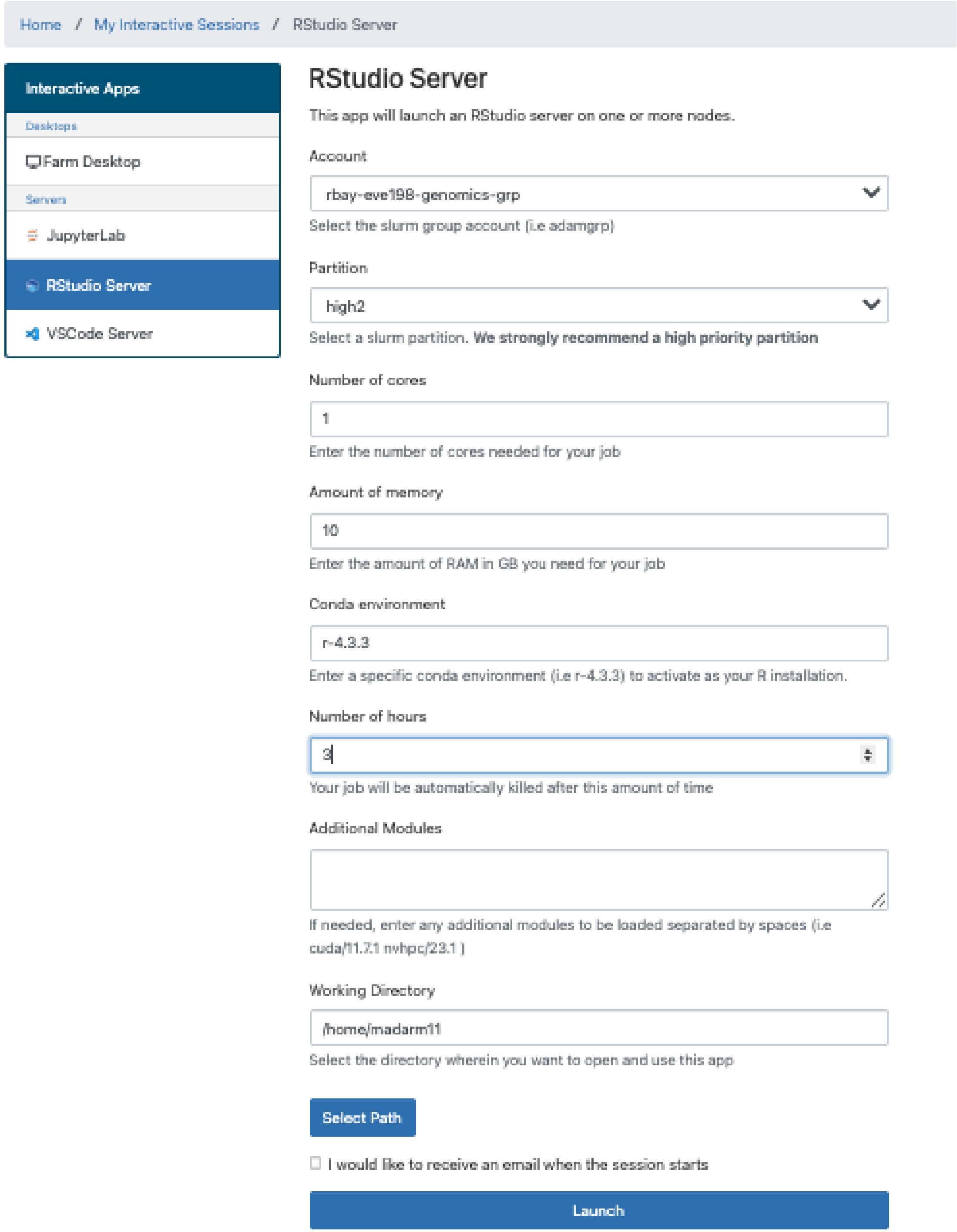
Intro to R!

Eve198

Week 5

Maddie Armstrong & Rachael Bay



RStudio Server

This app will launch an RStudio server on one or more nodes.

