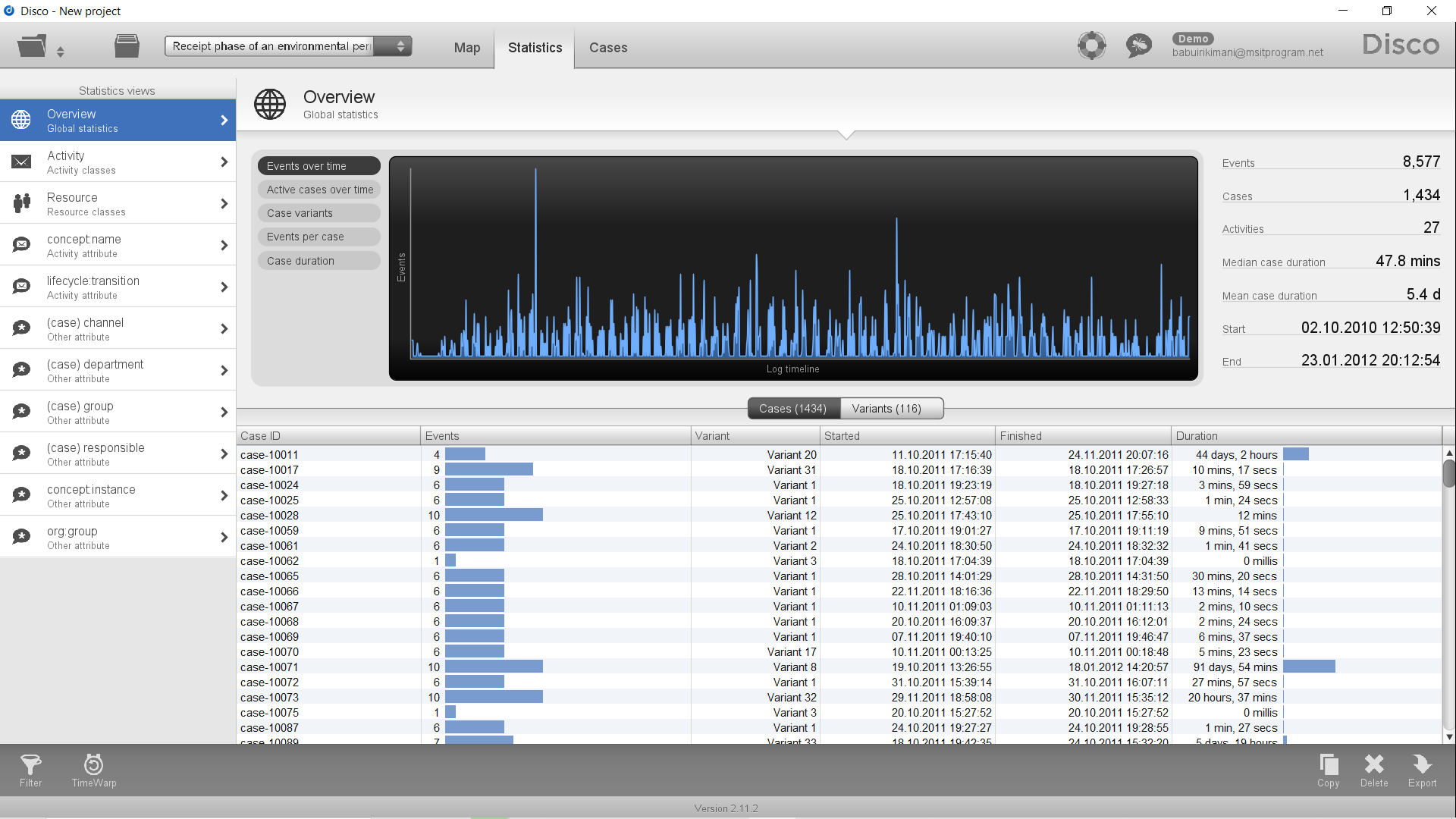
**DISCO**

Disco is simple, fast and easy to use. But in terms of functionality, coffers only smaller parts of this course. The advantages are easier and pleasant to use.

Open the event log ('Receipt phase of an environmental permit application process (\_WABO\_) CoSeLoG project.fbt') in Disco and switch to the 'Statistics' view.



Without switching to other views, use the statistics view to answer the following three subquestions:

1. **How many events are there on average per case?**

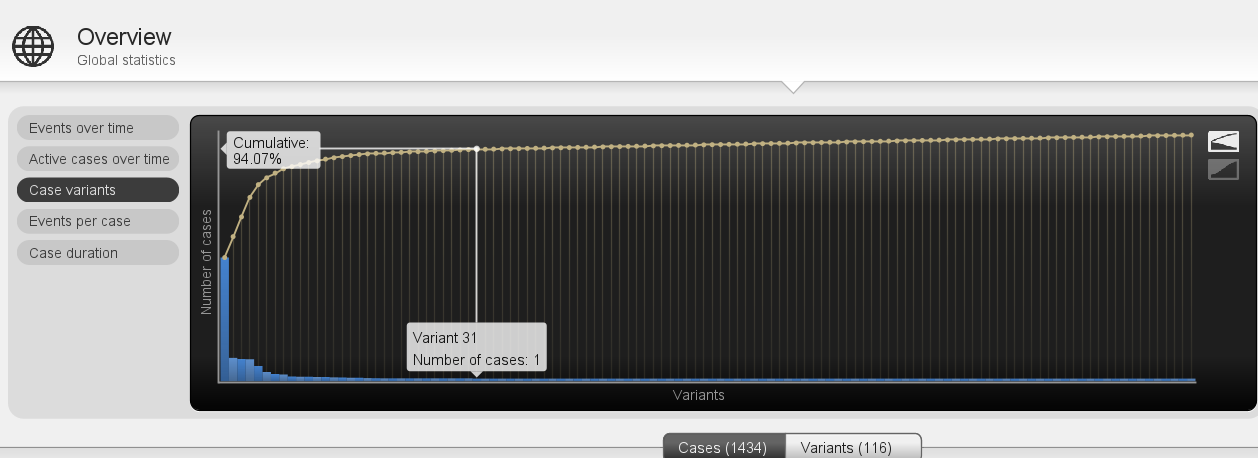
From above figure:

Events = 8577

Cases = 1434

No. of events per case = 8577 / 1434 = 5.981

1. **Can you indicate whether each case seems to be unique or whether many cases follow the same activity sequence?**



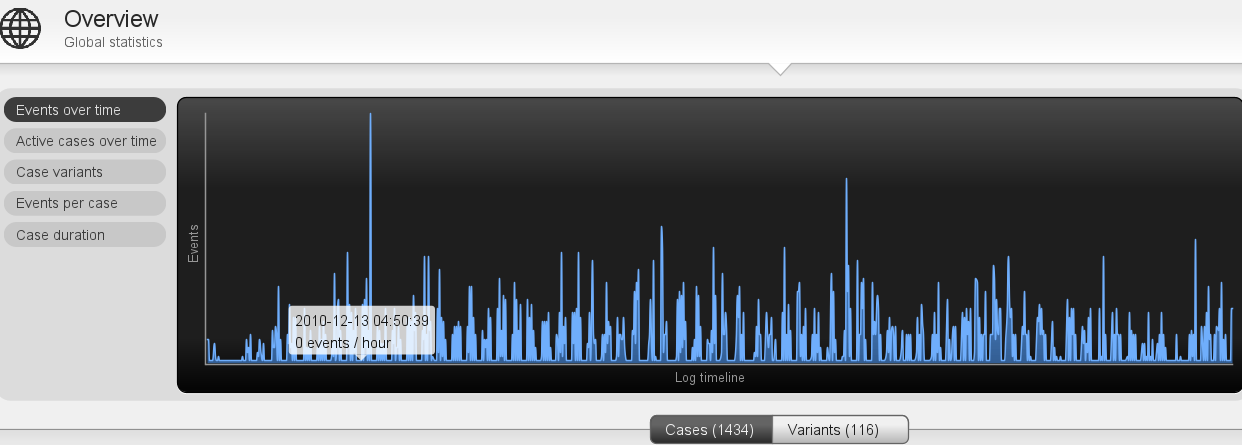
From above figure:

Total No. of variants = 116

From Variant 31 to last Variant, number of cases = 1 i.e. unique cases = 116 - 30 = 86

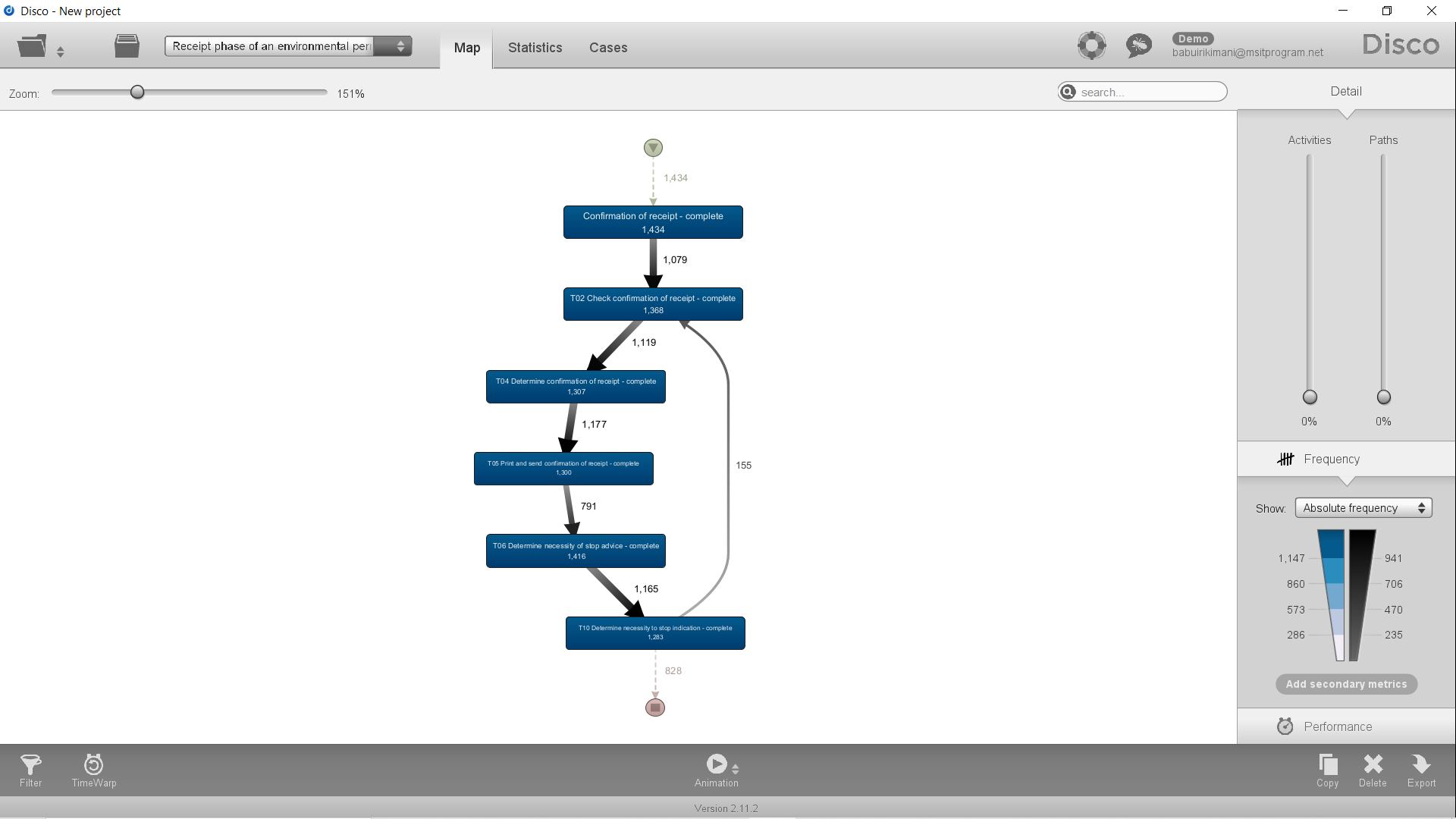
Remaining cases follow the same activity sequence (Variant 1 to Variant 30)

