DLI Accelerated Data Science Teaching Kit

Lecture 7.1 - What is Info Vis and Why it is Important

Information Visualization Crash Course

(AKA Information Visualization 101)

Chad Stolper
Software Engineer at Google
(graduated from Georgia Tech CS PhD)

What is Information Visualization?

Information Visualization

- “The use of **computer**-supported, **interactive**, **visual** representations of abstract data to **amplify cognition**.”

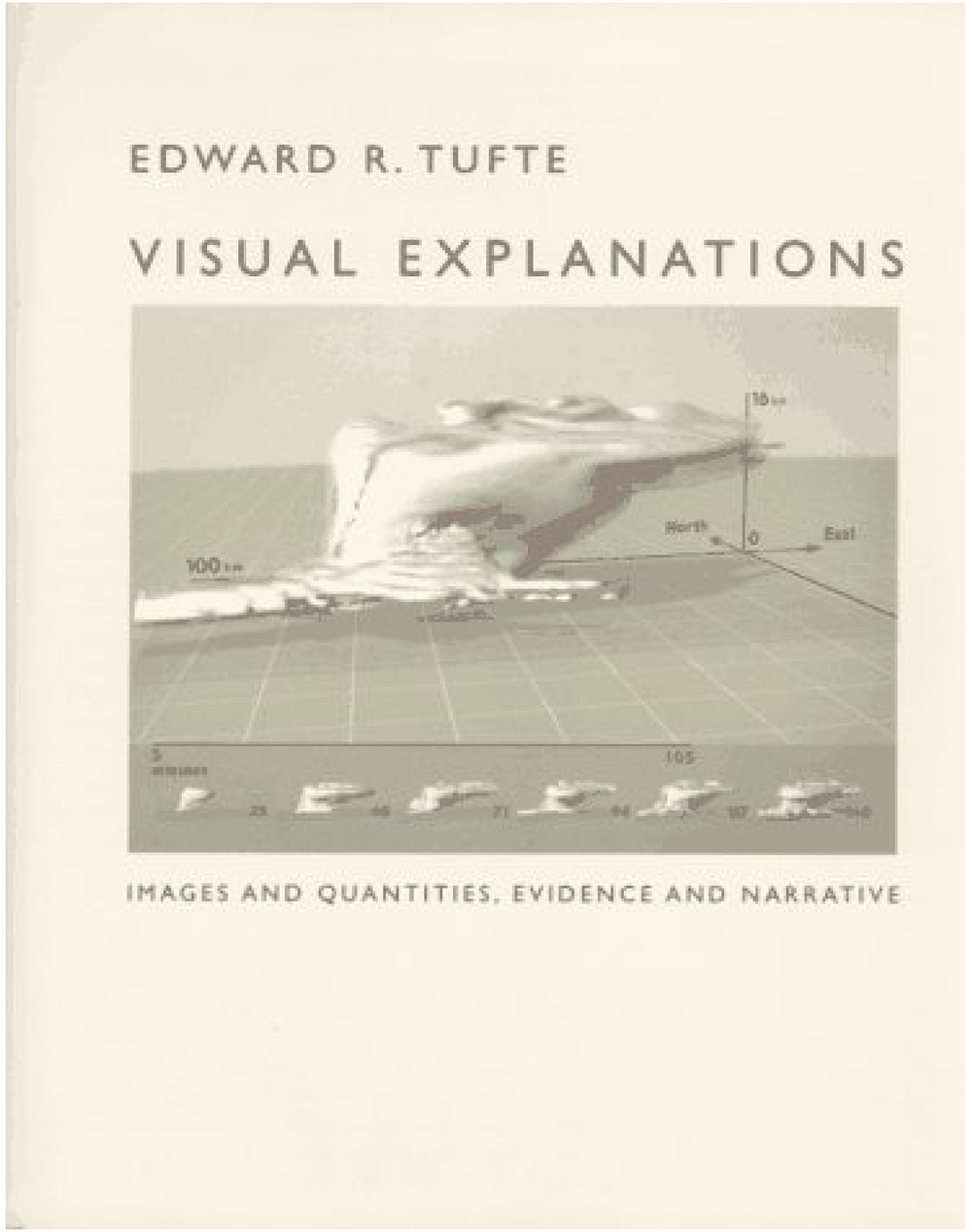
- Card, Mackinlay, and Shneiderman 1999

- **Communication**

- **Exploratory Data Analysis** (EDA)

Communication

(gone wrong)



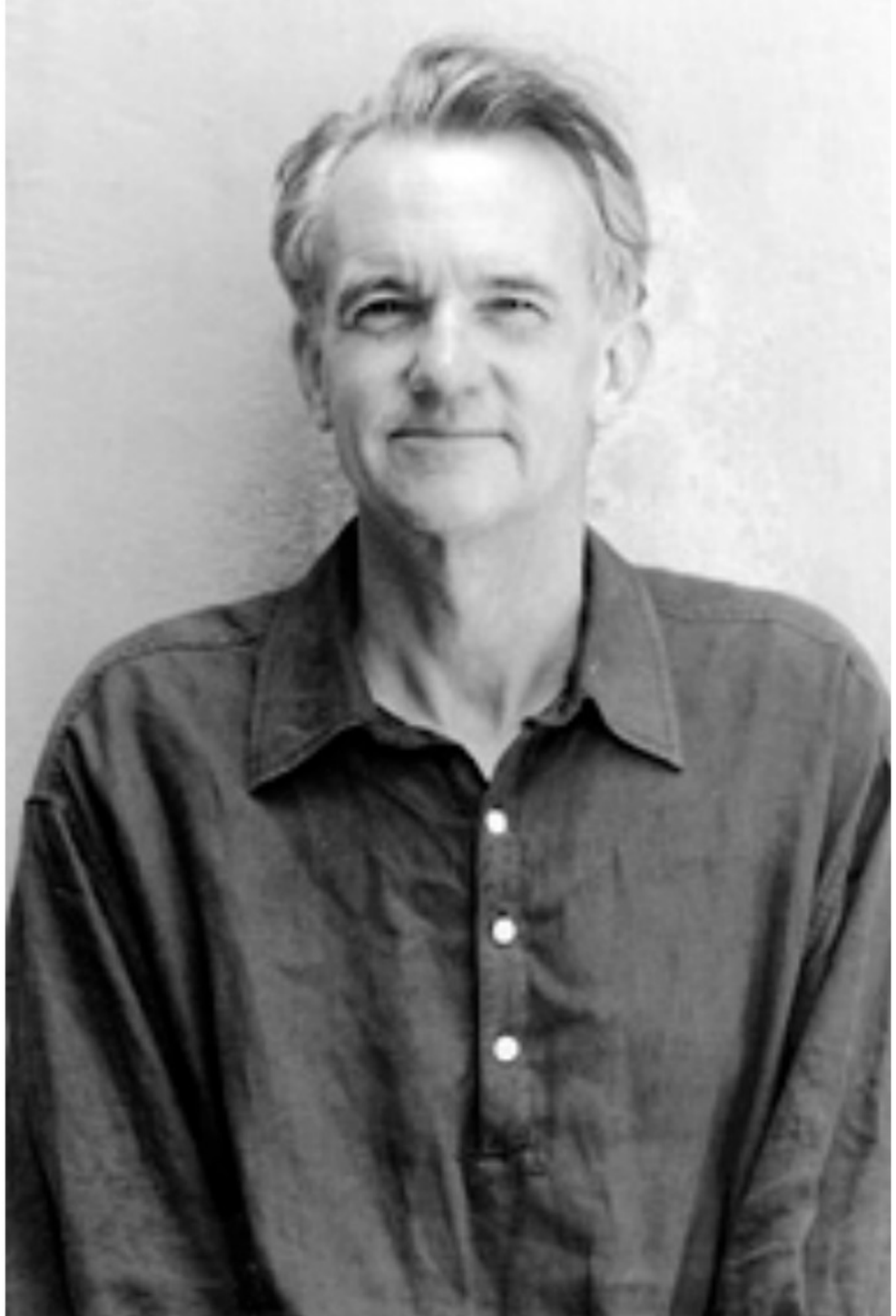
Tufte, E. R. (2012). Visual explanations: images and quantities, evidence and narrative. Cheshire, CT: Graphics Press.

