Optimization Analysis for Vendor Delivery Services

Final-Term report - BDM capstone Project

Submitted by

Bhavya Saxena

23DS2000108

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

Declaration Statement

I am working on a Project titled "Optimization analysis for vendor delivery services". I extend my appreciation to Shree Sai Facilities, 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 to the utmost extent of my knowledge and capabilities. The data has been gathered from primary sources 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 principles 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.

Signature of Candidate: (**Digital Signature**)

Name: Bhavya Saxena

Date: 4th October, 2023

1 Executive Summary and Title

This project centres around optimizing the operations of a vendor company associated with Blue Dart Couriers Ltd., situated in New Delhi, operating in the B2B segment of courier delivery services across India. The organization has their employees' delivering packets for Blue Dart and get paid on a per packet basis by them.

Organization Background: Shree Sai Facilities, A service provider company whose major client is Blue Dart Express limited. It is an Indian based organization started in 2012 and head quartered in New Delhi which has branches all over India such as NCR region, Jaipur, Ajmer. The organization has a turnover of nearly 4 crores with a man power of 140 employees. The organization sources/recruits man power and manages the payrolls, legal compliances of the staff for Blue Dart in of its various branches for delivery services.

Current Challenges: The organization is currently grappling with nominal profits in some of their branches, attributed to inefficient analysis of the correlation between the number of employees hired and the volume of packets received per branch. This obsolete analysis has resulted in suboptimal profit margins till now.

Analytical Approach: To address this issue, we have employed various analytical approaches, including graphical representations such as pie charts, bar graphs, and scatter plots. These methods have provided us with many valuable insights into the existing operational dynamics and the interdependence of different data variables.

Machine Learning Integration: Following a comprehensive analysis, we implemented a supervised learning model, Linear Regression model was chosen to be 'leveraged on the basis of the data analysis results. This model aims to determine the optimal number of employees per branch and the corresponding volume of packets to maximize branch profitability. There were 3 different regression models which were made to be leveraged for our analysis out of which one was selected as per the needs of the organization.

Website Integration: In the pursuit of seamless implementation and accessibility, our analytical insights and machine learning models have been integrated into a dedicated web platform. This website has been given to the vendor for a seamless usage of model whenever required, the website mentioned asks for the estimated number of packets from the user and using that predicts the number of employees they should keep in their branch to maximize profit.

Result/Outcomes: The anticipated results of this project include a refined understanding of operational dependencies, paving the way for the organization to enhance efficiency and subsequently increase overall profitability. By predicting the optimal employee-to-packet ratio, the organization can now streamline its workforce, leading to improved operational effectiveness.

In conclusion, this project offers a strategic roadmap to address the current challenges faced by the organization, providing data-driven solutions to optimize operations and drive long-term profitability.

2 Detailed Explanation of Analysis process

We can see the process of Analysis step-wise:

1) **First Look/Understanding Data**: To start with the process of Analysis, we first decided to Analyze the given data manually, to understand the variables and their respective value for the organization.
   While looking at the data manually, we could see the data revolved around the working and statistics of 4 branches – FAR, RMR, Okhla and JIA, which were the branches the organization needed an analysis on, as the owner thought that these branches had more potential than they are delivering on an average. To be sure that the data only contained these four branches as discussed with the organization, we uploaded the data in panda's library and printed the unique values.

```
In [19]: print(data2['BRANCH'].unique().tolist())

         ['FAR', 'OKHLA', 'JIA', 'RMR']
```

Image: Printing unique contents in BRANCH column

   After confirming the four branches, we started looking for key variables which directly or indirectly affected the profit of the organization. For this analysis we could see there were certain formulas on the "GPROFIT" (Profit) variable, by backtracking through these formulas we could see there were multiple variables which affected the profit.
   The variables which directly affected: -
   Expenses, Billing
   Variables indirectly affecting: -
   Packets, Employees, Months
   This analysis helped us in identifying our key variables, whose analysis further could be beneficial for our research.

2) **Data Cleaning:** Once the manual Analysis was done, the data was put in pandas to be analyzed further. While analyzing a bit more we found out some discrepancy in data which is discussed as follow.
   i) Serial Number column: The serial number column was an extra variable (Not useful in terms of analysis) which had blank values, because the organization had merged the rows which contained same branch and was not much useful for the analysis. So, when the data was uploaded to pandas its contents were showing as mostly Null values, therefore it was better for the analysis to remove this column.

   ii) Incentive column: This column had all values as 0, when asked why this was the case, the organization said they had halted the process of incentives for more than a year but hadn't removed the column from the format for future use. Thus, making this a column which was unnecessary for the analysis, so it was better to remove this and the above column using the drop method.

```
data = df.drop('SNO',axis=1)
data2 = data.drop('INCENTIVE',axis=1)
```

Image: code to drop SNO and INCENTIVE columns

iii) The third part of the data cleaning process was to unmerge the branch column as it was merged similar to the serial number column, but unlike "SNO" column this variable was important. The data was already uploaded to pandas at this point, so the unmerging was done using fillna function with method as "ffill". This basically fills all the null values using the above first not null entry.

3) Analysis: For the analysis part the key variables which were found in the first step were utilized to understand their impact on the profit variable.

