Hands on tutorial
A Regression Model in Azure ML

Tutorial #1 Questions:

- What are we trying to accomplish?
  - Given these dynamometer card readings, can we predict when the operation will fail?
  - Can we predict the Remaining Useful Life (RUL)?
- How can we use Azure Machine Learning to answer these questions?

Microsoft NEAL ANALYTICS

Scenario Data Overview :

well	date	idx	x	y
Alabama	8/16/2013 1:55	1	0.00074002	-5327.23698
Alabama	8/16/2013 1:55	2	0.15595834	-4476.758361
Alabama	8/16/2013 1:55	3	1.536036006	-3475.778052
Alabama	8/16/2013 1:55	4	3.779102308	-2636.548613
Alabama	8/16/2013 1:55	5	6.817062356	-1999.545532
Alabama	8/16/2013 1:55	6	10.7311654	-1622.668567
Alabama	8/16/2013 1:55	7	15.5308829	-1432.091513
Alabama	8/16/2013 1:55	8	21.09257393	-1346.745211
Alabama	8/16/2013 1:55	9	27.24284435	-1325.808833

- ~3.4 million rows of data
- 263 unique wells
- The data comprises of the following:
  - Date-Time at which the readings were taken
  - Idx - Reading number (1-60): During every up and down motion of the Pump Jack, 60 sets of reading will be taken at various positions
  - X - Position of the Pump Jack
  - Y – Load on the Pump Jack
- We then perform the Principal Component Analysis over this data which was covered during your Exploratory Data Analysis lecture

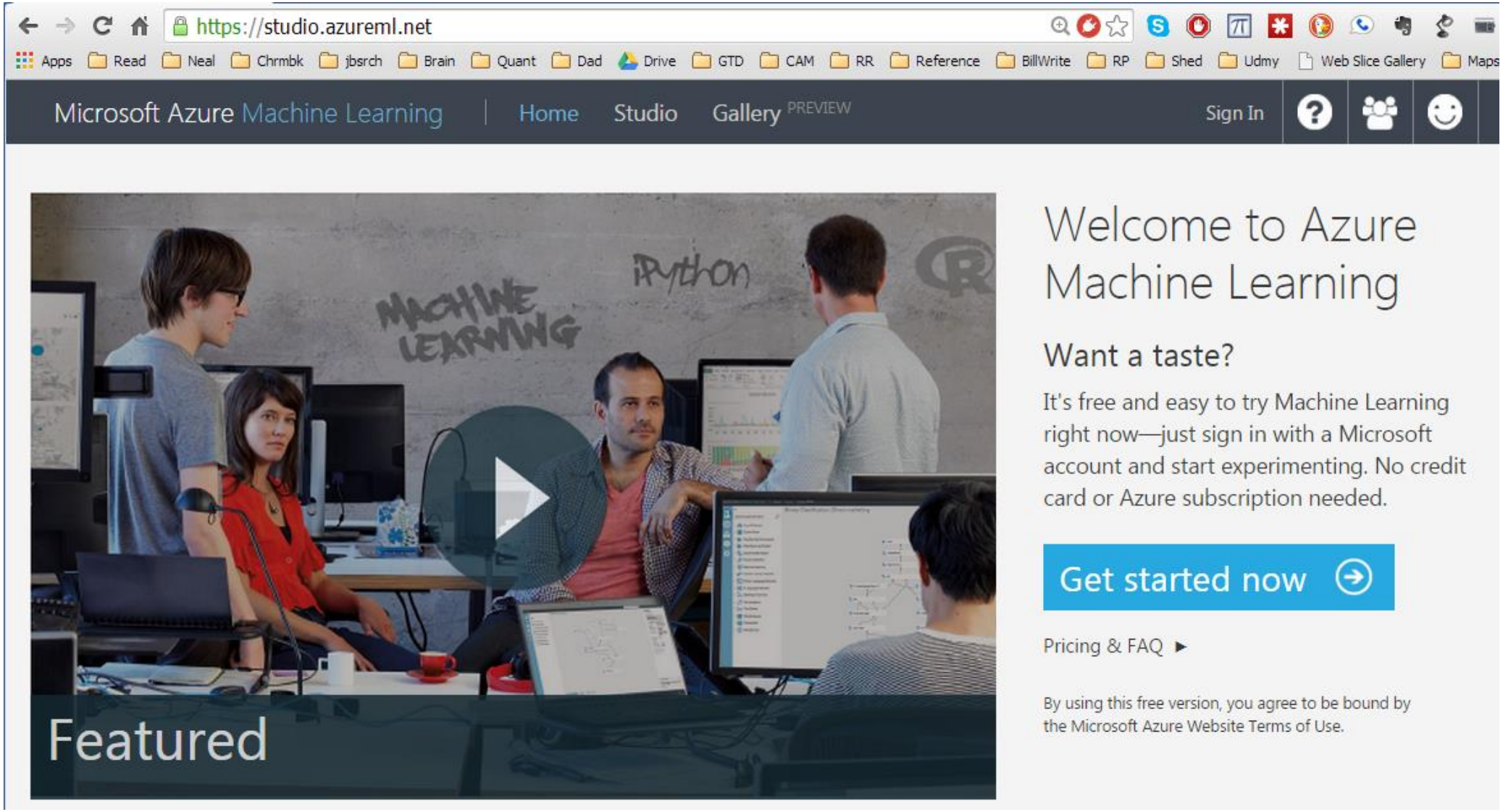
Scenario Data Overview:

- After the PCA was performed we ended up with~57K rows of data with 3 Principal Components (PC1, PC2, PC3)
- 263 unique wells
- Training data comprises of the following:
  - Date-Time at which the readings were taken
  - PC1
  - PC2
  - PC3
  - Assignments

well	date	PC1	PC2	PC3	Assignments	RUL
Alabama	8/16/2013 0:00	-8.243340971	2.918070697	2.023029116	35	191
Alabama	8/16/2013 0:00	-8.243340971	2.918070697	2.023029116	35	191
Alabama	9/3/2013 0:00	-8.495327766	2.311736767	2.154066737	35	173
Alabama	9/6/2013 0:00	-8.148383208	2.731807776	2.341956526	35	170
Alabama	9/6/2013 0:00	-8.148383208	2.731807776	2.341956526	35	170
Alabama	9/13/2013 0:00	-8.206877815	2.58376763	2.282912974	35	163
Alabama	9/13/2013 0:00	-8.206877815	2.58376763	2.282912974	35	163
Alabama	11/5/2013 0:00	-8.640142409	2.146508336	2.971266786	16	110
Alabama	11/5/2013 0:00	-8.640142409	2.146508336	2.971266786	16	110
Alabama	12/12/2013 0:00	-3.139291754	7.206502009	-2.257469677	34	73
Alabama	12/12/2013 0:00	-3.139291754	7.206502009	-2.257469677	34	73

- Response Variable - Remaining Useful Life (RUL), or Time to Failure (TTF)

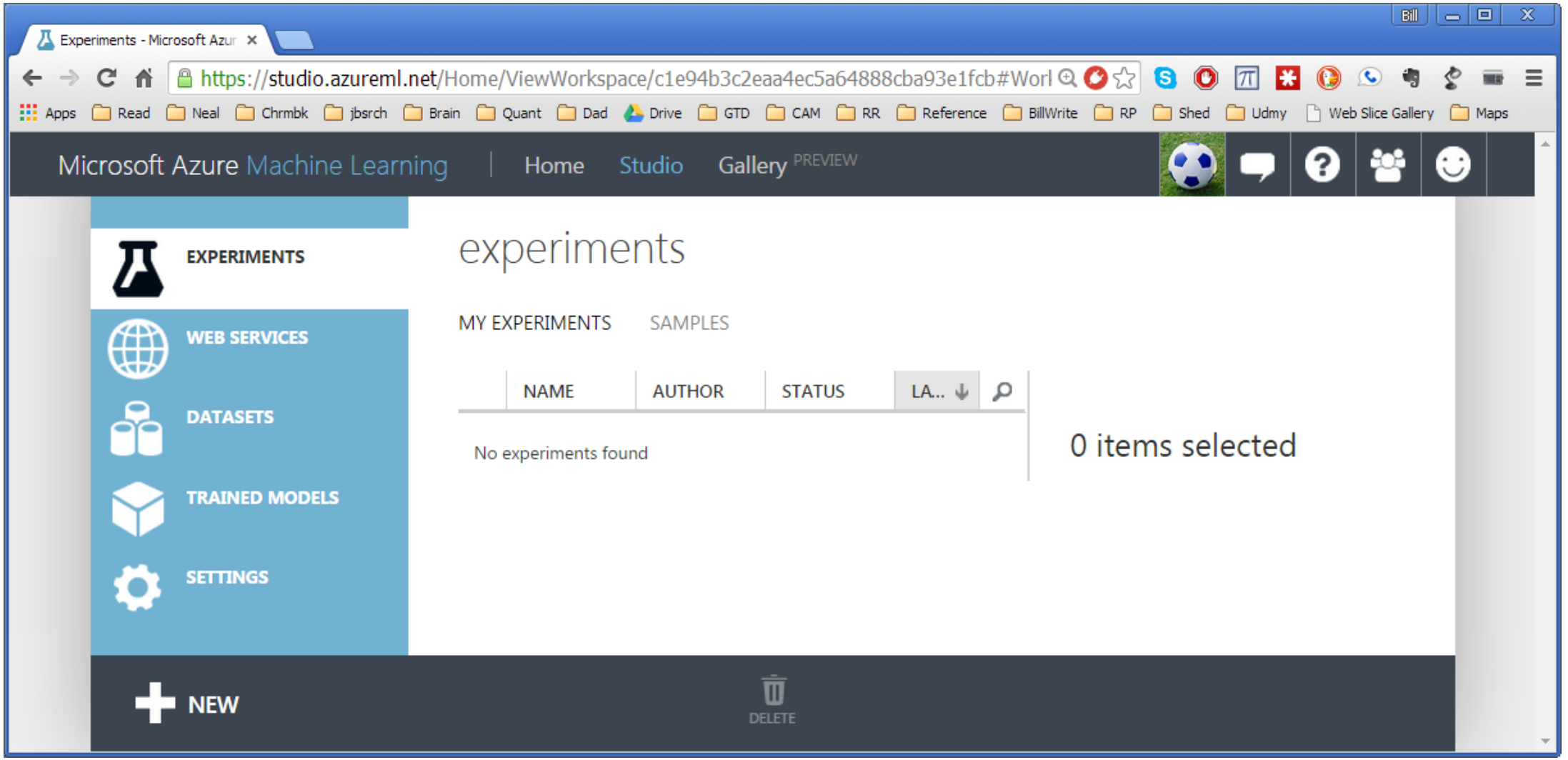
Step 1 : Go to https://studio.azureml.net/



Microsoft

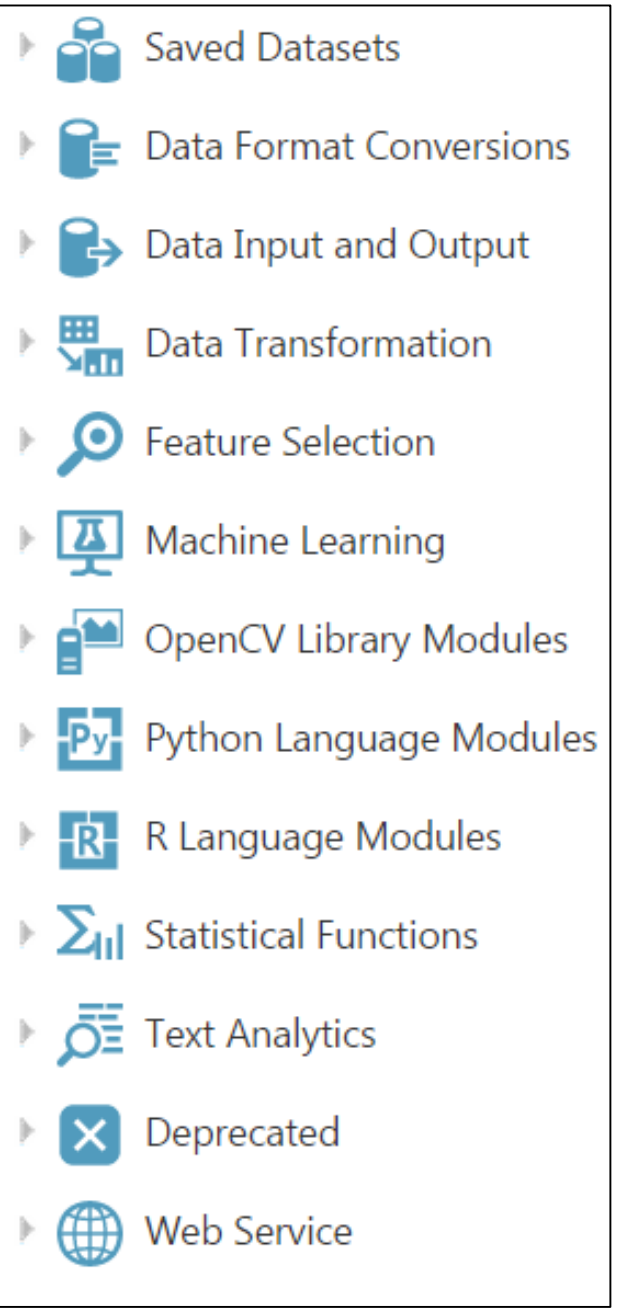
NEAL ANALYTICS

Step 2 : Log in to your account



Microsoft

NEAL ANALYTICS



Step 3.1 : Create New Experiment

- AML modelling … a checklist approach
  - ❑ Create new experiment

Microsoft NEAL ANALYTICS

Step 3.2 : Create a new experiment

- A word on modelling tradition and documentation
  - If you've worked with people building models you may have had a taste of several different modelling traditions

Mathematical people	Engineering people	Statistics people	Machine learning people
• "Everything is topological algebra" • Math tradition is to "do" modelling in equations in math environments (ex. Matlab) creating files organized (or not) in folders on hard drives	• "Everything is equations and programs" • Engineering tradition is to "do" modelling starting from equations and using programs organized in folders on hard drives.	• "I use SAS" or "I use SPSS" or "I use R" • Statistical tradition is to organize models in data and program files on computers.	• Machine learning people organize models in data and program files on computers.

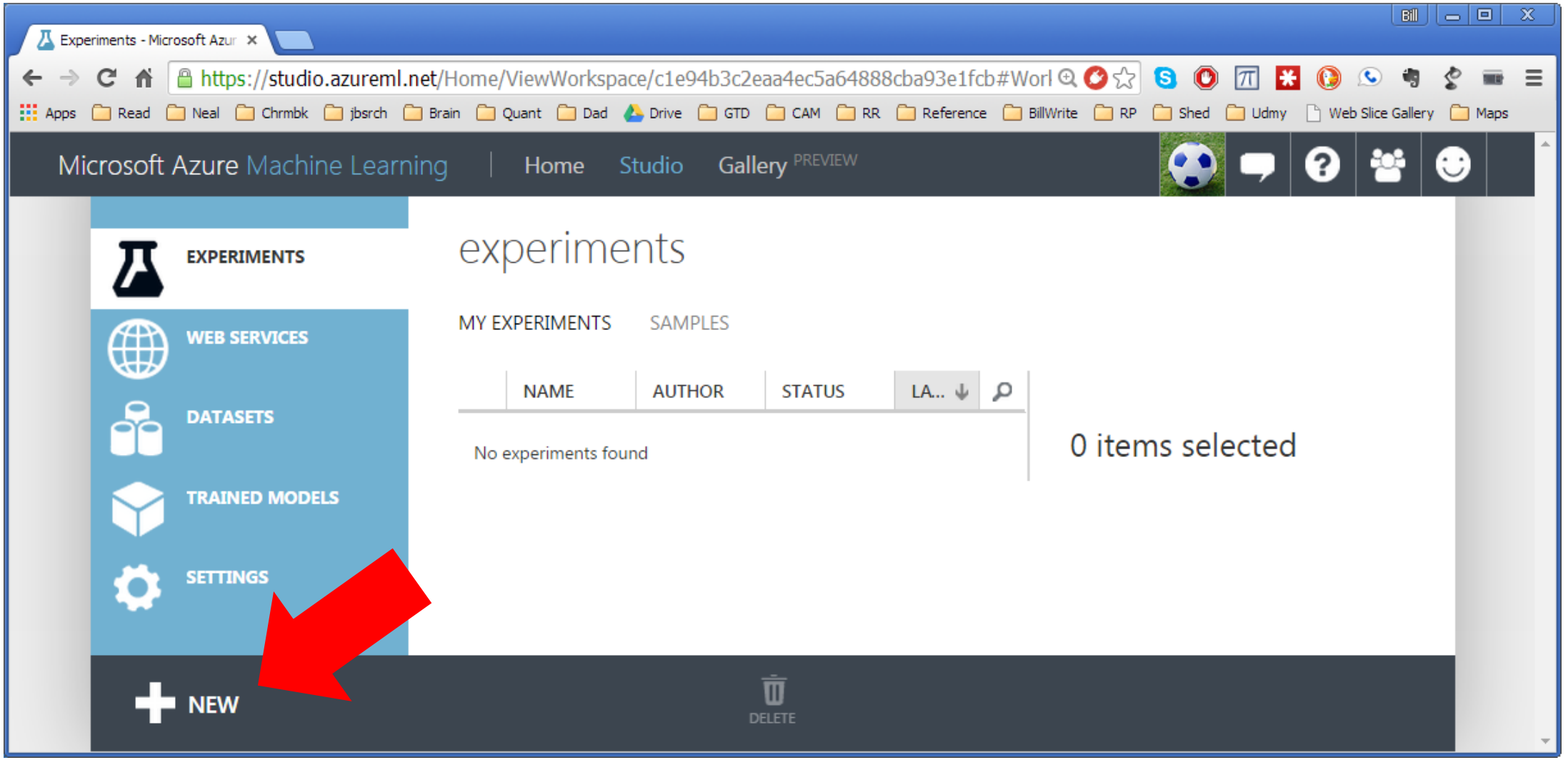
Step 3.3 : Create a new experiment (Recap)

Azure Machine Learning people "Everything is drag and drop"

Models are organized inside "experiments" that pull together data and programs into step by step documents.

- So, to "do" a model in Azure Machine Learning we create an experiment … document
  - Experiments are stored in the Azure Cloud
    - All your work is stored neatly in one place
    - Azure does not allow you to "scatter" files in random folders
  - Experiments neatly organize all the data and technical components of your modelling work
- Why does this matter?
  - Azure Machine Learning is an end-user tool
  - A "graphical user interface for modelling"

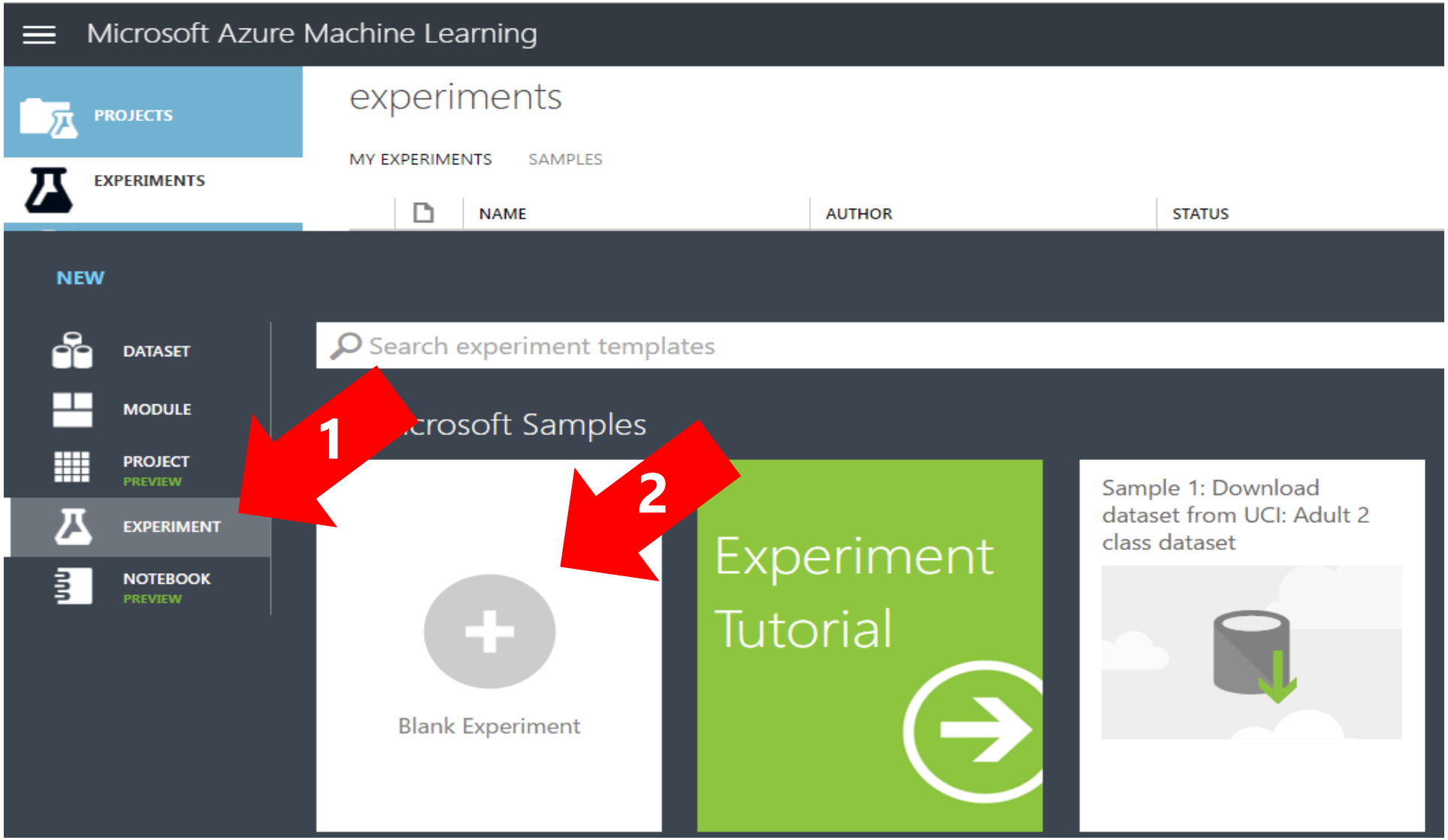
Step 3.4 : Create a new experiment



Microsoft

NEAL ANALYTICS

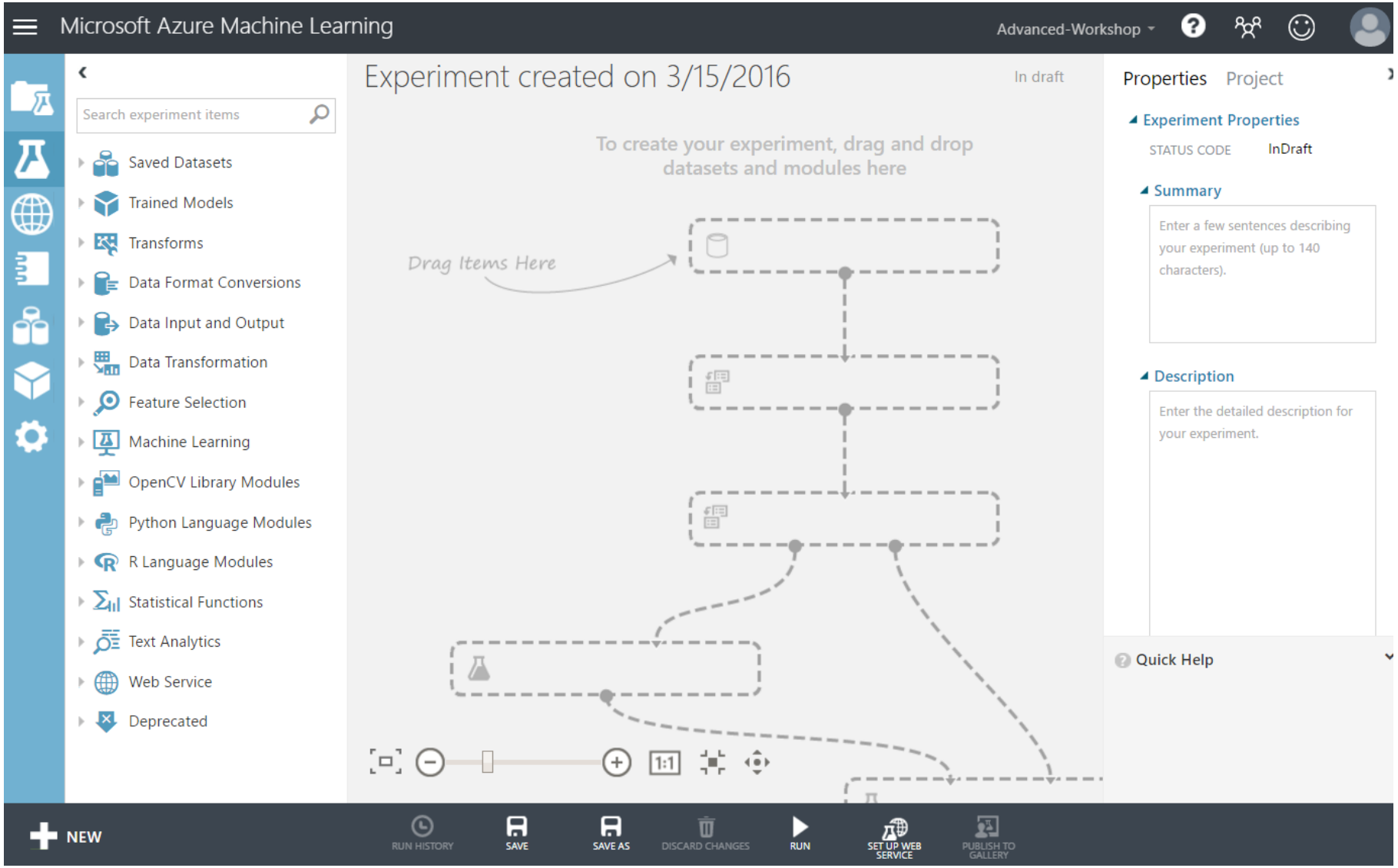
Step 3.5 : Create a new experiment



Microsoft

NEAL ANALYTICS

Step 3.6 : Slow down and look at the experiment



Look but **do not touch … yet**

As soon as we drag the first building block to the experiment …

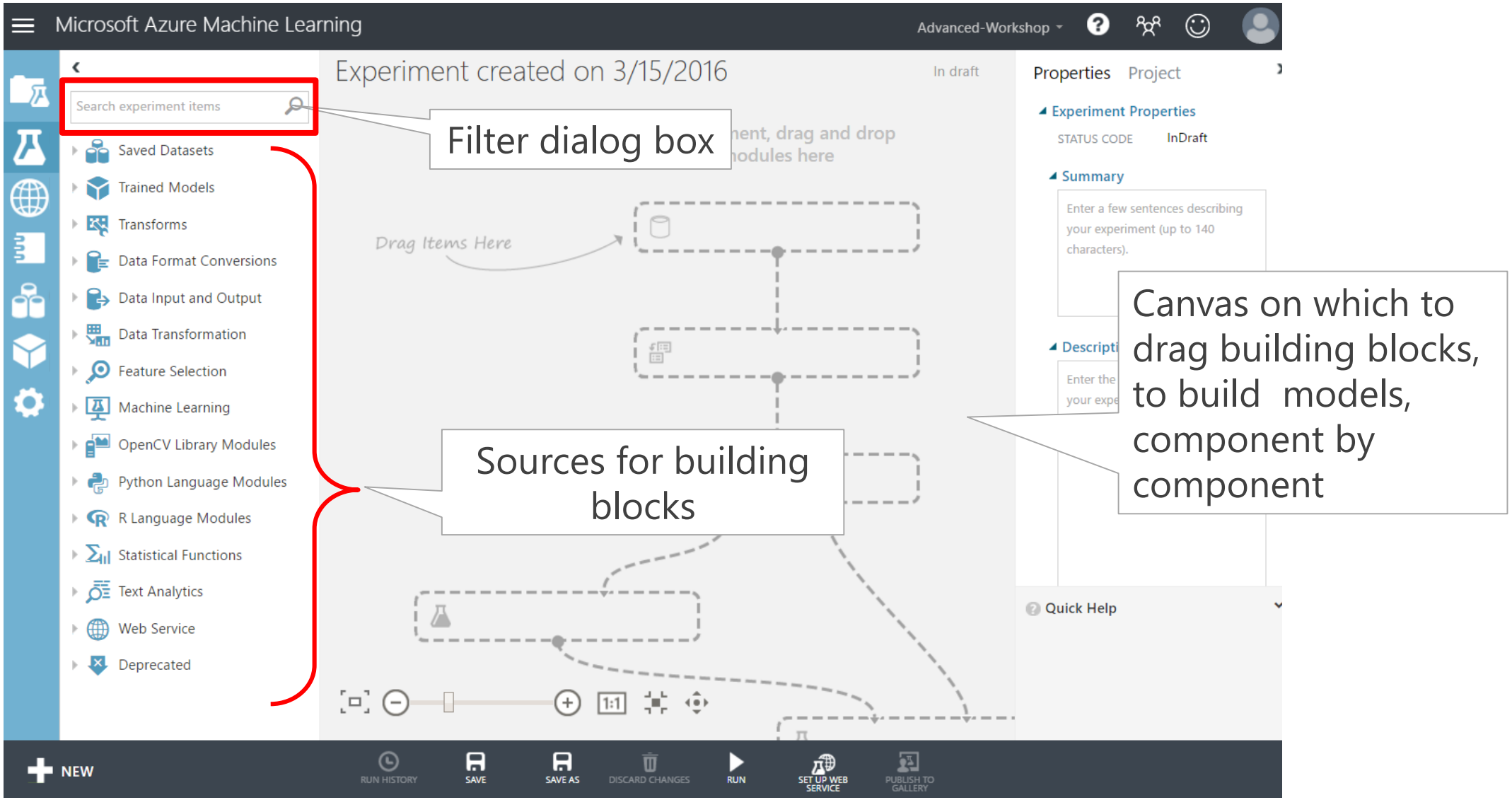
All the guides disappear

So let's look them over now

Microsoft

NEAL ANALYTICS

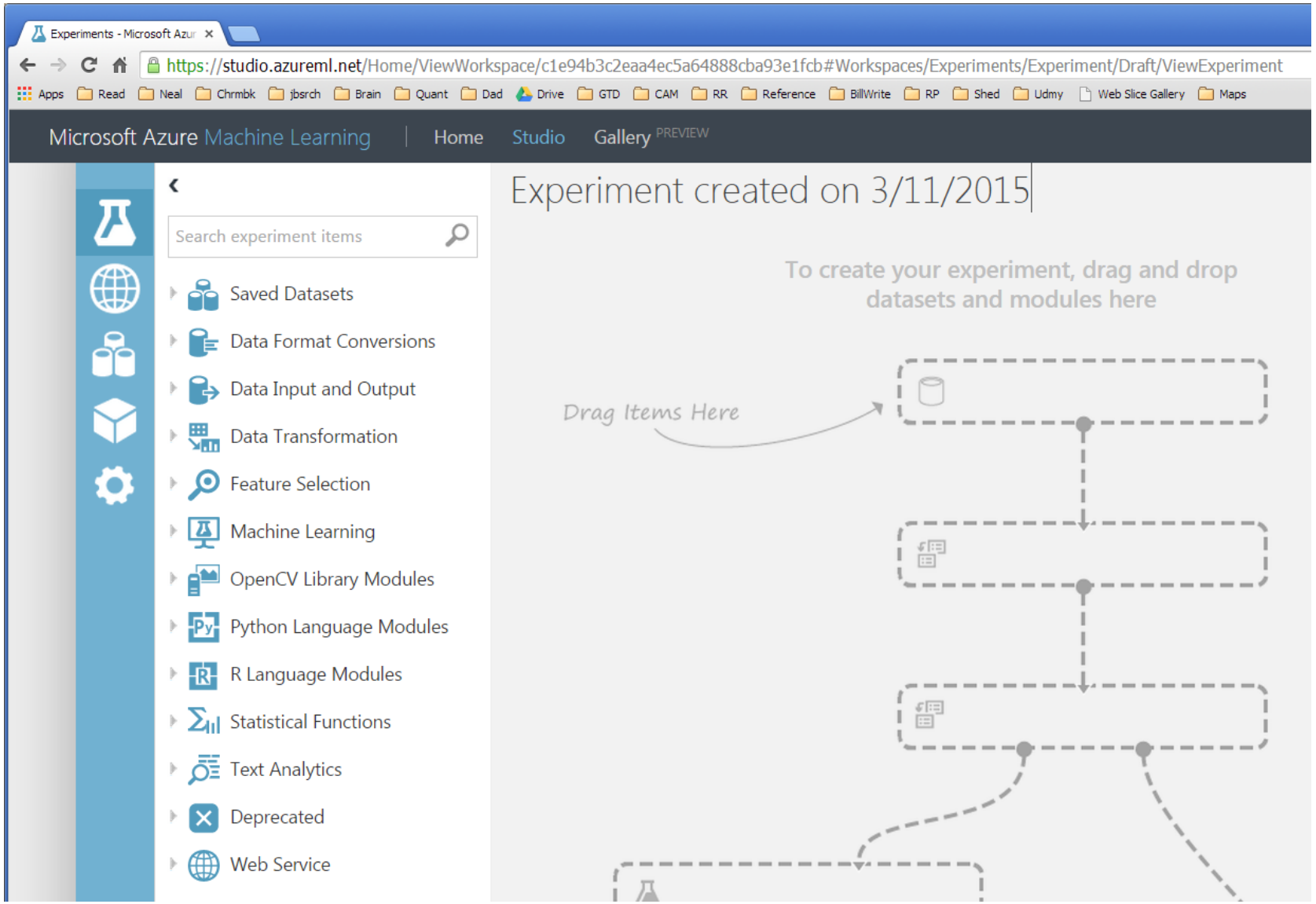
Step 3.7 : Slow down and look at the experiment



Microsoft

NEAL ANALYTICS

Step 3.8 : This is a “drag and drop” environment



Azure Machine Learning is to modelling as …

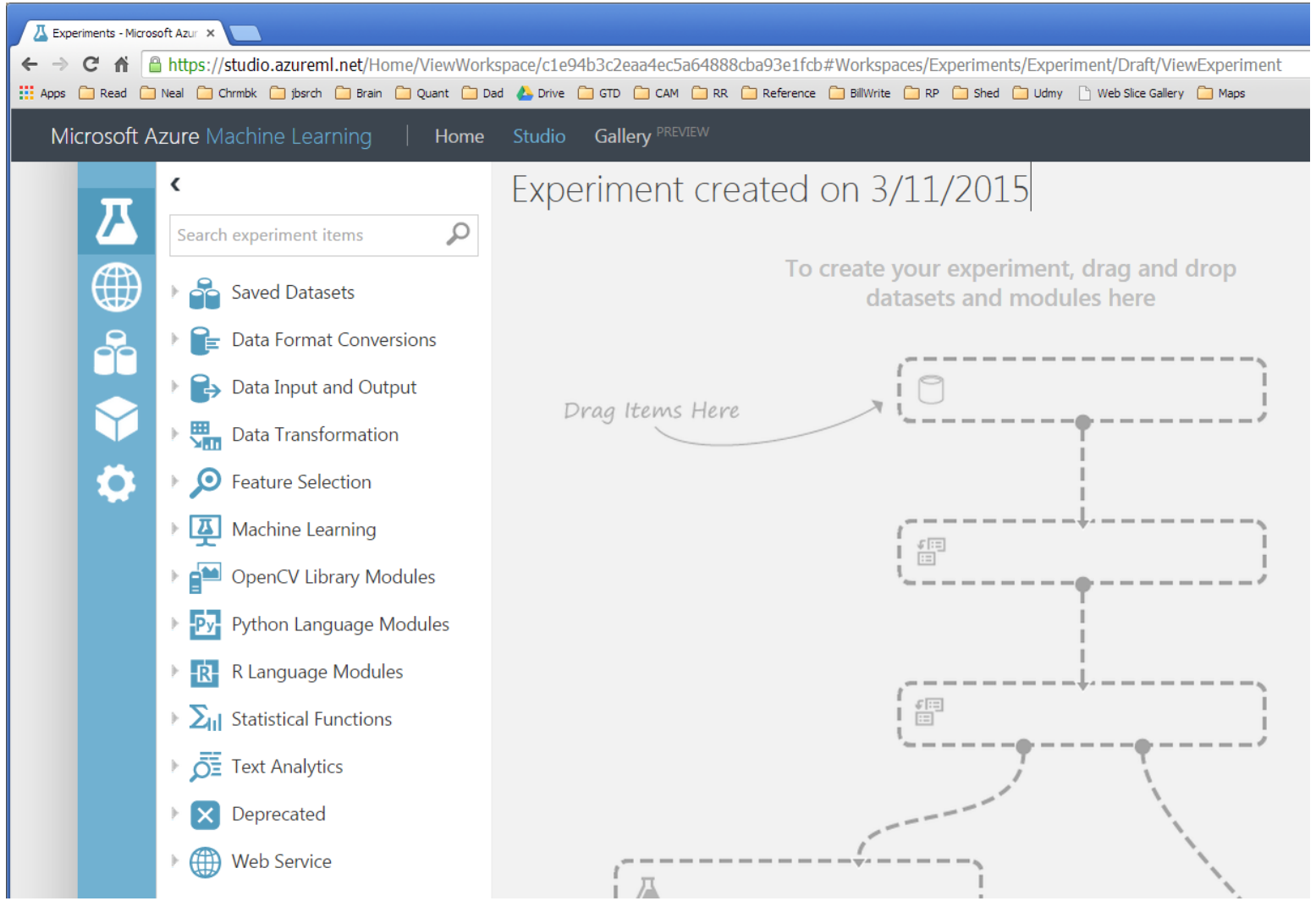
Inserting shapes into PowerPoint is to drawing

= Modelling by dragging and dropping

Microsoft

NEAL ANALYTICS

Step 3.9 : End users can do this …



Azure Machine Learning is a *disruptive* (easy to use) modelling environment

Try BEFORE you buy … and you …

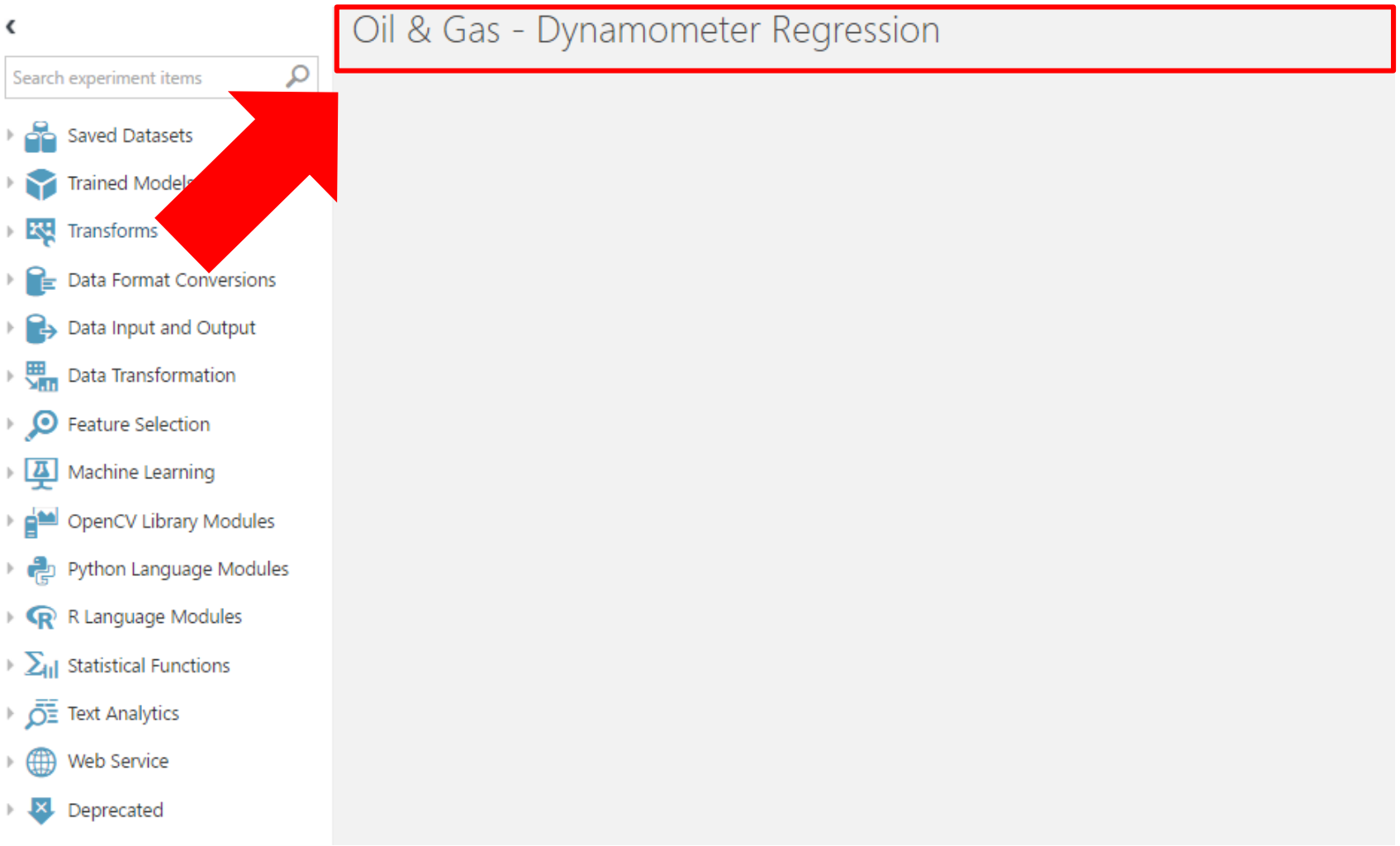
- Ask forgiveness … not permission

Never have to buy

- "Rent" instead

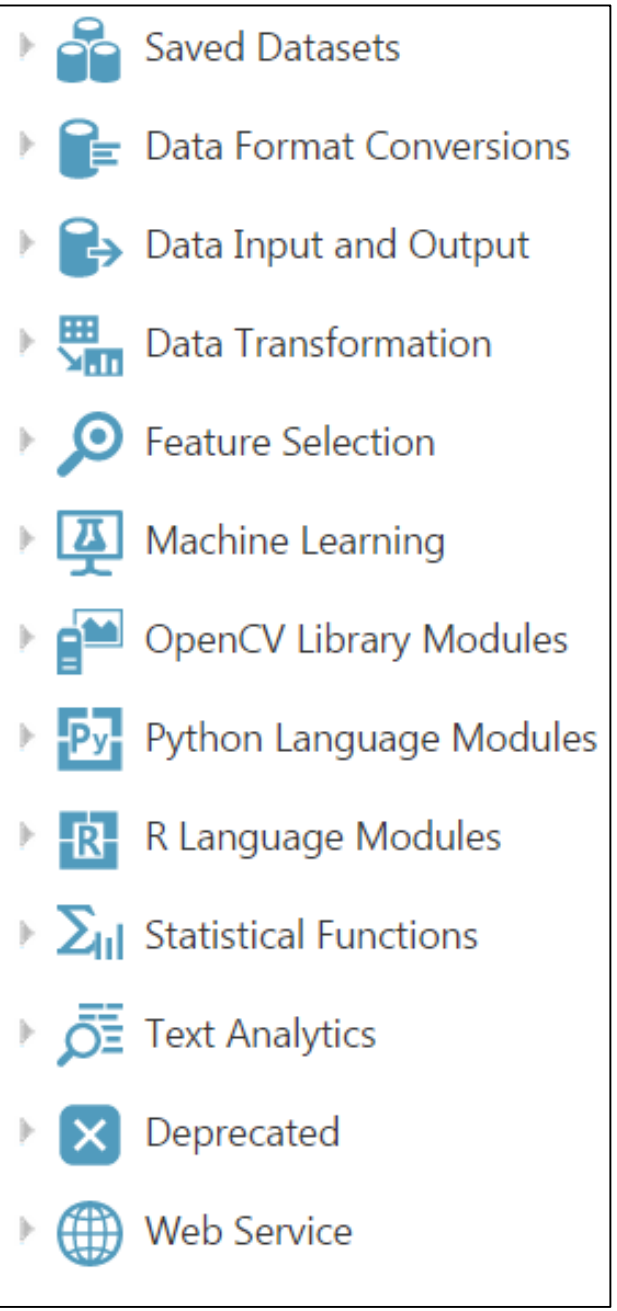
Microsoft NEAL ANALYTICS

Step 3.10 : Name your experiment



- Click on the title box at the top that says "Experiment Created on ...."
- Give the experiment the following title: "Oil & Gas - Dynamometer Regression"

Microsoft NEAL ANALYTICS



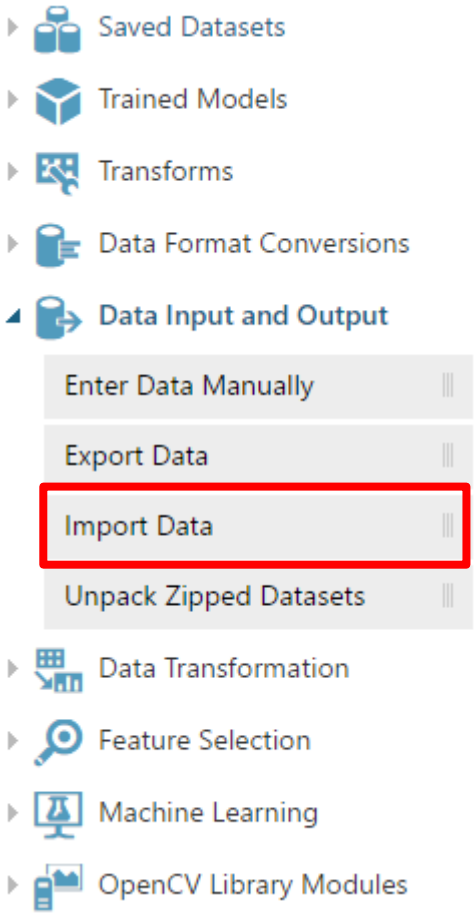
Step 4.1 : Import the dataset

- AML modelling … a checklist approach
  - ☑ ~~Create new experiment~~
  - ☐ Import data set

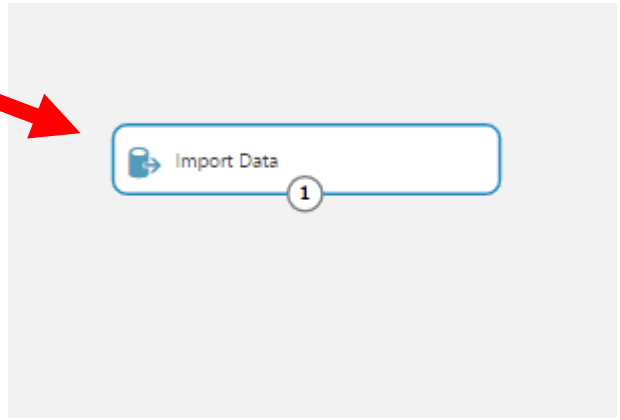
Microsoft NEAL ANALYTICS

Option 1 - Import the training dataset from the Azure Blob Storage. If option 1 did not work, skip to option 2.

Step 4.2 : Q? How to import data from Azure Blob?

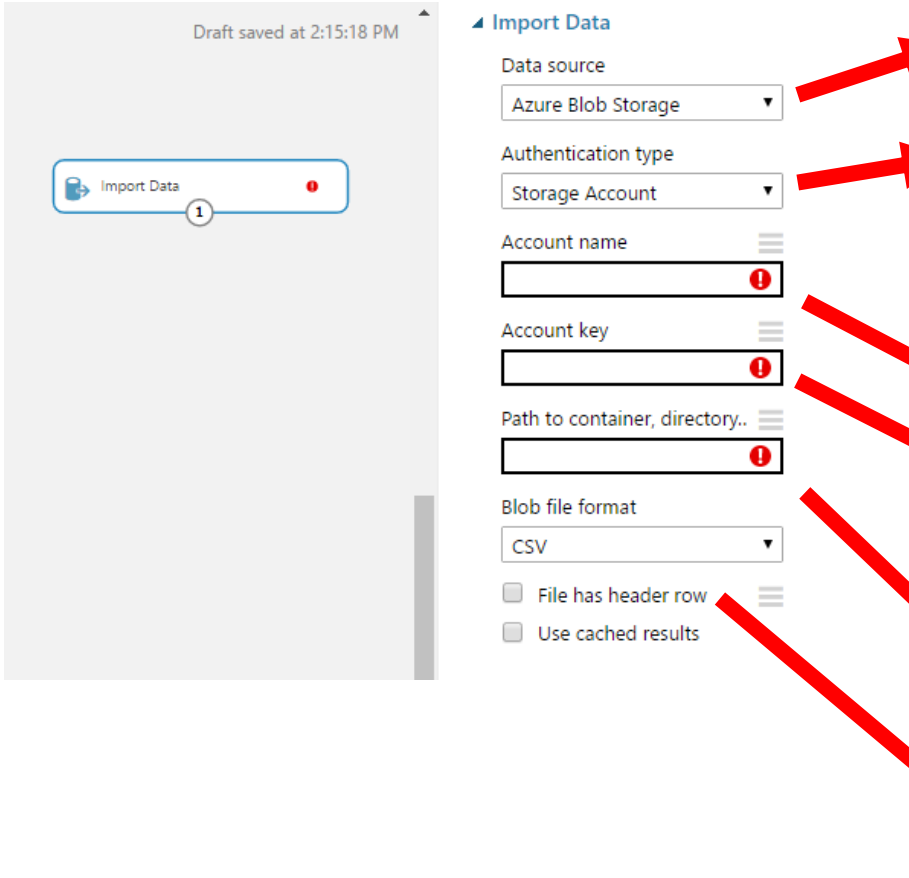


- Open “Data Input and Output” from the navigation pane at the left
- Drag “Import Data” to the canvas
  - “Import Data” loads data from sources such as the Web, Azure SQL, Windows Azure Blob storage, etc



Microsoft NEAL ANALYTICS

Step 4.3 : Q? How to import data from Azure Blob?



- Machine Learning
  - Click on Import Data
  - Chose "Azure Blob Storage" for Data source
  - Chose "Storage Account" for Authentication type
  - Copy and Paste the following information without the quotes in the Import Data module
    - Account Name – "**nealworkshop**"
    - Account key – "**RER9c7kfM1e67p7p7gI+TbkE5Y6aIzURg4PQc9lew+l8O+ZfU58gFjNgBW/WQm0u8N0YZQUG+wIaIzfWKxyIjA==**"
    - Path to container – **"/oilandgas/Dynamometer Card Reading Data RUL.csv**"
    - Check the File has header row check box

Microsoft NEAL ANALYTICS

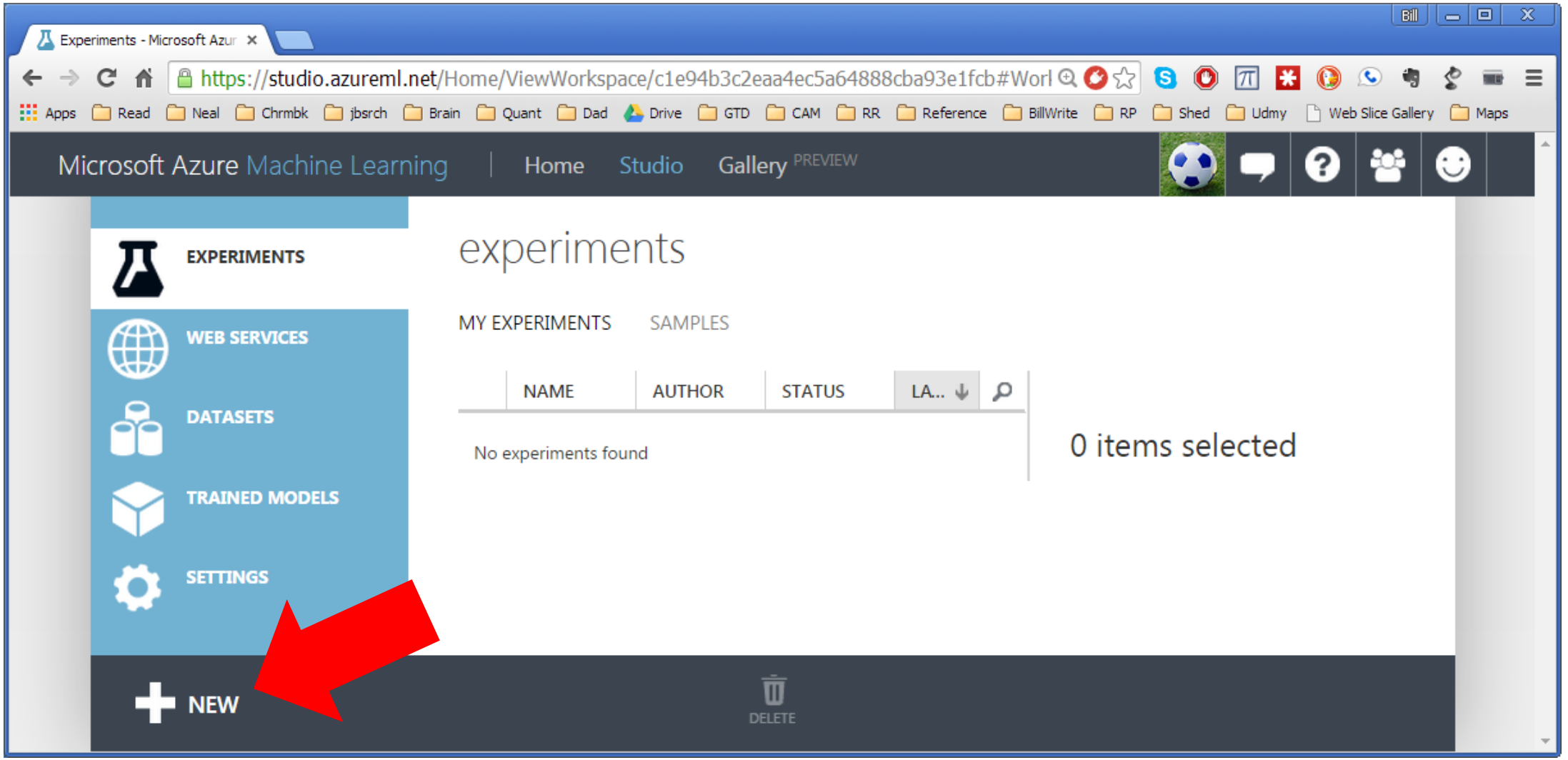
Step 4.4 : Q? How to import data from Azure Blob?

Microsoft NEAL ANALYTICS

If Option 1 worked, move on to Step 4.14

Option 2 - Import the training dataset from the saved datasets.

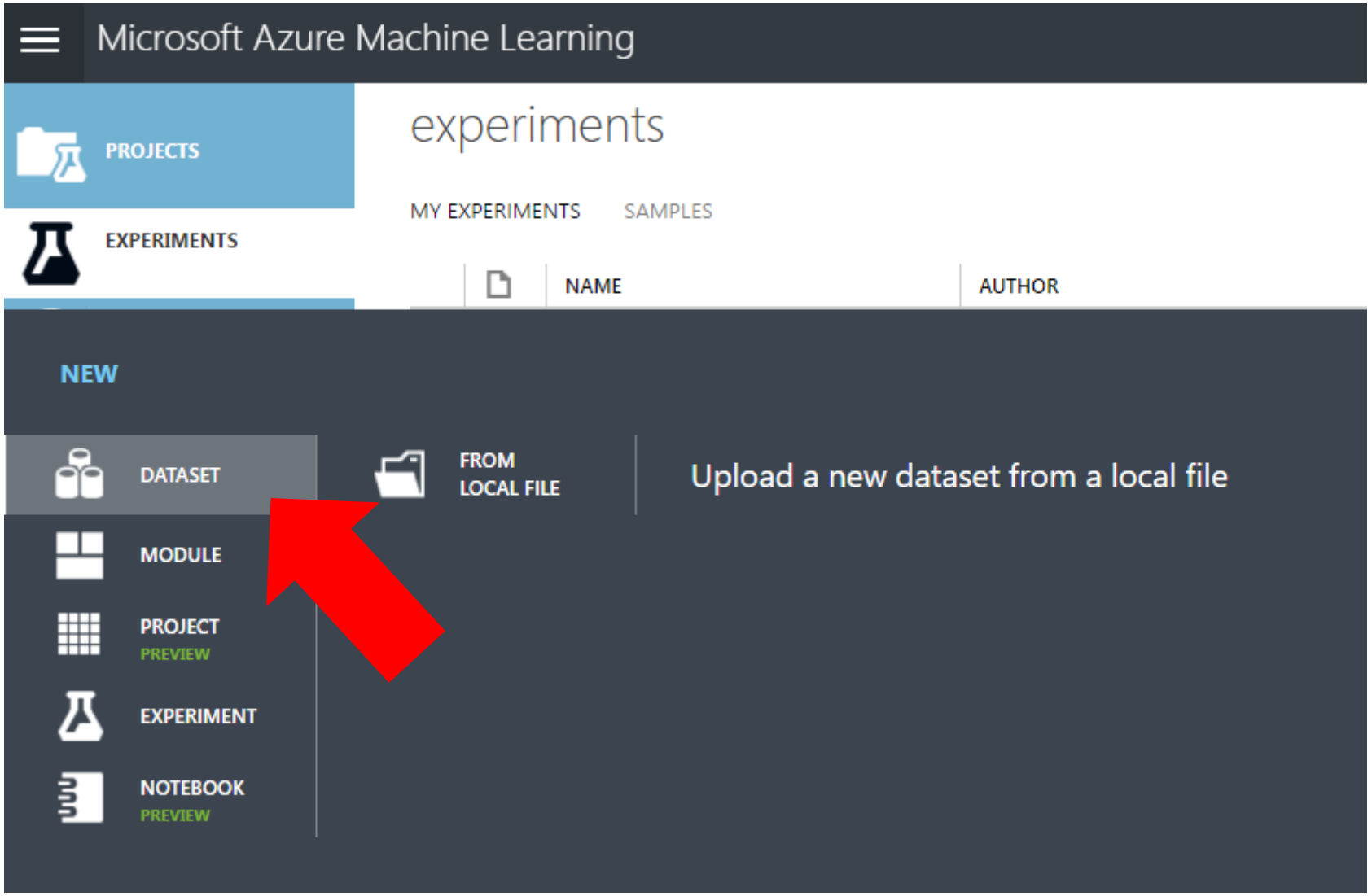
Step 4.5 : Import this tutorial's training dataset



Microsoft

NEAL ANALYTICS

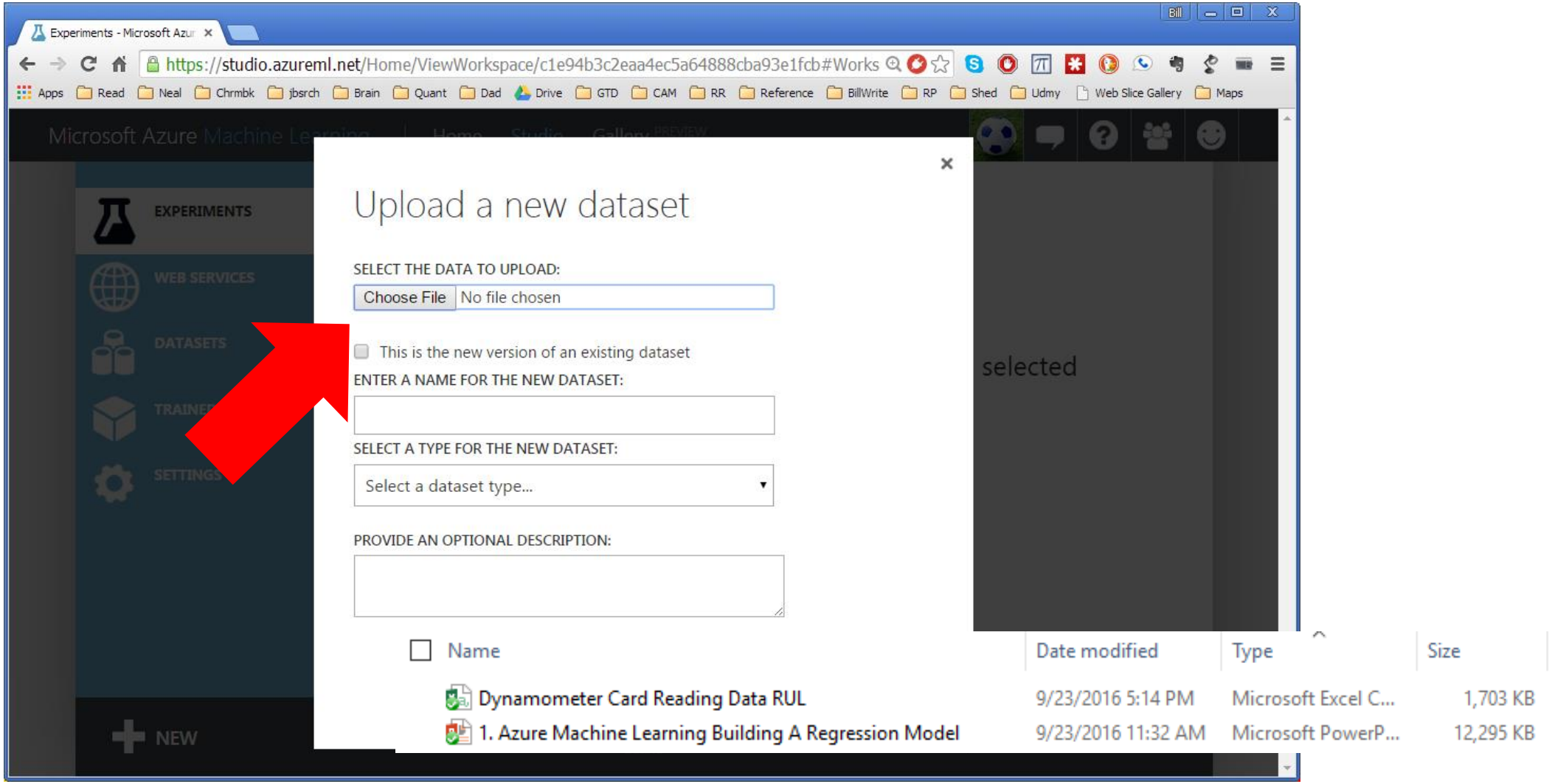
Step 4.6 : Import this tutorial's training dataset



Microsoft

NEAL ANALYTICS

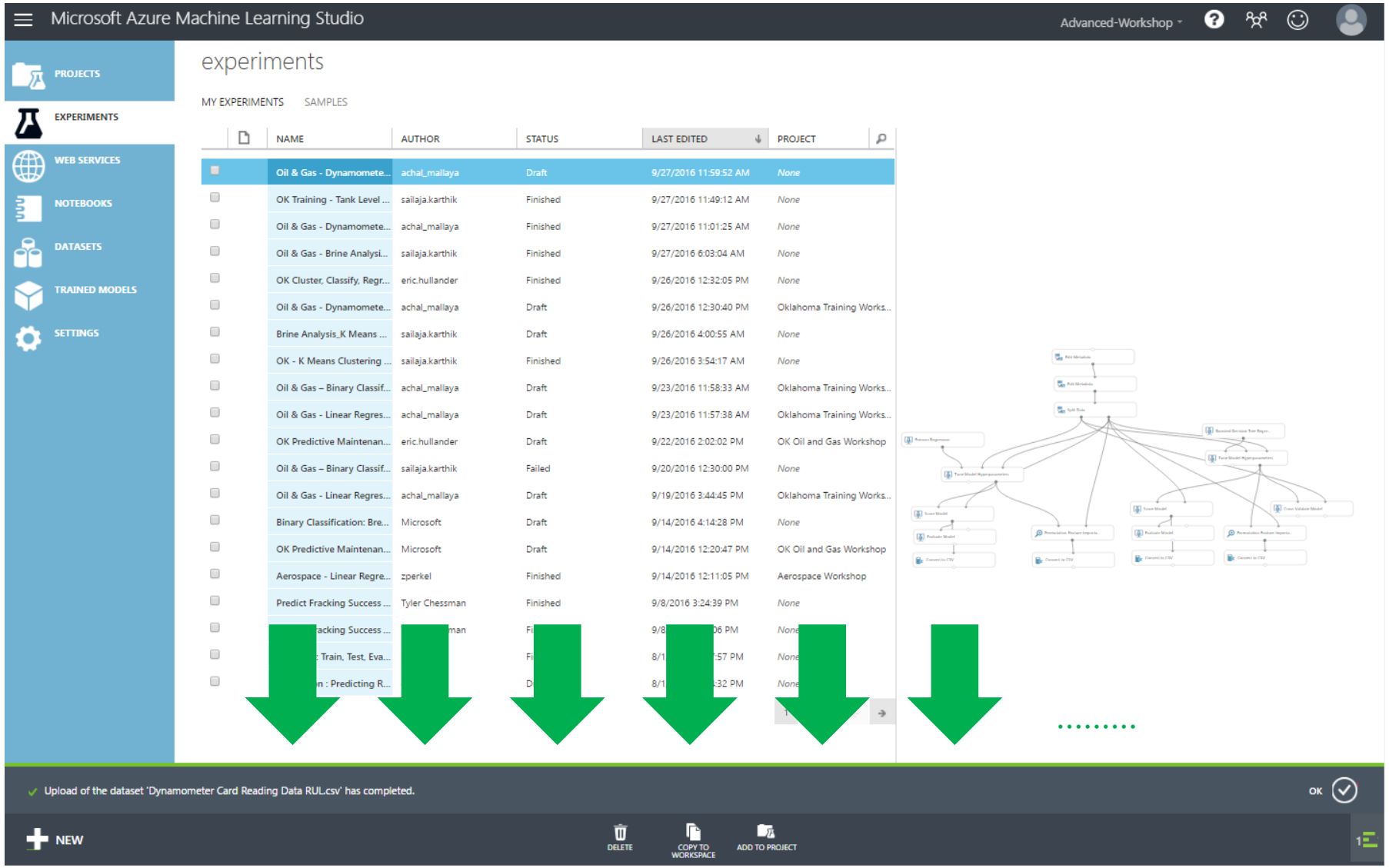
Step 4.7 : Import this tutorial's training dataset



