



Sinh hoạt Khoa học hàng tuần tại bệnh viện Thông Nhât



Data visualization (Trực quan hóa dữ liệu)

Lecture 1: Introduction

Nov 19 2025

Phuc-Loi Luu, PhD

Email: loilp@bvtm.org.vn

Zalo: 0901802182



Viện Nghiên cứu Ứng dụng Khoa học Sức khỏe và Lão hóa – Bệnh viện Thông Nhât

Organization of the lecture

 luuloi	Update README.md	cfeaf4e · now	 7 Commits
 Book	Add files via upload	3 minutes ago	
 Lecture_00	Create Note.txt	4 hours ago	
 README.md	Update README.md	now	

 README  

Data Visualization with R

Starting from 13:30 to 15 Wed (19/11/2025 - 25/03/2026)

Location: Auditorium at level 6th

[00.0 Introduction to Data Visualizaion \[Loi\] - 19/11/2025](#)

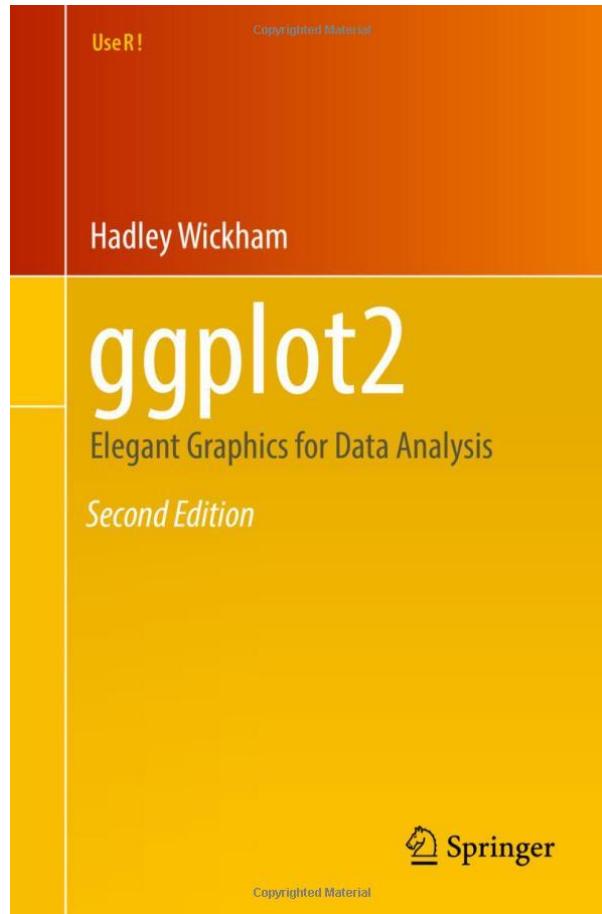
- [PDF](#)
- [Book](#)

[00.1 Practice [Dan] - 02/11/2025]

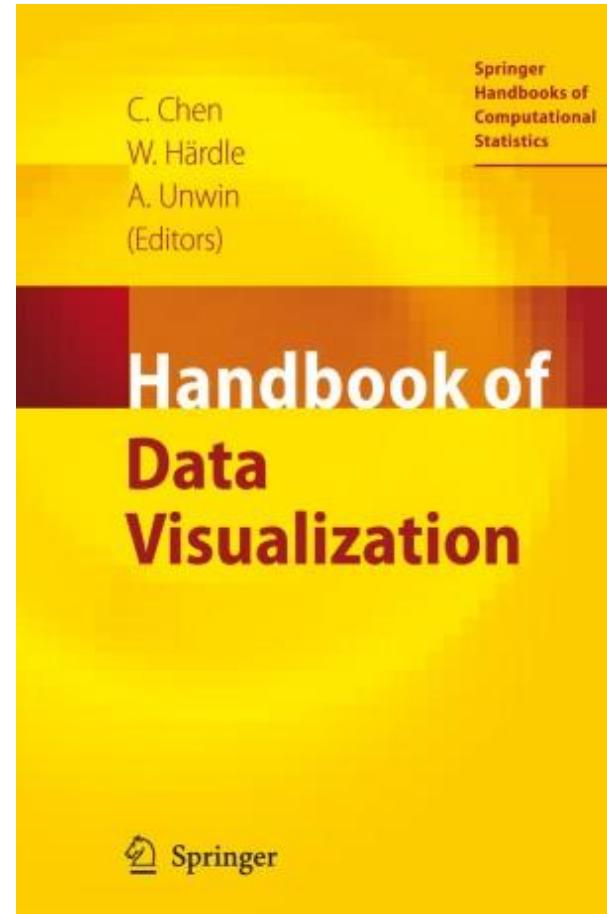
- [PDF](#)
- Installation R & Rstudio
- Basic R
- Introduction to Google Colab
- Practice:  [Open in Colab](#)

https://github.com/luuloi/Data_Visualization_With_R/

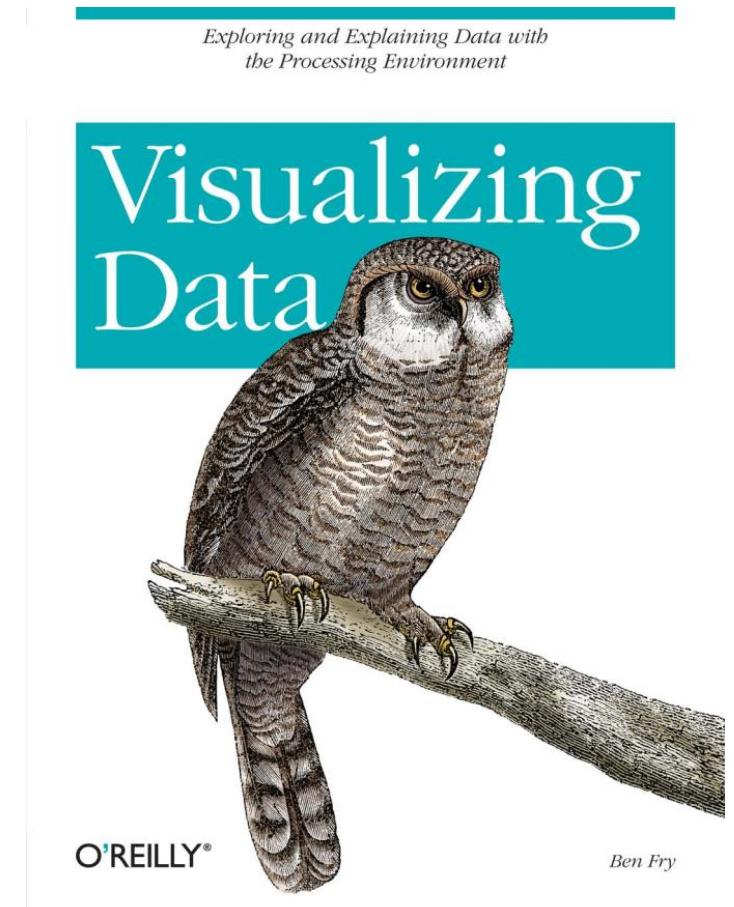
Books, slides, exercises and solutions to exercises



<https://ggplot2-book.org/index.html>
(GGP)



<https://link.springer.com/book/10.1007/978-3-540-33037-0>
(HDV)



<https://www.iaacblog.com/programs/books-about-data-visualization/>
(VD)

Roadmap of today lecture

- Overview of basic plots
- Types of plots
- Visualizing Images in pathology
- Why Do We Visualize Data?
- Student Projects

Basic Plots

- 1. Table
- 2. Bar/pie/waterfall plot
- 3. Boxplot
- 4. Histogram/density/violin
- 5. Variant of boxplot/histogram/density/violin/cloud
- 6. Parallel plot and its variants
- 7. Heatmap
- 8. Dendogram
- 9. Venn diagram
- 10. Upset plot
- 11. Forest plot
- 12. Survival plot
- 13. Volcano and MA plot
- 14. Scatter plot
- 15. Dot plot
- 16. Time series plot
- 17. Spatial/ Map/Location plot
- 18. Hexabin/contour plot
- 19. ROC curve
- 20. RMSE
- 21. Circos plot
- 22. Hexanbin plot
- 23. Maths plot
- 24. Radar plot

Pie + Bar + water fall + time plot

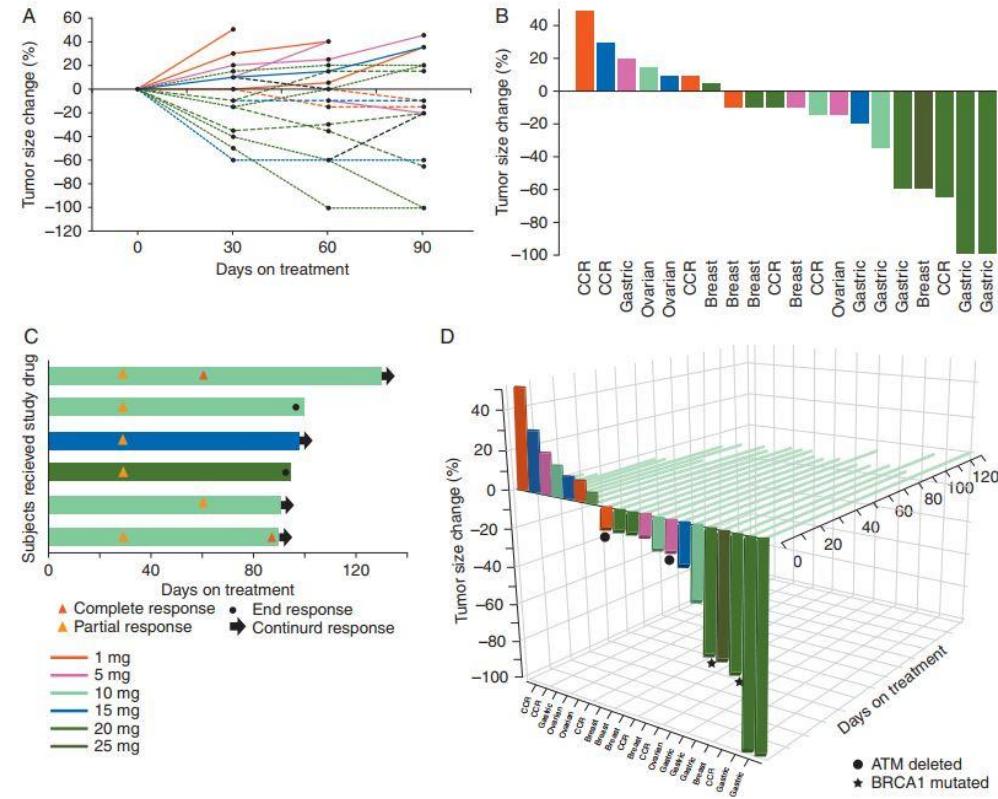
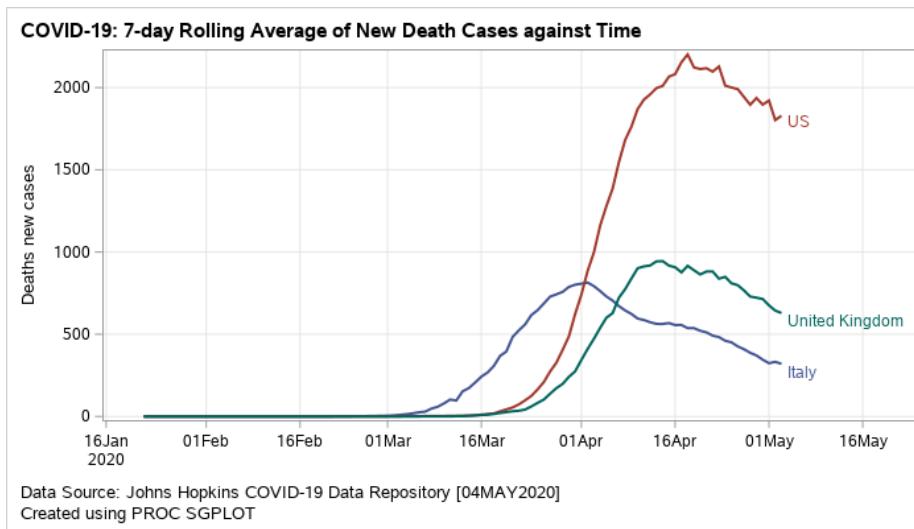
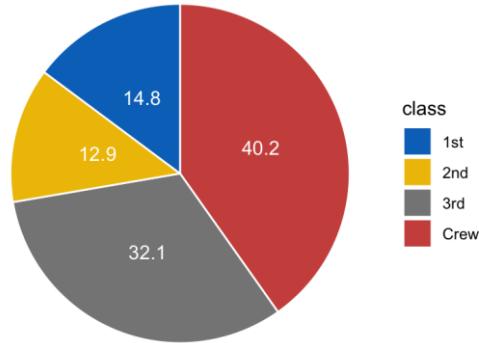
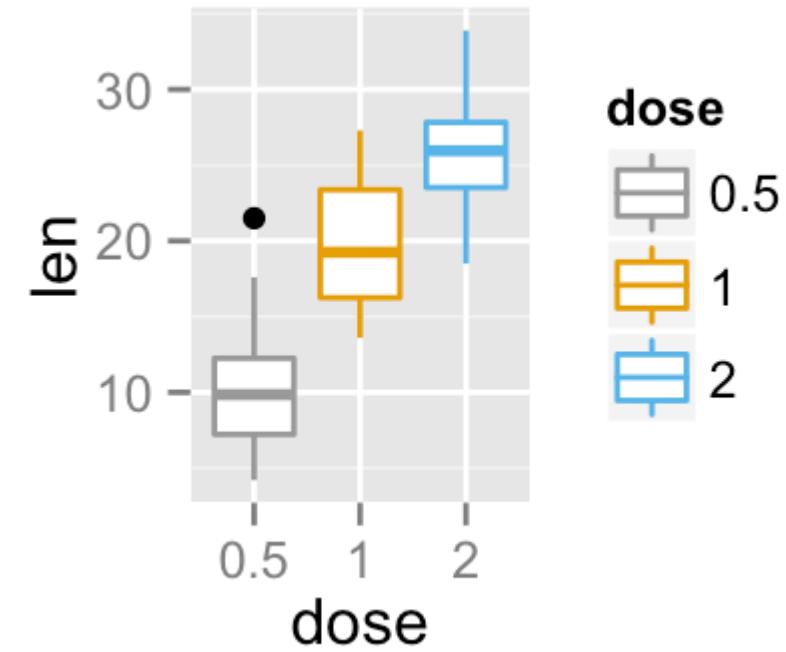
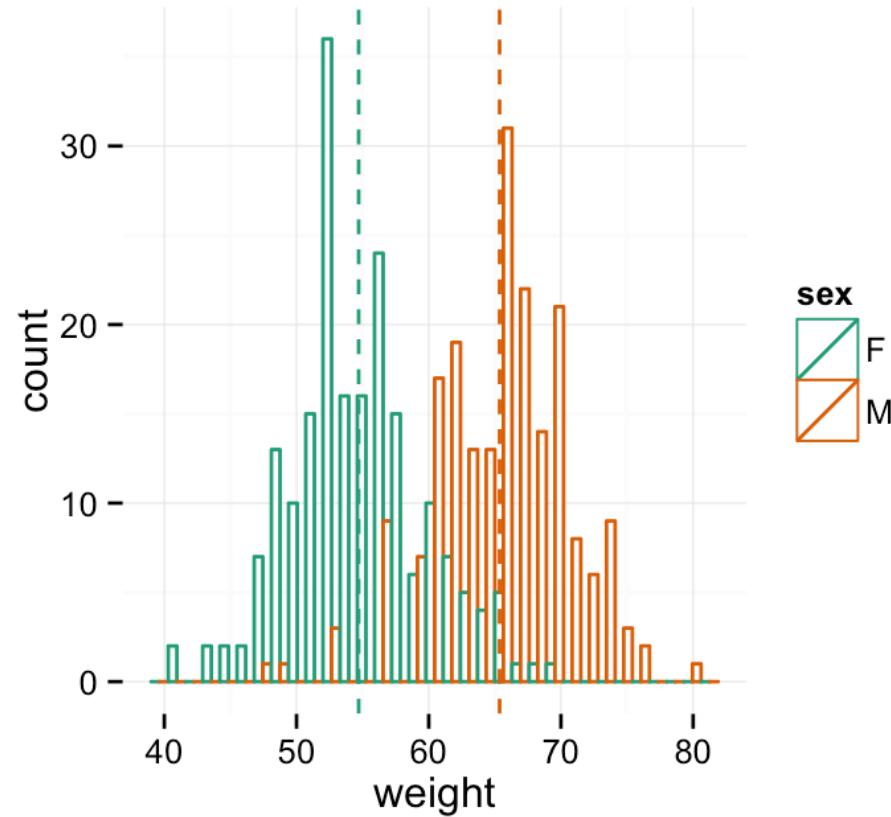
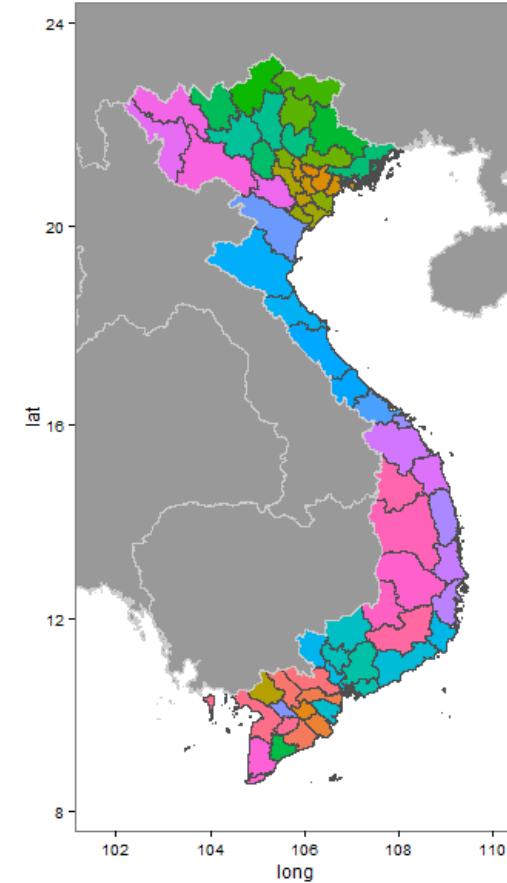
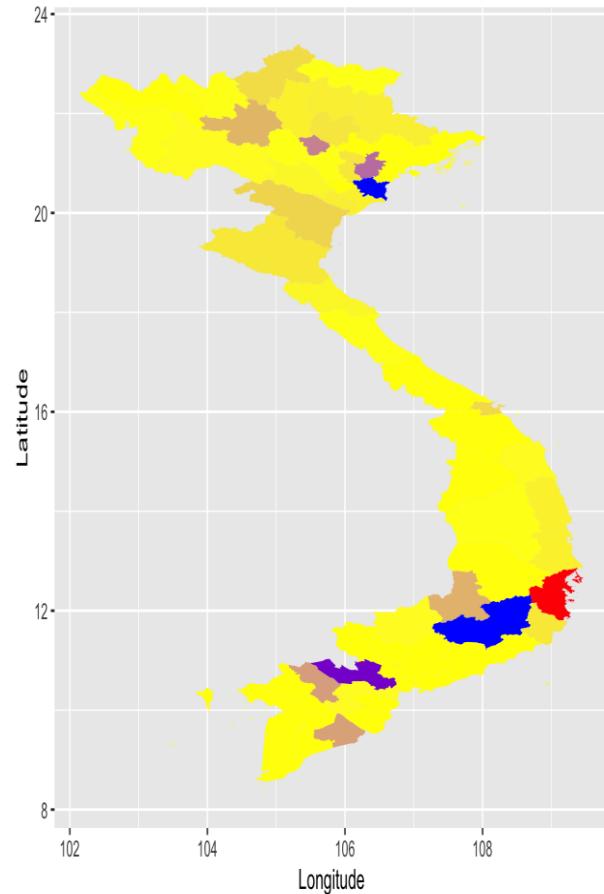


