|  |
| --- |
|  |
| Developer Journey: Analytics driven Customer Behavior Ranking for Retail Promotions using POS data (WIP) |
| This Journey helps Developers and SMEs how to create and Run a promotion using POS data |
|  |
|  |
|  |

|  |
| --- |
|  |

Contents

[1 Overview and Goal 2](#_Toc517448454)

[2 Flow (Sample - To update) 2](#_Toc517448455)

[3 Included Components 2](#_Toc517448456)

[4 Featured Technologies 3](#_Toc517448457)

[5 Watch the Video 3](#_Toc517448458)

[6 Steps 3](#_Toc517448459)

[6.1 Sign-up for Watson studio 3](#_Toc517448460)

[6.2 Create IBM Cloud services 4](#_Toc517448461)

[6.3 Create the Jupyter Notebook 5](#_Toc517448462)

[6.4 Add the data and configuration file 6](#_Toc517448463)

[6.5 Run the notebook 10](#_Toc517448464)

[6.6 View the results 11](#_Toc517448465)

[6.7 Dashboard 12](#_Toc517448466)

[7 Further enhancements: 12](#_Toc517448467)

[8 Troubleshooting 12](#_Toc517448468)

[9 License 12](#_Toc517448469)

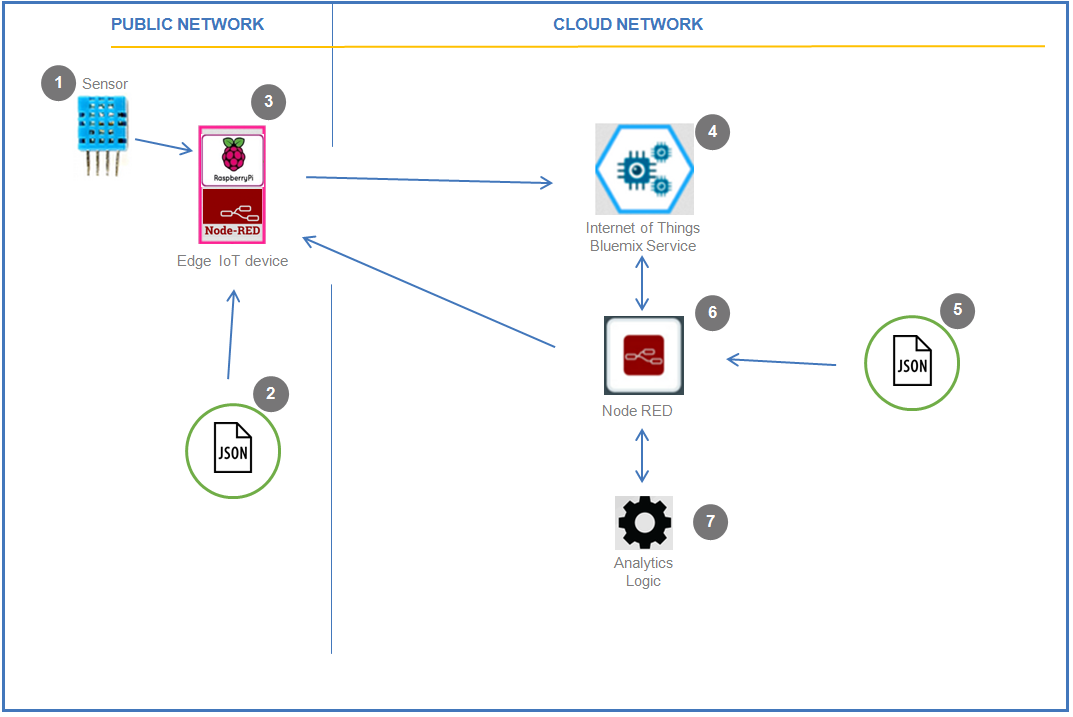
# Overview and Goal

Segment Customers based on their Point of Sale (PoS) buying behavior in Retail. Point of Sale data is captured at the counter in every Retail outlet either in FMCG or CPG.

