

Probability Basics

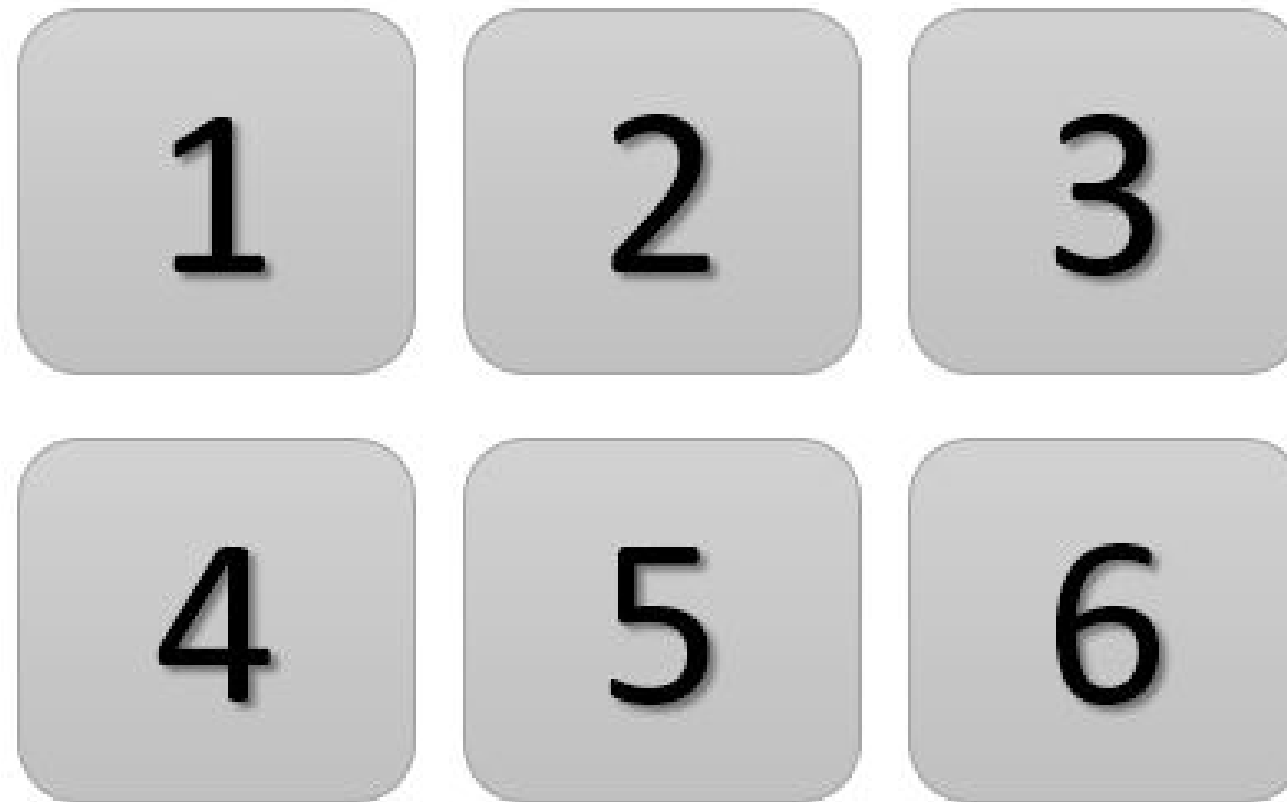
STATISTICAL SIMULATION IN PYTHON



Tushar Shanker
Data Scientist

Sample Space

Sample Space S : Set of all possible outcomes



Probability

Sample Space S : Set of all possible outcomes

Probability $P(A)$: Likelihood of event A

- $0 \leq P(A) \leq 1$
