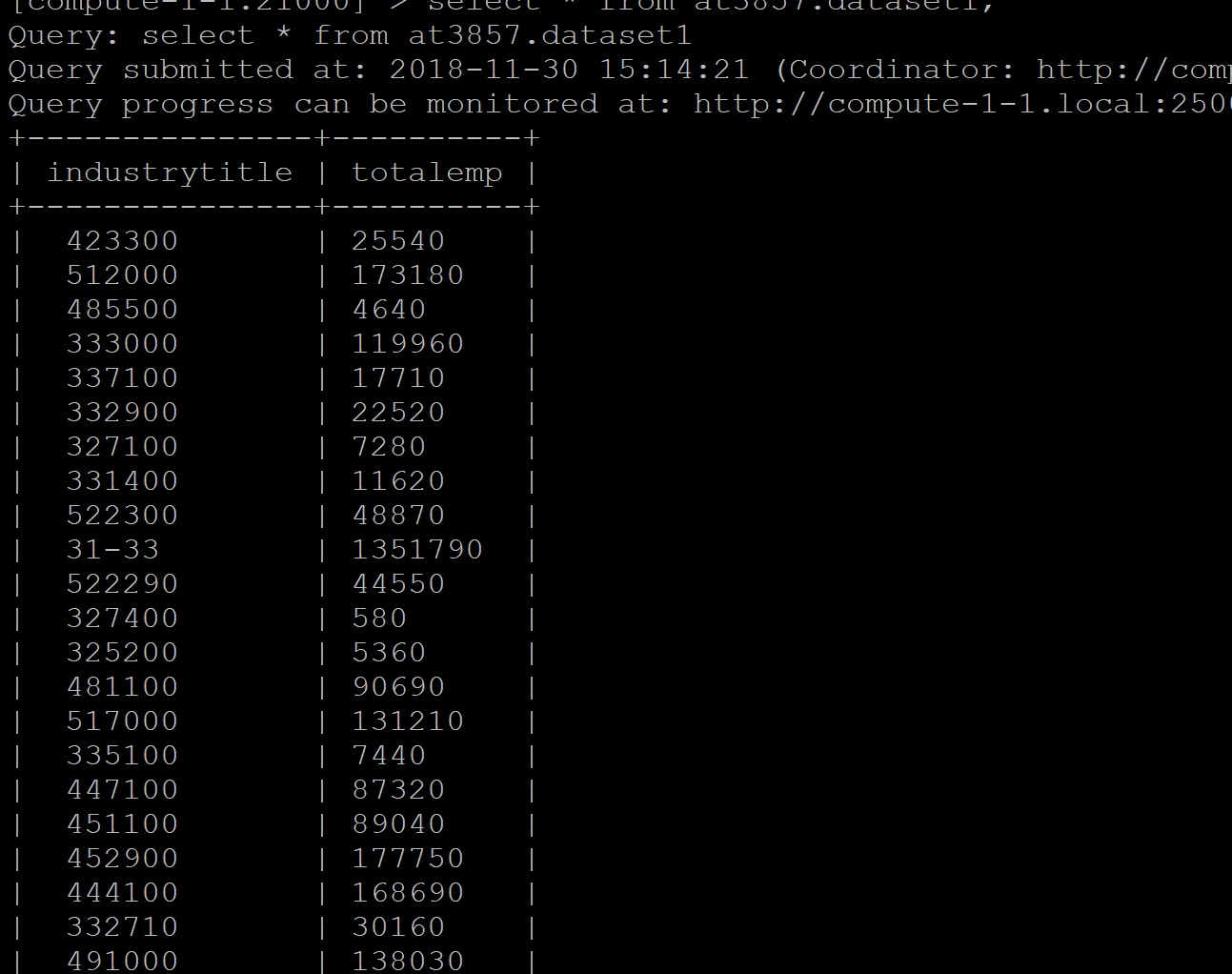
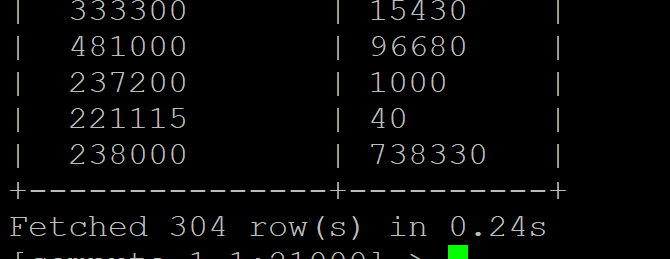
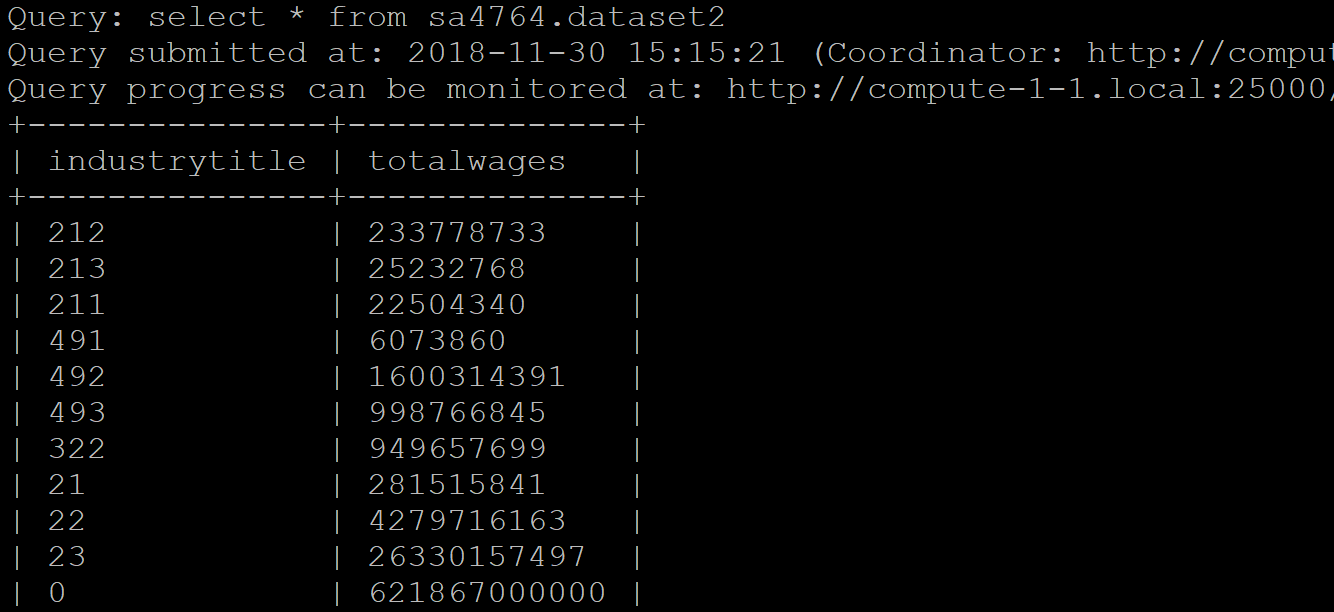
**CODE DROP 2**

