

Regression Project Proposal

T5 Bootcamp

Introduction

In this project we will explore the housing sale prices for King County. The data was collected for the period between May 2014 and May 2015. We will build a classification model that can predict the housing prices based on the given data.

Dataset

The dataset features about 21 columns and 21600 rows. The data was found on <https://www.kaggle.com/harlfoxem/housesalesprediction>

Questions

Which type of housing was more demanded small houses or bigger houses?

What are the most feature buyers are more interested in?

Problem statement

By using the housing sale prices between the period from May 2014 and May 2015 we will create a classification model to predict future prices.

Tools

NumPy
Pandas
Matplotlib
Seaborn
Logistic Regression Scikit-learn