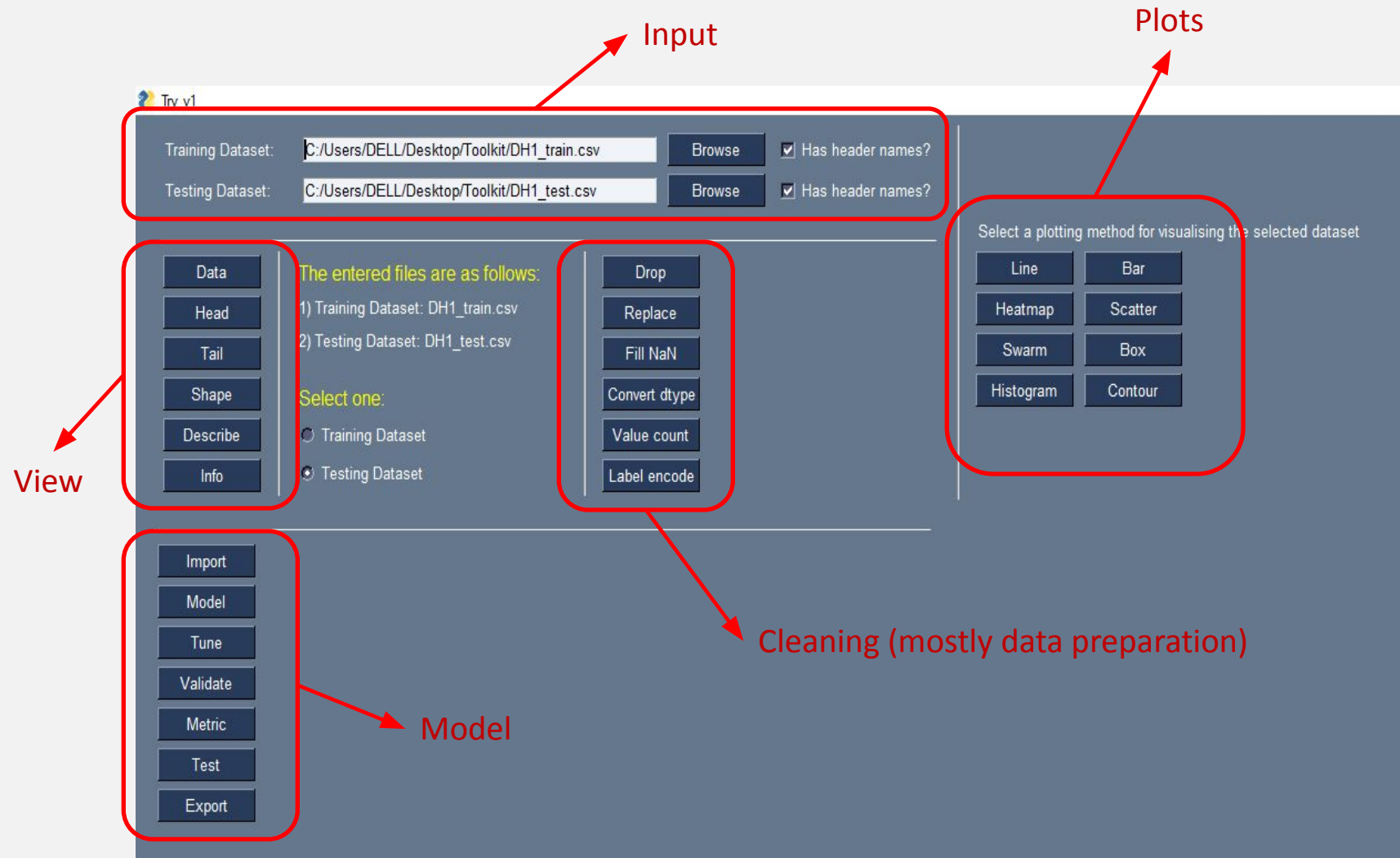
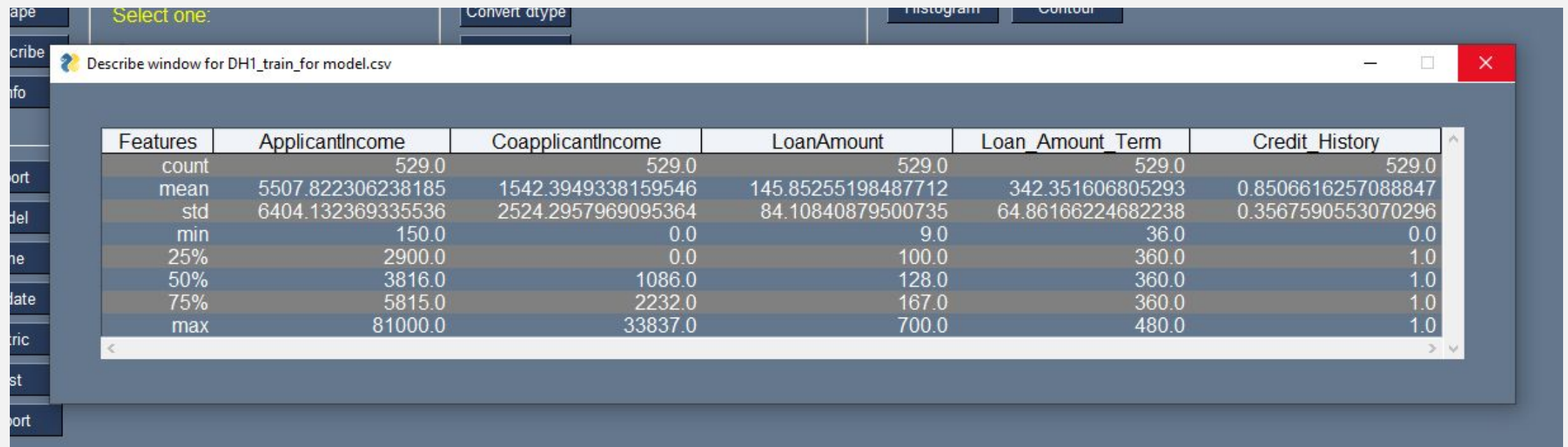