Figure 1 (A) Spider plot, (B) classical (2D) waterfall plot, (C) swimmer plot and (D) 3D waterfall plot.

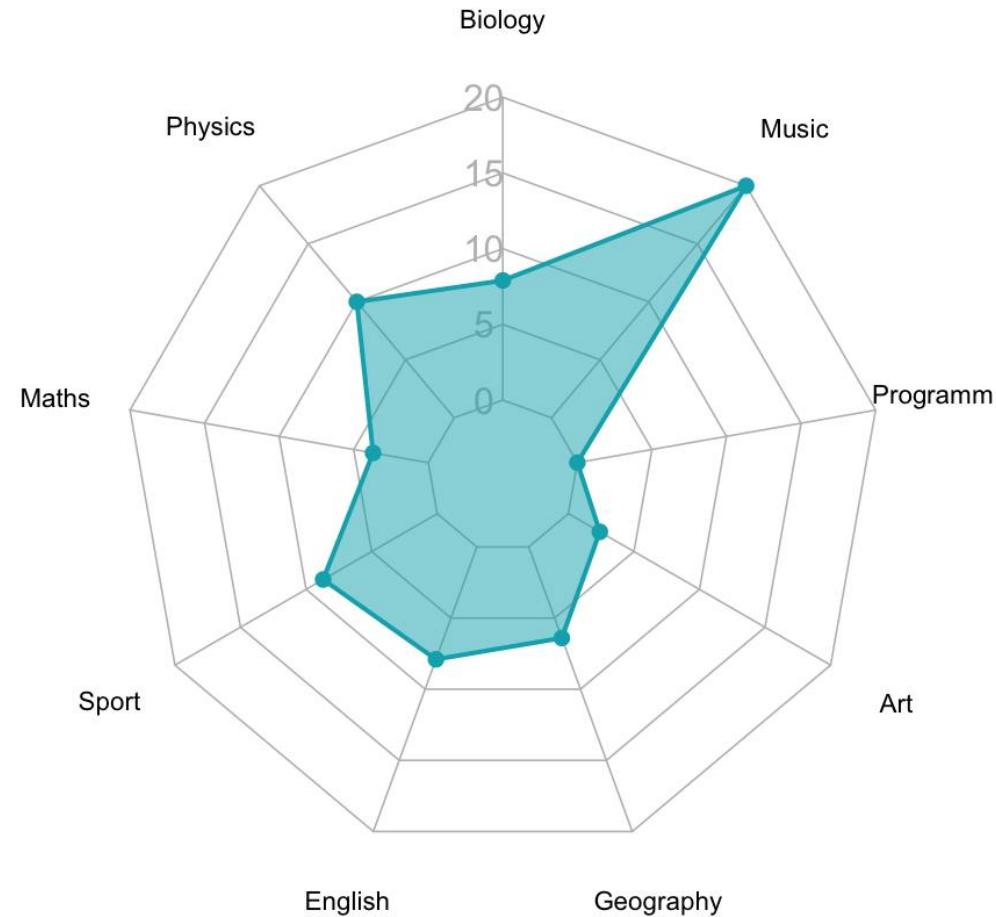
Boxplot and Histogram



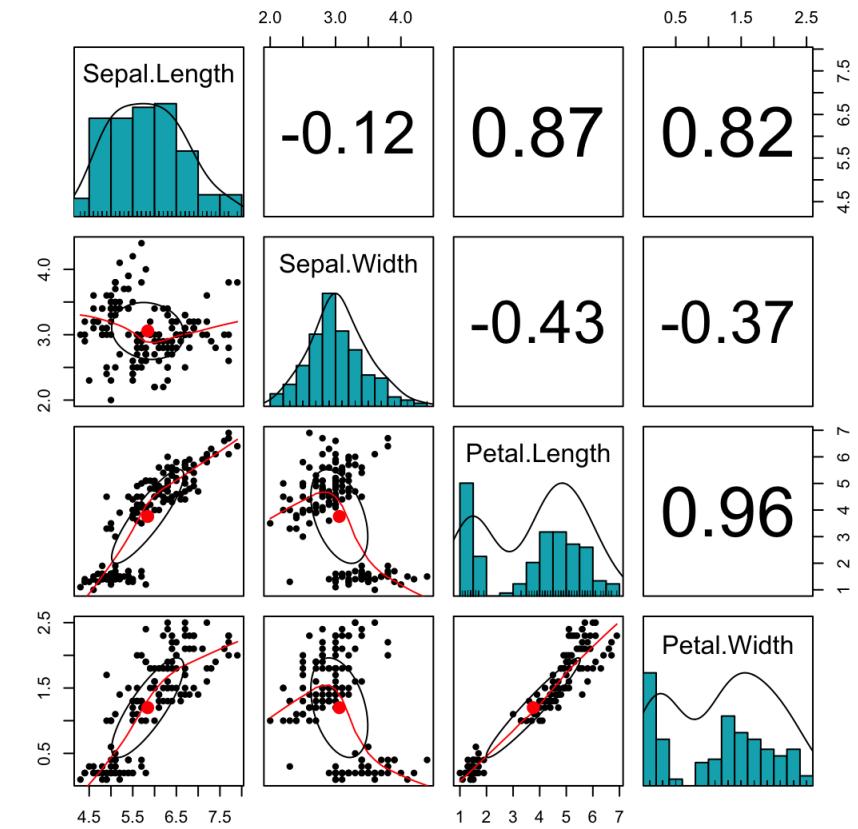
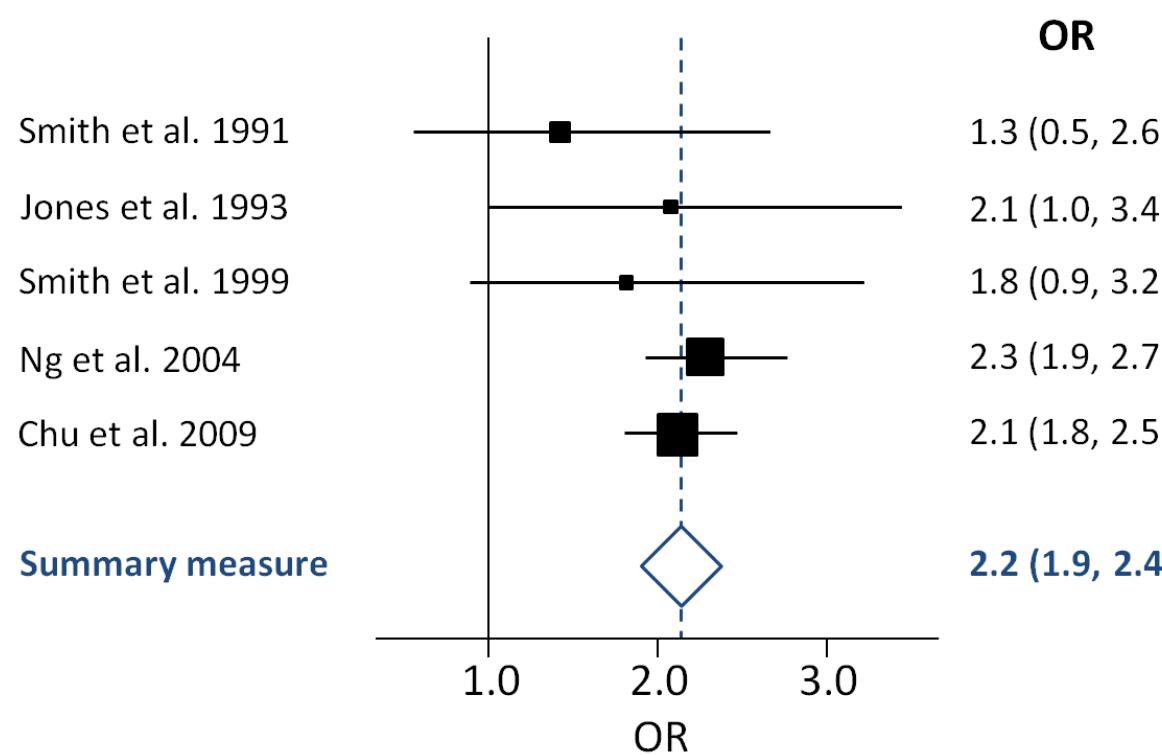
Map and location plots



