# Data-Driven Process Improvement of a LOCA Dispensing Station in an Optical Bonding Process

**Rui Oliveira**
ALGORITMI Research Centre,
LASI (Intelligent Systems Associated Laboratory),
School of Engineering, University of Minho, 4710-057
Guimarães, Portugal
Id6911@alunos.uminho.pt

**Paulo Sampaio**
Industrial Engineering and Management (IEM) R&D Group,
ALGORITMI Research Centre,
School of Engineering, University of Minho, 4710-057,
Braga, Portugal

**Júlio Viana**
IPC – Institute for Polymers and Composites, Department of Polymer Engineering,
LASI (Intelligent Systems Associated Laboratory),
University of Minho, 4804-533,
Guimarães, Portugal

## Abstract

With the fourth industrial revolution, the increased trend of digitalization of production processes and products resulted in more data regarding processes and products becoming available. For data-driven strategies like Lean Six Sigma to deal with the amount of data available today, it requires integrating a new set of tools and techniques from the Data Science research field. However, the fact that more data is available does not necessarily translate into more efficient and effective processes. For that, it is necessary to identify what is useful and can be transformed into value and what is not. Following the DMAIC cycle, the project presented in this article focuses on improving a production station of an Optical Bonding production line that produces flat panel displays for the automotive industry. The inclusion of Data Science tools and techniques tools in this improvement project helped to extract and analyze information from unstructured log files. The achieved results of this project show that the combination of data science tools with traditional process improvement tools and methodologies is very productive.

## Keywords
Optical Bonding process, Process Improvement, DMAIC, Data-driven analysis

## Introduction
As a consequence of the fourth industrial revolution, the increased trend of digitalization production processes and products resulted in the availability of more data regarding processes and products became available, be it for design, management, control or monitoring purposes (Cattaneo et al. 2018; Titmarsh et al. 2020). Business industries such as aerospace, semiconductor, or display devices are good examples of manufacturing industries that often involve complex production processes that create a larger variety and volumes of production data (Zheng et al. 2014).
However, the naive fact that more data is available does not mean that processes are becoming more efficient and effective. For that, it is necessary to take advantage of such availability and pinpoint what is useful and can be transformed into value and what is not. Furthermore, the unending development and usage of sensors and measurement techniques will continue to allow the collection and storage of large volumes of data of each individual equipment in any complex engineering systems, to discover new correlations of data and operating status(Coccia et al. 2021). As a

result, companies need to reinvent how processes are managed and how they can adjust to integrate information and physical data into intelligent workflows (Skalli et al. 2023).

An example of such is the Lean Six Sigma strategy, which, by following the DMAIC cycle and data-driven analysis, seeks to improve organizations operational efficiency and effectiveness (George 2003; Zwetsloot et al. 2018). However, due to the larger and wide variety of data volumes currently available, Lean Six Sigma has seen some of the traditional methods and tools it uses becoming less effective and in need to be complemented with a different set of tools that have become known as “data science” tools (Zwetsloot et al. 2018).

The work presented here is part of a larger improvement initiative of a multinational company located in Portugal that is a first-tier supplier of the automotive industry. Its main business area focuses on the development and production of car multimedia. The company product portfolio goes from navigation systems solutions, infotainment systems solutions, instruments systems and clusters, to steering sensors systems and control unit systems, among other products and solutions.

### 1.1 Objectives

The improvement project presented here follows the DMAIC methodology. It focuses on improving a production station of an Optical Bonding production line that produces flat panel displays for the main instrument cluster systems. Due to the nature of the data available, a data science platform was used to extract, transform and load the data. Using the platform’s graphic user interface (GUI), it was possible to perform rather simple analyses over the available data and discover improvement opportunities.
Some of the data presented here and related to the production process may be masked due to confidentiality issues.

## 2. Literature Review

Along with the fourth industrial revolution came great advantages and an equal share of new challenges. These challenges are not only limited to physical or digital settings and environments, as management philosophies, strategies and tactics must be readapted to new developments and ways of thinking and acting. Lean Six Sigma (LSS), a business process improvement strategy that resulted from the combination of Lean Manufacturing and Six Sigma and has been around for the past 20 years, is one of those management strategies facing new challenges (Pongboonchai-Empl et al. 2023).

