**Data Set Characteristics:**

There are ten baseline variables, age, sex, body mass index, average blood pressure, and six blood serum measurements were obtained for each individual patients, as well as the response of interest, a quantitative measure of disease progression one year after baseline.

In this assignment we are going to use pre-built dataset called Boston house-prices to answer all the question and here is the information about all the variables available in that dataset.

Number of Instances: 442 Number of Attributes: First 10 columns are numeric predictive values Target Column 11 is a quantitative measure of disease progression one year after baseline Attribute Information Age in years

sex

(BMI): body mass index

bp average blood pressure

s1 tc, T-Cells (a type of white blood cells)

s2 ldl, low-density lipoproteins

s3 hdl, high-density lipoproteins

s4 tch, thyroid stimulating hormone

s5 ltg, lamotrigine

s6 glu, blood sugar level

**Perform the given set of Instructions on the loaded dataset(Diabetes) using Pandas Library:**

Load the Pandas library and convert the loaded dataset into Dataframe for further execution.

(1)Get the name of all of the column stored in the Dataframe.

(2)Find the shape of the dataframe and perform Basic summary of variables in Dataframe.

(3)Select on one of the variable to be index of the Dataframe.

(4)What is Series in Dataframe and how it is differentiated from Numpy Arrays.

(5)Check if there are any missing values in the created Dataframe. Also check for the datatype of all the variable stored in Dataframe.

(6)Slice the data from 3 row to 7 row and from 2 column to 10 column.Check if any variable is not useful for your analysis of predicting whether patient is having Diabetes or not and drop that particular column that is not needed.

(7)Create a DataFrameGroupBy object BMI: Body Mass Index and saving it into new Dataframe named BMI. Calculate Groupbby Count, Mean and Variance for Dataframe BMI.

(8)What is the difference between unique() and nunique() function of Dataframe. Find unique() and nunique() for bp average blood pressure in the Dataframe.

(9)Get the list of all columns and sort the whole dataframe according to a specific column of your choice.

In [1]:



**import** pandas **as** pd

**import** numpy **as** np

**from** sklearn.datasets **import** load\_diabetes

diabetes **=** load\_diabetes()

diabetes\_df**=**pd.DataFrame(diabetes.data)

diabetes\_df

Out[1]:

|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 0.038076 | 0.050680 | 0.061696 | 0.021872 | -0.044223 | -0.034821 | -0.043401 | -0.002592 | 0.019908 | -0.017646 |
| **1** | -0.001882 | -0.044642 | -0.051474 | -0.026328 | -0.008449 | -0.019163 | 0.074412 | -0.039493 | -0.068330 | -0.092204 |
| **2** | 0.085299 | 0.050680 | 0.044451 | -0.005671 | -0.045599 | -0.034194 | -0.032356 | -0.002592 | 0.002864 | -0.025930 |
| **3** | -0.089063 | -0.044642 | -0.011595 | -0.036656 | 0.012191 | 0.024991 | -0.036038 | 0.034309 | 0.022692 | -0.009362 |
| **4** | 0.005383 | -0.044642 | -0.036385 | 0.021872 | 0.003935 | 0.015596 | 0.008142 | -0.002592 | -0.031991 | -0.046641 |
| **5** | -0.092695 | -0.044642 | -0.040696 | -0.019442 | -0.068991 | -0.079288 | 0.041277 | -0.076395 | -0.041180 | -0.096346 |
| **6** | -0.045472 | 0.050680 | -0.047163 | -0.015999 | -0.040096 | -0.024800 | 0.000779 | -0.039493 | -0.062913 | -0.038357 |
| **7** | 0.063504 | 0.050680 | -0.001895 | 0.066630 | 0.090620 | 0.108914 | 0.022869 | 0.017703 | -0.035817 | 0.003064 |
| **8** | 0.041708 | 0.050680 | 0.061696 | -0.040099 | -0.013953 | 0.006202 | -0.028674 | -0.002592 | -0.014956 | 0.011349 |
| **9** | -0.070900 | -0.044642 | 0.039062 | -0.033214 | -0.012577 | -0.034508 | -0.024993 | -0.002592 | 0.067736 | -0.013504 |
| **10** | -0.096328 | -0.044642 | -0.083808 | 0.008101 | -0.103389 | -0.090561 | -0.013948 | -0.076395 | -0.062913 | -0.034215 |
| **11** | 0.027178 | 0.050680 | 0.017506 | -0.033214 | -0.007073 | 0.045972 | -0.065491 | 0.071210 | -0.096433 | -0.059067 |
| **12** | 0.016281 | -0.044642 | -0.028840 | -0.009113 | -0.004321 | -0.009769 | 0.044958 | -0.039493 | -0.030751 | -0.042499 |
| **13** | 0.005383 | 0.050680 | -0.001895 | 0.008101 | -0.004321 | -0.015719 | -0.002903 | -0.002592 | 0.038393 | -0.013504 |
| **14** | 0.045341 | -0.044642 | -0.025607 | -0.012556 | 0.017694 | -0.000061 | 0.081775 | -0.039493 | -0.031991 | -0.075636 |
| **15** | -0.052738 | 0.050680 | -0.018062 | 0.080401 | 0.089244 | 0.107662 | -0.039719 | 0.108111 | 0.036056 | -0.042499 |
| **16** | -0.005515 | -0.044642 | 0.042296 | 0.049415 | 0.024574 | -0.023861 | 0.074412 | -0.039493 | 0.052280 | 0.027917 |
| **17** | 0.070769 | 0.050680 | 0.012117 | 0.056301 | 0.034206 | 0.049416 | -0.039719 | 0.034309 | 0.027368 | -0.001078 |
| **18** | -0.038207 | -0.044642 | -0.010517 | -0.036656 | -0.037344 | -0.019476 | -0.028674 | -0.002592 | -0.018118 | -0.017646 |
| **19** | -0.027310 | -0.044642 | -0.018062 | -0.040099 | -0.002945 | -0.011335 | 0.037595 | -0.039493 | -0.008944 | -0.054925 |
| **20** | -0.049105 | -0.044642 | -0.056863 | -0.043542 | -0.045599 | -0.043276 | 0.000779 | -0.039493 | -0.011901 | 0.015491 |
| **21** | -0.085430 | 0.050680 | -0.022373 | 0.001215 | -0.037344 | -0.026366 | 0.015505 | -0.039493 | -0.072128 | -0.017646 |
| **22** | -0.085430 | -0.044642 | -0.004050 | -0.009113 | -0.002945 | 0.007767 | 0.022869 | -0.039493 | -0.061177 | -0.013504 |
| **23** | 0.045341 | 0.050680 | 0.060618 | 0.031053 | 0.028702 | -0.047347 | -0.054446 | 0.071210 | 0.133599 | 0.135612 |
| **24** | -0.063635 | -0.044642 | 0.035829 | -0.022885 | -0.030464 | -0.018850 | -0.006584 | -0.002592 | -0.025952 | -0.054925 |
| **25** | -0.067268 | 0.050680 | -0.012673 | -0.040099 | -0.015328 | 0.004636 | -0.058127 | 0.034309 | 0.019199 | -0.034215 |
| **26** | -0.107226 | -0.044642 | -0.077342 | -0.026328 | -0.089630 | -0.096198 | 0.026550 | -0.076395 | -0.042572 | -0.005220 |
| **27** | -0.023677 | -0.044642 | 0.059541 | -0.040099 | -0.042848 | -0.043589 | 0.011824 | -0.039493 | -0.015998 | 0.040343 |
| **28** | 0.052606 | -0.044642 | -0.021295 | -0.074528 | -0.040096 | -0.037639 | -0.006584 | -0.039493 | -0.000609 | -0.054925 |