Edward Tufte

An American statistician and professor emeritus of political science, statistics, and computer science at Yale University.

He is noted for his writings on information design and as a pioneer in the field of data visualization.

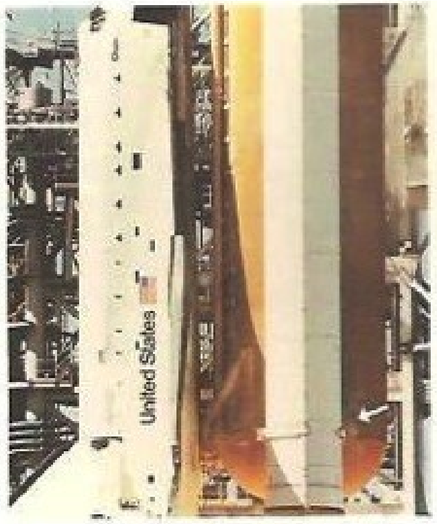
-Wikipedia



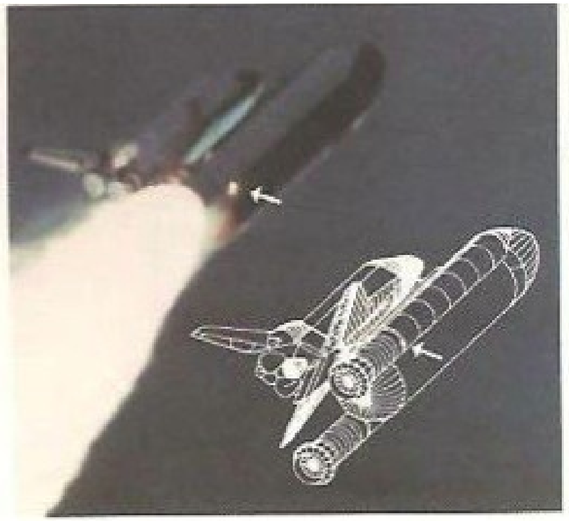
Space Shuttle Challenger

January 28, 1986

Morning Temperature: 31°F



Less than 1 second after ignition, a puff of smoke appeared at the aft joint of the right booster, indicating that the O-rings burned through and failed to seal. At this point, all was lost.

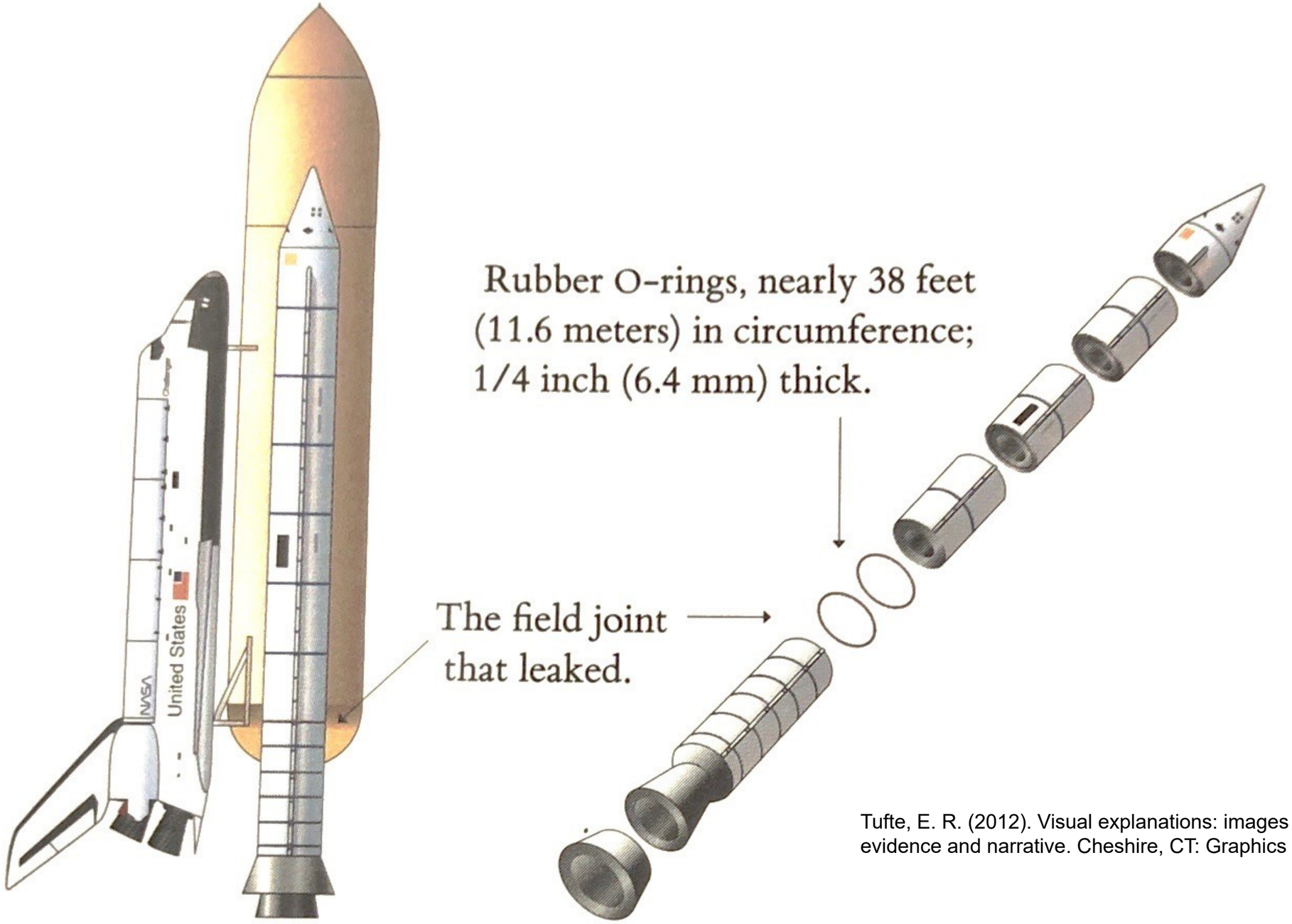


On the launch pad, the leak lasted only about 2 seconds and then apparently was plugged by [illegible] and insulation as the shuttle rose, flying through rather strong cross-winds. Then 58.788 seconds ignition, when the Challenger was 6 miles up, a flicker of flame emerged from the leaky joint seconds, the flame grew and engulfed the fuel tank (containing liquid hydrogen and liquid [illegible] That tank ruptured and exploded, destroying the shuttle.

