Manufacturing Analytics With Weather

Thursday, May 14, 2020

Manufacturers need to quickly identify the reasons why there are high amounts of scrap rate to save money and deliver quality product. This demonstration shows how weather was the key driver leading to reducing scrap rate using statistical analysis and dashboards using Cloud Pak For Data (CP4D).

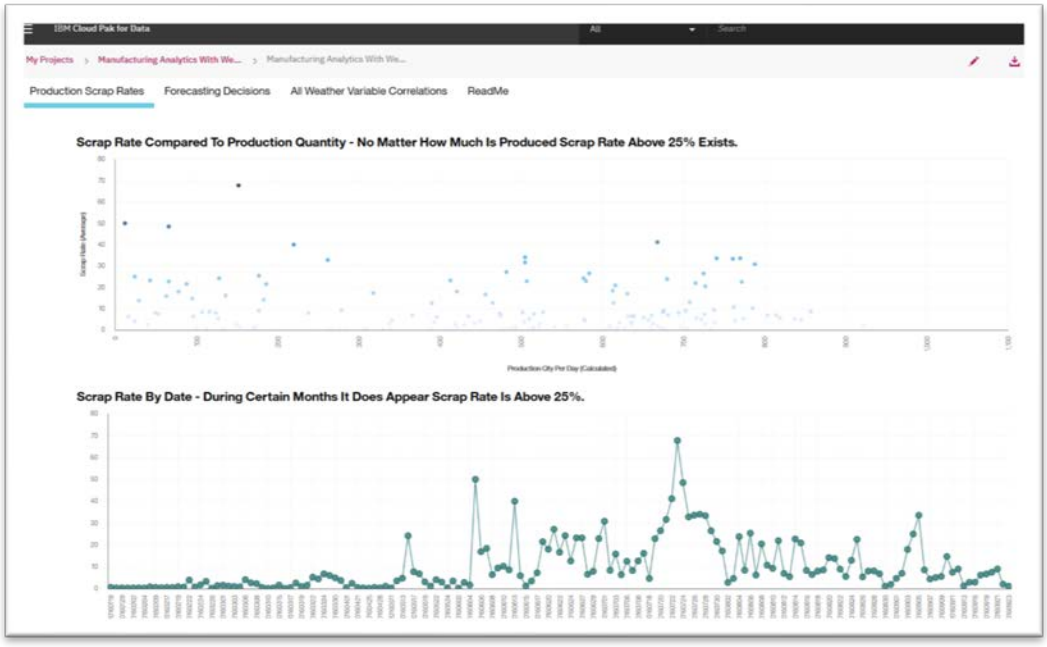
What's Included?

- 6 csv files
  - ***Prod_Data.csv*** – mocked up sample historical production data
  - ***WeatherHistory.csv*** – mocked up sample historical weather data
  - ***Scrap_Data.csv*** – mocked up sample historical scrap rate
  - ***All Weather and Manufacturing Data.csv*** – mocked up historical weather and production data to determine what the weather was at a particular point in time when scrap occurred.
  - ***Manufacturing Weather Forecasts.csv*** – mocked up weather forecast data
  - ***Scrap Rate By Time.csv*** – mocked up historical scrap rate over time
- SPSS Modeler Stream 18.1.1
  - ***Scrap_Analysis_Daily_Official.str***
- 1 Embedded Dashboard
  - ***Manufacturing_Analytics_With_Weather_Dashboard.json***
- Recording on IBM Community page. **WATCH THIS FIRST!!!**

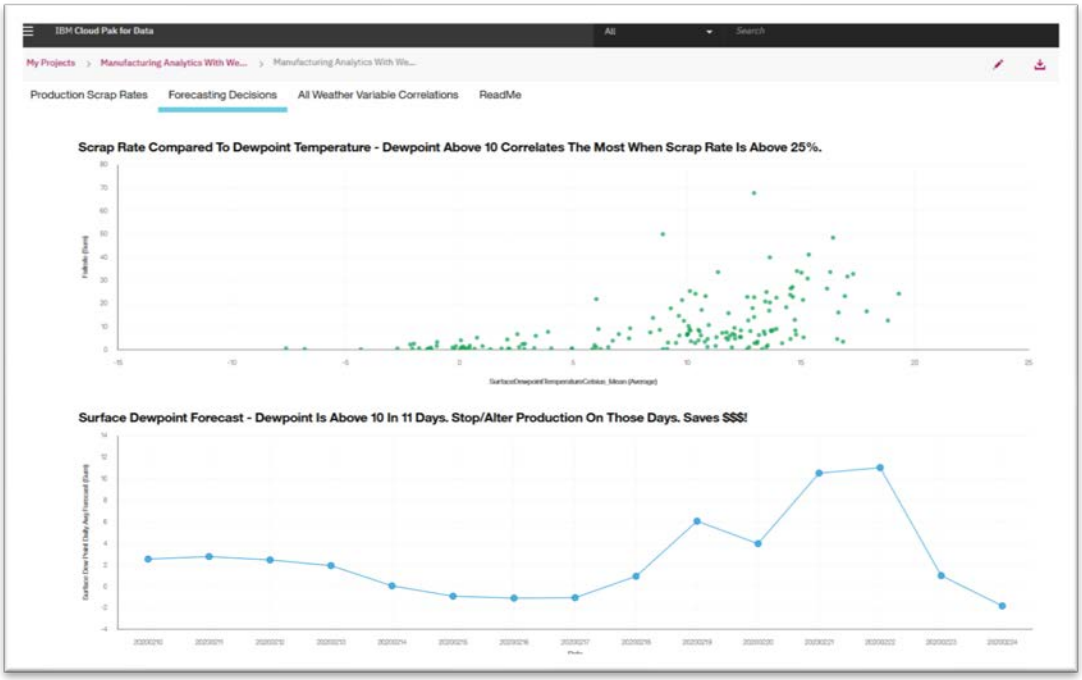
How does it work?

The dashboard shows that unless one considers the use of weather and statistical analysis, it is difficult to determine the reasons why large amounts of scrap rate exist.

The first tab shows that scrap is not a matter of how much you produce, nor is it really about time of year although it does appear that the summer and spring months are times of the year where higher scrap rate exists but we still don't know why.

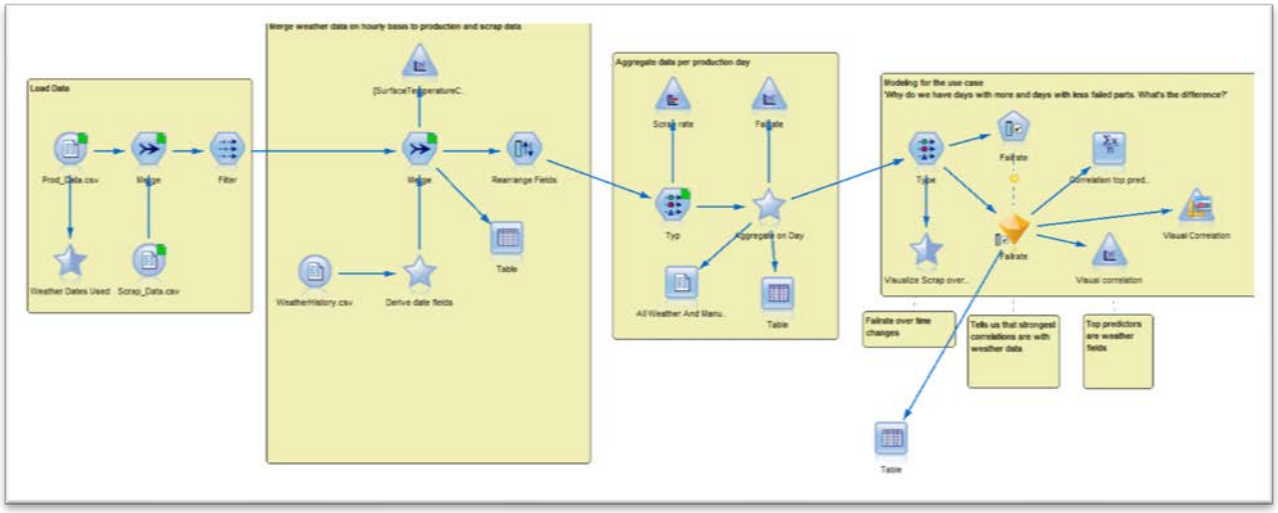


The second tab shows the high correlation of dewpoint to scrap rate and the third tab some other correlation results. Back on the second tab, we see that the forecast calls for dewpoint above 10, therefore, the schedule should be changed to make another product that day or re-allocate resources where more production can be done.

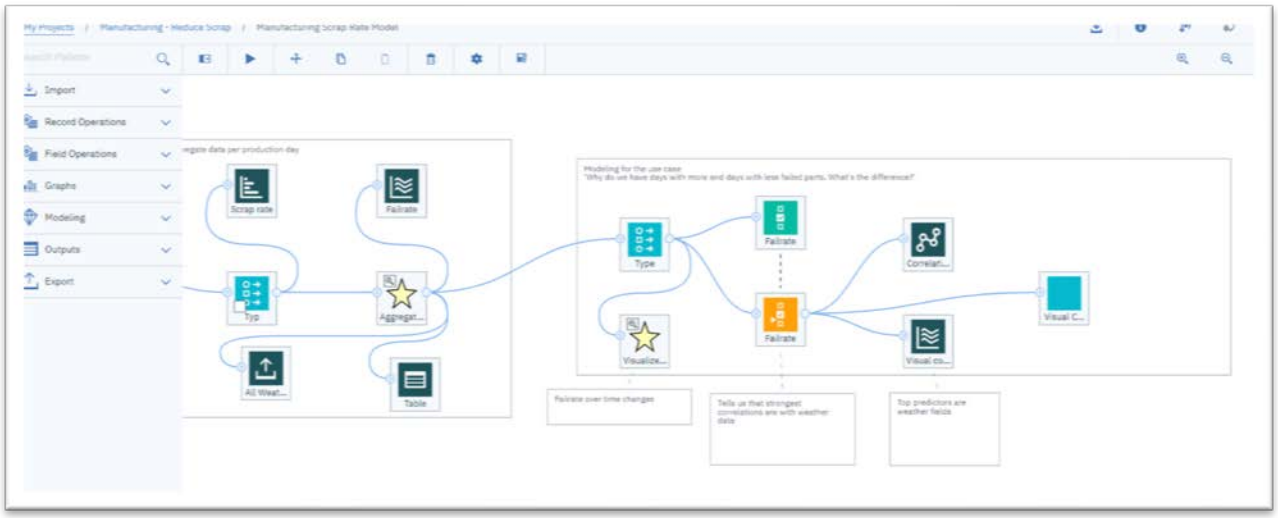


If you have SPSS Modeler Desktop 18.1.x or greater, then, you'll be able to see how the data was combined and correlation analysis was built. If you purchase **Premium Cloud Pak For Data Modeler Flows** then you'll be able to build your SPSS Modeler streams in CP4D or upload your streams from desktop to CP4D!

This was the initial SPSS Modeler stream developed using the desktop version (SPSS Modeler additional)



Here's the model above imported into our project as a Modeler Flow in CP4D. (Premium)



At the end of this README are steps to run the model in SPSS Modeler Desktop and as a CP4D Modeler Flow. The output of the model (both are the same) is **All Weather and Manufacturing Data.csv**, the merged data set that provides what the weather was for that day when scrap occurred. **Scrap Rate By Time.csv** and **Manufacturing Weather Forecast.csv** are used in the dashboard to show the existing problem with scrap rate , the upcoming 15 day forecast and how Dewpoint correlates the most to scrap rate. The remaining csv files are used in the model as inputs to **All Weather and Manufacturing Data.csv**. The SPSS Modeler/Modeler Flow also provides additional visualizations and correlation analysis that was used as a guide when the dashboard was built.