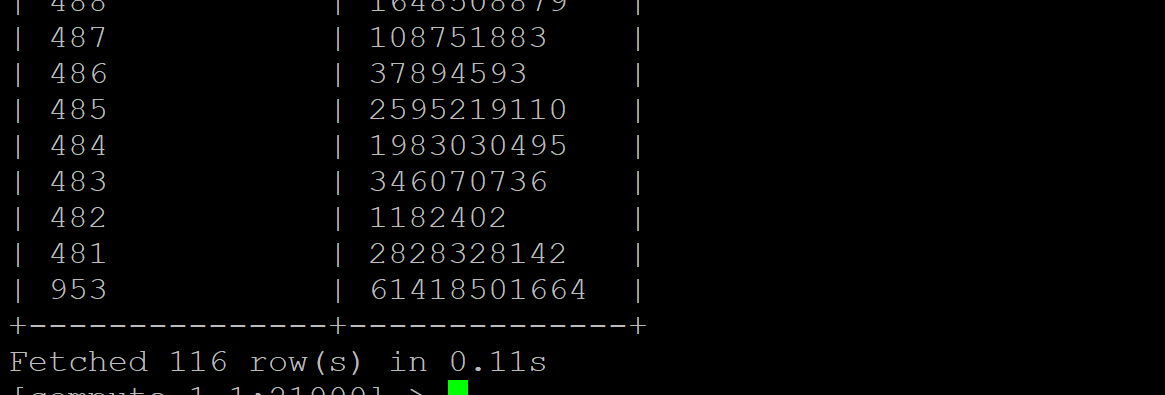
select \* from at3857.dataset1;





select \* from sa4764.dataset2;





create table akanksha\_naics  as (select substr(industrytitle,1,3) as industrycode,totalemp from dataset1 order by industrytitle ASC);

create table sumit\_naics  as (select substr(trim(industrytitle),1,2) as industrytitle ,totalwages from sa4764.dataset2  order by industrytitle ASC);

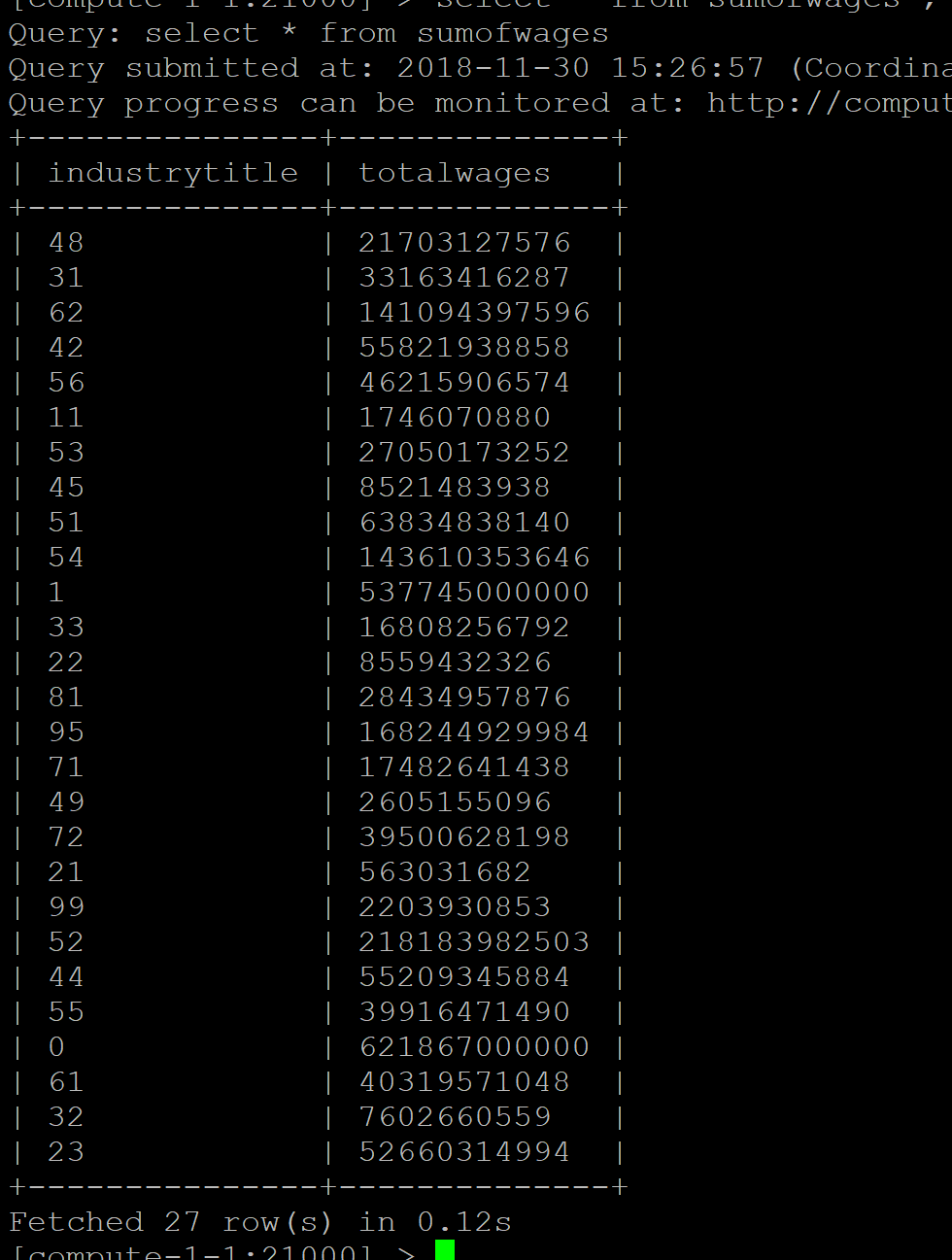
the above tables were created to format naics code as per industry title (for reference <https://www.naics.com/search/> )

create table sumofwages as select trim(industrytitle) as industrytitle ,sum(totalwages) as TotalWages from sumit\_naics group by industrytitle;

