```
import numpy as np
import pandas as pd

df = pd.read_csv('/content/diabetes.csv')
null_columns = ['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness',
'Insulin', 'BMI', 'Age', 'Outcome']
df[null_columns] = df[null_columns].replace({0:np.nan})
df.to_csv('/content/diabetes.csv', index=False)
df.head(20)


mean_value=df['Pregnancies'].mean()
print('Mean Value: '+str(mean_value))

null_columns = ['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness',
'Insulin', 'BMI', 'Age', 'Outcome']
df[null_columns] =df[null_columns].fillna(mean_value)
df.head(20)


from sklearn.datasets import load_diabetes
from sklearn.model_selection import KFold
import numpy as np
import pandas as pd

diabetes = load_diabetes()

X_diabetes = diabetes.data
y_diabetes = diabetes.target

scaler = StandardScaler()
X_scaler = scaler.fit_transform(X_diabetes)

cnt = 0
n_splits = 10
kf = KFold(n_splits=n_splits, shuffle=True, random_state=42)
for train_index, test_index in kf.split(X_scaler, y_diabetes):
    print(f'Fold:{cnt}, Train set: {len(train_index)}, \
    Test set:{len(test_index)}')
    cnt += 1


from sklearn.datasets import load_diabetes
from sklearn.model_selection import KFold, cross_val_score
from sklearn.tree import DecisionTreeRegressor
import numpy as np
import pandas as pd
from sklearn.preprocessing import StandardScaler
import math
```

```
diabetes = load_diabetes()

X = diabetes.data
y = diabetes.target

scaler = StandardScaler()
X_scaler = scaler.fit_transform(X)

cnt = 0
n_splits = 10
kf = KFold(n_splits=n_splits, shuffle=True, random_state=42)
for train_index, test_index in kf.split(X_scaler, y):
    print(f'Fold:{cnt}, Train set: {len(train_index)}, \
    Test set:{len(test_index)}')
    cnt += 1

def rmse(mse):
    return math.sqrt(abs(mse))

score = cross_val_score(DecisionTreeRegressor(random_state= 42), X, y, cv=kf,
scoring="neg_mean_squared_error")
print(f'Scores for each fold: {score}')
rmse(score.mean())
```