

The Art of Data Storytelling

The Essential Data Science Skill
Everyone Needs

Taught by

Isaac Reyes, Data Scientist



TEACHING



Lecturer in Statistics



**Australian
National
University**

Lecturer in Statistics



TED Presenter

CURRENT ROLE

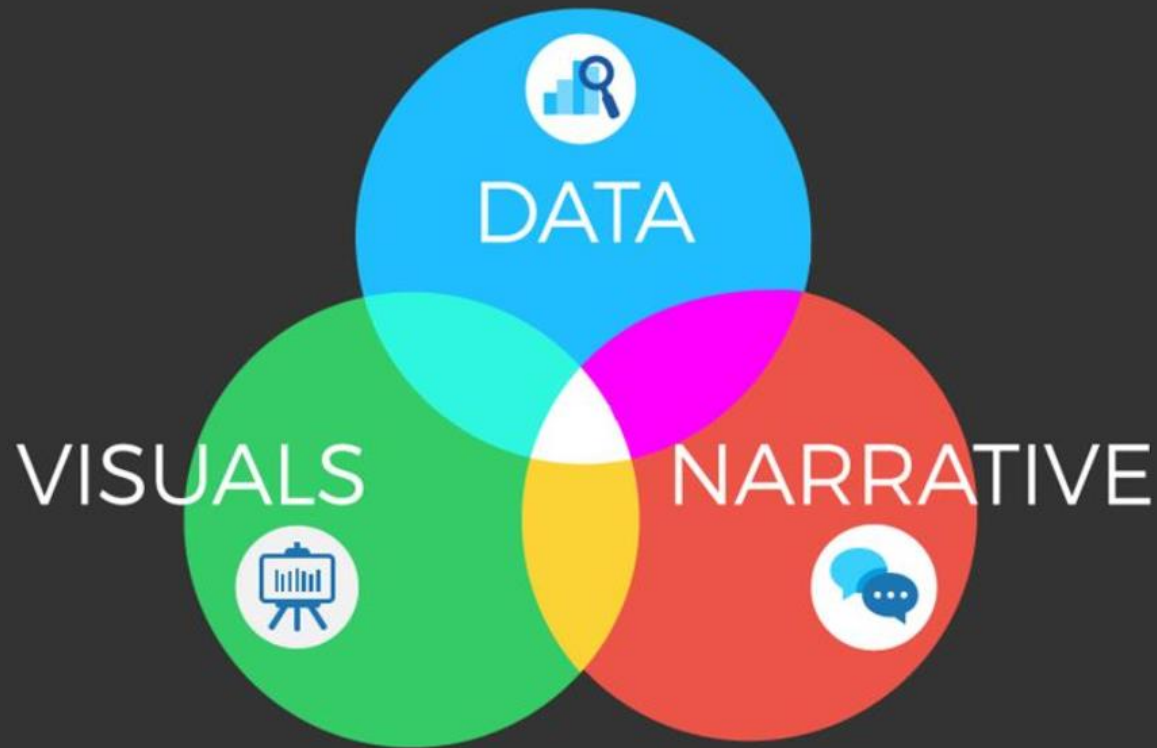


Head of Data Science



Lead Trainer

KEYS TO DATA STORYTELLING



AUDIENCE

KEYS TO DATA STORYTELLING

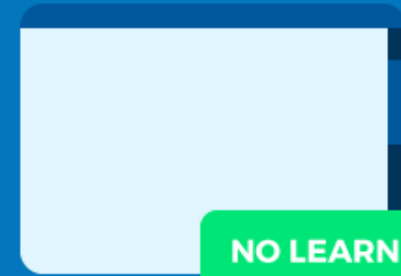
A row of stylized, dark gray silhouettes of people's heads and shoulders, representing an audience. The silhouettes are of various heights and shapes, suggesting a diverse group of individuals.

AUDIENCE

WHAT YOU PRESENT TO YOUR AUDIENCE



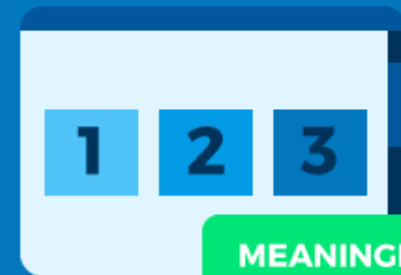
WHAT REALLY HAPPENS



NO LEARNING



FRAGMENTED LEARNING



MEANINGFUL LEARNING

Creating Audience Focused Data Stories

- Think about the most appropriate story medium for your audience.
- Select metrics and visualizations that **suit your audience**



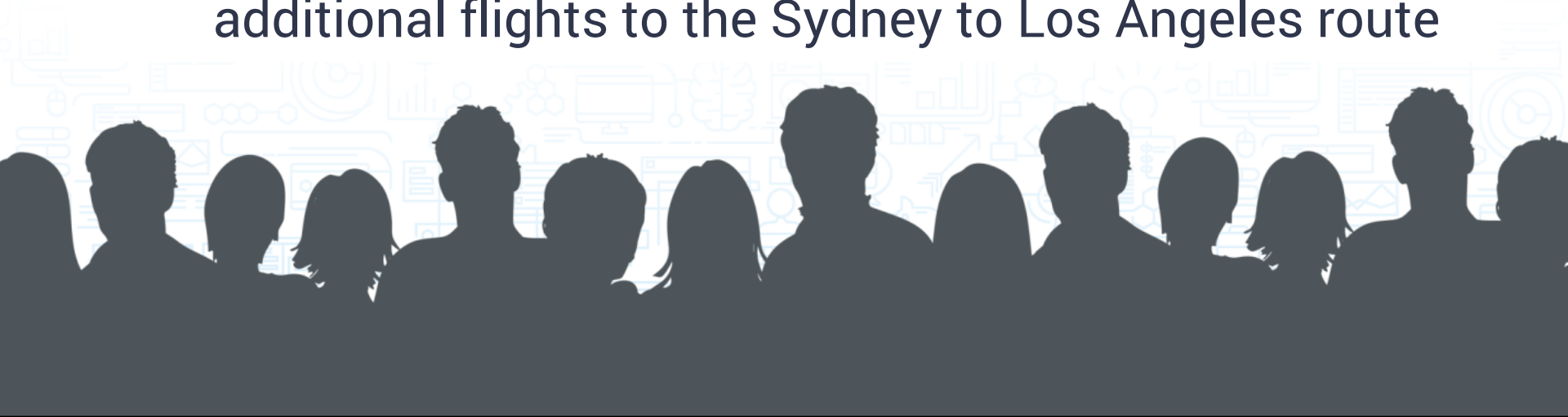
Have a clear storytelling goal

- Super important – Decide what the audience needs to know and focus your entire presentation around communicating this ‘big takeaway’ clearly.
- Be bold! You probably know the data better than anyone else in the company.



Let's make some assumptions...

