Assignment for Data Analyst position at Stora Enso

You’ve received a two .csv files as an attachment containing sample data from a hypothetical procurement process (see section The Data for details). Your task is to use this data set to answer the questions below and prepare a report of your findings.

Please return your report as a PDF file to [juha.byman@storaenso.com](mailto:juha.byman@storaenso.com) and attach all your code and plots. Please send your plots embedded in the report or separately in PDF or PNG format.

The Data

The database contains two tables: CASE\_TABLE and EVENT\_TABLE. CASE\_TABLE contains invoice line item info, EVENT\_TABLE contains all the event logs for invoice line item. All the tables contain data for the purchasing documents that were generated in year 2017.

We’d prefer that you complete the assignment using SQL, R, and/or python but feel free to use any tools that you think are suitable for the job.

Task 1.

* Analyse order channels, what is the share of invoices and spend per channel. Visualise and explain your observations

In Table CaseTable, 1st col is invoice ref. number, 2nd col is buyer, 3rd col is sidn, 4th col is seller, 5th col is item, 6th col is quantity, 7th col is channel. Here should upload the table, check how many and what kind of channels there are. Calculate # sum of invoices per channel, # of spend=5X6 per channel. Visualize as chart.

* Analyse lead time for our procurement process. What are your observations of the results?

In Table EventTable, 1st col is invoice number, 2nd col is stages of invoice, 3rd col is date and time, 4th col is buyer. Check time between stages of each invoice, sum them up and get the total length. Calculate statistical means, max, min (extract zeros).

Task 2.

* Open question. Did you find anything interesting in the dataset? Would you provide any recommendation how to optimize the process?

Check which vendors have quickest and shortest procurement process

Check which vendors are the most frequently used, and which get the most contracts.

Task 3.

* Open Modelling, python, etc -file and answers four questions.