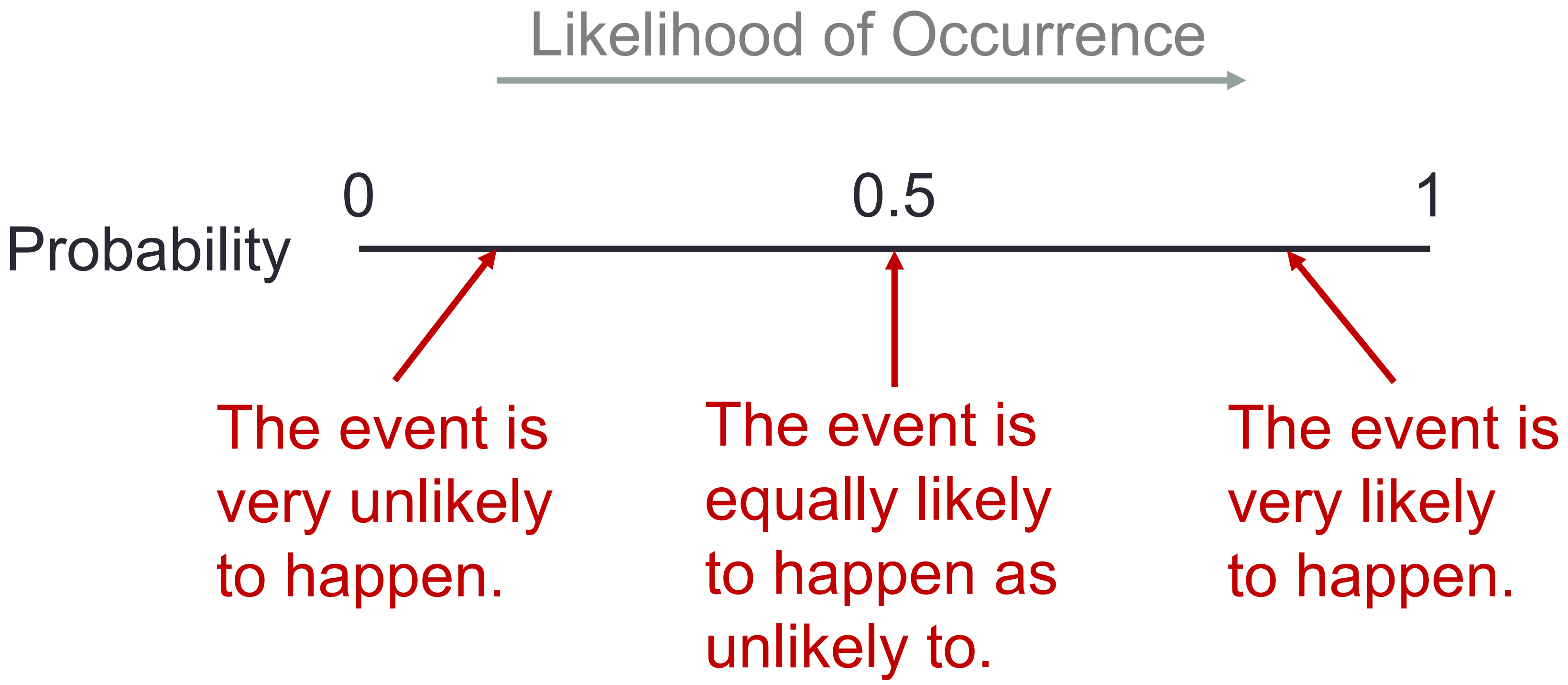
PROBABILITY

Analytics Primer

BASICS

Probability

- The **probability** that an event happens is a numerical measure of the likelihood of that event's occurrence.



Probability

- The **probability** that an event happens is a numerical measure of the likelihood of that event's occurrence.
  - Probabilities are numbers between 0 and 1.
  - Percentages are numbers between 0 and 100.

- **Sample space**: the collection of **all** possible outcomes in a random process.
  - Sum of all probabilities for an experiment must sum to 1.

