A close-up of a logo

Description automatically generated

**Examination Assignment Cover Page**

|  |  |
| --- | --- |
| **Module Code** | **MS5107** |
| Student name and ID **Jacky Kumar and (24259938)** |  |

In submitting this script, I am aware that it is my responsibility to adhere to the examination guidelines. Please tick (double click->… or Y/N) for the following:

|  |  |  |
| --- | --- | --- |
|  | **Yes** | **No** |
| I have read the module owners description/expectations for this exam assignment and the submission process. | Checkmark with solid fill |  |
| I have read the exam assignment guidelines. | Checkmark with solid fill |  |
| I am aware of what the University of Galway plagiarism policy entails. | Checkmark with solid fill |  |
| I have saved the files for submission (e.g., .docx, .xlsx) **following strictly the format and naming** **required** (e.g., **24259938\_MS5107\_ExamAssignment.docx, 24259938\_MS5107\_ExamAssignment\_Q1.xlsx,**  **24259938\_MS5107\_ExamAssignment\_Q2.xlsx**). | Checkmark with solid fill |  |

**Declaration for this Assignment Submission*:***

*I hereby declare that the work submitted is entirely my own work. It* has *not been taken from the work of others, except to the extent that such work has been cited and acknowledged within the text of my work. This work is not done in whole or in part by a machine or through Generative Artificial Intelligence, such as ChatGPT or else. I have not and will not allow anyone to copy my work thereby helping a peer to complete their assignment.*

**Q1. Production Optimization Report**

**To: Management of Celtic Candles Inc.**

**From: [Jacky Kumar]**

1. **Objective**

Finding the best manufacturing schedule for Celtic Candles' holiday candle product line is the aim of this investigation. Maximizing earnings while abiding by particular restrictions on resource availability, display space, and business regulations is the aim. The optimization methodology, the outcomes, and their business-related interpretation are described in this study.

**2. Model Formulation**

**Decision Variables:**  
Five choice factors serve as the foundation of the optimization model and indicate how many of each type of candle should be produced:

* **Large Jars** : Number of large jar candles to produce.
* **Small Jars** : Number of small jar candles to produce.
* **Large Pillars** : Number of large pillar candles to produce.
* **Small Pillars** : Number of small pillar candles to produce.
* **Votive Packs** : Number of votive candle packs to produce.

**Objective Function:**

A screenshot of a table

Description automatically generated

Maximizing the overall profit from the production and sale of the candles is the goal. Each product's profit per unit is as follows:

* Large Jars: €0.25
* Small Jars: €0.20
* Large Pillars: €0.24
* Small Pillars: €0.21
* Votive Packs: €0.16

Thus, the total profit function is:

Maximize Profit=0.25x1+0.20x2+0.24x3+0.21x4+0.16x5

**Constraints:**

The following list of limitations applies to the production process:

1. **Resource Limitations:**  
   The total amount of available resources (wax, fragrance, and wick) limits the production of candles. Each candle type consumes resources as follows:
   * **Wax:** 0.5x1+0.25x2+0.5x3+0.25x4+0.3125x5≤200,0000
   * **Fragrance:** 0.24x1+0.12x2+0.24x3+0.12x4+0.15x5≤100,0000
   * **Wick:** 0.43x1+0.22x2+0.58x3+0.33x4+0.8x5≤250,0000.
2. **Display Space Requirements:**

The following requirements must be met, and the merchant has set aside display space for several candle varieties:

* + **Large Jars and Large Pillars:** 0.48x1+0.23x3≥2 times.
  + **Small Jars and Small Pillars:** 0.24x2+0.23x4≥1.50. times
  + **Votive Packs:** 0.26x5≥1.00 times.

1. **Solver Setup**

The model was implemented in MS Excel using the Solver tool. The following settings were applied:

* **Objective:** Maximize total profit
* **Changing Variables:** x1,x2,x3,x4,x5.
* **Constraints:** All resource and display space were added.

The **Simplex LP** method was selected to solve the linear optimization problem.

**4. Results**

After running the Solver, the optimal production quantities for each product were determined as follows:

**Optimal Decision Variable Values:**

* **Large Jars** : 35,417 units
* **Small Jars** : 326,443 units
* **Large Pillars** : 0 units
* **Small Pillars** : 361,858 units
* **Votive Packs** : 32,693 units

**Maximum Profit:**  
The total profit achieved with this production plan is **€155,364**.

A screenshot of a computer

Description automatically generated

1. **Sensitivity Analysis**

The Solver's sensitivity report provides valuable insights into how changes in resources and profit margins impact the solution.