Desktop UI

- The desktop application of the toolkit looks like this.
- This is the **Main-window**.
- This window is divided into different sets of buttons, which are dedicated for specific group of operations.
- For e.g. the 6 buttons on the left(Data – Info) are for viewing the dataset related information. Hence, they are specified under **View** group



- For e.g. if I click on Describe button, from main-window, this pop-up appears which shows the statistical description of the selected dataset.
- Similarly, if I press Data, Head, Tail, Info buttons, a table pop-up window appears showing the corresponding information.



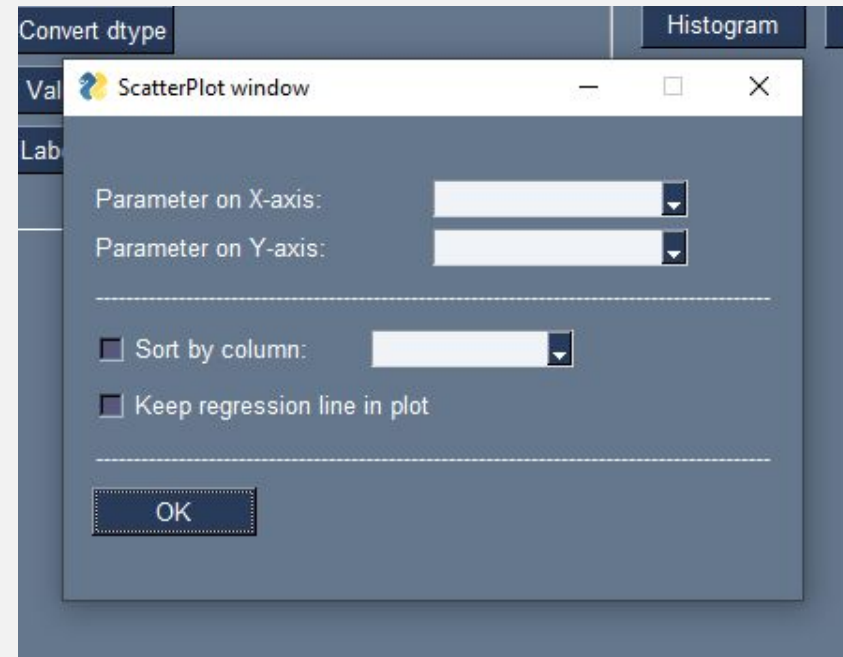
Describe window for DH1_train_for model.csv

Features	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History
count	529.0	529.0	529.0	529.0	529.0
mean	5507.822306238185	1542.3949338159546	145.85255198487712	342.351606805293	0.8506616257088847
std	6404.132369335536	2524.2957969095364	84.10840879500735	64.86166224682238	0.3567590553070296
min	150.0	0.0	9.0	36.0	0.0
25%	2900.0	0.0	100.0	360.0	1.0
50%	3816.0	1086.0	128.0	360.0	1.0
75%	5815.0	2232.0	167.0	360.0	1.0
max	81000.0	33837.0	700.0	480.0	1.0

