Sanger Advanced Flow Cytometry Course

Data Analysis

Part 1

wellcome sanger institute

@sangercytometry

Flow Cytometry Data Analysis Using R

Bridging the gap between the wet and dry labs

1 hour whistle stop tour of R

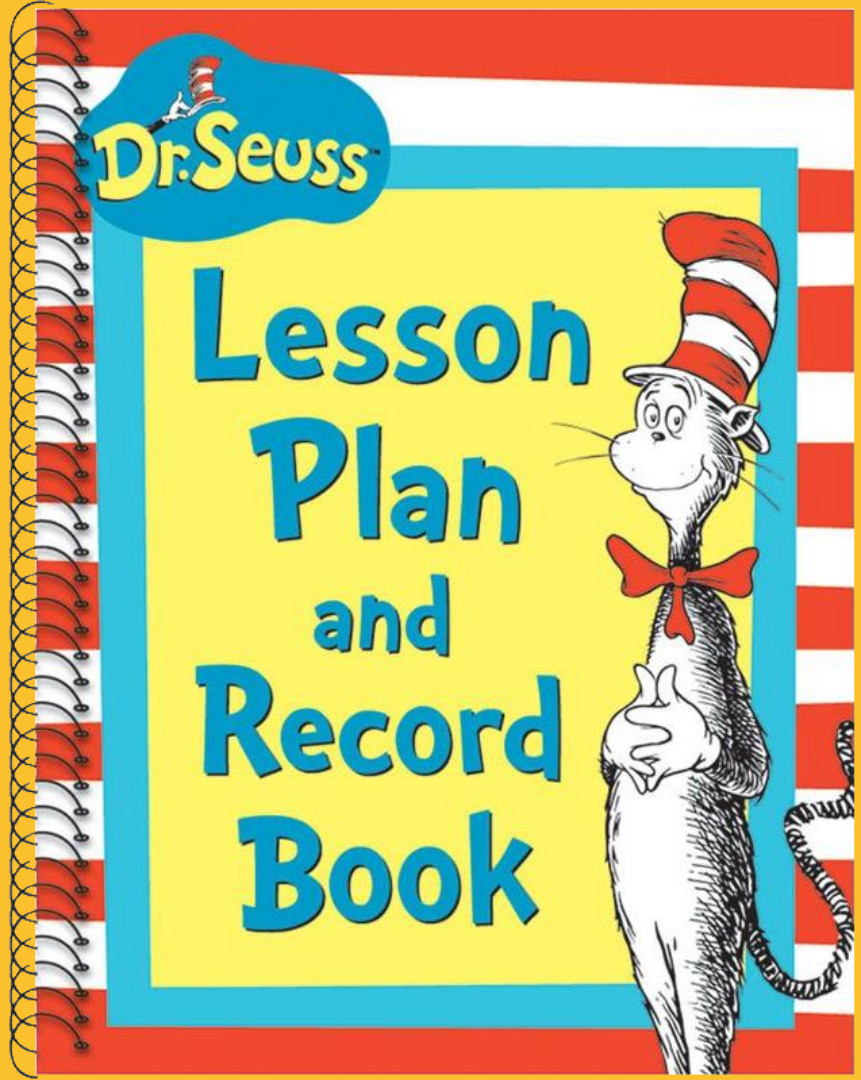
2 hours of applying R to flow cytometry

This is NOT a full on R course, it is a “reveal” of how to analyse flow cytometry data in R.

There are 1000’s of R courses online and the University of Cambridge provides several introductory courses: https://www.training.cam.ac.uk/bioinformatics/event/2601305

wellcome sanger institute

@sangercytometry



- What is R
- Using Rstudio
- Learn R basics

Tea break

- Quick FlowJo analysis (Chris only)
- What is in a FCS file
- How to load a FCS file into R
- Data transformation
- Basic statistics
- Basic figures
- Gating
- Data clean up
- Auto-gating
- Workflows
- TSNE and flowSOM
- Analyse your data