**Key Findings:**

1. **Resource Availability:**
   * Since the restrictions on fragrance and wax are binding, these resources are used to their fullest potential. Making them more accessible might result in more revenue.
   * Wick's present availability is adequate because it is not a binding limitation.
2. **Profit Margins:**
   * The answer is significantly impacted by changes in the little jars' profit margin. Their large manufacturing volume is the cause of this.
   * Huge pillars and votive packs are produced in relatively little or no quantities, their profit margins are less affected.
3. **Allowable Ranges:**

* Permitted increases and decreases for the goal function's coefficients are displayed in the sensitivity report. The ideal solution won't change if these coefficients are changed within these ranges.

1. **Business Interpretation**

Profits are maximized while all resource, display space, and business constraints—such as the requirement for more jars than pillars—are met by the ideal manufacturing plan. The system offers useful insights into resource usage and product prioritization, and it complies with Celtic Candles' contractual and strategic needs.

**Key Insights**

1. **Small Jars and Small Pillars:**

Their effective use of resources and comparatively high profitability, these two goods take center stage in the production schedule. The majority of the industrial output is represented by small jars and small pillars together, demonstrating their capacity to optimize profit under the existing limitations. Their dominating role in the production mix is further supported by their lower display space requirements.

1. **Large Pillars:**

The production strategy completely excludes large pillars. The main reason for this is because their resource requirements—particularly for wick—outweigh their profit-making potential. Under the current resource limits, they are less practicable due to their smaller profit margin when compared to small jars and small pillars.

1. **Votive Packs:**

Although they make up a little portion of the total profit, votive packets are part of the ideal production plan. The business ought to reconsider their place in the product line in light of their decreased profitability and comparatively large resource consumption. They might be a more lucrative choice if they raised their pricing or used less resources.

1. **Resource Utilization:**

The current approach makes full use of wax and aroma, demonstrating that these are the most important limiting resources. Increased production and higher profitability may be possible if these restrictions are loosened through supplier negotiations or alternative sourcing techniques.

**Recommendations**

In order to further maximize profitability, the study of the ideal production plan identifies important areas for strategic modifications and improvement. The following suggestions are put up for Celtic Candles in light of the Solver model's insights:

1. **Focus on Small Jars and Small Pillars**

Since they are very profitable and use resources efficiently, little jars and pillars dominate the production plan and make up a large portion of the total profit. The production strategy should continue to place a high premium on these goods. Celtic Candles should look into ways to better improve their production processes given their demand and capacity to generate maximum profits under the current limits. This can entail optimizing production processes, enhancing inventory control for associated resources, and promoting these goods to increase their profitability.

1. **Reevaluate Votive Packs**

Despite fulfilling the contractual criteria, the production plan's inclusion of votive packets only gives modest profit margins in comparison to other items. The contribution of the votive pack to profitability should be thoroughly examined by the business. It could be beneficial to think about shifting the resources currently used for votive packs to more lucrative products like small jars and small pillars if their production cannot be made more resource-efficient or their pricing changed for greater returns. As an alternative, votive packs may see a rise in sales if their attractiveness is improved by package upgrades or grouping with other high-margin items.

1. **Optimize Resource Supply**

The production plan's major restricting variables are the limitations on wax and fragrance. When these resources are used to their full potential, the company's potential to make more money is immediately limited. Celtic Candles ought to look at ways to get more of these materials. This can entail finding new suppliers, negotiating advantageous terms with current ones, or investigating less expensive substitute materials that are just as high-quality.

1. **Explore Price Adjustments**

Their inferior profitability in relation to their significant resource consumption, large pillars are excluded from the ideal production plan. Votive packets also make a little money. To increase these items' profitability and competitiveness, the business should think about changing their pricing policies. Their profitability within the product line may be enhanced by price increases that are commensurate with their value proposition and production costs, bolstered by focused marketing initiatives.   
Celtic Candles can optimize profitability while adhering to contractual responsibilities and resource limitations by putting these suggestions into practice. Additionally, these actions will give you the flexibility to adjust to future shifts in resource availability or demand.

**Conclusion**

An efficient plan for increasing sales of Celtic Candles' holiday product line is presented in this paper. The business can meet all contractual and operational requirements while turning a profit of **€155,364** by adhering to the suggested production plan. Profitability could be further increased by making additional changes to pricing, product design, and resource availability.

