

Lecture 02 – Tools for Visualization

Today's Learning Objectives:

- 1. Gain familiarity with the basic tools we'll be using this semester (GIMP, R & RStudio, ParaView, git/Github, and Peer Review)**

- 2. Practice independently performing basic data analysis and visualizations in a way that communicates ideas clearly.**

- 3. Practice constructive criticism on peer work.**

Gnu Image Manipulation Program (GIMP)

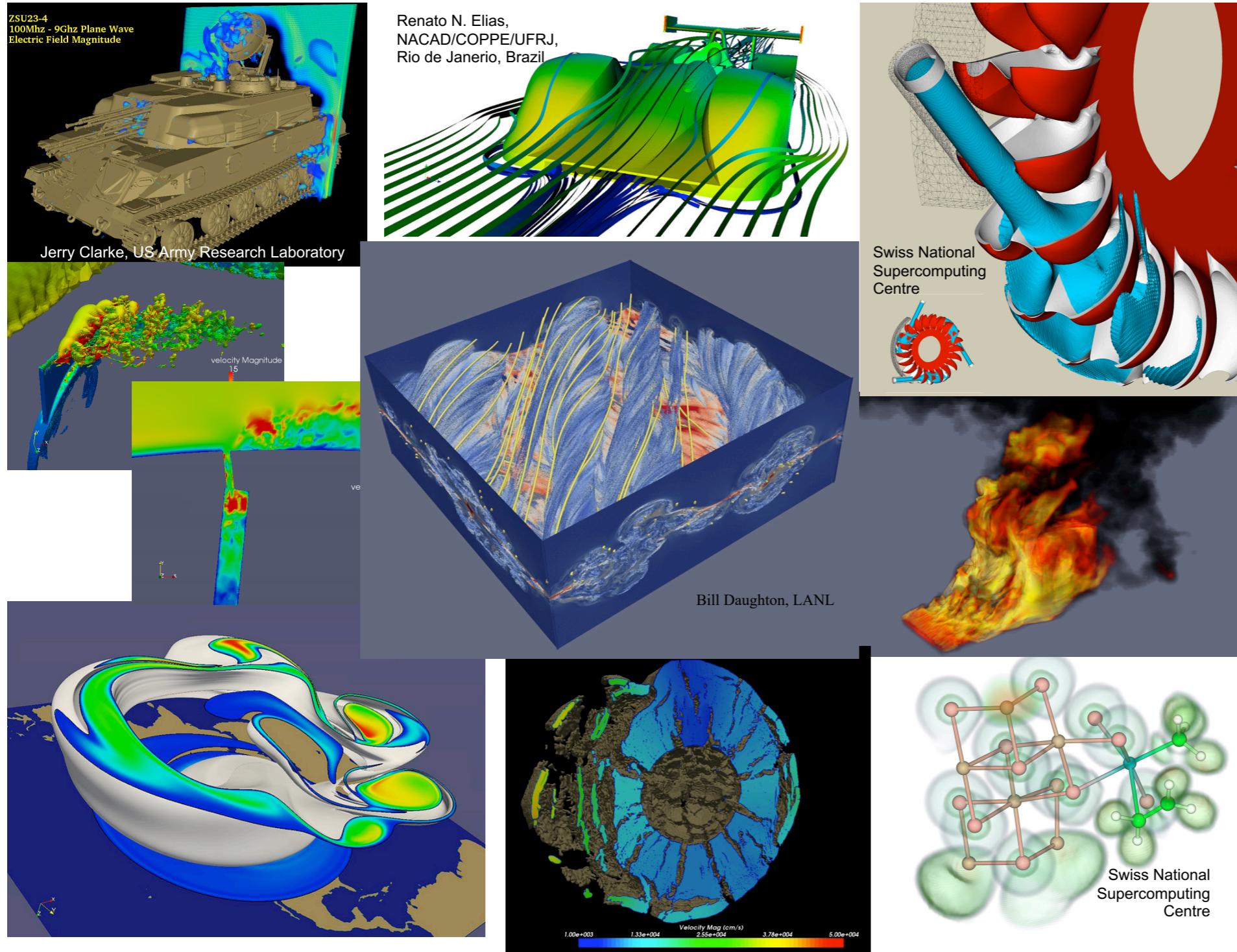
<https://www.gimp.org/tutorials/>



ParaView (Kitware)



https://www.paraview.org/Wiki/The_ParaView_Tutorial



R and RStudio



Studio®

<https://github.com/rstudio/hex-stickers/>

Many tutorials are available.

