

```
/* SAS Graph Examples */
```

```
proc import datafile =  
    "p:\qac\qac201\students\section3\rkabacoff\salaries.csv"  
    out=salaries dbms=csv replace;  
run;  
proc contents data=salaries;  
run;
```

#Variable	Type	Len	Format	Informat
1caseno	Char	4	\$4.	\$4.
3discipline	Char	3	\$3.	\$3.
2rank	Char	11	\$11.	\$11.
7salary	Num	8	BEST12.	BEST32.
6sex	Char	8	\$8.	\$8.
5yrsservice	Num	8	BEST12.	BEST32.
4yrssincephd	Num	8	BEST12.	BEST32.

```
/* Saving results */
```

```
ods _all_ close;  
ods html  
    path='C:\Users\rkabacoff\OD\qac201 applied data analysis\'(url=none)  
    file='sastest.html' gtitle;
```

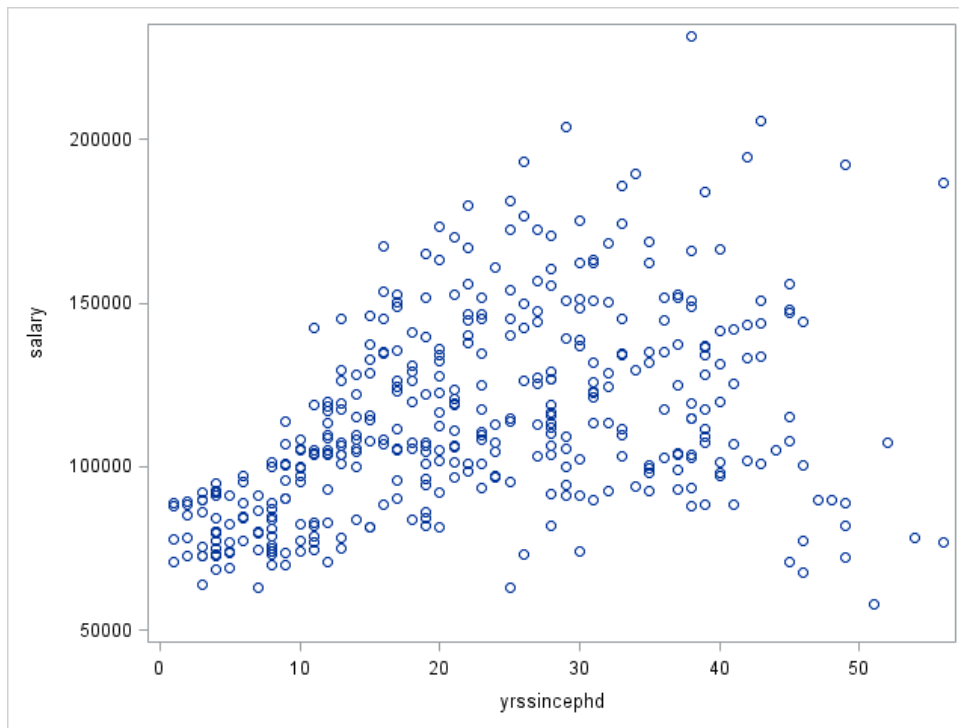
```
/* graphics code */
```

```
ods _all_ close;  
ods listing;
```

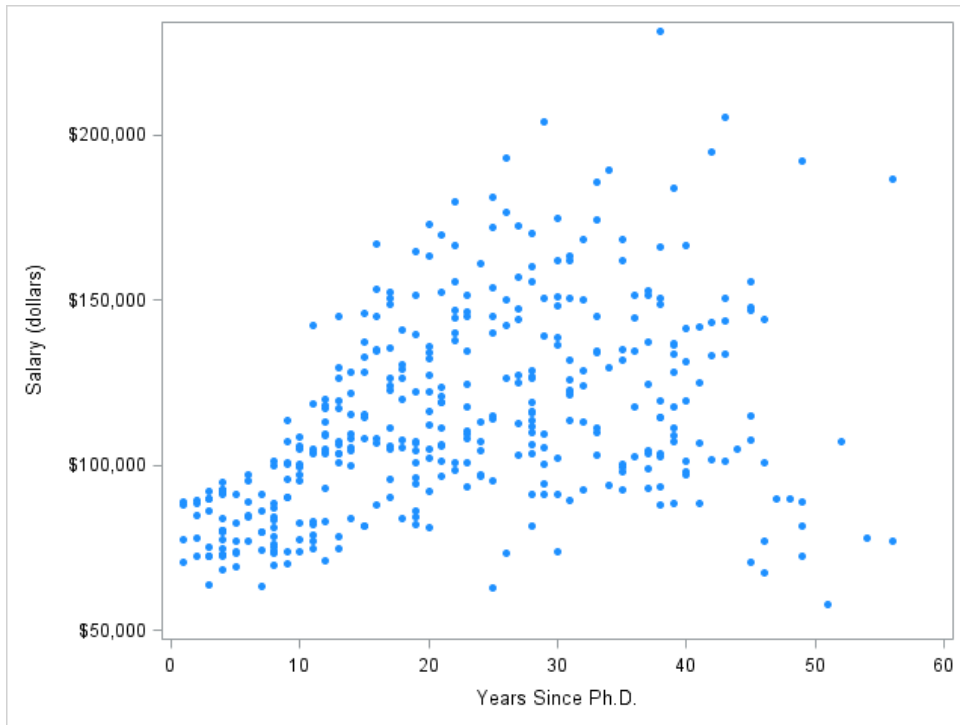
saves png files to location.

```
/* scatter plots */
```

```
proc sgplot data=salaries;  
  scatter x=yrssincephd y=salary;  
run;
```



