1. **Recommended Approach for Management**

* **Evaluate which of the simulated strategies yields the highest** **overall conversion rate while considering the balance between exploration (learning) and exploitation (earning).**

*Recommendation:*

For the given conversion rates of (A= 9%, B = 15%, C = 10%), the Epsilon-Greedy strategy with an epsilon of 0.2 yields the highest overall conversion rate while considering the balance between exploration and exploitation, with a total conversion rate of 14.49% over 100 days (rounds).

*Observations:*

This can mainly be attributed to the design of the strategy which incorporates exploitation almost immediately due to the high value of epsilon.

A site is chosen based on the best performing site up till a given round. Because the outcomes differ significantly, the best performing arm is chosen majority of the times, hence choosing the optimal outcome majority of the times.

This is almost 0.3% better than the greedy strategy employed.

Due to the high value of epsilon, exploitation is favoured over exploration and hence it converges to the best site fairly quickly ensuring maximum exploitation, choosing the best outcome (including the learning day).

This reduces the number of days wasted in learning and then earning (as in the greedy algorithm).

Even when the conversion rates of sites change to (A= 9%, B = 11%, C = 10%), the MAB’s produce a better performance compared to traditional methods. Exploration and exploration become much more important in this scenario due to the reduced differences in outcomes.

1. **ϵ-Greedy Performance with Altered Exploitation Rate:**

* **Analyze how the ϵ-Greedy algorithm performs when the exploitation rate is reduced to 50%. Discuss the trade-off between exploration and exploitation and its effect on overall conversions.**

*Observations:* When the exploitation rate is reduced from 80% to 30%, the overall conversion rate drops by about 1%.

This can be mainly attributed to the increase in the probability of exploration rounds, which may lead to sub-optimal decisions (Allocating traffic to site A or C which have lower conversion rates compared to site B).

*Recommendations:* In the given simulation, we have a pre-knowledge of the conversion rates which remain relatively stable for all the sites and the best site (Site B) has a significantly higher conversion rate compared to the rest. Under such a situation where we have prior knowledge of the best decision, it would be sub-optimal to allocate a lower % to exploitation. Hence, a lower epsilon is recommended.

But, in real-world scenarios, where the best decision is unknown, and the outcomes are not as constant and vary over time or if the outcomes are not significantly different, allocating a higher weightage for exploration in the initial phase can help improve the overall outcomes significantly.

1. **Impact on Softmax with Increased Weight for Better Performing Arms:**

* **Observe the changes in the Softmax strategy outcomes when the weight for better-performing arms is doubled. Explain how this adjustment affects the selection probability and the speed of convergence to the best-performing version.**

*Explanation:* The softmax algorithm assigns probabilities to each arm based on their estimated values, and arms are then chosen probabilistically according to these probabilities.

The softmax algorithm introduces a stochastic element into the arm selection process, allowing for exploration of suboptimal arms while still favoring arms with higher estimated mean rewards. The temperature parameter (*Tau*) controls the level of exploration, and adjusting it can influence the balance between exploration and exploitation. Higher temperatures encourage more exploration, while lower temperatures prioritize exploitation of the arms with higher estimated rewards.

*Observations:* For the Softmax algorithm, the initial weight (Tau) has been taken as 0.25 as initial weight probability during the exploitation phase is given as 25%. In the second iteration, the weight (Tau) has been reduced to 0.125. If the weight for better performing arm is doubled, Tau is reduced by half (assuming the weight here refers to (rate/tau) => 2\*rate/tau = rate/tau/2)

Reduction of Tau signifies more weight for the best performing arm hence promoting exploitation. It can be observed that the total conversion rate increases from 11.75% to 12.21%. Also, It can be observed that the exploration of the other sub-optimal arms A and C is reduced.

*Recommendations*: In context of the given simulation a higher Tau allows for more exploration and hence the probability of selecting a sub optimal arm is relatively high compared to a simulation with a lower Tau. The convergence to best outcome is slower for high values of tau compared to lower values.

In our context, because the outcomes are stable, a lower value of tau would be recommended. If the outcomes were more indeterministic, or not significantly different, a high value of tau would be favoured.

1. **Impact of Reduced Conversion Rate for Site B on MABs:**

* **Assess how the results and recommendations change if site B's conversion rate is modified to 11% instead of 15%. Discuss the implications for the effectiveness of MABs compared to traditional A/B/C tests, particularly in scenarios where the difference in performance between options is less pronounced.**

*Observations:* As a consequence of reducing the conversion rate for site B from 15% to 11%, it can be observed that there in an increase in exploration compared to the former scenario.

*Recommendation:* Under the new changes, it can be observed that the epsilon-greedy algorithm performs much better than the greedy algorithm. Epsilon-Greedy would be the recommended strategy in this scenario.

*Effectiveness of MAB’s in comparison to A/B/C:* In this scenario, where the outcomes for the 3 sites are not vastly different, we can clearly observe that the MAB’s outperform the traditional methods due to the explore-exploit strategies. They help implement the best solution at a given time and the learning like process helps make a choice that optimizes outcomes.

Another important advantage of the MAB’s is that unlike traditional methods, there is no need to wait for learning (exploration) to take action (exploitation). The exploration feature designed into the algorithms help take immediate action (exploit) with the learning (exploration) continuing on the side.

Under scenarios where the outcomes for the various sites are significantly different, traditional greedy strategies after conducting an A/B/C test may produce the best and consistent results.