Dung Dinh

Oracle Corporation | [Company address]

Oracle Analytics Cloud

Assignment

# Pre-requisite Workshop

* You need to have an OAC account with DV Author role at least
* If you don’t have it, please contact Thomas’s team to create new for you.

# Data Sources

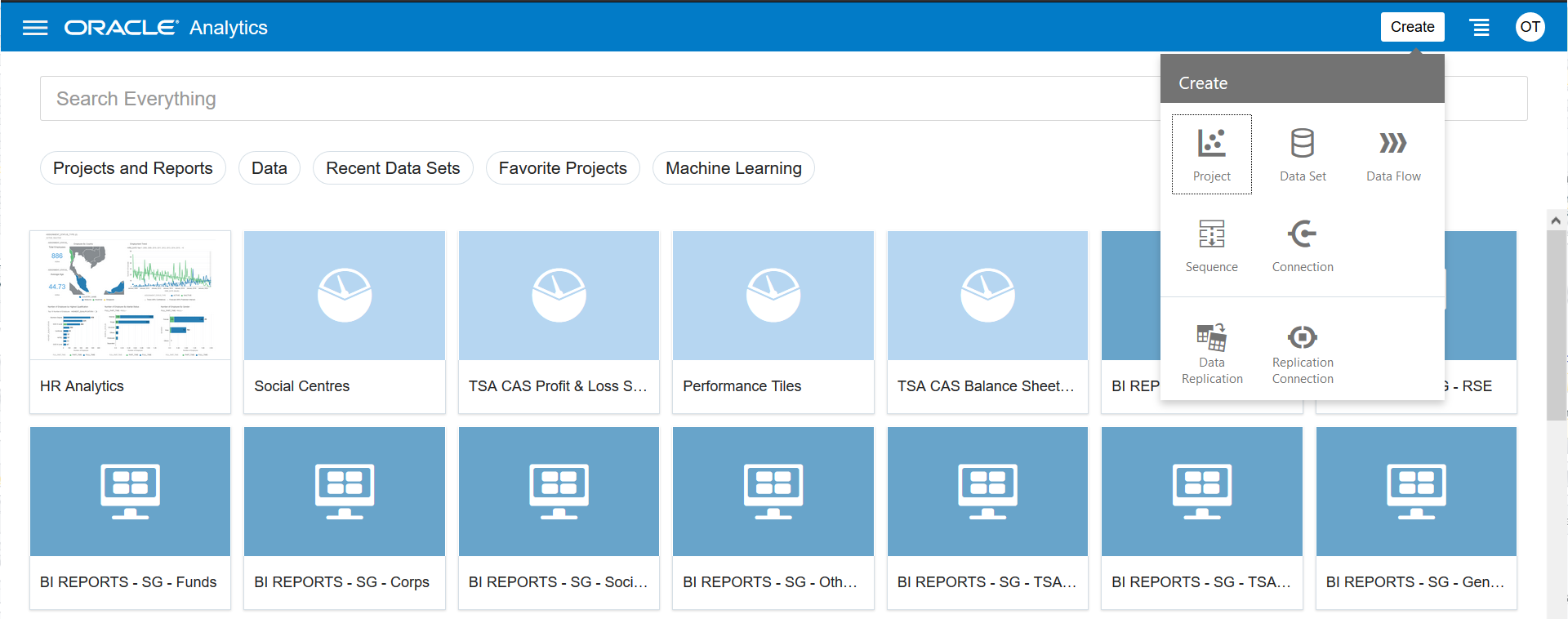
* The following links are sample datasets for OAC Workshop training. You are able to download them to do OAC assignments
* Calendar Time dataset
  + <https://objectstorage.us-ashburn-1.oraclecloud.com/n/id66dobbdxlj/b/OAC_Training/o/Calendar%20Time.csv>
* Human Resource Data Set
  + <https://objectstorage.us-ashburn-1.oraclecloud.com/n/id66dobbdxlj/b/OAC_Training/o/HR-Employee-Attrition.csv>
  + <https://objectstorage.us-ashburn-1.oraclecloud.com/n/id66dobbdxlj/b/OAC_Training/o/HR_Turnover_Data-Updated.xlsx>
* Financial Dataset
  + <https://objectstorage.us-ashburn-1.oraclecloud.com/n/id66dobbdxlj/b/OAC_Training/o/Financial%20-%20Updated.xlsx>

# Data Preparation

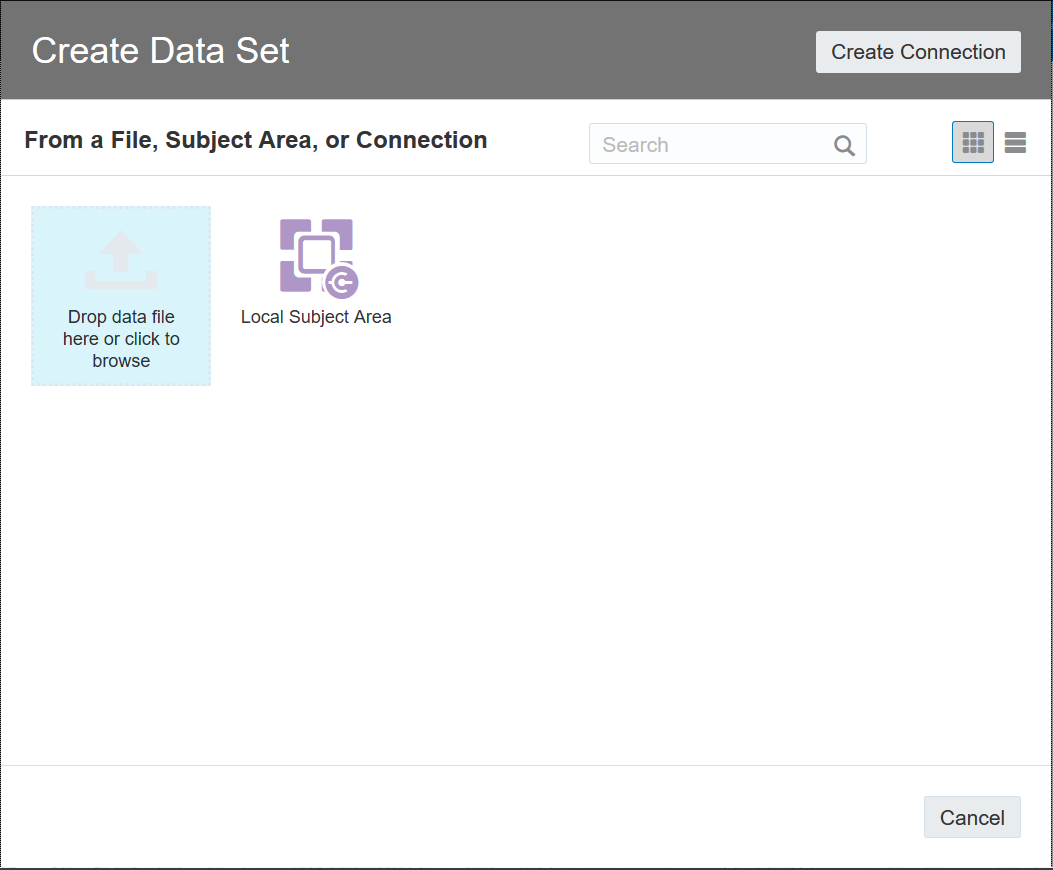
* Log in to OAC Cloud platform by accessing the URL <https://idcs-7c160756791f45afba69457a0f26df3e.identity.oraclecloud.com/ui/v1/signin>
* Upload external Datasets to OAC. HR team use Human Resource datasets and Finance team use the Financial dataset what you downloaded

## Financial Dataset Preparation

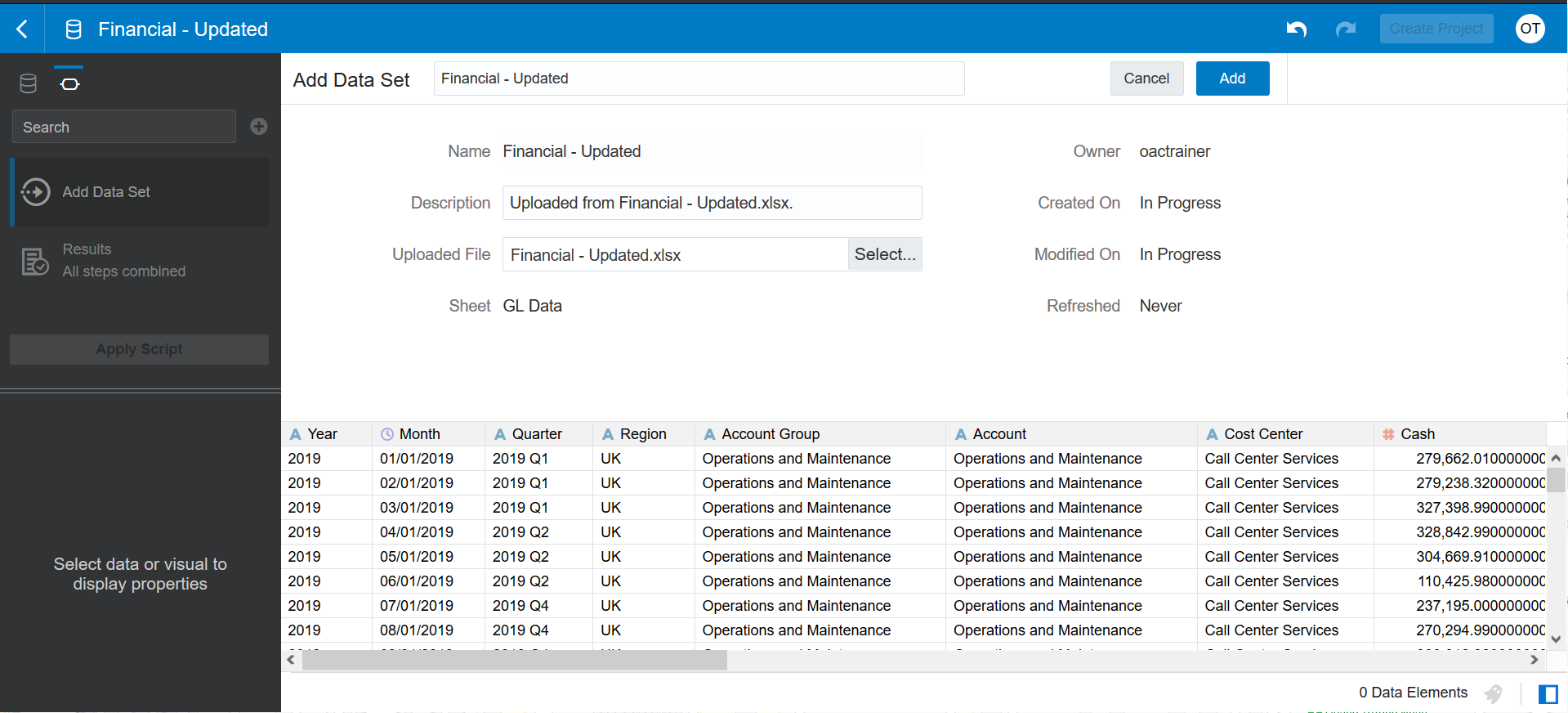
* Click on Create button on top-right corner 🡪 Data Set



