Chase Rawson

DATA 510.1OL - Introduction to Data Analytics

December 1, 2022

HDMA Washington State Home Loans, 2016

Finance Industry

The following data set I chose contains a total of 4,999 observations of mortgages compared with demographic information of the purchaser type, loan purpose, and enforcing agency name. My data set comes from the Consumer Financial Protection Bureau, which is “dedicated to making sure you are treated fairly by banks, lenders, and other financial institutions.” (“On your side through life’s financial moments,” n.d.) Also, the CFPB aims to,

Make consumer financial markets work for consumers, responsible providers, and the economy as a whole. We protect consumers from unfair, deceptive, or abusive practices and take action against companies that break the law. We arm people with the information, steps, and tools that they need to make smart financial decisions. (“The Bureau,” n.d.)

As the Home Mortgage Disclosure Act (HMDA) requires, this agency must provide data to show housing market issues. It allows public officials to make specific decisions based on the data. All these observations listed were compiled from the Consumer Financial Protection Bureau’s statistics on home loans from the state of Washington. Our mission statement, as a third-party watchdog group, is to ensure transparency and determine whether the Consumer Financial Protection Bureau and other government agencies are true to their word on assisting lenders with clarity, providing the organization ensures fair and equal treatment between financial institutions and lenders. Brent Ambrose, the Faculty Chair in Real Estate at Penn State’s Smeal College of Business, states, “It’s important that policy makers realize that there is a role for further scrutiny in the mortgage industry to try to bring more transparency to it.” (Auman-Bauer & Ward, 2020).

I chose this data set because I want to discover different demographics of the lenders, such as race or sex, to determine any improvements in mortgage policy regarding equality in Washington. This dataset provides enough evidence to compare the loan and lender relationships, using a sample of 4,999 different lenders in 2016. With this dataset, I believe that all the answers I am looking for come with other KPIs that can be created from this data.

**Application Approval Rate By Race KPI**

* KPI from a customer perspective.
* I chose this KPI because I want to compare trends in application approval rates throughout Washington by the applicant’s race. I am looking to discover how many loans were approved by each race out of the 4,999 loans listed in the data set.
* This KPI is being measured to determine if there are any inequalities when it comes to a specific race being qualified or unqualified for a loan.
* This measure is based on the percentage of home buyers qualified for a loan for each race.
* The target I hope to achieve with my data is to have the approval rating for each race as equal as possible.
* All raw data will come from Kaggle from the “HDMA Washington State Home Loans, 2016” dataset. In the dataset, I will use the column titled applicant\_race\_name\_1 and the column titled denial\_reason\_name\_1 to determine the approval rate for each race. The results will be shown on a stacked bar graph, with each bar showing each race’s approval and unapproved rates.
* This KPI will relate to the firm’s goal by whether or not the CFPB is justified in protecting their clients from financial institutions. Differentiating percentages of approval ratings by each race will determine whether or not the CFPB is true to its mission statement.

**Application Approval Rate By Sex KPI**

* KPI from a customer perspective.
* I chose this KPI because I want to compare trends in application approval rates throughout Washington by the applicant’s sex. I am looking to discover how many loans were approved by each sex out of a sample of 4,999 loans listed in the data set.
* This KPI is being measured to determine if there are any inequalities regarding each sex being qualified or unqualified for a loan.
* This measure is based on the percentage of home buyers qualified for a loan for each sex.
* All raw data will come from Kaggle from the “HDMA Washington State Home Loans, 2016” dataset. In the dataset, I will use the column titled applicant\_sex\_name and the column titled denial\_reason\_name\_1 to determine the approval rate for each sex. The results will be shown on a stacked bar graph, with each bar showing each sex’s approved and unapproved rates.
* This KPI will relate to the firm’s goal by whether or not the CFPB is justified in protecting their clients from financial institutions. Differentiating percentages of approval ratings by each sex will determine whether or not the CFPB is true to its mission statement.