- Let's assume that we are consultants who are brought in to look at the following airline data.
- The audience includes the COO of the airline (the decision maker)
- The COO wants a recommendation on whether to add additional flights to the Sydney to Los Angeles route



File

Home

Insert

Page Layout

Formulas

Data

Review

View

Tell me what you want to do

Paste

Cut

Copy

Format Painter

Clipboard

Calibri

11

Font

Alignment

General

\$

%

‚

0.00

0.00

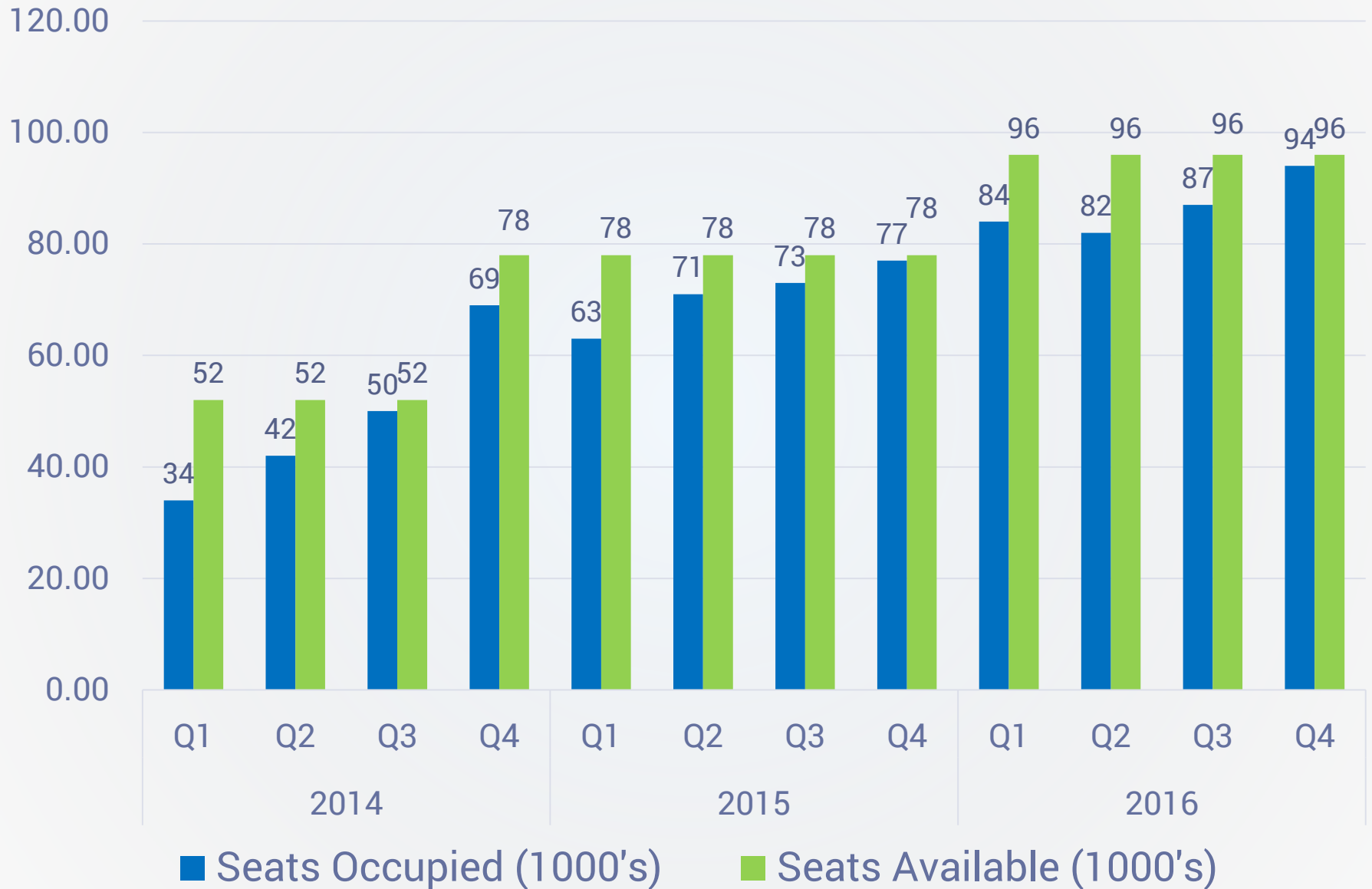
Number

Cor
Form

O20

	A	B	C	D	E	F	G	H
1	Sydney to Los Angeles Route							
2								
3			Seats Occupied (1000's)	Seats Available (1000's)				
4	2014	Q1	34	52				
5		Q2	42	52				
6		Q3	50	52				
7		Q4	69	78				
8	2015	Q1	63	78				
9		Q2	71	78				
10		Q3	73	78				
11		Q4	77	78				
12	2016	Q1	84	96				
13		Q2	82	96				
14		Q3	87	96				
15		Q4	94	96				
16								
17								

Sydney to Los Angeles Route

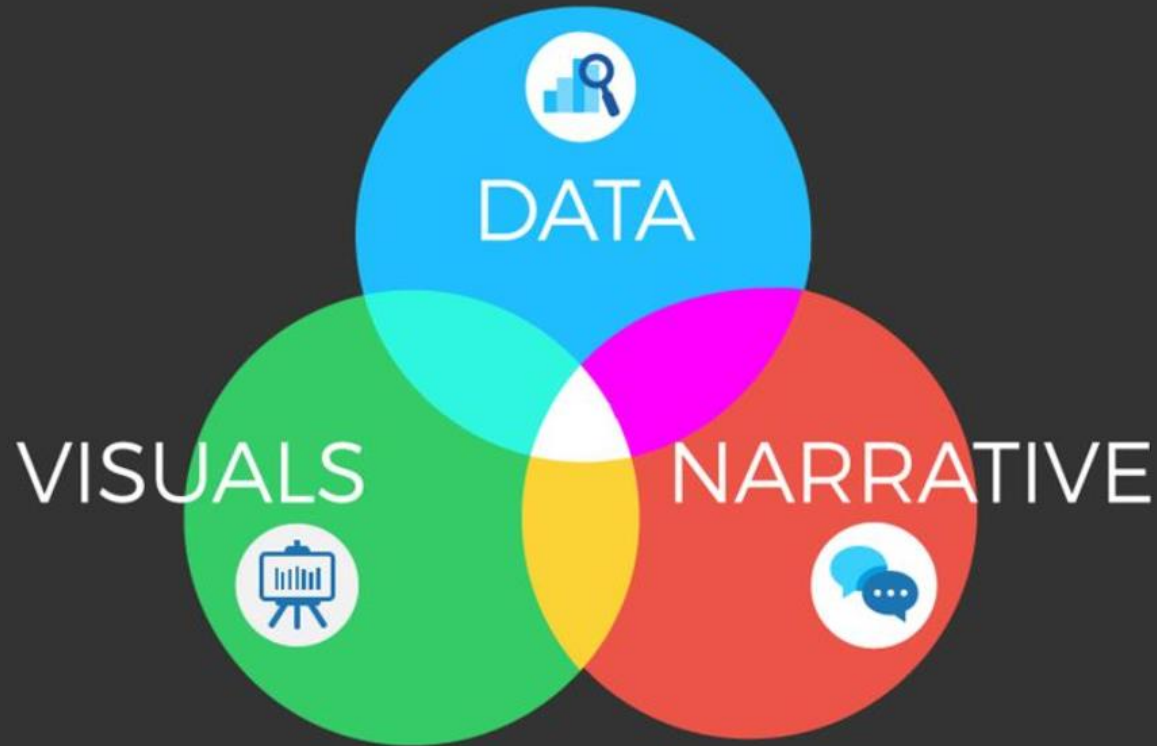


What are some of the insights?

- Seat demand is seasonal, peaking in Q4 each year
- Historically, when new seat capacity is added, seat occupancy grows to fill the extra capacity
- It is likely that total demand was not met with enough seats in Q4 2015 and Q4 2016.



KEYS TO DATA STORYTELLING



AUDIENCE



KEYS TO DATA STORYTELLING



Data

- ⬠ Pre-process and clean the data properly
- ⬠ Don't assume that the variables in front of you are necessarily the most useful variables for the problem at hand
- ⬠ Think carefully about the most appropriate metric