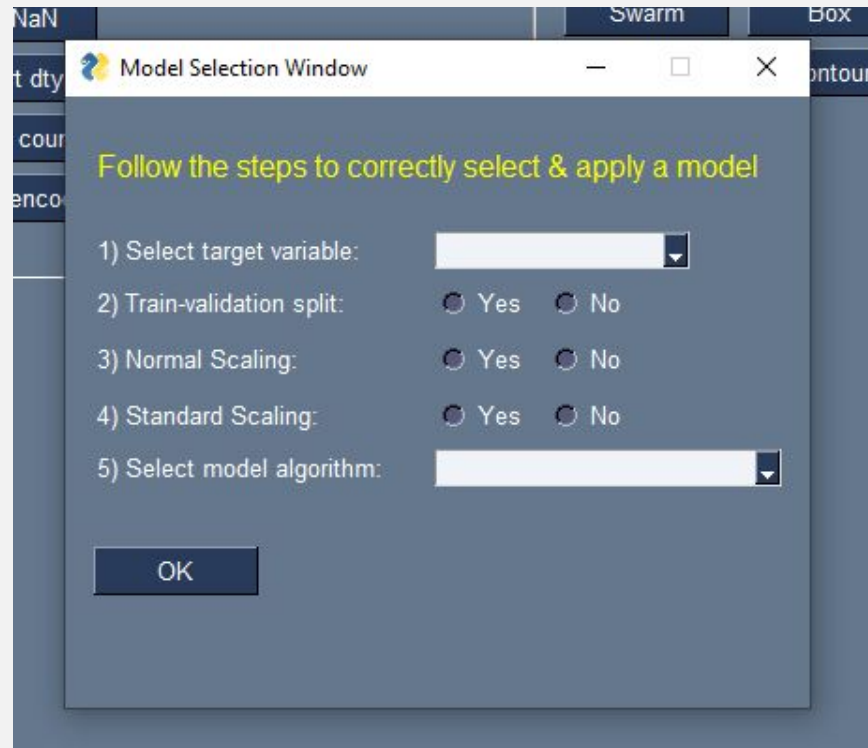
- For E.g. if I select the **Drop** button from main-window, this pop-up window appears. Similar pop-up windows appear after clicking other buttons in this group (**Cleaning group**)
- I can choose suitable options here and click on the corresponding button.
- That drop operation then gets completed.

The screenshot shows a 'Drop Selection Window' dialog box with a title bar containing a logo and window controls. The main content area is divided into two sections: 'ROWS' and 'COLUMNS', both with yellow headers. The 'ROWS' section is active, showing the question 'What do you wish to drop?' and 'Selected Dataset: DH1_train_for model.csv'. It lists four options for dropping rows: 'All NaN', 'NaN in Column:' (with a dropdown menu), 'Threshold NaN:' (with a text input field), and 'Index value:' (with a text input field). A 'Drop Row' button is at the bottom of this section. The 'COLUMNS' section is currently inactive, showing the prompt 'Select the column you wish to drop:' with a dropdown menu and a 'Drop Column' button below it. The background shows a sidebar with buttons like 'Browse', 'Drop', 'Repla', 'Fill N', 'Convert', 'Value c', and 'Label er'.

- For E.g. if I select the **Scatter** button from main-window, this pop-up window appears. Similar pop-up windows appear after clicking other buttons in this group (**Plots group**)
- I can choose suitable options here and click on the button.
- The plot is then shown in a new chrome tab.





- For E.g. if I select the **Model** button from main-window, this pop-up window appears. Similar pop-up windows appear after clicking other buttons in this group (**Model group**)
- I can choose suitable options here and click on the button.
- The selected model then gets created.



Web UI

User Authentication window





Toolkit Logo

Username



Password

Log In



Menu / Project Selection window (New project)

 	
Toolkit Logo	Menu
New Open	

Menu / Project Selection window (New project)

