

Project Based Learner Series

Data Wrangling and Descriptive Statistics



Tom Grant

Principal Analytical Training Consultant
Global Academic Programs



Project Based Learner Series

Meet the Instructors



Cristina
Anton



Rachel
McLawhon



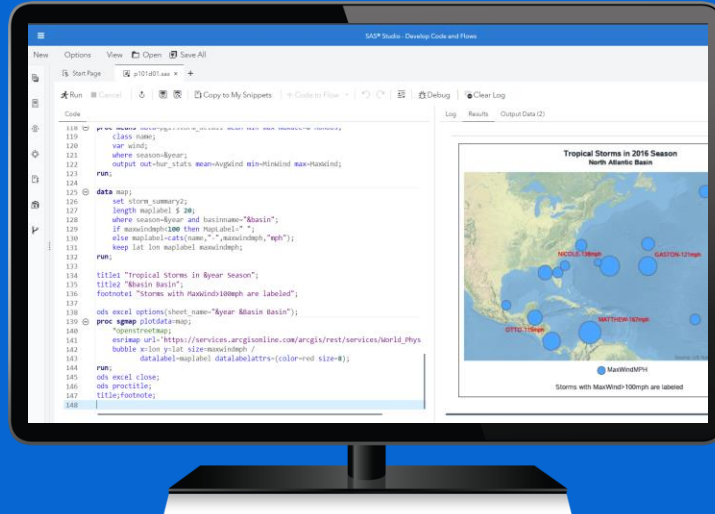
Jackie
Johnson



Tom
Grant

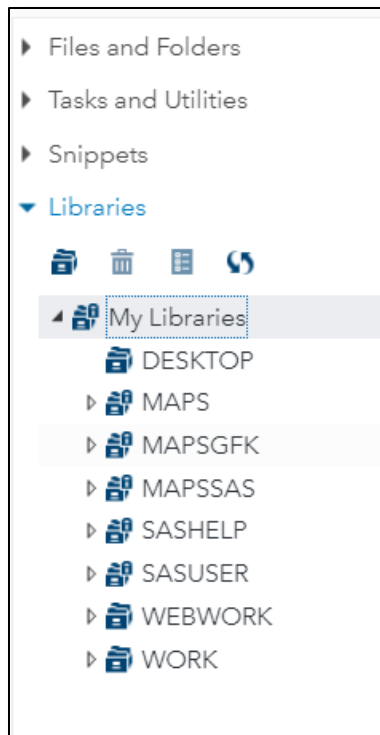
SAS Studio

Development Environment



- SAS Studio provides users with the ability to
- Use data processing, graphing, and analysis
 - Interact with a point-and-click interface to SAS that writes programs automatically.
 - Write SAS programs, which consist of commands in the form of DATA steps and PROC (or procedure) steps.

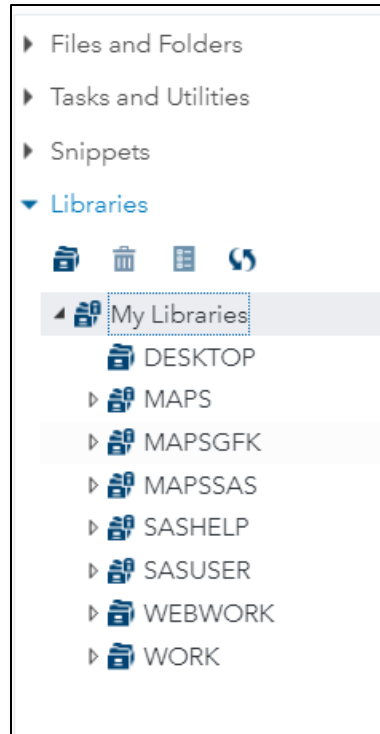
Libraries in SAS Studio



At the start of a SAS session, there are already a few predefined libraries available. Two of these libraries include:

- ❖ SAS Help Library
 - includes sample data for use
- ❖ Work Library
 - houses temporary data sets
 - contents deleted at the end of the SAS session