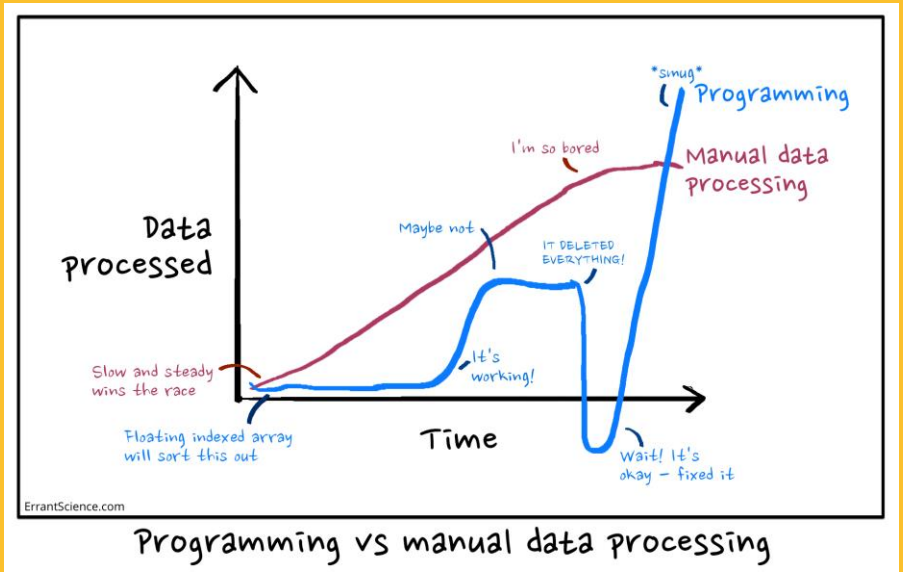
wellcome sanger institute

@sangercytometry

R basics

wellcome sanger institute

@sangercytometry



Programming vs manual data processing

Programming languages allow you to write a set of text instructions which a computer understands and can execute

This allows you

Automate tasks

Do complex analysis

It is more flexible and customizable than using software programs

“R is a language and environment for statistical computing and graphics.”

R is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes:

- an effective data handling and storage facility,
- a suite of operators for calculations on arrays, in particular matrices,
- a large, coherent, integrated collection of intermediate tools for data analysis,
- graphical facilities for data analysis and display either on-screen or on hardcopy, and
- a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.

- There are 13170 packages on CRAN (Comprehensive R Archive Network).
- There are 1562 packages on Bioconductor.
- 54 specific to flow cytometry.

wellcome sanger institute

@sangercytometry

Overview

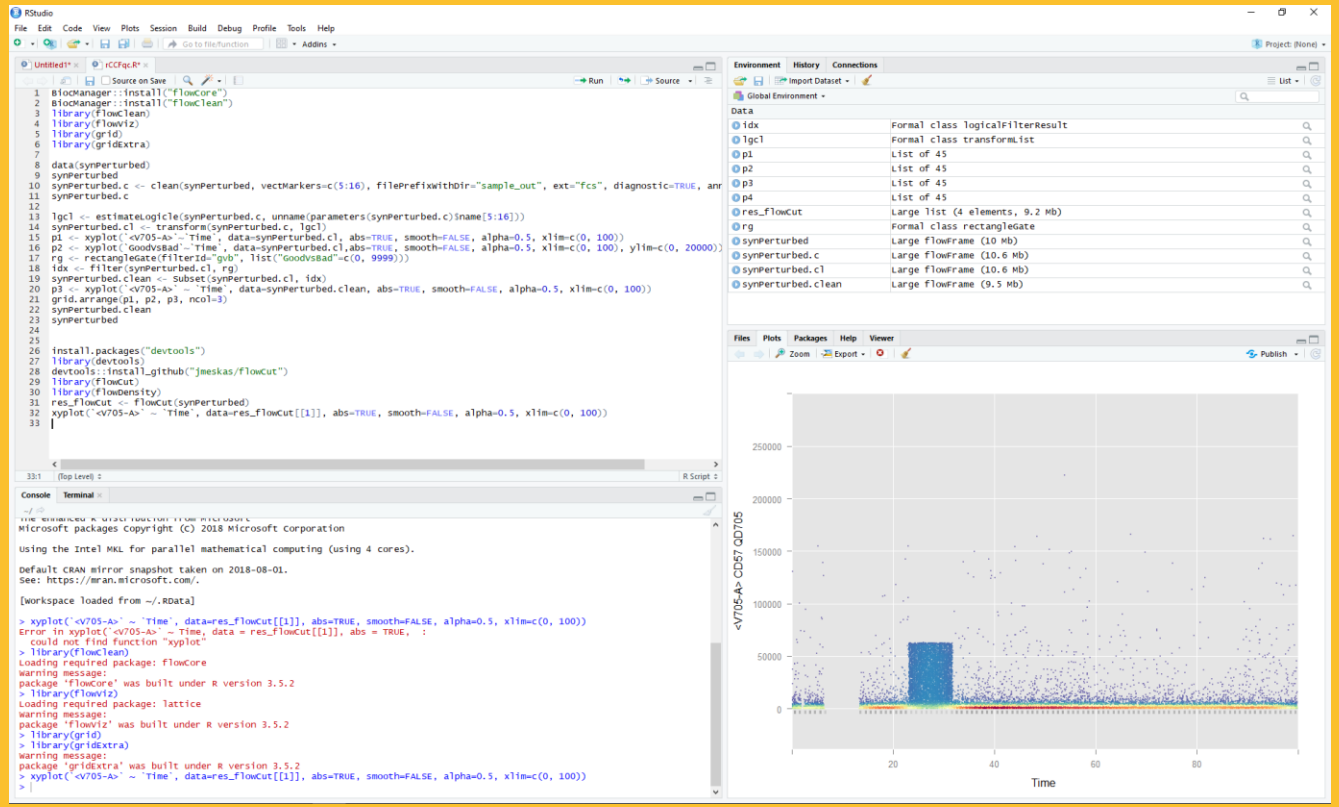
- Interface
- Entering basic commands
- Using scripts
- Accessing help
- Functions
- Plotting
- Importing data

