SAUDE +

PROCESS AND BUSINESS ANALYSIS

Dinis Fernandes – 20221848
Dinis Gaspar – 20221869
Luis Mendes – 20221949
Sara Ferrer – 20221947

Introduction

Saude+ is a mid-sized private medical clinic. The clinic has recently expressed concern over some of their key performance indicators. Administration is worried that their patients are receiving worse service due to inefficient processes. Saude+ is also conscious that it may be providing unequal care to patients based on patient profile items like age, region, insurance state or gender. For these reasons, Saude+ has brought in a Process mining team to dive deep into the inner workings of its operations with the goal to evaluate and manage these concerns.

Theoretical Background

Companies are collecting all that they can inside of their business, that includes the process of serving clients and all that happens in the back end to have the customer satisfied (something that doesn't always occur), to perform analysis on these millions of records manually nowadays has become extremely difficult. That is why solutions like Celonis have gained such importance to automate this workflow of continuous business analysis performance, called process mining. To do it is all that is needed is the event log, a table that usually comes from the ERP system of the companies that has a Case ID (to identify each flow of interactions that the customer had with the business), an Activity (the name of each activity that is performed along the way), and a timestamp (to register when it happened); this table can be combined with more tables to add more information about the customer or process; it is processed and loaded into the system in a process called ETL (Extract, Transform and Load) and the software starts to work.

After the table is uploaded the software starts to discover the relationship between activities to understand what the flow of the current process is, aiming to generate the as-is model, to do so it uses an improved version of the α-algorithm. With the as-is model it is possible to analyze the current workflow and see the flaws of the business that should be improved in a process called conformance checking, finalizing with a root cause analysis that should reveal what is causing the deviation inside the flow (these causes are just correlation, not the actual causes of the problem, to conclude that an in-depth analysis is needed but its not able to be performed with our current data. correlation ≠ causality).

All will be done with post-mortem data (data from the past), inside the Celonis software, finalizing with a dashboard analysis to visually represent what is happening and responding to the manager's concern about the clinic.

Project Goals

This project aims to perform a comprehensive analysis of Saúde+, a mid-sized private medical clinic, and their current operations. The clinic's current processes will be evaluated with the goal of finding inconsistencies and deviations from the ideal process. This information will give context to the clinic as to which areas require revision as they are currently leading to a worsening of service efficiency and subsequent loss of service quality, which will ultimately affect the clinic's long-term ability to remain relevant.

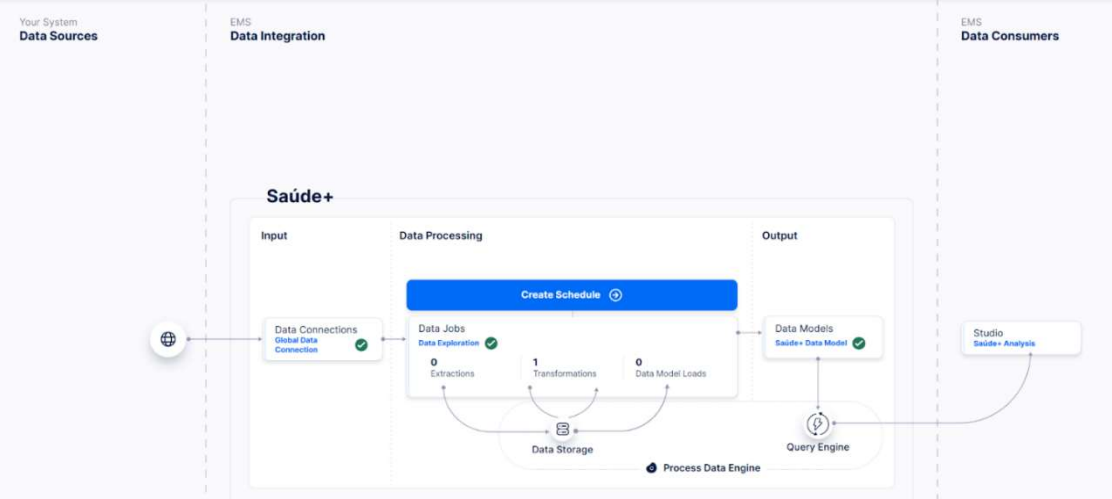
Additionally, this project will also evaluate whether the clinic is providing unequal service to patients based on customer profile. We will analyze whether older patients are less satisfied with the service and additionally if they are potentially struggling with the clinic's digital systems. Patients' regions will also be investigated to check for discrepancies in the care given to patients of different regions, for example if patients from the islands have more struggles with system errors. An assessment of the impact of patient gender will also be performed to discern if a patient's gender can lead to them being offered differing levels of service. Lastly, we will also assess whether insured customers are getting better care than uninsured customers.

These are concerns that Saude+ management has expressed to our team. To help them keep track of these concerns, in addition to the analysis contained in this report, we will provide them with a dashboard developed using Celonis. Within this dashboard, management will be able to look at their processes and their variants. Additionally, we have created dashboards containing KPIs and visualizations which managers can use to obtain context on the current situation of the clinic and use this context to support business decision-making.

