Case: House Price Prediction

For this exercise you will use the King County dataset (dataset included) to predict house prices

Exercise

1. Explore the Data

Explore and visualize the data. Produce a number of charts and maps that provide insight and understanding of the dataset.

2. Predict House Prices

Create a model that predicts house prices. Try to get a prediction that is both accurate and reliable. You can use any algorithm, combination of algorithms or advanced feature engineer techniques. However you have to motivate your choices.

3. Validate Your Model

Validate the quality of your model using multiple metrics and validation criteria. Explain the strengths and weaknesses of your model.

4. Productize Your Model

Create a REST service that applies the model. The REST service predicts the value of a house given any combination of parameters. The service should also properly handle incorrect data or missing values.

Assessment Criteria

We apply the following assessment criteria.

- Apply coding best practices
- Follow proper ML methodology
- Do sufficient data exploration
- Apply necessary data preparation
- Show us your knowledge of algorithms

Don't spend too much time optimizing your model. Model performance is not the key criterium.

Preferred Technologies

We prefer the following technologies for your solution

- Python, Pandas, Scikit learn and Numpy
- Visualization: Plotly, PyPlot or Seeborn

Expected Deliveries

We expect the following deliveries:

- A Jupyter notebook containing your solution
- A word document explaining the choices you made in your solution

NOTE: put your deliveries in a zip file, as raw Python code may not make it through our firewall