Code editor

Console

Workspace

Files, plots, packages, and help



wellcome sanger institute

@sangercytometry

Work through the handout

wellcome sanger institute

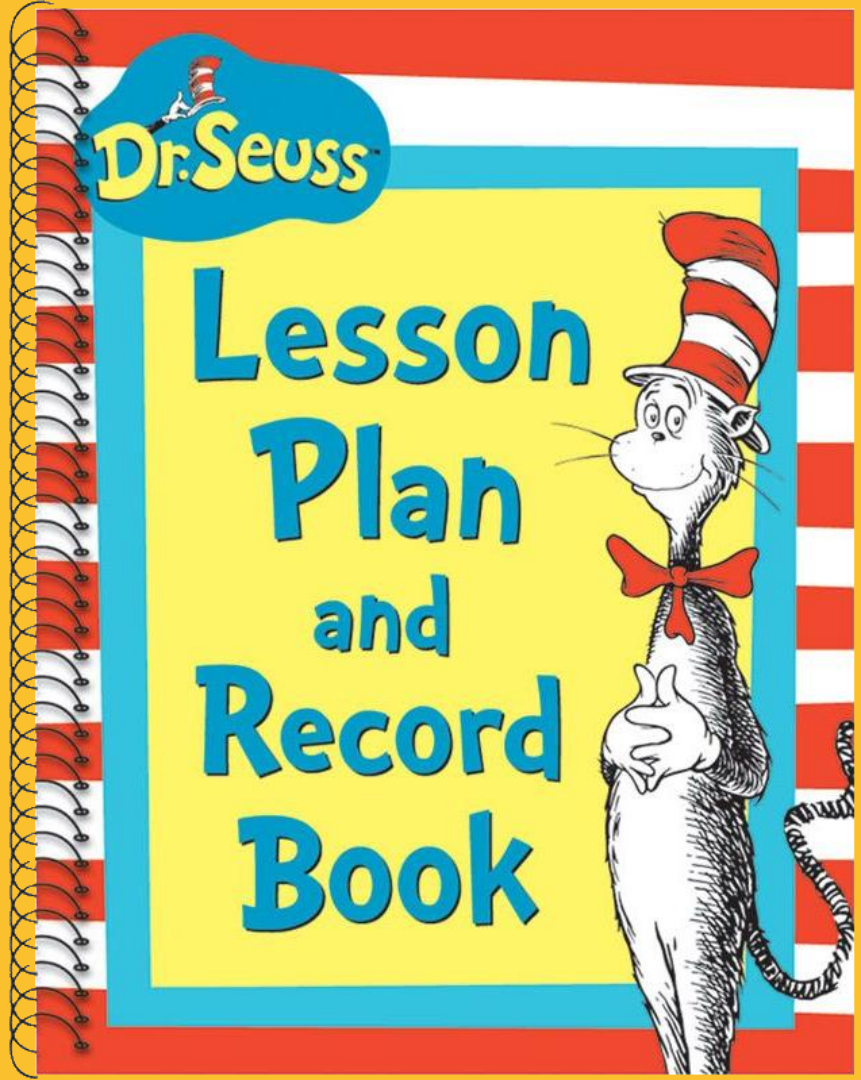
@sangercytometry

Sanger Advanced Flow Cytometry Course

Flow cytometry data analysis
Part 2

wellcome sanger institute

@sangercytometry



