

MODULE III

TALKING ABOUT DATA

INFO 590: Applied Data Science
Summer 2017

DISCUSSING DATA

Why are we (the data scientists) showing this data?

The receiver does not need all the complexity

Charts can never stand on their own

Titles, axes, legends and explanation

DISCUSSING DATA

Analysis: all the data

Visualization: only the most important data

Communication: all the important data

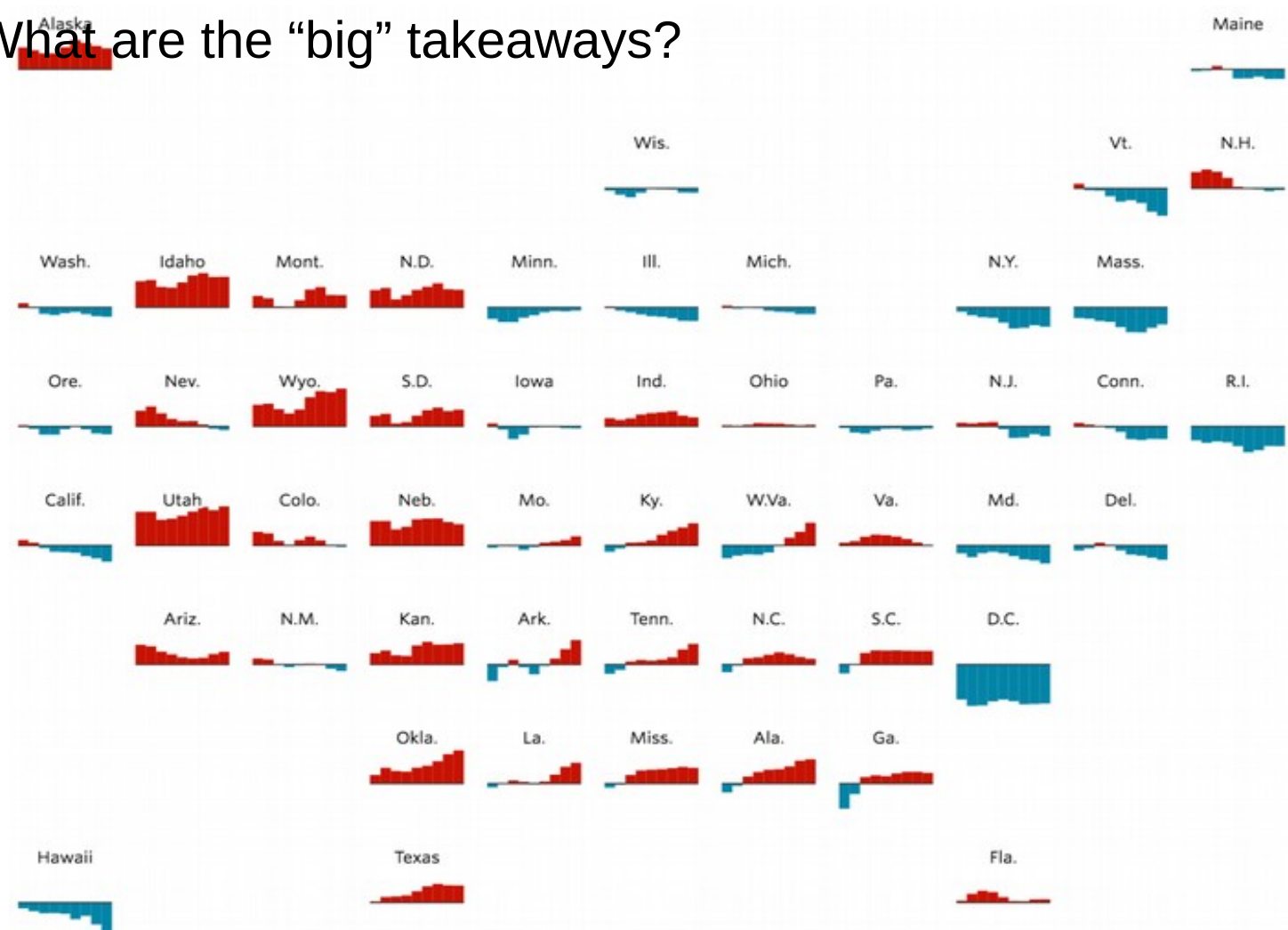
EXAMPLE

Ex: State Voting Trend Map

What is the overall trend?

