Retail Analytics with Weather

Thursday, May 14, 2020

Easily visualize which retail locations, products, and weather conditions have higher predicted revenue uplift in sales based on season to improve product inventory planning and improve marketing campaign effectiveness.

What's Included?

- 4 csv files
  - **Historical Sales.csv** – sample mocked up retailer sales data from the IBM Cognos Analytics sample data known as Go Sales.
  - **Historical Weather.csv** – sample mocked up weather data for 38 postal codes in the US for 6 historical dates.
  - **All Retail and Weather Data.csv** – Historical Sales.csv and Historical Weather.csv were merged into one file. Each row represents a sale and what the weather was at the time the sale occurred.
  - **Retail and Weather Model Data.csv** – This csv is the result of pulling historical weather and sales into SPSS Modeler 18.1.x (or CP4D Modeler Flow) to generate a scored data source that includes the predicted $/% revenue uplift metric coming from a Random Forest model.
- 1 SPSS Modeler Stream 18.1.x (Get a trial to SPSS Modeler here.)
  - Retail Model.str
- 1 CP4D Dashboard
  - Retail Weather Dashboard.json

How does it work?

1) Use the dashboard to visualize the correlation between top and bottom stores, products, and weather condition by season leveraging historical sales and weather to help an analyst determine where the most revenue uplift exists when doing marketing or inventory forecasting.
2) Examine the SPSS Model and CP4D Modeler Flow to see how you can ingest and develop code free models (or bring your code if you like) and prepared data comprised of historical retail sales and historical weather, in this case, to create a metric called Revenue Uplift representing the potential upside.

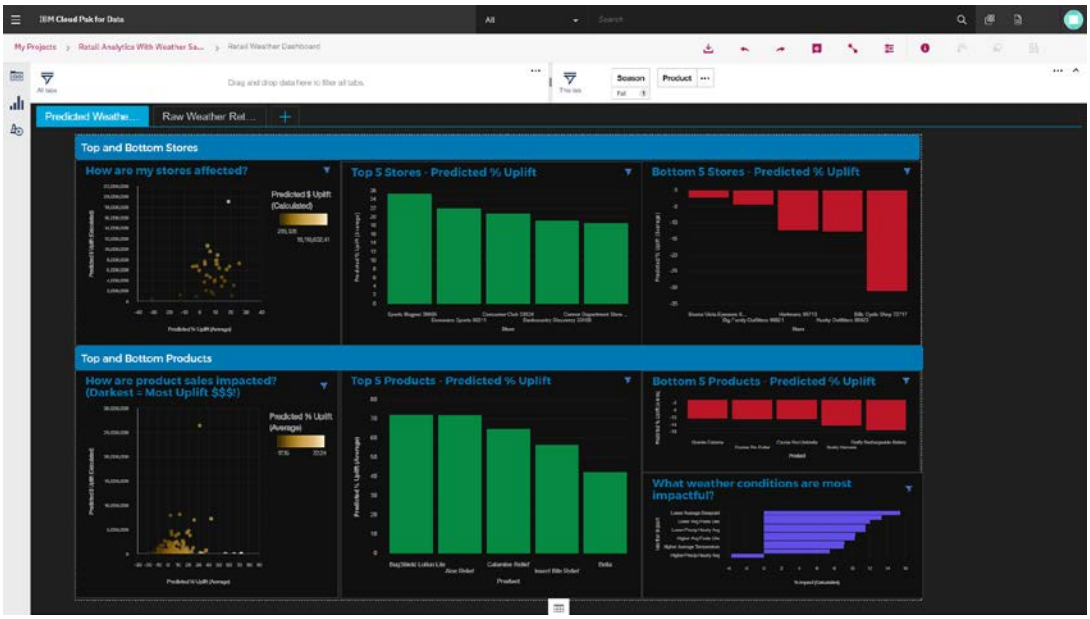
Learn About The Dashboard Showing Predicted $/% Revenue Uplift By Store, Product and Weather Condition. And, Historical Sales and Weather.

The dashboard answers the questions:

- Which products do we promote by store and weather condition for an upcoming marketing campaign or upcoming inventory planning session?
- Which products do we discontinue?
- Which products and stores are our top performers?
- What was the weather when a sale occurred?
- How does weather impact sales?

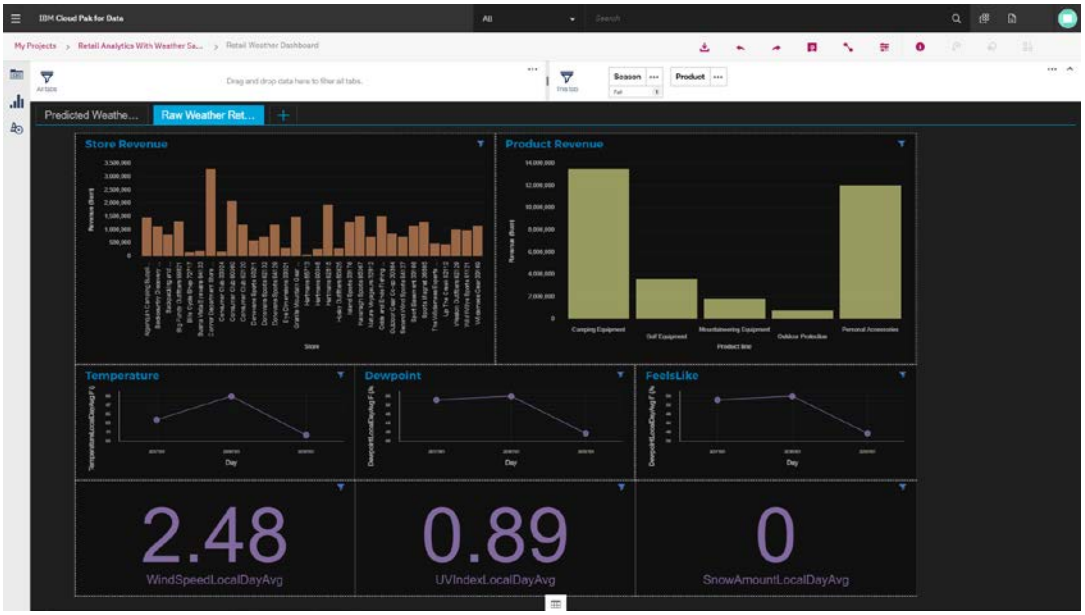
The Predicted Weather Retail Impact Tab will depict weather conditions contributing to sales and predicted % uplift/decrease by weather condition, retailer and product.
When you click on a retail location (or product or weather condition), the entire tab of this dashboard is filtered on your selection.
Change the season (upper right) to view the impact Summer, Spring and Fall have on sales.

The metric Predicted %/$ Uplift was created using a predictive model prior analysis. (See SPSS Modeler/Modeler Flows Below)



Tab 2

The second tab allows you to conduct top down analysis and when filtered on a specific view, be able see the historical weather conditions when the sale(s) took place.



Learn About The Retailer Revenue Uplift Prediction Model Using SPSS Modeler Desktop & Cloud Pak for Data Modeler Streams

All marketers look to identify which optimal combination of stores, products and weather conditions to target in a marketing campaign. All would like to run analytics on and predict revenue uplift based on historical sales and weather on the platform versus other ways of analysis that don't leverage the power of statistical analysis. Since competition is rife, finding the best combination is imperative. Most stores carry very limited inventory and need to be smart about how to promote what they do have in stock. The dashboard above shows a scored data set (*Retail and Weather Model Data.csv*) that contains the metric Predicted $/% Uplift guiding analysts to a smarter place to start when looking to conduct a campaign or make decisions on inventory and forecast.

The SPSS Modeler Stream below follows the (CRISP-DM) methodology of Data Mining and Data Science to:

1) Develop a revenue uplift predictive model (Business Understanding)
2) Ingests Historical Weather Company Weather Data and historical sales (Data Understanding and Data Preparation)
3) Build and Evaluate a Random Forest Model to predict revenue %/$ uplift (Modeling and Evaluation)
4) Build Predictive Dashboards so analysts can quickly identify optimal store, product, and weather combinations for a marketing campaign, inventory plan, or forecast. (Deployment)