ETL

In ETL we first extract the data, confirm and perform preliminary analysis and then load the data model to create dashboards with more in-depth analysis.

The first step is the connection with the Data Source, where we loaded both datasets, the Case table, with information about every case, such as gender, region, insurance, etc, and the other dataset, the event log, that comes from the ERP of the company.



The next step was to add a data Job named Data Exploration to help us do an initial analysis and check the data.

```
-- 1) Initial visualizations of the tables
Select * from "Case_table";
Select * from "Event_Log_Final_1";

Select count(*) from "Event_Log_Final_1";

-- 2) Primary key cheking
Select "Case_table"."CASE_ID"
From "Case_table"
Where "Case_table"."CASE_ID" is null;

Select "Event_Log_Final_1"."CASE_ID"
From "Event_Log_Final_1"
Where "Event_Log_Final_1"."CASE_ID" is null;

-- 3) Missing values
Select count(*)
From "Case_table"
Where "Case_table"."PATIENT_AGE" is null
      or "Case_table"."PATIENT_GENDER" is null
      or "Case_table"."PATIENT_REGION" is null
      or "Case_table"."HAS_INSURANCE" is null
      or "Case_table"."TOTAL_COST" is null
      or "Case_table"."SATISFACTION_SCORE" is null
      or "Case_table"."NUM_PREVIOUS_APPOINTMENTS" is null
      or "Case_table"."CANCELLATION_HISTORY" is null;

-- Only column that allow (and have) missing values is chronic_condition
Select count(*)
From "Case_table"
Where "Case_table"."CHRONIC_CONDITION" like 'None';

Select Count(*)
From "Event_Log_Final_1"
Where "Event_Log_Final_1"."ACTIVITY" is null
      or "Event_Log_Final_1"."TIMESTAMP" is null;

-- 4) Case ID analysis
Select count(*), count(distinct("Event_Log_Final_1"."CASE_ID"))
From "Case_table" join "Event_Log_Final_1"
      on "Case_table"."CASE_ID" = "Event_Log_Final_1"."CASE_ID";

Select Count("Case_table"."CASE_ID")
From "Case_table"
Group by "Case_table"."CASE_ID"
Having Count("Case_table"."CASE_ID") > 1;

with count_table as (Select "Event_Log_Final_1"."CASE_ID", Count("Event_Log_Final_1"."CASE_ID") as n_cases
                     From "Event_Log_Final_1"
                     Group by "Event_Log_Final_1"."CASE_ID")
Select min(n_cases), round(avg(n_cases),2), max(n_cases)
From count_table;

Select count(distinct("Event_Log_Final_1"."ACTIVITY"))
From "Event_Log_Final_1";
```

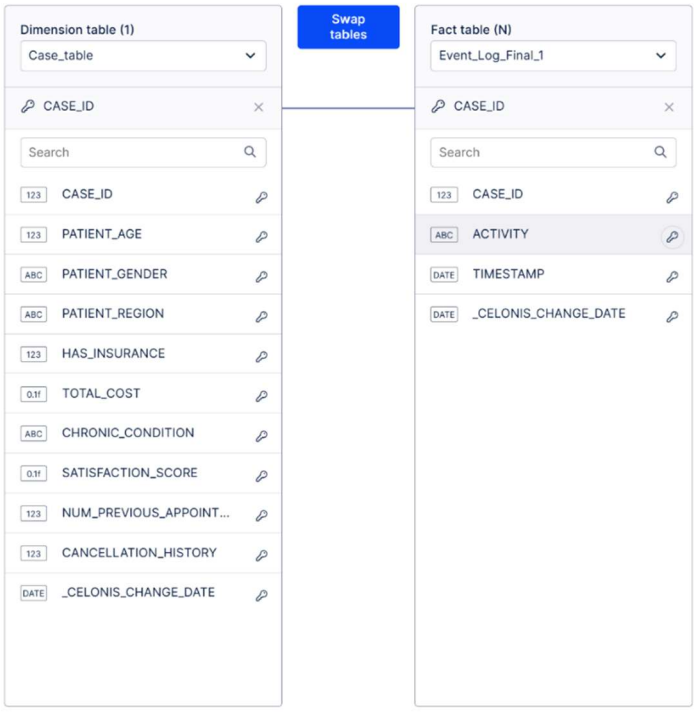
The process was:

1. See the composition of each of the tables.
2. Check if primary keys do not have missing values.
3. Check if there are no missing values in any column both tables.

We concluded that there are 5055 activities, 15 different unique activities and 1190 cases. There are no repeated cases. In a single case the minimum number of activities is 2, the maximum is 8 activities, and the average is 4,25 activities.

We would like to note that we disabled this transformation, so it does not run every time we load the data model.

