Effective Use of Figures in Research Papers

Marinka Zitnik

marinka@cs.stanford.edu

Today's Lecture

1) Why figures matter

2) Figures in science

3) How to design effective figures

4) Tools, tips, and guidelines

Disclaimer: The suggestions and remarks in this presentation are based on personal research experience. Research practices and approaches vary. Exercise your own judgment regarding the suitability of the content.

Today's Lecture

1) Why figures matter

2) Figures in science

3) How to design effective figures

4) Tools, tips, and guidelines

Why do Figures Matter?

- Figures are often the first part of research papers examined by editors and your peers

- Informative and well-designed figures:
  - Convey facts, ideas, and relationships far more clearly and concisely than text
  - Provide a means for discovering/quantifying patterns, trends, and comparisons
  - Help the audience better understand the objective and results of your research

Design once, reuse many times: Reuse figures from papers for posters, talks, proposals, etc.

RESEARCH ARTICLE SUMMARY

YEAST GENETICS

A global genetic interaction network maps a wiring diagram of cellular function

Michael Costanzo,* Benjamin VanderSluis,* Elizabeth N. Koch,* Anastasia Baryshnikova,* Carles Pons,* Guihong Tan,* Wen Wang, Matej Usaj, Julia Hanchard, Susan D. Lee, Vicent Pelechano, Erin B. Styles, Maximilian Billmann, Jolanda van Leeuwen, Nydia van Dyk, Zhen-Yuan Lin, Elena Kuzmin, Justin Nelson, Jeff S. Piotrowski, Tharan Srikumar, Sondra Bahr, Yiqun Chen, Raamesh Deshpande, Christoph F. Kurat, Sheena C. Li, Zhijian Li, Mojca Mattiazzi Usaj, Hiroki Okada, Natasha Pascoe, Bryan-Joseph San Luis, Sara Sharifpoor, Emira Shuteriqi, Scott W. Simpkins, Jamie Snider, Harsha Garadi Suresh, Yizhao Tan, Hongwei Zhu, Noel Malod-Dognin, Vuk Janjic, Natasa Przulj, Olga G. Troyanskaya, Igor Stagljar, Tian Xia, Yoshikazu Ohya, Anne-Claude Gingras, Brian Raught, Michael Boutros, Lars M. Steinmetz, Claire L. Moore, Adam P. Rosebrock, Amy A. Caudy, Chad L. Myers,† Brenda Andrews,† Charles Boone†

INTRODUCTION: Genetic interactions occur when mutations in two or more genes combine to generate an unexpected phenotype. An extreme negative or synthetic lethal genetic interaction occurs when two mutations, neither lethal individually, combine to cause cell death. Conversely, positive genetic interactions occur when two mutations produce a phenotype that is less severe than expected. Genetic interactions identify functional relationships between genes and can be harnessed for biological discovery and therapeutic target identification. They may also explain a considerable component of the undiscovered genetics associated with human diseases. Here, we describe construction and analysis of a comprehensive genetic interaction network for a eukaryotic cell.

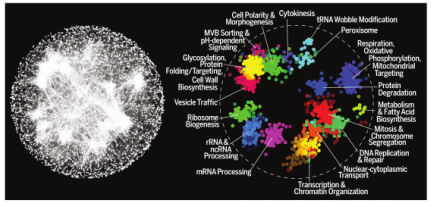
RATIONALE: Genome sequencing projects are providing an unprecedented view of genetic variation. However, our ability to interpret genetic information to predict inherited phenotypes remains limited, in large part due to the extensive buffering of genomes, making most individual eukaryotic genes dispensable for life. To explore the extent to which genetic interactions reveal cellular function and contribute to complex phenotypes, and to discover the general principles of genetic networks, we used automated yeast genetics to construct a global genetic interaction network.

RESULTS: We tested most of the ~6000 genes in the yeast *Saccharomyces cerevisiae* for all possible pairwise genetic interactions, identifying nearly 1 million interactions, including ~550,000 negative and ~350,000 positive interactions, spanning ~90% of all yeast genes. Essential genes were network hubs, displaying five times as many interactions as nonessential genes. The set of genetic interactions or the genetic interaction profile for a gene provides a quantitative measure of function, and a global network based on genetic interaction profile similarity revealed a hierarchy of modules reflecting the functional architecture of a cell. Negative interactions connected functionally related genes, mapped core bioprocesses, and identified pleiotropic genes, whereas positive interactions often mapped general regulatory connections associated with defects in cell cycle progression or cellular proteostasis. Importantly, the global network illustrates how coherent sets of negative or positive genetic interactions connect protein complex and pathways to map a functional wiring diagram of the cell.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aaf1420

CONCLUSION: A global genetic interaction network highlights the functional organization of a cell and provides a resource for predicting gene and pathway function. This network emphasizes the prevalence of genetic interactions and their potential to compound phenotypes associated with single mutations. Negative genetic interactions tend to connect functionally related genes and thus may be predicted using alternative functional information. Although less functionally informative, positive interactions may provide insights into general mechanisms of genetic suppression or resiliency. We anticipate that the ordered topology of the global genetic network, in which genetic interactions connect coherently within and between protein complexes and pathways, may be exploited to decipher genotype-to-phenotype relationships. ■

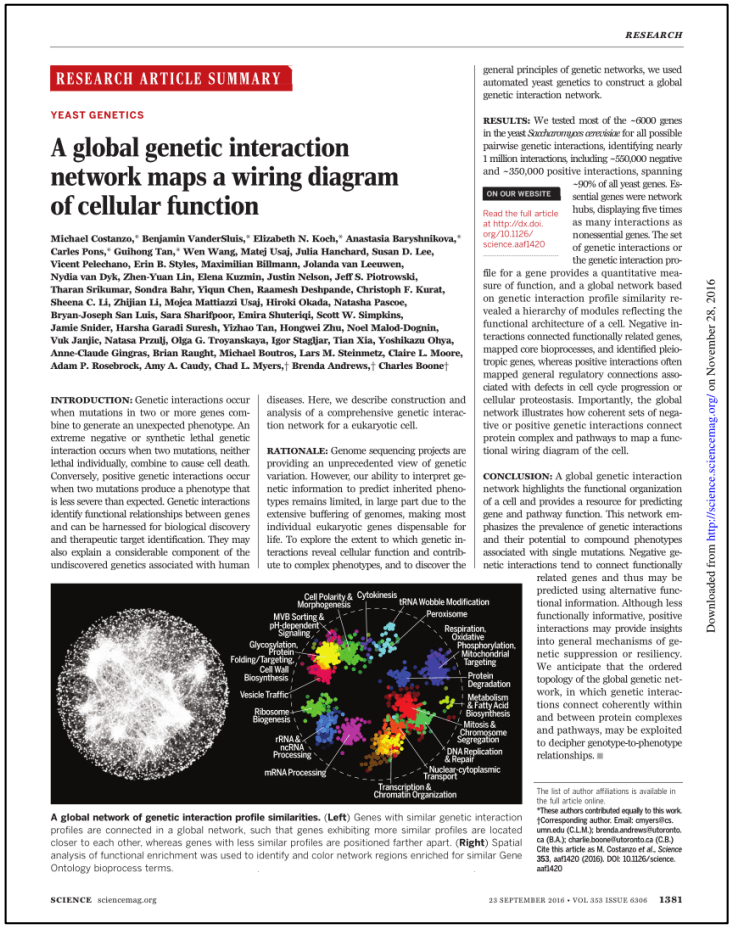


A global network of genetic interaction profile similarities. (**Left**) Genes with similar genetic interaction profiles are connected in a global network, such that genes exhibiting more similar profiles are located closer to each other, whereas genes with less similar profiles are positioned farther apart. (**Right**) Spatial analysis of functional enrichment was used to identify and color network regions enriched for similar Gene Ontology bioprocess terms.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: cmyers@cs.umn.edu (C.L.M.); brenda.andrews@utoronto.ca (B.A.); charlie.boone@utoronto.ca (C.B.)
Cite this article as M. Costanzo *et al.*, *Science* 353, aaf1420 (2016). DOI: 10.1126/science.aaf1420

Figure taken from: Costanzo et al. Science 353.6306 (2016).

Promote research ideas and make them accessible to other scientists



RESEARCH

RESEARCH ARTICLE SUMMARY

YEAST GENETICS

A global genetic interaction network maps a wiring diagram of cellular function

Michael Costanzo,* Benjamin VanderSluis,* Elizabeth N. Koch,* Anastasia Baryshnikova,* Carles Pons,* Guihong Tan,* Wen Wang, Matej Usaj, Julia Hanchard, Susan D. Lee, Vicent Pelechano, Erin B. Styles, Maximilian Billmann, Jolanda van Leeuwen, Nydia van Dyk, Zhen-Yuan Lin, Elena Kuzmin, Justin Nelson, Jeff S. Piotrowski, Tharan Srikumar, Sondra Bahr, Yiqun Chen, Raamesh Deshpande, Christoph F. Kurat, Sheena C. Li, Zhijian Li, Mojca Mattiazzi Usaj, Hiroki Okada, Natasha Pascoe, Bryan-Joseph San Luis, Sara Sharifpoor, Emira Shuteriqi, Scott W. Simpkins, Jamie Snider, Harsha Garadi Suresh, Yizhao Tan, Hongwei Zhu, Noel Malod-Dognin, Vuk Janjic, Natasa Przulj, Olga G. Troyanskaya, Igor Stagljar, Tian Xia, Yoshikazu Ohya, Anne-Claude Gingras, Brian Raught, Michael Boutros, Lars M. Steinmetz, Claire L. Moore, Adam P. Rosebrock, Amy A. Candy, Chad L. Myers,† Brenda Andrews,† Charles Boone†

INTRODUCTION: Genetic interactions occur when mutations in two or more genes combine to generate an unexpected phenotype. An extreme negative or synthetic lethal genetic interaction occurs when two mutations, neither lethal individually, combine to cause cell death. Conversely, positive genetic interactions occur when two mutations produce a phenotype that is less severe than expected. Genetic interactions identify functional relationships between genes and can be harnessed for biological discovery and therapeutic target identification. They may also explain a considerable component of the undiscovered genetics associated with human diseases. Here, we describe construction and analysis of a comprehensive genetic interaction network for a eukaryotic cell.

RATIONALE: Genome sequencing projects are providing an unprecedented view of genetic variation. However, our ability to interpret genetic information to predict inherited phenotypes remains limited, in large part due to the extensive buffering of genomes, making most individual eukaryotic genes dispensable for life. To explore the extent to which genetic interactions reveal cellular function and contribute to complex phenotypes, and to discover the general principles of genetic networks, we used automated yeast genetics to construct a global genetic interaction network.

RESULTS: We tested most of the ~6000 genes in the yeast *Saccharomyces cerevisiae* for all possible pairwise genetic interactions, identifying nearly 1 million interactions, including ~550,000 negative and ~350,000 positive interactions, spanning ~90% of all yeast genes. Essential genes were network hubs, displaying five times as many interactions as nonessential genes. The set of genetic interactions or the genetic interaction profile for a gene provides a quantitative measure of function, and a global network based on genetic interaction profile similarity revealed a hierarchy of modules reflecting the functional architecture of a cell. Negative interactions connected functionally related genes, mapped core bioprocesses, and identified pleiotropic genes, whereas positive interactions often mapped general regulatory connections associated with defects in cell cycle progression or cellular proteostasis. Importantly, the global network illustrates how coherent sets of negative or positive genetic interactions connect protein complex and pathways to map a functional wiring diagram of the cell.

CONCLUSION: A global genetic interaction network highlights the functional organization of a cell and provides a resource for predicting gene and pathway function. This network emphasizes the prevalence of genetic interactions and their potential to compound phenotypes associated with single mutations. Negative genetic interactions tend to connect functionally related genes and thus may be predicted using alternative functional information. Although less functionally informative, positive interactions may provide insights into general mechanisms of genetic suppression or resiliency. We anticipate that the ordered topology of the global genetic network, in which genetic interactions connect coherently within and between protein complexes and pathways, may be exploited to decipher genotype-to-phenotype relationships. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aaf1420

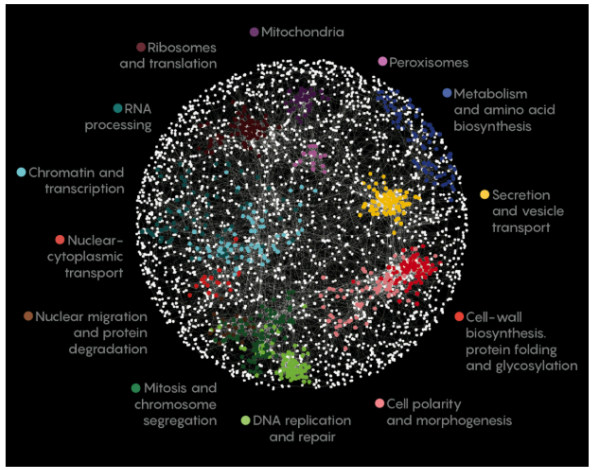
A global network of genetic interaction profile similarities. (**Left**) Genes with similar genetic interaction profiles are connected in a global network, such that genes exhibiting more similar profiles are located closer to each other, whereas genes with less similar profiles are positioned farther apart. (**Right**) Spatial analysis of functional enrichment was used to identify and color network regions enriched for similar Gene Ontology bioprocess terms.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: cmyers@cs.umn.edu (C.L.M.); brenda.andrews@utoronto.ca (B.A.); charlie.boone@utoronto.ca (C.B.)
Cite this article as M. Costanzo *et al.*, *Science* 353, aaf1420 (2016). DOI: 10.1126/science.aaf1420

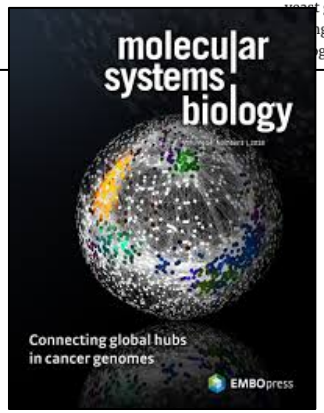
Different authors, different papers, different journals

Promote your research among general audience and media

Quanta magazine | Physics | Mathematics | Biology | Computer Science | All Articles

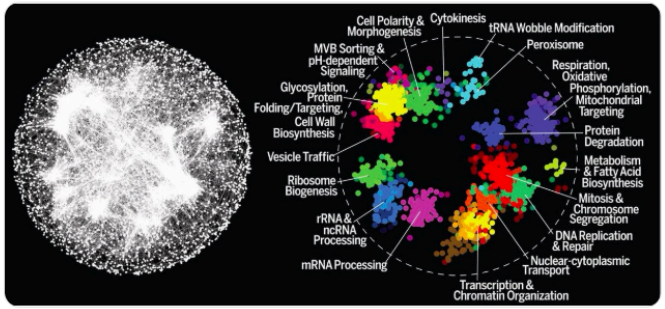


This figure maps the interactions among various genes (represented as dots) in the yeast genome. Genes with linked effects are connected by lines; genes with more highly correlated effects are closer together. The color of the dots corresponds to the biological processes and organelles in which the genes are involved.



Dryad @datadryad · 26 Dec 2016

Featured data from @sciencemagazine: **A global genetic interaction network maps a wiring diagram of cellular function** dx.doi.org/10.5061/dryad.…

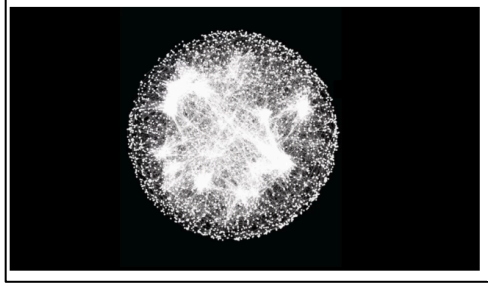


MOTHERBOARD
TECH BY VICE

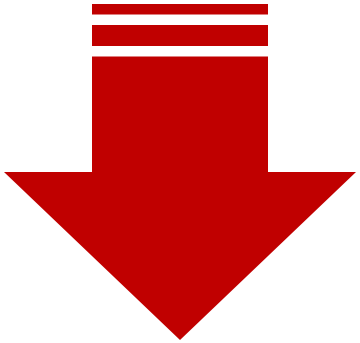
It Took 15 Years to Map Every Gene Interaction in a Yeast Cell

Understanding how thousands of individual yeast genes interact in pairs could expose the underlying genetic bases of human diseases.

SHARE | TWEET



Effective figures improve your papers



Maximize impact, boost citation count, stand out among your peers

Today's Lecture

1) Why figures matter

2) Figures in science

3) How to design effective figures

4) Tools, tips, and guidelines

Two Types of Papers with Different Visual Structure

1) Core CS conference papers:

- KDD, NeurIPS, ICML, ICLR, AAAI, etc.

2) Interdisciplinary journal papers:

- Nature, Science, PNAS, etc.

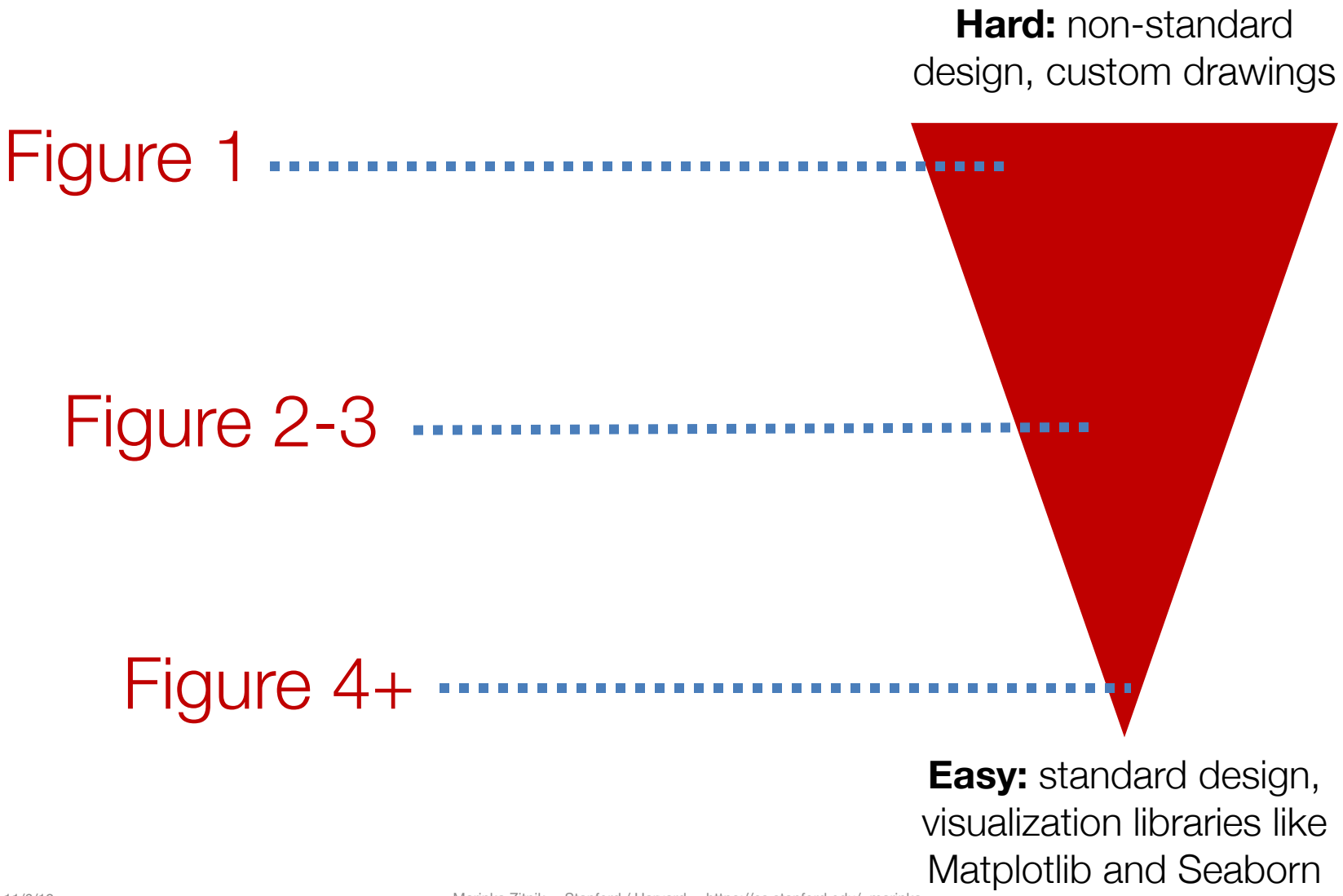
Core CS Conference Papers

The focus is on the development of new methods and their evaluation and comparison on benchmark datasets

Core CS Conference Papers: Visual Structure

- Figure 1: Key methodological contribution
  - Focus on most important information
  - **Impress your audience!**
    - Is your method/system the fastest, the largest, the most accurate?
    - What is the hard problem that your method solves?
    - What makes your method different from related work?
- Figure 2-3: Overview and algorithmic details
  - Inputs + Data transformation + Outputs
  - Show details about data transformations:
    - Graph convolutions, neural architectures, etc.
- Figure 4+: Results

Core CS Conference Papers:
Visual Structure



Examples:
Core CS Conference Papers

Abstract

Supervised learning on molecules has incredible potential to be useful in chemistry, drug discovery, and materials science. Luckily, several promising and closely related neural network models invariant to molecular symmetries have already been described in the literature. These models learn a message passing algorithm and aggregation procedure to compute a function of their entire input graph. At this point, the next step is to find a particularly effective variant of this general approach and apply it to chemical prediction benchmarks until we either solve them or reach the limits of the approach. In this paper, we reformulate existing models into a single common framework we call Message Passing Neural Networks (MPNNs) and explore additional novel variations within this framework. Using MPNNs we demonstrate state of the art results on an important molecular property prediction benchmark; these results are strong enough that we believe future work should focus on datasets with larger molecules or more accurate ground truth labels.