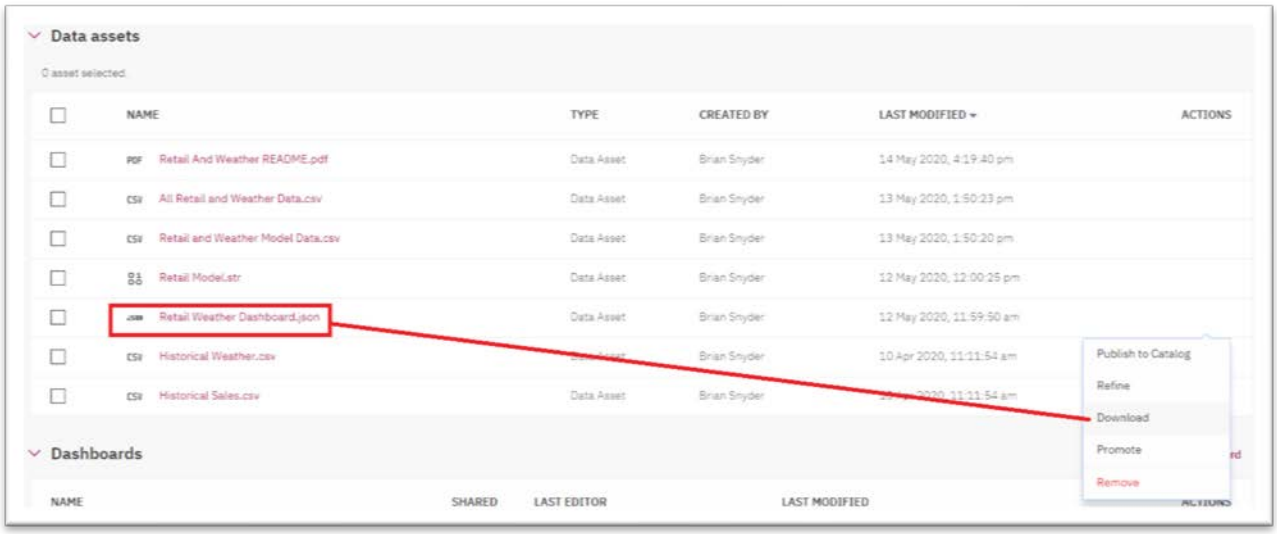
Prerequisites

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

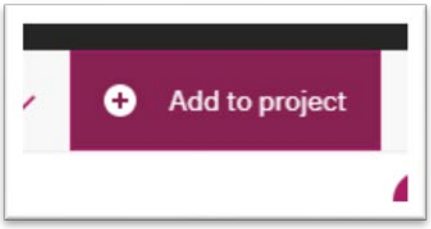
Service/Software	**Required Info**
Watson Studio	Importing data science assets to an analytics project. See Installing Watson Studio.
(Optional) SPSS Modeler 18.1.x	(Optional) If user wants to investigate the model (**Retail Model.str**) they can use SPSS Modeler Desktop to do so. It's a code free (R 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. Get a trial to SPSS Modeler here.
(Optional) Modeler Flow	(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.

Importing the Dashboard

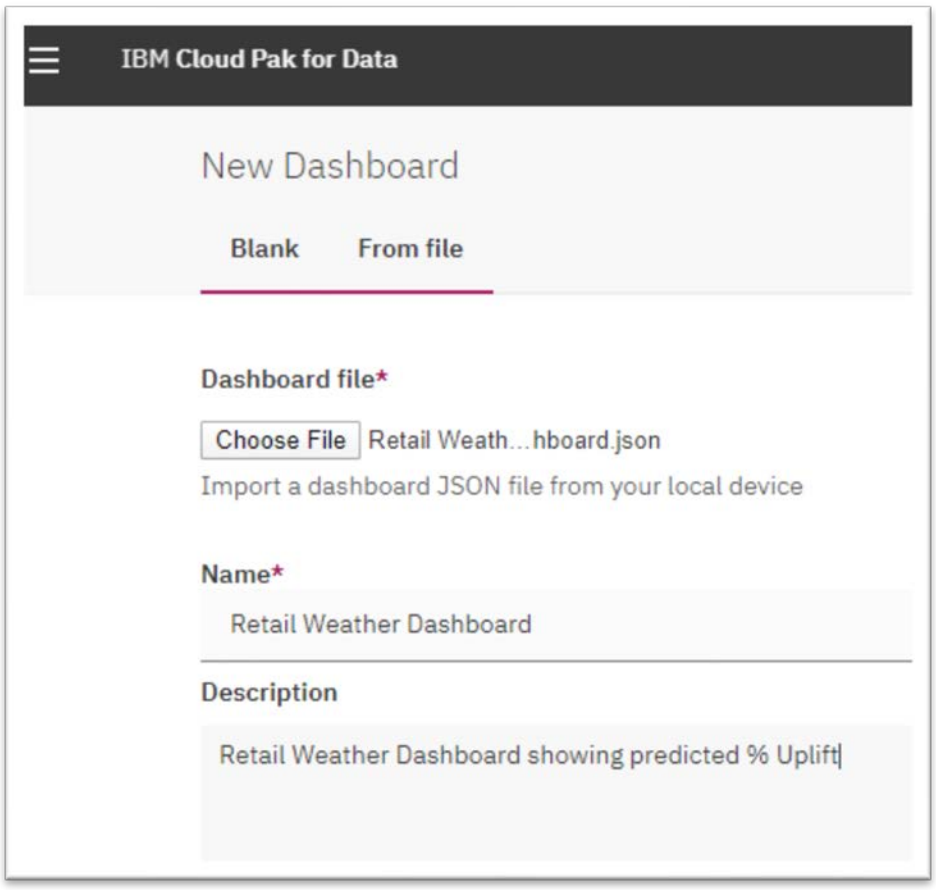
1) Go to your Data Assets and choose to Download **Retail Weather Dashboard.json**



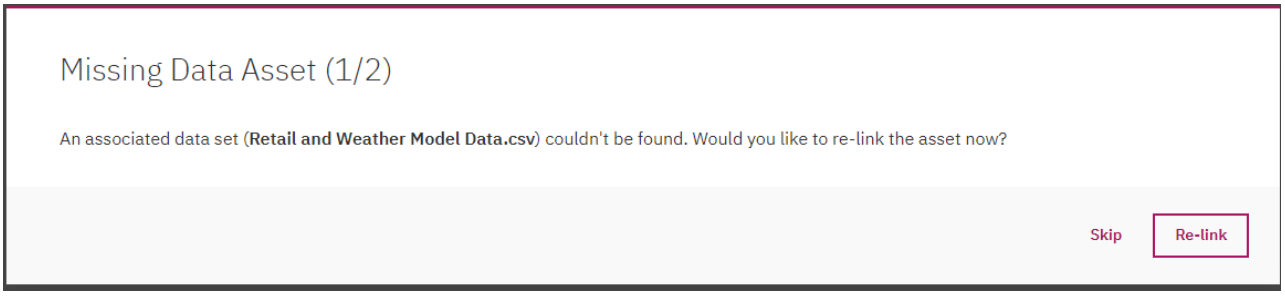
2) To Import the dashboard, In the upper right portion of your screen, click **Add to project**



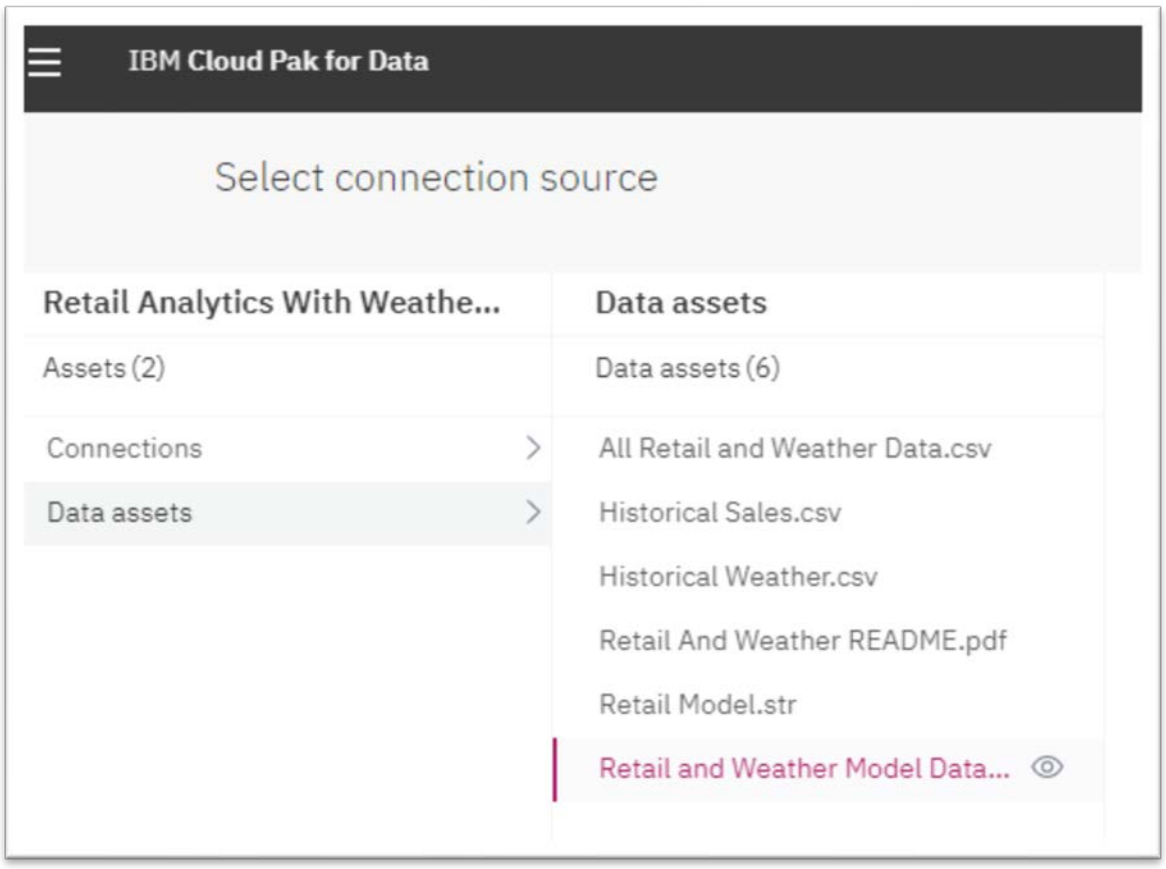
3) Select *Dashboard*, choose **From File**.
   Choose the file you downloaded called **Retail Weather Dashboard.json**
   Type in the name **Retail Weather Dashboard.**
   Add the description **Retail Weather Dashboard.**
   Click **Create**



