**Appendix**

**Column Choice:**

* ‘ViolentCrime1’ and ‘PropertyCrime1’ were created to replace missing values for these variables with the median values for the state. For example, if Dallas had a missing value for Violent Crime, we would take the median value of Violent Crime for the state of Texas, and use that as the data point for Dallas. ‘TotalCrime’ is the new column, equaling the sum of ‘PropertyCrime1’ and ‘ViolentCrime1’ for a city. This effectively gives us the total crimes per 100,000 citizens. We then used ‘TotalCrime’ to determine the overall safety of each city.
* To determine the urbanization of a school, we used the ‘Locale’ column. We used codes 11 (Large city), 12 (medium city) and 21 (suburb of a large city) that we thought would satisfy Maria’s requirements. Column ‘Locale2’ supposedly offered additional information on urbanization, was left blank in our data set, thus it was unusable.
* For program choice, we looked at 10 columns: each starting with ‘CIP11’ or ‘CIP52’ and ending with either ‘CERT1’, ‘CERT2’, ‘ASSOC’, ‘CERT4’, or ‘BACHL’. ‘CIP11’ refers to an Information Technology or Computer Science degree, and ‘CIP52’ refers to Management of Information Systems degree. The suffixes of the columns indicate Certificate of less than one academic year-CERT1,Certificate of at least one but less than two academic years-CERT2,Associate degree-ASSOC,Bachelor's degree-BACHL
* Our final dataset features 43 schools that made the cut based on Maria’s requirements. Our final dataset has columns for school name, city and state name, tuition, retention, public/private ownership, and potential earnings. Please note that we made visualizations based on the final, ranked dataset and all three original files (after being cleaned).

**Visualization explanations**

* Schools that fit Maria’s requirements and rank highly on her preferences

Displays the overall score given to each score based on how we scored each of her preferences. The max score of the preferences was 10, only obtained by UCLA. The drop-off of scores between UCLA and the next tier of schools is larger than the drop-off from these schools and those below them. UCLA is not only the highest scoring, but significantly above its competition, leading to a confident selection for Maria.

* State crime rate

Different colors designate between the states, whereas the size of the circles display the average total crime rate within that state. The data is filtered on the states that meet all of the requirements Maria gave us, reflecting our list of proposed schools that only included those from Colorado and California.

* Total crime and Locale by City

Different colors designate between the type of locale, with this specific graphic only showing the chosen colleges’ cities – with large cities and suburbs of large cities. Size of the circles displays the average total crime rate within those cities. This is a drilled-down look within California, the state with the lower crime rate, in an attempt to find the most suitable city for Maria’s school.

* Institutions in Los Angeles

Upon seeing that Los Angeles provided the safest option with the lowest crime rate, this tree map shows the relationship between potential earnings and tuition for schools located in LA. Each school is also denoted with 1, 2, or 3 – whether the school is private (for profit/non-profit) and public. The data shows a lighter color relative to the size of the rectangle for UCLA, denoting that it is a high value school when considering return on investment of tuition.

* University of California-Los Angeles School with programs offered

Within what was deemed to be the highest value, this graphic verifies the offering of the specific programs Maria was seeking in her school – either an IT or MIS degree

* Kauffmann Metro Area Rankings for Entrepreneurship 2014

To find the cities with the most consistently high entrepreneurship rankings, the 2014 data is displayed with the size of the circle being highest ranking and 2015 data with the darker being the highest ranking. This map shows that the most desired cities include Los Angeles, Denver, Miami, etc. This verifies the pursuit into LA or Denver for our school suggestions since the cities are consistently among the highest ranked for entrepreneurs.

* Average Startup Density and Average rate of new entrepreneurs

Similar to the previous map, this map displays the average rank of the two years of Entrepreneurship data. But this map further shows the relationship between the entrepreneurship data and Startup density data, with the intention of finding cities that are highly ranking in startup density as well. Note that the three Californian cities and Denver that we chose are all highly ranking in each of these metrics.

* Full and Part-Time Retention rates

Within the desired cities detailed previously, this treemap displays the schools with the highest retention of students – part-time vs. full-time. Several schools have equal/highly similar full-time retention numbers, including our top choice UCLA. It is set apart by also being highly ranked in part-time retention rate as well.

* Tuition for Public, Private (for profit), Private (non-profit) institutions

The three colors define the institution type, with red representing the public institutions, Maria’s preference. This eliminates several schools from consideration, notably leaving UCLA as an option. The size of the rectangle in the treemap is defined by the size of the rectangle. Though UCLA is among the highest of the public institutions, it is on the lower end of the spectrum when considering all of the schools measured in the chosen cities.

* Average earnings vs average tuition

In a continued attempt to find schools with the highest value in regards to tuition and potential earnings, this bar chart gave us our last four schools to choose and displays UCLA as the second highest in potential earnings and among the lowest in average tuition, signifying its high value financially.