



Process Mining Tutorial

Adapted from
fluxicon.com/academic/material/

Edited by Andrea Vandin, SSSA

Goals of this tutorial

- Understand the phases of **process mining** analysis
- Be able to **get started and play** around with your own data

Outline

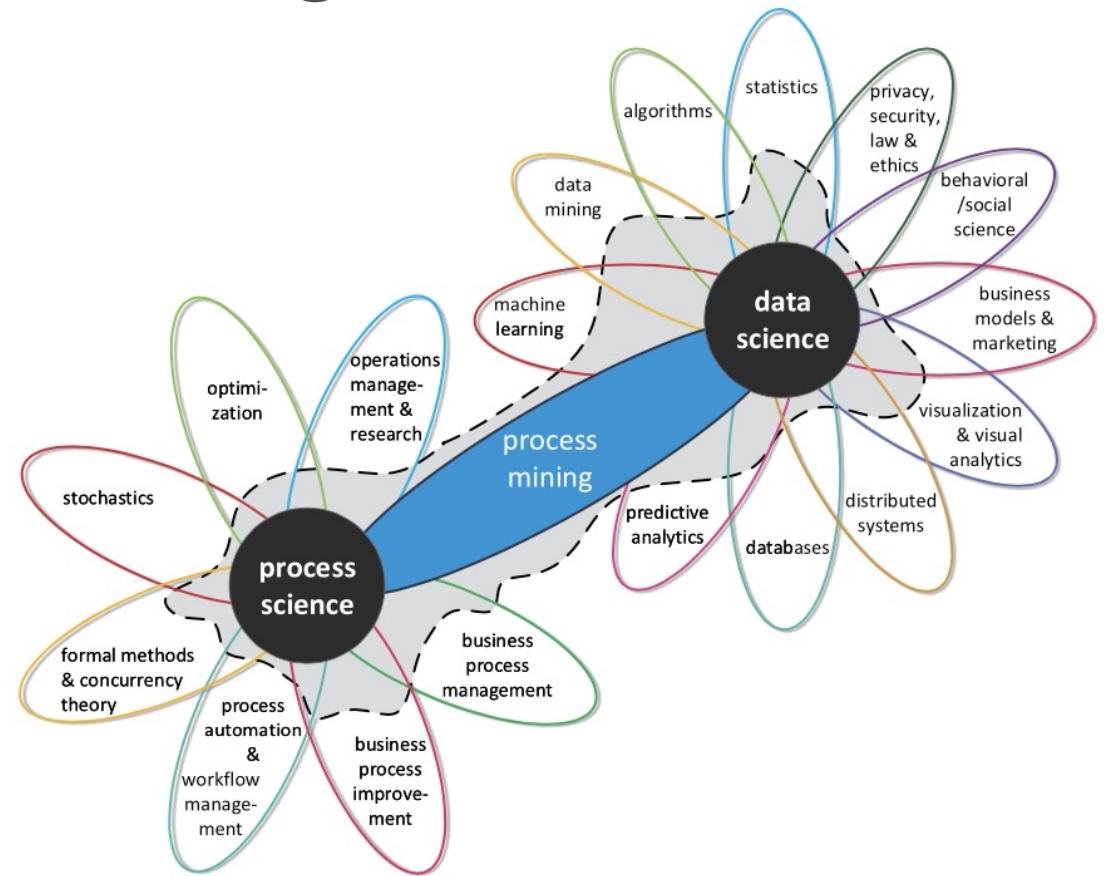
- 0. What is Process mining?**
 - 1. Example Scenario
 - 2. Roadmap
 - 3. Hands-on Session
 - 4. Take-away Points

What is Process Mining?

Process mining is a **family of techniques linking data science and process management to support the analysis of operational processes based on event logs.**

The goal of process mining is to turn event data into insights and actions. Process mining is an integral part of data science, fueled by the availability of data and the desire to improve processes.

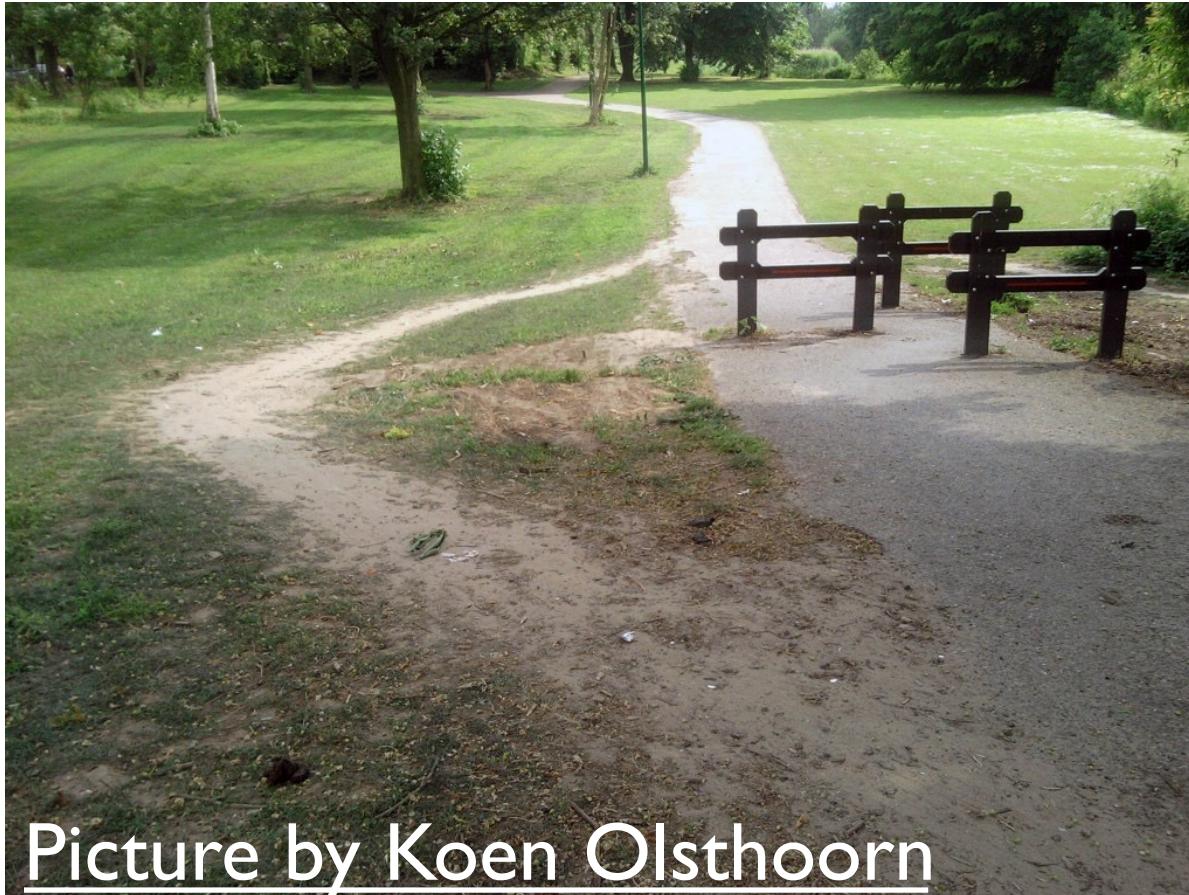
E.g., *process mining* uses event data to automatically discover a process model by observing events recorded by some enterprise system



Van Der Aalst, W., et al. (2011, August).

Process mining manifesto. In *International Conference on Business Process Management* (pp. 169-194). Springer

What can we do with Process Mining?



Picture by Koen Olsthoorn

With PM we can
find the actual
process is
different from
the one we
designed!

What can we do with Process Mining?



Picture by Koen Olsthoorn

With PM we can
find the actual
process is
different from
the one we
designed!

Our traces are the process logs...

We will see a simple example (purchasing process)

However, PM has been applied to several domains, e.g., the **legal** one

Challenges in Legal Process Discovery*

Hugo A. López¹

Department of Computer Science
University of Copenhagen**

Abstract. One of the main promises of process conformance is the opportunity to align normative processes (i.e. how the process *should* behave) and event logs (i.e. how does the process *actually* behaves). Results of conformance checking are valid as long as normative processes correspond to actual norms. Recent developments advocate the use of Natural Language Processing (NLP) to process model discovery from texts. We present a series of challenges in textual process discovery that limit its applicability to real norms. The challenges emerges from experiences with legal practitioners in the digitalization of administrative processes in Danish and Italian municipalities, and they need to be solved in order to provide accurate normative processes that reflect the intent of laws.

Key words: Natural Language Processing, Normative Processes, Process Discovery, Process Conformance

Business Process Compliance using Reference Models of Law

Hugo A. López^{1,3} , Søren Debois² , Tijs Slaats¹ , and Thomas T. Hildebrandt¹ 

¹ Software, Data, People & Society Section
Department of Computer Science
Copenhagen University, Denmark
`{hala,slaats,hilde}@di.ku.dk`

² Computer Science Department, IT University of Copenhagen, Denmark
`debois@itu.dk`

³ DCR Solutions A/S, Denmark

Abstract. Legal compliance is an important part of certifying the correct behaviour of a business process. To be compliant, organizations might hard-wire regulations into processes, limiting the discretion that workers have when choosing what activities should be executed in a case. Worse, hard-wired compliant processes are difficult to change when laws change, and this occurs very often. This paper proposes a model-driven approach to process compliance and combines a) reference models from laws, and b) business process models. Both reference and process models

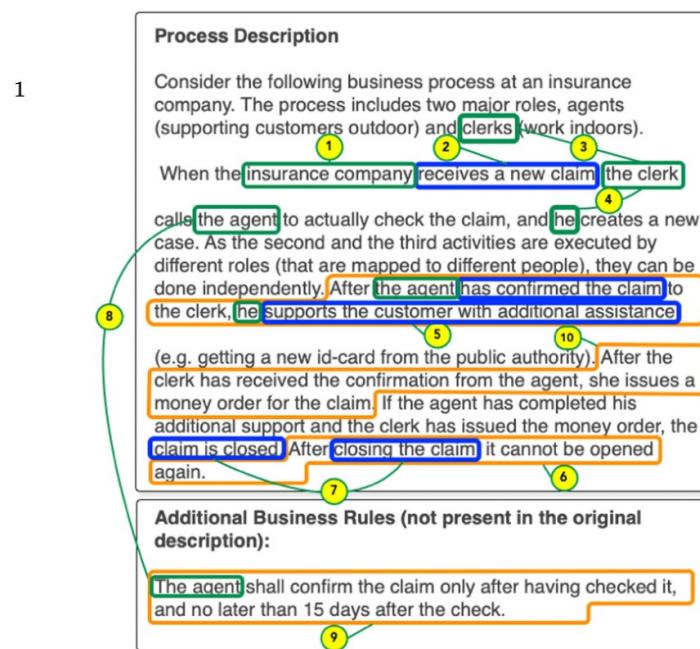
The process of a law?

Does my process comply with law?

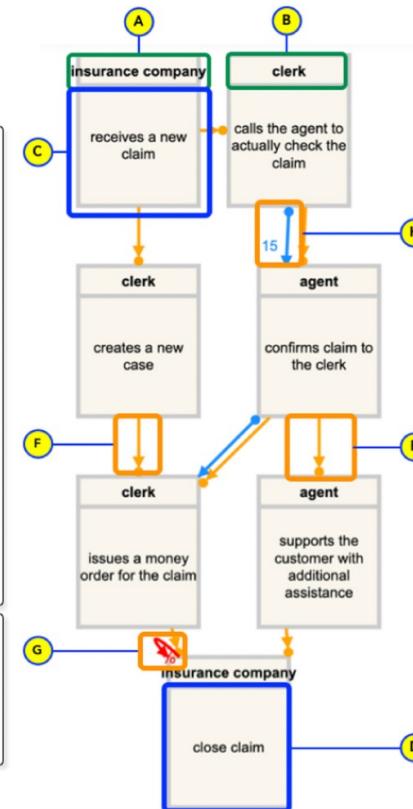
We will see a simple example (purchasing process)

However, PM has been applied to several domains, e.g., the **legal** one

Declarative Process Discovery: Linking Process and Textual Views



(a) Textual views



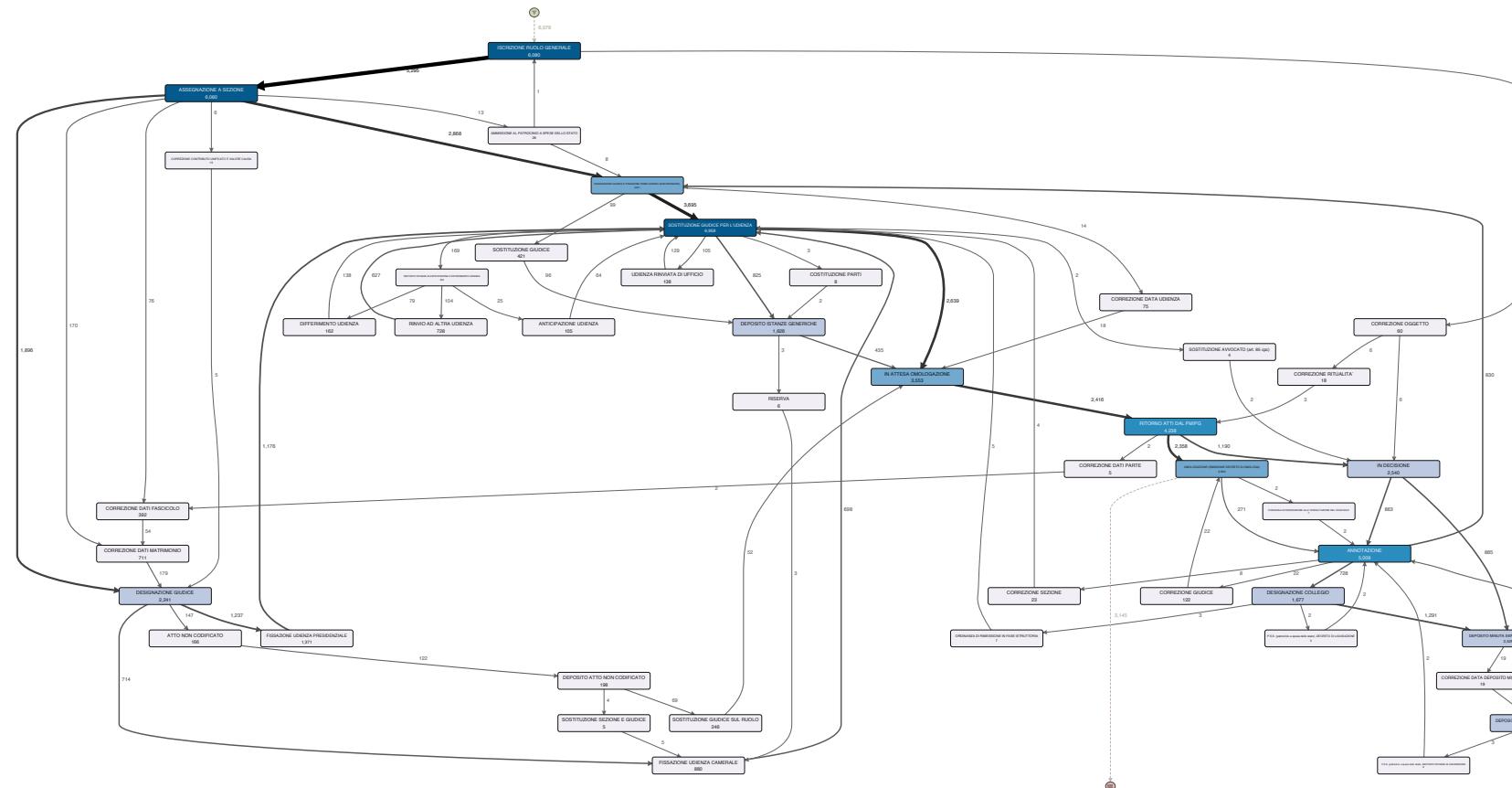
(b) Model view as a DCR graph

We will see a simple example (purchasing process)
However, PM has been applied to several domains, e.g., the **legal** one

How efficient is my law court?