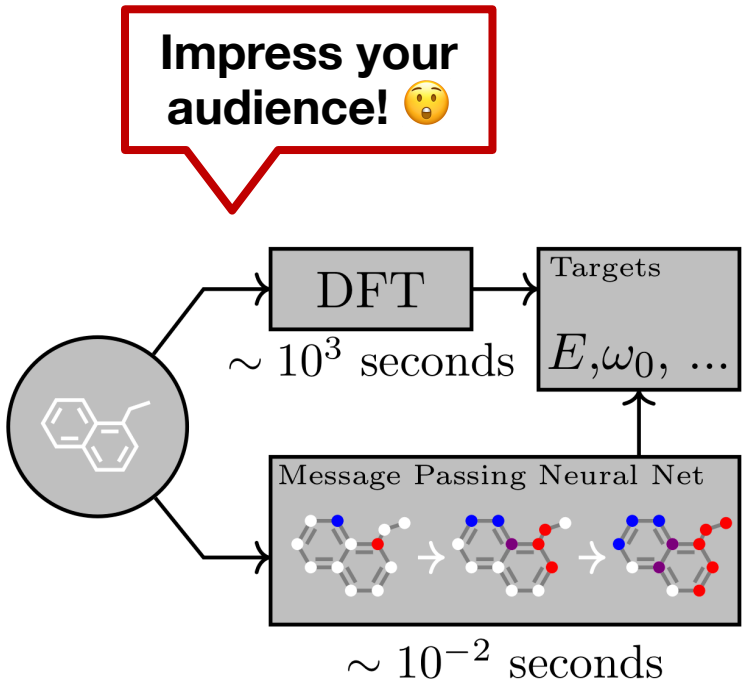


Figure 1. A Message Passing [illegible] Network predicts quantum properties of an organic [illegible] by modeling a computationally expensive [illegible]

Focus on key information: "Our method is so fast! Our paper should be published at ICML!"

Gilmer et al., Neural Message Passing for Quantum Chemistry, ICML, 2017.

Abstract

Large cascades can develop in online social networks as people share information with one another. Though simple reshare cascades have been studied extensively, the full range of cascading behaviors on social media is much more diverse. Here we study how *diffusion protocols*, or the social exchanges that enable information transmission, affect cascade growth, analogous to the way communication protocols define how information is transmitted from one point to another. Studying 98 of the largest information cascades on Facebook, we find a wide range of diffusion protocols – from cascading reshares of images, which use a simple protocol of tapping a single button for propagation, to the ALS Ice Bucket Challenge, whose diffusion protocol involved individuals creating and posting a video, and then nominating specific others to do the same. We find recurring classes of diffusion protocols, and identify two key counterbalancing factors in the construction of these protocols, with implications for a cascade's growth: the effort required to participate in the cascade, and the social cost of staying on the sidelines. Protocols requiring greater individual effort slow down a cascade's propagation, while those imposing a greater social cost of not participating increase the cascade's adoption likelihood. The predictability of transmission also varies with protocol. But regardless of mechanism, the cascades in our analysis all have a similar reproduction number ($\approx$1.8), meaning that lower rates of exposure can be offset with higher per-exposure rates of adoption. Last, we show how a cascade's structure can not only differentiate these protocols, but also be modeled through branching processes. Together, these findings provide a framework for understanding how a wide variety of information cascades can achieve substantial adoption across a network.

Impress your audience! 😲

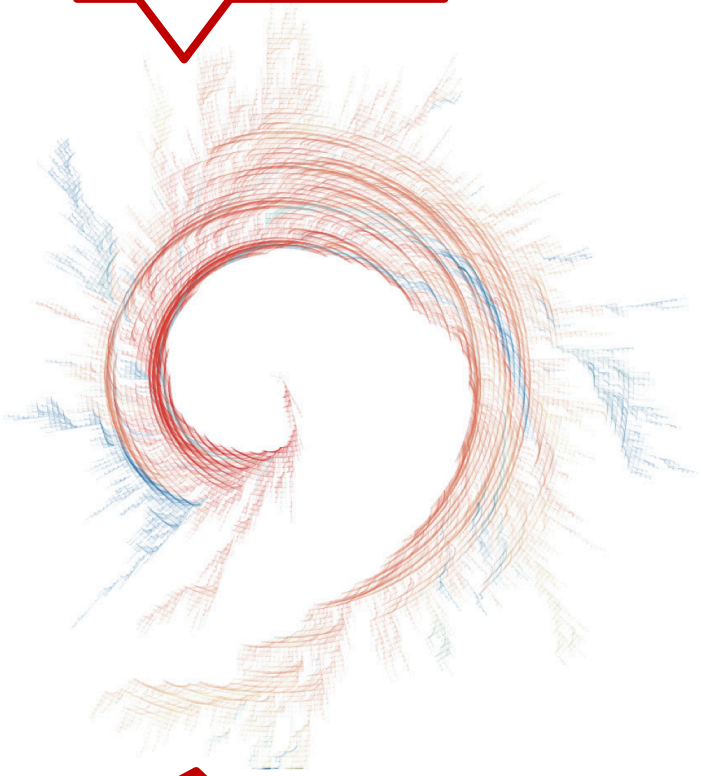


Figure 1: The d[illegible]n tree of a cascade with a volunteer diffusion [illegible] individuals posted music from an

"Cascades can be so large! Despite that, we know how to study them! Our paper should be published at ICWSM!"

Cheng et al., Do Diffusion Protocols Govern Cascade Growth?, ICWSM, 2018.

ABSTRACT

Cascades of information-sharing are a primary mechanism by which content reaches its audience on social media, and an active line of research has studied how such cascades, which form as content is reshared from person to person, develop and subside. In this paper, we perform a large-scale analysis of cascades on Facebook over significantly longer time scales, and find that a more complex picture emerges, in which many large cascades *recur*, exhibiting multiple bursts of popularity with periods of quiescence in between. We characterize recurrence by measuring the time elapsed between bursts, their overlap and proximity in the social network, and the diversity in the demographics of individuals participating in each peak. We discover that content virality, as revealed by its initial popularity, is a main driver of recurrence, with the availability of multiple copies of that content helping to spark new bursts. Still, beyond a certain popularity of content, the rate of recurrence drops as cascades start exhausting the population of interested individuals. We reproduce these observed patterns in a simple model of content recurrence simulated on a real social network. Using only characteristics of a cascade's initial burst, we demonstrate strong performance in predicting whether it will recur in the future.

Keywords: tion diffusion; m

Impress your audience! 😲

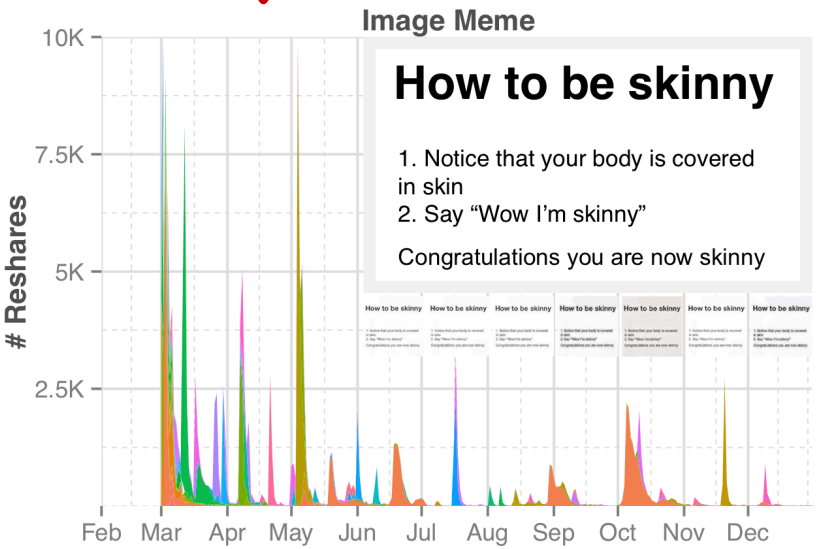


Figure 1: An example of a image meme that has recurred, or resurfaced in popularity multiple times, sometimes as a continuation of the same copy, and sometimes as a new copy of the same meme (example copies are shown as thumbnails). This recurrence appears as multiple peaks in the plot of reshares as a function of time.

Focus on most important information: Figure 1 answers question asked by the title

"Cascades can be so complex! Despite that, we know how to study them! Our paper should be published at WWW!"

Cheng et al., Do Cascades Recur?, WWW, 2016.

Abstract

Fairness in machine learning has predominantly been studied in static classification settings without concern for how decisions change the underlying population over time. Conventional wisdom suggests that fairness criteria promote the long-term well-being of those groups they aim to protect. We study how static fairness criteria interact with temporal indicators of well-being, such as long-term improvement, stagnation, and decline in a variable of interest. We demonstrate that even in a one-step feedback model, common fairness criteria in general do not promote improvement over time, and may in fact cause harm in cases where an unconstrained objective would not. We completely characterize the delayed impact of three standard criteria, contrasting the regimes in which these exhibit qualitatively different behavior. In addition, we find that a natural form of measurement error broadens the regime in which fairness criteria perform favorably. Our results highlight the importance of measurement and temporal modeling in the evaluation of fairness criteria, suggesting a range of new challenges and trade-offs.

Impress your audience! 😲

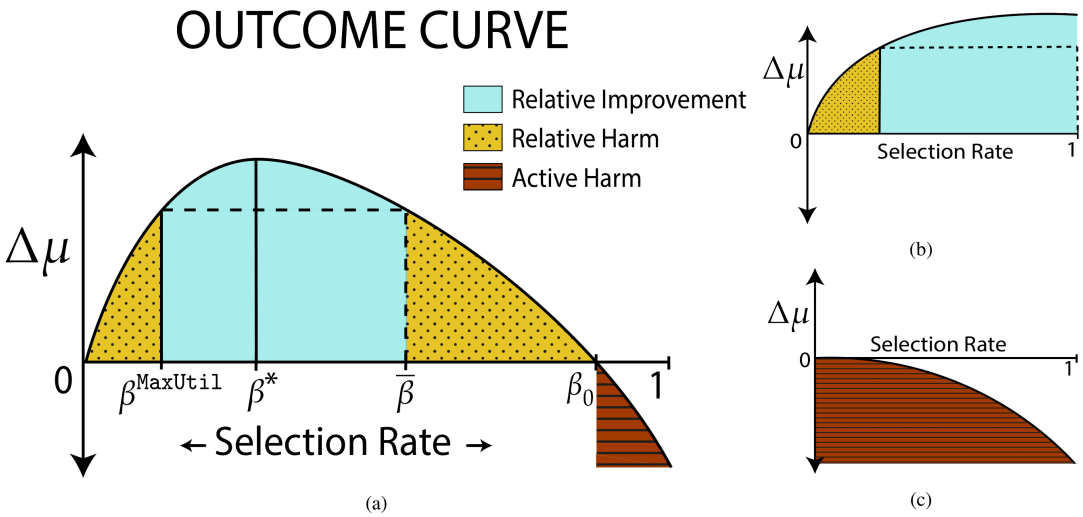


Figure 1. The above figure shows the *outcome curve*. The horizontal axis represents the selection rate for the population; the vertical axis represents the mean change in score. (a) depicts the full spectrum of outcome regimes, and colors indicate regions of active harm, relative harm, and no harm. In (b): a group that has much potential for gain, in (c): a group that has no potential for gain.

Focus on key information: Delayed impact of FML is not well-understood. Here we show a complete characterization of delayed impact.

Liu et al., Delayed Impact of Fair Machine Learning?, ICML, 2018. *(Best paper award)*

ABSTRACT

Deep learning models for graphs have achieved strong performance for the task of node classification. Despite their proliferation, currently there is no study of their robustness to adversarial attacks. Yet, in domains where they are likely to be used, e.g. the web, adversaries are common. Can deep learning models for graphs be easily fooled? In this work, we introduce the first study of adversarial attacks on attributed graphs, specifically focusing on models exploiting ideas of graph convolutions. In addition to attacks at test time, we tackle the more challenging class of poisoning/causative attacks, which focus on the training phase of a machine learning model. We generate adversarial perturbations targeting the *node's features* and the *graph structure*, thus, taking the dependencies between instances in account. Moreover, we ensure that the perturbations remain *unnoticeable* by preserving important data characteristics. To cope with the underlying discrete domain we propose an efficient algorithm NETTACK exploiting incremental computations. Our experimental study shows that accuracy of node classification significantly drops even when performing only few perturbations. Even more, our attacks are transferable: the learned attacks generalize to other state-of-the-art node classification models and unsupervised approaches, and likewise are successful even when only limited knowledge about the graph is given.

Impress your audience! 😲

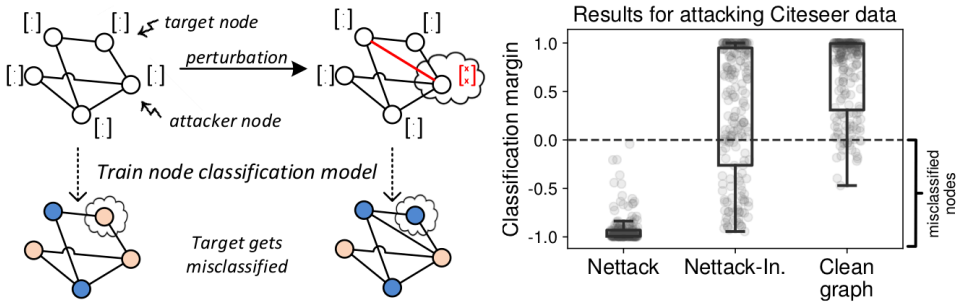


Figure 1: Small perturbations of the graph structure and node features lead to misclassification of the target.

Focus on key information: Yes, graph-based models for deep learning can be easily fooled. Here we show how devastating attacks can be.

Zugner et al., Adversarial Attacks on Neural Networks for Graph Data, KDD, 2018. *(Best paper award)*

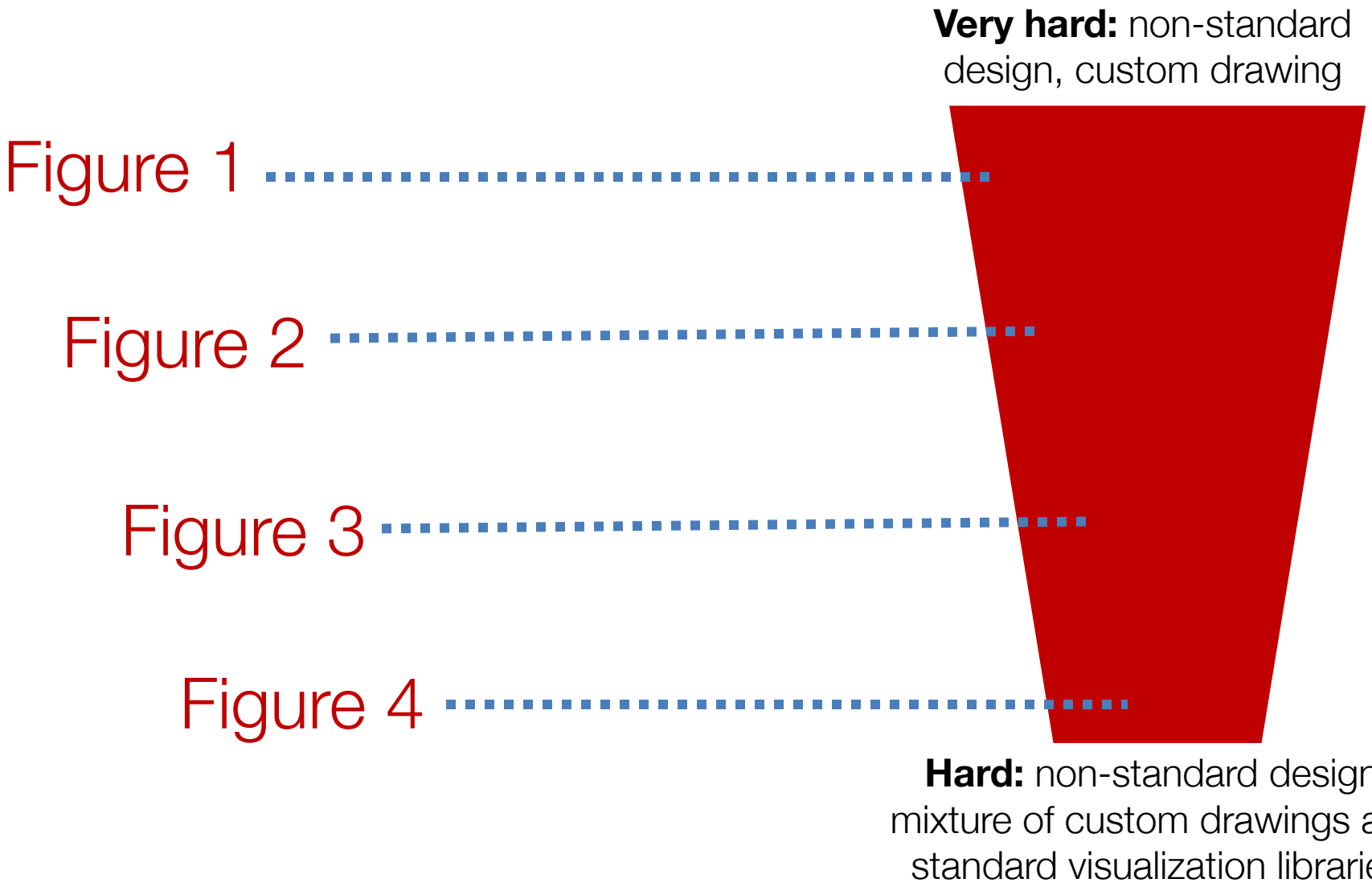
Interdisciplinary Journal Papers

The focus is on new scientific insights and demonstrating the importance of those insights to advance science

Interdisciplinary Journal Papers: Visual Structure

- Figure 1: Dataset, approach and key result
  - **Impress your audience!**
- Figure 2: Key result, detailed and unpacked
- Figure 3: Orthogonal evidence supporting results
- Figure 4: Orthogonal evidence supporting results
- Supplementary Figures: Methodological contributions, algorithms, robustness analyses

Interdisciplinary Journal Papers: Visual Structure



Examples: Interdisciplinary Journal Papers

RESEARCH ARTICLE

BIG DATA

Quantitative analysis of population-scale family tree with millions of relatives

Joanna Kaplanis,[1,2]* Assaf Gordon,[1,2]* Tal Shor,[3,4] Omer Weissbrod,[5] D iger,[4]
Mary Wahl,[1,2,6] Michael Gershovits,[2] Barak Markus,[2] Mona Sheikh,[2]
Melissa Gymrek,[1,2,7,8,9] Gaurav Bhatia,[10,11] Daniel G. MacArthur,[7,9,10]
Alkes L. Price,[10,11,12] Yaniv Erlich[1,2,3,13,14]†

Family trees have vast applications in fields as diverse as genetics, anthropology, and economics. However, the collection of extended family trees is tedious and usually relies on resources with limited geographical scope and complex data usage restrictions. We collected 86 million profiles from publicly available online data shared by genealogy enthusiasts. After extensive cleaning and validation, we obtained population-scale family trees, including a single pedigree of 13 million individuals. We leveraged the data to partition the genetic architecture of human longevity and to provide insights into the geographical dispersion of families. We also report a simple digital procedure to overlay other data sets with our resource.

Figures provide a visual story for the abstract

Kaplanis et al., Quantitative analysis of population-scale family trees with millions of relatives, *Science*, 2018.