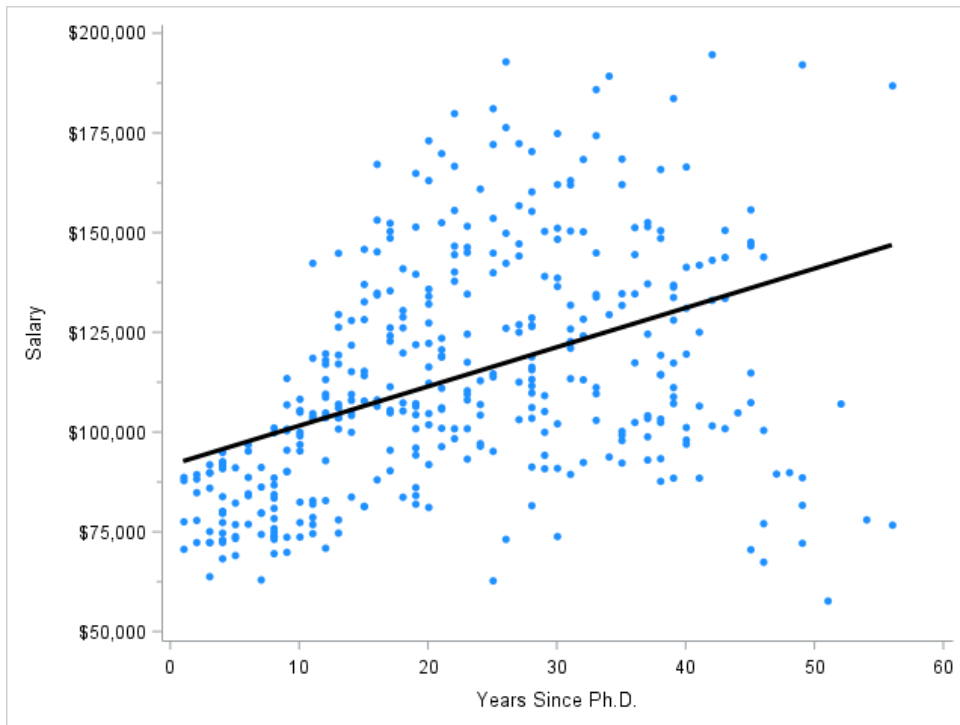
```
proc sgplot data=salaries;  
  scatter x=yrssincephd y=salary /  
    markerattrs=(symbol=circlefilled size=6 color=dodgerblue);  
  xaxis label = "Years Since Ph.D." values=(0 to 60 by 10);  
  yaxis label = "Salary (dollars)" ;  
  format salary dollar.;  
  title 'Nine Month Salary by Experience';  
run;
```



```

proc sgplot data=salaries noborder noautolegend;
  reg x=yrssincephd y=salary /
    lineattrs=(color=black thickness=3)
    markerattrs=(symbol=circlefilled size=6 color=dodgerblue);
  xaxis label = "Years Since Ph.D." values=(0 to 60 by 10);
  yaxis label = "Salary" values=(50000 to 200000 by 25000) minor;
  format salary dollar.;
  title 'Nine Month Salary by Experience';
  footnote 'With line of best fit';
run;

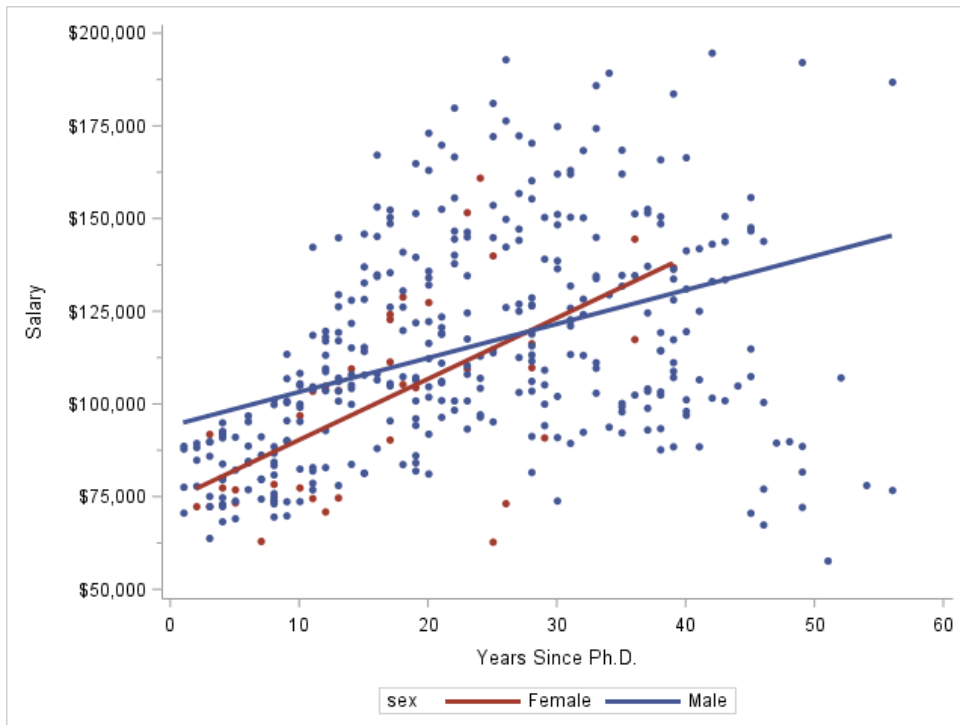
```



```

proc sgplot data=salaries noborder;
  reg x=yrssincephd y=salary / group=sex
      lineattrs=(thickness=3)
      markerattrs=(symbol=circlefilled size=6);
  xaxis label = "Years Since Ph.D." values=(0 to 60 by 10);
  yaxis label = "Salary" values=(50000 to 200000 by 25000) minor;
  format salary dollar.;
  title 'Nine Month Salary by Experience';
  footnote 'With line of best fit';
run;

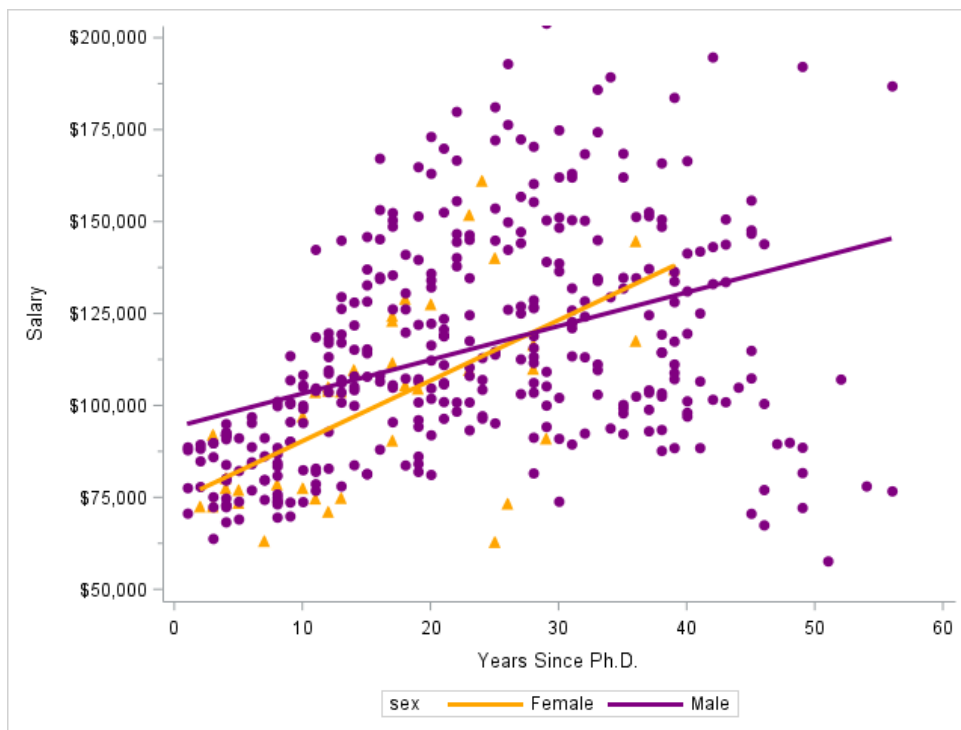
```



```

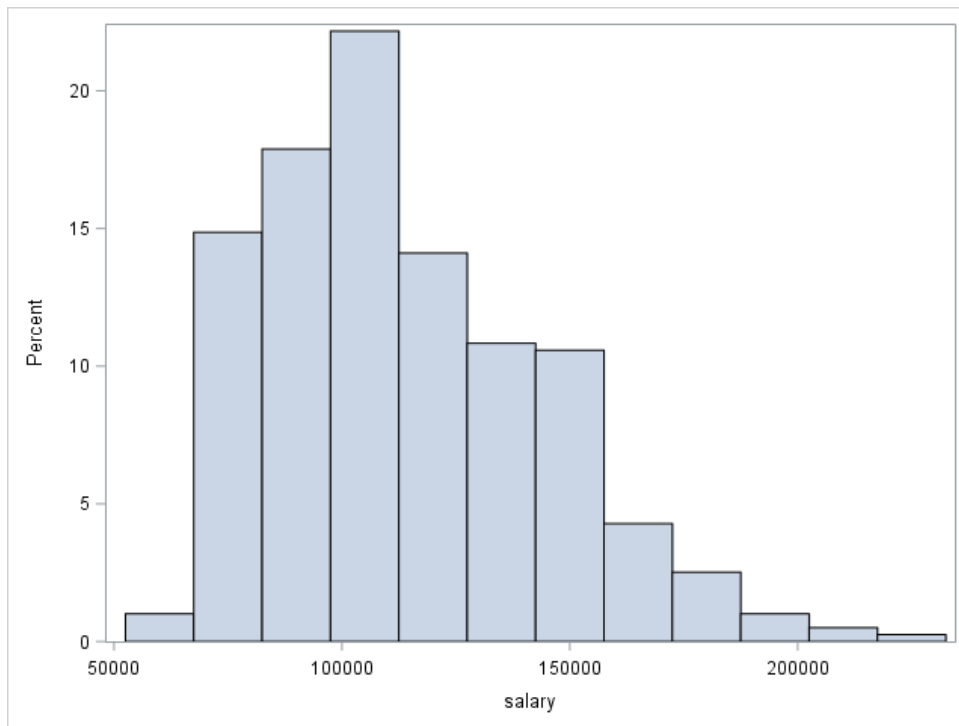
ods graphics / attrpriority=none;
proc sgplot data=salaries noborder;
styleattrs datacontrast=(purple orange)
datasymbols=(circlefilled trianglefilled)
datalinepatterns=(solid);
reg x=yrssincephd y=salary / group=sex
lineattrs=(thickness=3)
markerattrs=(size=8);
axis label = "Years Since Ph.D." values=(0 to 60 by 10);
yaxis label = "Salary" values=(50000 to 200000 by 25000) minor;
format salary dollar.;
title 'Nine Month Salary by Experience';
footnote 'With separate best fit lines';
run;

