Faculty of Engineering

INDUSTRIAL ENGINEERING

Undergraduate Thesis – Second Semester 2020

[201012] Development of a PSO metaheuristic for the menu planning problem with DCE preferences analysis

Sara Campuzano Grajales[a,e], Juliette Marina Martínez Bermúdez[a,b,e], Andrea Carolina Rodríguez Moreno[a,e],

Juan Carlos García[c,e], Gloria Bernal[d,e]

[a]*Student of Industrial Engineering*
[b]*Student of Economics*
[c]*Professor, Director of Undergraduate Thesis, Department of Industrial Engineering*
[d]*Professor, Director of Undergraduate Thesis, Department of Economics*
[e]*Pontificia Universidad Javeriana, Bogotá, Colombia*

Abstract

Menu planning is an optimization problem that seeks to minimize the cost of food inputs for its preparation while satisfying nutritional needs. This project will be applied in Colegio Mayor de San Bartolomé, where over the years the nutritional content of the lunch provided to students has not been analyzed in depth, neither a distinction has been made between the portions supplied to different age groups and their nutritional needs have not been fulfilled. On the other hand, it is a subsidized school with limited resources; therefore, it is pertinent to improve their efficiency in their expenditures, which can be achieved by optimally minimizing their costs. In order to address this problem and understand the tradeoff in food preferences faced by sixth to eleven grade students enrolled in the school, as well as the effect that the provision of nutritional information has on them, a DCE (Discrete Choice Experiment) is designed. Then, a linear optimization mathematical model is proposed as an approximation to solve the menu planning problem however, it is shown that this particular case exposes characteristics of an NP-hard problem; thus, a metaheuristic such as PSO (Particle Swarm Optimization) is highly convenient. This metaheuristic aims to minimize both the nutritional deviation concerning the requirement and the cost of food supplies and generate a varied menu solution, contemplating a time horizon of 15 days. Therefore, an algorithm considering the PSO, the estimated preferences, and the restrictions specified by the school, is developed, generating solutions for the previously stipulated age ranges according to the result of hierarchical clustering. Finally, an application designed under the ISO / IEC 25010 standard facilitates the entry of data and the visualization of the results. The project is carried out under the DMAIC methodology following the corresponding steps.

Keywords: menu planning problem, nutrition policy, nutritional preferences, nutritional information, discrete choice experiment, econometrics, optimization, multi-objective model, mathematical model, NP-hard problem, metaheuristics, particle swarm optimization, clustering,

1. Justification and approach to the problem

Adequate nutrition is essential in early childhood, childhood, and preadolescence since in these periods the nutritional status has a significant effect on physical development that is generated during puberty (Soliman et al., 2014). Likewise, nutrition plays an important role in the development and prevention of metabolic diseases (Bennett et al., 2015). Additionally, during the aforementioned period of life, there are patterns of consumption, tastes, and preferences in the choice of food, where a low level of education, poor knowledge about child nutrition, and low socioeconomic status of the household accentuate the likelihood of having nutrition problems (Yeasmin & Islam, 2016). Altogether, proper nutrition for children regardless of their family status or school resources should be guaranteed.

Some studies link obesity, through the Body Mass Index (BMI) with demographic variables, prices of unhealthy food, and participation in subsidies for the low-income populations. In addition, poverty is positively associated with obesity as low-income consumers tend to eat caloric and low-priced foods, to reduce food costs (Zhang et al., 2011). Besides, according to a quasi-experiment to evaluate a nutrition education program, it was shown that nutrition education in low-income children and parents generates positive changes in the diet, suggesting that attractive methods engaging parents could help their young children to make healthy eating decisions (Blitstein et al., 2016).

Moreover, according to the UNICEF 2019 report on the state of the world's children, millions are eating well below or well above the required calories. Diets tend to be poor in nutrients, thus increasing the risk factor that is related to the appearance of diseases. Table 1 shows the food situation of people in the world segmented by Latin America and the Caribbean and Colombia.

Table 1.Hunger and malnutrition in the world (in millions of people and percentage), 2018.

	Undernourished people	Chronic malnutrition in children under five years of age	Overweight in children under five years of age	Obesity in adults
World	789.1 (10.7%)	154.6 (22.9%)	40.6 (6%)	640.9 (12.8%)
Latin America and the Caribbean	40.7 (6.4%)	5.9 (11%)	3.7 (7%)	96.1 (22.8%)
Colombia	3.4 (7.1%)	5.2 (10.8%)	3.0 (6.3%)	9.0 (18.7%)

Source: Perspectives on the Right to Adequate Food and Nutrition, FIAN Colombia. (2018).

The main problems in terms of nutrition in Colombia are overweight, malnutrition, and lack of micronutrients. For example, children that suffer from overweight can lead to early-onset type 2 diabetes, stigma, and depression. In Colombia, the excess weight in school-age children has reached 24.4% and the delay in sizes, which refers to the lack of growth in expected average sizes of children between 5 and 12 years old, has been 7.4%. This limits and conditions the intellectual production of the population (MinSalud; DPS; INS; ICBF; UNAL, 2018), which should be a concern for Colombians in general, and schools in particular.

A good part of the daily feeding of children in Colombia occurs during study hours at school. Colombian children generally spend at least 5 to 6 hours a day in educational institutions (MEN, 2002), a day that includes the usual lunchtime and one or two snacks a day. In some cases, the School Feeding Program (PAE, Programa de Alimentación Escolar), a strategy created by the Ministry of National Education, provides children with food during the school day in order to positively impact learning processes and reduce absenteeism. This program has been successful in some schools, but it has also presented persistent failures in some regions, such as poor nutritional quality or even lack of food delivery due to corruption (Semana, 2019). These failures specifically affect the poorest schools or regions of the country, because that is precisely where there are more institutional weaknesses and where there are more beneficiaries of the program.

Finding a method that allows specifying the correct way in which schools should spend their resources to meet the nutritional needs of their students could help improve the current conditions in which food menus are delivered to students in the country. Taking into account student's preferences for particular bundles of food and information about nutritional requirements enriches this aim; as students will consume food that is more appealing according to sensory aspects and its way of preparation (Cattaneo & Proserpio, 2016). We focus on one low-middle income school in Bogotá, Colegio Mayor de San Bartolomé, which does not have a system that arranges meals accounting for nutritional requirements, student´s preferences, or cost-effectiveness. The study was carried out accounting for meals records provided by the school, and the participation of 270 students of the school in sixth to eleven grade.

Beyond the nutritional content of the food, we complement the analysis by account for students' preferences for food. Evidence suggests that motivations and attitudes towards a certain type of food generate a significant impact on its consumption (Bastamipour et al., 2013) and the provision of nutritional information has an impact by increasing the consumption of healthy foods (Burton et al., 2006). Therefore, we take into account the student's preferences for food in this analysis, as well as the impact that nutritional information has on those preferences.

On the other hand, the nutritional recommendation stipulated by Resolution 3803 of 2016 establishes values according to the particular group to which each person belongs, according to the period of life, gender, and physical activity. These factors are not taken into account in the school, since specific standard portions have not been established for each group. Calculations made based on three age groups (10 to 14 years and 15 to 19 years) that contain the target student population indicate that the recommended amount of calories is lower than that currently served in the school. Additionally, they exceed the recommended protein and carbs percentage for all groups and are fat deficient concerning the recommended percentage. It should be noted that they currently have an approximate cost of 23,354,697 COP in three weeks, that is, 15 days of lunch. The calculations and the nutritional components for the current school menu are referenced in Annex 2.

In this research, we use optimization concepts, operations research, and programming to create an application capable of making input purchase decisions and menu stipulation, guaranteeing nutritional requirements, and reducing costs for the Colegio Mayor de San Bartolomé. Likewise, it is necessary to evaluate the preferences of the students under an econometric estimation to determine possible characteristics to consider in the planning of the menu. Subsequently, it is pertinent to analyze the impact that the proposed solution would generate through a Before and After Evaluation that compares the current scenario with the one proposed, thus establishing the impact of the solution.

The problem described leads us to ask the following research questions:

What are the preferences for food faced by sixth to eleven grade students at the educational community of the Colegio Mayor de San Bartolomé considering the effect of providing nutritional information, measured by an econometric model?

How to improve nutrition in the lunch menu of high school students, taking the Colegio Mayor de San Bartolomé as a case study, through an optimization model that reduces costs related to food and contributes to meeting nutritional needs, taking into account the food preferences of the educational community?

2. Literature Review

Planning is a task that involves finding a sequence of events that lead to a specific task with certain constraints for subsequent execution. In particular, the problem of menu planning is to find an optimal combination of foods that achieve structural and nutritional characteristics for a sequence of days. In this way, the Ministry of National Education (MEN) establishes in Resolution 16432 a standard minute with food and nutritional aspects for each food supplement according to the age groups defined in the dietary guidelines for the Colombian population. Subsequently, a professional in nutrition and dietetics must develop 21-day menu cycles taking into account the availability of food and eating habits, and performing an analysis of calories, nutrients, and preparation guides.

Therefore, current research has a metaheuristic approach to this problem, as it seeks to find the best combination of elements that can satisfy every constraint (Ngo et al., 2016). Thus, a healthy and balanced diet provides an adequate combination of energy and nutrients for the body. Consequently, for the optimization and resolution of this problem, different methodologies have been applied, (see Annex 3). Then, one of the most popular quantitative techniques for the election of diets is linear programming, which was used for the first time in 1959 where the main objective was to minimize the total cost considering the satisfaction of all nutritional requirements. (Ngo et al., 2016). Other authors adopted this method for solving the problem of finding inconsistencies since linear programming restricts elements such as food variety and user preferences. Similarly, the complicated interrelationship of constraints makes nutritional balance difficult to achieve (Anderson & Earle, 1983). In other studies where user preferences were considered, not every nutritional requirement was reached.

There are some computational intelligence techniques that help reduce menu planning time and make the job easier. For example, genetic algorithms can explore various spaces for solving a problem and search for the best within a set of possible solutions considering different perspectives. So, it is proved that with the corresponding restrictions, the evolutionary approach can generate a set of economic and healthy menus (Moreira, Wanner, Martins, & Sarubbi, 2018). Despite this, the genetic algorithm was executed in a problem in Indonesia in order to alleviate infant mortality due to malnutrition resulting in a precision of 97,87%, at the same time that the binary swarm optimization algorithm (BPSO) was implemented with a precision of 99,14% (Ayu & Muhamad, 2019). Certainly, the BPSO has proven its effectiveness in solving this problem and is proven in multiple areas, as well as the implementation of its variants.

Furthermore, the Particle Swarm Optimization (PSO) algorithm along with its multiple variations has been developed in problems cataloged as an “NP-hard problem”, because it is a population-based competitive algorithm, efficient with numerical optimization problems (Wu et al., 2014). Also, easy implementation, high quality of solutions, efficiency, and speed of convergence are strengths of PSO (Guo et al., 2014). Thus, it was used to generate healthy lifestyle recommendations that fit people’s parameters (Fister et al., 2017) and for older people, with the help of the collection of their personal health information (Feier et al., 2017). Also to solve the problem of dietary planning for adults by meeting nutritional needs and meeting the required calories, also, variety in dishes was taken into account through the appropriate proportion of groups in each meal (Porras et al., 2019), among others. However, the standard algorithm has some shortcomings, including premature convergence, this is why it must have promising exploration capacity. Additionally, the PSO has also been combined with traditional optimization algorithms and some evolutionary optimization algorithms in order to pay back their weakness and take advantage of the characteristics that these algorithms propose (Zhang et al., 2015).

Moreover, menu planning models can be broken down into two classes, single-objective problems, and multi-objective problems. Usually, when treating multi objectives models, linear programming of objectives or *Goal Programming* (GP) is used, where, to find the best possible diets, a weighted sum of unwanted deviations is minimized, minimizes the highest amount of weighted unwanted deviations, or minimizes the weighted sum of the above two target functions. Therefore, it is ideal to formulate a GP model that handles qualitative values for the unwanted deviation of the nutrient variable through Fuzzy Sets and thus handles the problems of scaling and incommensurability (Gerdessen & De Vries, 2015). On the other hand, exploring other methods that include multi-objective functions, is the Technological University of South China, where (Chen, 2019) adopted a specific model based on the multiobjective evolutionary algorithm with dual-level files, demonstrating better results than other single-level algorithms.

In addition to the traditional menu planning problem and its aforementioned variants, it is possible to consider the preferences of the users at the same time that the feasibility of introducing healthy foods is determined. The above is possible by performing a DCE (Discrete choice experiment) that provides preferences to hypothetical options before being implemented. For example, this analysis was implemented in a low-income community in popular diners in Peru (Buttorff et al., 2015), through a multi-profile and better-worse version of DCE, where participants chose their favorite and least favorite menu of three hypothetical options. Preferences were examined and the results indicated that healthy additions to meals were feasible, since participants were willing to pay more for it. In addition, previous studies have shown the importance of the proportion of information on more nutritional food choices (Koç & Van Kippersluis, 2017). Although preferences are taken into account when generating nutritional diet recommendations using PSO (Pop et al., 2013), those are mostly individual recommendations and nutritional preferences are not established through a DCE. To the best of our knowledge, this is the first study in Colombia that solves the menu planning problem with PSO for a study case taking into account group nutritional preferences using a DCE.

