



# Predicting Win Scenarios in PUBG with IBM Watson Studio

Hands-On Workshop | Digital Summit '18

## Miracle Innovation Labs

Miracle Software Systems, Inc.

## Predicting Win Scenarios in PUBG with IBM Watson Studio

### Introduction

This document contains a step-by-step process for creating a Machine Learning model in IBM Watson Studio and will teach you the end to end process for ML model building to predict the winning probability of PUBG users.

This guide was prepared by [Miracle's Innovation Labs](#).

### Pre-Requisites

All attendees must have their workstation (with Internet) to participate in the lab (Both PC and MAC are compatible). The following pre-requisites will help you to make the Hands-on Lab experience easier.

- Active email ID for registering with IBM Cloud to get access for IBM Watson Studio

### Technology Involved

- IBM Cloud
- IBM Watson Studio

## Lab Steps

So, let us get started with the application!

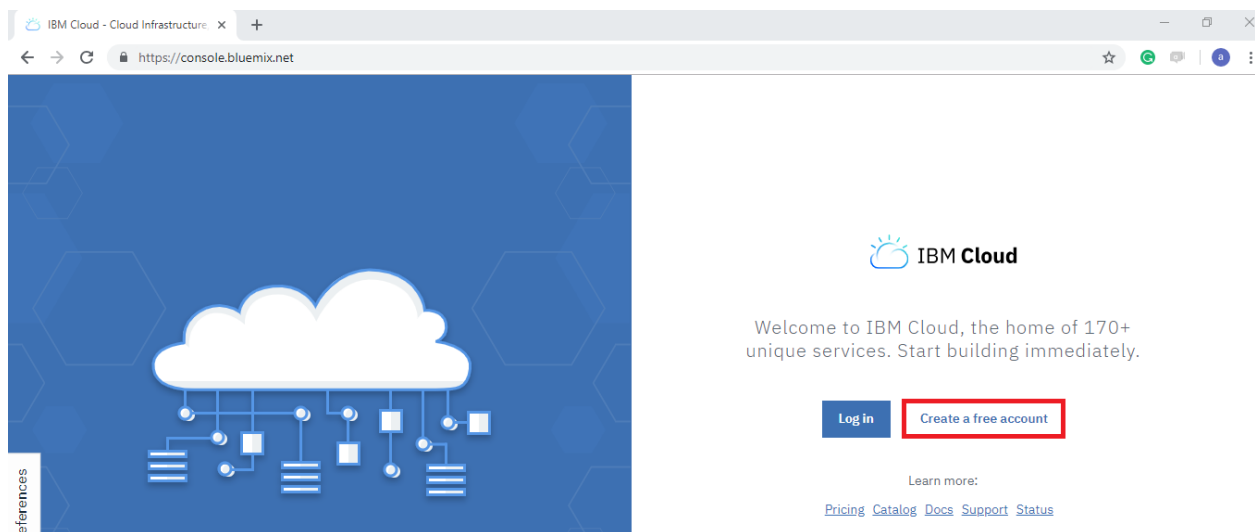
The following steps will outline how you can create a Machine Learning model using Watson Studio. This application helps in predicting the Winning percentage of PUBG users based on the historical data of players.

### Step #1 | Creating Watson Studio in IBM Cloud


The first step is to make sure that we have access to the IBM Cloud with either a free trial option (or) a paid subscription.

Login to IBM Cloud at <http://bluemix.net> (or) register today at, <https://console.bluemix.net/registration>

If you are a new user, click on **Create a free account**.



Now, you need to provide the details to the fields that are marked as required and click on **Create Account**.



Sign up for an IBMid and create your IBM Cloud account

## Build on IBM Cloud for free with no time restrictions

**Guaranteed free development with Lite plans**  
Develop worry-free and at no cost with cap based Lite plan services for as long as you like.

**Start on your projects right away**  
Skip entering your credit card info and get working in just a few short steps.

**Get a \$200 credit when you upgrade**  
After you upgrade to a Pay-As-You-Go account, you can use the credit to try new services or scale your projects. The credit is valid for 1 month and can be used with any of our IBM Cloud offerings.

**Ready to get started? Sign up today!**

Email\*

First Name\*

Last Name\*

Country or Region\*

India

Password\*

IBM may use my contact data to keep me informed of products, services and offerings:

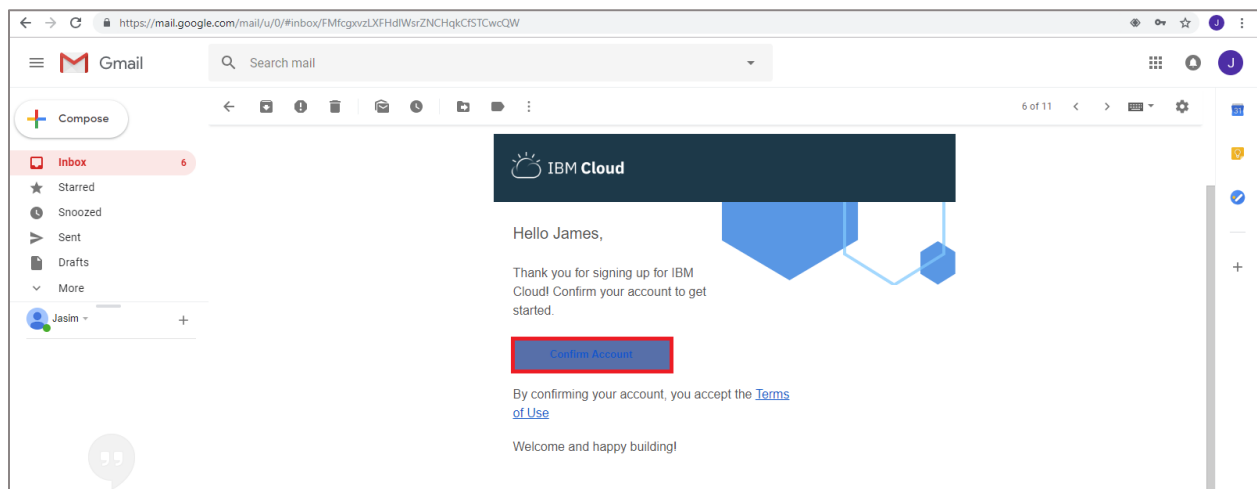
☒ by email

You can withdraw your marketing consent at any time by sending an email to [netapp@us.ibm.com](mailto:netapp@us.ibm.com). Also you may unsubscribe from receiving marketing emails by clicking the unsubscribe link in each such email. More information on our processing can be found in the [IBM Privacy Statement](#).

By submitting this form, I acknowledge that I have read and understand the IBM Privacy Statement. I accept the product [Terms and Conditions](#) of this form.

Create Account

After creating account, confirmation mail will be sent to the registered mail id. Click on **Confirm Account** and then login to your Bluemix Account.



Now, login to your Bluemix Account.

## Log in to IBM

Don't have an IBMid?  
[Create an account](#)

[Log in with your company credentials \(SSO\)](#)

Need help? [Contact the IBM Help Desk](#)

IBMid

[Forgot IBMid?](#)

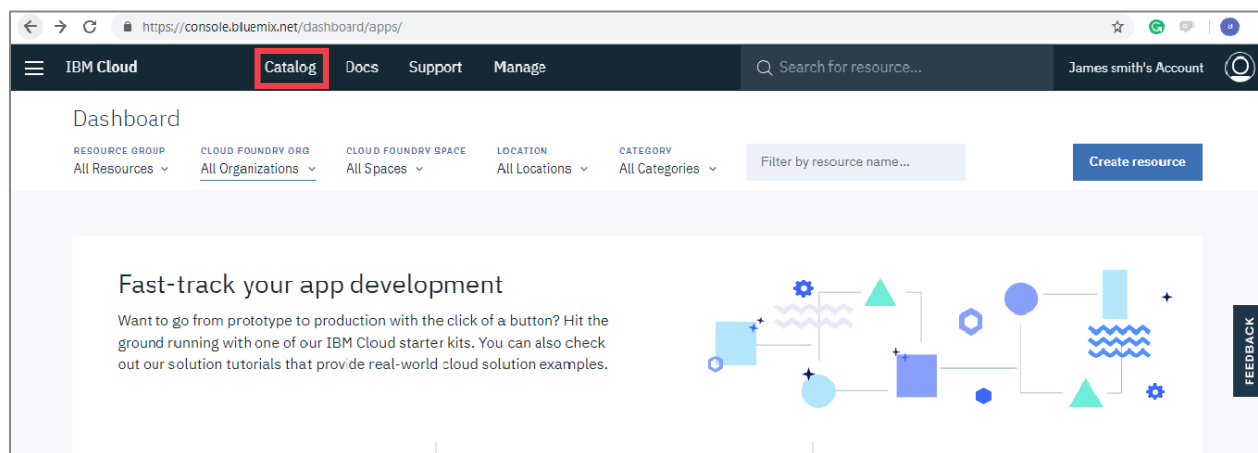
Password

[Forgot password?](#)

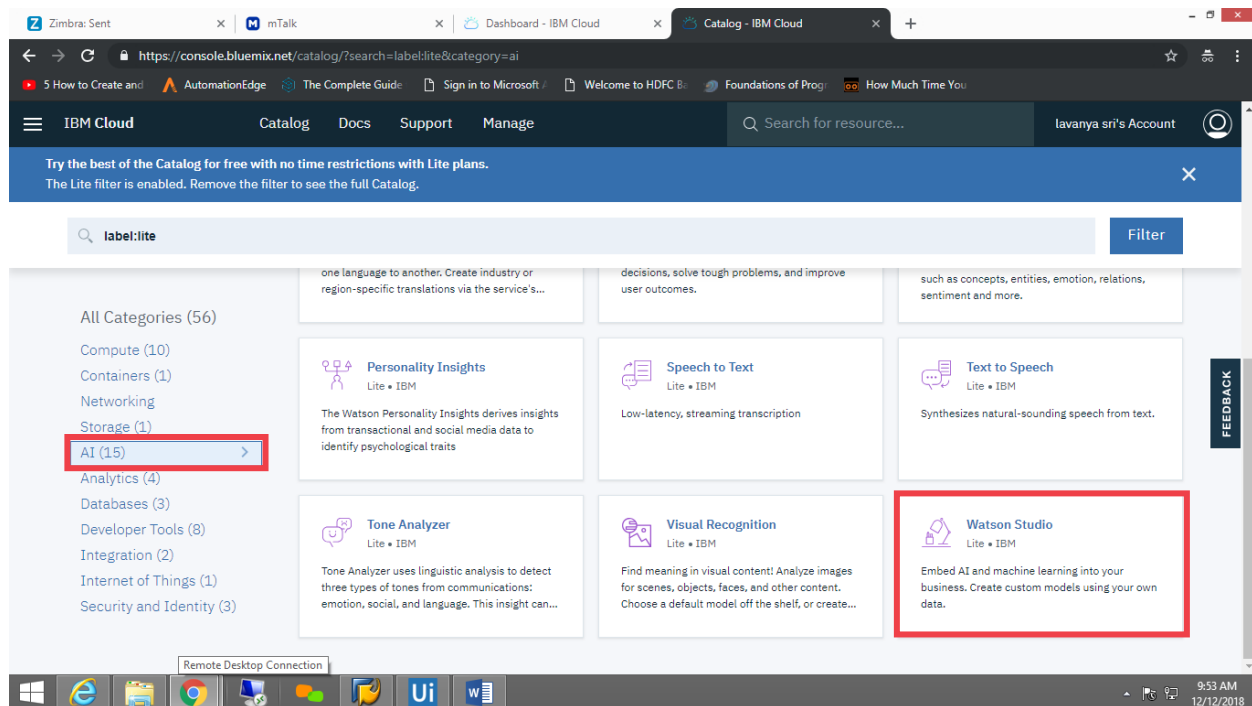
☐ Remember me [?](#)

Log in

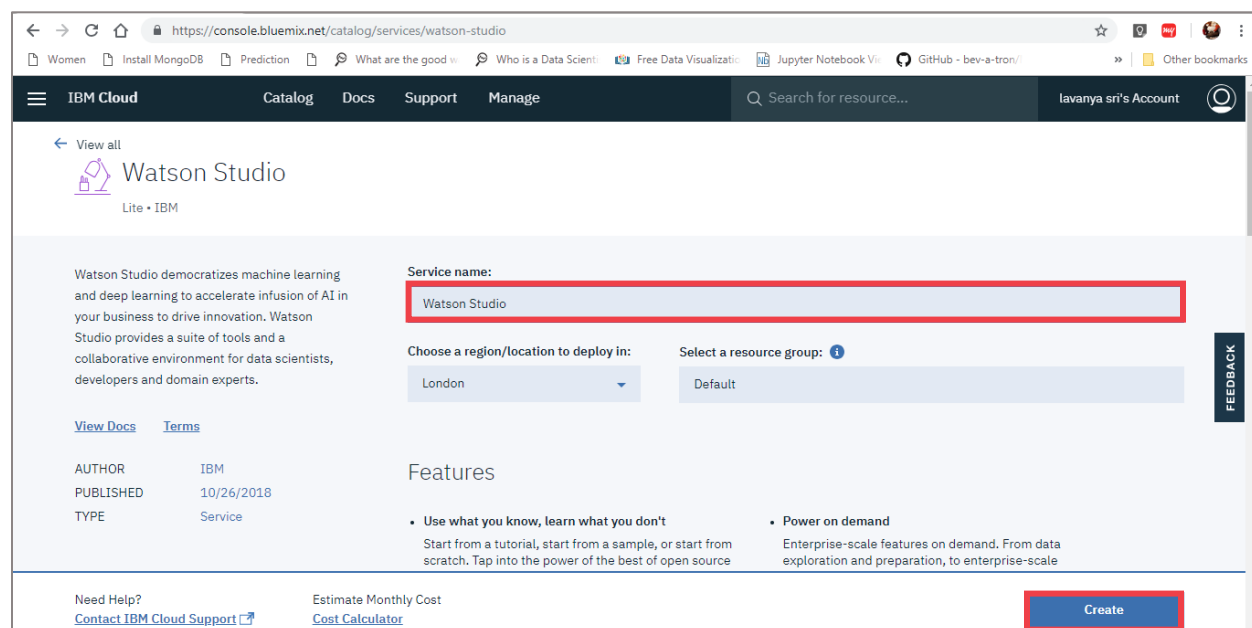
You can see the dashboard as shown below. Now, click on **Catalog** option.



In Catalog section, under the **AI category** you will be able to see all the Watson Services. Click on **Watson Studio** tile for building the application.



Provide a unique name for the service instance in the **Service name** field. For example, type **Watson Studio**. Leave the default values for the other options and click on **Create**.

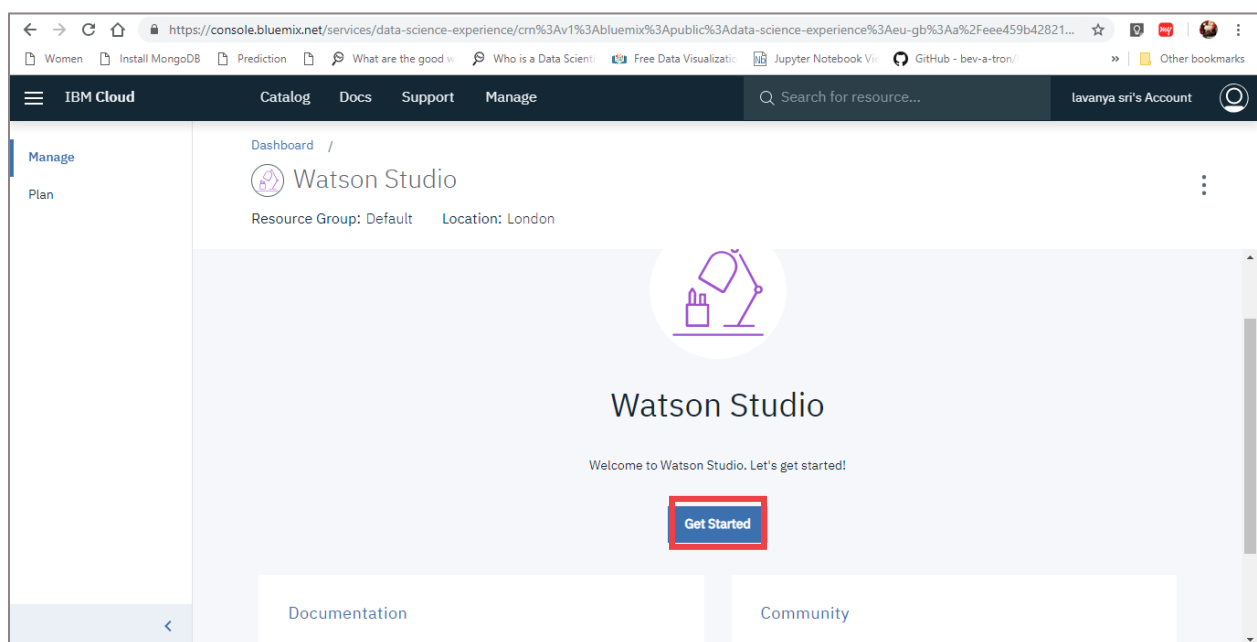


**Note** - Creation of service may take up to a minute or two.

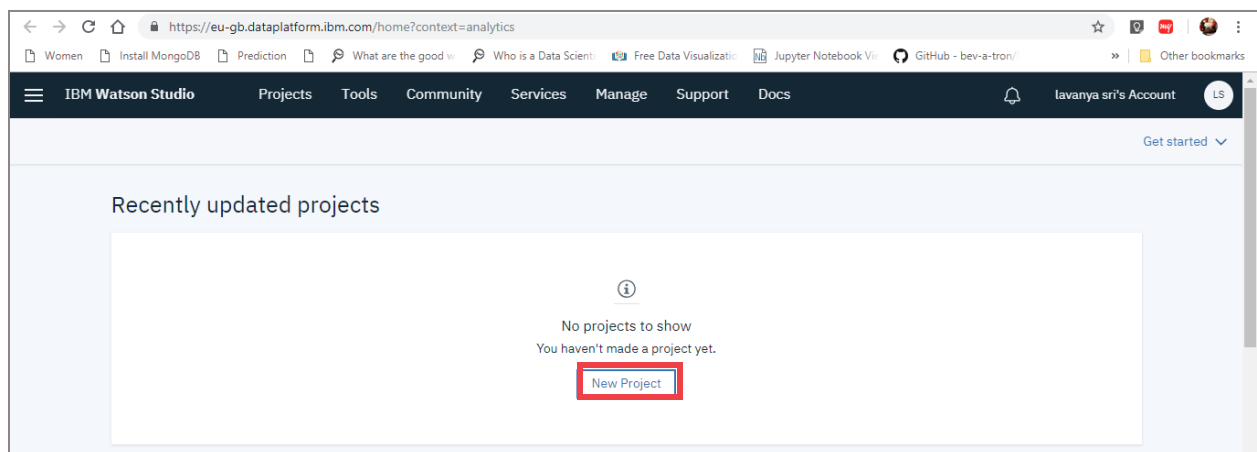
Once the service is created successfully we will see a screen as shown below. Click on **Get Started** button to work with the Watson Studio.

