**A Python Script to Split a Shapefile in to Several Shapefiles**

****Abstract****

Keywords: Jupytor notebook, Python Script

Split function is very important in the field of surveying and GIS application to split shapefiles in to several shapefiles. This can be done by using several software and online platforms. But using python script we can do easily this without using more options of software. The project was design to split shapefiles in to several shapefiles using python script. This was carried out in Jupytor notebook environment in Anaconda. Anaconda is free and open source software, such that this method has several advantages and those could be achieved in this project.

# **Table of Contents**

Chapter 1-Introduction…………………………………………...1

# **Background……………………………………………….1**

* 1. Objectives………………………………………………...2
  2. Chapter outlines…………………………………………..2

Chapter 2- Methodology and stepwise results………................3-10

Chapter 3- Evaluation of Project Result…………………….11-12

Chapter 4- Conclusion and Future Developments………………13

Chapter 5-Appendix………………………………………….14-15

Chapter 1

# Introduction

## Background

There are so many spatial tools and methods that are based in Python. When we learn GIS customization and programming, we also encounter numerous GIS customization applications.

An easy-to-use, nontopological format called a shapefile is used to store the geometric coordinates and attribute data of geographic features. A shapefile's geographic elements can be represented by points, lines, or polygons. As a mostly open protocol enabling data interchange among Esri and other GIS software packages, it is created and governed by Esri. Geospatial data in vector format are often stored in a shapefile format.

There are 3 key files associated with any and all shapefiles:

* .shp: the file that contains the geometry for all features.
* .shx: the file that indexes the geometry.
* .dbf: the file that stores feature attributes in a tabular format.

These files need to have the same name and to be stored in the same directory (folder) to open properly in a GIS, R or Python tool.We can use several python platforms such that Anaconda, QGIS, Colab etc to Split a shapefile in to several shapefiles. This project can be performing simple split tool for shpefiles by using Jupitor notebook.

This method is easily and many advantages for larger datasets and possible to accsess easily.

## Objectives

This project has the following as objectives:

* Coding a python script using Jupytor notebook, A Python Script to Split a shapefile in to several shapefiles.

## Chapter Outlines

**Chapter 1**: Discusses the detailed background of the project. Then it gives the objectives that are expected to be achieved by the study.

**Chapter 2:** Methodology is presented in this chapter and step wise results also mentioned.

**Chapter 3:** The evaluation of the final results is described within

**Chapter 4:** Conclusion and Future Developments

**Chapter 5:** Final chapter which consists of the appendices required to elaborate the related environment of the project to users with interest.