Microsoft

NEAL ANALYTICS

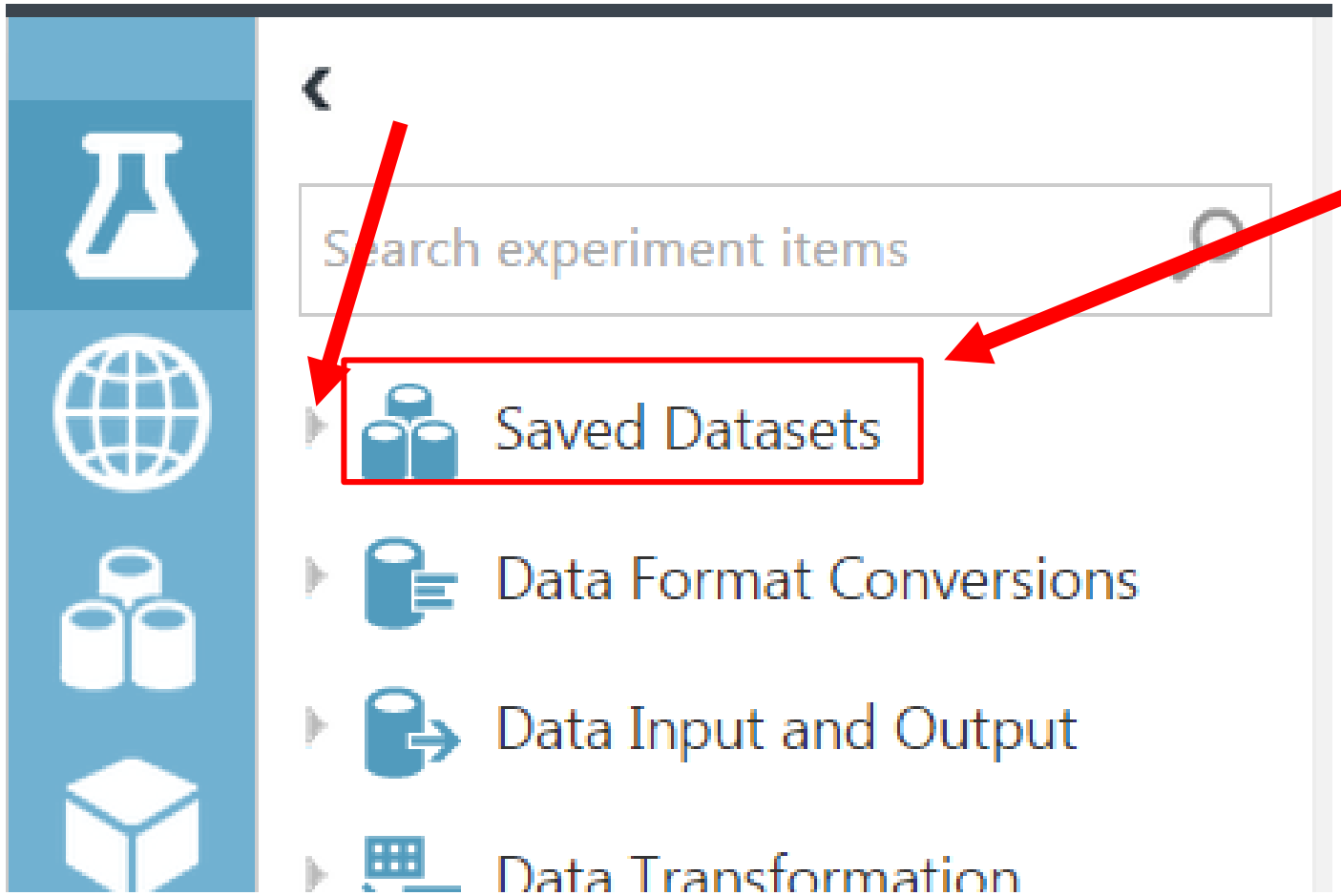
Step 4.8 : Import this tutorial's training dataset



Microsoft

NEAL ANALYTICS

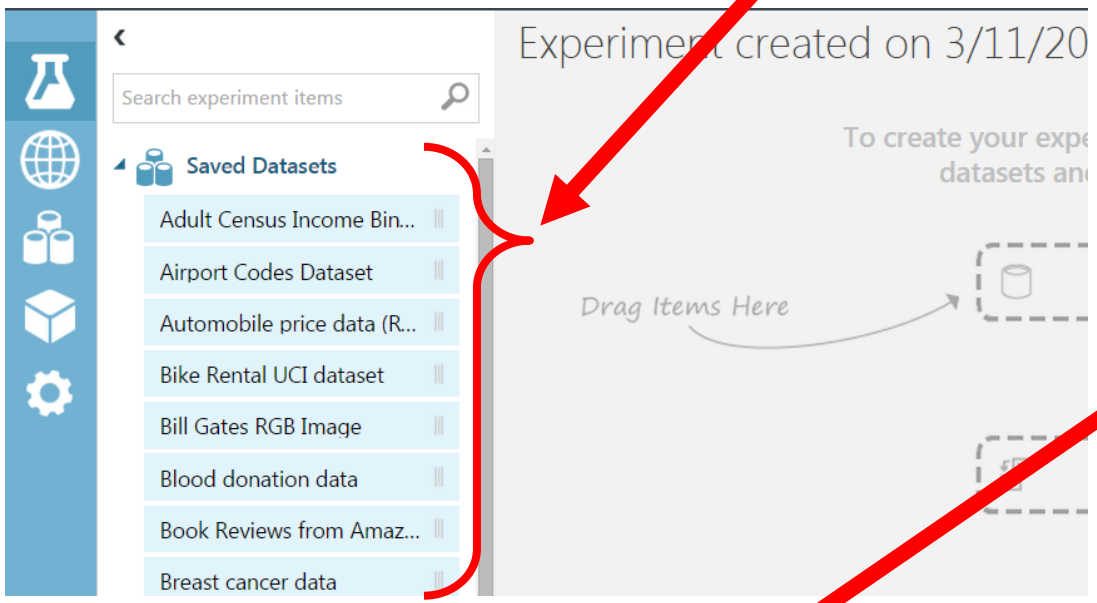
Step 4.9 : Open “Saved Datasets”



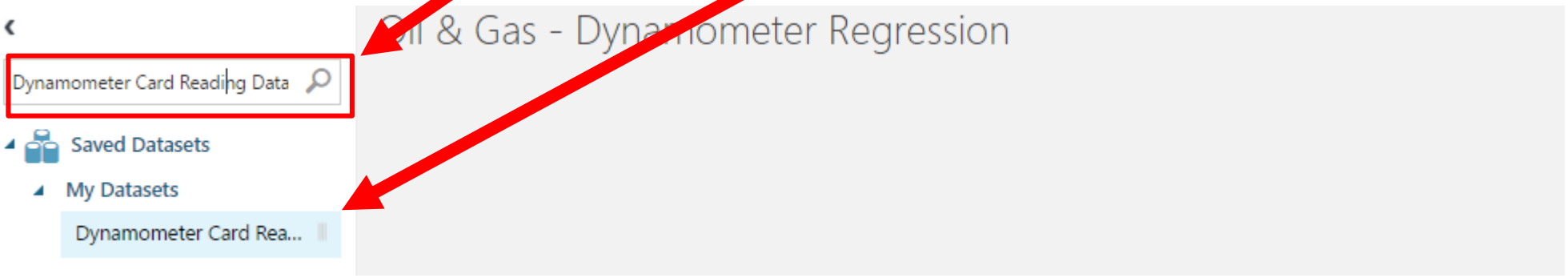
- By clicking on the triangle at the left
of “Saved Datasets”

Microsoft NEAL ANALYTICS

Step 4.10 : Take a second to notice the MANY datasets



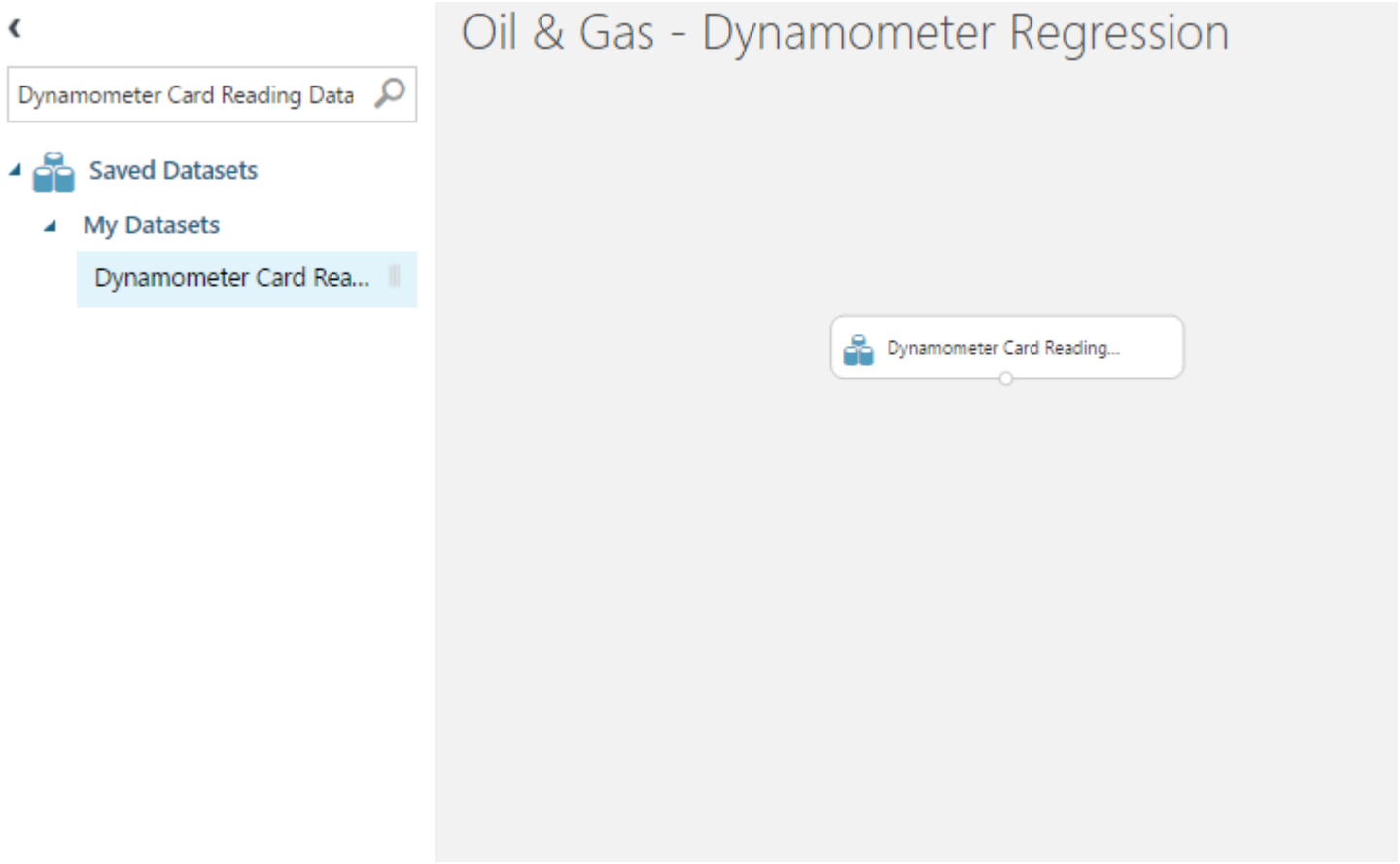
- To filter to the data set for this tutorial …
- Type “Dynamometer Card Reading Data RUL” in the “Search experiment items” dialog box
- The data set list will reduce to our data set for this tutorial



Microsoft

NEAL ANALYTICS

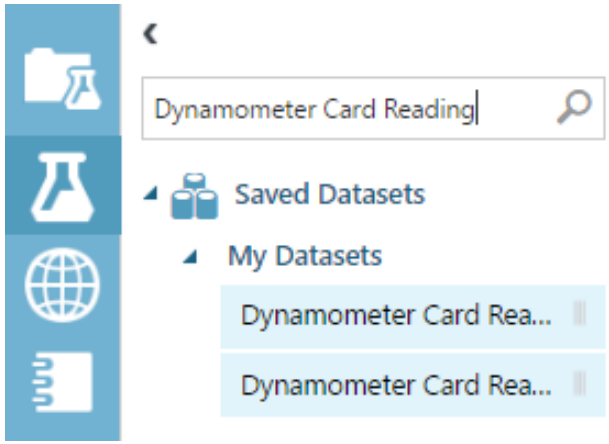
Step 4.11 : Drag the data set to the experiment



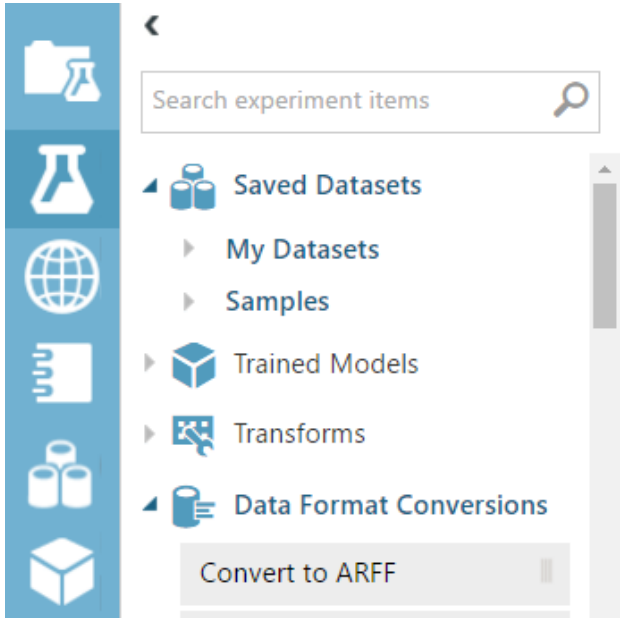
- *Note* when you drag the first element of your model to the canvas … all the guides disappear
- Now, where are all the tools that were at the left Azure Machine Learning?
- They are still there, … but we need to un-filter to see them

Microsoft NEAL ANALYTICS

Step 4.12 : Backspace over “Predictive”



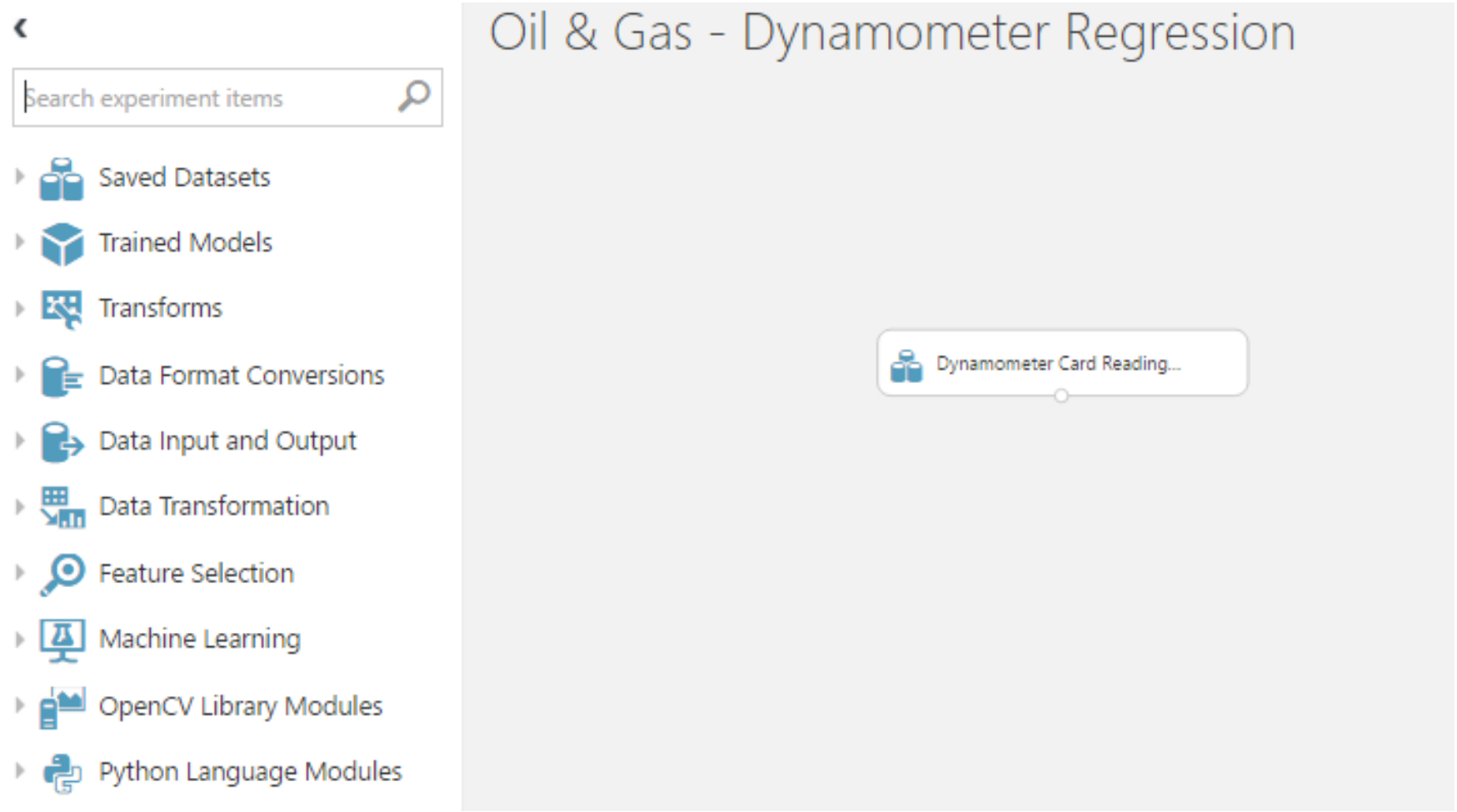
- Backspace to remove “Dynamometer Card Reading Data RUL” from the “Search experiment items” dialog



- Then click the triangle at left of “Saved Datasets” to close the dataset list

Microsoft NEAL ANALYTICS

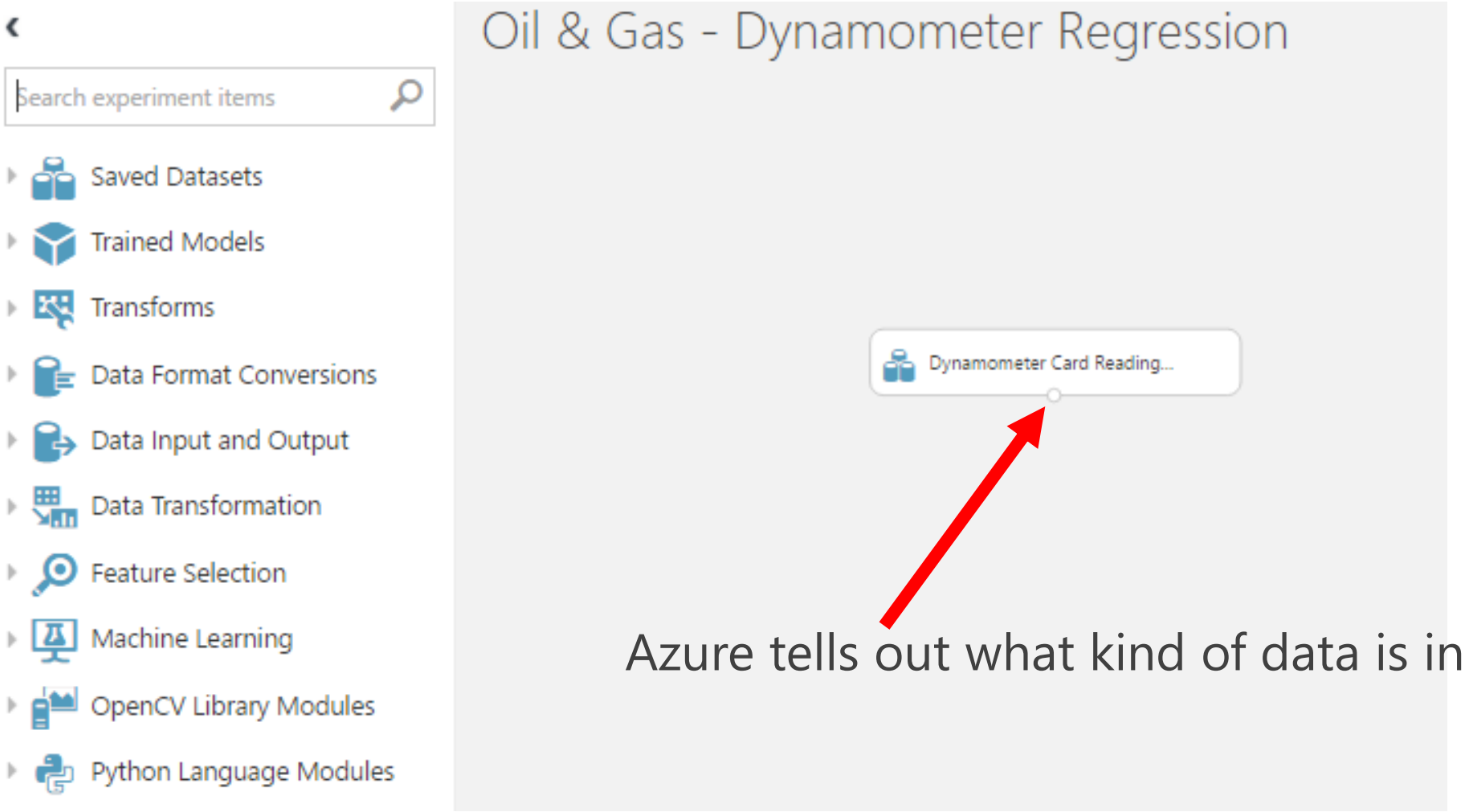
Step 4.13 : Admire your data set living in your experiment



Microsoft

NEAL ANALYTICS

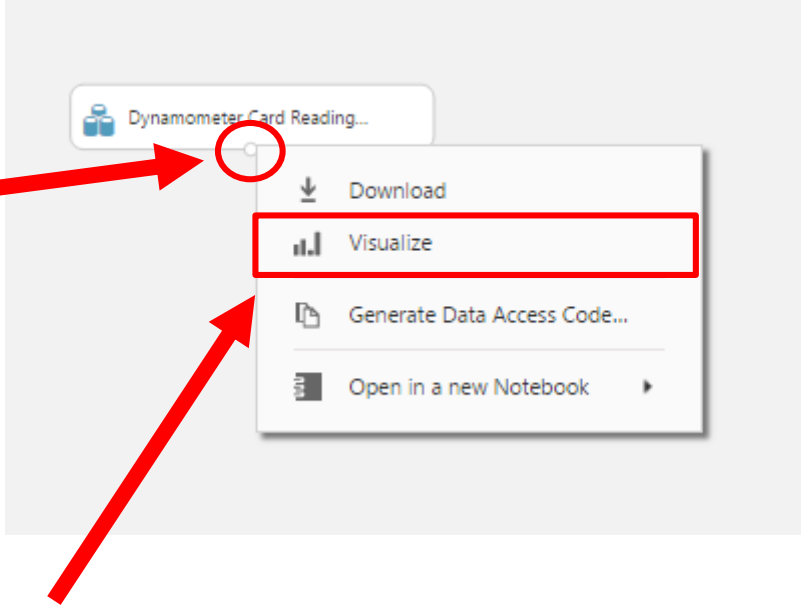
Step 4.14 : Hover your mouse over the bottom-middle circle on the data set



Microsoft NEAL ANALYTICS

Step 4.15 : Visualize the data set

Right click on the bottom-middle circle of the module

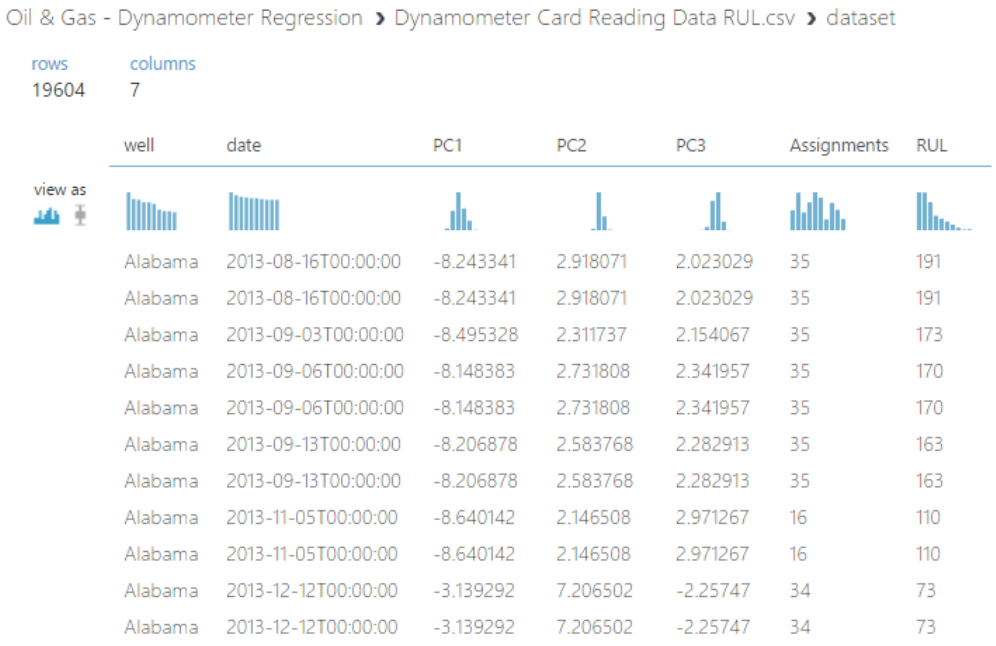


Then click “Visualize”

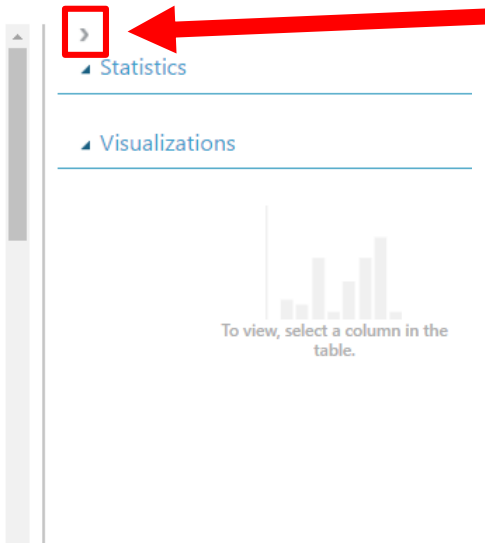
Microsoft

NEAL ANALYTICS

Step 4.16 : Now adjust the window so you can see all columns of the data



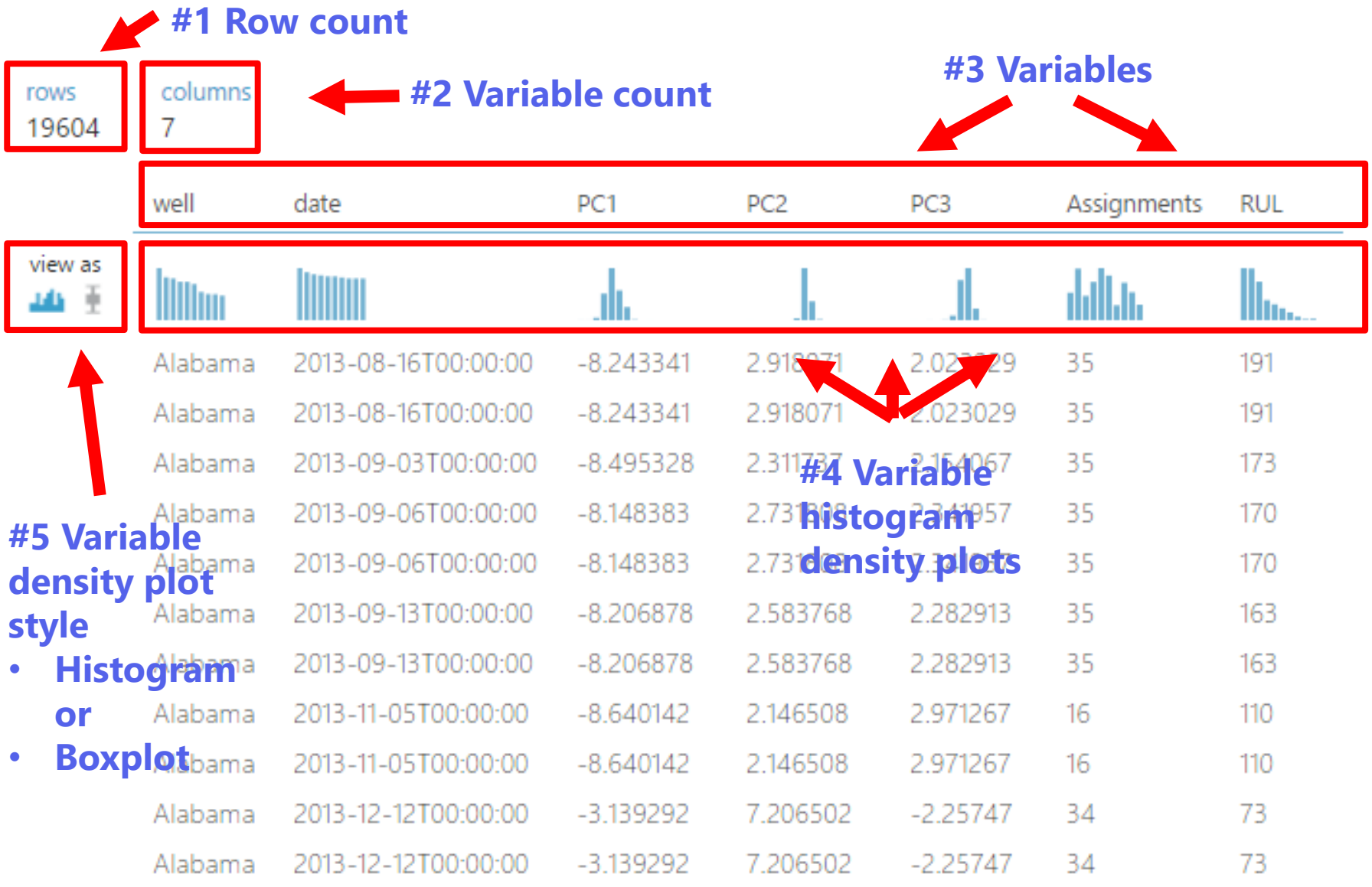
well	date	PC1	PC2	PC3	Assignments	RUL
Alabama	2013-08-16T00:00:00	-8.243341	2.918071	2.023029	35	191
Alabama	2013-08-16T00:00:00	-8.243341	2.918071	2.023029	35	191
Alabama	2013-09-03T00:00:00	-8.495328	2.311737	2.154067	35	173
Alabama	2013-09-06T00:00:00	-8.148383	2.731808	2.341957	35	170
Alabama	2013-09-06T00:00:00	-8.148383	2.731808	2.341957	35	170
Alabama	2013-09-13T00:00:00	-8.206878	2.583768	2.282913	35	163
Alabama	2013-09-13T00:00:00	-8.206878	2.583768	2.282913	35	163
Alabama	2013-11-05T00:00:00	-8.640142	2.146508	2.971267	16	110
Alabama	2013-11-05T00:00:00	-8.640142	2.146508	2.971267	16	110
Alabama	2013-12-12T00:00:00	-3.139292	7.206502	-2.25747	34	73
Alabama	2013-12-12T00:00:00	-3.139292	7.206502	-2.25747	34	73



- Click the “close” arrow
- Don’t worry about “losing” the visualization window.
  - Click on any variable name and the window re-opens.