Ongoing collaboration with

Vittoria Caponecchia, Daniele Licari, Andrea Vandini
Francesca Chiaromonte Giovanni Comande'

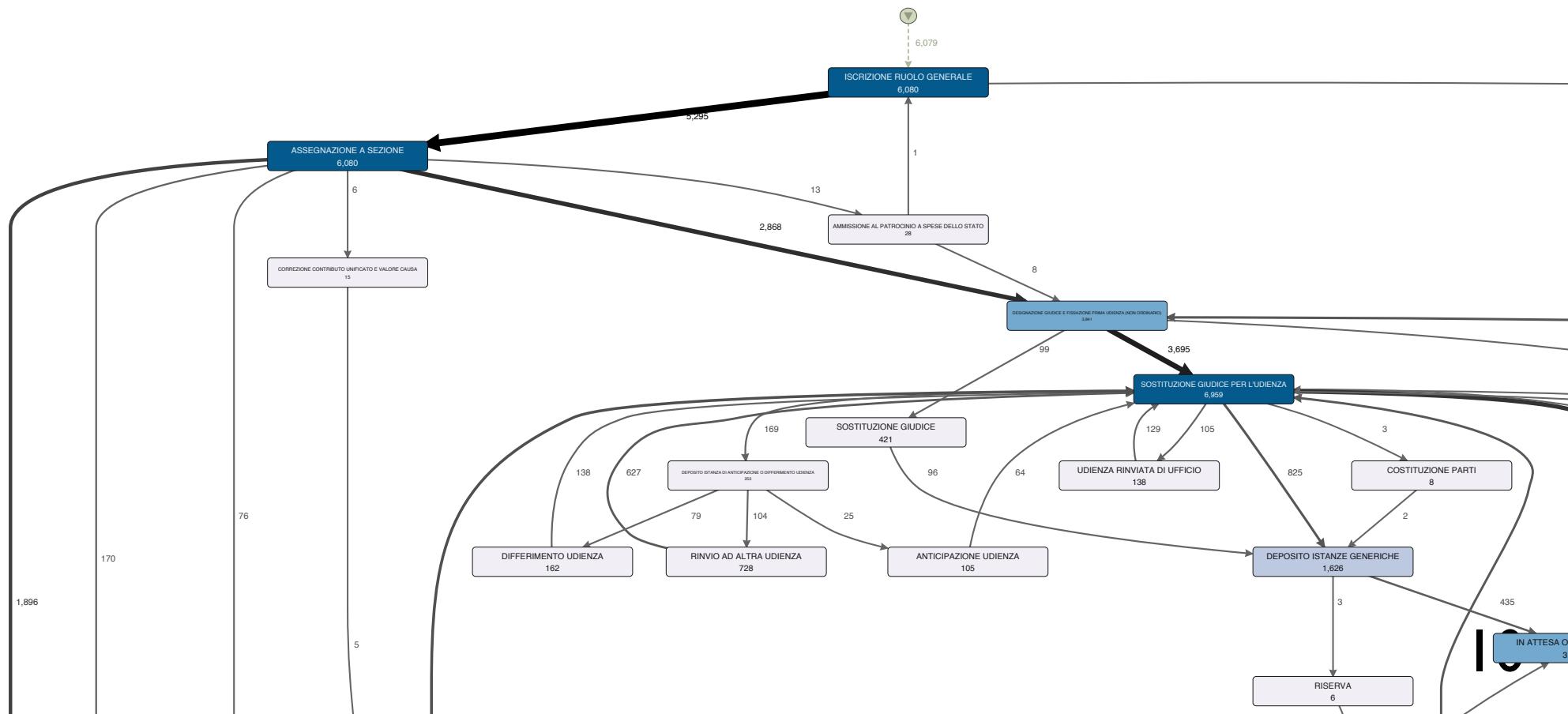


We will see a simple example (purchasing process)

However, PM has been applied to several domains, e.g., the **legal** one

How efficient is my law court?

Ongoing collaboration with
Vittoria Caponecchia, Daniele Licari, Andrea Vandin
Francesca Chiaromonte Giovanni Comande'



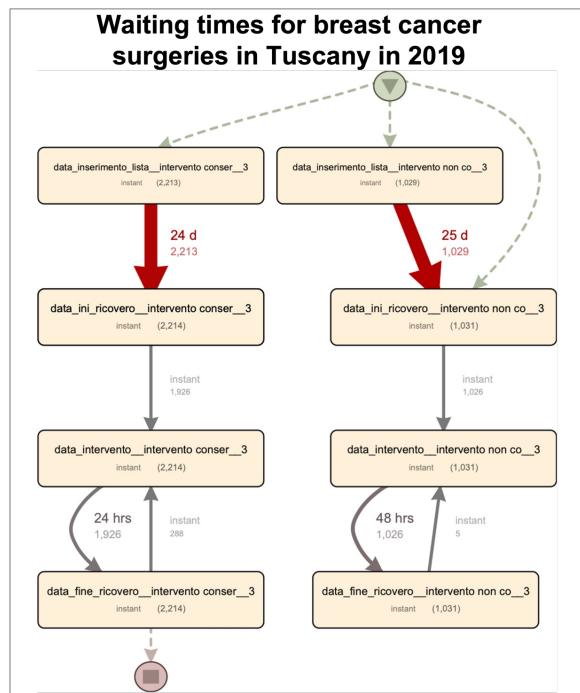
We will see a simple example (purchasing process)
 However, PM has been applied to several domains, e.g., the **healthcare** one

Process Mining and Clinical Pathways: an application to Breast cancer data in Tuscany

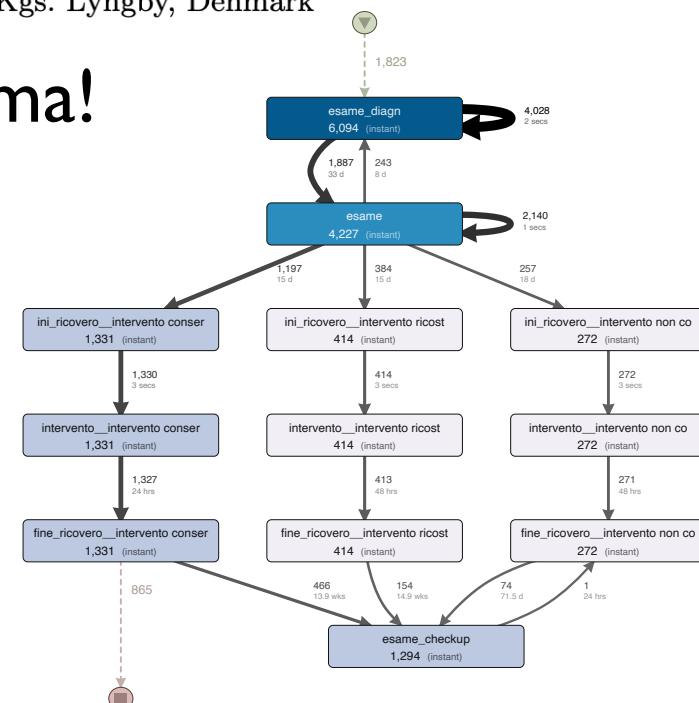
Francesca Ferrè¹, Chiara Seghieri¹, Andrea Burattin², and Andrea Vandin^{1,2}

¹ Sant'Anna School of Advanced Studies, Pisa, Italy

² DTU Technical University of Denmark, Kgs. Lyngby, Denmark



Now also Sima!



We will see a simple example (purchasing process)

However, PM has been applied to several domains, e.g., the **security** one

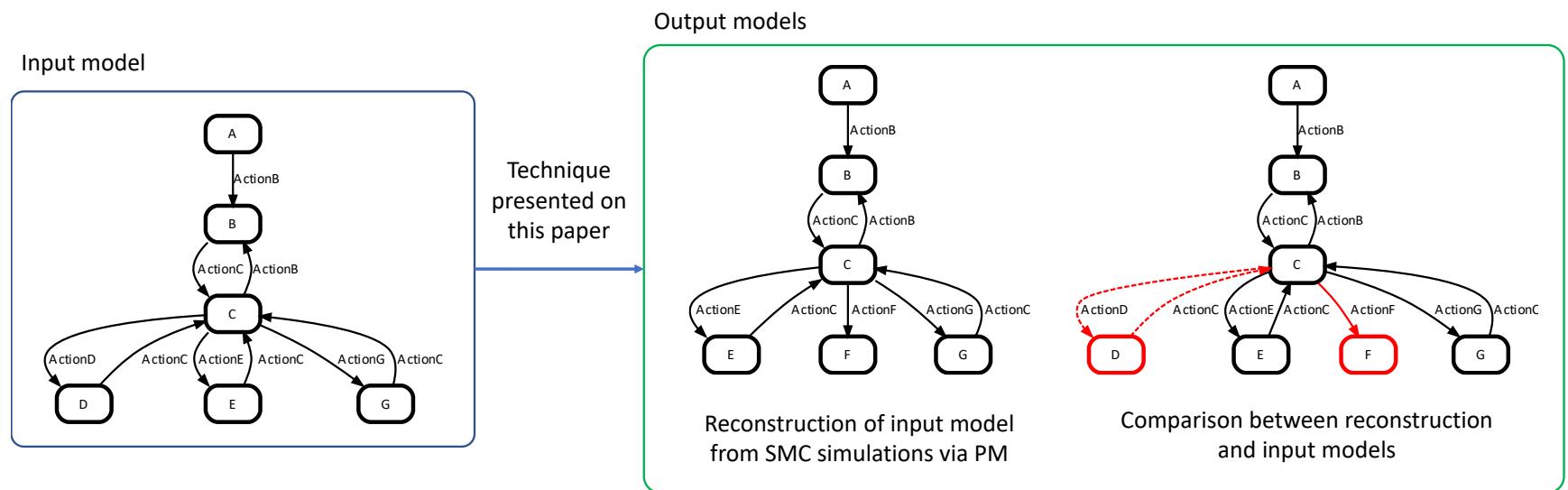
White-box validation of quantitative dynamic product lines by statistical model checking and process mining

Roberto Casaluce^a, Andrea Burattin^b, Francesca Chiaromonte^{a,c}, Alberto Lluch Lafuente^b, Andrea Vandin^{a,b,*}

^a*Institute of Economics and EMbeDS, Sant'Anna School of Advanced Studies, Pisa, Italy.*

^b*DTU Technical University of Denmark, Lyngby, Denmark.*

^c*Dept. of Statistics and Huck Institutes of the Life Sciences, Penn State University, USA*

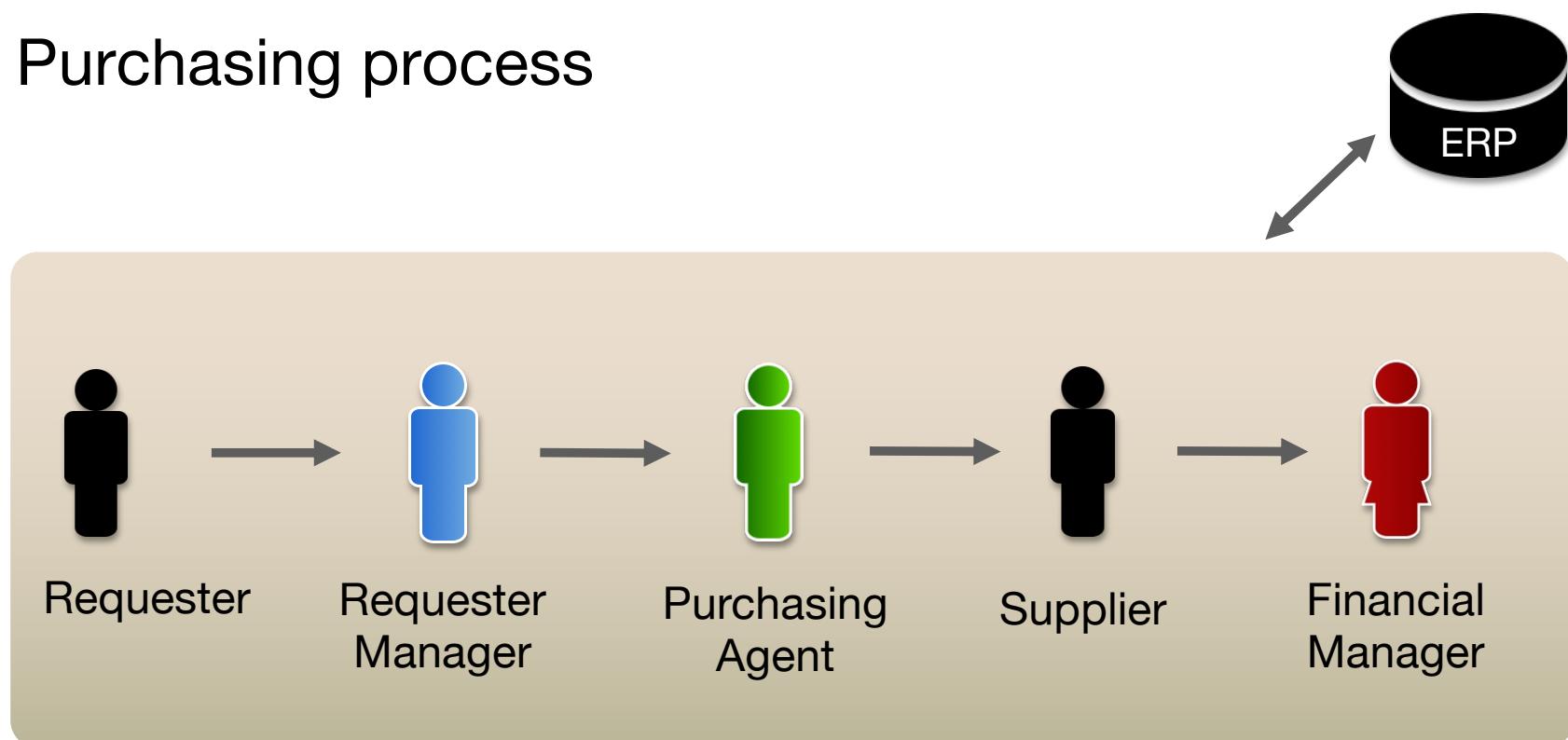


Outline

0. What is Process mining?
1. **Example Scenario**
2. Roadmap
3. Hands-on Session
4. Take-away Points

Example Scenario

Purchasing process



Problems

1. Inefficient operations
2. Need to demonstrate compliance
3. Complaints about process duration

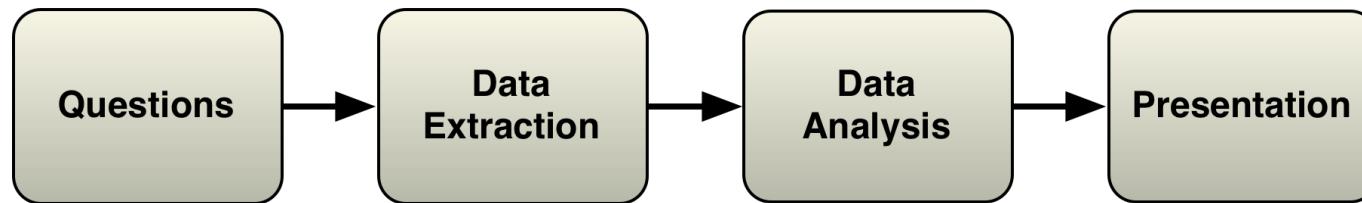
Analysis Goals

1. Understand the process in detail
2. Check whether there are deviations from the payment guidelines
3. Control performance targets
 - Complete process in at most 21 days