Chapter 2

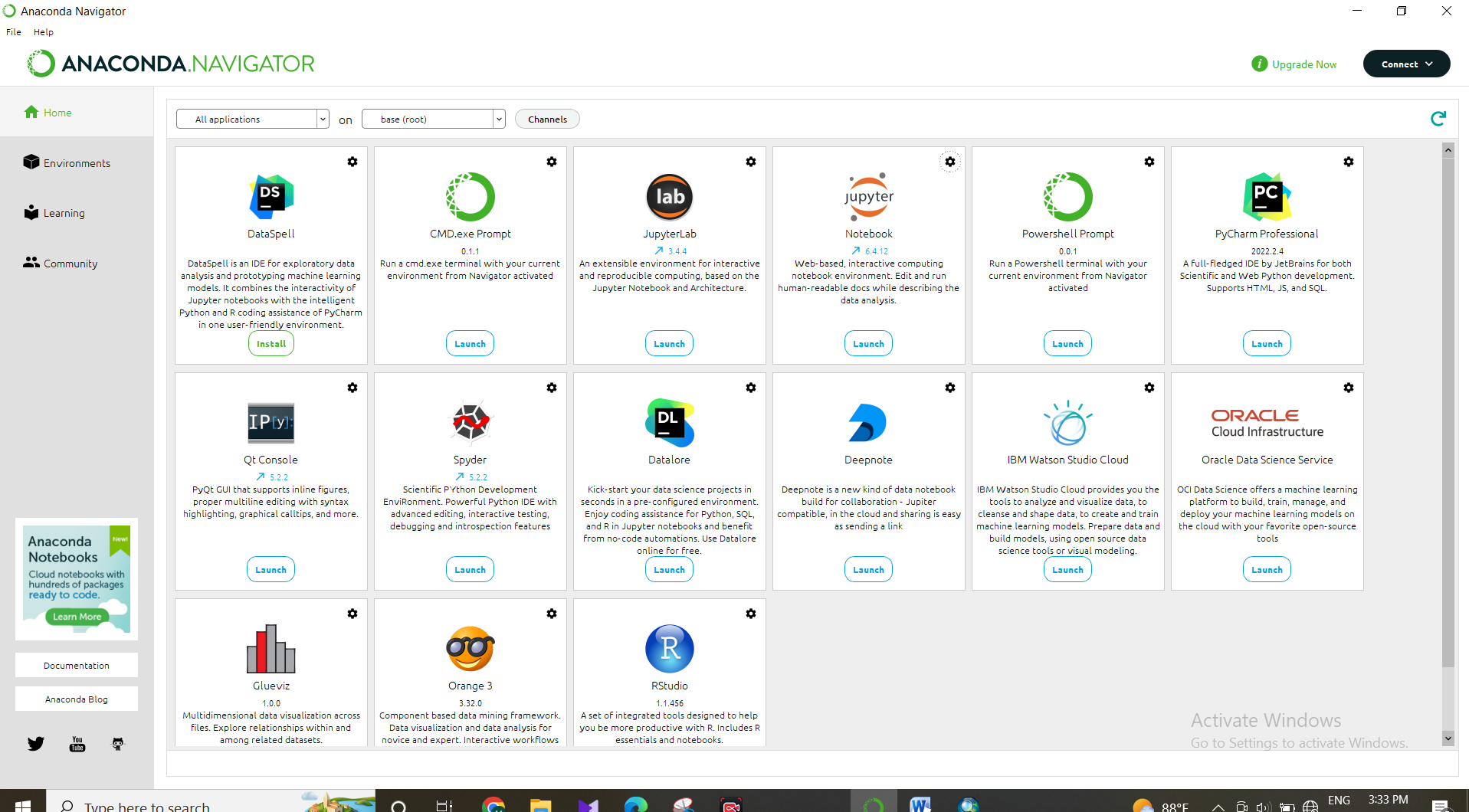
# Methodology and step wise results

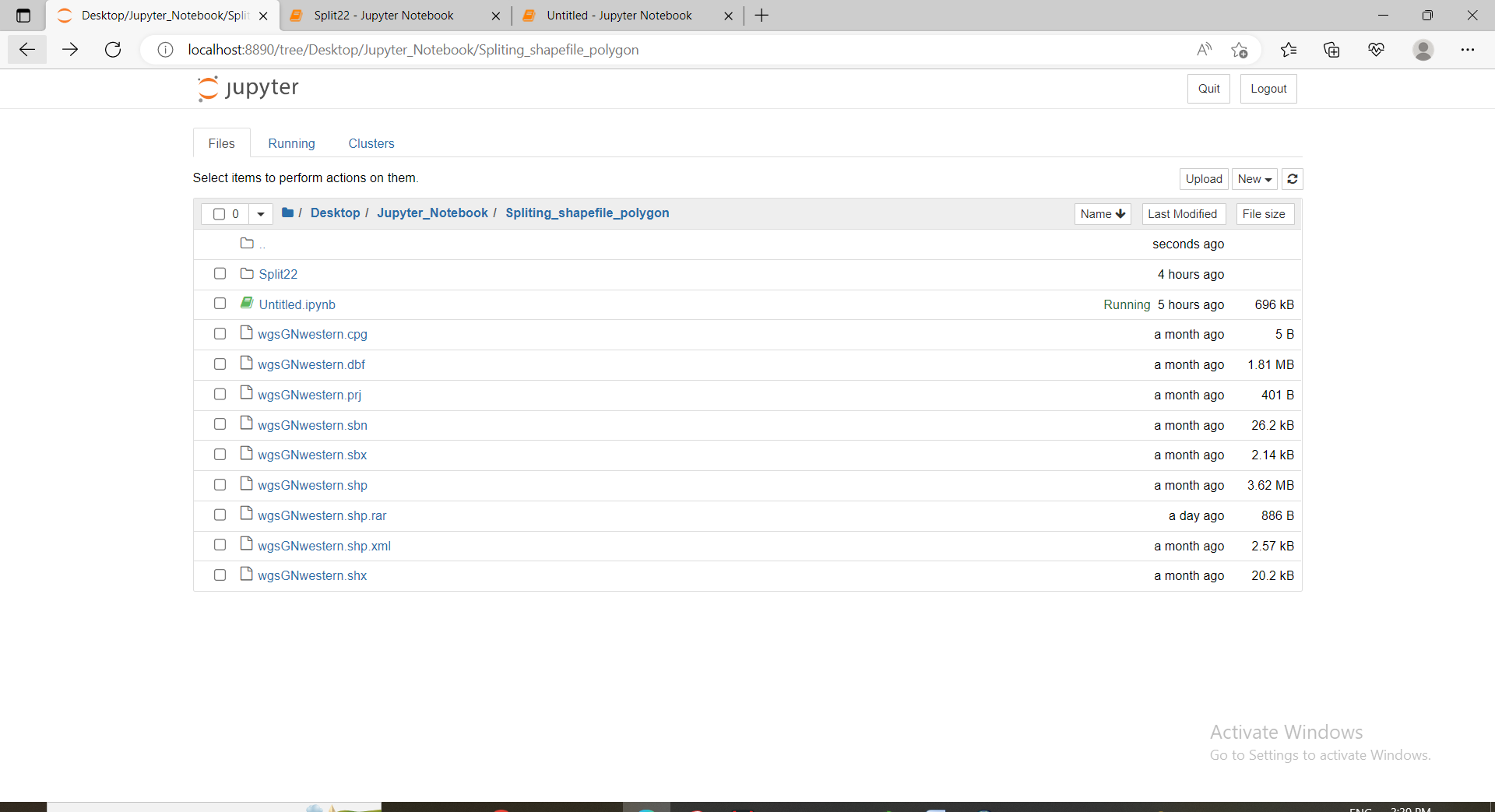


It will use the geopandas library to work with vector data in Python. It will also use matplotlib.pyplot to plot data. It can use the GeoPandas function read file to import shapefiles (). Keep in mind that it instruct Python to search for the read file () function in the geopandas library by calling the function with gpd.read file ().

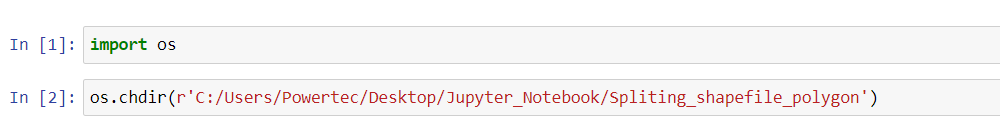
Below it saw how to quickly plot shapefiles using geopandas plotting. The geopandas plotting is a great option for quickly exploring our data. However it is less customizable than matplotlib plotting. Below it will learn how to create the same map using matplotlib to setup the axes and how to split in to shapefiles as well.

Here western province shape file had used to split in to shape files for several districts. Shape file contain with 2515 of GN divisions and it is easy to use find the district and split in to district shape files by using python script.

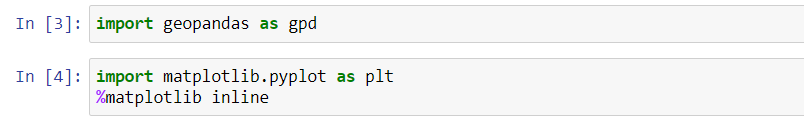
* Firstly Anacondar Navigator was opened.
* Launch the Jupytor Notebook and create a working folder.



* Locate the folder and mention the working folder that all files and everything locate in this specific folder.



* Then import geopandas and matplotlib packages.



