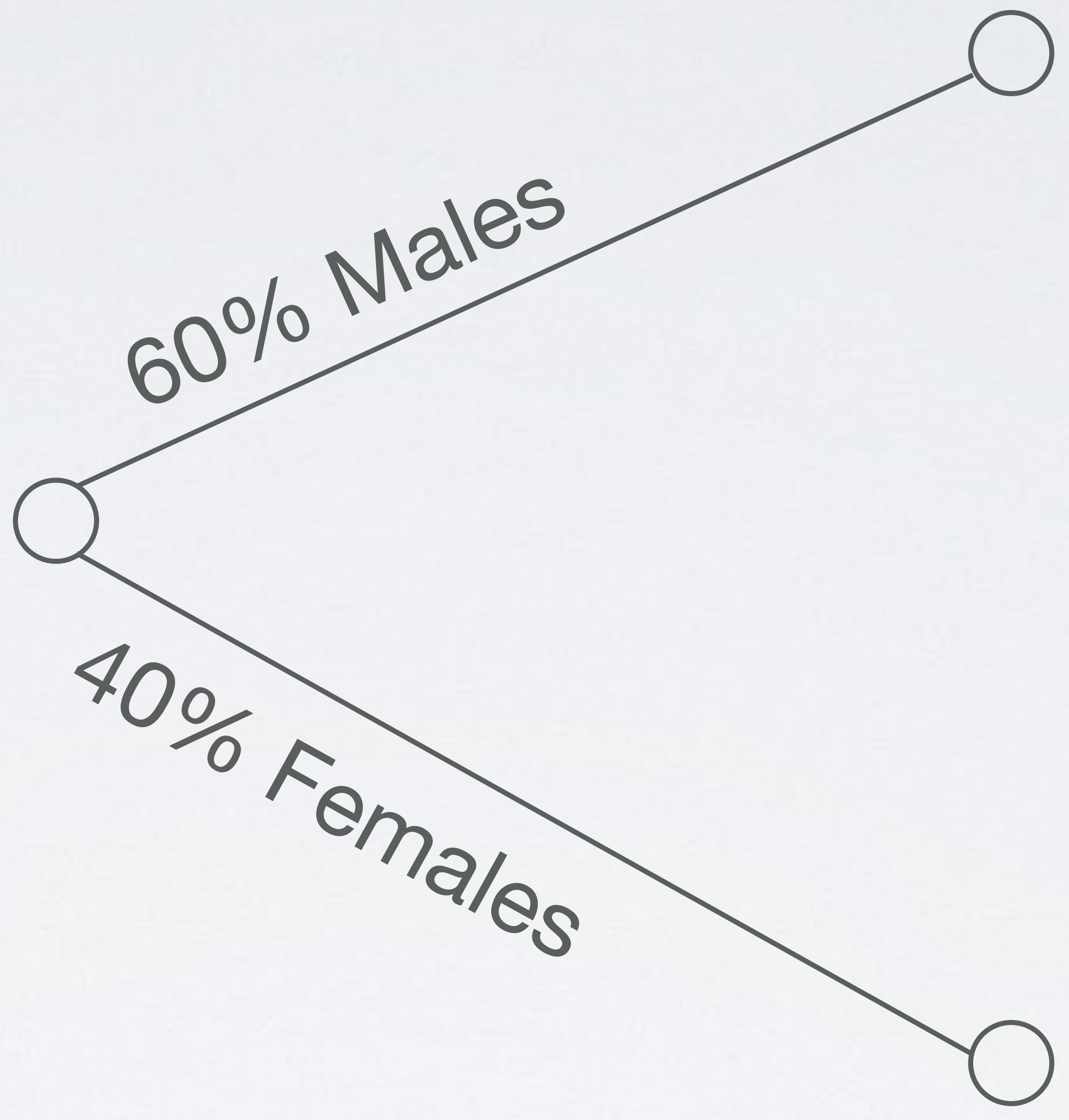
Make predictions with your bare hands

Example:

Suppose an apparel eCommerce site has 60% male customers and 40% female customers. The females buy jeans or skirts in equal numbers; the boys only buy jeans. A random customer buys jeans from the apparel company's website. What is the probability this customer is female?

