Query 8

Find the average Prevailing Wage for each Job for each Year (take part time and full time separate)

Platform used: PIG

Input file:

/user/hive/warehouse/project.db/h1b\_app2

Explanation:

Step1) Segregated the data into Full time and Part time bags using FILTER command.

Step2) Filtered them into multiple bags based on year.

Step3) Grouped each of those bags by job\_tile.

Step4) Found the average prevailing wage for each of those bags.

Strp5) Stored the output in the hdfs.

Pig file:



Output file:

