Visualise research and data

Mattias Vesterlund, KTH Library
(with slides from Sara Lind, KTH Library)

Agenda

- Data and research data
- Questions to ask of any visualisation
- The need for visualisations during data analysis
- How to select the right graph for your data
- Interactivity
- Design and layout
- Colour
- Posters
- Accessible visualisation
- Key takeaways
- Tools and resources

What is data?

Definition Wikipedia (computer science):

Data requires interpretation to become information. Digital data is data that is represented using the binary number system of ones (1) and zeros (0), instead of analog representation. In modern (post-1960) computer systems, all data is digital.

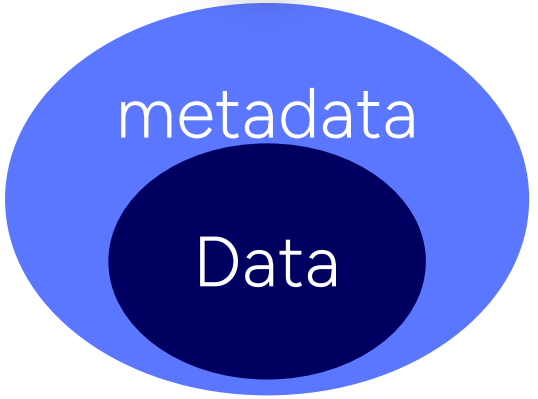
Definition according to EOSC:

reinterpretable digital representation of information in a formalized manner suitable for communication, interpretation, or processing.

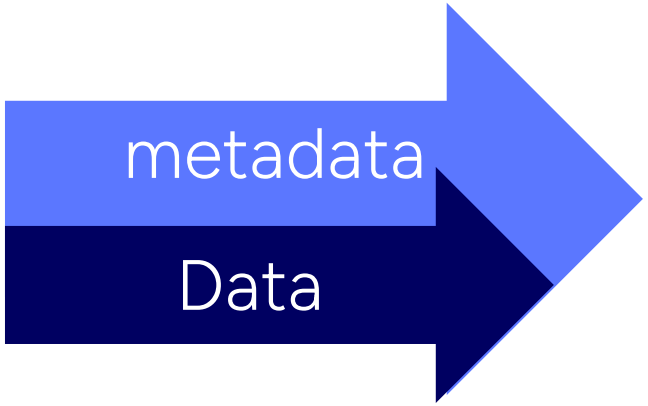
Bit sequence = 1 byte data

What is metadata?

Metadata = data about data, often the information needed to interpret data in a meaningful context, (technical, administrative, descriptive metadata)



Data-object



Data-flow

Big Data, Little Data, No Data: Scholarship in the Networked World
Christine L. Borgman, MIT Press (2017)

Good questions to ask before DataViz
(from 5W + 1H)

- What?
  - My Research Data
  - Our analysis workflow
  - A generalizable concept
  - An observed relationship
- Why?
  - To verify quality (QC)
  - To explore
  - To explain (to whom?)
- How?
  - What type of plot or diagram can I use to convey this?
- What aspect of the data?
  - Distribution
  - Comparison
  - Relationship
  - Composition

What?

- Data
  - Raw data
  - Transformed or normalised data
  - Transcribed data
  - Random data
  - Statistical calculations
  - Approximations, estimates or projections

- Conceptual
  - Graphical abstract
  - Infographics
  - Descriptive diagrams

Why?

- **Purpose** – What are you trying to achieve?
  - Quality control
  - Explore the data for analysis purposes
  - Explain the data or the results of an analysis
- **Audience**
  - Myself
  - Other researchers in my field (e.g. my PI)
  - Other researchers in a separate field
  - General public
- **Context**
  - Slide show presentation
  - Conference poster
  - Journal article

A QC example

- What?
  - Data on gene expression levels (**numerical values**) for 1000 genes in different cancer cell lines of different types
- Why?
  - Quality control – ensure that gene expression on an overall level (**distribution**) is similar.
- How?
  - Boxplot to see distribution of the numerical value (gene expression) for each cell line.

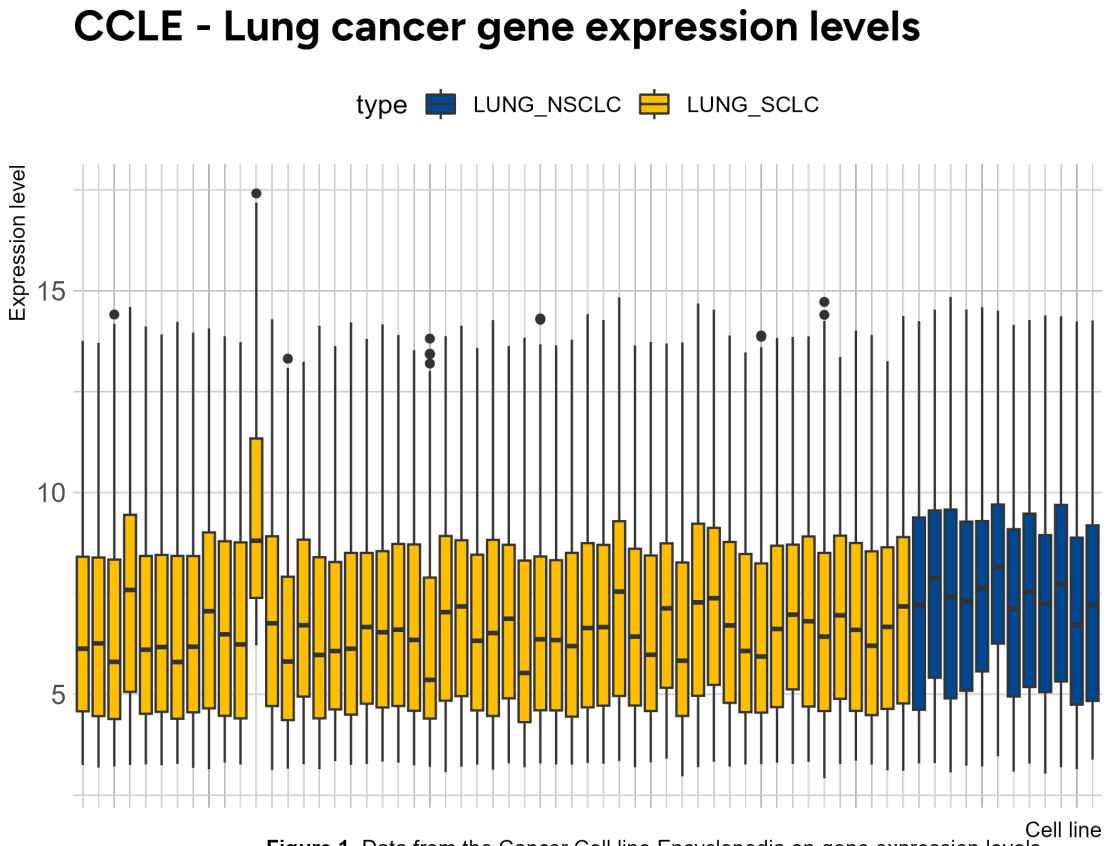


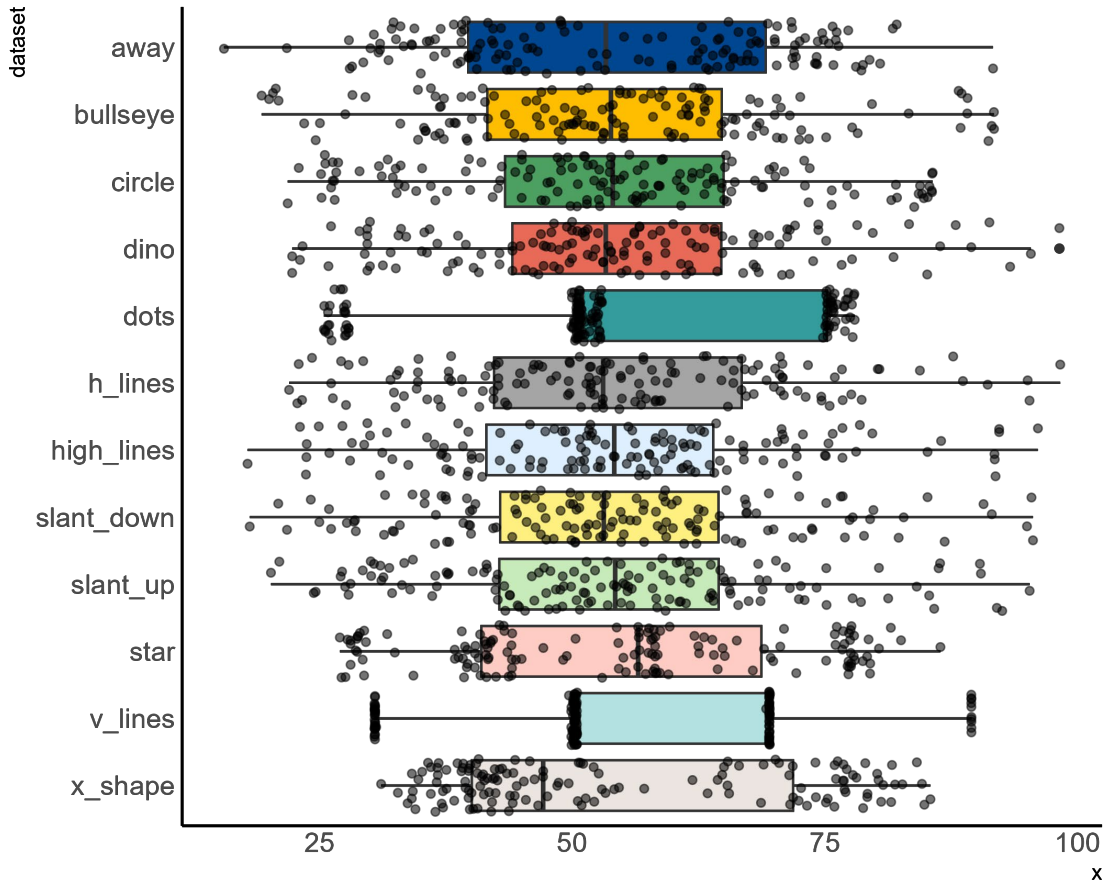
Figure 1. Data from the Cancer Cell line Encyclopedia on gene expression levels (RNA) for Small Cell Lung Cancer and Non-Small Cell Lung Cancer cell lines.

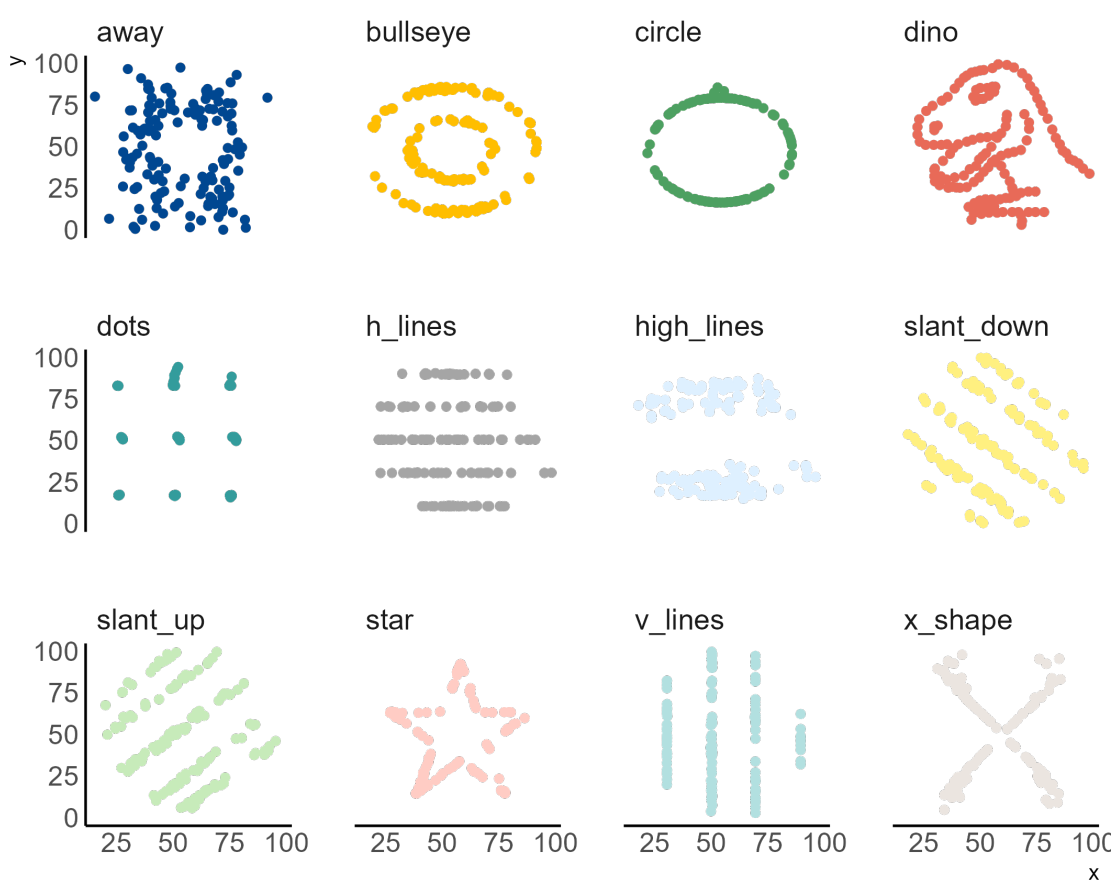
The danger of summary statistics

	dino_x	dino_y	away_x	away_y	h_lines_x	h_lines_y	...
1	55.38	97.18	32.33	61.41	53.37	90.21	...
2	51.54	96.03	53.42	26.19	52.80	90.09	...
3	46.15	94.49	63.92	30.83	47.05	90.46	...
4	42.82	91.41	70.29	82.53	42.45	89.51	...
5	40.77	88.33	34.12	45.73	42.70	90.44	...
6	38.72	84.87	67.67	37.11	32.38	90.14	...
7	35.64	79.87	53.26	97.48	32.53	70.16	...
8	33.08	77.56	63.51	25.10	33.37	70.46	...
9	28.97	74.49	67.98	80.96	32.66	70.05	...
10	26.15	71.41	67.37	29.72	22.96	70.43	...
11	23.08	66.41	15.56	80.07	27.15	70.21	...
12	22.31	61.79	71.79	71.07	26.17	70.50	...