Finally, it is essential to offer free-to-use technological tools that provide information on an adequate intake in an agile and precise way. Therefore, in this work, we develop a solution that provides better menu planning in the school "Colegio Mayor de San Bartolomé", for high school students. In this case, we apply a PSO metaheuristic to the menu planning problem in order to satisfy nutritional requirements while reducing total cost. Additionally, taking into account the preferences of the students to determine the characteristics that should be included in the construction of the menu, a DCE will be elaborated, which also allows estimating the effect of providing nutritional education on the preferences of the students. Considering the variety of menus, group preferences, a multi-period planning horizon, and differentiated portions according to anthropometric characteristics, this case is considered a problem that is difficult to solve through optimization methods such as linear programming; however, we will take as a reference our mathematical model with simpler characteristics to compare our solution with a reduced optimum.

3. Objectives

To develop an application for the school menu planning based on a PSO metaheuristic with multiple objectives of cost reduction and satisfaction of nutritional needs, considering the dilemmas in food preferences faced by the educational community.

1. Estimate the nutritional preferences of the educational community through the econometric estimation of a DCE, in the presence and absence of nutritional information.
2. Design, implement, and evaluate a mathematical model based on the PSO metaheuristic for cost reduction and satisfaction of nutritional needs, taking into account the nutritional preferences of the educational community.
3. Design an application for easier data input, execution, and display of results of the model based on the PSO metaheuristic.
4. Contrast the cost and nutritional efficacy of the current state of the educational institution, with the results obtained by the model.

4. Methodology

1. Estimate the nutritional preferences of the educational community through the econometric estimation of a DCE, in the presence and absence of nutritional information

1.1. Identification of attributes and assignment of levels

We use the Discrete choice experiment (DCE) in order to elicit student's preferences for hypothetical meal bundles following the attributes used in the best-worst DCE executed in Peru as reference (Buttorff et al., 2015). Six attributes were selected to measure the multi-dimensional trade-off in student preferences for price, protein, soup, variety in carbohydrates, salad, and dessert. The choices within the attributes (the levels) are made as realistic as possible. The attributes, levels as well as the definition and translation in Spanish are summarized in table 2.

Table 2. Menu characteristics and levels used in the experiment.

Menu characteristics	Levels	Definition	Spanish translation
Price	6,000, 8,000, 10,000	The price is approximately between 1.58 USD and 2.64 USD. The current price of the school lunch is contained in this range.	Colombian pesos (Colombian currency)
Protein	Fish, white meat, red meat, vegetable protein	It is the type of protein that the dish has.	Proteína
Soup	Vegetable soup, vegetable cream, no soup	If the lunch does not have soup or the type included.	Sopa
Variety of carbs	1,2,3	Number of kinds of carbohydrates in a single dish.	Variedad carbohidratos
Salad	Fruit salad, cold salad, hot salad	Type of salad.	Ensalada
Dessert	Homemade, non-homemade	Type of dessert, where homemade means prepared in the cafeteria or bought otherwise.	Postre

1.2. Experimental Design

Due to the number of attributes and levels considered in our experiment, the number of all possible combinations of attribute levels results in a full factorial design matrix with 648 rows, which implies a combination of 209,628 pairs. In order to reduce the number of choice sets, we use D-efficient, a statistically efficient design available as a module in Stata which guarantees minimum variation around the parameter estimates by minimizing the estimated standard errors in a multinomial logit (MNL) model. We produced 18 scenarios which are split into two versions of nine scenarios. This will be the number of choice sets per block to avoid fatigue in students. As it is a binary choice set framework, students may select the preferred bundle out of two possible in every scenario (see figure 2).

Furthermore, the random utility theory provides the theoretical basis for the DCE (McFadden, 1974). In this case, the model and the equation derived are shown below, in table 3.

Table 3. DCE Model applied to nutritional preferences.

Description	Equation
In this case, student *n* is assumed to choose between J alternative lunches, choosing the one associated with the highest utility. Thus, individual n will choose lunch *i* over *j* if the equation is satisfied. Where U is the utility for a given lunch.	$U_{ni} > U_{nj} \ \forall i \neq j \ \in$ (1)
However, the utility of any given lunch is not directly observable. The choices are modeled in terms of probabilities. Therefore, the probability (P) of an individual n who chooses lunch *i* over *j* can be estimated.	$P_{ni} = [U_{ni} > U_{nj}] \ \forall i \neq j \ \in$ (2)
The utility (U) associated with a particular lunch is made up of the deterministic component V_{in}, which is a function of m lunch observed attributes (x_1 ,..., x_m), and ε_{in}, the unobserved lunch attributes and the individual-level variations. The utility to individual n associated with lunch *i* combined with the treatment interaction can be specified as follows in Equation 3 (the base variables for the treatment interactions are omitted).	$U_{in} = V_{in} + \varepsilon_{in} = \beta_0 + \sum_{k=1}^{m} (\beta_k * A_k) + T_1 * \sum_{k=1}^{m-1} (\beta_{1k} * A_{1k}) + \varepsilon_{in}$ (3)
The student choice between two lunches shown in each scenario *s* is X_{ns}. Where the estimated coefficients β_k, represent the preference for each attribute level A_k. The sub-index 1 is associated with the treatment T, γ_n represents the observed and unobserved heterogeneity within individuals captured by the fixed effects attribute, and ε_s is the idiosyncratic term.	$[X_{ns}] = \sum_{k=1}^{m} (\beta_k * A_k) + T_1 * \sum_{k=1}^{m-1} (\beta_{1k} * A_{1k}) + \gamma_n + \varepsilon_s$ (4)

In the previously stated model, a positive sign in the coefficients indicates that an attribute has a positive effect on the take-up of a given lunch and higher values indicate the importance of the attribute. Hence, we expect significant coefficients for some attributes that could be used as restrictions in the model to solve the menu planning problem considering the specific preferences of this study's case.

Moreover, the interaction coefficients β_{1k} refer to the causal effect of receiving nutritional information on the choices for the treatment group. The reference group for T_1 is the control group, which did not receive any information. Hence, we expect significant coefficients on these interactions, which means, providing nutritional information modifies nutritional preferences and may therefore be relevant to students' healthy lunch choices.

1.3. Data collection

To examine the role of information on nutritional preferences, we used a one-time, online survey (see Annex 4). The first part of the survey provided information through a video regarding nutritional recommendations for the treatment group (see Annex 7), which mainly emphasizes the importance of a varied diet including all kinds of proteins, properties of vegetables and fruits and to avoid processed foods. Another section of the survey provides nine choice sets to elicit student's preferences for nutrition in lunch (see Figure 2). The last section of the survey collected anthropometric information and a question related to willingness to pay for variety, relevant input data needed in upcoming sections. The survey was sent via email by the school who agreed to participate and let us manage the responses with confidentiality conditions.

Figure 2. Example choice set shown to respondents. Note: Prices in Colombian Pesos.

Additionally, to assess that students fully understand the activity a survey pilot was carried out with students belonging to the same age group to which the experiment is directed (between 10 and 19 years old). The responders expressed that the video had an adequate length of time (3.6 minutes) and was clear. Likewise, they reported that they fully understood the instructions to answer the questions regarding the DCE.

Subsequently, the survey was sent to 946 students from the school. They were divided into random samples for the four variations of the survey, block 1 control and treatment, and block 2 control and treatment. To achieve 95% confidence given a population of 270 attenders to the school restaurant, 159 answers were required. Consequently, the study started on June 26th and ended on July 17th of 2020 where 169 students responded completely to the questionnaire, and this data is used to estimate the preferences of the school students.

2. Design, implement, and evaluate a mathematical model based on the PSO metaheuristic for cost reduction and satisfaction of nutritional needs, taking into account the nutritional preferences of the educational community

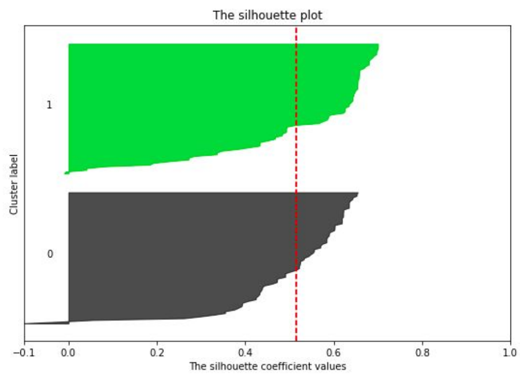
2.1. Nutritional requirements

For the collection and processing of the data, in the first place, a survey was sent to a representative sample of students, who were the same study subjects in the DCE of the previous paragraph. In this survey, the students answered questions about their basic anthropometric data, such as their height, age, weight, among others. By having a representative sample, the data was considered for the rest of the population when determining their nutritional requirements. In second place, in this project, by only taking into account lunch, the information in Resolution 3803 of 2016 had to be transformed proportionally to what the lunch represents nutritionally in the caloric intake of a day. According to the Colombian Institute of Family Welfare (ICBF, Instituto Colombiano de Bienestar Familiar), lunch should represent 35% of a day's caloric intake.

Lastly, to obtain reliable information for the food databases, the 2018 ICBF TCAC tables, among other official ICBF documents, were considered. These tables consider traditional preparations of the country and provide the information for each preparation in macro and micronutrients. The databases used in the application consider around 250 food preparations, which were extracted from these tables. The databases from the application also consider the price of each of these foods. These were consulted on official websites of potential suppliers of the school, such as Corabastos.

Due to differences of nutritional requirements between age groups, it is important to separate students to serve different portions according to their nutritional needs. To determine the number of groups and its characteristics we use clustering methods. The first method chosen is Agglomerative Hierarchical clustering, where every observation starts as its cluster and then are successively merged. The linked criteria used is ward, which determines the merge strategy, in this case, minimizes the sum of squared differences within all clusters. The variables used to make the cluster were the weight and age of the 153 students who answered this information in the survey. The clustering is made using SciPy, an open-source Python library (Hierarchical clustering — SciPy v1.5.3 Reference Guide, 2020).

Figure 3. Silhouette analysis for Hierarchical Agglomerative Clustering on sample data with two clusters.



To assess the number of clusters we use the silhouette analysis to study the separation distance between resulting clusters using scikit-learn, another open-source Python library (Pedregosa et al., 2011). This measure has a range of [-1, 1], where the coefficient near to positive 1 indicates that the sample is far from its neighboring clusters. The closest coefficient to 1 is generated with two clusters with a score of 0.52, the Figure 3 shows this result.

We also corroborate this result with scikit-learn using the k-means algorithm, which clusters data separating samples in n groups of equal variance and minimizing the sum of squares within the cluster. The selected value is again two clusters with a silhouette score of 0.52. The previous mentioned calculus and all graphs regarding clustering are shown in Annex 6.

Hence, using the first cluster results we take a mean for the age of 16.4 and 12.7 years old for groups 1 and 2 respectively, and a mean weight of 57.1 and 43.1 for groups 1 and 2 respectively. For the implementation of the project, we use group 1 as students age between 10 and 14 years old and for group 2 students between 15 and 19 years old. This means, there will exist two identical menu combinations with different portions which satisfy the nutritional needs of each particular group.

2.2. The mathematical model for linear optimization

In order to model this problem the first feature identified were the food categories that every lunch should be composed by, such as soup, rice, proteín, vegetable, grain, energétic, juice, salad, and dessert. This information is stored within the set K, which is defined from item 1 (soup) to item 9 (dessert). As a direct restriction from the school, every day should be displayed two different options for protein, salad, and juice. Due to this particular restriction, the different combinations of possible lunches among these options were proposed. Each of these combinations was stored in set A, which is defined from item 1 to item 8, which are shown in table 4.

Table 4. Lunch combinations

Lunch combinations	
Lunch 1	Protein 1, Salad 1, Juice 1
Lunch 2	Protein 2, Salad 1, Juice 1
Lunch 3	Protein 1, Salad 2, Juice 1
Lunch 4	Protein 1, Salad 1, Juice 2
Lunch 5	Protein 2, Salad 2, Juice 1
Lunch 6	Protein 2, Salad 1, Juice 2
Lunch 7	Protein 1, Salad 2, Juice 2
Lunch 8	Protein 2, Salad 2, Juice 2

Consequently, the information for the number of days the menu should be planned for was also defined among a set called J, which goes from 1 to 15 elements. Furthermore, the food database was also stored onto a set called I, which is defined from 1 to 248 elements, every single one of these elements belongs to a food category defined in the K set. Every element in the I set has its own nutritional information of macronutrients (defined in parameter $Macro_{im}$ per 1 gr of food), which is why set M was defined with these three elements: protein, carbohydrates, and fat content.

