

## Filtering on start activities

In general, PM4Py offers two methods to filter a log or a dataframe on start activities. In the first method, a list of start activities has to be specified. On the activities that are contained in the list, the filter is applied on. In the second method, a decreasing factor is used. An explanation can be inspected by clicking on the button below.

```
from pm4py.algo.filtering.log.start_activities import start_activities_f
log_start = start_activities_filter.get_start_activities(log)
filtered_log = start_activities_filter.apply(log, ["S1"]) #suppose "S1"
```

Figure – 1

```
from pm4py.algo.filtering.pandas.start_activities import start_activitie
log_start = start_activities_filter.get_start_activities(dataframe)
df_start_activities = start_activities_filter.apply(dataframe, ["S1"],
                                                    parameters={start_activities_f
                                                                    start_activities_f
```

Figure -2

First of all, it might be necessary to know the starting activities. Therefore, code snippets are provided. Subsequently, an example of filtering is provided. The first snippet is working with log object, the second one is working on a dataframe. log\_start is a dictionary that contains as key the activity and as value the number of occurrence.

```
from pm4py.algo.filtering.log.start_activities import start_activities_f
log_af_sa = start_activities_filter.apply_auto_filter
(log, parameters={start_activities_filter.Parameters.DECR
```

Figure -3

```
from pm4py.algo.filtering.pandas.start_activities import start_activitie
df_auto_sa = start_activities_filter.apply_auto_filter
(dataframe, parameters={start_activities_filter.Parameter
```

Figure - 4

As mentioned earlier, there is also a method that aims to keep the frequent start activities. Again, the first snippet is about a log object, the second is about a dataframe object. The default value for DECREASING\_FACTOR is 0.6.

