Storytelling with Data

Module 10: Infographics — Demonstration and discussion

Scott Spencer
Faculty and Lecturer
Columbia University

Agenda

Upcoming deliverable

Today’s objectives

More on uncertainty

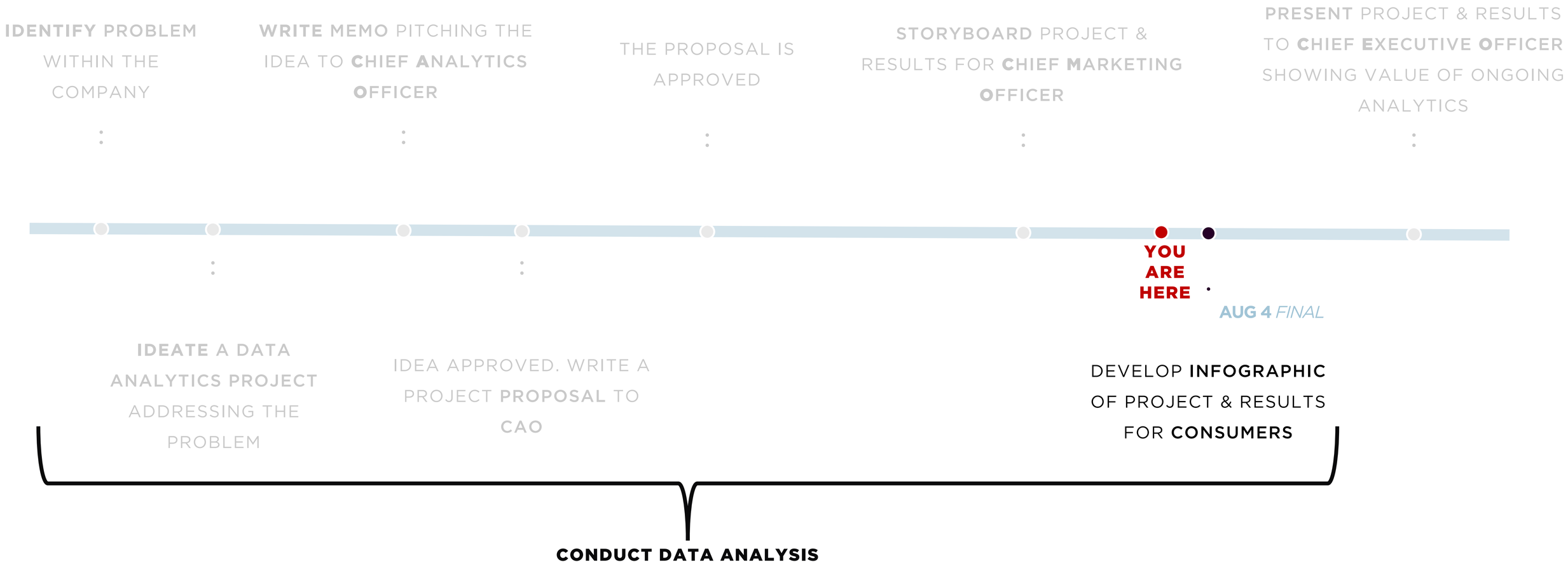
Information graphics

Questions or suggestions?

Upcoming deliverables

Upcoming deliverables

Information graphic — reframe your story, this time building off the messages you built for the marketing team in order to craft an infographic that displays the results of the analytic work in a way that is accessible, engaging, and exciting for a **general or consumer audience**.
And: present your critique of an exemplary graphic.

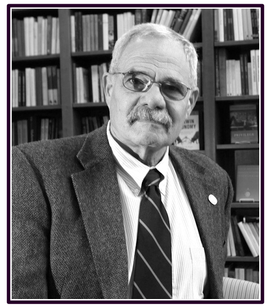


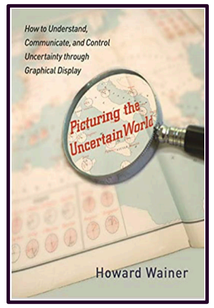
Today's Objectives

Objectives

1. More on communicating uncertainty with data visualization
2. Discussion on infographics
3. Obtaining feedback to improve your infographic

More on uncertainty





Picturing the Uncertain World

Wainer

He is is an American statistician, past principal research scientist at the Educational Testing Service, adjunct professor of statistics at the Wharton School of the University of Pennsylvania, and author, known for his contributions in the fields of statistics, psychometrics, and statistical graphics.