The next step was to create a data model with the tables, the data types of each variable were correct, so nothing was changed, then we join the case/facts table (“Case_table”) with the event/activity table ("Event_Log_Final_1") on the “CASE_ID”, creating this data model:

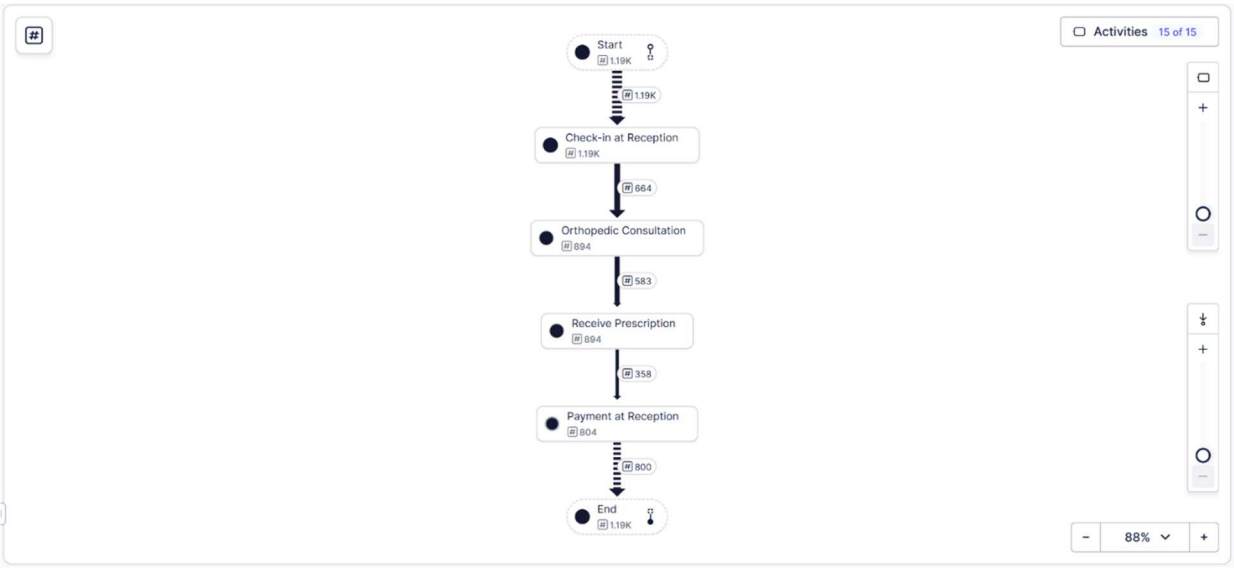


After these checks and analysis, we are ready to start a more in-depth analysis and to create our dashboards to present to the management of Saúde+.

Dashboard Analysis

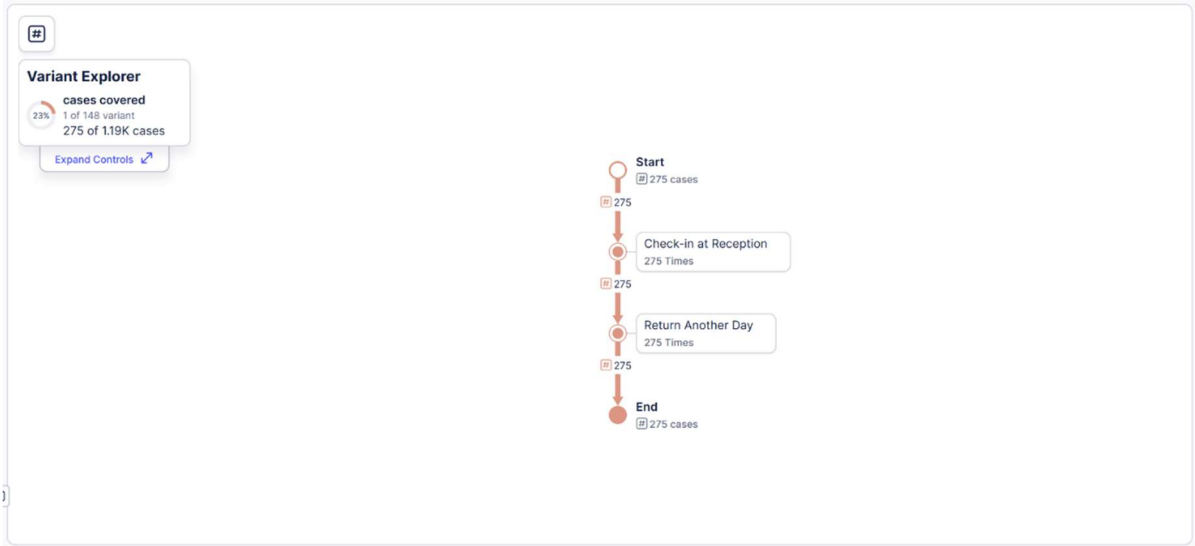
Process Discovery

Process Explorer:



- All cases start with the check-in at reception, as expected.
- Orthopedic consultation only happens in 75% of cases.
- The patient made a complaint in 16% of cases.
- Only 68% of cases end with payment at Reception
- Possible undesired activities:
    - Return Another day (23%)
    - Waiting for doctor (6%)
    - Missed turn (2%)
    - Complaint (16%)
    - System error (only 5 cases)
    - Consultation cancelled (4%)
    - Incomplete Information (only 10 cases)

