



# Process Mining Metrics

Anne Rozinat, Fluxicon



A young boy with dark, wavy hair is looking out of a window. His face is pressed against the glass, and his hands are resting on the windowsill. The window is covered in raindrops, and the view outside is blurred, showing green foliage and a dark, overcast sky. The word "Boring?" is written in large, white, bold letters across the center of the image.

**Boring?**









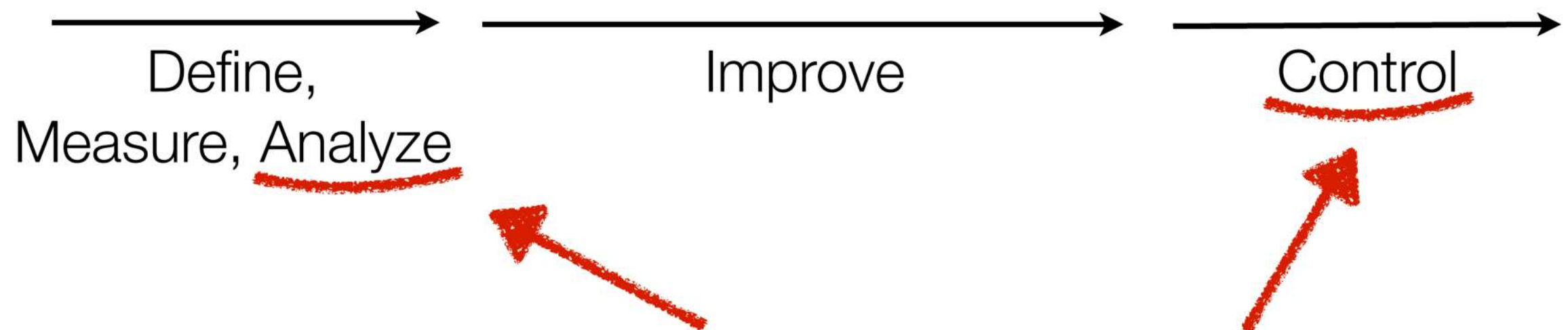
**So, Why Do We  
Need Metrics?**



**metric:** [me-trik]

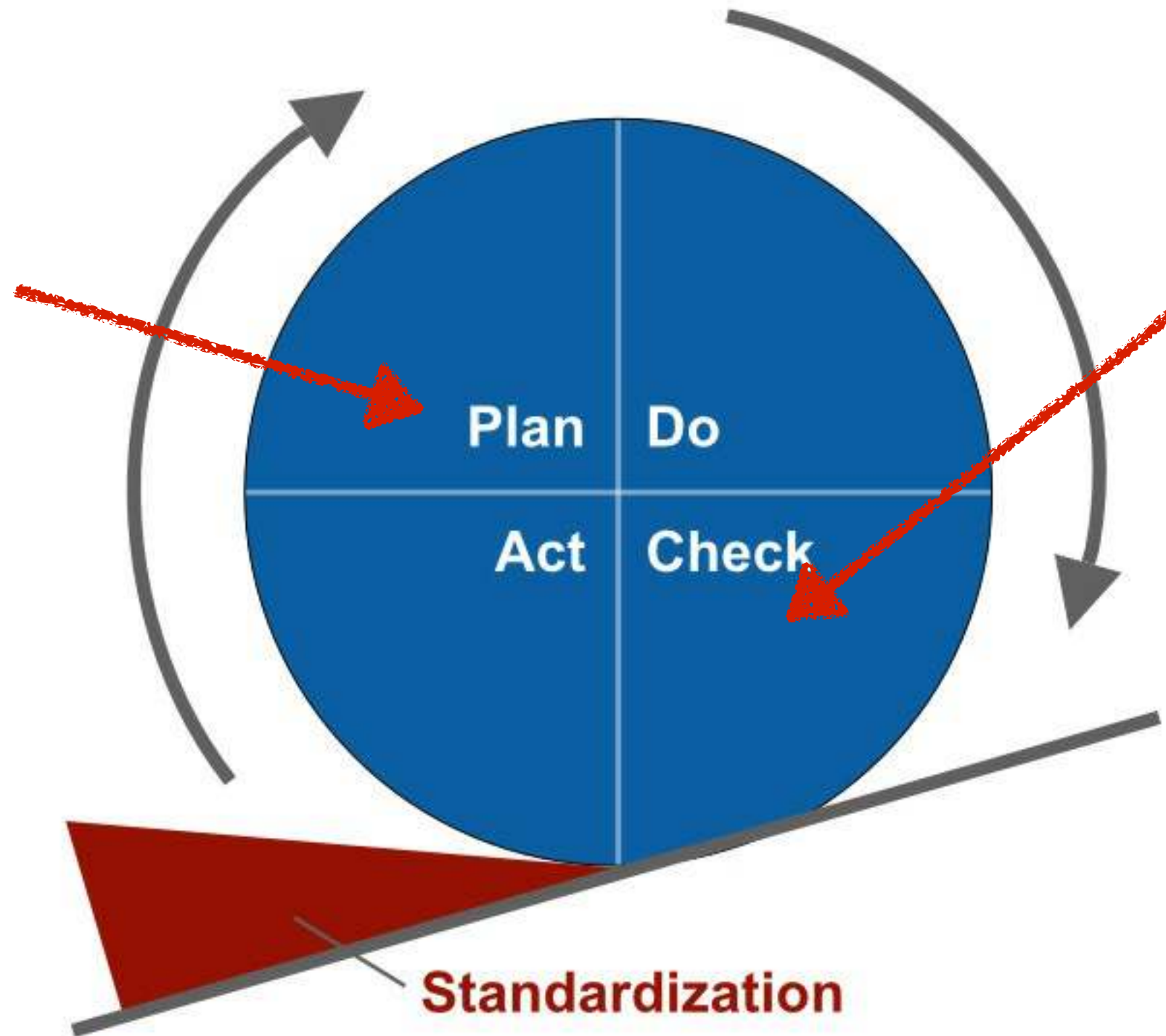
“A standard for measuring or evaluating something”







# Continuous Improvement





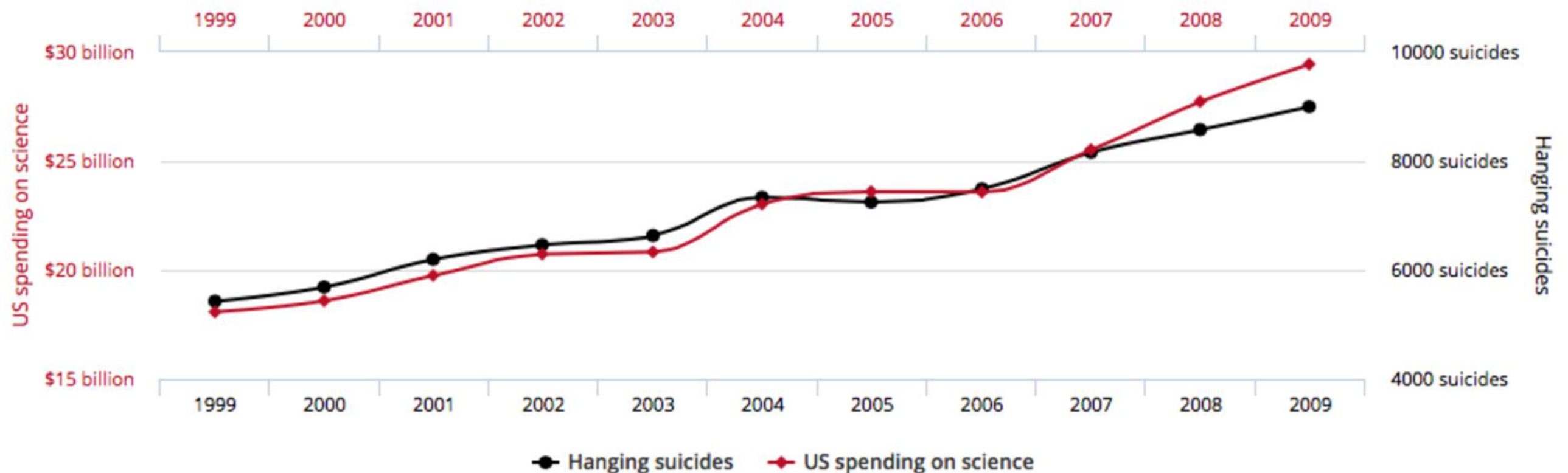
# Problem 1: Deception





# US spending on science, space, and technology correlates with Suicides by hanging, strangulation and suffocation

Correlation: 99.79% ( $r=0.99789126$ )



Data sources: U.S. Office of Management and Budget and Centers for Disease Control & Prevention

tylervigen.com

Spurious Correlations

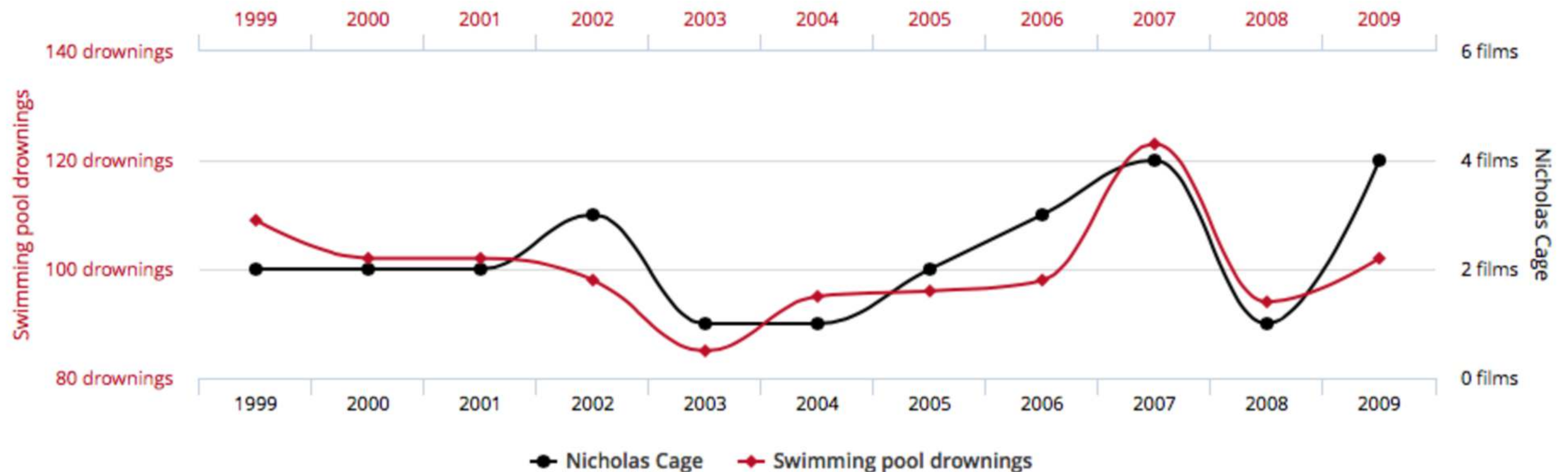


# Number of people who drowned by falling into a pool

correlates with

## Films Nicolas Cage appeared in

Correlation: 66.6% ( $r=0.666004$ )



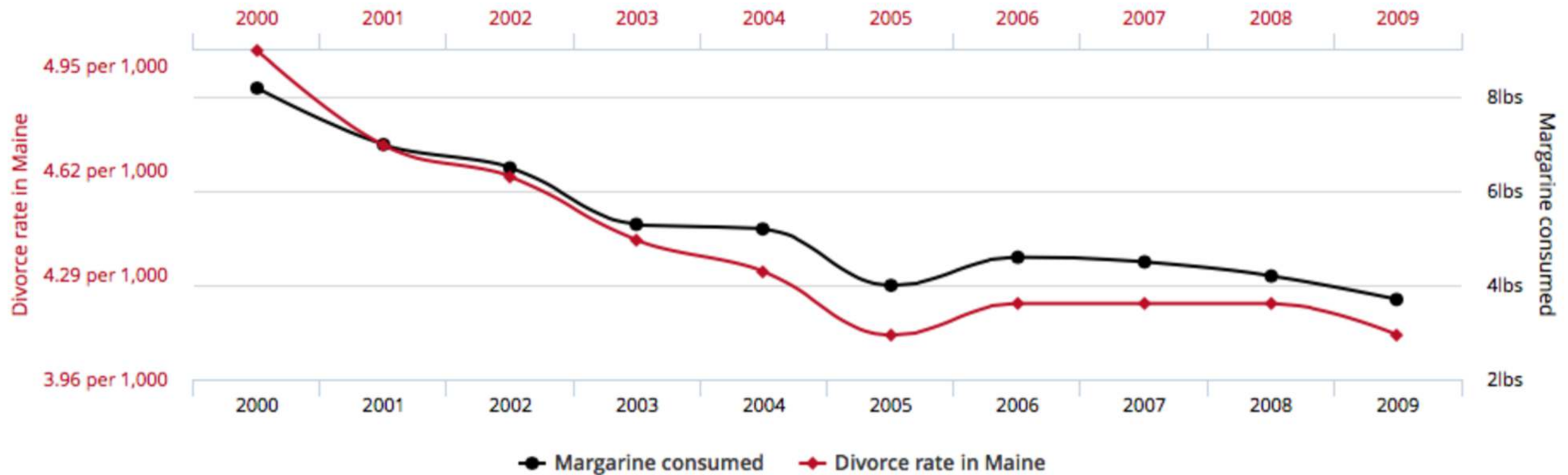
tylervigen.com

Data sources: Centers for Disease Control & Prevention and Internet Movie Database

