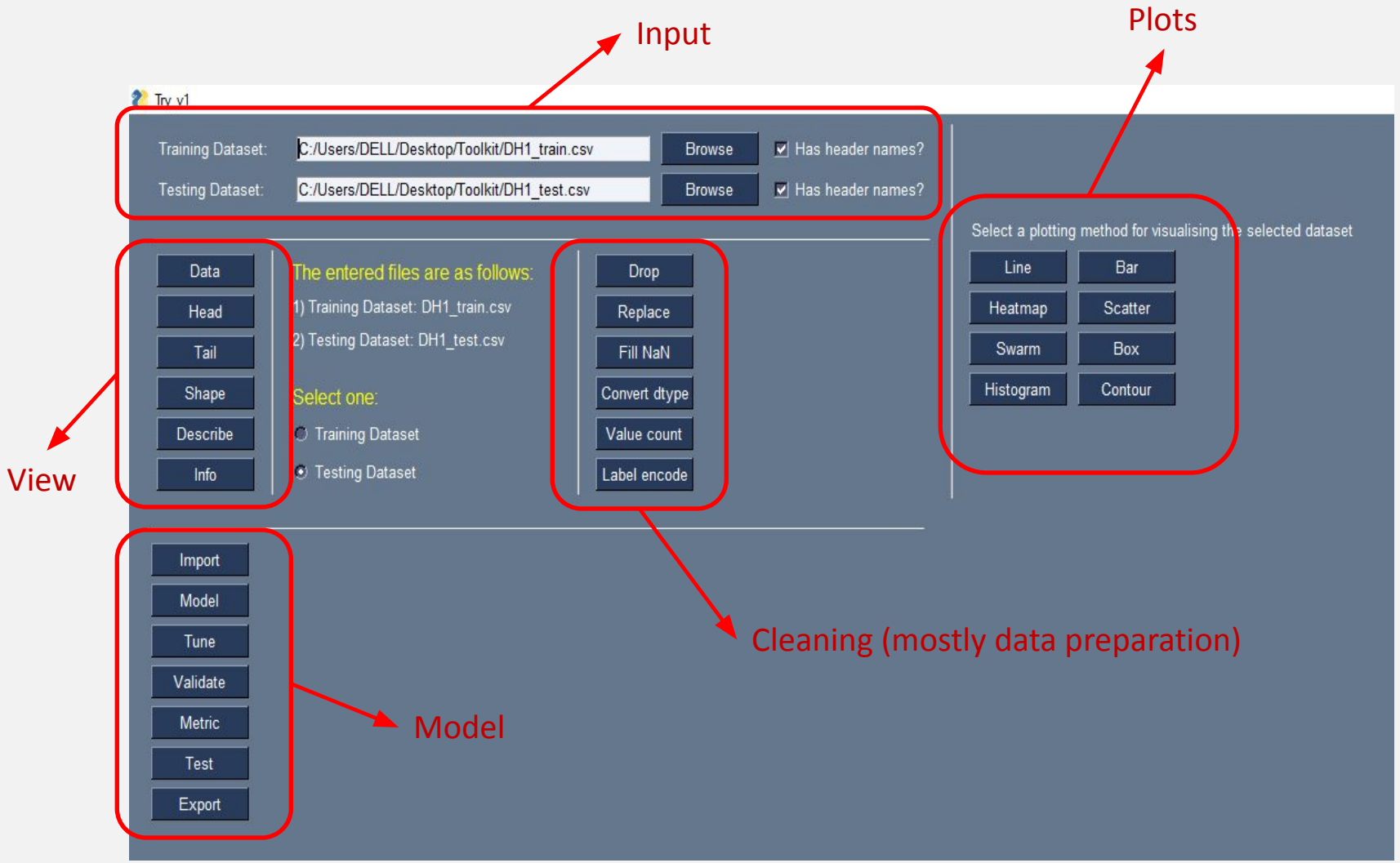
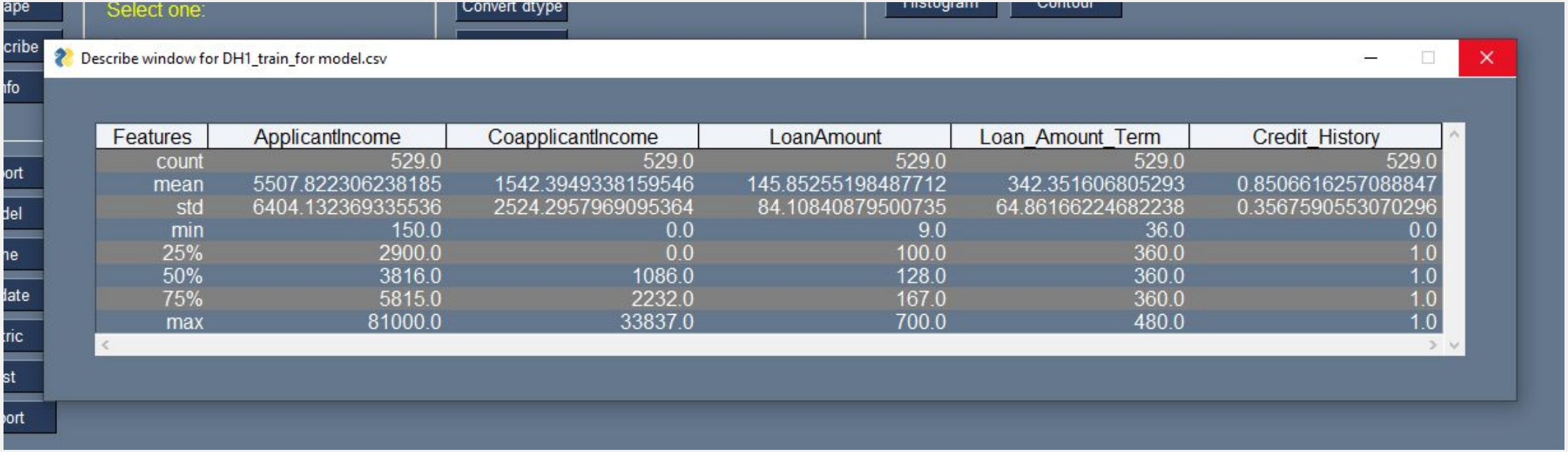
Desktop UI

- 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.

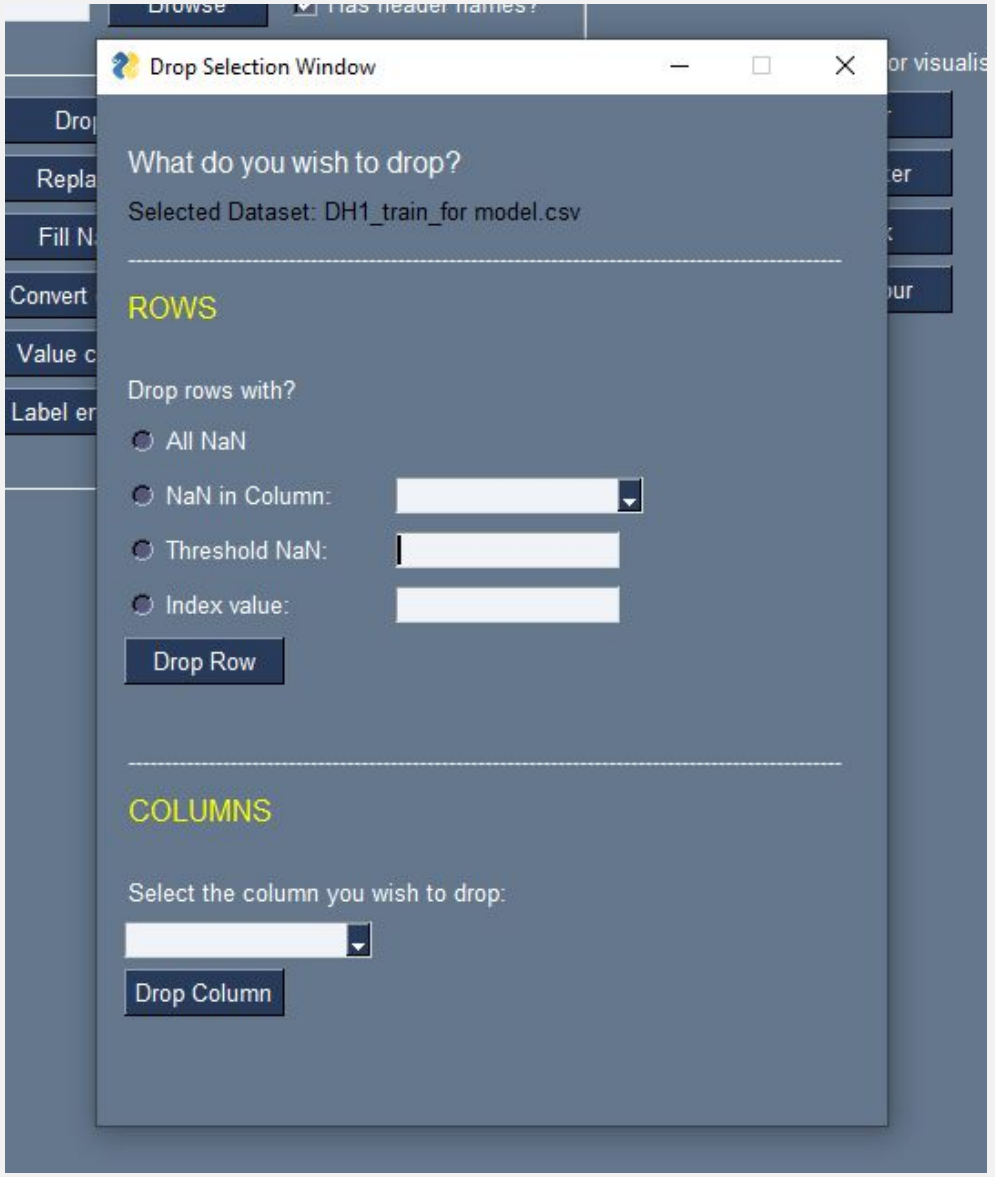


Features	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History
count	529.0	529.0	529.0	529.0	529.0
mean	5507.822306238185	1542.3949338159546	145.85255198487712	342.351606805293	0.8506616257088847
std	6404.132369335536	2524.2957969095364	84.10840879500735	64.86166224682238	0.3567590553070296
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.

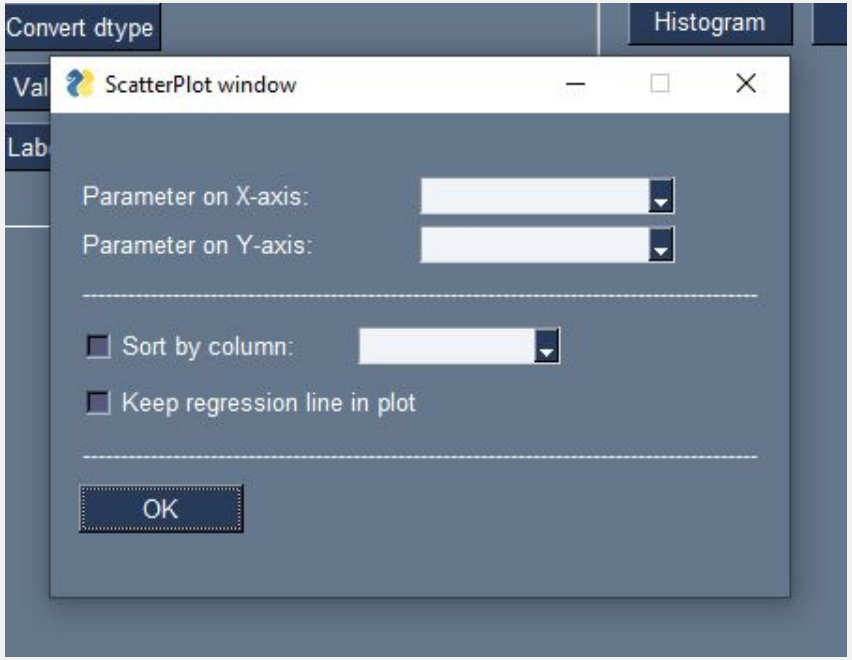
- That drop operation then gets completed.



- 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.