1. To start with the analysis, we uploaded the data in pandas to get a general view and understanding of the variables, then we tried to analyze the averages and min/max value for all the variables, for which we utilized describe () function. Using this, we were able to get a general idea and range of each variable.

```
print(data2.describe(include='all'))

       BRANCH   MONTH  Employees     MANDAYS       WAGES         SALARY  \
count      80      80  80.000000   80.000000   80.000000      80.000000
unique      4      20        NaN         NaN         NaN            NaN
top       FAR  Apr-22        NaN         NaN         NaN            NaN
freq       20       4        NaN         NaN         NaN            NaN
mean      NaN     NaN  17.198750  446.637500  619.005125  270856.205250
std       NaN     NaN   5.140047  133.723808   46.458224   58147.580675
min       NaN     NaN  10.000000  269.000000  538.460000  172832.500000
25%       NaN     NaN  14.000000  360.250000  627.707500  234351.875000
50%       NaN     NaN  15.000000  383.500000  642.500000  247362.500000
75%       NaN     NaN  19.500000  520.000000  642.500000  314056.795000
max       NaN     NaN  31.000000  816.000000  696.040000  439383.360000
```

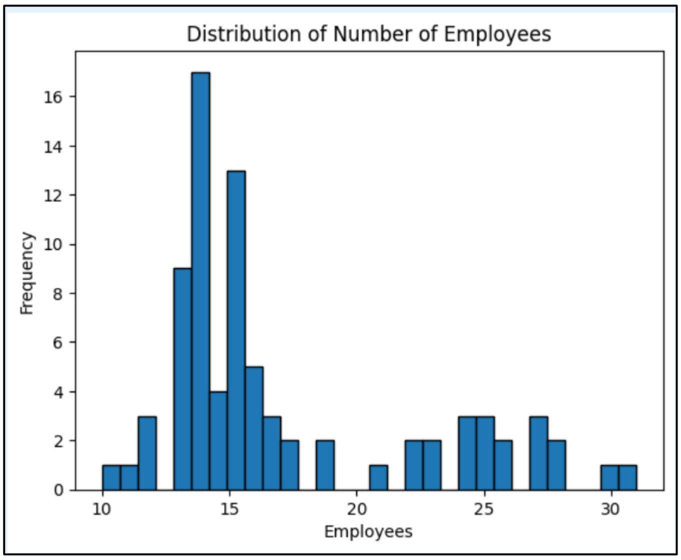
Image: coded snippet of describe function

2. The above analysis helped us get some general idea on the variables, but this analysis was not able to tell us about the working of the different branches. To understand the current working of branches, we were able to utilize the concept of pivot tables. Using pivot tables, we were able to see the average values of our key variables identified in the first step (Understanding Data) for each of our branches. By analyzing the pivot tables (table 1 and 2 in descriptive statistics section), we were able to identify that the branch – FAR (Faridabad) was the most profitable branch for the organization till now, it could be seen that due to its higher number of employees and packets (Nearly twice as other branches) it had a higher amount of profit.

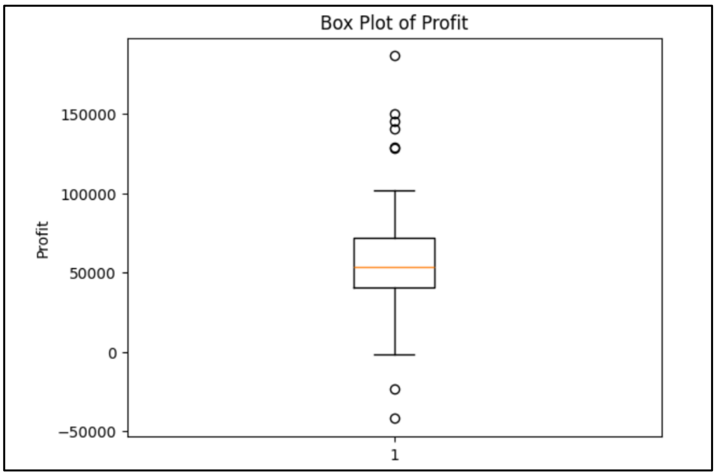
3. After the Analysis using pivot tables, there was a need of descriptive analysis of our key variables to further understand them. For this, Jupyter Notebook was used in which the data was loaded through pandas and visual plots were created using matplotlib (Pandas and matplotlib are libraries of python).

i) The graph below (Graph1) shows the frequency trend of the number of employees in a branch in any given month by the organization, this trend tells us that the organization tries to keep as much employees as they can hire but currently the number averages around 16-17.

ii) The Box-plot of the "GPROFIT" column (Graph2) in our dataset tells us that the organization has in some months reached to a good amount of profit of even above ₹1,50,000 but it averages around ₹50,000, in the boxplot we can even see that the organization sometime face losses and that too has gone up to the amount of 50,000 which the organization should plan on avoiding in the future.



Graph1: Frequency Distribution of Employees



Graph2: Box-Plot of "GPROFIT" (Profit)

4. Correlation matrix – Identifying Independent Variables

The major part of our analysis was to figure out which variables to utilize for our model building, we figured out the key variables which directly or indirectly affect our profit in the first part of our analysis section, but to verify which variables to utilize as independent variables among these we utilized the concept of correlation matrix. First, we created a general correlation of all the numerical variables in the dataset, but to focus on the key variables identified we created a separate correlation of only the key variables identified to be able to choose the independent variables amongst them.

<u>**Choosing Independent Variables**</u> - By looking at the below correlation matrix we can see our four key variables which affect our profit either directly or indirectly, among these variables we can see that we can categorize these variables into two sets as mentioned below -

i) First set as (Employees and EXPENSES) – as Expenses is directly affected by Employees as the number of employees being higher or lower affect the cost affect variables such as Salary, PF etc. indirectly affecting the EXPENSES.

ii) The second set as (Packets and Billings) – as Billings is directly calculated from the number of packets multiplied by per packet rate.