This Code pattern is a walkthrough of the steps involved in using this PoS data to Segment Customers based on their buying behavior. Segmentation techniques are used to Rank Customers. The ranked Customer data is further used to generate Dashboards that delivers insights on the Customer buying Metrics.

IBM cloud storage, Python notebooks in IBM Data science experience, IBM Cognos will be used to implement the above Use case

# Flow (Sample - To update)



</doc/images/iea\_arch\_flow.png>

1. Steps

# Included Components

* **[IBM Cloud]**(https://console.bluemix.net/catalog/): IBM's innovative cloud computing platform or IBM Cloud in short (formerly Bluemix) combines platform as a service (PaaS) with infrastructure as a service (IaaS) and includes a rich catalog of cloud services that can be easily integrated with PaaS and IaaS to build business applications rapidly.
* **[IBM Data Science Experience]**(https://www.ibm.com/bs-en/marketplace/data-science-experience): Analyze data using Python, Jupyter Notebook and RStudio in a configured, collaborative environment that includes IBM value-adds, such as managed Spark.

# Featured Technologies

* **[Analytics]**(https://developer.ibm.com/code/technologies/analytics?cm=IBMCode-\_--\_-featured\_technologies-\_-analytics):Finding patterns in data to derive information.
* **[Data Science]**(https://developer.ibm.com/code/technologies/data-science?cm=IBMCode-\_--\_-featured\_technologies-\_-data-science):Systems and scientific methods to analyze structured and unstructured data in order to extract knowledge and insights.
* **[Cognos??]()**

