

Paper SAS4398-2020
Sailing Your Data Lake with Self-Service Data Prep
Peter Baquero, SAS Institute Inc

ABSTRACT

Did the crystal clear analytics promised by your data lake never materialize? Have you been left with a murky data swamp, fearing what might be lurking beneath the surface? There are many reasons that the promise of data lakes hasn't come to fruition. One of those reasons is the lack of a single comprehensive tool to connect, integrate, transform, and govern the data in a data lake. This tool should be usable by citizen data scientists to deliver true self-service data preparation. It's easy to understand why there hasn't been a single tool that connects, integrates, transforms, and governs data in a data lake. By their nature, data lakes are not homogenous environments. Any tool working with data in a data lake needs to be able to work with structured, semi-structured, unstructured, and binary data. Integration is the next challenge. Baseline data integration should include basic table manipulation such as joining, appending, subsetting, and so on. Integration should be GUI-driven and not require advanced coding. During and after the data integration process, the data should be transformed to meet data quality standards. Both integration and transformation should be able to leverage artificial intelligence (AI) and machine learning to facilitate and streamline the processes. Finally, the tool should have data governance so that users can know what data they are working with. Self-service data preparation is the tool that can help make data lake dreams a reality.

INTRODUCTION

Data lakes held the promise of revolutionizing enterprise level data analytics and reporting. Gartner was an early critic of the hype, predicting that "...60 percent of big data projects will fail to go beyond piloting and experimentation, and will be abandoned."¹ Gartner analyst Nick Heudecker thinks that estimate is too conservative. The real number – closer to 85 percent². So what went wrong with data lakes? The idea is good – store all your data in a single repository in various states of transformation to facilitate access by consumers. The failure point for data lakes comes when users try to access the data. Users need a single tool to read a variety of data formats, understand the data they are working with, and easily perform complex data transformations. Let's take a look at how SAS Data Preparation is that tool by making data access, data profiling, data integration, data governance, and enterprise collaboration a breeze.

DATA LAKES – PROMISES WERE MADE

¹ <https://www.gartner.com/en/newsroom/press-releases/2015-09-15-gartner-says-business-intelligence-and-analytics-leaders-must-focus-on-mindsets-and-culture-to-kick-start-advanced-analytics>

² <http://www.globalbigdataconference.com/news/141927/85-of-big-data-projects-fail-but-your-developers-can-help-yours-succeed.html>

Let's start by digging deeper into the specific challenge data lakes were supposed to address. To do that we need to take a look at a standard enterprise data lifecycle for new analytic or reporting endeavors. Typically, data requests originate when an analyst has an idea or question that requires data to be answered. The analyst will submit a data request to their IT department. The data request is picked up by a data engineer. The engineer connects to the various enterprise data sources, performs any necessary integration and transformation, and produces an analytic base table that is consumed by the analyst. This process is summarized in Figure 1.

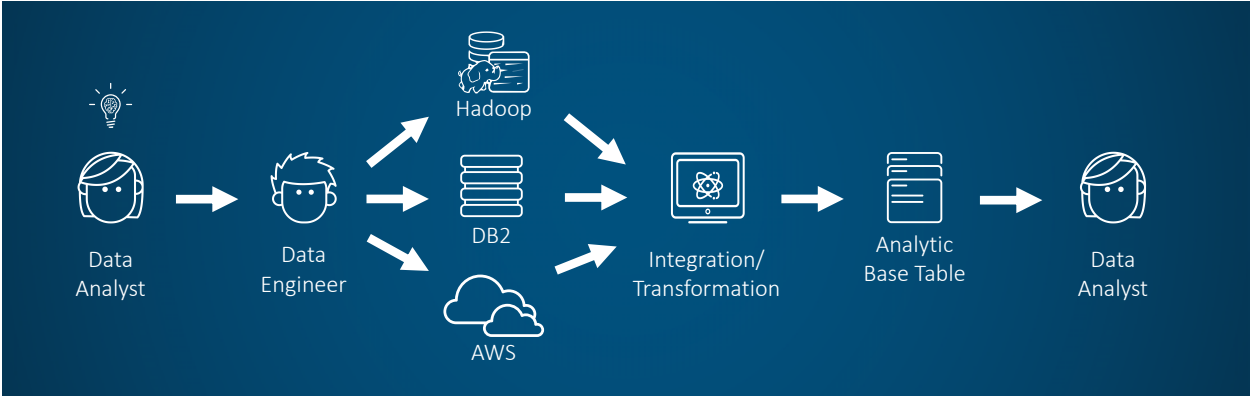


Figure 1 Standard Enterprise Data Lifecycle for New Analytic or Reporting Request

Data lakes were supposed to short circuit this process by giving data analysts access to source files in various states of transformation. Data lakes failed because they did not account for the difference in skill set between a data analyst and a data engineer.

ENTERPRISE KNOWLEDGE WORKERS

The Enterprise Data Lifecycle in Figure 1 has two main actors; data engineers and data analysts. These personas have different skill sets because they work with the data in fundamentally different ways. Data engineers are going to be very technically savvy. They are going to know the differences between different SQL languages and the nuances of data storage in different RDBMSs. They won't have familiarity with the data beyond where it sits, where it needs to go, and how it has to look at the end. They are not risk modelers, biostatisticians, or economists, so they don't have domain expertise necessary to evaluate how fit for use the data is to address the specific question being asked.

Data analysts, on the other hand, do have that domain expertise. They are the ones who work with the data directly to answer specific questions. Data analysts really are an amalgamation of many personas; report builder, business analyst, data scientist. They may have varying levels of technical expertise. Some analysts may be comfortable building reports only in a drag and drop reporting environment. Other analysts may be comfortable writing code in various languages for various purposes.

Since they data engineer and the data analyst are working with the data in different capacities, their preferred interaction model will differ as well. Data engineers will want to touch the data as few times as possible to build out slowly changing enterprise class data flows. The data analyst on the other hand will want to continuously touch the data as they go through a highly iterative, agile process of answering their question. Figure 2 gives a great overview of the split between these personas.

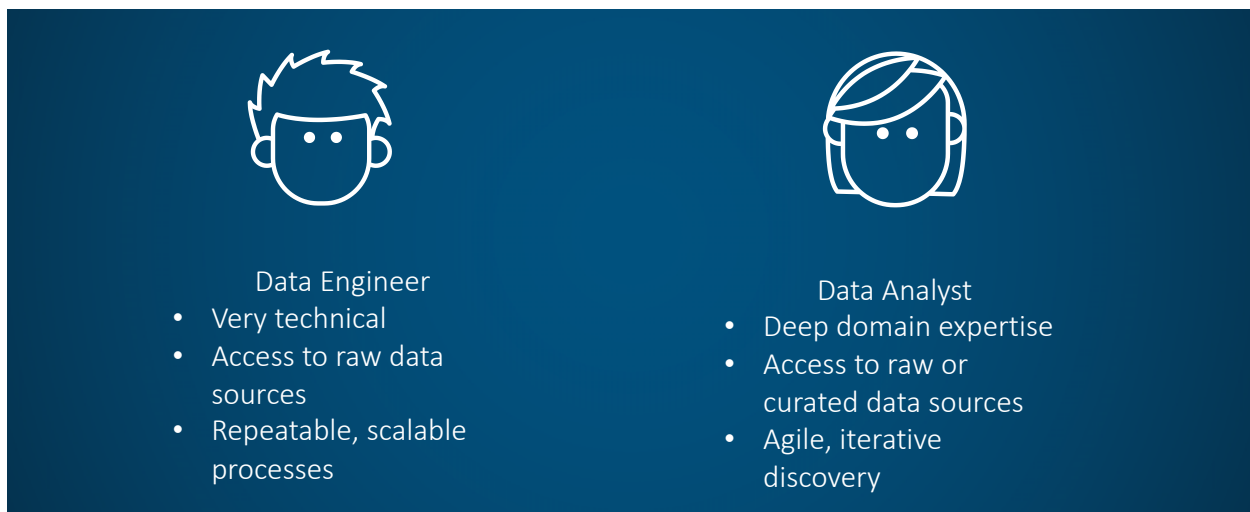


Figure 2 Enterprise Knowledge Worker Personas

The split in skill sets between a data engineer and a data analyst makes it very clear why data lakes failed: there wasn't a self-service data preparation tool that could bridge the gap of the broad amalgam of skill sets under the data analyst persona and the data engineer persona. The success of data lakes depended on the ability of data analysts to not only access the data, but to also understand the data and perform the necessary, sometimes substantial, transformation work to create an analytic base table. Let's take a look at how SAS Data Preparation is that tool.

SAS DATA PREPARATION – PICKING UP WHERE DATA LAKES STOPPED

The four main areas where a tool can help a data analyst working with a data lake are access, data profiling, data integration, data governance, and enterprise collaboration. Let's take a look at each area individually.

