Working with R in the RE

Emily Perry and Eleni Christodoulou

Research Engagement Manager

11th March 2025

Data security

- This training session will include data from the GEL Research Environment
- As part of your IG training you have agreed to not distribute these data in any way
- If you are joining virtually, you are not allowed to:
    - Invite colleagues to watch this training with you
    - Take any screenshots or videos of the training
    - Share your webinar link (we will remove anyone who is here twice)

Presenters

Emily Perry
Research Engagement Manager

Eleni Christodoulou
Solutions Specialist - LifeBit

Questions

All your microphones are muted

Use the Zoom Q&A to ask questions

Upvote your favourite questions: if we are short on time we will prioritise those with the most votes

Helpers

Magdalena Drożdż
Bioinformatician - Research Services

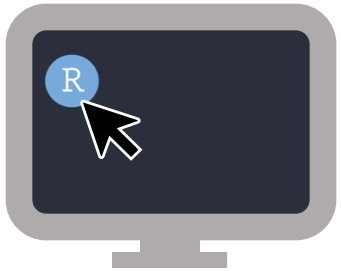
Hamzah Syed
Solutions Manager - Lifebit

Agenda

1. Introduction and admin
2. Working with RStudio on the HPC
3. Plotting in R
4. Working with R libraries
5. Working with RStudio in CloudOS interactive sessions
6. Query clinical data with R
7. Help and questions

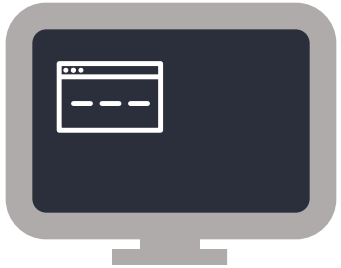
2. Working with RStudio on the HPC

RStudio in the RE



Opens a default version of R

Can launch an HPC job



Choose your preferred R version

A job on the inter queue

```
bsub -q inter -P <your_project_code> -R
rusage[mem=1000] –M 1000 –n 1 -Is bin/bash
```

Do some work

```
bjobs
```
(to find job number)

```
bkill <job number>
```

https://re-docs.genomicsengland.co.uk/hpc_queues/#the-inter-queue

Rstudio demo

Camtasia 2024 File Edit Modify Text View Export Window Help Fri 17 Jan 13:28

Amazon WorkSpaces

Computer
eperry's Home
Link to emily
Old Firefox Data
Airlock
CloudOS Academic
CloudOS Discovery Forum
CloudOS Internal
Desktop.Rproj
Document Viewer
Emacs
Ensembl
Firefox
Git GUI
GVim
IGV Browser
IVA
Labkey
LibreOffice 7.6
Open Targets
Panel App
Participant Explorer
R
RE Messages
Research Environment Documentation
Research Registry
RStudio
Terminal Emulator
Text Editor
Visual Studio Code
Welcome Pack
Trash

Genomics england

Applications Places System

Fri Jan 17, 13:28

3. Plotting in R

No GUI on the HPC

```
png("my_image.png")
```

>

```
Unable to start device PNG
or
Unable to open connection
to X11 display
```

Use an X Virtual Frame Buffer

```
> module load
R/<version>
> xvfb-run -a R
```

```
png("my_image.png")
plot(1)
dev.off()
```

.bashrc

```
alias R='xvfb-run -a
R'
```

```
png("my_image.png")
plot(1)
dev.off()
```

.bashrc

```
function r_headless(){
module purge
module load lang/R/<version>
xvfb-run -a Rscript $1
}
```

```
png("my_image.png")
plot(1)
dev.off()
```

```
r_headless script.R
```

Use an X Virtual Frame Buffer

```
> module load
R/<version>
> xvfb-run -a R
```

```
library(ggplot2)
fig_1 <- ggplot(*some_data*)
ggsave(fig_1, type = "cairo",
file="fig_1.<extension>")
```

.bashrc

```
alias R='xvfb-run -a
R'
```

```
library(ggplot2)
fig_1 <- ggplot(*some_data*)
ggsave(fig_1, type = "cairo",
file="fig_1.<extension>")
```

