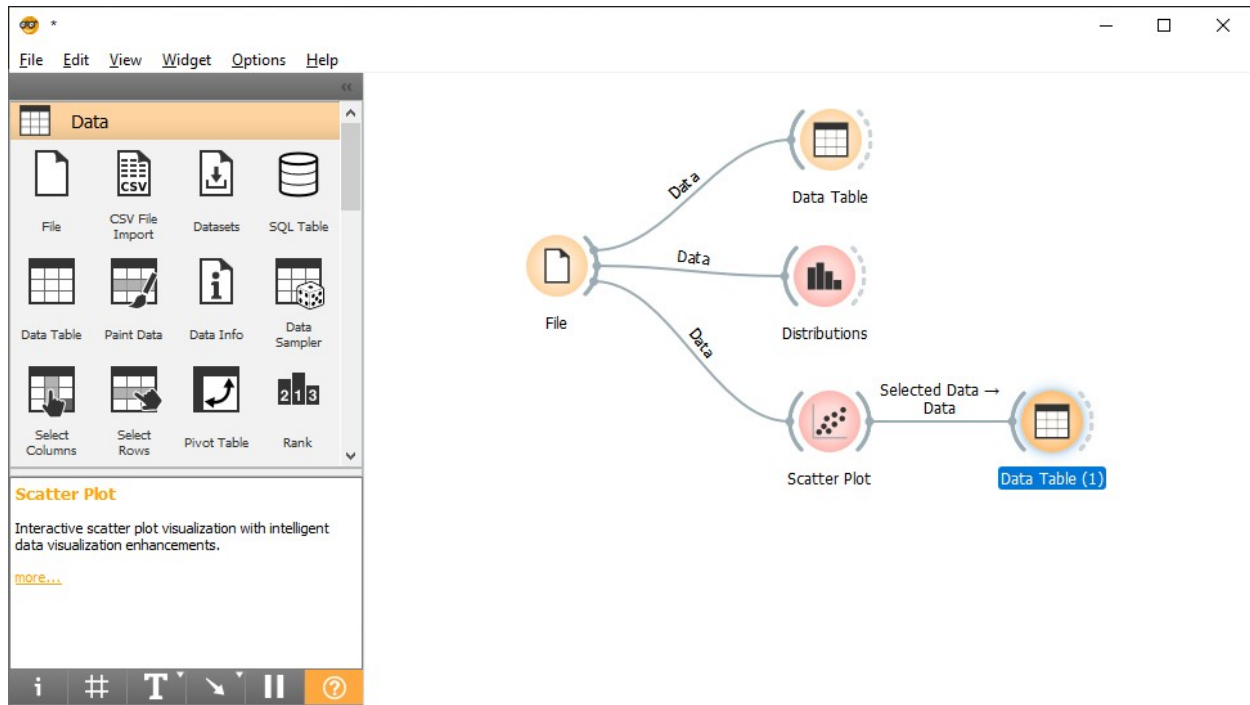


Lesson 2: Data Workflows

<https://www.youtube.com/watch?v=lb-x36xqJ-E>



Exercise: In the Distribution widget, set the “Bin Width” to smallest setting. Looking at the distributions for each variable, which variable gives the best spread between the types of irises?

Exercise: In the Scatter Plot widget, how do think “Find Informative Projections” is scoring pairs of variables. Try using the Paint Data widget connected to the Scatter Plot widget to test your theory.

My guess is that they are using VizRank: <https://file.biolab.si/biolab/supp/bi-vizrank/details.html>. We will talk about this later.

(Note: They let their certificate expire. If want to see this webpage, you have to add an exception.)