ChE 488 Project Proposal
Matthew Grattan and George Ho

1. ## Diet Problem

We will find the minimum cost, feasible diet (subject to nutrition constraints) for several countries (this is a simple LP, for each country). We will compare and contrast the resulting diets.

- The diet problem is a simple LP: we are less interested in the method of solution, and more interested in the actual answer to the optimization problem.
- Possible data sources:
  - Global food prices:
    - https://www.kaggle.com/jboysen/global-food-prices
    - https://public.opendatasoft.com/explore/dataset/global-food-prices-database-wfp/table/
  - Nutrition information:
    - https://catalog.data.gov/dataset/mypyramid-food-raw-data-f9ed6
    - https://data.world/adamhelsinger/food-nutrition-information

2. ## Protein Classification with Support Vector Machines

We will perform image classification of 16 different species of proteins using SVMs.

- Support vector machines have a strong foundation in optimization theory (with strong mathematical guarantees on the maximally separating hyperplane) and a good track record with image classification. We would be applying them to a relevant and impactful problem.
- Possible data sources:
  - https://www.kaggle.com/c/human-protein-atlas-image-classification/data

3. ## Scheduling Problem

We will investigate a weakened form of the scheduling problem (as applied to class scheduling at Cooper), investigating whether or certain simplifying assumptions still lead to (sufficiently) interesting solutions.

- Possible simplifying assumptions include: constraining conflicting classes to conflict entirely, and constraining professors to
- Data sources are limited, so the easiest and most viable thing to do would be to create a toy dataset.