Outline

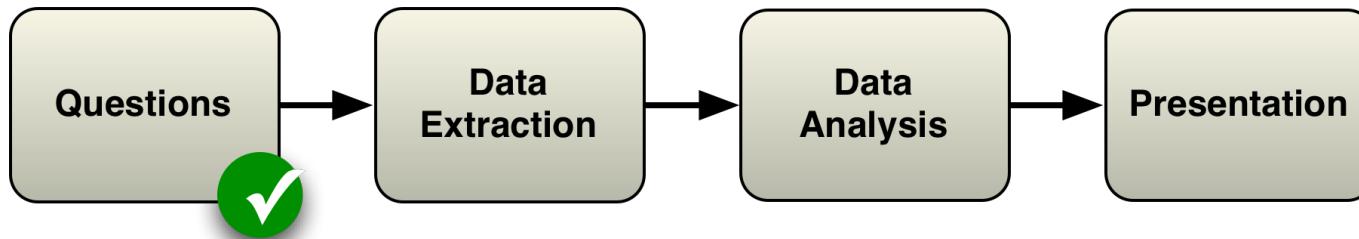
- 0. What is Process mining?
- 1. Example Scenario
- 2. Roadmap**
- 3. Hands-on Session
- 4. Take-away Points

Roadmap



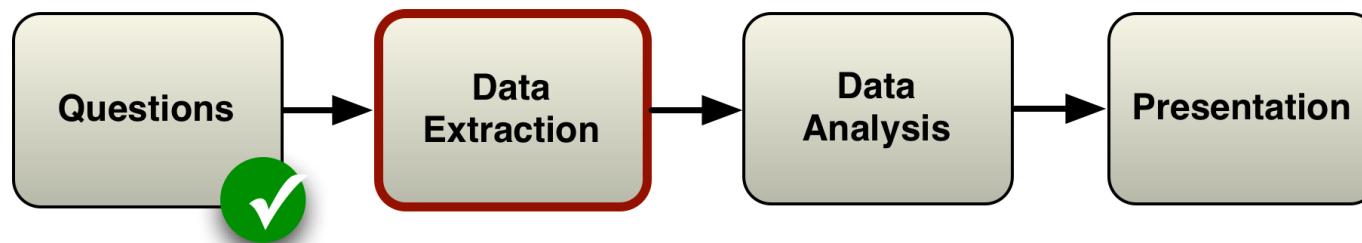
- | | | | |
|--------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none">• Determine questions• Process scope• Which IT systems | <ul style="list-style-type: none">• Via DB administrator• CSV file or database extract | <ul style="list-style-type: none">• Extract 'As-is' process• Answer questions | <ul style="list-style-type: none">• Present results (e.g., report, presentation, workshop etc.) |
|--------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|

Roadmap



1. How does the process actually look like?
2. Are there deviations from the prescribed process?
3. Do we meet the performance targets?

Roadmap

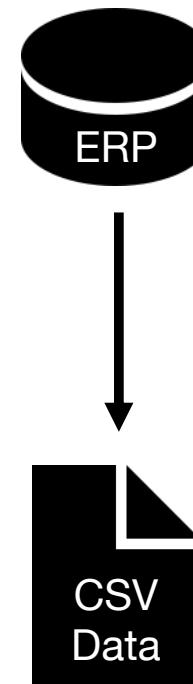


Data Extraction

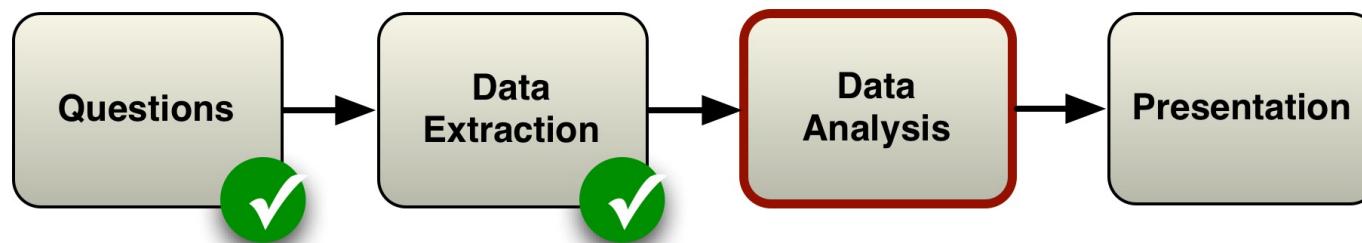
IT staff extracts history logs (from the Enterprise Resource Planning system, ERP) in a CSV file

Such CSV file is the starting point for our tutorial

- Please download it from our wiki (Slides page)
- Or here: [PurchaseExample.csv](#)

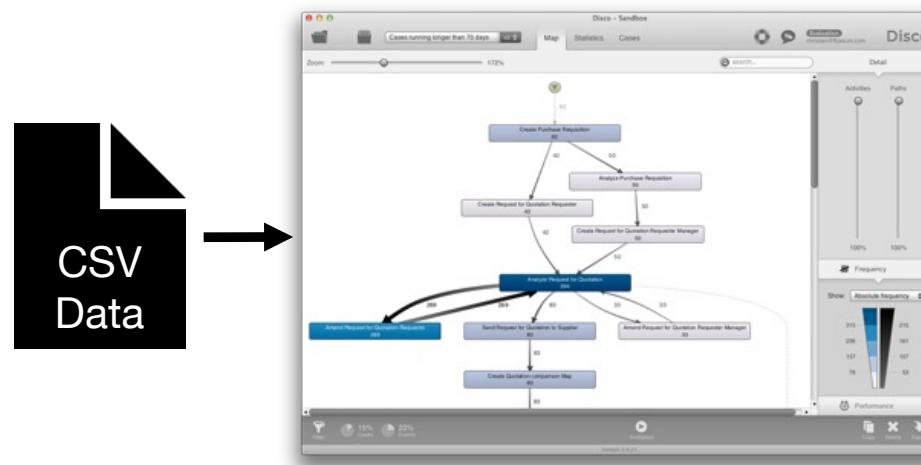


Roadmap



Data Analysis

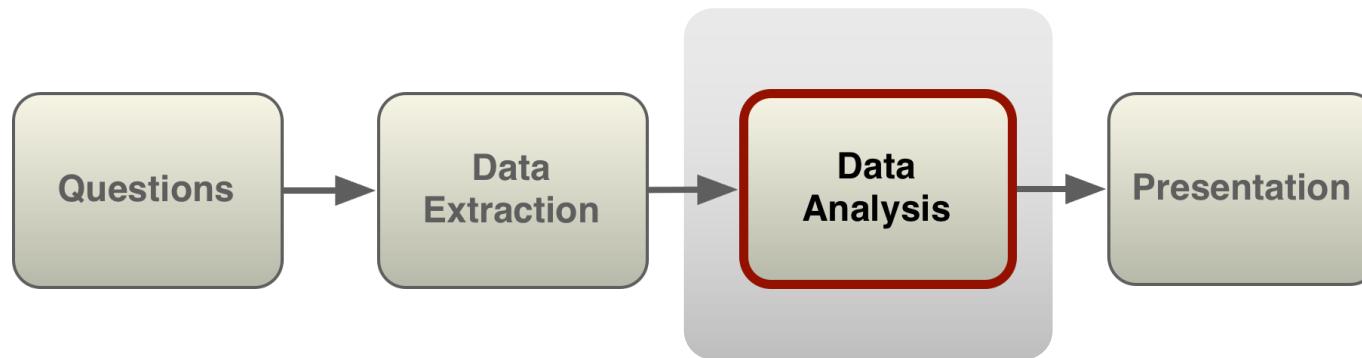
We use the process mining tool Disco to perform the data analysis



Download from fluxicon.com/disco

Sant'Anna is Academic Partner: free license with @santannapisa.it email

Roadmap



Focus of this session

Outline

0. What is Process mining?
1. Example Scenario
2. Roadmap
- 3. Hands-on Session**
4. Take-away Points

Hands-on Session

Let's get started!



Step 1 - Inspect Data

Open **PurchasingExample.csv** file in python/pandas (or Excel) and inspect its contents

- Every row corresponds to one event
- You find information on Case IDs, Activities, Start and end times, Resources, Roles

If using excel...



Screenshot of Microsoft Excel showing a CSV file named "PurchasingExample.csv". The table contains 36 rows of data with columns A through F.

	A	B	C	D	E	F
1	Case ID	Start Timestamp	Complete Timestamp	Activity	Resource	Role
2	1	2011/01/01 00:00:00.000	2011/01/01 00:37:00.000	Create Purchase Requisition	Kim Passa	Requester
3	2	2011/01/01 00:16:00.000	2011/01/01 00:29:00.000	Create Purchase Requisition	Immanuel Karagianni	Requester
4	3	2011/01/01 02:23:00.000	2011/01/01 03:03:00.000	Create Purchase Requisition	Kim Passa	Requester
5	1	2011/01/01 05:37:00.000	2011/01/01 05:45:00.000	Create Request for Quotation	Kim Passa	Requester
6	1	2011/01/01 06:41:00.000	2011/01/01 06:55:00.000	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
7	2	2011/01/01 08:16:00.000	2011/01/01 08:26:00.000	Create Request for Quotation	Alberto Duport	Requester
8	4	2011/01/01 08:39:00.000	2011/01/01 09:00:00.000	Create Purchase Requisition	Fjodor Kowalski	Requester
9	2	2011/01/01 09:34:00.000	2011/01/01 09:38:00.000	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
10	5	2011/01/01 09:49:00.000	2011/01/01 10:35:00.000	Create Purchase Requisition	Esmana Liubiata	Requester
11	2	2011/01/01 10:16:00.000	2011/01/01 10:21:00.000	Amend Request for Quotation	Christian Francois	Requester
12	2	2011/01/01 11:15:00.000	2011/01/01 11:48:00.000	Analyze Request for Quotation	Magdalena Predutta	Purchasing Agent
13	6	2011/01/01 11:20:00.000	2011/01/01 11:37:00.000	Create Purchase Requisition	Christian Francois	Requester
14	1	2011/01/01 11:43:00.000	2011/01/01 12:09:00.000	Send Request for Quotation to Supplier	Karel de Groot	Purchasing Agent
15	1	2011/01/01 12:32:00.000	2011/01/01 16:03:00.000	Create Quotation comparison Map	Magdalena Predutta	Purchasing Agent
16	2	2011/01/01 12:33:00.000	2011/01/01 12:39:00.000	Amend Request for Quotation	Esmana Liubiata	Requester
17	2	2011/01/01 13:28:00.000	2011/01/01 13:38:00.000	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
18	7	2011/01/01 14:05:00.000	2011/01/01 15:00:00.000	Create Purchase Requisition	Esmana Liubiata	Requester
19	8	2011/01/01 14:27:00.000	2011/01/01 15:17:00.000	Create Purchase Requisition	Fjodor Kowalski	Requester
20	2	2011/01/01 15:18:00.000	2011/01/01 15:40:00.000	Send Request for Quotation to Supplier	Francois de Perrier	Purchasing Agent
21	2	2011/01/01 15:55:00.000	2011/01/01 16:43:00.000	Create Quotation comparison Map	Karel de Groot	Purchasing Agent
22	9	2011/01/01 16:17:00.000	2011/01/01 16:34:00.000	Create Purchase Requisition	Tesca Lobes	Requester
23	6	2011/01/01 17:32:00.000	2011/01/01 17:45:00.000	Create Request for Quotation	Alberto Duport	Requester
24	8	2011/01/01 18:00:00.000	2011/01/01 18:07:00.000	Create Request for Quotation	Tesca Lobes	Requester
25	6	2011/01/01 18:39:00.000	2011/01/01 18:55:00.000	Analyze Request for Quotation	Magdalena Predutta	Purchasing Agent
26	4	2011/01/01 18:45:00.000	2011/01/01 18:51:00.000	Analyze Purchase Requisition	Maris Freeman	Requester Manager
27	4	2011/01/01 18:56:00.000	2011/01/01 18:58:00.000	Create Request for Quotation	Heinz Gutschmidt	Requester Manager
28	8	2011/01/01 19:04:00.000	2011/01/01 19:27:00.000	Analyze Request for Quotation	Francois de Perrier	Purchasing Agent
29	6	2011/01/01 19:47:00.000	2011/01/01 19:55:00.000	Amend Request for Quotation	Penn Osterwalder	Requester
30	4	2011/01/01 19:58:00.000	2011/01/01 20:19:00.000	Analyze Request for Quotation	Francois de Perrier	Purchasing Agent
31	8	2011/01/01 20:21:00.000	2011/01/01 20:34:00.000	Amend Request for Quotation	Tesca Lobes	Requester
32	6	2011/01/01 20:55:00.000	2011/01/01 21:28:00.000	Analyze Request for Quotation	Francois de Perrier	Purchasing Agent
33	4	2011/01/01 21:06:00.000	2011/01/01 21:14:00.000	Amend Request for Quotation	Nico Ojenbeer	Requester
34	8	2011/01/01 21:35:00.000	2011/01/01 22:01:00.000	Analyze Request for Quotation	Magdalena Predutta	Purchasing Agent
35	4	2011/01/01 22:01:00.000	2011/01/01 22:24:00.000	Analyze Request for Quotation	Francois de Perrier	Purchasing Agent
36	6	2011/01/01 22:13:00.000	2011/01/01 22:28:00.000	Amend Request for Quotation	Anne Olwada	Requester

