Interactive Visual Analysis with Galaxy Charts

Aysam Guerler

Galaxy Team – Taylor Lab, 2014
Why Visualize?

- Quick check: did it work?
- Exploration and hypothesis generation
- Sharing/publishing
Anscombe’s Quartet
What is Galaxy Charts?

Use Galaxy → Create Tabular Results → Visualize with Galaxy Charts
Make a new chart (1 of 4)

Wait for the upload to complete. Select your **Dataset** and click on the **Visualization Icon** then select **Charts**.
Name your chart **Unclustered Heatmap**.
Select a chart type

Double click on the **Heatmap** icon.
Select data columns

At first click on **Row labels** and select **Column 2**. Then, click on **Draw**.
Select your **Dataset** and click on the **Visualization Icon** then select **Charts**.
Give your chart a name

Name your chart **Clustered Heatmap**.
Select a new chart type

- Area charts
  - Regular (NVD3)
  - Expanded (NVD3)
  - Stream (NVD3)
  - Pie chart (NVD3)

- Data processing (requires 'charts' tool from Toolshed)
  - Histogram (NVD3)
  - Discrete Histogram (jqPlot)
  - Box plot (jqPlot)
  - Clustered Heatmap (Custom)

Double click on the **Clustered Heatmap** icon.
Select data columns

At first click on **Row labels** and select **Column 2**. Then, click on **Draw**.
Use the mouse wheel or your touch pad to zoom into the highlighted area.
**Enlarged view**

**Tooltips** popup if you move the mouse pointer over a box. Here the interaction between B4143 and B3295 is highlighted. Click on **Editor** again to further customize this chart.
Go to the **Configuration** tab.
Heatmap specific options are **highlighted**. Feel free to set **axis labels** or other options.
Define a URL template

Paste a database URL into the template URL field and add the \_LABEL\_ tag. You may use http://www.ncbi.nlm.nih.gov or any other database. Click on Draw to redraw the chart.
Data points linked to web sources

Double click on a box and the browser will open two new tabs using the previously defined URL template.
Cluster selection and analysis

Select one element from each highlighted row. What are the corresponding protein functions?
Identified protein categories

- Chemotaxis
- RNA Polymerase
- Flagella
- Chaperone

Please return to the Editor.
Make a new chart (3 of 4)

Select your **Dataset** and click on the **Visualization Icon** then select **Charts**.
Give your chart a name

Name your chart **Score Histogram**.
Analyze the score distribution

Double click on the **Histogram** icon and click on **Draw**.
Give your chart a name

Click on **Draw**.
Click on **Screenshot** and select **Save as PNG**. Finally, return to the **Editor** again.
Select your Dataset and click on the Visualization Icon then select Charts.
Name your chart **Discrete Histogram**.
Analyze the protein distribution

- Area charts
  - Regular (NVD3)
  - Expanded (NVD3)
  - Stream (NVD3)
  - Pie chart (NVD3)
- Data processing (requires 'charts' tool from Toolshed)
  - Histogram (NVD3)
  - Discrete Histogram (jqPlot)
  - Box plot (jqPlot)
  - Clustered Heatmap (Custom)

Double click on the **Discrete Histogram** icon.
Add more data

Click on Add Data.
Select a second data group

At first click on **Observations** and select **Column 2**. Then, click on **Draw**.
Which proteins have most interactions?

Done with Part I.
Scratchbook
Activate the **Scratchbook** by clicking on the above icon.
Activate the Scratchbook

Click on Saved Visualizations.
Activate the Scratchbook

Select a Visualization and repeat the process by selecting **Saved Visualizations** again.
Scratchbook for multiple charts

Resize all visualizations so they fit into the screen.
Summary

- Galaxy Charts supports **>16 visualization types**
- Integrating **custom visualizations** is easy
- Our **goal** is to handle **>10k data points**
Thank you.