File

Home

Insert

Page Layout

Formulas

Data

Review

View

Tell me what you want to do

Paste

Cut

Copy

Format Painter

Clipboard

Calibri

11

Font

Alignment

General

\$

%

‚

0.00

0.00

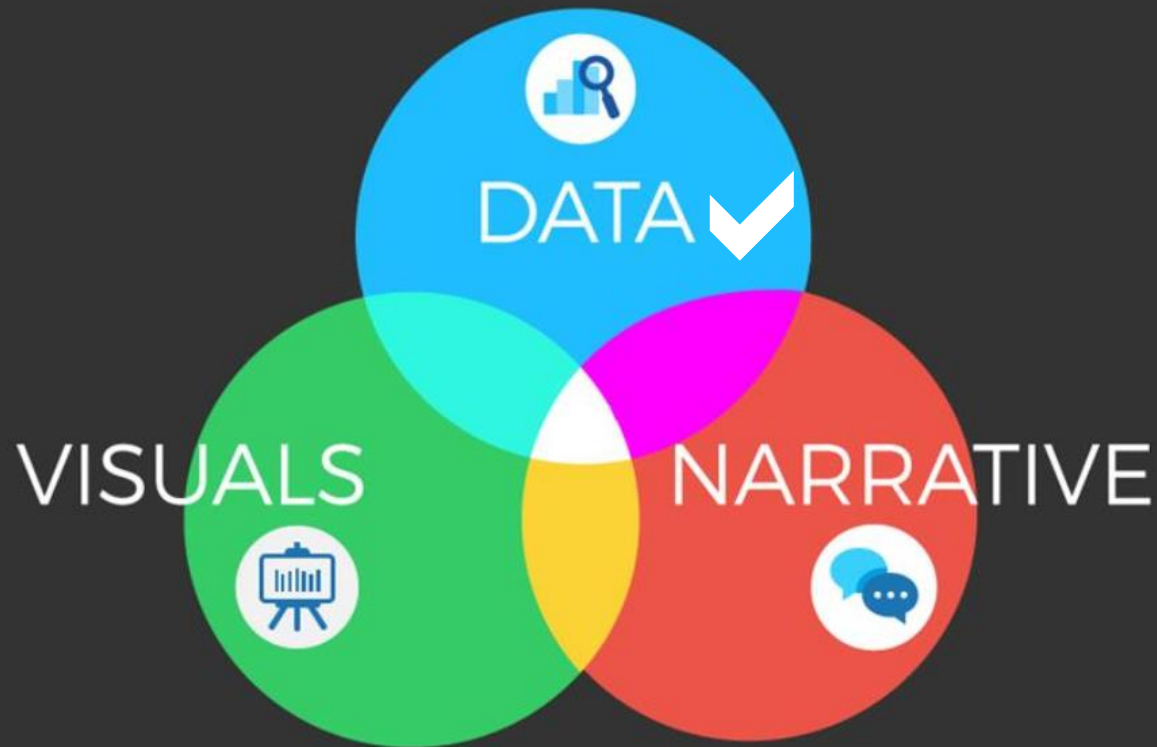
Number

Cor
Form

O20

	A	B	C	D	E	F	G	H
1	Sydney to Los Angeles Route							
2								
3			Seats Occupied (1000's)	Seats Available (1000's)				
4	2014	Q1	34	52				
5		Q2	42	52				
6		Q3	50	52				
7		Q4	69	78				
8	2015	Q1	63	78				
9		Q2	71	78				
10		Q3	73	78				
11		Q4	77	78				
12	2016	Q1	84	96				
13		Q2	82	96				
14		Q3	87	96				
15		Q4	94	96				
16								
17								