So, it would be wise to choose a single independent variable from each of these sets.

```
In [13]:  num_data = data2.select_dtypes(include='number')
          num_data[['Employees','PACKETS','EXPENSES','BILLING','GPROFIT']].corr()
```

Out[13]:

	Employees	PACKETS	EXPENSES	BILLING	GPROFIT
Employees	1.000000	0.856638	0.995597	0.960413	0.392401
PACKETS	0.856638	1.000000	0.843270	0.910701	0.652146
EXPENSES	0.995597	0.843270	1.000000	0.959536	0.377283
BILLING	0.960413	0.910701	0.959536	1.000000	0.622792
GPROFIT	0.392401	0.652146	0.377283	0.622792	1.000000

Image: correlation matrix of key variables identified

By looking at the correlation matrix we can see that from set 1 Employees variable has a higher chance for affecting the profit than "EXPENSES", similarly in second set PACKETS column seems to be the better candidate to be an independent variable. Hence, these two variables seemed to be the optimal feature variables at this point, with 'GPROFIT' being our target variable.

<u>**Why This Method of Analysis is better**</u> – The reason for which our method of analysis using Jupyter Notebook, Pivot tables, Python libraries is considered better as using this analysis we were able to draw many conclusions on our given variables, visually as well as statistically.

4. Model Selection and Creation – Linear Regression:

Following the identification of independent variables in the previous step, the feature and target variables were determined as follows –

- Feature variables: Employees, Packets
- Target Variable: GPROFIT

i) Upon analysing our feature and target variables, we observed that our dataset comprised both target and independent variables. This characteristic made a supervised learning model suitable for our case. After evaluating various supervised learning algorithms, Linear Regression emerged as the most fitting choice.

ii) For implementation and research purposes, the initial model was created using the specified variables to predict profit based on the given number of employees and packets. The first step in model creation involved splitting the data into training and testing sets. To achieve this, the 'train_test_split' library from 'sklearn.model_selection' was employed. However, a challenge arose at this point – determining the optimal random state for our split.

- To address this, a for loop was implemented, iterating over all possible random states from 0 to 10,000. We compared the model scores for each iteration and selected the random state with the highest score for our data split.
- Subsequently, the model was fitted with the training data and evaluated by obtaining its score on the test data.

iii) Despite the successful implementation of the Linear Regression model, discussions with the business revealed its limited usability. The business could easily estimate profit when provided with the values of the two variables. Consequently, a re-evaluation of the existing model or the exploration of alternative models became imperative at this point.
Further consultations with the organization highlighted that they were in a need of predicting the optimal number of employees to enhance profitability. The current approach involved maintaining employee numbers in a branch as per Blue Dart's request, adversely affecting profit. In light of this, the Employees variable emerged as the optimal target variable, with the Packets variable serving as the independent variable.

iv) After the discussions and further research, new set of feature and target variables were selected for the purpose:

- Feature variables: Packets
- Target Variable: Employees

The new model was created looking at the specified needs of the organization, which took number of packets as the input and predicted the number of employees needed to get these packets delivered.

But there was still one aspect which was left, that aspect was "profitability". Our model was able to predict the number of employees needed to get these packets delivered but this was being done without looking at the aspect of profitability.

v) To put in the aspect of profitability the data which was fed in the model was changed, instead of just using all of our data for the training of our model, the data was filtered to take in only those rows where our profit was higher than a certain margin (above 40,000), so that the model is trained on data in which the organization has actually gained some profit.
Using the above fix, we were able to optimize our model to think and predict the number of employees which could better the organizations profit.

```
# Features to use for prediction
feature_to_predict = 'Employees'
feature_for_prediction = 'PACKETS'

filtered_data = num_data[num_data['GPROFIT'] > 40000]
```

Image: filtering data on basis of profit to optimize our model

5. Website Integration:

i) The next challenge was to make the model available for the organization at any given time, for this a website was the best option. To host the model on an open URL on the net, there was a need for two things:

   - UI of the website
   - Model integration in site

ii) To solve the above problems various technologies were utilized such as: HTML, CSS, Flask, sklearn

   - HTML and CSS were used in the creation of website's UI which was kept minimalistic as it was a functional website, which just provided the organization the means to utilize our model.
   - Flask, a web framework for Python, was instrumental in developing the backend of the website. It facilitated managing the interaction between the user interface and the predictive model. Flask's simplicity and flexibility allowed for the seamless incorporation of machine learning models into the web application.
   - The predictive model, developed using the sklearn library, was integrated into the Flask application. sklearn provided the tools for data preprocessing, model training, and predictions, making it a valuable asset in the machine learning pipeline.

iii) After the creation of the website along with the integration of model on a local host, the website was hosted on PythonAnywhere.com.
PythonAnywhere is an online platform that provides an integrated development environment (IDE) and web hosting environment for Python applications.
Using PythonAnywhere, the model was deployed on a website and now can be used on any device.

Website Link - https://bhav.pythonanywhere.com/

6. Discussion with Organization

A formal discussion and showcase of our project were conducted with the organization to present the analysis, model development, and website integration. The presentation aimed to highlight key findings, the utility of the predictive model, and the enhanced features aligned with organizational objectives.