Since it is needed to know which $i \in I$ foods will appear in the lunch a∈A on day j∈J, a binary decision variable is defined as X_{iaj}. Y_{ij} is also defined as a binary decision variable that activates when the food $i \in I$ appears on the day j∈J, no matter what lunch combination it appears in, equation 5 shows the relationship between these two variables. This last variable helps to maintain control over the variety component, see equation 6.

$$\sum_{a \varepsilon A} X_{iaj} \leq Y_{ij} * \mu \quad \forall i \in I, \ \forall j \in J \qquad (5)$$

$$\sum_{i \in I} \sum_{l=j}^{j+4} Y_{ij} \leq 1 \quad \forall j \in J : j \leq |J| - 4 \qquad (6)$$

Every lunch will be composed of eight out of the nine K categories, with a specific portion for every category. The portion for the K category is a decision variable (Por_k) which has an upper bound (UB_k - parameter) and lower bound (LB_k - parameter), see equation 7 and 8. Even though this is a decision variable for the category, the given portion of food $i \in I$ that belongs (binary parameter σ_{ik}) to the category K, in the lunch a∈A on day j∈J ($PorA_{i,a,j}$) is restrained to be the same as Por_k when $\sigma_{ik} = 1$, see equation 9. $PorA_{i,a,j}$ also has a close relationship with X_{iaj} because $PorA_{i,a,j}$ can be greater than 1 when X_{iaj} is 1, see equation 10 and 11.

$$por_k \leq UB_k \qquad \forall k \in K \ (7)$$

$$por_k \geq LB_k \qquad \forall k \in K \ (8)$$

$$PorA_{iaj} * \sigma_{ik} \geq Por_k - \mu(1 - X_{iaj}) \quad \forall i \in I, \ \forall a \in A, \ \forall j \in J, \forall k \in K \ : \sigma_{ik} = 1 \qquad (9)$$

$$X_{iaj} \leq PorA_{iaj} \quad \forall i \varepsilon I, \ \forall a \in A, \quad \forall j \varepsilon J \qquad (10)$$

$$PorA_{iaj} \leq \mu * X_{iaj} \quad \forall i \varepsilon I, \ \forall a \in A, \ \forall j \varepsilon J \qquad (11)$$

In this order of ideas, every lunch has nutritional information conformed by the nutritional component of all the foods in the lunch. This information is stored in a decision variable called $Info_{maj}$, and is determined in equation 12.

$$Info_{maj} = \sum_{i \in I} Macro_{im} * PorA_{iaj} \qquad \forall a \in A, \ \forall j \in J, \forall m \in M \qquad (12)$$

Every lunch should have one food per category (see equation 13), except categories 4, 5, and 6, in which the school needs one to always be excluded (see equation 14 and 15). On one hand, subset E represents the categories that always have to be present in a lunch (1,2,3,7,8 and 9). On the other hand, subset F represents the categories that could not be present at lunch (4,5, and 6).

$$\sum_{i \in I} X_{iaj} * \sigma_{ik} = 1 \quad \forall a \in A, \forall j \in J, \ \forall k \in E \qquad (13)$$

$$\sum_{i \in I} X_{iaj} * \sigma_{ik} \leq 1 \quad \forall a \in A \ \forall j \in J, \ \forall k \in F \qquad (14)$$

$$\sum_{k=4}^{6}\sum_{i\in I} X_{iaj} * \sigma_{ik} = 2 \quad \forall j \in J, \quad \forall a \in A \qquad (15)$$

Another requirement for the lunch composition is that some categories should be the same food in every lunch of the same day, see equation 16. These categories are stored subset Shared and are 1,2,4,5,6, and 9. Consequently, the rest of the categories are explained in equations 17 through 22 to accomplish the lunch combinations shown in table 4.

$$X_{iaj} * \sigma_{ik} = X_{ibj} * \sigma_{ik} \quad \forall i \in I,\ \forall j \in J,\ \forall a \in A,\ \ \forall b \in A,\ \forall k \in Shared \qquad (16)$$
$$X_{i1j} * \sigma_{i3} = X_{iaj} * \sigma_{i3} \quad \forall i \in I,\ \forall j \in J,\ \forall a \in A : a = 3,4,7 \qquad (17)$$
$$X_{i2j} * \sigma_{i3} = X_{iaj} * \sigma_{i3} \quad \forall i \in I,\ \forall j \in J,\ \forall a \in A : a = 5,6,8 \qquad (18)$$
$$X_{i1j} * \sigma_{i7} = X_{iaj} * \sigma_{i7} \quad \forall i \in I,\ \forall j \in J,\ \forall a \in A : a = 2,4,6 \qquad (19)$$
$$X_{i3j} * \sigma_{i7} = X_{iaj} * \sigma_{i7} \quad \forall i \in I,\ \forall j \in J,\ \forall a \in A : a = 5,7,8 \qquad (20)$$
$$X_{i1j} * \sigma_{i8} = X_{iaj} * \sigma_{i8} \quad \forall i \in I,\ \forall j \in J,\ \forall a \in A : a = 2,3,5 \qquad (21)$$
$$X_{i4j} * \sigma_{i8} = X_{iaj} * \sigma_{i8} \quad \forall i \in I,\ \forall j \in J,\ \forall a \in A : a = 6,7,8 \qquad (22)$$

Lastly, the two different options of protein on the same day should be from a different animal which is why several subsets about the animal the proteins came from were defined. These subsets are chicken (b), beef (r), fish (p), pork (c), turkey (pav) and vegetable (v) , see equation 23 through 28.

$$X_{i,1,j} + X_{i,2,j} \leq 1 \quad \forall j \in J,\ \forall i \in I(b) \qquad (23)$$
$$X_{i,1,j} + X_{i,2,j} \leq 1 \quad \forall j \in J,\ \forall i \in I(r) \qquad (24)$$
$$X_{i,1,j} + X_{i,2,j} \leq 1 \quad \forall j \in J,\ \forall i \in I(p) \qquad (25)$$
$$X_{i,1,j} + X_{i,2,j} \leq 1 \quad \forall j \in J,\ \forall i \in I(c) \qquad (26)$$
$$X_{i,1,j} + X_{i,2,j} \leq 1 \quad \forall j \in J,\ \forall i \in I(pav) \qquad (27)$$
$$X_{i,1,j} + X_{i,2,j} \leq 1 \quad \forall j \in J,\ \forall i \in I(v) \qquad (28)$$

Every element of the I set has nutritional information about macronutrients but also has information about their own cost. In order to store this information, the parameter P_i was defined as the Price for the food $i \in I$ for 1gr. The cost incurred for lunch a∈A on day j∈J is stored in a decision variable called $PrecioA_{a,j}$ is determined in equation 29.

$$Precio_{aj} = \sum_{i \in I} P_i * PorA_{iaj} \quad \forall a \in A,\ \forall j \in J \qquad (29)$$

The students for whom the menu has been planned have nutritional requirements in terms of macronutrients at lunch, which are stored under parameter Req_m. Every lunch aims to compliance this requirement, however, since this requirement is a number and not a range, is very unlikely that $Info_{maj}$ take the exact value of Req_m, which is why equation 30 is made for slacks. On one hand, $Hmenos_{maj}$ is a decision variable that represents the deviation below the requirement. On the other hand, $Hmas_{maj}$ is a decision variable that represents the deviation above the requirement.

$$Info_{maj} - Req_m + Hmenos_{maj} - Hmas_{maj} = 0 \quad \forall j \in J,\ \forall a \in A,\ \forall m \in M \qquad (30)$$

In the same way, the school has a WTP (*Willingness to pay),* which is a parameter for the model. Equation 31 shows the same restriction as equation 30 above but in terms of costs. Decision variables for money deviation slacks are also defined as $HPmenos_{aj}$ and $HPmas_{aj}$.

$$Precio_{aj} - WTP + HPmenos_{aj} - HPmas_{aj} = 0 \quad \forall j \in J,\ \forall a \in A \qquad (31)$$

Since the objective function is a weighted multi-objective, the objectives have to be in standard terms in order to be added. Due to this nature, all the parameters and decision variables that have a direct relationship with the nutritional or cost objective should be scaled into a standard value, between 0 to 1. In order to do this, the method feature scaling is used. For instance, equation 32 shows the expression used to find the scaled value for the nutritional requirement ($ReqS_m$), in the same way equation 33 does the same but to find the scaled value for the Willing to pay (WTPS). Both of these equations use parameters for their possible minimum and maximum values.

$$ReqS_m = \frac{(Req_m) - MinV_m}{MaxV_m - MinV_m} \quad \forall m \in M \qquad (32)$$

Where:

$MinV_m$: minimum possible value for the macronutrient $m \in M$ in a single lunch

$MaxV_m$: maximum possible value for the macronutrient $m \in M$ in a single lunch

$$WTPS = \frac{(WTP) - MinVP}{MaxVP - MinVP} \qquad (33)$$

Where:

$MinVP$: Minimum cost possible for one lunch

$MaxVP$: Maximum cost possible for one lunch

Feature scaling method is also used to scale variable decisions, such as the nutritional deviation, both above ($PromSMas_{maj}$, see equation 34) and below ($PromSMenos_{maj}$, see equation 35) to the requirement, per macronutrient $m \in M$, in the lunch a∈A on day j∈J. In the same way, is scaled the deviation for the cost, both above ($PromPSmas_{aj}$, see equation 36) and below ($PromSMenos_{maj}$, see equation 37) to the WTP, in lunch a∈A on day j∈J.

$$PromSMas_{maj} = \frac{(Req_m + Hmas_{maj}) - MinV_m}{MaxV_m - MinV_m} \quad \forall m \in M, \quad \forall j \in J, \forall a \in A \quad (34)$$

$$PromSMenos_{maj} = \frac{(Req_m - Hmenos_{maj}) - MinV_m}{MaxV_m - MinV_m} \quad \forall m \in M, \quad \forall j \in J, \ \forall a \in A \quad (35)$$

$$PromPSmas_{aj} = \frac{(WTP + HPmas_{aj}) - MinVP}{MaxVP - MinVP} \quad \forall j \in J, \ \ \forall a \in A \quad (36)$$

$$PromPSmenos_{aj} = \frac{(WTP - HPmenos_{aj}) - MinVP}{MaxVP - MinVP} \quad \forall j \in J, \ \ \forall a \in A \quad (37)$$

In the same way equation 30 and 31 were defined, its scaled dual is also defined, both for the nutritional slack (see equation 38) and the cost slack (see equation 39).

$$DSMenos_{maj} + PromSMenos_{maj} = ReqS_m \quad \forall j \varepsilon J, \ \forall m \varepsilon M, \ \forall a \varepsilon A \quad (38)$$

$$PromSMas_{maj} - DSMas_{maj} = ReqS_m \quad \forall j \varepsilon J, \ \forall m \varepsilon M, \ \forall a \varepsilon A \quad (39)$$

Since a deviation from the WTP can only be above or below the WTP, equations 40 through 43 are defined, in order to ensure only one decision variable is activated. Equation 40 and 41 does this for the normal slack, and equation 42 and 43 for the scaled slack. Both $BinWTP_{aj}$ and $Bin2_{aj}$ are defined as binary decision variables.

$$HPMas_{aj} \leq BinWTP_{aj} * \mu \quad \forall a \varepsilon A, \ \ \forall j \varepsilon J \qquad (40)$$

$$HPMenos_{aj} \leq (1 - BinWTP_{aj}) * \mu \quad \forall a \varepsilon A, \ \ \forall j \varepsilon J \qquad (41)$$

$$DSWTPMas_{aj} \leq Bin2_{aj} * \mu \quad \forall a \varepsilon A, \ \ \forall j \varepsilon J \qquad (42)$$

$$DSWTPMenos_{aj} \leq (1 - Bin2_{aj}) * \mu \quad \forall a \varepsilon A, \ \ \forall j \varepsilon J \qquad (43)$$

Finally, scaled decisions variables are defined in terms of the day, without taking into consideration a specific lunch, or if its deviation is below or above the requirement. In order to do so, the scaled deviation variable from equation 38 ($DSmenos_{maj}$) and equation 39 ($DSmas_{maj}$) is added in all of its lunch combinations, see equation 44 for the nutritional deviation, and equation 45 for the cost deviation.

$$DStotal_{mj} = \tfrac{1}{2} * \sum_{a \in A} \tfrac{1}{|A|} * DSmenos_{maj} + \tfrac{1}{2} * \sum_{a \in A} \tfrac{1}{|A|} * DSmas_{maj} \quad \forall m \in M, \ \forall j \in J \qquad (44)$$

$$DSWTPtotal_j = \tfrac{1}{2} * \sum_{a \in A} \tfrac{1}{|A|} DSWTPmas_{aj} - \tfrac{1}{2} * \sum_{a \in A} \tfrac{1}{|A|} DSWTPmenos_{aj} \qquad \forall j \in J \qquad (45)$$

In order to propose a multi-objective function of the model, we use a weighted goal programming function, where weights are assigned to the deviation of each objective from its respective goal, two goals were considered. First, to minimize the total deviation of each macronutrient compared to the recommended value according to the Resolution 3803 of 2016 and, second, minimize the total deviation of the cost above the WTP (this is the same as indirectly maximizing the total deviation of the cost under the WTP). Willingness to pay (WTP), is defined as the maximum average cost per lunch the school will accept to pay to implement the menu. That is, the average cost the school is currently paying in food inputs to make lunches. Additionally, $W1$ and $W2$ are the constants representing the importance of the corresponding components of the fitness function, where $W1 + W2 = 1$. Methods to determine the weighting values are not implemented because W1 and W2 are defined by the school which establishes a value of 0.6 and 0.4 respectively.

