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

A Mid-term 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:

EXECUTIVE SUMMARY AND TITLE

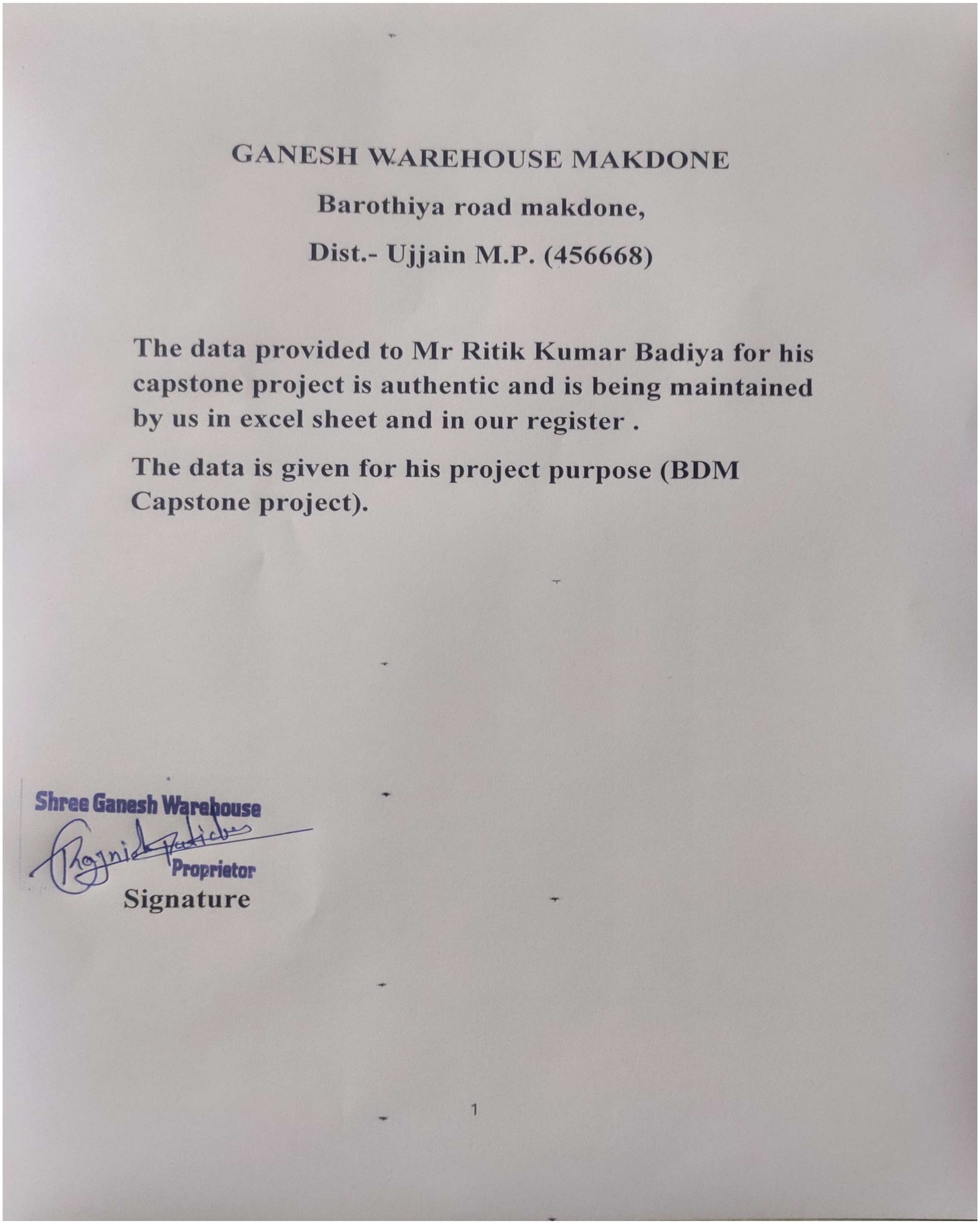
This project centers on a warehouse situated in the village of Makdone, specializing in the storage and distribution of Chana, Soybeans, and Dhana. Operating as a hybrid B2C and B2B enterprise, it caters directly to the agricultural requirements of the local community and also manages certain government contracts. Our primary emphasis within this project will be on the B2C aspect of the business.