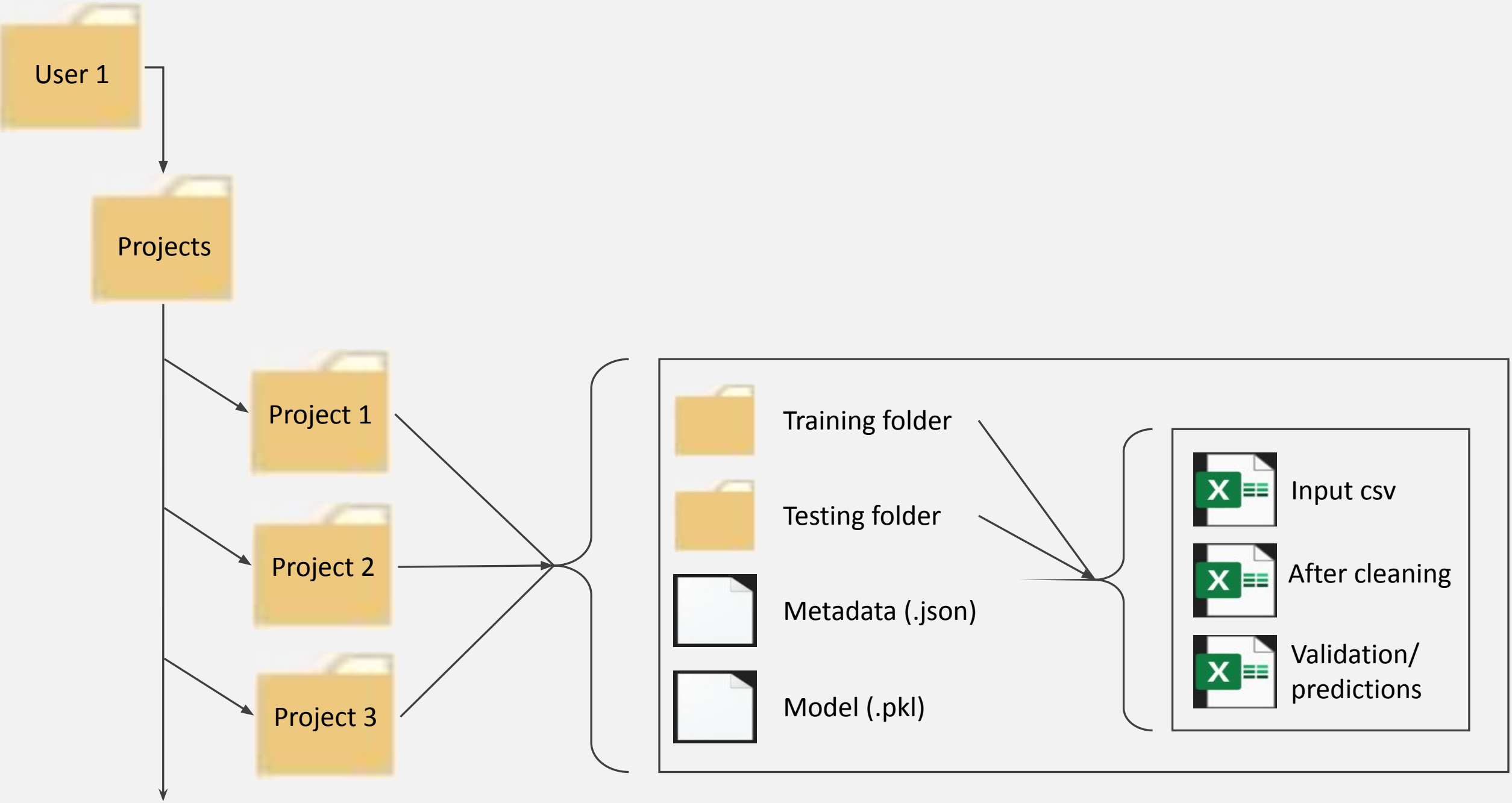
 	
Toolkit Logo	Menu
New Open	<div><p>New Project id:</p><input type="text" value="Project 4"/><p>Add Description:</p><div><div></div><div></div><div></div></div></div> <div>OK</div>

Menu / Project Selection window (Open project)



 						
Toolkit Logo	Menu					
<p>New</p> <p>Open</p>	<div><div><p>Project list:</p><table border="1"><tbody><tr><td>Project 1</td></tr><tr><td>Project 2</td></tr><tr><td>Project 3</td></tr></tbody></table></div><div><p>Project description:</p><table border="1"><tbody><tr><td>This is a demonstration project. The name of this project is "Project 2".</td></tr><tr><td>This was the entered description while creating this project.</td></tr></tbody></table></div></div> <p>OK</p>	Project 1	Project 2	Project 3	This is a demonstration project. The name of this project is "Project 2".	This was the entered description while creating this project.
Project 1						
Project 2						
Project 3						
This is a demonstration project. The name of this project is "Project 2".						
This was the entered description while creating this project.						

Brief explanation of the folder structure

- The folder structure aims at storing (saving) the projects and the performed operations in those projects.
- A major folder, named 'Projects' will contain all the projects folders (Project 1, Project 2, Project 3, etc).
- A new project folder (say, Project x), will be automatically created when the user clicks "OK" in the Project selection window (New project).
- These individual project folders will contain the metadata file, training folder (if training set was imported), testing folder (if testing set was imported), and the published model file (if model was published).
- The training folder will contain csv files for input, after cleaning operations, and after validation. the testing folder will contain csv files for input, after cleaning operations, and final predictions.
- These mentioned csv files will be automatically created & saved in the corresponding folders after the respective operation is successfully performed.