KEYS TO DATA STORYTELLING



AUDIENCE



KEYS TO DATA STORYTELLING





**“Above all else,
show the data”
- Edward Tufte**

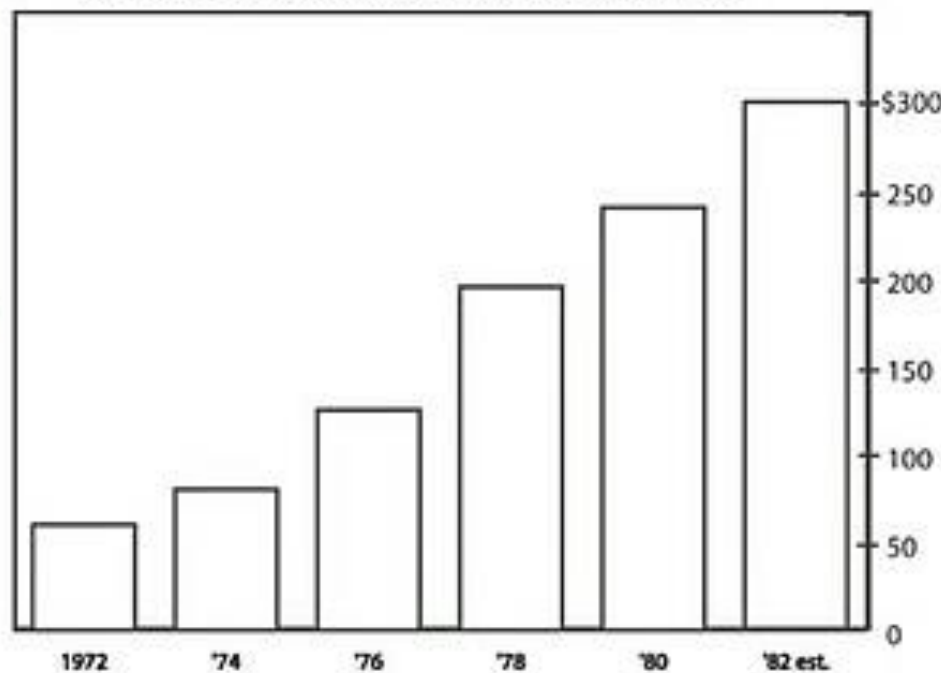
MONSTROUS COSTS

Total House and Senate
campaign expenditures,
in millions



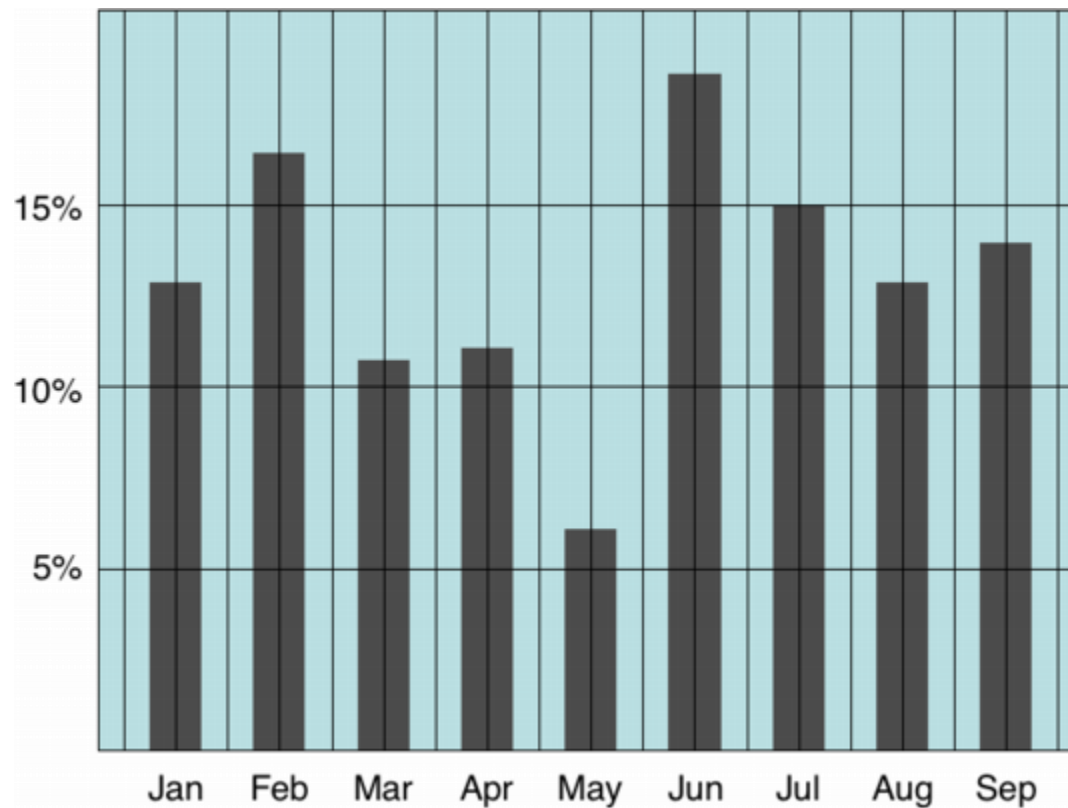
MONSTROUS COSTS

Total House and Senate campaign expenditures, in millions

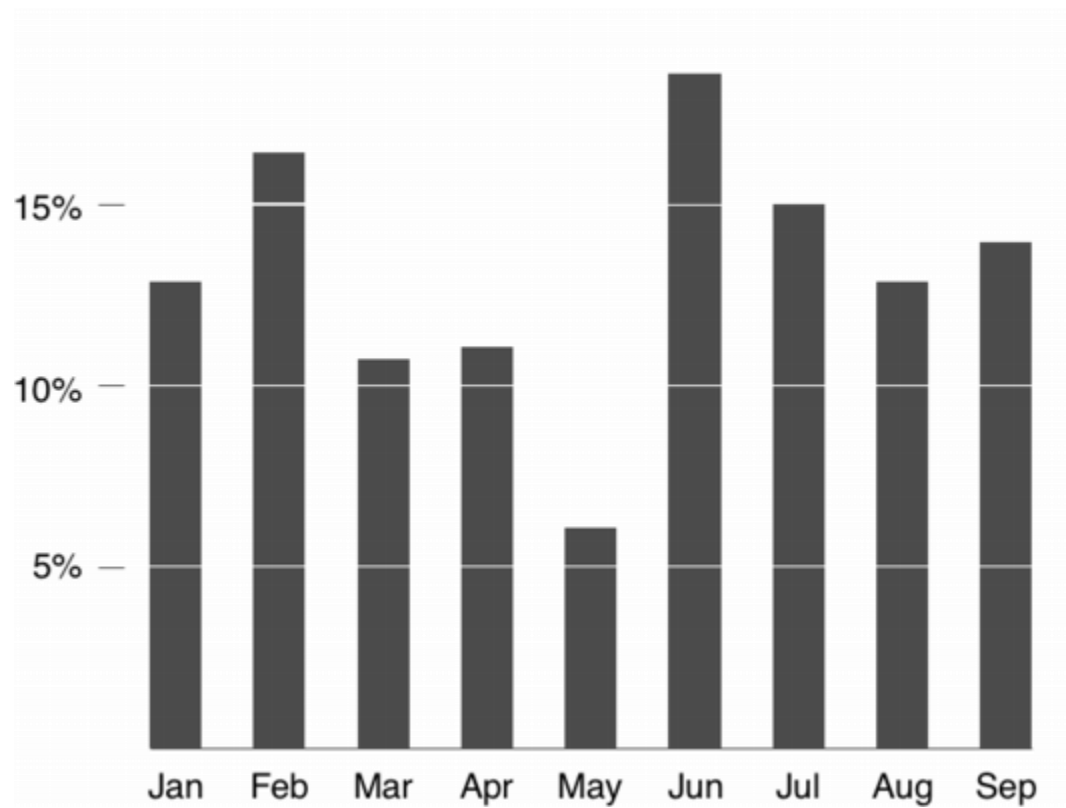


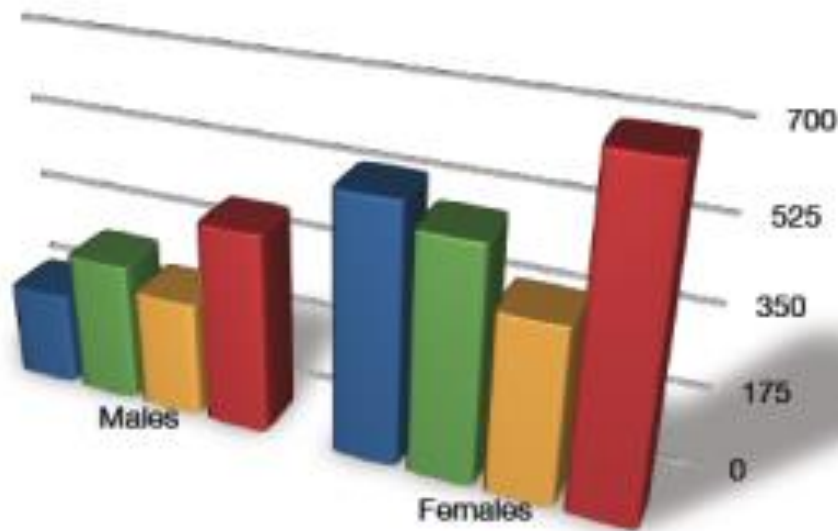
$$\text{Data-ink ratio} = \frac{\text{Data-ink}}{\text{Total ink used to print the graphic}}$$

Low data-ink ratio



High data-ink ratio

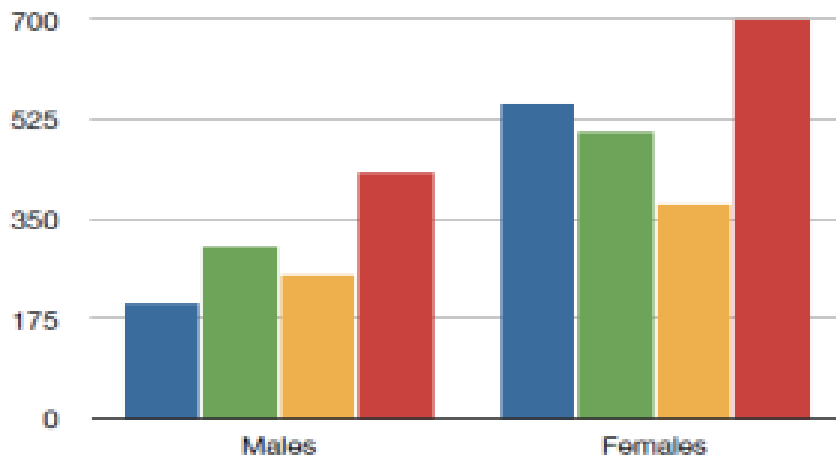




- Low data-ink ratio



■ 0-\$24,999 ■ \$25,000+ ■ 0-\$24,999 ■ \$25,000+



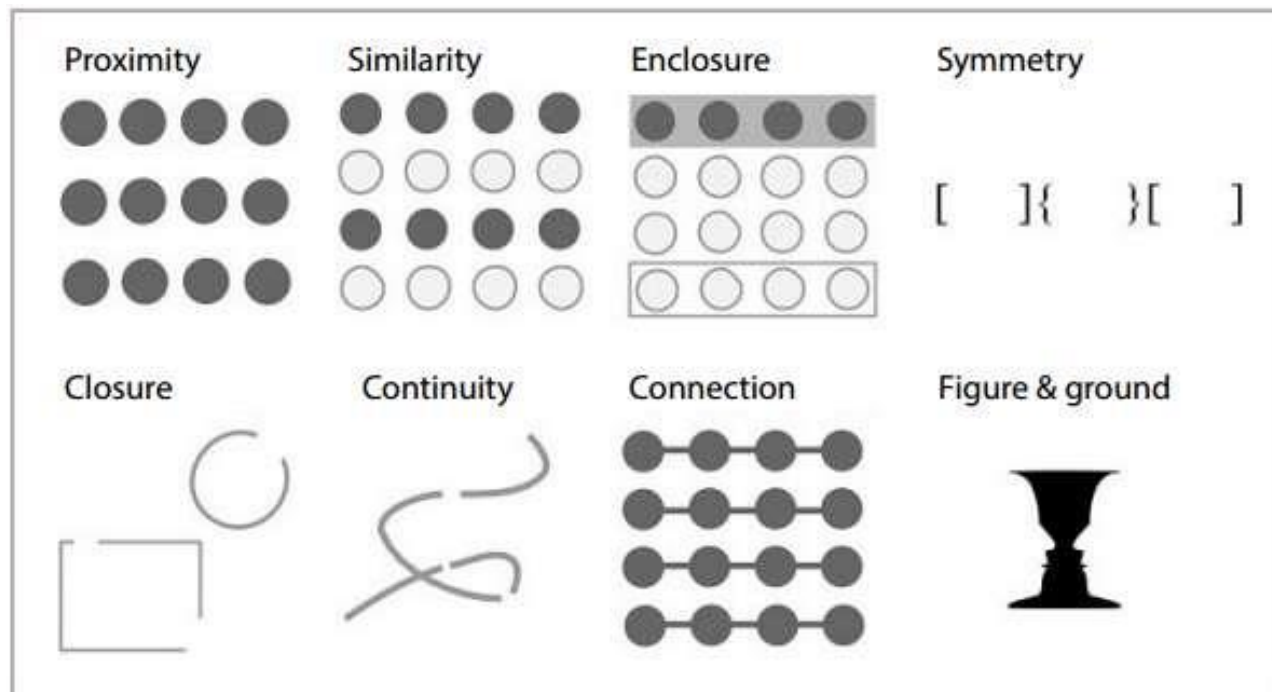
- High data-ink ratio



■ 0-\$24,999 ■ \$25,000+ ■ 0-\$24,999 ■ \$25,000+

The Gestalt Laws of Visual Perception

Gestalt is a psychology term which means "unified whole". It refers to theories of **visual perception** developed by German psychologists in the 1920s. These theories attempt to describe how people tend to organize visual elements into **groups** or *unified wholes* when certain principles are applied. These principles are:



Gestalt Laws: Proximity

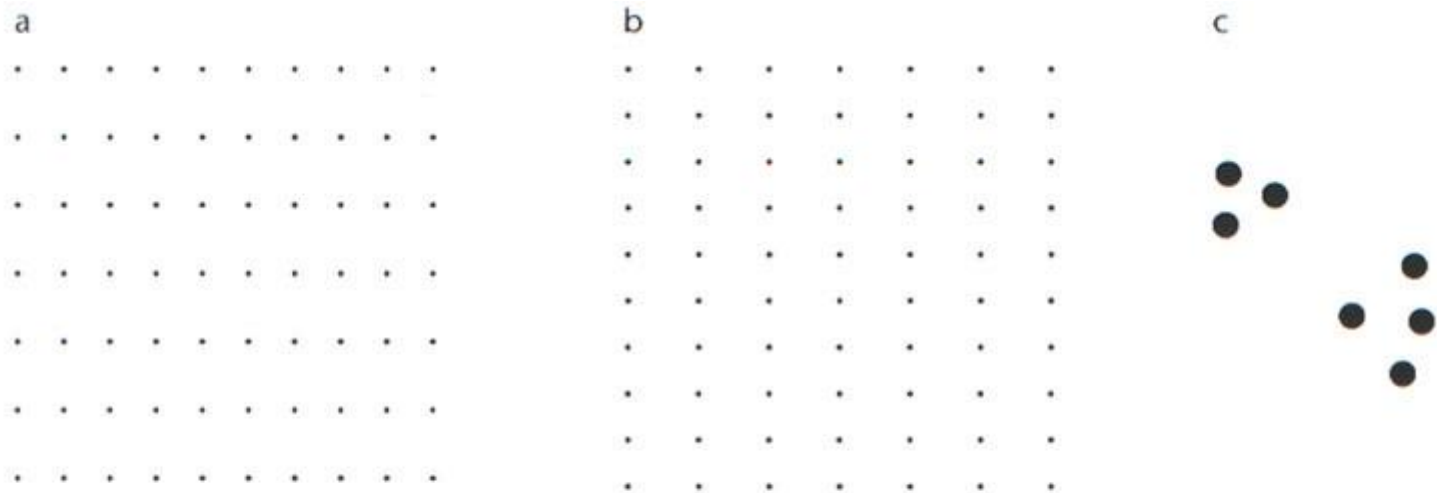
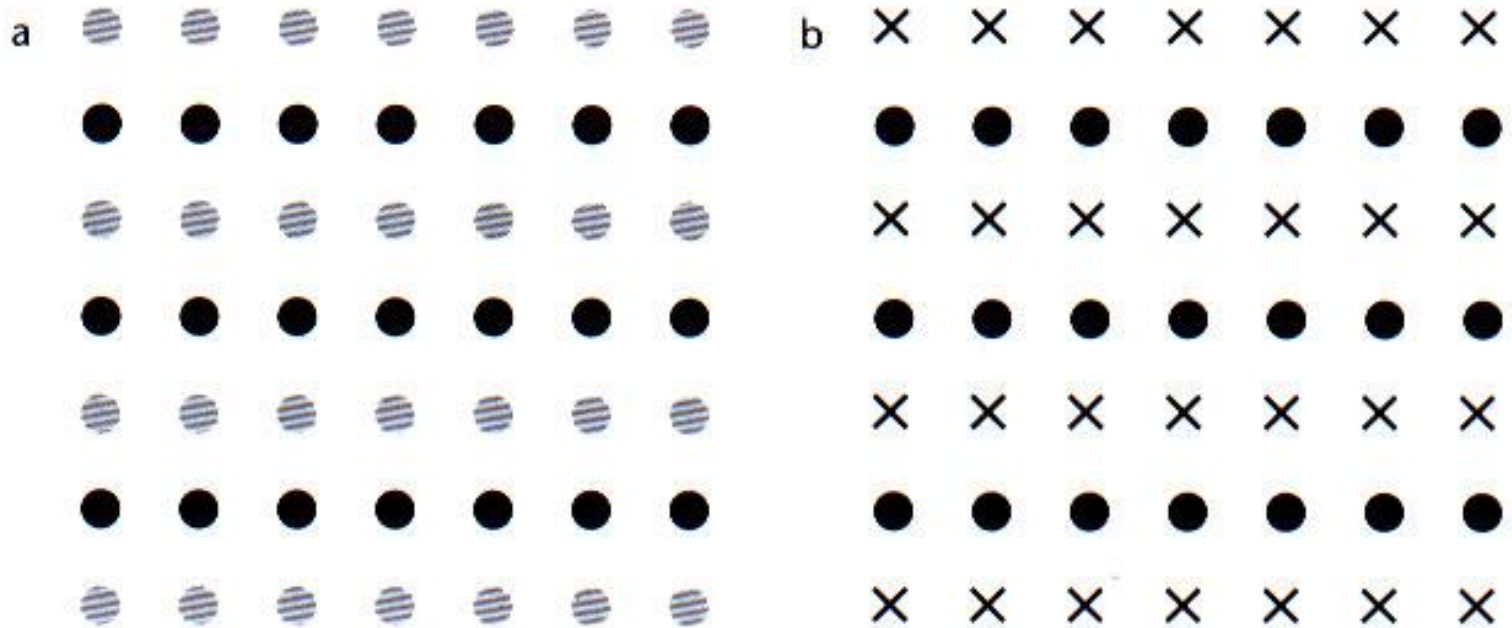


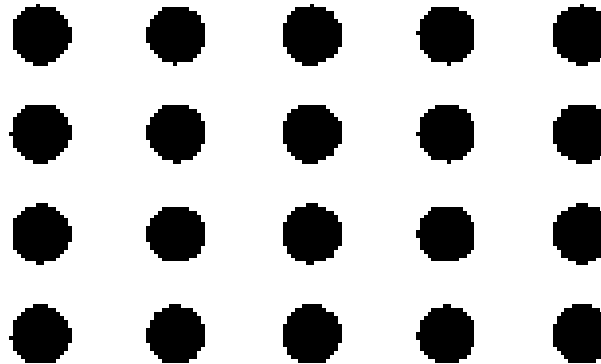
Figure 6.2 Spatial proximity is a powerful cue for perceptual organization. A matrix of dots is perceived as rows on the left (a) and columns on the right (b). In (c), because of proximity relationships, we perceive two groupings of dots.