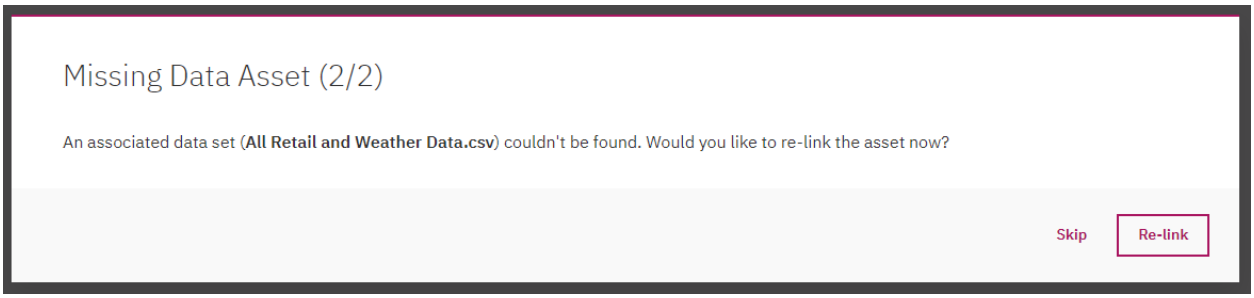
4) You'll be prompted to point to the correct missing data asset for **Retail and Weather Model Data.csv**. Click Re-link.