**Q2. Predictive Modelling for Term Deposit Subscriptions**

**Introduction**

By forecasting consumer reactions, Western Alliance Bank hopes to maximize its telemarketing efforts for term deposits. Finding the clients who are most likely to subscribe can help the bank better manage resources, lower marketing expenses, and increase subscription rates. The creation of a predictive model is described in this paper along with its methodology, the justification for the selected approach, and practical conclusions drawn from the analysis.

We have organized the model creation and analysis process using the CRISP-DM (Cross-Industry Standard Process for Data Mining) architecture.

**Business Understanding**

The goal of the business problem is to increase the effectiveness of telemarketing by forecasting whether a customer would sign up for a term deposit (yes) or not (no). If this project is successful, it will result in:

* 1. Enhanced high-potential customer targeting.
  2. Lowered marketing expenses by steering clear of low-probability clients.
  3. Increased return on investment for telemarketing efforts.

**Key Questions**

1. What are the customer characteristics most strongly associated with term deposit subscriptions?
2. Can we predict with reasonable accuracy which customers will subscribe?
3. How can these predictions translate into actionable business strategies?

**Data Understanding**

Campaign-specific, financial, and customer demographic data are included in the dataset included. This comprises:

1. Age, marital status, and level of education are examples of demographic data.

A screenshot of a computer

Description automatically generated

1. Account balance, loan status (personal and housing loans), and default history are examples of financial indicators.

A screenshot of a computer

Description automatically generated

1. Campaign Features: Call duration, type of contact, and prior campaign results.

A calendar with numbers and letters

Description automatically generated

1. Indicating subscription status (yes for subscribers, nay for non-subscribers), y is the target variable.



**Data Preparation Process**

**Understand the Dataset Structure**

Prior to data preparation:

Examine the file to learn about the data kinds and variables.

y is the target variable (yes/no, binary).

Predictor variables include both category (e.g., employment, marital status, education) and numerical (e.g., age, balance, duration).

Look for possible outliers, inconsistent data, and missing numbers.

**Handle Missing Values**

**Why?**

Missing data can distort model training and affect predictions.

There is no blank and missing data in the data set.

**Encode Categorical Variables**

**Why?**

Categorical variables need to be converted into numeric representations because machine learning algorithms need numerical inputs.

**Steps:**

* **One-Hot Encoding:** Give each category in a variable its own binary column (dummy variable).  
    
  **Marital\_single:** 1 if single, else 0 is an example of how marital (single, married, divorced) is transformed.  
    
  marital\_married: 0 otherwise, 1 if married.  
    
  marital\_divorced: 0 otherwise, 2 if divorced.
* **Remove One Dummy Variable:** Remove one category (the reference category) to prevent multicollinearity.  
    
  Drop marital\_single if marital\_married, marital\_divorced, and marital\_single are already formed.

**Handle Class Imbalance**

**Why?**

The target variable y is probably unbalanced in the dataset because there are more yes (subscriptions) than no (non-subscriptions). The model may be skewed toward the majority class due to class imbalance.  
  
  
**Oversampling:** Use synthetic methods such as SMOTE (Synthetic Minority Oversampling Technique) to increase the number of yes cases.  
  
**Undersampling:** To provide a balanced dataset, reduce the number of no cases by deleting examples at random.  
  
**Class Weights:** Make use of algorithms that give minority classes (like yes) higher weights.

**Final Prepared Dataset**

After these steps, the dataset is ready for modeling. The final structure will look like this:

| **Age** | **Balance** | **Duration** | **marital\_married** | **marital\_divorced** | **Target (y)** |
| --- | --- | --- | --- | --- | --- |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 35 | 1200 | 200 | 1 | 0 | 0 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 42 | 3000 | 150 | 0 | 1 | 1 |

**Modelling**

**Chosen Model: Logistic Regression, Classification Tree (decision tree)**

**Classification Tree Model was selected as the primary model due to:**

1. **Creation of Root Nodes:**

* First, the feature that best divides the dataset into groups is chosen.
* Splits are assessed using a splitting criterion such as the Gini Index or Information Gain.

1. **Splitting Recursively:**

* Every split subset turns into a new node.
* Until a stopping condition (such as a minimum number of samples per node or a maximum depth) is satisfied, the algorithm divides these nodes recursively.

1. **Nodes of leaves:**

* A final class label is applied to these terminal nodes of the tree.
* A class, such as yes or no in the case of term deposit subscriptions, is represented by each leaf node.