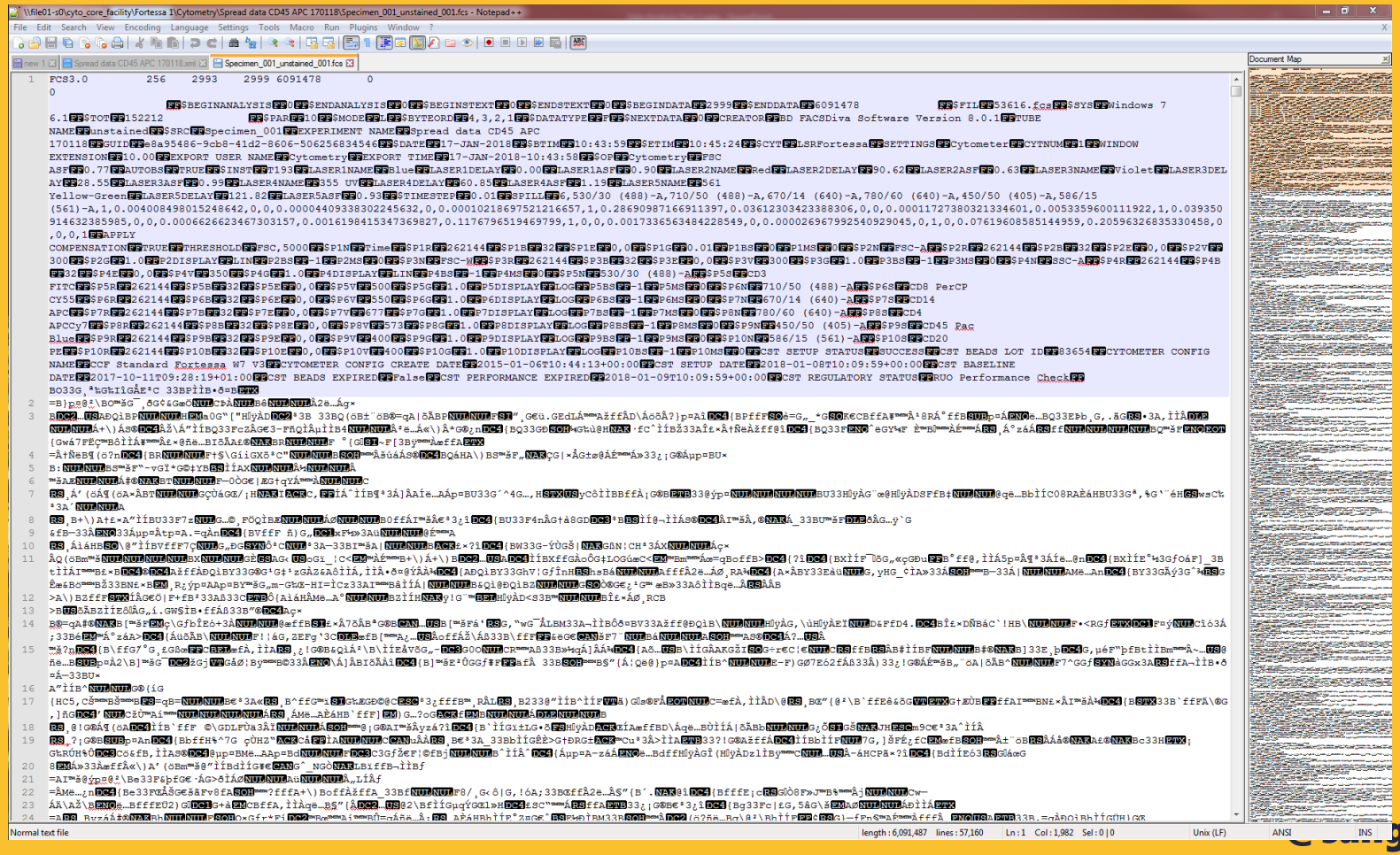
- Install R and RStudio (IDE)
- What is R
- Using Rstudio
- Learn R basics

Tea break

- Quick FlowJo analysis (Chris only)
- What is in a FCS file
- How to load a FCS file into R
- Data transformation
- Basic statistics
- Basic figures
- Gating
- Data clean up
- Auto-gating
- Workflows
- TSNE and flowSOM
- Analyse your data