The most dangerous equation

De Moivre's equation:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad \therefore \quad \sigma_{\bar{x}} < \sigma$$

σ the measure of the variability of a population (its standard deviation).

$\sigma_{\bar{x}}$ the variation of averages of subsets of the population.

n the number of observations in each subset

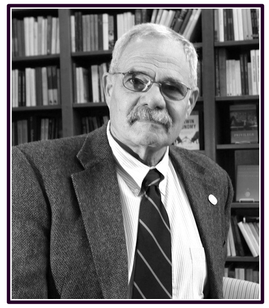
Why so dangerous?

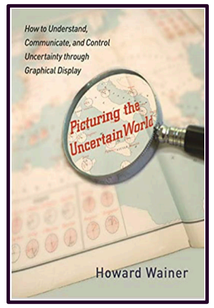
Extreme length of time during which ignorance of it has caused confusion

Wide breadth of areas that have been misled

Seriousness of the consequences that ignorance has caused

Wainer, Howard. *The Most Dangerous Equation* in Chapter 1, Picturing the Uncertain World. Princeton University Press, 2009. 1–18. Print.





Picturing the Uncertain World

Wainer

He is is an American statistician, past principal research scientist at the Educational Testing Service, adjunct professor of statistics at the Wharton School of the University of Pennsylvania, and author, known for his contributions in the fields of statistics, psychometrics, and statistical graphics.

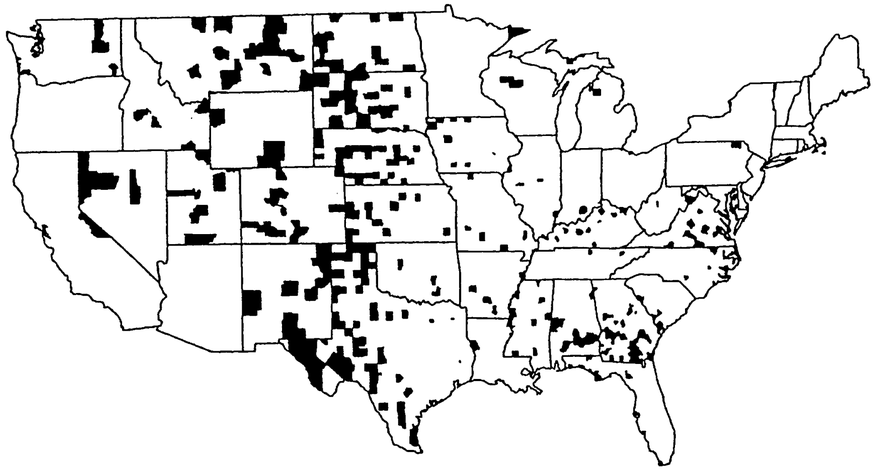


Figure 1.1.
Lowest kidney cancer death rates. The counties of the United States with the lowest 10% age-standardized death rates for cancer of kidney/urethra for U.S. males, 1980–1989 (from Gelman and Nolan, 2002, p. 15, reprinted with permission).