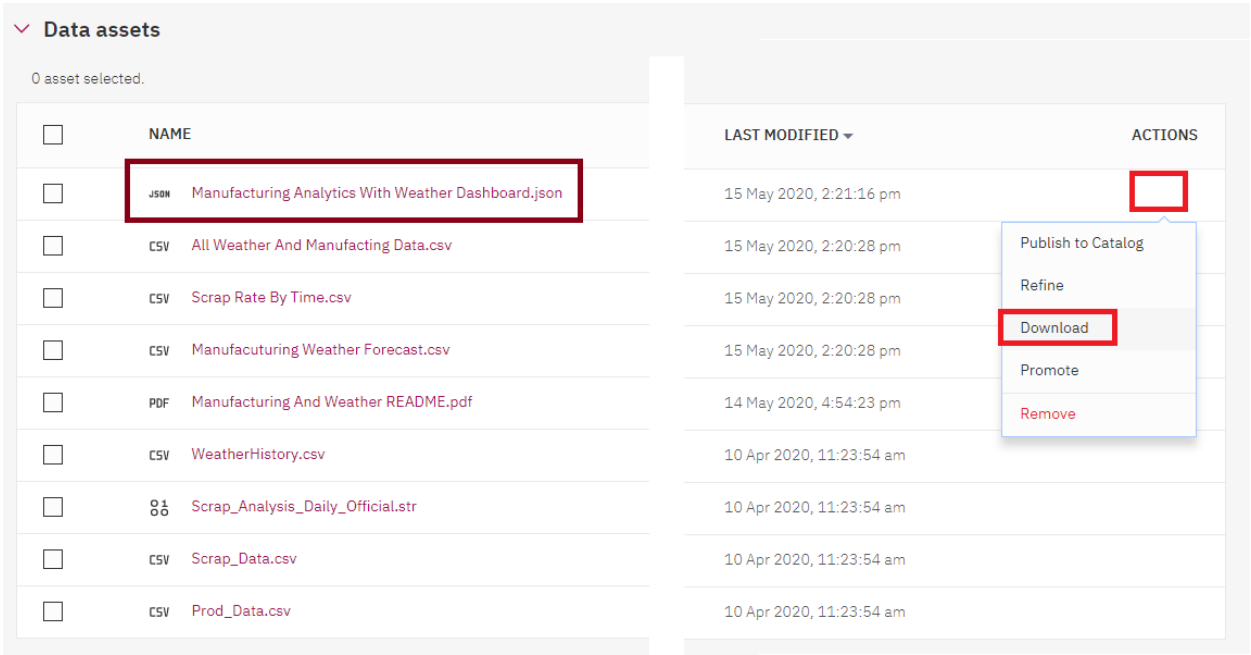
Prerequisites

Required services: To use the industry accelerators, you must install one or more of the following services on IBM® Cloud Pak for Data

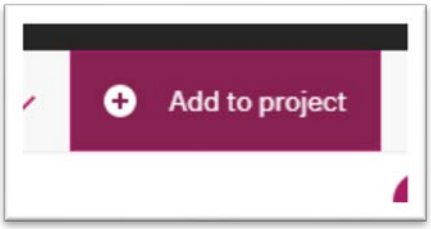
Service/Software	Required Info
Watson Studio	Importing data science assets to an analytics project. See Installing Watson Studio.
(Optional) SPSS Modeler 18.1.x	(Optional) If user wants to investigate the model (**Scrap_Analysis_Daily_Official.str**) they can use SPSS Modeler Desktop to do so. It's a code free (R and Python also) desktop data science tool that can be imported into CP4D as a Modeler Flow. Modeler Flows are not required to use SPSS Modeler Desktop with CP4D.
(Optional) Modeler Flows	(Optional) Used as a code free data mining and data science tool within CP4D. Users can build streams like those using SPSS Modeler Desktop as Modeler Flows in CP4D. SPSS Modeler Desktop is not required to use CP4D Modeler Flows. CP4D Modeler Flows can be built independent of SPSS Modeler Desktop. It's important to know that you can import SPSS Modeler streams as CP4D Modeler Flows into CP4D. Get a trial to SPSS Modeler here.

Importing the accelerator

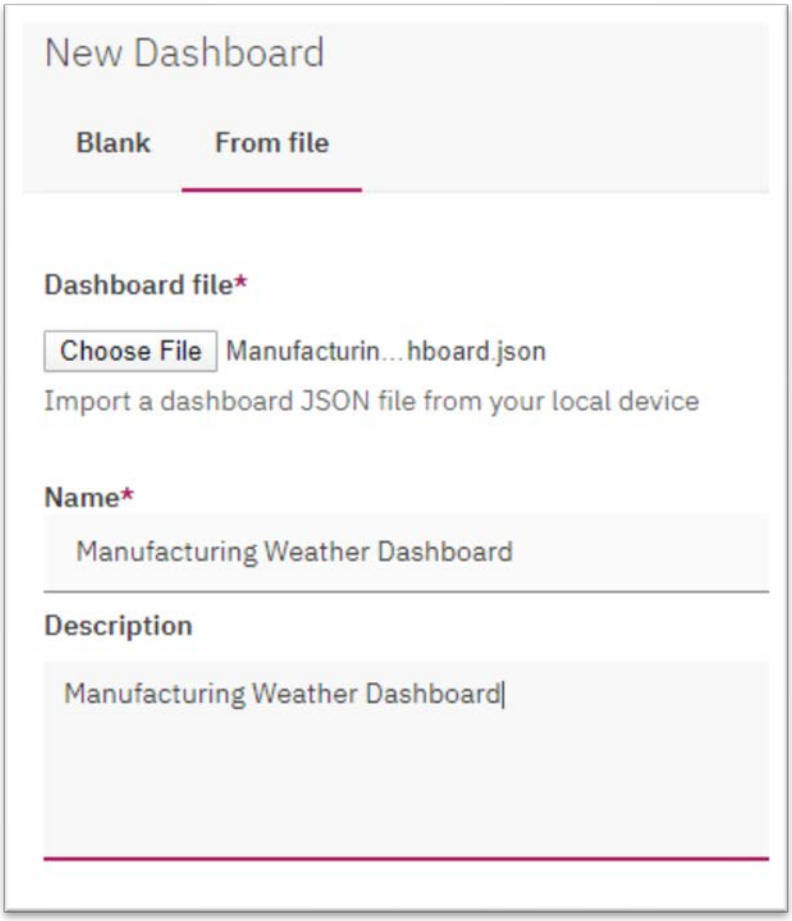
1) Go to your Data Assets and choose to Download ***Manufacturing_Analytics_With_Weather_Dashboard.json***



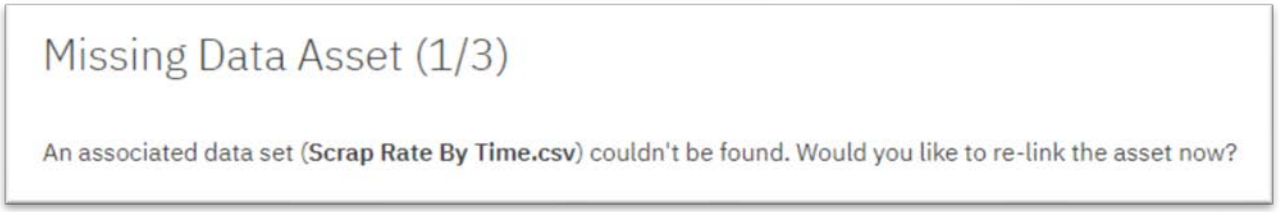
2) To Import the dashboard, In the upper right portion of your screen, click **Add to project**



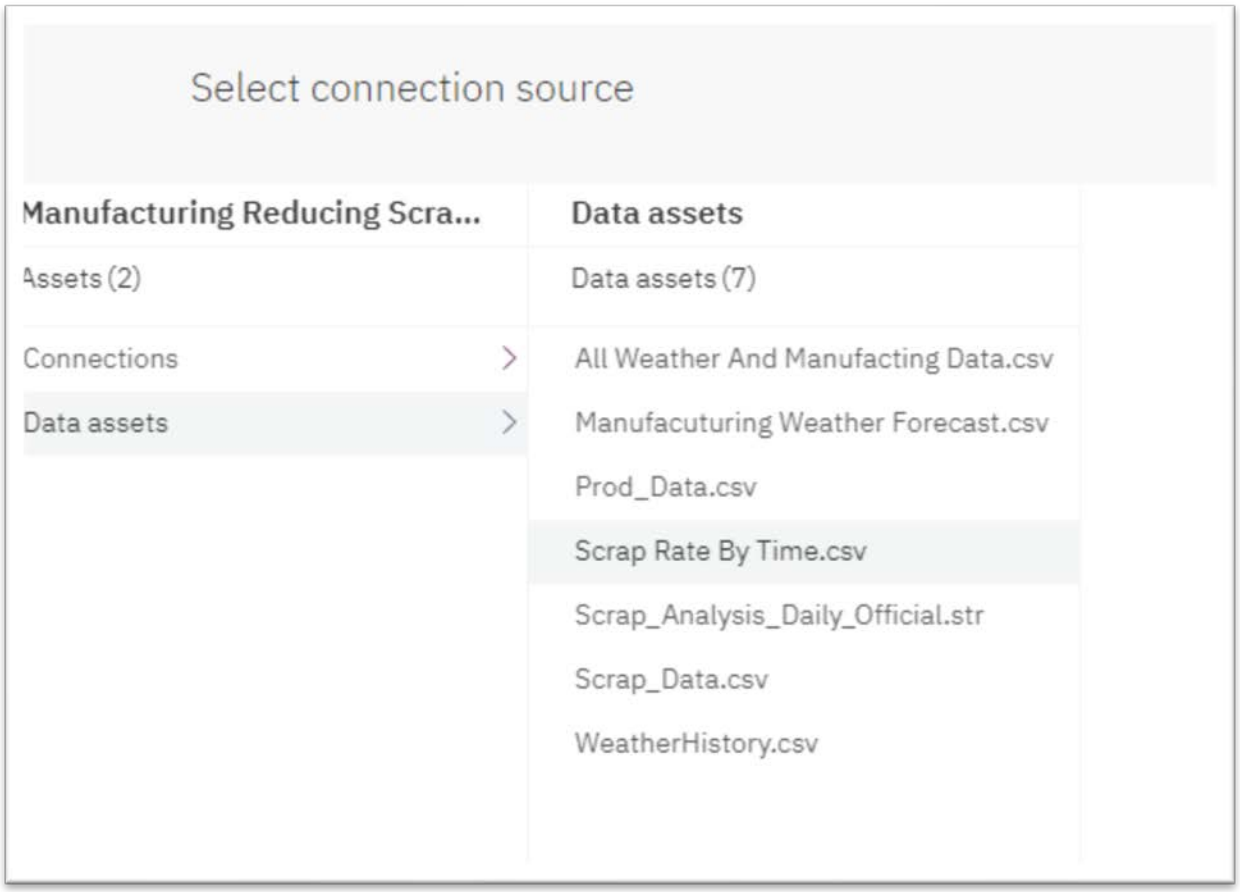
3) Select *Dashboard*, choose **From File**.
Choose the file you downloaded called **Manufacturing_Weather_Dashboard.json**
Enter the name **Manufacturing Weather Dashboard.**
Add the description **Manufacturing Weather Dashboard**.
Click **Create**



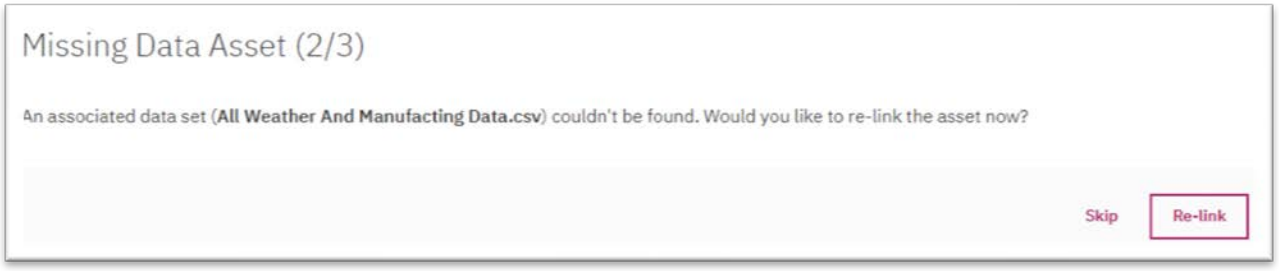
4) You'll likely be prompted to point to the correct files for **Scrap Rate By Time.csv**. Click Re-link.