If using pandas (python)...

```
logs.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 9119 entries, 0 to 9118
Data columns (total 6 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Case ID          9119 non-null    int64  
 1   Start Timestamp  9119 non-null    datetime64[ns]
 2   Complete Timestamp 9119 non-null    datetime64[ns]
 3   Activity         9119 non-null    object  
 4   Resource         9119 non-null    object  
 5   Role              9119 non-null    object  
dtypes: datetime64[ns](2), int64(1), object(3)
memory usage: 498.7+ KB
```

logs						
	Case ID	Start Timestamp	Complete Timestamp	Activity	Resource	Role
0	1	2011-01-01 00:00:00	2011-01-01 00:37:00	Create Purchase Requisition	Kim Passa	Requester
1	2	2011-01-01 00:16:00	2011-01-01 00:29:00	Create Purchase Requisition	Immanuel Karagianni	Requester
2	3	2011-01-01 02:23:00	2011-01-01 03:03:00	Create Purchase Requisition	Kim Passa	Requester
3	1	2011-01-01 05:37:00	2011-01-01 05:45:00	Create Request for Quotation	Kim Passa	Requester
4	1	2011-01-01 06:41:00	2011-01-01 06:55:00	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
...
9114	1284	2011-10-14 13:53:00	2011-10-14 14:07:00	Pay Invoice	Pedro Alvares	Financial Manager
9115	1448	2011-10-14 13:56:00	2011-10-14 14:24:00	Settle Dispute With Supplier	Karalda Nimwada	Financial Manager
9116	1941	2011-10-14 14:05:00	2011-10-14 14:18:00	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
9117	1448	2011-10-14 14:24:00	2011-10-14 14:24:00	Authorize Supplier's Invoice payment	Karalda Nimwada	Financial Manager
9118	1448	2011-10-14 15:16:00	2011-10-14 15:31:00	Pay Invoice	Karalda Nimwada	Financial Manager

If using pandas (python)...

```
logs.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 9119 entries, 0 to 9118
Data columns (total 6 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Case ID          9119 non-null    int64  
 1   Start Timestamp  9119 non-null    datetime64[ns]
 2   Complete Timestamp 9119 non-null    datetime64[ns]
 3   Activity         9119 non-null    object  
 4   Resource         9119 non-null    object  
 5   Role              9119 non-null    object  
dtypes: datetime64[ns](2), int64(1), object(3)
memory usage: 498.7+ KB
```

```
logs[logs["Case ID"]==1]
```

	Case ID	Start Timestamp	Complete Timestamp	Activity	Resource	Role
0	1	2011-01-01 00:00:00	2011-01-01 00:37:00	Create Purchase Requisition	Kim Passa	Requester
3	1	2011-01-01 05:37:00	2011-01-01 05:45:00	Create Request for Quotation	Kim Passa	Requester
4	1	2011-01-01 06:41:00	2011-01-01 06:55:00	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
12	1	2011-01-01 11:43:00	2011-01-01 12:09:00	Send Request for Quotation to Supplier	Karel de Groot	Purchasing Agent
13	1	2011-01-01 12:32:00	2011-01-01 16:03:00	Create Quotation comparison Map	Magdalena Predutta	Purchasing Agent
36	1	2011-01-01 22:44:00	2011-01-01 23:13:00	Analyze Quotation Comparison Map	Immanuel Karagianni	Requester
37	1	2011-01-01 23:13:00	2011-01-01 23:13:00	Choose best option	Tesca Lobes	Requester
42	1	2011-01-02 01:22:00	2011-01-02 09:20:00	Settle Conditions With Supplier	Francois de Perrier	Purchasing Agent
60	1	2011-01-02 09:58:00	2011-01-02 10:10:00	Create Purchase Order	Karel de Groot	Purchasing Agent
68	1	2011-01-02 14:09:00	2011-01-02 14:43:00	Confirm Purchase Order	Sean Manney	Supplier
97	1	2011-01-02 20:49:00	2011-01-03 03:37:00	Deliver Goods Services	Sean Manney	Supplier
137	1	2011-01-03 11:20:00	2011-01-03 11:21:00	Release Purchase Order	Elvira Lores	Requester
176	1	2011-01-03 19:09:00	2011-01-03 19:10:00	Approve Purchase Order for payment	Karel de Groot	Purchasing Agent
221	1	2011-01-04 00:54:00	2011-01-04 00:54:00	Send Invoice	Kiu Kan	Supplier
276	1	2011-01-04 15:08:00	2011-01-04 15:13:00	Release Supplier's Invoice	Karalda Nimwada	Financial Manager
277	1	2011-01-04 15:13:00	2011-01-04 15:13:00	Authorize Supplier's Invoice payment	Karalda Nimwada	Financial Manager
278	1	2011-01-04 15:22:00	2011-01-04 15:31:00	Pay Invoice	Pedro Alvares	Financial Manager

If using pandas (python)...

```
logs.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 9119 entries, 0 to 9118
Data columns (total 6 columns):
 #   Column           Non-Null Count   Dtype  
--- 
 0   Case ID          9119 non-null    int64  
 1   Start Timestamp  9119 non-null    datetime64[ns]
 2   Complete Timestamp 9119 non-null    datetime64[ns]
```

```
logs[logs["Case ID"]==2]
```

	Case ID	Start Timestamp	Complete Timestamp	Activity	Resource	Role
1	2	2011-01-01 00:16:00	2011-01-01 00:29:00	Create Purchase Requisition	Immanuel Karagianni	Requester
5	2	2011-01-01 08:16:00	2011-01-01 08:26:00	Create Request for Quotation	Alberto Duport	Requester
7	2	2011-01-01 09:34:00	2011-01-01 09:38:00	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
9	2	2011-01-01 10:16:00	2011-01-01 10:21:00	Amend Request for Quotation	Christian Francois	Requester
10	2	2011-01-01 11:15:00	2011-01-01 11:48:00	Analyze Request for Quotation	Magdalena Predutta	Purchasing Agent
14	2	2011-01-01 12:33:00	2011-01-01 12:39:00	Amend Request for Quotation	Esmana Liubiata	Requester
15	2	2011-01-01 13:28:00	2011-01-01 13:38:00	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
18	2	2011-01-01 15:18:00	2011-01-01 15:40:00	Send Request for Quotation to Supplier	Francois de Perrier	Purchasing Agent
19	2	2011-01-01 15:55:00	2011-01-01 16:43:00	Create Quotation comparison Map	Karel de Groot	Purchasing Agent
39	2	2011-01-01 23:33:00	2011-01-01 23:44:00	Analyze Quotation Comparison Map	Esmana Liubiata	Requester
40	2	2011-01-01 23:44:00	2011-01-01 23:44:00	Choose best option	Fjodor Kowalski	Requester
47	2	2011-01-02 02:15:00	2011-01-02 15:07:00	Settle Conditions With Supplier	Magdalena Predutta	Purchasing Agent
77	2	2011-01-02 16:21:00	2011-01-02 16:31:00	Create Purchase Order	Magdalena Predutta	Purchasing Agent
94	2	2011-01-02 20:23:00	2011-01-02 20:29:00	Confirm Purchase Order	Esmeralda Clay	Supplier
123	2	2011-01-03 03:15:00	2011-01-04 14:53:00	Deliver Goods Services	Esmeralda Clay	Supplier
312	2	2011-01-04 23:01:00	2011-01-04 23:02:00	Release Purchase Order	Elvira Lores	Requester
342	2	2011-01-05 09:03:00	2011-01-05 09:04:00	Approve Purchase Order for payment	Karel de Groot	Purchasing Agent
367	2	2011-01-05 14:50:00	2011-01-05 14:50:00	Send Invoice	Esmeralda Clay	Supplier
409	2	2011-01-06 05:04:00	2011-01-06 05:07:00	Release Supplier's Invoice	Pedro Alvares	Financial Manager
411	2	2011-01-06 05:38:00	2011-01-06 05:58:00	Settle Dispute With Supplier	Karalda Nimwada	Financial Manager
413	2	2011-01-06 05:58:00	2011-01-06 05:58:00	Authorize Supplier's Invoice payment	Pedro Alvares	Financial Manager
415	2	2011-01-06 06:21:00	2011-01-06 06:32:00	Pay Invoice	Karalda Nimwada	Financial Manager

Step 2 - Import Data

Load **PurchasingExample.csv** in Disco

Assign columns as follows:

- Case ID → Case ID
- Start and Complete Timestamp → Timestamp
- Activity → Activity
- Resource → Resource
- Role → Other

Click ‘Start import’

Disco – Tutorial

Case ID column is used

X Case    

	Case ID	Start Timestamp	Complete Timestamp	Activity	Resource	Role
1	1	2011/01/01 00:00:00.000	2011/01/01 00:37:00.000	Create Purchase Requisition	Kim Passa	Requester
2	2	2011/01/01 00:16:00.000	2011/01/01 00:29:00.000	Create Purchase Requisition	Immanuel Karagianni	Requester
3	3	2011/01/01 02:23:00.000	2011/01/01 03:03:00.000	Create Purchase Requisition	Kim Passa	Requester
4	1	2011/01/01 05:37:00.000	2011/01/01 05:45:00.000	Create Request for Quotation	Kim Passa	Requester
5	1	2011/01/01 06:41:00.000	2011/01/01 06:55:00.000	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
6	2	2011/01/01 08:16:00.000	2011/01/01 08:26:00.000	Create Request for Quotation	Alberto Duport	Requester
7	4	2011/01/01 08:39:00.000	2011/01/01 09:00:00.000	Create Purchase Requisition	Fjodor Kowalski	Requester
8	2	2011/01/01 09:34:00.000	2011/01/01 09:38:00.000	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
9	5	2011/01/01 09:49:00.000	2011/01/01 10:35:00.000	Create Purchase Requisition	Esmana Liubiata	Requester
10	2	2011/01/01 10:16:00.000	2011/01/01 10:21:00.000	Amend Request for Quotation	Christian Francois	Requester
11	2	2011/01/01 11:15:00.000	2011/01/01 11:48:00.000	Analyze Request for Quotation	Magdalena Predutta	Purchasing Agent
12	6	2011/01/01 11:20:00.000	2011/01/01 11:37:00.000	Create Purchase Requisition	Christian Francois	Requester
13	1	2011/01/01 11:43:00.000	2011/01/01 12:09:00.000	Send Request for Quotation to Supplier	Karel de Groot	Purchasing Agent
14	1	2011/01/01 12:32:00.000	2011/01/01 16:03:00.000	Create Quotation comparison Map	Magdalena Predutta	Purchasing Agent
15	2	2011/01/01 12:33:00.000	2011/01/01 12:39:00.000	Amend Request for Quotation	Esmana Liubiata	Requester
16	2	2011/01/01 13:28:00.000	2011/01/01 13:38:00.000	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
17	7	2011/01/01 14:05:00.000	2011/01/01 15:00:00.000	Create Purchase Requisition	Esmana Liubiata	Requester
18	8	2011/01/01 14:27:00.000	2011/01/01 15:17:00.000	Create Purchase Requisition	Fjodor Kowalski	Requester
19	2	2011/01/01 15:18:00.000	2011/01/01 15:40:00.000	Send Request for Quotation to Supplier	Francois de Perrier	Purchasing Agent
20	2	2011/01/01 15:55:00.000	2011/01/01 16:43:00.000	Create Quotation comparison Map	Karel de Groot	Purchasing Agent
21	9	2011/01/01 16:17:00.000	2011/01/01 16:34:00.000	Create Purchase Requisition	Tesca Lobes	Requester
22	6	2011/01/01 17:32:00.000	2011/01/01 17:45:00.000	Create Request for Quotation	Alberto Duport	Requester
23	8	2011/01/01 18:00:00.000	2011/01/01 18:07:00.000	Create Request for Quotation	Tesca Lobes	Requester
24	6	2011/01/01 18:39:00.000	2011/01/01 18:55:00.000	Analyze Request for Quotation	Magdalena Predutta	Purchasing Agent
25	4	2011/01/01 18:45:00.000	2011/01/01 18:51:00.000	Analyze Purchase Requisition	Maris Freeman	Requester Manager
26	4	2011/01/01 18:56:00.000	2011/01/01 18:58:00.000	Create Request for Quotation	Heinz Gutschmidt	Requester Manager
27	8	2011/01/01 19:04:00.000	2011/01/01 19:27:00.000	Analyze Request for Quotation	Francois de Perrier	Purchasing Agent
28	6	2011/01/01 19:47:00.000	2011/01/01 19:55:00.000	Amend Request for Quotation	Penn Osterwalder	Requester
29	4	2011/01/01 19:59:00.000	2011/01/01 20:10:00.000	Analyze Request for Quotation	Francois de Perrier	Purchasing Agent

Cancel File encoding: UTF-8 Use quotes  Ready to start import. Start import

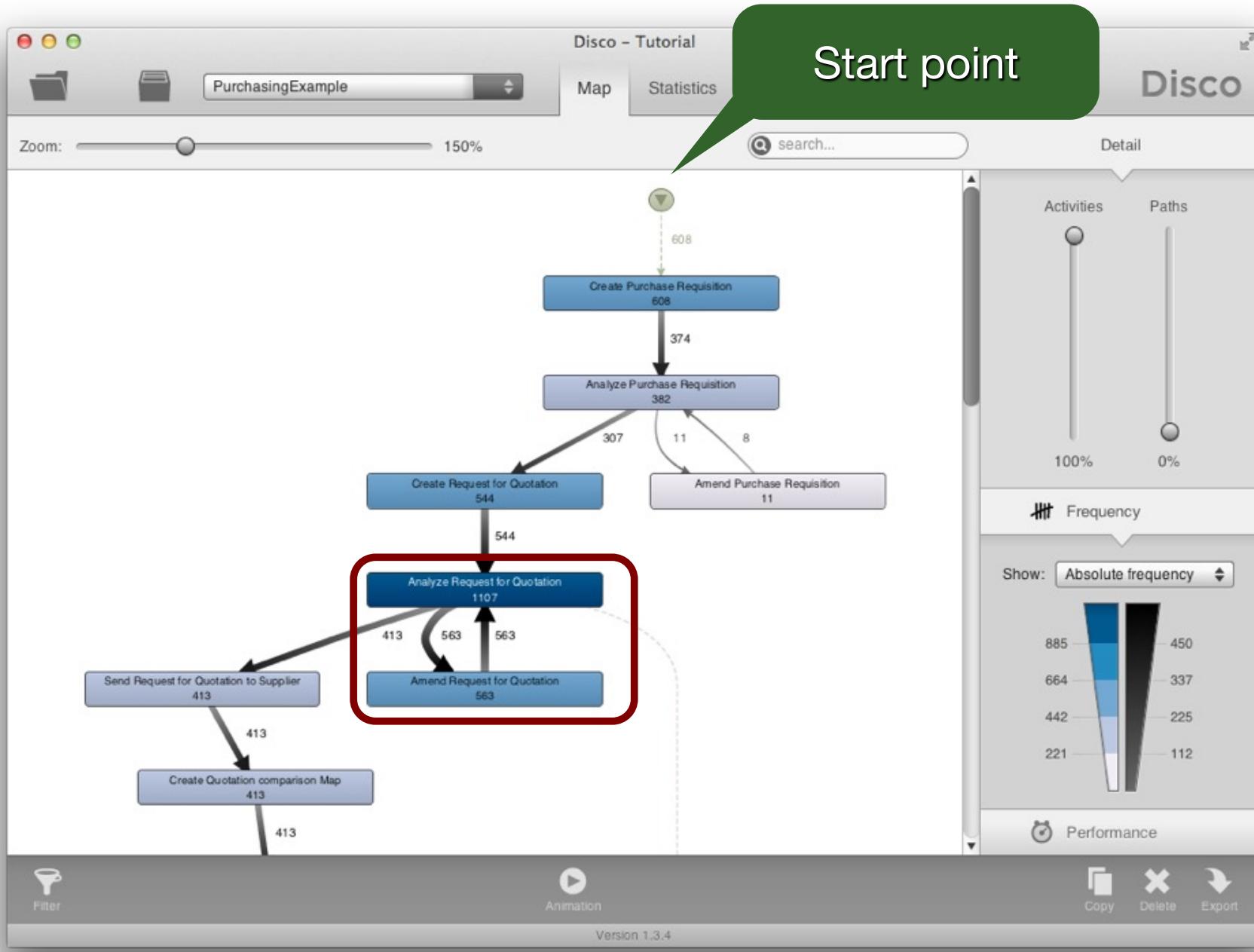
Step 3 - Inspect Process

Look at the resulting process model

- Numbers in rectangles are activity frequencies
- Number at arcs is frequency of connection

→ You see the main process flows

- All 608 cases start with activity ‘Create Purchase Requisition’
- Lots of changes were made (amendments)!



