



# Interactive Scatterplots

Michel Wermelinger  
The Open University, UK



<http://tiny.cc/iScatter>

[michel.wermelinger@open.ac.uk](mailto:michel.wermelinger@open.ac.uk)

## Motivation

Data charts are everywhere, but students and the public can't easily explore and check the presented data and statistics because most charts are static. Even when interactive, like Google Charts, they often just allow to pan, zoom, and inspect the x/y values of a data point.

The iScatter tool allows to create highly interactive **web-based scatterplots** and **bubble charts** (circle area shows a third variable), two chart types widely used in STEM. The subject-agnostic example below shows how to **explore a dataset with a few clicks**, without the learning curve of professional statistical software. iScatter has more features than described here: logarithmic scales, pan, zoom, etc.

iScatter is **freely available** from the URL above, which also provides an interactive guided tour, and details on how to define your datasets with simple spreadsheets and how to configure your own charts.

## Main features

Flexible, multivariate, 'semantic' **data model**

- metadata describes variables and measurement levels, e.g. ratio
- model based on overlapping subsets, *not* disjoint data series ⑥
- iScatter defines one subset per distinct value of a variable ④
- you select individual points to define a custom subset ⑧

All chart elements are **interactive**, all changes are **animated** ⑧

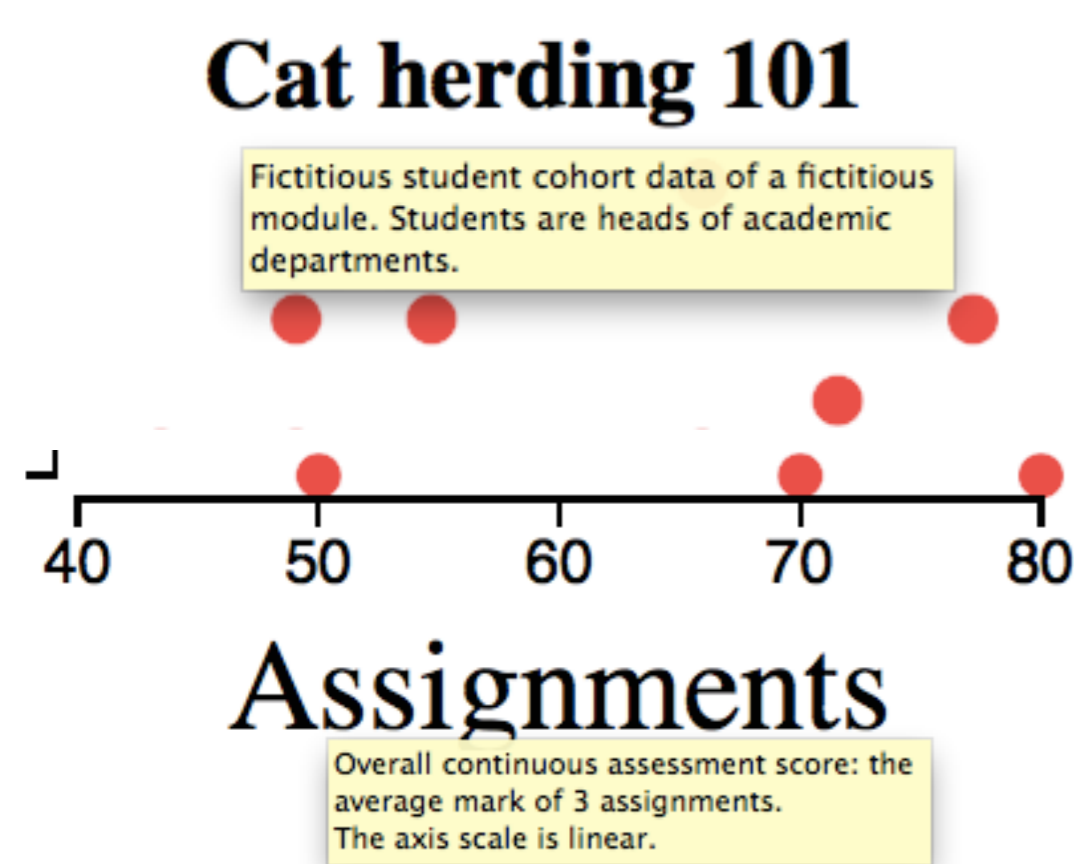
- hover for information on demand, without cluttering chart ①④⑥⑦
- click to get pop-up menu or dialog to change chart ②③④⑤

