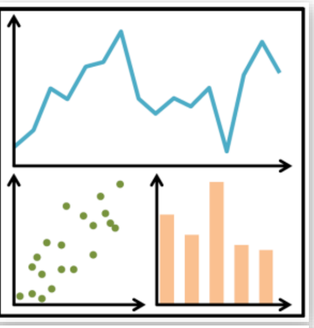
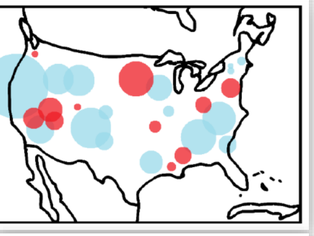
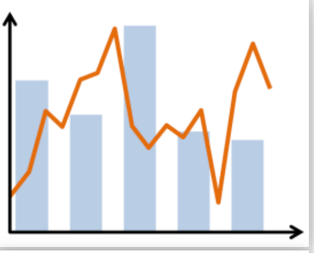
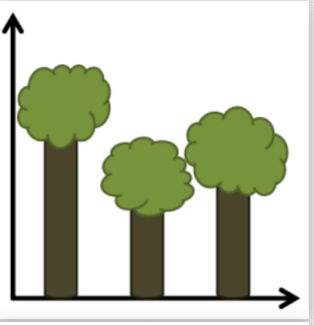
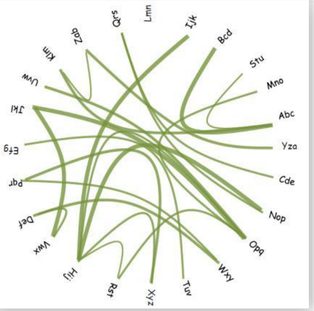






Lecture 2:
Design Rules of Thumb

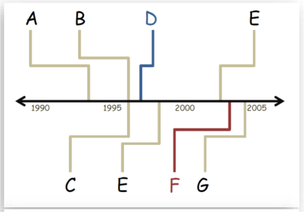
CS 7250

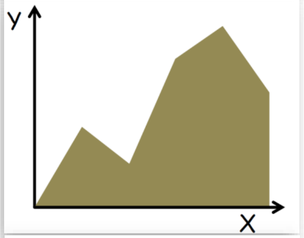
SPRING 2020

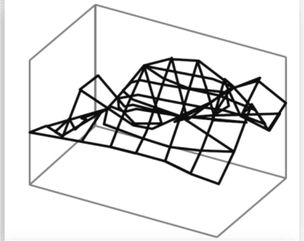
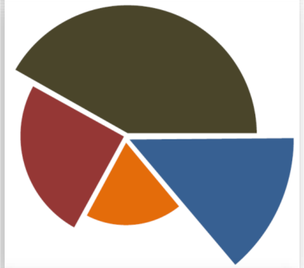
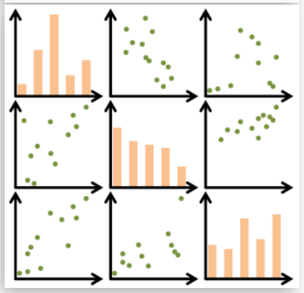
Prof. Cody Dunne

NORTHEASTERN UNIVERSITY

Slides and inspiration from Michelle Borkin, Krzysztof Gajos, Hanspeter Pfister, Miriah Meyer, Jonathan Schwabish, and David Sprague

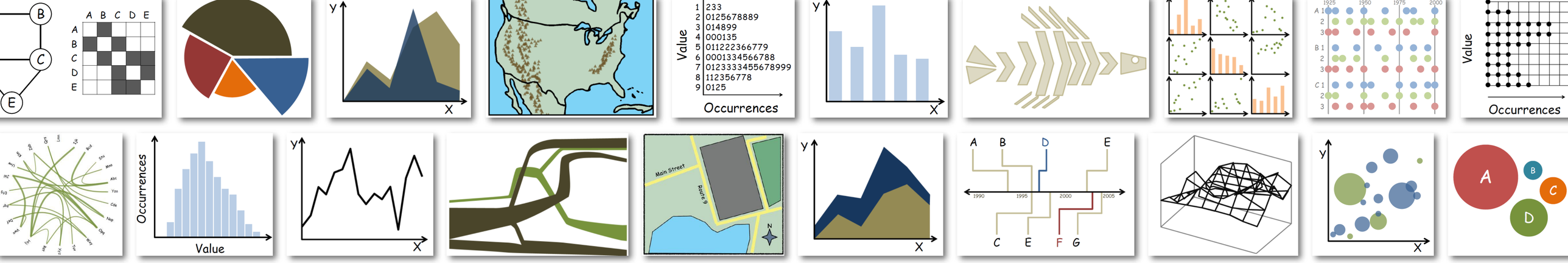




Previously, on CS 7250…

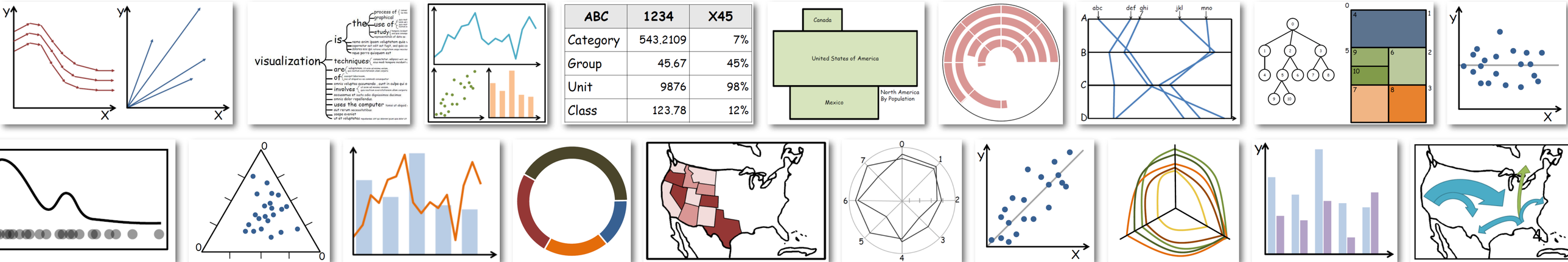
What is visualization anyway?

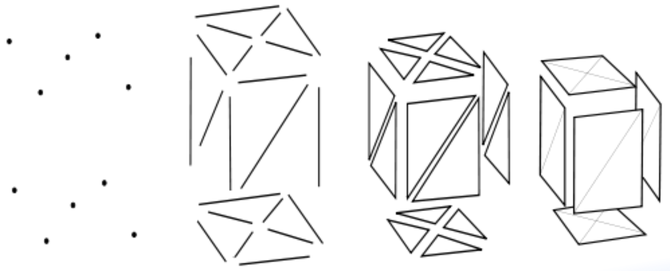


(static or interactive)

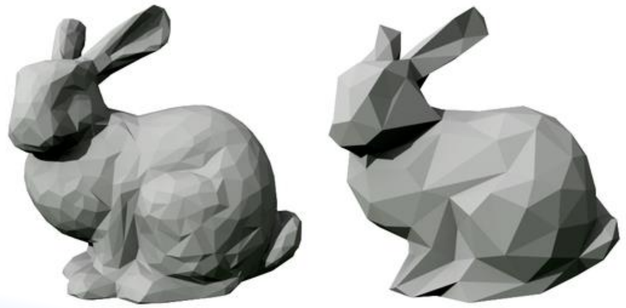
(abstract or spatial)

visualization: the visual representation of data to reinforce human cognition





computer graphics



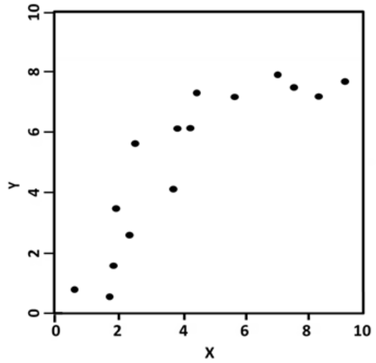
HCI

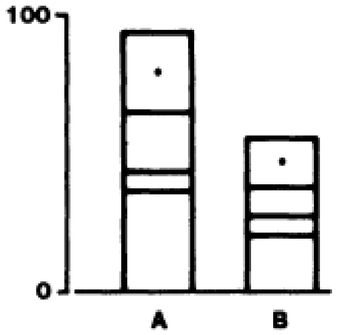
design

visualization

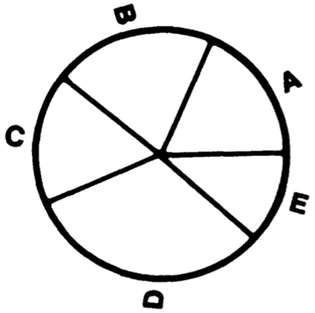
psychology

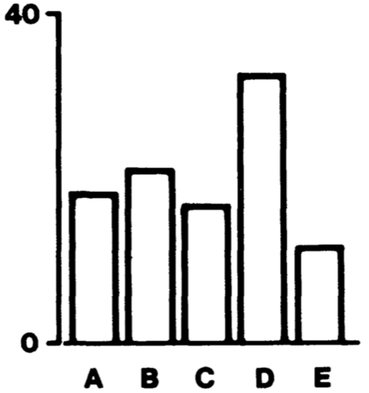
art





statistics





Staff Introductions

Cody Dunne
Assistant Professor
Instructor

Sara Di Bartolomeo, CS
PhD
TA

Gayathri Raj
UG
Service-Learning TA

Course Homepage

https://canvas.instructure.com/courses/1781732

- If you don't have an account on our Canvas yet: https://canvas.instructure.com/enroll/CMAPDM
- Use your name as known by the registrar and your @husky.neu.edu email.