- $P(S) = 1$ *eg.* $P(H) + P(T) = 1$

Probability

Sample Space S : Set of all possible outcomes

Probability $P(A)$: Likelihood of event A

- $0 \leq P(A) \leq 1$
- $P(S) = 1$ eg. $P(H) + P(T) = 1$



Mutually Exclusive Events

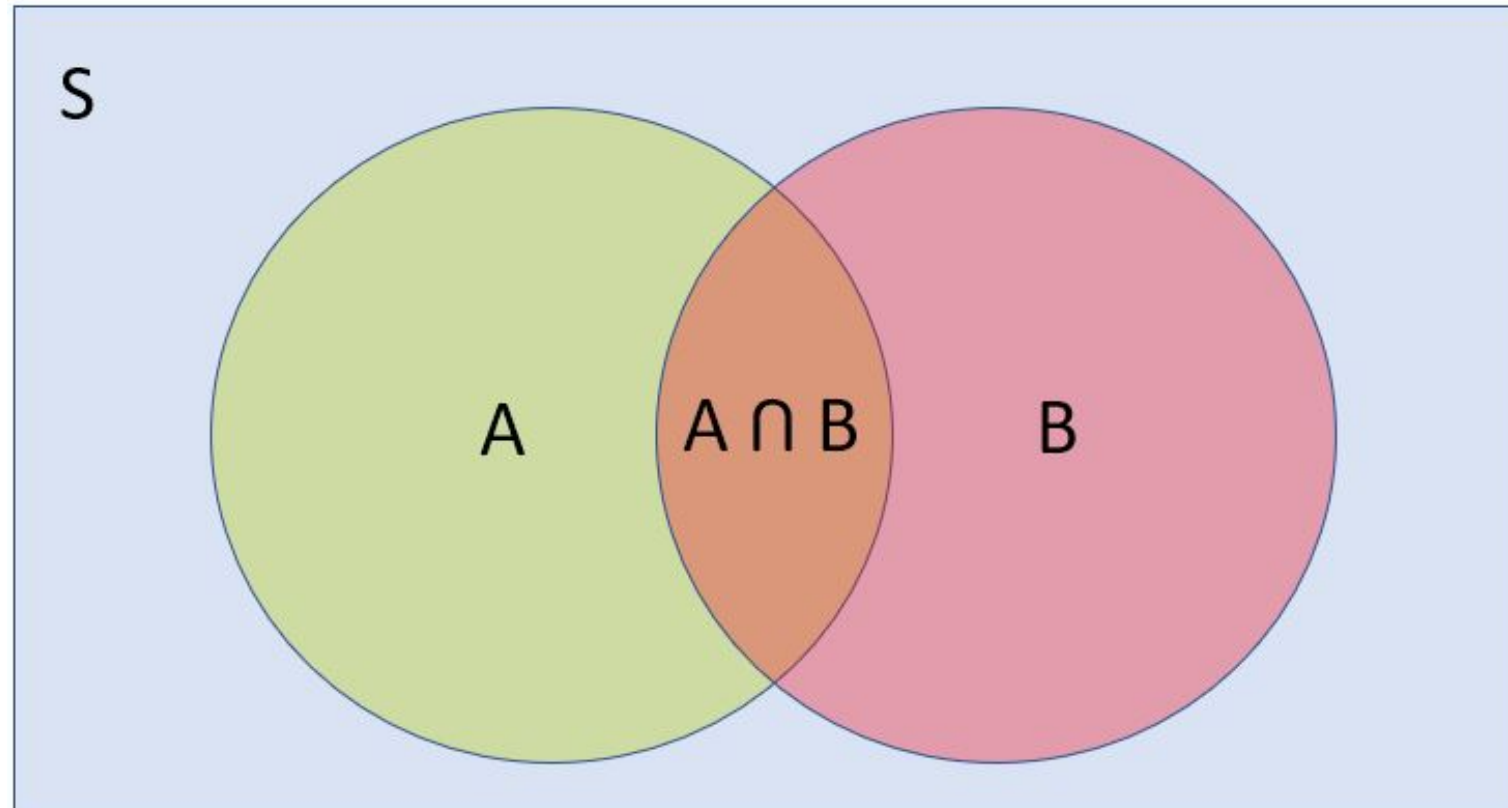
Sample Space S : Set of all possible outcomes