Positive Feedback: The organization provided positive feedback on various aspects of the project:

- **Utility of the Model:** The predictive model's potential in predicting the number of employees needed for optimal profitability was recognized as a valuable tool for decision-making.

- **Website Interface:** The functional and minimalist UI design of the website received positive feedback for its user-friendly nature and ease of navigation.

- **Integration with PythonAnywhere:** The decision to host the website on PythonAnywhere.com was appreciated for ensuring accessibility across devices without the need for local installations.

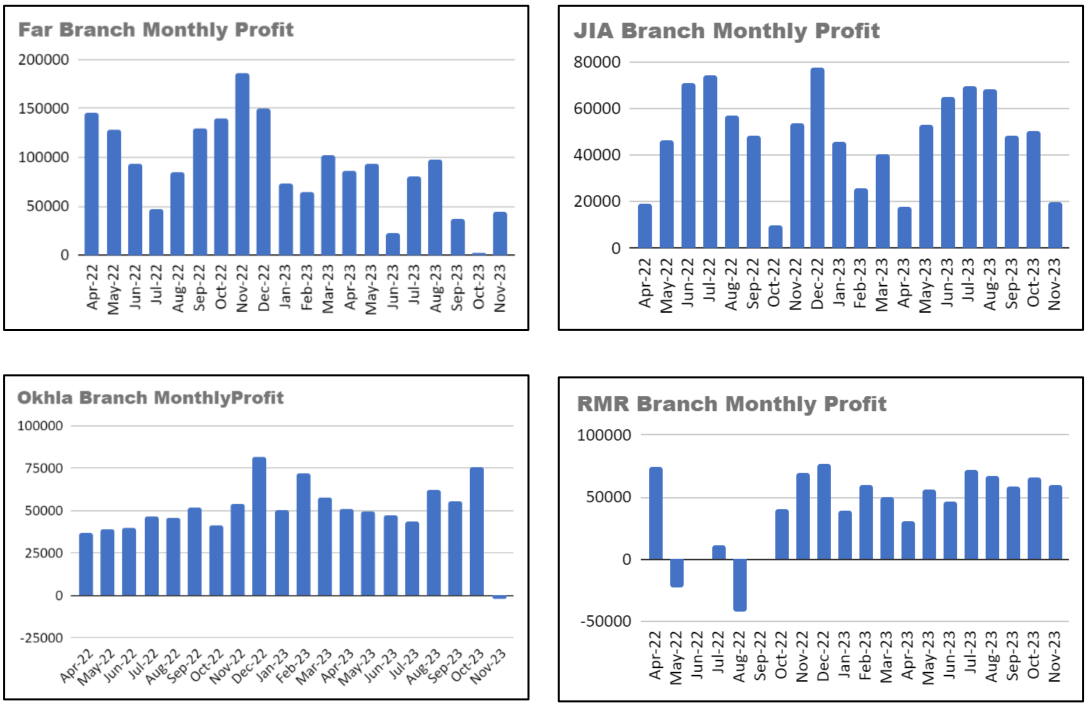
Overall, the organization expressed satisfaction with the collaborative approach, emphasizing the alignment of the project outcomes with their operational needs. The positive feedback serves as a testament to the successful integration of data analysis, machine learning, and web development to address the organization's specific challenges and goals.

3 Result And Findings

During our research of the business's processes and its data we were able to derive many useful insights which has helped them enhance their profitability as well as help them understand the current performance as well as the potential in their business.

The results are as follows:

1. **Branch-Wise Profit Analysis:**



Images: Monthly profit of each given branch

i) **FAR Branch:** Viewing the monthly data of FAR branch, it is easy to see that the organization was gaining good profit in the year 2022 from this branch and also it can be seen that it had achieved the maximum amount of profit from this branch itself in the month of November, however we can see that the profit steadily declined in the upcoming year 2023. The cause of decline of profit seems to be the decrease in number of packets of each consecutive month while the number of employees increased slightly or remained, which increased the expenses but the billing remained the same.

ii) **JIA Branch:** In Jia Branch we see a profit capped at "80,000", significantly lower than what we see in FAR branch. This is due to the employee difference in both the branches as FAR branch has an average of – 25 employees in a year, whereas Jia branch only has 16. Apart from that, we can see that the profit dipped in some of the months such as 'Oct-22' and 'Apr-23', it can be seen that the dip in these two months attributed to low number of packets in those months, while the number of employees remained same as in other months. This led to an increase in expenses yet the billing (which attributes to profit) was low.

iii) **Okhla Branch:** Viewing the monthly profit, it can be seen that Okhla branch is the steadiest branch amongst all of the branches, as the monthly profit of this branch has been constantly above 30,000 and also has peaked up to 80,000 in the month of December 2022. The reason for the constant performance of this branch has been its number of employees, in the dataset it can be seen that the number of employees in the branch was kept at a low number of an average of '13' and not hastily increasing this number even when the number of packets increased as this trend is seen in other branches.

iv) **RMR Branch:** 'RMR' branch showed the most variance in its profit over the months as compared to the other branches during that same timeframe. It can be seen that this branch had a huge dip in the end of the year 2022 and is the only branch giving a loss to the organization in the months of May and August, while giving near zero profit in the months of June and September 2022. It can be seen that it regained its ability to give out profit in the month of October in 2022.
However, the dip in profit and losses of this branch in the year 2022 was attributed to a very poor maintenance of the employee to packet ratio, which can be seen in its lower performing months with the number of employees being on a higher side and the number of packets being on a lower side. For example: in the month of August,2022 the organization faced a loss of "42,000" the number of packets in that month was nearly "24,000" whereas the employees at that time was 17 which is considerably higher than the number of packets the branch had received, resulting in the organizations worst performing month and branch.

v) **General Branch Analysis:** By viewing the data of all the branches along with the above visualizations, a trend can be seen in the organization's profit depending upon the month of the year. It can be seen that all of the branches peaked in their profits at the end of the year 2022 (November, December), this could be attributed to the months being the most holiday/occasion populated months of the year. As people send their friends/relatives gifts and letters it could be seen that the number of packets and hence the profit peaked around that time. When discussed this with the organization they approved of this hypothesis and said that in this year they were hoping to see the same spike in the month of December 2023.