# Watch the Video

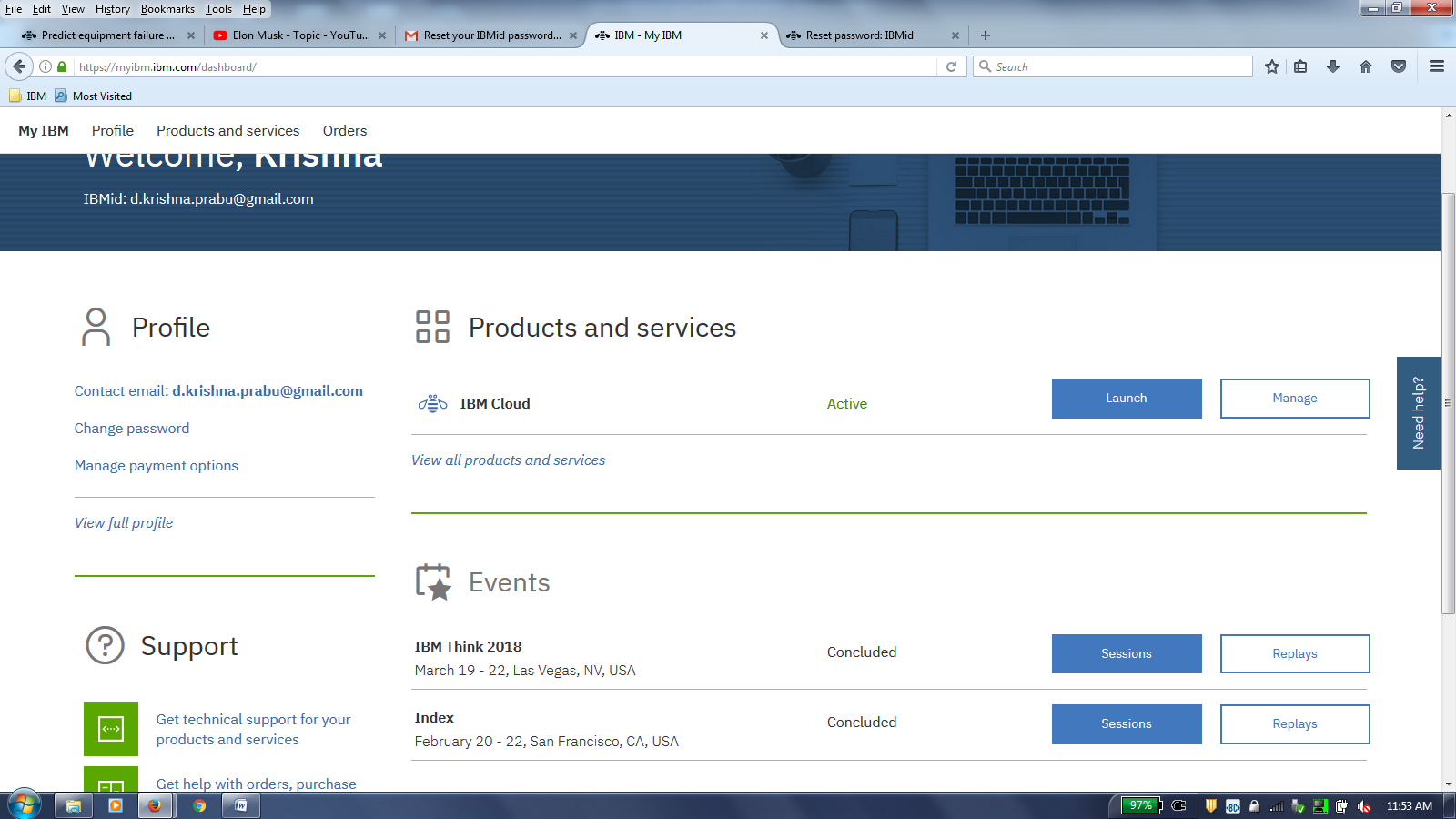
<Video - WIP>

# Steps

1. Signup for Watson Studio
2. Create IBM Cloud services
3. Create the Jupyter Notebook
4. Add the data and configuration file
5. Run the Notebook
6. View the results
7. Save Segmentation data
   1. Update Dashboard steps