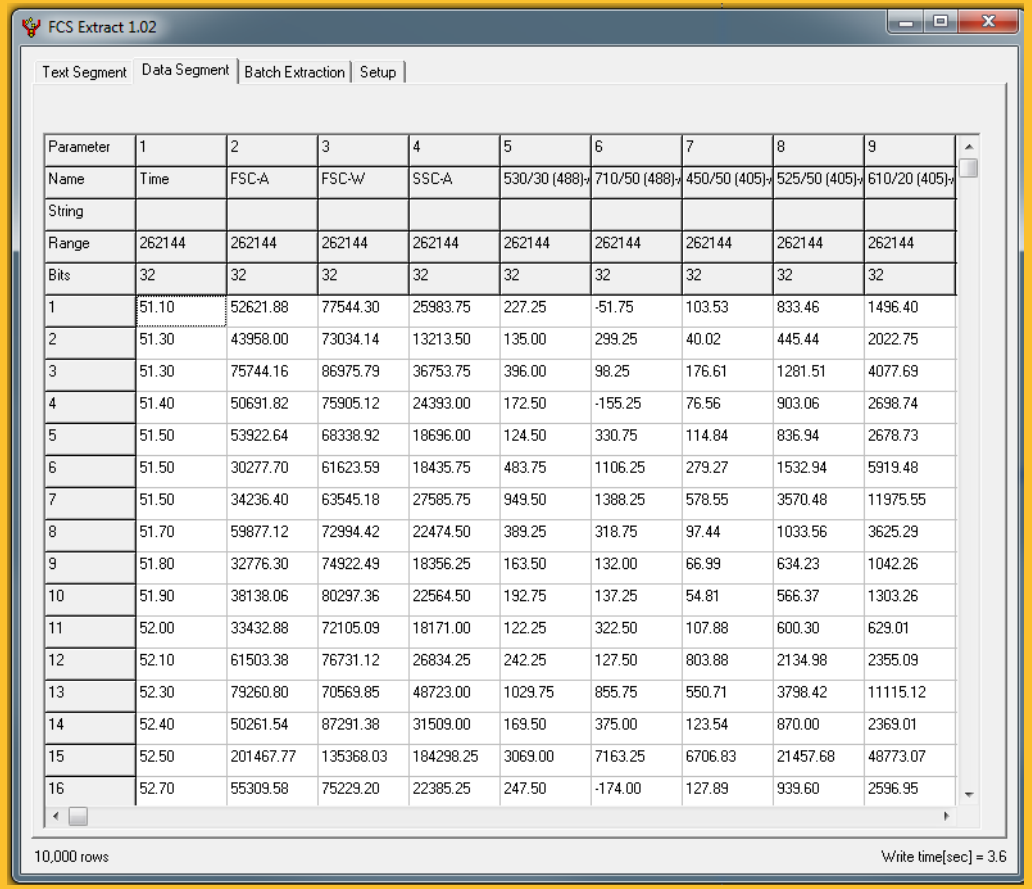
wellcome sanger institute

@sangercytometry



wellcome sanger institute

gercytometry



Parameter	1	2	3	4	5	6	7	8	9
Name	Time	FSC-A	FSC-W	SSC-A	530/30 (488)-A	710/50 (488)-A	450/50 (405)-A	525/50 (405)-A	610/20 (405)-A
String									
Range	262144	262144	262144	262144	262144	262144	262144	262144	262144
Bits	32	32	32	32	32	32	32	32	32
1	51.10	52621.88	77544.30	25983.75	227.25	-51.75	103.53	833.46	1496.40
2	51.30	43958.00	73034.14	13213.50	135.00	299.25	40.02	445.44	2022.75
3	51.30	75744.16	86975.79	36753.75	396.00	98.25	176.61	1281.51	4077.69
4	51.40	50691.82	75905.12	24393.00	172.50	-155.25	76.56	903.06	2698.74
5	51.50	53922.64	68338.92	18696.00	124.50	330.75	114.84	836.94	2678.73
6	51.50	30277.70	61623.59	18435.75	483.75	1106.25	279.27	1532.94	5919.48
7	51.50	34236.40	63545.18	27585.75	949.50	1388.25	578.55	3570.48	11975.55
8	51.70	59877.12	72994.42	22474.50	389.25	318.75	97.44	1033.56	3625.29
9	51.80	32776.30	74922.49	18356.25	163.50	132.00	66.99	634.23	1042.26
10	51.90	38138.06	80297.36	22564.50	192.75	137.25	54.81	566.37	1303.26
11	52.00	33432.88	72105.09	18171.00	122.25	322.50	107.88	600.30	629.01
12	52.10	61503.38	76731.12	26834.25	242.25	127.50	803.88	2134.98	2355.09
13	52.30	79260.80	70569.85	48723.00	1029.75	855.75	550.71	3798.42	11115.12
14	52.40	50261.54	87291.38	31509.00	169.50	375.00	123.54	870.00	2369.01
15	52.50	201467.77	135368.03	184298.25	3069.00	7163.25	6706.83	21457.68	48773.07
16	52.70	55309.58	75229.20	22385.25	247.50	-174.00	127.89	939.60	2596.95

