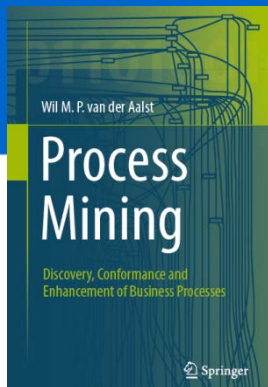


*Process Mining: Data Science in Action*

# Comparative Process Mining Using Process Cubes

prof.dr.ir. Wil van der Aalst  
[www.processmining.org](http://www.processmining.org)



**TU/e**

Technische Universiteit  
**Eindhoven**  
University of Technology

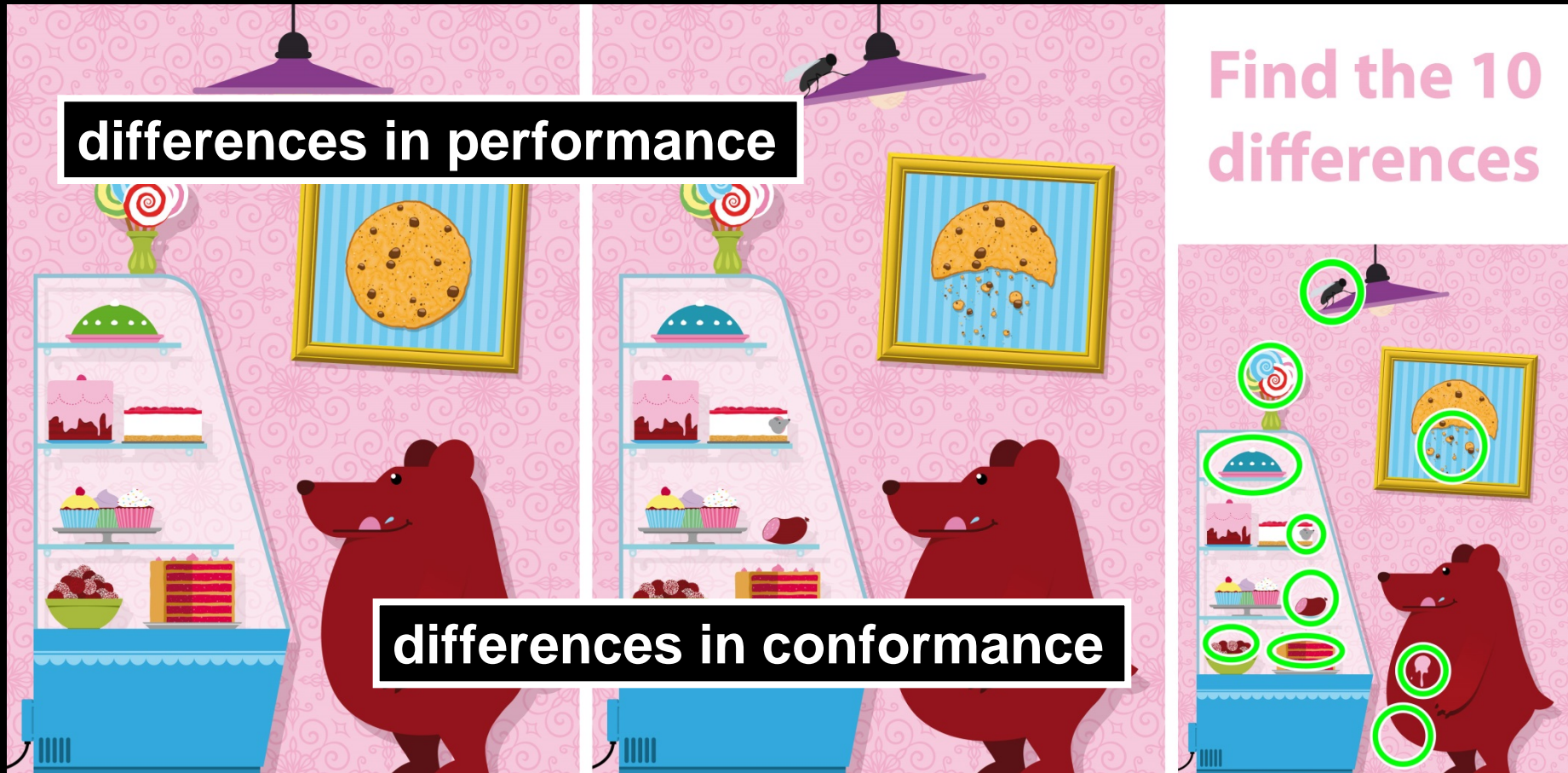
**Where innovation starts**



**differences in performance**

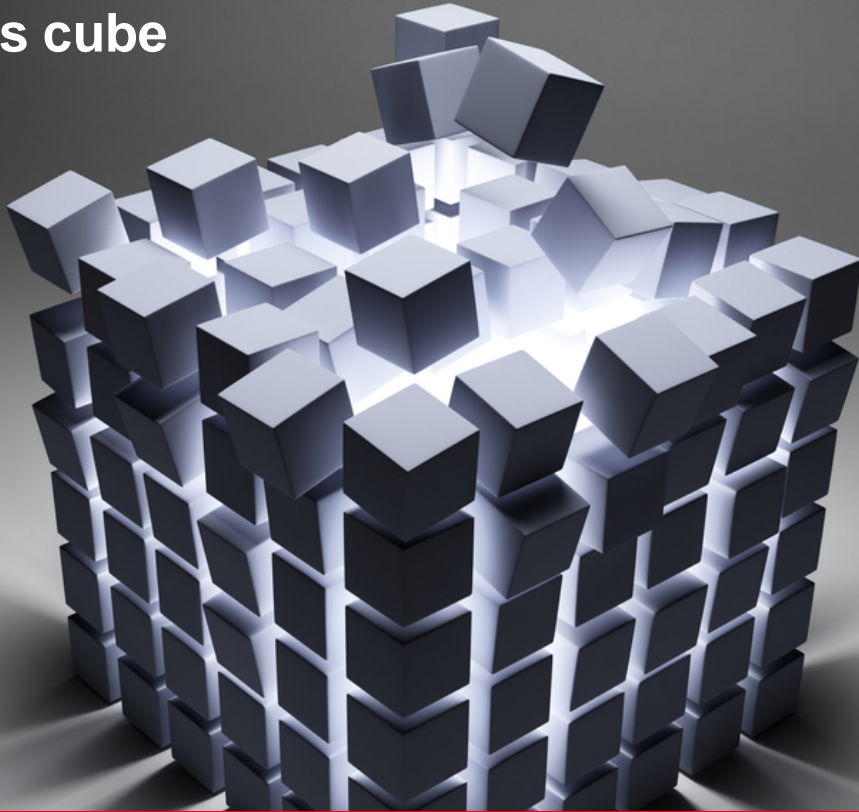
**differences in conformance**

**Find the 10  
differences**





**process cube**



**dimensions**

- **time**
- **department**
- **location**
- **amount**
- **gender**
- **level**
- **priority**
- **type**
- **compliant**
- ...

**compare event data and processes  
along different dimensions**

# Input data mining

**select (predictor)  
variables**

**select response  
variable**  
(in case of supervised learning)

attribute 1	attribute 2	attribute 3	...	attribute k
$V_{11}$	$V_{12}$	$V_{13}$	...	$V_{1k}$
$V_{21}$	$V_{22}$	$V_{23}$	...	$V_{2k}$
$V_{31}$	$V_{32}$			$V_{3k}$
$V_{41}$	$V_{42}$			$V_{4k}$
...	...			...

**every row is an  
instance of some  
(un)supervised  
learning problem**

# Input process mining

**select attribute  
representing the case**

**select attribute  
representing the  
activity**

attribute 1	attribute 2	attribute 3	...	attribute k
$V_{11}$	$V_{12}$	$V_{13}$	...	$V_{1k}$
...				...

