NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

BC2410 – PRESCRIPTIVE ANALYTICS
GROUP PROJECT REPORT

Making Salad Stop great again

Seminar Class 1, Group 6 (Tuesday - 2.30PM)

Prepared for: Prof Tang Qinshen

Name	Matriculation Number
Cao Thuc Anh	U2010720G
Ernest Ang Cheng Han	U1921310H
Jalvin Nai Guang Jun	U1810925F
Jordan Low Cheng Lin	U2010804E
Liu Qingyi	U1921143C

Table of Contents

Executive Summary

SaladStop! is the first and largest healthy food chain in Asia, that also provides all customers with full access to the nutritional information and serving size of all their menu items.

Apart from its menu items with preset choices of ingredients, SaladStop also provides a Create-Your-Own (CYO) menu, which allows customers to select out of 87 ingredients to form their own salad. However, this leads to problems as it is computationally hard for customers to come up with a menu that accurately matches their needs within the short time frame of queuing and ordering. Hence, to resolve this, we have built a web application using a linear optimisation model, for customers to input their requirements (e.g. nutritional needs, budget, allergies) and the application will then output the ideal list of ingredients for customers to order in their salad.

Furthermore, due to the constant change in population's tastes and preferences as well as the cost of ingredients, there will be changes in demand for ingredients to be ordered. Hence, we have used machine learning to forecast demand and built a stochastic optimisation model using the forecasted demand values to optimize quantities of ingredients to order for SaladStop. This will allow SaladStop to maximize their profits while taking into account the constant change in demand.

1. **Problem**

a. <u>Salad Stop's current limitations</u>

Currently when a customer enters Salad Stop they are faced with a whole slew of ingredients to choose from, starting from 4 different bases followed by 4 wraps, 2 grain bowls, 57 toppings and 19 dressings. Salad Stop also provides a calorie and nutritional breakdown for all its ingredients and while the average health-conscious consumer would appreciate such information, it is too much for any normal human being to process.

Additionally, due to the large variety of ingredients, fluctuations in the demand of each ingredient is also to be expected. These changes in demand can be caused by consumers changing taste and preferences as well as the changes in the cost of ingredients. Currently, Salad Stop is able to only rely on historical demand to make preparations for the day. This method of predicting demand is generally inaccurate and would lead to a fall in customer satisfaction when the ingredients they want are not available.