Regardless of the industry or business, LSS focuses on achieving better process performance by exploring the process data's underlying relations and using that same data and its hidden insights to improve processes (Zheng et al. 2014). One of the challenges that the fourth industrial revolution (or industry 4.0, or simply I4.0) puts to LSS is related to the digitalization of business processes (Kregel et al. 2021). As traditional LSS projects usually rely on datasets with 30–1000 observations (Kregel et al. 2021; Zwetsloot et al. 2018) and a limited number of process variables/ parameters, current datasets surpass all of those numbers either in observation records or number of process variables/parameters. Furthermore, LSS relies on structured datasets where information is properly arranged in a tabular manner, and one can identify each process variable and each process record. The issue is that many datasets are built and stored in unstructured formats, holding various information like time, date, user, operation number, process parameter value, operation steps, program command lines, sensors readings, information from others preceding and proceeding processes, among other possibilities, all written sequentially.

Many companies open their way into the Industry 4.0 concept by starting with the digital transformation of their process. For example, they start to implement sensors and other smart technologies in search of faster processes and more significant cost reduction, along with the ability to optimize quality with real-time data analysis, which therefore leads to the creation of large amounts of data (Butt 2020; Pongboonchai-Empl et al. 2023; Veile et al. 2019), and in many cases of unstructured data from various sources (Clancy et al. 2023). Nevertheless, the digitalization of processes data through the implementation of efficient and functional information technologies allowed companies to extract key valuable insights from their operations, aiding them to start making even more data-driven decisions (Bag et al. 2020; Clancy et al. 2023; Dhamija & Bag 2020).

Besides the mentioned physical information technologies, other technologies such as Machine Learning (ML), Artificial Intelligence, Data Mining, Predictive Analytics, Pattern Matching, Cluster Analysis, Data Visualization, Graph Analysis, Simulation, Neural Network and Multivariate Statistics (Nader 2022) are examples of technologies and tools that contributed for that extraction of key valuable insights, which consequently benefits LSS strategies. For example, Big data analytics (BDA) and Data Mining (DM) techniques can be used to process large amounts of unstructured data and help to answer much more complex questions and problems (Hoerl et al. 2014; Kregel et al. 2021). How Big Data and Data Mining can be employed in LSS is one of the current trends in the field of LSS/6-Sigma research (Antony et al. 2019; Kregel et al. 2021; Sony et al. 2020).

Nonetheless, all this buzz around these potentialities and benefits that such technologies like DM and BDA can bring to LSS, conceal the fact that all this commotion of processes being monitored and controlled by sensors and other measures, leads to the fact that it is gathered more data and it is actually used, which may result in lost opportunities for process discovery and improvement (Koppel & Chang 2021), if not worst, adding noise to the improvement initiates that could in fact deliver benefits. As (Gandomi & Haider 2015) stated that the big discussion and promotion initiatives around Big Data, Predictive Analytics and others, disguises the fact that the analytics-ready structured data forms only a small part of big data. This means that although it is seen as inefficient work, the pre-processing data to enable data-driven decision-making is fundamental if one wishes to extract useful knowledge from the collected data (Dahlström 2022).

The underlying fact is that every data-driven strategy requires valuable data to achieve its purposes, regardless of the data size, structure and format. Citing Titmarsh et al. (2020), "*From the quality engineering perspective, problems can never be solved unless discovered. Therefore, providing the historical data that cover components and processes' behaviour constitutes the basis of identifying problems and beginning the quality improvement procedure.*". Since LSS is a data-driven strategy to improve processes and solve problems, it is of the utmost importance that data logs uphold relevant information regarding processes, products and services, otherwise having sensors and other information technologies extracting and collecting data is just another type of waste.

## Improvement project

An Optical Bonding production process involves bonding two or more optical components, namely a flat panel display and cover glass, using an Optical Clear Adhesive (OCA) with a refractive index that matches both components optical index, filling the space between both components. Such application of this type of bonding process not only improves optical characteristics of the full assembly stack (flat panel display, adhesive and cover glass), but as mechanical ones by increasing the whole assembly's resistance to shocks, vibrations, moisture, humidity and thermal variation, as well as other electrical benefits ((Abrahamson et al. 2019; Bahadur et al. 2011; Oliveira et al. 2023).
In the specific case of this project, the production process uses a low viscosity Liquid Optical Clear Adhesive (LOCA) for bonding the two components. This type of process, also known as Wet Optical Bonding process, involves six main steps:

- Materials Preparations - Which includes treatment and inspection of the display and cover glass surfaces;
- DAM Dispensing - An OCA portion is dispensed on the display surface, according to a specific pattern and over the edge of the bonding area, which will act as a containment barrier (also known as DAM);
- Filling Dispensing - Also according to a specific pattern, the LOCA is dispensed on top of the display surface and within the bonding area;
- Assembly – Once the display unit is in the assembly position, a cover glass is put on top of the display panel with the LOCA. This task is performed by a robotic arm with high precision that thoroughly places the cover glass on top of the LOCA so that no air gets trapped between the LOCA and the cover glass (Bahadur et al. 2011; Campbell, 2016), applying some pressure on the cover glass in order to help spread the LOCA around the bonding area and also forcing out any air inside the assembly stack and contributing for a complete bonding contact between all components (Cruz et al. 2017; Yeh et al. 2013);
- Curing - Finally, the full assembled product goes through a curing process that finishes the bonding of all components and seals the LOCA (Campbell 2016; Mozdzyn & Rudolph 2011).