```

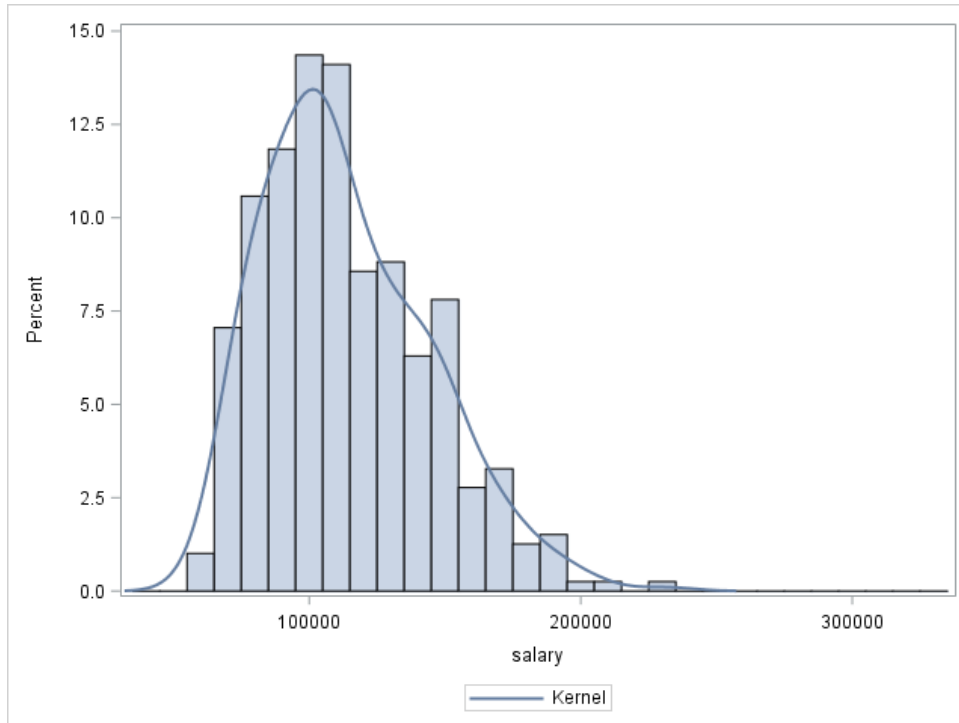


```
/* histograms */
```

```
proc sgplot data=salaries;  
  histogram salary;  
run;
```