**Logistic Regression**

Logistic Regression is the second selected model beacase of:

* **Interpretability:** It gives precise information on how each predictor affects the likelihood of subscribing.
* **Suitability for Binary Classification**: Logistic regression is made especially for yes/no binary outcomes.
* **Efficiency:** Even with a big dataset, computational efficiency

**A screenshot of a computer

Description automatically generatedModel Evaluation and results**

**Classification Tree**

* Confusion Matrix:
  + True Negatives (0 predicted as 0): 3951
  + False Positives (0 predicted as 1): 49
  + False Negatives (1 predicted as 0): 272
  + True Positives (1 predicted as 1): 249
* Metrics:
  + Accuracy: 92.90%
  + Specificity: 0.9877
  + Sensitivity (Recall): 0.4779
  + Precision: 0.8356
  + F1 Score: 0.6081
* Error Rates:
  + Class 0: 1.23%
  + Class 1: 52.21%
  + Overall Error: 7.10%

**A screenshot of a computer

Description automatically generated**

**Logistic Regression**

* **Confusion Matrix**:
  + True Negatives (0 predicted as 0): **3914**
  + False Positives (0 predicted as 1): **86**
  + False Negatives (1 predicted as 0): **343**
  + True Positives (1 predicted as 1): **178**
* **Metrics**:
  + Accuracy: **90.51%**
  + Specificity: **0.9785**
  + Sensitivity (Recall): **0.3416**
  + Precision: **0.6742**
  + F1 Score: **0.4535**
* **Error Rates**:
  + Class 0: **2.15%**
  + Class 1: **65.83%**
  + Overall, Error: **9.49%**

**Comparison:**

**1. Accuracy:** The classification tree model performs better than the logistic model (90.51%) with 92.90% accuracy.   
  
**2. Specificity:** Compared to the clasification model (0.9877), the logistic model (0.9785) is somewhat more accurate in identifying Class 0.

**3. Sensitivity (Recall):** The classification model (0.4779) outperforms the logistic model (0.3416) in recognizing Class 1.

**4. Precision:** There are less false positives in the classification model (0.8356) compared to the logistic model (0.6742).

**5. F1 Score:** A better balance between precision and recall is suggested by the second model's (0.6742) superior performance over the first (0.4535).

**6. Error Rates:** In comparison to the logistic model (65.83%), the classification model has a much lower Class 1 error rate (52.21%), as well as lower overall and Class 0 error rates.

**Key Findings**

**Most Significant Predictors:**  
  
**Length:** Longer calls have a strong correlation with subscriptions.  
  
**Poutcome:**  Customers are more inclined to subscribe if they have had positive experiences with prior campaigns.  
  
**Balance:** A larger account balance is linked to a higher chance of subscribing.  
  
**Job and Age:** Professionals in their middle years are typically more receptive to term deposits.

**Customer Profile for Subscriptions:**

• Age range: 30 to 50 years.   
• Stability of finances: high balance, no defaults.  
• Campaign Engagement: Longer call durations and favorable responses to prior campaigns.

**Recommendations**

1. **Concentrate on Marketing**

Prioritize focusing on clients who are more likely to sign up for term deposits in order to increase marketing effectiveness:

* **Target Demographics:** Since they are more inclined to invest in term deposits, concentrate on middle-aged clients (30–50 years old) who have larger account balances.
* **Campaign Engagement:** Longer call durations indicate a higher level of interest in term deposits from customers. To increase subscription rates, focus your efforts on interacting with these clients.
* **Past Campaign Success:** Due to their responsiveness, customers with a track record of successful interactions (such as successful campaigns in the past) ought to be given preference.

1. **Maximize the Timing of Your Campaign**

To identify the times with the highest subscription rates, examine the monthly contact data (month variable). By scheduling campaigns during the most productive months, you may maximize the timeliness of your telemarketing efforts.

1. **Examine More Complex Models**

Advanced ensemble techniques that are better at capturing intricate patterns and interactions between variables, such as Random Forest or Gradient Boosting, should be taken into consideration for improved prediction performance.

1. **Track the Performance of the Model**

To keep the model accurate, keep updating and retraining it with fresh campaign data. Frequent performance evaluations guarantee that the model adjusts to evolving consumer habits and continues to be a dependable instrument for targeting initiatives.

**Conclusion**

A reliable and understandable method for forecasting term deposit subscriptions is the logistic regression model. The bank may considerably increase the effectiveness of its telemarketing campaigns by concentrating on important factors including duration, outcome, and balance.  
  
**Next Steps:**

* Use the model to target customers in real time while running campaigns.
* Utilize the model's insights to improve customer segmentation and customize advertising.
* To improve accuracy and business results even more, keep looking into new features and sophisticated models.

By ensuring data-driven decision-making, this strategy enables Western Alliance Bank to optimize the return on investment of its telemarketing initiatives.