Hall of Fame or Hall of Shame

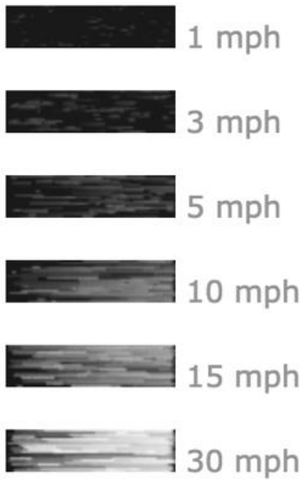
Prof. Krzystzof Gajos

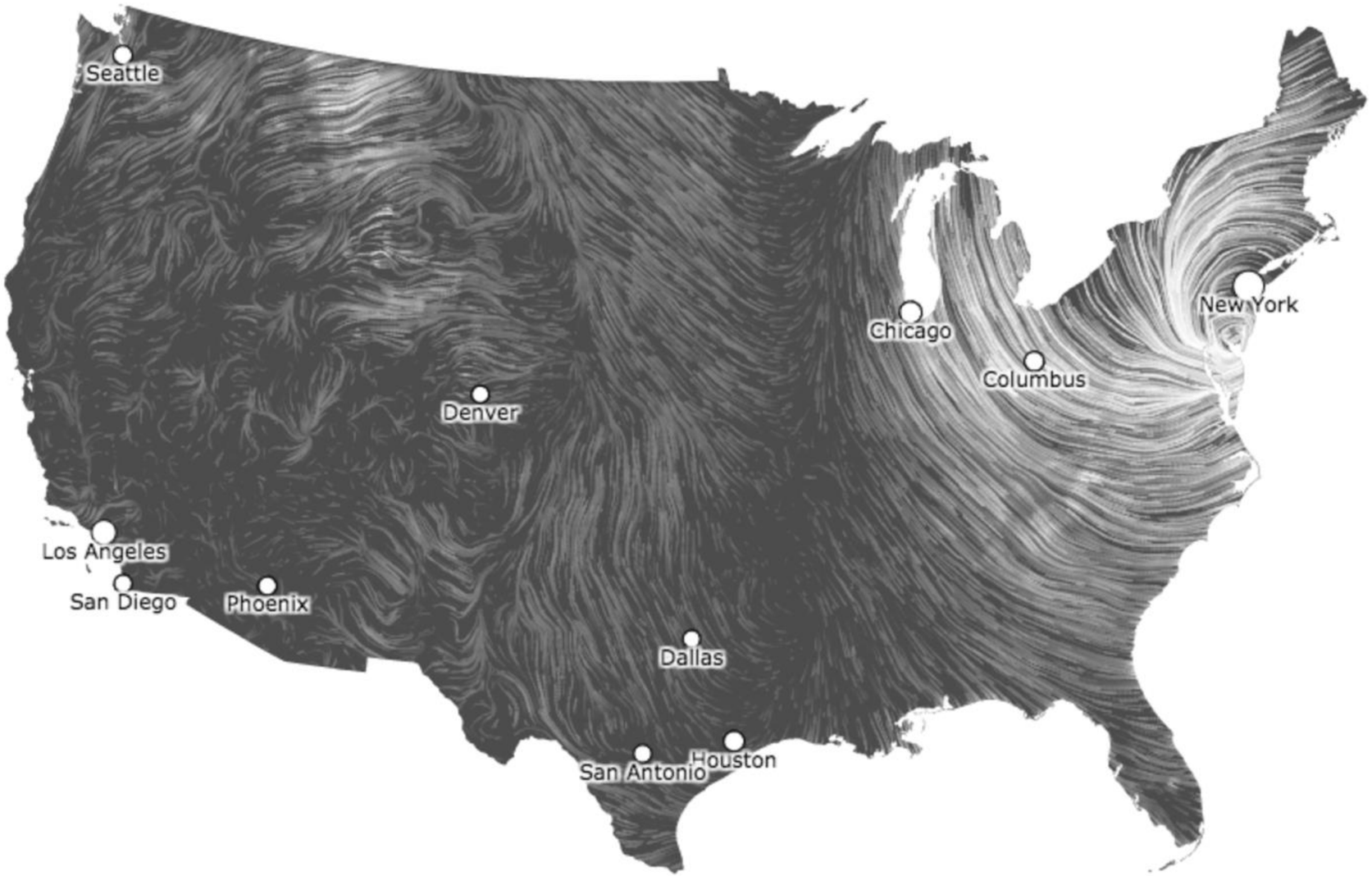
wind map

past patterns <previous next>

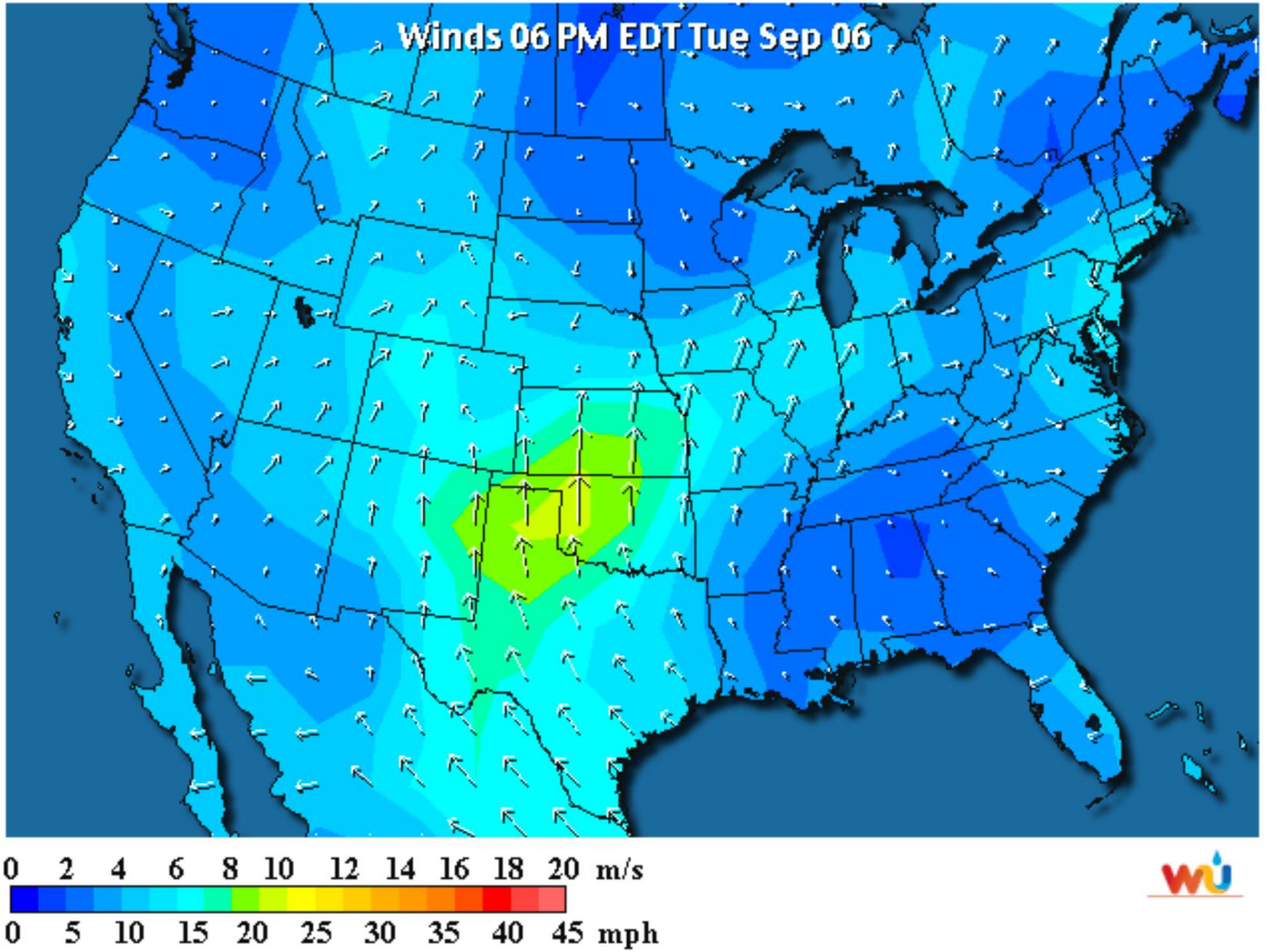
October 29, 2012
8:59 pm EST
(time of forecast download)

top speed: **45.1 mph**
average: **9.4 mph**



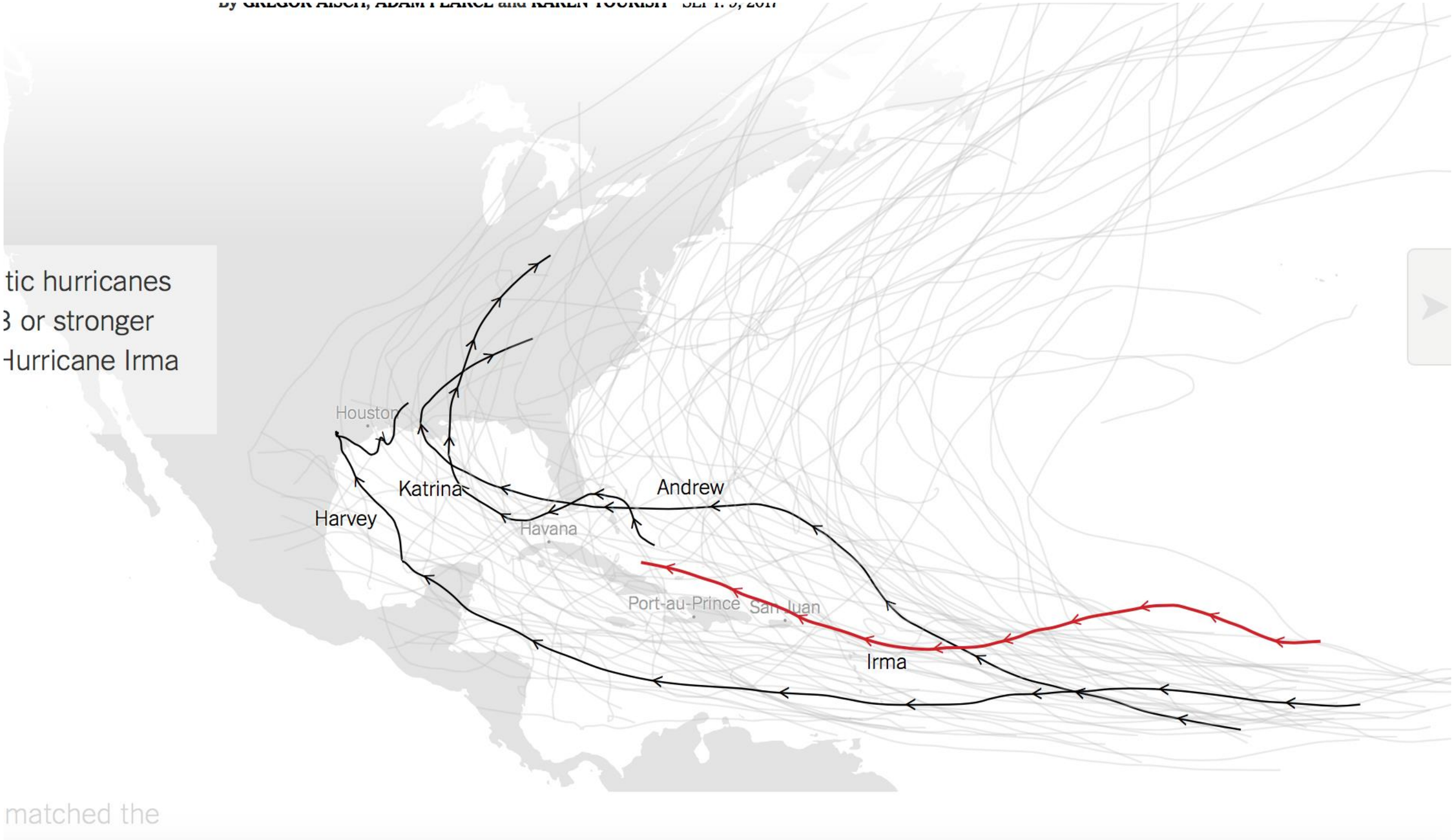


http://hint.fm/wind/gallery/oct-29.js.html


Winds 06 PM EDT Tue Sep 06
0 2 4 6 8 10 12 14 16 18 20 m/s
0 5 10 15 20 25 30 35 40 45 mph

By GREGOR AISCH, ADAM PEARCE and KAREN YOURISH SEPT. 9, 2017

tic hurricanes
3 or stronger
Hurricane Irma



matched the

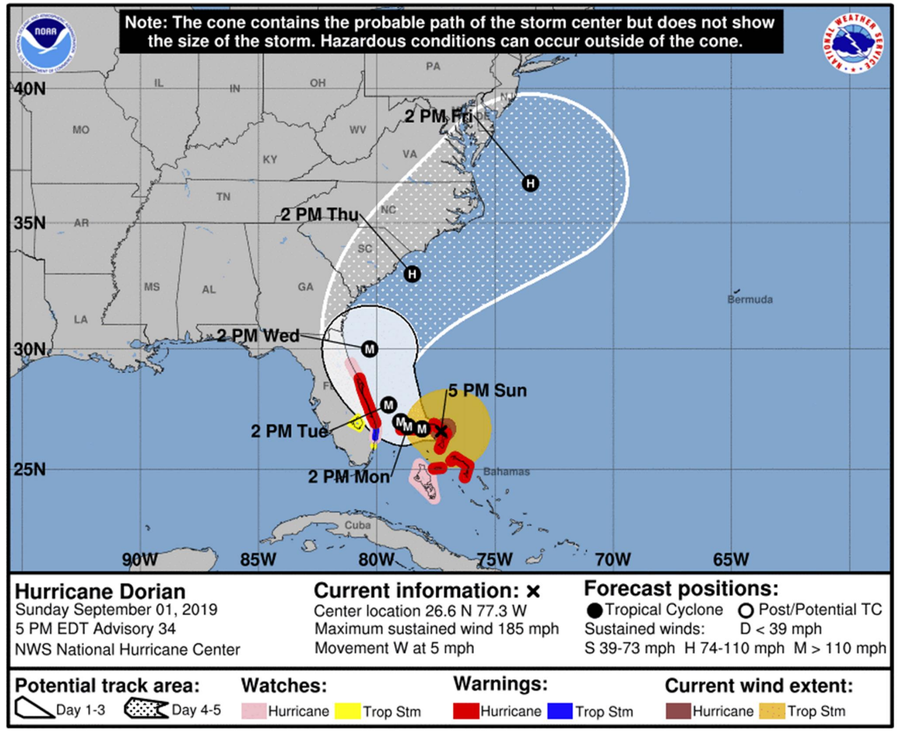
Kathryn Busemeyer - https://www.nytimes.com/interactive/2017/09/09/us/hurricane-irma-records.html

Matthew Kolosick - https://twitter.com/blkahn/status/905141924939649024/photo/1

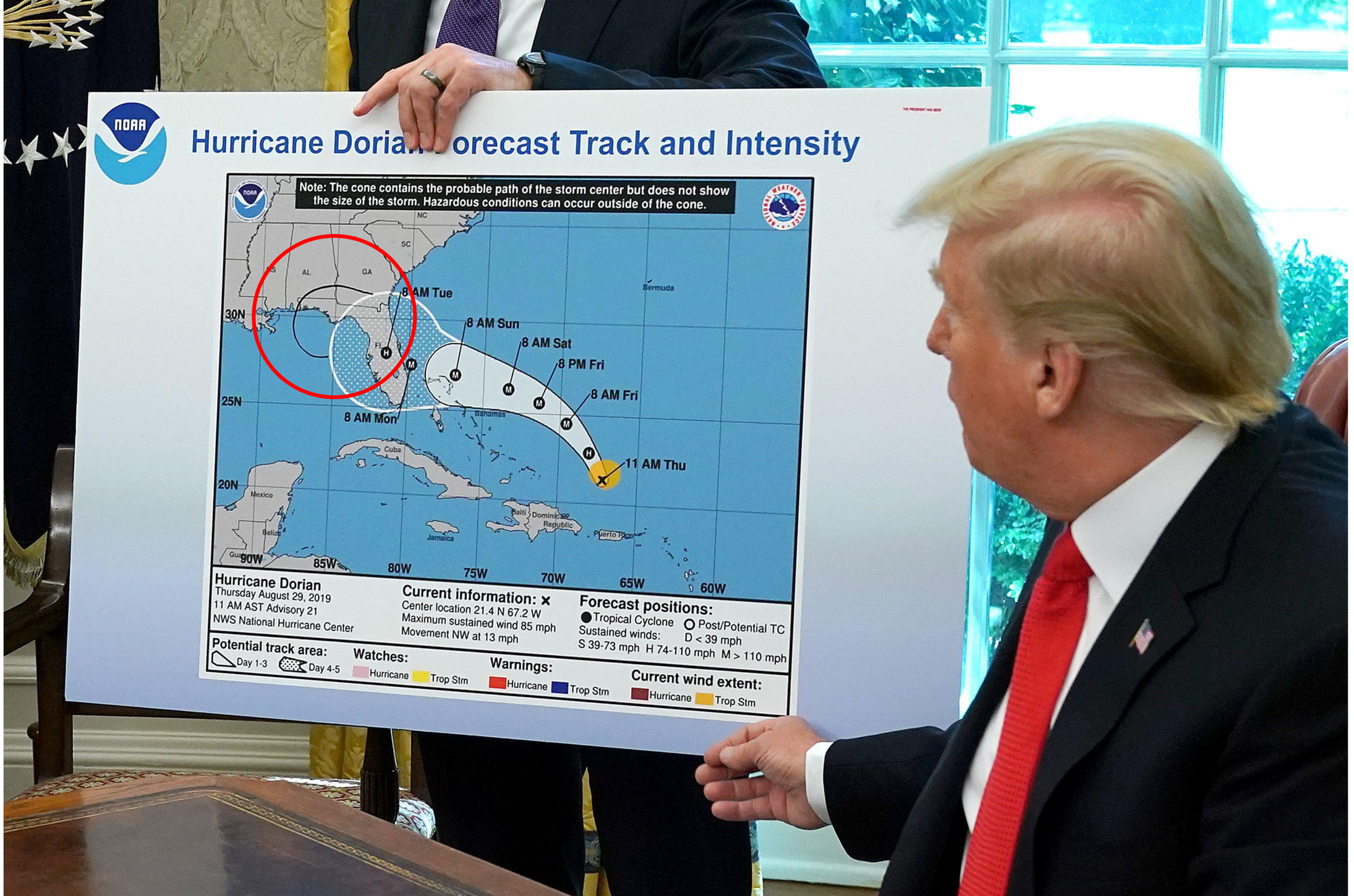
In-Class Redesign — Hurricane Funnels (continued…)

