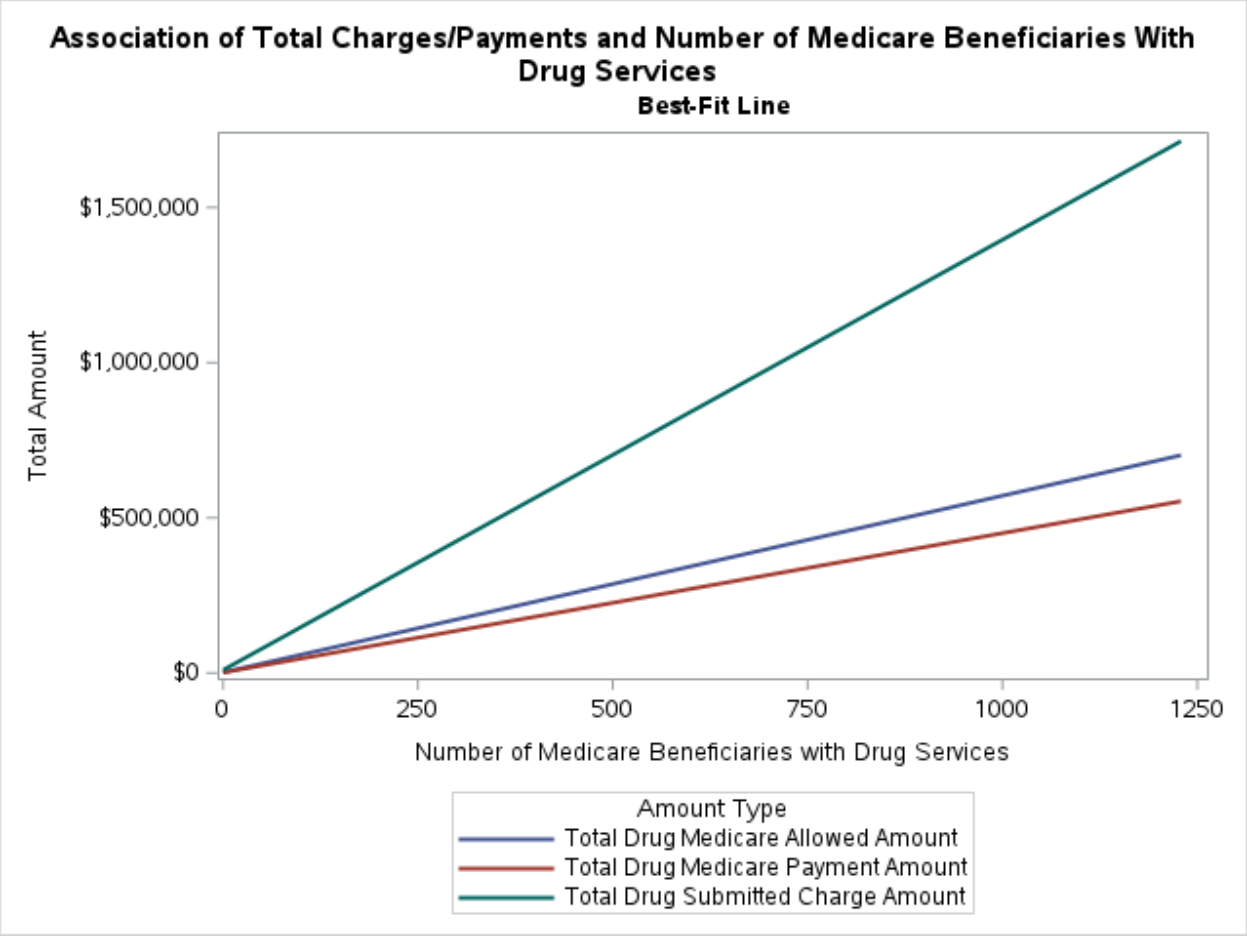
Exercise 1.
Code:

```
data cms_payment;
 set npi.cms_providers_la;
 keep npi total_drug_unique_benes total_drug_medicare_payment_amt;
run;

data cms_allowed;
 set npi.cms_providers_la;
 keep npi total_drug_unique_benes total_drug_medicare_allowed_amt;
run;

data cms_sub;
 set npi.cms_providers_la;
 keep npi total_drug_unique_benes total_drug_submitted_chrg_amt;
run;

data cms_append;
 set cms_payment (in=in_pay rename=(total_drug_medicare_payment_amt = amount))
     cms_allowed (in=in_allow rename=(total_drug_medicare_allowed_amt = amount))
     cms_sub (in=in_sub rename=(total_drug_submitted_chrg_amt = amount));
 if in_pay then amount_type = "Total Drug Medicare Payment Amount";
 else if in_allow then amount_type = "Total Drug Medicare Allowed Amount";
 else if in_sub then amount_type = "Total Drug Submitted Charge Amount";
run;

proc sort data=cms_append;
 by npi;
run;

proc sgplot data=cms_append;
   title1 "Association of Total Charges/Payments and Number of Medicare Beneficiaries With
Drug Services";
   title2 "Best-Fit Line";
   label amount_type = "Amount Type";
   reg y=amount x= total_drug_unique_benes / group=amount_type nomarkers;
   xaxis label="Number of Medicare Beneficiaries with Drug Services";
   yaxis label="Total Amount";
   format amount dollar15.;
run;
```