Figure 1 provides an example model of the production line under study during this project. Following that, full assemble products are submitted to a full inspection process, where an operator looks for several possible defects, some related to both the bonding process like air bubbles trapped inside the assembly, insufficient bonding process (LOCA not covering all bonding area) or possible foreign particles inside the assembly, and other defects related to product like scratches, dents, blurs, among other possible defects.

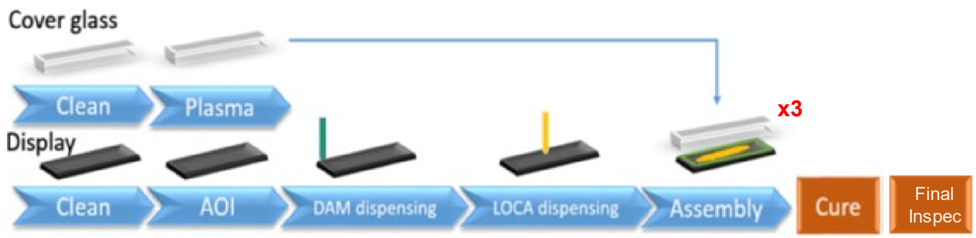


Figure 1. Display optical bonding production process (Adapted from: (Cruz et al., 2017))

The production line under evaluation considered in this project is very similar to the one mentioned in Figure 1, where the major difference is regarding the Assembly step. In the production line considered in this project, the Assembly step is executed in three assembly stations, each one with a robotic arm with high precision.
Since the curing station is the last production step before final inspection and it is performed by a single oven that has not presented any issues on the production line under study or any of the other productions lines with equal equipment, the process owner suggested that the first improvement project would focus on the assembly station. From that first improvement project, and among other results, a decrease of almost 70% in the internal rejection rate of defective assemblies associated with that production line was achieved. Therefore, it was requested by the company board that the next improvement project focus on the preceding station of the production line, the filling dispensing station, in order to improve the production line further.

### 3.1 Define phase - Project Charter and objectives

Following the company board's request, the proposed project concerns the filling dispensing station. The main purpose of this station is to dispense a specific amount of LOCA with a specific pattern and in a predetermined location/area of the display surface, within the designated bonding area. This dispensing process is performed using an electromechanical system with high precision with a single nozzle, discharging the LOCA over the designated location at a constant rate and speed. Figure 2 is a graphical representation of the described process.


Figure 2. An example of a filling dispensing/coating method (Adapt from: https://www.dexerials.jp/en/news/2020/news20019.html)

Along with the Process Owner and the Process Engineering team, several Gemba walks were performed to the production line to evaluate the filling station. One issue identified right at the beginning was that the station software was recording production information in daily unstructured log files and sometimes the information regarding a single dispensing operation could be divided between 2 logs. Another concern that the Process Engineering team reported was the possibility for the logs not containing information regarding all relevant variables, such as the record of the deposition pressure at each dispensation. A change in the dispensing pressure could cause variations in the dispensing quantity of the LOCA and the dispensing pattern.

Therefore, and having into consideration the current status of the production station, it was established by the project team with the production process owner that the project would have e maximum duration of 6 months and that the following objectives would have to be pursued:

1. Confirm that all relevant parameters and corresponding data were being recorded on the station production logs;
2. Improve station process output:

These objectives were established considering that most of the time spent on this project would be on studying the station software and logs to confirm that the desirable information was being recorded, where, in what conditions and whether it was possible to extract it. Also, by the time this project started in April 2021, the production line rejection rate was around 1.15% and stable. According to the Process Engineering team's experience, this station was usually stable, even when the production line was not working in a stable manner.

### 3.2 Measure phase - Data gathering and structure

The first months of this improvement project were spent analyzing the station log records and the information available within them. Soon some expectancies were confirmed, mainly that there was only one set of files where production information was being recorded. However, not all desirable information was being recorded.

Moreover, as previously explained, the production station software recorded the information in unstructured log files and in one file per calendar day, which means that a new log was being created every day, leading to a situation where information regarding the filling operation for a specific display unit could be divided between 2 logs. Figure 3 and Figure 4 present brief examples of how the logs were being created and the log information structure (both images have been softened due to confidentiality).