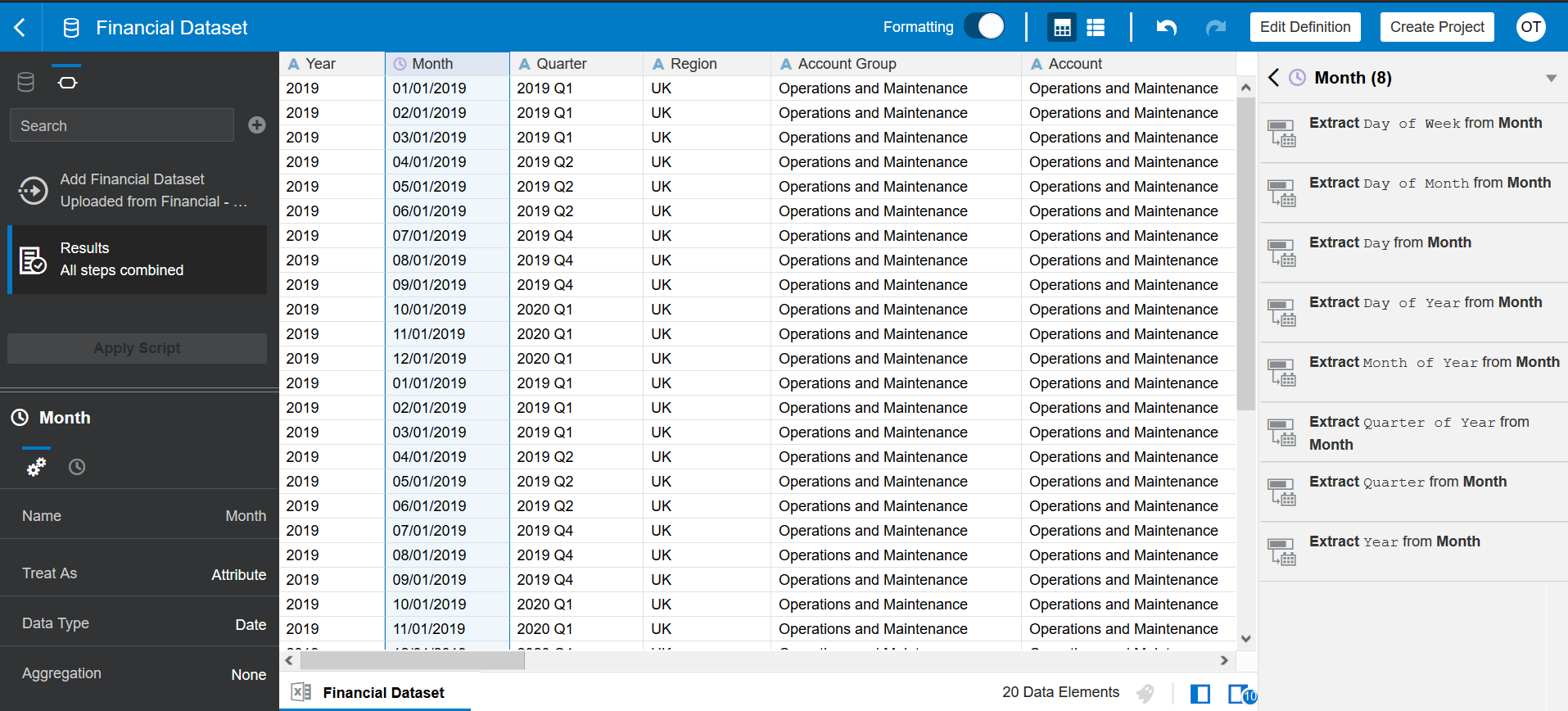
* Create Data Set popup appears 🡪 Choose Drop data file here or click to browse



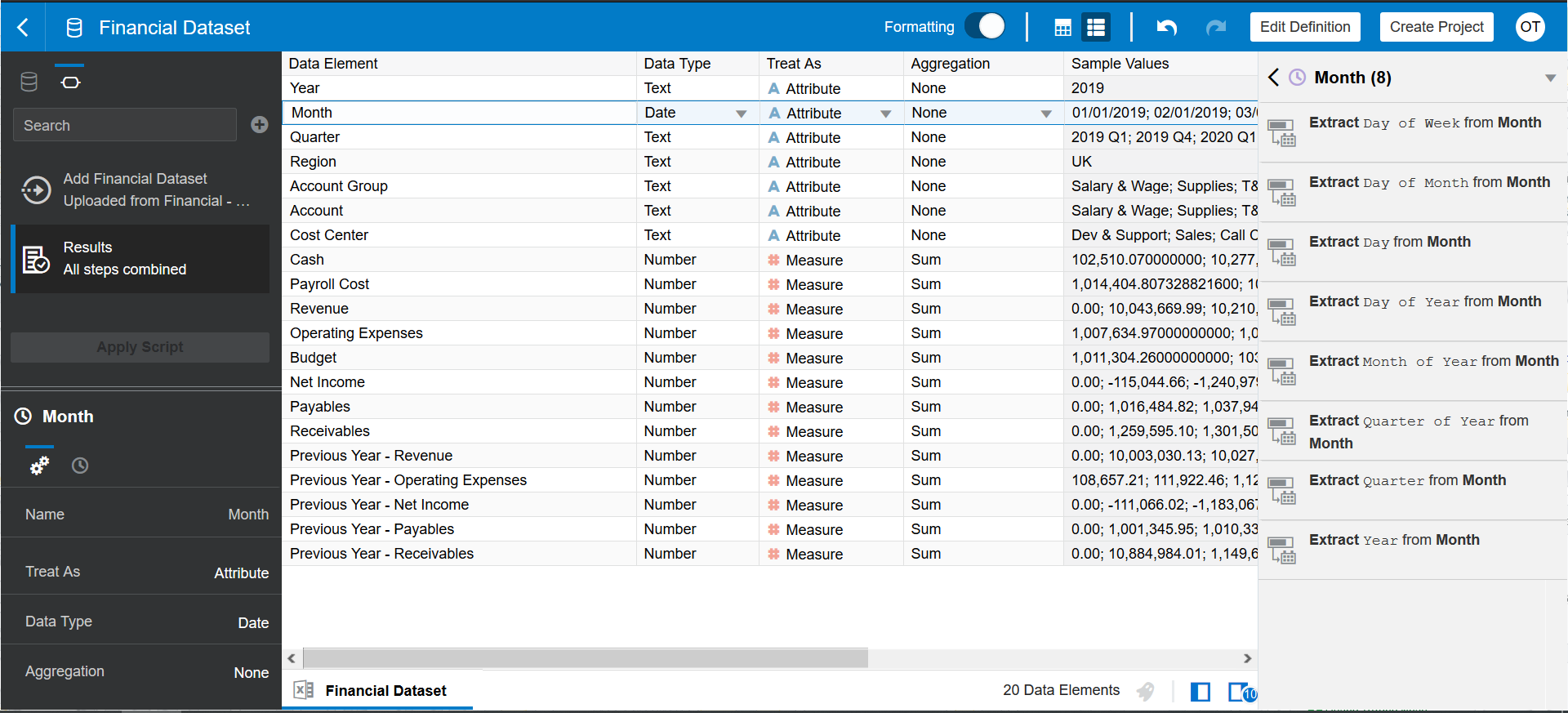
* Browse to your dataset in local location such Financial dataset



* Enter name of dataset: Financial Dataset
* Click Add button



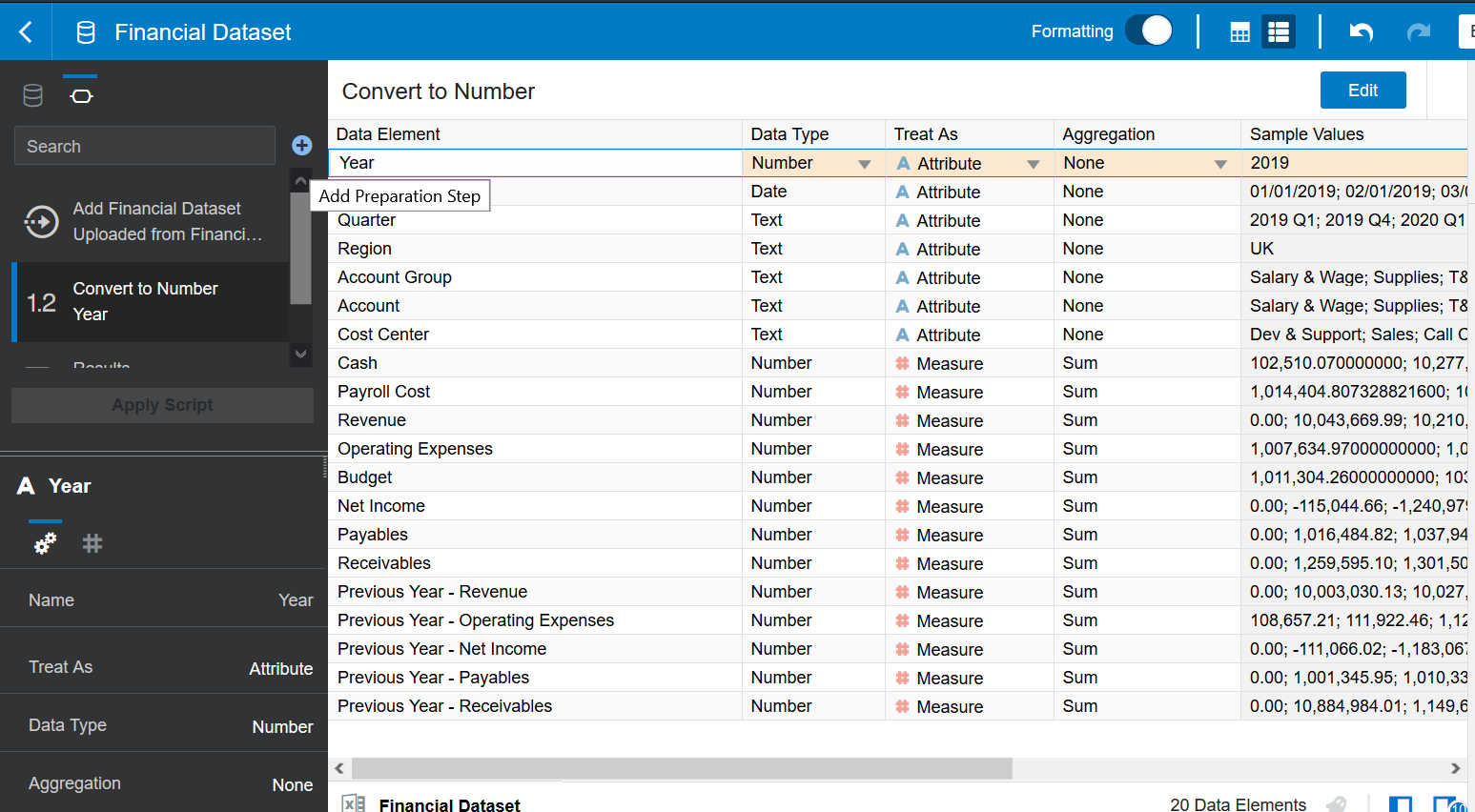
* Click on the Metadata icon to review metadata of dataset