## Step #2 | Create a Project in Watson Studio

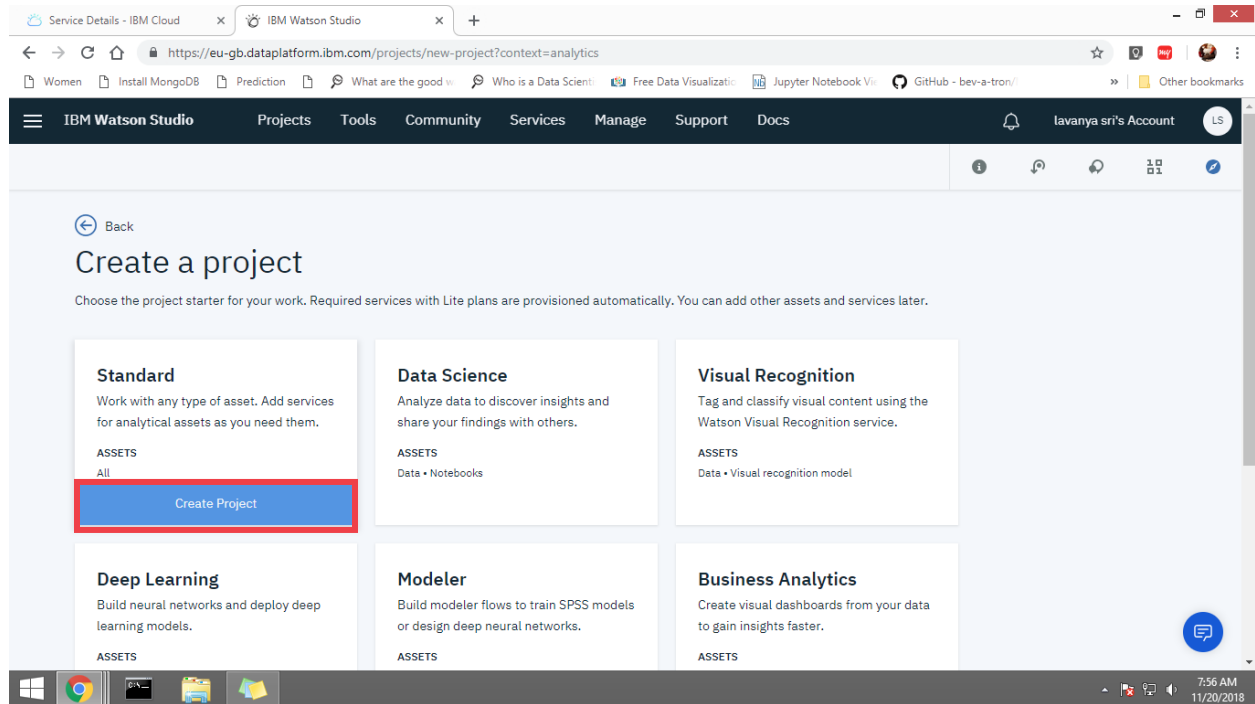
Now we have landed into **IBM Watson Studio** platform where we can create our models using different tools and services.



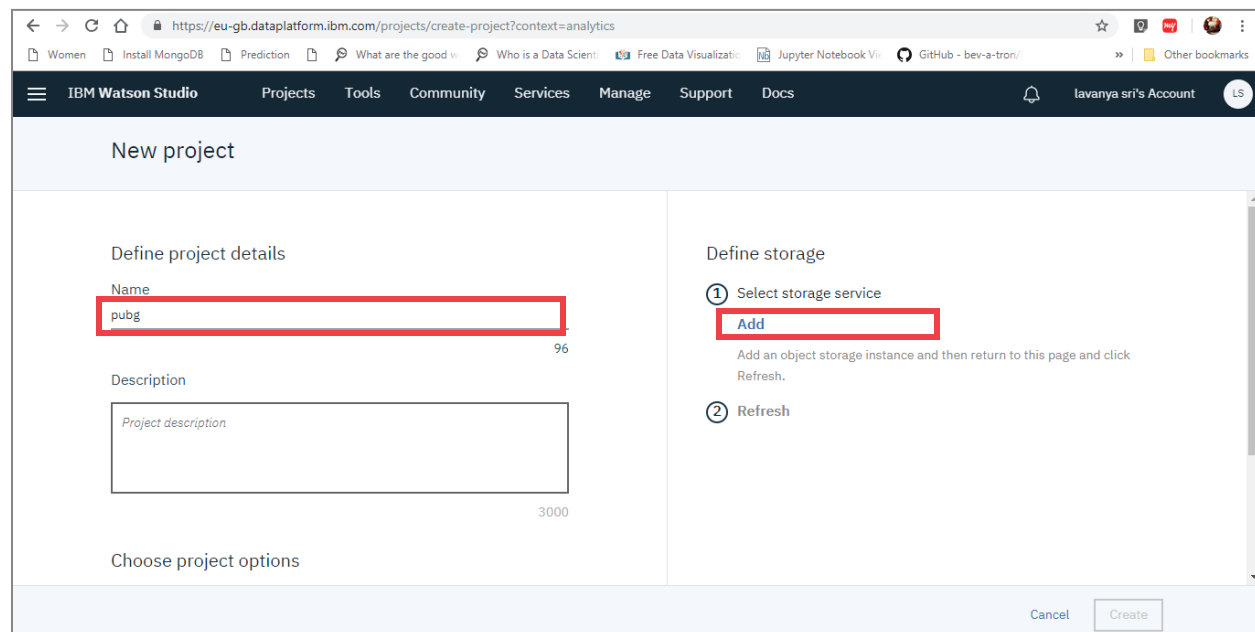
Initially, we have to create a project. Click on the **New Project** button.



Once we click on New Project we can see number of options such as Standard, Data Science, etc. Please select Standard and click on **Create Project**.

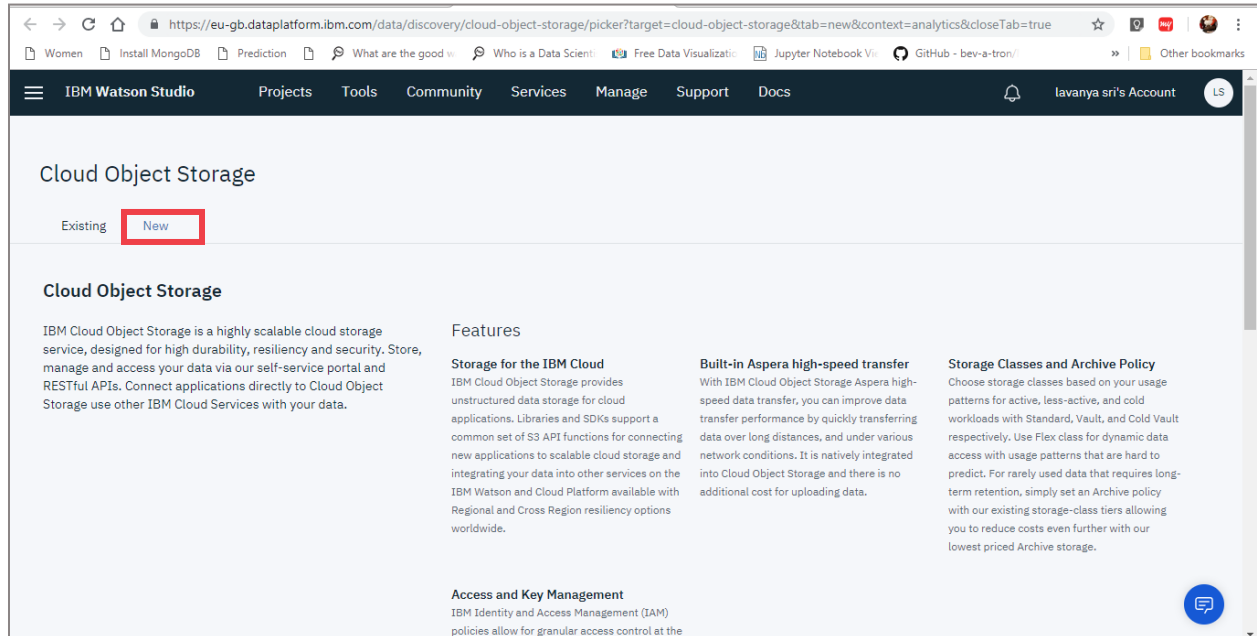


Provide a **name** to the project. Here we need to add Object storage instance of IBM Cloud to our project. For that click on **Add**.

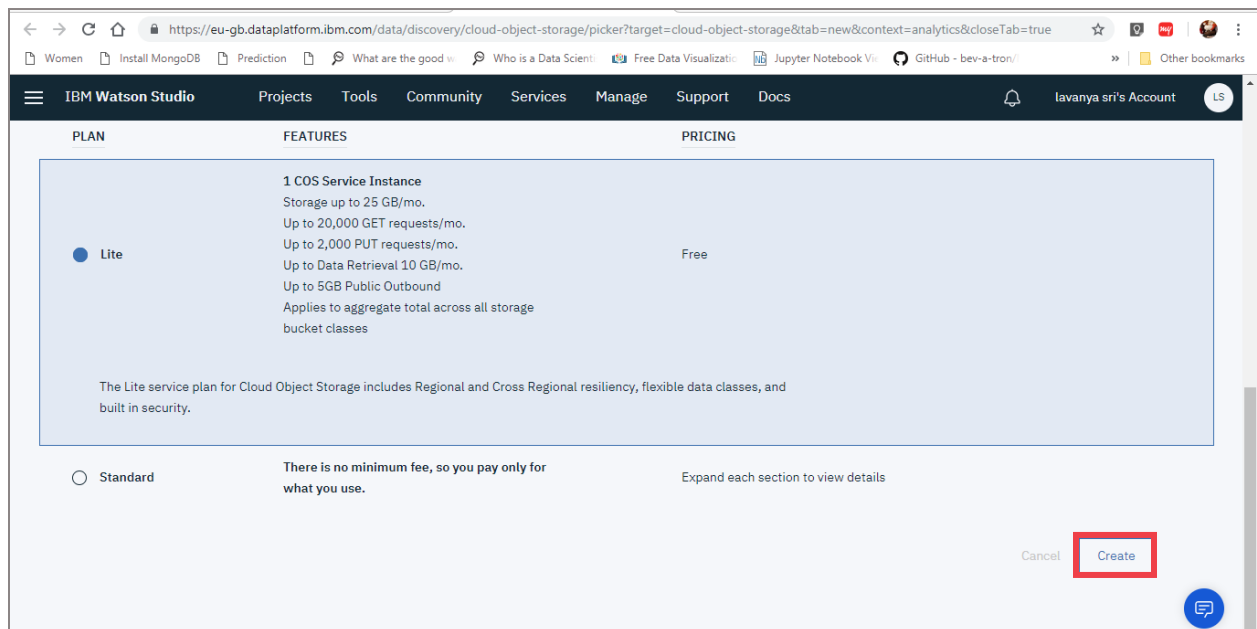




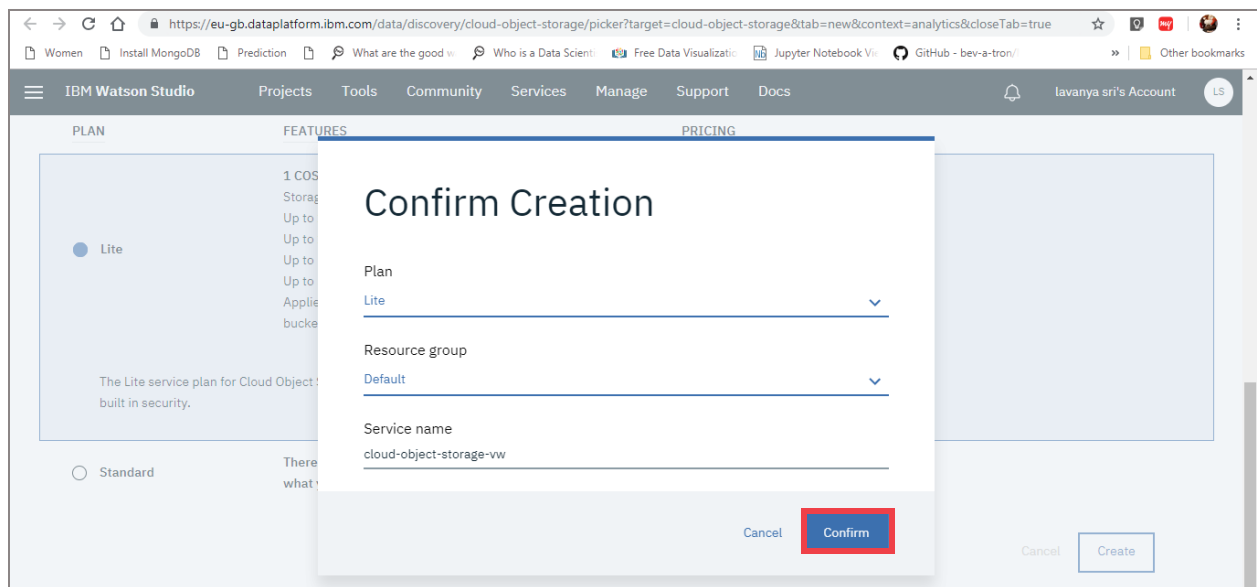
Here, we are creating a **New** Cloud Object Storage instance to our project.



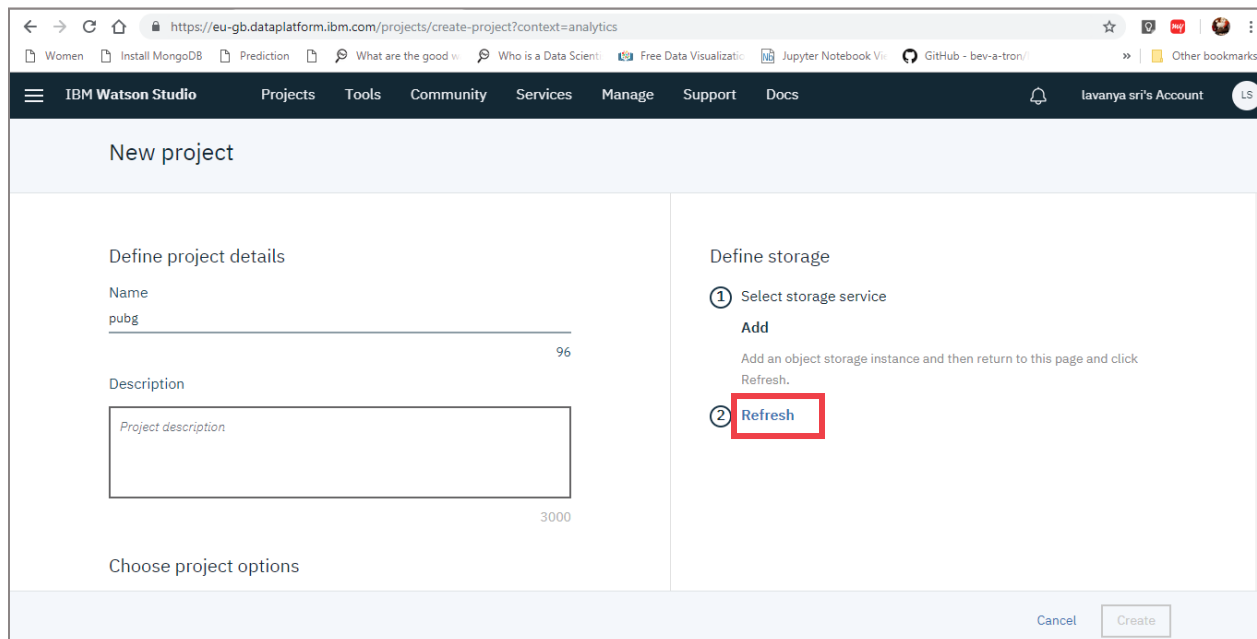
Choose the **Lite** plan option and click on **Create**.