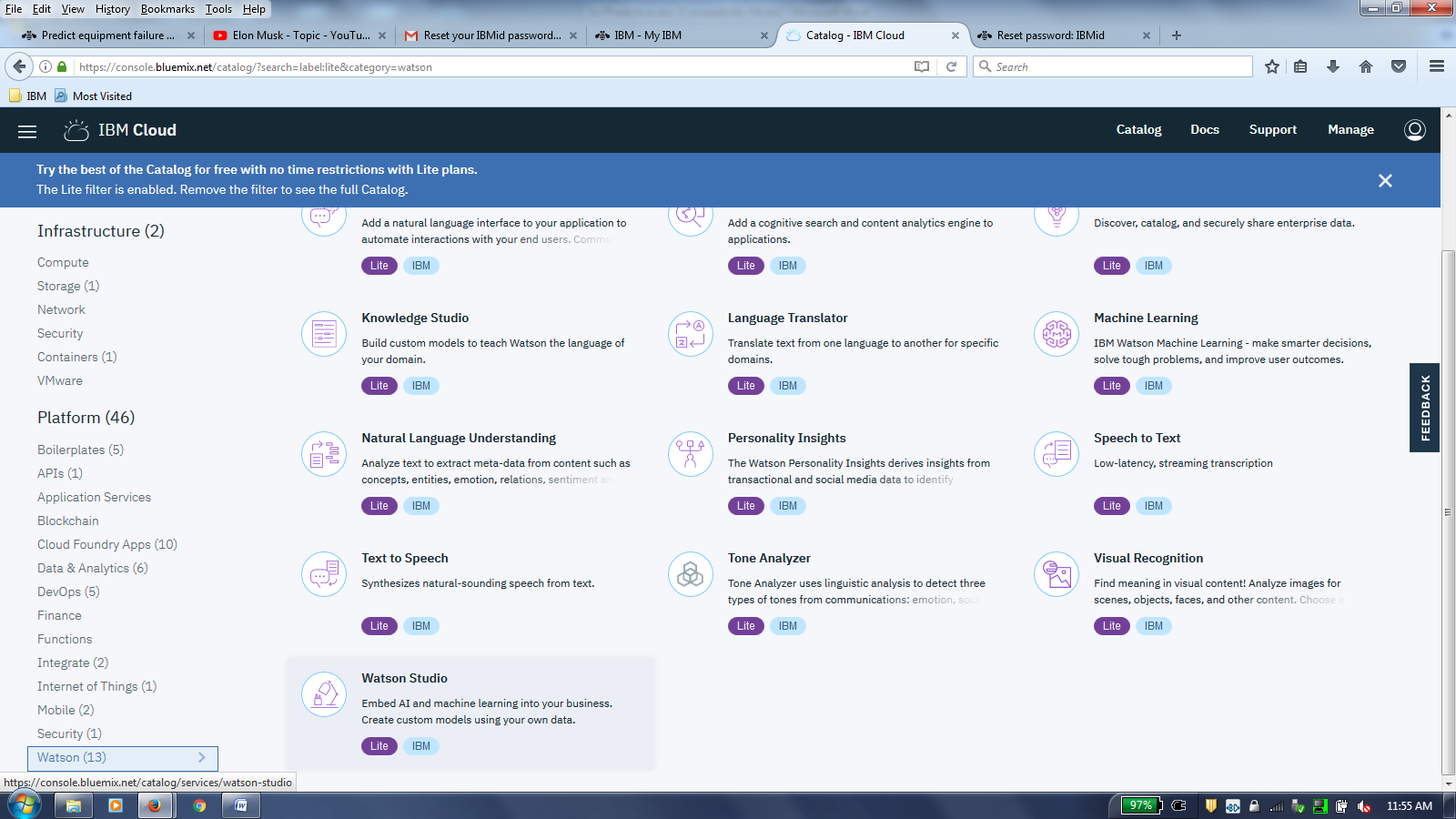
## Sign-up for Watson studio

Register at: <https://www.ibm.com/cloud/watson-studio>



**IBM Cloud -> Launch**

**Catalog -> Watson -> Watson Studio**

****

## Create IBM Cloud services

**Check Repository access:**

[https://github.com/IBM/commerce\_pos\_analytics](https://github.com/IBM/commerce_pos_analytics%20)

Download necessary files for Code + Configuration + Sample Data

[https://github.com/IBM/commerce\_pos\_analytics/blob/master/notebook/watson\_iotfailure\_prediction.ipynb](https://github.com/IBM/iot-predictive-analytics/blob/master/notebook/watson_iotfailure_prediction.ipynb)

<https://github.com/IBM/commerce_pos_analytics/blob/master/notebook/customer_segmentation_promo.ipynb>

[https://github.com/IBM/commerce\_pos\_analytics/blob/master/data/iot\_sensor\_dataset.csv](https://github.com/IBM/iot-predictive-analytics/blob/master/data/iot_sensor_dataset.csv)

[https://github.com/IBM/commerce\_pos\_analytics/blob/master/data/iot\_sensor\_dataset.txt](https://github.com/IBM/iot-predictive-analytics/blob/master/data/iot_sensor_dataset.txt)

## Create the Jupyter Notebook

First create a new project in Watson Studio. Follow the detailed steps provided in the [IBM online documentation for Watson Studio Project creation](https://datascience.ibm.com/docs/content/analyze-data/creating-notebooks.html), or watch a video on using [Watson Studio to create a project](https://youtu.be/QSttEjcHtl0).

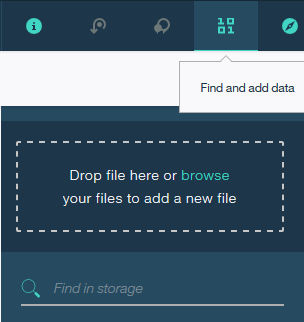
In [Watson Studio](http://dataplatform.ibm.com/):

Use the menu on the top to select Projects and then Default Project. Click on Add notebooks (upper right) to create a notebook.

