

HALFBACK
HIGHLY AVAILABLE SMART FACTORIES IN THE CLOUD



Fonds européen de développement régional
(FEDER)
Europäischer Fonds für regionale Entwicklung
(EFRE)



Process Mining

Agnès Braud – agnes.braud@unistra.fr



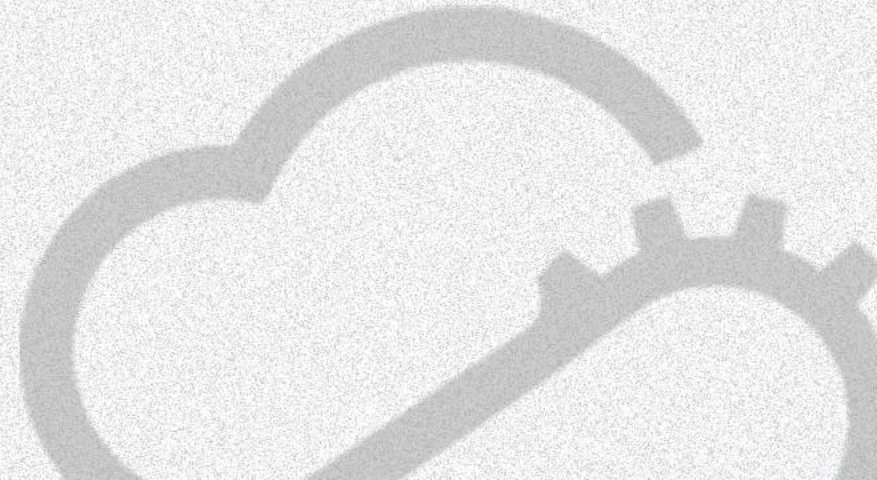
Workshop 13/06/2017

According to Gartner (2015), 72 percent of manufacturing industry's data is unused due to the complexities involved with variables, such as pressure, temperature and time.

"Predicts 2016: Opportunities Abound for the Factory to Reach Its Potential", Gartner, 2015

Examples of manufacturing data

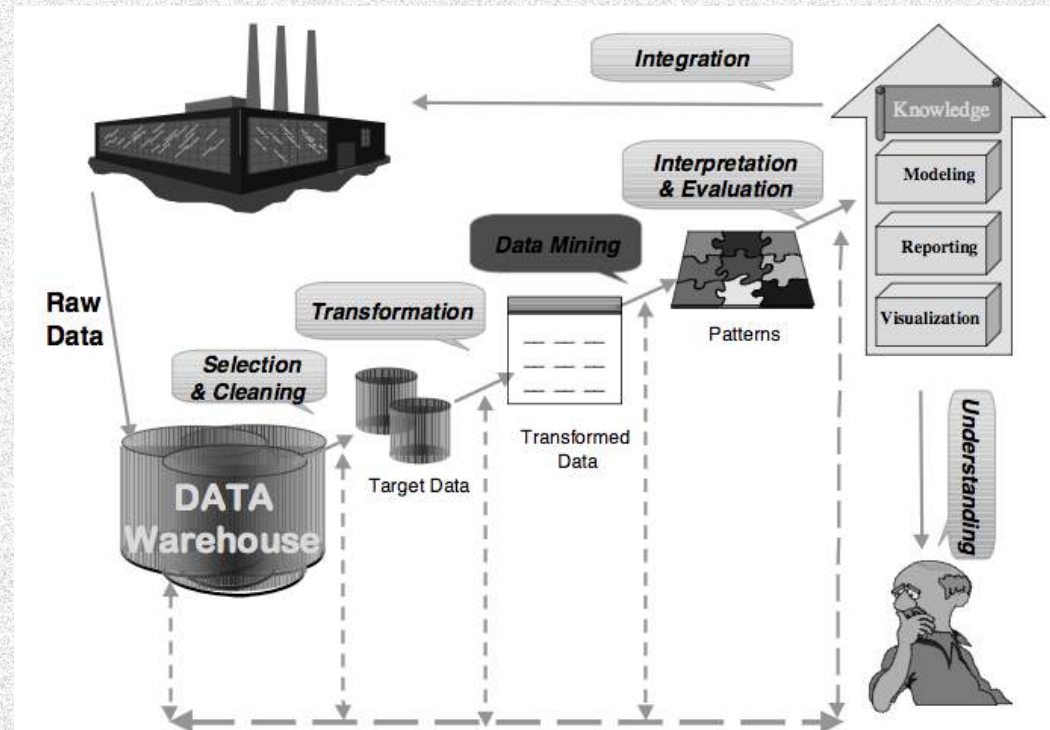
- Sensors data
- Error logs
- Maintenance reports
- ...



