

```

print("Pima Data Set") print("This study is about Pima Indians, a community of American Indians in
Arizona, which have an unusual high incidence and prevalence of diabetes mellitus. The data set has 9
variables, (see Table below), and the typical machine learning task is to predict the class variable of whether
the subject has been diagnosed as diabetic or not.") 1. Number of times pregnant [npreg] 2. Plasma glucose
concentration [glu] 3. Diastolic blood pressure (mm Hg) [bp] 4. Triceps skin fold thickness (mm) [skin] 5.
[2-Hour serum insulin (mu U/ml)] 6. Body mass index (weight in kg/(height in m) 2) [bmi] 7. Diabetes
pedigree function[ped] 8. Age (years)[age] 9. Diabetes (No or Yes) [Type] Table : Variables in the Pima
data set. From description at UCI Machine Learning Repository.") print(" Following the book entitled :
print("Here Date Set is build taken the results of insulin test each 2-Hours serum insulin (mu U/m),
and by consequence this variable numbered 5 doesnt appear at the following analysis ") print("Data Set
pimatraining : 200*8 pimaesting : 332*8 MachineLearningSupervisedProblem : Classification")print(" Bibliografat
FinnrupNielsen.DataMiningwithPython(Workingdraft)29, 2017

```