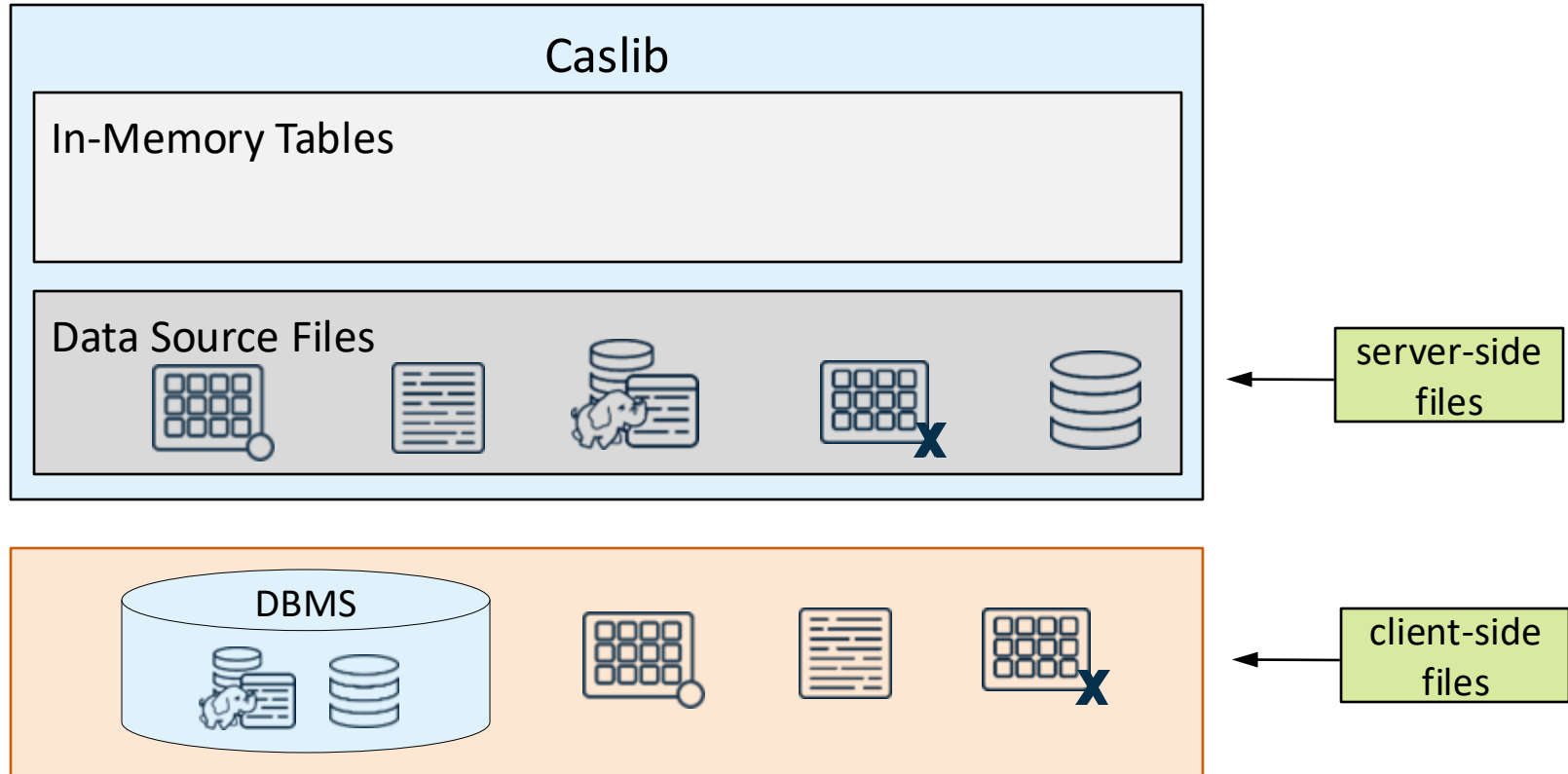
Guidelines for Libraries



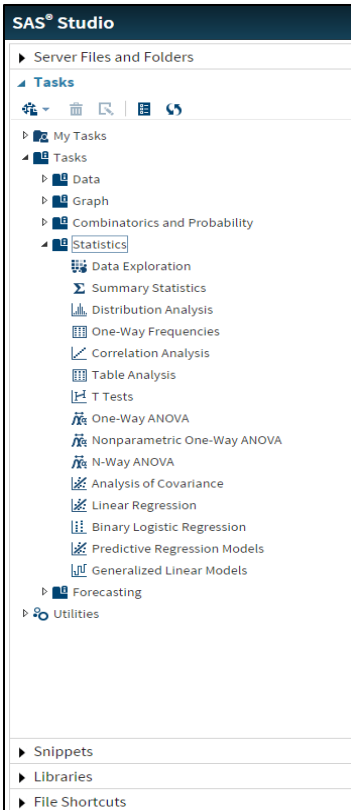
When creating a library, there are a few naming conventions that must be followed. Library names must

- ❖ have an eight-character maximum
- ❖ start with a letter or underscores
- ❖ continue with letters, numbers, or underscores.

