Here is an Rstudio school project

Building regression model to attempt to explain prices of the house.

The steps in the **Part 1** are the following:

Predict the house price in thousands of dollars using the given features/ variables

* Loading and studying the dataset
* Conducting exploratory data analysis for the variables
* Perform a correlation analysis between variables.
* Visualizing the correlation for the variables to identify which pairs of variables are correlated
* Implement linear regression with all variables and intercept to estimate the coefficients of the model defined in the previous task to predict the prices of the house.
* Improve model
* Plot of the original vs. the fitted values
* Findings and results.

**Part 2**

Detect whether a given patient is healthy or suffers from hyperthyroidism or suffers from hypothyroidism

* Loading and studying the dataset
* Conducting exploratory data analysis for all variables
* Preparing the data for further analysis. Normalizing the data using the min-max method.
* Using stratified random sampling, dividing the data for training and testing as 70% of the data for the training set and 30% of the data for the test set.
* Conducting classification with two classification models (decision trees and k-nearest neighbor) to predict the class labels.
* Constructing for each classifier a model that neither overfits nor underfits the data.
* 8. Redoing the KNN classification for the original data (without scaling).
* 9. Detailed conclusion on results.

**Part 3**

find similarities of different types of wine by grouping them via clustering.

* Loading and studying the dataset
* Conducting exploratory data analysis for the variables
* Using three different methods: Elbow method, Silhouette method and Gap statistic method
* Use the optimal number of clusters for the clustering
* Evaluate results and performance of your model
* Detail results