1. **What is the main observation that can be made from the 'Events over time' graph?**

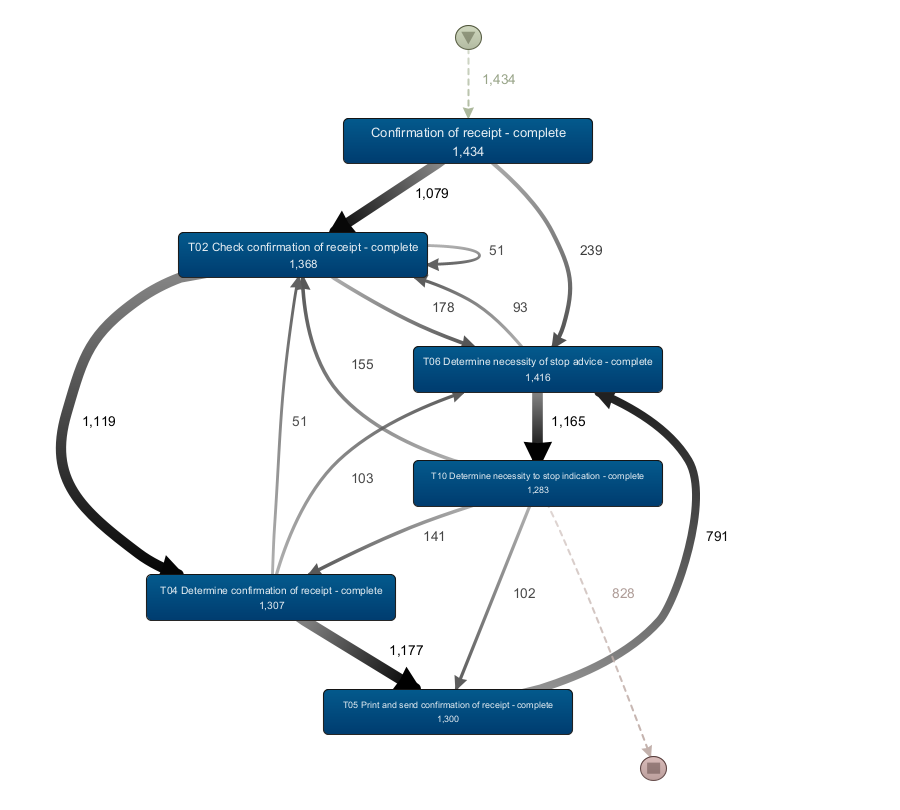


Most of the days, Number of events per hour is 0. Apart from some outliers, others are having nearly same number of events per hour.

While still in Disco, switch to the 'map' view to display a process map.

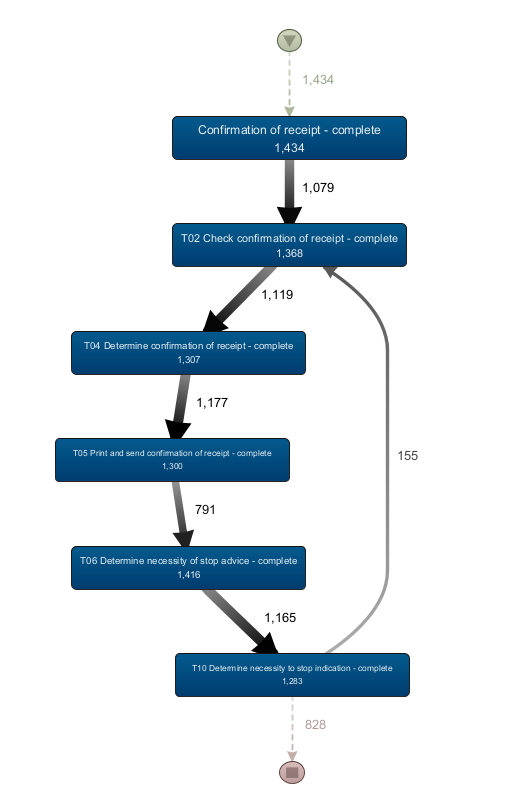


Using the map view, change the activity and path detail settings in order to create a comprehensible process map (e.g. a process map that could be printed on one A4 or letter paper or shown on a single computer screen while still being readable in full).

1. **Discuss this process map, what is the main process?** 

T02 Check Confirmation of Receipt - Complete is the main process(Activity 0%, Paths 20%)

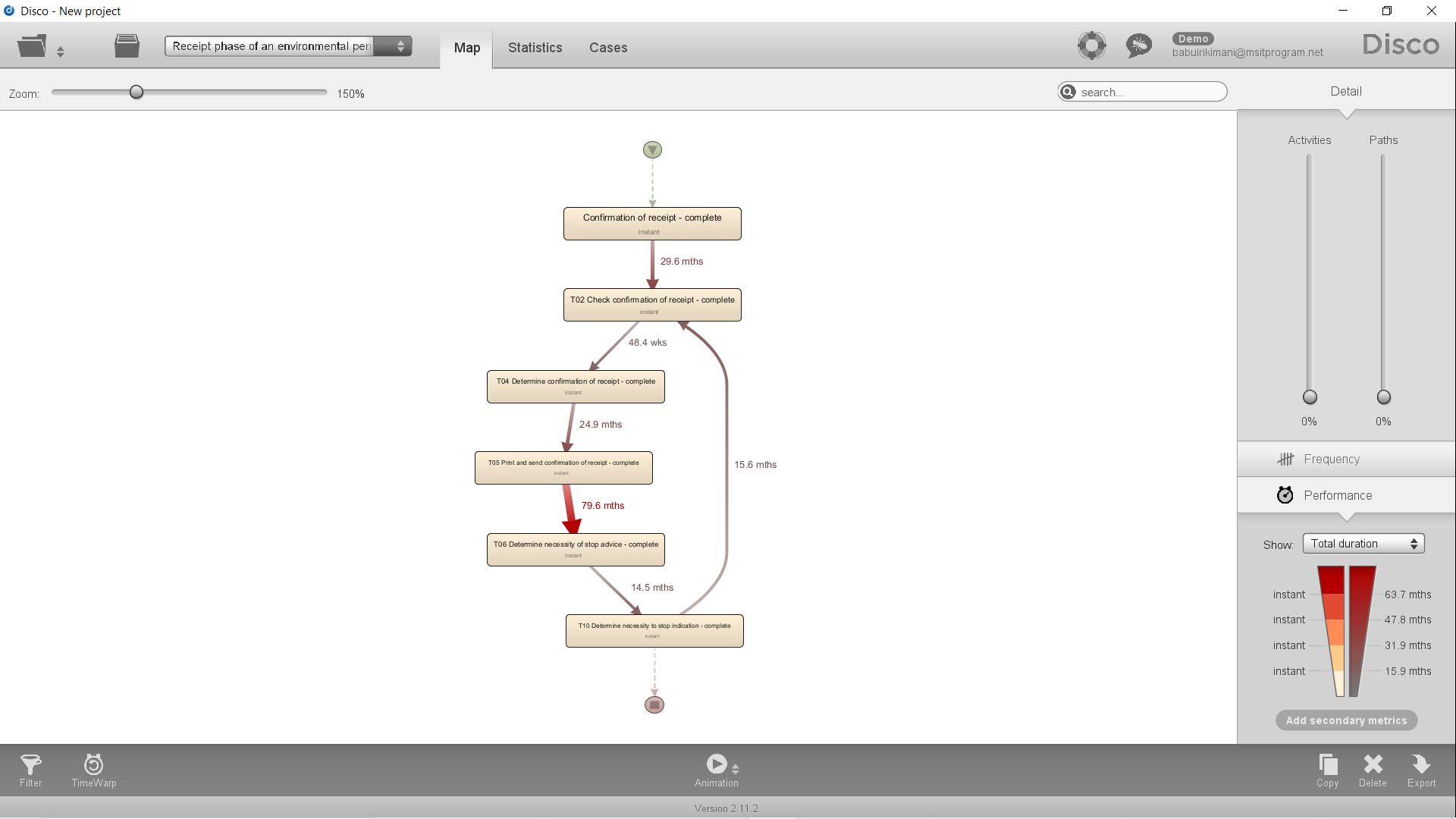
1. **Which activities and paths between activities are frequent?**



Activity 0% and Paths 0%

T02, T04, T05, T06 and T10 are frequent

While still in Disco, and while using the same process map (e.g. do not change the activity and path settings), switch to the performance projection.



**Discuss where the process takes most time, e.g. where there are possibilities for improvement. Relate these times (of the bottlenecks) to the time spent in other parts of the process. In other words, discuss how severe the bottleneck is with respect to the time spent on other activities.**

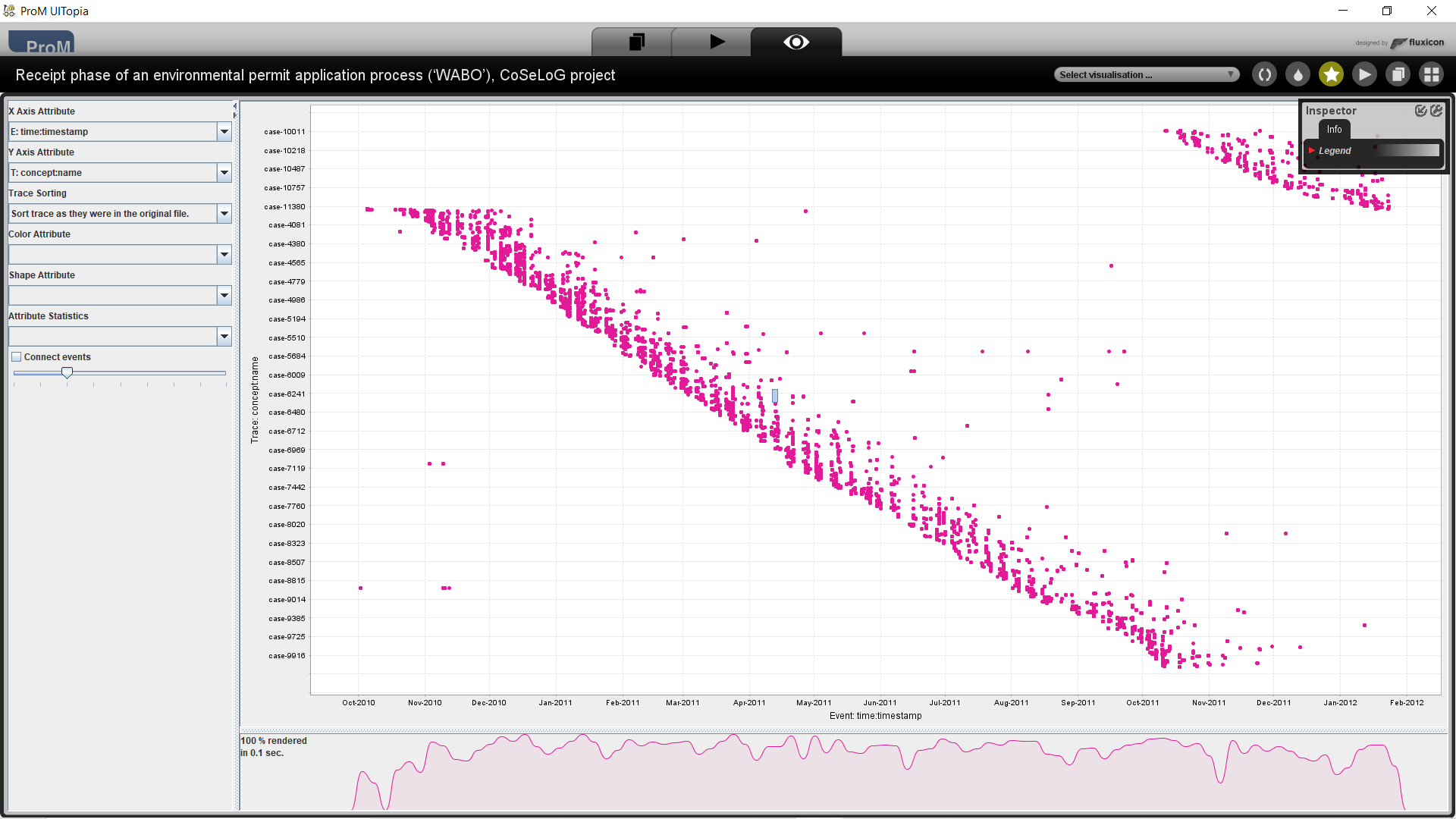
**Also explicitly mention the performance metric chosen (e.g. total, mean, median, or max) and why you have chosen this setting.**

The performance metric chosen is Total Duration to know about the time for each and every process. There is a bottleneck from activity T05 (Print and send confirmation of receipt) to T06 (Determine necessity of stop advice) i.e. 79.6 months which is more time taking compared to other processes. It takes more time to complete which results in delay of other processes.