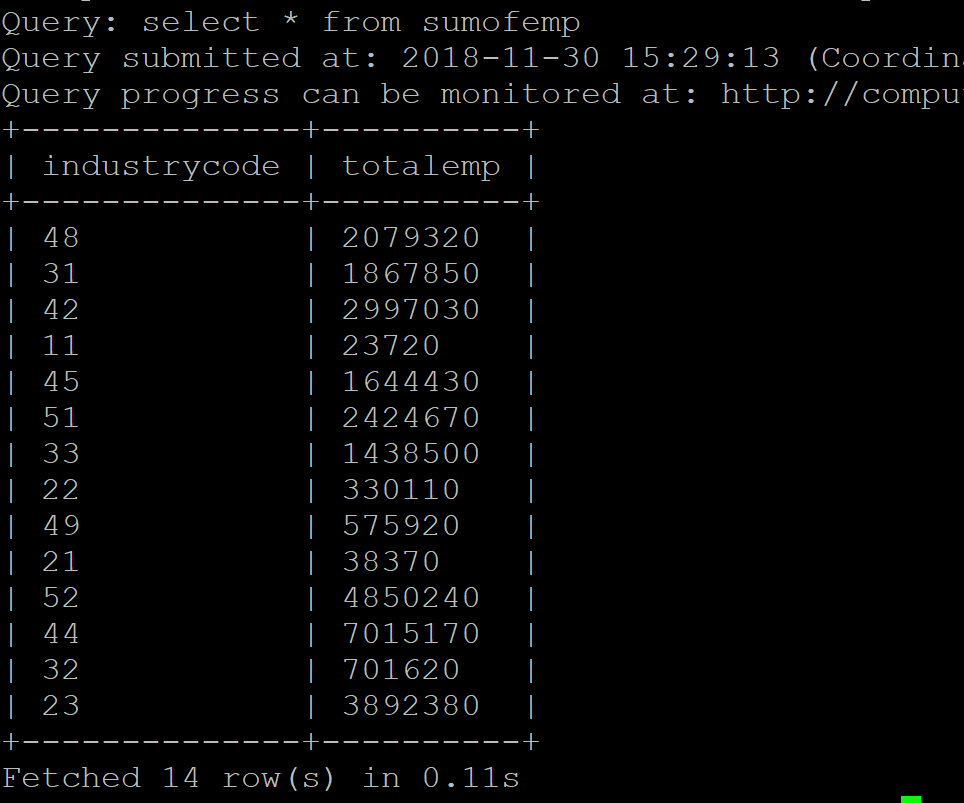
create table sumofemp as Select trim( industrycode) as industrycode ,sum(totalemp) as TotalEmp from akanksha\_naics group by industrycode;

the above tables perform aggregation functions on similar industry codes

Select \* from sumofwages

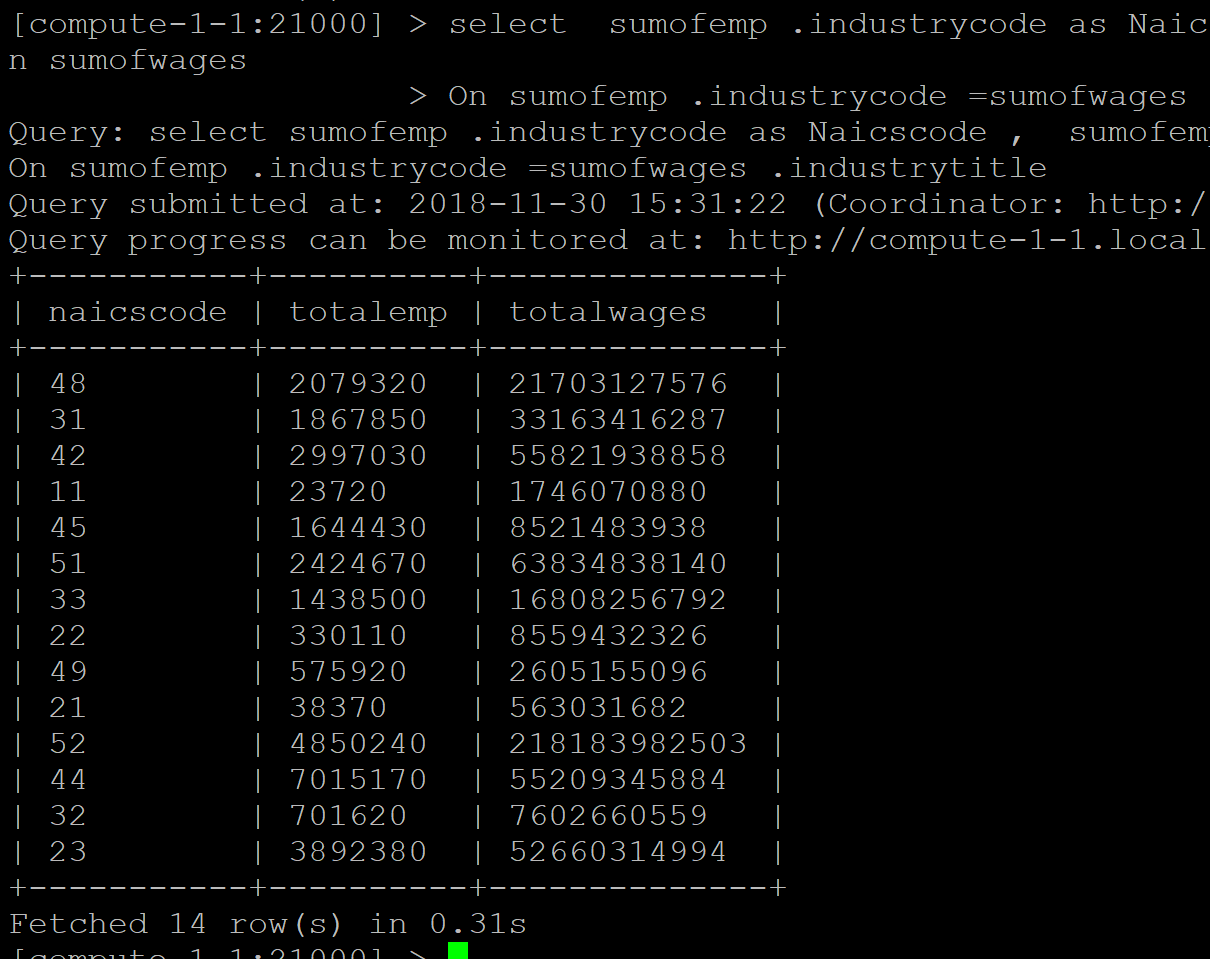


select \* from sumofemp;



select  sumofemp .industrycode as Naicscode ,  sumofemp.totalemp as totalemp , sumofwages.totalwages as totalwages from sumofemp  join sumofwages

