Q1) what is meant by Posterior Sampling?
Posterior sampling, also known as **Bayesian sampling** or **Thompson sampling**, is a technique used in Bayesian statistics and machine learning to make decisions or predictions by sampling from the posterior distribution of a model's parameters.

Key Concepts:

1. **Prior Distribution**: This represents the initial beliefs or knowledge about the parameters before any data is observed. It's a probability distribution that reflects what you know (or assume) about the parameters before seeing any evidence.
2. **Likelihood**: This represents the probability of observing the data given a specific set of parameters. It's derived from the model and the data.
3. **Posterior Distribution**: This is the updated belief about the parameters after observing the data. It combines the prior distribution with the likelihood using Bayes' Theorem:

   $$\text{Posterior} = (\text{Likelihood} \times \text{Prior}) / \text{Evidence}$$

   The posterior distribution reflects what is known about the parameters after considering the data.
4. **Posterior Sampling**: This involves drawing samples from the posterior distribution. Each sample represents a plausible set of parameter values given the observed data. These samples can then be used for various purposes, such as making predictions, estimating uncertainty, or making decisions.

Example in Bayesian Inference:

Consider a simple case where you're trying to estimate the probability of a coin landing heads (parameter θ):

1. **Prior**: Before flipping the coin, you might assume that any value of θ (the probability of heads) is equally likely, so you use a uniform prior.
2. **Data (Likelihood)**: You flip the coin several times and observe the results (e.g., 7 heads in 10 flips).
3. **Posterior**: After observing the data, you update your beliefs about θ using Bayes' Theorem, resulting in a posterior distribution that might now favor values of θ closer to 0.7.
4. **Posterior Sampling**: You can then draw samples from this posterior distribution to estimate the probability of heads in future flips, or to make decisions in a probabilistic context (e.g., in a reinforcement learning problem).

Application in Reinforcement Learning (Thompson Sampling):

In the context of reinforcement learning, particularly in multi-armed bandit problems, **Thompson sampling** is a strategy where an agent samples from the posterior distribution of the expected rewards for each action (arm) and then selects the action that maximizes the sampled reward. This method balances exploration and exploitation by naturally exploring actions with uncertain outcomes and exploiting actions with known, high rewards.