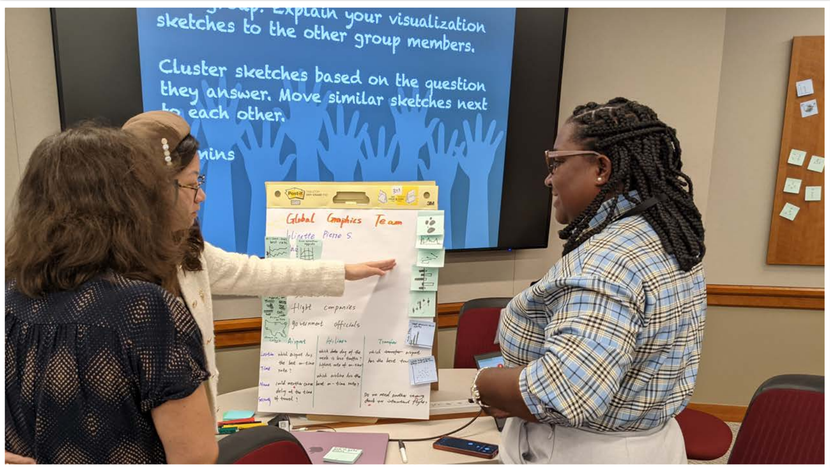
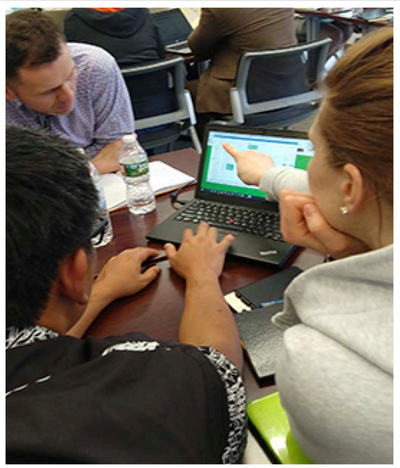
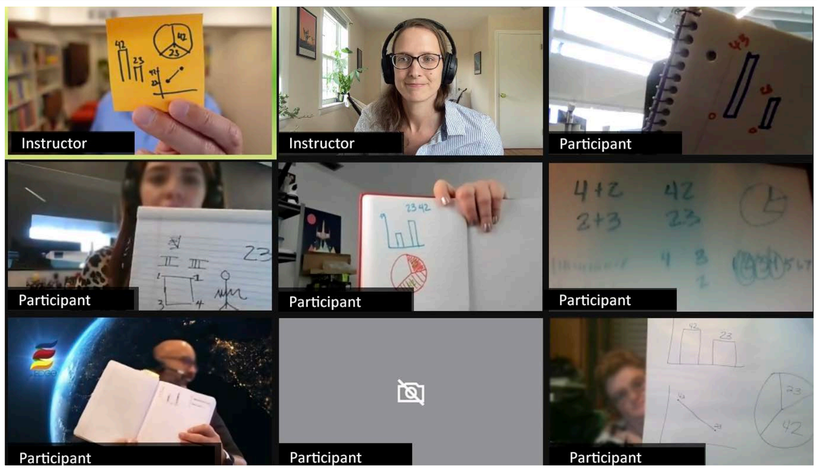
DESIGN SPRINTS FOR VISUALIZATION

CAROLINA NOBRE, JOHANNA BEYER







UNIVERSITY OF TORONTO

HARVARD School of Engineering and Applied Sciences

Carolina Nobre

Assistant Professor at the University of Toronto

Instructor for UofT Course

CSC2537 : Data Visualization

Johanna Beyer

Research Scientist at Harvard University

Co-Instructor for Harvard Course

CS 171: Data Visualization

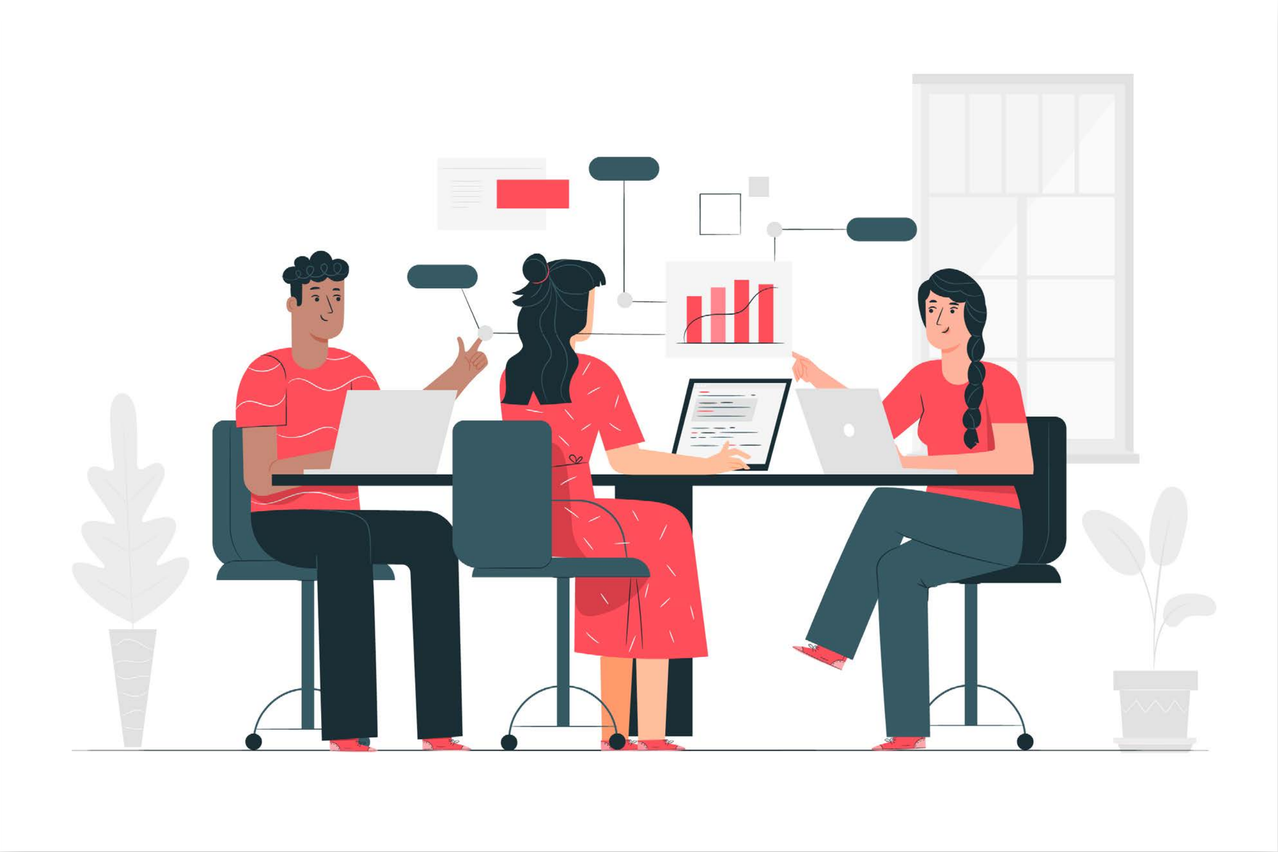
AGENDA

- Introduction to Design Sprints
- Hand on Design Sprint!
- Tips for Practitioners and Teachers
- Wrap-Up

PART I - INTRODUCTION TO VIS DESIGN SPRINTS

DESIGN SPRINTS

- Guided 5-Day Process
- Small Teams
- Quick build and test ideas



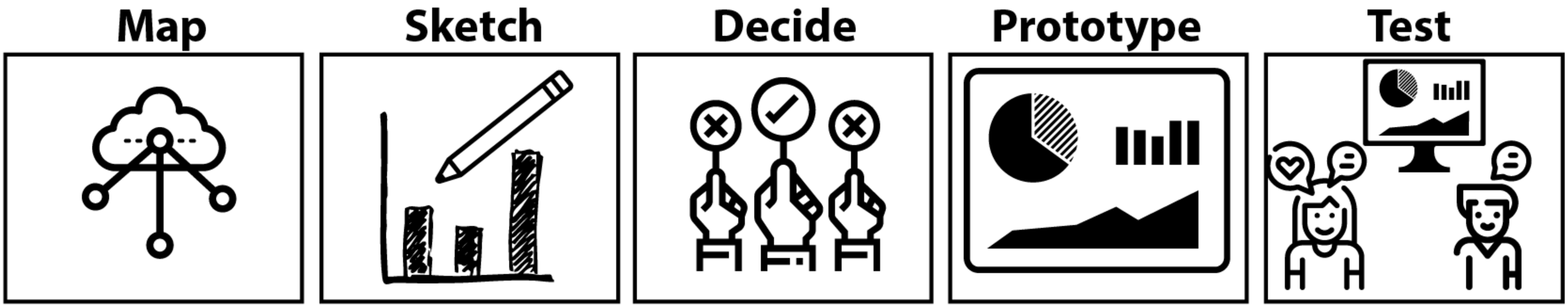
WHY VISUALIZATION DESIGN SPRINTS?

- Well-defined & time constrained
- Learner-Centered

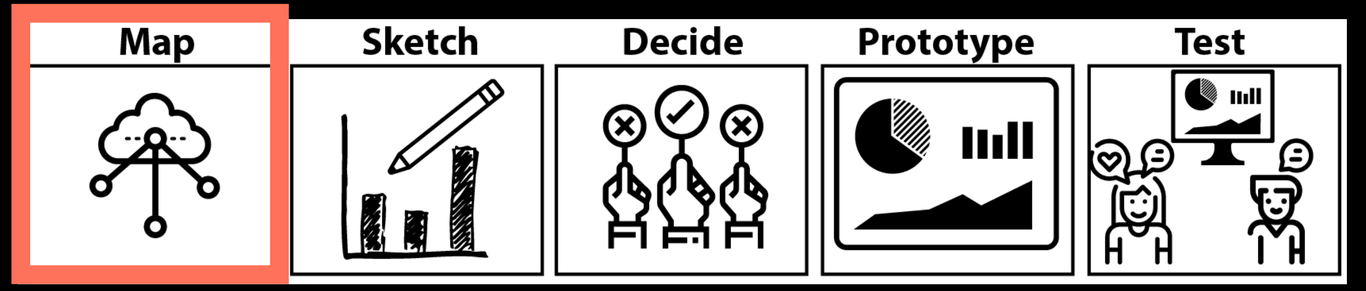
PRACTITIONERS VS TEACHING

- Design sprints can be adjusted for different settings
- We will point out specific tips.

5-STEP DESIGN SPRINT

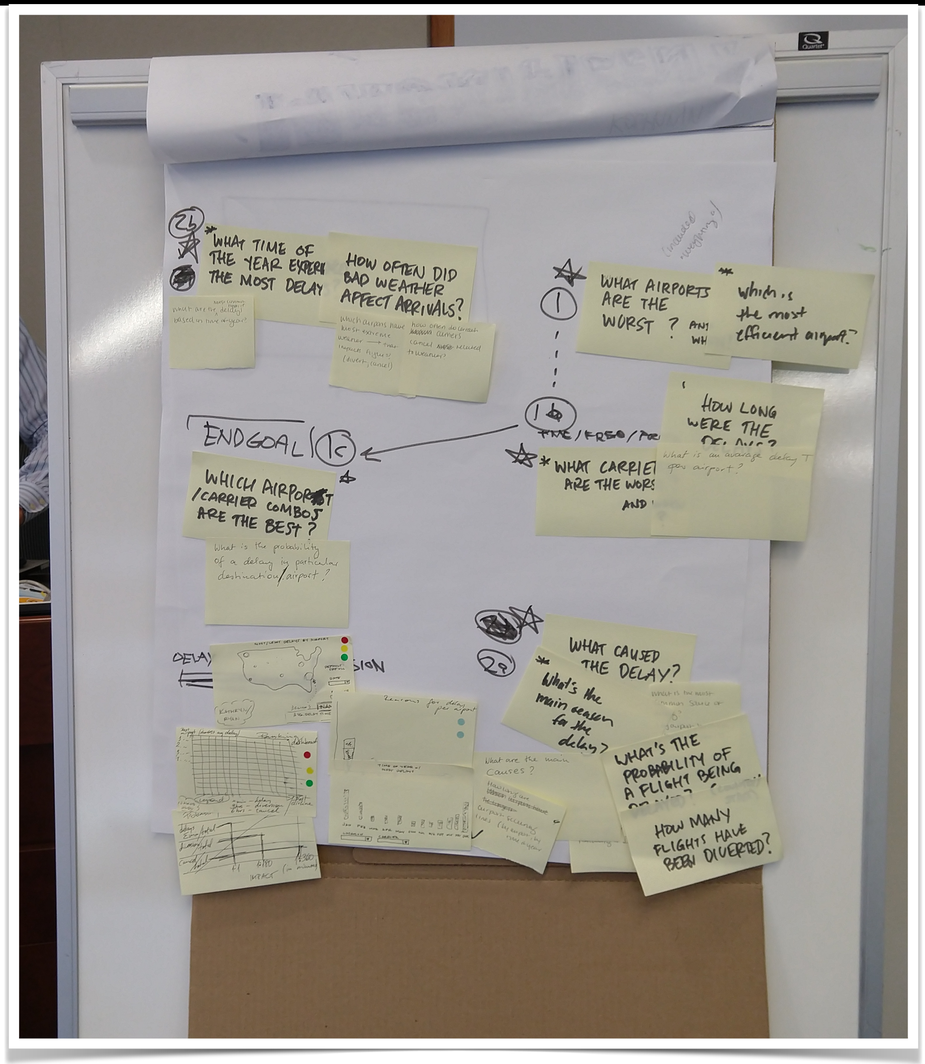


5-STEP DESIGN SPRINT

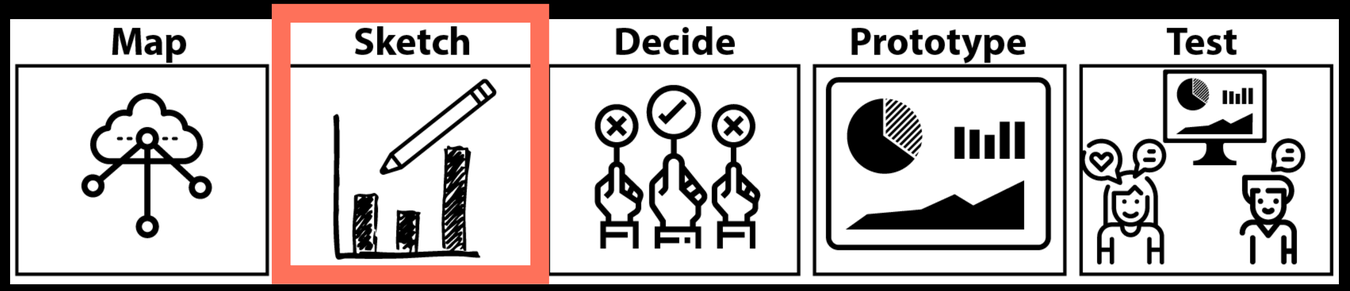


1. Map

- Data Types
- What is the data?
- Who is the audience?
- What is the question?