As the shuttle exploded and broke up at approximately 73 seconds after launch, the two booster rockets crisscrossed and continued flying wildly. The right booster, identifiable by its failure plume, is now to the left of its non-defective counterpart.

The flight crew of Challenger 51-L. Front row, left to right: [illegible] Smith, pilot; Francis R. (Dick) Scobee, commander; Ronald [illegible] Back row: Ellison S. Onizuka, S. Christa McAuliffe, George [illegible] Judith A. Resnik.

Tufte, E. R. (2012). Visual explanations: images and quantities, evidence and narrative. Cheshire, CT: Graphics Press.



Tufte, E. R. (2012). Visual explanations: images and quantities, evidence and narrative. Cheshire, CT: Graphics Press.

Most Watched Science Experiment

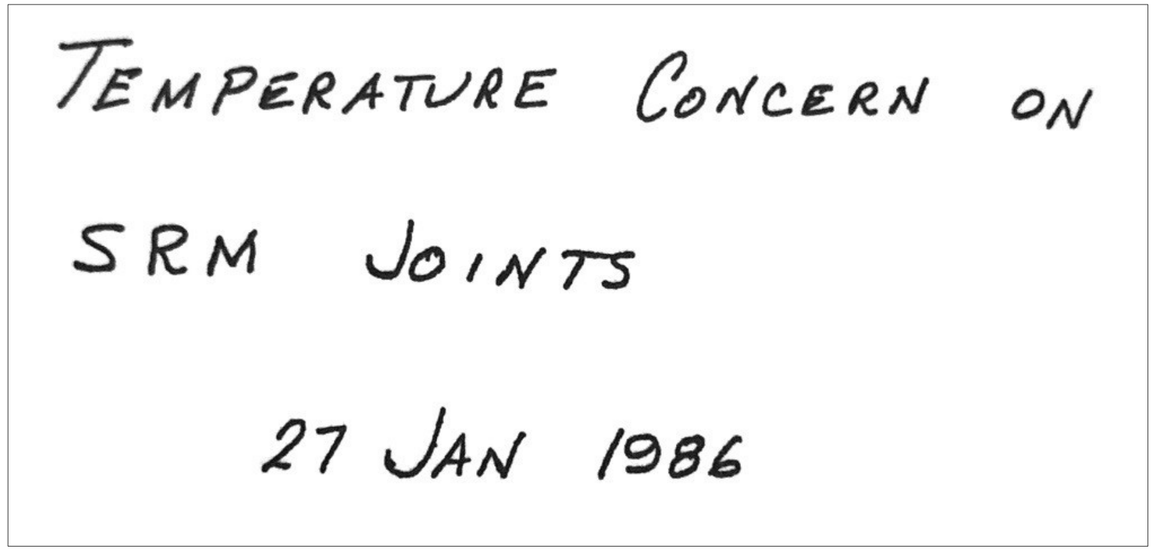
Richard Feynman, Physics Nobel laureate explained how rubber became rigid in cold temperate

YouTube video:
https://youtu.be/6Rwcbsn19c0

Video originally from: http://www.FeynmanPhysicsLectures.com

How did this happen?

Engineers at Morton Thiokol, the rocket maker, presented on the day before and recommended not to launch.



Tufte, E. R. (2012). Visual explanations: images and quantities, evidence and narrative. Cheshire, CT: Graphics Press.

CONCLUSIONS :

- TEMPERATURE OF O-RING IS NOT ONLY PARAMETER CONTROLLING BLOW-BY

  SRM 15 WITH BLOW-BY HAD AN O-RING TEMP AT 53°F
  SRM 22 WITH BLOW-BY HAD AN O-RING TEMP AT 75°F
  FOUR DEVELOPMENT MOTORS WITH NO BLOW-BY WERE TESTED AT O-RING TEMP OF 47° TO 52°F

  DEVELOPMENT MOTORS HAD PUTTY PACKING WHICH RESULTED IN BETTER PERFORMANCE

- AT ABOUT 50°F BLOW-BY COULD BE EXPERIENCED IN CASE JOINTS

- TEMP FOR SRM 25 ON 1-28-86 LAUNCH WILL BE 29°F 9 AM
  38°F 2 PM

- HAVE NO DATA THAT WOULD INDICATE SRM 25 IS DIFFERENT THAN SRM 15 OTHER THAN TEMP

RECOMMENDATIONS :

- O-RING TEMP MUST BE ≥ 53°F AT LAUNCH

  DEVELOPMENT MOTORS AT 47° TO 52°F WITH PUTTY PACKING HAD NO BLOW-BY SRM 15 (THE BEST SIMULATION) WORKED AT 53°F

- PROJECT AMBIENT CONDITIONS (TEMP & WIND) TO DETERMINE LAUNCH TIME

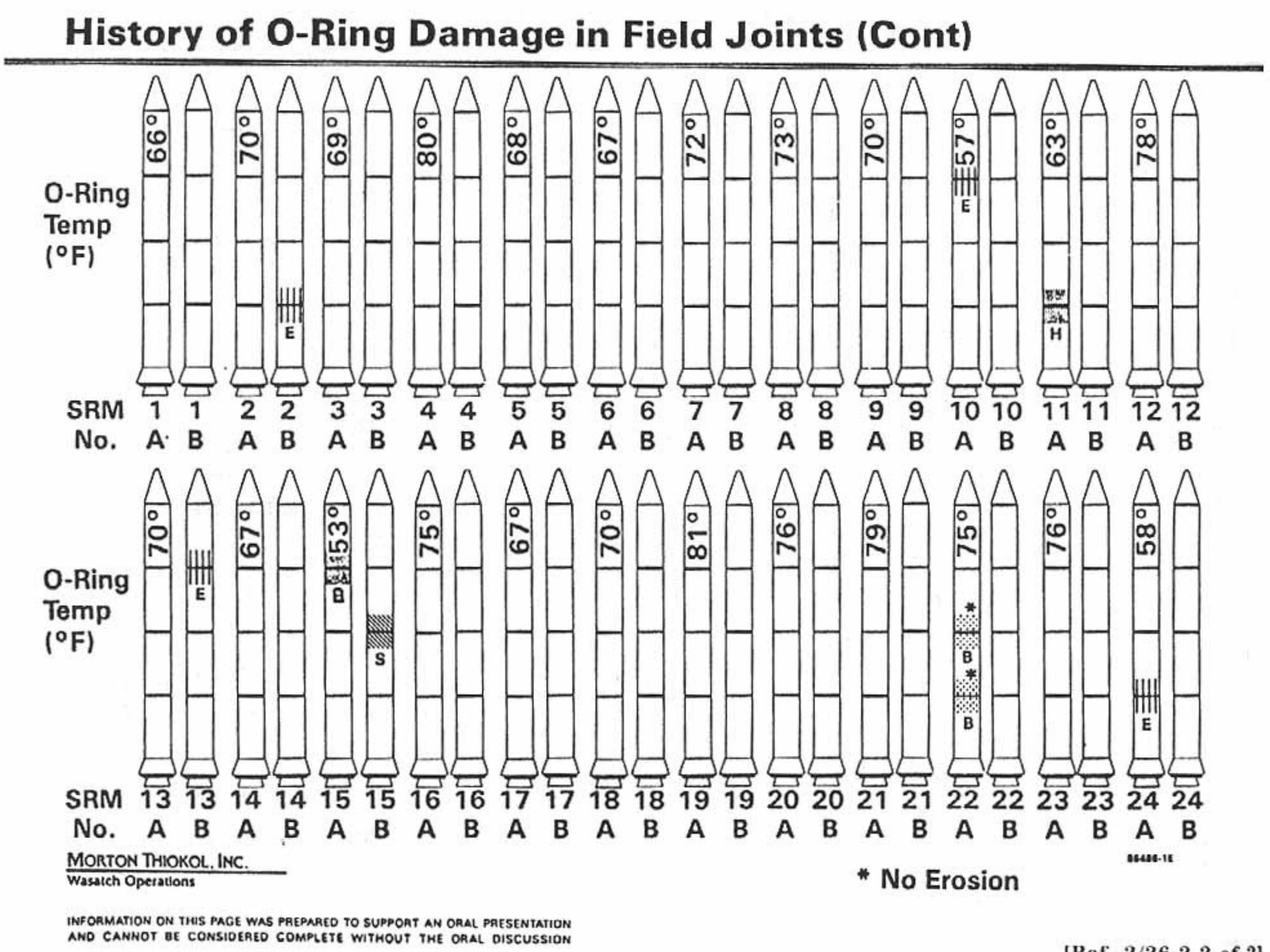
RECOMMENDATIONS :

