Address Linkage Key with AnalyticsIQ
Milestone 1 Presentation

Project Owners :

Warren Smith

Jin Wang

Project Team :

Kat McElveen, Team Lead

Saritha Gudala, Technical Writer

Kat Greer, Technical Specialist

Byron Smith, Technical Specialist

IT7993 Capstone, Spring 2021

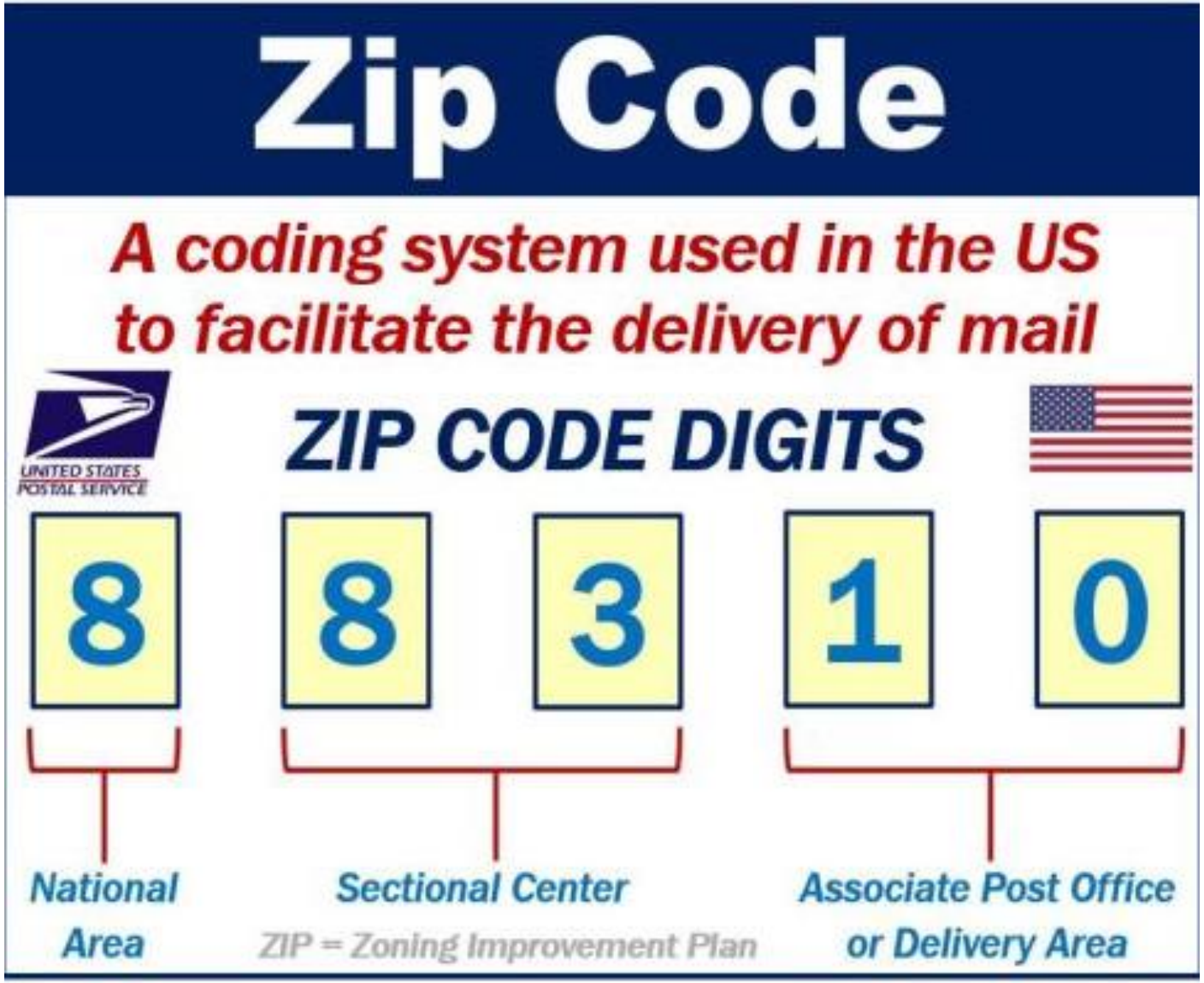
February 26, 2021

Project Overview

- Develop a tool/set of tools to compute the delivery point (DPBC) of a given address following the guidelines of USPS.
- The tool will be initially developed and tested for properly formatted addresses.
- As part of further improvements, it will be tested for improper address format and/or for the presence of certain types of delivery points such as high-rise buildings.