Step 3 - Inspect Process

It's important to be able to adjust the level of detail for the process map

Move up the 'Activities' slider down to lowest position (0%)

- Only the activities from the most frequent process variant are shown

Disco – Tutorial

PurchasingExample

Map Statistics Cases Enterprise anne@fluxicon.com Disco

Zoom: 49% search...

Activities Paths

Frequency

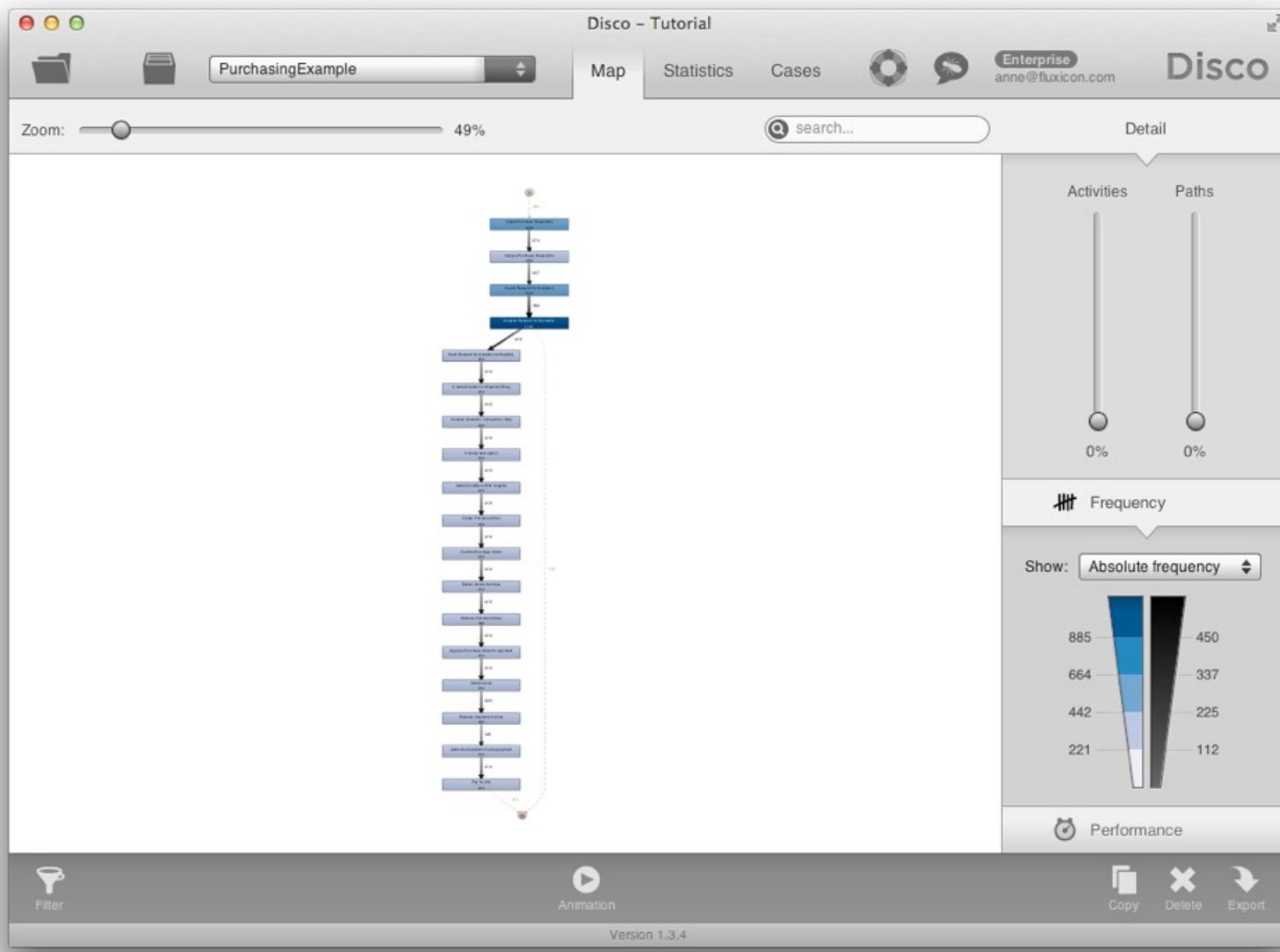
Show: Absolute frequency

Performance

Filter Animation

Version 1.3.4

Copy Delete Export



```

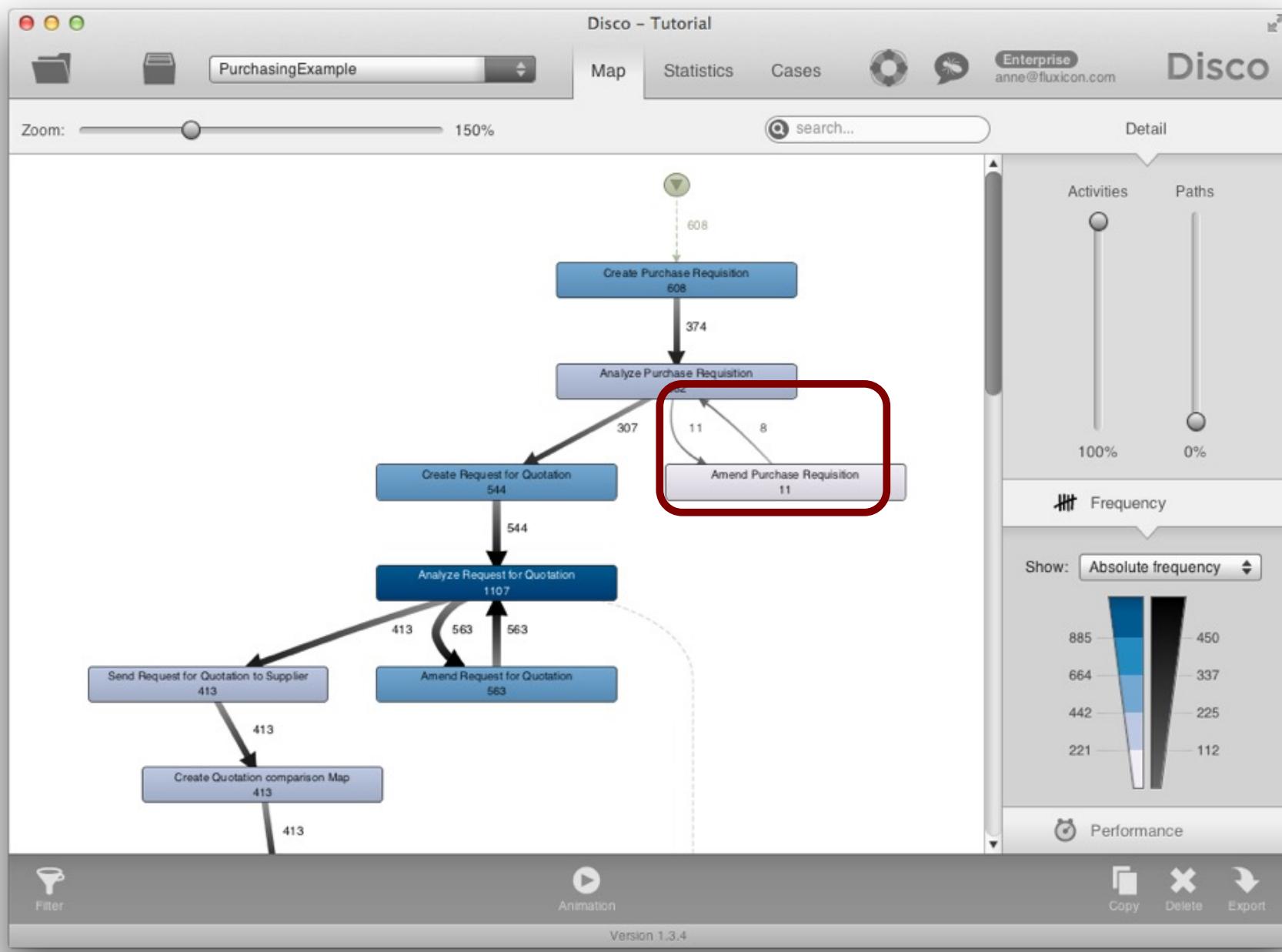
graph TD
    A[Initiate New Purchase] --> B[Identify New Supplier]
    B --> C[Check Supplier Availability]
    C --> D[Place Purchase Order]
    D --> E[Check Purchase Status]
    E --> F[Check Supplier Availability]
    F --> G[Place Purchase Order]
    G --> H[Check Purchase Status]
    H --> I[Check Supplier Availability]
    I --> J[Place Purchase Order]
    J --> K[Check Purchase Status]
    K --> L[Check Supplier Availability]
    L --> M[Place Purchase Order]
    M --> N[Check Purchase Status]
    N --> O[Check Supplier Availability]
    O --> P[Place Purchase Order]
    P --> Q[Check Purchase Status]
    Q --> R[Deliver Goods]
  
```

Step 3 - Inspect Process

Gradually move the 'Activities' slider up to 100% again until all activities are shown

- Even infrequent activities such as 'Amend Purchase Requisition' are shown

You'll notice that 11 cases are flowing in to 'Amend Purchase Requisition' but only 8 are moving out - Where are the other 3?

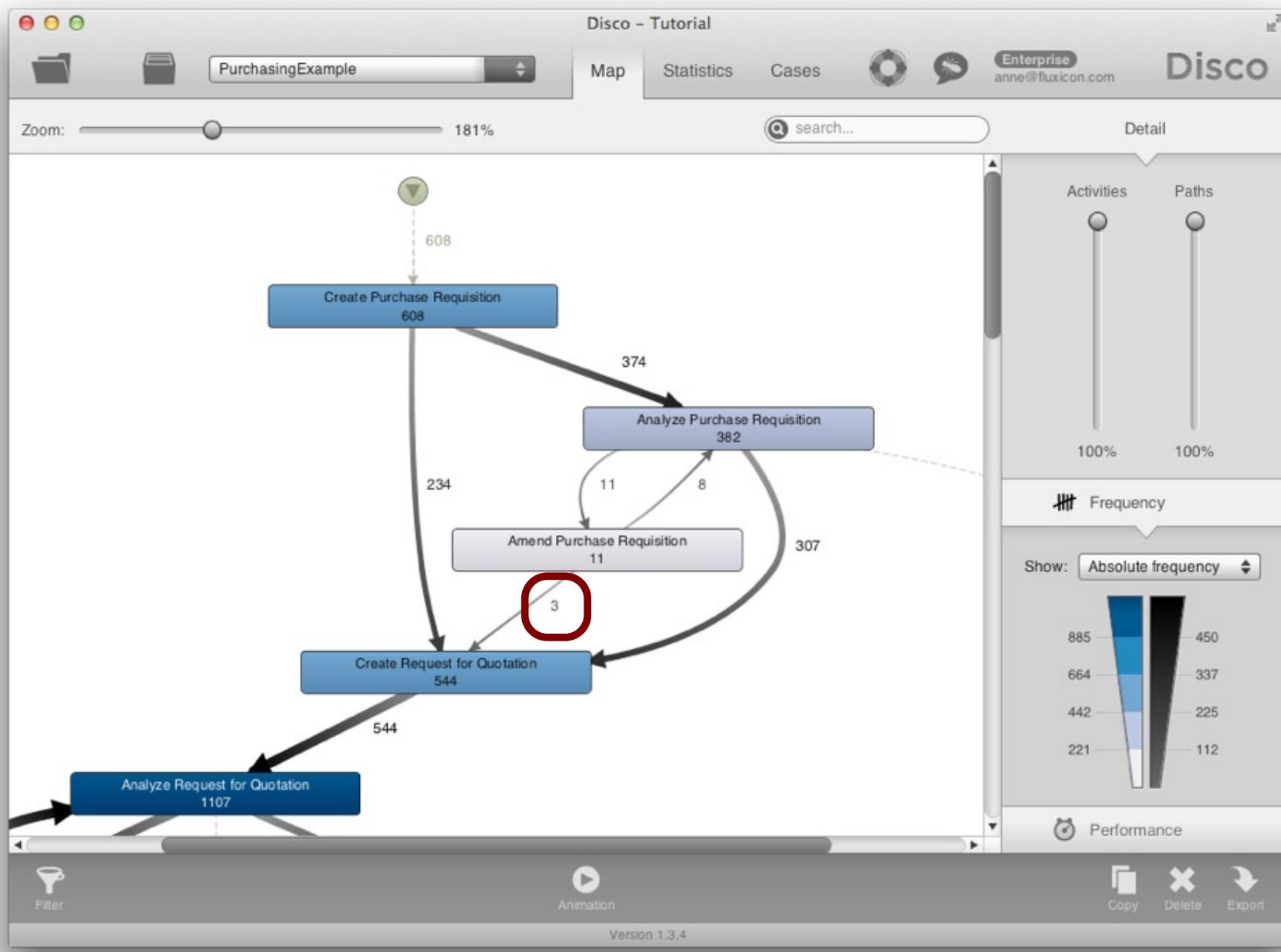


Step 3 - Inspect Process

Move up the ‘Paths’ slider up to the top

You now see a 100% detailed picture of the executed process

- The 3 missing cases move from ‘Amend Purchase Requisition’ to ‘Create Request for Quotation’



Step 4 - Inspect Statistics

Look at ‘Statistics’ tab to see ‘Overview’ information about the event log

- 9,119 events were recorded for 608 cases
- Timeframe is January - October 2011

The ‘Case duration’ is typically up to 15 or 16 days, but some cases take very long (more than 70 or 80 days!)

Disco – Tutorial

PurchasingExample Map Statistics Cases

Enterprise anne@fluxicon.com Disco

Statistics views

Overview Global statistics

- Activity Activity classes
- Resource Resource classes
- Role Other attribute

Events over time Active cases over time Case variants Events per case **Case duration** Case utilization Mean activity duration Mean waiting time

Number of cases Case duration

up to 81 days, 8 hours Number of cases: 9

	Events	Started	Finished	Duration
1	17	01.01.2011 00:00:00	04.01.2011 15:31:00	3 days, 15 hours
2	22	01.01.2011 00:16:00	06.01.2011 06:32:00	5 days, 6 hours
3	18	01.01.2011 02:23:00	07.01.2011 16:31:00	6 days, 14 hours
4	19	01.01.2011 08:39:00	06.01.2011 21:32:00	5 days, 12 hours
5	19	01.01.2011 09:49:00	06.01.2011 23:04:00	5 days, 13 hours
6	21	01.01.2011 11:20:00	06.01.2011 13:50:00	5 days, 2 hours
7	22	01.01.2011 14:05:00	07.01.2011 09:53:00	5 days, 19 hours
8	19	01.01.2011 14:27:00	07.01.2011 16:04:00	6 days, 1 hour
9	15	01.01.2011 16:17:00	04.01.2011 05:56:00	2 days, 13 hours
10	18	02.01.2011 02:52:00	10.01.2011 16:16:00	8 days, 13 hours
11	6	02.01.2011 05:12:00	03.01.2011 19:57:00	1 day, 14 hours
12	25	02.01.2011 05:31:00	10.01.2011 16:02:00	8 days, 10 hours
13	25	02.01.2011 07:47:00	10.01.2011 06:49:00	7 days, 23 hours

Cases (608) Variants (70)

Case ID Events Started Finished Duration

Filter Copy Delete Export

Version 1.3.4

Step 5 - Inspect Cases

Select ‘Cases’ tab to inspect variants and individual service instances

- The third most frequent process variant ends after ‘Analyze Purchase Requisition’ (ca. 10.36% of all cases follow this pattern)

→ Why are so many requests abrupt?