	mean_x	mean_y	std_dev_x	std_dev_y	corr_x_y
away	54.266	47.835	16.77	26.94	-0.064
bullseye	54.269	47.831	16.769	26.936	-0.069
circle	54.267	47.838	16.76	26.93	-0.068
dino	54.263	47.832	16.765	26.935	-0.064
dots	54.26	47.84	16.768	26.93	-0.06
h_lines	54.261	47.83	16.766	26.94	-0.062
high_lines	54.269	47.835	16.767	26.94	-0.069
slant_down	54.268	47.836	16.767	26.936	-0.069
slant_up	54.266	47.831	16.769	26.939	-0.069
star	54.267	47.84	16.769	26.93	-0.063
v_lines	54.27	47.837	16.77	26.938	-0.069
x_shape	54.26	47.84	16.77	26.93	-0.066

The danger of summary statistics

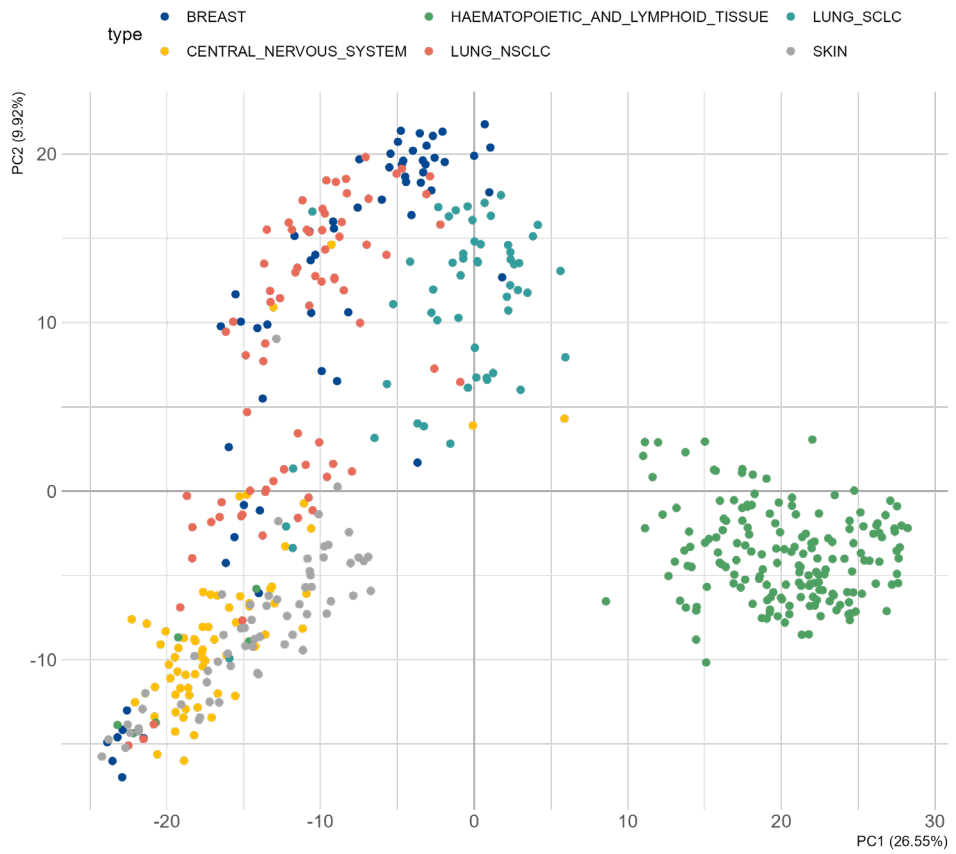




Dimensionality reduction for larger datasets

Principal Component Analysis

CCLE PC1 vs PC2

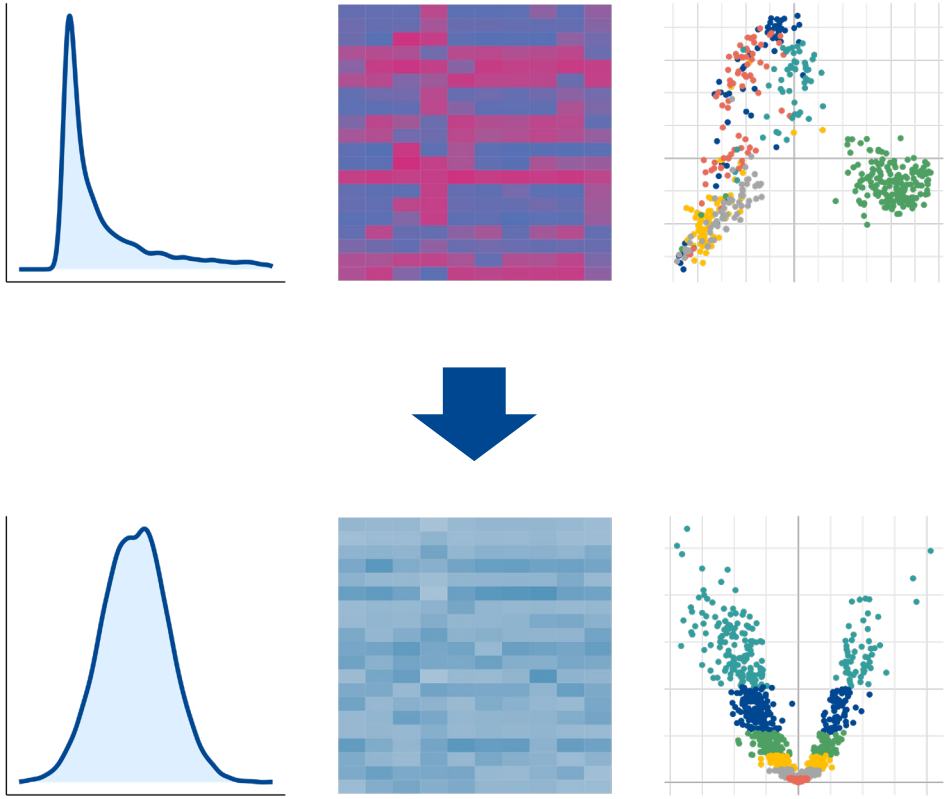


Uniform Manifold Approximation and Projection (UMAP)

CCLE UMAP

Exploring data with visualisations

- Try out different visualisations
- Look at your data from many angles
- Follow how transformations impact your data by visualising it.
  Think before -> after, but include intermediary steps too
- Interactivity can be a great help, especially for "big" data.



How?

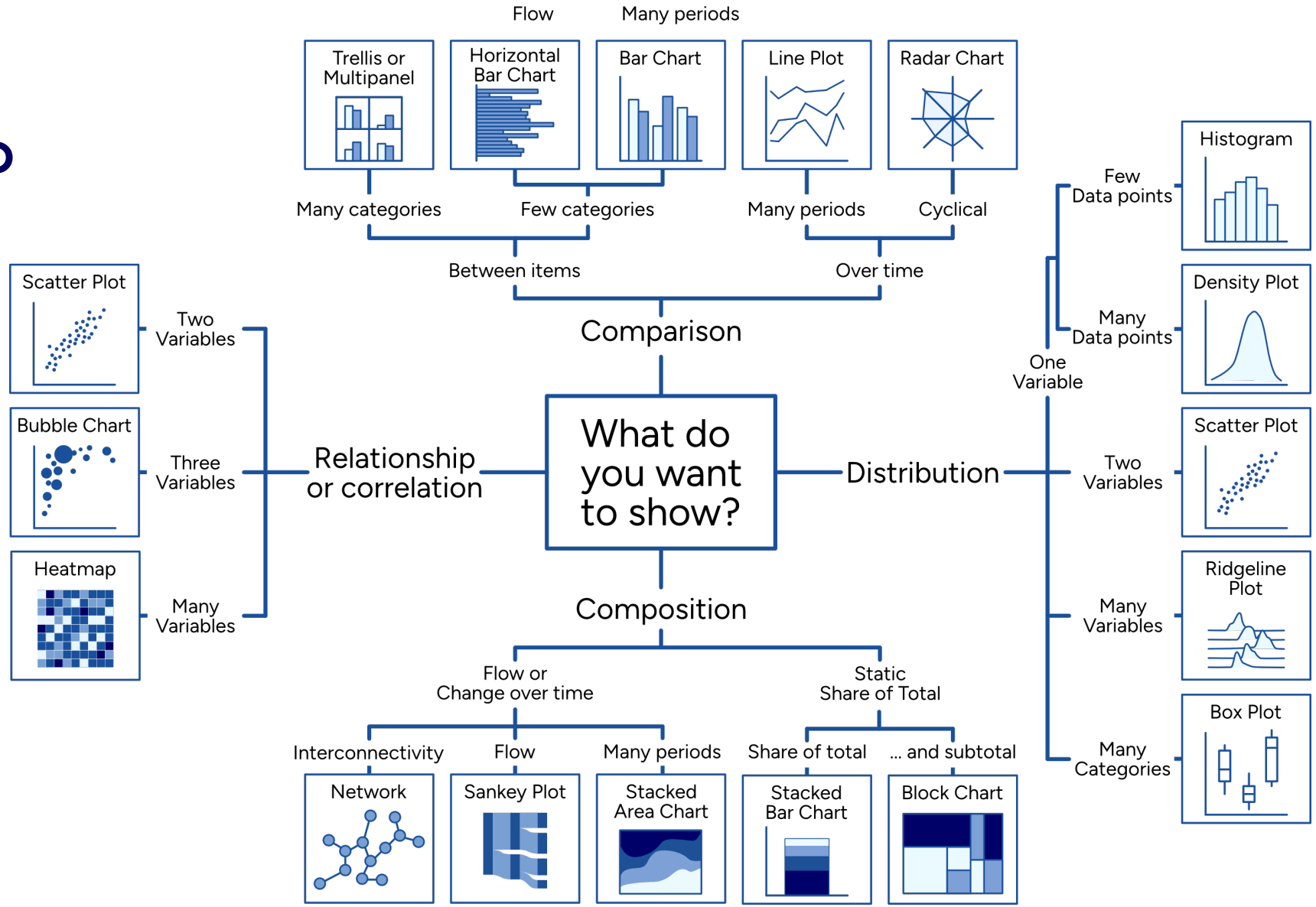
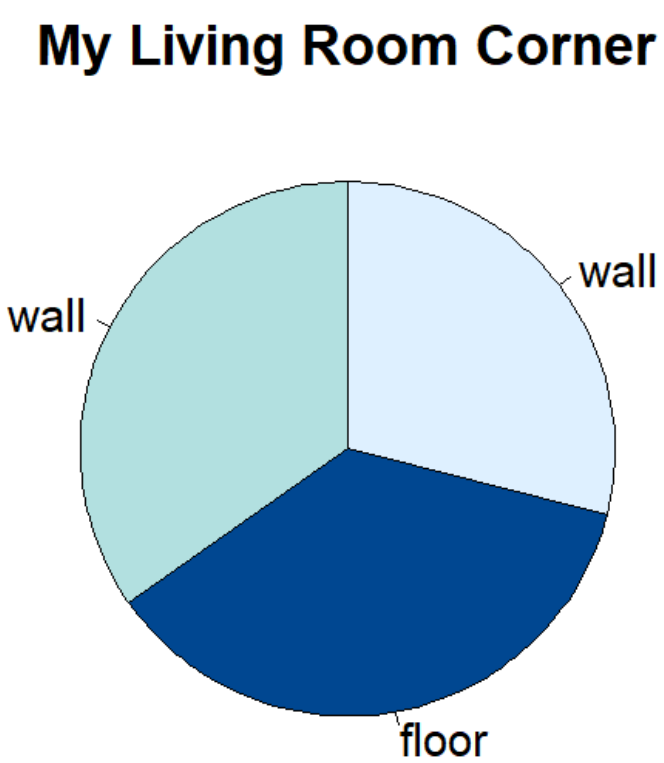
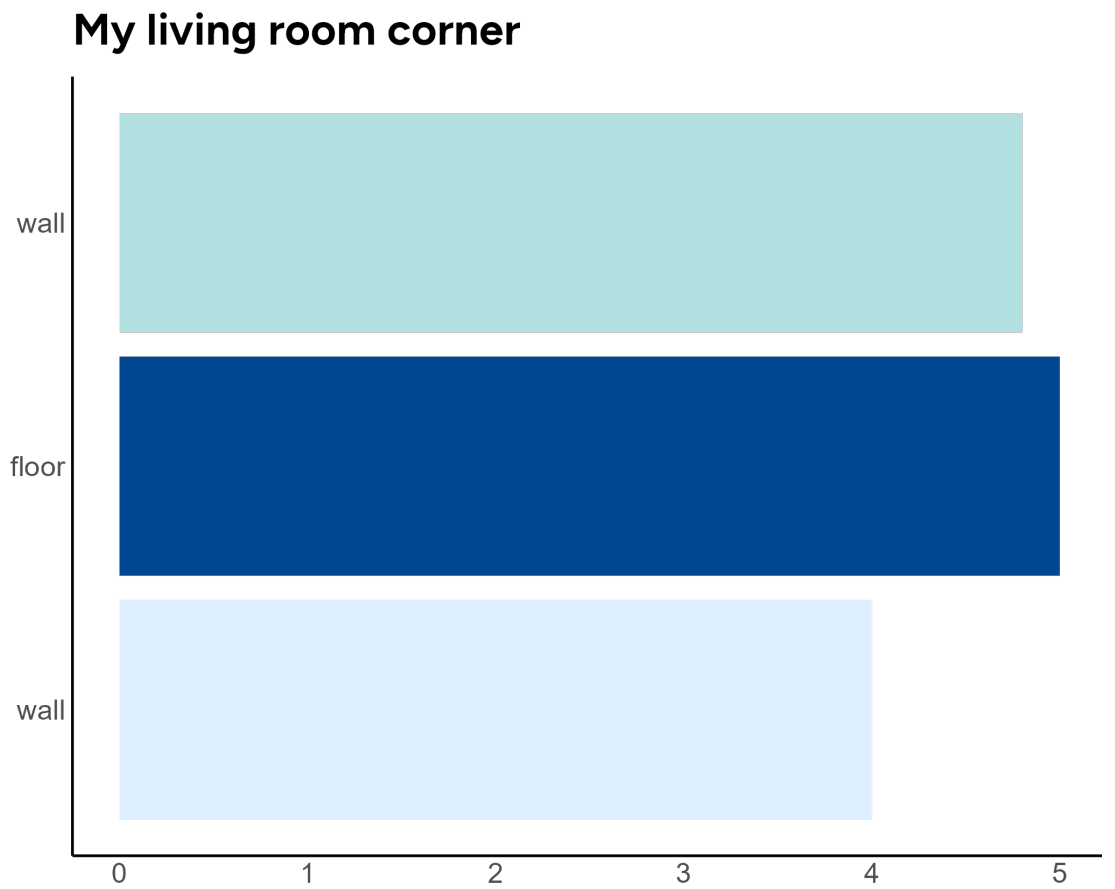


