Demo: Exploring Associations Using the Scatter Plot Task

Use the Scatter Plot task to examine the association between the response variable **SalePrice** and predictor variables in our **ameshousing3** data. We also want to see the general shape of each association. We'll start by generating a scatter plot to see whether there's an association between **SalePrice** and **Above Ground Living Area**.

1. In the Navigation pane, select **Tasks and Utilities**.
2. Expand **Tasks**.
3. Expand **Graph** and open the **Scatter Plot** task.
4. Select the **stat1.ameshousing3** table.
5. Assign **Gr_Liv_Area** to the X axis role and assign **SalePrice** to the Y axis role.
6. On the APPEARANCE tab, expand the **FIT CURVES** property and select **Regression** to add a regression fit to the scatter plot.
7. Click **Run**.

Generated Code

```
ods graphics / reset width=6.4in height=4.8in imagemap;

proc sgplot data=STAT1.AMESHOUSING3;
   reg x=Gr_Liv_Area y=SalePrice / nomarkers;
   scatter x=Gr_Liv_Area y=SalePrice /;
   xaxis grid;
   yaxis grid;
run;

ods graphics / reset;
```

Multiple Scatter Plots

1. Expand **Statistics** and open the **Data Exploration** task to plot multiple correlation plots simultaneously.
2. The **stat1.ameshousing3** table should be selected.
3. Select **Lot_Area**, **Gr_Liv_Area**, **Garage_Area**, **SalePrice**, **Basement_Area**, and **Deck_Porch_Area** as the Continuous variables.
4. On the PLOTS tab, clear the check box to output a Scatter plot matrix.
5. Select the **Regression scatter plots** option, and select **SalePrice** as the response variable.
6. Click **Run**.

Generated Code

```
options validvarname=any;
ods noproctitle;
ods graphics / imagemap=on;

/* Regression scatter plot macro */
%macro regressionScatterplot(xVar=, yVar=, title=, groupVar=);
        proc sgscatter data=STAT1.AMESHOUSING3;
                plot (&yVar)*(&xVar) / %if(&groupVar ne %str()) %then
                        %do;
                                group=&groupVar legend=(sortorder=ascending) %end;
                reg;
                title &title;
        run;
```

```
        title;
%mend regressionScatterplot;

%regressionScatterplot(xVar=Lot_Area, yVar=SalePrice,
        title="SalePrice vs Lot_Area");
%regressionScatterplot(xVar=Gr_Liv_Area, yVar=SalePrice,
        title="SalePrice vs Gr_Liv_Area");
%regressionScatterplot(xVar=Garage_Area, yVar=SalePrice,
        title="SalePrice vs Garage_Area");
%regressionScatterplot(xVar=SalePrice, yVar=SalePrice,
        title="SalePrice vs SalePrice");
%regressionScatterplot(xVar=Basement_Area, yVar=SalePrice,
        title="SalePrice vs Basement_Area");
%regressionScatterplot(xVar=Deck_Porch_Area, yVar=SalePrice,
        title="SalePrice vs Deck_Porch_Area");
```

The SAS Studio Data Exploration task limits the number of continuous variables to six and writes individual scatter plots to output. To plot more than five variables simultaneously in a panel plot, use PROC SGSCATTER.