Variant Explorer:

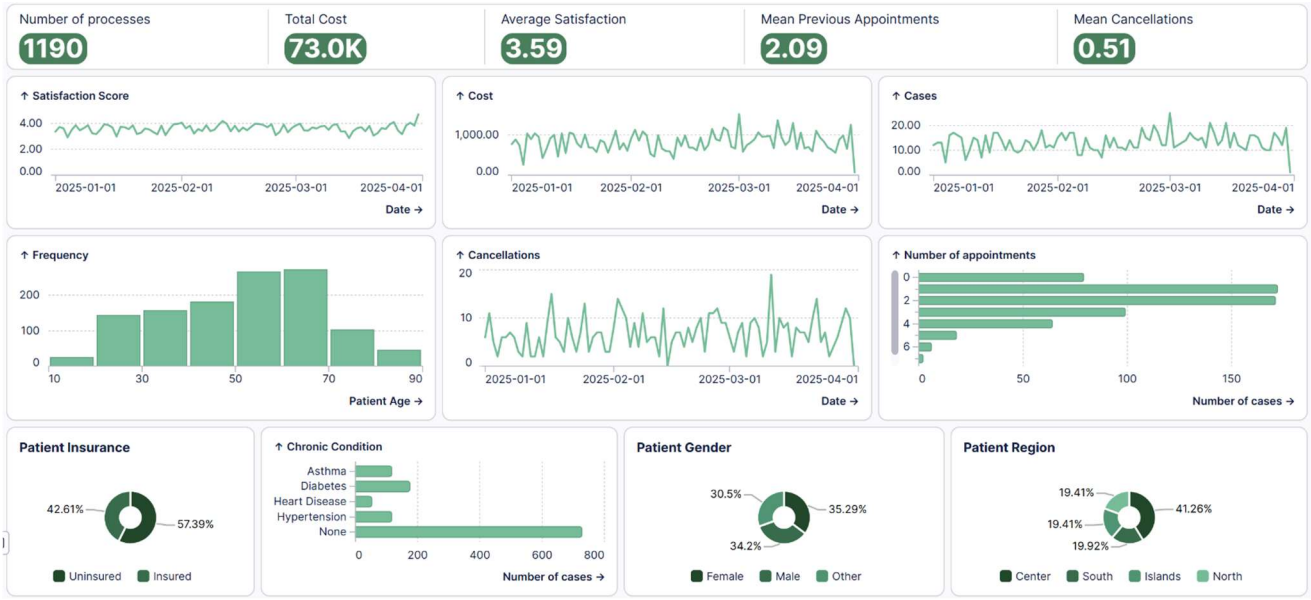


- There are 148 unique variants of the process.
- We have 1190 cases.
- The “Happy” Path, the most common process variant which accounts for 23% of all cases, showcases a process flow where the patient checks in at the reception and then has to return another day, and thus has the shortest throughput time of all projects at just 6 minutes. This appears to indicate some sort of issue with fulfilling the appointment.
- The second most common variant, which represents 17% of cases, shows what is an expected path of a patient at the clinic (check-in, consultation, prescription and payment) and has an average throughput time of 42 minutes.
- All other variants only account for a maximum of 5% of cases each.
- There are variants with throughput times as long as 2 hours, which model patients with exams, complaints, etc. which add a lot of time, whereas the shortest is 6 minutes and models situations where the patient actually has no consultation, either because they miss their turn or because they get told to return on another day (happy path) and thus have to spend minimal time at the clinic.
- There are variants where patients check in and receive exam results right away, without performing an exam or where they perform exams and don’t receive the results before leaving.
- There were 21 cases where the patient missed their turn.
- Variant 7 displays a path where the patient receives a prescription without going to the appointment first.

Generic Conclusions:

- Most variants of the process follow logical paths, even if undesirable.
- There is an inconsistency in some (49) cases where the consultation is cancelled after it happens.