**select attributes representing  
additional properties of the  
event or case (optional)**

**select attribute  
representing the  
time**

**select attribute  
representing the  
information (optional)**

# Loading a CSV file into Disco

Disco - New project

Academic  
w.m.p.v.d.aalst@tue.nl

Disco

Case ID

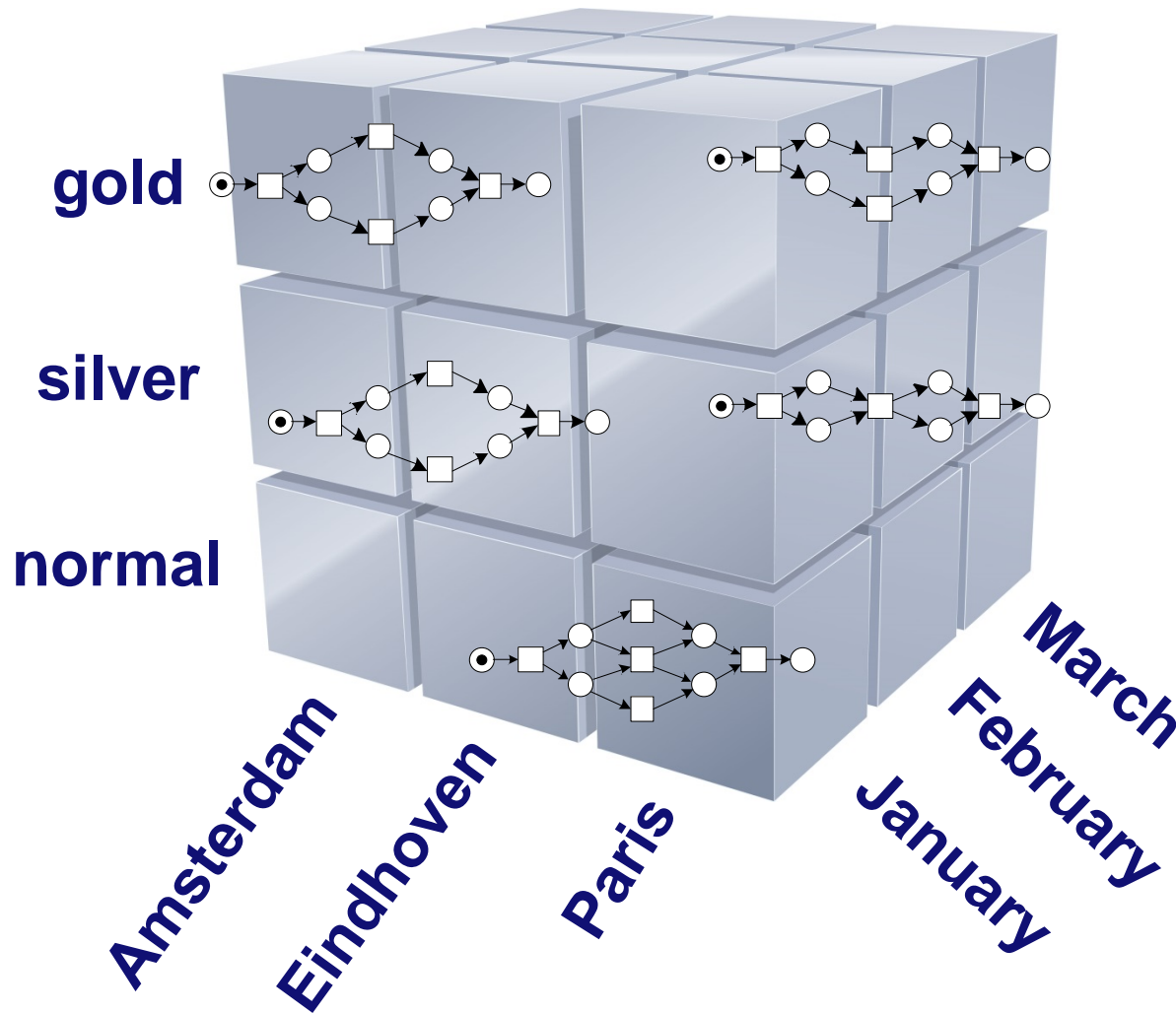
column is used

Case

Case ID	Activity	Resource	Complete Timestamp	AANTALV	AANVSPEC	ABW_COD	ABW_OMSCHR	AGBCOD	CBV_VERRCO	CBV_VERROM	CLUSTERNAAM	DBCSPEC	Divis
121	PROCINST.1	KLASSE 3B 613000	GYNAECOLOGIE H5Z	2005/05/14 01:00:00.000									
122	PROCINST.1	LIGDAGTARIEF 40014	GYNAECOLOGIE H5Z	2005/05/14 01:00:00.000									
123	PROCINST.1	VERV.CONSLT 411100	Polikliniek Verlosk.-Gyn.	2005/11/09 00:00:00.000									
124	PROCINST.2	VERV.CONSLT 411100	Polikliniek Verlosk.-Gyn.	2006/07/12 01:00:00.000		GYN	ZF	Poliklinisch ziekenfo...	7	411100	VERV.CONSLT	HERHAAL CONSULT	GYN
125	PROCINST.3	TARIEF CONS. 419100	Polikliniek Verlosk.-Gyn.	2005/07/13 01:00:00.000	1	GYN	NZ	Poliklinisch niet-ziek...	7	411100	VERV.CONSLT	HERHAAL CONSULT	GYN
126	PROCINST.3	TARIEF CONS. 419100	Polikliniek Verlosk.-Gyn.	2005/07/13 01:00:00.000	1	GYN	NZ	Poliklinisch niet-ziek...	7	419100	TARIEF CONS.	EERSTE ADMINISTR. C...	GYN
127	PROCINST.4	VERV.CONSLT 411100	Polikliniek Verlosk.-Gyn.	2005/02/24 00:00:00.000	1	GYN	ZF	Poliklinisch ziekenfo...	7	411100	VERV.CONSLT	HERHAAL CONSULT	GYN
128	PROCINST.4	CYTOLECTOC. 355201	Pathologie	2005/02/24 00:00:00.000	1	GYN	ZF	Poliklinisch ziekenfo...	7	355201	CYTOLECTOC.	GYNAECOLOGISCHE C...	GYN
129	PROCINST.4	TARIEF CONS. 419100	Polikliniek Verlosk.-Gyn.	2005/02/24 00:00:00.000	1	GYN	ZF	Poliklinisch ziekenfo...	7	419100	TARIEF CONS.	EERSTE ADMINISTR. C...	GYN
130	PROCINST.4	TEL.CON.S. KO 415100	Polikliniek Verlosk.-Gyn.	2005/03/17 00:00:00.000	1	GYN	ZF	Poliklinisch ziekenfo...	7	415100	TEL.CON.S. KO	TELEFONISCH CONSU...	GYN