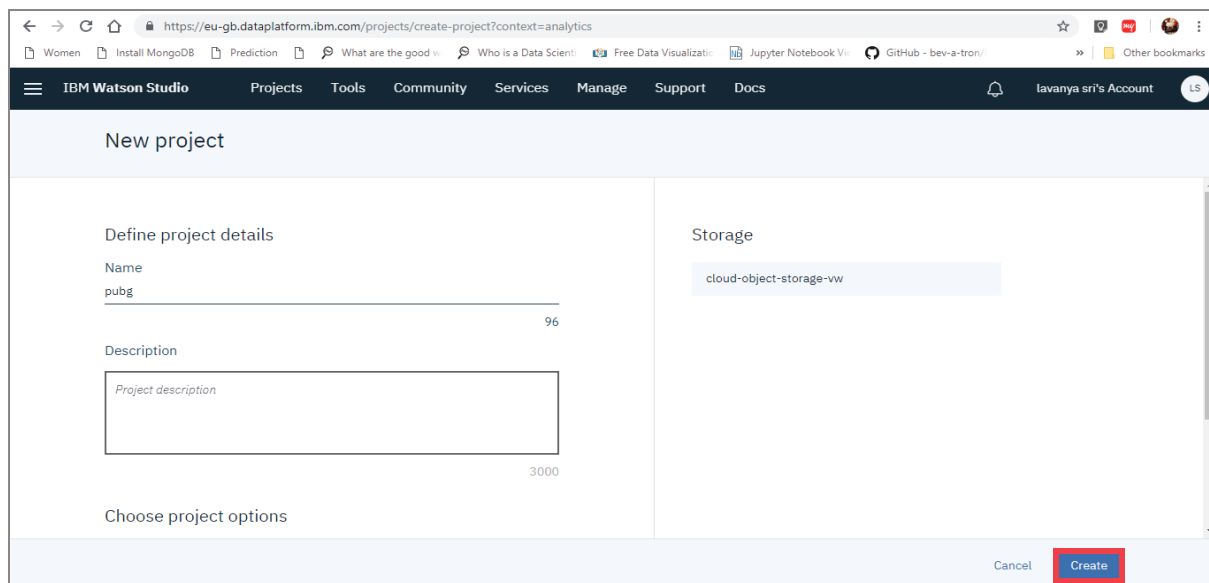
Click on **Confirm** to create the instance.



Click on **Refresh**, so we can see the storage instance as shown below.

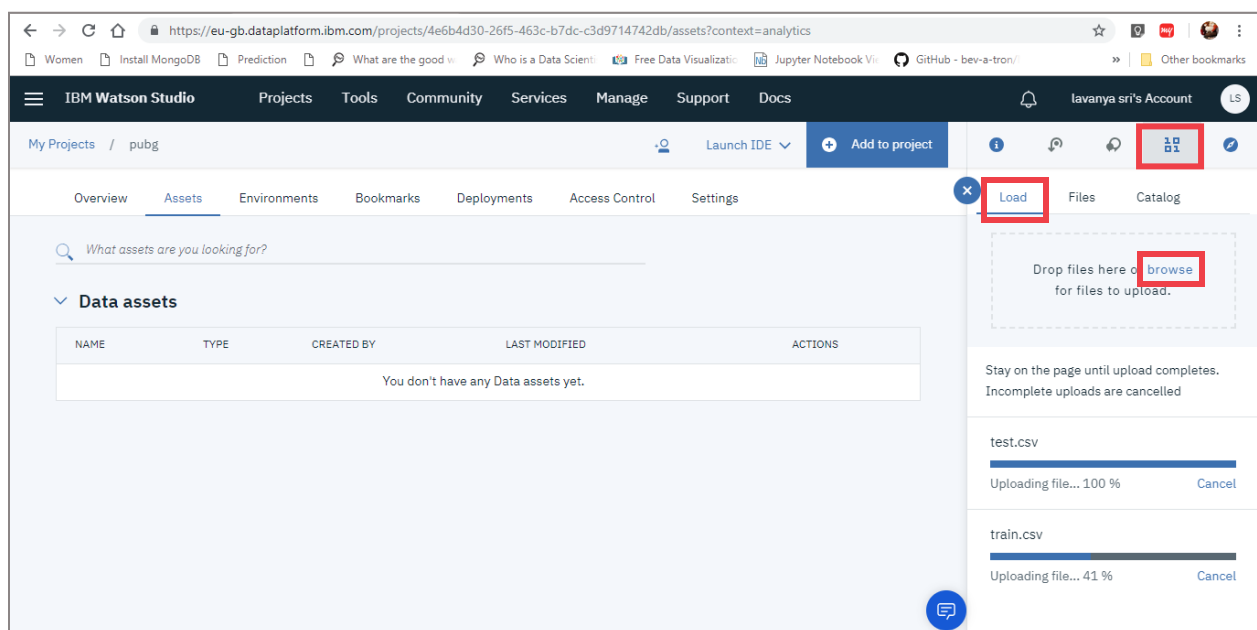


Click on **Create**.

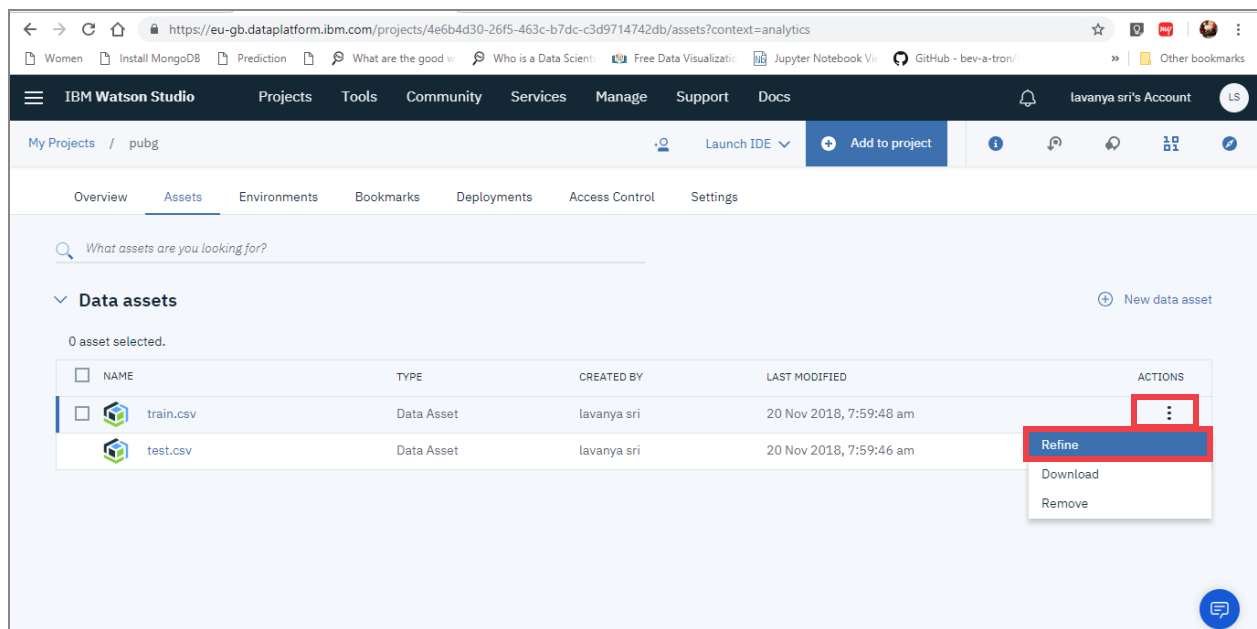


## Step #3 | Importing Datasets and Performing Data Refinement

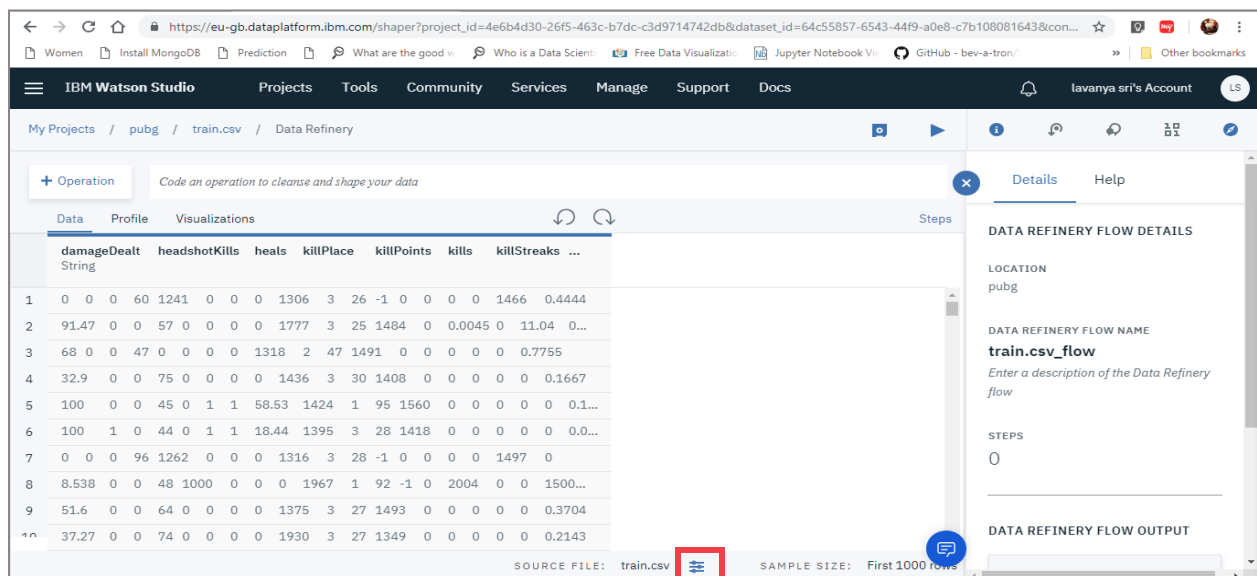
Import the PUBG datasets into our project, click on **browse** and get the datasets. To build and test a model we need both train and test datasets.



Once the datasets are loaded into data assets as shown, we need to refine them. So, click on the 3 vertical dots under Actions and select Refine for the **train.csv**.



Click on the symbol which is highlighted below, to change the format of dataset.



From the dropdown menu of Field delimiter select **tab (\t)** option and click on **Apply**.

The screenshot shows the IBM Watson Studio interface. On the left, the 'Format options' panel is open. It has a checkbox 'First row contains column headers' which is checked. Under 'Encoding', 'UTF-8' is selected. Under 'Field delimiter', 'Tab (\t)' is selected and highlighted with a red box. Under 'Quote character', 'None' is selected. Under 'Escape character', 'None' is selected. At the bottom of this panel, the 'Apply' button is highlighted with a red box. On the right, the 'Raw data' table is visible, showing 1000 rows. The columns are: damageDealt, headshotKills, heals, killPlace, killPoints, kills, killStreaks, and longestKill. The first row of data is: 558.6, 0, 4, 11, 1265, 3, 2, 237.5.

The data is formatted now, click on the symbol highlighted below to save the changes.

The screenshot shows the IBM Watson Studio interface with the 'Data Refinery' tab selected. A red box highlights the 'Save and Run flow' button, which is a play icon. Below the button, the 'Data' tab is active, showing a table with 1000 rows. The columns are: damageDealt String, headshotKills String, heals String, killPlace String, killPoints String, kills String, killStreaks String, and longestKill String. The first row of data is: 0, 0, 0, 60, 1241, 0, 0, 0.

Click on **Save and Run flow**. Now, we will be having a new dataset with name **train.csv\_shaped.csv** which is formatted as shown below.

The screenshot shows the IBM Watson Studio Data Refinery interface. The left sidebar contains 'DATA REFINERY FLOW DETAILS' with fields for 'LOCATION' (pubg), 'DATA REFINERY FLOW NAME' (train.csv\_flow), and 'STEPS' (0). The main area shows 'DATA REFINERY FLOW OUTPUT' with a preview of the 'train.csv\_shaped.csv' dataset. A 'Save and Run flow' button is highlighted in red at the bottom right.

Click on **Continue Working**. Then move back to the main page of project.

The screenshot shows the IBM Watson Studio Data Refinery interface with a 'What's next?' dialog box. The dialog box contains the text: 'Your Data Refinery flow is currently running. You can view its progress on the Summary and Runs page. When the flow completes, you can view its output from there too.' The 'Continue Working' button is highlighted in red. In the background, a data table is visible with columns for 'Data', 'Profile', and 'Visualizations'.

Data	Profile	Visualizations
damageDealt String	headshotKills String	
1 0	0	
2 91.47	0	
3 68	0	
4 32.9	0	
5 100	0	
6 100	1	
7 0	0	
8 8.538	0	
9 51.6	0	
10 37.27	0	

Here, we can see **train.csv\_shaped.csv** file in the Data assets.

The screenshot shows the IBM Watson Studio interface. The top navigation bar includes 'IBM Watson Studio', 'Projects', 'Tools', 'Community', 'Services', 'Manage', 'Support', and 'Docs'. The user is logged in as 'lavanya sri's Account'. The main content area is titled 'My Projects / pubg' and shows a search bar with the text 'What assets are you looking for?'. Below the search bar, there are two sections: 'Data assets' and 'Data Refinery flows'. The 'Data assets' section shows a table with 4 assets: 'test.csv\_shaped.csv', 'train.csv\_shaped.csv', 'train.csv', and 'test.csv'. The 'Data Refinery flows' section shows a table with 1 flow: 'test.csv\_flow'.

NAME	TYPE	CREATED BY	LAST MODIFIED	ACTIONS
test.csv_shaped.csv	Data Asset	lavanya sri	20 Nov 2018, 8:04:31 am	⋮
train.csv_shaped.csv	Data Asset	lavanya sri	20 Nov 2018, 8:03:40 am	⋮
train.csv	Data Asset	lavanya sri	20 Nov 2018, 7:59:50 am	⋮
test.csv	Data Asset	lavanya sri	20 Nov 2018, 7:59:47 am	⋮

NAME	TYPE	CREATED BY	LAST MODIFIED	ACTIONS
test.csv_flow	Data Refinery flow	lavanya sri	20 Nov 2018, 8:04:17 am	⋮

So, now click on the 3 vertical dots under Actions and repeat the same process to refine the test.csv until we generate a **test.csv\_shaped.csv** as shown.

Once this is done, click on **Add to Project**.

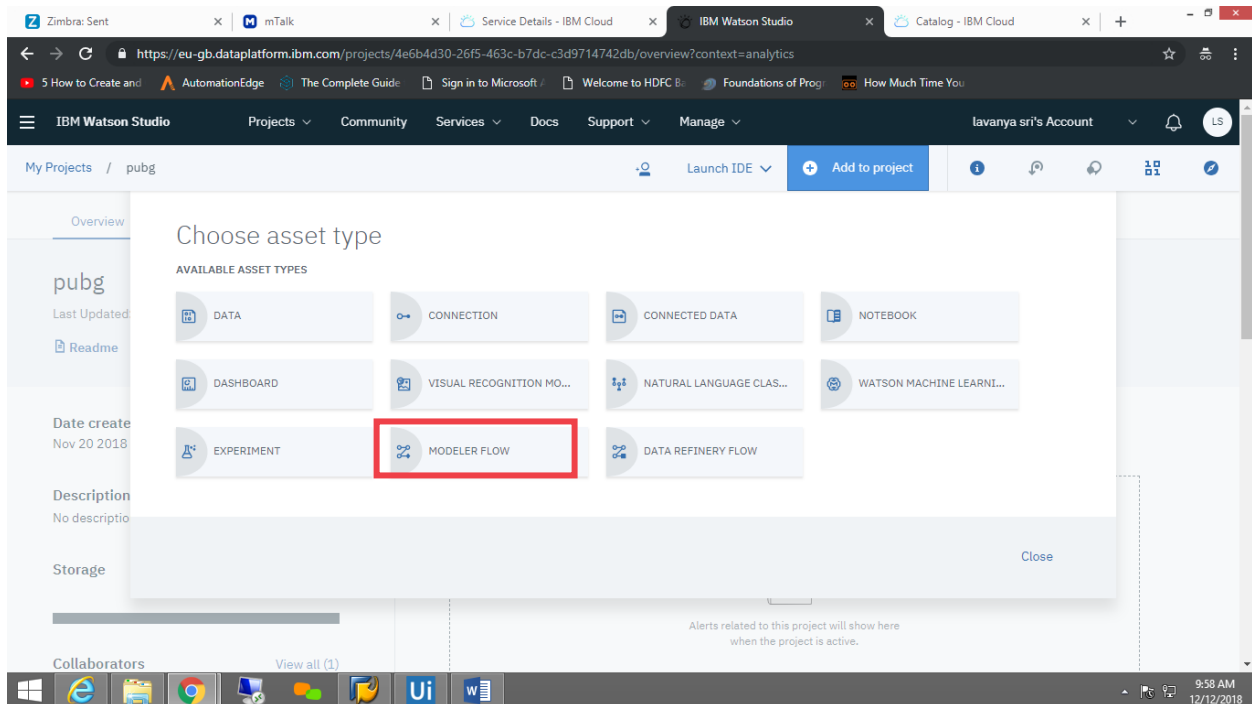
This screenshot is similar to the previous one, but with red annotations. A red box highlights the 'Add to project' button in the top navigation bar. Another red box highlights the 'ACTIONS' column header in the 'Data assets' table, and a third red box highlights the three vertical dots (⋮) in the 'ACTIONS' column for the first row, 'test.csv\_shaped.csv'.

