# Optimizing Revenue Streams: A Holistic Study of Grain Weight Change and Charging Practices

**Final report for the BDM capstone Project**

Submitted by

Name: Ritik Kumar Badiya

Roll number: 21f3001509

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

# Contents

# Declaration

I am working on a Project titled " Optimizing Revenue Streams: A Holistic Study of Grain Weight Change and Charging Practices". I extend my appreciation to Ganesh Warehouse for providing the necessary resources that enabled me to conduct my project.

I hereby assert that the data presented and assessed in this project report is genuine and precise tothe utmost extent of my knowledge and capabilities. The data has been gathered from primarysources and carefully analyzed to assure its reliability.

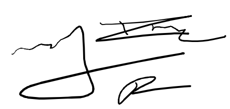
Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the information of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I understand that all recommendations made in this project report are within the context of the academic project taken up towards course fulfillment in the BS Degree Program offered by IIT Madras. The institution does not endorse any of the claims or comments.

Name : Ritik Kumar Badiya



Date : 15/11/2023

# Executive summary

The project centers on a longstanding warehouse in Makdone, owned by Mansingh Patidar for over 25 years, specializing in storing and distributing Chana, Soybeans, and Dhana. With an annual turnover of 12 lakhs, the warehouse caters to local agricultural needs and manages government contracts. The primary focus is on the B2C aspect. Challenges include fluctuating grain weights impacting rental fees, the decision to maintain a permanent labor force, and choosing between chana and wheat.

The warehouse faces issues with charging practices, specifically the timing of fees linked to grain removal. To address this, statistical analyses correlating days with grain weight loss are underway, aiming to refine charging practices based on initial or reduced weight, factoring in accrued interest. The second challenge involves exploring the feasibility of a permanent labor force, assessing competitive offers for consistency and efficiency. The third challenge revolves around choosing between chana and wheat. The analysis involves quantifying potential additional income from chana, considering overhead costs.

The project aims to improve decision-making by exploring various charging practices, such as evaluating the impact of considering the initial weight or factoring in rent interest. The overarching objective is to enhance efficiency and accuracy in warehouse operations.

# Analysis Process

**Step 01: Preprocessing**

During preprocessing, the raw data is cleaned using Excel to ensure that it is free from errors and missing values. This is a crucial step to create a reliable dataset for further analysis. Null data points are removed to enhance the quality of the dataset.

**Step 02: Variable Selection and Generation of New Variables**

In this step, careful consideration is given to selecting the most relevant variables for the analysis. The chosen variables include 'item,' 'days,' 'weight_in,' 'weight_out,' 'bags_count,' 'charge_per_kg_per_bag,' and 'charge.' These variables are considered essential for solving the problem at hand.

Additionally, new variables are generated to provide more insights:

1. **Loss:** Calculated as the difference between 'weight_in' and 'weight_out.' This variable represents the loss in weight during the specified period.
2. **Loss per Bag:** Calculated by dividing 'loss' by the number of bags, providing a normalized measure of loss.
3. **Log Days:** The logarithm of the 'days' variable. This transformation may be applied to handle situations where the relationship between 'days' and other variables is better represented on a logarithmic scale.
4. **Weighted Charging:** A variable that takes into account the product of 'weight_in,' 'days,' and 'charge_per_kg_per_bag.' This variable may represent a weighted measure of charging based on the input weight.

**Step 03: Statistical Analysis**

In this phase, we conducted two statistical tests to assess the coefficient of correlation within the dataset.

**Test 01: Correlation Coefficient between Log_Days and Total_Weight_Loss**

- **Null Hypothesis ($H_0$):** The correlation coefficient (ρ) between log_days and total_weight_loss is equal to 0 .
- **Alternate Hypothesis ($H_a$):** The correlation coefficient (ρ) between log_days and total_weight_loss is not equal to 0.

For this test, we calculated the correlation coefficient (ρ) using the formula.

$$\rho(X,Y) \qquad = E\,\frac{(X-\bar{x})(Y-\bar{y})}{\sigma x.\sigma y}$$

$$= \frac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{(\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2)}}$$

$$= \frac{\Sigma(x_i - 1.75)(y_i - 2.32)}{\sqrt{(\sum(x_i - 1.75)^2 \sum(y_i - 2.32)^2)}}$$

$= 0.733$

Here, X represents 'log_days,' Y represents 'loss/bag,' .