131	PROCINST.5	ELEKTROCARD. 330001B	Algemeen Lab Klinische Chemie	2005/01/10 00:00:00.000									
132	PROCINST.5	THORAX 2R 386002	Radiologie	2005/01/10 00:00:00.000									
133	PROCINST.5	COUPE INZAGE 355111	Pathologie	2005/01/10 00:00:00.000									
134	PROCINST.5	GY...											
135	PROCINST.5	CIT...											
136	PROCINST.5	1E...											
137	PROCINST.5	TAR...											
138	PROCINST.5	AS...											
139	PROCINST.5	AB...											
140	PROCINST.5	RH-D CENTRIF 370606	Algemeen Lab Klinische Chemie	2005/01/10 00:00:00.000									
141	PROCINST.5	AANNAME LAB 370000	Algemeen Lab Klinische Chemie	2005/01/10 00:00:00.000									
142	PROCINST.5	HAEMOGLOB. S 370701S	Algemeen Lab Klinische Chemie	2005/01/10 00:00:00.000									
143	PROCINST.5	LEUCO ELEC S 377121S	Algemeen Lab Klinische Chemie	2005/01/10 00:00:00.000									
144	PROCINST.5	TROMBO S 370715S	Algemeen Lab Klinische Chemie										
145	PROCINST.5	AANNAME LAB 370000	Algemeen Lab Klinische Chemie										
146	PROCINST.5	ALBUMINE 378453A	Algemeen Lab Klinische Chemie										
147	PROCINST.5	ALK.FOSFAT. 370423	Algemeen Lab Klinische Chemie										
148	PROCINST.5	BILI. GECON. 370401	Algemeen Lab Klinische Chemie										
149	PROCINST.5	BILI TOTAAL 370401C	Algemeen Lab Klinische Chemie										
150	PROCINST.5	CALCIUM 377498A	Algemeen Lab Klinische Chemie	2005/01/10 00:00:00.000									
151	PROCINST.5	G-GLUT-TRANS 372417	Algemeen Lab Klinische Chemie	2005/01/10 00:00:00.000									
152	PROCINST.5	GLUCOSE 370402	Algemeen Lab Klinische Chemie	2005/01/10 00:00:00.000									
153	PROCINST.5	KALIUM POTEN 370443	Algemeen Lab Klinische Chemie	2005/01/10 00:00:00.000									
154	PROCINST.5	CREATININE S 370419S	Algemeen Lab Klinische Chemie	2005/01/10 00:00:00.000									
155	PROCINST.5	LDL-KINMET. 370410C	Algemeen Lab Klinische Chemie	2005/04/13 00:00:00.000									