5 min

https://canvas.instructure.com/courses/1781732/assignments/13386302



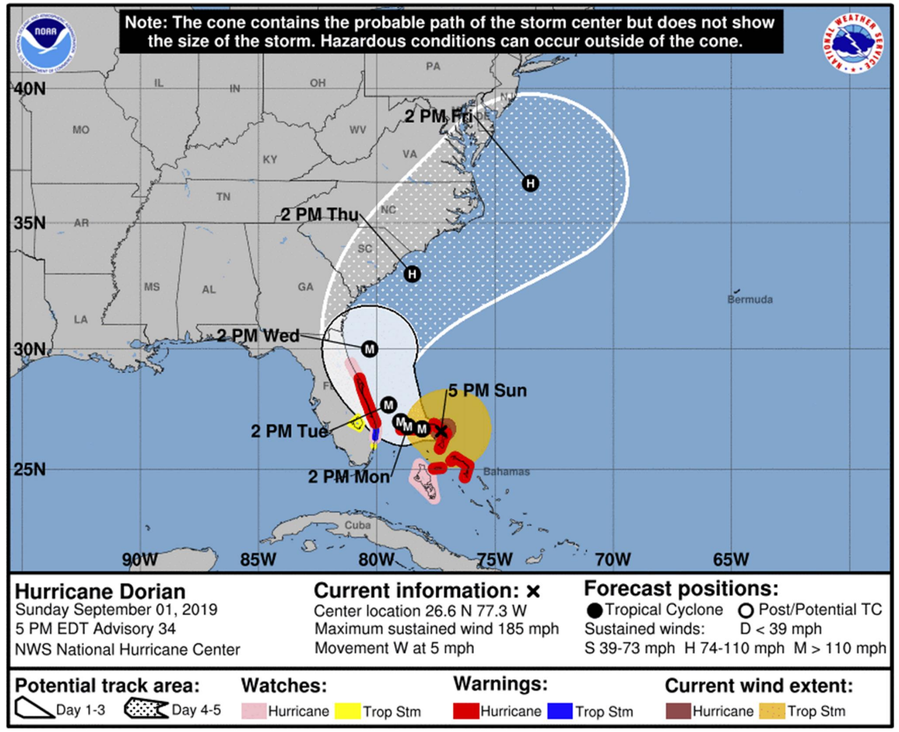
https://www.nhc.noaa.gov/refresh/graphics_at5+shtml/155815.shtml?cone#contents


Hurricane Dorian Forecast Track and Intensity
Note: The cone contains the probable path of the storm center but does not show the size of the storm. Hazardous conditions can occur outside of the cone.
NC
SC
AL
GA
FL
Bermuda
Bahamas
Cuba
Haiti
Dominican Republic
Puerto Rico
Jamaica
Mexico
Belize
Guatemala
30N
25N
20N
90W
85W
80W
75W
70W
65W
60W
8 AM Tue
8 AM Sun
8 AM Sat
8 PM Fri
8 AM Fri
8 AM Mon
11 AM Thu
Hurricane Dorian
Thursday August 29, 2019
11 AM AST Advisory 21
NWS National Hurricane Center
Current information:
Center location 21.4 N 67.2 W
Maximum sustained wind 85 mph
Movement NW at 13 mph
Forecast positions:
Tropical Cyclone
Post/Potential TC
Sustained winds:
D < 39 mph
S 39-73 mph H 74-110 mph M > 110 mph
Potential track area:
Day 1-3
Day 4-5
Watches:
Hurricane
Trop Stm
Warnings:
Hurricane
Trop Stm
Current wind extent:
Hurricane
Trop Stm

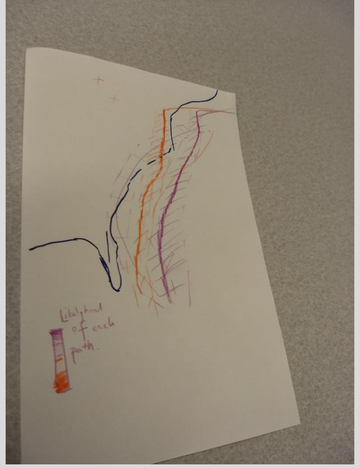
In-Class Redesign — Hurricane Funnels (continued…)

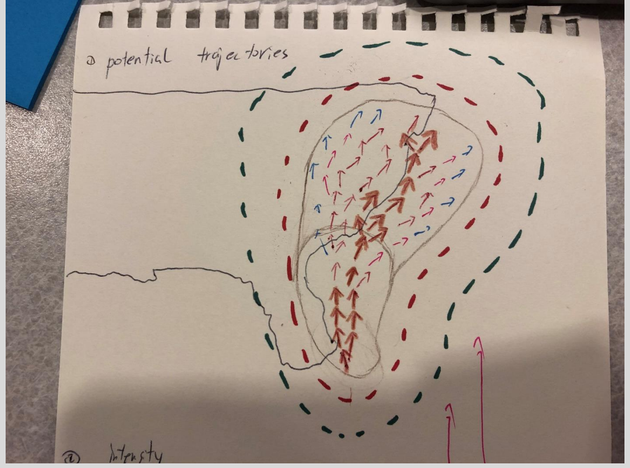
5 min

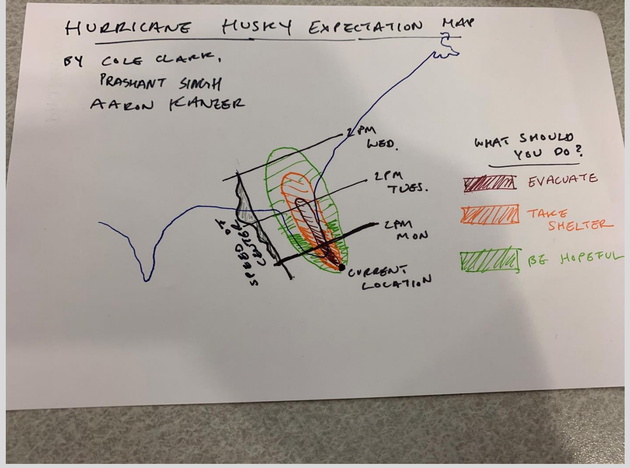
https://canvas.instructure.com/courses/1781732/assignments/13386302

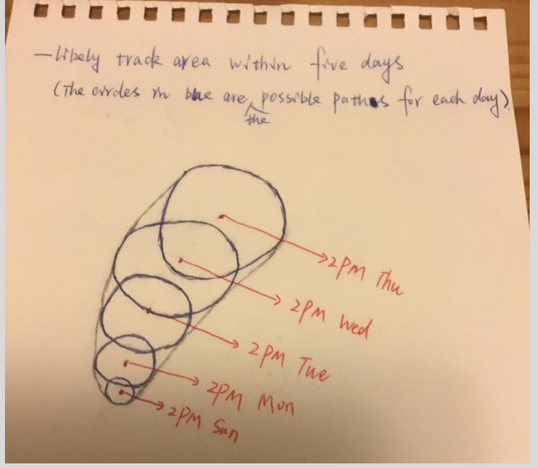


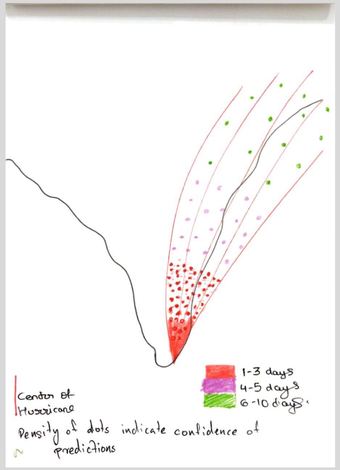
https://www.nhc.noaa.gov/refresh/graphics_at5+shtml/155815.shtml?cone#contents

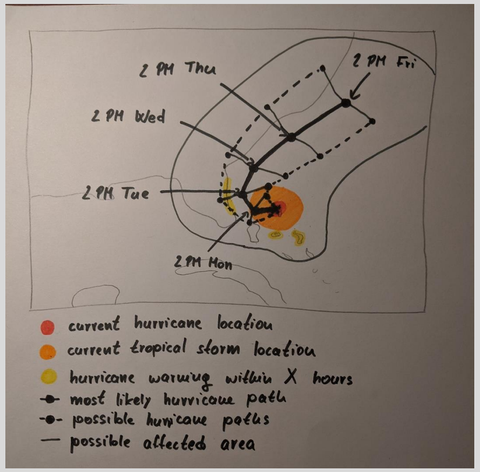

likelihood of each path

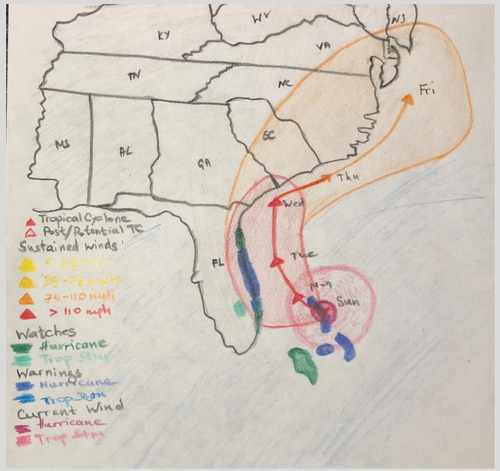

a potential trajectories
intensity


HURRICANE HUSKY EXPECTATION MAP
BY COLE CLARK,
PRASHANT SINGH
AARON KANZER
2PM WED.
2PM THURS.
2PM MON
CURRENT LOCATION
LEADING TO LANDFALL
WHAT SHOULD YOU DO?
EVACUATE
TAKE SHELTER
BE HOPEFUL

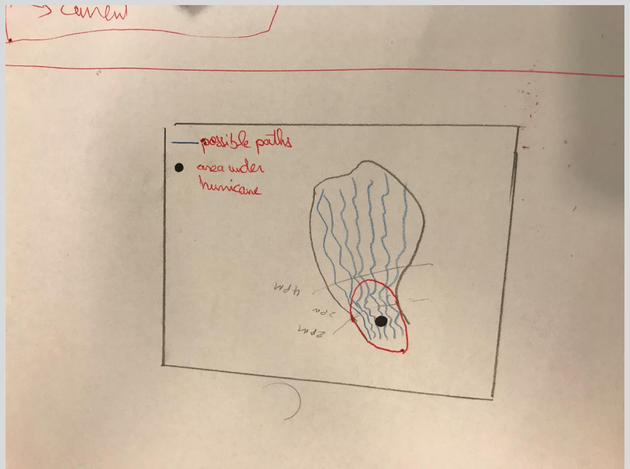

-likely track area within five days
(The circles in blue are possible paths for each day)
2PM Thu
2PM Wed
2PM Tue
2PM Mon
2PM Sun

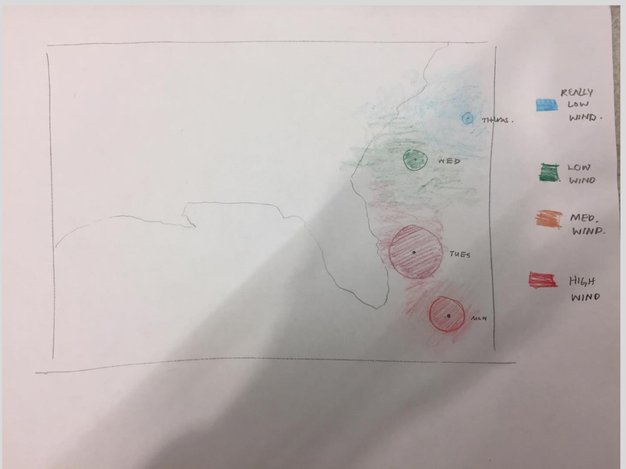

1-3 days
4-5 days
6-10 days
Center of Hurricane
Density of dots indicate confidence of predictions


2 PM Thu
2 PM Fri
2 PM Wed
2 PM Tue
2 PM Mon
current hurricane location
current tropical storm location
hurricane warning within X hours
most likely hurricane path
possible hurricane paths
possible affected area


KY
WV
VA
MD
TN
NC
MS
AL
GA
SC
FL
Fri
Thu
Wed
Tue
Mon
Sun
Tropical Cyclone
Post/Potential TC
Sustained winds:
39-73 mph
74-110 mph
> 110 mph
Watches
Hurricane
Warnings
Hurricane
Current Wind
Hurricane


Possibility of getting hit
Radius shows the intensity
High impact
Medium impact
Low impact