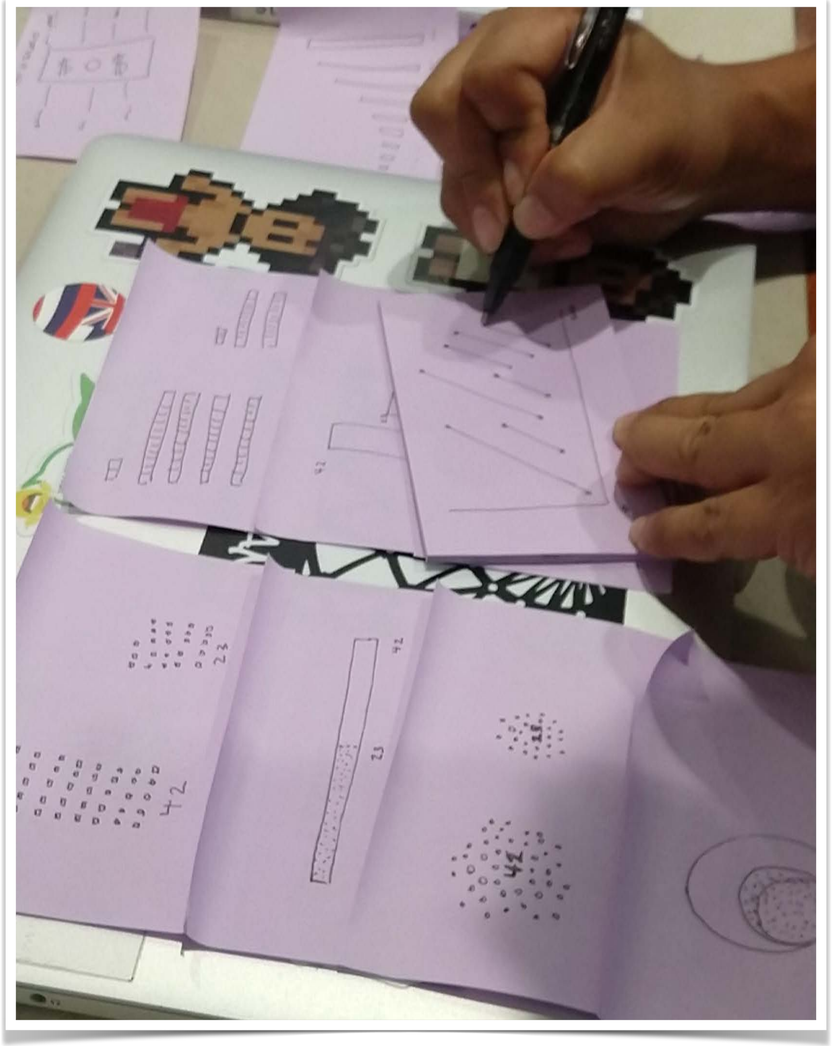


5-STEP DESIGN SPRINT

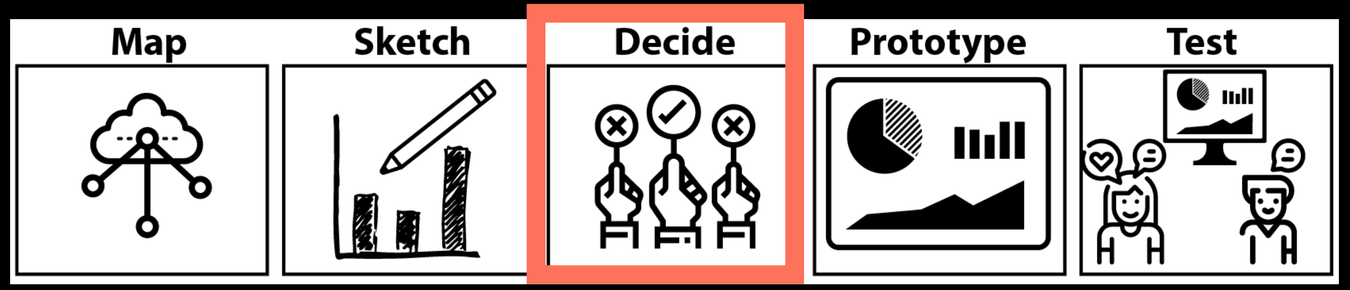


2. Sketch

- Pen & Paper
- Done individually

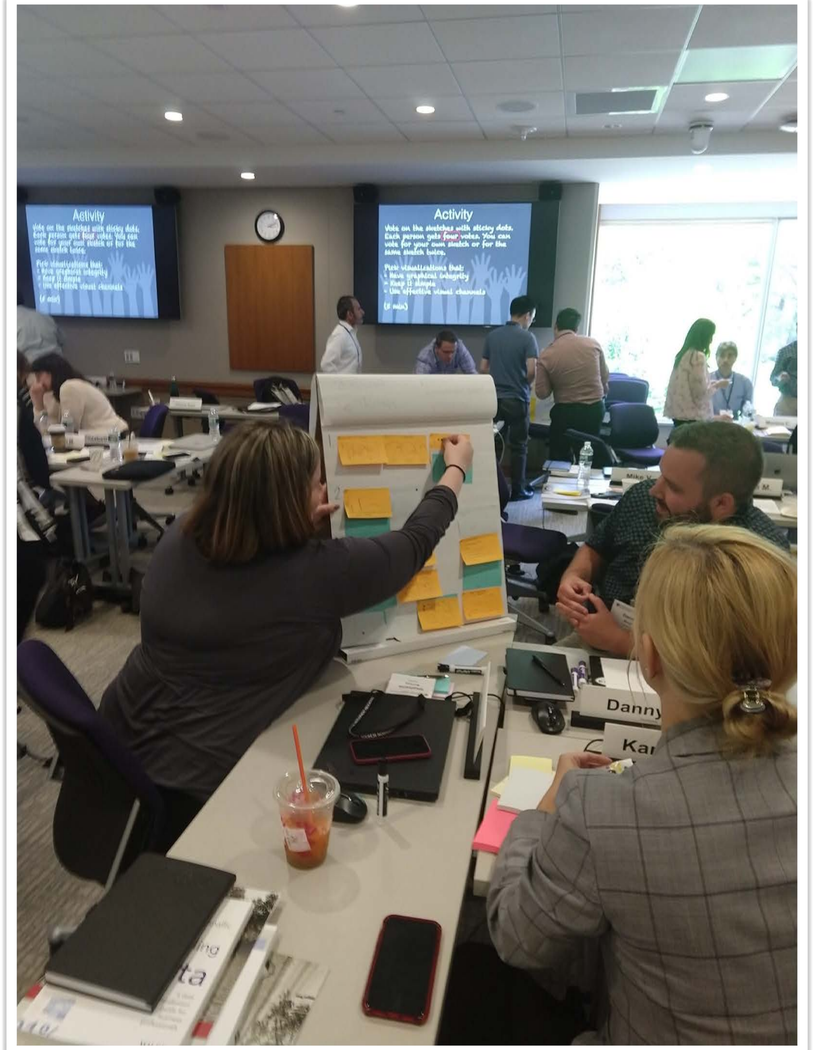


5-STEP DESIGN SPRINT

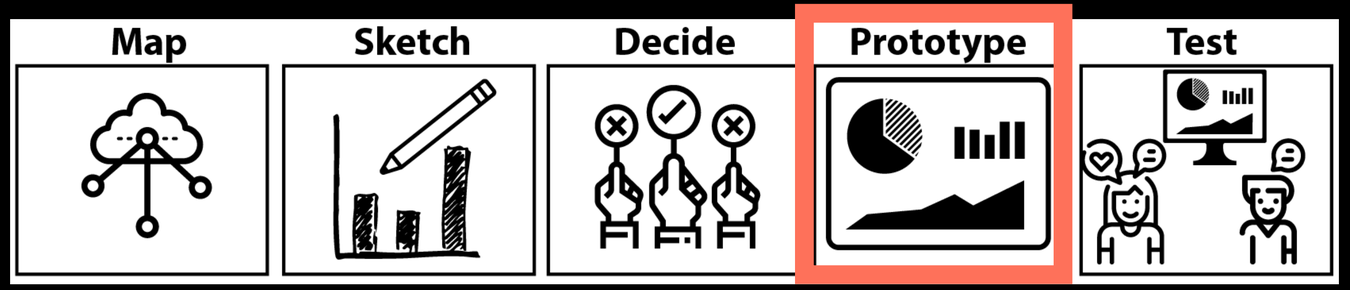


3. Decide

- Present Sketches
- Cluster & Group Sketches
- Discuss
- Vote for Best Designs

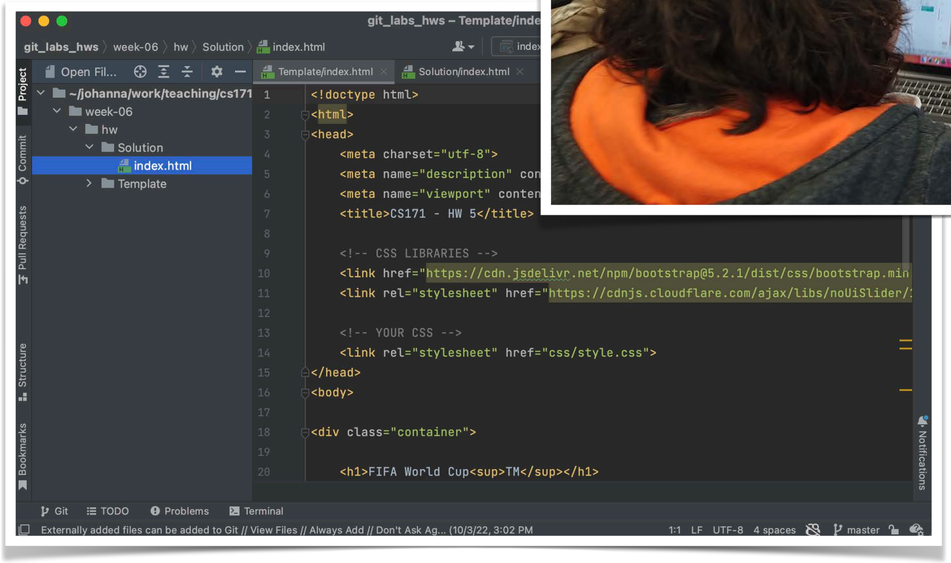


5-STEP DESIGN SPRINT

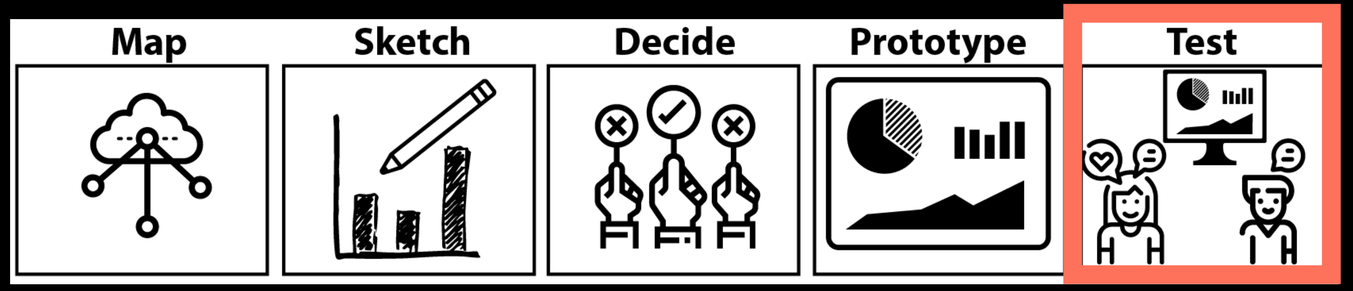


4. Prototype

- Implement ideas quickly



5-STEP DESIGN SPRINT



5. Test

- Qualitative user eval

PART II - HANDS ON DESIGN SPRINT

TEAMS & DATASET

FORM TEAMS : 2-3 MINUTES

FORM GROUPS OF 3-5 PEOPLE, PREFERABLY ONE TEAM PER TABLE.

TEAMS & DATASET

TEAM NAMES - 1 MINUTE

GIVE YOURSELVES A TEAM NAME. BE CREATIVE!

WRITE YOUR TEAM NAME ON THE TOP OF YOUR TEAM SHEET.

