

HOL Guide for Enterprise Risk Analysis

USING DATAMEER, HDINSIGHT, TRENDMICRO DEEP SECURITY
AND CHEF

AVYAN CONSULTING CORP | AZUREMARKETPLACE@AVYANCONSULTING.COM

HOL Guide for Enterprise Risk Analysis

The purpose of this section is to capture all changes made to the content of the document.

Contact for Enquiries and Proposed Changes

If you have any questions regarding this document, please contact:

Email Address

azuremarketplace@avyanconsulting.com

1 Table of Contents

1. Overview	3
2. How to deploy this solution	3
3. How to configure the components	7
1. Datameer	7
2. TrendMicro	7
4. Signing into Datameer UI	8
5. Configure Datameer to Fetch Data from Azure Storage	11
6. Link, Clean and Prepare the Data	15
7. Perform Analysis to Identify Outliers	26
8. Logging in to the TrendMicro DSM	45
1. Server name	45
2. Server login	45
9. Perform policy configuration on the TrendMicro DSM	46
10. Exercises	48
1. Datameer – Visualize the Data	48
2. TrendMicro – Malware test	48
11. Visualize the Data	48
12. Malware Test	52
1. Generating Malware alert in the computer	52
2. Dashboard – Malware Alert	53
3. Malware Alert verification	53
13. References, Attachments & Definitions (Respective track leads)	54
1. References	54

HOL Guide for Enterprise Risk Analysis

1 Overview

The purpose of this document is to provide the step-by-step instructions of deploying and configuring the Enterprise Risk Analysis using Datameer Business Intelligence and TrendMicro DeepSecurity security solution and lab exercises.

The exercises includes creation of credit fraud risk awareness using sample (representative) data, building powerful Infographics of the Datameer and the security intelligence in the malware detection of the TrendMicro DeepSecurity

2 How to deploy this solution

This section will provide you the details of how to deploy this solution in the Microsoft Azure

- 1) Go to the below link available in the Github

<https://github.com/AvyanConsultingCorp/azure-quickstart-templates/tree/master/datameer-trend-chef-businessintelligence>

- 2) Click on the “Deploy to Azure” in the page, this will take you to the page where you need to provide the parameters



- 3) Provide the custom parameters for the solution accordingly and click “Next”

HOL Guide for Enterprise Risk Analysis

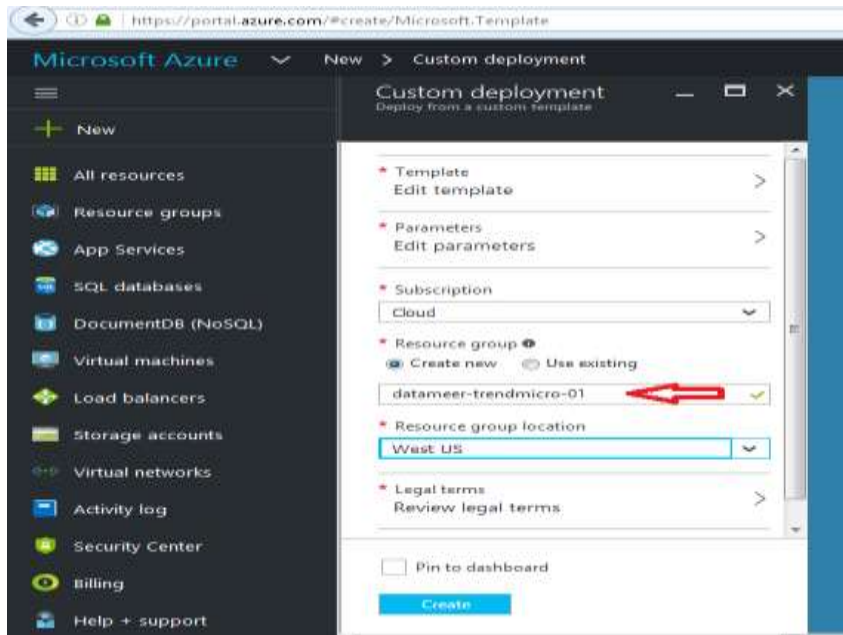
The screenshot shows the Microsoft Azure portal interface for a custom deployment. The left sidebar contains navigation links for 'New', 'All resources', 'Resource groups', 'App Services', 'SQL databases', 'DocumentDB (NoSQL)', 'Virtual machines', 'Load balancers', 'Storage accounts', 'Virtual networks', 'Activity log', 'Security Center', 'Billing', and 'Help + support'. The main content area is titled 'Custom deployment' and 'Parameters'. The 'Parameters' tab is active, showing a list of parameters to be configured. The 'Subscription' dropdown is highlighted with a red arrow, indicating it is the next step in the process. Other parameters include 'Template', 'Resource group', 'Resource group location', 'Legal terms', and various HDInsight cluster parameters like 'LOCATION', 'HDCLUSTERTYPE', 'HDCLUSTERNAME', 'HDCLUSTERLOGNUSERNAME', 'HDCLUSTERLOINPASSWORD', 'HDSSHUSERNAME', 'HDSSHPASSWORD', and 'HDISTORAGEACCOUNT'.

- 4) You need to select the subscription you want to deploy this solution

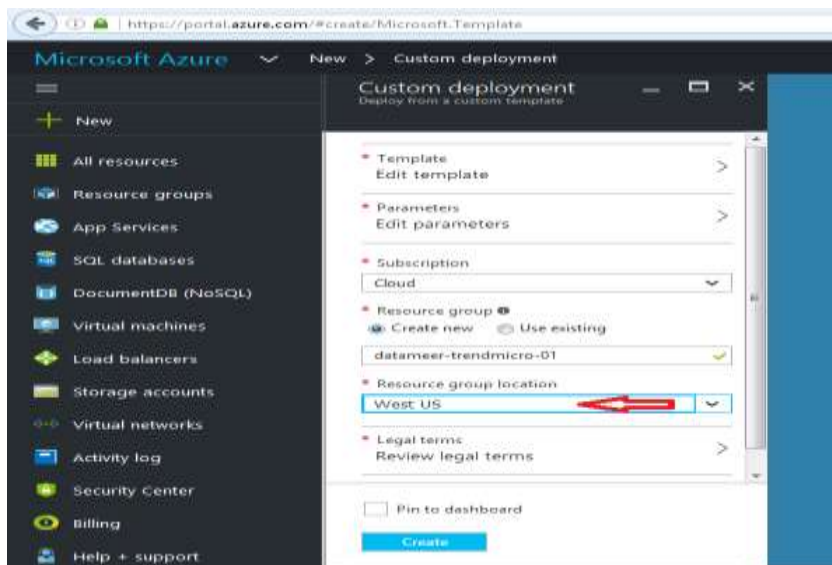
This screenshot shows the same 'Custom deployment' parameters page as the previous one, but with the 'Subscription' dropdown menu open. A red arrow points to the dropdown, indicating the user should select a subscription. The 'Resource group' is set to 'Create new', and the 'Resource group location' is set to 'West US'. The 'Legal terms' are set to 'Review legal terms'. The 'Create' button is visible at the bottom.

HOL Guide for Enterprise Risk Analysis

- 5) Either you can create a new “Resource Group” or use the existing resource group to deploy this solution

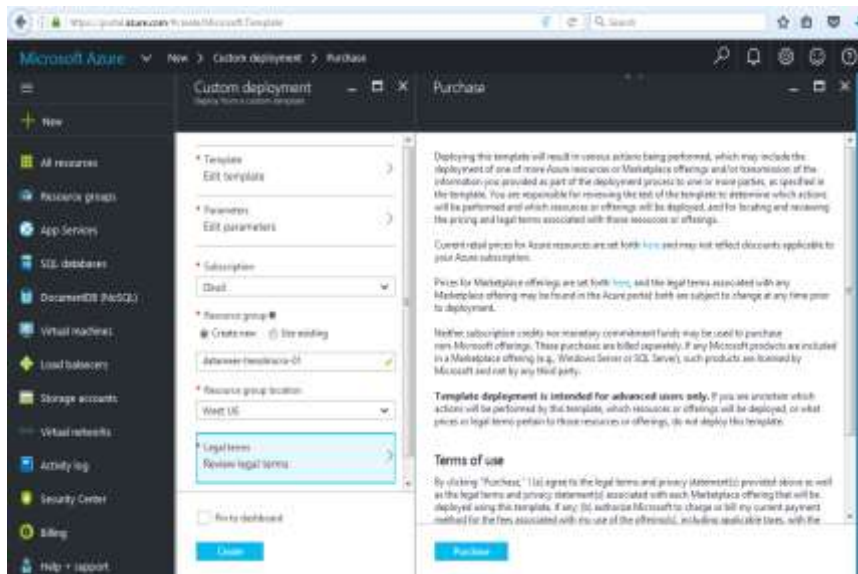


- 6) Select your choice of Region to deploy this solution,

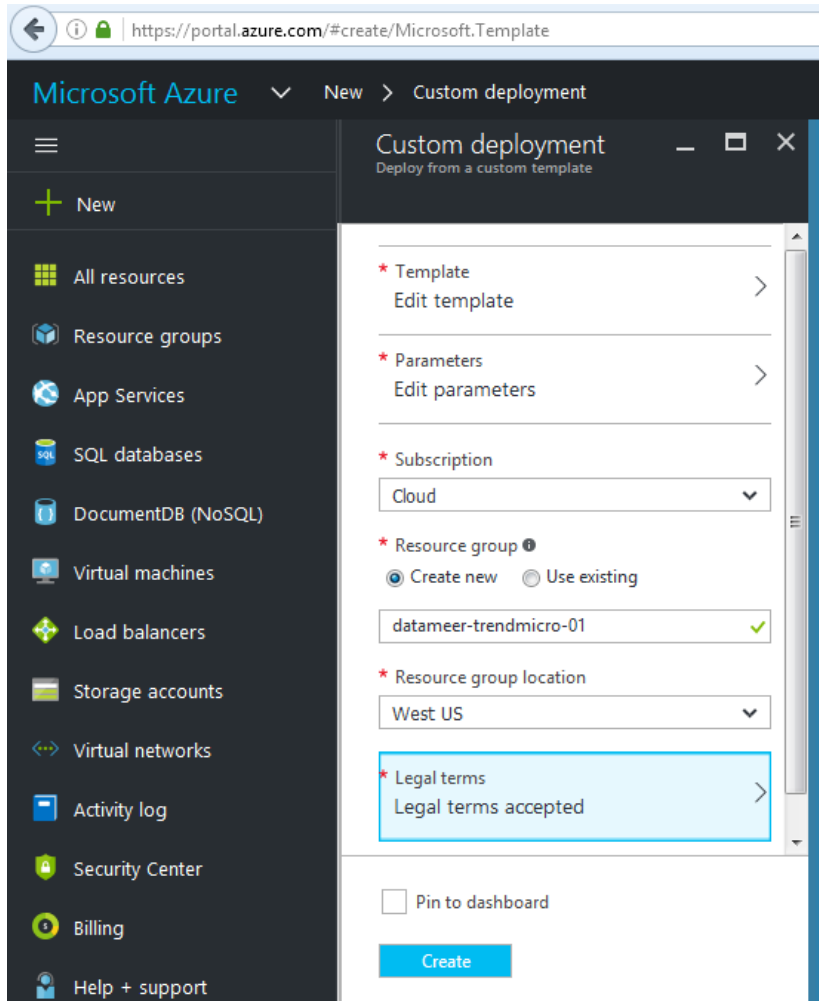


HOL Guide for Enterprise Risk Analysis

- 7) Accept the legal terms to deploy the products from the Azure marketplace which includes, Datameer, TredndMicro and Chef. Click on the “Purchase” button for the same.



- 8) Click on the “Create” button to start deploy the solution now



3 How to configure the components

3.1 Datameer

Datameer is the product used for the Big Data Analysis. It can be used many types of data and can connect to different data sources like storage, database etc. In this solution, the data (.csv) from the azure blog storage will be used too identify the Fraud detection using the credit card. The below sections will provide the details of the configuration of the data in the Datameer for the Big Data Analysis

3.2 TrendMicro

TrendMicro is the industry leading security product, which has the capabilities of

HOL Guide for Enterprise Risk Analysis

- Anti-Virus/Anti-malware detection and prevention.
- Web reputation
- Host based firewall
- Host based Intrusion detection and prevention
- File Integrity monitoring
- Log Inspection

TrendMicro DeepSecurity is an agent based security solution which will help the organisations to comply with all their security requirements.

This solution, showcases the Anti-Malware capabilities of the TrendMicro deepSecurity and below sections will provide the details of the configuration on the same.

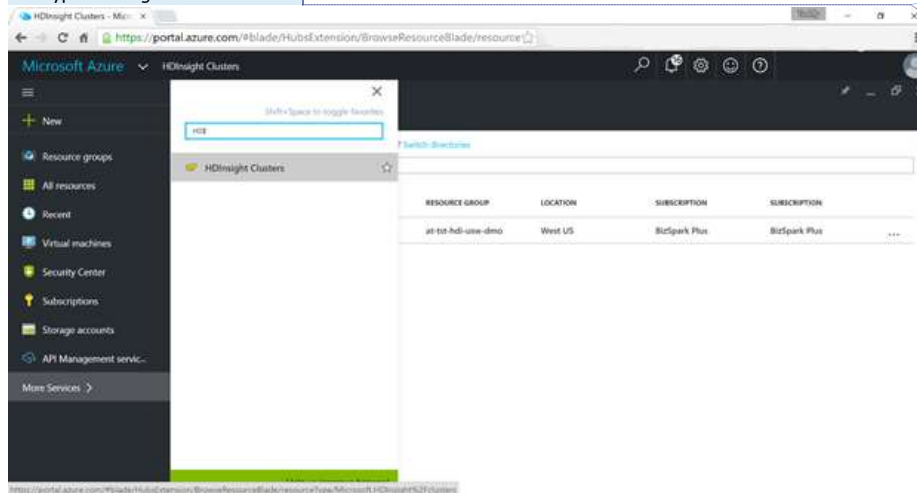
4 Signing into Datameer UI

The fraud analysis is performed with the Business Analytics components of the solution, and namely Datameer and Azure HDInsight. All steps are executed in the Datameer UI. There are two ways that you can use to access the Datameer UI:

- Accessing it directly via the UI URL
- Accessing it via the Azure Management Portal

For the purposes of this HOL we will access the Datameer UI from the Azure Management Portal. Follow these steps:

1. In Azure Management Portal (<http://portal.azure.com>) click on *More Services* in the left-side navigation and type *HDInsight* in the filter box

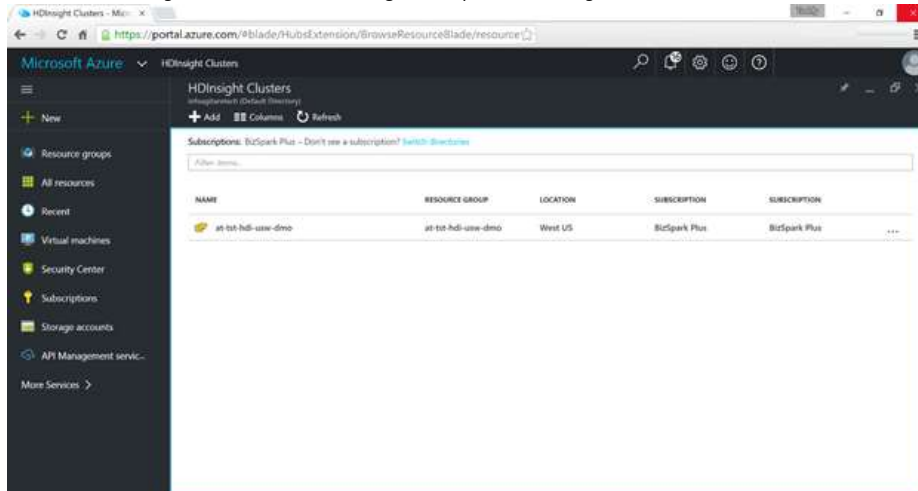


Commented [GP1]: Instead we should ask them to open the resource group and then navigate to the HDInsight cluster.....

