

## Question 1

(a)

Import log-flat.xes to ProM. Click the eye symbol to view resource.



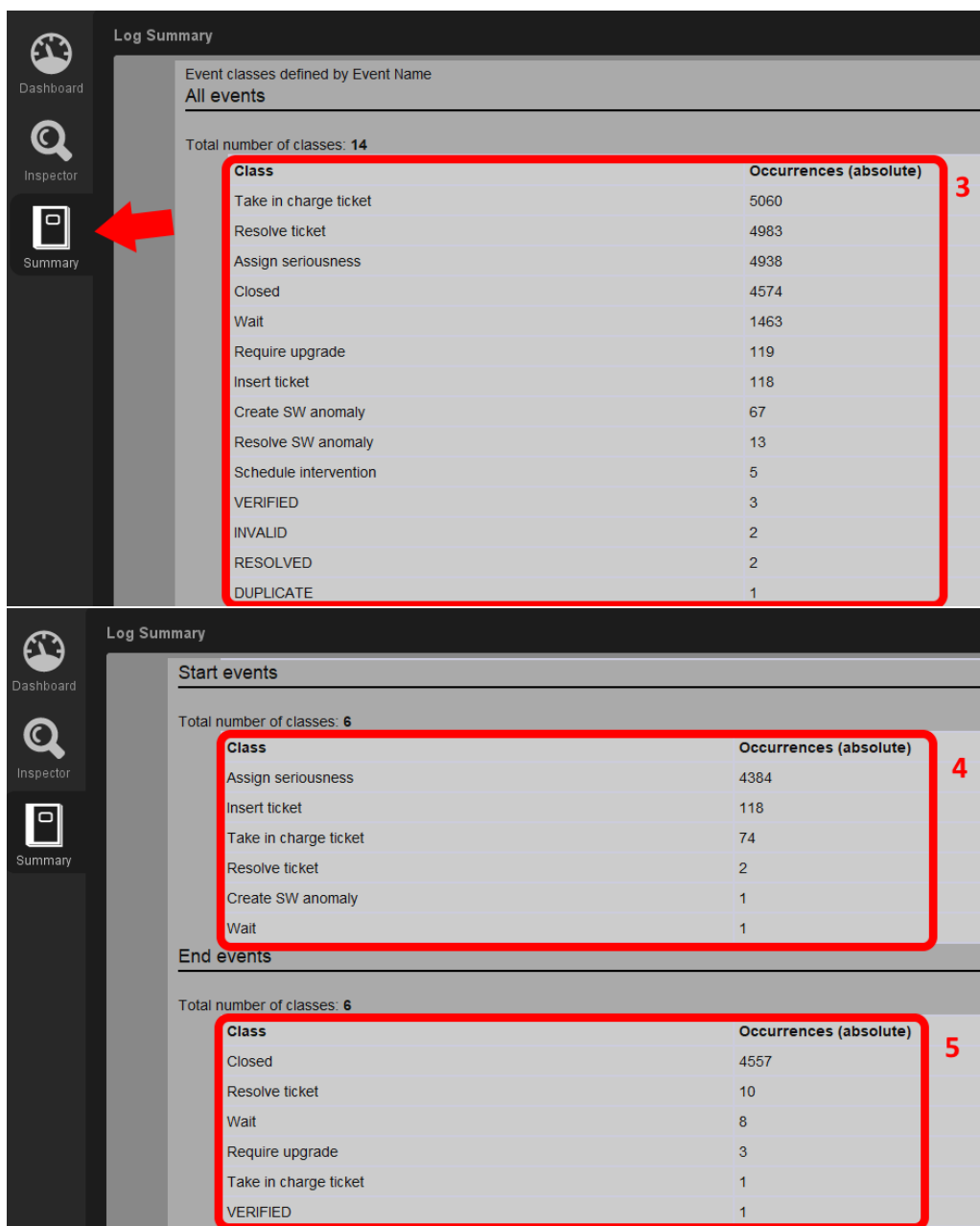
This results in the following view:



From it we can read

- the time period (1) covered by the event log, which is from 13.01.2010 to 03.01.2014
- the number of cases, events and activities of the log (2), being 4580, 21348 and 14 respectively (note that activities appear in ProM as event classes)