Available Data Files



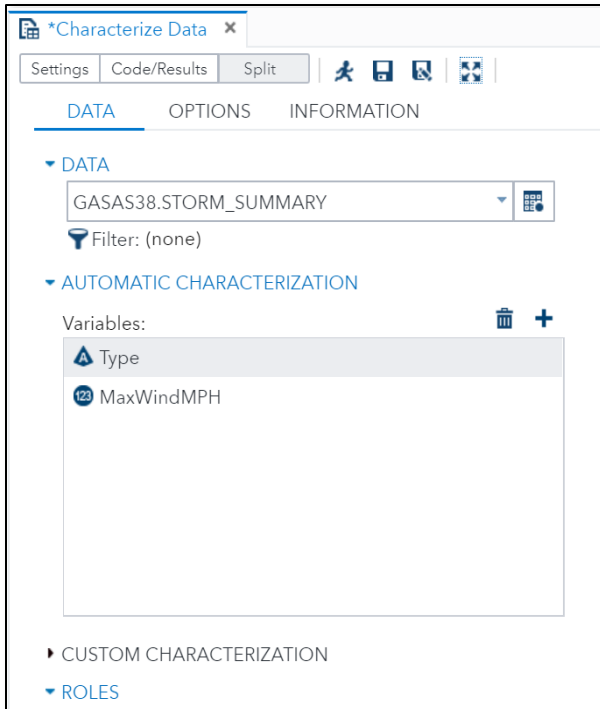
What Is a Task?



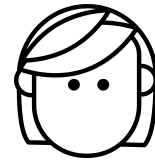
A task is a specific type of analysis, report, or data manipulation that you can perform against data in a project.



Characterize Data Task



Creates a summary
report of tables and
graphs to describe the
variables









Let's take a look at what's available in SAS Viya for Learners SAS Studio

The screenshot displays the SAS Studio web interface. At the top is a blue header bar with the text "SAS® Studio - Develop Code and Flows" and search, notification, and help icons. Below the header is a menu bar with "New", "Options", "View", "Open", and "Save All". The left sidebar contains a "Tasks" panel with a search bar and a list of task categories: "SAS Tasks" (highlighted) and "My Tasks". Under "SAS Tasks", there is a "Standard" category with sub-items: "Econometrics", "Forecasting", "Optimization and ..." (highlighted), "Prepare Data", "Statistical Process ...", "Statistics", and "Visualize Data". The main content area shows the "Start Page" with a "GET STARTED" section containing links for "Program in SAS", "Build a flow", "Import data", and "Query data". A "NEW" badge highlights "Explore new features in SAS Studio". Below this is a "LEARN" section with the text "Learn SAS Studio - videos, tutorials, and training". The right side of the main content area features a "Welcome to SAS STUDIO" message, describing it as "A rich environment to develop code and design flows." and a "Programs" section with the text "Write programs using the robust SAS language." and a diagram of a data flow. At the bottom of the interface, there is a status bar with "Start Page", "Recover (10)", "Submission (0)", and a help icon.

Demo

Tasks in SAS Studio

Let's take a look at what's available in SAS Viya for Learners Visual Analytics



SAS® Visual Analytics - Explore and Visualize





Choose Data

Home


Import data

< Back

Results: 1-7 of 7

Name	★	Library	Date Modified	Modifi
 HOMEEQUITY	☆	HELPDATA	Jul 19, 2024 10:26 PM	--
 HOME_EQUITY	☆	CASUSER(tom....	Aug 7, 2024 2:25 PM	Tom.G
 VS_BANK	☆	SVSO	Jul 19, 2024 10:25 PM	sasbo
 VS BANK	☆	TLINDATA	Jul 19, 2024 10:20 PM	sasbo

7 of 7

 HOME_EQUITY
CASUSER(tom.grant@sas.com)
Details Columns (18)
▼ Properties
Asset type: In-memory data
Date modified: Aug 7, 2024 2:25 ...