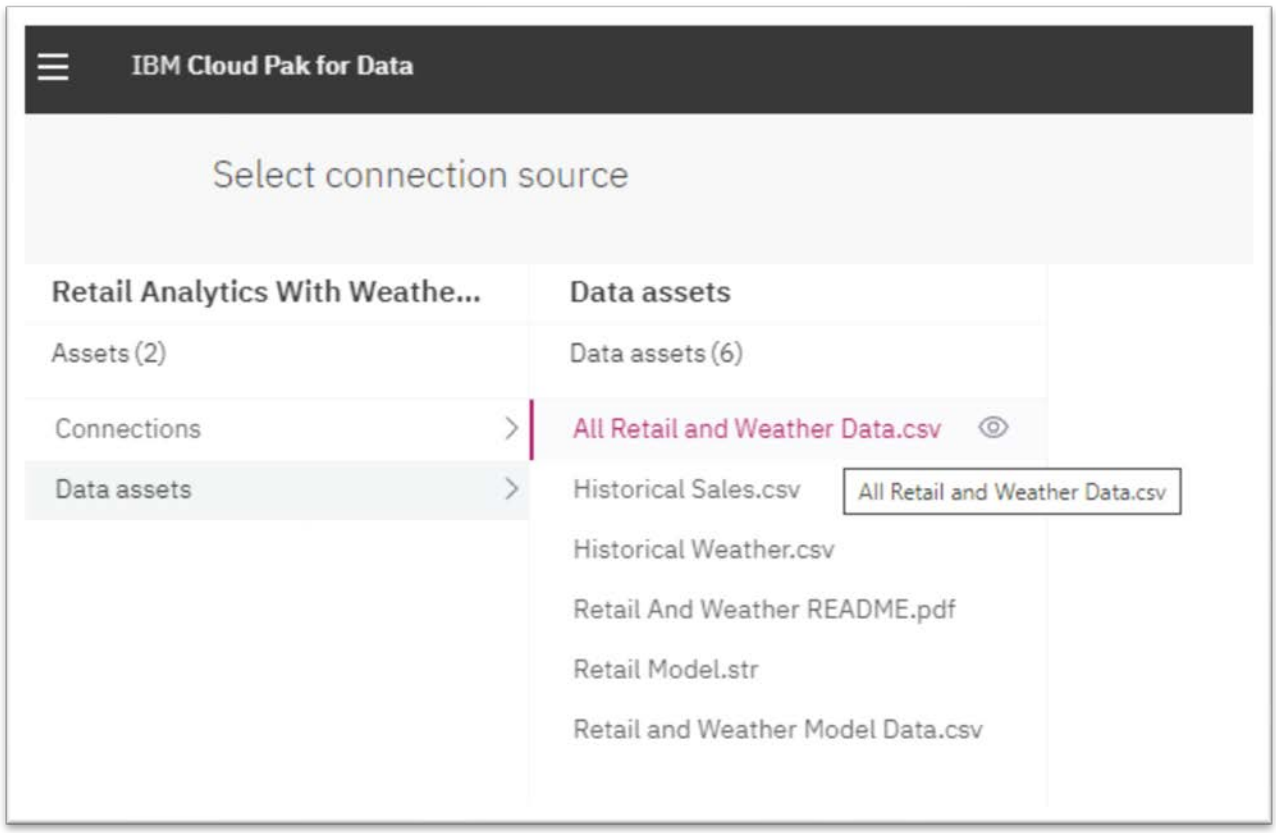


5) Choose **Retail and Weather Model Data.csv**

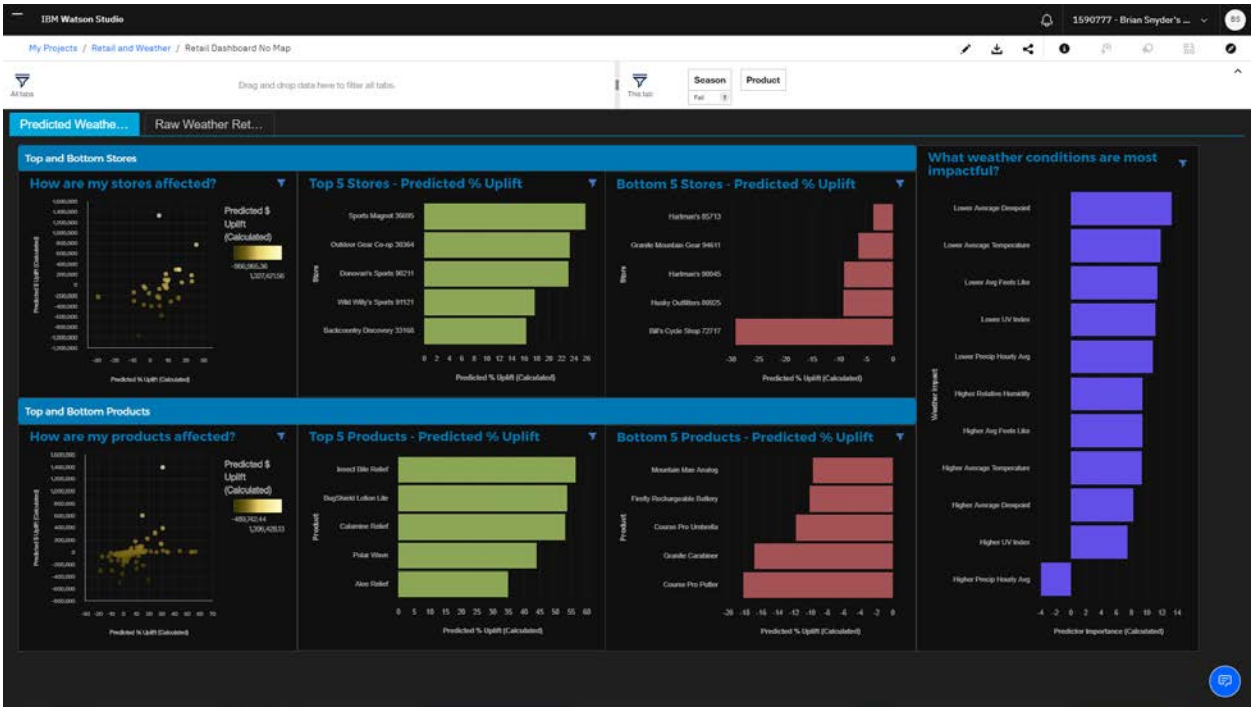


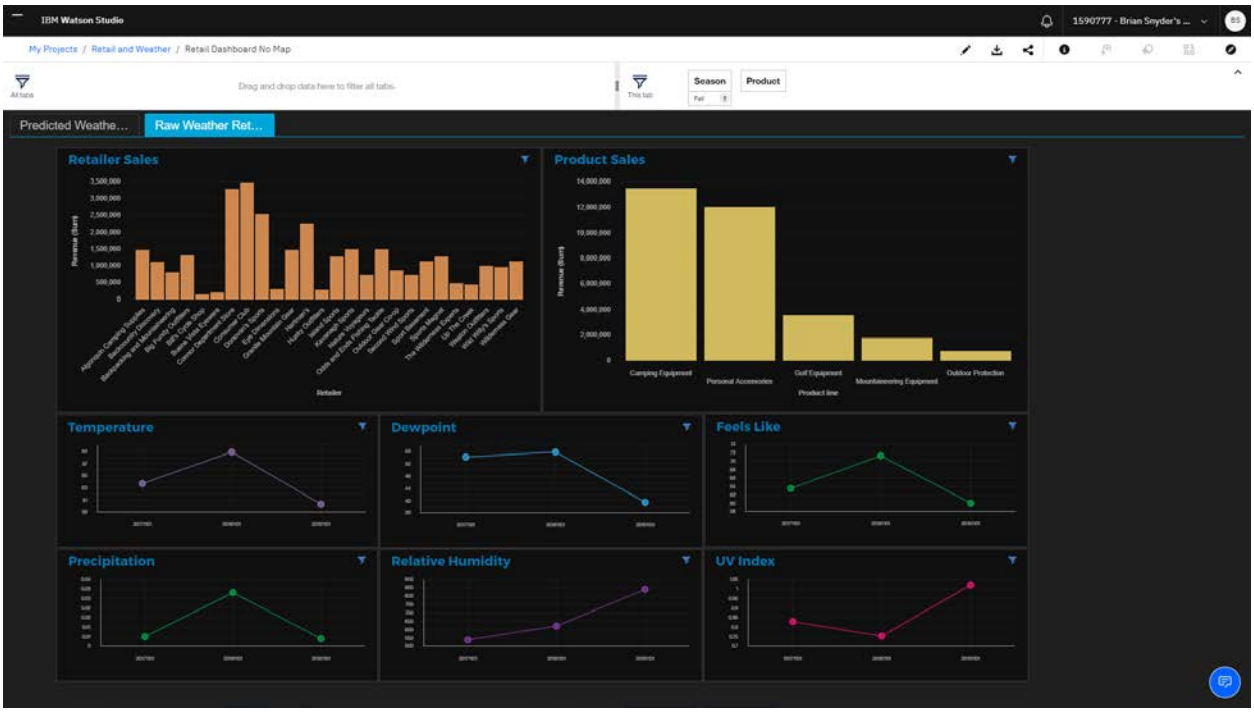
6) Next, you'll be promoted to point to the other required csv for the dashboard called **All Retail and Weather Data.csv** Click **Re-link**



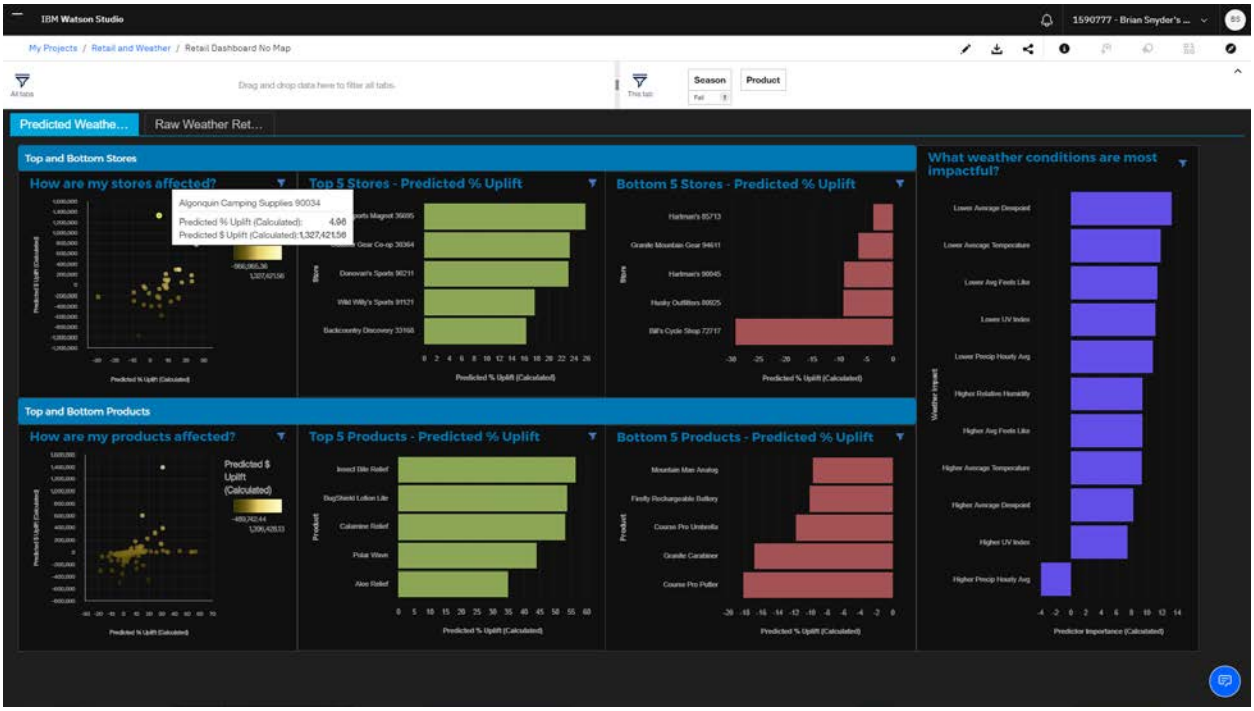


7) The dashboard should open and look as such on the 1st and 2nd tabs:



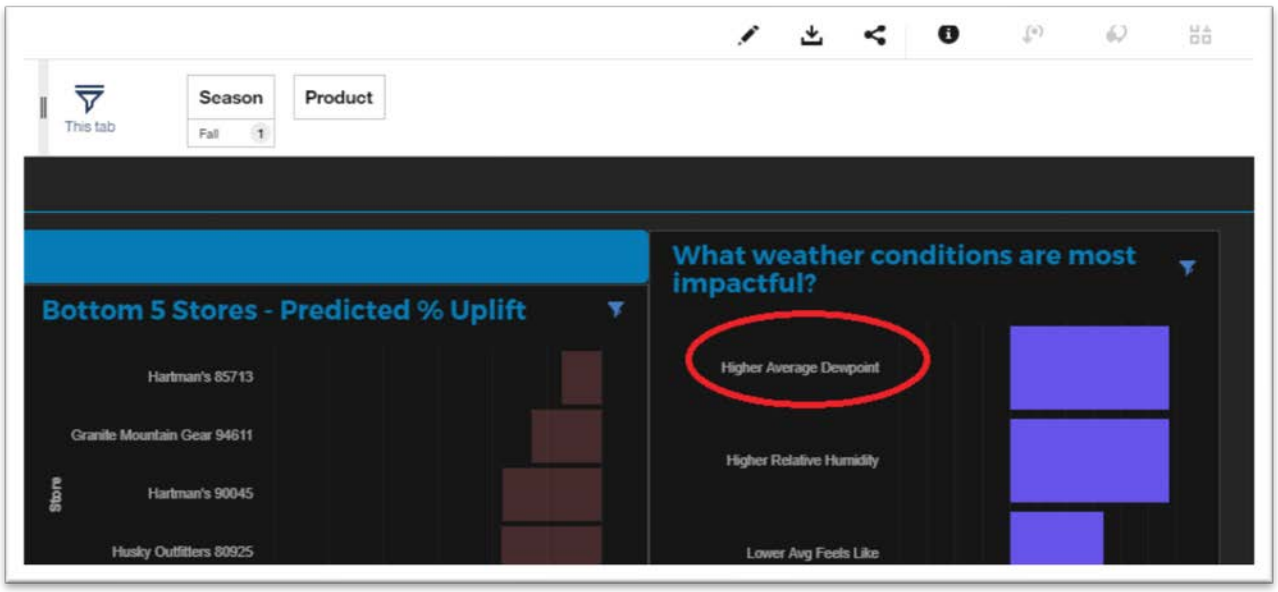


8) The dashboard is interactive. On the first tab, click on the store with the highest predicted % : **Algonquin Camping Supplies**.