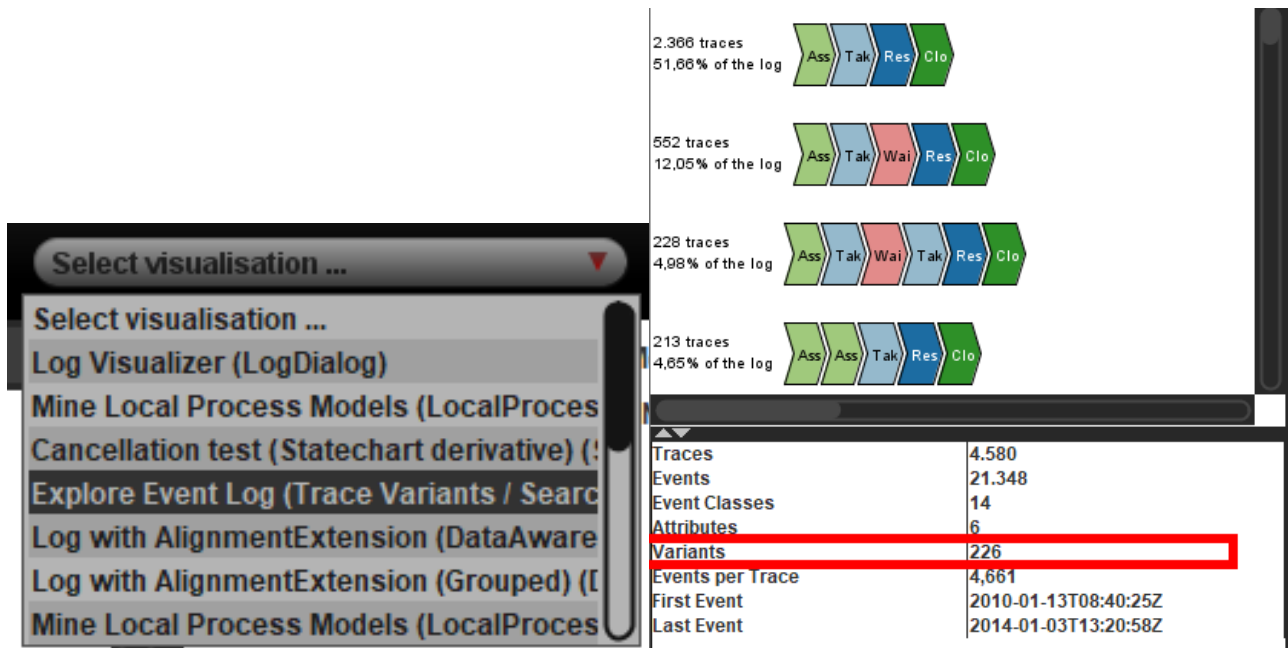
To gain more information on the activities we click the summary tab in the left which results in the following view:



From (3) we get a table of occurrence frequencies for each activity. From (4) we get a table of occurrence frequencies for each start activity. From (5) we get a table of occurrence frequencies for each end activity.

To determine the number of unique trace variants we click on 'Select visualization' and select 'Explore Event Log'

Under this view all trace variants are listed and some further information is given. From this we learn that there are 226 unique variants in the event log.

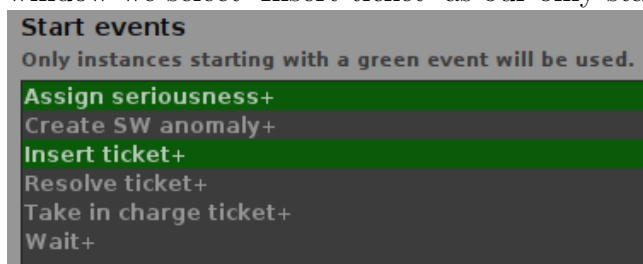


### To-do: Discussion

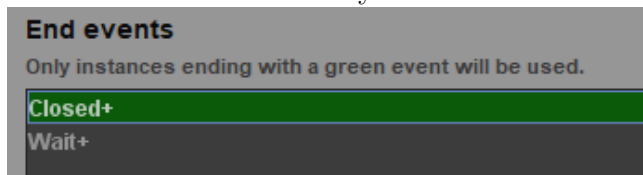
98% of tickets taken in charge are resolved ( $\frac{4983}{5060}$ ) The variants seem to be quite diverse (1:20 ratio of cases to variants, although distributed very unevenly)