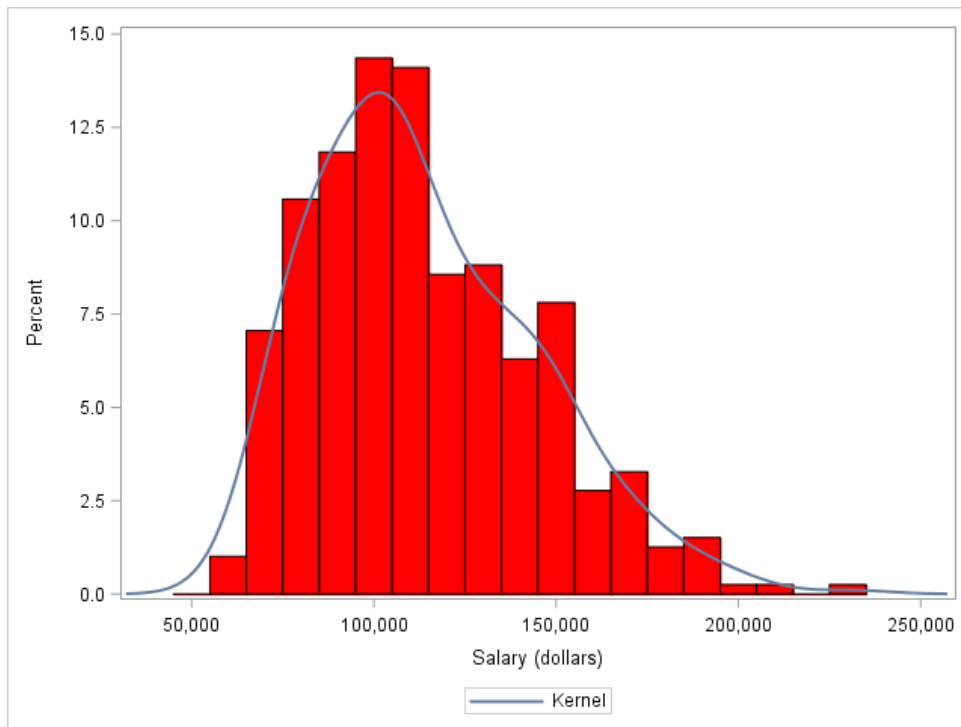


```
proc sgplot data=salaries;  
  histogram salary/ nbins=30;  
  density salary / type = kernel;  
run;
```





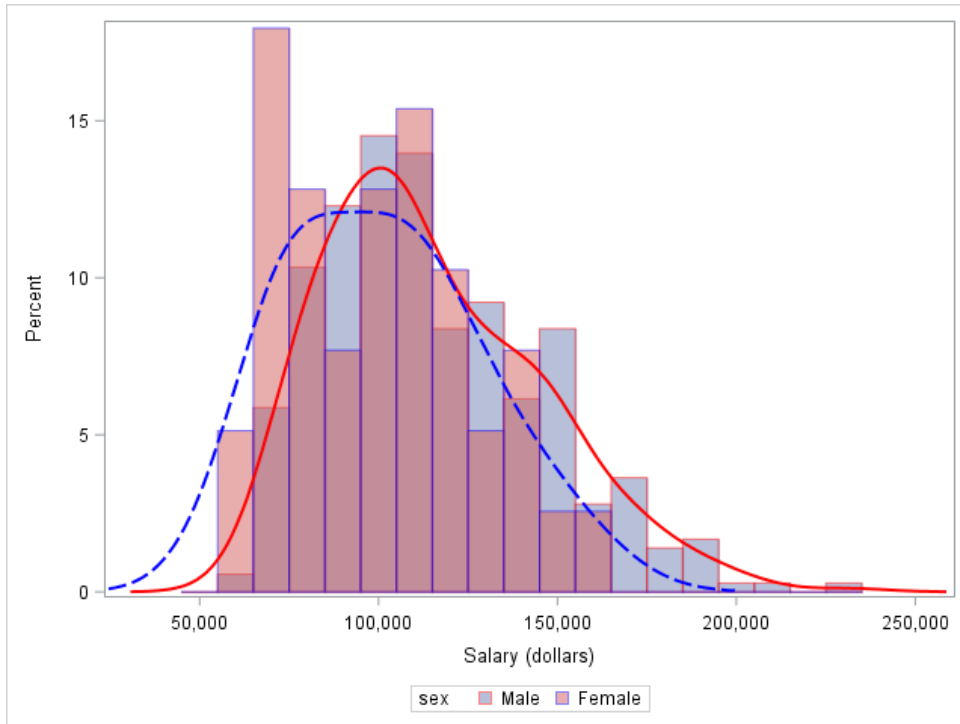
```
proc sgplot data=salaries;  
  histogram salary/ binwidth=10000 fillattrs=(color=red);  
  density salary / type = kernel ;  
  xaxis label="Salary (dollars)";  
  format salary comma.;  
  title 'Nine month salary distribution';  
run;
```



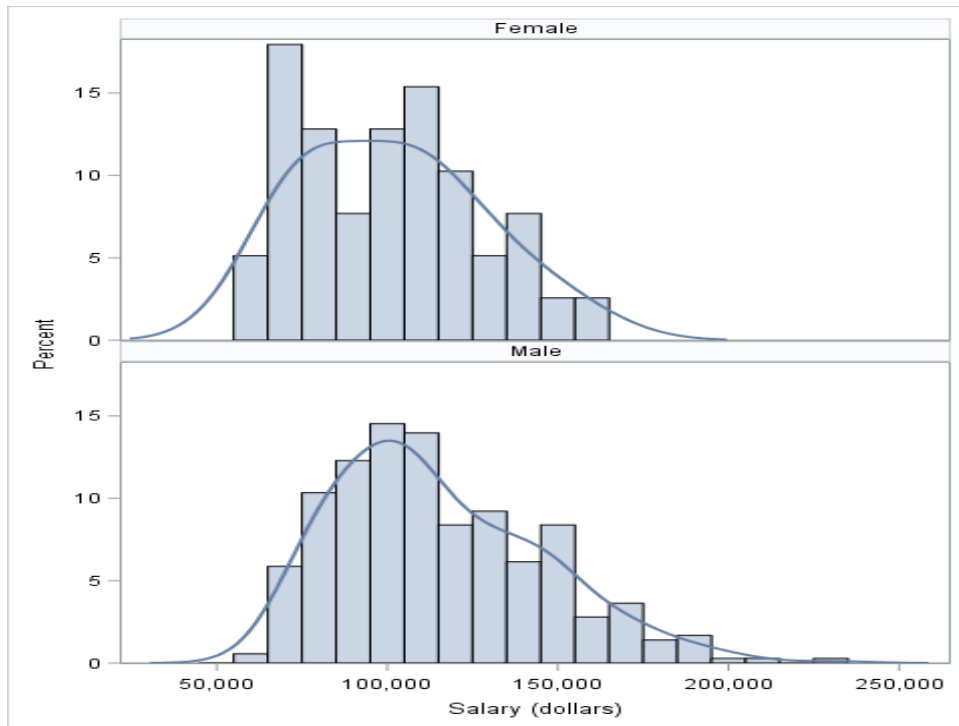
```

proc sgplot data=salaries;
  styleattrs datacontrastcolors=(red blue);
  histogram salary/ group=sex transparency=0.5 binwidth=10000;
  density salary / group=sex type = kernel ;
  xaxis label="Salary (dollars)";
  format salary comma.;
  title 'Nine month salary distribution by sex';
run;