**PROM**

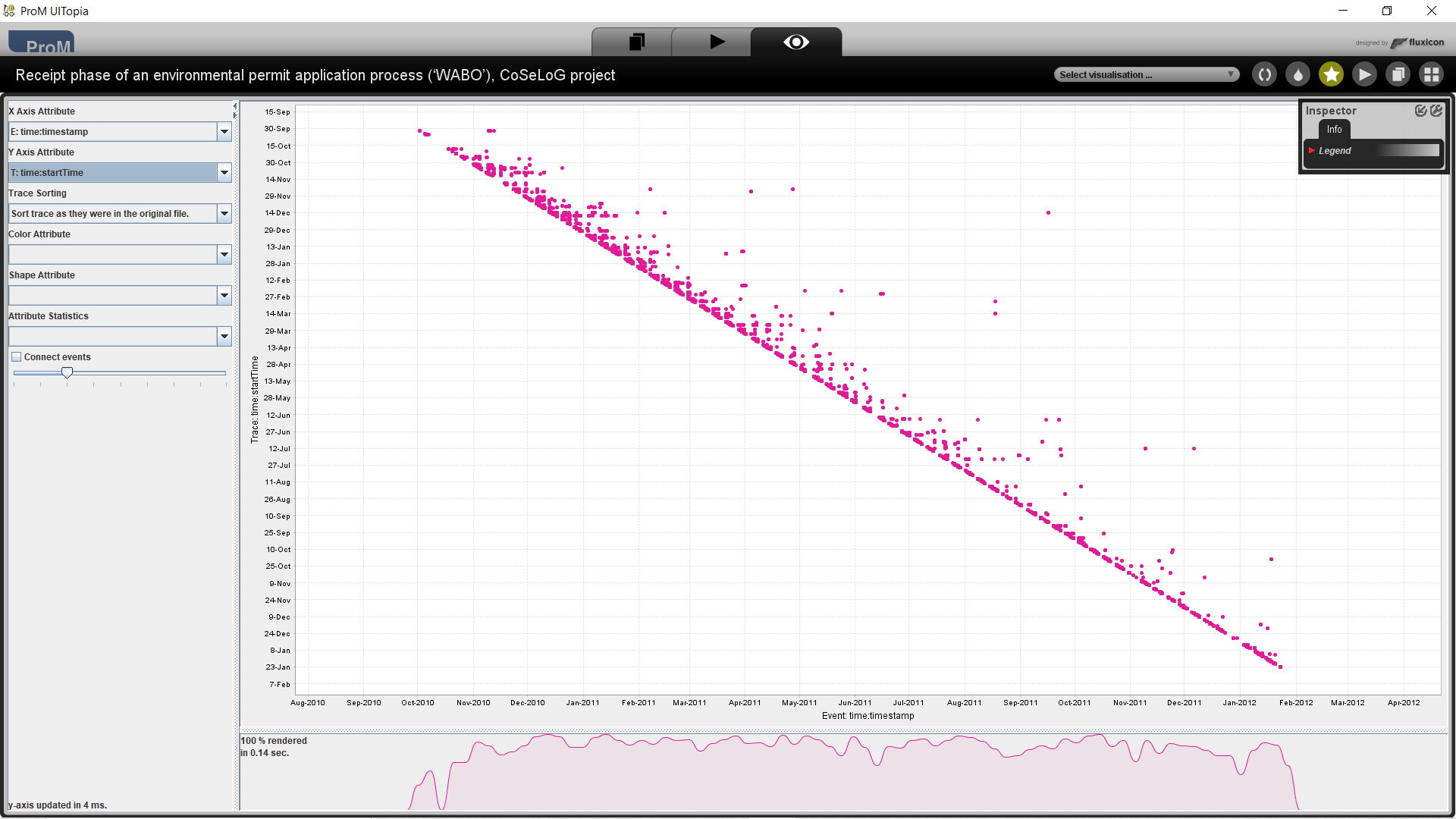
ProM is a very extensive tool which allows for dozens of different model types. It allows for many different input formats and it allows for many different types of process mining.

Now load the original event log in ProM. Visualize the event log using the Dotted Chart or XDottedChart visualizer (by pressing the 'eye'-icon with the event log selected and switching to the Dotted Chart or XDottedChart visualizer).



Using the Dotted Chart, answer the following questions:

1. **Is the arrival rate of new cases constant? If not, when are there fluctuations? If yes, how can we see this from the Dotted Chart?**



From above figure we can say that the arrival rate of new cases constant

1. **Can you observe a change in the global process?**

As there is only one colour in the dotted chart so there is no change in the global process

You are now asked to discover a Petri net on the event log. However, the unfiltered event log results in an incomprehensible Petri net. Therefore, you are allowed to run the 'Filter log using simple heuristics' plug-in *once* on the original event log to discover a Petri net on the filtered event log.