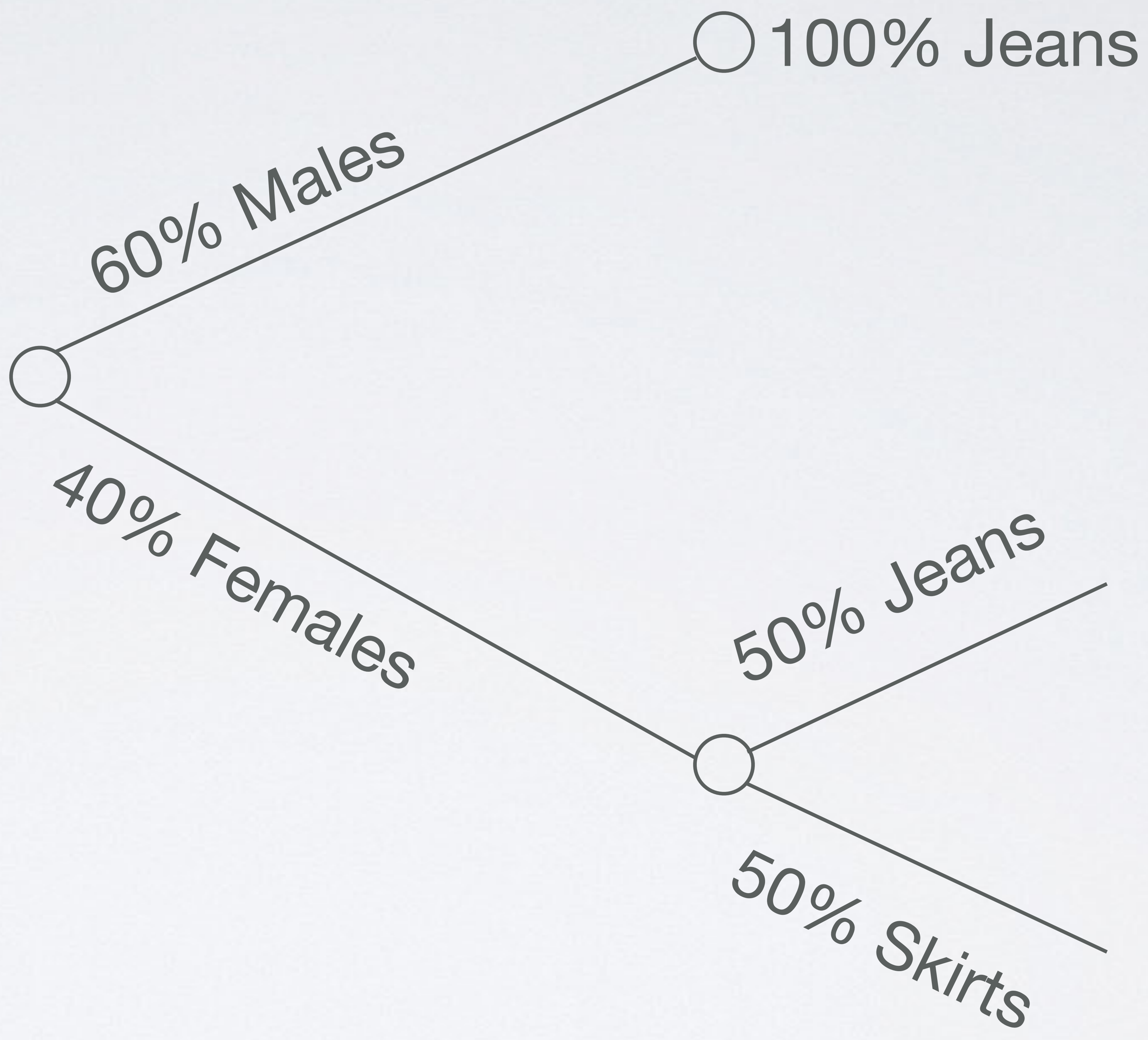
Example:

Suppose an apparel eCommerce site has 60% male customers and 40% female customers. The females buy jeans or skirts in equal numbers; the boys only buy jeans. A random customer buys jeans from the apparel company's website. What is the probability this customer is female?