2. **Identification of Most Profitable Branch (FAR):**

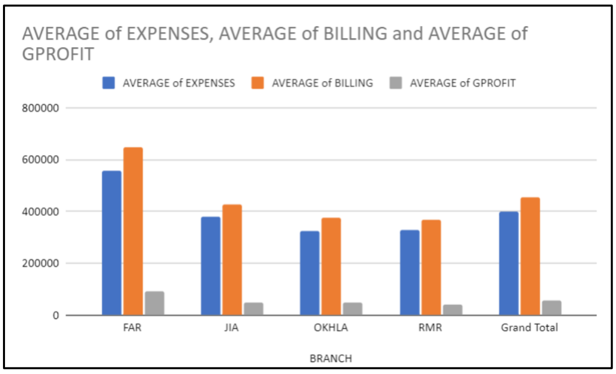


Image: Bar graph depicting Expenses, Billing and GPROFIT of each branch

- Through the use of a range of analytical tools like Seaborn, Matplotlib, Pandas, and Pivot Tables, the analysis revealed that the Faridabad (FAR) branch consistently outperformed others in terms of profitability for Shree Sai Facilities.
- **Result**: Visual representations including graphs and pivot tables effectively showcased the performance patterns across various branches, unequivocally highlighting FAR as the top performer with superior profit margins, which later was found out to be attributed to a large number of employees along with a higher and steadier number of packets compared to the other branches.

3. Identification of Independent Variables:

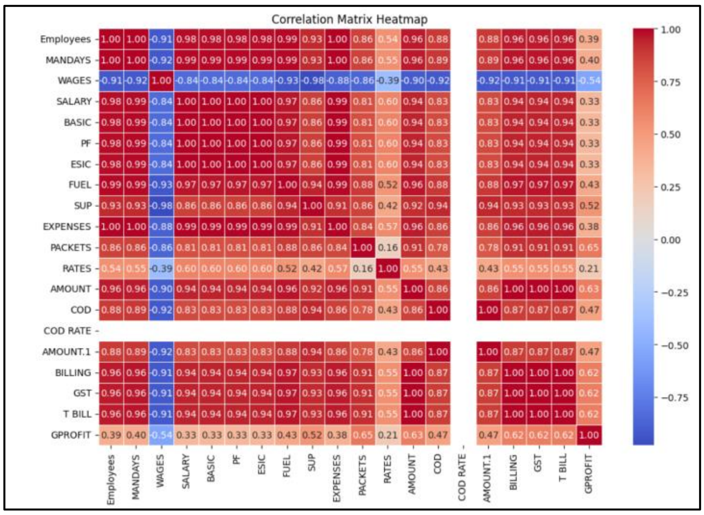


Image: Correlation matrix key variables

- Utilizing advanced correlation matrices and the visual aid of the heatmap showcased above, our analysis was successful in identifying key independent variables pivotal in shaping profitability.
- Upon conducting correlation analysis, our findings delineated two distinct sets of variables: Employees and Expenses constituted one set, while Packets and Billings comprised the other.

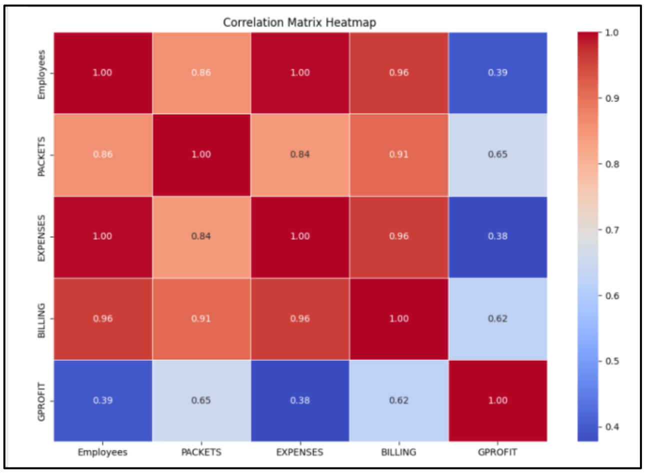


Image: Heatmap of identified key variables

- **Result**: Delving deeper into the correlation strengths, our selection process zeroed in on two paramount independent variables for subsequent modeling: **Employees** and **Packets**. These variables emerged as the most influential factors warranting further investigation and integration into our predictive models.

