

## **BIG DATA:**

What and how organizations are collecting leading indicators and the benefits they receive

---

Kent Szalla  
General Manager

# Plan for Today

- This session is a case study of many organizations
- Learning Objectives
  1. What metrics organizations are using to analyze leading indicator data they collect.
  2. What is data governance.
  3. How to leverage data governance for maximum analytical output and benefit.
- Brief view into the new world of ML and AI

# Level Set - Leading Indicators



Leading indicators are pre-incident measurements, as opposed to lagging indicators, which are measurements collected after an incident occurs.

For example, a flat tire is a lagging indicator because the blowout already has occurred, but an inspection that notes the poor quality of the tire and prevents a blowout from taking place is a leading indicator.

Kyle W. Morrison. "Get started. Select leading indicators to help measure safety." Health+Safety, Feb 1, 2014

# Common Leading Indicators

The leaves of deciduous trees, like maples and poplars, do often to turn upward before heavy rain. The leaves are actually reacting to the sudden increase in humidity that usually precedes a storm. Leaves with soft stems can become limp in response to abrupt changes in humidity, allowing the wind to flip them over.

Jamie McLeod. "Can Leaves Predict a Storm?" Farmer's Almanac | Weather, Jun 8, 2009

# Common Leading Indicators

- Inspections, audits, observations
- Safe and unsafe
- At risk and not at risk
- Severity of the observation
- Antecedent, precursors

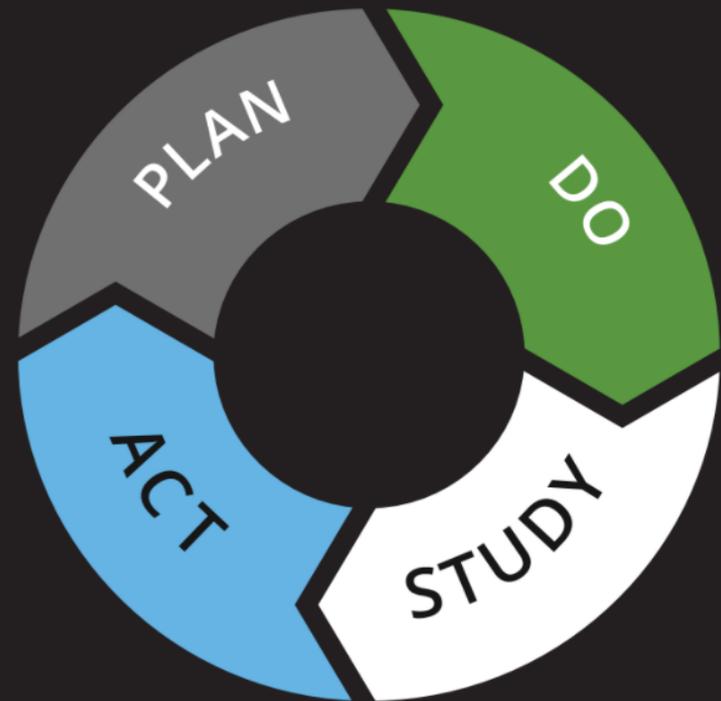
# How to collect the data?

- PLAN**
- Purpose
  - Expectations
  - Data Use Plan
  - Communication

- DO**
- Inspection Strategy
  - Observe
  - Initial Correction

- STUDY**
- Periodic Review
  - Identify Gaps & Trends
  - Measure Progress

- ACT**
- Accountability
  - Feedback
    - \*Positive & Negative
  - Develop Action Plans
  - Data-driven decisions



# How to collect the data?

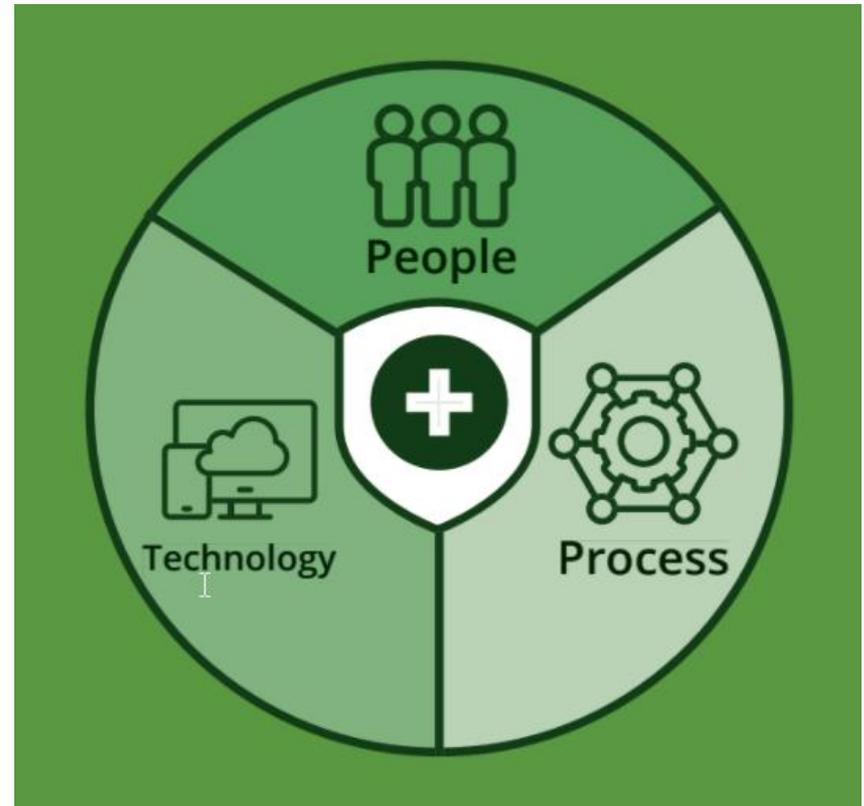
- Understand current state of your organization
- Set up your program to meet the need
  - Baby steps
  - Set goals and expectations
  - KISS (part of everyday process)
  - Electronic
- Measure (and coach)
  - People doing the collection
  - Measure culture
  - Risks observed
- Ensure the data collected is actionable
- Build in data governance



# Data Governance

Data governance is a process to ensure data meets precise standards and business rules as it is entered into a system. Data governance enables businesses to exert control over the management of data assets. This process encompasses the people, process, and technology that is required to ensure that data is fit for its intended purpose.

*Experian Information Solutions, Inc.*



# Data Governance

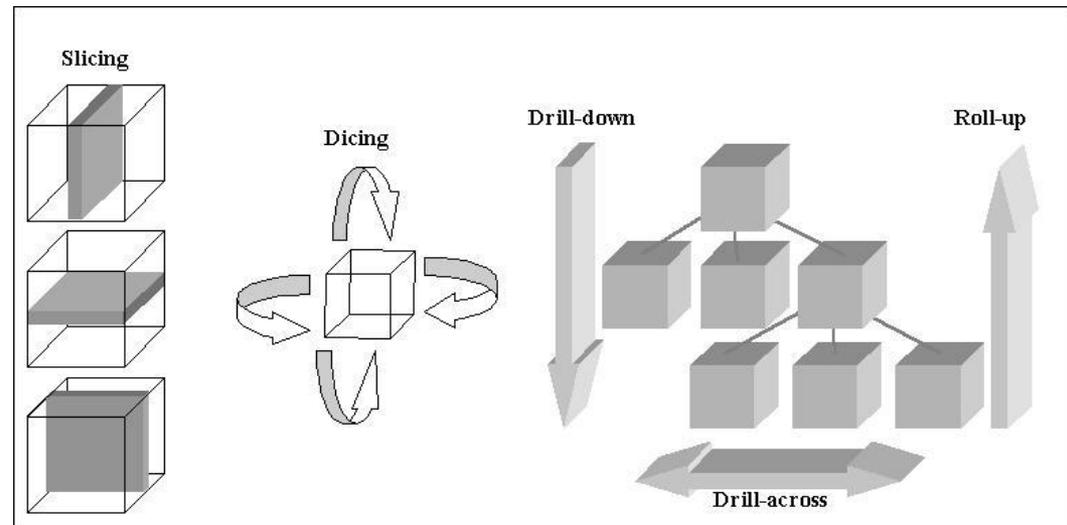


# Data Governance

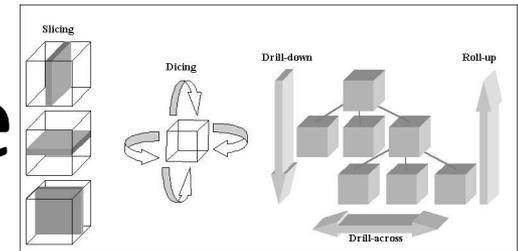
- Plan for your outputs
- Build simple processes for the outputs
- Measure and adjust
- Ensure one truth in the data

# Data Governance

- How do you want to see the data organized
  - By department
  - By region
  - By business unit
  - By observed party
  - By project
  - By work area
  - By floor
  - By....

