

1.

The assumptions, or inductive bias, in this project included which additional features were chosen, how the data was split, which polynomial degrees were tested, and ultimately the degree selected and applied to the test set. The additional features chosen included, Anchorage unemployment rates, national construction materials inflation rate, and the national prime interest rates. The data was split into thirds. Data point one was assigned to the training set, then data point two to the validation set, and then data point three to the test set, and then the continuing of this pattern until all data points were in one of the three sets. The validation set was modeled against the training set with polynomial regressions of degree 1, 2, & 3. The selected function chosen for the test set was degree 1, as the results showed it had the best performance overall in the training against the validation set step while also producing a positive number of forecasted homes.

2.

Forecasted number of homes per year:

2019 3092.4026378923336  
2020 3080.3001270923087  
2021 3068.197616292284  
2022 3056.095105492259  
2023 3043.9925946922413  
2024 3031.8900838922164  
2025 3019.7875730921915  
2026 3007.6850622921665  
2027 2995.5825514921416  
2028 2983.4800406921167

3.

Validation set against training set:

Mean squared error for degree 1 : 4127599.483108724 Variance score for degree 1 : 0.4392809212594684  
Mean squared error for degree 2 : 12236596.678003373 Variance score for degree 2 : -0.6622962679125739  
Mean squared error for degree 3 : 121384998160486.12 Variance score for degree 3 : -16489700.730994252

Test set against training set:

Mean squared error for degree 1 : 4062501.436776918 Variance score for degree 1 : 0.4464181342305183