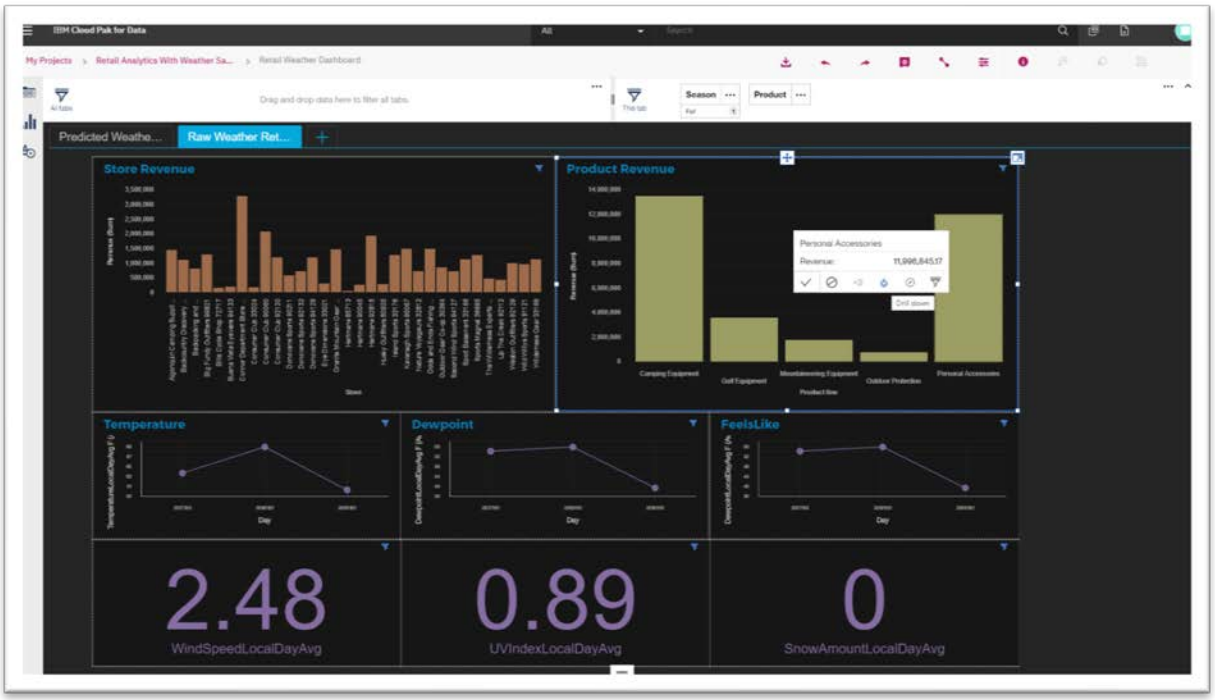


9) Next, click on **Cat Eye**, a product predicted to have the one of the highest predicted $ uplift in the fall.

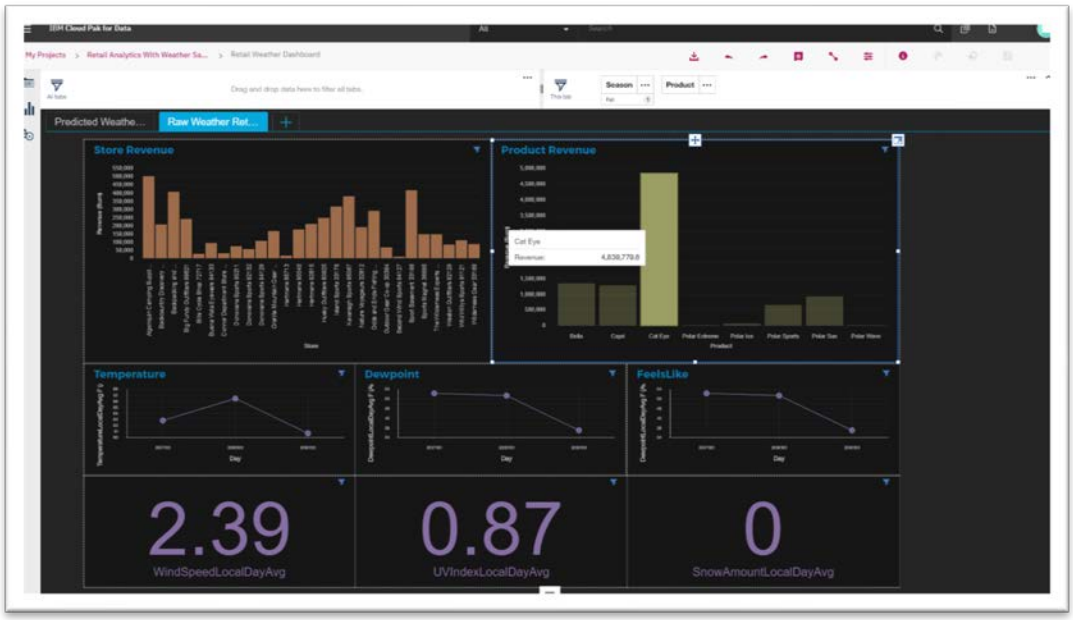
10) Looks like we should focus marketing on this store and product in the fall when the weather conditions have higher than average dewpoint in the forecast. Or, stock the shelves with this product and store with these weather conditions.



11) On the second tab, this is how most do top down analysis to find where their promotions and marketing efforts should reside by focusing on the top sellers from the past. In the end, you'll see you may have identified a top selling product but not the one likely to generate the most uplift based on season.

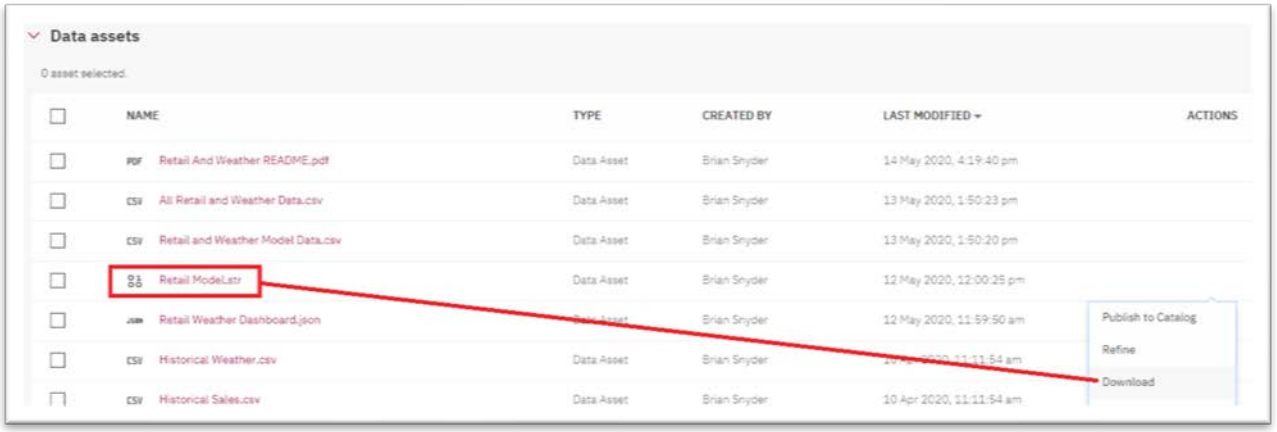


12) Right click to drill down from Product Line to Product Type to Product. Click on the top selling product to see what the weather conditions were on the days this top seller sold.


IBM Cloud Pak for Data
My Projects
Retail Analytics With Weather Sa...
Retail Weather Dashboard
Season
Product
Predicted Weathe...
Raw Weather Ret...
Store Revenue
Product Revenue
Cat Eye
Revenue: 4,838,779.0
Temperature
Dewpoint
FeelsLike
Day
2.39
WindSpeedLocalDayAvg
0.87
UVIndexLocalDayAvg
0
SnowAmountLocalDayAvg

13) If you have SPSS Modeler Desktop 18.1.x or greater, you can explore the model (*Retail Model.str*) on your own (Download SPSS Modeler here).

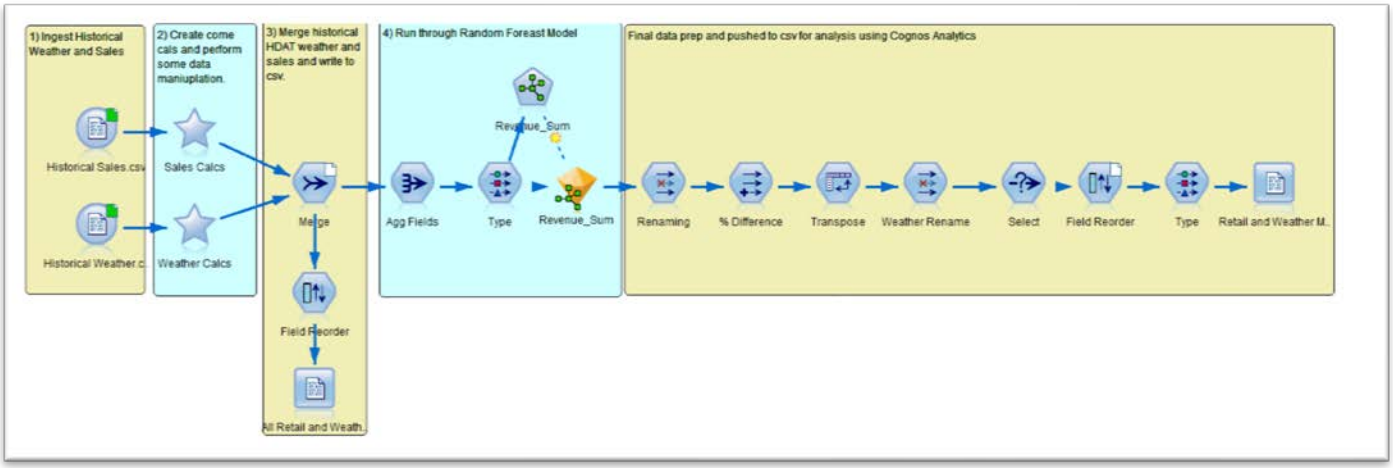
Download the file *Retail Model.str* from your Data Assets.