Although the station logs did not possess all relevant information, it was clear that these logs could be transformed in order to retrieve some information that would allow to perform some analysis.

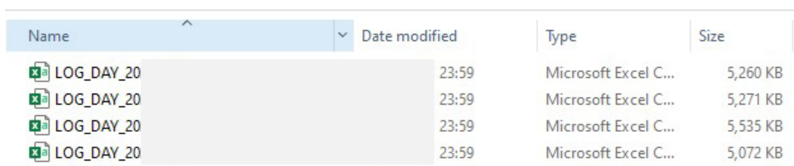


Figure 3. Examples of daily production log files

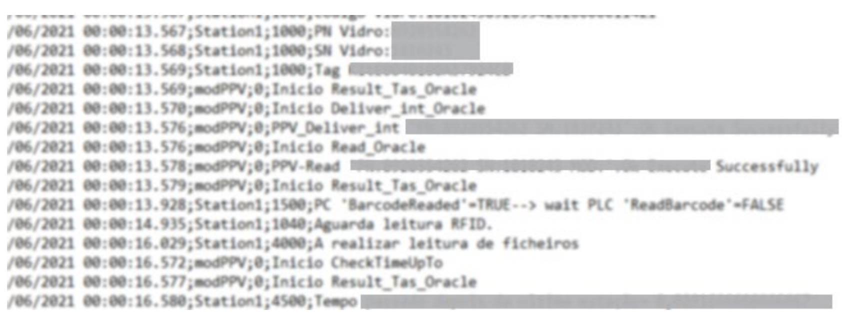


Figure 4. Example of a Log structure

Using the RapidMiner Studio®, a data science software that provides data mining and machine learning procedures, an Extract, Transform and Loading (ETL) process was built in order to automatically gather all log files and extract information such as timestamp, display unit ID and serial number, dispensing time per display unit, total process time (since display unit enters the dispensing station until it leaves station), to which assembly station was sent afterward and other information. Figure 5 presents a summary of the ETL built.

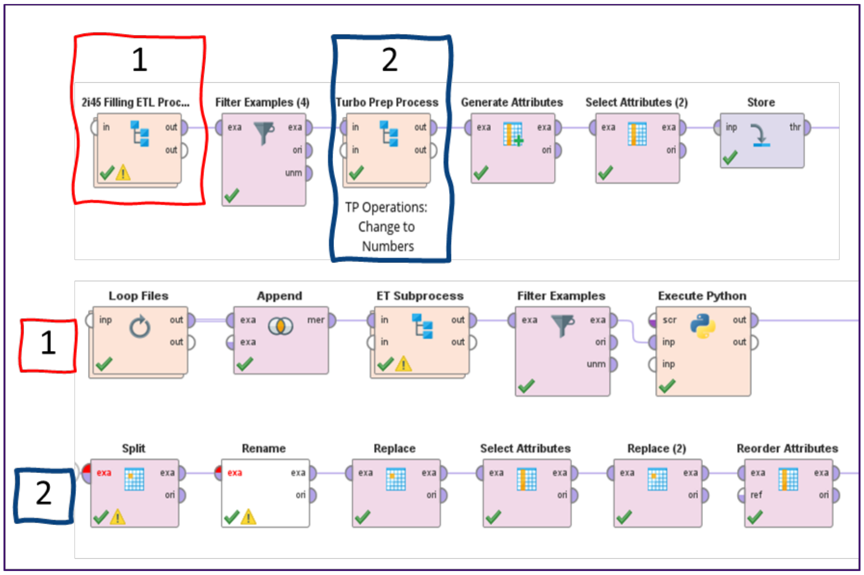


Figure 5. ETL (Extract, Transform and Load) process using RapidMiner Studio®

The ETL process presented in Figure 5 has two sub-processes. The first subprocess, marked as 1 in the red box, is a subprocess to append all logs files, perform some text transformations and, with the help of some Python® coding, finishes the extraction of data from the set of logs available. The second subprocess, marked as 2 in the blue box, performs the remaining necessary transformations of data in order to present all information in a structured tabular format. The final result from this ETL process, as it can be seen in Figure 6, is a structured log in a tabular format with 8561 records and 11 attributes/variables, from which 6 are considered descriptive variables(date tag, time tag, transportation kit tag, display unit model tag and display unit serial number), meaning that the other 6 variables were considered process variables.