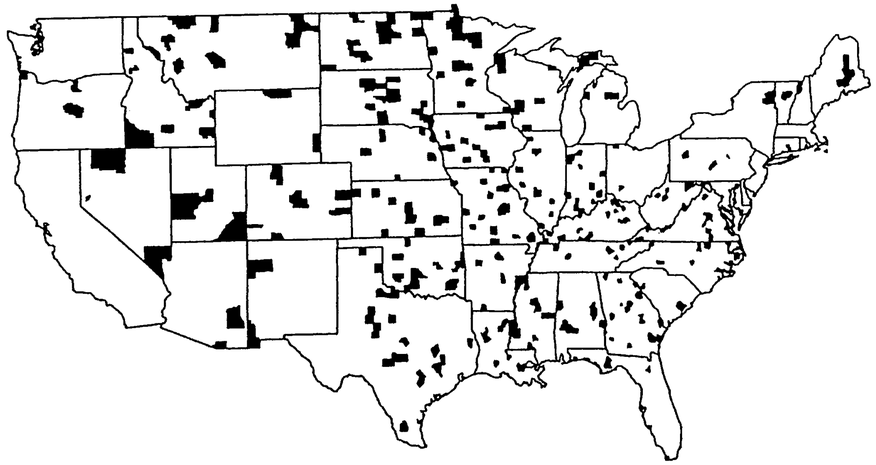
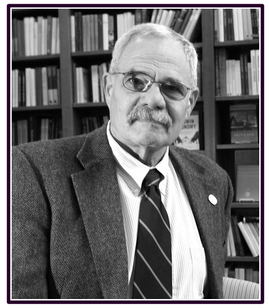
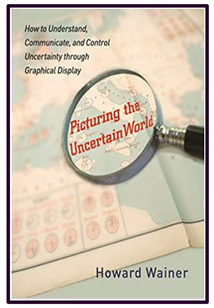


Figure 1.2.
Highest kidney cancer death rates. The counties of the United States with the highest 10% age-standardized death rates for cancer of kidney/urethra for U.S. males, 1980–1989 (from Gelman and Nolan, 2002, p. 14, reprinted with permission).

Wainer, Howard. *The Most Dangerous Equation* in Chapter 1, Picturing the Uncertain World. Princeton University Press, 2009. 1–18. Print.

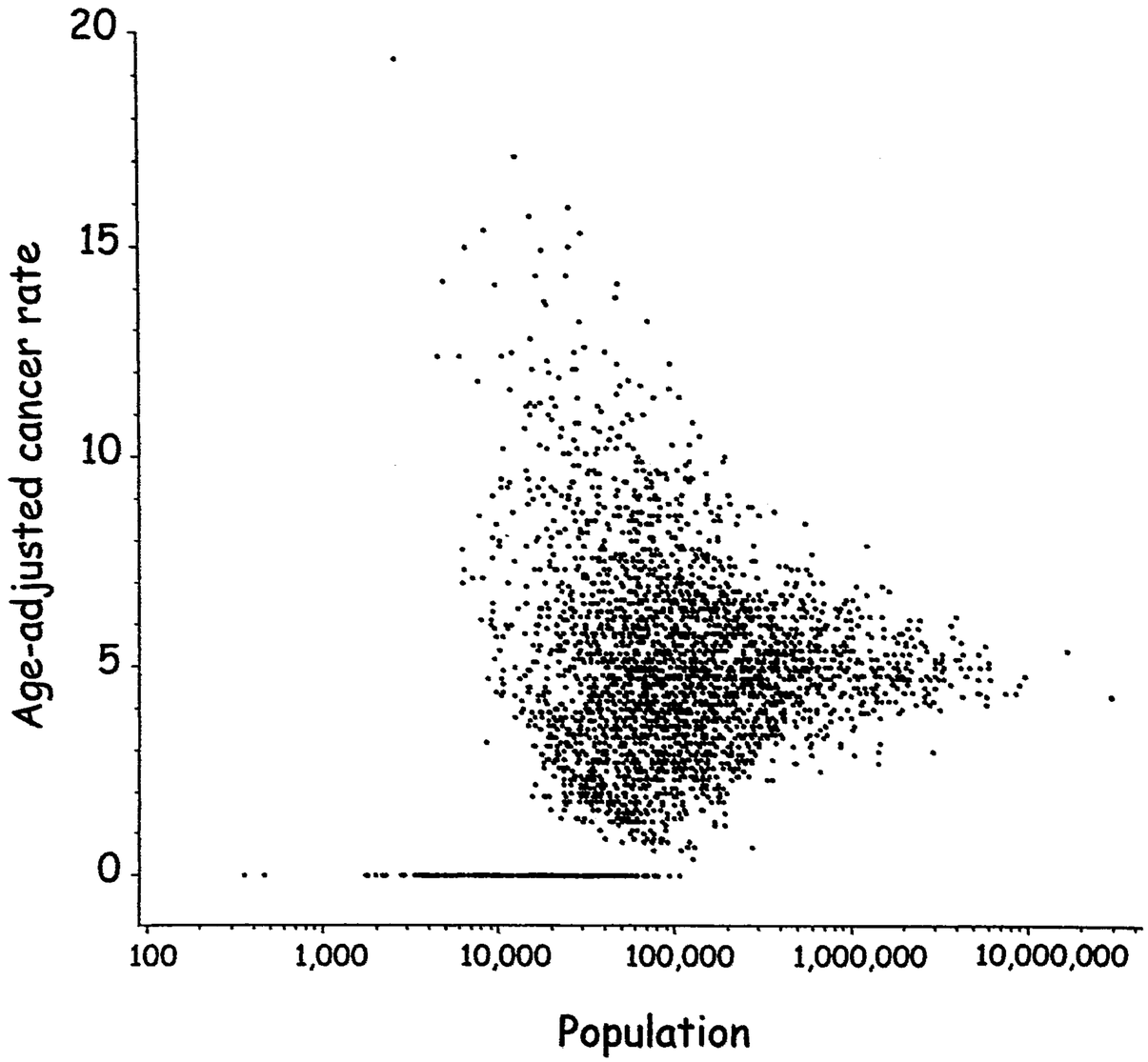




Picturing the Uncertain World

Wainer

He is is an American statistician, past principal research scientist at the Educational Testing Service, adjunct professor of statistics at the Wharton School of the University of Pennsylvania, and author, known for his contributions in the fields of statistics, psychometrics, and statistical graphics.



Wainer, Howard. *The Most Dangerous Equation* in Chapter 1, Picturing the Uncertain World. Princeton University Press, 2009. 1–18. Print.

BARUCH FISCHHOFF

Communicating Uncertainty
Fulfilling the Duty to Inform

Scientists are often hesitant to share their uncertainty with decisionmakers who need to know it. With an understanding of the reasons for their reluctance, decisionmakers can create the conditions needed to facilitate better communication.

Communicating Uncertainty: Fulfilling the Duty to Inform

Fischhoff

He is a Professor in Department Engineering and Public Policy and Institute for Politics and Strategy at Carnegie Mellon University. A graduate of the Detroit Public Schools, he holds a BS in mathematics and psychology from Wayne State University and an MA and PhD in psychology from the Hebrew University of Jerusalem.

Good decisions rely on knowledge of uncertainty

Scientists are often **hesitant to share their uncertainty** with decisionmakers who need to know it. With an **understanding of the reasons for their reluctance**, decisionmakers can create the conditions needed to facilitate better communication.

Failure to express uncertainty has negative value

Communicating knowledge can worsen results if it induces unwarranted confidence or is so hesitant that other, overstated claims push it aside.

Quantifying uncertainties aid verbal expression

Concern: people will misinterpret quantities of uncertainty, inferring more precision than intended. **Response**: Most people like getting quantitative information on uncertainty, from them can get the main message, and **without them are more likely to misinterpret verbal expressions of uncertainty**.

Posing clear questions guide understanding

Concern: people cannot use probabilities. **Response**: **laypeople can provide high-quality probability judgments, if they are asked clear questions and given the chance to reflect on them**.

Communicating uncertainty protects credibility

Concern: credible intervals may be used unfairly in performance evaluations. **Response**: **probability judgments give us more accuracy about the information**; i.e., won't be too confident or lack enough confidence.

Fischhoff, Baruch. "Communicating Uncertainty: Fulfilling the Duty to Inform." Issues in Science and Technology 28.4 (2012): 63–70. Print.

When (ish) is My Bus? User-centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems

Matthew Kay, University of Washington, mjskay@uw.edu
Tara Kola, Computer Science, tarakola@cs.washington.edu
Jessica R. Hullman, Information School, jhullman@uw.edu
Sean A. Munson, HCDE, samunson@uw.edu
Julie A. Kientz, Computer Science & Engineering, jakientz@cs.washington.edu

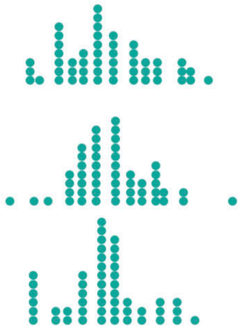
When (ish) Is My Bus?