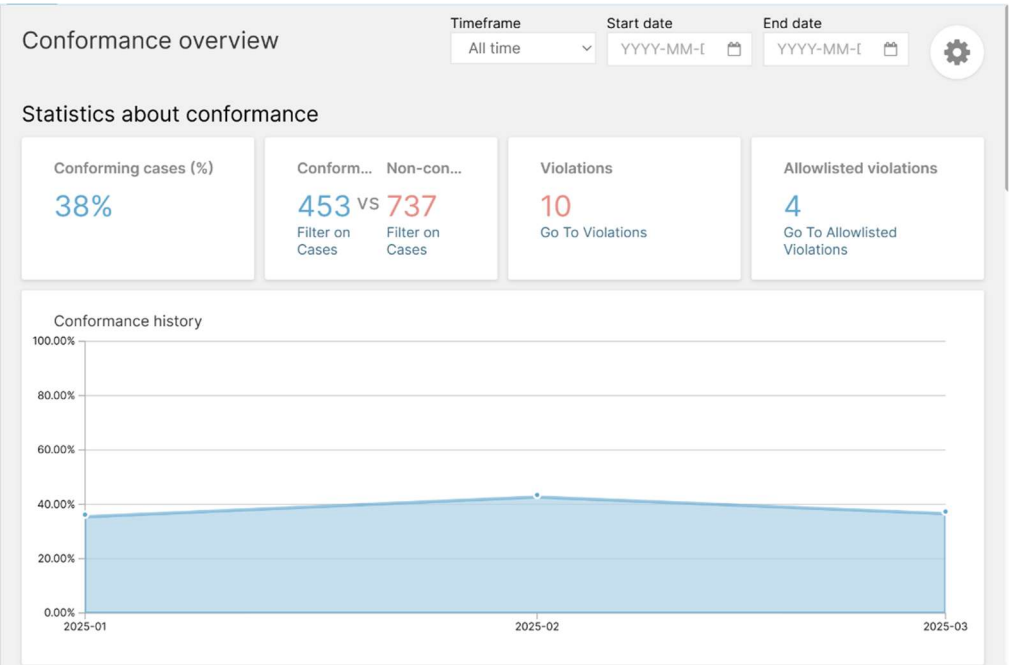
Business



Business Overview page:

- There are 1190 cases.
- The total cost is 73 000.
- Average satisfaction score is 3.59 and holds a similar level through time, with some oscillations.
- On average, patients have had 2 previous appointments in the last 3 months.
- On average, patients have had 0.5 appointments cancelled in the last 3 months, which can be roughly translated to 1 in 2 customers having a cancelled appointment.
- There are customers of all ages, but nearly 50% of customers are between the ages of 50 and 70, and only roughly 2% are under the age of 20.
- Patients between 40 and 50 years old are the least satisfied, with an average satisfaction score of 2.55. There is a clear difference in patient satisfaction for patients younger than 40 and older than 40, 4.17 average satisfaction score for patients under 40 vs 3.38 for patients over 40. Additionally, people older than 40 are having difficulties and are having to return to another day very often (it's the most common process variant for this segment), in comparison with younger people that their most common process variant is to receive a normal appointment.
- We can conclude that daily costs and number of cases have some variation but hold constant over time, with a peak on March 1st, perhaps with an extremely slight upward trend.
- 57% of cases include insured customers.
- The gender distribution is almost identical, 34,2% male, 35,3% female and 30,5% other.
- Most cases come from customers from the center of Portugal, 41%, the distribution is nearly identical across the other 3 regions at around 19/20% for the North, South and Islands.

