# Visualization & Visual Analytics 1

Angus Forbes

creativecoding.evl.uic.edu/courses/cs424

# Sketches

Wide range of scores, between 14 and 77. (Minimum score was a 9, maximum a 90).

# Sketches

Many people chose similar datasets:

- how often do I use my phone (10 people)
- how often do I watch netflix (3 people)
- how much do I drink (3 people)
- what are my daily activities (4 people)

# Sketches

And similar visualization choices:

- circle packing layouts: 15
- bar charts : 13
- maps : 9
- sunburst charts : 8
- pie charts : 8
- clocks : 4
- world clouds : 3
- people getting fat : 2

# Sketches

*Content* was vague...

I.e., a visualization that displayed information about Netflix only showed the genre, but left out interesting information about *what* specific content was being watched. Or some visualizations looked at, say, Snapchat, but didn't provide any detail of what the actual content of the interaction was.

# Sketches

*Context* was vague...

I.e., A visualization that showed that you used Snapchat on certain days would be much more interesting if the visualization gave some insight into *why* you were using Snapchat, *how* you were using, what you felt or thought about when using it, etc.

# Sketches

## *Context* was vague...

Or, a visualization about how many people got on or off certain subway stops left out interesting contextual information about the *meaning* of those subway stops – Are some stops in wealthier or poorer areas? Are some stops close to where people work? Are some times of day related to traffic to sporting events? What are the demographics of the people getting on/off each stop?

# Sketches

## *Good vs not so good...*

- Good visualizations were correlated with more complex or sophisticated data collection

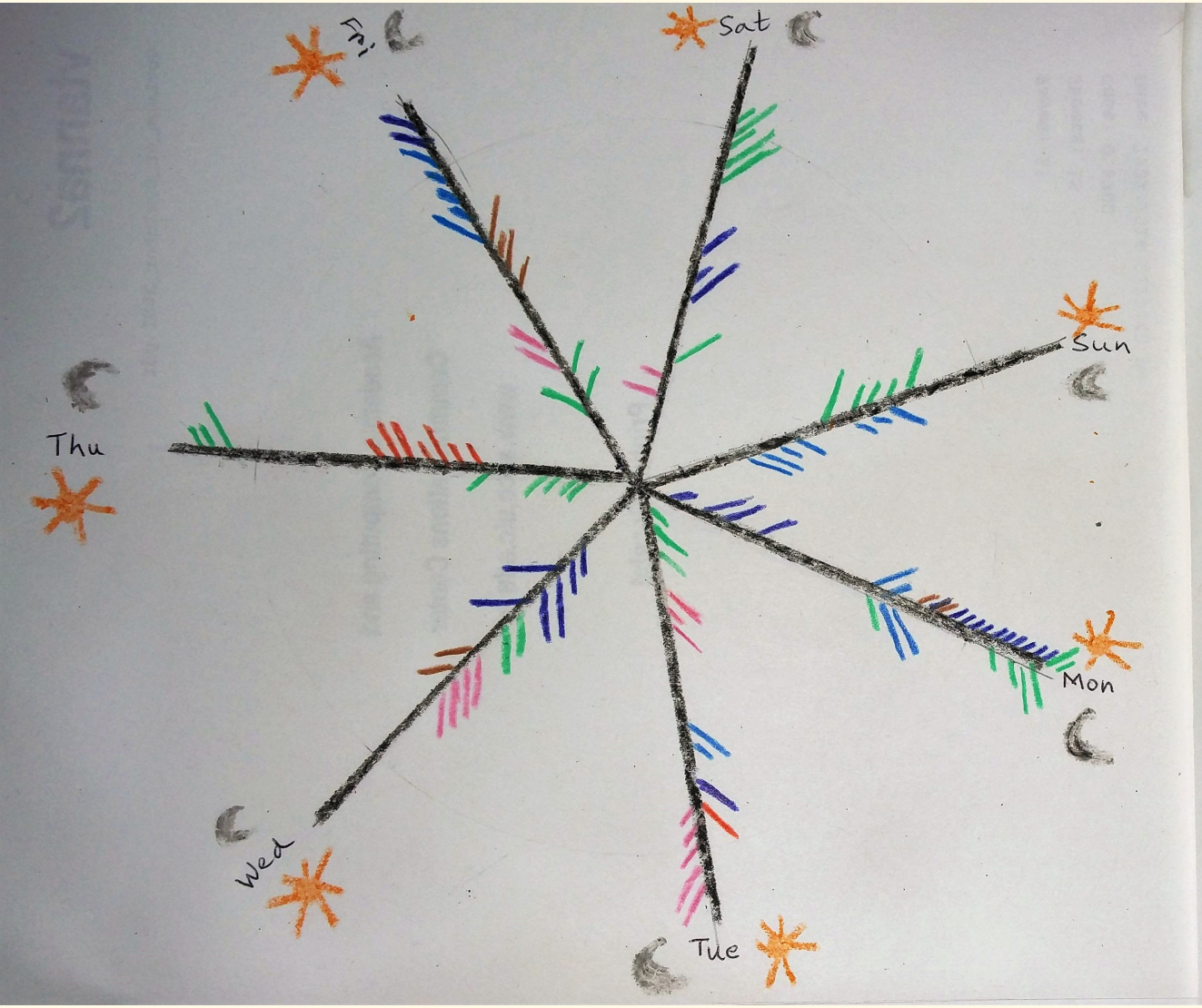
- Good visualizations seemed to incorporate the actual data

- The not so good visualizations weren't simply drawn poorly, they failed to communicate the purpose of the visualization and the relevance of the data

# Sketches

Something to think about - whether in the field of design or any other field (business, data science, research in general, etc) - what makes something interesting? and why?

Another question to ask yourself is: why *this* visualization instead of another one? what does the viewer get out seeing the data in this particular way?