- O-RING TEMP MUST BE ≥ 53°F AT LAUNCH

  DEVELOPMENT MOTORS AT 47° TO 52°F WITH PUTTY PACKING HAD NO BLOW-BY SRM 15 (THE BEST SIMULATION) WORKED AT 53°F

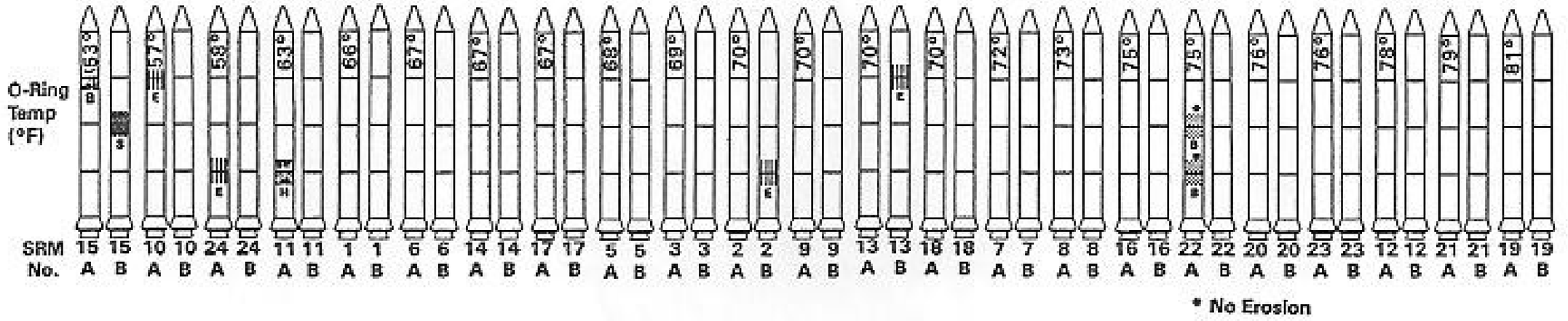
- PROJECT AMBIENT CONDITIONS (TEMP & WIND) TO DETERMINE LAUNCH TIME

Tufte, E. R. (2012). Visual explanations: images and quantities, evidence and narrative. Cheshire, CT: Graphics Press.

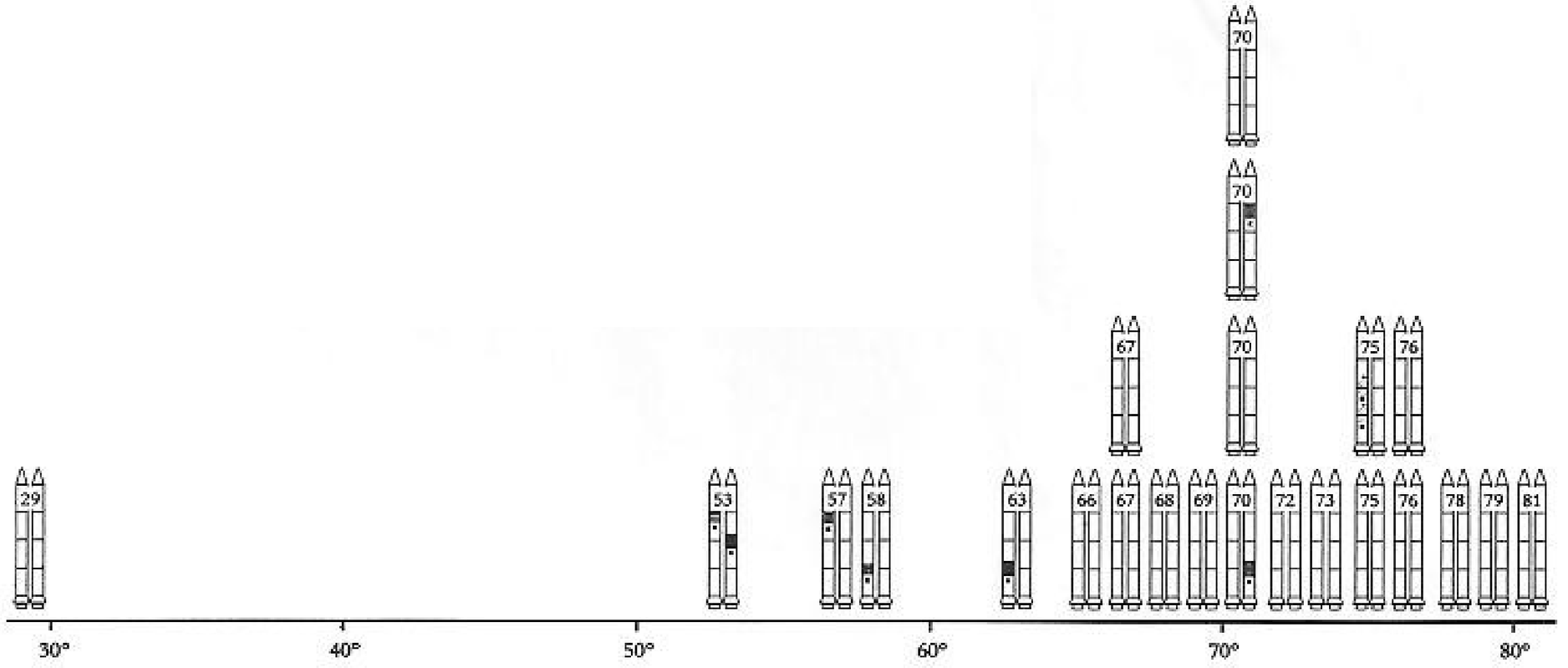


[Ref. 2/26-2 2 of 3]

Tufte, E. R. (2012). Visual explanations: images and quantities, evidence and narrative. Cheshire, CT: Graphics Press.