Figure 1

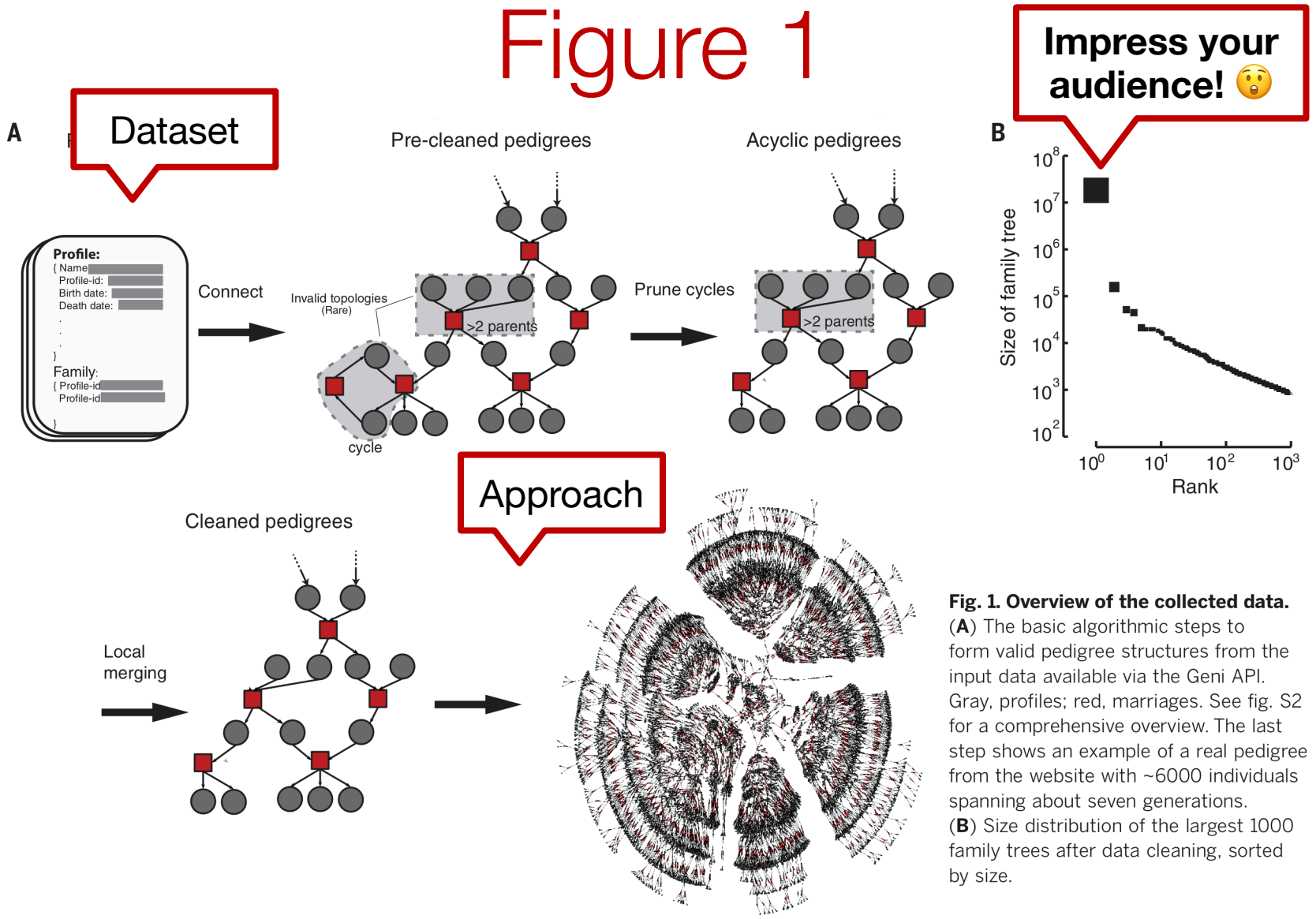


Fig. 1. Overview of the collected data. (**A**) The basic algorithmic steps to form valid pedigree structures from the input data available via the Geni API. Gray, profiles; red, marriages. See fig. S2 for a comprehensive overview. The last step shows an example of a real pedigree from the website with ~6000 individuals spanning about seven generations. (**B**) Size distribution of the largest 1000 family trees after data cleaning, sorted by size.

Kaplanis et al., Quantitative analysis of population-scale family trees with millions of relatives, *Science*, 2018.

Figure 2

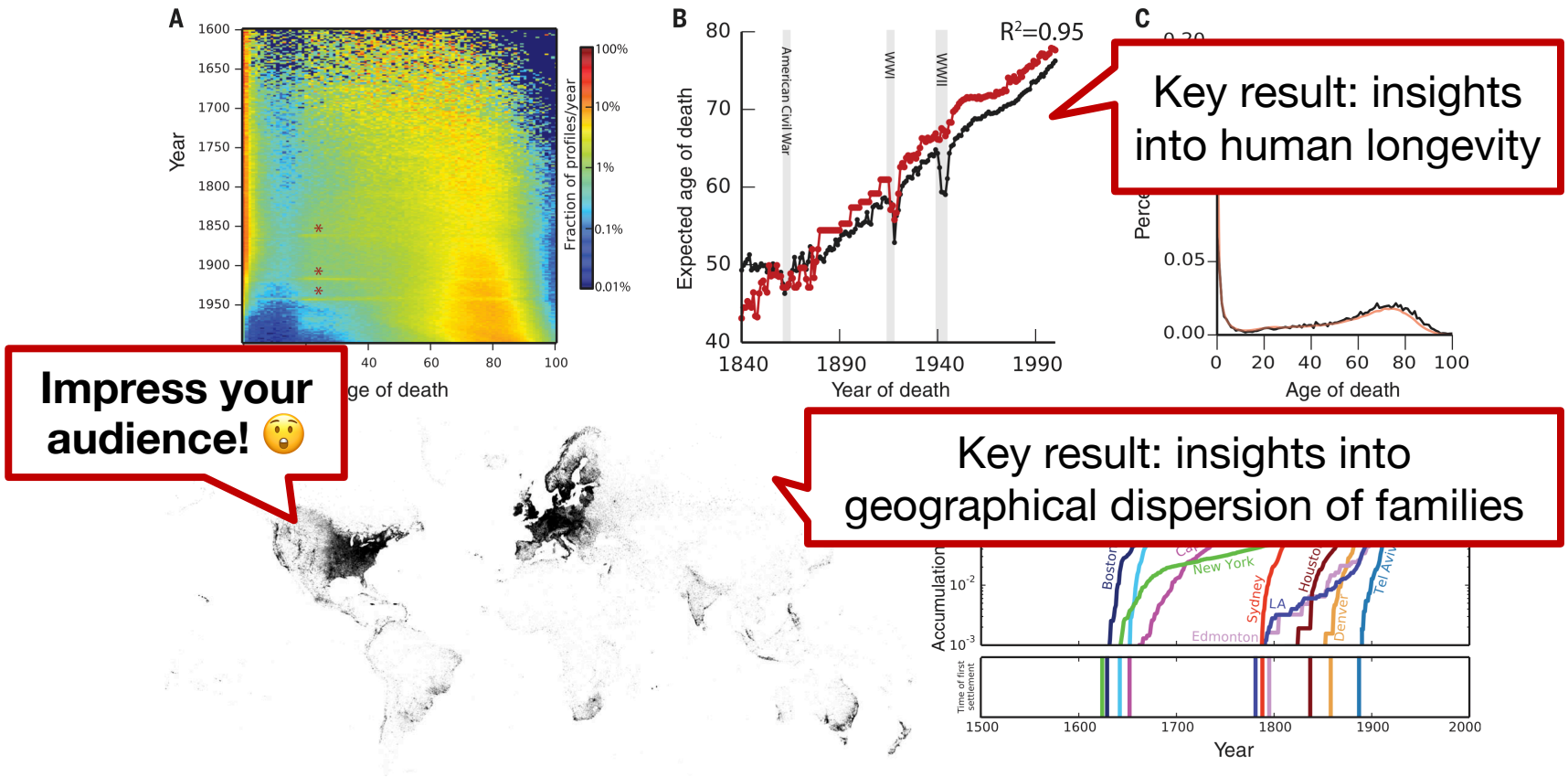


Fig. 2. Analysis and validation of demographic data. (**A**) Distribution of life expectancy per year. Colors correspond to the frequency of profiles of individuals who died at a certain age for each year. Asterisks indicate deaths at military age in the Civil War and First and Second World Wars. (**B**) Expected life span in Geni (black) and the Oeppen and Vaupel study [red (*27*)] as a function of year of death. (**C**) Comparison of the life-span distributions versus Geni (black) and HMD (red). See also fig. S5A. (**D**) Geographic distribution of the annotated place-of-birth information. Every pixel corresponds to a profile in the data set. (**E**) Validation of geographical assignment by historical trends. Top: Cumulative distribution of profiles since 1500 for each city on a logarithmic scale as a function of time. Bottom: Year of first settlement in the city.

Kaplanis et al., Quantitative analysis of population-scale family trees with millions of relatives, *Science*, 2018.

Figure 3

Further analyses supporting key result

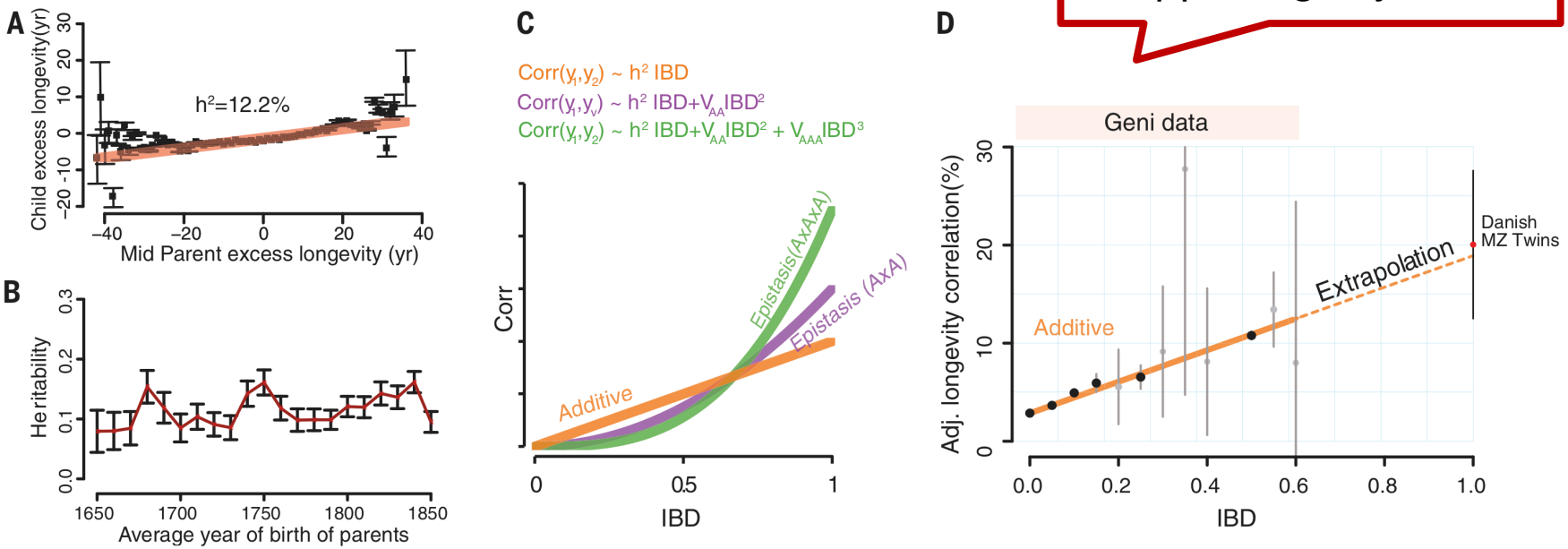
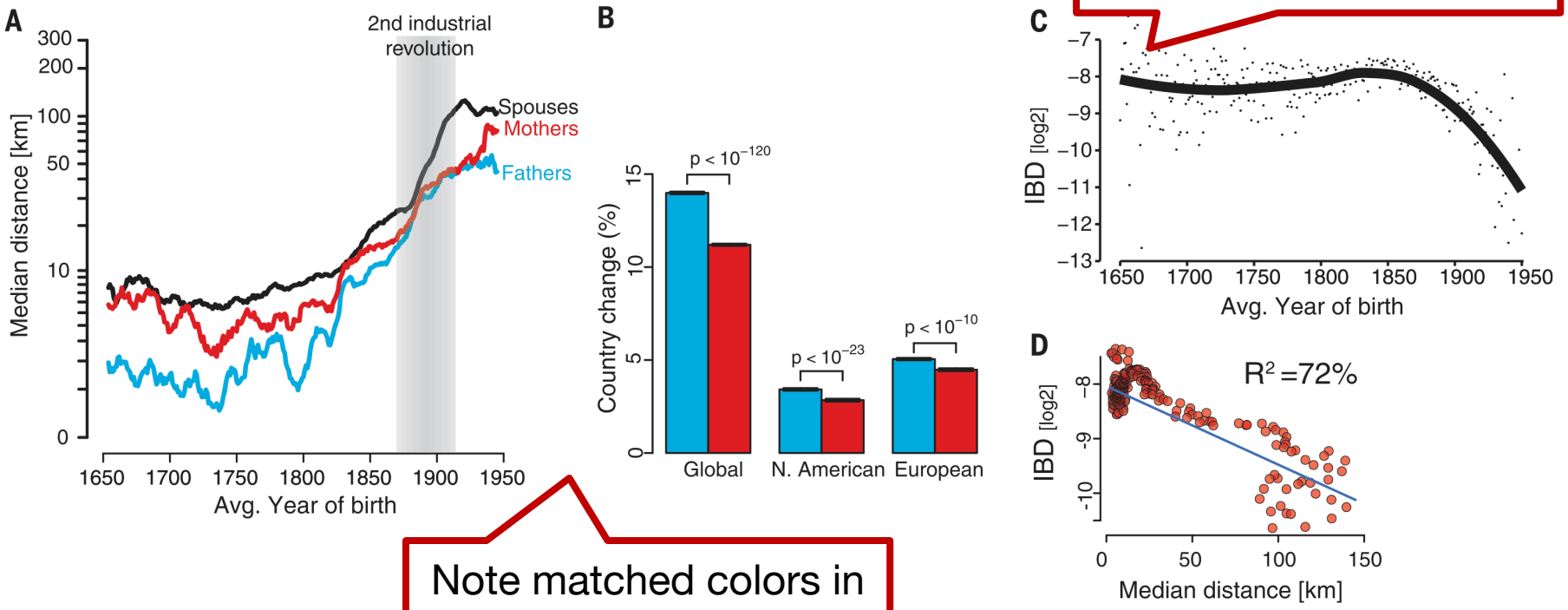


Fig. 3. The genetic architecture of longevity. (**A**) Regression (red) of child longevity on its mid-parent longevity (defined as difference between age of death and expected life span). Black squares, average longevity of children binned by the mid-parent value; gray bars, estimated 95% confidence interval (CI). (**B**) Estimated narrow-sense heritability (red) with 95% confidence intervals (black bars) obtained by the mid-parent design stratified by the average decade of birth of the parents. (**C**) Correlation of a trait as a function of IBD under strict additive (h^2, orange), squared (V_{AA}, purple), and cubic (V_{AAA}, green) epistasis architectures after dormancy adjustments. (**D**) Average longevity correlation as a function of IBD (black circles) grouped in 5% increments (gray: 95% CI) after adjusting for dominancy. A dashed line denotes the extrapolation of the models toward monozygotic twins from the Danish Twin Registry (red circle).

Kaplanis et al., Quantitative analysis of population-scale family trees with millions of relatives, *Science*, 2018.

Figure 4

Further analyses supporting key result



Note matched colors in panels A and B

Fig. 4. Analysis of familial dispersion of father-offspring places of birth (cyan), mother-offspring (red), and marital radius (black) as a function of time (average year of birth). (**B**) Rate of change in the country of birth for father-offspring (cyan) or mother-offspring (red) stratified by major geographic areas. (**C**) Average IBD ($\log_2$) between s a function of average year of birth. Individual dots represent the measured average per year; the black line denotes the smooth trend using locally weighted regression. (**D**) IBD of couples as a function of marital radius. Each dot represents a year between 1650 to 1950. The blue line denotes the best linear regression line in log-log space.

Kaplanis et al., Quantitative analysis of population-scale family trees with millions of relatives, *Science*, 2018.

COMPUTER SCIENCE

Human-level performance in 3D multiplayer games with population-based reinforcement learning

Max Jaderberg*†, Wojciech M. Czarnecki*†, Iain Dunning†, Luke Marris, Guy Lever, Antonio Garcia Castañeda, Charles Beattie, Neil C. Rabinowitz, Ari S. Morcos, Avraham Ruderman, Nicolas Sonnerat, Tim Green, Louise Deason, Joel Z. Leibo, David Silver, Demis Hassabis, Koray Kavukcuoglu, Thore Graepel

Reinforcement learning (RL) has shown great success in increasingly complex single-agent environments and two-player turn-based games. However, the real world contains multiple agents, each learning and acting independently to cooperate and compete with other agents. We used a tournament-style evaluation to demonstrate that an agent can achieve human-level performance in a three-dimensional multiplayer first-person video game, *Quake III Arena* in Capture the Flag mode, using only pixels and game points scored as input. We used a two-tier optimization process in which a population of independent RL agents are trained concurrently from thousands of parallel matches on randomly generated environments. Each agent learns its own internal reward signal and rich representation of the world. These results indicate the great potential of multiagent reinforcement learning for artificial intelligence research.

Jaderberg et al., Human-level performance in 3D multiplayer games with population-based reinforcement learning, *Science*, 2019.