Hence, the multi-objective function is represented as follows:

$$min\ z = W1 * \sum_{m\varepsilon M} \sum_{j\varepsilon J} \frac{1}{|M|} DStotal_{mj} + W2 * \sum_{j\varepsilon J} DSWTPtotal_{j} \quad (46)$$

2.3. PSO Design

For the design of the particle, the number of days for which it is required to set up a menu is defined, this value can be any integer between 1 to 15 days. Additionally, the number of possible combinations of lunch alternatives (8) and the food categories that make up a lunch (9) are considered. In each of these categories, for each of the 8 possible lunch combinations, for each of the n days, the position and velocity information is available. Moreover, in table 5 the pseudocode for the PSO metaheuristic is shown. The velocity and position are calculated in each of the food categories that make up a lunch, only on the day with the worst performance of each particle.

Table 5. Velocity and position functions

```
SUBPROCESS PSO
    Initialization();

    FOR particle = 1 TO n DO
        CalculateFF();
    END FOR

    WHILE (Iterations <= MaxIterations AND Time <= Tmax) DO
          FOR particle = 1 TO n DO
              FOR day = 1 TO k DO
                  IF WorstDay = TRUE THEN
                      CalculateVelocity();
                      CalculatePosition();
                  END IF
                  CategoryExclusion();
                  AdjustPortion();
              END FOR

              CalculateFF();
              CalculatePersonal Best();
              CalculateGlobal Best();
          END FOR
          Time = Time + timer();
          Iterations = Iterations + 1;
    END WHILE
END SUBPROCESS
```

First of all, the velocity function used for the PSO is shown on equation 47.The values of C1 and C2 are fixed at 0.5 for the calculation. Second of all, equation 48 presents the discretization of the PSO, by only considering the integer part of the velocity result. Lastly, equation 49 shows the expression used to find the position for each particle.

velocity = w*vAnt + (C1*rand1 * (Pbest-position)) + (C2*rand2 * (gBest-position)) (47)

veloc = Int(Velocity) (48)

Position = Position + veloc (49)

Where:

- vAnt represents the previous velocity for the particle
- Pbest represents the personal best for the particle
- gBest represents the global best between all particles

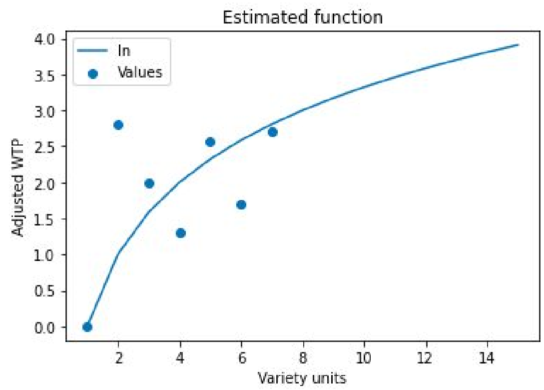
2.3.1. Objective function

The quality of a particle is assessed with a fitness function, FF, consisting of three components. The first two components are described above in Equation (46) regarding the nutritional and the cost component. Besides, the variety component has been introduced as the third component in Equation 50 below, to give an estimate of how diverse the solution is. This component has a range of [0, 1], where 1 means diverse and 0 means not diverse at all, thus it makes the FF to have more variety.

$$min\ z = W1 * \sum_{m \varepsilon M} \sum_{j \varepsilon J} \frac{1}{!M!} DStotal_{mj} + W2 * \sum_{j \varepsilon J} DSWTPtotal_{j} - \sum_{k \varepsilon K} \frac{1}{!K!} VariedadSP_{k} \quad (50)$$

The variety component is not weighted by any constant because is designed as a reward for the FF, also, internalize diminishing marginal returns exposed by a logarithmic function, which is the selected utility function to represent the data's behavior (Figure 4) and hence, represents the student preferences for variety taking into account different rewards for the FF for every marginal increment of variety in the solution. The mentioned estimated function is obtained from the question included in the previously explained questionnaire, where students reported the WTP for every additional unit of variety in a 7 day hypothetical scenario where it was possible to pay to eat different lunches.

Figure 4. Estimated function for the adjusted WTP of variety units



2.4. Restrictions

Returning to the methodology specified above on the DCE but this time taking the parameters of the entire sample on average. Student's preferences are estimated through the following specified model to generate the corresponding restriction in the algorithm.

$$[X_{sn}] = \sum_{k=1}^{m} (\beta_k * A_k) + \gamma_n + \varepsilon_s \quad (51)$$

A conditional (fixed-effects) logistic regression is made (Analysis et al., 2010) where the reference group for each attribute corresponds to a vegetable protein, no soup, three varieties of carbohydrates, hot salad, and non-homemade dessert. We estimate the attribute for 'price' as a continuous variable to calculate the willingness to pay. The results are specified in Annex 7 where the coefficients of the fish and meat levels for the attribute are significant, as well as, soup and cream for the soup attribute, one kind of carbohydrates, and cold salad. All the mentioned coefficients are significant at 99% confidence. Nonetheless, due to restriction of the number of foods available in the data base for soups and carbohydrates, and the allowed number of salads stipulated for the school, we use only the found significant estimation related to the levels of the attribute protein and is then used as a restriction.

To build the restriction, the WTP is estimated through the *wtp* command in *Stata* which implements the method Krinsky Robb. The mentioned estimations for the significant level-attributes are shown in Annex 7. It is found that responders are willing to sacrifice 4,669 and 3,747 COP to have fish and white meat respectively, instead of vegetable protein. Thus, the proportion between these two components is used to build the restriction which specifies that the times fish is served has to be 25% greater than the times chicken is served in the menu.

Additionally, both the school and the country have traditional dishes that do not meet the 9 food categories proposed for the project. However, these are very important dishes in Colombian culture. This project considers using special dishes in the menu output, subject to only one special dish per solution. The choice of this plate is not considered within the metaheuristic but it is considered within the algorithm. Although the choice of this dish does not depend on the position and velocity of a population of particles, it does take into account its impact on the FF in order to choose the dish that best suits the menu chosen by the PSO.

3. **Design an application for easier data input, execution, and display of results of the model based on the PSO metaheuristic.**

We consider Microsoft Excel tool as the main tool for the design and programming of the application, as well as its VBA (Visual Basic for Application) programming component for the development of metaheuristics. The choice of this tool is motivated by its compatibility since it is considered one of the most used spreadsheet programs and rarely encounters problems with sending a file, in addition to being easy to access for all users. On the other hand, it is a tool known and understood by all team members. Finally, Microsoft Excel has an effective and friendly interface that allows it to carry out activities in a simple and practical way.

Furthermore, for the final design of the application we rely on the update of the ISO 9126 standard for the application of design standards and good programming practices. This update corresponds to ISO / IEC 25010: 2011, currently in force. ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission), specialized systems for worldwide standardization, together with other international organizations, governmental and non-governmental, incoordination, participate in the work of the ISO / IEC 25010: 2011.

A high-quality application is essential to provide value and avoid potential negative consequences for stakeholders This can be achieved by defining the necessary and desired quality characteristics associated with the goals and objectives of the administrators of the school, who will make use of it. In addition, this includes quality characteristics related to the software and data system, as well as the impact it generates.

This international standard defines:

A. "A quality in use model composed of five characteristics that are related to the result of the interaction when a product is used in a particular context of use"(ISO/IEC, 2011).
B. “A product quality model composed of eight characteristics that relate to the static properties of the software and the dynamic properties of the computer system” (ISO/IEC, 2011).

The quality models in this International Standard can be used to identify relevant quality characteristics that can later be used to establish requirements, the satisfaction criteria, and corresponding measures. Furthermore, the characteristics and sub-characteristics provide consistent terminology to specify, measure and evaluate the quality of the application. Likewise, the established quality requirements can be compared to verify their integrity.

In this case, it refers to the application that generates a menu for a periodicity of 15 days, and that meets the nutritional characteristics at the lowest possible price, including the preference of the users. The above encompasses a human-computer relationship as it will be used by those in charge of generating this menu of the school.

3.1 Development of the application:

In order to create a system that conforms to the characteristics and sub-characteristics of the ISO / IEC 25010: 2011 standard, as well as the needs and objectives of the school, we developed an application where a user can obtain a menu of lunches for a period of 15 days, meeting the nutritional requirements at the lowest possible cost, and taking into account the preferences of the users. This is achieved by the development of a PSO metaheuristic, which has the following input parameters

- A database that includes a large part of the foods (cooked) and nutritional information (energy (KCal), protein (g), lipids (g) and carbohydrates) from the Colombian Food Composition Table (TCAC) 2018 (TCAC, Tabla de Composición de Alimentos Colombianos, 2018). Additionally, each food is associated with a price both raw and cooked; this last piece of information is essential to be able to calculate the objective function of the algorithm, since as it is a multi-objective function, the price and the nutritional deviation must be in the same units. Therefore, to know the cooked price of the food we use a conversion factor to change the price from raw to cooked.

Furthermore, each food belongs to a specific category, where an option number is assigned according to the amount of elements that exist in it. Food can be classified into a category as follows: Category 1: Soup, Category 2: Rice, Category 3: Protein, Category 4: Vegetable, Category 5: Grain, Category 6: Energetic, Category 7: Salad, Category 8: Pulp (Juice) and Category 9: Dessert. Also, there is Category 10: Specials,this category is specially designed for those foods that are considered special since they contain several categories (from 2 to 7) in it. This is because they are foods with a large amount of nutritional components and they meet the requirements of these categories, these foods can refer to hamburgers, ajiaco, lasagna, etc.

Also, an additional element that a food may or may not have is a classification that only applies to the categories "Soup" and "Protein", since these are the categories in which one element predominates over another according to the preference of users. Thus, for the category of soup, foods can be classified depending on whether it is soup (ID = 1) or cream of soup (ID = 2). On the other hand, protein is classified as follows: chicken (ID = 1), meat (ID = 2), fish (ID = 3), pork (ID = 4), goat (ID = 5), turkey (ID = 6) or vegetable (ID = 7).

- Nutritional parameters such as:
  - Calories per kilogram (per day): it is the amount of calories per weight in kg that a user requires per day. This requirement is associated with both the age of the students and the type of physical activity they perform.
  - Percentage of calories per lunch: it is the percentage of calories that a lunch should assume per day and that a child should consume, taking into account the other meals of the day such as breakfast, dinner and ounces.
  - Total fat percentage: it is the percentage of fat (g) that a lunch should have to meet the nutritional requirements.
  - Total carbohydrate percentage: it is the percentage of carbohydrate (g) that a lunch should have to meet the nutritional requirements.
  - Protein percentage: is the percentage of protein (g) that a lunch should have to meet nutritional requirements.

- Exclusive parameters of the PSO:
  - Maximum time (seconds): is the maximum number of seconds in which you want to obtain a solution.
  - Population size and number of iterations: number of particles and iterations that will be involved in the algorithm.
  - Number of days: is the number of days for which the user wants to generate a solution.
  - Nutritional part weighting: since the objective function of the metaheuristic is a multi-objective function, weights are given to the nutritional part and the price part.

These input data can be modified as required by the user and in order to do so, the application requests an input password, so only the indicated people could make these modifications. Additionally, the application allows the user to add as many foods as they want, taking into account the information associated with it, also respecting the logic of the algorithm.

Finally, in order to present a friendly and easy-to-use interface for the users of the school, the application has 5 spreadsheets to present all the necessary information: Instructions, Menu, Base, Parameters, Requirements and Results. The purpose of spreadsheet is as follows:

- **Instructions:** proposes to guide the user in the use of the application, presenting the information contained in each spreadsheet along with its purpose and presenting the names of documents that contain the nutritional references that we took into account in the project
- **Menu:** The menu for 15 days is presented, specifying the food and the portion to be served for each group. Additional shows the average oh the price, calories, grams of protein, grams of carbohydrate and grams of fat.
- **Parameters:** The nutritional parameters are presented, which are stipulated according to the recommendation indicated in resolution 3803 of 2016 of the Ministry of Health. These parameters are the same metaheuristic input data.
- **Requirements:** Representative anthropometric data collected through surveys are presented. Relating physical activity to the recommended caloric requirement.
- **Results:** The nutritional results and the average price of the menu by group are presented. Additionally, the basic foods and the quantities necessary to implement the 15-day menu are specified. The approximate costs of the same are summarized.

4. Contrast the cost and nutritional efficacy of the current state of the educational institution, with the results obtained by the model

The school presents an initial lunch solution for a period of 15 days where the food is presented according to its category and a portion of grams pre-established per food for every day and for all students, regardless of their weight, age and physical activity. Additionally, the school provides us with a list of supplies that must be purchased per month with their respective cost in Colombian pesos. In this way, with the information provided and in order to determine the nutritional and cost efficacy that the school presents as an initial solution, this solution is compared with the one provided by the PSO metaheuristic.

Therefore, through the metaheuristic we intend to reduce the deviation from the nutritional requirements necessary for a child's lunch, taking into account their age, weight and type of physical activity. Also, we intend to generate this solution at the lowest possible cost and taking into account the preference of users and a variety in the 15-day menu. Furthermore, to determine the nutritional and cost efficacy of the proposed solution, different performance tests are carried out to ensure that the design achieves the performance requirements. Then, these performance tests include variation in input data, comparison with a basic mathematical model, and comparison with the initial solution.