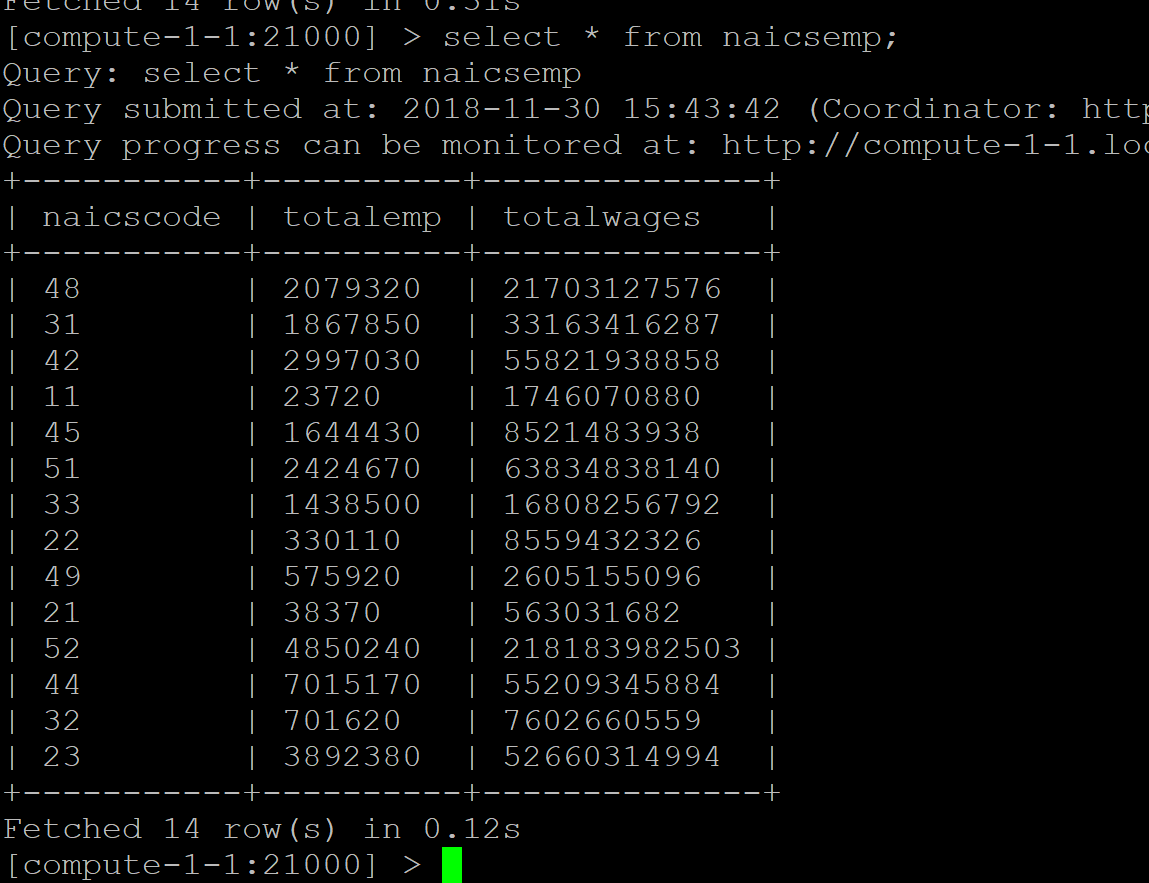
On sumofemp .industrycode =sumofwages .industrytitle;



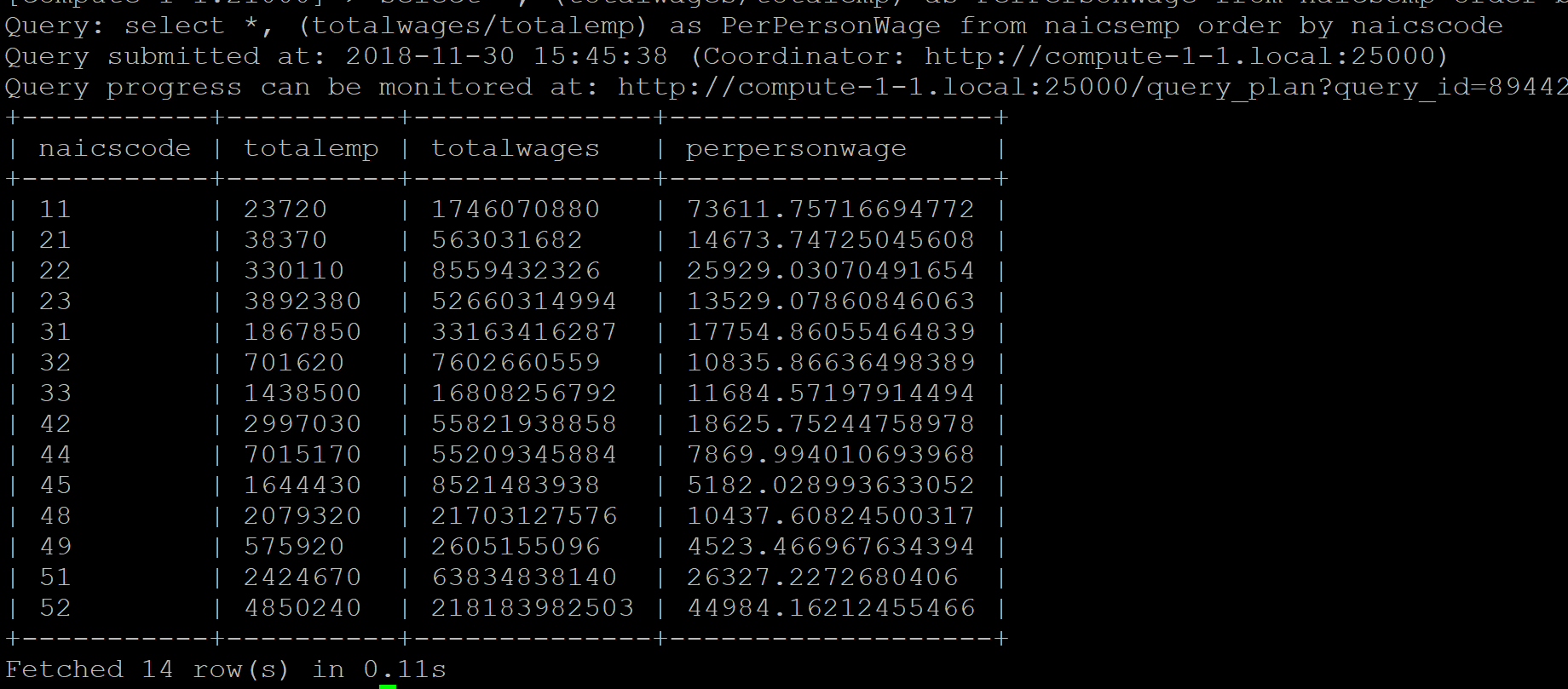
create table naicsemp as select sumofemp .industrycode as Naicscode , sumofemp.totalemp as totalemp , sumofwages .totalwages as totalwages from sumofemp join sumofwages