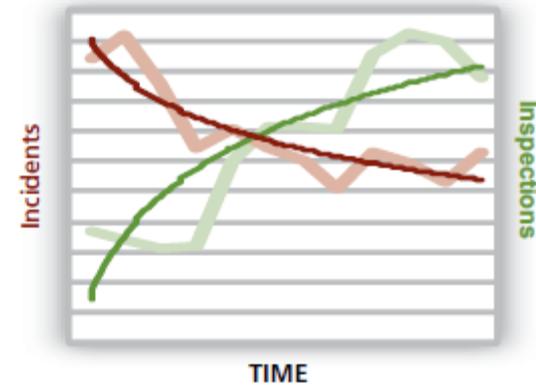
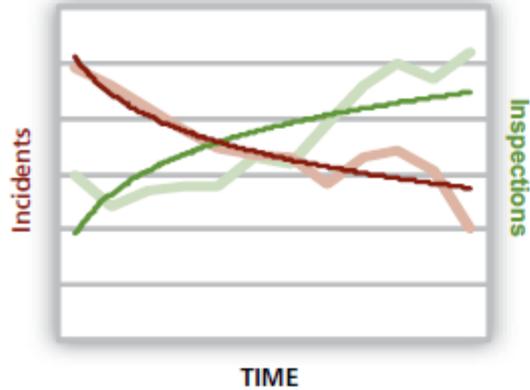
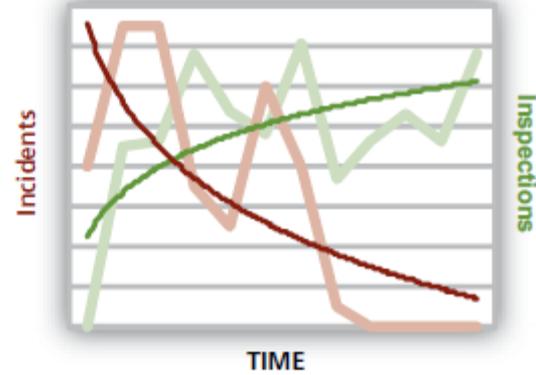
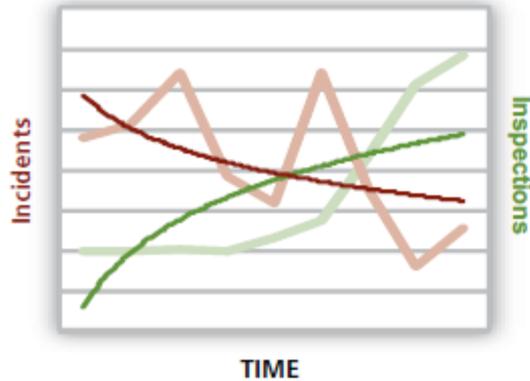


# Data Governance



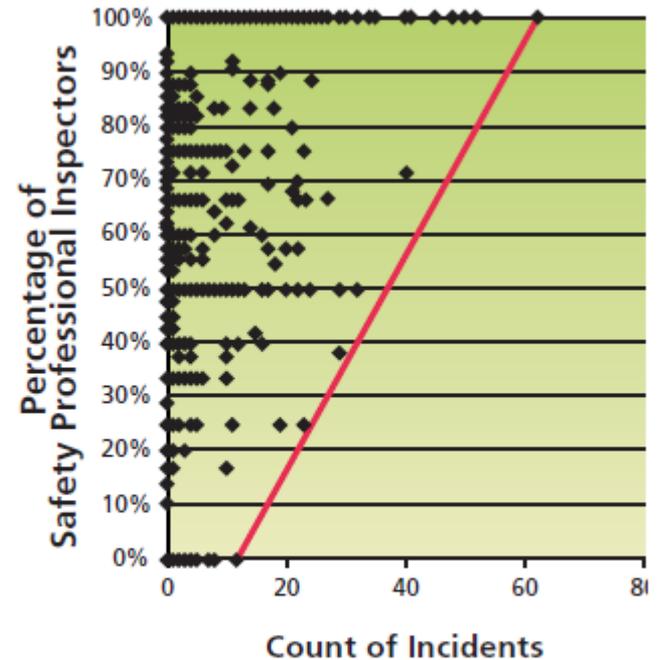
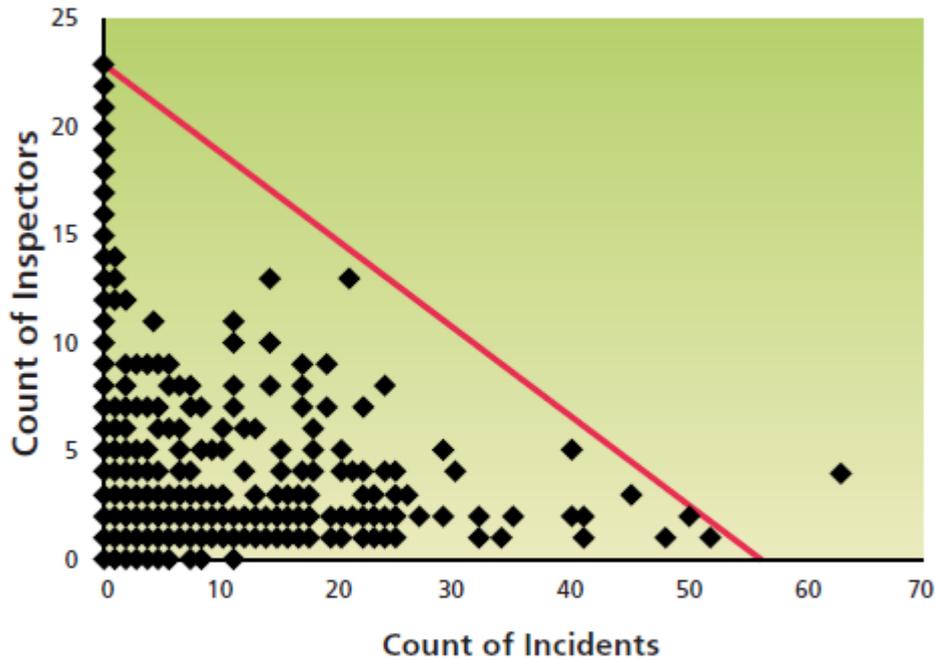
- How do you want to see the risks
  - By category
  - By subcategories
  - By sub-subcategories
- How do you want to measure and coach people doing the observations
  - Are they observing the right things?

# Considerations (the 4 Safety Truths)



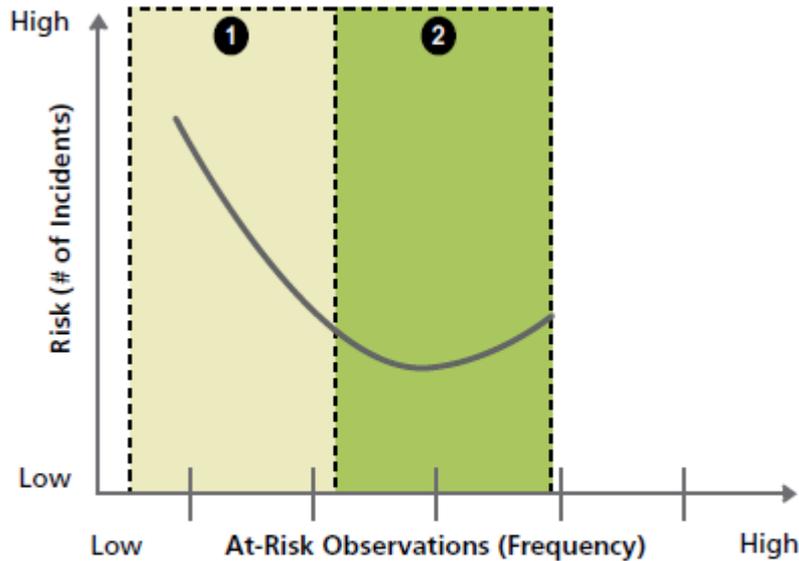
1. Incidents decrease as inspections increase