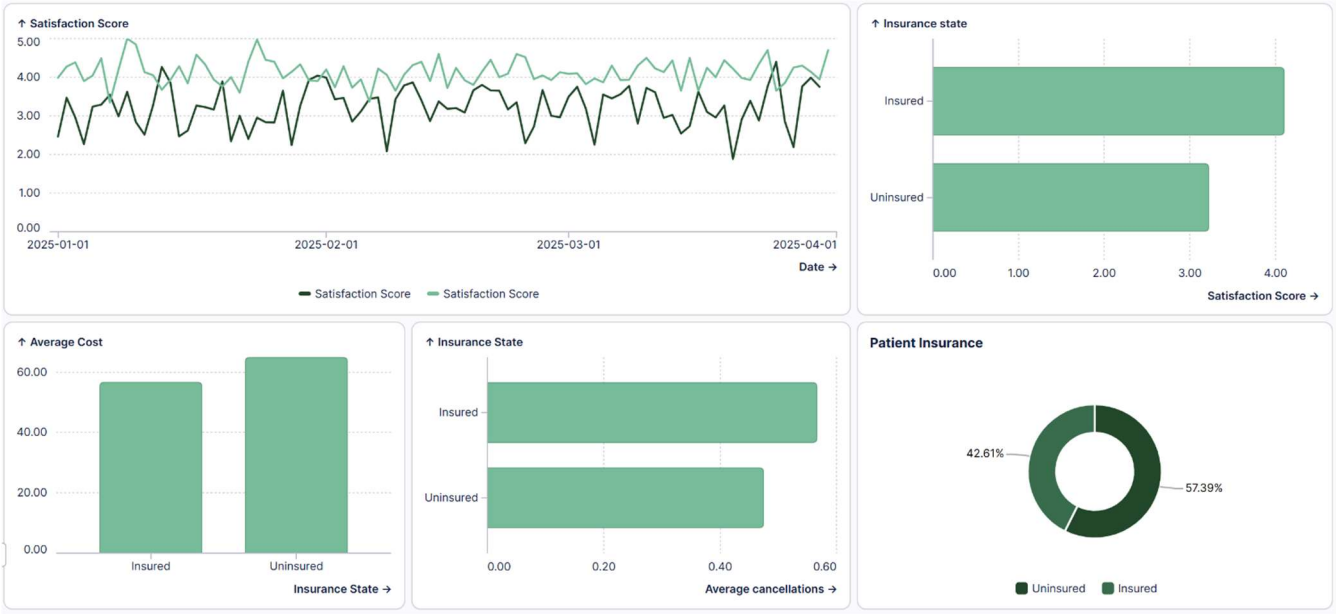
Conformance Checking



- 38 % of the cases are conforming, with the addition of 4 violations to the allowlist. The percentage of conforming cases is consistent throughout the three months of available data, with a very slight peak in February.
- We have 10 violations. This leaves us with 737 non-conforming cases with an average time of 35.8 minutes, a 13.8-minute decrease over conforming cases, and around 4.0 steps, a 0.7 step decrease over conforming cases. Thus, we can conclude that the non-conforming cases have shorter throughput time on average, yet they are not desired because they consist of undesired activities, such as the patient having to return another day, having their consultation cancelled, missing their turn, among other negative outcomes for both the patient and the clinic.
- The allowed violations are: Perform exams; Payment Via app, Receive exam results, Schedule New consultation. These are activities that while not being a part of the target process model, are expected activities of the clinic's normal and efficient operation.
- 23% of the total cases which translate to 275 cases return another day without having the consultation, of course the time and steps of these cases are smaller, but still a very undesired activity. It usually happens to older patients, 55+ years old. They are also patients, from the center region, people without insurance and from female patients.
- Most complaints are made by people without insurance, and by Males, people with hypertension also file more complaints. It translates to 191, 16% cases. And the steps are 5.6, which is more than the target process model, and takes 11 minutes longer.
- Another violation we have is check-in at reception directly followed by receiving a prescription, which does not make sense, but it's part of 8% of the cases, 98 of them specifically. This happens to people with insurance, and to male patients.

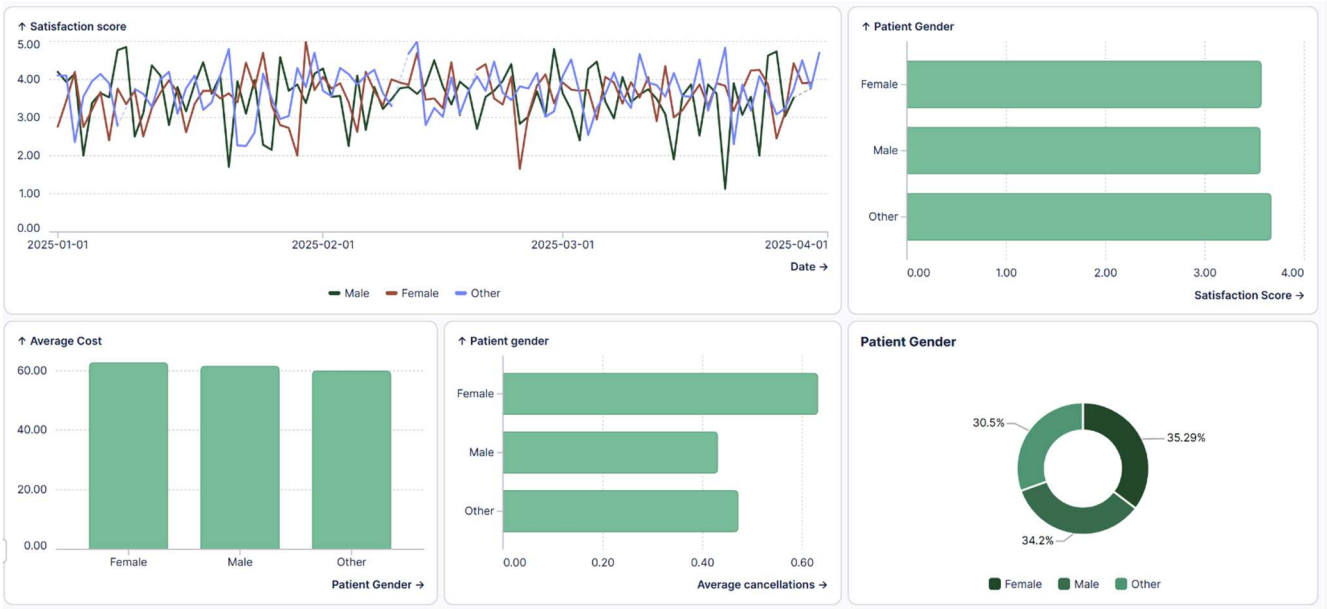
- Waiting for the Doctor also is an undesired activity, which happens in 77 cases, 6% and adds 11 minutes to the throughput time and 0.9 steps to the process. This is concerning because it happens to people with asthma and heart diseases and to people with insurance.
- People with insurance also have more consultations cancelled. With 6.1 steps and 4.0% of the cases.
- We have another strange violation, that is Payment at receptions followed by payment at reception, this happens in 31 cases, and adds 3.4 minutes to the process, and is more common for people with hypertension and asthma, and insured patients.
- People that miss their turn are all women between the ages 35 and 45 with insurance. The processes with this violation have 2.0 steps and a throughput time of 5.8 minutes and they are 2.0% of the cases.
- People with insurance and older women also have their information incomplete, 1% of the cases.
- And the system error is indeed for people from the islands with 4 and 1 from the north and they have no insurance, and this adds 13.3 minutes to the process. There are 5 cases with this in total.