TEAMS & DATASET

PICK A DATASET - 5 MINUTES

Name: **Space Exploration Logs**

- **Attributes:**
  1. Mission ID
  2. Astronaut Names
  3. Launch Date
  4. Destination Planet/Moon
  5. Duration of Mission (in days)
  6. Scientific Discoveries
  7. Spacecraft Used
  8. Space Weather Conditions
  9. Communication Protocol
  10. Mission Objectives

Name: **Fantasy RPG Character Stats**

- **Attributes:**
  1. Character Name
  2. Race
  3. Class
  4. Level
  5. Health Points
  6. Attack Damage
  7. Armor Rating
  8. Magic Abilities
  9. Inventory Items
  10. Quest Achievements

Name: **Wildlife Migration Patterns**

- **Attributes:**
  1. Species Name
  2. Tracking ID
  3. Migration Start Date
  4. Migration End Date
  5. Distance Traveled (in miles)
  6. Environmental Factors (e.g., temperature)
  7. Breeding Ground Location
  8. Migration Route Map
  9. Predation Incidents
  10. Conservation Status

PICK ONE OF THE DATASETS AT YOUR TABLE

MAP

20 MINUTES

SKETCH DECIDE PROTOTYPE TEST

ANSWER THE FOLLOWING QUESTIONS

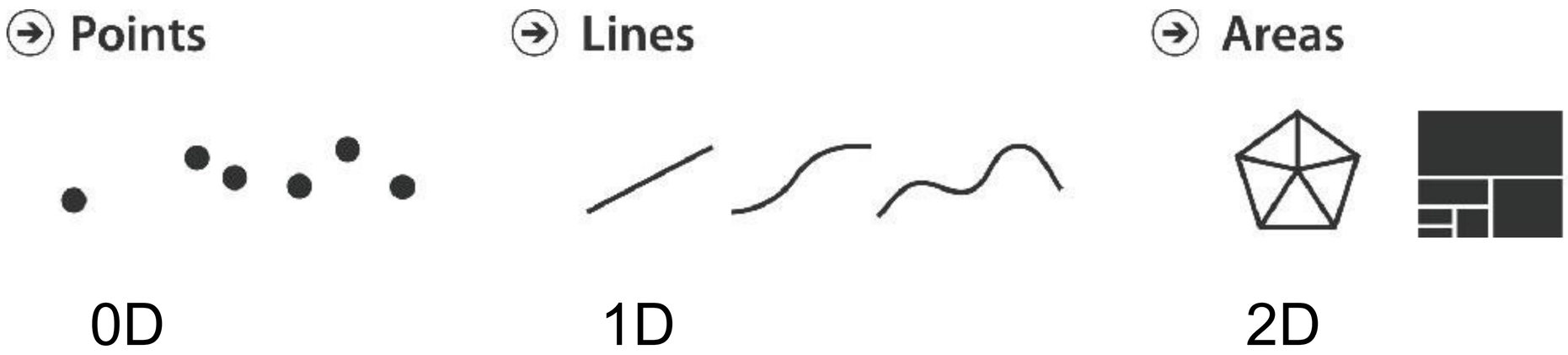
- What is in the data?
- Who is the audience?
- What is the question your data will answer?

USE STICKY NOTES!

SKETCH

DECIDE PROTOTYPE TEST

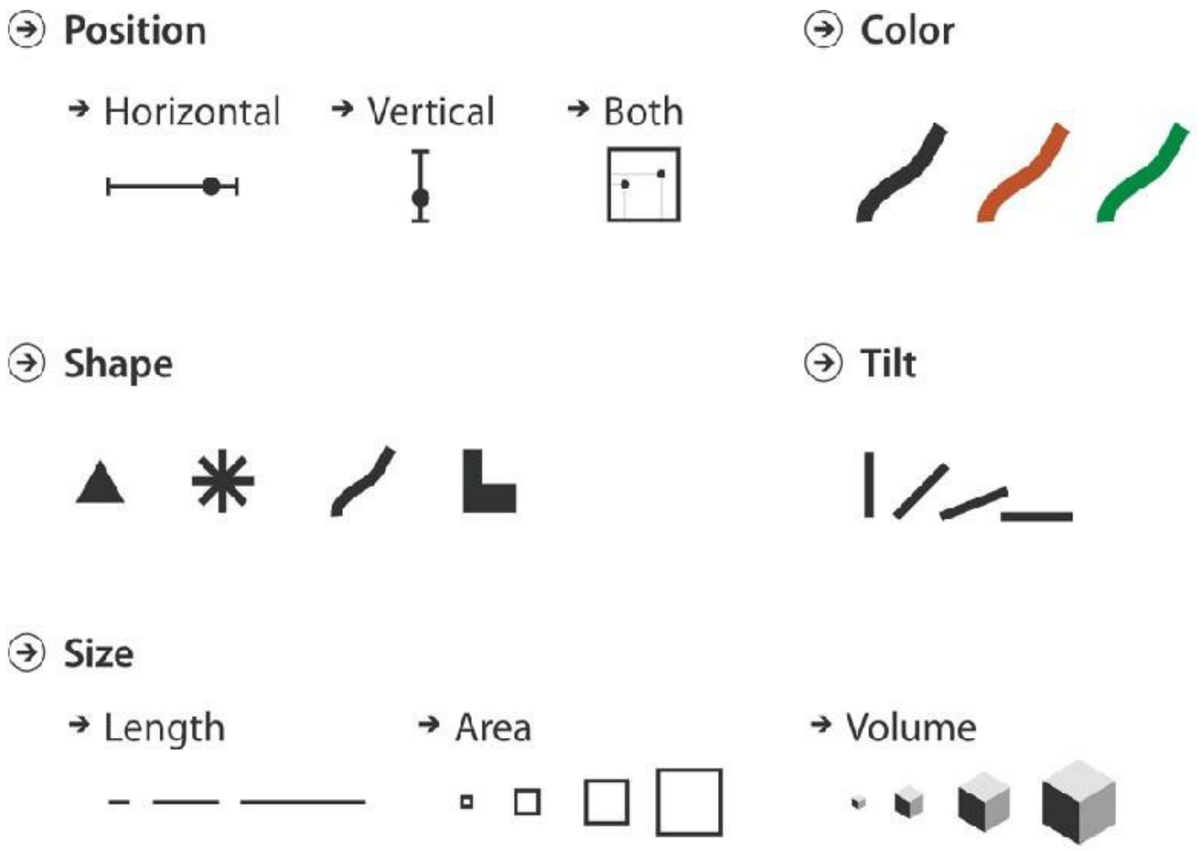
MARKS AND CHANNELS



SKETCH

DECIDE PROTOTYPE TEST

MARKS AND **CHANNELS**



Types of Channels

Magnitude Channels

How much?

Channels

Position

Length

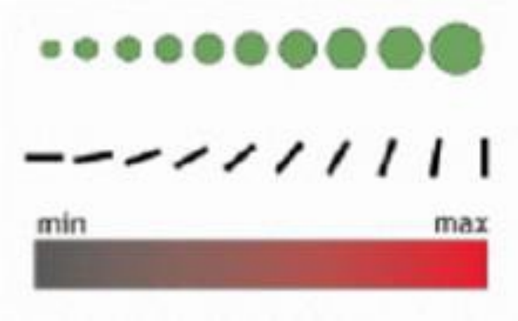
Saturation

Ordinal Data

- Obey a < relationship
- Small, medium, large

Quantitative Data

- Can do arithmetic on them
- 10 cm, 23 F, etc.



Types of Channels

Identity Channels

What? Where?

Channels

Shape

Color (Hue)

Spatial

Categorical Data

- Are = or ≠ to other values
- Apples, Oranges, Bananas,...

Channels: Expressiveness Types and Effectiveness Ranks

Magnitude Channels: Ordered Attributes

- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

Most — Effectiveness — Least

Same: Color luminance, Color saturation

Same: Curvature, Volume (3D size)

Identity Channels: Categorical Attributes

- Spatial region
- Color hue
- Motion
- Shape

Visual vocabulary

Designing with data

There are so many ways to visualise data - how do we know which one to pick? Use the categories across the top to decide which data relationship is most important in your story, then look at the different types of chart within the category to form some initial ideas about what might work best. This list is not meant to be exhaustive, nor a wizard, but is a useful starting point for making informative and meaningful data visualisations.

FT graphic: Alan Smith; Chris Campbell; Ian Bott; Liz Faunce; Graham Parrish; Billy Ehrenberg-Shannon; Paul McCallum; Martin Stabe
Inspired by the Graphic Continuum by Jon Schwabish and Severino Ribecca

ft.com/vocabulary

Deviation

Emphasise variations (+/-) from a fixed reference point. Typically the reference point is zero but it can also be a target or a long-term average. Can also be used to show sentiment (positive/neutral/negative).

Example FT uses
Trade surplus/deficit, climate change

Diverging bar
A simple standard bar chart that can handle both negative and positive magnitude values.

Diverging stacked bar
Perfect for presenting survey results which involve sentiment (eg disagree/neutral/agree).

Spine
Splits a single value into two contrasting components (eg male/female).

Surplus/deficit filled line
The shaded area of these charts allows a balance to be shown – either against a baseline or between two series.

Correlation

Show the relationship between two or more variables. Be mindful that, unless you tell them otherwise, many readers will assume the relationships you show them to be causal (i.e. one causes the other).

Example FT uses
Inflation and unemployment, income and life expectancy

Scatterplot
The standard way to show the relationship between two continuous variables, each of which has its own axis.

Column + line timeline
A good way of showing the relationship between an amount (columns) and a rate (line).

Connected scatterplot
Usually used to show how the relationship between 2 variables has changed over time.

Bubble
Like a scatterplot, but adds additional detail by sizing the circles according to a third variable.

XY heatmap
A good way of showing the patterns between 2 categories of data, less effective at showing fine differences in amounts.

Ranking

Use where an item's position in an ordered list is more important than its absolute or relative value. Don't be afraid to highlight the points of interest.

Example FT uses
Wealth, deprivation, league tables, constituency election results

Ordered bar
Standard bar charts display the ranks of values much more easily when sorted into order.

Ordered column
See above.

Ordered proportional symbol
Use when there are big variations between values and/or seeing fine differences between data is not so important.

Dot strip plot
Dots placed in order on a strip are a space-efficient method of laying out ranks across multiple categories.

Slope
Perfect for showing how ranks have changed over time or vary between categories.

Lollipop
Lollipops draw more attention to the data value than standard bar/column and can also show rank and value effectively.

Bump
Effective for showing changing rankings across multiple dates. For large datasets, consider grouping lines using colour.

Distribution

Show values in a dataset and how often they occur. The shape (or 'skew') of a distribution can be a memorable way of highlighting the lack of uniformity or equality in the data.

Example FT uses
Income distribution, population (age/sex) distribution, revealing inequality

Histogram
The standard way to show a statistical distribution - keep the gaps between columns small to highlight the 'shape' of the data.

Dot plot
A simple way of showing the change or range (min/max) of data across multiple categories.

Dot strip plot
Good for showing individual values in a distribution, can be a problem when too many dots have the same value.

Barcode plot
Like dot strip plots, good for displaying all the data in a table, they work best when highlighting individual values.

Boxplot
Summarise multiple distributions by showing the median (centre) and range of the data

Violin plot
Similar to a box plot but more effective with complex distributions (data that cannot be summarised with simple average).

Population pyramid
A standard way for showing the age and sex breakdown of a population distribution; effectively, back to back histograms.

Cumulative curve
A good way of showing how unequal a distribution is: y axis is always cumulative frequency, x axis is always a measure.

Frequency polygons
For displaying multiple distributions of data. Like a regular line chart, best limited to a maximum of 3 or 4 datasets.

Beeswarm
Use to emphasise individual points in a distribution. Points can be sized to an additional variable. Best with medium-sized datasets

Change over Time

Give emphasis to changing trends. These can be short (intra-day) movements or extended series traversing decades or centuries: Choosing the correct time period is important to provide suitable context for the reader.

Example FT uses
Share price movements, economic time series, sectoral changes in a market

Line
The standard way to show a changing time series. If data are irregular, consider markers to represent data points.

Column
Columns work well for showing change over time - but usually best with only one series of data at a time.

Column + line timeline
A good way of showing the relationship over time between an amount (columns) and a rate (line).

Slope
Good for showing changing data as long as the data can be simplified into 2 or 3 points without missing a key part of story.

Area chart
Use with care – these are good at showing changes to total, but seeing change in components can be very difficult.

Candlestick
Usually focused on day-to-day activity, these charts show opening/closing and high/low points of each day.

Fan chart (projections)
Use to show the uncertainty in future projections - usually this grows the further forward to projection.

Connected scatterplot
A good way of showing changing data for two variables whenever there is a relatively clear pattern of progression.

Calendar heatmap
A great way of showing temporal patterns (daily, weekly, monthly) – at the expense of showing precision in quantity.

Priestley timeline
Great when date and duration are key elements of the story in the data.

Circle timeline
Good for showing discrete values of varying size across multiple categories (eg earthquakes by continent).

Vertical timeline
Presents time on the Y axis. Good for displaying detailed time series that work especially well when scrolling on mobile.

Seismogram
Another alternative to the circle timeline for showing series where there are big variations in the data.

Streamgraph
A type of area chart; use when seeing changes in proportions over time is more important than individual values

Magnitude

Show size comparisons. These can be relative (just being able to see larger/bigger) or absolute (need to see fine differences). Usually these show a 'counted' number (for example, barrels, dollars or people) rather than a calculated rate or per cent.

Example FT uses
Commodity production, market capitalisation, volumes in general

Column
The standard way to compare the size of things. Must always start at 0 on the axis.

Bar
See above. Good when the data are not time series and labels have long category names.

Paired column
As per standard column but allows for multiple series. Can become tricky to read with more than 2 series.

Paired bar
See above.

Marimekko
A good way of showing the size and proportion of data at the same time – as long as the data are not too complicated.

Proportional symbol
Use when there are big variations between values and/or seeing fine differences between data is not so important.

Isotype (pictogram)
Excellent solution in some instances – use only with whole numbers (do not slice off an arm to represent a decimal).

Lollipop
Lollipop charts draw more attention to the data value than standard bar/column – does not have to start at zero (but preferable).

Radar
A space-efficient way of showing value of multiple variables– but make sure they are organised in a way that makes sense to reader.

Parallel coordinates
An alternative to radar charts – again, the arrangement of the variables is important. Usually benefits from highlighting values.

Bullet
Good for showing a measurement against the context of a target or performance range.

Grouped symbol
An alternative to bar/column charts when being able to count data or highlight individual elements is useful.

Part-to-whole

Show how a single entity can be broken down into its component elements. If the reader's interest is solely in the size of the components, consider a magnitude-type chart instead.

Example FT uses
Fiscal budgets, company structures, national election results

Stacked column/bar
A simple way of showing part-to-whole relationships but can be difficult to read with more than a few components.