On sumofemp .industrycode =sumofwages .industrytitle;

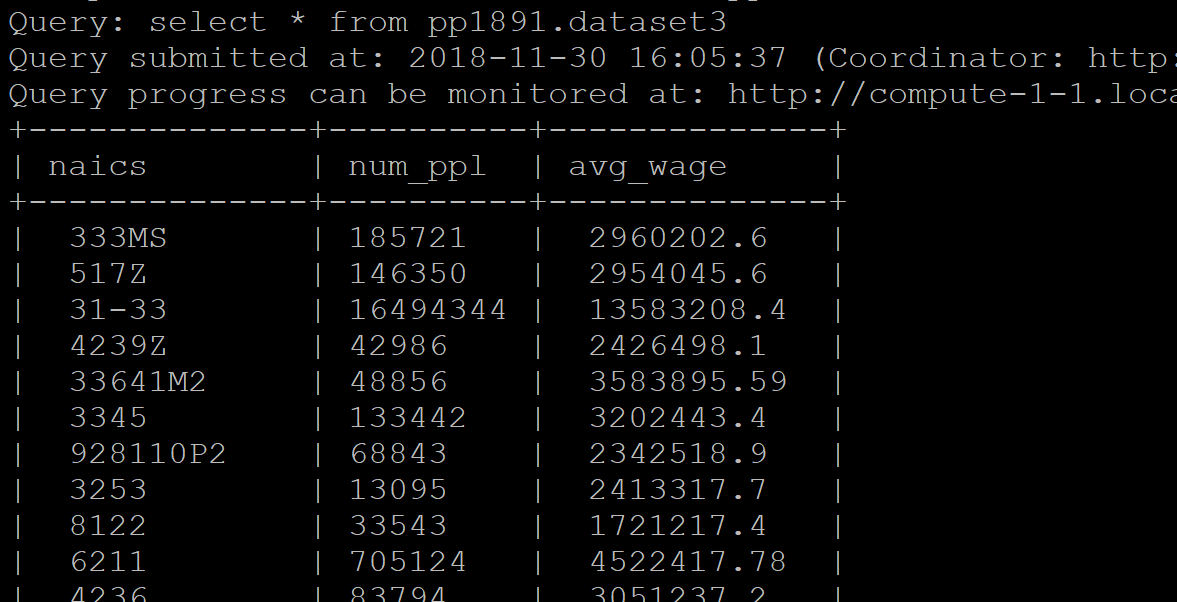
Select \* from naicsemp;

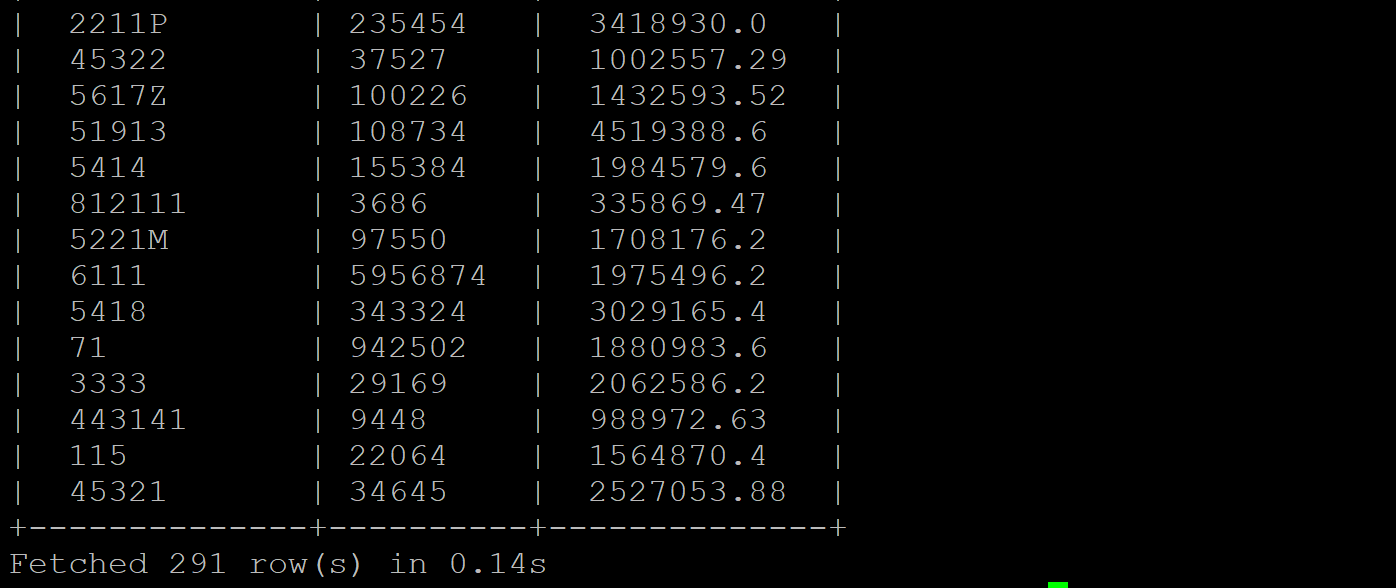


select \*, (totalwages/totalemp) as PerPersonWage from naicsemp order by naicscode;



