The NSF funds research and education in science and engineering, through grants, contracts, and cooperative agreements. It accounts for about 20 percent of federal support to academic institutions for basic research. The type of activities funded by NSF and the purposes of NSF funding vary. Consequently, the degree of NSF responsibility for and the management control of such activities also varies. Therefore, it is beneficial and interesting to track the NSF award history and explore the change of NSF funding policy over years and also evaluate the performance of research institute/principal investigator on the progress of the NSF awards.

The NSF releases all its grant data in XML which contains information such as award effective period, funding division, award amount and details about the investigator. In this project, data from 2000 to 2016 were obtained from NSF website. A simple parser for NSF XML files was created to retrieve a spread with the details about each grant.

From here, there are several questions can be answered by digging into the data. Some examples may include:

• Which state/institution/principal investigator is most awarded by NSF, regarding number of awards and the amount of funding

• What are the fields or subfields that are more likely to be awarded large about of funding and how does this rule change over the decades

It is then possible to create another parser to parse the grant IDs from XML file of Pubmed, which contains the information of the grant ID. By matching the grant ID with the NSF grant ID, it will enable the track down of the paper either published under the grant or further track the paper cite the paper of interest.