Image inspired by Nadieh Bremer

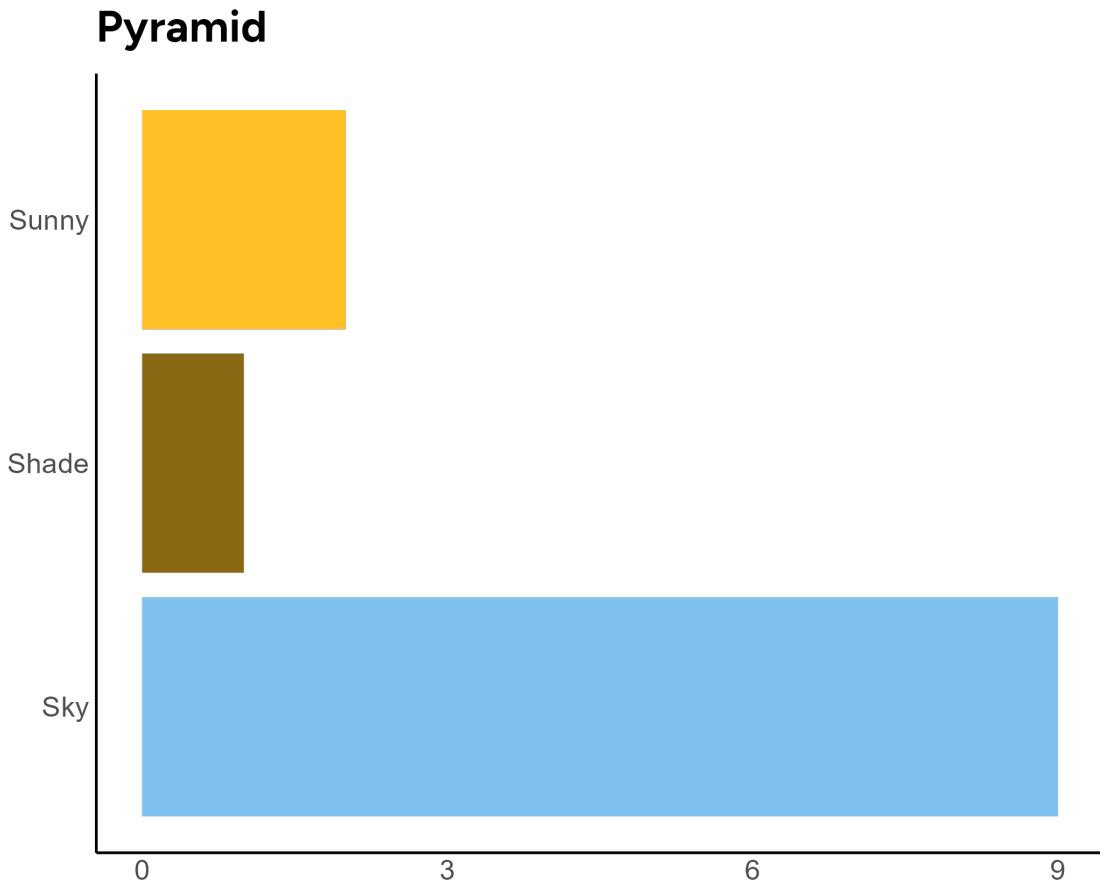
Data-to-Viz DataVizcatalogue.com

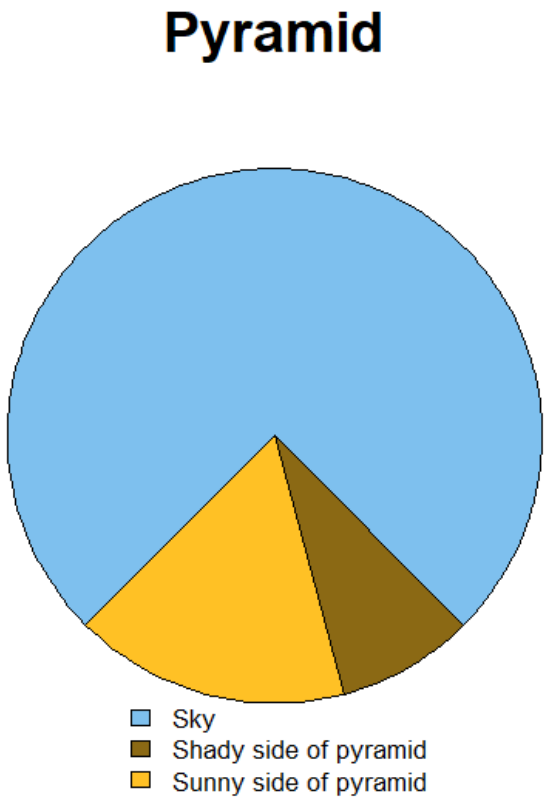
Pie Charts (maybe avoid these)





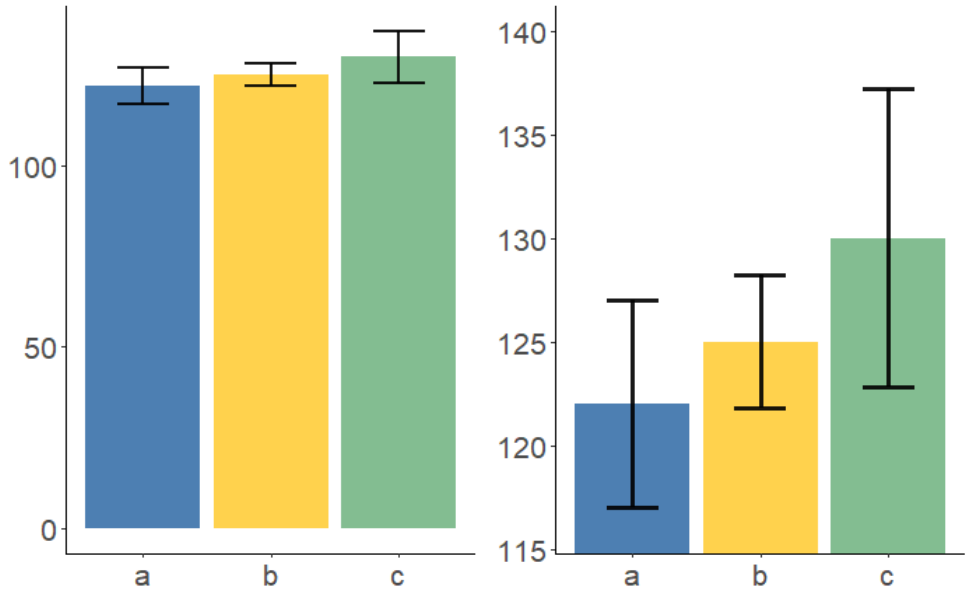
Pie Charts (maybe avoid these)





Pie Charts (maybe avoid these)

- But bar graphs are not without problems:



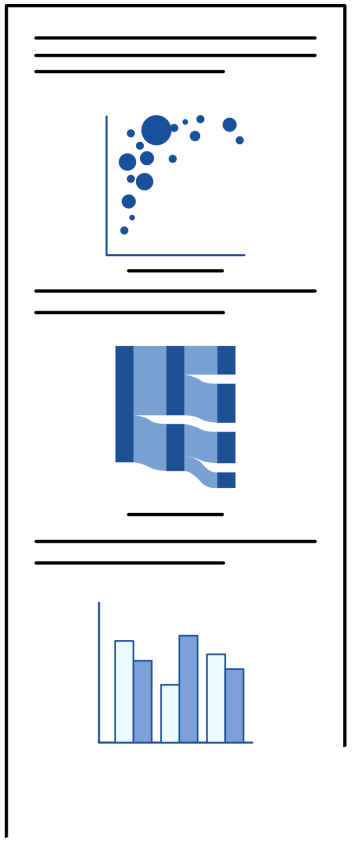
Further reading:

- Quantifying Data Distortion in Bar Graphs in Biological Research
- Bad bar charts distort data — and pervade biology

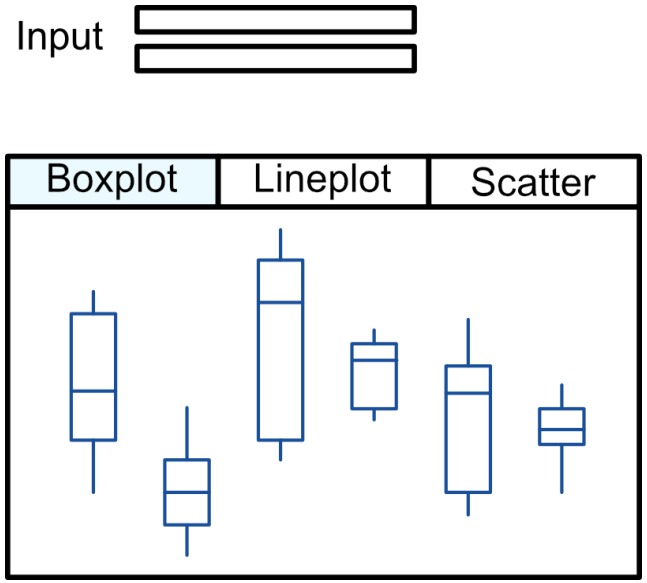
Interactivity

Interactivity - examples

Linear flow



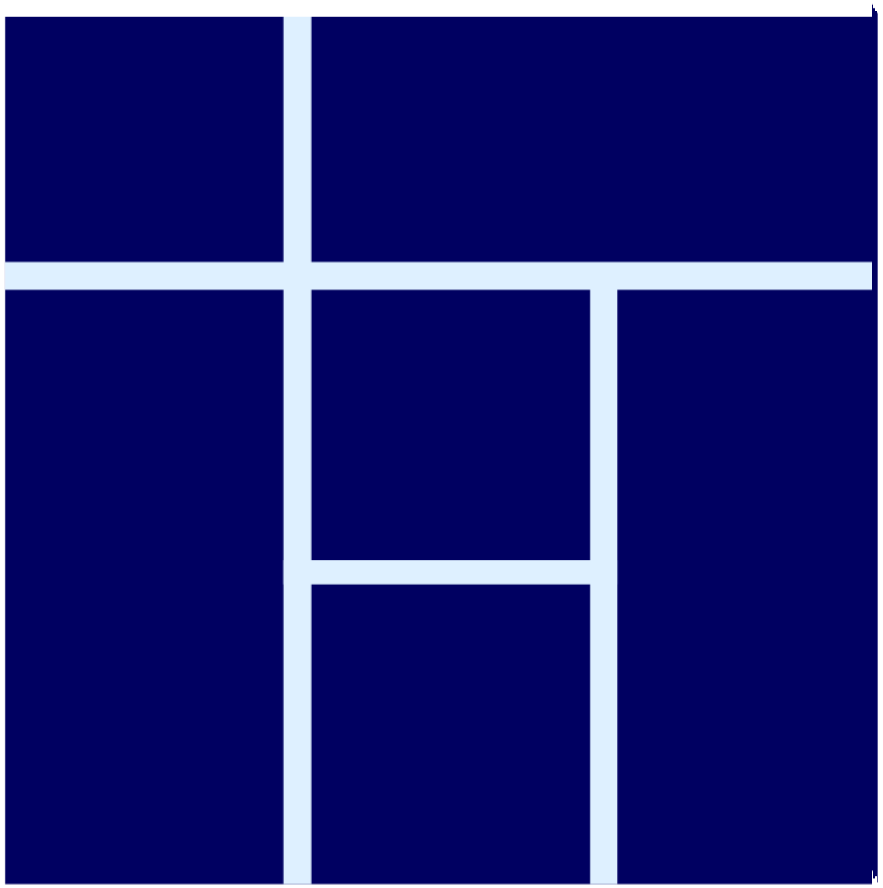
Dashboard



(for example code see: github:Dataviz2025)

Design and layout

Design principles



Layout

Graphics inspiration from Karolinska Institutet Universitetsbiblioteket

Design on paper first

Colour

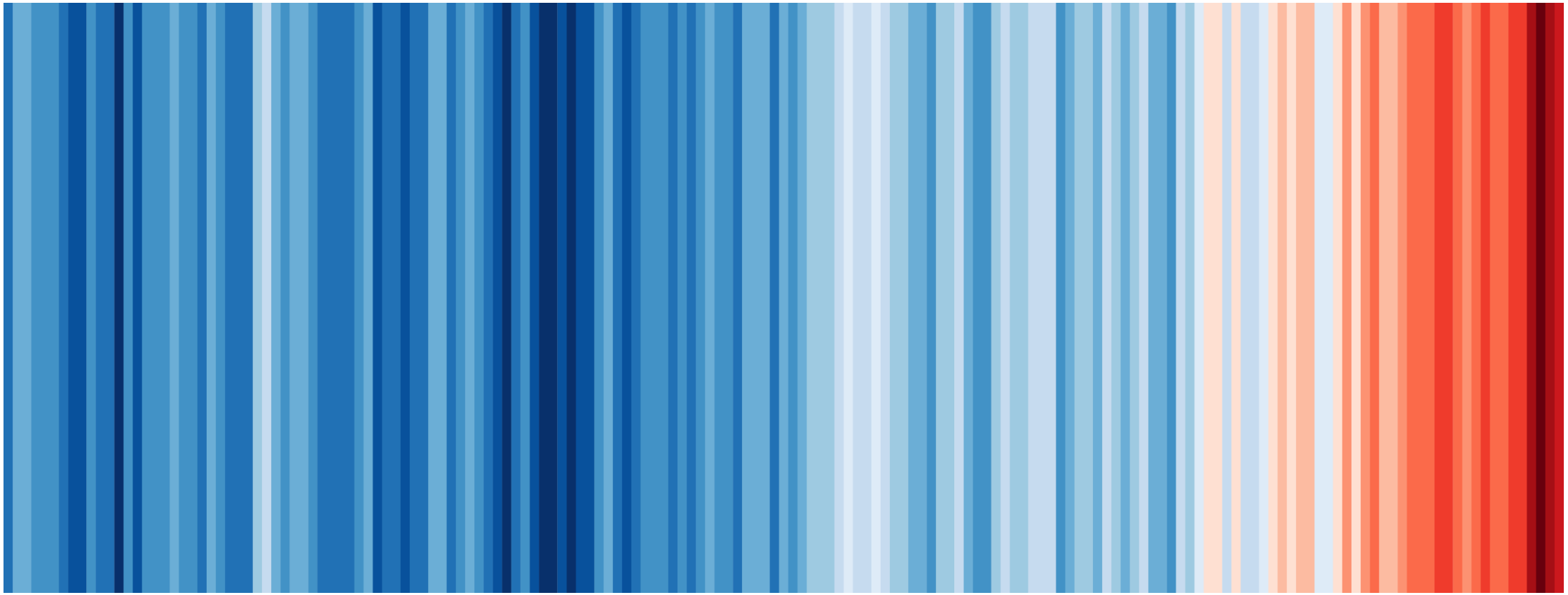
A question you always need to ask – Do I need colour?