# Considerations (the 4 Safety Truths)

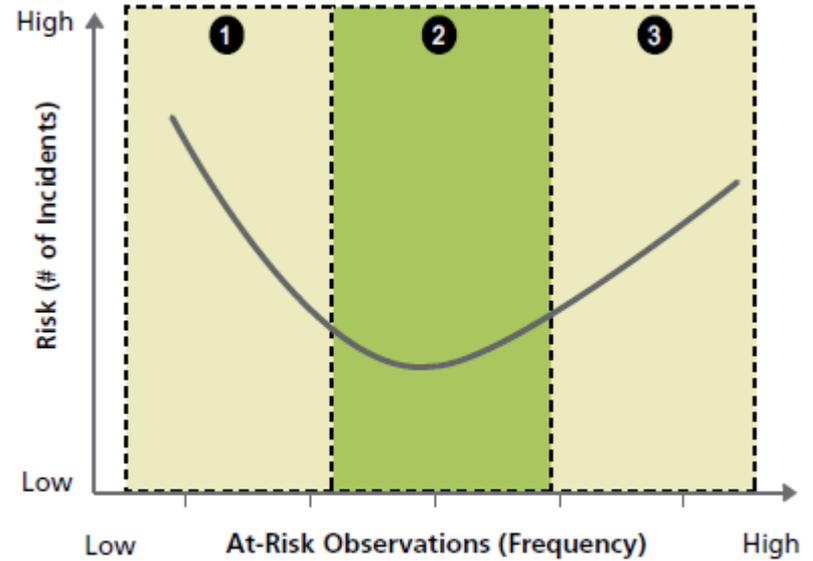


2. More inspectors, and more diverse inspectors, results in few incidents

# Considerations (the 4 Safety Truths)



3. Too many 100% safe inspections predicts unsafe worksite



4. Too many unsafe inspections predicts unsafe worksite

# Common Outputs / Metrics

Remember we discussed earlier...

- Measure (and coach)
  - People doing the collection
  - Measure culture
  - Risks observed

# Measuring

- People and Culture
  - Participation
  - Safe per Inspection
  - At-risk per inspection
  - Severity of observations
  - At-risks with comments
  - All safe inspections
  - What is being observed (easy stuff only?)
- Observed entities
  - Weighted % Safe (with severity)
  - Per inspection averages
  - Open issue closure rate
  - At risks with comments

Pictures are  
better...