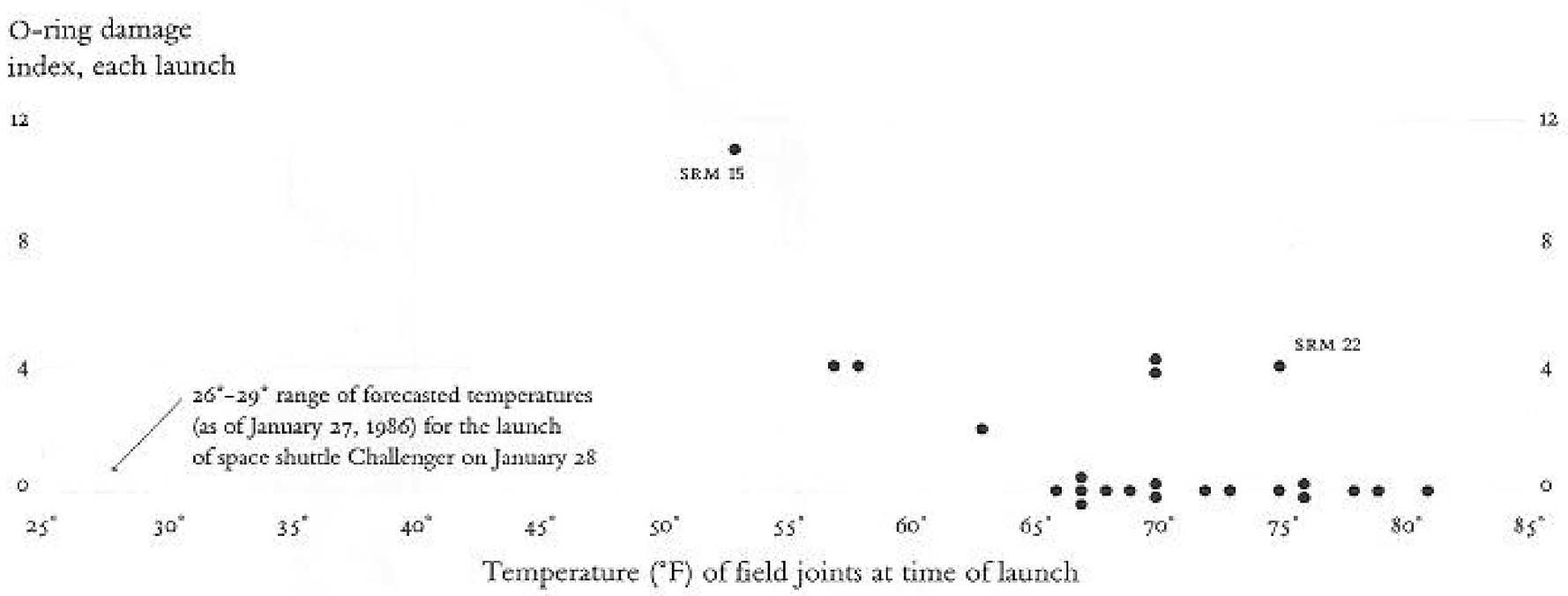
Tufte, E. R. (2012). Visual explanations: images and quantities, evidence and narrative. Cheshire, CT: Graphics Press.



Tufte, E. R. (2012). Visual explanations: images and quantities, evidence and narrative. Cheshire, CT: Graphics Press.

Flight	Date	Temperature °F	Erosion incidents	Blow-by incidents	Damage index	Comments
51-C	01.24.85	53°	3	2	11	Most erosion any flight; blow-by; back-up rings heated.
41-B	02.03.84	57°	1		4	Deep, extensive erosion.
61-C	01.12.86	58°	1		4	O-ring erosion on launch two weeks before Challenger.
41-C	04.06.84	63°	1		2	O-rings showed signs of heating, but no damage.
1	04.12.81	66°			0	Coolest (66°) launch without O-ring problems.
6	04.04.83	67°			0	
51-A	11.08.84	67°			0	
51-D	04.12.85	67°			0	
5	11.11.82	68°			0	
3	03.22.82	69°			0	
2	11.12.81	70°	1		4	Extent of erosion not fully known.
9	11.28.83	70°			0	
41-D	08.30.84	70°	1		4	
51-G	06.17.85	70°			0	
7	06.18.83	72°			0	
8	08.30.83	73°			0	
51-B	04.29.85	75°			0	
61-A	10.30.85	75°		2	4	No erosion. Soot found behind two primary O-rings.
51-I	08.27.85	76°			0	
61-B	11.26.85	76°			0	
41-G	10.05.84	78°			0	
51-J	10.03.85	79°			0	
4	06.27.82	80°			?	O-ring condition unknown; rocket casing lost at sea.
51-F	07.29.85	81°			0	

Tufte, E. R. (2012). Visual explanations: images and quantities, evidence and narrative. Cheshire, CT: Graphics Press.



Tufte, E. R. (2012). Visual explanations: images and quantities, evidence and narrative. Cheshire, CT: Graphics Press.

So, communication is **extremely important.**

Visualization can help with that – **communicate ideas and insights.**

Visualization can also help with
Exploratory Data Analysis (EDA)

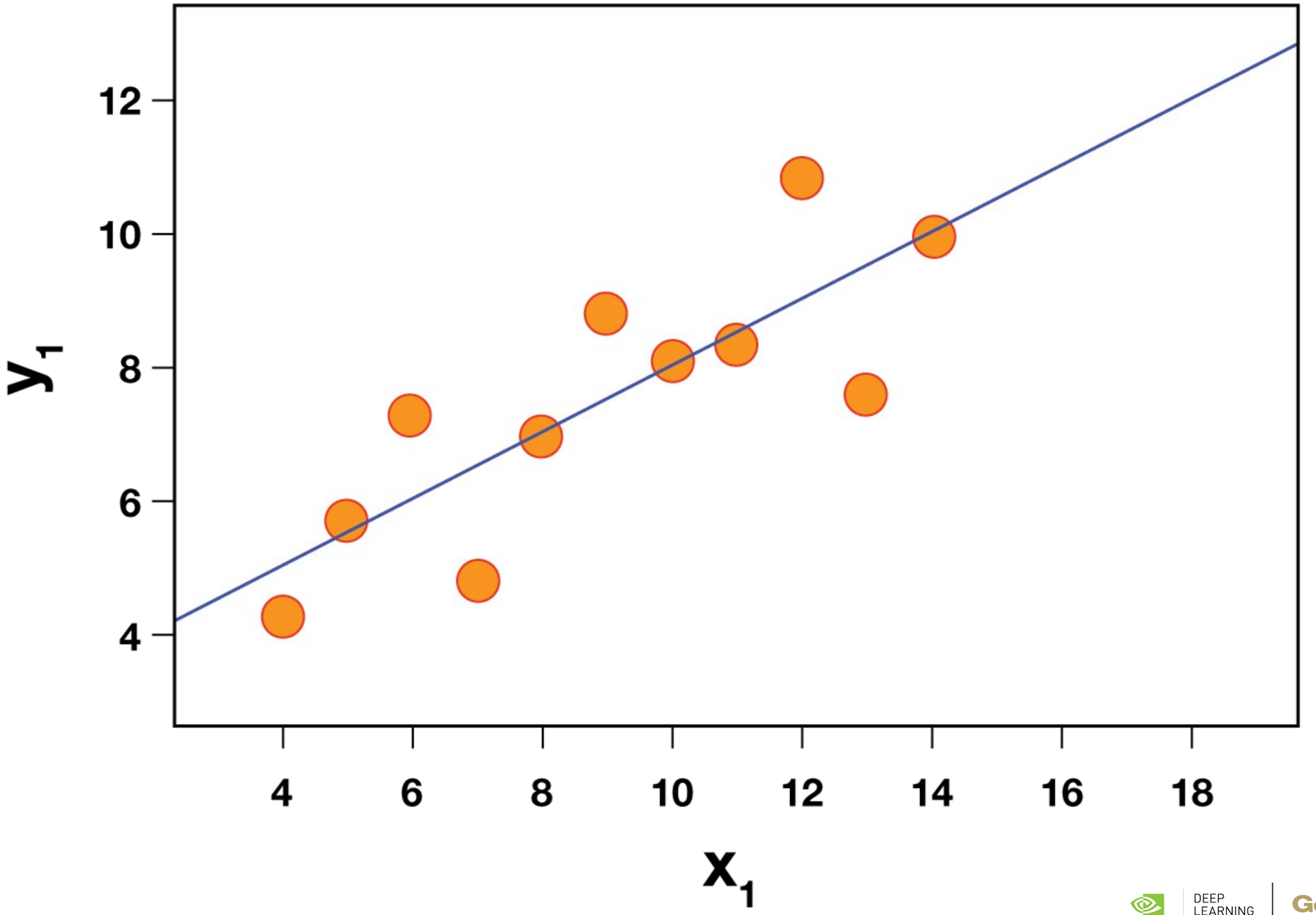
But why do you need to explore data at all???