4. Model Performance and Optimization:

i) Feature Selection:

```
# Features to use for prediction
feature_to_predict = 'Employees'
feature_for_prediction = 'PACKETS'
```

- The feature to be predicted (**Employees**) and the feature for prediction (**PACKETS**) were identified from the filtered dataset by various analytical techniques mentioned in the above sections. Identification of these variables was a crucial step to identify and build a working Linear Regression model.

ii) Data Preprocessing:

```
# Data preprocessing - standardize numerical features
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
```

- Numerical features were standardized using the **StandardScaler** to ensure uniformity in feature scales, a crucial step for many machine learning algorithms

iii) Data Filtering - Refinement for Profitability:

```
filtered_data = num_data[num_data['GPROFIT'] > 40000]
```

- A critical enhancement involved refining the model to focus specifically on profitability. By filtering the data based on a profit threshold (above ₹40,000), the model was optimized to predict the number of employees needed for scenarios where the organization has achieved significant profits.
- To ensure the robustness of our model, we filtered the dataset to include only data points where the gross profit (**GPROFIT**) exceeded ₹40,000. This step aimed to focus on instances with significant profit margins.

.

iv) Train-Test Split:

```
# Train-test split
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.2, random_state=5257)
```

- The dataset was split into training and testing sets using an 80-20 ratio. This partitioning ensured that the model's performance could be evaluated on unseen data.
- Random State Optimization **-** The model underwent an optimization process to determine the optimal random state for data split. A for loop iterated over potential

random states, and the highest scoring state was chosen to enhance the model's accuracy.

v) **Model Training:**

```
# Creating and training the model
model3 = LinearRegression()
model3.fit(X_train, y_train)
```

- Linear Regression model (**model3**) was instantiated and trained using the training data. Linear Regression was chosen for its simplicity and interpretability, which aligns with our objective of maximizing profit.

vi) **Model Evaluation:**

```
model_score = model3.score(X_test, y_test)
print(f"score : {round(model_score,2)}")S

score : 0.92
```

- The trained model was evaluated on the testing set to assess its performance in predicting the number of employees required. The model achieved a coefficient of determination (R^2) of **model score** on the test data, indicating the proportion of the variance in the target variable (**Employees**) that is predictable from the input variable (**PACKETS**).

vii) **Result:**

```
# Make predictions for new data
new_data = [[30000]]
new_data_scaled = scaler.transform(new_data)
predicted_employees = model3.predict(new_data_scaled)
print(f"Predicted Employees for new data: {predicted_employees[0]}")

Predicted Employees for new data: 17.178055210796007
```

- The trained model demonstrated promising predictive performance, as indicated by the coefficient of determination (**model score**) obtained during evaluation. This suggests that the model is capable of effectively estimating the number of employees needed to maximize profit based on the provided number of packets.
- Using the model, we were able to predict the number of employees which were needed in an organization-based on the number of packets entered by the user.

5. **Web Platform Utilization:**

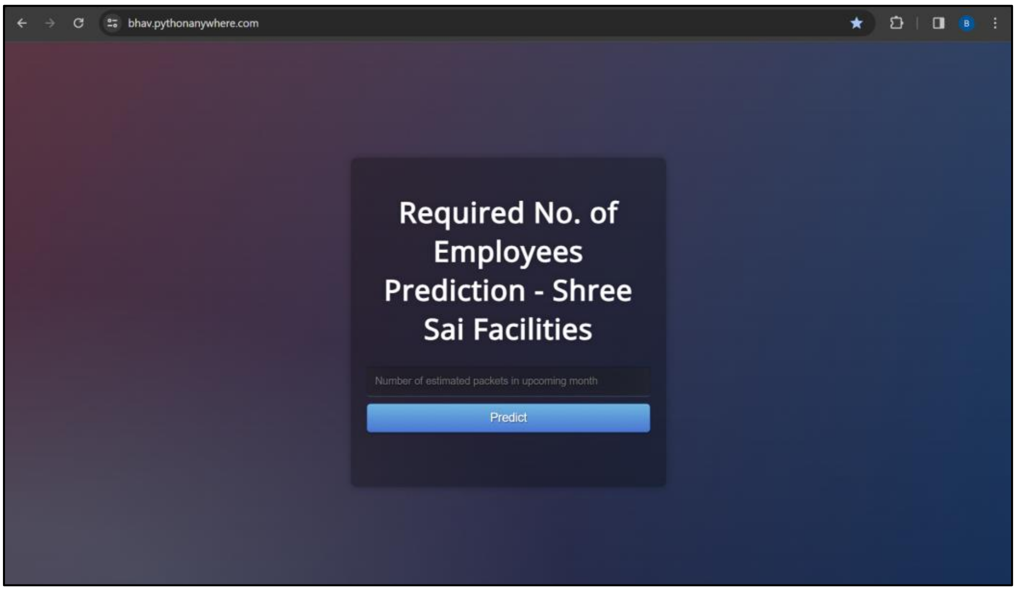


Image: https://bhav.pythonanywhere.com

- **Motive**: Once, the model was created the predictions of different inputs were shared with the organization for feedback, during this phase the problem of a lengthy procedure was identified, the organization contacting us with their input and then us executing the model on a local host and then finally providing the organization with the result seemed quite lengthy and inefficient for the organization.
  Due to this problem the organization might back out after some time as they wouldn't want to waste their precious time just trying to contacting us. To solve this issue website integration seemed to be the best solution, the website being hosted on a free-to-use website domain (pythonanywhere.com) seemed to be the perfect choice.
  The motive behind integrating a URL based website with the model, was the easy usability of our model, by website integration the model could be anywhere and at any time used by the organization.
- **User-Friendly Interface:** The website's UI design, implemented using HTML and CSS, received positive feedback for its simplicity and functionality. The organization commended the ease with which they could input packet estimates and obtain optimized employee predictions.
- **Flask Integration:** Flask, the Python web framework, played a crucial role in developing the backend of the website. Its flexibility allowed seamless interaction between the UI and the predictive model, ensuring a smooth user experience.
- **PythonAnywhere Deployment:** The decision to host the website on PythonAnywhere.com was a strategic one, providing accessibility without the need for local installations. The online platform enabled the deployment of the predictive model, making it available for usage on any device with internet access.
- **Result:** The website hosting our predictive model is now accessible via the following link: https://bhav.pythonanywhere.com/. By deploying the model on PythonAnywhere, we have made it available for utilization on any device with internet access.