0000 – 1159 hrs

1200 – 2359 hrs

long yawn

short yawn

Studying

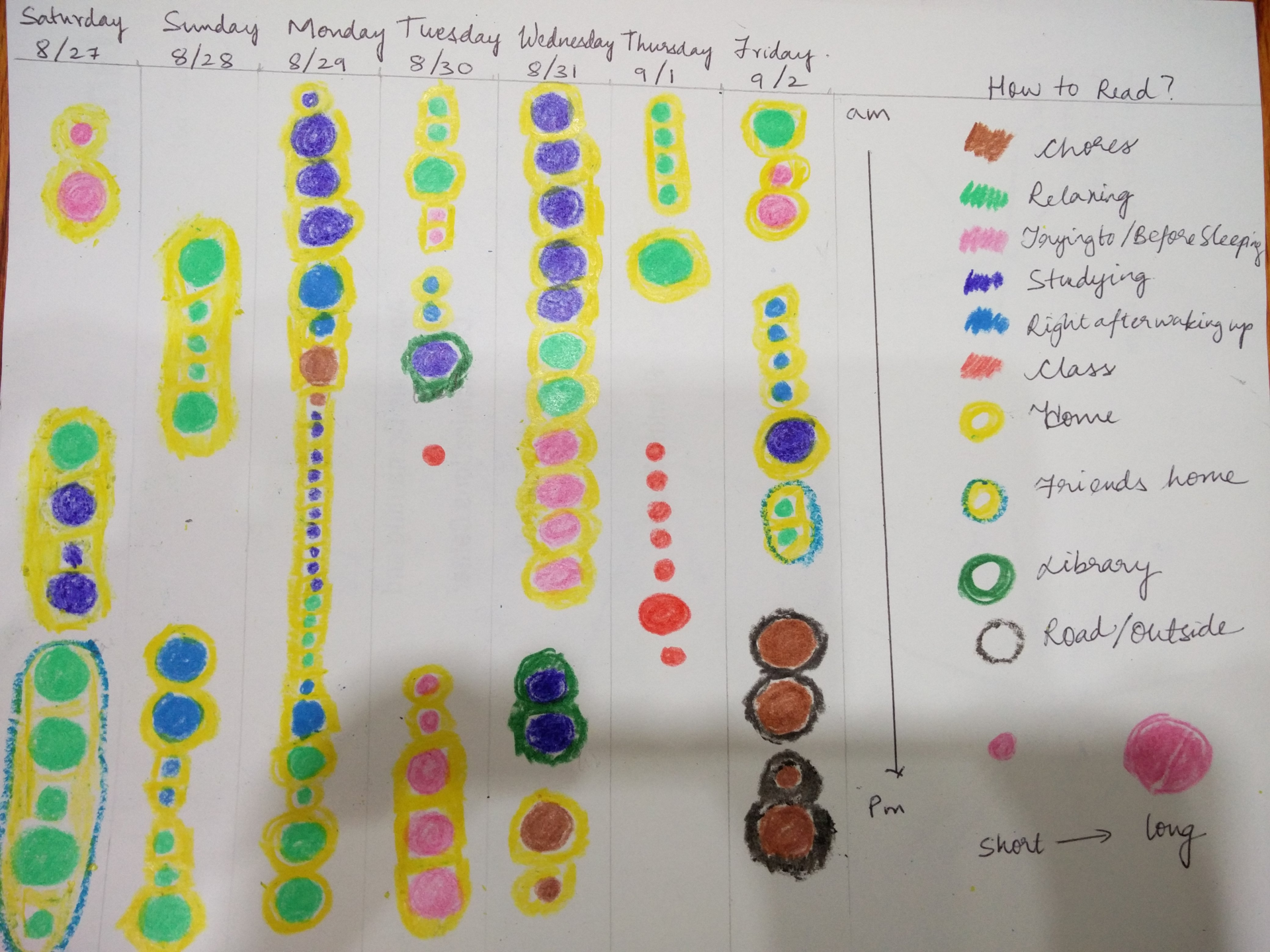
Class

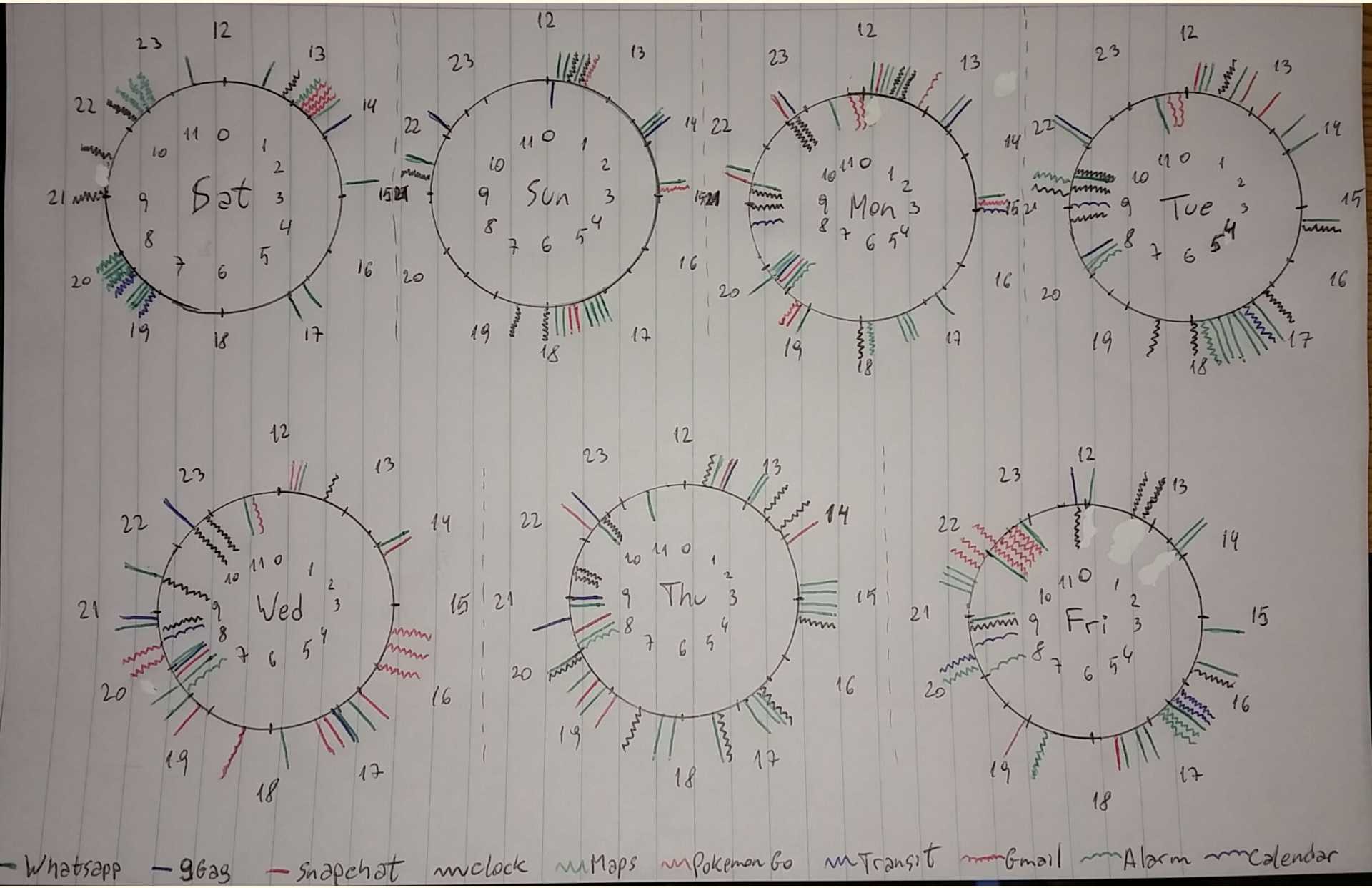
Right after waking up

Trying to / before sleeping

Studying

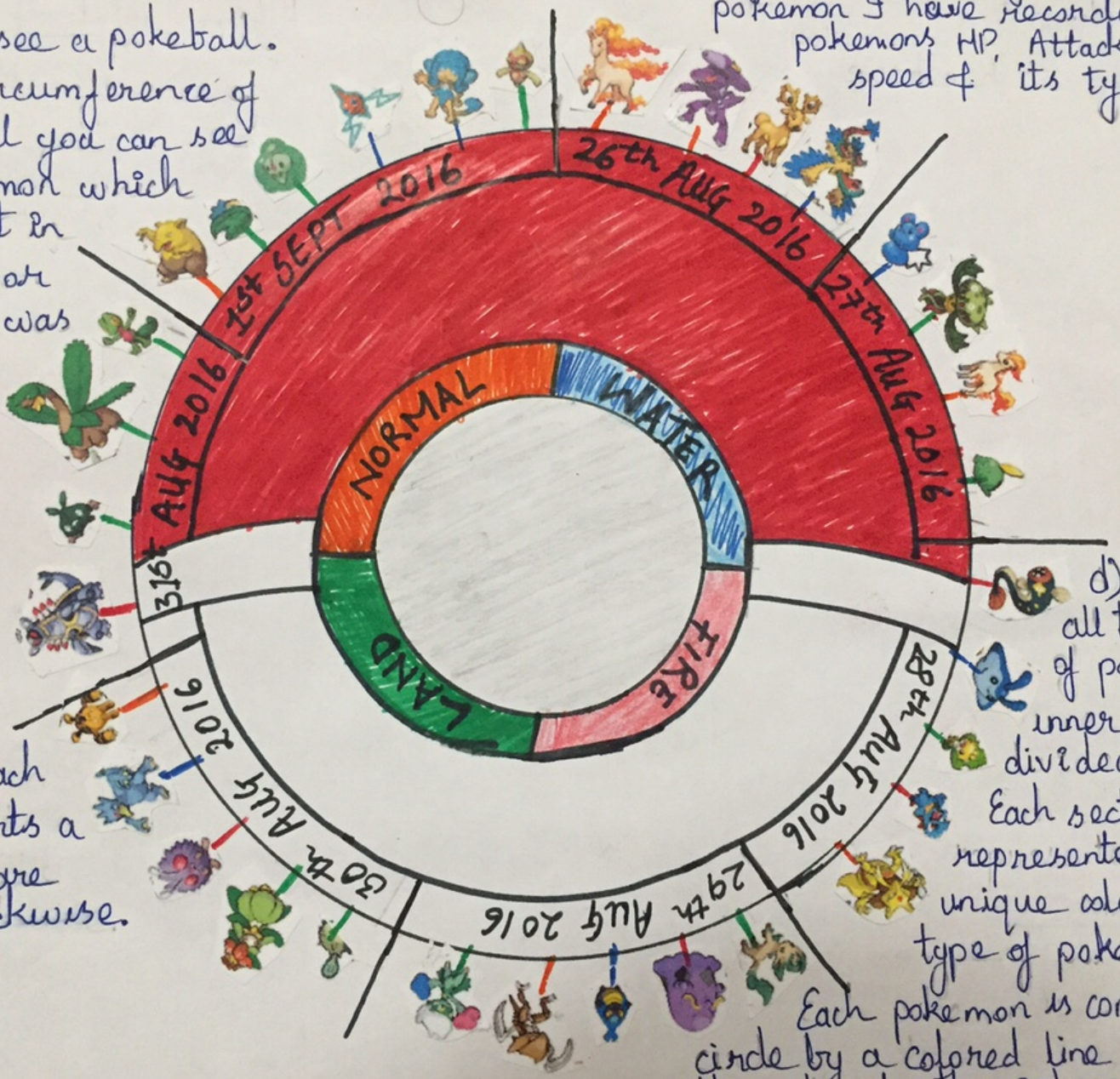
Relaxing


Saturday 8/27
Sunday 8/28
Monday 8/29
Tuesday 8/30
Wednesday 8/31
Thursday 9/1
Friday 9/2
am
Pm
How to Read?
Chores
Relaxing
Trying to / Before Sleeping
Studying
Right after waking up
Class
Home
Friends home
Library
Road / Outside
short → long


Sat
Sun
Mon
Tue
Wed
Thu
Fri
Whatsapp
9Gag
Snapchat
Clock
Maps
Pokemon Go
Transit
Gmail
Alarm
Calendar

## (1) Initial state:

a) You can see a pokeball. Along the circumference of the pokeball you can see all the pokemon which were caught in the 7 days for which data was collected.




b) Pokeball is divided into 7 arcs & each arc represents a date. Dates are ordered clockwise.

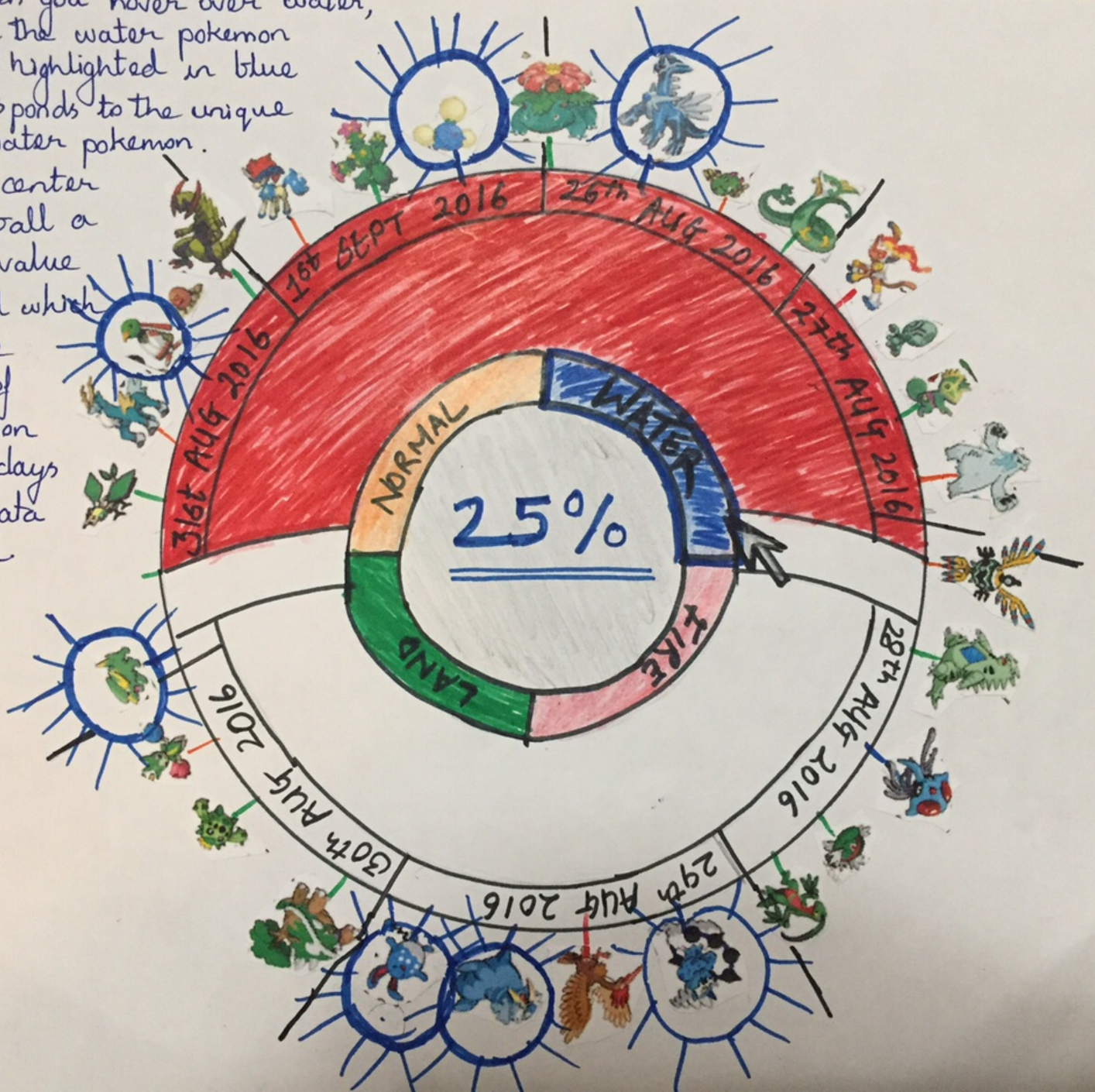
c) Each arc contains all the pokemon caught on that day. For each pokemon I have recorded the pokemons HP, Attack, defence speed & its type.

d) Based on all the types of pokemon, the inner circle is divided into sections. Each section is represented by a unique color for each type of pokemon.

Each pokemon is connected to the circle by a colored line representing the color for the pokemon type.