wellcome sanger institute

@sangercytometry

Let’s load some data ☺

wellcome sanger institute

@sangercytometry

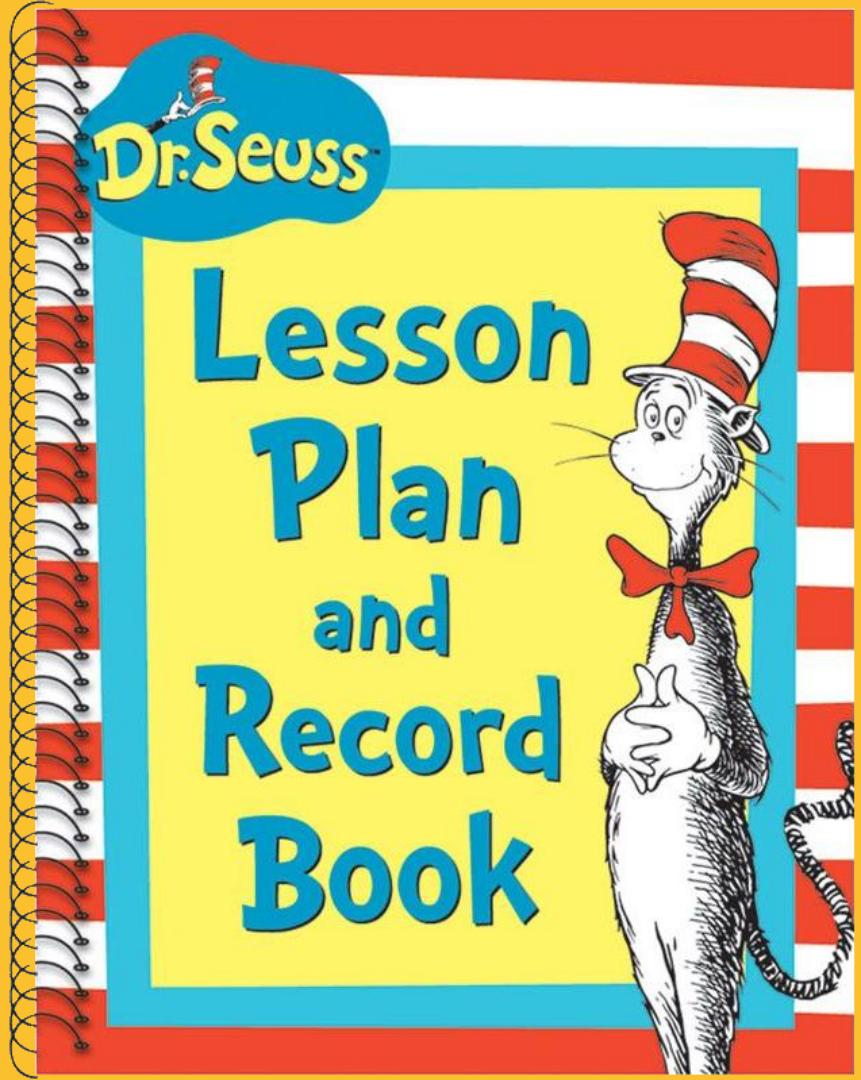
Sanger Advanced Flow Cytometry Course

Data analysis

Part 3

wellcome sanger institute

@sangercytometry



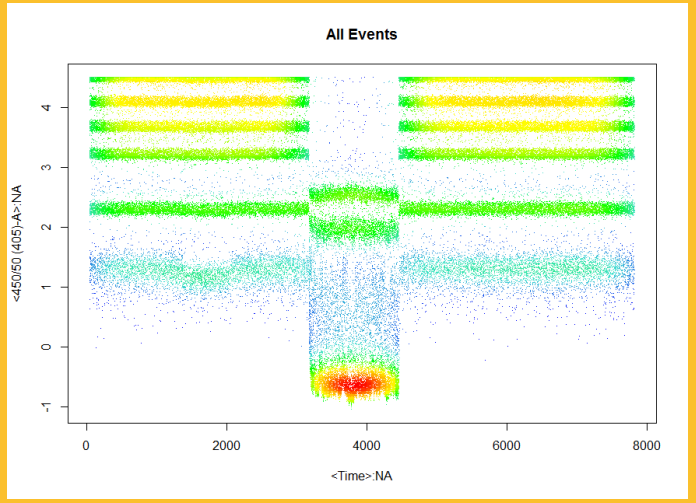
- Install R and RStudio (IDE)
- What is R
- Using Rstudio
- Learn R basics

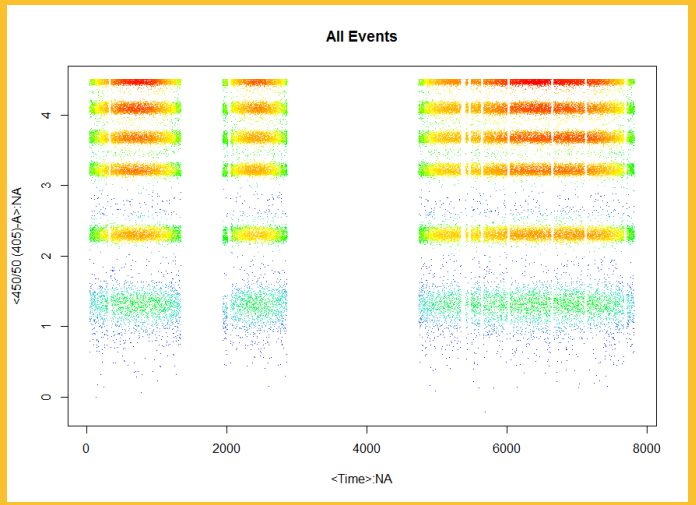
Tea break

- What is in a FCS file
- Quick FlowJo analysis (Chris only)
- How to load a FCS file into R
- Data transformation
- Basic statistics
- Basic figures
- Gating
- Data clean up
- Auto-gating
- Workflows
- TSNE and flowSOM
- Analyse your data