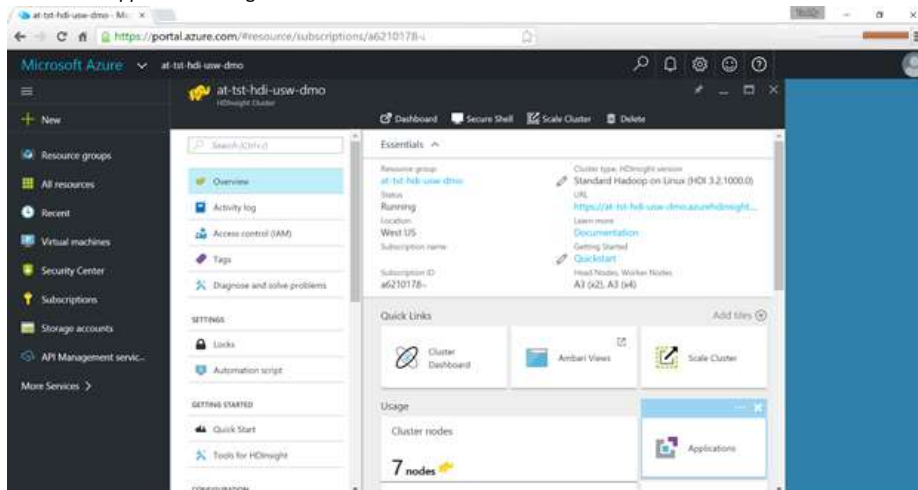
Commented [TM2]: Do we have specific name for the resource group?

HOL Guide for Enterprise Risk Analysis

2. Click on the *HDInsight Clusters* in the resulting list to open the *HDInsight Clusters Blade*



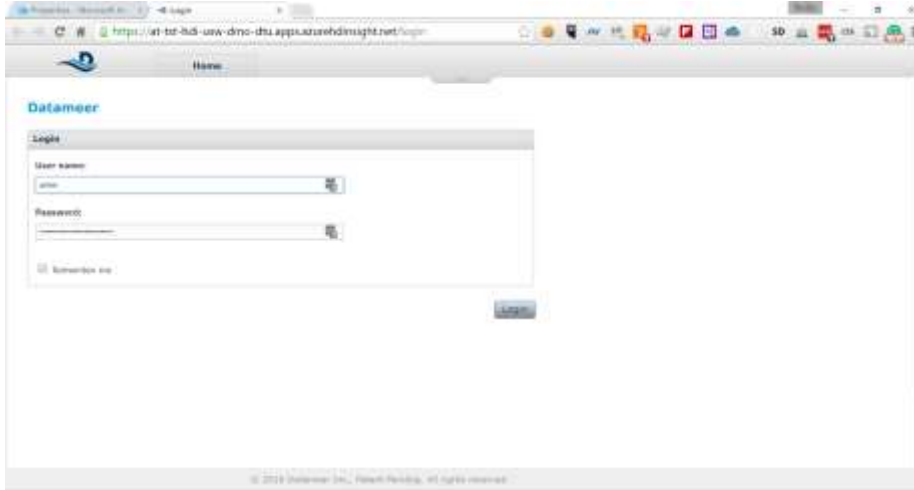
3. Click on the HDInsight Cluster from the Avyan Consulting Enterprise Risk Analysis Solution and on the right click on the *Applications* widget



You will see Datameer as the only application in the list.

HOL Guide for Enterprise Risk Analysis

- Click on the *Portal* link on the right of the Datameer application to load the Datameer UI



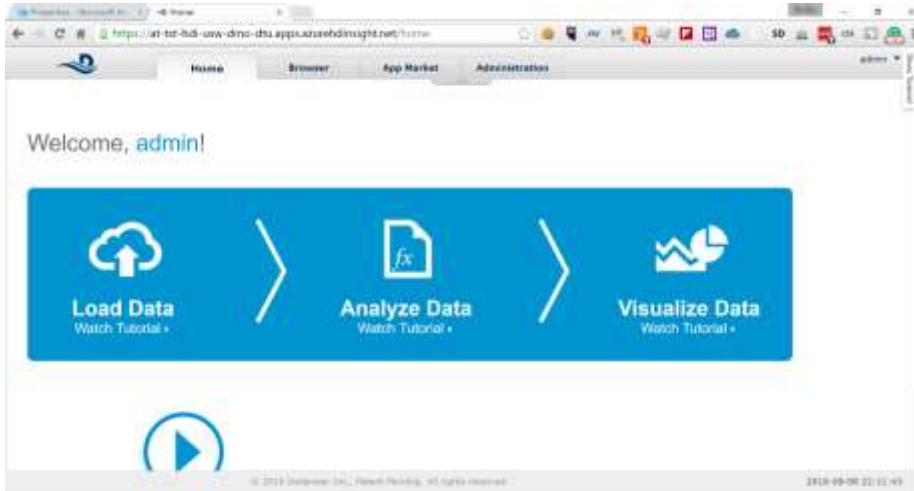
You will be prompted to sign in using your Datameer username and password

- Sign into Datameer UI using the following default credentials:

username: *admin*

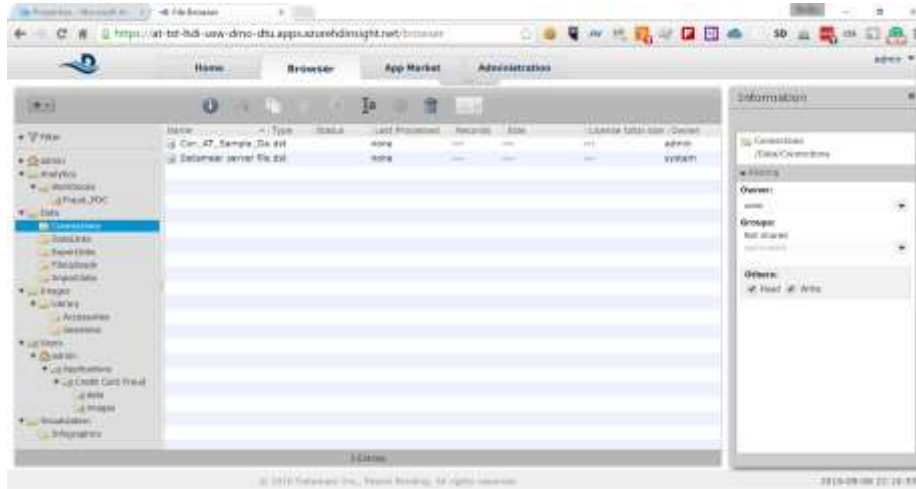
password: *admin*

You will see the Welcome Screen for Datameer and an introduction video will pop up



HOL Guide for Enterprise Risk Analysis

6. Close the introduction video pop-up and click on the Browse tab



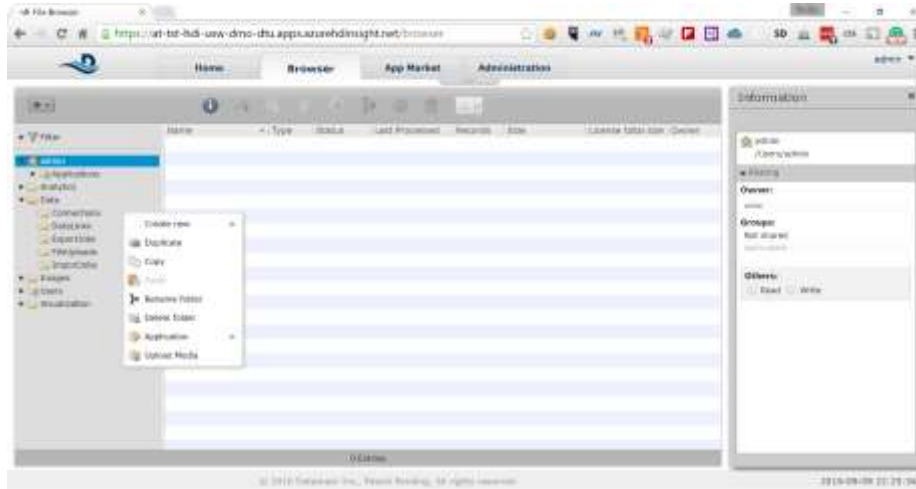
5 Configure Datameer to Fetch Data from Azure Storage

Datameer has more than 65 connectors built in, that allow various systems as data sources. For the purpose of this HOL we will use the Azure Storage connector and fetch the data from there. The assumption is that you have storage account data that contains the transaction data.

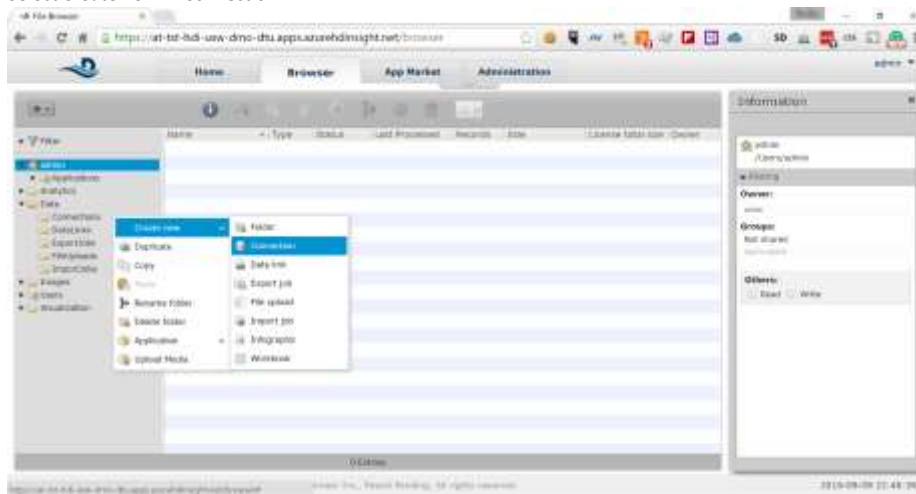
In order to configure Datameer to fetch the data from Azure Storage account you need to go through the following steps:

HOL Guide for Enterprise Risk Analysis

1. Expand the *Data* node in the left-side navigation and right-click on *Connections*

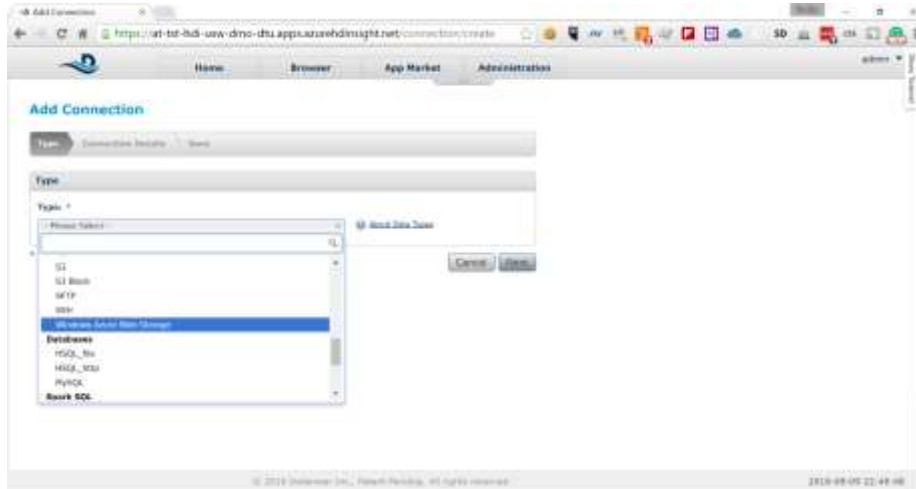


2. Select *Create new* -> *Connection*

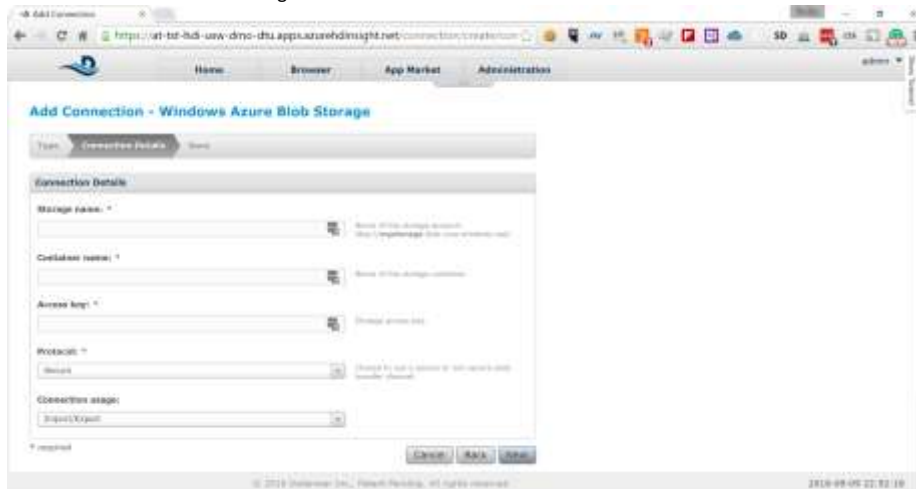


HOL Guide for Enterprise Risk Analysis

3. In the *Type* drop-down, scroll down to *File* section and select *Windows Azure Blob Storage*



4. Click *Next* and fill in the following information on the next screen



Storage Name: The name of the storage account where your data is **TBD for the HOL**

Container Name: The name of the container where your data is **TBD for the HOL**

Access Key: The key used to access the above storage account **TBD for the HOL**

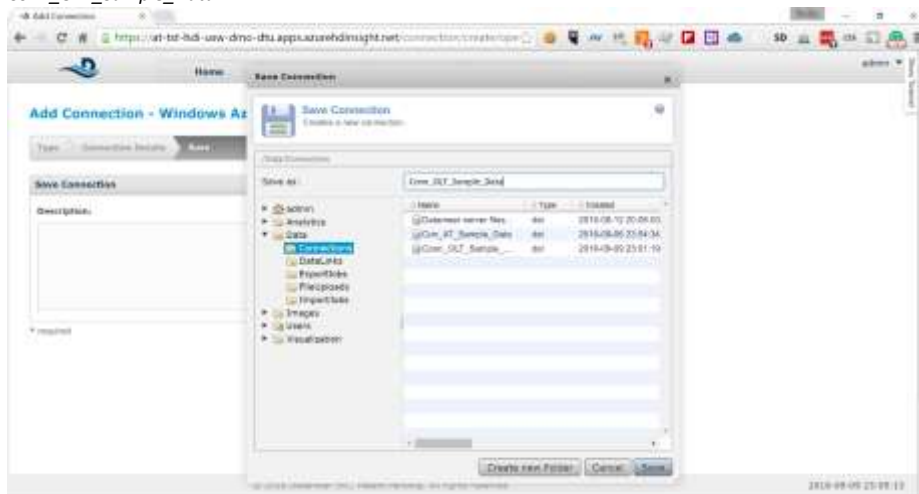
Leave the default values for *Protocol* and *Connection usage*.

5. Click *Next* and on the next screen type the following description for the connection:
"Online Transactions Sample Data"

HOL Guide for Enterprise Risk Analysis

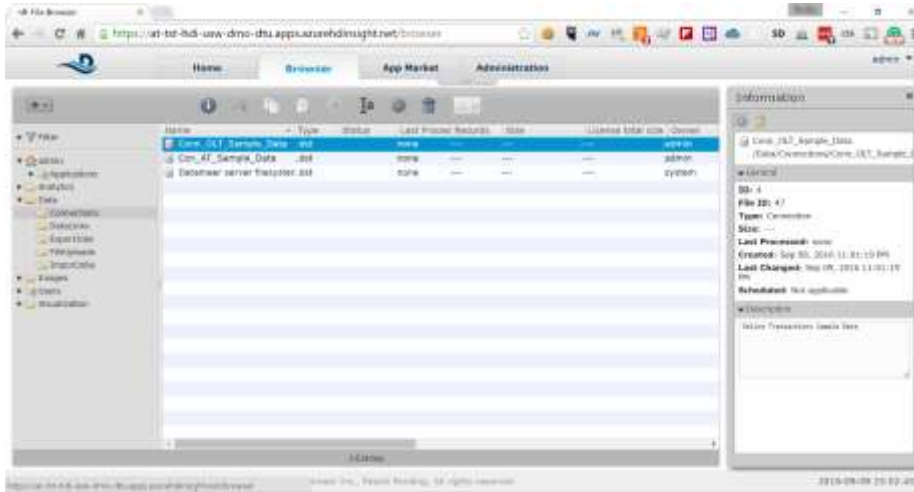


- Click **Save** to save the connection and type the following name in the *Save as* field:
Conn_OLT_Sample_Data



HOL Guide for Enterprise Risk Analysis

- Click **Save** again and you will see the new connection in the list



Now you have Datameer configured to look for data in the specified Azure Storage account and you can start creating your analysis.