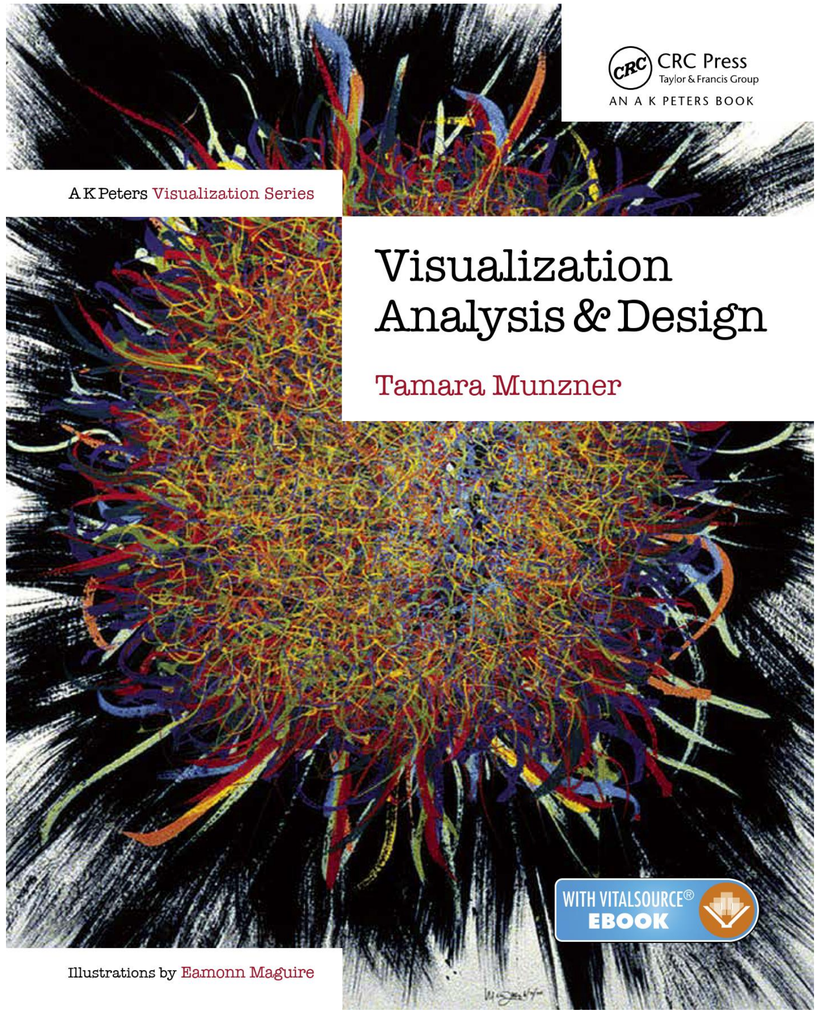

possible paths
area under hurricane


REALLY LOW WIND.
LOW WIND
MED. WIND.
HIGH WIND
WED
TUES

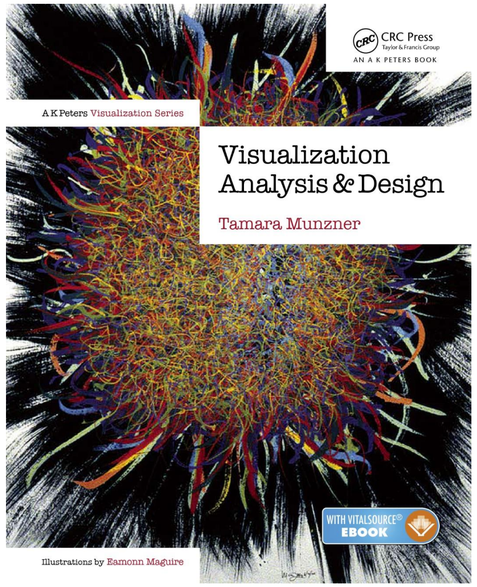
READING QUIZ

The Nested Model for Visualization Development

TEXTBOOK



Additional "recommended" books as resources in syllabus



“Nested Model”

Domain situation
Observe target users using existing tools

Example

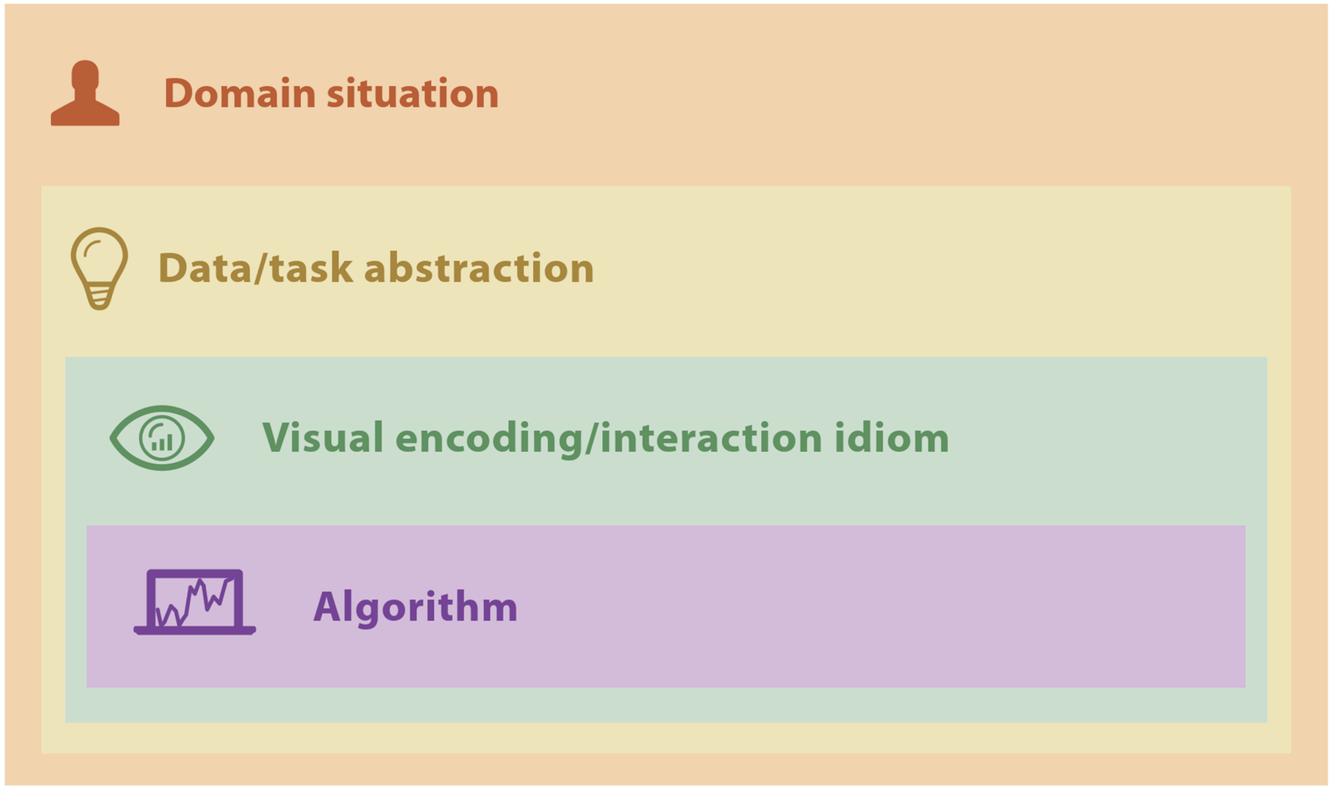
FAA (aviation)

What is the busiest time of day at Logan Airport?

Map vs. Scatter Plot vs. Bar

[Tamara Munzner]

Nested Model



Nested Model

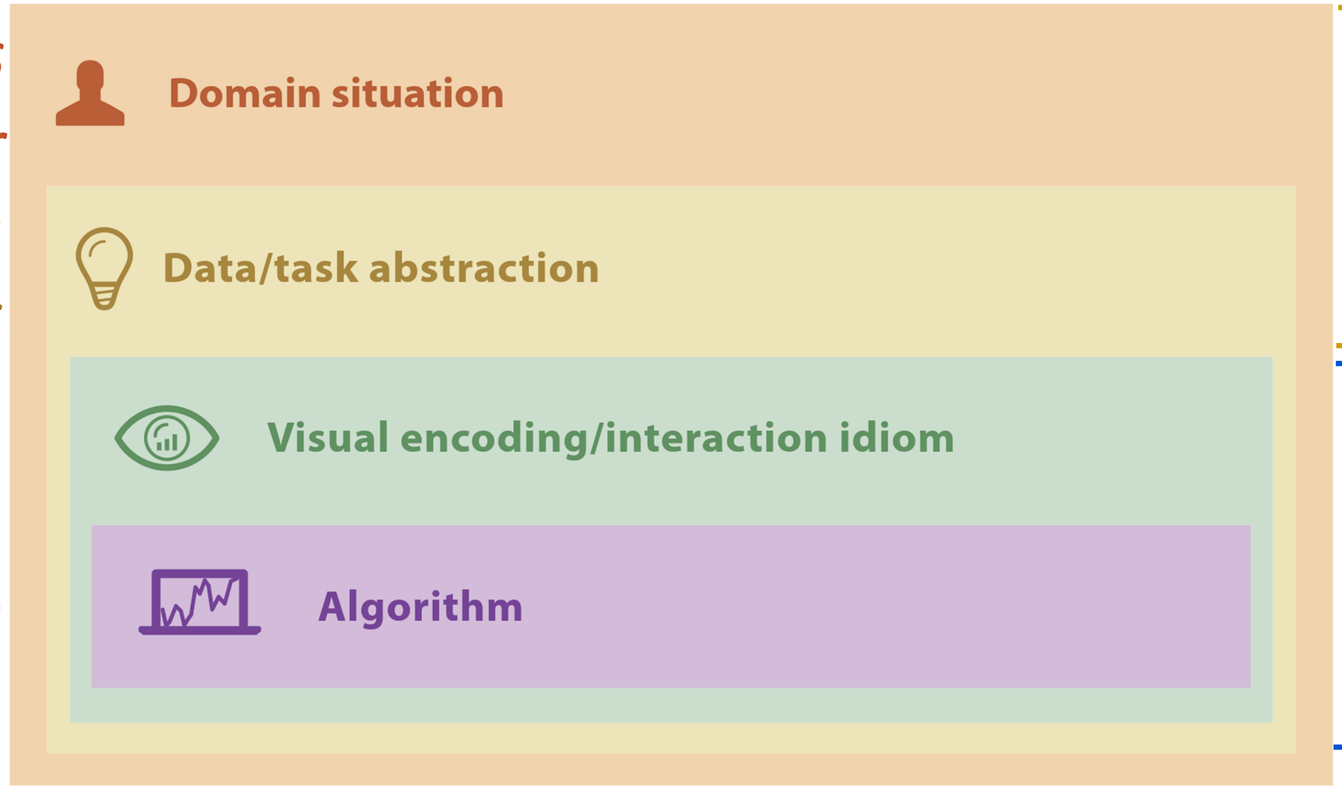
Human-centered design

Designer understands user

Abstract domain tasks

Visualization design

Implementation



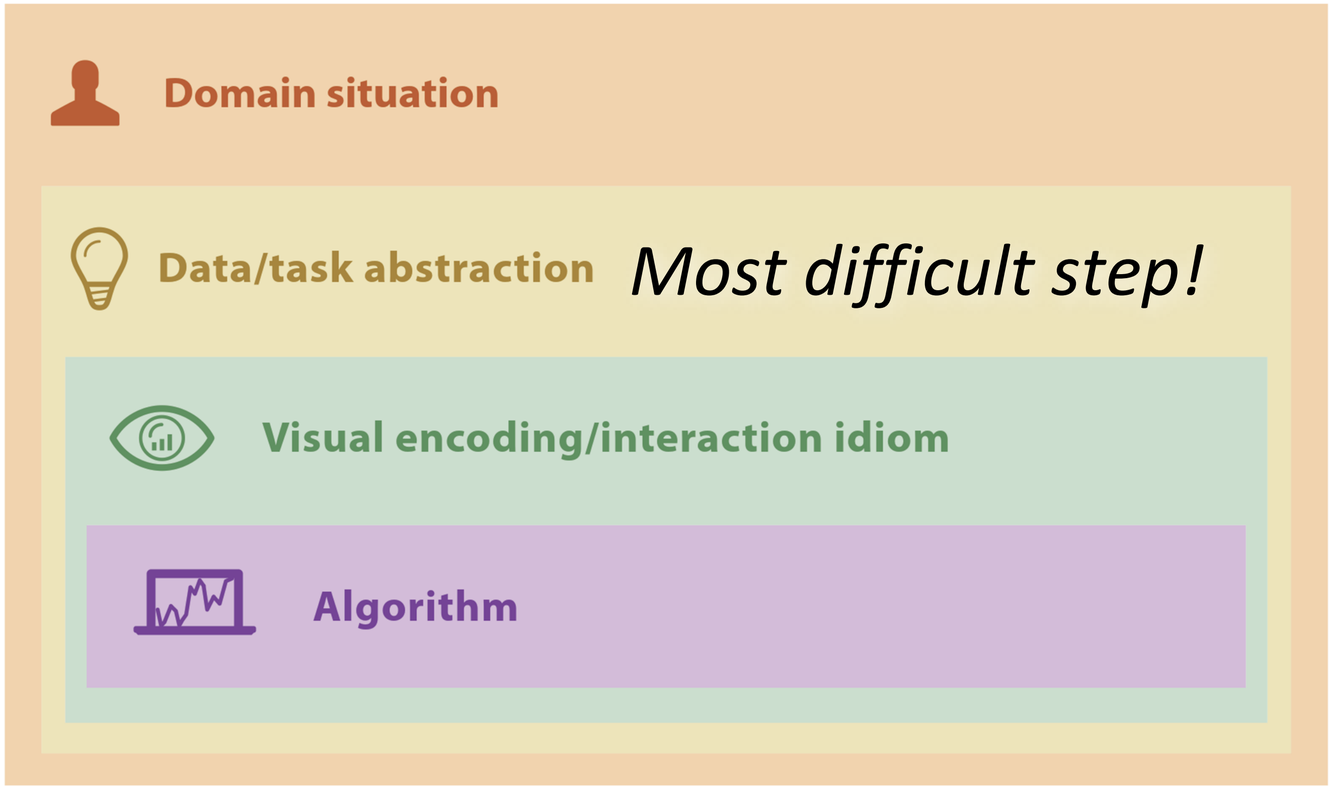
Identified

Designed

Nested Model

Design Study

TOP-DOWN
"problem-driven"

