==========

Readme for reproducibility submission of paper ID [sidm521]

A) Source code info
Repository: https://github.com/umich-dbgroup/foofah
Programming Language: Python & C++
Additional Programming Language info: Python 2.7
Compiler Info: g++ (for compiling C++ script in ./foofah/foofah_libs/foofah_utils.cpp)
Packages/Libraries Needed:
Boost.Python, python modules: numpy, tabulate, cherrypy, editdistance, python-Levenshteinm matplotlib

B) Datasets info
Repository: ./foofah/tests/data (also available in https://github.com/markjin1990/foofah_benchmarks)

C) Hardware Info
We are pretty flexible in hardware requirement. We have tested our system in both Mac and Linux computers.

D) Experimentation Info
Recommended:

Through Reprozip:

1) Install Reprounzip with your prefered component: reprounzip-docker (used by us), reprounzip-vagrant, reprounzip-vistrails

2) Download foofah_experiment.rpz

3) Setup Foofah reprozip file

```
$ reprounzip docker setup foofah_experiment.rpz foofah
```

4) Run Foofah experiments:

```
$ reprounzip docker run foofah
```

The experiments might take around 14 hours. All 6 figures in the experiment section of our paper can be downloaded using the following command once the experiments are complete.

5) Download the experiment results

Find the docker container id of the container created by reprounzip

```
$ docker ps -al
```

Donwload specific experiment result file:

```
$ docker cp [container-id]:[file-full-path] .
```

[file-full-path] is the full path of generated experimental results, which include 6 figures:

figure_11_a.png
figure_11_b.png
figure_11_c.png
figure_12_a.png
figure_12_b.png
figure_12_c.png

You can find the full paths of these files along with other test result files using the following command

```
$ reprounzip -v showfiles foofah_experiment.rpz
```

Note: Ideally, we should use the download command of reprounzip as follows to download the generated files:

```
$ reprounzip docker download foofah --all
```

But it always gave “Can't get output file” error, which we failed to identify the causes due to the time limit.

Alternative:

D1) Scripts and how-tos to prepare the software for system

1) Install Boost.Python on Mac or on Linux

2) In ./foofah, run:

```
$ python setup.py install
```

This will install all python modules and compile ./foofah/foofah_libs/foofah_utils.cpp

D2) Scripts and how-tos for all experiments executed for the paper

1) In ./foofah/tests, run:

```
$ ./run_experiments.sh
```

The experiments might take around 14 hours. All 6 figures in the experiment section of our paper will be in ./foofah/tests/figures once the experiments are complete. Other experiment results will be in ./foofah/tests/test_result

==========