Add Cancel

Let's take a look at what's available in SAS Viya for Learners

Visual Analytics

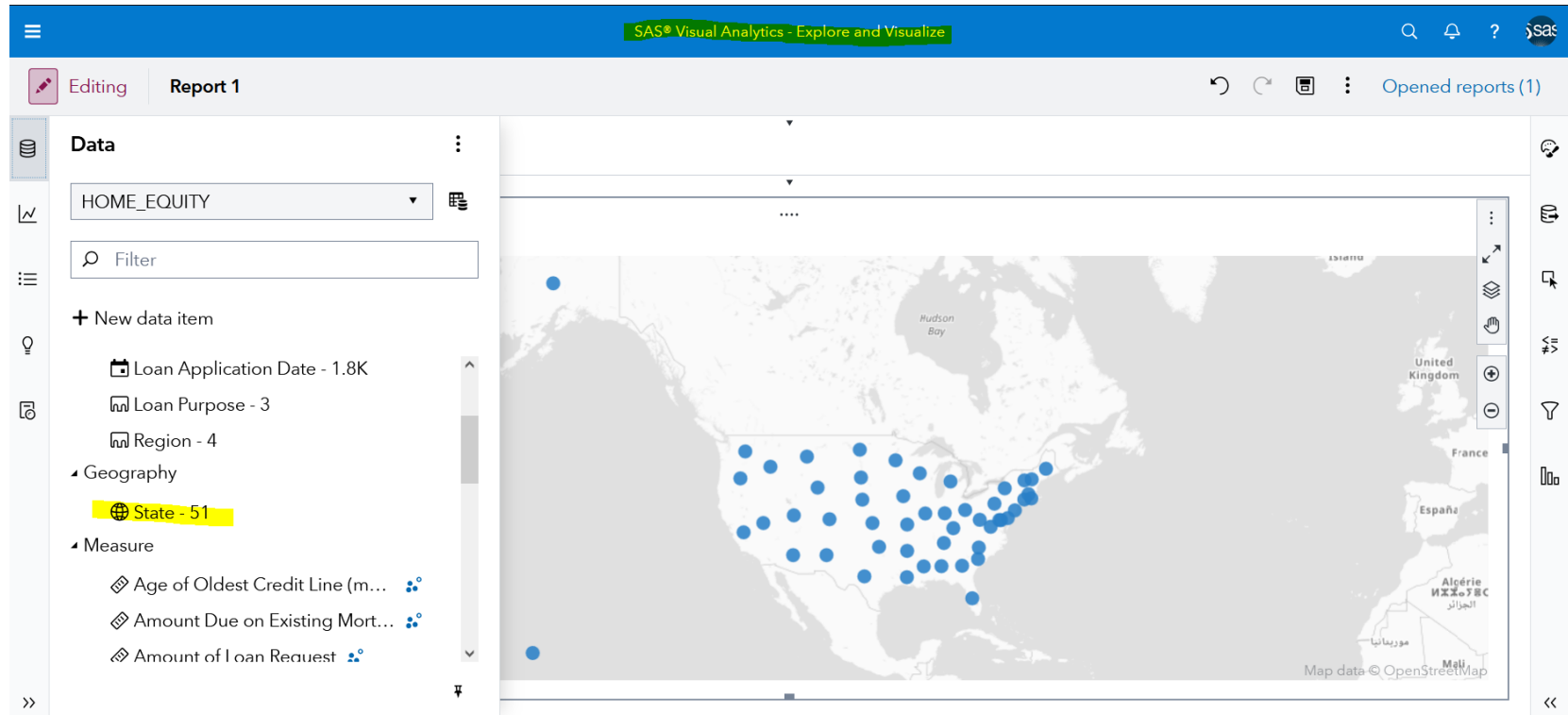
The screenshot shows the 'Choose Data' dialog in SAS Visual Analytics. The search bar contains 'Home' and shows 'Results: 1-7 of 7'. The table below lists the search results:

Name	★	Library	Date Modified	Modified By
HOME_EQUITY	☆	HELPDATA	Jul 19, 2024 10:26 PM	--
HOME_EQUITY	☆	CASUSER(tom....	Aug 7, 2024 2:25 PM	Tom.G
VS_BANK	☆	SVSO	Jul 19, 2024 10:25 PM	sasboc
VS_BANK	☆	TIINDATA	Jul 19, 2024 10:20 PM	sasboc

The 'HOME_EQUITY' dataset is selected. The details panel on the right shows the dataset's columns, including 'Loan Status (BAD)' and 'Amount of Loan Request (LOAN)'.

