Executive Summary:

For my project, I created a machine learning model that predicts individual physicians total annual Medicare prescription drug cost. The data uses over 50 features that Medicare gives for each doctor. Different features pertain to the specific health demographics of physician’s patient population, number of prescriptions prescribed, physician’s specialty and much more.

            The results of this project are important because health care and particularly pharmaceutical costs in the United States are massive. Largely driven by the costs of brand name drugs there are many financial and economic reasons to monitor physician drug costs. As the project shows, for some physicians there is a wide disparately between their actual annual drug costs and the predicted annual drug costs. This is quite relevant to both Medicare and the insurance companies that provide ‘Medicare Advantage Plans’. Medicare Advantage Plans are HMO’s where people contract their Medicare to a private insurance company.

These plans, in an effort to control costs and improve quality add and eliminate doctors when deemed appropriate. It is important for insurance companies to recognize doctors with high costs relative to what they should be as they have the ability to terminate contracts with network physicians, or to work with doctors to decrease their costs. With all the variables that affect the profile of medications specific physicians prescribe, machine learning models are an excellent way to make fair predictions to see which doctors are the true outliers. It is of course important to recognize there are individual circumstances where physician’s drugs costs could be outlier’s and have reasonable explanations that weren’t accounted for within the many variables included in the model.

After concluding my project, I had a coefficient of determination of 65%. This indicates that 65% of the total annual drug costs could be concluded from the provided variables. It’s importat to realize that among other things, part of the variance is the physicians prescribing habits. This is a large factor and is what we are taking our outliers to be suggestive of.

As a means of figuring out how effective the model was, I calculated the percentage of physicians who’s total annual prescription drug cost were no higher than a specific percentage above our prediction. Of note over 83% of physicians were no higher than 50% above our prediction

|  |  |
| --- | --- |
| Percentage of predicted annual prescription costs | Doctors with drug costs below threshold |
| Within 120% of predicted prescription costs | 62.7% |
| Within 140% of predicted prescription drug costs | 78.1% |
| Within 150% of predicted prescription drug costs | 83.2% |
| Within 175% of predicted prescription drug costs | 91.1% |
| Within 200% of predicted prescription drug costs | 94.8% |

## **Featured Notebooks/Analysis/Deliverables**

Medicare\_Rx\_Slides.PPT

Medicare\_Rx\_Code.ipynb

Medicare\_Rx\_Report.docx