ACCESS

It all starts with data access. SAS Data Preparation can access many of the leading cloud vendor platforms as well as traditional on premise data sources. SAS has specifically designed ACCESS interfaces and data connectors to Amazon Redshift, Google BigQuery, Hadoop, MongoDB, Salesforce, Oracle, PC Files, Impala, Snowflake, Vertica, and many more³. ODBC and JDBC SAS/ACCESS interfaces add further extensibility. Wherever your data sits, chances are SAS Data Preparation can access it with a dedicated SAS/ACCESS interface. For the purposes of this paper let's assume our data lake is in Amazon S3.

Now that we can get to our data, can we read it? Amazon claims that they "...can store virtually any kind of data in any format."⁴ Being able to access virtually any kind of data in any format is a big challenge for a self-service data preparation tool. But SAS Data

³ <https://support.sas.com/en/software/sas-access.html#1740662f-92d7-419d-9502-2855e7055a8d>

⁴ <https://aws.amazon.com/s3/faqs/>

Preparation has us covered! In addition to being able to access data in any of the relational SAS/ACCESS interfaces, SAS Data Preparation can also access PDF files, CSV files, ORC files, Parquet files, Excel workbooks, and social media feeds.

After you configure your connection to S3 through SAS Data Explorer, you will be able to navigate your S3 bucket like any other folder structure. Figure 3 shows the files in we will be working with in our S3 bucket on AWS. Figure 4 shows the same files in SAS Data Explorer.

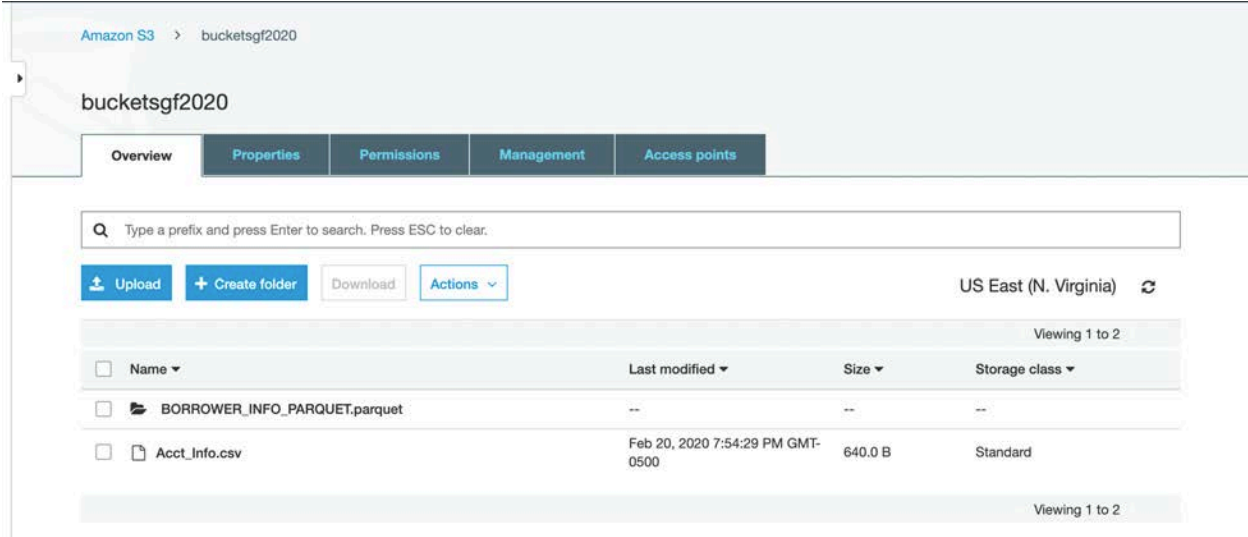


Figure 3 Files in S3

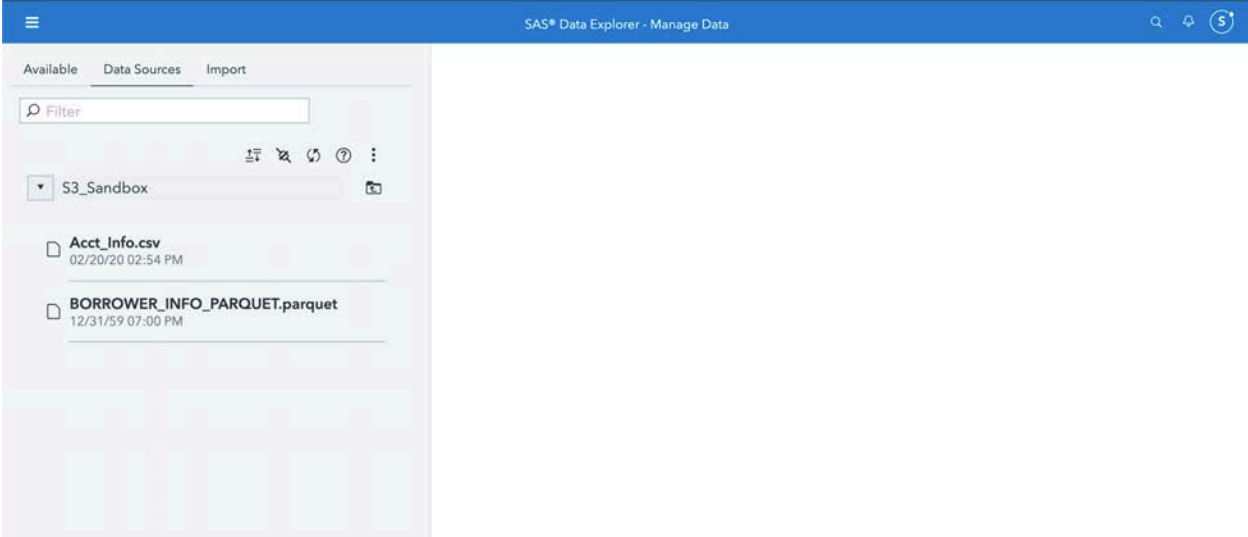


Figure 4 S3 Files in SAS Data Explorer

We will work with two sample files for the rest of this paper; Acct_Info and Borrower_Info. Acct_Info is in csv format and has all the account information for credit card holder. Borrower_Info is in parquet format and has all the demographic information for the credit card holders. We can use these two files to answer business questions about the relationship between credit limit and age, or the geographical concentration of credit card

holders, or the geographical concentration of credit exposure, or any other number of questions. Clicking the lightning bolt icon in SAS Data Explorer will load these files from S3 into memory as shown in Figure 5.

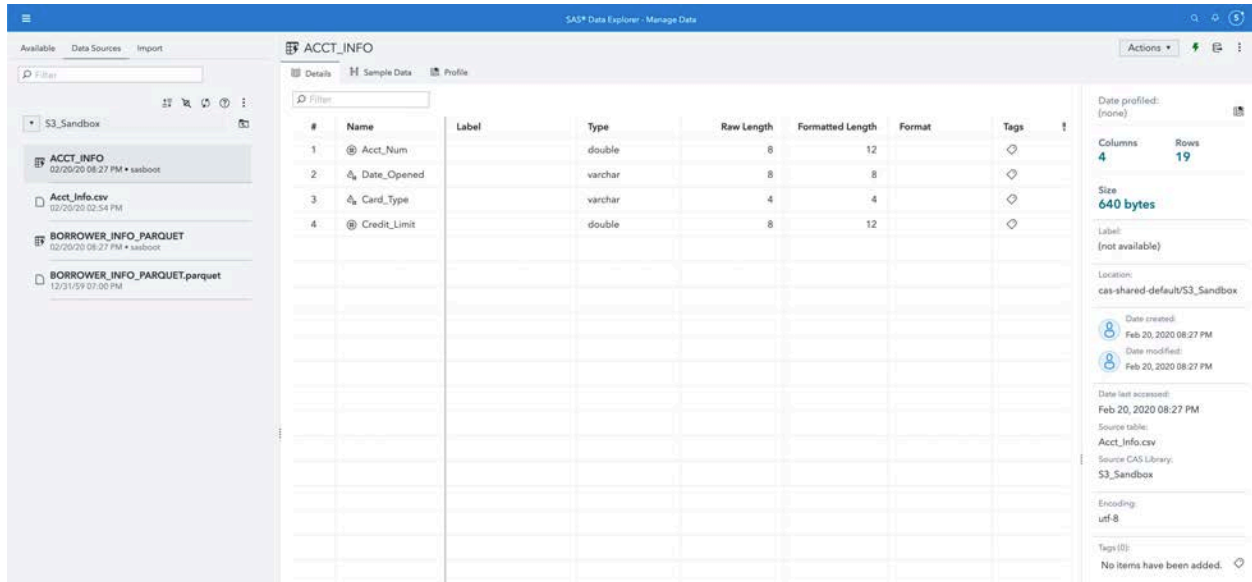


Figure 5 S3 Files Loaded Directly to Memory

DATA PROFILING

Now that we can get to our data and read it into our analytic platform, we have to evaluate its fit for use for our specific question. For the purposes of this paper we want to create an analytic base table to explore the relationship between borrower age and credit limit.