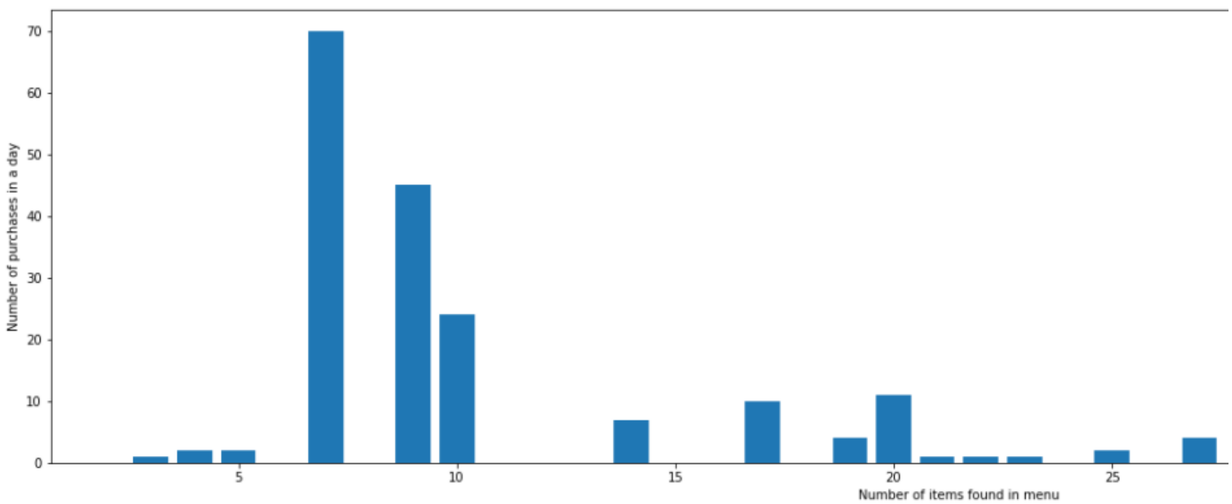
b. <u>Research findings</u>

Studies have shown that complicated options do little to help the decision making process (*HBR: To keep your customers, Keep It Simple* 2014). In the study conducted by Harvard Business Review, it is found that the best tool for measuring consumer-engagement efforts is the "Decision Simplicity Index". Fundamentally, this index indicates that businesses need to be able to make it easy to navigate their options, earn trust from their consumers about the information shown and help consumers evaluate choices.

More specifically to Salad Stop, it is found that "configuration overload" is also another factor which would affect customer purchase decisions (Matzler et al., 2007). The provision of default-settings have been shown to lower the configuration complexity and enable customers to settle on a decision more comfortably without diverting to coping mechanisms against configuration overload which would essentially put off or delay the purchase.

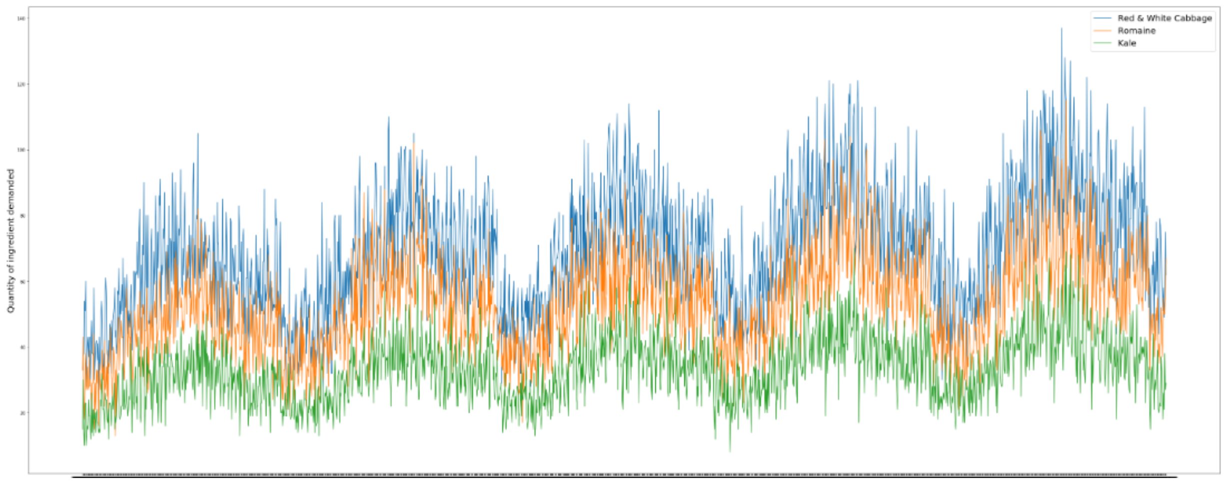
c. Data analytics findings

Based on a research paper referenced, we have generated a dataset to find out the relationship between the number of choices and how it influences the purchase probability of a consumer.



From the following graph, we observe that restaurants with a low number of items on their menu encounter a lower number of purchases. On the other end of the spectrum, restaurants with a high number of items on their menu also encounter a low number of purchases. The optimal number of items on the menu that receives the highest number of purchases seem to be 7 items.

We have also generated an artificial dataset on the ingredients demand for SaladStop using a store item demand dataset in Kaggle.



From the graph above, we can then observe that there are fluctuations in the demand for each of the different ingredients.

2. Our Proposed Solution

a. General overview

Our project would thus focus on two different aspects of the business.

Firstly, we would develop an optimization model to help customers who are not able to decide which ingredients to pick in their customized salad bowl based on their personal nutritional needs. We have also developed a web app proof of concept (POC).