Spurious Correlations

# Divorce rate in Maine correlates with Per capita consumption of margarine

Correlation: 99.26% ( $r=0.992558$ )



tylervigen.com

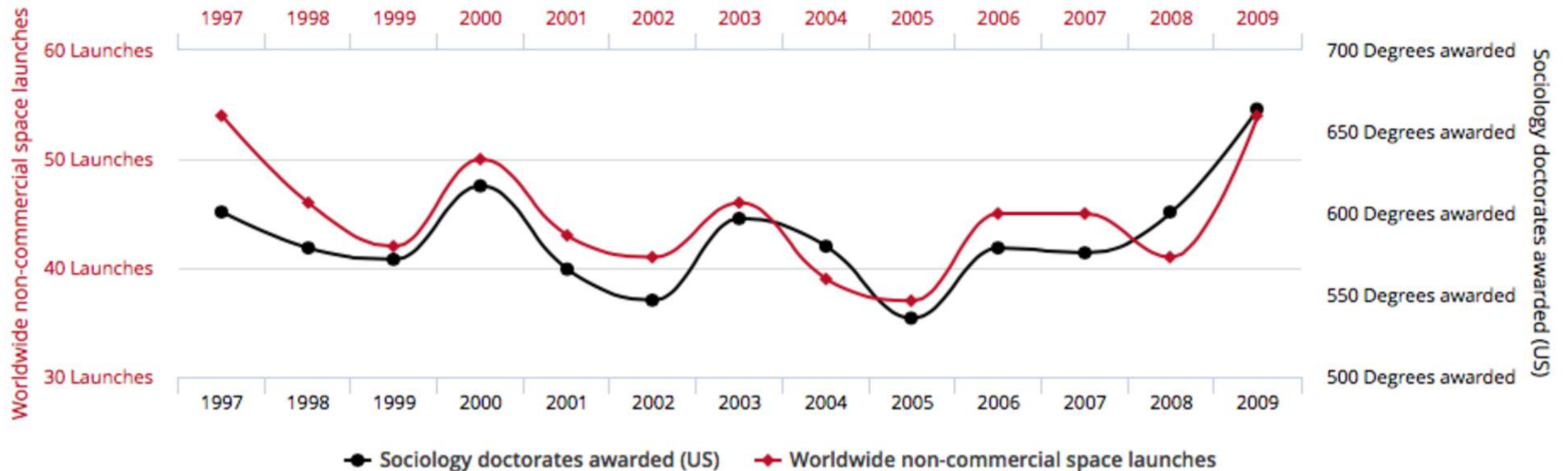
Data sources: National Vital Statistics Reports and U.S. Department of Agriculture

Spurious Correlations



# Worldwide non-commercial space launches correlates with Sociology doctorates awarded (US)

Correlation: 78.92% ( $r=0.78915$ )



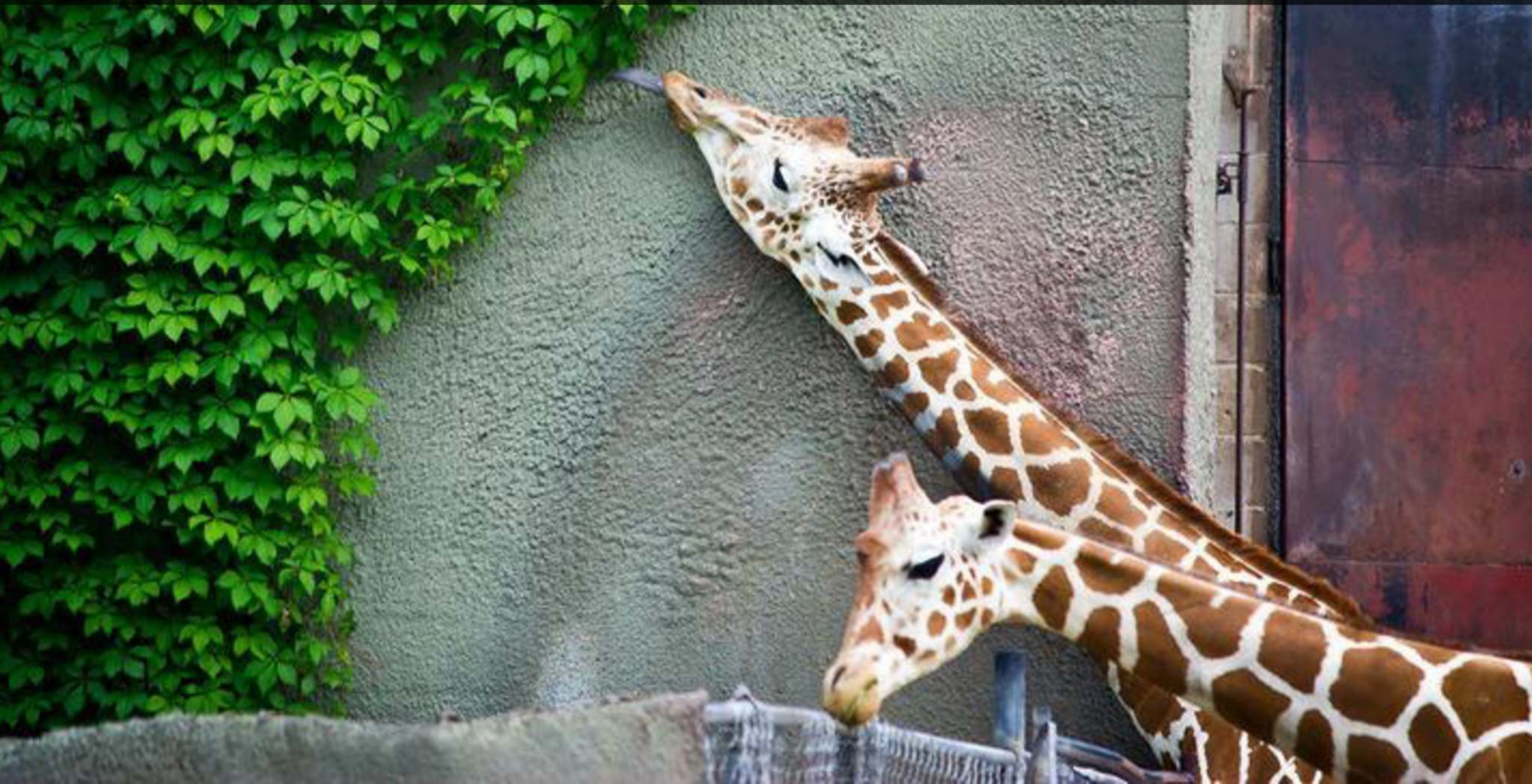
Data sources: Federal Aviation Administration and National Science Foundation