Figure 1

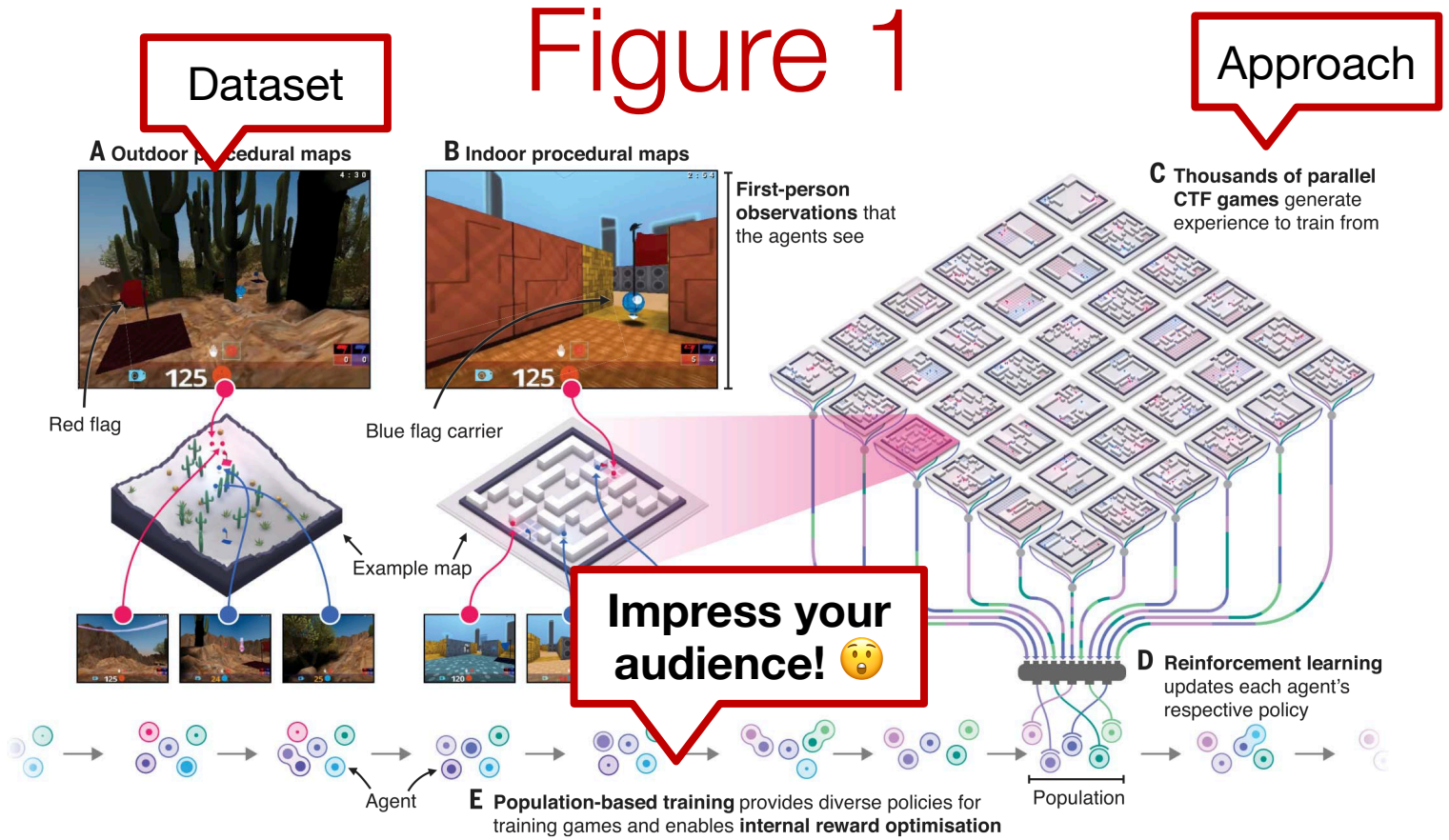


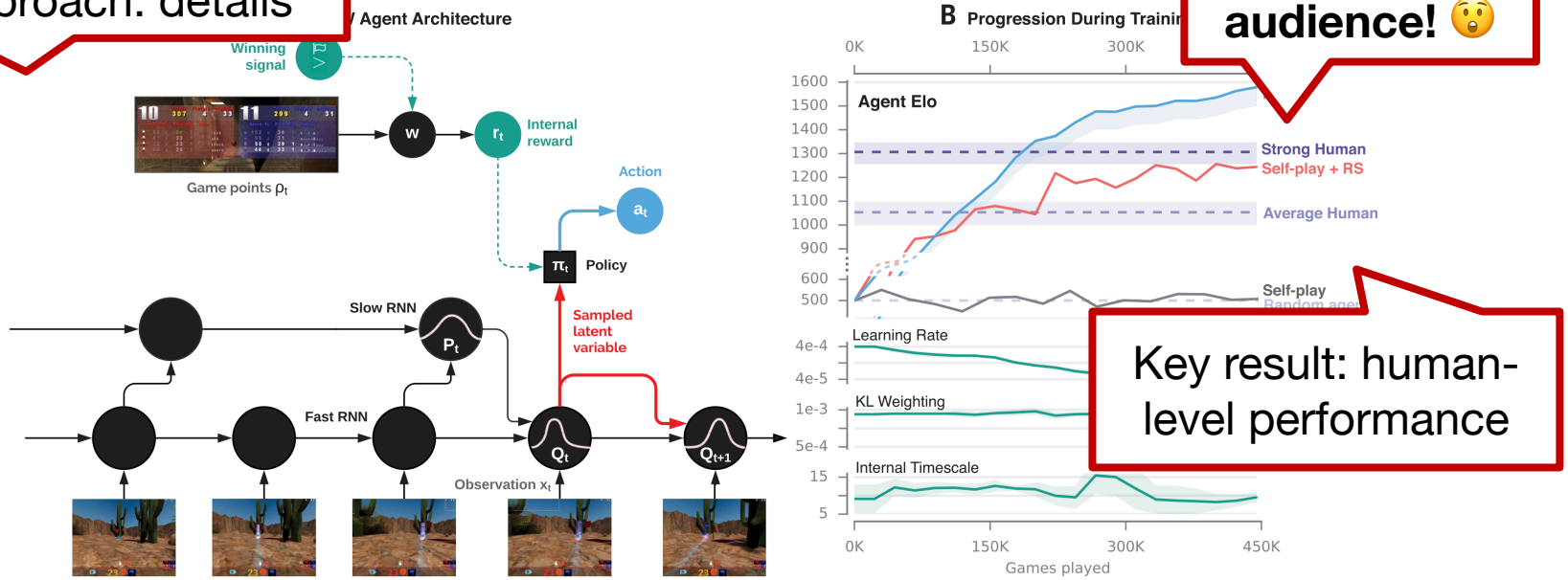
Fig. 1. CTF task and computational training framework. (**A** and **B**) Two example maps that have been sampled from the distribution of (A) outdoor maps and (B) indoor maps. Each agent in the game sees only its own first-person pixel view of the environment. (**C**) Training data are generated by playing thousands of CTF games in parallel on a diverse distribution of procedurally generated maps and (**D**) used to train the agents that played in each game with RL. (**E**) We trained a population of 30 different agents together, which provided a diverse set of teammates and opponents to play with and was also used to evolve the internal rewards and hyperparameters of agents and learning process. Each circle represents an agent in the population, with the size of the inner circle representing strength. Agents undergo computational evolution (represented as splitting) with descendents inheriting and mutating hyperparameters (represented as color). Gameplay footage and further exposition of the environment variability can be found in movie S1.

Jaderberg et al., Human-level performance in 3D multiplayer games with population-based reinforcement learning, *Science*, 2019.

Figure 2

Approach: details

Impress your audience! 😲



Key result: human-level performance

Fig. 2. Agent architecture and benchmarking. (**A**) How the agent processes a temporal sequence of observations x_t from the environment. The model operates at two different time scales, faster at the bottom and slower by a factor of τ at the top. A stochastic vector-valued latent variable is sampled at the fast time scale from distribution $\mathbb{Q}_t$ on the basis of observations x_t. The action distribution π_t is sampled conditional on the latent variable at each time step t. The latent variable is regularized by the slow moving prior $\mathbb{P}_t$, which helps capture long-range temporal correlations and promotes memory. The network parameters are updated by using RL according to the agent's own internal reward signal r_t, which is obtained from a learned transformation **w** of game points ρ_t. **w** is optimized for winning probability through PBT, another level of training performed at yet a slower time scale than that of RL. Detailed network architectures are described in fig. S11. (**B**) (Top) The Elo skill ratings of the FTW agent population throughout training (blue) together with those of the best baseline agents by using hand-tuned reward shaping (RS) (red) and game-winning reward signal only (black), compared with human and random agent reference points (violet, shaded region shows strength between 10th and 90th percentile). The FTW agent achieves a skill level considerably beyond strong human subjects, whereas the baseline agent's skill plateaus below and does not learn anything without reward shaping [evaluation procedure is provided in (*28*)]. (Bottom) The evolution of three hyperparameters of the FTW agent population: learning rate, Kullback-Leibler divergence (KL) weighting, and internal time scale τ, plotted as mean and standard deviation across the population.

Jaderberg et al., Human-level performance in 3D multiplayer games with population-based reinforcement learning, *Science*, 2019.

Figure 3

Further analyses supporting key result

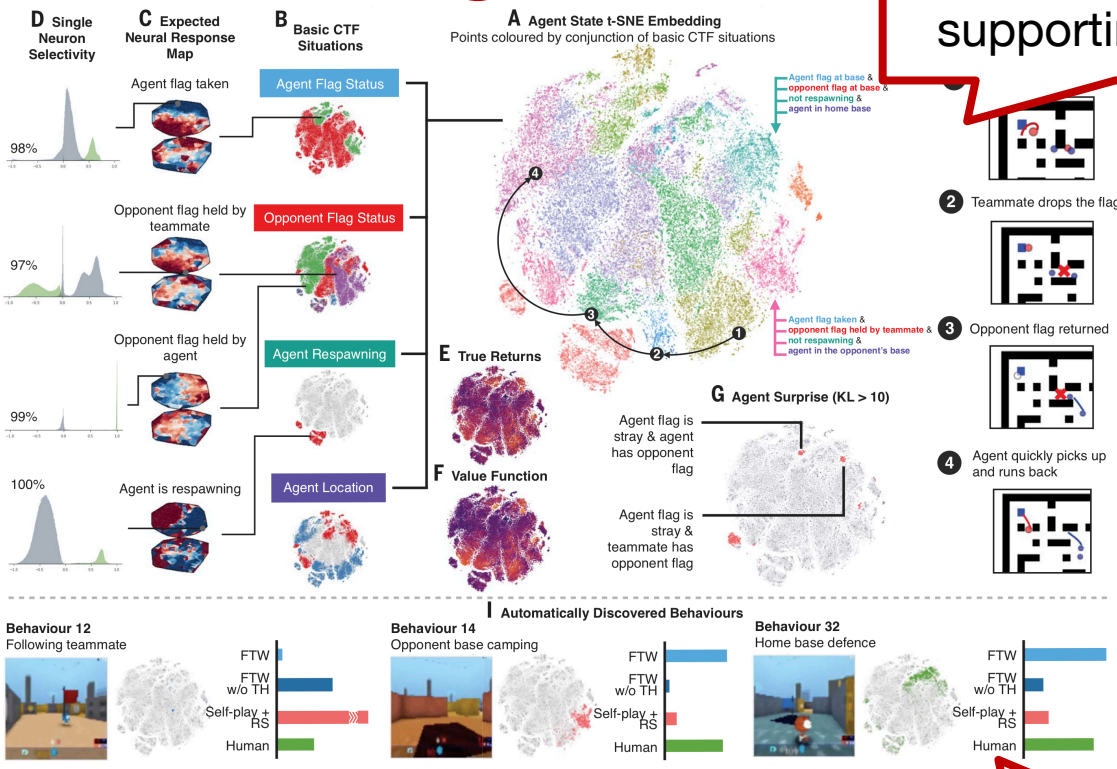


Fig. 3. Knowledge represent… behavioral analysis. (**A**) The … situations and show the predictive accuracy of this neu… true return of the agent's internal reward sig… prediction, its value function (orange denote… denotes low value). (**G**) Regions where the a… representation diverges (red), the agent's su… KL between the agent's slow– and fast–tim… (*28*). (**H**) The four-step temporal sequence … "opponent base camping." (**I**) Three automa… behaviors of agents and corresponding regions in the t-SNE embedding. (Right) Average occurrence per game of each behavior for the FTW agent, the FTW agent without temporal hierarchy (TH), self-play with reward shaping agent, and human subjects (fig. S9).

Note how observations x_t are visualized, i.e., maps of base camps

Note the use of matched colors

Jaderberg et al., Human-level performance in 3D multiplayer games with population-based reinforcement learning, *Science*, 2019.

Figure 4

Further analyses supporting key result

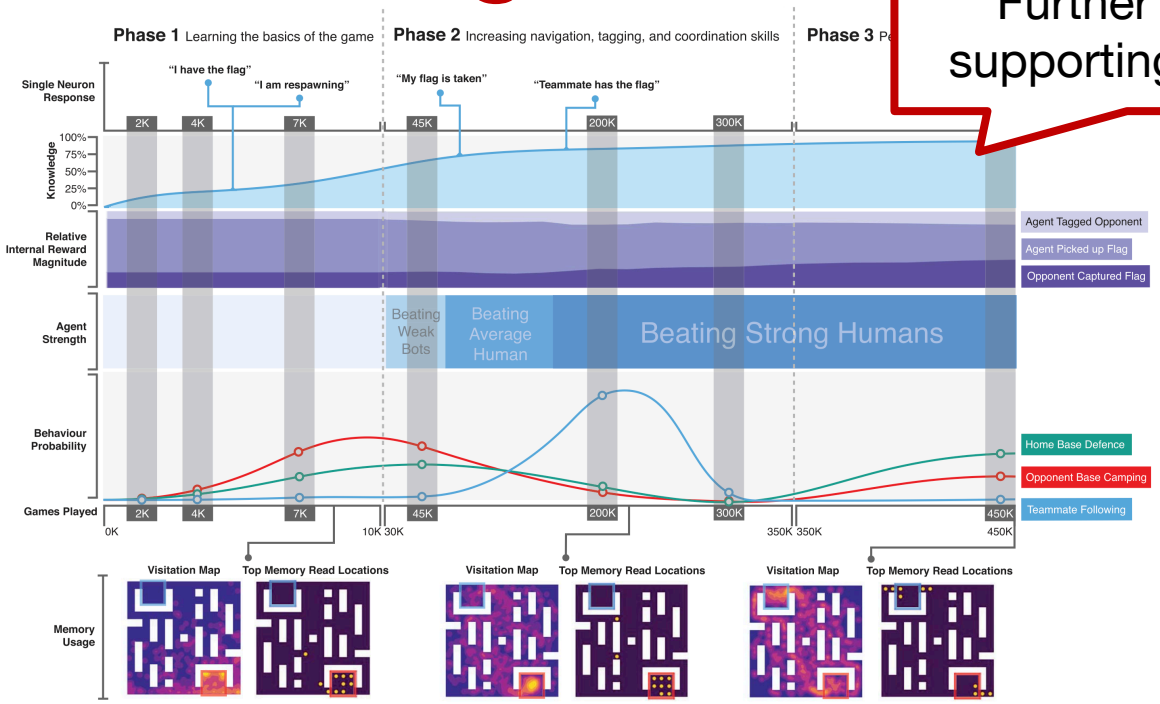


Fig. 4. Progression of agent during training. Shown is the development of knowledge representation and behaviors of the FTW agent over the training period of 450,000 games, segmented into three phases (movie S2). "Knowledge" indicates the percentage of game knowledge that is linearly decodable from the agent's representation, measured by average scaled AUCROC across 200 features of game state. Some knowledge is compressed to single-neuron responses (Fig. 3A), whose emergence in training is shown at the top. "Relative internal reward magnitude" indicates the relative magnitude of the agent's internal reward weights of 3 of the 13 events corresponding to game points ρ. Early in training, the agent puts large reward weight on picking up the opponent's flag, whereas later, this weight is reduced, and reward for tagging an opponent and penalty when opponents capture a flag are increased by a factor of two. "Behavior probability" indicates the frequencies of occurrence for 3 of the 32 automatically discovered behavior clusters through training. Opponent base camping (red) is discovered early on, whereas teammate following (blue) becomes very prominent midway through training before mostly disappearing. The "home base defense" behavior (green) resurges in occurrence toward the end of training, which is in line with the agent's increased internal penalty for more opponent flag captures. "Memory usage" comprises heat maps of visitation frequencies for (left) locations in a particular map and (right) locations of the agent at which the top-10 most frequently read memories were written to memory, normalized by random reads from memory, indicating which locations the agent learned to recall. Recalled locations change considerably throughout training, eventually showing the agent recalling the entrances to both bases, presumably in order to perform more efficient navigation in unseen maps (fig. S7).

Jaderberg et al., Human-level performance in 3D multiplayer games with population-based reinforcement learning, *Science*, 2019.

Evolution of resilience in protein interactomes across the tree of life

Marinka Zitnik[a], Rok Sosič[a], Marcus W. Feldman[b,1], and Jure Leskovec[a,c,1]

[a]Department of Computer Science, Stanford University, Stanford, CA 94305; [b]Department of Biology, Stanford University, Stanford, CA 94305; and [c]Chan Zuckerberg Biohub, San Francisco, CA 94158

Contributed by Marcus W. Feldman, December 18, 2018 (sent for review October 19, 2018; reviewed by Edoardo Airoldi and Aviv Bergman)

Phenotype robustness to environmental fluctuations is a common biological phenomenon. Although most phenotypes involve multiple proteins that interact with each other, the basic principles of how such interactome networks respond to environmental unpredictability and change during evolution are largely unknown. Here we study interactomes of 1,840 species across the tree of life involving a total of 8,762,166 protein–protein interactions. Our study focuses on the resilience of interactomes to network failures and finds that interactomes become more resilient during evolution, meaning that interactomes become more robust to network failures over time. In bacteria, we find that a more resilient interactome is in turn associated with the greater ability of the organism to survive in a more complex, variable, and competitive environment. We find that at the protein family level proteins exhibit a coordinated rewiring of interactions over time and that a resilient interactome arises through gradual change of the network topology. Our findings have implications for understanding molecular network structure in the context of both evolution and environment.

Zitnik et al., Evolution of resilience in protein interactomes across the tree of life, *PNAS*, 2019.

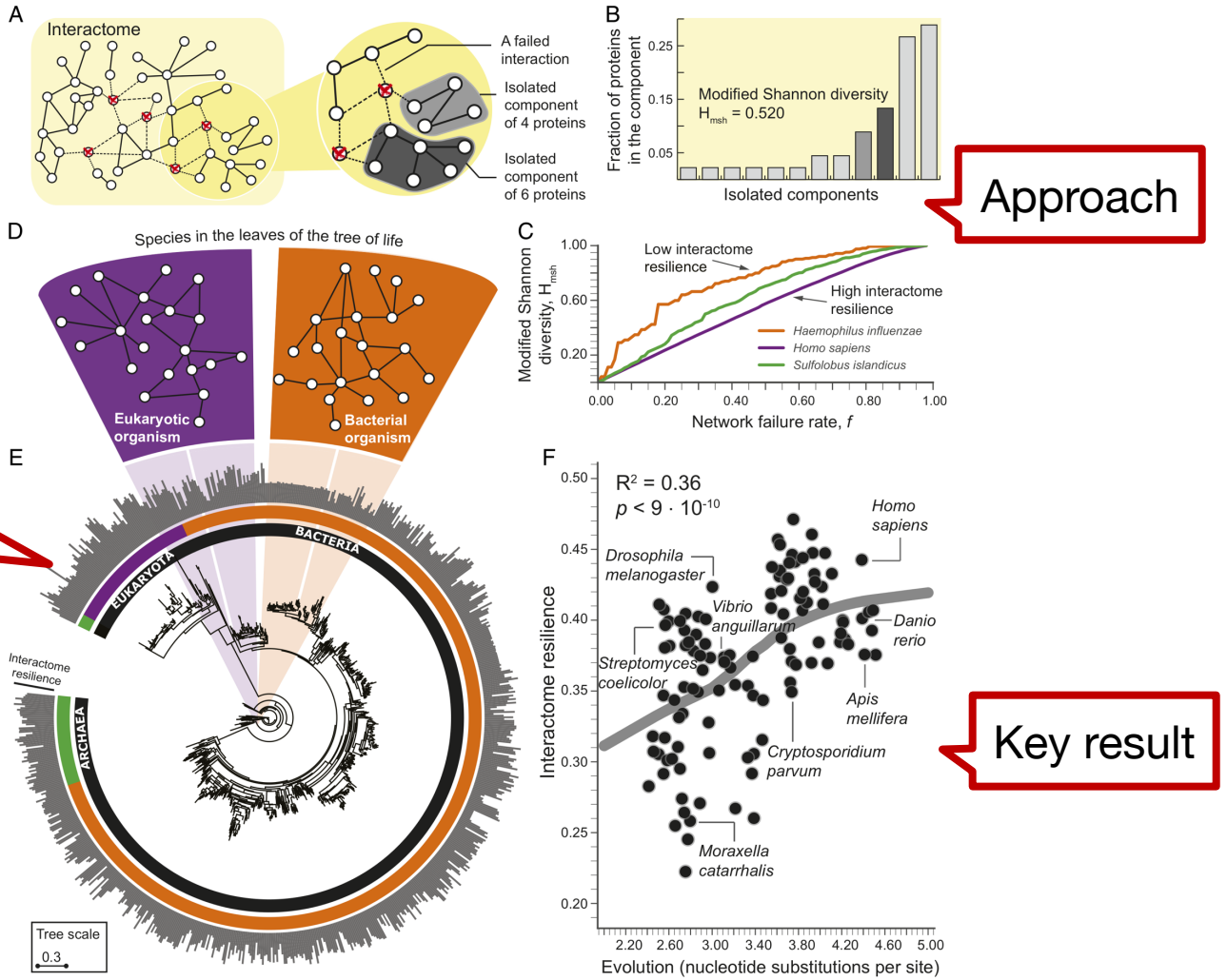
Figure 1

Dataset

Approach

Impress your audience! 😲

Key result



Zitnik et al., Evolution of resilience in protein interactomes across the tree of life, *PNAS*, 2019.