Next, we would also use stochastic forecasting to predict the quantity of each salad ingredient to be ordered to maximize Saladstop's profits.

b. Salad ingredients selection

i. Customer pain points

Given the large number of 87 ingredients for customers to pick from, based on our survey conducted, customers face difficulty making the most optimum choice based on their own requirements. This is because it is computationally hard for them to come up with an order that truly matches their nutrition requirements and budget.

Furthemore, from our research and analysis conducted, the results have shown that when there are too many options in the menu, the probability of purchase is lower. We will build a model based on our artificially generated dataset to be able to predict the purchase probability based on the number of options given to consumers. Through the model, we estimate the purchase probability to be the highest at 8 - 10 menu options.

Hence, with a salad selector optimiser, we can help customers select a salad that meets their requirements, giving them fewer options and increasing the likelihood of purchase.

ii. Other considerations

Some customers will also have dietary and allergy requirements (e.g vegan, vegetarian, no nuts). Our salad selection model will also be able to take these considerations into account when selecting ingredients for the salad.

c. Linear optimization model

Generate salad based on user's input:

Min $9.9 \times s + 11.9 \times t + \sum_{i=45}^{68} x_i \times p_i$

s.t $s + t = 1$ (iv. 1)

$\sum_{i=11}^{44} x_i = 7$ (iv. 2)

$\sum_{i=87}^{69} x_i = 2$ (iv. 3)

$\boldsymbol{x} \geq 0$ (iv. 4)

$\mathbf{min\_nutri} \leq \mathbf{nutri}' \times \boldsymbol{x} \leq \mathbf{max\_nutri}$; with min_nutri & max_nutri are vectors of levels of nutritions like protein, fat set by customers, nutri is a matrix of dimension (number of ingredient x number of nutritions) (iii. 1)

$M \times r_1 \geq \sum_{i \in Vg} x_i$, $M \times r_2 \geq \sum_{i \in Vt} x_i$, $M \times r_3 \geq \sum_{i \in G} x_i$, $M \times r_4 \geq \sum_{i \in D} x_i$, $M \times r_5 \geq \sum_{i \in N} x_i$, $M \times r_6 \geq \sum_{i \in S} x_i$; with Vg, Vt, G, D, N and G are sets of indices of vegan, vegetarian, gluten, dairy, nuts and sugar dishes, i.e. Vg = {1, 2, 3, 4, 9, 10,..., 82}, S = {20, 72, 75, 87}; M is a large number that is at least equal to the sum of x on the RHS; r_i with i = 1,...,6 is the dietary requirements corresponding to the sets of sum of x (iii. 2)

$\sum_{i=45}^{68} x_i \geq \text{min\_pr}$, with min_pr is the minimum number of premium topping (iii. 3)

$9.9 \times s + 11.9 \times t + \sum_{i=45}^{68} x_i \times p_i \leq \text{budget}$ (iii. 4)

Details on Salad Stop: Link: https://www.saladstop.com.sg/cyo/

1. Choose either premium base or standard base (standard base includes wrap and grain bowls)
2. Choose 2 dressing
3. Choose 7 standard toppings (even if you chose premium base)
4. Can add as many premium toppings as you want but will have additional charges

i. Objective function

The objective function will be customer centric, to minimize cost paid by the customer while meeting user input constraints when selecting the ingredients required for the salad. This objective function is preferred because the problem stems from the customer side rather than the business side. Whereby customers are confused about what to order due to the presence of too many choices. If we are optimizing the choices customers should make, then we should minimize the cost of the salad paid by the customer as the customer would want to pay as little as possible for their salad.

ii. Decision variables

The selection of an ingredient in the salad ([nx1] binary variables) : x where n = 87
The standard base selection (binary variable) : s
The premium base selection (binary variable) : t

iii. Core constraints

The core constraints are customer centric to focus on the needs of the customer etc, essentially the constraints for the nutritional requirements

1. Ensure that the selection of ingredients meets nutrition requirements of user
2. Ensure that the dietary needs of user is met
3. Ensure that the number of premium toppings meet user requirements
4. Ensure that the total cost of the salad is within the user's budget

iv. Secondary constraints

Being customer centric, the secondary constraints focuses on the other needs and wants of the customers that we have added into the model