| Row No. | Date | Time | KitTag | Display | SerialNumber | DispTimePe... | ProcessTime | TotalDispTi... | MaxTimeall... | RobotToSend | TimeToLast... |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 17/06/2021 | 00:01:18.983 | E0040150C5... | 8928554262 | 1699557 | 2 | 30 | 1167192 | 99999999 | 2 | 101.00000 |
| 2 | 17/06/2021 | 00:02:50.489 | E0040100A3... | 8928554262 | 1699276 | 2 | 30 | 1167194 | 99999999 | 3 | 45 |
| 3 | 17/06/2021 | 00:03:30.338 | E0040100A3... | 8928554262 | 1692706 | 3 | 30 | 1167197 | 99999999 | 1 | 55.00000 |
| 4 | 17/06/2021 | 00:07:35.182 | E0040100A3... | 8928554262 | 1699142 | 2 | 30 | 1167199 | 99999999 | 2 | 45 |
| 5 | 17/06/2021 | 00:08:15.027 | E0040100A3... | 8928554262 | 1683928 | 2 | 30 | 1167201 | 99999999 | 3 | 56.00000 |
| 6 | 17/06/2021 | 00:08:54.199 | E0040150C5... | 8928554262 | 1743328 | 3 | 30 | 1167204 | 99999999 | 1 | 48.00000 |
| 7 | 17/06/2021 | 00:09:42.625 | E0040100A3... | 8928554262 | 1743317 | 3 | 29 | 1167207 | 99999999 | 2 | 46.00000 |
| 8 | 17/06/2021 | 00:10:22.003 | E0040150BE... | 8928554262 | 1743212 | 2 | 30 | 1167209 | 99999999 | 3 | 56.00000 |
| 9 | 17/06/2021 | 00:11:00.622 | E0040100A3... | 8928554262 | 1743148 | 1 | 30 | 1167210 | 99999999 | 1 | 65.00000 |
| 10 | 17/06/2021 | 00:11:45.734 | E0040100A3... | 8928554262 | 1743144 | 3 | 36 | 1167213 | 99999999 | 2 | 80.00000 |
| 11 | 17/06/2021 | 00:12:24.653 | E0040150C5... | 8928554262 | 1743399 | 3 | 30 | 1167216 | 99999999 | 3 | 83.00000 |
| 12 | 17/06/2021 | 00:13:04.386 | E004010081... | 8928554262 | 1745269 | 1 | 30 | 1167217 | 99999999 | 1 | 84.00000 |
| 13 | 17/06/2021 | 00:13:49.308 | E0040100A3... | 8928554262 | 1703725 | 3 | 34 | 1167220 | 99999999 | 2 | 89.00000 |
| 14 | 17/06/2021 | 00:14:28.577 | E0040100A3... | 8928554262 | 1699161 | 3 | 30 | 1167223 | 99999999 | 3 | 89.00000 |
| 15 | 17/06/2021 | 00:15:08.192 | E00401002B... | 8928554262 | 1683442 | 2 | 30 | 1167225 | 99999999 | 1 | 90 |
| 16 | 17/06/2021 | 00:15:51.330 | E0040100A3... | 8928554262 | 1703115 | 2 | 34 | 1167227 | 99999999 | 2 | 94.00000 |
| 17 | 17/06/2021 | 00:17:37.043 | E0040100A6... | 8928554262 | 1703713 | 2 | 96 | 1167229 | 99999999 | 3 | 160.00000 |
| 18 | 17/06/2021 | 00:19:22.820 | E0040100A3... | 8928554262 | 1703503 | 3 | 95 | 1167232 | 99999999 | 1 | 227.00000 |
| 19 | 17/06/2021 | 00:20:07.220 | E0040100A3... | 8928554262 | 1701643 | 2 | 34 | 1167234 | 99999999 | 2 | 202.00000 |
| 20 | 17/06/2021 | 00:20:53.740 | E0040100A3... | 8928554262 | 1701661 | 2 | 37 | 1167236 | 99999999 | 3 | 221.00000 |
| 21 | 17/06/2021 | 00:21:33.499 | E0040100A3... | 8928554262 | 1701839 | 2 | 30 | 1167238 | 99999999 | 1 | 186.00000 |
| 22 | 17/06/2021 | 00:22:12.892 | E0040100A3... | 8928554262 | 1701742 | 1 | 30 | 1167239 | 99999999 | 2 | 120.00000 |
| 23 | 17/06/2021 | 00:22:52.817 | E0040150C5... | 8928554262 | 1701975 | 1 | 30 | 1167240 | 99999999 | 3 | 115.00000 |
| 24 | 17/06/2021 | 00:23:31.813 | E0040100A3... | 8928554262 | 1701822 | 2 | 30 | 1167242 | 99999999 | 1 | 108 |
| 25 | 17/06/2021 | 00:24:11.409 | E0040100A3... | 8928554262 | 1701758 | 2 | 30 | 1167244 | 99999999 | 2 | 107.00000 |