**Statistics** are computed per subset and per variable ④

- up to 11 basic stats: median, mean, standard deviation, etc. ④
- metadata prevents meaningless stats like mean of ordinal variable
- statistics automatically updated when chart changes ⑧
- statistics shown visually with lines, to ease comparison ⑤

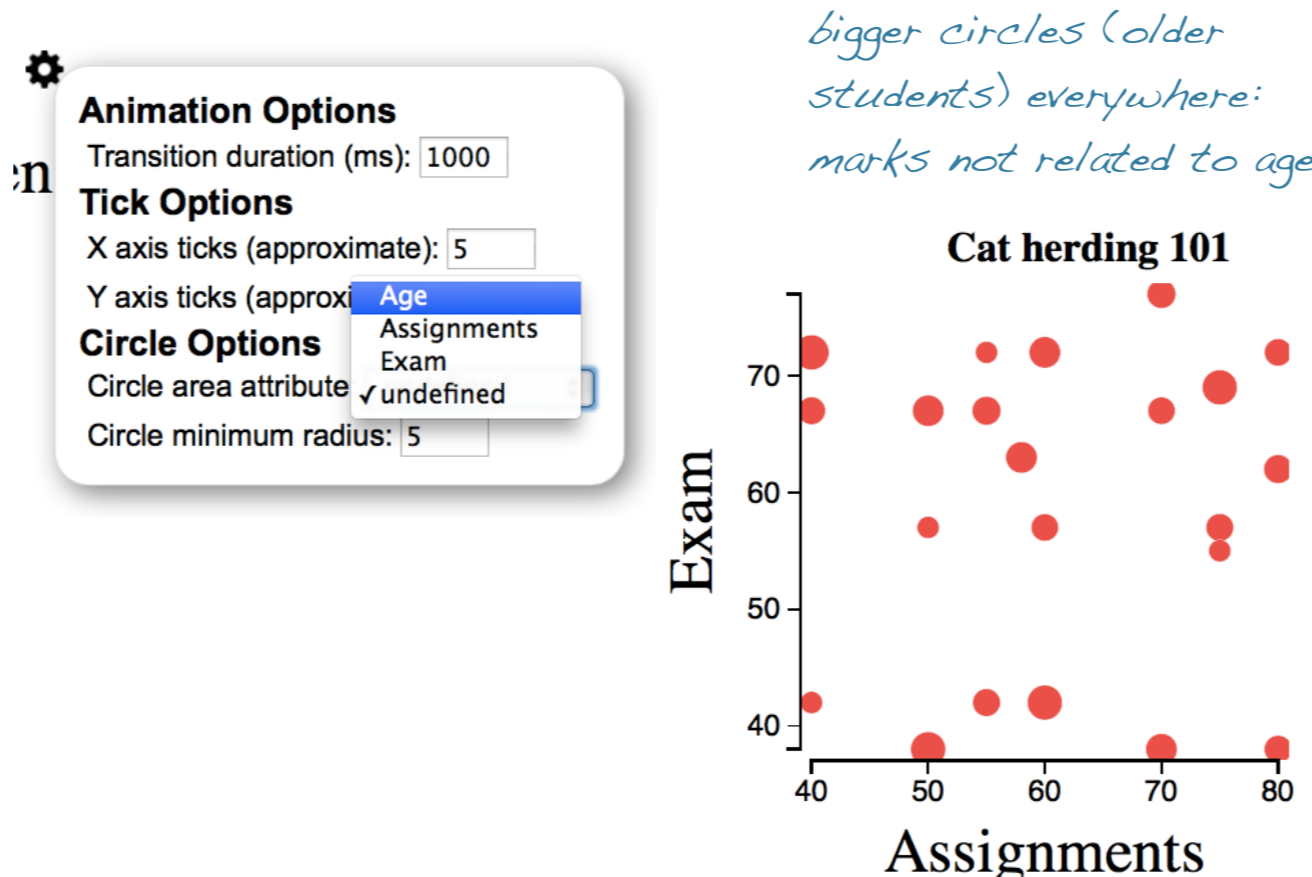
### ① How was the dataset obtained?