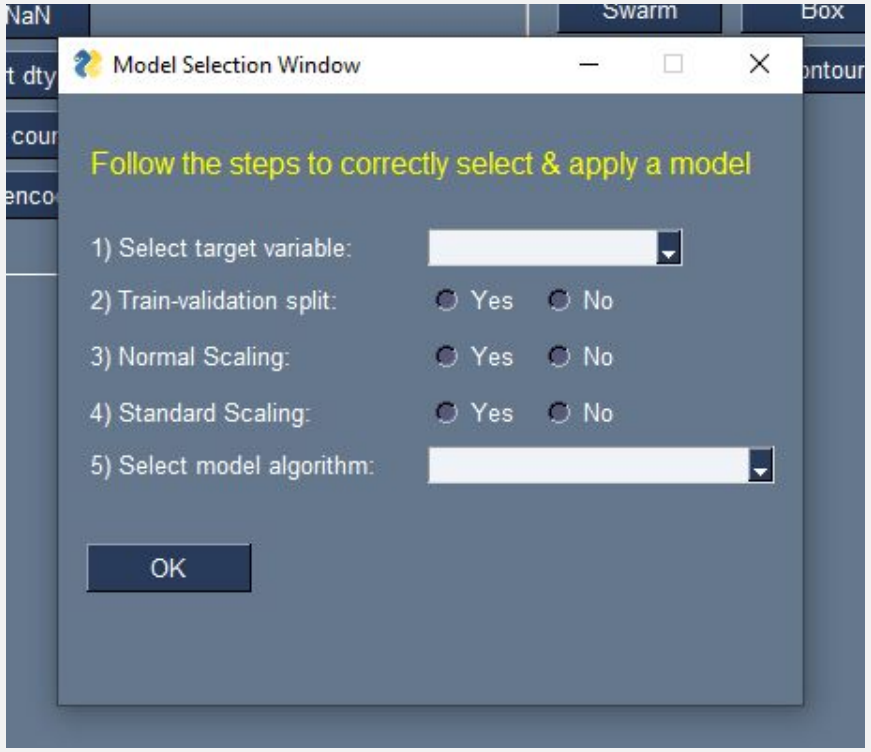
- The plot is then shown in a new chrome tab.



- 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**)

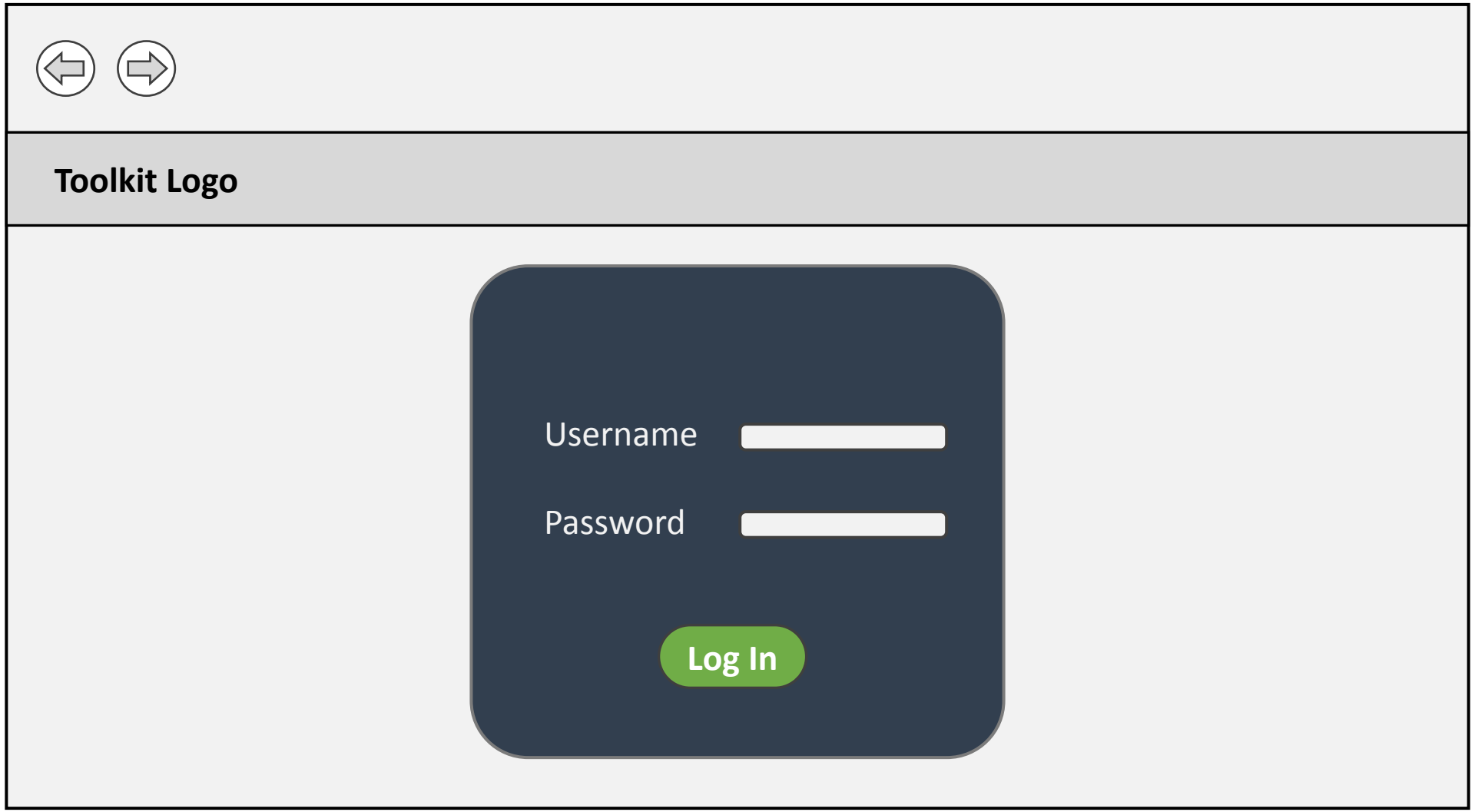
- I can choose suitable options here and click on the button.

- The selected model then gets created.



Web UI

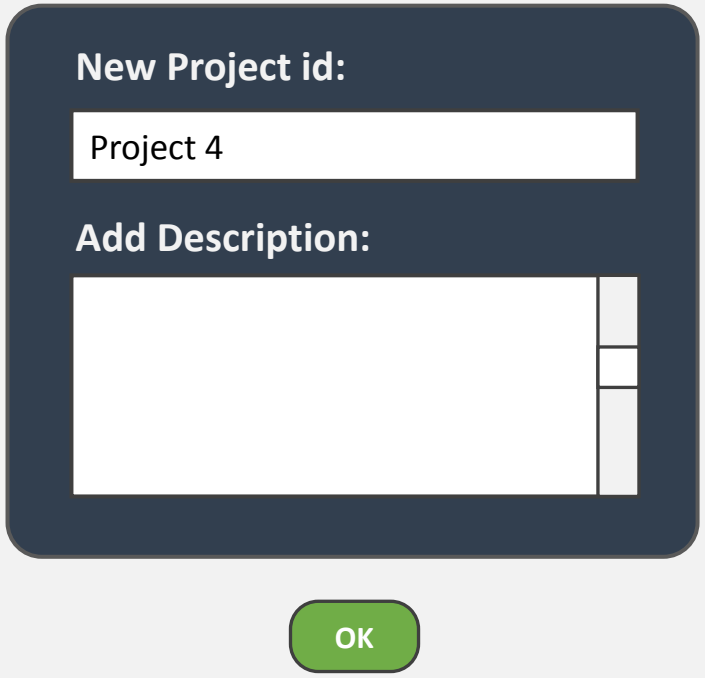
User login window



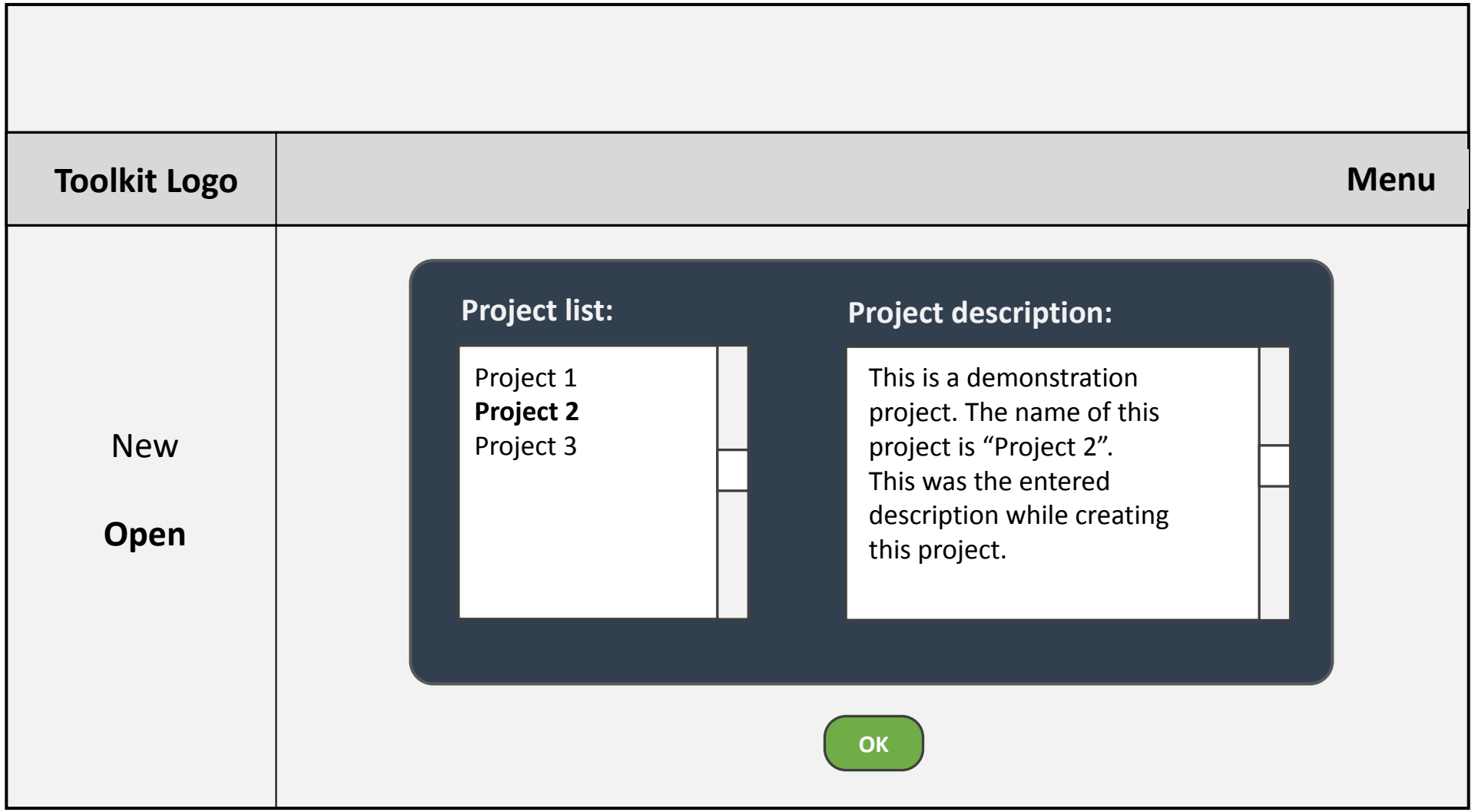
Menu

From this window, this will also act like a button. (when clicked, leads to this Menu window)

Menu (New project)

