

# Web scraping and Google API Email Notification System

Mir Mohammed Ali

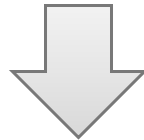
# What is the purpose of this project?

This was a fun little project I decided to put together because I wanted to experiment with different tools and how they clicked together. I enjoyed making this project because I learned so much about the different types of services mentioned in this project.

What is the MVP (minimum viable product) of this personal project?

---

Web scraping Project



The MVP is to get timely information from a website in the form of email notifications.

## Overview of Services Used to Setup and Accomplish Goal

---

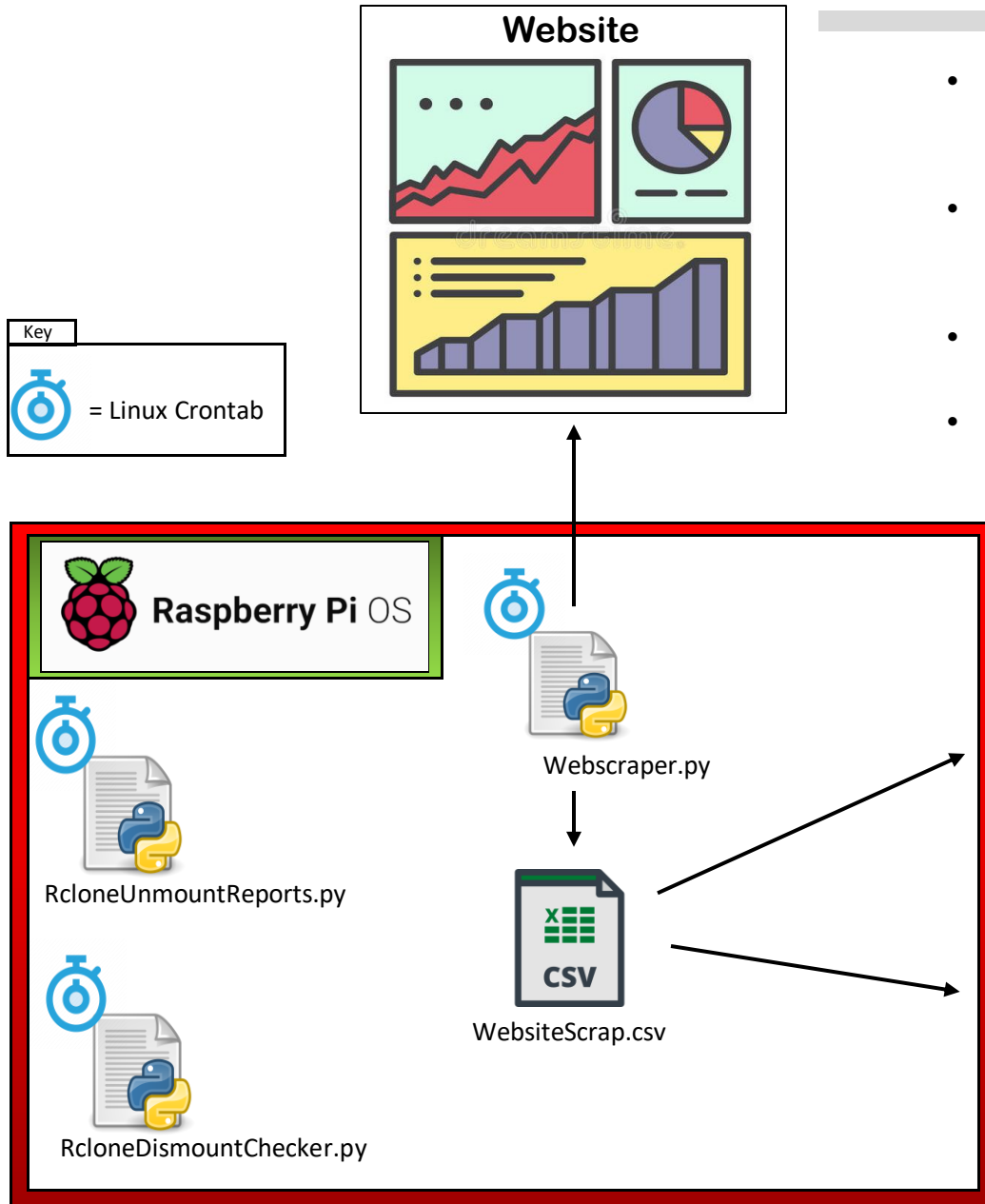


## Why use and setup these particular services?

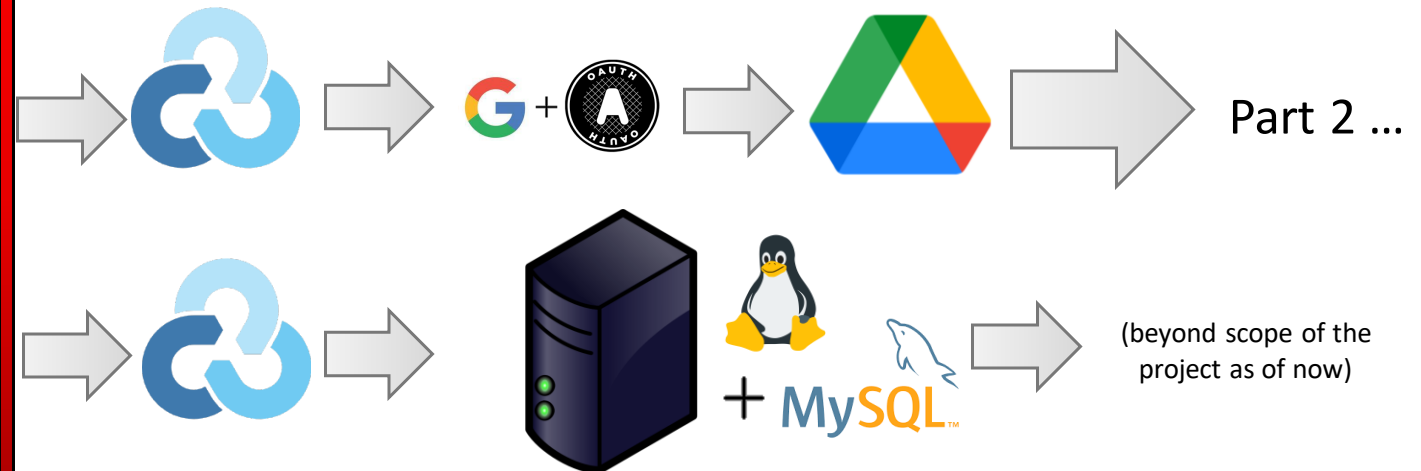
---

- Runs consistently as long as it scrapes websites sparingly
- Gives you a flexibility on how you want to manage your data
- Serves as a fun project to get exposure to different tools
- Uses Google Drive API which is free
- Gmail Push Notifications can reach many types of devices
- Easy to create and deploy as a personal automation