```

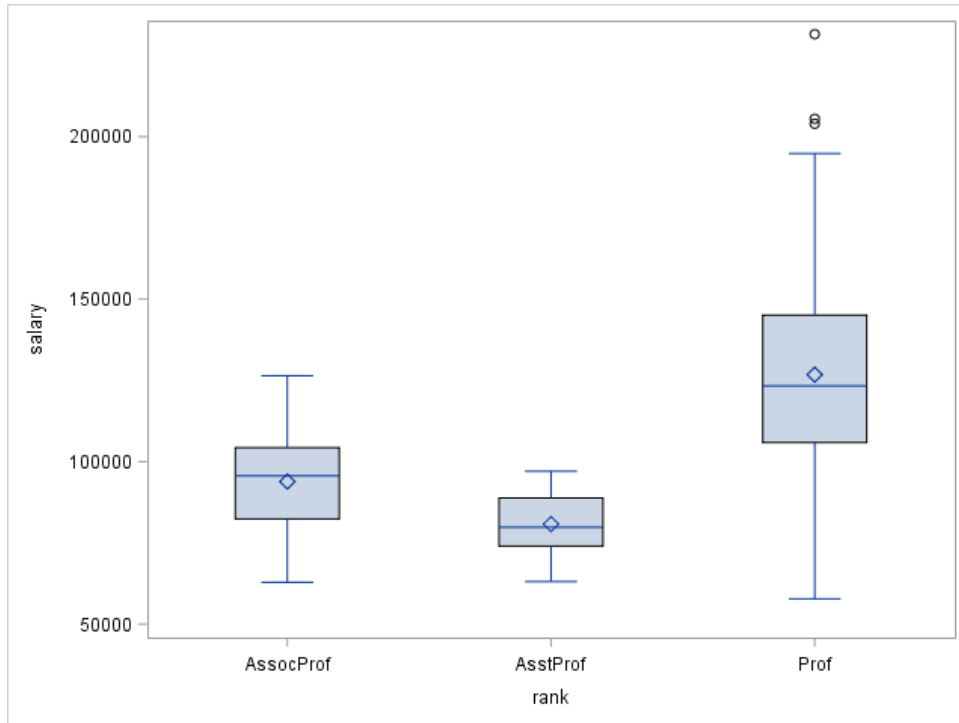


```
proc spanel data=salaries noautolegend;
  panelby sex/ rows=2 novarname;
  histogram salary/ binwidth=10000;
  density salary / type = kernel;
  format salary comma.;
  title 'Nine month salary distribution by sex';
  label salary = "Salary (dollars)";
run;
```

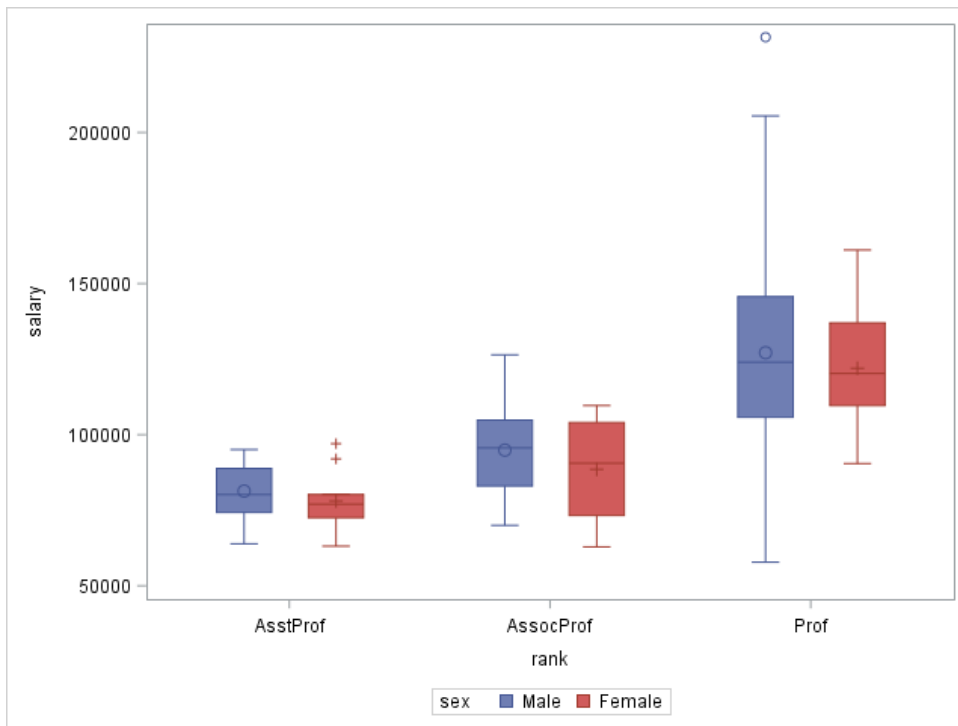


```
/* box plots */
```

```
proc sgplot data=salaries;  
  vbox salary / category=rank;  
run;
```



