### DATA PREPARATION

- The dataset ‘federalfunds.txt’ contains 2,666,153 rows, with each row describing one dollar amount allocation event (grant) from one CFDA program in one zip code in one year.
- The data spans years 2005-2009 with similar number of records in each year.
- The dataset shows program funds allocated across 335 unique CFDA codes.
- The dataset shows program funds allocated across 31,517 unique zip codes.

- The data set ‘federalfunds.txt’ contains a sample of 2,390 CFDA programs. The official list of CFDA programs can be found at the CFDA.gov website which shows 3,210 unique CFDA codes. We named the dataset ‘CFDAPrograms’.
- Based on the official list of 2,390 CFDA programs, there are 150 Agencies managing between 102 and 1 CFDA programs.
- Our analysis of agency descriptions in the dataset ‘CFDAPrograms’ has revealed that there are 45 clearly identifiable departments overseeing between 1 and 28 agencies, and between 1 and 514 CFDA programs. Departments are identifiable by the first two digits of the CFDA program.
- We have developed an algorithm that makes departments identifiable, through a department code, for most individual CFDA programs. Based on the first two digits of the CFDA program.
- After this analysis, 34 CFDA programs for the official ‘federalfunds.txt’ dataset remained unattached with a department, accounting for less than 2% of the programs listed in the ‘CFDAPrograms’ dataset.

- We worked to determine how many unique CFDA codes could be matched to an identifiable federal department.
- Using our department matching algorithm, we were able to match 7,663 CFDA codes with one of the 45 identifiable departments, which amounts to 85% of all CFDA codes appearing in the data set ‘federalfunds.txt’.
- The remaining 15% of unmatched CFDA codes amount to 3% of all grant spending.
- Finally, we attached the department code to the department that is most identifiable.

- Suggested Categories for the 1306 CFDA numbers are not necessarily the best fit for the analysis.
- One CFDA program code does not map uniquely to a category code does not map uniquely to.
- Therefore, we choose to use Department based classifications for our analysis rather than category classification. For most departments there is a predominant category but for all (Department of Agriculture, for example).

### EXPLORATORY DATA ANALYSIS

- Preparing the final table for analysis:
  - Initially, we reduced the size of federalfunds by focusing on only the columns that we needed (AUDIT YEAR, STATE, CFDA AMOUNT) and dropping by AUDIT YEAR, STATE, CFDA to create table ‘stateAudits’.
  - In order to get the department code to join to the we created a table, ‘SAI’, that joined the table with department code with SAI by the CFDA numbers.
  - Finally, we joined SAI with the given state information giving us our table ready for analysis.

### EDUCATION DESCRIPTIVE STATISTICS

- For each state and for each year, 2005-2009, we analyzed Per Capita Income and Spending on Education per Capita.

- In 2005 the highest per capita income was CT with $50,226 (per capita spending was $608.63).
- The lowest was MS with $22,675 (per capita spending was $689.20).
- The average was $34,667.
- In 2005 the highest educational funding per capita was SD that had $5,060 (per capita income was $33,772).
- The lowest was NV with $29,801 (per capita income was $38,637).
- The average was $35,448.

- In 2006 the highest per capita income was CT with $54,191 (per capita spending was $533.75) and the lowest was MS with $27,711 (per capita spending was 1,071.41).
- The average was $36,021.
- In 2006 the highest educational funding per capita was SD that had $28,861.18 (per capita income was $35,205).
- The lowest was NV with $20,068 (per capita income was $38,637).
- The average was $29,960.

- In 2007 the highest per capita income was CT with $54,191 (per capita spending was $549.59) and the lowest was MS with $29,237 (per capita spending was 1,010.16).
- The average was $38,567.
- In 2007 the highest educational funding per capita was SD that had $5,375.97 (per capita income was $38,729).
- The lowest was NV with $19,884.65 (per capita income was $45,159).
- The average was $29,960.

- In 2008 the highest per capita income was CT with $61,322 (per capita spending was $457.05) and the lowest was MS with $30,563 (per capita spending was $953.04).
- The average was $40,138.
- In 2008 the highest educational funding per capita was SD that had $6,315.82 (per capita income was $41,311).
- The lowest was NV with $20,10.22 (per capita income was $30,449).
- The average was $1,312.20.

- In 2009 the highest per capita income was CT with $61,042 (per capita spending was $391.77) and the lowest was MS with $35,504 (per capita spending was $893.21).
- In 2009 the highest educational funding per capita was SD that had $5,513.74 (per capita income was $39,524) and the lowest was NV with $35,353.89 (per capita income was $41,415).
- The average was $1,354.43.

### CLUSTER ANALYSIS

- The data does not reveal an obvious causal relationship as to how states actually received funding, which required us to dig deeper.
- Cluster analysis was chosen to be the model of choice because it would allow us to see the variation between the states income and the variation between the state’s educational funding.
- A cluster analysis was performed separately for each year, 2005-2009.
- The clusters were formed based on the following dimensions: Per Capita Personal Income and Educational Spending Per Capita.

### ZIP CODE LEVEL ANALYSIS

- Similar to the state level analysis we examined zip codes based upon income and funding allocated to education.
- We again ran a cluster analysis on the data from 2005 to see if the zip code data would follow the trend that appeared on the state level where higher than average incomes received more funding than did lower income states.
- Due to the variation in the incomes and funding per zip code outliers needed to be removed and therefore any zip code that had an extraordinarily high per capita Adjusted Gross Income (AGI) funds were removed, as were the zip codes with funding outside the normal range.
- This left us with approximately 6,600 aggregated records of yearly spending data by the on CFDA programs offered by the Department of Education per zip code.

- This cluster shows the following trend:
  - Those with the greatest AGIs have the most funding allocated to their educational programs as opposed to those zip codes with the lesser AGIs.

### RECOMMENDATIONS

- Increase educational funding for states/municipalities with below average Personal Income/AGI:
  - In our statistical analysis of the data we found that those states with higher than average incomes are receiving the largest amounts of educational funding. We recommend that this be changed to the reverse where those states that have lower than average incomes receive more educational funding.
  - The states that were continuously low in education only have about 5% of their funding allocated to education, whereas as states such as South Dakota, who has the highest per capita education spending had 12% on average.
  - In addition to this education is competing with health and human services for the largest amount of funding.
  - Creating a better category system with machine learning algorithms.