Radar plot

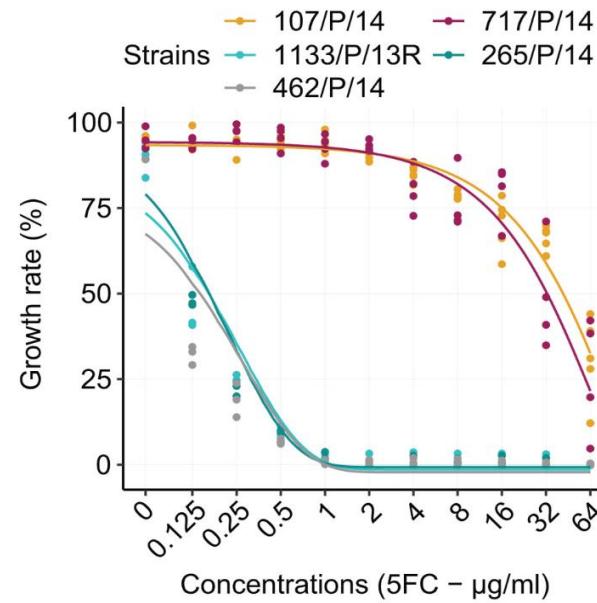


Forest plot vs matrix plot

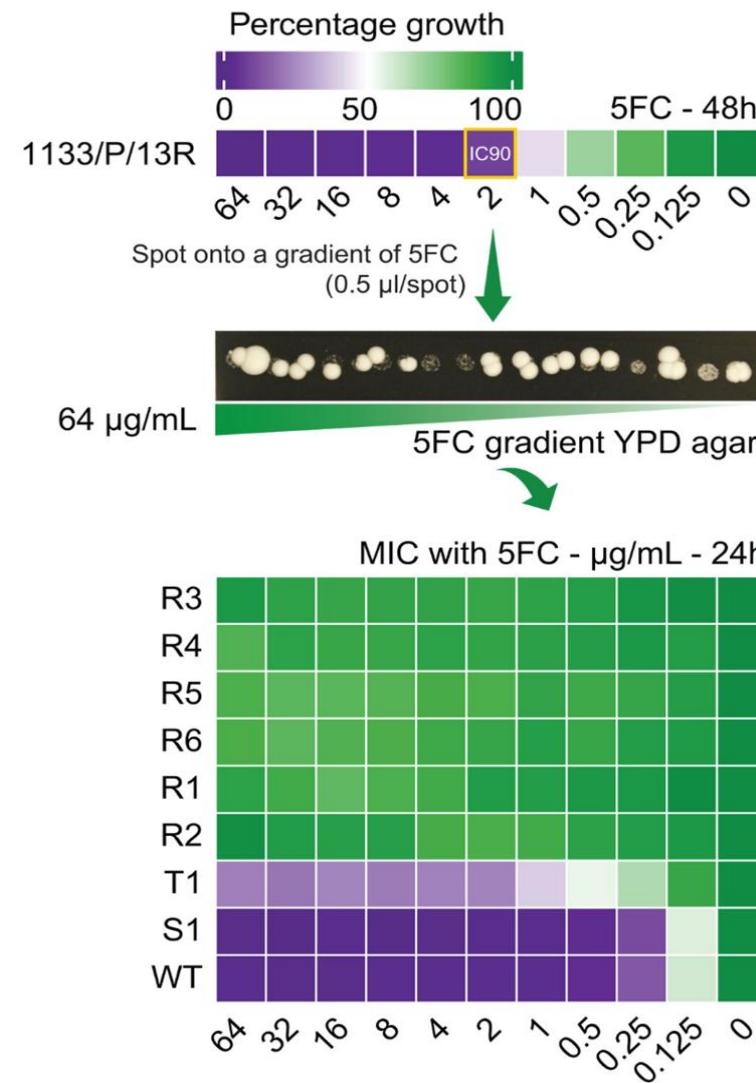


In vitro evolution of 5FC resistance in *C. auris*

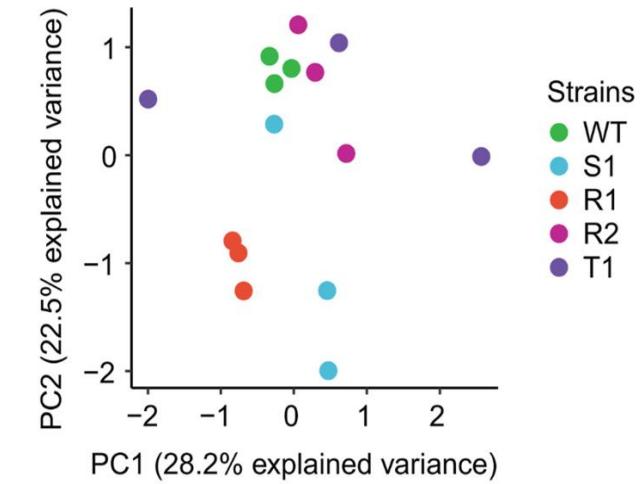
A



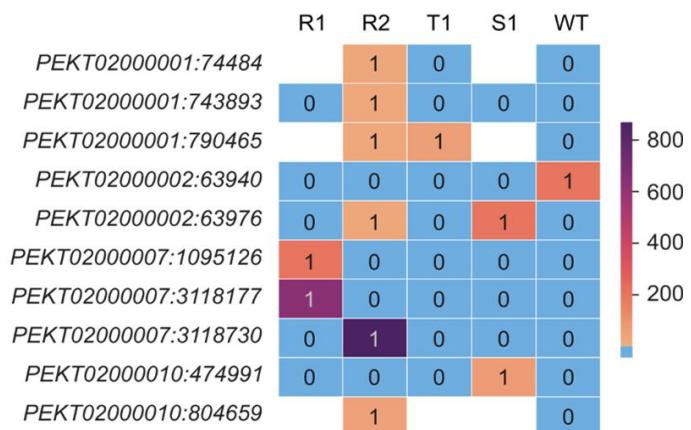
B



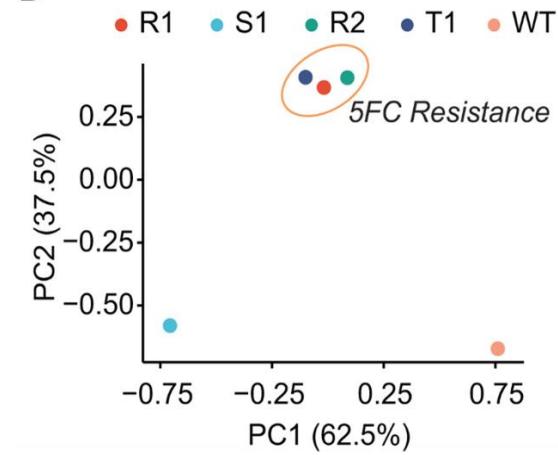
C



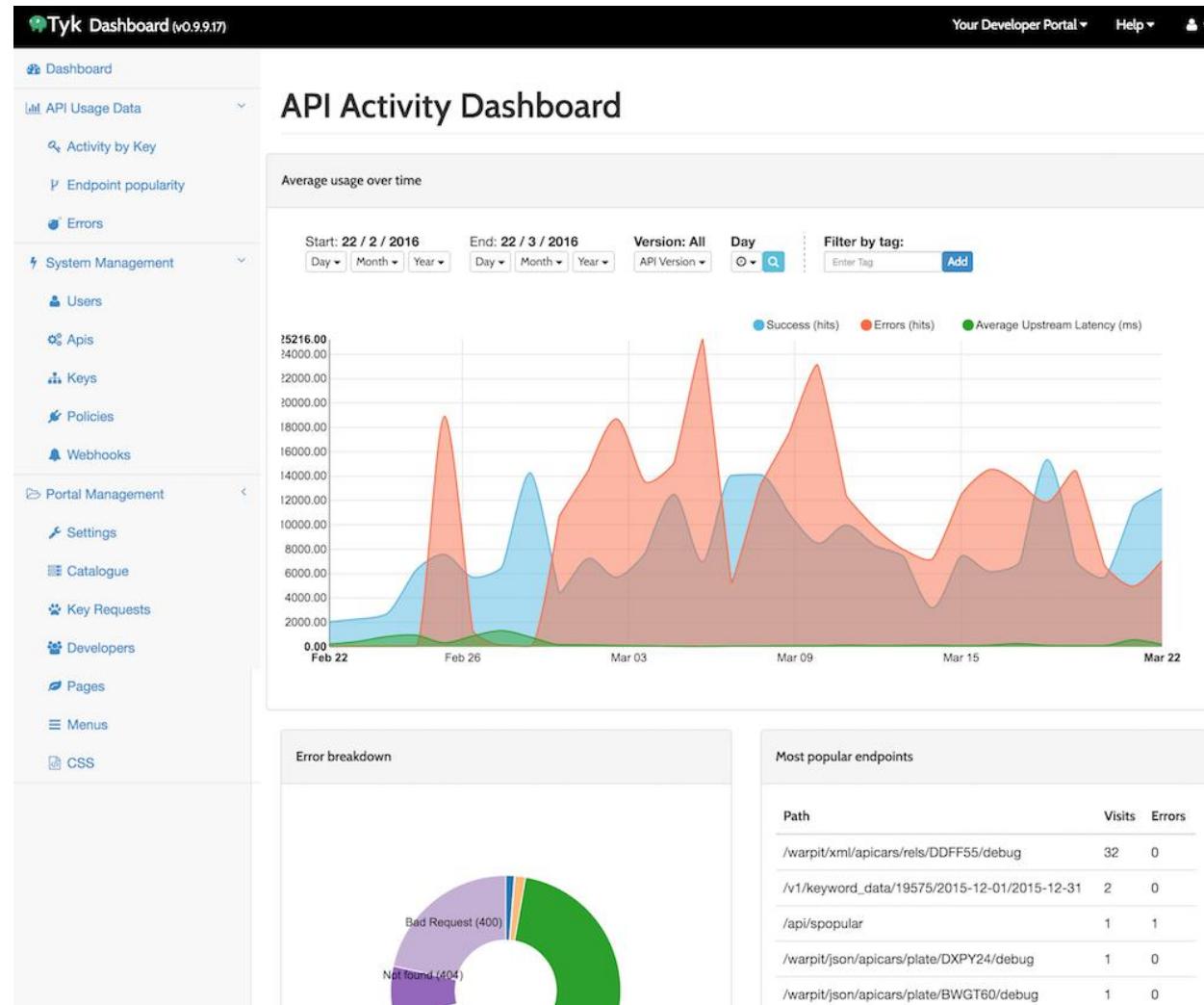
E



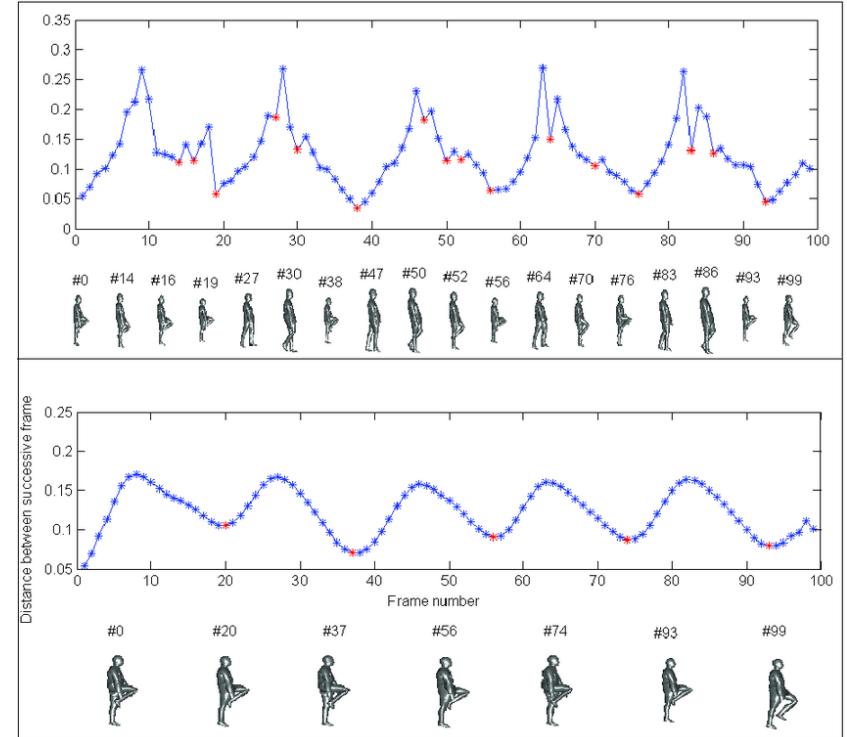
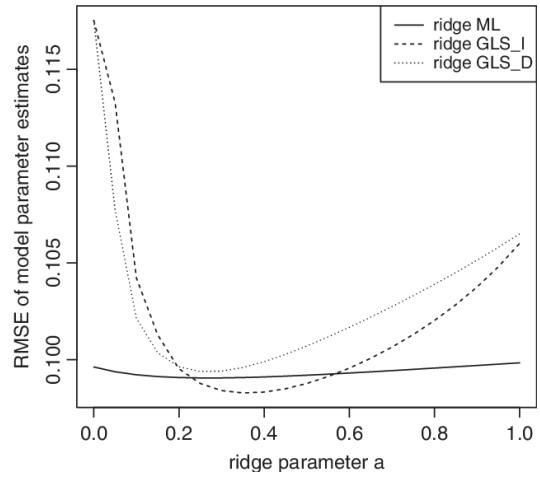
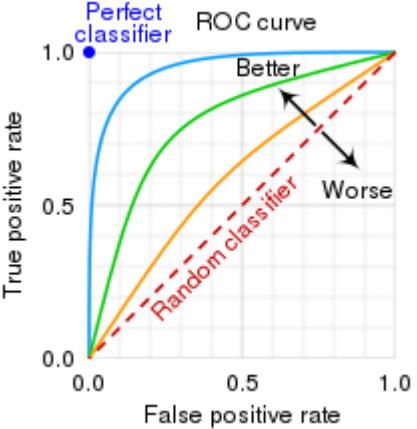
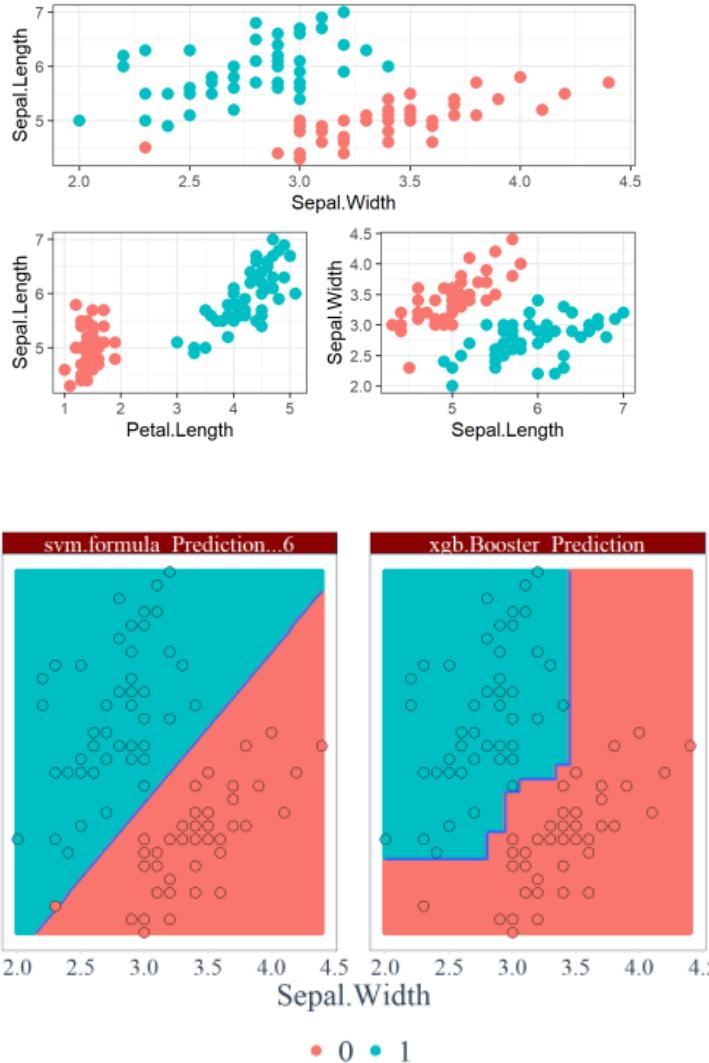
D



Dashboard with Tyk/Tableau

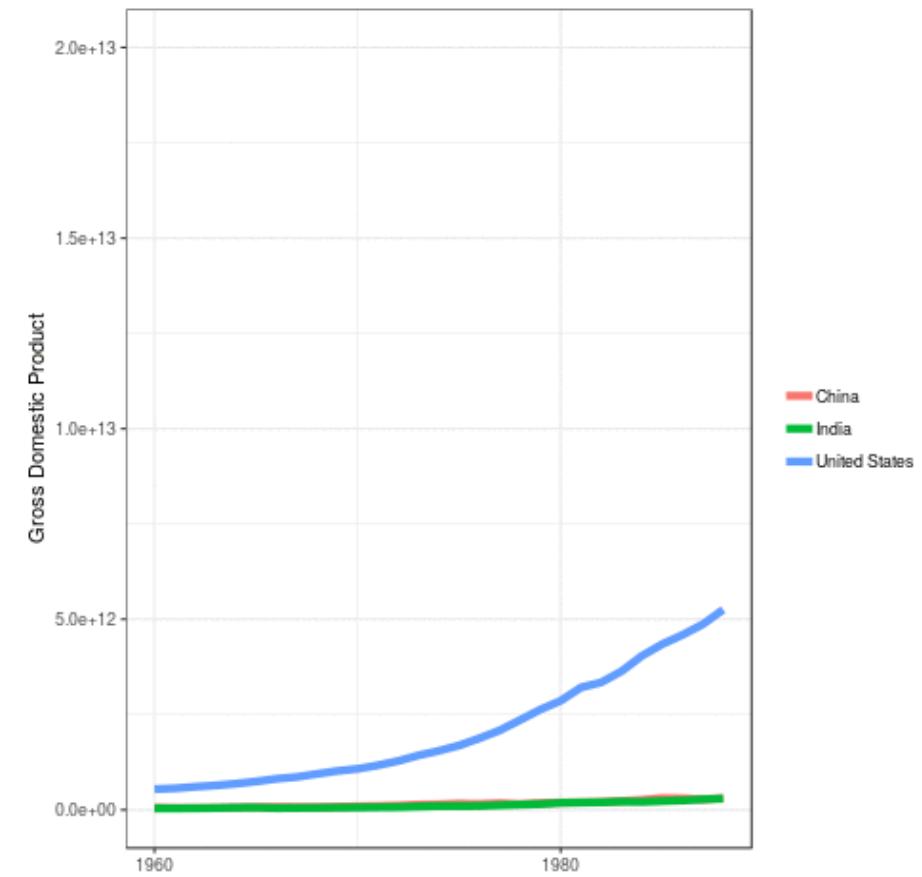
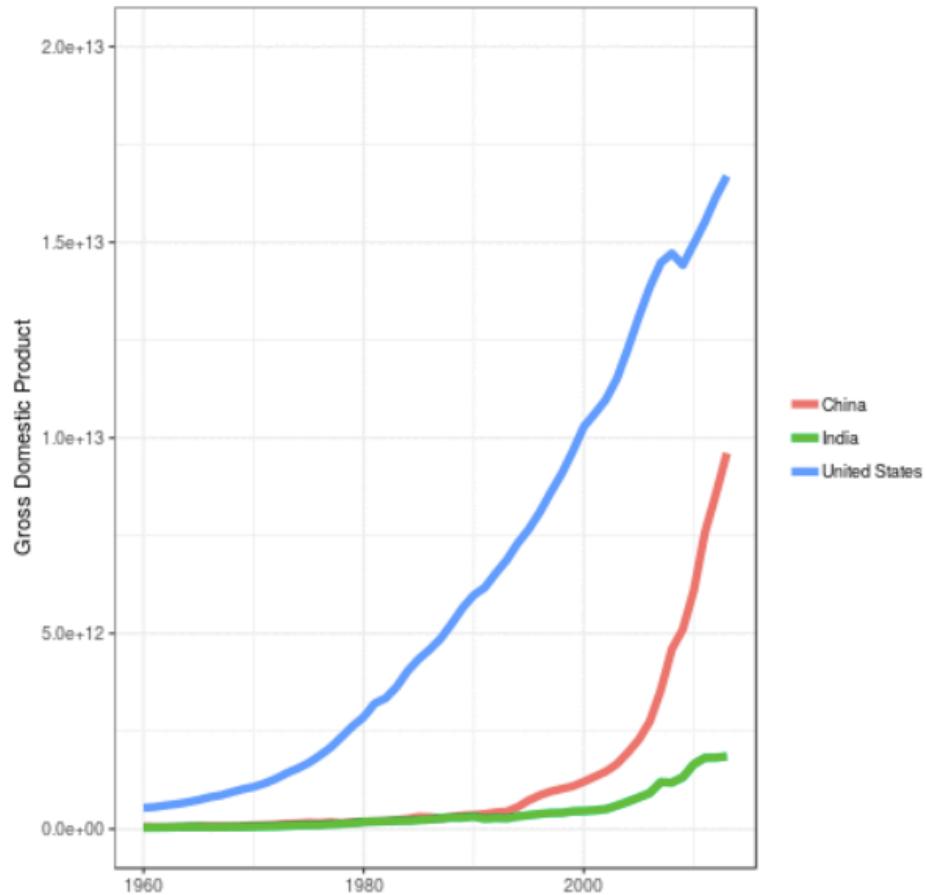


Data visualization for machine learning

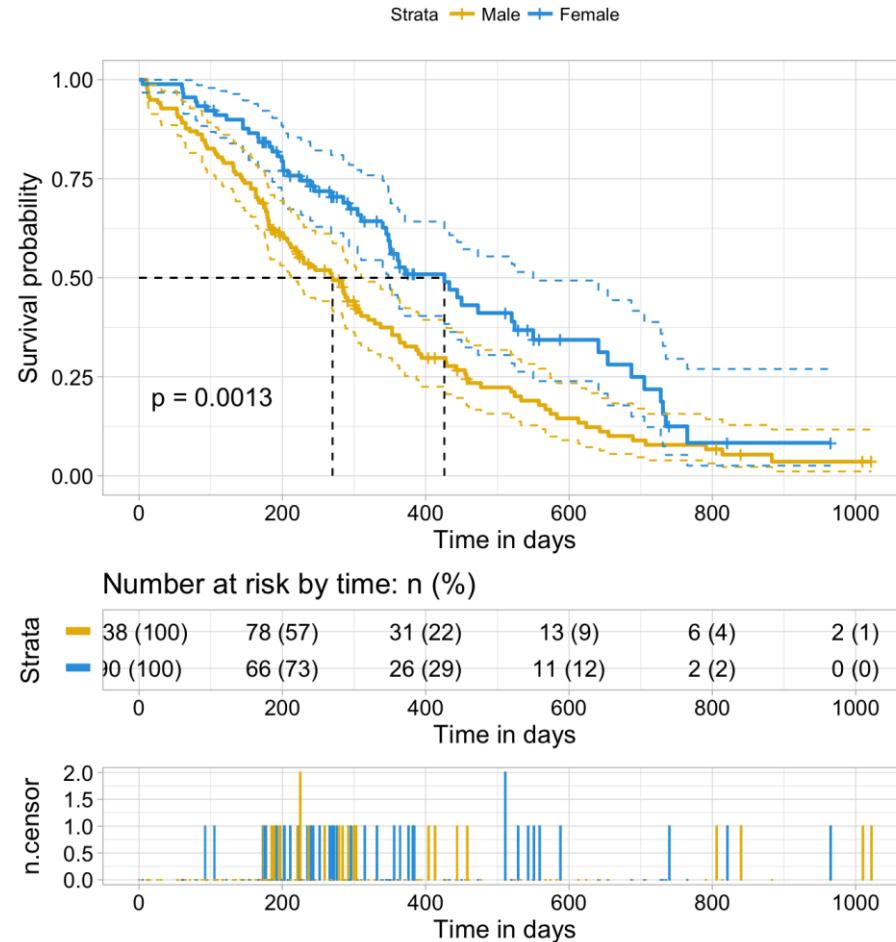


Speed curve smoothing: (top) speed curve before smoothing, (bottom) speed curve after smoothing.
<https://hal.archives-ouvertes.fr/tel-01094740/file/RimSlamaThesis2014.pdf>