Account

rbay-eve198-genomics-grp

Select the slurm group account (i.e adamgrp)

Partition

high2

Select a slurm partition. **We strongly recommend a high priority partition**

Number of cores

1

Enter the number of cores needed for your job

Amount of memory

10

Enter the amount of RAM in GB you need for your job

Conda environment

r-4.3.3

Enter a specific conda environment (i.e r-4.3.3) to activate as your R installation.

Number of hours

3

Your job will be automatically killed after this amount of time

Additional Modules

If needed, enter any additional modules to be loaded separated by spaces (i.e cuda/11.7.1 nvhpc/23.1)

Working Directory

/home/madarm11

Select the directory wherein you want to open and use this app

☐ I would like to receive an email when the session starts

What is R? Rstudio?

Base R

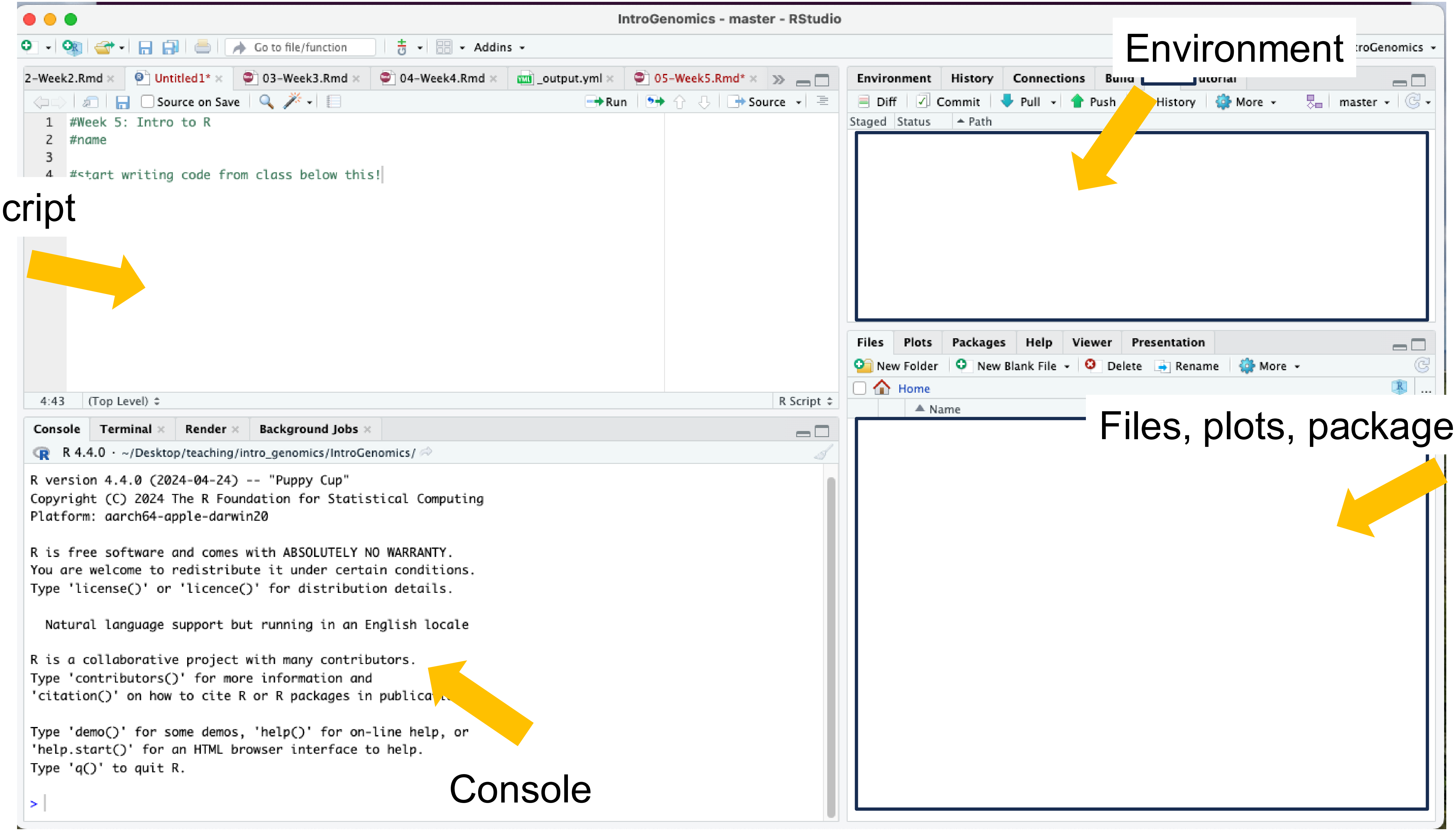
- Programming language
- Command-line interface where code is run