4.1 Comparison of both solutions using an impact evaluation through before-and-after evaluation:

It is important to analyze the nutritional effect and the price effect caused by the implementation of the algorithm. This analysis can be done through impact evaluations, a particular type of evaluation, which purpose is to answer questions of cause and effect. "Unlike general evaluations, which can answer many types of questions, impact evaluations are structured around one particular type of question: What is the impact (or causal effect) of a program on an outcome of interest?" Gertler et al., 2011). To answer the previous question, it is necessary to establish what would have happened to the students without the program.

Based on the above, the main role of impact evaluation is to produce evidence on the effectiveness of the application for school users who will have associated benefits thanks to this. Furthermore, this can be achieved through a chain. which "sets out a logical, plausible outline of how a sequence of inputs, activities, and outputs for which a project is directly responsible interacts with behavior to establish pathways through which impacts are achieved" (Gertler et al., 2011). Then, the steps of the chain are: 1. Inputs 2. Activities 3. Outputs 4. Outcomes 5. Final outcomes.

The scope of the project considers only the implementation of the application on links in the chain that consider the supply, that is, the school administrators, but does not include a measurement of the results on the students after implementing it (that is, the demand side), that is, until the step of "Output (3)". In this way, the inputs refer to a lunch minute established for 15 days that the school gives us mainly and that is currently implemented in it, we also take into account the price list associated for each of the foods in this minute. Then, as a change activity, we propose the application based on a PSO metaheuristic that generates a menu for 15 days, reducing the deviation from nutritional requirements at the lowest possible cost, taking into account the preference of users.

Finally, in order to evaluate the outputs that the application generates, we take into account a before and after evaluation. Therefore, this study "measures outcomes in a group of participants before introducing a product or other intervention, and then again afterward. Any change in the results is attributed to the product or the intervention " (England, 2020). Then, in this study we intend to understand the impact generated by the solution of the metaheuristic compared to the solution proposed by the school, impact on nutritional compliance as well as on the price generated by lunch.

5. Engineering design component

Design statement

Design a DCE that estimates the food preferences faced by the educational community of the Colegio Mayor de San Bartolomé, evaluating the effect of providing nutritional information.

Design a PSO metaheuristic that finds a solution that reduces food-related costs and contributes to meeting nutritional needs, considering the dilemmas in food preferences faced by the educational community.

Design a tool that, using the aforementioned metaheuristics, determines the food inputs necessary to meet the nutritional requirements of the children of the school, the recipes for each day of the week, and the incurring costs.

Design process

In figures 5 to 7 you can find the flow diagram of the program. As can be seen, the algorithm takes into account more different functions than the PSO metaheuristic. For example, the choice of preferences and special foods are not taken into account within the metaheuristics. Instead of this, the PSO metaheuristic chooses a menu within the 9 proposed categories, and with the best PSO particle, modifications are made to include the preferences in every day of this particle, and include the special food on the day with the worst performance of the FF. These changes are made taking into account that it is the best possible change according to the FF.

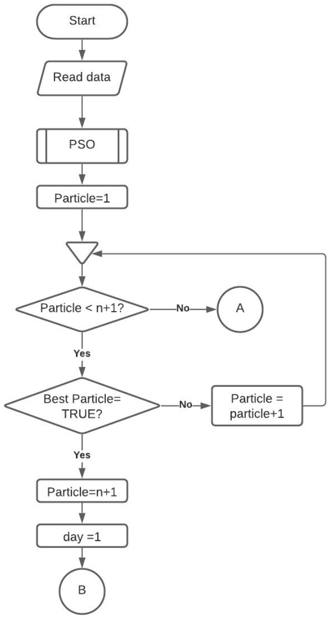


Figure 5. Part 1 of the flow diagram

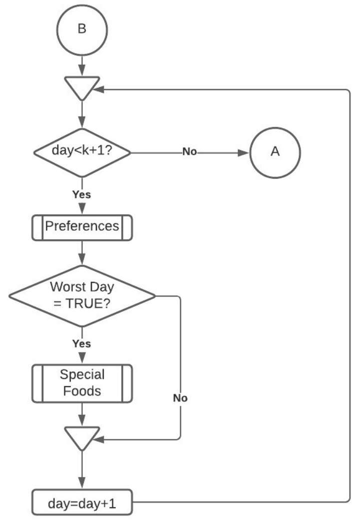


Figure 6. Part 2 of the flow diagram

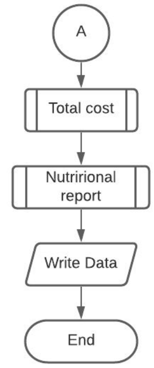


Figure 7. Part 3 of the flow diagram

Performance Requirement

The output file of the application is a list of food inputs with unit prices, total prices, recipes to cook each day, the portions to be served by cluster group, and the specification of the nutritional content of the menu, all this for a period of 15 days. The solution also ensures that each person will have two proteins, juice and salad options each day. On the other hand, the metaheuristic is expected to run in 30 minutes or less. This means that the solution suggested by the tool would be the best found in those 30 minutes.

Performance Tests

- **Mathematical model**

Figures 8 and 9 show an exponential trend in the model's run time as the number of elements in some sets increases. In both figures, the vertical axis shows the time in seconds that the running model lasted. Figure 8 shows the analysis of what affects increasing a lunch for a single day. On the other hand, figure 9 shows the analysis of what affects increasing one day to lunch. It should be noted that running the model for 15 days and 8 lunches was not possible due to the necessary execution time. The tests that had a time of less than 3,600 seconds were run 10 times, and the average of these times can be found in the graph. The tests with a running time greater than 3,600 seconds were run 3 times.

Figure 8. Time when increasing lunch combinations, with one day

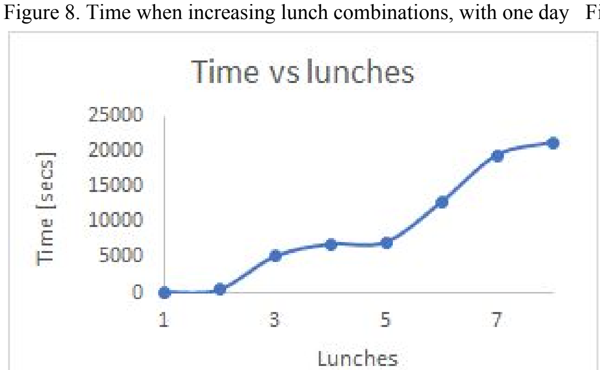
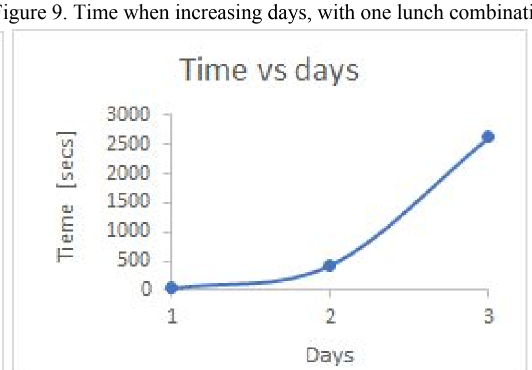


Figure 9. Time when increasing days, with one lunch combination



These tests were run with the weights of W1 = 0.6 (nutritional compliance) and W2 = 0.4 (price minimization). Table 6 shows the results for a day and a lunch. However, in the search for better nutritional compliance, especially in the protein macronutrient, a weight of W1 = 0.7 and W2 = 0.3 was also tested. In these last conditions the price increased dramatically, but the nutritional compliance in the 3 macronutrients was very close to zero. Due to 4,795 COP is minor than WTP, and because macronutrient compliance improved significantly, use this weight ratio is recommended rather than W1 = 0.6 and W2 = 0.4.

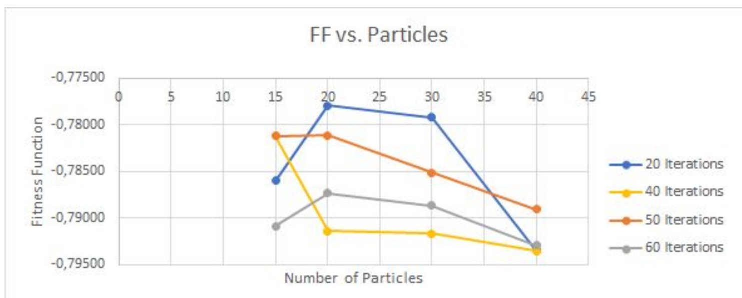
Table 6. Comparison between two weights relationships in the mathematical model

Weights relationship	Objective Function	Proteins deviation	Carbohydrates deviation	Fats deviation	Average cost
60 / 40	-0.006903776	60.9344	0	1.16418	2,558.38 COP
70 / 30	-0.003743046	11.0566	0	0	4,795.44 COP

- **PSO**

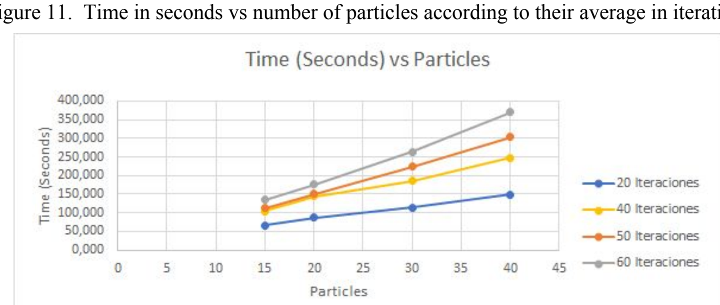
The purpose of testing the algorithm (see Annex 8) is to check how the individuals in the population react to each other, to see if they improve their solution, the Personal Best correction of the best particle, the running time of the algorithm, the best population size setting and the best number of iterations. Therefore, to analyze the best configuration of the population size and the number of iterations, the values of both were varied and several runs were made to test and average the randomness and the improbable behavior. In this way, in figure 10 we present how the population size and the number of iterations affect the fitness value for the case of a menu for 15 days. To analyze the behavior of PSO, the population sizes considered were 15, 20, 30 and 40, with 20, 40, 50 and 60 iterations. Also,for each of these combinations, a number of 10 tests have been performed.

Figure 10. Fitness function vs number of particles according to their average in iterations



In Figure 10, a trend can be seen in all the iterations that increasing the population size generally improves the adequacy of the solutions generated, except for the smallest population size considered (15 particles), when it increases to a bigger size (20 particles), this is because there can be a high component of randomness with few particles. Furthermore, it can be observed that all the iterations tend to coincide their best version of the fitness function with the maximum size of the population considered (in this case 40). On the other hand, it would be logical to think that increasing the number of iterations generally improves the quality of the solutions, however, there is still a level of randomness and indeterminacy present, as for the tests with 50 and 60 iterations, the averages of the solutions were not always the best. The best average of the graph has been found with a population of 40 and a number of 40 iterations.

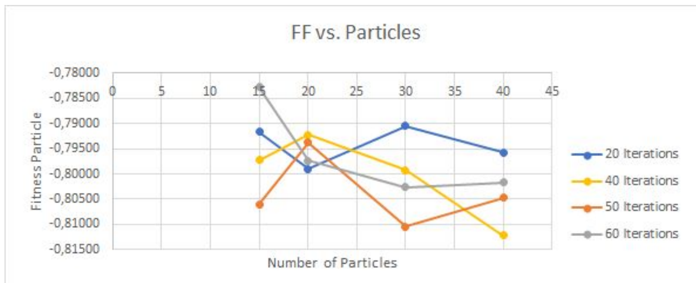
Figure 11. Time in seconds vs number of particles according to their average in iterations



From Figure 11, it is noted that all the execution time values are approximately linearly variable with the population size, and that increasing the number of iterations always increases the execution time. It can be seen that the minimum time to generate a solution is 65,526 seconds (1,092 minutes) with a configuration of 15 particles and 20 iterations (FF: -0,786) and the maximum time to generate a solution is 370,324 seconds (6,172 minutes) with a configuration of 40 particles and 60 iterations (FF: -0.790). For the configuration of 40 particles and 40 iterations, which is the one with the best fitness function (-0.794), it takes a time of 248.331 seconds (4.14 minutes).

In the previous tests, only the scenarios that produced a PSO solution were taken into account, without taking into account the preferences of the users and the implementation of a special meal in one day per 15-day solution. Below we present the scenarios analyzed above, including a variation in the objective function, since it is included the preferences of the users towards fish over chicken, in addition to including a special meal in one day of the entire solution:

Figure 12. Fitness function vs number of particles according to their average in iterations including preferences and special foods



Comparing these scenarios with those that do not contain special foods or user preferences, it can be seen that in Figure 12 there is no trend in the iterations. This is due to the fact that including the preferences of the users and the special food there is a change in the value of the objective function, altering what was reported under the single PSO scenario. However, something to note is that the best average of the graph has been found with a population of 40 and a number of 40 iterations, as in the scenario above. Thus, taking this combination of number of particles and iterations, a good solution will be found in both cases.

