

Toolbox

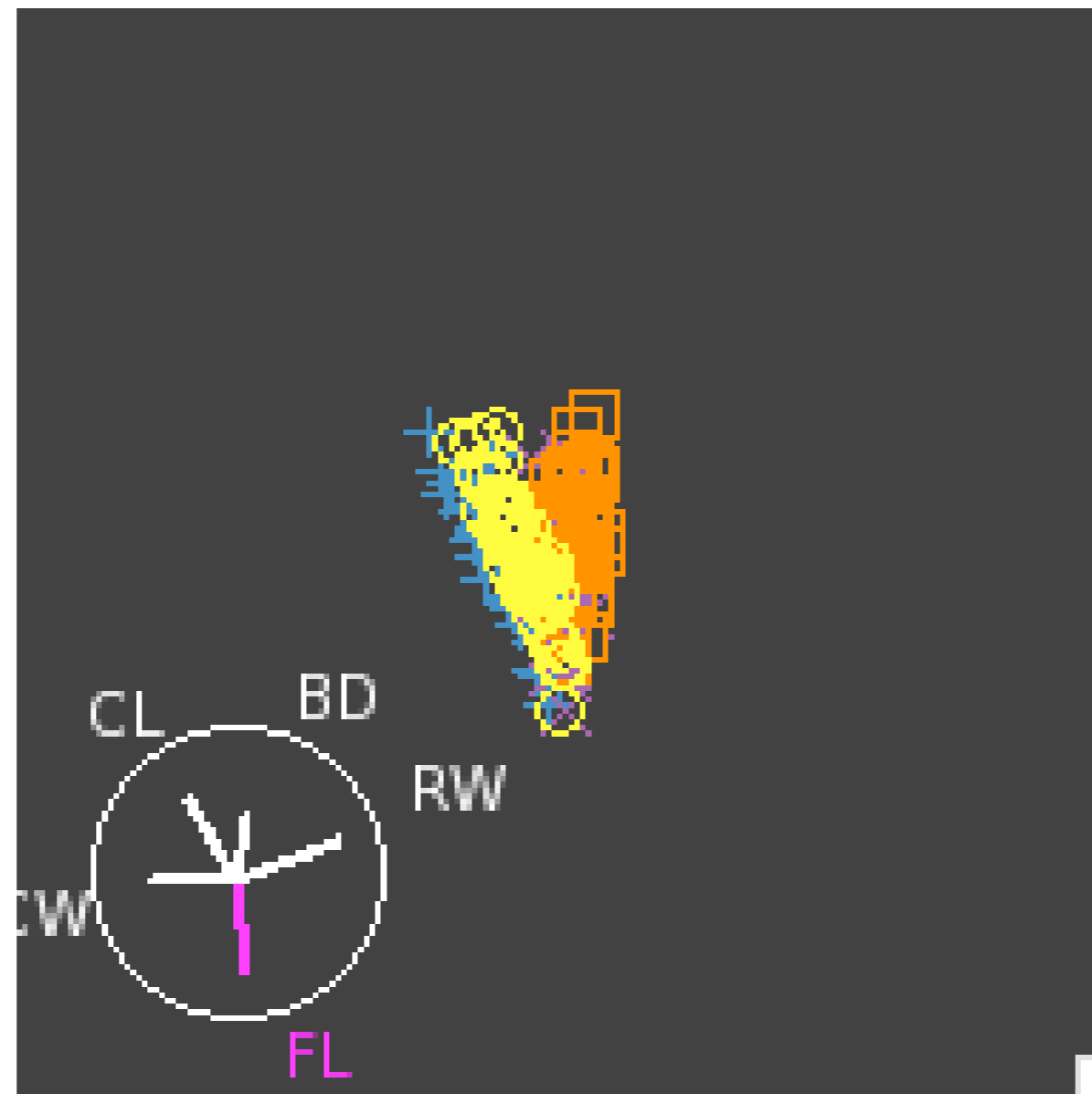
A set of direct manipulations that
we apply to a set of plot types