For over 25 years, the warehouse has been a steadfast presence in the village of Makdone, under the ownership of Mansingh Patidar. With an annual turnover of approximately 12 lakhs, the warehouse boasts a storage capacity of 10,000 quintals of grains, housed within a spacious area spanning 140*40 sqft. Presently, the warehouse is efficiently managed by a single manager, with no permanent labor force.

The primary challenge confronting the organization revolves around the charging practices, specifically the timing of charging customers. The observed fluctuations in grain weight pose a potential impact on the final rental fee. Currently, charges are applied when customers remove their grains, typically when the weight is lower. The organization also grapples with the decision of whether to maintain a permanent labor force. Additionally, a dilemma exists in choosing between chana and wheat, as chana demands greater care, yet offers a higher charge rate.

To address the central issue, we are conducting various statistical analyses correlating the number of days with the loss in weight of grains. This approach aims to determine whether we should base our calculations on the initial weight (90kg) or factor in the reduced weight, while also considering the accrued interest from the day the grains were stored. Through these analyses, we seek to gain insights that will guide our decision-making process and enhance the accuracy of our charging practices.In order to tackle the second challenge, we are exploring the possibility of securing a labor force for permanent employment. We will seek competitive offers, considering the potential to offer higher compensation in exchange for consistently excellent work, as compared to non-permanent labor. Our analysis will focus on evaluating the viability and adequacy of the received offers, ensuring that they align with our expectations for efficient and flawless work. In addressing the third challenge, our approach involves quantifying the potential annual additional income from opting for chana instead of wheat. Simultaneously, we will assess the extra overhead costs associated with managing chana. By analyzing the financial implications of this choice, we aim to make an informed decision that balances the increased revenue from chana with the corresponding overhead expenditures, ensuring a comprehensive understanding of the trade-off involved.

PROOF OF ORIGINALITY OF DATA



GANESH WAREHOUSE MAKDONE

Barothiya road makdone,

Dist.- Ujjain M.P. (456668)

The data provided to Mr Ritik Kumar Badiya for his capstone project is authentic and is being maintained by us in excel sheet and in our register .

The data is given for his project purpose (BDM Capstone project).

Shree Ganesh Warehouse

Proprietor

Signature

1

Fig-01 : Letter From Organisation

Fig-02 : Field Visit (B2B building)

Fig-03 : Meeting With Manager

Fig-04 : Out Side of Warehouse (B2B building)

Fig-05 : B2C warehouse

Fig-06 : Meeting Inside Office

Fig-07 : Office

META DATA

Meta Data about warehouse	
Location	Makdone Village, Barothiya Road, Ganesh Warehouse
Owner	Mansingh Patidar
Business Type	Both Business-to-Business (B2B) and Business-to-Consumer (B2C)
Annual Turnover	₹12 Lakhs
Storage Capacity	10,000 Quintals
Warehouse Floor Area	140 feet by 40 feet
Manager	Rajnish Patidar

Meta Data about sheet.xlsx	
File Name	Sheet1.xlsx
Author	Rajne
Duration	21/02/2020 to 31/10/2023
Number of Attributes	14

Description of Attributes	
Name	Name of Customer
Item	Item to be stored
Sdate	Starting Date of Storage
Edate	Last Date of Storage
Days	Total Days
Bags Count	Count of Bags of Grains
Total Weight In	Total weight at the starting date
Total Weight Out	Total weight at the last date
Labour Cost	Total labor cost
Charge	Total rental charge
Rate	Market Rate of Grains
Rent per Bag	Rental cost per bag
Stack	Location of bags in the warehouse

Brainstorming questions	
Parameter	**Value**
Number of Labours Needed	5 persons
Minimum Pay	for 105 days
Labour Cost	₹3 per kg
Initial Packing of Bags	90 kg