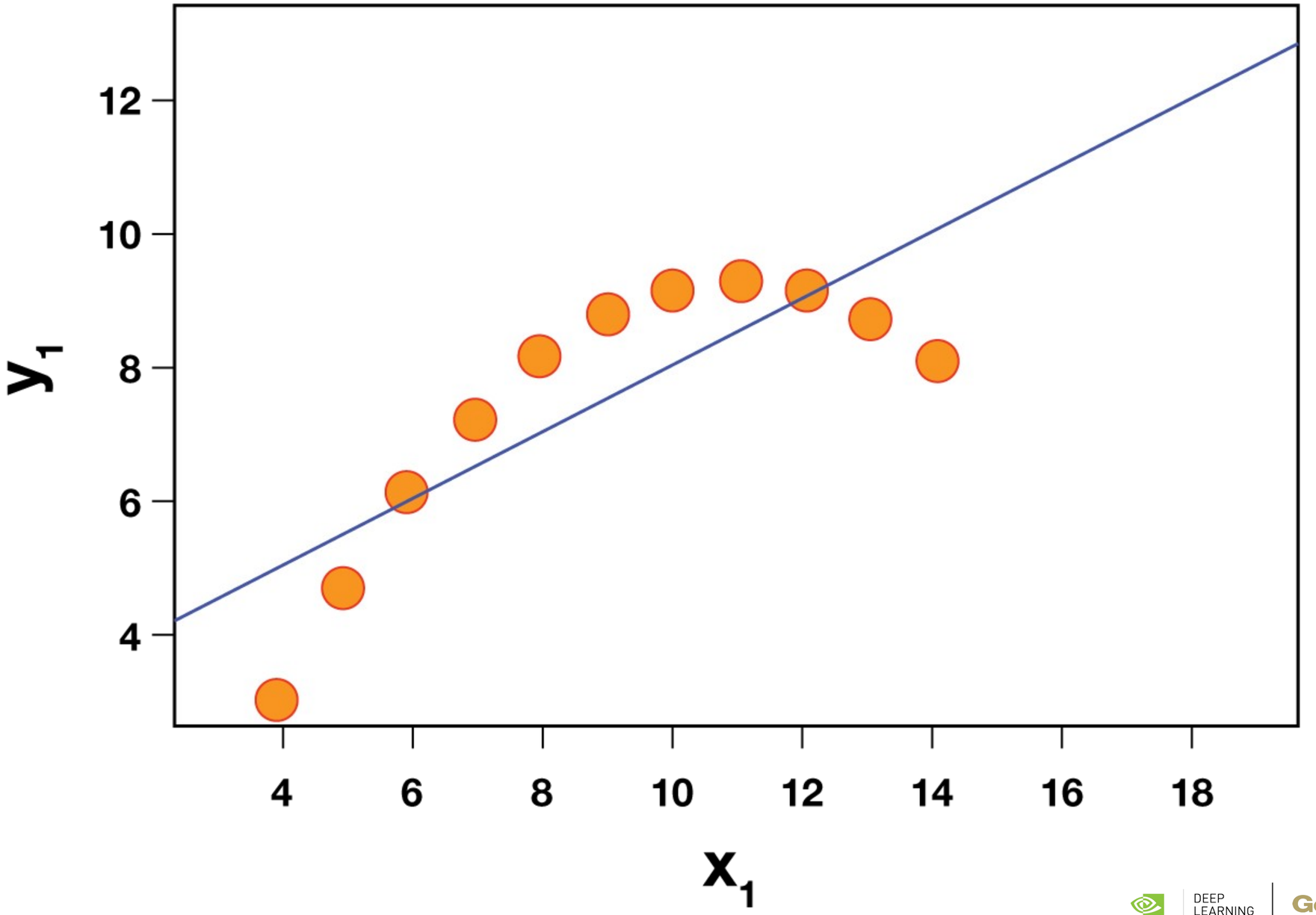
“There are three kinds of lies: lies, damned lies, and statistics.”