Microsoft NEAL ANALYTICS

Step 4.17 : Now look at the data for 4 attributes



Microsoft NEAL ANALYTICS

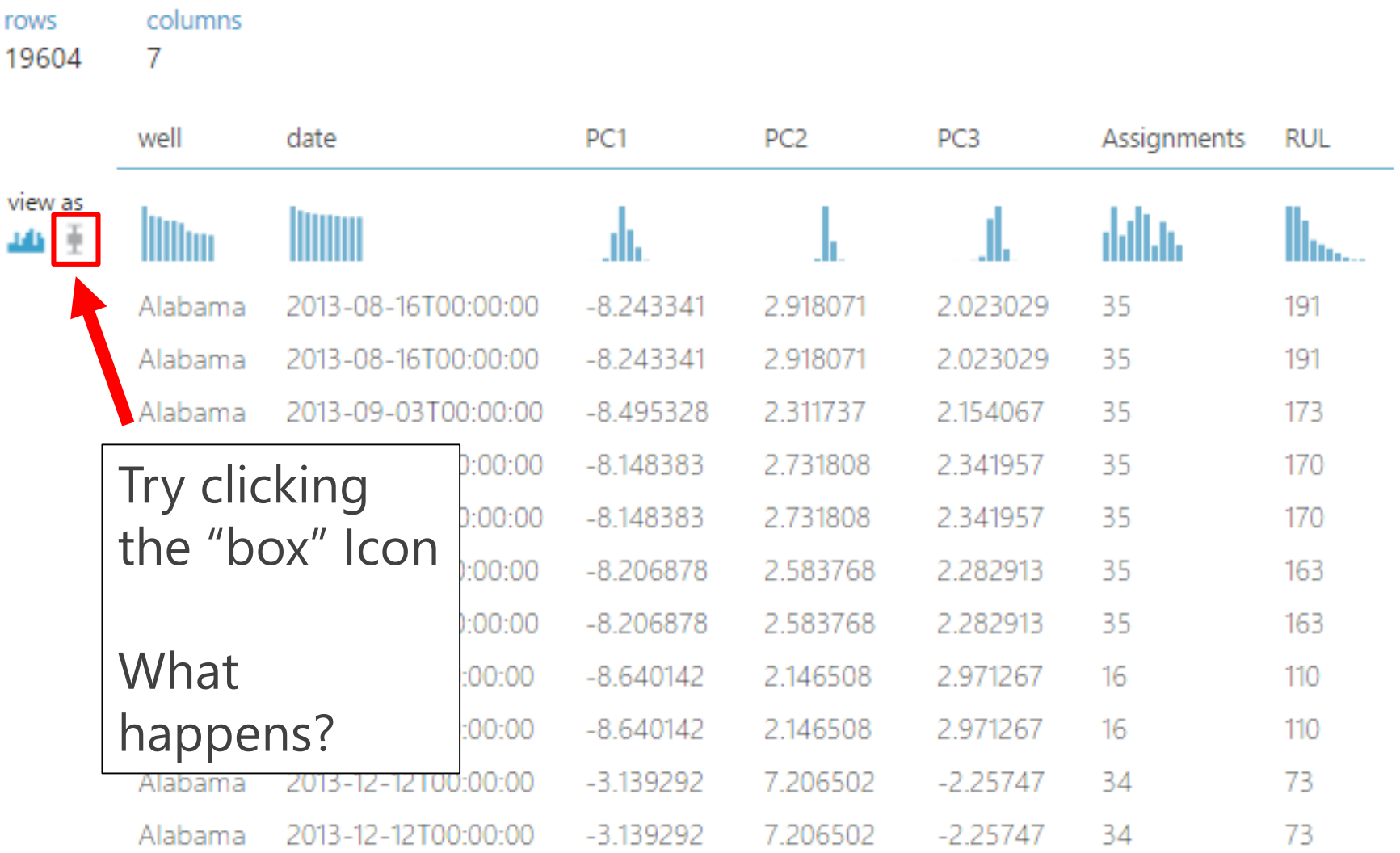
Aside Why does this matter?

- Because Azure Machine Learning is an end user tool it …
  - Does as much data science work as it can, for users
    - Example: In one click on the data set, we were able to inspect the data
      #1 Row count,
      #2 Variable Count,
      #3 Variable Names, and see
      #4 the distribution of our data
    - In particular, we can see the frequency distribution of different sensor variable

Microsoft

NEAL ANALYTICS

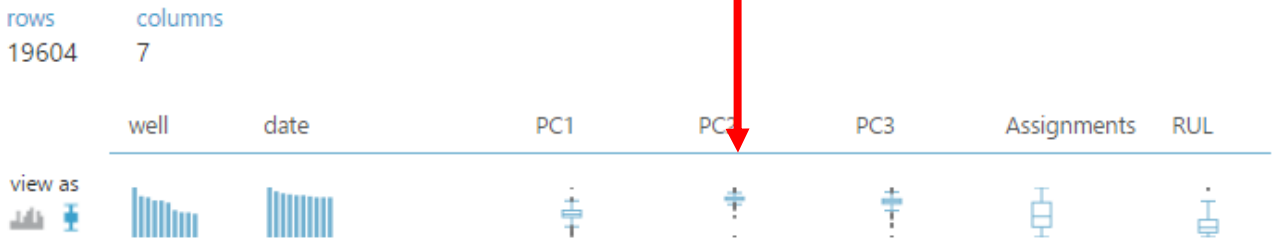
Step 4.18 : Now let's change to box plot view

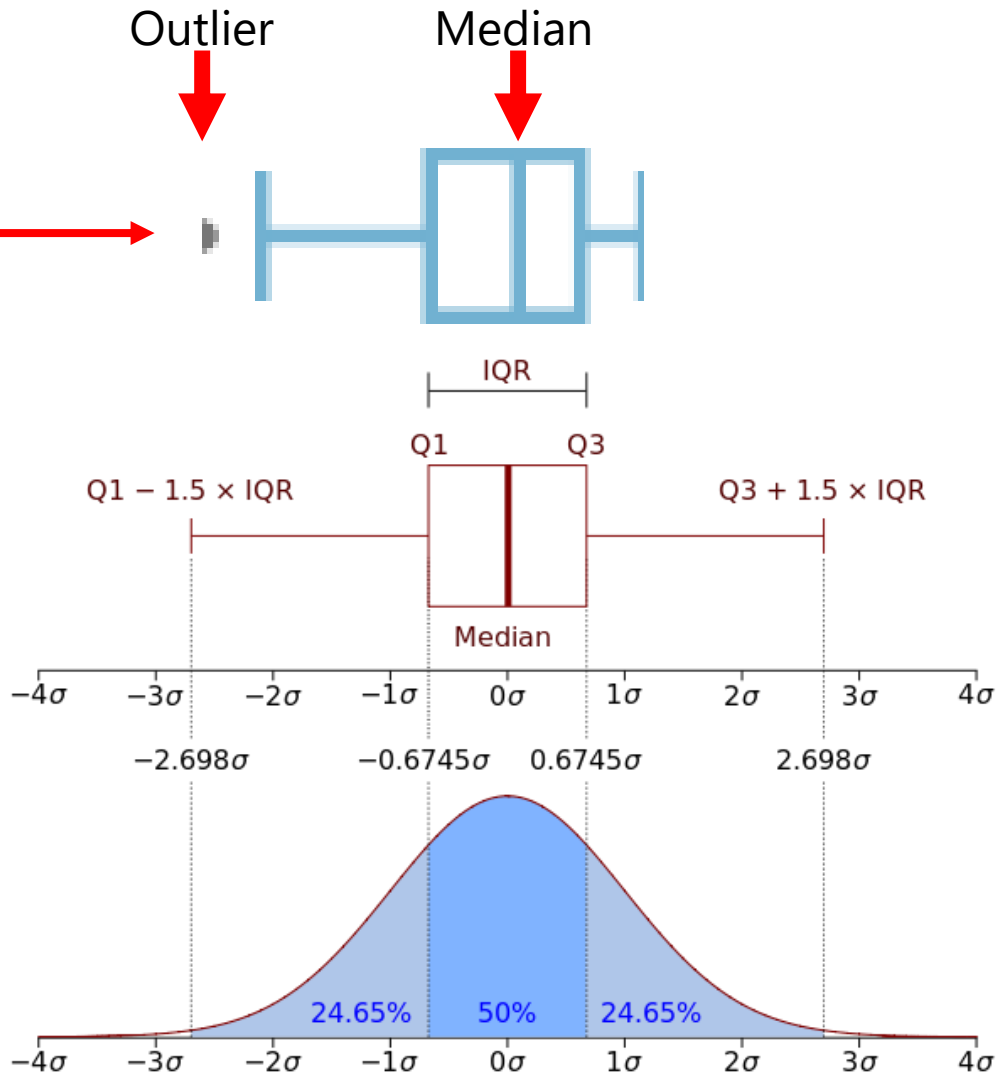


Microsoft

NEAL ANALYTICS

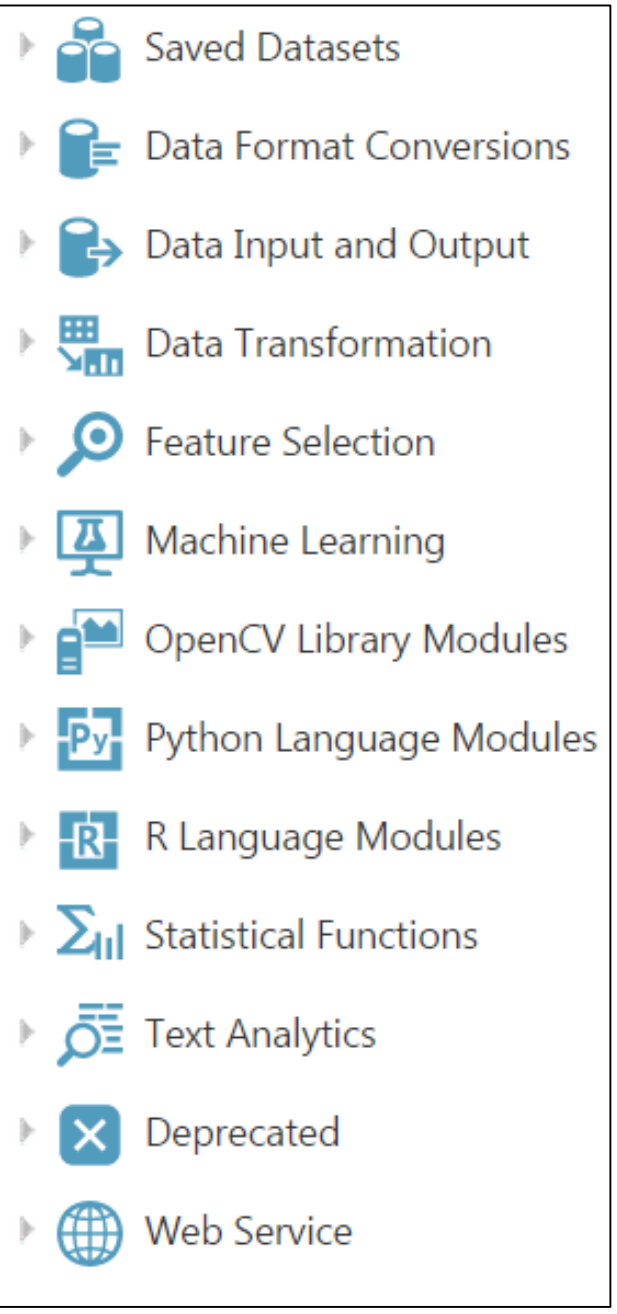
Step 4.19 : Look at the box plots of the data





Microsoft

NEAL ANALYTICS



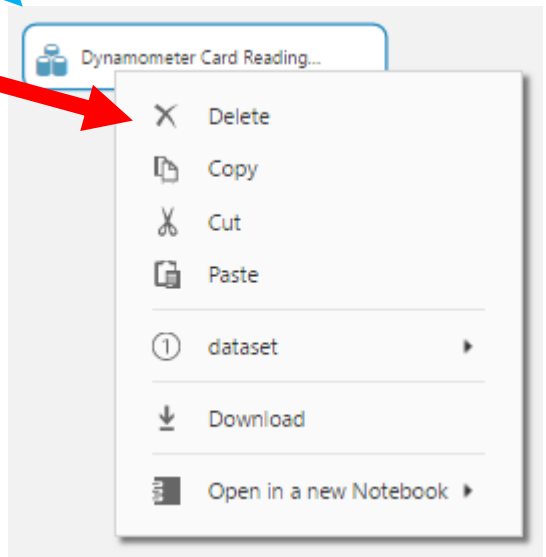
Step 5.1 : In addition to data what are the other components we need for an AML model?

- AML modelling … a checklist approach
  - ~~Create new experiment~~
  - ~~Import data set~~
  - Basic understanding of the Azure Machine Learning Canvas

Microsoft NEAL ANALYTICS

Step 5.2 : Azure Machine Learning Canvas

- Exercise
  - Now delete the data set the Azure Machine Learning Way
    - Step 1: Click on the dataset module
    - Step 2: While the module has a blue outline, right click
    - Step 3: Select “delete”
    - Step 4: Try Ctrl-Z to undo
      - Ctrl-Z does not work
- AML Canvas Lessons
  - There is no “undo” on the canvas
  - Work with your models patiently

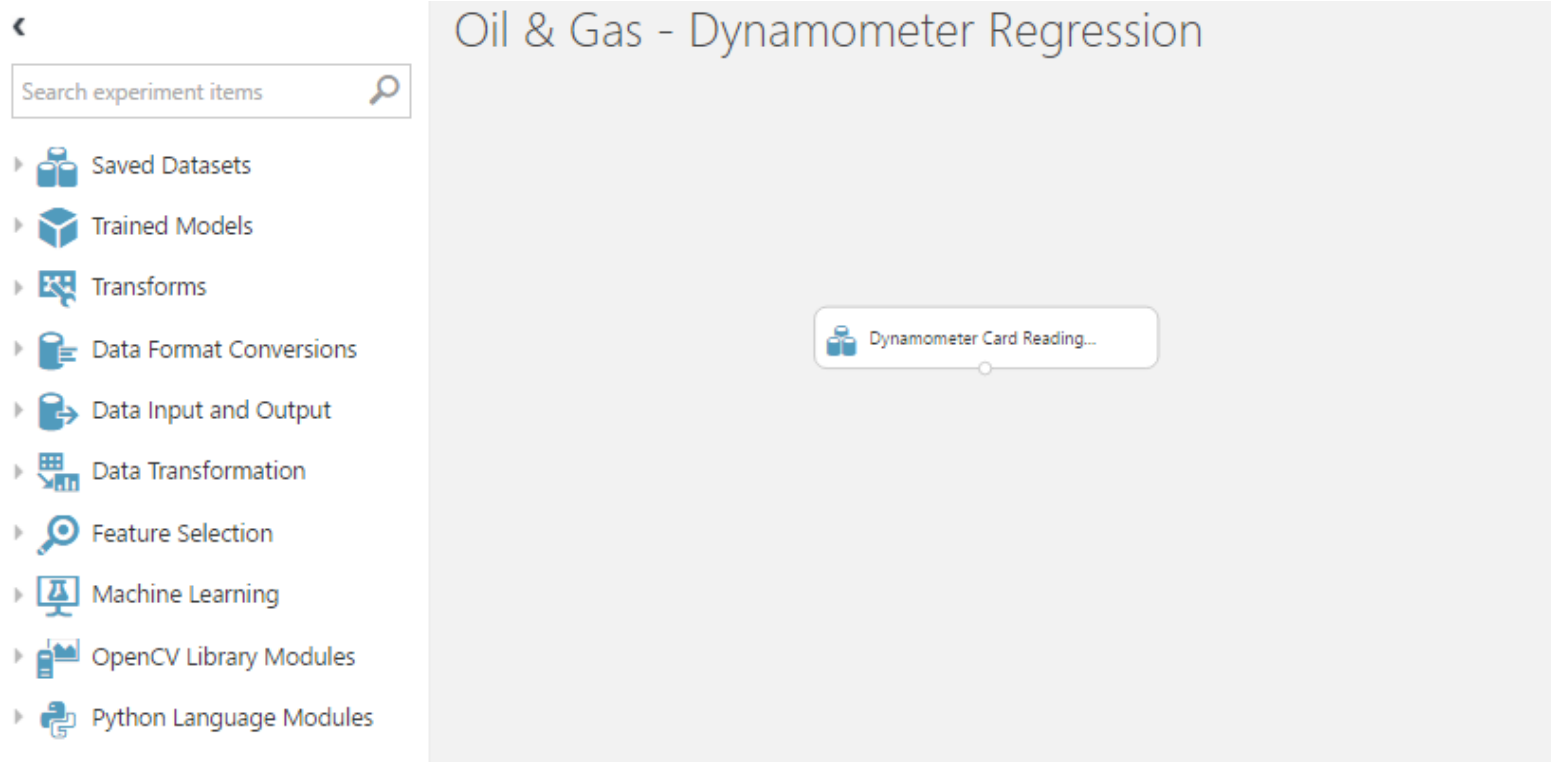


Microsoft

NEAL ANALYTICS

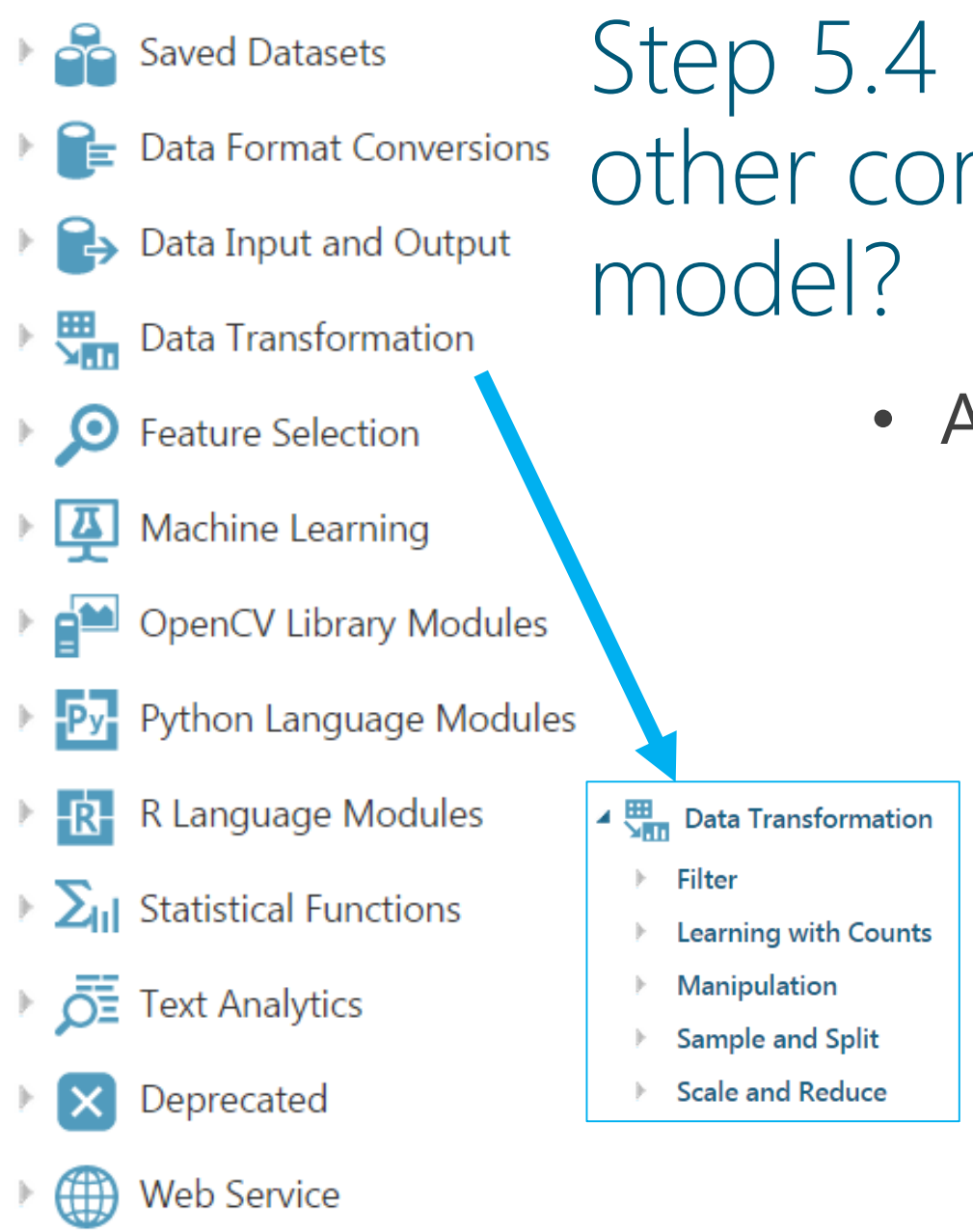
Step 5.3 : Azure Machine Learning Canvas

- Now add your data back to the empty Azure Machine Learning Canvas
  - Use Option 1 – Use Import Data module to import data from the Azure Blob
  - If Option 1 failed, then use Option 2



Microsoft NEAL ANALYTICS

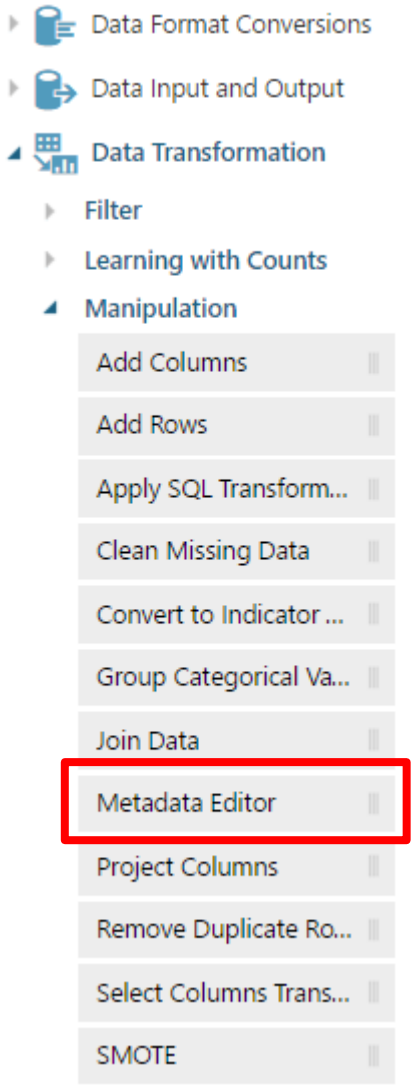
Step 5.4 : In addition to data what are the other components we need for an AML model?



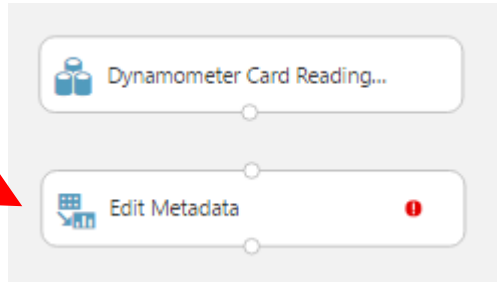
- AML modelling … a checklist approach
  - ~~☐ Create new experiment~~
  - ~~☐ Import data set~~
  - ~~☐ Basic understanding of the Azure Machine Learning Canvas~~
  - ☐ Data transformation

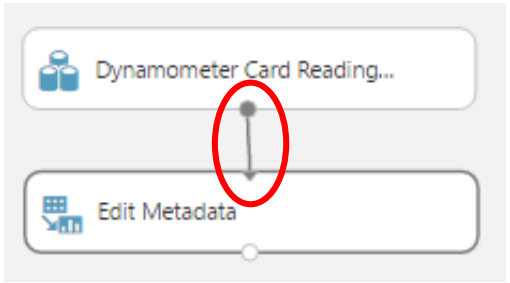
Microsoft NEAL ANALYTICS

Step 5.5 : Q? Edit Metadata



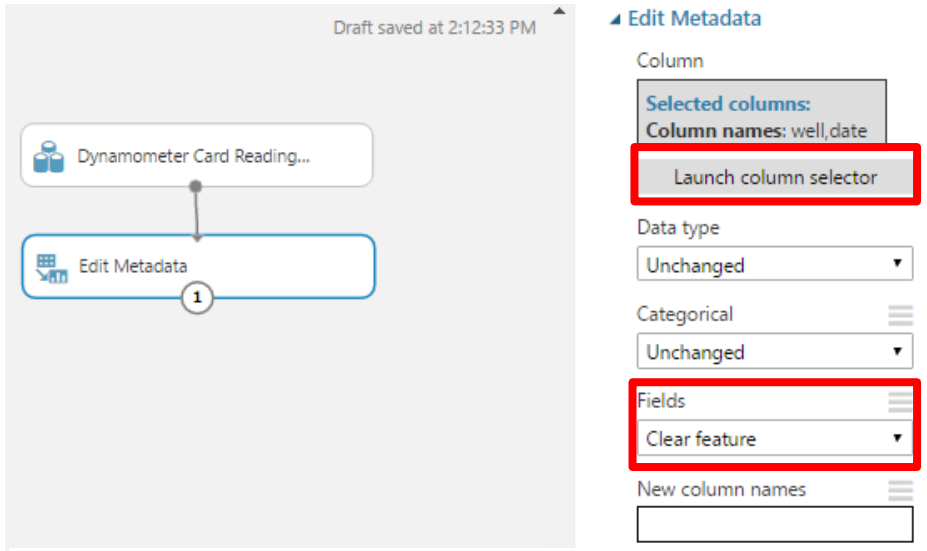
- Open “Data Transformation” and then “Manipulation” from the navigation pane at the left
- Drag “Edit Metadata” to the canvas
  - “Edit Metadata” edits metadata associated with the columns in a dataset
- Next, click and hold on the bottom middle circle of your “Dataset” module
- While holding down the mouse button, drag the line to the top middle circle of “Edit Metadata” module



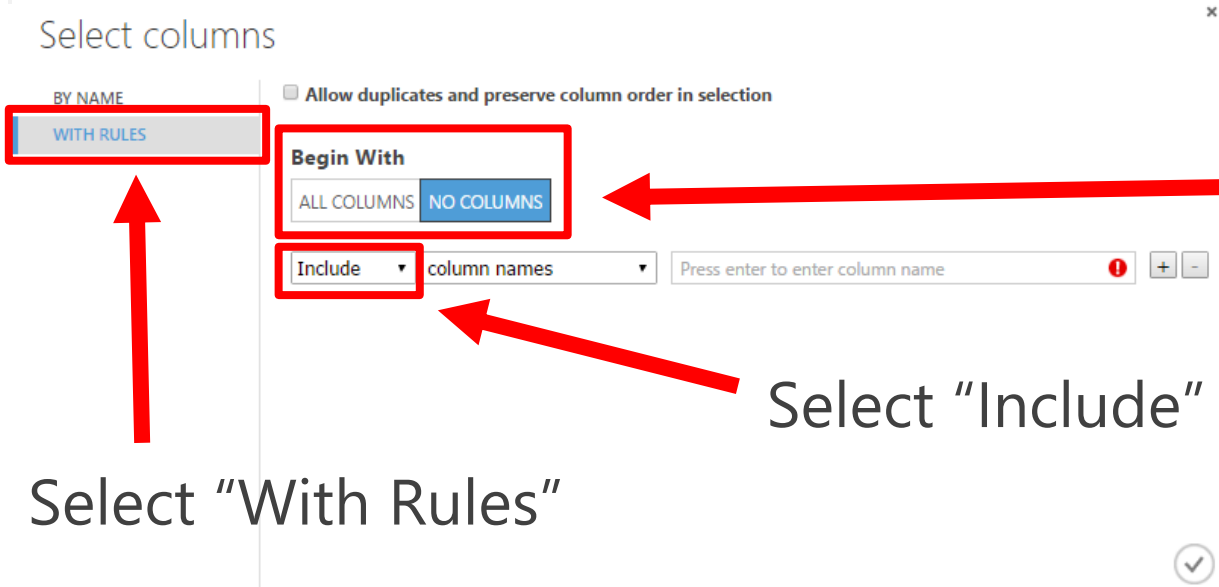


Microsoft NEAL ANALYTICS

Step 5.6 : Q? Edit Metadata



- Machine Learning
  - Click on Metadata Editor
  - Chose “Clear feature” from the drop down menu under Fields
  - Then click on “Launch column selector”



Select “No Columns”

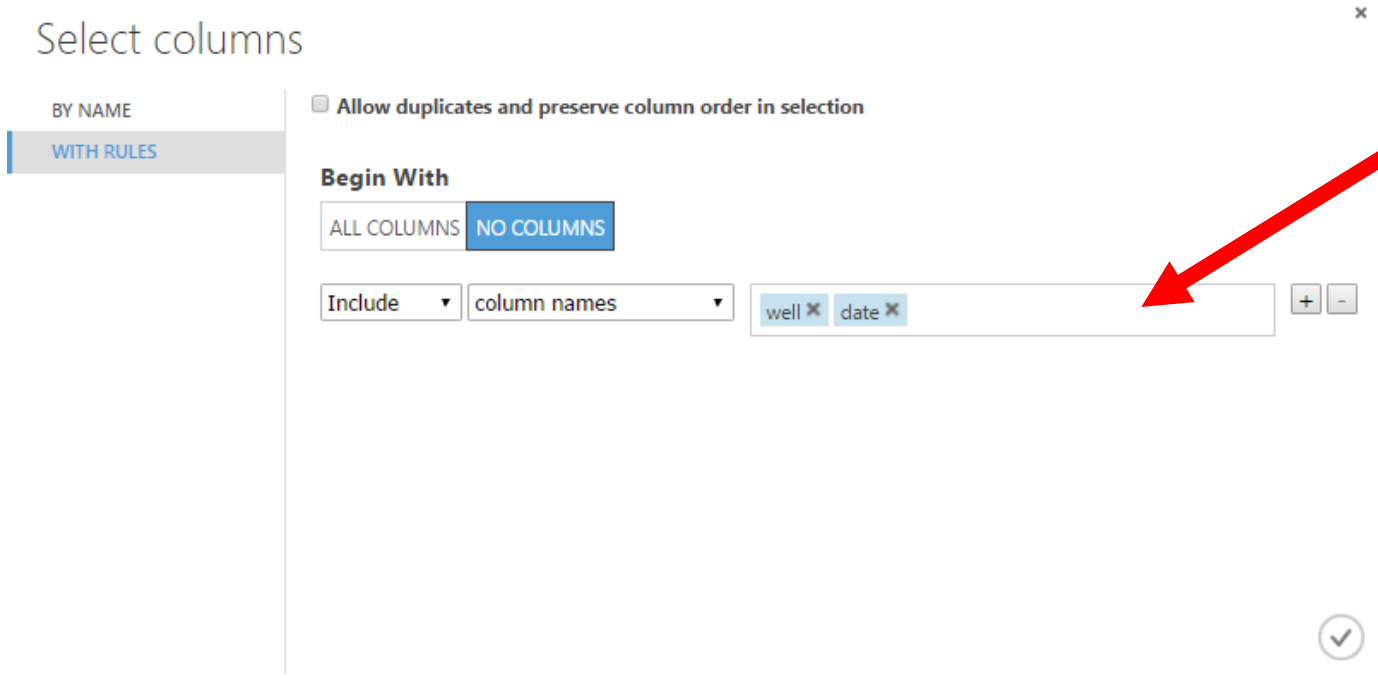
Select “Include” from the drop down menu

Select “With Rules”

Microsoft

NEAL ANALYTICS

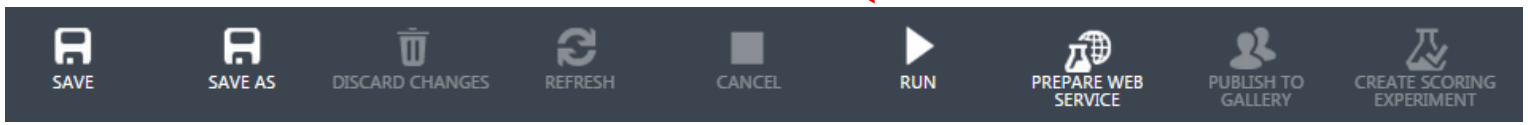
Step 5.7 : Q? Edit Metadata



- Click here and a drop down list appears
- Select one variable at a time
  - well
  - date