ExampleSet (8,561 examples, 0 special attributes, 11 regular attributes)

Figure 6. Example of a structured log after the ETL process containing 8561 records and 11 variables

From this moment onwards, some initial measures and analysis were performed using the "Results Tab" available on the RapidMiner® graphical user interface (GUI). From the standard statistics that the GUI has available, it was possible to understand how the 8561 produced units divided themselves over 7 consecutive production days, how these were distributed over the following assembly stations, among other statistics. Also, it was possible to identify 40 display units that during these 7 days were processed more than once in this production line, meaning that for some reason, these display units were involved in some rework processes, recovered and used again for production.

Figure 7. Example of the RapidMiner® graphical user interface (GUI)

### 3.3 Analyze phase

As usually happens in the improvement projects developed under the DMAIC cycle, the measure and the analyze phases overlap, as both phases analyze the available data. As previously mentioned, of the 11 variables that the project team could extract from the production logs, 6 were considered process variables. From the analysis and measurement made to the other 6 variables, three were considered by the project team as irrelevant as they did not present any relevant information or any correlation with the remaining variables. The process experts engineers assigned one of those variables to a function that the station could perform but was not being used. From the remaining variables, 3 of them were further analyzed: "Process time" – The time between a display unit entering the dispensing station until it leaves the station; "Time to last Station" – The total time passed between the moment that the display unit has been released from the inspection station preceding the DAM dispensing station, which precedes filling dispensing station, until the moment that the same display unit is released to one of the assembly stations; "RobotToSend" – To which of the following three assembly stations was sent each display unit, identifying each assembly station by the robot number, this is, assembly station number 1 has the Robot1, assembly station number 2 has the Robot2 and assembly station number 3 has the Robot3.

A Pareto analysis performed on the "Process Time" variable revealed that 55,19% of the analyzed production either had a processing time of 30 seconds, 18,6% had a processing time of 29 seconds, 4,8% had a processing time of 36 seconds and 3,41% had a processing time of 37 seconds. As shown in Figure 8, these 4 process timings represent 82% of the entire sample.

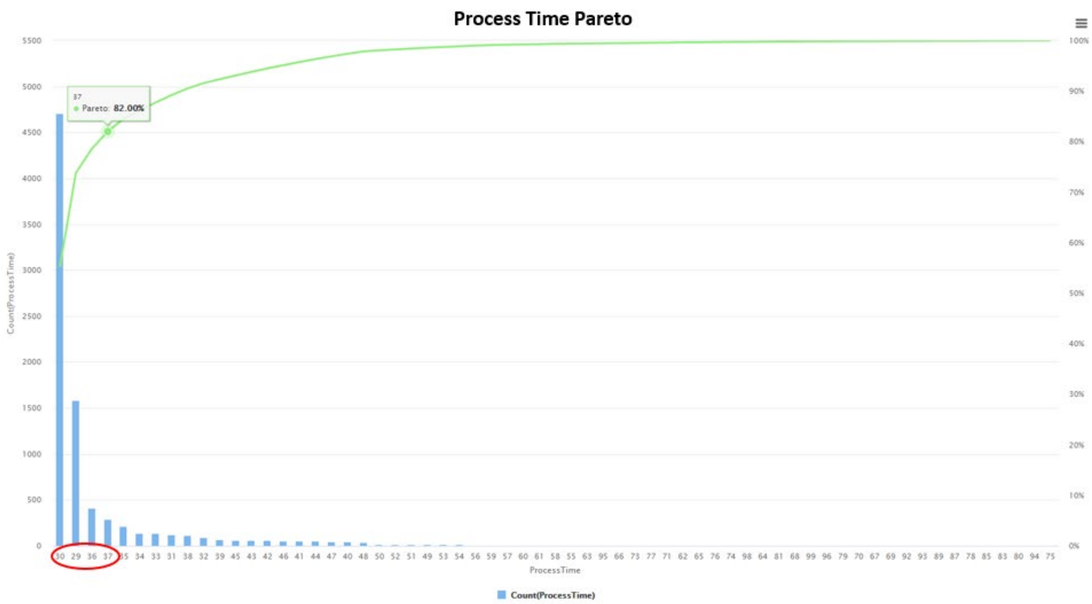


**Figure 8. Pareto Chart of the Process Time variable of the filling station**

This information raised some questions because although the foreseen processing time for that would be around 30 seconds, the fact that the number of display units with processing times of 36 and 37 seconds was 5 times higher than display units with 31 seconds of process times, showed that for some reason some parts were being processed and then stopped inside the dispensing station until they were able to move to next production subprocess, one of the 3 assembly stations. This "Process time" variable is crossed with the "RobotToSend" variable to understand if the assembly station had any influence over this variable. Figure 9, Figure 10 and Figure 11 show 3 bar charts of the "Process Time" variable, one for each assembly station destination, this is the assembly station with Robot 1, the assembly station with Robot 2 and the assembly station with Robot 3.