Toolbox: plot types

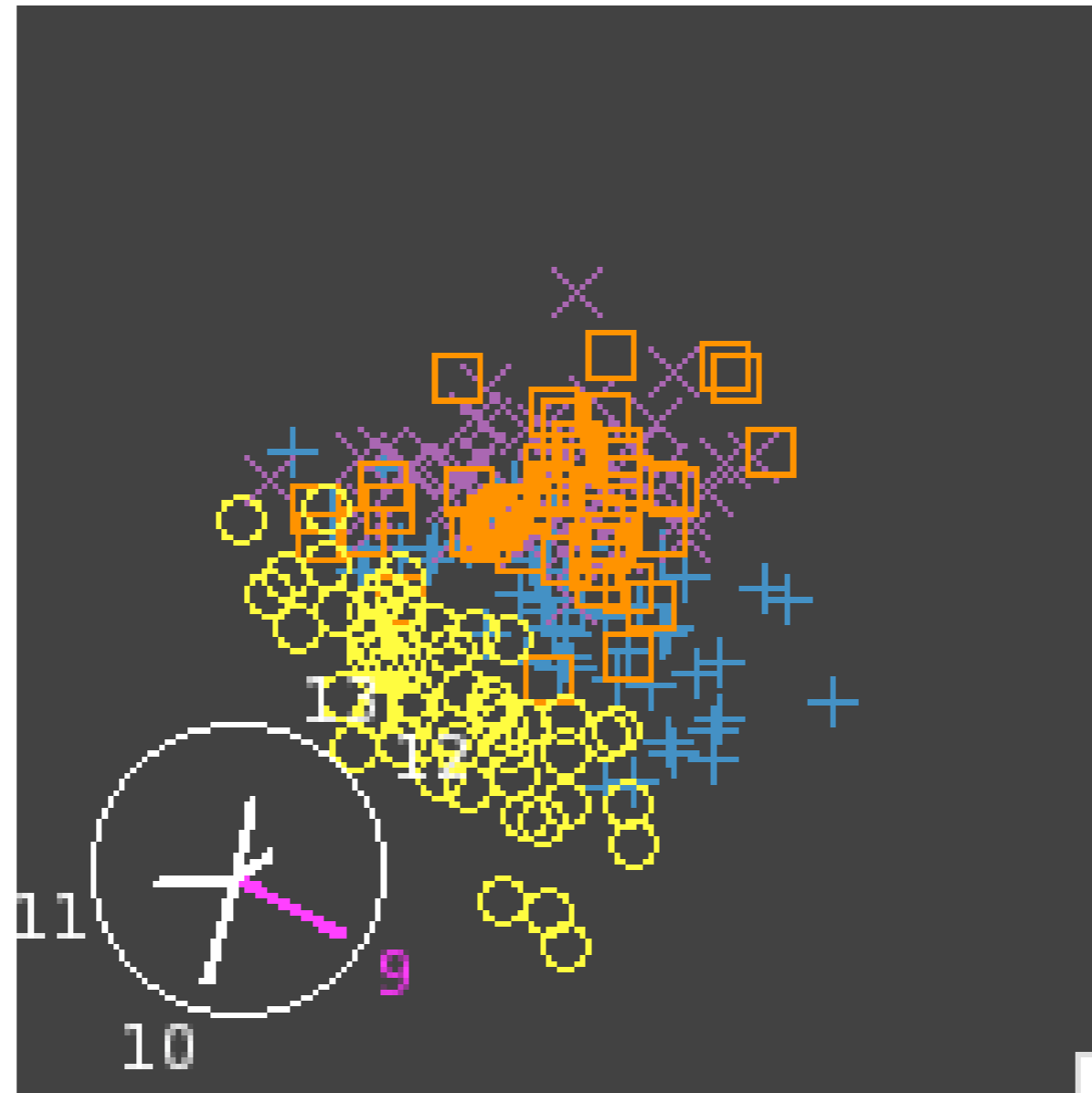
Dotplot	
Histogram	
Bar chart	
Scatterplot	
Parallel coordinate plot	
Scatterplot matrix	