1. **Clearly indicate which settings you have used for the 'Filter log using simple heuristics' plug-in.**

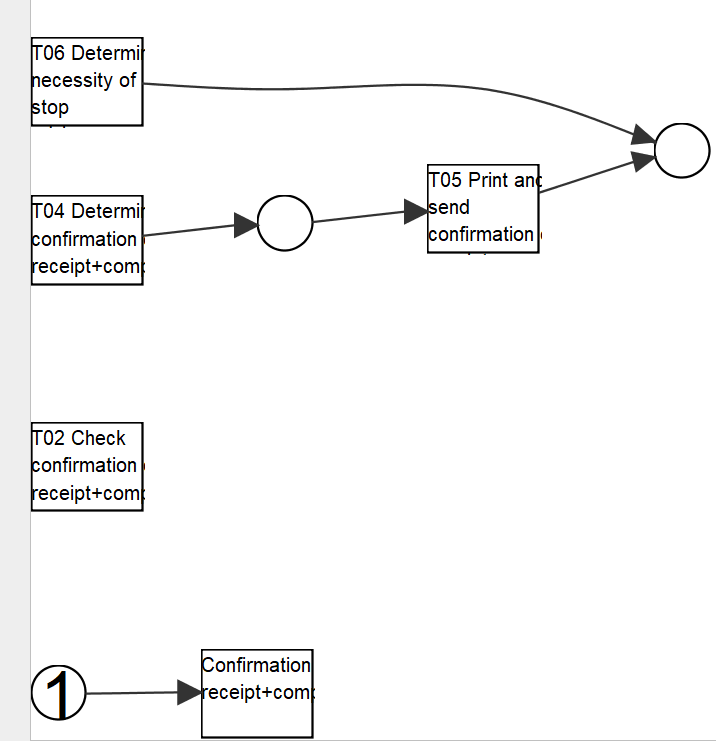
I have followed the default settings to convert the event log to filtered event log using plug-in called ‘Filter log using simple heuristics’

1. **Explicitly motivate the filtering settings chosen, why did you pick this percentage or selection of activities?**

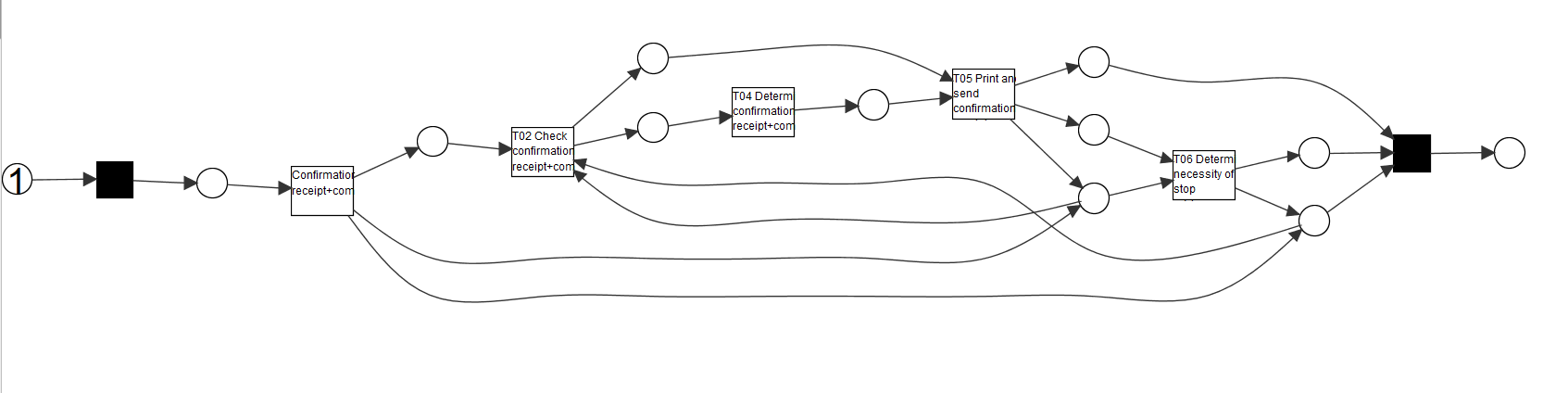
Based on the important activities from the process map of the same event log.

1. **Discuss and argue which plug-in (or chain of plug-ins) you have used to discover a Petri net, for instance by comparing two or more plug-in results and arguing why one of the Petri nets is better.**

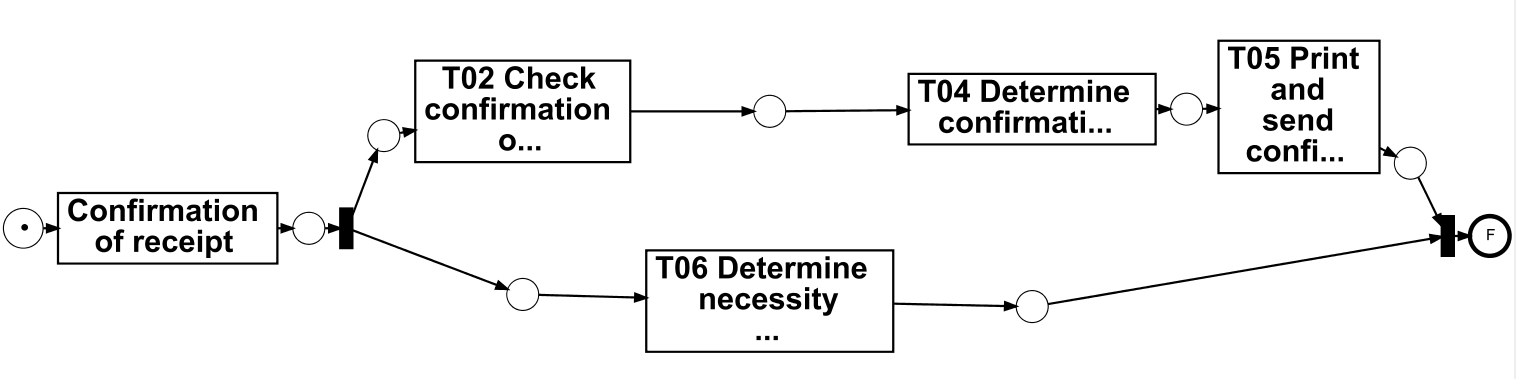
Petri Net using Alpha-algorithm:



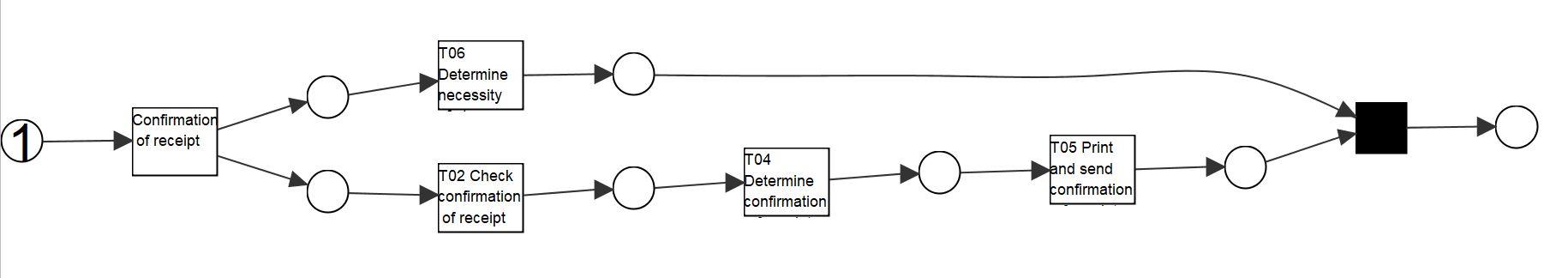
Petri Net using ILP:



Petri Net using Heuristics Miner:



Petri Net using Inductive Miner:



Alpha Miner is not well connected.ILP, Heuristics and Inductive Miner are well connected. But ILP is more elaborated.

1. **Explain the (best) Petri net: what is the main process and what are notable parts of the Petri net?**

The best Petri Net is done with Inductive Miner as it is structured, concise, precise and well connected

The main process is T02 Check confirmation of receipt

The notable parts of the Petri Net are:

Confirmation of receipt

T02 Check confirmation of receipt

T04 Determine confirmation

T05 Print and send confirmation

T06 Determine necessity

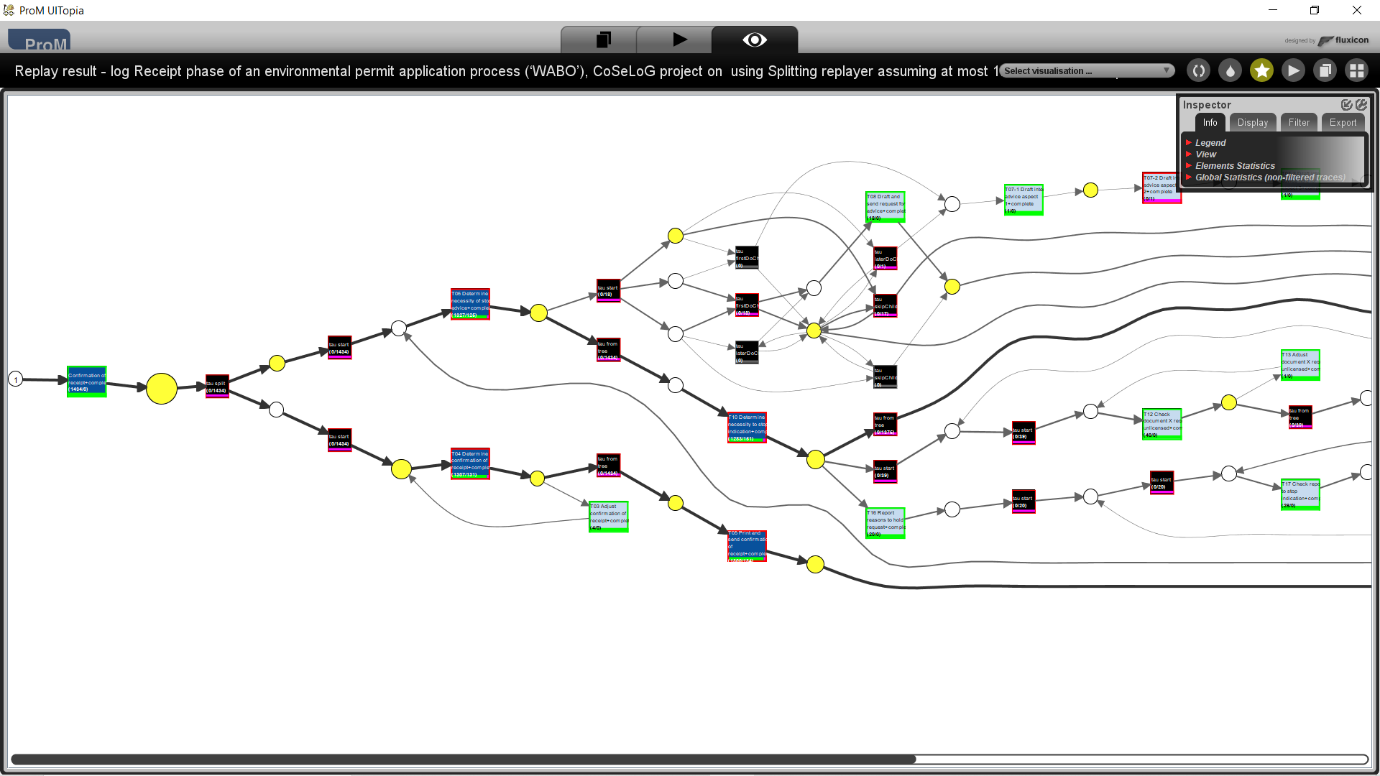
Note that this question requires you to experiment with different filtering settings and discovery plug-ins. You are not required to describe *everything* you have tried but found unsuccessful. Only describe the successful combination of plug-ins and its result(s) and argue why your final result is 'good'.

**Suggested list of plug-ins or plug-in chains to produce a Petri net:**

* Mine for a Petri Net using Alpha-algorithm
* Mine for a Petri Net using ILP
* Mine for a Heuristics Net using Heuristics Miner *followed by* Convert Heuristics net into Petri net
* Mine for a Petri net with Inductive Miner

The organization has a process model that describes the 'should be' process (i.e. a normative process model). Load the file 'normativeModel.pnml' into ProM and apply conformance checking on this process model, and on the full unfiltered original event log.

