Prototype business use cases

We have provided a bulleted list of the steps taken in arriving at the 4 data sets that we utilized in our prototype.

* Our first step was in selection of USDA data from the data.gov website as this is the lighthouse agency that is being used by the GSA in their creation of a Center of Excellence for Data Visualization and Analytics.
* Our second step was to utilize data sets that would be of interest to current USDA users and demonstrate the value of our analysis even though this was just a prototype. We focused on the USDA Farm Service Agency (FSA) because of our current experience in supporting the design and development of the Enterprise Data Warehouse.
* The 4 biggest areas of analytics at USDA FSA are Crop Acreage, Commodities, Payments and Programs. Initial analysis of the 500+ data sets showed that there was not enough Commodities data for us to do proper analytics; we therefore focused on the other 3 areas. These 3 areas follow an annual calendar cycle where Crop Acreage is a focus from May 15-July 15, payments are a focus from August 1 to October 1 and Programs from October 1 to January 1.
  + **Crop Acreage** - From the start of the fiscal year on October 1 to July 15, farms must report on the number and types of crops being grown on their farm. As we approach the final two months of this period, the FSA executives at the state and county level focus on the reporting trends and on reporting compliance to ensure that all farms in their areas of jurisdiction are compliant. Analytics and visualization of this data is the focus at all levels of the FSA during this period as the data from Crop Acreage Reporting is the seed to a number of other FSA activities.
  + **Payments** - From August 1 to October 1, the data generated from the Crop Acreage Reporting and other programs (Farm Loans, Conservation, etc.) are used to generate payments from the FSA to farms. These payments form the basis of another set of data analytics and visualization as to the types of payments made and to analysis of the payments by demographic criteria and location criteria.
  + **Programs** - Data from the payments is used to perform analytics on each of the programs and measure their success across demographics and location. This allows funding for the next fiscal year to be allocated based on the success of each program. The data also allows us to see which demographic groups who are eligible for programs may be underrepresented and allows FSA to develop outreach programs for them.
* Using the prototype tool we conducted further detail analysis on data sets in the 3 areas listed above. We used the R tool to correlate various sets of data and develop hypotheses on what the data was showing us. We verified our findings with USDA personnel at FSA to gauge their interest in learning more about what knowledge could be gained from this information. The data analytics that showed the greatest promise have been utilized in our prototype.