Output:



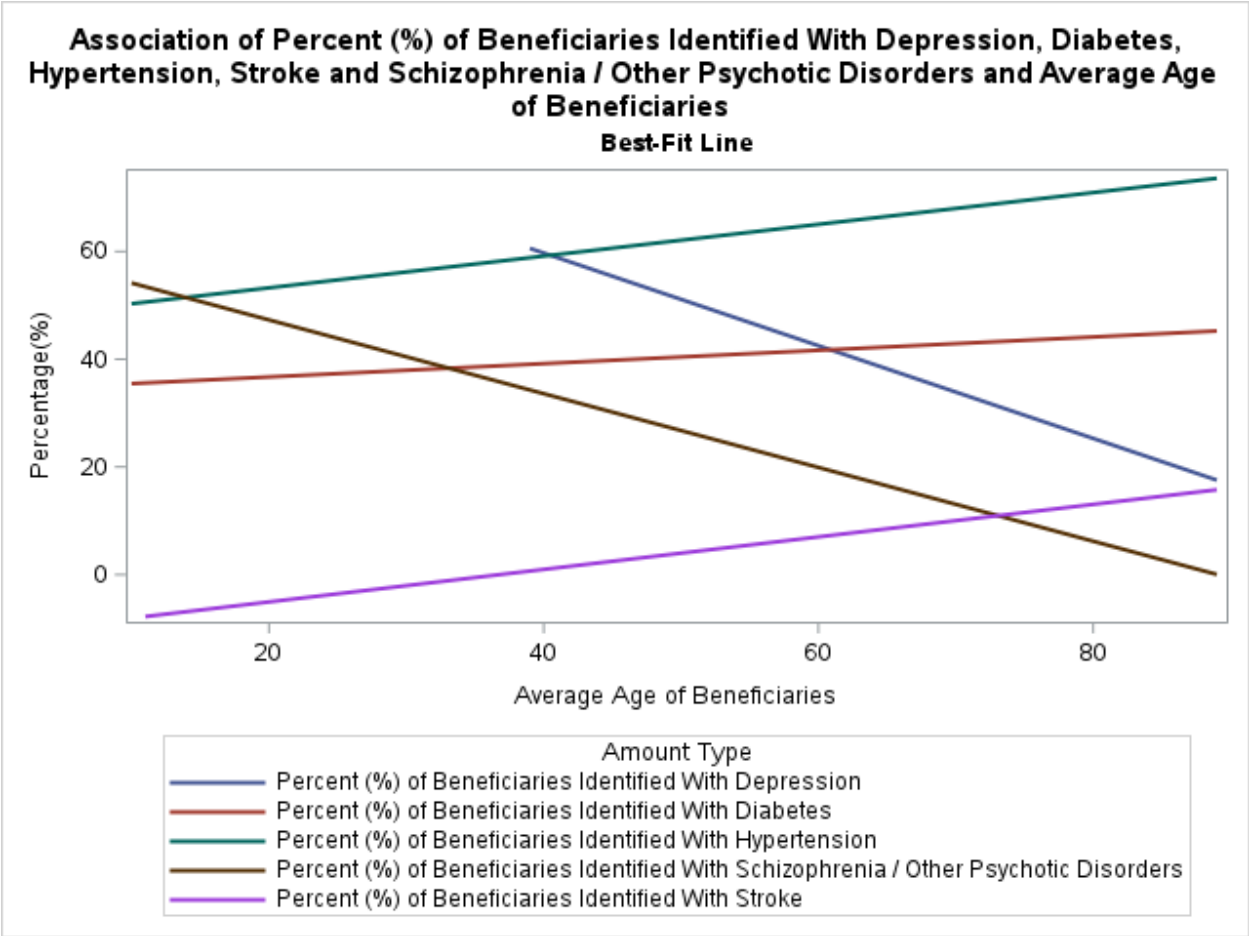
Exercise 2.
Code:

```
data cms_dep;
 set npi.cms_providers_la;
 keep npi beneficiary_average_age beneficiary_cc_depr_percent;
run;

data cms_dia;
 set npi.cms_providers_la;
 keep beneficiary_average_age beneficiary_cc_diab_percent;
run;

data cms_hyp;
 set npi.cms_providers_la;
 keep beneficiary_average_age beneficiary_cc_hypert_percent;
run;

data cms_str;
```

```
 set npi.cms_providers_la;
 keep beneficiary_average_age beneficiary_cc_strk_percent;
run;

data cms_sch;
 set npi.cms_providers_la;
 keep beneficiary_average_age beneficiary_cc_schiot_percent;
run;

data cms_append;
 set cms_dep (in=in_dep rename=(beneficiary_cc_depr_percent = amount))
     cms_dia (in=in_dia rename=(beneficiary_cc_diab_percent = amount))
     cms_hyp (in=in_hyp rename=(beneficiary_cc_hypert_percent = amount))
     cms_str (in=in_str rename=(beneficiary_cc_strk_percent = amount))
     cms_sch (in=in_sch rename=(beneficiary_cc_schiot_percent = amount));
 if in_dep then amount_type = "Percent (%) of Beneficiaries Identified With Depression";
 else if in_dia then amount_type = "Percent (%) of Beneficiaries Identified With Diabetes";