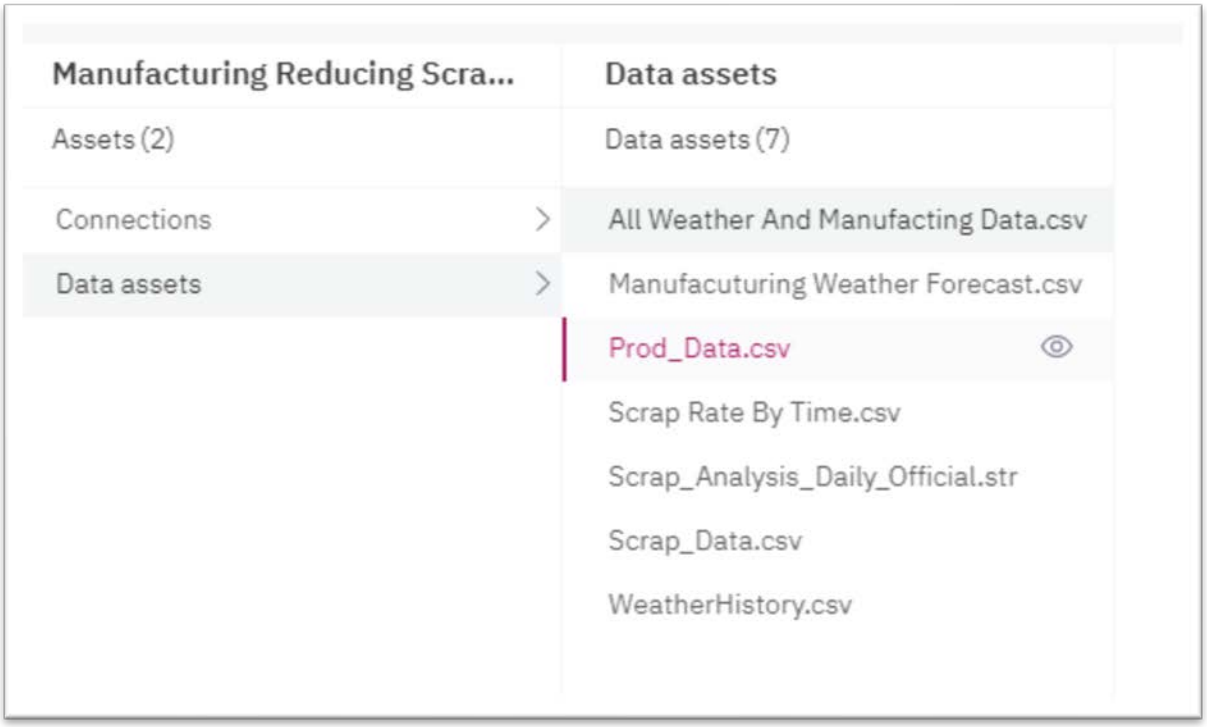


5) Choose **Scrap Rate By Time.csv.csv**

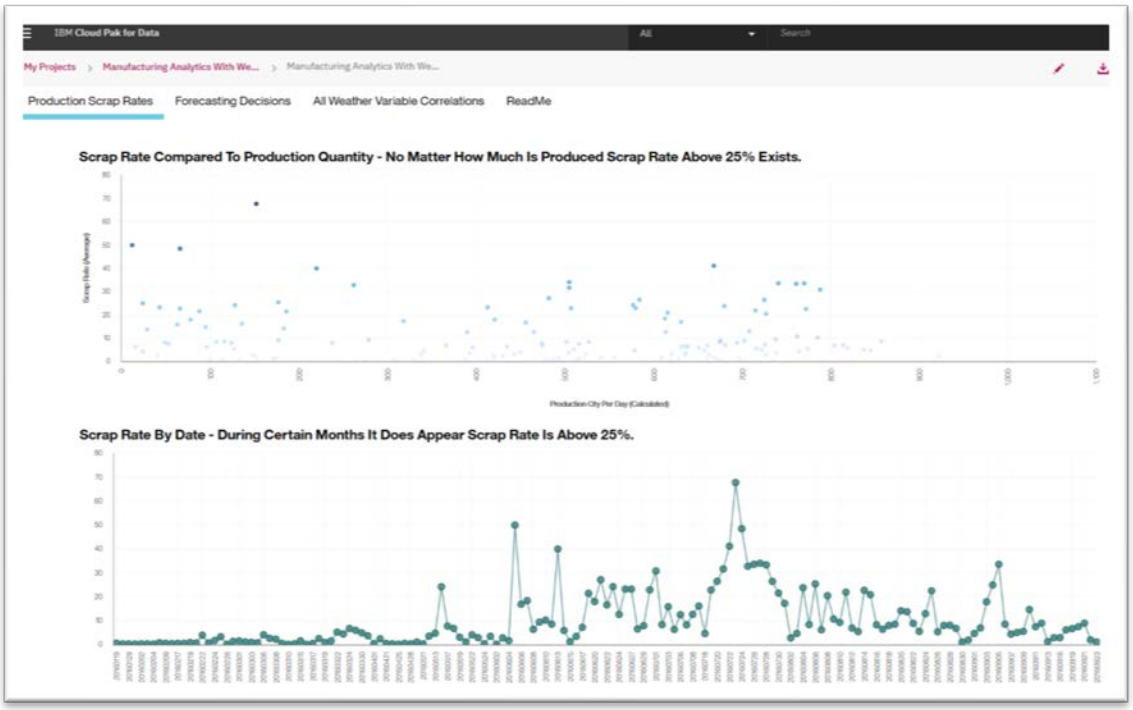


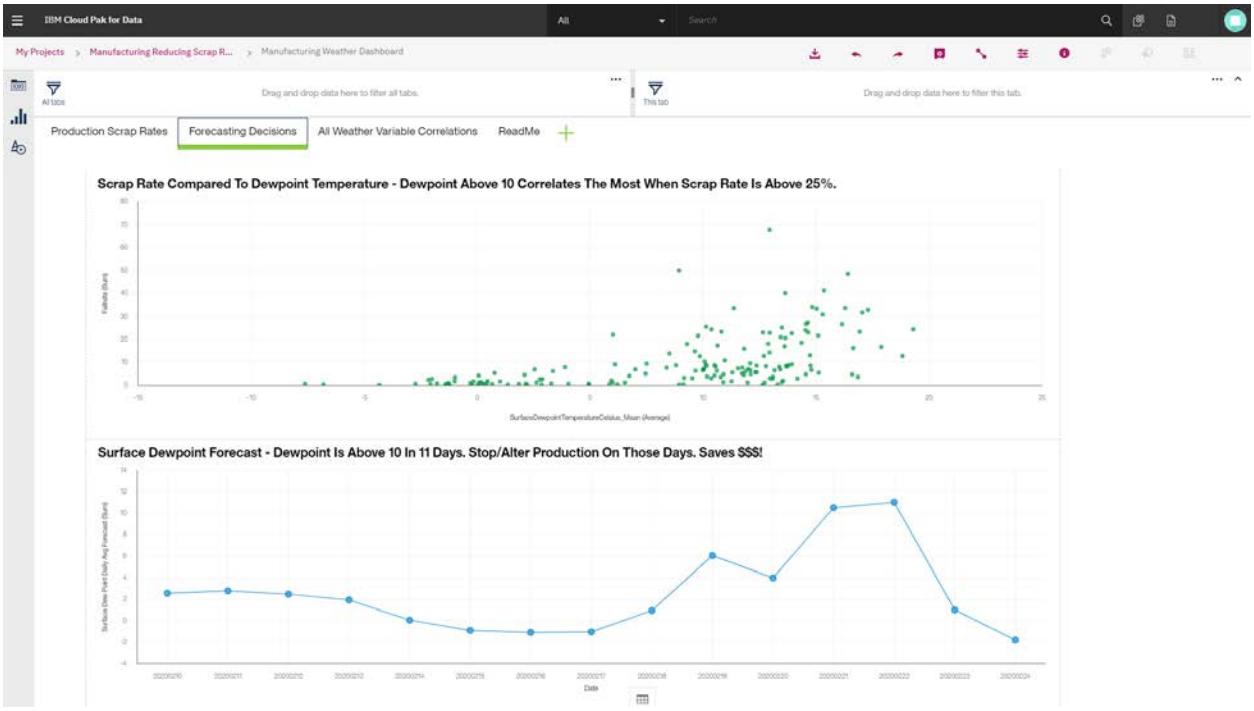
6) Next, you'll be prompted to point to the second required csv for the dashboard **All Weather And Manufacturing Data.csv** Click **Re-link**

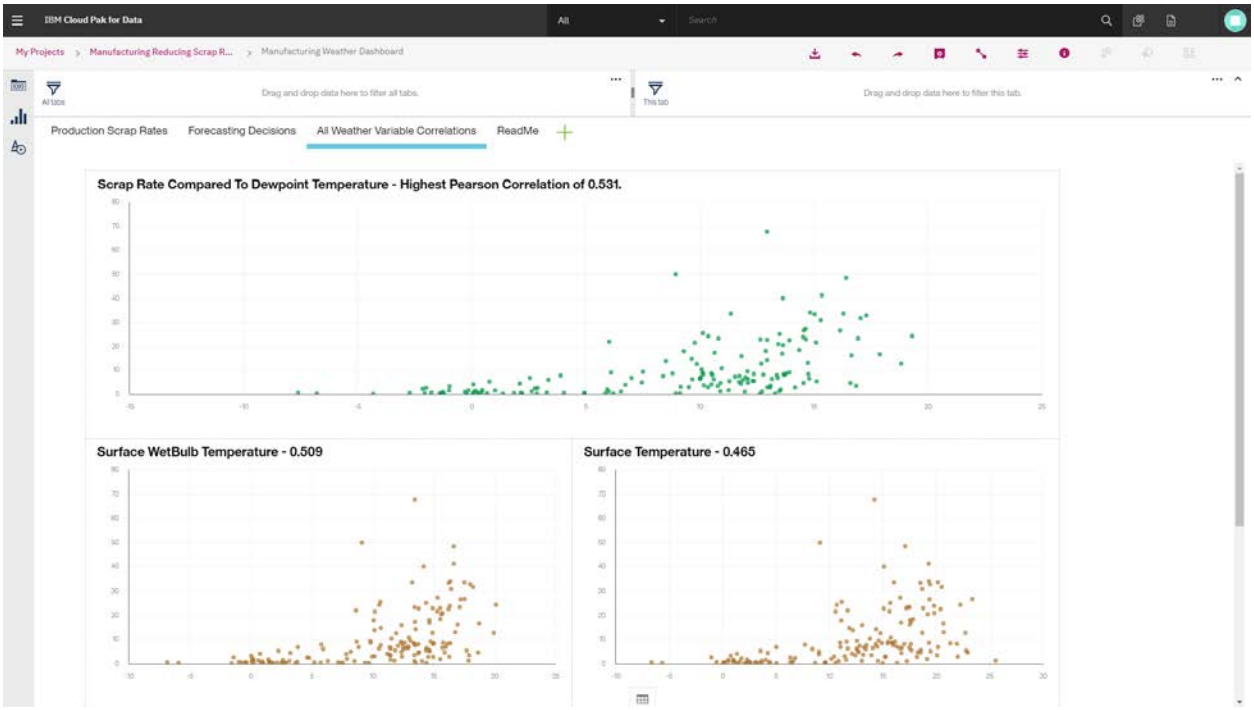


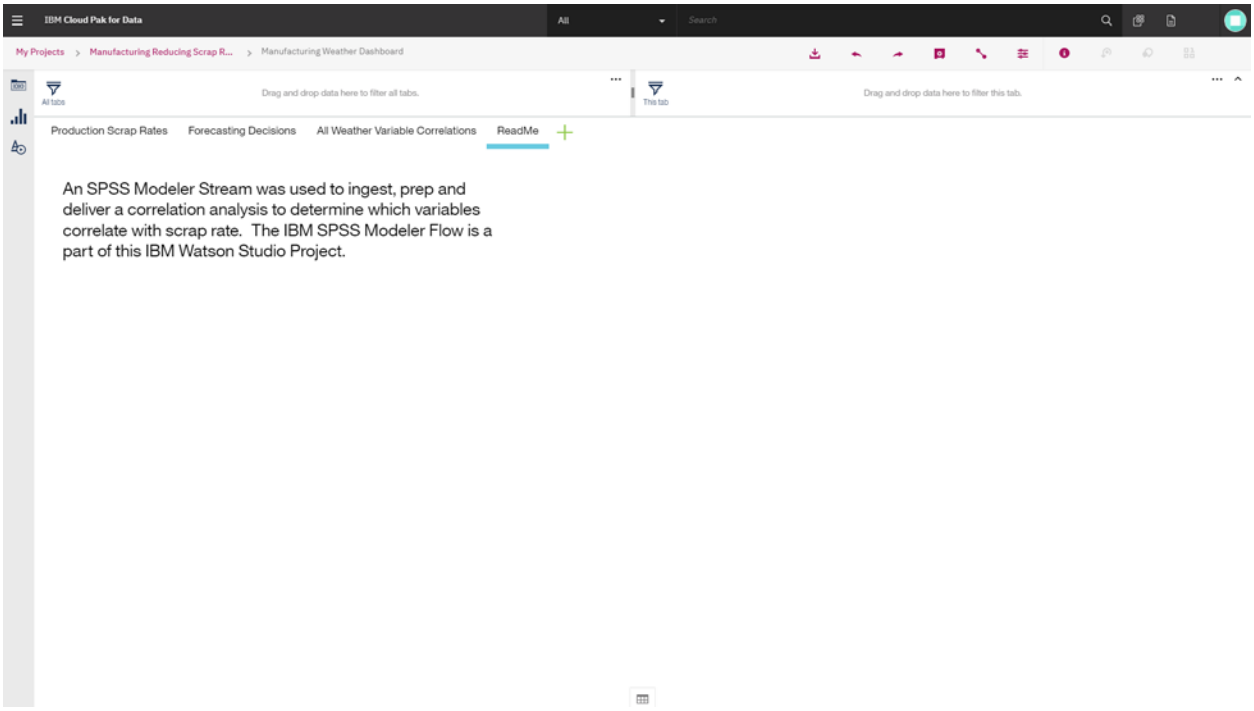


7) Re-link the last csv called **Manufacturing Weather Forecast.csv**
------No snapshots taken----- See above
8) The dashboard should open and look as such with each slide telling the story how Dewpoint was shown to statistically correlate the most with scrap rate when the dewpoint was above 1



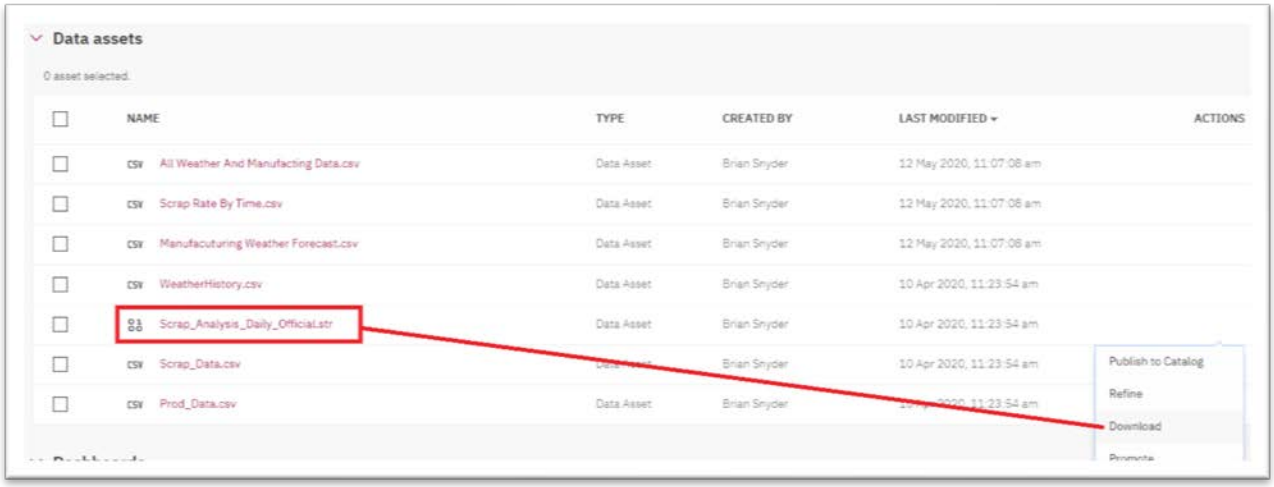

IBM Cloud Pak for Data
All
Search
My Projects
Manufacturing Reducing Scrap R...
Manufacturing Weather Dashboard
All tabs
Drag and drop data here to filter all tabs.
This tab
Drag and drop data here to filter this tab.
Production Scrap Rates
Forecasting Decisions
All Weather Variable Correlations
ReadMe
Scrap Rate Compared To Dewpoint Temperature - Dewpoint Above 10 Correlates The Most When Scrap Rate Is Above 25%.
Failure (Sum)
SurfaceDewpointTemperatureCelsius_Mean (Average)
Surface Dewpoint Forecast - Dewpoint Is Above 10 In 11 Days. Stop/Alter Production On Those Days. Saves $$$!
Surface Dew Point Daily Avg Forecast (Sum)
Date