Kay

He is an Assistant Professor of Information at UMSI, and works in human–computer interaction and information visualization.

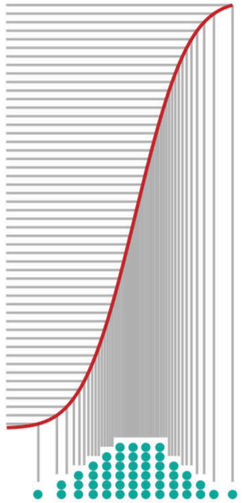
Probability density of Normal distribution

To generate a discrete plot of this distribution, we could try taking **random draws** from it. However, **this approach is noisy**: it may be very different from one instance to the next.



Instead, we use the **quantile function (inverse CDF)** of the distribution to generate "draws" from evenly-spaced quantiles.

We plot the quantile "draws" using a Wilkinsonian dotplot, yielding what we call a **quantile dotplot**: a consistent discrete representation of a probability distribution.



By using quantiles we facilitate interval estimation from frequencies: e.g., knowing there are 50 dots here, if we are willing to miss our bus **3/50** times, we can count **3 dots** from the left to get a one-sided **94% (1 – 3/50) prediction interval** corresponding to that risk tolerance.

Kay, Matthew et al. "When (Ish) Is My Bus? User-Centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems." New York, New York, USA: ACM Press, 2016. 5092–5103. Web.
Kay, Matthew, https://github.com/mjskay/when-ish-is-my-bus/blob/master/quantile-dotplots.md

Group work

We want information graphics to …

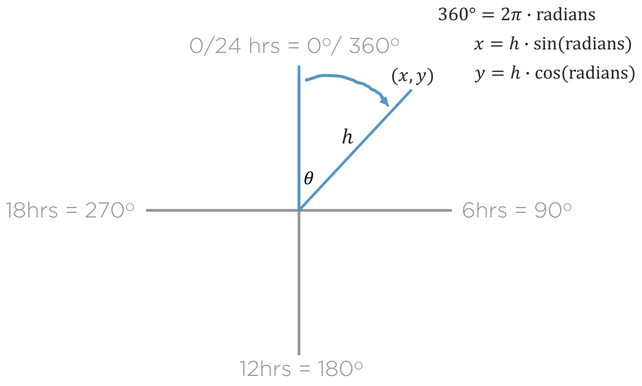
Tell a complete story where the purpose is to inform, entertain or persuade the audience. It should:

simple, focused messages
new, surprising information
credible data sources

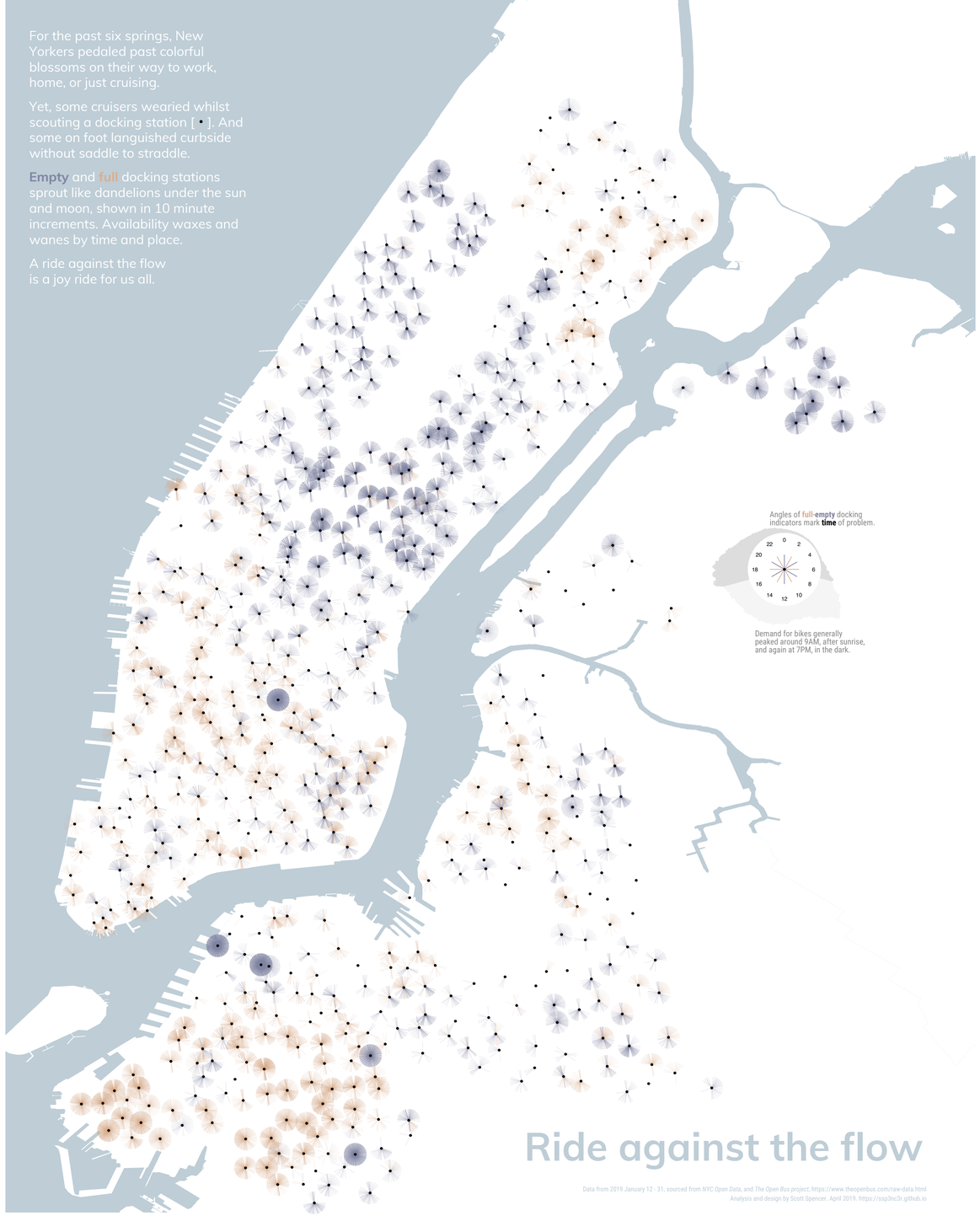
visually coherent, integrated
use comparisons for context, meaning
principles of information design, organized

CitiBike example, continued

Maths to create data encodings at each dock station



Opacity is the lack of transparency. Opacity was lowered so that a single unavailable bike or full station would not be very noticeable but several markings at that time (on different days) would increase opacity, making the marking brighter and thus increase its visual importance.



Basic math can help when making custom graphics. If you'd like a refresher on basic algebra, geometry, trigonometry: Simmons, George F. *Precalculus Mathematics in a Nutshell*. Barnes & Noble Books, 1987. Print.

Let’s look ahead

For Next Week, Module 11:

Agenda next week

Next deliverables, *final* information graphic **AND** critique of an infographic you select to share.

Combining written and visual communication of analytics with verbal delivery.

The minimum

Doumont, Jean-Luc. *Effective Oral Presentations*, in *Trees, Maps, and Theorems*. Principiæ, 2009. Print.

Consider his process of preparing an oral presentation and how oral presentation shares complements written and visual communication.

Tufte, Edward R. *The Cognitive Style of PowerPoint: Pitching Out Corrupts Within*, in *Beautiful Evidence*. Graphics Press, 2006. Print.

Think about the issues he raises of information presented in the form of projected slides.

Duarte, Nancy. *There's always room to improve*, in Chp. 8, *Resonate: Present Visual Stories That Transform Audiences*. Wiley, 2010. Print.

Consider how she organizes and designs the visual component of a presentation, and how it relates to oral delivery.