**Purchaser Type by Race KPI**

* KPI from a customer perspective.
* I chose this KPI because I want to compare trends in the purchaser type of the loan by the applicant’s race. I am looking to discover how many applicants of each race chose Fannie Mae, Freddie Mac, Ginnie Mae, or another type of loan purchaser.
* This KPI is being measured to determine which financial institution the applicant uses to finance their home by each race.
* This measure is based on the percentage of each financial institution involved in the home-buying process in the state of Washington from each race.
* All raw data will come from Kaggle from the “HDMA Washington State Home Loans, 2016” dataset. In the dataset, I will use the columns titled purchaser\_type\_name and applicant\_race\_name\_1.
* This KPI will relate to the firm’s goal of protecting consumers against discriminatory practices against financial institutions. One way to protect home buyers is by ensuring fair competition in the state. Institutions controlling a majority of state loans for a specific race will show that the CFPB is dishonest with its customers.

**Purchaser Type by Sex KPI**

* KPI from a customer perspective.
* I chose this KPI because I want to compare trends in the purchaser type of the loan by the applicant’s sex. I am looking to discover how many applicants of each sex chose Fannie Mae, Freddie Mac, Ginnie Mae, or another type of loan purchaser.
* This KPI is being measured to determine which financial institution the applicant uses to finance their home by each sex.
* This measure is based on the percentage of each state of Washington financial institution involved in the home-buying process from each sex.
* All raw data will come from Kaggle from the “HDMA Washington State Home Loans, 2016” dataset. In the dataset, I will use the column titled purchaser\_type\_name and applicant\_sex\_name.
* This KPI will relate to the firm’s goal of protecting consumers against discriminatory practices against financial institutions. One way to protect home buyers is by ensuring fair competition in the state. Institutions controlling a majority of state loans for a specific sex will show that the CFPB is dishonest with its customers.

**Loan Purpose by Race KPI**

* KPI from a customer perspective.
* I chose this KPI to determine why someone wants to file a loan application, whether it be purchasing a house, refinancing, or home improvement. Using these choices, I am looking to discover how many applicants of each race chose their loan purpose.
* This KPI is being measured to determine the demographics of those purchasing homes, refinancing, or home improvement.
* This measure is based on the percentage of each loan purpose by applicant’s race.
* All raw data will come from Kaggle from the “HDMA Washington State Home Loans, 2016” dataset. In the dataset, I will use the columns titled loan\_purpose\_name and applicant\_race\_name\_1.
* This KPI will relate to the firm’s goal of providing financial tips to loan applicants. Depending on interest rates, it would be responsible for the CFPB to assist applicants with navigating “their financial choices and shop for the deal that works best for them.” (“The Bureau”, n.d.). Although the CFPB is not a loan servicer like Fannie Mae, it must also ensure that financial institutions provide fair treatment to their clients without discrimination by race. Without the CFPB assistance, not as many loan applicants want to refinance their homes.

**Loan Purpose by Sex KPI**

* KPI from a customer perspective.
* I chose this KPI to determine why someone wants to file a loan application, whether it be purchasing a house, refinancing, or home improvement. Using these factors, I am looking to discover how many applicants of each sex chose each factor.
* This KPI is being measured to determine the demographics of those purchasing homes, refinancing, or home improvement.
* This measure is based on the percentage of each loan purpose by applicant’s sex.
* All raw data will come from Kaggle from the “HDMA Washington State Home Loans, 2016” dataset. In the dataset, I will use the column titled loan\_purpose\_name and applicant\_sex\_name.
* This KPI will relate to the firm’s goal of providing financial tips to loan applicants. Depending on interest rates, it would be responsible for the CFPB to assist applicants with navigating “their financial choices and shop for the deal that works best for them.” (“The Bureau”, n.d.). Although the CFPB is not a loan servicer like Fannie Mae, they must also ensure that financial institutions provide fair treatment to their clients without discrimination by sex. Without the CFPB assistance, not as many loan applicants want to refinance their homes.