The t-statistic was determined using formula

$$t = \frac{\rho\sqrt{n-2}}{\sqrt{1-\rho^2}}$$

$$t = \frac{0.733\sqrt{1009-2}}{\sqrt{1-0.733^2}}$$

$$= 35.55$$

With a significance level of 9.99 and degrees of freedom set at 1007 (total data points being 1009), a critical t-statistic value of 2.58 was needed. Our objective was to compute the t-statistic and compare it to the critical value to determine whether the null hypothesis is acceptable or rejected. So we can reject the null hypothesis since (35.55 > 2.58).

**Test 02: Correlation Coefficient between Days and Total_Weight_Loss**

- **Null Hypothesis ($H_0$):** The correlation coefficient (ρ) between days and total_weight_loss is equal to 0 .
- **Alternate Hypothesis ($H_a$):** The correlation coefficient (ρ) between days and total_weight_loss is not equal to 0.

For this test, we calculated the correlation coefficient (ρ) using the formula.

$$\rho(X,Y) \quad = E\,\frac{(X-\mu_x)(Y-\mu_y)}{\sigma x.\sigma y}$$

$$= \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2)}}$$

$$= \frac{\Sigma(x_i - 1.75)(y_i - 212.70)}{\sqrt{(\sum(x_i - 1.75)^2 \sum(y_i - 212.70)^2)}}$$

$$= 0.703$$

Here, X represents 'log_days,' Y represents 'loss/bag,' .

The t-statistic was determined using formula

$$t = \frac{\rho\sqrt{n-2}}{\sqrt{1-\rho^2}}$$

$$= \frac{0.703\sqrt{1009-2}}{\sqrt{1-0.703^2}}$$

$$= 32.868$$

With a significance level of 9.99 and degrees of freedom set at 1007 (total data points being 1009), a critical t-statistic value of 2.58 was needed. Our objective was to compute the t-statistic and compare it to the critical value to determine whether the null hypothesis is acceptable or rejected. So we can reject the null hypothesis since (32.868 > 2.58).

By conducting these two tests, we can conclude that there is a significant correlation between the variable "days" and the "loss in weight." Therefore, we are now inclined to proceed with assuming a linear relationship between the number of days and the corresponding weight loss. While the relationship showed strength when considering the logarithm of days, for further analysis, we will focus on the linear association between days and weight loss.

**Step 04: Consideration of Two Scenarios**

- **Scenario 01: Weighted Charging** In this scenario, the analysis considers revenue at the end, utilizing 'weight_in' instead of 'weight_out.' This scenario may provide insights into the impact of using the initial weight for charging calculations.
- **Scenario 02: Interest Charging** This scenario introduces the consideration of interest on current charging from the date when the grain was kept inside. It explores the impact of incorporating interest into the charging practices.

**Step 05: Comparing Two approaches**

To compare two approaches, we aim to determine an equivalent monthly interest rate (I). In simpler terms, we utilize the linear regression method to find the best-fit (I) that optimizes the solution. Once we have equivalent (I) values, we can compare them with the second approach, where interest is taken into account. By comparing the interest values, we can make informed decisions on which approach to adopt.

Here is the procedure:

**i) current revenue pattern:**

*current_charging(CC) = weight_out(WO) * days * charge_per_kg_per_bag(rate)*

**ii) two new approaches:**

*a) weighted_charging(WC) = current_charging + days * loss * charge_per_kg_per_bag* [consideringweight_in(more weight) ]

*b) intrest_charging(IC) = current_charging + (current_charging * days * (I/30) )/100*

[current charging + interest, I is monthly interest] to compare two methods try to find best I that fits weighted_charging:

→ *WC = IC*

→ *CC + days * loss * rate = CC + CC * days * (I/30) * (1/100)*

→ *WO * rate = CC * I * (1/3000)*

→ *loss * rate = WO * days * rate * I * (1/3000)*

→ *loss = WO * days * I * (1/3000)*

from line fitting (y=mx) m is given by:

m=(Σy) / (Σx)

so here,

$$I = \left(\frac{\Sigma loss}{\Sigma(days \cdot WO)}\right) \cdot \left(\frac{1}{3000}\right)$$

$$\Sigma loss = 71491.7$$

$$\Sigma(days \cdot WO) = 767506073.5$$

$$\left(\frac{71491.7}{767506073.5}\right) \cdot \left(\frac{1}{3000}\right)$$

Which implies: I = 0.279pm
=(3.35pa).

**Problem Solving Steps for Permanent Employee Compensation:**

**Step 1: Define Variables**

- In this initial step, it's crucial to identify and define the key variables that play a role in the compensation scenario. These variables include:
    1. Start date (sdate): The date when the employee begins their service.
    2. End date (edate): The date marking the end of the employment period.
    3. Bags count: A quantitative measure of the work output, representing the number of bags handled by the employee.