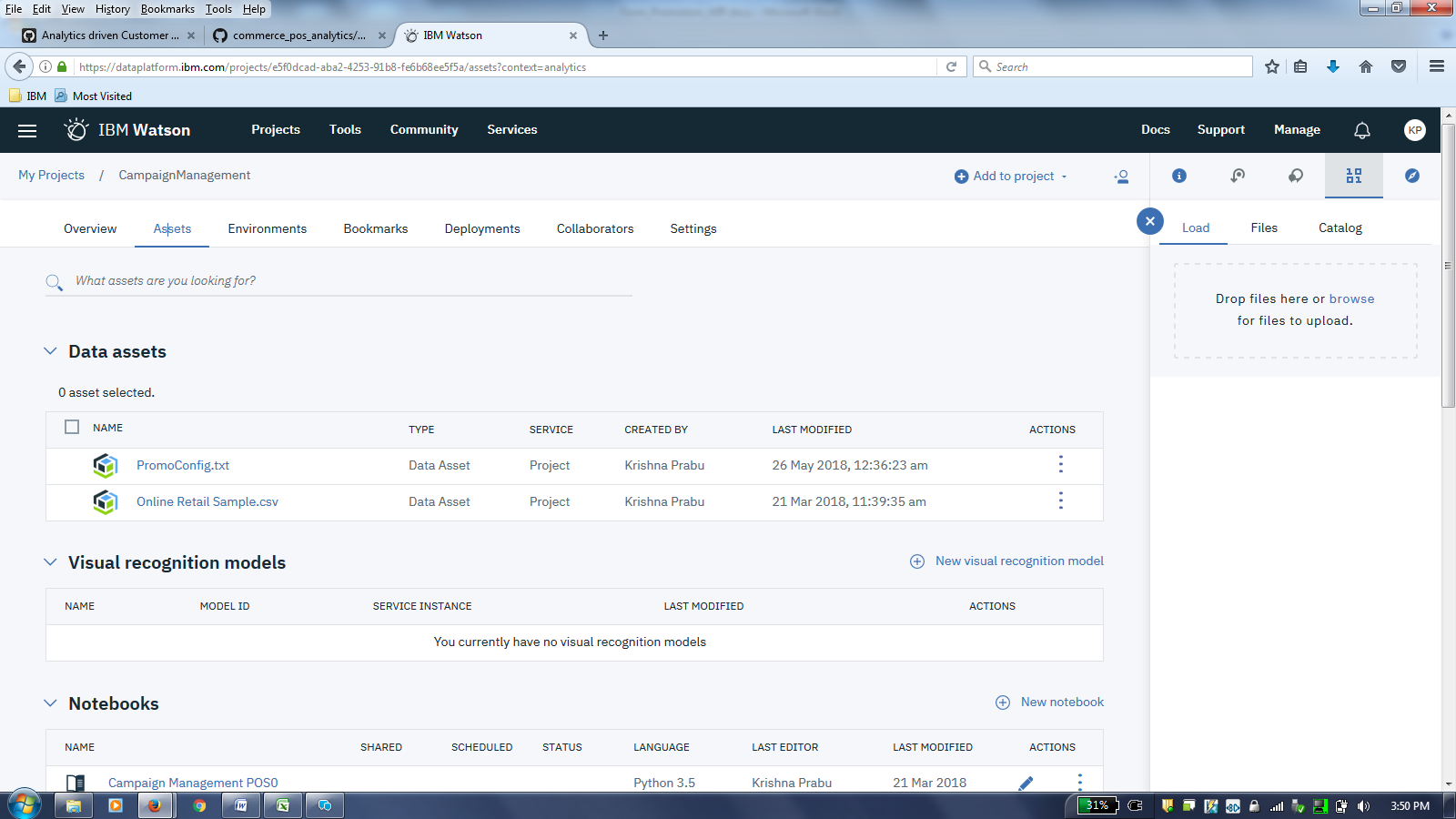
* Select the From URL tab.
* Enter a name for the notebook.
* Optionally, enter a description for the notebook.
* Enter this Notebook URL: <https://github.com/IBM/commerce_pos_analytics/blob/master/notebook/customer_segmentation_promo.ipynb>
* Select the free Anaconda runtime.
* Click the Create button.
* Upload the sample .json, .txt Watson Studio configuration file to Watson Studio Object storage from URL below:  
  https://github.com/IBM/commerce\_pos\_analytics/blob/master/data/Online Retail Sample.csv

<https://github.com/IBM/commerce_pos_analytics/blob/master/configuration/PromoConfig.txt>