Gestalt Laws: Similarity

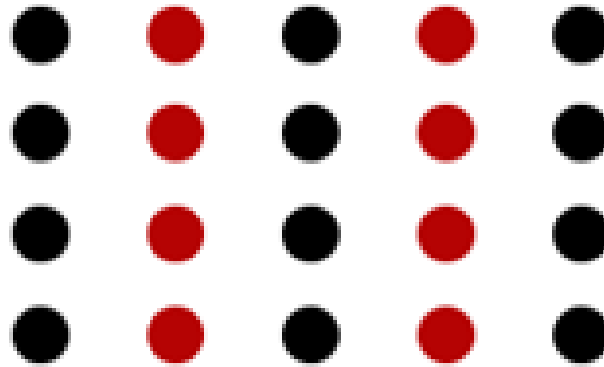


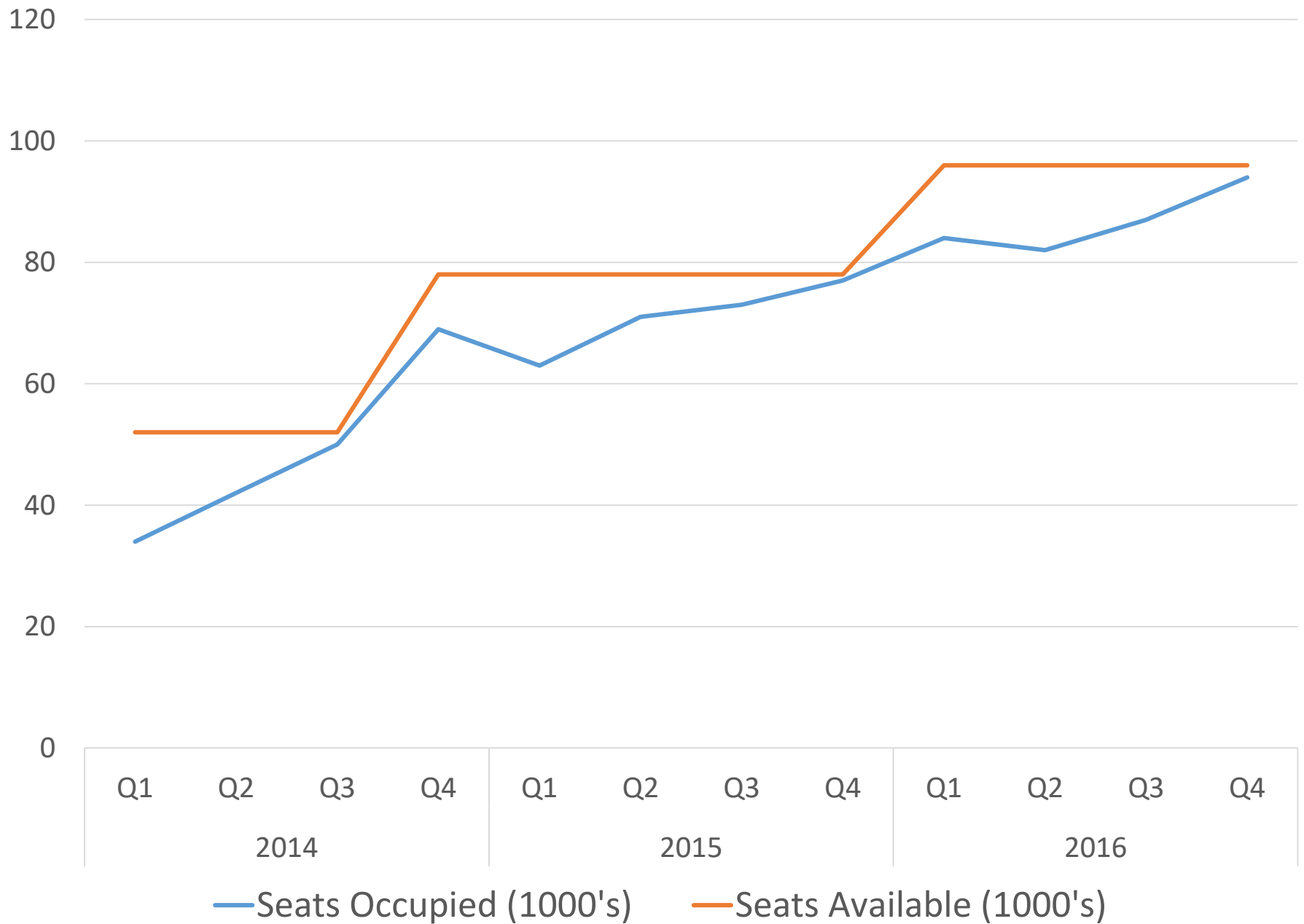
Ware C. (2004). *Information Visualization: Perception for Design*. Elsevier Inc: San Francisco, CA, p.189.