In addition to the algorithm's own tests, anthropometric tests were taken into account, that is, varying the anthropometric input parameters to verify that the algorithm was measurable and achievable from different perspectives. In this way, 3 different scenarios were taken into account, where the age range of the children to be analyzed varied: 10-12 years, 13-15 years and 16-18 years. According to these scenarios, the amount of calories required per day also changed depending on the average weight of the sample. Then, for these tests, a population size of 40 individuals, 40 iterations and 10 tests were taken into account for each scenario.

Table 7. Menu solution data according to an age range and its caloric requirements

Age	Time (Seconds)	FF	Protein Deviation	Carbohydrate Deviation	Fat Deviation	Cost
10-12	274	-0.837	10.03	4.18	-3.51	4,473
13-15	276	-0.835	9.96	3.94	-2.92	4,400
16-18	272	-0.821	9.36	3.72	-4.82	4,486

According to Table 7, similar values can be evidenced in all scenarios both for deviations from nutritional requirements and for cost. Additionally, the time it takes to execute the algorithm was quite similar in all scenarios, on average it takes 274 seconds (4.56 minutes). On the other hand, the lowest cost of lunch is given for students who are between 13 and 15 years old, while the best objective function is given for the scenario in which students are between 10 and 12 years old. In conclusion, the algorithm will behave in the same way, regardless of the input data provided, that is, it will take into account a minimization of the deviation in terms of nutritional requirements at the lowest possible cost.

A test is carried out in which the PSO is executed 15 times just for one lunch. The joint solution exhibits good performance in variety, given the natural randomness of the heuristic, however, as it is not a solution that contemplates the 15 days as a whole, the preference restriction is violated. Additionally, it has a greater nutritional deviation in carbohydrates and proteins, in fat it has a lower deviation. As for the price, this solution is approximately 2,000 COP on average above the PSO solution. This test highlights the importance of contemplating the 15 days simultaneously in the PSO.

- **Comparison of Optimization Model vs PSO**

Table 8 shows the disaggregated results for both PSO metaheuristic and the mathematical model. The information on the macronutrients is given as the deviation of grams from the requirement, which is why these three results will have a better performance the closer to zero their values are. On the other hand, the Objective function is given in deviation units and the average price is the average price of lunches in COP, Colombian Pesos. This table also analyzes two possible scenarios, the first is a scenario where only the best lunch is analyzed, so only values are obtained for a single lunch in a single day. In the second scenario, 8 possible combinations of lunches in a single day are analyzed. It should be noted that it was not possible to run scenarios with more variables due to the NP-hard characteristic of this problem.

Table 8. Comparison between PSO metaheuristics and mathematical model

		Proteins deviation	Carbohydrates deviation	Fats deviation	Average cost
One day and one lunch	**PSO**	22	25	-8	6,853.19 COP
	MATHEMATICAL MODEL	60.93	0	1.16	2,558.38 COP
One day and eight lunches	**PSO**	16	19	-2	7,841.31 COP
	MATHEMATICAL MODEL	61.05	1.62	4.33	2,979.37 COP

Restrictions

The main restrictions of the program are mentioned below:

- To perform the DCE, a sample of at least 55 people is required for a confidence level of 90%. However, the sample was expanded to 169 people to reach a confidence level of 95%.
- It is necessary to know the respective price of the food supplies. The application already has a database of more than 250 food supplies. However, if the user wants to have more options will need to add each food, with its price and nutritional information. It is also needed to update the food prices everytime the provider changes them.
- This problem assumes that all food supplies are always available.
- It is necessary to know age, weight and other variables of a representative sample of students for the well functioning of the algorithm.
- In this design, the periodicity of purchase will not be taken into account.
- In this design, two different protein, juice and salad options will be proposed for each day for fifteen days, that is, for three calendar weeks.
- This application considers meeting the nutritional requirement of macronutrients, and does not take into account information or micronutrient requirements at any time.
- Each student will belong to a diet profile that will determine the portions they need for each lunch component. This profile depends on different factors, such as: weight, gender, level of physical activity, age, among others. This diet group was determined with the help of hierarchy clustering.
- The portion servings of a food category in the same diet group will be the same for every day.
- This application is designed to be used in a non-personalized way for each student, therefore, it is not possible to include personal preferences. Thus, it is impossible to include in the program, for example, the specific tastes of a child regarding certain vegetables, since the preferences to be considered correspond to the general population.
- The algorithm chooses meals for each of the nine proposed categories, which are: soup, rice, protein, grain, vegetable, energetic, salad, juice and dessert.
- The two protein options presented on the same day must be from a different animal. For example, if the first protein option is chicken nuggets, the second option cannot come from a chicken. It must be a preparation of pork, beef, fish, turkey, etc.
- Among the categories of vegetable, grain and energetic, one should always be excluded. The one that least affects the fitness function of the particle is always excluded.
- Food categories that offer two options on the daily menu should always be different from each other.
- The algorithm takes into account special meals which are compound meals that do not fall into the initially proposed categories. Within these special meals is taken into account the hamburger, lasagna, tamal, etc. This special food is only presented once in the solution.
- The algorithm takes into account the preferences of the population for the elaboration of the menu. These were determined by the result of the DCE.
- The algorithm takes into account the number of times a certain food appears on the menu to guarantee its variety.
- This project will be carried out according to Colombian standards. The nutritional information of the project is based on the following documents:

- ○ Resolution 3803 of 2016: Recommendations for energy and nutrient intake (RIEN by its acronym in Spanish, Recomendaciones de ingesta de energía y nutrientes) for the Colombian population. (Ministry of health, 2016).
- ○ Characterization and Nutritional Composition of 18 traditional preparations in the Colombian population. (ICBF, 2017)
- ○ Colombian Food Composition Table (TCAC by its acronym in Spanish, Tabla de Composición de Alimentos Colombianos). (ICBF, 2018).

5.6 Compliance with the standard

- **SIX SIGMA DMAIC**

In this project the SIX SIGMA DMAIC standard was used, which according to Gutiérrez (2009) consists of carrying out five phases for quality in the optimization of processes. These phases are:

1. **Define:** In this phase, the main nutritional problems faced by the educational community of this school were discussed with their administration. It was defined that although the nutritional problems were not alarming in their community, since there are no students in a serious state of malnutrition, they do state that it is important to improve the meeting of nutritional requirements, at the same time that the school invests less money in buying supplies for the student's lunches.
2. **Measure**: In this phase, the school provided the menu they handled in the month of March, along with a price list for each of the foods. In order to carry out the nutritional measurements of the menu, the information on the macronutrients of each of the foods was consulted in the TCAC, an official colombian document for this matter. To determine the nutritional information for each dish, it is necessary to add the grams of each macronutrient per food On the other hand, in resolution 3803 of 2016 RIEN (Ministry of health, 2016) the nutritional requirements of children of high school age were consulted. Then, the average nutritional information of the lunches of the current situation of the school was compared with the requirements that must be met for each macronutrient. This is analyzed in depth on the before and after evaluation.
3. **Analyze:** The school states that it has never had the resources to consult a professional on the nutritional aspect of lunches. Until now, the decision on which meals to offer had been based solely on the experience of the kitchen manager and the supplies they had access to. On the other hand, the portion that was offered to each student was not standard. Each student decided approximately the amount of food they wanted. In the absence of a deeper analysis behind all these decisions, it is natural that the school failed to meet the nutritional requirements, and that they are driving a high price per lunch.
4. **Improve:** Through the application designed for the school, a PSO metaheuristic is used that considers the nutritional requirements of the students and the price of food supplies in order to produce a menu that improves the menu that the school currently provides its students.
5. **Control:** In this project the control phase was not reached, however recommendations were made to the school so that it can use the application and can maintain this tool over time. One of these recommendations is that the school periodically give students nutrition lectures, since the DCE found that children's nutritional preferences change when having nutritional information for what it's best for them.

- **Standard ISO / IEC 25010: 2011**

To guarantee that the final design of the application complies with the standard declared in the project as well as with the ISO / IEC 25010: 2011 Standard, Mrs. Nadia Lozada Torres, Administrative Director and Strategic Management of the school, who will make frequent use of the application, was asked to evaluate the main aspects established in the standard through an online survey. In this way, the survey was divided into two main sections, Quality in use Model and Product Quality Model, where different characteristics were evaluated:

Quality in use model: effectiveness, efficiency, satisfaction, freedom from risk and context coverage.
Within the "satisfaction" characteristic, we evaluate other sub-characteristics that the standard presents, such as: Utility, Trust, Pleasure and Comfort. Likewise, within the characteristic of "freedom from risk and context" we evaluate the following sub-characteristics: Economic risk mitigation and Health and safety risk mitigation. Finally, within the "Context coverage" characteristic, we evaluate the sub-characteristics of: Context coverage and Flexibility.

Product Quality model: functional suitability, reliability, performance efficiency, usability, security, compatibility, maintainability and portability.
The detail of each section at the survey, characteristic and sub-characteristic is found in Annex 9.

6. Results

6.1. Discrete Choice Experiment results

Table 9 shows the average marginal effects calculated from predictions of a previously estimated conditional logit model of Equation 4 (Annex 7), which includes the component regarding the treatment interaction. The average marginal effect significant at 99%, 95% and 99% of confidence for the treatment interaction corresponds to the attribute-level of price, white meat and red meat respectively. The attribute-level price has a negative sign, which suggests that students prefer low price meals over high price meals. When a lunch increases its price by 1,000 COP, the probability of choosing that lunch decreases by 3,26 percentage points. Nonetheless, the treatment interaction coefficient has a positive sign, which means that the previous probability decreases by 2.37 percentage points, and this means that students in the treatment group pay relatively less attention to the price attribute.

Moreover, the probability of choosing a lunch with white meat respect to a lunch with vegetal protein increments by 13.32 percentage points, but in presence of information the previous probability decreases to 2.45 percentage points, and the probability of choosing a lunch with red meat with respect to a lunch with vegetal protein increases by 9.73 percentage points, nonetheless it decreases by 9.33 percentage points for the treatment group. Thus, it is receiving more attention the vegetable protein level-attribute compared to their pairs. Overall, this suggests that students in the treatment group might have extra considerations for healthier food and less consideration than the control group for the price of the lunch. Evidently, provision of information on nutritional recommendation has an impact over student's nutritional preferences.

Additionally, the WTP for each significant level-attribute and the significant interactions for the treatment group are calculated, and shown in Annex 7.

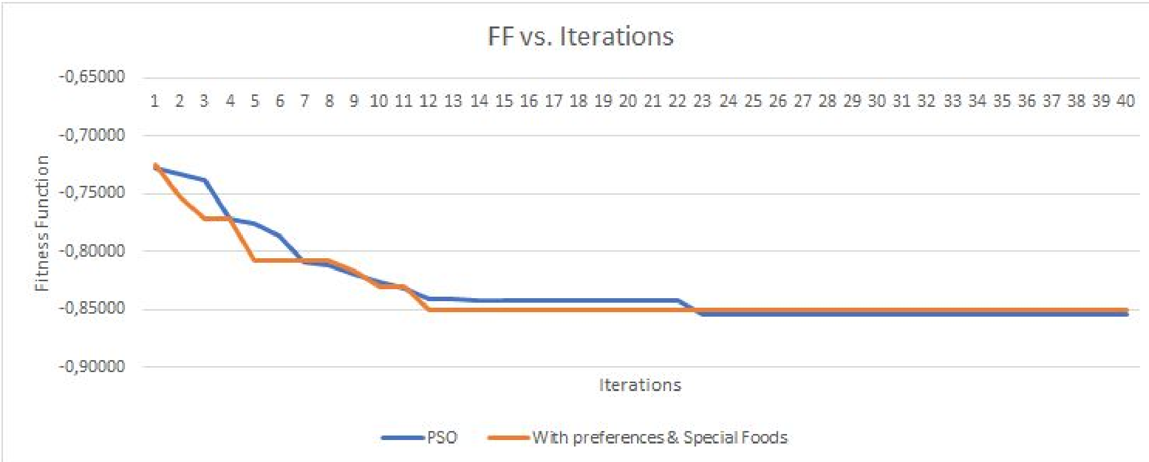
Table 9. Average marginal effects estimate for the conditional logit model with interactions.