HOME_EQUITY
CASUSER(tom.grant@sas.com)
Details Columns (18)
③ **Loan Status (BAD)**
Index: 0, Length: 12 formatted, 8 raw
③ **Amount of Loan Request (LOAN)**
Index: 1, Format: DOLLAR, Length: 12 for...

Let's take a look at what's available in SAS Viya for Learners Visual Analytics



Let's take a look at what's available in SAS Viya for Learners

Visual Analytics

The screenshot displays the SAS Visual Analytics interface. The top navigation bar includes a menu icon, the title 'SAS® Visual Analytics - Explore and Visualize', and search, notification, and help icons. Below the navigation bar, the 'Editing' tab is active, and the report is titled 'Report 1'. The left sidebar shows a list of objects under the 'Objects' section, including 'Text input', 'Analytics', 'Automated explanation' (highlighted in yellow), 'Automated prediction', 'Forecasting', 'Network analysis', 'Path analysis', 'Text topics', 'Containers', 'Precision container', 'Prompt container', and 'Scrolling container'. The main content area is divided into three panels. The first panel, titled 'What are the characteristics of Loan Status?', shows a text-based explanation: '0 is more common at 80.05% (4.8K of 6K). 1 is less common at 19.95%. The three most related factors are Debt to Income Ratio, Number of Delinquent Credit Lines, and Value of Current Property.' The second panel, titled 'What factors are most related to Loan Status?', displays a horizontal bar chart showing the top factors: Debt to Income Ratio (blue), Number of Delinquent Credit Lines (yellow), Value of Current Property (purple), Number of Derogatory Credit Lines (orange), Age of Oldest Credit Line (teal), Number of Recent Credit Lines (red), and Amount of Loan Request (green). The third panel, titled 'What is the relationship between Loan Status and Debt to Income Ratio?', shows a bar chart comparing the Debt to Income Ratio for Loan Status 0 (blue) and Loan Status 1 (grey). The chart indicates that the average Debt to Income Ratio is 33 when Loan Status is 0, with a minimum of 0.72 and a maximum of 46. The average for Loan Status 1 is 19.95%.

What are the characteristics of Loan Status?
0 is more common at 80.05% (4.8K of 6K). 1 is less common at 19.95%.
The three most related factors are Debt to Income Ratio, Number of Delinquent Credit Lines, and Value of Current Property.

What factors are most related to Loan Status?