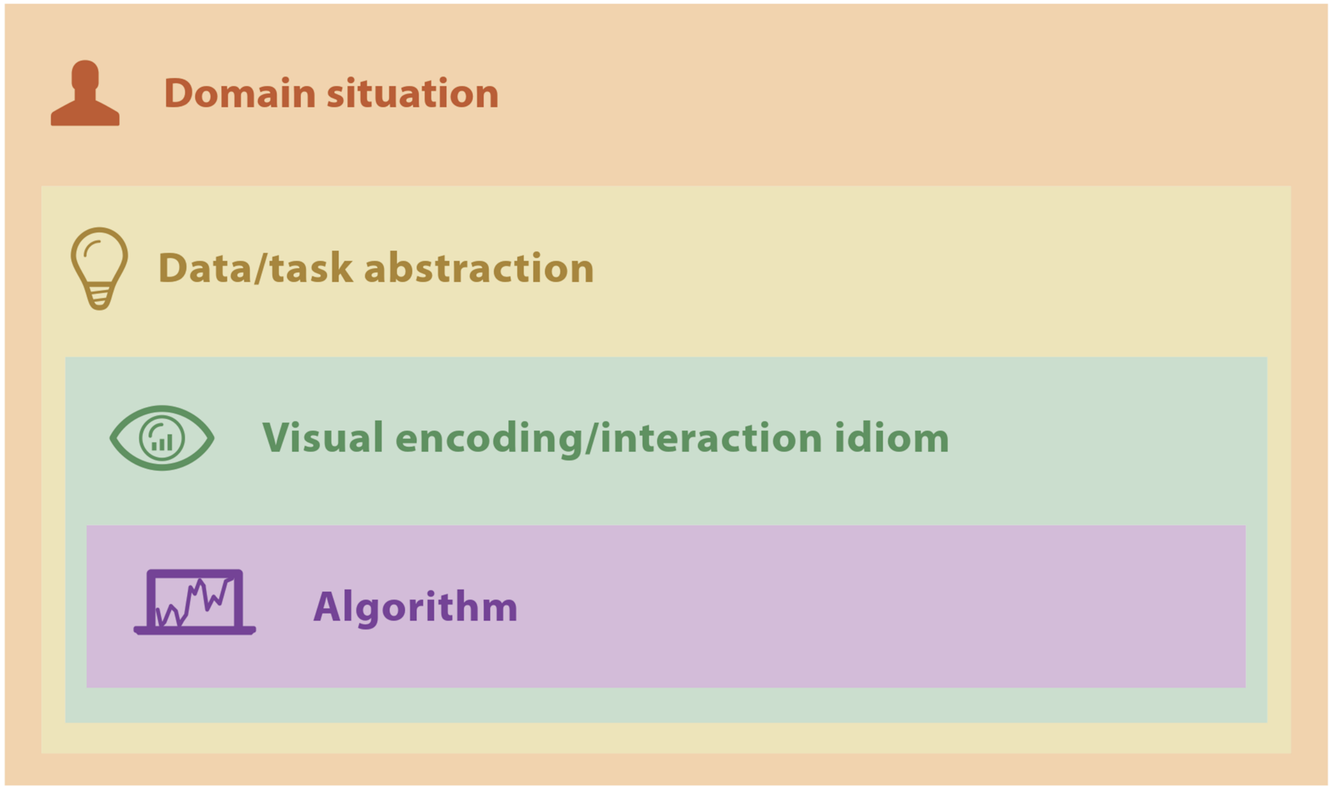


Technique

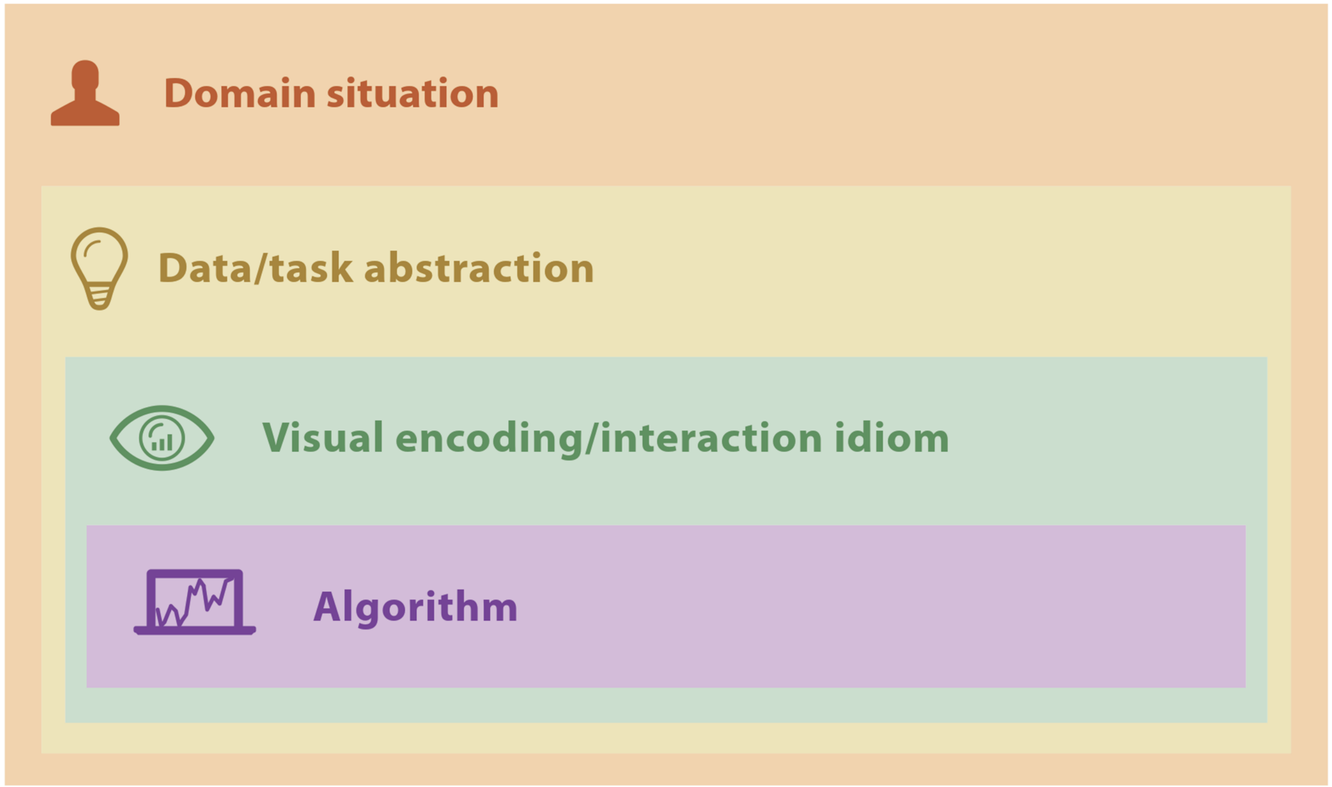
BOTTOM-UP
"technique-driven"

Nested Model

Mistakes propagate through model!



Threats to Validity



Threats to Validity

✓ *Final Project validation*

Domain situation

Data/task abstraction

Visual encoding/interaction idiom

Algorithm

Final project follow-up

PROJECTS

(Using the nested model via *design study "lite" methodology*)

Project Overview

The goal of the final project is to expose students to the real-world research and design process of doing an in-depth visualization design and implementation. As part of the project each group of ~3 students will need to work with real datasets, complete a task analysis, conduct a fact-finding interview, design and implement an interactive web-based visualization, solicit and incorporate feedback, and communicate the final project and results through a webpage, write a research paper, demo video, and in-class presentation. We will be following the Design Study "Lite" Methodology advocated by Borkin et al. (2017) .

Aside from the final project deadline itself, there will be a series of intermediate assignment deadlines incorporated in order to keep students on-track with the workload. Separate documents will be posted and provided to students with additional information and requirements for each assignment deadline. These assignments are worth a substantial portion of the student's final grade.

Detailed Requirements

Visualization

The interactive web-based visualization must:

- *Technology*
  - Be created using D3.
  - Not use dashboard generation tools for the final product, e.g., Tableau Web Player, R Shiny.
  - Be able to run on your web page.
  - The visualization is responsive, appropriately resizing to changes in the browser dimensions.
  - Be usable with both Chrome AND Firefox.
  - Your code passes the W3 validator.
  - There are no errors in the browser console in Firefox or Chrome.
  - The visualization should update within 1 second (ideal target 100ms) of any user interaction, even if to only display partial results, and remain responsive to user inputs during any calculations.
    - You may need to use parallelization e.g. Web Workers and advanced rendering e.g. WebGL to achieve this, and/or choose smaller datasets. Try to avoid needing WebGL if at all possible.
    - Your visualization will be evaluated on a desktop PC with these specifications: CPU: Intel Core i7-6700K @ 4GHz (4 cores, 8 threads), RAM: 32GB, GPU: NVIDIA GeForce GTX 1080, Network:

Service-Learning Projects

Why are we doing service learning?

Design Study "Lite" Methodology (Borkin et al. 2017)

- Design studies are a growing and valuable research area.
- Real-world data visualization experience.
- Visualization for exploration and communication.
- A more realistic experience of creating visualizations, and doing work in general.
- Teaches design, interview, evaluation, communication, and feedback techniques difficult to replicate in a classroom.
- Higher-stakes deliverables.
- Professional development.
- Make a positive impact in the community.
- Publication?

Service-Learning Projects

What are the challenges?

- Real-world data is messy and difficult to gather and process.
- Partners may not have clear goals and expectations.
- There is communication and scheduling overhead, inc. for teaching staff to differentiate assignment grading if necessary.
- Project areas may be too predefined.
- Project areas may be too ambiguous.
- May not actually make a meaningful impact.
- Reduces time for white-room technical education.
- More ambiguous expectations and grading challenges.
- Possible variation in student workload.
- Students may not know they are signing up for Service-Learning in advance (common problem with our registrar).

Service-Learning Projects

Who to blame for getting you into this?

Examples of Successful Course Projects

(Albeit with different requirements)

Project Example — CerebroVis

CerebroVis: Designing an Abstract yet Spatially Contextualized Cerebral Artery Network Visualization

Aditeya Pandey, Harsh Shukla, Geoffrey S. Young, Lei Qin, Amir A. Zamani, Liangge Hsu, Raymond Huang, Cody Dunne, and Michelle A. Borkin

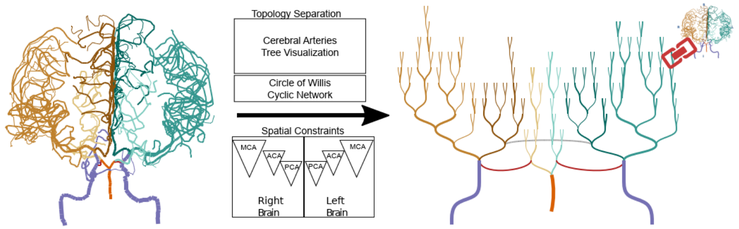


Fig. 1: CerebroVis is a novel network visualization for cerebral arteries. CerebroVis uses a abstract topology-preserving visual design which is put in spatial context by enforcing constraints on the network layout. Here we show the conversion of an almost symmetrical healthy human brain cerebral artery network from a 2D isosurface visualization (left) to CerebroVis (right). Each artery has the same categorical color in both views (see Sec. 3 for a legend).

Abstract—Blood circulation in the human brain is supplied through a network of cerebral arteries. If a clinician suspects a patient has a stroke or other cerebrovascular condition they order imaging tests. Neuroradiologists visually search the resulting scans for abnormalities. Their visual search tasks correspond to the abstract network analysis tasks of browsing and path following. To assist neuroradiologists in identifying cerebral artery abnormalities we designed CerebroVis, a novel abstract—yet spatially contextualized—cerebral artery network visualization. In this design study, we contribute a novel framing and definition of the cerebral artery system in terms of network theory and characterize neuroradiologist domain goals as abstract visualization and network analysis tasks. Through an iterative, user-centered design process we developed an abstract network layout technique which incorporates cerebral artery spatial context. The abstract visualization enables increased domain task performance over 3D geometry representations, while including spatial context helps preserve the user's mental map of the underlying geometry. We provide open source implementations of our network layout technique and prototype cerebral artery visualization tool. We demonstrate the robustness of our technique by successfully laying out 61 open source brain scans. We evaluate the effectiveness of our layout through a mixed methods study with three neuroradiologists. In a controlled experiment our study participants used CerebroVis and a conventional 3D visualization to examine real cerebral artery imaging data and to identify a simulated intracranial artery stenosis. Participants were more accurate at identifying stenoses using CerebroVis (odds ratio 2.5, absolute risk difference 13%). More broadly, we discuss the applications of our design approach to a general design paradigm we call *Abstraction with Context*. A free copy of this paper, the evaluation stimuli and data, and source code are available at osf.io/e5sxt

Index Terms—Network Visualization, Spatial Context, Abstract Design, Flow Network, Medical Imaging, Cerebral Arteries.

1 Introduction

Arteries in the human brain form a network of blood flow, and a blockage or leakage in this network can lead to life-threatening cerebrovascular conditions such as a stroke or aneurysm. Strokes alone are the fifth leading cause of death as well as a leading cause of serious long-term disability in the United States, and is globally the second leading cause of death after heart disease [32]. Early detection and diagnosis of these conditions is essential for effective life-saving treatment. Conventional diagnostics rely on an expert neuroradiologist identifying vascular abnormalities through examination of medical images (e.g., CTA, MRA). This data is commonly rendered in 3D in order to assist the doctor with identification of the abnormalities. However, prior research indicates that existing representations of the 3D cerebral arteries—e.g., isosurface, volume rendering, and Maximum Intensity Projection (MIPS)—introduce visual artifacts and task performance challenges such as overplotting/occlusion [19], false impression of geometry [19], and excessive artery bends.

In this design study, we present a novel 2D visualization of the cerebral artery system designed to assist doctors in the identification of cerebrovascular abnormalities. Inspired by existing visualization research which has demonstrated the effectiveness of 2D representations for spatial search tasks in other medical imaging cases, e.g., cardiovascular arteries [6] and connectomics [33], we present a novel 2D abstract representation of the cerebral arteries. To our knowledge, this is the first attempt to approach the cerebrovascular diagnostics tasks faced by neuroradiologists from the perspective of network science and using an abstract 2D visual encoding.