Hover the mouse on the chart/axis titles for descriptions of the data set and variables.  
*NB: screen shots don't show mouse pointer.*



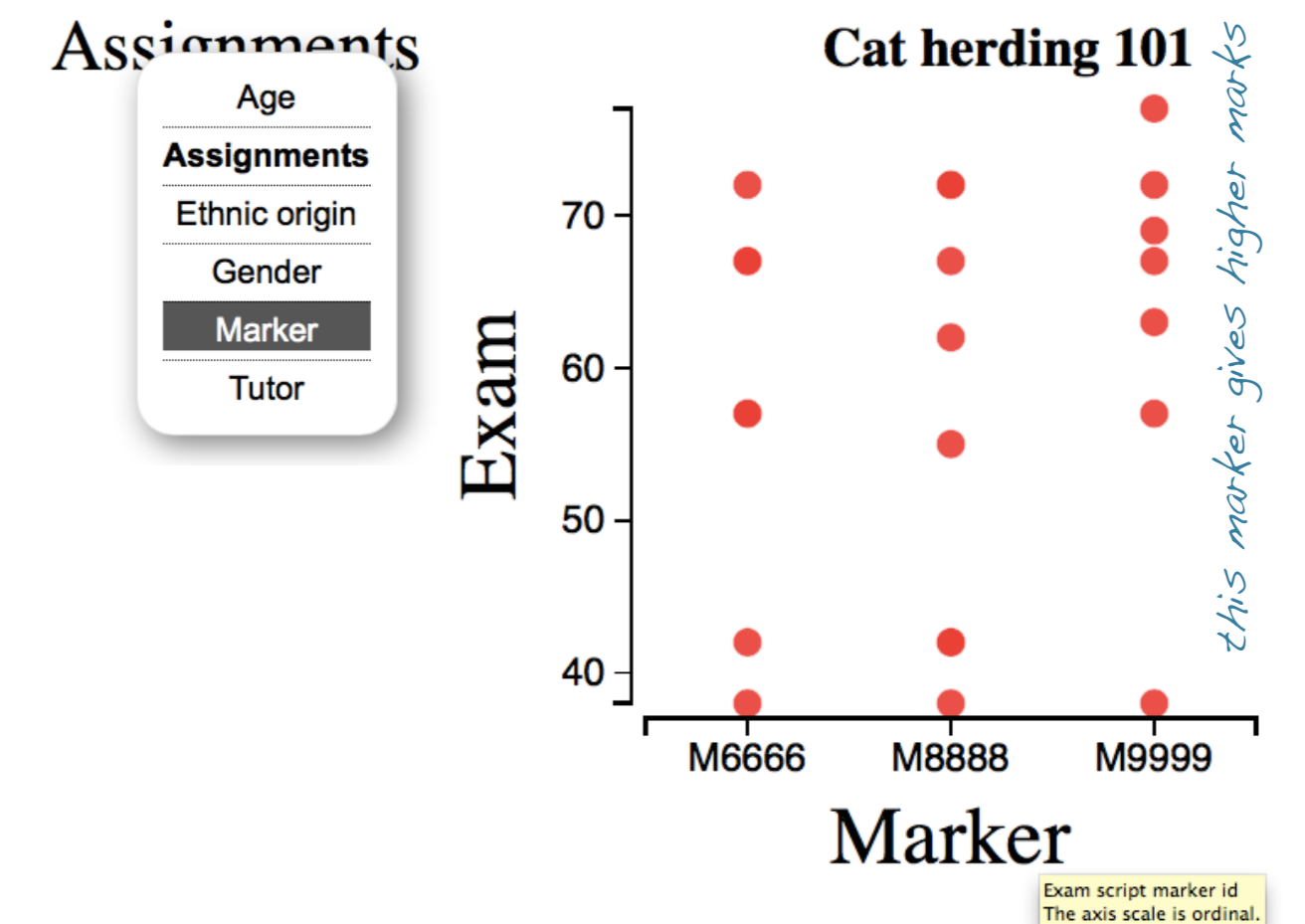
### ② Do older students perform better?