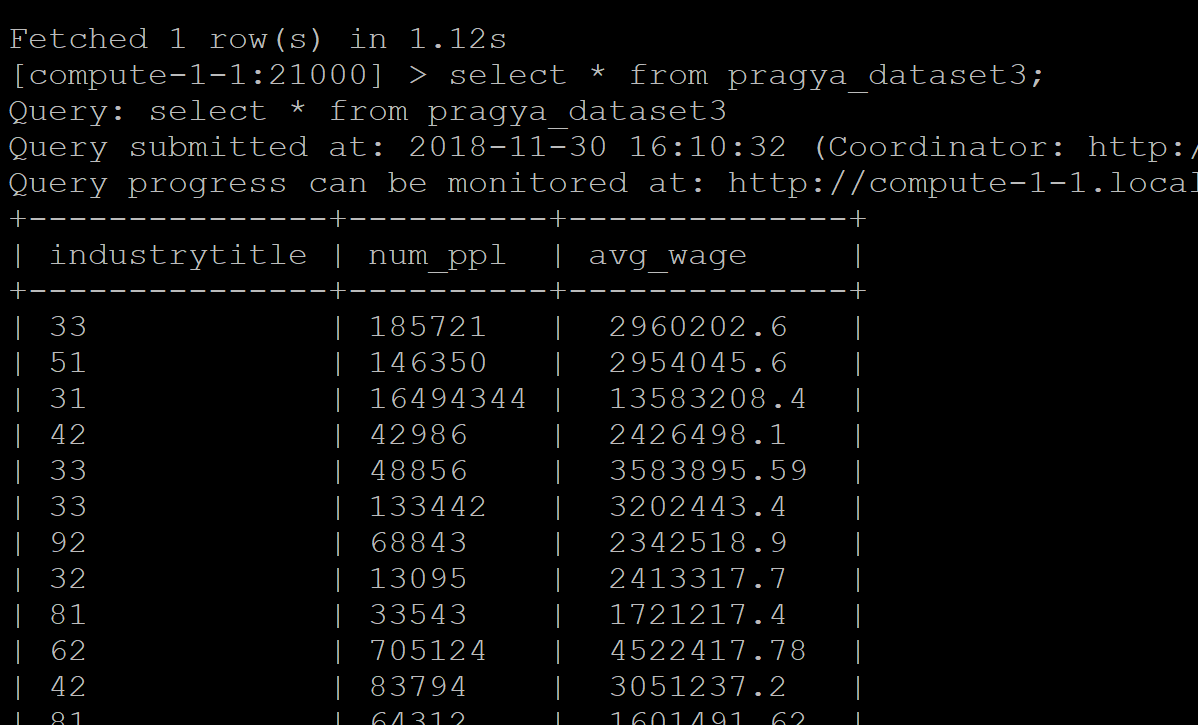
select \* from pp1891.dataset3;