Insurance Analysis:



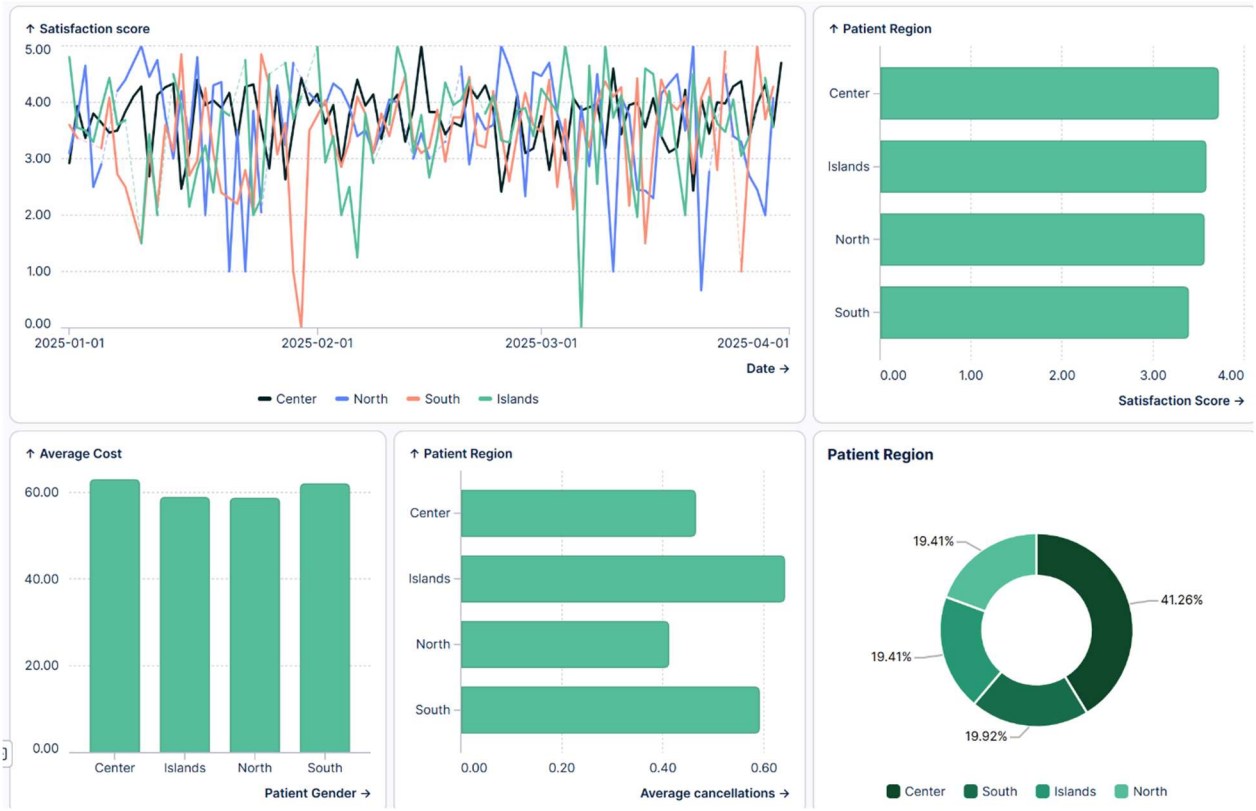
- Insured Patients have a higher average satisfaction score (4.1), when compared with uninsured patients (3.22). We can also see that this gap is consistent throughout time, obviously with some fluctuation, but the average daily satisfaction level is higher for insured patients for almost the full period of available data, uninsured patients only have (marginally) higher scores on 3 days.
- As is expected due to insurance coverage, insured patients have higher average cancellation rates and lower average costs than uninsured patients. This is likely a result of the comfort insured patients have due to the coverage.
- It's also worth mentioning that the common process variant for uninsured patients occurs 40% of the time and consists of the patient getting told to return another day, whereas for insured patients the most common variant happens 25% of the time and is an expected path.
- With all this in mind, while it's not absolutely confirmed and further analysis with more detailed data should be carried out, there appears to be an issue where uninsured patients are not being offered equal care to insured patients.

Gender Analysis:



- The average satisfaction level is similar across all gender options, with “other” being slightly higher than the “male” and “female”. All three are around 3.6.
- The average cost is also similar across all gender options with female patients having slightly higher costs followed my male and other patients. All three are around 60€.
- There is a noticeable difference in the number of cancelled appointments. In the last 3 months, female patients have cancelled, on average, 0.63 appointments, whereas other patients have cancelled 0.47 and male patients 0.43.
- The most common process variant is identical and similarly common across all gender options representing around 20-25% for all three. It describes an undesired path where the patient is told to return another day.
- Considering all this information, we don’t believe that gender is a key factor in patients being offered varying levels of care.