# STATs

## Inspections

# Inspections

6,007

% All Safe  
**39.3%**

% All Unsafe 14.1%

Inspections / Inspector: **32.0**

# Inspectors

188

# Observations

292,549

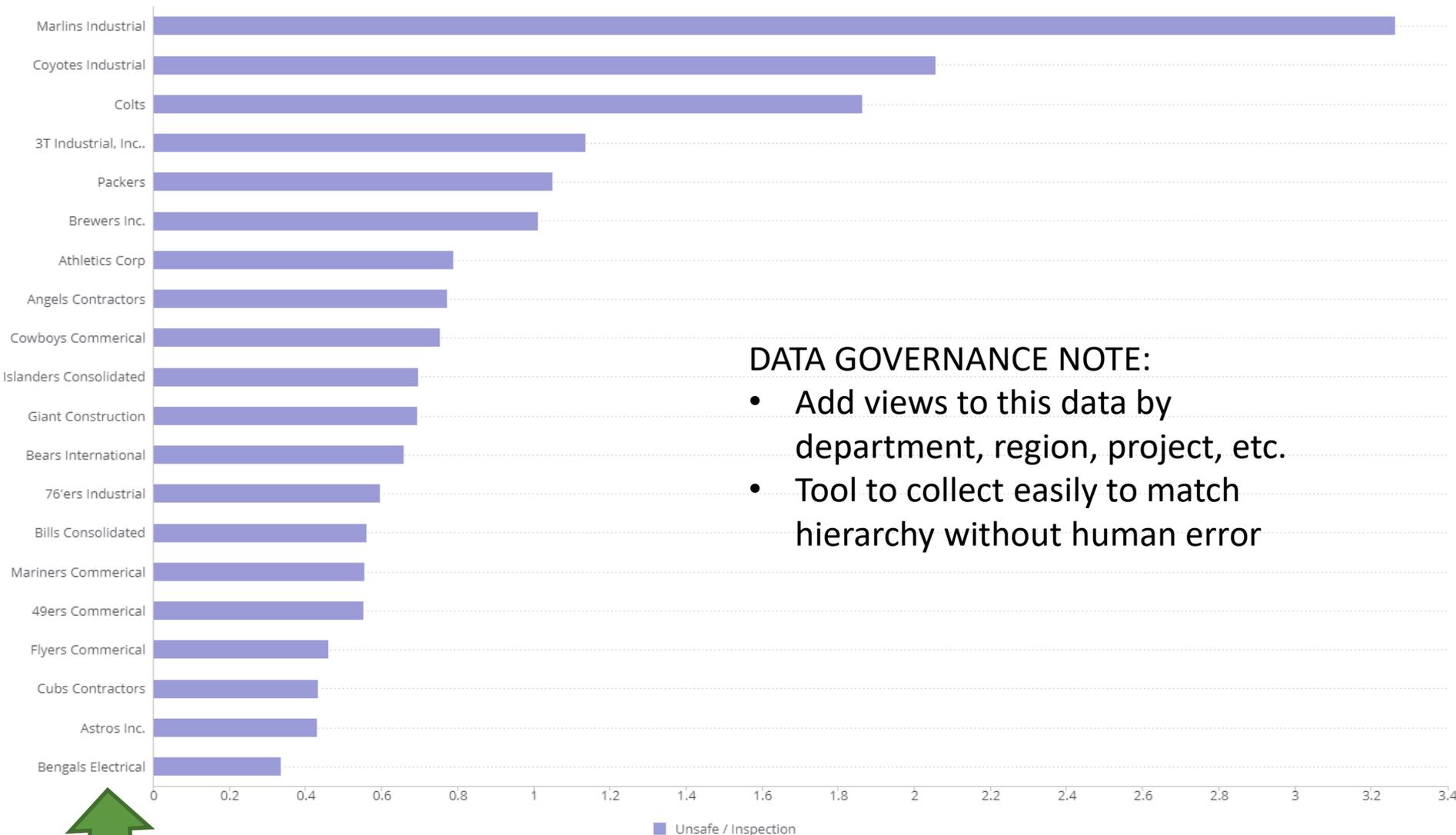
# Unsafe

10,298

Unsafe / Inspection: **1.7**

# Incidents

284

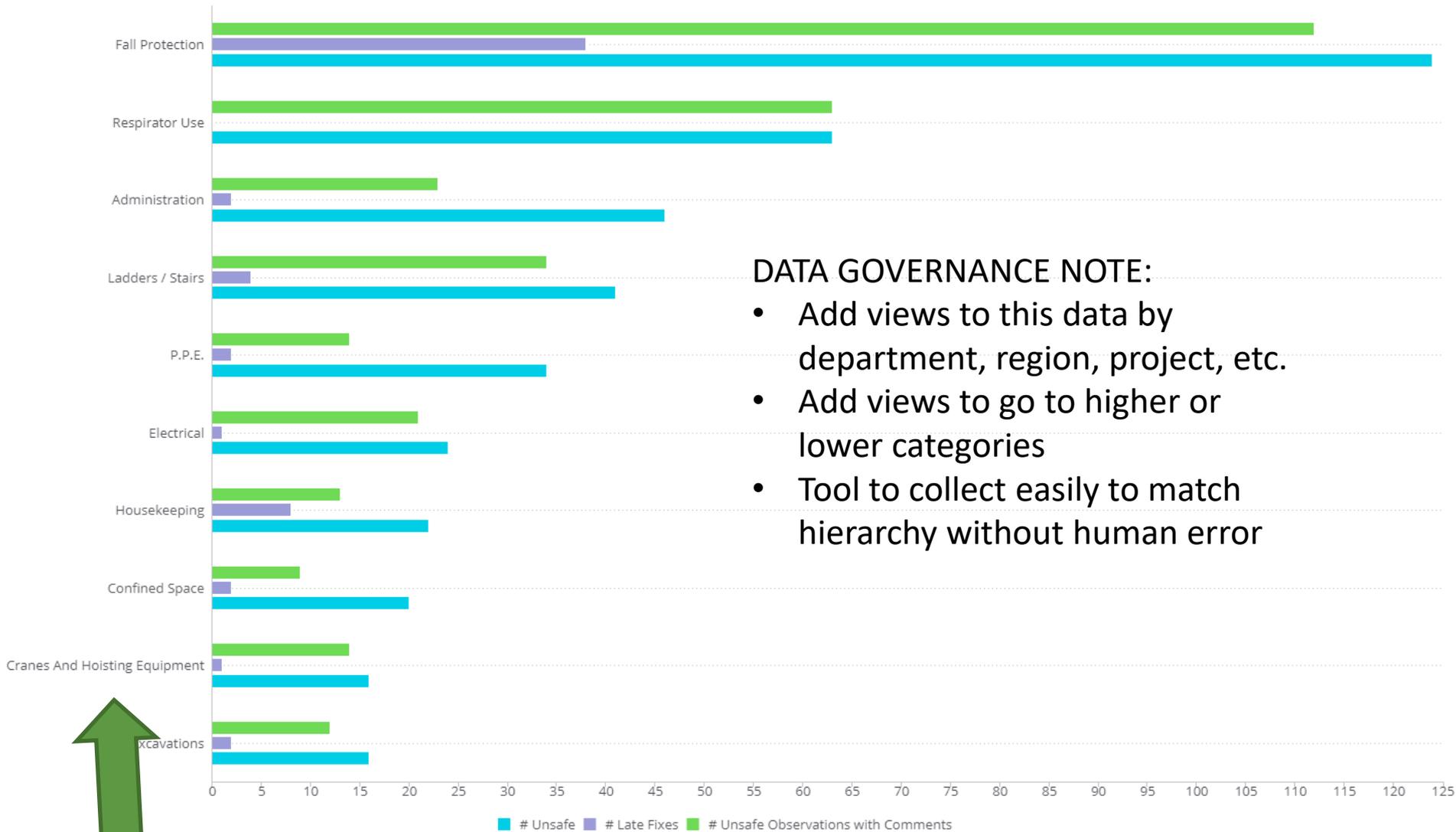


**DATA GOVERNANCE NOTE:**

- Add views to this data by department, region, project, etc.
- Tool to collect easily to match hierarchy without human error



**OBSERVED  
ENTITY**

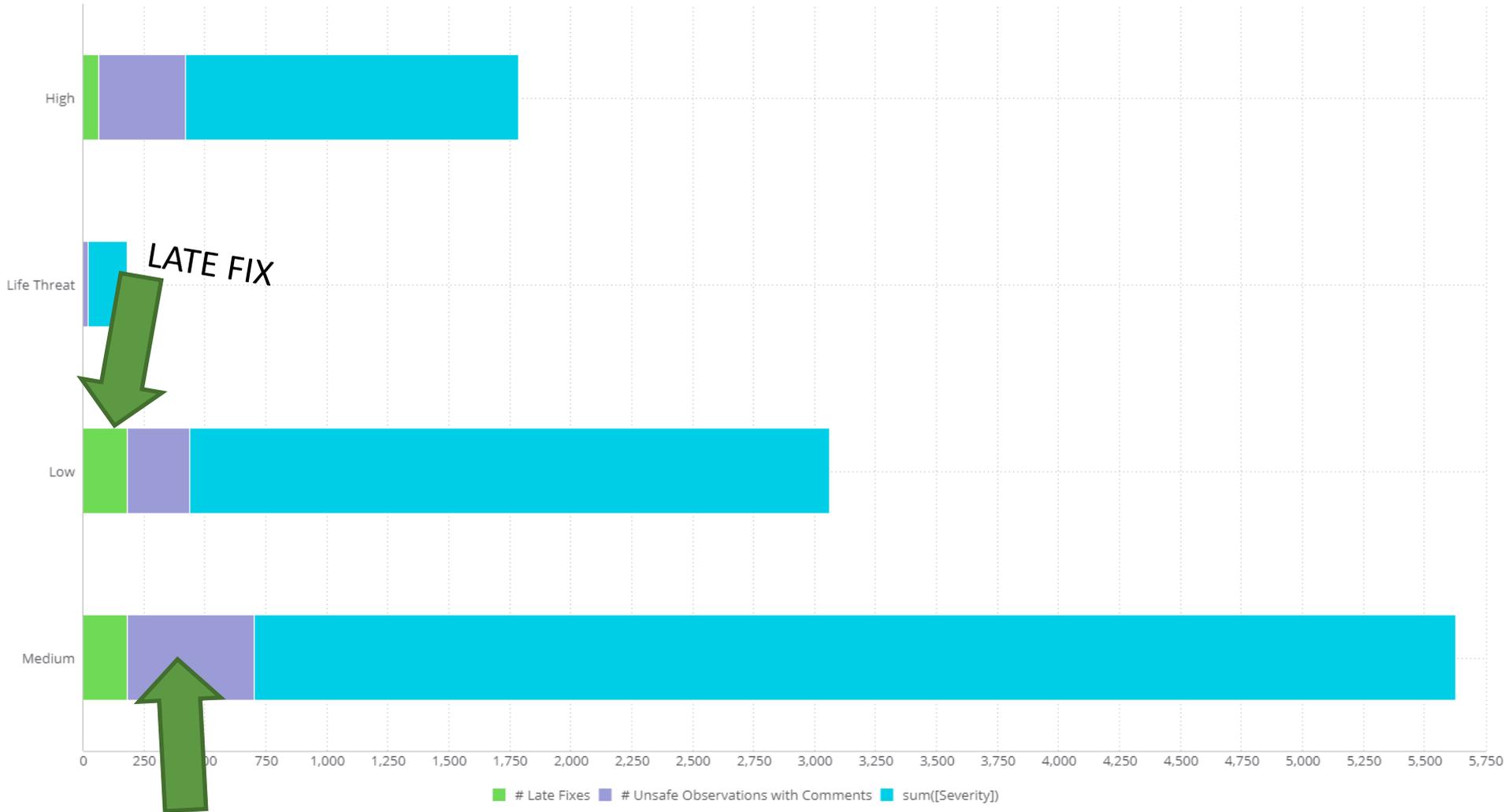


**DATA GOVERNANCE NOTE:**

- Add views to this data by department, region, project, etc.
- Add views to go to higher or lower categories
- Tool to collect easily to match hierarchy without human error