tylervigen.com

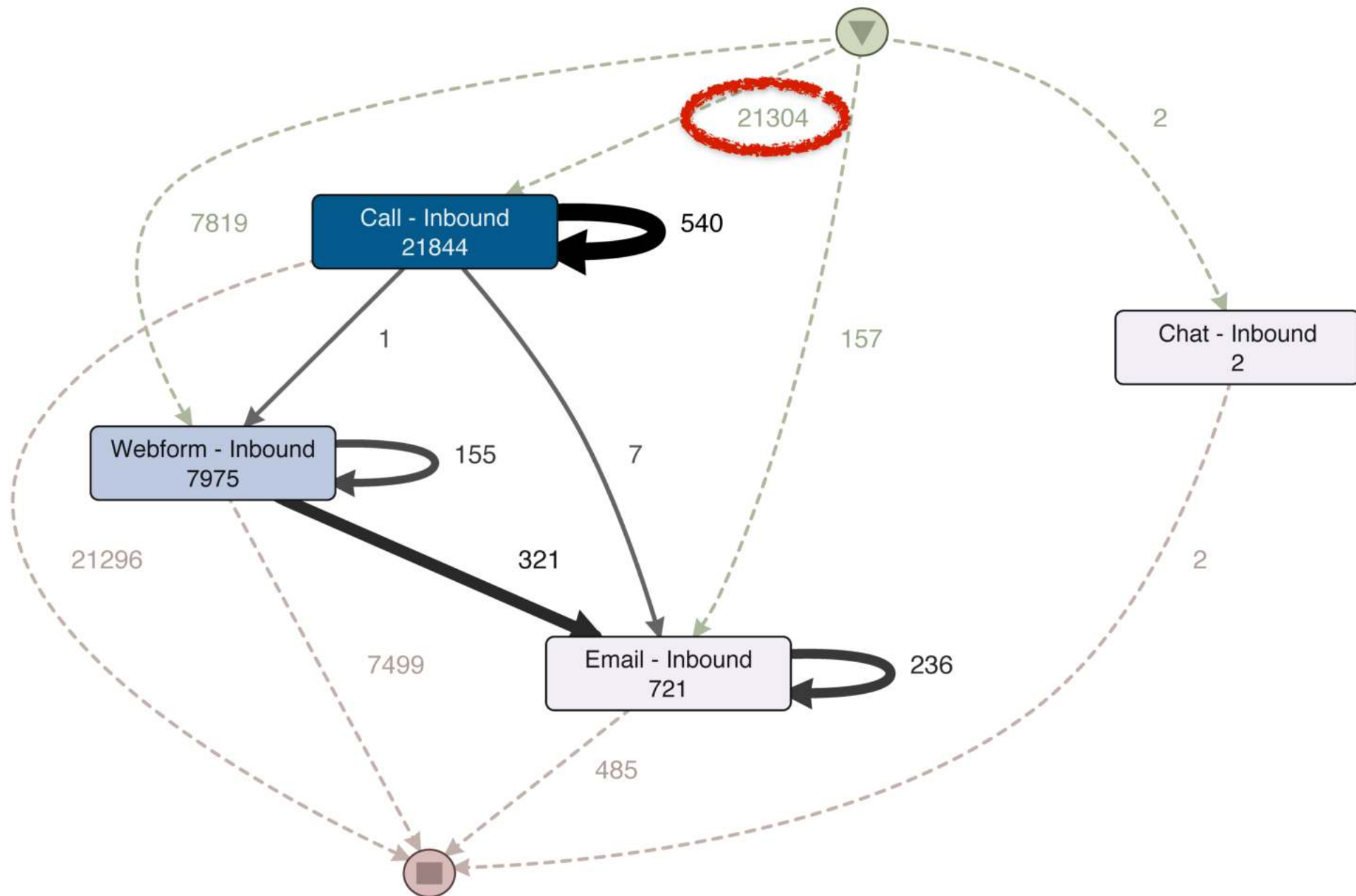
Spurious Correlations



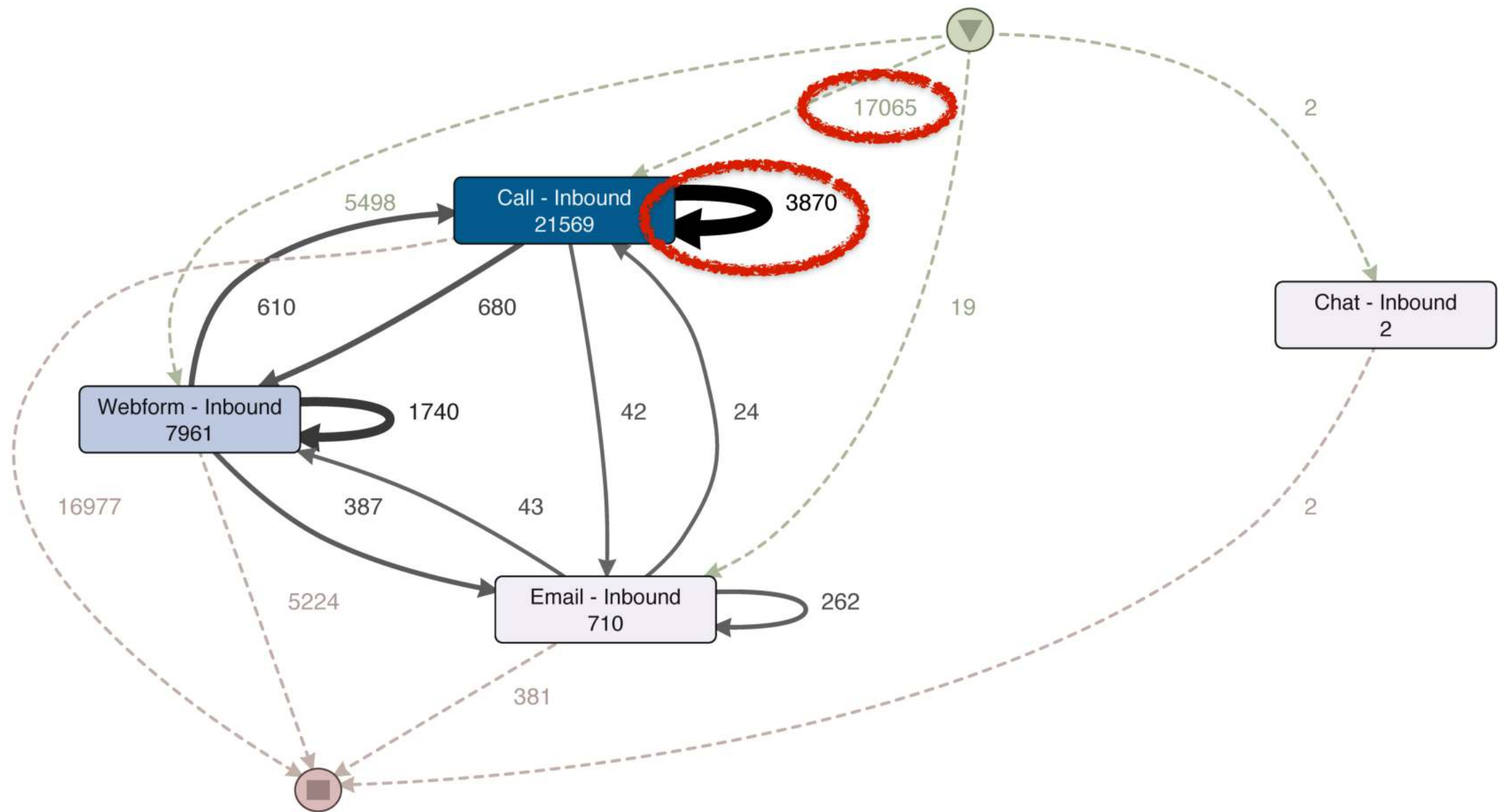
# Problem 2: You Get What You Measure







**98% First Contact Resolution rate**



**82% First Contact Resolution rate**

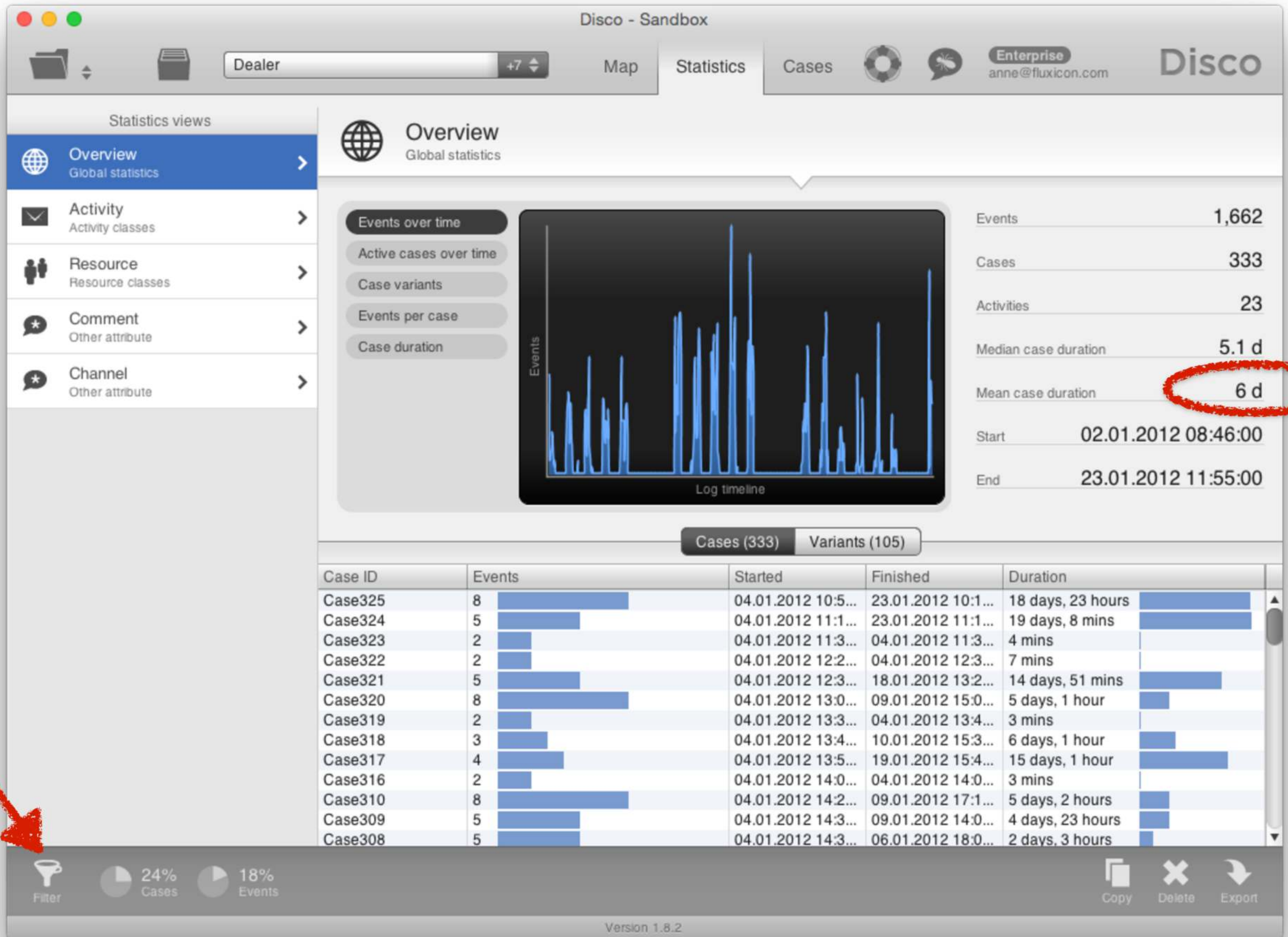


# Problem 3:

## Tip Of The Iceberg











Filter settings for: Dealer

Active filters



Timeframe

Filters by timestamp



Follower

Filters by subsequences



Attribute

Removes events by attribute



Performance

Filters cases by performance

click to add filter...



Timeframe

Filters by timestamp



January 2012

Mon	Tue	Wed	Thu	Fri	Sat	Sun
26	27	28	29	30	31	1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5



00 : 00 : 00

Keep cases:



- ☐ Contained in timeframe
- ☐ Intersecting timeframe
- ☒ Started in timeframe
- ☐ Completed in timeframe
- ☐ Trim to timeframe

January 2012

Mon	Tue	Wed	Thu	Fri	Sat	Sun
26	27	28	29	30	31	1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5



00 : 00 : 00



Reset



Undo changes



Recipe

Cancel

Copy and filter

Apply filter





# Inspecting The Quality of Metrics



- 
- A photograph of three young goats in a field. The goat on the left is dark-colored, while the two on the right are light-colored with some brown markings. They are standing in a field of dry grass and some blue plastic debris. The text is overlaid on the image in a large, white, sans-serif font.
- 1. Validity**
  - 2. Reproducibility**
  - 3. Stability**
  - 4. Analyzability**

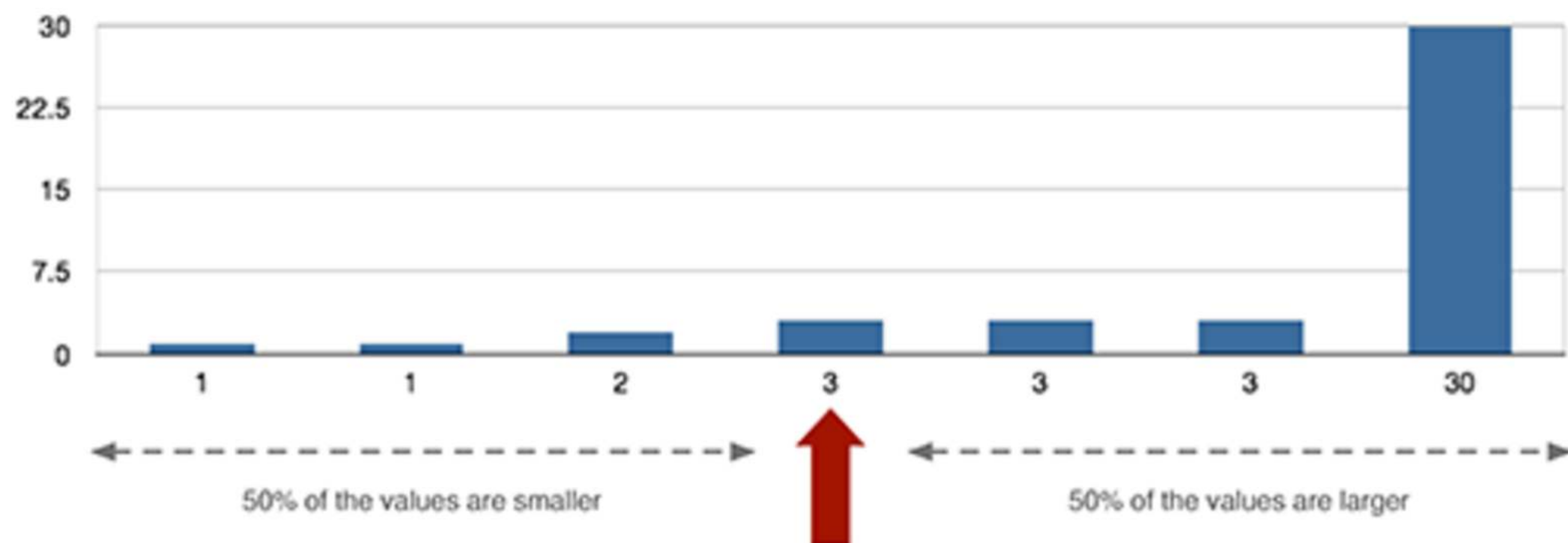
**Validity:** Sufficient correlation of the measured value to the actual property to measure