1. Ensure that either standard base is selected or premium base selected. Cannot be neither selected or both selected
2. Ensure that only exactly 7 toppings unless the user wants more will be chosen
3. Ensure that only exactly 2 dressings unless the user wants more will be chosen
4. Ensure that x is non-negative

v. Results

With this linear optimization model, we will be able to select a set of ingredients to form a salad for the users and calculate the total profits made for each salad for the business.

d. <u>Ingredients order quantity</u>

i. Business pain points

There is a constant change in the demand of ingredients to be purchased due to the changing population taste and preferences as well as the cost of ingredients. Hence, it is difficult to determine the future demand of ingredients and the optimal purchase quantities of ingredients and selling prices so as to maximize Saladstop's expected profit, subject to a total space constraint for the storage of our ingredients.

ii. Other considerations

We will have to consider the demand of ingredients when using the stochastic model. Hence, we modeled the demand data for the ingredients using a kaggle demand dataset. Thereafter, to get the current demand of ingredients, we used a LSTM neural network to predict the demand levels of ingredients and use those future forecast as inputs to the stochastic model to get the optimal order quantity in a further period

Furthermore, by ordering the optimized quantity of ingredients, we will be able to reduce food waste created, aligning with SaladStop's goal of being an eco-friendly organization.

e. Stochastic optimization model

Generate quantities of ingredients to import by Salad Stop and the expected amount of profits

Max $-\sum_{i=1}^{87} c_i x_i + \frac{1}{K}\sum_{k=1}^{K} p_i t_{ki}$, with c_i & p_i as the cost and selling price of each ingredient

s.t $t_{ki} \leq x_i$ (iii.1)

$t_{ki} \leq d_{ki}$ (iii.1, C.iii.2)

$\sum_{i=1}^{87} s_i x_i \leq C$, with C as total space and s_i as the amount of space occupied by an unit of ingredient (iii.3)

$\boldsymbol{x} \geq$ min_order (iii.4)

i. Objective function

The objective function will be to maximize the average daily profit to be earned by Salad Stop based on the number of ingredients to be ordered.

ii. Decision variables

Amount of ingredients to order ([nx1] integer variables) : $\boldsymbol{x}$ where n = 87

Matrix of sale quantity of the ingredient for each number of days in the dataset ([nxK] integer variables) : $\boldsymbol{t}$ where n = 87, K = number of days

iii. Core constraints

Constraints based on the ingredient price, cost, demand, space and total space of the store to determine import quantities for each ingredient

1. Constraint to linearize term min[px, pd] where px is the expected amount of quantity to order and pd is the expected amount of demand (Reference: Newsvendor model formula)
2. Ensure that the amount of ingredients to order for each day is less or equal to the expected demand
3. Ensure that the space based on amount of ingredients to order for each day is less or equal than total space available
4. Ensure that the amount of ingredients to order is at least equal to the minimum order per day recorded (due to the estimated pricing, some ingredients' selling prices are less than their costs)

iv. Results

With the combined method of machine learning to predict the future demand and the stochastic optimisation model, we will be able to obtain the daily number of ingredients to be ordered as well as the average daily profit for Salad Stop.

f. Interactive API proof of concept

A web app POC was developed to allow potential customers to provide their nutritional, dietary and cost-related preferences and requirements. These requirements collected would be channeled into the linear ingredient optimization model as constraints. After the inputs, the customer would be able to generate a salad with a list of ideal ingredients that fulfill their dietary needs at the click of a button. The customers would then be able to use the list generated to place their orders effortlessly, knowing that the end product would be a salad that fits the preferences stated.

3. Implementation Feasibility

Our team believes it is highly feasible to implement our linear optimization model on SaladStop!'s business. The model and proof of concept web program can be easily implemented onto SaladStop's current online ordering system, namely the "Create-your-own" section. As the customer inputs his/her nutritional, dietary and monetary preferences, the model will generate an ideal salad with ingredients selected to fit the customer. A link could be then provided to direct the customer straight to the check-out page.

a. Potential limitations

One potential limitation with regards to implementation of the ingredient optimizer for customers at SaladStop!'s physical locations could be the issue of having electronic point-of-sales systems for the customers to enter their preferences.