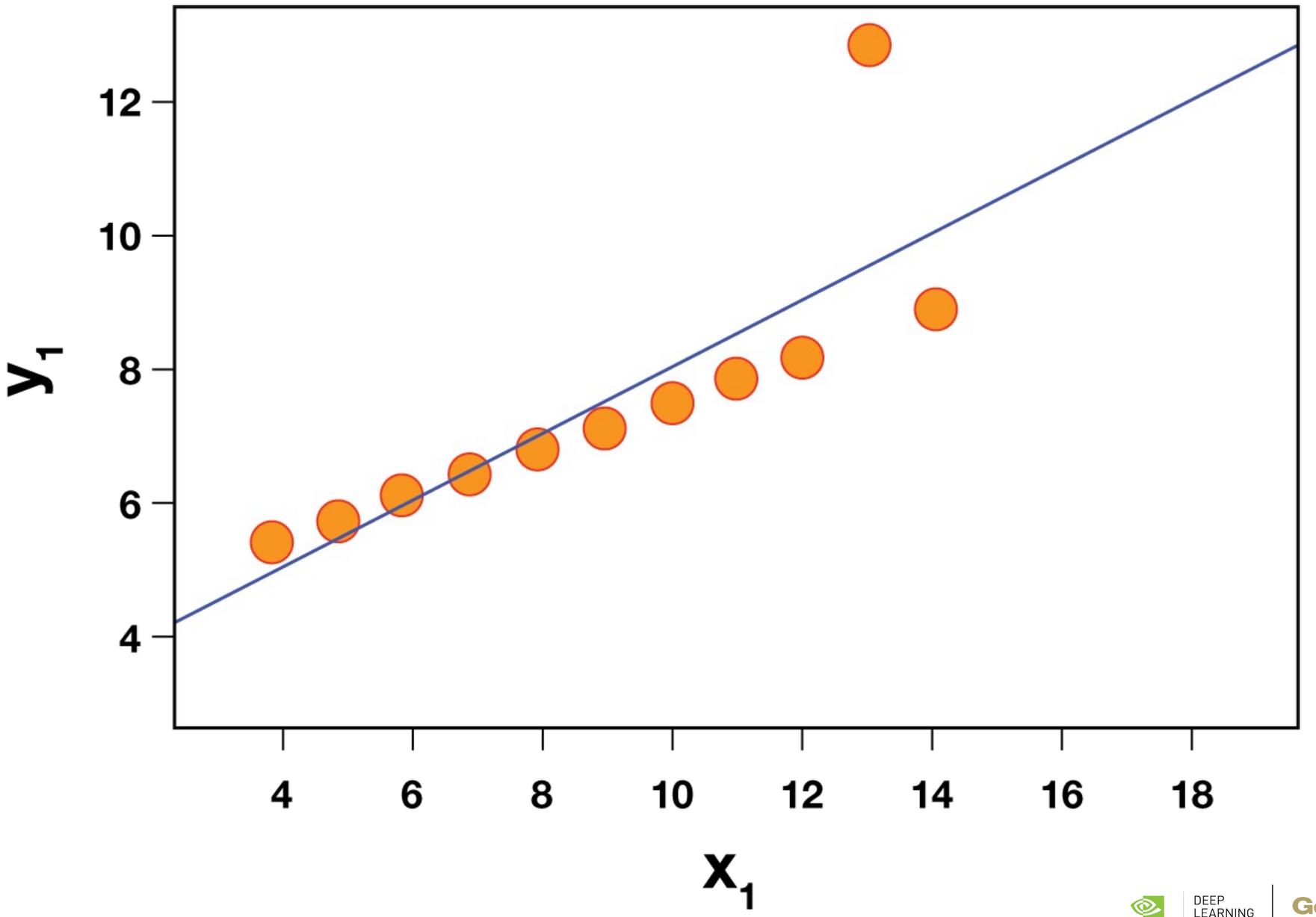
-Various Attributions

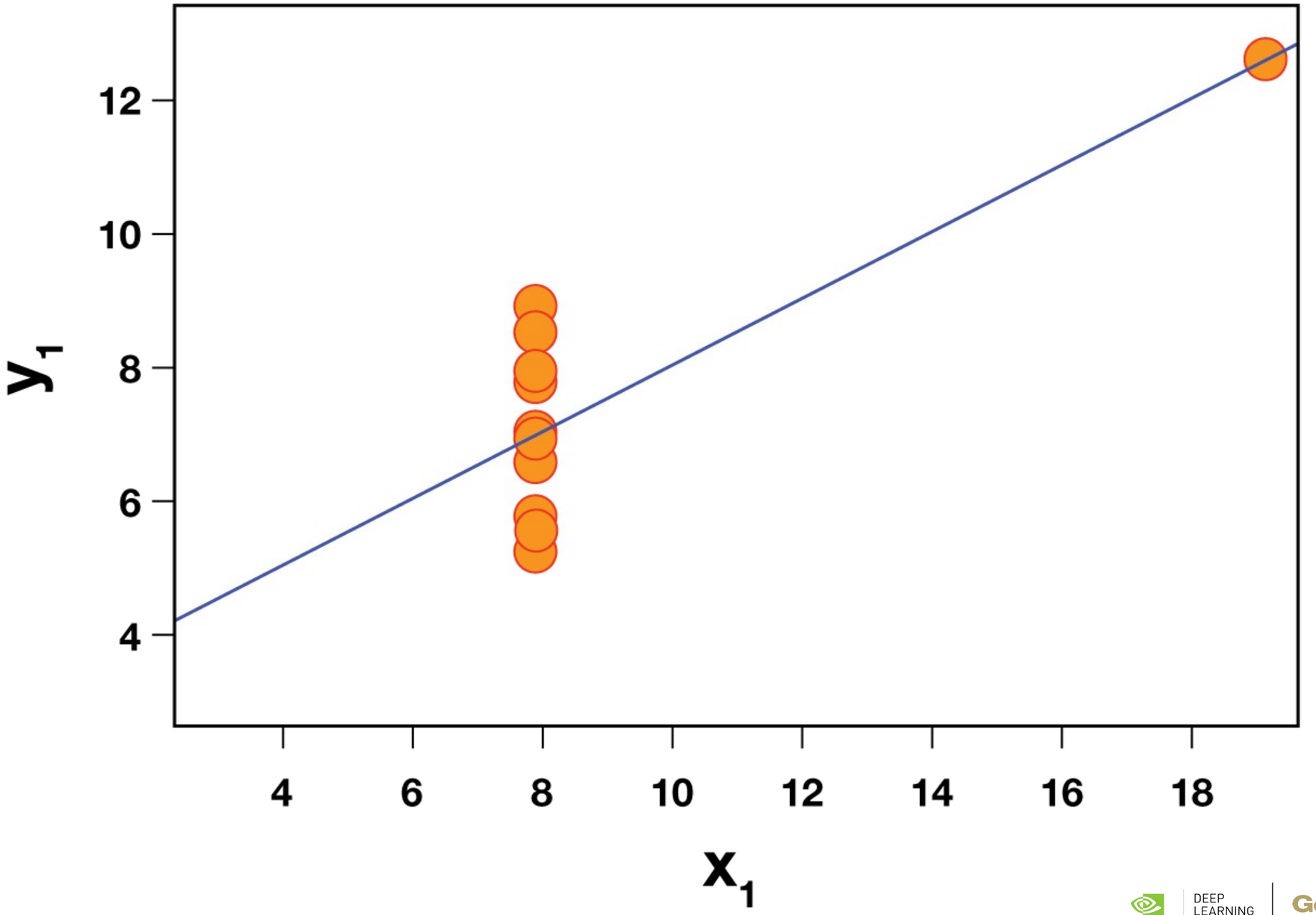
Mystery Data Set

Property	Value
mean(x)	9
variance (x)	11
mean(y)	7.5
variance (y)	4.122
correlation (x,y)	0.816
Linear Regression Line	y = 3 + 0.5x