**Reproducibility:** The measurement process can be repeated at any time and will lead to the same results

**Stability:** The measured value is only correlated to the actual property to measure

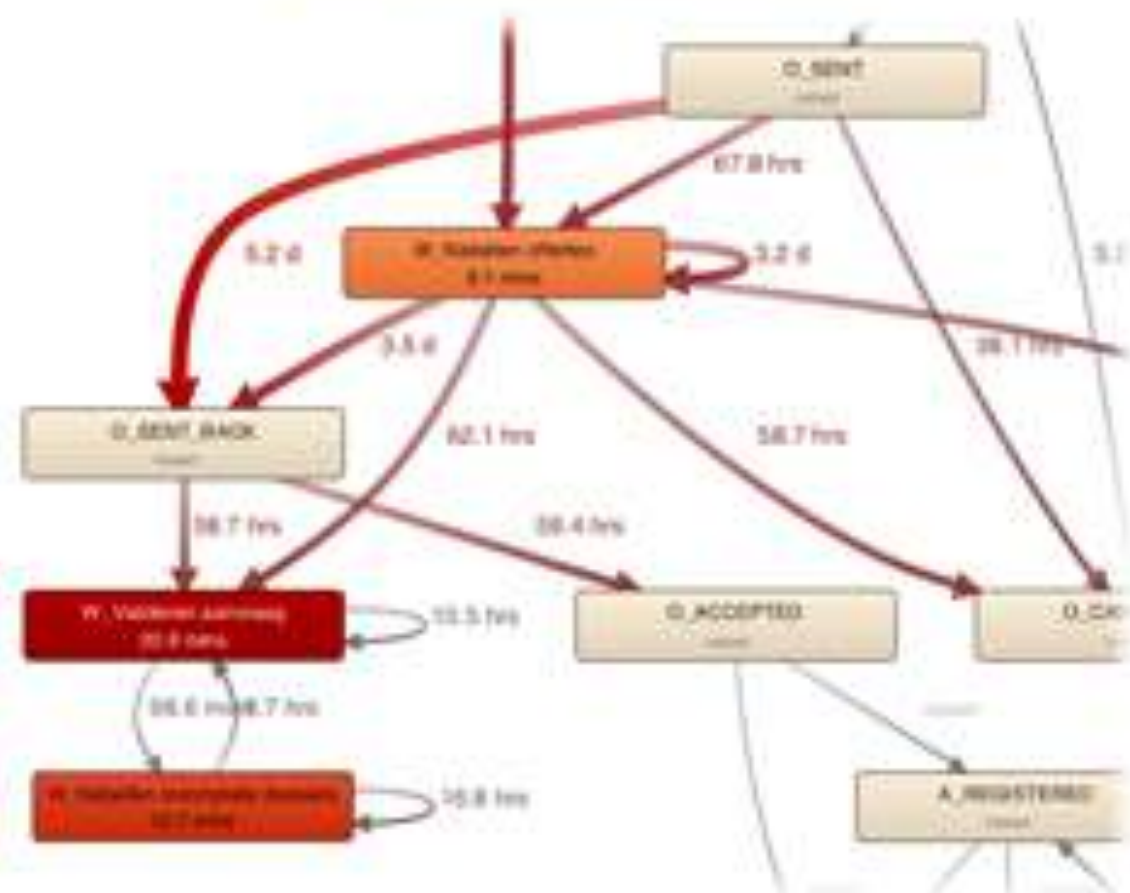




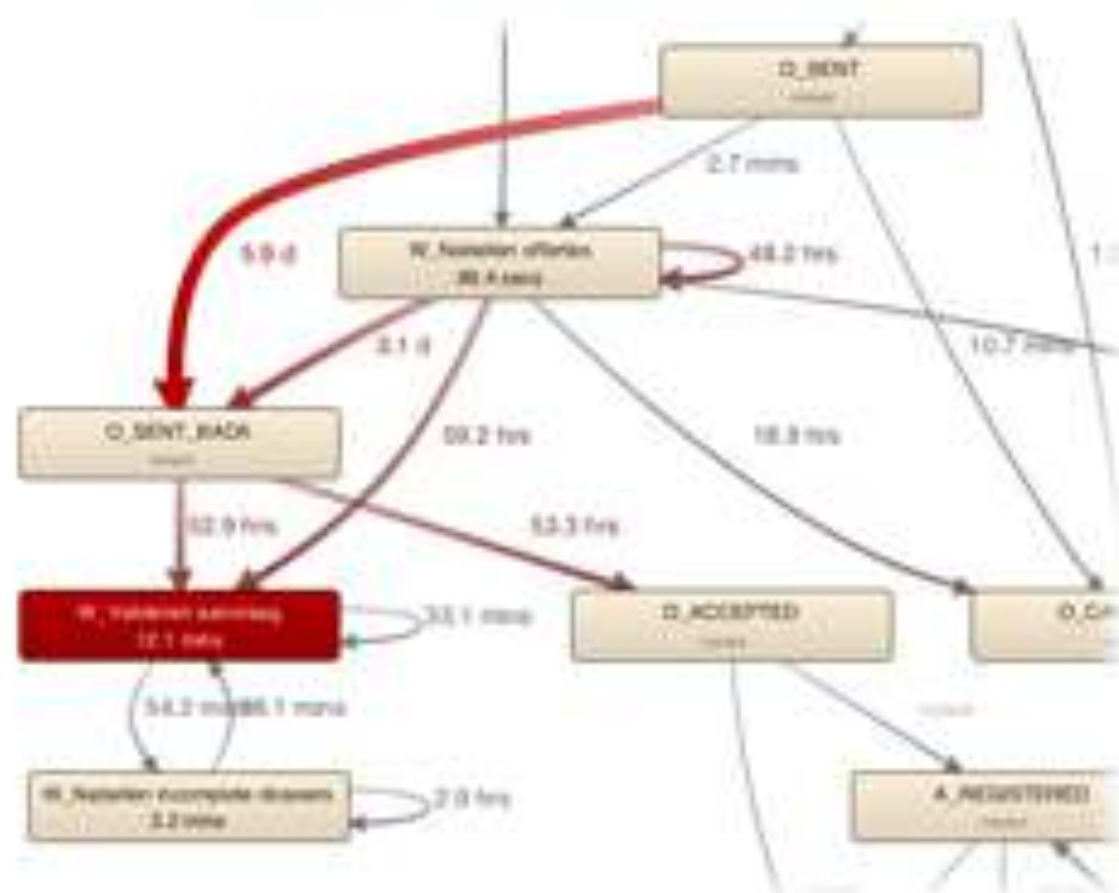
**Median = 3**

$$\begin{aligned}\text{Mean} &= (1 + 1 + 2 + 3 + 3 + 3 + 30) / 7 \\ &= 6.14\end{aligned}$$

## Mean Performance



## Median Performance





**Analyzability:** The degree of comparability of two measured values

A hedgehog is the central figure, wearing a white dress shirt and a dark tie. It is positioned in the center of the frame, looking directly at the camera. The background is a dark, textured surface, possibly a wall or a large rock, with some greenery visible at the top. The lighting is soft, highlighting the hedgehog's spines and the texture of its clothing.

**(1) Nominal**

**(2) Ordinal**

**(3) Interval**

**(4) Ratio Scale**



Metric Scale	Example
Nominal	“red”, “green”, and “blue” → No comparison

Metric Scale	Example
Nominal	“red”, “green”, and “blue” → No comparison
Ordinal	“Night”, “Dawn”, “Noon”, “Afternoon”, “Evening” → Order, no equal spacing



Metric Scale	Example
Nominal	“red”, “green”, and “blue” → No comparison
Ordinal	“Night”, “Dawn”, “Noon”, “Afternoon”, “Evening” → Order, no equal spacing
Interval	Temperature measured in Celsius or Fahrenheit → Equal spacing, comparing differences possible

Metric Scale	Example
Nominal	“red”, “green”, and “blue” → No comparison
Ordinal	“Night”, “Dawn”, “Noon”, “Afternoon”, “Evening” → Order, no equal spacing
Interval	Temperature measured in Celsius or Fahrenheit → Equal spacing, comparing differences possible
Ratio	Temperature measured in Kelvin → Absolute 0, “twice as much” comparisons possible





**So, what ARE  
Process Mining  
Metrics?**

Process Area	Metric
Number of steps	<ul style="list-style-type: none"><li>• Average/median events per case</li></ul>