Events

- An **event** is a collection of one or more outcomes from a process whose result cannot be predicted with certainty.
- The probability of an event A is denoted, $P(A)$.
- Example: When rolling a fair dice, what is the probability of rolling a 6?

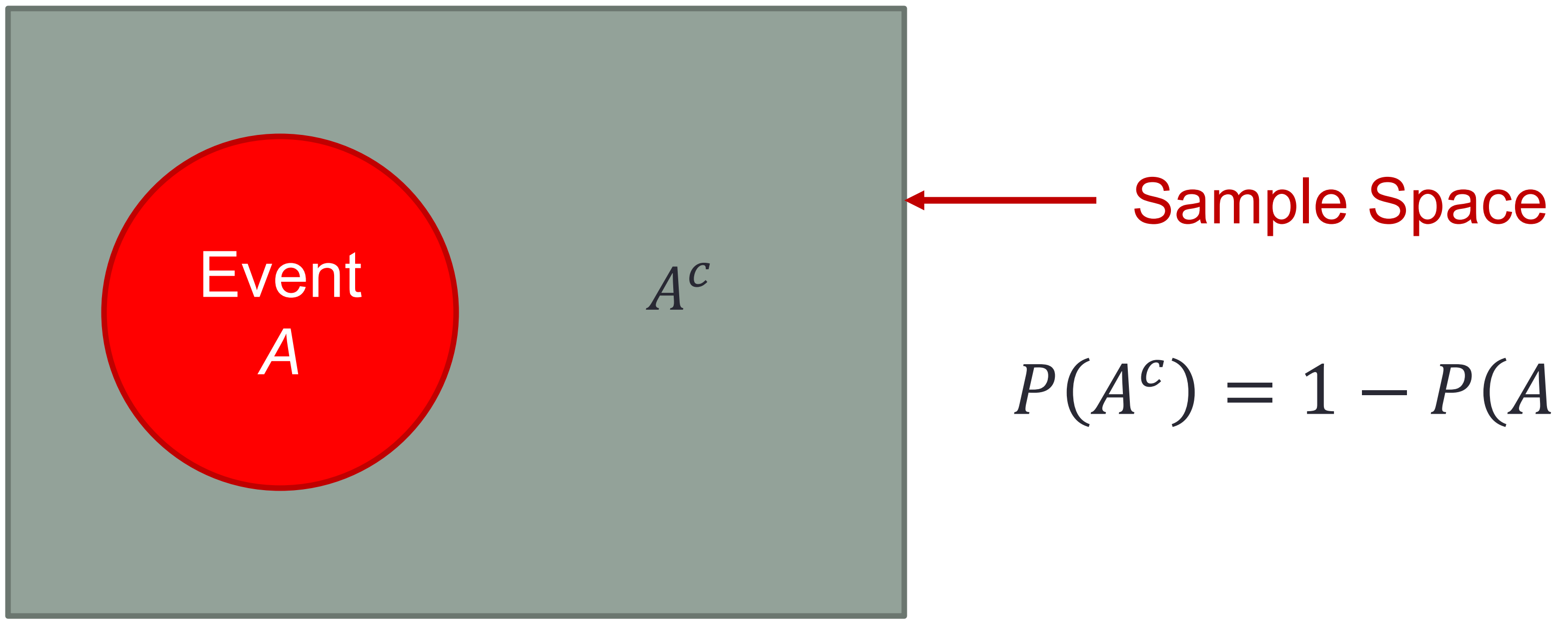
BASIC PROBABILITY RULES

Basic Relationships

- There are some basic probability relationships that still can be used to calculate the probability of an event occurring:
  - Complement of an Event
  - Union of Two Events
  - Intersection of Two Events
  - Mutually Exclusive Events

