Behavior in banking transactions will be analyzed for changes in language within age groups, gender, and geographical locations in the southeast United States. These changes will be examined against bankruptcy data to determine if there is a correlation between behavior within a bank and future financial success.
Roles and responsibilities

Project PI will be the primary data creator, and manage data produced and stored from the project. All data produced from the project is expected to be openly shared once the project funding is completed via online websites and repositories such as ICPSR.

Expected data

Expected project data includes data tables collected from openly available US Census Data. This data will be analyzed using Excel and R software. Data will also be collected from public available datasets such as ICPSR and NYC Open Data. Collected project data will be approximately 1-2 Terabytes in size.

Period of data retention

Project data will retained throughout the course of the project on the PIs campus data services. Once the project is complete data will be submitted to ICPSR, or Figshare if the data is not accepted by ICPSR. Once project is complete, data will be scrubbed of personally identifiable information.

Data format and dissemination

Project data is expected to be shared with the scientific community via peer-reviewed journal articles, presentations at academic conferences, poster presentations and white papers in a variety of venues. The primary data format for data will be tabular spreadsheet data. Manuscript files, presentation materials, graphic images of summary data, and short explanatory videos describing the project are also expected to be produced.

Data storage and preservation of access

ICPSR is a non-profit organization committed to preserving social science research data. It is located at the University of Michigan and currently houses over 3,000 data sets from all areas of social science research. Figshare is a data sharing site created in 2012 that allows researchers to openly share datasets. It has over 1,000 datasets in all areas of the sciences and social sciences. Project data expected to be contributed to ICPSR include anonymized project data, summary data, an explanatory note(s) for data, and graphics and display items.

Additional possible data management requirements

Campus project data requirements include classifying final project data. This action will be done, as well as ensure storage of final project data is commensurate with the appropriate data classification.