

ANLY500 Spring 2018 Course Project

Ask a question and let your data answer it!

There are several open databases available for researches around the world. The main purpose of this project is to make you familiar with how to use a database to answer some questions and resolve issues related to companies. Your task is to develop a research question or identify a challenge from following datasets. Then use the appropriate analysis techniques you've learned throughout the course to answer your question or provide a solution to your challenge.

The datasets are in first come first serve order so choose your desired dataset before it is taken by other teams. You may choose your desired dataset by **February 20, 2018**. If any team does not choose the topic, they will be assigned the remaining datasets.

Topics:

- 1) Bay Area bike trip: <http://www.bayareabikeshare.com/open-data>
- 2) San Francisco Fire Department Service Calls Data (8 groups: Economy, City management, Transportation, Public safety, Health and Social services, Energy and Environment, Housing and Buildings, City Infrastructure, Culture and Recreation) : <https://data.sfgov.org/>
- 3) IRS dataset for year 2015 : <https://www.irs.gov/uac/soi-tax-stats-annual-extract-of-tax-exempt-organization-financial-data>
- 4) NYPD datasets (more than 10 projects) : https://data.cityofnewyork.us/browse?Dataset=Information_Agency=Police+Department+%28NYPD%29

Timeline and deadlines

Date	Task	Notes
February 20	Submit a clearly stated research question and dataset description	Should include • Link to data • General description of data
March 6	Project Proposal due	
March 20	Project status	Prepare a presentation about the steps you have taken
April 10	Project presentations	
April 17	Project Report	Final report on your project

Honor Code

You may not discuss this project in any way with anyone outside your team, beside the professor. Failure to abide by this policy may result in a 0 for all teams involved.