**Applied Business Intelligence Final Report**

Ian Listopad

**Data Sourcing:**

I Chose to do research and ultimately answer questions about population migration within the United States. This includes things like where are people moving to, when did the most people collectively move, what year had the lowest amount of people moving. To answer these questions, I employed the help of the US Census Bureau. That had a dataset that showed all the population related census information for every census year since 1910.

**Data Preparation and Integration:**

For my Data prep I chose to use Python, simply because it is what I am most comfortable with. I have used it in my professional life for things other than data analytics, so it is what I feel is the easiest to use.

For this I ended up dropping some of the columns that I did not feel were necessary for my questions, such as the columns that dealt with representatives. I Chose not to drop the null lines as they all dealt with region, which is something I will want to look at in my PowerBI models. Fortunately, this was mostly a clean dataset, seeing as it was prepared and published by the Census Bureau rather than say real time data from customers.

**Data Analysis:**

For my analysis I took a hybrid approach using both Python and Power BI. I used Python for more of the individual trends like overall change by year and percent change between 1910 and 2020. Python is better suited for this because by default it doesn’t do anything extra to the data whereas Power BI does summarization and is ultimately suited better for visualizations of overall data in my case.

I did use Power BI for just that. Power BI helped me see the bigger overall picture in terms of what states had the most sum of percent change. I also used a new type of chart, a Sankey chart. This allows me to filter by year and see where the population of the total United States is going via line weights. This is a good representation of where exactly the residents of the United States are living. Along with that we can also use Power BI to easily see and filter on trendlines as well in a more user-friendly way.

Using Python and Power BI we can see that people are moving throughout the country in such a way that they are moving to the south and west. Along with this we see that there is a good mix of people moving away and too bigger states. This all mostly stays true for both time periods, 1910-2020 and even later like 1980-2020 or 2000-2020. We also see a change in political populations. For 2000-2020 we see that people are overwhelmingly moving to states that are more conservative or middle learning such as Utah, Idaho, and Texas. This could be a shift in the political climate mixed with things like weather climate.

**Key Findings:**

Using Python we can see that by far the largest state in terms of population is California, which is not a major surprise. We also see that Wyoming, although growing steadily is the smallest state in terms of population.

Along with python I have employed Power BI to create a dashboard. From this we get some interesting information. Like for example using the Pie chart we can see that Nevada has had the largest sum of percent change in terms of population, this translates to Nevada having the largest overall Percentage change in population with a total of 3691.89%. In last overall place is North Dakota with a whopping 35% change from 1910 to 2020.

Now we also see from Python that the West Region was the most growing Region in the Unite states between 1910 and 2020. Which makes sense because 4 of the top 5 states with the overall largest growth were from the West (Nevada, California, Arizona, and Alaska). Some other interesting information that we see is that in 2020 Utah had the highest percent change. The next highest region with regards to total change in population is the South. This goes to show that people are moving out of the Midwest and Northeast regions at a faster pace than they are moving in. This could be due to a few factors. Climate is the big one that comes to mind first, people are moving to warmer places. We also see that when we change our timeframe to 2000-2020, we can see that people are overwhelmingly slowing rates from liberal states to more conservative states. For example, for that timeframe Utah, Idaho, Texas, and Arizona, and Florida were the 5 of the 6 fastest growing states. With Nevada being number one and seemingly being a swing state for the presidential elections from 2000-2020. This all could be a shift in the political climate in the United States, but it could also be a culture shift. Take the fact that California only grew 16% in the same timeframe, nearly 40 points beneath Nevada.

Overall People are moving into the South and West at a faster rate than the Midwest and Northeast and people are moving from Liberal states to more conservative or middle learning states.