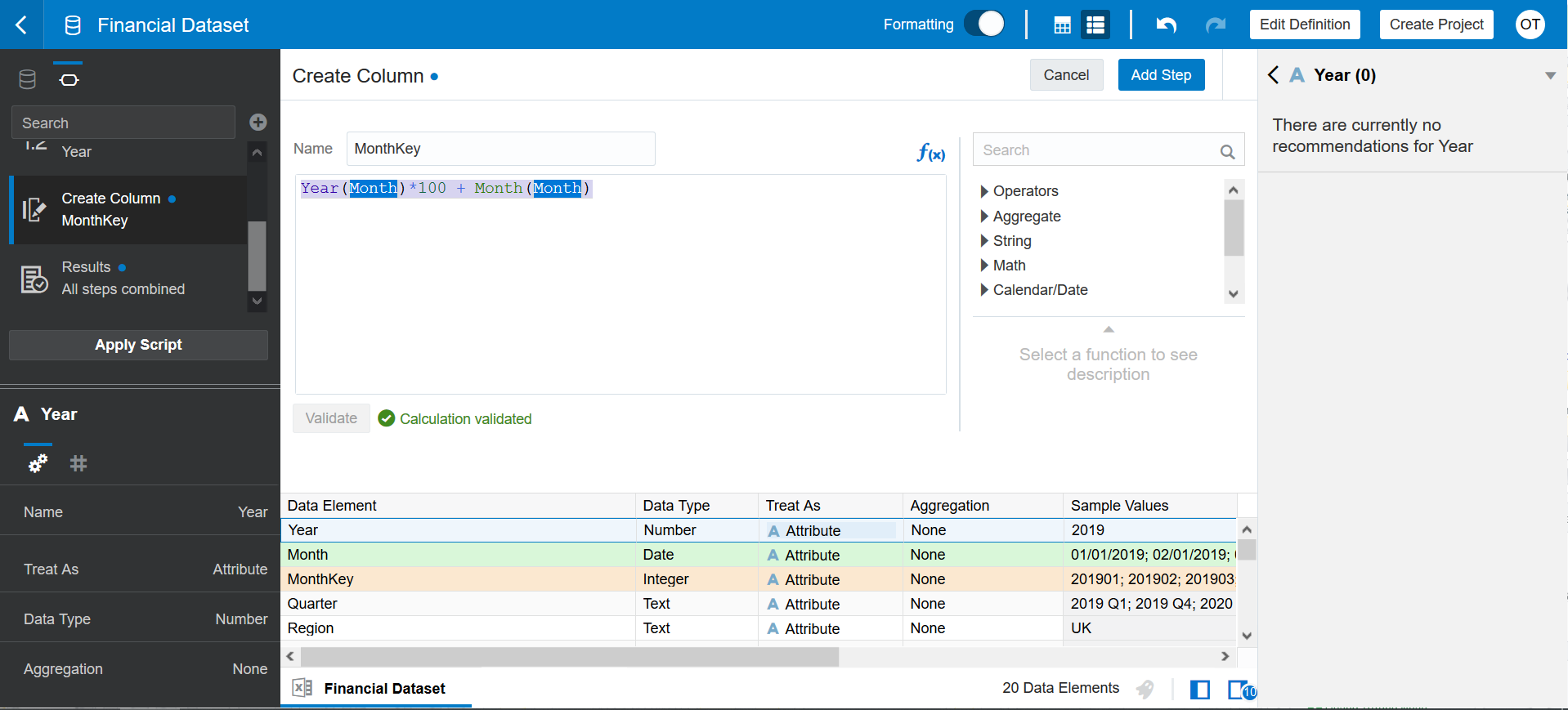
* Update the Data Type, Properties and Aggregation functions accordingly

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Data Type** | **Treat As** | **Aggregation** |
| Year | Number | Attribute | None |
| Month | Date | Attribute | None |
| Quarter | Text | Attribute | None |
| Region | Text | Attribute | None |
| Account Group | Text | Attribute | None |
| Account | Text | Attribute | None |
| Cost Center | Text | Attribute | None |
| Cash | Number | Measure | Sum |
| Revenue | Number | Measure | Sum |
| Operating Expenses | Number | Measure | Sum |
| Budget | Number | Measure | Sum |
| Net Income | Number | Measure | Sum |
| Payables | Number | Measure | Sum |
| Receivables | Number | Measure | Sum |
| Previous Year - Revenue | Number | Measure | Sum |
| Previous Year - Operating Expenses | Number | Measure | Sum |
| Previous Year - Net Income | Number | Measure | Sum |
| Previous Year - Payables | Number | Measure | Sum |
| Previous Year - Receivables | Number | Measure | Sum |

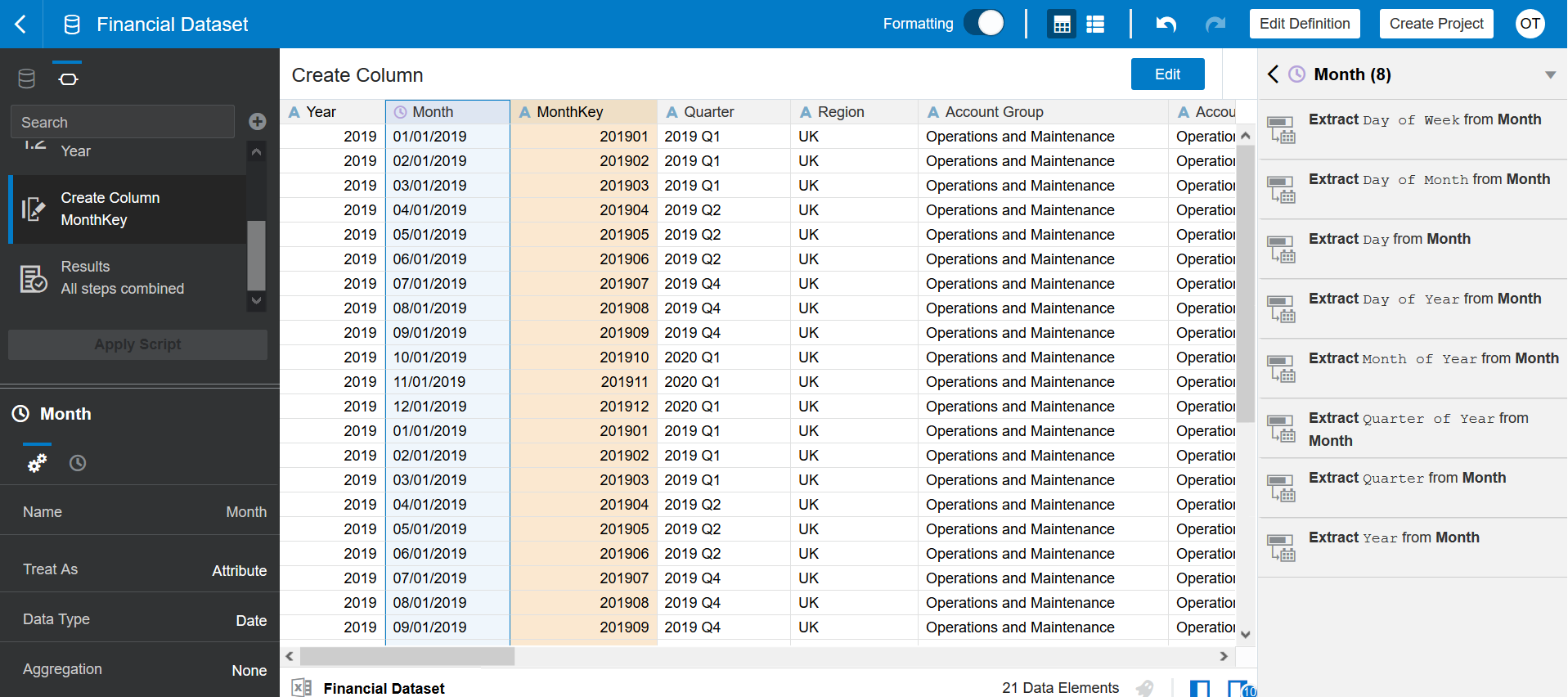
* Click Apply Script. You need to run Apply Script after you make any change on dataset. Otherwise, your changes will not be updated.
* Create a new column name MonthKey. The column is extracted from the Month column by using Expression, it is Integer datatype and YYYYMM format.
* Click Plus icon to Add Preparation Step



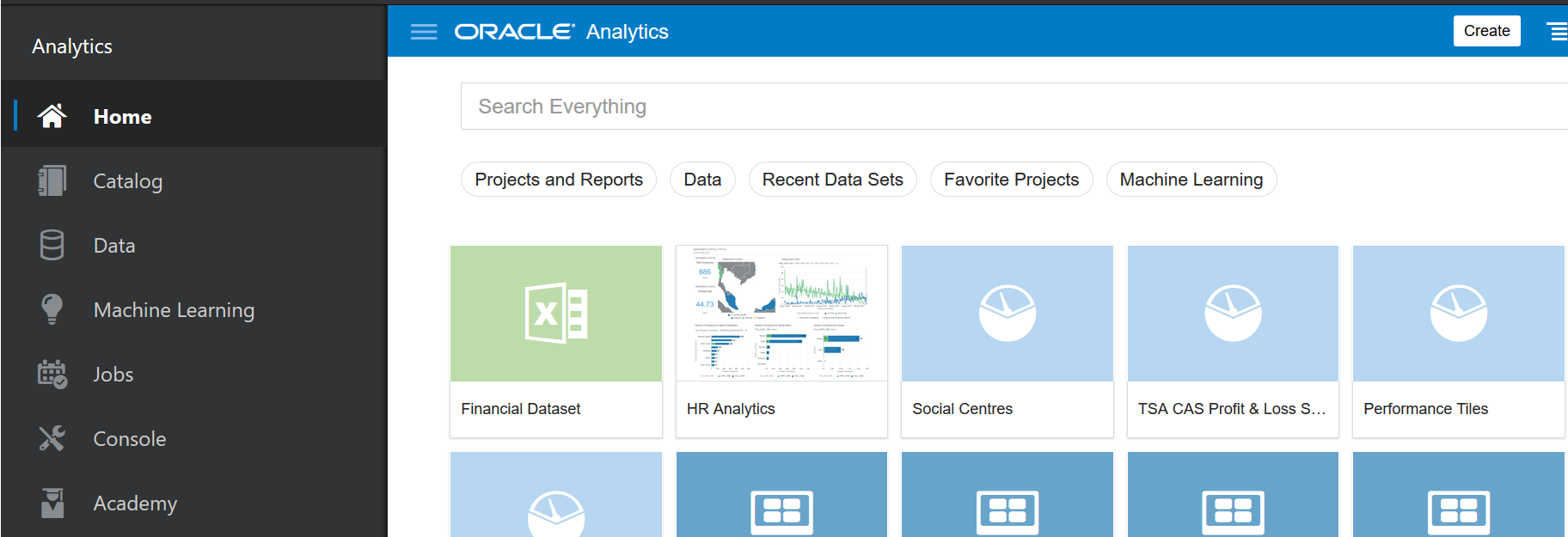
* + Enter name of new column = MonthKey
  + Build expression: Year(Month)\*100 + Month(Month)
  + Click Validate to make sure there is no invalid syntax
  + Click Add Step
  + Click Apply Script