Click on the gear to change settings. Select age for the circle area.



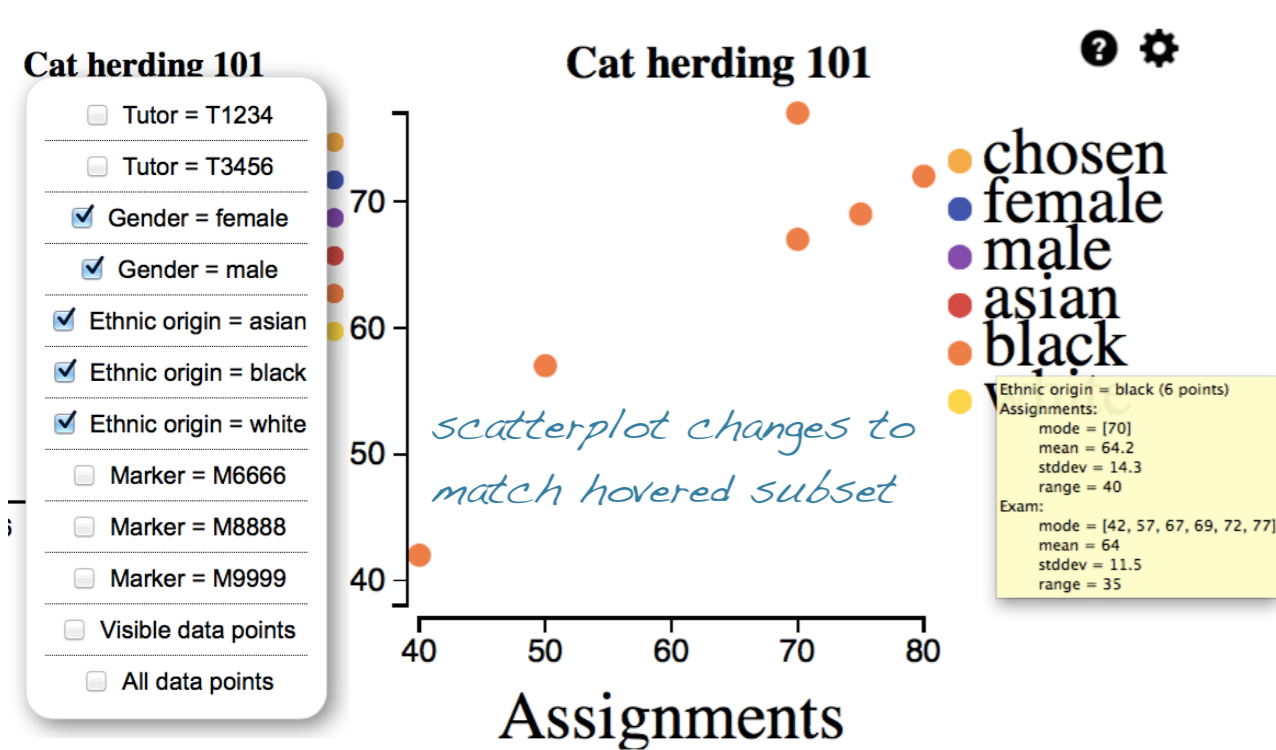
### ③ Is some marker lenient or severe?

Click on the x axis title to select the marker variable.