SAS Data Explorer gives us a better understanding of the data before we even load it into SAS Data Preparation for any transformation work. Figure 5 shows some of information SAS Data Explorer makes available for all datasets. This is a great place to perform a sanity check on the data to make sure you have everything you need at a high level. Tables in SAS Data Explorer have three tabs; Details, Sample Data, and Profile.

The Details tab lists all the available metadata for the table including column names, types, and lengths. This is a great place to determine if you have the right columns to answer your business question and see if those columns are in an appropriate format. For our purposes we need fields that represent a unique account, the account’s credit limit, and the age of the account holder. We can see ACCT_INFO has Acct_Num and Credit_Limit. SAS Data Explorer also gives us the data types. We can see that both Acct_Num and Credit_Limit are numeric.

Figure 6 shows the information we have in the Sample Data tab of SAS Data Explorer. We have switched to the BORROW_INFO_PARQUET table to ensure we have our third required data point, customer name. In the Sample Data tab we can see actual values from our table. We see that we have a Name field and we have an Acct_Num field, so we can join back to our ACCT_INFO table.

Acct_Num	Name	address	age	AGE_RNG	atmuser	behave_scr	City
3483357	Dwayne Johnson	2590 SW 175TH AVE	45	36 - 45		380	San Diego
72127	Tom Brady	RR 2 BOX 2014	35	26 - 35		450	San Diego
9243828	Zsa Zsa Gabor	945 NW CIRCLE AP...	46	46 - 55		220	Tucson
8123180	Jennifer Aniston	244 SW RIMROCK ...	55	46 - 55		220	Provo
7611353	Steve Martin	6150 SW KING BLV...	31	26 - 35		260	Phoenix
6164356	Betty White	63533 J D ESTATES...	39	36 - 45		460	Provo
7538391	Andy Warhol	2718 SE PINE LN	32	26 - 35	N	500	Denver
1175348	Bob Ross	5410 SW 202ND AVE	43	36 - 45	N	380	San Diego
2748056	Linda Ronstadt	819 S COMSTOCK ...	30	26 - 35		500	Portland
5399690	Mickey Rourke	25 EVANS ST APT 1	35	26 - 35		300	Los Angeles
9696195	George Clooney	18020 SW HART DR	45	36 - 45		420	San Jose
2029997	Claire Danes	14225 SW 22ND AVE	36	36 - 45		380	San Francisco
4206913	Tom Selleck	2060 SW 187TH AVE	59	>55		420	San Diego
6830786	Shirley Temple	5824 SW BYTN-HLS...	34	26 - 35	Y	220	Phoenix
6164356	Betty White	63533 J D ESTATES...	39	36 - 45		460	Provo
9696195	George Clooney	18020 SW HART DR	45	36 - 45		420	San Jose
2029997	Claire Danes	14225 SW 22ND AVE	36	36 - 45		380	San Francisco
4206913	Tom Selleck	2060 SW 187TH AVE	59	>55		420	San Diego
7611353	Steve Martin	6150 SW KING BLV...	31	26 - 35		260	Phoenix

Figure 6 Borrow Info Sample Data

Figure 7 shows us what we will find in the third and final tab of SAS Data Explorer, the Profile tab. There is a lot of very useful information here. First, we see that Acct_Num is not unique. Name is also not unique but this might be ok if individuals have more than one credit card account. Since Acct_Num and Name are not unique in the same exact percentages, it would seem that we have duplicate records in our data. We can also see that the total number of unique Acct_Nums and Names is 14. We'll want to check the ACCT_INFO table for duplicates. Let's go ahead and load it into SAS Data Preparation!

Column	Unique	Null	Blank	Pattern Count	Mean	Median	Mode	Standa...	Standa...	Minimum	Maximum
AGE_RNG	21.00% (4)			2			36 - 45			26 - 35	>55
Acct_Num	75.49% (14)				5,475...	6,164...		2,972...	681,9...	72,127.00	9,696,195.00
City	47.37%			8			San Di...			Denver	Tucson
Name	71.69% (14)			13						Andy Warhol	Zsa Zsa Ga...
address	75.49% (14)			12						14225 SW ...	RR 2 BOX ...
age	64.17%				40.79	39.00	45.00	9.13	2.10	30.00	59.00
atmuser	64.67% (10)			1			N			N	Y
behave_scr	42.11%				371.05	380.00		95.27	21.86	220.00	500.00

Figure 7 Borrower Info Profile

Figure 8 shows ACCT_INFO loaded into SAS Data Preparation. We were worried that there were duplicates as we saw in BORROWER_INFO. In the Transformations panel on the left we see a transformation called Remove Duplicates. We can deploy that transformation simply by double clicking it. Figure 9 shows the details of the Remove Duplicates transformation in the bottom panel beneath the dataset and also the transformation as a step in the Plan in the panel on the right. Figure 10 shows the results of running the Remove Duplicates transformation. We now have 14 rows in the table, which ties out to what we saw in the Profile tab in SAS Data Explorer. We just deduped a dataset with only three mouse clicks and without writing a single line of code! Let's do the same thing for BORROWER_INFO.

SAS® Data Studio - Prepare Data

Plan 1

Table Profile Metadata

ACCT_INFO

Result rows: 100

Acct_Num	Date_Opened	Card_Type	Credit_Limit
1175348	12/25/78	VISA	10000
2748056	03/08/83	VISA	15000
5399690	10/12/08	VISA	10000
9696195	11/15/93	VISA	20000
2029997	03/09/09	MC	15000
4206913	02/15/87	VISA	30000
6164356	01/23/92	MC	35000
9696195	11/15/93	VISA	20000
2029997	03/09/09	MC	15000
4206913	02/15/87	VISA	30000
7611353	04/09/06	VISA	45000
6830786	07/28/82	VISA	50000
3483357	02/02/02	MC	45000
72127	06/12/98	VISA	27000
9243828	09/04/04	VISA	32000
8123180	10/03/03	VISA	17000
7611353	04/09/06	VISA	45000
6164356	01/23/92	MC	35000
7538391	05/06/97	MC	12000

Add a transform to the plan.

Plan

Name: Plan 1

Modified: 02/20/20 08:54 PM

Add a transform to the plan.

Transforms

- Column Transforms
 - Change case
 - Convert column
 - Manage columns
 - Remove
 - Rename
 - Split
 - Trim whitespace
- Custom Transforms
 - Calculated column
 - Code
- Data Quality Transforms
 - Casing
 - Field extraction
 - Gender analysis
 - Identification analysis
 - Match and cluster
 - Matchcodes
 - Parsing
 - Remove duplicates
 - Standardize
- Multi-input Transforms
 - Append
 - Join
- Row Transforms
 - Analytic partitioning
 - Filter
 - Transpose
 - Unique identifier

Figure 8 ACCT_INFO Loaded into SAS Data Preparation

SAS® Data Studio - Prepare Data

Plan 1

Table Profile Metadata

ACCT_INFO

The session table is not current to the plan. Run the plan to update the table. Result rows: 100

Acct_Num	Date_Opened	Card_Type	Credit_Limit
1175348	12/25/78	VISA	10000
2748056	03/08/83	VISA	15000
5399690	10/12/08	VISA	10000
9696195	11/15/93	VISA	20000
2029997	03/09/09	MC	15000
4206913	02/15/87	VISA	30000
6164356	01/23/92	MC	35000
9696195	11/15/93	VISA	20000
2029997	03/09/09	MC	15000
4206913	02/15/87	VISA	30000
7611353	04/09/06	VISA	45000
6830786	07/28/82	VISA	50000
3483357	02/02/02	MC	45000
72127	06/12/98	VISA	27000
9243828	09/04/04	VISA	32000
8123180	10/03/03	VISA	17000
7611353	04/09/06	VISA	45000
6164356	01/23/92	MC	35000
7538391	05/06/97	MC	12000

1. Remove Duplicates

Run

Remove duplicates across all columns

Select columns to group by:

Acct_Num

Plan

Name: Plan 1

Modified: 02/29/20 08:54 PM

1. Remove Duplicates

Figure 9 Remove Duplicates Transformation

The screenshot displays the SAS Data Studio interface for a data preparation task. On the left, the 'Transforms' sidebar is open, showing various transformation options. The 'Remove duplicates' option is highlighted. The main workspace shows a data table named 'ACCT_INFO (session)'. The table has four columns: 'Acct_Num', 'Date_Opened', 'Card_Type', and 'Credit_Limit'. Below the table, the '1. Remove Duplicates' transformation is configured. The 'Remove duplicates across all columns' checkbox is checked, and 'Acct_Num' is selected as the column to group by. The 'Run' button is visible next to the configuration. On the right, the 'Plan' pane shows the transformation step in a sequence.