What are the outliers?

What are the “big” takeaways?



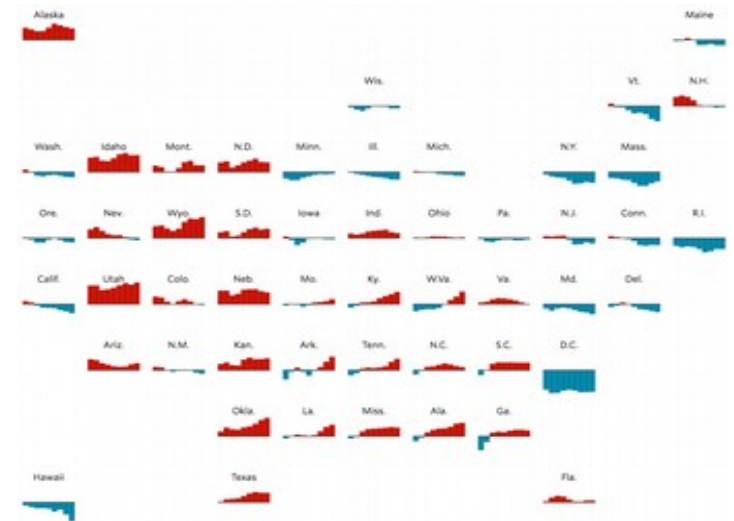
EXAMPLE

Ex: State Voting Trend Map

Rise in Republican voting in “south”, reverse in North East and New England.

Ohio, Michigan and Pennsylvania and Iowa vote approximately the same as the country - Michigan, Iowa and Pennsylvania trend towards more blue; Ohio more red