**Step 2: Table Slicing and Rejoining**

- Recognizing the need for labor at both the start and end dates is essential. Since there's only one entry for both dates, a careful process of table slicing and rejoining is undertaken.
- The objective is to create a dataset where there are distinct entries for both the start and end dates, with the 'bags count' value replicated for each date.

- This results in the creation of two key variables: 'date' and 'bags count,' facilitating further analysis.

**Step 3: Derive New Variable**

- Introduce a new variable to enhance the dataset and provide additional insights:
    1. Labour cost (labour_cost): Calculated by multiplying the 'bags count' by a predetermined rate, in this case, 3. This variable represents the monetary cost associated with the labor input.

**Step 4: Calculate Monthly Labour Costs**

- Move on to analyzing the labor cost at different points throughout the year. This involves breaking down the annual labor cost into monthly segments.
- The aim is to understand how the labor cost fluctuates over the months and to identify any patterns or trends.
- Calculate the average monthly labor cost, which provides a baseline for assessing the proposed pay structure.

**Step 5: Analyze Proposed Pay**

- With a comprehensive understanding of the labor cost dynamics, critically evaluate the proposed pay structure.
- Consider economic factors, including industry standards, company budget constraints, and employee expectations.
- Determine whether the proposed pay is economically feasible and competitive in the market.
- Make informed decisions on whether adjustments are necessary to ensure a fair and acceptable compensation structure for the employees.

# Results and Finding

**Results of considering linear relations:**

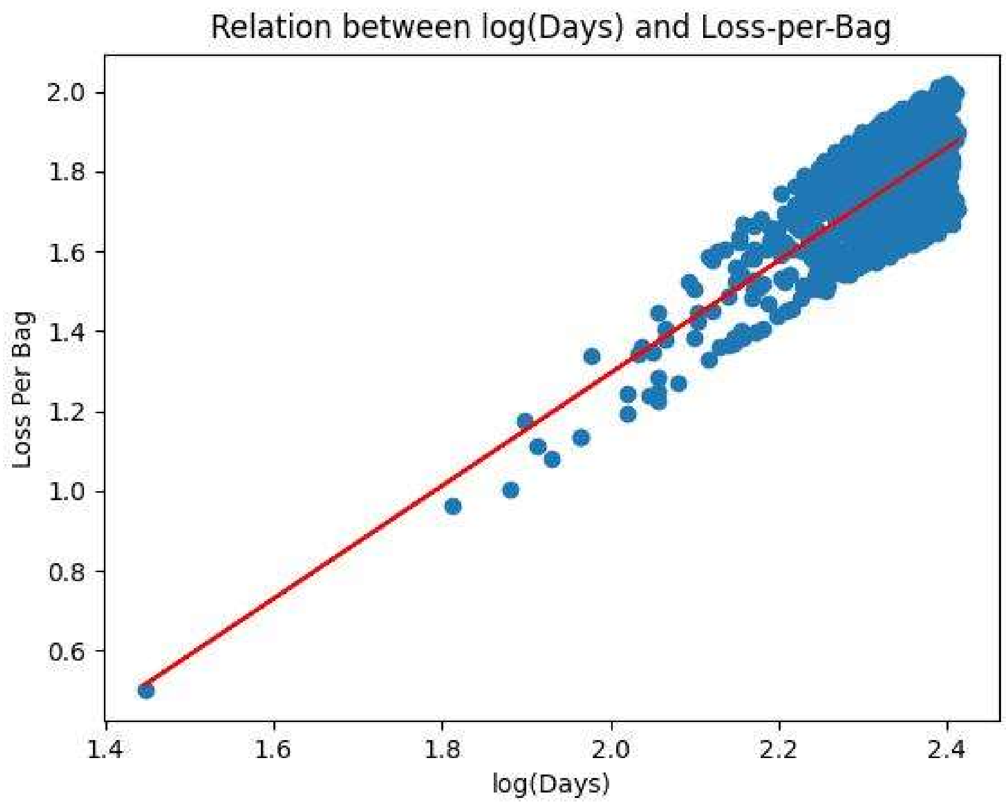



**Fig-01 : Relation between log(Days) and Loss-per-Bag**

We conducted a comprehensive analysis to explore the correlation between the logarithm of the number of days and the loss-per-bag metric. The results of our analysis yielded a correlation coefficient of 0.73. Moreover, we applied a significance test at a 9.99% confidence level and found evidence to reject the null hypothesis, indicating that there is a significant correlation between the number of days and the loss-per-bag metric.