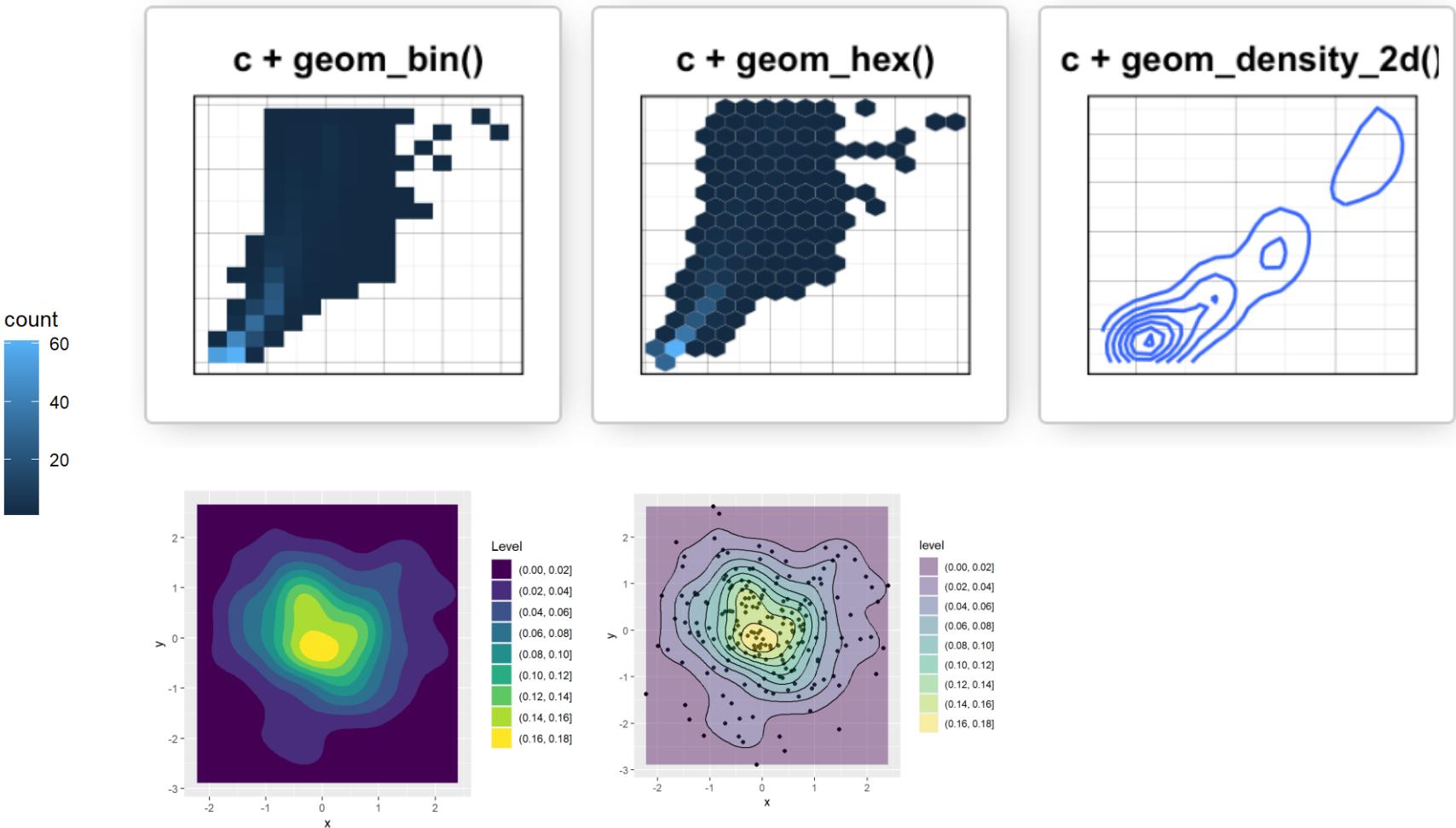
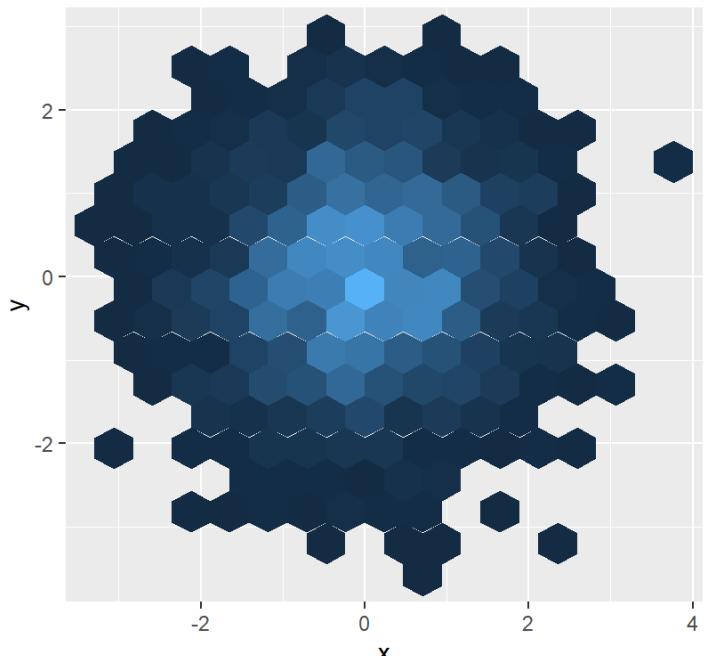
Amination plot



Survival plot (Kaplan-Meier Curve)



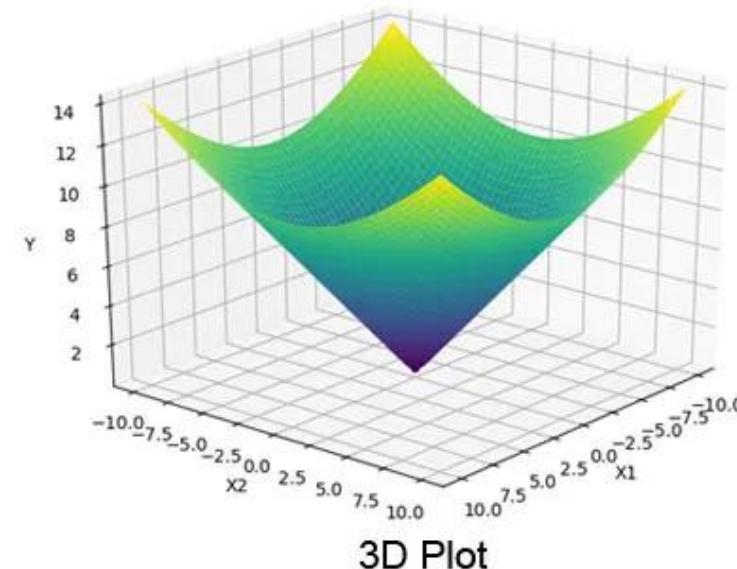
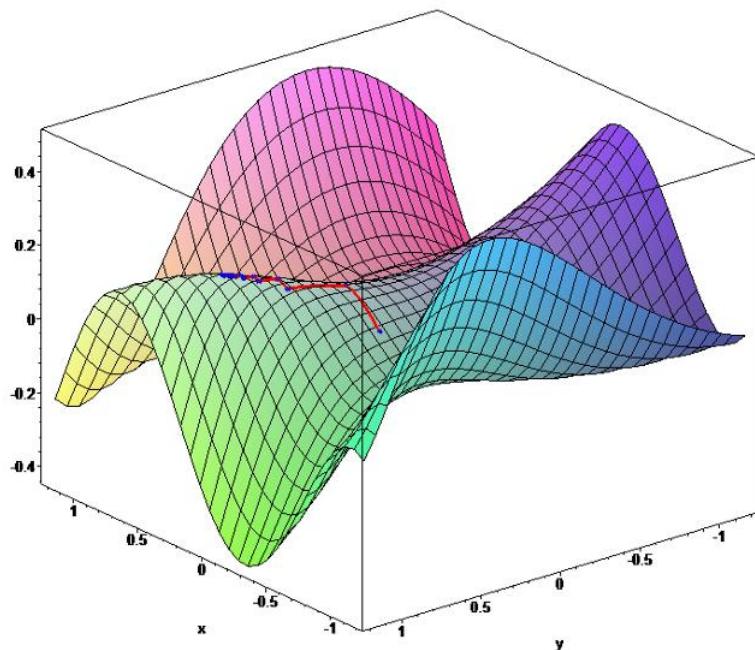
Binning and contour plot



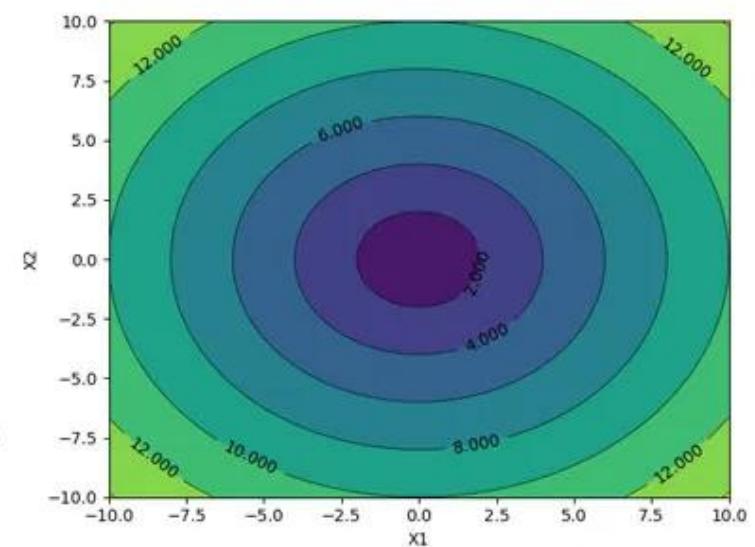
<https://r-charts.com/correlation/hexbin-chart-ggplot2/>

<https://r-charts.com/correlation/contour-plot-ggplot2/>
https://vincenzocoia.com/post/contour_plots/

3D contour plot for Gradient Descent



3D Plot

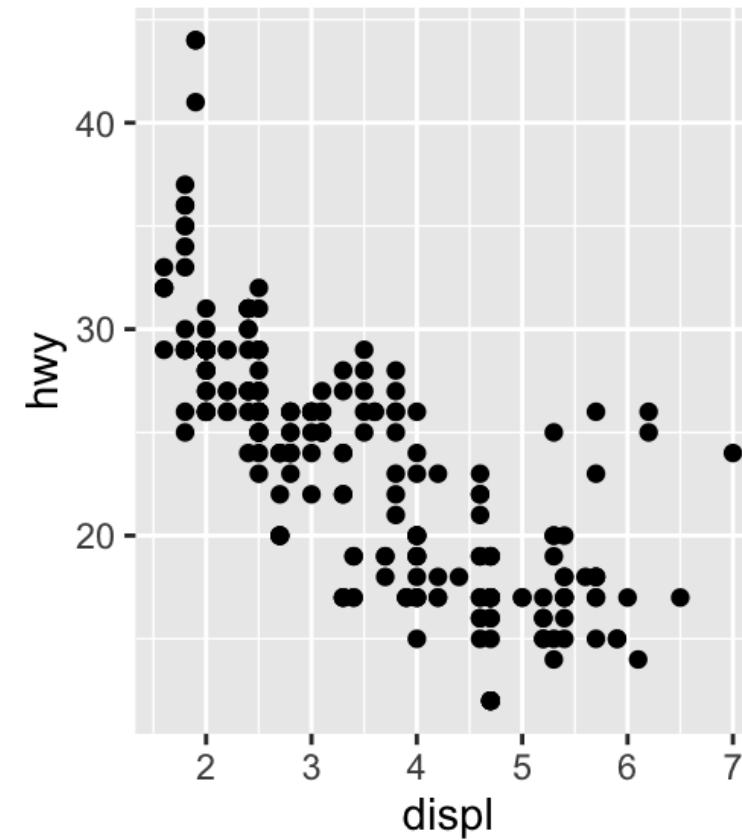
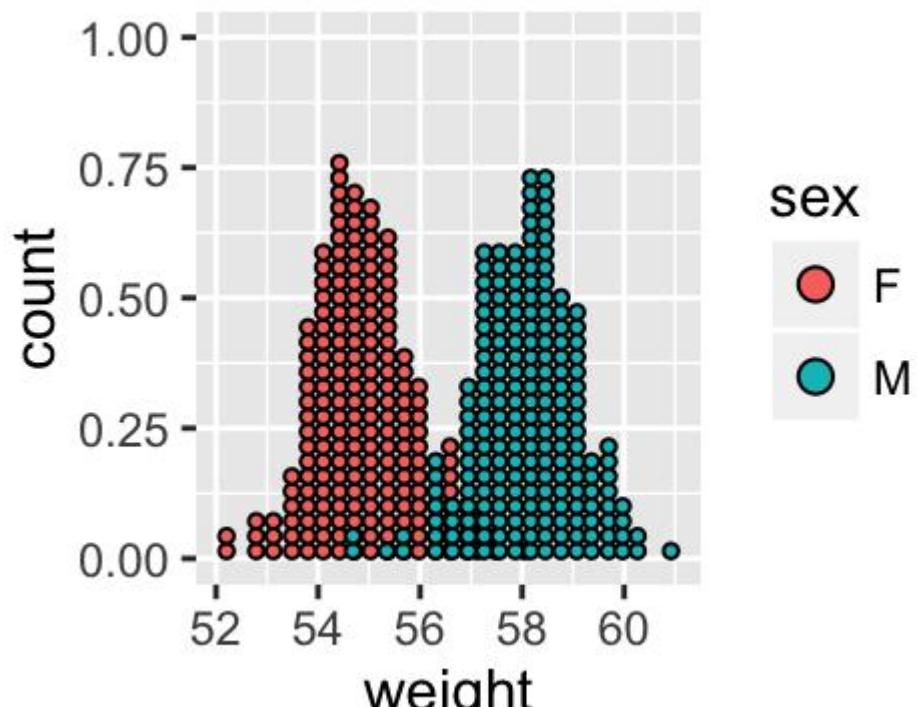


