Project 1 Proposal – Group 4

**Project Title**

One (Healthy) Recipe to Rule Them All

**Team Members**

1. Lakna Premasinghe
2. Foluke Daramola
3. Alyssa Hondrade

**Project Description**

The goal of the project is to compare recipes from two popular recipe websites, Spoonacular and Food.com, and identify “healthy” recipes using two market-implemented measures: Nutri-Scores [[Link]](https://www.mangiu.ch/index/) and Weight Watchers (WW) Smart Points [[Link]](https://www.weightwatchers.com/us/how-it-works/smartpoints). The project will investigate recipe popularity, meal types, and cuisines.

Scope

The nutritional values of interest will be reduced to the minimum required for calculating the “health” score: **calories**, **saturated fat**, **sugar**, and **protein**.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Calories | Total Fat | Saturated fat | Sugar | Sodium | Protein | Fibre |
| Spoonacular | X | X | X | X | X | X | X |
| Food.com | X | X | X | X | X | X | - |
| Nutri-Score | X | - | OPT | OPT | OPT | OPT | OPT |
| WW formula | X | - | X | X | - | X | - |

The scope of the meal types will be narrowed down to **breakfast**, **lunch**, and **dinner**.

**Research Questions**

1. Are more popular / higher-rated recipes healthier? What is the health rating of the highest-rated recipes?
2. What meal type (i.e., breakfast, lunch, or dinner) have the healthiest/unhealthiest (percentage) recipes? What is the most popular ingredient for each meal?
3. Which cuisine has the healthiest recipes?

**Datasets**

1. Spoonacular API [[Link]](https://spoonacular.com/food-api/docs#Search-Recipes-Complex)
2. Kaggle – Food.com Recipes and Interactions [[Link]](https://www.kaggle.com/datasets/shuyangli94/food-com-recipes-and-user-interactions?select=RAW_recipes.csv)

**Task Breakdown**

Tue, 29 Aug 2023

1. Team admin.
   1. Ensure all team members have the required API keys and datasets.
   2. Ensure all team members have access to the GitHub repository.
2. Define strategy and metrics: identify attributes, how many of each, etc.
3. Build a data retrieval plan: variables and for-loops required.
4. Begin retrieving the data.

Wed, 30 Aug 2023 – Retrieve the data.

* Convert Spoonacular API requests to dataframes.
* Merge data frames.

Thu, 31 Aug – 01 Sep 2023 – Assemble and clean the data.

* Stage 1 assemble/clean:
  + Remove rows without minimum values for the Nutri-Score calculation.
  + Calculate the Nutri-Score and create updated dataframe.
* Stage 2 assemble/clean:
  + Convert Spoonacular aggregate-likes score to a 5-point system.
  + Calculate Weight Watchers Smart Points for Food.com (this data is available in the Spoonacular dataset), given the following formula [[Link]](https://www.watcherspoint.com/weight-watchers-smart-points-calculator):

Sat/Sun/Mon, 02–04 Sep 2023 – Analysis & Data Visualisation

* Are more popular / higher-rated recipes healthier?
  + Sort the data by the highest ratings and identify the top recipes.
  + Isolate the Nutri-Score and WW Smart Point column to identify trends.
  + Create a scatter plot of popularity vs each scoring system.
* What is the health rating of the highest-rated recipes?
  + Of the highest-rated, create a visualisation of the health ratings.
  + Note any observations.
* What meal type (i.e., breakfast, lunch, or dinner) have the healthiest/unhealthiest (percentage) recipes? What is the most popular ingredient for each meal?
  + Filter the data by the meal type.
  + Create a plot of the health scores per meal type.
  + If applicable, identify the most popular ingredient for each.
* Which cuisine has the healthiest recipes?
  + Filter the data by cuisine.
  + Create a scatter plot of the WW Smart Points, distinguishing cuisine by colour.
  + Of the “healthy” cuisines, if identifiable, select a few cuisines for further analysis. Otherwise, select the cuisine with the most recipes.
  + Create a plot that best answers the question.

Tue, 05 Sep 2023 – Conclusions + Create Slide Deck

* Summarise findings and identify key points for the presentation.
* Create the slide deck and finalise the README file.
* Identify presentation allocation.

Wed, 06 Sep 2023 – Presentation Practice