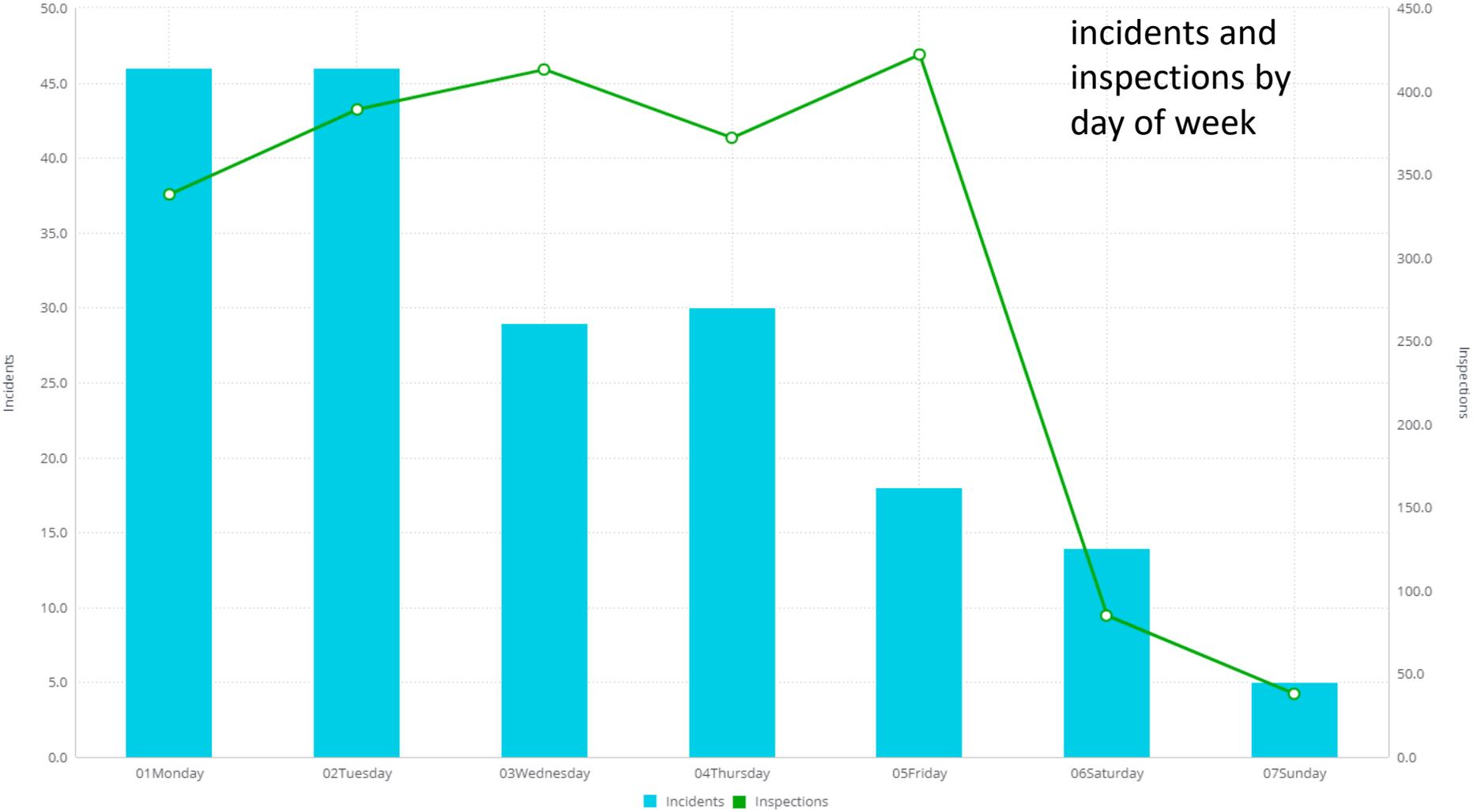


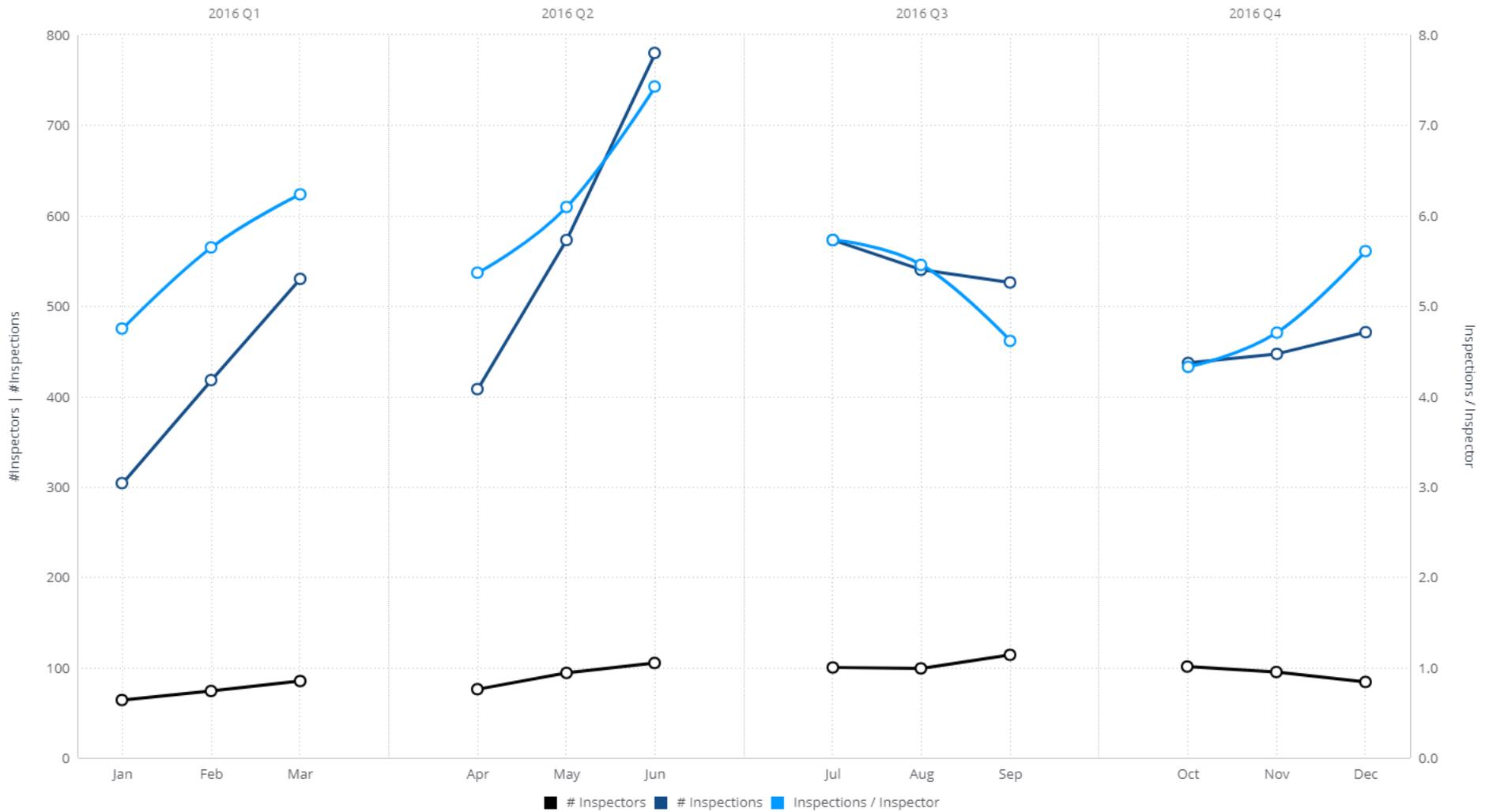
**CATEGORIES  
OBSERVED**

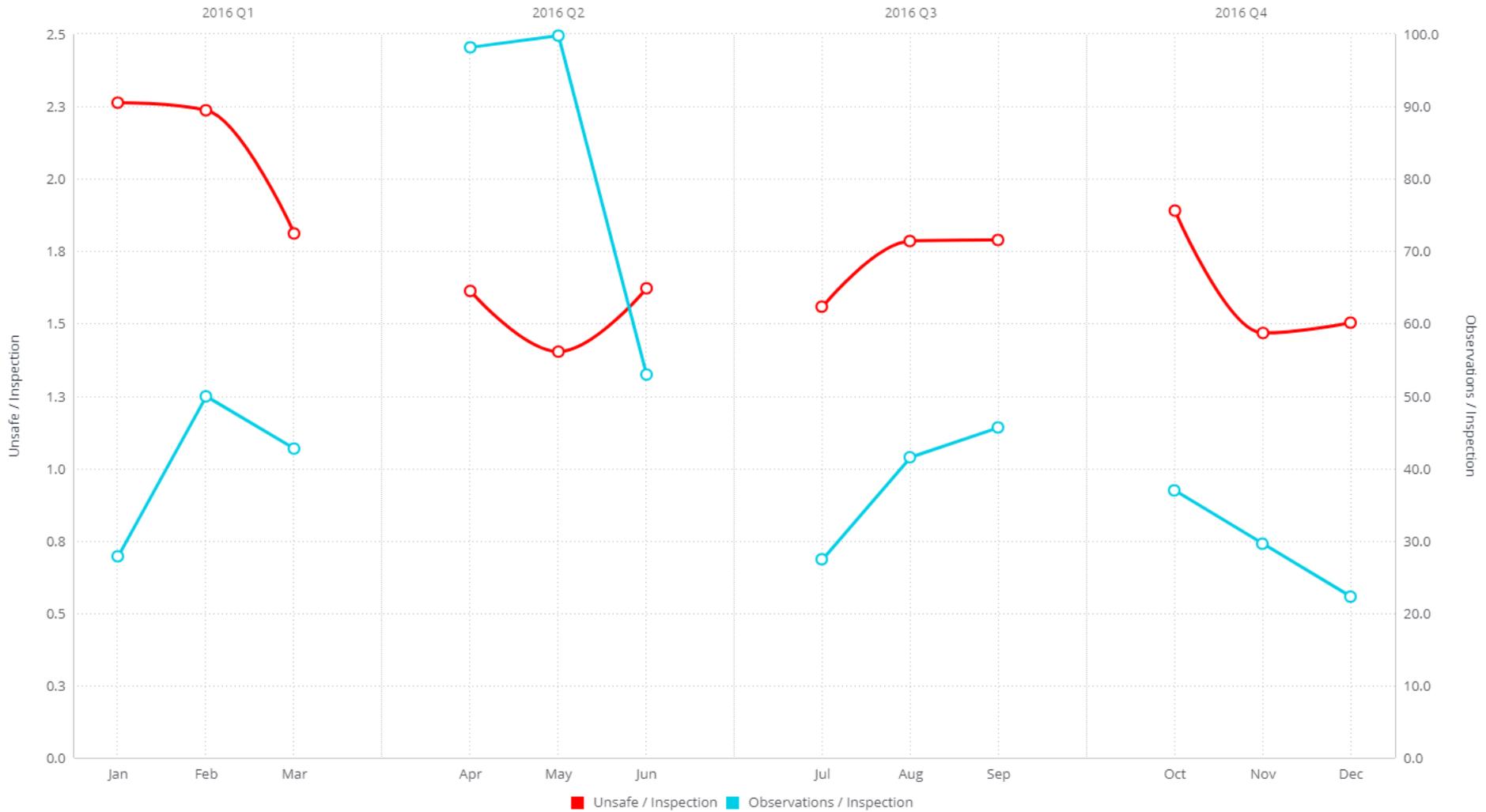


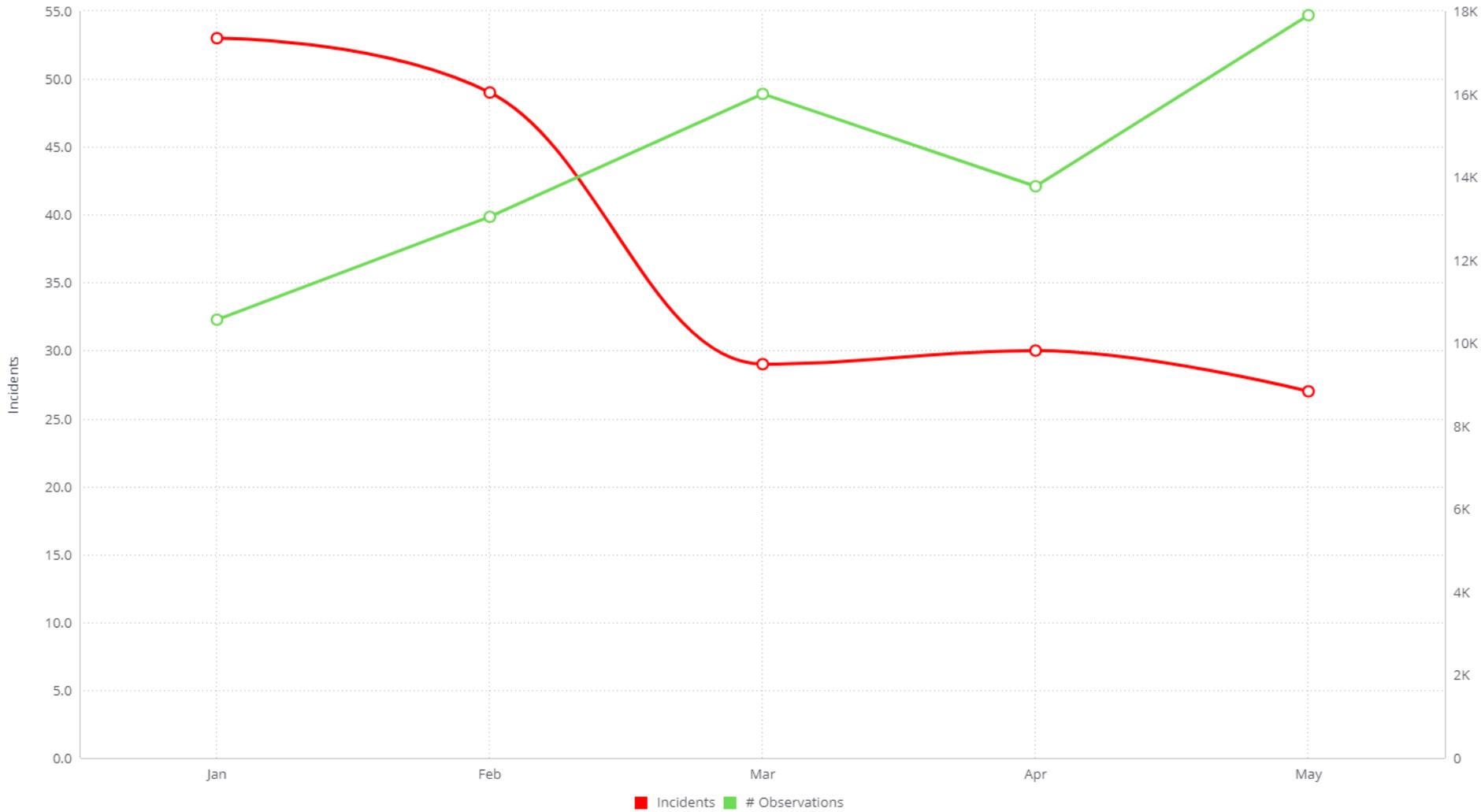
COMMENTS

# Correlation between incidents and inspections by day of week





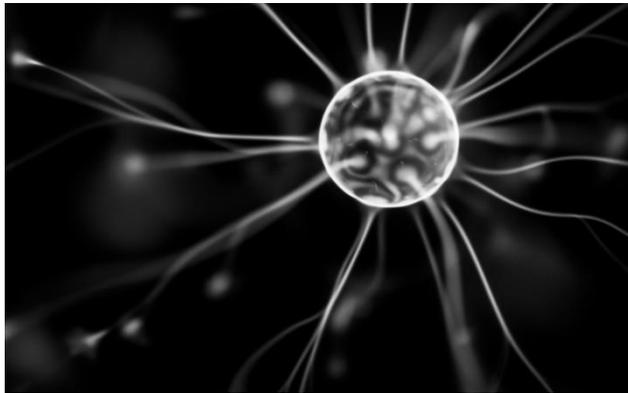




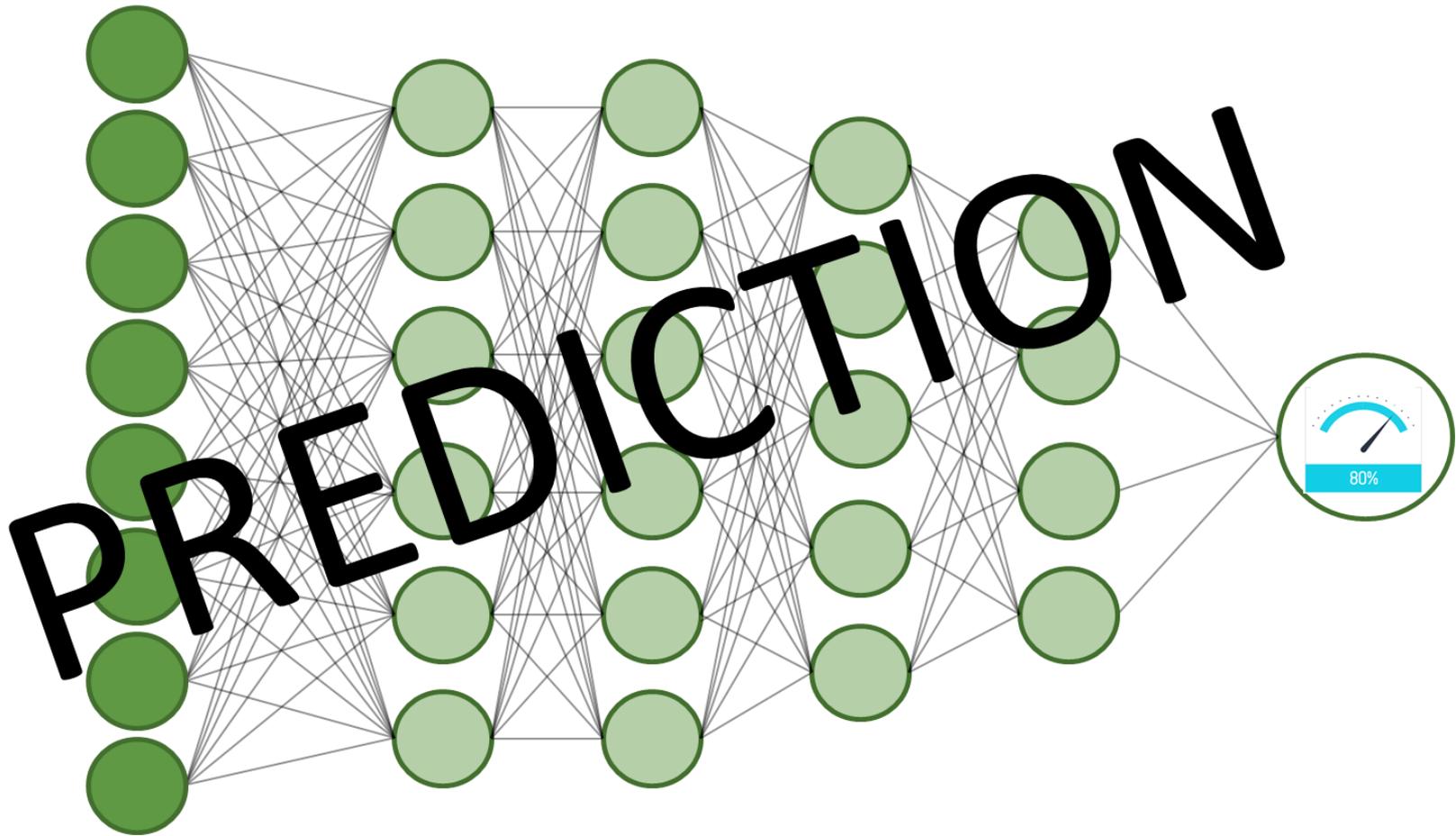
Month	Observations / Inspection	Safe / Inspection	Unsafe / Inspection	# Inspections	# Observations	# Safe	# Unsafe	% Unsafe Resolved Immediately	# Observers	# Incidents	% All Safe	# First Marked Not Corrected	# Comments
Jan-17	27.93	27.00	0.92	109116	3047072	2946442	100630	88.34%	16242	854	57.07%	11729	135015
Feb-17	28.21	27.30	0.91	108282	3054647	2956062	98585	88.98%	16400	754	56.40%	10864	129691
