Steps:

1. Import Data
2. Clean Data
3. Split data into training and test set
4. Create a model. Selecting an algorithm to analyse the data. Libraries provides algorithms e.g scikit learn
5. Train the model
6. Make predictions
7. Evaluate predictions and improve. Change algorithm or parameters

Model is trained, saved and then used. This is called model persistence.

Libraries used

* Numpy
* Pandas
* MatPlotLib
* Sci-kit Learn

Machine learning projects are mostly developed in Jupyter notebook.

**Actual Project**

Task: We have an online music store. When users sign up, we take their age and gender and recommend them music albums they would like to increase sales.

Very basic. Data is such that men at a certain age range like a music genre and same for women.

Data did not require cleaning.

Split data into training and test set in the ratio 80:20.

Used functions from sklearn DecisionTreeClassifier, train\_test\_split. And accuracy\_score module to split the data into training and test set, train a decision tree classifier and to calculate the accuracy of the model.