"Too often, we ask how we can use color in our visualizations when we should be asking why we are using it."

Rotschild, C. 2020. Color in Data Visualization: Less How, More Why
https://connorrothschild.github.io/v1/color-in-data-vis/

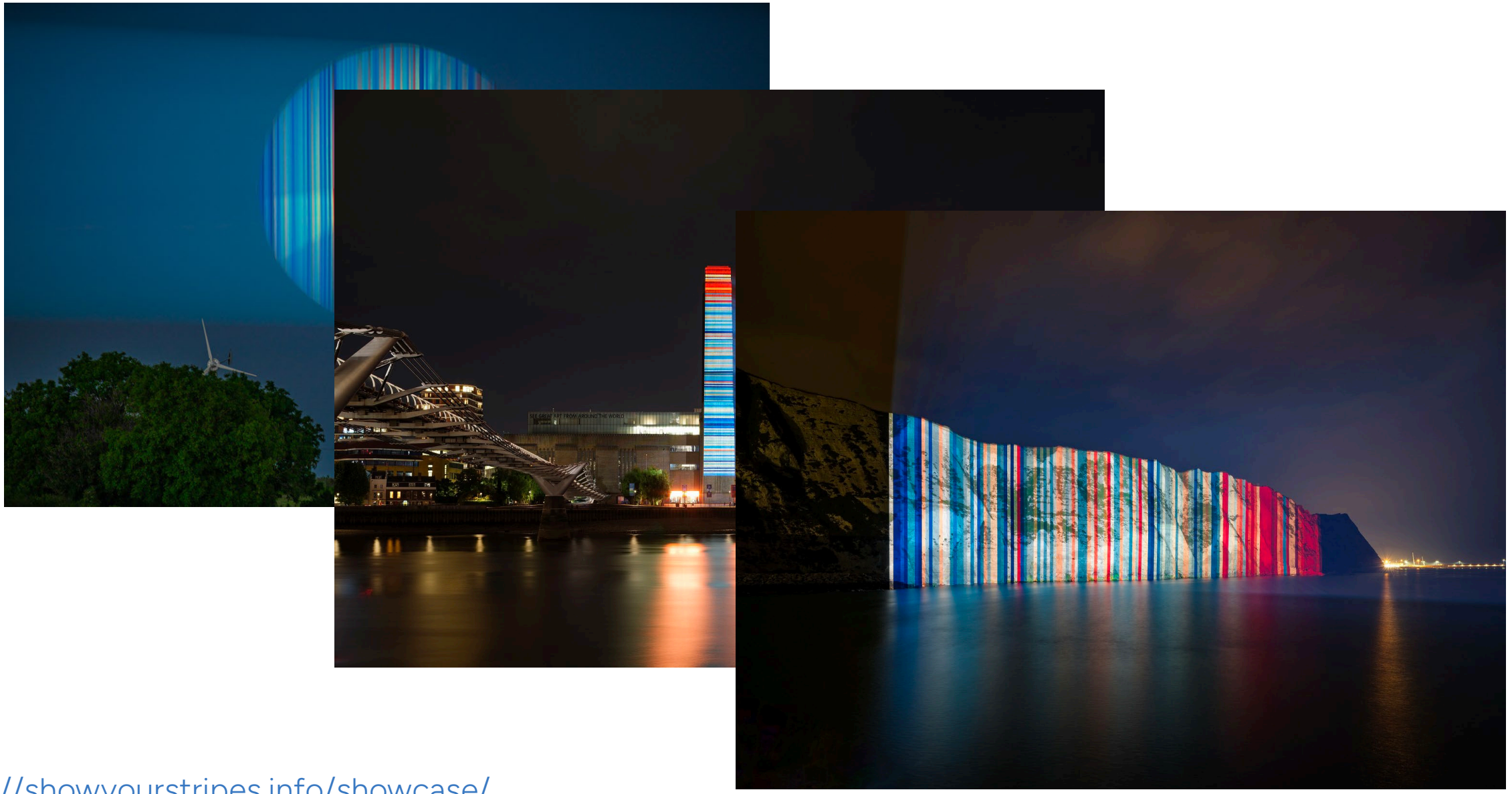
Photo by Steve Johnson on Unsplash

Colours have meaning



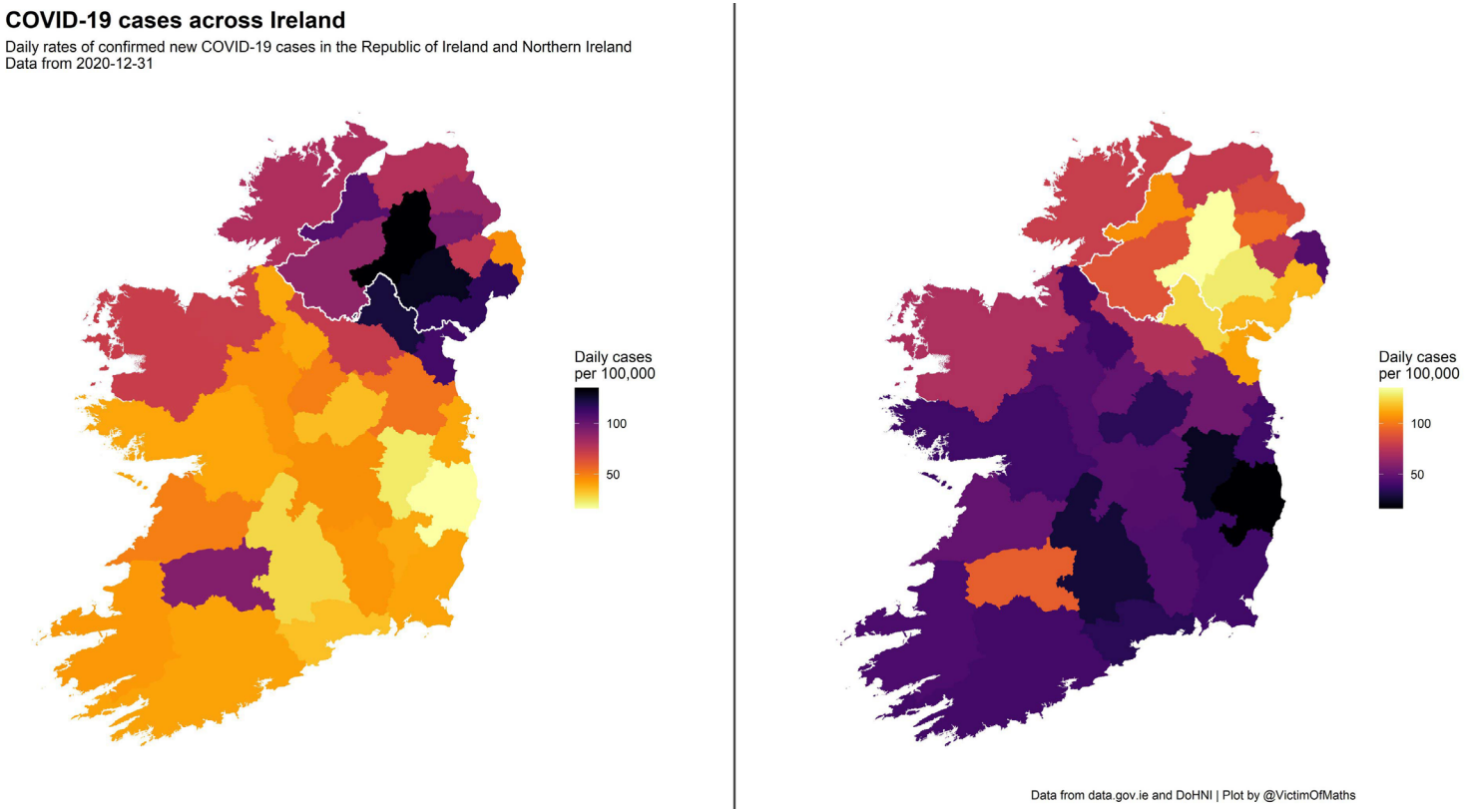
Colours can have a culturally conditioned meaning

Photo credit: University of Reading
showyourstripes.info/showcase/



https://showyourstripes.info/showcase/

Colours and density

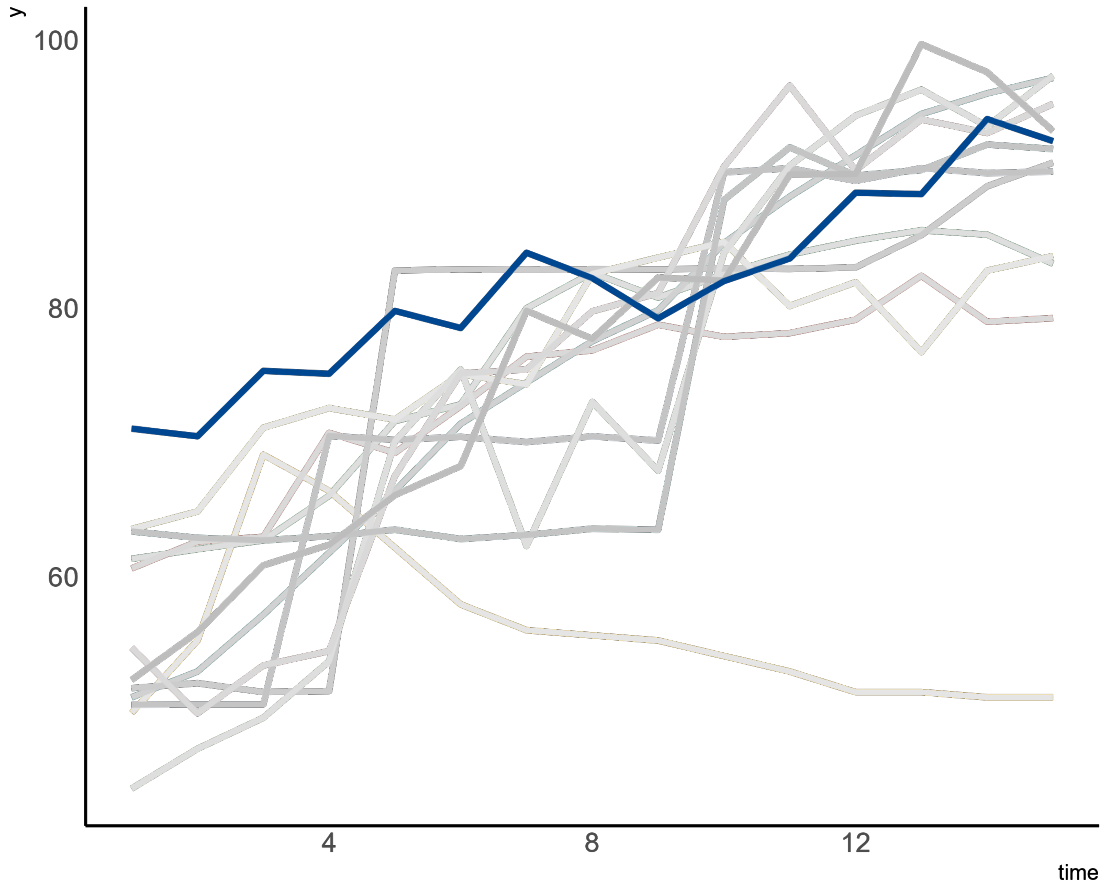


The colour shade is culturally conditioned to show density (light low, dark high)

Bildkälla: Angus, Colin (2021): CoVid Plots and Analysis. The University of Sheffield. Dataset. https://doi.org/10.15131/shef.data.12328226 CC BY SA

Make grey your best friend

Direct attention (but don’t decieve)

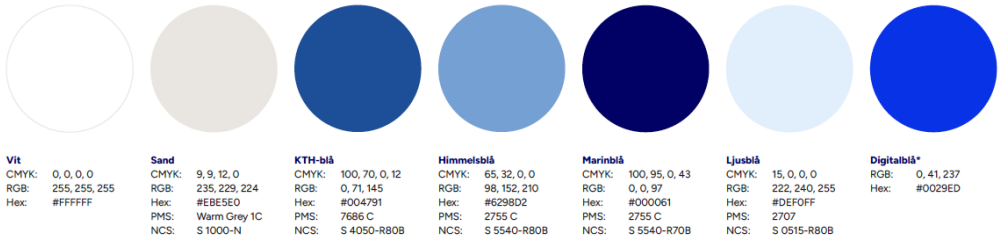


Where colours don't exist

Source: David S. Goodsell, RCSB Protein Data Bank.
doi: 10.2210/rcsb_pdb/goodsell-gallery-030

Use the pre-defined colours of your organisation

Main colours (use more)



Vit
CMYK: 0, 0, 0, 0
RGB: 255, 255, 255
Hex: #FFFFFF

Sand
CMYK: 9, 9, 12, 0
RGB: 235, 229, 224
Hex: #EBE5E0
PMS: Warm Grey 1C
NCS: S 1000-N

KTH-blå
CMYK: 100, 70, 0, 12
RGB: 0, 71, 145
Hex: #004791
PMS: 7686 C
NCS: S 4050-R80B

Himmelsblå
CMYK: 65, 32, 0, 0
RGB: 98, 152, 210
Hex: #6298D2
PMS: 2755 C
NCS: S 5540-R80B

Marinblå
CMYK: 100, 95, 0, 43
RGB: 0, 0, 97
Hex: #000061
PMS: 2755 C
NCS: S 5540-R70B