Complement of an Event

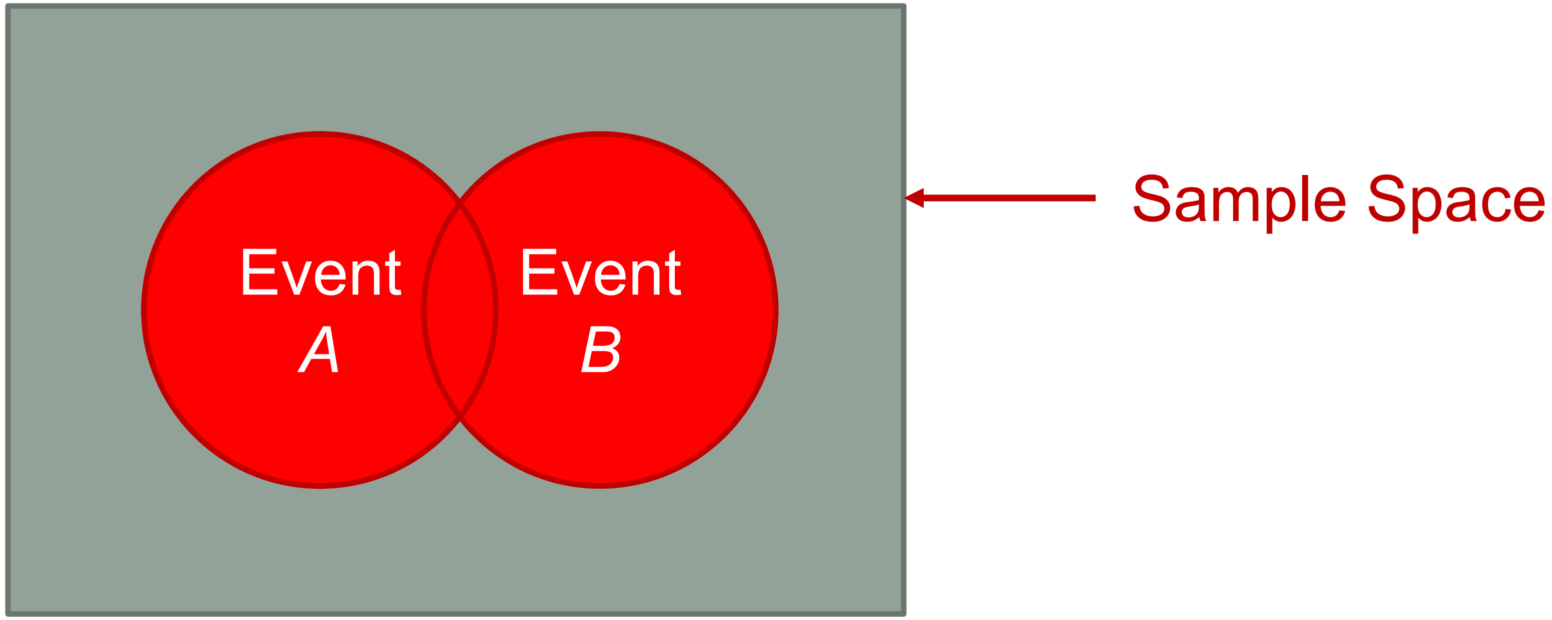
- The **complement** of an event *A* is defined to be the event consisting of all sample points that are not in *A*.
- The complement of *A* is denoted by either A^c or $\bar{A}$.