Workflow Selection window



Toolkit Logo

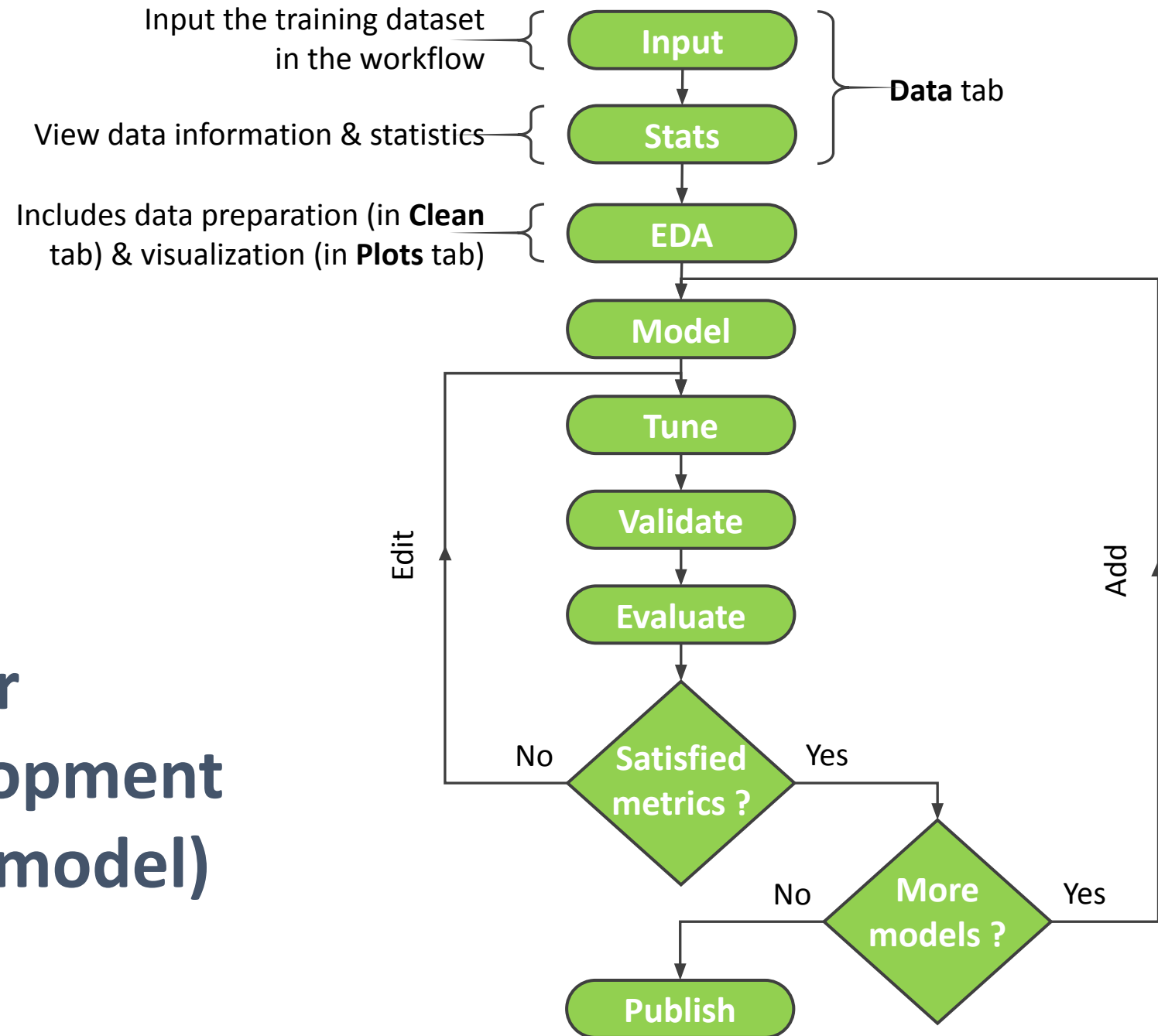
Project 2

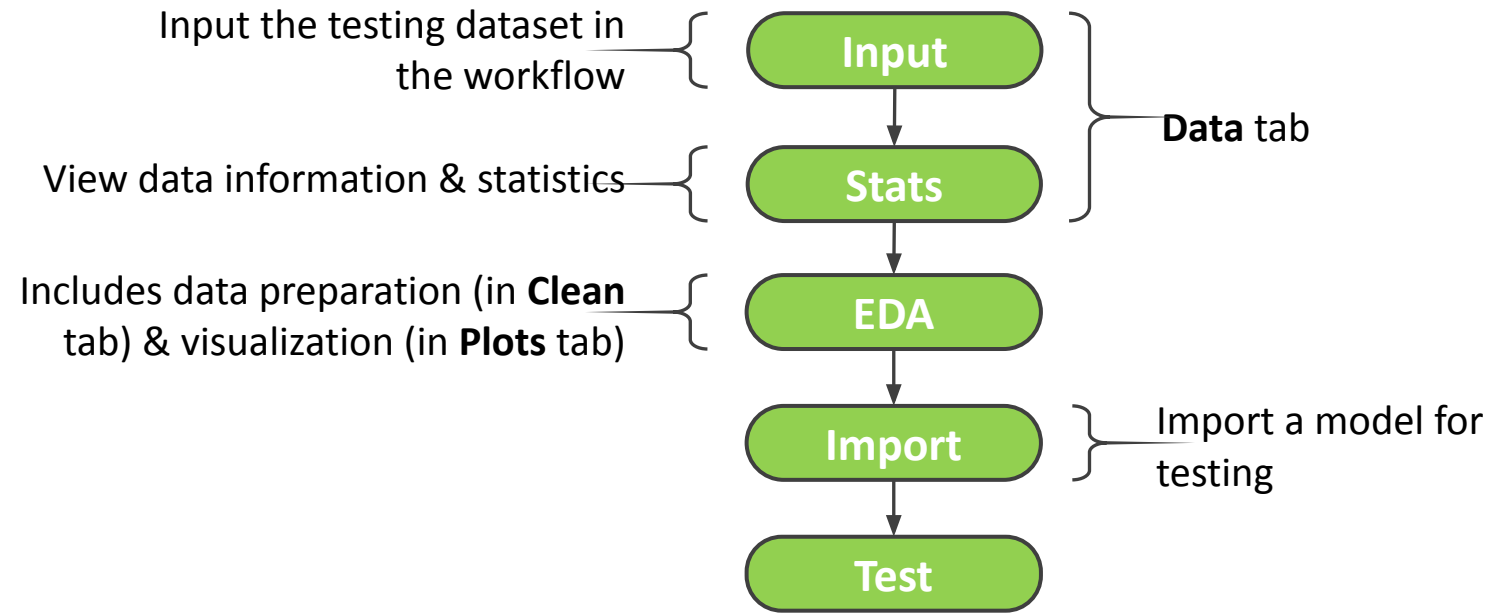
Select workflow

MODEL DEVELOPMENT
(Train)