The screenshot shows the RStudio interface. The top bar has tabs for "RStudio" and "Project: (None)". The left sidebar includes "Go to file/function", "Addins", and a "Files" section. The main area has tabs for "Environment", "History", and "Connections". The "Environment" tab shows variables: "alldata2" (681 obs. of 20 variables), "values" (681 obs. of 3 variables), and "Labels" (density: 1000, dia: 0.1). The "Plots" tab displays a 3D scatterplot titled "Cost of Transport" with axes "CR", "Wo", and "Cost of Transport". The plot shows a dense cloud of purple points and several outliers in orange and yellow. The "Console" tab at the bottom shows R code and its execution results. It includes a warning about the "viridis" package and a successful library loading command.

```
237 alldata2$Freq,alldata2$Wo,alldata2$COT)
238 V2","V3")
239 t of Transport")
240 s(20,alpha=1,begin=0,end=1,direction=1,option="C"),values,c(0,3*10^8),labels,55,3,0.8,0.08)
241 , adj=-0.1, "A", cex=1.2)
242
243 alldata_pinch2$Comp,alldata_pinch2$Wo,alldata_pinch2$COT)
244 V2","V3")
245 st of Transport")
246 s(20,alpha=1,begin=0,end=1,direction=1,option="C"),values,c(0,3*10^8),labels,25,2,0.85,0.05)
247 , adj=-0.1, "B", cex=1.2)
248
249 s(20,alpha=1,begin=0,end=1,direction=1,option="C"),alldata_pinch2$COT,c(0,3*10^8),nticks=4,3)
250
251 alldata2$Freq,alldata2$Comp,alldata2$Work)
252 V2","V3")
253 k")
254 s(20,alpha=1,begin=0,end=1,direction=1,option="A"),values,c(0,20000),labels,45,4,0.85,0.07)
255 , adj=-0.1, "C", cex=1.2)
256
257
258 alldata_pinch2$Freq,alldata_pinch2$Wo,alldata_pinch2$Work)
246:116 # Cost of Transport and Work
Cost of Transport and Work
Cost of Transport
<-->
Error in viridis(20, alpha = 1, begin = 0, end = 1, direction = 1, option = "C") : could not find function "viridis"
library(viridis)
Loading required package: viridisLite
new3Dscatterplot(viridis(20,alpha=1,begin=0,end=1,direction=1,option="C"),values,c(0,3*10^8),labels,25,2,0.85,0.05)
>
```



Sign up for Github:

Github: <https://github.com>

Use your Chapman email address!

Pick a password that is different than your Chapman password

Install git:

git download: <https://git-scm.com/downloads>

Put it in default directory

Activate in RStudio:

Go to Preferences > Git/SVN

Check “Enable version control interface for RStudio projects”



Start a New Project in RStudio (Method 1 – Easiest)

Go to Github and in “Repositories” click the green “New Repository” button.

Pick a unique repository name (this will also be a folder name on your computer).

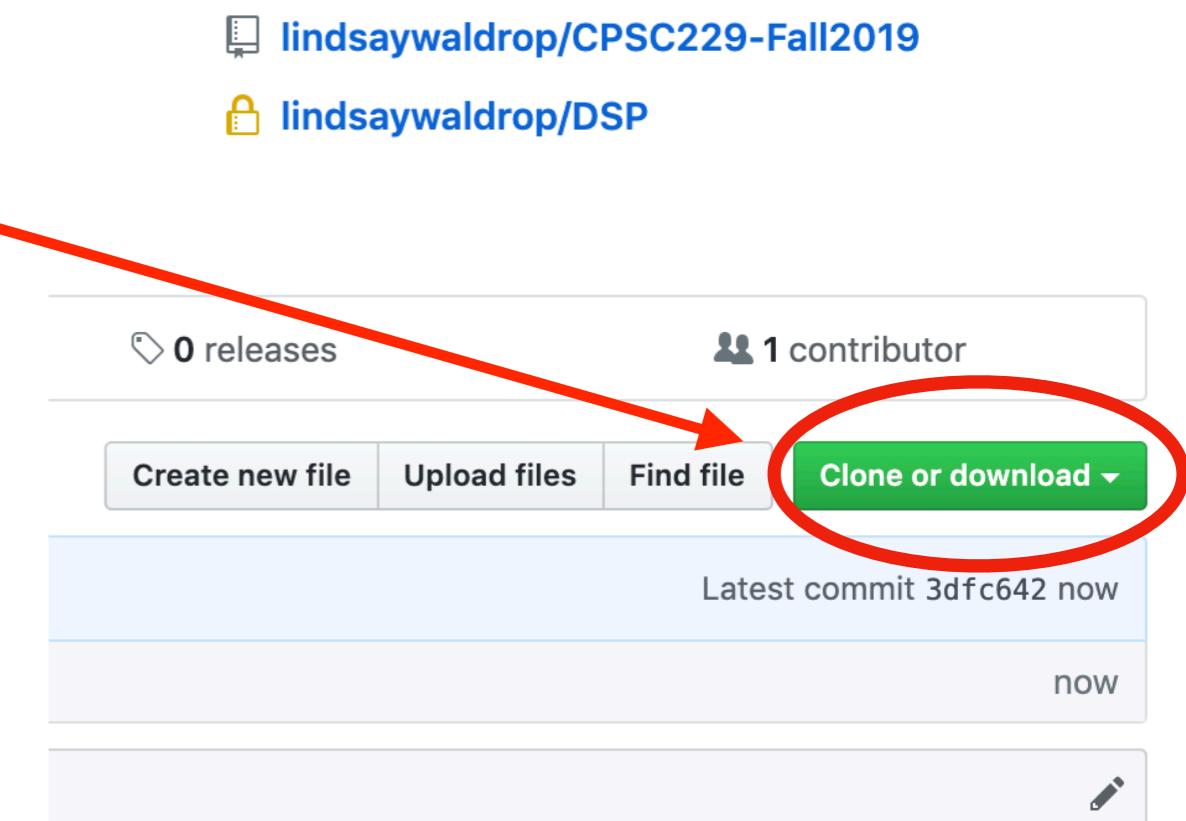
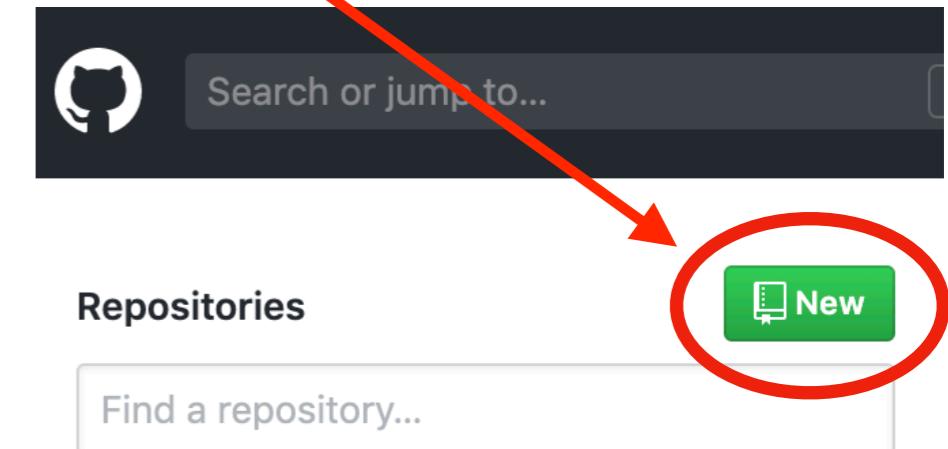
Click “Initialize this repository with a README file” then “Create Repository”

