Lesson 8 Hands-On
Directions
This Hands-On will be graded, so make sure you complete each part. When you are done, please submit one document with all of your findings for grading.

Description
For this hands on, you'll be exploring the DASK documentation further. Go to this website and read the overview section, then pick at least two sections under USER INTERFACE and SCHEDULING to examine. Please answer the following questions after examining these features:

- What do you think the most useful DASK feature is?
- Why is the advent of DASK so important?
- About what would you like to learn more?

Dask is an open-source library for parallel computing written in Python. As Python grows more popular in data analytics and general programming. Dask also includes many of the python packages that are familiar like Django, pandas, NumPy, matpoltlib, and flask which we have also previously covered. It also integrates using Juypter Notebook which also supports kernels and is open source as well. Dask is an important tool for a data scientist.

The User Interface includes several that we have already worked in like DataFrames and Arrays in Python and there are so much more like Bags, Machine Learning, XArray, Delayed, and Futures.

The Scheduling section has collections of an array, bag, and dataframe into a task graph into schedules of synchronous, threaded, distributed, and multiprocessing.

The most useful Dask feature would be “through its parallel computing features, Dask allows for rapid and efficient scaling of computation (Why and How to Use Dask with Big Data | by Admond Lee | Towards Data Science).” Dask increases the Pandas ability in performance and scalability.

Dask is a cool, easier, and quicker way to work in Python on an interface similar to Jupyter Notebook. I would like to continue to learn more about it.