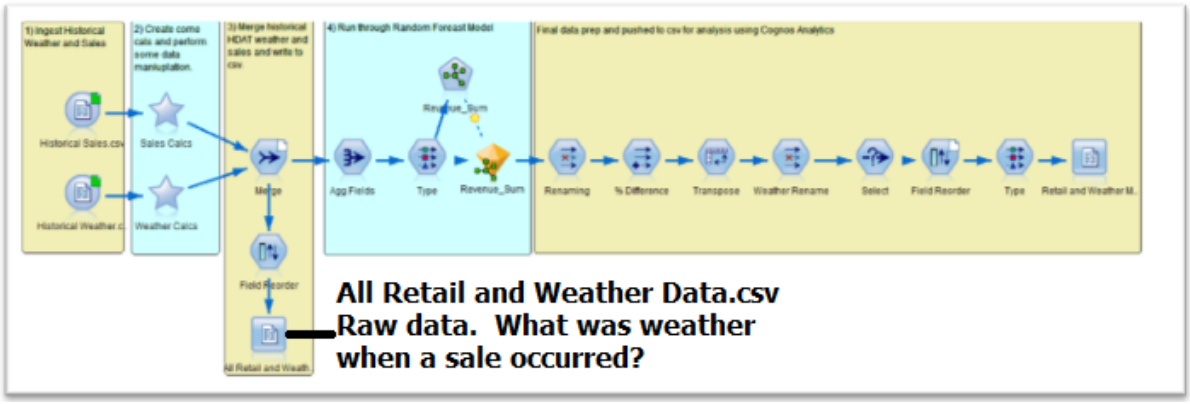
Cloud Pack For Data, has **Modeler Flows** as a premium service as well that provide the same type of environment except using CP4D.

Both are not required, but are useful at data mining and data science without the need to code in R or Python.

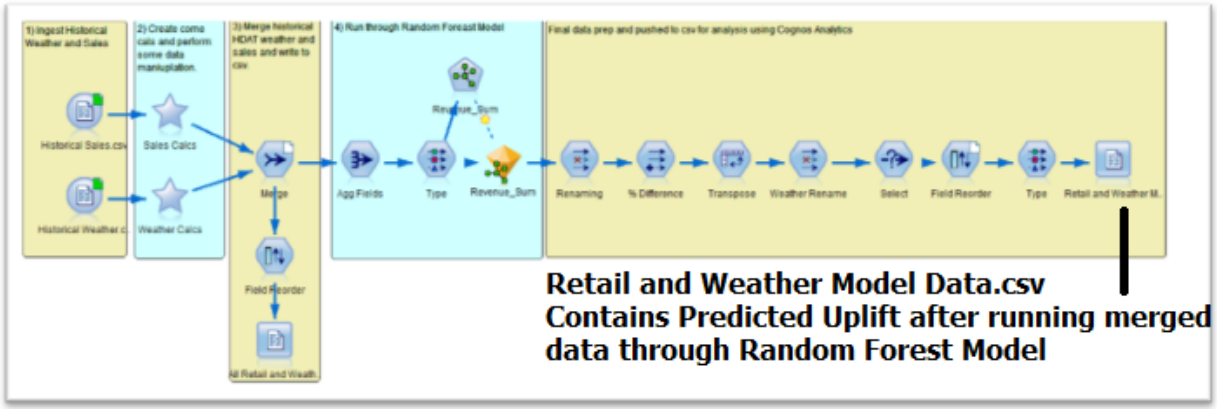
Here's the model (*Retail Model.str*) in SPSS Modeler Desktop (Download here).



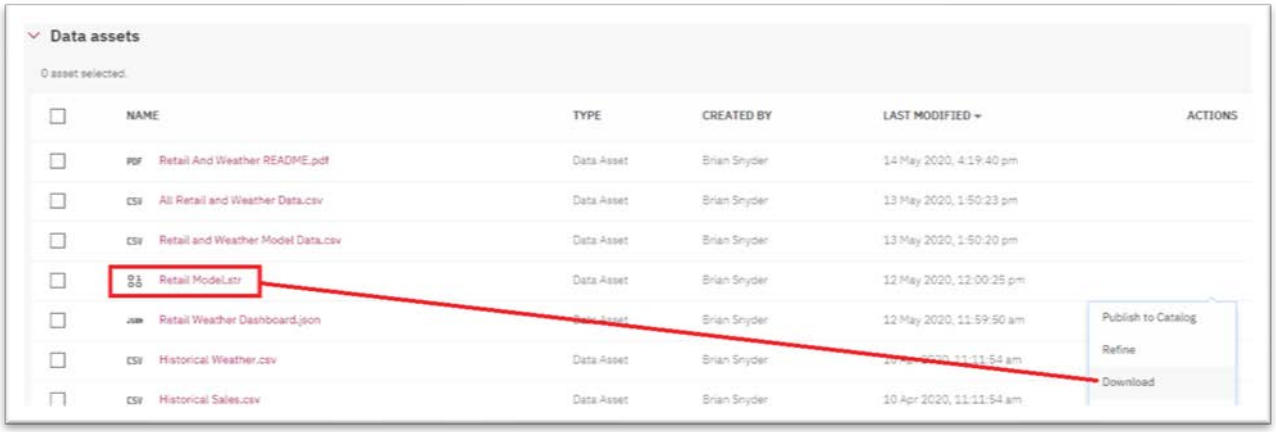
14) You can click and run this node to generate the csv file named *All Retail and Weather Data.csv*. This file is used on the 2nd tab of the dashboard. It's the merged result of historical weather and historical sales that depicts what the weather was the day a sale occurred. Most times, you'd likely be pulling and writing to database tables, but in this case, we're ingesting and writing csv files. You can overwrite the *Historical Sales.csv* and *Historical Weather.csv* which in turn will update *All Retail and Weather Data.csv* when executed. All 3 files should be added to your Data assets tab on CP4D.


1) Ingest Historical Weather and Sales
2) Create some calc and perform some data manipulation.
3) Merge historical HDAT weather and sales and write to csv
4) Run through Random Forest Model
Final data prep and pushed to csv for analysis using Cognos Analytics
Historical Sales.csv
Sales Calcs
Historical Weather c...
Weather Calcs
Merge
Field Reorder
All Retail and Weath...
Revenue_Sum
Agg Fields
Type
Revenue_Sum
Renaming
% Difference
Transpose
Weather Rename
Select
Field Reorder
Type
Retail and Weather M...
All Retail and Weather Data.csv
Raw data. What was weather
when a sale occurred?

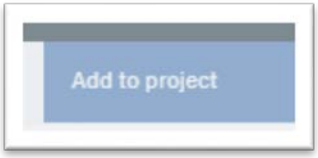
15) When executing from the far node on the right, you're ingesting the *Historical Weather.csv* and *Historical Sales.csv* file, creating some calculations, merging the files together, running it through a Random Forest Model, performing a few more data prep steps and write to file *Retail and Weather Model Data.csv* the source of the 1st tab of the dashboard.