* To upload these files in Watson Studio object storage,
  + Go to "My Projects -> Your Project Name"
  + Click on the Find and add data icon on top ribbon
  + Select the file and upload one by one

[](https://github.com/IBM/iot-predictive-analytics/blob/master/doc/images/ipredict_upload_file_sample.png)

Now you must be able to see the uploaded files listed under "My Projects -> Your Project Name -> Assets" table



## Add the data and configuration file

Fix-up configuration parameter .json file name and values:

Go to the Notebook in Watson Studio by navigating to "My Projects -> CustomerSegmentation" Under Assets tab, under Notebooks section you will find the Notebook you just imported Click on the Click to Edit and Lock icon to edit the notebook in Jupyter notebook in Watson Studio

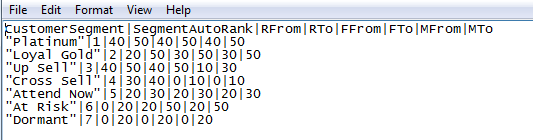
For more details on Creating, Editing and sharing notebooks in IBM Watson Studio refer to [Notebooks Watson Studio documentation](https://datascience.ibm.com/docs/content/analyze-data/notebooks-parent.html)

You can now update the variables that refer to the .json configuration file name in the Python - Jupyter Notebook. This step is necessary only if you had changed the name of the sample .json configuration file you had uploaded earlier for any reason.

The default .json configuration file, you uploaded earlier works without any changes with the Sample data supplied.

But, if you have a data file with different column names and wanted to customise the model to use these column names you can do so. Below are the steps to configure the .json configuration file to train the Predictive models using your custom data file.

1. Download the [PromoConfig.txt configuration file](https://github.com/IBM/iot-predictive-analytics/blob/master/configuration/iotpredict_config.json) to your Computer local folder
2. Open a local copy of the .json file in text editor like notepad and edit the [Watson Studio configuration .json file](https://github.com/IBM/iot-predictive-analytics/blob/master/configuration/iotpredict_config.json)
3. Update the column values to suit your requirements and save the PromoConfig.txt file. Retain the rest of the format and composition of the .txt file
4. Delete the copy of 'PromoConfig.txt' in Watson Studio data store if one is already uploaded by you earlier.
5. Now upload your local edited copy of ' PromoConfig.txt ' by following the steps in section 6.3 above.



The descriptions of the columns that can be configured are as below.

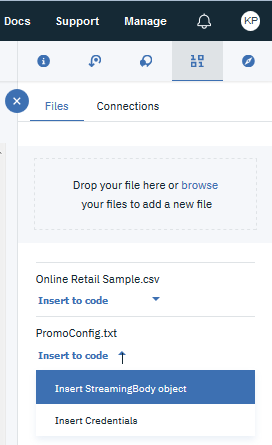
i. Customer Segment: Name of the Customer segement tobe assigned  
ii. Segment Auto Rank: Corresponding Ranking for the Customer segment. Used only if the Customer segmentation is run as AUTO  
iii. RFrom, RTo: Range of Recency value in dates. This is a numeric field with range from 0-100. Used only if the Customer segmentation mode is run as MANUAL

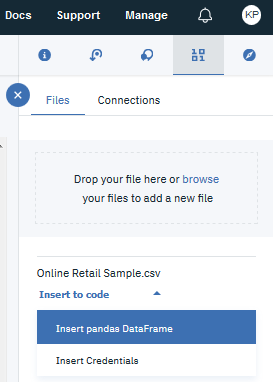
iv. FFrom, FTo: Range of Frequency value in dates. This is a numeric field with range from 0-100. Used only if the Customer segmentation mode is run as MANUAL

v. MFrom, MTo: Range of Shopping value. This is a numeric field with range from 0-100. Used only if the Customer segmentation mode is run as MANUAL

* The cell 3.1.2 of the Jupyter Notebook has a function definition which is shown for illustration purposes.  
  These details that have user specific security details are striked out in the screenshots shown below.  
  This function will need to be recreated with your user specific access credentials ang target data object.  
  In order to do that first delete all pre existing code in cell 3.1.2 of the notebook.

