### Context

Health care in the United States is provided by many distinct organizations. Health care facilities are largely owned and operated by private sector businesses. 58% of US community hospitals are non-profit, 21% are government owned, and 21% are for-profit. According to the World Health Organization (WHO), the United States spent more on healthcare per capita ($9,403), and more on health care as percentage of its GDP (17.1%), than any other nation in 2014. Many different datasets are needed to portray different aspects of healthcare in US like disease prevalences, pharmaceuticals and drugs, Nutritional data of different food products available in US. Such data is collected by surveys (or otherwise) conducted by Centre of Disease Control and Prevention (CDC), Foods and Drugs Administration, Center of Medicare and Medicaid Services and Agency for Healthcare Research and Quality (AHRQ). These datasets can be used to properly review demographics and diseases, determining start ratings of healthcare providers, different drugs and their compositions as well as package informations for different diseases and for food quality. We often want such information and finding and scraping such data can be a huge hurdle. So, Here an attempt is made to make available all US healthcare data at one place to download from in csv files.

### Content

* **Nhanes Survey (National Health and Nutrition Examination Survey)** - The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The survey is unique in that it combines interviews and physical examinations. NHANES is a major program of the National Center for Health Statistics (NCHS). NCHS is part of the Centers for Disease Control and Prevention (CDC) and has the responsibility for producing vital and health statistics for the Nation. The NHANES interview includes demographic, socioeconomic, dietary, and health-related questions. The examination component consists of medical, dental, and physiological measurements, as well as laboratory tests administered by highly trained medical personnel. The diseases, medical conditions, and health indicators to be studied include: Anemia, Cardiovascular disease, Diabetes, Environmental exposures, Eye diseases, Hearing loss, Infectious diseases, Kidney disease, Nutrition, Obesity, Oral health, Osteoporosis, Physical fitness and physical functioning, Reproductive history and sexual behavior, Respiratory disease (asthma, chronic bronchitis, emphysema), Sexually transmitted diseases, Vision. 10000 individuals are surveyed to represent US statistics.  
  Five files in this datasets represent current recent Nhanes data -  
    