- Do people not know what they can buy?
- We can find this back in the process map, too

Disco – Tutorial

PurchasingExample

Map Statistics Cases

Enterprise anne@fluxicon.com Disco

search...

Variants (70) Cases (63)

423 Case with 2 events

Events 2
Start 02.03.2011 07:36:00
Duration 12 hours, 42 mins
Active time 8.4 %

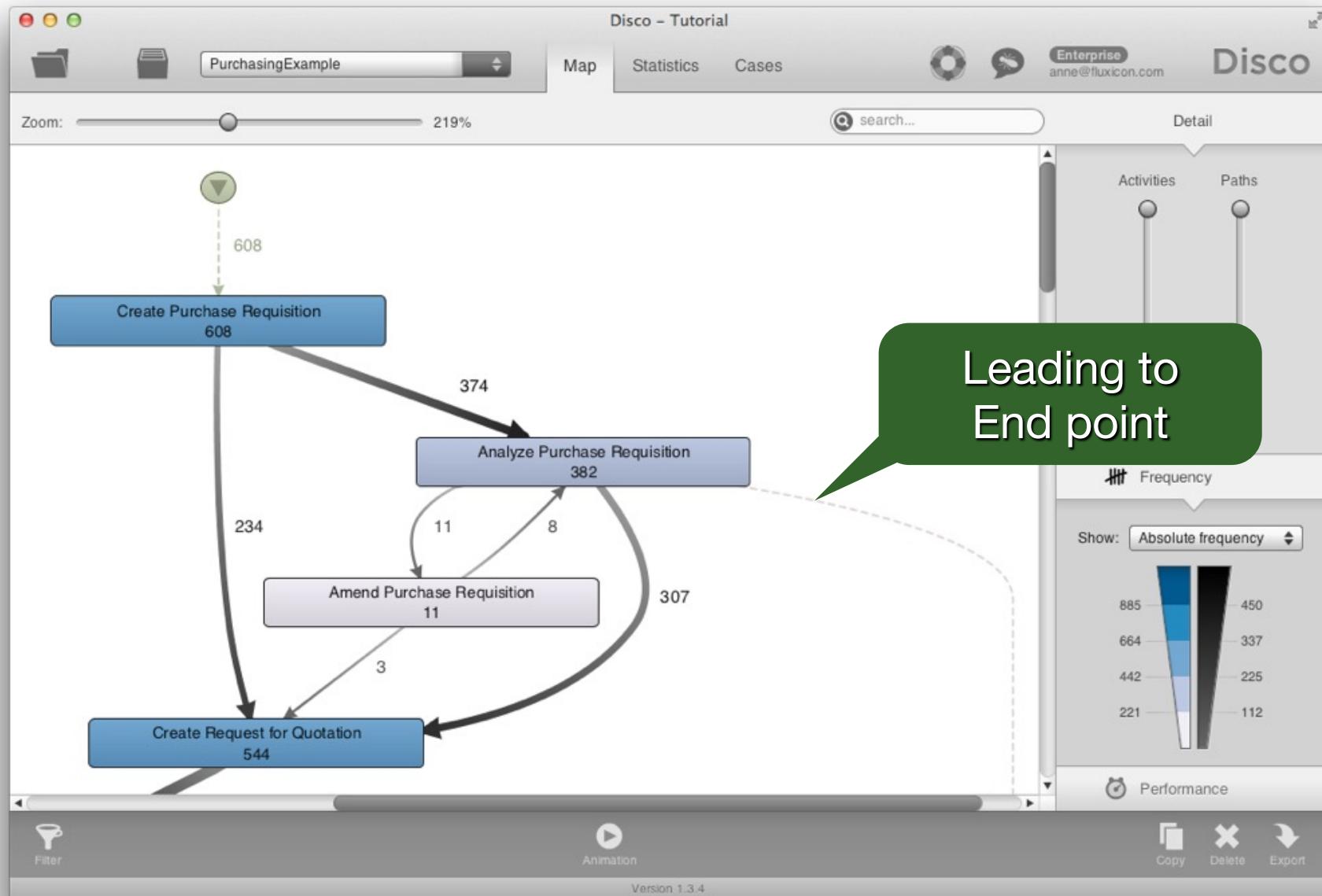
Graph Table

	Activity	Resource	Date	Time	Duration	Role
1	Create Purchase Requisition	Clement Duchot	02.03.2011	07:36:00	55 mins	Requester
2	Analyze Purchase Requisition	Heinz Gutschmidt	02.03.2011	20:09:00	9 mins	Requester

Filter

Version 1.3.4

Copy Delete Export



Results so far...

Original Questions:

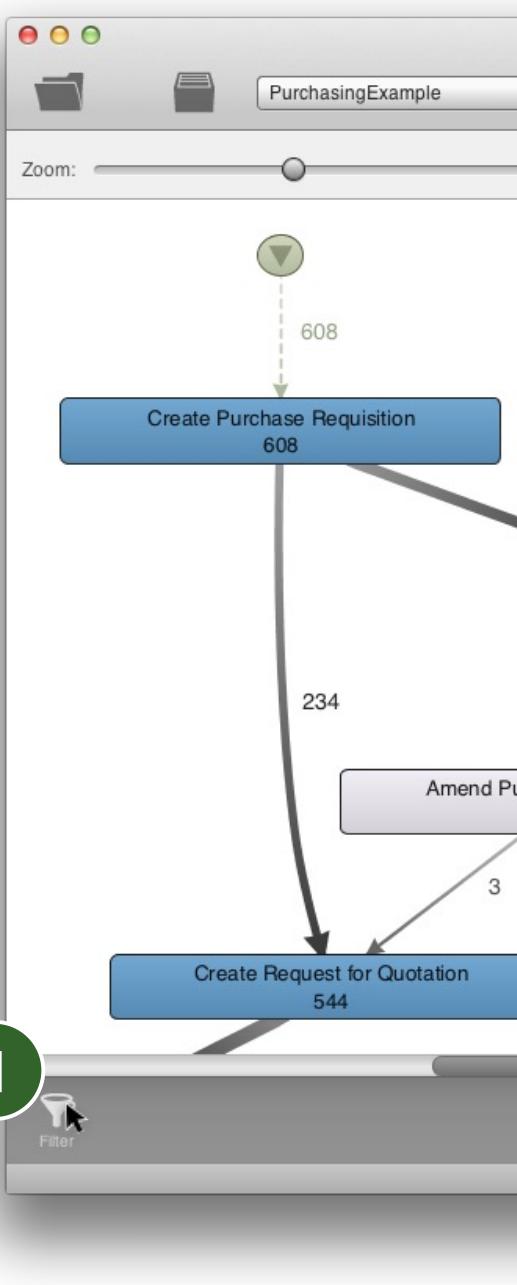
- ➊ 1. How does the process actually look like?
 - Objective process map discovered
 - Lots of amendments and stopped requests:
Update of purchasing guidelines needed
- ➋ 2. Are there deviations from the prescribed process?
- ➌ 3. Do we meet the performance targets?
 - Not by all (some take longer than 21 days):
Is there a bottleneck in the process? -> Next

Step 6 - Filter on Performance

Click on the Filter symbol in the lower left corner and add a Performance filter

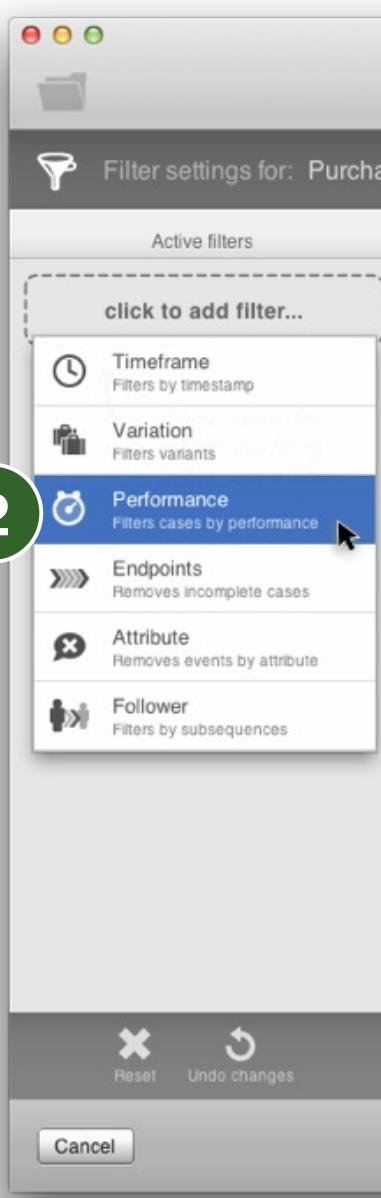
- Select 21 days as lower boundary
- You'll see that ca. 15% of the purchase orders take longer than 21 days

Press 'Apply filter' to focus only on those cases that take longer than 21 days



1

The screenshot shows a process flow diagram with three main steps: 'Create Purchase Requisition' (608), 'Amend Pu' (234), and 'Create Request for Quotation' (544). A green circle labeled '1' highlights the 'Filter' button at the bottom left of the visualization window.



2

The screenshot shows the 'Filter settings for: PurchasingExample' dialog. It lists several filter types under 'Active filters':

- click to add filter...
- Timeframe (Filters by timestamp)
- Variation (Filters variants)
- Performance** (Filters cases by performance) - This option is selected and highlighted with a blue background and a mouse cursor icon.
- Endpoints (Removes incomplete cases)
- Attribute (Removes events by attribute)
- Follower (Filters by subsequences)

At the bottom of the dialog are 'Reset' and 'Undo changes' buttons, and a 'Cancel' button.

Log filter

In this view you can config
are an essential tool for c
your event log data for an

On the left, you can build
Here are some suggestio
your log:

Variation filter

Your log has a hig
complex process
variation.

Performance filter

The cases in your
help you focus yo
criteria.

Disco – Tutorial

Filter settings for: PurchasingExample

Active filters

Performance
Filters cases by performance

click to add filter...

Performance
Filters cases by performance

Filter cases by: Case duration



Short cases Long-running cases

21 days 109 days, 9 hours

Minimum duration Maximum duration

15% of cases

Use cases running longer than 21 days.

