1.what is the process for loading a dataset from an external source?

Ans.There are several different ways to populate the dataset.for example you can use LINO to SQL to query the data base and load the results into the dataset for more information see LINO to SQL.

2.How can use pandas to read JSON files?

Ans.To read the files we use read ison() function and through it we pass the path to the JSON file we want to read.once we dataset that it returns data frame.

3.Describe the significance of DASK?

Ans.Dask emphasize the following virtues ,familiar,provides parallelized numpy array and pandas data frame object ,flexible, provides a task scheduling interface for more custom workload.

4.Describe the functions of DASK?

Ans.Dask is a free and open source library for parallel computing in python dask help

you scale your data science and machine learning workshops dask makes it easy to works with numpy pandas.

5.Describe Cassandra's features?

Ans.Apache Cassandra Is an open source user available distributed NOSQL DBMS which is designed to handle large amounts of data across Many servers.