CODE FOR DATA SCIENCE (DIABETES PREDICTION SYSTEM):

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn import metrics
%matplotlib inline

data=pd.read_csv("C:\\Users\\indra\\Documents\\csvfiles\\diabetes.csv")
data.head(10)

#checking the number of rows and columns in the file
data.shape

#checking for null values in the data
data.isnull().values.any()

data.rename(columns={'DiabetesPedigreeFunction':'DPF','BloodPressure':'BP'},
inplace=True)
data.head(5)

data.describe()

data.corr()
```

```
corrmat=data.corr()
top_corr_features=corrmat.index
plt.figure(figsize=(10,10))
g=sns.heatmap(data[top_corr_features].corr(),annot=True,cmap="RdYlGn")



print("Number of zeros in Glocose : ",data[data["Glucose"]==0].shape[0])
print("Number of zeros in Blood Pressure : ",data[data["BP"]==0].shape[0])
print("Number of zeros in Skin Thickness :
",data[data["SkinThickness"]==0].shape[0])
print("Number of zeros in Insulin : ",data[data["Insulin"]==0].shape[0])
print("Number of zeros in BMI : ",data[data["BMI"]==0].shape[0])
print("Number of zeros in Diabetes Prediction Factor :
",data[data["DPF"]==0].shape[0])
print("Number of zeros in Age : ",data[data["Age"]==0].shape[0])


data["Glucose"]=data["Glucose"].replace(0,data["Glucose"].mean())
data["BP"]=data["BP"].replace(0,data["BP"].mean())
data["SkinThickness"]=data["SkinThickness"].replace(0,data["SkinThickness"].
mean())
data["Insulin"]=data["Insulin"].replace(0,data["Insulin"].mean())
data["BMI"]=data["BMI"].replace(0,data["BMI"].mean())
data["DPF"]=data["DPF"].replace(0,data["DPF"].mean())
data["Age"]=data["Age"].replace(0,data["Age"].mean())


print("Number of zeros in Glocose : ",data[data["Glucose"]==0].shape[0])
```

```
print("Number of zeros in Blood Pressure : ",data[data["BP"]==0].shape[0])
print("Number of zeros in Skin Thickness :
",data[data["SkinThickness"]==0].shape[0])
print("Number of zeros in Insulin : ",data[data["Insulin"]==0].shape[0])
print("Number of zeros in BMI : ",data[data["BMI"]==0].shape[0])
print("Number of zeros in Diabetes Prediction Factor :
",data[data["DPF"]==0].shape[0])
print("Number of zeros in Age : ",data[data["Age"]==0].shape[0])

#counting the total individual Outcomes
positive_outcome=len(data.loc[data["Outcome"]==1])
negative_outcome=len(data.loc[data["Outcome"]==0])
(positive_outcome,negative_outcome)

y=np.array([positive_outcome,negative_outcome])
mylabels=["Diabetic people (268)","Non-diabetic people (500)"]
plt.pie(y,labels=mylabels,colors=["orange","yellow"])
plt.title("Number of diabetic and Non-diabetic persons")
plt.show()

df={'Diabetic':positive_outcome,'Non-diabetic':negative_outcome}
A=list(df.keys())
B=list(df.values())
plt.bar(A,B,width=0.2)
plt.title("Number of diabetic and Non-diabetic persons")
plt.show()
```

```
X=data.drop(columns=["Outcome"])
Y=data["Outcome"]
X_test,X_train,Y_test,Y_train=train_test_split(X,Y,test_size=0.30,random_state
=10)

model=RandomForestClassifier(random_state=10)
model.fit(X_train,Y_train.ravel())
pred=model.predict(X_test)
pred

acc=metrics.accuracy_score(Y_test,pred)
print("ACCURACY OF THE MODEL : ",acc)

def prediction_calculator(n):
   for i in range(n):
      print("\n ENTER THE DETAILS FOR PERSON : ",(i+1))
      Age_ip=input("\nAge : ")
      Gender=input('Gender (f/F/m/M): ')
      if Gender=='f' or Gender=='F':
          Preg_ip=input("Number of Pregnancies : ")
      else:
         Preg_ip=0
      Bmi_ip=input("BMI : ")
      Glucose_ip=input("Glucose level : ")
      Insulin_ip=input("Insulin level : ")
      Bp_ip=input("BP level : ")
      St_ip=input("Skin Thickness : ")
```

```
    Dpf_ip=input("Diabetes prediction factor : ")


c=np.array([Preg_ip,Glucose_ip,Bp_ip,St_ip,Insulin_ip,Bmi_ip,Dpf_ip,Age_ip])
    c_rs=c.reshape(1,-1)
    pred=model.predict(c_rs)

    if pred==1:
       print("DIABETIC PERSON !!")
    else:
       print("NON-DIABETIC PERSON :)")


no_of_people=int(input("\n ENTER NUMBER OF PEOPLE : "))
prediction_calculator(no_of_people)
```

DONE BY:

YASHWANT SAIARJUN.S.V

VUTUKURI ESWAR CHAND

YARRAM ABILASH REDDY

KARRI VENKATA NAGA SAI VAMSI