NAME	TYPE	CREATED BY	LAST MODIFIED	ACTIONS
test.csv_shaped.csv	Data Asset	lavanya sri	20 Nov 2018, 8:04:31 am	⋮
train.csv_shaped.csv	Data Asset	lavanya sri	20 Nov 2018, 8:03:40 am	⋮
train.csv	Data Asset	lavanya sri	20 Nov 2018, 7:59:50 am	⋮
test.csv	Data Asset	lavanya sri	20 Nov 2018, 7:59:47 am	⋮

NAME	TYPE	CREATED BY	LAST MODIFIED	ACTIONS
test.csv_flow	Data Refinery flow	lavanya sri	20 Nov 2018, 8:04:17 am	⋮

## Step #4 | Creating MODELER FLOW and Preparing the Data for Model Building

Select **MODELER FLOW** option here, in order to create a Machine Learning model.



Provide a **name** to the modeler flow for example: PUBG which we are going to build and click on **create**.



https://eu-gb.dataplatform.ibm.com/canvas/new-canvas?projectGuid=4e6b4d30-26f5-463c-b7dc-c3d9714742db&context=analytics

IBM Watson Studio Projects Tools Community Services Manage Support Docs

Name\*  
PUBG 46

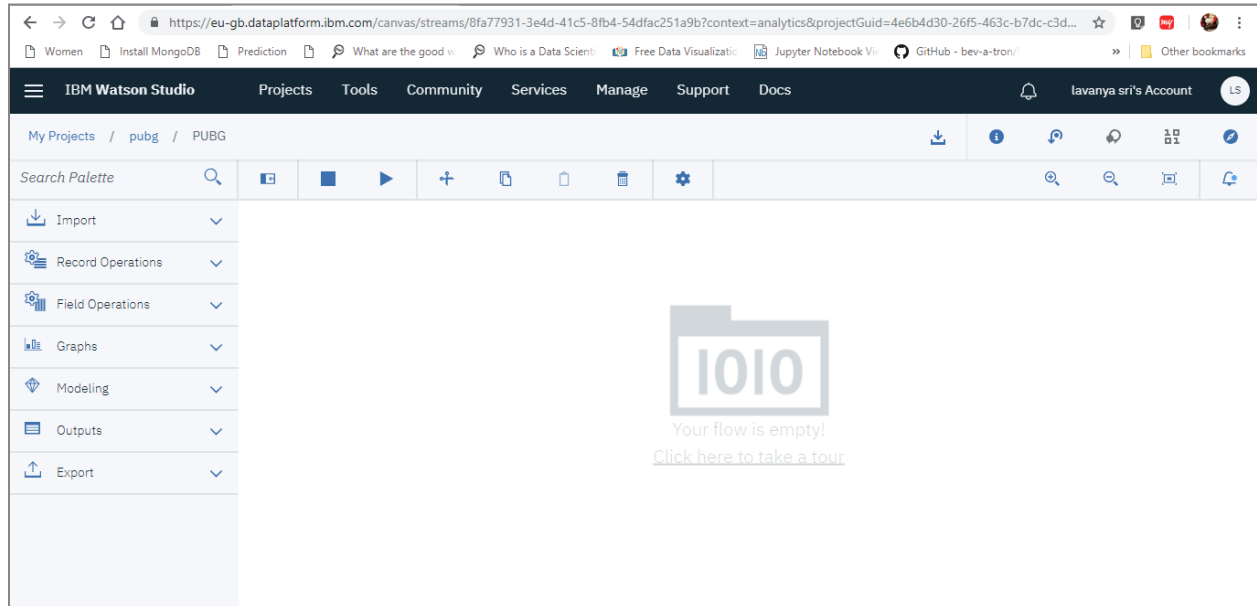
Description  
Type description here. 500

Select flow type  
☒ Modeler Flow ☐ Neural Network Modeler BETA

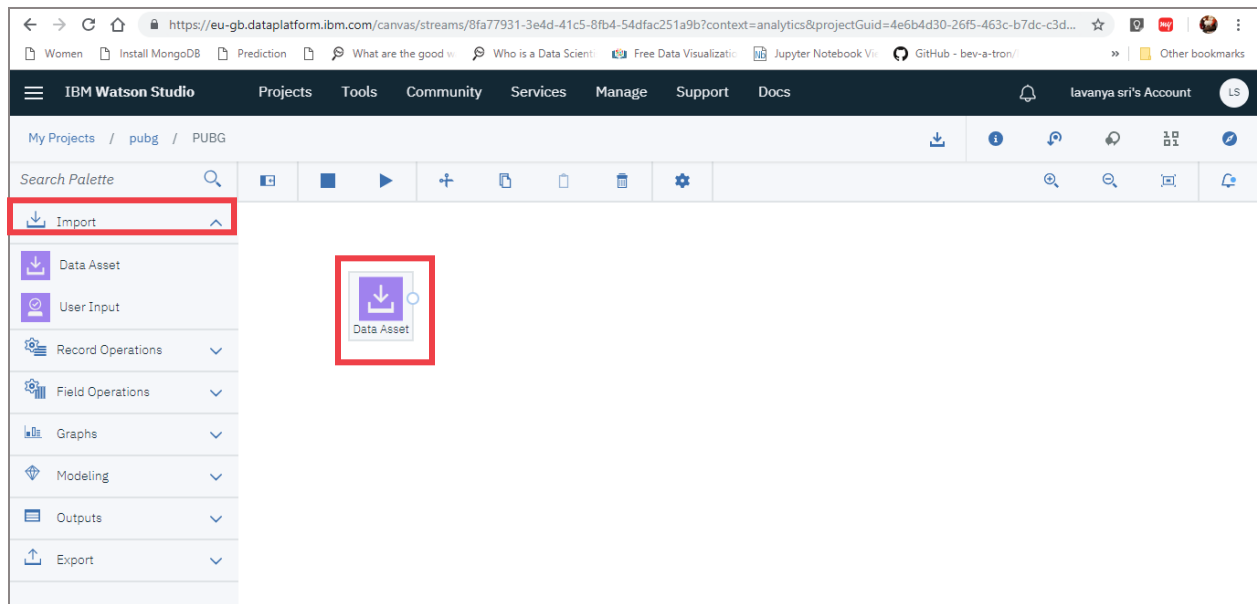
Runtime  
☒ IBM SPSS Modeler ☐ Scala Spark BETA

Cancel Create

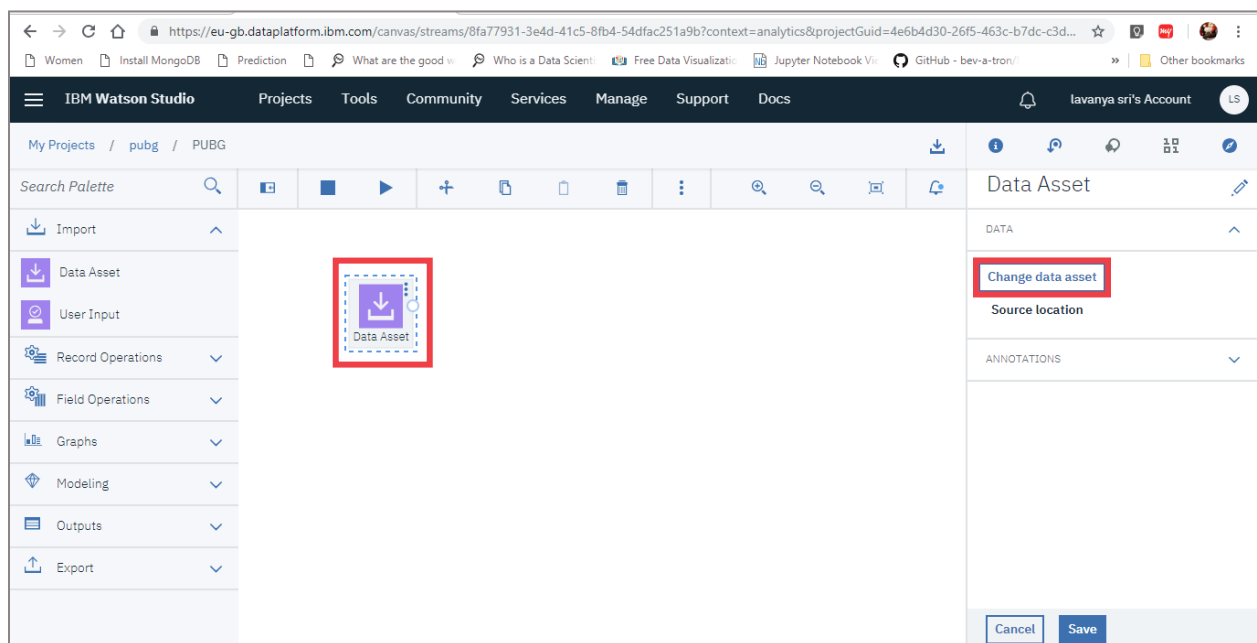
This is the **IBM SPSS MODELER** workplace where we will build the ML model.



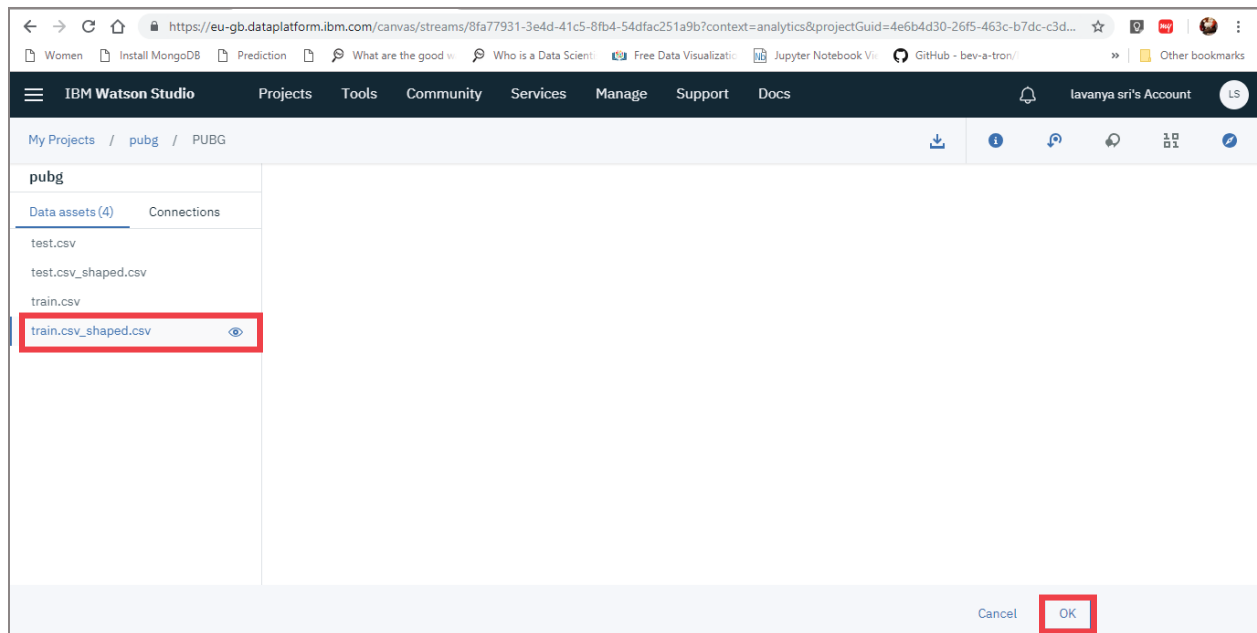
From the **Import palette** (on left side) select **Data Asset** node and just Drag and Drop it into the white space (Canvas).



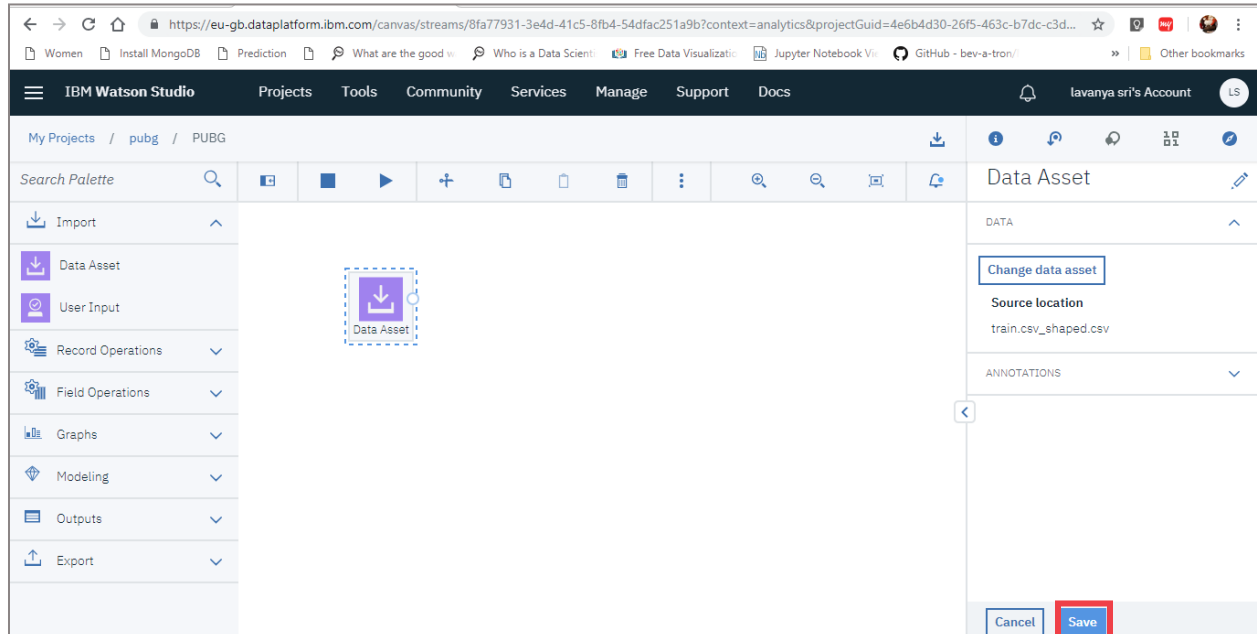
Double-click on the **Data Asset** node and click on **Change data asset** option.



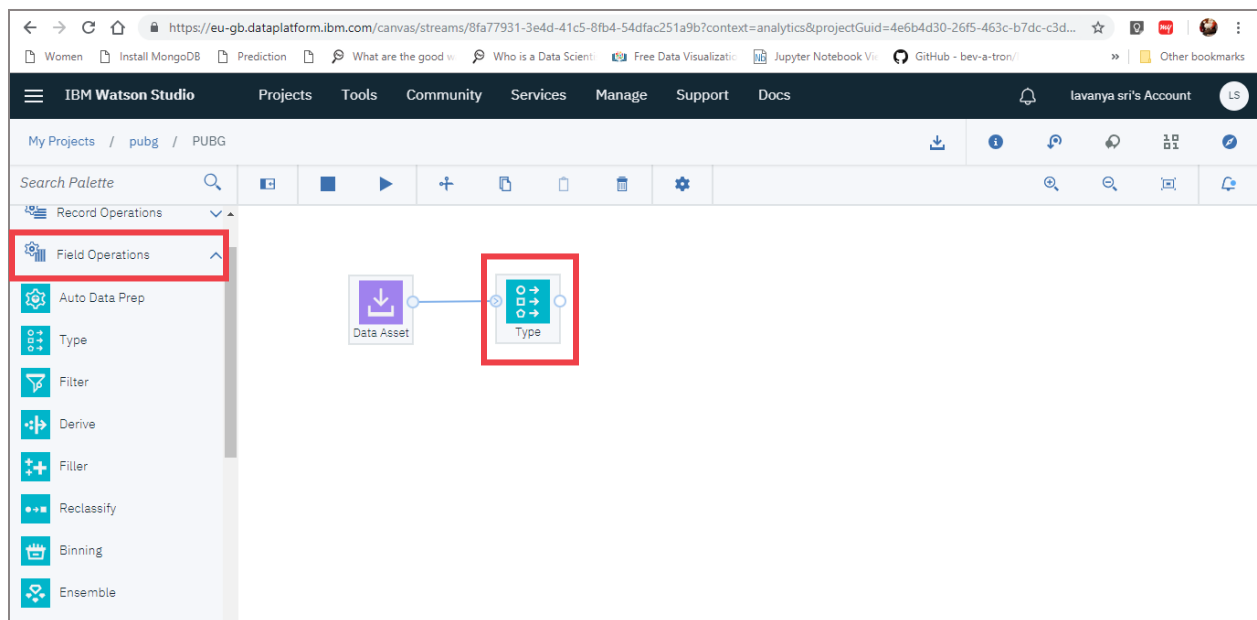
Select **train.csv\_shaped.csv** and click on **OK**.



Now click on **Save** to add our dataset successfully to the data asset node.