In this paper we first offer a novel framing of cerebral arteries using network theory. Next, we characterize the domain goals and present them as network analysis tasks. In an iterative user-centered design with

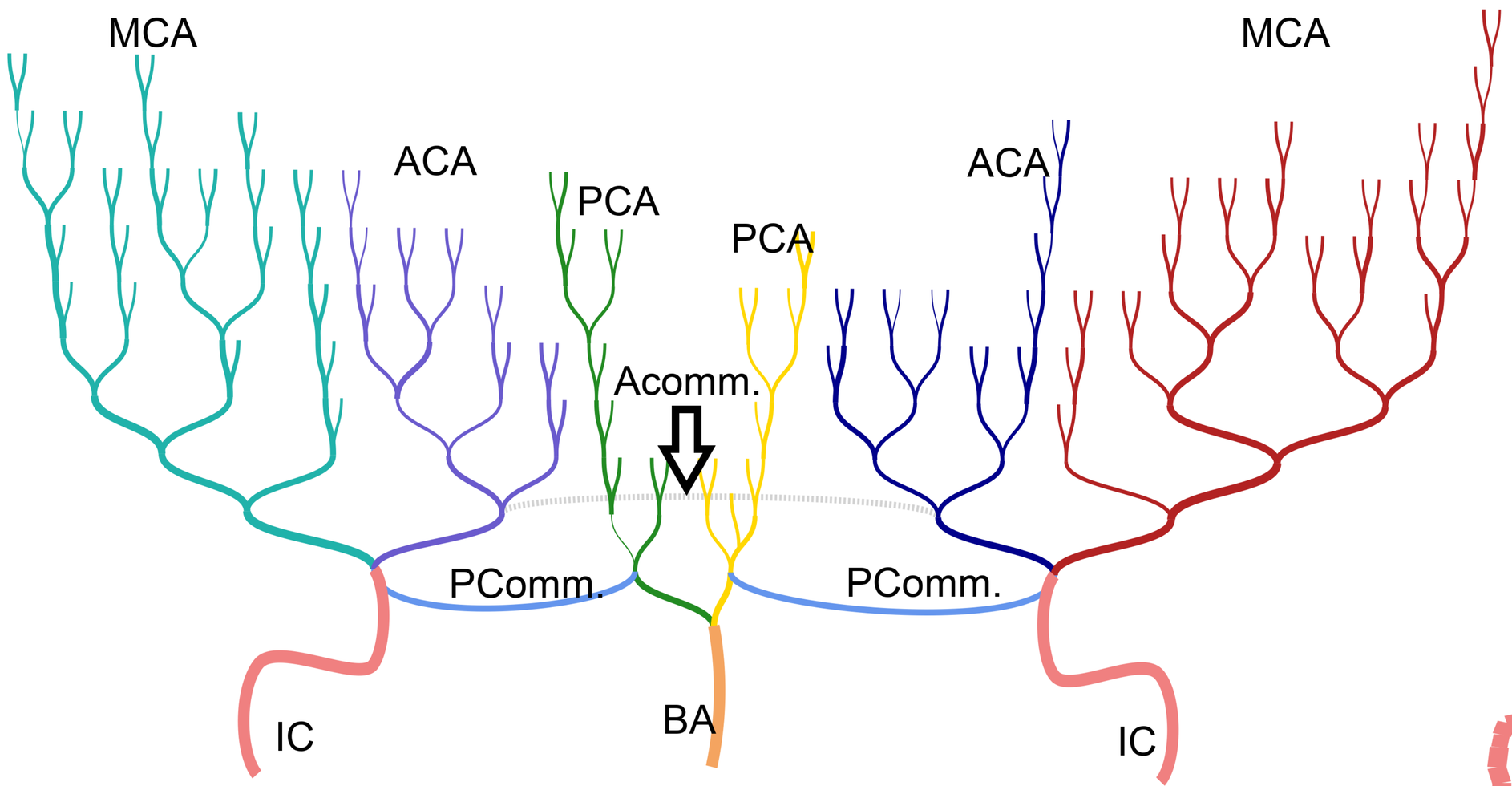
- *Aditeya Pandey, Harsh Shukla, Cody Dunne, and Michelle Borkin are with Northeastern University. E-mail: {pandey.ad, shukla.h}@husky.neu.edu, {c.dunne, m.borkin}@northeastern.edu*
- *Geoffrey S. Young, Amir A. Zamani, Liangge Hsu, and Raymond Huang are with Brigham and Women's Hospital. E-mail: {gsyoung, azamani, lhsu1, ryhuang}@bwh.harvard.edu*
- *Lei Qin is with the Dana-Farber Cancer Institute. E-mail: lqin2@partners.org*

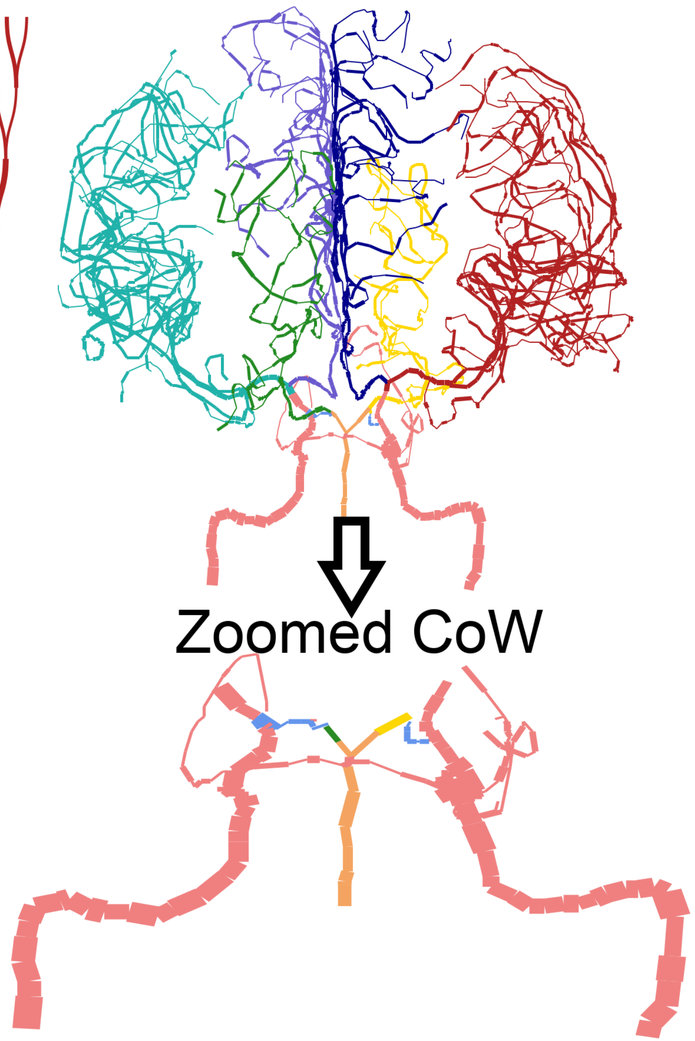
Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

CS 7260 Fall 2017:
Visualization for
Network Science

Pandey et al. VIS 2019

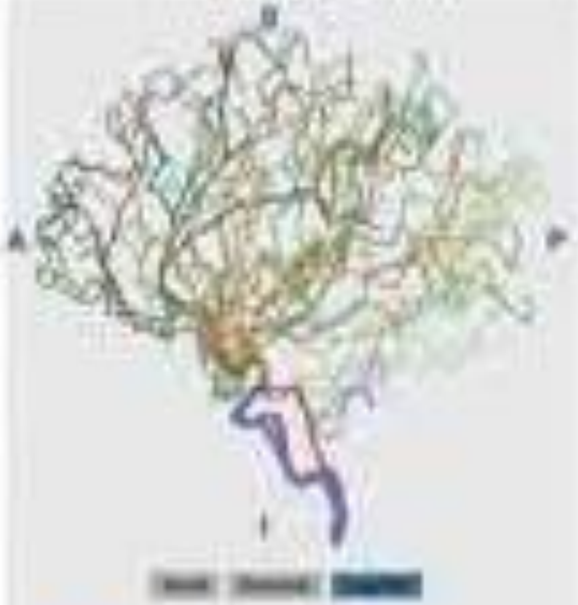
Project Example — CerebroVis





Pandey et al. VIS 2019

2D Projection

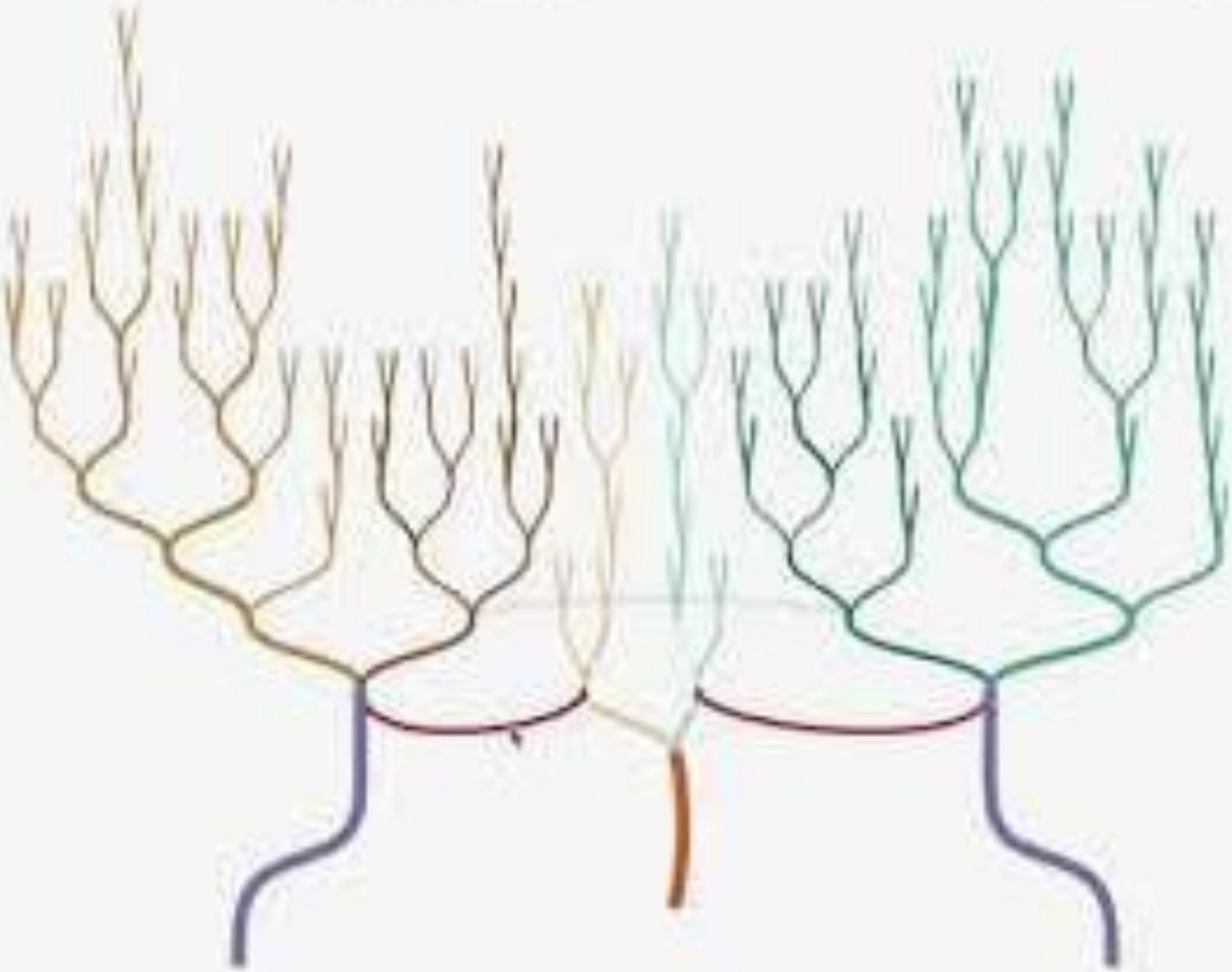


Default

Radiology

Blood Flow

CerebroVis



PROJECT EXAMPLE — WWOVIS

Close and Distant Reading via Named Entity Network Visualization: A Case Study of Women Writers Online

Sarah Campbell [*] Zheng-yan Yu [†] Sarah Connell [‡] Cody Dunne [§]

Northeastern University

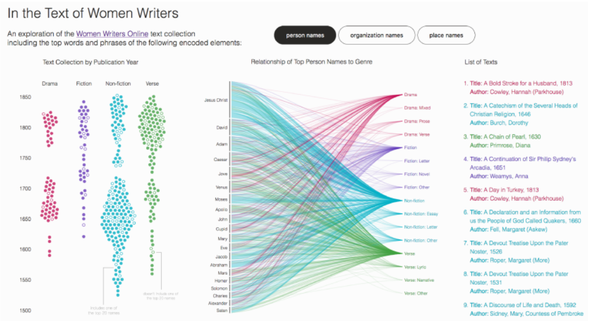


Figure 1: Three linked visualizations showing a named entity network queried from the Women Writers Online textbase. Left: a beeswarm visualization of the genre of each publication by year. Middle: a bipartite node-link visualization of the top 20 named entities connected to the genres of texts they reside in. Right: a list of texts that include at least one top 20 named entity, ordered alphabetically and linked to the full text. Marks are colored categorically by genre: drama is pink ●, fiction is purple ●, non-fiction is blue ●, and verse is green ●. Empty circles show texts that do not include any of the top 20 named entities, e.g. ○.

ABSTRACT

Close reading and distant reading are widely used in digital humanities and can benefit from information visualizations. Digital humanities scholars have curated numerous TEI-encoded textual collections which provide the data necessary for blending both close and distant reading – however we do not have tools to support general users in conducting these blended analyses. In this paper we focus on one such collection: Women Writers Online (WWO). We contribute the design and implementation of a multiple coordinated view network visualization to facilitate close and distant reading in WWO and a transparent view into our iterative design process to help guide future designers and humanists in applying our approach to other textual collections.

Index Terms: Applied computing—Education—Digital libraries and archives; Human-centered computing—Visualization—Visualization application domains—Information visualization; Human-centered computing—Visualization—Visualization techniques—Graph Drawings;

[*]e-mail: campbell.sar@husky.neu.edu
[†]e-mail: yu.zheng@husky.neu.edu
[‡]e-mail: sa.connell@northeastern.edu
[§]e-mail: c.dunne@northeastern.edu

1 INTRODUCTION

Close and distant reading are two important tools in the digital humanities toolbox which can both benefit from information visualization [6]. **Close reading** is the traditional method for literary criticism. Several visualizations have been developed to support close reading, but existing approaches can benefit from adding supplementary **named entity** information, especially acting persons and their relationships [6]. **Distant reading**, introduced by Moretti [11], alternatively focuses on an abstract view of global features of one or more texts. Network/graph visualizations can be particularly useful for examining relationships between these features and texts for corpus analysis [6]. We designed an interactive visualization to support a blend of close and distant reading – both explorations at scale and text-level investigation.