(b)

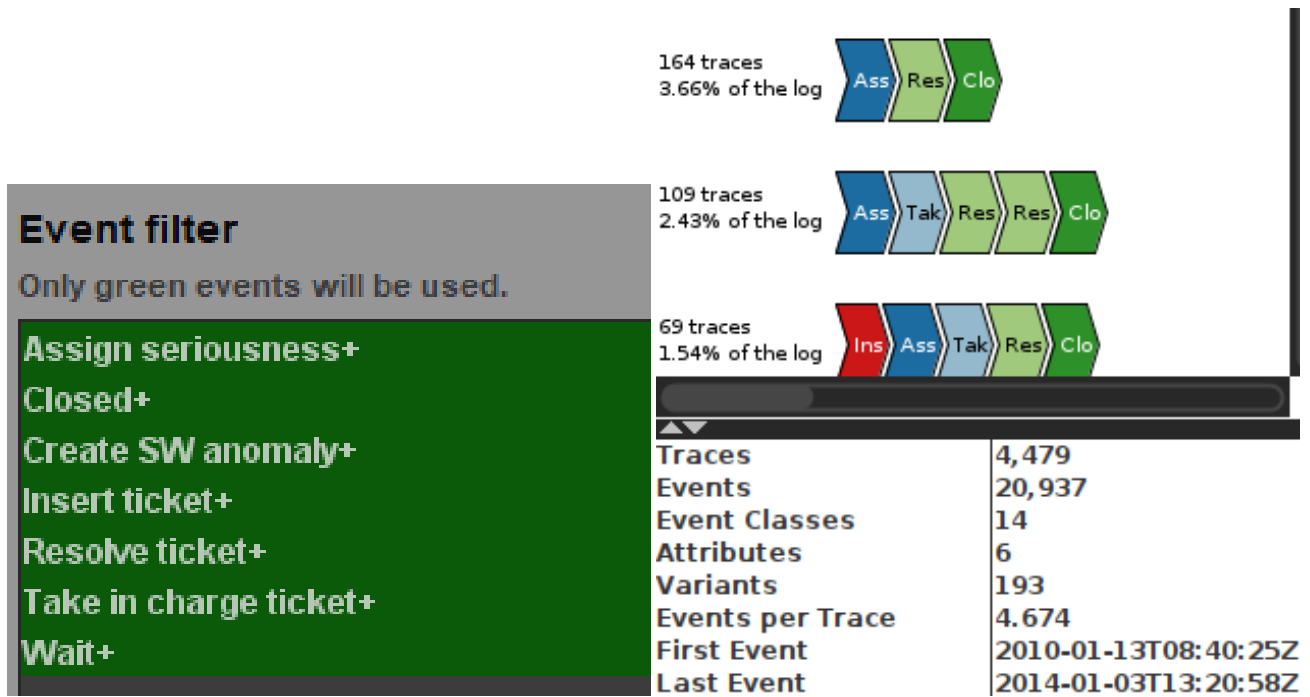
From the introduction we learned that every trace has to start with 'Insert ticket' or 'Assign seriousness' and ends with 'Closed'. Therefore every trace that does not begin/end with these events must have started/ended outside of our observed time period, making it incomplete. To filter out incomplete traces we go on the 'Actions' tab, select 'Filter Log using Simple Heuristics' and press 'Start'. In the first dialogue we just click 'Next'. In the next dialogue window we select 'Insert ticket' as our only start event and click 'Next'.



For the end events we only select 'Closed' and click 'Next'.



Since we do not want to filter out any other events we select 100% of events in Event Filter and click 'Finish'. We now get the following log:



(c)

1. As in (a) we inspect the overview of *log-complete* to find 114 cases, 655 events and 7 activities. Also just like in (a) we select the visualization 'Explore Event Log' to find out there are 20 unique trace variants.



2. Just like in (a) we look at 'Summary' to find a table of activities along with their frequency of occurrence:



3.