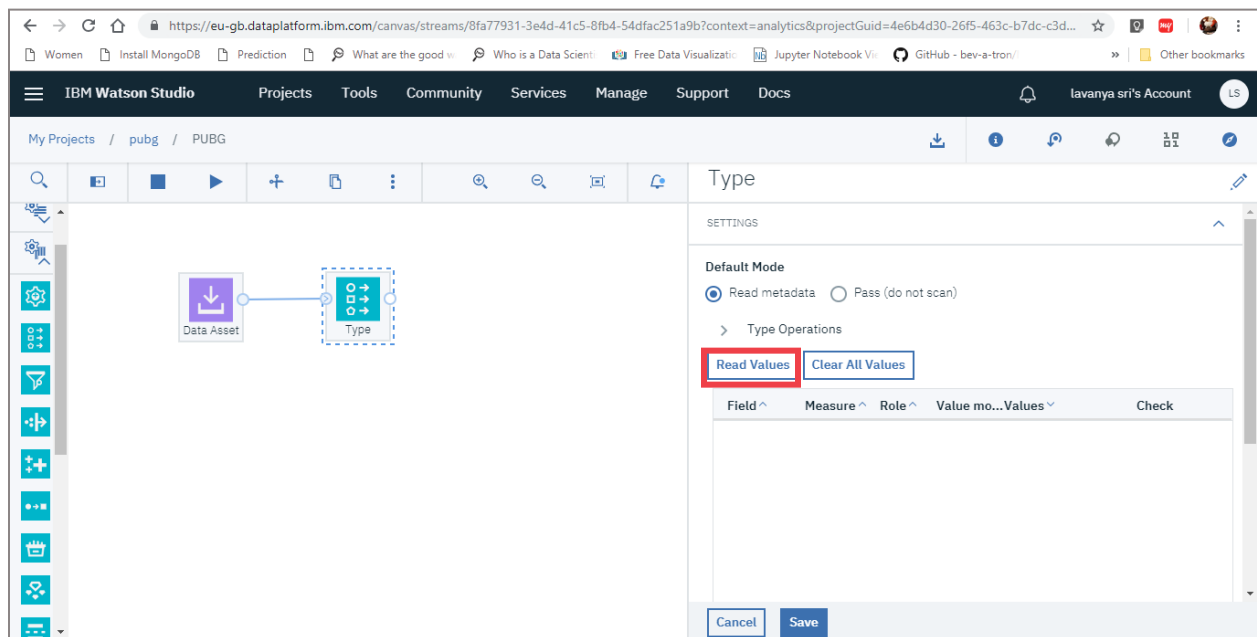


Now, click on the **Field operations palette**, drag and drop the **Type** node to the canvas and join the Data Asset node to Type node through the small rounds which are adjacent to each other.



Double click on the **Type** node and choose **Read Values** in order to read all the values from our train data.

This will take a while to process.



As we have to predict the **WinPlace Percentage** we need to make the role of it as Target variable from **Input** variable by just opting, the **Target** option from the drop down list.

The screenshot shows the IBM Watson Studio interface. In the center, a 'Data Asset' is connected to a 'Type' node. On the right, the 'Type' settings panel is open. Under 'Default Mode', 'Read metadata' is selected. Under 'Type Operations', 'Read Values' and 'Clear All Values' buttons are visible. A table lists fields with their measures and roles. The 'winPlace...' field is highlighted with a red box, and its 'Role' dropdown menu is open, showing 'Target' as the selected option.

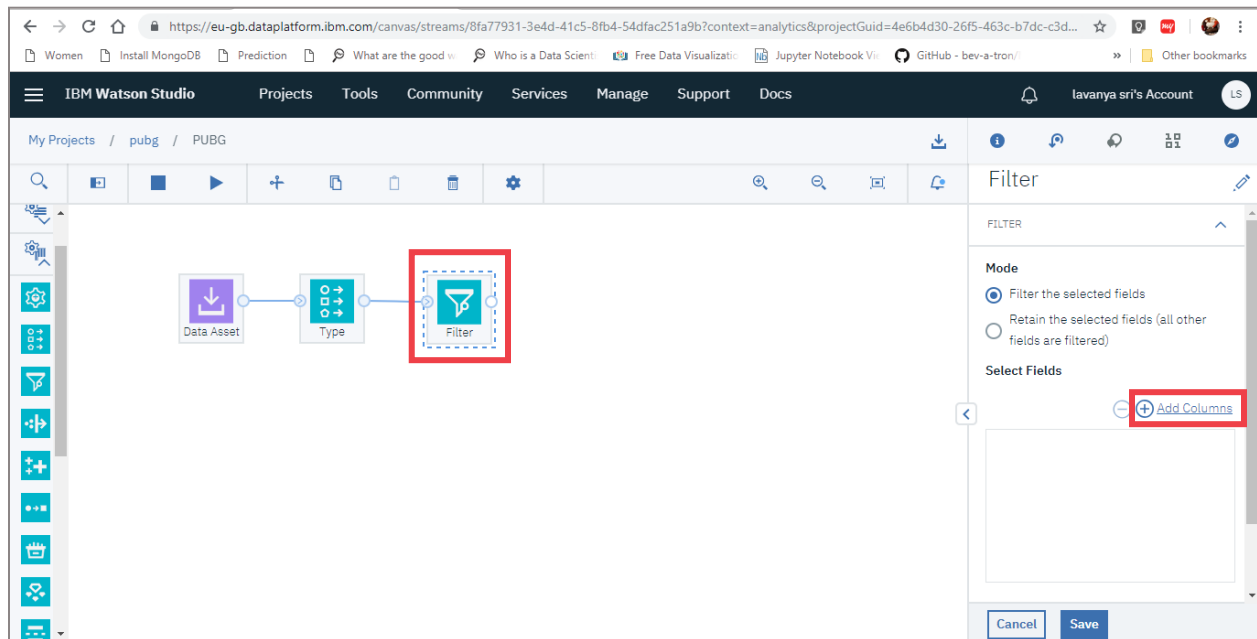
Field	Measure	Role	Value mo...Values	Check
winPlace...	Continuous	Target	Specify 0.0, 1.0	None
longestKill	Continuous	Input	Specify 0.0, 624.2	None
matchType	Continuous	None	Specify 1, 3	None
rideDista...	Continuous	Split	Specify 0.0, 17210.0	None
...	...	...	...	...

Then Click on **Save**.

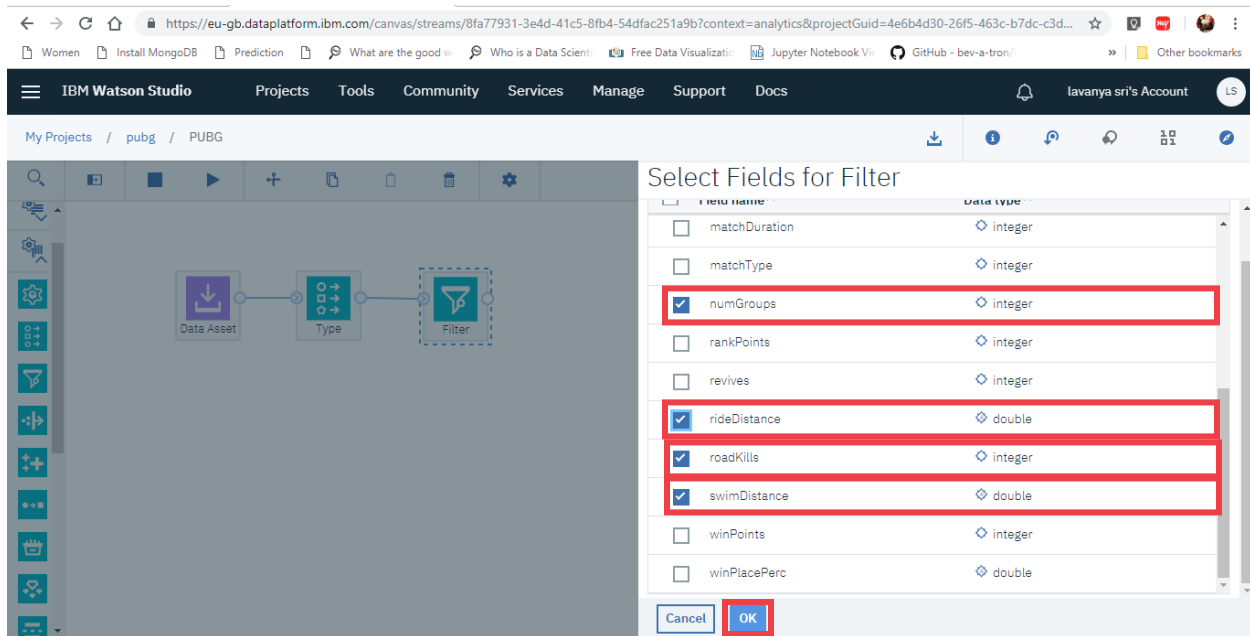
The screenshot shows the IBM Watson Studio interface. In the center, a 'Data Asset' is connected to a 'Type' node. On the right, the 'Type' settings panel is open. Under 'Default Mode', 'Read metadata' is selected. Under 'Type Operations', 'Read Values' and 'Clear All Values' buttons are visible. A table lists fields with their measures and roles. The 'winPlace...' field is highlighted with a red box. At the bottom of the panel, the 'Save' button is highlighted with a red box.

Field	Measure	Role	Value mo...Values	Check
winPlace...	Continuous	Target	Specify 0.0, 1.0	None
longestKill	Continuous	Input	Specify 0.0, 624.2	None
matchType	Continuous	Input	Specify 1, 3	None
rideDista...	Continuous	Input	Specify 0.0, 17210.0	None
...	...	...	...	...

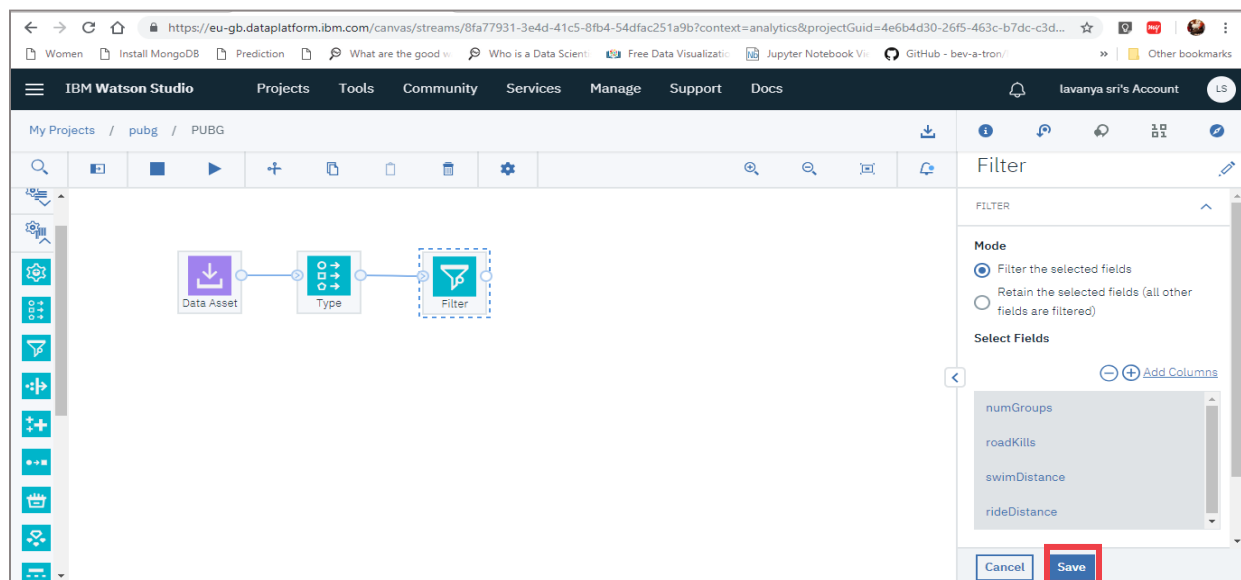
Now add a **Filter node** from the **Field operations palette** and join it to the **Type node**. Also, click on Add Columns, in order to filter out all the unnecessary fields from our dataset.



Filter out **numGroups**, **rideDistance**, **roadKills**, **swimDistance** from our dataset by selecting them. Click on **Ok** after selection.

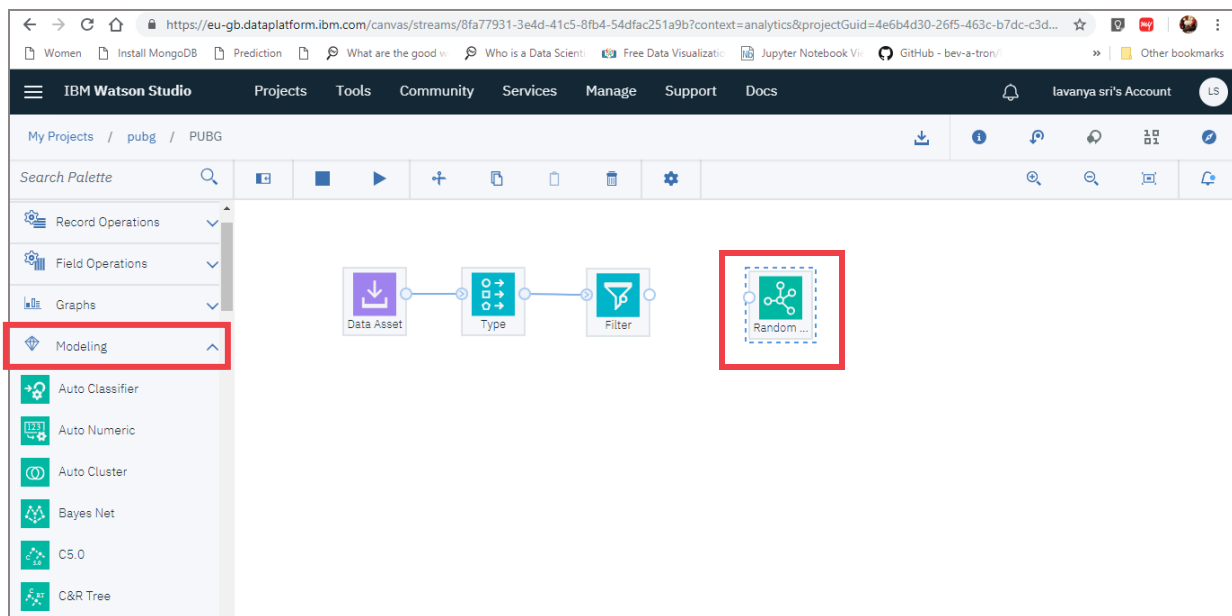


Click on **Save** to make the changes

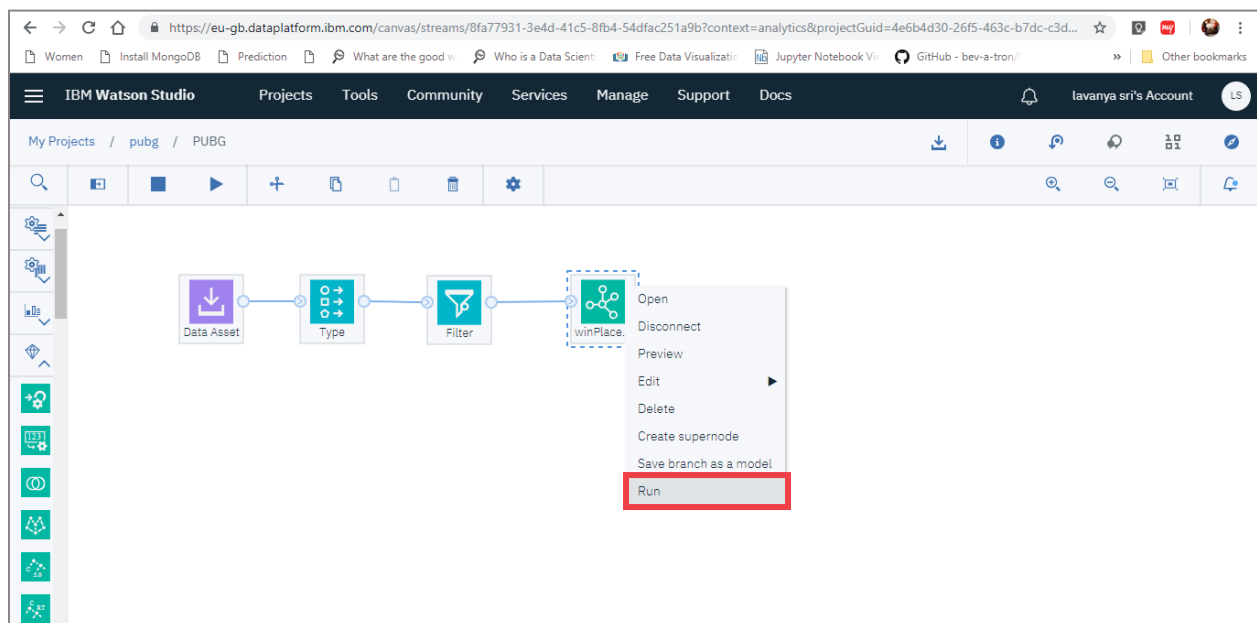


## Step #5 | Model Building

Choose an algorithm from the **Modeling** palette. Drag and drop **Random Forest** model which is an apt model for our dataset and combine it to the Filter node.