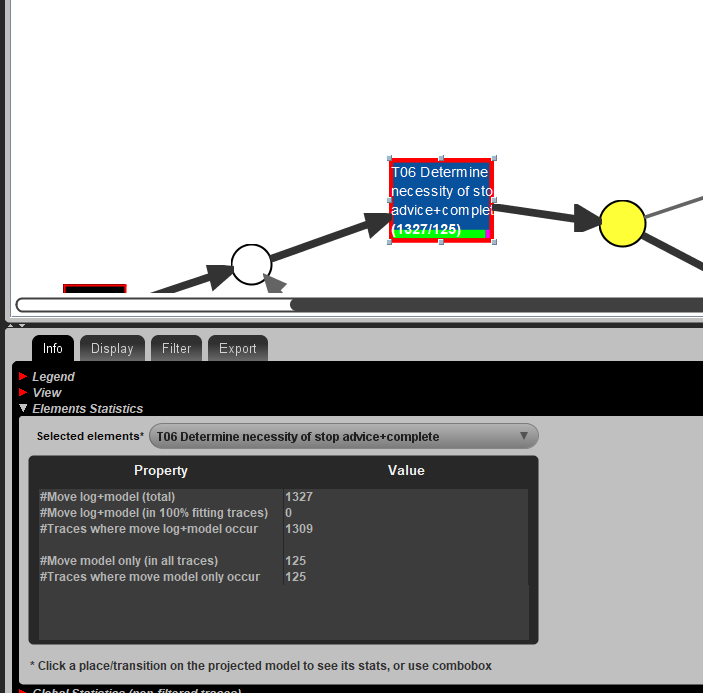
1. **Include a screenshot of the part of the normative process model, with the conformance information projected onto it, that shows where most of the deviations occur.**



1. **What is the replay fitness (the 'trace fitness' statistic) of the event log on the normative process model?**

Replay fitness is nothing but the resulting model can reproduce the behaviour in the log.

1. **Select the transition 'T06 Determine necessity of stop advice+complete' (on the top left of the model) and discuss its element statistics: how many times is the transition executed correctly and how many times incorrectly?**



Out of 1327 times, 1309 times transition executed correctly

1. **Using the element statistics of transition 'T06 Determine necessity of stop advice+complete', what can you say about the (in)correct execution of this activity?**

Total move log+model traces = 1327

Due to incorrect execution of this activity there is red colour when compared to green colour for other activities

**Instructions to align the process model with the event log:**

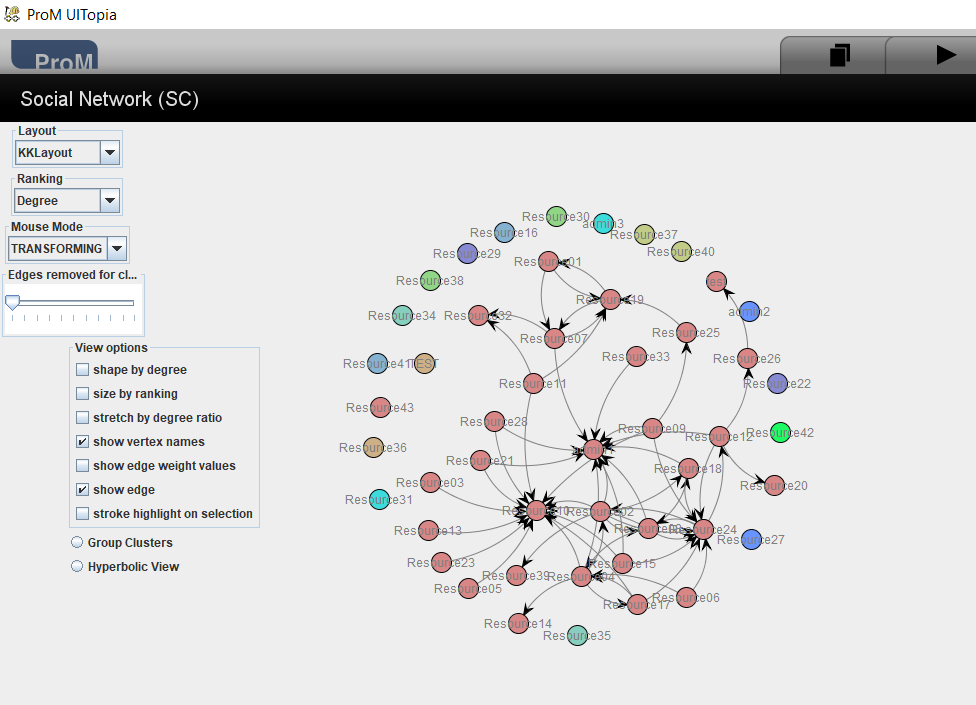
1. Import the normative model using the 'PNML Petri net files' importer.
2. Select the imported normative Petri net and the event log, start the plug-in called ‘Replay a Log on Petri Net for Conformance Analysis’ (not the variant with performance!), and click 'yes' in the 'No Final Marking' pop-up.
3. Select the 'sink' place on the left (note: do not select '0-sink' etc.) and click the button 'Add Place >>' to add the place 'sink' to the candidate final marking list. Now click 'Finish'.
4. Click 'Finish' in the mapping wizard.
5. Click 'No, I've mapped all necessary event classes' to indicate that some events are not present in the normative model.
6. Now click 'Next' and 'Finish'. The normative process model is shown with conformance information projected onto it.

If you followed these instructions exactly you do not need to mention these steps in your answer.

More information regarding this conformance technique is provided in lecture 4.7: 'Aligning observed and modeled behavior' (and to a lesser extend in the lectures 4.3 through 4.6).

The final analysis you have to perform on the original event log is a resource analysis, e.g. looking at the user behavior in the event log.

1. **Use the plug-in 'Mine for a Subcontracting Social Network'. Note that subcontracting means that if individual *j* frequently executed an activity in-between two activities executed by individual *i*, then individual *i* subcontracted work to individual *j*. Answer the following question using this view: Can two or more groups of users be distinguished? Explicitly discuss the settings you have used in the resulting visualization.**



1. **Again use one of the two Dotted Chart plug-ins. For the XDottedChart change the component type to 'org:resource'. If you use the Dotted Chart visualizer change the 'Y Axis Attribute' to 'C: Resource classifier' and the color attribute to 'C: Activity Classifier'. Answer the following two questions using this view:**

****

1. **Are all users executing activities from the start of the event log, or are some users joining later?**

Not all the users are executing the activities from the start of the event log.

1. **Are users mainly executing particular activities or are most users executing most of the activities?**

Most of the users are executing most of the activities.