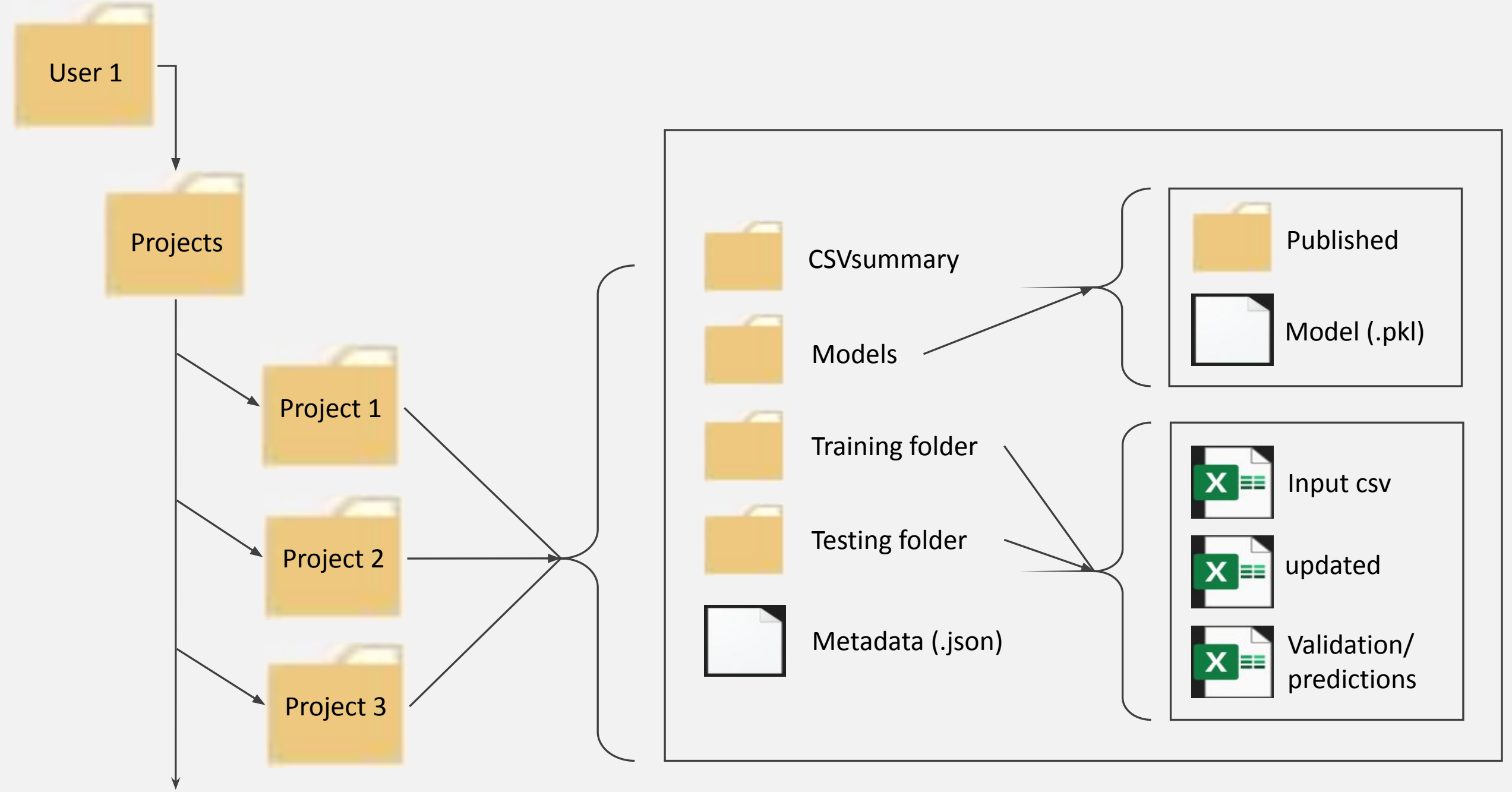
Toolkit Logo	Menu
New Open	

Menu (Open project)

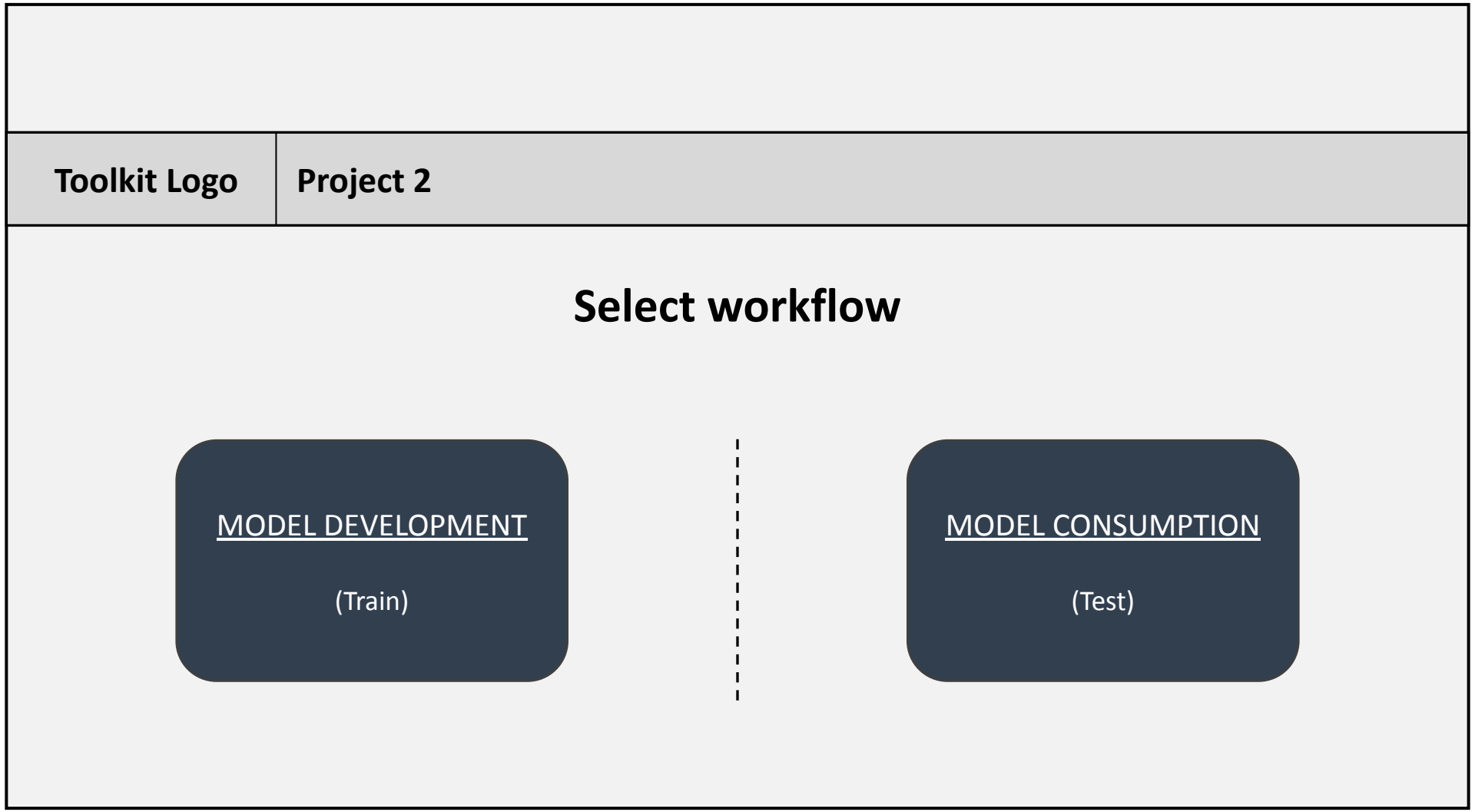


Brief explanation of the folder structure

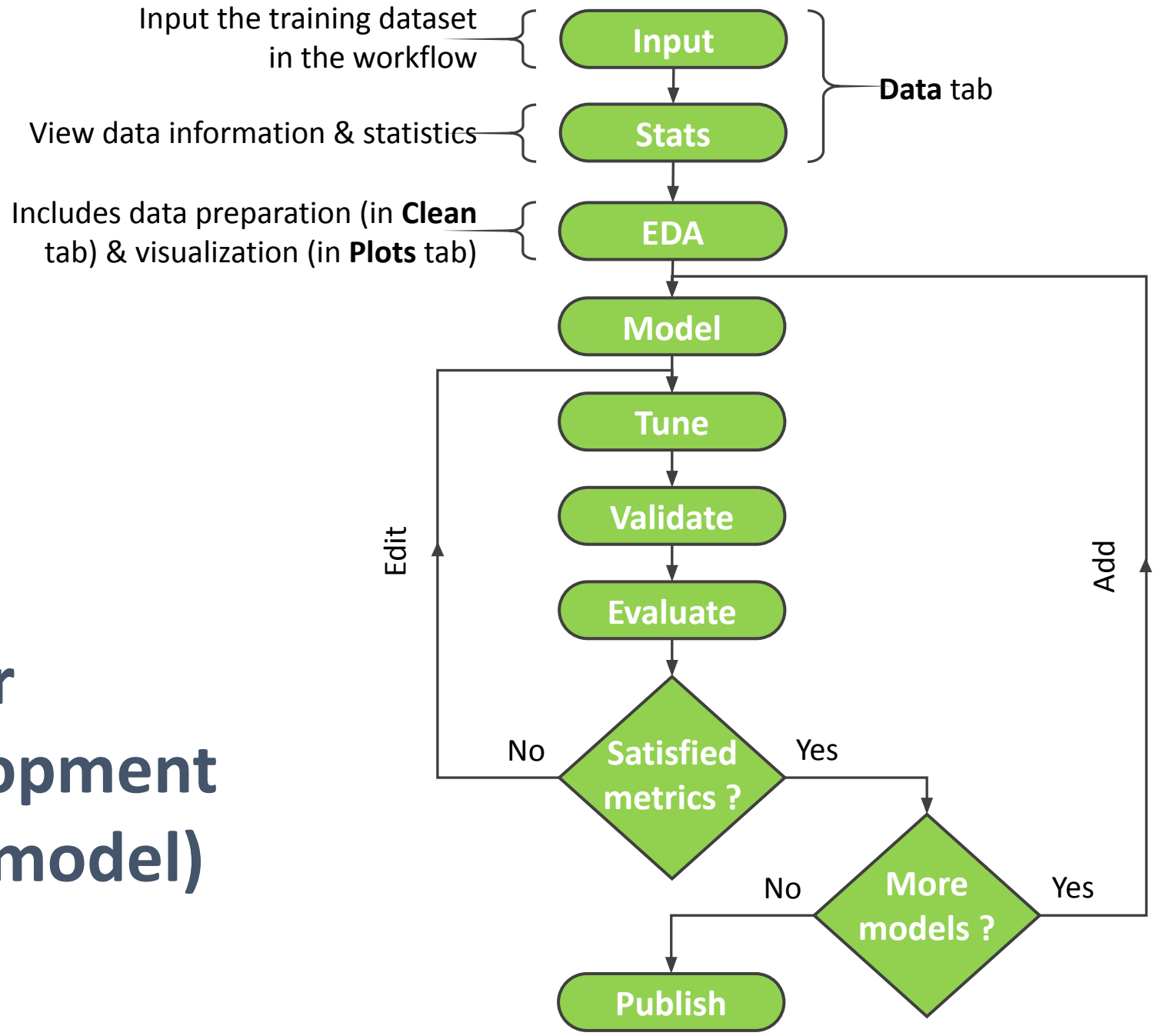
- The folder structure aims at storing (saving) the projects and the performed operations in those projects.
- A major folder, named ‘Projects’ will contain all the projects folders (Project 1, Project 2, Project 3, etc).
- A new project folder (say, Project x), will be automatically created when the 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.