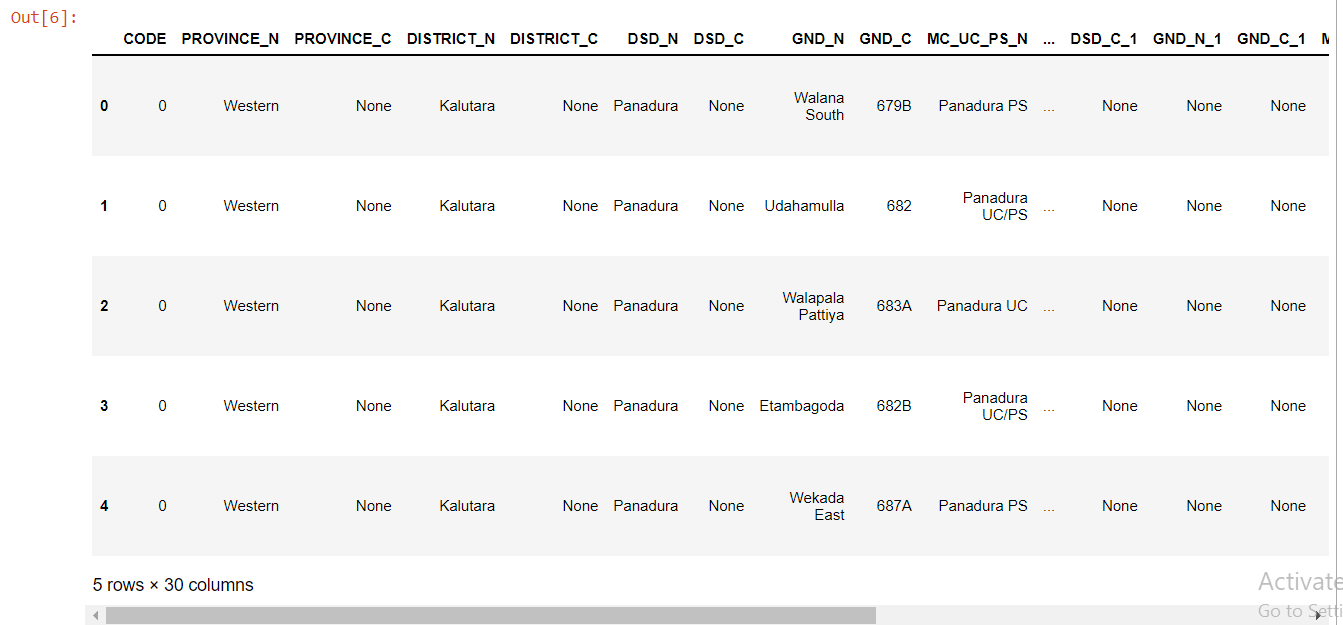
* Call to the geopandas and read the shapefile.



* Take a look at to the first 5 rows of the attribute table.



**Output of this:**



## Make the plot with legend.

## 

## Output of this:

## 

## The shape files contained specific border boundaries for the district shape files that Kalutara, Colombo, Gampaha. We had gone through each of the border boundaries.