$$P(A^c) = 1 - P(A)$$

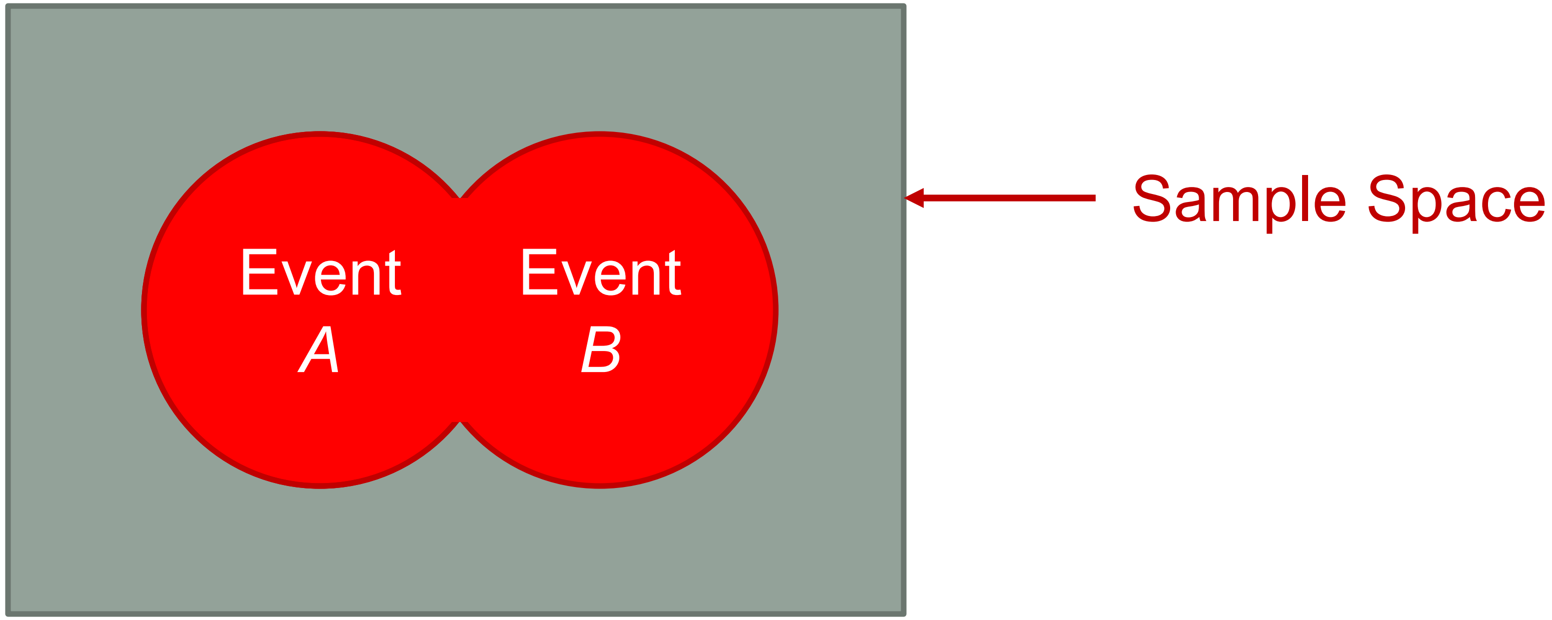
Union of Two Events

- The **union** of an event *A* and an event *B* is the event containing all sample points that are in *A* or *B* or both.
- The union of *A* and *B* is denoted by $A \cup B$.