## Case I: Complexity of a monitor defect.

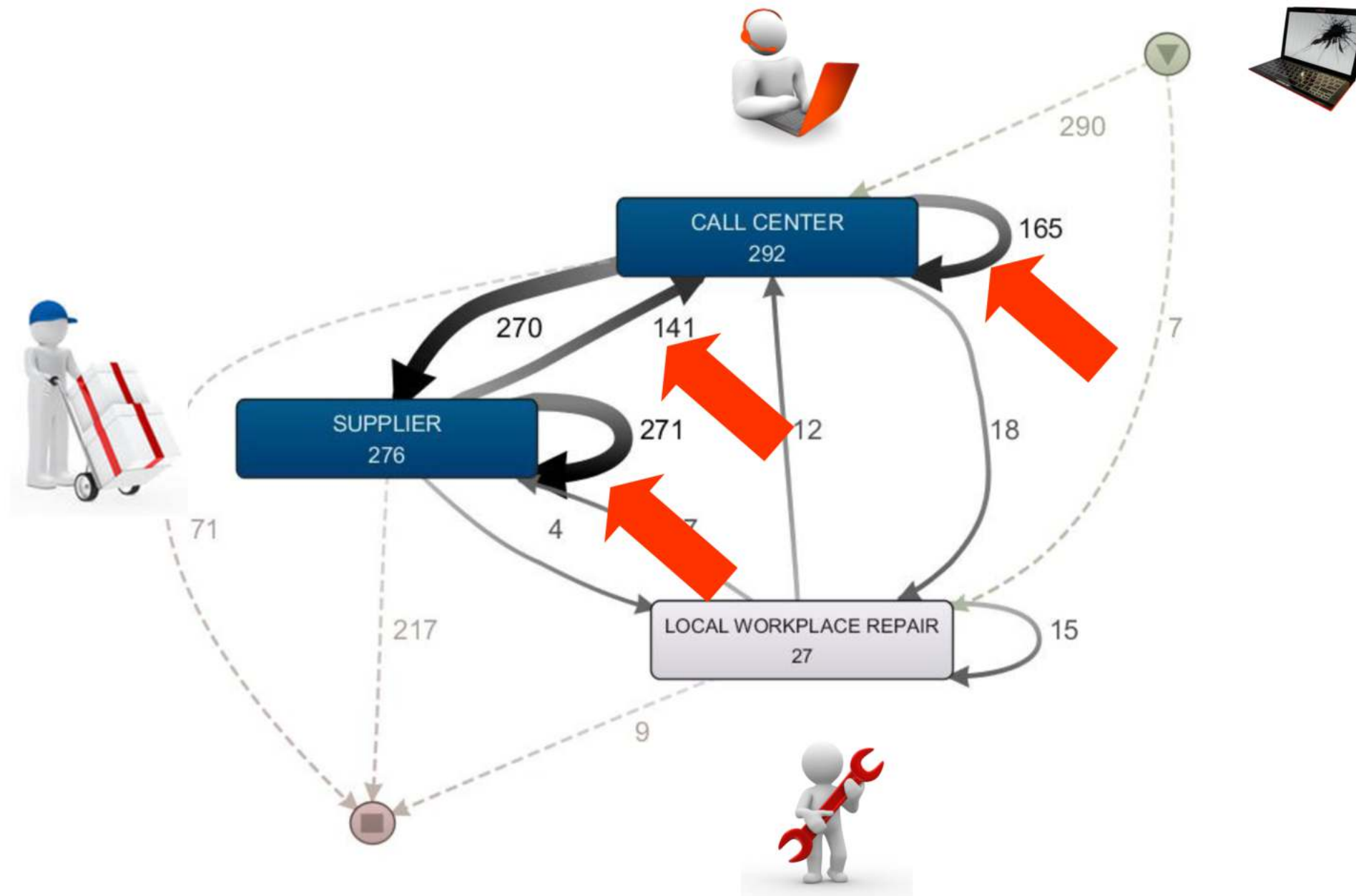


**11**steps



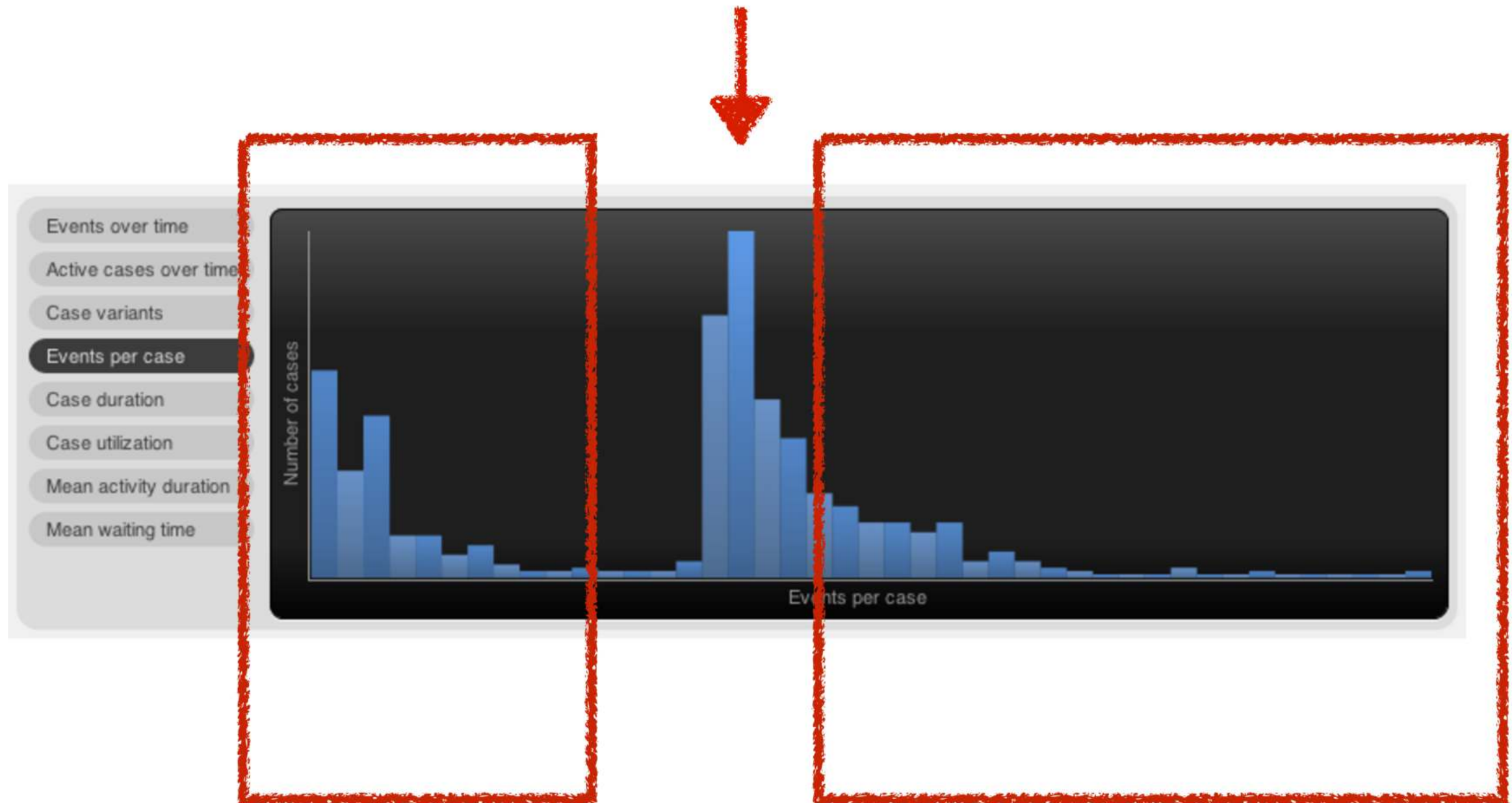


# Can you spot the rework and potential waste?



Process Area	Metric
Number of steps	<ul style="list-style-type: none"><li>• Average/median events per case</li><li>• % of cases within a certain range</li></ul>

# Distribution of events per case

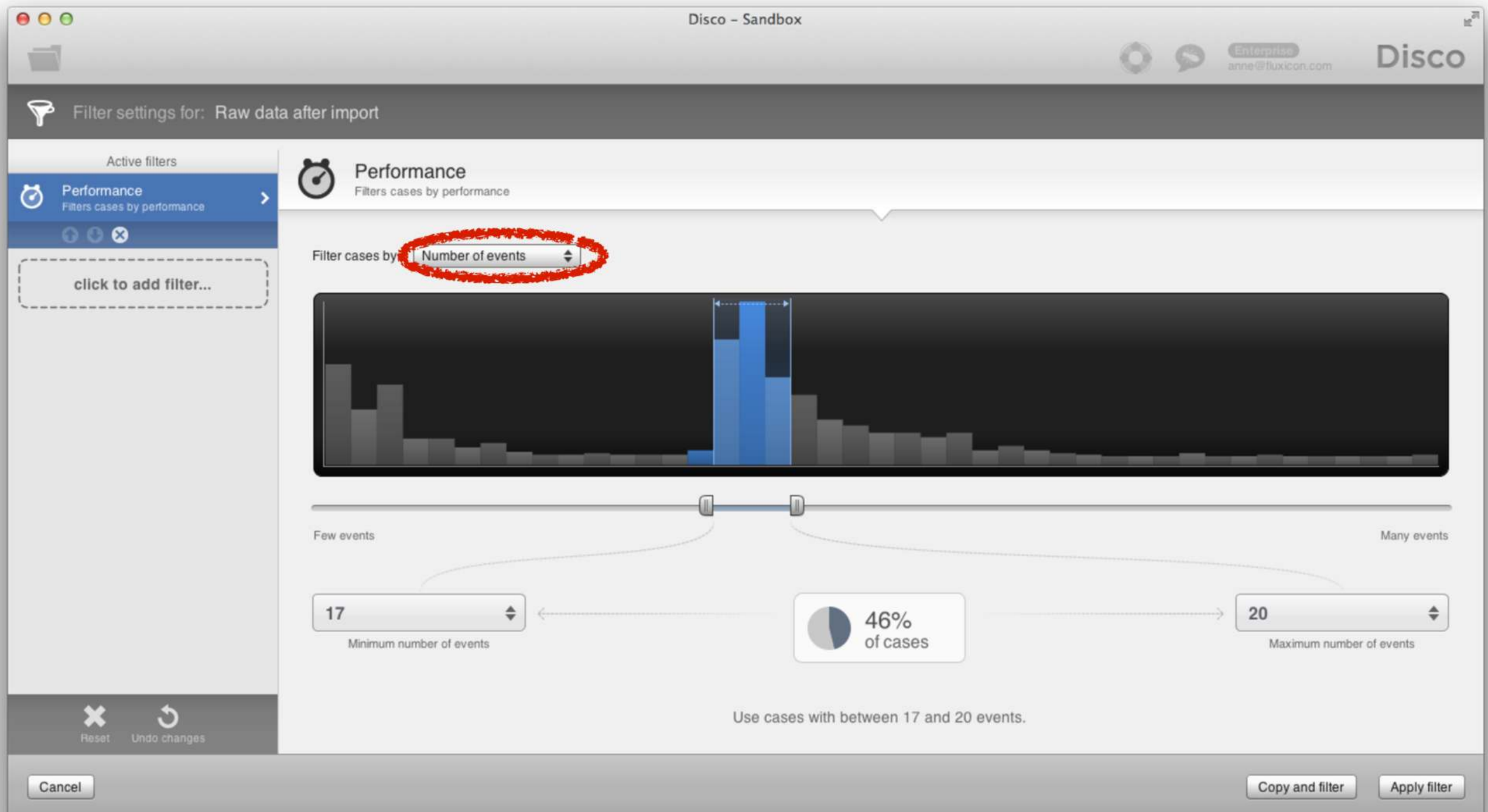


Unnecessary cases

Unnecessary steps



# % of Cases falling into desired range



Process Area	Metric
Number of steps	<ul style="list-style-type: none"><li>• Average/median events per case</li><li>• % of cases within a certain range</li></ul>
Variability	<ul style="list-style-type: none"><li>• No. of variants</li></ul>

### Cases (50)

Case360

Case1254  
10 events

Case1257  
14 events

Case357  
12 events

Case344  
13 events

Case1296  
12 events

Case91  
14 events

Case70  
12 events

Case64  
13 events

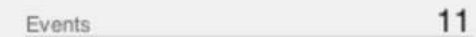
Case1221  
13 events

Case1220  
18 events

Case30  
10 events

Case 24

### Case with 11 events



Start 28.11.2011 09:27:00

Duration 50 days, 23 hours

Graph Table

[illegible]



Disco - Sandbox

Callcenter +7

Map Statistics Cases

Enterprise  
anne@fluxicon.com

Disco

search...

Variants (36)

Cases (50)

Case360  
Case with 11 events

Complete log  
All cases (50)

Variant 1  
6 cases (12%)

Variant 2  
3 cases (6%)

Variant 3  
3 cases (6%)

Variant 4  
3 cases (6%)

Variant 5  
2 cases (4%)

Variant 6  
2 cases (4%)

Variant 7  
2 cases (4%)

Variant 8  
1 case (2%)

Variant 9  
1 case (2%)

Variant 10  
1 case (2%)

Variant 11  
1 case (2%)

Variant 12

Filter

3% Cases 6% Events

Disco - Sandbox

Dealer +7

Map Statistics Cases

Enterprise  
anne@fluxicon.com

Disco

search...

Variants (367)

Cases (1205)

Case1244  
Case with 11 events

Complete log  
All cases (1205)

Variant 1  
166 cases (13.8%)

Variant 2  
69 cases (5.73%)

Variant 3  
58 cases (4.81%)

Variant 4  
57 cases (4.73%)

Variant 5  
50 cases (4.15%)

Variant 6  
36 cases (2.99%)

Variant 7  
35 cases (2.9%)

Variant 8  
31 cases (2.57%)

Variant 9  
20 cases (1.66%)

Variant 10  
19 cases (1.58%)

Variant 11  
14 cases (1.16%)

Variant 12

Case1245  
13 events

Case1246  
7 events

Case1247  
12 events

Case1248  
11 events

Case1249  
4 events

Case369  
2 events

Case368  
2 events

Case367  
2 events

Case366  
3 events

Case365  
3 events

Case364  
2 events

Case363

Case1244  
Case with 11 events

Events 11

Start 15.12.2011 08:41:00

Duration 32 days, 3 hours

Graph Table

	Activity	Resource	Date	Time	Comment	Channel
1	Order created	Value 1	15.12.2011	08:41:00	Value 1	Dealer
2	Shipment via logistics partner	Value 2	15.12.2011	08:48:00	Value 2073	Dealer
3	Product received	Value 17	19.12.2011	16:30:00	Value 2074	Dealer
4	Repairshop X - Checking	Value 17	19.12.2011	16:30:00	Value 2075	Dealer
5	Repairshop X - OK	Value 5	20.12.2011	08:24:00	Value 2076	Dealer
6	Request signed L1	Value 2	21.12.2011	10:07:00	Value 1	Dealer
7	Request signed L2	Value 6	21.12.2011	12:31:00	Value 1	Dealer
8	Request signed L3	Value 7	21.12.2011	13:19:00	Value 1	Dealer
9	Amount modified	Value 8	16.01.2012	11:17:00	Value 2077	Dealer
10	Special Refund issued	Value 8	16.01.2012	11:18:00	Value 2078	Dealer
11	Warehouse	Value 8	16.01.2012	12:21:00	Value 1	Dealer

Filter 90% Cases 81% Events

Copy Delete Export

Version 1.8.2

Process Area	Metric
Number of steps	<ul style="list-style-type: none"><li>• Average/median events per case</li><li>• % of cases within a certain range</li></ul>
Variability	<ul style="list-style-type: none"><li>• No. of variants</li><li>• Variant to cases ratio</li></ul>

Small no. of variants covering many cases ...  
.. or many different and mostly unique cases?

$$\text{Variability in \%} = \frac{\# \text{ Variants}}{\# \text{ Cases}} * 100$$



$$\text{Variability in \%} = \frac{\# \text{ Variants}}{\# \text{ Cases}} * 100$$

Example with highest variability:  
**20 unique cases** (i.e., 20 variants)

$$\begin{aligned}\text{Variability in \%} &= \frac{20}{20} * 100 \\ &= 1 * 100 \\ &= \mathbf{100\%}\end{aligned}$$

$$\text{Variability in \%} = \frac{\# \text{ Variants}}{\# \text{ Cases}} * 100$$