 else if in_hyp then amount_type = "Percent (%) of Beneficiaries Identified With Hypertension";
 else if in_str then amount_type = "Percent (%) of Beneficiaries Identified With Stroke";
 else if in_sch then amount_type = "Percent (%) of Beneficiaries Identified With Schizophrenia /
Other Psychotic Disorders";
run;

proc sort data=cms_append;
 by npi;
run;

proc transpose
 data=npi.cms_providers_la
 out=cms_long (rename=(Col1=amount _LABEL_ = amount_type))
 name=at;
 by npi beneficiary_average_age;
 var beneficiary_cc_depr_percent beneficiary_cc_diab_percent beneficiary_cc_hypert_percent
beneficiary_cc_strk_percent beneficiary_cc_schiot_percent;
run;

proc sgplot data=cms_long;
   title1 "Association of Percent (%) of Beneficiaries Identified With Depression, Diabetes,
Hypertension, Stroke and Schizophrenia / Other Psychotic Disorders and Average Age of
Beneficiaries ";
   title2 "Best-Fit Line";
   label amount_type = "Amount Type";
   reg y=amount x=beneficiary_average_age / group=amount_type nomarkers;
   xaxis label="Average Age of Beneficiaries";
```

```
   yaxis label="Percentage(%)";

run;
```

Output:



Exercise 3.