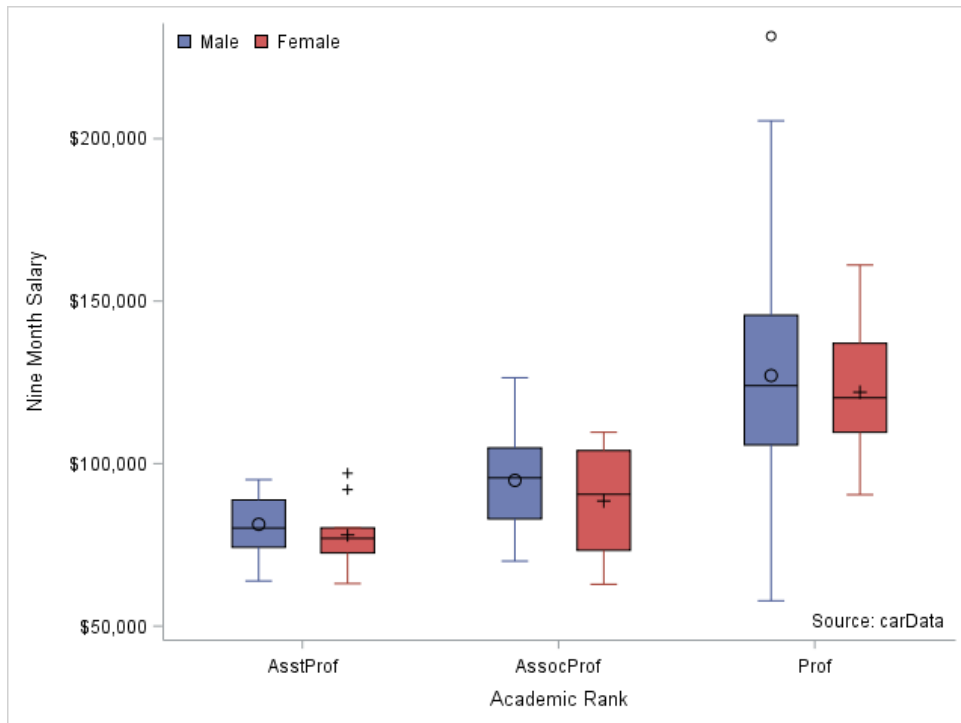
```
proc sgplot data=salaries;  
  vbox salary / category=rank group=sex;  
  xaxis type=discrete values=('AsstProf' 'AssocProf' 'Prof');  
run;
```



```

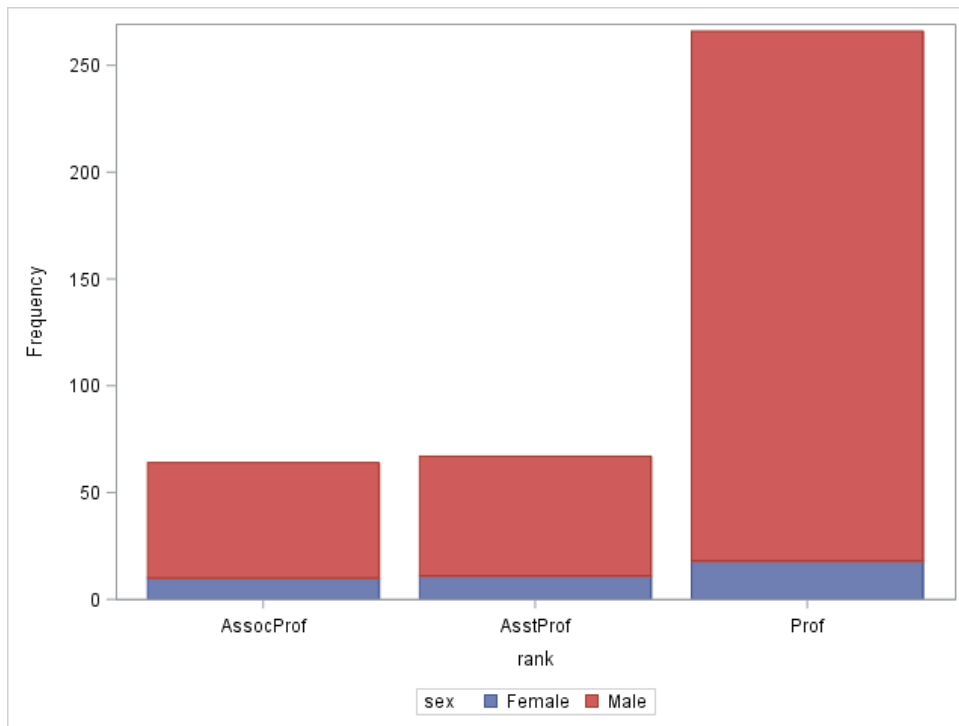
proc sgplot data=salaries noborder;
  vbox salary / category=rank group=sex
    lineattrs=(color=black) medianattrs=(color=black)
    meanattrs=(color=black) outlierattrs=(color=black);
  xaxis label="Academic Rank" type=discrete
    values=('AsstProf' 'AssocProf' 'Prof');
  yaxis label="Nine Month Salary";
  format salary dollar.;
  title 'Gender, Rank, and Salary';
  keylegend /location=inside position=topleft noborder title="";
  inset 'Source: carData'/postion=bottomright;
run;

```

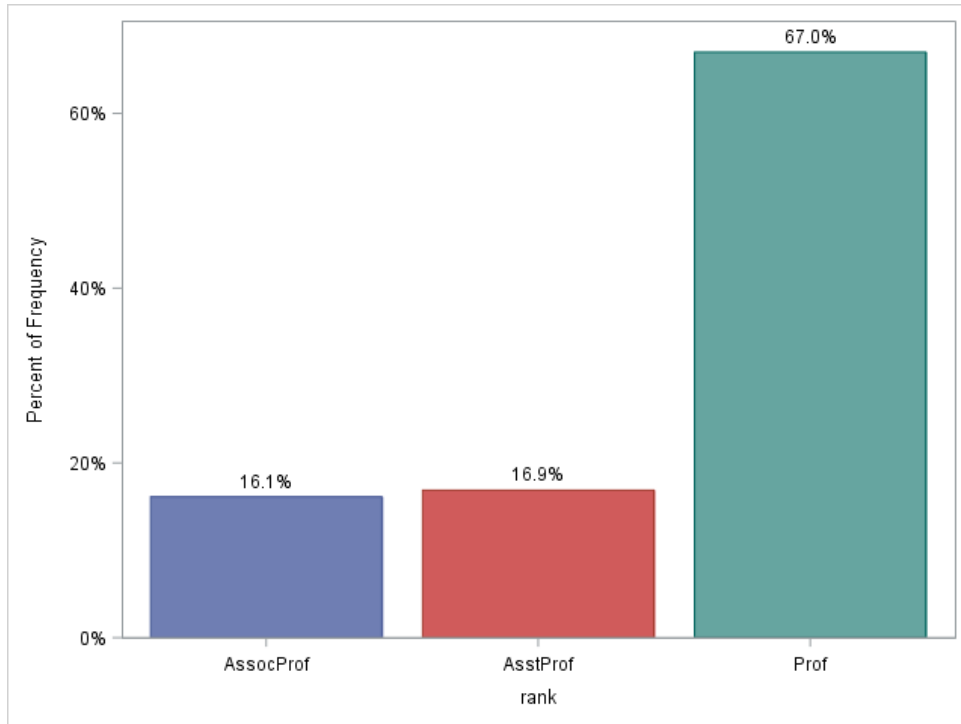


```
/* bar plots */
```

```
proc sgplot data=salaries;  
  vbar rank/group=sex;  
run;
```



```
proc sgplot data=salaries noautolegend;  
  vbar rank/stat=pct group=rank datalabel datalabelattrs=(size=9pt);  
run;
```