Contour Plot

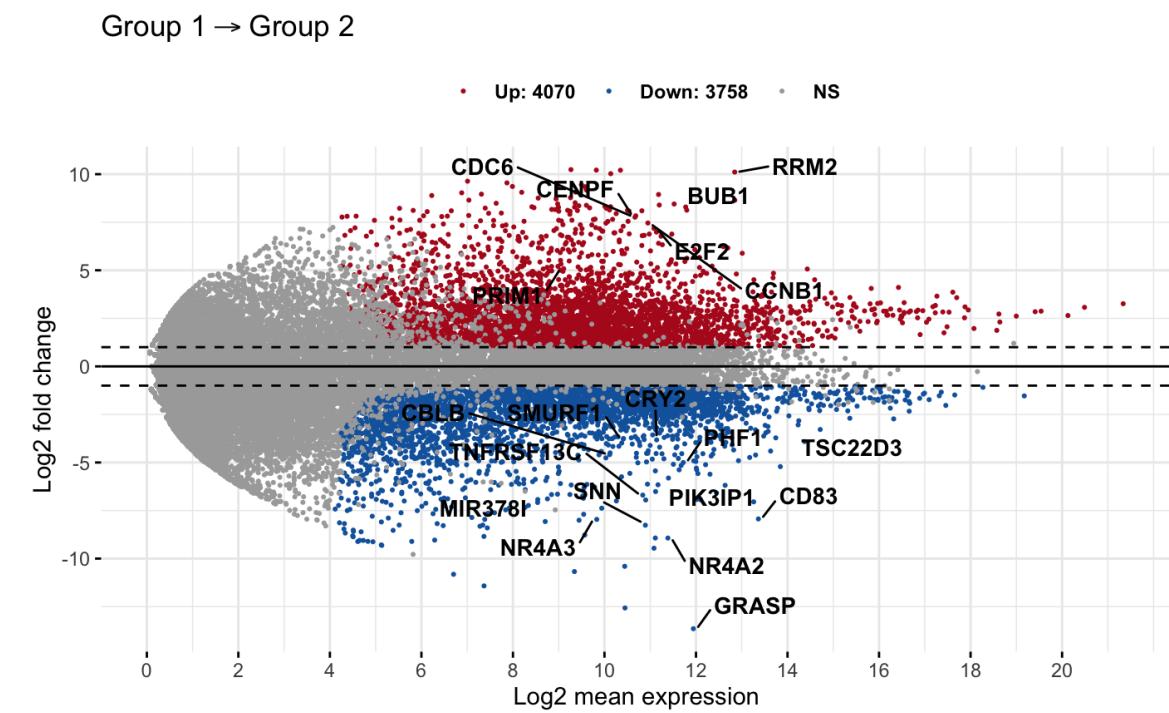
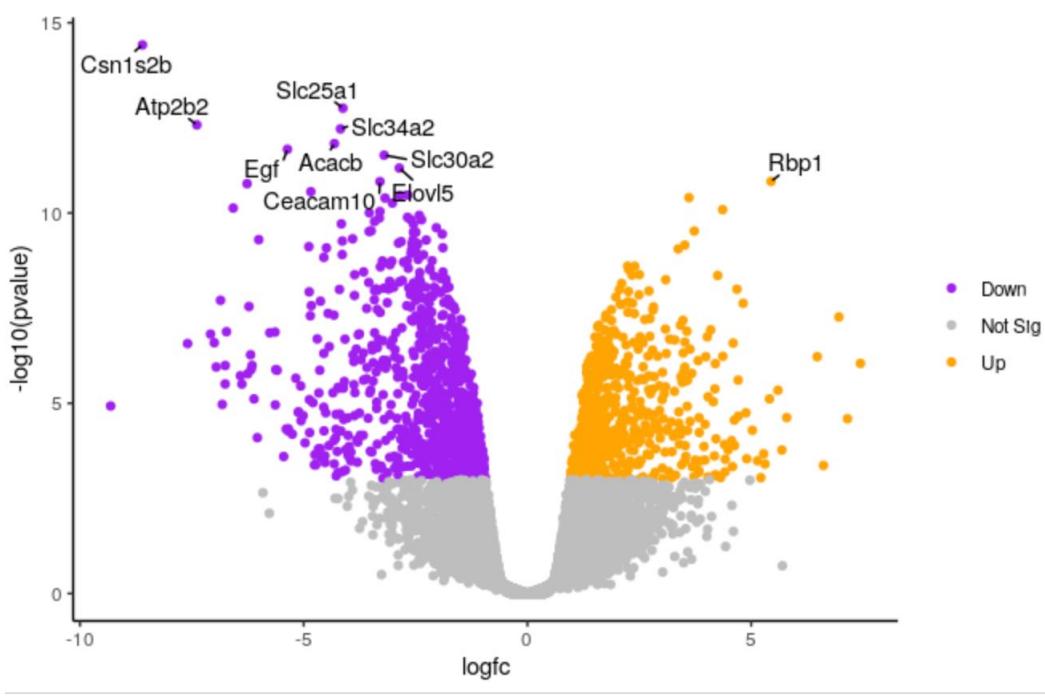
<https://medium.com/@jackyfeng530/machine-learning-gradient-descent-8681c40d27c3>

<http://www.adeveloperdiary.com/data-science/how-to-visualize-gradient-descent-using-contour-plot-in-python/>

Dotplot and jitter plot



Volcano vs MA (Bland–Altman) plot



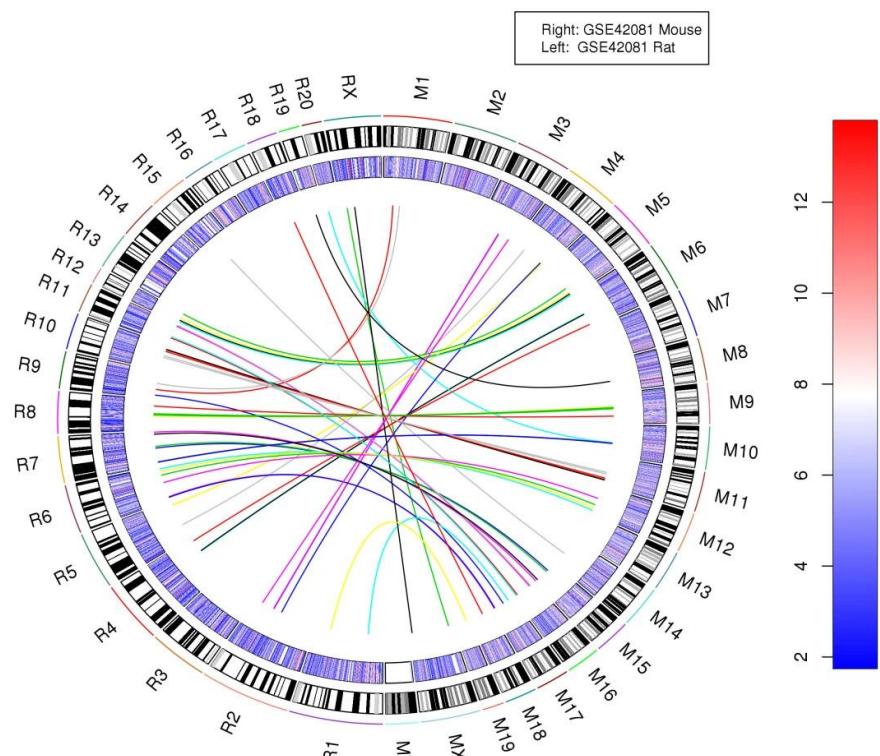
<http://onetipperday.sterding.com/2012/11/t-test-volcano-plot.html>

https://biocorecrg.github.io/CRG_RIntroduction/volcano-plots.html

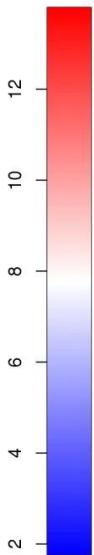
<https://rpkgs.datanovia.com/ggpubr/reference/ggmaplot.html>

Circos plot

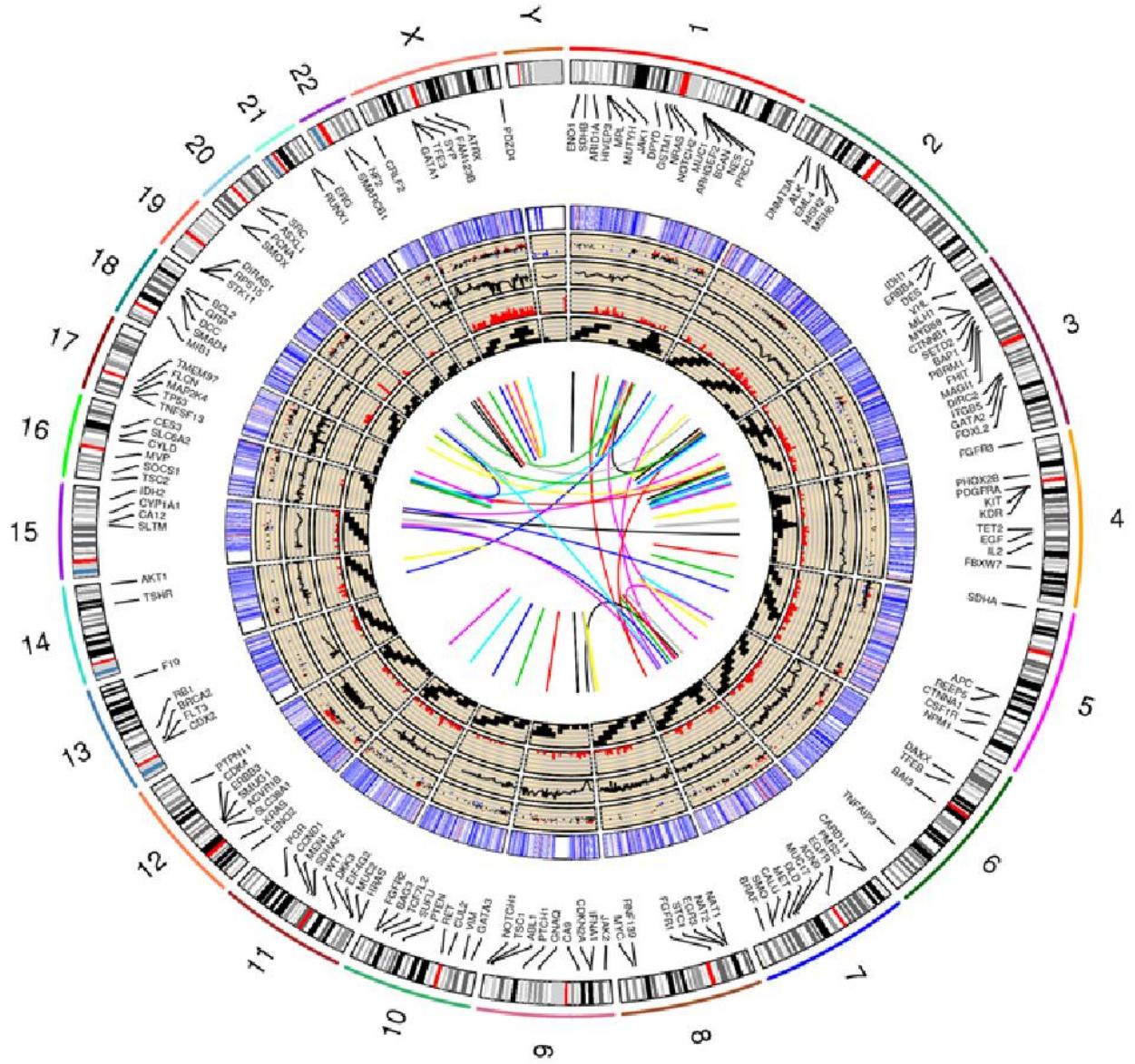
RCircos Demo: Mouse and Rat Gene Expression



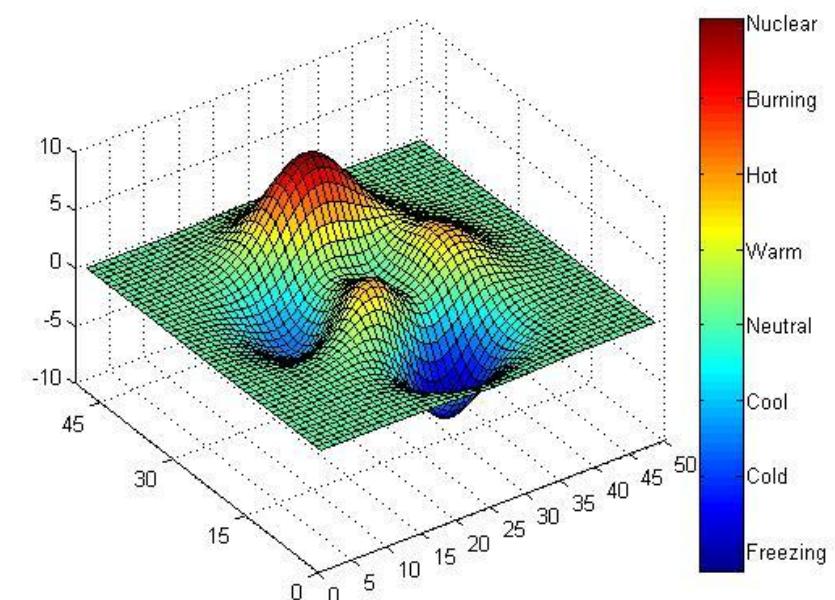
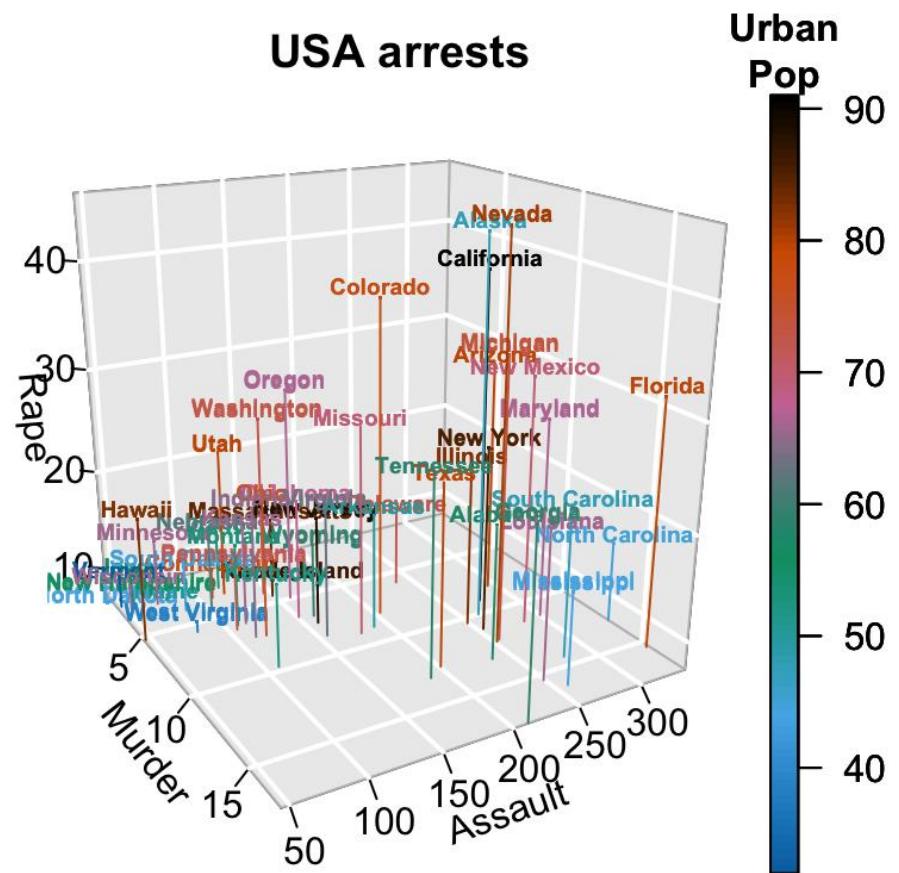
Right: GSE42081 Mouse
Left: GSE42081 Rat



RCircos 2D Track Plot with Human Genome



3D plot



<http://www.sthda.com/english/wiki/impressive-package-for-3d-and-4d-graph-r-software-and-data-visualization>

<https://stackoverflow.com/questions/43330727/how-to-plot-three-dimension-data-or-four-dimension-data-to-filled-contour-plot-on-a-3d-surface>

Linked Views for Visual Exploration

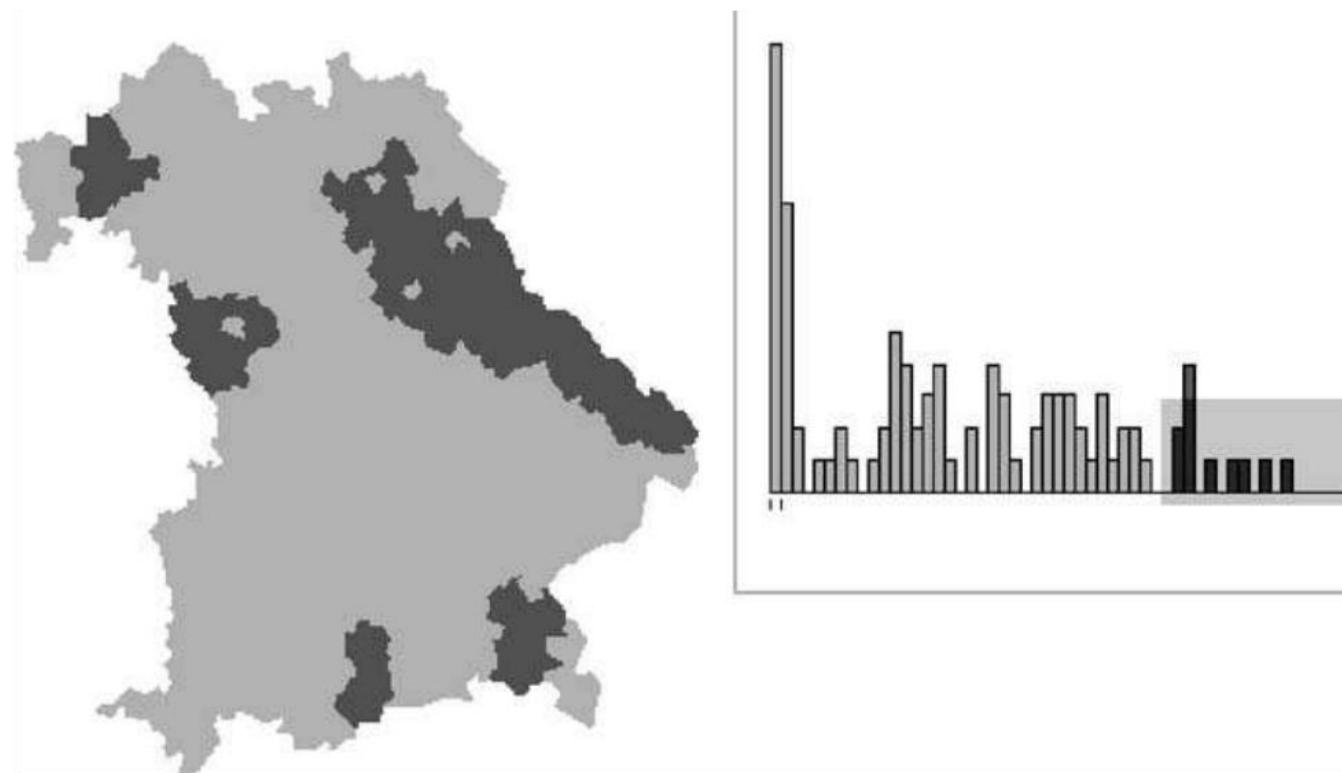


Figure 8.3. Land usage on the county level in Bavaria. Highlighted are those counties with a high percentage of forestry