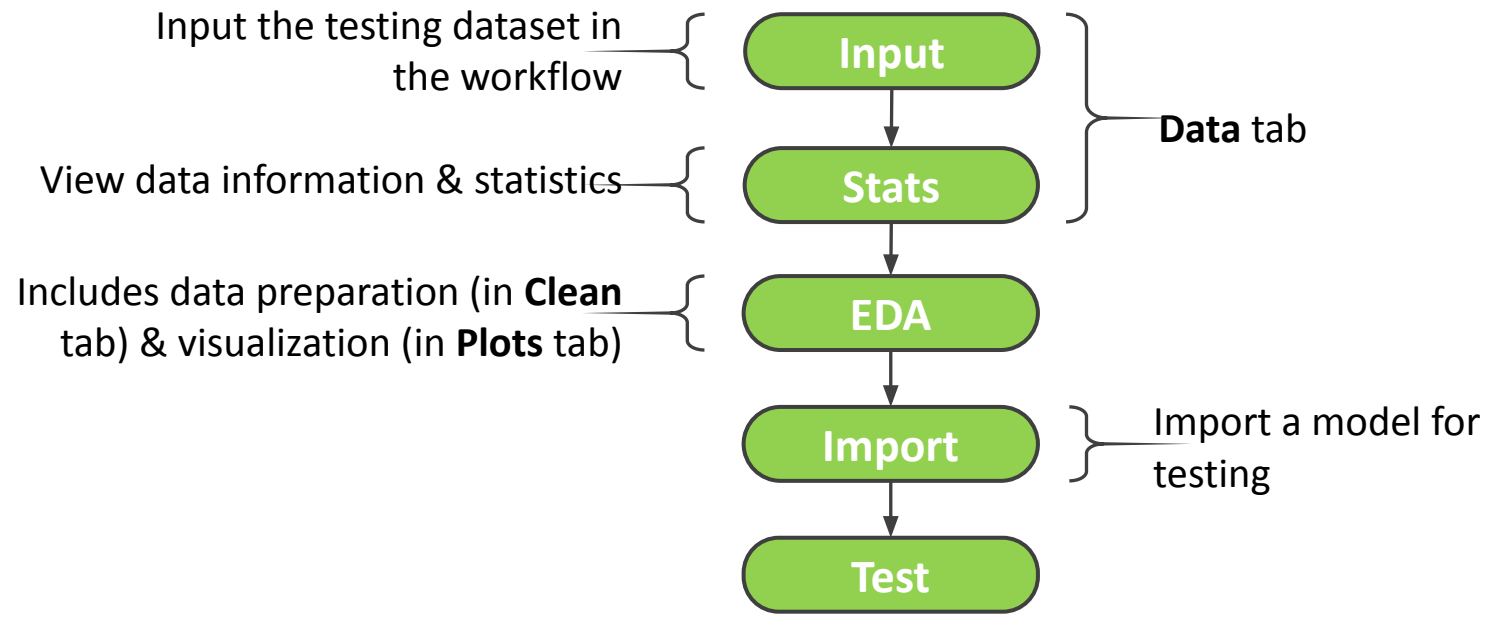

User 1
Projects
Project 1
Project 2
Project 3
CSVsummary
Models
Training folder
Testing folder
Metadata (.json)
Published
Model (.pkl)
Input csv
updated
Validation/
predictions

Workflow Selection window



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

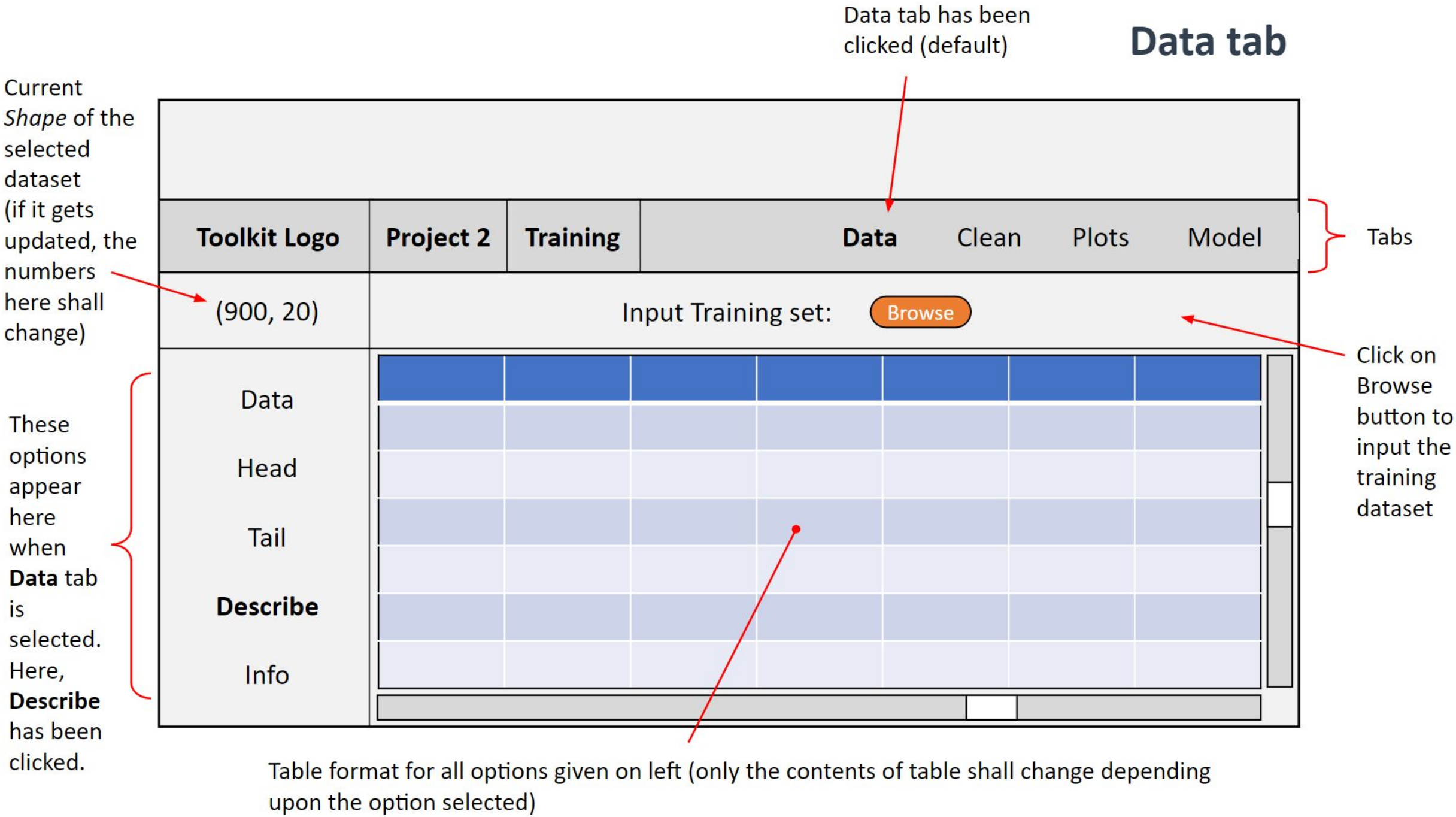




Work-flow for Model-consumption (Import model & test)

Model-development workflow UI

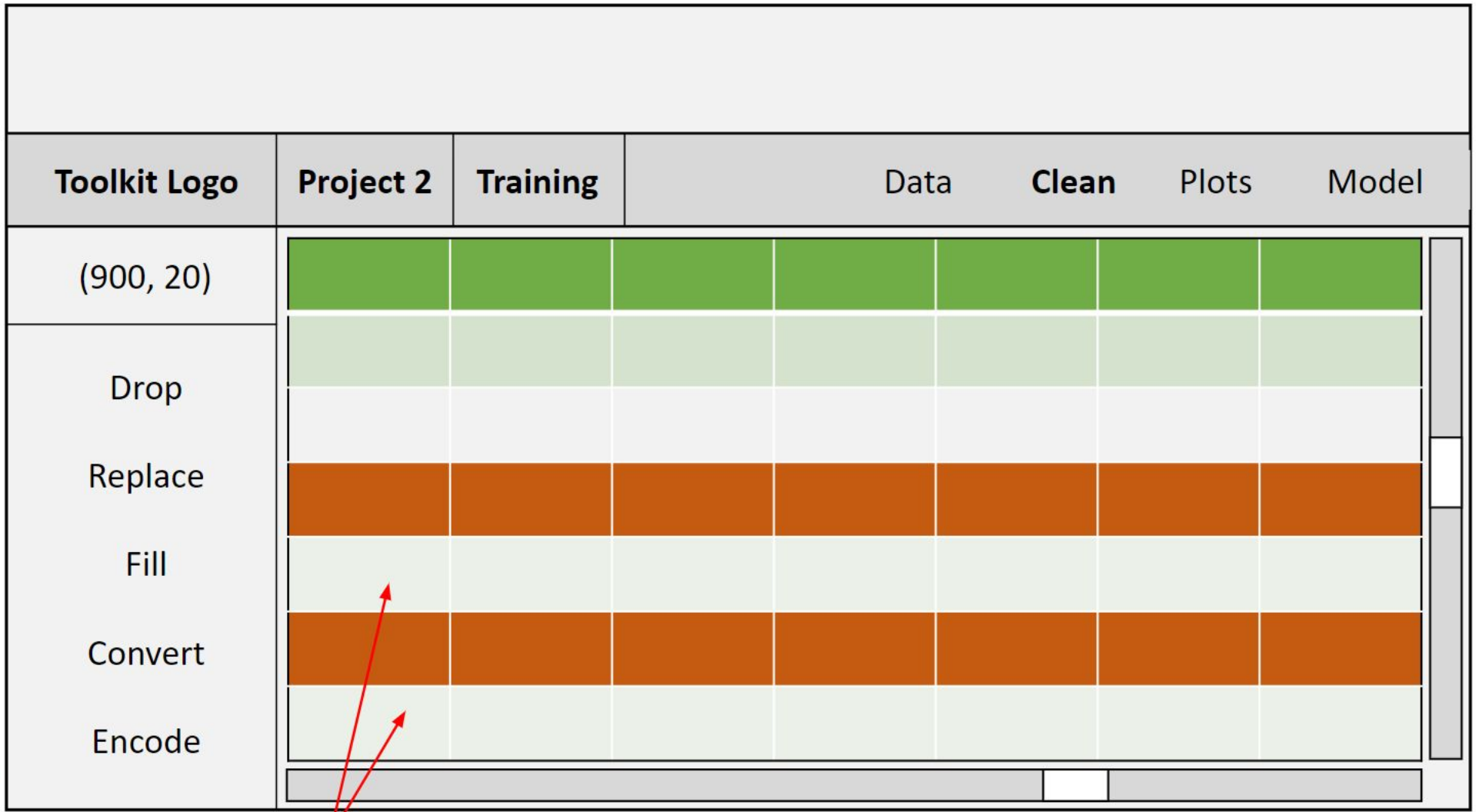
Data tab


Data tab has been clicked (default)
Current *Shape* of the selected dataset (if it gets updated, the numbers here shall change)
Toolkit Logo
Project 2
Training
Data
Clean
Plots
Model
Tabs
(900, 20)
Input Training set:
Browse
Click on Browse button to input the training dataset
Data
Head
Tail
Describe
Info
These options appear here when **Data** tab is selected. Here, **Describe** has been clicked.
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



Clean options when selected opens a specific tab at the same place

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.

Options in bold indicate that they were selected earlier. Tab shown her only for representation purpose

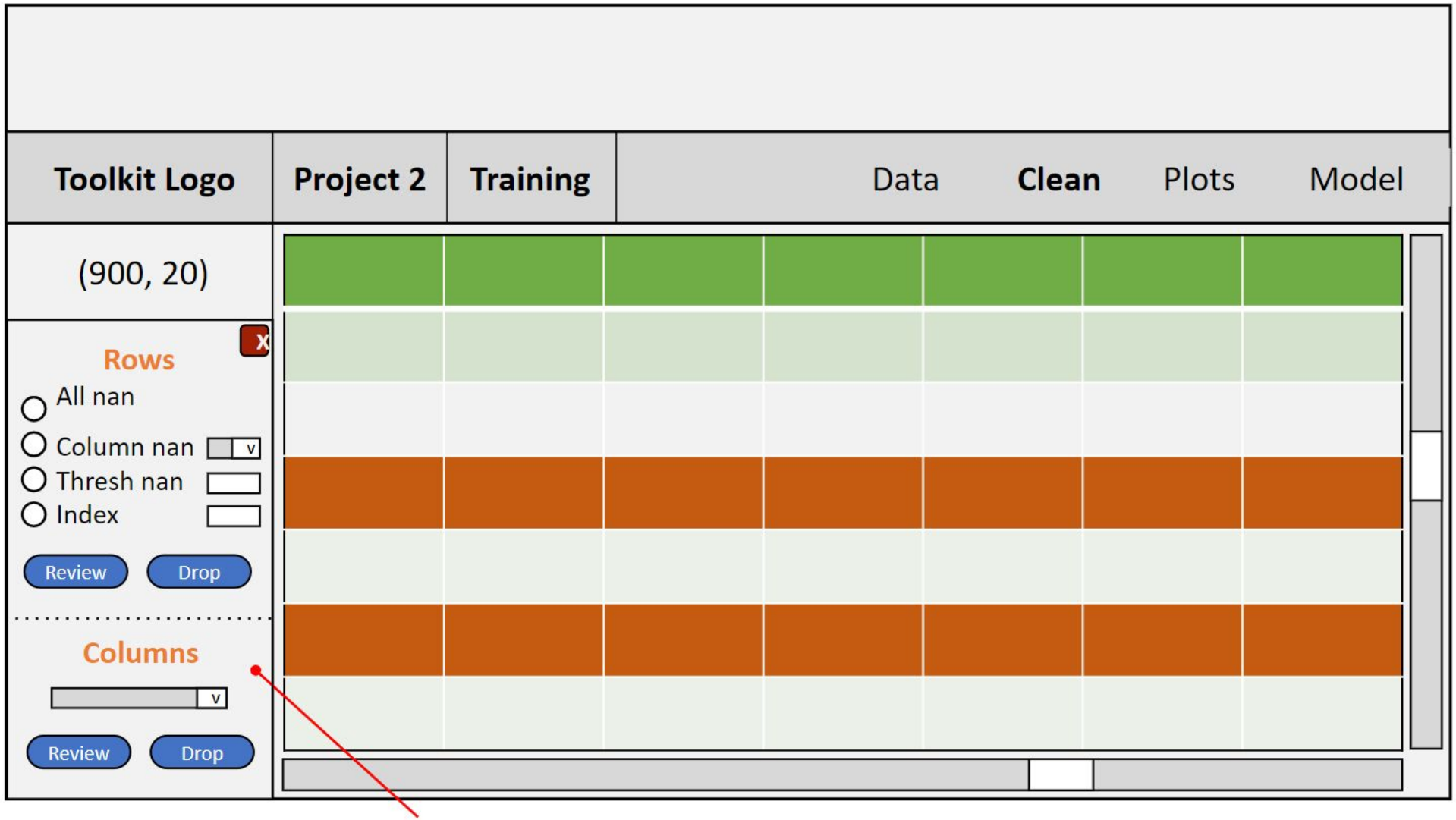
Drop

Replace

Fill

Convert

Encode



Drop option was selected earlier, hence **this particular** tab has opened here. (Select the operations of choice & hit the respective blue button, the corresponding changes will be reflected in the table alongside)