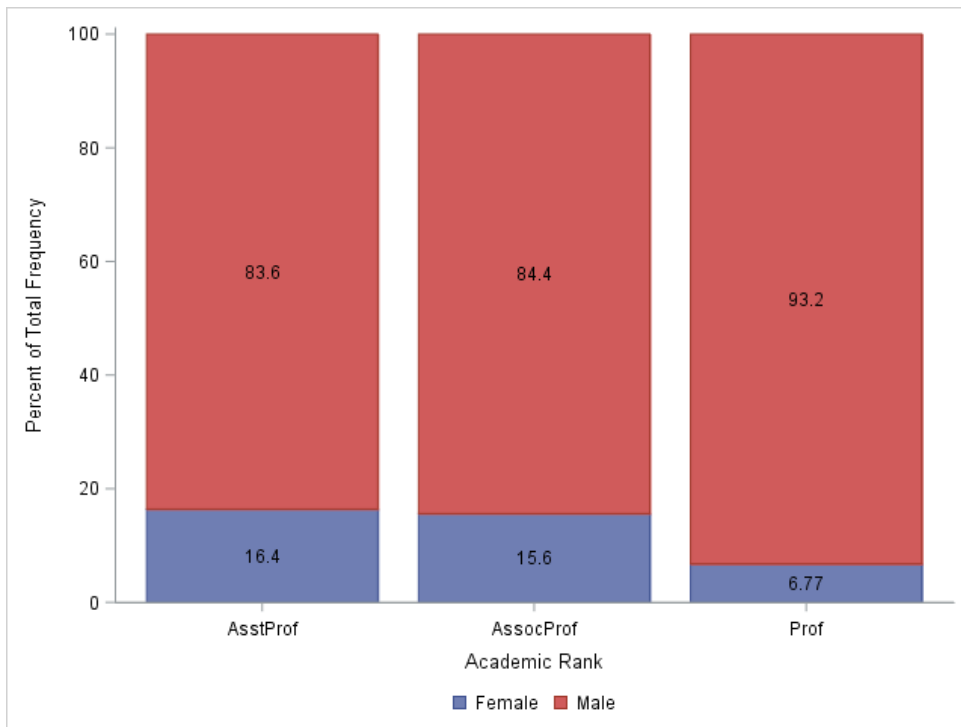




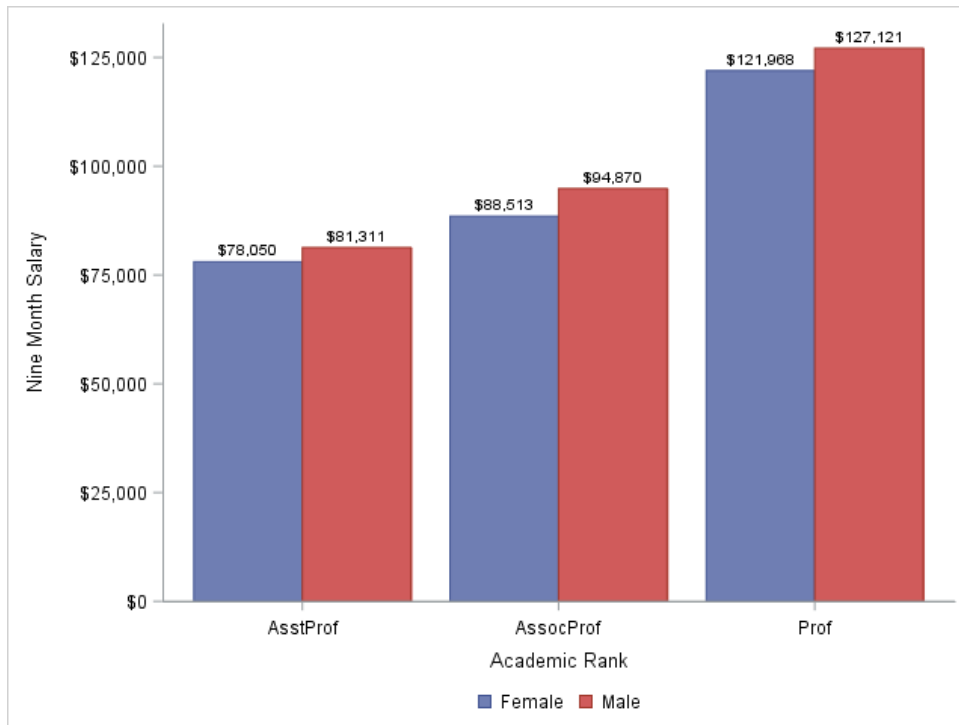
```

proc sort data=salaries out=salaries;
  by rank;
run;
proc freq data=salaries noprint;
  by rank;
  tables sex / out=FreqOut;
run;
proc sgplot data=FreqOut noborder;
  vbar rank / response=Percent group=sex groupdisplay=stack
  seglabel seglabelattrs=(size=9pt) seglabelformat=best4.1;
  xaxis label="Academic Rank" type=discrete
  values=('AsstProf' 'AssocProf' 'Prof');
  keylegend /noborder title="";
  title '100% Stacked Bar Chart';
run;