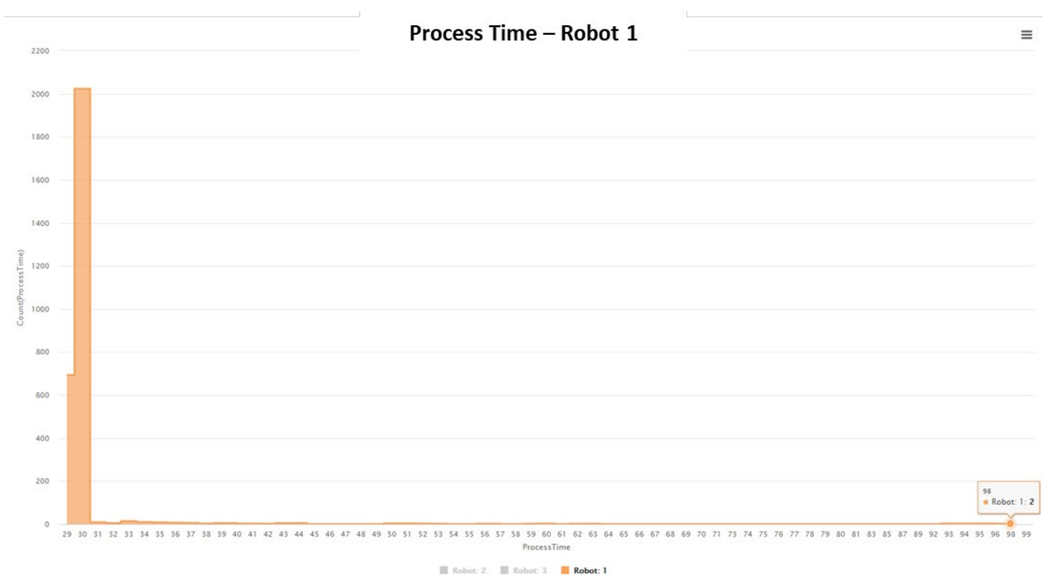


Figure 9. Process Time bar chart of the units processed in the filling station and sent to assembly number 1

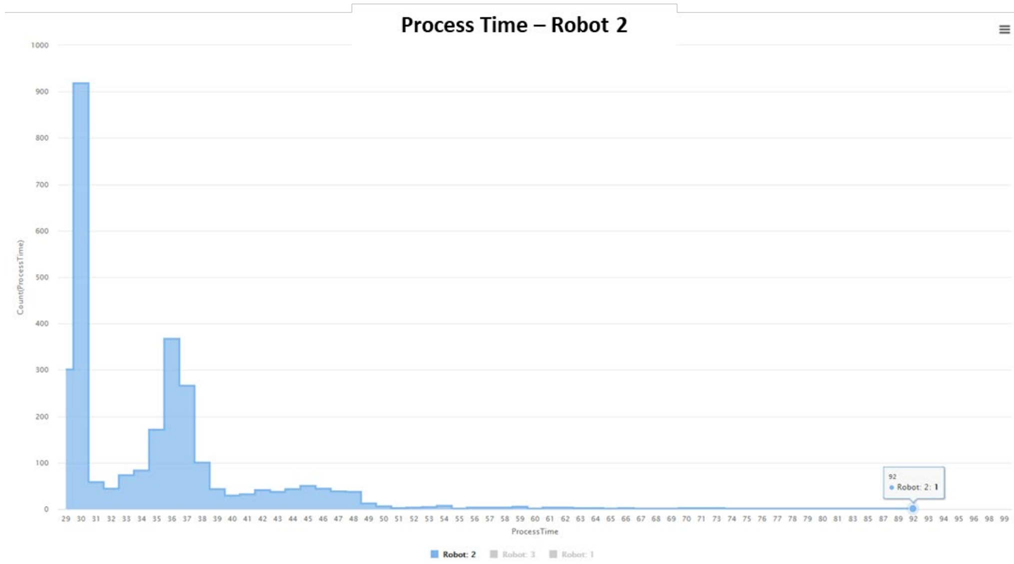


Figure 10. Process Time bar chart of the units processed in the filling station and sent to assembly number 2


Figure 11. Process Time bar chart of the units processed in the filling station and sent to assembly number 3

As can be seen from the previous three figures, the large majority of the display units with processing time of 36 and 37 seconds were being sent to the assembly station with Robot 2. Process experts investigated this situation further and found that it was related to a set of rules introduced on the software managing the entire production line.

