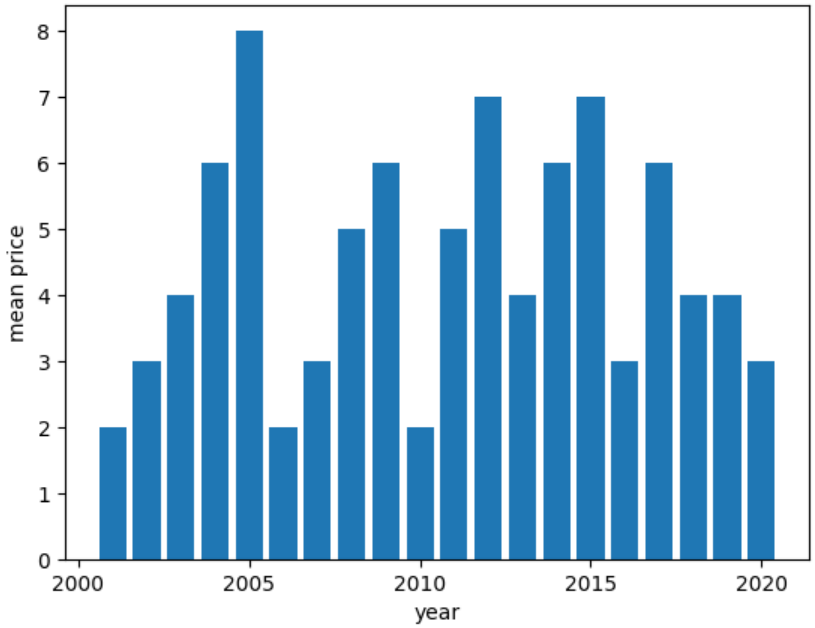
EE 381 Project

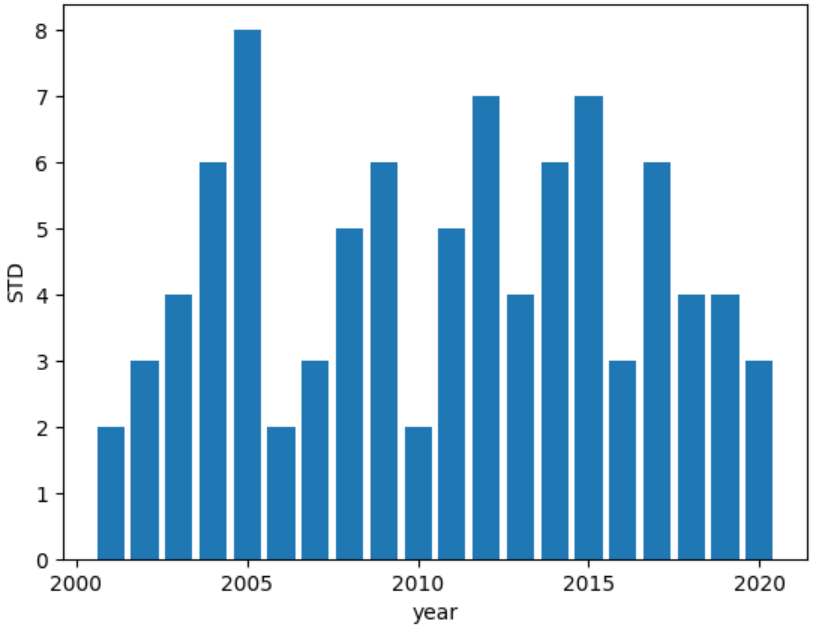
In this project we will numerically calculate the mean, standard deviation and probabilities of a large dataset using python.

The dataset used in this project consists of residential units' sale prices over a period of 20 years in a particular state. The dataset is provided as a separate CSV file 'Sales_01_20'. Please note that the data provided is not sorted.

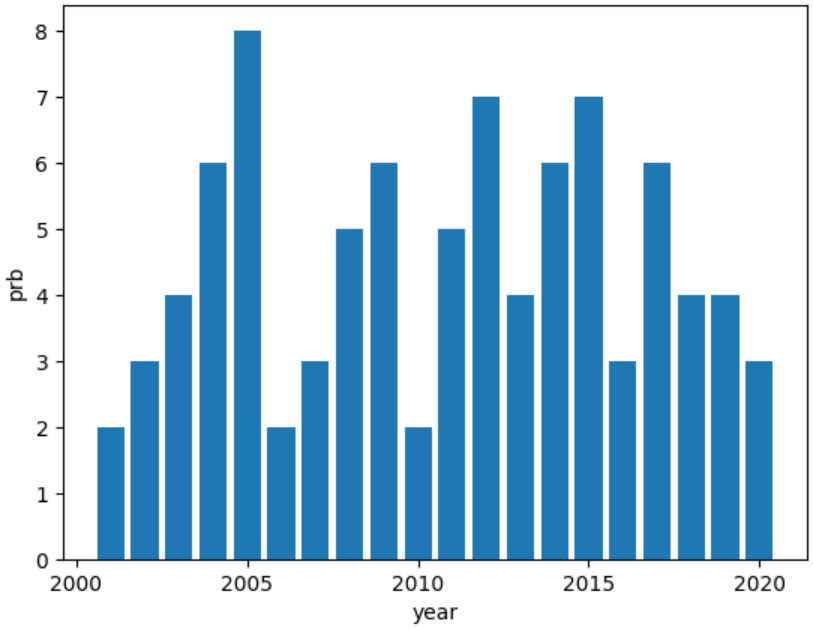
Calculate the yearly mean and standard deviation (STD) values of sale prices from 2001 to 2020. Plot the results using two bar graphs, an example is shown below.

A report template is provided, please fill the report template file, save as pdf and submit.





We would like to calculate the yearly probability of sale price ranging from $200,000 to $300,000 inclusive. This can be accomplished by counting number of houses sold for a price within that range divided by the total number of houses sold that year. An example result is provided below.



Report submission instruction:

Please submit **two files** as your report, these include a project report (pdf) created using the report template provided and your pycharm code file (.py) Please do not submit MS word format report file. Please note that correct results reported using an incorrect code will not receive any grade.