| **29** | 0.067136 | 0.050680 | -0.006206 | 0.063187 | -0.042848 | -0.095885 | 0.052322 | -0.076395 | 0.059424 | 0.052770 |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **412** | 0.074401 | -0.044642 | 0.085408 | 0.063187 | 0.014942 | 0.013091 | 0.015505 | -0.002592 | 0.006209 | 0.085907 |
| **413** | -0.052738 | -0.044642 | -0.000817 | -0.026328 | 0.010815 | 0.007141 | 0.048640 | -0.039493 | -0.035817 | 0.019633 |
| **414** | 0.081666 | 0.050680 | 0.006728 | -0.004523 | 0.109883 | 0.117056 | -0.032356 | 0.091875 | 0.054724 | 0.007207 |
| **415** | -0.005515 | -0.044642 | 0.008883 | -0.050428 | 0.025950 | 0.047224 | -0.043401 | 0.071210 | 0.014823 | 0.003064 |
| **416** | -0.027310 | -0.044642 | 0.080019 | 0.098763 | -0.002945 | 0.018101 | -0.017629 | 0.003312 | -0.029528 | 0.036201 |
| **417** | -0.052738 | -0.044642 | 0.071397 | -0.074528 | -0.015328 | -0.001314 | 0.004460 | -0.021412 | -0.046879 | 0.003064 |
| **418** | 0.009016 | -0.044642 | -0.024529 | -0.026328 | 0.098876 | 0.094196 | 0.070730 | -0.002592 | -0.021394 | 0.007207 |
| **419** | -0.020045 | -0.044642 | -0.054707 | -0.053871 | -0.066239 | -0.057367 | 0.011824 | -0.039493 | -0.074089 | -0.005220 |
| **420** | 0.023546 | -0.044642 | -0.036385 | 0.000068 | 0.001183 | 0.034698 | -0.043401 | 0.034309 | -0.033249 | 0.061054 |
| **421** | 0.038076 | 0.050680 | 0.016428 | 0.021872 | 0.039710 | 0.045032 | -0.043401 | 0.071210 | 0.049769 | 0.015491 |
| **422** | -0.078165 | 0.050680 | 0.077863 | 0.052858 | 0.078236 | 0.064447 | 0.026550 | -0.002592 | 0.040672 | -0.009362 |
| **423** | 0.009016 | 0.050680 | -0.039618 | 0.028758 | 0.038334 | 0.073529 | -0.072854 | 0.108111 | 0.015567 | -0.046641 |
| **424** | 0.001751 | 0.050680 | 0.011039 | -0.019442 | -0.016704 | -0.003819 | -0.047082 | 0.034309 | 0.024053 | 0.023775 |
| **425** | -0.078165 | -0.044642 | -0.040696 | -0.081414 | -0.100638 | -0.112795 | 0.022869 | -0.076395 | -0.020289 | -0.050783 |
| **426** | 0.030811 | 0.050680 | -0.034229 | 0.043677 | 0.057597 | 0.068831 | -0.032356 | 0.057557 | 0.035462 | 0.085907 |
| **427** | -0.034575 | 0.050680 | 0.005650 | -0.005671 | -0.073119 | -0.062691 | -0.006584 | -0.039493 | -0.045421 | 0.032059 |
| **428** | 0.048974 | 0.050680 | 0.088642 | 0.087287 | 0.035582 | 0.021546 | -0.024993 | 0.034309 | 0.066048 | 0.131470 |
| **429** | -0.041840 | -0.044642 | -0.033151 | -0.022885 | 0.046589 | 0.041587 | 0.056003 | -0.024733 | -0.025952 | -0.038357 |
| **430** | -0.009147 | -0.044642 | -0.056863 | -0.050428 | 0.021822 | 0.045345 | -0.028674 | 0.034309 | -0.009919 | -0.017646 |
| **431** | 0.070769 | 0.050680 | -0.030996 | 0.021872 | -0.037344 | -0.047034 | 0.033914 | -0.039493 | -0.014956 | -0.001078 |
| **432** | 0.009016 | -0.044642 | 0.055229 | -0.005671 | 0.057597 | 0.044719 | -0.002903 | 0.023239 | 0.055684 | 0.106617 |
| **433** | -0.027310 | -0.044642 | -0.060097 | -0.029771 | 0.046589 | 0.019980 | 0.122273 | -0.039493 | -0.051401 | -0.009362 |
| **434** | 0.016281 | -0.044642 | 0.001339 | 0.008101 | 0.005311 | 0.010899 | 0.030232 | -0.039493 | -0.045421 | 0.032059 |
| **435** | -0.012780 | -0.044642 | -0.023451 | -0.040099 | -0.016704 | 0.004636 | -0.017629 | -0.002592 | -0.038459 | -0.038357 |
| **436** | -0.056370 | -0.044642 | -0.074108 | -0.050428 | -0.024960 | -0.047034 | 0.092820 | -0.076395 | -0.061177 | -0.046641 |
| **437** | 0.041708 | 0.050680 | 0.019662 | 0.059744 | -0.005697 | -0.002566 | -0.028674 | -0.002592 | 0.031193 | 0.007207 |
| **438** | -0.005515 | 0.050680 | -0.015906 | -0.067642 | 0.049341 | 0.079165 | -0.028674 | 0.034309 | -0.018118 | 0.044485 |
| **439** | 0.041708 | 0.050680 | -0.015906 | 0.017282 | -0.037344 | -0.013840 | -0.024993 | -0.011080 | -0.046879 | 0.015491 |
| **440** | -0.045472 | -0.044642 | 0.039062 | 0.001215 | 0.016318 | 0.015283 | -0.028674 | 0.026560 | 0.044528 | -0.025930 |
| **441** | -0.045472 | -0.044642 | -0.073030 | -0.081414 | 0.083740 | 0.027809 | 0.173816 | -0.039493 | -0.004220 | 0.003064 |

442 rows × 10 columns