  Nhanes20052006.csv   
    
  Nhanes20072008.csv   
    
  Nhanes20092010.csv   
    
  Nhanes20112012.csv  
    
  Nhanes20132014.csv
  + **Data fields' description** -
    - Nhanes20052006.csv - [Demographic](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Demographics&CycleBeginYear=2005), [Dietary](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Dietary&CycleBeginYear=2005), [Examinations](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Examination&CycleBeginYear=2005), [Laboratory](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Laboratory&CycleBeginYear=2005)
    - Nhanes20072008.csv - [Demographic](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Demographics&CycleBeginYear=2007), [Dietary](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Dietary&CycleBeginYear=2007), [Examinations](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Examination&CycleBeginYear=2007), [Laboratory](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Laboratory&CycleBeginYear=2007)
    - Nhanes20092010.csv - [Demographic](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Demographics&CycleBeginYear=2009), [Dietary](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Dietary&CycleBeginYear=2009), [Examinations](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Examination&CycleBeginYear=2009), [Laboratory](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Laboratory&CycleBeginYear=2009)
    - Nhanes20112012.csv - [Demographic](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Demographics&CycleBeginYear=2011), [Dietary](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Dietary&CycleBeginYear=2011), [Examinations](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Examination&CycleBeginYear=2011), [Laboratory](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Laboratory&CycleBeginYear=2011)
    - Nhanes20132014.csv - [Demographic](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Demographics&CycleBeginYear=2013), [Dietary](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Dietary&CycleBeginYear=2013), [Examinations](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Examination&CycleBeginYear=2013), [Laboratory](http://https:/wwwn.cdc.gov/nchs/nhanes/search/variablelist.aspx?Component=Laboratory&CycleBeginYear=2013)
* **US Drugs datasets** - FDA provides a database for searching all the published drugs and all the unpublished drugs on their website, This database provides all the information about package of drugs and compositions of drugs their NDC codes. Description of variables for this datasets are as follows -
  + Drugs\_product (current and unfinished)
    - PRODUCTID - Id of the product
    - PRODUCTNDC - National drug code of the product
    - PRODUCTTYPENAME - Type of the product
    - PROPRIETARYNAME - Proprietary name of the product
    - PROPRIETARYNAMESUFFIX - Proprietary name Suffix
    - NONPROPRIETARYNAME - Non- proprietary (common name) of the product
    - DOSAGEFORMNAME - Dosage information
    - ROUTENAME - Route of taking drugs (Oral / Injections)
    - STARTMARKETINGDATE - Date on which marketing for the drug has started
    - ENDMARKETINGDATE - Date on which the marketing for the drug has stopped
    - MARKETINGCATEGORYNAME - Marketing category name
    - APPLICATIONNUMBER - Application number for registering drug
    - LABELERNAME - Labeler name
    - SUBSTANCENAME - Names of the substances in drug
    - ACTIVENUMERATORSTRENGTH - Strength of the drug
    - ACTIVEINGREDUNIT - Unit of strength
    - PHARM\_CLASSES - Pharmaceutical class of the drugs
    - DEASCHEDULE - DEA schedule
  + Drugs Package (current and unfinished)
    - PRODUCTID - Id of the product
    - PRODUCTNDC - National drug code of the product
    - NDCPACKAGECODE National drug code of the package
    - PACKAGEDESCRIPTION - description of the p[ackage
* **Nutritions Data from USDA** - Whenever we buy a packaged food product, we find the nutritional fact written on it. United States Department of Agriculture Agricultural Research Service’s Food composition database. This database contains all kinds food products available in US and provides description of their nutritions. This dataset is web scrapped and converted into a csv file. Variables are self-explanatory names yet the descriptions can be found at this link - variables descriptions -( All values are per 100 grams) -
  + Data fields' description -
    - NDB\_No - Nutrition database number
    - Shrt\_Desc - Short description
    - Water\_(g) - water in grams per 100 grams
    - Energ\_Kcal - Energy in Kcal
    - Protein\_(g) - Protein
    - LipidTot(g) - Total Lipid
    - Ash\_(g) - Ash
    - Carbohydrt\_(g) - Carbohydrate, by difference
    - FiberTD(g) - Fiber, total dietary
    - SugarTot(g) - Total Sugars
    - Calcium\_(mg) - Calcium
    - Iron\_(mg) - Iron
    - Magnesium\_(mg) - Magnesium
    - Phosphorus\_(mg) - Phosphorus
    - Potassium\_(mg) - Potassium
    - Zinc\_(mg) - Zinc
    - Copper\_(mg) - Copper
    - Manganese\_(mg) - Manganese
    - Selenium\_(æg) - Selenium
    - VitC(mg) - Vitamin C, total ascorbic acid
    - Thiamin\_(mg) - Thiamin
    - Riboflavin\_(mg) - Riboflavin
    - Niacin\_(mg) - Niacin
    - PantoAcid(mg) - Pantothenic acid
    - VitB6(mg) - Vitamin B6
    - FolateTot(æg) - Folate, total
    - FolicAcid(æg) - Folic acid
    - FoodFolate(æg) - Folate, food
    - FolateDFE(æg) - Folate, DFE
    - CholineTot (mg) - Choline, total
    - VitB12(æg) - Vitamin B-12
    - VitAIU - Vitamin A, IU
    - VitARAE - Vitamin A, RAE
    - Retinol\_(æg) - Retinol
    - AlphaCarot(æg) - Carotene, alpha
    - BetaCarot(æg) - Carotene, beta
