Data Analytics

Practice 7 – Advanced Use cases

Time series analyzation | Process mining

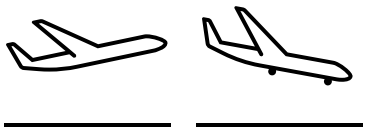
Tasks:

1) Time series analyzation

Context: You work in an airline, and you collected data about the passengers per month from 1949 until 1960. You have a meeting later that day with your boss.

Note: All files can be found within the Moodle-course. You are free to use any tools you think that are suitable for the following tasks (PowerPoint, Python, etc.).

a) Have a look at the dataset you were provided with. Are there any anomalies in this dataset? Can you find any trends in the dataset?
b) Are there any seasonal trends you can make out? How does the month impact the number of tickets sold?
c) Now think about a role in the airline you want to represent (Head of sales, Fleet manager, etc.). Also think about a message you want to communicate to your boss, to improve the performance of your department in the coming year.
d) Create a meaningful visualization of your message (Charts, text, slides (max. 2-3), etc.). Try to also include an explanation on why you think this trend is going to continue, based on the previous years. In the end your message to your boss should clearly stand out.



2) Process Mining

Your company is selling tickets for music events. However not every customer is able to participate at a given event. To get their ticket refunded they must open a new refund request in your support section. This request will be processed by multiple employees and all events will be documented in a log file together with a timestamp and a cost for said action.

You are provided with such a log file and your manager asks you to create a model of this process.

Note: The “ticket_process.csv” can be found within your Moodle-Course

a) Note down all possible states a request can go through. Are they following a certain pattern? What states of the process can be done parallel and which ones must be sequential?
b) Try to find out how many times each state has been run through. Also note down how many times a state has transitioned to another step within the process. (Note: The beginning of the process and the ending of the process should both be included in your observations)
c) Based on your observations in tasks a) and b) try to create a visualization in form of a chronological model (for example with BPMN) of this process. What is our biggest time loss and how could we improve this process?