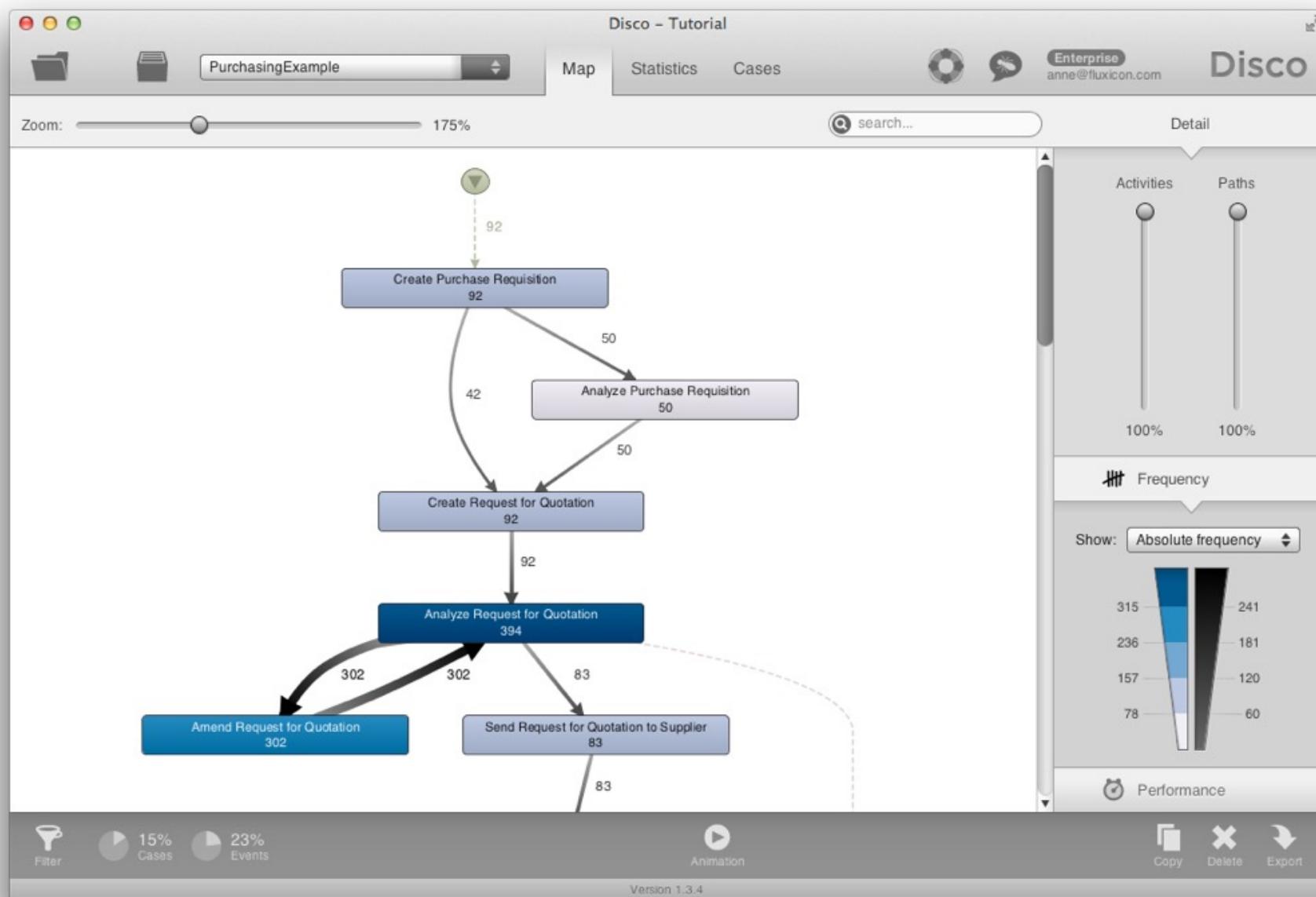
Reset Undo changes

Cancel Copy and filter Apply filter

Step 7 - Spot Bottlenecks

The filtered process map shows the process flow for the 92 (15%) ‘slow’ cases

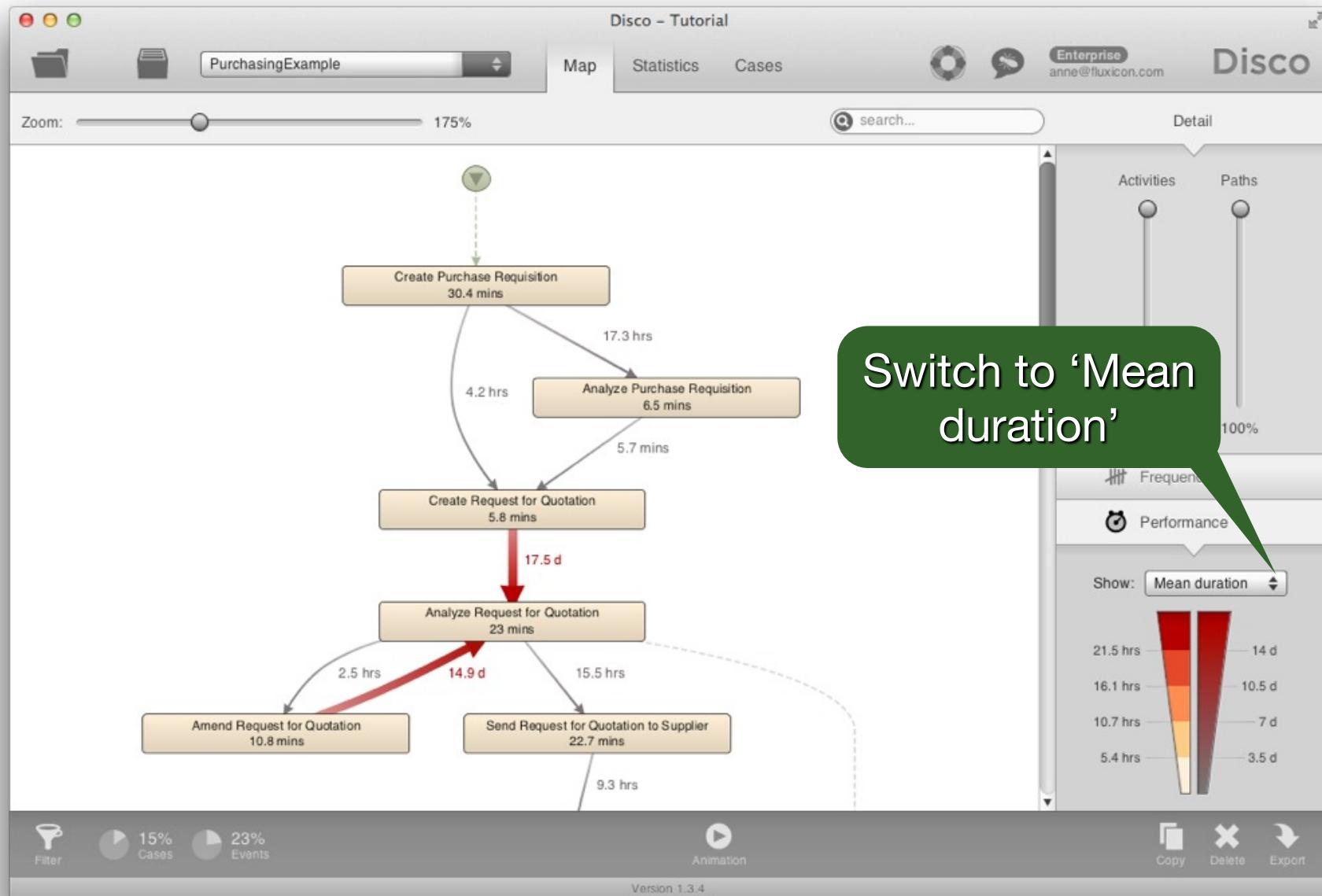
- On average 3 amendments per case!
 - 92 cases, 302 amendments...



Step 7b - Spot Bottlenecks

Switch to ‘Performance’ view

- ‘Total duration’ shows the high-impact areas
- Switch to ‘Mean duration’:
 - On average it takes **more than 14 days** to return from the rework loop to the normal process
 - What about min and max duration?



Step 8 - Animate Process

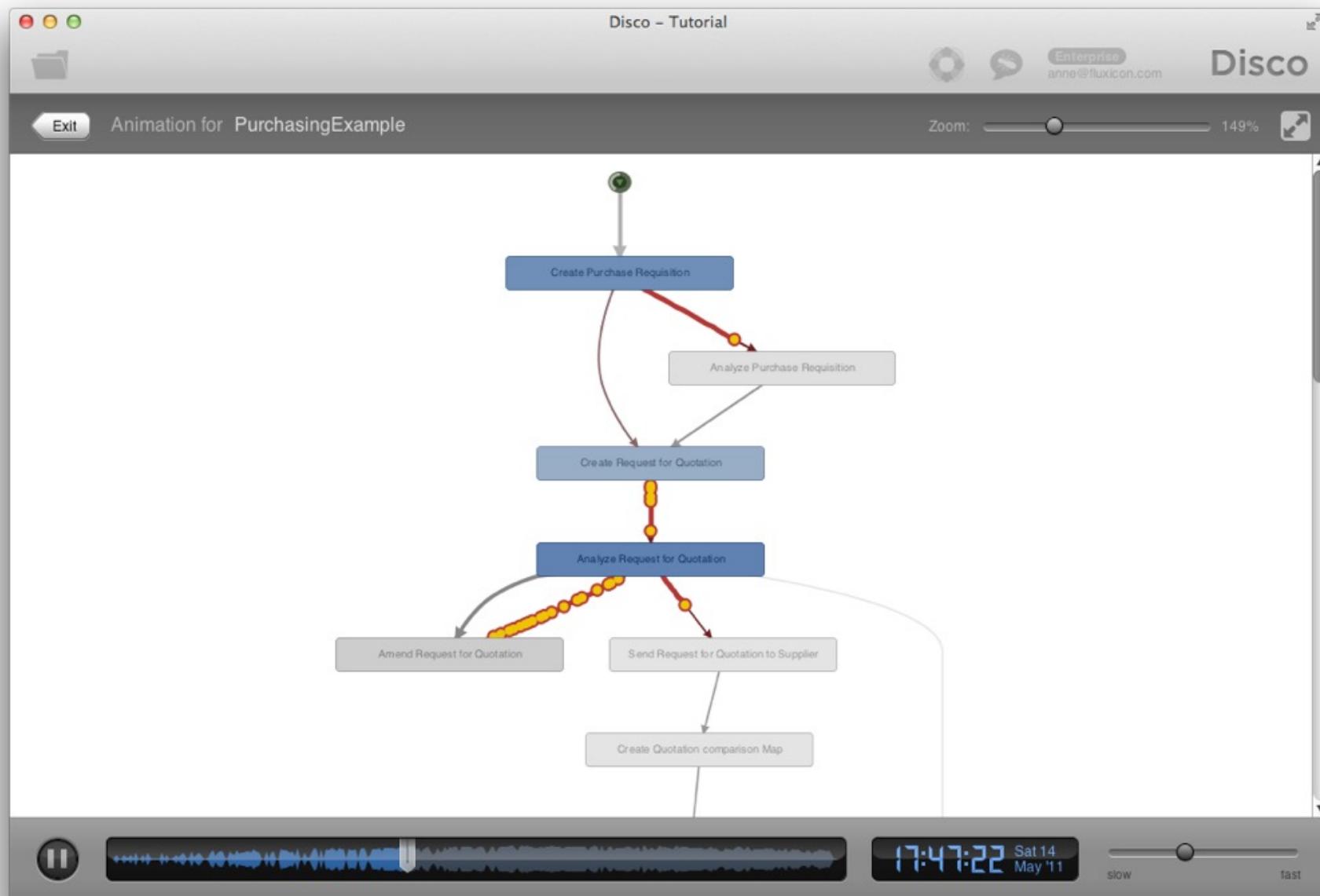
Visualize bottleneck:

Press ▶ button to start animation

Observe how purchase orders move through the process

Drag needle to the end of the timeline

- observe how the most used paths get thicker and thicker

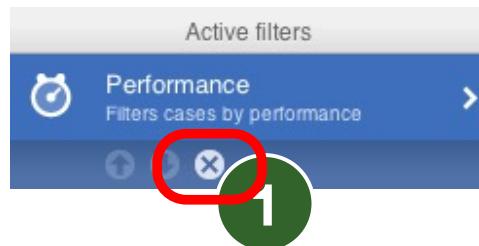


Results so far...

- ➊ 1. How does the process actually look like?
 - Objective process map discovered
 - Lots of amendments and stopped requests:
Update of purchasing guidelines needed
- ➋ 2. Are there deviations from the prescribed process? -> **Next**
- ➌ 3. Do we meet the performance targets?
 - Not by all (some take longer than 21 days)
 - The 'Analyze Request for Quotation' activity is a huge bottleneck: **Process change is needed**

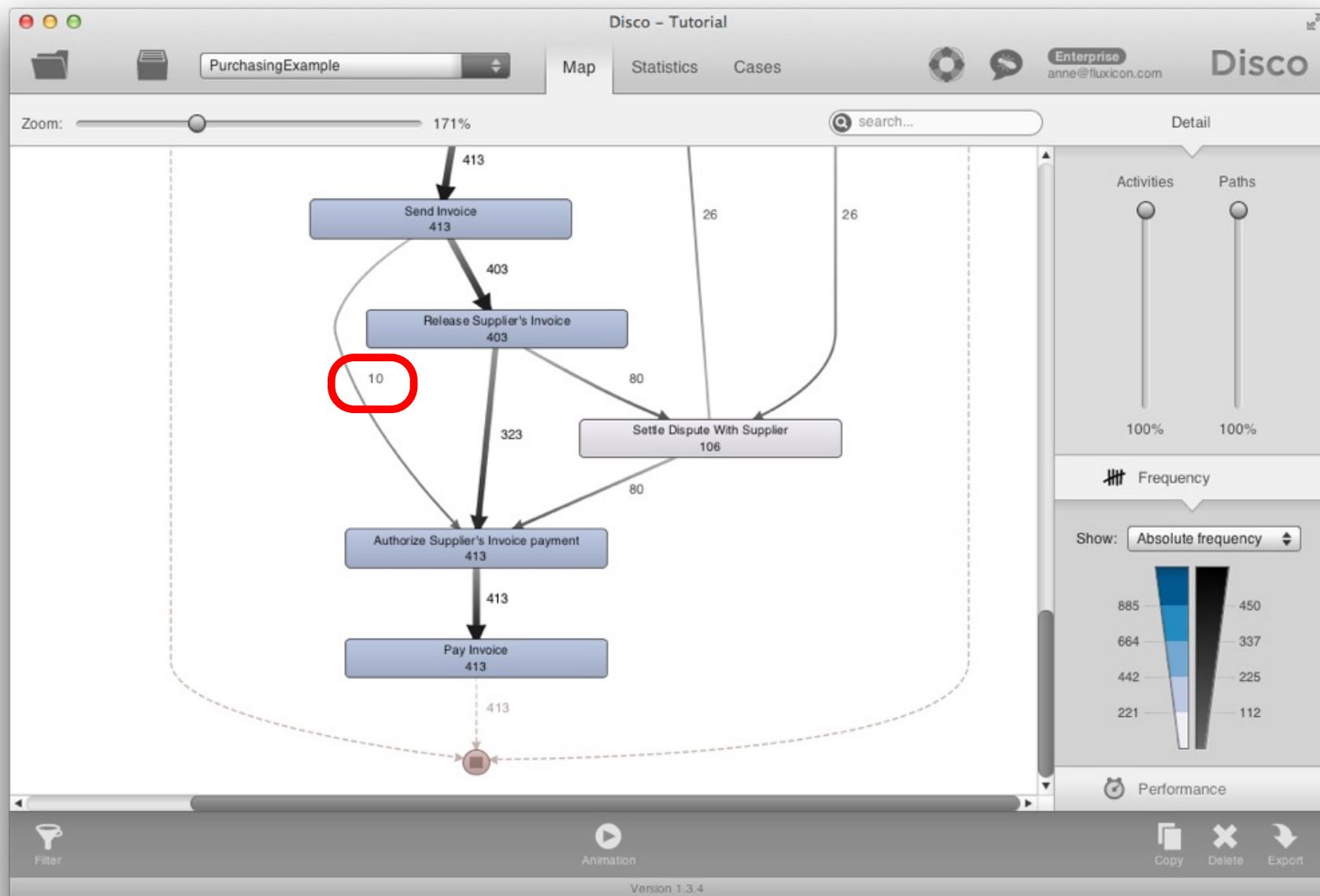
Step 9 - Compliance Check

Exit the animation, return to Filter settings, and remove performance filter



Switch back to Frequency Map view and scroll to end of the process

- 10 cases skip the mandatory ‘Release Supplier’s Invoice’ activity!

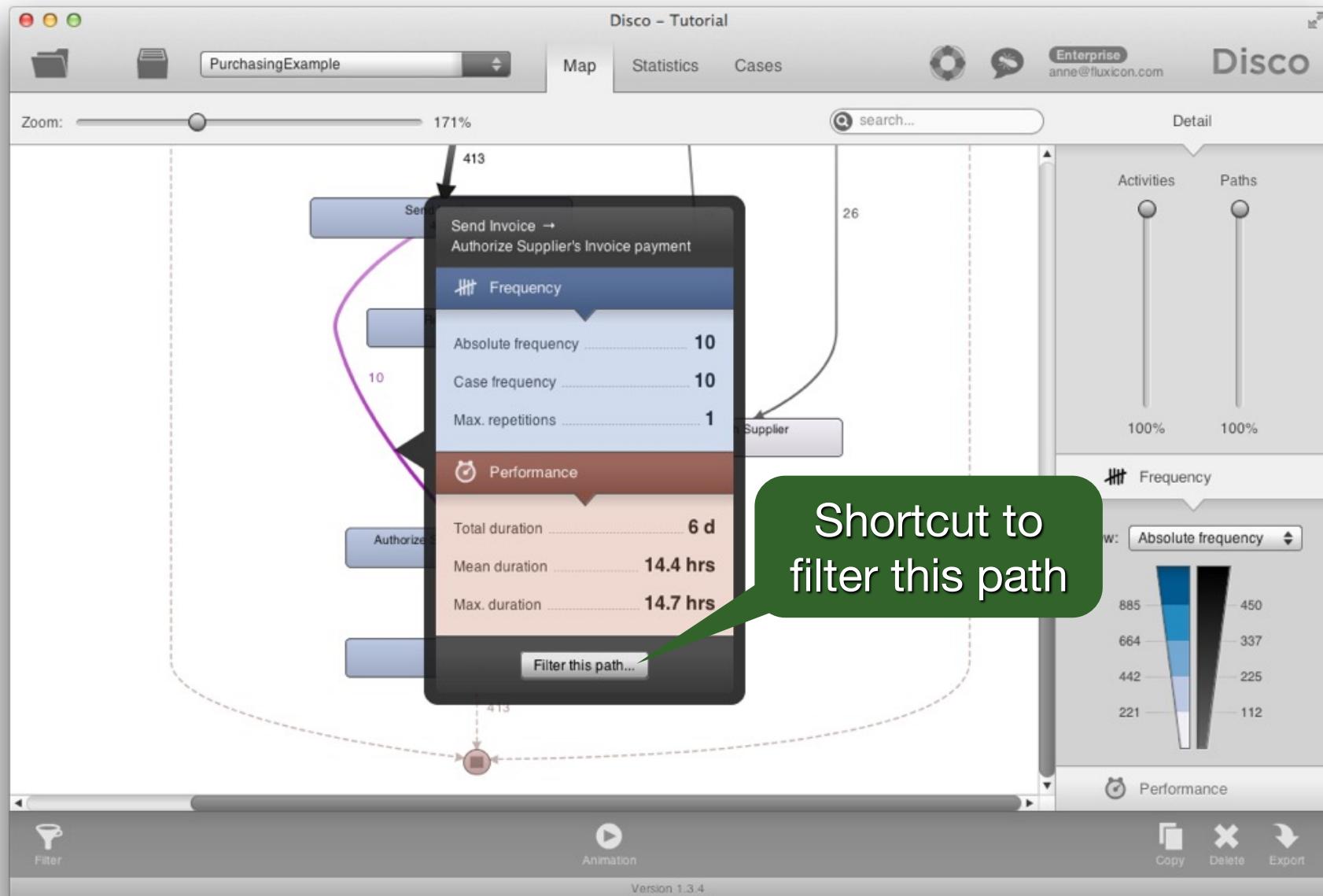


Step 9 - Compliance Check

Drill down: Click on the path from ‘Send invoice’ to ‘Authorize Supplier’s Invoice payment’ and press ‘Filter this path...’

