Report: Determining the Optimal Bench Size for Cashiers at ApnaMart Using Monte Carlo Simulation

Submitted by: Riswana Bano
Roll no.: 22b1038
Date: 02-05-25

1. Objective

To ensure uninterrupted operations across ApnaMart stores in Ranchi, it's crucial to have an adequate number of standby (bench) cashiers who can step in when regular cashiers are absent. The goal is to determine the optimal number of bench cashiers required to cover absenteeism effectively, ensuring that 95% of operational days are adequately staffed, to ensure safety, increase efficiency and bring them into use.

2. Data Available and Preprocessing

Data Source: Attendance records for each store over a specified period for 31 days.

Key Metrics:

Absent Days: Number of days a cashier was absent due to reasons like sick leave, casual leave, etc.

Operational Days: Total number of days the store was operational.

Probability of absent days for each stored by calculated by: -

Absence Probability Calculation:

Absence Probability = {Absent Days/Operational Days}

This metric provides the likelihood of a cashier being absent on any given day for each store.

All the data processed and calculated is done in excel file.

3. Data Transformation: Converting Attendance to Numeric Values

To facilitate quantitative analysis and simulation, the categorical attendance data was transformed into numeric values i.e.,

Present: Assigned a value of `1`

Absent (including Sick Leave, Week Off, Casual Leave, Privileged Leave): Assigned a value of `0`

This binary representation simplifies the computation of absence probabilities and is a common practice in data analysis to handle categorical variables.

4. Monte Carlo Simulation: Rationale and Implementation

Why Monte Carlo Simulation?

For this type of problem optimization methods like Monte Carlo simulation is fitted because Monte Carlo simulation is a statistical technique that utilizes random sampling to model and understand the impact of risk and uncertainty in prediction and forecasting models. It's particularly useful when dealing with systems influenced by random variables and when analytical solutions are complex or infeasible.

Implementation Steps:

1. Simulation Parameters:

- **Capturing Variability in Absenteeism:**

  Monte Carlo simulations rely on random sampling to model uncertainty. By simulating 10,000 days, we can capture a wide range of possible absentee scenarios, including rare events. This extensive sampling ensures that our model accounts for the inherent variability in daily cashier attendance.
- **Ensuring Statistical Reliability**

  The accuracy of Monte Carlo simulations improves with the number of iterations. Specifically, the standard error of the estimate decreases proportionally to the inverse square root of the number of simulations. Therefore, increasing the number of iterations from 1,000 to 10,000 reduces the standard error by approximately a factor of $\sqrt{10}$, enhancing the reliability of our percentile estimates.
- **Balancing Precision and Computational Efficiency:**

  While more iterations can lead to more precise estimates, they also require more computational resources. In our case, 10,000 simulations provided a good balance, offering sufficient precision for decision-making without excessive computational time.
- **Industry Standards and Best Practices:**

  In various fields, including finance and project management, running 10,000 simulations is a common practice for Monte Carlo analyses. This number is often sufficient to achieve stable and reliable results, especially when estimating percentiles like the 95th percentile, which is critical for planning purposes.
- **Focus on Extreme Percentiles:**

Estimating extreme percentiles, such as the 95th percentile, requires a larger number of simulations to ensure accuracy. With 10,000 iterations, we can confidently determine the number of bench cashiers needed to cover absenteeism on 95% of operational days, thereby minimizing the risk of understaffing.

Number of Simulations: 10,000 days to represent a wide range of possible scenarios. 10,000 days of possible scenario was generated in excel.

Stores: Each store's absenteeism is simulated based on its calculated absence probability.

2. Simulation Process:

For each simulated day:

Generate a random number for each store.

If the random number is less than the store's absence probability, mark the cashier as absent for that day.

Count the total number of absent cashiers across all stores for each simulated day.

3. Result Analysis:

Compiled the distribution of total absent cashiers over the 10,000 simulated days.

Determined the 95th percentile of this distribution to identify the number of absent cashiers that would not be exceeded on 95% of the days.

5. Results

95th Percentile: The simulation indicates that on 95% of days, the number of absent cashiers does not exceed 7.

Recommendation: Maintaining a bench of 7 standby cashiers can ensure that 95% of operational days are adequately staffed, minimizing disruptions due to absenteeism.

6. Excel Workbook Details

An Excel workbook has been prepared to provide a transparent view of the calculations and simulations:

Sheet 1: Absence Probability

Lists each store with its corresponding absent days, operational days, and calculated absence probability.

Sheet 2: Monte Carlo Simulation and Final Recommendation

Contains the simulation results for 10,000 days, showing the number of absent cashiers per day.

The **optimal bench size** is the **smallest number of extra staff** that ensures absenteeism is covered on **≥95% of the days** in at least **95% of the months**. Based on the simulation, the optimal size in our case was **7 staff.**

Table 1 95% Percentile

Metric	Value
95th Percentile (Recommended Bench Size)	7

7. Conclusion

By leveraging Monte Carlo simulation, ApnaMart can make data-driven decisions to determine the optimal number of bench cashiers required. This approach accounts for the inherent uncertainty in daily absenteeism and ensures that the majority of operational days are covered, thereby maintaining service quality and operational efficiency.