IST 718 - Big Data Analytics Project Checkpoint 2

By: Brian Taylor, Matthew Stockhaus & Sathish Rajendiran

Status report:

- Specification - What problem are you attempting to solve?
    - Due to COVID-19 changing social and professional constructs, people have moved for various reasons to different parts of the country. From the news, online publications, and personal experiences, we can see first-hand that this has had a significant impact on the overall housing market. Our goal with this project is to use our selected datasets to see over a 3-year period, focusing on the last year particularly, the trends where people are moving from and where they are moving to, to gauge which areas of the country are seeing population increases and the associated impacts on housing costs.
- Observation - What have you observed thus far in the data?
    Observing the data for our project, we have found that COVID-19 has impacted the population and housing market through the following:
    - Zillow data:
        - Home inventory has decreased since summer of 2019, but has sharply declined since May 2020, now reaching its lowest point in the last 4 years.
        - Home Median List Price has increased significantly since January 2020, plateaued in Fall 2020, and has gone back up since January 2021.
        - Home Values have continued to increase since 2012 steadily and have reached an all-time high in 2020 in the last 20 years. For example, Home Values for 2021 are highest between SFO and LA in California on the west coast, and DC to Boston on the east coast when reviewing our heatmap.
    - Employment data:
        - Employment to Population Ratio was on a steady increase from 2009 to 2020 and then declined from roughly 61% to 51% in 2020, reaching an all-time low in the last 20 years.
        - Unemployment rate had been steadily declining from 2009 to early 2020, then increased from roughly 3.75% to over 14.5% in 2020, reaching an all-time high in the last 20 years.
    - Mortgage data:
        - Downward trend in both 15 year and 30-year fixed mortgage rates since 2018, reaching all-time lows in 2020-2021 in the last 20 years.
- Analysis - What modeling techniques will you use?
    - Forecast t home values, inventory, median list price, as well as potential mortgage rates for 15 and 30 year loans.
        - Prophet
        - ARIMA / SARIMAX
        - ACF and PACF
        - MLP / Perceptron
- Recommendation - What work do you still have to do?
    - Further Model development / refinement
    - Further development of interactive heatmap / map dashboard showing price / inventory across the US.