Deepali Pravali

307

Assignment - 4

Date : / /20

Title - White python code that loads any dataset from & does some basic data cleaning, add component on dataset.

Problem statement :- TO python code for that loads any data set & plot the graph.

pre-lab :- A basic understanding of computer programming terminologies, A basic understanding of any of Programming language will helpp in understanding the python programming language & datascience concepts.

Theory :-

1) step 1 Acquring data -
you need to acquire the data you need to obtain the source material before analyzing or acting on it.

2 step 2 - A :- exploring data
Put data together that you need.

step 2 - B Pre-Prossesing Data.
The raw data that you get from source are never in format you need to

Poor

Perform analysis on. There are two main goals in the data pre-processing step. The first is to clean the data to address data quality issues.

4 Step 3:- Analyzing Data :-
Analysing data now that you have ~~your~~ data nicely prepared, next step is to analyze the data. Data analysis involves building a model from your data, which is called in data the input data is used by the analysis technique to build a model.

5 Step 4:- communicating Result -
reporting insights the fourth step in our data science process is reporting the insights gained from our awareness their is very importance step to communic your insight & make a change shape based on your audience & should not be taken lightly.

- Dropping ~~col~~ansir a Data frame.
often you'll find not all the categories of data useful to you.

Date : / /20

for e.g you might have a dataset containing student information but want to four on containing student information Pandas provided a handy way of removing unwanted column or rows from a Dataframe with the drope() fron lets look at a simple e.g where dedrop and of columns from a data frame. let us now see how we can handle missing values using pandas.

Conclusion :- Thus, student can implement notebook for (perform step3 & step 4 data science steps for any data (IPL-data) by using python. tooks like pandas, matplottib, numpy etc.

Pooja