Descriptive Statistics

Chana:

- **Mean (Average):** The average price of Chana is approximately 1.83.
- **Median (50th Percentile):** The middle value of the dataset is around 1.85.
- **Standard Deviation:** The prices exhibit a relatively low amount of variability, with a standard deviation of approximately 0.11.
- **Minimum:** The lowest recorded price for Chana is 1.21.
- **Maximum:** The highest recorded price is 2.01.
- **Range:** The price range, from the minimum to the maximum, is about 0.80.
- **25th Percentile (Q1):** 25% of the prices fall below 1.77.
- **50th Percentile (Q2/Median):** 50% of the prices fall below 1.85.
- **75th Percentile (Q3):** 75% of the prices fall below 1.91.

Wheat:

- **Mean (Average):** The average price of Wheat is around 1.65.
- **Median (50th Percentile):** The middle value of the dataset is approximately 1.66.
- **Standard Deviation:** Wheat prices show a higher degree of variability compared to Chana, with a standard deviation of approximately 0.19.
- **Minimum:** The lowest recorded price for Wheat is 0.90.
- **Maximum:** The highest recorded price is 2.03.
- **Range:** The price range is more extensive, with a value of about 1.13.
- **25th Percentile (Q1):** 25% of the prices fall below 1.51.
- **50th Percentile (Q2/Median):** 50% of the prices fall below 1.66.
- **75th Percentile (Q3):** 75% of the prices fall below 1.79.

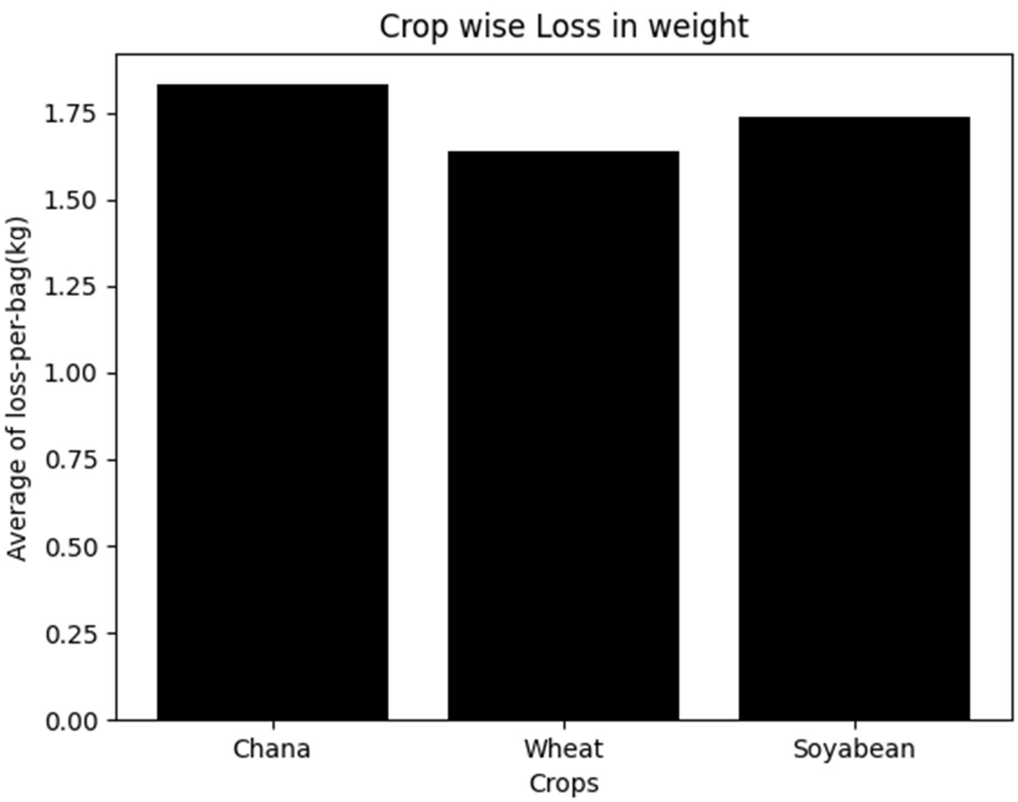
Soybean:

- **Mean (Average):** The average price of Soybean is around 1.74.
- **Median (50th Percentile):** The middle value of the dataset is approximately 1.75.
- **Standard Deviation:** Soybean prices show a moderate level of variability, with a standard deviation of approximately 0.18.
- **Minimum:** The lowest recorded price for Soybean is 0.0.
- **Maximum:** The highest recorded price is 2.02.
- **Range:** The price range is quite broad, with a value of about 2.02.