Acct_Num	Date_Opened	Card_Type	Credit_Limit
1175348	12/25/78	VISA	10000
2748056	03/08/83	VISA	15000
5399690	10/12/08	VISA	10000
9696195	11/15/93	VISA	20000
2029997	03/09/09	MC	15000
4206913	02/15/87	VISA	30000
6830786	07/28/82	VISA	50000
3483357	02/02/02	MC	45000
72127	06/12/98	VISA	27000
9243828	09/04/04	VISA	32000
8123180	10/03/03	VISA	17000
7611353	04/09/06	VISA	45000
6164356	01/23/92	MC	35000
7538391	05/06/97	MC	12000

Figure 10 ACCT_INFO After Remove Duplicates Transformation

Figure 11 shows BORROWER_INFO loaded into SAS Data Preparation. We still want to remove the duplicates just like we did for ACCT_INFO, but this time we are going to take a different approach. For ACCT_INFO we used the Remove Duplicates transformation from the Transformation pane. This time we are going to the automated transformation suggestions features.

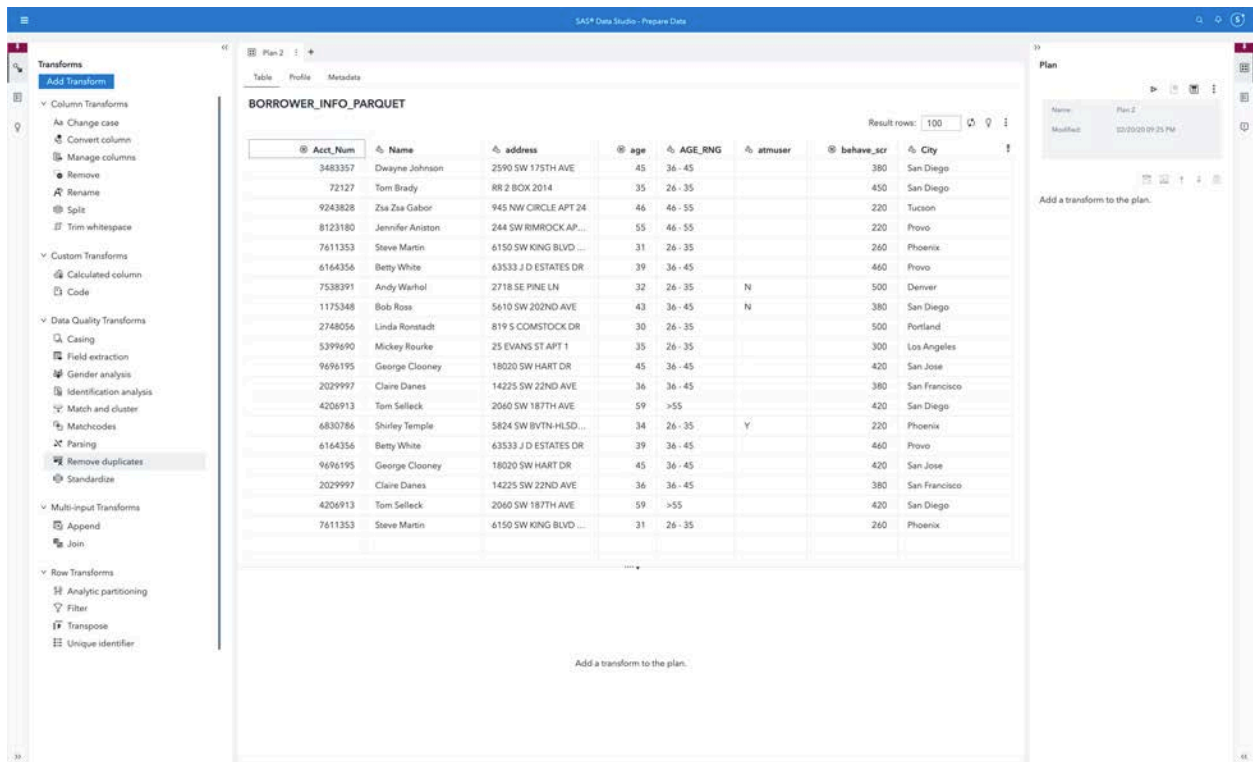


Figure 11 BORROWER_INFO Loaded Into SAS Data Preparation

Clicking on the light bulb icon on the left hand pane as seen in Figure 11 will take us to the Suggestions panel. SAS Data Preparation transformation suggestions deploys ten out of the box models against a dataset to determine which transformations, if any, would be appropriate to perform on the dataset. Clicking the double triangle in the Suggestions panel will bring up the complete list of all currently deployed models as shown in Figure 12. Clicking the Get Suggestions button will deploy all the models against all the columns in the dataset and return a set of suggested transformations. The results for our dataset are show in Figure 13.