IBM Cloud Pak for Data
All
Search
My Projects
Manufacturing Reducing Scrap R...
Manufacturing Weather Dashboard
All tabs
Drag and drop data here to filter all tabs.
This tab
Drag and drop data here to filter this tab.
Production Scrap Rates
Forecasting Decisions
All Weather Variable Correlations
ReadMe
Scrap Rate Compared To Dewpoint Temperature - Highest Pearson Correlation of 0.531.
Surface WetBulb Temperature - 0.509
Surface Temperature - 0.465

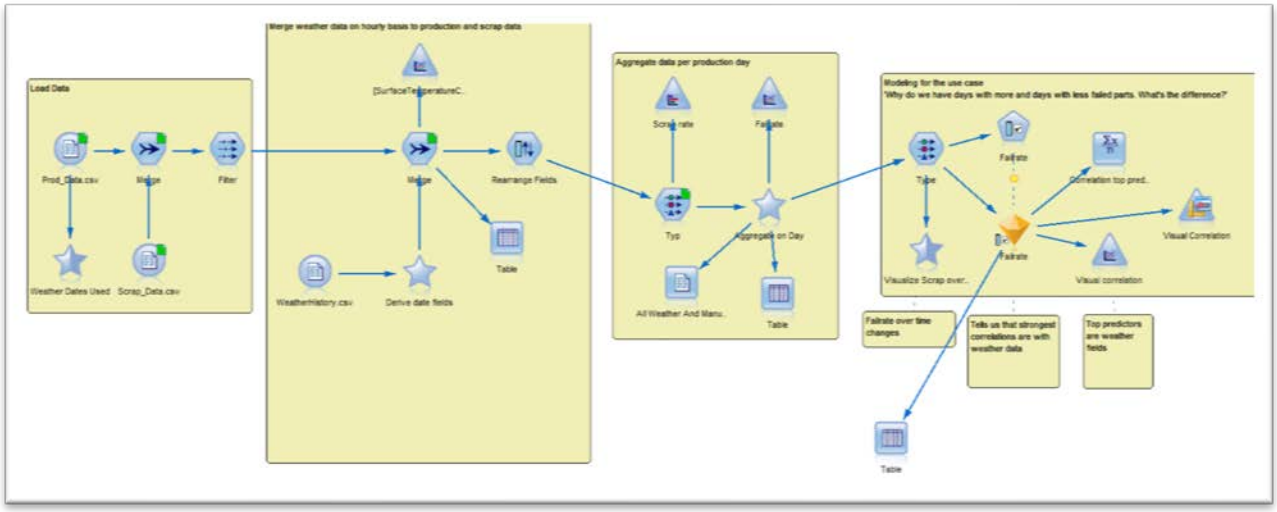


Working With SPSS Modeler 18.1.x

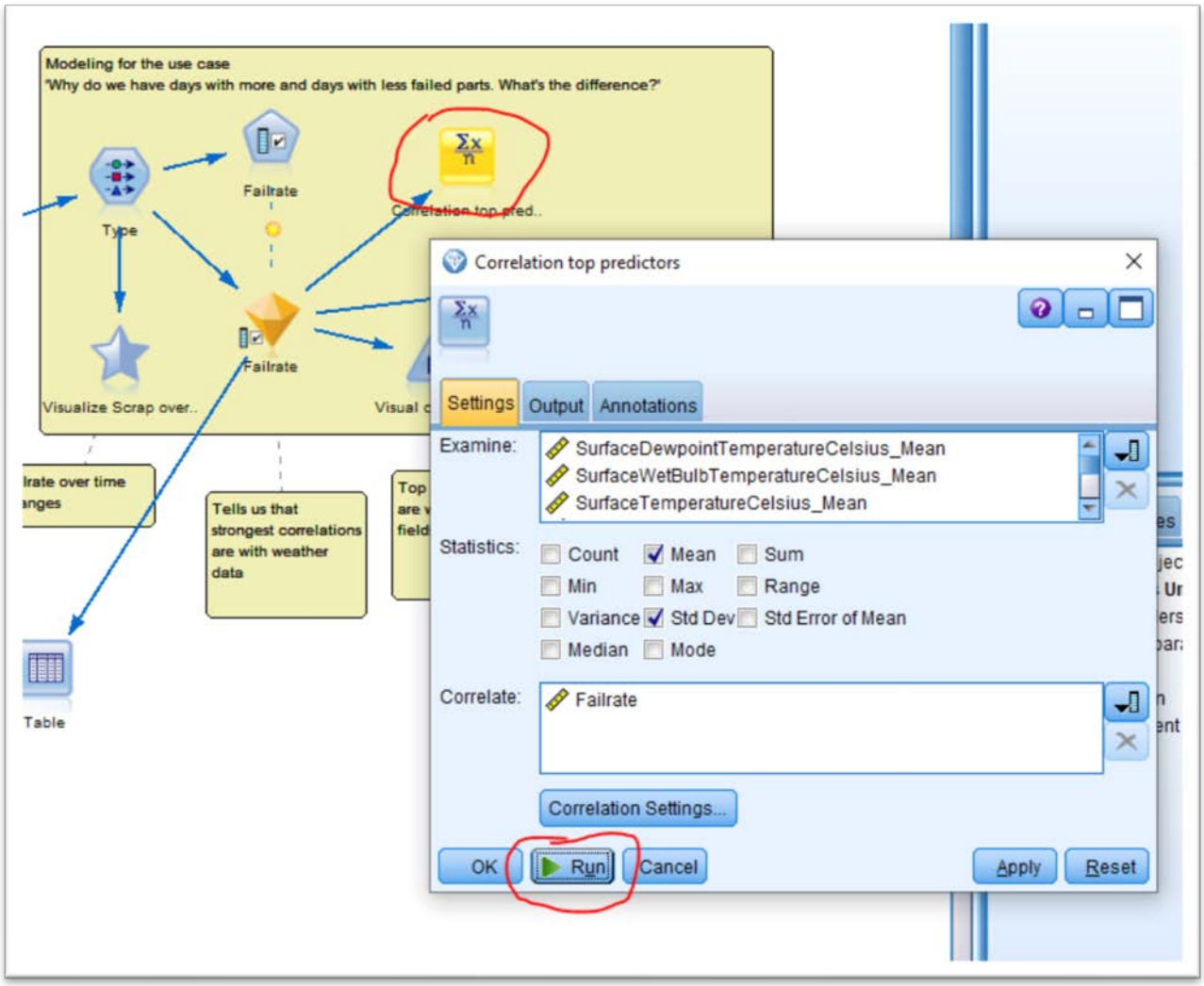
1) Download file **Scrap_Analysis_Daily_Official.str** from your Data Assets**.**



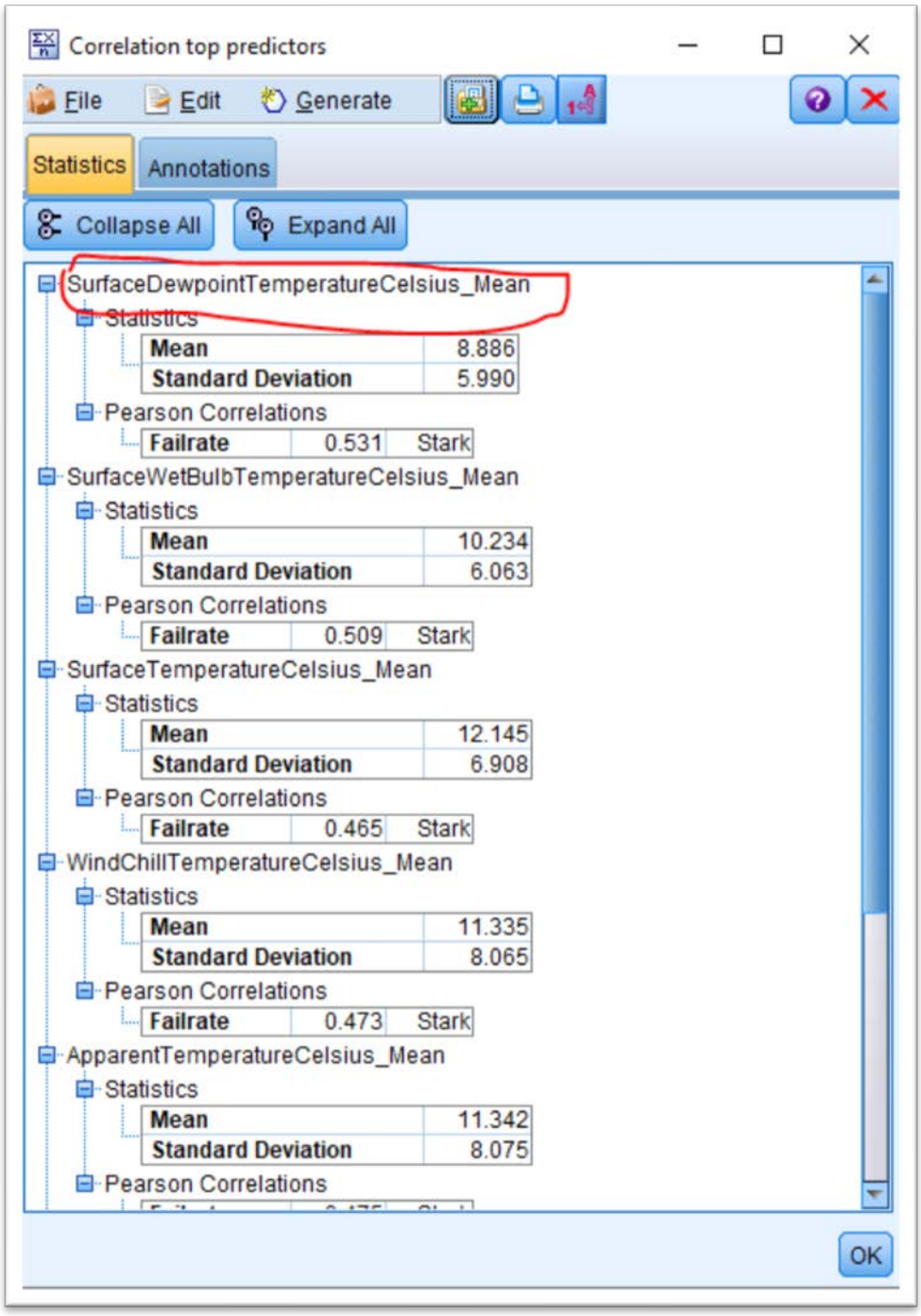
2) Open SPSS Modeler Desktop. Open the file **Scrap_Analysis_Daily_Official.str**


Load Data
Prod_Data.csv
Merge
Filter
Weather Dates Used
Scrap_Data.csv
Merge weather data on hourly basis to production and scrap data
[SurfaceTemperatureC..
Merge
Rearrange Fields
Table
WeatherHistory.csv
Derive date fields
Aggregate data per production day
Scrap rate
Fail rate
Typ
Aggregate on Day
All Weather And Manu..
Table
Modeling for the use case
'Why do we have days with more and days with less failed parts. What's the difference?'
Table
Fail rate
Correlation top pred..
Visual Correlation
Visualize Scrap over..
Fail rate
Visual correlation
Fail rate over time changes
Tells us that strongest correlations are with weather data
Top predictors are weather fields
Table

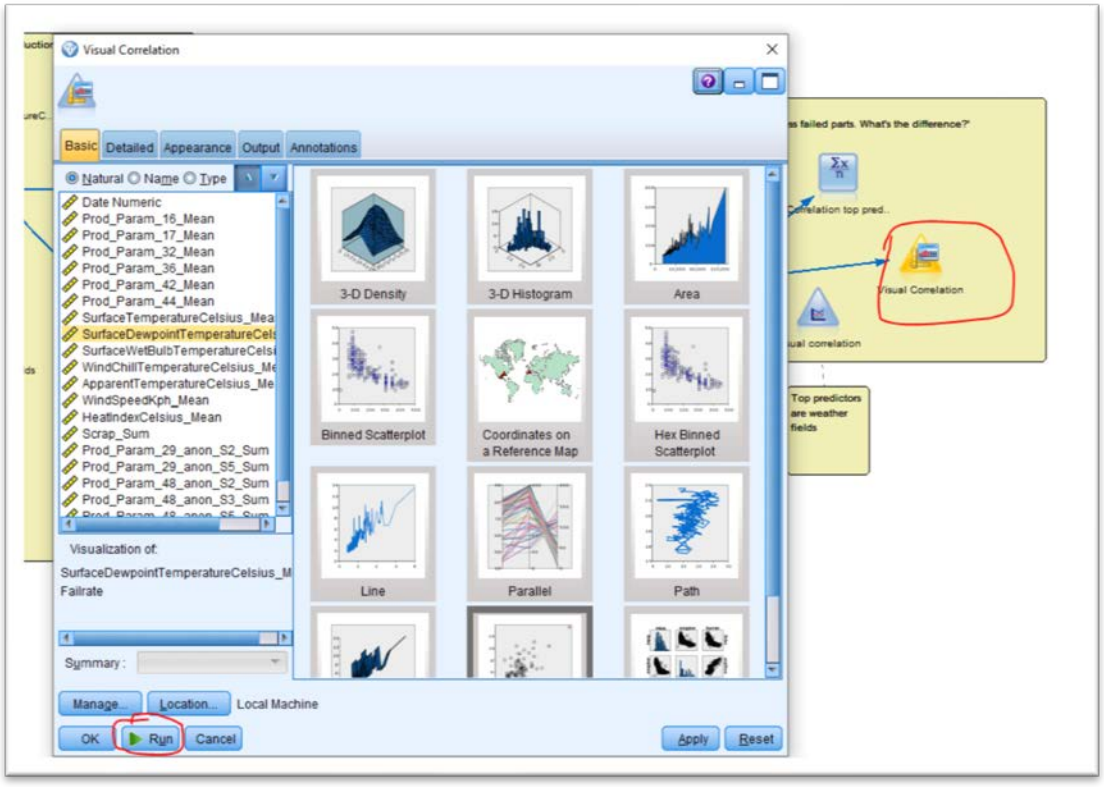
3) To view which weather variable correlates the most to failure rate, double click and Run the **Correlation top predictors** node.