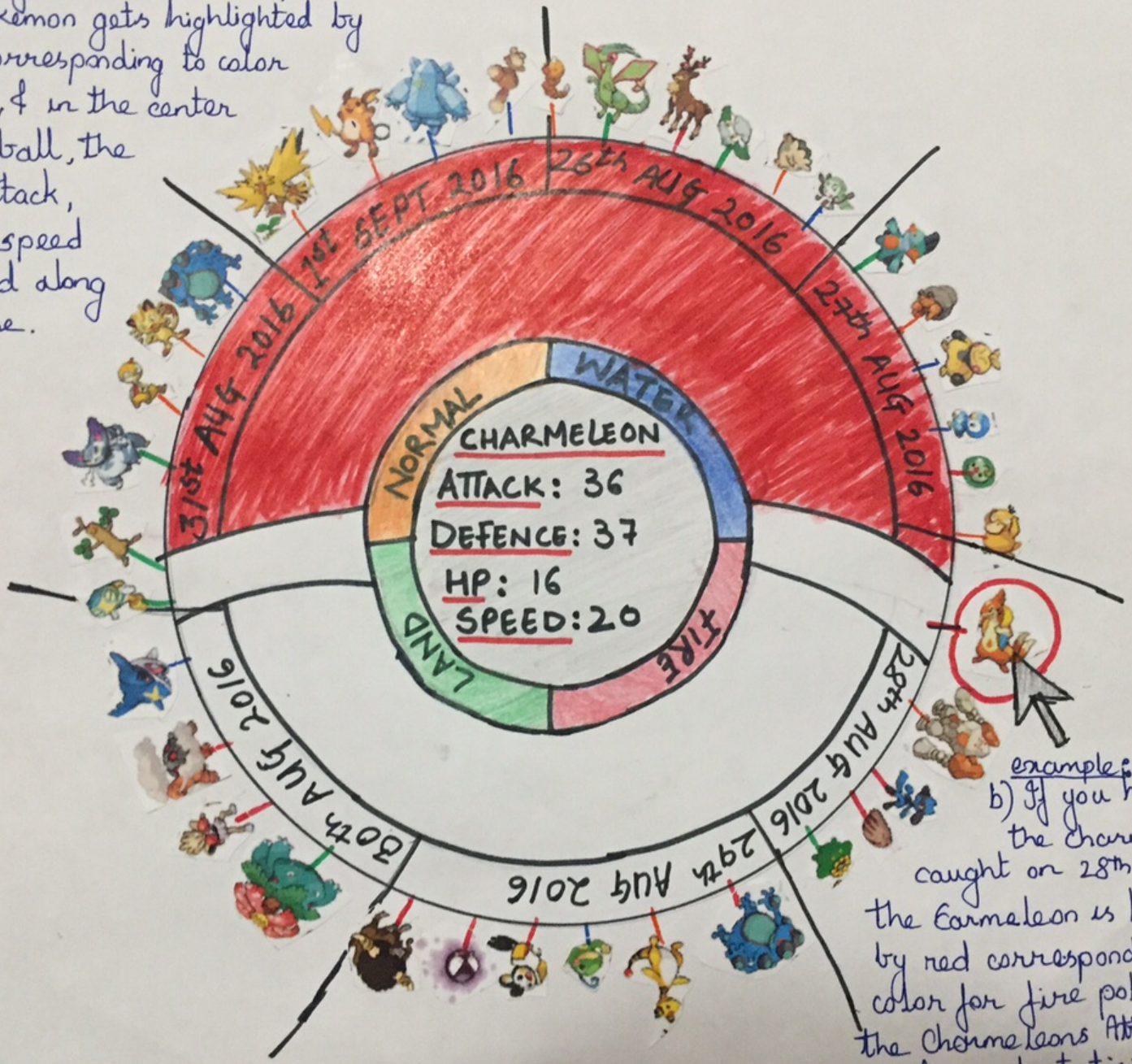
(2) When you hover over water, all the water pokemon get highlighted in blue which corresponds to the unique color for water pokemon.

Also in the center of the pokeball a percentage value is displayed which indicates the percentage of water pokemon caught in 7 days for which data was collected



3

a) When you hover over a pokemon the pokemon gets highlighted by the color corresponding to color for its type, & in the center of the pokeball, the pokemons attack, defence, HP, speed get displayed along with its name.




example:

b) If you hover over the Charmeleon caught on 28th August 2016, the Charmeleon is highlighted by red corresponding to the color for fire pokemon. Also the Charmeleons Attack, defence HP & speed statistics are shown

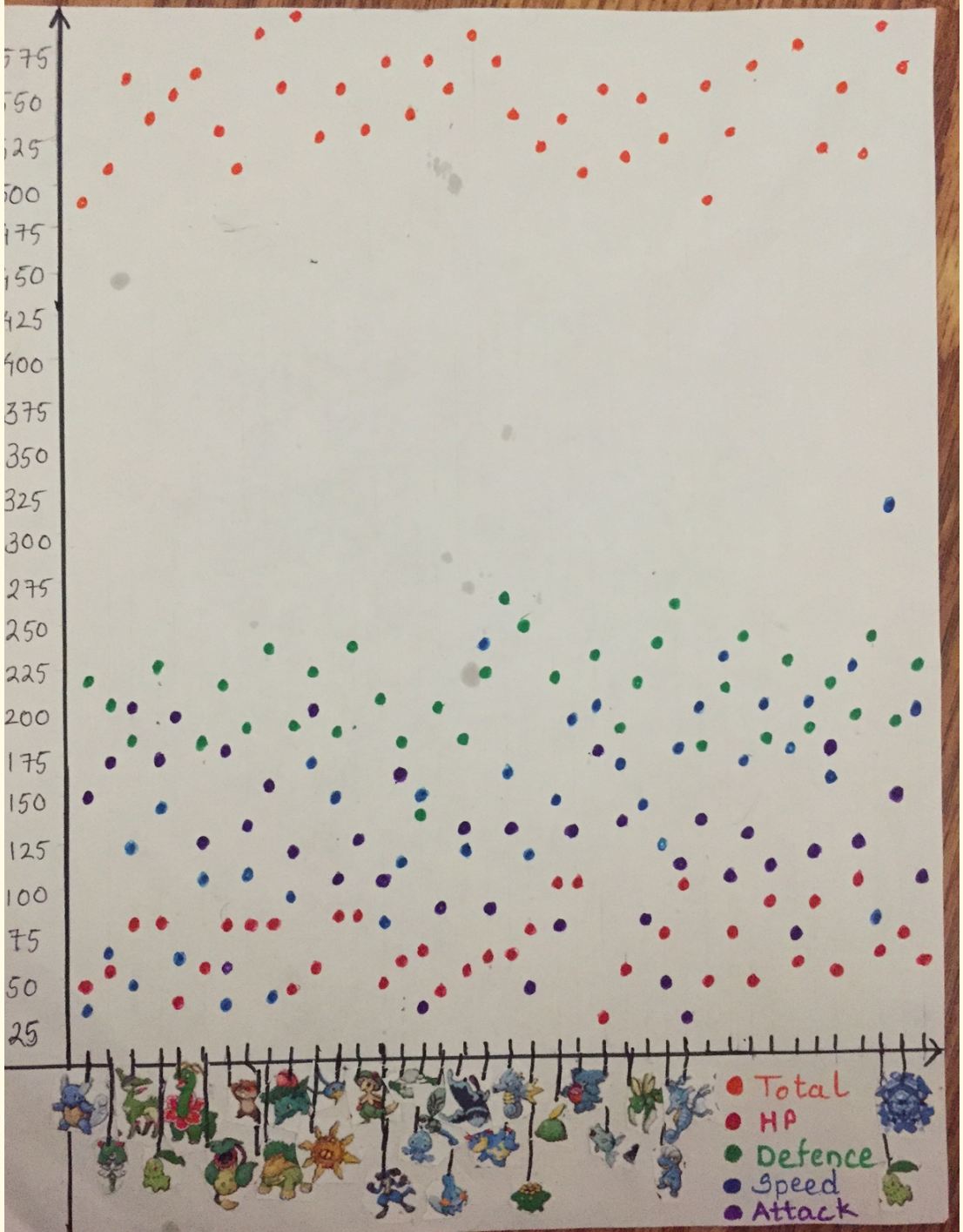
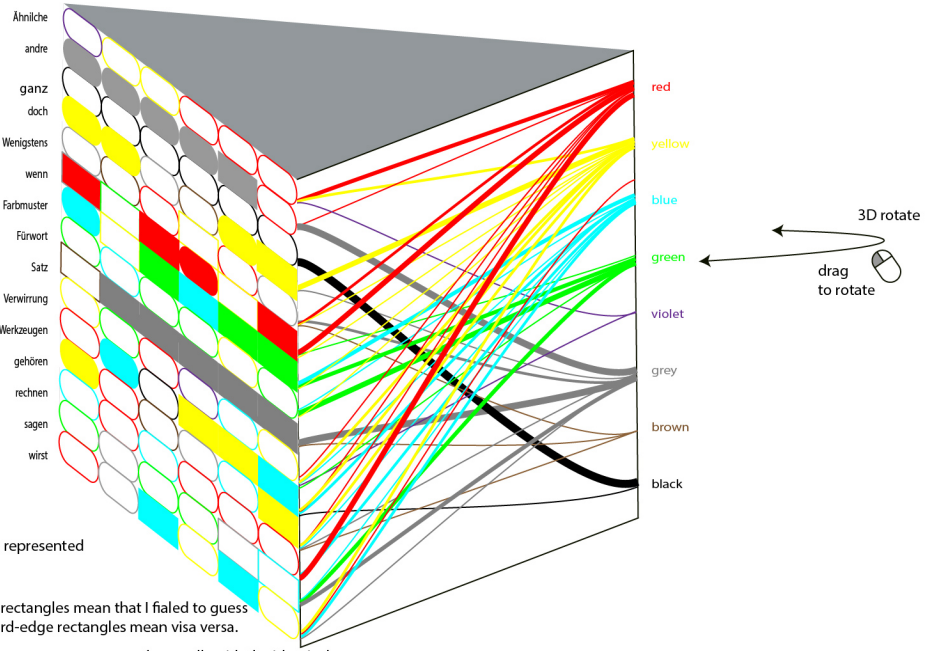

575
550
525
500
475
450
425
400
375
350
325
300
275
250
225
200
175
150
125
100
75
50
25
Total
HP
Defence
Speed
Attack

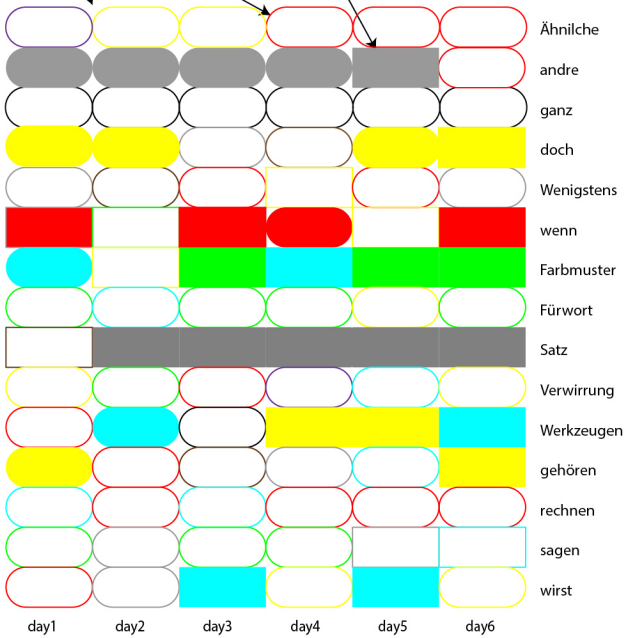
Figure - 2




the associated color is represented by the storke tone

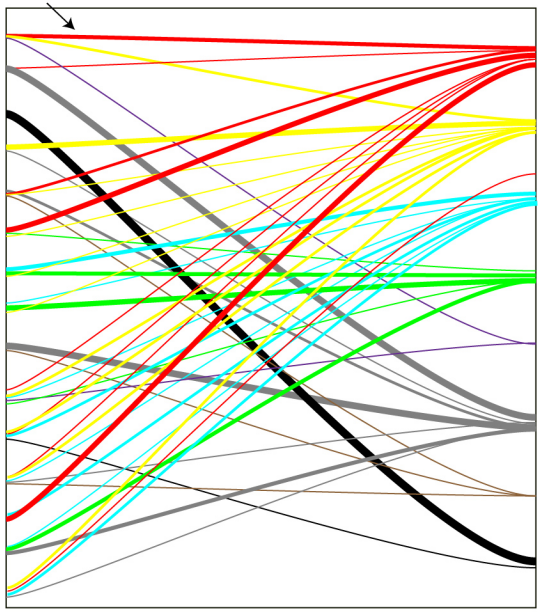
round corner rectangles mean that I fialed to guess meanings. Hard-edge rectangles mean visa versa.

