Print

Demo: Exploring Ames Housing Data Using SAS Studio Tasks

Use the Table Analysis task to generate plots and tables for the categorical variables in the **ameshousing3** data set. Then use the Distribution Analysis task to generate plots and descriptive statistics for the continuous variables.

Generating Plots and Tables for Categorical Variables Using the Table Analysis Task

1. In the Navigation pane, select **Tasks and Utilities**.
2. Expand **Tasks**.
3. Expand **Statistics** and select the **Table Analysis** task.
4. On the DATA tab, click the **Select a table** icon and select the **stat1.ameshousing3** table.
5. Assign the following variables to the Row variables role. Use the Ctrl key to select multiple variables.
    - **House_Style**
    - **Overall_Qual**
    - **Overall_Cond**
    - **Year_Built**
    - **Heating_QC**
    - **Central_Air**
    - **Fireplaces**
    - **Mo_Sold**
    - **Yr_Sold**
    - **Garage_Type_2**
    - **Foundation_2**
    - **Masonry_Veneer**
    - **Lot_Shape_2**
6. On the OPTIONS tab, select **Cell** under Percentages, and select **Frequencies and percentages** under Cumulative.
7. Under STATISTICS, clear the **Chi-square statistics** check box.
8. Click **Run**.

Generated Code

```
ods noproctitle;

proc freq data=STAT1.AMESHOUSING3;
   tables  (House_Style Overall_Qual Overall_Cond Year_Built Heating_QC
            Central_Air Fireplaces Mo_Sold Yr_Sold Garage_Type_2 Foundation_2
            Masonry_Veneer Lot_Shape_2) / norow nocol plots(only)=(freqplot);
run;
```

Obtaining Descriptive Statistics for Continuous Variables Using the Distribution Analysis Task

1. Open the **Distribution Analysis** task under Statistics. Notice that the **stat1.ameshousing3** table is already selected. SAS Studio displays the last data set that was used.
2. Assign the following continuous variables to the Analysis variables role. Use the Ctrl key to select multiple variables.
    - **Lot_Area**
    - **Gr_Liv_Area**
    - **Bedroom_AbvGr**
    - **Garage_Area**
    - **SalePrice**

    - **Basement_Area**
    - **Full_Bathroom**
    - **Half_Bathroom**
    - **Total_Bathroom**
    - **Deck_Porch_Area**
    - **Age_Sold**
    - **Log_Price**
3. On the OPTIONS tab, select **Add normal curve**, **Add kernel density estimate**, and **Add inset statistics.**
4. Expand **Inset Statistics** and select **Mean** and **Standard deviation** in addition to the default, **Number of observations**.
5. Click **Run** to submit the generated code.

Generated Code

```
ods noproctitle;
ods graphics / imagemap=on;

/*** Exploring Data ***/
proc univariate data=STAT1.AMESHOUSING3;
   ods select Histogram;
   var Lot_Area Gr_Liv_Area Bedroom_AbvGr Garage_Area SalePrice Basement_Area
       Full_Bathroom Half_Bathroom Total_Bathroom Deck_Porch_Area Age_Sold Log_Price;
   histogram Lot_Area Gr_Liv_Area Bedroom_AbvGr Garage_Area SalePrice
       Basement_Area Full_Bathroom Half_Bathroom Total_Bathroom Deck_Porch_Area
       Age_Sold Log_Price / normal kernel;
   inset n mean std / position=ne;
run;
```