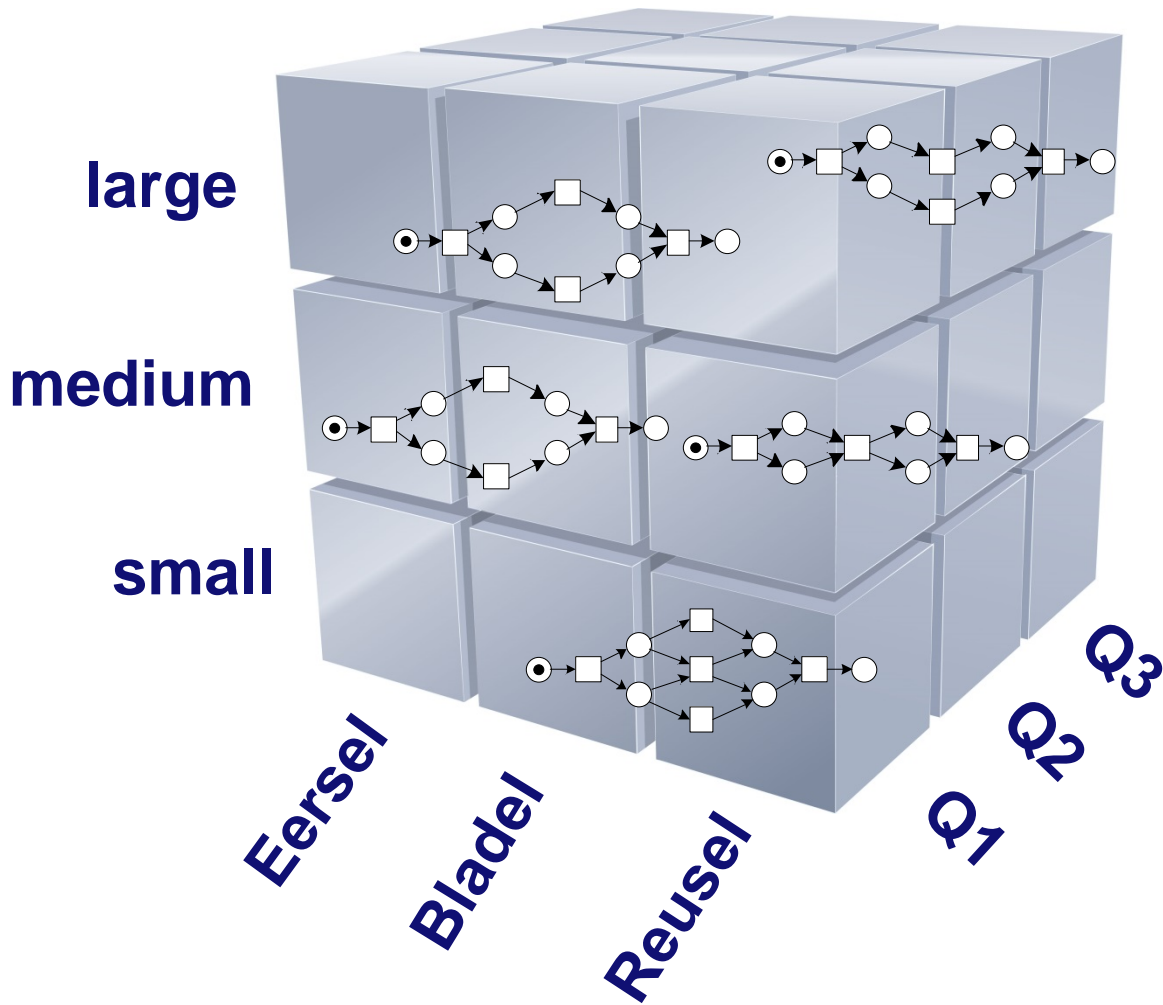
multiple mappings possible

comparing groups of cases/events

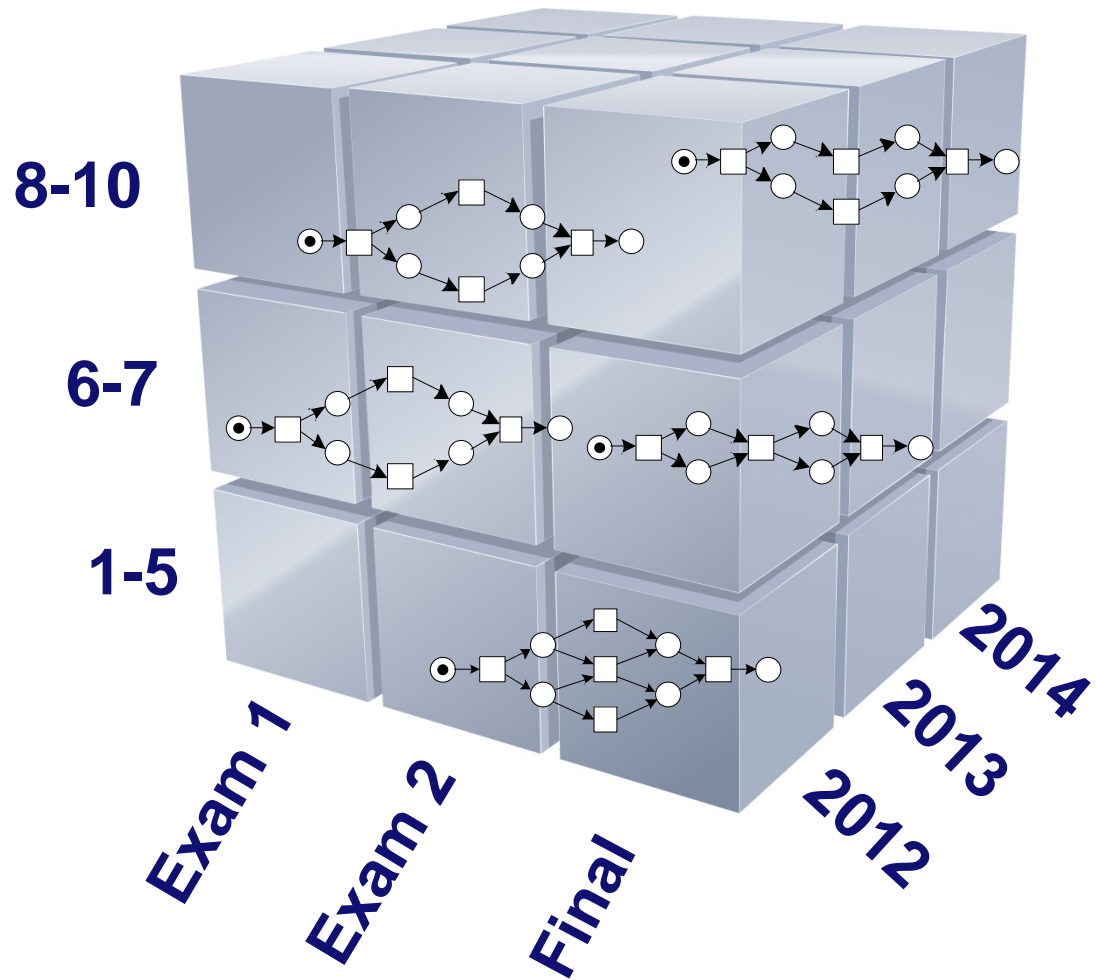