NH and Georgia have changed dramatically



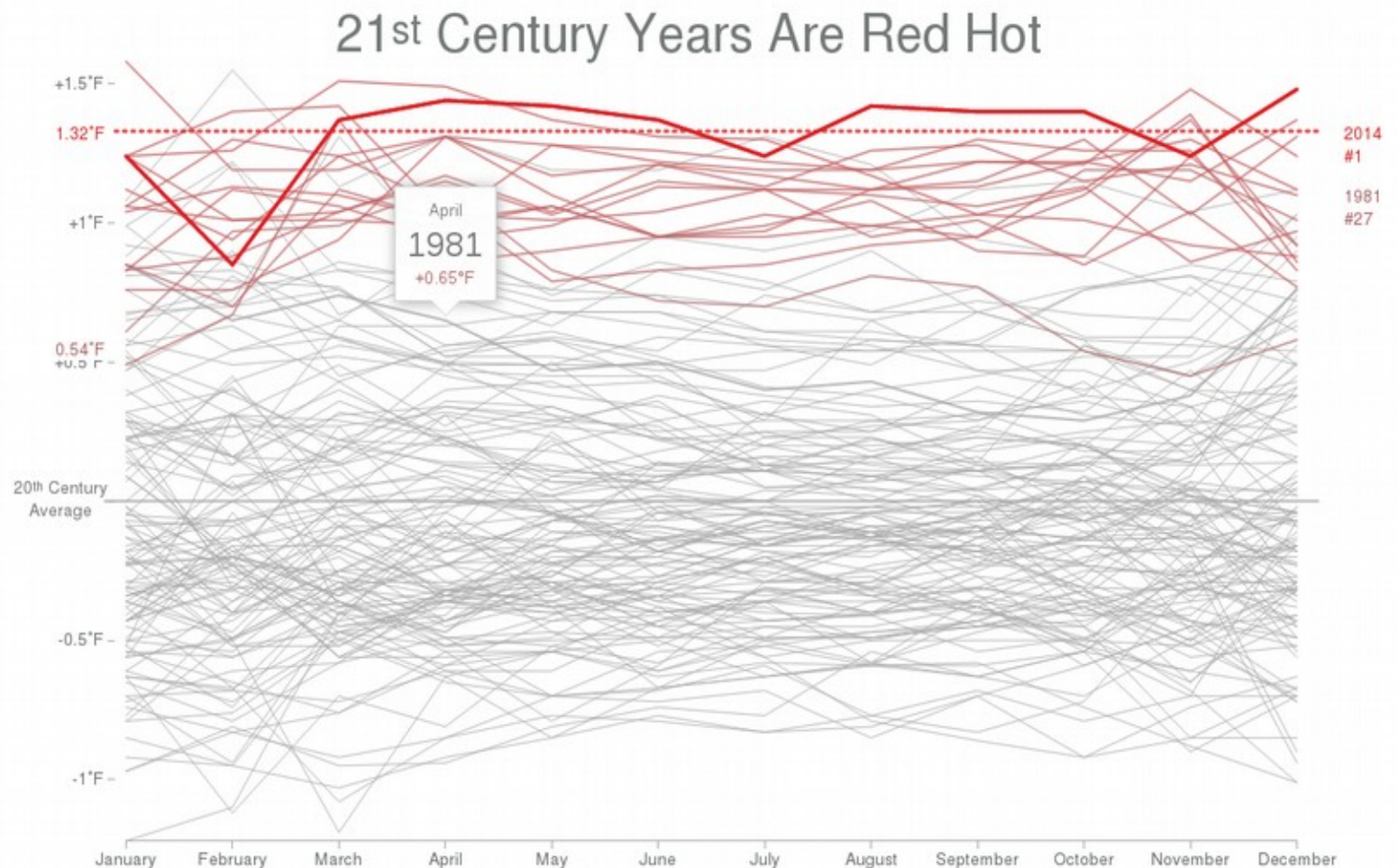
EXAMPLE

Ex: Hottest years of 21st century

What is the overall trend?

What are the outliers?

What are the “big” takeaways?



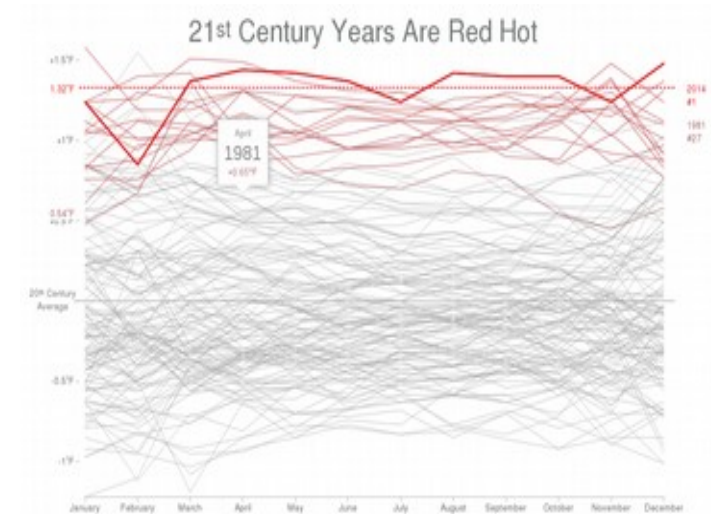
EXAMPLE

Ex: Hottest years of 21st century

2014 was the hottest year in the 21st century (1.32 degrees above average)

16 of the 17 warmest years on record have occurred after 2000

2003 (+1.10 degrees),
2005 (+1.17 degrees),
and 2010 (+1.26 degrees) were also record breaking years