wellcome sanger institute

@sangercytometry





Data Clean-Up

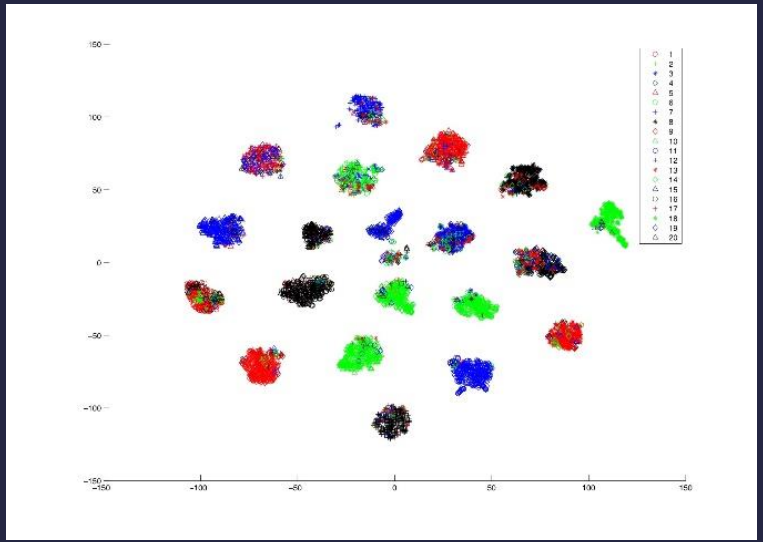
Time vs Fluorescence

flowCut
flowClean
flowAI
flowQC

wellcome sanger institute

@sangercytometry

Laurens van der Maaten & Geoffrey Hinton



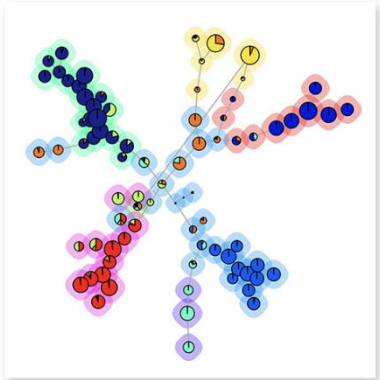
T-SNE

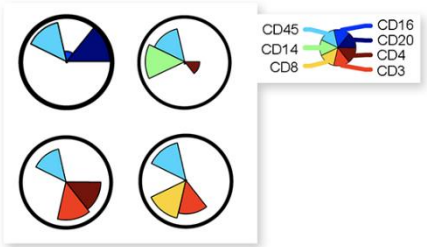
wellcome sanger institute

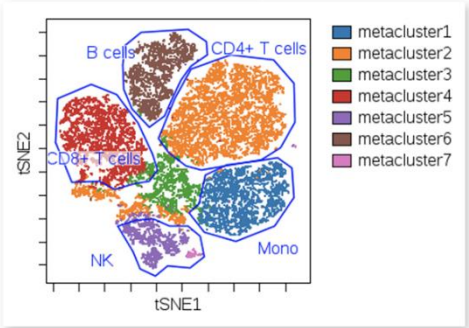