Probability $P(A)$: Likelihood of event A

- $0 \leq P(A) \leq 1$
- $P(S) = 1$
 - $P(H) + P(T) = 1$
- For mutually exclusive events A and B :
 - $P(A \cap B) = 0$
 - $P(A \cup B) = P(A) + P(B)$

Probability

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



Using Simulation for Probability Estimation

Steps for Estimating Probability:

1. Construct sample space or population.
2. Determine how to simulate one outcome.
3. Determine rule for success.
4. Sample repeatedly and count successes.
5. Calculate frequency of successes as an estimate of probability.

Let's practice!

STATISTICAL SIMULATION IN PYTHON

More Probability Concepts

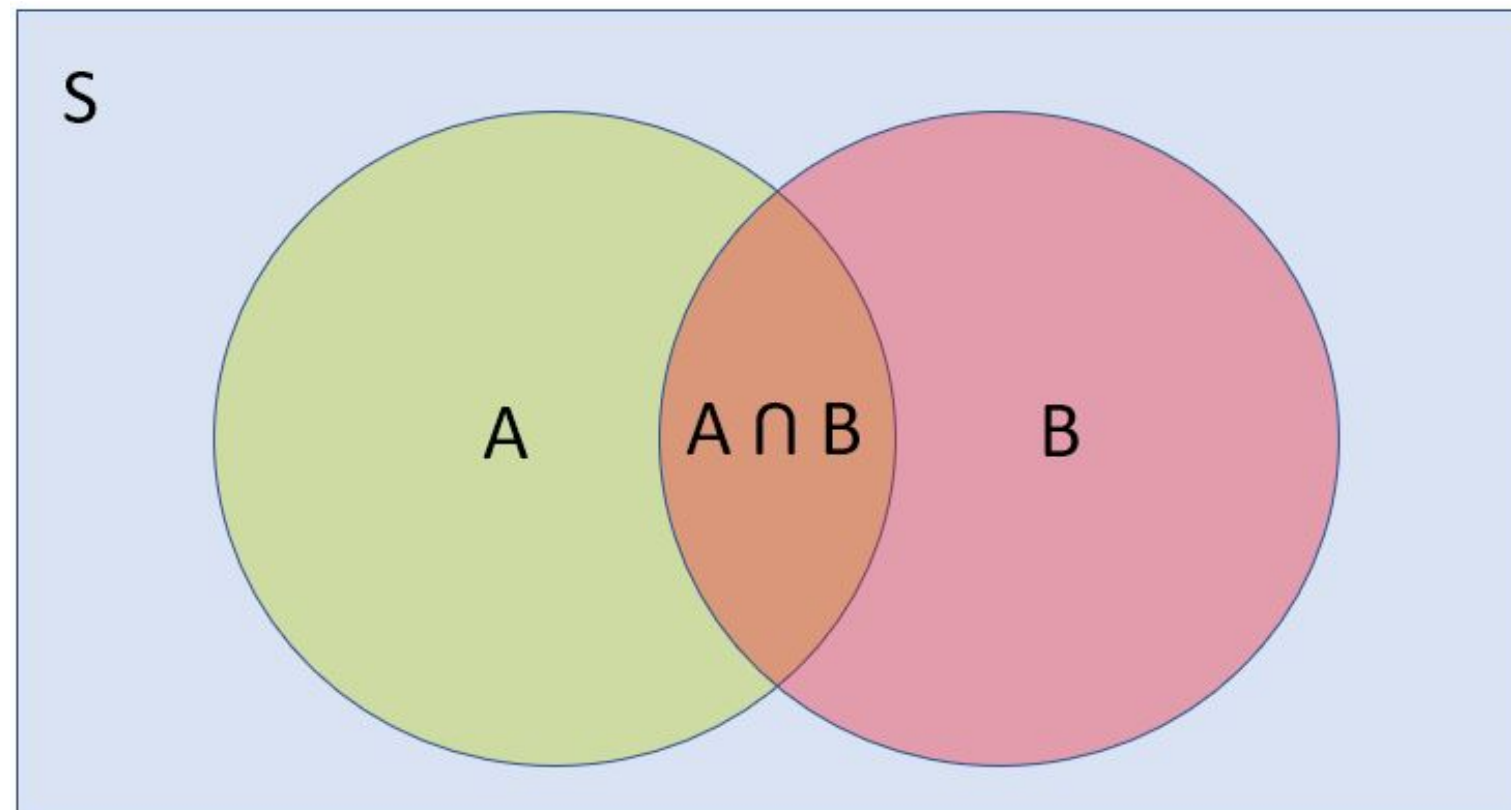
STATISTICAL SIMULATION IN PYTHON



Tushar Shanker
Data Scientist

Conditional Probability

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



Conditional Probability

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

Bayes Rule

- Conditional Probability

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

- **Bayes' rule:** $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

Independent Events

- Independent Events

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

- Conditional Probability: $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$

Solar Panels & Clean Vehicles

- Number of houses = 150

	Solar Panels	No Solar Panels	
Hybrid / EV	30/150	50/150	
No Hybrid / EV	10/150	60/150	
			150

Solar Panels & Clean Vehicles

$$P(\text{Solar}) = P(\text{Solar} \cap \text{Hybrid, EV}) + P(\text{Solar} \cap \text{No Hybrid, EV}) = \frac{30}{150} + \frac{10}{150} = \frac{40}{150}$$

	Solar Panels	No Solar Panels	
Hybrid / EV	30/150	50/150	80/150
No Hybrid / EV	10/150	60/150	70/150
	40/150	110/150	150/150

Solar Panels & Clean Vehicles

$$P(\text{Solar}|\text{Hybrid, EV}) = \frac{P(\text{Solar} \cap \text{Hybrid, EV})}{P(\text{Hybrid, EV})} = \frac{30}{80} = 0.375$$

	Solar Panels	No Solar Panels	
Hybrid / EV	30/150	50/150	80/150
No Hybrid / EV	10/150	60/150	70/150
	40/150	110/150	150/150

Let's practice!