Mining data

- The goal: making value out of huge amount of data
- An iterative process

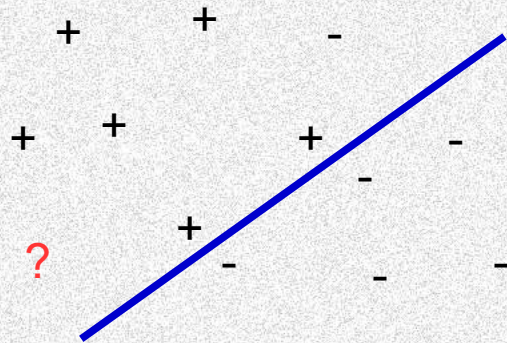
U. Fayyad: Taming the Giants and the Monsters: Mining Large Databases for Nuggets of Knowledge (1998)



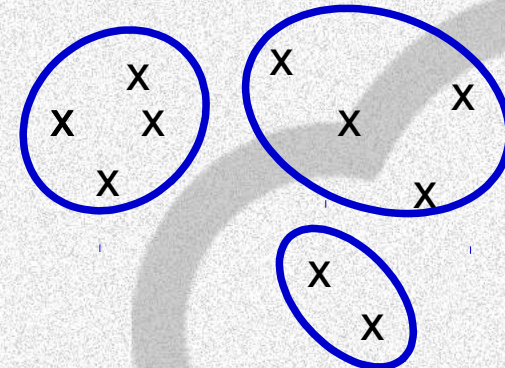
Mining data

- Different tasks

Predictive: data are labelled,
the goal is to learn a
model to predict the
label of **a new case**



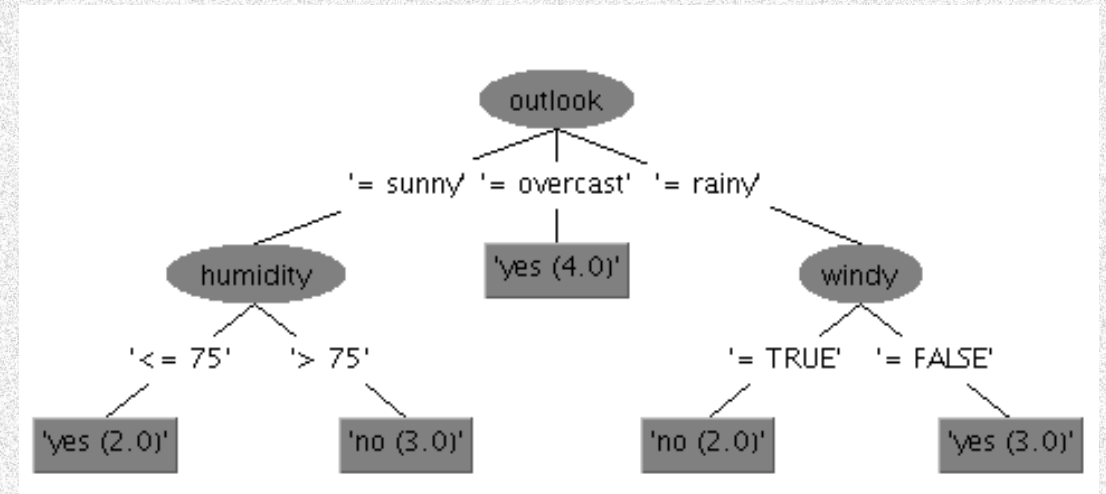
Descriptive: data are not labelled,
the goal is to find
regularities characterising
the set of data