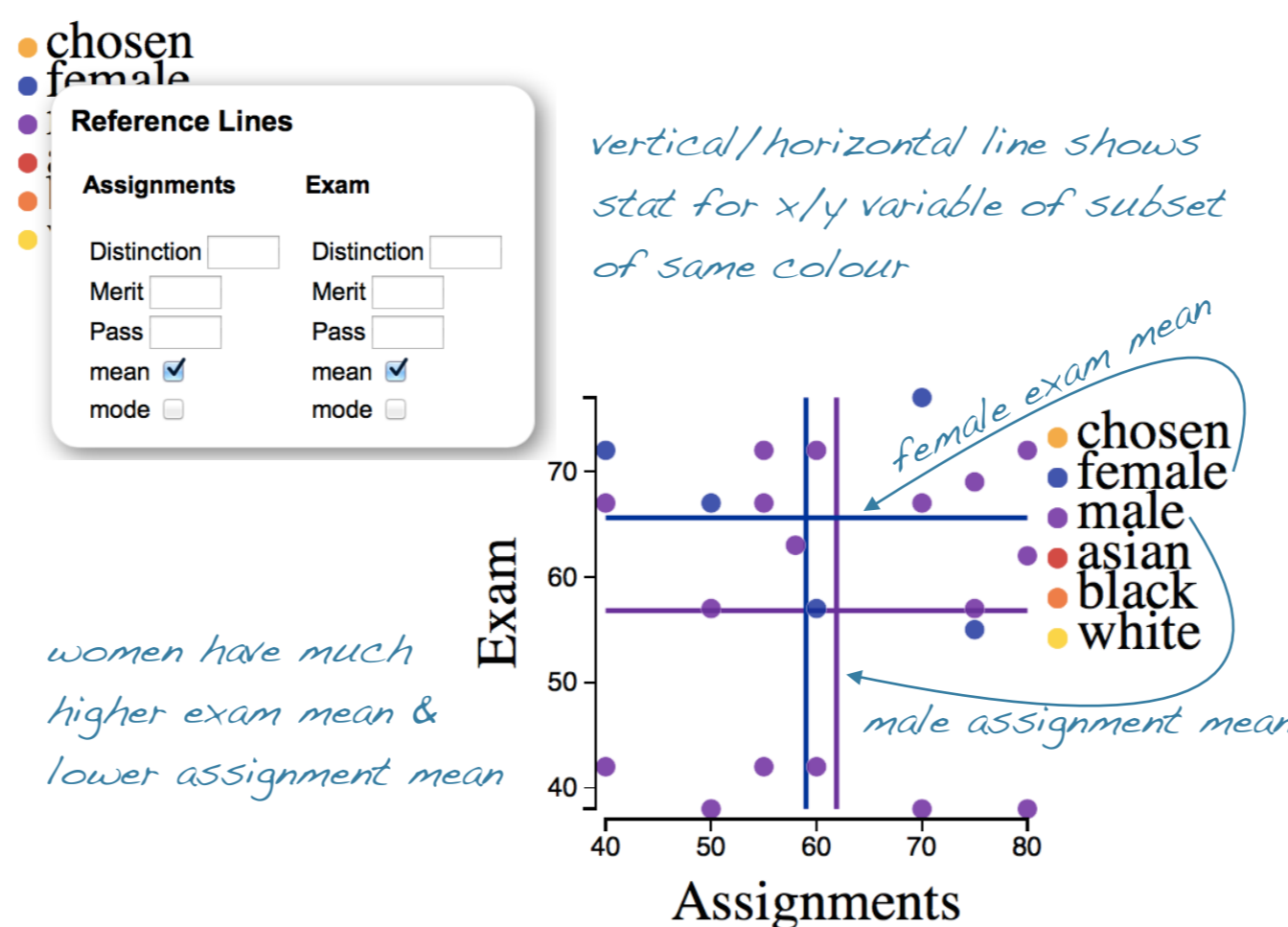
### ④ How have minorities performed?

Click on the chart title to select some pre-computed subsets. Hover on the legend of a subset to see its statistics.