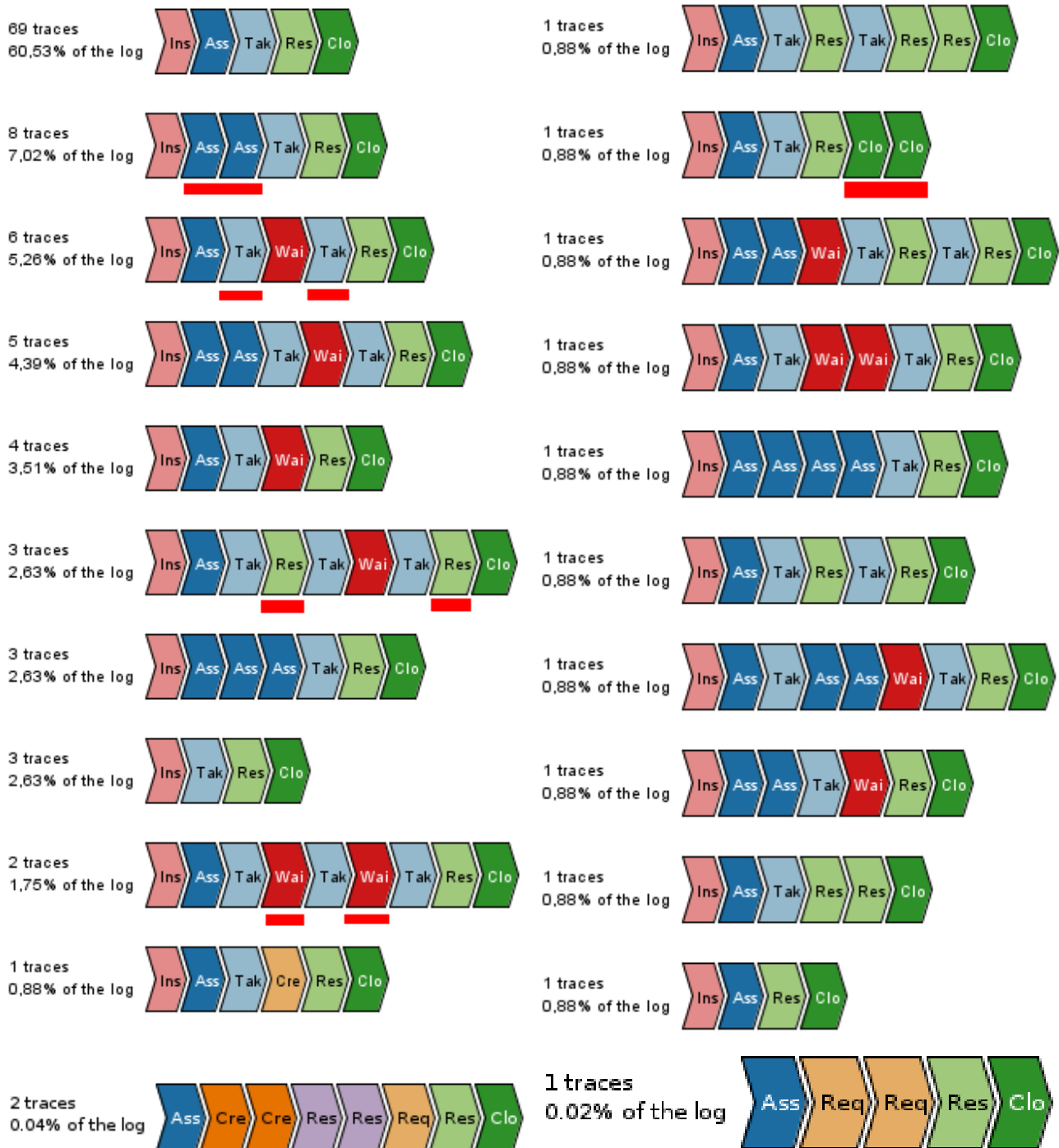
4.

5. To find out which activities appear more than once in at least one trace we again take a look at the 'Explore Event Log' visualization:

Assignment 1

Business Process Intelligence

June 29, 2022



The tasks that appear more than once (as marked in the visualization) are

- Ass: Assign seriousness
- Tak: Take in charge ticket
- Res: Resolve ticket
- Wai: Wait
- Clo: Closed
- Cre: Create SW Anomaly
- Req: Require Update

**(d)**

1. We choose Pie Charts for these visualization. For *Ticket type* distribution we choose **TICKET TYPE** as dimension:

```
"case_table_csv"."TICKET TYPE"
```

and **COUNT(TICKET TYPE)** as KPI:

```
COUNT("case_table_csv"."TICKET TYPE")
```