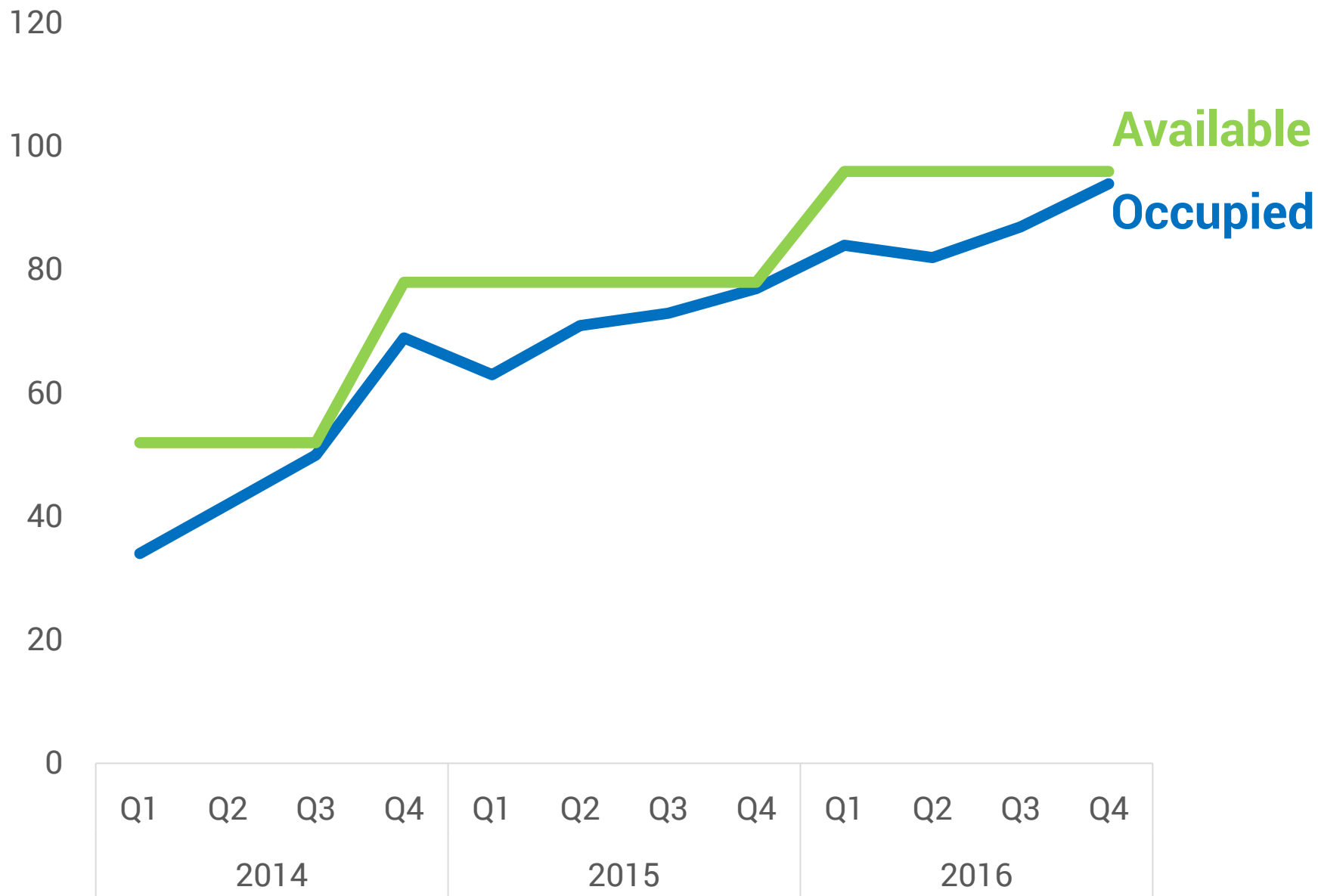
Gestalt Laws: Similarity

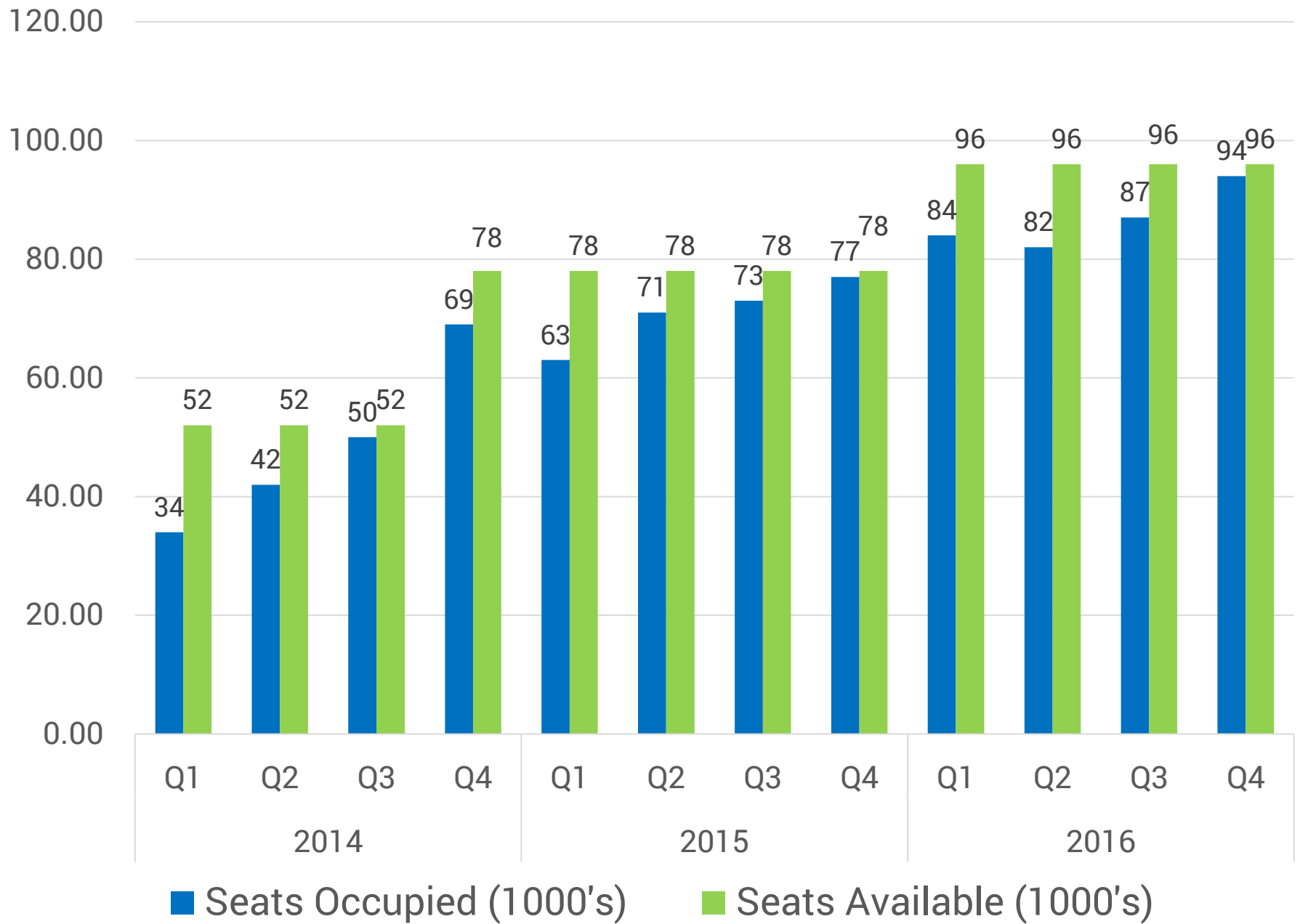


Gestalt Laws: Similarity

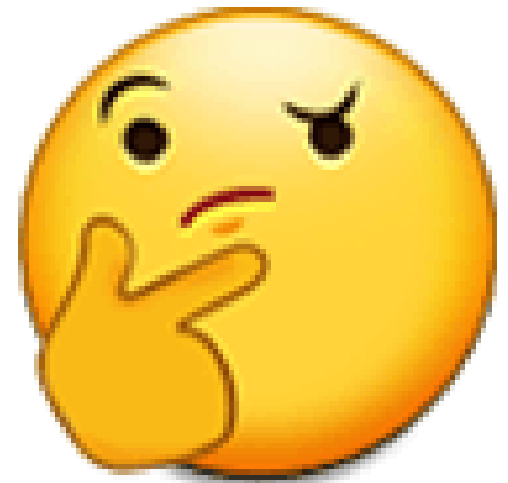








What we have so far is just a chart, not a data story.