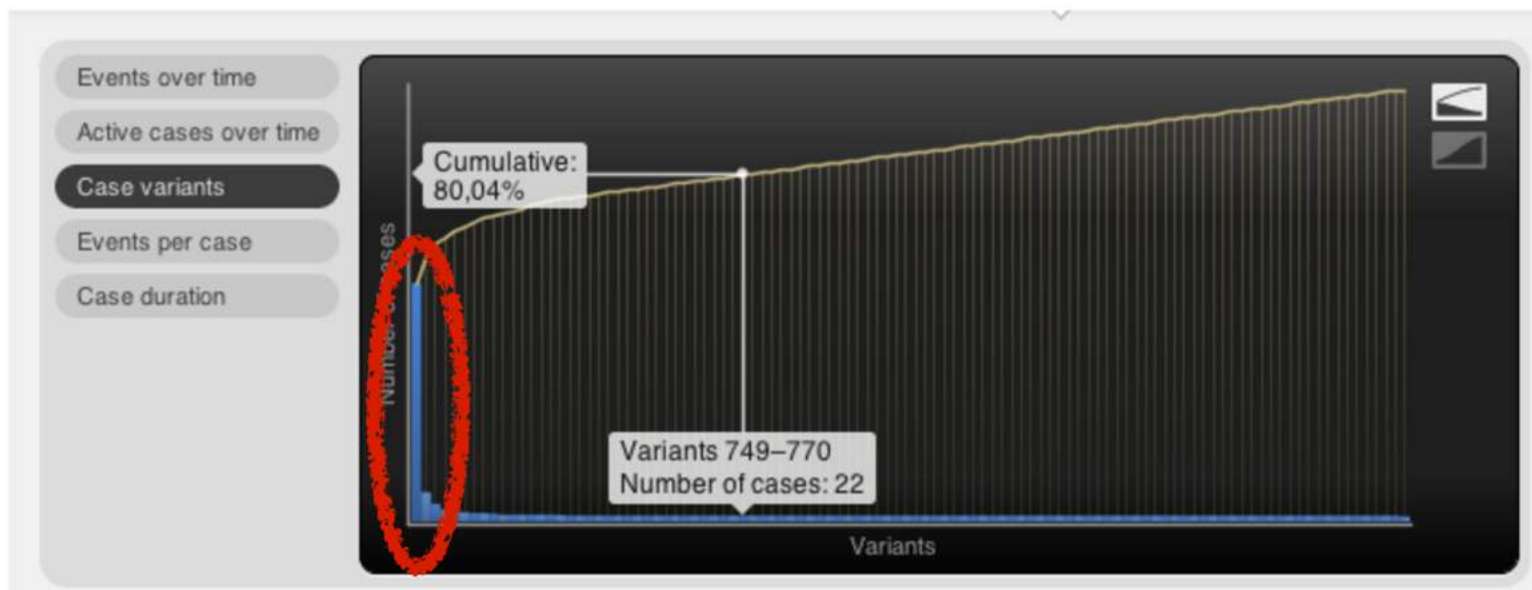
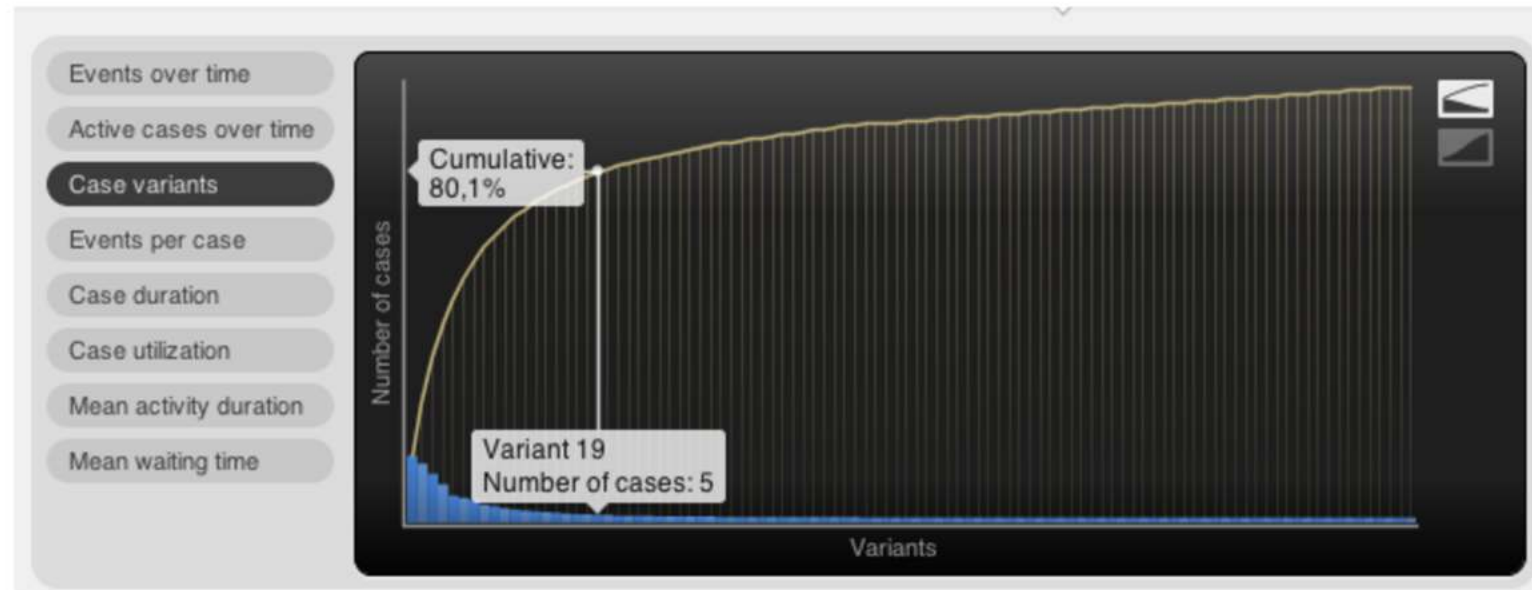
Example with low variability:  
20 cases following **two different variants**

$$\begin{aligned}\text{Variability in \%} &= \frac{2}{20} * 100 \\ &= 0.1 * 100 \\ &= \mathbf{10\%}\end{aligned}$$

Process Area	Metric
Number of steps	<ul style="list-style-type: none"><li>• Average/median events per case</li><li>• % of cases within a certain range</li></ul>
Variability	<ul style="list-style-type: none"><li>• No. of variants</li><li>• Variant to cases ratio</li><li>• Distribution of variants</li></ul>



# Process variant distribution

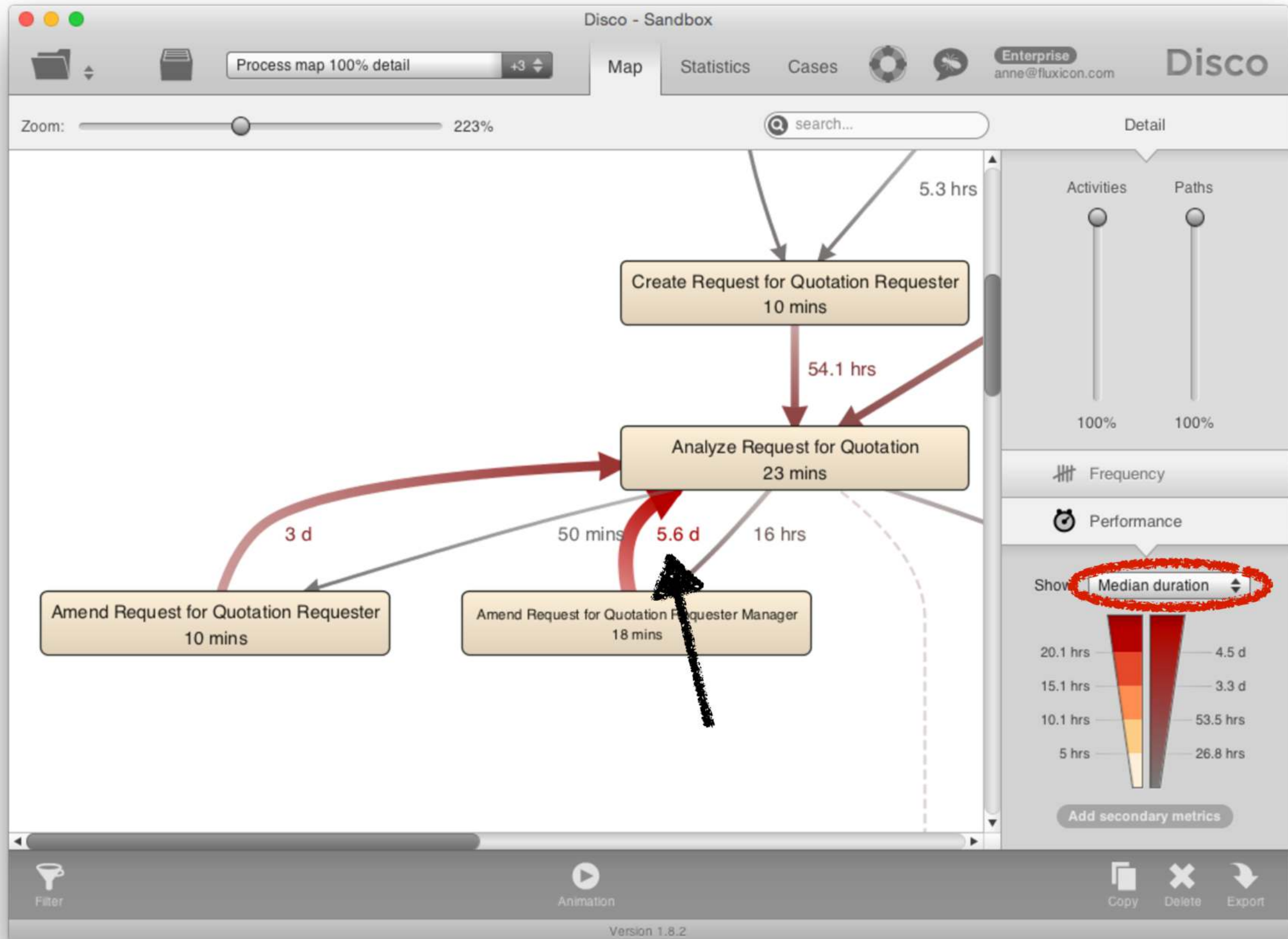


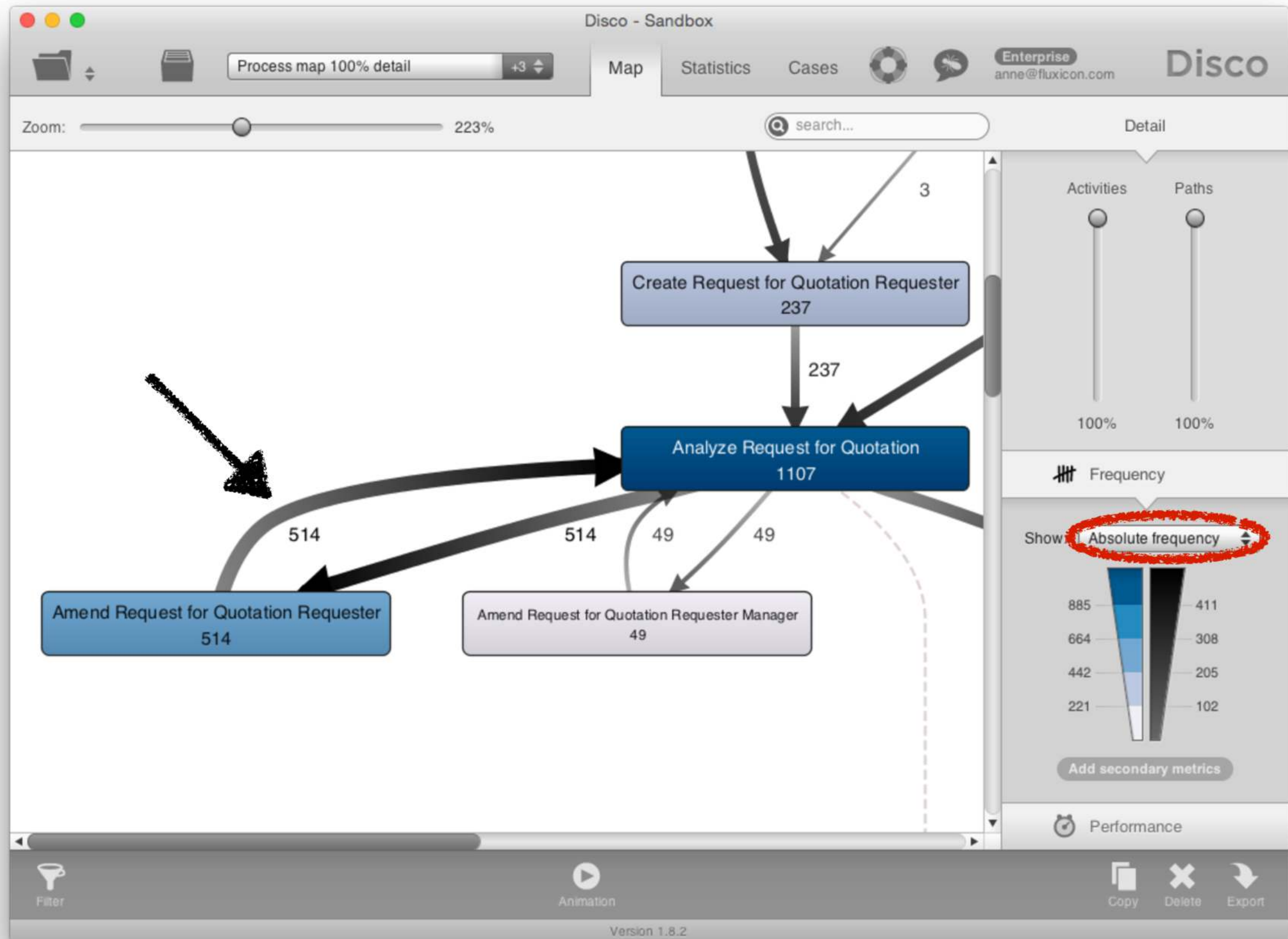
Process Area	Metric
Number of steps	<ul style="list-style-type: none"><li>• Average/median events per case</li><li>• % of cases within a certain range</li></ul>
Variability	<ul style="list-style-type: none"><li>• No. of variants</li><li>• Variant to cases ratio</li><li>• Distribution of variants</li></ul>
Repetitions	<ul style="list-style-type: none"><li>• No. of cases with repetitions</li><li>• Top activities causing rework</li></ul>