Note: The .pynb file that you imported have code with dummy credentials for illustration purposes.  
This needs to be replaced by your user specific function with your own access credentials.  
The steps below explain that.

1. In section 3.1.2 of Jupyter Notebook (not this README file), Insert (replace) your own Object storage file credentials to read the 'PromoConfig.txt' configuration file  
   

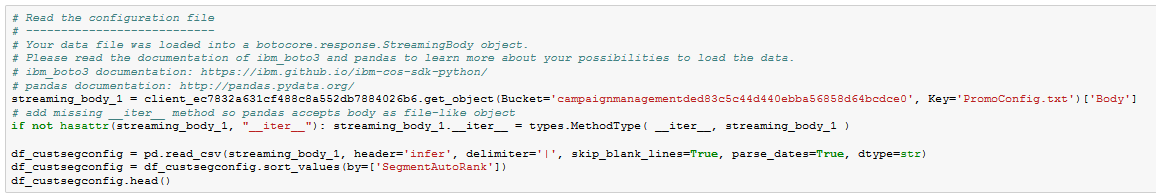


1. This step will auto generate a function that reads the data followed by a call to the function as below:

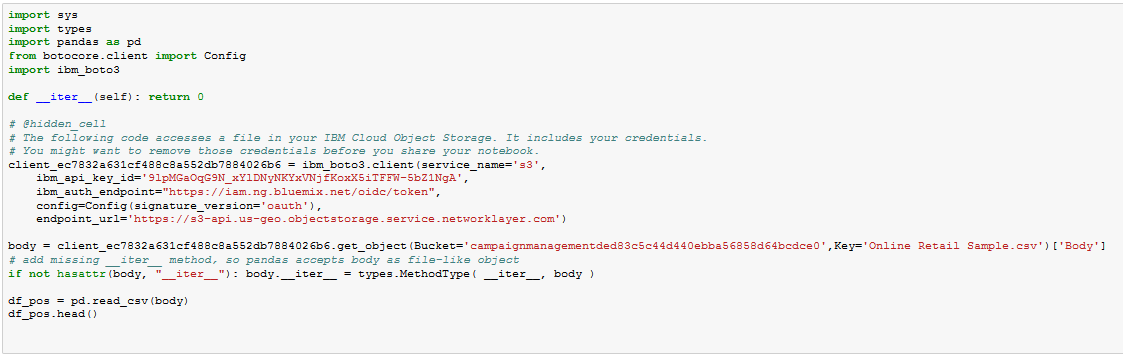
Refer to screen shot above for details.  
For more details, revisit the documentation help links provided in beginning of section 5.2.2

#### Add the data and configuration to the notebook

Use Find and Add Data (look for the 10/01 icon) and its Connections tab. You must be able to see your database connection created earlier. From there you can click Insert to Code under the 'Data connection' list and add ibm DBR code with connection credentials to the flow.

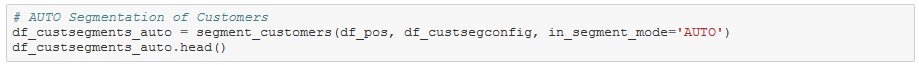


Note: If you don't have your own data and configuration files, you can reuse our example in the "Read IoT Sensor data from database" section. Look in the /data/iot\_sensor\_dataset.csv directory for data file.



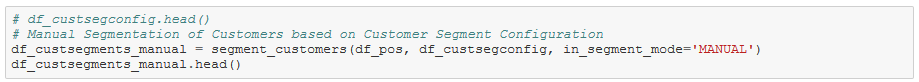
The Customer segmentation can be run in 2 modes

1. By setting the mode to “AUTO”



In AUTO mode Customer segmentation algorithm will automatically do the segmentation and ranking based on Customer Buying behavior

1. By setting the mode to “MANUAL”



In MANUAL mode, the range of values mentioned for Customer metrics will be used for applying the segmentation.

## Run the notebook

When a notebook is executed, what is actually happening is that each code cell in the notebook is executed, in order, from top to bottom.