A kind of predictive model : decision trees

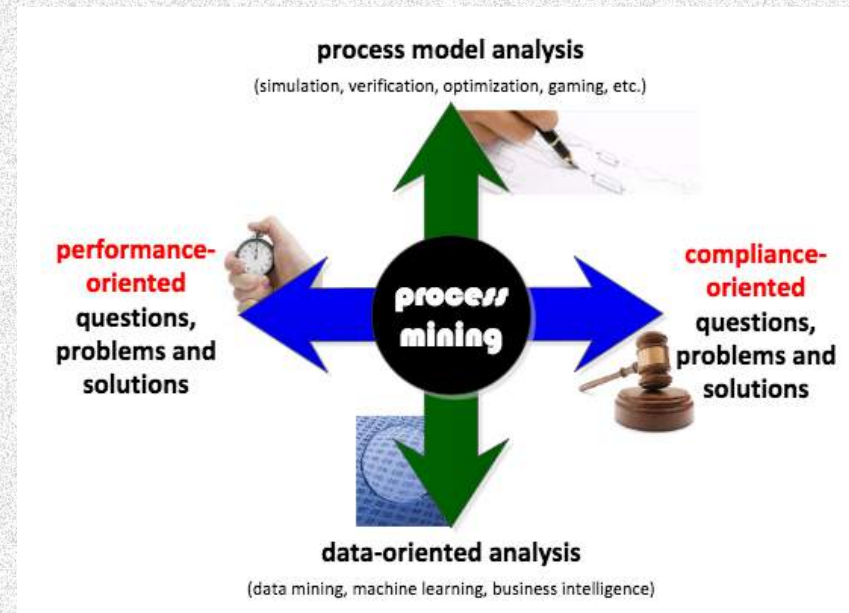
(generated with the Weka software by the University of Waikato)

No.	outlook Nominal	temperature Numeric	humidity Numeric	windy Nominal	play Nominal
1	sunny	85.0	85.0	FALSE	no
2	sunny	80.0	90.0	TRUE	no
3	overcast	83.0	86.0	FALSE	yes
4	rainy	70.0	96.0	FALSE	yes
5	rainy	68.0	80.0	FALSE	yes
6	rainy	65.0	70.0	TRUE	no
7	overcast	64.0	65.0	TRUE	yes
8	sunny	72.0	95.0	FALSE	no
9	sunny	69.0	70.0	FALSE	yes
10	rainy	75.0	80.0	FALSE	yes
11	sunny	75.0	70.0	TRUE	yes
12	overcast	72.0	90.0	TRUE	yes
13	overcast	81.0	75.0	FALSE	yes
14	rainy	71.0	91.0	TRUE	no



Process Mining

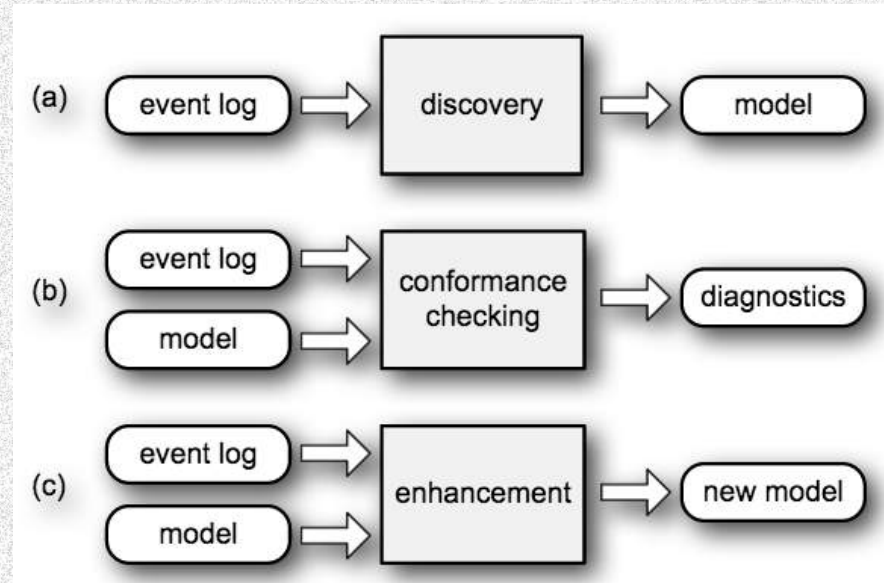
can be seen as the link between data mining and model analysis