Coefficients	Estimates dy/dx	Std. error	P-values
Price	-3.26x 10^{-5}	5.70x 10^{-6}	0.000
Fish	0.1218	0.0287	0.000
White meat	0.1332	0.0313	0.000
Red meat	0.0973	0.0334	0.004
Vegetable soup	0.0789	0.0269	0.003
Vegetable cream	0.1001	0.0242	0.000
One kind of carbohydrates	0.0829	0.0246	0.001
Two kinds of carbohydrates	-0.0127	0.0239	0.596
Fruit salad	0.0209	0.0243	0.390
Cold salad	0.0528	0.0236	0.025
Homemade dessert	0.0233	0.0176	0.186
Interactions			
Price*treatment	2.37x 10^{-5}	8.62x 10^{-6}	0.006
Fish*treatment	-0.0381	0.0427	0.372
White meat*treatment	-0.1087	0.0462	0.019
Red meat*treatment	-0.1906	0.0495	0.000
Vegetable soup*treatment	-0.0337	0.0398	0.397
Vegetable cream*treatment	0.0059	0.0363	0.870
One kind of carbohydrates*treatment	0.0357	0.0366	0.329
Two kinds of carbohydrates*treatment	0.0493	0.0356	0.166
Fruit salad*treatment	-0.0158	0.0364	0.664
Cold salad*treatment	0.0593	0.0357	0.097
Homemade dessert*treatment	-0.0210	0.0263	0.425

6.2. Results of the metaheuristic:

To evaluate the results of the algorithm, the input parameters with the best performance were taken into account according to the tests performed, that is, 40 iterations and a population size of 40 individuals. Then, Figure 13 shows the evolution of the global best through the iterations of the metaheuristics. The blue line shows the observed performance of the PSO, while the orange line shows the observed performance of the final solution, which considers special foods and preferences, outside the PSO. In this graph it is observed that the FF converges around iteration 15 in both cases. It is also observed that the global best, taking into account both the PSO and the final solution, has a very similar result. Analyzing this graph, it is evident that the metaheuristics entered a local optimum, and since it did not have a mutation component or a more exploratory component, the best solution of the PSO was not able to get out of this local optimum.

Figure 13. Evolution of the FF through iterations



Consequently, in order to analyze the behavior and proper functioning of the metaheuristics, it is decided to verify that all the foods in the database are analyzed by the particles and that from an objective function, it effectively chooses the best food as part of the solution. Table 10 shows evidence that all foods were considered by the population particles, it was not necessarily chosen as part of the best solution, but it was evaluated and may have been preserved or eliminated according to the behavior of the objective function.

Table 10. Number of times a food was chosen by any of the particles through the PSO

Option	Food	Number of times it was chosen
1	Bread (Pan)	520
2	Spaghetti Soup Pasta (Pasta Sopa Spaguetti)	389
3	White rice cooked without salt (Arroz blanco cocido sin sal)	378
.	.	.
246	Mute corn envuelto (Envuelto maíz mute)	4
247	Pasta (Pasta)	3
248	Potato (Papa)	3

In this way, it can be seen that *bread* tends to be chosen more times for its nutritional characteristics and for its price, while the opposite happens with *potatoes*. Thus, it is guaranteed that all foods were evaluated within the metaheuristics, that is, its operation is correct. However, taking into account that the PSO converges on a solution after a certain number of iterations, it is possible that it has remained in a local optimum, so it is considered as a point of improvement.

6.3. Comparison with the initial solution of the school

In our case, for the generation of a solution we took into account a maximum time of 1800 seconds (30 minutes), a size of 40 particles and 40 iterations, a solution for 15 days (school requirement) and a weight for the nutritional part of 0.6 (suggestion user). Also, we took into account the characteristics of the second group of students who are between the ages of 15 and 19, their average weight and the type of physical activity that the majority of users perform. According to the solution presented to the school in Annex 10, it can be seen that the variety requirements are met, by presenting a maximum of 4 times per food in the entire solution, except for the category of "Rice" and "Salad" because these categories present few options in the database. Additionally, in terms of preferences, it can be seen that the fish comes out 14 times in the entire solution, while the chicken comes out only 5 times, thus complying with the requirements of the educational community.

In comparison to the initial solution presented by the school, significant differences can be seen in fitness function and in the average cost for lunch of both solutions. Therefore, Table 11 shows the data of the initial solution versus the data provided by the application:

Table 11. Comparison of the average cost per menu, average Fitness Function per menu and Fitness Function of the best particle of the initial solution of the school vs the solution of the PSO

	Average cost	Average - FF Lunch	Personal Best (Particle)
Initial solution	7,566.044	0.046	-0.263
Metaheuristic solution	4,579.819	0.003	-0.842

Since the objective function of the project is to minimize the deviation of the nutritional requirements at the lowest possible cost, we can see significant results in the Fitness Function of the lunches on average generated by the PSO compared to the school solution, in 0.049 points as well as in the Fitness Function of the best particle in 0.579 points. Likewise, a decrease in the average price per lunch of 2,986,225 COP is evident. Moreover, taking into account the deviation of the nutritional requirements in terms of proteins, carbohydrates and fats, Table 12 shows the results presented by the PSO metaheuristic, including preferences and special meals for a 15-day menu:

Table 12. Average cost and nutritional requirements per menu of the PSO

	Average - Protein	Average - Carbs	Average - Fats	Average Cost
Metaheuristic solution	42.360g	129.184g	28.597g	4,579.819g

For the population of students we are analyzing, an approximate amount of 40,586g of protein, 148,814g of carbohydrates and 36,076g of fat is required for a one-day lunch. According to the results of the solution proposed by the algorithm, there is a deviation in port of 1,774g, in carbohydrates of -19,630g and in fat of -,7,481g. This at a cost of 4579, 819 COP on average for lunch. On the other hand, Table 13 presents the initial results proposed by the school for a 15-day menu:

Table 13. Average cost and nutritional requirements per menu of the initial school solution

	Average - Protein	Average - Carbs	Average - Fats	Average Cost
Initial solution	60.096g	173.238g	23.630g	7,566.044g

For the population of students we are analyzing, an approximate amount of 60,069 of protein, 148,814g of carbohydrates and 36,076g of fat is required for a one-day lunch. According to the results of the initial solution of the school, there is a deviation in proteins of 19,480g, in carbohydrates of 24,424g and in fat of -12,447g. This at a cost of 7566, 044 COP on average per lunch.

With the results obtained by the metaheuristics, a decrease in the average price per lunch of 2,986,225 COP Colombian pesos can be evidenced, as well as a achievement in percentage of the grams of proteins, carbohydrates and fats requires per lunch as follows:

- 95.81% in compliance with the required grams of protein for a lunch versus 67.57% established by the school.
- 86.81% in compliance with the required grams of carbohydrates for a lunch versus 74.57% established by the school.
- 79.27% in compliance with the required grams of fat for lunch versus 65.50% established by the school.

Achieving nutritional requirements is important for the students' diet as this contributes to their health and adequate academic performance. Additionally, minimizing the cost of lunch per student contributes to the financial needs of the school. According to the above and taking into account the amount of proteins, carbohydrates and fats that the analyzed population requires per day and the results presented by the school in comparison with the solution provided by the metaheuristic, it can be seen that the PSO solution has a less deviation in terms of nutritional requirements and presents a lower average price per lunch. In conclusion, the school is recommended to constantly use the application according to the needs it presents.

Evaluation of the final design of the application

To evaluate the final design of the application, it was presented to Mrs. Nadia Lozada, Administrative Director and Strategic Management of the school, who will make frequent use of it. Therefore, a survey was presented where the characteristics and sub-characteristics of the ISO / IEC 25010: 2011 standard were evaluated on a scale from 0 to 5 where 0 is "completely disagree" and 5 is "completely agree". The results of the survey are in Annex 9.

- **Quality in use model**

Regarding the results, for the section "Quality model in use" a rating of 4,545 was obtained out of a total of 5 points. Then, we obtained a rating of 5 in terms of efficiency, confidence, pleasure, comfort, economic risk mitigation, and health and safety mitigation. Likewise, we obtained a rating of 4 in terms of Effectiveness, Satisfaction, Utility, Context Coverage and Flexibility. Therefore, as points of improvement for these last aspects, a shorter execution time can be taken into account to provide a good solution or even present greater flexibility to include specific aspects of other needs presented by the school that arise at a certain time after the initial evaluation.

- **Product Quality Model**

On the other hand, for the "Product Quality Model" section we obtained a rating of 4,875 out of a total of 5 points. Then, we obtained a rating of 5 in terms of Functional suitability, Performance efficiency, Usability, Reliability, Security, Maintainability and Portability. Similarly, we obtained a rating of 4 in terms of Compatibility. Therefore, as points of improvement for this last aspect, it can be taken into account to make this application in different tools that are more compatible with the needs of the school or even in the computer console, and that does not depend on excel to execute the solution, that is, not limit the solution to a single tool.

6.4. Impact measurement:

The development of an application for the generation of a 15-day menu with the help of the PSO metaheuristic has proven to be efficient and significant compared to an initial solution presented by the school, where the nutritional requirements were not achieve in the most appropriate way and where a relatively high average cost per lunch was presented. The implementation and development of the degree work has proven to generate financial, social and operational impact in this educational community. Firstly, this work achieves a reduction in costs going from a value of 23’354,697 COP to a value of 15’938,332 COP generated by the algorithm. Secondly, compliance with the nutritional requirement for students increases on average 28.24% in protein, 12.24% in carbohydrates and 13.77% in fat. Finally, in the operational field, the application offers flexibility and a runtime of approximately 4.56 minutes, which considerably reduces planning time and menu effectiveness.

7. Conclusions and recommendations

In this project we present an algorithm that includes a PSO metaheuristic for the menu planning problem in Colegio Mayor de San Bartolomé, which aims to minimize nutritional deviations and the average cost per lunch, also it offers a varied menu solution for a time horizon of 15 days. This algorithm provides a better solution compared to the current one of the school, improving the nutritional fulfillment and saving 7’416,365 COP in food supplies.

Additionally, to solve the problem we take into account the estimated nutritional preferences of the educational community through a *Discrete Choice Experiment*, which also analyzes the effect of providing nutritional information on these preferences. The results show that the provision of information changes the students preferences, and these changes are consistent with an adequate nutrition, so it is recommended to provide nutritional information to the entire educational community to impact their preferences, and therefore, benefit them through a better diet choice.

Evidently, the linear mathematical model gives an optimal solution, however regarding this particular problem, we evidence that the execution time for one day and 8 lunches exceeds 4 hours, being a longer time than proposed, in addition to not satisfying the required variety by the school. On the other hand, according to the performance tests carried out, it can be seen that the best combination of input parameters for both the algorithm and the PSO corresponds to a population of 40 particles and 40 iterations, which generate an average execution time of 10 and 4.14 minutes respectively. In addition, it is evident that the algorithm works correctly by varying the input anthropometric data, so it is possible to change the sample population and obtain an adequate solution regarding the needs. Besides, the average solution regarding the joint solution for 15 days has a better nutritional and cost performance compared to a single solution day.

However, we found that the metaheuristic stagnates at a local optimum after approximately 15 iterations, thus it is proposed for future research to explore mutation in the algorithm, or even validate a hybridization between different methodologies to mitigate this problem. Among recommendations for future research, we propose to consider the impact and evaluation of an implemented solution in the school, include micronutrients within the nutritional deviation component of the PSO Fitness Function and explore the possibility to express all its components in monetary terms as well as implement methods to generate solution sets and determine the optimal weighting such as pareto fronts using scalarization methods.

8. Annex table

		Access link
Annex 1	Glossary	https://drive.google.com/file/d/1VGi3NEXh9R4B0gmMKmFQd34TAPuxjm34/view?usp=sharing
Annex 2	Minutes, costs and nutritional estimates of San Bartolomé lunches	https://drive.google.com/file/d/1GsiKv_fV9MA3FcIkPGXAEq9Tpuvx8JUd/view?usp=sharing
Annex 3	Literature review table	https://drive.google.com/file/d/12HOE4MoqUerTtK-hIMmv8Qz6PdhOaamY/view?usp=sharing
Annex 4	Survey with DCE to students	https://drive.google.com/file/d/13tLJRQB5a7RLu75UNTPbwhVhy_oyZx7M/view?usp=sharing
Annex 5	Mathematical model	https://drive.google.com/file/d/1r41KnzQTpnUkueKRfrryew8MoPG4Q4ds/view?usp=sharing
Annex 6	Clustering	https://drive.google.com/file/d/1xEVig3xGZShSL9Pa9h0dKN8k5t8C2jae/view?usp=sharing
Annex 7	Preference estimates	https://drive.google.com/file/d/1Hu08S10nOodWkHkknTYyx_5kCgJ-b9sT/view?usp=sharing
Annex 8	Performance tests for PSO	https://drive.google.com/file/d/1HgD6clB8OEQcw4a4Y7fenYBG65ZcwhYj/view?usp=sharing
Annex 9	Evaluation of the final application	https://drive.google.com/file/d/1crJRokTuSA0obRJYMVAlfYZ1IMLk33ks/view?usp=sharing
Annex 10	VBA Application - Menu Planning San Bartolomé	https://drive.google.com/file/d/1ktPOlTkbHLxS22UWsnmM4RJ7TJZGAbB7/view?usp=sharing

References

Al, A. O. E. T. (2012). *MEASURING CHILDREN ' S FOOD PREFERENCES : USING*. https://doi.org/10.1111/j.1745-459X.2012.00391.x

Arne Risa Hole, (2015). *"DCREATE: Stata module to create efficient designs for discrete choice experiments"* Statistical Software Components S458059, Boston College Department of Economics, revised 25 Aug 2017.

Anderson, A. M., & Earle, M. D. (1983). Diet planning in the third world by linear and goal programming. *Journal of the Operational Research Society*, *34*(1), 9–16. https://doi.org/10.1057/jors.1983.2

Análisis De Situación De Salud (ASIS) Dirección de Epidemiología y Demografía. (2018). https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/PSP/asis-nacional-2017.pdf

Analysis, M., Poor, I., & Trip, E. (2010). Clogit. 1–15.