.bashrc

```
function r_headless(){
module purge
module load lang/R/<version>
xvfb-run -a Rscript $1
}
```

```
library(ggplot2)
fig_1 <- ggplot(*some_data*)
ggsave(fig_1, type = "cairo",
file="fig_1.<extension>")
```

```
r_headless script.R
```

Plotting demo

Camtasia 2024 File Edit Modify Text View Export Window Help Fri 17 Jan 13:28

Amazon WorkSpaces

Computer
eperry's Home
Link to emily
Old Firefox Data
Airlock
CloudOS Academic
CloudOS Discovery Forum
CloudOS Internal
Desktop.Rproj
Document Viewer
Emacs
Ensembl
Firefox
Git GUI
GVim
IGV Browser
IVA
Labkey
LibreOffice 7.6
Open Targets
Panel App
Participant Explorer
R
RE Messages
Research Environment Documentation
Research Registry
RStudio
Terminal Emulator
Text Editor
Visual Studio Code
Welcome Pack
Trash

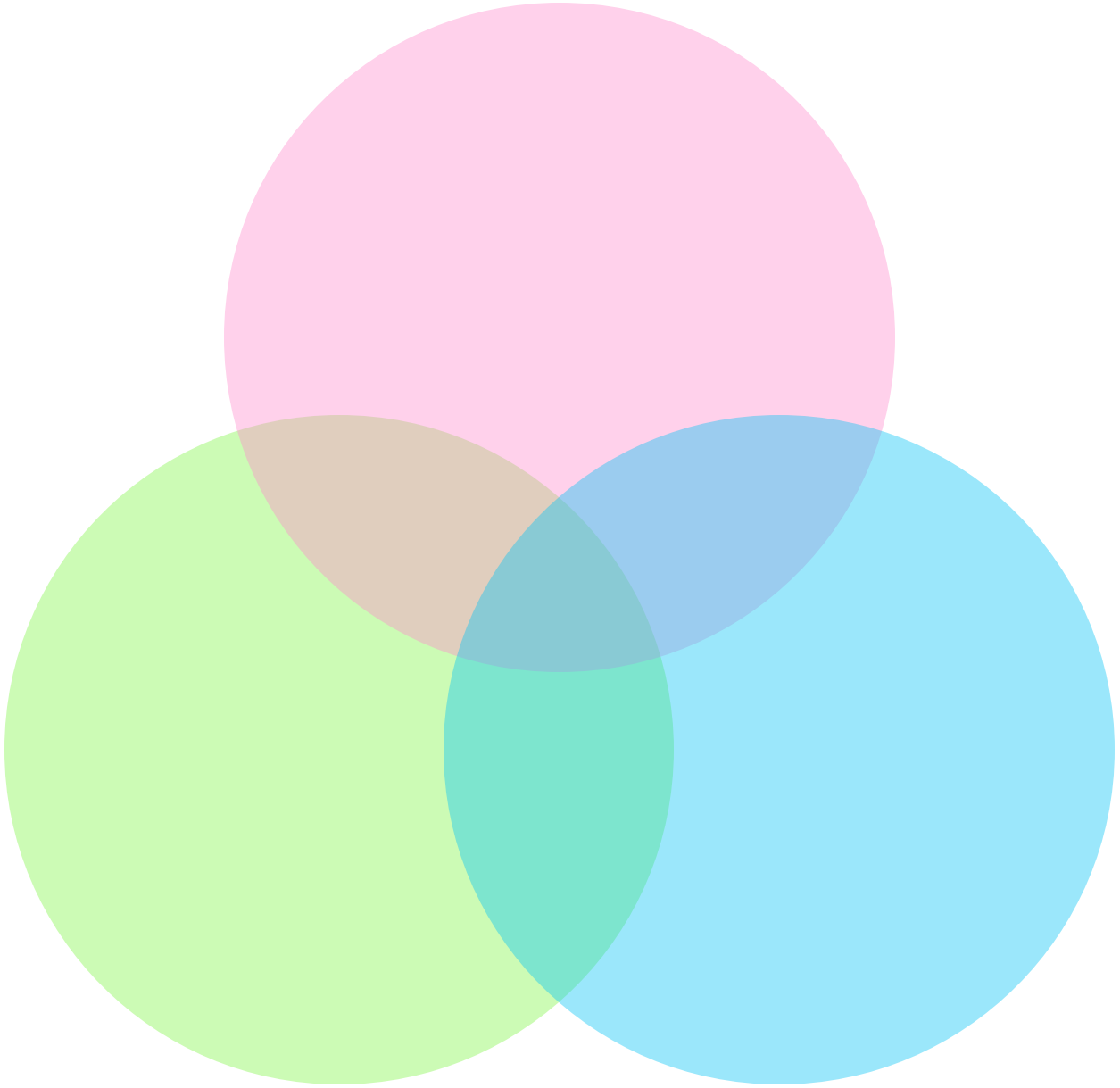
Genomics england

Applications Places System

Fri Jan 17, 13:28

4. Working with R libraries

R packages

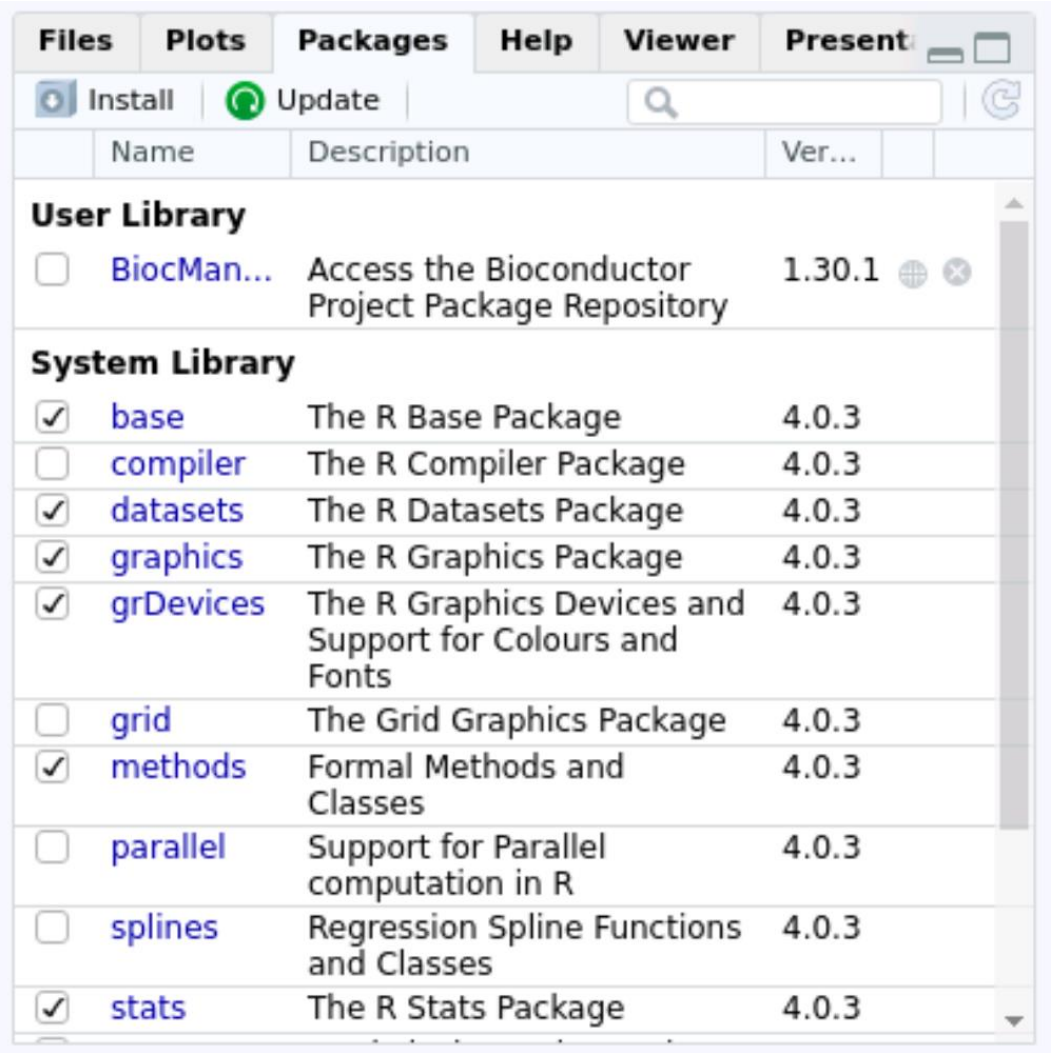


R/3.6.3
R/4.2.1
R/4.3.3

https://re-docs.genomicsengland.co.uk/r_packages/

Loading libraries

```
library(library_name)
```



https://re-docs.genomicsengland.co.uk/r_packages/#installing-r-packages-from-cran

Bioconductor

