**DataVisor Batch Data Upload Guide**

1. Please make sure your technical account manager (TAM) has whitelisted your IPs where you will be uploading the data from.
2. You will receive 3 emails from your TAM:
   1. Email 1: s3\_ferry.zip
   2. Email 2: config.config.zip
   3. Email 3: password to unzip the config.config.zip
3. How to use the s3\_ferry script to upload data to your dedicated DataVisor data bucket on AWS S3:
   1. Make sure you have python 2.7 installed on the machine where you run the script to upload data.
   2. Install additional python packages and upgrade to the latest version:

Install pip using command: sudo easy\_install pip

Then use below command to upgrade or install additional packages

sudo pip install FileChunkIO awscli boto –-upgrade

sudo pip install boto

* 1. Unzip the s3\_ferry.zip in email 1 to your local directory.
  2. Use the password in email 3 to unzip the config.config.zip in email 2, after which you will have a config.config file.
  3. Move the config.config file to the ./src under the unzipped s3\_ferry folder, so that it stays in the same directory with the upload.py script:

$ ls .

DatavisorUploadManager.py

README.client.en (or README.client.chinese)

upload.py

Config.config

CommonUtil.py

* 1. Use the upload.py script to upload your data. Command line syntax:

python ./upload.py -i <input dir or file path>

* 1. The script will return 0 on success, or non-zero on error.
  2. Notes:
     1. The script only uploads new file/directories that have not be uploaded previously.
     2. The script supports upload in parallel.
     3. If the connection is broken, it will restart from where it left off.
     4. Please contact DataVisor if you wish to submit different directory/file format or need additional feature.