Linked Views for Visual Exploration

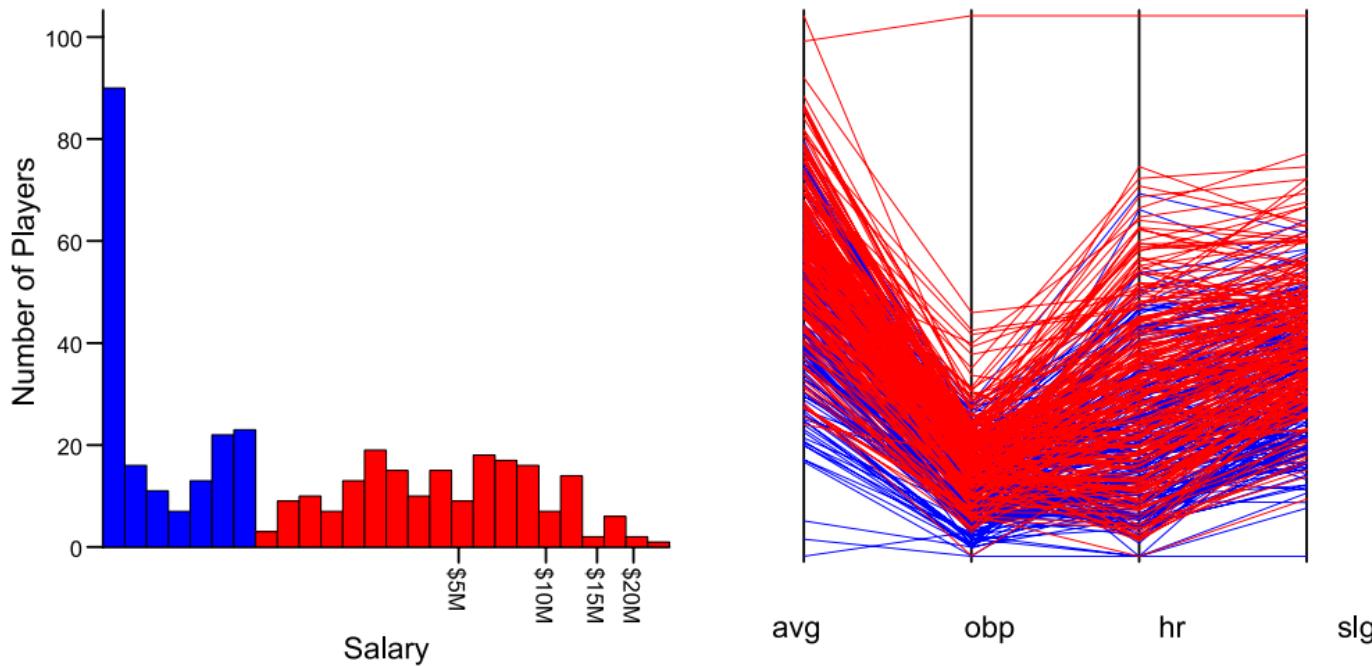
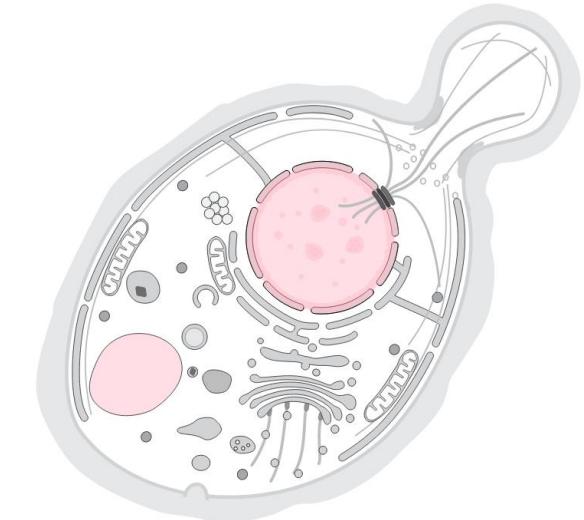
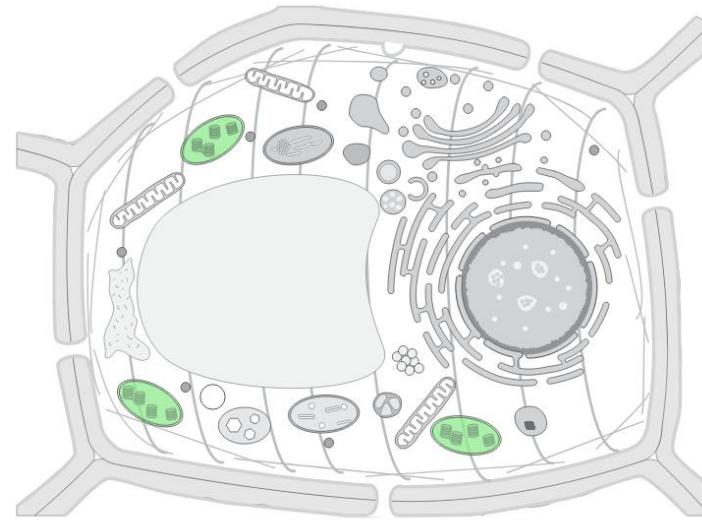
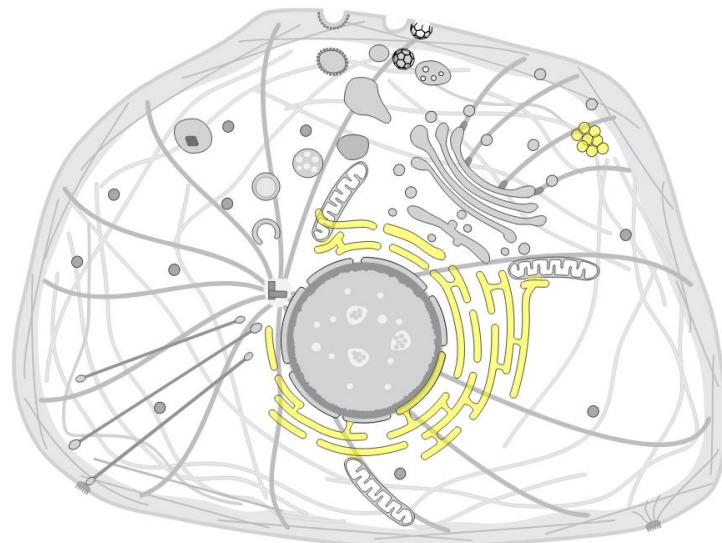


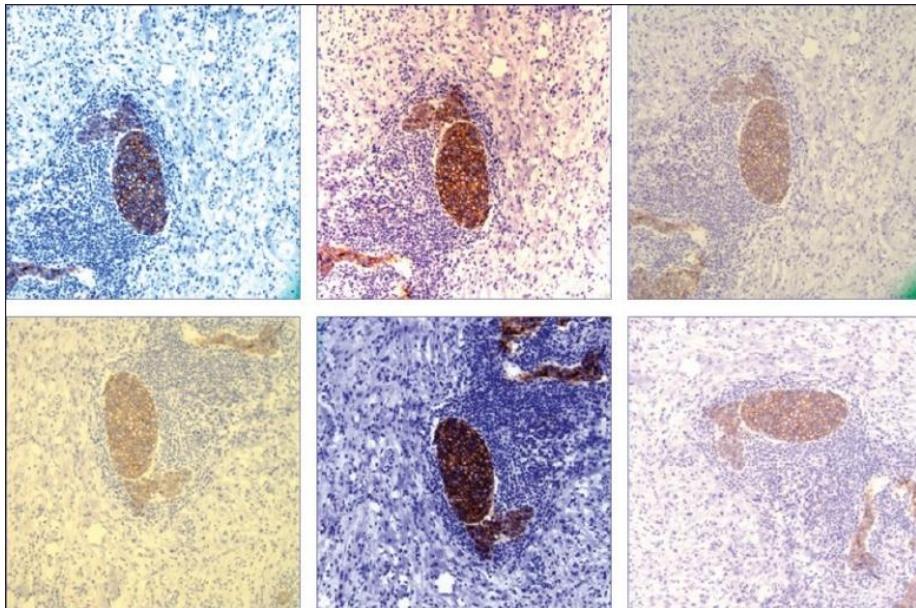
Figure 9.5. Faceting and parallel coordinates. For the 2004 season, the *left view* shows a histogram of salary on a log scale. Players earning over \$1,000,000 were selected and are shown in the linked *parallel coordinates plot* as *dashed line*. The parallel coordinates view shows four major measures of performance: batting average, on-base percentage, home-run rate, and slugging average

Draw Cell



<https://github.com/svalvaro/drawCell>

Visualizing Images in pathology

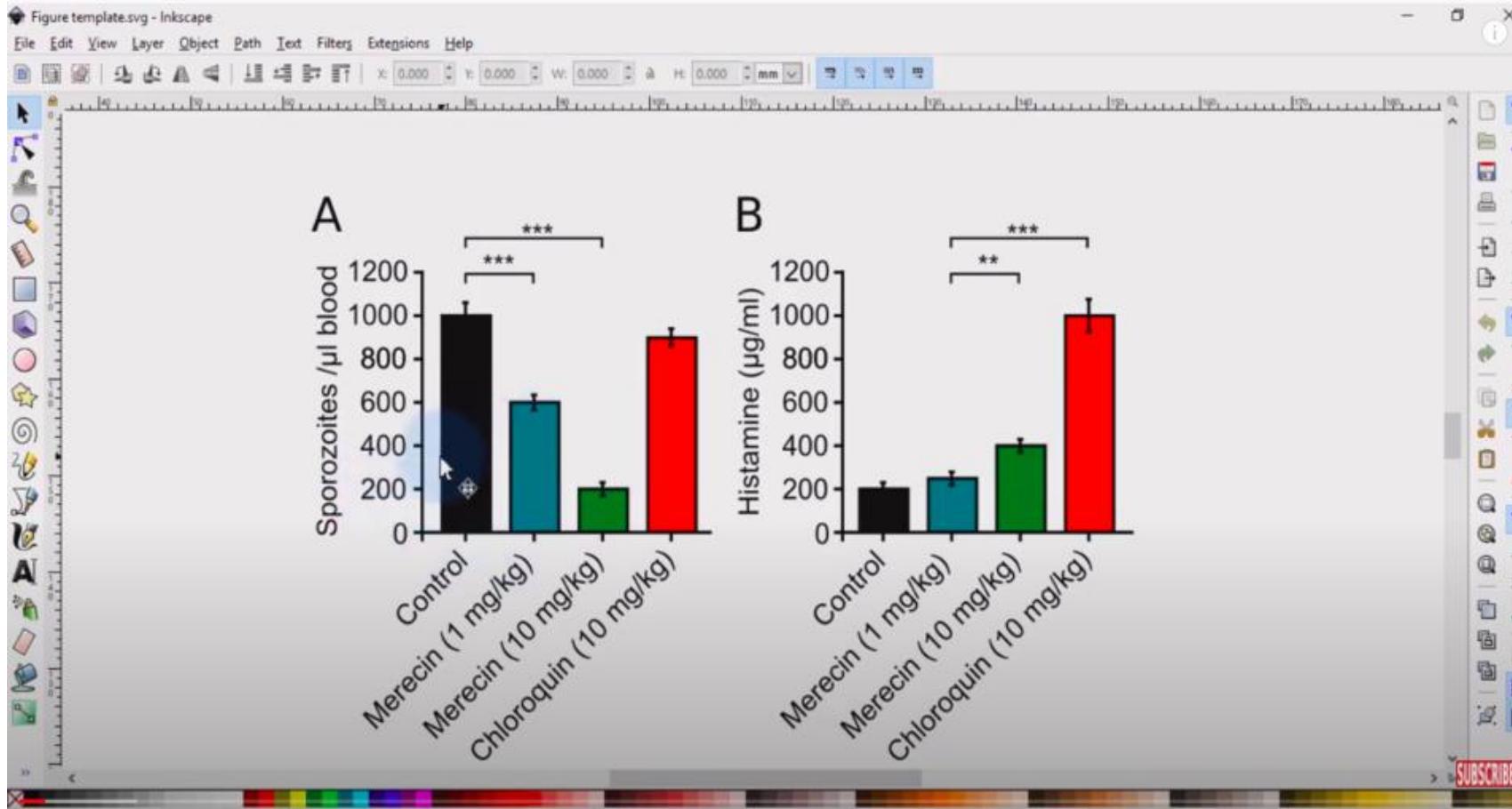


Different digital images of the same region on a glass slide photographed at the same magnification by six different pathologists, each using similar microscopes and the same attached digital cameras (HER-2/neu immunohistochemical stain)



Whole slide images help create a “virtual multiheaded microscope” that supports interactive education (Image courtesy of BioImagene)

Inkscape for scientists: Making a data figure for a publication



Why Do We Visualize Data?

2	5	8	4	4	4	8	5	6	6	4	2	7	6	2	5	8	4	7	4	4	2	6	8	1	6
3	1	2	2	8	7	6	1	8	5	1	2	4	5	1	4	8	6	3	7	3	3	3	7	8	1
1	6	4	7	1	8	8	2	4	3	3	2	3	5	4	4	2	7	3	8	5	1	5	7	6	3
1	1	7	8	3	6	3	3	7	7	7	2	4	3	1	5	1	7	8	7	4	3	4	5		
3	2	7	4	6	6	8	3	3	8	4	7	5	8	2	7	4	2	2	1	5	3	6	2	1	8

2	5	8	4	4	4	8	5	6	6	4	2	7	6	2	5	8	4	7	4	4	2	6	8	1	6
3	1	2	2	8	7	6	1	8	5	1	2	4	5	1	4	8	6	3	7	3	3	3	7	8	1
1	6	4	7	1	8	8	2	4	3	3	2	3	5	4	4	2	7	3	8	5	1	5	7	6	3
1	1	7	8	3	6	3	3	7	7	7	2	4	3	1	5	1	7	8	7	4	3	4	5		
3	2	7	4	6	6	8	3	3	8	4	7	5	8	2	7	4	2	2	1	5	3	6	2	1	8

Why Do We Visualize Data?

TABLE 1.1 Table with four groups of numbers: What do they tell you?

Group A		Group B		Group C		Group D	
x	y	x	y	x	y	x	y
10.00	8.04	10.00	9.14	10.00	7.46	8.00	6.58
8.00	6.95	8.00	8.14	8.00	6.77	8.00	5.76
13.00	7.58	13.00	8.74	13.00	12.74	8.00	7.71
9.00	8.81	9.00	8.77	9.00	7.11	8.00	8.84
11.00	8.33	11.00	9.26	11.00	7.81	8.00	8.47
14.00	9.96	14.00	8.10	14.00	8.84	8.00	7.04
6.00	7.24	6.00	6.13	6.00	6.08	8.00	5.25
4.00	4.26	4.00	3.10	4.00	5.39	19.00	12.50
12.00	10.84	12.00	9.13	12.00	8.15	8.00	5.56
7.00	4.82	7.00	7.26	7.00	6.42	8.00	7.91
5.00	5.68	5.00	4.74	5.00	5.73	8.00	6.89

Explore Data

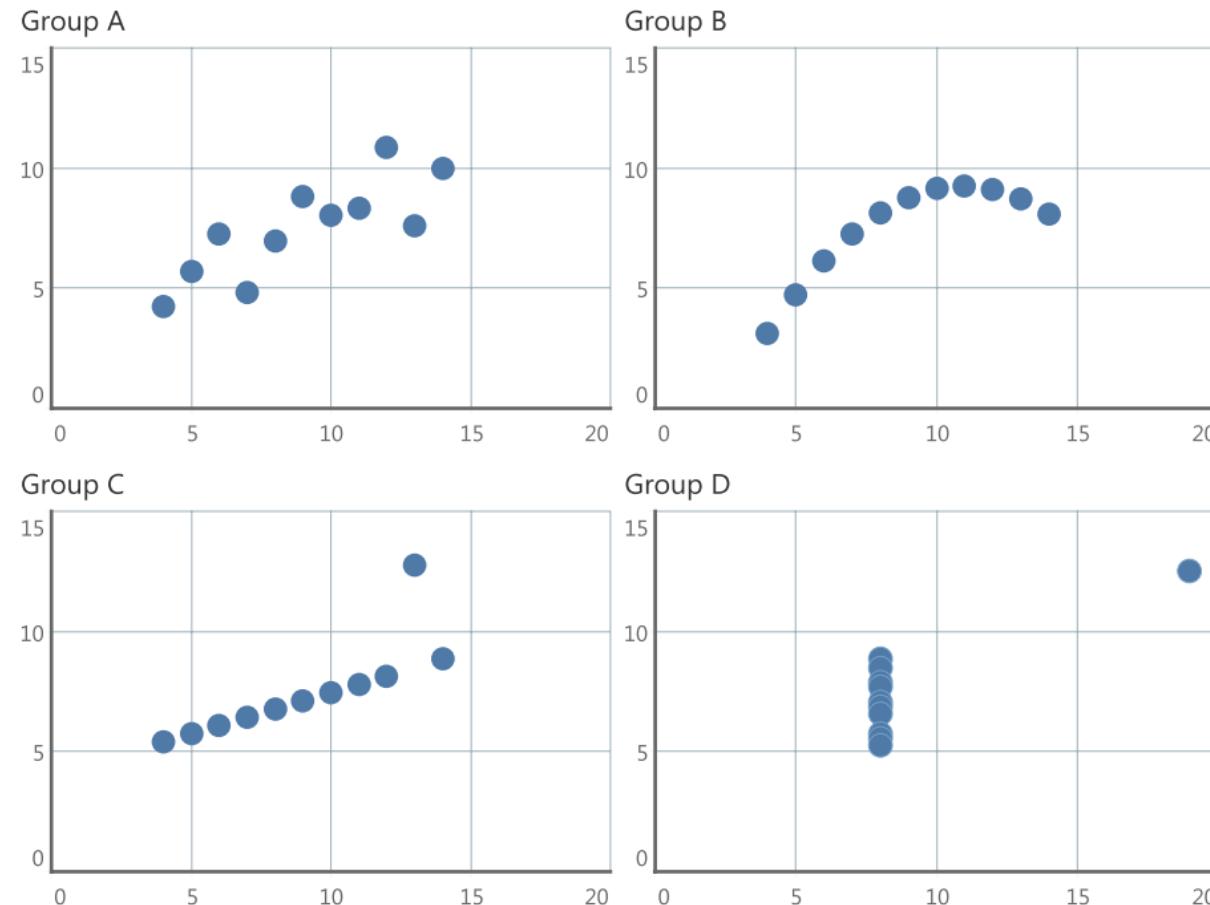


FIGURE 1.1 Now can you see a difference in the four groups?