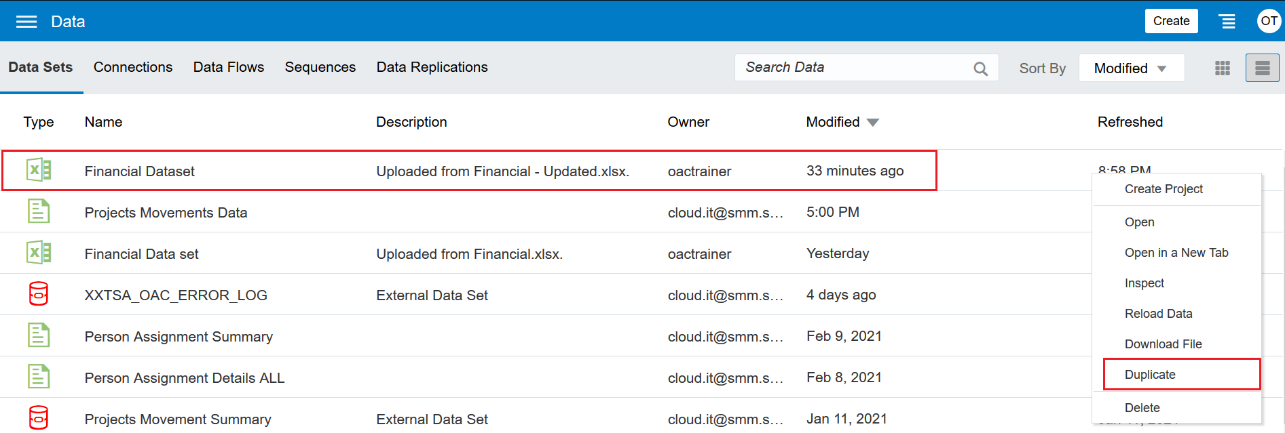
* Click on Data icon to view data



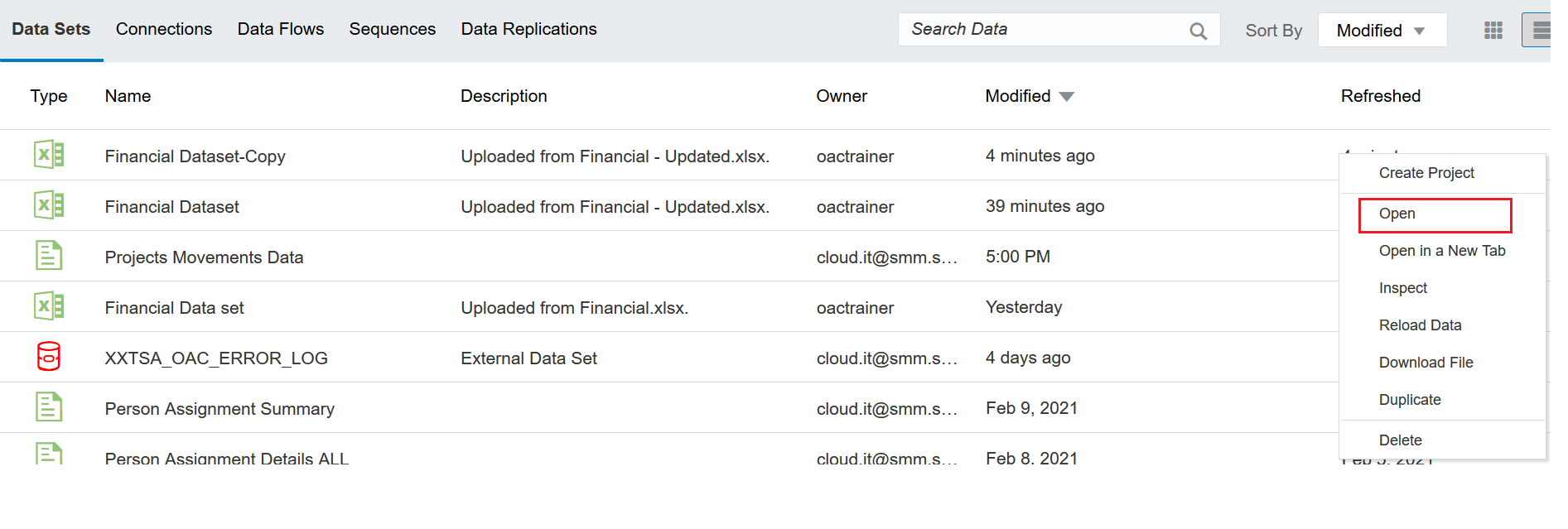
* After you upload a dataset, the AI/ML engine of OAC run and provide recommendation for extracting data from existing column such as date, geography, email, phone, currency,…
* Based on Financial dataset, OAC knows Month column which is Date, so it recommended some extraction from Month column.
  + Day Of Week
  + Day Of Month
  + Day Of Year
  + Month Of Year
  + Quarter of Year
  + Quarter
* Click on a hamburger icon on top-left corner (Navigator) 🡪 Data



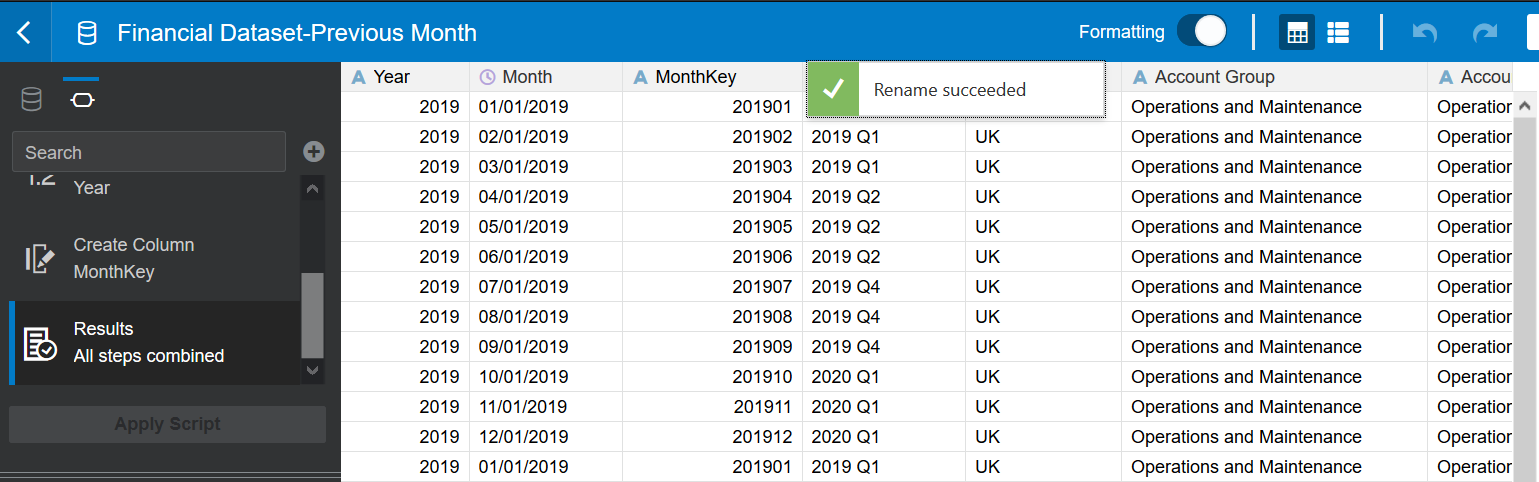
* Choose Financial Dataset 🡪 Duplicate



* A new entry is created but it does not create a new data source, it only creates the new entry in OAC pointing to the same data source
* Click to open the new dataset



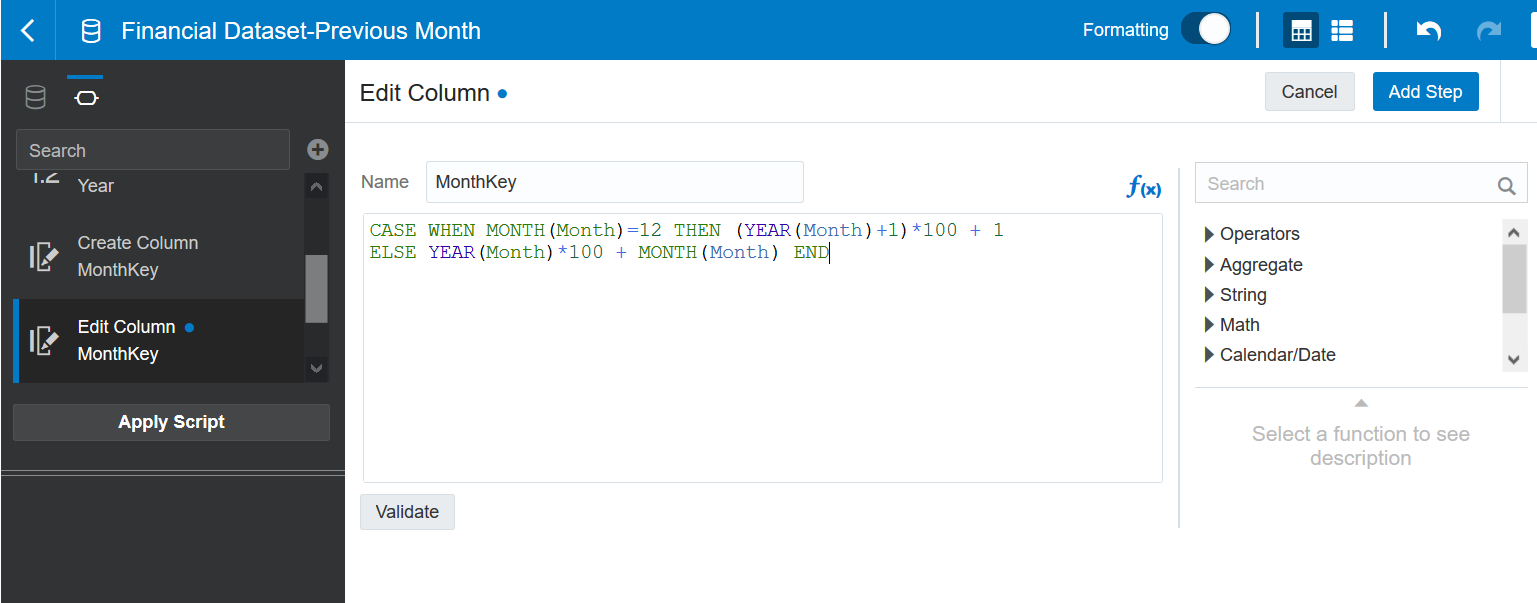
* Change name of new dataset to Financial Dataset-Previous Month. This dataset is used to build a few of Time analysis such **Month Over Month**