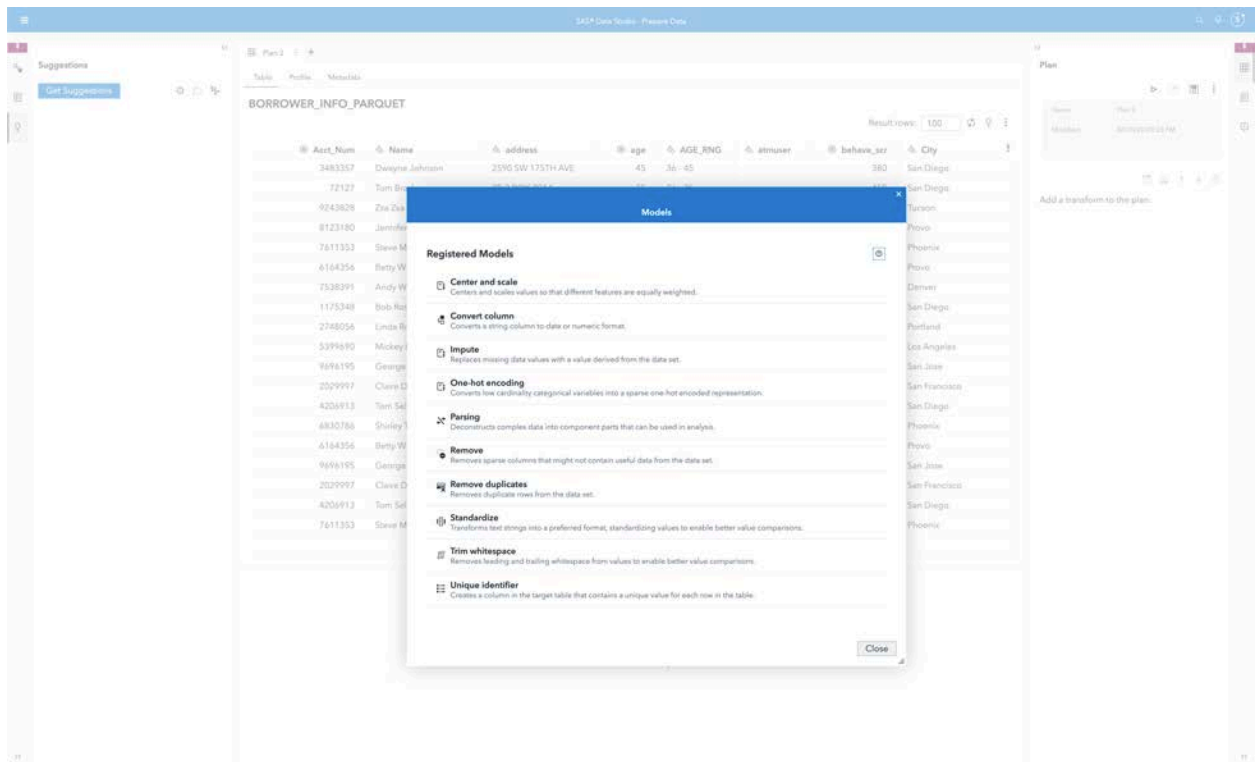


Figure 12 Deployed Transformation Models in SAS Data Preparation

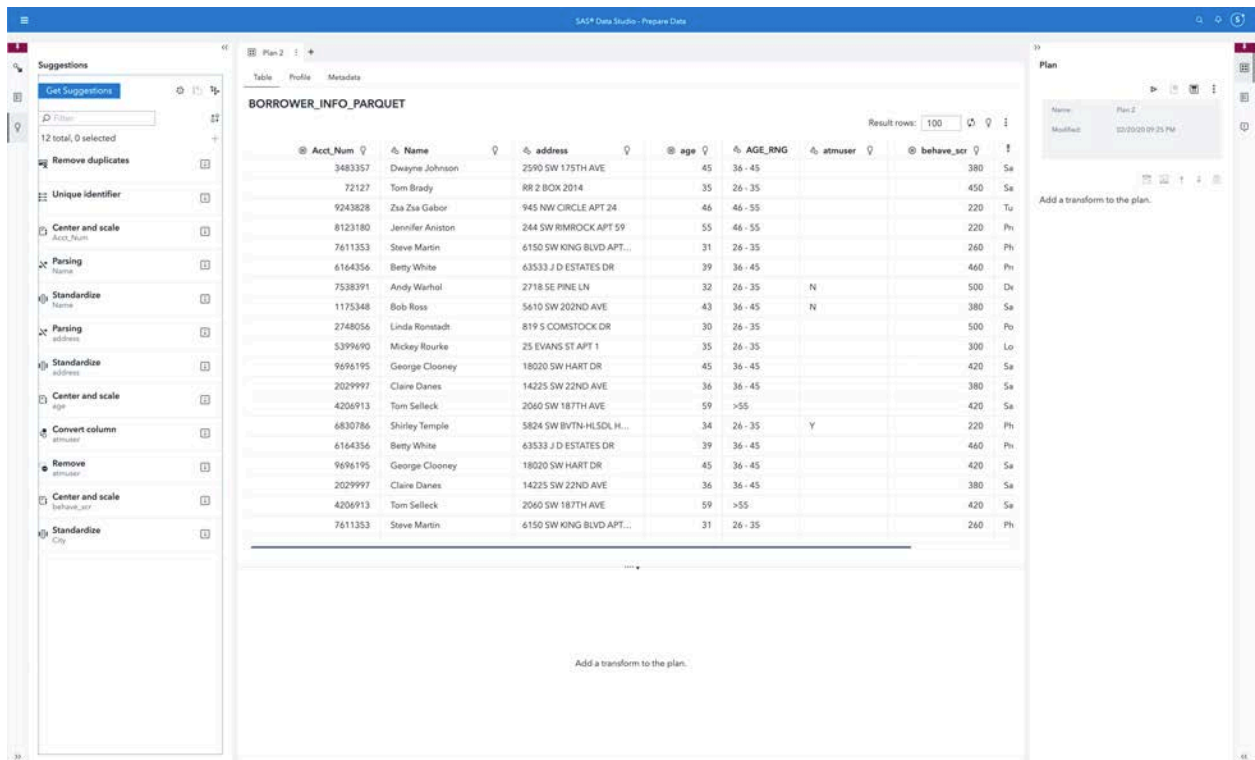


Figure 13 SAS Data Preparation Suggested Transformations

Even if we had not seen that there were duplicates in our unique identifier, Acct_Num, SAS Data Preparation transformation models detected that there were duplicates in the dataset and the very first suggestion is to remove duplicates. This is the same Remove Duplicates transformation and process that we saw with Acct_Info. Double clicking the suggestion opens the details of the transformation in the bottom pane beneath the dataset and adds the Removed Duplicates transformation as a step in the Plan panel on the right. Running the transformation removes the duplicates and leaves us with fourteen records, the same number we had in the ACCT_INFO table. The final result is show in Figure 14.

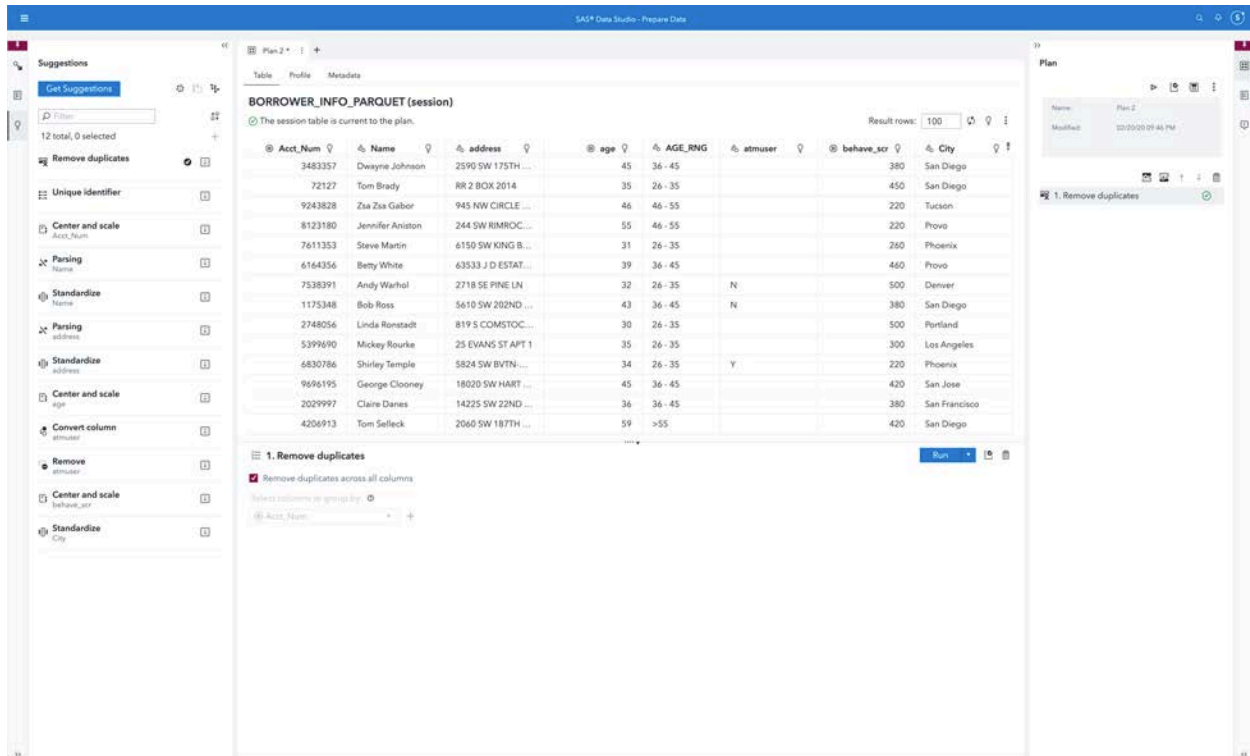


Figure 14 Borrower Info After Remove Duplicates Transformation Suggestion

So now we were able to leverage AI models to detect duplicates in our dataset and recommend a transformation to remove those duplicates. Again, we did this without writing a single line of code. SAS Viya has powerful data profiling tools accessible from when we start looking at the data in its source environment to when we load the data into the SAS Data Preparation to being manipulation. These tools are drag and drop giving even the most novice user comfort that they are being used correctly.

We've completed an initial review of our data and transformed it, so it's ready for the next big milestone... integration!

DATA INTEGRATION

Data integration can be particularly daunting to the uninitiated business analyst. The variety of joins and conditions and the fear of their incorrect implementation can be enough to scare away all but the most seasoned knowledge workers. The case could be made that this is one of the most critical skill set gaps between a data engineer and a data analyst. But fear not, SAS Data Preparation is here to help!

We are looking to understand the relationship between customer age and credit limit. We know we need to bring the data in ACCT_INFO and BORROWER_INFO together to do that. We know that we will do that using the Acct_Num field in both datasets. To do this in SAS Data Preparation, let's continue with our work in the BORROWER_INFO. Going back to the Transformations panel and clicking on Join under Multi-Input Transformations bring up the Join Tables dialog box as show in Figure 15.

