

# Document on how to run data migration script

This document covers how to run and setup the script that migrates data from the raw\_data landing zone into the analysis schema.

---

## setup

There is a provisions file kept with the script called provisions.sh that will install all package dependencies that are required on Linux Ubuntu 14.04 environment. It can be run as a shell script from the command line.

```
> ./provisions.sh
```

All software dependencies will be installed for you.

---

## Running the script

The script will be run by executing a python script with the following command

```
>python get_data.py
```

This job can be run on cron by issuing the following command in a crontab file

```
>crontab -e # to create cron file
```

### crontab file

This will run the get\_data.py script every minute and migrate new data to the analysis schema in the database.

```
* * * * * python get_data.py
```

## Test

The output of the landing table should be identical to a sql query in the same directory that the *getdata.py* containing folder is in called *mysql*. The output of this job should be the same as the query, except *packetid* and other fields that *get\_data.py* generates cannot be generated by SQL.