those cells with the identical color to the correct guess are mareked by solid colors

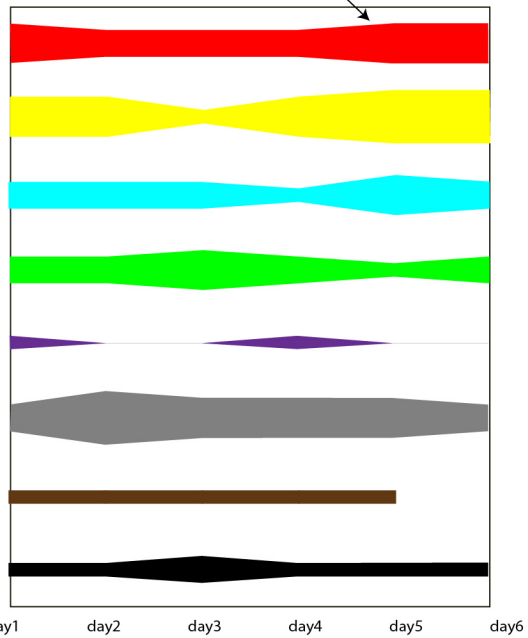


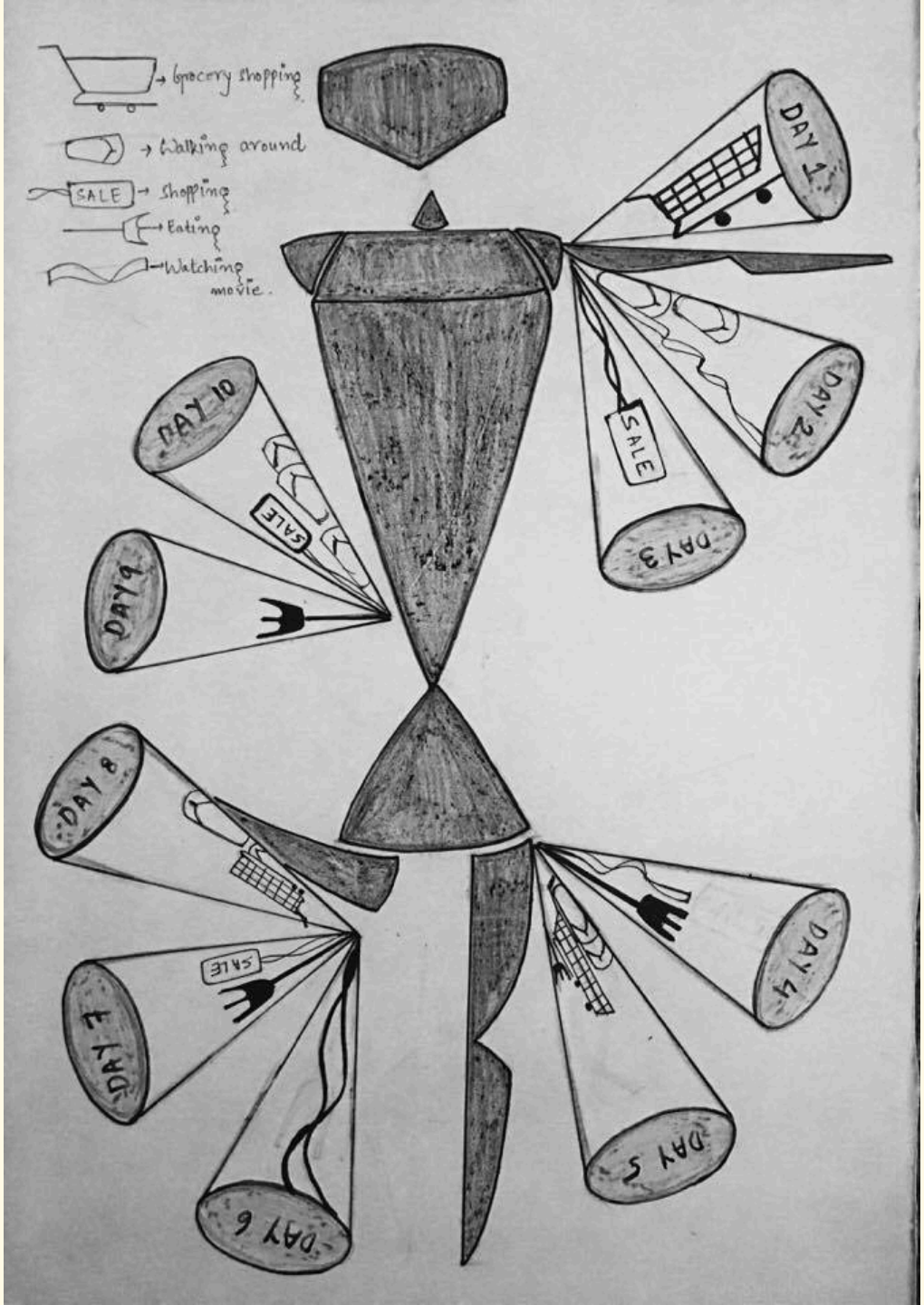

stroke indicates its number of recurrence. e.g. 3pt = 3 times associated with the term

hovering pointer over a color will decrease the opacity of the others, so the selected color will stands out.

