Dear Students,

Here is some information essential for the first 4 lessons of the Processing and interpretation of neuroscience data.

# Important – to do prior first lecture

We kindly request all students to install Oracle's VirtualBox, a free software compatible with most platforms using x86/x64 CPUs. You can download it here:

* https://www.virtualbox.org/wiki/Downloads.

For guidance on using VirtualBox, please refer to the following resources:

* https://www.virtualbox.org/manual/

https://kb.hosting.com/docs/configuring-a-virtual-machine-for-local-testing, and

* https://www.youtube.com/watch?v=nvdnQX9UkMY.

Additionally, please download the prepared Virtual Machine. Note that it is approximately 10GB in size, so ensure you have sufficient space (20GB on a drive). VM download link:

* https://filesender.aco.net/?s=download&token=2b1e953a-5a08-484c-9401-680cad8364fe

Once VirtualBox is installed and You download the files, please open the VirtualBox and import the OVA file.

Some guide: https://www.youtube.com/watch?v=PJdsjpZmMMo

Pass for VM is the same as user name. Please check if it starts, we can do some troubleshooting at the beginning.

## MAC users:

Please verify your processor. If it is an ARM version, such as the M1, rather than an x86, VirtualBox will not function for you. In this case, please contact us so we can guide you on how to install the necessary materials. Alternatively, consider collaborating with someone who has a Linux or Windows machine.

# Additional materials

Please have a look at the provided links for additional materials, as well as the attached cheat sheets. This is provided together in the VM link.

# Lectures overview:

## 21.10

* Introduction to R
* Practical’s, basic data assessment and analysis: means, medians, t-test etc.
* Visual analytics: basic plots

## 28.10

* Introduction to bulk sequencing and analysis methods
* Dimensional reduction: PCA and UMAP
* Clustering methods: k-means, k-nn

## 4.11

* Introduction to Seurat
* Preprocessing of scRNAseq

## 11.11

* Clustering in Seurat
* DEG interpretation

# Additional materials

## Rstudio usage

* General introduction: https://www.youtube.com/watch?v=eA6kf4yJXGw&list=WL&index=26&t=332s
* Basics: https://www.youtube.com/watch?v=FY8BISK5DpM
* Looking at data: https://choonghyunryu.github.io/dlookr/
* data explorer library: https://cran.r-project.org/web/packages/DataExplorer/vignettes/dataexplorer-intro.html#qq-plot
* tables: https://www.danieldsjoberg.com/gtsummary/

## GGplot in R

* Datacamp: https://www.youtube.com/watch?v=YxKr2a-Y1WE&list=PLjgj6kdf\_snaBCTJEi53DvRVgOuVbzyku
* Hadley talk: https://www.youtube.com/watch?v=ZdPNBF6GWBw

## Tidyverse

* general intro: https://www.youtube.com/watch?v=KsBBRHAgAhM
* Dplyr: https://www.youtube.com/watch?v=Bg4qxVNaDck
* Data wrangling in Dplyr: https://www.youtube.com/watch?v=XcK4chr2jws
* intro: youtube.com/watch?v=FV5mPG1uLJk&t=37s&pp=ygUSdGVhY2hpbmcgdGlkeXZlcnNl

## General statistics

* statistics: https://www.youtube.com/watch?v=dYJLUvo0Q6g
* variable selection: https://www.youtube.com/watch?v=XMZ\_0aloVAc&list=PL6Gy8EsbrMF-lsnMywDYt73yrzm5HaGtG

## Dimensional reduction

* Eigenvectors: https://www.youtube.com/watch?v=9CT0jnem4vM&t=18s
* Eigenvectors and values math: https://www.youtube.com/watch?v=JtcNe--fsyA&t=5s
* t-SNE: https://www.youtube.com/watch?v=Bl85EfTiGj0
* t-SNE Explained: https://www.youtube.com/watch?v=NEaUSP4YerM&list=PL6Gy8EsbrMF-lsnMywDYt73yrzm5HaGtG&index=7
* PCA: https://www.youtube.com/watch?v=FgakZw6K1QQ&list=PL6Gy8EsbrMF-lsnMywDYt73yrzm5HaGtG&index=6
* PCA - math; https://www.youtube.com/watch?v=S51bTyIwxFs&t=909s
* PCA - standardization: https://www.youtube.com/watch?v=dh8aTKXPKlU&t=601s
* PCA - interpretation: https://www.youtube.com/watch?v=BiuwDI\_BbWw
* UMAP math: https://www.youtube.com/watch?v=jth4kEvJ3P8&list=WL&index=18

## Clustering

* k-means: https://www.youtube.com/watch?v=4E\_DFMt60rc
* k-NN: https://www.youtube.com/watch?v=48RqX4HTtCE

## Seurat

* Basics of Seurat: https://www.youtube.com/watch?v=G3Cg7vGpctg&list=PL6Gy8EsbrMF-lsnMywDYt73yrzm5HaGtG&index=44
* https://biostatsquid.com/seurat-objects-explained