What is Minimally Viable Product?

- A tool or set of tools which allow me to compute the delivery point code of a given address given:
  1) A “clean” address (as defined by publication 28)
  2) ZIP and ZIP+4 appended
  3) A dataset which contains Highrise addresses



Project Progress Summary

- Assess the overall project progress in terms of the final deliverables specified in the project plan. Summarize overall accomplishments as a group.
  - Project team had initial kick off meeting to gather requirements.
  - Team members reviewed USPS publications to become familiar the USPS guidelines.
  - Team members installed and configured software to be able to work with the data collaboratively.
  - Team members began working with provided data and USPS guidelines to create the tool(s).

Milestone 1 Deliverables Status Update

Complete:

- Research USPS address formatting and delivery point barcode (DPBC) assignment rules
- Analysis of tools
- Setup and installation of remote tools and development environment

In Progress:

- General rule #1 development

Changed:

- None

Delivery Point Barcode Rules (Primary)

1. General Rule

Address: 1234 MAIN ST (PO BOX 44, RR 1 BOX 154, HC 1 BOX 1264)
DPBC: 34 (44, 54, 64)

Use last two digits. Print code characters in DPBC representing last two digits of primary street number (or post office box, rural route box, or highway contract route number).

2. No Numbers

Address: MAIN St (RR 1, HC 1)
DPBC: 99 (99, 99)

Use 99. Print code characters in DPBC representing last two digits of primary street number (or PO Box, rural route, or highway contract route number).

3. Single Digits

Address: 8 MIAN St (PO BOX 1, RR 1 BOX 2, HC 1 BOX 3)
DPBC: 08 (01, 02, 03)

Add leading zero. Print code characters in DPBC representing leading zero and single digit.

8. Leading/Embedded Alphas

Address: 23S41 MAIN ST (23S4 MAIN ST, 2W3S1 MAIN ST, MAINS ST, C8INT)
DPBC: 11 (04, 01, 01)

Print code characters in DPBC representing last two digits to right of alphas. If single digit to right of alphas, add leading zero.

9. Slashes (/)

Address: 123/4 MAIN ST (PO BOX ¼, RR 1 BOX 123/124/125, H 3 BOX 11/13)
DPBC: 23 (03, 23, 07)

Print code characters in DPBC representing 99 whenever a slash appears directly next to numeric in the primary street number.

10. Other Embedded Symbols

Address: 1.23 MAIN ST (PO BOX 1-3, RR 1 BOX 1.23, HC 3 BOX 11*7)
DPBC: 23 (03, 23, 07)