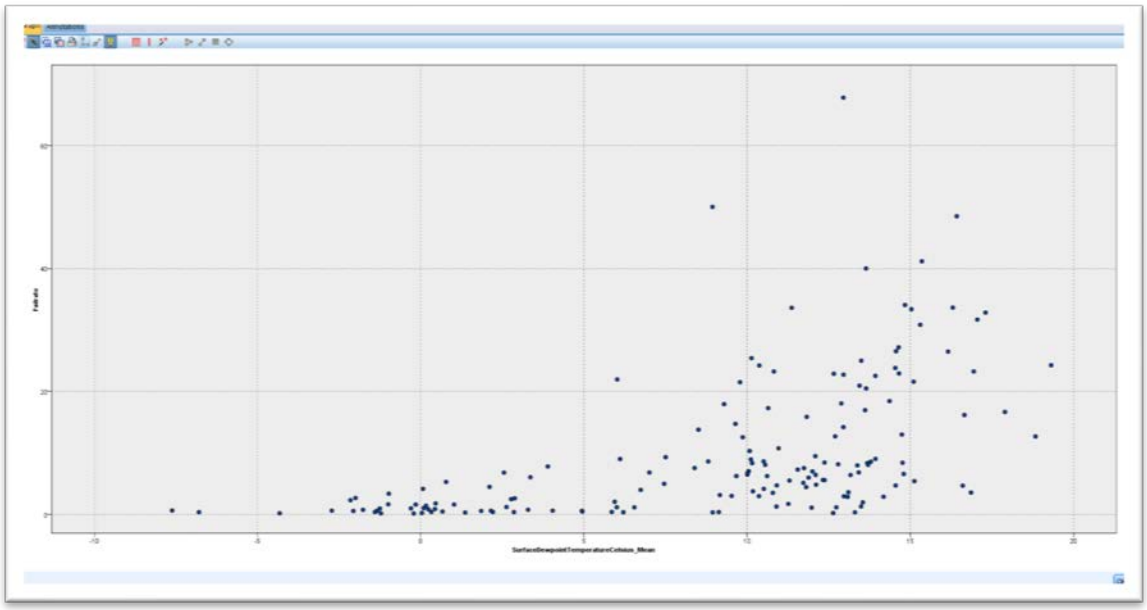
4) Notice that **SurfaceDewpointTemperatureCelcius_Mean** correlates the most and is the leading contributor to scrap rate with a confidence score of 53.1%.



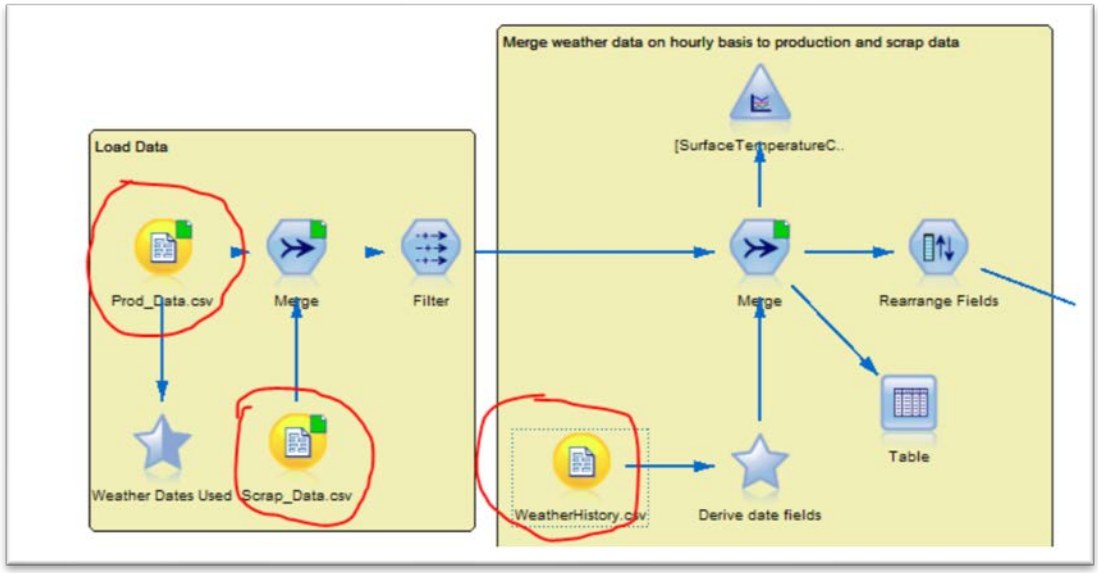
5) Next, let's see the comparison of Dewpoint Vs Scrap Rate. This is the visualization we built into the dashboard. On the third tab we show the remaining that have lower confidence levels. Double click in **Visual Correlation** and click **Run.**



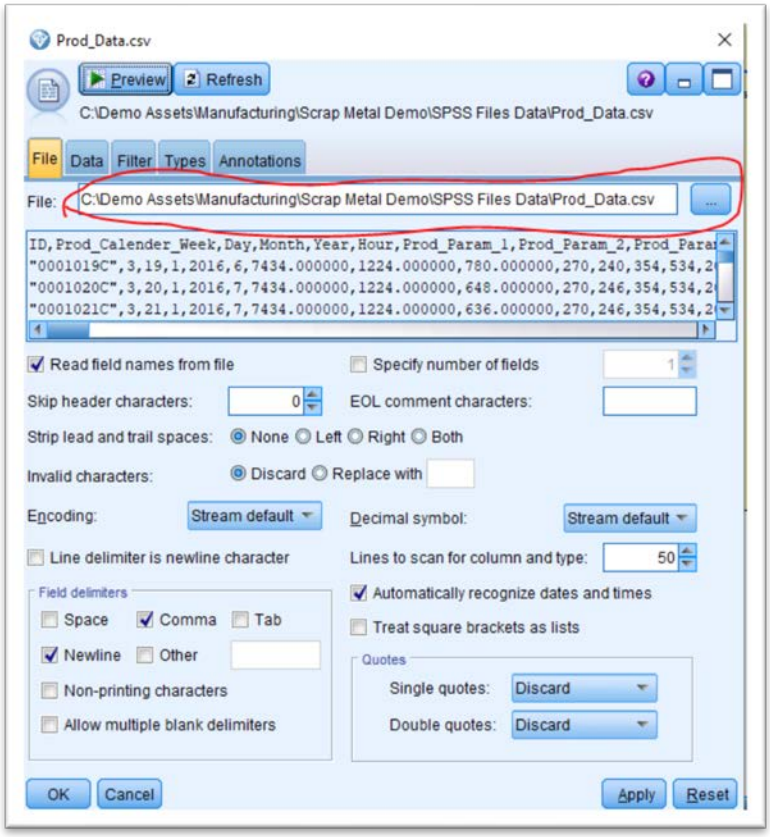
This is the same chart we added to the dashboard.



NOTE: You'll need to repoint the csv files to their new location

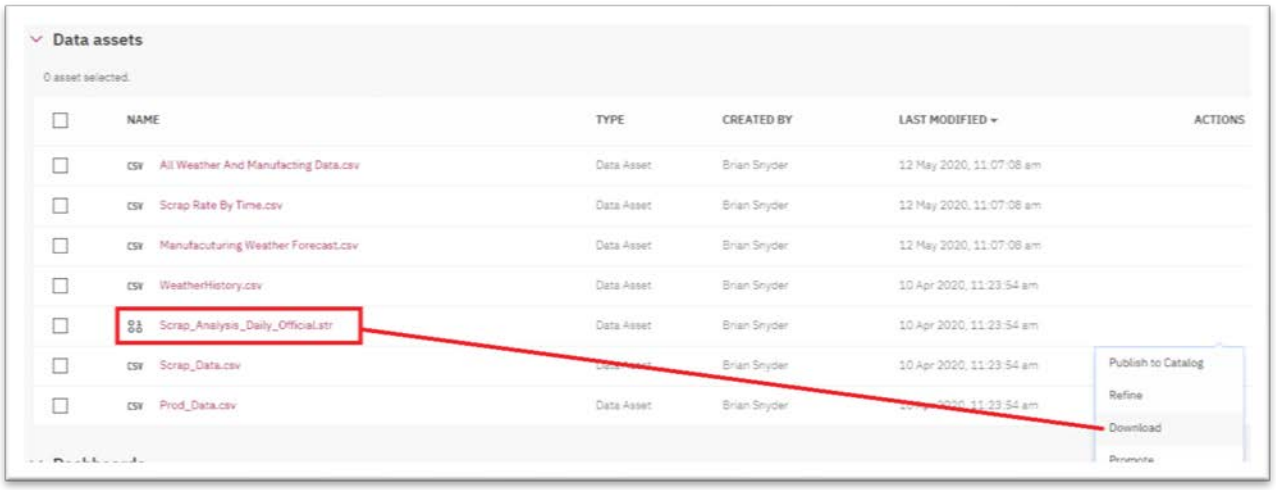


Just double click on the node and choose the file location. Here's a snapshot of one of them.

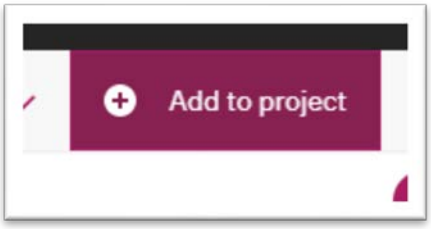


Working with Modeler Streams

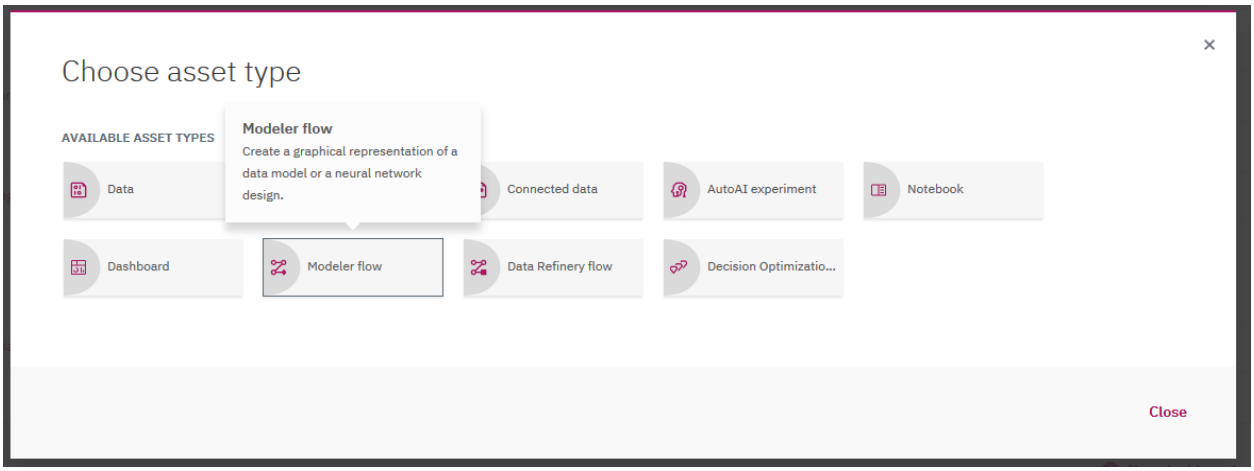
1) Download file **Scrap_Analysis_Daily_Official.str** from your Data Assets**.**