@sangercytometry

Sofie Van Gassen

flowSOM

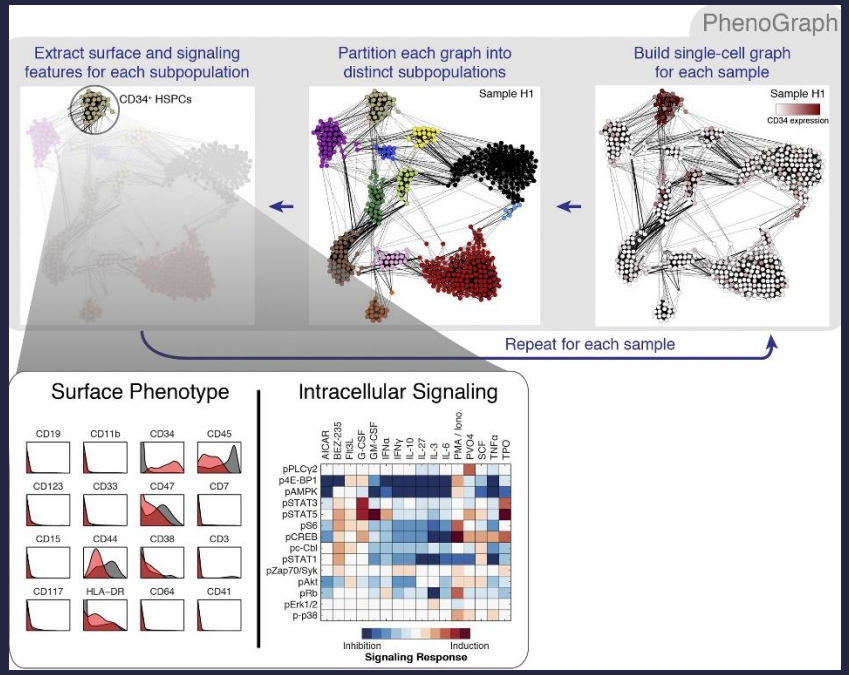


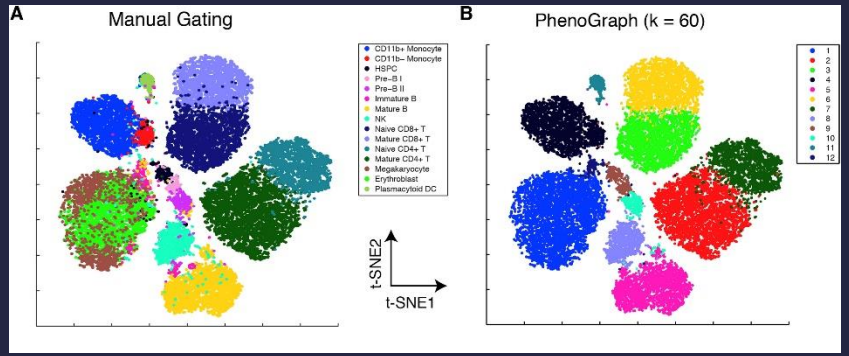




wellcome sanger institute

@sangercytometry

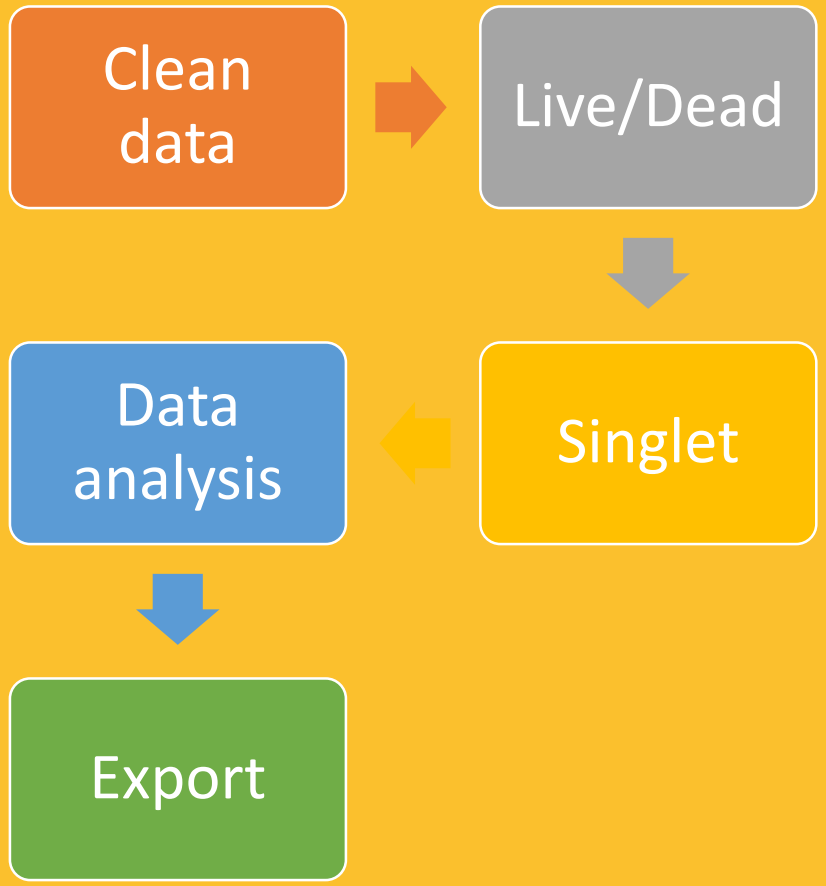




Phenograph

wellcome sanger institute

@sangercytometry



Workflow

I am only taking easy questions

wellcome sanger institute

@sangercytometry