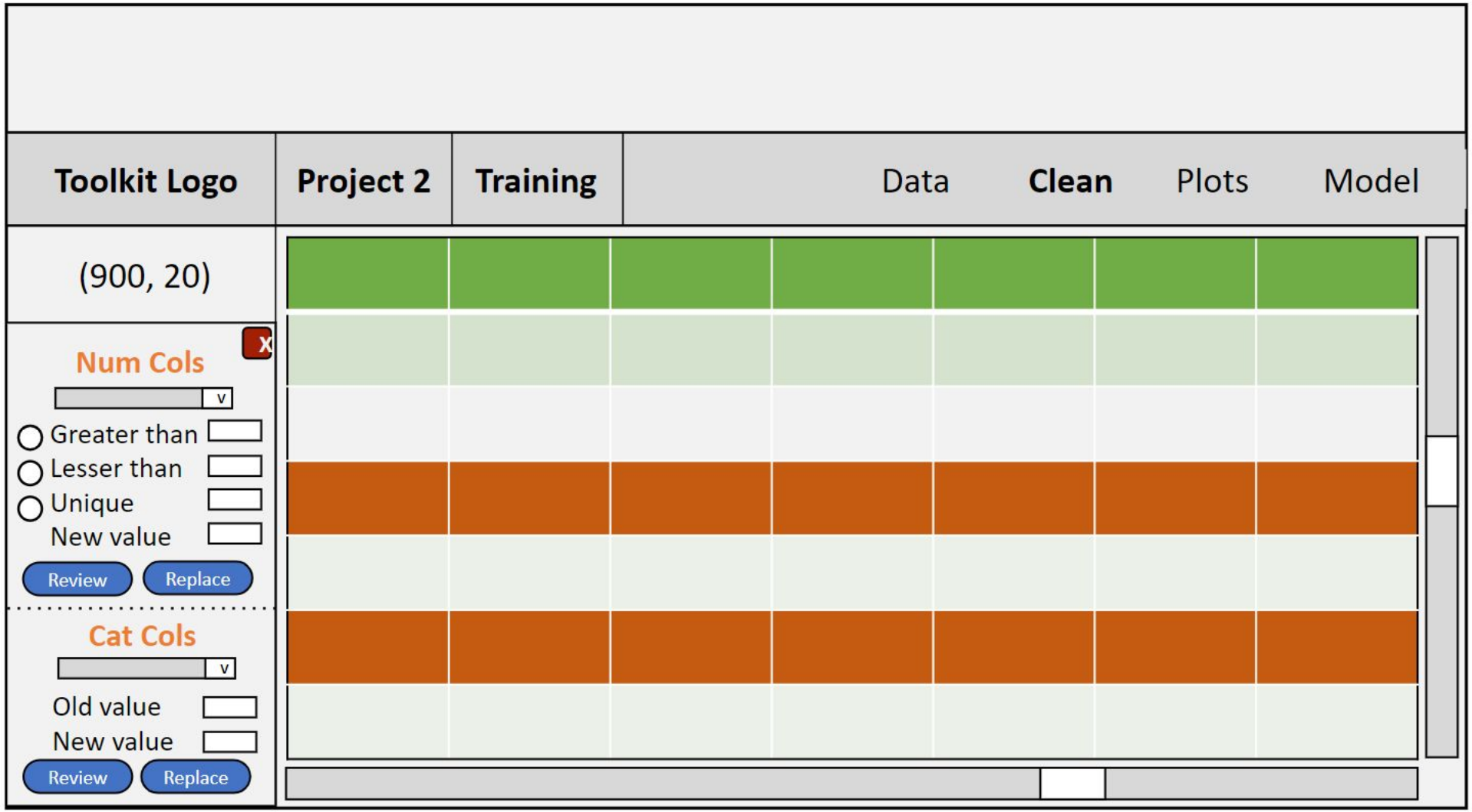
Drop

Replace

Fill

Convert

Encode



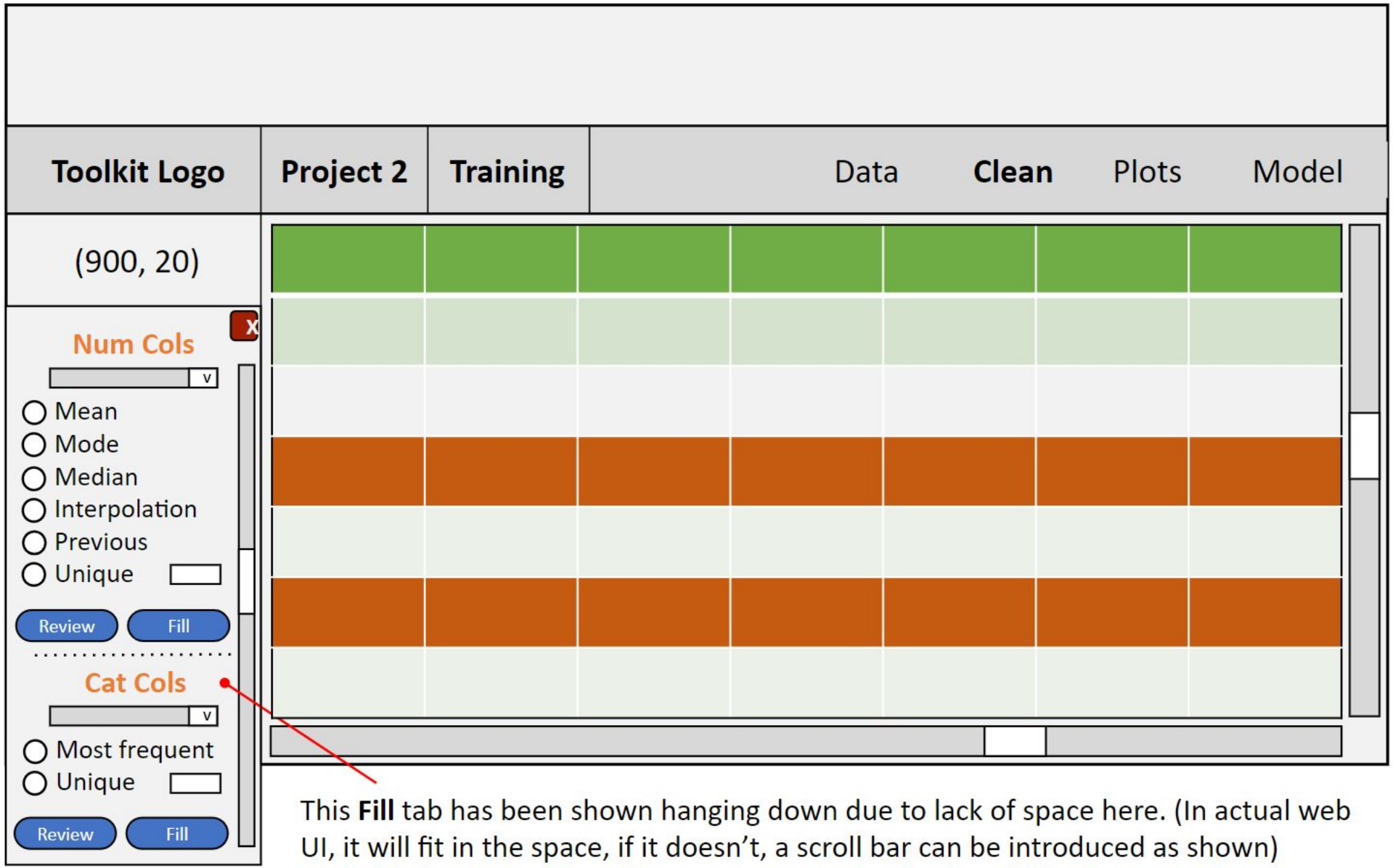
Drop

Replace

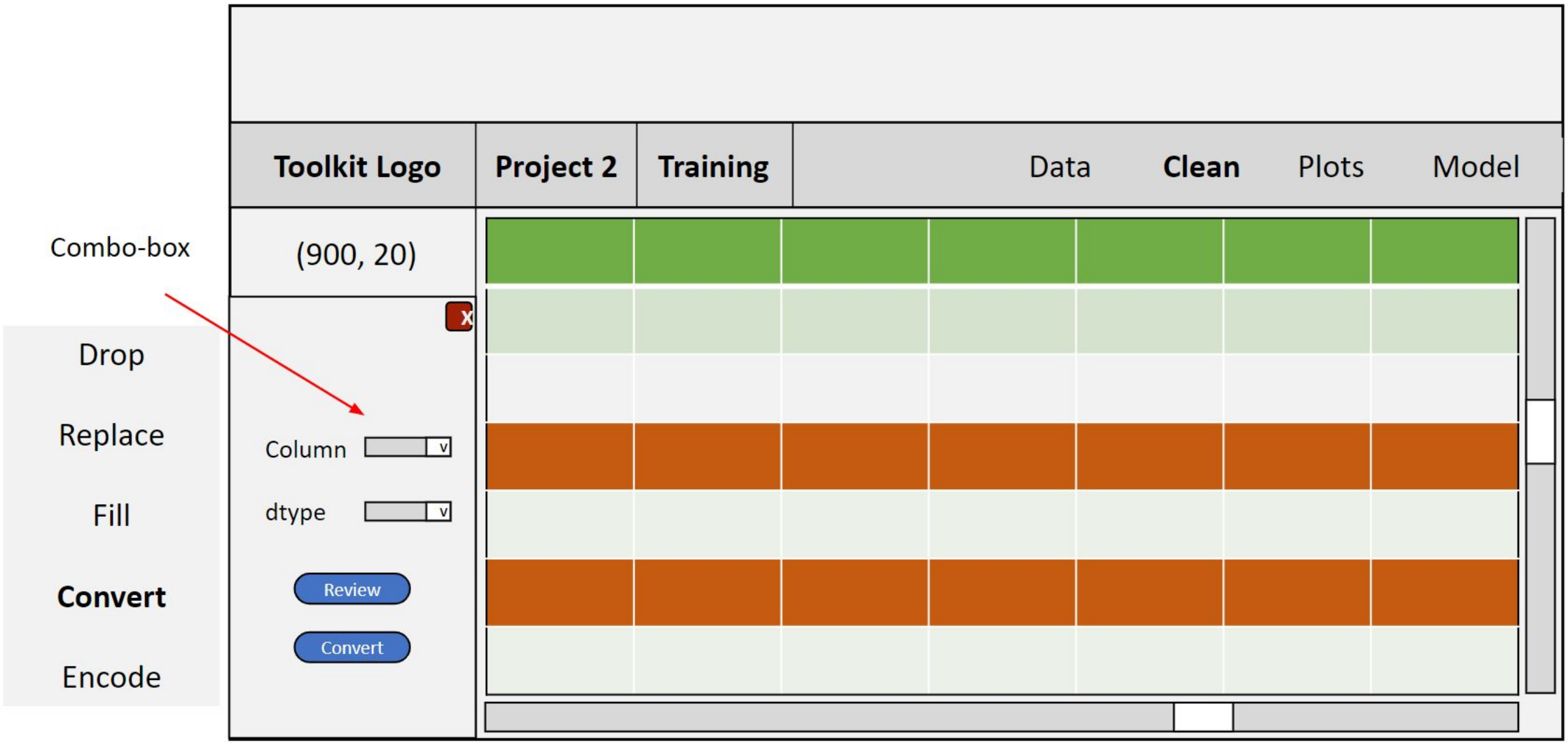
Fill

Convert

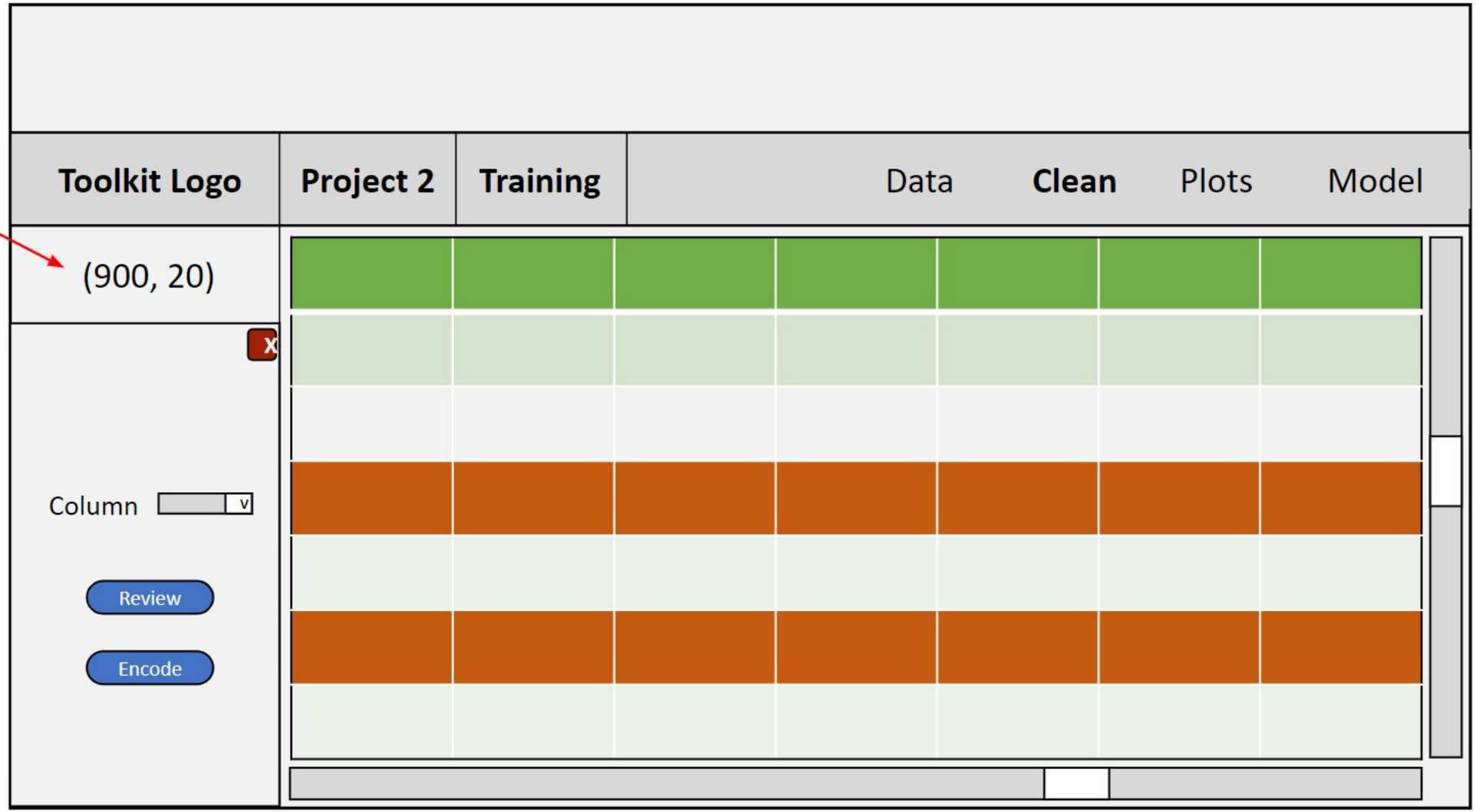
Encode



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)


Toolkit Logo
Project 2
Training
Data
Clean
Plots
Model
(900, 20)
Combo-box
Drop
Replace