Following the correlation analysis, we proceeded to conduct a regression analysis to further understand the relationship between these variables. The regression analysis unveiled an intercept of -1.53 and a slope of 1.41. These coefficients provide quantitative insights into the linear relationship between the logarithm of the number of days and the associated loss-per-bag. In essence, the intercept represents the estimated loss-per-bag when the logarithm of days is zero, while the slope denotes the rate of change in the loss-per-bag for each unit increase in the logarithm of days. This analytical approach enhances our understanding of the nuanced quantitative dynamics between the two variables under consideration.

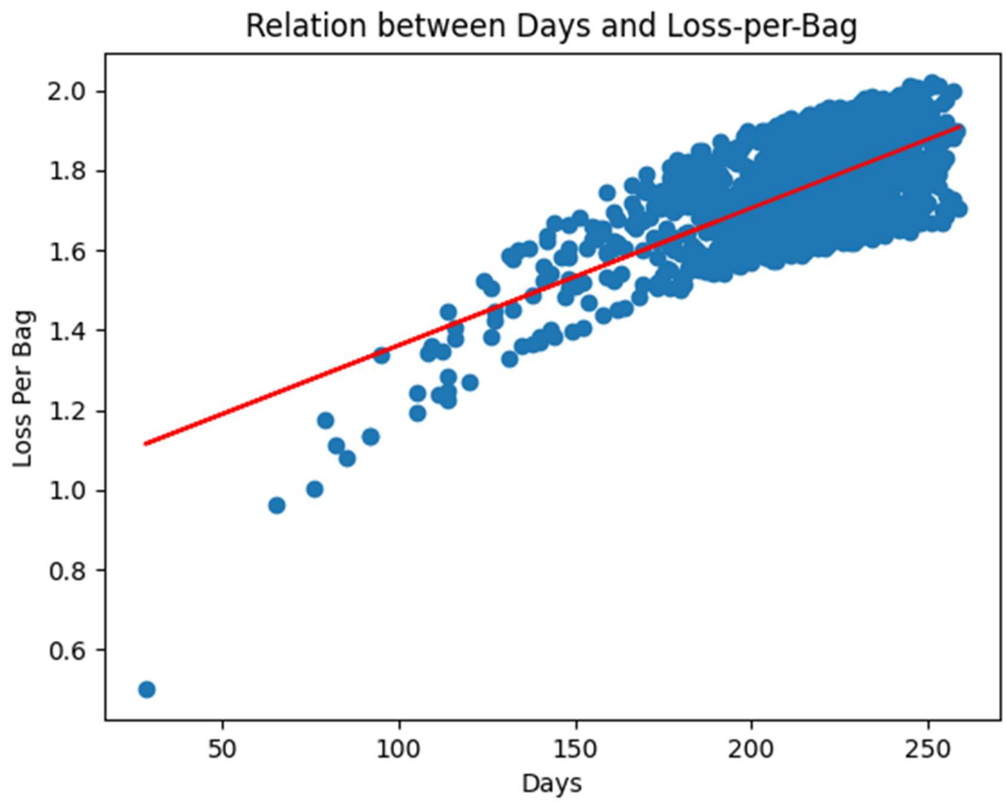



**Fig-02 : Relation between Days and Loss-per-Bag**

We performed an extensive examination to investigate the connection between the the duration in days and the loss-per-bag metric. The analysis produced a correlation coefficient of 0.70. Additionally, employing a significance test at a 9.99% confidence level led to the rejection of the null hypothesis, indicating a noteworthy correlation between the number of days and the loss-per-bag metric.

Subsequent to the correlation analysis, a regression analysis was conducted to delve deeper into the relationship between these variables. The regression analysis revealed an intercept of -1.01 and a slope of 0.0034. These coefficients offer quantitative insights into the linear association between the number of days and the corresponding loss-per-bag. Essentially, the intercept represents the estimated loss-per-bag when the logarithm of days is zero, while the slope signifies the rate of change in the loss-per-bag for each unit increase in the logarithm of days. This analytical approach enhances our comprehension of the nuanced quantitative dynamics between the two variables under examination.