6 Link, Clean and Prepare the Data

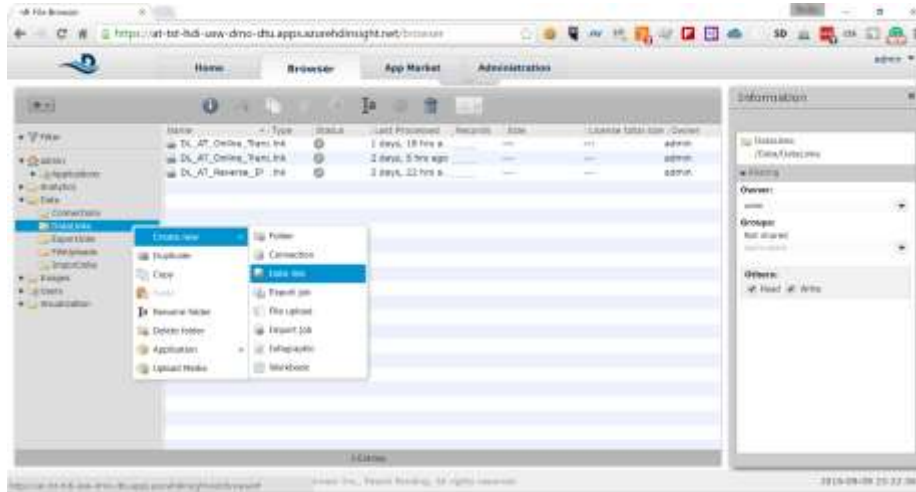
Before we start our analysis we need to tell Datameer which data exactly we want to analyze and make sure that it is in the correct format. The sample data we provided has the following two fields that need to be fixed before it is usable for analysis:

- The *timestamp* field is in ISO-8601 format, which needs to be converted into date/time field that Datameer can understand. We can do this conversion while we are linking the data.
- The *purchase_amount* field is a money field that is interpreted as a *STRING* by Datameer. We need to convert this to *FLOAT* in order to be able to do calculations. We will do that using Datameer formulas once we start our analysis.

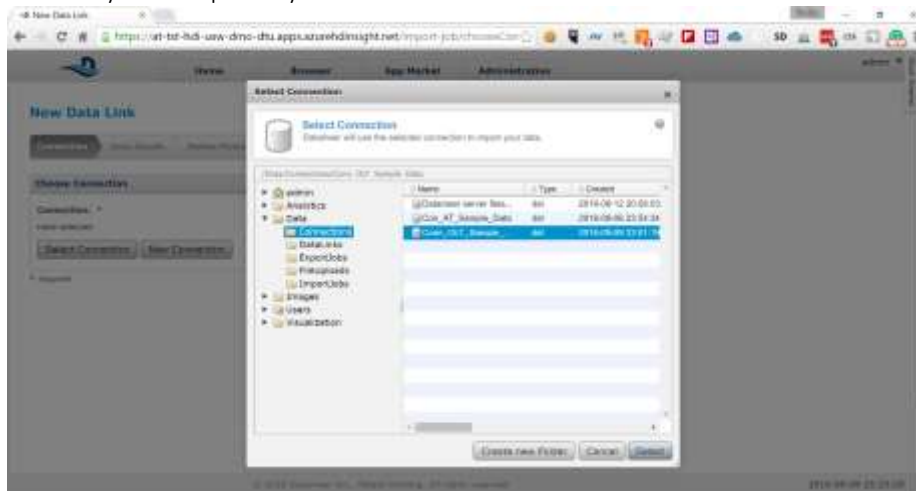
Here are the steps to link the data for analysis.

HOL Guide for Enterprise Risk Analysis

1. Right-click on the *DataLinks* node in the left-side navigation and select *Create new -> Data link*



2. On the next screen click on the *Select Connection* button and select the *Conn_OLT_Sample_Data* connection you created previously



3. Click on the *Select* button in the pop-up. Keep the default value *CSV/TSV* in the *File Type* drop down and click *Next*

HOL Guide for Enterprise Risk Analysis

© 2018 Splunk Inc., All Rights Reserved. 2018-09-05 25:27:03

4. On the next screen type the following in the *File Or Folder* field:
/samples/online-transactions-cc_masked.csv

© 2018 Splunk Inc., All Rights Reserved. 2018-09-05 25:28:30

Scroll down to the bottom, keeping the default values for the rest of the fields, and click on *Next*

HOL Guide for Enterprise Risk Analysis

5. Datameer pre-fetches a representative sample of the data and shows it on the next screen

[illegible]

- Click on the down-arrow for the field type under *timestamp* and change the type from *STRING* to *DATE*

[illegible]

- The dates in the timestamp are automatically marked in red because Datameer cannot parse the ISO-8601 date by default and an input field appears under the field type drop-down

HOL Guide for Enterprise Risk Analysis

9. On the next screen keep the default value for *Trigger* and click on *Next*

New Data Link

Connections | Data Sources | Refresh Sample Data | Subscribes | Save

Refresh Sample Data

Trigger:

☒ Manually
☐ On a schedule

This determines how and when to refresh the sample data used for analysis. Manually refreshes sample data when you click the Refresh Sample Data button. On a schedule refreshes sample data at the specified frequency. Set the frequency and time of day.

Advanced

* required

Cancel Back Next

© 2018 Splunk Inc., All rights reserved. 2018-09-05 25:48:31

10. On the next screen type a meaningful description for the DataLink and click on *Save*

New Data Link

Connections | Data Sources | Refresh Sample Data | Subscribes | Save

Save Data Link

Description:

Data link for the active transactions sample link

Generate name:

☒ Generate sample name(s) after save

Notifications

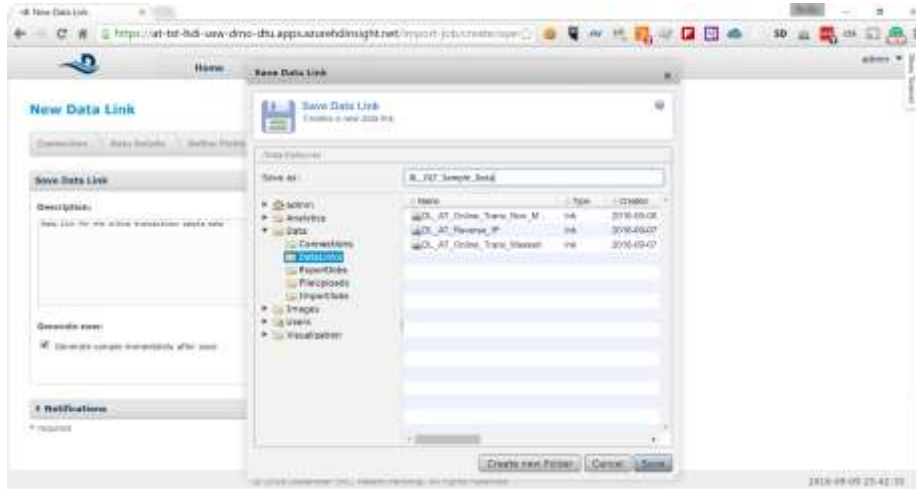
* required

Cancel Back Save

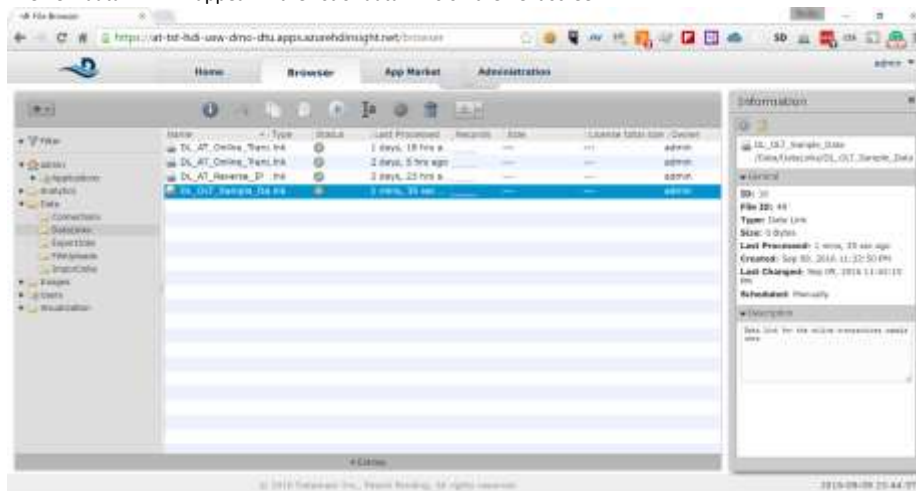
© 2018 Splunk Inc., All rights reserved. 2018-09-05 25:41:49

HOL Guide for Enterprise Risk Analysis

11. Type the following name in the *Save as* field for the DataLink and click on the *Save* button



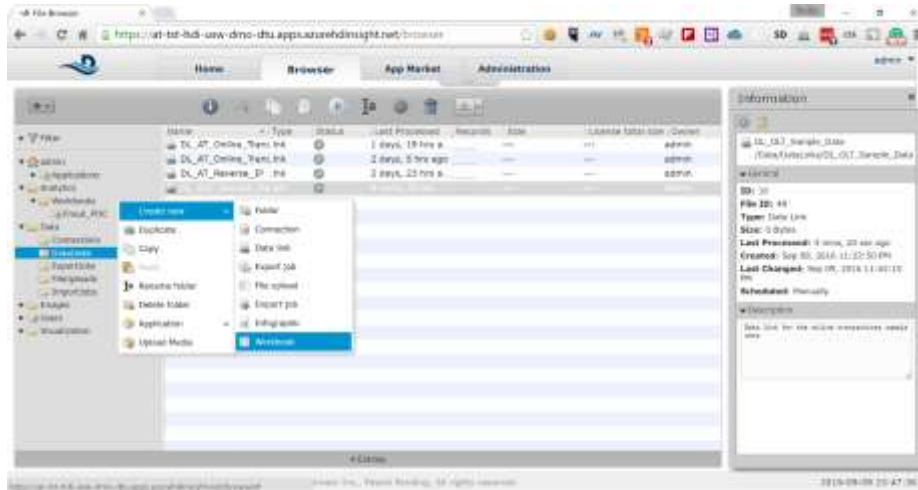
12. The new data link will appear in the list of data links on the next screen



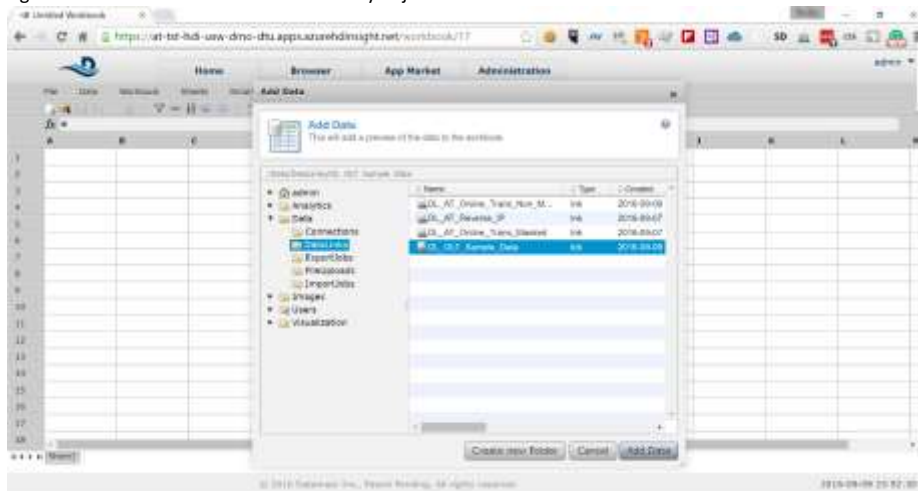
Now you have your dataset linked and have done some preliminary clean-up of the data. Next we will create a workbook where we will finish cleaning up our data and do our analysis. Here are the steps.

13. Expand the *Analytics* node in the left-side navigation, right-click on the *Workbook* node and select *Create new -> Workbook*

HOL Guide for Enterprise Risk Analysis

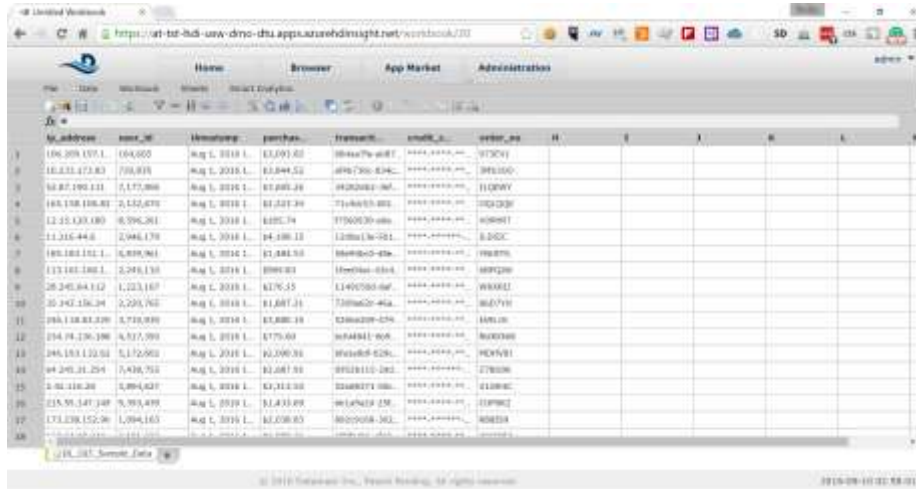


14. A new workbook is created and a pop-up window is shown asking you to select the dataset you want to use for analysis. Expand the **Data** node in the pop-up navigation and click on the **DataLinks** node. In the right-side window select the data link that you just created.



HOL Guide for Enterprise Risk Analysis

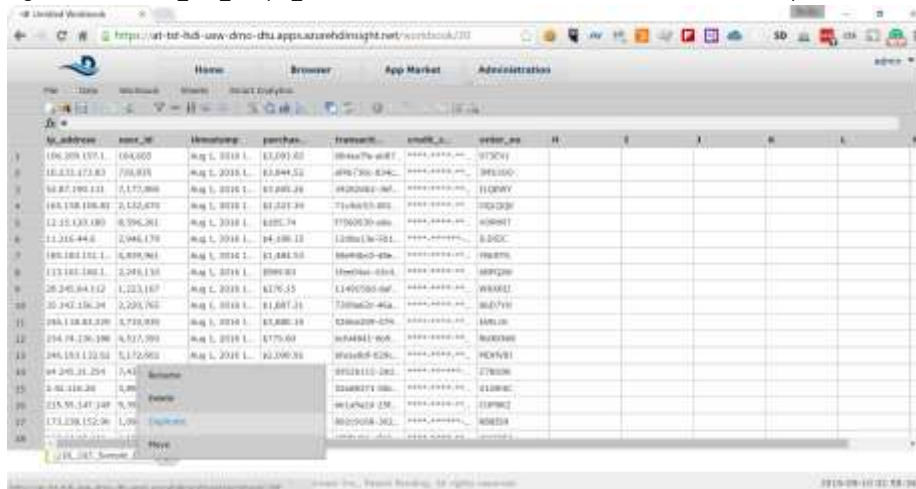
- Click on **Add Data** to load a sample of the dataset in the workbook sheet



Id	Address	Name	HomePhone	Purchase	Transaction	Credit	Order
1	106.209.107.1	064,660	Aug 1, 2018 L	13,093.60	0844764887	****,****,****	9730V1
2	10.233.173.83	708,818	Aug 1, 2018 L	13,844.52	0847306-834C	****,****,****	385100
3	10.87.190.131	5,173,868	Aug 1, 2018 L	17,885.28	0420380-067	****,****,****	1109WY
4	168.158.108.83	2,152,878	Aug 1, 2018 L	11,527.39	7744455-882	****,****,****	282330
5	12.15.133.180	8,796,281	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
6	11.215.44.8	2,946,178	Aug 1, 2018 L	94,186.12	1288136-021	****,****,****	1,102C
7	189.183.182.1	4,839,961	Aug 1, 2018 L	17,885.28	0844764887	****,****,****	9730V1
8	173.183.182.1	2,246,138	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
9	26.240.94.112	1,223,167	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
10	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
11	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
12	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
13	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
14	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
15	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
16	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
17	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
18	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
19	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
20	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801

The UI you are presented with is very similar to Excel and uses the same concepts.