R 4.33

```
options(
BIOCONDUCTOR_CONFIG_FILE =
"https://artifactory.aws.gel.ac:443/artifactory/bioconductor.org-cache/config.yaml"
)

library("BiocManager")
BiocManager::install("<package_name>")
library(<package_name>)
```

R 3.6.0 to 4.3.2

```
.libPaths(c( .libPaths(), "/tools/aws-workspace-apps/ce/R/4.0.2"))
library("BiocManager")
BiocManager::install("<package_name>")
library(<package_name>)
```

https://re-docs.genomicsengland.co.uk/r/#installing-and-configuring-packages-from-bioconductor

CRAN/Bioconductor demo

Chrome File Edit View History Bookmarks Profiles Tab Window Help Fri 10 Mar 10:55

Inuvika OVD | Software available on the HPC

re.extge.co.uk/ovd/

Raise A Ticket | RE docs | CloudRE-docs | JIRA Confluence | Multi Factor Auth | Change Password | bob | SharePoint | RE1 | RE1 UG | brs_git | Profit.co | product board

Terminal - eperry@corp.gel.ac@phpgridzlogn003:/gel_data_resources/software_catalogues/R_catalogue

File Edit View Terminal Tabs Help

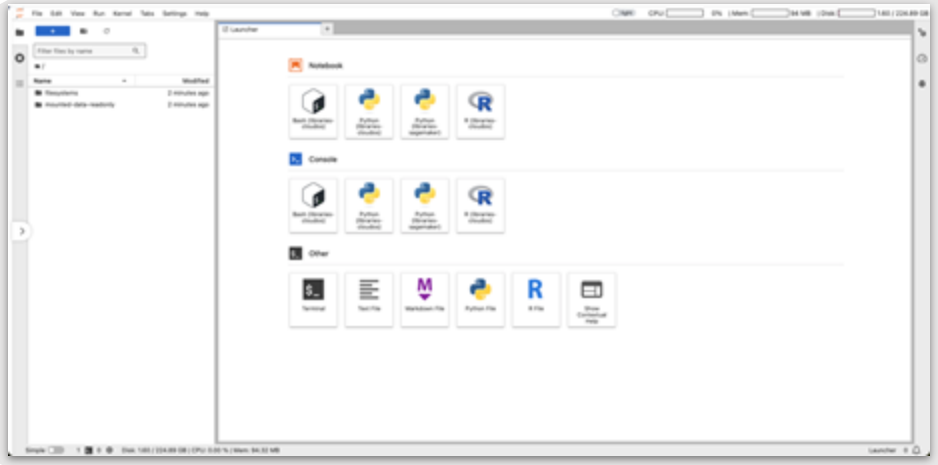
```
.2            |
| 21 | testpy2_7_12nopypi                  | /resources/conda/miniconda3/envs/testpy2_7_12nopypi                  | pybedtools     | 0.8.
