Tiana Worthy

CIS 4610 Special Projects

**Mastering Tableau (David Baldwin):**

The attached Tableau packaged workbook contains several worksheets and dashboards employing the concepts found in the book Mastering Tableau. The dataset used for analysis was “gse\_sf2019c\_loans.accdb” from the FHFA website, <https://www.fhfa.gov/DataTools/Downloads/Pages/Public-Use-Databases.aspx>. That dataset contains mortgage data combined with Census information. The focus of analysis was demographic mortgage access and terms. The questions asked are as follows:

What are the differences in interest rates for each demographic population?

What is the price of a loan for each demographic population?

Is there correlation between the price of loan and income?

What is the average interest rate, borrower income, and loan-to-value for each demographic in each state?

What is the average interest rate in each state?

At what age do first-time home buyers enter the market for each demographic population?

There are dashboards containing regression analysis, quartiles, and correlation. All statistical operations were performed using the R connector in Tableau. There are many calculated fields for cleaning the dataset, and one join to a .csv file with descriptive names for postal codes. The workbook contains a 500,000 row data extract from the dataset, which is originally 4.5 million rows and was too large and cumbersome to work with.

**Analyzing Police Data with Python and Pandas (Data Camp):**

Also attached is a Jupyter Notebook file containing data analysis of police traffic stops in Burlington, VT using Python, Pandas, and Matplotlib. There are fewer visualizations but more data cleaning and transforming operations performed including changing datatypes, creating new data frames, checking for and eliminating nulls. This analysis is much less involved the Tableau visualizations simply for lack of time to dig deeper.