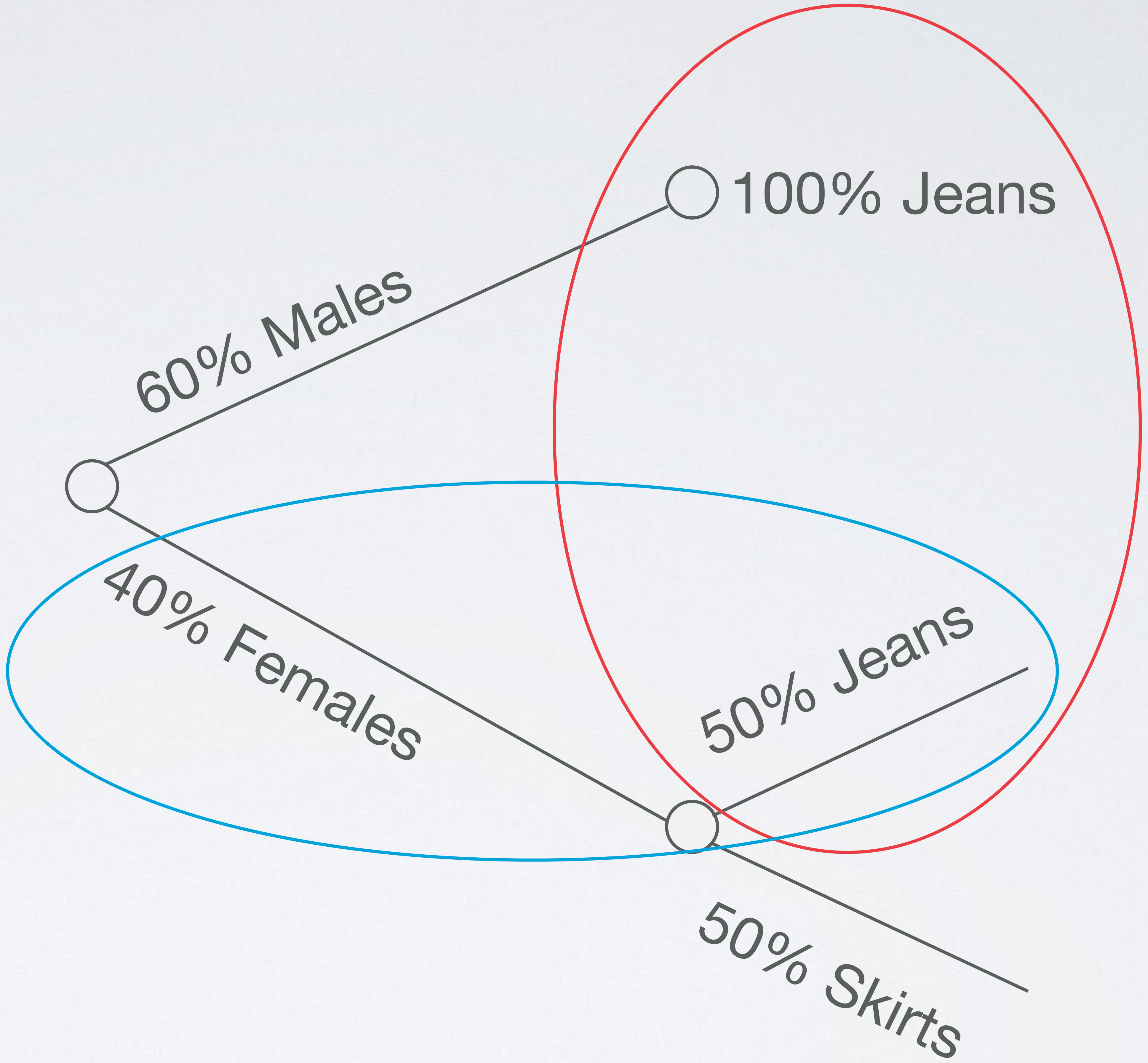
Example:

Suppose an apparel eCommerce site has 60% male customers and 40% female customers. The females buy jeans or skirts in equal numbers; the boys only buy jeans. A random customer buys jeans from the apparel company's website. What is the probability this customer is female?