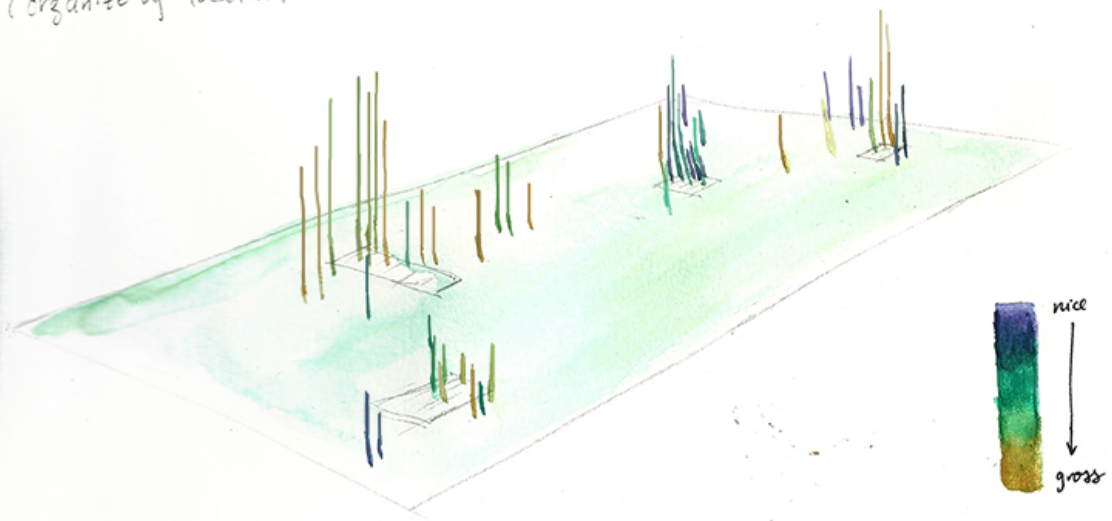



thickness indicates the population of the color in the corresponding day

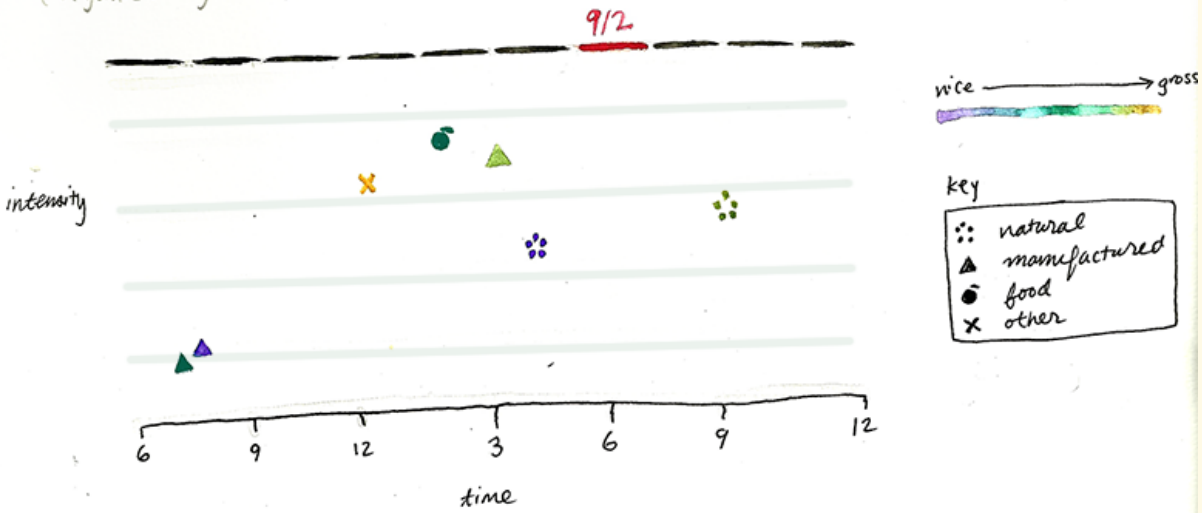


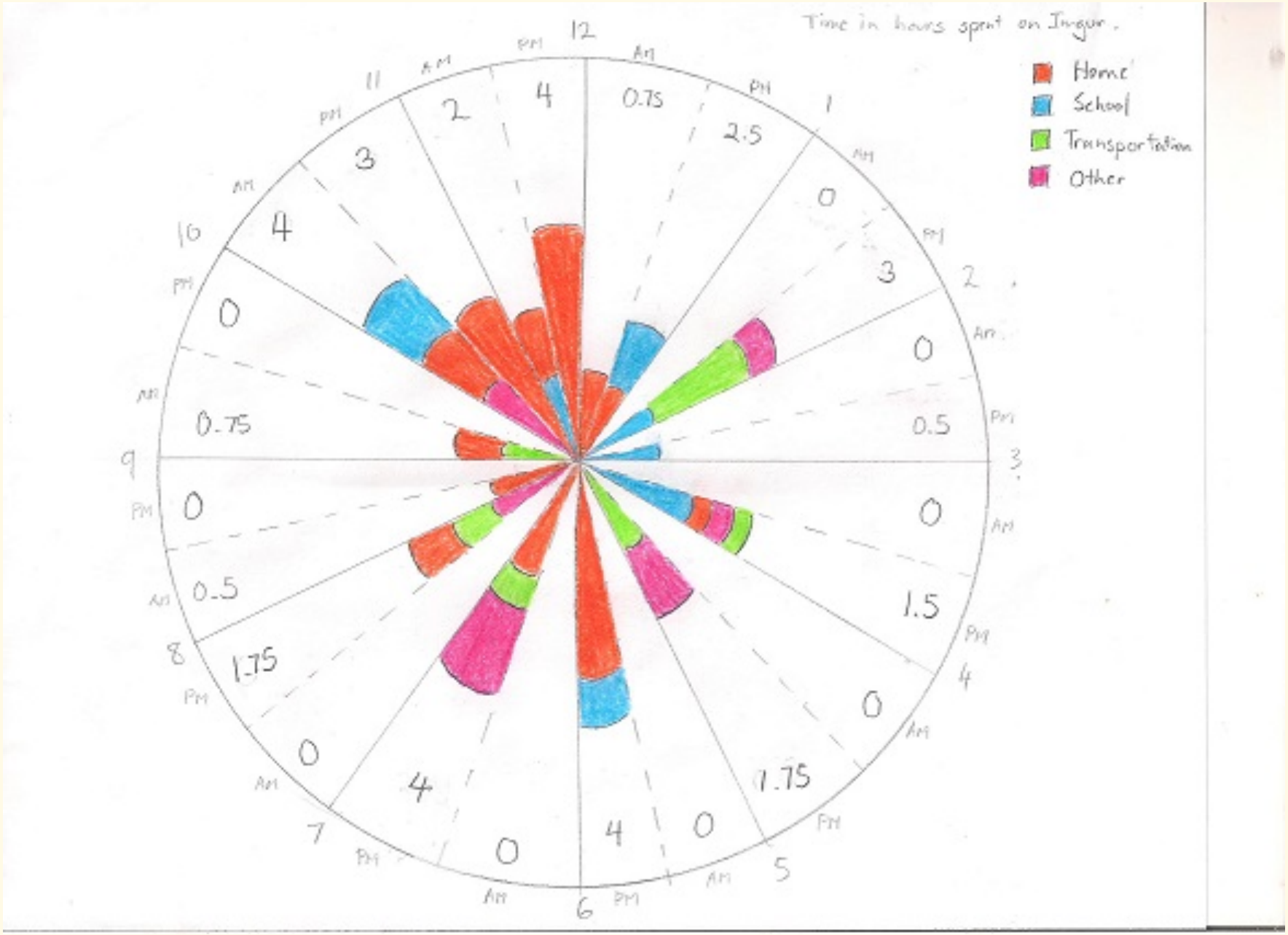

→ Grocery shopping
→ Walking around
SALE → Shopping
→ Eating
→ Watching movie.
DAY 1
DAY 2
SALE
DAY 3
DAY 10
SALE
DAY 9
DAY 8
SALE
DAY 7
DAY 6
DAY 5
DAY 4

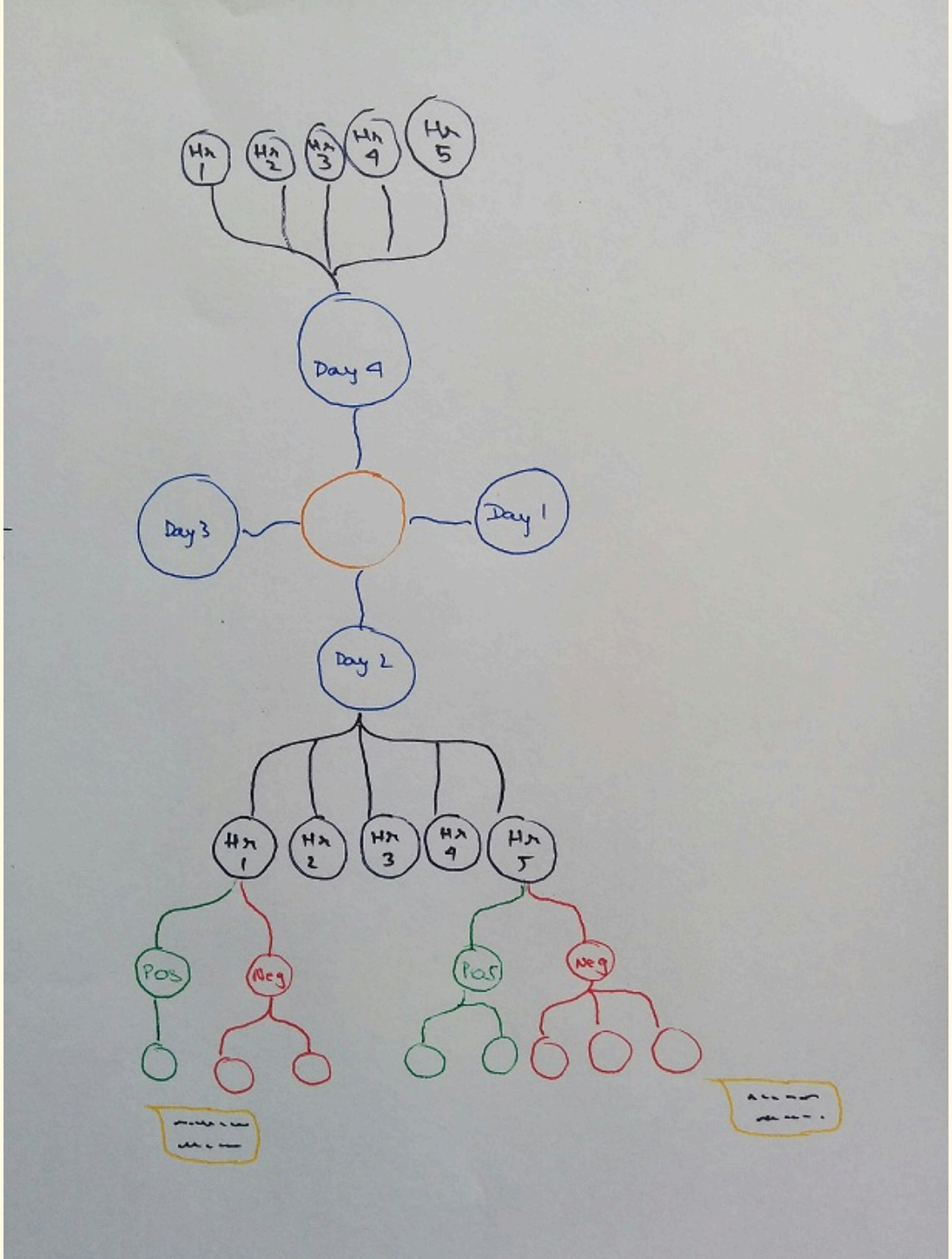
(organize by location)

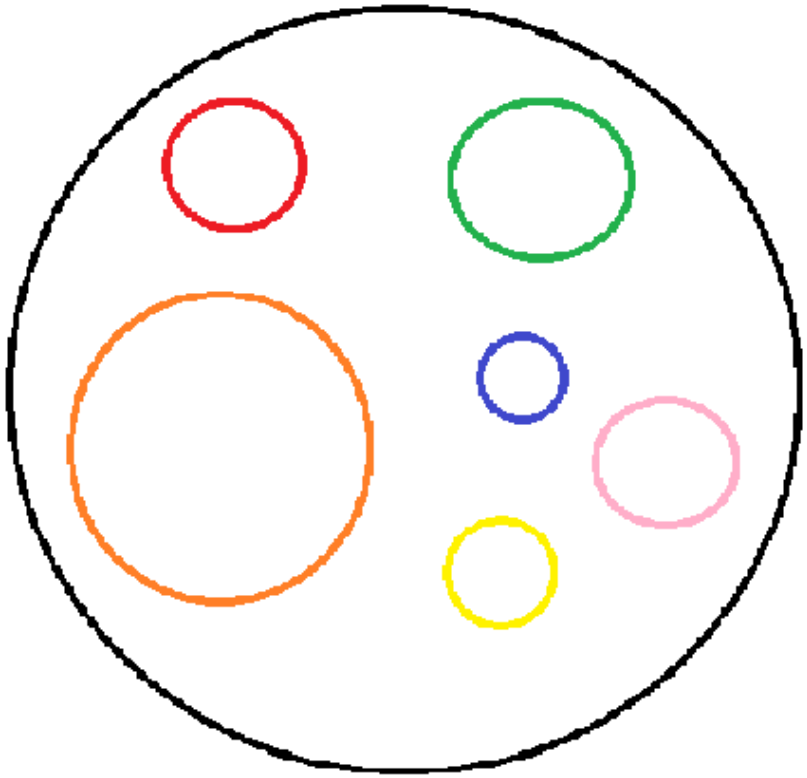



(organize by time)