W. van der Aalst et al.: Process Mining in the Large : A Tutorial (2013)

Process Mining

Three types



W. van der Aalst et al.: Process Mining Manifesto (2012)

Process Mining

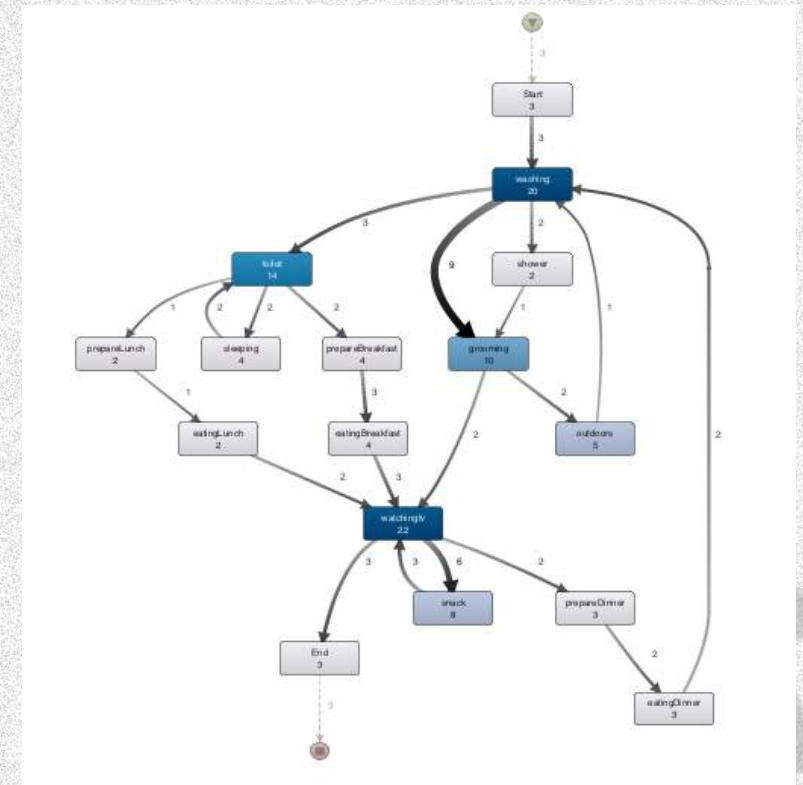
Example of dataset about daily living of several individuals (4TU.Centre for Research Data – processed with software Disco by Fluxicon)

	Case ID	Activity	Start Timestamp	Complete Timestamp
40	21	prepareDinner	2012/11/12 22:35:21.000	2012/11/12 22:37:55.000
41	21	eatingDinner	2012/11/12 22:35:21.000	2012/11/12 22:40:55.000
42	21	washing	2012/11/12 22:53:46.000	2012/11/12 22:53:54.000
43	21	watchingtv	2012/11/12 22:54:10.000	2012/11/12 23:34:43.000
44	21	toilet	2012/11/12 23:36:45.000	2012/11/12 23:45:55.000
45	21	watchingtv	2012/11/12 23:46:23.000	2012/11/13 01:27:41.000
46	21	End	2012/11/13 01:27:41.000	2012/11/13 01:27:41.000
47	31	Start	2012/11/13 01:29:06.000	2012/11/13 01:29:06.000
48	31	washing	2012/11/13 01:29:06.000	2012/11/13 01:29:39.000
49	31	grooming	2012/11/13 01:29:52.000	2012/11/13 01:29:56.000
50	31	sleeping	2012/11/13 01:30:32.000	2012/11/13 08:53:50.000
51	31	prepareBreakfast	2012/11/13 08:56:37.000	2012/11/13 09:00:26.000
52	31	eatingBreakfast	2012/11/13 09:04:54.000	2012/11/13 09:10:12.000
53	31	washing	2012/11/13 09:11:46.000	2012/11/13 09:11:49.000

