Lab 11 (Lab 5 Makeup) – Intro to Glue

Glue is an interactive data visualization tool. It's particular strength is allowing you to compare subsets of data across multiple plots. The guts of the Glue visualization tool are written in Python, and you can export the python syntax that makes the plot as a python script for any plot that you make.

You may find that an exploration in Glue is a useful starting point for plots for your final project, though note that you should manipulate axis labels to be more descriptive and should add legends if you choose to include plots generated in glue in your final project.

Quick Tips
- To launch glue, type "glue" at a command line
- Upload data files to Glue by dragging and dropping into the "Data Collection" box
- Create plots by dragging data from the "Data collection" box into the main plot window. A dropdown menu will appear with various plot options types
- Change the variables (columns) being plotted, the scaling of axes, limits, etc. using the "Plot options" box at lower left
- Select subsets of data using the data selection tool, which appears inside all plots in the bar at the top as as a blue band with red points on top of it
- To add a new subset, go to Data Manager→ New subset. A new subset will appear in the Subsets menu in the "Data Collection" box
- Subsets can be turned on or off in any given plot by unchecking the box corresponding to them in the "plot layers" box at center left
- To change the data corresponding to a given subset, click on it in the data collection menu, then click over to the plot where you want to do the selecting and use the data selection tool to choose new data
- To change the color, transparency, etc. for a given subset of data, click on it in the "plot layers" box and a menu will appear with various options.
- The full dataset will always appear as a plot layer so that you can apply a given operation to all subsets of data simultaneously.
- To color points according to a categorical variable, use data subsets with different assigned colors. To color points according to a continuous variable, use the "color" dropdown menu for that plot layer.
- To size points in a scatterplot according to another variable, use the "Points" menu in the Plot layer and the "Size" dropdown menu.

Assignment

Your task: Create a plot using glue that captures at least FOUR dimensions of the exoplanet database data (four columns) in a single plot, and that you think tells an interesting and important story about the data. It can be related to something that you looked at for Project 2, or you can come up with a wholly new data visualization.

When you are satisfied with the plot in Glue, export it as a python script (.py file). Open it with a text editor and change the axes labels to be more descriptive (scroll to the bottom), plus add a title and a legend as necessary if you have time. To execute the script and generate your final plot, type python scriptname.py in a terminal.

You'll have until 2:45 to generate your plot, add it to the google slides for today linked from the Moodle page, and practice your presentation. We will come back together to present your results and have a full class discussion at 2:45.

Presentation Guidelines

The member of your lab group who is the *__least__* comfortable with presenting should present the plot. Practice a systematic introduction that includes all major features of the plot (axes, scales, colors, sizes) and describes the big picture/takeaway message about the data.

The lab member who is the *__most__* comfortable presenting should field 1-2 questions from the rest of the class. In order to practice focused listening, restate the question that was asked in your own words (e.g. "Steve has asked why we chose to use a log scale for the x axis but not the y axis"), and ask for clarification where necessary (e.g. "are you asking why we chose to divide the mass into just two bins?").

The criteria for assessment are given on the next page. Please review them carefully before beginning your plot design.

Grading criteria:

	4 points	7 points	10 points
Plot design	The plot fell short in more than one aspect of dimensionality or design	The plot feel a bit short in dimensionality or design	The plot captured at least four data dimensions, and was formatted and labeled according to our criteria for graphical excellence
Plot context	The plot design was sufficiently flawed so as to confuse or obscure the story of the data	The coherence of the story could have been improved in minor ways with tweaks to the plot design	The plot tells an interesting story where all four data dimensions are relevant
Presenting Plot	Presenter missed describing one or more major features of the plot or did not interpret the big picture	Presenter missed describing one or more minor features of the plot or takeaway message	Presenter described all of the major features of the plot and the big picture takeaway
Fielding questions	Presenter did not attempt to summarize or clarify the question(s) or did not answer politely	Presenter fell a little short in terms of requesting clarification, restating the question, or answering clearly and thoroughly	Presenter asked for clarification if necessary, restated the question, and gave a clear, polite and thoughtful response

Total Points: _____/40