- **25th Percentile (Q1):** 25% of the prices fall below 1.66.
- **50th Percentile (Q2/Median):** 50% of the prices fall below 1.75.
- **75th Percentile (Q3):** 75% of the prices fall below 1.86.

Comparison:

- **Mean:** Chana has the highest average price, followed by Soybean and Wheat.
- **Variability:** Wheat prices exhibit the highest variability (standard deviation), while Chana has the least.
- **Range:** Soybean prices have the widest range, indicating a larger spread of prices.
- **Median:** Chana and Soybean have similar median values, while Wheat is slightly lower.
- **Quartiles:** Chana has the highest 75th percentile (Q3), indicating that a significant portion of its prices are relatively high.



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 Test 01, the correlation coefficient between 'Log Days' and 'Loss' is calculated. This test assesses whether there is a linear relationship between the logarithm of the days and the loss in weight.

In Test 02, the correlation coefficient between 'Days' and 'Loss' is calculated. This test investigates the linear relationship between the actual number of days and the loss in weight.

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: Linear Relation Consideration

After establishing a linear relationship between 'Days' and 'Loss,' the analysis extends to investigating linear relations in both scenarios (Weighted Charging and Interest Charging). This involves exploring whether a linear model adequately captures the relationships in these scenarios.

Step 06: Finding the Best Interest Rate (I)

weighted_charging = current_charging + days * loss * charge_per_kg_per_bag

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

we will now fit a linear curve from which we will try to find out I so that intrest_charging can define weighted_charging. on equating both sides:

loss = days * weight_out * I/3000

comparing to: y=mx,

form linear regressing we know: $\Sigma y = m\ \Sigma x$

we can say that: $\Sigma loss = ((\Sigma(days * weight\_out))/3000) * I$

which implies : $I = \Sigma loss /((\Sigma(days * weight\_out))/3000)$

Step 07: Decision-Making

Based on the analysis, make a decision on which charging practice to consider, taking into account the expected interest rate I.

Justification for Linear Regression:

Linear regression is employed due to the observed linear relationship between loss in weight and days, providing a means to compare the two charging practices. The analysis leverages the linearity assumption to derive insights into the relationships between variables and make informed decisions.