STATISTICAL SIMULATION IN PYTHON

Data Generating Process

STATISTICAL SIMULATION IN PYTHON

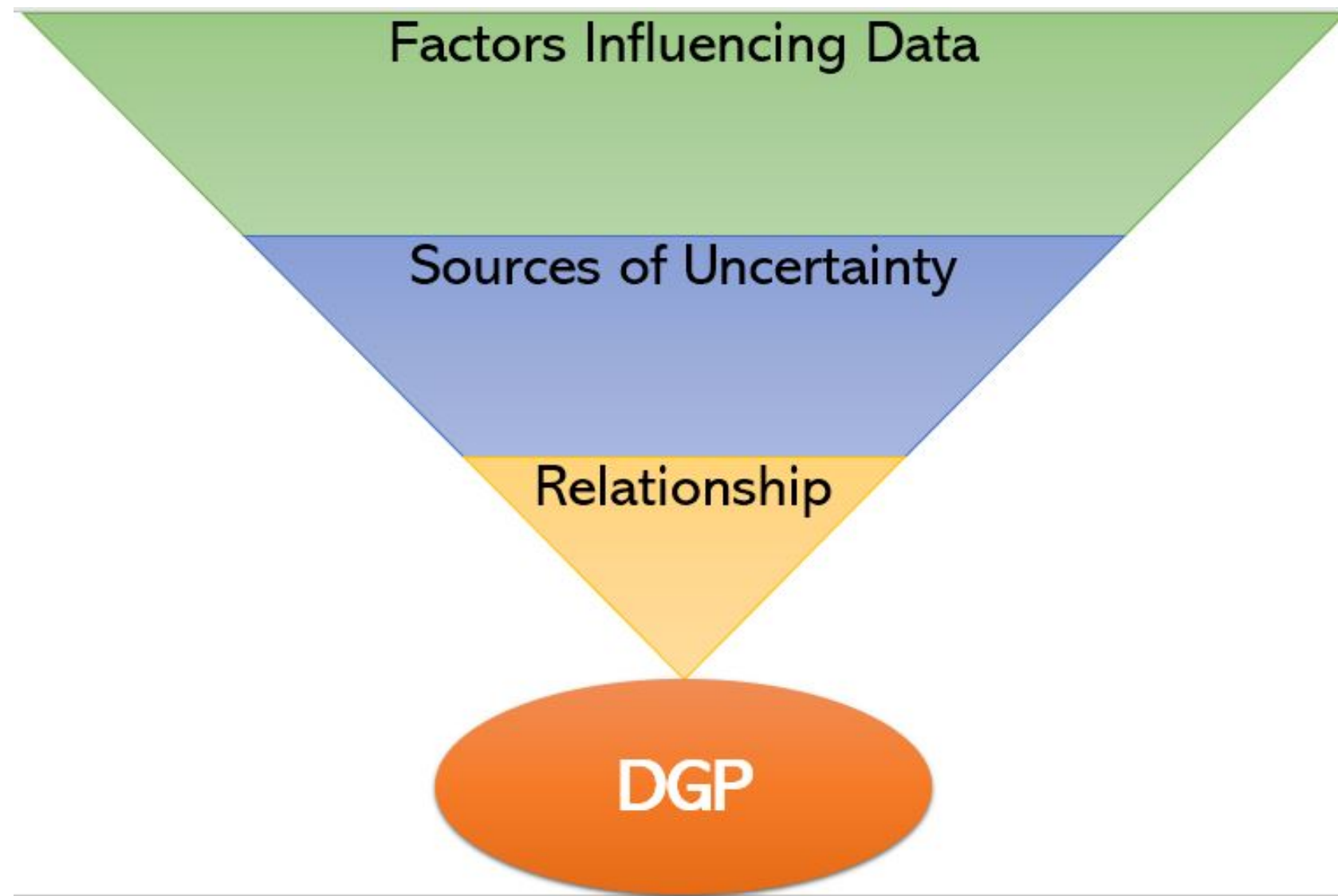


Tushar Shanker
Data Scientist

Simulation Steps

1. Define Possible Outcomes for Random Variables.
2. Assign Probabilities.
3. Define Relationships between Random Variables.

