



AIMLCPFDS/DSECLPFDS - Python Exercises for Practice (S4 to S6)

NumPy:

1. Write a NumPy program to find the number of elements of an array, length of one array element in bytes and total bytes consumed by the elements.
2. Write a NumPy program to find the set difference of two arrays. The set difference will return the sorted, unique values in array1 that are not in array2.
3. Write a NumPy program to compute the cross product of two given vectors.
4. Write a NumPy program to compute the determinant of a given square array.
5. Write a NumPy program to compute the eigenvalues and right eigenvectors of a given square array.
6. Write a NumPy program to multiple three matrices each of 3*3 dimension.
7. Explore the Linear Algebra support provided by NumPy.

Pandas and Visualization:

1. Open a Jupyter notebook and import pandas, NumPy, matplotlib, seaborn and Sklearn.
2. Load the inbuilt *iris* dataset. (Download this dataset from the internet).
3. Peek into this data using `head()`, `info()` and glance over some statistics using `describe()`.
4. Plot the distribution of all numerical features and the categorical target using matplotlib and observe the plots.
5. Plot a “feature pair-wise” scatter plot to see how the numerical features are correlated to each other and print out the pairwise correlation coefficients between the numerical features.

This is known as EDA (exploratory data analysis) and it is recommended for any good data scientist to explore data thoroughly before moving on to the training/modelling phase.