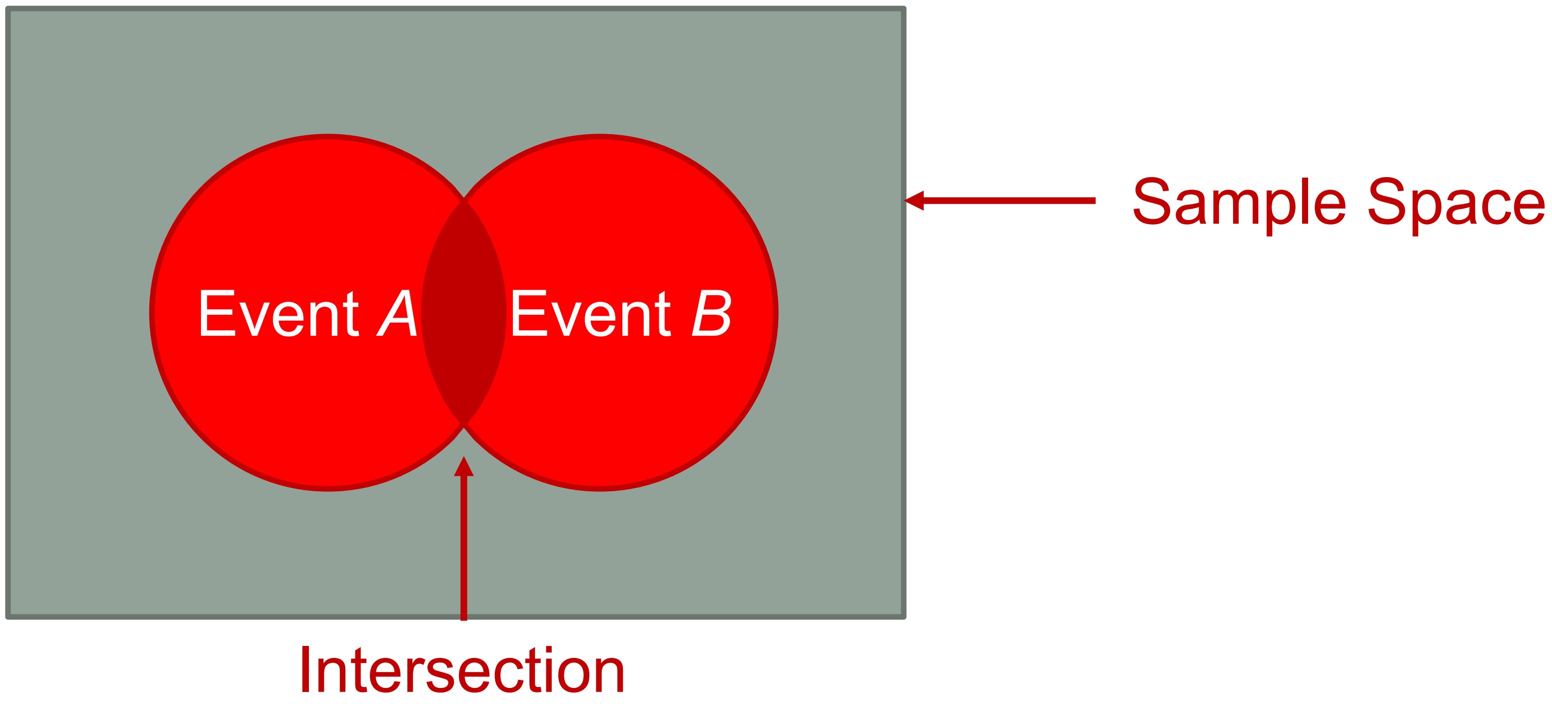
Union of Two Events

- The **union** of an event *A* and an event *B* is the event containing all sample points that are in *A* or *B* or both.
- The union of *A* and *B* is denoted by $A \cup B$.



Intersection of Two Events

- The **intersection** of an event *A* and an event *B* is the event containing all sample points that are in **both** *A* and *B*.
- The intersection of *A* and *B* is denoted by $A \cap B$.



Addition Law

- The **addition law** provides a way to compute the union of events *A* and *B*:

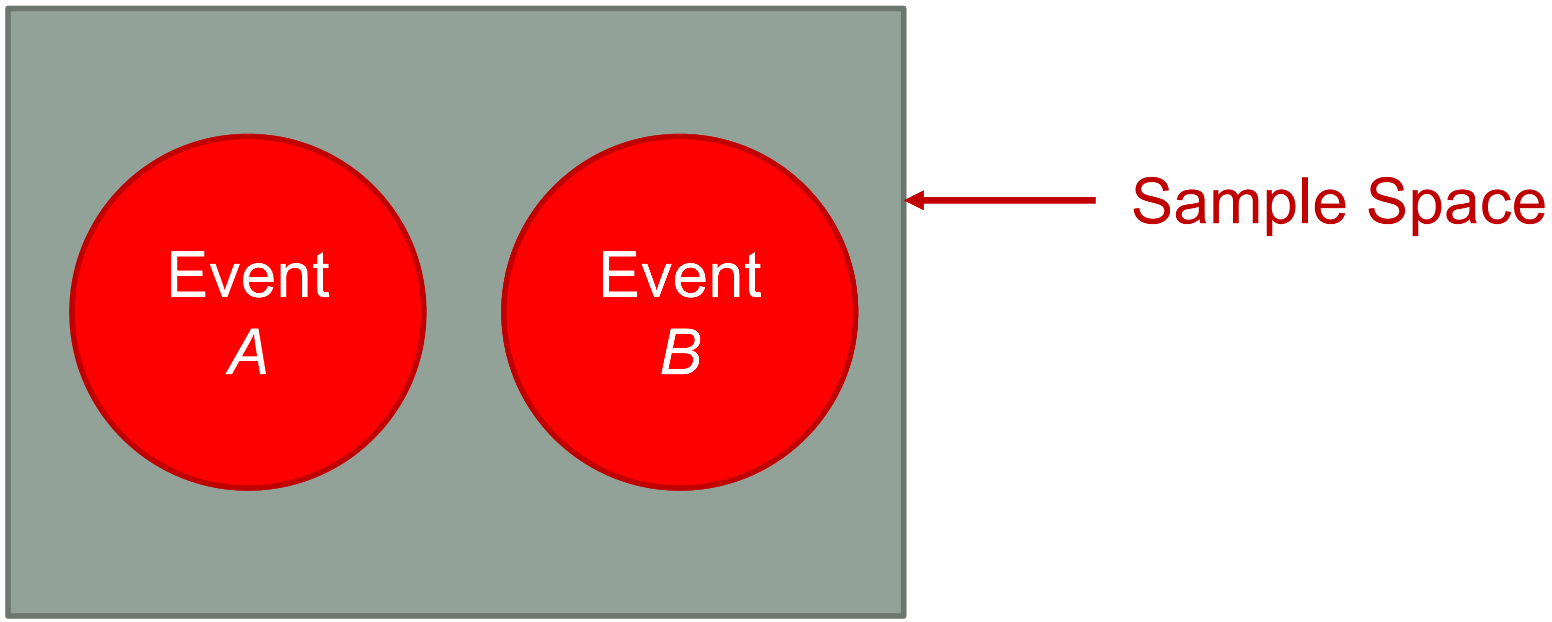
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

OR

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

Mutually Exclusive Events

- Two events are **mutually exclusive** if the events have no sample points in common – do not intersect.
- This also means that the events cannot both occur. If one event occurs, the other cannot.



Addition Law – Mutually Exclusive Events

- The **addition law** provides a way to compute the union of events *A* and *B*:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

If two events are mutually exclusive they do not intersect.

$$P(A \cup B) = P(A) + P(B)$$

Conditional Probabilities

- The probability of an event given that another event has occurred is called a **conditional** (or **joint**) **probability**.
- The conditional probability of *A* given that *B* has already occurred is denoted by $P(A|B)$.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Multiplication Law

- The **multiplication law** provides a way to compute the probability of the intersection of two events as long as you know the conditional probabilities:

$$P(A \cap B) = P(A|B) \times P(B)$$

OR

$$P(A \cap B) = P(B|A) \times P(A)$$

Independent vs. Dependent

- **Independent event** – two events are independent if the occurrence of one event doesn't influence the probability of the occurrence of the other event

- **Dependent event** – two events are dependent if the occurrence of one event impacts the probability of the occurrence of the other event.

Independent vs. Dependent

- **Independent event** – two events are independent if the occurrence of one event doesn’t influence the probability of the occurrence of the other event

- **Dependent event** – two events are dependent if the occurrence of one event impacts the probability of the occurrence of the other event.
  - **Mutually Exclusive** – Special case of dependent events; Two events are mutually exclusive if the occurrence of one event precludes the occurrence of a second event.

Independent Events

- If the probability of an event *A* is not changed by the existence of event *B*, then the two events are called **independent**.

$$P(A|B) = P(A) \quad \text{OR} \quad P(B|A) = P(B)$$

Multiplication Law – Independent Events

- The **multiplication law** provides a way to compute the probability of the intersection of two events as long as you know the conditional probabilities:

$$P(A \cap B) = P(A) \times P(B)$$

ONLY IF EVENTS ARE INDEPENDENT

Conditional vs. Marginal Probabilities

- **Marginal probabilities** can be thought of as unconditional probabilities – just probabilities of events without any condition.
- For example, let’s look at promotion of people at a company with advanced degrees vs. those who don’t have them.

