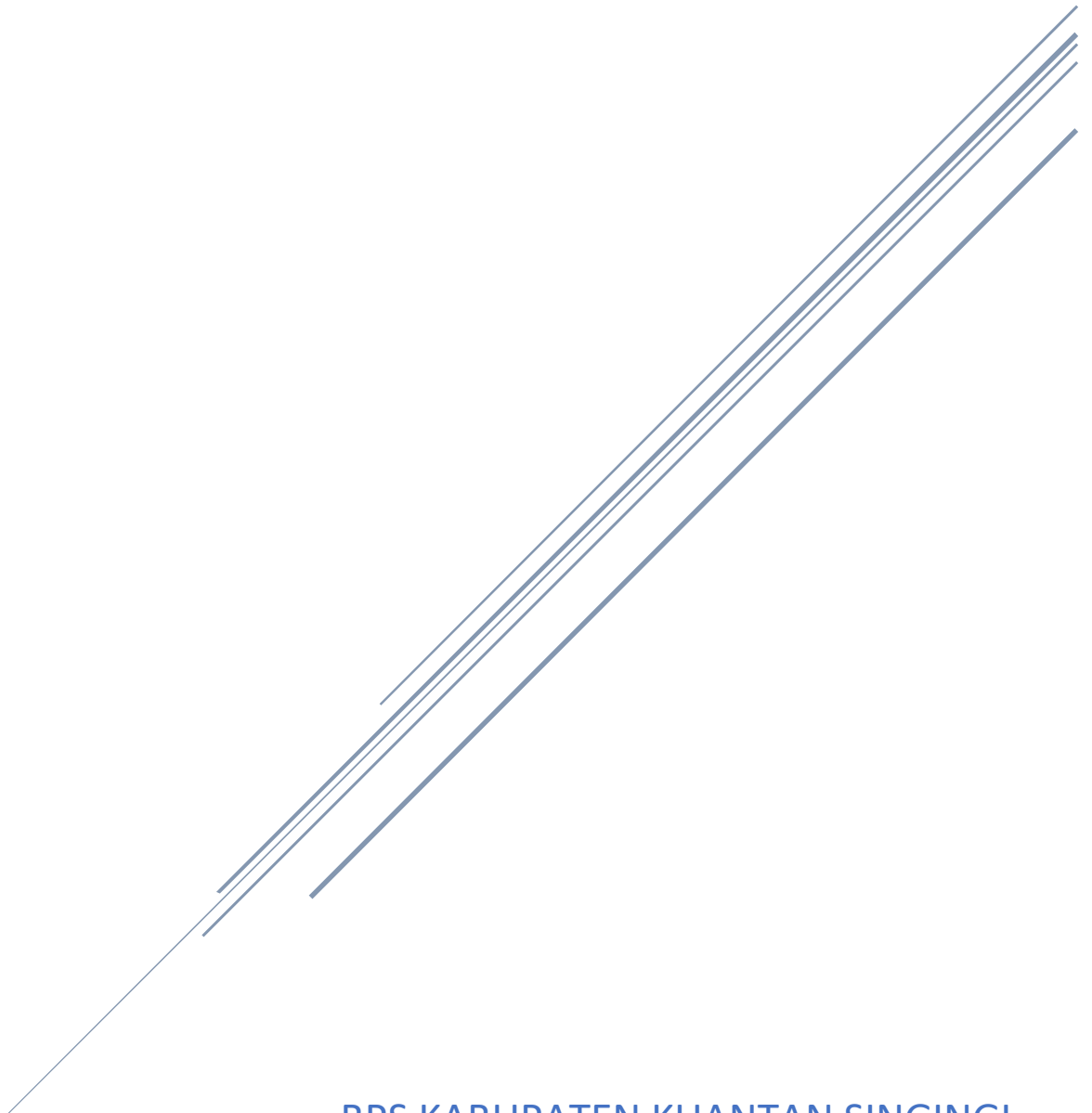


III.B.6. LAPORAN MANIPULASI DATA

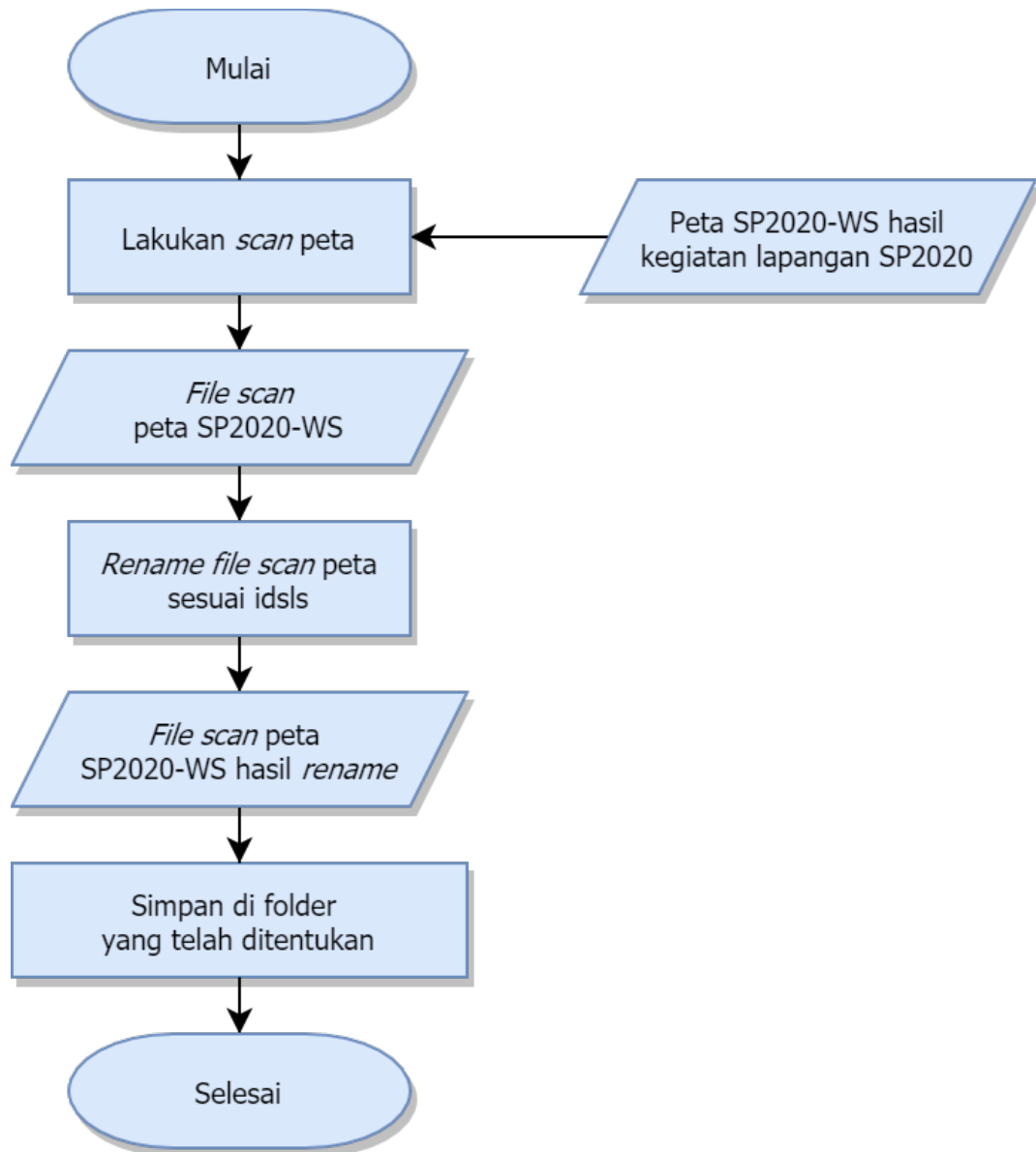
Rename File Hasil Scan Peta SP2020-WS



BPS KABUPATEN KUANTAN SINGINGI
Januari-Maret 2021

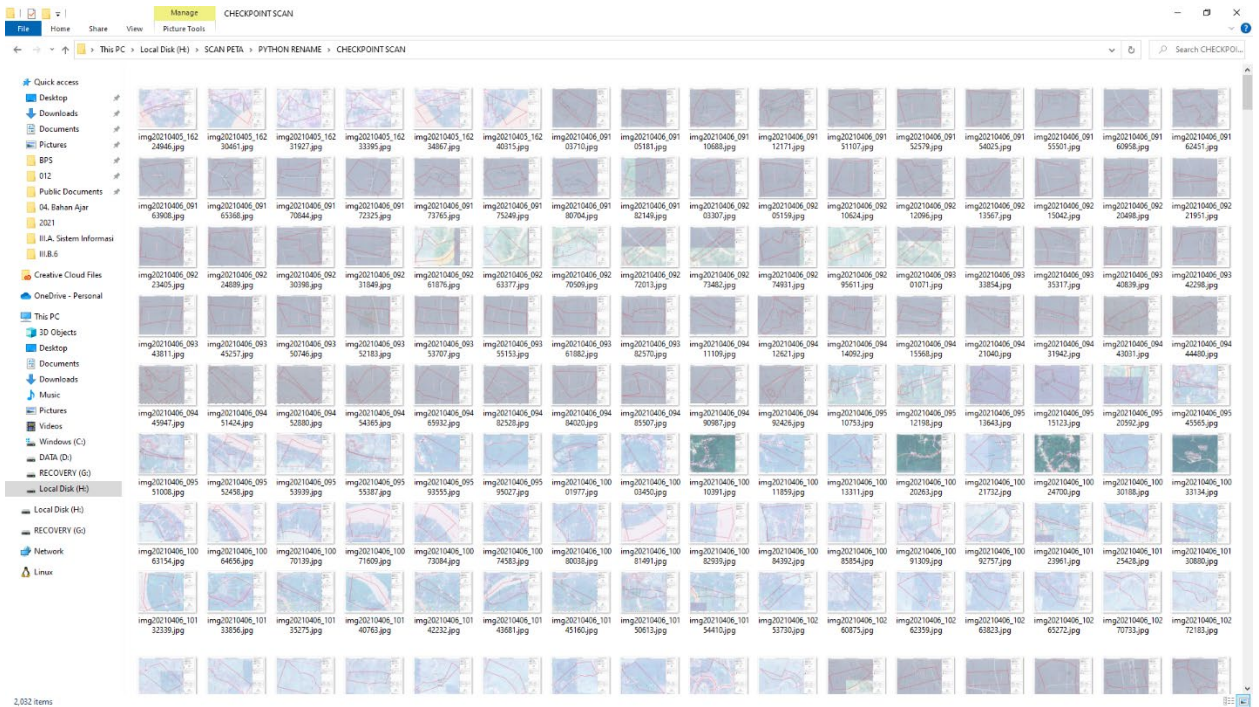
A. Data Asal

Sensus Penduduk (SP2020) menghasilkan peta SP2020-WS hasil lapangan yang dilengkapi titik bangunan, nomor bangunan, dan perbaikan batas SLS (jika ada). Untuk merubah peta kertas tersebut menjadi peta digital, tahapan pertama yang harus dilakukan adalah melakukan scan terhadap seluruh peta tersebut.



Gambar 1. Alur Proses Scan Peta SP2020-WS

Hasil dari kegiatan *scan* tersebut berupa *file scan* peta SP2020-WS yang penamaan *file*-nya masih acak.



Gambar 2. Hasil *Scan* Peta SP2020-WS

B. Kegiatan Manipulasi Data

File hasil *scan* peta SP2020-WS masih memiliki nama *file* yang acak sesuai output dari *scanner*. Sedangkan untuk lanjut ke tahap georeferensi, penamaan file harus berdasarkan idsls. Ada berbagai cara untuk melakukan *rename file* hasil *scan*. Salah satu *tool* yang dapat digunakan adalah menggunakan *optical character recognition* (OCR) seperti tesseract dan bahasa pemrograman python. Kelebihan tool ini dibanding tool seperti Bulk Rename Utility adalah *rename file* dapat dilakukan secara otomatis oleh sistem dengan membaca idsls yang ada disudut kanan atas *file scan* peta SP2020-WS.

