What's in Your Food?
A Data-Driven Nutrient Analysis

By Alina Yildir

Please read more at

Introduction

According to **Canada's Food Guide**, maintaining a balance of nutrients such as **protein**, **fibre**, **carbohydrates**, and **potassium** contributes to **better overall health**. At the same time, reducing the intake of **saturated fat**, **trans fat**, **sugar**, **sodium**, and **cholesterol** plays a key role in minimizing **health risks**.
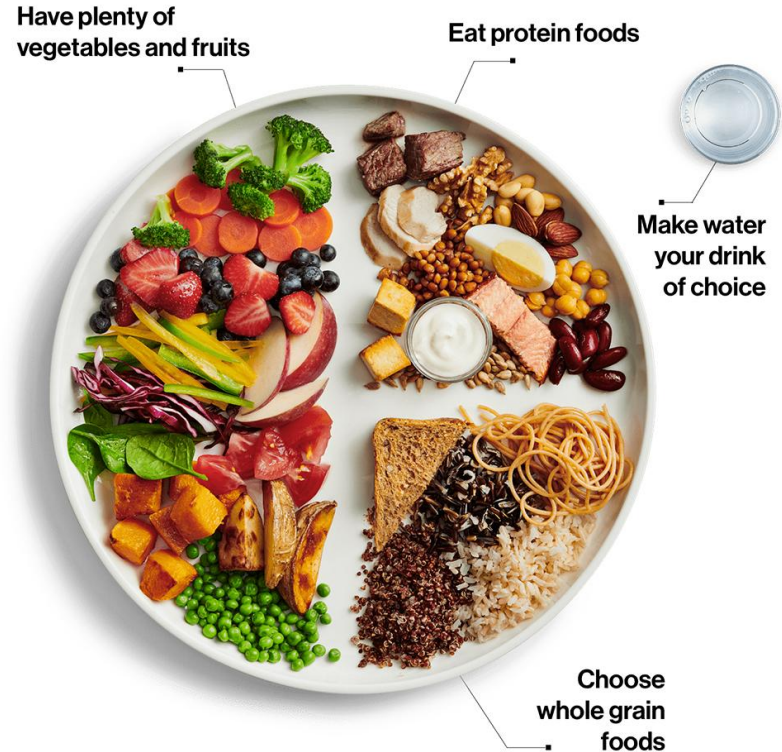

Have plenty of vegetables and fruits

Eat protein foods

Make water your drink of choice

Choose whole grain foods

Image source: Canada's Food Guide, Government of Canada. Retrieved from https://food-guide.canada.ca/en/.

Objectives

The objective is to analyze food categories, subcategories, and individual foods to identify the **highest and lowest levels of the 12 key nutrients** from the **Nutrition Facts Table**, compare **protein-to-fat ratios** across subcategories, and evaluate **nutrient density per calorie** for **macronutrients (carbohydrates, protein, and fat)**.

Nutrition Facts
Valeur nutritive

Per 1 cup (250 mL)
pour 1 tasse (250 mL)

		% Daily Value* % valeur quotidienne*
Calories 110		
Fat / Lipides 0 g		0 %
Saturated / saturés 0 g + Trans / trans 0 g		0 %
Carbohydrate / Glucides 26 g		
Fibre / Fibres 0 g		0 %
Sugars / Sucres 22 g		22 %
Protein / Protéines 2 g		
Cholesterol / Cholestérol 0 mg		
Sodium 0 mg		0 %
Potassium 450 mg		13 %
Calcium 30 mg		2 %
Iron / Fer 0 mg		0 %

*5% or less is **a little**, 15% or more is **a lot**
*5 % ou moins c'est **peu**, 15 % ou plus c'est **beaucoup**

Image source: Nutrition Facts Tables, Government of Canada. Retrieved from https://www.canada.ca/en/health-canada/services/food-nutrition/nutrition-labelling/nutrition-facts-tables.html.

Data Preprocessing

Dataset: **The Canadian Nutrient File** (Health Canada) – provides nutrient data for 1,000+ commonly consumed foods, covering 19 key nutrients.

Original Format: 17 separate CSV files (one per food category).

Preprocessing Steps (applied to each file individually):
- Removed unnamed rows/columns containing only missing values.
- Flattened multi-level headers while preserving relevant subheadings.
- Selected 12 key nutrients based on **Nutrition Facts Table** guidelines.
- Standardized column names to align with Health Canada's format.
- Replaced "tr" (trace amounts) and "N/A" (no suitable value available) with 0.
- Added the missing + Trans (g) column where necessary.
- Retained only relevant columns and reordered them for consistency.
- Normalized nutrient values to be per 100g instead of per serving size for standardization.

Final Merged Dataset: 1,098 foods, 16 columns (Food Name, Calories, Fat, Carbohydrates, Protein, Sodium, Iron, etc.).

Please read more at
https://github.com/yildiramdsa/nutrient_composition_of_common_foods_in_canada_analyzing_the_canadian_nutrient_file/blob/main/notebooks/data_preprocessing.ipynb.

Question 1. Which food categories contain the highest and lowest levels of the 12 key nutrients listed in the Nutrition Facts Table?

Question 2. Within a selected food category, which subcategories have the highest and lowest levels of these nutrients?