R studio

- User interface/Integrative development environment using the R programming language
- Visualization tools for analysis and figure making

https://wiingy.com/learn/r-studio/difference-between-r-programming-and-rstudio/

Outside of this class you can get R studio for your desktop here: https://posit.co/download/rstudio-desktop/



```
#Week 5: Intro to R
#name

#start writing code from class below this!
```

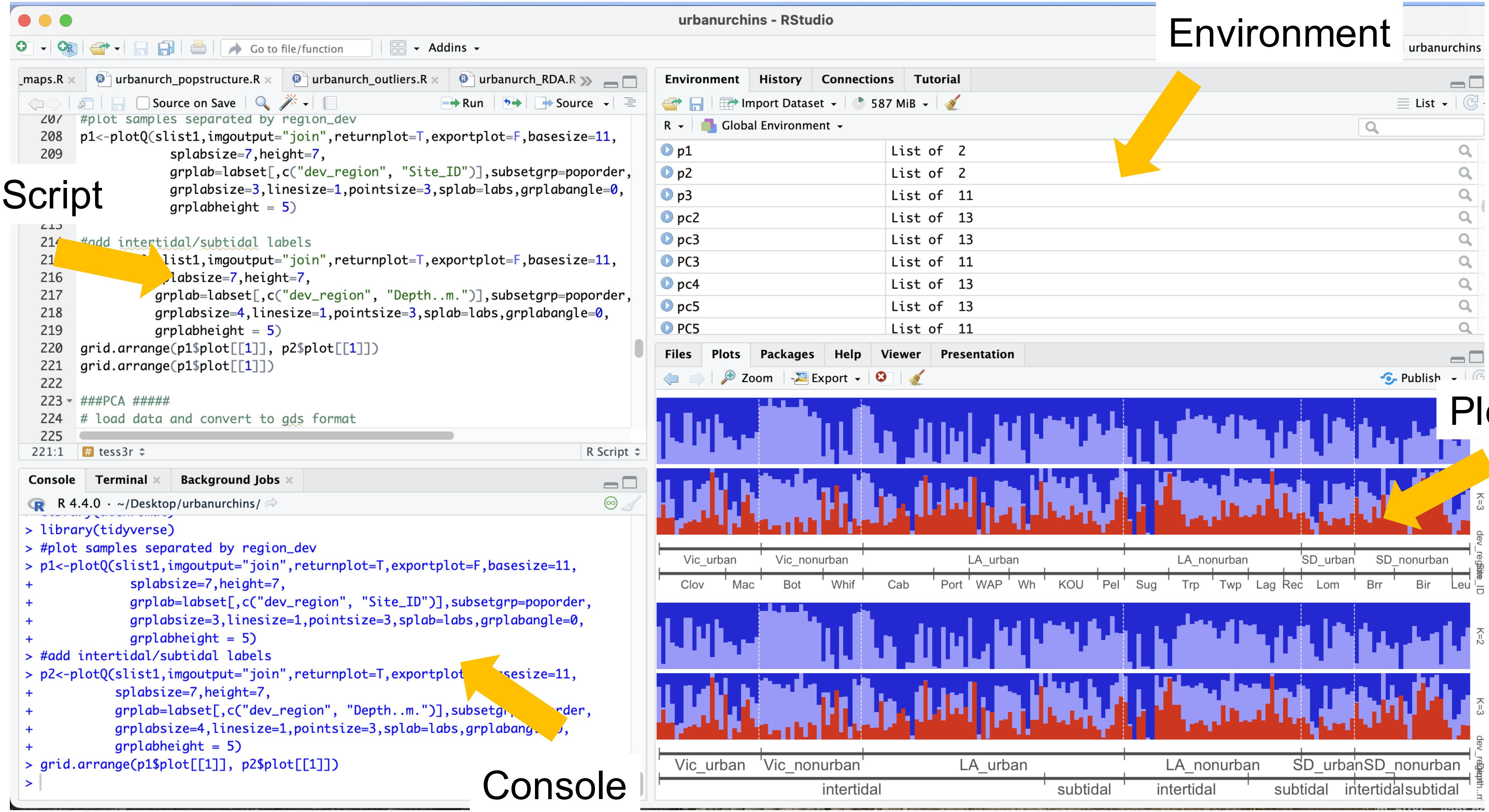
```
R version 4.4.0 (2024-04-24) -- "Puppy Cup"
Copyright (C) 2024 The R Foundation for Statistical Computing
Platform: aarch64-apple-darwin20

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

Environment

Script

Files, plots, packages!

Console



Environment

Script

Plots!

Console

