

Food Nutrition Analysis

Project Scope

The project was proposed to analyze the Macro nutrition (Protein, Carbs and Fats) on vegetarian food items. The Project document starts with the Scope of the project, followed by objective of the project, Source and Limitation, Query used (SQL), and a conclusion that talks about the value that can be fetched from the dataset.

Objective

This project is to clean the dataset that contains nutrition facts of Both non-vegetarian and vegetarian foods and prepare a dataset that can help to customize a diet plan packed with appropriate nutrients.

Data Source and Limitation:

The source of the data used in the project is fetched from Kaggle and based on analysis it was identified that the nutrition facts are not completely accurate but can provide a valuable insight on nutrition to formalize a vegetarian diet based on the transformation goal.

Code:

```
Use DA;

Select * from NutionFact_Dataset;

/*Rename Nutrition Fact Dataset to a short abbreviation*/

exec sp_rename 'NutionFact_Dataset', 'NF'

Select * from NF; /*Table Name Changed Successfully*/

/* Data Cleaning - Creating a table with only Veg Food Items */

SELECT * into Veg_NF from NF where Category != 'Meat, Poultry' OR Category != 'Fish, Seafood'
OR Category != 'Drinks,Alcohol, Beverages' or Food != '%Eggs%' OR Food != 'Beef' or Food != ' Yolks';
```

```

/* Viewing Data */
select * from Veg_NF;
select * from Veg_NF where Protein is NULL or Carbs is NULL;
select Food, carbs from Veg_NF where Carbs is NULL;

/* Data Cleaning - Updating Null Carbs to Zero as major of the Null carbs are fat sources
with negligible carbs in nature*/
Update Veg_NF set carbs = 0 where Carbs is null;
Update Veg_NF set Fiber = 0 where Fiber is null;

select * from Veg_NF where Calories = 4 * Carbs AND Protein is NULL;
select * from Veg_NF where Calories < 4 * Carbs AND Protein is NULL;

/* Due to Round off data the carbs multiplied by 4 value is passing calorie value which
is technically wrong
Hence setting the Nutrition value to double and subtracting 0.5 for such value to
calculate Protein*/

/* To check the Datatypes of Each Fields */
Exec sp_columns Veg_NF;

Alter Table Veg_NF Alter Column Carbs Decimal(10,2);
Alter Table Veg_NF Alter Column Protein Decimal(10,2);

Update Veg_NF set Carbs = carbs - 0.51 where Calories < 4 * Carbs AND Protein is NULL;

/*Updating Pineapple, Lemon Juice data and removing rest of the incorrect data
considering importance */

Update Veg_NF Set Carbs = 16 where Food = 'Pineapple'
Update Veg_NF Set Protein = 0.6 where Food = 'Pineapple'
Update Veg_NF Set Fat = 0.1 where Food = 'Pineapple'
Update Veg_NF Set Calories = 67 where Food = 'Pineapple'

Delete Veg_NF where Calories < 4 * Carbs AND Protein is NULL;

/* Fixing Null Protein */
Update Veg_NF set Protein = 0 where Calories = 4 * Carbs AND Protein is NULL;

/*setting a value 0 for fats in vegetables, fruits, and puffed rice as most of the
vegetables/Fruits and rice
will be negligible in fats considering the units given in the data set*/

Update Veg_NF set Fat = 0
where Protein is NULL And Category = 'Vegetables';

/*Create a SP to create a Fat and Carbs per 100g*/

Update Veg_NF set Fat = 0
where Protein is NULL And Category = 'Fruits';

Update Veg_NF set Fat = 0
where Protein is NULL And Food = 'Puffed rice';

/*setting a value 0 for protein as Butter, margarine, mayonise are fat sources with
negligible proteins*/

```

```

Update Veg_NF set protein = 0
where Protein is NULL And Category = 'Fats, Oils, Shortenings';

select * from veg_nf where Fat is NULL;

/* Removing 3 items from deserts with insufficient data */
Delete Veg_NF where Protein is NULL And Category = 'Desserts, sweets';

/* Now calculating Protein for the 12 Null with Fats and Carbs*/

update Veg_NF set protein = (Calories - ((4 * carbs) + (9 * Fat))) / 4 where Protein is
null;

/* Now calculating Fats for the rest of Fat with Protein and Carbs*/

update Veg_NF set Fat = (Calories - ((4 * carbs) + (4 * protein))) / 9 where Fat is null;

/* Create and calulate calculating Protein per 100g */
Alter table Veg_NF ADD Protein_Per_100g decimal (10,2);

update Veg_NF set Protein_Per_100g = (Protein / [Unit (Grams)]) * 100;

/*Validating and removing data based on incorrect information*/

select * from Veg_NF where Protein < 0 or Fat < 0 or carbs < 0;
Delete Veg_NF where food = 'French-fried';

select * from Veg_NF
where Calories !< ((4*Protein) + (9*Fat) + (4*Carbs)) + 20 OR
Calories !> ((4*Protein) + (9*Fat) + (4*Carbs)) - 20;

Delete Veg_NF where Category = 'Desserts, sweets' AND Calories !< ((4*Protein) + (9*Fat)
+ (4*Carbs)) + 20 OR
Calories !> ((4*Protein) + (9*Fat) + (4*Carbs)) - 20;

Delete Veg_NF where
food = 'Potatoes Mashed with milk and butter' OR
food = 'Potatoes, pan-tried' OR
food = 'Scalloped with cheese potatoes' OR
food = 'Potato chips';

Update Veg_NF Set Carbs = 60 where Food = 'Soybeans'
Update Veg_NF Set Protein = 72 where Food = 'Soybeans'
Update Veg_NF Set Fat = 40 where Food = 'Soybeans'
Update Veg_NF Set Calories = 888 where Food = 'Soybeans'

/*Create a carbs and Fats per 100g of Food*/

Alter Table Veg_NF Add Carbs_per_100g Decimal(10,2);
update Veg_NF set Carbs_per_100g = (Carbs/[Unit (Grams)]) * 100;

Alter Table Veg_NF Add Fat_per_100g Decimal(10,2);
update Veg_NF set Fat_per_100g = (Fat/[Unit (Grams)]) * 100;

select * from Veg_NF where fat_per_100g = 0.00;

```

Conclusion:



Nutrition
Dataset_Cleaned.csv

From the above cleaned Dataset, a dietician can fetch Macro Nutrients on most of the locally available vegetarian Food Items to customize a diet plan based on the Customer's Fitness Goals.