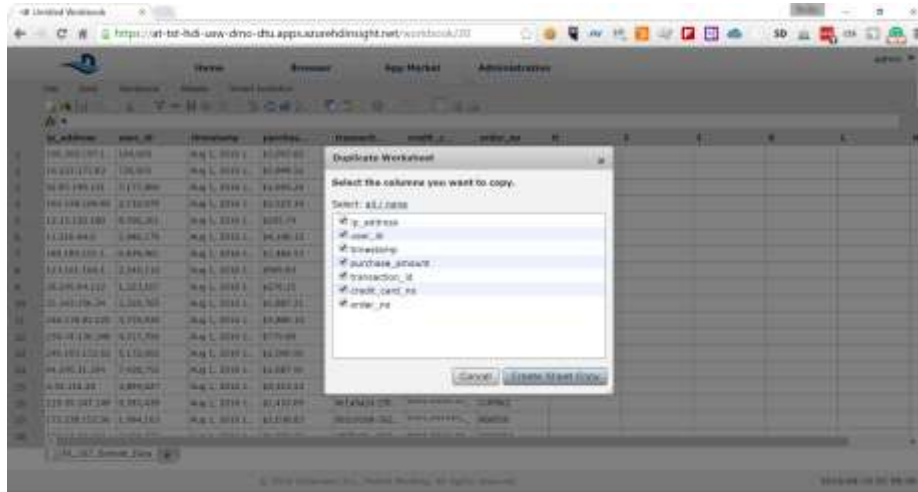
- Right-click on the **DL_OLT_Sample_Data** sheet at the bottom of the screen next and select **Duplicate**



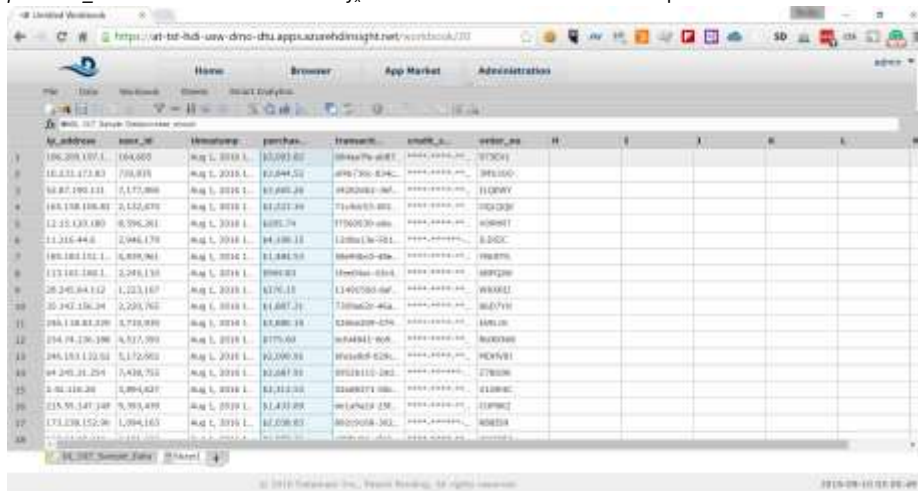
Id	Address	Name	HomePhone	Purchase	Transaction	Credit	Order
1	106.209.107.1	064,660	Aug 1, 2018 L	13,093.60	0844764887	****,****,****	9730V1
2	10.233.173.83	708,818	Aug 1, 2018 L	13,844.52	0847306-834C	****,****,****	385100
3	10.87.190.131	5,173,868	Aug 1, 2018 L	17,885.28	0420380-067	****,****,****	1109WY
4	168.158.108.83	2,152,878	Aug 1, 2018 L	11,527.39	7744455-882	****,****,****	282330
5	12.15.133.180	8,796,281	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
6	11.215.44.8	2,946,178	Aug 1, 2018 L	94,186.12	1288136-021	****,****,****	1,102C
7	189.183.182.1	4,839,961	Aug 1, 2018 L	17,885.28	0844764887	****,****,****	9730V1
8	173.183.182.1	2,246,138	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
9	26.240.94.112	1,223,167	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
10	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
11	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
12	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
13	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
14	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
15	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
16	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
17	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
18	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
19	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801
20	26.240.94.112	2,229,762	Aug 1, 2018 L	1,895.74	7790030-489	****,****,****	408801

HOL Guide for Enterprise Risk Analysis

17. Keep all of the fields selected in the pop-up and click on the *Create Sheet Copy* button



18. A new copy of the sheet is created that contains all of the data from the original sheet. Click on the *purchase_amount* column to enable the f_x field for that column available on top of the sheet



19. Type the following in the f_x field and press *Enter*

FLOAT(SUBSTITUTEALL(SUBSTR(#DL OLT Sample Data!purchase amount;1);",";""))

The formula strips the \$ (dollar) sign in front of the amount, removes all commas and converts the string to FLOAT. Now you can use numeric functions to perform calculations on the field.

HOL Guide for Enterprise Risk Analysis

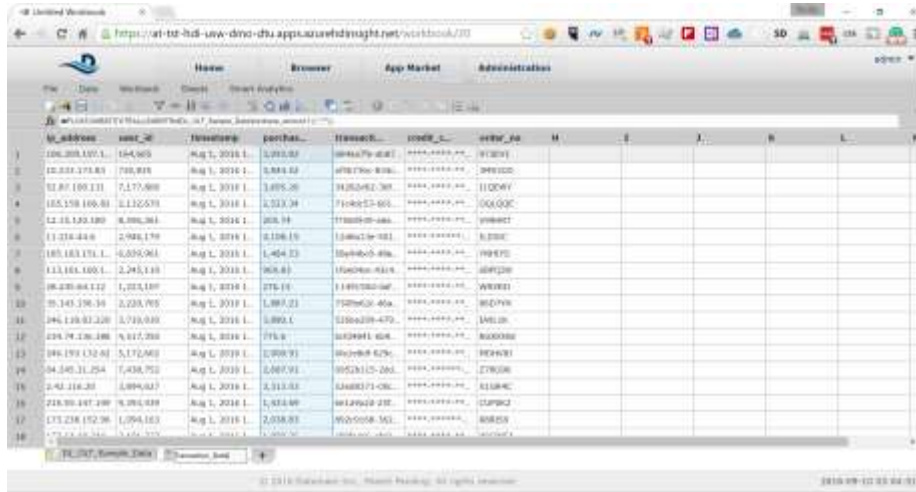
Id	Address	Name	Date	Amount	Status	Other Fields
1	106.209.107.1	106.209.107.1	Aug 1, 2018	1,000.00	Completed	...
2	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
3	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
4	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
5	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
6	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
7	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
8	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
9	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
10	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
11	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
12	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
13	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
14	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
15	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
16	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
17	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
18	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
19	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
20	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...

- Right-click on the sheet name at the bottom of the screen to show the context menu for *Sheet1* and select *Rename*

Id	Address	Name	Date	Amount	Status	Other Fields
1	106.209.107.1	106.209.107.1	Aug 1, 2018	1,000.00	Completed	...
2	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
3	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
4	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
5	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
6	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
7	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
8	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
9	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
10	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
11	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
12	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
13	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
14	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
15	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
16	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
17	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
18	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
19	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...
20	10.233.173.83	10.233.173.83	Aug 1, 2018	1,000.00	Completed	...

HOL Guide for Enterprise Risk Analysis

21. Rename *Sheet1* to *Transaction_Data*



Account	Date	Transaction	Merchant	Transaction ID	Other Info
106,000,107.1	Aug 1, 2018	1,213.50	WALMART STORE	WALMART STORE	WALMART STORE
10,000,000.00	Aug 1, 2018	1,213.50	WALMART STORE	WALMART STORE	WALMART STORE
10,000,000.00	Aug 1, 2018	1,213.50	WALMART STORE	WALMART STORE	WALMART STORE
10,000,000.00	Aug 1, 2018	1,213.50	WALMART STORE	WALMART STORE	WALMART STORE
10,000,000.00	Aug 1, 2018	1,213.50	WALMART STORE	WALMART STORE	WALMART STORE
10,000,000.00	Aug 1, 2018	1,213.50	WALMART STORE	WALMART STORE	WALMART STORE
10,000,000.00	Aug 1, 2018	1,213.50	WALMART STORE	WALMART STORE	WALMART STORE
10,000,000.00	Aug 1, 2018	1,213.50	WALMART STORE	WALMART STORE	WALMART STORE
10,000,000.00	Aug 1, 2018	1,213.50	WALMART STORE	WALMART STORE	WALMART STORE
10,000,000.00	Aug 1, 2018	1,213.50	WALMART STORE	WALMART STORE	WALMART STORE

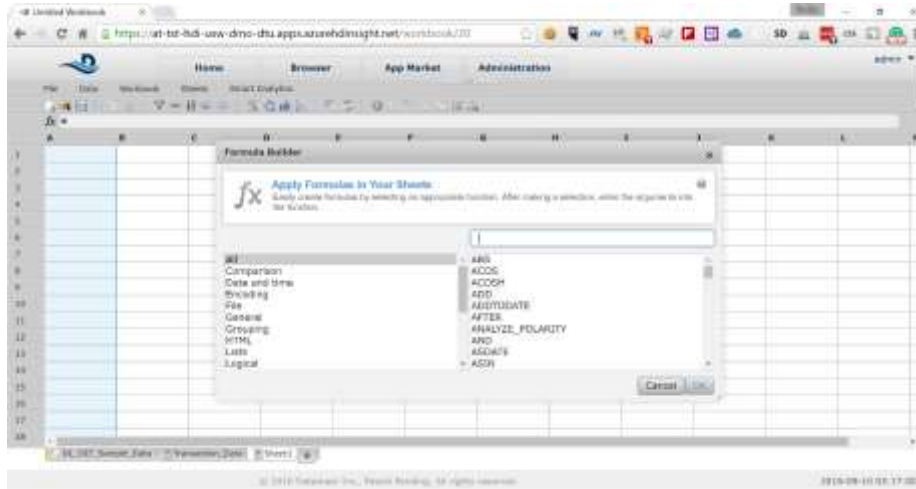
With this we are done with the clean-up of our data and are ready to perform our analysis.

7 Perform Analysis to Identify Outliers

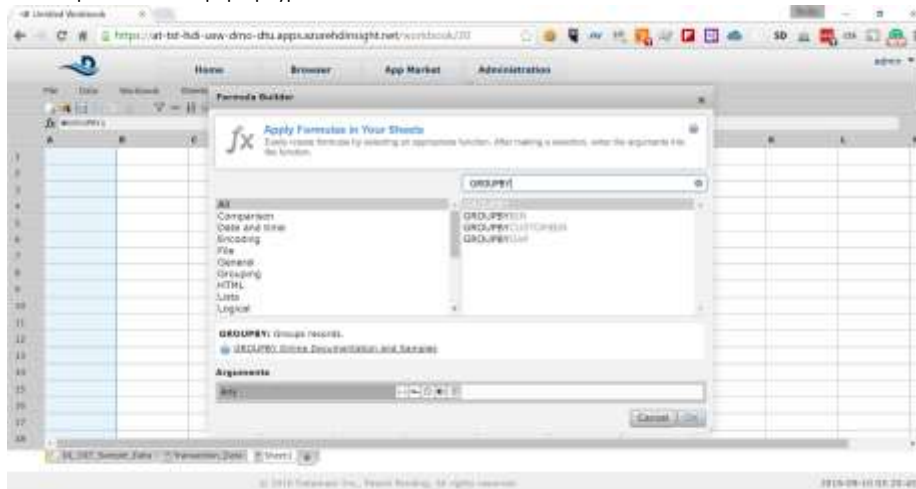
The goal of our analysis is to identify unusual purchasing patterns that deviate from a well-established norm. If we notice something unusual this may be sign that fraud may be committed. For the purpose of this HOL we will be looking for period during the month, in which the transactions significantly deviate from the normal patters during the rest of the month. Here the steps:

HOL Guide for Enterprise Risk Analysis

1. Click on the + sign next to the *Transaction_Data* sheet to create an empty sheet for analysis

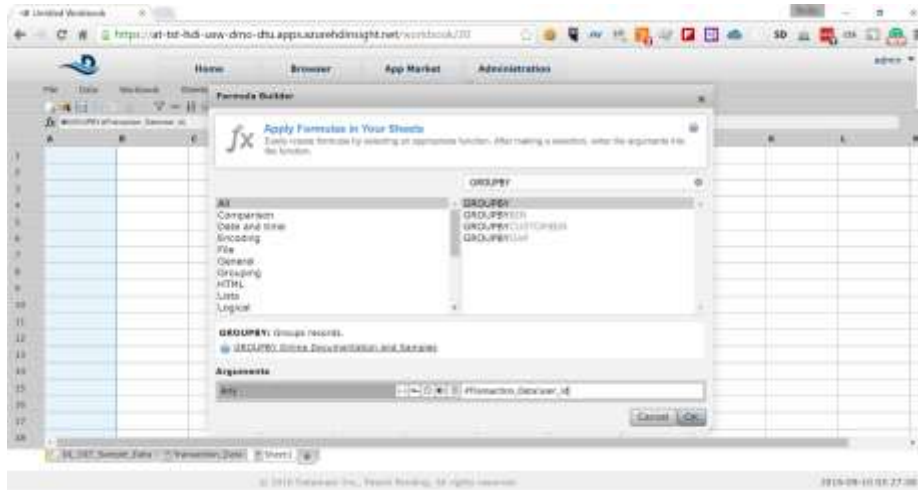


2. In the input field in the pop-up type *GROUPBY* to filter the functions and select *GROUPBY* from the list

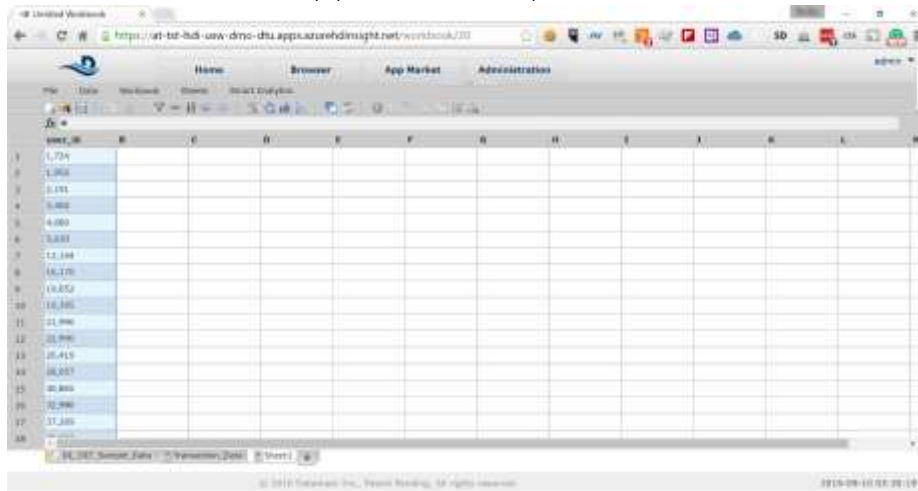


3. In the *Arguments* input field at the bottom of the pop-up type *#Transaction_Data!user_id* to group the data by user identifier and click on the *OK* button