C. Hasil Manipulasi Data

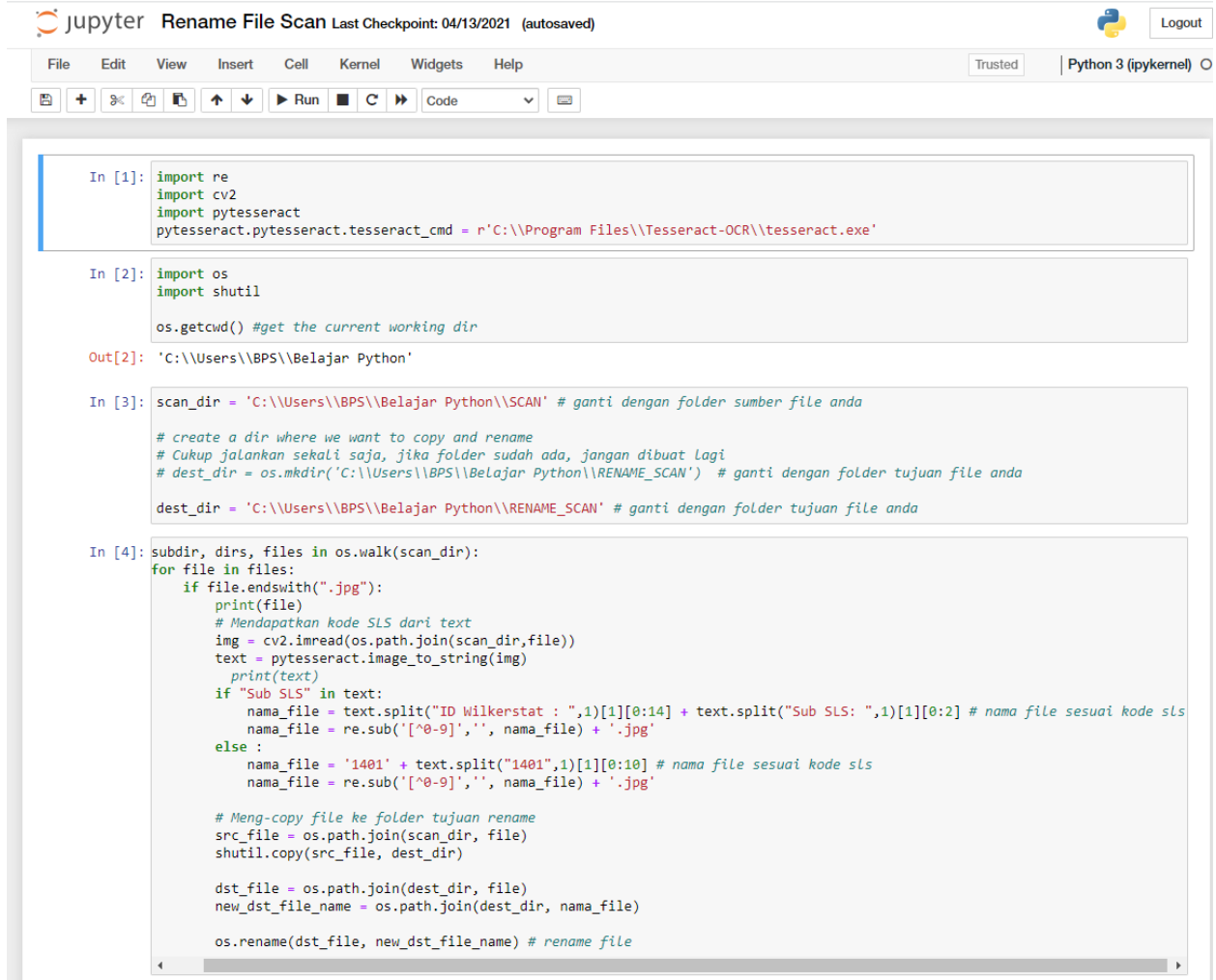
Hasil dari kegiatan manipulasi data adalah *file scan* peta SP2020-WS yang nama *file*-nya sudah berupa idsls. Terdapat 2061 *file scan* peta SP2020-WS yang nama *file*-nya telah berhasil di-*rename* untuk dapat dijadikan *input* tahapan selanjutnya yaitu georeferensi.



Gambar 3. Hasil *Rename File Scan* Peta SP2020-WS

D. Script/Command yang Digunakan

Script yang digunakan untuk melakukan *rename file scan* peta SP2020-WS menggunakan python dan library pyesseract, re, dan cv2. *Script* lengkap dapat diakses di *file* Rename File Scan.ipynb yang sudah kami lampirkan.



Jupyter Rename File Scan Last Checkpoint: 04/13/2021 (autosaved) Python 3 (ipykernel)

```
In [1]: import re
import cv2
import pytesseract
pytesseract.pytesseract.tesseract_cmd = r'C:\\Program Files\\Tesseract-OCR\\tesseract.exe'
```

```
In [2]: import os
import shutil

os.getcwd() #get the current working dir
```

```
Out[2]: 'C:\\Users\\BPS\\Belajar Python'
```

```
In [3]: scan_dir = 'C:\\Users\\BPS\\Belajar Python\\SCAN' # ganti dengan folder sumber file anda

# create a dir where we want to copy and rename
# Cukup jalankan sekali saja, jika folder sudah ada, jangan dibuat lagi
# dest_dir = os.mkdir('C:\\Users\\BPS\\Belajar Python\\RENAME_SCAN') # ganti dengan folder tujuan file anda

dest_dir = 'C:\\Users\\BPS\\Belajar Python\\RENAME_SCAN' # ganti dengan folder tujuan file anda
```

```
In [4]: subdir, dirs, files in os.walk(scan_dir):
for file in files:
    if file.endswith(".jpg"):
        print(file)
        # Mendapatkan kode SLS dari text
        img = cv2.imread(os.path.join(scan_dir,file))
        text = pytesseract.image_to_string(img)
        print(text)
        if "Sub SLS" in text:
            nama_file = text.split("ID Wilkerstat : ",1)[1][0:14] + text.split("Sub SLS: ",1)[1][0:2] # nama file sesuai kode sls
            nama_file = re.sub('[^0-9]', '', nama_file) + '.jpg'
        else :
            nama_file = '1401' + text.split("1401",1)[1][0:10] # nama file sesuai kode sls
            nama_file = re.sub('[^0-9]', '', nama_file) + '.jpg'

        # Meng-copy file ke folder tujuan rename
        src_file = os.path.join(scan_dir, file)
        shutil.copy(src_file, dest_dir)

        dst_file = os.path.join(dest_dir, file)
        new_dst_file_name = os.path.join(dest_dir, nama_file)

        os.rename(dst_file, new_dst_file_name) # rename file
```

Gambar 4. Script Rename File Scan Peta SP2020-WS