1             |
| 22 | testpy2_7_12nopypi                  | /resources/conda/miniconda3/envs/testpy2_7_12nopypi                  | bedtools       | 2.29
.2            |
| 23 | testpy2_7_12pypi                    | /resources/conda/miniconda3/envs/testpy2_7_12pypi                    | pybedtools     | 0.8.
1             |
| 24 | testpy2_7_12pypi                    | /resources/conda/miniconda3/envs/testpy2_7_12pypi                    | bedtools       | 2.29
.2            |
[eperry@corp.gel.ac@phpgridzlogn003 conda_catalogue]$ cd ../
[eperry@corp.gel.ac@phpgridzlogn003 software_catalogues]$ cd R_catalogue/
[eperry@corp.gel.ac@phpgridzlogn003 R_catalogue]$ ./query_catalogue.sh bedtools Biobase
| Library   | vs    | R_VS   |
|----------|------|-------|
[eperry@corp.gel.ac@phpgridzlogn003 R_catalogue]$ ./query_catalogue.sh Biobase
|    | Library   | vs     | R_VS   |
|---:|:----------|:------|:------|
|  0 | Biobase   | 2.50.0 | 4.0.2  |
|  1 | Biobase   | 2.46.0 | 3.6.1  |
|  2 | Biobase   | 2.46.0 | 3.6.2  |
|  3 | Biobase   | 2.50.0 | 4.0.0  |
|  4 | Biobase   | 2.50.0 | 4.0.3  |
|  5 | Biobase   | 2.54.0 | 4.1.0  |
[eperry@corp.gel.ac@phpgridzlogn003 R_catalogue]$
```

