Read
API Covid
OWID

OWID

Amazon S3

A - *coviddaily.csv*

A - *coviddaily.csv*   B - *top10Data.csv*

B - *top10Data.csv*   C - *top10Clusters.csv*

A
.CSV

Save...
A-covidDaily.csv

Load...
A-covidDaily.csv

A
.CSV

B
.CSV

Save...
B-top10Data.csv

Load...
B-top10Data.csv

B
.CSV

C
.CSV

Save...
C-top10Clusters.csv

Execute
py Script

.py

Execute
R Script

.R

Execute
R Script

.R

Heroku Pipeline
Schedule 1
Read API and Save CSV to S3

Script
Read API COVID

1) ***Read API***
3) Transform to ***CSV***
4) Save ***CSV*** in S3 = ***A-covidDaily.csv***
5) Look for ***errors*** and ***send E-mail***

Heroku Pipeline
Schedule 2
Read CSV + Clean
+ Filter + Save CSV

Script PRE-PROCESSING

1) Read ***A-covidDaily.csv***
2) Clean + add extra info + filter top 10
3) Save ***B-top10Data.csv***
4) Look for ***errors*** and ***send E-mail***

Heroku Pipeline
Schedule 3
Read CSV + Profiling
+ Cluster + Save CSV

Script PROFILING + CLUSTERS

1) Read ***B-top10Data.csv***
2) Profiling + Clusters
3) Save ***C-top10Clusters.csv***
4) Look for ***errors*** and ***send E-mail***

.CSV   ***A - coviddaily.csv****:* Dinamic daily file with API daily covid data

.CSV   ***B-top10Data.csv:*** Covid Daily data clean, with extra info and top10 filtered.

.CSV   ***C-top10Clusters.csv:*** ...¿?