Data Generating Process



Cricket



India

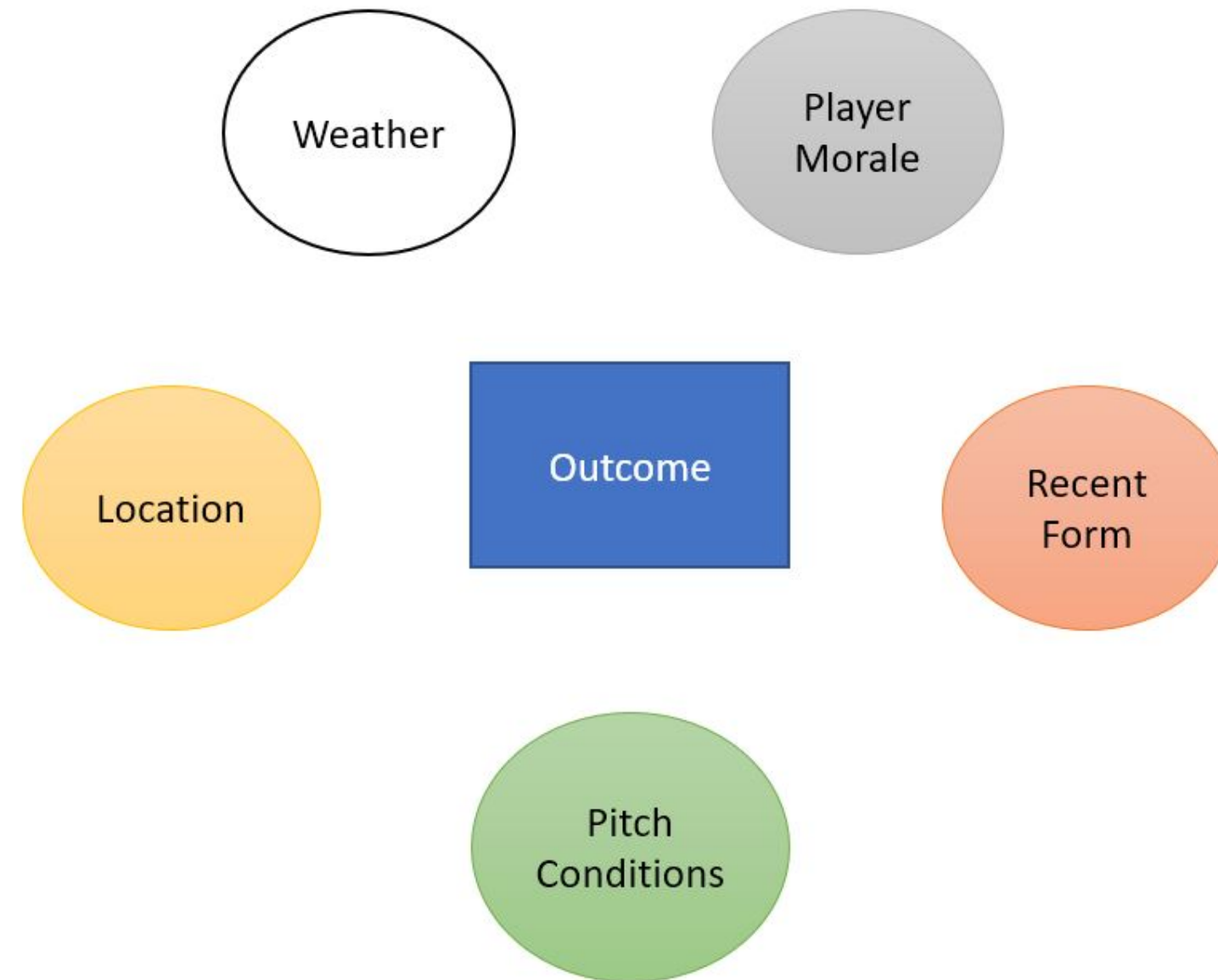
VS