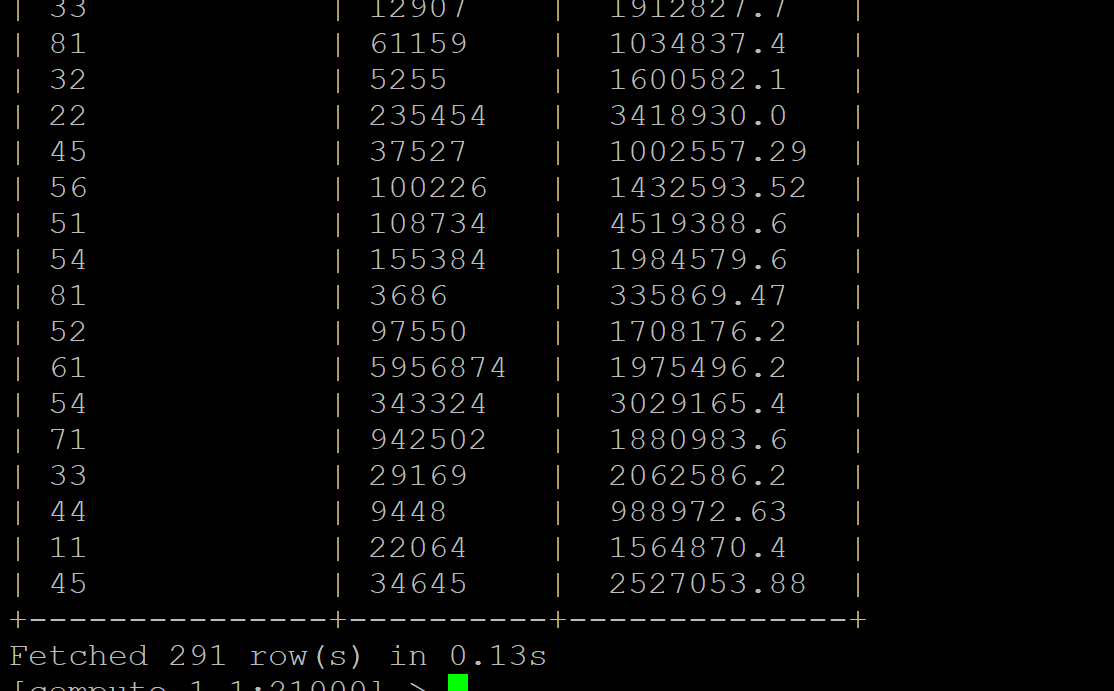




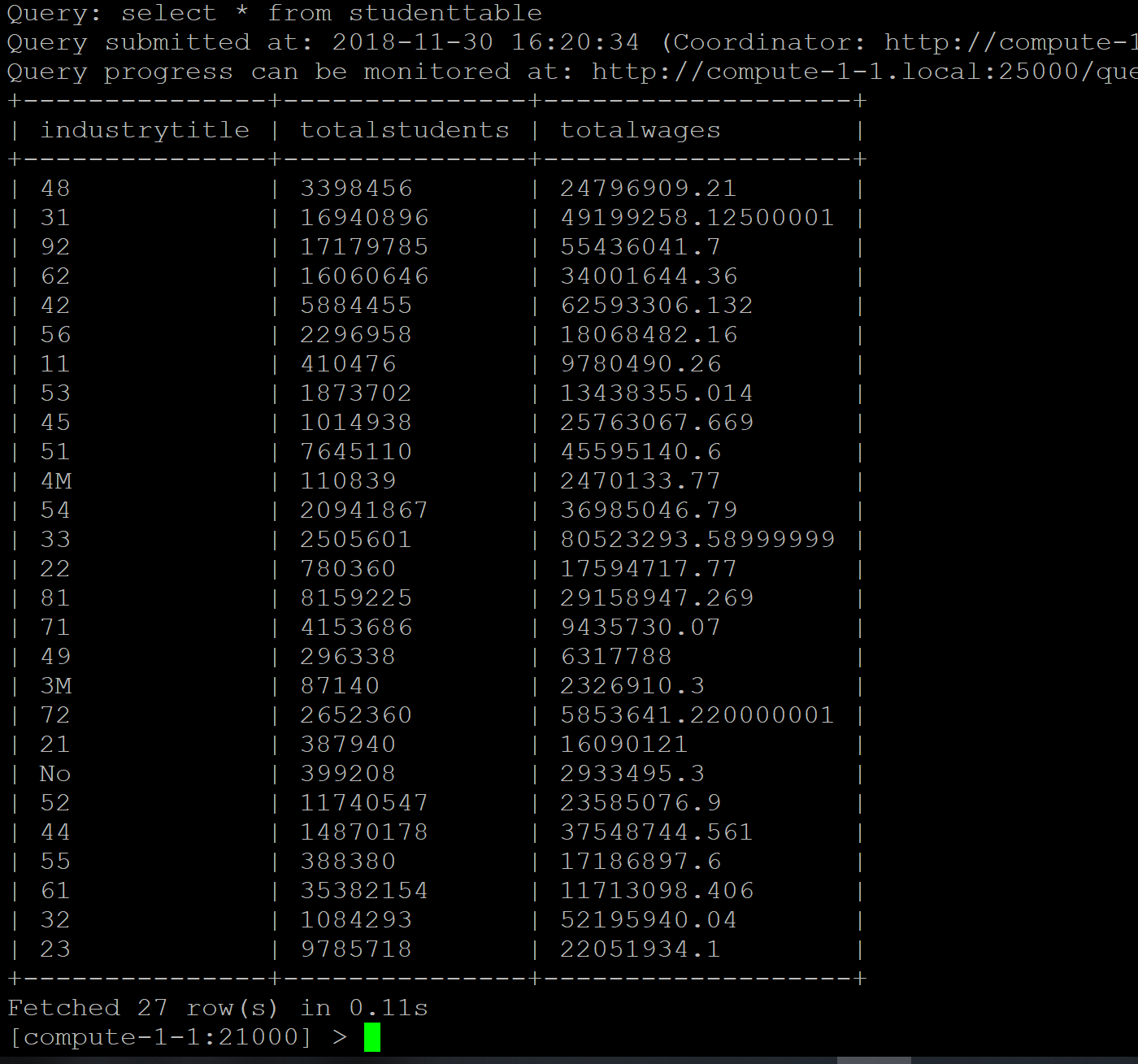
create table pragya\_dataset3  as (select substr(trim(naics),1,2) as industrytitle ,num\_ppl, avg\_wage from pp1891.dataset3  order by naics ASC);

select \* from pragya\_dataset3;

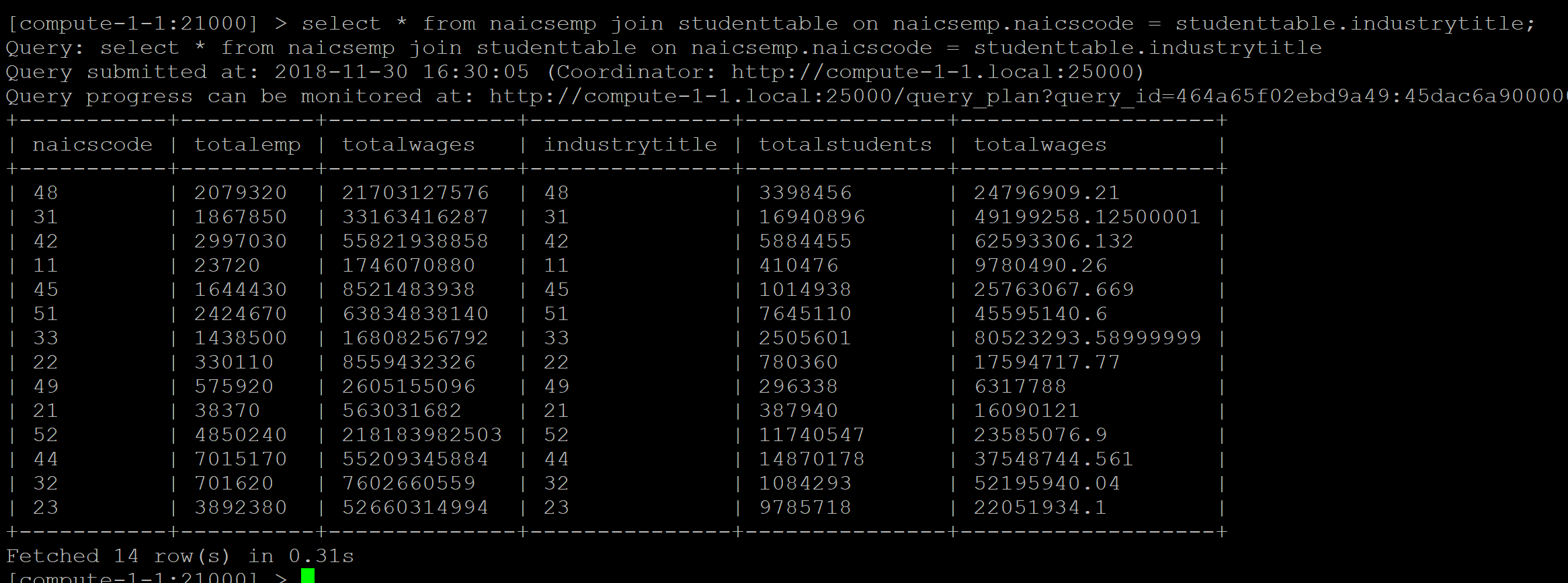




create table studenttable as select trim(industrytitle) as industrytitle ,sum(num\_ppl) as totalstudents ,sum (cast(avg\_wage as double))as TotalWages from pragya\_dataset3 group by industrytitle;