Question 3. Within a selected subcategory, which individual foods have the highest and lowest levels of these nutrients?

https://public.tableau.com/app/profile/alina.yildir/viz/AData-DrivenNutrientAnalysis/AData-DrivenNutrientAnalysis2

https://public.tableau.com/app/profile/alina.yildir/viz/AData-DrivenNutrientAnalysisPerServing/AData-DrivenNutrientAnalysis

Question 4. Which food subcategories have the highest and lowest protein-to-fat ratios?

https://public.tableau.com/app/profile/alina.yildir/viz/AData-DrivenNutrientAnalysis/AData-DrivenNutrientAnalysis2
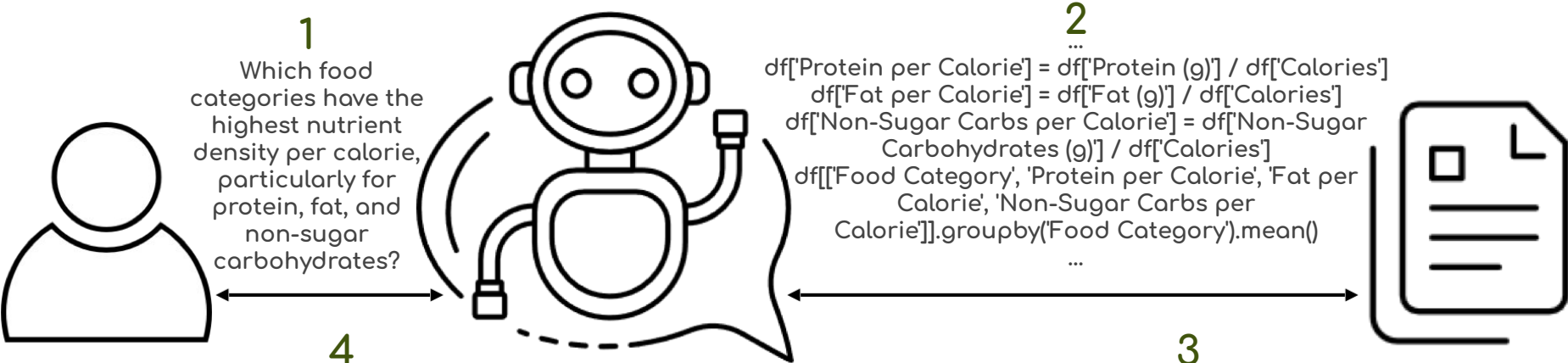
Question 5. Which food subcategories have the highest nutrient density per calorie, particularly for macronutrients (carbohydrates, protein, and fat)?

https://public.tableau.com/app/profile/alina.yildir/viz/AData-DrivenNutrientAnalysis/AData-DrivenNutrientAnalysis2

https://yildiramdsa-nutrient-composition--csv-chatbotcsv-chatbot-kdmzcd.streamlit.app

Database Chatbots: Interacting with CSV Data

1

Which food categories have the highest nutrient density per calorie, particularly for protein, fat, and non-sugar carbohydrates?

2

```
...
df['Protein per Calorie'] = df['Protein (g)'] / df['Calories']
df['Fat per Calorie'] = df['Fat (g)'] / df['Calories']
df['Non-Sugar Carbs per Calorie'] = df['Non-Sugar Carbohydrates (g)'] / df['Calories']
df[['Food Category', 'Protein per Calorie', 'Fat per Calorie', 'Non-Sugar Carbs per Calorie']].groupby('Food Category').mean()
...
```

4

Protein Density: The 'Fish and Shellfish' category has the highest protein density per calorie, making it an excellent choice for high-protein diets. This is followed by 'Meat and Poultry', and 'Eggs and Egg Dishes'.

Fat Density: The 'Fats and Oils' category leads with the highest fat density per calorie, followed by 'Eggs and Egg Dishes' and 'Legumes, Nuts and Seeds'.

Non-Sugar Carbohydrate Density: 'Breads, Cereals and Other Grain Products' rank highest, making them ideal for diets that focus on high fiber and complex carbohydrates. This is followed by 'Vegetables and Vegetable Products' and 'Snacks'.

3

...	
Food Category	Protein per Calorie
Baked Goods	0.015058
Beverages	0.003933
...	
Food Category	Fat per Calorie
Baked Goods	0.015058
Beverages	0.003933
...	
Food Category	Non-Sugar Carbs per Calorie
Baked Goods	0.095682
Beverages	0.078777
...	

Summary

- Identified food categories with the highest and lowest levels of 12 key nutrients.
- Analyzed subcategories within each food category to determine nutrient variation.
- Examined individual foods within subcategories to highlight the most and least nutrient-dense options.
- Compared food subcategories based on protein-to-fat ratios.
- Evaluated nutrient density per calorie for macronutrients (carbohydrates, protein, and fat).

References

Government of Canada. (n.d.). *Canada's food guide*. Retrieved from https://food-guide.canada.ca/en/.

Government of Canada. (n.d.). *Nutrition facts tables*. Retrieved from https://www.canada.ca/en/health-canada/services/food-nutrition/nutrition-labelling/nutrition-facts-tables.html.

Government of Canada. (n.d.). *Canadian nutrient file*. Retrieved from https://open.canada.ca/data/en/dataset/a289fd54-060c-4a96-9fcf-b1c6e706426f.

Thank You for Your Attention!

Please read more at
https://medium.com/@yildir.a.mdsa/whats-in-your-food-a-data-driven-nutrient-analysis-e3a0f7a5c553.