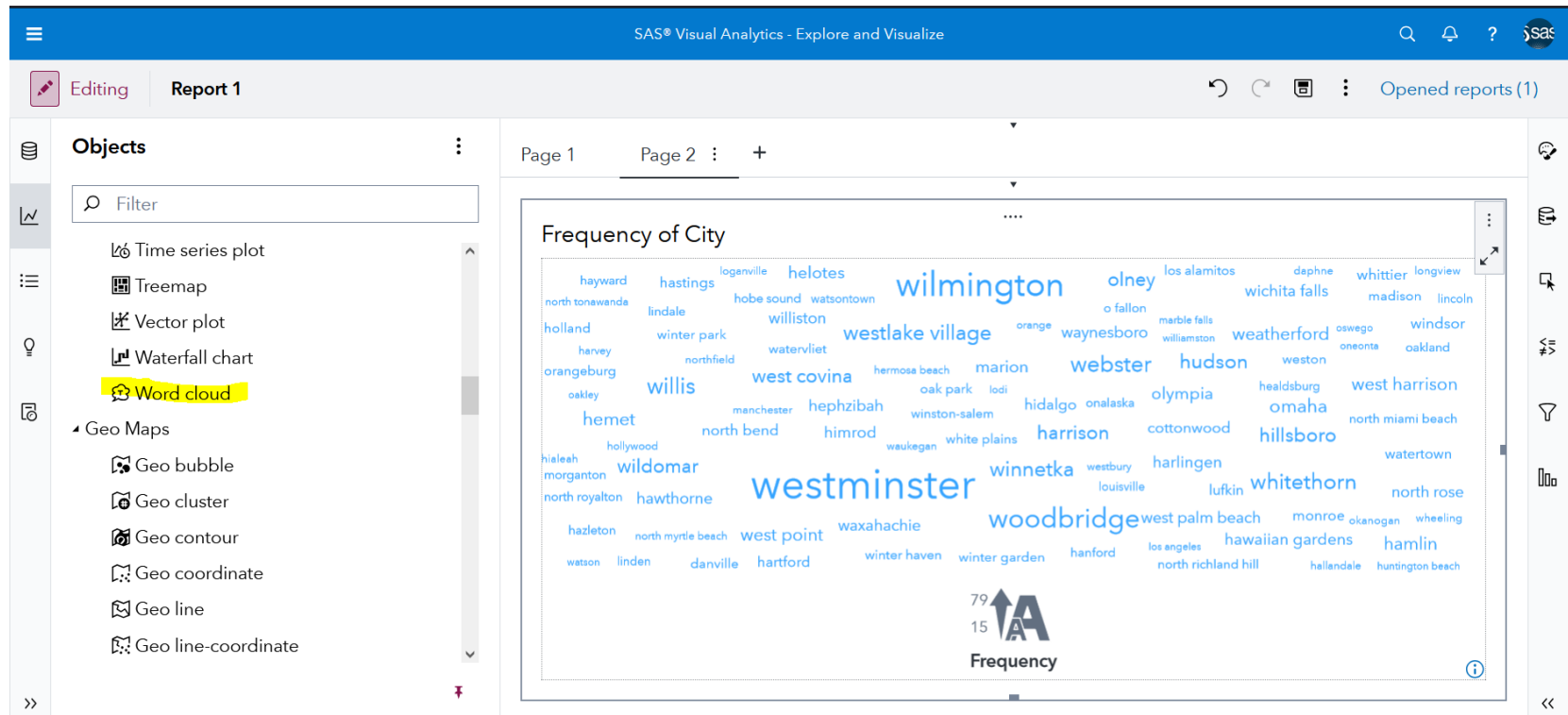
Factor	Relative Importance (Color)
Debt to Income Ratio	Blue
Number of Delinquent Credit Lines	Yellow
Value of Current Property	Purple
Number of Derogatory Credit Lines	Orange
Age of Oldest Credit Line	Teal
Number of Recent Credit Lines	Red
Amount of Loan Request	Green

What is the relationship between Loan Status and Debt to Income Ratio?

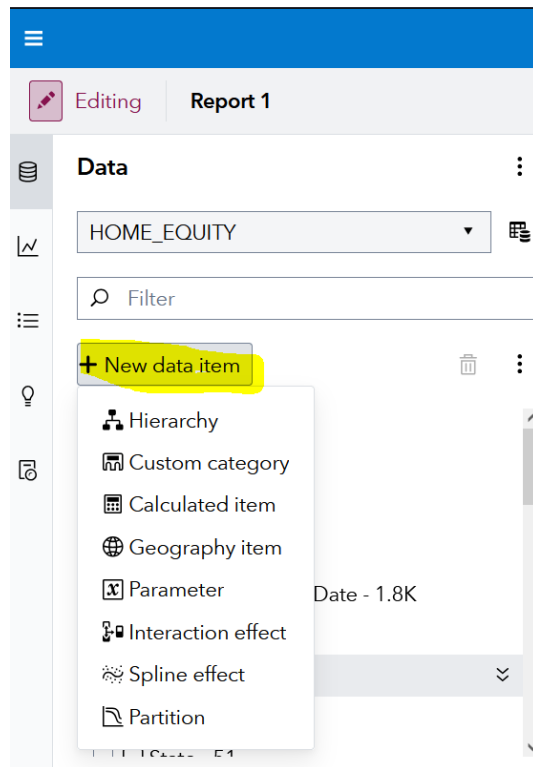
The average Debt to Income Ratio is 33 when Loan Status is 0, with a minimum of 0.72 and a maximum of 46. The average for Loan Status 1 is 19.95%.