Results and Finding

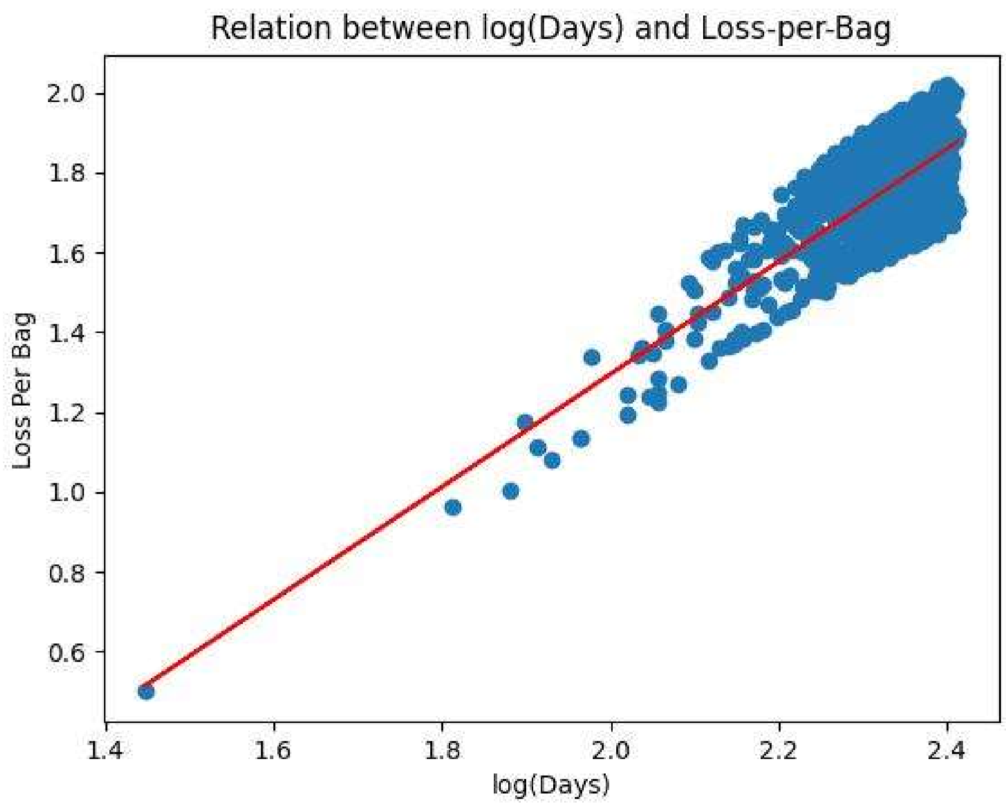


Fig-08 : Relation between log(Days) and Loss-per-Bag

We conducted a thorough analysis to determine the correlation between the logarithm of the number of days and the loss-per-bag metric. The calculated correlation coefficient was found to be 0.73, providing strong support for our hypothesis that there exists a significant correlation between these two variables. Following the correlation analysis, we proceeded to fit a linear curve to the data. The regression analysis revealed an intercept of -1.53 and a slope of 1.41, shedding light on the quantitative relationship between the logarithm of days and the associated loss-per-bag.