MODEL CONSUMPTION
(Test)

Work-flow for Model-development (Build a new model)





Work-flow for Model-consumption (Import model & test)

Model-development workflow UI

Data tab

Data tab has been clicked

Current *Shape* of the selected dataset (if it gets updated, the numbers here shall change)

These options appear here when **Data** tab is selected. Here, **Describe** has been clicked.

Toolkit Logo	Project 2	Training	DataClean		PlotsModel		}	Tabs																																		
(900, 20)	Input Training set: <input type="button" value="Browse"/>																																									
Data Head Tail Describe Info	<table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>																																									

Click on Browse button to input the training dataset

Table format for all options given on left (only the contents of table shall change depending upon the option selected)

Brief explanation of the previous slide(Data tab)

- The main objective of this tab is to input (browse) & view the training dataset. Statistical information can also be viewed.
- The shape (as shown on the left side, below the toolkit logo), indicates the current (or updated) no. of rows and no. of columns in the selected dataset. The format is *(no. of rows, no. of columns)*.
- The Data option when clicked, shows the entire dataset in the table format alongside. This option is selected by default. So that, when the dataset is browsed successfully, the data is readily displayed without user needing to click on this option.
- The Head option when clicked, shows only the first 5 rows of the dataset in the table format alongside. Similarly for Tail option, it shows the last 5 rows of the dataset.
- The Describe option when clicked, shows the statistical information of the dataset in the table format alongside.
- The Info option when clicked, shows the columns related info (names, dtypes, count/percent of nan values) for the dataset in the table format alongside

Clean tab

Toolkit Logo	Project 2	Training	Data Clean		PlotsModel		
(900, 20)							
Clean options when selected opens a specific tab at the same place	Drop						
	Replace						
	Fill						
	Convert						
	Encode						

Automatically highlighted rows indicate the presence of 1 or more *nan* values in them

Brief explanation of the previous slide(Clean tab)

- The main objective of this tab is data preparation (cleaning, converting, encoding, etc)
- The options for data preparation are at left. Any of the option when clicked, opens a sub-window (related to the clicked option) at the same place. This is clearly shown in the upcoming slides.
- The user can choose the appropriate options from this sub-window and click on the appropriate button. The changes will be reflected in the table format alongside. The user can again go to Data tab and select what information to be displayed.
- For each sub-window, an exit close button (top-right corner of sub-window) is provided to go back to the options display. This is more clearly shown in the upcoming slides.

Drop

Replace

Fill

Convert

Encode

- Drop
- Replace
- Fill**
- Convert
- Encode

Toolkit Logo

(900, 20)

Num Cols

v

☐ Mean

☐ Mode

☐ Median

☐ Interpolation

☐ Previous

☐ Unique

Review

Fill

Cat Cols

v

☐ Most frequent

☐ Unique

Review

Fill

Project 2

Training

Data Clean

PlotsModel

This **Fill** tab has been shown hanging down due to lack of space here. (In actual web UI, it will fit in the space, if it doesn't, a scroll bar can be introduced as shown)

Plots tab

Toolkit Logo	Project 2	Training	Data Clean	Plots	Model
Line Scatter Heatmap Box Histogram Contour Swarm Bar	The selected plot shall be displayed here.				

These options appear here when **Plots** tab is selected. Note that there is no *shape* display under this tab

Brief explanation of the previous slide(Plots tab)

- The main objective of this tab is data visualization (plots, heatmap, histogram, etc)
- The options for types of plots are at left. Any of the option when clicked, opens a sub-window (related to the clicked option) at the same place. This is clearly shown in the upcoming slides.
- The user can choose the appropriate options from this sub-window and click on the appropriate button. The plot will be then displayed in the space alongside.
- For each sub-window, an exit close button (top-right corner of sub-window) is provided to go back to the options display. This is more clearly shown in the upcoming slides.