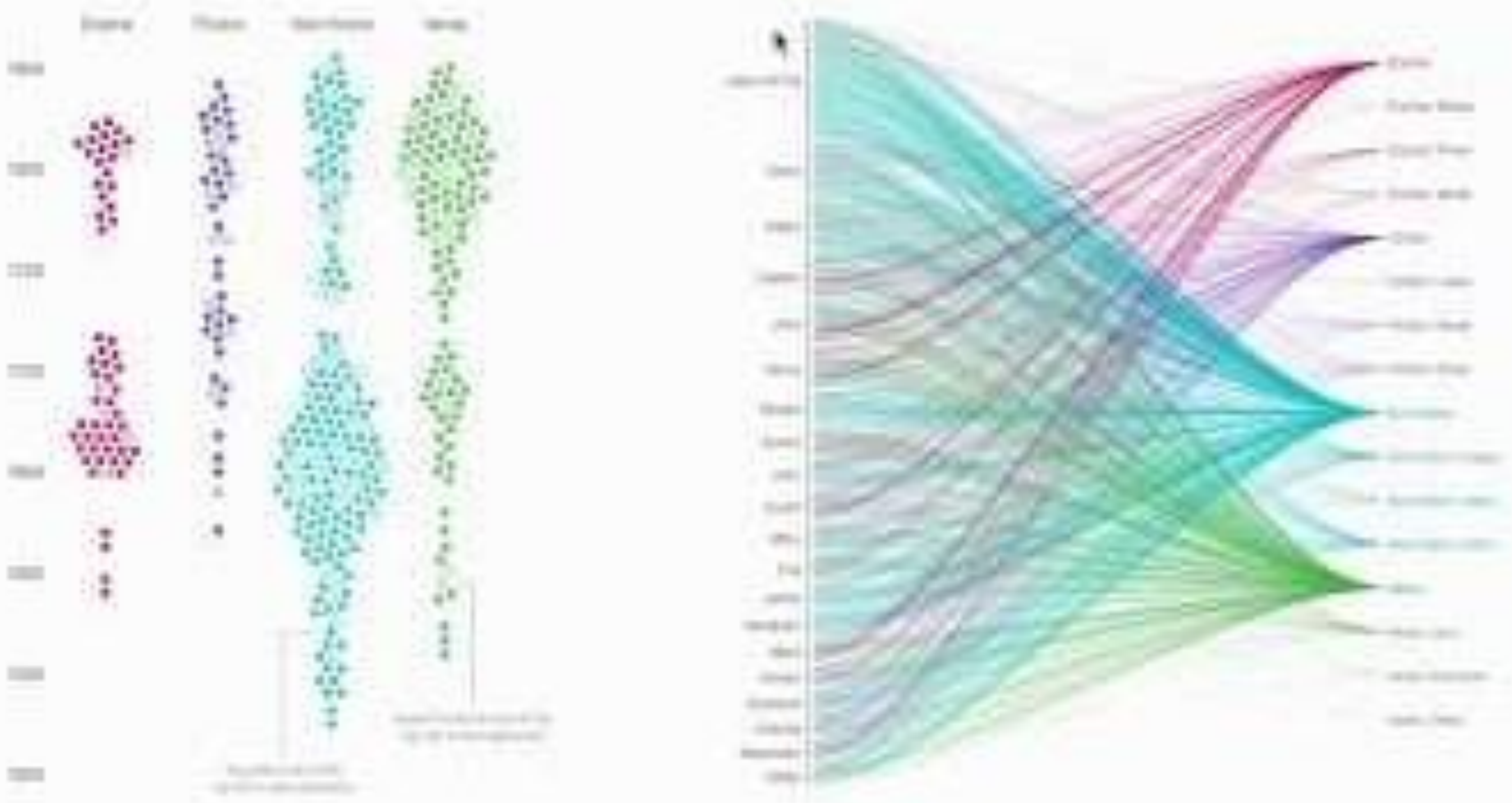
This paper focuses on the application of our visualization approach in service to the Women Writers Project (WWP). The WWP is a long-term digital humanities research project at Northeastern University, devoted to early modern women's writing and electronic text encoding. The goal of the project is to bring texts from pre-Victorian women writers out of the archive and make them more accessible to a wide audience of teachers, students, scholars, and the general user. We focus on the WWP's major textual collection, Women Writers Online (WWO). WWO is a full-text collection of early women's writing in English. It currently includes full transcriptions, encoded following the standards of the Text Encoding Initiative (TEI), of 407 texts published between 1526 and 1850. In addition to the collection's broad chronological framing, the texts in WWO also represent a very diverse set of genres, ranging from prophecies, religious meditations, petitions, and recipe books to

DS 4200 SPRING 2018: INFORMATION PRESENTATION & VISUALIZATION

Cambpell et al. VIS4DH 2018

including the top words and phrases of the following encoded elements:

person names | organization names | place names



List of Texts

EXAMPLE OF A SUCCESSFUL *DIFFERENTIATED* COURSE PROJECT

(Requires prior instructor approval to waive / alter requirements)

Project Example — Diversiform Timelines

Evaluating the Effect of Timeline Shape on Visualization Task Performance

Sara Di Bartolomeo, Aditeya Pandey, Aristotelis Leventidis, David Saffo, Uzma Haque Syeda, Elin Carstensdottir, Magy Seif El-Nasr, Michelle Borkin, Cody Dunne
Northeastern University
(dibartolomeo.s | pandey.ad | saffo.d | syeda.u)@husky.neu.edu, elin@ccs.neu.edu,
(magy | m.borkin | c.dunne)@northeastern.edu

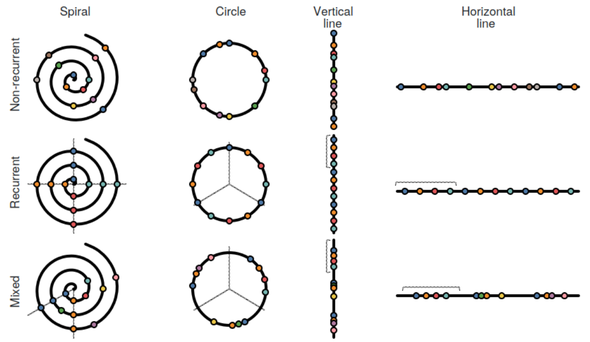


Figure 1. We evaluate the effect on task performance of four timeline shapes (left to right) across three types of temporal event sequence data (top to bottom). The images are simplified versions of the stimuli that we used in our experiment. Each dot on a timeline represents an event and has a specific categorical color to highlight where the dataset has recurrent events. Dashed lines highlight the recurrent intervals or a set of recurrent events.

ABSTRACT
Timelines are commonly represented on a horizontal line, which is not necessarily the most effective way to visualize temporal event sequences. However, few experiments have evaluated how timeline shape influences task performance. We present the design and results of a controlled experiment run on Amazon Mechanical Turk ($n = 192$) in which we evaluate how timeline shape affects task completion time, correctness, and user preference. We tested 12 combinations of four shapes — horizontal line, vertical line, circle, and spiral — and three data types — recurrent, non-recurrent, and mixed event sequences. We found good evidence that timeline shape meaningfully affects user task completion time but not correctness and that users have a strong shape preference. Building on our results, we present design guidelines for creating effective timeline visualizations based on user task and data types. A free copy of this paper, the evaluation stimuli and data, and code are available at https://osf.io/qr5yu/

Author Keywords
Timelines; Temporal Event Sequences; Information Visualization; Controlled Experiments

CCS Concepts
•**Human-centered computing** → **Human computer interaction (HCI);** *Visualization design and evaluation;* Information visualization;

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

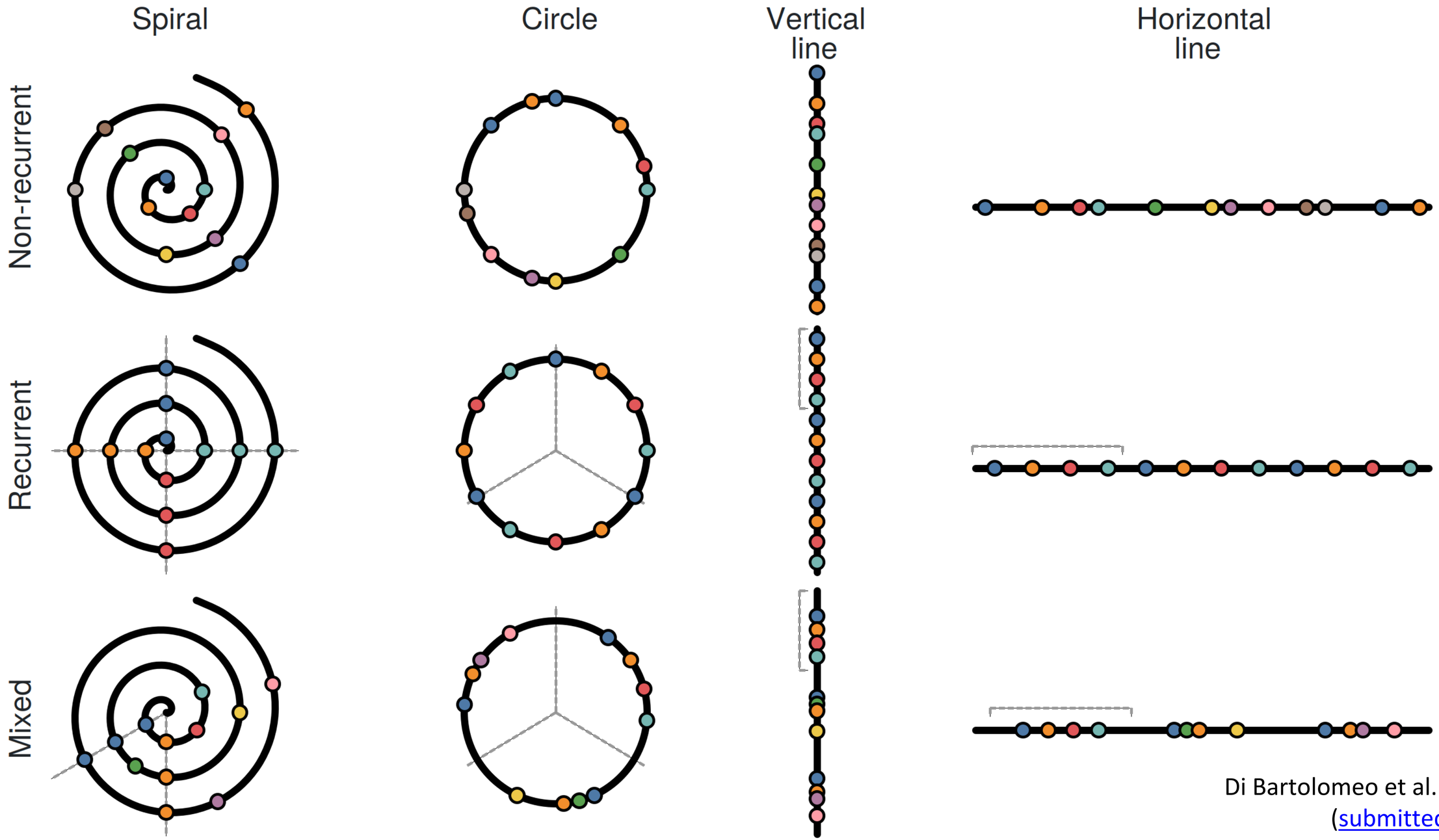
ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376237

INTRODUCTION
A timeline is a visual representation of a series of events in time. The use of timelines dates back to 17th century [32] when Joseph Priestly designed a visualization that showed the rise and fall of empires in Europe's history. In the modern era, timelines have become prevalent in our daily lives as the de facto representation to show financial trends, weather

CS 7340 Fall 2018: Theory and Methods in Human Computer Interaction

Di Bartolomeo et al. CHI 2020 (submitted version)

Project Example — Diversiform Timelines



Di Bartolomeo et al. CHI 2020
(submitted version)

Project Ideas:
VIS + X

Where X = (ML | SEC | NLP | HCC | GAM | NS | SYS | …)

Potential Venue: IEEE VIS 2020 Short Papers

Deadline ~June 13, 2020

Short Papers Call for Participation

IEEE Vis 2019 solicits submissions in a short paper format. Short papers draw from the same paper types and topics as full papers of VAST, Infovis and Scivis, ranging from theoretical to applied research contributions.

Short papers welcome submissions describing original work with focused and concise research contributions, incremental work such as follow-up extensions or evaluations of existing methods, or exploratory work. We provide example papers below.

The short paper submission deadline is June 13, 2019, creating an opportunity to showcase late-breaking research results. This deadline is one week after the June 6 notification for conditional acceptance of full papers, enabling the distillation and focused expression of work that otherwise might not be part of VIS.

Submission Guidelines:

Short papers can be up to four pages plus up to one page of references. Submissions will be in the VGTC conference two-column format, in line with the IEEE VIS formatting guidelines. Short papers may be single-blind or double-blind - reviewers' identities are not revealed. Authors may choose whether to anonymize their submission or not.

Short papers must be submitted through the Precision Conference System.

Example Papers:

- Max Zeyen, Tobias Post, Hans Hagen, James Ahrens, David Rogers, and Roxana Bujack. Color interpolation for non-Euclidean color spaces. SciVis 2018 Short Papers.
- Anke Friederici, Habib Toye, Ibrahim Hoteit, Tino Weinkauf, Holger Theisel, and Markus Hadwiger. A Lagrangian Method for Extracting Eddy Boundaries in the Red Sea and the Gulf of Aden. SciVis 2018 Short Papers.
- Oliver Reiter, Marcel Breeuwer, Eduard Gröller, and Renata Georgia Raidou. Comparative visual analysis of pelvic organ segmentations. EuroVis 2018 Short Papers.