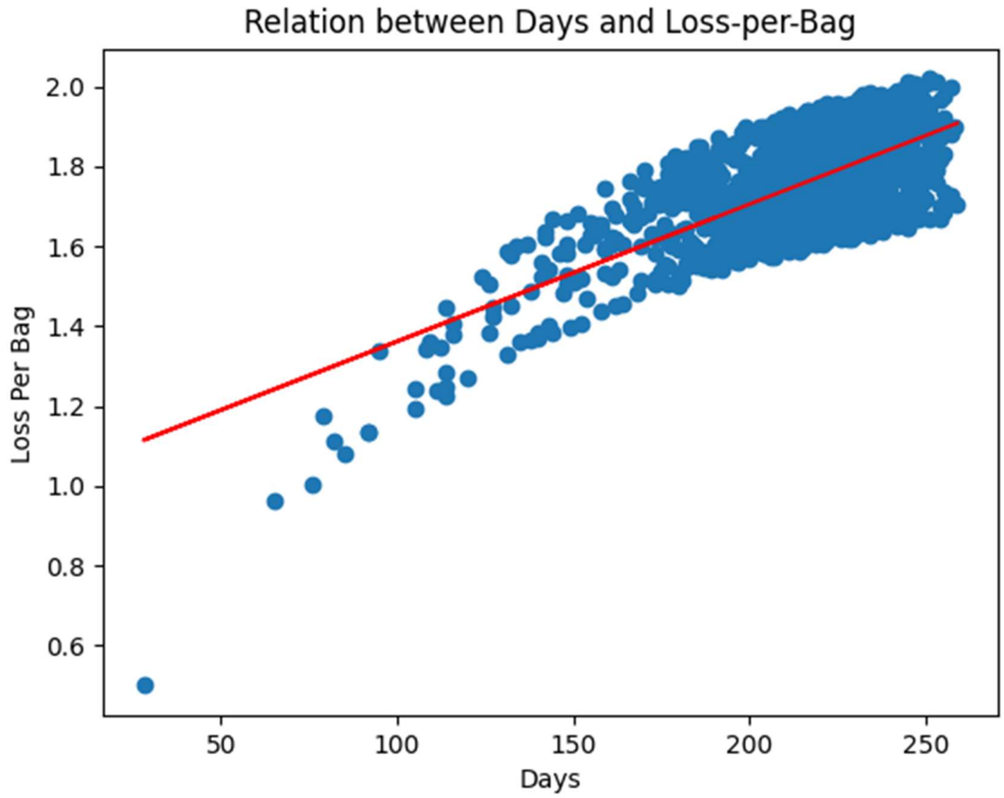
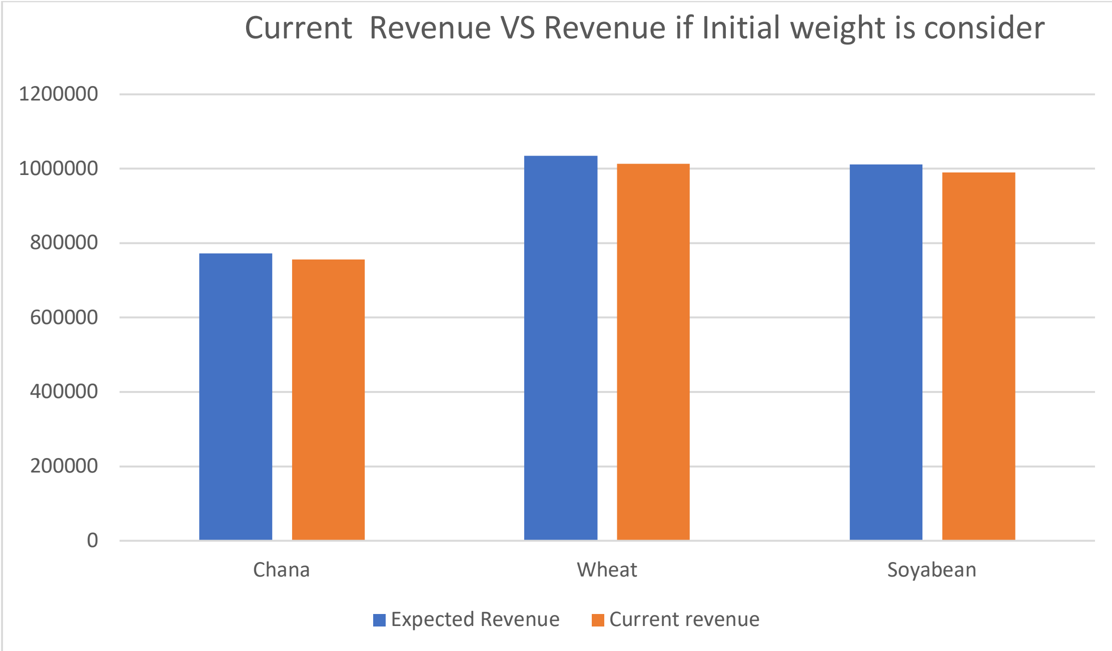


Fig-09 : Relation between Days and Loss-per-Bag

We conducted a thorough analysis to determine the correlation between the variable "Days" and the corresponding "loss-per-bag," revealing a substantial correlation coefficient of 0.70. This finding strongly supports our initial hypothesis that there is a significant correlation between these two factors. It's noteworthy that this correlation slightly strengthens when we substitute the variable "Days" with its logarithmic transformation, although for the purposes of our initial assumption, we will adhere to the original representation.

Subsequent to establishing the correlation coefficient, we proceeded to fit a linear curve to the data. The resulting linear model is characterized by an intercept of 1.01 and a slope of 0.0034. These parameters provide valuable insights into the relationship between "Days" and "loss-per-bag," shedding light on the underlying trends in the dataset.



Based on the provided illustration, it is evident that considering the initial weight instead of the final weight leads to a substantial increase in revenue. The increments for each crop are as follows:

Chana: 15,838

Wheat: 21,191

Soybean: 22,130

The cumulative increase over the two-year period amounts to 59,159, signifying a significant total revenue gain. We will also consider other charging practice and compare them with these two.