Ayu, L., & Muhamad, S. (2019). ScienceDirect ScienceDirect Optimum Nutrition Intake from Daily Dietary Recommendation for Indonesian Children Using Binary Particle Swarm Optimization Algorithm. *Procedia Computer Science*, *157*, 16–24. https://doi.org/10.1016/j.procs.2019.08.136

Bastamipour, M, Mohamadi, Ehsan, Pourashraf, Y. (2013). Investigating local foods consumption in Ilam. *Advances in Environmental Biology*. http://www.aensiweb.com/aeb/2013/1956…

Bennett, B. J., Hall, K. D., Hu, F. B., Mccartney, A. L., & Roberto, C. (2015). Nutrition and the science of disease prevention: A systems approach to support metabolic health. *Annals of the New York Academy of Sciences*, *1352*(1), 1–12. https://doi.org/10.1111/nyas.12945

Blitstein, J. L., Cates, S. C., Hersey, J., Montgomery, D., Shelley, M., Hradek, C., Kosa, K., Bell, L., Long, V., Williams, P. A., Olson, S., & Singh, A. (2016). Adding a Social Marketing Campaign to a School-Based Nutrition Education Program Improves Children's Dietary Intake: A Quasi-Experimental Study. *Journal of the Academy of Nutrition and Dietetics*, *116*(8), 1285–1294. https://doi.org/10.1016/j.jand.2015.12.016

Burton, S., Creyer, E. H., Kees, J., & Huggins, K. (2006). *Attacking the Obesity Epidemic: The Potential Health Benefits of Providing Nutrition Information in Restaurants*. *96*(9), 1669–1675. https://doi.org/10.2105/AJPH.2004.054973

Buttorff, C., Trujillo, A. J., Diez-Canseco, F., Bernabe-Ortiz, A., & Miranda, J. J. (2015). Evaluating consumer preferences for healthy eating from Community Kitchens in low-income urban areas: A discrete choice experiment of Comedores Populares in Peru. Social Science and Medicine, 140, 1–8. https://doi.org/10.1016/j.socscimed.2015.06.033

Cattaneo, M. L. C., & Proserpio, V. B. C. (2016). *School children preferences for fish formulations: The impact of child and parental food neophobia*. *July*, 1–8. https://doi.org/10.1111/joss.12224

Chen, X.-C., Wei, T.-R., Guo, J.-H., Zhou, W.-X., Song, A., Chen, W.-N., & Zhang, J. (2019). Multiobjective Evolutionary Algorithm with Double-level Archives for Nutritional Dietary Decision Problem. *2019 9th International Conference on Information Science and Technology (ICIST)*, *2018*, 419–426. https://doi.org/10.1109/icist.2019.8836901

Choubey, D. K., Kumar, P., Tripathi, S., & Kumar, S. (2020). Performance evaluation of classification methods with PCA and PSO for diabetes. *Network Modeling Analysis in Health Informatics and Bioinformatics*, *9*(1). https://doi.org/10.1007/s13721-019-0210-8

England, P. H. (30 de Enero de 2020). *GOV.UK*. Obtenido de https://www.gov.uk/guidance/before-and-after-study-comparative-studies

Feier, B., B, I., Cristea, C., & S, R. (2017). *International Conference on Advancements of Medicine and Health Care through Technology; 12th - 15th October 2016, Cluj-Napoca, Romania*. *59*(October 2016), 107–110. https://doi.org/10.1007/978-3-319-52875-5

Fister, D., Fister, I., & Rauter, S. (2017). Generating eating plans for athletes using the particle swarm optimization. *CINTI 2016 - 17th IEEE International Symposium on Computational Intelligence and Informatics: Proceedings*, 193–198. https://doi.org/10.1109/CINTI.2016.7846402

Gerdessen, J. C., & De Vries, J. H. M. (2015). Diet models with linear goal programming: Impact of achievement functions. *European Journal of Clinical Nutrition*, *69*(11), 1272–1278. https://doi.org/10.1038/ejcn.2015.56

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact Evaluation in Practice.* https://openknowledge.worldbank.org/handle/10986/25030

Guo, W., Liu, G., Chen, G., & Peng, S. (2014). A hybrid multi-objective PSO algorithm with local search strategy for VLSI partitioning. *Frontiers of Computer Science*, *8*(2), 203–216. https://doi.org/10.1007/s11704-014-3008-y

Gutiérrez, H., Dela Vara, R., Perroy, R., & Careas, S. (2009). CONTROL ESTADÍSTICO DE LA CALIDAD Y 6 SIGMA. In Mc Graw Hill (Vol. 1). (424- 431) https://doi.org/10.1017/CBO9781107415324.004

Hierarchical clustering (scipy.cluster.hierarchy) — SciPy v1.5.3 Reference Guide. (2020). https://docs.scipy.org/doc/scipy/reference/cluster.hierarchy.html

Instituto Colombiano de Bienestar Familiar, & Colombia, U. N. de. (2017). Caracterización y composición nutricional de 18 preparaciones tradicionales en la población colombiana. In *Icbf*. https://www.icbf.gov.co/sites/default/files/caracterizacion-composicion-18-preparaciones-tradicionales-icbf1_0.pdf

Instituto Colombiano de Bienestar Familiar, & Colombia, U. N. de. (2018). Tabla de Composición de Alimentos Colombianos. In *Icbf*. https://www.icbf.gov.co/sites/default/files/tcac_web.pdf

ISO, International Organization for Standardization. (2001). ISO/ IEC 9126-1:2001, Software Engineering—Product Quality, Part 1: Quality Model.

ISO/IEC. (2011). *Online Browsing Platform (OBP)*. Obtenido de https://www.iso.org/obp/ui/#iso:std:iso-iec:25010:ed-1:v1:en

López, N. (2017). ¿El programa de alimentación escolar contribuye a mejorar el rendimiento académico?: Evaluación del PAE en Colombia. *Vniversitas Económica*, *17*(11), 1–44. https://drive.google.com/file/d/0BwCZbd8XuyfWZkZUX0dreHhwOGs/view

Koç, H., & Van Kippersluis, H. (2017). Thought for Food: Nutritional Information and Educational Disparities in Diet. https://www.journals.uchicago.edu/doi/abs/10.1086/694571

Kushner, Robert F. (2018). "Valoración y tratamiento de la obesidad." Harrison. Principios de Medicina Interna, 20e Eds. J. Larry Jameson, et al. New York, NY: McGraw-Hill, http://accessmedicina.mhmedical.com/content.aspx?bookid=2461&sectionid=213018712

Macmillan, R., & Page, J. (2009). A Game you can't Win. *Contexts*, *8*(3), 70–72. https://doi.org/10.1525/ctx.2009.8.3.70

Maldonado, Carlos. (2016). METAHEURÍSTICAS Y RESOLUCIÓN DE PROBLEMAS COMPLEJOS 1,2 METAHEURISTICS AND SOLVING COMPLEX PROBLEMS. *Revista Colombiana de Filosofía de la Ciencia.* 16. 169-185. https://www.researchgate.net/publication/311205182_METAHEURISTICAS_Y_RESOLUCION_DE_PROBLEMAS_COMPLEJOS_12_METAHEURISTICS_AND_SOLVING_COMPLEX_PROBLEMS

Marrero, A., Segredo, E., & Leon, C. (2019). *On the automatic planning of healthy and balanced menus*. 71–72. https://doi.org/10.1145/3319619.3326776

MEN. (2002). *DECRETO NÚMERO 1850 DE 2002*. *2002*(44). https://www.mineducacion.gov.co/1621/article-103274.html

MEN. (2015). Resolución 16432. https://www.icbf.gov.co/cargues/avance/docs/resolucion_mineducacion_16432_2015.htm

Ministerio de salud. (2016). Resolución número 3803 de 2016 - Recomendaciones de ingesta de energía y nutrientes (RIEN) para la población colombiana. *22 De Agosto*, 26. https://www.icbf.gov.co/cargues/avance/docs/resolucion_minsaludps_3803_2016.htm

Ministerio de salud. (2016). Recomendaciones de ingesta de energía y nutrientes (RIEN) para la población colombiana. Generalidades, usos y razones para su implementación. Documento técnico. https://www.icbf.gov.co/cargues/avance/docs/resolucion_minsaludps_3803_2016.htm

Ministerio de salud. (2016). Recomendaciones de ingesta de energía y nutrientes (RIEN) para la población colombiana. Generalidades, usos y razones para su implementación. Documento resumen. https://www.icbf.gov.co/cargues/avance/docs/resolucion_minsaludps_3803_2016.htm

MinSalud; DPS; INS; ICBF; UNAL. (2015). Encuesta Nacional de la Situación Nutricional ENSIN 2015. Nota de Política. In *Ministerio de Salud y Protección Social, Departamento Administrativo para la Prosperidad Social, Instituto Nacional de Salud, Instituto Colombiano de Bienestar Familiar, Universidad Nacional de Colombia* (Vol. 1). https://www.prosperidadsocial.gov.co/temporales/Encuesta%20Nacional%20de%20la%20Situación%20Nutricional%20-%20ENSIN%202015.pdf

Moreira, R. P. C., Wanner, E., Martins, F. V. C., & Sarubbi, J. F. M. (2018). An Evolutionary Mono-Objective Approach for Solving the Menu Planning Problem. *2018 IEEE Congress on Evolutionary Computation, CEC 2018 - Proceedings*. https://doi.org/10.1109/CEC.2018.8477888

Mousa, A. A., El-Shorbagy, M. A., & Abd-El-Wahed, W. F. (2012). Local search based hybrid particle swarm optimization algorithm for multiobjective optimization. *Swarm and Evolutionary Computation*, *3*, 1–14. https://doi.org/10.1016/j.swevo.2011.11.005

Ngo, H. C., Cheah, Y. N., Goh, O. S., Choo, Y. H., Basiron, H., & Kumar, Y. J. (2016). A review on automated menu planning approaches. *Journal of Computer Science*, *12*(12), 582–596. https://doi.org/10.3844/jcssp.2016.582.596

Pop, C. B., Chifu, V. R., Salomie, I., Cozac, A., & Mesaros, I. (2013). Particle Swarm Optimization-based method for generating healthy lifestyle recommendations. *Proceedings - 2013 IEEE 9th International Conference on Intelligent Computer Communication and Processing, ICCP 2013*, 15–21. https://doi.org/10.1109/ICCP.2013.6646074

Porras, E. M., Fajardo, A. C., & Medina, R. P. (2019). Solving dietary planning problem using particle swarm optimization with genetic operators. *ACM International Conference Proceeding Series*, 55–59. https://doi.org/10.1145/3310986.3311005

Ramírez, J. D., & Palacios, J. (1981). Glosario de términos piagetianos. *Infancia y Aprendizaje*, *4*(sup2), 123–143. https://doi.org/10.1080/02103702.1981.10821906

Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

Semana. (2019). *Denuncias sobre el Programa de Alimentación Escolar (PAE) en 10 departamentos del país*. https://www.semana.com/educacion/articulo/denuncias-sobre-el-programa-de-alimentacion-escolar-pae-en-10-departamentos-del-pais/638301

Soliman, A., De Sanctis, V., & Elalaily, R. (2014). Nutrition and pubertal development. *Indian Journal of Endocrinology and Metabolism*, *18*, S39–S47. https://doi.org/10.4103/2230-8210.145073

Sufahani, S., & Ismail, Z. (2014). A new menu planning model for Malaysian secondary schools using optimization approach. *Applied Mathematical Sciences*, *8*(149–152), 7511–7518. https://doi.org/10.12988/ams.2014.49725.

UNICEF. (2019). *Children, food and nutrition. The state of the world children 2019. Executive summary.* https://www.unicef.org/reports/state-of-worlds-children-2019

Yeasmin, S., & Islam, K. (2016). A comparative study of health, nutritional status, and dietary pattern of primary school going and dropout slum children in Dhaka City, Bangladesh. *Asian Journal of Medical Sciences*, *7*(4), 59–63. https://doi.org/10.3126/ajms.v7i4.14558

World Health Organization. (2012). How to Conduct a Discrete Choice Experiment for Health Workforce Recruitment and Retention in Remote and Rural Areas: A user guide with case studies. https://www.who.int/hrh/resources/dceguide/en/

Wu, G., Qiu, D., Yu, Y., Pedrycz, W., Ma, M., & Li, H. (2014). Expert Systems with Applications Superior solution guided particle swarm optimization combined with local search techniques. *EXPERT SYSTEMS WITH APPLICATIONS*, *41*(16), 7536–7548. https://doi.org/10.1016/j.eswa.2014.06.005

Zhang, Q., Chen, Z., Diawara, N., & Wang, Y. (2011). Prices of Unhealthy Foods, Food Stamp Program Participation, and Body Weight Status Among U.S. Low-Income Women. *Journal of Family and Economic Issues*, *32*(2), 245–256. https://doi.org/10.1007/s10834-010-9228-x

Zhang, Y., Wang, S., & Ji, G. (2015). A Comprehensive Survey on Particle Swarm Optimization Algorithm and Its Applications. *Mathematical Problems in Engineering*, *2015*. https://doi.org/10.1155/2015/931256