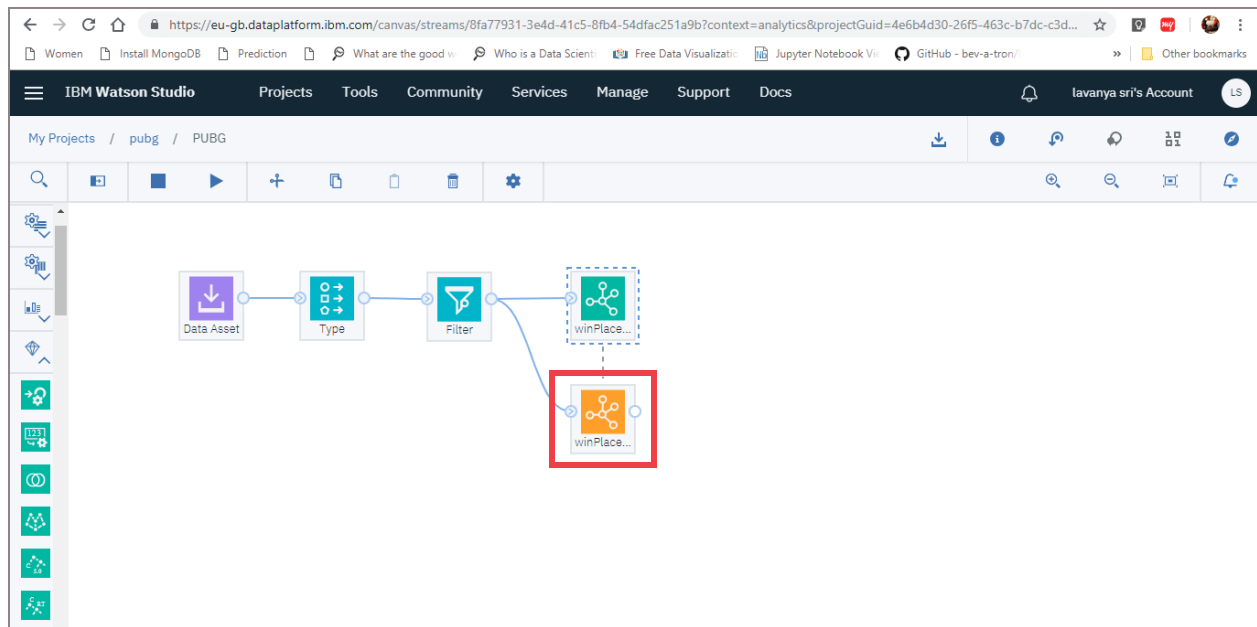


Name of Random forest will be changed to the Target variable name which is WinPlacePerc. Click on the three vertical dots on the node and click on **Run** to generate a machine learning model.

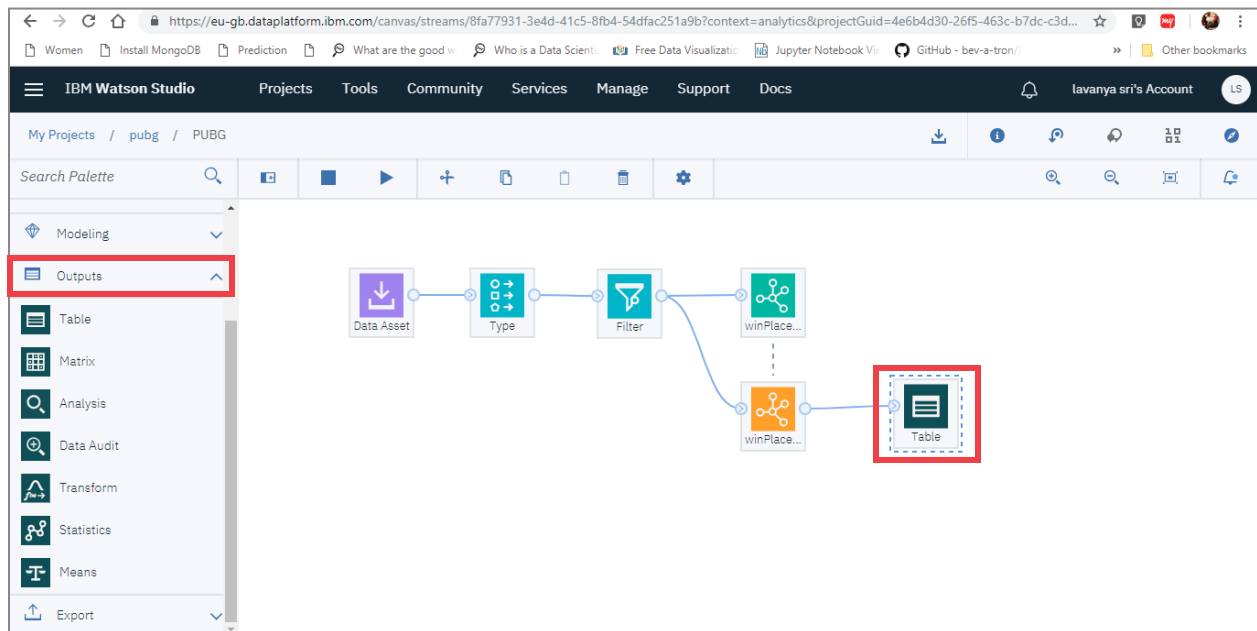


The node in orange color is the generated **model** for our dataset, with which we can predict the WinPlacePerc for new data.

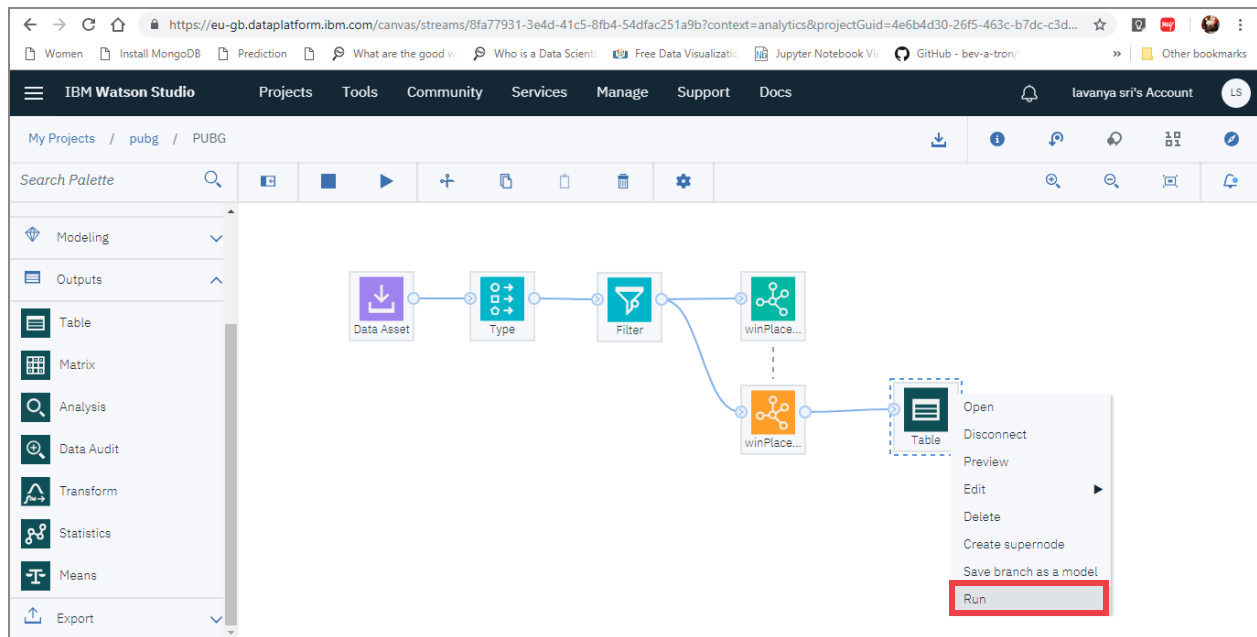




We can view the predicted values by using a Table node from Output palette.



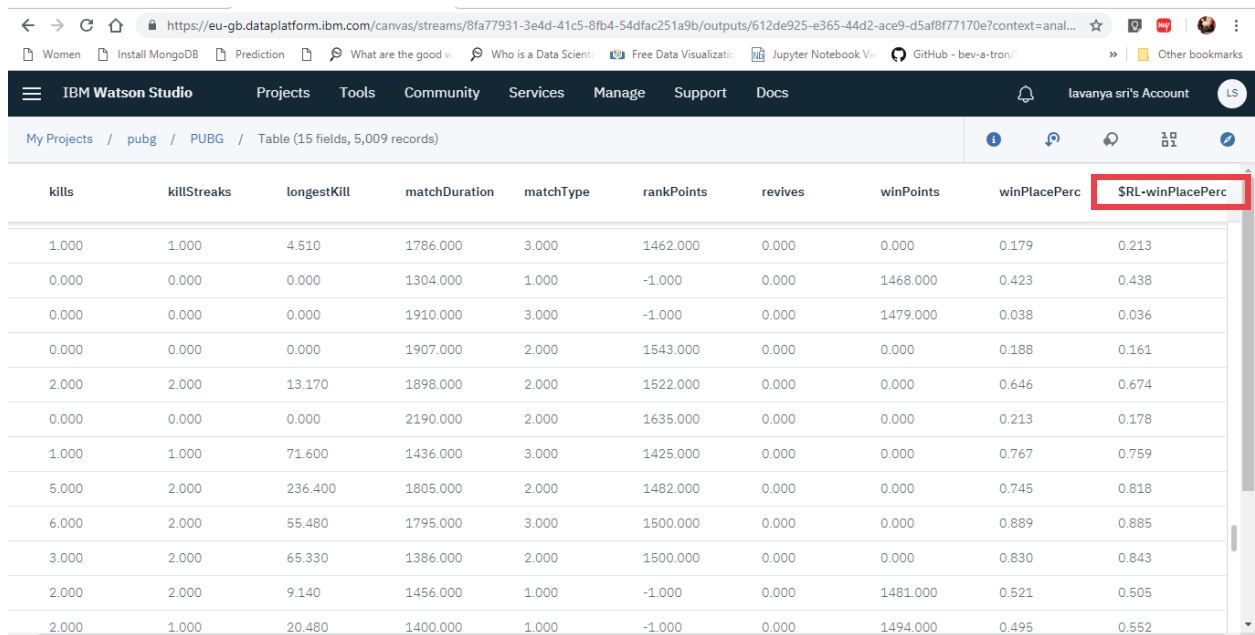
Click on **Run**, to view the newly generated values as well.



Like this we will be having a Preview of the results,

https://eu-gb.dataplatform.ibm.com/canvas/streams/8fa77931-3e4d-41c5-8fb4-54dfac251a9b/outputs/2a0fa8b6-4f2a-4106-88e5-c5f859e262d5?context=anal...

Go to page 2 and Scroll on to the End. The last column values (**\$RL-WinPlacePerc**) are the newly generated values for the given input data by the model.

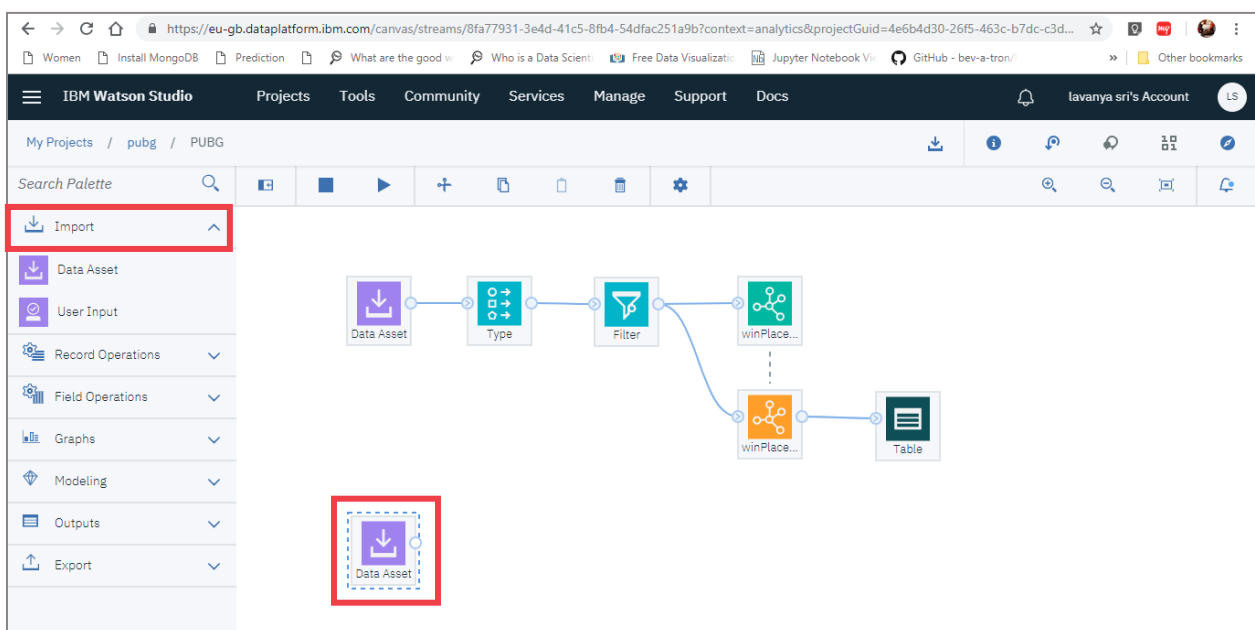


kills	killStreaks	longestKill	matchDuration	matchType	rankPoints	revives	winPoints	winPlacePerc	SRL-winPlacePerc
1.000	1.000	4.510	1786.000	3.000	1462.000	0.000	0.000	0.179	0.213
0.000	0.000	0.000	1304.000	1.000	-1.000	0.000	1468.000	0.423	0.438
0.000	0.000	0.000	1910.000	3.000	-1.000	0.000	1479.000	0.038	0.036
0.000	0.000	0.000	1907.000	2.000	1543.000	0.000	0.000	0.188	0.161
2.000	2.000	13.170	1898.000	2.000	1522.000	0.000	0.000	0.646	0.674
0.000	0.000	0.000	2190.000	2.000	1635.000	0.000	0.000	0.213	0.178
1.000	1.000	71.600	1436.000	3.000	1425.000	0.000	0.000	0.767	0.759
5.000	2.000	236.400	1805.000	2.000	1482.000	0.000	0.000	0.745	0.818
6.000	2.000	55.480	1795.000	3.000	1500.000	0.000	0.000	0.889	0.885
3.000	2.000	65.330	1386.000	2.000	1500.000	0.000	0.000	0.830	0.843
2.000	2.000	9.140	1456.000	1.000	-1.000	0.000	1481.000	0.521	0.505
2.000	1.000	20.480	1400.000	1.000	-1.000	0.000	1494.000	0.495	0.552

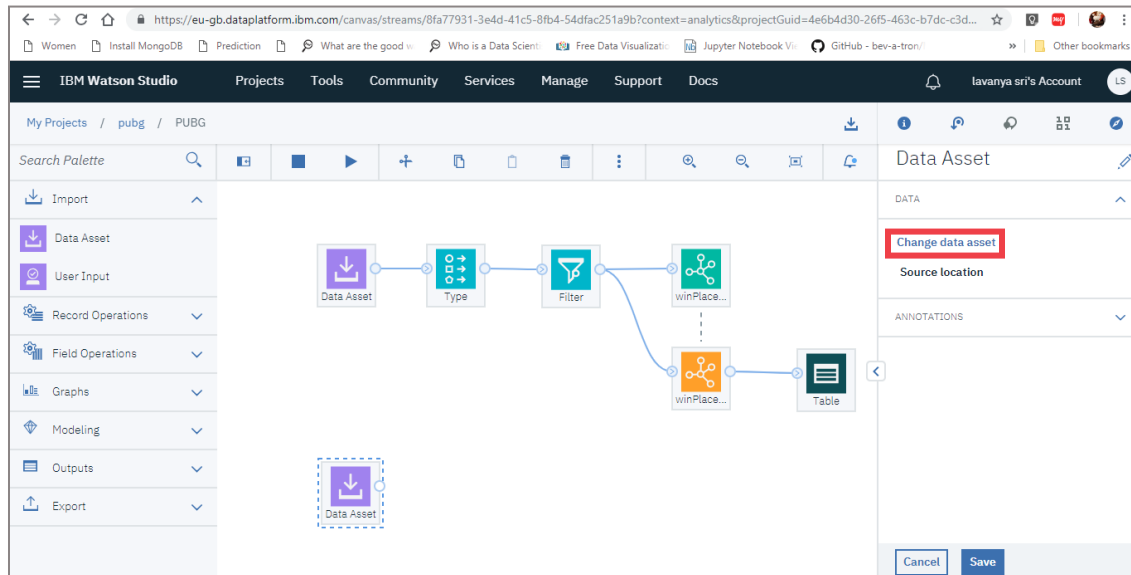
## Step #6 | Testing the Model

We will feed the Test dataset to the generated model above by creating the stream again as below,

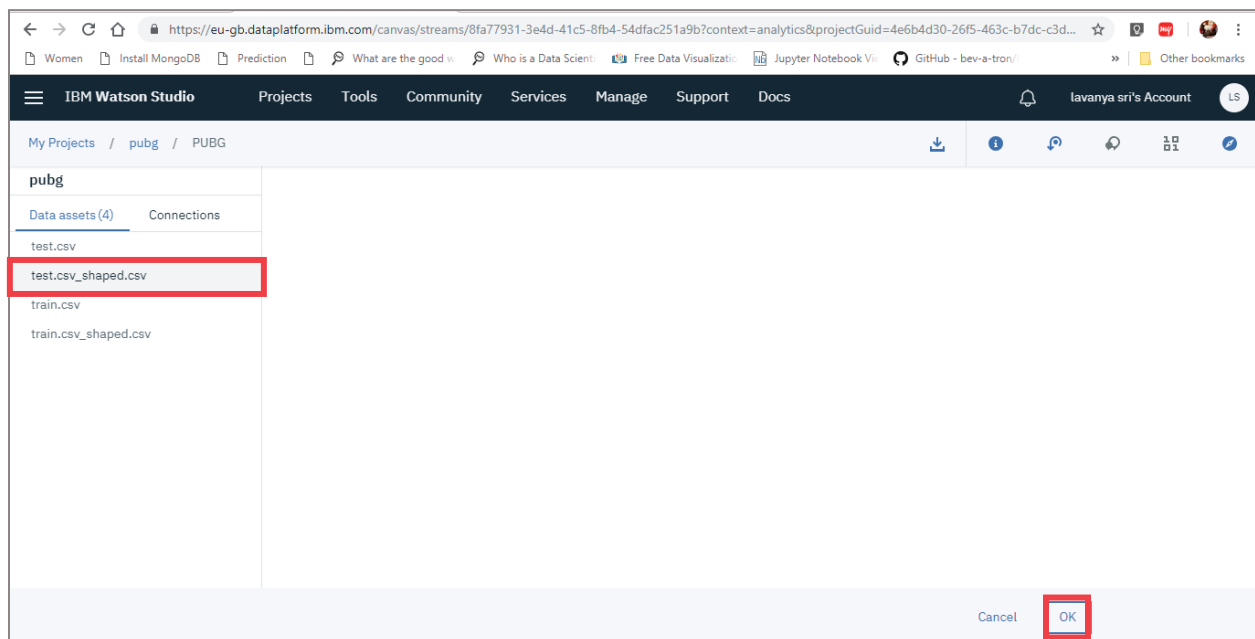
From the **Import palette** select **Data Asset** node, drag and drop it into the white space (Canvas).