Click the “Clone or download” button and copy the URL to your clipboard.

Return to RStudio and from Files select “New Project...”

Select “Version Control” option and then “git”

Copy the URL from Github into the repository URL box and then click through to create a new project!



Protip: If you want to clone an existing repository, follow the same instructions, just use the existing repository’s URL instead of creating a new one!

Start a New Project in RStudio (Method 2 – Harder)

From RStudio and from Files select “New Project...” and then “New Directory” and “New Project”.

Be sure to select “Create a git repository”. Click through to start your new project.

Add some code/files. Commit these changes.

Go to Github and in “Repositories” click the green “New Repository” button.

Give your repository on Github THE SAME NAME as your RStudio project (this will help sync them).

UNSELECT “Initialize this repository with a README file” then “Create Repository.”

Click the “Clone or download” button and copy the URL to your clipboard.

Return to RStudio and click the little cog under the “Git” tab, select “Shell.”



Click the “Clone or download” button and copy the URL to your clipboard.

Enter the following commands:

```
> git remote add origin URLYOUPIED  
> git remote -v  
> git push --set-upstream origin master
```

Now you can push and pull from Github!

Hands-on Practice

Get in pairs based on the type of plotting program that you enjoy working with the most (R, Python, whatever). This will require hands-on work.

Go to the following Github page and download data set:

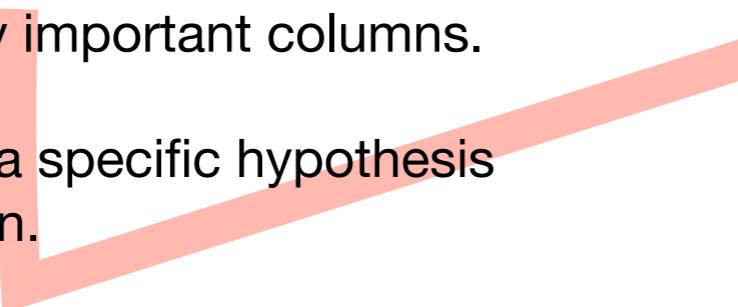
<https://github.com/fivethirtyeight/data/tree/master/nba-elo>

With your partner, create two visualizations that answer the following questions with the data on NBA Elo ratings:

1. Do NBA Elo ratings reliably predict the outcome of a game?

2. Do NBA Elo ratings reliably predict the point differential by which a game should be won or lost?

Tips:

- Start by importing and looking at the data set. Identify important columns.
 - Come up with a specific hypothesis for each question.
 - Think of a plot that could show your hypothesis.
 - Create a plot with a script.
 - Give to your partner for input.
- 

Lecture 02 – Tools for Visualization

Action Items for Next Time:

Homework:

1. Complete Data Carpentry R tutorial section “Intro to R” (if you need R refresher)
2. Look through GIMP Beginner Tutorial.
3. Complete Software Carpentry’s Version Control with git tutorial (optional)
4. Submit powerpoint slide on Canvas by Mon 11 am.

Reading for next time:

1. Ware Chapter 1