Step 3: Exploratory Data Analysis

Exploratory Data Analysis (EDA), also known as Data Exploration, is a step in the Data Analysis Process, where a number of techniques are used to better understand the dataset being used.

3.1) Understanding Your Variables

3.1.1) Head of the dataset
3.1.2) The shape of the dataset
3.1.3) List types of columns
3.1.4) Info of the dataset
3.1.5) Summary of the dataset

3.1.1) Head of the Dataset

```
# Display first five records
diabetes_dataset.head()

# Display last five records
diabetes_dataset.tail()

# Disply random records
diabetes_dataset.sample(5)
```

3.1.2) The Shape of Dataset

```
 # Numbers of rows and columns
diabetes_dataset.shape
```

3.1.3) List types of columns

```
# List types of all columns
diabetes_dataset.dtypes
```

3.1.4) Info of the Dataset

```
# Checking for null values
diabetes_dataset.info()

# Statistical Summary
diabetes_dataset.describe()
```

Step 4: Split the data frame in X & Y

```
target_name = 'Outcome'

# Separate object for target feature
y = diabetes_dataset[target_name]

# Separate obhect for input feature
X = diabetes_dataset.drop(target_name,axis=1)
```