“Anscombe’s Quartet”—in the paper “Graphs in Statistical Analysis” in 1973.

Discover the trends of data

TABLE 1.2 What are the trends in sales?

Category	2013 Q1	2013 Q2	2013 Q3	2013 Q4	2014 Q1	2014 Q2	2014 Q3	2014 Q4
Furniture	\$463,988	\$352,779	\$338,169	\$317,735	\$320,875	\$287,934	\$319,537	\$324,319
Office Supplies	\$232,558	\$290,055	\$265,083	\$246,946	\$219,514	\$202,412	\$198,268	\$279,679
Technology	\$563,866	\$244,045	\$432,299	\$461,616	\$285,527	\$353,237	\$338,360	\$420,018
Category	2015 Q1	2015 Q2	2015 Q3	2015 Q4	2016 Q1	2016 Q2	2016 Q3	2016 Q4
Furniture	\$307,028	\$273,836	\$290,886	\$397,912	\$337,299	\$245,445	\$286,972	\$313,878
Office Supplies	\$207,363	\$183,631	\$191,405	\$217,950	\$241,281	\$286,548	\$217,198	\$272,870
Technology	\$333,002	\$291,116	\$356,243	\$386,445	\$386,387	\$397,201	\$359,656	\$375,229

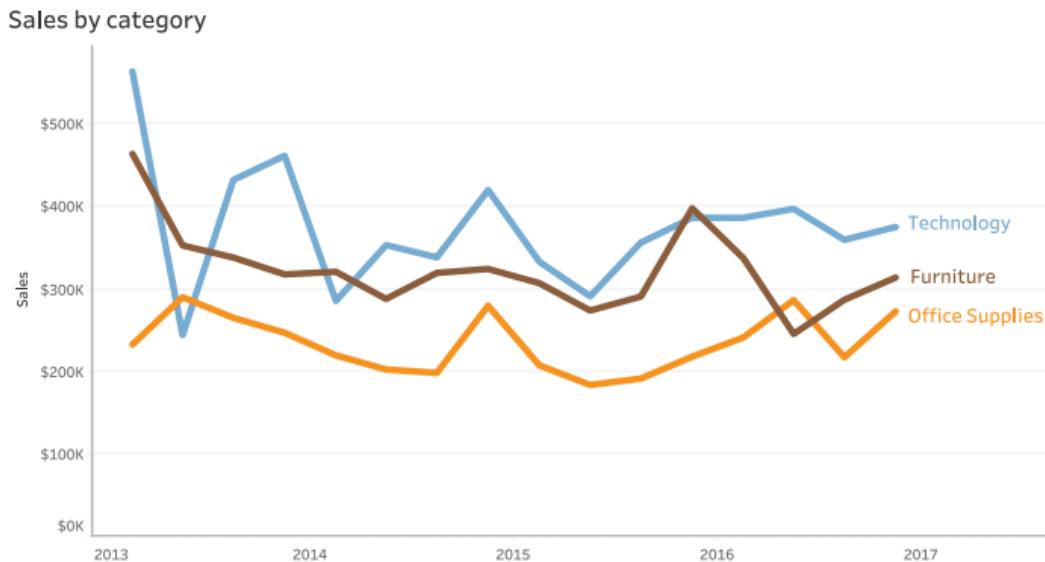
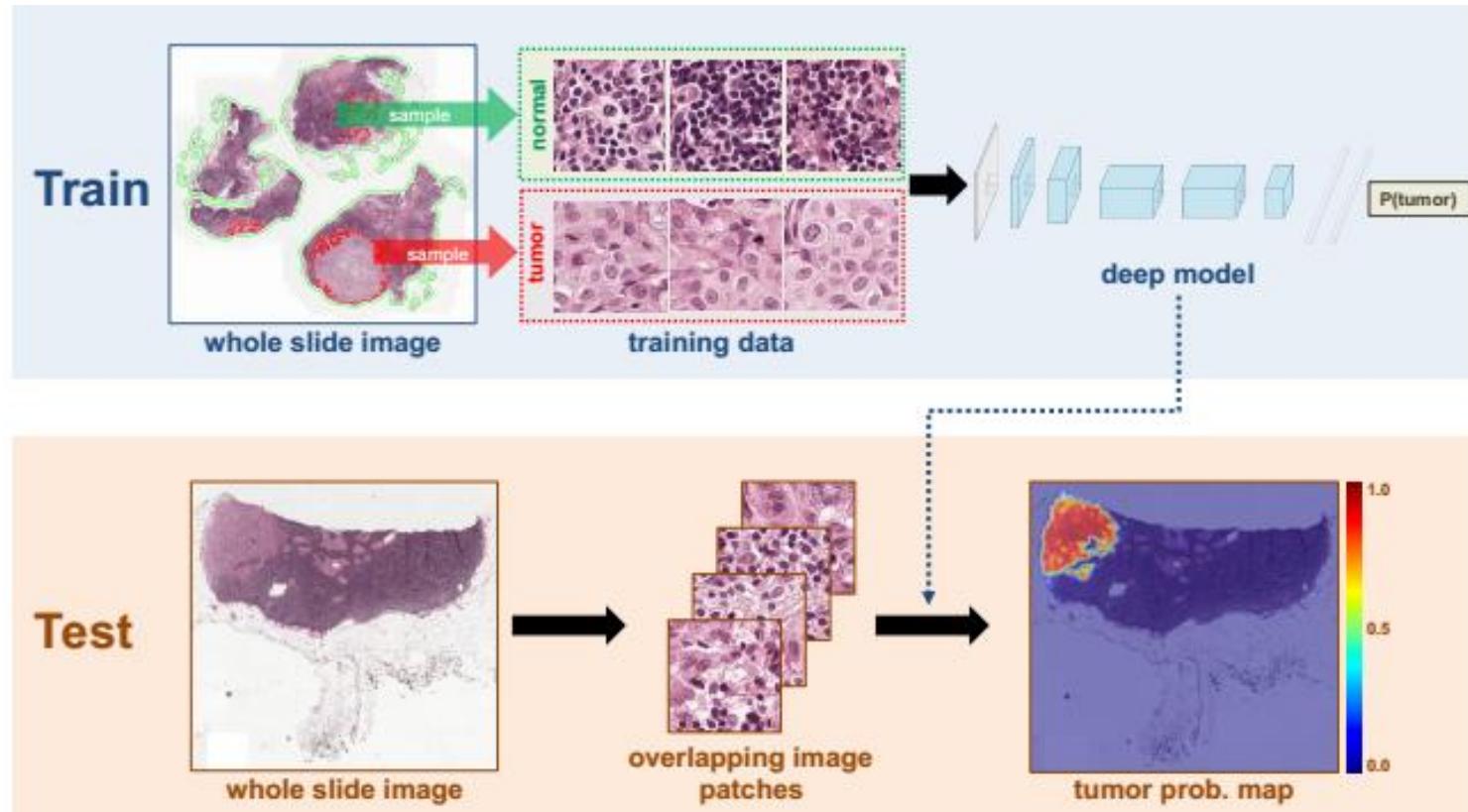
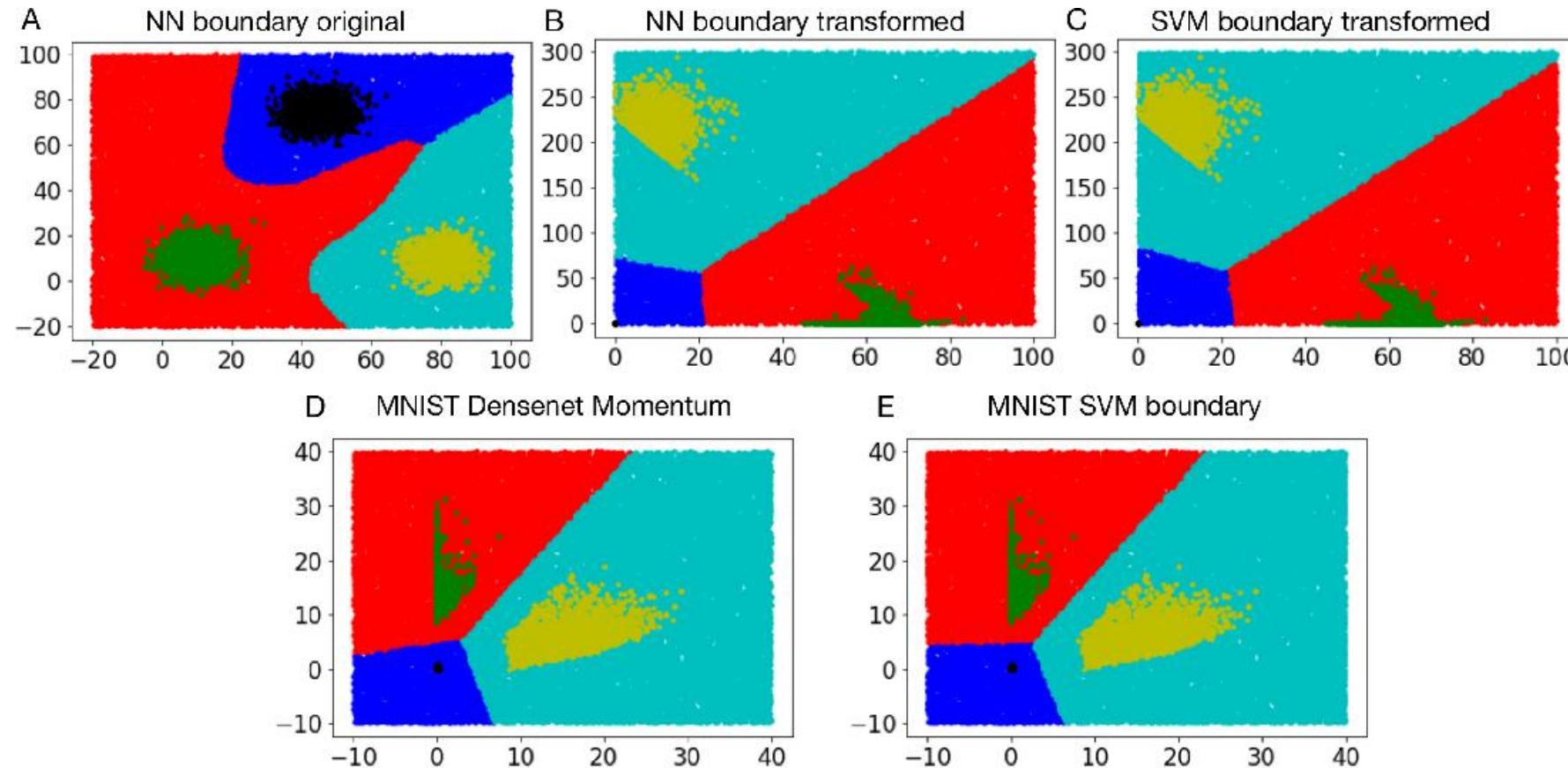


FIGURE 1.2 Now can you see the trends?

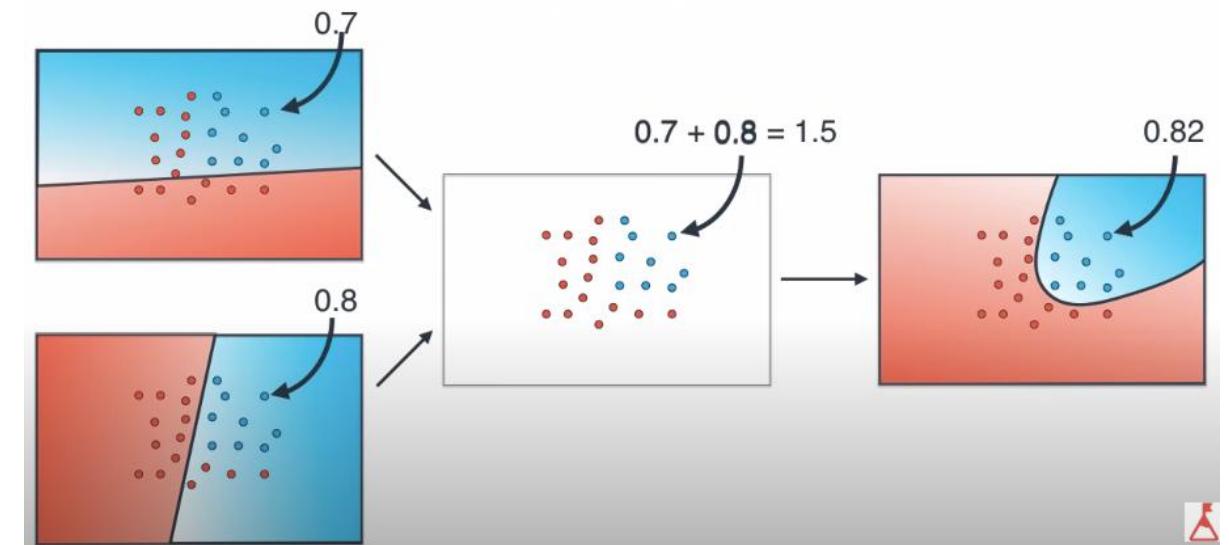
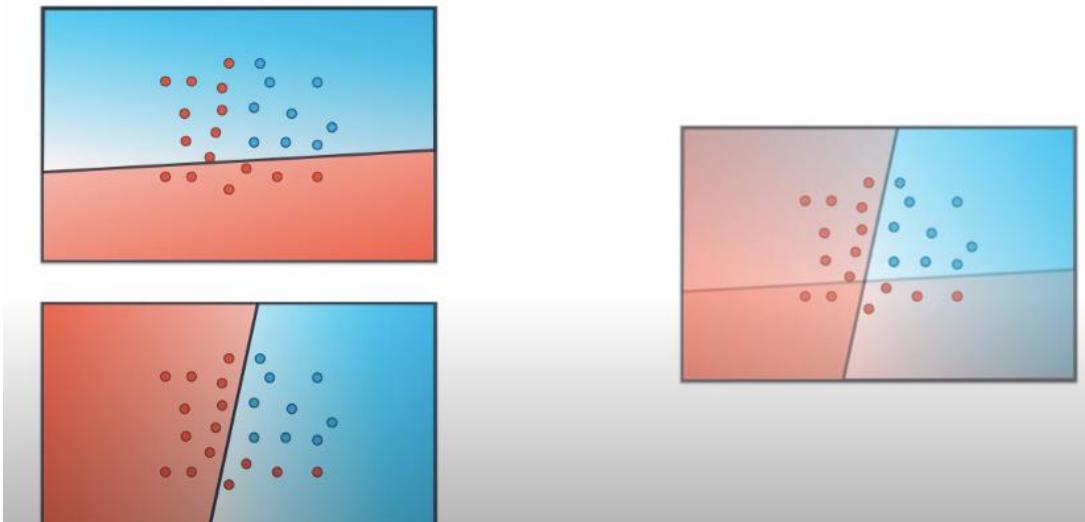
Neural network Interpretable



