

```
from google.colab import drive
drive.mount('/content/drive/')
```

Drive already mounted at /content/drive/; to attempt to forcibly remount, call drive.mount(

```
import time
```

```
# starting time
start = time.time()
```

```
import pandas as pd
```

```
import time
```

```
# starting time
startp = time.time()
pandas = pd.read_csv('/content/drive/MyDrive/lending company/accepted_2007_to_2018Q4.csv')
endp = time.time()
# total time taken
print(f"Runtime of the program is {endp - startp}")
```

```
/usr/local/lib/python3.7/dist-packages/IPython/core/interactiveshell.py:2882: DtypeWarni
  exec(code_obj, self.user_global_ns, self.user_ns)
Runtime of the program is 56.44224762916565
```

```
del(pandas)
```

```
pip install dask[dataframe]
```

```
Requirement already satisfied: dask[dataframe] in /usr/local/lib/python3.7/dist-packages (
Requirement already satisfied: toolz>=0.7.3 in /usr/local/lib/python3.7/dist-packages (fr
Requirement already satisfied: fsspec>=0.6.0 in /usr/local/lib/python3.7/dist-packages (f
Requirement already satisfied: pandas>=0.23.0 in /usr/local/lib/python3.7/dist-packages (f
Requirement already satisfied: numpy>=1.13.0 in /usr/local/lib/python3.7/dist-packages (f
Requirement already satisfied: partd>=0.3.10 in /usr/local/lib/python3.7/dist-packages (f
Requirement already satisfied: pytz>=2017.3 in /usr/local/lib/python3.7/dist-packages (f
Requirement already satisfied: python-dateutil>=2.7.3 in /usr/local/lib/python3.7/dist-pa
Requirement already satisfied: locket in /usr/local/lib/python3.7/dist-packages (from pa
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-packages (from
```

```
import dask.dataframe as dd
# starting time
startd = time.time()
```

```
dask = dd.read_csv(' /content/drive/MyDrive/lending company/accepted_200/_to_2018Q4.csv ' )  
endd = time.time()  
# total time taken  
print(f"Runtime of the program is {endd - startd}")
```

Runtime of the program is 0.27263951301574707

```
del(dask)
```

```
pip install modin[ray]
```

```

Requirement already satisfied: click>=7.0 in /usr/local/lib/python3.7/dist-packages (from
Collecting colorful
  Downloading colorful-0.5.4-py2.py3-none-any.whl (201 kB)
  |████████████████████████████████████████| 201 kB 45.1 MB/s
Collecting opencensus
  Downloading opencensus-0.7.13-py2.py3-none-any.whl (127 kB)
  |████████████████████████████████████████| 127 kB 65.4 MB/s
Collecting aioredis<2
  Downloading aioredis-1.3.1-py3-none-any.whl (65 kB)
  |████████████████████████████████████████| 65 kB 3.4 MB/s
Requirement already satisfied: prometheus-client>=0.7.1 in /usr/local/lib/python3.7/dist
Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-packages (from
Collecting aiohttp-cors
  Downloading aiohttp-cors-0.7.0-py3-none-any.whl (27 kB)
Requirement already satisfied: jsonschema in /usr/local/lib/python3.7/dist-packages (from
Collecting gpustat
  Downloading gpustat-0.6.0.tar.gz (78 kB)
  |████████████████████████████████████████| 78 kB 5.0 MB/s
Collecting aiohttp
  Downloading aiohttp-3.7.4.post0-cp37-cp37m-manylinux2014_x86_64.whl (1.3 MB)
  |████████████████████████████████████████| 1.3 MB 49.9 MB/s
Collecting py-spy>=0.2.0
  Downloading py_spy-0.3.8-py2.py3-none-manylinux_2_5_x86_64.manylinux1_x86_64.whl (3.1
  |████████████████████████████████████████| 3.1 MB 39.8 MB/s
Collecting hiredis
  Downloading hiredis-2.0.0-cp37-cp37m-manylinux2010_x86_64.whl (85 kB)
  |████████████████████████████████████████| 85 kB 3.6 MB/s
Collecting async-timeout
  Downloading async_timeout-3.0.1-py3-none-any.whl (8.2 kB)
Requirement already satisfied: typing-extensions>=3.6.5 in /usr/local/lib/python3.7/dist
Requirement already satisfied: chardet<5.0,>=2.0 in /usr/local/lib/python3.7/dist-packa
Collecting multidict<7.0,>=4.5
  Downloading multidict-5.1.0-cp37-cp37m-manylinux2014_x86_64.whl (142 kB)
  |████████████████████████████████████████| 142 kB 57.1 MB/s
Collecting yarl<2.0,>=1.0
  Downloading yarl-1.6.3-cp37-cp37m-manylinux2014_x86_64.whl (294 kB)
  |████████████████████████████████████████| 294 kB 49.5 MB/s
Requirement already satisfied: idna>=2.0 in /usr/local/lib/python3.7/dist-packages (from
Requirement already satisfied: nvidia-ml-py3>=7.352.0 in /usr/local/lib/python3.7/dist-p
Requirement already satisfied: psutil in /usr/local/lib/python3.7/dist-packages (from g
Collecting blessings>=1.6
  Downloading blessings-1.7-py3-none-any.whl (18 kB)
Requirement already satisfied: google-api-core<2.0.0,>=1.0.0 in /usr/local/lib/python3.
Collecting opencensus-context==0.1.2
  Downloading opencensus_context-0.1.2-py2.py3-none-any.whl (4.4 kB)
Requirement already satisfied: googleapis-common-protos<2.0dev,>=1.6.0 in /usr/local/li
Requirement already satisfied: google-auth<2.0dev,>=1.21.1 in /usr/local/lib/python3.7/c
Requirement already satisfied: setuptools>=40.3.0 in /usr/local/lib/python3.7/dist-packa
Requirement already satisfied: pyasn1-modules>=0.2.1 in /usr/local/lib/python3.7/dist-pa
Requirement already satisfied: cachetools<5.0,>=2.0.0 in /usr/local/lib/python3.7/dist-p
Requirement already satisfied: rsa<5,>=3.1.4 in /usr/local/lib/python3.7/dist-packages (
Requirement already satisfied: pyparsing>=2.0.2 in /usr/local/lib/python3.7/dist-package
Requirement already satisfied: pyasn1<0.5.0,>=0.4.6 in /usr/local/lib/python3.7/dist-pac
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packa