Process Area	Metric
Number of steps	<ul style="list-style-type: none"><li>• Average/median events per case</li><li>• % of cases within a certain range</li></ul>
Variability	<ul style="list-style-type: none"><li>• No. of variants</li><li>• Variant to cases ratio</li><li>• Distribution of variants</li></ul>
Repetitions	<ul style="list-style-type: none"><li>• No. of cases with repetitions</li><li>• Top activities causing rework</li></ul>
Case durations	<ul style="list-style-type: none"><li>• Average/median case duration</li><li>• % of cases outside of SLA</li></ul>

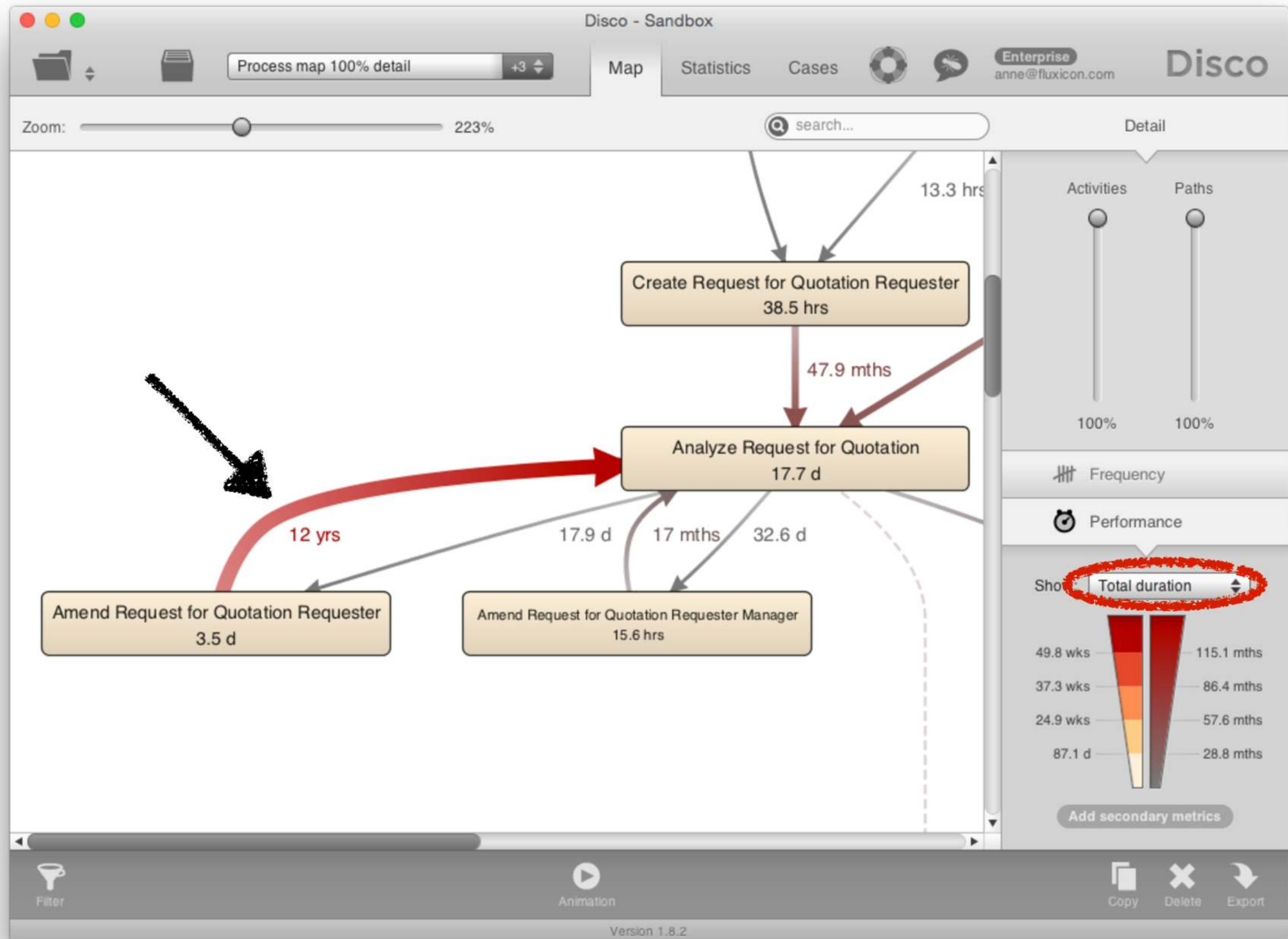


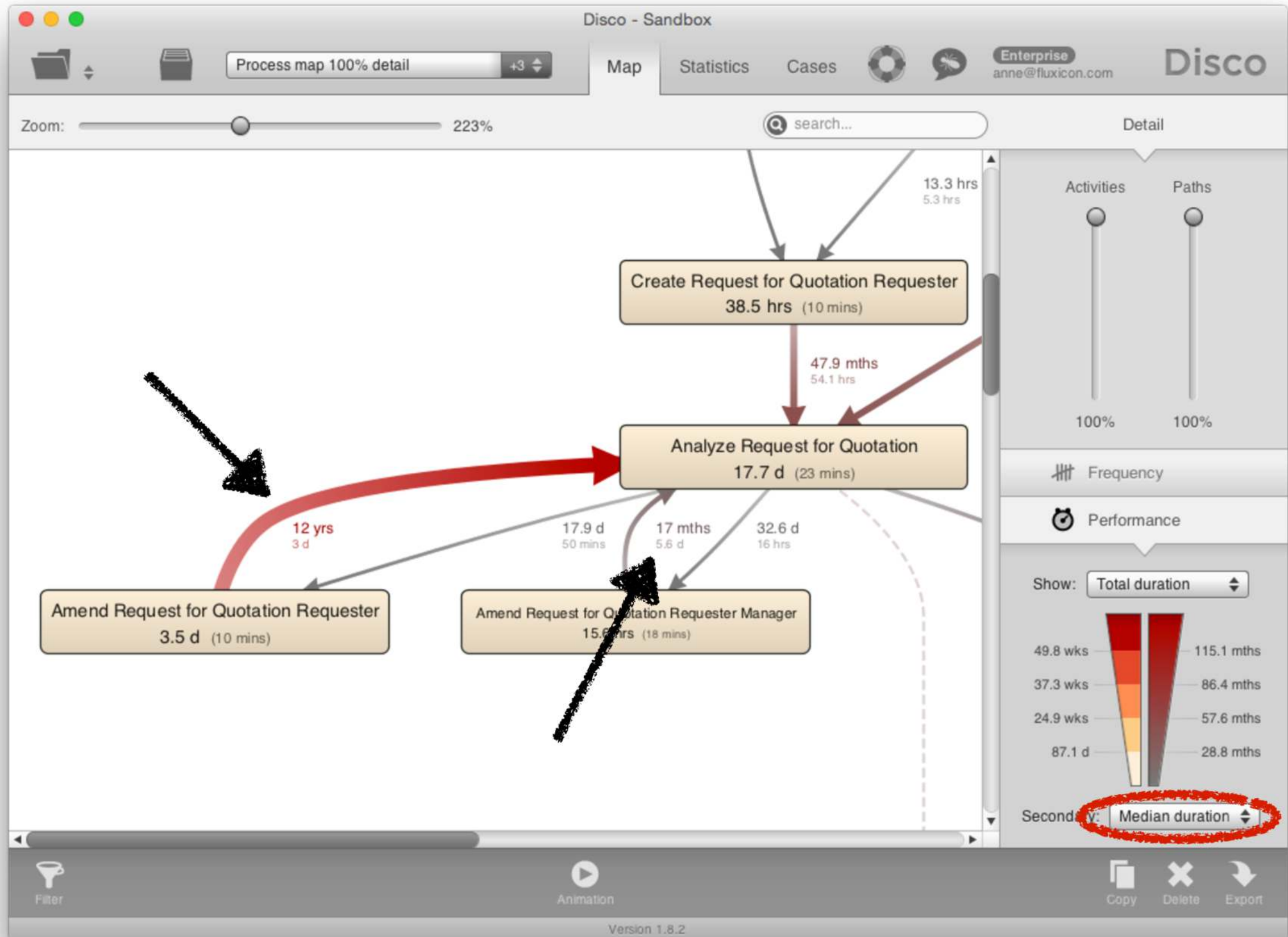
Process Area	Metric
Number of steps	<ul style="list-style-type: none"><li>• Average/median events per case</li><li>• % of cases within a certain range</li></ul>
Variability	<ul style="list-style-type: none"><li>• No. of variants</li><li>• Variant to cases ratio</li><li>• Distribution of variants</li></ul>
Repetitions	<ul style="list-style-type: none"><li>• No. of cases with repetitions</li><li>• Top activities causing rework</li></ul>
Case durations	<ul style="list-style-type: none"><li>• Average/median case duration</li><li>• % of cases outside of SLA</li></ul>
Waiting times	<ul style="list-style-type: none"><li>• Total (cumulative) delays between activities</li><li>• Difference between well-performing and slow cases</li></ul>