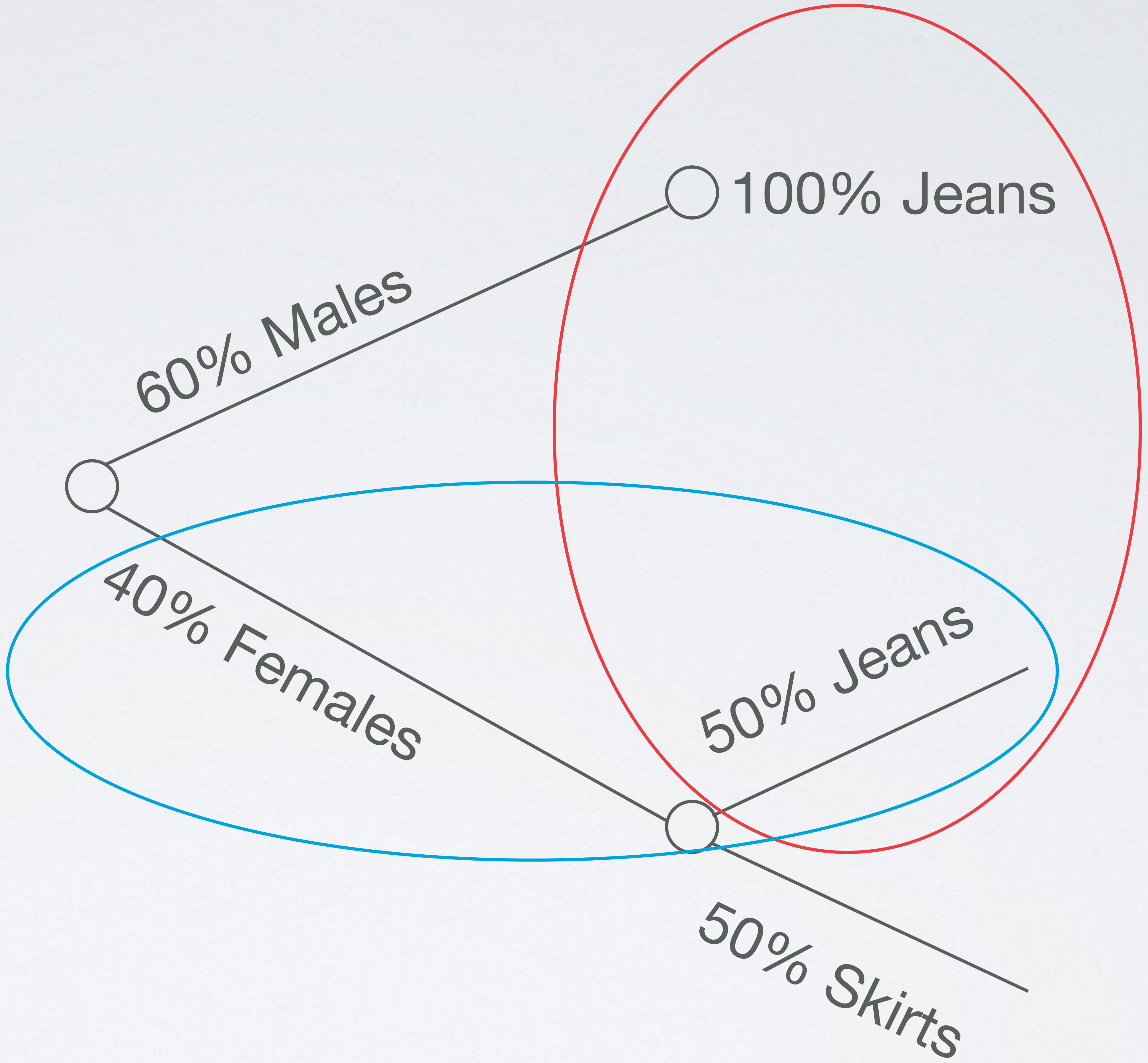
Example:

Suppose an apparel eCommerce site has 60% male customers and 40% female customers. The females buy jeans or skirts in equal numbers; the boys only buy jeans. A random customer buys jeans from the apparel company's website. What is the probability this customer is female?



Example:

Suppose an apparel eCommerce site has 60% male customers and 40% female customers. The females buy jeans or skirts in equal numbers; the boys only buy jeans. A random customer buys jeans from the apparel company's website. What is the probability this customer is female?