VS Code

Applications | [R_catalogue - File Man... | Terminal - eperry@corp.... | [README.md — R_catal... | 10:55

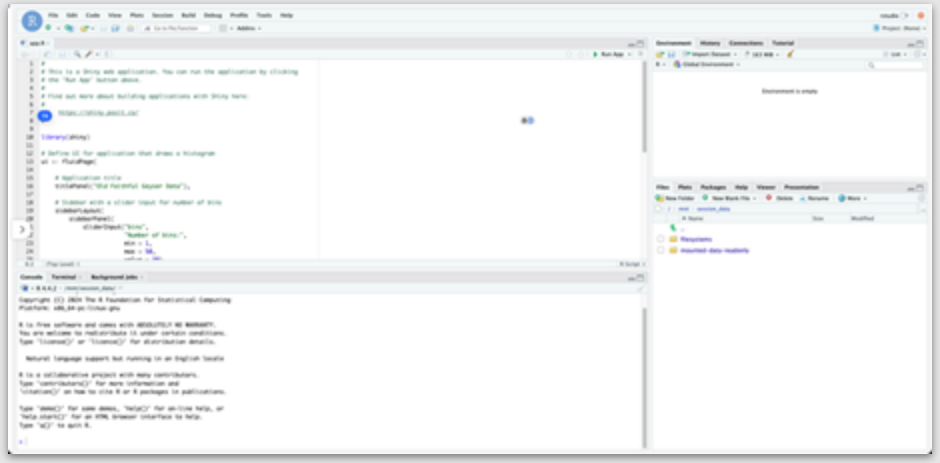
5. Working with RStudio in CloudOS interactive sessions