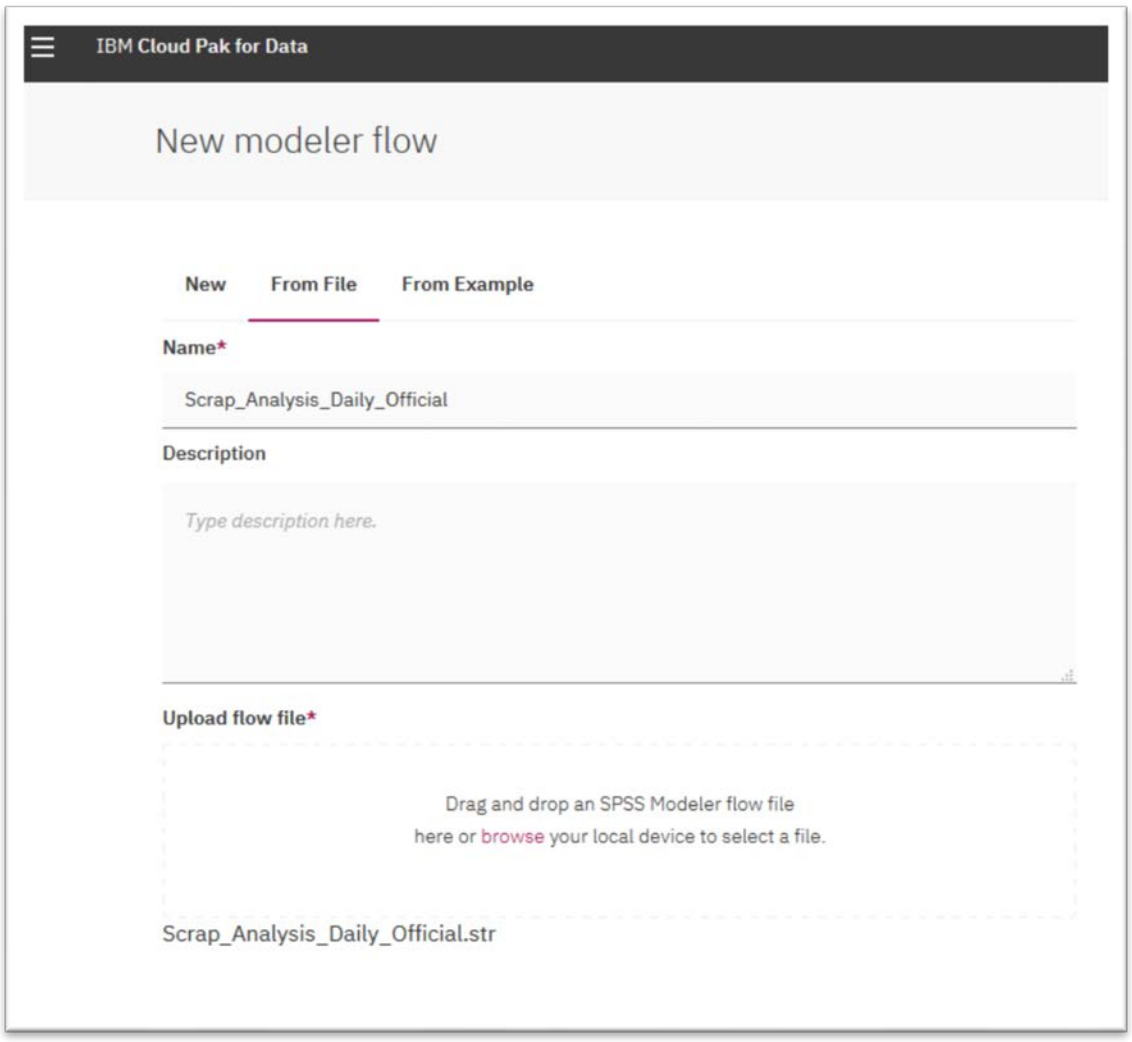
1) To add a Modeler Flow, click **Add to project**



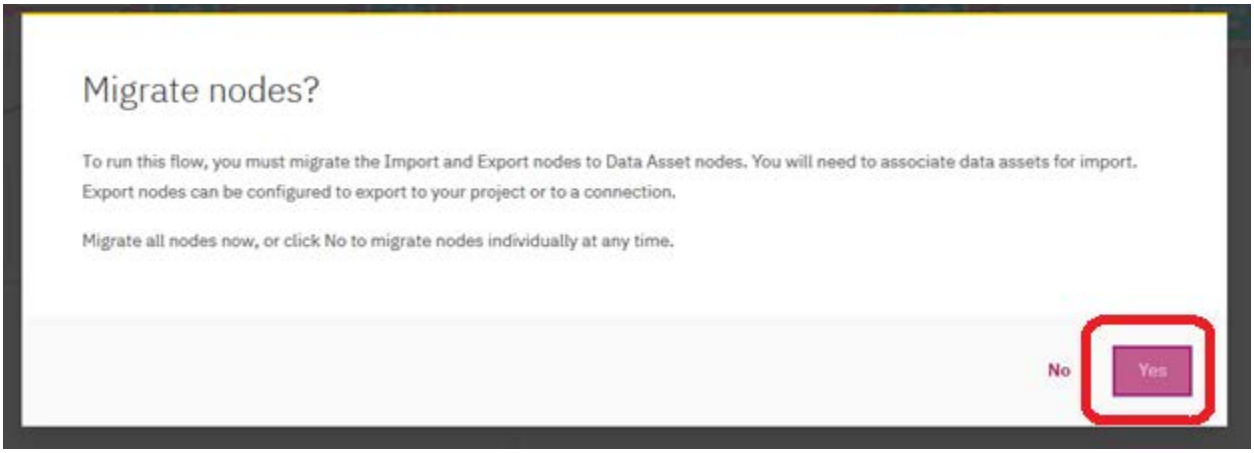
2) Choose **Modeler Flow**



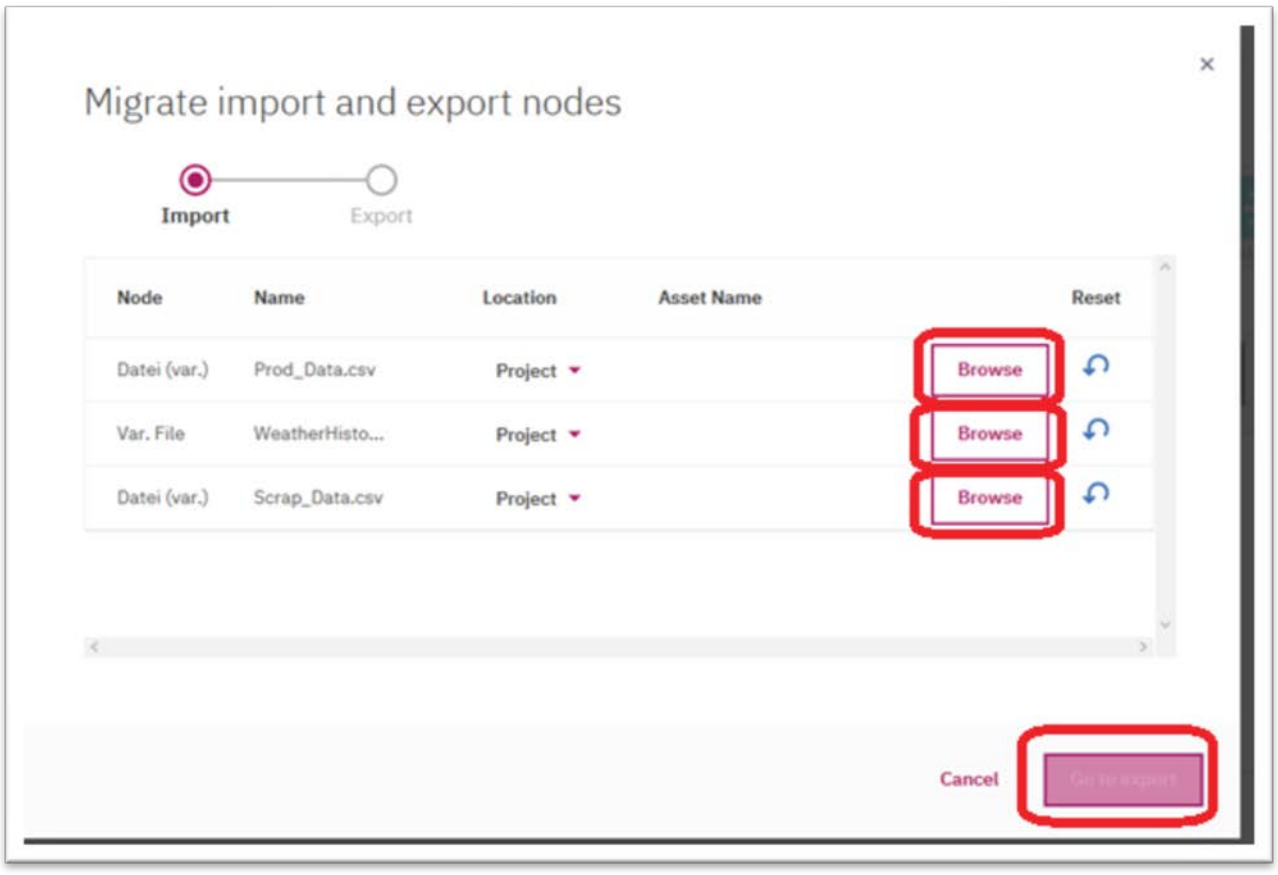
3) Choose From File and use the file labeled *Scrap_Analysis_Daily_Official.str*



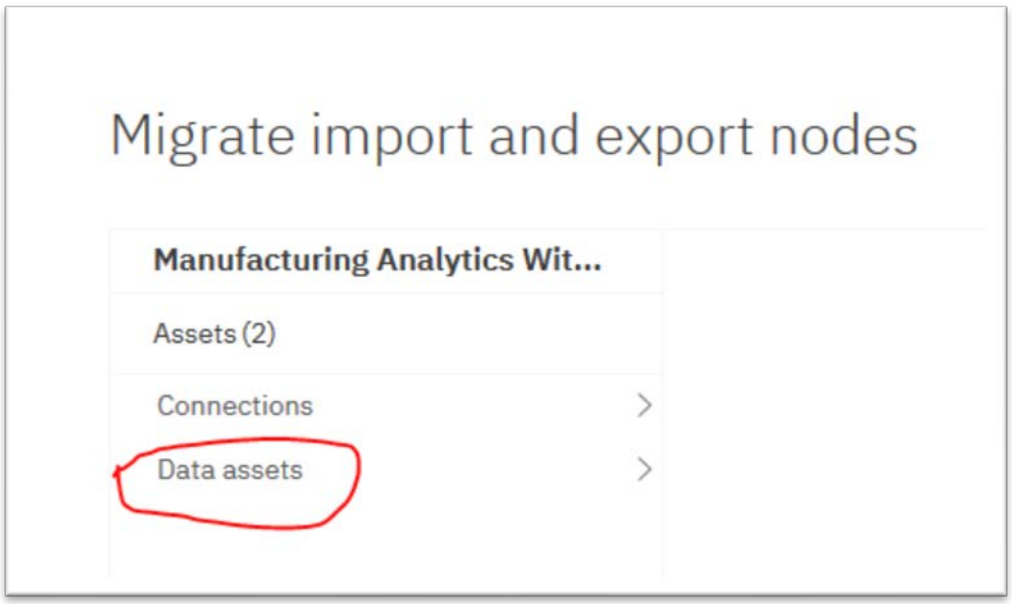
4) Choose **Yes** when prompted *Migrate Nodes?*

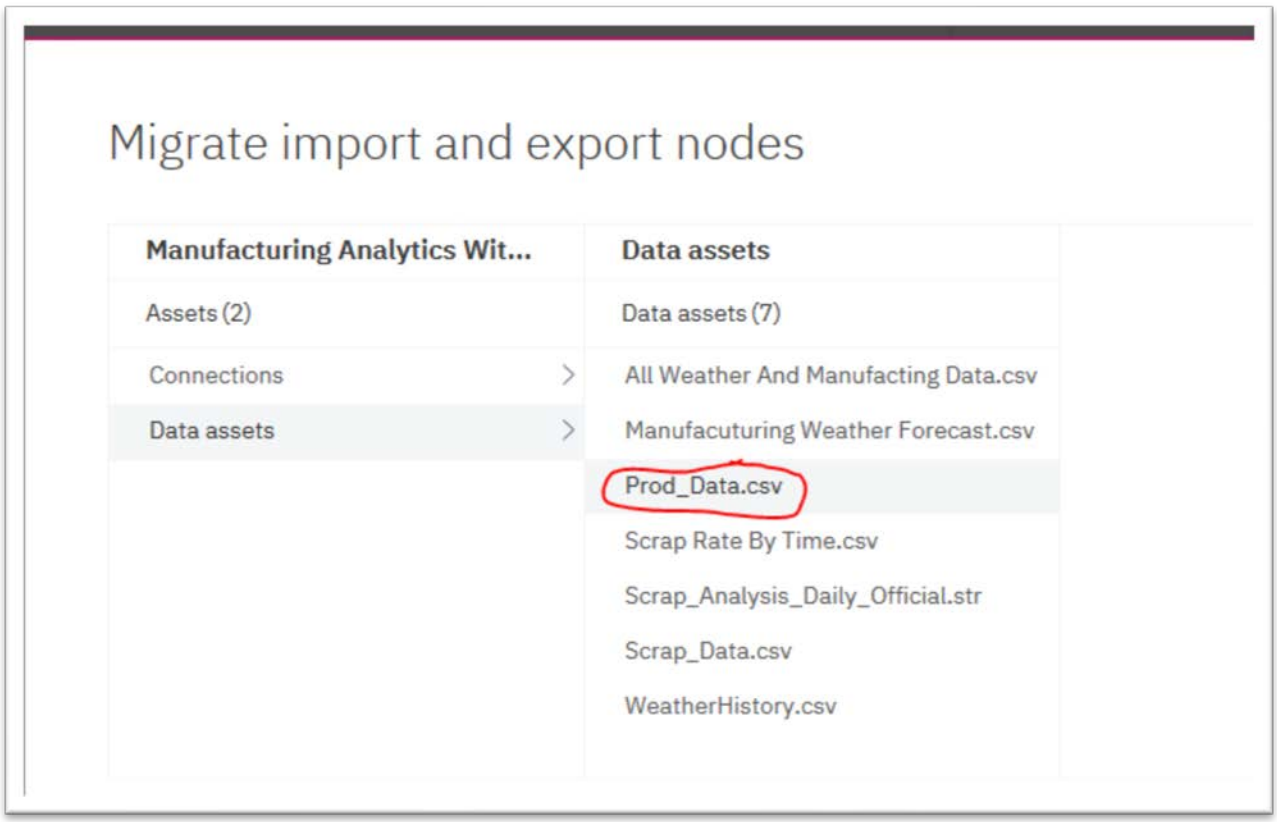


5) Find the corresponding csv file name like we did for the dashboard. First we ingest the import data and then point to the export files generated by the model.