**Taking into account different cases:**

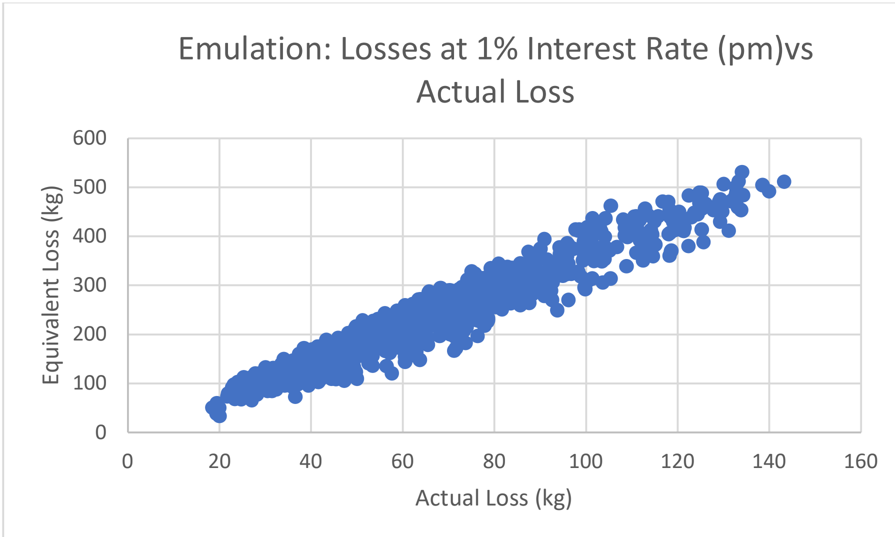



**Fig-03 : Actual weight loss of grains vs loss in loss should have if it should be equivalent to 1 % interest per month**

Examining the provided graph, it becomes evident that with a standard interest rate of 12% per annum, the observed weight loss in grams needs to be approximately five times greater than the actual loss. In practical terms, if there is a total weight loss of, for instance, 40 kg, the equivalent loss considering the initial weight would need to be 200 kg to align with a 12% per annum interest rate. This comparison highlights the substantial impact of the interest rate factor on perceived losses in weight.

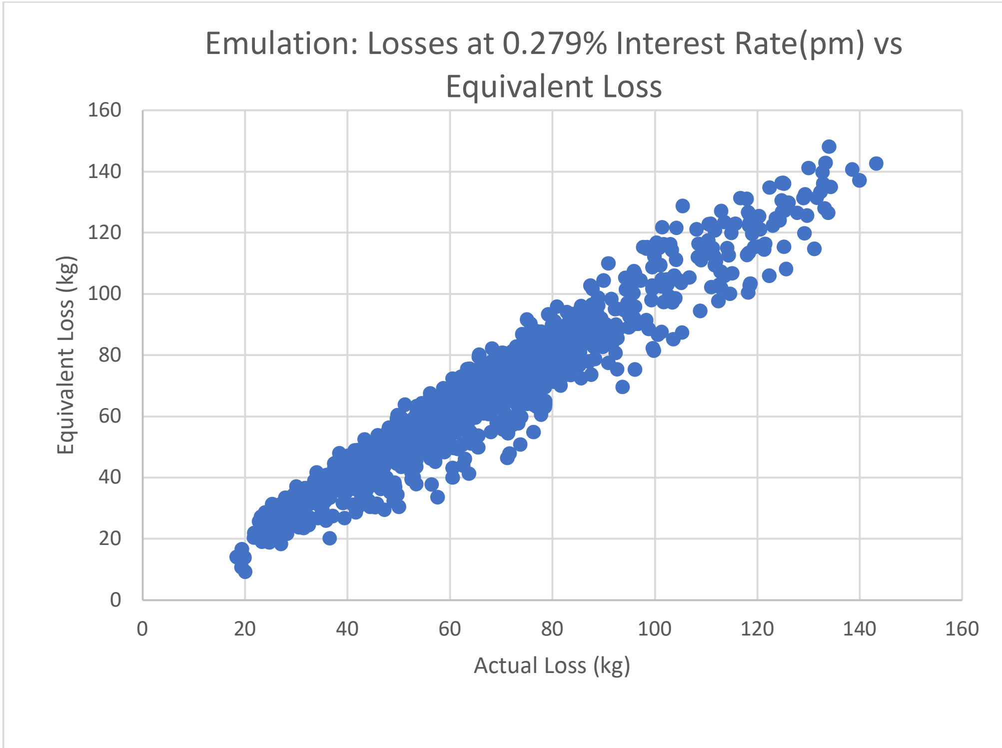



**Fig-04 : Actual weight loss of grains vs loss in loss should have if it should be equivalent to 0.28 % interest per month**

In this illustration, we have taken into account the derived interest rate, denoted as I (0.279% per month). Notably, the observed loss aligns closely with this derived interest rate. Therefore, if we were to consider this loss in conjunction with the initial weight, it would yield an equivalent interest rate of I (0.279% per month). This suggests a strong correlation between the observed loss and the derived interest rate, reinforcing the reliability of the calculated interest rate in reflecting the weight loss scenario.