HOL Guide for Enterprise Risk Analysis

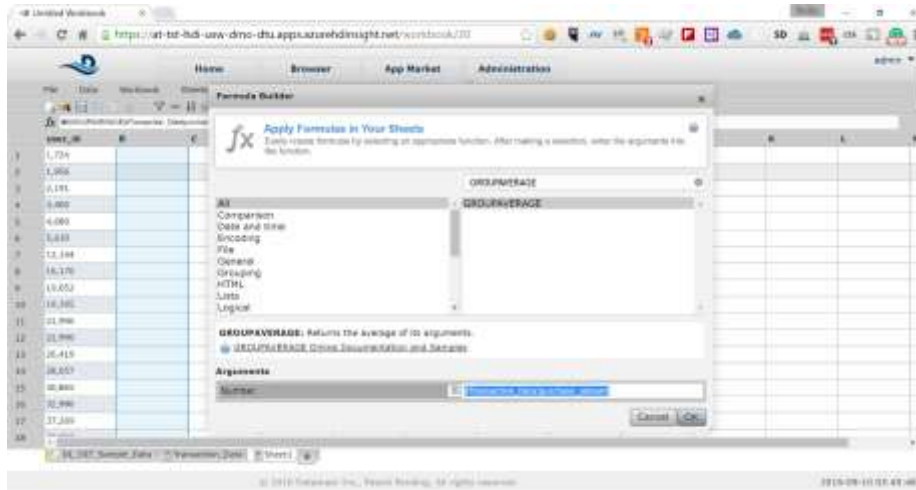


4. The first column of the sheet will be populated with list of unique user identifiers



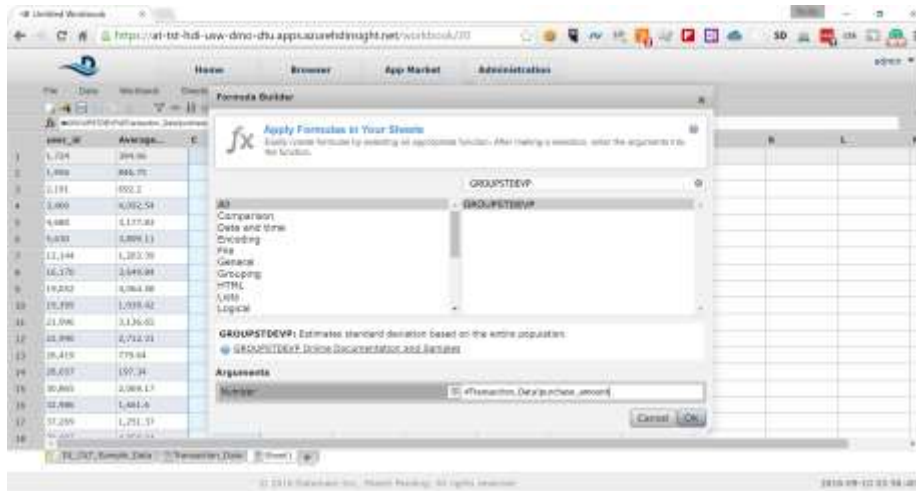
5. Click on the second column to show the functions pop-up again and type **GROUPAVERAGE** in the filter box and select the **GROUPAVERAGE** function. In the *Arguments* field type *#Transaction_Data!purchase_amount*

HOL Guide for Enterprise Risk Analysis



This will calculate the average purchase amount for each of the users.

- Click on the third column to show the function pop-up again and type **GROUPSTDEV** and select the **GROUPSTDEV** function

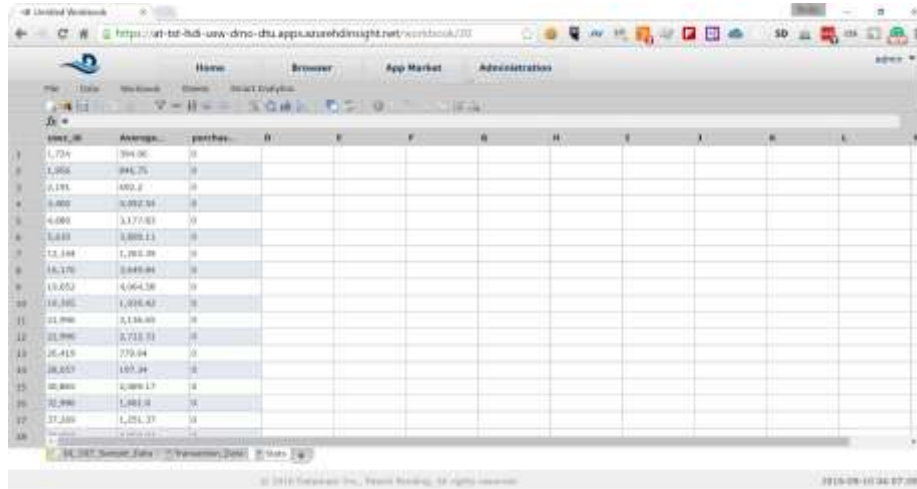


In the Arguments field type `#Transaction_Data!purchase_amount` to calculate the standard deviation for the *purchase_amount* field

- Right-click on the sheet name and select **Rename** from the context menu to rename the sheet. Choose the following name for the sheet:

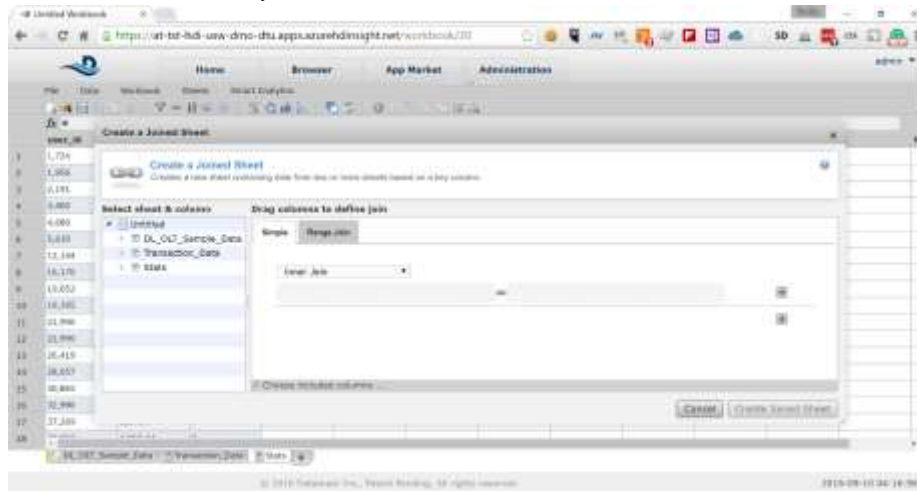
HOL Guide for Enterprise Risk Analysis

Stats



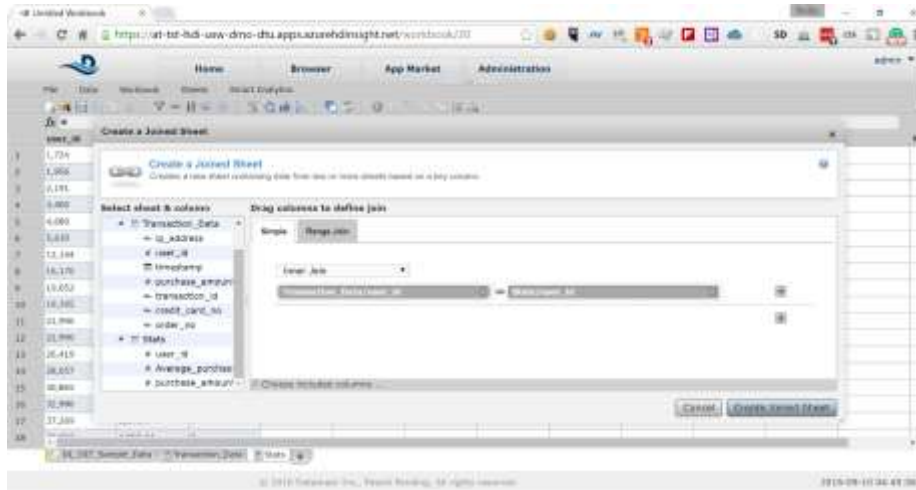
user_id	Average	pathes													
1,734	394.36	0													
1,954	945.75	0													
2,191	452.2	0													
3,002	3,032.54	0													
4,000	1,177.83	0													
5,433	1,888.11	0													
11,144	1,765.08	0													
16,176	2,445.84	0													
19,052	4,964.18	0													
18,345	1,405.42	0													
21,996	2,156.50	0													
21,996	2,718.71	0													
26,415	779.04	0													
28,057	185.34	0													
30,840	4,389.17	0													
32,990	1,861.0	0													
37,340	1,251.37	0													

- Next we need to join the transaction data for each user with the statistical data for each user to determine how much particular transaction differentiates from the common norm. From the menu bar select **Data -> Join** to create a joined sheet



- Expand the **Transaction_Data** node in the pop-up navigation tree and drag the **user_id** field to the right. Do the same with the **user_id** field from the **Stats** node.

HOL Guide for Enterprise Risk Analysis



Click on the *Create Joined Sheet* button to create the joined sheet.

- The resulting sheet will show the joined data from both *Transaction_Data* and *Stats* sheets. For convenience let's rename few of the columns. Right-click and rename the columns as below:

Transaction_Data	Stats	Transaction_Data	Transaction_Data	Transaction_Data	Transaction_Data	Transaction_Data	Transaction_Data	Stats Av	Stats po	Stats	Stats	Stats
1	1,734	71,161,35.50	Aug 1, 2010 0	304.30	54851ac-c96	0000-0000-00	0000-0000-00	304.30	0	0	0	0
2	1,858	14,137,215,149	Aug 1, 2010 1	946.75	21121276-946	0000-0000-00	0000-0000-00	946.75	0	0	0	0
3	1,171	115,113,47.33	Aug 1, 2010 0	993.7	2004606-9-0	0000-0000-00	0000-0000-00	993.7	0	0	0	0
4	5,002	108,81,83,115	Aug 1, 2010 1	4,000.94	9431a10-901	0000-0000-00	0000-0000-00	4,000.94	0	0	0	0
5	4,000	95,173,108,135	Aug 1, 2010 0	3,177.81	908515e-980	0000-0000-00	0000-0000-00	3,177.81	0	0	0	0
6	1,433	5,339	65,25,36,63	Aug 1, 2010 1	1,809.11	943d75-62c	0000-0000-00	1,809.11	0	0	0	0
7	11,144	12,144	189,182,51,226	Aug 1, 2010 0	1,303.89	84d4d7b-73d	0000-0000-00	1,303.89	0	0	0	0
8	16,176	16,176	95,1,161,236	Aug 1, 2010 1	1,644.84	114e47b-3e8	0000-0000-00	1,644.84	0	0	0	0
9	19,652	19,652	131,86,248,238	Aug 1, 2010 0	4,964.58	ed86d79-8a0	0000-0000-00	4,964.58	0	0	0	0
10	18,302	18,302	47,62,231,079	Aug 1, 2010 1	1,439.47	5d90905-80c	0000-0000-00	1,439.47	0	0	0	0
11	21,990	21,990	117,213,133,21	Aug 1, 2010 0	5,133.89	8187243b-188	0000-0000-00	5,133.89	0	0	0	0
12	21,990	22,990	144,112,115,2	Aug 1, 2010 0	5,712.31	9d95d99-3ac	0000-0000-00	5,712.31	0	0	0	0
13	26,415	26,415	44,71,92,219	Aug 1, 2010 0	779.64	6123066d-471	0000-0000-00	779.64	0	0	0	0
14	26,657	26,657	215,76,86,115	Aug 1, 2010 0	197.39	8c25f8a-2498	0000-0000-00	197.39	0	0	0	0
15	32,860	32,860	228,87,238,18	Aug 1, 2010 0	2,868.17	88d4118f-88d	0000-0000-00	2,868.17	0	0	0	0
16	32,860	32,860	95,99,91,117	Aug 1, 2010 1	1,891.9	911724e-9f8	0000-0000-00	1,891.9	0	0	0	0
17	37,380	37,380	117,173,165,2	Aug 1, 2010 0	1,254.37	ca8f34d-50ac	0000-0000-00	1,254.37	0	0	0	0
18	37,380	37,380	10,91,91,91	Aug 1, 2010 0	1,254.37	ca8f34d-50ac	0000-0000-00	1,254.37	0	0	0	0

For convenience let's rename few of the columns. Right-click and rename the columns as below:

Transaction_Data.user_id -> *user_id*

Transaction_Data.purchase_amount -> *purchase_amount*

Stats.Average_purchase_amount -> *average_purchase_amount*

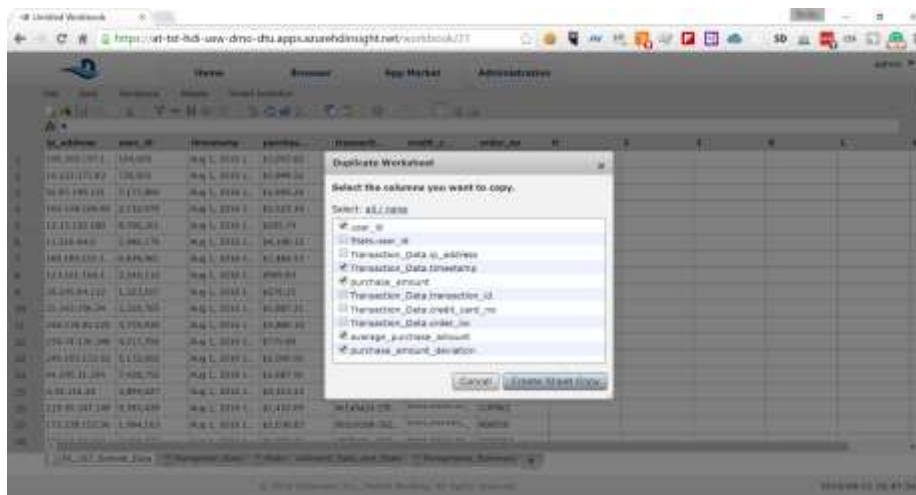
Stats.purchase_amount_Stdev -> *purchase_amount_deviation*

HOL Guide for Enterprise Risk Analysis

Also, right-click on the sheet name and rename it to *Joined_Data_and_Stats*

- Next, we will identify the outliers by creating a copy of the joined data and filtering it. Right-click on the *Joined_Data_and_Stats* sheet and select *Duplicate*. We will select only the data we need and ignore the rest. In the pop-up select only the following fields:

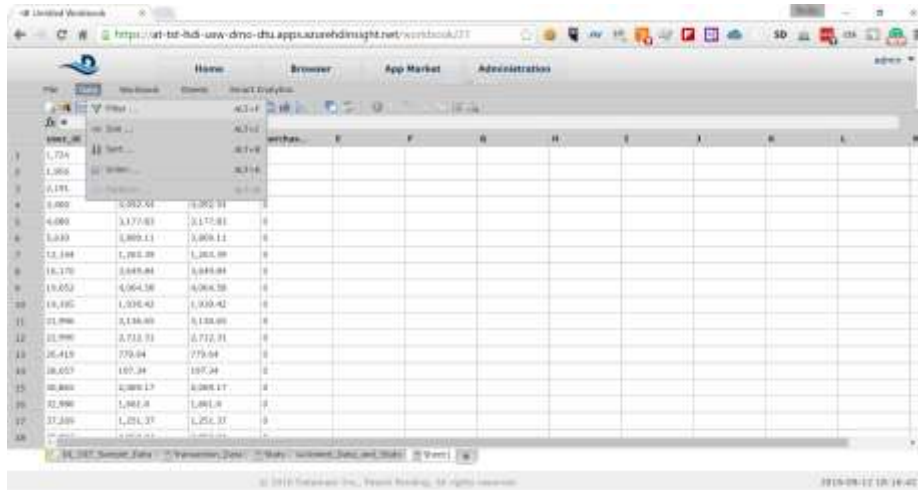
user_id
purchase_amount
average_purchase_amount
purchase_amount_deviation



Click on *Create Sheet Copy* button

- Right-click on the *Transaction_Data.timestamp* field and rename it to *timestamp* only.
- For the purpose of our analysis we will consider transactions with deviation two times more than standard deviation as outliers. In the new sheet select *Data -> Filter* from the menu.

