Website link:

<https://www.bls.gov/oes/current/oes_research_estimates.htm>

I downloaded the data for all sectors

Steps to copy:

scp at3857@access.cims.nyu.edu:/home/at3857/oes\_research\_2016\_allsectors.csv /home/at3857

hdfs dfs -put oes\_research\_2016\_allsectors.csv /user/at3857/project

hdfs dfs -chmod a+x /user/at3857/project

Inputfile path: /user/at3857/project/oes\_research\_2016\_allsectors.csv

Outputfile path: /user/at3857/project/projectoutput/part-00000

Florida state path:

Inputfile path: /user/at3857/project/oes\_research\_2016\_allsectors.csv

Output : /user/at3857/projectFlorida/output/part-00000

Reducer :

/user/at3857/project/laborstatsreducer.py

Mapper for florida:

/user/at3857/project/maptry2.py

Mapper for New York:

/user/at3857/project/maptry.py

IMPALA TABLEs:

New York state:

create external table dataset1 (industrytitle string, totalemp int)

row format delimited fields terminated by ','

location '/user/at3857/project/projectoutput/';

Florida state:

create external table dataset1\_florida(industrytitle string, totalemp int)

row format delimited fields terminated by ','

location '/user/at3857/ projectFlorida /output/';