	Adv. Degree - YES	Adv. Degree - NO	Total
Promoted	288	36	324
Not Promoted	672	204	876
Total	960	240	1200

Conditional vs. Marginal Probabilities

- **Marginal probabilities** can be thought of as unconditional probabilities – just probabilities of events without any condition.
- For example, let's look at promotion of people at a company with advanced degrees vs. those who don't have them.

	Adv. Degree - YES	Adv. Degree - NO	Total
Promoted	288	36	324
Not Promoted	672	204	876
Total	960	240	1200

$$P(Adv.Degree) = \frac{960}{1200} = 0.8$$

Conditional vs. Marginal Probabilities

- **Marginal probabilities** can be thought of as unconditional probabilities – just probabilities of events without any condition.
- For example, let's look at promotion of people at a company with advanced degrees vs. those who don't have them.

	Adv. Degree - YES	Adv. Degree - NO	Total
Promoted	288	36	324
Not Promoted	672	204	876
Total	960	240	1200

$$P(Promotion) = \frac{324}{1200} = 0.27$$

Conditional vs. Marginal Probabilities

- **Marginal probabilities** can be thought of as unconditional probabilities – just probabilities of events without any condition.
- For example, let's look at promotion of people at a company with advanced degrees vs. those who don't have them.

	Adv. Degree - YES	Adv. Degree - NO	Total
Promoted	288	36	324
Not Promoted	672	204	876
Total	960	240	1200

$$P(Promotion|No\ Adv.Degree) = \frac{36}{240} = 0.15$$

Conditional vs. Marginal Probabilities

- **Marginal probabilities** can be thought of as unconditional probabilities – just probabilities of events without any condition.
- For example, let's look at promotion of people at a company with advanced degrees vs. those who don't have them.

	Adv. Degree - YES	Adv. Degree - NO	Total
Promoted	288	36	324
Not Promoted	672	204	876
Total	960	240	1200

$$P(Promotion|Adv.Degree) = \frac{288}{960} = 0.30$$

Example

Number of Credit Cards Per Age Group

	Age Group				
Credit Cards	20 – 29	30 – 39	40 – 49	50+	**Total**
0	56	24	33	97	**210**
1 – 2	182	273	187	387	**1029**
3 – 4	147	358	413	212	**1130**
5 – 6	65	195	154	157	**571**
7 – 8	32	101	98	88	**319**
9+	10	67	123	11	**211**
Total	**492**	**1018**	**1008**	**952**	**3470**

Example

- Determine the probability of the following:
  1. Person is between the age of 20-29 and owns 3-4 credit cards

  2. Person is between the age of 20-29 or owns 3-4 credit cards

Example

- Determine the probability of the following:
  1. Person is between the age of 20-29 and owns 3-4 credit cards

$$\frac{147}{492} \times \frac{492}{3470} = \frac{147}{3470} = 0.0424$$

  2. Person is between the age of 20-29 or owns 3-4 credit cards

$$\frac{492}{3470} + \frac{1130}{3470} - \frac{147}{3470} = \frac{1475}{3470} = 0.4251$$

Example

- Determine the probability of the following:

3. Person owns 5-6 credit cards

4. Person owns at least one credit card

Example

- Determine the probability of the following:

3. Person owns 5-6 credit cards

$$\frac{571}{3470} = 0.1646$$

4. Person owns at least one credit card

$$1 - \frac{210}{3470} = \frac{3260}{3470} = 0.9395$$

Example

- Determine the probability of the following:

5. Person owns 1-2 credit cards given they are between the age of 30-39

6. Person is above the age of 40 given they own 9 or more credit cards

Example

- Determine the probability of the following:

5. Person owns 1-2 credit cards given they are between the age of 30-39

$$\frac{273}{1018} = 0.2682 \qquad \text{OR} \qquad \frac{\left(\frac{273}{3470}\right)}{\left(\frac{1018}{3470}\right)} = 0.2682$$

6. Person is above the age of 40 given they own 9 or more credit cards

$$\frac{123 + 11}{211} = 0.635$$