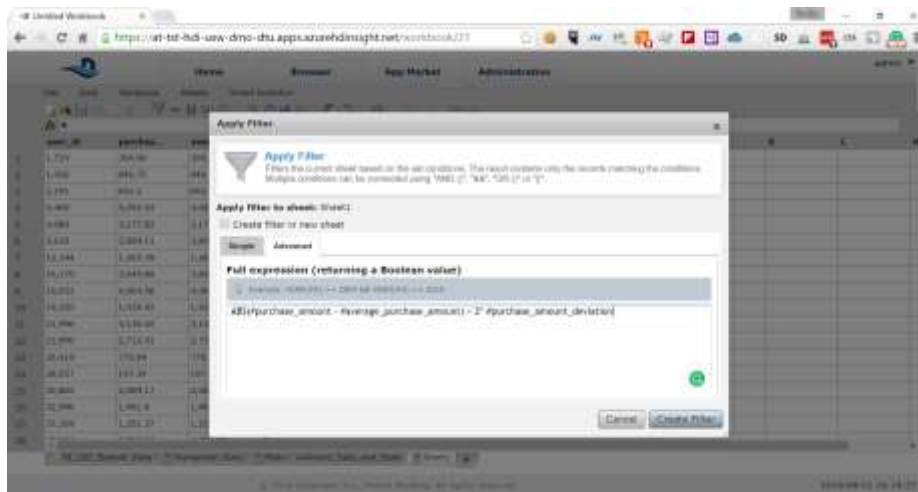
HOL Guide for Enterprise Risk Analysis



The screenshot shows a Datasheet view in the Datameer application. The table has columns labeled 'purchase_amount' and 'average_purchase_amount'. The data is organized into rows, with the first row containing values 1,729 and 1,729. The table is currently empty, indicating that the filter has been applied successfully.

14. Select the *Advanced* tab in the pop-up and type the following formula:

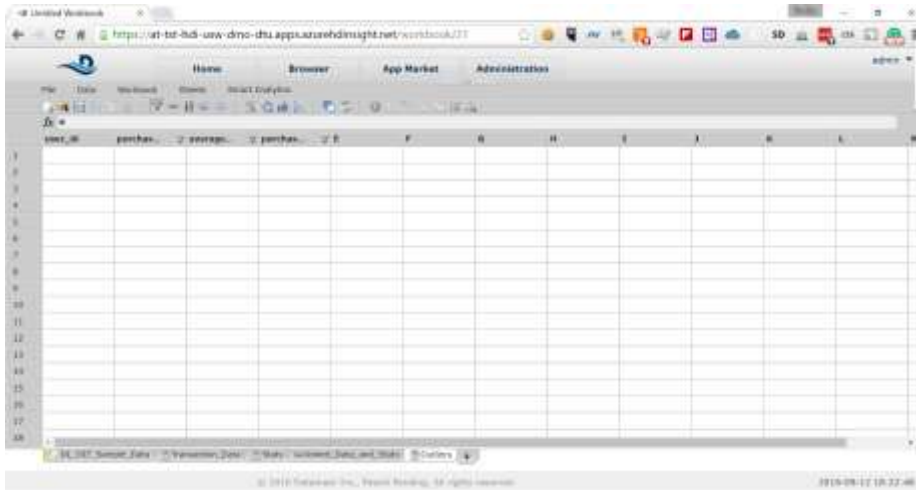
*$ABS(\#purchase_amount - \#average_purchase_amount) > 2 * \#purchase_amount_deviation$*



The screenshot shows the 'Apply Filter' dialog box in the Datameer application. The 'Advanced' tab is selected, and the 'Full expression (returning a Boolean value)' field contains the formula: $ABS(\#purchase_amount - \#average_purchase_amount) > 2 * \#purchase_amount_deviation$. The 'Create filter in new sheet' checkbox is checked, and the 'Create Filter' button is highlighted.

15. The resulting sheet may be empty because the representative sample that Datameer has selected may not have transactions that are considered outliers. Right-click on the sheet name and rename it to *Outliers*

HOL Guide for Enterprise Risk Analysis



16. Finally, we would like to create a summary of the data that we would like to visualize. Let's start with summary of the *Transaction_Data*. Create new sheet and in the formula pop-up select the *GROUPBY* function



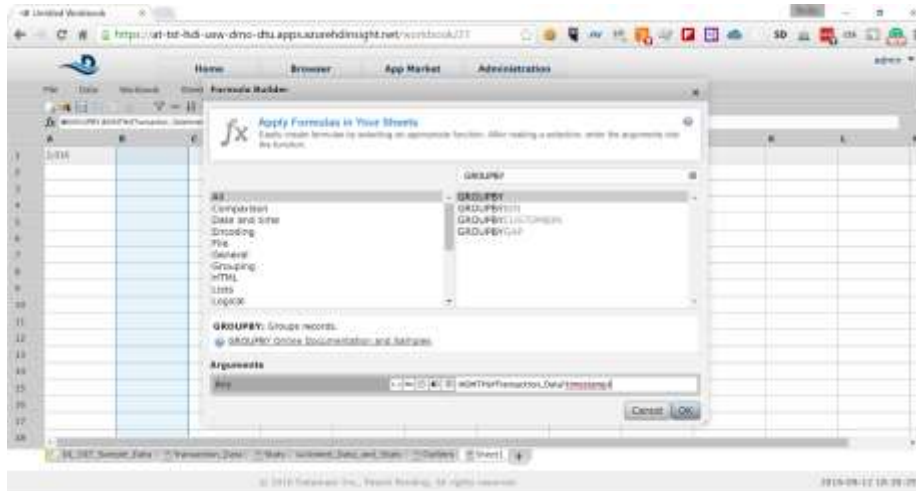
In the *Arguments* field type the following formula:

`YEAR(#Transaction_Data!timestamp)`

17. Click on the next column and in the formula pop-up select again the *GROUPBY* function and paste the following formula in the *Arguments* field:

HOL Guide for Enterprise Risk Analysis

MONTH(#Transaction_Data!timestamp)



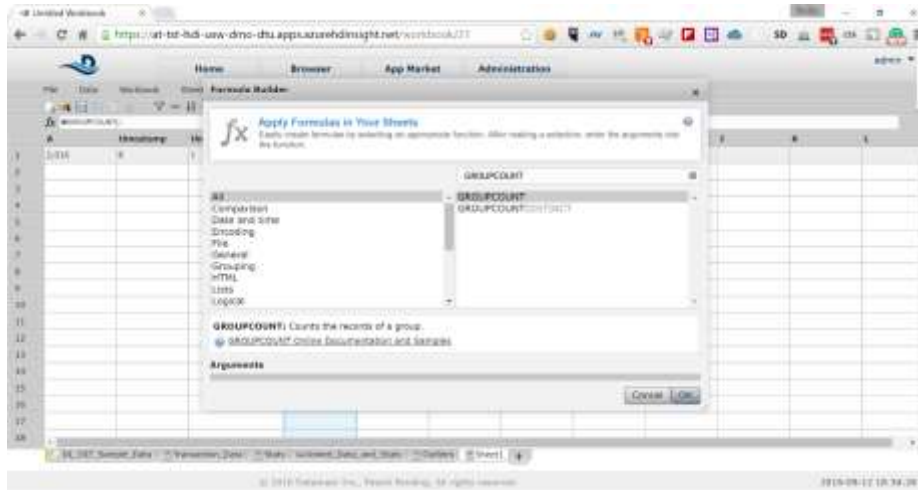
- Click on the third column and in the formula pop-up select again the *GROUPBY* function and paste the following formula in the *Arguments* field:

DAY(#Transaction_Data!timestamp)



- Click on the fourth column and in the formula pop-up select the *GROUPCOUNT* function and click the *OK* button

HOL Guide for Enterprise Risk Analysis



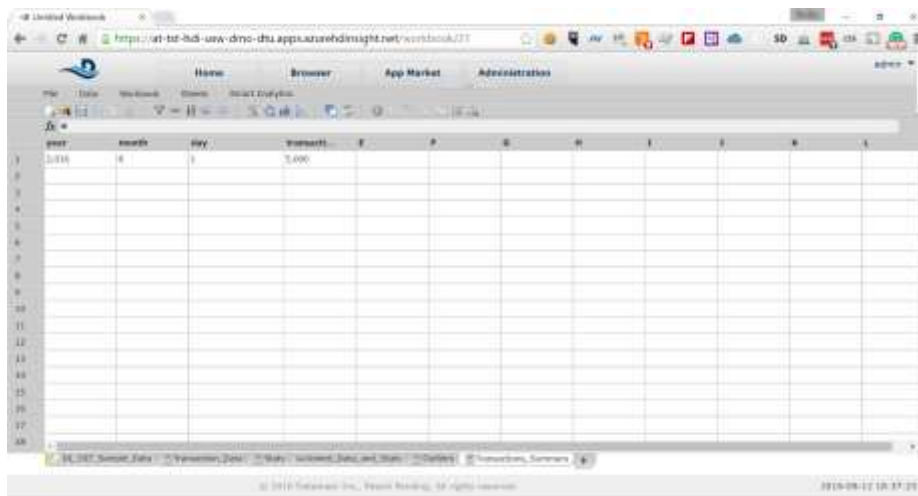
20. We have created summary sheet for our transaction data. Rename the field names as follows:

year

month

day

transactions_count

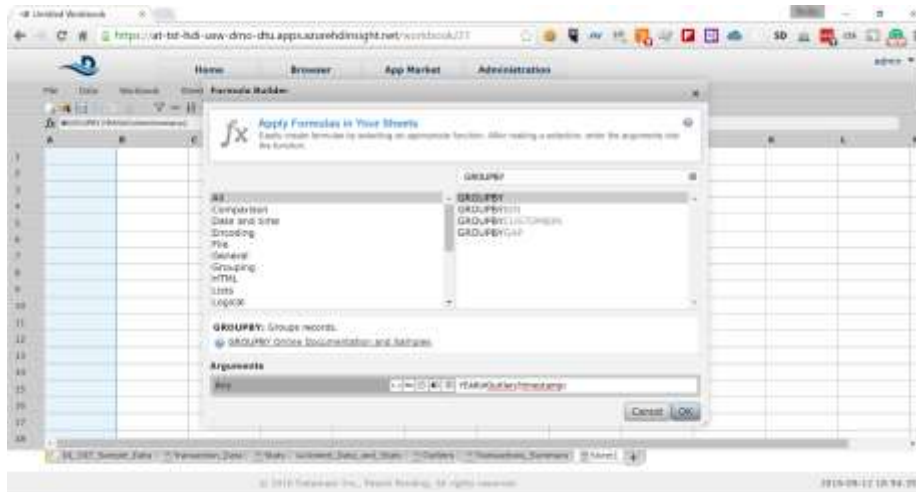


Also, rename the sheet to *Transactions_Summary*

HOL Guide for Enterprise Risk Analysis

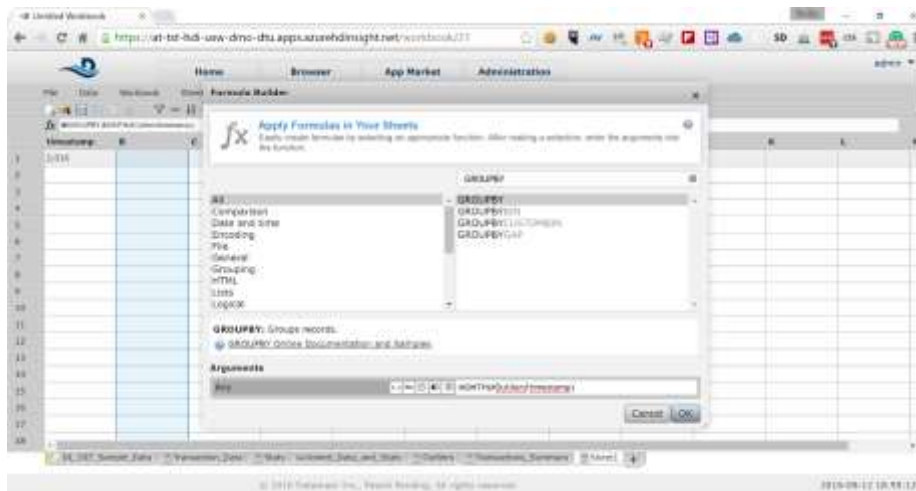
21. Let's create similar summary for the outliers. Create new sheet and in the formula pop-up select the *GROUPBY* function. In the *Arguments* field type the following formula:

YEAR(#Outliers!timestamp)



22. Click on the next column and in the formula pop-up select again the *GROUPBY* function and paste the following formula in the *Arguments* field:

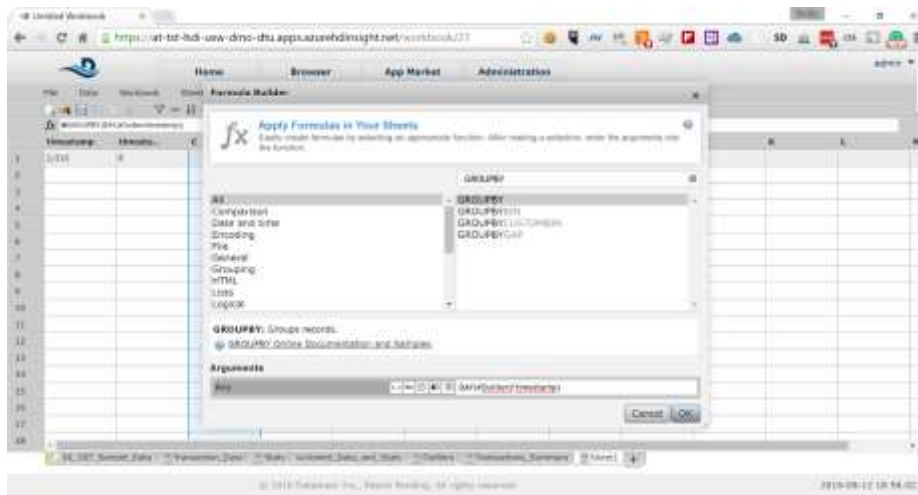
MONTH(#Outliers!timestamp)



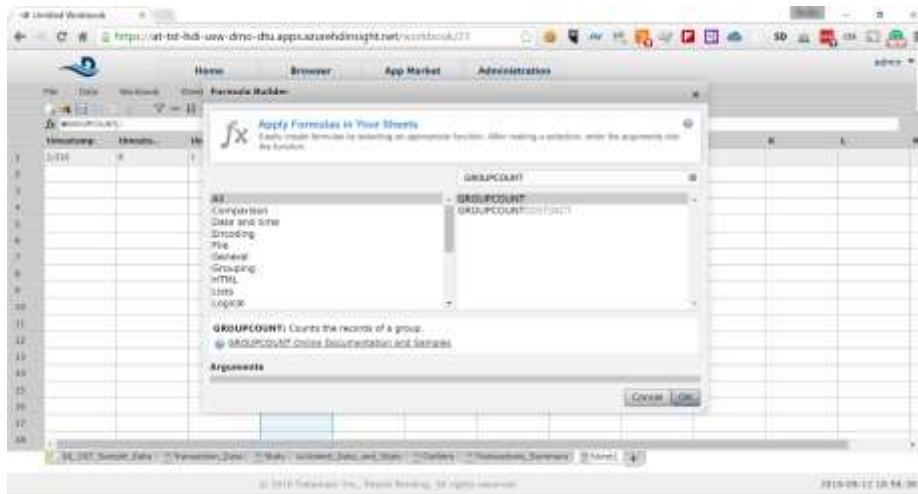
HOL Guide for Enterprise Risk Analysis

23. Click on the third column and in the formula pop-up select again the *GROUPBY* function and paste the following formula in the *Arguments* field:

DAY(#Outliers!timestamp)



