# Ground water Prediction by Logistic Regression a Python solution

## Steps

### Data preparation

Import necessary libraries like numpy, matplotlib, pandas, and sklearn.

Read the dataset 'Groundwater.csv' into a pandas DataFrame.

Visualize the count of different situations using a count plot.

Preprocess the data, split it into training and testing sets, and perform feature scaling.

Train a logistic regression classifier, predict on the test set, and evaluate its performance using a confusion matrix and accuracy scores.

A screenshot of a computer

Description automatically generated

A graph with different colored bars

Description automatically generated with medium confidence

## Improving Accuracy

Attempt to improve accuracy by encoding categorical variables and splitting data differently.

Re-train the logistic regression classifier and evaluate its performance again.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

## Visualization

Visualize data using count plots and heatmaps.

Create pie charts to represent data distribution and situation percentages.

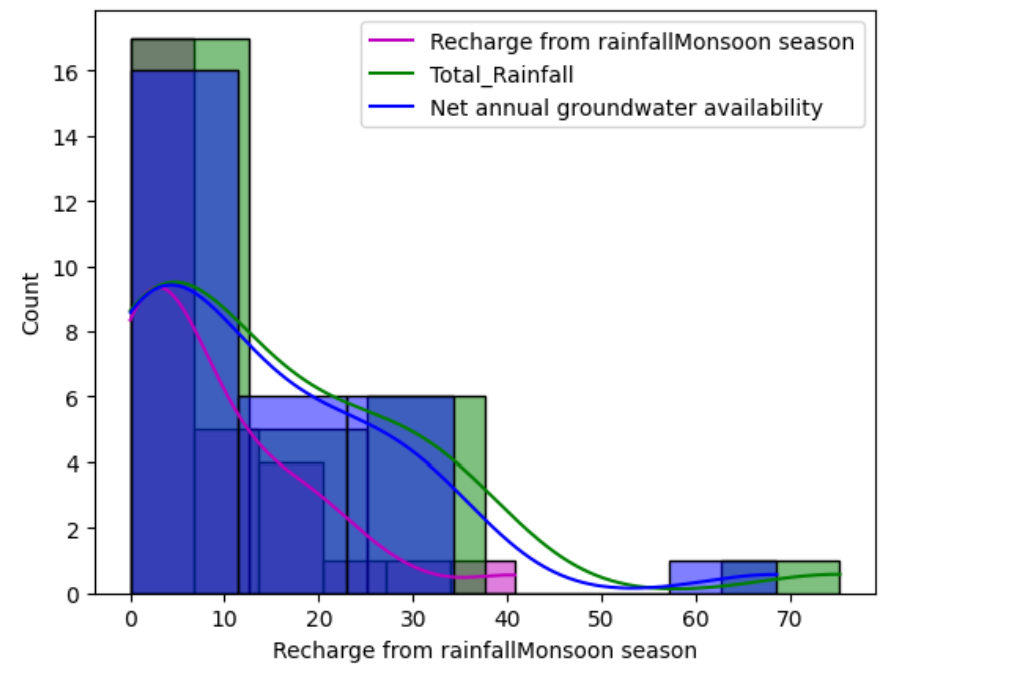
Plot scatter plots and distribution plots for various variables.

A screenshot of a computer

Description automatically generated

A pie chart with different colored circles

Description automatically generated



A graph of water quality

Description automatically generated

## A graph with green and blue lines Description automatically generated’’

## Decision Tree Classifier

Split data, scale features, and train a decision tree classifier.

Print accuracy scores and feature importances.

## Random Forest Classifier

Train random forest classifiers with different configurations and evaluate their performance.