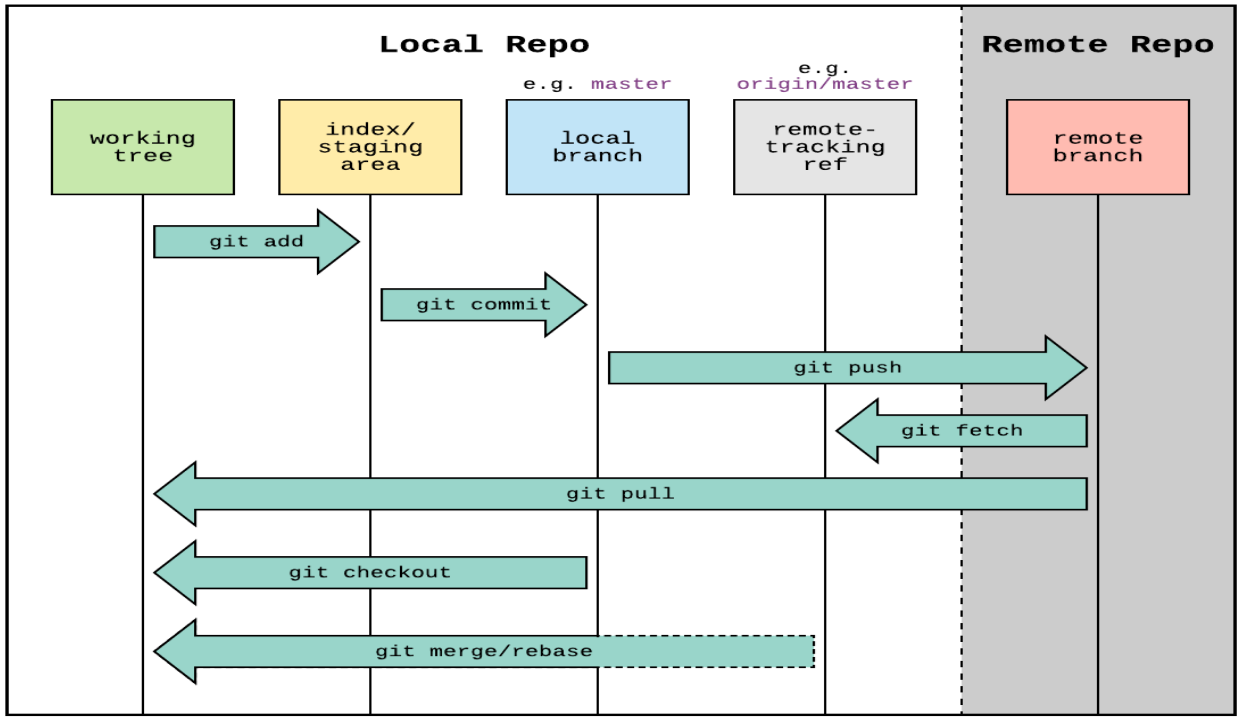
Use last two digits to right of the symbol. Print code characters in DPBC representing last two digits to the right of all symbols (except slashes), such as periods and hyphens appearing in primary street numbers. If single digit to right, add leading zero.