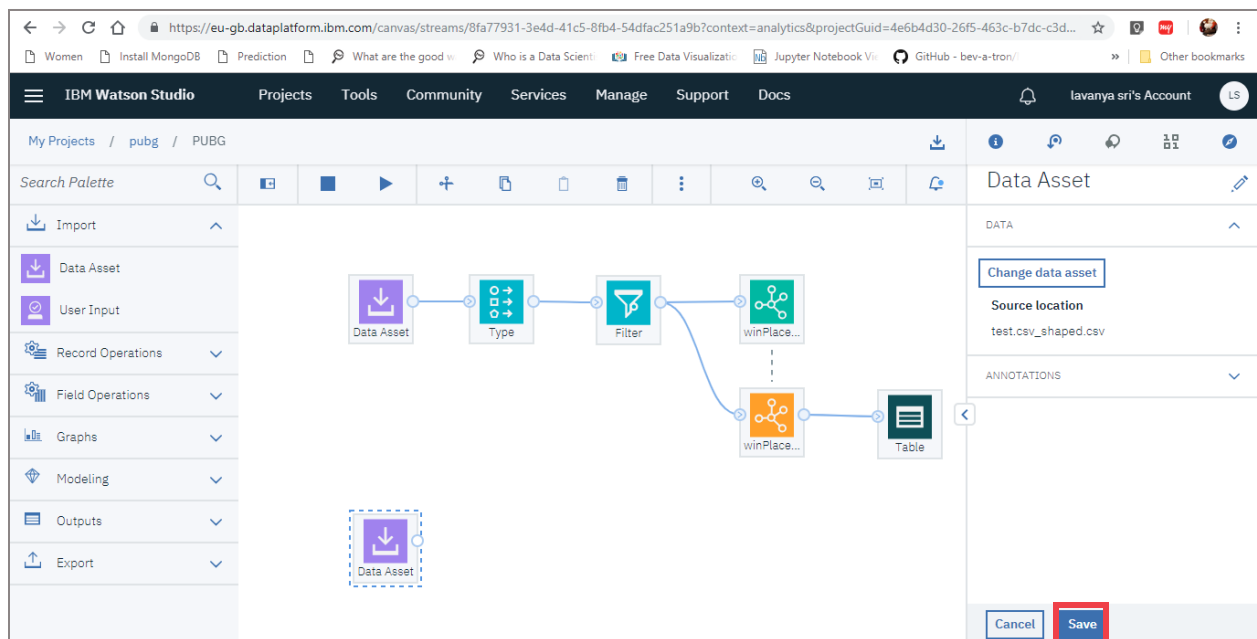
Double-click on the **Data Asset node** and click on **Change data asset** option.



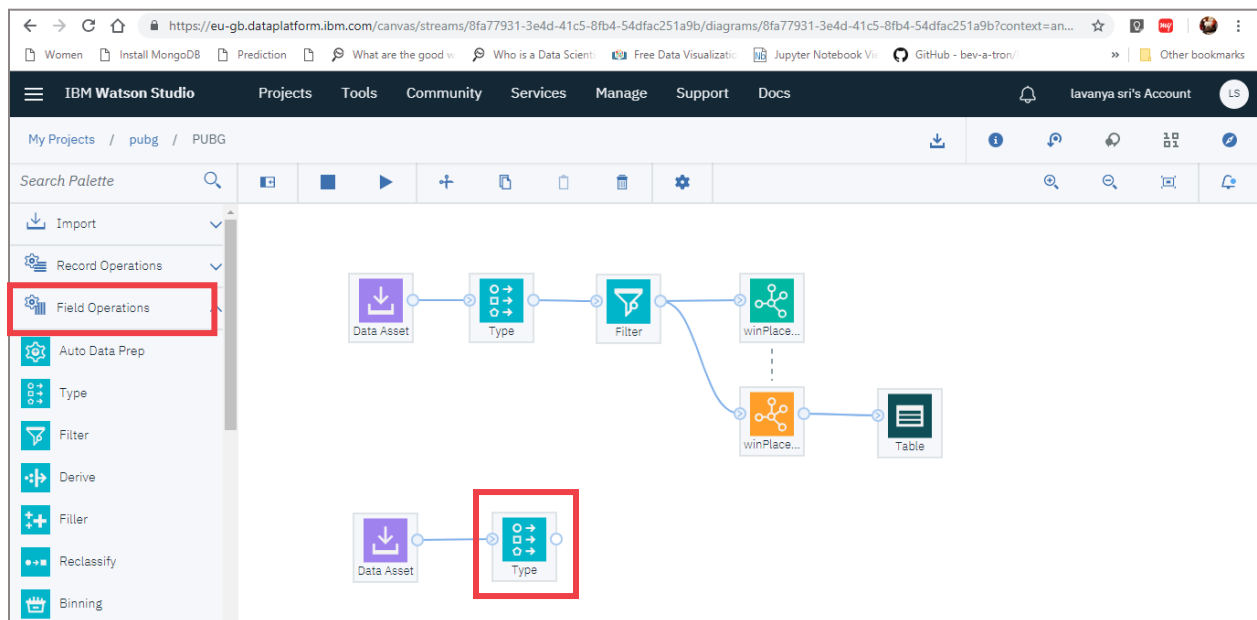
Select **test.csv\_shaped.csv** and click on **Ok**.



Now click on **Save** to successfully add our dataset to the data asset node.



Now click on the **Field operations palette**, drag and drop the **Type** node to the canvas. Join the Data Asset node to Type node through the small rounds which are adjacent to each other.



Double click on the **Type node** and choose **Read Values** in order to read all the values from our train data.

This will take a while to process.

The screenshot shows the IBM Watson Studio interface. The top navigation bar includes 'IBM Watson Studio', 'Projects', 'Tools', 'Community', 'Services', 'Manage', 'Support', and 'Docs'. The user is logged in as 'lavanya sri's Account'. The main workspace displays a workflow with three nodes: 'Data Asset', 'Type', and 'Filter'. The 'Type' node is selected, and the 'Type Operations' panel is open on the right. In this panel, the 'Default Mode' is set to 'Read metadata', and the 'Read Values' button is highlighted with a red box. Below the buttons, there is a table with columns: 'Field', 'Measure', 'Role', 'Value mo...Values', and 'Check'. The 'Save' button is visible at the bottom right of the panel.

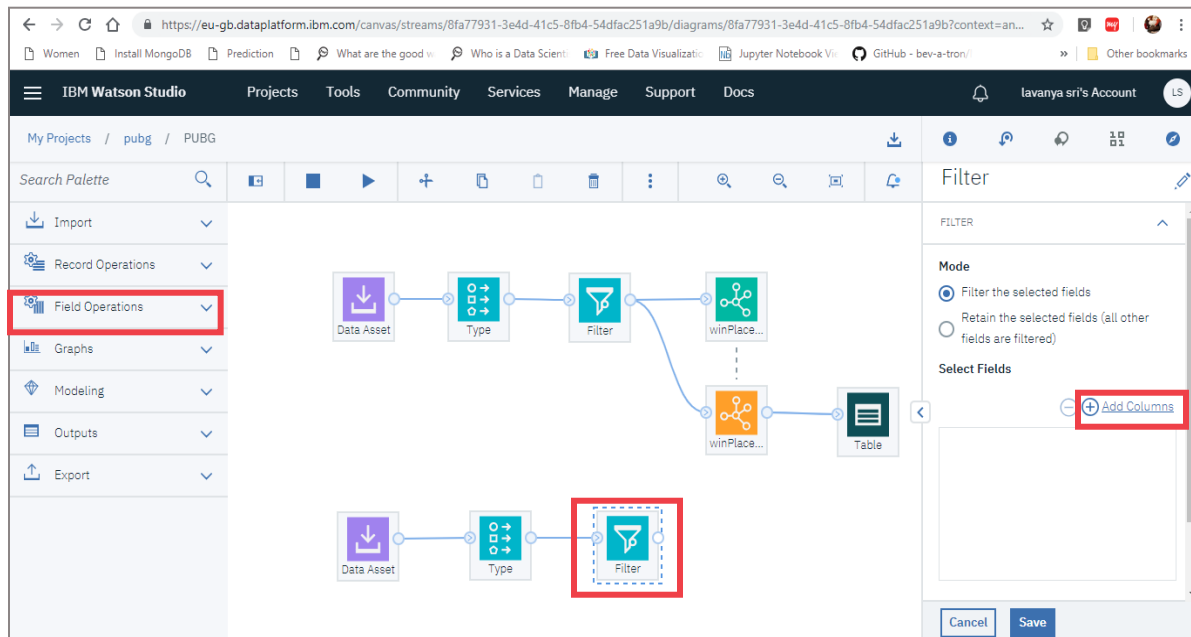
No more changes are needed here as no need of target variable with testing data as we will predict the given test input with the prebuilt model built using the training data. Click on **Save**.

The screenshot shows the same IBM Watson Studio interface as the previous one, but now the 'Type Operations' panel displays a list of fields with their measures and roles. The 'Read Values' button is still highlighted. The table in the panel contains the following data:

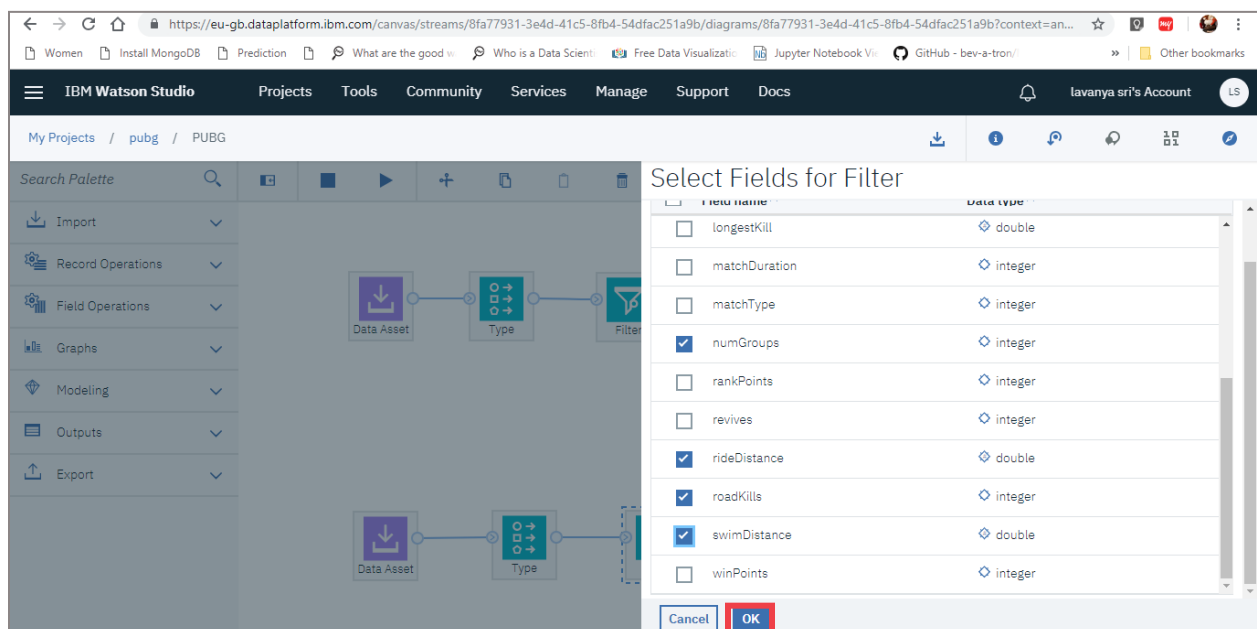
Field	Measure	Role	Value mo...Values	Check
killPoints	Continuou	Input	Specify 0, 1999	None
killPlace	Continuou	Input	Specify 1, 99	None
matchDu...	Continuou	Input	Specify 599, 2186	None
killStreaks	Continuou	Input	Specify 0, 4	None
winPoints	Continuou	Input	Specify 0, 1747	None
swimDist...	Continuou	Input	Specify 0.0, 557.8	None
roadKills	Continuou	Input	Specify 0, 2	None

The 'Save' button is now highlighted with a red box at the bottom right of the panel.

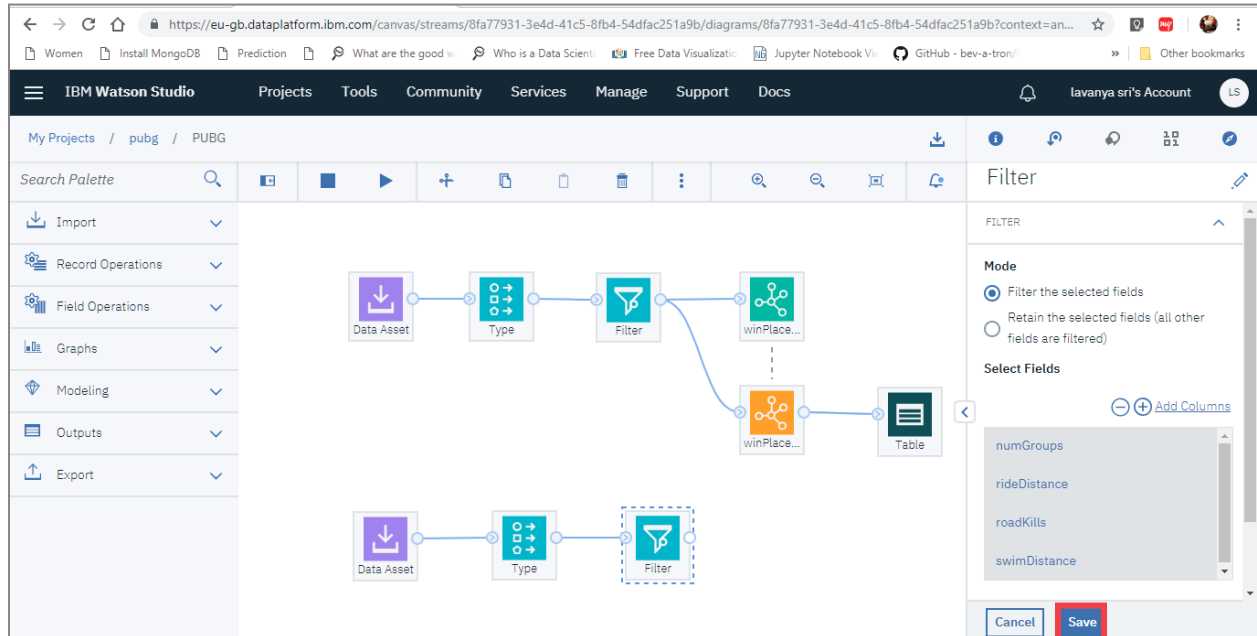
Add a **Filter** node from the **Field operations** palette. Join it to the **Type** node. Also, click on **Add Columns**, in order to filter out all the unnecessary fields from our dataset.