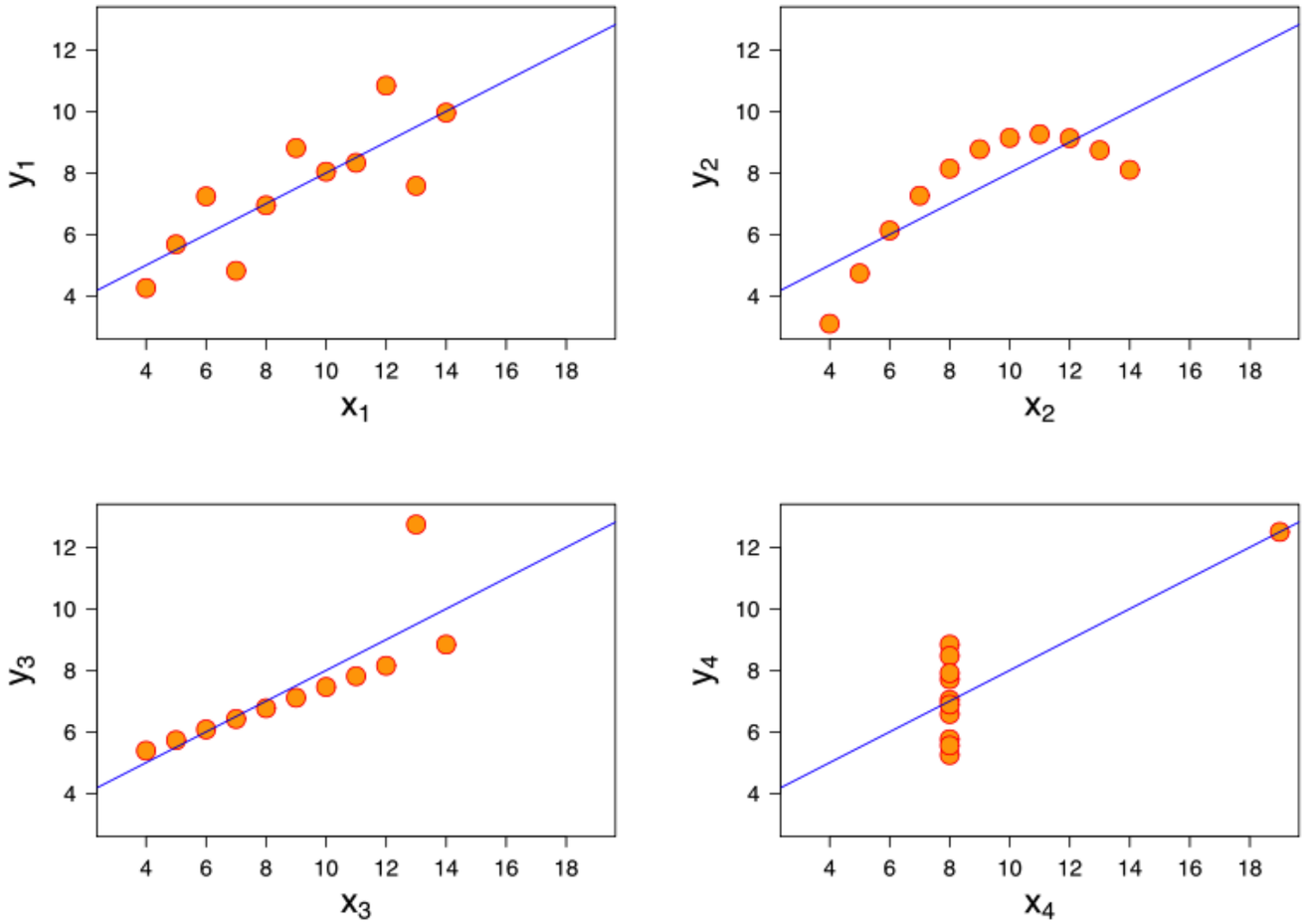






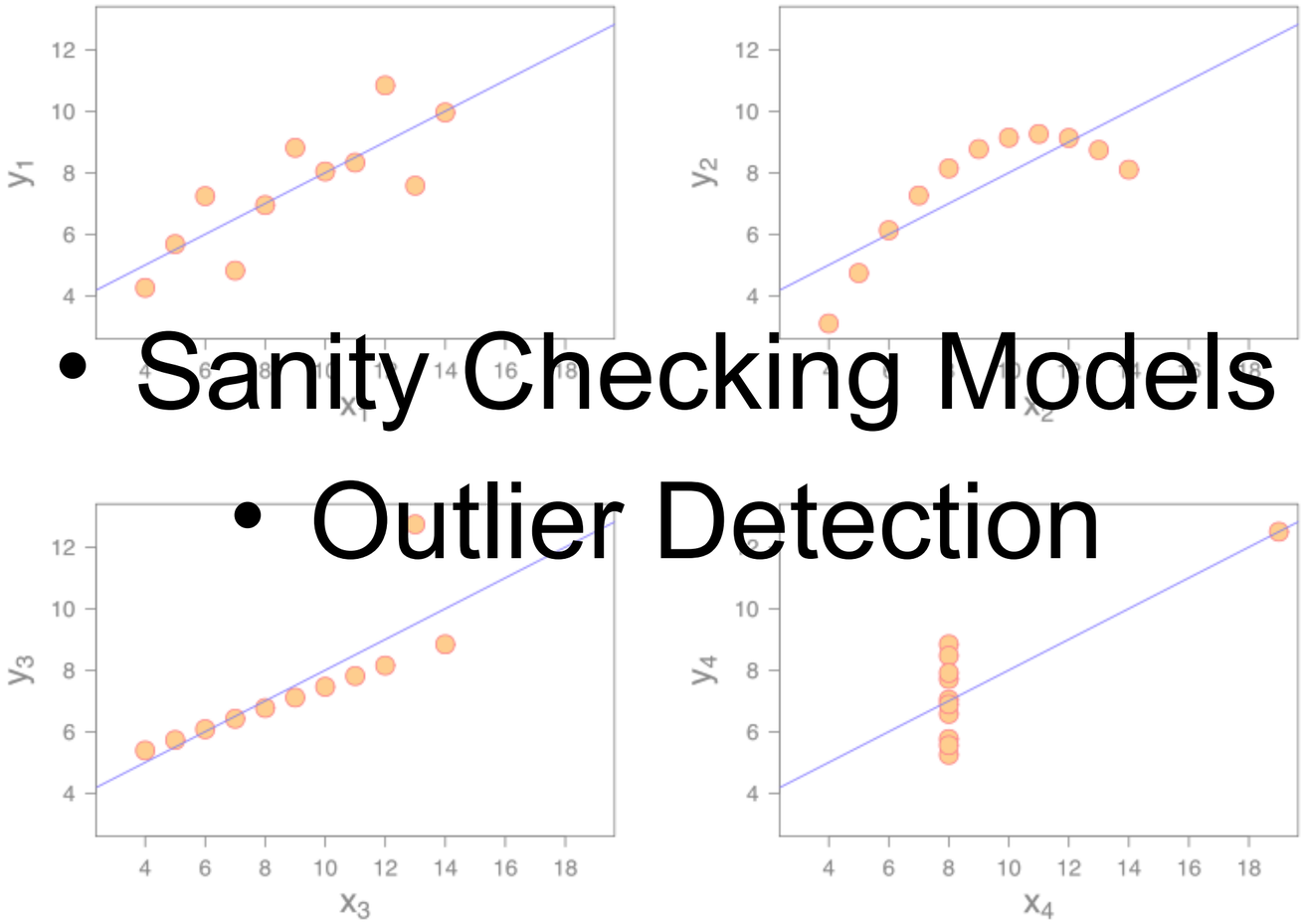


Anscombe’s Quartet



https://en.wikipedia.org/wiki/Anscombe%27s_quartet

Anscombe's Quartet



- Sanity Checking Models
  - Outlier Detection

DEEP
LEARNING
INSTITUTE

PRAIRIE VIEW
A&M UNIVERSITY

DLI Accelerated Data Science Teaching Kit

Thank You