* Click on MonthKey column 🡪 Edit and update the expression accordingly

CASE WHEN MONTH(Month)=12 THEN (YEAR(Month)+1)\*100 + 1

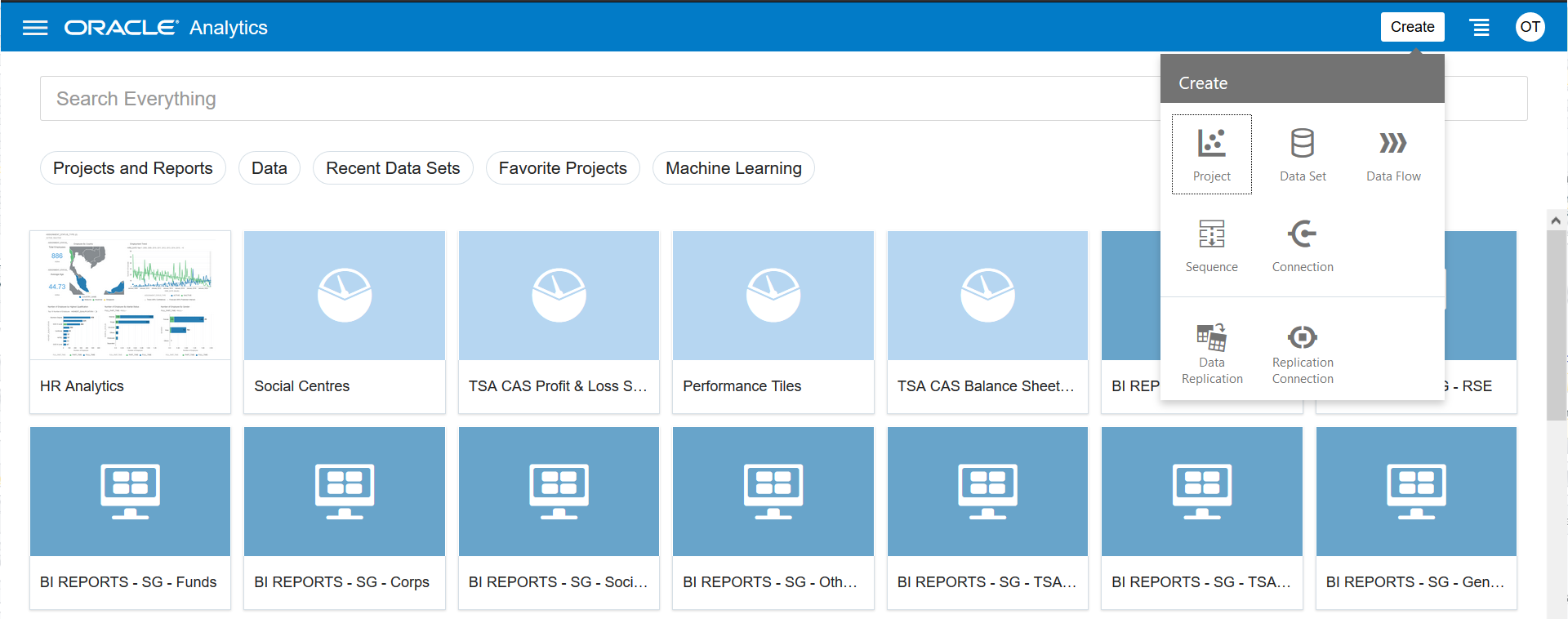
ELSE YEAR(Month)\*100 + MONTH(Month) +1 END



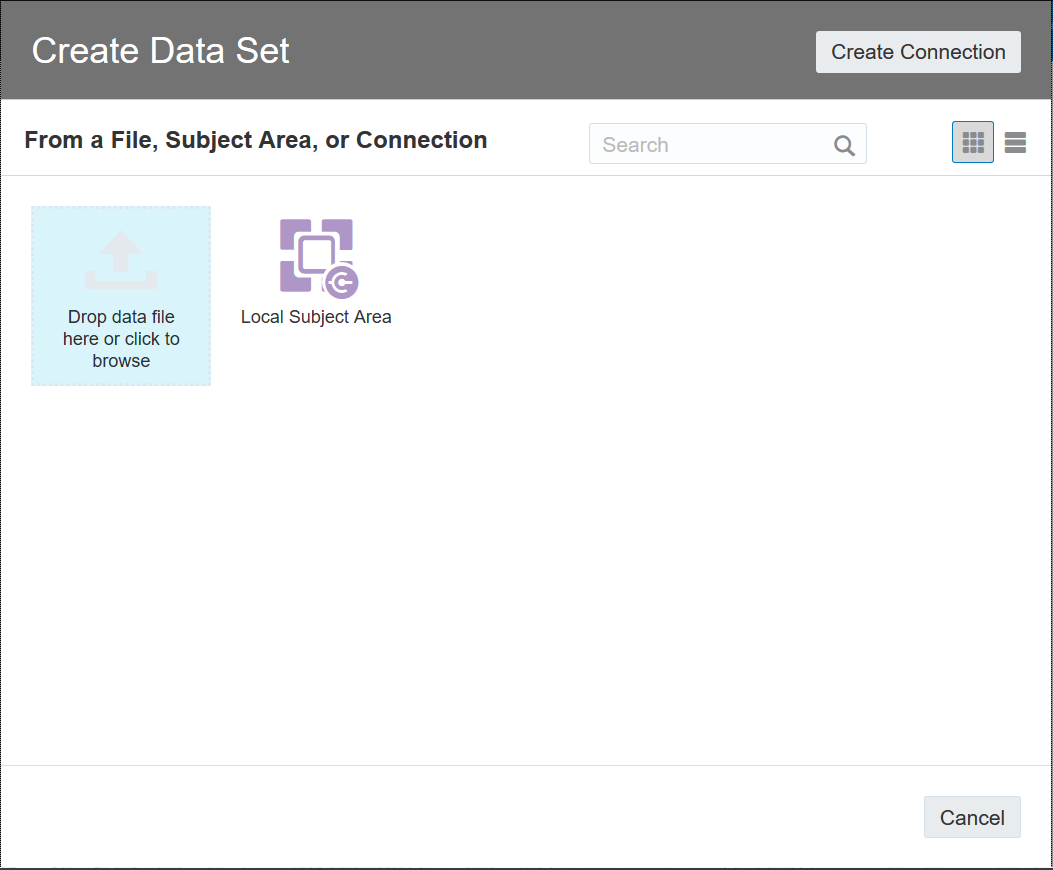
* Click Add Step, and then Apply Script

## HR Dataset Preparation

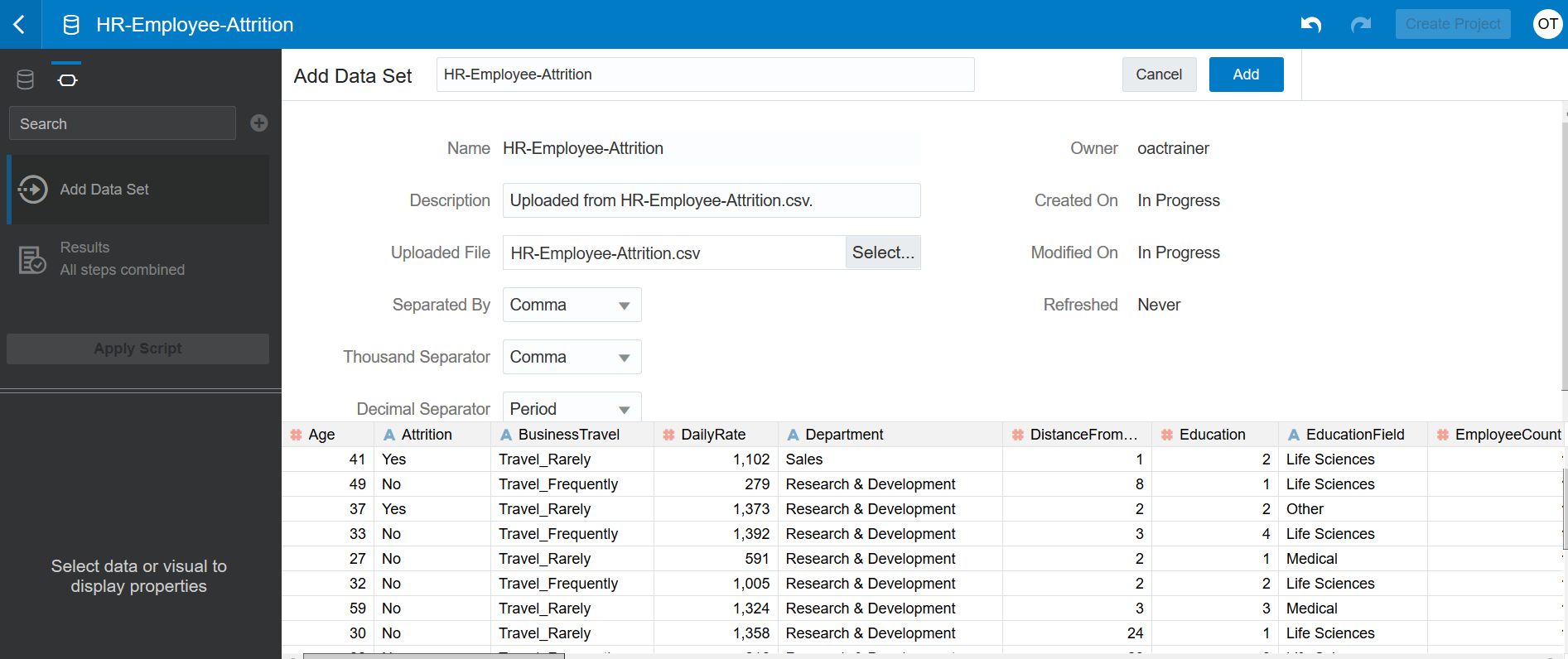
* Click on Create button on top-right corner 🡪 Data Set