As shown on the dashboard, each KPI is represented by a visual. Using a random sample of 4,999 mortgages out of 466,000, we can see that there is equality between sex! Although there is a vast difference in the sample size (3008 are male, 1113 are female applicants), there is substantial similarity in the agency that oversees their mortgage process, the purpose of why they applied for a loan, and the lender they chose to finance their loan. It should be noted that these KPIs should not depend on the demographics of Washington State; we are only trying to determine if there are any discrepancies in the data provided and determine any equal opportunities for the applicants from that standpoint. Therefore, I could consider each of these KPIs to be on target, as each agency, including the CFPB, ensures fair and equal treatment between financial institutions and lenders.

We also see that equality between races is different. For example, as seen in the Loan Purpose by Race KPI, there is a lower than average percentage of those Asian, Native Hawaiian, and African Americans using the funds for home improvement purposes, lower than 5%. Also, there is a lower-than-average percentage of those Asian, Native Hawaiian, and American Indians using the funds for refinancing purposes. This can also be shown on the Purchaser Type and Agency Name by Race KPIs. Because the Loan Purpose KPI is the one showing more equal treatment, I would only consider that KPI to be on target, as each agency ensures fair and equal treatment between financial institutions and lenders, respecting their use of the funds.

With the original dataset, there were two noticeable changes that I had to make. The first change was to my sample size. The original dataset showed over 466,000 rows of data consisting of Washington state residents applying for mortgages. I shortened the sample to 4,999 rows because I believed the excel file did not save any changes due to how much data was collected. This turned out to be false, as the file extension was .csv instead of .xlsx. Nevertheless, the dashboard shows the results of the 4,999 rows of data collected.

The second change concerns the Application Approval Rate by Race and Sex KPIs. I noticed when creating the dashboards that the data collected within the column titled denial\_reason\_name\_1 was not showing and may have been deleted. As I discovered that other agencies conduct financial oversight between the lenders (Purchaser Type) and the applicants, I decided that my last two KPIs should analyze which agencies conduct that oversight. Therefore, the bottom four KPIs shown above and the two KPIs below are shown in the dashboard.

**Agency Name by Race**

* KPI from a customer perspective.
* I chose this KPI to determine which agencies oversee the lending process by race.
* This KPI is being measured to determine the demographics of applicants the CFPB, HUD, or another agency oversees.
* This measure is based on the percentage of each agency name by applicant’s race.
* All raw data will come from Kaggle from the “HDMA Washington State Home Loans, 2016” dataset. In the dataset, I will use the column titled agency\_name and applicant\_race\_name.
* This KPI will relate to the firm’s goal of ensuring equal treatment between applicants and lenders. Although these agencies are not lenders, it is crucial to determine whether any discrimination or favoritism has occurred.

**Agency Name by Sex**

* KPI from a customer perspective.
* I chose this KPI to determine which agencies oversee the lending process by sex.
* This KPI is being measured to determine the demographics of applicants the CFPB, HUD, or another agency oversees.
* This measure is based on the percentage of each agency name by applicant’s sex.
* All raw data will come from Kaggle from the “HDMA Washington State Home Loans, 2016” dataset. In the dataset, I will use the column titled agency\_name and applicant\_sex\_name.
* This KPI will relate to the firm’s goal of ensuring equal treatment between applicants and lenders. Although these agencies are not lenders, it is crucial to determine whether any discrimination or favoritism has occurred.

As defined by the Legal Information Institute a Cornell University, “A mortgage involves the transfer of an interest in land as [security](https://www.law.cornell.edu/wex/secured_transactions" \o "reference on security" \t "https://www.law.cornell.edu/wex/_self) for a loan or other obligation. It is the most common method of financing [real estate](https://www.law.cornell.edu/wex/real_property" \o "reference on real estate" \t "https://www.law.cornell.edu/wex/_self) transactions. The mortgagor is the party transferring the interest in land. The mortgagee, usually a financial institution, is the provider of the loan or other interest given in exchange for the security interest.” (Legal Information Institute, n.d.). With mortgages, there are regulations enforced by certain agencies, most commonly by the Consumer Financial Protection Bureau and the Department of Housing and Urban Development, that ensures that applicants aren’t discrimminated against by purchaser, such as Fannie Mae, Freddie Mac, or Ginnie Mae. As determined by the sample of data, the differences in inequalities among each purchasers shows a concern towards the commitment of racial equality.