Figure 2

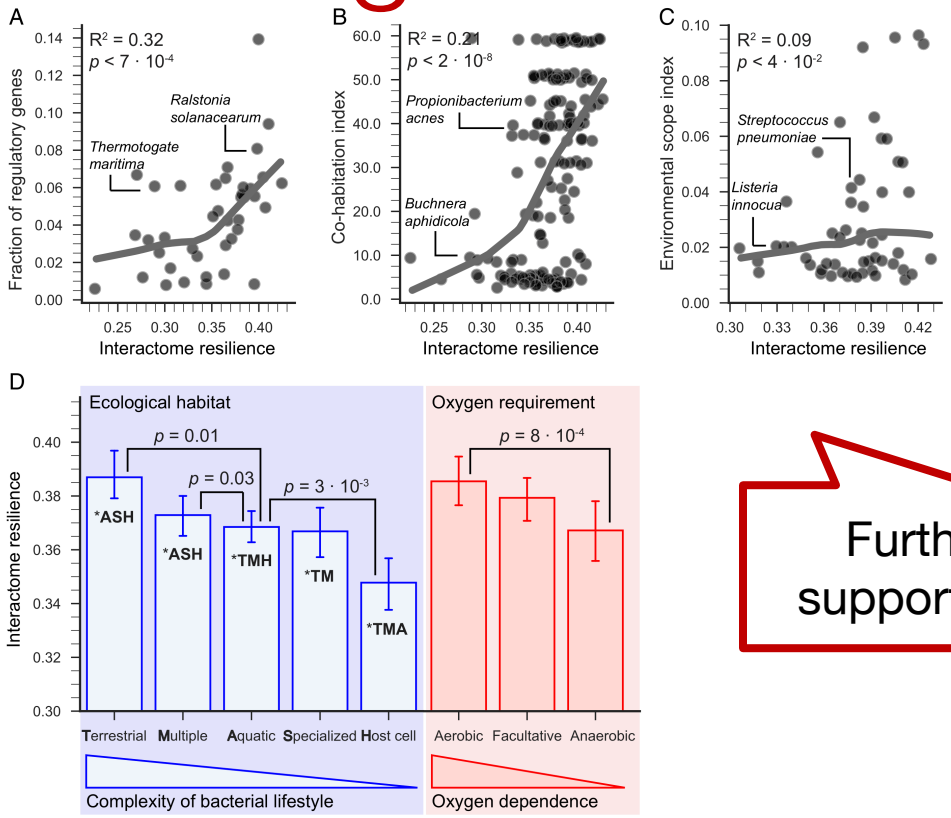
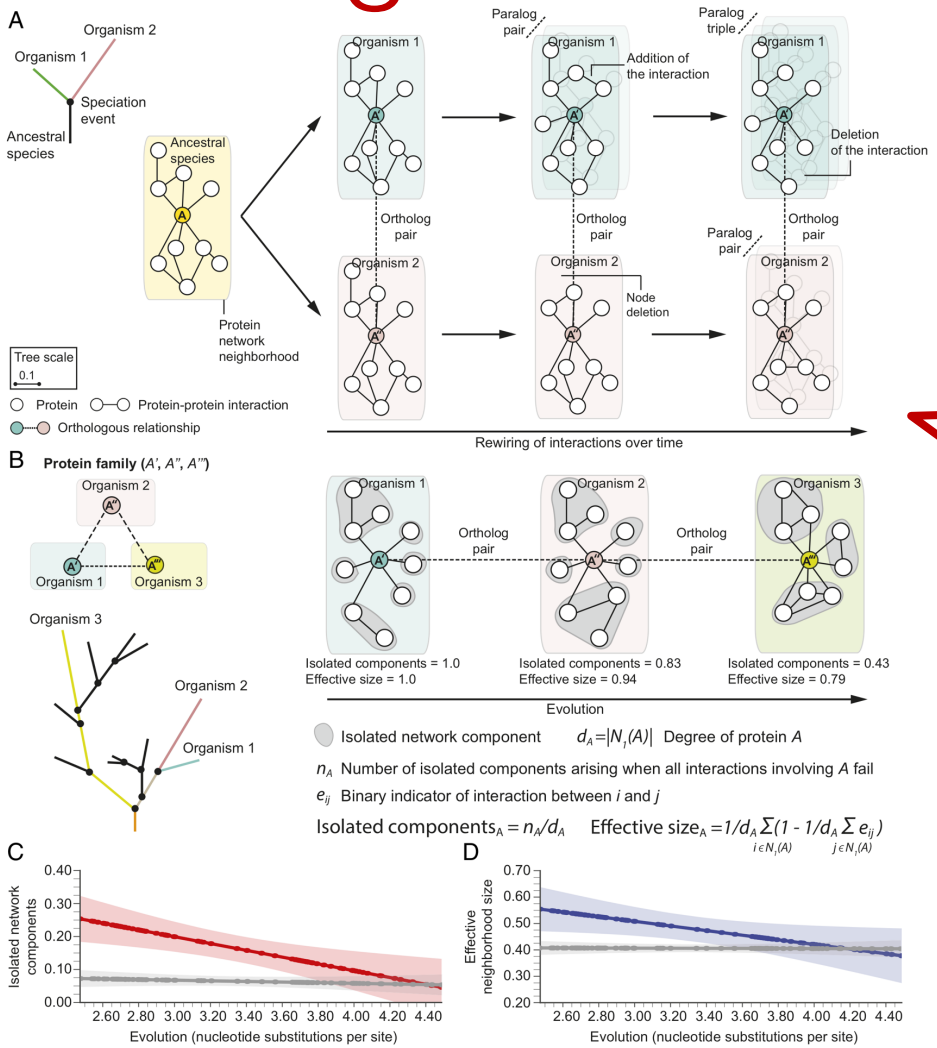


Fig. 2. Bacteria with more resilient interactomes survive in more complex, variable, and competitive environments. We use ecological information for 287 bacterial species (32) to examine the relationship between species' interactome resilience and their ecology (*SI Appendix*, section S4). (*A*) Interactome resilience positively correlates with the fraction of regulatory genes in bacteria, an established indicator of environmental variability of species' habitats (32) ($R^2 = 0.32$). (*B* and *C*) For environmental viability of a species, we use a cohabitation index that records how many organisms populate each environment in which the species is viable (i.e., the level of competition in each viable environment) and an environmental scope index that records a fraction of the environments in which the species is viable (i.e., species' environmental diversity) (32). The resilience of the interactome positively correlates with the level of cohabitation encountered by bacteria ($R^2 = 0.21$), and bacteria with resilient interactomes tend to thrive in highly diverse environments ($R^2 = 0.09$). (*D*) Terrestrial bacteria have the most resilient interactomes ($P = 7 \cdot 10^{-3}$), and host-associated bacteria have the least resilient interactomes ($P = 4 \cdot 10^{-6}$). In bacteria, interactome resilience is indicative of oxygen dependence. Aerobic bacteria have the most resilient interactomes ($P = 8 \cdot 10^{-4}$), followed by facultative and the anaerobic bacteria. Error bars indicate 95% bootstrap confidence interval; *P* values denote the significance of the difference of the means according to a Mann–Whitney *U* test.

Further analyses supporting key result

Zitnik et al., Evolution of resilience in protein interactomes across the tree of life, *PNAS*, 2019.

Figure 3



Approach

Key results

Zitnik et al., Evolution of resilience in protein interactomes across the tree of life, *PNAS*, 2019.

Figure 4

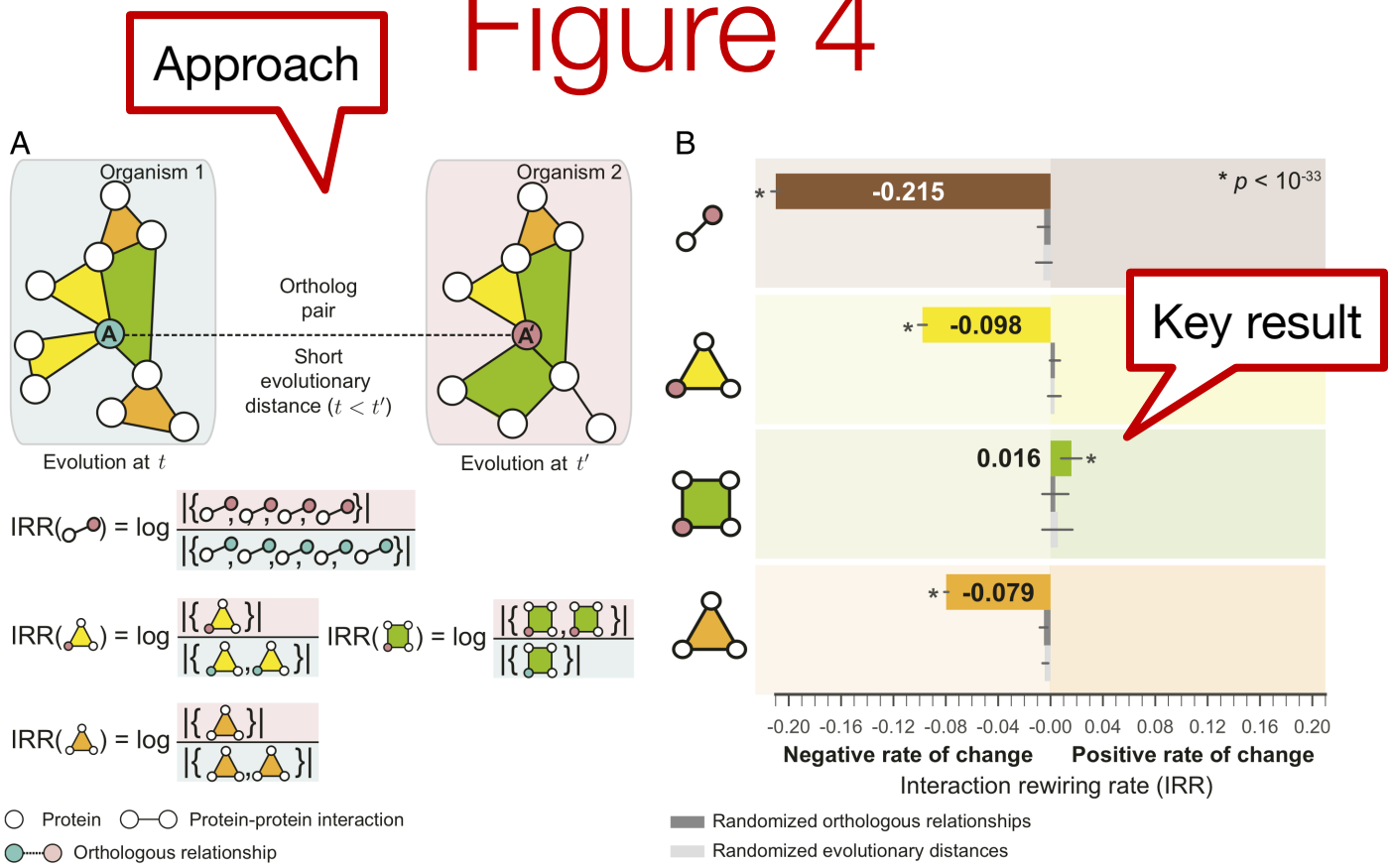


Fig. 4. The rewiring rate of interactions in local protein neighborhoods varies with the topology of network motifs. (*A*) Interaction rewiring rate (IRR) measures the fold change between the probability of observing a particular network motif in the network neighborhood of protein A' and the probability of observing the same motif in the neighborhood of an evolutionarily younger orthologous protein A. A positive (negative) rate of change indicates the motif becomes more (less) common over time (*SI Appendix*, section S7). Shown are the rewiring rates for interactions (i.e., edges; the number of interactors of A' vs. A), triangle motifs touching the orthologous protein (yellow), square motifs touching the orthologous protein (green), and triangle motifs in the protein network neighborhood (orange). (*B*) Square motifs become more common in protein neighborhoods during evolution ($P < 10^{-33}$), which is supported by a range of biological evidence (18, 37, 38). However, triangle motifs become less common over time ($P < 10^{-33}$ for both types of triangle motifs). Gray bars indicate random expectation (*SI Appendix*, section S7), either for random orthologous relationships (dark gray) or for random evolutionary distances (light gray); error bars indicate 95% bootstrap confidence interval; and *P* values denote the significance of the difference of IRR distributions using a two-sample Kolmogorov–Smirnov test.

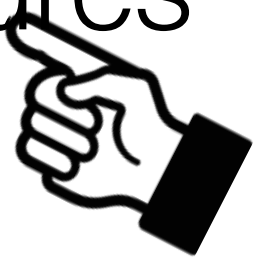
Zitnik et al., Evolution of resilience in protein interactomes across the tree of life, *PNAS*, 2019.

Today's Lecture

1) Why figures matter

2) Figures in science

3) How to design effective figures



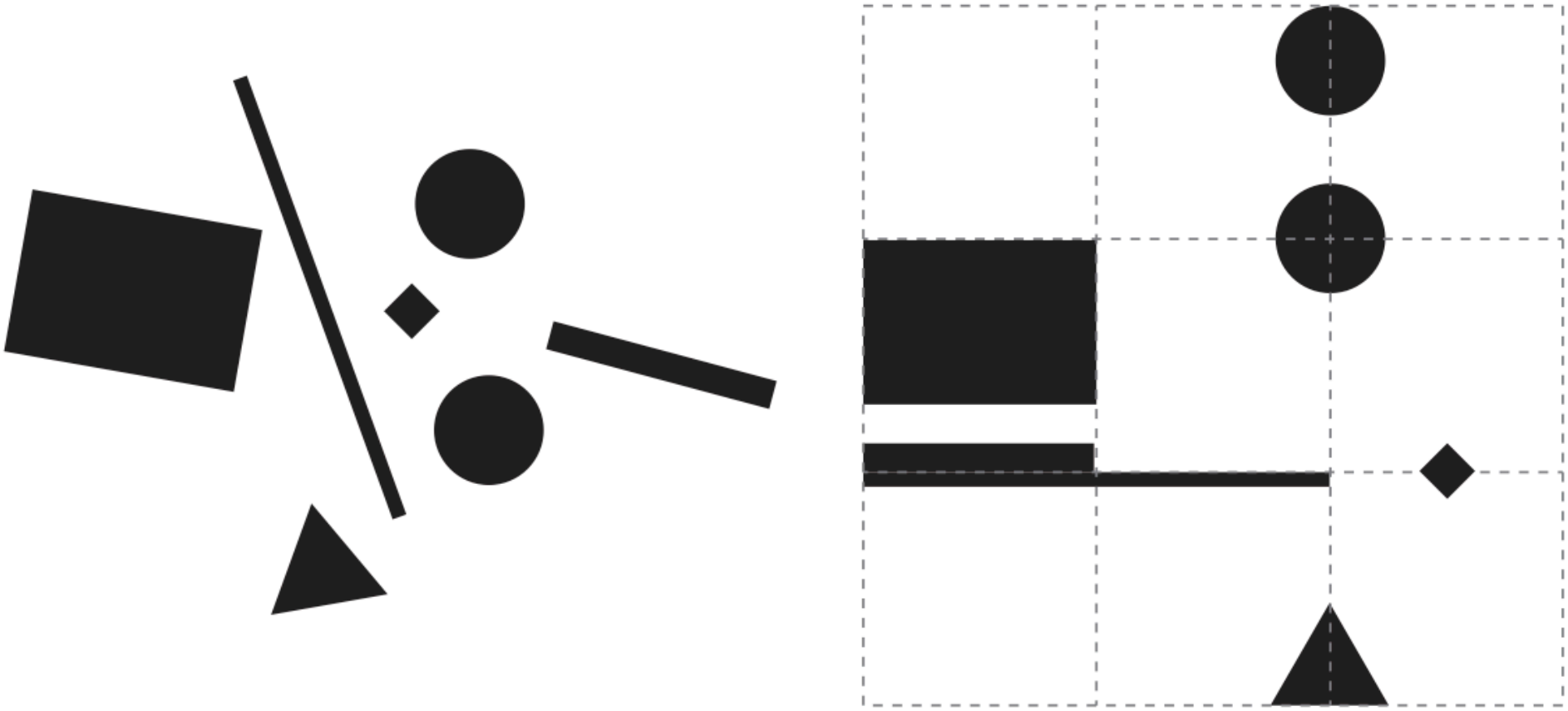
4) Tools, tips, and guidelines

Principle #1: Design figures for the audience (not for you)

Before your design figures think about:

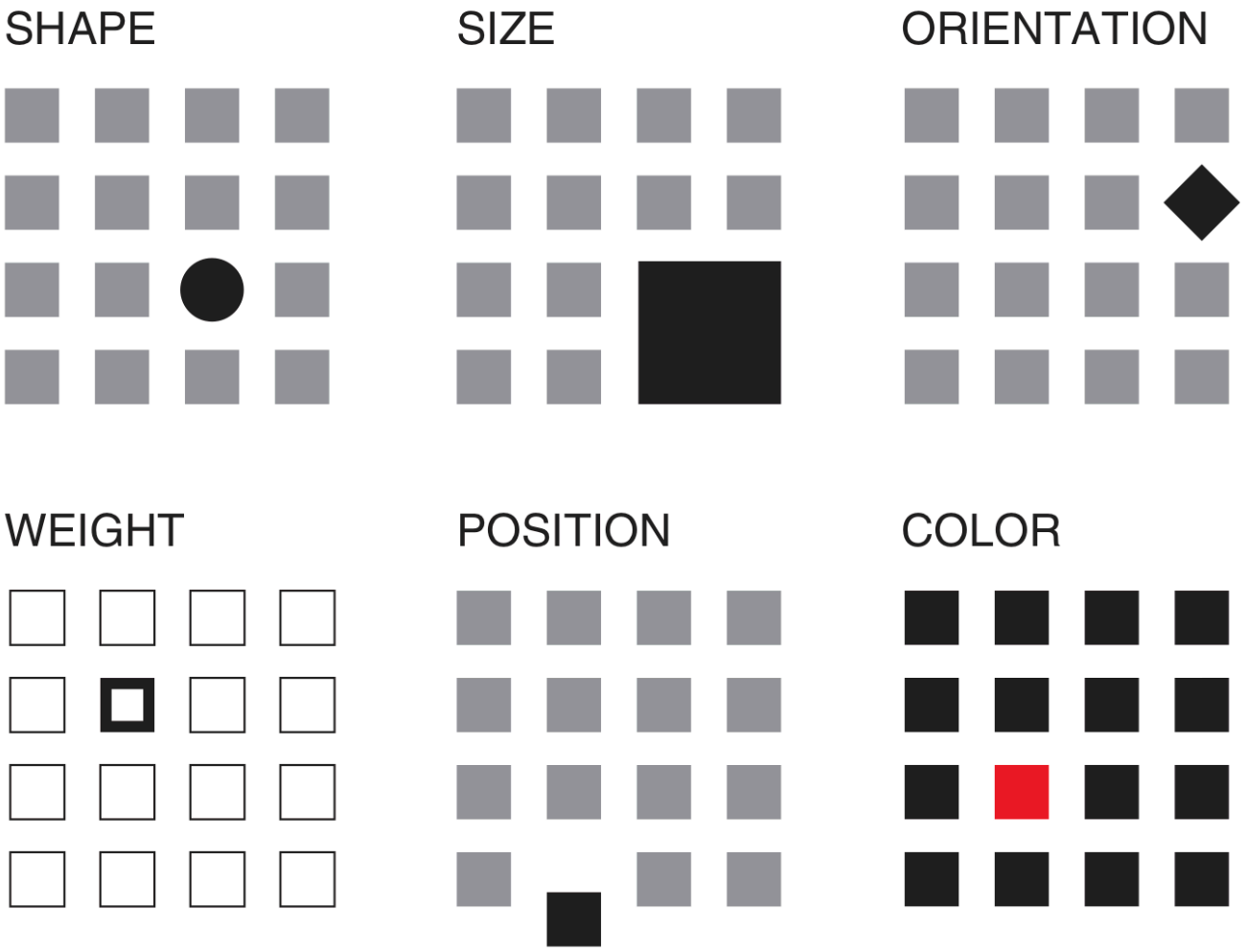
1) Make-up of the audience:
   - Will a figure appear in a specialized journal?
   - Is a figure aimed at a broad readership?
2) Background knowledge of the audience:
   - Audience may not know what you know
   - Figures should provide all the information necessary for the audience to fully comprehend them
3) Disciplinary conventions:
   - Graphical conventions and norms exist in each field

Principle #2: Design a clear visual structure with pleasant symmetries



Rolandi *et al.,* A brief guide to designing effective figures for the scientific paper. Advanced Materials 23.38 (2011)

Principle #3: Use visual contrast, but keep figures simple



Rolandi *et al.,* A brief guide to designing effective figures for the scientific paper. Advanced Materials 23.38 (2011)

Principle #4: Use readable and legible typography

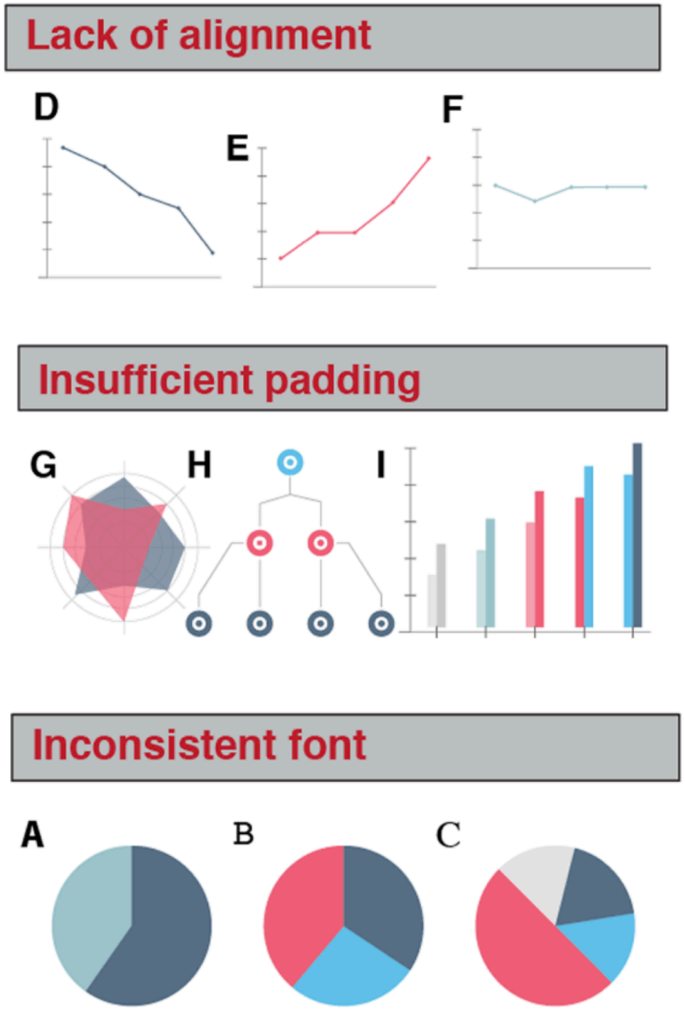
Adequate readability due to high value contrast

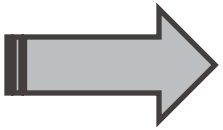
Inadequate readability due to low value contrast