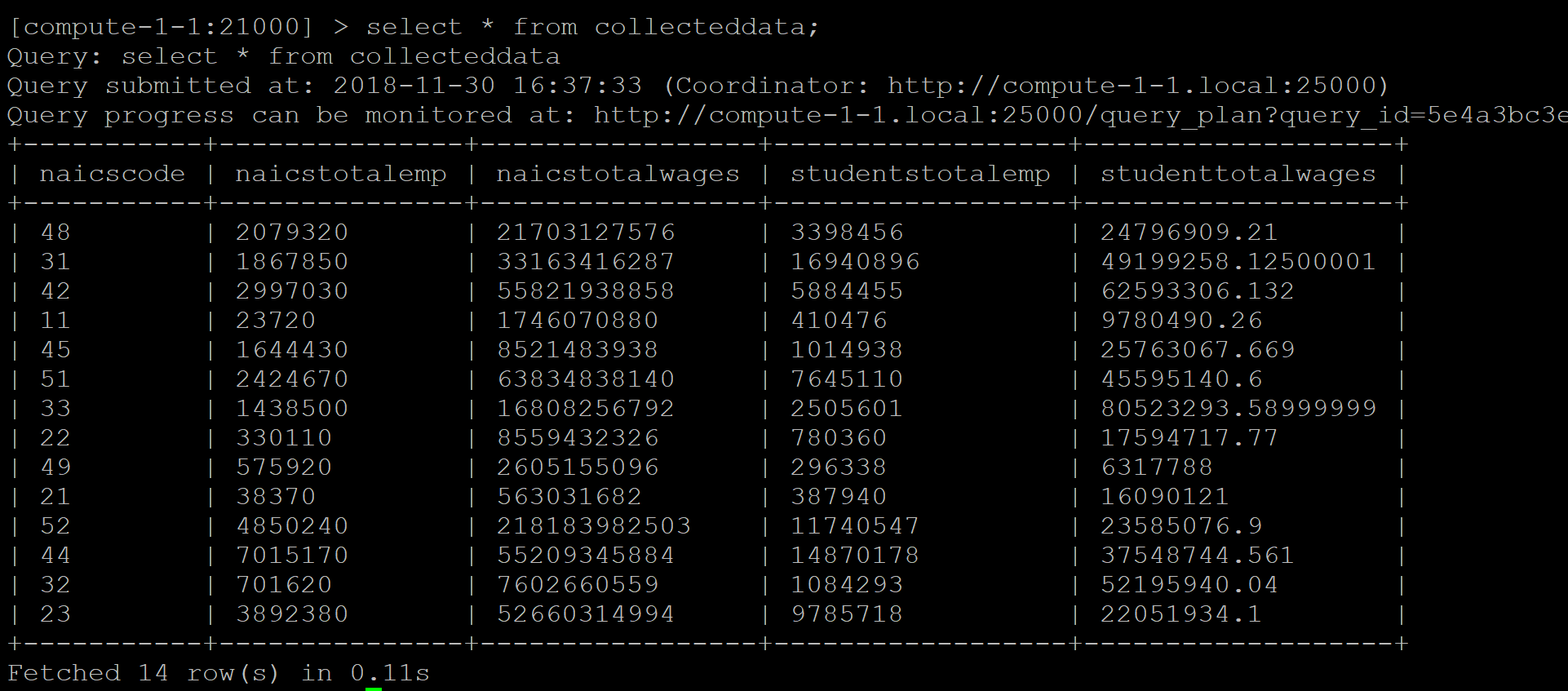


select \* from naicsemp join studenttable on naicsemp.naicscode = studenttable.industrytitle;



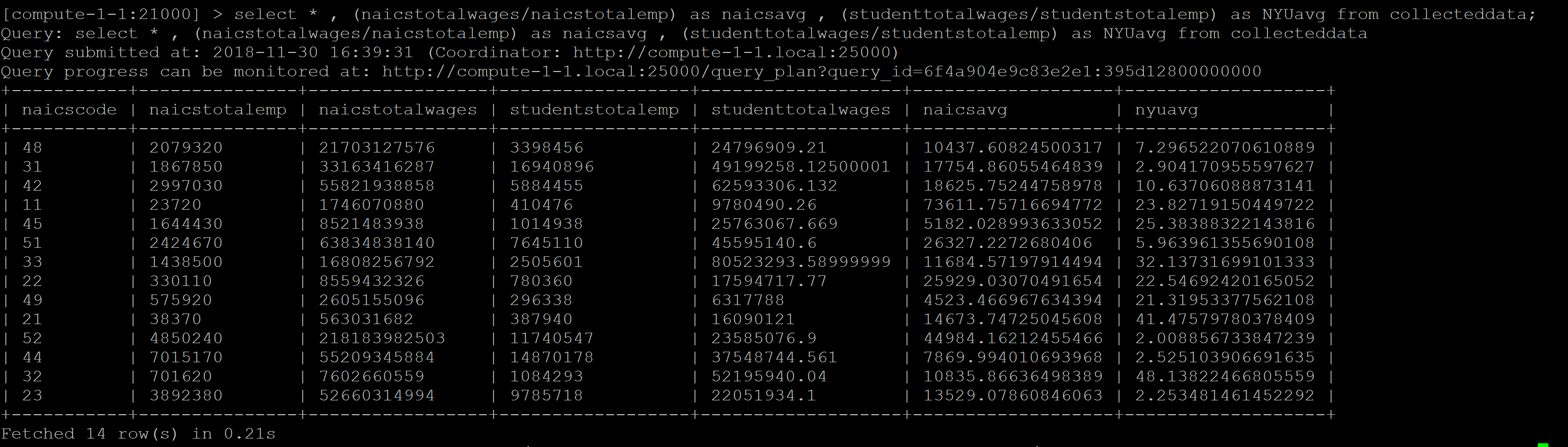
Create table collecteddata as select naicsemp.naicscode as naicscode, naicsemp.totalemp as naicstotalemp, naicsemp.totalwages as naicstotalwages , studenttable.totalstudents as studentstotalemp, studenttable.totalwages as studenttotalwages from naicsemp join studenttable on naicsemp.naicscode = studenttable.industrytitle;

Select \* from collecteddata;



select \* , (naicstotalwages/naicstotalemp) as naicsavg , (studenttotalwages/studentstotalemp) as NYUavg from collecteddata;

Create table bundling as select \* , (naicstotalwages/naicstotalemp) as naicsavg , (studenttotalwages/studentstotalemp) as NYUavg from collecteddata;



create table bundling2 as select dataset1.industrytitle , dataset1.totalemp , sa4764.dataset2.totalwages ,pp1891.dataset3.num\_ppl , pp1891.dataset3.avg\_wage from dataset1 join sa4764.dataset2 join pp1891.dataset3 on trim(dataset1.industrytitle) = trim(sa4764.dataset2.industrytitle) and trim(sa4764.dataset2.industrytitle) = trim(pp1891.dataset3.naics); 