Fill
Convert
Encode
Column
dtype
Review
Convert

Note that the *shape* will be updated after performing any of the operations under **EDA** tab



Drop

Replace

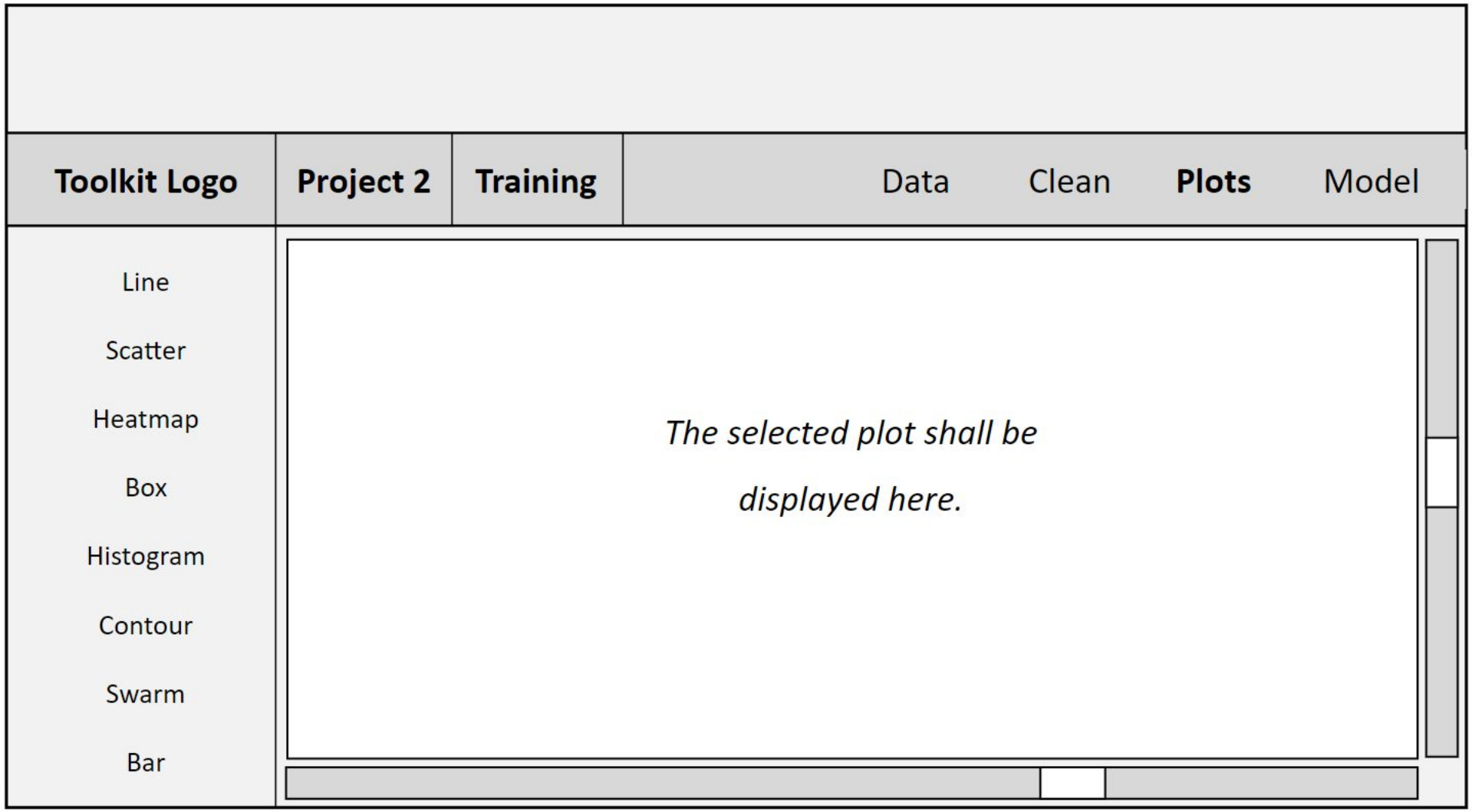
Fill

Convert

Encode

Plots tab

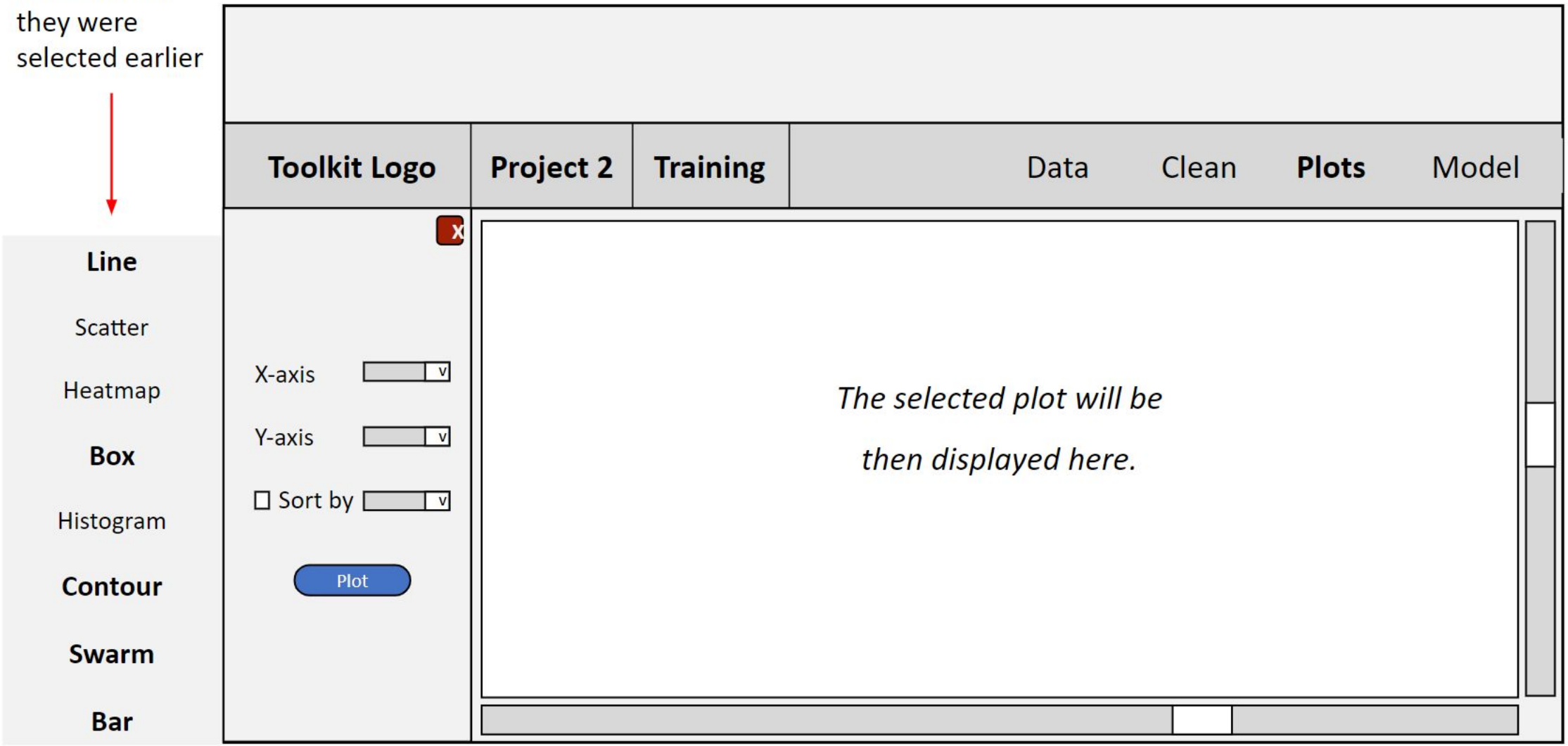
These options appear here when **Plots** tab is selected. Note that there is no *shape* display under this tab