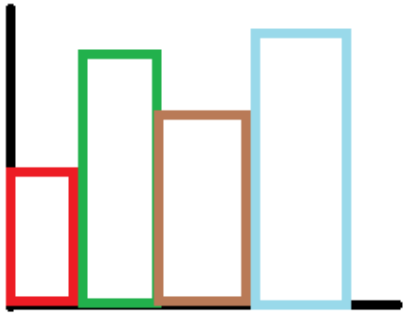


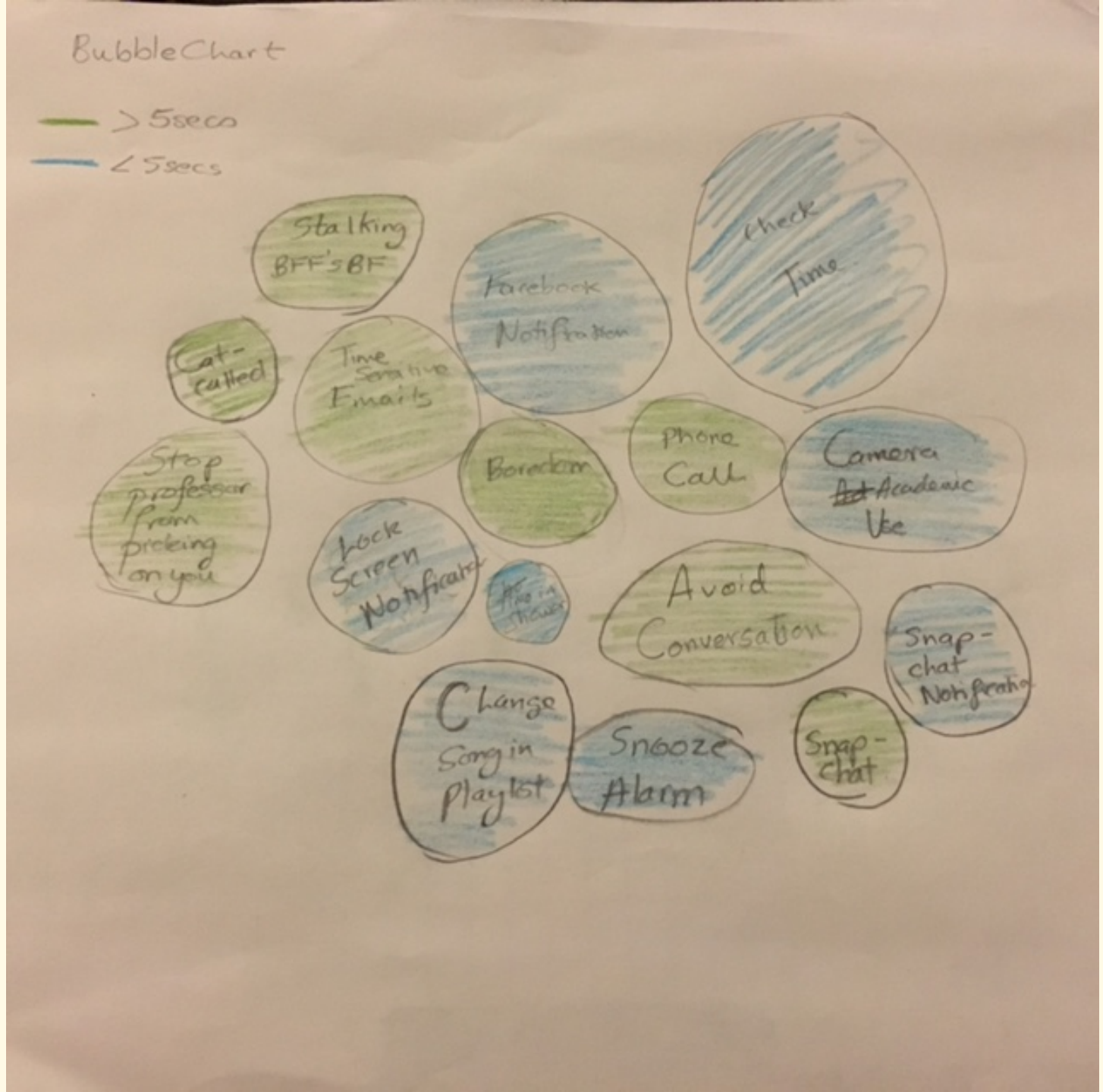

Time in hours spent on Imgur.
Home
School
Transportation
Other
12 AM
0.75
2.5
1 AM
0
3
2 AM
0
0.5
3 AM
0
1.5
4 AM
0
1.75
5 AM
0
4
6 PM
4
0
7 PM
0
4
8 PM
1.75
0
9 PM
0
0.5
10 PM
0
0.75
11 PM
4
3
12 PM
2
4


Hr 1
Hr 2
Hr 3
Hr 4
Hr 5
Day 4
Day 3
Day 1
Day 2
Hr 1
Hr 2
Hr 3
Hr 4
Hr 5
Pos
Neg
Pos
Neg



Each circle is a different day that I monitored my usage of apps. Every circle but the black one, its size is based on the total usage. Clicking on each circle will result in a popup with bar graph.




Bubble Chart
> 5secs
< 5secs
Stalking BFF's BF
Facebook Notification
Check Time
Cat-calling
Time Sensitive Emails
Boredom
Phone Call
Camera Academic Use
Stop professor from picking on you
Lock Screen Notifications
Avoid Conversation
Snap-chat Notifications
Change Song in Playlist
Snooze Alarm
Snap-chat

Oscar Salinas
CS 424

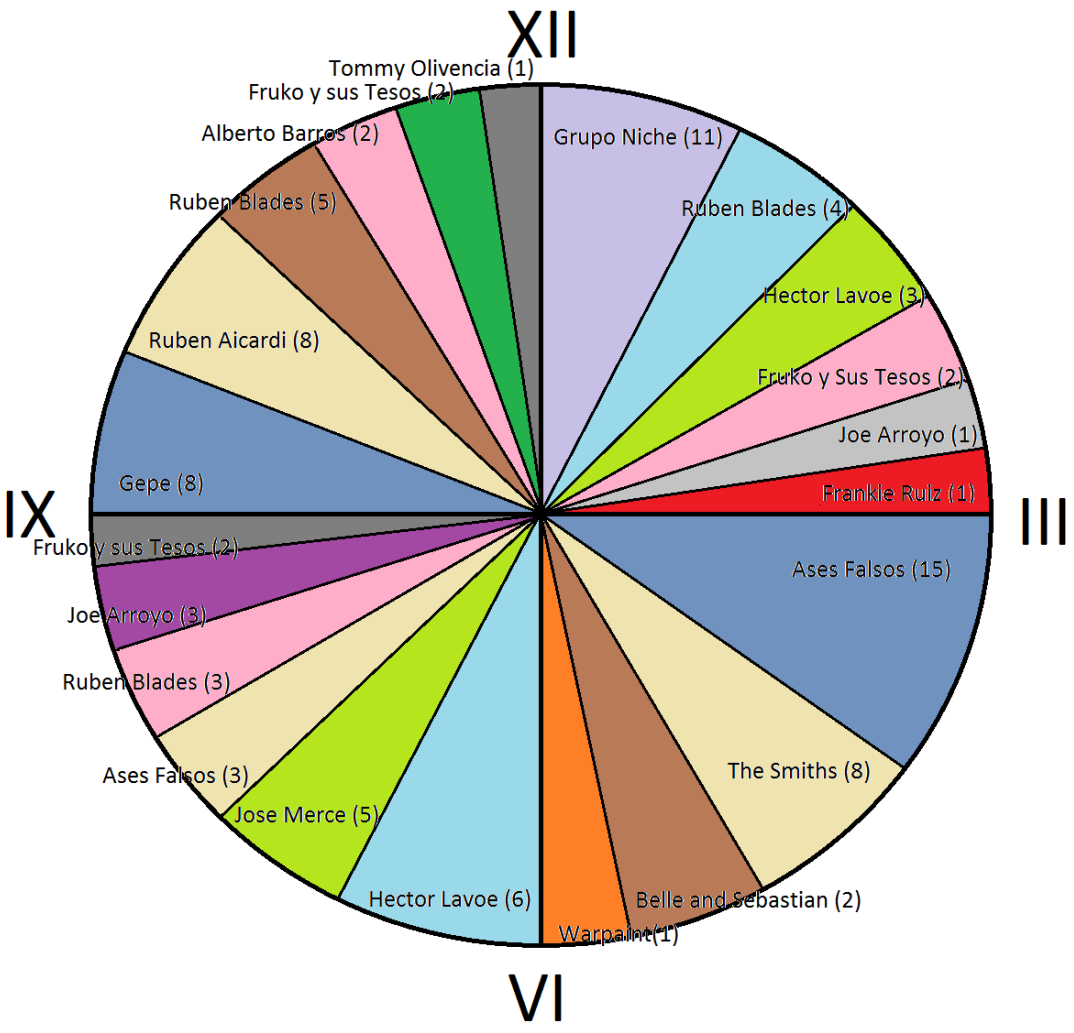


# The Music I listened to on Friday September 2nd

## From 12 pm to 12 am

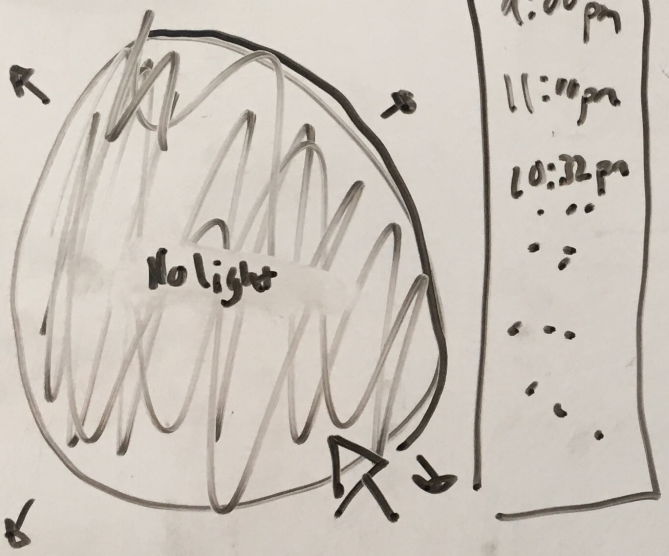
Split into 4 hour chunks, the artists I listened to are listed in order according to the number of songs played by each artist (in parentheses)

(not to scale)