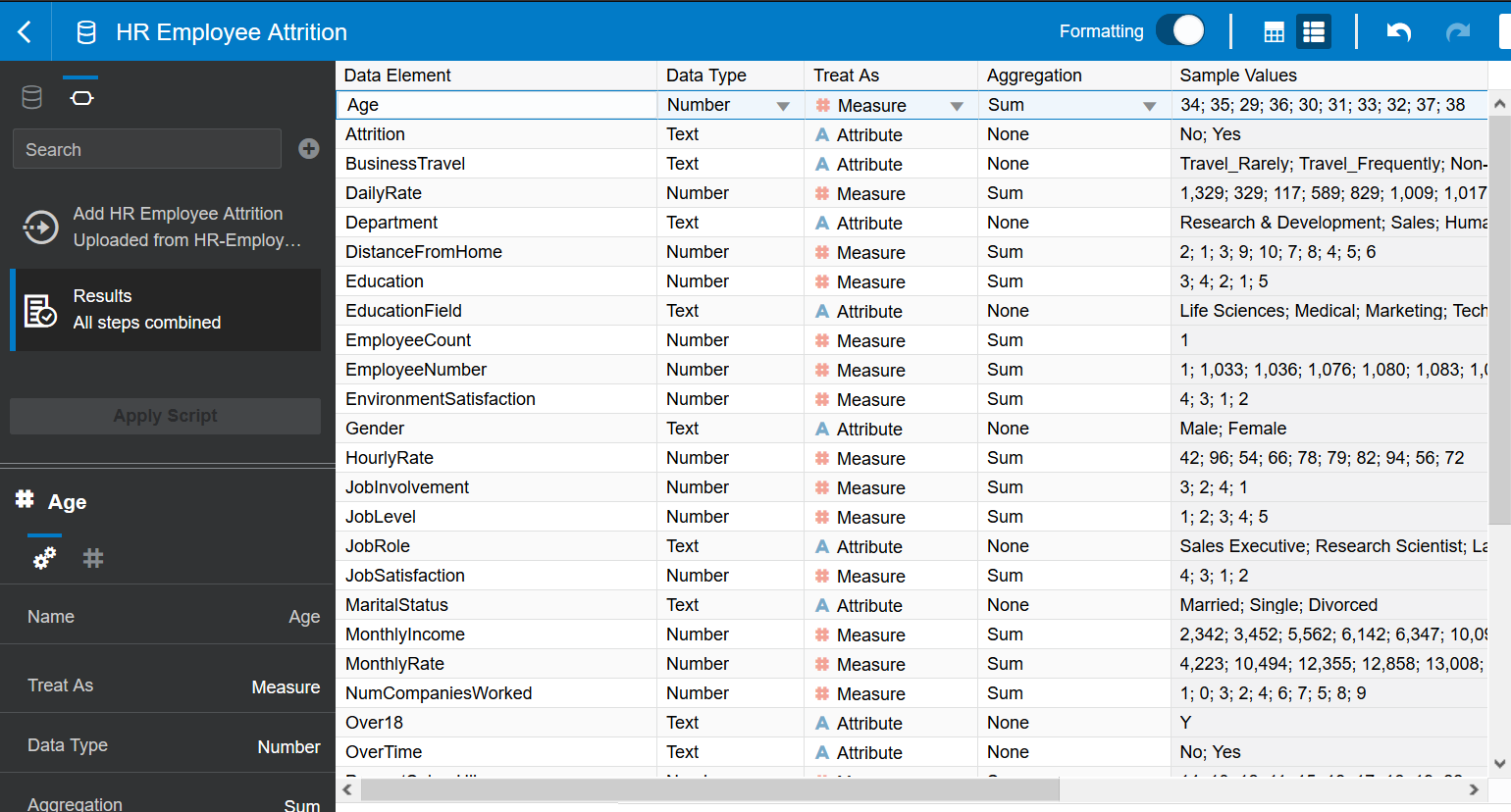
* Create Data Set popup appears 🡪 Choose Drop data file here or click to browse



* Browse to your dataset HR-Employee-Attrition.csv in local location

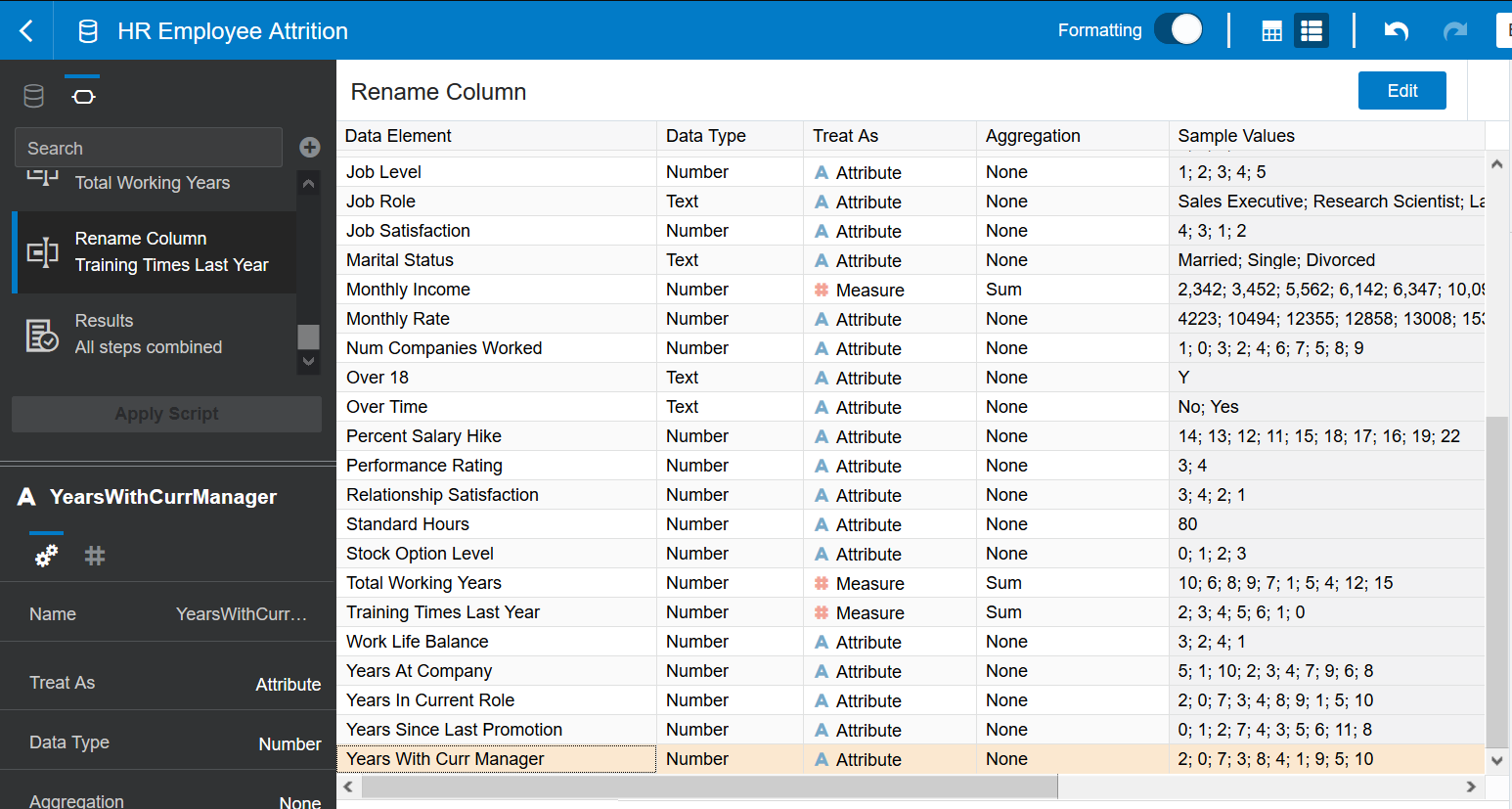


* Rename dataset HR Employee Attrition, and then Add
* Click on Metadata icon



* Rename and update the Data Type, Properties and Aggregation functions accordingly

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Column Name** | **New Column Name** | **Data Type** | **Treat As** | **Aggregation** |
| Age |  | Number | Attribute | None |
| Attrition |  | Text | Attribute | None |
| BusinessTravel | Business Travel | Text | Attribute | None |
| DailyRate | Daily Rate | Number | Attribute | None |
| Department | Department | Text | Attribute | None |
| DistanceFromHome | Distance From Home | Number | Measure | Sum |
| Education | Education | Number | Attribute | None |
| EducationField | Education Field | Text | Attribute |  |
| EmployeeCount | Employee Count | Number | Measure | Sum |
| EmployeeNumber | Employee Number | Number | Attribute |  |
| EnvironmentSatisfaction | Environment Satisfaction | Number | Attribute |  |
| Gender | Gender | Text | Attribute |  |
| HourlyRate | Hourly Rate | Number | Attribute |  |
| JobInvolvement | Job Involvement | Number | Attribute |  |
| JobLevel | Job Level | Number | Attribute |  |
| JobRole | Job Role | Text | Attribute |  |
| JobSatisfaction | Job Satisfaction | Number | Attribute |  |
| MaritalStatus | Marital Status | Text | Attribute |  |
| MonthlyIncome | Monthly Income | Number | Measure | Sum |
| MonthlyRate | Monthly Rate | Number | Attribute |  |
| NumCompaniesWorked | Num Companies Worked | Number | Attribute |  |
| Over18 | Over 18 | Text | Attribute |  |
| OverTime | Over Time | Text | Attribute |  |
| PercentSalaryHike | Percent Salary Hike | Number | Attribute |  |
| PerformanceRating | Performance Rating | Number | Attribute |  |
| RelationshipSatisfaction | Relationship Satisfaction | Number | Attribute |  |
| StandardHours | Standard Hours | Number | Attribute |  |
| StockOptionLevel | Stock Option Level | Number | Attribute |  |
| TotalWorkingYears | Total Working Years | Number | Measure | Sum |
| TrainingTimesLastYear | Training Times Last Year | Number | Measure | Sum |
| WorkLifeBalance | Work Life Balance | Number | Attribute |  |
| YearsAtCompany | Years At Company | Number | Attribute |  |
| YearsInCurrentRole | Years In Current Role | Number | Attribute |  |
| YearsSinceLastPromotion | Years Since Last Promotion | Number | Attribute |  |
| YearsWithCurrManager | Years With Curr Manager | Number | Attribute |  |



* Apply Script
* Add Preparation Step 🡪 Add a new column
* Upload HR\_Turnover\_Data-Updated.xlsx for Turnover analysis
* Choose the sheet New Joiners
* Enter the name HR Employee New Joiners, then Click Add
* Add Preparation Step 🡪 Add new column DateKey by using express

(YEAR(Date of Join)\*100+MONTH(Date of Join))\*100+DAYOFMONTH(Date of Join)

* Upload HR\_Turnover\_Data-Updated.xlsx for Turnover analysis
* Choose the sheet Leavers
* Enter the name HR Employee Leavers, then Click Add
* Add Preparation Step 🡪 Add new column DateKey by using express