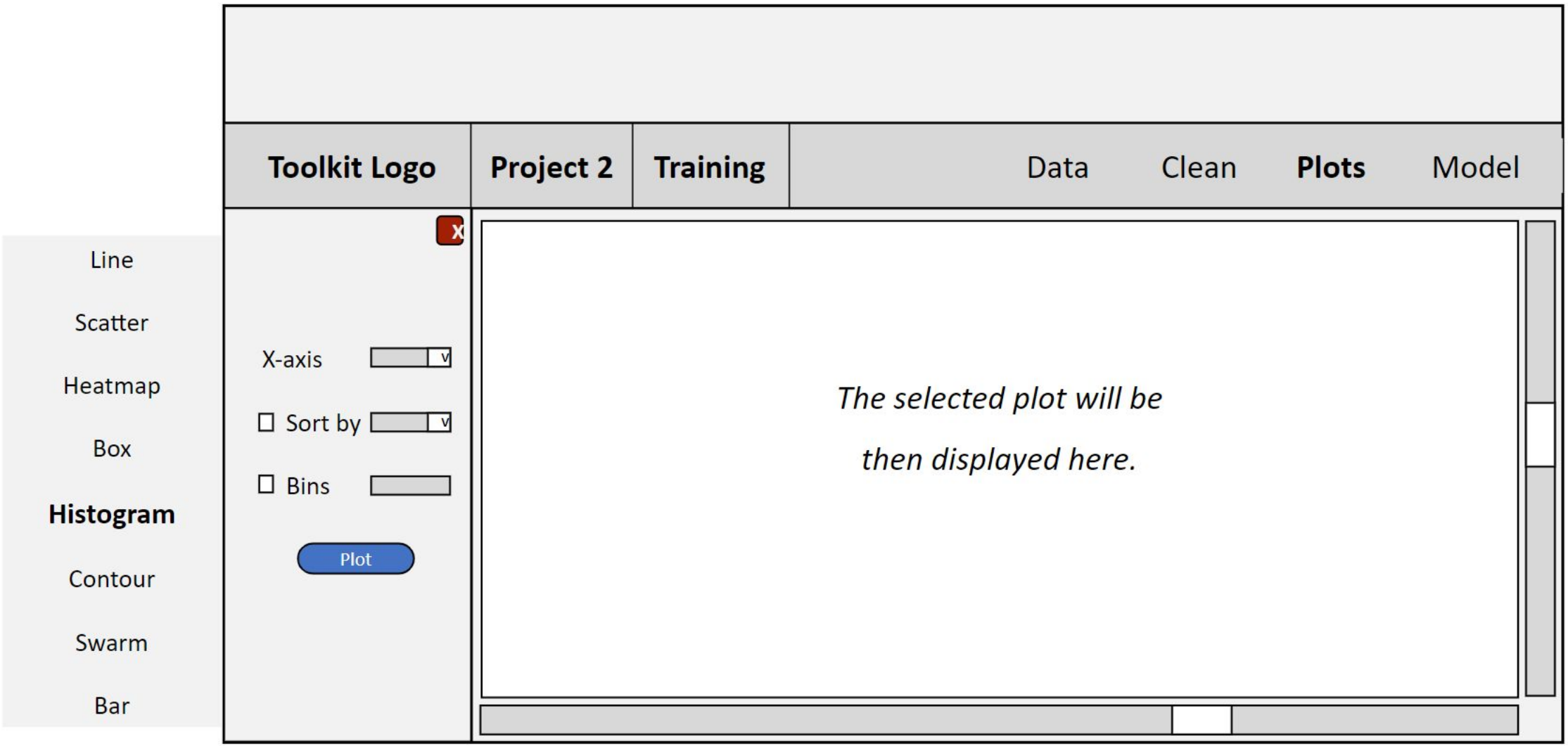


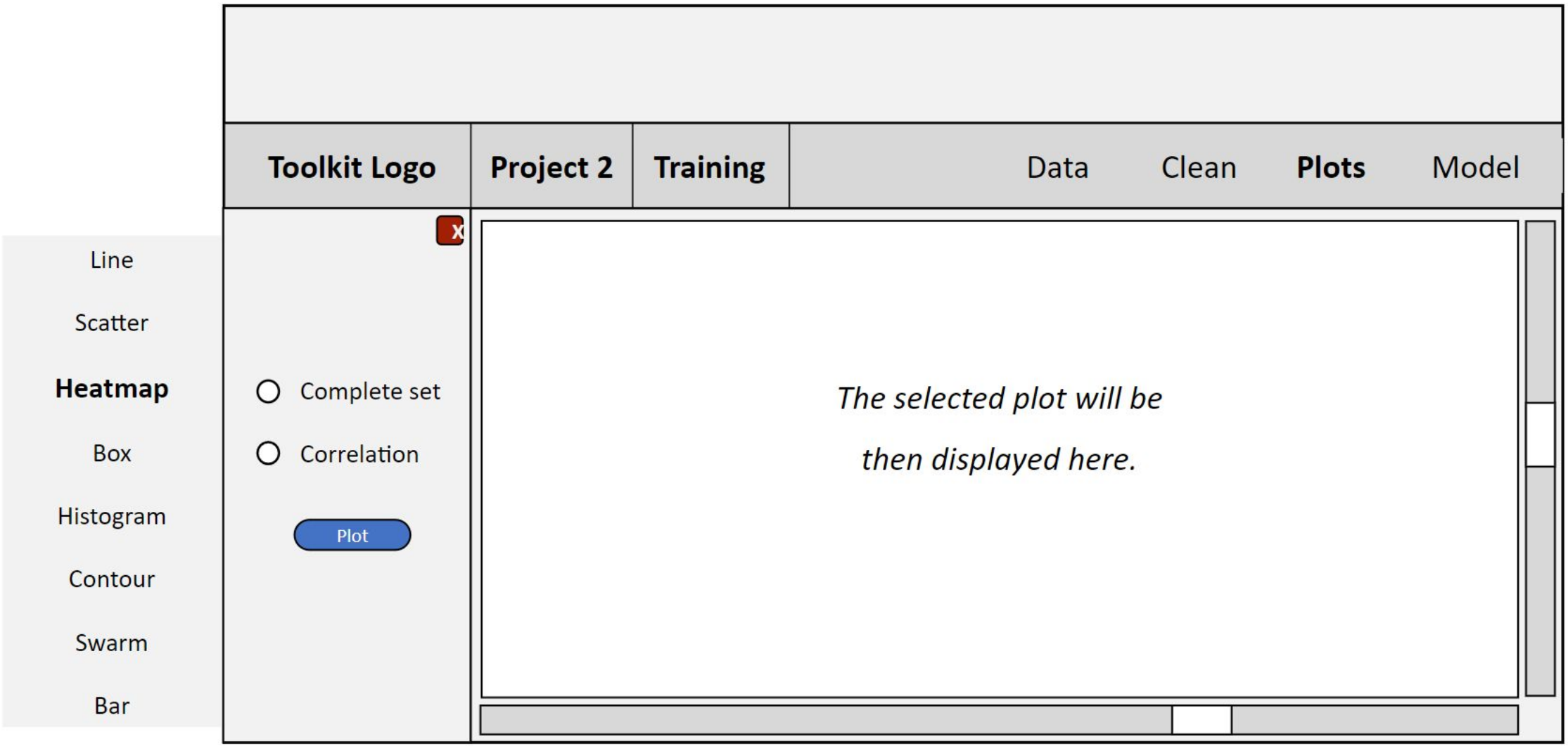
Brief explanation of the previous slide(Plots tab)

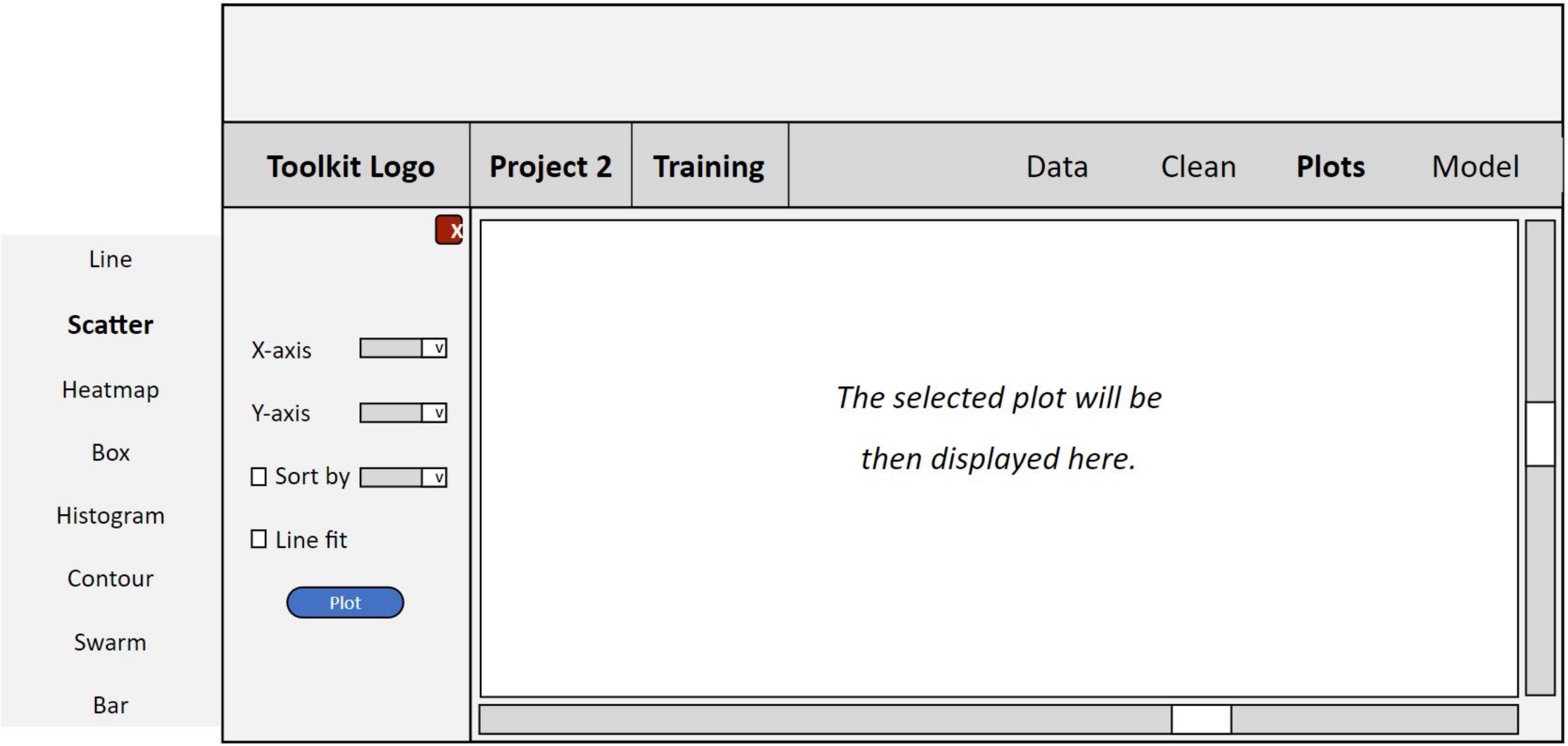
- The main objective of this tab is data visualization (plots, heatmap, histogram, etc)
- The options for types of plots are at left. Any of the option when clicked, opens a sub-window (related to the clicked option) at the same place. This is clearly shown in the upcoming slides.
- The user can choose 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.