Step 1: Next, click on the check circle

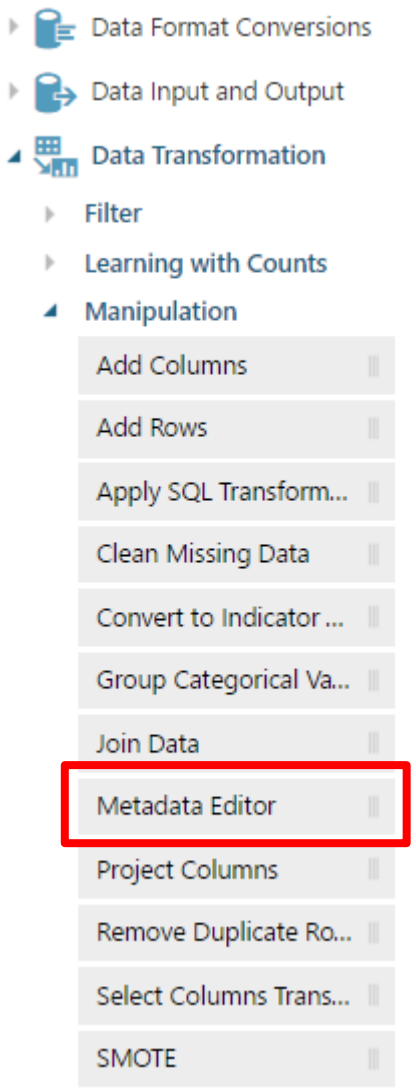
Step 2: Click “Run”



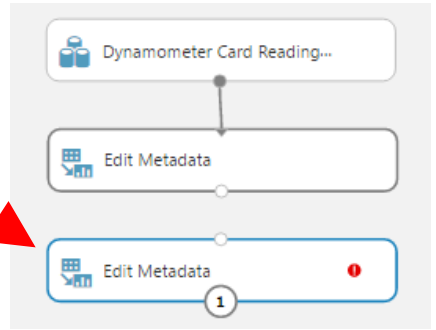
Microsoft

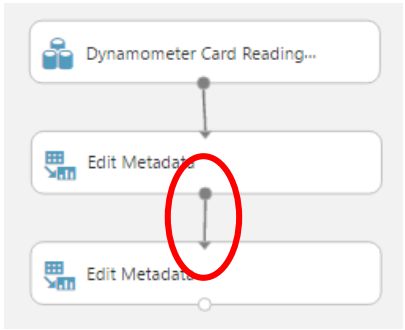
NEAL ANALYTICS

Step 5.8 : Q? Edit Metadata



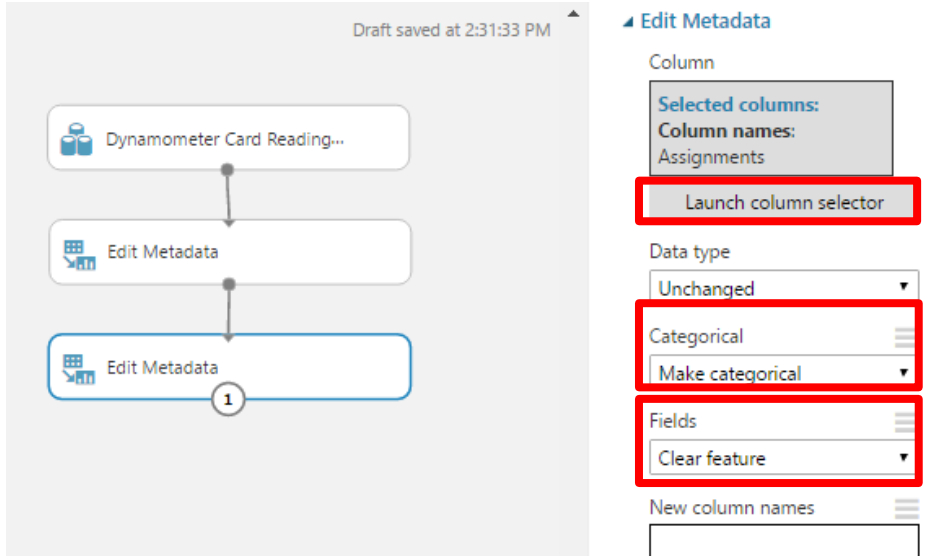
- Open “Data Transformation” and then “Manipulation” from the navigation pane at the left
- Drag “Edit Metadata” to the canvas
  - “Edit Metadata” edits metadata associated with the columns in a dataset
- Next, click and hold on the bottom middle circle of your “Edit Metadata” module
- While holding down the mouse button, drag the line to the top middle circle of “Edit Metadata” module



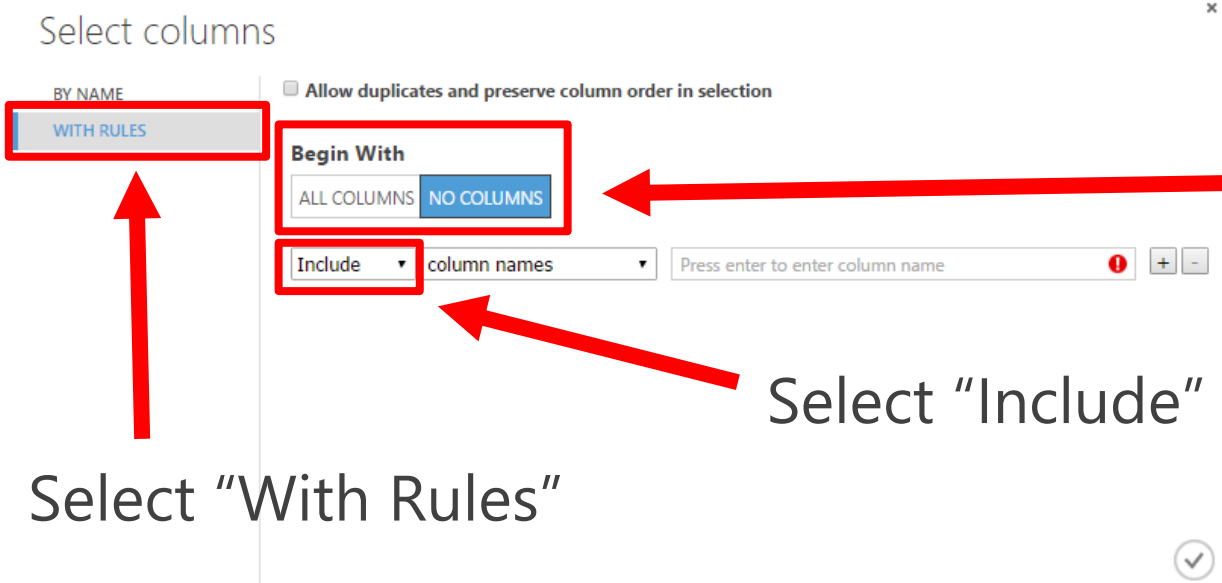


Microsoft NEAL ANALYTICS

Step 5.9 : Q? Edit Metadata



- Machine Learning
  - Click on Metadata Editor
  - Chose “Clear feature” from the drop down menu under Fields
  - Set Categorical to “Make Categorical”
  - Then click on “Launch column selector”



Select “No Columns”

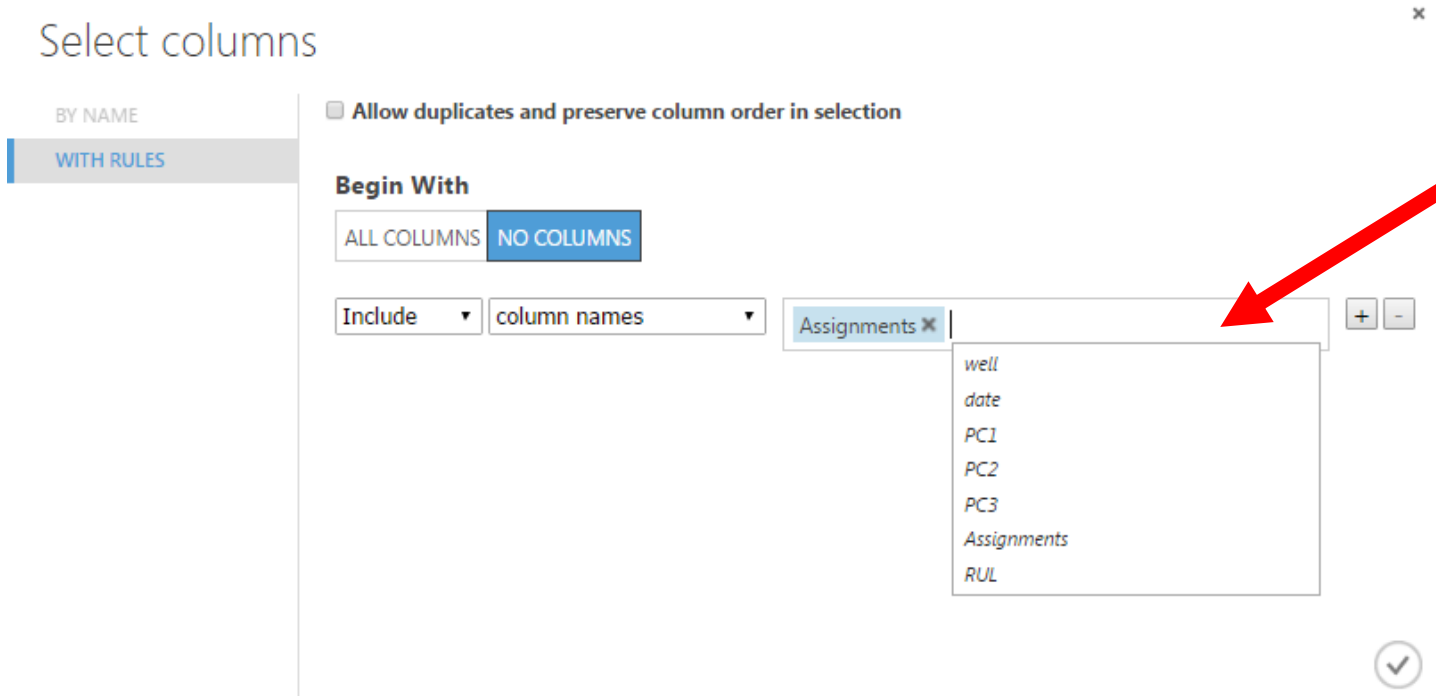
Select “Include” from the drop down menu

Select “With Rules”

Microsoft

NEAL ANALYTICS

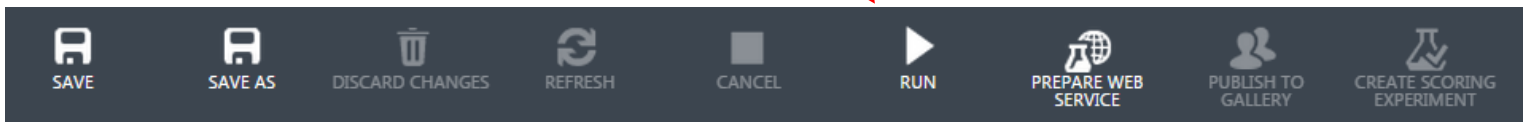
Step 5.10 : Q? Edit Metadata



- Click here and a drop down list appears
- Select one variable at a time
  - Assignments

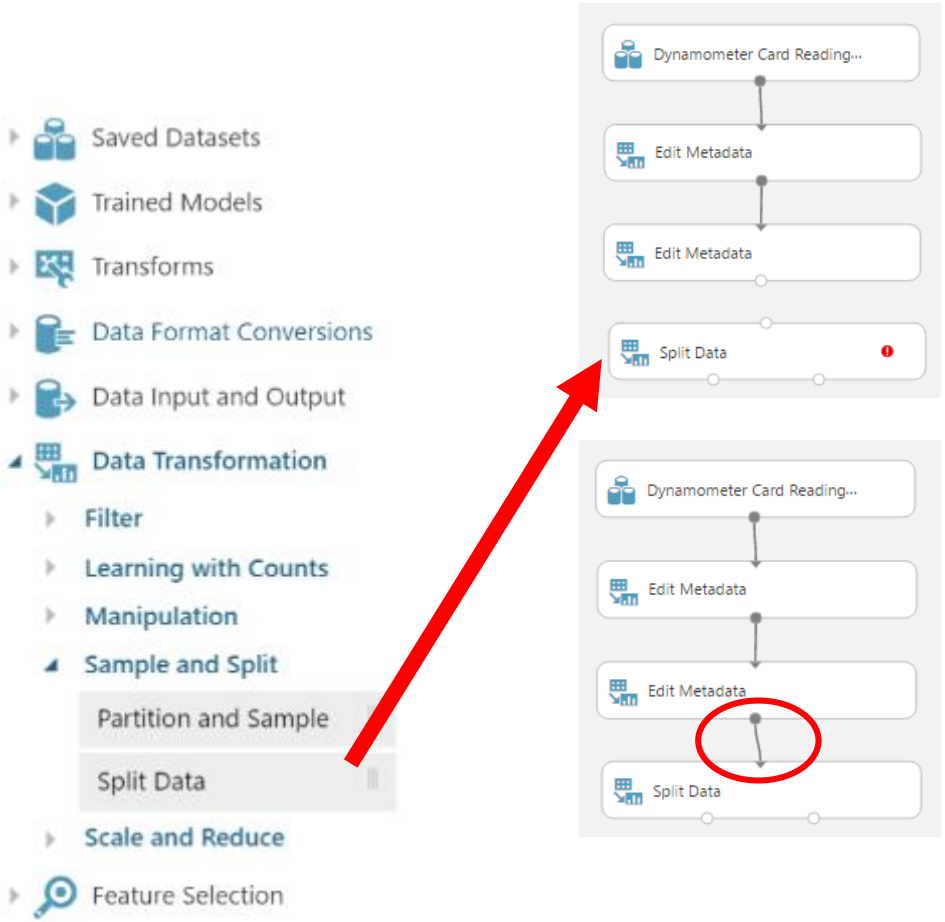
Step 1: Next, click on the check circle

Step 2: Click “Run”



Microsoft NEAL ANALYTICS

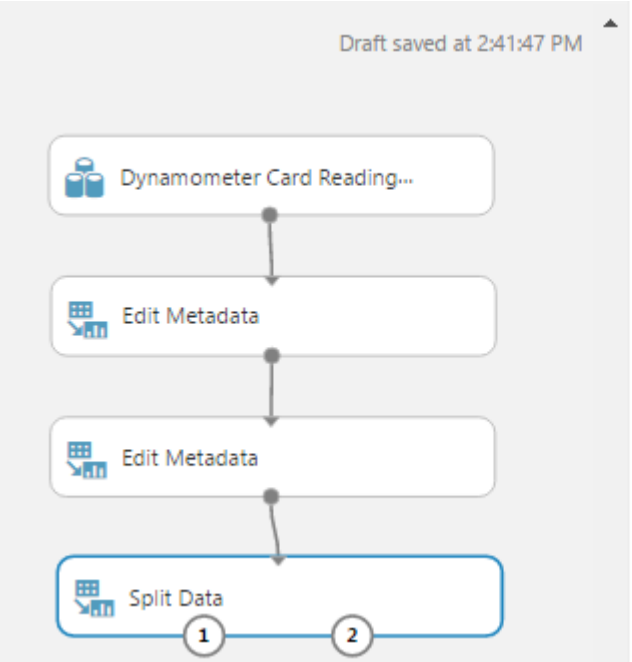
Step 5.11 : Q? How to split the dataset for training and scoring?

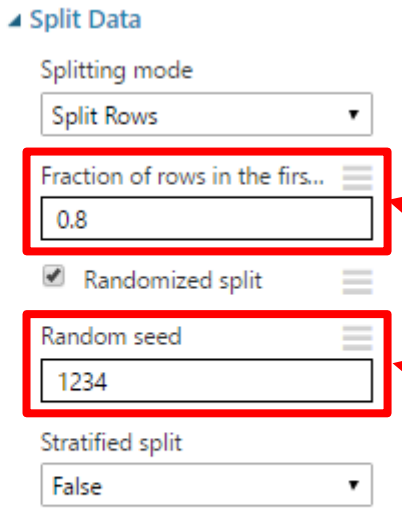


- Open “Data Transformation” and then “Sample and Split” from the navigation pane at the left
- Drag “Split Data” to the canvas
  - “Split Data” splits the rows of the dataset into 2 distinct sets
- Next, click and hold on the bottom left circle of your “Edit Metadata” module
- While holding down the mouse button, drag the line to the top middle circle of “Split Data” module

Microsoft NEAL ANALYTICS

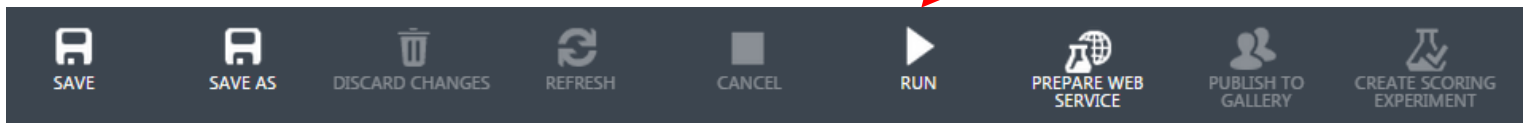
Step 5.12 : Q? How to split the dataset for training and scoring?





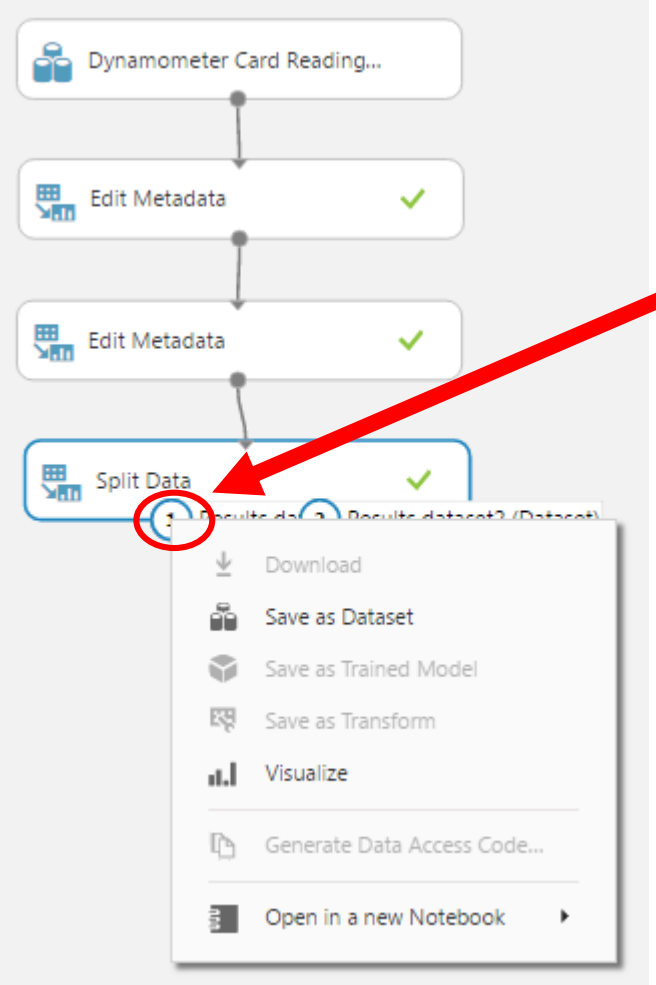
- Click on “Split Data” on your canvas, and make sure it has a blue outline
- Change the “Fraction of rows in the first output dataset” to 0.8
- Set the Random seed to 1234

Next, Click “Run”



Microsoft NEAL ANALYTICS

Step 5.13 : Q? How to split the dataset for training and scoring?



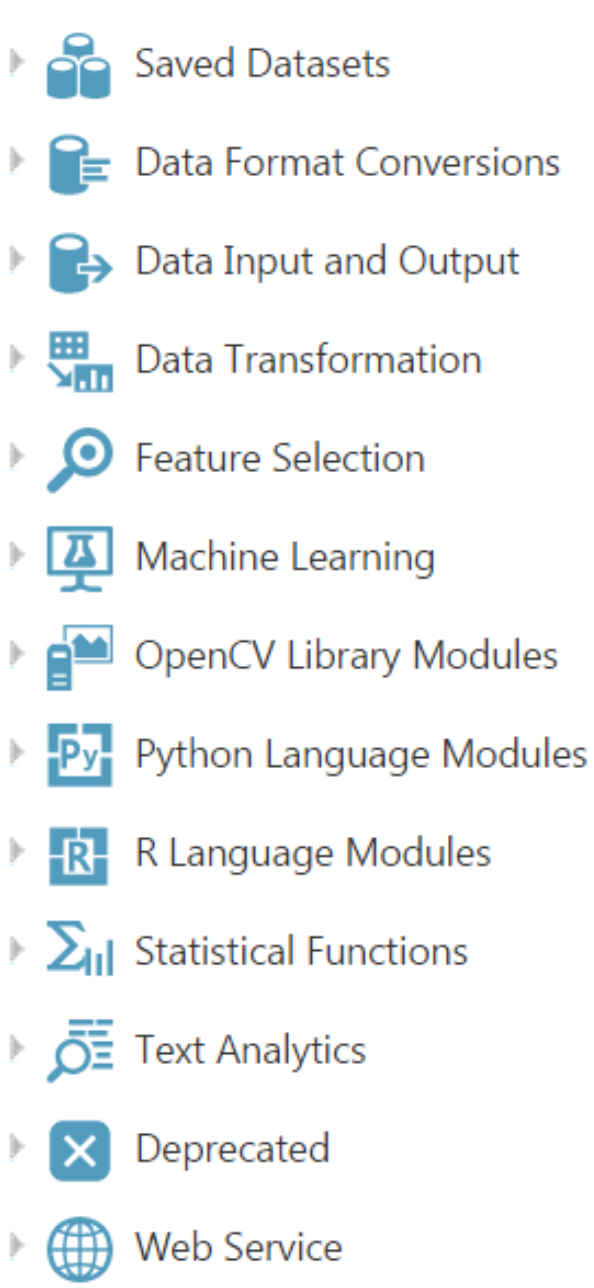
Right click on the bottom-left circle in “Split Data” and select “Visualize”

rows 15683 columns 7

view as

well	date	PC1	PC2	PC3	Assignments	RUL
Mooshie	2014-06-30T00:00:00	-3.475528	2.271826	2.588896	27	120
RobinHood	2013-12-12T00:00:00	-4.499741	-0.716241	-3.652891	0	115
Strsky	2013-09-29T00:00:00	-10.229244	4.505863	-0.833009	35	184
Koby	2014-05-01T00:00:00	-12.597254	4.843114	6.97637	16	151
Kaly	2014-09-22T00:00:00	2.935428	8.755442	-2.064127	8	116
Skitty	2013-11-03T00:00:00	6.851085	15.864168	-2.050981	37	79
Ike	2015-01-07T00:00:00	6.611832	-8.507007	-2.124907	38	18
Marti	2014-11-21T00:00:00	8.443667	3.012456	-3.437652	10	158
Strsky	2014-03-29T00:00:00	-10.008209	3.834503	1.782888	16	3

Microsoft NEAL ANALYTICS

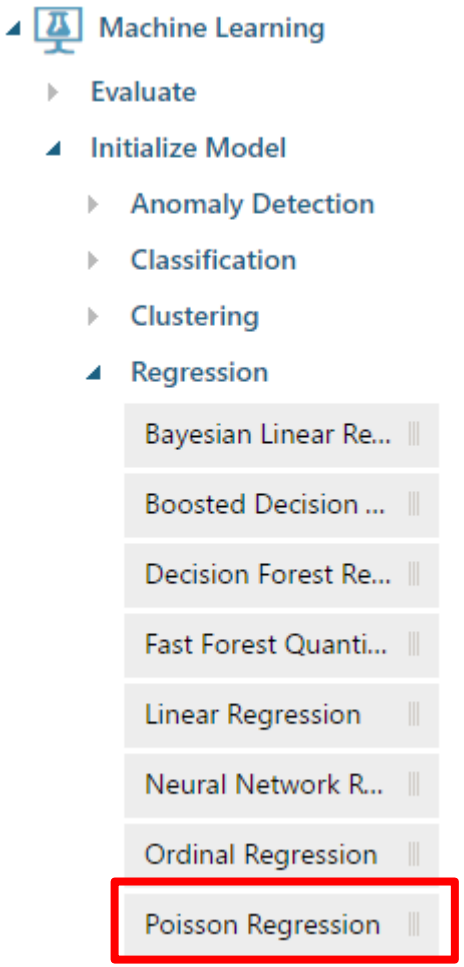


Step 6.1 : AML Modeling Components

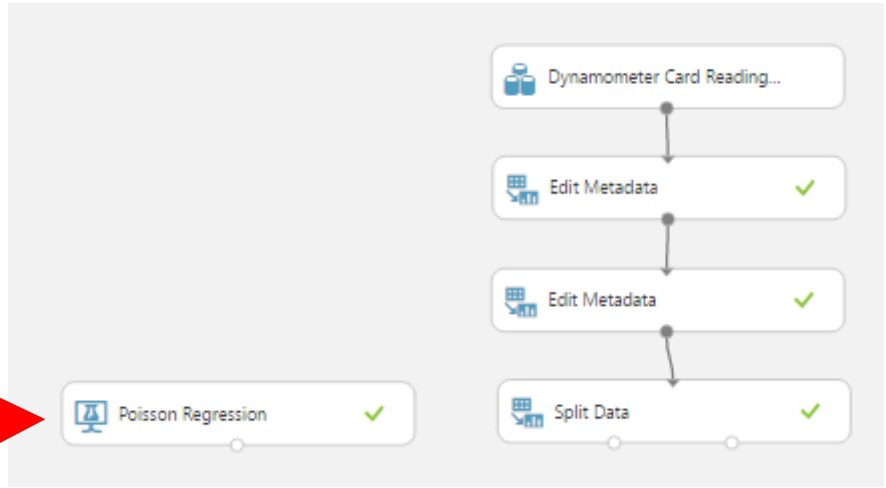
- AML modelling … a checklist approach
  - ☑ ~~Create new experiment~~
  - ☑ ~~Import data set~~
  - ☑ ~~Basic understanding of the Azure Machine Learning Canvas~~
  - ☑ ~~Data transformation~~
  - ☐ Machine Learning

Microsoft NEAL ANALYTICS

Step 6.2 : Q? How to develop the Machine Learning Algorithm?



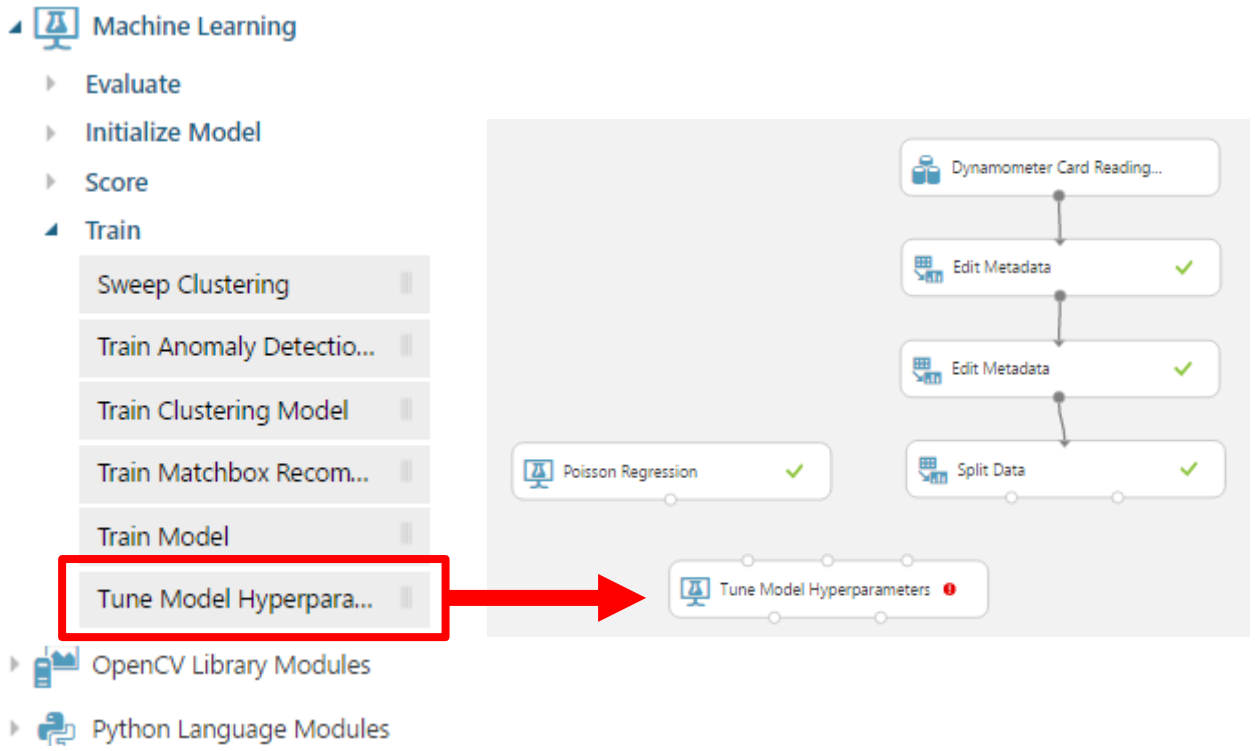
- Open “Machine Learning” -> “Initialize Model” -> “Regression” from the navigation pane at the left
- Drag “Poisson Regression” to the canvas
  - “Poisson Regression” Creates a regression model that assumes data has a Poisson distribution



Microsoft

NEAL ANALYTICS

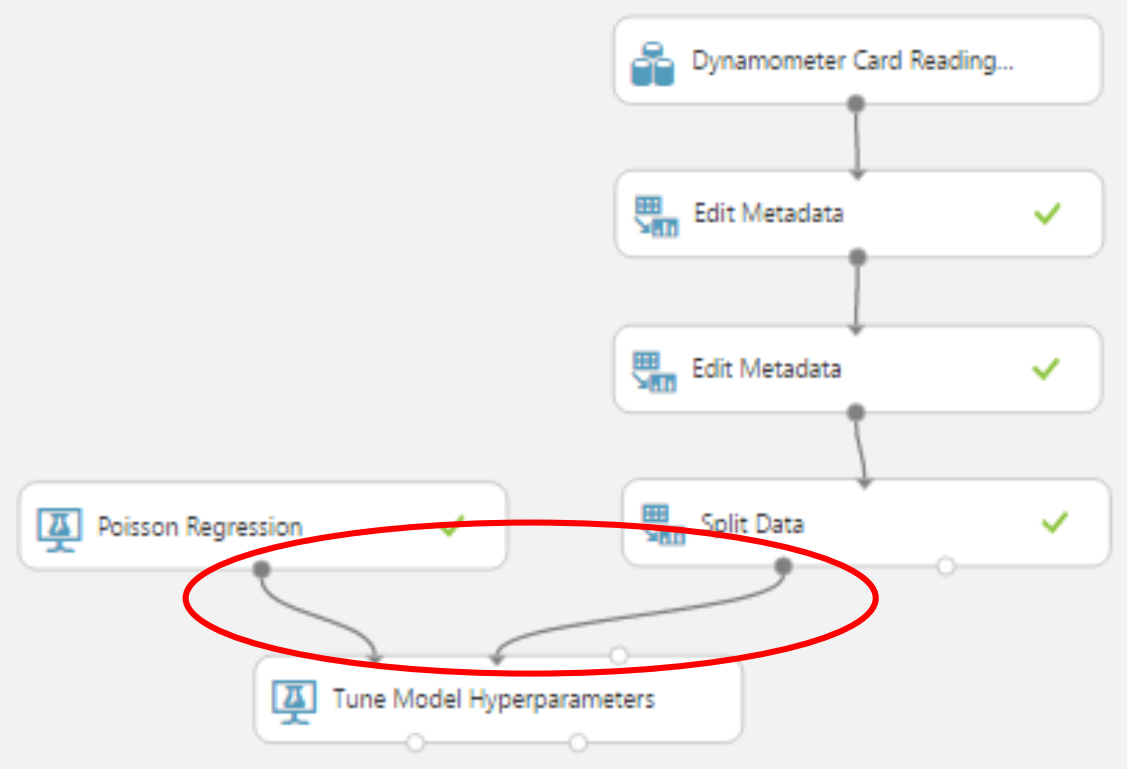
Step 6.3 : Q? How to develop the Machine Learning Algorithm?