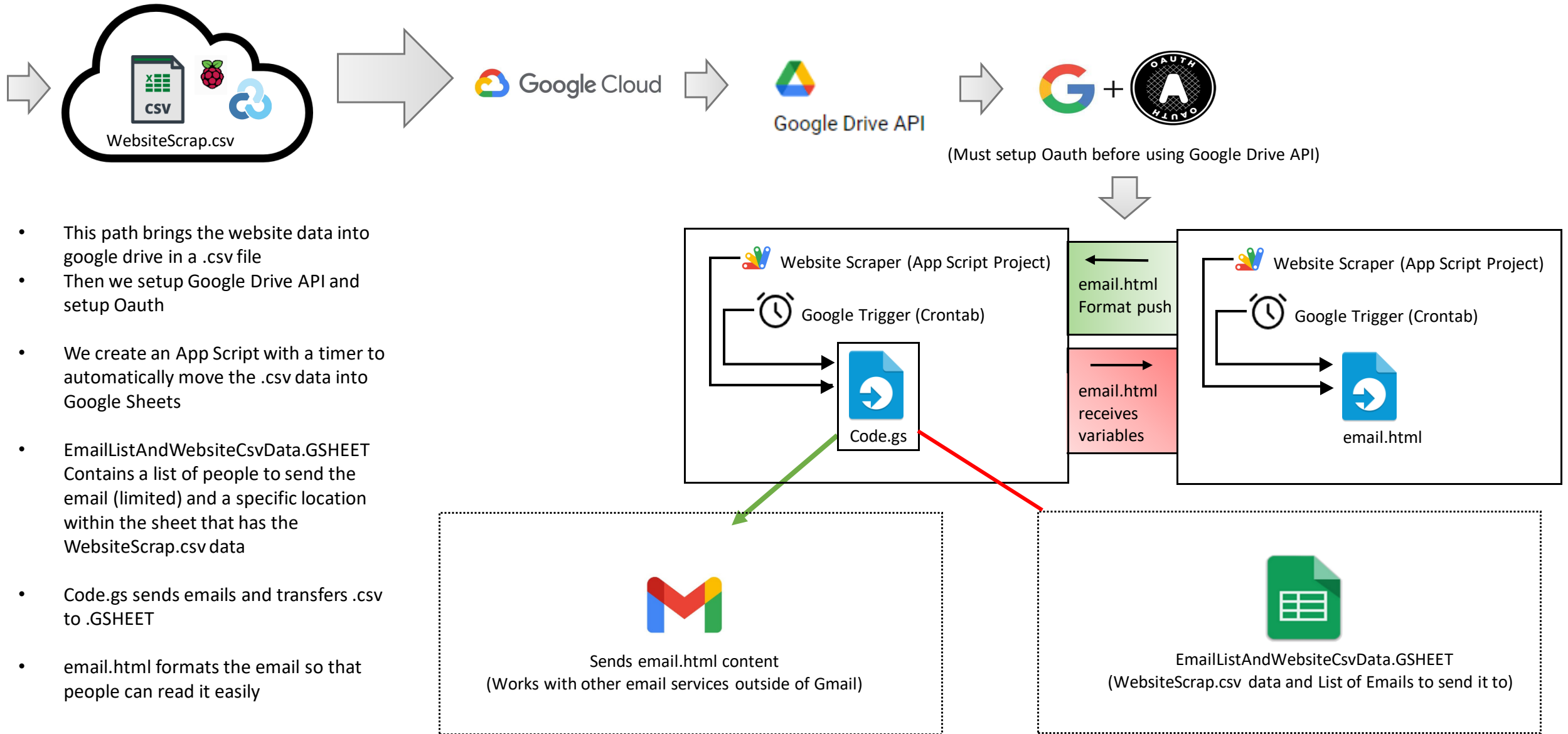
# How everything works together Part 1



- RcloneUnmountReports.py creates a report anytime Rclone disconnects with essential information for troubleshooting
- RcloneDismountChecker.py attempts to remounts Rclone whenever it is dismounted. It stops after three successive attempts of failure.
- Webscraper.py uses Pandas to gather tabular data which is saved in a .csv file
- (Future thoughts) We can even store past data in a home server using MySQL if we should keep track of the data for insights



# How everything works together Part 2



- This path brings the website data into google drive in a .csv file
- Then we setup Google Drive API and setup Oauth
- We create an App Script with a timer to automatically move the .csv data into Google Sheets
- EmailListAndWebsiteCsvData.GSHEET Contains a list of people to send the email (limited) and a specific location within the sheet that has the WebsiteScrap.csv data
- Code.gs sends emails and transfers .csv to .GSHEET
- email.html formats the email so that people can read it easily

## Conclusion

---

- This was a fun project that I had completed
- It works well as a small application for picking up simple consistent data (weather data, world records data, etc.)
- This was all experimental and was a process of trial and error it seems to work well, but it can always improve
- It relies heavily on several services working properly and is more like glue from point to point
- Great for small personal web scraping projects