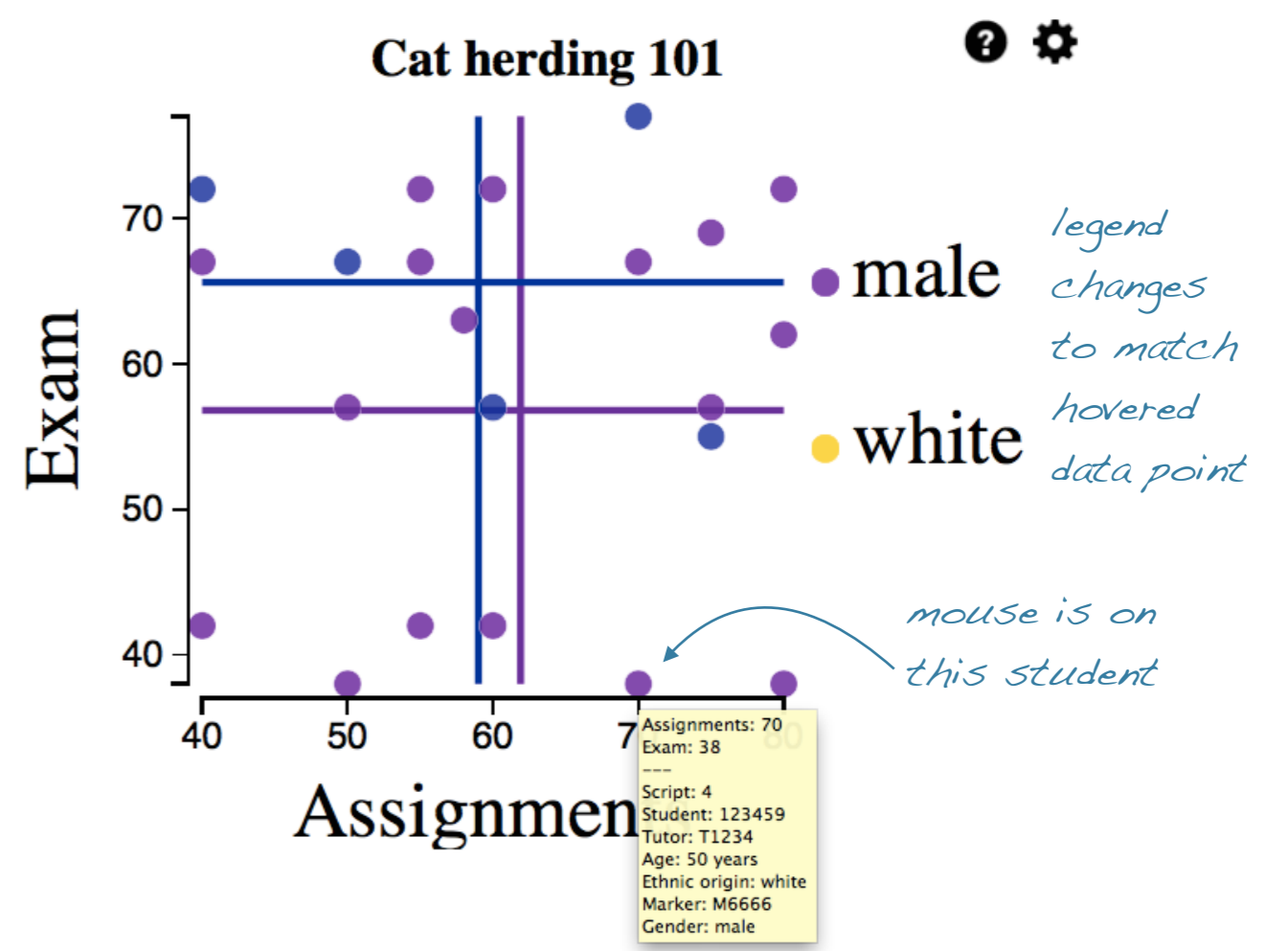
### ⑤ Have women better mean scores?

Click on the female subset. Select the mean for each variable. Repeat for male subset.