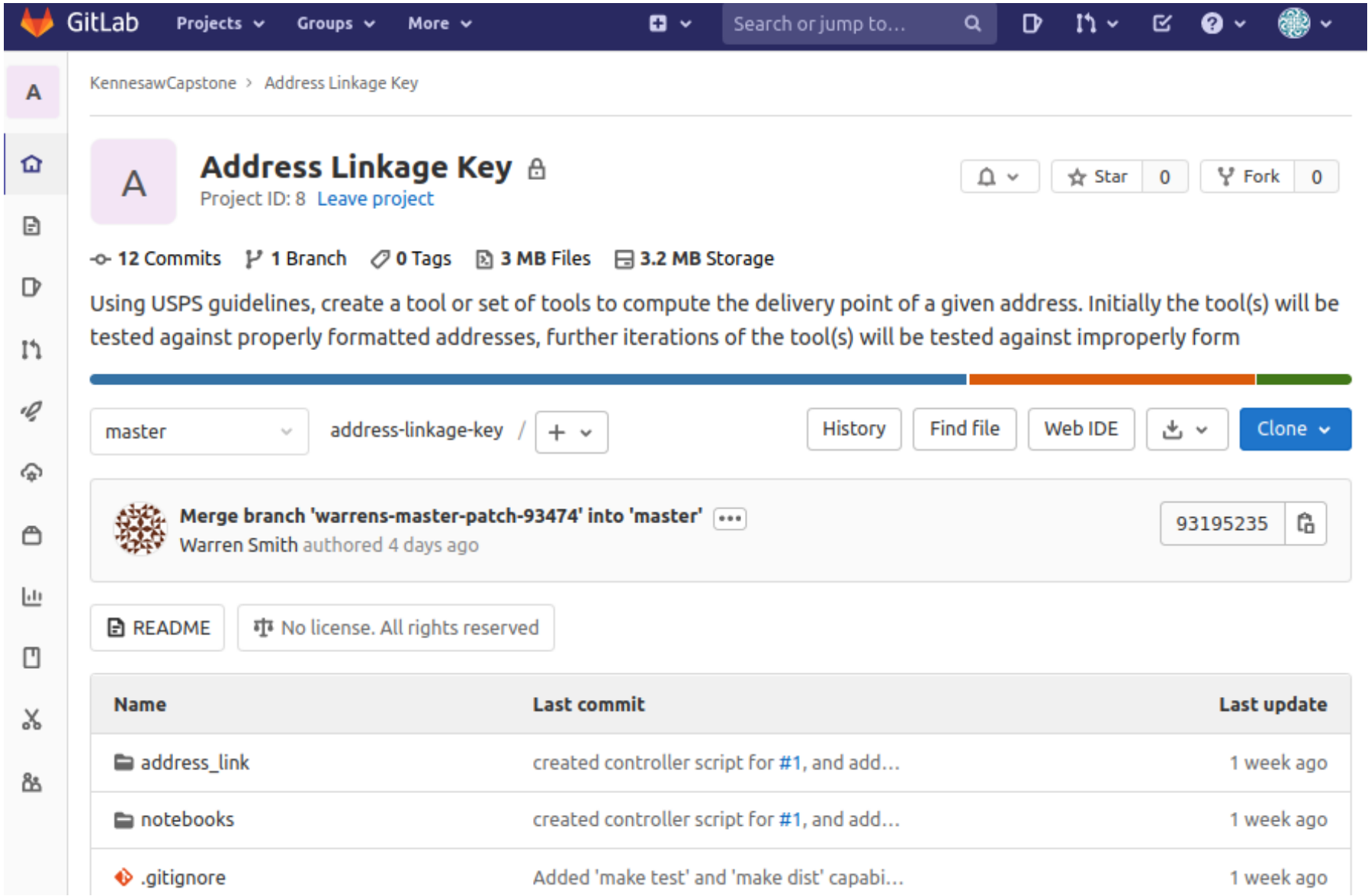
Required Tools and Technology

- Git
- Tortoise Git
- Putty Suite (SSH key generation)
- GitLab
- Docker
- Jupyter Notebook
- PySpark

Getting Connected – Source Control

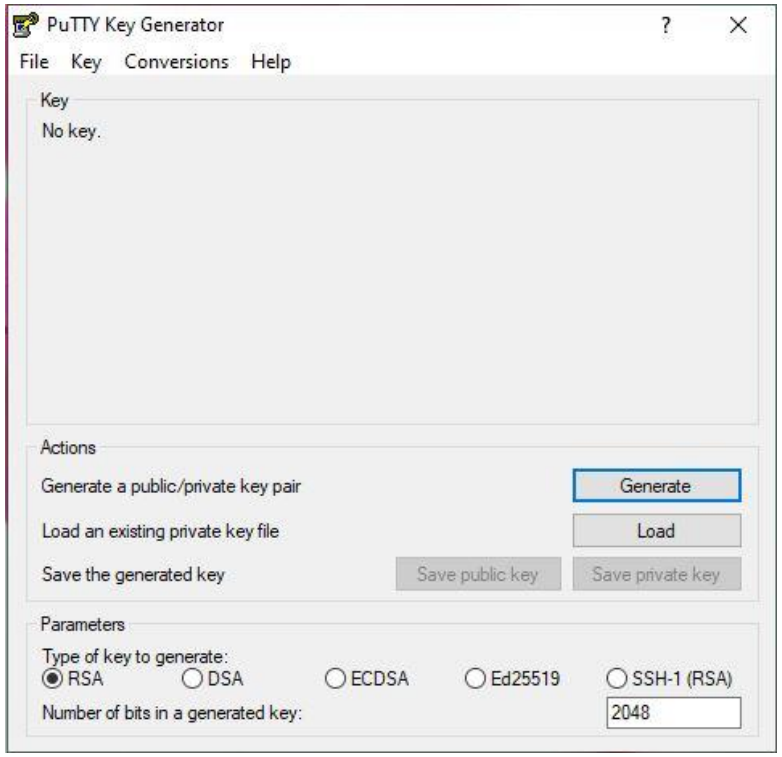
- Git is a distributed version-control system for tracking changes in any set of files.
- Git is built for the world's largest software implementations (Linux).
- As part of project implementation, we are using a subset of git features.
- TortoiseGit is a Windows UI on top of Git.
- GitLab is a Web UI on top of Git.