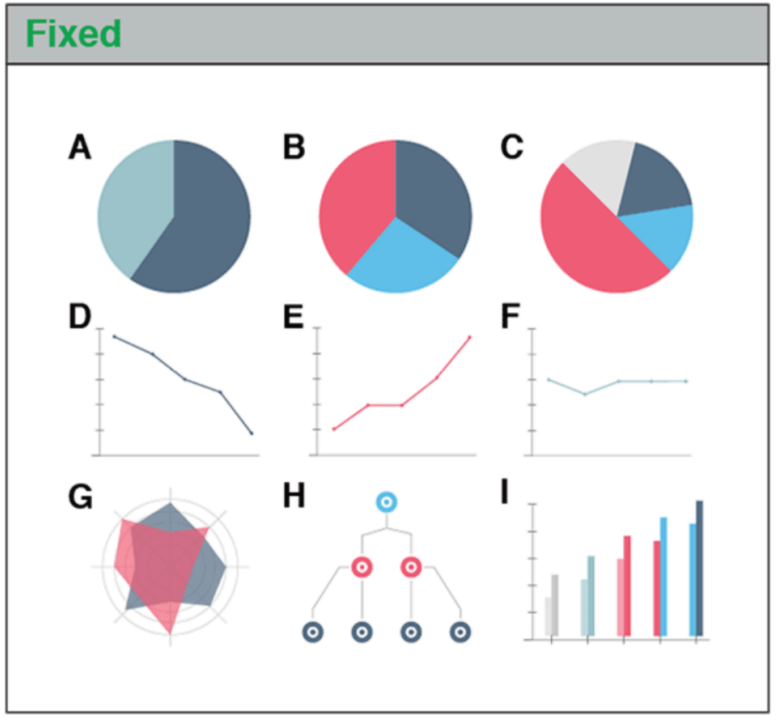
Inadequate readability due to patterned background

Rolandi *et al.,* A brief guide to designing effective figures for the scientific paper. Advanced Materials 23.38 (2011)

Principle #5: Be consistent, align panels and use sufficient padding







Source: Jean Fan, Harvard

The Good, Bad, and Ugly

Good

Hello
World

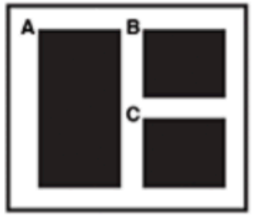
Bad

Hello
World

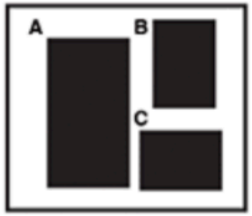
Ugly

Hello
World

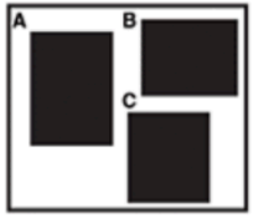
Good



Bad



Ugly



Good

Bad

Ugly

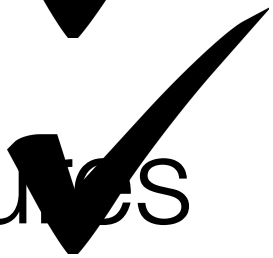
Source: Jean Fan, Harvard

Today's Lecture

1) Why figures matter

2) Figures in science

3) How to design effective figures



4) Tools, tips, and guidelines

Key Rules to ALWAYS Follow

1) Save raw data and results to a tsv/csv/binary file:
   - Your figures will need multiple rounds of editing
2) Read in the data and design figures

Important: Save figures as PDF or other vector format:

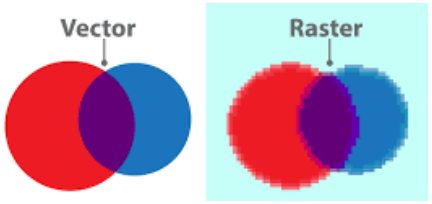
- You might need to use multiple tools to draw a figure
  - Example:
    1. First, use seaborn to draw a clustermap
    2. Then, export clustermap as PDF
    3. Finally, use Adobe Illustrator to annotate the clustermap
  - Example:
    1. First, use D3.js to layout a network
    2. Then, export the network as PDF
    3. Finally, use Adobe Illustrator to show node features and node labels

Why shouldn't you use raster formats (e.g., JPG, GIF, PNG, TIF)?

Raster images:

- Use a fixed number of colored pixels and can't be dramatically resized (pixilation, distortion issues)
- When saved, they cannot be reopened and edited!

Vector images (e.g., PDF, EPS, AI, SVG):



- Remain editable!
- You can open them in Illustrator and edit text or any other element within the graphic
- Can be converted to a raster image but not vice-versa
- plt.savefig('myfig.pdf')

Only use raster format for web, Github repo, etc.

Tools, Software & Frameworks

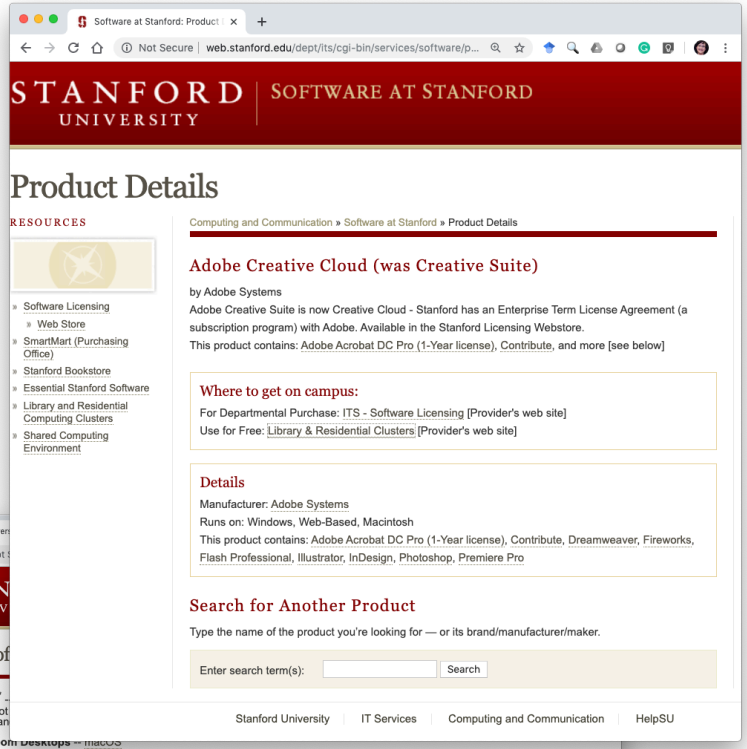
Tools, Software, and Frameworks

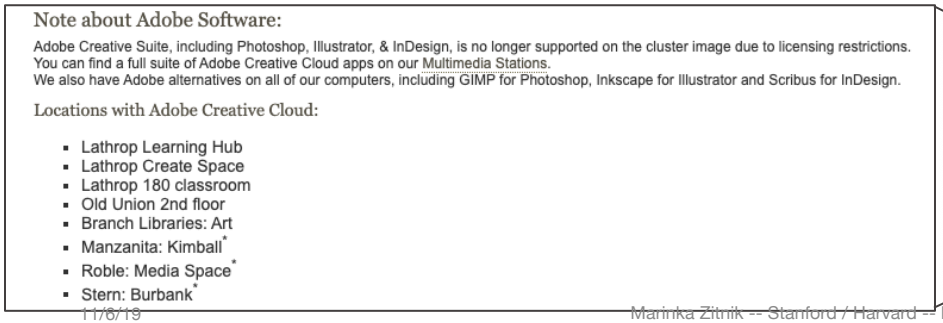
- Adobe Illustrator
  - Adobe Creative Cloud
- LaTeXiT
  - chachatelier.fr/latexit
- Matplotlib
  - matplotlib.org
- Seaborn
  - seaborn.pydata.org
- Bokeh
  - bokeh.pydata.org
- D3.js
  - d3js.org
- GeoPandas
  - geopandas.org
- Google Charts
  - developers.google.com/chart
- Circos
  - circos.ca
- gnuplut
  - gnuplot.info
- TikZ
  - texample.net/tikz
- Plotly
  - plot.ly/python
- missingno
  - github.com/ResidentMario/missingno
- billboard.js
  - naver.github.io/billboard.js
- Squaire.js
  - wsj.github.io/squaire

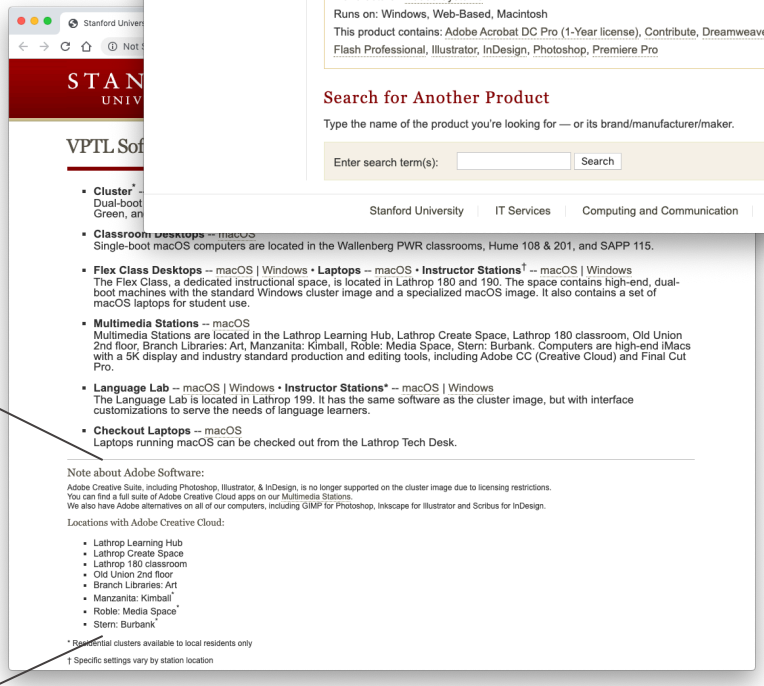
Adobe Illustrator and Alternatives

- Where to get on campus:
  - For departmental purchase
  - **Use for Free:** Stanford Library & Residential Clusters

- Free alternatives:
  - Inkscape, https://inkscape.org
  - GIMP, https://www.gimp.org
  - Boxy-SVG, https://boxy-svg.com







How to get from a JS vis to an effective figure?

Three steps:

1) Use a JS library from two slide ago and generate a visualization

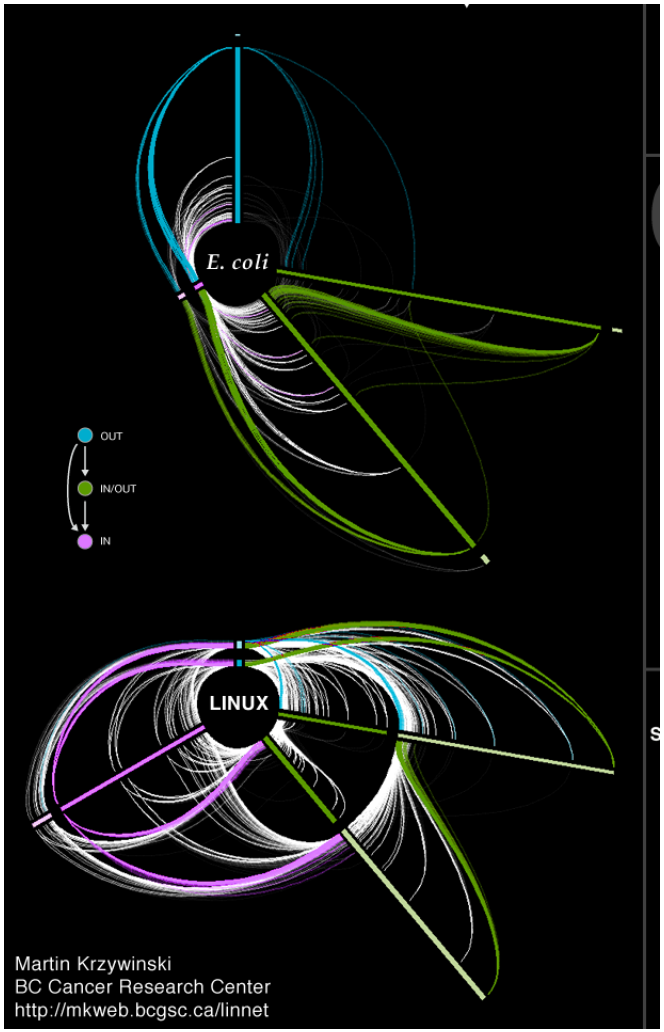
2) Generate a PDF file from HTML:
   - stackoverflow.com/questions/18191893/generate-pdf-from-html-in-div-using-javascript

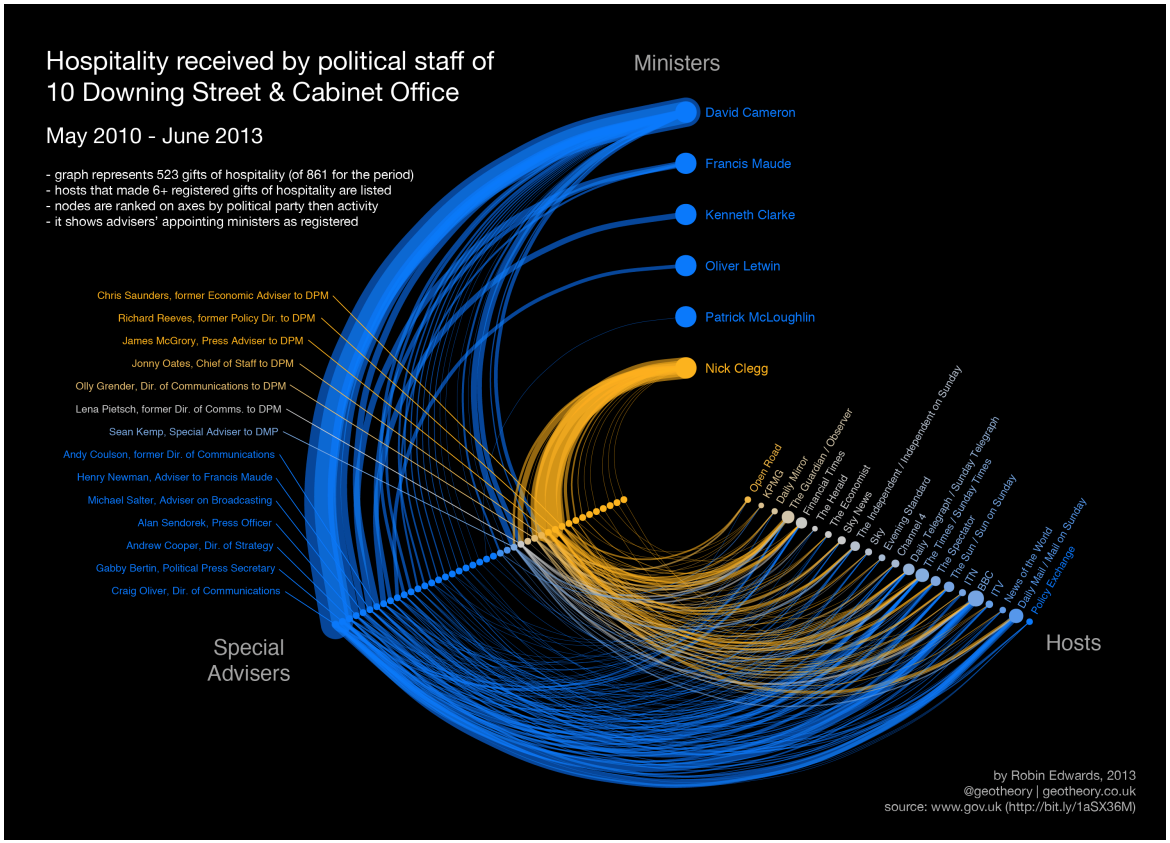
3) Open the PDF in Illustrator and make further edits:
   - Change colors
   - Add labels and annotations
   - Add new visual elements, e.g., insets, logos
   - Combine with other graphics to get a multi-panel figure

Tools for Network & Relational Data

- Gephi, gephi.org
- Graphviz, graphviz.org
- NetworkX, networkx.github.io
- JSNetworkX, jsnetworkx.org
- igraph, igraph.org/python
- sigma.js, sigmajs.org
- Cytoscape, cytoscape.org
- Hive plots, hiveplot.com

Hive Plots





Nodes are mapped to and positioned on radially distributed linear axes — this mapping is based on network structural properties. Edges are drawn as curved links.

How to draw networks with features, labels, weights, directions?

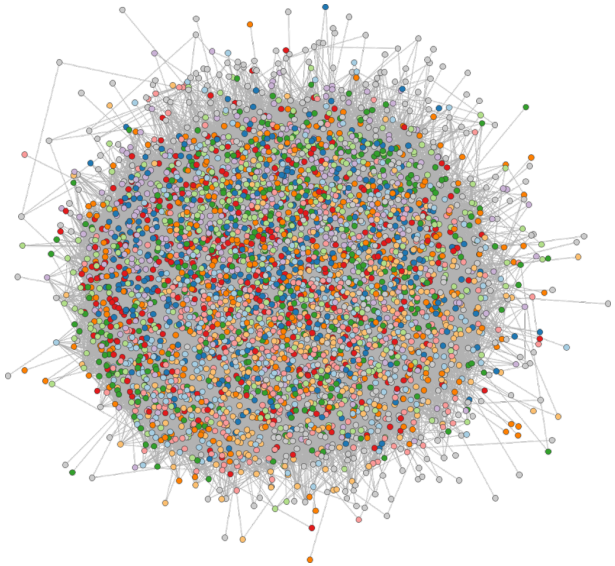
Four steps:

1) Use NetworkX to create a network with metadata:
   - nx.set_node_attributes(G, {0: {'attr1': 20, 'attr2': 'nothing'}, 1: {'attr2': 3}})
   - nx.set_edge_attributes(G, attrs = {(0, 1): {'attr1': 20, 'attr2': 'nothing'}, (1, 2): {'attr2': 3}})
2) Write the network in Gephi's GEXF format:
   - nx.write_gexf(G, "net.gexf")
3) Use Gephi to layout, color, visualize, annotate the net:
   - Can select and then edit any subgraph based on any combination of metadata
4) Export the network as a PDF figure from Gephi

How to select a network layout?

Goal: No hairballs in your papers!

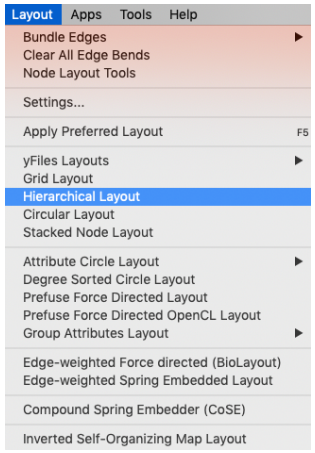
- Rule of thumb: Can visualize networks with <1,000 nodes
- Unless networks have special structure or have custom network layouts



Hairballs can be pretty, but are they useful? What we need is insight. Not a picture!

