**Executive Summary**

The objective for our project was to find the best school and program for Maria based on her requirements and preferences in where she wants to attend. We narrowed down a list of schools based on her requirements, then chose from that list using rankings based on her preferences. We wanted to give Maria at least 10 options to choose from that would allow her to be happy with her choice of school and successful in her future.

The 3 raw data files we were given to work with were quite messy. The crime data had missing values for a couple cities, including the names of cities themselves. When we encountered missing data for crime, we would replace it with the median value of that crime type for that 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. For missing city names, we simply used the metro area’s name. Again, this issue was not very significant and this step only had to be done a few times, and half of the crime data was filtered out anyway due to Maria’s safety instructions.

On the other hand, the College Scorecard had lots of missing data that could have been useful to us. For example, percentile earnings potential for the schools was largely blank. For Maria’s earnings potential preference, we had to use mean earnings, which is a less robust statistic and is more vulnerable to outliers.

There was one point of confusion in the Metro data. Maria indicated she wanted the top 25% of best entrepreneurial cities so that she could do an internship at a start up during school. A z-score of .6745 denotes the 75th percentile (cut off for the top 25%) in a normal population. However, there are cities in the Metro data file that are in the top 25% of the cities listed (10 out of 40) based on ranking, but have a z score lower than .6745 (San Diego in 2014 and 2015 and San Antonio is 2015). To give Maria more options, we decided to still take these cities as they were in the top 25% of the 40. Additionally, the Metro data did not have information on startup density after 2013, did not have Opportunity Share data until 2009, and did not have new entrepreneur rates until 2007. Because the missing information was uniformly absent for all cities, we did not feel the need to clean this at all, and felt comfortable using the raw data. Lastly for Metro data, Tableau could not automatically locate 9 cities. We had to manually add the latitude and longitude of these cities to accurately show entrepreneurial rankings across the country.

Maria did not specify what degree she wanted to pursue, so to give her the most options we expanded our search for all degrees, from a 1 year degree to a 4 year Bachelor’s degree. Based on this, she has 40 schools to choose from that offer a degree in MIS, and 30 schools that offer a computer science degree. This gives her options in what major and degree she chooses. We also saw that on the final data set, only four cities made the cut based on her requirements.

Comparing the tuition fee and the earnings after the completion of the program against each other, we find that Los Angeles Trade Technical College has the lowest average tuition fee among all the colleges shortlisted for Maria. Average earnings of the student after completion of the course is highest for University of Southern California. Overall we see a trend higher is the tuition fee of the college higher is the average earnings of the student.

Given the retention rate is the crucial criteria for Maria while selecting a college, Virginia Sewing Machine and Learning Center has the highest average full time retention rate, University of San Diego has the highest part time retention rate. California College San Diego, has the least full time and part time retention rates.

After all of our work and research, we conclude that Maria will be happy and successful in choosing to study either Computer Science or MIS at the University of California- Los Angeles. Los Angeles offers safety through low crime rates and opportunity through a high number of start ups in a big city setting. UCLA will offer lots of academic support for Maria in striving to have her finish her degree there. UCLA is a highly respected, accredited, public school and is an affordable option for Maria, with high potential earnings for her upon graduation.