SUBMIT YOUR WORK

Papers
Short Papers
Posters
Tutorials
Workshops
Panels
Doctorial Colloquium
Application Spotlights

IMPORTANT INFO

Open Practices
Video Previews

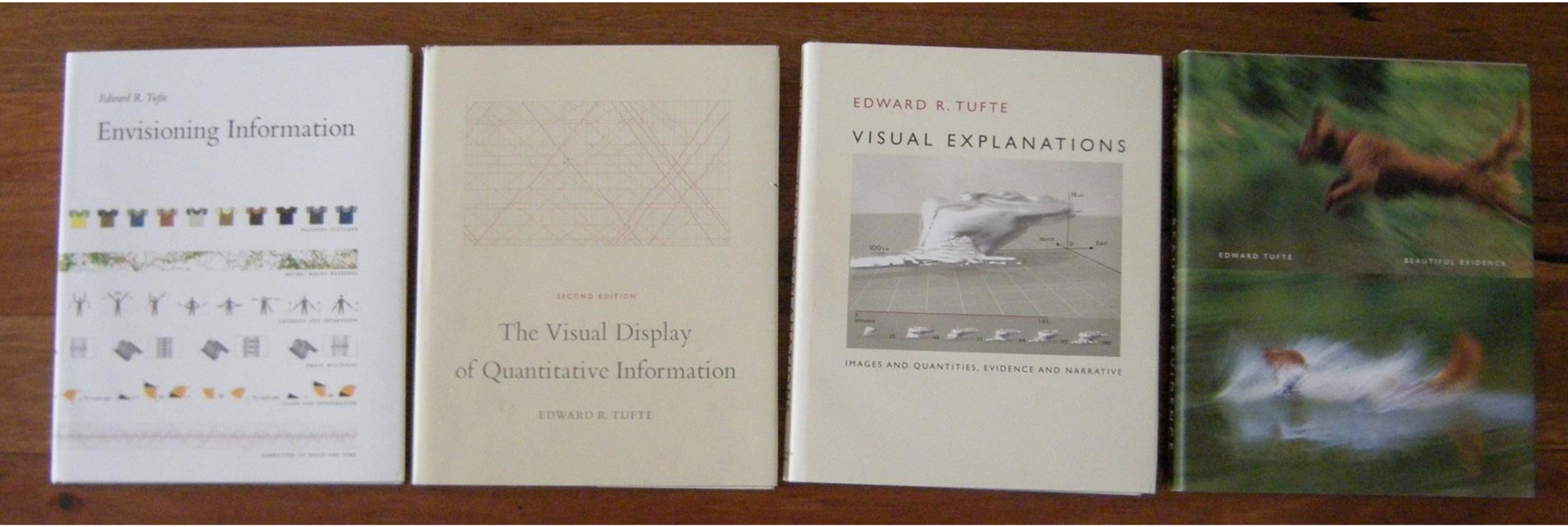
PROJECTS

In-class project pitches: M 2020.01.27

What questions do you have for me?

DESIGN & RULES OF THUMB

Edward Tufte



Tufte will be doing one of his one-day courses in Boston on 10/29, 30, and 31 2018.
$220 for students includes these books.
https://www.edwardtufte.com/tufte/courses

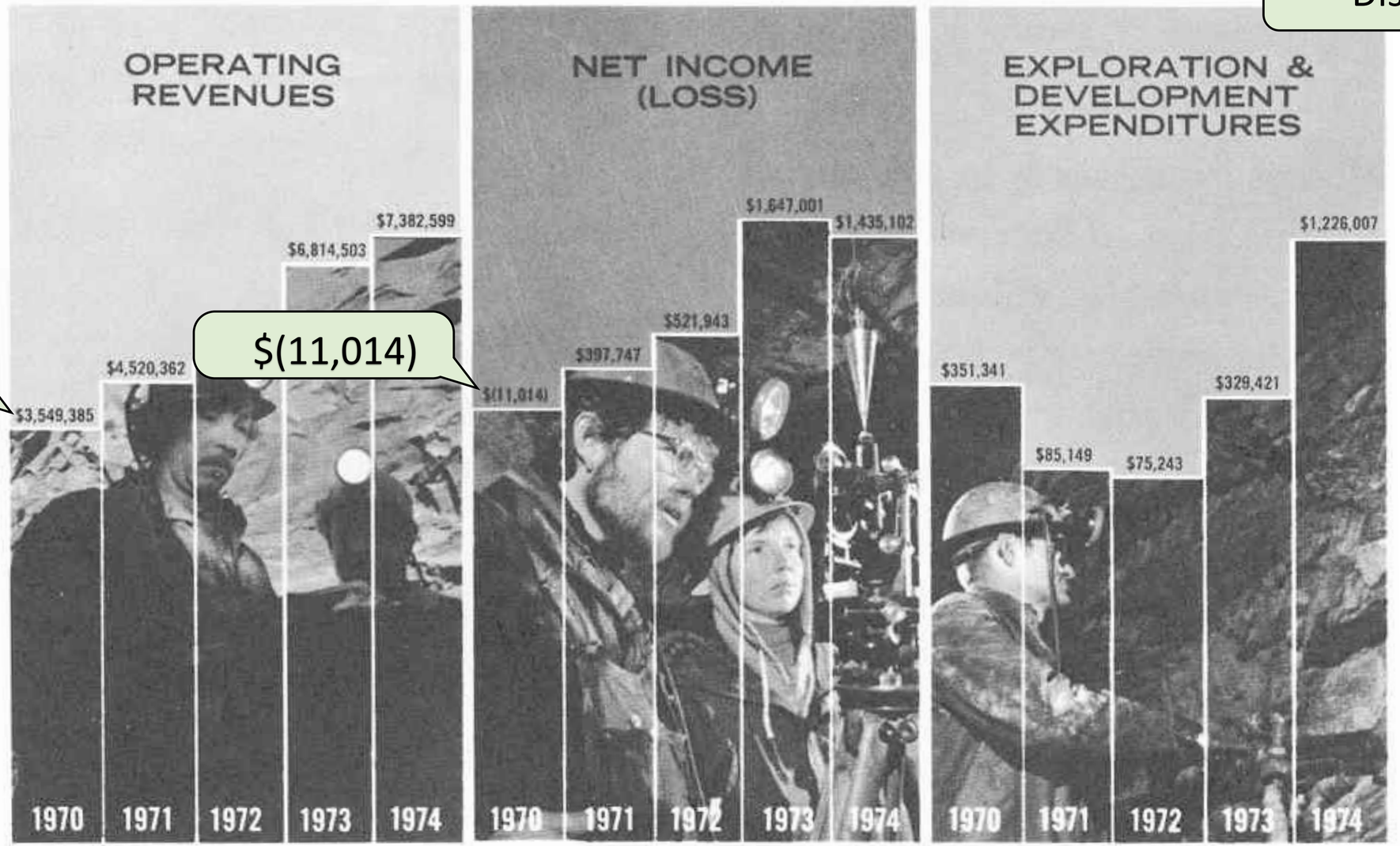
“Graphical Integrity”

“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.”

(Axes and axis labels, titles, annotations, legends, etc.)

Tufte, “Visual Display of Quantitative Information” (1983)

"Distorted Scales"



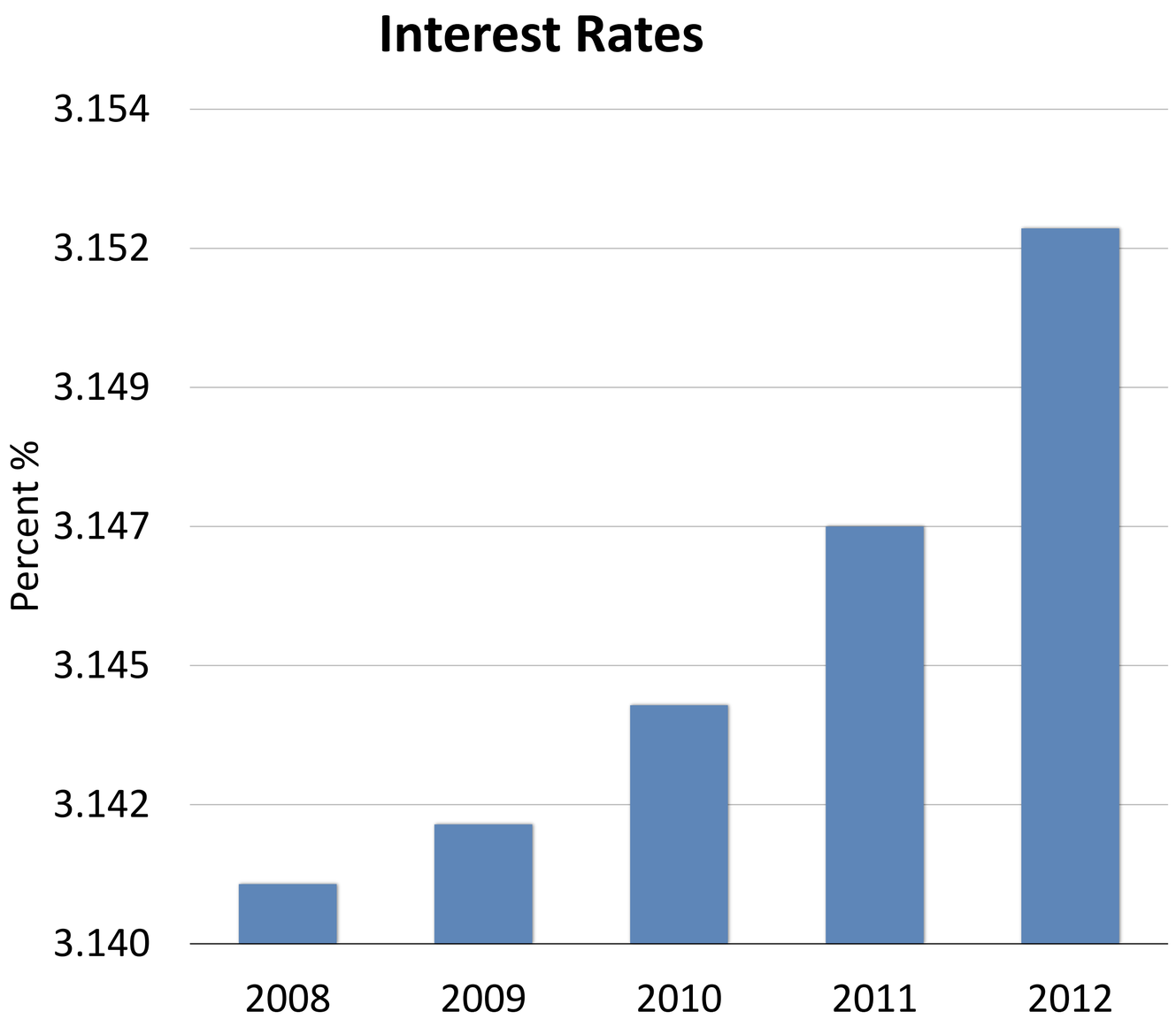
$3,549,385

$(11,014)

y-axis baseline?!

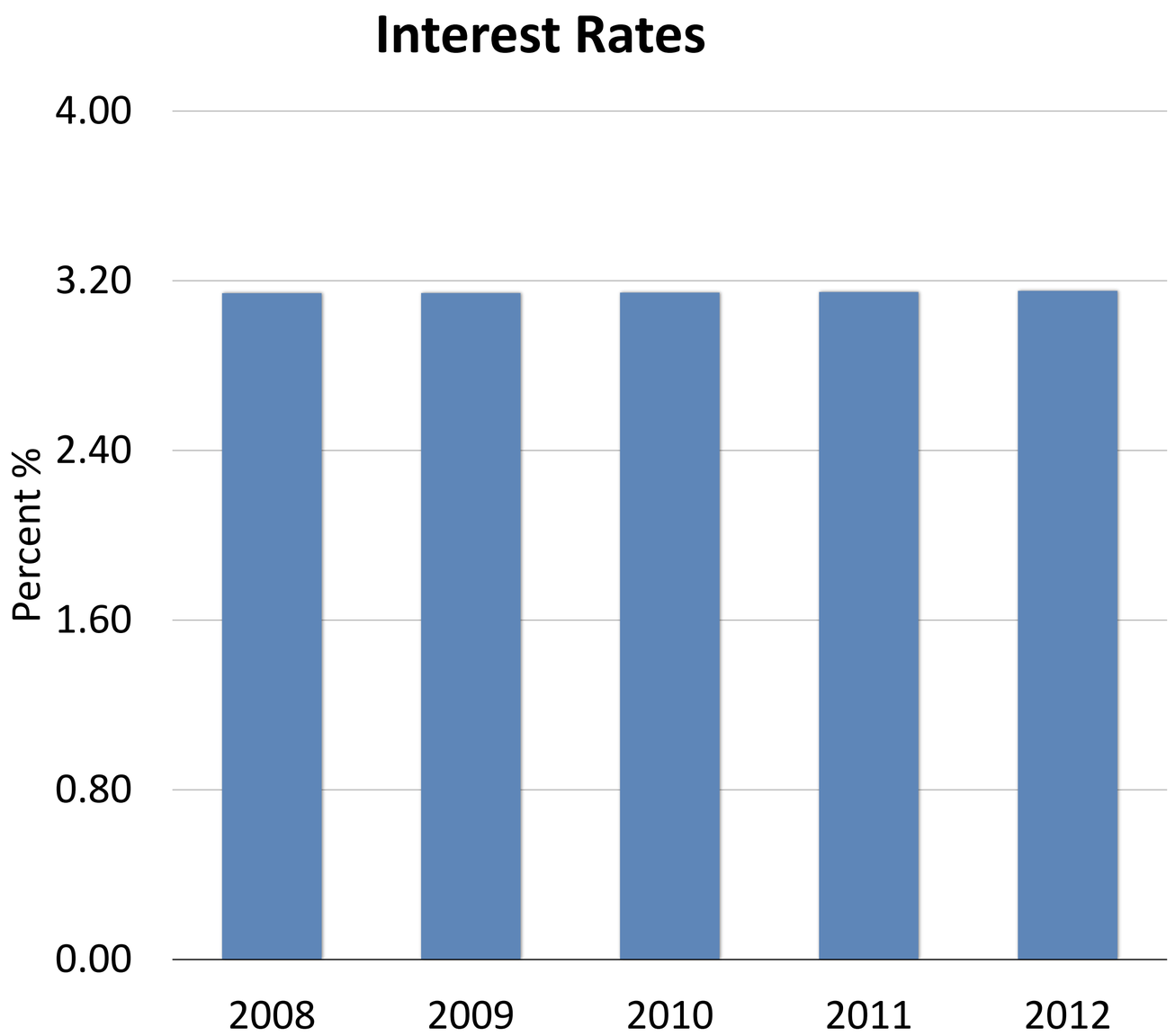
"Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data."

Tufte, "Visual Display of Quantitative Information" (1983)



"Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data."

Based on http://data.heapanalytics.com/how-to-lie-with-data-visualization



CONTEXT!

"Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data."

Based on http://data.heapanalytics.com/how-to-lie-with-data-visualization

Upcoming Assignments & Communication

https://canvas.instructure.com/courses/1781732

If you don't have an account on our Canvas yet:
https://canvas.instructure.com/enroll/CMAPDM
Use your name as known by the registrar and your @husky.neu.edu email.

Look at the upcoming assignments and deadlines (12:01am)!

Everyday Required Supplies:

- 5+ colors of pen/pencil
- White paper
- Laptop and charger
- Table tent

Use Canvas Discussions for general questions, email the instructor/TAs for questions specific to you.