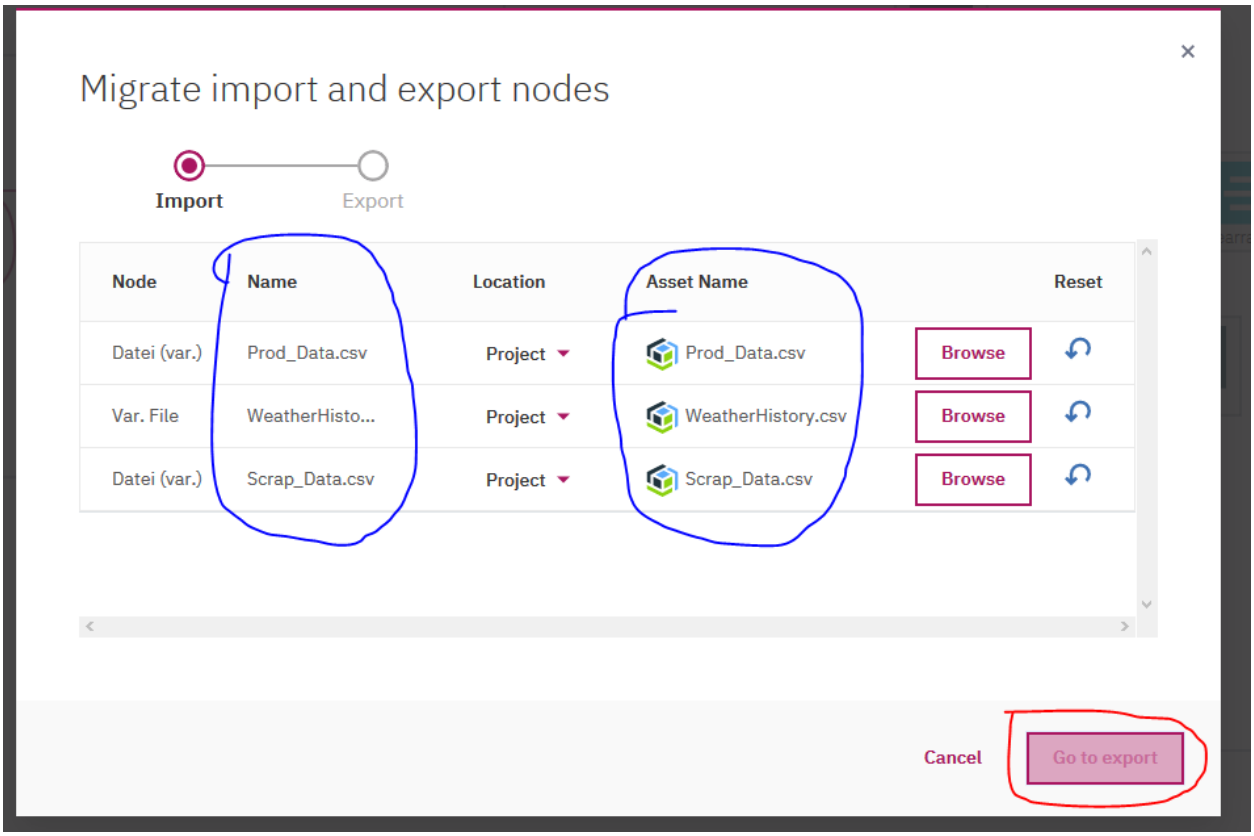


6) Choose Data Assets to find the corresponding file name.

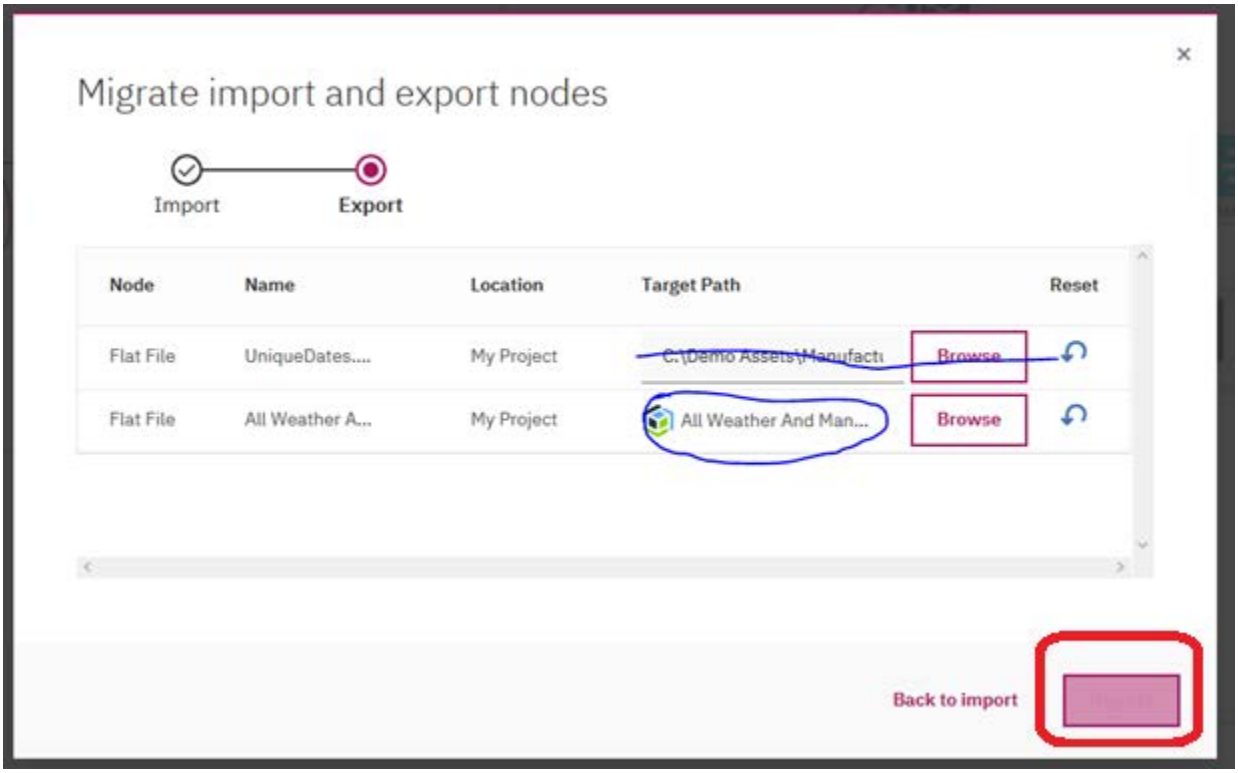




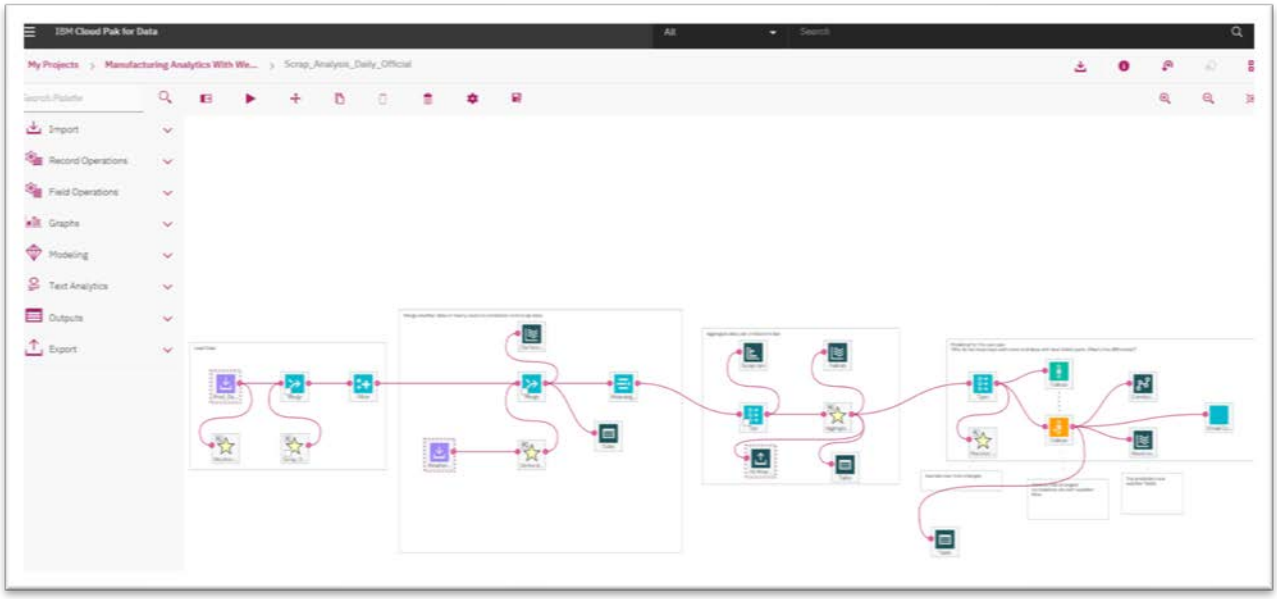
7) This is what it should like once all 3 have been populated. Click **Go To Export**



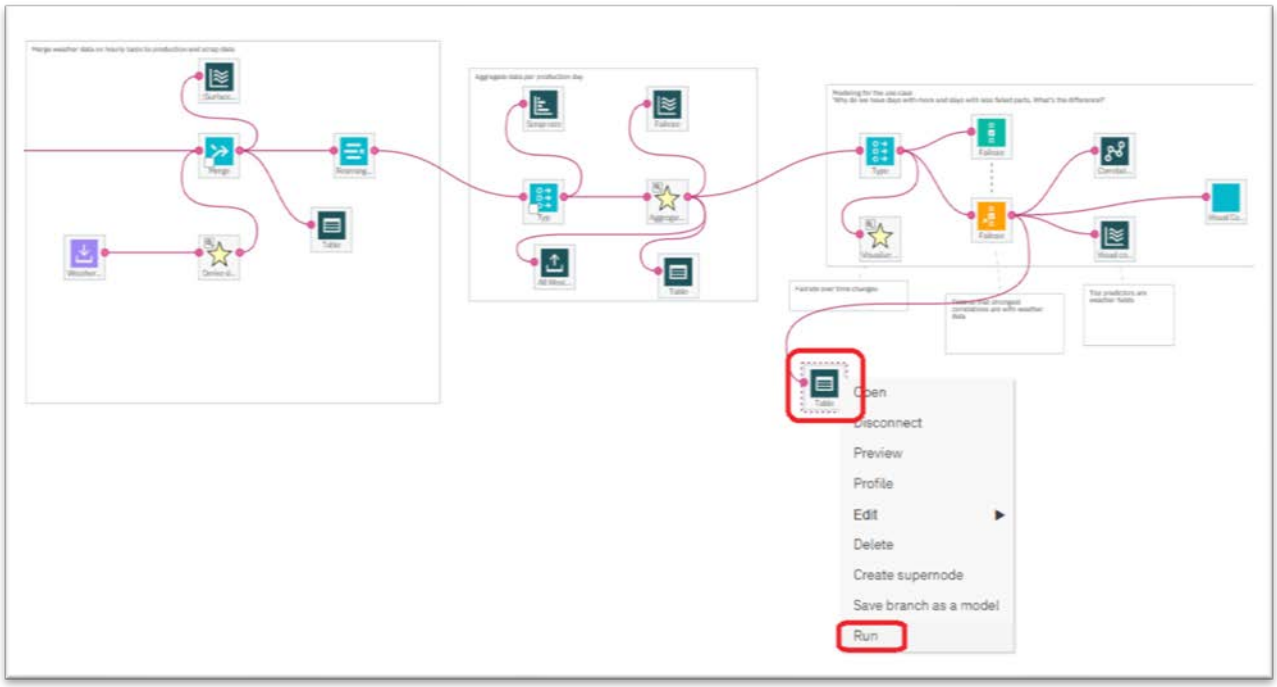
8) We're only choosing one file here. We're not utilizing the other. Choose **All Weather an Manufacturing Data.csv** only. Click **Migrate**



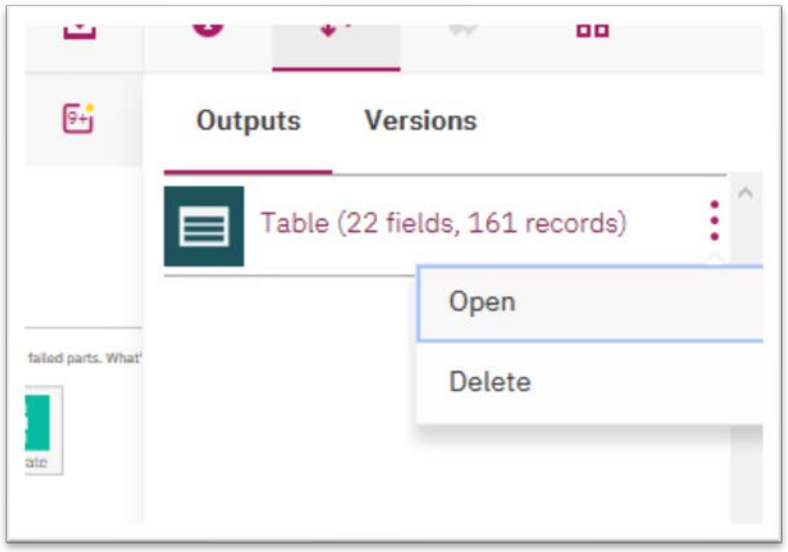
9) Here's what the model looks like once it loads.