Marimekko
A good way of showing the size and proportion of data at the same time – as long as the data are not too complicated.

Pie
A common way of showing part-to-whole data – but be aware that it's difficult to accurately compare the size of the segments.

Donut
Similar to a pie chart – but the centre can be a good way of making space to include more information about the data (eg total).

Treemap
Use for hierarchical part-to-whole relationships; can be difficult to read when there are many small segments.

Voronoi
A way of turning points into areas – any point within each area is closer to the central point than any other centroid.

Arc
A hemicycle, often used for visualising parliamentary composition by number of seats.

Gridplot
Good for showing % information, they work best when used on whole numbers and work well in small multiple layout form.

Venn
Generally only used for schematic representation.

Waterfall
Can be useful for showing part-to-whole relationships where some of the components are negative.

Spatial

Aside from locator maps only used when precise locations or geographical patterns in data are more important to the reader than anything else.

Example FT uses
Population density, natural resource locations, natural disaster risk/impact, catchment areas, variation in election results

Basic choropleth (rate/ratio)
The standard approach for putting data on a map – should always be rates rather than totals and use a sensible base geography.

Proportional symbol (count/magnitude)
Use for totals rather than rates – be wary that small differences in data will be hard to see.

Flow map
For showing unambiguous movement across a map.

Contour map
For showing areas of equal value on a map. Can use deviation colour schemes for showing +/- values

Equalised cartogram
Converting each unit on a map to a regular and equally-sized shape – good for representing voting regions with equal value.

Scaled cartogram (value)
Stretching and shrinking a map so that each area is sized according to a particular value.

Dot density
Used to show the location of individual events/locations – make sure to annotate any patterns the reader should see.

Heat map
Grid-based data values mapped with an intensity colour scale. As choropleth map – but not snapped to an admin/political unit.

Flow

Show the reader volumes or intensity of movement between two or more states or conditions. These might be logical sequences or geographical locations.

Example FT uses
Movement of funds, trade, migrants, lawsuits, information; relationship graphs.

Sankey
Shows changes in flows from one condition to at least one other; good for tracing the eventual outcome of a complex process.

Waterfall
Designed to show the sequencing of data through a flow process, typically budgets. Can include +/- components.

Chord
A complex but powerful diagram which can illustrate 2-way flows (and net winner) in a matrix.

Network
Used for showing the strength and inter-connectedness of relationships of varying types.

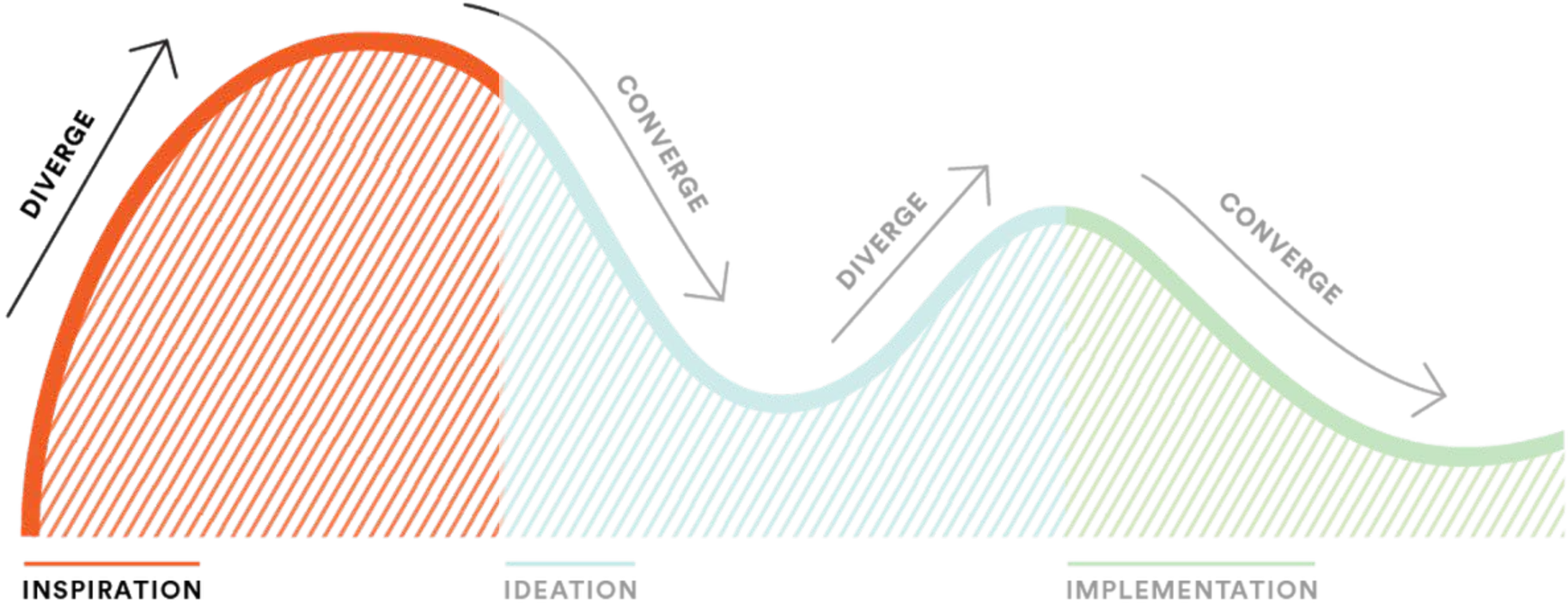
FT

© Financial Times 2016-2019
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

SKETCH

DECIDE PROTOTYPE TEST

20 MINUTES



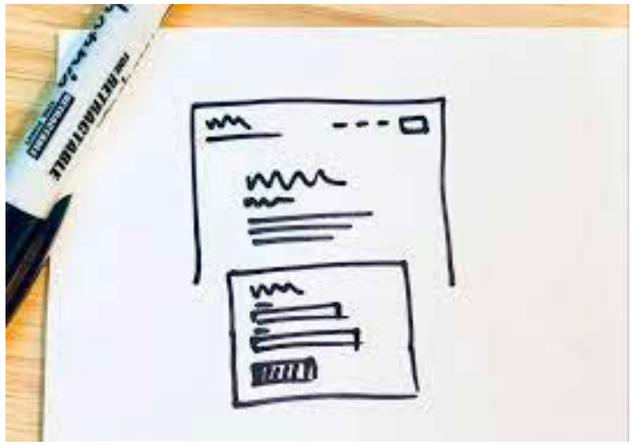
Human-centered design process by IDEO. Credit: IDEO

SKETCH

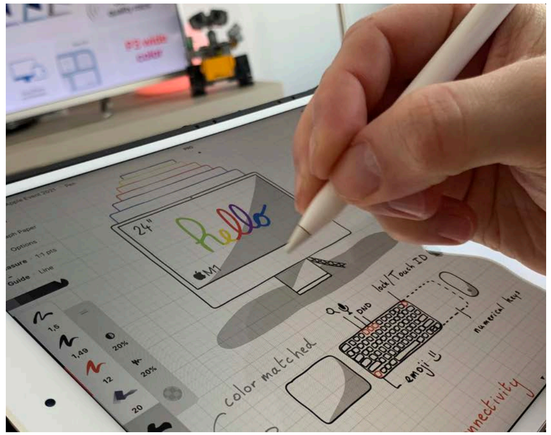
DECIDE PROTOTYPE TEST

20 MINUTES

SKETCH SEVERAL OPTIONS (INDIVIDUALLY) WITH THE MEDIUM OF YOUR CHOICE!



PEN + PAPER



DIGITAL SKETCH / MIRO BOARD

LEGOS

SKETCH

DECIDE PROTOTYPE TEST

UPLOAD SKETCHES : 5 MINUTES

TAKE A PHOTO AND UPLOAD YOUR SKETCHES TO YOUR GROUP BOARD ON MIRO

SKETCH

DECIDE PROTOTYPE TEST

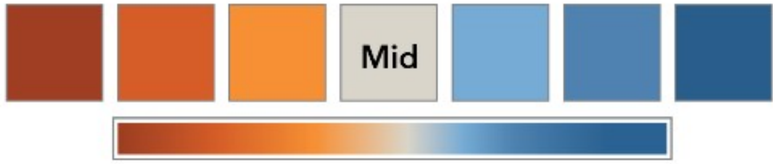
UPLOAD SKETCHES : 5 MINUTES

SHOW AND TELL

The Use of Color in Data Visualization

Sequential
color is ordered from low to high

Diverging
two sequential colors with a neutral midpoint



Categorical
contrasting colors for individual comparison

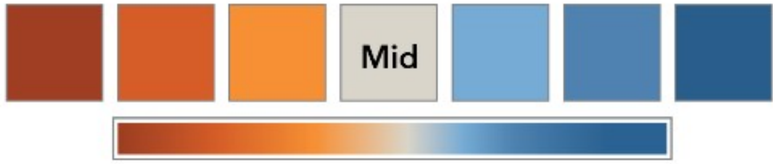
Ordinal Data
Objective:
Distinction + Sense of order / Magnitude

Source: *The Big Book of Dashboards* (Figure 1.16)

The Use of Color in Data Visualization

Sequential
color is ordered from low to high

Diverging
two sequential colors with a neutral midpoint



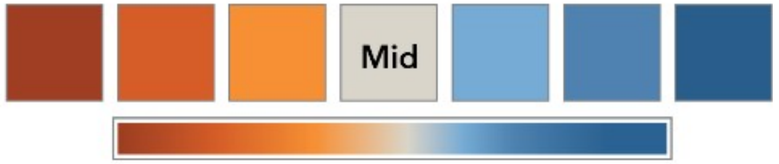
Categorical
contrasting colors for individual comparison

Ordinal Data
With a midpoint

The Use of Color in Data Visualization

Sequential
color is ordered from low to high

Diverging
two sequential colors with a neutral midpoint



Categorical
contrasting colors for individual comparison

Nominal Data
Objective: Contrast

Source: *The Big Book of Dashboards* (Figure 1.16)

The Use of Color in Data Visualization

Data Legibility

Sequential
color is ordered from low to high

Diverging
two sequential colors with a neutral midpoint

Categorical
contrasting colors for individual comparison

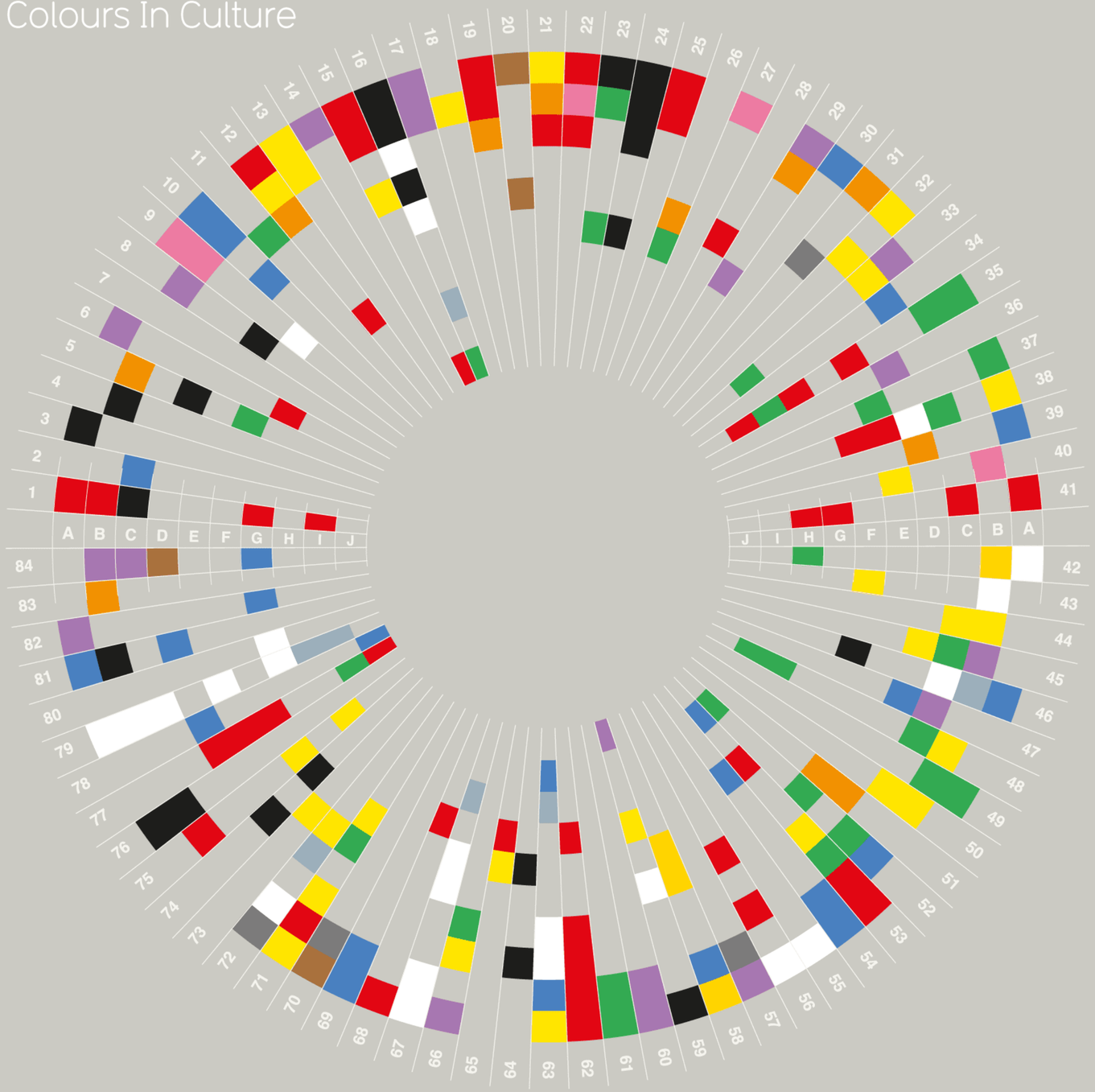
Editorial Salience

Highlight
color used to highlight something

Alert
color used to get reader's attention

Source: *The Big Book of Dashboards* (Figure 1.16)