Mar-17	29.40	28.48	0.92	123609	3633726	3520149	113577	88.27%	17278	824	57.10%	13322	150184
Apr-17	29.96	29.08	0.88	111882	3351562	3253344	98218	88.06%	16889	741	57.80%	11727	137784
May-17	30.74	29.52	1.23	120936	3717923	3569509	148414	91.58%	17720	883	58.75%	12490	148059
Jun-17	30.14	29.25	0.89	121698	3667443	3559067	108376	88.18%	17723	730	58.14%	12815	151988
Jul-17	29.47	28.58	0.88	112824	3324538	3224699	99839	88.18%	17059	631	58.68%	11798	139402
Aug-17	28.91	28.02	0.89	126190	3648535	3535778	112757	88.48%	17223	661	58.43%	12986	156394

# Humans are human

## Computers are powerful



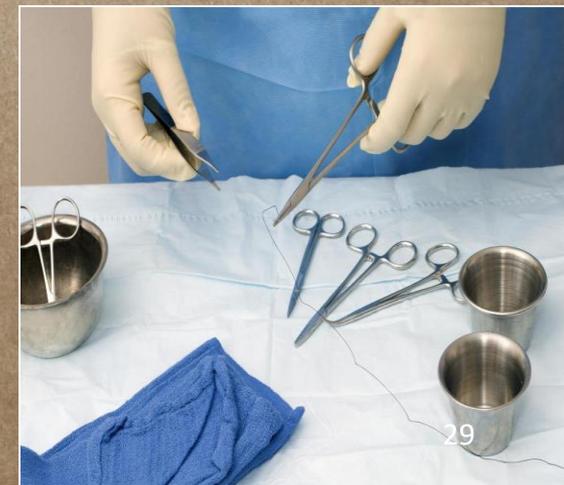
Like microscopes that see viruses undetectable by the human eye, computers can find patterns that humans are incapable of seeing



# Let's start with a true story...



SUN	MON	TUES	WEDS	THUR	FRI	SAT