Ljusblå
CMYK: 15, 0, 0, 0
RGB: 222, 240, 255
Hex: #DEF0FF
PMS: 2707
NCS: S 0515-R80B

Digitalblå*
RGB: 0, 41, 237
Hex: #0029ED

Secondary colours (use when needed)

Mörkgrön
CMYK: 75, 5, 75, 70
RGB: 13, 74, 33
Hex: #0D4A21

Mörk turkos
CMYK: 80, 45, 40, 57
RGB: 28, 67, 76
Hex: #1C434C

Mörk tegel
CMYK: 10, 100, 60, 57
RGB: 120, 0, 26
Hex: #78001A

Mörkgul
CMYK: 25, 65, 100, 20
RGB: 166, 89, 0
Hex: #A65900

Mörkgrå
CMYK: 0, 0, 0, 90
RGB: 50, 50, 50
Hex: #323232

Bruten svart*
RGB: 33, 33, 33
Hex: #212121

Grön
CMYK: 72, 13, 75, 0
RGB: 77, 160, 97
Hex: #4DA060

Turkos
CMYK: 72, 17, 38, 7
RGB: 51, 156, 156
Hex: #339C9C

Tegel
CMYK: 5, 70, 64, 0
RGB: 232, 106, 88
Hex: #E86A58

Gul
CMYK: 0, 30, 95, 0
RGB: 255, 190, 0
Hex: #FFBE00

Grå
CMYK: 0, 0, 0, 45
RGB: 165, 165, 165
Hex: #A5A5A5

Bruten vit*
RGB: 252, 252, 252
Hex: #FCFCFC

Ljusgrön
CMYK: 27, 0, 36, 0
RGB: 199, 235, 186
Hex: #C7EBBA

Ljus turkos
CMYK: 34, 0, 16, 0
RGB: 178, 224, 224
Hex: #B2E0E0

Ljus tegel
CMYK: 0, 28, 18, 0
RGB: 255, 204, 196
Hex: #FFCCC4

Ljusgul
CMYK: 0, 4, 39, 0
RGB: 255, 240, 176
Hex: #FFF0B0

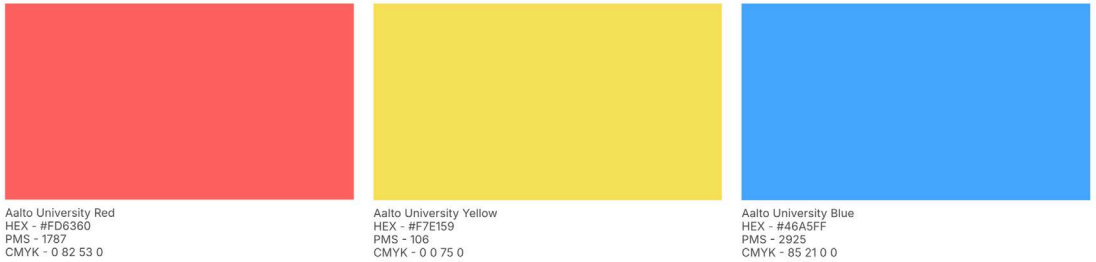
Ljusgrå
CMYK: 0, 0, 0, 15
RGB: 230, 230, 230
Hex: #E6E6E6

* Bruten svart och bruten vit används enbart som textfärg på webb för att förbättre läsbarhet och tillgänglighet på skärm.

[KTH Visual Identity]

Use the pre-defined colours of your organisation

Brand colours

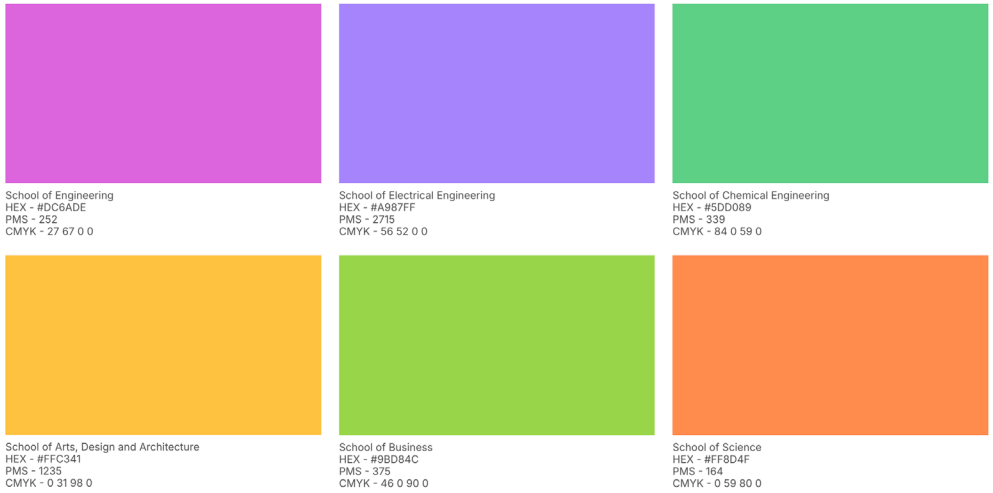


Aalto University Red
HEX - #FD6360
PMS - 1787
CMYK - 0 82 53 0

Aalto University Yellow
HEX - #F7E159
PMS - 106
CMYK - 0 0 75 0

Aalto University Blue
HEX - #46A5FF
PMS - 2925
CMYK - 85 21 0 0

School colours



School of Engineering
HEX - #DC6ADE
PMS - 252
CMYK - 27 67 0 0

School of Electrical Engineering
HEX - #A987FF
PMS - 2715
CMYK - 56 52 0 0

School of Chemical Engineering
HEX - #5DD089
PMS - 339
CMYK - 84 0 59 0

School of Arts, Design and Architecture
HEX - #FFC341
PMS - 1235
CMYK - 0 31 98 0

School of Business
HEX - #9BD84C
PMS - 375
CMYK - 46 0 90 0

School of Science
HEX - #FF8D4F
PMS - 164
CMYK - 0 59 80 0

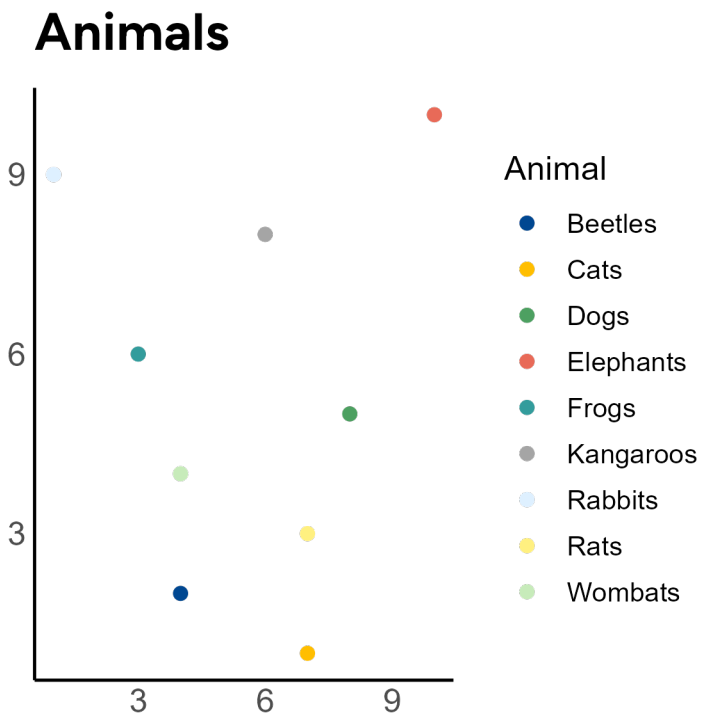
[Aalto Colour Palette]

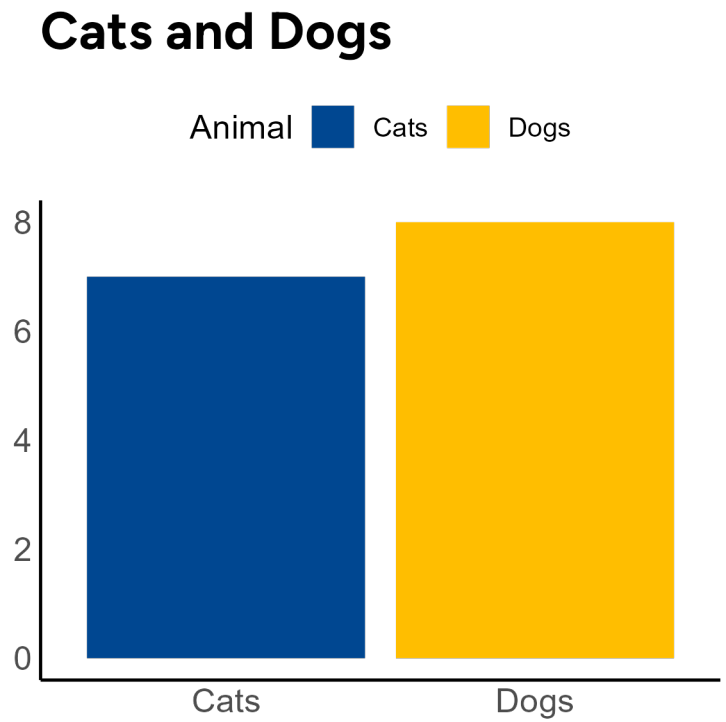
Use a colour theme

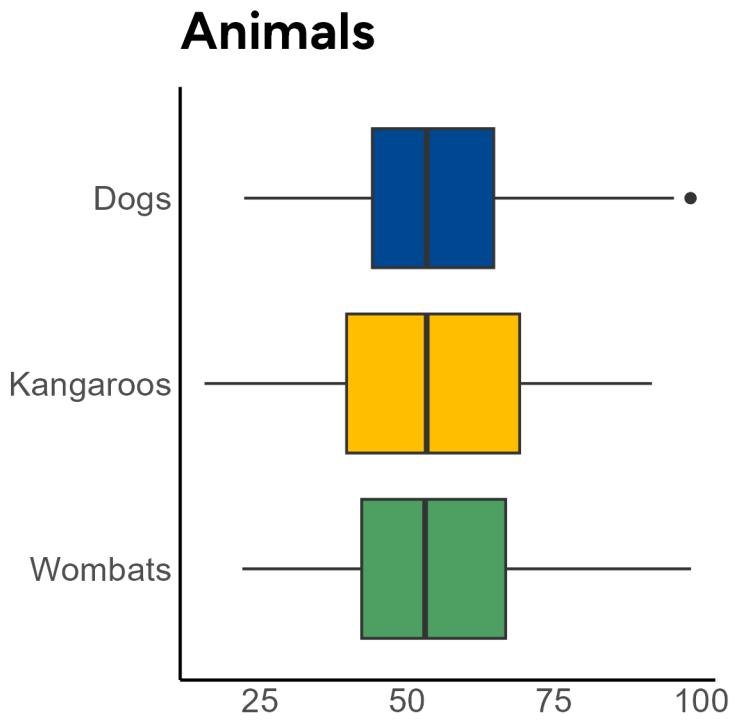
And stick to it

If the colour palette doesn't align with your data, use other colours. Make a palette that emphasizes your data while letting brand colours be present in the layout.
Be consistent with your chosen theme.

