**Project 1**

**Prepping For The Benjamins**

Team Members: Elizabeth Duenas, John Hawkins, Gustavo Mendoza

Our focus was led by the primary question of, “Does getting a college degree mean you will earn more money?” We all agreed that we had the preconceived notion that for the most part, we all go to school to earn more money or get a better job. If asked with the question during a conversation, “Prove your point?” we wanted to let the data speak for itself.

Based on the data obtained from the Census, we were able to create heat maps for poverty, no high school diploma, high school diploma and college degree. Looking at the poverty throughout the US, we can observe that there is a higher concentration around some major cities like Dallas, Los Angeles, Seattle, some around the Texas border and along the Mississippi River. We also see that West Virginia and Kentucky are the heaviest in poverty out of all. When comparing people with high school diploma vs no high school diploma we noticed a big change when it comes to obtaining a diploma. We see that there is a high rate of students obtaining a high school diploma in the northeast portion of the United States. This includes states such as New York, Pennsylvania, Illinois and Iowa. California was a stagnate especially in Los Angeles. Finally looking at the college degrees, we can see that there is a higher rate of people in bigger cities like Houston, Los Angeles and Miami. Comparing the rate of students obtaining high school rate to college you can really see that most students migrate to bigger cities. Looking at the income per capita, we can see that there is more money within bigger city like we see in Houston, Los Angeles, and Seattle. What was odd was that Kentucky and West Virginia were still generating money.

Digging deeper into the data, for all 50 states we wanted to see the composition makeup for education level. The education levels were broken into these three categories: No high school diploma, High School Diploma and College Degrees. Using a stacked bar chart, the total makeup of the US states did not demonstrate any jarring numbers. The graph appeared uniformed with an even makeup of the various education levels.

Using the scatterplot and calculating R value we compared poverty rate with those not having a high school diploma. This scatterplot showed a strong uphill (positive) linear relationship. The Rvalue was calculated at .7805. The following slide demonstrates how both per capita and poverty rates when compared to education levels follow a normal distribution curve. Most of the data is tightly clustered around the mean. The mean for no high school diploma is 11.10 and 13.70 for poverty rate. We would need more data in order to be able to make a conclusion because as of right now we cannot. All that we can say is that our R value explains 60% of variation in our data model. Viewing the second graph showing the distribution curve we can see that the max and min values are more clearly visible. For per capita rate we have a max of 50832.0 and min of 22500.0. Given that P values are 1.4 e^-11 and 3.9 e^18 when R value is high and P value is low, means that your model explains a lot of variation within the data and is significant.

While the data reflects that higher levels of educational attainment will result in higher income, there is further work to do in this area to determine if this is indeed the only predictor of higher income. Among our group there is anecdotal evidence that within certain industries educational attainment does not equate to income. One such example is the Oil and Gas industry, we decided to take a quick look and see if we could make a case for this. Looking at data for Midland County, Texas the per capita income is $38,545, the average income for all other counties in the US is $26,039. Without further data available one assumption is that industry plays a large role in determining what an individual’s level of income will be. Another factor that was not included in our analysis due to a lack of available data is what part the cost of attaining higher education plays in effectively reducing income, i.e. student loan debt. Current estimates are that the average outstanding student loan in the US is $37,000. It would be interesting to study this further.

Although we did not find any definitive proof that getting a higher education means more money, we did see interesting trends we would like to explore further. For the areas with high education and high income, how debt burdened are those people? How do trade jobs, if mapped for the US compare to the areas with high college degrees? There is more that is left unanswered and should leave us all second guessing if going to college really is the best way to make more money. Our data definitely did not show that college is always the winner.