**Directions for setting up the Linux server to run this project**

git clone <https://github.com/Anderson-Lab/capstone-spring-2018-team-7.git> from the home directory, otherwise you will have to change many filepaths

Install Node.js

sudo apt-get install nodejs

sudo apt-get install npm

npm install google-play-scraper

npm install app-store-scraper

npm install command-line-args

Install MongoDB

<https://docs.mongodb.com/manual/tutorial/install-mongodb-on-debian/>

Install Python 2.7

Install pip

Install python packages

pip install pymongo

pip install pandas

pip install sys

pip install sklearn

pip install datetime

pip install json

pip install os

pip install re

pip install demjson

pip install bson

Install crontab with crontab -e

0 3 \* \* \* bash ~/capstone-spring-2018-team-7/collect\_data/pull\_apple.sh

5 3 \* \* \* bash ~/capstone-spring-2018-team-7/collect\_data/pull\_android.sh

0 5 \* \* \* mongod

1 10 \* \* \* python ~/capstone-spring-2018-team-7/import\_new\_json.py

55 14 \* \* \* python ~/capstone-spring-2018-team-7/findOutliers.py

0 15 \* \* \* bash ~/capstone-spring-2018-team-7/cayce.sh

Change absolute filepaths in python scripts

Line 323 of findOutliers.py

Line 15 and 38 of import\_all\_json.py

Line 11 and 35 of import\_new\_json.py

Line 11 of load\_stoplist.py

All other filepaths are relative to the home directory

To set up MongoDB with past data:

Run python import\_all\_json.py

Run python load\_stoplist.py

Login info for Google Cloud Server:

Username: [cofccapstoneteam7@gmail.com](https://mail.google.com/mail/?view=cm&fs=1&tf=1&to=cofccapstoneteam7@gmail.com)

Password: datascience

Use Instructions:

Click upper left drop down menu and go to VM instances.

Select the linuxvm option from the vm instances.

Click SSH