Line

Scatter

Heatmap

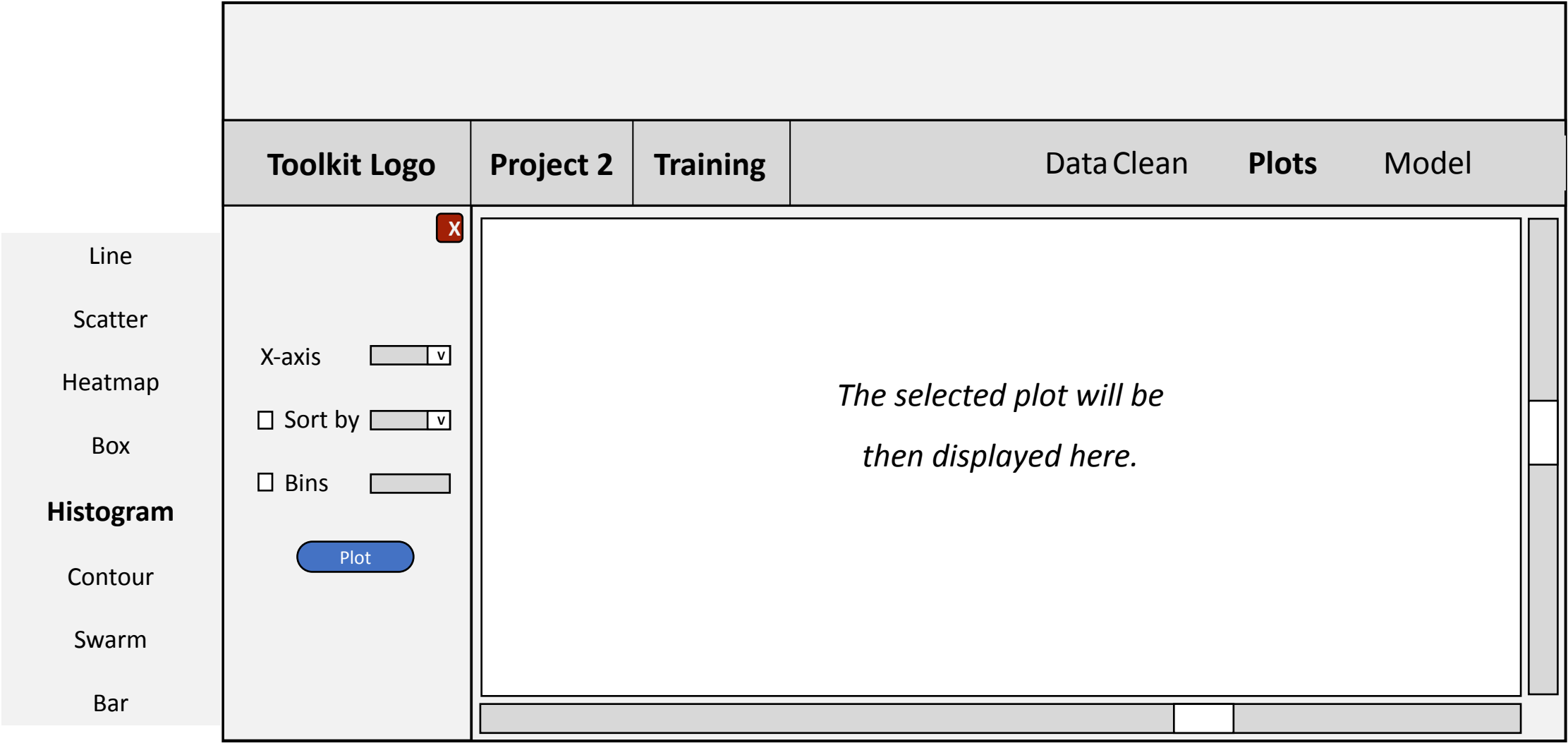
Box

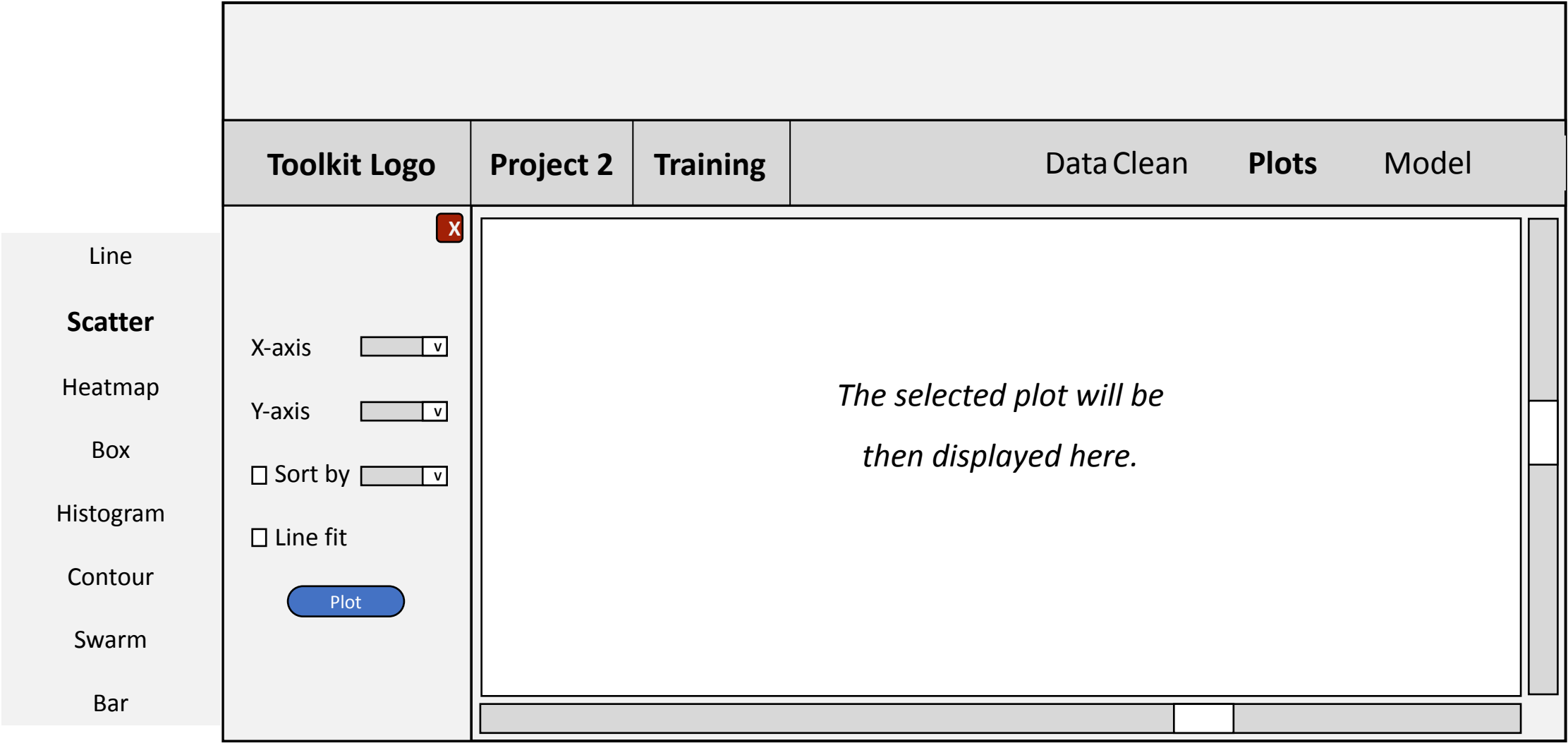
Histogram

Contour

Swarm

Bar





Model tab

Toolkit Logo

Project 2

Training

Data Clean

PlotsModel

Models:

Model 1

Model 2

Model 3

Add

Edit

Delete

Compare

Model Parameters:

1) Target variable:

v

2) Validation split:

☐ Yes

☐ No

3) Re-shuffle split:

☐ Yes

☐ No

4) Algorithm:

v

5) Normalization:

☐ Yes

☐ No

6) Standardization:

☐ Yes

☐ No

Algorithm Parameters:

1) Parameter 1:

2) Parameter 2:

3) Parameter 3:

Fit model

Validate

Evaluation metrics:

Progress bar

0

100

Publish

Test

Add a new model

Edit the selected model

Delete the selected model

Compare selected models

Mandatory step

Directly takes to Data tab of Testing workflow

Brief explanation of the previous slide(Model tab)

- The main objective of this tab is model building, tuning, validation, evaluation (using metrics) & publish (export model).
- In Model parameters part, the selection of options is inter-related. For e.g. if the user selects Yes for Validation-split option, only then the options for Re-shuffle split will be activated. Until then, the options for shuffle split will be faded and cannot be selected.
- Similarly, if the selected algorithm is not affected by feature scaling, the Normalization & Standardization options will be faded (or deactivated, user wont be able to select them). Refer this [link](#) for detailed explanation of feature-scaling (Normalization & Standardization)
- On the same lines, if the user selects Yes for Normalization, the options for Standardization will be faded. They can be re-activated by clicking No for Normalization. Similar in case of Standardization.