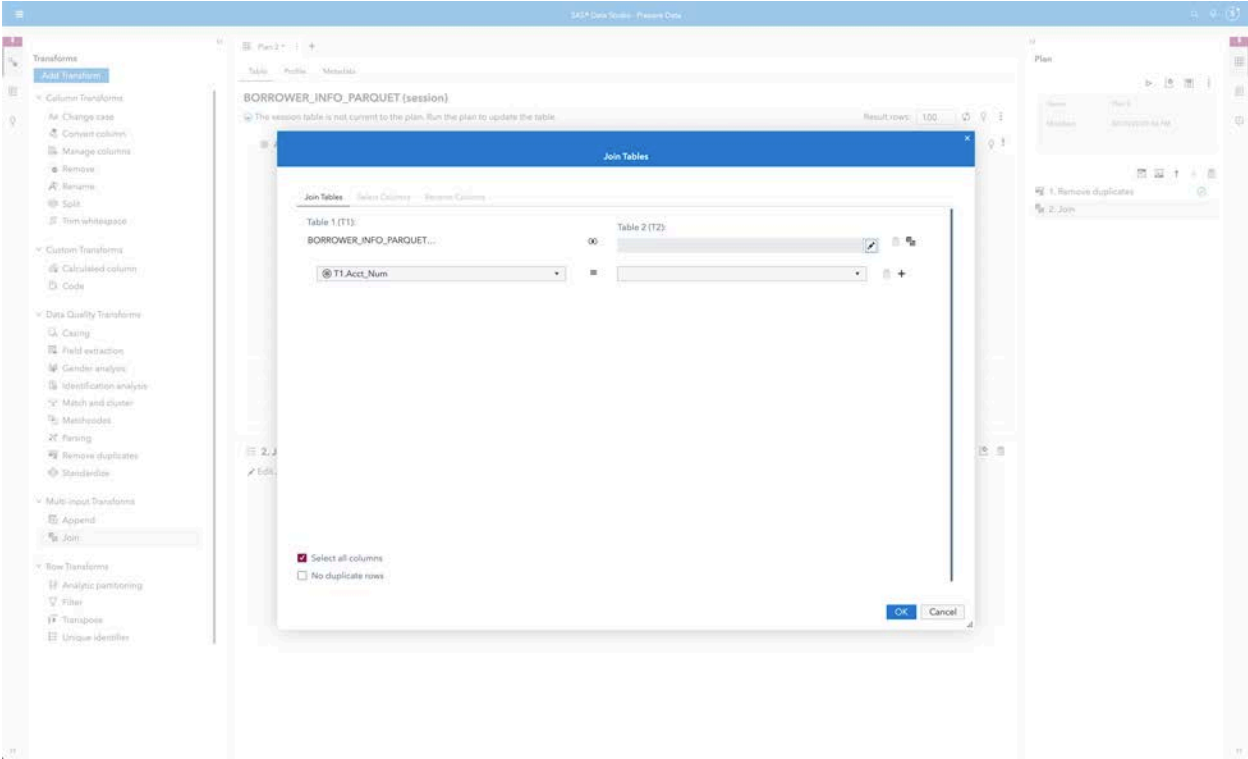


Figure 15 SAS Data Preparation Initial Join Tables Dialog Box

Clicking on the Venn Diagram in the middle gives us the option of Inner, Left, Right, and Full joins. We can click the edit icon to open the SAS Data Explorer interface to select our second table, which will be ACCT_INFO as shown in Figure 16.

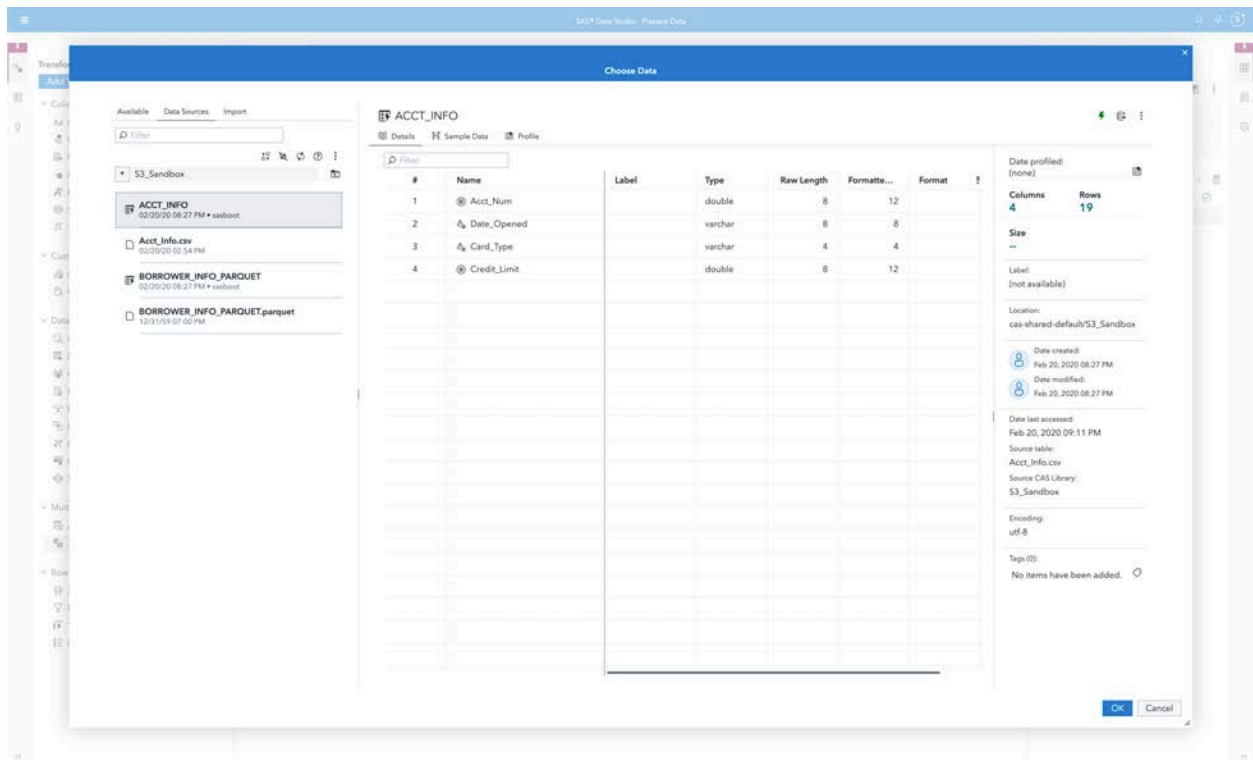


Figure 16 Table 2 SAS Data Explorer Selection Dialogue

Clicking ok automatically populates the table and join condition as shown in figure 17.

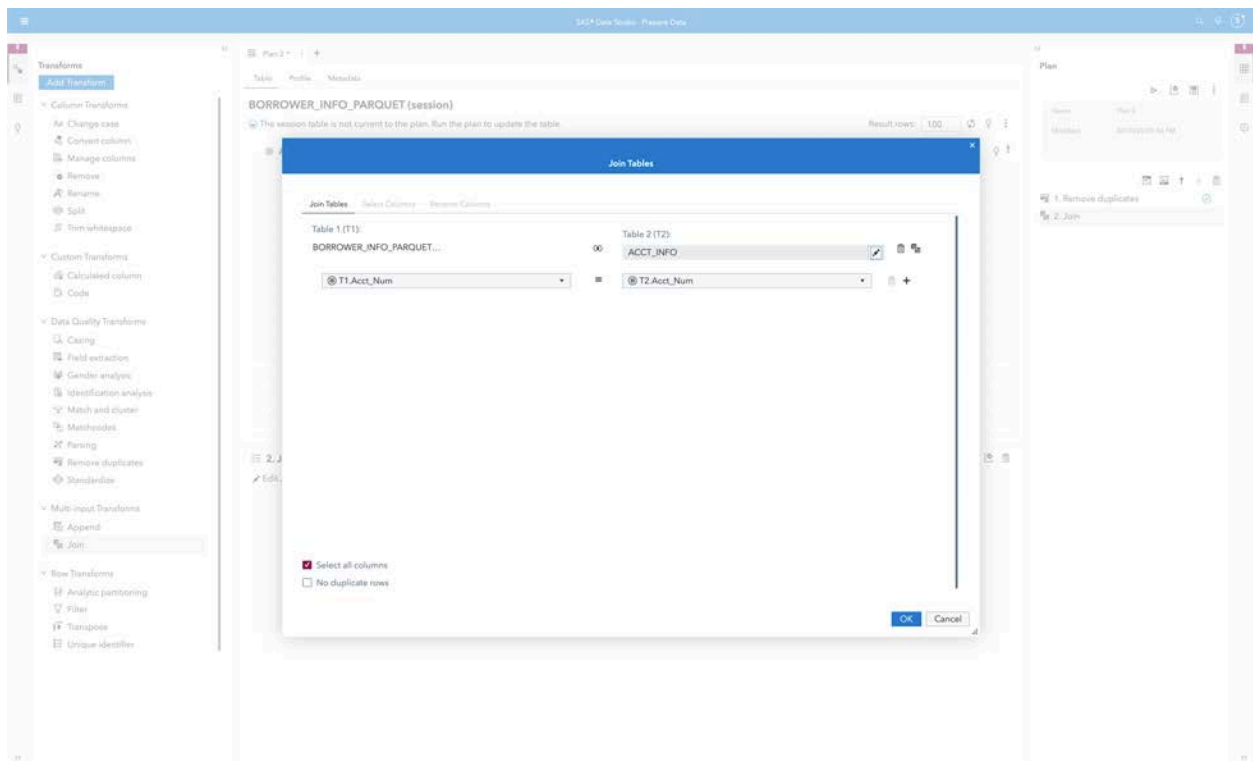


Figure 17 SAS Data Preparation Finalized Join Tables Dialog Box

Clicking OK closes the Join Table dialog box. Running the transformation joins the tables. The result is shown in Figure 18.