Model wise, the team foresees that it may be repetitive for loyal customers who consistently patronize SaladStop!. Assuming the customers' dietary, nutritional and monetary preferences do not change, it is likely that the model will provide the same few ingredients (with minimal variations due to factors such as ingredient stock), and as such provide a monotonous dining experience. One easy fix could be to introduce an element of randomness in the linear ingredient optimizer such that various combinations of ingredients that fit identical constraints would be provided to the customers thereby reducing the number of repeat recommendations.

Regarding the stochastic model, one of its limitations in this case would be the input dataset. Particularly, the dataset used in our proposed stochastic and LSTM model has a level of risk, which has some known frequency of occurrence and a clear trend. If the real dataset collected by SaladStop! varies without a clear frequency of occurrence, the stochastic model and LSTM might not be able to fit themselves to the context. In such cases of ambiguity, another approach could be to try a robust optimization model or twisting some hyperparameters of LSTM model such as number of hidden layers, activation function.

b. Usefulness of new information

Given that SaladStop! collects information regarding how effective the linear ingredient optimizer has been for the customers, as well as the most popular ingredients as per customers' preferences, SaladStop! can make use of this information to create menu items with generalized classifications of the nutritional preferences - such as for "Gym Addicts" who often require high amounts of protein and low fat.

Next, the stochastic model would provide highly accurate information regarding how SaladStop! should purchase its ingredients. Having a clear understanding of demand would allow the ingredients to be ordered with utmost timeliness and hence ensuring freshness as well as minimizing waste. This would maximize the taste of the salads and hence customer satisfaction as well as reducing waste in resources as the perishable ingredients would be utilized within its shelf lives.

c. Progress tracking

The most straightforward way to track the effectiveness and progress of the linear ingredient selection model would be to track the number of orders that follow the recommendations provided. Depending on the success of this model, SaladStop! can then proceed accordingly. The next step could be to collect reviews and feedback on the recommendations from customers who have used the model. Adjustments could then be made to better the model in terms of its recommendations returned to suit customers better.

Our team acknowledges that it is important to track customers' satisfaction with the ingredients provided and garner constructive feedback to further optimize the model to better serve customers since it is easily tuned via the model constraints.
On the other hand, to track the stochastic model's progress, our team suggests to track how often SaladStop! has to make orders of their salad ingredients off-schedule since this would suggest an inaccuracy in the model. Should the ordering of ingredients be consistently timely, it would point towards a success in applying the stochastic model. More demand data should then be collected to improve the accuracy of the stochastic model.

4. Evaluation and Conclusion

a. Other recommendations

Constraints for Salad Optimizer

As salad stop is well known for it being one the first businesses in the world to fully carbon label their entire menu, our model can also potentially look to include the total carbon footprint of the salad and ingredients. From a business perspective, this would enable the more environmentally conscious customers to be fully aware of the carbon impacts of their meal, helping build on the environmentally friendly image of the business.

Sensitivity analysis

Sensitivity analysis could also be further done for saladstop to find the shadow price of each salad ingredient. This could help saladstop better price differentiate for each of their ingredients beyond just premium and standard. However, an additional caveat to conducting sensitivity analysis and finding the shadow price would be that from a business perspective, too much price differentiation might also turn customers away when they cannot see the value differentiation between products. If further sensitivity analysis is to be conducted and price differentiation implemented, it would be key to remain practical and introduce one or two more levels of price differentiation at most. Further differentiation can come in the form of recommended orders that would be priced cheaper than if normally purchased in a custom order.

Applicability of Models

The salad optimizer and stock optimiser models can also be easily modified to be used on other businesses. Subway, Wok Express, Pasta Express, Maki Sushi are all businesses with a similar model that would be able to make use of the Optimizers for their customers to get a recommendation based on macronutrients, calories, dietary requirements and price. The stock optimiser would also allow these businesses to predict and purchase the optimal quantity each day or week.