—	—	—	—	—	01	02
03	04	05	06	07	08	09
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	—	—



# The state of Artificial Intelligence



$$AI = ML + TD + DK$$

AI = Artificial Intelligence

ML = Machine Learning

TD = Test Data (aka BIG DATA)

DK = Domain Knowledge  
(subject matter expertise)

# The state of Artificial Intelligence

- AI is everywhere now
  - Google
  - Facebook
  - Tesla
- Tons of tools
  - Open source like TensorFlow
  - Paid like IBM Watson
- Data scientist profession on the rise

“Machine learning...is the next transformation...[it] will be the basis and fundamentals of every successful huge IPO win in 5 years.”  
– Eric Schmidt, Google Executive Chairman

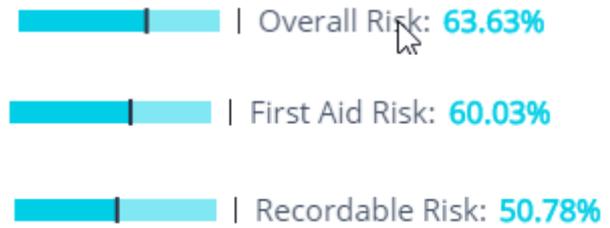
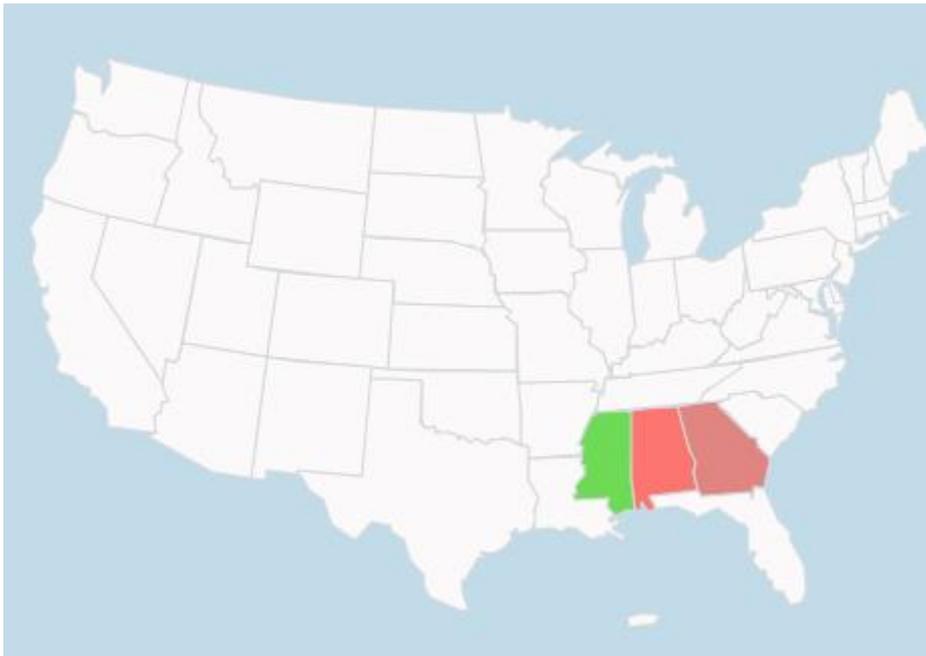