**Hertz has 8,650 rental locations and different types of customers.**



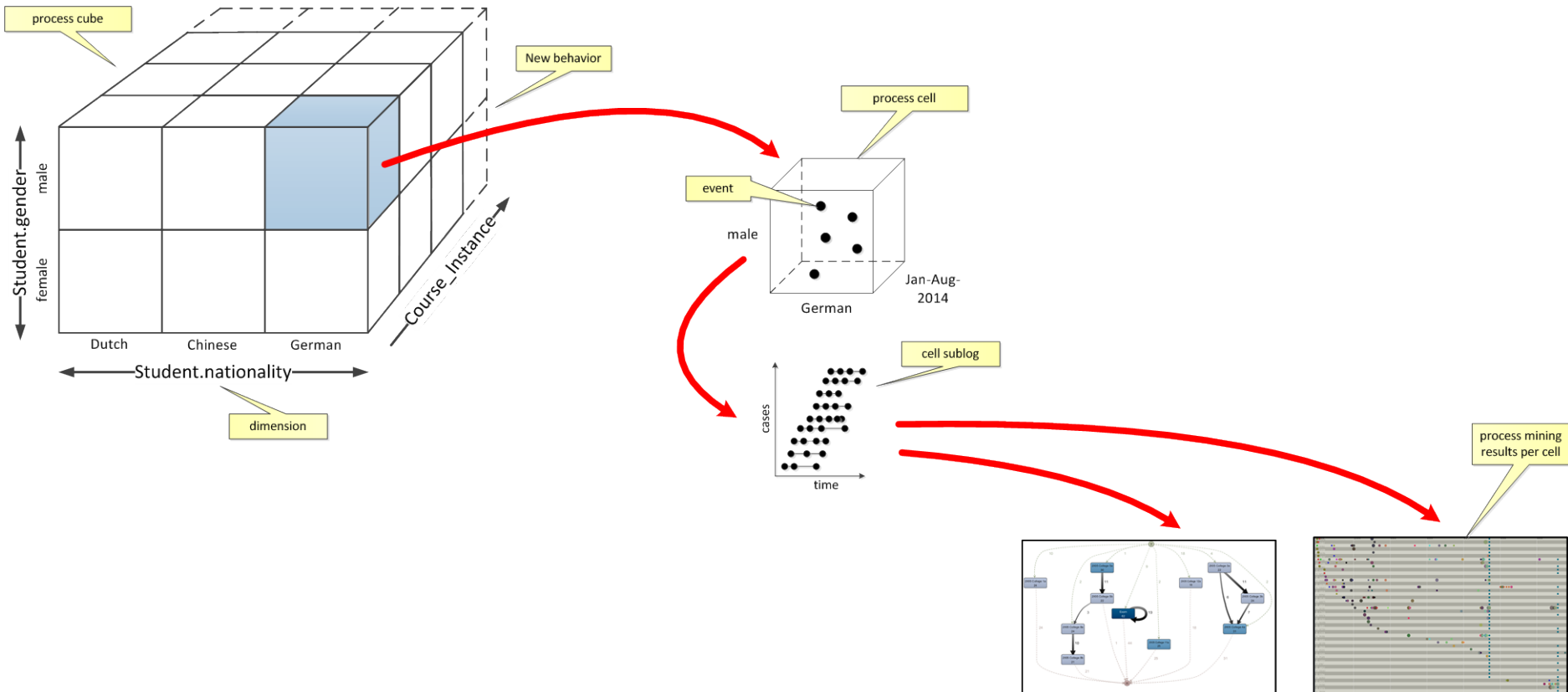