Process Mining

Process discovery

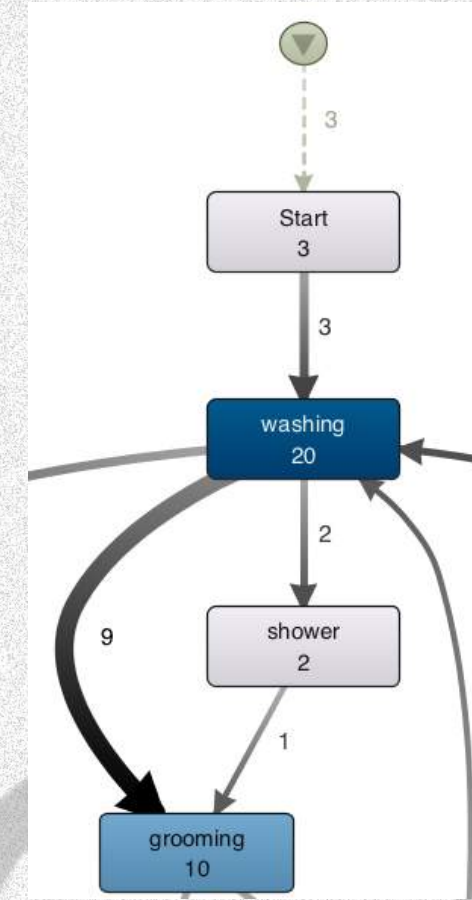
	Case ID	Activity	Start Timestamp	Complete Timestamp
40	21	prepareDinner	2012/11/12 22:35:21.000	2012/11/12 22:37:55.000
41	21	eatingDinner	2012/11/12 22:35:21.000	2012/11/12 22:40:55.000
42	21	washing	2012/11/12 22:53:46.000	2012/11/12 22:53:54.000
43	21	watchingtv	2012/11/12 22:54:10.000	2012/11/12 23:34:43.000
44	21	toilet	2012/11/12 23:36:45.000	2012/11/12 23:45:55.000
45	21	watchingtv	2012/11/12 23:46:23.000	2012/11/13 01:27:41.000
46	21	End	2012/11/13 01:27:41.000	2012/11/13 01:27:41.000
47	31	Start	2012/11/13 01:29:06.000	2012/11/13 01:29:06.000
48	31	washing	2012/11/13 01:29:06.000	2012/11/13 01:29:39.000
49	31	grooming	2012/11/13 01:29:52.000	2012/11/13 01:29:56.000
50	31	sleeping	2012/11/13 01:30:32.000	2012/11/13 08:53:50.000
51	31	prepareBreakfast	2012/11/13 08:56:37.000	2012/11/13 09:00:26.000
52	31	eatingBreakfast	2012/11/13 09:04:54.000	2012/11/13 09:10:12.000
53	31	washing	2012/11/13 09:11:46.000	2012/11/13 09:11:49.000



Process Mining

Process discovery

	Case ID	Activity	Start Timestamp	Complete Timestamp
40	21	prepareDinner	2012/11/12 22:35:21.000	2012/11/12 22:37:55.000
41	21	eatingDinner	2012/11/12 22:35:21.000	2012/11/12 22:40:55.000
42	21	washing	2012/11/12 22:53:46.000	2012/11/12 22:53:54.000
43	21	watchingtv	2012/11/12 22:54:10.000	2012/11/12 23:34:43.000
44	21	toilet	2012/11/12 23:36:45.000	2012/11/12 23:45:55.000
45	21	watchingtv	2012/11/12 23:46:23.000	2012/11/13 01:27:41.000
46	21	End	2012/11/13 01:27:41.000	2012/11/13 01:27:41.000
47	31	Start	2012/11/13 01:29:06.000	2012/11/13 01:29:06.000
48	31	washing	2012/11/13 01:29:06.000	2012/11/13 01:29:39.000
49	31	grooming	2012/11/13 01:29:52.000	2012/11/13 01:29:56.000
50	31	sleeping	2012/11/13 01:30:32.000	2012/11/13 08:53:50.000
51	31	prepareBreakfast	2012/11/13 08:56:37.000	2012/11/13 09:00:26.000
52	31	eatingBreakfast	2012/11/13 09:04:54.000	2012/11/13 09:10:12.000
53	31	washing	2012/11/13 09:11:46.000	2012/11/13 09:11:49.000



Process Mining Conformance checking