Sketch 1

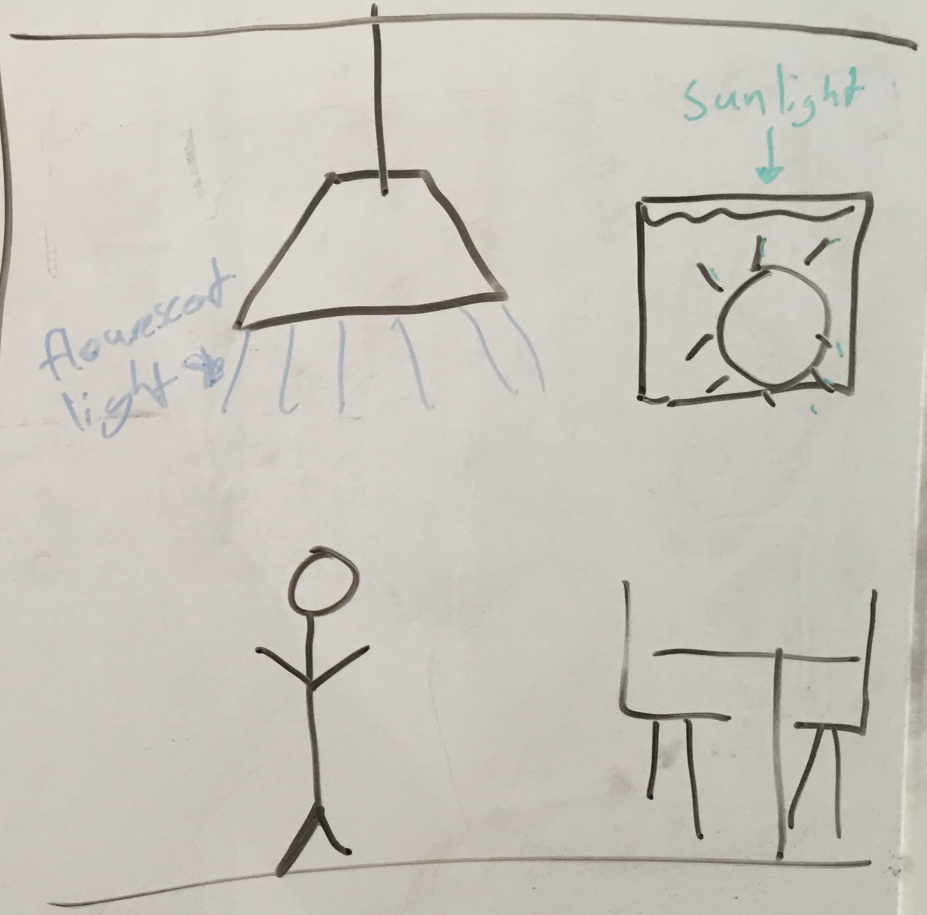
Times recorded as no light




9:00 pm
11:00 pm
10:32 pm

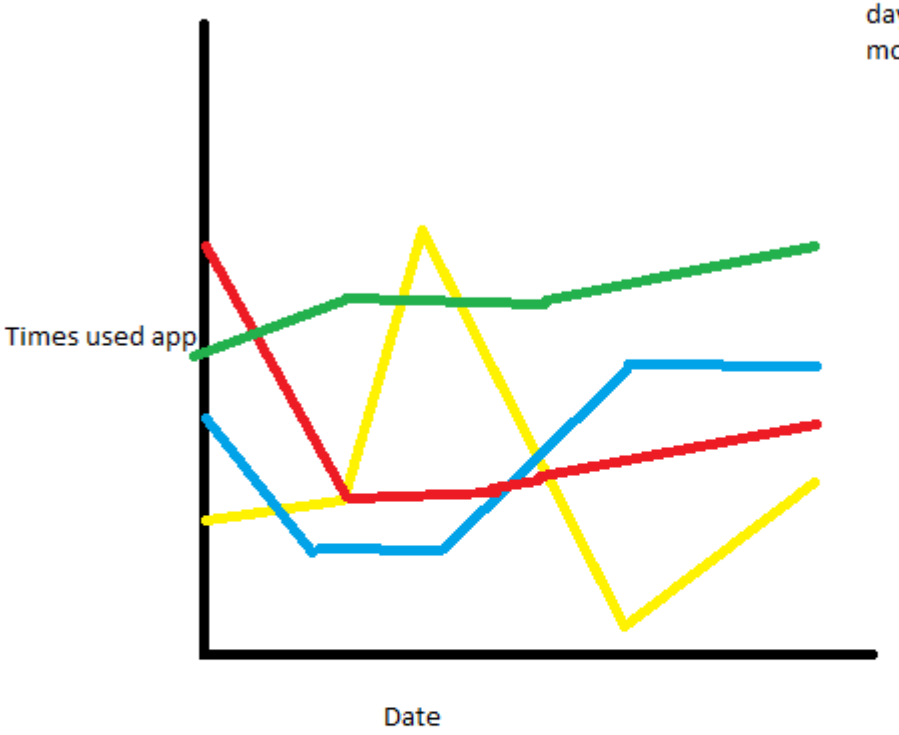
Switch Source

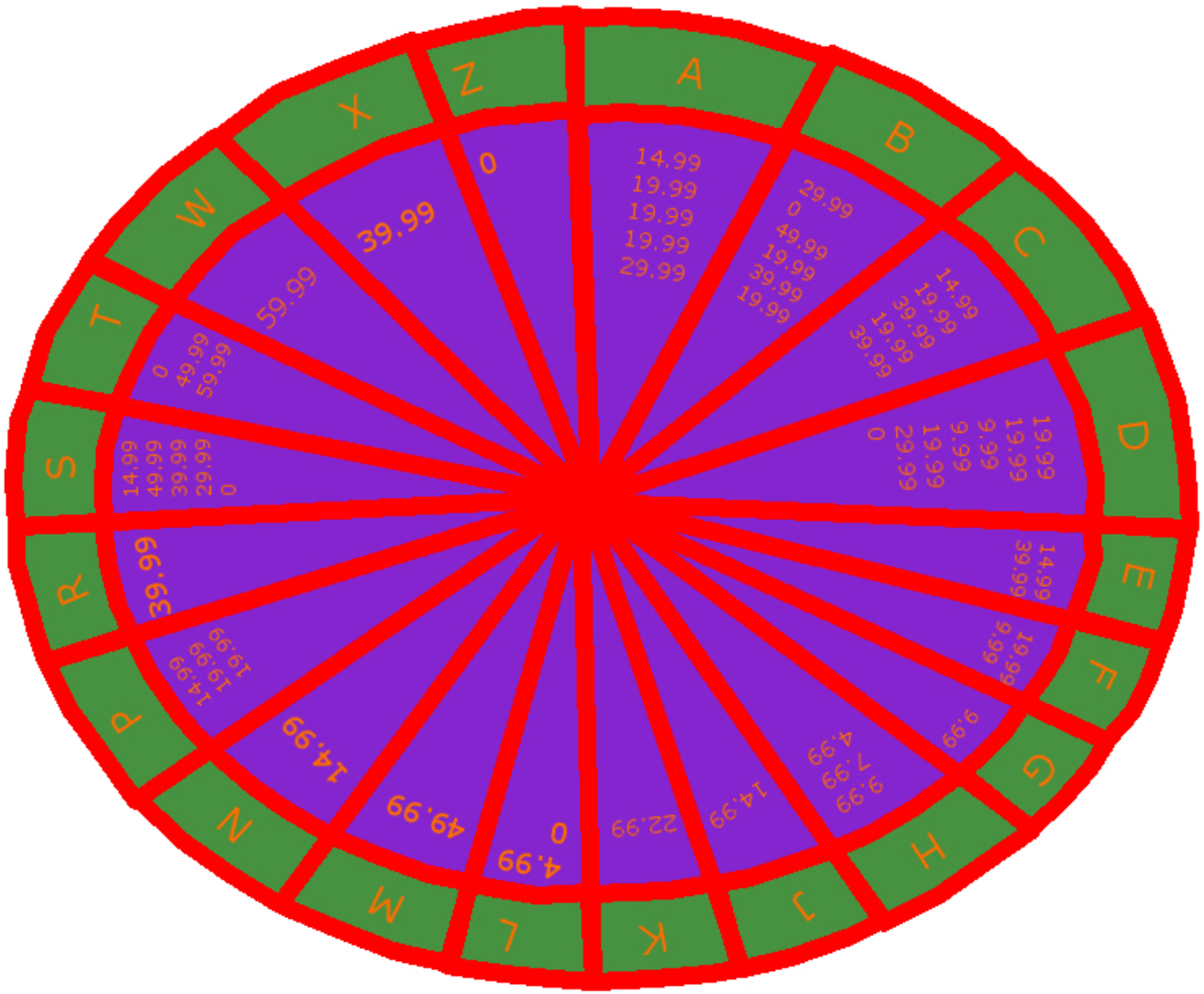
Sketch 2

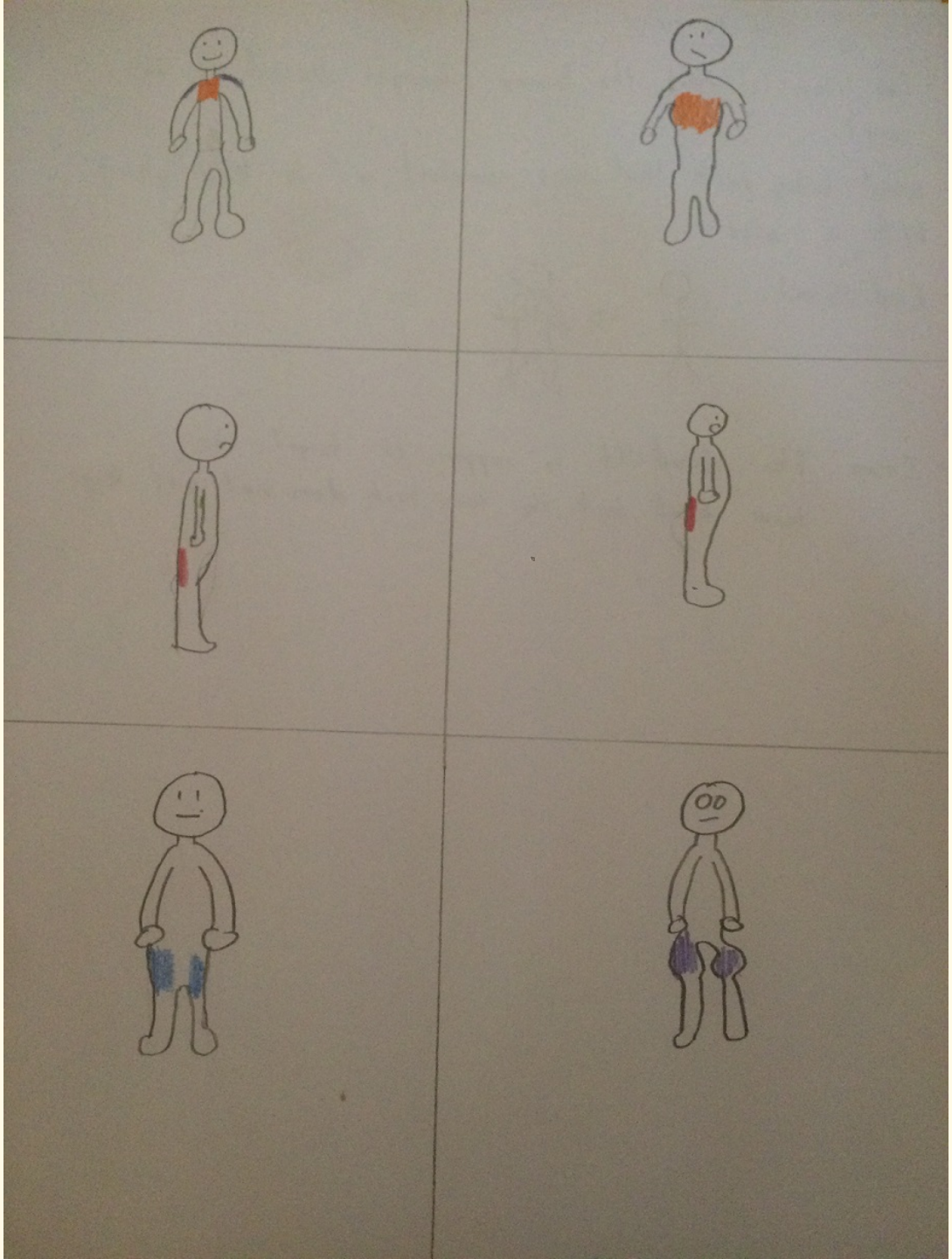


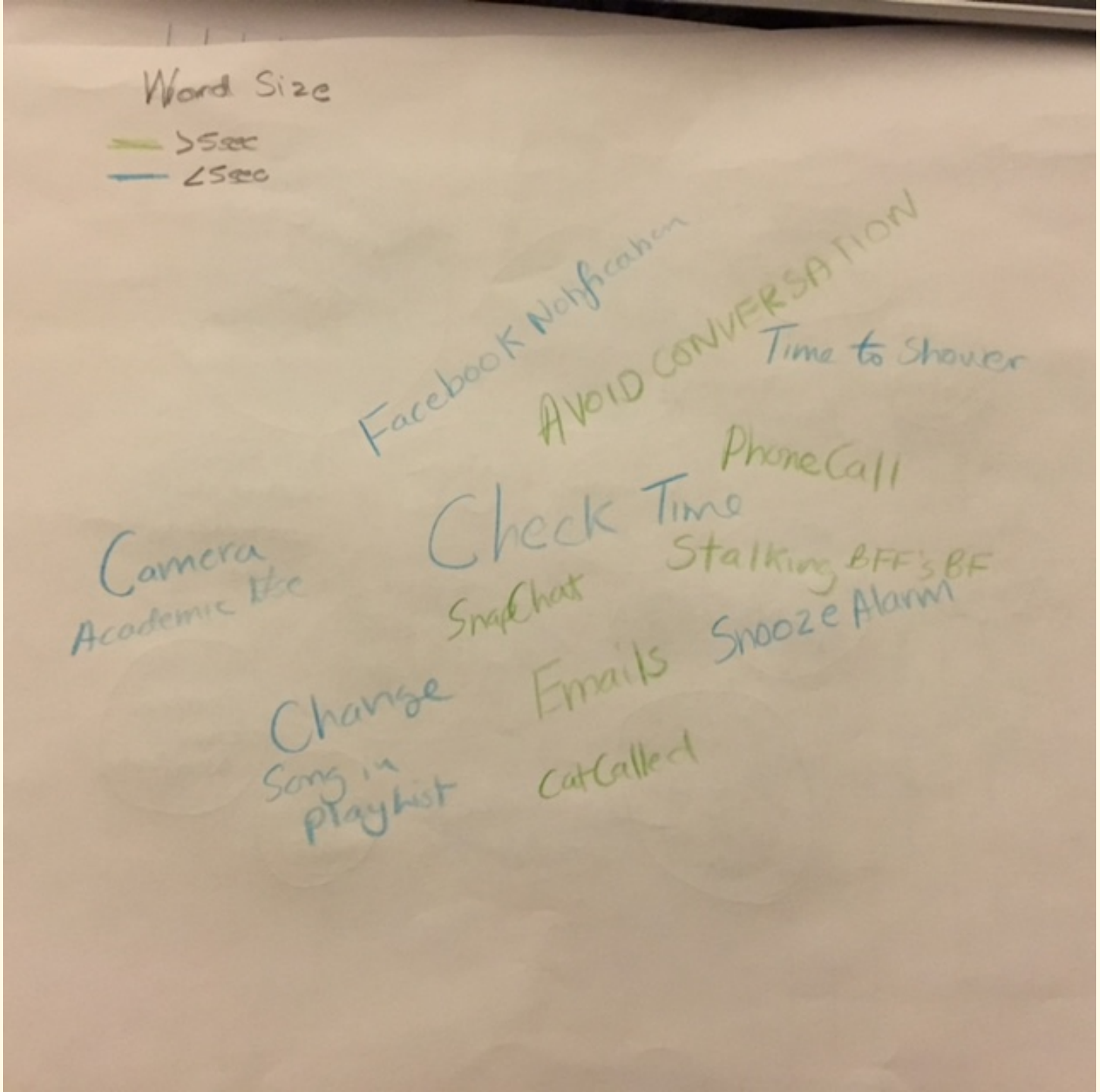

Swith Sources

This presentation will clearly show which day I use what app the most.




