1. **Reading data from .xls file**

Install package openpyxl:- pip install openpyxl

import openpyxl as xl  
  
wb = xl.load\_workbook('Employee\_Detail.xlsx')  
sheet = wb['Sheet1']  
sell\_value = sheet.cell(1,1)  
print(sheet.max\_row)  
for row in range (2,sheet.max\_row+1):  
 cell = sheet.cell(row,1)  
 print(cell.value)  
 newCol= cell.value \* 5  
 newCol\_cell = sheet.cell(row,5)  
 newCol\_cell.value=newCol  
wb.save('Employee\_Detail2.xlsx')

1. **Creating Bar Chart in to .xls**

import openpyxl as xl  
from openpyxl.chart import BarChart, Reference  
  
wb = xl.load\_workbook('Employee\_Detail.xlsx')  
sheet = wb['Sheet1']  
sell\_value = sheet.cell(1,1)  
print(sheet.max\_row)  
for row in range (2,sheet.max\_row+1):  
 cell = sheet.cell(row,1)  
 print(cell.value)  
 newCol= cell.value / 100  
 newCol\_cell = sheet.cell(row,5)  
 newCol\_cell.value=newCol  
value = Reference(sheet, min\_row=2, max\_row=sheet.max\_row, min\_col=5, max\_col=5)  
chart= BarChart()  
chart.add\_data(value)  
sheet.add\_chart(chart, 'g2')  
wb.save('Employee\_Detail.xlsx')

1. **Read data from CSV**

import sqlite3  
import csv  
  
conn = sqlite3.connect('example.db')  
c = conn.cursor()  
  
#Create table  
c.execute('''CREATE TABLE employee  
 (id integer PRIMARY KEY,  
 email text,  
 name text,  
 address text)''')  
  
with open('employee.csv', 'r') as f:  
 reader = csv.reader(f)  
 next(reader) # Skip the header row.  
 for row in reader:  
 print(row)  
 c.execute("INSERT INTO employee VALUES (?, ?, ?, ?)", row)  
conn.commit()  
# rows = c.execute("SELECT \* FROM employee ")  
# print(rows.fetchall())

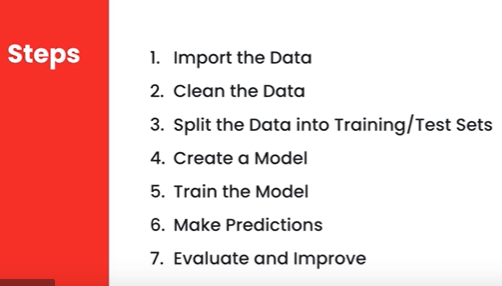
**4.Fetch/insert data from/into table**

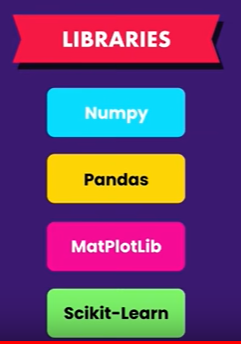
import sqlite3  
conn = sqlite3.connect('example.db')  
c = conn.cursor()  
  
symbol = 'Krishan'  
c.execute("SELECT \* FROM employee WHERE name = '%s'" % symbol)  
print(c.fetchall())  
  
#Larger example that inserts many records at a time  
employees = [('1133', 'abc@abc.com', 'Ramesh', 'Sector 11'),  
 ('1134', 'abc@abc.com', 'Shyam', 'Sector 12'),  
 ('1135', 'abc@abc.com', 'Gita', 'Sector 22'),  
 ]  
c.executemany('INSERT INTO employee VALUES (?,?,?,?)', employees)  
  
for row in c.execute('SELECT \* FROM employee ORDER BY name'):  
 print(row)

**5. Update data into table**

import sqlite3  
con = sqlite3.connect('example.db')  
  
  
def sql\_update(con):  
 cursorObj = con.cursor()  
 symbol = 'Ram'  
 cursorObj.execute("UPDATE employee SET address = 'HCL 126 noida' where name = '%s'" % symbol)  
 con.commit()  
  
  
sql\_update(con)  
c = con.cursor()  
c.execute("SELECT \* FROM employee where name='Ram' ")  
print(c.fetchall())

**6. Data analysis**





* Install Anaconda
* Download Anaconda for python3.7 from [www.Anaconda.com](http://www.Anaconda.com)
* It will install Jupyter, pandas etc..
* Open Terminal Window
* $jupyter notebook
* It will open browser window localhost:8888/tree
* Create new Python notebook
* Rename notebook to Employee\_Data

**Prediction for Employees**

import pandas as pd

from sklearn.tree import DecisionTreeClassifier

emplyee\_data = pd.read\_csv('Employee\_data.csv')X = emplyee\_data.drop(columns=['profile'])

Y = emplyee\_data['profile']

model = DecisionTreeClassifier()

model.fit(X,Y)

predictions = model.predict([ [21,1], [22,0] ])

predictions

**Calculate Prediction Score**

import pandas as pd

from sklearn.tree import DecisionTreeClassifier

from sklearn.model\_selection import train\_test\_split

from sklearn.metrics import accuracy\_score

emplyee\_data = pd.read\_csv('Employee\_data.csv')

X = emplyee\_data.drop(columns=['profile'])

Y = emplyee\_data['profile']

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X, Y, test\_size=0.2)

model = DecisionTreeClassifier()

model.fit(X\_train,Y\_train)

predictions = model.predict(X\_test)

score = accuracy\_score(Y\_test, predictions)

score

**Dump data as model**

import pandas as pd

from sklearn.tree import DecisionTreeClassifier

from sklearn.externals import joblib

emplyee\_data = pd.read\_csv('Employee\_data.csv')

X = emplyee\_data.drop(columns=['profile'])

Y = emplyee\_data['profile']

model = DecisionTreeClassifier()

model.fit(X,Y)

joblib.dump(model, 'employee-recommender.joblib')

**Load data from Dump**

import pandas as pd

from sklearn.tree import DecisionTreeClassifier

from sklearn.externals import joblib

# music\_data = pd.read\_csv('music.csv')

# X = music\_data.drop(columns=['gener'])

# Y = music\_data['gener']

# model = DecisionTreeClassifier()

# model.fit(X,Y)

model = joblib.load('employee-recommender.joblib')

predictions = model.predict([ [21,1], [22,0] ])

predictions