STATS 506 Problem Set 4 Question 3

Yunbin Peng

December 3, 2017

Author:
Yunbin Peng
Data Set Used:
Medicare Provider Utilization and Payment Data
https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/Physician-and-Other-Supplier2013.html
Reference Script:
Solution by Prof Shedden
http://dept.stat.lsa.umich.edu/~kshedden/Courses/Stat506/ps4/
Description:
This script use SQL in SAS to rank average cost per claim for different medical treatments, find top 10 most frequent provider types and total amount paid to providers for different provider type.

Part a

Use SAS script provided and get data set MEDICARE_PS_PUF.

```
* Save data into local drive;

libname mywork "L:\SAS\Homework4";
data mywork.medicare;
    set MEDICARE_PS_PUF;

run;
```

Part b

```
* compute total payment;

data medicare1;
    set MEDICARE_PS_PUF;
    totpay = line_srvc_cnt * average_medicare_payment_amt;

run;
```

Part c

part i

```
* table with different service with total cost and total count;
proc sql;

    create table medicare2 as
        select hcpcs_code, hcpcs_description, sum(line_srvc_cnt) as count,
sum(totpay) as s
        from medicare1
        group by hcpcs_code, hcpcs_description;
    quit;


* table with different service with average cost and total count;
proc sql;

    create table medicare3 as
        select hcpcs_description, count, (s/count) as average_cost
        from medicare2;
    quit;

* table of different service order by average cost;
proc sql;

    create table medicare4 as
        select hcpcs_description, count, average_cost
        from medicare3
        order by average_cost desc;
    quit;

proc print data=medicare4(obs = 5);

run;
```

Obs	hcpcs_description	count	average_cost
1	Sipuleucel-t, minimum of 50 million autologous cd54+ cells activated with pap-gm-csf, including leukapheresis and all other preparatory procedures, per infusion	371	24943.48
2	Factor ix (antihemophilic factor, purified, non-recombinant) per i. u.	89	23586.61
3	Balloon dilation and insertion of stent in leg artery	6916	8333.82
4	Implantable neurostimulator pulse generator, dual array, rechargeable, includes extension	98	8179.29
5	Removal of plaque and insertion of stent in leg artery	420	7961.25

Sipuleucel-t has the highest average cost for $24943

```
* table of different services with over 100000 times order by average cost;

proc sql;

    create table medicare5 as
        select hcpcs_description, count, average_cost
        from medicare4
        where count > 100000;
    quit;

proc print data=medicare5(obs=5);

run;
```

Obs	hcpcs_description	count	average_cost
1	Injection, pegfilgrastim, 6 mg	175634	2334.62
2	Heart artery bypass to repair one artery	100849	902.02
3	Repair of knee joint	390824	823.88
4	Replacement of thigh bone and hip joint prosthesis	166796	771.02
5	Injection, aflibercept, 1 mg	1028450	767.10

Pegfilgrastim injection has highest average cost of $2335 among services with over 100000 counts.

part ii

```
* restrict to individual providers;

proc sql;

    create table dai as
        select *
        from medicare1
        where nppes_entity_code = 'I';
    quit;

* table of different providers with total amount paid;

proc sql;

    create table dax as
        select npi, provider_type, sum(totpay) as s
        from dai
        group by npi, provider_type;
    quit;

* filter with over 1 millions charged;

proc sql;
```

```
    create table daz as
        select *
        from dax
        where s > 1000000;
    quit;

* create table of provider types with frequency;

proc sql;

    create table daz as
        select provider_type, count(provider_type) as n
        from daz
        group by provider_type
        order by n desc;
    quit;

proc print data=daz(obs=10);

run;
```

Obs	provider_type	n
1	Ophthalmology	1123
2	Hematology/Oncology	664
3	Radiation Oncology	357
4	Rheumatology	269
5	Dermatology	239
6	Cardiology	220
7	Medical Oncology	211
8	Internal Medicine	128
9	Diagnostic Radiology	90
10	Nephrology	83

Above table gives 10 most frequent provider types.

```
* compute average paid to providers for each provider type in descending
order;

proc sql;

    create table avg_type as
        select provider_type, mean(s) as average
        from dax
        group by provider_type
        order by average desc;
    quit;

proc print data=avg_type(obs=2);



* average paid in acsending order;
proc sql;
    create table avg_type1 as
        select *
        from avg_type
        order by average;
    quit;

proc print data=avg_type1(obs=2);
* Lowest 2: Certified Nurse Midwife, Mass Immunization Roster Biller
```

Obs	provider_type	average
1	Ophthalmology	339100.19
2	Hematology/Oncology	335557.23

Highest 2: Ophthalmology, Hematology/Oncology

Obs	provider_type	average
1	Certified Nurse Midwife	3400.19
2	Mass Immunization Roster Biller	3720.36

Lowest 2: Certified Nurse Midwife, Mass Immunization Roster Biller