```

```

import numpy as np
import modin.pandas as pd
# starting time

```

```

start = time.time()
modin = pd.read_csv('/content/drive/MyDrive/lending_company/accepted_2007_to_2018Q4.csv')
end = time.time()
# total time taken
print(f"Runtime of the program is {end - start}")

```

UserWarning: Ray execution environment not yet initialized. Initializing...
 To remove this warning, run the following python code before doing dataframe operations

```

import ray
ray.init()

```

UserWarning: The size of /dev/shm is too small (6308233216 bytes). The required size at Runtime of the program is 81.41269421577454

pip install ray

```

Requirement already satisfied: ray in /usr/local/lib/python3.7/dist-packages (1.6.0)
Requirement already satisfied: pyyaml in /usr/local/lib/python3.7/dist-packages (from ray)
Requirement already satisfied: redis>=3.5.0 in /usr/local/lib/python3.7/dist-packages (from ray)
Requirement already satisfied: grpcio>=1.28.1 in /usr/local/lib/python3.7/dist-packages (from ray)
Requirement already satisfied: attrs in /usr/local/lib/python3.7/dist-packages (from ray)
Requirement already satisfied: numpy>=1.16 in /usr/local/lib/python3.7/dist-packages (from ray)
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-packages (from ray)
Requirement already satisfied: msgpack<2.0.0,>=1.0.0 in /usr/local/lib/python3.7/dist-packages (from ray)
Requirement already satisfied: protobuf>=3.15.3 in /usr/local/lib/python3.7/dist-packages (from ray)
Requirement already satisfied: click>=7.0 in /usr/local/lib/python3.7/dist-packages (from ray)
Requirement already satisfied: six>=1.5.2 in /usr/local/lib/python3.7/dist-packages (from ray)

```

```

import ray
# Read a directory of files in remote storage.
startr = time.time()
ds = ray.data.read_csv('/content/drive/MyDrive/lending_company/accepted_2007_to_2018Q4.csv')
endr = time.time()
# total time taken
print(f"Runtime of the program is {endr - startr}")

```

(raylet) [2021-09-07 19:20:36,512 E 1086 1086] local_object_manager.cc:269: Failed to se
 Runtime of the program is 429.99610137939453

from tabulate import tabulate

```

table = [["Method", 'Loading Time'], ['Pandas', endp - startp], ['Dask', endd - startd], ['M

```

```
print(tabulate(table))
```

```
-----
Methode  Loading Time
Pandas   56.44224762916565
Dask     0.27263951301574707
Modin    81.41269421577454
Ray      429.99610137939453
-----
```

Dask took the least time to load the data set

✓ 0s completed at 1:20 AM

