**OSS LAB PRACTICE SESSION 2**

Name-Ujjawal Mandhani

Batch-F6

Enroll No.-9918103237

import urllib2

import pandas as pd

wiki = "https://en.wikipedia.org/wiki/List\_of\_state\_and\_union\_territory\_capitals\_in\_India"

page = urllib2.urlopen(wiki)

from bs4 import BeautifulSoup

soup = BeautifulSoup(page)

all\_tables=soup.find\_all('table')

print(all\_tables)

right\_table=soup.find('table', class\_='wikitable sortable plainrowheaders')

print(right\_table)

A=[]

B=[]

C=[]

D=[]

E=[]

F=[]

G=[]

for row in right\_table.findAll("tr"):

cells = row.findAll('td')

states=row.findAll('th') #To store second column data

if len(cells)==6: #Only extract table body not heading

A.append(cells[0].find(text=True))

B.append(states[0].find(text=True))

C.append(cells[1].find(text=True))

D.append(cells[2].find(text=True))

E.append(cells[3].find(text=True))

F.append(cells[4].find(text=True))

G.append(cells[5].find(text=True))

df=pd.DataFrame(A,columns=['Number'])

df['State/UT']=B

df['Admin\_Capital']=C

df['Legislative\_Capital']=D

df['Judiciary\_Capital']=E

df['Year\_Capital']=F

Mongo version 3.3.0

import os

import pandas as pd

import numpy as np

from IPython.core.display import display, HTML

import pymongo

from pymongo import MongoClient

print 'Mongo version', pymongo. version

client = MongoClient('localhost', 27017)

db = client.test

collection = db.people

collection.drop()

os.system('mongoimport -d test -c people dummyData.json')

df['Former\_Capital']=G

print(df)