GitLab
Projects
Groups
More
Search or jump to...
KennesawCapstone › Address Linkage Key
A
Address Linkage Key
Project ID: 8 Leave project
Star 0
Fork 0
12 Commits
1 Branch
0 Tags
3 MB Files
3.2 MB Storage
Using USPS guidelines, create a tool or set of tools to compute the delivery point of a given address. Initially the tool(s) will be tested against properly formatted addresses, further iterations of the tool(s) will be tested against improperly form
master
address-linkage-key /
History
Find file
Web IDE
Clone
Merge branch 'warrens-master-patch-93474' into 'master'
Warren Smith authored 4 days ago
93195235
README
No license. All rights reserved
Name
Last commit
Last update
address_link
created controller script for #1, and add...
1 week ago
notebooks
created controller script for #1, and add...
1 week ago
.gitignore
Added 'make test' and 'make dist' capabi...
1 week ago

Steps to generate keys using putty:

- Open the Putty Key Generator.
- Select “RSA” under type of keys to generate
- Select the “Generate” button under the actions pane.
- Save the keys by navigating to file option.



PySpark Jupyter Notebook

- Option 1: Docker container to initialize a PySpark Jupyter Notebook locally
- Option 2: PySpark Jupyter Notebook on JupyterHub server

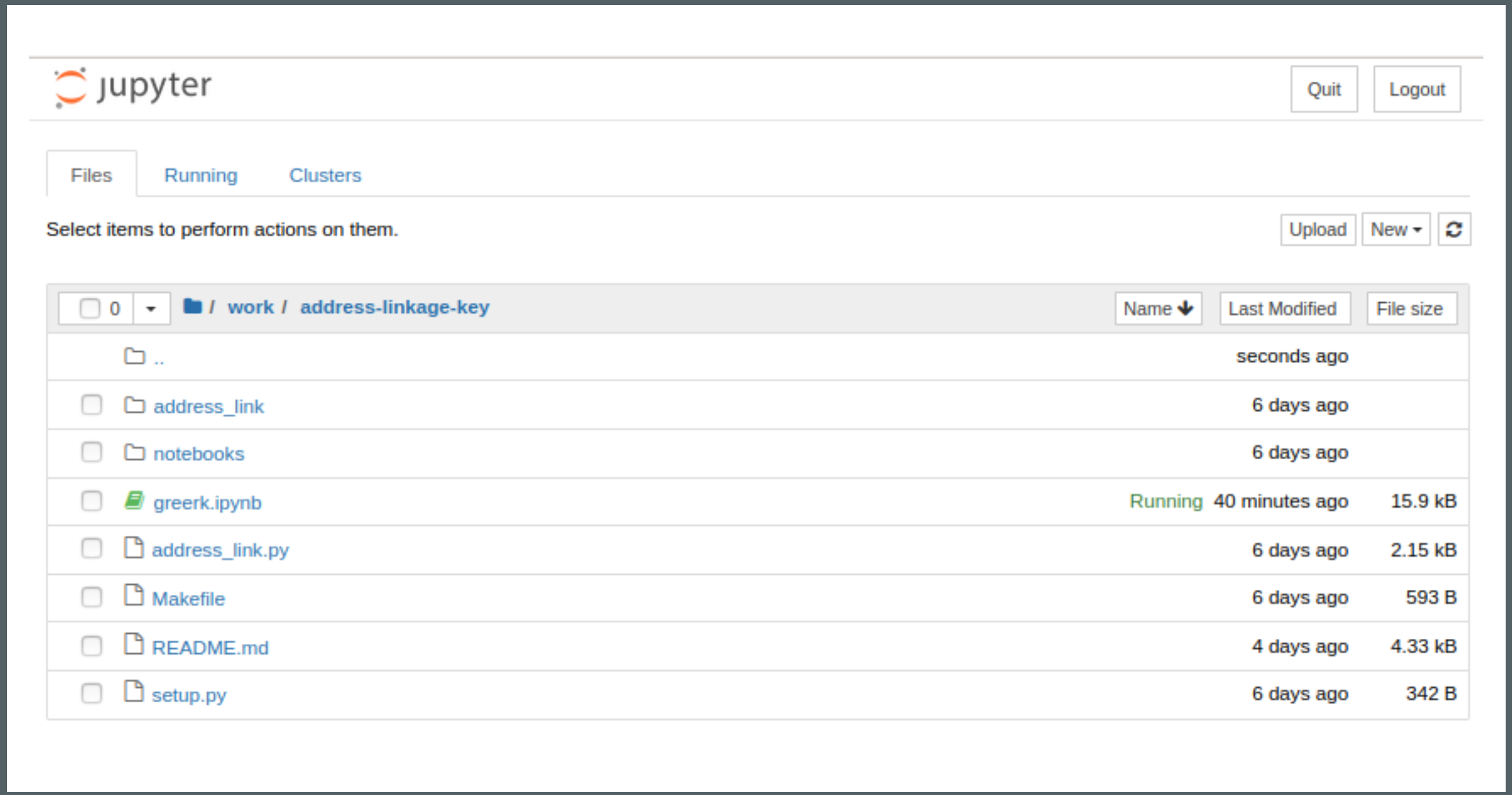
Steps for starting docker container:

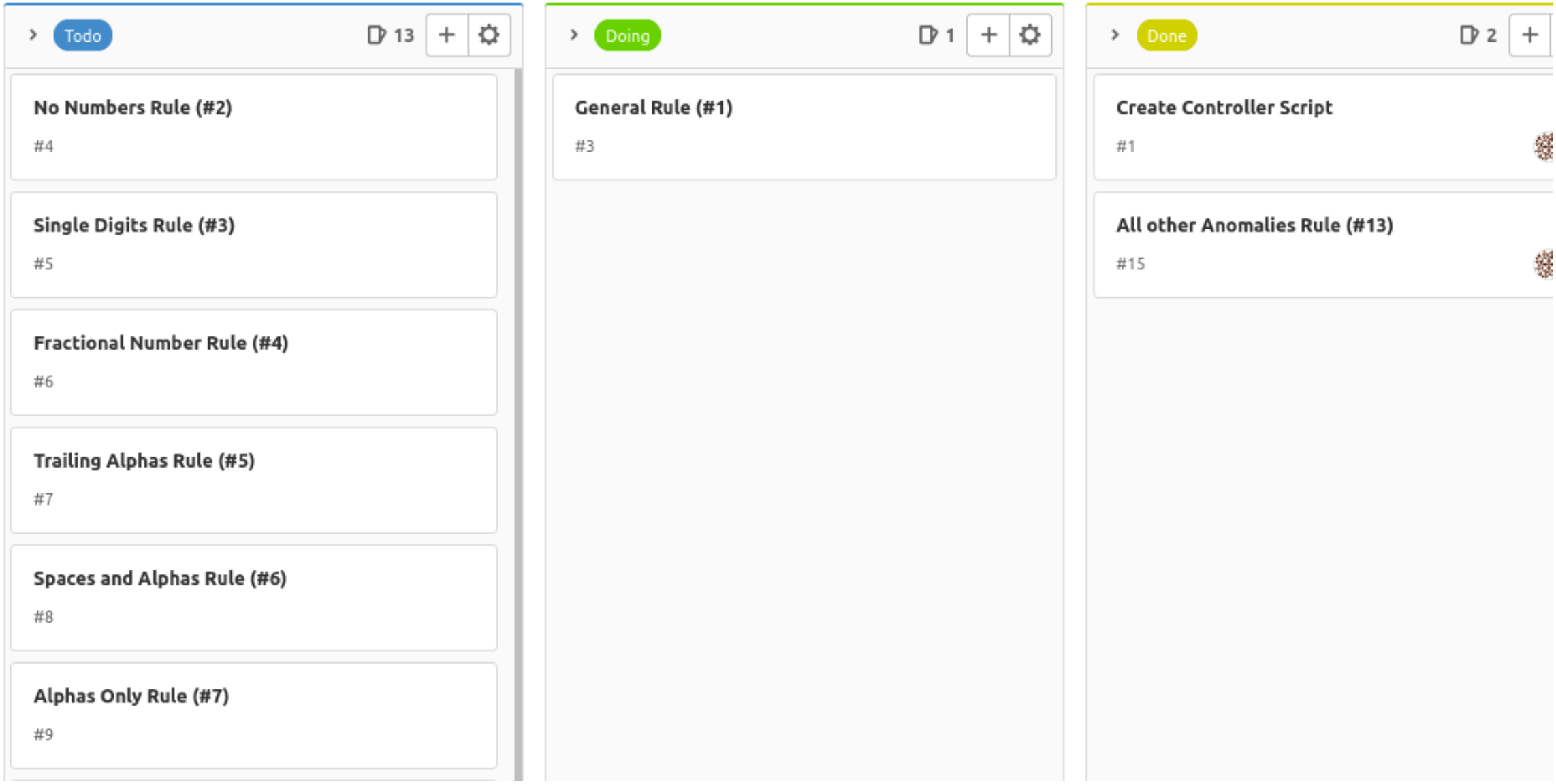
Link to access for installation : https://docs.docker.com/docker-for-windows/install/

