

Stat 506 - Problem 4

Zhen Qin

3.

Datasets and scripts are from <https://github.com/jensqin/Stat506> and <https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/medicare-provider-charge-data/physician-and-other-supplier.html>.

a.

Download the zip file [Medicare Physician and Other Supplier Data CY 2013](#). Then set the correct directory and save data by adding such code.

```
libname lib './';

/* save the file medicare.sas7bdat */
data lib.medicare;
set Medicare_PS_PUF;

run;
```

b.

Use a “data” statement to create a new variable called **totpay** that contains the amount of money paid from Medicare to the provider for the services represented in each line of data.

```
libname lib './';

/* load file */
data medicare;
set lib.medicare;

/* compute totpay */
data lib.mdc_ttp;
set medicare(keep = LINE_SRVC_CNT AVERAGE_MEDICARE_PAYMENT_AMT hcpcs_code line_srvc_cnt
hcpcs_description npi provider_type NPPES_ENTITY_CODE);
totpay = LINE_SRVC_CNT * AVERAGE_MEDICARE_PAYMENT_AMT;

run;
```

c.

i.

Using the data constructed in part (b), obtain the average cost.

```
libname lib './';

proc sql;

/* get the summation of cost */
create table sumcost as
select sum(totpay) as s, sum(LINE_SRVC_CNT) as n, MIN(HCPCS_DESCRIPTION) as description
from lib.mdc_ttp
group by HCPCS_CODE;

/* get the average of cost */
create table avgcost as
select s/n as avg, n, description
from sumcost
having n gt 100000
order by avg desc;

quit;

data lib.avgcost;
set avgcost;

run;
```

ii.

Restrict the data to individual providers.

Part 1

```

libname lib './';

proc sql;

/* compute total money with type */
create table typecost as
select sum(totpay) as s, MIN(PROVIDER_TYPE) as provtype
from lib.mdc_ttp
where NPPES_ENTITY_CODE = "I"
group by NPI;

/* get ordered count */
create table countcost as
select count(*) as n
from typecost
where s>1000000
group by provtype
order by n desc;

quit;

/* get the top 10 rows */
proc sql outobs=10;

create table top10cost as
select *
from countcost;

quit;

proc print data=top10cost;

run;

```

The output file is c2.lst.

Part 2

```
proc sql;

/* get average money for each type */
create table avgtype as
select mean(s) as a
from typecost
group by provtype
order by a desc;

quit;

data lib.avgtype;
set avgtype;

run;
```