Gephi

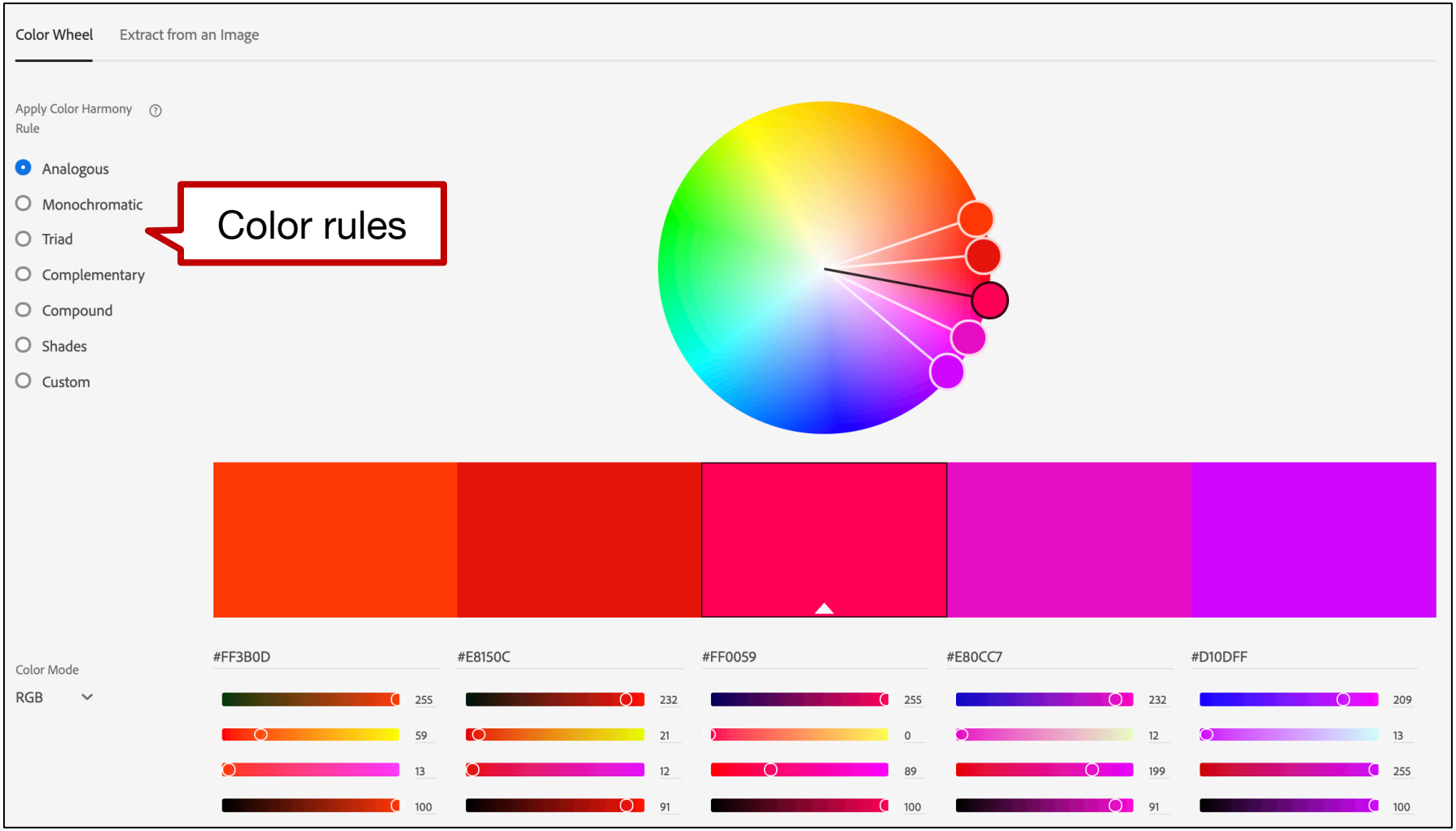
Cytoscape



Colors

Color Advice

Adobe color, https://color.adobe.com



Color Advice: Brewer Palettes

Brewer palettes: Color combinations selected for their special properties for use in data visualization

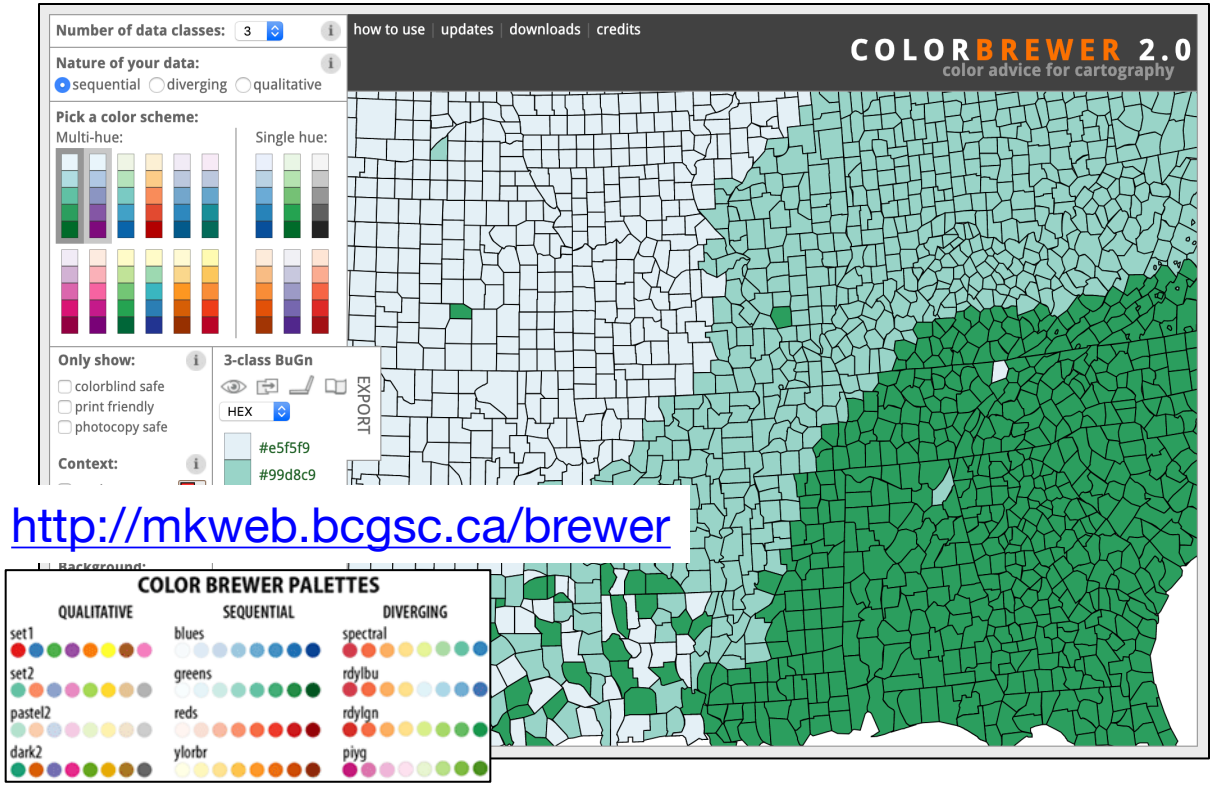
Color Brewer, http://colorbrewer2.org

3 types of palettes:

qualitative — colors do not have a perceived order

sequential — colors have a perceived order and perceived difference between successive colors is uniform

diverging — two back-to-back sequential palettes starting from a common color



http://mkweb.bcgsc.ca/brewer

Color palettes for color blindness, http://mkweb.bcgsc.ca/colorblind

Where to Get Ideas for Effective Figures?

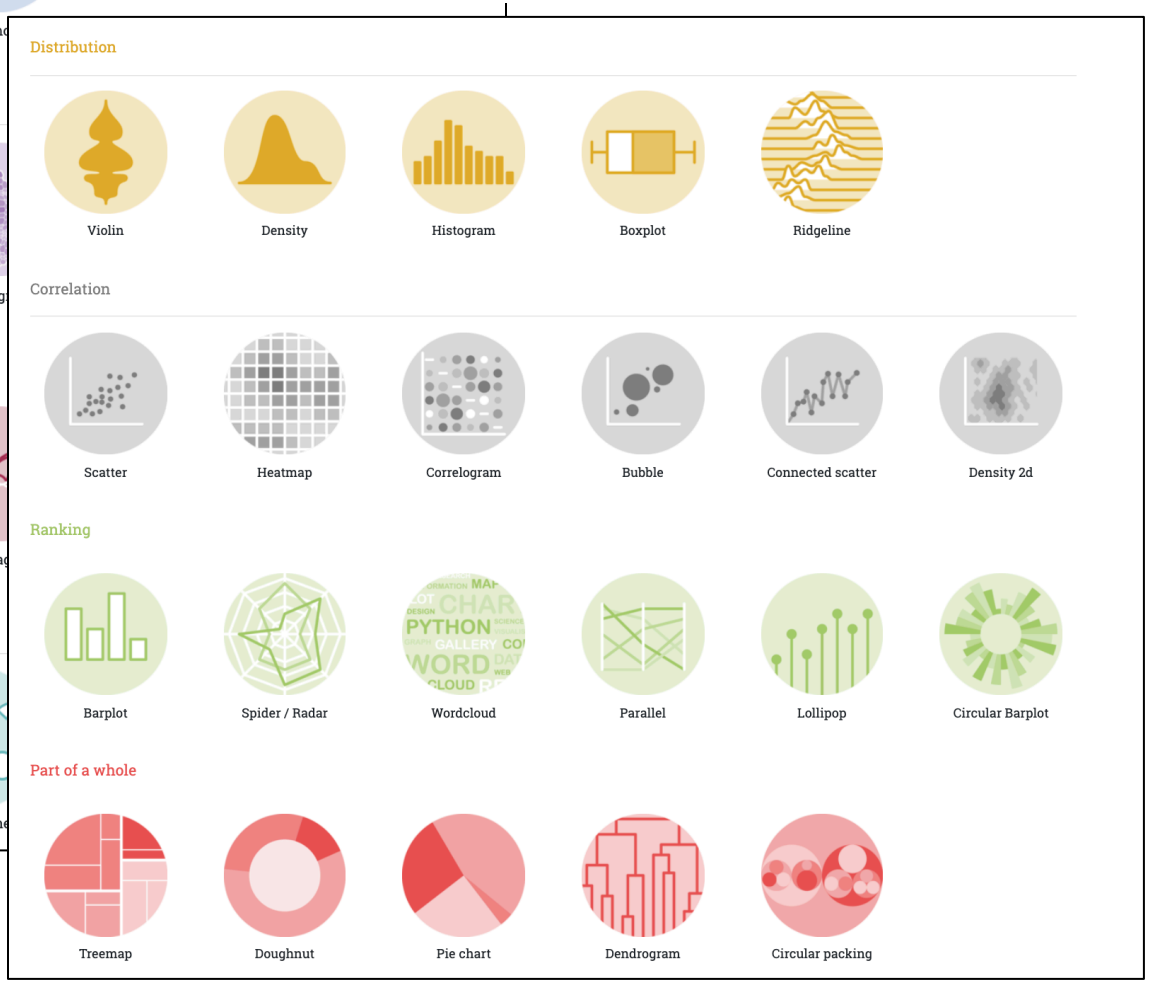
Where to get ideas for figures?

1) Papers published in last issues of Nature, Science, PNAS, Nature Methods, Nature Biotech, etc.
   - No need to read the papers, just look at figures!
2) Martin Krzywinski, mkweb.bcgsc.ca
   - Inventor of several popular visualization tools
   - Designed many Nature, Science, etc. covers
3) www.d3-graph-gallery.com
   - Gallery with hundreds of chart, graphs, geo, part-of-whole
   - Reproducible & editable source code!
4) developers.google.com/chart/interactive/docs/gallery
   - Over 30 chart types, including many non-standard ones
   - Tutorials and source code for every chart type!

Where to get ideas for figures?

www.d3-graph-gallery.com

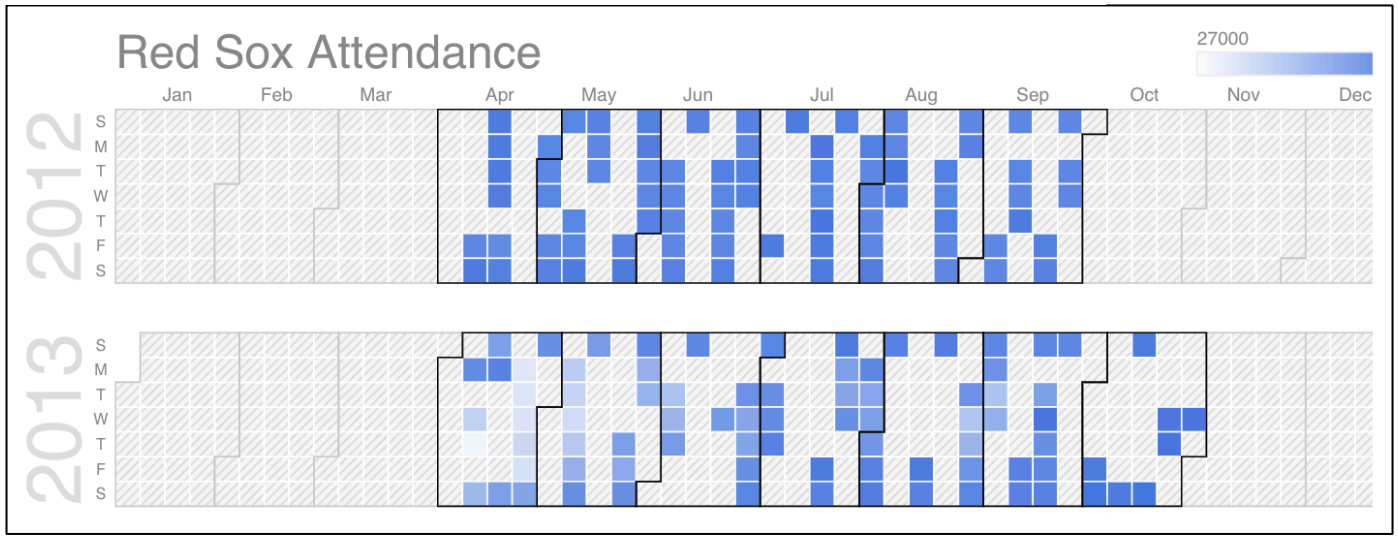
Many non-standard, but highly effective chart types. Source code!



Where to get ideas for figures?

https://developers.google.com/chart with source code!

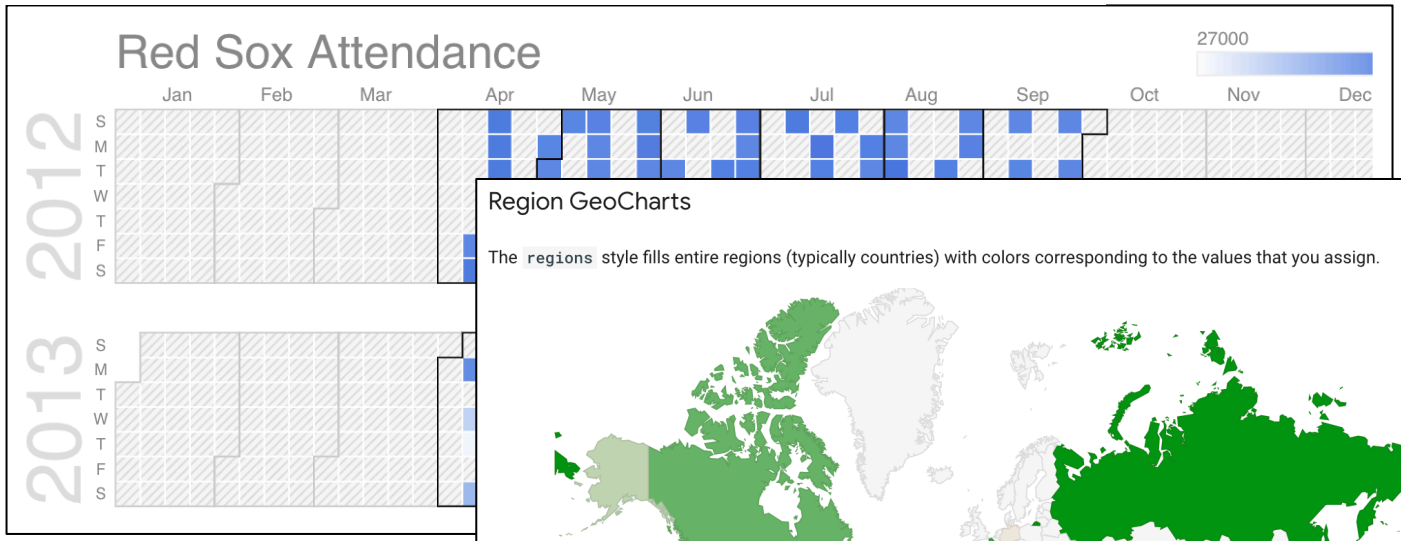
- Chart Types
- Chart Gallery
- Annotation Charts
- Area Charts
- Bar Charts
- Bubble Charts
- Calendar Charts
- Candlestick Charts
- Column Charts
- Combo Charts
- Diff Charts
- Donut Charts
- Gantt Charts
- Gauge Charts
- GeoCharts
- Histograms
- Intervals
- Line Charts
- Maps
- Org Charts
- Pie Charts
- Sankey Diagrams
- Scatter Charts
- Stepped Area Charts
- Table Charts
- Timelines
- Tree Map Charts
- Trendlines
- Waterfall Charts
- Word Trees
- Miscellaneous Examples

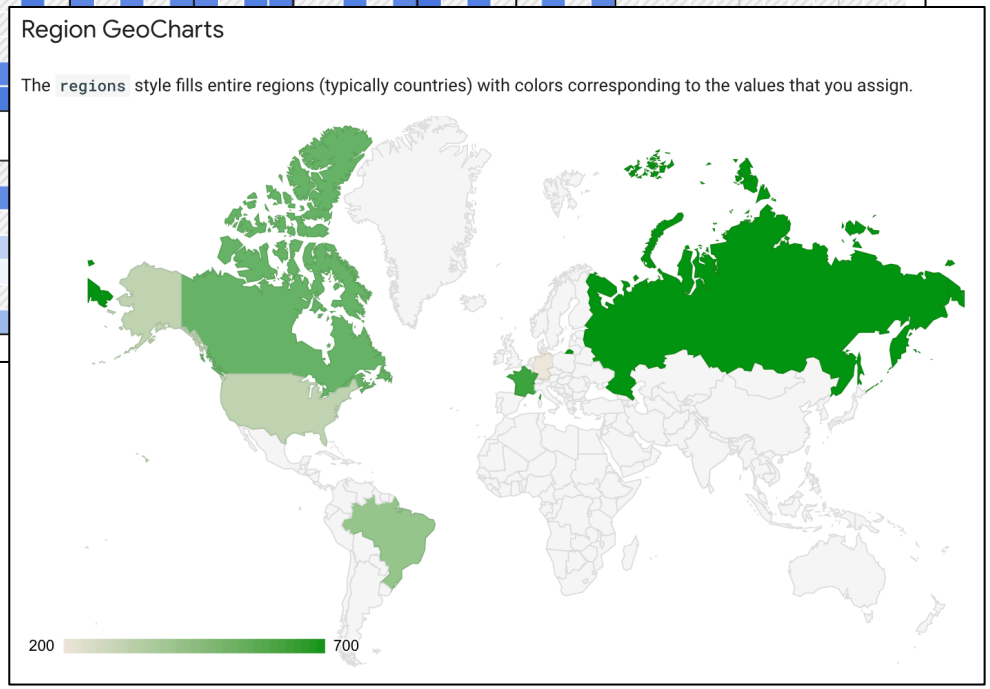


Where to get ideas for figures?

https://developers.google.com/chart with source code!

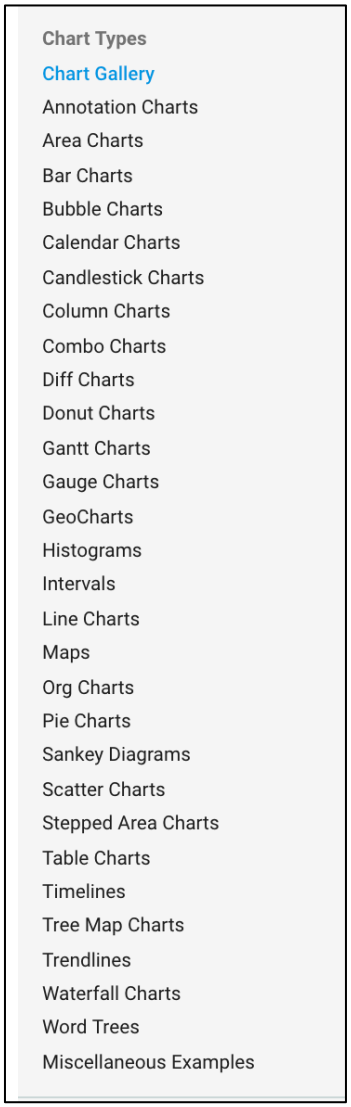
- Chart Types
- Chart Gallery
- Annotation Charts
- Area Charts
- Bar Charts
- Bubble Charts
- Calendar Charts
- Candlestick Charts
- Column Charts
- Combo Charts
- Diff Charts
- Donut Charts
- Gantt Charts
- Gauge Charts
- GeoCharts
- Histograms
- Intervals
- Line Charts
- Maps
- Org Charts
- Pie Charts
- Sankey Diagrams
- Scatter Charts
- Stepped Area Charts
- Table Charts
- Timelines
- Tree Map Charts
- Trendlines
- Waterfall Charts
- Word Trees
- Miscellaneous Examples