Neural network Interpretable

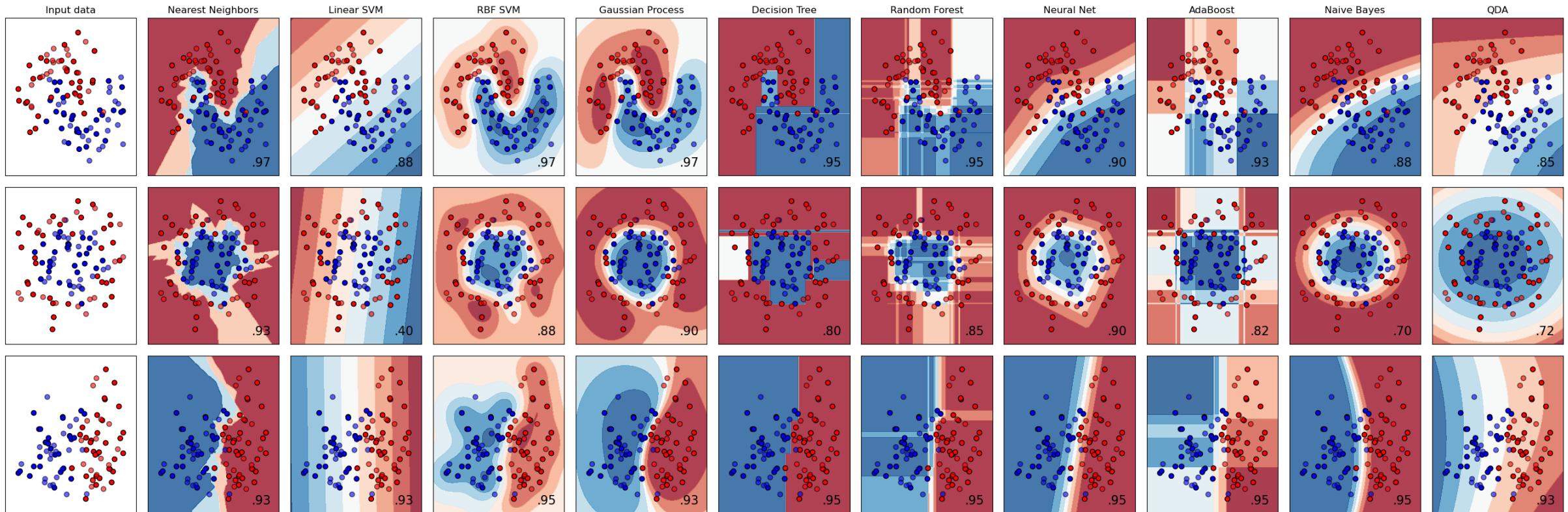


Neural network Interpretable



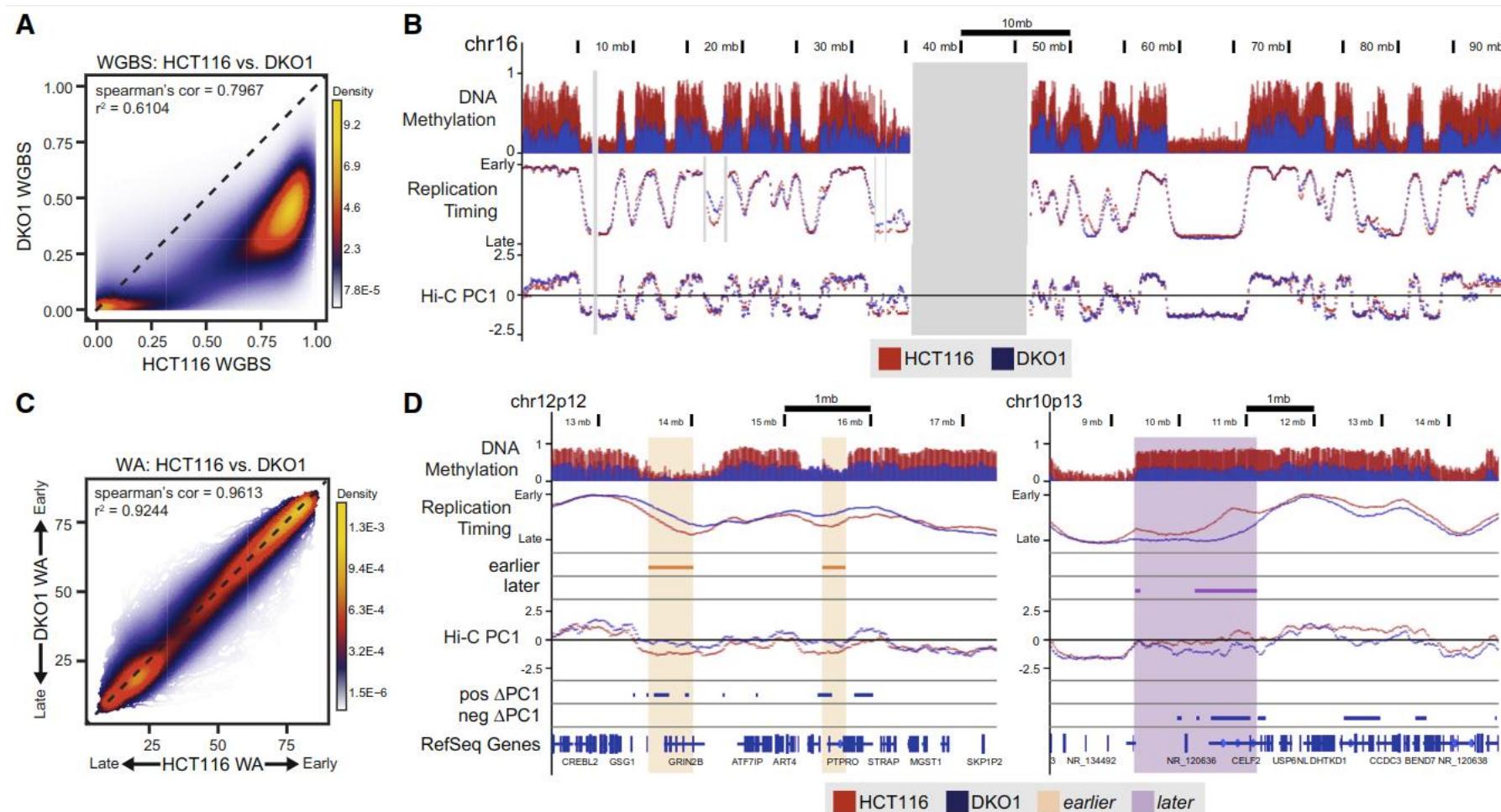
<https://www.youtube.com/watch?v=BR9h47Jtqyw>

Decision Boundary

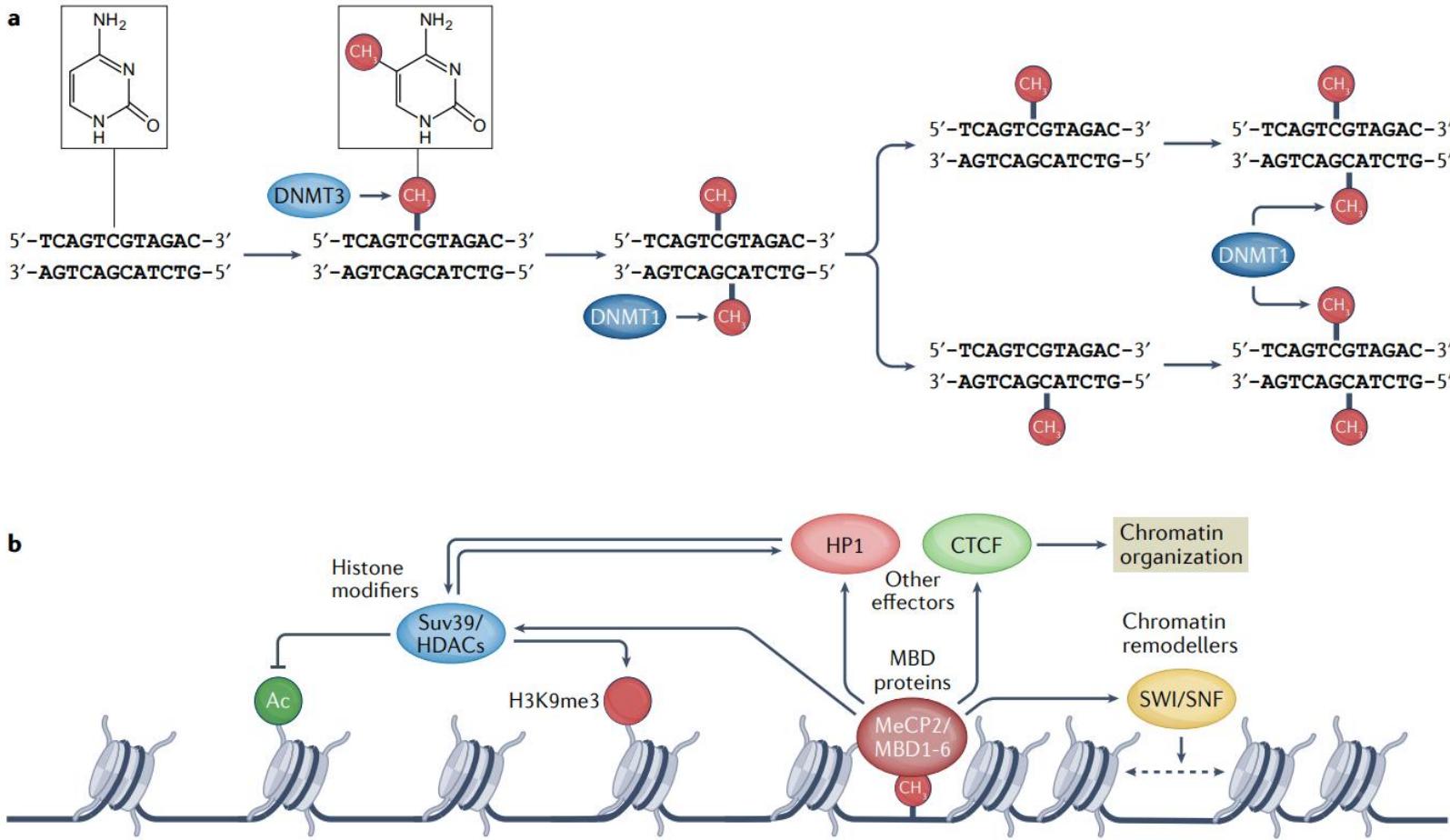


https://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html

Presenting a hypothesis

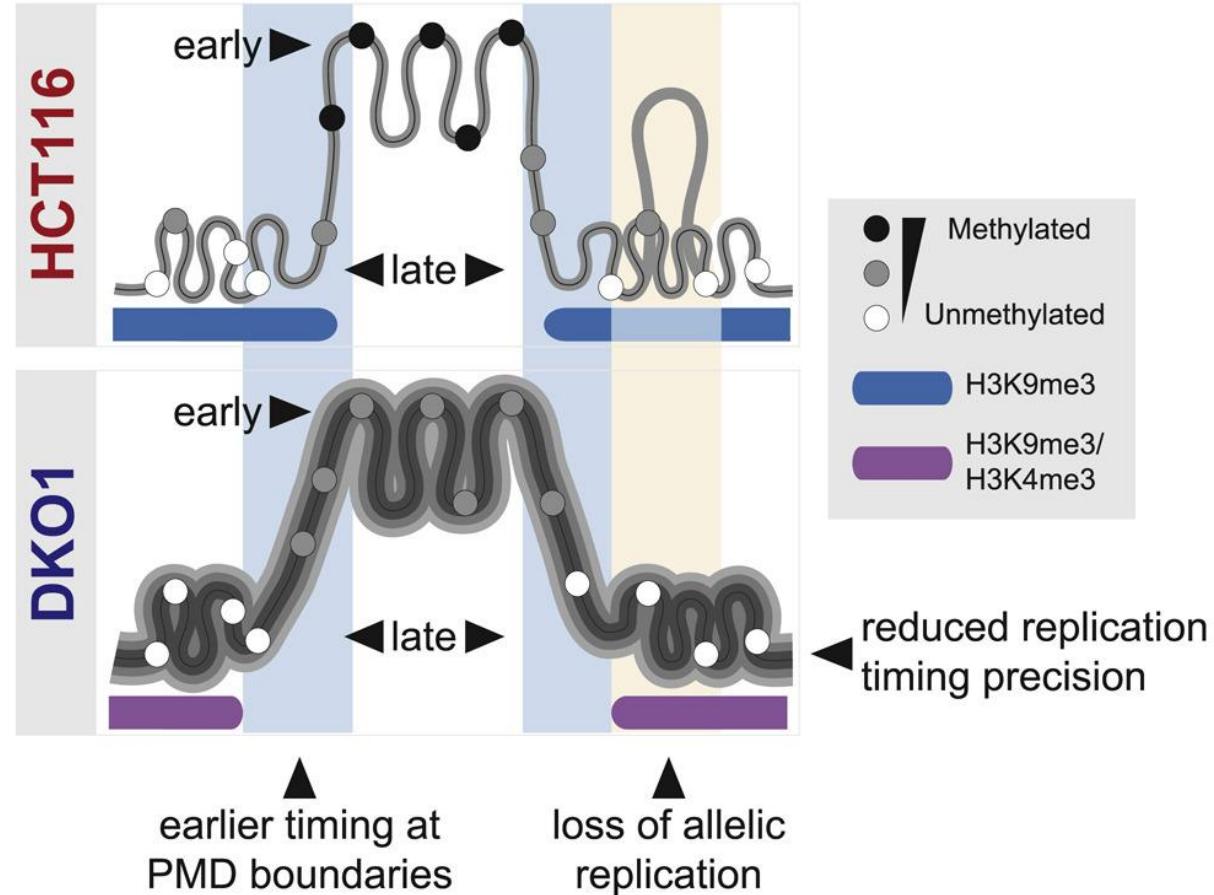


Summary scientific facts

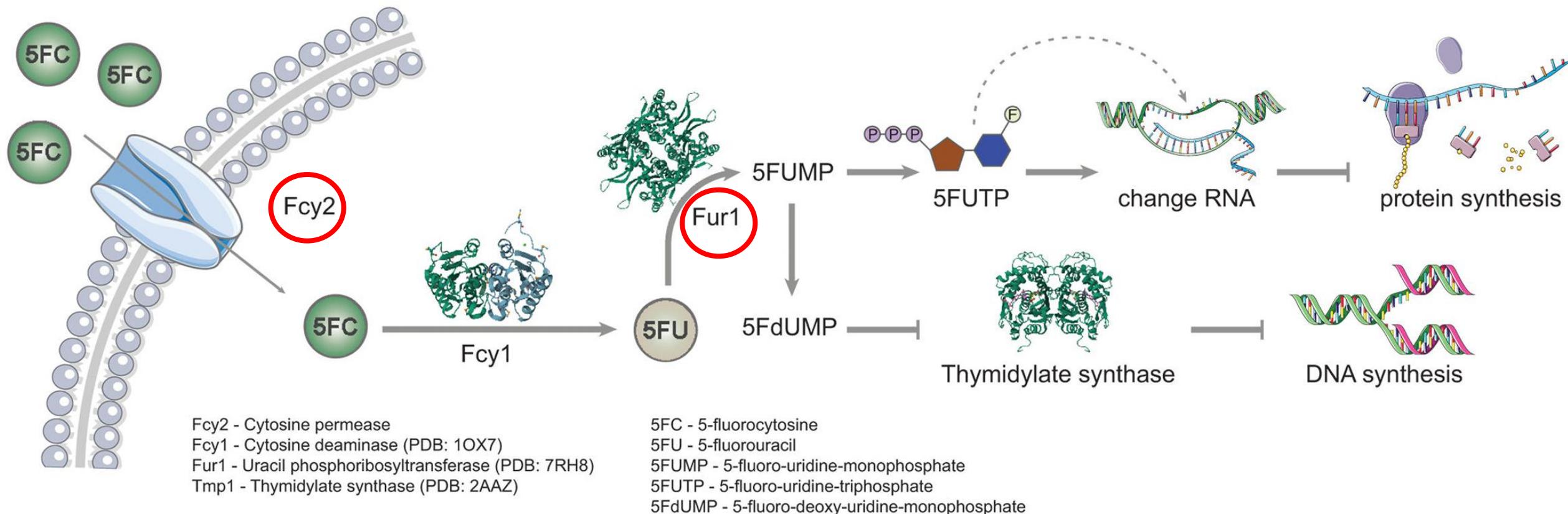


CTCF, CCCTC binding factor; HP1, heterochromatin protein 1; MeCP2, methyl-CpG-binding protein 2.

Graphical abstract



Rapid *in vitro* evolution of flucytosine resistance in *Candida auris*



Reference

- Data visualization
 - https://www.youtube.com/watch?v=3X3cjEqaPro&list=PLZo40sVm_w_4My97tv62aDITkOxOaJup7X&index=1
 - https://www.youtube.com/watch?v=Khj94kaJjWo&list=PLg91P4KVeUZbg28jiPq_qTgTCf6xQya-e
(<https://sas.uwaterloo.ca/~rwoldfor/courses/DataVisualization/Slides/2020/>)
 - <https://curran.github.io/dataviz-course-archive/>
- Machine learning visualization
 - <https://goodekat.github.io/presentations/2020-isugg-iml/slides.html#1>
 - <https://medium.com/analytics-vidhya/awesome-machine-learning-visualizations-5208f1617ec5>

Thank you for your listening
Q & A