24. Click on the fourth column and in the formula pop-up select the *GROUPCOUNT* function and click the *OK* button



HOL Guide for Enterprise Risk Analysis

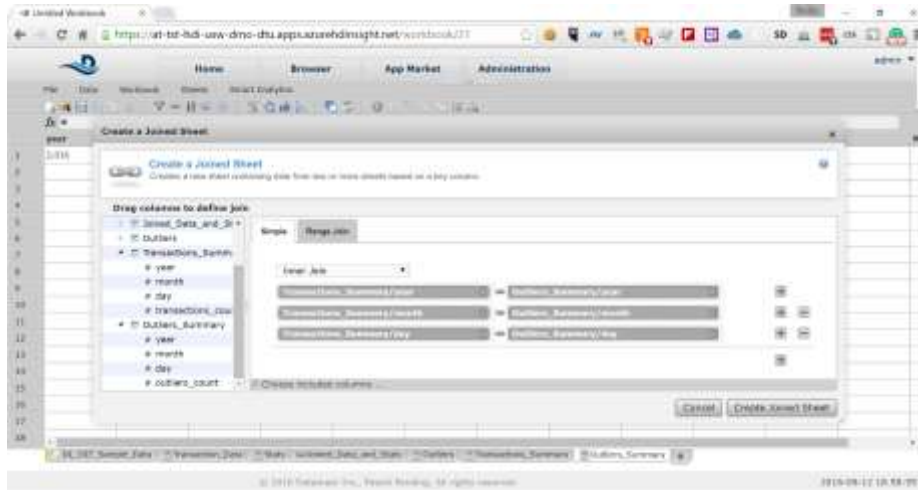
25. We have created summary sheet for our transaction data. Rename the field names as follows:
- year*
 - month*
 - day*
 - outliers_count*

year	month	day	outliers_count	F	G	H	I	J	K	L	M
2015	1	1	50,170								
2015	1	2	49,684								
2015	1	3	50,130								
2015	1	4	50,289								
2015	1	5	50,675								
2015	1	6	49,833								
2015	1	7	50,342								
2015	1	8	50,280								
2015	1	9	49,584								
2015	1	10	49,891								
2015	1	11	50,066								
2015	1	12	49,961								
2015	1	13	49,540								
2015	1	14	49,789								
2015	1	15	50,280								
2015	1	16	50,115								
2015	1	17	50,187								

Also, rename the sheet to *Outliers_Summary*

26. We need to join the two summary sheets to have the results available in a single sheet for visualization. Select *Data -> Join* and join the *Transactions_Summary* and *Outliers_Summary* sheets by year, month and date as on the picture below by clicking on the *Create Joined Sheet* button

HOL Guide for Enterprise Risk Analysis



Rename the joined sheet to *Joined_Summary*

27. Let's copy the joined sheet and remove the duplicate data from it. Right-click on the *Joined_Summary* sheet and select *Duplicate*. Select the following fields in the pop-up:

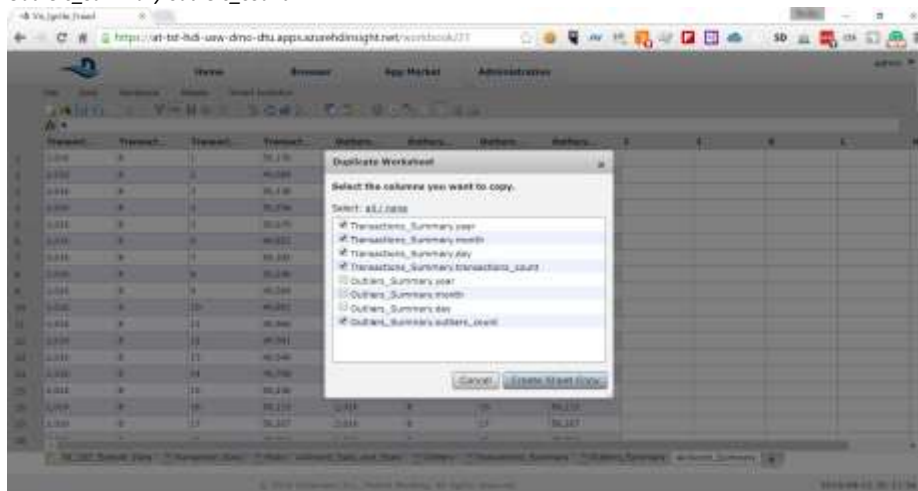
Transactions_Summary.year

Transactions_Summary.month

Transactions_Summary.day

Transactions_Summary.transactions_count

Outliers_Summary.outliers_count



HOL Guide for Enterprise Risk Analysis

28. Rename the sheet to *Summary* and the columns as follows:
- Transactions_Summary.year* -> *year*
 - Transactions_Summary.month* -> *month*
 - Transactions_Summary.day* -> *day*
 - Transactions_Summary.transactions_count* -> *transactions_count*
 - Outliers_Summary.outliers_count* -> *outliers_count*

The screenshot shows the Tableau Desktop interface. At the top, there's a navigation bar with tabs for Home, Browser, App Market, and Administration. Below this is a toolbar with various icons for file operations, data manipulation, and visualization. The main workspace displays a table of data with the following columns: year, month, day, framework, and authors_count. The authors_count column is highlighted with a black box and labeled 'authors_count (INTEGER)'. The table contains data for the years 2015 and 2016, with months ranging from 1 to 12. The authors_count values are 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, and 21. The bottom of the interface shows a status bar with the text '© 2015 Tableau Software, Inc. All rights reserved.' and a timestamp '2015-09-17 10:20'.

29. Click on the sixth column and cancel the formula pop-up. In the f_x field on top of the sheet type the following:

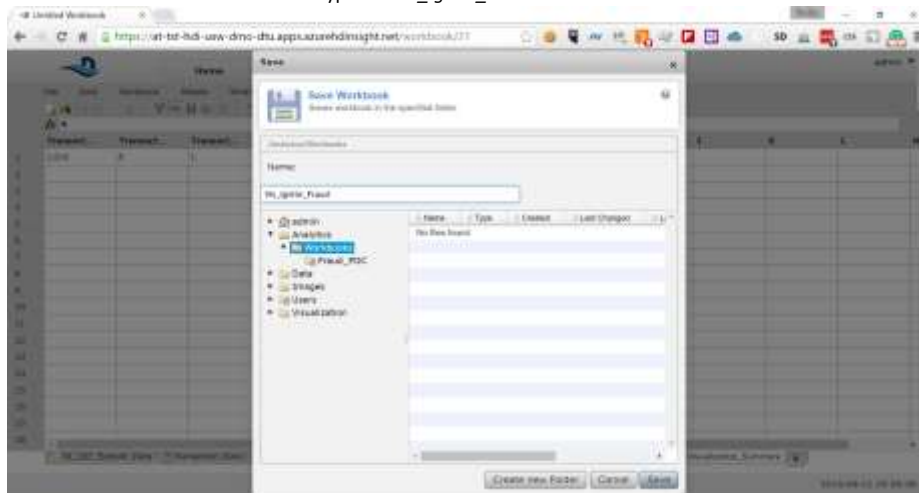
$$(\#Summary!outliers_count/\#Summary!transactions_count) * 100$$

HOL Guide for Enterprise Risk Analysis

year	month	day	transaction	referrer	appid
2014	8	1	50,470	1	0.0000000000000000
2014	8	8	40,852	1	0.0000000000000000
2014	8	11	40,846	1	0.0000000000000000
2014	8	16	50,178	8	0.0000000000000000
2014	8	19	40,842	1	0.0000000000000000
2014	8	20	40,842	1	0.0000000000000000
2014	8	27	50,237	1	0.0000000000000000

Also, rename the field to *outliers_to_transactions*

30. Select *File -> Save* from the menu and type the *Vis_Ignite_Fraud* in the *Name* field

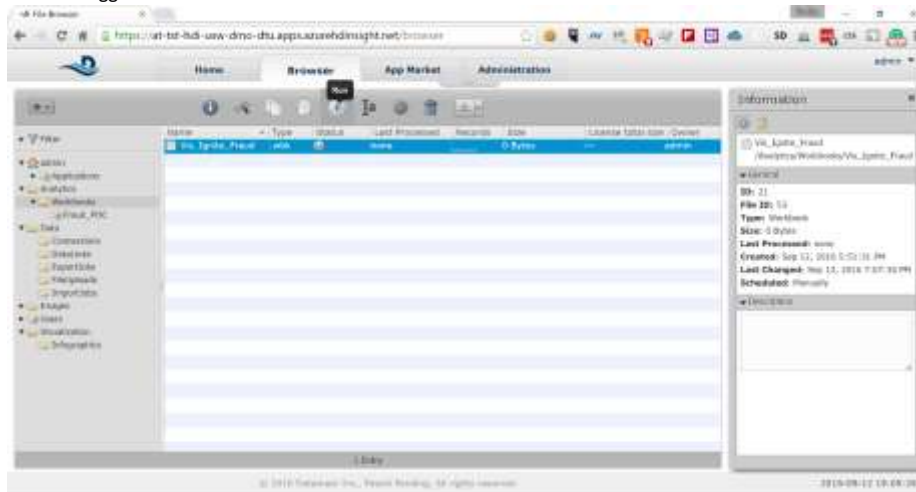


31. On the next screen keep the default values for all the fields. Scroll down and click on the *Save* button again

HOL Guide for Enterprise Risk Analysis

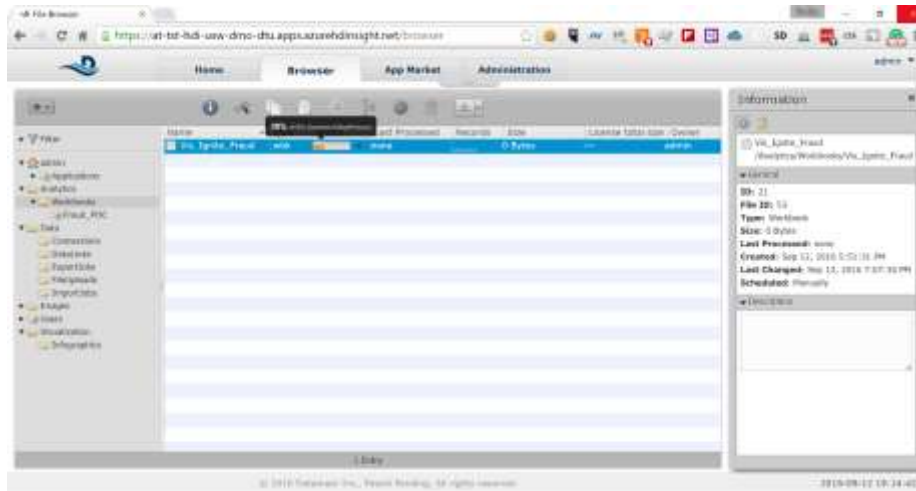


32. In the list of workbooks select the newly created workbook and click on the run button from the toolbar. This will trigger the calculation on the full data set



You will see updates in the *Status* column, showing you how the Hadoop job is progressing.

HOL Guide for Enterprise Risk Analysis



8 Logging in to the TrendMicro DSM

8.1 Server name

From the output section of the deployment you can get the URL for TrendMicroDSM, Splunk and Chef Server (Microsoft.Template)

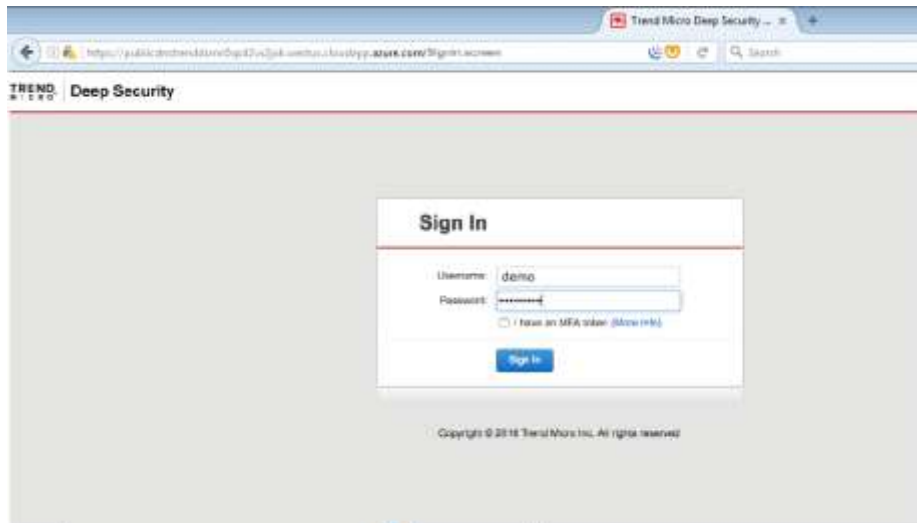


8.2 Server login

To login to TrendMicro DSM

- Paste the TrendMicro DSM URL in the browser
- Enter the **Username** and **Password** provided in the parameter section during the deployment

HOL Guide for Enterprise Risk Analysis



9 Perform policy configuration on the TrendMicro DSM

1. Changing the base policy

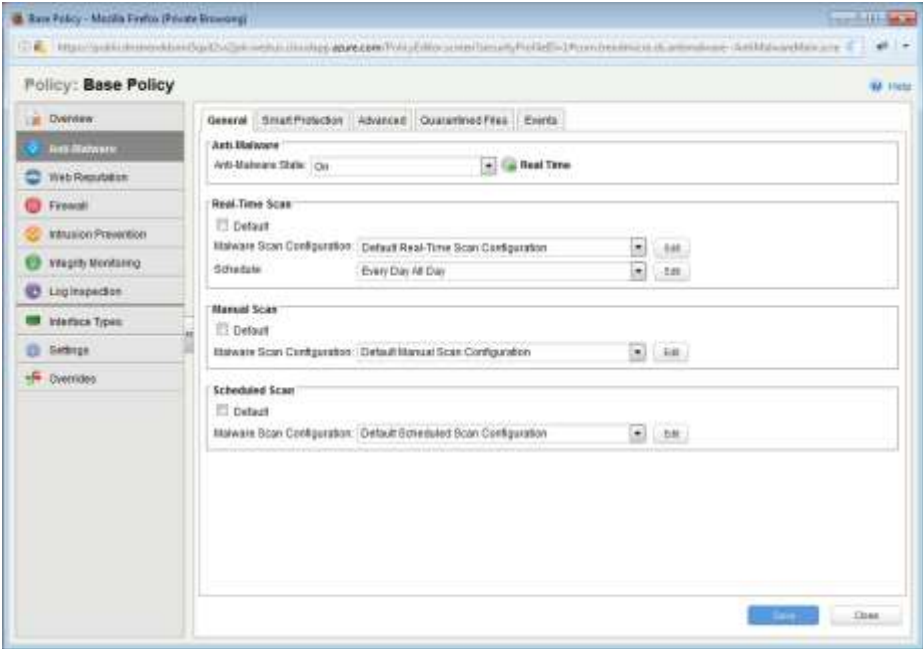
Go to policies->Base Policy



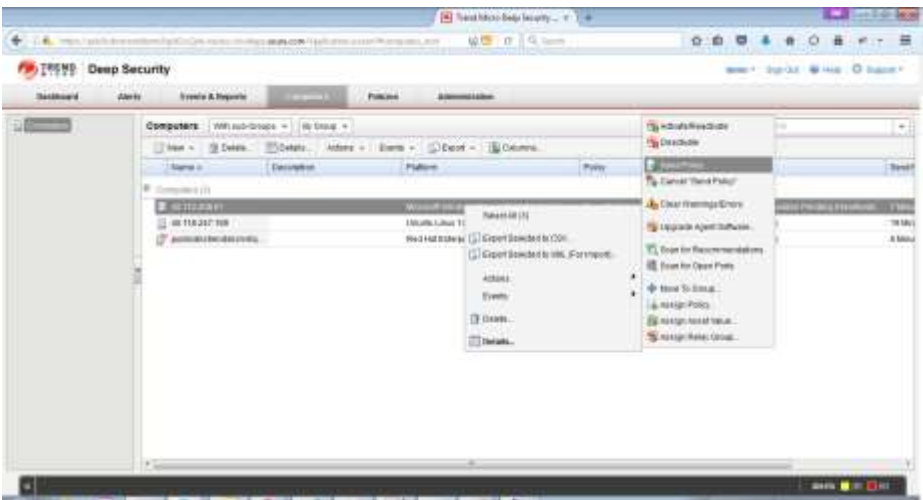
2. Enable Anti-Malware

Go to Anti-malware->Anti-Malware State->On
Click "Save"

HOL Guide for Enterprise Risk Analysis

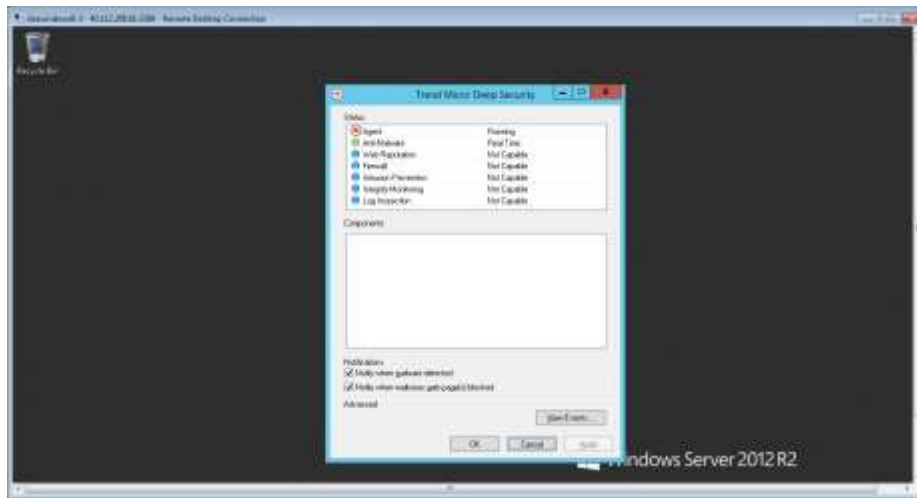


3. Applying policies to computer



HOL Guide for Enterprise Risk Analysis

4. Verifying policy in the computer



10 Exercises

10.1 Datameer – Visualize the Data

Datameer has powerful Infographics to Visualise the data. In this exercise, the data analysed in the above configuration will be displayed graphically.