Grand tour	
Guided tour	
Manual controls	

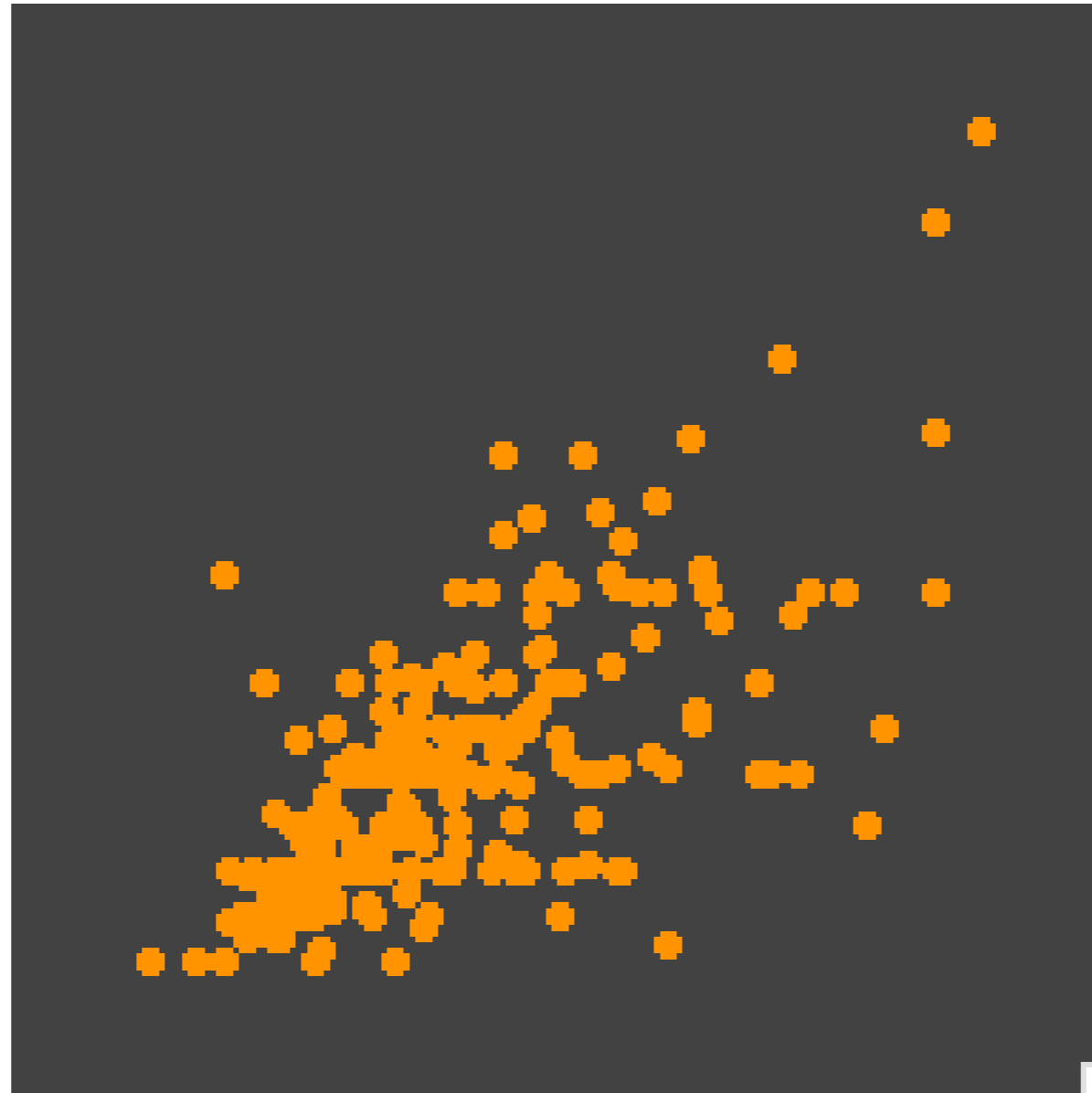
Grand tour



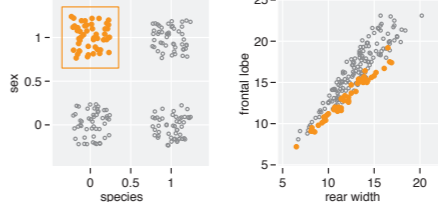
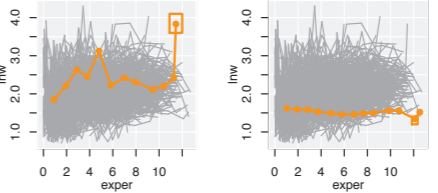
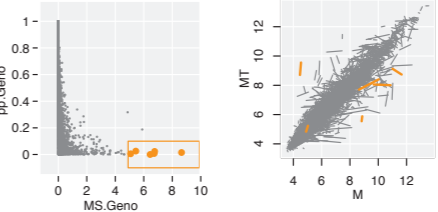
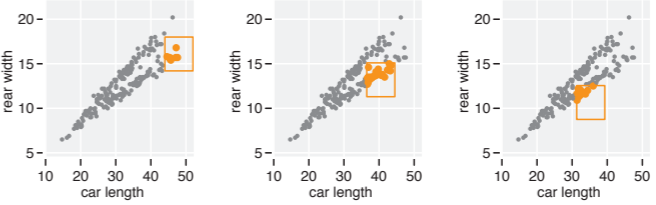
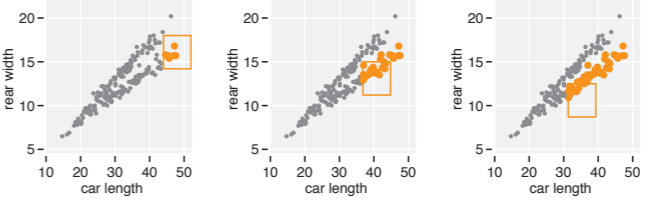
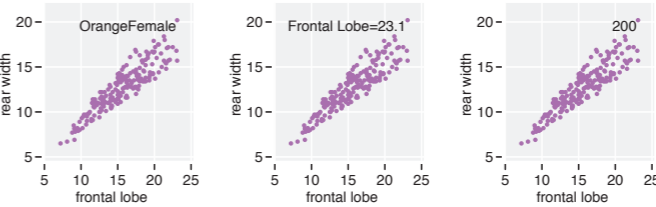

Guided tour



Manual controls



Toolbox: Manipulations

Brushing (one-to-one)	
Brushing (categorical variable)	
Brushing (one point to one line)	
Transient brushing	
Persistent painting	
Identification	
Scaling	

Your turn

1. Open the olive oils data

- a. Generate a barchart of area and a scatterplot of eicosenoic vs. linoleic. Highlight South-Apulia. How would you describe the marginal distribution in linoleic and eicosenoic acids for the oils from South Apulia relative to the other areas?
- b. Generate a grand tour of oleic, linoleic, arachidic, and eicosenoic. Describe the clustering in the 4D space in relationship to difference between the three regions.

2. Open the wages data, and use linking by subject to identify a man:

- a. whose wage is relatively high after 10 years of experience.
- b. who had a high salary at the start of his workforce experience and a lower salary subsequently.
- c. who has had a relatively high wage, and who has a graduate equivalency diploma.

Timeline

20 mins	Toolbox
30 mins	Missing values
45 mins	Supervised Classification
45 mins	Unsupervised Classification
30 mins	Inference

Break