Each code cell is selectable and is preceded by a tag in the left margin. The tag format is In [x]:. Depending on the state of the notebook, the x can be:

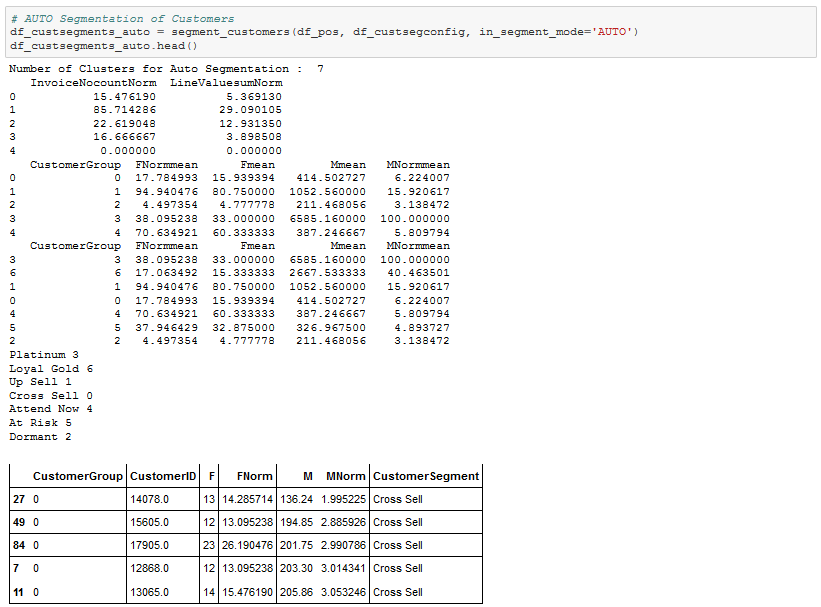
* A blank, this indicates that the cell has never been executed.
* A number, this number represents the relative order this code step was executed.
* A \*, this indicates that the cell is currently executing.

There are several ways to execute the code cells in your notebook:

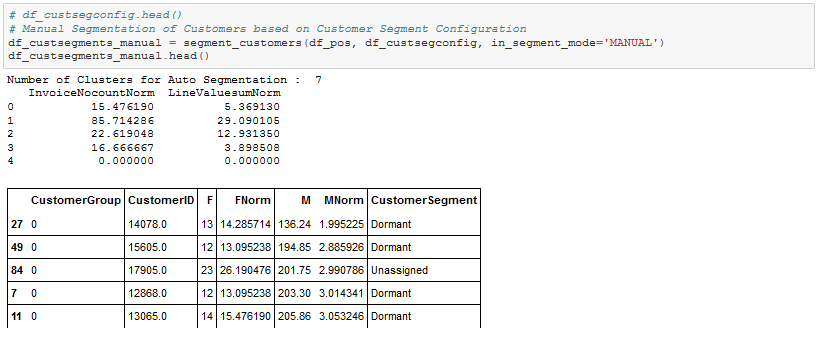
* One cell at a time.
  + Select the cell, and then press the Play button in the toolbar.
* Batch mode, in sequential order.
  + From the Cell menu bar, there are several options available. For example, you can Run All cells in your notebook, or you can Run All Below, that will start executing from the first cell under the currently selected cell, and then continue executing all cells that follow.
* At a scheduled time.
  + Press the Schedule button located in the top right section of your notebook panel. Here you can schedule your notebook to be executed once at some future time, or repeatedly at your specified interval.

## View the results

Results for AUTO Customer segmentation



Results for MANUAL Customer segmentation



## Dashboard

<to be merged with - Srikanth>

# Further enhancements:

Inputs from Raspberry Pi GPIO pins can be used for reading temperature from a external Temperature sensor attached

Outputs to Raspberry Pi GPIO pins can be used for triggering action

# Troubleshooting

See Debugging.md[https://github.com/IBM/commerce\_pos\_analytics/blob/master/DEBUGGING.md]

# License

Apache 2.0[https://github.com/IBM/commerce\_pos\_analytics/blob/master/LICENSE]

<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<End>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>