Code:

```
data family;
 set npi.cms_providers_la;
 if provider_type = "Family Practice";
run;

data totser (keep=totser);
 set family end=last;
```

```
 totser + total_services;
 if last then output;
run;

data familyna;
set family(keep=npi nppes_provider_last_org_name nppes_provider_first_name
total_services);
 if _n_ = 1 then set totser;
 pct_services = total_services/totser;
 format pct_services percent10.3;
run;

proc sort data=familyna;
 by descending pct_services;
run;

title "Family Practitioners with the Highest Percent Total Number of Services";
proc print data=familyna ;
 var npi nppes_provider_last_org_name nppes_provider_first_name pct_services;
run;
```

In this sample of Family Practice providers, the percentages representing the number of services provided by individual providers relative to the total number of services across all providers range from 0.005% to 17.251%.
The wide range demonstrates the significant variation in the contribution of services among family practitioners with some providers contributing a very small share of services. The biggest contributor is Dr. Leonardo Ortiz at 17.25% and the next biggest contributor is at 5.55% with Dr. Ralph Napolitano which highlights the substantial difference between the biggest contributors.

Exercise 4.

```
data family;
set npi.cms_providers_la;
if provider_type = "Family Practice";
run;

data psych;
set npi.cms_providers_la;
if provider_type = "Psychiatry";
run;

data emed;
set npi.cms_providers_la;
```

```
if provider_type = "Emergency Medicine";
run;

data totfam (keep=totben);
set family end=last;
totben + total_unique_benes;
if last then output;
run;

data totps (keep=totben);
set psych end=last;
totben + total_unique_benes;
if last then output;
run;

data totem (keep=totben);
set emed end=last;
totben + total_unique_benes;
if last then output;
run;

data familyto;
set family(keep=npi nppes_provider_last_org_name
nppes_provider_first_name provider_type total_unique_benes);
if _n_ = 1 then set totfam;
num_benes_relative_tot = total_unique_benes/totben;
format num_benes_relative_tot 10.3;
run;

data psyto;
set psych(keep=npi nppes_provider_last_org_name
nppes_provider_first_name provider_type total_unique_benes);
if _n_ = 1 then set totps;
num_benes_relative_tot = total_unique_benes/totben;
format num_benes_relative_tot 10.3;
run;

data emto;
set emed(keep=npi nppes_provider_last_org_name
nppes_provider_first_name provider_type total_unique_benes);
```

```
if _n_ = 1 then set totem;
num_benes_relative_tot = total_unique_benes/totben;
format num_benes_relative_tot 10.3;
run;


data mergedprov;
set familyto psyto emto;
run;

proc means data=mergedprov Median;
class provider_type;
run;
```

For "Emergency Medicine," the median value for "num_benes_relative_tot" is approximately 0.0028399.
For "Family Practice," the median value for "num_benes_relative_tot" is approximately 0.0032028.
For "Psychiatry," the median value for "num_benes_relative_tot" is approximately 0.0023218.

Exercise 5.

```
proc contents data=npi.cms_providers_la;
run;

data cms_deactivated;
 length NPI $10;
 informat NPPES_Deactivation_Date mmddyy10.;
 infile "~/my_files/data/NPPES_Deactivated_NPI_Report_20171010.csv" dsd;
 input  NPI $
        NPPES_Deactivation_Date;
 format NPPES_Deactivation_Date mmddyy10.;
run;

title "Contents of the Deactivation NPI Report Data Set";
proc contents data=cms_deactivated;
run;

proc sort data=npi.cms_providers_la;
   by NPI;
run;
```

```
proc sort data=cms_deactivated;
   by NPI;
run;

data combine;
  merge npi.cms_providers_la(in=a) cms_deactivated(in=b);
  by NPI;
  if a and b;
run;
```

There are 24 records of deactivation, the earliest deactivation being on 05/14/2015 and the latest being 10/09/2017

Exercise 6

```
proc sort data=npi.cms_providers_la;
   by NPI;
run;

proc sort data=cms_deactivated;
   by NPI;
run;

data updated_cms_providers;
   merge npi.cms_providers_la(in=a) cms_deactivated(in=b);
   by NPI;
   if a;
run;
```