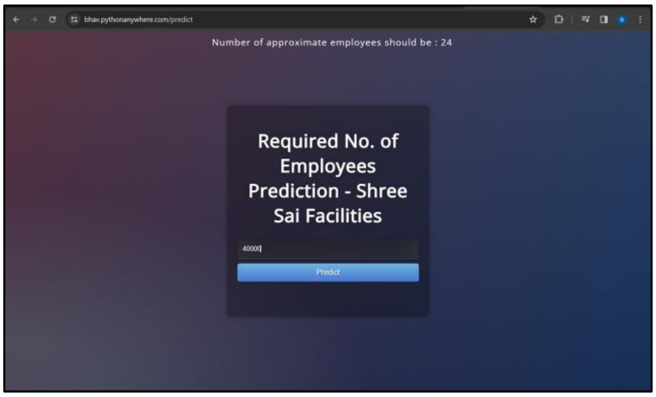


Image: Predicting number of employees using website

This integration not only showcases the practical application of the model but also underscores our commitment to providing convenient and efficient solutions for organizational needs. The website serves as a valuable tool for decision-makers, enabling them to leverage predictive analytics seamlessly in their operational processes.

6. **Positive Impact on Decision-Making:**
    - **Strategic Workforce Planning:** The optimized model, now focused on profitability, offers the organization insights into strategic workforce planning. By predicting the ideal number of employees based on packet estimates, Shree Sai Facilities can make informed decisions to maximize branch profitability.
    - **Operational Efficiency:** The data-driven approach to analysing dependencies between employees, packets, and profitability has led to a refined understanding of operational dynamics. This, in turn, empowers the organization to streamline its workforce for enhanced operational efficiency.
    - **Positive Feedback:** The results and analytics derived in search of model as well as the website integration were well appreciated by the organization and welcomed in their enhancement of their business processes.

4 Interpretation Of Results and Recommendations

This section acts a conclusion to the project and derives upon the previous section and integrates the recommendations on basis of our research to the organization.

The comprehensive analysis conducted on Shree Sai Facilities operational data has provided valuable insights into the dynamics of their business, particularly regarding branch profitability and workforce optimization. From the identification of key variables to the development and integration of predictive models, each step has contributed to a deeper understanding of the organization's challenges and opportunities. Here are the key interpretations of our findings and recommendations for Shree Sai Facilities:

1. **Branch-Wise Profit Analysis:**
    - The analysis revealed significant variations in profitability across different branches, with Faridabad (FAR) emerging as the top performer consistently. This indicates that there are underlying operational factors driving the profitability of each branch, such as employee management, packet volume, and expense control. The most important factor for which FAR branch gained the title of "most profitable branch" for the organization was a large number of employees along with a higher and steadier number of packets i.e. a good employee to packet ratio.
    - **Recommendation**: Shree Sai Facilities should conduct a detailed review of operational practices at each branch, particularly those underperforming, to identify specific areas for improvement. This may involve revising staffing levels, optimizing resource allocation, and implementing cost-saving measures to enhance profitability. Along with this the organization should keep in check the number of people being employed in their branches to maintain a good ratio between them, for which they can utilize our model integrated website as well.

2. **Identification of Independent Variables:**
    - The correlation analysis highlighted employees and packets as the most influential factors affecting profitability. This underscores the importance of strategic workforce planning and efficient packet management in maximizing branch profitability.
    - **Recommendation**: Shree Sai Facilities should prioritize the optimization of employee-to-packet ratios across all of its branches to ensure efficient resource utilization. This may involve implementing dynamic staffing strategies based on demand forecasting and streamlining packet handling processes to minimize overhead costs, for this the organization could keep a monthly check on the number of employees in their branches and the number of packets they're receiving so that if the ratio is not enough in a branch, they could either shift the employees to other branches or ask for more packets in a month.

3. **Model Performance and Optimization:**
    - The development and refinement of predictive models have provided Shree Sai Facilities with a powerful tool for decision-making. By accurately predicting the number of employees needed for optimal profitability, the organization can make informed workforce decisions aligned with their business objectives.
    - **Recommendation**: Continuous monitoring and refinement of the predictive model are essential to ensure its effectiveness over time. Shree Sai Facilities should get the model updated with new data and adjust parameters as needed to reflect changing market conditions and operational requirements either by contacting us or any data analytics organization or professional.

4. **Web Platform Utilization:**
    - The integration of the predictive model into a user-friendly web platform has enhanced accessibility and usability for Shree Sai Facilities. Decision-makers can now access real-time predictions on employee requirements from any device with internet access, facilitating faster and more informed decision-making.

- The website integration was quite applauded by the organization's decision makers as well as their supervisors especially for its easy to use, portable aspect, along with the potential to help in decision making.
- **Recommendation**: Shree Sai Facilities should promote the adoption of the web platform among key stakeholders and provide training on its usage to maximize its utility. Additionally, regular feedback from users should be solicited to identify areas for improvement and refinement.