Setting the path to required directory.

Ex : C:\Users\Address LinkageKey

Code to run docker: docker run -it --rm -p 8888:8888 -v ${PWD}:/home/jovyan/work jupyter/pyspark-notebook


jupyter
Quit
Logout
Files
Running
Clusters
Select items to perform actions on them.
Upload
New
0
/ work / address-linkage-key
Name
Last Modified
File size
..
seconds ago
address_link
6 days ago
notebooks
6 days ago
greerk.ipynb
Running 40 minutes ago
15.9 kB
address_link.py
6 days ago
2.15 kB
Makefile
6 days ago
593 B
README.md
4 days ago
4.33 kB
setup.py
6 days ago
342 B


Todo
13
No Numbers Rule (#2)
#4
Single Digits Rule (#3)
#5
Fractional Number Rule (#4)
#6
Trailing Alphas Rule (#5)
#7
Spaces and Alphas Rule (#6)
#8
Alphas Only Rule (#7)
#9
Doing
1
General Rule (#1)
#3
Done
2
Create Controller Script
#1
All other Anomalies Rule (#13)
#15

Reflections

- Getting acquainted to the CASS functionality of USPS to develop a tool/tools to determine DPBC (delivery point bar code).
- Gaining hands-on with new tools/technologies Git, PySpark, Putty generator for generation of SSH keys.
- Some of the tools are familiar and some are not so familiar; need to apply learned knowledge to better grasp new concepts and tools.
- After seeing the complexity and steps involved for Warren to configure out setup, access, security and testing environment it makes sense as to why the DevOps field is growing so rapidly. There are a lot of moving parts and it is very complex.

Milestone 2 Plans

- Review issues list and assign rules/tasks to team members.
- Aiming to develop a minimally viable product by end of Milestone 2.
- Expecting tool/tools will determine at least the delivery point bar code (DPBC) of a properly formatted address.
- Continue meeting weekly with AnalyticsIQ.

REFERENCES

1. USPS Publication 28 (Postal Addressing Standards)
2. USPS CASS Technical Guide

THANK YOU