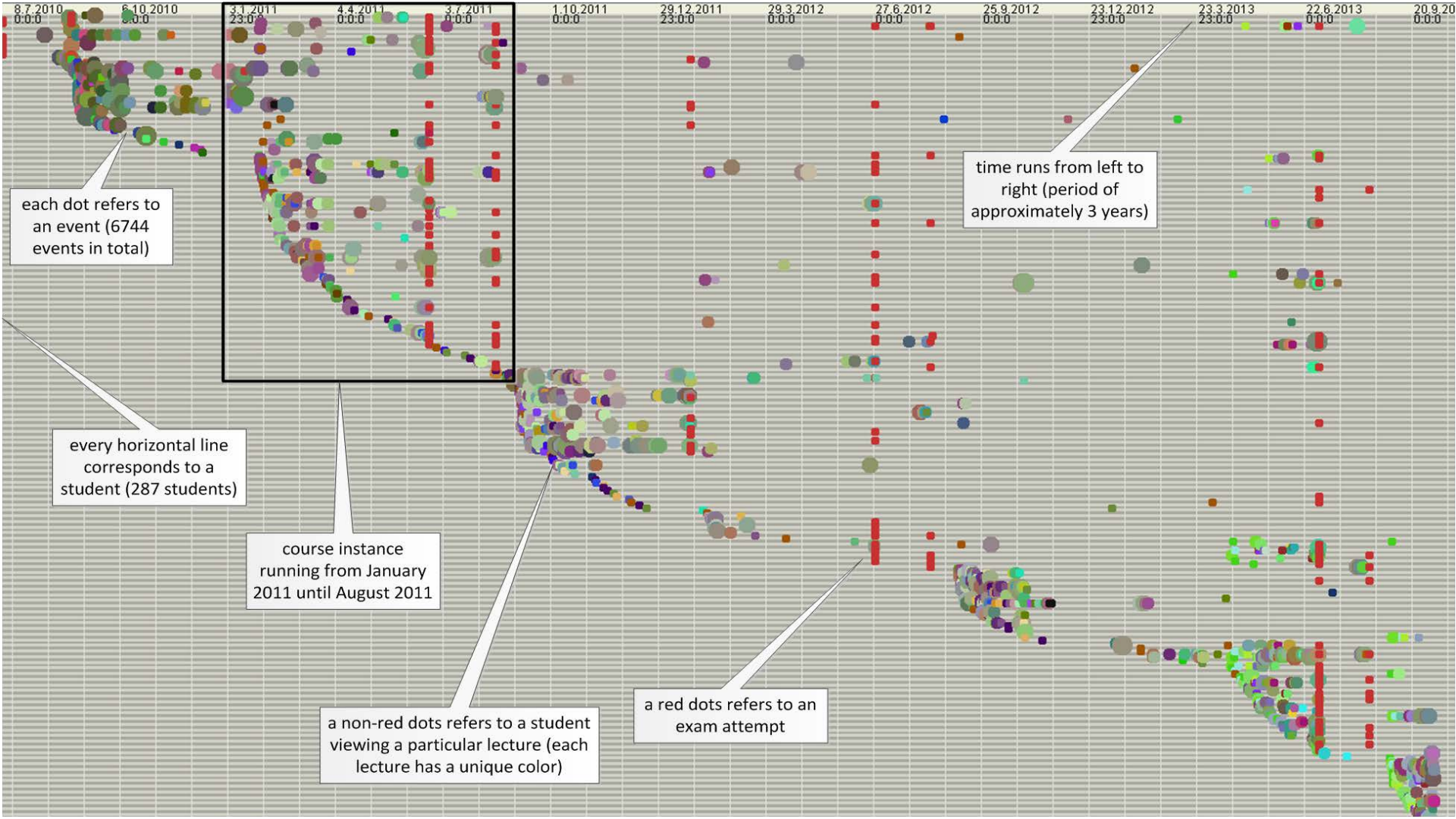
**All Dutch municipalities are handing out building permits within the boundaries set by the Dutch law.**