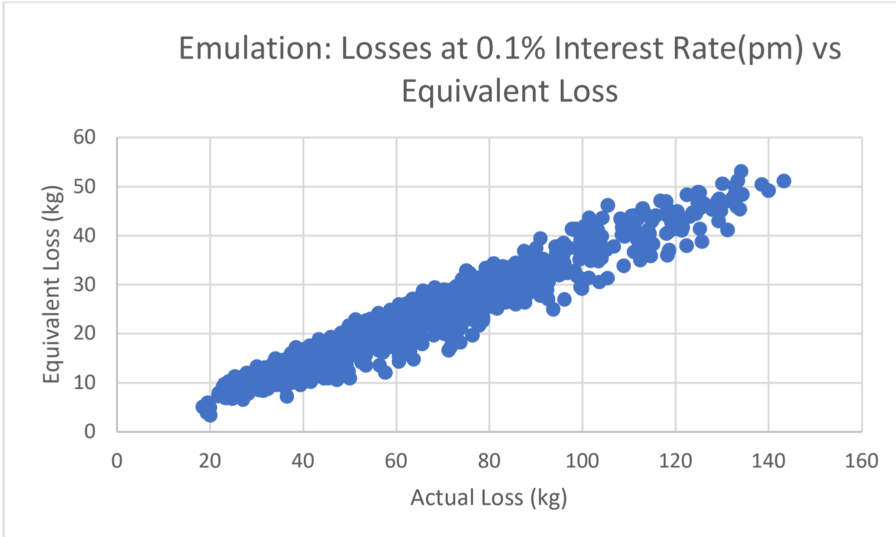



**Fig-05 : Actual weight loss of grains vs loss in loss should have if it should be equivalent to 0.1 % interest per month**

The diagram further reveals that even when considering a very low interest rate, such as 0.1% per month, there is a notable misalignment with the actual loss. This underscores the significance of this decision, as opting for such a low interest rate could result in an overall loss of revenue. The discrepancy between the low interest rate and the actual loss depicted in the diagram emphasizes the potential financial implications associated with choosing a rate that does not accurately reflect the true nature of the loss.

Indeed, based on the analysis of the three figures presented above, it is evident that the most accurate alignment with the actual weight loss occurs when the interest rate is set at I = 0.279% per month. This finding underscores the importance of selecting an appropriate interest rate, as it directly impacts the congruence between the modeled loss and the real-world weight loss scenario. The figure suggests that I = 0.279% per month provides the best fit, reflecting a closer resemblance between the calculated interest rate and the observed weight loss data.

**Comparision between current and proposed methods:**

We have now three prospectives:

a) Current revenue pattern
b) Revenue if Initial weight consider
c) Revenue if interest consider

Our analysis reveals that considering the initial weight results in an equivalent interest rate of (I*12) = 3.35% per annum. Additionally, we have taken into account I = 12% per annum, which represents the standard interest rate.

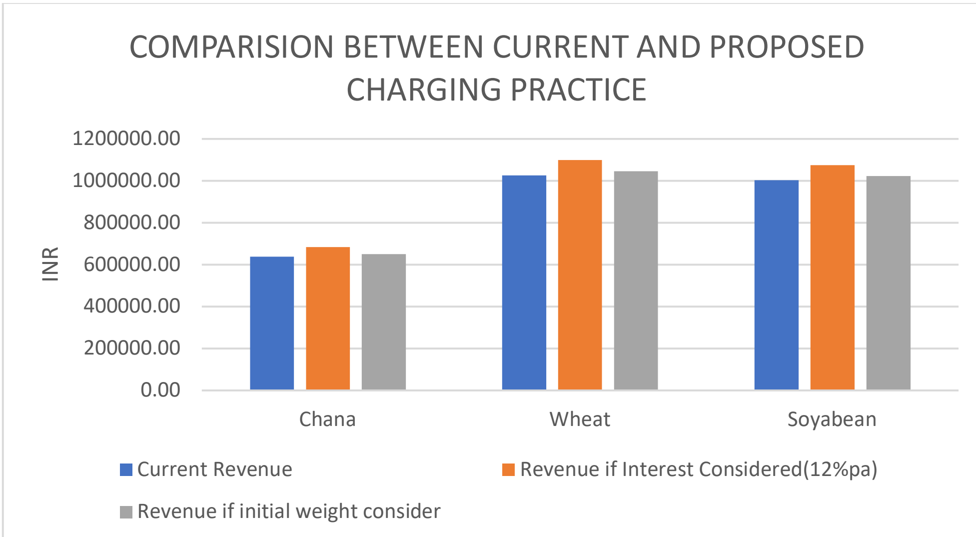



**Fig-06 : Comparision of total revenue(feb20-oct23) at different practices**

On comparing the revenue generated from three different crops—Chana, Wheat, and Soyabean—under three distinct scenarios: the current revenue, revenue with a 12% annual interest rate considered, and revenue with consideration of the initial weight.