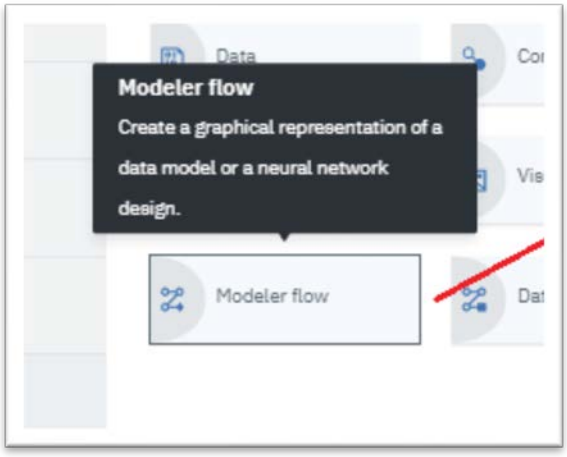


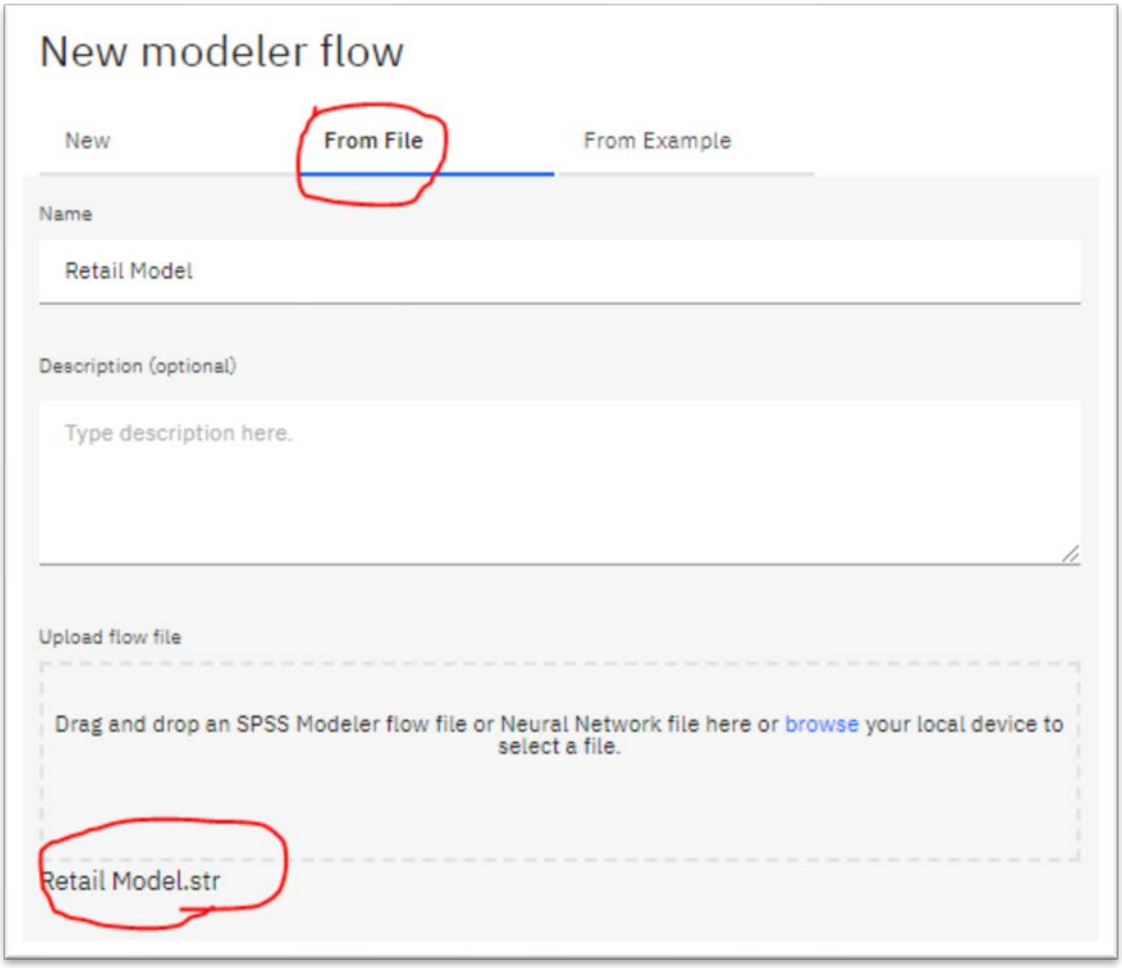
16) Download Retail Model.str from your Data Assets



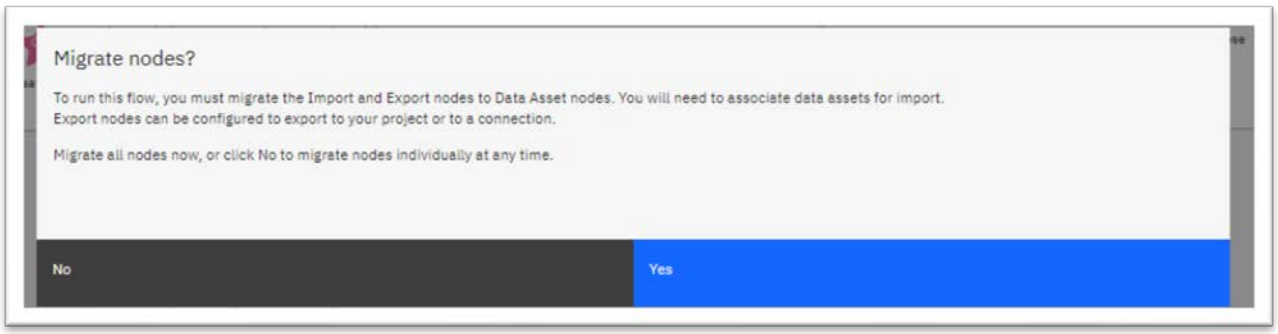
17) You can import the *Retail Model.str* into as a **Modeler Flow** in CP4D. (Available as Premium Service). This provides a web based version of SPSS Modeler Desktop for you to do your data mining using CP4D itself. To import the model, create a new **Modeler Flow** from a file. Choose *Retail Model.str* and click **Create.**