    - BetaCrypt(æg) - Cryptoxanthin, beta
    - Lycopene\_(æg) - Lycopene
    - Lut+Zea\_ (æg) - Lutein + zeaxanthin
    - VitE(mg) - Vitamin E (alpha-tocopherol)
    - VitDæg - Vitamin D (D2 + D3)
    - VitDIU - Vitamin D
    - VitK(æg) - Vitamin K (phylloquinone)
    - FASat(g) - Fatty acids, total saturated
    - FAMono(g) - Fatty acids, total monounsaturated
    - FAPoly(g) - Fatty acids, total polyunsaturated
    - Cholestrl\_(mg) - Cholesterol
    - GmWt\_1 - gram weight 1
    - GmWt\_Desc1 gram weight 1 descriptions
    - GmWt\_2 - gram weight 2
    - GmWt\_Desc2 - gram weight 2 description
* **Star rating of health care plans with HOS-CAHPS measures** - HOS CAHPS survey measures are the base of determining star rating of healthcare plan. Files related to star rating have two types of measures which are used to determine star rating of the healthcare plans - Part C and Part D. Part C is has three type of information 1. Chronic conditions (disease) 2. Tests and Vaccines 3. Member experience with healthcare plans. All variables starting with C01 to C32 are related to part C of the surveys. Similarly Part D of the survey is related to Drugs plans customer services. In data variables starting with D01 to D15 is related to part D. Surveys such as HOS CAHPS etc contains questions whose final standing results into C01 to C32, and D01 to D15 measures. Dataset has two star rating and measurements data released in fall 2015 and Spring 2016. Files description -
  + Starratingfall/spring2015C\_cutoff.csv - Contains information about different cut off used in determining star rating of part C measures.
  + Starratingfall/spring2016D\_cutoff - Contains information about different cut off used in determining star rating of part D measures.
  + Starratingfall/spring\_domain.csv - Contains information about domain rating of plans
  + Starratingfall/springhighperforming\_plans.csv - List of high performing plans
  + Starratingfal/springlowperforming\_plans.csv - List of low performing plans
  + Starratingfall/springmasterdata.csv - Contains information on all the measures of all plans
  + Starratingfall/springplansfinalstarrating.csv - Having information of star rating of healthcare plans
  + **Description** -
    - CONTRACT\_ID - Healthcare plan id
    - Organization Type - Type of the organizer - employer/demo/local cpp etc
    - Contract Name - Name of the contract
    - Organization Marketing Name - Self explanatory
    - Parent Organization - Healthcare provider
    - HD1: Staying Healthy: Screenings, Tests and Vaccines (domain)
      * C01: Breast Cancer Screening
      * C02: Colorectal Cancer Screening
      * C03: Annual Flu Vaccine
      * C04: Improving or Maintaining Physical Health
      * C05: Improving or Maintaining Mental Health
      * C06: Monitoring Physical Activity
      * C07: Adult BMI Assessment
    - HD2: Managing Chronic (Long Term) Conditions (domain)
      * C08: Special Needs Plan (SNP) Care Management
      * C09: Care for Older Adults – Medication Review
      * C10: Care for Older Adults – Functional Status Assessment
      * C11: Care for Older Adults – Pain Assessment
      * C12: Osteoporosis Management in Women who had a Fracture
      * C13: Diabetes Care – Eye Exam
      * C14: Diabetes Care – Kidney Disease Monitoring
      * C15: Diabetes Care – Blood Sugar Controlled
      * C16: Controlling Blood Pressure
      * C17: Rheumatoid Arthritis Management
      * C18: Reducing the Risk of Falling
      * C19: Plan All-Cause Readmissions
    - HD3: Member Experience with Health Plan (domain)
      * C20: Getting Needed Care
      * C21: Getting Appointments and Care Quickly
      * C22: Customer Service
      * C23: Rating of Health Care Quality
      * C24: Rating of Health Plan
      * C25: Care Coordination
    - HD4: Member Complaints and Changes in the Health Plan's Performance (domain)
      * C26: Complaints about the Health Plan
      * C27: Members Choosing to Leave the Plan
      * C28: Beneficiary Access and Performance Problems
      * C29: Health Plan Quality Improvement
    - HD5: Health Plan Customer Service (domain)
      * C30: Plan Makes Timely Decisions about Appeals
      * C31: Reviewing Appeals Decisions
      * C32: Call Center – Foreign Language Interpreter and TTY Availability
    - DD1: Drug Plan Customer Service
      * D01: Call Center – Foreign Language Interpreter and TTY Availability
      * D02: Appeals Auto–Forward
      * D03: Appeals Upheld
    - DD2: Member Complaints and Changes in the Drug Plan’s Performance
      * D04: Complaints about the Drug Plan
      * D05: Members Choosing to Leave the Plan
      * D06: Beneficiary Access and Performance Problems
      * D07: Drug Plan Quality Improvement
      * DD3: Member Experience with the Drug Plan
      * D08: Rating of Drug Plan
      * D09: Getting Needed Prescription Drugs
    - DD4: Drug Safety and Accuracy of Drug Pricing
      * D10: MPF Price Accuracy
      * D11: High Risk Medication
      * D12: Medication Adherence for Diabetes Medications
      * D13: Medication Adherence for Hypertension (RAS antagonists)
      * D14: Medication Adherence for Cholesterol (Statins)
      * D15: MTM Program Completion Rate for CMR
    - SNP - Are they offering special plans
    - Sanction Deduction - If sanction is deducted from last survey to this survey
    - 2016 Part C Summary - 2016 Part C rating
    - 2016 Part D Summary - 2016 Part D rating
    - 2016 Overall - 2016 Overall star rating of the plan
    - Rated-as - Category name
    - Highest rating - category -C/D/Overall for which rating is high
    - Rating - Star rating of the plan