- Open “Machine Learning” and then “Train” from the navigation pane at the left
- Drag “Tune Model Hyperparameters” module to the canvas
  - “Tune Model Hyperparameters” perform a parameter sweep on the model to determine the optimum parameter settings

Microsoft NEAL ANALYTICS

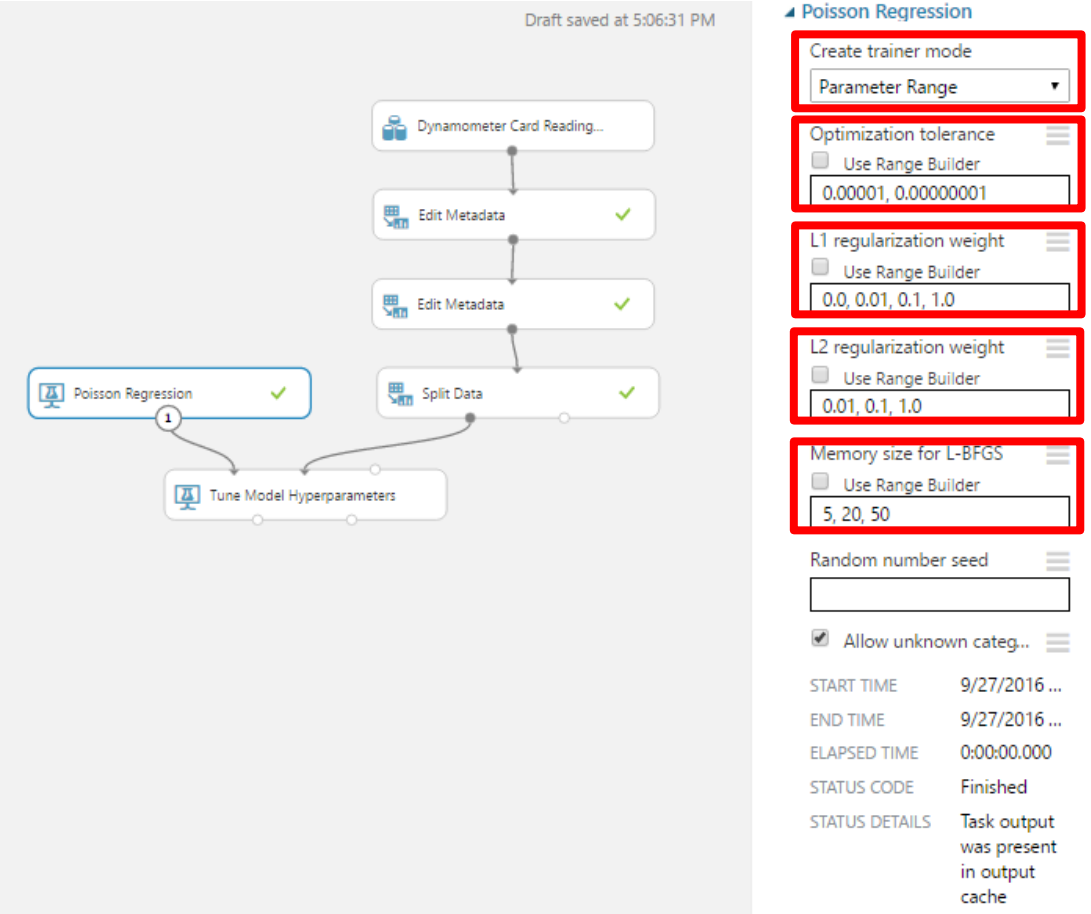
Step 6.4 : Q? How to develop the Machine Learning Algorithm?



- Next, click and hold on the bottom middle circle of your “Poisson Regression” module and drag the line to the top left circle of the “Tune Model Hyperparameters” module.
- Similarly, click and hold on the bottom left circle of your “Split Data” module and drag the line to the top right circle of the “Tune Model Hyperparameters” module.

Microsoft NEAL ANALYTICS

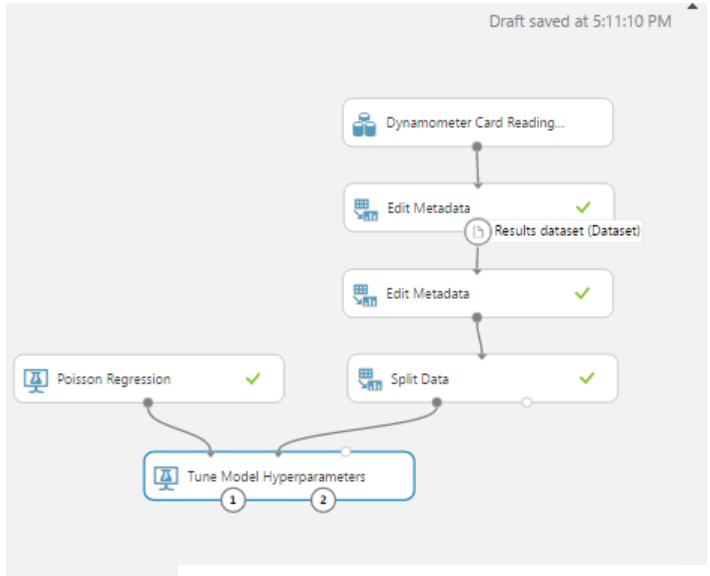
Step 6.5 : Q? How to develop the Machine Learning Algorithm

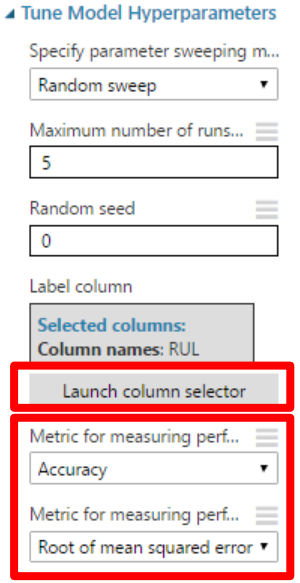


- Machine Learning
  - Chose “Parameter Range” from Create container mode
  - Set Optimization tolerance to “0.00001, 0.00000001”
  - Set L1 regularization weight to “0.0, 0.01, 0.1, 1.0”
  - Set L2 regularization weight to “0.01, 0.1, 1.0”
  - Set Memory size for L-BFGS to “5, 20, 50”

Microsoft NEAL ANALYTICS

Step 6.6 : Q? How to develop the Machine Learning Algorithm





- Machine Learning
  - Click on "Tune Model Hyperparameters"
  - Set Metric for measuring performance for classification to "AUC"
  - Set Metric for measuring performance for regression to "Root of mean squared error"
  - Then click on "Launch column selector"

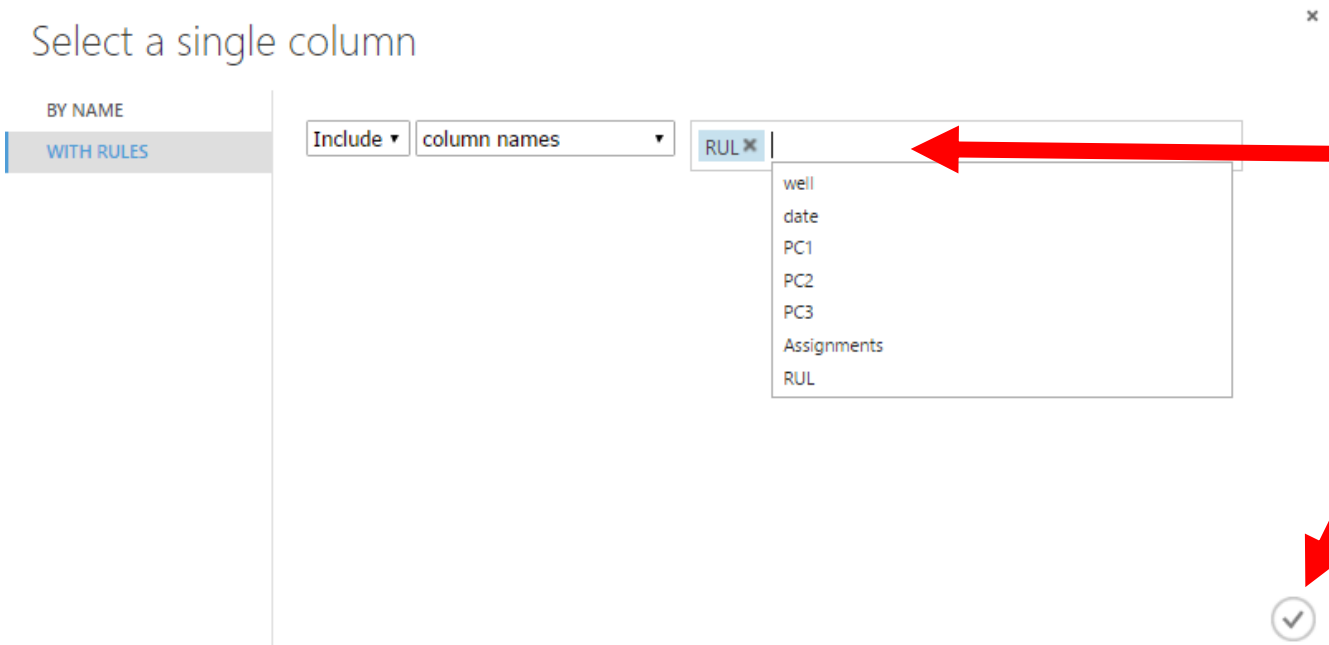
Select "With Rules"

Select "Include" from the drop down menu

Microsoft

NEAL ANALYTICS

Step 6.7 : Q? How to develop the Machine Learning Algorithm?



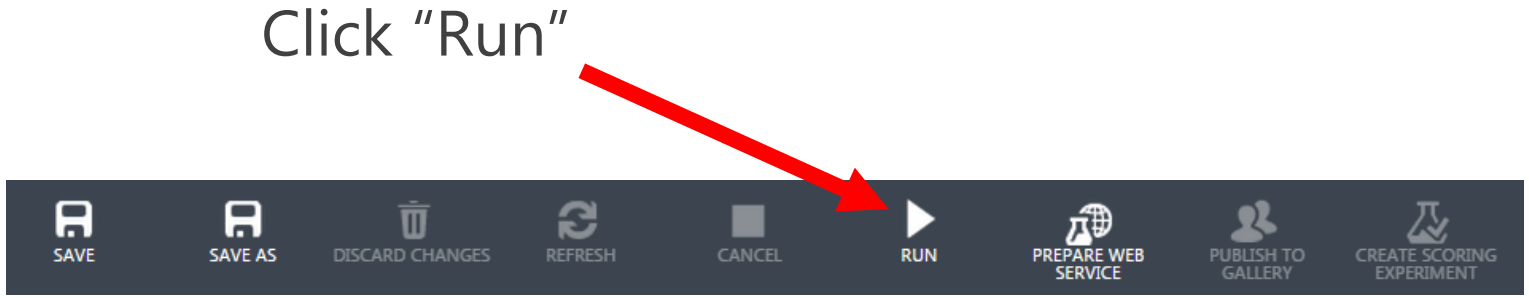
- Click here and a drop down list appears
- Select one variable at a time
  - RUL
- Next, click on the check circle

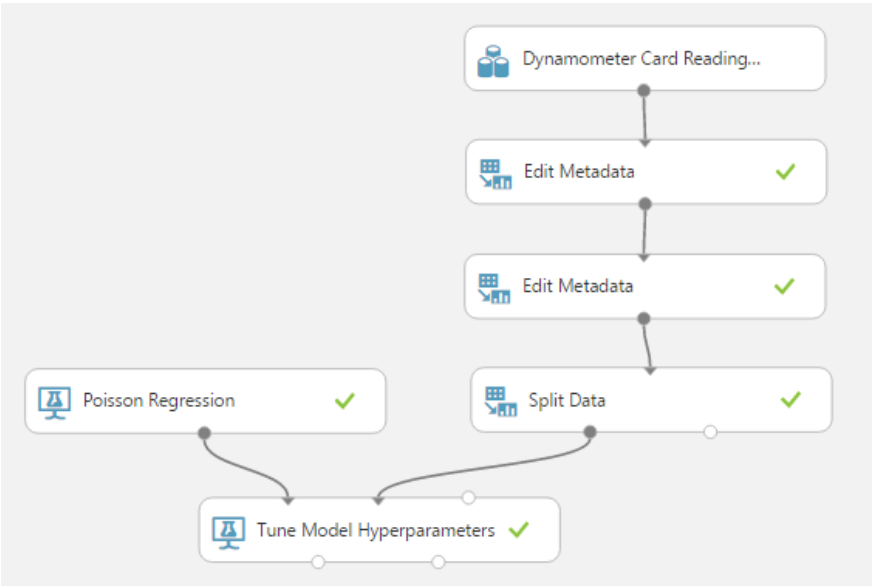
Microsoft

NEAL ANALYTICS

Step 6.8 : Run the Experiment

Click “Run”

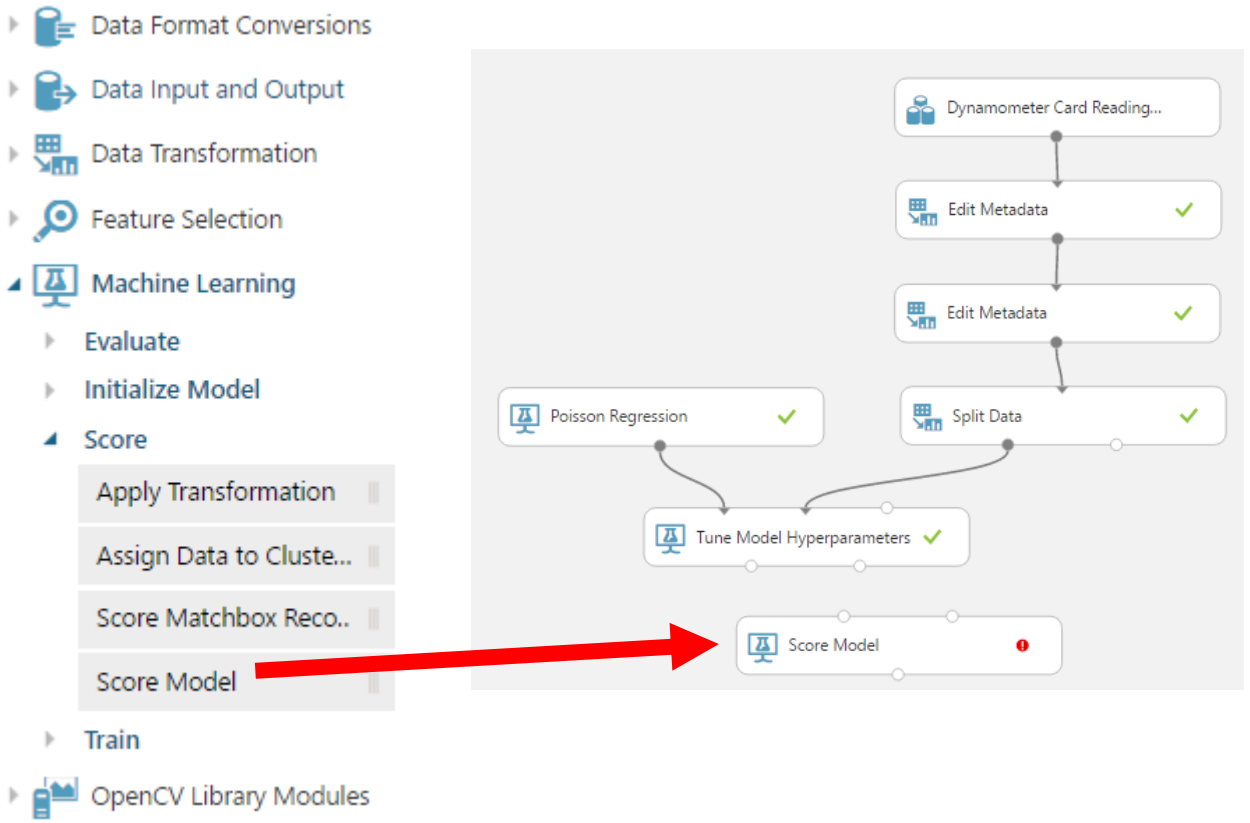




- After Azure Machine Learning “runs”, you will see a green check mark in all the modules
- All it took to train a machine learning model was drag and drop

Microsoft NEAL ANALYTICS

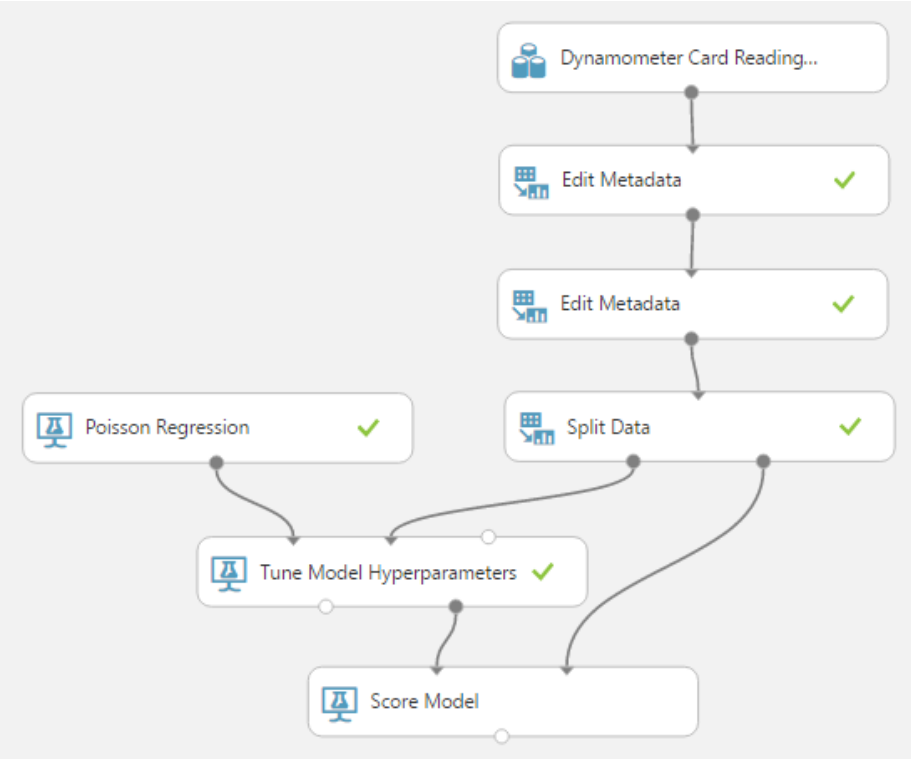
Step 6.9 : Add Score Model



- Open “Machine Learning” and then “Score” from the navigation pane at the left
- Drag two “Score Model” to the canvas
  - “Score Model” scores a trained classification or a regression model

Microsoft NEAL ANALYTICS

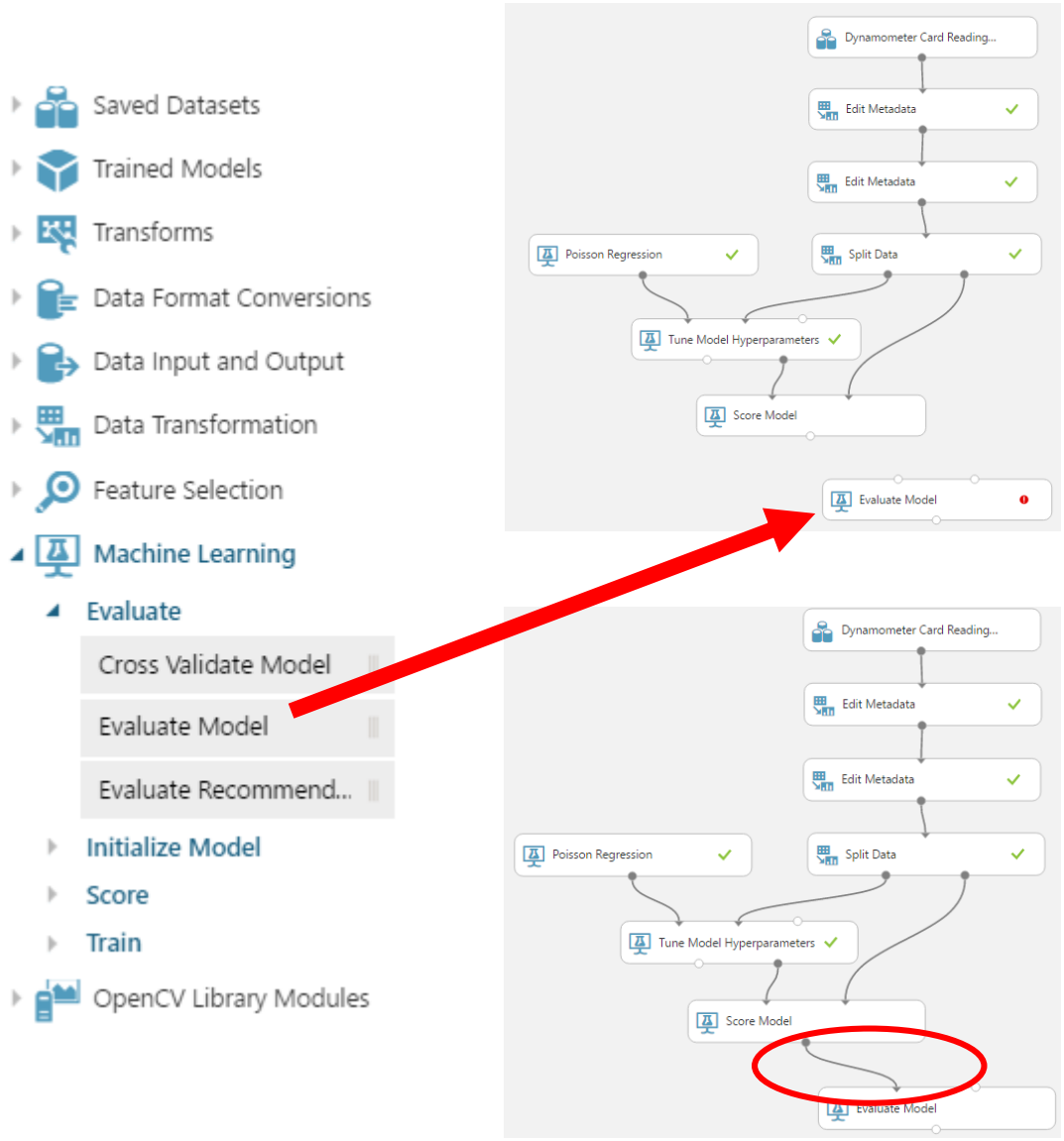
Step 6.10 : Q? How to develop the Machine Learning Algorithm?



- Next, click and hold on the bottom right circle of your “Tune Model Hyperparameters” module and drag the line to the top left circle of the “Score Model” module.
- Similarly, click and hold on the bottom right circle of your “Split Data” module and drag the line to the top right circle of the “Score Model” module.

Microsoft NEAL ANALYTICS

Step 6.11 : Add Evaluate Model

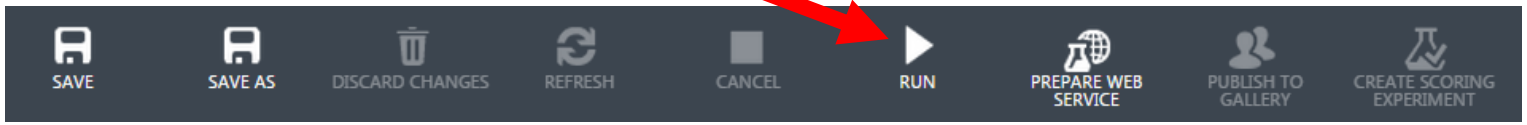


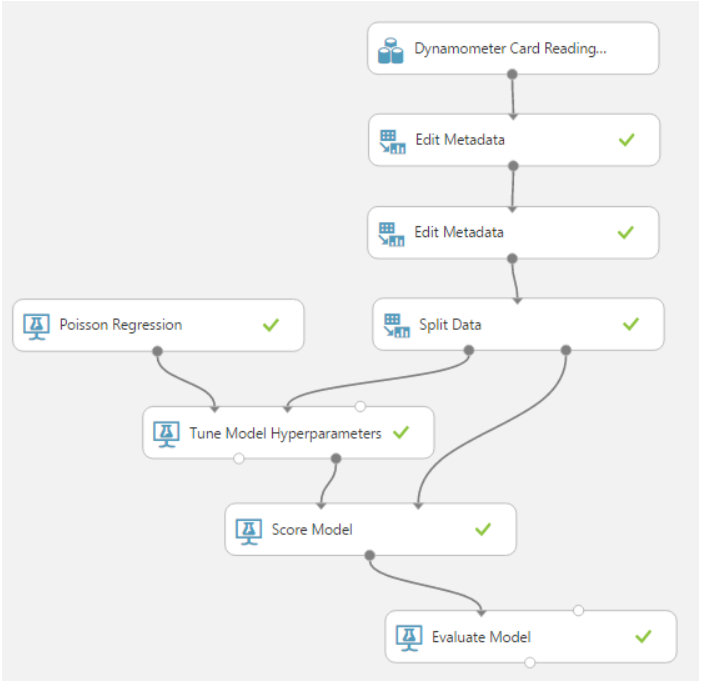
- Open “Machine Learning” and then “Evaluate” from the navigation pane at the left
- Drag “Evaluate Model” to the canvas
  - “Evaluate Model” evaluates the scored classification or regression model with standard metrics
- Next, click and hold on the bottom middle circle of your “Score Model” module
- While holding down the mouse button, drag the line to the top left circle of “Evaluate Model” module

Microsoft NEAL ANALYTICS

Step 6.12 : Run the Experiment

Click “Run”

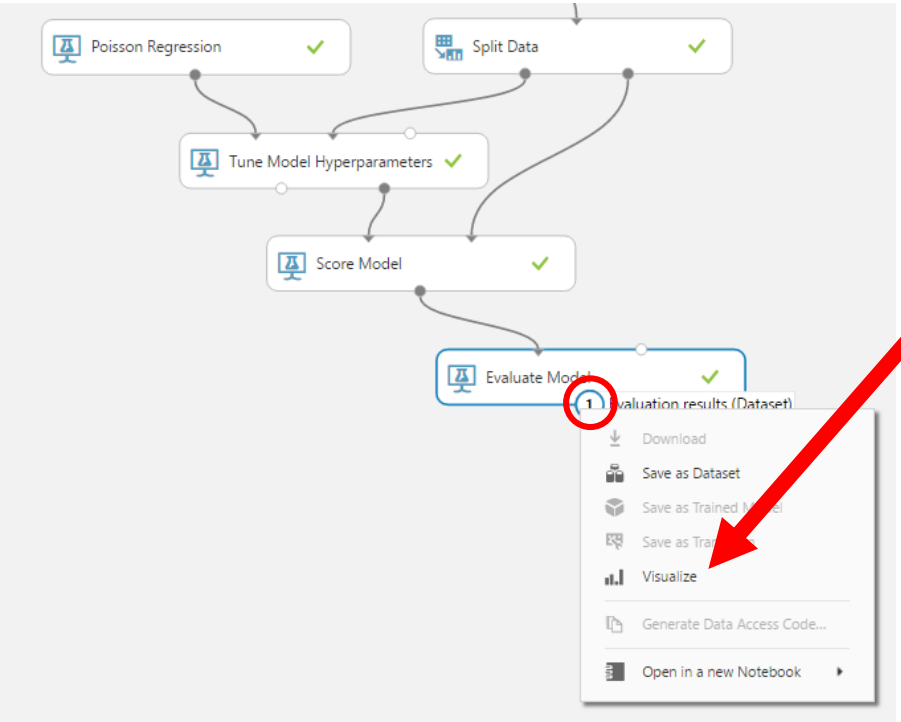




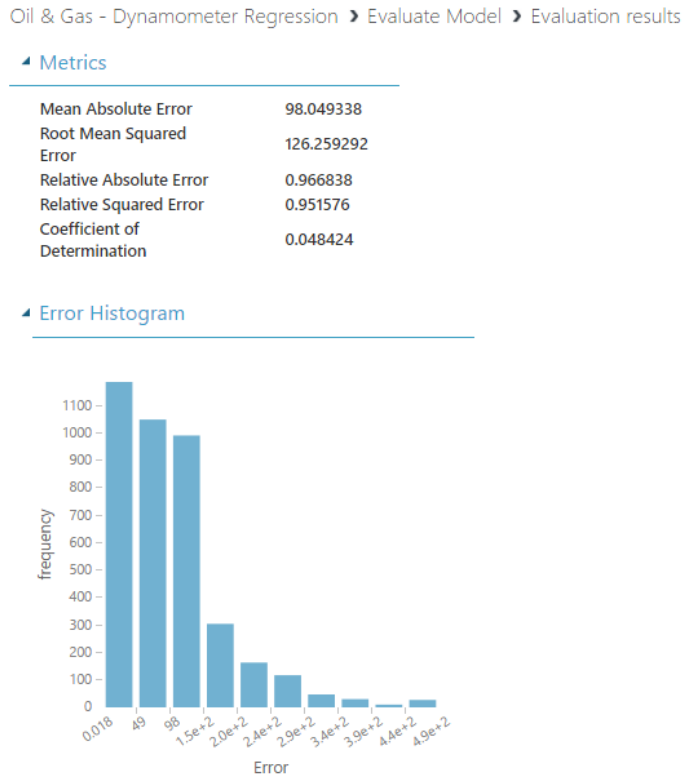
- After Azure Machine Learning “runs”, you will see a green check mark in all the modules
- All it took to train a machine learning model was drag and drop

Microsoft NEAL ANALYTICS

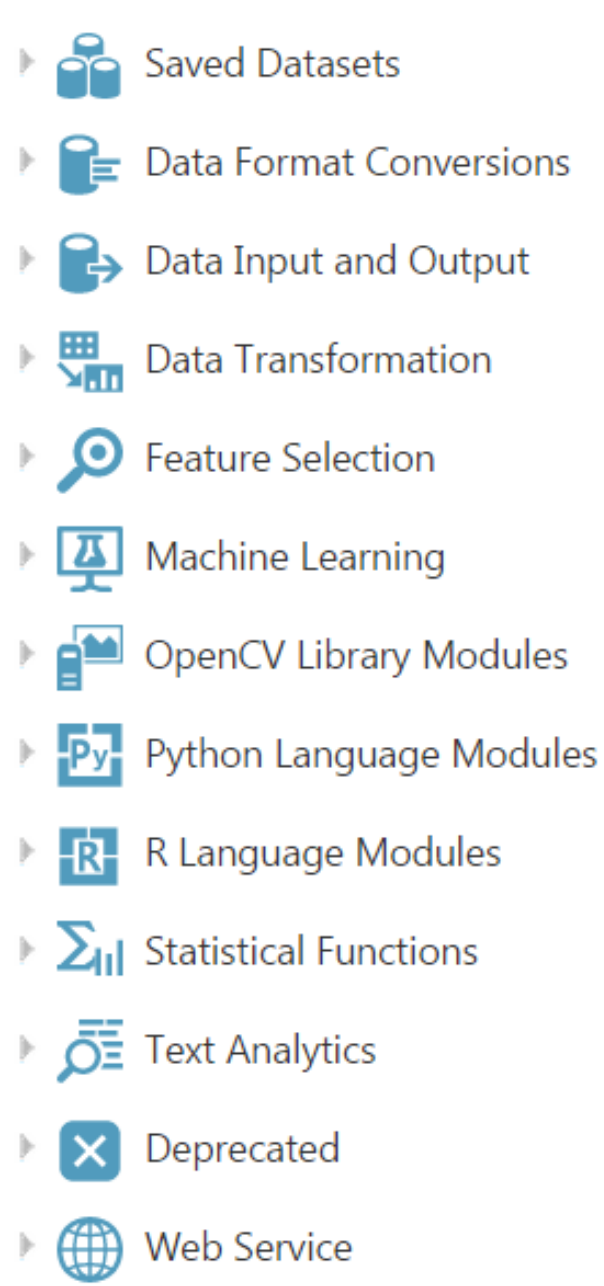
Step 6.13 : Visualize the Results



- Right click on the bottom-middle circle of “Evaluate Model” and select “Visualize”



Microsoft NEAL ANALYTICS



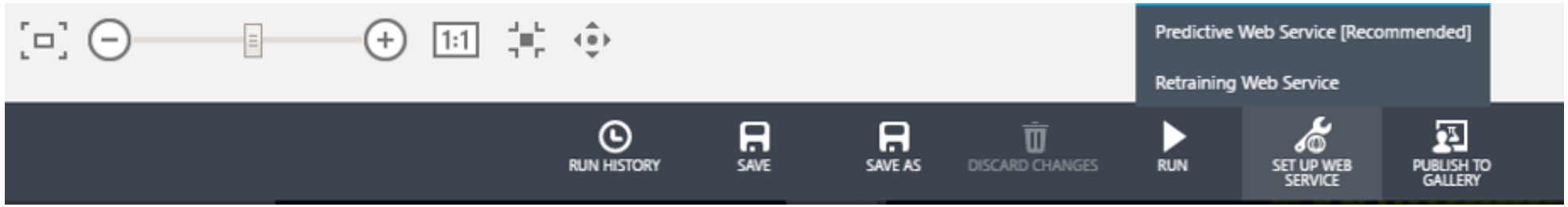
Step 7.1 : Deploy the experiment as a web service

- AML modelling … a checklist approach
  - ☑ ~~Create new experiment~~
  - ☑ ~~Import data set~~
  - ☑ ~~Basic understanding of the Azure Machine Learning Canvas~~
  - ☑ ~~Data transformation~~
  - ☑ ~~Machine Learning~~
  - ☐ Deploy the web service

Microsoft NEAL ANALYTICS

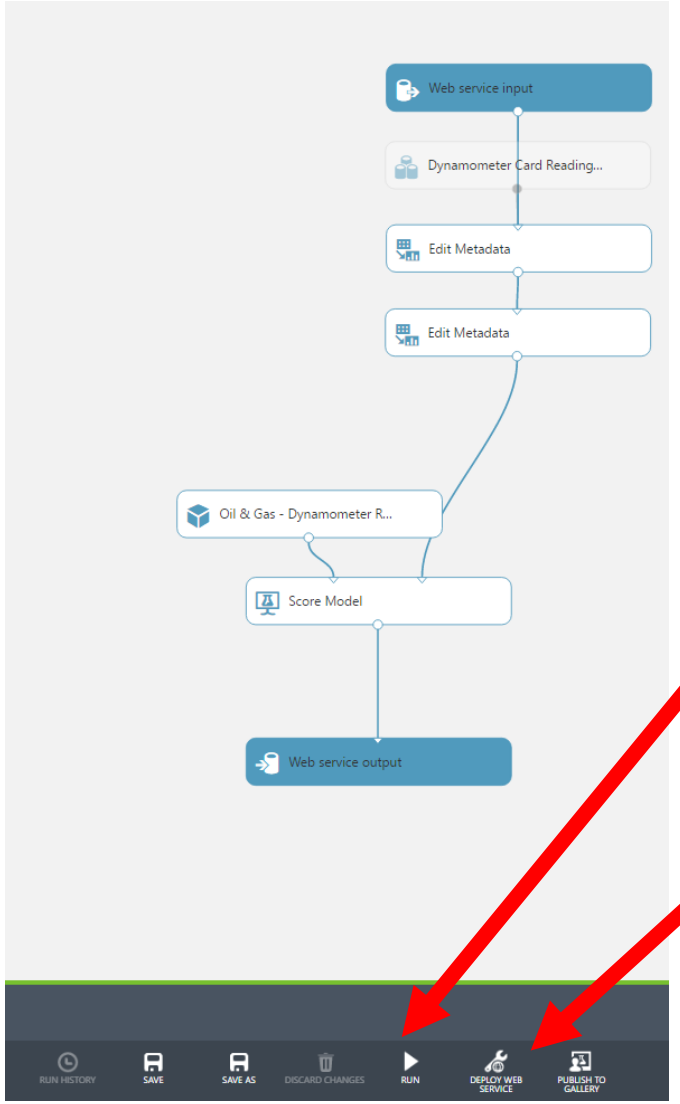
Step 7.2 : Deploy the web service

- Once you've trained your model, you're ready to use it to score new data.
- To do this, you convert your training experiment into a predictive experiment. By converting to a predictive experiment, you're getting your trained model ready to be deployed as a scoring web service.
- Users of the web service will send input data to your model and your model will send back the prediction results.
- Click Set Up Web Service, then select Predictive Web Service



Microsoft NEAL ANALYTICS

Step 7.3 : Deploy the web service



- To deploy your predictive experiment,
  - click **Run** at the bottom of the experiment canvas. Wait for the experiment to re-run
  - then click **Deploy Web Service**.
- The web service is set up and you are placed in the web service dashboard.