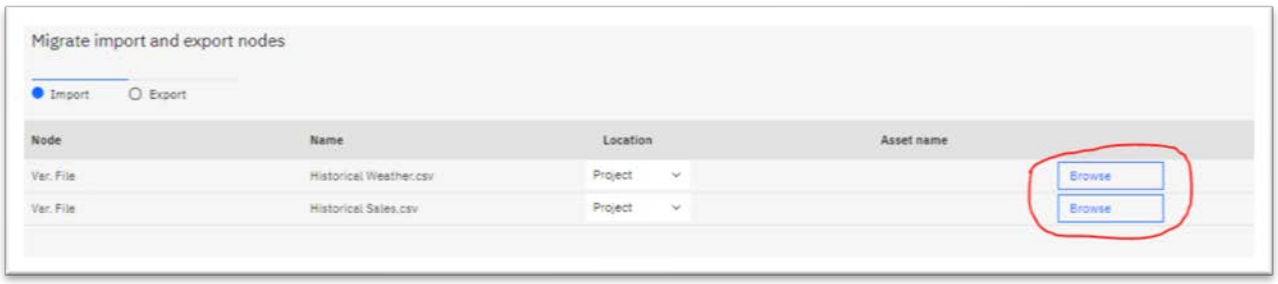


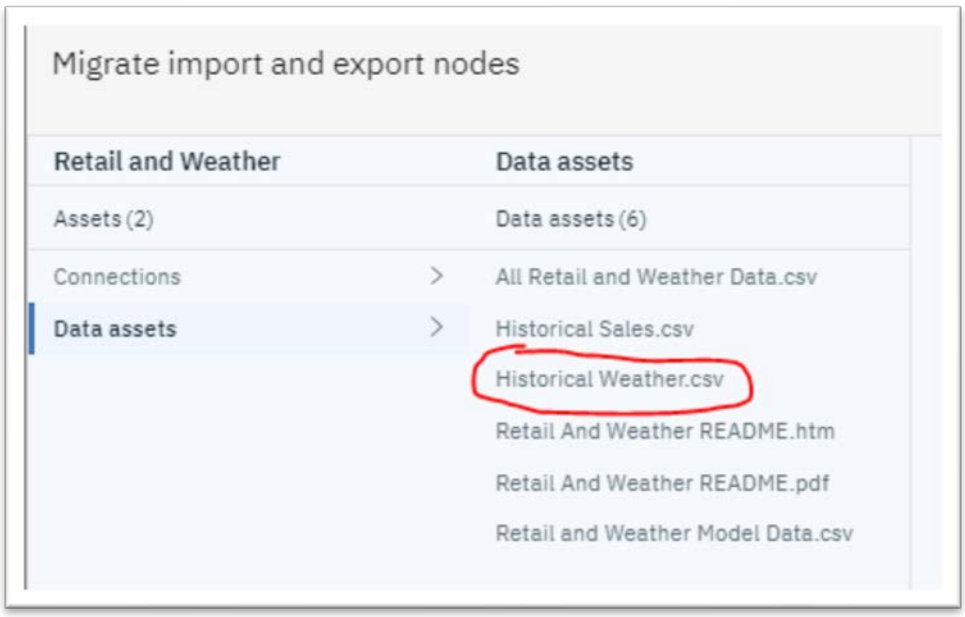


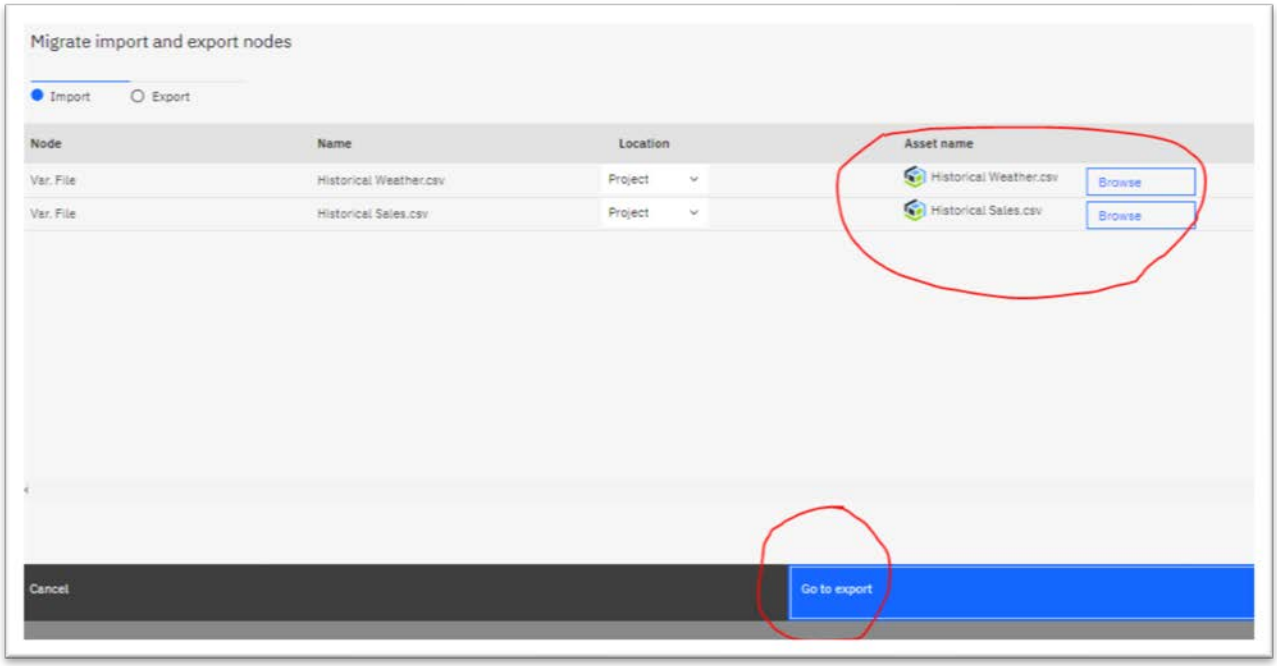


18) You'll be prompted to Migrate nodes? Click **Yes** and point the source nodes and export node to their respective Data asset you imported earlier. Click **Browse → Data Assets → Historical Weather.csv.** Click **Go To Export** when complete.

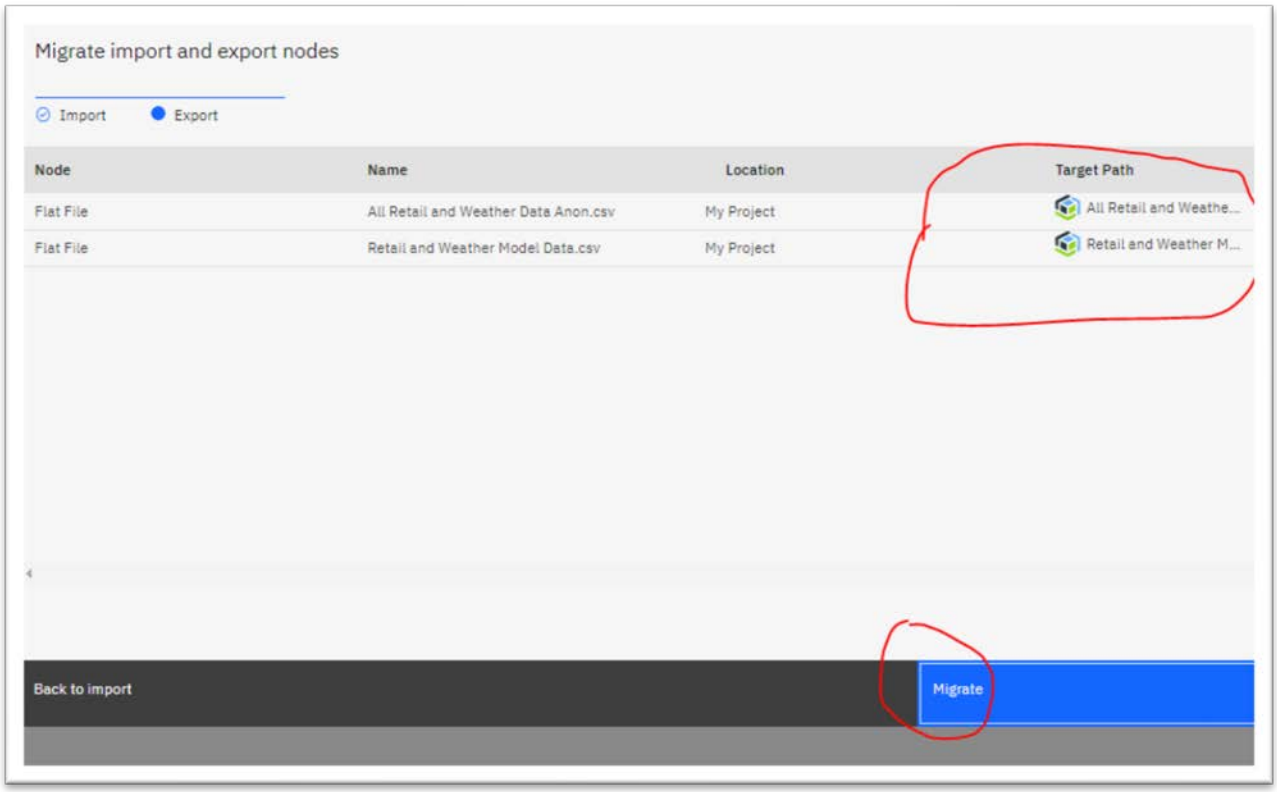

Migrate nodes?
To run this flow, you must migrate the Import and Export nodes to Data Asset nodes. You will need to associate data assets for import.
Export nodes can be configured to export to your project or to a connection.
Migrate all nodes now, or click No to migrate nodes individually at any time.
No
Yes


Migrate import and export nodes
Import
Export
Node
Name
Location
Asset name
Var. File
Historical Weather.csv
Project
Browse
Var. File
Historical Sales.csv
Project
Browse

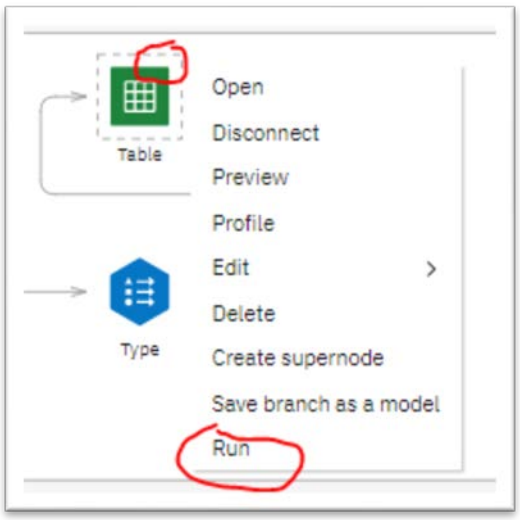

Migrate import and export nodes
Retail and Weather
Assets (2)
Connections
Data assets
Data assets
Data assets (6)
All Retail and Weather Data.csv
Historical Sales.csv
Historical Weather.csv
Retail And Weather README.htm
Retail And Weather README.pdf
Retail and Weather Model Data.csv



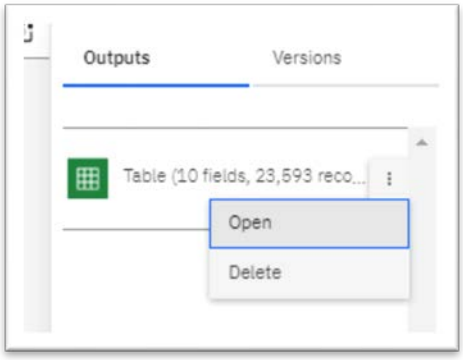
19) Seen here the mapping of the 2 export files used in the dashboard. Map *All Retail and Weather Data.csv* and *Retail and Weather Model Data.csv* to their respective match. Click **Migrate.**

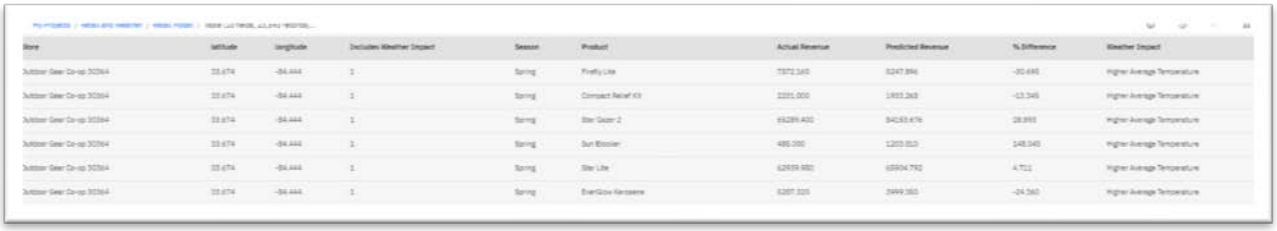


20) When complete, the stream will appear. Click on the **Table** on the far right and choose **Run.** This will ingest the 2 files called *Historical Weather.csv* and *Historical Sales.csv,* perform some data preparation, run the data through a Random Forecast Model, prepare the data bit more for analysis and export it to *Retail and Weather Data Model.csv* (the file used on the 1st tab of the dashboard)



21) Choose to **Open** the result. The field ***% difference*** represents the predicted uplift in the dashboard built off the value provided from the Random Forest model called Predicted Revenue by store, season, product, and weather condition.





For questions and inquiries please reach out to Brian Snyder, bsnyder@us.ibm.com