One decision that should be made by each enforcing agency is, if not already enforced, to to require each agency to notify each purchaser on whether each race is below or above the average percent of applicants of each state. For example, Ginnie Mae is conducting business in Washington State. Each enforcing agency should first provide Ginnie Mae with the racial backgroud of each applicant utilizing their services and second, have each agency enfore Ginnie Mae to commit towards equalizing their demographics based on all applicants in the state of Washington. Another decision, very similar to the first, but instead enforce those commitments towards the enforcing government agencies. Also, making sure each purchaser is actively interested in the diversity among their applicants is a must. Protecting applicants from discriminatory practices, as described by researchers at Northwestern University, “should continue, and efforts should be increased to ensure that all home seekers receive equal treatment regardless of their race or ethnic background.” (Hurd Anyaso, 2020). Lastly, it’s important for each enforcing agency to continue to follow all laws, and to realize their importance in providing honesty and transparency to the residents of the United States.

The process of building a dashboard was a new experience for me. As I started this class with minimal experience using Excel, learning new tools throughout each week as well as watching videos on YouTube of the dasboard building process, I feel very proud of my dasboard work. Before this class, I have used Excel just for it’s general purpose, viewing spreadsheets and collecting data and now, I feel like I’ve utilize Excel for someone more useful. The process of building the dashboard came with the ways I wanted to interpet my data, with each bar being stacked. With varying samples of race and sex, I wanted to visualize my data as having each bar be of equal length with varying color to show each purchaser, loan purpose, and agency comparing each bar to race and sex. Testing each header in the pivot chart to fulfill my wish of having stacked bar graphs was a challenge, but it eventually worked in the end.

An an Information Technology major with a Data Analytics concentration, I question how dashboards will help me in the future. If any reason, could be used in a financial standpoint by determining the costs of implementing IT infrastructure. Dashboards could also be used to conduct IT inventory management; collecting the numbers of devices already deployed such as the value, RAM size, and hard drive size. And maybe, I might utilize my concentration to work in the Data Analytics field.

**References**

Auman-Bauer, K., & Ward, M. (2020, July 23). *Uncovering racial disparities in mortgage markets*. Penn State University. Retrieved December 3, 2022, from <https://www.psu.edu/news/research/story/uncovering-racial-disparities-mortgage-markets/>

Hurd Anyaso, H. (2020, January 23). *Racial discrimination in mortgage market persistent over last four decades*. Northwestern Now. Retrieved December 3, 2022, from <https://news.northwestern.edu/stories/2020/01/racial-discrimination-in-mortgage-market-persistent-over-last-four-decades/>

Legal Information Institute. (n.d.). *Mortgage*. Legal Information Institute. Retrieved December 3, 2022, from <https://www.law.cornell.edu/wex/mortgage>

Miker400. (2018, March 13). *HDMA Washington State Home Loans, 2016*. Kaggle. Retrieved October 13, 2022, from <https://www.kaggle.com/datasets/miker400/washington-state-home-mortgage-hdma2016>

*On your side through life's financial moments.* Consumer Financial Protection Bureau. (n.d.). Retrieved October 12, 2022, from <https://www.consumerfinance.gov/?_gl=1%2A1g2acw5%2A_ga%2AMzQ3NDYzNDUyLjE2NjU1OTQxMjk.%2A_ga_DBYJL30CHS%2AMTY2NTU5NDEyOS4xLjEuMTY2NTU5NDE2MS4wLjAuMA..>

*The Bureau*. Consumer Financial Protection Bureau. (n.d.). Retrieved October 13, 2022, from <https://www.consumerfinance.gov/about-us/the-bureau/>

**Raw Dataset Chosen**

See attached dataset titled “HDMA Washington State Home Loans, 2016” or <https://www.kaggle.com/datasets/miker400/washington-state-home-mortgage-hdma2016>