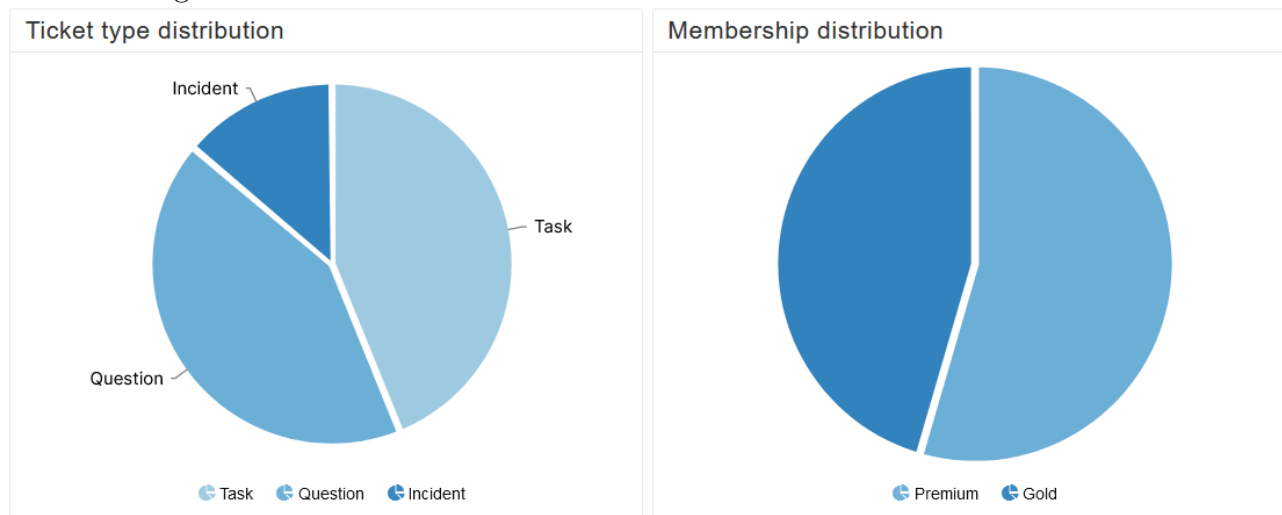
For *Membership* distribution we choose **MEMBERSHIP** as dimension:

```
"case_table_csv"."MEMBERSHIP"
```

and **COUNT(MEMBERSHIP)** as KPI:

```
COUNT("case_table_csv"."MEMBERSHIP")
```

The resulting distribution visualization can be seen below.



2. We obtained the column chart titled 'Total workload per ressource' by using

```
"event_table_csv"."RESOURCE"
```

as dimension and

```
COUNT("event_table_csv"."ACTIVITY")
```

as KPI. The x-Axis shows the resources 1-8 in ascending order, the y-Axis shows the summed number of activities handled by the given resource.

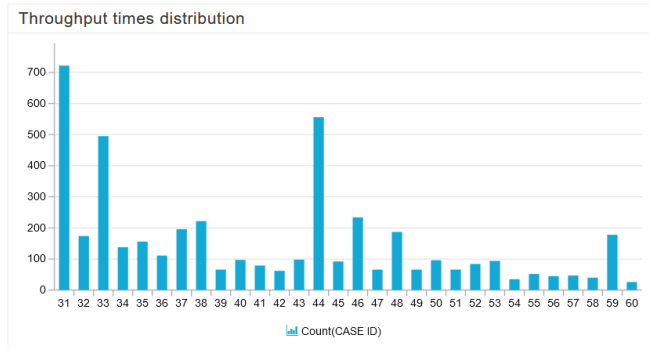
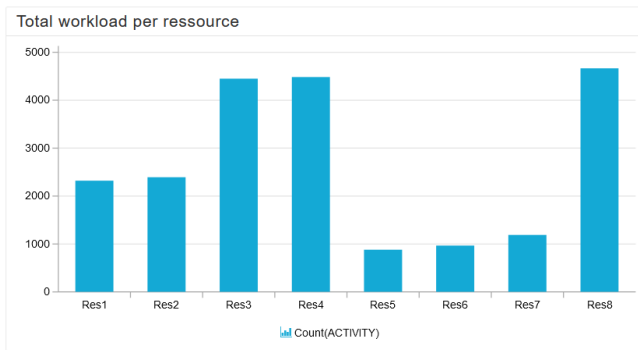
3. We created the column chart named 'Throughput times distribution' by selecting the total throughput time in days as our dimension:

```
CALC_THROUGHPUT(ALL_OCCURRENCE['Process Start'] TO ALL_OCCURRENCE['Process End'], REMAP_
```