Command line and GUI



jupyter

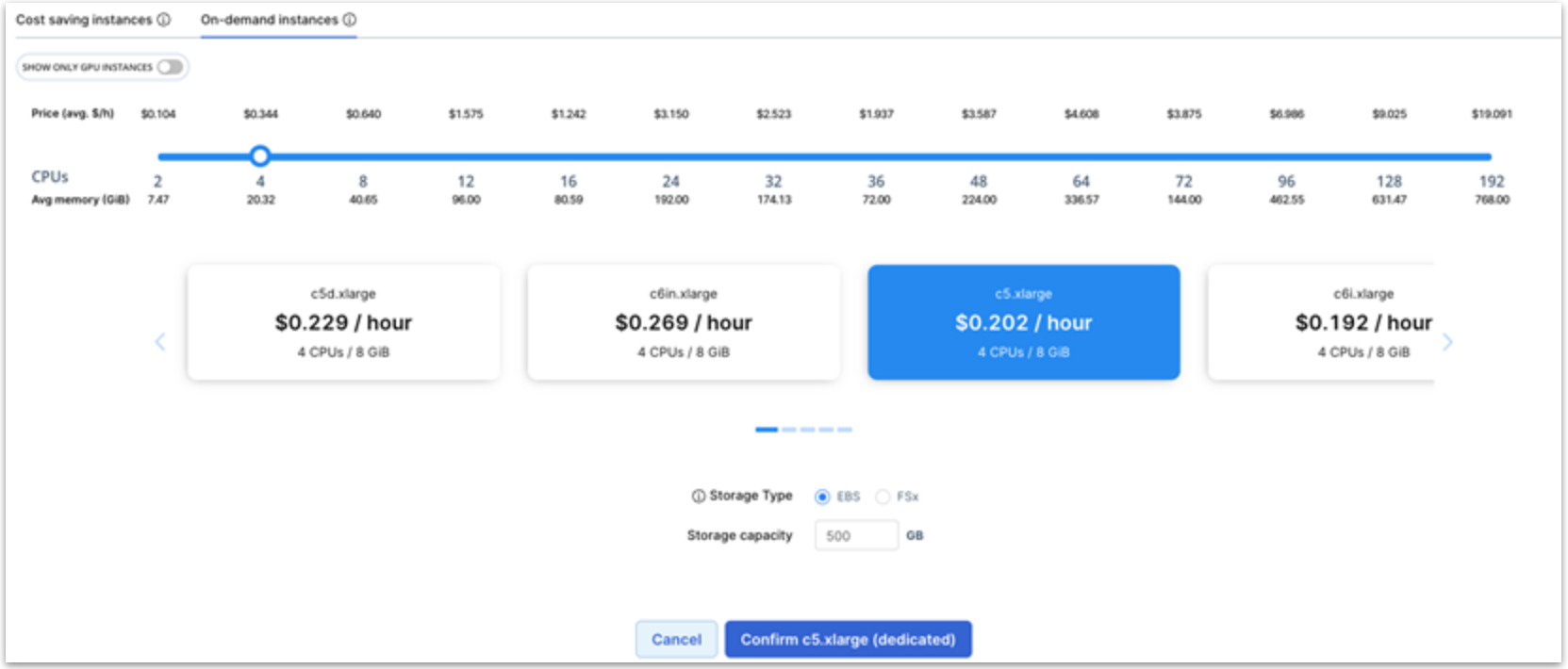
R Studio®



Compute options

Scale your compute resources to match your data size and analysis needs

GPUs available on demand



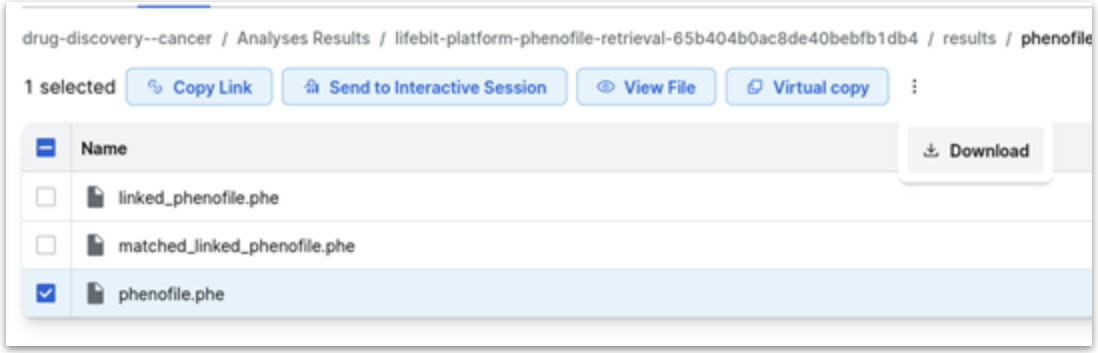
Installing packages

- Full Flexibility with installing packages from CRAN, Bioconductor and Conda
- Save Snapshots of environments