KEYS TO DATA STORYTELLING



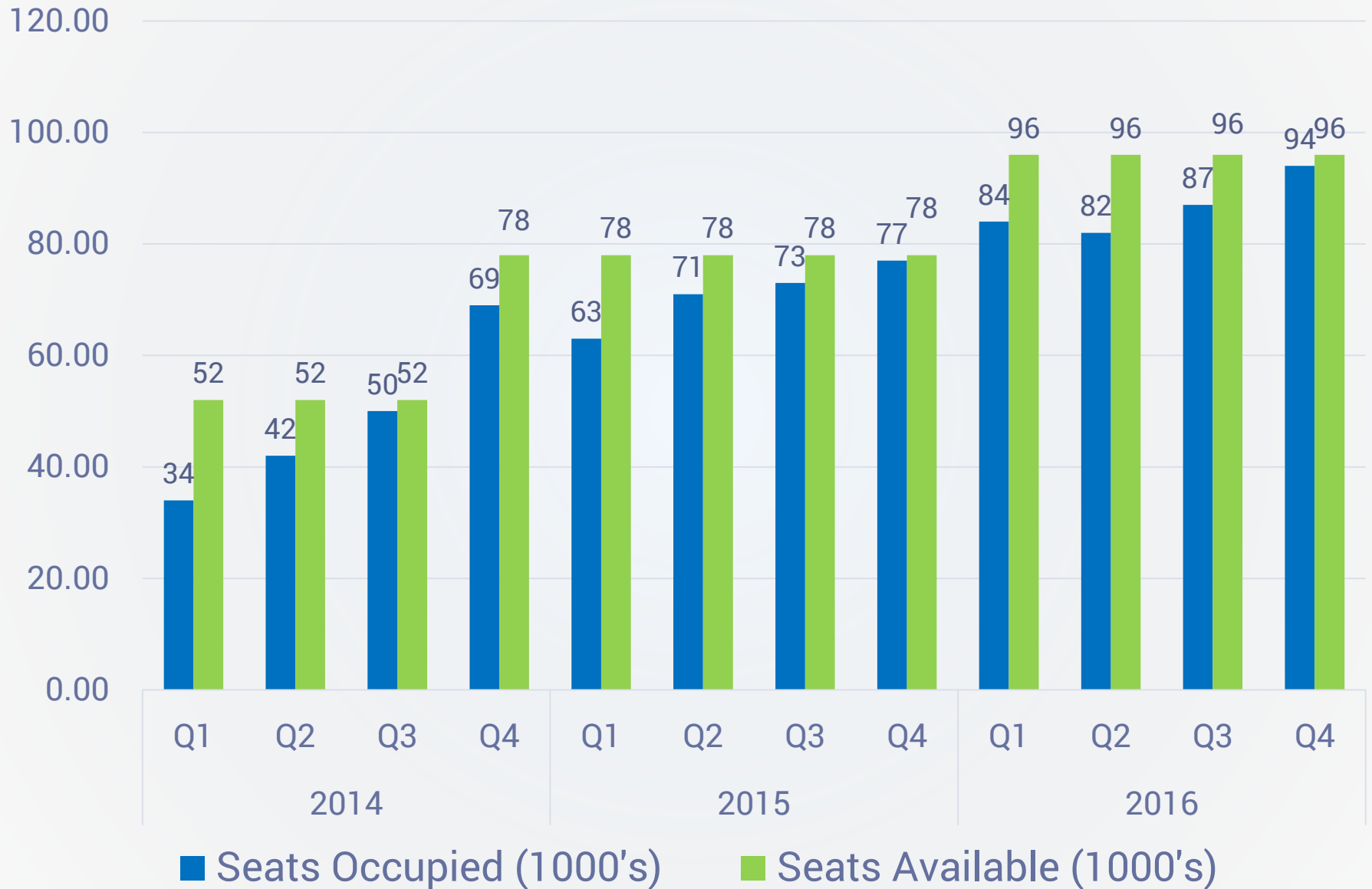
AUDIENCE



KEYS TO DATA STORYTELLING

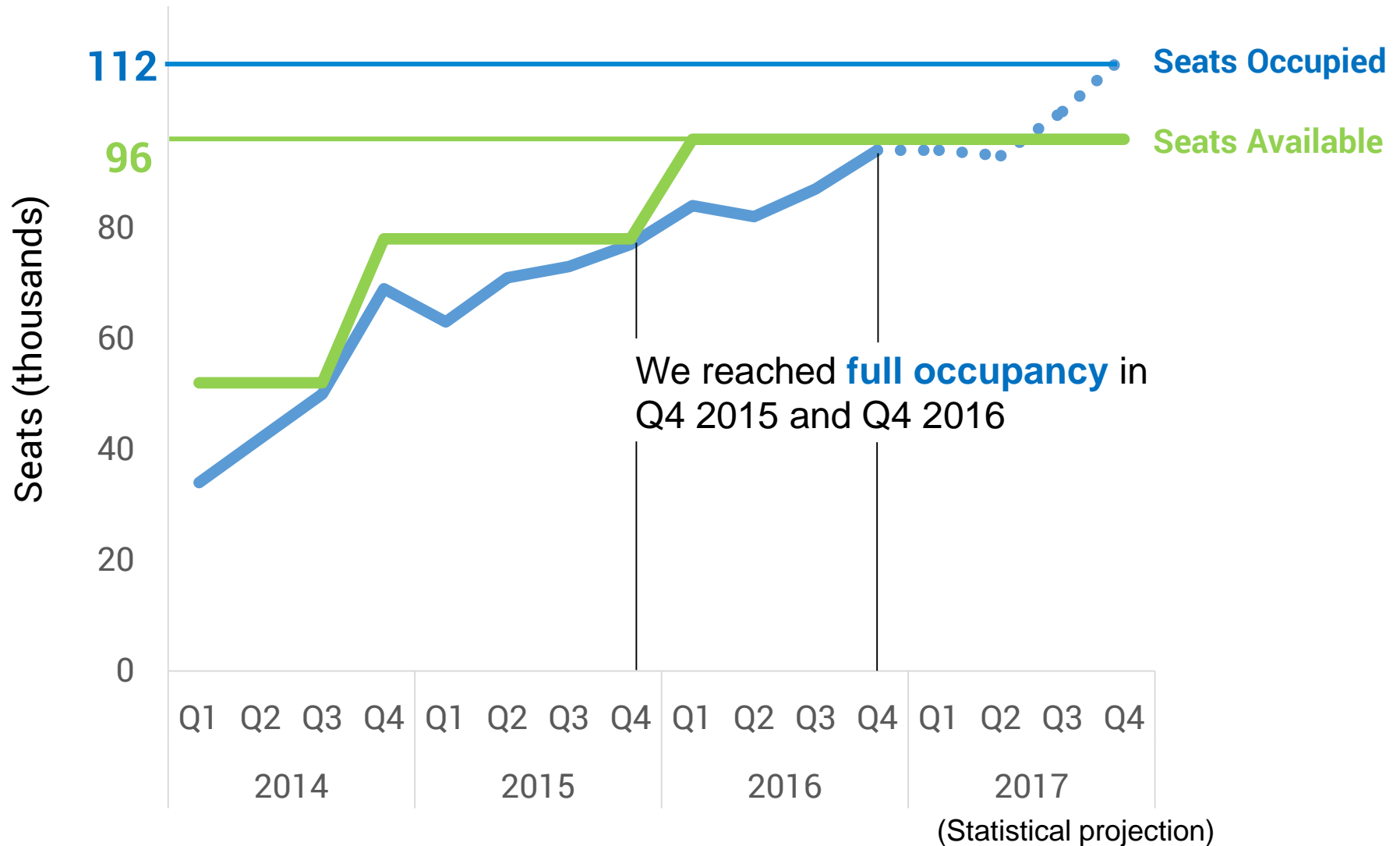


Sydney to Los Angeles Route



Please approve an extra **20,000 seats** for Q4 2017

If we don't add seats to Q4 '17, we risk leaving revenue on the table



Position



Length



Angle



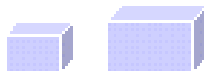
Slope



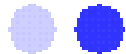
Area



Volume



Color



Density



More
Accurate

Less
Accurate

WHY?

Accuracy
Ranking of
Quantitative
Perceptual
Tasks
Estimated;
only pairwise
comparisons
have been
validated
(Cleveland &
McGill,1984)



Data Storytelling for Business

**Jan 23-24, Manila
Feb 27-28, Singapore**



Statistics for Business Applications

**Feb 13-15, Manila
Mar 28-30, Singapore**



Predictive Analytics for Data-Driven Decision Making and Profit Maximization

Apr 19-21, Manila



Predicting Customer Behavior and Generating Revenue with Machine Learning

Apr 24-26, Manila



Isaac Reyes -
Data Scientist
@isaac.reyes

- Home
- About
- Photos
- Likes
- Posts
- Create a Page



Liked

Following

Message

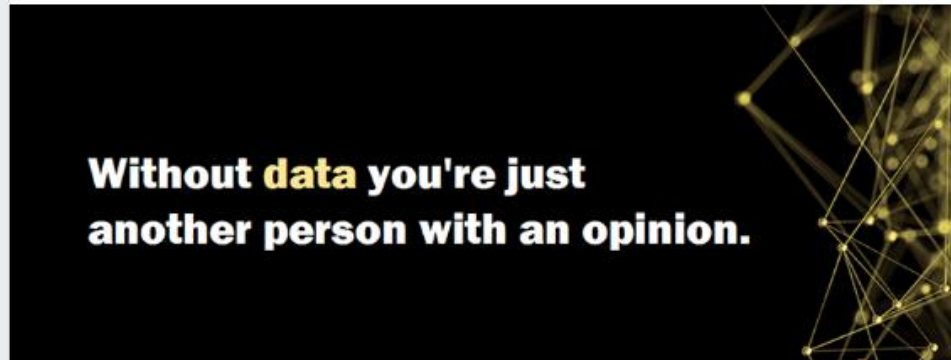
More

Send Email

Status Photo / Video

Write something on this Page...

Photos See All



Public Figure

Search for posts on this Page

Invite friends to like this Page

I teach the art and science of data storytelling.

629 Likes
Kai Zhao and 6 other friends like this

About See All

Typically replies within an hour
Message Now

NBC2 VIEWER VOTE

NBC-2.COM

HOW CONCERNED ARE YOU ABOUT THE ZIKA VIRUS?



Isaac Reyes - Data Scientist

Like This Page · 18 October 2016 · Edited ·

I'm 100% more concerned about this bar chart than Zika.

Tag photo Add location Edit

Like Comment Share

2.4k

Chronological

309 shares

44 comments



Dan Meyer Ouch... epic fail

Like · Reply · Message · 18 October 2016 at 09:52



Brandon Carlisle Go home bar chart, you're drunk!

Unlike · Reply · Message · 28 · 18 October 2016 at 18:54



Cleigh Yu pastilan. sa mauwat lang

See translation

Like · Reply · Message · 3 · 18 October 2016 at 19:01



John Peter Ibañez Uhhh is this accurate. Sarcasm*

Write a comment...



Q&A

DataSeer