Microsoft NEAL ANALYTICS

Step 7.4 : Deploy the web service

- Machine Learning Studio deploys the experiment as a web service and takes you to the dashboard for that web service.
  - You can return to the experiment by clicking View snapshot or View latest
  - Run a simple test of the web service by clicking the Test
  - There is also information here for creating applications that can access the web service

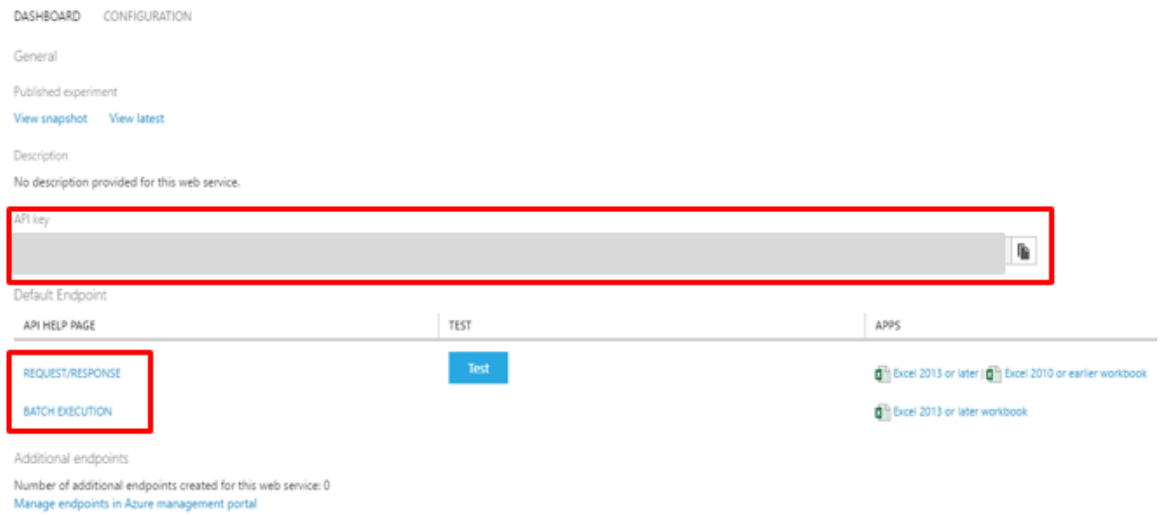
Microsoft NEAL ANALYTICS

Step 7.5 : Deploy the web service

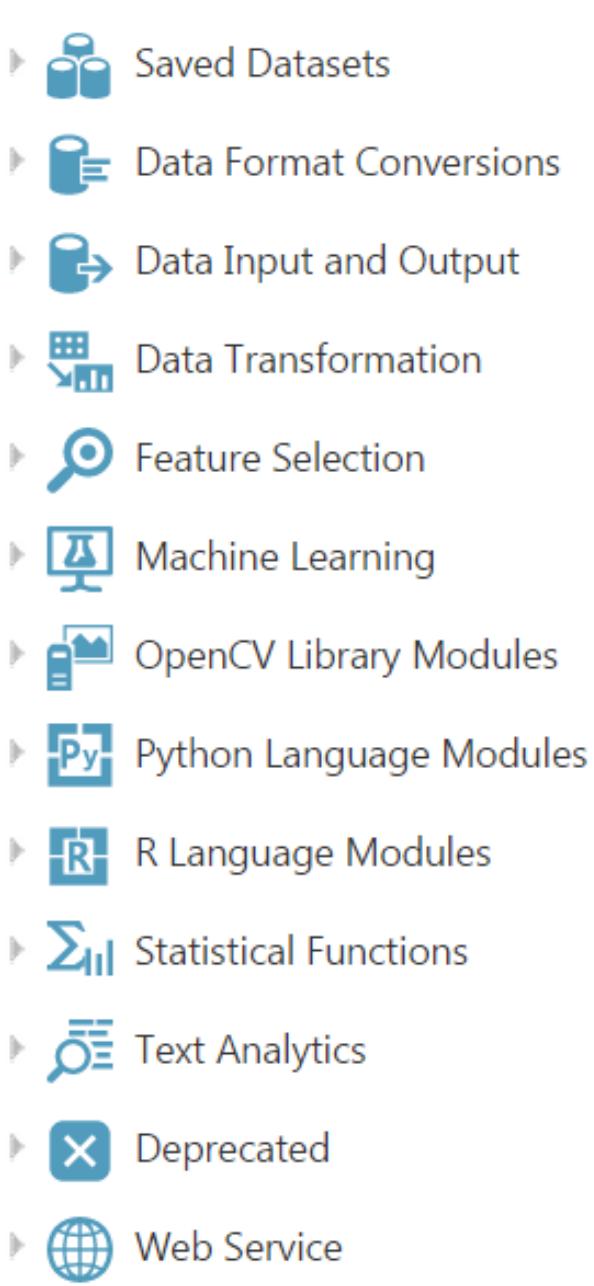
- An Azure Machine Learning web service can be consumed in two different ways
  - A Request-Response Service (RRS): is a low-latency, highly scalable web service used to provide an interface to the stateless models that have been created and deployed from an Azure Machine Learning Studio experiment. It enables scenarios where the consuming application expects a response in real-time.
  - A Batch Execution Service (BES) is a service that handles high volume, asynchronous, scoring of a batch of data records. A BES would be useful when responses are not needed immediately, such as for regularly scheduled scoring for individuals or internet of things (IOT) devices.

Step 7.6 : Deploy the web service

- There are four pieces of information that are needed to call either the RRS or BES service. Click a service to find the following links and information for both RRS and BES
  - The service API key, available on the services Dashboard
  - The service request URI, available on the API help page for the chosen service
  - The expected API request headers and body, available on the API help page for the chosen service
  - The expected API response headers and body, available on the API help page for the chosen service



Microsoft NEAL ANALYTICS

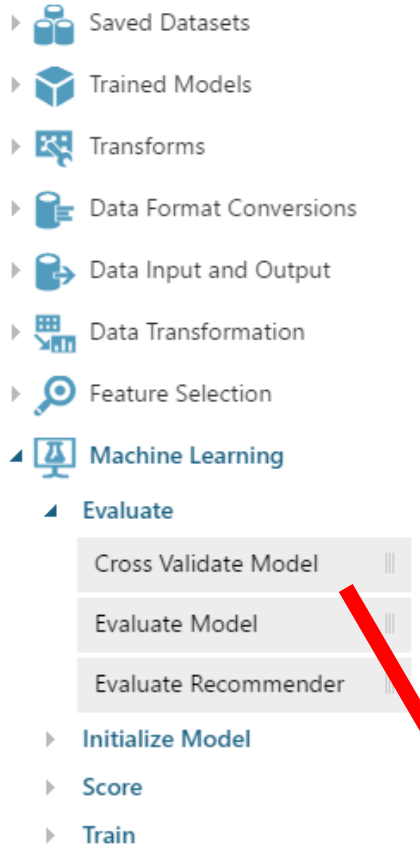


Step 8.1 : ADD CROSS VALIDATE MODEL MODULE TO YOUR EXPERIMENT

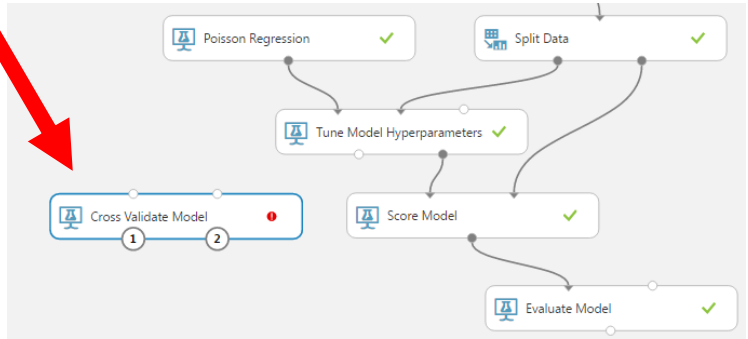
- AML modelling … a checklist approach
  - ☑ ~~Create new experiment~~
  - ☑ ~~Import data set~~
  - ☑ ~~Basic understanding of the Azure Machine Learning Canvas~~
  - ☑ ~~Data transformation~~
  - ☑ ~~Machine Learning~~
  - ☑ ~~Deploy the web service~~
  - ☐ Add Cross Validate Model

Microsoft NEAL ANALYTICS

Step 8.2 : Add Cross Validate Model



- Now that you have created a web service let's go back to our experiment and leverage cross validation to protect against overfitting.
- Open "Machine Learning" and then "Evaluate" from the navigation pane at the left
- Cross-Validate Model randomly splits the training data into 10 folds. A fold is a subset of the original dataset.

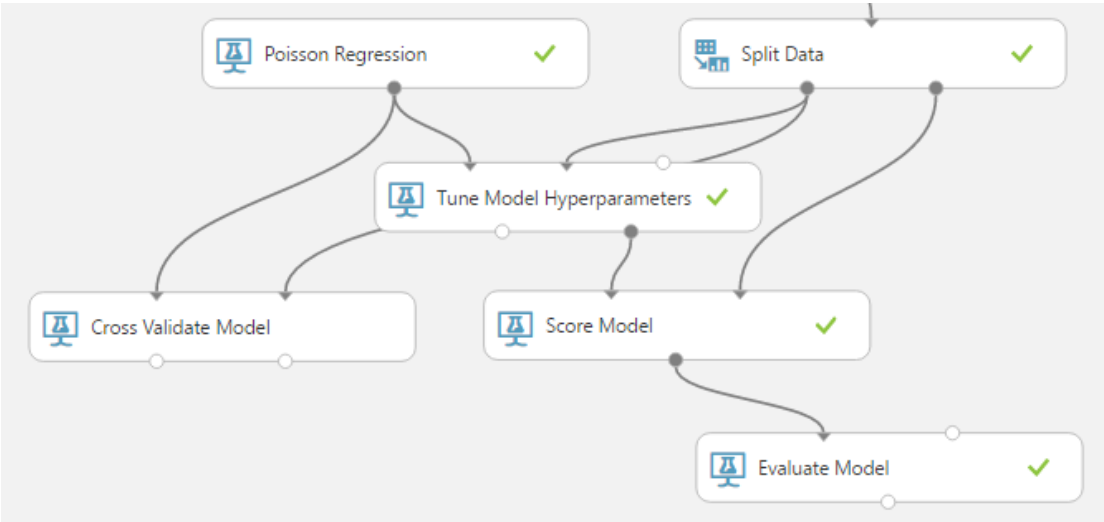


- In each iteration over the dataset, Cross-Validate Model uses one fold as a validation dataset, and uses the remaining folds to train a model.
- Each of the 10 models is tested against the data in all the other folds.

Microsoft NEAL ANALYTICS

Step 8.3 : Add Cross Validate Model

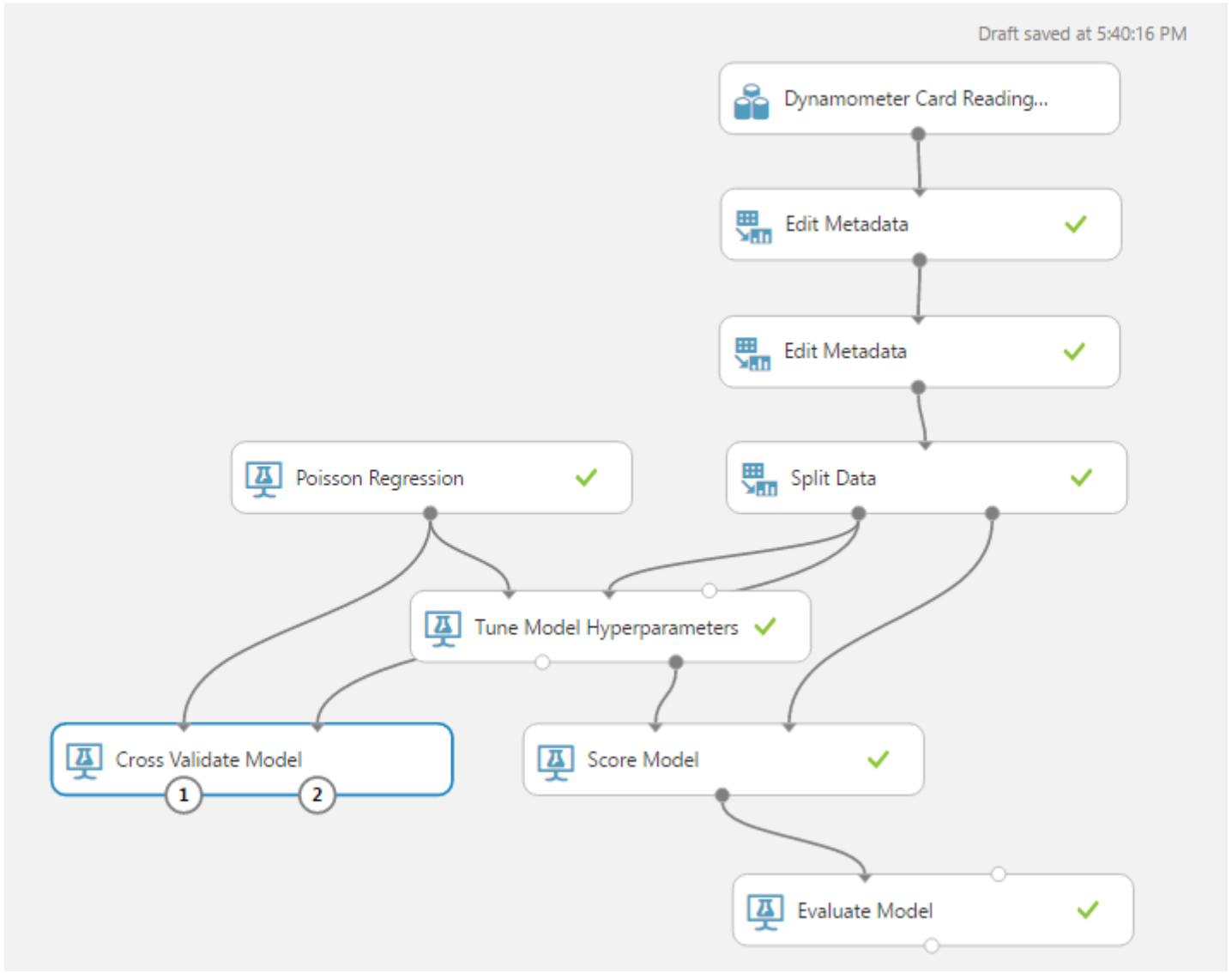
- Next, click and hold on the bottom middle circle of your “Poisson Regression” module and drag the line to the top left circle of the “Cross Validate Model” module.
- Similarly, click and hold on the bottom left circle of your “Split Data” module and drag the line to the top right circle of the “Cross Validate Model” module.

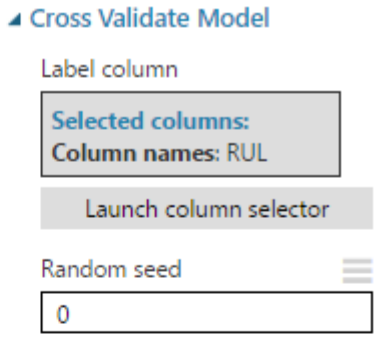


Microsoft

NEAL ANALYTICS

Step 8.4 : Add Cross Validate Model



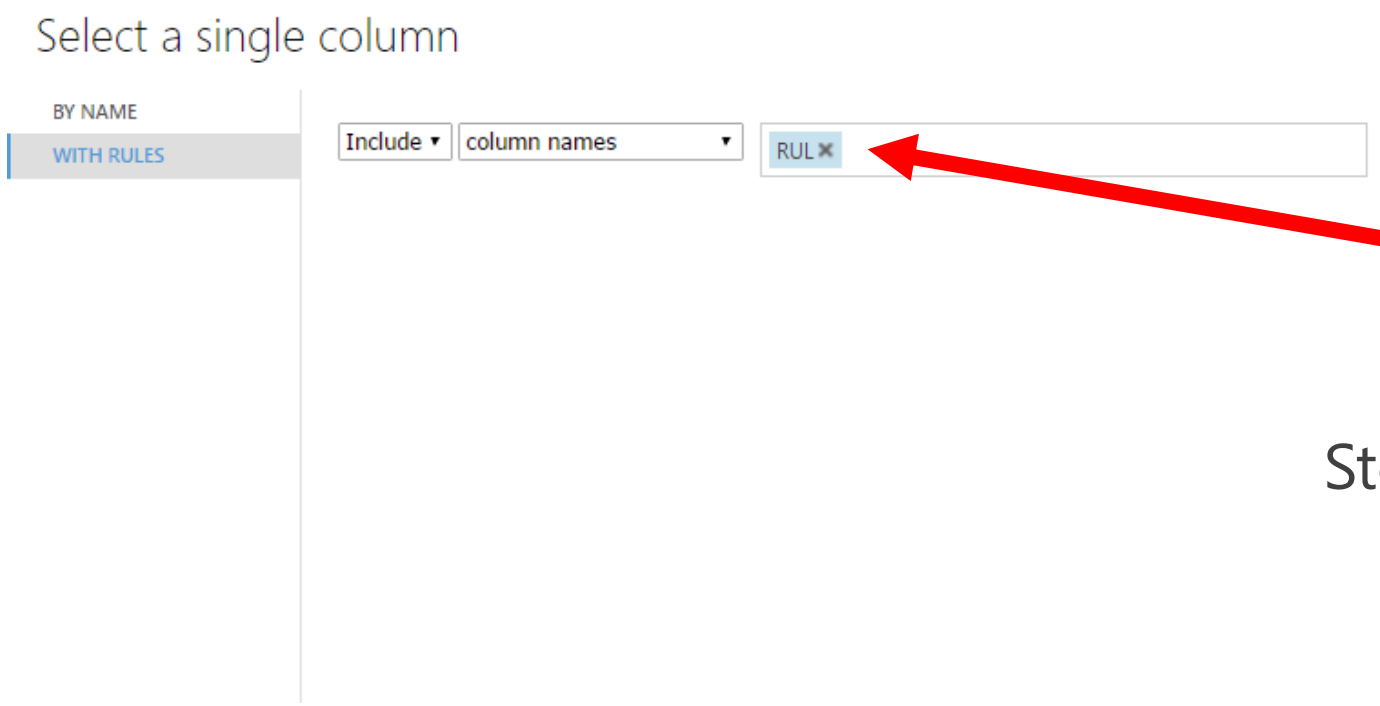


- Machine Learning
  - Click on Cross Validate Model
  - Then click on “Launch column selector”

Microsoft

NEAL ANALYTICS

Step 8.5 : Add Cross Validate Model



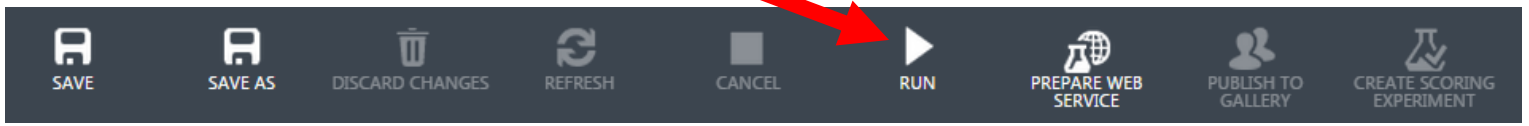
- Click here and a drop down list appears
- Select one variable at a time
  - RUL

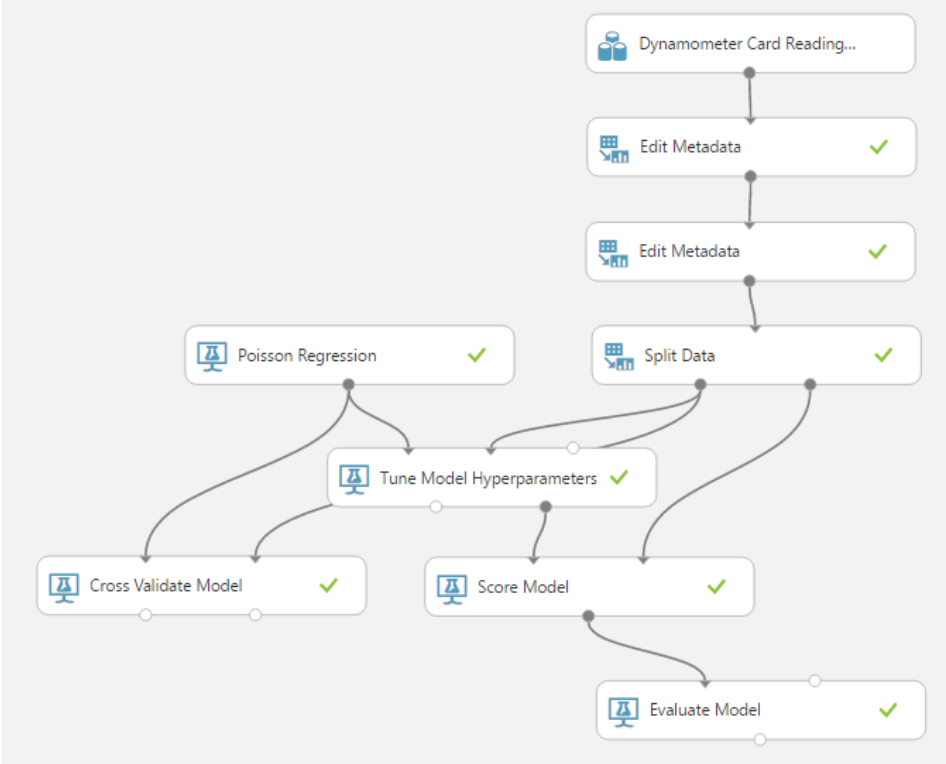
Step 1: Next, click on the check circle

Microsoft NEAL ANALYTICS

Step 8.6 : Run the Experiment

Click “Run”

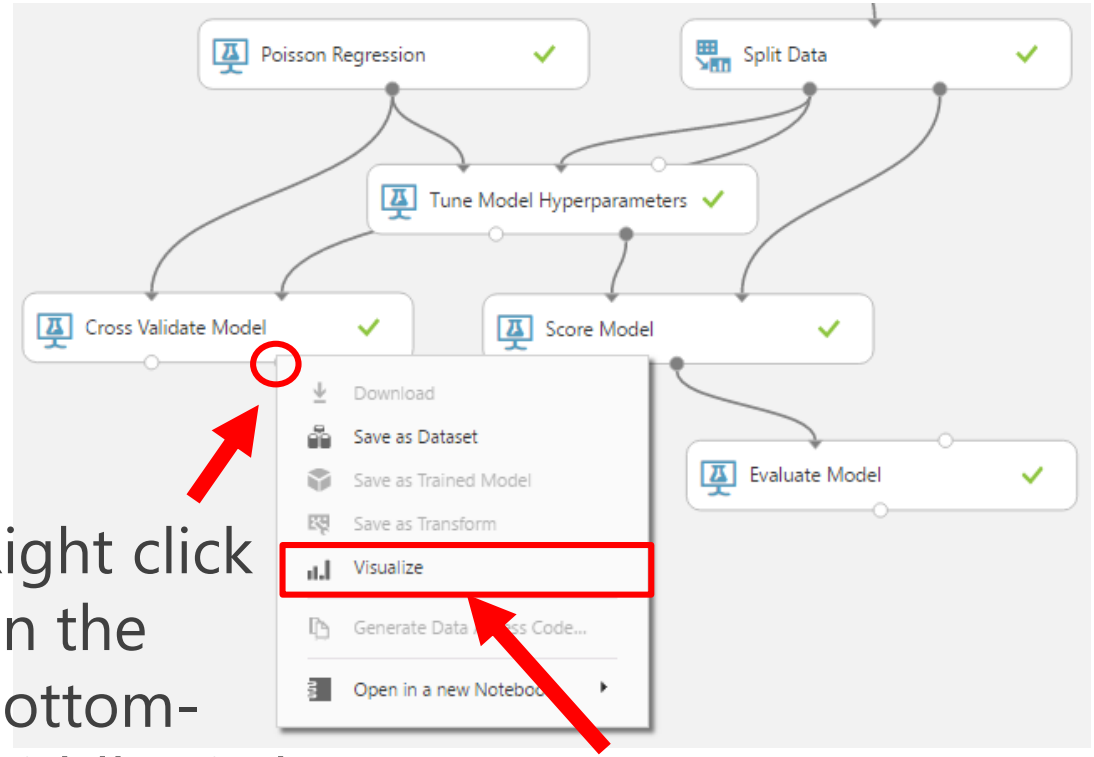




- After Azure Machine Learning “runs”, you will see a green check mark in all the modules
- All it took to train a machine learning model was drag and drop

Microsoft NEAL ANALYTICS

Step 8.7 : View the Results



Right click on the bottom-middle circle of the module

Then click "Visualize"

rows 12 columns 8

Fold Number	Number of examples in fold	Model	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error	Relative Squared Error	Coefficient of Determination
0	1568	Poisson Regression	95.254548	120.208794	0.959735	0.920502	0.079498
1	1568	Poisson Regression	98.259191	125.306205	0.947469	0.933576	0.066424
2	1568	Poisson Regression	96.44021	123.140742	0.963371	0.936875	0.063125
3	1568	Poisson Regression	94.807055	121.216648	0.969834	0.942225	0.057775
4	1568	Poisson Regression	95.269114	122.499986	0.968451	0.953006	0.046994
5	1568	Poisson Regression	97.005047	122.227326	0.966371	0.943058	0.056942
6	1569	Poisson Regression	98.558575	123.283834	0.995255	0.964106	0.035894
7	1569	Poisson Regression	99.281823	126.825869	0.951691	0.940865	0.059135
8	1568	Poisson Regression	98.030145	125.383687	0.960464	0.940604	0.059396
9	1569	Poisson Regression	97.13467	122.922173	0.963328	0.952806	0.047194
Mean	15683	Poisson Regression	97.004038	123.301526	0.964597	0.942762	0.057238
Standard Deviation	15683	Poisson Regression	1.544365	2.018065	0.012863	0.011947	0.011947

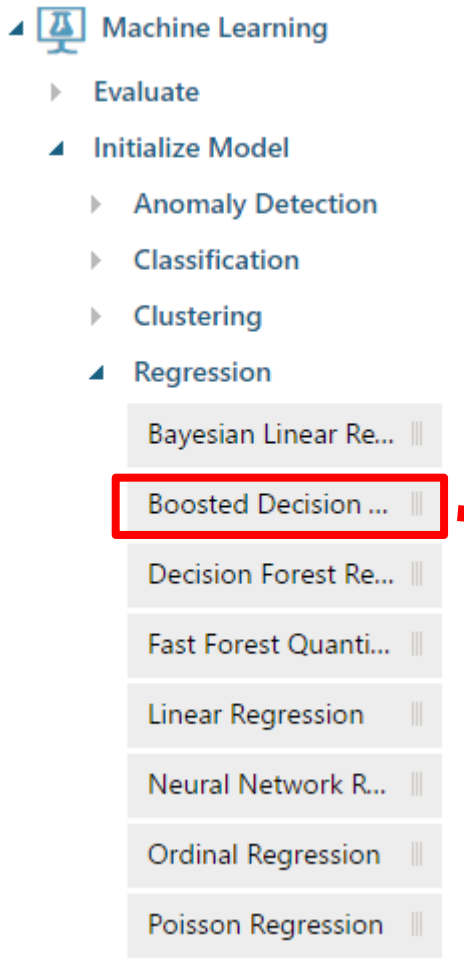
Microsoft NEAL ANALYTICS

Step 8.8 : View the Results

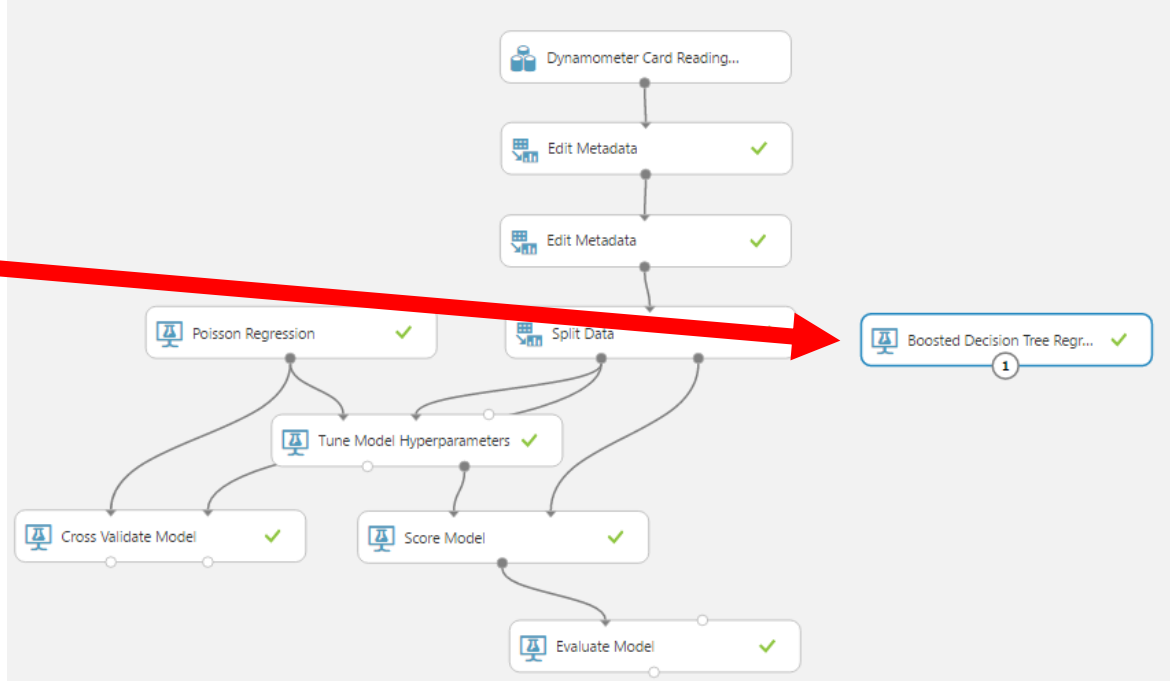
- A careful review and comparison of fold results can help you identify irregularities in the dataset or the model.
  - Fold number: An identifier for the current fold results.
  - Number of examples in fold: The count of rows assigned to each fold.
  - Model: The type of learner that was used in creating the models.
- The metrics that are provided depend on the type of model that you are evaluating.
  - Classification models: Precision, recall, F-score, AUC, average log loss, and training log loss metrics.
  - Regression models: Mean absolute error, root mean squared error, relative absolute error, and relative squared error metrics

If you have time left then let's try inserting new machine learning algorithm into the model and run the model. We can then compare the performance of the two models.