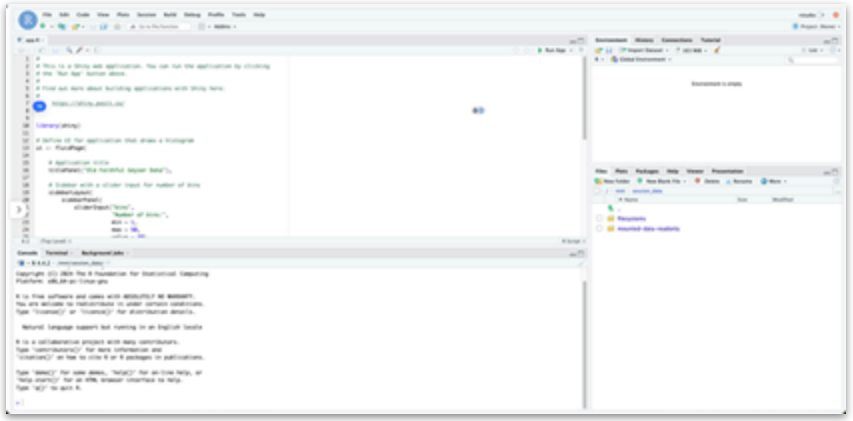
Develop, run and share code

- Write Scripts, Notebooks, Apps and more…
- Collaborate in real time with multiple users
- Share code with others in your workspace

Access your results



Interactive session > File Explorer > Download to VDI > Airlock



CloudOS demo

6. Query clinical data with R

LabKey

- Central database of:
  - Clinical data
  - Results of bioinformatics analysis
  - Locations of genomic files
- Point and click interface
- API

https://re-docs.genomicsengland.co.uk/labkey/

LabKey API

Combine queries between tables

Work in a variety of programming languages (support for Python and R) using SQL queries

Replicate queries between releases and analyses

Work locally and on the HPC

https://re-docs.genomicsengland.co.uk/labkey_api/

LabKey .netrc

- You can access the same data via the LabKey API as you can through other means
- You will need to configure access to the LabKey API with your username and password
    - In your home directory
    - On the HPC
- You do this by editing a file called .netrc

https://re-docs.genomicsengland.co.uk/labkey_api/

LabKey API demo

Camtasia 2024 File Edit Modify Text View Export Window Help Fri 17 Jan 13:52

Amazon WorkSpaces

Computer
eperry's Home
Link to emily
Old Firefox Data
Airlock
CloudOS Academic
CloudOS Discovery Forum
CloudOS Internal
Desktop.Rproj
Document Viewer
Emacs
Ensembl
Firefox
Git GUI
GVim
IGV Browser
IVA
Labkey
LibreOffice 7.6
Open Targets
Panel App
Participant Explorer
R
RE Messages
Research Environment Documentation
Research Registry
RStudio
Terminal Emulator
Text Editor
Visual Studio Code
Welcome Pack
Trash

Genomics england

Applications Places System

Fri Jan 17, 13:52

7. Getting help and questions

Getting help

Check our documentation:
https://re-docs.genomicsengland.co.uk/
Click on the documentation icon in the environment

Contact our Service Desk:
https://jiraservicedesk.extge.co.uk/plugins/servlet/desk

Training sessions

3rd Tuesday every month

Introduction to the RE

18/3 15/4 20/5

22/7 19/8 16/9

Materials from past training all online

https://re-docs.genomicsengland.co.uk/upcoming/

Training sessions

8/4	Working with python in the RE
13/5	Building cancer cohorts and survival analysis
10/6	Building rare disease cohorts with matching controls
8/7	Finding participants based on genotypes
9/9	Getting medical records for participants
14/10	What tools and workflows should I use to fulfil an overall goal?

https://re-docs.genomicsengland.co.uk/upcoming/

Materials from past training all online

Feedback

Thank you

Visit: https://re-docs.genomicsengland.co.uk/