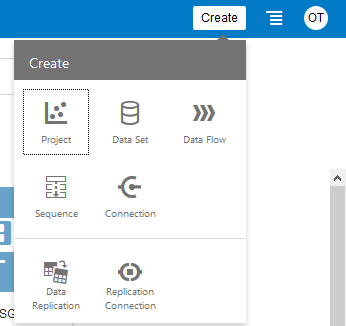
(YEAR(Date of Leaving)\*100+MONTH(Date of Leaving))\*100+DAYOFMONTH(Date of Leaving)

# Share Dataset

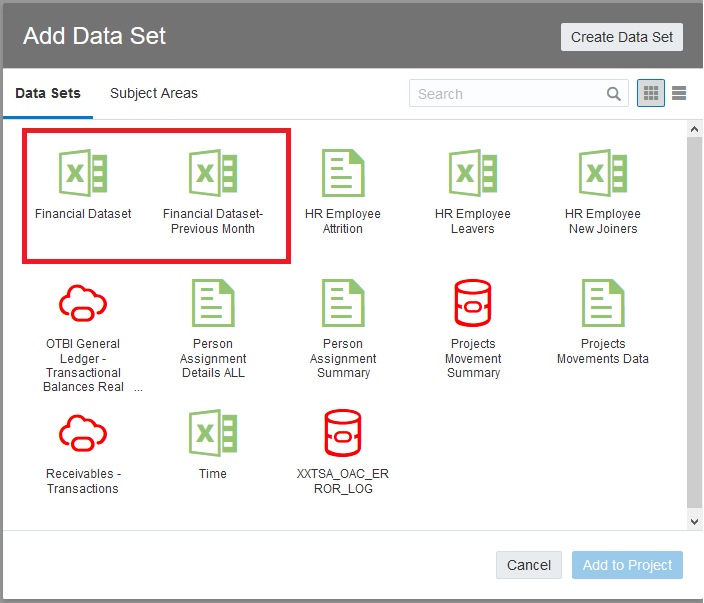
# Data Modelling

## Financial Data Model

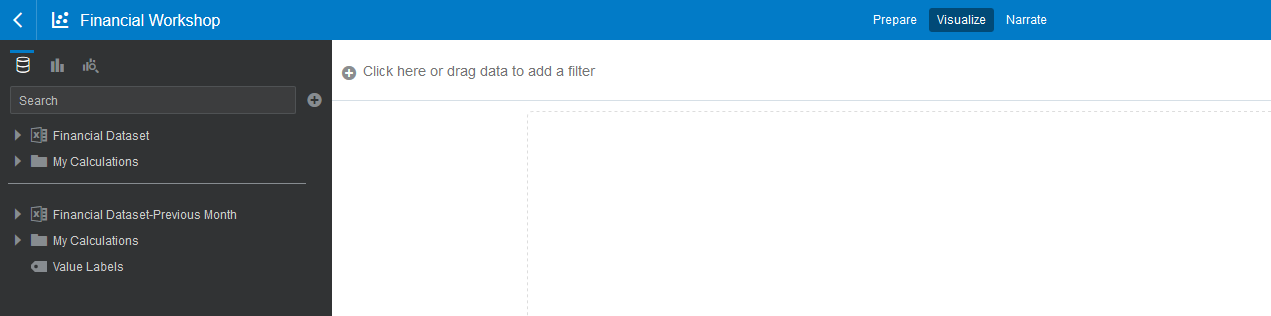
* On the right-top corner, click on Create 🡪 Project

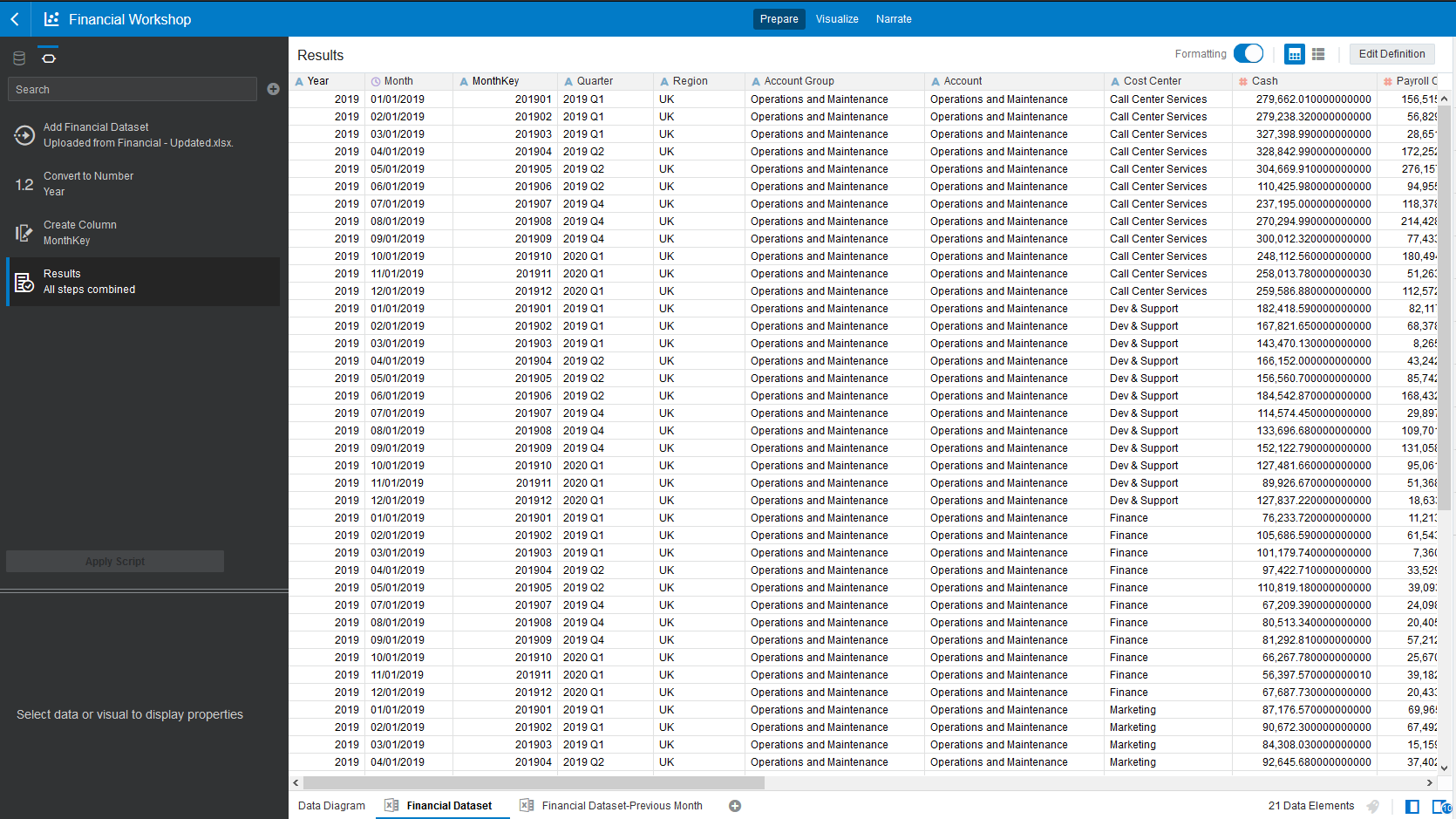


* Choose the Financial datasets: Financial Dataset and Financial Dataset – Previous Month

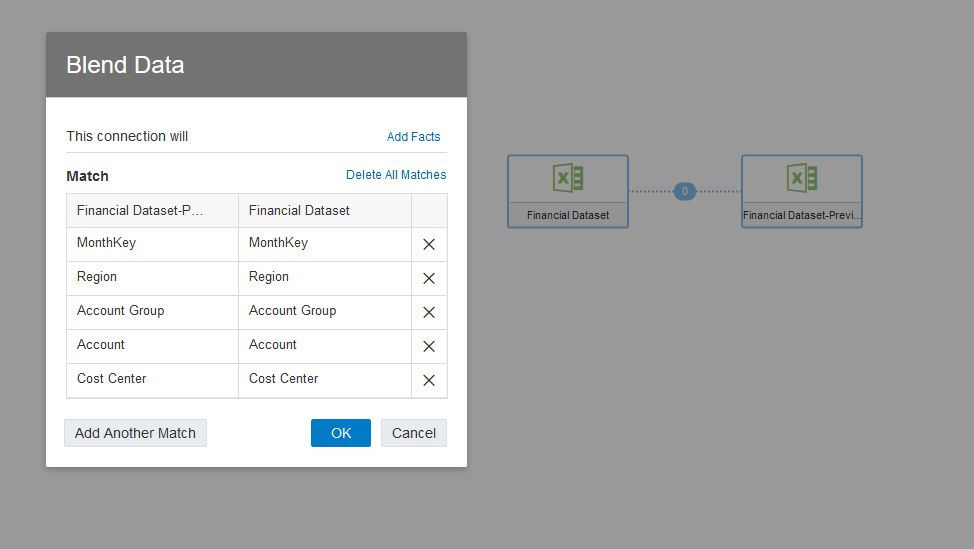


* Enter project name: **Financial Workshop**
* Click on Prepare tab

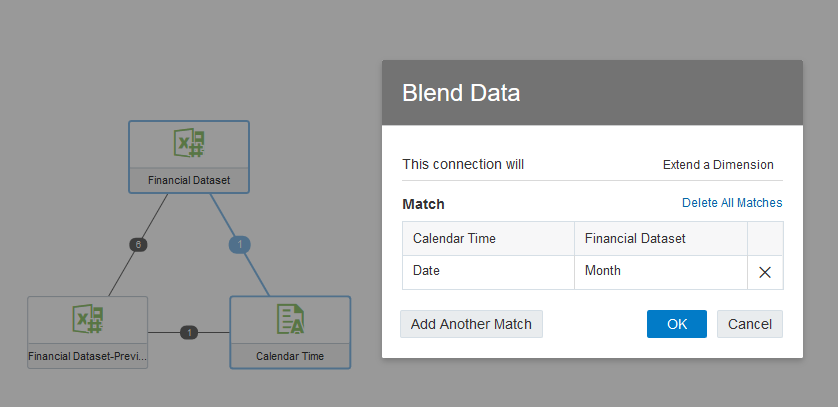




* Click on Data Diagram to build the relationship of two datasets. By default, there is no link, so we are creating the link based on MonthKey, Region, Account Group, Account and Cost Center columns
* Click on the dot line and Blend Data popup appears
* Click Add Another Match to add connection
* Click OK



* Click Add Data Set, then choose Calendar Time dataset. If Calendar Time dataset not existed, you should add new data set by uploading Calendar Time.csv file
* Click on the dot line between Financial Dataset and Calendar Time, creating the link