```
#plot samples separated by region_dev
p1<-plotQ(slist1,imgoutput="join",returnplot=T,exportplot=F,basesize=11,
            splabsize=7,height=7,
            grplab=labset[,c("dev_region", "Site_ID")],subsetgrp=poporder,
            grplabsize=3,linesize=1,pointsize=3,splab=labs,grplabangle=0,
            grplabheight = 5)

#add intertidal/subtidal labels
           list1,imgoutput="join",returnplot=T,exportplot=F,basesize=11,
          plabsize=7,height=7,
          grplab=labset[,c("dev_region", "Depth..m.")],subsetgrp=poporder,
          grplabsize=4,linesize=1,pointsize=3,splab=labs,grplabangle=0,
          grplabheight = 5)
grid.arrange(p1$plot[[1]], p2$plot[[1]])
grid.arrange(p1$plot[[1]])

###PCA #####
# load data and convert to gds format
```

p1	List of 2
p2	List of 2
p3	List of 11
pc2	List of 13
pc3	List of 13
PC3	List of 11
pc4	List of 13
pc5	List of 13
PC5	List of 11

```
> library(tidyverse)
> #plot samples separated by region_dev
> p1<-plotQ(slist1,imgoutput="join",returnplot=T,exportplot=F,basesize=11,
+             splabsize=7,height=7,
+             grplab=labset[,c("dev_region", "Site_ID")],subsetgrp=poporder,
+             grplabsize=3,linesize=1,pointsize=3,splab=labs,grplabangle=0,
+             grplabheight = 5)
> #add intertidal/subtidal labels
> p2<-plotQ(slist1,imgoutput="join",returnplot=T,exportplot        sesize=11,
+           splabsize=7,height=7,
+           grplab=labset[,c("dev_region", "Depth..m.")],subsetg          rder,
+           grplabsize=4,linesize=1,pointsize=3,splab=labs,grplabang     ),
+           grplabheight = 5)
> grid.arrange(p1$plot[[1]], p2$plot[[1]])
>
```

Important Terms

Object: a “thing” saved to your R environment , this can include

- Data
- Plot
- etc!

Function: something you can “do” to the data

Packages: tools built into R or downloaded that you can use to complete different analyses/make figures!

https://datacarpentry.github.io/R-genomics/01-intro-to-R.html#vectors_and_data_types

Data Types/Structures:

Data Types:

1. Character (names)
2. Numeric (numbers)
3. Logical (T/F)
4. Integer (2L for example)
5. Complex (imaginary #s)
6. Raw (not really used)

Data Structures:

1. Vector
2. List
3. Matrices
4. Factor
5. Data frame

https://datacarpentry.github.io/R-genomics/01-intro-to-R.html#vectors_and_data_types

Let’s get to coding!