Let's take a look at what's available in SAS Viya for Learners

Visual Analytics



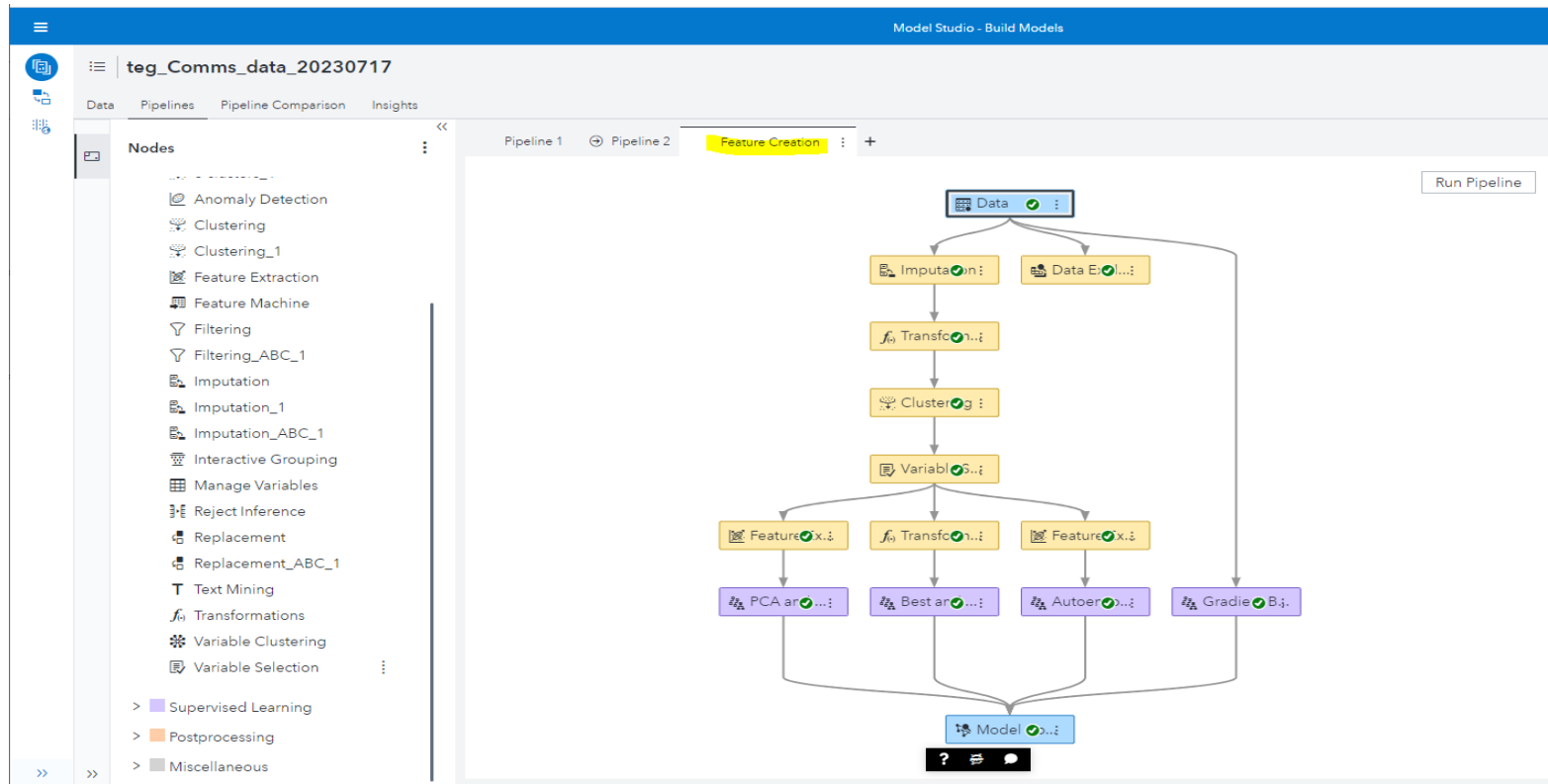
Let's take a look at what's available in SAS Viya for Learners Visual Analytics



Demo

Visual Analytics

Let's take a look at what's available in SAS Viya for Learners Model Studio



Demo

Model Studio

Discussion

How comfortable are YOU with letting a “computer” decide when and how to manipulate your data??



See you next time!