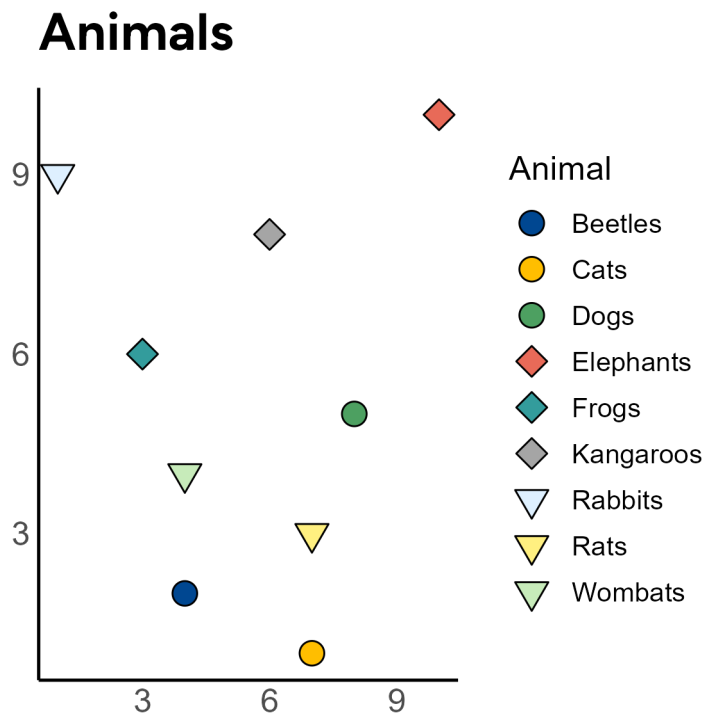
Consistent coloring

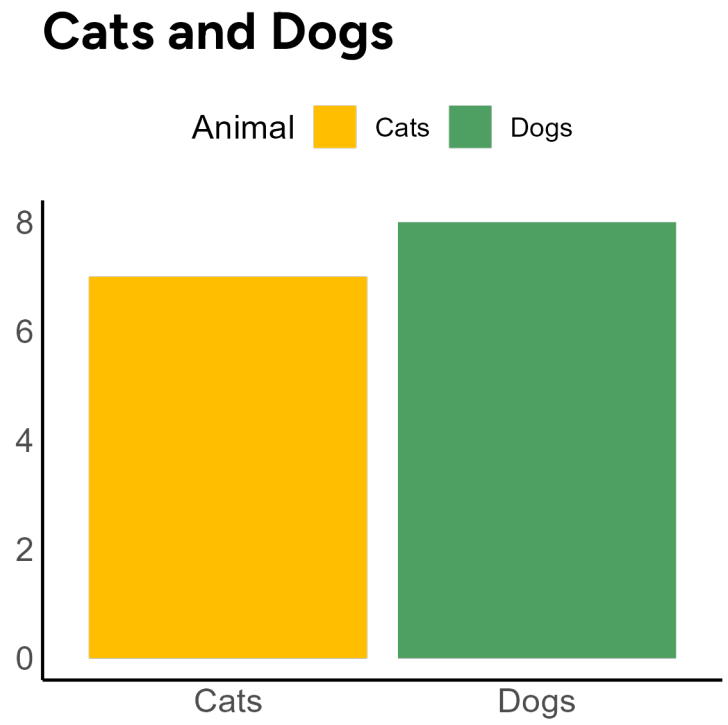


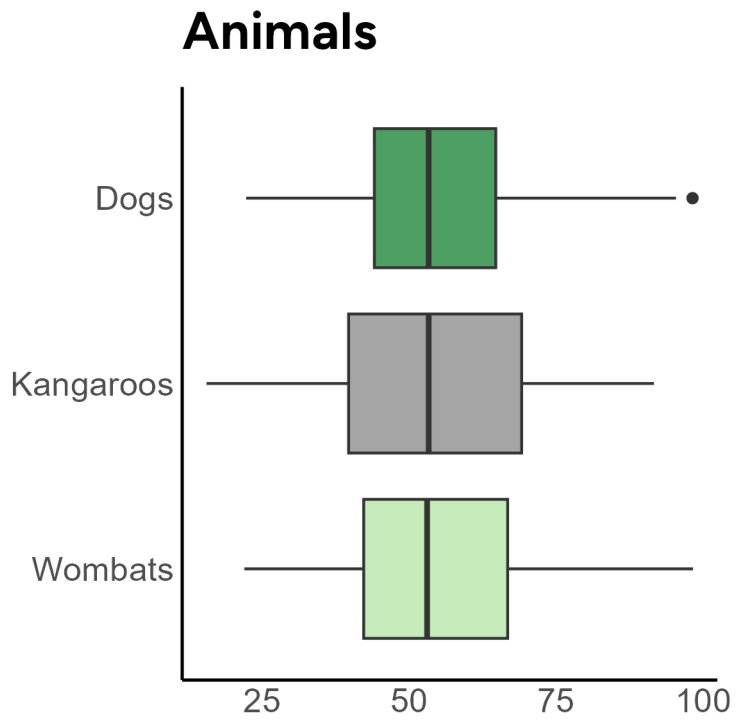




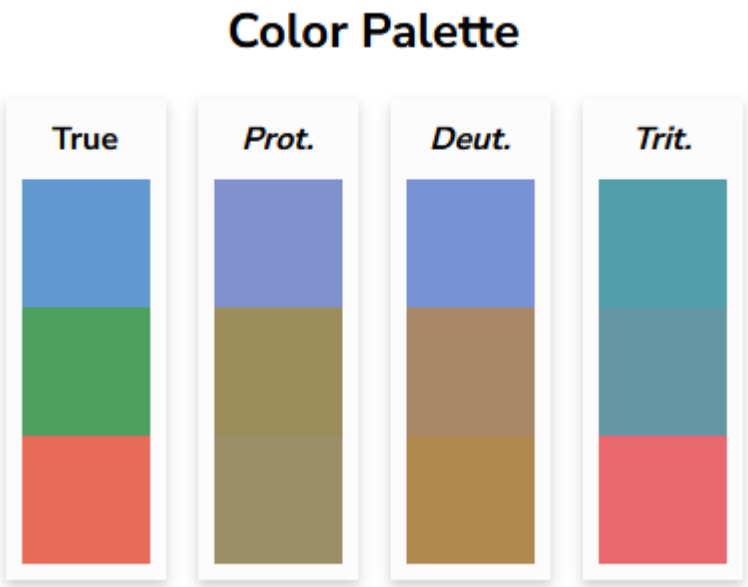
Consistent coloring



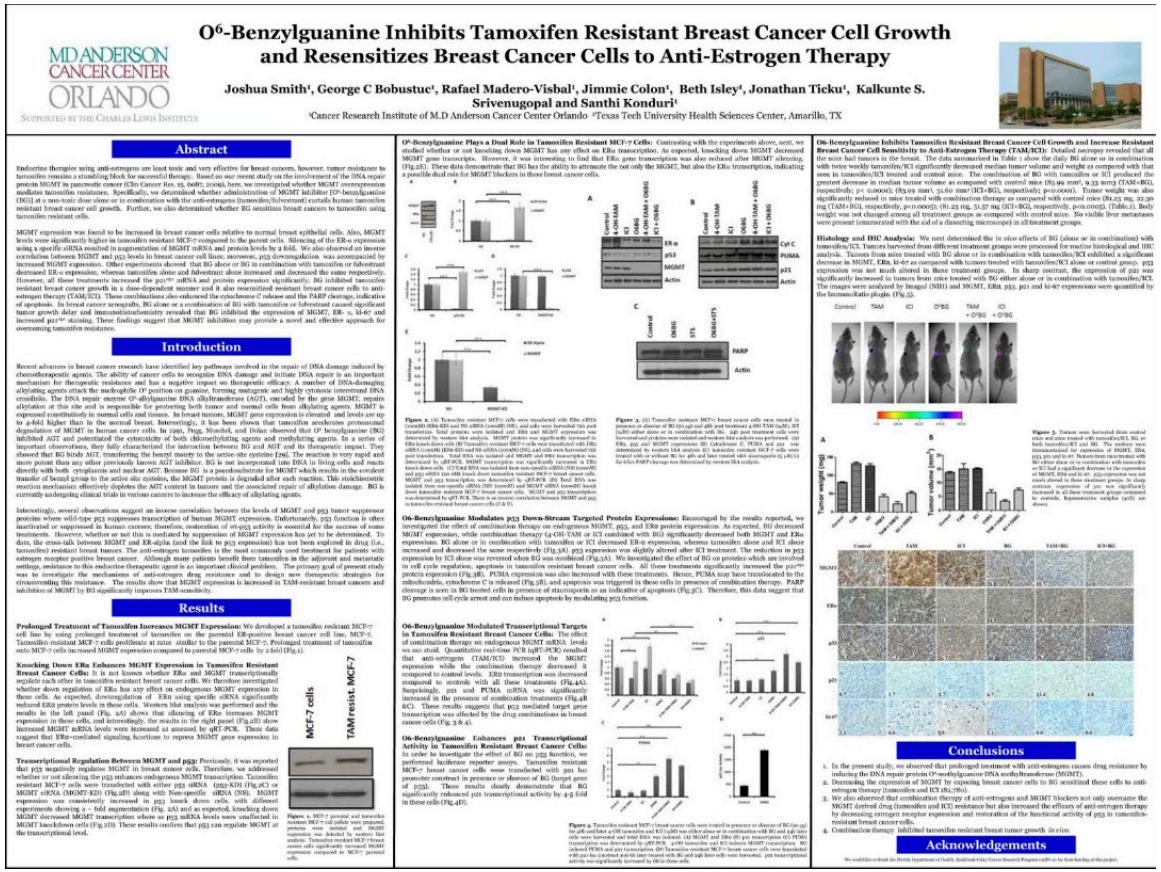




Mind the color blind!



Negative space is your friend



O6-Benzylguanine Inhibits Tamoxifen Resistant Breast Cancer Cell Growth and Resensitizes Breast Cancer Cells to Anti-Estrogen Therapy

Joshua Smith[1], George C Bobustuc[1], Rafael Madero-Visbal[1], Jimmie Colon[1], Beth Isley[1], Jonathan Ticku[1], Kalkunte S. Srivenugopal and Santhi Konduri[1]

[1]Cancer Research Institute of M.D Anderson Cancer Center Orlando [2]Texas Tech University Health Sciences Center, Amarillo, TX

Abstract

Endocrine therapies using anti-estrogens are least toxic and very effective for breast cancers, however, tumor resistance to tamoxifen remains a stumbling block for successful therapy. Based on our recent study on the involvement of the DNA repair protein MGMT in pancreatic cancer (Clin Cancer Res. 13, 6687, 2006), here, we investigated whether MGMT overexpression mediates tamoxifen resistance. Specifically, we determined whether administration of MGMT inhibitor [O6-benzylguanine (BG)] at a non-toxic dose alone or in combination with the anti-estrogens (tamoxifen/fulvestrant) curtails human tamoxifen resistant breast cancer cell growth. Further, we also determined whether BG sensitizes breast cancers to tamoxifen using tamoxifen resistant cells.

MGMT expression was found to be increased in breast cancer cells relative to normal breast epithelial cells. Also, MGMT levels were significantly higher in tamoxifen resistant MCF-7 compared to the parent cells. Silencing of the ER-α expression using a specific siRNA resulted in augmentation of MGMT mRNA and protein levels by 2 fold. We also observed an inverse correlation between MGMT and p53 levels in breast cancer cell lines; moreover, p53 downregulation was accompanied by increased MGMT expression. Other experiments showed that BG alone or BG in combination with tamoxifen or fulvestrant decreased ER-α expression, whereas tamoxifen alone and fulvestrant alone increased and decreased the same respectively. However, all these treatments increased the $p21^{WAF}$ mRNA and protein expression significantly. BG inhibited tamoxifen resistant breast cancer growth in a dose-dependent manner and it also resensitized resistant breast cancer cells to anti-estrogen therapy (TAM/ICI). These combinations also enhanced the cytochrome C release and the PARP cleavage, indicative of apoptosis. In breast cancer xenografts, BG alone or a combination of BG with tamoxifen or fulvestrant caused significant tumor growth delay and immunohistochemistry revealed that BG inhibited the expression of MGMT, ER-α, ki-67 and increased $p21^{WAF}$ staining. These findings suggest that MGMT inhibition may provide a novel and effective approach for overcoming tamoxifen resistance.

Introduction

Recent advances in breast cancer research have identified key pathways involved in the repair of DNA damage induced by chemotherapeutic agents. The ability of cancer cells to recognize DNA damage and initiate DNA repair is an important mechanism for therapeutic resistance and has a negative impact on therapeutic efficacy. A number of DNA-damaging alkylating agents attack the nucleophile O6 position on guanine, forming mutagenic and highly cytotoxic interstrand DNA crosslinks. The DNA repair enzyme O6-alkylguanine DNA alkyltransferase (AGT), encoded by the gene MGMT, repairs alkylation at this site and is responsible for protecting both tumor and normal cells from alkylating agents. MGMT is expressed constitutively in normal cells and tissues. In breast tumors, MGMT gene expression is elevated and levels are up to 4-fold higher than in the normal breast. Interestingly, it has been shown that tamoxifen accelerates proteasomal degradation of MGMT in human cancer cells. In 1991, Pegg, Moschel, and Dolan observed that O6 benzylguanine (BG) inhibited AGT and potentiated the cytotoxicity of both chloroethylating agents and methylating agents. In a series of important observations, they fully characterized the interaction between BG and AGT and its therapeutic impact. They showed that BG binds AGT, transferring the benzyl moiety to the active-site cysteine [1m]. The reaction is very rapid and more potent than any other previously known AGT inhibitor. BG is not incorporated into DNA in living cells and reacts directly with both cytoplasmic and nuclear AGT. Because BG is a pseudosubstrate for MGMT which results in the covalent transfer of benzyl group to the active site cysteine, the MGMT protein is degraded after each reaction. This stoichiometric reaction mechanism effectively depletes the AGT content in tumors and the associated repair of alkylation damage. BG is currently undergoing clinical trials in various cancers to increase the efficacy of alkylating agents.

Interestingly, several observations suggest an inverse correlation between the levels of MGMT and p53 tumor suppressor proteins where wild-type p53 suppresses transcription of human MGMT expression. Unfortunately, p53 function is often inactivated or suppressed in human cancers; therefore, restoration of wt-p53 activity is essential for the success of many treatments. However, whether or not this is mediated by suppression of MGMT expression has yet to be determined. To date, the cross-talk between MGMT and ER-alpha (and the link to p53 expression) has not been explored in drug (i.e., tamoxifen) resistant breast tumors. The anti-estrogen tamoxifen is the most commonly used treatment for patients with estrogen receptor positive breast cancer. Although many patients benefit from tamoxifen in the adjuvant and metastatic settings, resistance to this endocrine therapeutic agent is an important clinical problem. The primary goal of present study was to investigate the mechanisms of anti-estrogen drug resistance and to design new therapeutic strategies for circumventing this resistance. The results show that MGMT expression is increased in TAM-resistant breast cancers and inhibition of MGMT by BG significantly improves TAM-sensitivity.

Results

Prolonged Treatment of Tamoxifen Increases MGMT Expression: We developed a tamoxifen resistant MCF-7 cell line by using prolonged treatment of tamoxifen on the parental ER-positive breast cancer cell line, MCF-7. Tamoxifen-resistant MCF-7 cells proliferate at rates similar to the parental MCF-7. Prolonged treatment of tamoxifen onto MCF-7 cells increased MGMT expression as compared to parental MCF-7 cells by 2-fold (Fig.1).

Knocking Down ERα Enhances MGMT Expression in Tamoxifen Resistant Breast Cancer Cells: It is not known whether ERα and MGMT transcriptionally regulate each other in tamoxifen resistant breast cancer cells. We therefore investigated whether down regulation of ERα has any effect on endogenous MGMT expression in these cells. As expected, downregulation of ERα using specific siRNA significantly reduced ERα protein levels in these cells. Western blot analysis was performed and the results in the left panel (Fig. 2A) shows that silencing of ERα increases MGMT expression in these cells, and interestingly, the results in the right panel (Fig.2B) show increased MGMT mRNA levels were increased as assessed by qRT-PCR. These data suggest that ERα-mediated signaling functions to repress MGMT gene expression in breast cancer cells.

Transcriptional Regulation Between MGMT and p53: Previously, it was reported that p53 negatively regulates MGMT in breast cancer cells. Therefore, we addressed whether or not silencing the p53 enhances endogenous MGMT transcription. Tamoxifen resistant MCF-7 cells were transfected with either p53 siRNA (p53-KD) (Fig.3C) or MGMT siRNA (MGMT-KD) (Fig.3D) along with Non-specific siRNA (NS). MGMT expression was consistently increased in p53 knock down cells, with different experiments showing a ~ 6-fold augmentation (Fig. 3A) and as expected, knocking down MGMT decreased MGMT transcription whereas as p53 mRNA levels were unaffected in MGMT knockdown cells (Fig.3B). These results confirm that p53 can regulate MGMT at the transcriptional level.

O6-Benzylguanine Plays a Dual Role in Tamoxifen Resistant MCF-7 Cells: Contrasting with the experiments above, next, we studied whether or not knocking down MGMT has any effect on ERα transcription. As expected, knocking down MGMT decreased MGMT gene transcripts. However, it was interesting to find that ERα gene transcription was also reduced after MGMT silencing. (Fig.2E). These data demonstrate that BG has the ability to attenuate the not only the MGMT, but also the ERα transcription, indicating a possible dual role for MGMT blockers in those breast cancer cells.