A
B
C
D
E
F
G
H
J
K
L
M
N
P
R
S
T
W
X
Z
14.99
19.99
19.99
19.99
29.99
29.99
0
49.99
19.99
39.99
19.99
14.99
19.99
39.99
19.99
39.99
0
29.99
19.99
9.99
9.99
9.99
19.99
39.99
14.99
19.99
9.99
9.99
7.99
4.99
22.99
14.99
0
4.99
49.99
14.99
14.99
19.99
59.99
0
49.99
59.99
14.99
49.99
39.99
0
29.99
39.99
59.99
39.99




Word Size
>5sec
<5sec
Facebook Notification
AVOID CONVERSATION
Time to Shower
PhoneCall
Check Time
Camera
Academic Use
Stalking BFF's BF
SnapChat
Change
Song in
Playlist
Emails
Snooze Alarm
CatCalled

# Thoughts on Grading Projects

Your work determines how I grade. If I can't find anything that seems interesting or unique or challenging – if there's no *direction* or *intent* to your work – then it is difficult to grade.

I try to put myself in the eyes of a potential employer or committee member, who would do the same thing: Assess your interests and abilities (and probably in a much more judgmental way!).

# Thoughts on Grading Projects

The grades are not meant to be punitive or even judgmental. They are a type of feedback to help guide you to where you could push yourself further.

Think of your grade and any comments is more like a critique. In *any* project, there is room to expand and to explore further. And a much better way to get jobs, scholarships, etc, is to have a strong portfolio of work (code, projects, etc) – rather than relying on GPA.

# Thoughts on Grading Projects

Of course, if you don't have experience in visualization, you may not great marks on your first project. However, by the end of the class, after 15 weeks, your projects hopefully will be much more interesting and accomplished.

Your final grade is heavily dependent on getting better throughout these 15 weeks.

# What is Vis, and Why do it?

Visualization design *augments* human capabilities, leveraging our perceptual and cognitive abilities to enhance the process of reasoning.

# What is Vis, and Why do it?

Despite the numerous advances in technology, many problems:

- require a person to frame the problem, the task to be solved
- require human intervention to make sense of data
- require integration with other information sources
- require expert knowledge to explore the parameter space

# What is Vis, and Why do it?

And yet...

- very helpful to automatically keep track of data, especially large datasets

- to parse, curate, sort, filter, highlight, annotate, explore, share data, etc.

- take subsets of data and pass along it to other software / analysis processes

- to handle updates, edits, dynamics of data

# What is Vis, and Why do it?

Innovation in human culture relies largely on the use of external representations:

- Writing
- Charts / plots / graphs
- Maps
- Tables / matrices / spreadsheets
- Drawing / sketching
- Blueprints
- Photographs

# What is Vis, and Why do it?

Even common visual representations of data were *invented*.

- Tree diagrams: ~1150
- Line chart: ~1700
- Bar chart:  ~1780
- Scatter plot: ~1785
- Pie chart: ~1801
- Flow maps: ~1859
- Tree maps: ~1992

# What is Vis, and Why do it?

Different visualizations were *created* to fulfill a purpose, often for a particular dataset, and then, when effective, re-used for many other datasets.

Interactive data visualization allows us to develop new ways to *externalize* and thus to *reason* about all types of data.

# What is Vis, and Why do it?

Computers, data mining algorithms, layout algorithms, display technology, HCI techniques, immersive technology, VR/AR, and eventually Brain-Computer interfaces, etc. ...

... give us an infinite palette from which to create interactive externalizations that can help us reason, to find or highlight patterns in the world and in the data we collect to make sense of those patterns.

# What is Vis, and Why do it?

The complexity of our world *demands* new representations.

Advances in the data collecting via scientific instrumentation vastly exceeds our ability to make sense of it.

# What is Vis, and Why do it?

Roger Malina says that at the dawn of the information age there has been an "epistemological inversion" with regards to scientific practice.

Before: limited amounts of data; obvious ways to analyze it

Now: immense amounts of data; unclear how best to analyze

The practice of *all science* is data science.

# What is Vis, and Why do it?

Examples:

- up to 1,000,000 proteins in a single human cell, and over 2.7 trillion cells in the human body.

- over 100,000,000,000 neurons (100 billion) in the adult human brain, with 100,000,000,000,000 (100 trillion) different connections between them

- 10,000,000 nurse reports using EHR over the last 10 years

# What is Vis, and Why do it?

How do you visualize this data? What do you need to visualize it for?

What questions do you want to answer? How can you use this data to help you?

Understanding the *task* you want to accomplish will help you design the visualization

# What is Vis, and Why do it?

Ben Shneiderman wrote an early "manifesto" called The Eyes Have It about what makes a visualization good:

- Lets users **interact** to navigate the data
- Provides an overview of the data to provide **context**...
- ... Then lets the user get to "**details-on-demand**"
- Lets users **reorganize** the data to compare and analyze subsets of data

# What is Vis, and Why do it?

Although these are important principles, certain types of data make this difficult (and interesting).

- What if you can't get an overview?
- What if the data is dynamic?
- What if it's not clear how to compare different parts of the data?

# Your data projects

- What motivated you to choose this data?
- Why is it interesting to others?
- What does it tell you about yourself? About society?
- What visualization "task" does it try to fulfill?
- In what ways are you planning to represent the data?
- For each of these ways, what parts of the data are emphasized? Obscured?
- Do you have suggestions for your colleague on how to better represent this data?

# Vis topics?

Take a few minutes and think about and write down some research questions you have.

What kinds of datasets do they involve?

What visualization tools are used to display/ explore/analyze the data?

If you could design an ideal tool given any budget and technology, what would it look like and what would it do?

# To do:

- Coding portion of Project 1 is due next Monday
- Quiz on Thursday or next Tuesday (on D3.js and textbook chapters)

# What is Vis, and Why do it?

Visualization design *augments* human capabilities, leveraging our perceptual and cognitive abilities.

# What is Vis, and Why do it?

Despite advances in AI and ML, many problems:

- require a person to frame the problem, the task to be solved
- require human intervention to make sense of data
- require integration with other information sources
- require expert knowledge to explore the parameter space

# What is Vis, and Why do it?

And yet...

- very helpful to automatically keep track of data, especially large datasets

- to parse, curate, sort, filter, highlight, annotate, explore, share data, etc.

- take subsets of data and pass along it to other software / analysis processes

- to handle updates, edits, dynamics of data

# What is Vis, and Why do it?

Innovation in human culture relies largely on the use of external representations:

- Writing
- Charts / plots / graphs
- Maps
- Tables / matrices / spreadsheets
- Drawing / sketching
- Blueprints
- Photographs

# What is Vis, and Why do it?

Even common visual representations of data were *invented*.

- Tree diagrams: ~1150
- Line chart: ~1700
- Bar chart:  ~1780
- Scatter plot: ~1785
- Pie chart: ~1801
- Flow maps: ~1859
- Tree maps: ~1992

# What is Vis, and Why do it?

Different visualizations were *created* to fulfill a purpose, often for a particular dataset, and then, when effective, re-used for many other datasets.

Interactive data visualization allows us to develop new ways to *externalize* and thus to *reason* about all types of data.

# What is Vis, and Why do it?

Computers, data mining algorithms, layout algorithms, display technology, HCI techniques, immersive technology, VR/AR, and eventually Brain-Computer interfaces, etc. ...

... give us an infinite palette from which to create interactive externalizations that can help us reason, to find or highlight patterns in the world and in the data we collect to make sense of those patterns.

# What is Vis, and Why do it?

The complexity of our world *demands* new representations.

Advances in the data collecting via scientific instrumentation vastly exceeds our ability to make sense of it.

# What is Vis, and Why do it?

Roger Malina says that at the dawn of the information age there has been an "epistemological inversion" with regards to scientific practice.

Before: limited amounts of data; obvious ways to analyze it

Now: immense amounts of data; unclear how best to analyze

The practice of *all science* is data science.

# What is Vis, and Why do it?

Examples:

- up to 1,000,000 proteins in a single human cell, and over 2.7 trillion cells in the human body.

- over 100,000,000,000 neurons (100 billion) in the adult human brain, with 100,000,000,000,000 (100 trillion) different connections between them

- 10,000,000 nurse reports using EHR over the last 10 years

# What is Vis, and Why do it?

How do you visualize this data? What do you need to visualize it for?

What questions do you want to answer? How can you use this data to help you?

Understanding the *task* you want to accomplish will help you design the visualization

# What is Vis, and Why do it?

Ben Shneiderman wrote an early "manifesto" called The Eyes Have It about what makes a visualization good:

- Lets users **interact** to navigate the data
- Provides an overview of the data to provide **context**...
- ... Then lets the user get to "**details-on-demand**"
- Lets users reorganize the data to compare and analyze subsets of data

# What is Vis, and Why do it?

Although these are important principles, certain types of data make this difficult (and interesting).

- What if you can't get an overview?
- What if the data is dynamic?
- What if it's not clear how to compare different parts of the data?

# What is Vis, and Why do it?

In other words – there is lots of room for innovation in visualization:

- layout algorithms
- interaction techniques
- analysis techniques
- application to particular datasets & assisting “domain experts” with their tasks

# What is Vis, and Why do it?

Some Vis topics I think are interesting:

- Machine learning / unsupervised learning
- Social systems / social networks / socio-technical systems – (e.g. Why isn't there a good visualization of Facebook data or other social networks?)
- Integrating "hard science" and "soft sciences": sociology, psychology, political sciences.

# What is Vis, and Why do it?

Some Vis topics I think are interesting:

- How do design / aesthetic principles (form, composition, color, etc) help to augment transmission / memorability of data

- How can you effectively represent: bias, uncertainty, probability, interpretation, even ideology related to a dataset

# Your data projects

- What motivated you to choose this data?
- Why is it interesting to others?
- What does it tell you about yourself? About society?
- What visualization "task" does it try to fulfill?
- In what ways are you planning to represent the data?
- For each of these ways, what parts of the data are emphasized? Obscured?
- Do you have suggestions for your colleague on how to better represent this data?