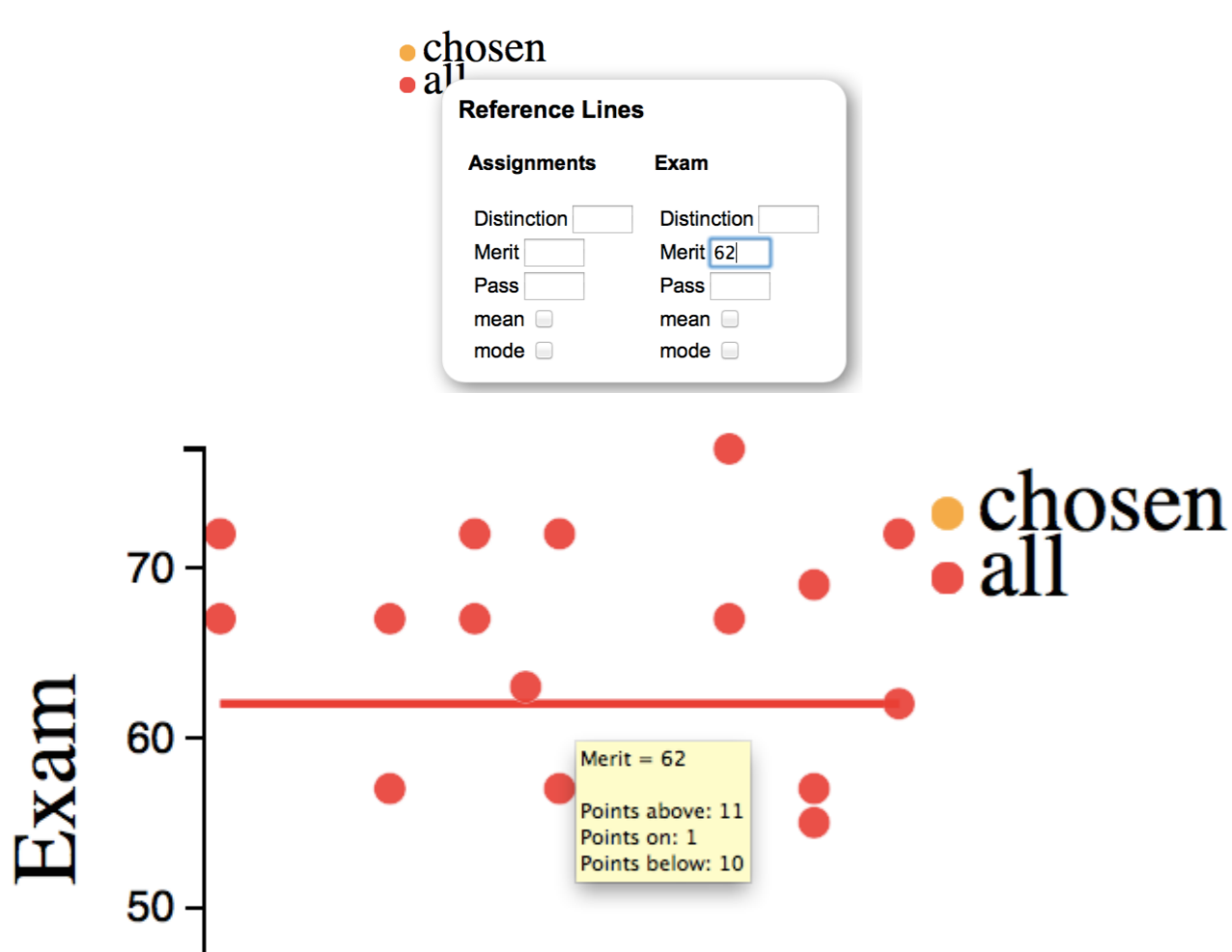
### ⑥ Look at borderline exam scripts

Hover on each student below the pass mark (40) to see their details, like the script id.



### ⑦ How many merits if threshold at 62?

Select 'all data points' like in ④. Click the 'all' legend, set the threshold. Hover on the line.



### ⑧ What is the mean assignment score of under-40 white students?

Select the age variable for the y axis, like in ③. Select the 'white' subset, like in ④. Select the assignment mean for the 'chosen' subset, like in ⑤. Click on white students (yellow circles) below 40 years old. As you do, they're added to 'chosen', its mean is updated, and the line moves accordingly. Hover on it to see the exact assignment mean of those selected so far.