and selecting

```
COUNT("case_table_csv"."CASE ID")
```

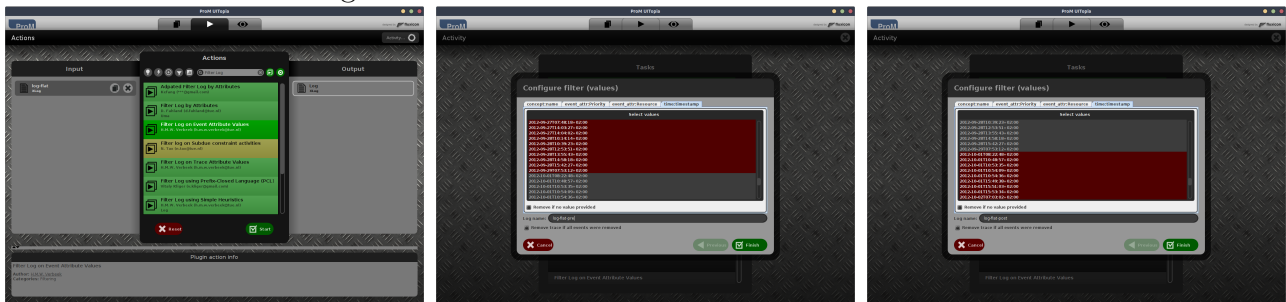
as our KPI. The x-Axis shows the throughput time in days and the y-Axis shows the number of cases with a given throughput time.



## Question 2

(a)

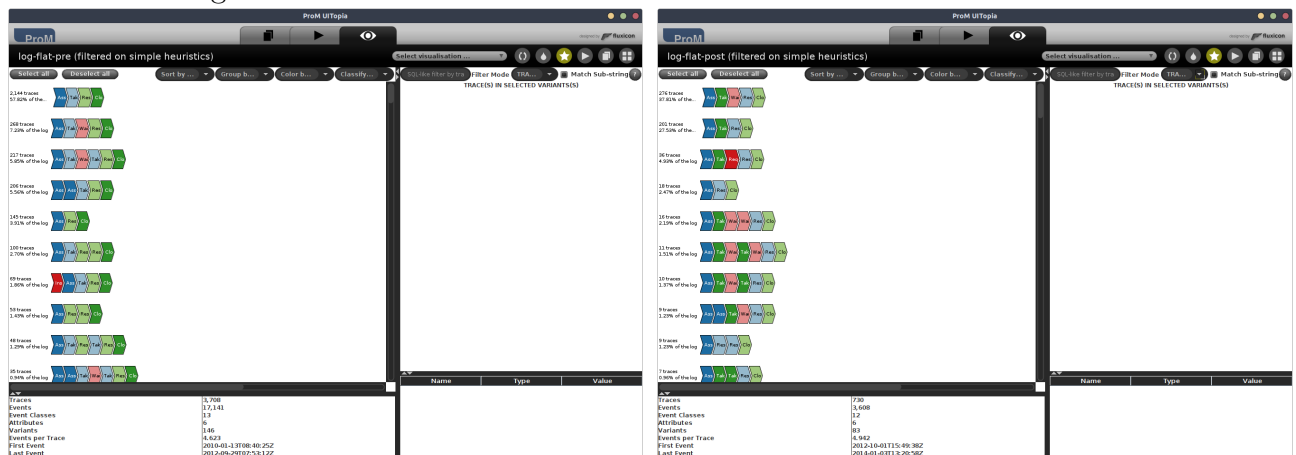
We can split the event log into two event logs using the plugin 'Filter Log on Event Attribute Values' with the following selections:



Afterwards, we can apply our filtering to only allow valid traces, as seen before.

For the event log pre 01.10.2012 we get 13 unique activities across 3708 cases, as a whole consisting of 6 variants.

For the event log post (including) 01.10.2012 we get 12 unique activities across 730 cases, as a whole consisting of 6 variants.





(a)

(b)

(c)

(d)

**Question 3**

**Question 4**

**Question 5**

**Question 6**