**Students  
making  
homework,  
assignments  
and exams.**

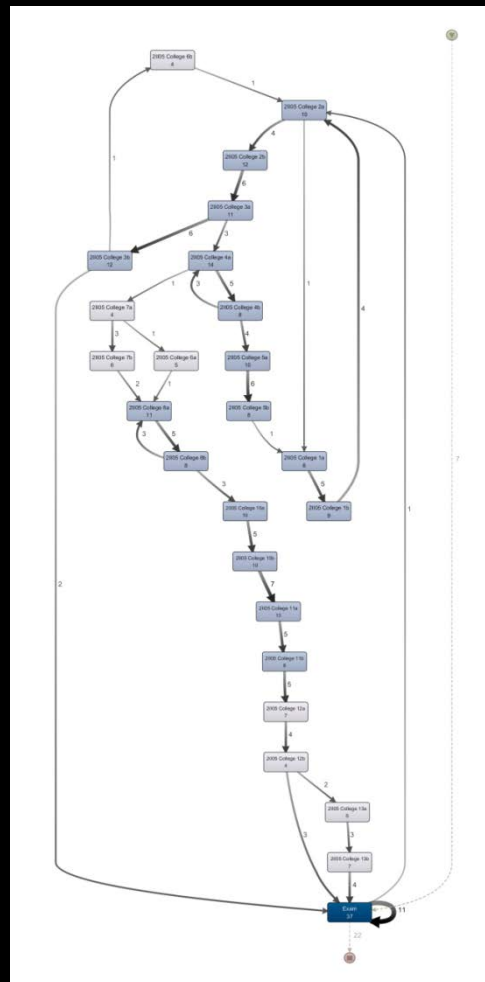
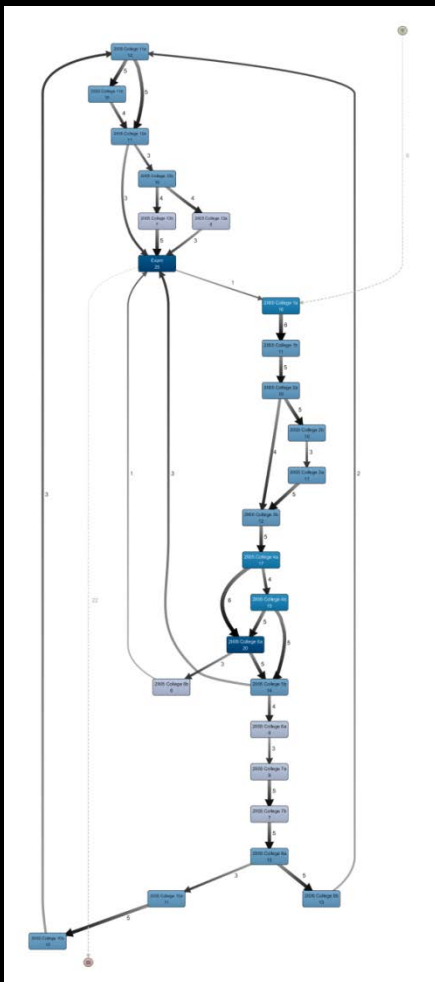
# Process cube: Basic idea is simple





