PROCESS MINING DATA REQUIREMENTS

3 key ingredients

The event log must contain at least 3 main elements: a case id, an activity and a timestamp column. The case id identifies the single process instance, the activity describes what happened and the timestamp traces when a given activity took place in this specific instance of process execution.

Process mining is a family of techniques and algorithms for process discovery and analysis. The main goal of process mining is extracting information about processes from data gathered in event logs. To perform a successful process mining project, the event log must meet some fundamental requirements in terms of structure and content.
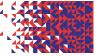
Data Structure

In order to show a realisitc process visual the event log should have one column with different case ids. There should be different activities and timestamps for each case id. Timestamps ought to have the same pattern for all process instances.

Project Goal

You should always keep the purpose of your project in mind. The data should contain dimensions corresponding to the main purpose of the analysis. Don't forget to add dimensions describing the data such as regions, warehouses, customers etc.

Human Readable Data

Process mining projects are cross-departemental and connect stakeholders coming from different backgrounds. It is very important that all data is not only accurate but also understandable to all participants.