```

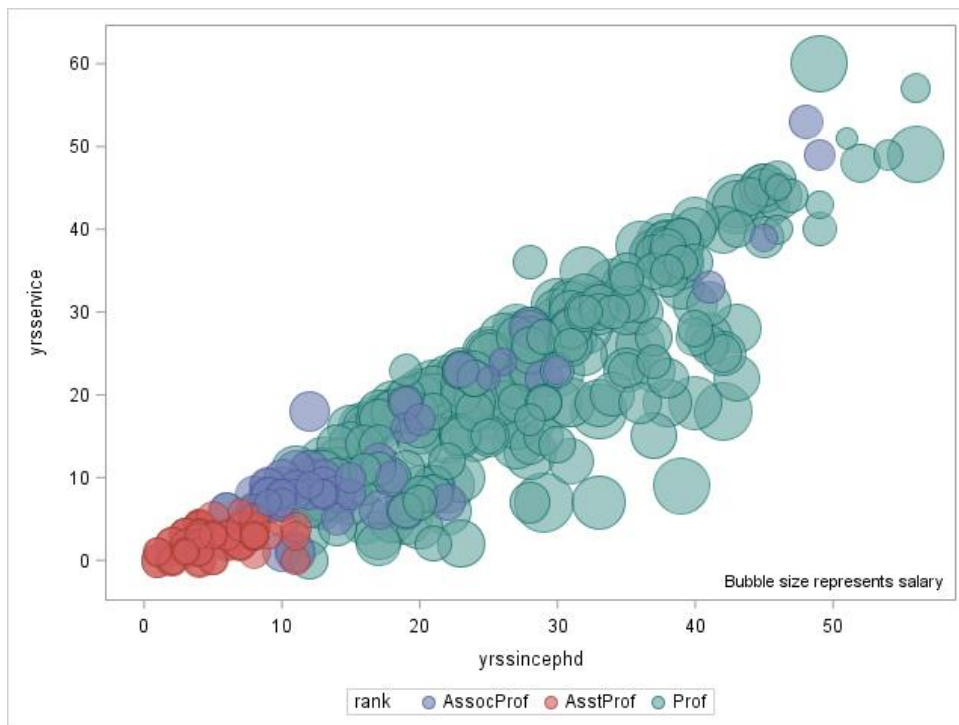


```
proc sgplot data=salaries noborder;
  vbar rank / response=salary group=sex groupdisplay=cluster
    stat=mean datalabel;
  xaxis label="Academic Rank" type=discrete
    values=('AsstProf' 'AssocProf' 'Prof');
  yaxis label="Nine Month Salary";
  format salary dollar.;
  title 'Mean Salary by Rank and Sex';
  keylegend /noborder title="";
run;
```



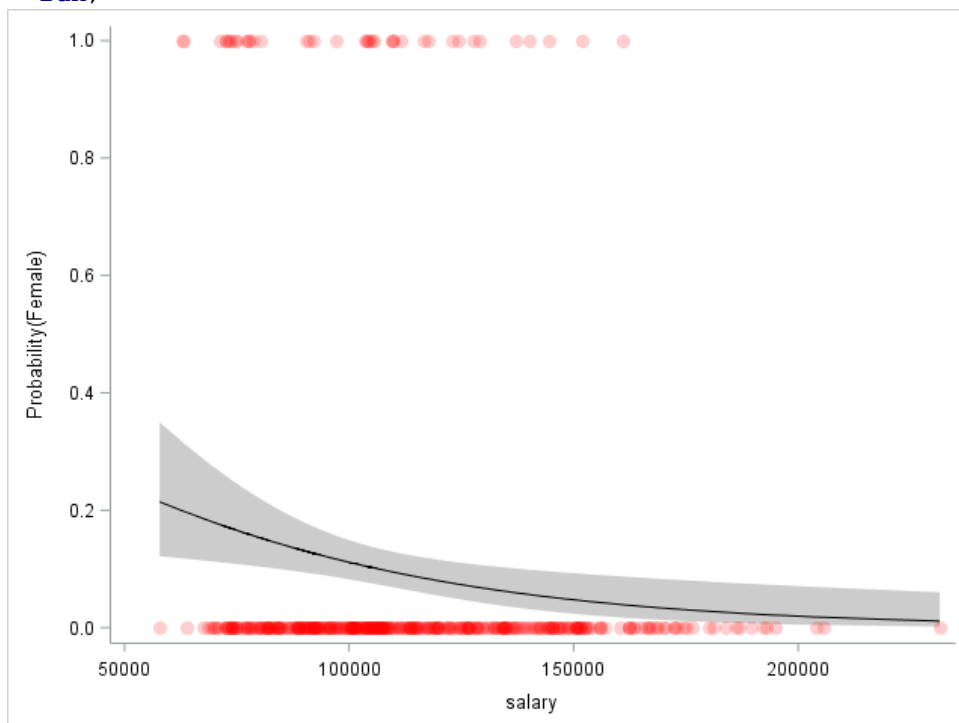
```
/* Bubble Plot */
```

```
proc sgplot data=salaries;  
  bubble x=yrssincephd y=yrsservice size=salary/  
  group=rank transparency=.4;  
  inset "Bubble size represents salary" / position=bottomright  
  textattrs=(size=8pt);  
run;
```



```
/* logistic regression */
```

```
proc logistic data=salaries
  plots=effect; /* use descending option for male */
class sex rank;
model sex = salary;
output out=new p=pred l=lower u=upper;
run;
data new;
  set new;
  if sex = 'Male' then sexcat = 0;
  if sex = 'Female' then sexcat = 1;
  drop sex;
run;
data new;
  set new;
  sex = sexcat;
run;
proc sort data=new out=new; by pred; run;
proc sgplot data=new noautolegend noborder;
  band x=salary upper=upper lower=lower/ fillattrs=(color=grey)
    transparency=.6;
  series x=salary y=pred / lineattrs=(color=black thickness=1);
  scatter x= salary y=sex/ markerattrs=(symbol=circlefilled size=10
    color=red) transparency=.8;
  yaxis label="Probability(Female)";
  title 'Relationship of Salary to Probability of Being Female';
  title2 'Predictions Based on Logistic Regression';
  footnote justify=right
    'Actual values, prediction line, and 95% confidence band';
run;
```



```

/* logistic regression - grouped */
proc logistic data=salaries; /* use descending option for male */
class sex rank;
model sex = salary rank;
output out=new p=pred l=lower u=upper;
run;
data new;
set new;
if sex = 'Male' then sexcat = 0;
if sex = 'Female' then sexcat = 1;
drop sex;
run;
data new;
set new;
sex = sexcat;
run;
proc sort data=new out=new; by rank pred; run;
ods graphics / attrpriority=none;
proc sgplot data=new noborder;
styleattrs datalinepatterns=(solid shortdash dash);
series x=salary y=pred / group=rank lineattrs=(thickness=2);
yaxis label="Probability ( Female )";
xaxis label="Nine Month Salary" values=(50000 to 200000 by 30000);
title 'Relationship of Salary to Probability of Being Female by Rank';
title2 'Predictions Based on Logistic Regression';
format salary dollar.;
keylegend / title="Academic Rank" location=inside position=topright
down=3 noborder;
run;

```