PASSED

Fitness of event  
log wrt idealized  
model is 0.37



FAILED

Fitness of event  
log wrt idealized  
model is 0.28



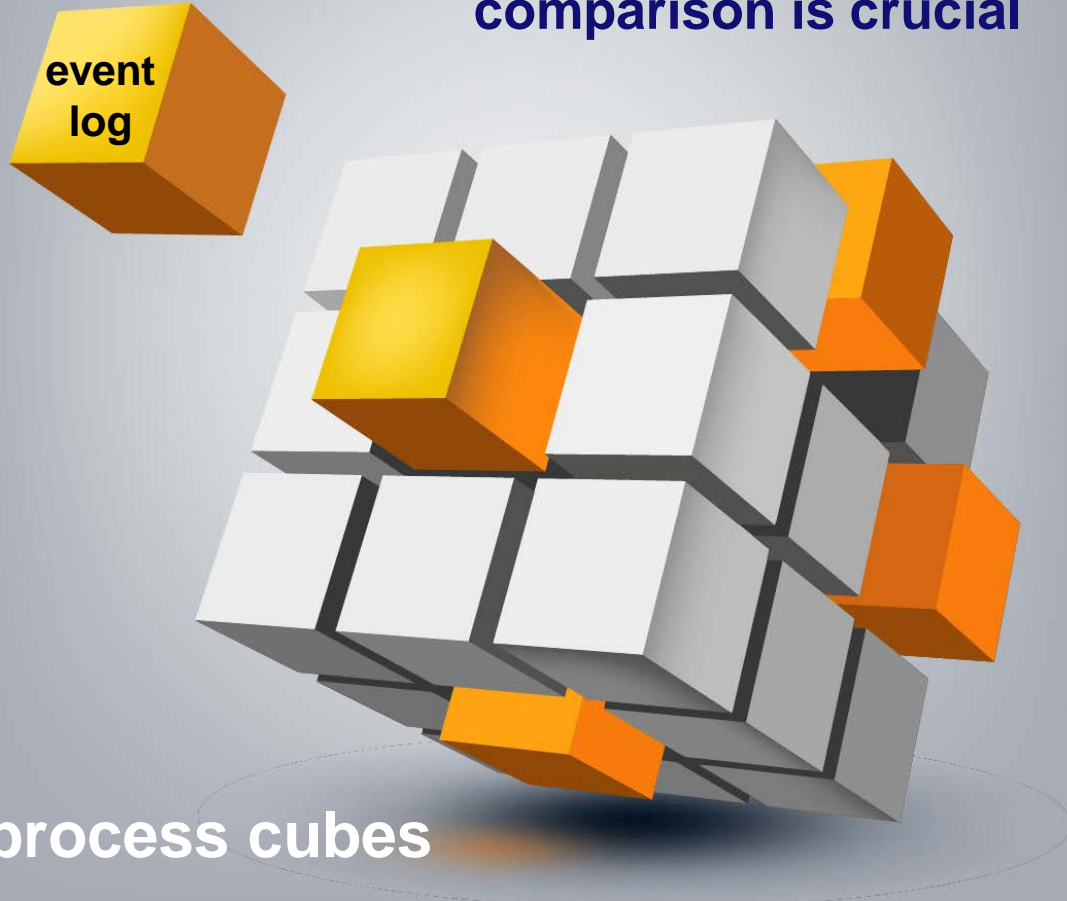


**data  
warehouse**

**Related, but often data is aggregated making it useless for process mining (events are lost).**

**Online Analytical Processing (OLAP)**

event data are multidimensional  
comparison is crucial





### *Part I: Preliminaries*

**Chapter 1**  
Introduction

**Chapter 2**  
Process Modeling and  
Analysis

**Chapter 3**  
Data Mining

### *Part III: Beyond Process Discovery*

**Chapter 7**  
Conformance  
Checking

**Chapter 8**  
Mining Additional  
Perspectives

**Chapter 9**  
Operational Support

### *Part II: From Event Logs to Process Models*

**Chapter 4**  
Getting the Data

**Chapter 5**  
Process Discovery: An  
Introduction

**Chapter 6**  
Advanced Process  
Discovery Techniques

### *Part IV: Putting Process Mining to Work*

**Chapter 10**  
Tool Support

**Chapter 11**  
Analyzing “Lasagna  
Processes”

**Chapter 12**  
Analyzing “Spaghetti  
Processes”

### *Part V: Reflection*

**Chapter 13**  
Cartography and  
Navigation

**Chapter 14**  
Epilogue

Wil M. P. van der Aalst

# Process Mining

Discovery, Conformance and  
Enhancement of Business Processes

 Springer

**For more details see:**

***W.M.P. van der Aalst. Process Cubes: Slicing, Dicing, Rolling Up and Drilling Down Event Data for Process Mining. In Asia Pacific Conference on Business Process Management (AP-BPM 2013), volume 159 of Lecture Notes in Business Information Processing, pages 1-22. Springer-Verlag, Berlin, 2013.***