1. **Chana:**
   - **Current Revenue:** 638,079.49
   - **Revenue with 12% Interest:** 684,203.56
   - **Revenue with Initial Weight Consideration:** 650,817.00
2. **Wheat:**
   - **Current Revenue:** 1,024,794.02
   - **Revenue with 12% Interest:** 1,099,465.35
   - **Revenue with Initial Weight Consideration:** 1,045,403.40
3. **Soyabean:**
   - **Current Revenue:** 1,001,826.65
   - **Revenue with 12% Interest:** 1,074,051.48

- **Revenue with Initial Weight Consideration:** 1,021,797.60

**Key Observations:**

- Considering a 12% annual interest rate generally leads to higher revenues across all crops compared to the current revenue.
- The revenue with initial weight consideration varies for each crop but tends to fall between the current revenue and the revenue with a 12% interest rate.

This analysis provides valuable insights into the potential financial impact of different considerations on crop revenue, allowing for informed decision-making in agricultural planning and management.

## Results for labours problems

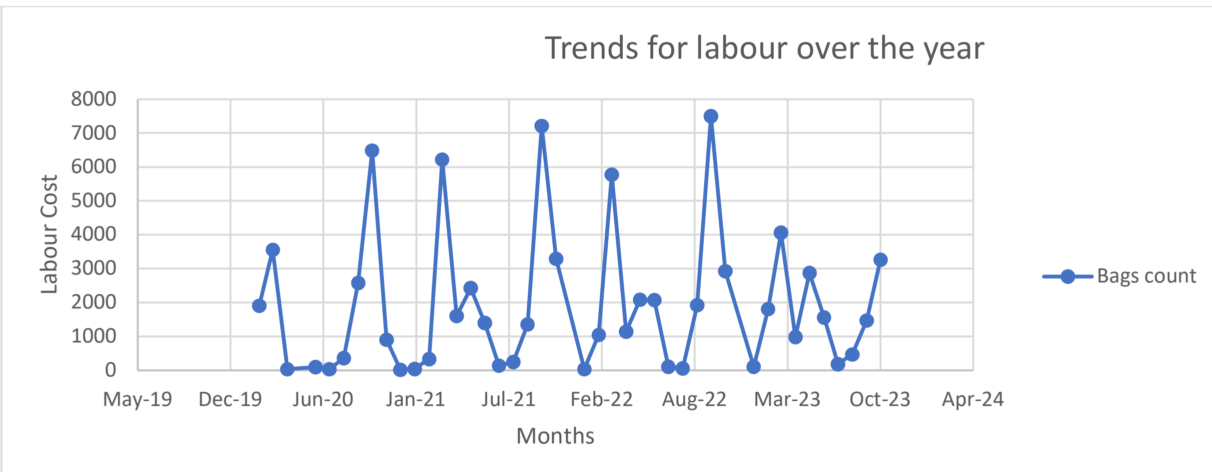



**Fig-07 : Trends of labour requirements at various months**

**Following points can be note:**

1. **Labour Charge:**
    - The "Labour charge" represents the associated labor costs in each month.
    - Like the bags count, the labor charge also varies widely, ranging from a minimum of 60 to a maximum of 22491 (in Oct-22).
2. **Trends:**
    - Looking at the trends, there are notable peaks in both bags count and labor charge, particularly in Oct-20, Mar-21, Oct-21, and Oct-22.
    - Certain months have a relatively stable low count and labor charge, while others show more significant variability.

3. **Seasonal Patterns:**

    - There seems to be some seasonality in the data, as certain months consistently show higher or lower values. For example, Oct-20 and Oct-21 both have high counts and charges.
    - The data also shows a drop in bag count and labor charge during the months of Apr-20, Dec-20, and Jan-22.

Upon closer examination, it's apparent that the months of October and March consistently experience a heightened demand for labor, suggesting potential peak periods or specific activities requiring more workforce involvement. In contrast, months like February, April, May, June, and November show a moderate, yet steady, demand for labor, indicating a sustained but not exceptionally high level of resource utilization.

Furthermore, months such as January, July, and August consistently demonstrate a lower demand for labor. This could be indicative of seasonal trends, holidays, or a generally slower pace of operations during these months.

Adding to the insights, the dataset reveals an average monthly cost of 5835, suggesting a baseline expenditure across the observed period. This average serves as a reference point for evaluating the variations in labor charges across different months, providing a clearer understanding of the overall financial implications associated with the fluctuating demand for labor.

