TEAM NAME: **Quadruples**

TEAM MEMBERS:

1. Amruta Folane (NetID: asf160130)
2. Ashish Mohapatra (NetID: axm160031)
3. Dhruv Sangvikar (NetID: dgs160230)
4. Rohit Sindhu (NetID: rks160030)

PROJECT TYPE: Custom

PROJECT TITLE: Correlation of Fast Food Joints density and chronic diseases in the USA.

DATA SOURCES:

1. US Department of Health and Human Services.
    a. Nutrition, Physical Activity, and Obesity - Behavioural Risk Factor Surveillance System
    b. Link: https://catalog.data.gov/dataset/nutrition-physical-activity-and-obesity-behavioral-risk-factor-surveillance-system/resource/fdb2306b-13cc-4925-9633-cd0030bf482b?inner_span=True
    c. Domain: Federal
    d. Description: This dataset includes data on adult's diet, physical activity, and weight status from Behavioural Risk Factor Surveillance System. This data is used for DNPAO's Data, Trends, and Maps.
2. US Department of Health and Human Services.
    a. Specific: Centres for Disease Control and Prevention.
    b. Link: https://catalog.data.gov/dataset/center-for-medicare-amp-medicaid-services-cms-medicare-claims-data
    c. Domain: Federal
    d. Description: This dataset includes Inpatient and Outpatient claims, Master Beneficiary Summary Files, and many other files. Indicators from this data source have been computed by personnel in CDC's Division for Heart Disease and Stroke Prevention (DHDSP). This is one of the datasets provided by the National Cardiovascular Disease Surveillance System. The system is designed to integrate multiple indicators from many data sources to provide a comprehensive picture of the public health burden of CVDs and associated risk factors in the United States.
3. Fast Food joint distribution in the United States.
    a. Link: http://www.fastfoodmaps.com/data.html
    b. Domain: Federal
    c. Description: Author wrote scrapers for other major fast food chains and produced a list of 50,000 locations in the US. Using the site, people can go to any location in the US and see the major fast food restaurants there. Author's intention wasn't to make it easier for people to find fast food, nor was it to criticize the ubiquity of fast food. Rather he was interested in the visualization, the graphical portrayal of information. These graphics convey something that is true about the US.

TOOLS NEEDED:

1. CSV to RDF converters:
    a. https://www.w3.org/wiki/ConverterToRdf

   b. https://www.w3.org/wiki/ConverterToRdf#CSV_.28Comma-Separated_Values.29
2. Apache Jena Fuseki
3. Google Charts APIs: https://developers.google.com/chart/
4. SPARQL Queries (for data manipulation)
5. Web Framework for application

BASIC SCHEDULE:

1. Team will meet for 4 hours every week (Mon-Tue 11am-1pm).
2. Individual tasks will be discussed, distributed and integrated.
3. Milestone 1:
   a. Date: 16th October, 2017
      i. Convert data to appropriate RDF formats
      ii. Make sure the tools, data and approach of the project is on the right track.
      iii. APIs are identified. Read up how to use them.
4. Milestone 2:
   a. Date: 24th October, 2017
      i. Check readiness for first checkpoint of presentation.
5. Milestone 3:
   a. Date: 6th November, 2017
      i. First version of output with all integrations and queries working.

EXPECTED RESULTS:

The correlation is established showing the density distribution of fast food joints with health diseases. Also, the relationship between Medicare claims and heart related diseases is shown. The result will be output using Google Charts API, marking the geographical locations with the data obtained from the RDF datasets.

BRAINSTORM SESSION:

On brainstorming for different project ideas, each one of us come with the following respective ideas:

1. Amruta Folane:
   Prices of houses in an area in relation with the salaries of the people in the same area.
2. Ashish Mohapatra:
   Imdb movie ratings in relation with their box office collection.
3. Dhruv Sangvikar:
   Correlation between Museums locations and their artifacts category, history and other related information.
4. Rohit Sindhu:
   Restaurants and their menus, and how a particular recipe on the menu is prepared in terms of raw materials, etc.