Switch to Cases view to see the 10 cases

- Actionable result: We can either change the operational system to prevent the violation or provide targeted training



Disco – Tutorial

PurchasingExample

Map Statistics Cases

Enterprise anne@fluxicon.com Disco

search...

Variants (6)

- Complete log All cases (10) >
- Variant 1 3 cases (30%) >
- Variant 2 2 cases (20%) >
- Variant 3 2 cases (20%) >
- Variant 4 1 case (10%) >
- Variant 5 1 case (10%) >
- Variant 6 1 case (10%) >

Cases (10)

	Cases	Events
4	19 events	
5	19 events	
150	17 events	
170	17 events	
182	25 events	
413	17 events	
424	18 events	
547	21 events	
812	16 events	
940	16 events	

170 Case with 17 events

Events 17
Start 22.01.2011 22:35:00
Duration 11 days, 7 hours
Active time 19.32 %

Graph Table

Activity	Resource	Date	Time	Duration	Role
1 Create Purchase Requisition	Miu Hanwan	22.01.2011	22:35:00	45 mins	Requester
2 Analyze Purchase Requisition	Francis Odell	23.01.2011	15:11:00	5 mins	Requester
3 Create Request for Quotation	Heinz Gutschmidt	23.01.2011	15:21:00	2 mins	Requester
4 Analyze Request for Quotation	Francois de Perrier	25.01.2011	23:48:00	37 mins	Purchasing
5 Send Request for Quotation to Supplier	Karel de Groot	26.01.2011	02:45:00	21 mins	Purchasing
6 Create Quotation comparison Map	Karel de Groot	26.01.2011	03:25:00	3 hours, 22 mins	Purchasing
7 Analyze Quotation Comparison Map	Esmara Liubiata	26.01.2011	13:28:00	32 mins	Requester
8 Choose best option	Tesca Lopes	26.01.2011	14:00:00	0 millis	Requester
9 Settle Conditions With Supplier	Karel de Groot	26.01.2011	20:26:00	14 hours, 1 min	Purchasing
10 Create Purchase Order	Karel de Groot	27.01.2011	22:40:00	10 mins	Purchasing
11 Confirm Purchase Order	Esmerala Clay	28.01.2011	07:12:00	7 mins	Supplier
12 Deliver Goods Services	Esmerala Clay	31.01.2011	00:41:00	1 day, 8 hours	Supplier
13 Release Purchase Order	Kim Passa	01.02.2011	16:35:00	1 min	Requester
14 Approve Purchase Order for payment	Magdalena Predutta	02.02.2011	00:15:00	1 min	Purchasing
15 Send Invoice	Sean Manney	02.02.2011	14:17:00	0 millis	Supplier
16 Authorize Supplier's Invoice payment	Karalda Nimwada	03.02.2011	04:57:00	0 millis	Financial M
17 Pay Invoice	Pedro Alvares	03.02.2011	05:49:00	6 mins	Financial M

Filter 1% Cases 2% Events

Copy Delete Export

Version 1.3.4

Results so far...

- ➊ 1. How does the process actually look like?
 - Objective process map discovered
 - Lots of amendments and stopped requests:
Update of purchasing guidelines needed
- ➋ 2. Are there deviations from the prescribed process? -> **Yes, training or system change needed**
- ➌ 3. Do we meet the performance targets?
 - Not by all (some take longer than 21 days)
 - The 'Analyze Request for Quotation' activity is a huge bottleneck: **Process change is needed**

Step 10 - Organizational View

Last Step: We seek an alternative view on the data to visualize the organizational flow

Go to ‘Project view’ and press ‘Reload’:



Set ‘Activity’ column to ‘Other’ and configure ‘Role’ column as ‘Activity’

Disco – Tutorial

Enterprise anne@fluxicon.com

Role

column is used

X  Activity   

Name: Role

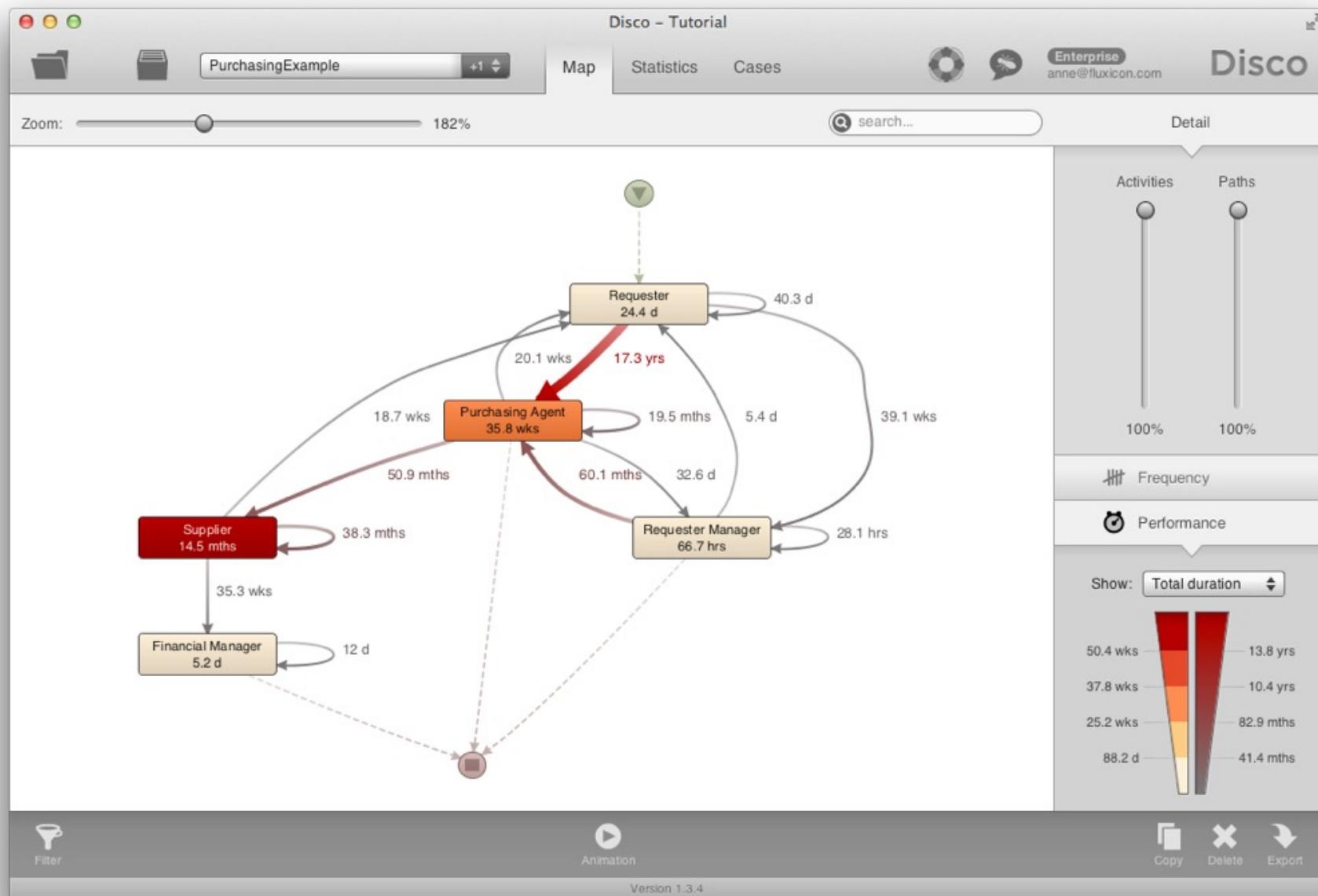
	Case ID	Start Timestamp	Complete Timestamp	Activity	Resource	Role
1	1	2011/01/01 00:00:00.000	2011/01/01 00:37:00.000	Create Purchase Requisition	Kim Passa	Requester
2	2	2011/01/01 00:16:00.000	2011/01/01 00:29:00.000	Create Purchase Requisition	Immanuel Karagianni	Requester
3	3	2011/01/01 02:23:00.000	2011/01/01 03:03:00.000	Create Purchase Requisition	Kim Passa	Requester
4	1	2011/01/01 05:37:00.000	2011/01/01 05:45:00.000	Create Request for Quotation	Kim Passa	Requester
5	1	2011/01/01 06:41:00.000	2011/01/01 06:55:00.000	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
6	2	2011/01/01 08:16:00.000	2011/01/01 08:26:00.000	Create Request for Quotation	Alberto Duport	Requester
7	4	2011/01/01 08:39:00.000	2011/01/01 09:00:00.000	Create Purchase Requisition	Fjodor Kowalski	Requester
8	2	2011/01/01 09:34:00.000	2011/01/01 09:38:00.000	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
9	5	2011/01/01 09:49:00.000	2011/01/01 10:35:00.000	Create Purchase Requisition	Esmara Liubiata	Requester
10	2	2011/01/01 10:16:00.000	2011/01/01 10:21:00.000	Amend Request for Quotation	Christian Francois	Requester
11	2	2011/01/01 11:15:00.000	2011/01/01 11:48:00.000	Analyze Request for Quotation	Magdalena Predutta	Purchasing Agent
12	6	2011/01/01 11:20:00.000	2011/01/01 11:37:00.000	Create Purchase Requisition	Christian Francois	Requester
13	1	2011/01/01 11:43:00.000	2011/01/01 12:09:00.000	Send Request for Quotation to Supplier	Karel de Groot	Purchasing Agent
14	1	2011/01/01 12:32:00.000	2011/01/01 16:03:00.000	Create Quotation comparison Map	Magdalena Predutta	Purchasing Agent
15	2	2011/01/01 12:33:00.000	2011/01/01 12:39:00.000	Amend Request for Quotation	Esmara Liubiata	Requester
16	2	2011/01/01 13:28:00.000	2011/01/01 13:38:00.000	Analyze Request for Quotation	Karel de Groot	Purchasing Agent
17	7	2011/01/01 14:05:00.000	2011/01/01 15:00:00.000	Create Purchase Requisition	Esmara Liubiata	Requester
18	8	2011/01/01 14:27:00.000	2011/01/01 15:17:00.000	Create Purchase Requisition	Fjodor Kowalski	Requester
19	2	2011/01/01 15:18:00.000	2011/01/01 15:40:00.000	Send Request for Quotation to Supplier	Francois de Perrier	Purchasing Agent
20	2	2011/01/01 15:55:00.000	2011/01/01 16:43:00.000	Create Quotation comparison Map	Karel de Groot	Purchasing Agent
21	9	2011/01/01 16:17:00.000	2011/01/01 16:34:00.000	Create Purchase Requisition	Tesca Lobes	Requester
22	6	2011/01/01 17:32:00.000	2011/01/01 17:45:00.000	Create Request for Quotation	Alberto Duport	Requester
23	8	2011/01/01 18:00:00.000	2011/01/01 18:07:00.000	Create Request for Quotation	Tesca Lobes	Requester
24	6	2011/01/01 18:39:00.000	2011/01/01 18:55:00.000	Analyze Request for Quotation	Magdalena Predutta	Purchasing Agent
25	4	2011/01/01 18:45:00.000	2011/01/01 18:51:00.000	Analyze Purchase Requisition	Maris Freeman	Requester Manager
26	4	2011/01/01 18:56:00.000	2011/01/01 18:58:00.000	Create Request for Quotation	Heinz Gutschmidt	Requester Manager
27	8	2011/01/01 19:04:00.000	2011/01/01 19:27:00.000	Analyze Request for Quotation	Francois de Perrier	Purchasing Agent
28	6	2011/01/01 19:47:00.000	2011/01/01 19:55:00.000	Amend Request for Quotation	Penn Osterwalder	Requester
29	4	2011/01/01 19:58:00.000	2011/01/01 20:10:00.000	Analyze Request for Quotation	Francois de Perrier	Purchasing Agent

Cancel File encoding: UTF-8 Use quotes  Ready to start import. Start import

Step 10 - Organizational View

Instead of the activity flow, we are now looking at how the process moves through different *roles* in the organization

- Inefficiencies can often be found at the borders of organizational units
- Clearly, the Purchasing agents are causing the biggest delays in the process!



Close the loop

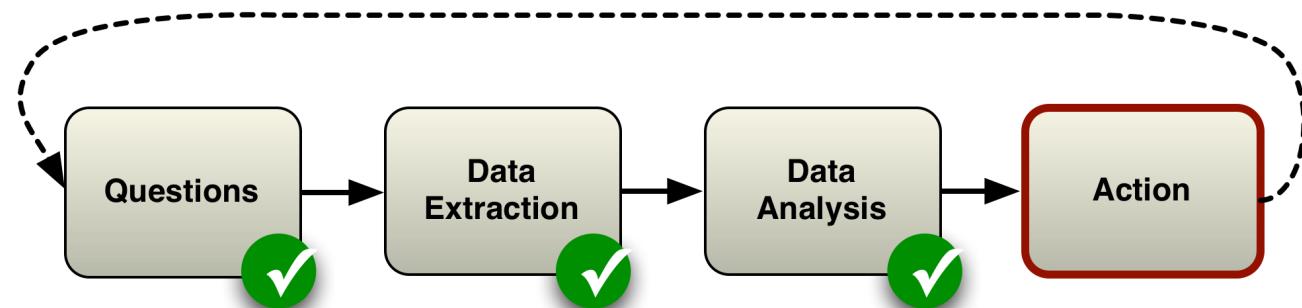
We have learned

- How to play with Disco to discover a process model

We have found

- opportunities to improve the process

Now: Take action and verify results



Bonus

Import data again and configure both
‘Activity’ and ‘Role’ column as ‘Activity’

- Can you see what happens now?

Outline

0. What is Process mining?
1. Example Scenario
2. Roadmap
3. Hands-on Session
- 4. Take-away Points**

Take-away Points

Real processes are often more complex than you would expect

There is no one “right” model:

- You can take multiple views on the same data

Process mining is an explorative, interactive activity

Questions?

Want to use these techniques in your research?

andrea.vandin@santannapisa.it