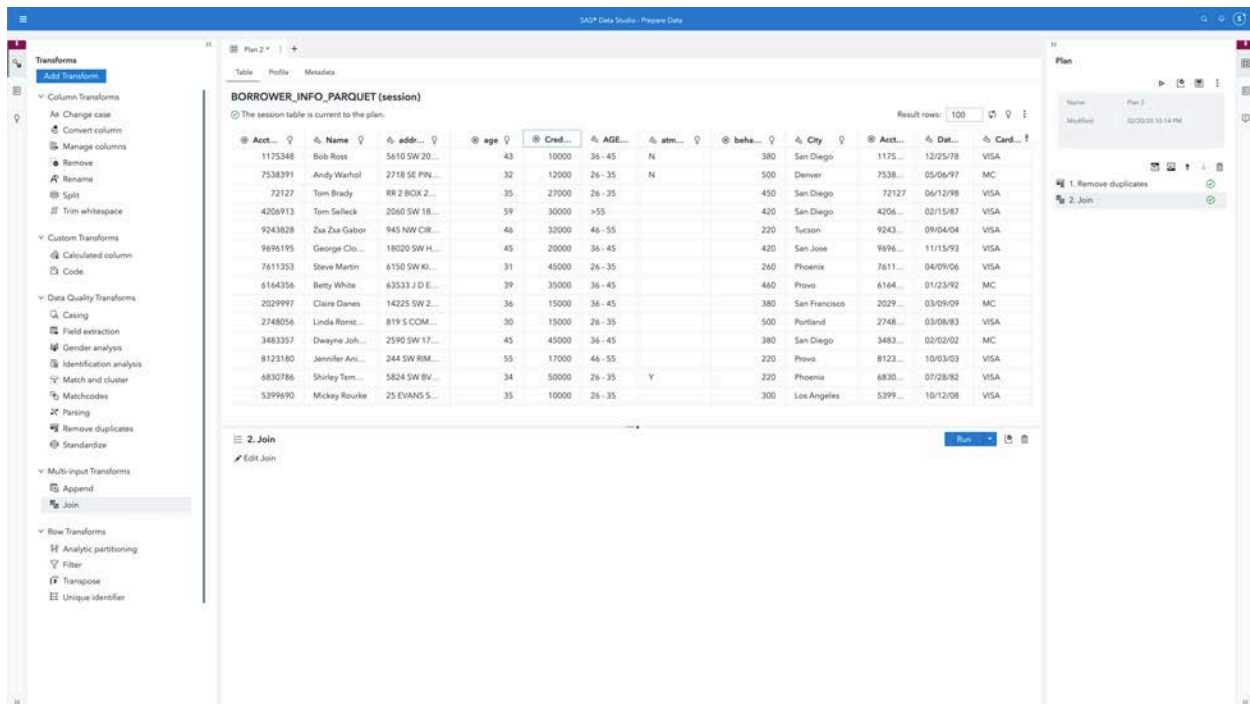


Figure 18 SAS Data Preparation Completed Join Tables Transformation

And voila! In less than a half dozen mouse clicks we have performed a perfect table join. All without writing a single line of code.

We can save our Data Plan to be rerun at a later date without having to recreate it. We can also save our resulting dataset as Analytic Base Table back to S3 in Parquet format as shown in Figure 19.

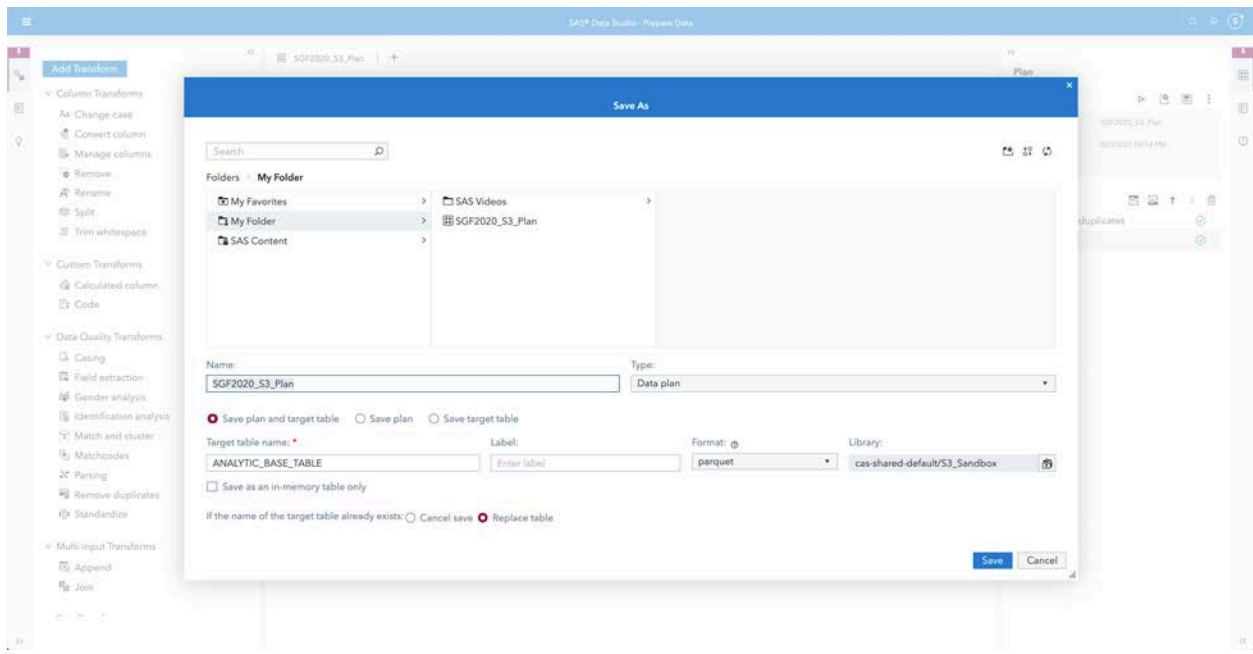


Figure 19 Saving The Analytic Base Table To S3

DATA GOVERNANCE

Data Governance is key to helping data analysts understand what data they are working with. SAS Data Preparation has embedded tools that enable Data Governance.

The first is SAS Data Lineage Viewer, which allows us to see and understand the relationships between objects in SAS Viya. Figure 20 shows SAS Data Lineage Viewer for our resulting dataset, the Analytic Base Table. Here we can clearly see the relationship between the Analytic Base Table, the Data Prep Plan, our S3 data lake, ACCT_INFO, and BORROWER_INFO. This visual representation makes it easier for a data analyst to understand the interdependencies between objects and evaluate upstream and downstream impact analysis on data flows.

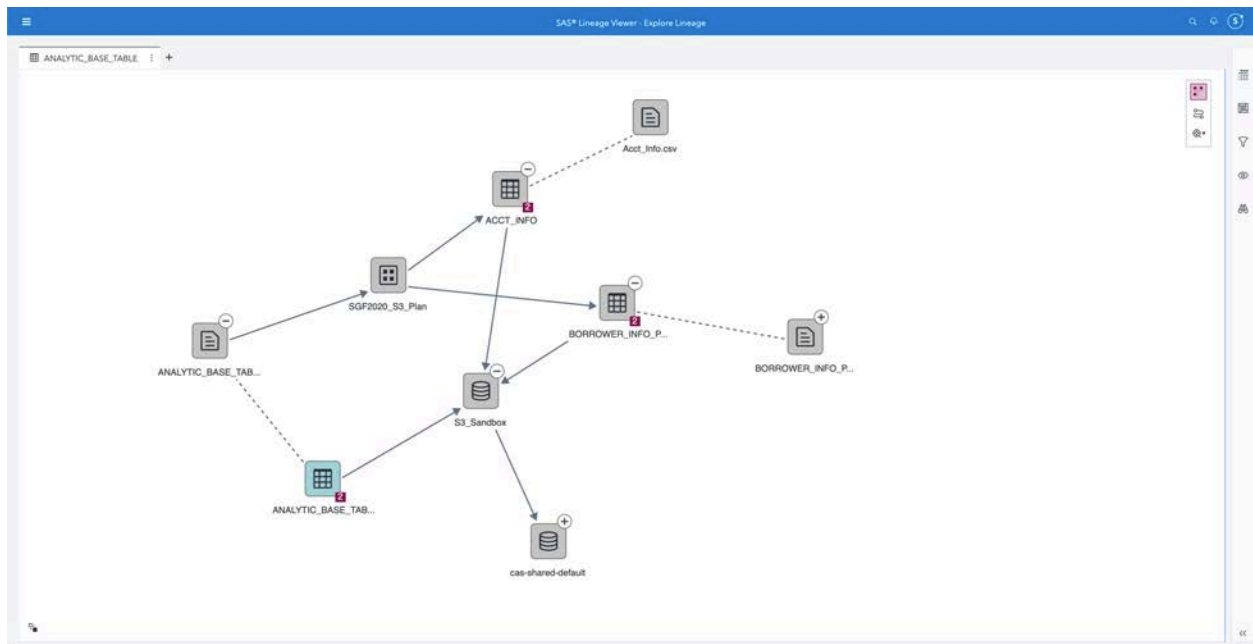


Figure 20 Analytic Base Table in SAS Lineage Viewer

The next tool SAS Data Preparation has for data governance is the ability to add labels to columns and tags to datasets. Column labels allow users to give additional context to specific data elements within a dataset. Dataset tags allow users to add keywords that give other users a better understanding of what that dataset should be used for.