Options in bold indicate that they were selected earlier

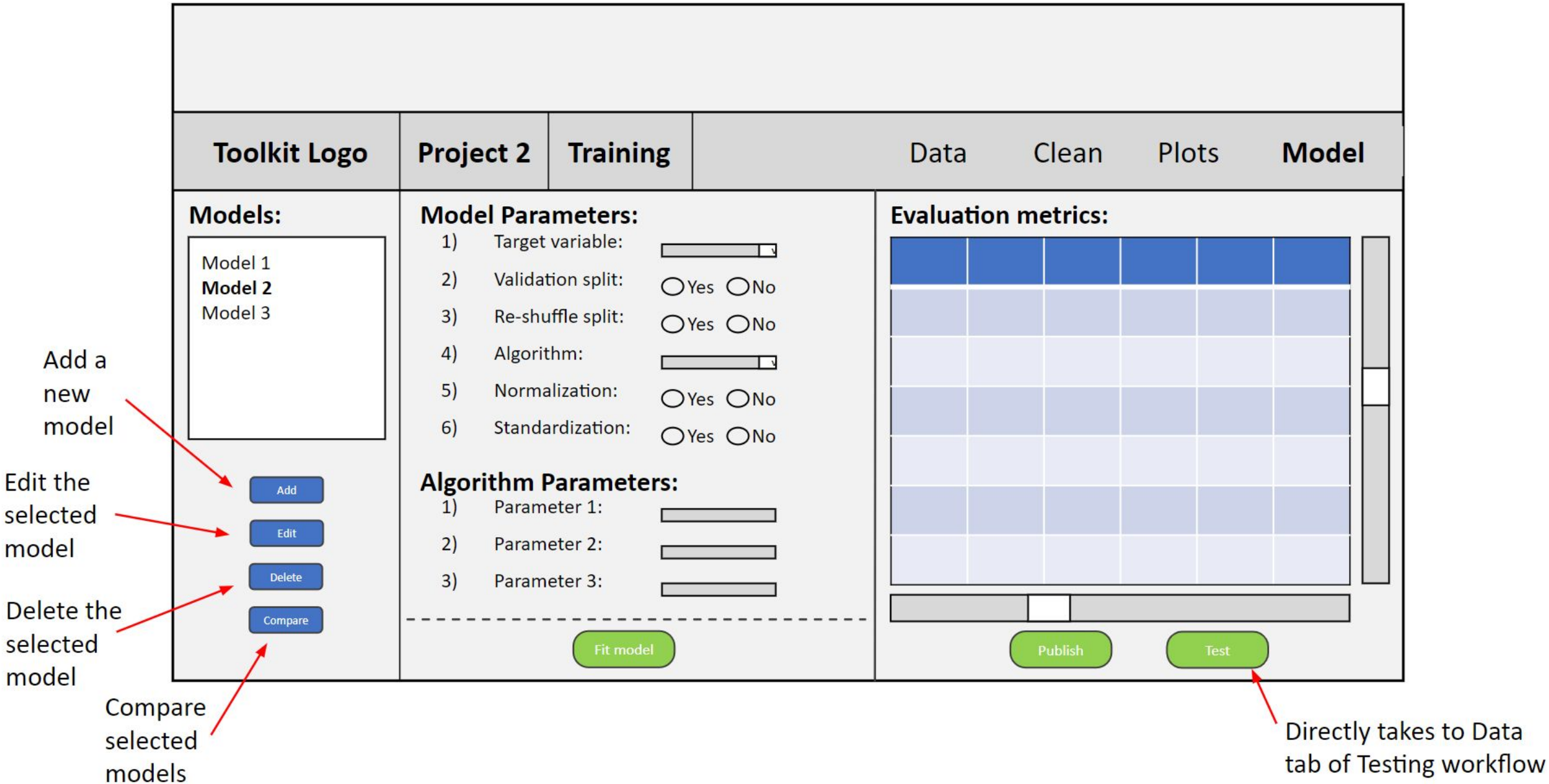



Toolkit Logo
Project 2
Training
Data
Clean
Plots
Model
X
Line
Scatter
Heatmap
Box
Histogram
Contour
Swarm
Bar
X-axis
Sort by
Bins
Plot
The selected plot will be
then displayed here.


Toolkit Logo
Project 2
Training
Data
Clean
Plots
Model
X
Line
Scatter
Heatmap
Box
Histogram
Contour
Swarm
Bar
Complete set
Correlation
Plot
The selected plot will be
then displayed here.


Toolkit Logo
Project 2
Training
Data
Clean
Plots
Model
X
Line
Scatter
Heatmap
Box
Histogram
Contour
Swarm
Bar
X-axis
Y-axis
Sort by
Line fit
Plot
The selected plot will be
then displayed here.

Model tab


Toolkit Logo
Project 2
Training
Data
Clean
Plots
Model
Models:
Model 1
Model 2
Model 3
Add
Edit
Delete
Compare
Model Parameters:
1) Target variable:
2) Validation split: Yes No
3) Re-shuffle split: Yes No
4) Algorithm:
5) Normalization: Yes No
6) Standardization: Yes No
Algorithm Parameters:
1) Parameter 1:
2) Parameter 2:
3) Parameter 3:
Fit model
Evaluation metrics:
Publish
Test
Add a new model
Edit the selected model
Delete the selected model
Compare selected models
Directly takes to Data tab of Testing workflow

Brief explanation of the previous slide(Model tab)

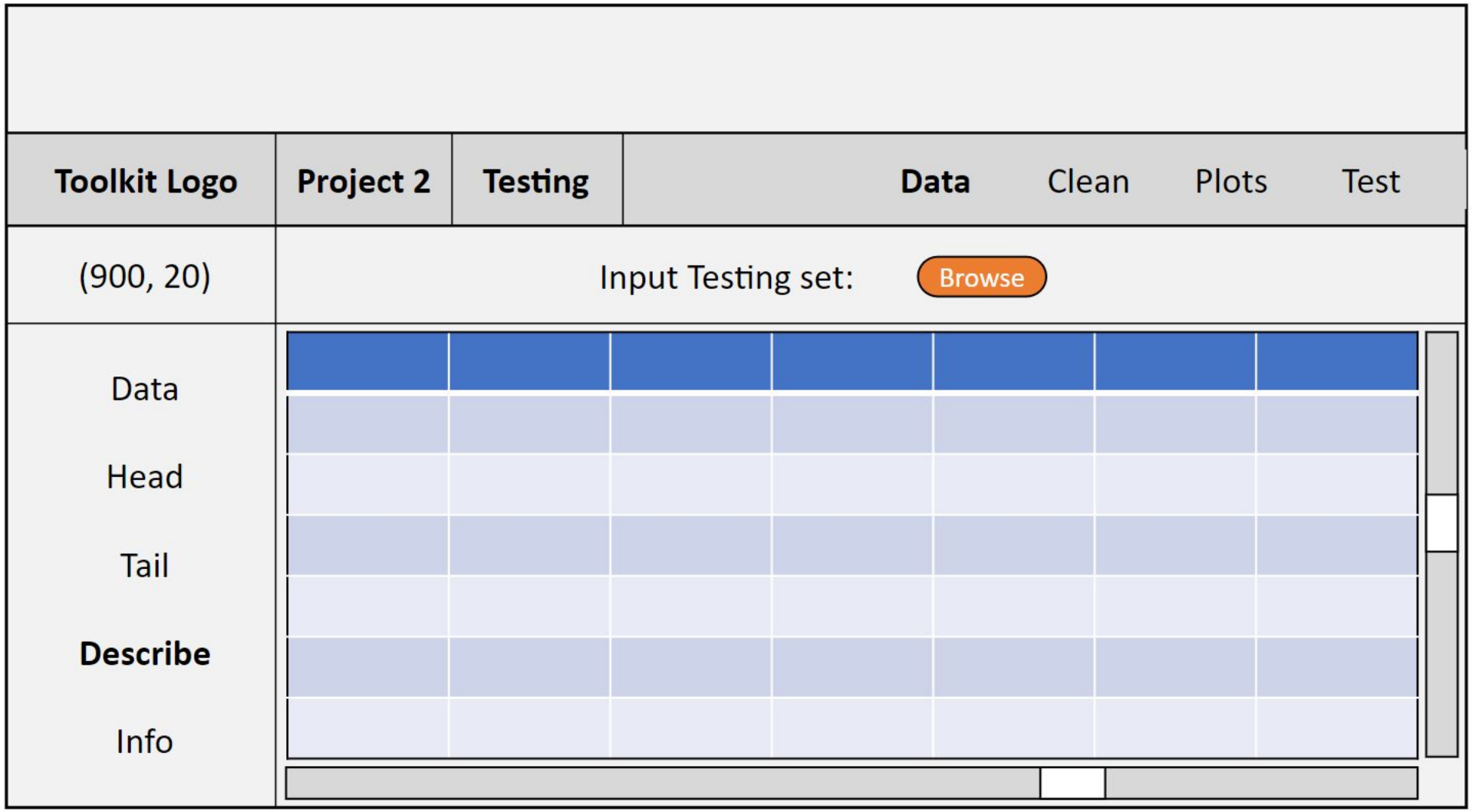
- The main objective of this tab is model building, tuning, validation, evaluation (using metrics) & publish (export model).
- In Model parameters part, the selection of options is inter-related. For e.g. if the user selects Yes for Validation-split option, only then the options for Re-shuffle split 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
Data
Clean
Plots
Test
(900, 20)
Input Testing set:
Browse
Data
Head
Tail
Describe
Info

Clean tab

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

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)

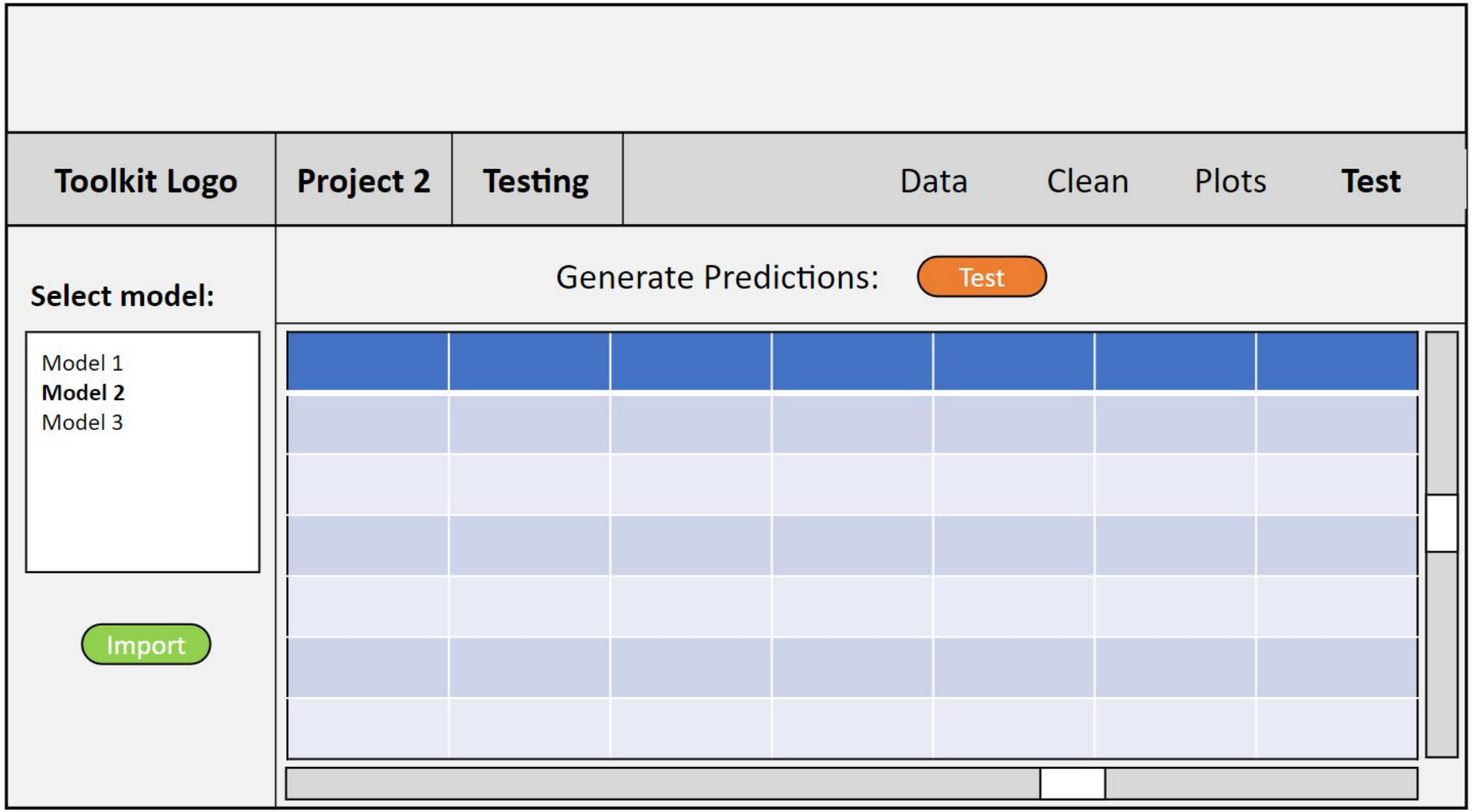
Plots tab

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)

Test tab


Toolkit Logo
Project 2
Testing
Data
Clean
Plots
Test
Select model:
Generate Predictions:
Test
Model 1
Model 2
Model 3
Import