Colours In Culture



A Western / American
B Japanese
C Hindu
D Native American
E Chinese
F Asian
G Eastern European
H Arab
I African
J South American

1 Anger
2 Art / Creativity
3 Authority
4 Bad Luck
5 Balance
6 Beauty
7 Calm
8 Celebration
9 Children
10 Cold
11 Compassion
12 Courage
13 Cowardice
14 Cruelty
15 Danger
16 Death
17 Decadence
18 Deceit
19 Desire
20 Earthy
21 Energy
22 Erotic
23 Eternity
24 Evil
25 Excitement
26 Family
27 Femininity
28 Fertility
29 Flamboyance
30 Freedom
31 Friendly
32 Fun
33 God
34 Gods
35 Good Luck
36 Gratitude

37 Growth
38 Happiness
39 Healing
40 Healthy
41 Heat
42 Heaven
43 Holiness
44 Illness
45 Insight
46 Intelligence
47 Intuition
48 Religion
49 Jealousy
50 Joy
51 Learning
52 Life
53 Love
54 Loyalty
55 Luxury
56 Marriage
57 Modesty
58 Money
59 Mourning
60 Mystery
61 Nature
62 Passion
63 Peace
64 Penance
65 Power
66 Personal power
67 Purity
68 Radicalism
69 Rational
70 Reliable
71 Repels Evil
72 Respect
73 Royalty
74 Self-cultivation
75 Strength
76 Style
77 Success
78 Trouble
79 Truce
80 Trust
81 Unhappiness
82 Virtue
83 Warmth
84 Wisdom

Yellow
Grey
Gold
Silver

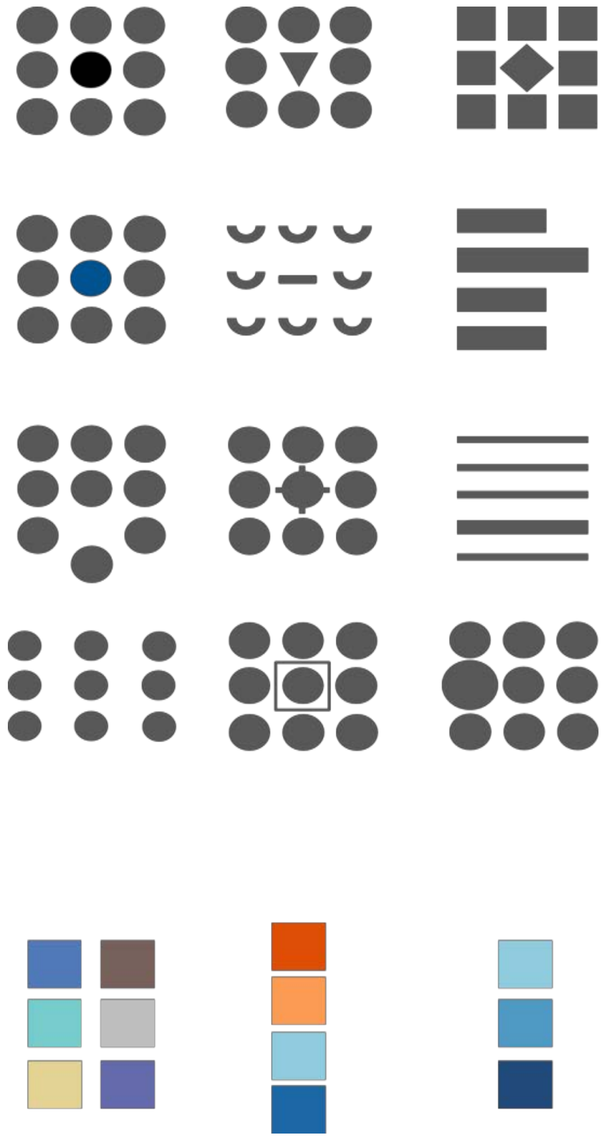
source: Pantone, ColorMatters & web sources

PRE-ATTENTIVE ATTRIBUTES

System 1: Pre-Attentive	System 2: Attentive
Parallel processing; multiple visual elements perceived simultaneously	*Serial processing*
Quick; Operates automatically, with little or no effort and no sense of voluntary control	*Slow;* allocates attention to the effortful mental activities that demand it, including complex computations
Unconscious; generates impressions, intuitions, intentions, and feelings	*Conscious;* engaged as needed to solve more complex problems

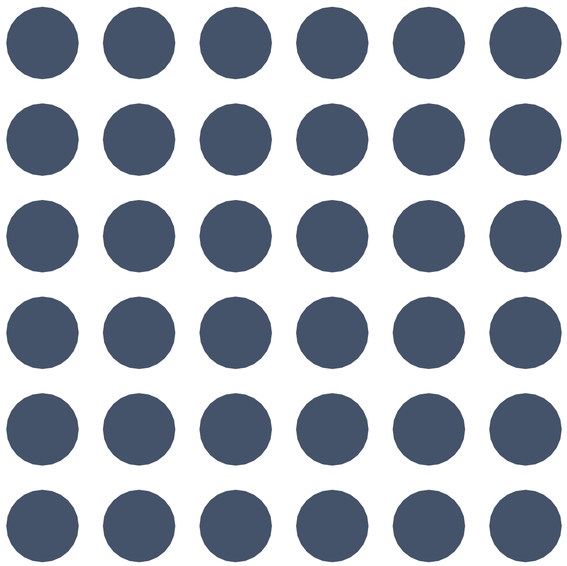
Pre-Attentive Attributes

- Color
  - Hue
  - Value/Intensity/Lightness
- Size
- Shape
- Length & Width
- Texture
- Elongation
- Orientation
- Spatial Grouping
- Order
- Spatial Position
- Enclosure
- Convex/Concave

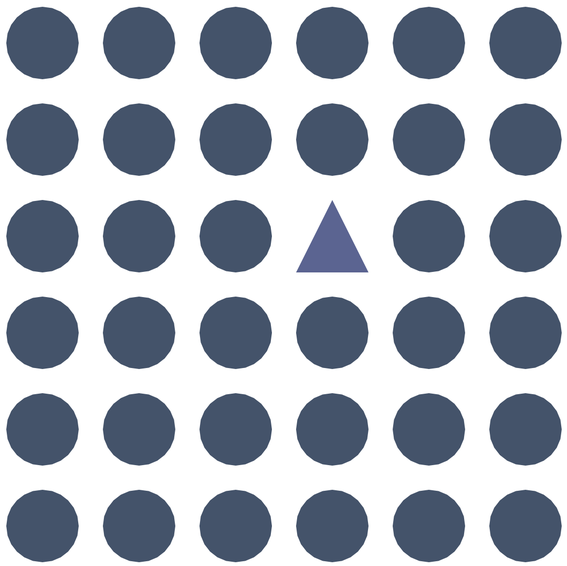


Ware, 2013

Shape



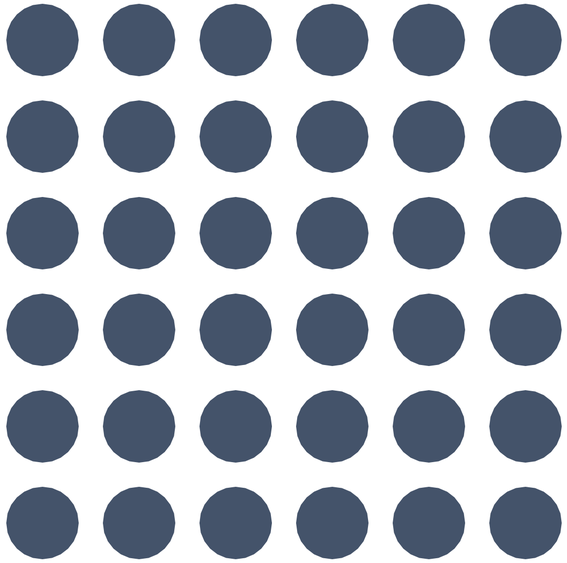
Shape



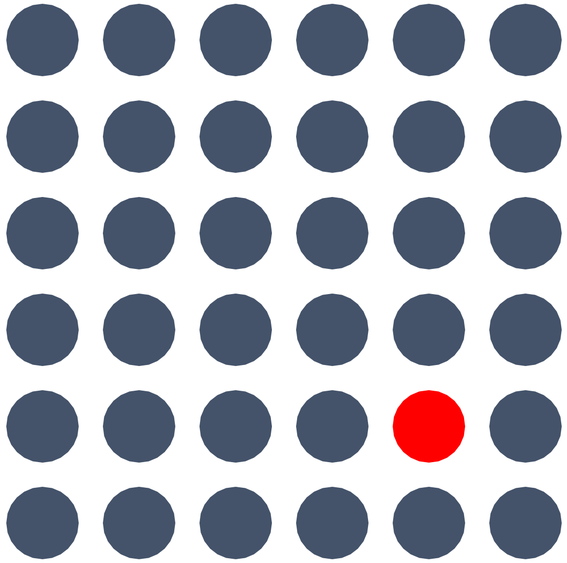
Orientation

Orientation

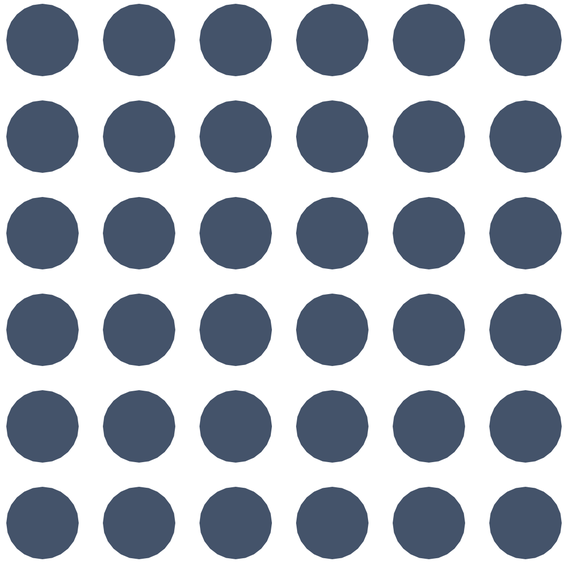
Color (Hue)



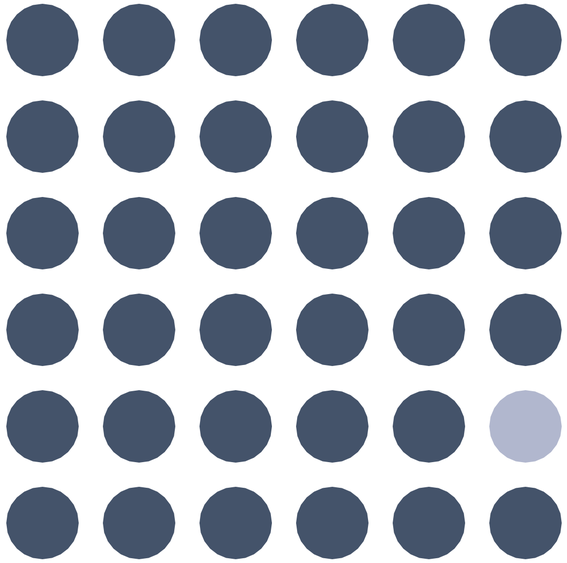
Color (Hue)