O6-Benzylguanine Modulates p53 Down-Stream Targeted Protein Expressions: Encouraged by the results reported, we investigated the effect of combination therapy on endogenous MGMT, p53, and ERα protein expressions. As expected, BG decreased MGMT expression, while combination therapy (4-OH-TAM or ICI combined with BG) significantly decreased both MGMT and ERα expressions. BG alone or in combination with tamoxifen or ICI decreased ER-α expression, whereas tamoxifen alone and ICI alone increased and decreased the same respectively (Fig.3A). p53 expression was slightly altered after ICI treatment. The reduction in p53 expression by ICI alone was reversed when BG was combined (Fig.3A). We investigated the effect of BG on proteins which are involved in cell cycle regulation, apoptosis in tamoxifen resistant breast cancer cells. All these treatments significantly increased the $p21^{WAF}$ protein expression (Fig.3B). PUMA expression was also increased with these treatments. Hence, PUMA may have translocated to the mitochondria, cytochrome C is released (Fig.3B), and apoptosis was triggered in these cells in presence of combination therapy. PARP cleavage is seen in BG treated cells in presence of staurosporin as an indicative of apoptosis (Fig.3C). Therefore, this data suggest that BG promotes cell cycle arrest and can induce apoptosis by modulating p53 function.

O6-Benzylguanine Modulated Transcriptional Targets in Tamoxifen Resistant Breast Cancer Cells: The effect of combination therapy on endogenous MGMT mRNA levels was our next study. Quantitative real-time PCR (qRT-PCR) resulted that anti-estrogens (TAM/ICI) increased the MGMT expression while the combination therapy decreased it compared to control levels. ERα transcription was decreased compared to controls with all these treatments (Fig.4A). Surprisingly, p21 and PUMA mRNA was significantly increased in the presence of combination treatments (Fig.4B &C). These results suggests that p53 mediated target gene transcription was affected by the drug combinations in breast cancer cells (Fig. 3 & 4).

O6-Benzylguanine Enhances p21 Transcriptional Activity in Tamoxifen Resistant Breast Cancer Cells: In order to investigate the effect of BG on p53 function, we performed luciferase reporter assays. Tamoxifen resistant MCF-7 breast cancer cells were transfected with p21 luciferase construct in presence or absence of BG (target gene of p53). These results clearly demonstrate that BG significantly enhanced p21 transcriptional activity by 4-5 fold in these cells (Fig.4D).

O6-Benzylguanine Inhibits Tamoxifen Resistant Breast Cancer Cell Growth and Increase Resistant Breast Cancer Cell Sensitivity to Anti-Estrogen Therapy (TAM/ICI): Detailed necropsy revealed that all the mice had tumors in the breast. The data summarized in Table 1 show the daily BG alone or in combination with twice weekly tamoxifen/ICI significantly decreased median tumor volume and weight as compared with that seen in tamoxifen/ICI treated and control mice. The combination of BG with tamoxifen or ICI produced the greatest decrease in median tumor volume as compared with control mice (BG 99 mm³, 9.33 mm3 (TAM+BG), respectively; $p < 0.0001$; (83.09 mm³, 32.60 mm³ (ICI+BG), respectively; $p=0.0006$). Tumor weight was also significantly reduced in mice treated with combination therapy as compared with control mice (84.23 mg, 22.30 mg (TAM+BG), respectively, $p=0.0005$; (91.23 mg, 51.27 mg (ICI+BG), respectively, $p=0.0005$). (Table.1). Body weight was not changed among all treatment groups as compared with control mice. No visible liver metastases were present (enumerated with the aid of a dissecting microscope) in all treatment groups.

Histology and IHC Analysis: We next determined the in vivo effects of BG (alone or in combination) with tamoxifen/ICI. Tumors harvested from different treatment groups were processed for routine histological and IHC analysis. Tumors from mice treated with BG alone or in combination with tamoxifen/ICI exhibited a significant decrease in MGMT, ERα, ki-67 as compared with tumors treated with tamoxifen/ICI alone or control group. p53 expression was not much altered in these treatment groups. In sharp contrast, the expression of p21 was significantly increased in tumors from mice treated with BG either alone or in combination with tamoxifen/ICI. The images were analyzed by ImageJ (NIH) and MGMT, ERα, p53, p21 and ki-67 expressions were quantified by the ImmunoRatio plugin. (Fig.5).

Conclusions

1. In the present study, we observed that prolonged treatment with anti-estrogens causes drug resistance by inducing the DNA repair protein O6-methylguanine DNA methyltransferase (MGMT).
2. Decreasing the expression of MGMT by exposing breast cancer cells to BG sensitized these cells to anti-estrogen therapy (tamoxifen and ICI 182,780).
3. We also observed that combination therapy of anti-estrogens and MGMT blockers not only overcame the MGMT-derived drug (tamoxifen and ICI) resistance but also increased the efficacy of anti-estrogen therapy by decreasing estrogen receptor expression and restoration of the functional activity of p53 in tamoxifen-resistant breast cancer cells.
4. Combination therapy inhibited tamoxifen resistant breast tumor growth in vivo.

Acknowledgements

We would like to thank the Florida Department of Health, Bankhead-Coley Cancer Research Program 09B01-01 for their funding of this project.

This is one of the hardest things...

How to use colours in data visualisations

Sara Lind, Mattias Westerlund – KTH Royal Institute of Technology

Conclusion

Intro text. Ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique.

Method

Body text. Ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga.

Data analysis

Body text. Ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga.

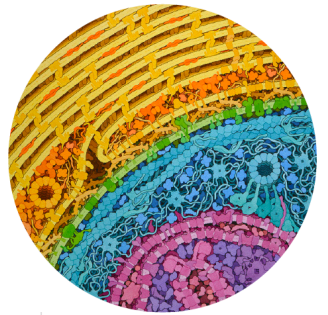


Figure 1. David S. Goodsell, RCSB Protein Data Bank. doi: 10.2210/rcsb_pdb/goodsell-gallery-030

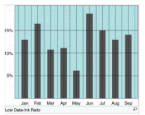


Table 1.

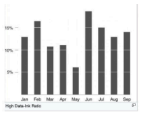


Table 2

References

Body text. Ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga.

Read publication

Sara Lind,
Communications officer
saral6@kth.se

Mattias Vesterlund,
Research data coordinator
matves29@kth.se

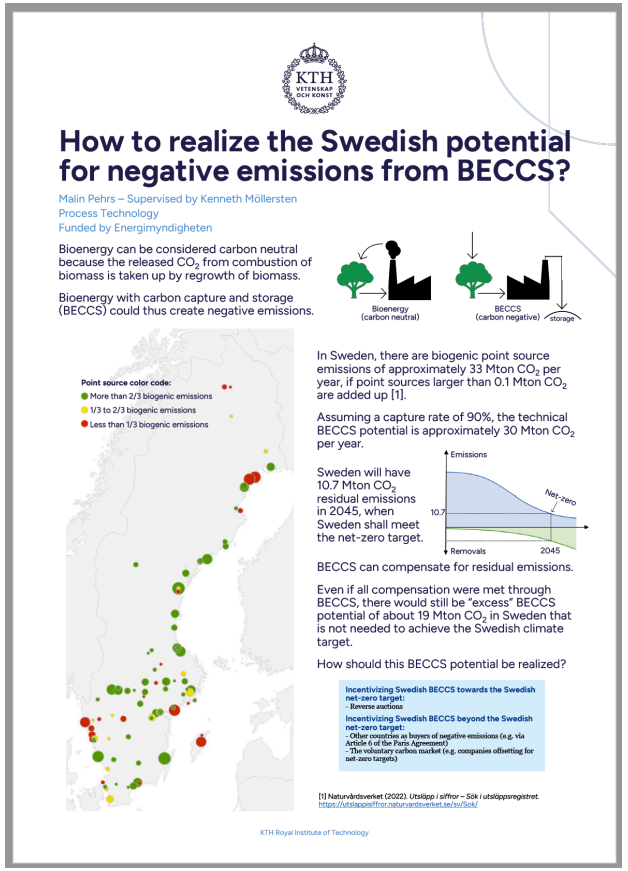
KTH Royal Institute of Technology

Posters

Design to help the reader

- Make it obvious the order to read
- Emphasize what you want to say
- The poster is a conversation starter, nothing else!
- The visuals and graphs are key elements
- Make it obvious what your research is about, even at long distance

Look at good examples and templates



How to realize the Swedish potential for negative emissions from BECCS?

Malin Pehrs – Supervised by Kenneth Möllersten
Process Technology
Funded by Energimyndigheten

Bioenergy can be considered carbon neutral because the released CO_2 from combustion of biomass is taken up by regrowth of biomass.

Bioenergy with carbon capture and storage (BECCS) could thus create negative emissions.

Bioenergy (carbon neutral) BECCS (carbon negative) storage

Point source color code:
- More than 2/3 biogenic emissions
- 1/3 to 2/3 biogenic emissions
- Less than 1/3 biogenic emissions

In Sweden, there are biogenic point source emissions of approximately 33 Mton CO_2 per year, if point sources larger than 0.1 Mton CO_2 are added up [1].

Assuming a capture rate of 90%, the technical BECCS potential is approximately 30 Mton CO_2 per year.

Sweden will have 10.7 Mton CO_2 residual emissions in 2045, when Sweden shall meet the net-zero target.

Emissions Net-zero 10.7 Removals 2045

BECCS can compensate for residual emissions.

Even if all compensation were met through BECCS, there would still be "excess" BECCS potential of about 19 Mton CO_2 in Sweden that is not needed to achieve the Swedish climate target.

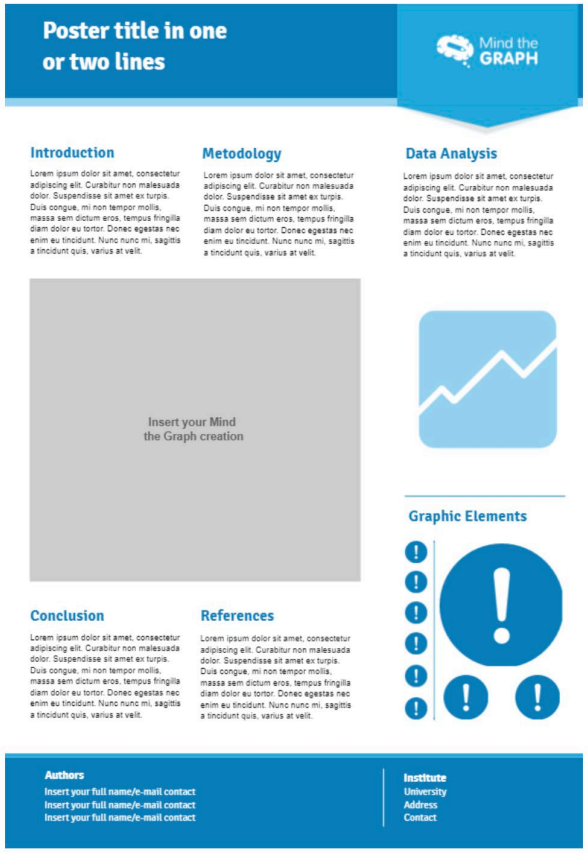
How should this BECCS potential be realized?

Incentivizing Swedish BECCS towards the Swedish net-zero target:
- Reverse auctions

Incentivizing Swedish BECCS beyond the Swedish net-zero target:
- Other countries as buyers of negative emissions (e.g. via Article 6 of the Paris Agreement)
- The voluntary carbon market (e.g. companies offsetting for net-zero targets)

[1] Naturvårdsverket (2022). *Utsläpp i siffror – Sök i utsläppsregistret.* https://utslappisiffror.naturvardsverket.se/sv/Sok/

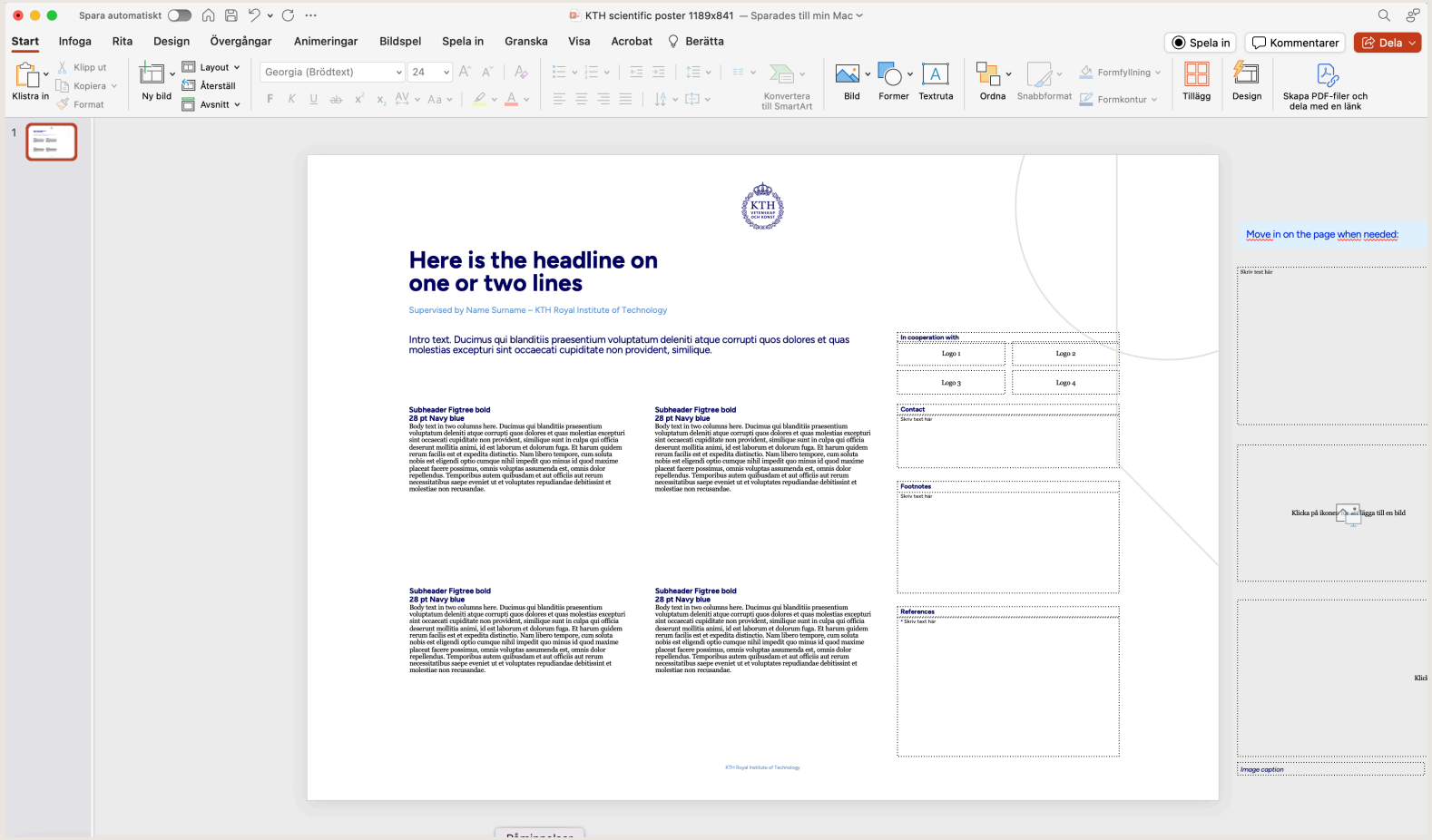
KTH Royal Institute of Technology



Examples from Mind the Graph
https://mindthegraph.com/infographic-templates/format/posters/