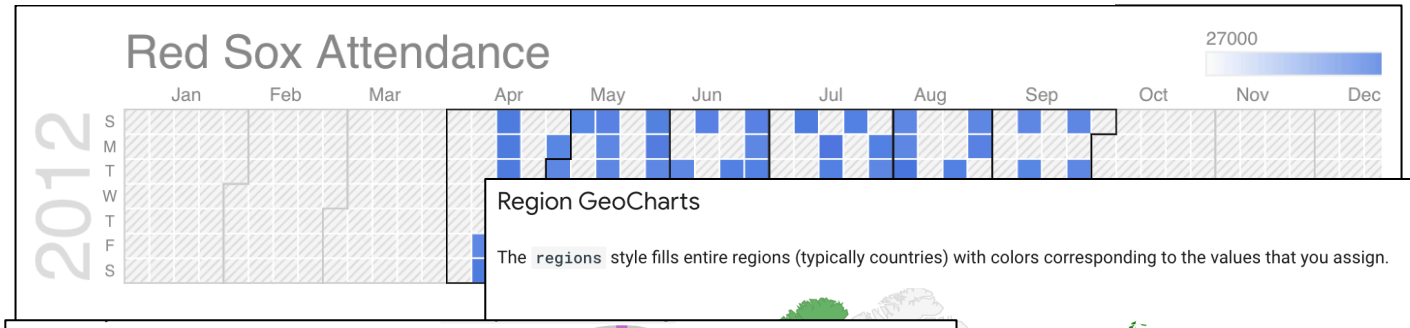


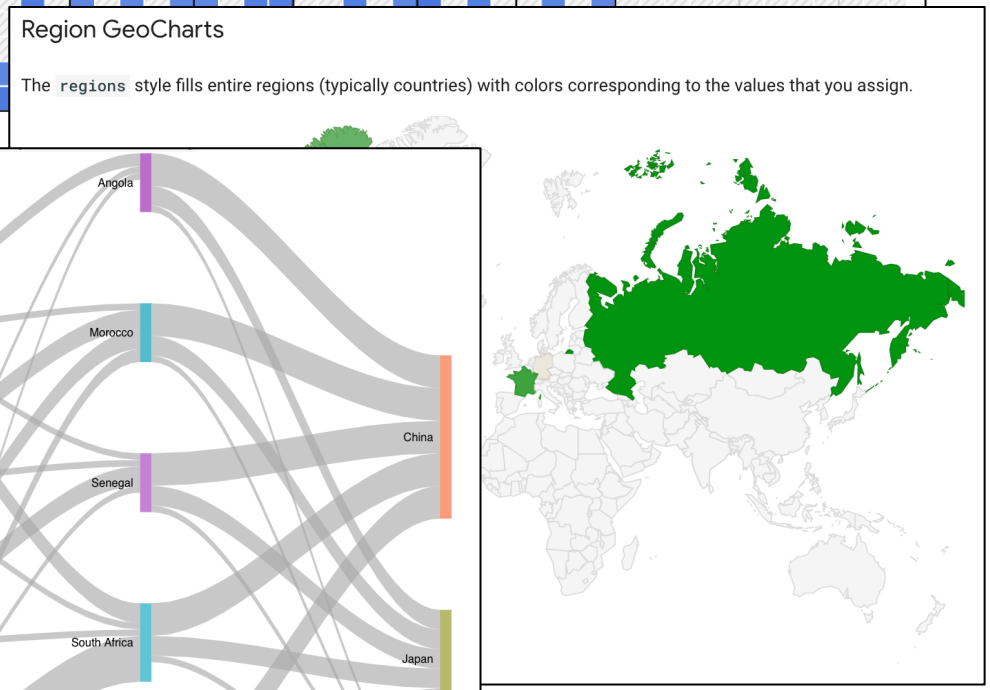


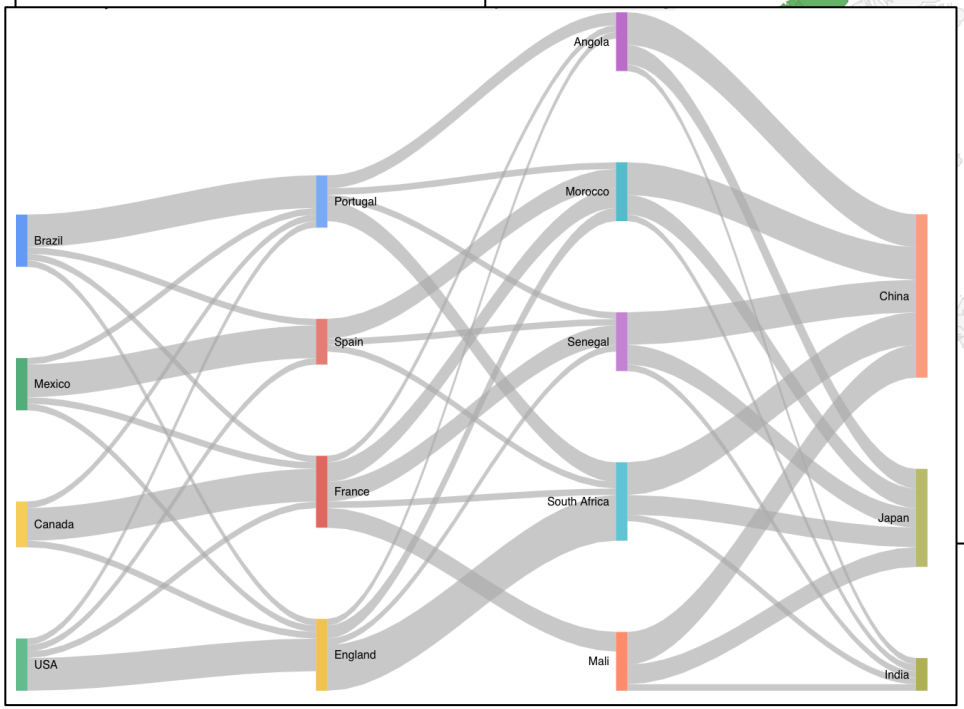
Where to get ideas for figures?

https://developers.google.com/chart with source code!



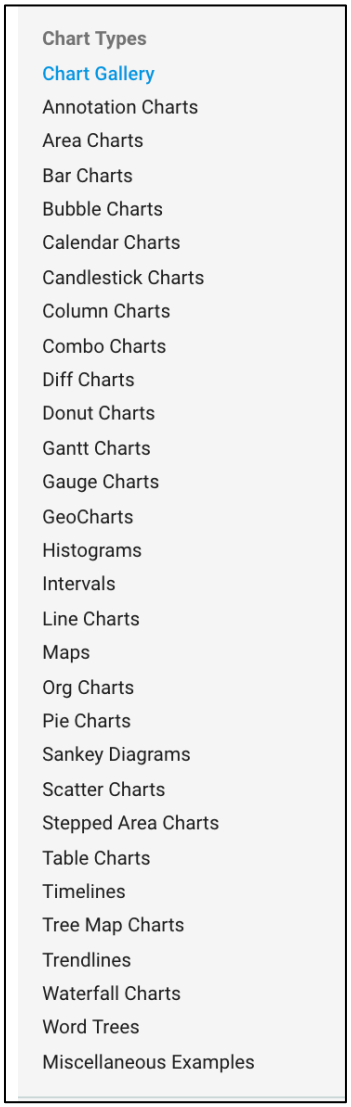


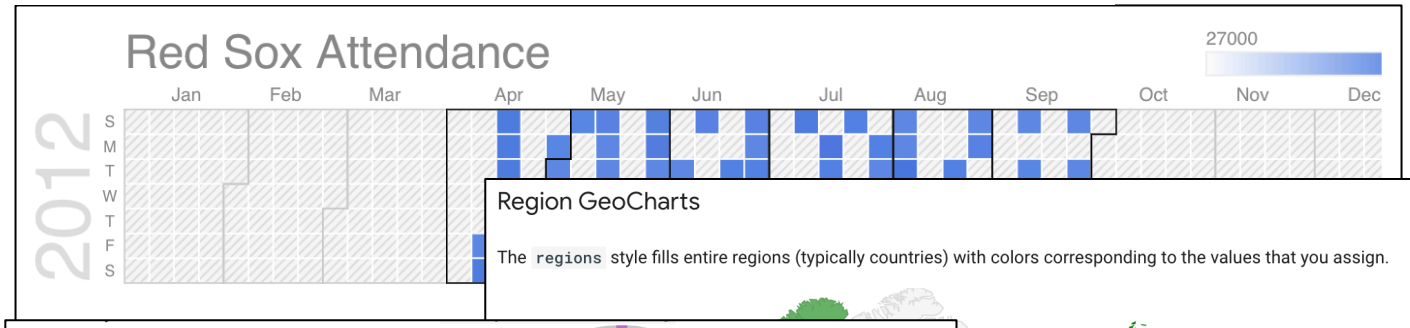


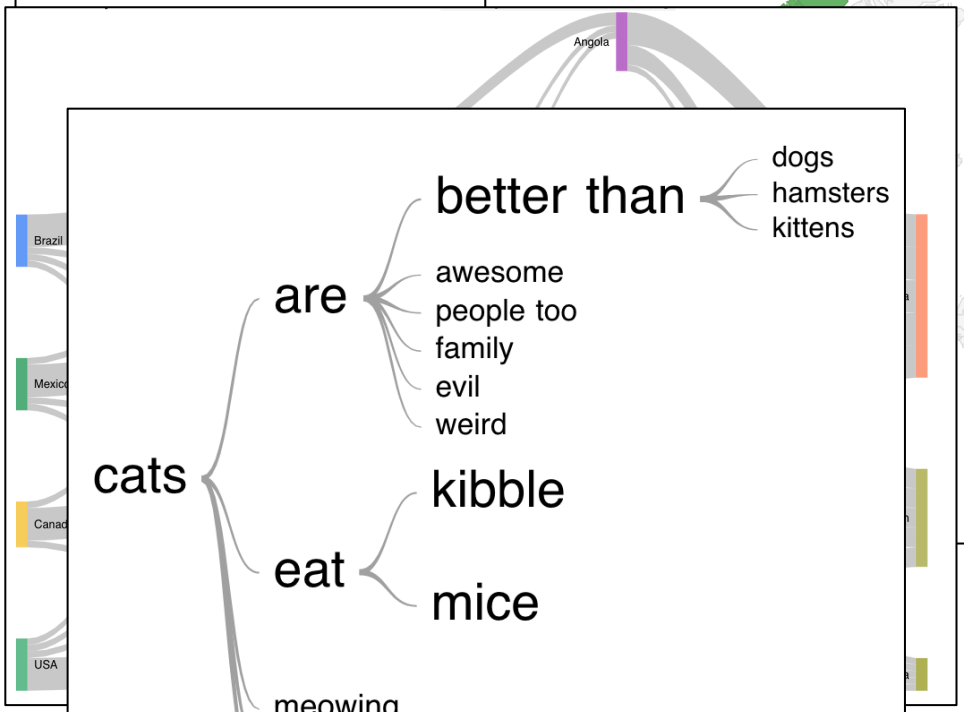


Where to get ideas for figures?

https://developers.google.com/chart with source code!

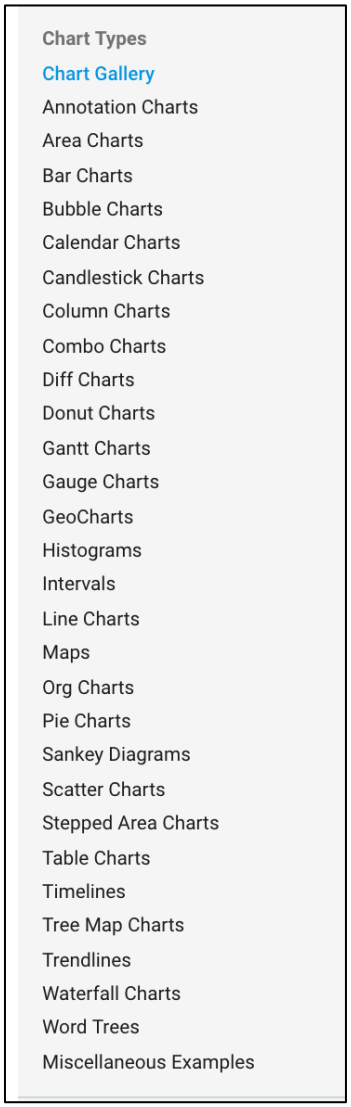


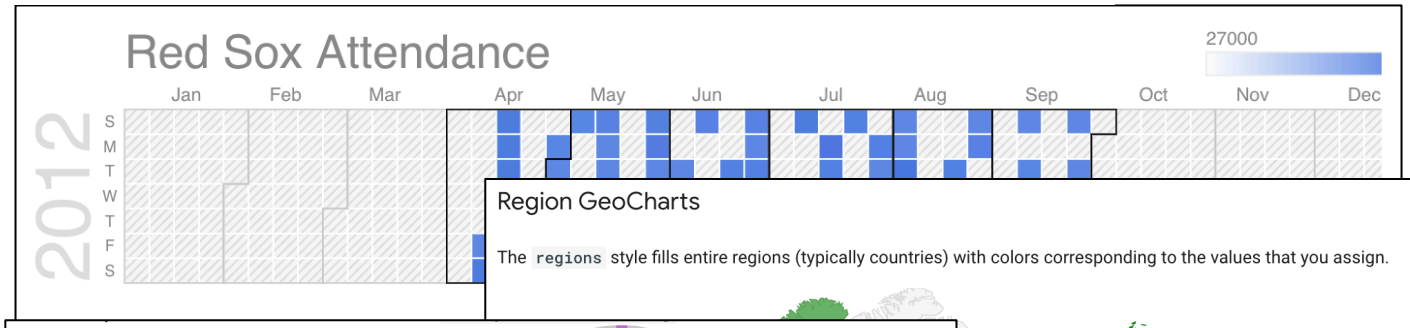


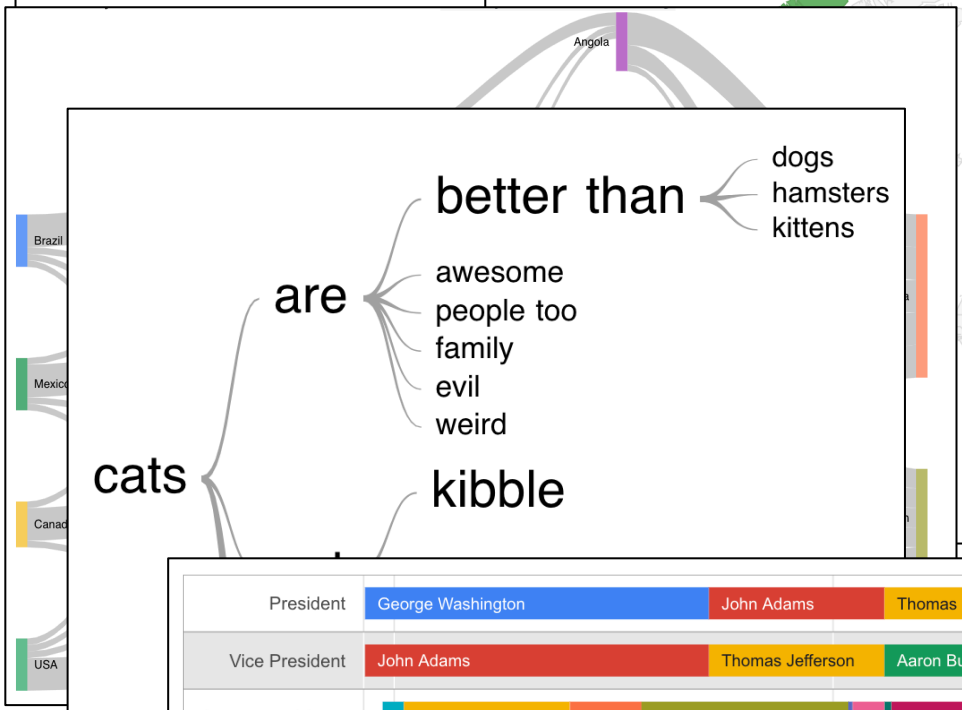


Where to get ideas for figures?

https://developers.google.com/chart with source code!



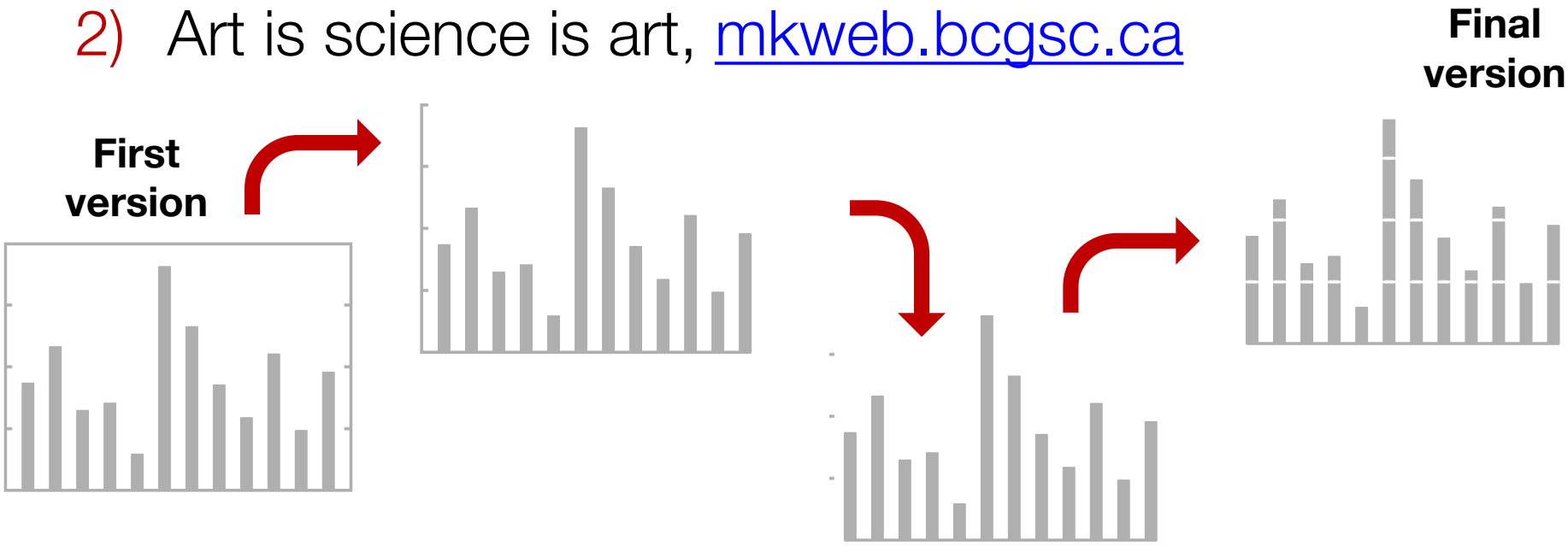




Guidelines

Guidelines #1

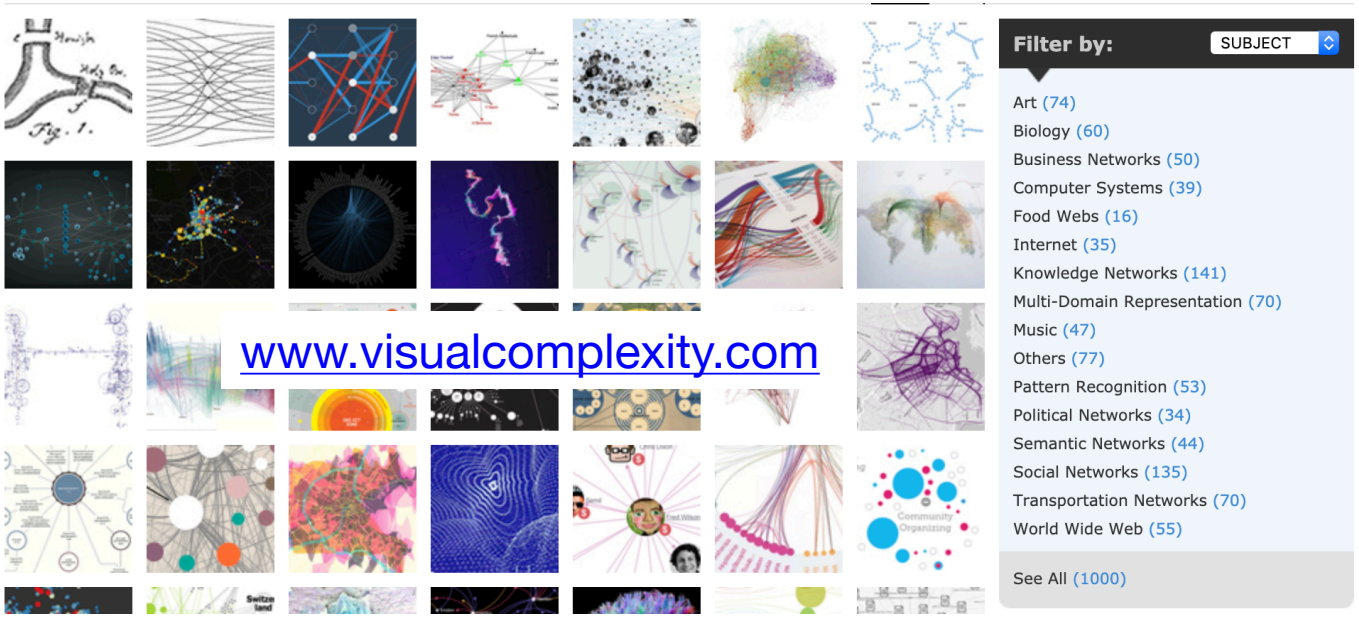
1) Tufte's design rules:
   - sealthreinhold.com/school/tuftes-rules
   - Data-to-ink-ratio: Maximize data-ink and erase as much non-data-ink as possible (avoid chart junk)
2) Art is science is art, mkweb.bcgsc.ca



Guidelines #2

3) Google's principles for designing charts:
   - material.io/design/communication/data-visualization.html
     - Principles: Be honest, Lend a helping hand, Delight users, Give clarity of focus, Embrace scale, Provide structure

4) Manuel Lima, Design Lead @ Google:



Today's Lecture

1) Why figures matter ✓

2) Figures in science ✓

3) How to design effective figures ✓

4) Tools, tips, and guidelines ✓

Three Takeaway Messages

1) Figures are often the first part of research papers examined by editors and your peers

2) Well-designed figures convey facts, ideas, and relationships far more clearly/concisely than text

3) Focus on effectively conveying complex information rather than on attention-getting decoration