

Graph Analytics Project Proposal (2021/22)

Guidelines

The project is a group activity and counts 40% towards the final mark. The project focuses on a business problem presented by Nokia, which is described below and requires you to use the methods learned throughout the semester to make the inference of a Graph and provide analytical insights. The delivery deadline is the 31st of May. Deliveries include:

1. **Report**, 5-page document that summarizes the main findings and. The report should include the following sections: 1) Introduction; 2) Data & Methods; 3) Results; 4) Discussion; 5) Conclusions; 6) Bibliography.
2. **Presentation**, a 10-minute pitch about the outcomes of the project. Groups must share the presentation on Moodle by the delivery deadline.

Nokia will support the project. During Lecture 6, members from Nokia will be available during the last 15 minutes of Lecture 6 for a Q&A session focusing on questions you might have about the data, problem context, or expected outcomes. For any questions regarding the project's dataset and business context, reach out to Mariel (mariel.barbachan@nokia.com).

Finally, the deliveries will be assessed by the IMS teaching staff and Nokia officials. Relevant elements for assessment include the novelty of the application, clarity of presentation, relatedness to the curricular unit, and business relevance of outcomes.

Below you can find a short description of the business context, available data, and the project-specific details.

Introduction

Order to Cash (OTC) is the business process that contemplates all events from the order processing system, from receiving the order to payment. Currently, Nokia processes and stores OTC records from operations in 180 countries. The Figure below shows an example of an OTC high-level path.



The analysis of OTC records offers an opportunity to extract powerful business insights at different scales. For instance, at a higher level, it allows to map similarities between countries, partners, and

clients depending on their purchases and OTC processes generated, allowing the identification of commonalities and new opportunities (e.g., recommendation of new products to clients). At a lower level, it provides a fine-grain track record to compare business processes in different countries, involving other partners, and identify inefficient procedures that could represent cost savings and better service.

Data availability

The data consists of one document with data in tabular format. The document includes information covering one year of activities from Celonis, containing the process's ID, the action, and a timestamp.

Table 1 Sample Activity

| ACTIVITY | TIMESTAMP | SALES_DOC | SALES_DOC_ITEM | COUNTRY |
|------------------------------------|---------------------|------------|----------------|---------|
| Record Goods Issue | 2021-04-08 15:17:29 | 1127382723 | 000010 | PT |
| Create Delivery | 2021-04-06 11:44:53 | 1127382723 | 000010 | PT |
| Create Intercompany invoice | 2021-04-09 14:10:22 | 1127382723 | 000010 | PT |
| Create Invoice | 2021-06-29 23:06:07 | 1127382723 | 000010 | PT |

Moreover, students will have access to data for a full year of OTC processes in the following list of countries:

- Portugal estimated length of 1748578 rows
- Germany estimated length of 6291969 rows
- Brazil estimated length of 1837252 rows
- Thailand estimated length of 2755592 rows

The data is available through Moodle and is organized in one document per country.

Project Description/Business Problem — The OTC Process Graph

OTC processes can be conceptualized as a weighted directed graph, where nodes represent the possible events and edges the expected (i.e., average) time transitions between events. Two critical efficiency metrics are the rework rate (r) and the overall throughput time (t). The former

corresponds to cycles and loops in the graph, while the latter corresponds to the average time taken to complete the process.

Considering the business impact of analyzing the two abovementioned efficiency metrics from the OTC graph, can you characterize its structure and motifs (e.g., the existence of cycles and their context in the overall graph) and derive business insights for the improvement of the process?

The following is a list of project narrative suggestions that can be observed in this project:

- Find the optimal process paths as a subgraph representing the process with the lowest possible overall throughput time from beginning to end. Characterize these processes and quantify their relevance. What insights or recommendations for improvement can you deliver?
- Compare the efficiency of the OTC process graphs between a group of countries. Can you quantify the similarities and differences and derive valuable insights?