Use templates

- But alter to make your research clear



Tools for layout and collaboration

Free* software to use

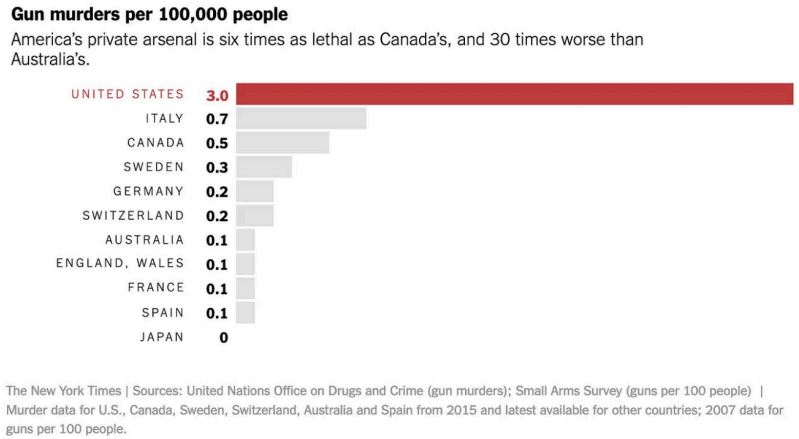
- PowerPoint (almost always included in business accounts such as universities)
- Google Docs (or Slides)
- Canva
- Adobe Express
- Quarto
- Affinity (by Canva)

*Free versions, but premium (paid) features might be even more useful

Accessible visualisation

Two important things about accessible viz

- Add Alt-texts



If I was posting this chart from the *New York Times* article "How to Reduce Shootings" on social media I would include a link to the article and write this alt text:

alt= "**Bar chart** of **gun murders per 100,000 people** where **America's murder rate is 6 times worse than Canada, and 30 times Australia**"

Amy Cesal. 2020. https://medium.com/nightingale/writing-alt-text-for-data-visualization-2a218ef43f81

And contrast

In KTH Visual identity manual all allowed combinations are marked

[Contrast checker]

Grön	Mörkgrön	Grön	Ljusgrön	Sand	Vit	Svart
Mörkgrön	1:1	3,2:1**	7,9:1*	8,3:*	10,4:1*	2:1
Grön	3,2:1**	1:1	2,5:1	2,6:1	3,2:1**	6,5:1*
Ljusgrön	7,9:1*	2,5:1	1:1	1:1	1,3:1	16:1*
Sand	8,3:*	2,6:1	1:1	1:1	1,2:1	16,8:1*
Vit	10,4:1*	3,2:1**	1,3:1	1,2:1	1:1	21,5:1*
Svart	2:1	6,5:1*	16:1*	16,8:1*	21,5:1*	1:1

Turkos	Mörk turkos	Turkos	Ljus turkos	Sand	Vit	Svart
Mörk turkos	1:1	3,3:1**	7,5:1*	8,6:1*	10,7:1*	2:1
Turkos	3,3:1**	1:1	2,3:1	2,6:1	3,3:1**	6,4:1*
Ljus turkos	7,5:1*	2,3:1	1:1	1,1:1	1,4:1	14,6:1*
Sand	8,6:1*	2,6:1	1,1:1	1:1	1,2:1	16,8:1*
Vit	10,7:1*	3,3:1**	1,4:1	1,2:1	1:1	21,5:1*
Svart	2:1	6,4:1*	14,6:1*	16,8:1*	21,5:1*	1:1

Tegel	Mörk tegel	Tegel	Ljus tegel	Sand	Vit	Svart
Mörk tegel	1:1	3,7:1*	8,1:1*	9,3:1*	11,6:1*	1,8:1
Tegel	3,7:1*	1:1	2,2:1	2,5:1	3,2:1*	6,6:1*
Ljus tegel	8,1:1*	2,2:1	1:1	1,1:1	1,4:1	14,7:1*
Sand	9,3:1*	2,5:1	1,1:1	1:1	1,2:1	16,8:1*
Vit	11,6:1*	3,2:1*	1,4:1	1,2:1	1:1	21,5:1*
Svart	1,8:1	6,6:1*	14,7:1*	16,8:1*	21,5:1*	1:1

Gul	Mörkgul	Gul	Ljusgul	Sand	Vit	Svart
Mörkgul	1:1	3,1:1**	4,5:1*	4,2:1**	5,2:1*	4,1:1**
Gul	3,1:1**	1:1	1,5:1	1,3:1	1,7:1	12,6:1*
Ljusgul	4,5:1*	1,5:1	1:1	1,1:1	1,1:1	18,3:1*
Sand	4,2:1**	1,3:1	1,1:1	1:1	1,2:1	16,8:1*
Vit	5,2:1*	1,7:1	1,1:1	1,2:1	1:1	21,5:1*
Svart	4,1:1**	12,6:1*	18,3:1*	16,8:1*	21,5:1*	1:1

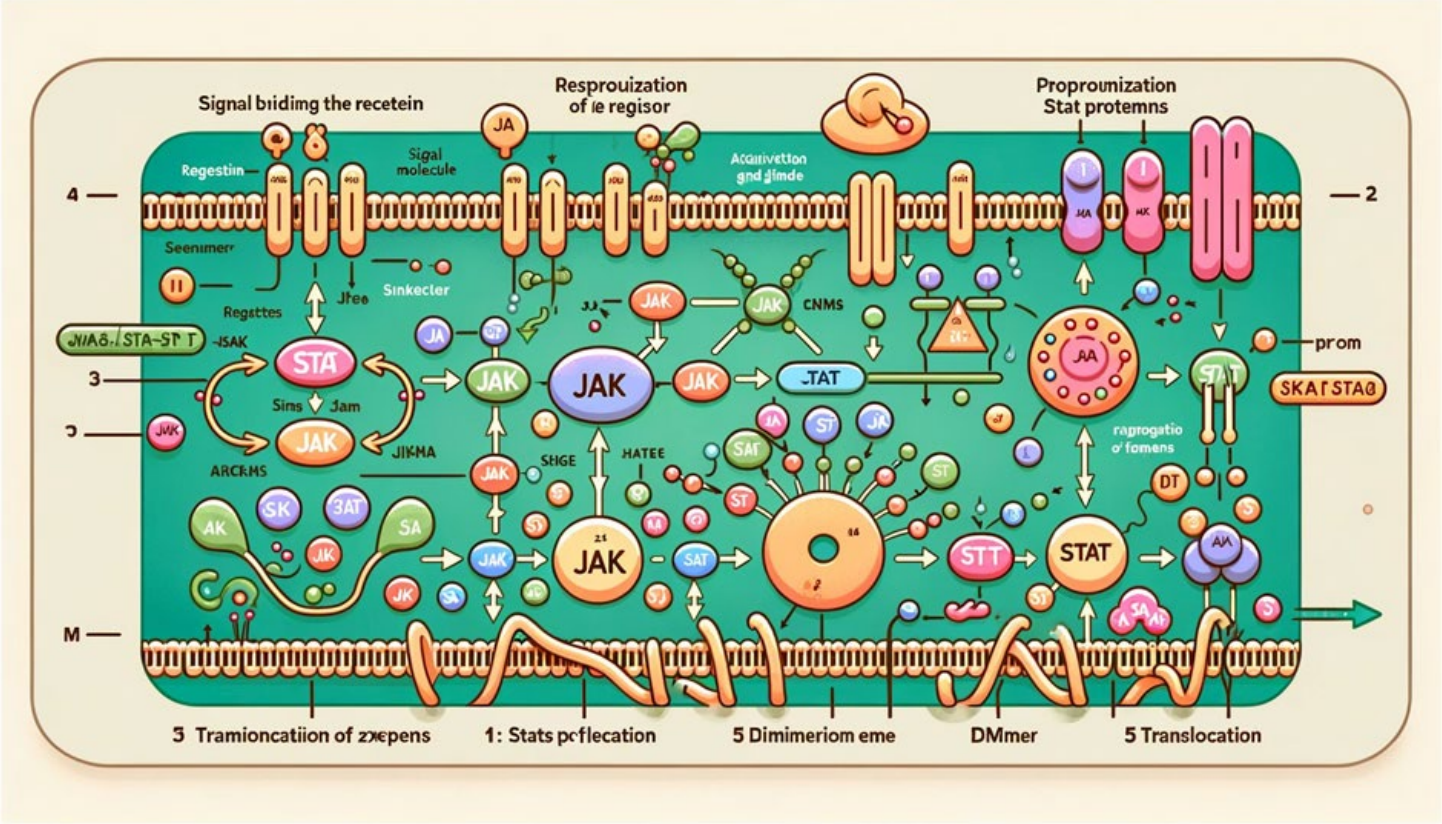
Grå	Mörkgrå	Grå	Ljusgrå	Sand	Vit	Svart
Mörkgrå	1:1	5,2:1*	10,3:1*	10,3:1*	12,8:1*	1,6:1
Grå	5,2:1*	1:1	2:1	2:1	2,5:1	8,5:1*
Ljusgrå	10,3:1*	2:1	1:1	1:1	1,2:1	16,8:1*
Sand	10,3:1*	2:1	1:1	1:1	1,2:1	16,8:1*
Vit	12,8:1*	2,5:1	1,2:1	1,2:1	1:1	21,5:1*
Svart	1,6:1	8,5:1*	16,8:1*	16,8:1*	21,5:1*	1:1

Key take-aways

Think about where you put your research data

- Using a webservice to analyse your data usually includes uploading the data to the service.
- This is essentially sending your data to a separate server, likely in a different country.
- If you are working with sensitive data, or data with personal information this can be deeply problematic.
- If you are unsure: Ask for advice from the relevant support function at your research institute.
  For KTH: researchdata@kth.se
  For Aalto: researchdata@aalto.fi

AI can be helpful... and deceitful



- This illustration was published and peer-reviewed (now retracted)
- It contains multiple errors made by AI
- This was one of many graphs/illustrations with errors

Blog post by Elisabeth Bik, 2024. Source
From: Guo X, Dong L and Hao D (2024) Cellular functions of spermatogonial stem cells in relation to JAK/STAT signaling pathway. *Front. Cell Dev. Biol.* 11:1339390. doi: 10.3389/fcell.2023.1339390 (RETRACTED) CC-BY

Key take-aways

- Help the viewer by directing attention
- Use consistent coloring and mind the color blind
- Reduce complexity if possible - Less is usually more
- Avoid circular representations of data
- Design to get attention
- Be creative about reaching out

Tools and Resources

Links to useful resources

- DataVizcatalogue.com
  Contains links to many additional resources and galleries
- Data-to-Viz
  Useful decision trees for selecting graphics
- Ktheme for R and kth-quarto
  Unofficial KTH colorschemes and layouts for R and quarto
- Visual storytelling psychology by Nadieh Bremer
  Educational slides on the psychology behind visualisation
- InfraVis National Infrastructure for DataViz
  Provides support to researchers (a lot of it free)
- Affinity
  From Canva, nowadays free for everyone.
- Adobe Illustrator
  Academic Licenses available through your university typically.
- Adobe Express or Canva
  Free version available. Great for working with layout.
- Python graph gallery
  Contains lots of examples on how to make plots in python.
- R Graph gallery
  Contains lots of examples on how to make plots in R.
- Plotly
  Making graphs interactive. For R, python, Matlab and more.
- Shiny
  Web framework for making interactive apps. R and python.
- Datawrapper color guide
  What to consider when choosing colors for data visualisation
- Colour tools:
  coolors.co: Fast palette generator.
  palett.es: Generate a palette from one color
  Color-blindness tool: Check your palettes.
  Contrast checker: Check your contrast
- Tikz for LaTeX
  Create vector-graphics in LaTeX.

Links to useful resources

- Icons – for infographics and graphical abstracts:
  bioicons: Life science
  Scidraw: Life science
  bioart (NIH): Health and Life science
  Wikimedia Commons: Everything CC-BY
- D3.js graph gallery
  Contains lots of examples on how to make plots in D3.js
- OriginLab
  GUI based with plugin apps for data analysis and visualisation. Popular in engineering.
- Inkscape
  Free alternative to Illustrator and Affinity Designer.

Agenda

Monday 9th of February

10.00	FAIR data and how to check fairness of your data
11.00	An introduction to Data Management Plans
12.00	Supporting good research management practices with Electronic Lab Notebooks
14.00	Where will my data end up? - Selecting a high quality repository for the data underlying your published results

Tuesday 10th of February

10.00	Automated large-scale text analysis
11.00	Visualize your research data
13.00	BRIGHT Data Catalog - Developing A Research Data Management Infrastructure
13.30	Introduction to Data Publishing and Zenodo
14.00	Where's the data to your article? Learn how to prepare a Data Availability Statement

Wednesday 11th of February

10.00	Searching for research data
11.00	FAIR by Default: Building Reproducible Digital Research Projects
12.00	Introduction to OpenRefine
12.30	Licensing research data for reuse
13.00	Responsible Preservation of Research Data & Data Available on Unreasonable Request

Full schedule at: www.nordiclovedataweek.org