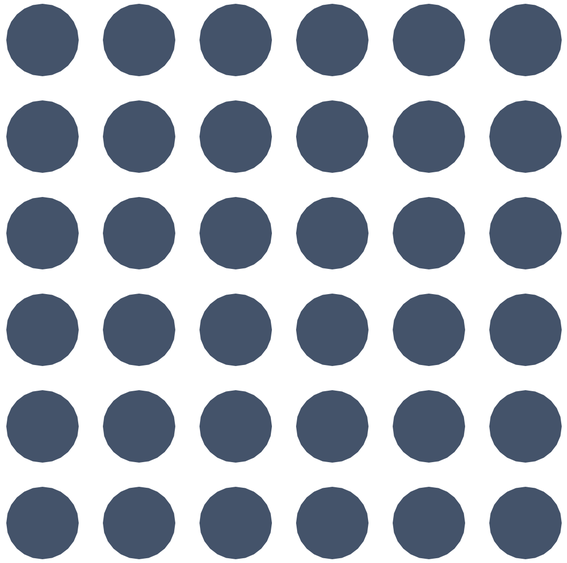
Intensity



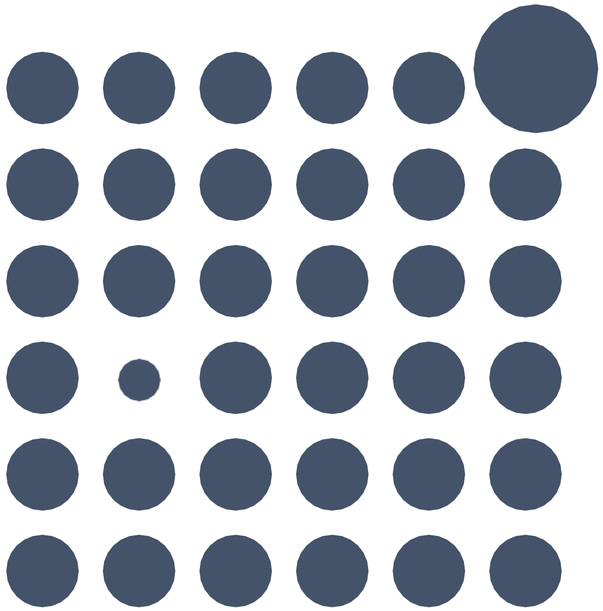
Intensity



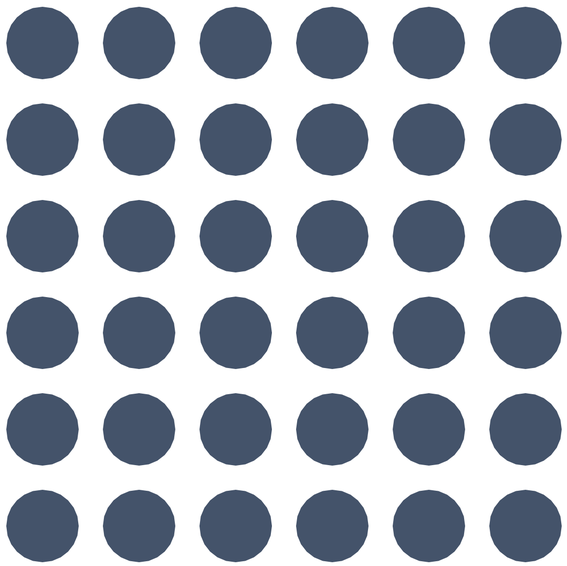
Size



Size

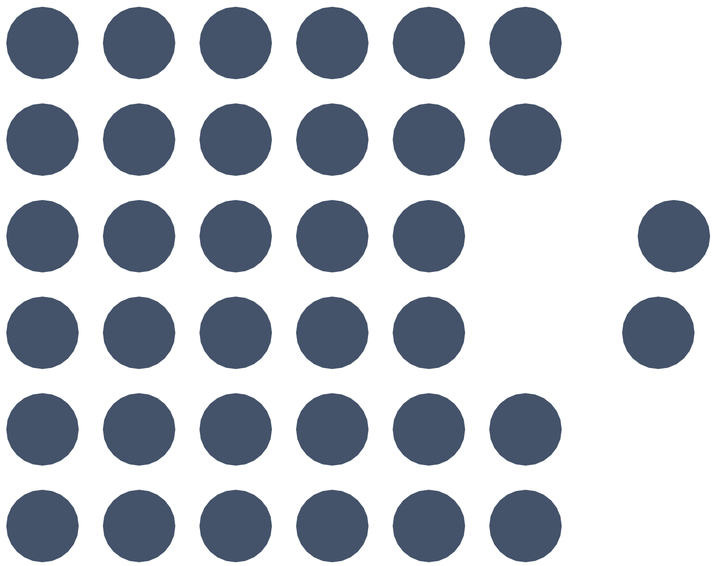


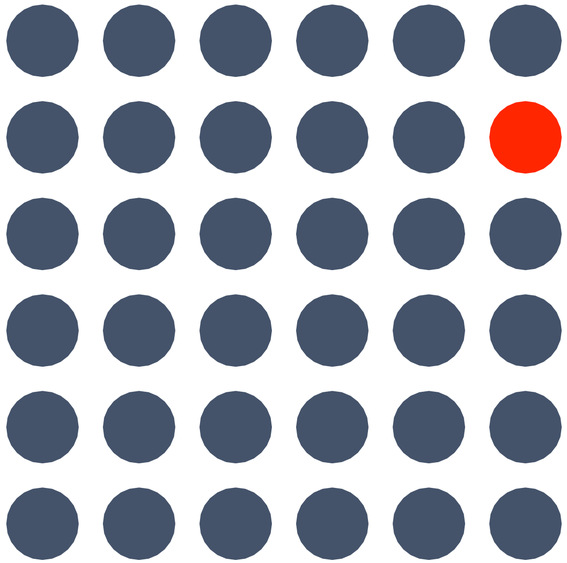
Position

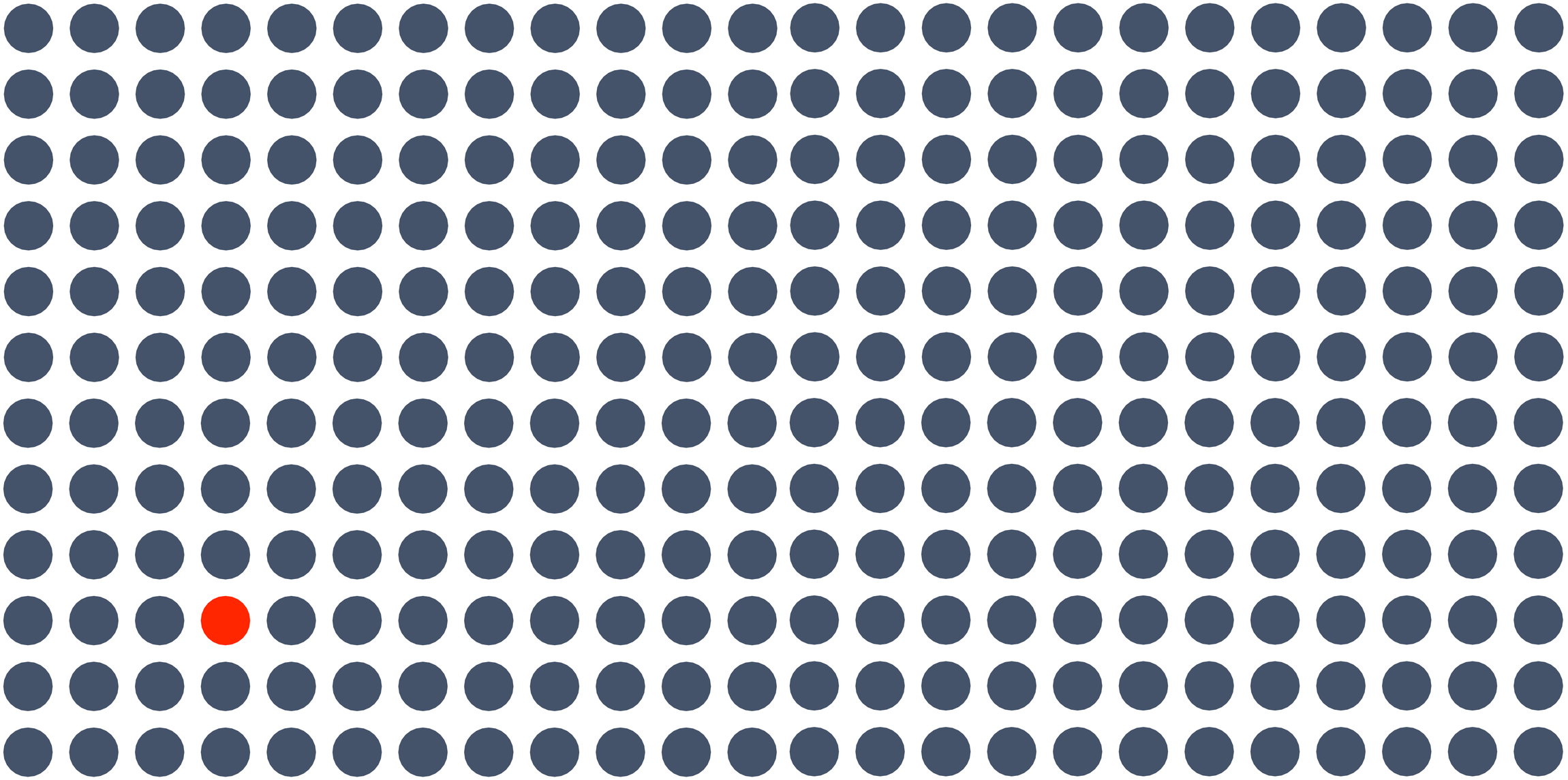


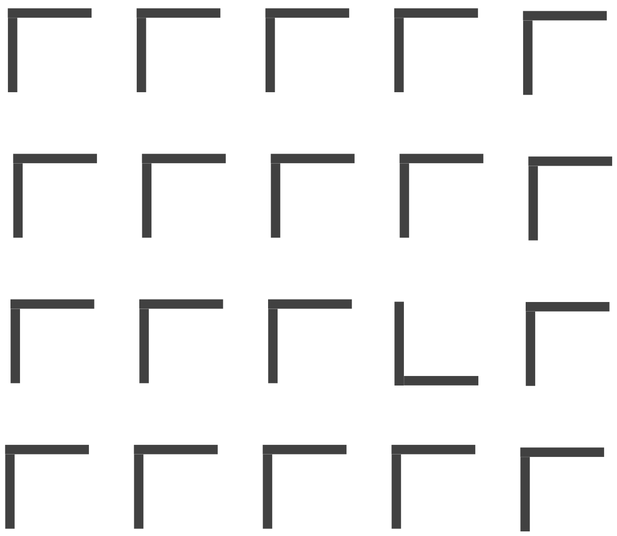
Position

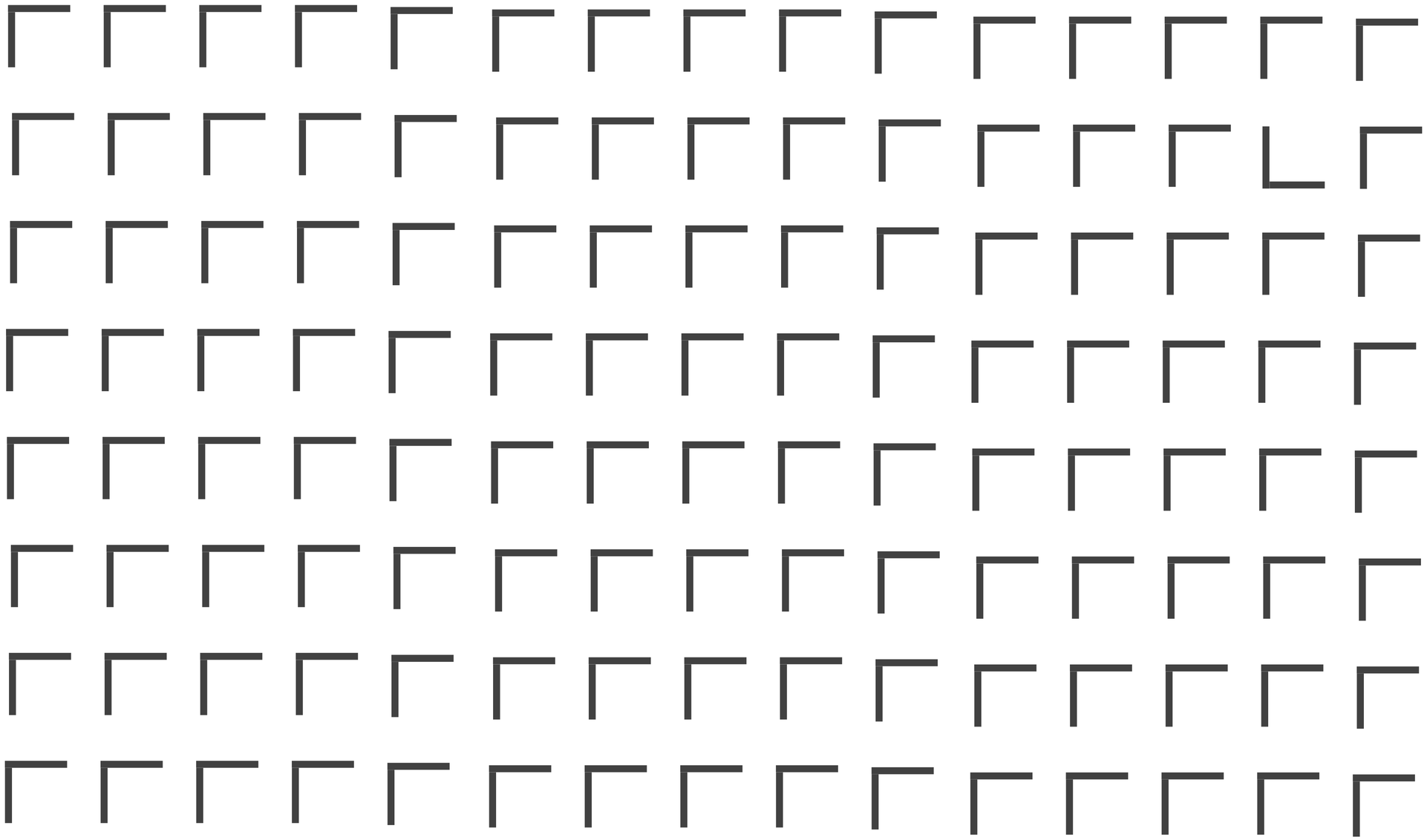
Position







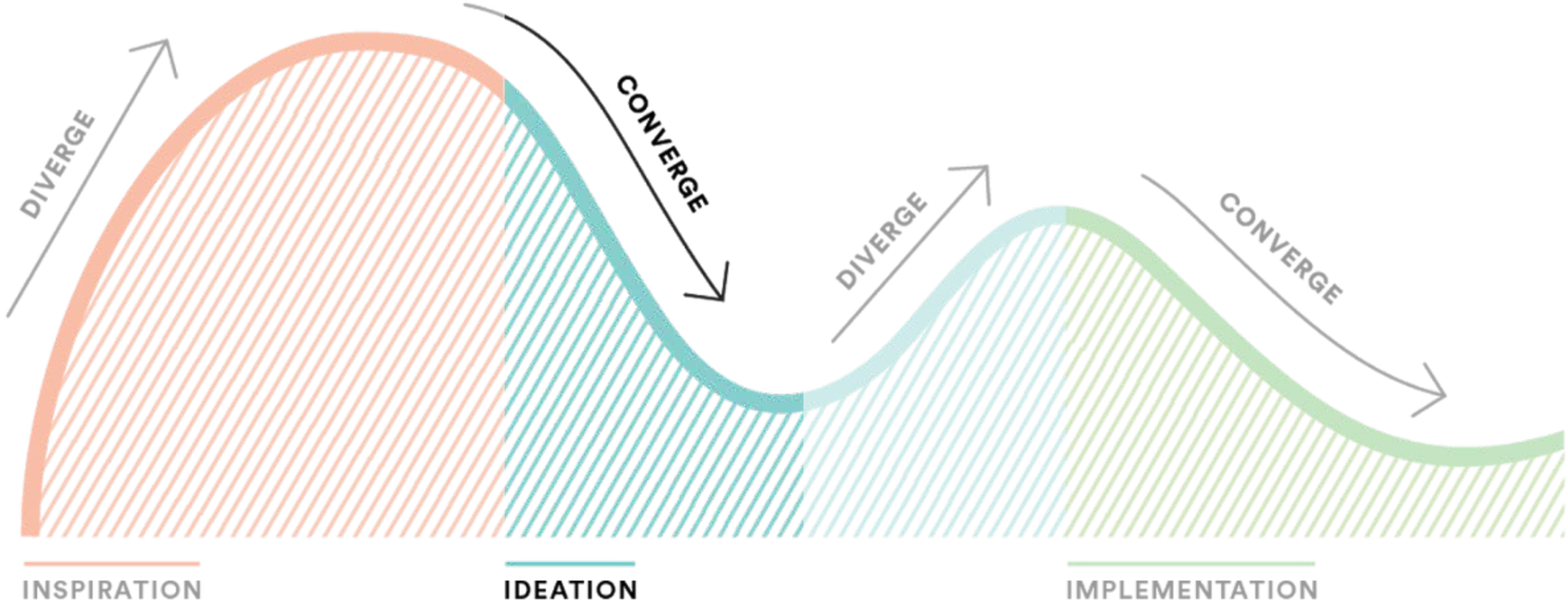




DECIDE

PROTOTYPE TEST

15 MINUTES



Human-centered design process by IDEO. Credit: IDEO

DECIDE

PROTOTYPE TEST

15 MINUTES

DECIDE ON THE BEST SKETCH

- Share sketches within your group
- Cluster & Group Sketches by similarity/questions they answer
- Discuss Pros/Cons of different designs

DECIDE

PROTOTYPE TEST

2 MINUTES

VOTE FOR THE BEST DESIGN **INDIVIDUALLY**!