10.2 TrendMicro – Malware test

TrendMicro has security intelligence built-in to protect the systems against the malwares. In this exercise, showcases the TrendMicro DSM malware detection capability

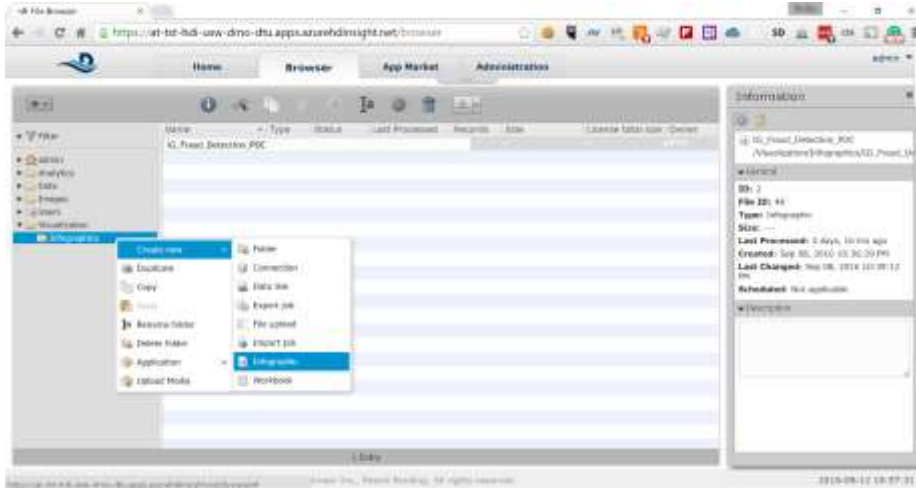
Play with the exercises and enjoy

11 Visualize the Data

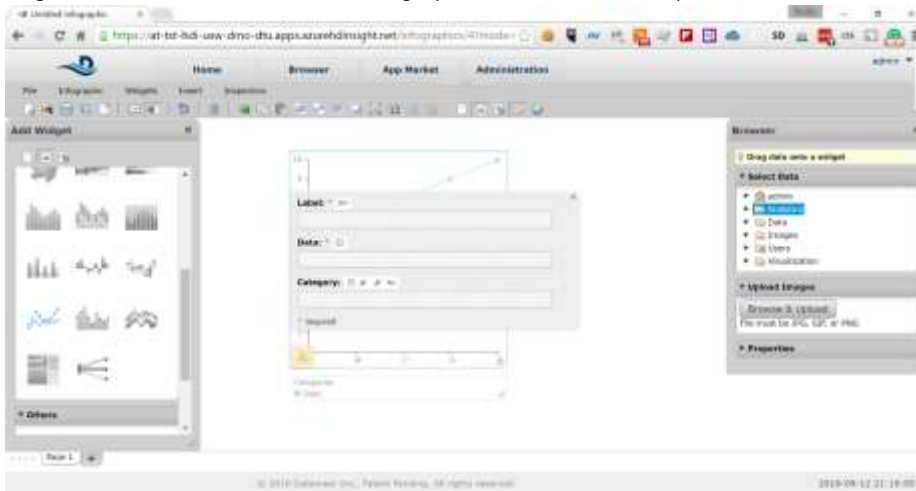
The last exercise in this HOL is to visualize the data and identify certain days when the irregular transactions have spiked. To do that we will use the following steps:

1. In Datameer's Browser view expand the Visualization node and right click on Infographics -> Create New -> Infographic

HOL Guide for Enterprise Risk Analysis

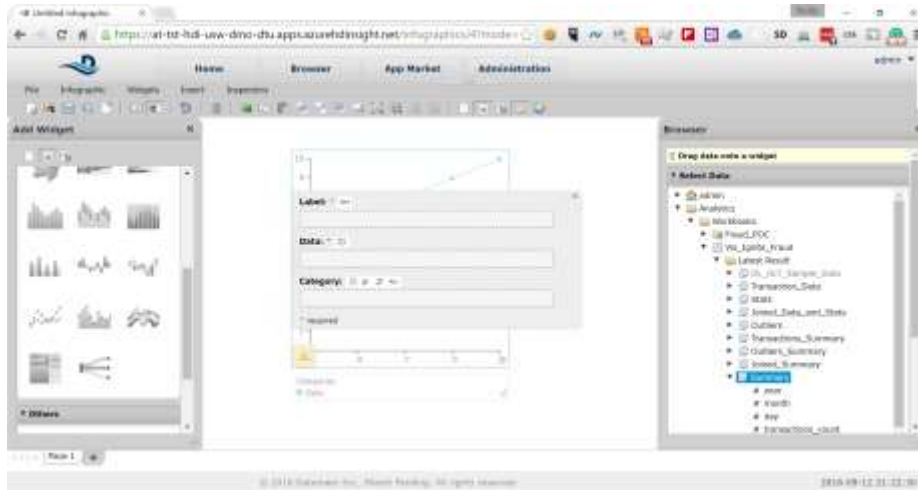


2. Drag the *Line and Area Chart* from the *Add Widget* pane on the left to the work pane in the middle

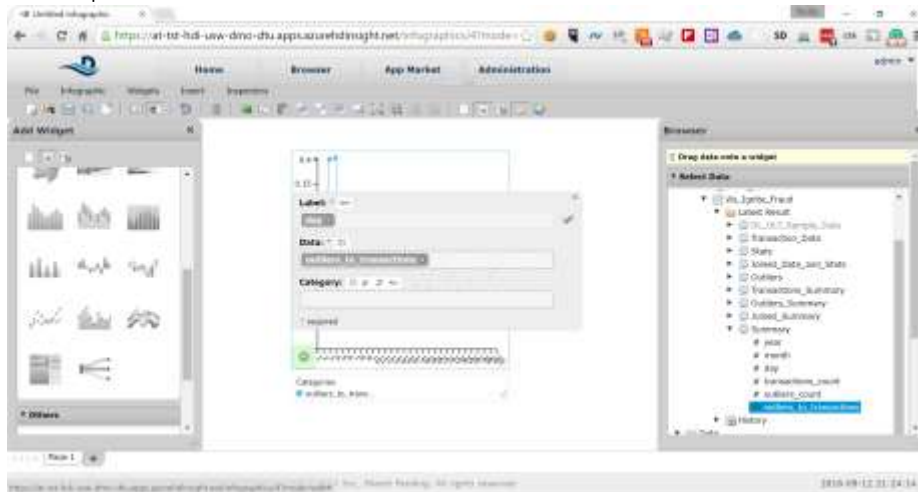


3. In the *Browser* pane expand *Analytics* node and then *Workbooks* -> *Vis_Ignite_Fraud* -> *Latest Results* -> *Summary*

HOL Guide for Enterprise Risk Analysis



4. Drag the `day` field to the `Label` input field and the `outliers_to_transactions` field to the `Data` input field in the Work pane



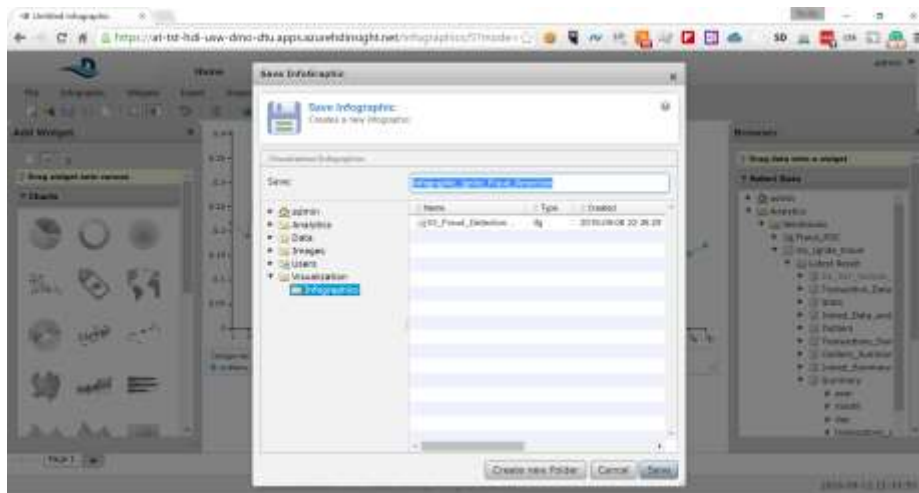
5. Select Infographic -> View from the menu to present the infographic. You can easily see that on the 3rd and 4th day of the month the outliers significantly spiked, which is a sign of something unusual going on those two days

HOL Guide for Enterprise Risk Analysis



6. Select Infographic -> Edit from the Menu and then File -> Save. Type the following in the Name field:

Infographic_Ignite_Fraud_Detection
and click on the Save button



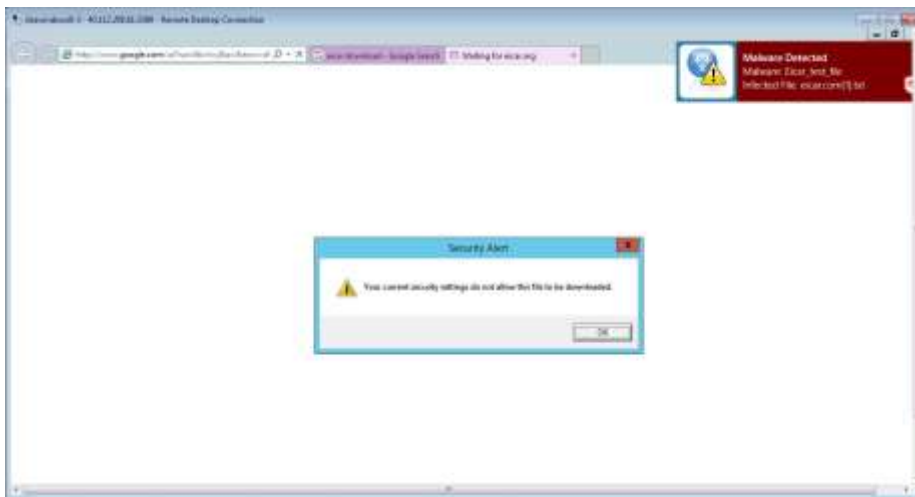
12 Malware Test

12.1 Generating Malware alert in the computer

The Malware test can be performed by going to the url

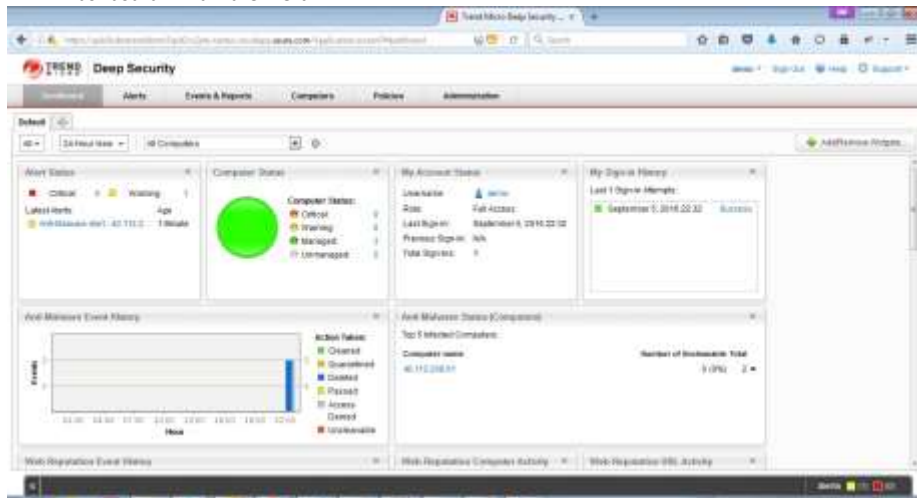
<https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&sqi=2&ved=0ahUKEwj6n6H1-IXPAhUSzGMKHZMGC5AQFggmMAE&url=http%3A%2F%2Fwww.eicar.org%2Fdownload%2Ffeicar.com.txt&usg=AFQjCNE8DvVI7BE5Nd2hg1zNDTP6hNjclA&bvm=bv.132479545,d.cGc>

this is eicar malware test

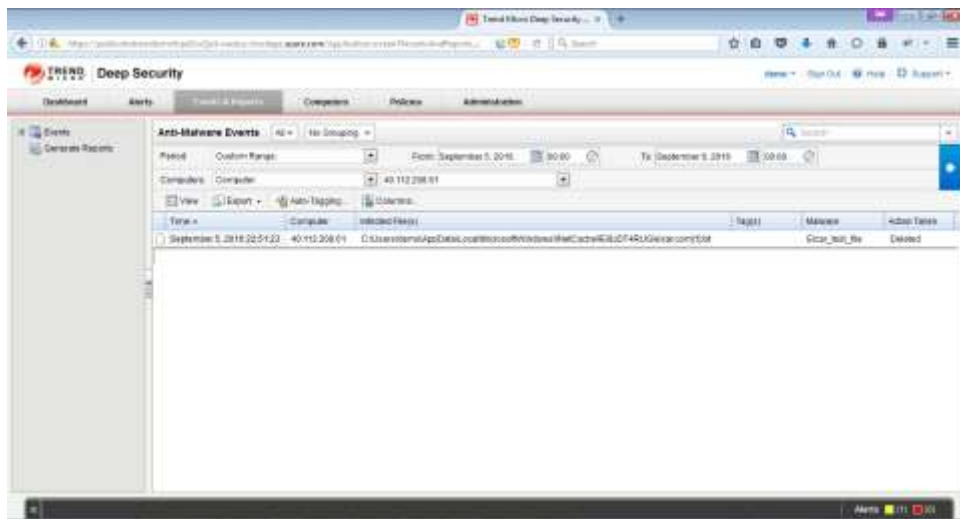


HOL Guide for Enterprise Risk Analysis

12.2 Dashboard – Malware Alert



12.3 Malware Alert verification



13References, Attachments & Definitions (Respective track leads)

13.1 References

No.	Document Title	Link/ Attachment	Comments
1			