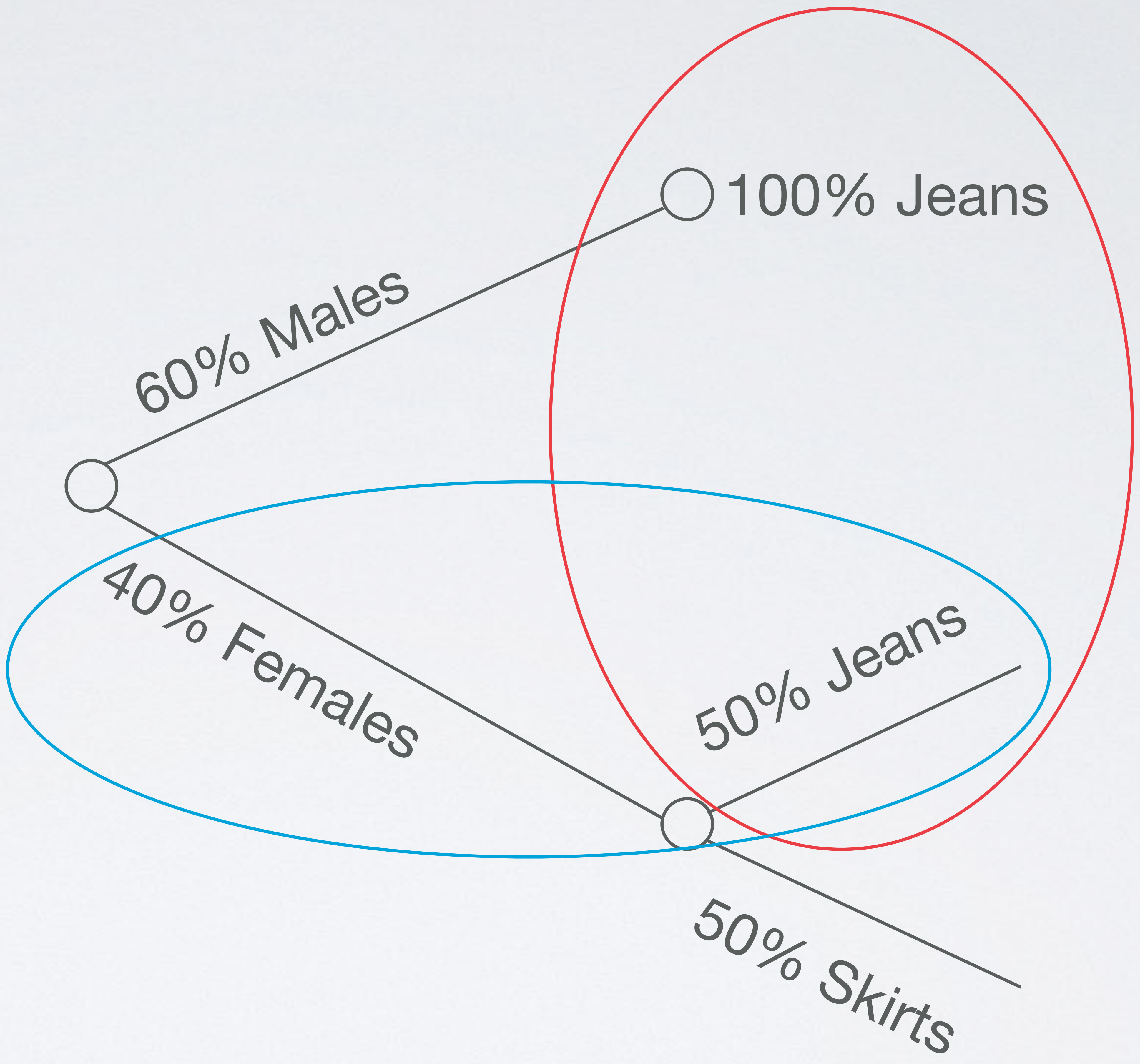


100% of males and 50% of females buy jeans which is 80% of the total website traffic (total possible outcomes).

20% of all traffic are females that buy jeans.

Example:

Suppose an apparel eCommerce site has 60% male customers and 40% female customers. The females buy jeans or skirts in equal numbers; the boys only buy jeans. A random customer buys jeans from the apparel company's website. What is the probability this customer is female?



100% of males and 50% of females buy jeans which is 80% of the total website traffic (total possible outcomes).

20% of all traffic are females that buy jeans.

$$\frac{20\%}{80\%} = \frac{1}{4} = 25\%$$

And lastly, a word about you…