Region Analysis:



- The average satisfaction is highest for patients from the center (3.72), then it is very similar for the Islands and North regions (3.58 and 3.56) and it is lowest for patients from the South region (3,39).
- The average cost is similar across regions with the center and South having slightly higher values, just above 60€ and Islands and North having slightly lower values just below 60€.
- Within the last 3 months patients from the Islands and South regions have cancelled significantly more appointments, with an average around 0.6 for both, whereas patients from the Center and North regions have cancelled 0.46 and 0.41, on average respectively. This may be due to patients from the South and the Islands having to cancel their appointments because they won't be able to make it to the clinic on the consultation date.
- The center of Portugal is the region that has, as the most common case variant, a situation where the patient enters the clinic and leaves immediately to return another day, something that doesn't happen for the rest of the regions where the most common case variant describes a situation where patients enter and have a normal appointment.
- With this information, we can say that it's possible that patients from the center are being offered better care than patients from other regions, in particular South and Islands. A deeper analysis could confirm or deny this.

System Issue Analysis:

Case ID	Patient Age	Patient Region	Timestamp	Activity	Patient Gender
103	31	Islands	+057138-02-12T06:40:00.000Z	System Error	Male
1191	31	Islands	+057144-09-12T10:40:00.000Z	System Error	Other
578	40	Islands	+057121-05-19T02:00:00.000Z	System Error	Male
876	55	North	+057189-06-08T00:40:00.000Z	System Error	Other
283	60	Islands	+057091-05-06T15:20:00.000Z	System Error	Other
217	60	Center	+057193-10-02T19:20:00.000Z	Incomplete Information	Male
359	60	North	+057113-02-08T13:20:00.000Z	Incomplete Information	Male
1150	60	South	+057123-05-14T02:00:00.000Z	Incomplete Information	Female
54	65	Islands	+057162-09-29T10:00:00.000Z	Incomplete Information	Other
100	70	North	+057170-08-16T21:20:00.000Z	Incomplete Information	Female
192	70	Center	+057179-11-10T02:40:00.000Z	Incomplete Information	Female
310	70	South	+057024-12-10T04:00:00.000Z	Incomplete Information	Female
400	70	Center	+057067-02-24T20:00:00.000Z	Incomplete Information	Other
1234	70	Islands	+057108-01-14T11:20:00.000Z	Incomplete Information	Other
1316	70	North	+056974-05-24T08:40:00.000Z	Incomplete Information	Male

- To analyze system issues (error or lack of information) present in the event log, we created an action flow within Celonis which, when run, will email a csv file containing case ID, patient age, patient age, patient region, the activity (system error or incomplete information) and the timestamp, the most recent version of this file can be seen above. With this information, we can then see if any of these factors have a direct impact on system issues.
- We can see that ‘Incomplete Information’ only occurs for patients aged 60 or more, this could be due to these older patients struggling with the clinic’s digital systems leading to a lack of information required. For ‘System Error’ there is a more balanced age distribution.
- We can conclude that so far gender has no impact on these issues.
- Region has no clear impact on issues of ‘Incomplete Information’, but from the 5 cases of ‘System Error’ that were present 4 were from patients from the Islands, this could indicate that patients from this region are experiencing more and more frequent issues with the clinic’s system than patients from other regions. Another error occurred to a patient from the North, but as this is a single case it’s hard to know if there is an actual relation or if it was just a coincidence.

Final recommendations

Saude+ management should be worried about process deviations, there are more undesired paths than desired ones for the clients, that is a clear indicator that the clinic is not operating efficiently and immediate action is needed to maintain their clients in the long run, more specific concerns about: patients having to return on another day, waiting for doctor, complaining, facing system errors, having their consultation cancelled, having incomplete information. It's crucial to understand why this is happening, what the customers are feeling and their needs.

The manager's concerns about unequal service were correct in some cases:

Customer Segments:

- The insurance group is receiving better treatment, with most customers attaining a good service while the ones without insurance must return another day, 40% of the time. The gender really doesn't contribute to different treatment. It is worth mentioning that the "others" group have slightly higher satisfaction scores. Finally, as mentioned above, it's possible that patients from the center are getting better service than patients from other regions, in particular the South and the Islands.

Patient Profile:

- Patients younger than 40 years old, roughly 27% of all cases, show higher satisfaction (4.17 average satisfaction score for patients under 40 vs 3.38 for patients over 40), the system interface should be checked to see if it is too complex for older people since they are coming to the clinic and having to return another day.

Conclusion

Saude+ management expressed concerns over inconsistent process deviations within the patient's process of attending an appointment, declining patient satisfaction levels, as well as potentially providing unequal care to patients based on their gender, region, age or whether they are insured or not.

Using Celonis, we performed a full Process Intelligence pipeline to attend to these concerns. We were able to confirm that there are various process deviations which can lead to worse service and more dissatisfied customers. We also verified that insured patients are being offered better care than insured patients, older patients are less satisfied and may struggle more with the clinic's systems and that patients from regions outside the center, in particular the South and Islands may also be provided with worse care. However, we concluded that patient satisfaction, daily costs and number of cases have held relatively constant throughout the three-month window of our analysis, and finally we concluded that patients were not being offered different levels of care based on their gender.

To conclude, with the insights from this report we believe that Saude+ management can improve their operations with a key focus on providing equal care to all patients. This should, in turn, lead to better service and more satisfied patients.