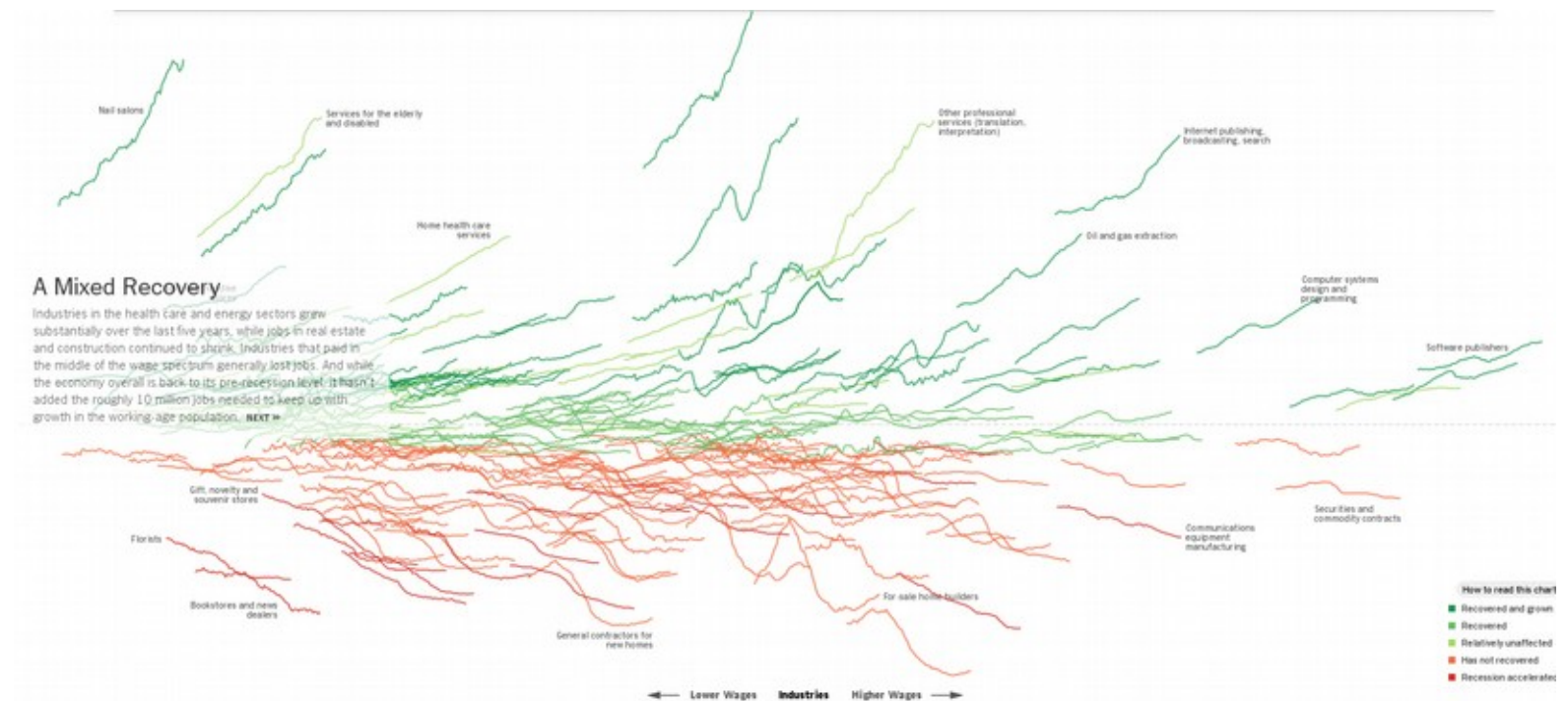
EXAMPLE

Ex: Job recovery

What is the overall trend?

What are the outliers?

What are the “big” takeaways?



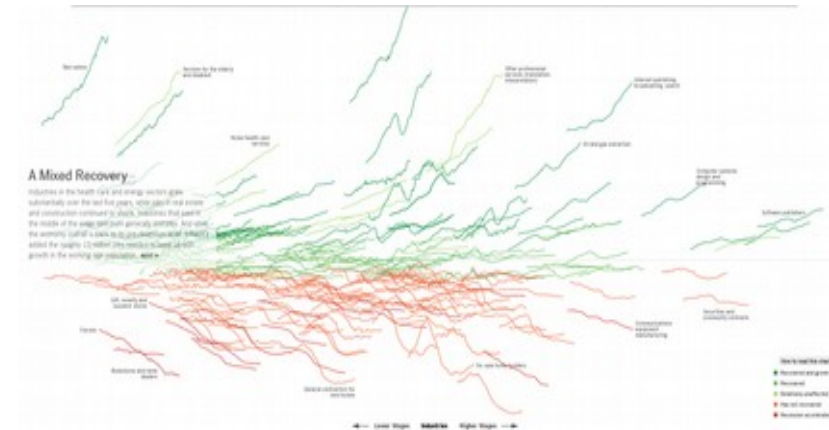
EXAMPLE

Ex: Job recovery

Computer, software and internet related jobs are growing (software publishers added 100k jobs, programming added 750k, internet publishing added 100k)

“Old way” of phones, construction, books is dying (200k less general contractors, 110k less telecom, 50k less in books)

Health care, elderly care, and services also on the rise



DISCUSSING DATA

Charts can never stand on their own

Use words to drive home key points not made by the graphic, provide numbers give a more full picture