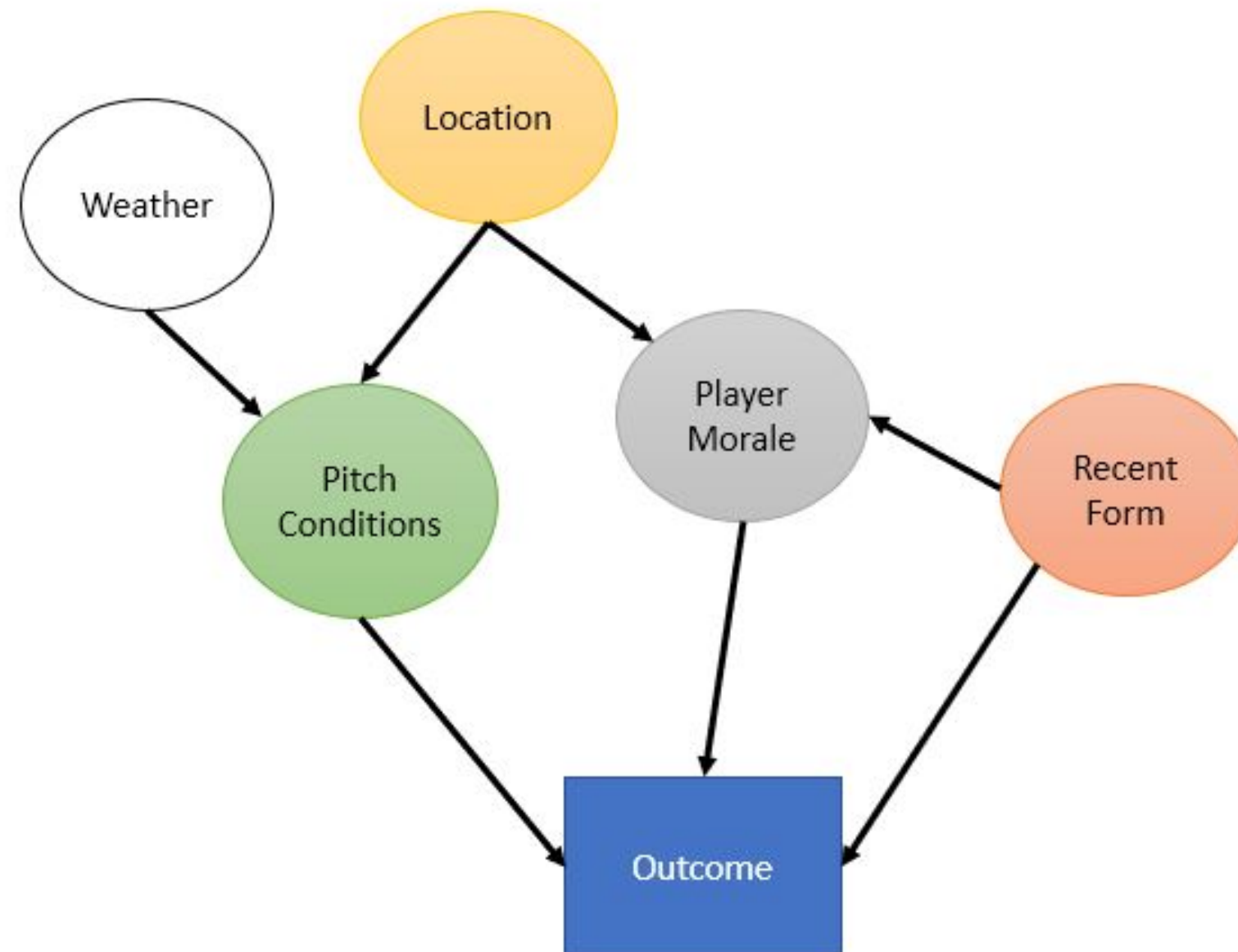
England

¹ Source: Wikipedia

Cricket



Cricket



Let's practice!