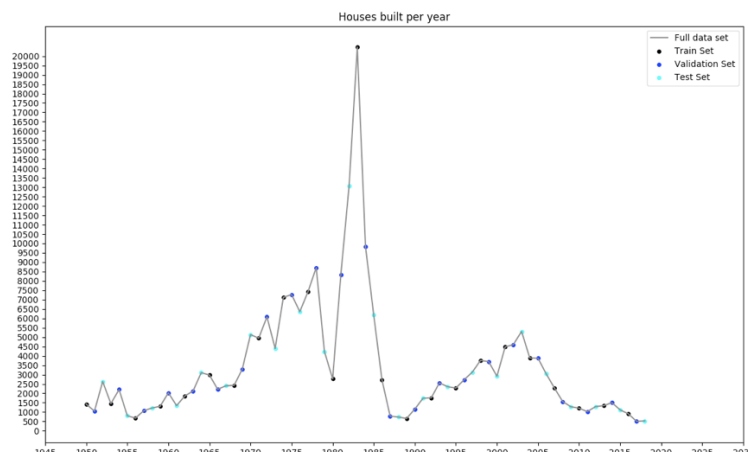


Figure 1

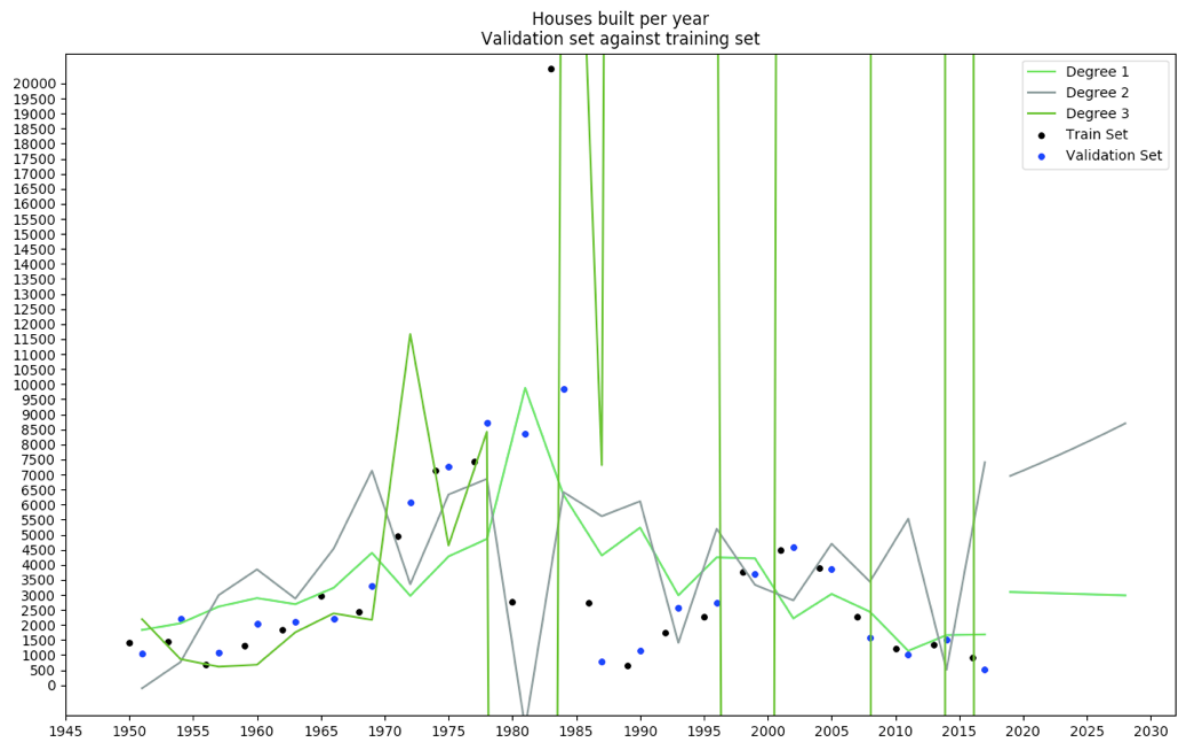


Figure 2

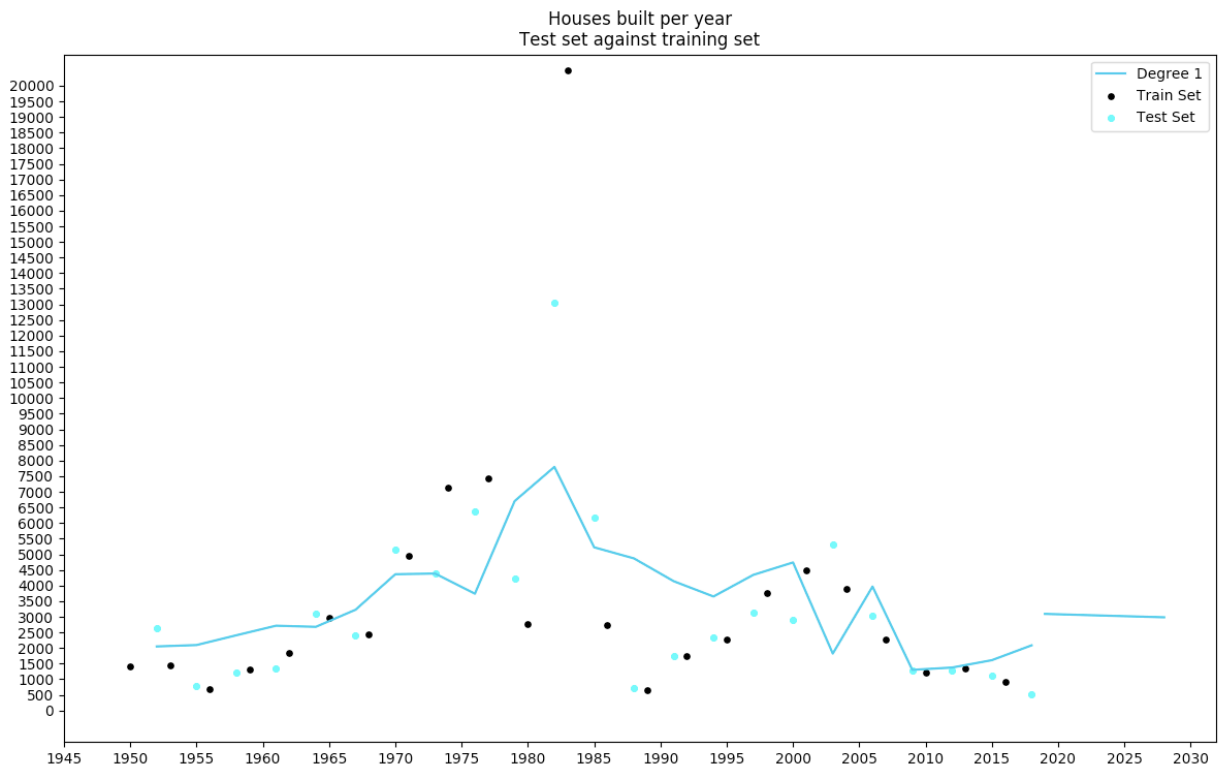


Figure 3