5. **Positive Impact on Decision-Making:**
    - The project has already demonstrated a positive impact on decision-making at Shree Sai Facilities, empowering the organization to make data-driven decisions for enhanced operational efficiency and profitability.
    - **Recommendation**: Shree Sai Facilities should continue to leverage data analytics and predictive modelling to drive continuous improvement in their business processes. By fostering a culture of data-driven decision-making, the organization can stay ahead of the competition and achieve sustainable growth in the long term.

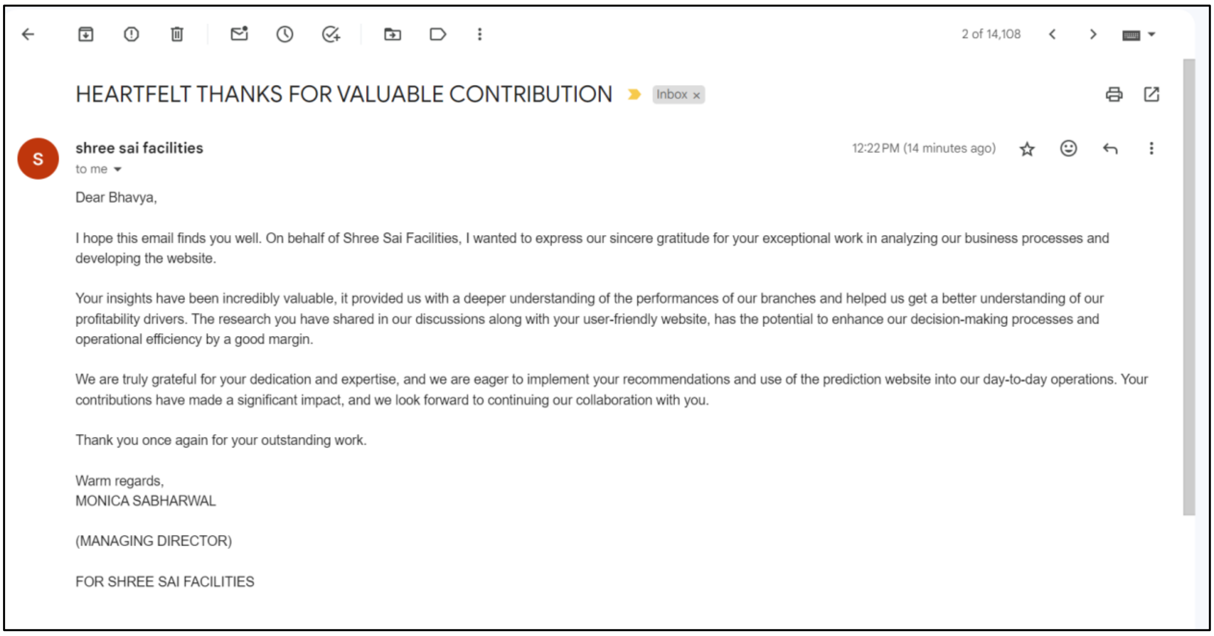
6. **Extended support by us:**
    - We are committed to providing continuous support to the organization beyond the initial integration phase. We will try our best to remain actively engaged in the upcoming months, offering comprehensive guidance and monitoring to ensure the smooth operation of the website and maximize its effectiveness in enhancing the organization's day-to-day processes and profitability.
    - Our extended support encompasses several key areas:
        - Technical Guidance and Monitoring of the Website
        - Strategic Consultation
        - Performance Evaluation
        - Training and Knowledge Transfer
        - Proactive Communication

      By offering extended support in these key areas, we are committed to maximizing the value of the partnership and driving sustained success for the organization.

In conclusion, the project has provided Shree Sai Facilities with actionable insights and tools to optimize their operations and drive long-term profitability. By implementing the recommendations outlined above and embracing a data-driven approach to decision-making, the organization can position itself for success in the competitive courier delivery services market.

5 Feedback

The comprehensive analysis, model optimization, and positive feedback from the organization affirm the success of the project in addressing the initial challenges and providing a valuable tool for informed decision-making. The proof of which can be seen in the mail (shown below) received by us from the head of the organization.


2 of 14,108

HEARTFELT THANKS FOR VALUABLE CONTRIBUTION Inbox ×

shree sai facilities
to me

12:22 PM (14 minutes ago)

Dear Bhavya,

I hope this email finds you well. On behalf of Shree Sai Facilities, I wanted to express our sincere gratitude for your exceptional work in analyzing our business processes and developing the website.

Your insights have been incredibly valuable, it provided us with a deeper understanding of the performances of our branches and helped us get a better understanding of our profitability drivers. The research you have shared in our discussions along with your user-friendly website, has the potential to enhance our decision-making processes and operational efficiency by a good margin.

We are truly grateful for your dedication and expertise, and we are eager to implement your recommendations and use of the prediction website into our day-to-day operations. Your contributions have made a significant impact, and we look forward to continuing our collaboration with you.

Thank you once again for your outstanding work.

Warm regards,
MONICA SABHARWAL

(MANAGING DIRECTOR)

FOR SHREE SAI FACILITIES

Image: Feedback by the organization's MD

6 Important Links

- GitHub link for python code - https://github.com/sam12321/IITM---Vendor-Delivery-Analysis
- Website link - https://bhav.pythonanywhere.com/