STATISTICAL SIMULATION IN PYTHON

eCommerce Ad Simulation

STATISTICAL SIMULATION IN PYTHON

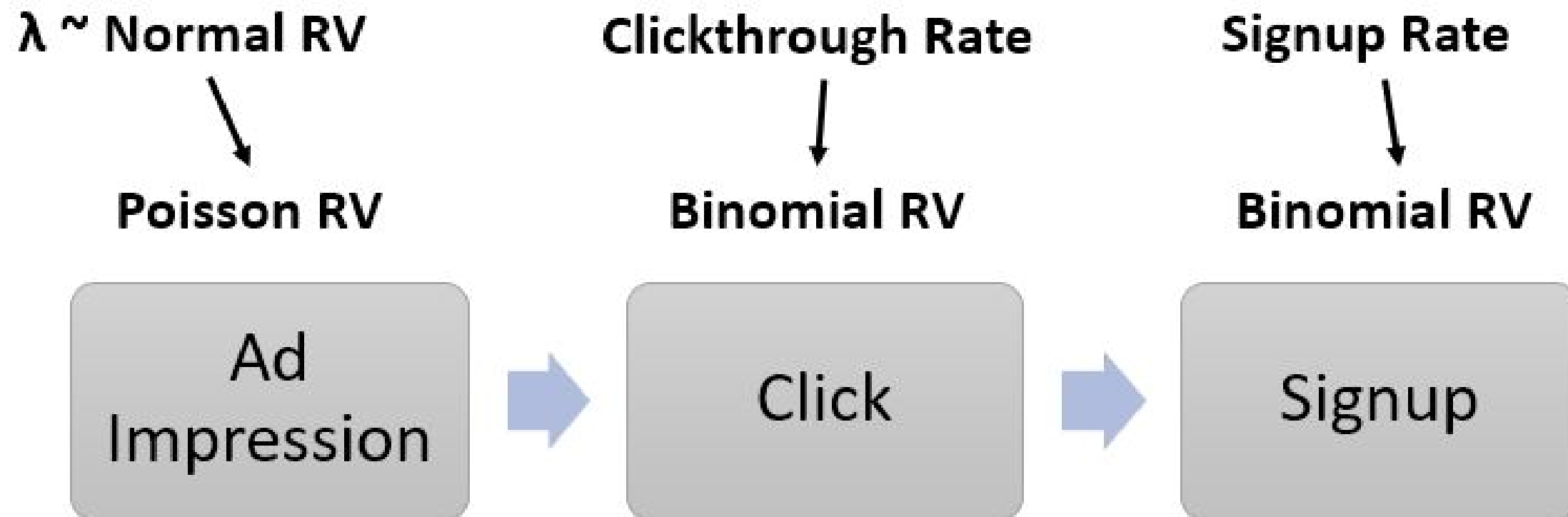


Tushar Shanker
Data Scientist

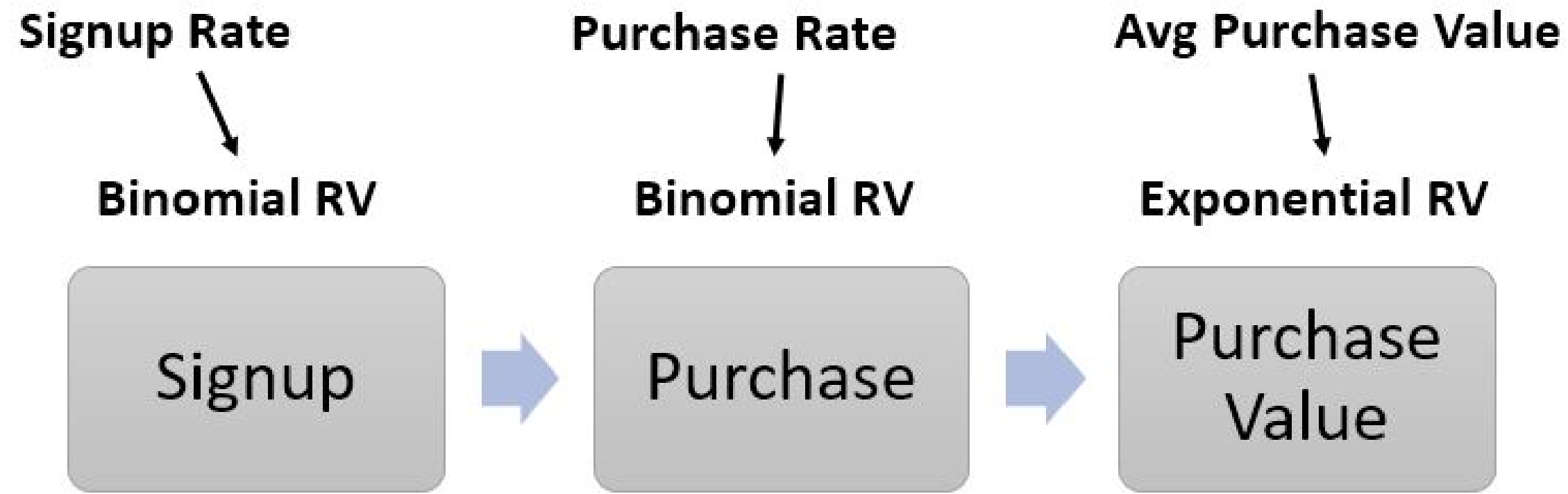
eCommerce Funnel



Signup Flow



Purchase Flow



Let's practice!

STATISTICAL SIMULATION IN PYTHON