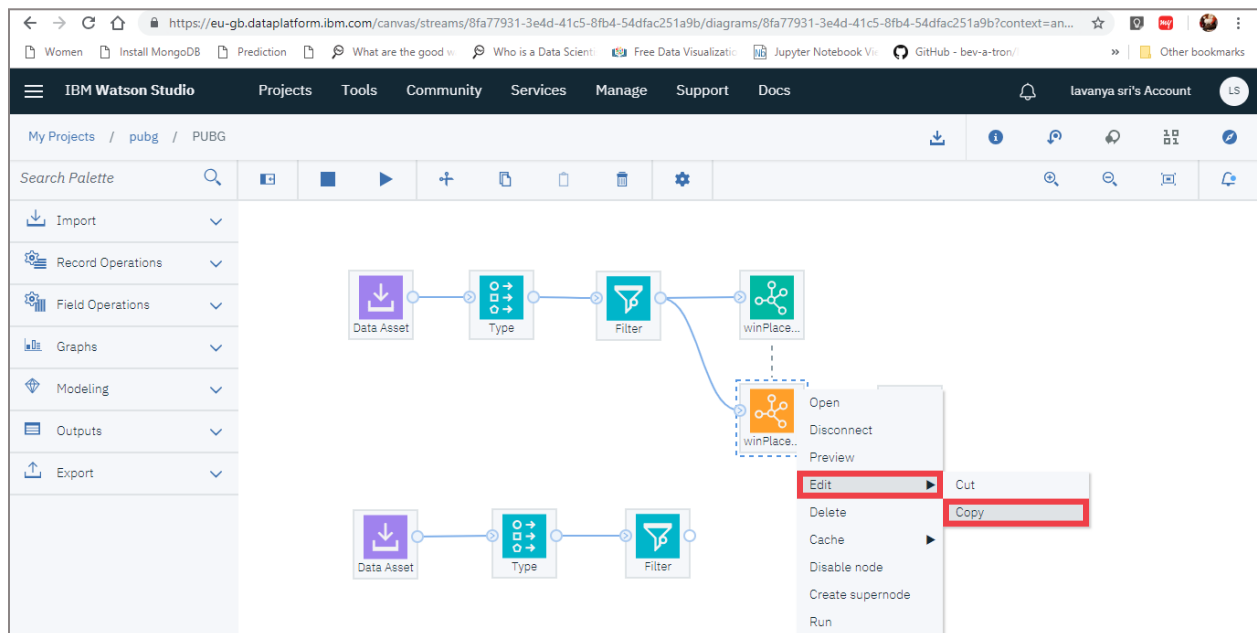
Filter out **numGroups**, **rideDistance**, **roadKills**, and **swimDistance** from our dataset by selecting them. Click on **Ok** after selection.



Click on **Save** to make the changes.

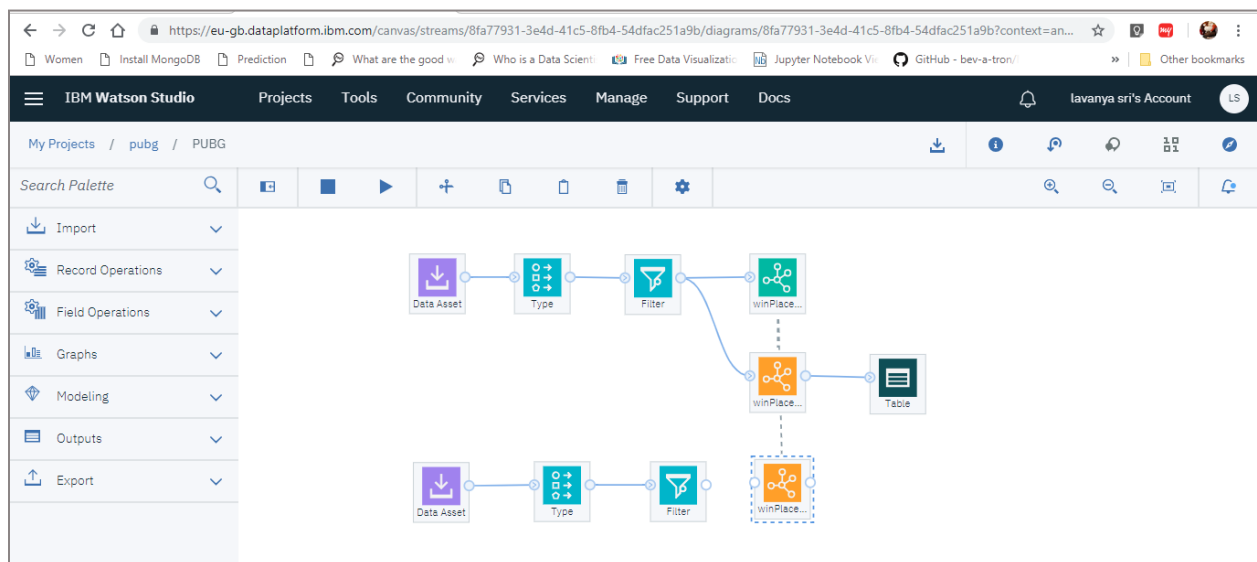
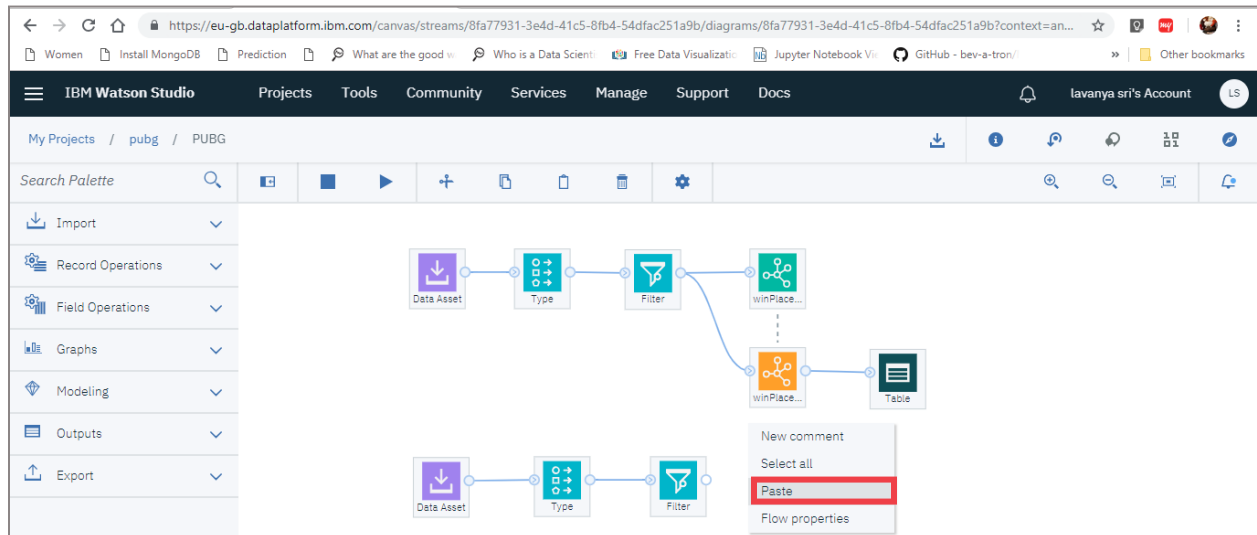


To test the created model, we will give the test data as input to the pre-built model. Hover on the generated model (orange color) then you will be able to see three dots, click on it. Choose **Edit** and select **Copy**.

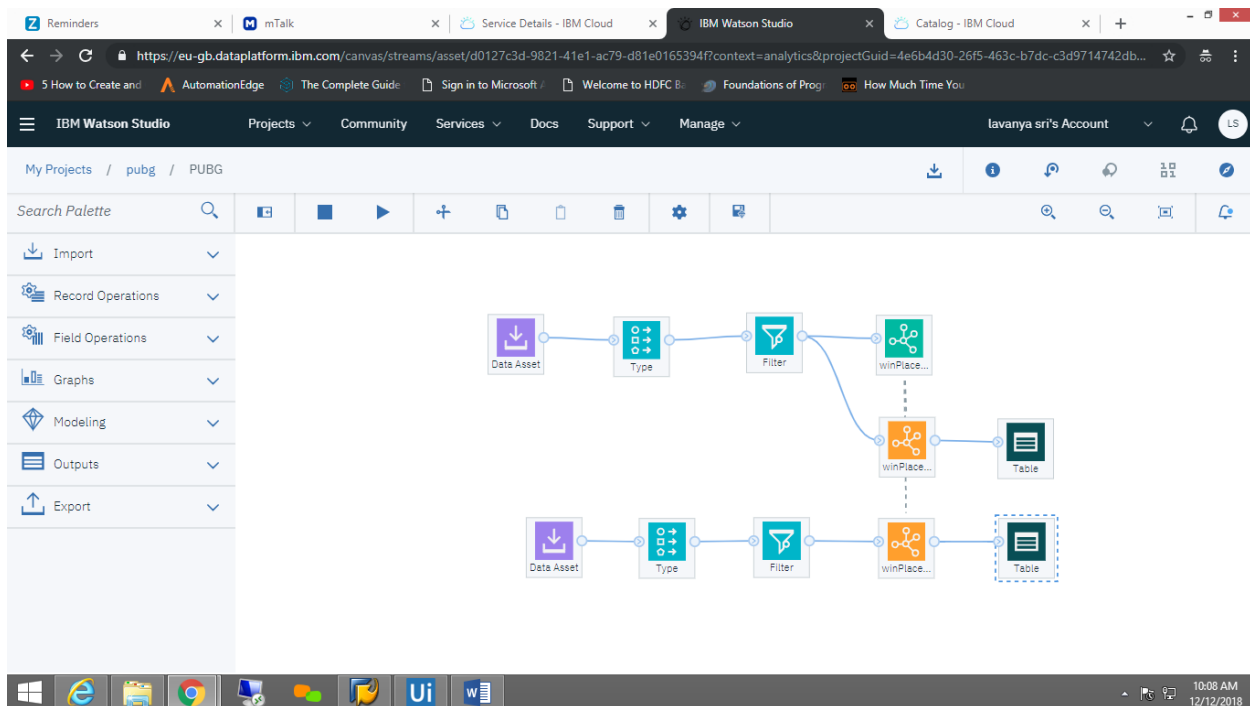




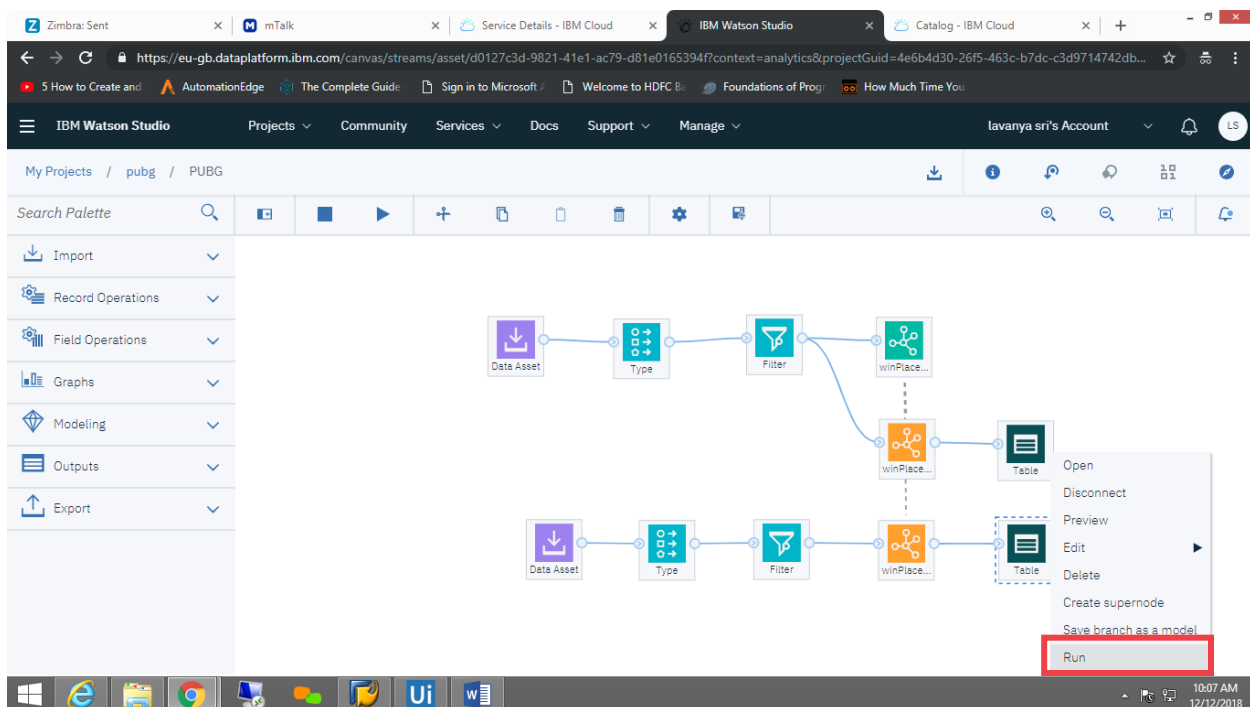
**Paste** the model next to the filter node and combine them.



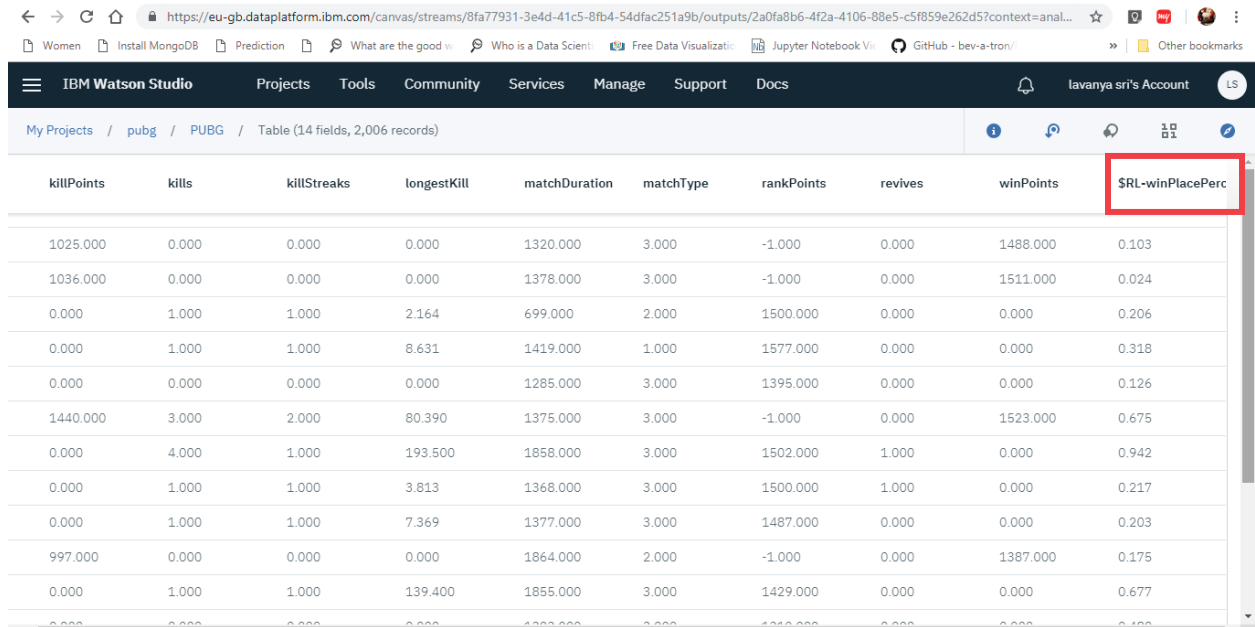
We can view the predicted values by using a **Table** node from Output palette.



Using the three dots of table node, select the **Run** option to view the newly generated values.



The last column values (**\$SRL-winPlacePerc**) are the newly predicted winplaceperc values for the given input data from the model.



The screenshot shows the IBM Watson Studio interface. The top navigation bar includes 'IBM Watson Studio', 'Projects', 'Tools', 'Community', 'Services', 'Manage', 'Support', and 'Docs'. The user's account is 'lavanya sri's Account'. The main content area displays a table with 14 fields and 2,006 records. The table columns are: killPoints, kills, killStreaks, longestKill, matchDuration, matchType, rankPoints, revives, winPoints, and \$SRL-winPlacePerc. The \$SRL-winPlacePerc column is highlighted with a red box. The table data is as follows:

killPoints	kills	killStreaks	longestKill	matchDuration	matchType	rankPoints	revives	winPoints	\$SRL-winPlacePerc
1025.000	0.000	0.000	0.000	1320.000	3.000	-1.000	0.000	1488.000	0.103
1036.000	0.000	0.000	0.000	1378.000	3.000	-1.000	0.000	1511.000	0.024
0.000	1.000	1.000	2.164	699.000	2.000	1500.000	0.000	0.000	0.206
0.000	1.000	1.000	8.631	1419.000	1.000	1577.000	0.000	0.000	0.318
0.000	0.000	0.000	0.000	1285.000	3.000	1395.000	0.000	0.000	0.126
1440.000	3.000	2.000	80.390	1375.000	3.000	-1.000	0.000	1523.000	0.675
0.000	4.000	1.000	193.500	1858.000	3.000	1502.000	1.000	0.000	0.942
0.000	1.000	1.000	3.813	1368.000	3.000	1500.000	1.000	0.000	0.217
0.000	1.000	1.000	7.369	1377.000	3.000	1487.000	0.000	0.000	0.203
997.000	0.000	0.000	0.000	1864.000	2.000	-1.000	0.000	1387.000	0.175
0.000	1.000	1.000	139.400	1855.000	3.000	1429.000	0.000	0.000	0.677

Hurrah!! With this lab you were able to create a modeler flow using IBM Watson Studio for PUBG dataset.

For any questions regarding the lab please feel free to reach out to [innovation@miraclesoft.com](mailto:innovation@miraclesoft.com). We hope you enjoyed creating Machine Learning models with us 😊