# The state of Artificial Intelligence

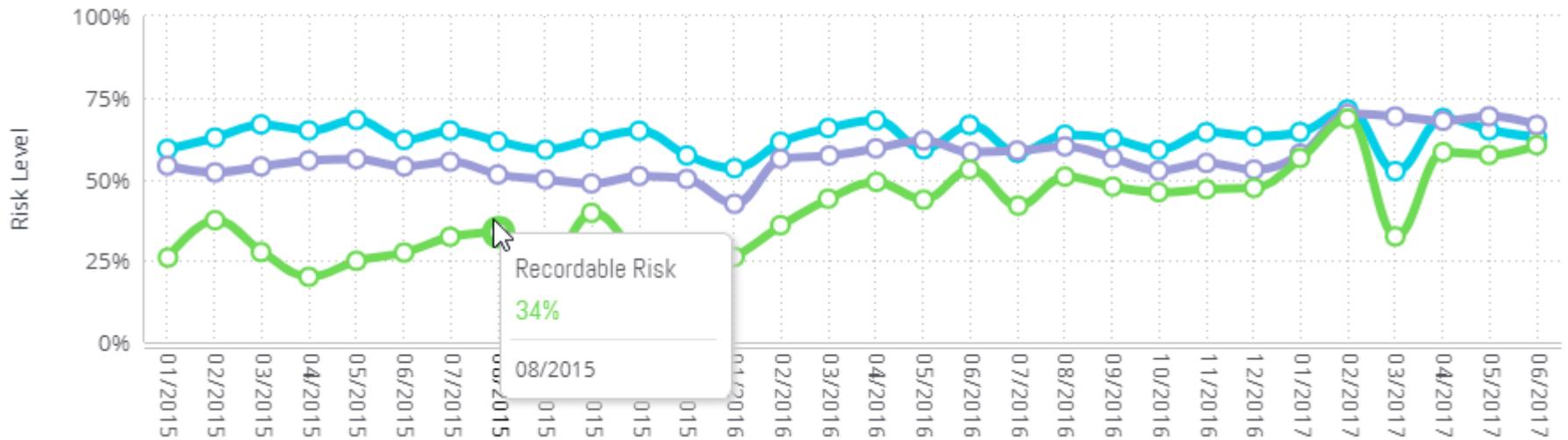
- AI is everywhere now
  - Google
  - Facebook
  - Tesla
- Tons of tools
  - Open source like TensorFlow
  - Paid like IBM Watson
- Data scientist profession on the rise

“Machine learning...is the next transformation...[it] will be the basis and fundamentals of every successful huge IPO win in 5 years.”  
– Eric Schmidt, Google Executive Chairman

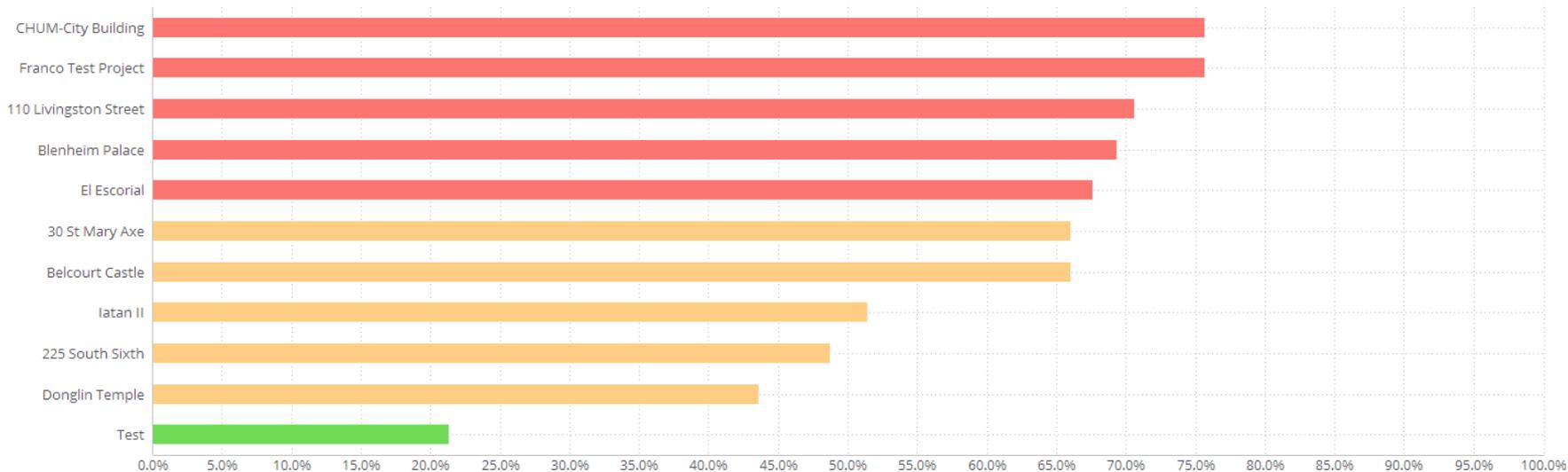




Overall Risk First Aid Risk Recordable Risk



Project Risk



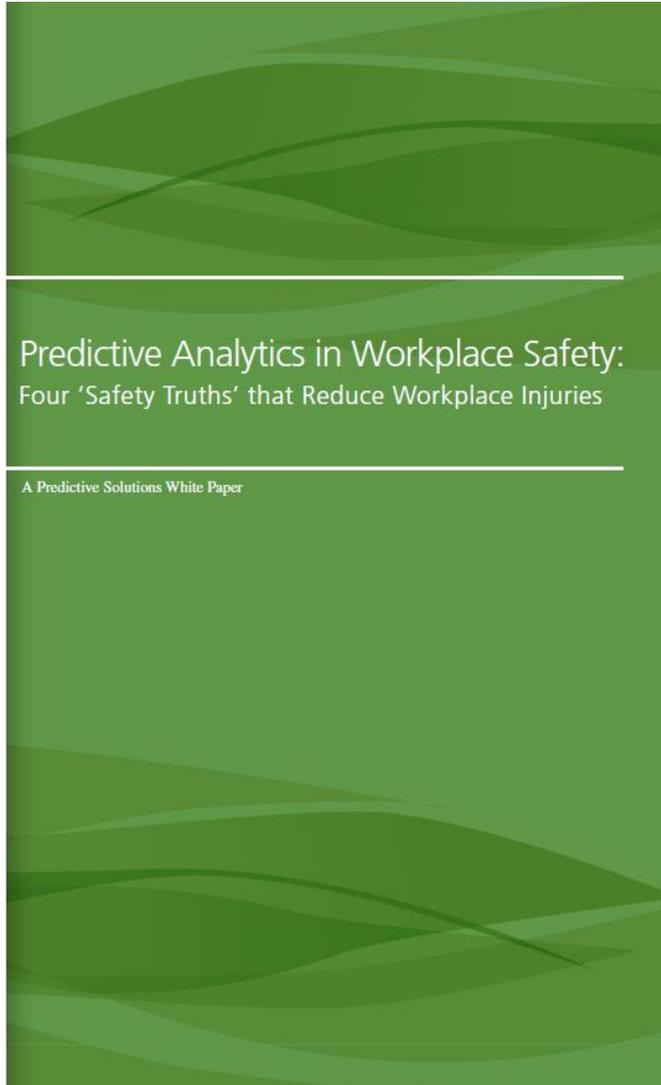
Risk Severity

Project	First Aid Risk Level	Recordable Risk Level
110 Livingston Street	66.4%	67.4%
225 South Sixth	66.3%	48.9%
30 St Mary Axe	68.2%	48.9%
Belcourt Castle	50.7%	48.9%
Blenheim Palace	68.2%	54.4%
CHUM-City Building	68.2%	66.5%
Donglin Temple	33.0%	46.4%
El Escorial	47.3%	59.2%
Franco Test Project	66.4%	69.9%
Iatan II	66.2%	6.3%
Test	35.3%	1.9%

Risk Category

Category	Type	Risk Level
Body Part	Arm	63.8%
	Back	53.6%
	Eye	46.7%
	Lower	35.7%
Mechanism	Caught In	44.5%
	Contact	58.5%
	Exposure	28.9%
	Slip/Trip	49.5%
	Strain	50.7%
	Struck	39.8%
Nature	Abrasion	41.6%
	Contusion	53.5%
	Foreign Object	52.8%
	Laceration	55.9%
	Strain	51.5%

# Questions?



[kszalla@predictivesolutions.com](mailto:kszalla@predictivesolutions.com)

412-809-1888 x1841

[www.predictivesolutions.com](http://www.predictivesolutions.com)

# Thank You