Pay it forward

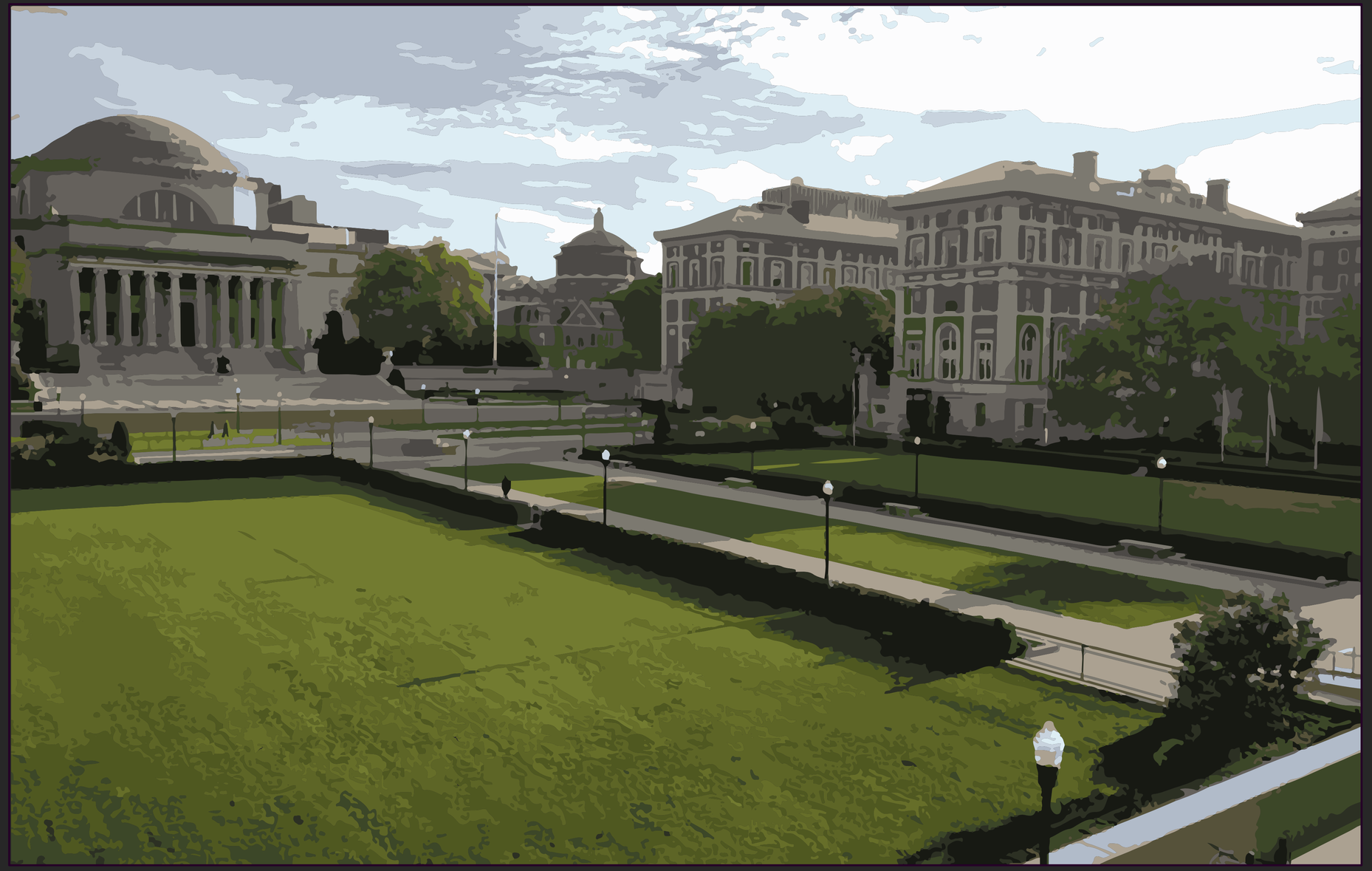
Next steps?

Did you enjoy choosing your own project? How might you plan to deepen the analysis in a way you'd like to showcase in your portfolio?

Letter to your future colleagues?

Having had the experience of taking this course, what advice do you have for your future colleagues about to begin this course? In other words, what would you want your younger self to know?

See you next week!



Spencer, Scott. *View From Butler*. 27 July 2019. Illustration