## 

## Output of this:

## 

## It has specified the length of the shape file.

## 

## Output of this:

## 

## Next one has started with for loop to specify the length of the shape file.

## 

## Output of this:

## 

## By using the first row which is the zero was taken.

## 

## Output of this:

## 

## Put a name to the file in order to name the shapefile. Index is i.

## 

## Output of this: Getting split shapefiles to given specified folder path. It contains specific polygons.

## 

## Then use the command glob in order to get all the shapefiles. And put the folder that I created which are boundaries.

## 

## 

## Get all the shapefiles or all the files that ended in (.shp) extention from the folder.

## 

## Output of this:

## 

## By using for loop basically plot the each of polygons.

## 

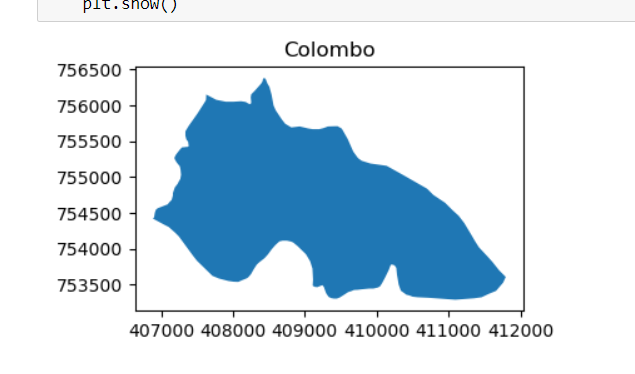
Chapter 3

# Evaluation of Project Result

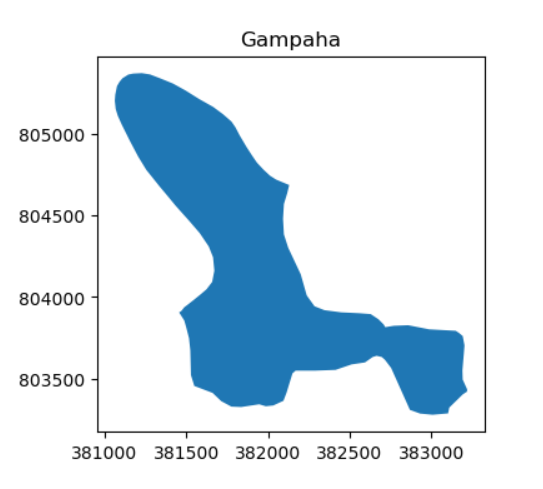


The project's results are shown and discussed in the order that they were output by the Python script. For easy identification, some outcomes of coding are mentioned above with the coding also.

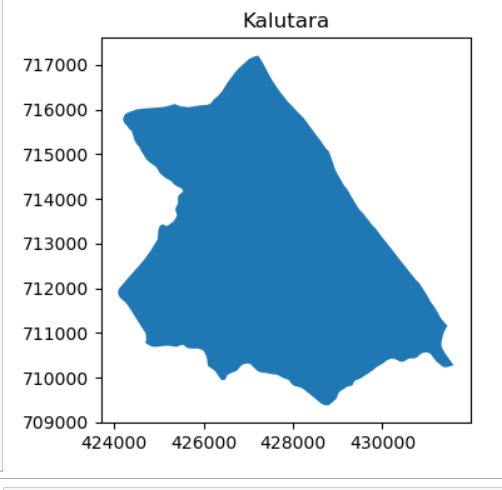
* **Splited Colombo shapefile.**



* **Splited Gampaha shapefile.**



* **Splited Kalutara shapefile.**



Chapter 4

# Conclusion and Future Developments

This project will only help to split the shapefiles. It is possible to access from anywhere in the world from any device. In this work its need to install some packages. The python script was helped to develop this program using Jupytor notebook.

The future development of these tools can be directed to many paths. It would provide an aesthetic support to the user, if they prefer alike. On top of all, if the two plugins are made standalone programs, which have the ability to run in different operating systems and mobile platforms, it would benefit the most.

Chapter 6

# Appendix



In this coding attempt, I learned about the Jupytor Notebook specific python classes and their methods, parameters in the methods and how to implement them in a script.

## Appendix

import os

os.chdir(r'C:/Users/Powertec/Desktop/Jupyter\_Notebook/Spliting\_shapefile\_polygon')

import geopandas as gpd

import matplotlib.pyplot as plt

%matplotlib inline

Shapefile=gpd.read\_file('C:/Users/Powertec/Desktop/Jupyter\_Notebook/Spliting\_shapefile\_polygon/wgsGNwestern.shp')

shapefile.head()

fig, ax = plt.subplots(figsize = (18,10))

shapefile.plot(ax =ax, column = 'DISTRICT\_N',legend = True, legend\_kwds = {'loc':'upper left'})

plt.show()

shapefile.loc[[0]]

len(shapefile)

for i in range(len(shapefile)):

print(i)

shapefile.loc[0,'DISTRICT\_N']

for i in range(len(shapefile)):

# print(i)

name = shapefile.loc[i,'DISTRICT\_N']

shapefile.loc[[i]].to\_file('C:/Users/Powertec/Desktop/Jupyter\_Notebook/Spliting\_shapefile\_polygon/Split22/Boundaries/'+name+'.shp')

from glob import glob

shapefile\_names=glob('C:/Users/Powertec/Desktop/Jupyter\_Notebook/Spliting\_shapefile\_polygon/Split22/Boundaries/\*.shp')

shapefile\_names

for i in shapefile\_names:

shp = gpd.read\_file(i)

fig, ax = plt.subplots(figsize = (4,4))

shp.plot(ax=ax)

plt.title(shp.loc[0,'DISTRICT\_N'])

plt.show()