Further uses of these models could be used in other businesses as well in order to help lower costs to customers and solve their own multiple newsvendor problem as well.

b. Conclusion

In conclusion our project managed to achieve two things; One, a proof of concept to address a common customer pain point at saladstop, to help customers find the cheapest options available yet still meeting their requirements for macronutrients, calories and other dietary requirements. Two, through the use of machine learning and Stochastic optimisation, accurately predict the demand and suggest a recommended amount of each ingredient to saladstop to purchase and prepare for the upcoming day and week.

Appendices

Appendix A: Explanation of Variables

References

Create your own. Create Your Own – SaladStop! (n.d.). Retrieved Mar 30, 2022, from https://www.saladstop.com.sg/cyo/

Wang, Y., Luo, X., & Lin, Z. (2019, December 12). *Too little or too much seller assortment: The effects on buyers' purchase probabilities in a food sharing platform*. SSRN. Retrieved March 30, 2022, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3493371

New York University Stern School of Business Auditing Spring 2022. (n.d.). Retrieved March 30, 2022, from https://www.stern.nyu.edu/sites/default/files/assets/documents/Spring%202022%2011am%20Audit%20Syllabus.pdf

Store item demand forecasting challenge. Kaggle. (n.d.). Retrieved April 5, 2022, from https://www.kaggle.com/competitions/demand-forecasting-kernels-only/data?select=train.csv

A Data-Driven Approach to Reduce Food Waste. (n.d.). Retrieved April 10, 2022, from https://research.sabanciuniv.edu/42477/1/10402428.Sancaktaro%C4%9Flu__Afsin.pdf

Matzler, K., Waiguny, M., & Fuller, J. (2007, September 25). *"Spoiled for Choice: Consumer Confusion in Internet-Based Mass Customization."* Business Perspectives. Retrieved April 17, 2022, from https://www.businessperspectives.org/images/pdf/applications/publishing/templates/article/assets/1867/im_en_2007_03_Matzler.pdf

To keep your customers, Keep It Simple. Harvard Business Review. (2014, August 1). Retrieved April 17, 2022, from https://hbr.org/2012/05/to-keep-your-customers-keep-it-simple

Appendix A: Explanation of Variables

menu.csv

Variable Name	Details
Ingredient	Ingredient to be selected to form the salad
Ingredient type	Type of ingredient (e.g standard base, premium base, standard topping, premium topping, dressing)
Price	Selling price of ingredient in premium topping as additional costs is required when premium topping is added.
Serving Size	Serving size of each ingredient to be added in grams (g).
Cost of Goods Sold (COGS)	Cost of purchasing each unit of ingredient. Unit is specific in next column
per (g, pcs, unit)	Value of unit of ingredient at cost price (e.g. 100g)
COGS_per_serving	Cost of purchasing one serving of each ingredient
Calories	Number of calories of one serving of each ingredient in kcal
Carbohydrates	Amount of carbohydrates in one serving of each ingredient in grams.
Protein	Amount of protein in one serving of each ingredient in grams.
Fat	Amount of fats in one serving of each ingredient in grams.
Sugar	Amount of sugar in one serving of each ingredient in grams.
Sodium	Amount of sodium in one serving of each ingredient in grams.
Fibre	Amount of fibre in one serving of each ingredient in grams.
Cholesterol	Amount of cholestrol in one serving of each ingredient in grams.
Carbon footprint	The amount of carbon footprint produced to serve the ingredient (i.e. through delivery processes,...)
Vegan	Binary variable, whether the ingredient is suitable for vegan (1) or not (0)
Vegetarian	Binary variable, whether the ingredient is suitable for vegetarian (1) or not (0)
Gluten	Binary variable, whether the ingredient is gluten (1) or not (0)

Dairy	Binary variable, whether the ingredient is dairy (1) or not (0)
Nuts	Binary variable, whether the ingredient is nut (1) or not (0)
Spicy	Binary variable, whether the ingredient is spicy (1) or not (0)
Sources	Reference source for the COGS of the ingredients