Exemplos de Bases de Dados

GoAmazon Dataset - https://www.arm.gov/campaigns/amf2014goamazon

DataOne dataset https://www.dataone.org/data

Dados do Governo Brasileiro http://dados.gov.br

Portal da Biodiversidade Brasileira https://portaldabiodiversidade.icmbio.gov.br/

USGS data Portal http://data.usgs.gov/

Google public datasets. http://www.google.com/publicdata/directory

Large datasets publicly available. https://www.quora.com/Where-can-I-find-large-datasets-open-to-the-public

Yahoo WebScope http://webscope.sandbox.yahoo.com/

Data.gov: U.S. Government Data - http://www.data.gov/

Freebase - http://www.freebase.com/

Yelp - http://www.yelp.com/developers/documentation

Numerous APIs from Google (e.g., Maps, Freebase, YouTube, etc.) https://developers.google.com/apis-explorer/#p/

Trulia, Zillow: real estate listing sites (http://developer.trulia.com) (http://www.zillow.com/howto/api/APIOverview.htm)

Numerous graph datasets (large and small): SNAP https://snap.stanford.edu/data/index.html

Dados de filmes: Rotten Tomatoes, IMDB - http://developer.rottentomatoes.com/ http://stackoverflow.com/questions/1966503/does-imdb-provide-an-api

The Free 'Big Data' Sources Everyone Should Know - http://www.datasciencecentral.com/profiles/blogs/the-free-big-data-sources-everyone-should-know

Quandl - a dataset search engine for time-series data. http://www.quandl.com/

A collection of links to various datasets. http://aws.amazon.com/datasets/

Microsoft Academic Search http://datamarket.azure.com/dataset/mrc/microsoftacademic

Retrosheet: MLB statistics (Game/Play logs) - http://www.retrosheet.org/

Various geophysical datasets - http://maps.ngdc.noaa.gov/viewers/geophysics/

Social trends http://www.pewsocialtrends.org/category/datasets/

Academic torrents (terabytes) - http://academictorrents.com/

Data Science Initiative - Microsoft Research http://research.microsoft.com/en-us/projects/data-science-initiative/default.aspx

CKAN (Dados Abertos de Governo – http://ckan.org) , \Quandl (Dados Financeiros - https://www.quandl.com), Kaggle datasets. (https://www.kaggle.com/datasets), Microsoft Azure Market Place (http://datamarket.azure.com/browse/data)