Step 8.9 : Adding the Boosted Decision Tree Regression



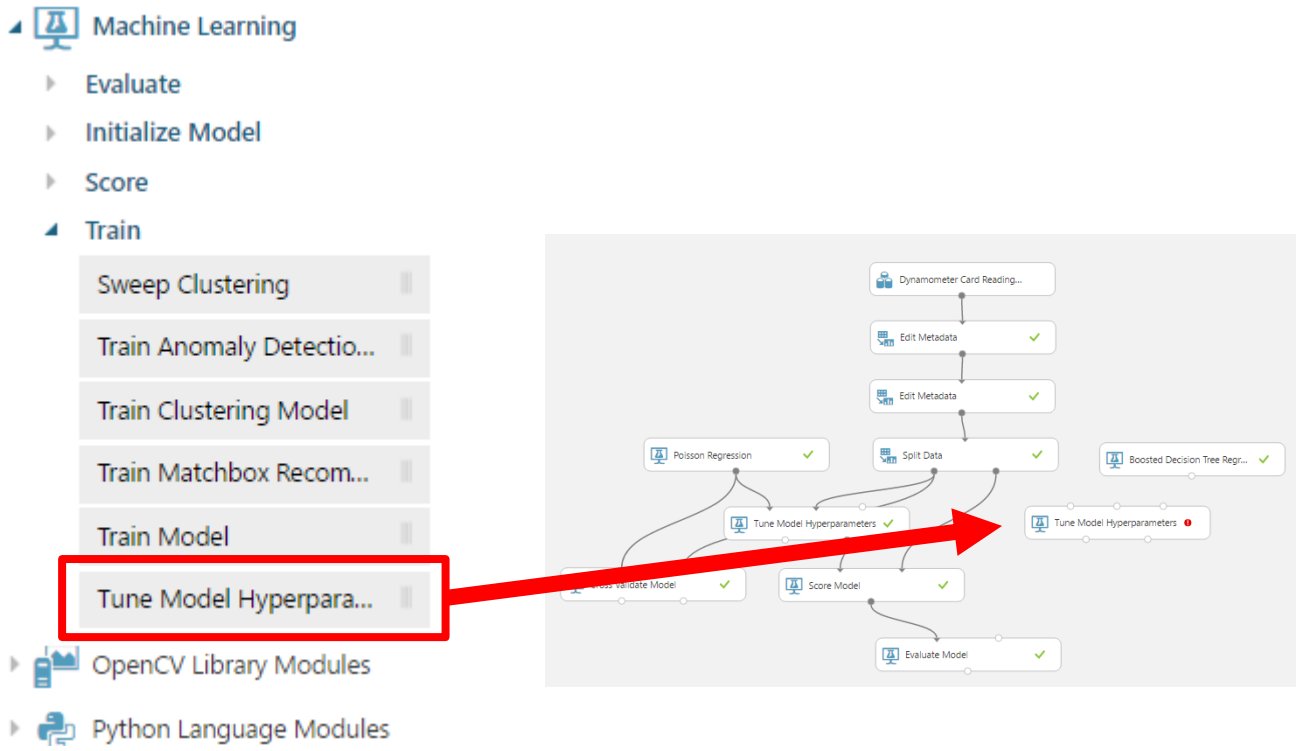
- Open “Machine Learning” -> “Initialize Model” -> “Classification” from the navigation pane at the left
- Drag “Boosted Decision Tree Regression” to the canvas



Microsoft

NEAL ANALYTICS

Step 8.10 : Q? How to develop the Machine Learning Algorithm?

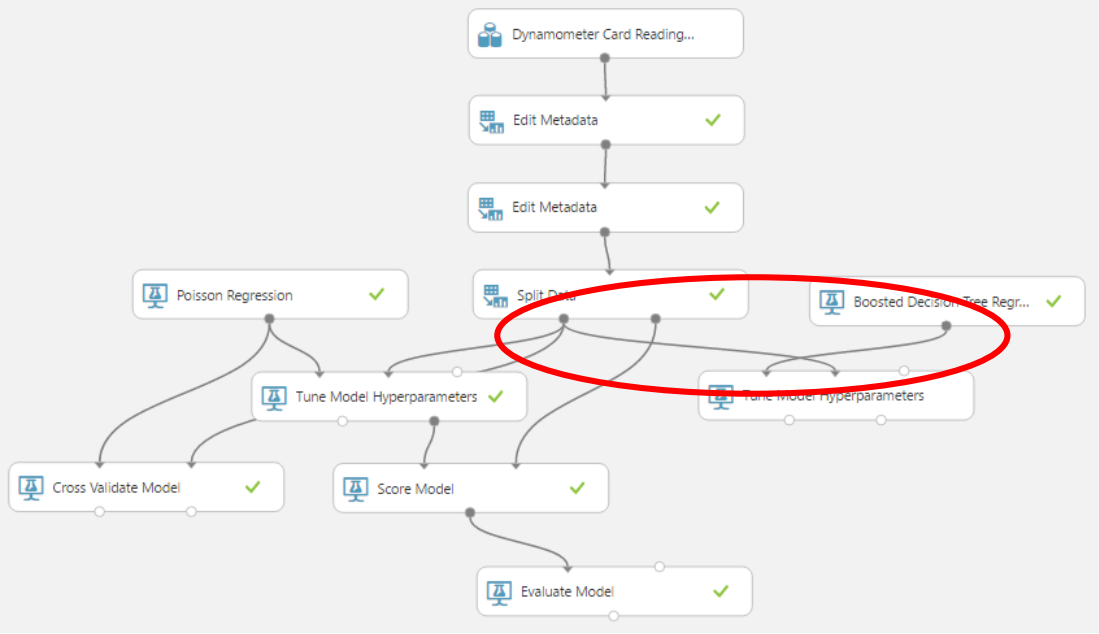


- Open “Machine Learning” and then “Train” from the navigation pane at the left
- Drag “Tune Model Hyperparameters” module to the canvas
  - “Tune Model Hyperparameters” perform a parameter sweep on the model to determine the optimum parameter settings

Microsoft

NEAL ANALYTICS

Step 8.11 : Q? How to develop the Machine Learning Algorithm?

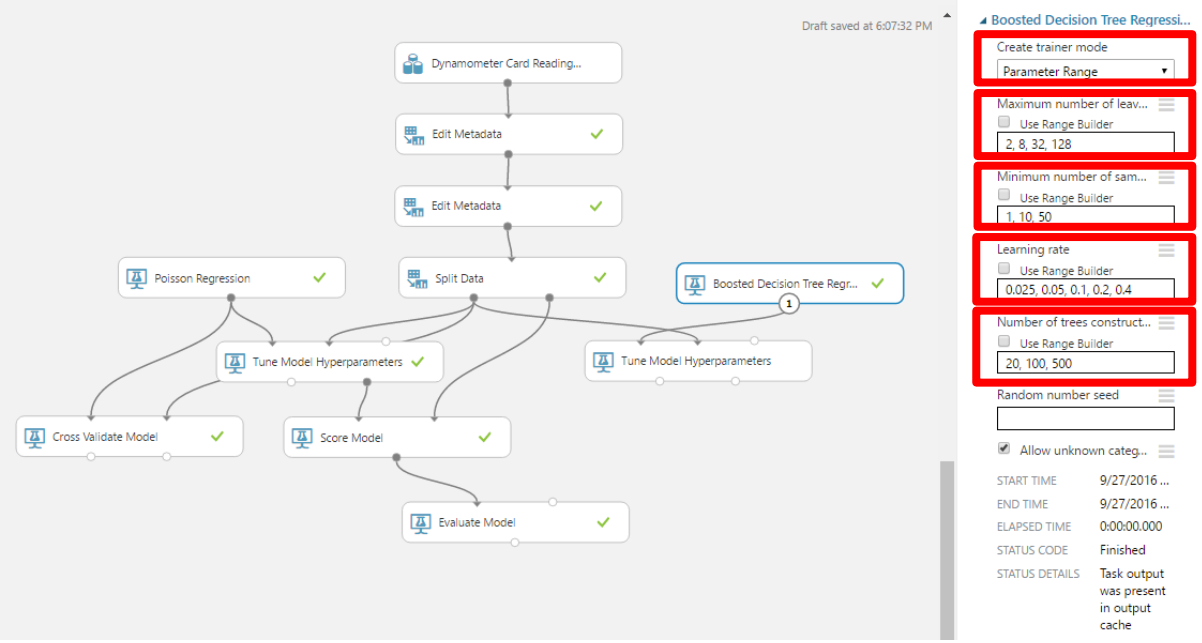


- Next, click and hold on the bottom middle circle of your “Boosted Decision Tree Regression” module and drag the line to the top left circle of the “Tune Model Hyperparameters” module.
- Similarly, click and hold on the bottom left circle of your “Split Data” module and drag the line to the top right circle of the “Tune Model Hyperparameters” module.

Microsoft

NEAL ANALYTICS

Step 8.12 : Q? How to develop the Machine Learning Algorithm

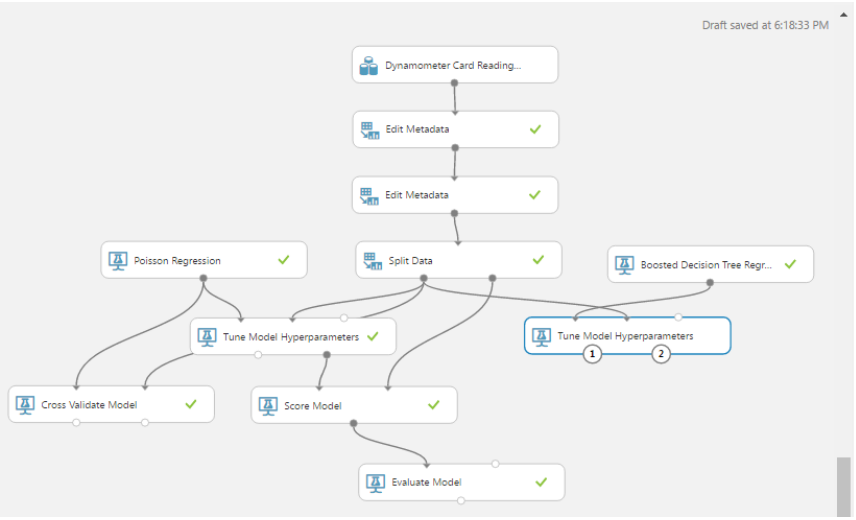


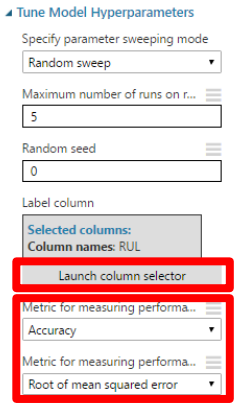
- Machine Learning
  - Chose “Parameter Range” from Create container mode
  - Set Maximum number of leaves per tree to “2, 8, 32, 128”
  - Set Minimum number of samples per leaf mode to “1, 10, 50”
  - Set Learning rate to “0.025, 0.05, 0.1, 0.2, 0.4”
  - Set Number of trees constructed to “20, 100, 500”

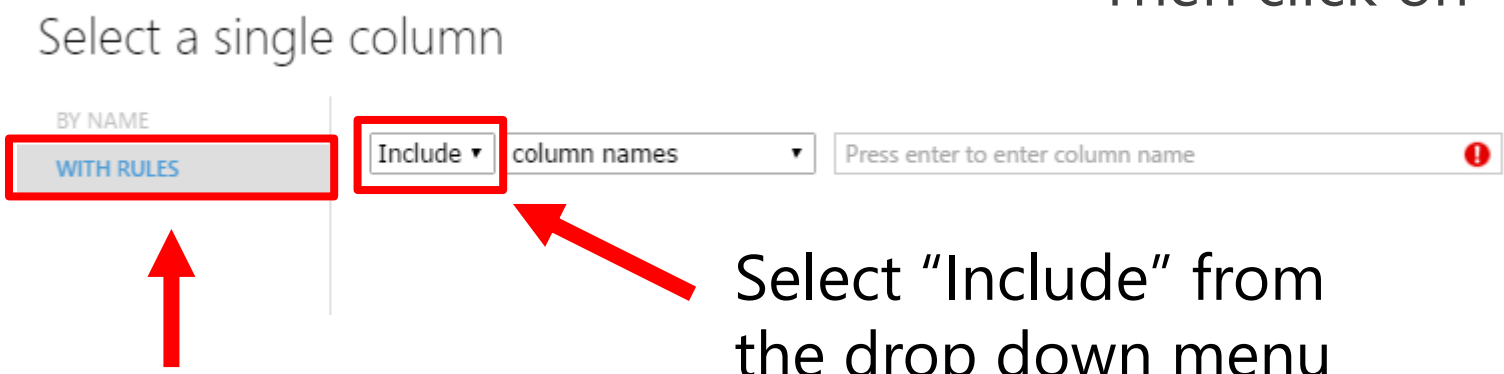
Microsoft NEAL ANALYTICS

Step 8.13 : Q? How to develop the Machine Learning Algorithm

- Machine Learning
  - Click on “Tune Model Hyperparameters”
  - Set Metric for measuring performance for classification to “AUC”
  - Set Metric for measuring performance for regression to “Root of mean squared error”
  - Then click on “Launch column selector”







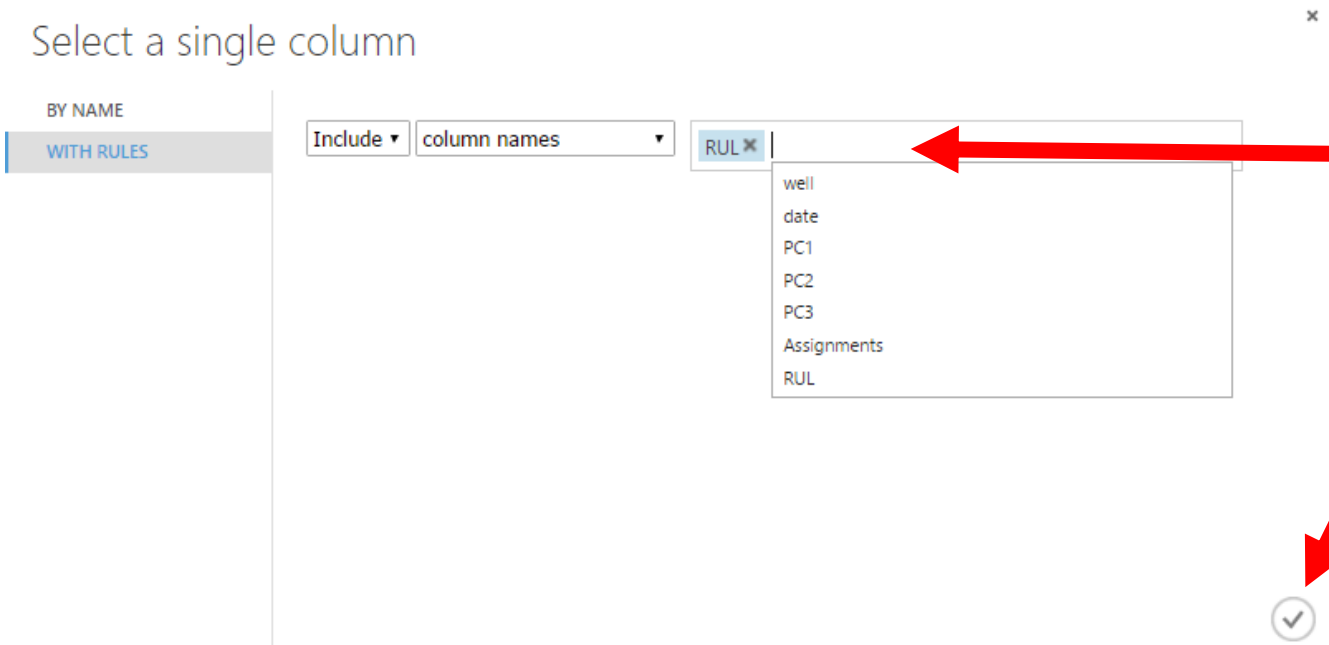
Select “With Rules”

Select “Include” from the drop down menu

Microsoft

NEAL ANALYTICS

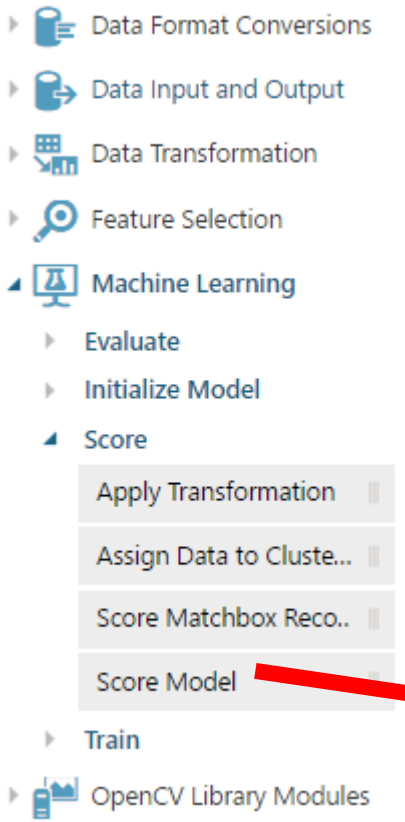
Step 8.14 : Q? How to develop the Machine Learning Algorithm?



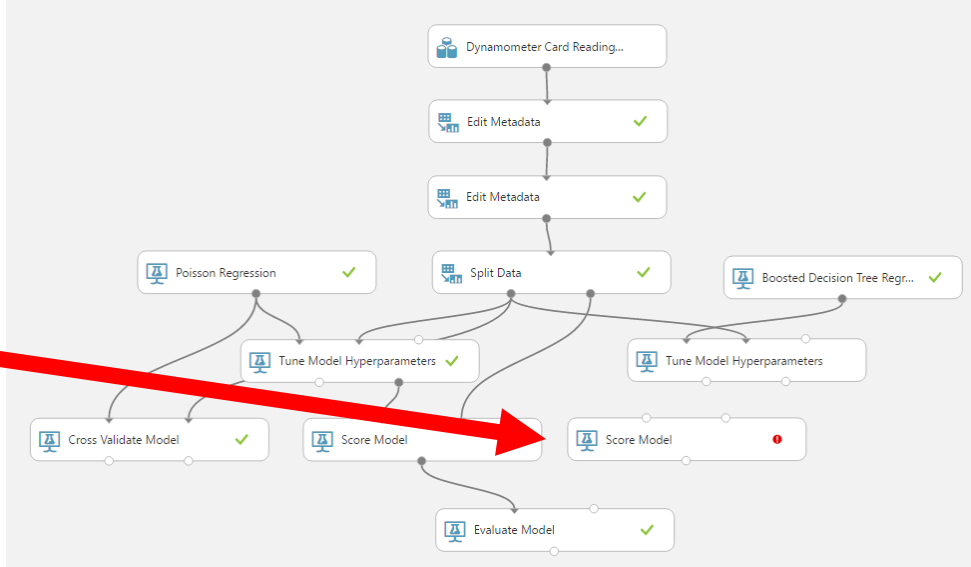
- Click here and a drop down list appears
- Select one variable at a time
  - RUL
- Next, click on the check circle

Microsoft NEAL ANALYTICS

Step 8.15 : Add Score Model

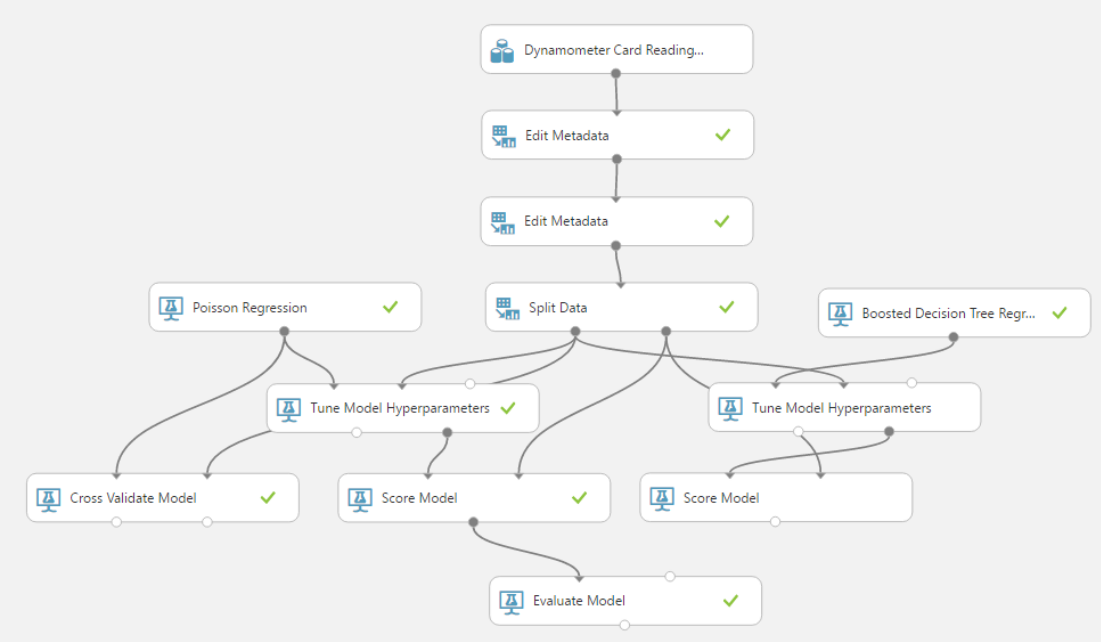


- Open “Machine Learning” and then “Score” from the navigation pane at the left
- Drag two “Score Model” to the canvas
  - “Score Model” scores a trained classification or a regression model



Microsoft NEAL ANALYTICS

Step 8.16 : Q? How to develop the Machine Learning Algorithm?

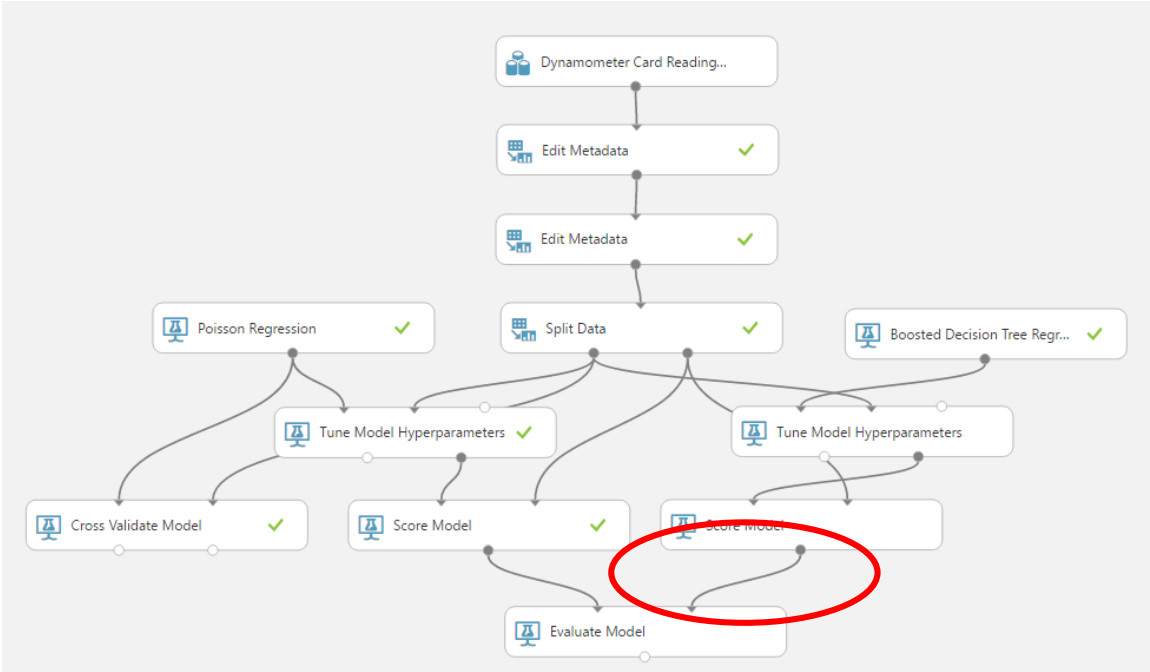


- Next, click and hold on the bottom right circle of your “Tune Model Hyperparameters” module and drag the line to the top left circle of the “Score Model” module.
- Similarly, click and hold on the bottom right circle of your “Split Data” module and drag the line to the top right circle of the “Score Model” module.

Microsoft NEAL ANALYTICS

Step 8.17 : Connect to the Evaluate Model

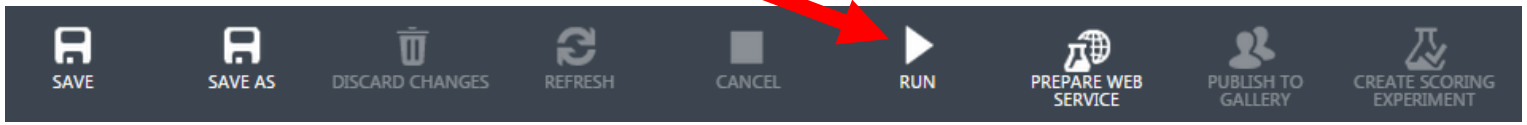
- Next, click and hold on the bottom middle circle of your “Score Model” module
- While holding down the mouse button, drag the line to the top left circle of “Evaluate Model” module

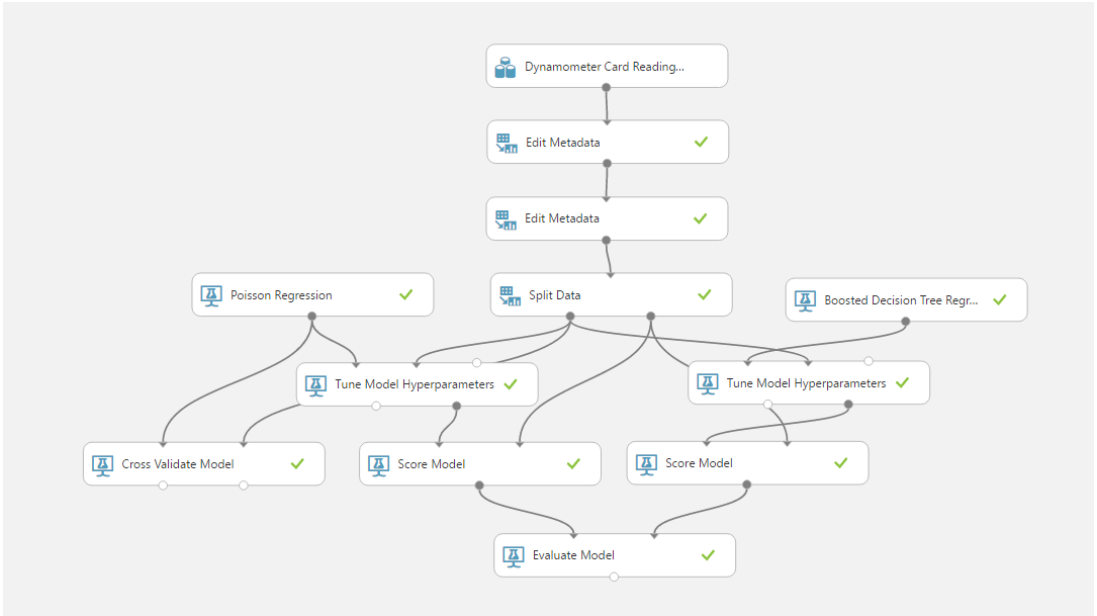


Microsoft NEAL ANALYTICS

Step 6.18 : Run the Experiment

Click “Run”



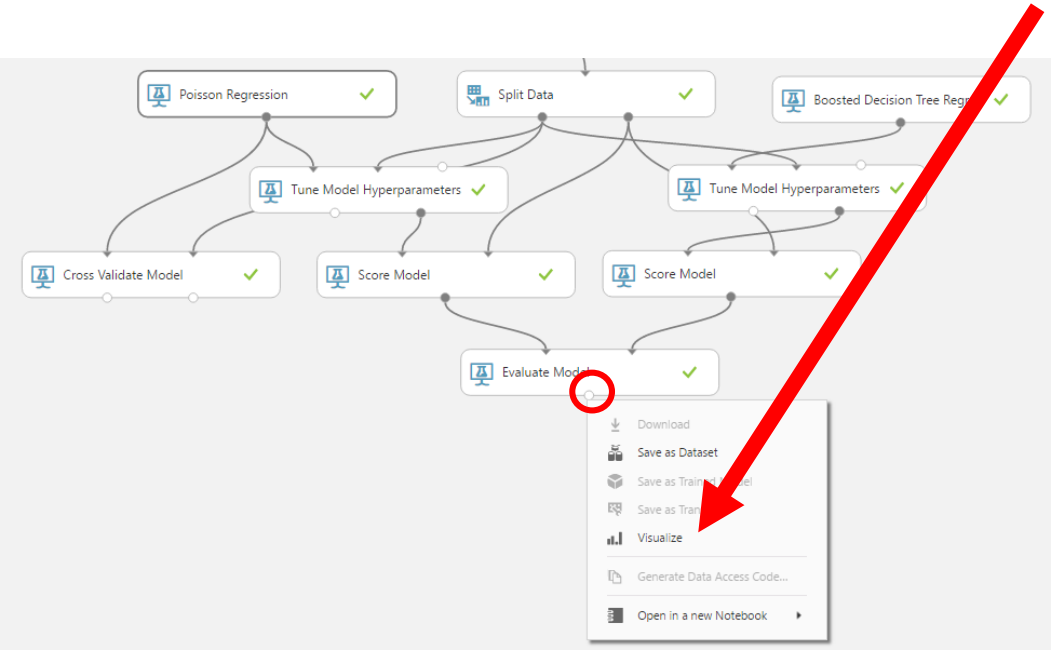


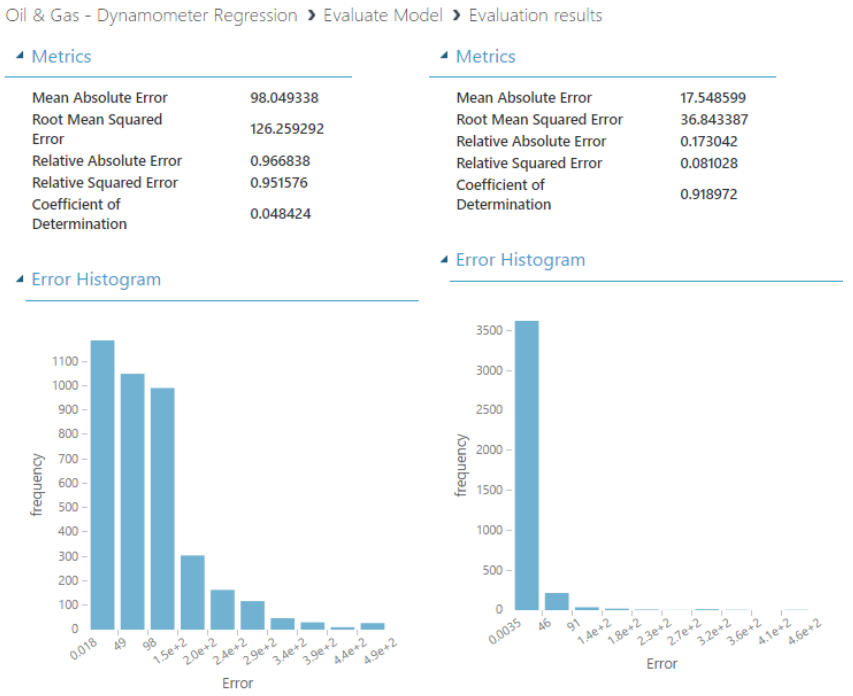
- After Azure Machine Learning “runs”, you will see a green check mark in all the modules
- All it took to train a machine learning model was drag and drop

Microsoft NEAL ANALYTICS

Step 8.19 : Visualize the Results

- Right click on the bottom-middle circle of “Evaluate Model” and select “Visualize”





Microsoft

NEAL ANALYTICS

Step 8.20 : Interpret the Results

- You can now compare between the two models using the following model performance metrics:
  - Mean Absolute Error
  - Root Mean Squared Error
  - Relative Absolute Error
  - Coefficient of Determination
- From the metrics displayed on the left, we can conclude that the model on the right is performing better than that on the left.

Microsoft NEAL ANALYTICS