And that brings us to the final need any self-service data preparation need when working with a data lake, enterprise collaboration.

ENTERPRISE COLLABORATION

One of the challenges inherent in a data lake environment is the creation of duplicate datasets by data analysts. This happens because analysts across the organization don't know what others are working on. Sometimes data analysts on the same team don't know what others are working on!

We saw the beginnings of how SAS Data Preparation helps prevent this situation with some of the governance tools. SAS Viya has additional tools to specifically address the need for enterprise collaboration. The first is SAS Drive, shown in Figure 21. All the work a data analyst does is stored here. Data Analysts can share their work, including specific Data Prep Plans, with other analysts and groups as shown in Figure 22.

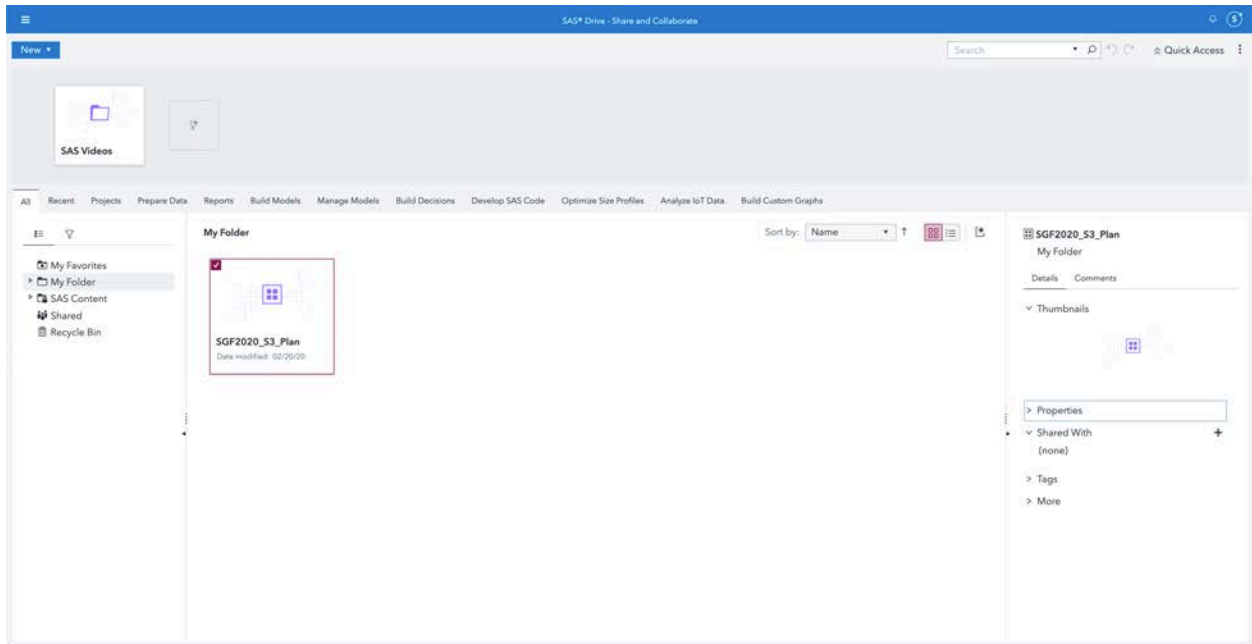


Figure 21 SAS Drive - Share and Collaborate

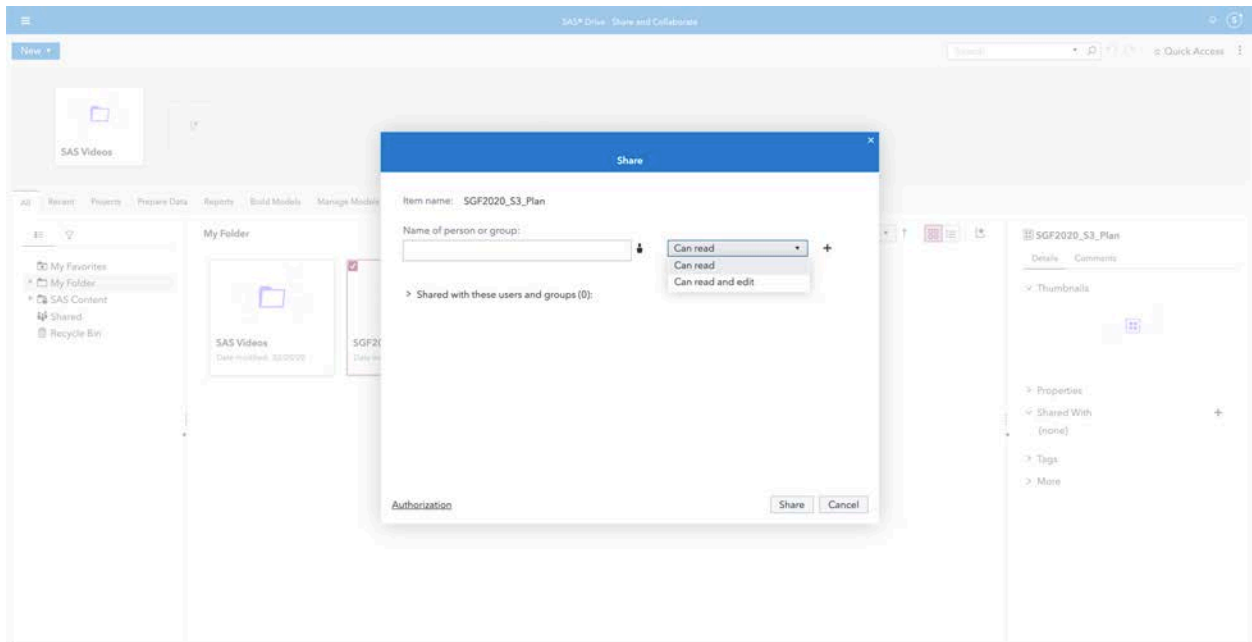


Figure 22 Share a Data Prep Plan With a User or Group

We can see how all the artifacts a data analyst could create through their workflows would be stored on SAS Drive. From the initial Data Prep Plan all the way through to the model of report. And all of it is shareable throughout the enterprise with column labels and dataset tags.

CONCLUSION

We've seen how the promise of data lakes was their ultimate undoing. Storing data in various formats and various stages of transformation in a single location to facilitate access fell short of actually accelerating analytics and reporting. This is due to the difference in skill sets between a data engineer and a data analyst.

The data engineer is very technical and would require considerable lead time to create an analytic base table that is ready for consumption by the data analyst. The data engineer would require a significant number of requirements up front because of their lack of subject matter expertise around the data.

The data analyst on the other hand will have deep subject matter expertise around the data but likely less technical expertise. They would know whether the data is fit for use to answer the specific question they trying to address. But they wouldn't necessarily have the technical wherewithal to perform the manipulations necessary for the remediation. Especially if the remediation required going back to source systems and integrating files in various format.

And so data lakes languished. Data analysts didn't have a tool that enabled them to access data in various formats, profile and understand the source data, perform necessary transformations and integrations, and share their work. Data kept getting dropped in the lake, but fewer and fewer users were able to engage with it in a meaningful way.

Until now! We have seen how SAS Data Preparation addresses all of these challenges. Dedicated SAS/ACCESS interfaces make getting to your data in any repository a breeze. SAS Data Explorer is able to read and profile data in a wide variety of formats. SAS Data Preparation bridges the deeper skill gaps between a data engineer and a data analyst with drag and drop transformations and AI driven transformation suggestions. SAS Lineage Viewer, column labels, and dataset tags provide fundamental data governance capabilities. And finally SAS Drive enables enterprise collaboration.

SAS Data Preparation makes working with your data lake a breeze!

REFERENCES

Gartner Newsroom. Accessed February 15, 2020. <https://www.gartner.com/en/newsroom/press-releases/2015-09-15-gartner-says-business-intelligence-and-analytics-leaders-must-focus-on-mindsets-and-culture-to-kick-start-advanced-analytics>

Global Big Data Conference. Accessed February 15, 2020. <http://www.globalbigdataconference.com/news/141927/85-of-big-data-projects-fail-but-your-developers-can-help-yours-succeed.html>

SAS/ACCESS Support Page. Accessed February 15, 2020. <https://support.sas.com/en/software/sas-access.html#1740662f-92d7-419d-9502-2855e7055a8d>

Amazon Web Services S3 FAQs. Accessed February 15, 2020. <https://aws.amazon.com/s3/faqs/>

RECOMMENDED READING

- *SAS Data Explorer 2.5: User's Guide – Making Data Available to CAS*
<https://go.documentation.sas.com/?docsetId=datahub&docsetTarget=n08ysqrqe3rexzn1c5gwo61noixg.htm&docsetVersion=2.5&locale=en#n1nbcyfxc28i1n1ae1nan7f20m1>

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Peter Baquero
SAS Institute Inc.
Peter.Baquero@sas.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.