The Data Set

Our data set is a combination of the 2000 and the 2010 US census data, and we focused on the Maryland-specific statistics. Maryland has 24 counties and Baltimore city, which is treated as a county equivalent for census purposes. Each county has 10 main reported attributes, and over 70 if sub-attributes are included. The main values are total population, population of one race, white population, black or African population, American Indian population, Asian population, Native Hawaiian population, other single race population, population of two or more races, and finally the Hispanic/Latino population of any race. Sub-attributes include values for all possible combinations of races.

We also used the diversity indices for each US state and each Maryland County from the 2010 census data. The diversity index is calculated as the percentage of time two randomly picked people from a geographical area would fall into two different racial groups.

All the relevant tables can be found at the FactFinder webpage of the US Census Bureau (http://factfinder2.census.gov/main.html).

The Tool

We used Spotfire for all our visual analysis. Spotfire is a useful tool but we felt it could be improved in certain areas. The first suggestion we have is a feature to edit and combine data tables with the same set of columns after they have been separately imported into Spotfire. We especially needed this to compute the difference between the 2010 census results and the 2000 census results. Since we could not do this through Spotfire, we did it manually using formulas in Microsoft Excel.

Another suggestion we have is to look at decreasing the time Spotfire takes to start up. It seems that Spotfire has to load the data found in the Excel files every time it starts up but some Excel files we used were not very big and the time it took to load those files seemed to be longer than it should be.

We also felt that Spotfire should have been able to add labels to interactive shape maps in ESRI format. People unfamiliar with the geography of a region will need to interact with the map to know the names of the regions, which makes this map unsuitable for presentation.

One final suggestion we have is to extend the functionality of the filter window. The way it works now for string values is that we can either choose to include all the entries for a data table column or pick which entries to include one by one. We think it would be useful to have a feature that specifies including all the entries of a column except one. While working on this project we had a column for the geographical areas in Maryland and we wanted to include all the counties but remove the state of Maryland, which was just a summation of all the data found in the colonies. To do this we needed to individually select all the counties in Maryland in the filter window.
The Headlines

I. Baltimore City Population Has Been on a Steady Decline for the Past Two Decades

Baltimore has continued its total population decline that goes back to 1990. The changes in the census from 1990 to 2000 show that the total population of the city of Baltimore decreased by a staggering 11.5% and our visualization of the changes in the census from 2000 to 2010, shown below, reveal that the total population of the city of Baltimore dropped another 4.6%.

The above visualization is a treemap of Maryland census data where the hierarchy is determined by geographical area, the color is determined the total population change (in percent), and the size is determined by the total population. The treemap clearly shows in red that the city of Baltimore is the only geographical area of Maryland that had a percentage decline in total population from 2000 to 2010.

This leads us to conclude that despite being an industrial city, Baltimore has seen a lot of its population leave the city in recent years, mainly due to the enormous decline of the white population in the area and the lack of the Hispanic/Latino population moving in. We speculate that this could be because of the crime rate that is higher than the national average or the recent negative decline in job growth.

II. Hispanics Drive Maryland Population Increase by Doubling State-wide and Nearly Quadrupling in Some Counties

An overwhelming result we found while analyzing the census data is that total population changes that were positive were almost entirely driven by Hispanic/Latinos moving into Maryland. The visualization shown below is a scatter plot where the x-axis is the change in the population of Hispanics/Latinos, the y-axis is the change in total population, and each circle represents a county in Maryland. The size of the circle is determined by the total population in the county.
The visualization above presented interesting insights upon closer inspection. There are three dear outliers in the data presented above: Baltimore City, Montgomery County, and Prince George’s County. Montgomery County observed a huge increase in total population and more than half of that was the result of Hispanics/Latinos moving into the area. Despite the fact that Montgomery Country seems to lie far away from the other counties in the scatter plot it still fits the linear trend that counties with Hispanic/Latino people moving in saw increases in their total population.

The more interesting outliers are Prince George’s County and Baltimore City. Prince George’s County saw a huge increase in Hispanics/Latinos moving into the area but the total population only changed by a smaller amount due to the staggering decline of the white population in the area from 2000 to 2010. Baltimore City saw a much smaller increase in Hispanics/Latinos moving in but a significant decline in its white population so it is the only county that does not fit the trend of positive population change when observing Hispanics/Latino people moving into the area.

On further investigation, our research discovered some speculations that the decline of the white population in Maryland was due to a national trend of the older white population aging and settling in non-industrial areas, while the under-18 population is increasingly coming from a mixed race background. Our visual analysis revealed that this is not necessarily true. The visualization below shows just how the decrease in white population plotted against the increase in mixed race population in every county. The decline in white population is more dramatic for Prince George’s County and Baltimore City, but the increase of the mixed race population does not nearly balance this decline.
One interesting anomaly with the data for changes in the Hispanic/Latino population is the effect it had on certain counties. Big counties like Montgomery County saw big increases in population but relative to the size of the county the changes were more modest. The visualization below is a treemap where the hierarchy is determined by geographical area, the color is determined by the Hispanic/Latino population change (in percent), and the size is determined by the total population of Hispanics/Latinos given by the 2010 census data. The treemap below clearly shows that Frederick County saw a four-fold increase in the Hispanic/Latino population from ten years ago!
Maryland is the Fourth Most Diverse State in the 2010 Census, and Montgomery County is its Most Diverse County

The last important insight we gained from the 2010 Census data is information on the diversity indices. The results below show that the most diverse state, by far, is Hawaii but that Maryland is the fourth most diverse state behind Hawaii, Texas, and Nevada. Also, Montgomery County is the most diverse county in Maryland.
On a final note, despite the high diversity indices of some counties, and the total increase in the multiracial population in the state, the multiracial population only accounts for 2.9% of the total population of Maryland. It is, however, an increase from the 2000 census, when it only accounted for 2.0% of the total.