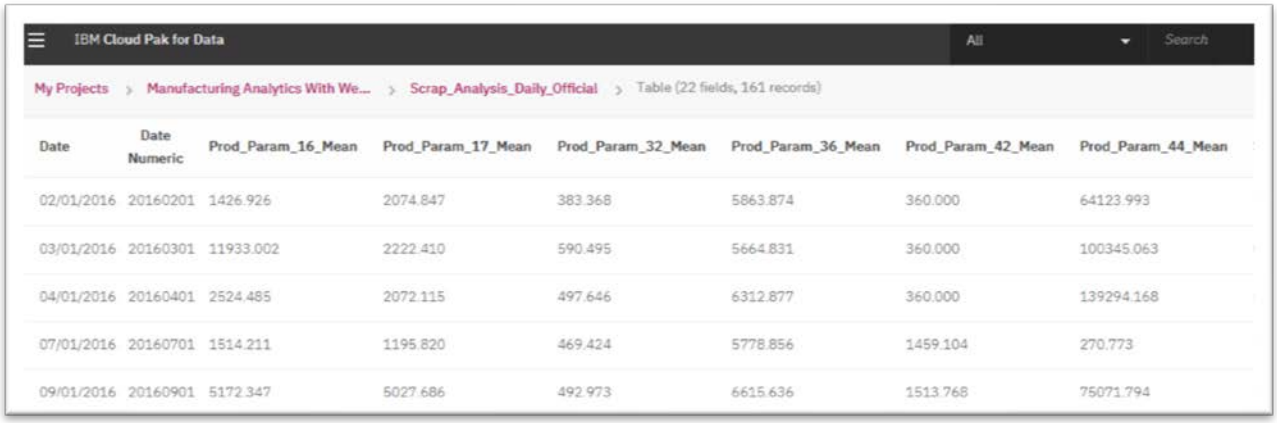
10) Modeler Flows go from left to right. To ingest the import files, merge them, add some calculations and view it, choose the Table node's triple dots , choose **Run**



11) In the upper right portion of your screen, you can choose **Open** to view the merged results of production line data and the historical weather when activity was collected.

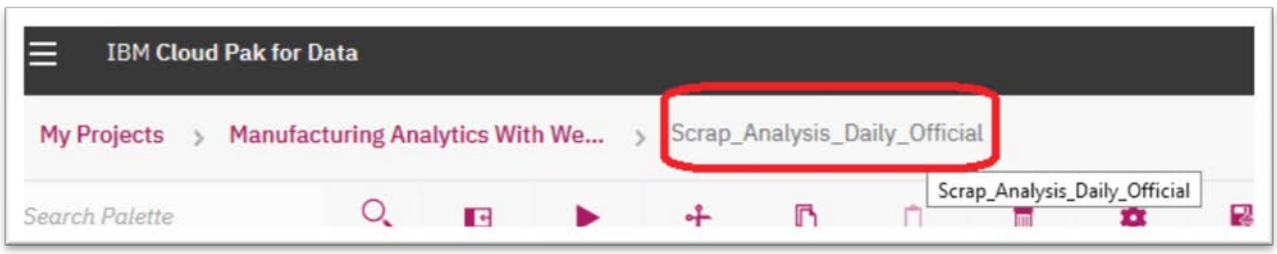


12) The results.

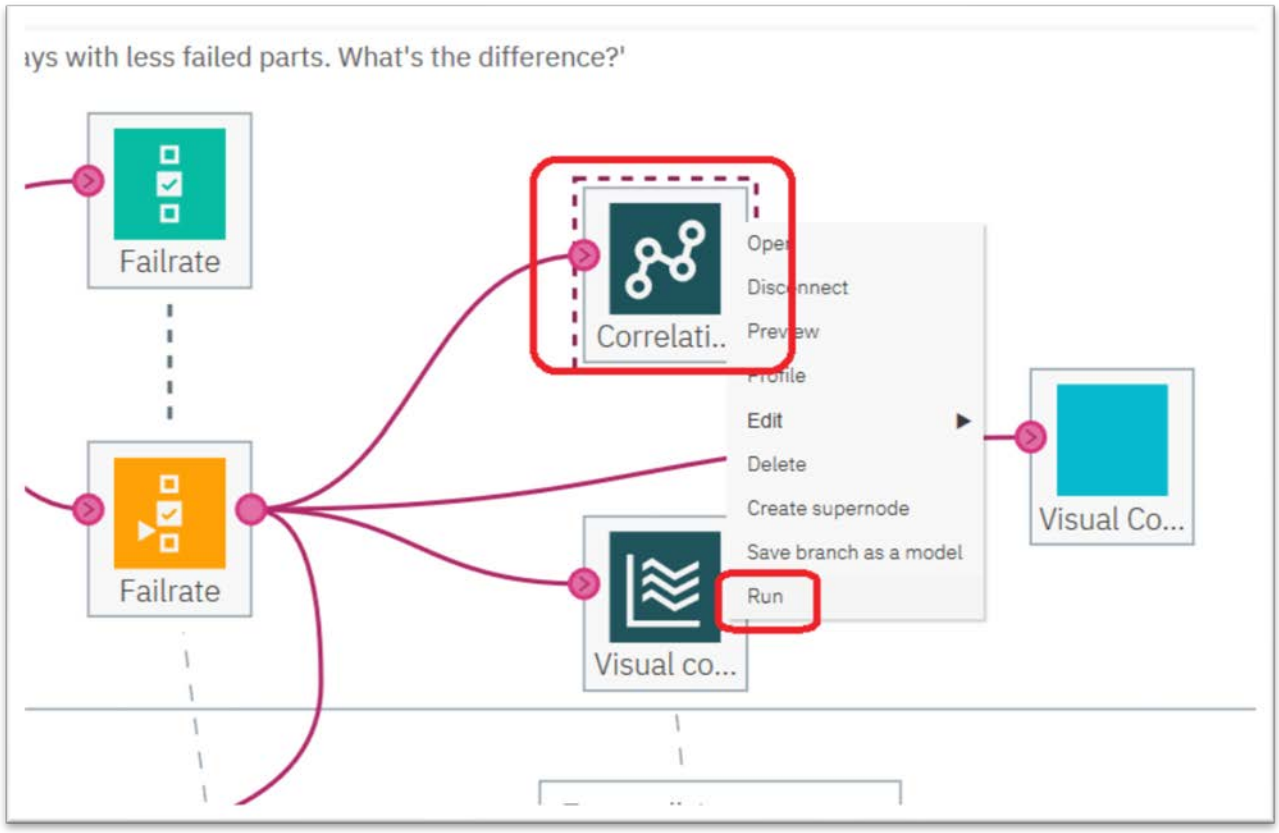


Date	Date Numeric	Prod_Param_16_Mean	Prod_Param_17_Mean	Prod_Param_32_Mean	Prod_Param_36_Mean	Prod_Param_42_Mean	Prod_Param_44_Mean
02/01/2016	20160201	1426.926	2074.847	383.368	5863.874	360.000	64123.993
03/01/2016	20160301	11933.002	2222.410	590.495	5664.831	360.000	100345.063
04/01/2016	20160401	2524.485	2072.115	497.646	6312.877	360.000	139294.168
07/01/2016	20160701	1514.211	1195.820	469.424	5778.856	1459.104	270.773
09/01/2016	20160901	5172.347	5027.686	492.973	6615.636	1513.768	75071.794

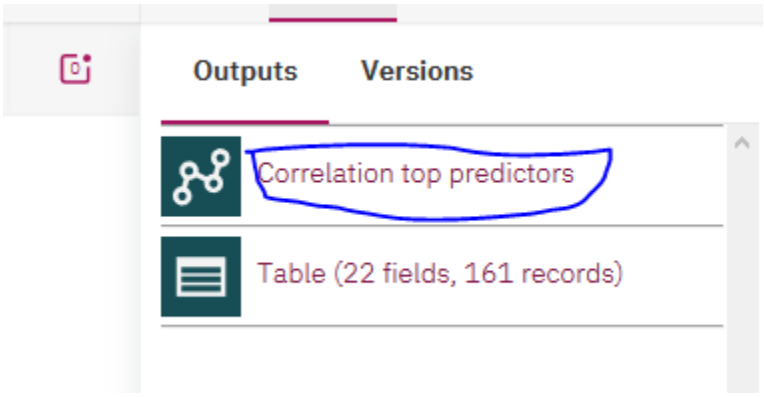
13) To return to the model, please click *Scrap_Analysis_Daily_Official* in the breadcrumb trail at the top left portion of your screen.



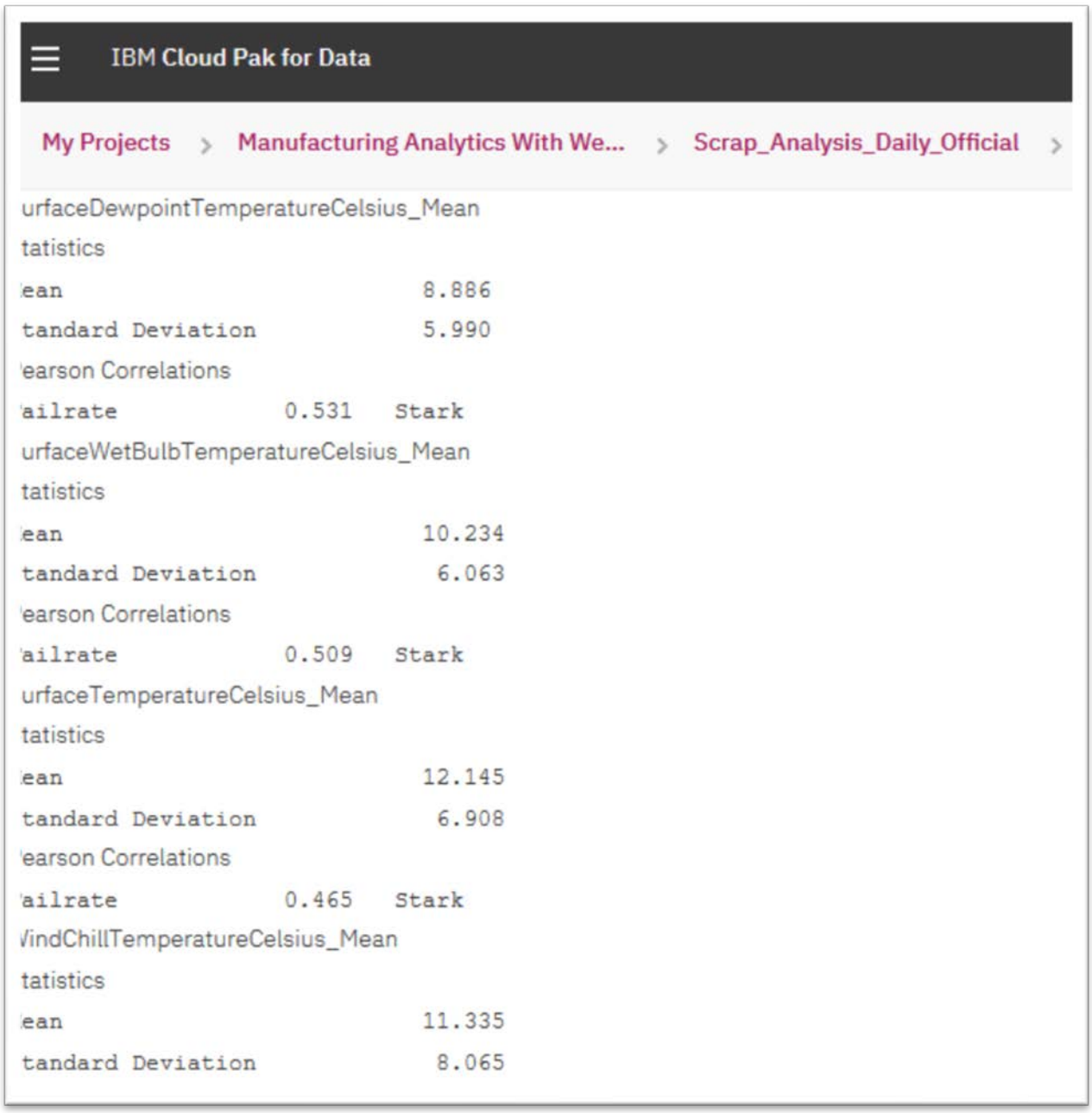
14) To view the correlation analysis like we did in SPSS Modeler, click the **Correlation** node and choose **Run**



15) Once execution has completed, click **Correlation top predictors** to view the results.



16) Here's the result. Notice **SurfaceDewpointTemperatureCelcius_Mean** correlated the most when comparing failure rate to all weather variables with a confidence score of 53.1%



Thanks for using this demo asset with weather.

If you have any questions, please reach out to bsnyder@us.ibm.com.