- Now, the Algorithm parameters part, this part will be activated only after the Algorithm is selected in the Model parameters part. The parameters related to the selected Algorithm will be displayed here with their default values. If the Algorithm is changed, the parameters here will be changed. (There is a different set of parameters for different algorithms, only some of them match)
- It should be noted that in actual web UI, there should be enough space to accommodate at least 6 such parameters. Here only 3 are shown for representation.

...Continue

- When the button Fit model is clicked, a display message should be displayed whether the fit was successful or not.
- When the button Validate is clicked, and if it is successful, all evaluation metric scores should be readily updated in the Evaluation metrics section.
- When the button Publish is clicked, a .pkl file of the model is created, a display message should prompt that it is done successfully.

Model-consumption workflow (UI)

Data tab

Toolkit Logo	Project 2	Testing	DataClean PlotsTest				
(900, 20)	Input Testing set: <div>Browse</div>						
Data Head Tail Describe Info							

Clean tab

Toolkit Logo	Project 2	Testing	Data Clean	PlotsTest
--------------	-----------	---------	-------------------	-----------

Same as functions under the **Clean** tab
as shown for model-development
workflow.

The only change is as shown above.
(**Test** tab instead of **Model** tab)

Plots tab

Toolkit Logo	Project 2	Testing	Data Clean	Plots	Test
--------------	-----------	---------	------------	--------------	------

Same as functions under the **Plots** tab
as shown for model-development
workflow.

The only change is as shown above.
(**Test** tab instead of **Model** tab)

Toolkit Logo	Project 2	Testing	Data Clean	PlotsTest
<div>Import model: <input type="button" value="Browse"/></div> <div>Generate: <input type="button" value="Test"/></div>				