From an interview with the organization's owner, it has been disclosed that the optimal number of required laborers during active work periods ranges between 5 to 6 individuals. However, given this operational scenario, the budgetary allocation of 6000 per month appears insufficient to accommodate permanent employment for this workforce.

This information sheds light on the financial challenges the organization might face in maintaining a consistent and adequately staffed workforce throughout the year. The misalignment between the expected budget and the actual labor demand highlights a potential need for a more flexible employment arrangement, such as hiring seasonal or part-time workers to align with fluctuations in workload.

# Recommandations

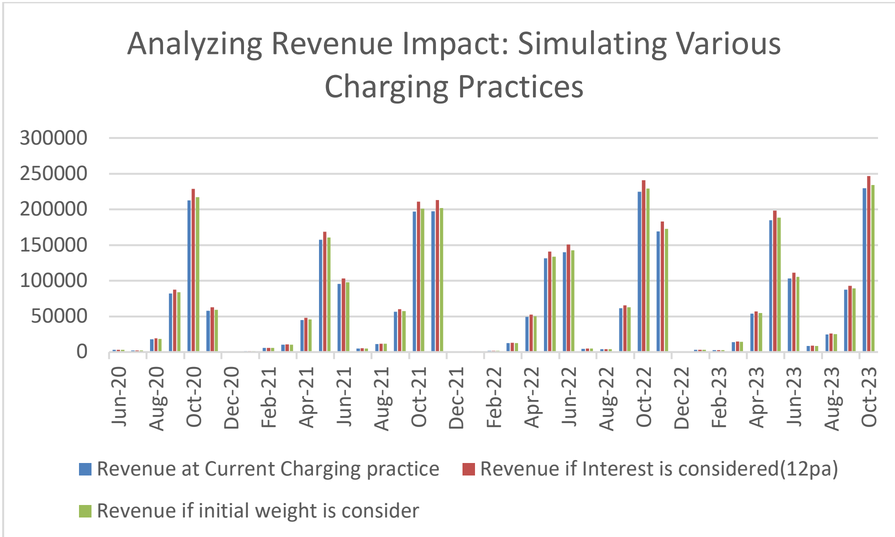



**Fig-08 : Analyzing Revenue Impact**

**Key observations:**

**Total Revenue at Current Practice(base line ):** 2,664,700.159

**Total Revenue at Interest Consideration:** 2,857,720.391 (+7.25% compared to current practice)

**Total Revenue at Initial Weight Consideration:** 2,718,018 (+2.00% compared to current practice)

These percentages represent the relative revenue increase or decrease compared to the revenue generated under the current charging practice.

**Suggestions for charging practice:**

Based on our findings and analysis, we recommend revising the current charging practice. Currently, the system does not account for the initial weight or interest. When incorporating the initial weight, the return is calculated to be (3.35%pa). However, if interest is taken into consideration, the return is expected to be even higher(12%pa).

To optimize the charging practice, we propose incorporating interest on the rent amount if paid at the end date. This adjustment would not only better reflect the financial dynamics but also ensure a more accurate and potentially increased return on investment. Implementing this change aligns the charging practice with a more comprehensive understanding of the financial aspects involved, contributing to a fairer and more profitable system for warehouse.

**Suggestions for labour:**

Considering the uneven demand for labor and a net lesser requirement for a consistent workforce, it is advisable to adopt a strategy of managing irregular or temporary labor. This approach is recommended due to the following reasons:

1. **Fluctuating Demand:** The inconsistent demand for labor suggests that maintaining a full-time, regular workforce may result in periods of underutilization, leading to inefficiencies and potentially poor workforce productivity.
2. **Cost Efficiency:** Employing regular staff during periods of low demand could lead to higher fixed labor costs and, subsequently, less competitive salaries for the regular employees. Managing irregular labor allows for better cost control during periods of reduced workload.
3. **Flexibility:** Utilizing irregular or temporary labor provides the flexibility needed to scale the workforce up or down according to the specific demands of the business. This adaptability is crucial in optimizing resource utilization and maintaining operational efficiency.
4. **Salary Considerations:** By managing irregular labor, the organization can tailor salary structures to align with actual work hours or project-based requirements. This can result in more cost-effective compensation practices.

In conclusion, adjusting the labor management strategy to incorporate irregular or temporary labor is recommended to better align with the dynamic nature of demand. This approach ensures greater flexibility, cost efficiency, and the ability to adapt to changing operational needs, ultimately contributing to a more resilient and responsive workforce management model.