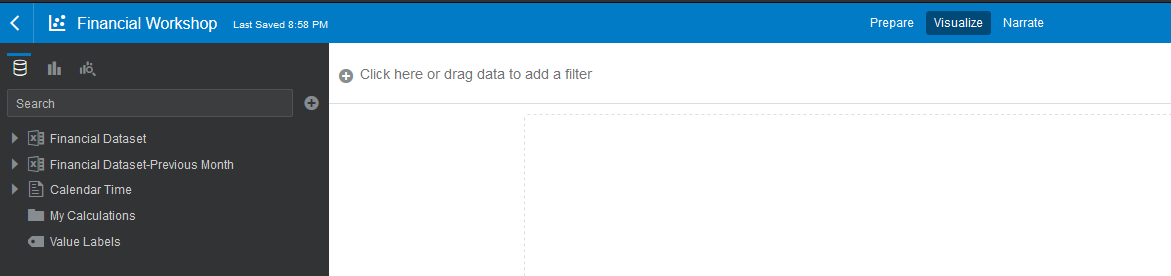


* Click Save icon to save the project

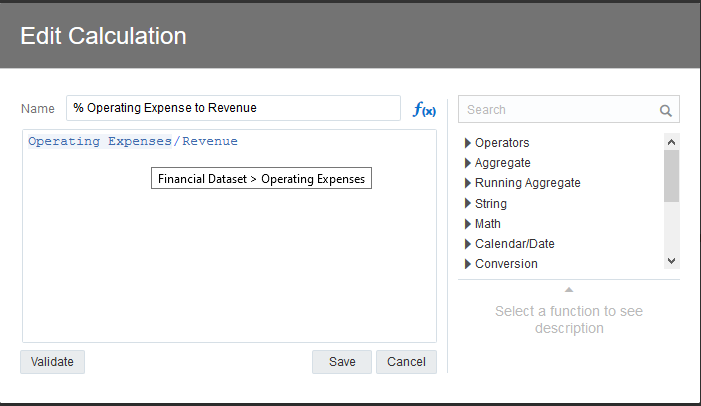
# Data Stories

## Financial Workshop

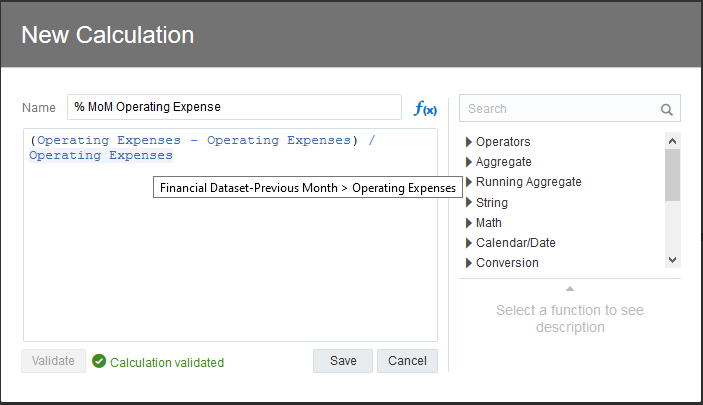
* Open the project Financial Workshop
* Click on Visualize tab



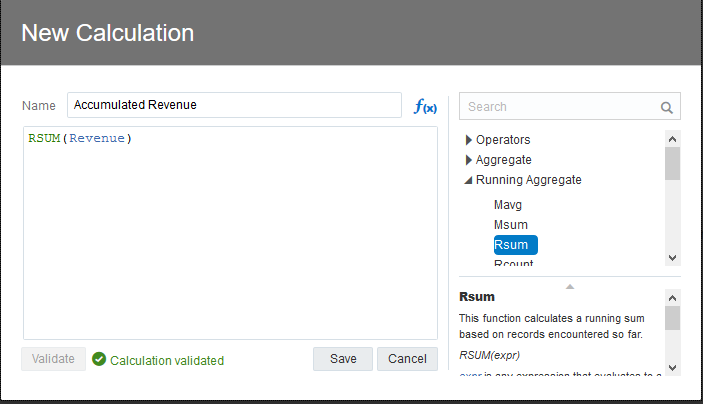
* Right-click on My Calculations 🡪 Add Calculation
  + % Operating Expense to Revenue = Operating / Revenue



* You should choose the columns from Financial Dateset
* Add new calculation: **% MoM Operating Expense**
  + **Expression:** (Operating Expense of Month (N) – Operating Expense of Month(N-1)) / Operating Expense of Month(N-1)

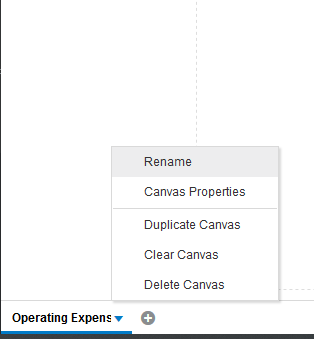


* Add new calculation: **% MoM Receivables**
  + **Expression:** (Receivables of Month (N) – Receivables of Month(N-1)) / Receivables of Month(N-1)
* Add new calculation: **% MoM Payables**
  + Expression: (Payables of Month (N) – Payables of Month(N-1)) / Payables of Month(N-1)
* Add new calculation: % MoM Gross Revenue
  + Expression: (Revenue of Month (N) – Revenue of Month(N-1)) / Revenue of Month(N-1)
* Add new calculation: Accumulated Gross Revenue
  + Expression: Running SUM of all revenue base on Time

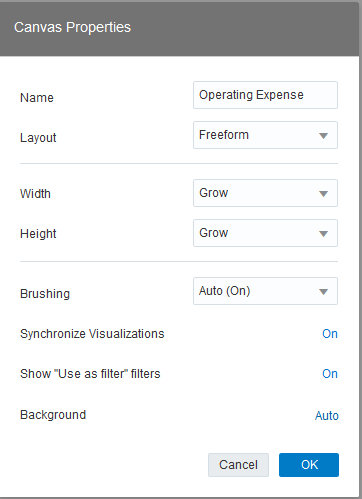


### Operating Expense canvas

* Rename Canvas 1 to Operating Expense
* Click on the triangle icon of canvas 🡪 Canvas Properties



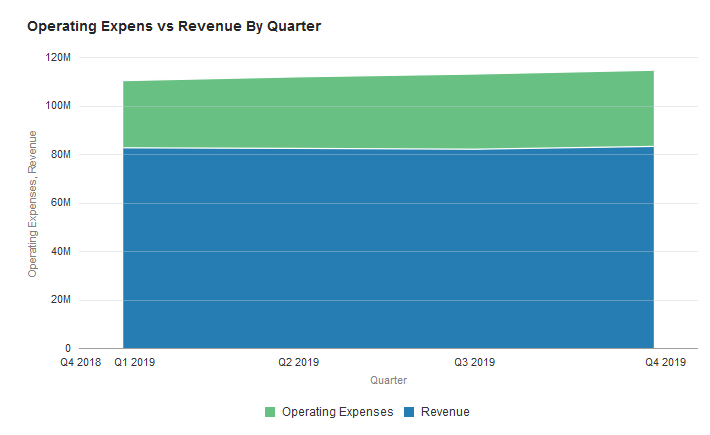
* Choose Free form in Layout dropdown list



* Click OK

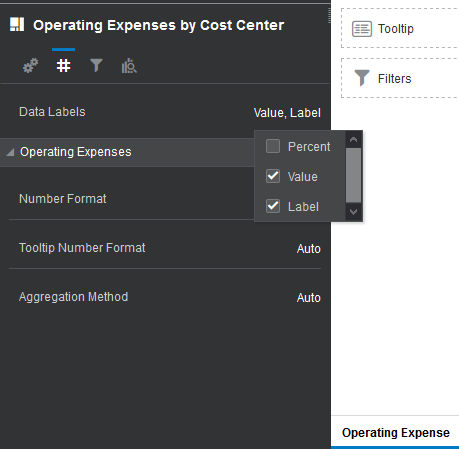
#### Operating Expense vs Revenue By Quarter

* Drag and drop Operating Expense, Revenue and Quarter from dataset into visualization area
* Choose chart type: **Stacked Area chart**
* Values (Y Axis) should include Operating Expense and Revenue measures
* Category (X Axis) should include Quarter
* Change title of chart: **Operating Expense vs Revenue By Quarter**

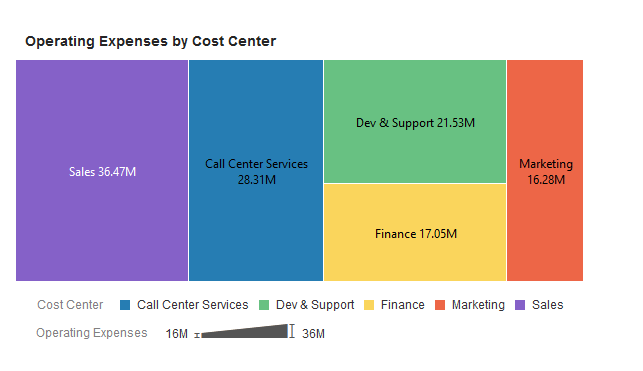


#### Operating Expense By Cost Center

* Drag and drop Operating Expense and Cost Center from dataset into visualization area
* Choose chart type: **Treemap chart**
* Values (box size) : Operating Expense
* Color : Cost Center
* Change title of chart: **Operating Expenses By Cost Center**
* Legend: Bottom
* Data Labels should display Label and Value

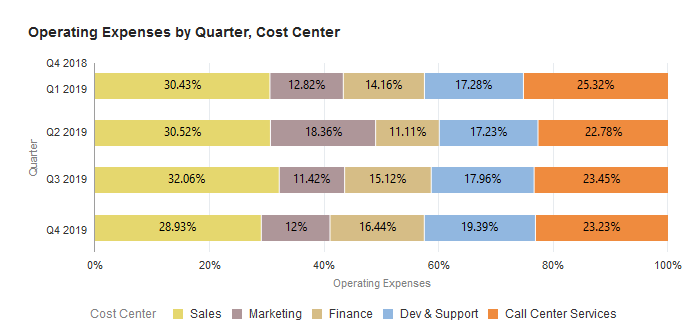


* Number Format: Number
  + - Abbreviate = On



#### Operating Expense By Cost Center vs Quarter

* Drag and drop Operating Expense, Cost Center Quarter from dataset into visualization area
* Choose chart type: **Horizontal 100% Stacked Bar chart**
* Values (Y-Axis) : Operating Expense
* Category (X-Axis): Quarter
* Color : Cost Center
* Change title of chart: **Operating Expenses By Cost Center vs Quarter**
* Legend: Top



#### Operating Expense Change By Month

* Drag and drop Operating Expense and Month from dataset into visualization area
* Choose chart type: **Waterfall chart**
* Values (Y-Axis) : Operating Expense
* Category (X-Axis): Quarter
* Color : Cost Center
* Change title of chart: **Operating Expenses By Cost Center vs Quarter**

Legend: Top