- Each person gets 3 votes (silver dot stickers)
- You can vote for your own sketch
- You can use all three votes on the same design if you wish

PROTOTYPE TEST

15 MINUTES

REFINE THE CHOSEN SKETCHES INTO A DASHBOARD

- Consider the questions from the map phase.
- Consider any interaction the actual implementation would have.
- Carefully choose the color encodings and other visual channels.
- Incorporate any feedback from the sketch discussions.

PROTOTYPE TEST

UPLOAD PROTOTYPE : 5 MINUTES

TAKE A PHOTO AND UPLOAD YOUR PROTOTYPE TO YOUR GROUP BOARD ON MIRO

TEST : DEMO

5 MINUTES

WE WILL NOW DEMO A THINK ALOUD TEST (LINK)

TEST

2 MINUTES

CHOOSE ONE PERSON TO MOVE TO THE TABLE TO YOUR LEFT

THIS PERSON WILL BE THE “TESTER"

TEST : THINK -ALOUD PROTOCOL

5 MINUTES

CONDUCT A THINK ALOUD PROTOCOL TO TEST YOUR DESIGN.

TEST : FOLLOW-UP

5 MINUTES

ASK YOUR TESTER ANY QUESTIONS THAT YOU MAY STILL HAVE.

PROTOTYPE TEST

10 MINUTES : SHOW AND TELL

SHOW AND TELL

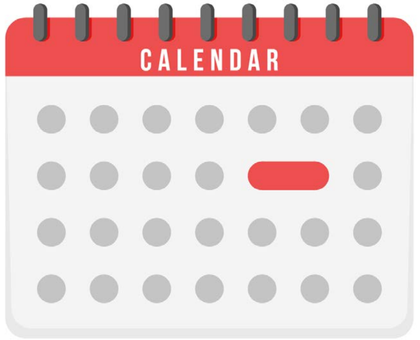
PART III - TIPS FOR PRACTITIONERS AND TEACHERS

TIPS FOR PRACTITIONERS

- RESERVE CONSECUTIVE DAYS
- MAKE SURE EVERYONE IS FOCUSED AND COMMITTED ON SPRINT
- ROLES IN A DESIGN SPRINT:
  - FACILITATOR
  - DESIGNER, ENGINEER, MARKETER
  - PRODUCT CHIEF (DECIDER)

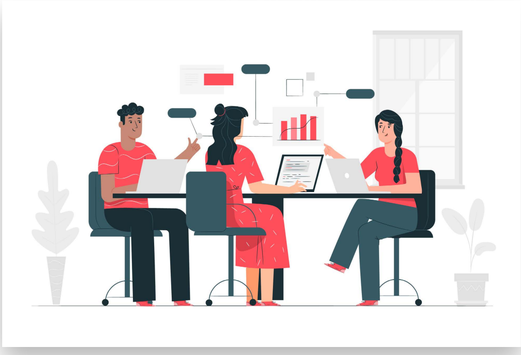
TIPS FOR TEACHERS: ADJUST LENGTH, SCOPE & DIFFICULTY

Length:



Modality:

Online /
In-Person:



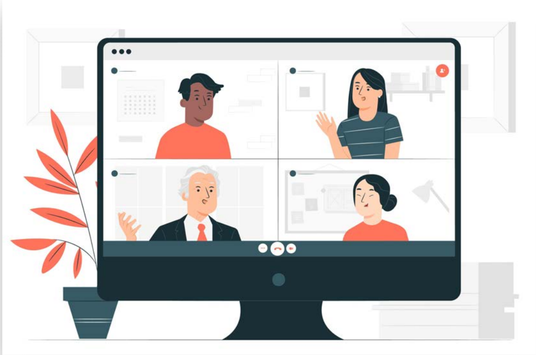
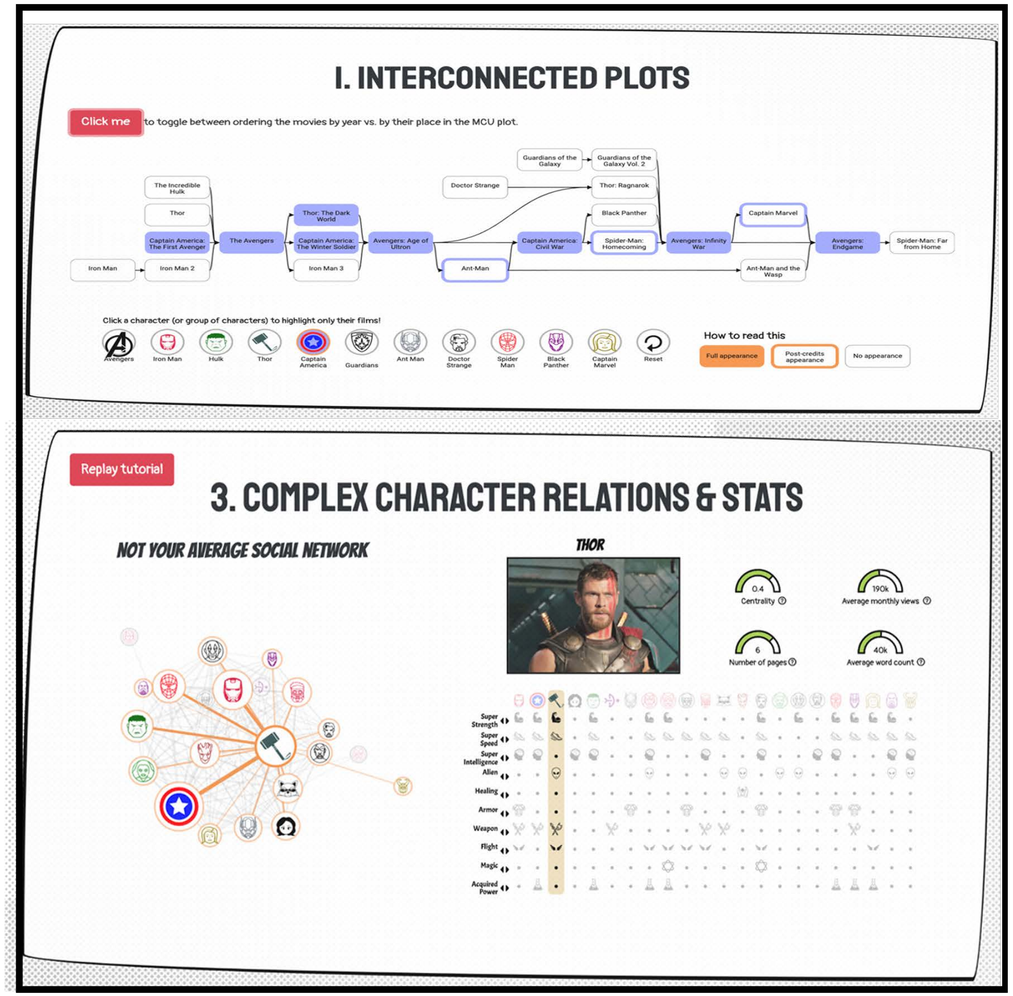
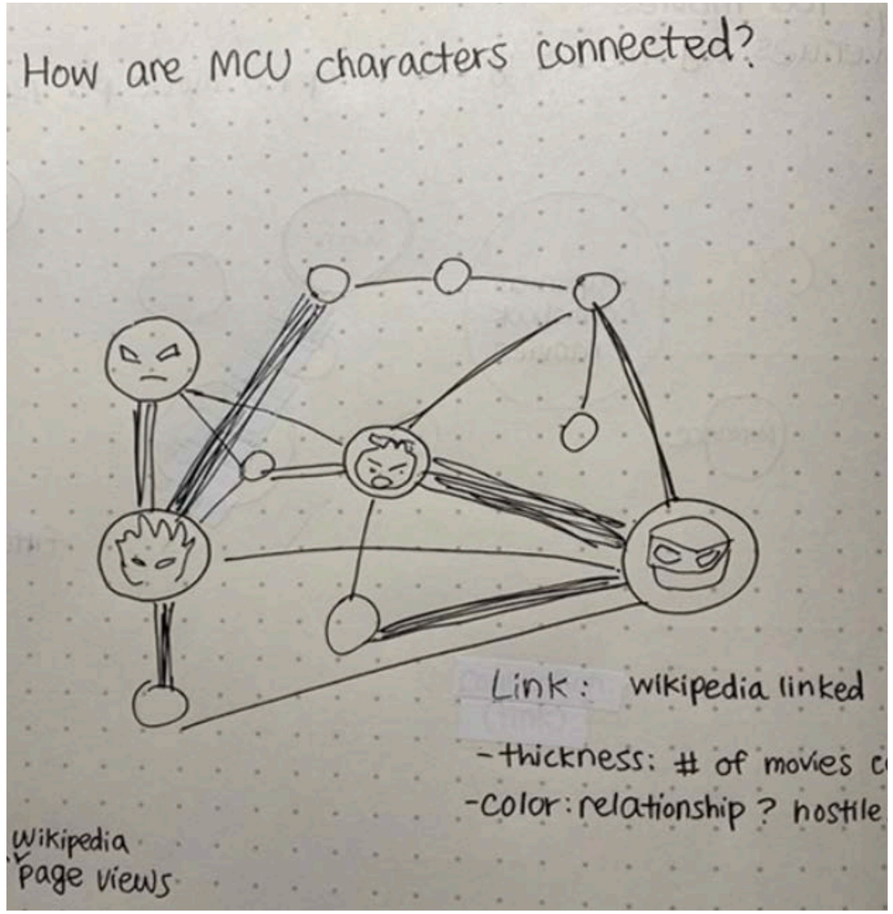


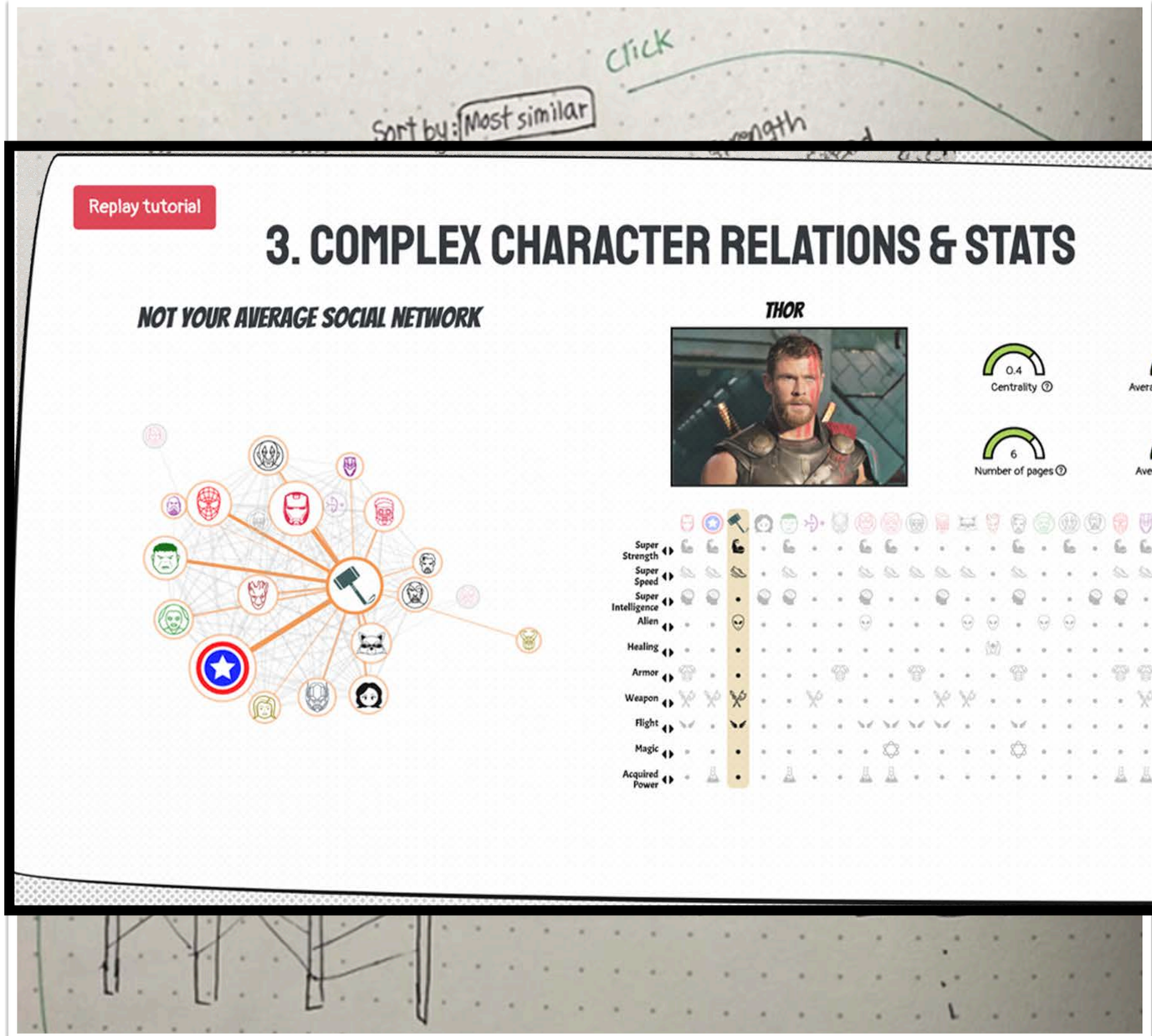
Image by catalyststuff, slides and brgfx on Freepik