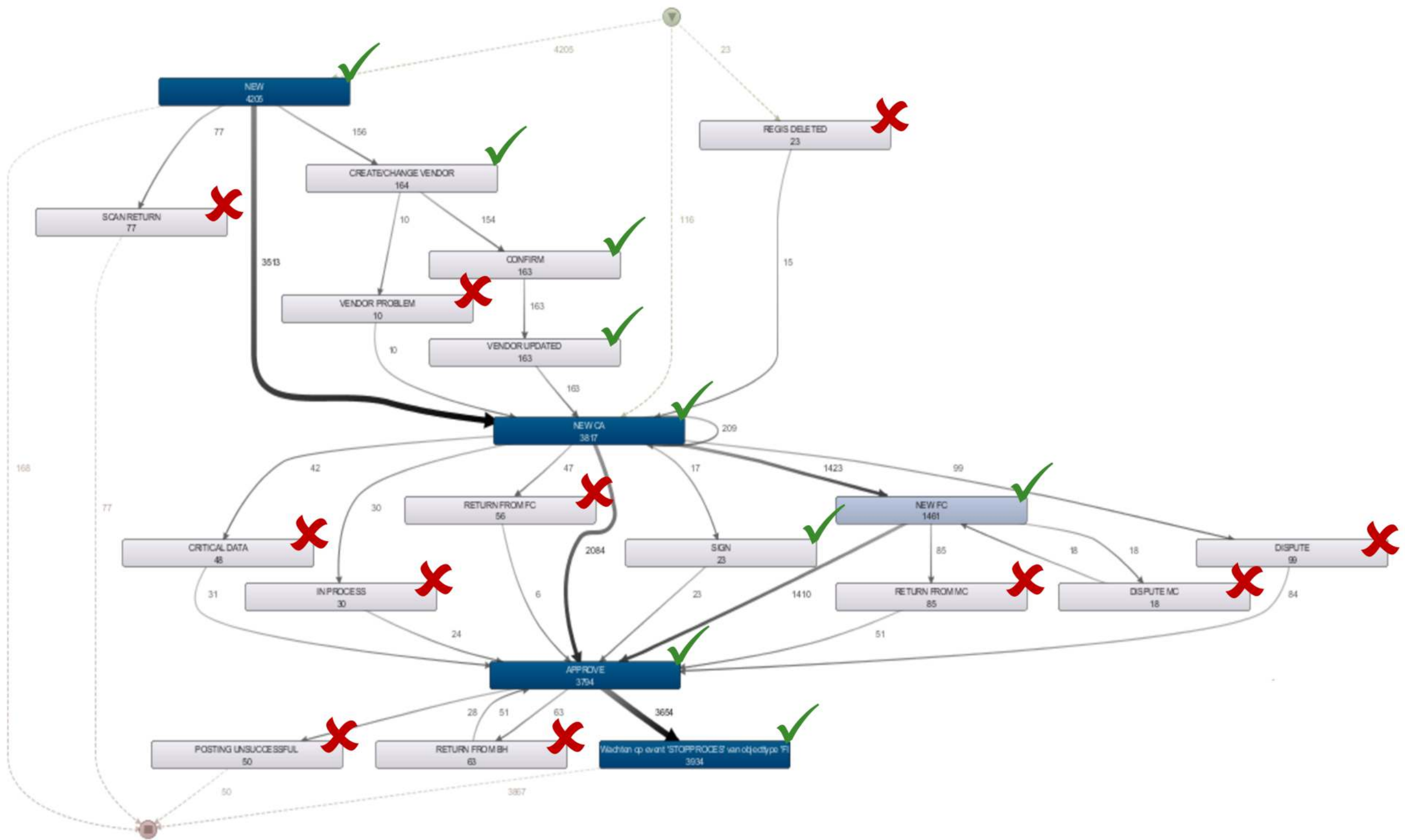






Process Area	Metric
Number of steps	<ul style="list-style-type: none"> <li>• Average/median events per case</li> <li>• % of cases within a certain range</li> </ul>
Variability	<ul style="list-style-type: none"> <li>• No. of variants</li> <li>• Variant to cases ratio</li> <li>• Distribution of variants</li> </ul>
Repetitions	<ul style="list-style-type: none"> <li>• No. of cases with repetitions</li> <li>• Top activities causing rework</li> </ul>
Case durations	<ul style="list-style-type: none"> <li>• Average/median case duration</li> <li>• % of cases outside of SLA</li> </ul>
Waiting times	<ul style="list-style-type: none"> <li>• Total (cumulative) delays between activities</li> <li>• Difference between well-performing and slow cases</li> </ul>
Value-add vs. Non-value-add activities	<ul style="list-style-type: none"> <li>• Ratio value-add vs. non-value add activities in total</li> <li>• Top non-value adding activities and root causes</li> </ul>





Process Area	Metric
Fallout	<ul style="list-style-type: none"><li>• Number or % of cases not reaching a milestone (especially in sales and ordering processes: revenue leakages)</li></ul>

Process Area	Metric
Fallout	<ul style="list-style-type: none"><li>• Number or % of cases not reaching a milestone (especially in sales and ordering processes: revenue leakages)</li></ul>
Compliance violations	<ul style="list-style-type: none"><li>• Not allowed paths (e.g., skipped approvals)</li><li>• Violated Segregation of Duties</li><li>• Violated SLAs</li><li>• Suspiciously short activities (that normally take time)</li></ul>



Process Area	Metric
Fallout	<ul style="list-style-type: none"><li>• Number or % of cases not reaching a milestone (especially in sales and ordering processes: revenue leakages)</li></ul>
Compliance violations	<ul style="list-style-type: none"><li>• Not allowed paths (e.g., skipped approvals)</li><li>• Violated Segregation of Duties</li><li>• Violated SLAs</li><li>• Suspiciously short activities (that normally take time)</li></ul>
Organizational view	<ul style="list-style-type: none"><li>• Groups handling most of the work</li><li>• Hand-over patterns</li><li>• Delays between groups</li><li>• Interdependencies between processes</li><li>• Concentration (specialization) vs. Decentralization of tasks</li></ul>

Process Area	Metric
Fallout	<ul style="list-style-type: none"><li>• Number or % of cases not reaching a milestone (especially in sales and ordering processes: revenue leakages)</li></ul>
Compliance violations	<ul style="list-style-type: none"><li>• Not allowed paths (e.g., skipped approvals)</li><li>• Violated Segregation of Duties</li><li>• Violated SLAs</li><li>• Suspiciously short activities (that normally take time)</li></ul>
Organizational view	<ul style="list-style-type: none"><li>• Groups handling most of the work</li><li>• Hand-over patterns</li><li>• Delays between groups</li><li>• Interdependencies between processes</li><li>• Concentration (specialization) vs. Decentralization of tasks</li></ul>
Workload over time	<ul style="list-style-type: none"><li>• Number of cases in the process over time</li><li>• Seasonal patterns and impacts on other processes</li></ul>

Process Area	Metric
Fallout	<ul style="list-style-type: none"> <li>• Number or % of cases not reaching a milestone (especially in sales and ordering processes: revenue leakages)</li> </ul>
Compliance violations	<ul style="list-style-type: none"> <li>• Not allowed paths (e.g., skipped approvals)</li> <li>• Violated Segregation of Duties</li> <li>• Violated SLAs</li> <li>• Suspiciously short activities (that normally take time)</li> </ul>
Organizational view	<ul style="list-style-type: none"> <li>• Groups handling most of the work</li> <li>• Hand-over patterns</li> <li>• Delays between groups</li> <li>• Interdependencies between processes</li> <li>• Concentration (specialization) vs. Decentralization of tasks</li> </ul>
Workload over time	<ul style="list-style-type: none"> <li>• Number of cases in the process over time</li> <li>• Seasonal patterns and impacts on other processes</li> </ul>
Unused paths	<ul style="list-style-type: none"> <li>• Unused paths (available in application / specified process but never used) may provide opportunity to reduce maintenance costs</li> </ul>

Process Area	Metric
Fallout	<ul style="list-style-type: none"> <li>• Number or % of cases not reaching a milestone (especially in sales and ordering processes: revenue leakages)</li> </ul>
Compliance violations	<ul style="list-style-type: none"> <li>• Not allowed paths (e.g., skipped approvals)</li> <li>• Violated Segregation of Duties</li> <li>• Violated SLAs</li> <li>• Suspiciously short activities (that normally take time)</li> </ul>
Organizational view	<ul style="list-style-type: none"> <li>• Groups handling most of the work</li> <li>• Hand-over patterns</li> <li>• Delays between groups</li> <li>• Interdependencies between processes</li> <li>• Concentration (specialization) vs. Decentralization of tasks</li> </ul>
Workload over time	<ul style="list-style-type: none"> <li>• Number of cases in the process over time</li> <li>• Seasonal patterns and impacts on other processes</li> </ul>
Unused paths	<ul style="list-style-type: none"> <li>• Unused paths (available in application / specified process but never used) may provide opportunity to reduce maintenance costs</li> </ul>
Automation opportunities	<ul style="list-style-type: none"> <li>• Manual steps that can be automated</li> </ul>



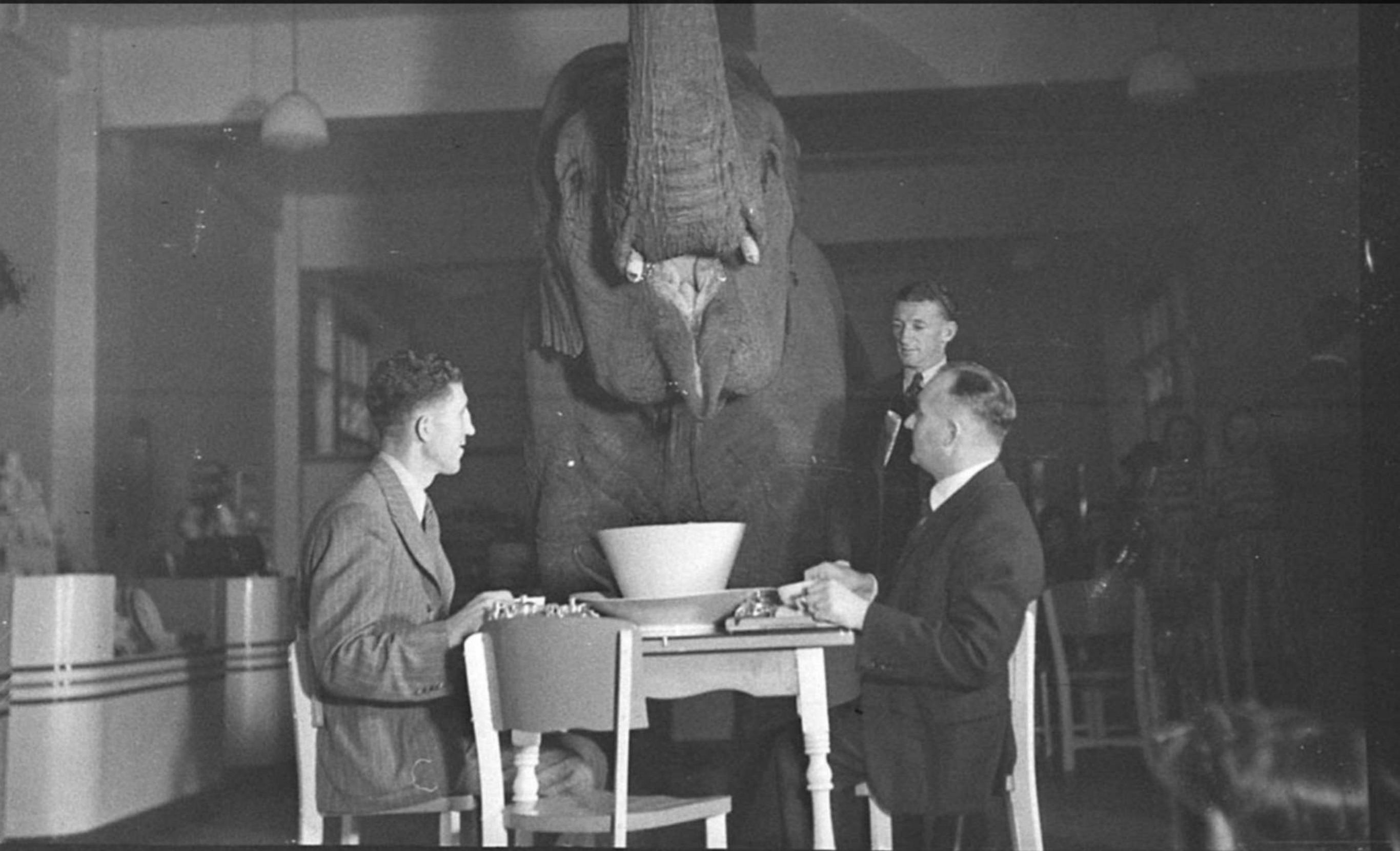
Process Area	Metric
Fallout	<ul style="list-style-type: none"> <li>• Number or % of cases not reaching a milestone (especially in sales and ordering processes: revenue leakages)</li> </ul>
Compliance violations	<ul style="list-style-type: none"> <li>• Not allowed paths (e.g., skipped approvals)</li> <li>• Violated Segregation of Duties</li> <li>• Violated SLAs</li> <li>• Suspiciously short activities (that normally take time)</li> </ul>
Organizational view	<ul style="list-style-type: none"> <li>• Groups handling most of the work</li> <li>• Hand-over patterns</li> <li>• Delays between groups</li> <li>• Interdependencies between processes</li> <li>• Concentration (specialization) vs. Decentralization of tasks</li> </ul>
Workload over time	<ul style="list-style-type: none"> <li>• Number of cases in the process over time</li> <li>• Seasonal patterns and impacts on other processes</li> </ul>
Unused paths	<ul style="list-style-type: none"> <li>• Unused paths (available in application / specified process but never used) may provide opportunity to reduce maintenance costs</li> </ul>
Automation opportunities	<ul style="list-style-type: none"> <li>• Manual steps that can be automated</li> </ul>
Open cases analysis	<ul style="list-style-type: none"> <li>• How long have incomplete cases been open?</li> <li>• At which stage are most cases waiting?</li> </ul>



**Some Last Words  
On The Way**



# People are afraid





**You need their  
participation**





**You may not see  
everything**







**Thank You!**