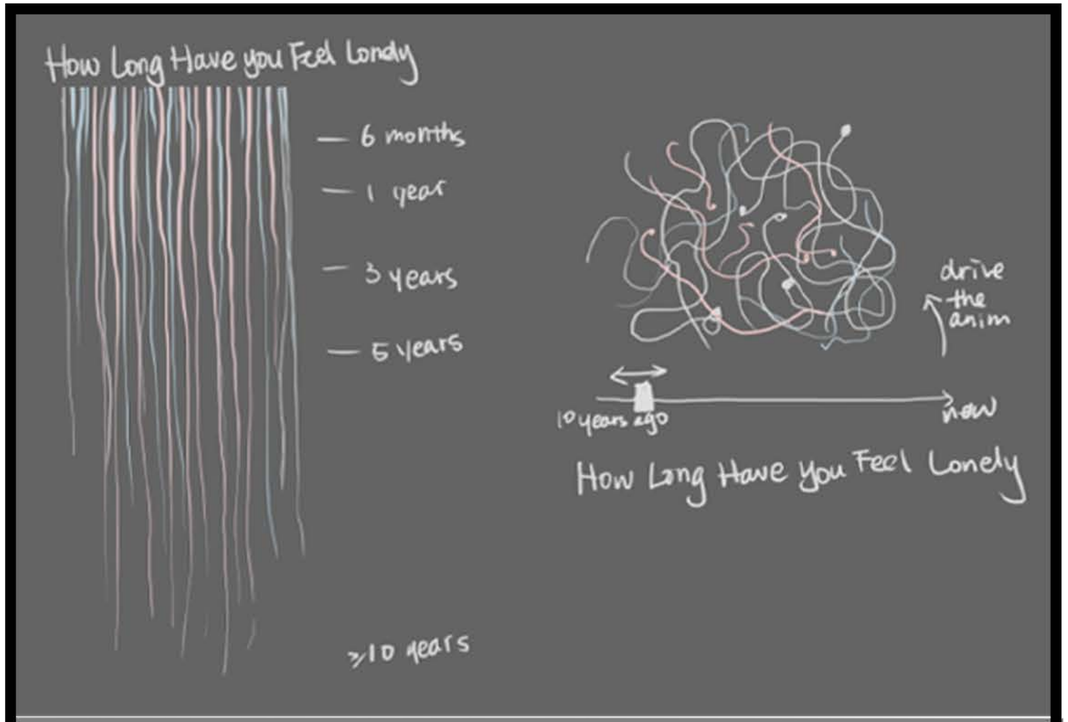
EXAMPLE UNDERGRAD CLASS

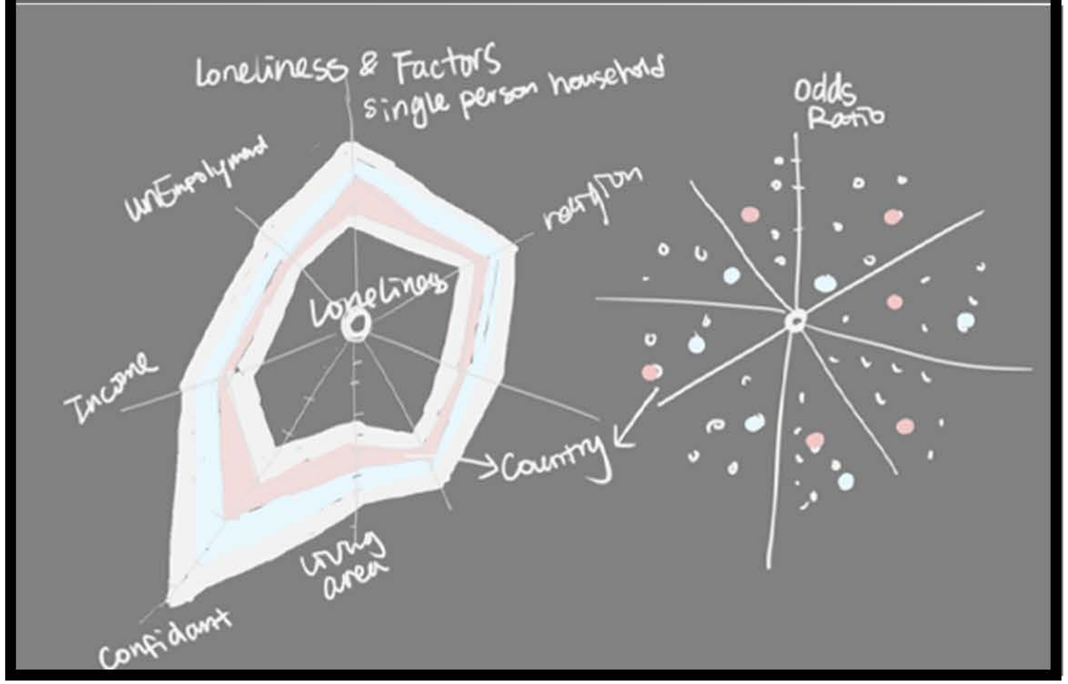
- FINAL PROJECT
  - LAST 5 WEEKS OF CLASS
  - FREE TEAM AND DATASET CHOICE
  - WEEKLY SUBMISSION OF PROJECT NOTEBOOK
  - STORYTELLING WEBSITE WITH D3
-




How are MCU characters connected?
Link: wikipedia linked
-thickness: # of movies c
-color: relationship? hostile
Wikipedia
Page views

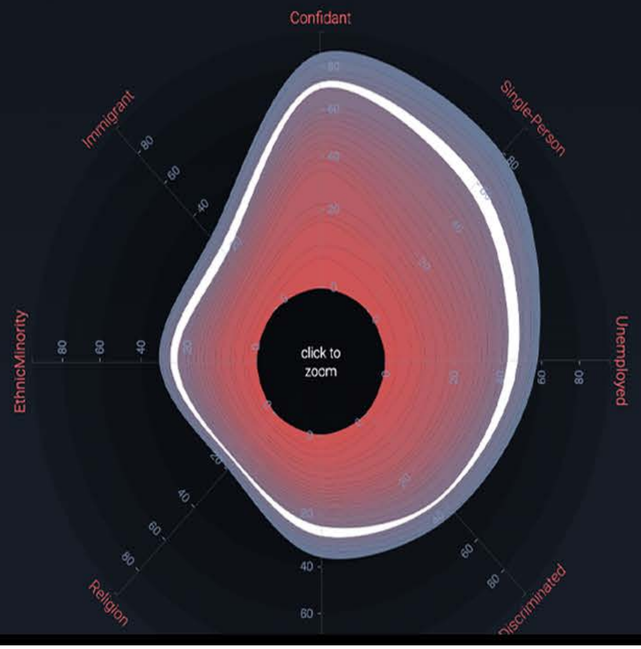

click
Sort by: Most similar
Replay tutorial
3. COMPLEX CHARACTER RELATIONS & STATS
NOT YOUR AVERAGE SOCIAL NETWORK
THOR
0.4
Centrality
190k
Average monthly views
6
Number of pages
40k
Average word count
Super Strength
Super Speed
Super Intelligence
Alien
Healing
Armor
Weapon
Flight
Magic
Acquired Power





LONELINESS

a social story



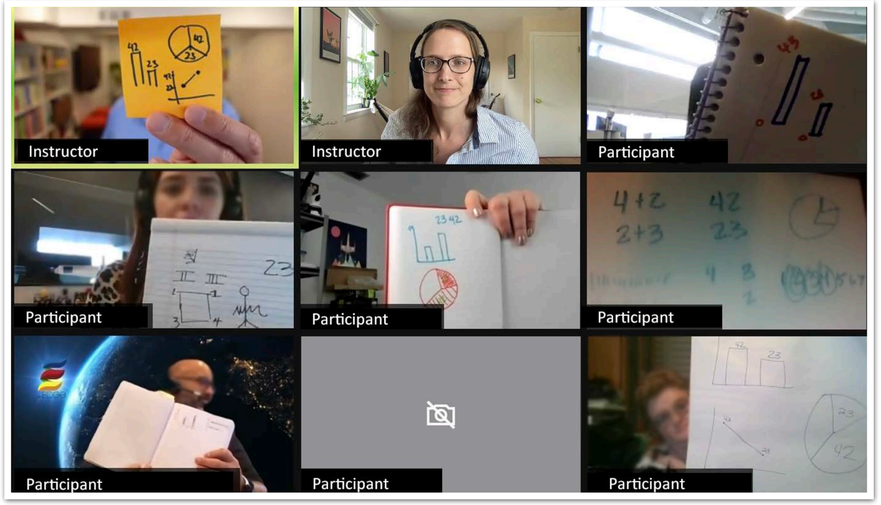
If we summarize from these data what factors contribute more to loneliness. We can find that **loneliness is most related with having no confidant, living alone and being unemployed**. This result is not surprising. Research shows, among three levels of social ties, **confidant relationship is most salient to emotional-well being**. But maintaining other levels of social connections can also alleviate loneliness:

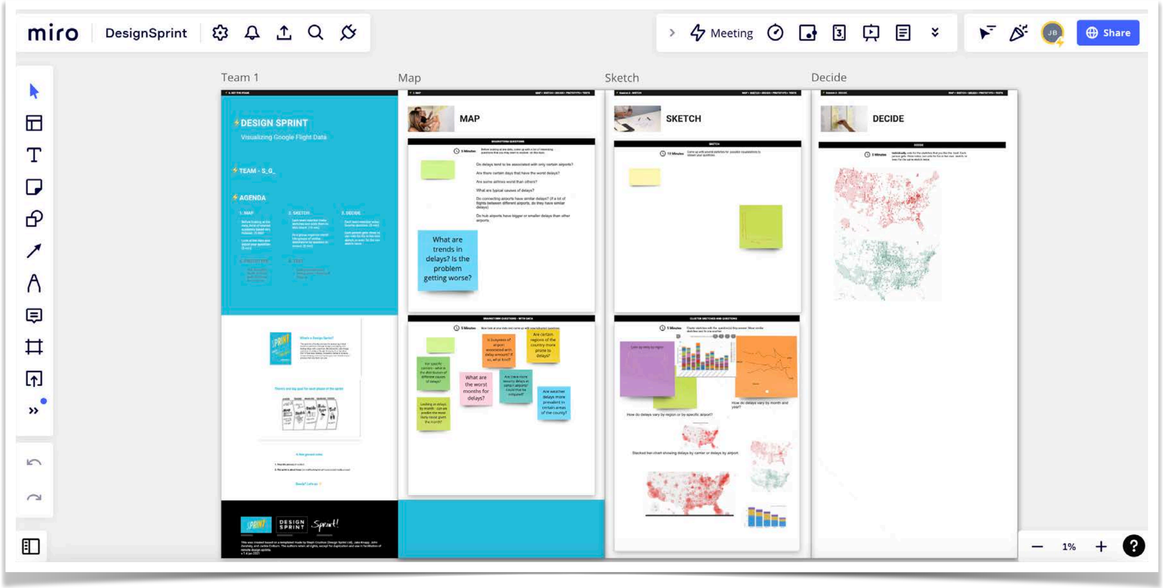
Top 3 Factors Contributing to Loneliness in

Norway

EXAMPLE ONLINE CLASS

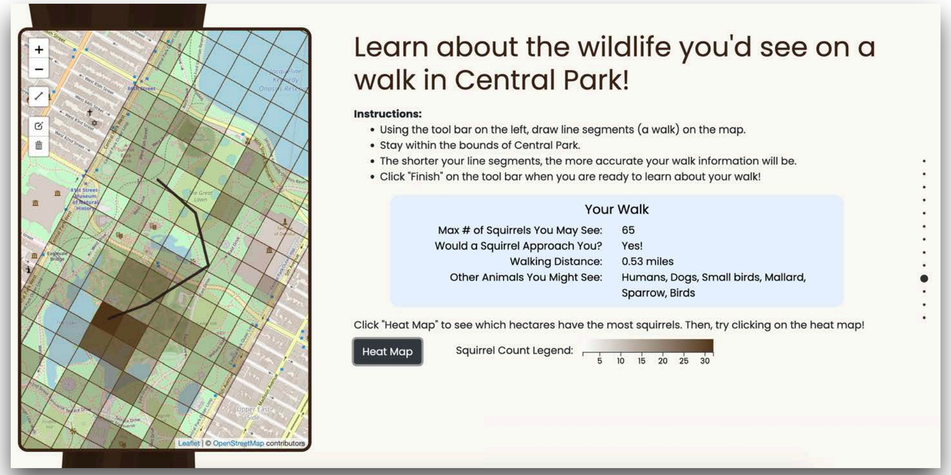
- ZOOM BREAK-OUTS AT END OF CLASS
- 1 GOOGLE DOC PER TEAM
- 1 MIRO BOARD PER CLASS
- WEBSTORM ‘CODE WITH ME’





TEACHING WITH VIS DESIGN SPRINTS

- STUDENTS CAN IMMEDIATELY APPLY WHAT THEY LEARNED IN CLASS
- PROJECTS ADHERE BETTER TO DESIGN GUIDELINES
- STUDENT ENGAGEMENT IS HIGH, PROJECT QUALITY IS HIGH



DESIGN SPRINTS FOR VISUALIZATION

CAROLINA NOBRE, JOHANNA BEYER

jbeyer@g.harvard.edu
cnobre@cs.toronto.edu

Design Sprint Materials

Tutorial Slides, Example datasets, instructions, class activities, grading rubrics, example student projects).

https://github.com/VCG/vis_design_sprints_tutorial_2023

UNIVERSITY OF TORONTO

HARVARD
School of Engineering and Applied Sciences