Another analysis was performed on the Time to Last Station variable, which represents the total time passed between the moment that the display unit has been released from the inspection station preceding the DAM dispensing station and filling dispensing station until the moment that the same display unit was sent to one of the assembly stations.

The initial statistics regarding this variable showed that the lowest time registered was 45 seconds, the average was around 117 seconds, the standard deviation was around 92.8 seconds, and the maximum time registered was 3600 seconds. A bar chart regarding this variable can be seen in Figure 12. However, the facts that the standard deviation was extremely high and there were values higher than 900 seconds (the equivalent of a shift break) led the project team to investigate this situation further.

As a result of new investigations, it was found that the root cause for such situation was related to the matter that software managing the entire production line having a rule that for every display unit with a “Time to Last Station” value equal or higher than 15 minutes (900 seconds), the software would automatically reject those units. In order words, display units that take more than 15 minutes to reach from the display inspection station until one of the assembly stations are considered of no good use for assembly.

It was also found that the staff operating the production line knew about this issue and, whenever possible, would prefer to stop the production line entirely and not to load the production line with new display units, avoiding a higher number of automatically rejected parts. The downside of this decision is that every time they wanted to restart production, it would require them to perform a new production line setup.

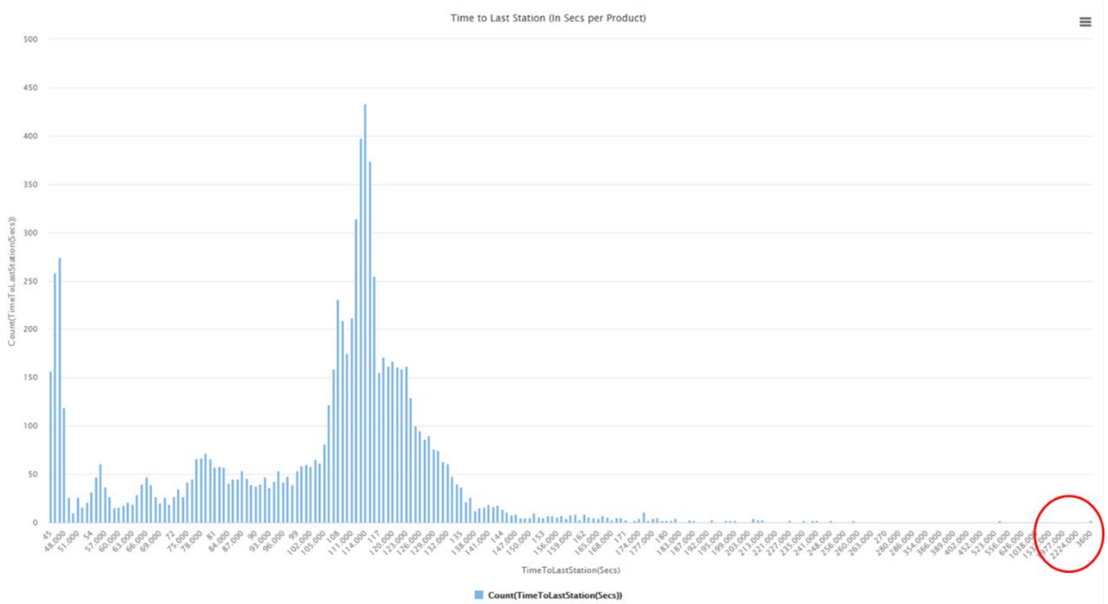


Figure 12. Time to Last Station (In secs per product) variable bar chart

## 3.4 Improve phase

From the performed analyses, three improvement actions were defined. The first improvement action was related to identifying new process parameters that the process engineering team considered relevant for a more effective control of the filling dispensing station. More specifically, they consider it was critical to start controlling parameters such as: Monitoring the dispensing quantity of the LOCA; Measurement and monitoring of the LOCA thickness dispensed over each display unit; Measuring and monitoring the LOCA viscosity; and monitoring of the LOCA dispensing pressure. To start controlling and monitoring these parameters, the Process Engineering team had to request the intervention of the station manufacturer in order to increase the number of sensors installed in the dispensing station and to perform the necessary changes on the station software to start registering these parameters.

The second improvement action was related to the production line managing software in order to improve and balance the distribution of the displays unit over the 3 assembly stations (Robots), avoiding that the majority of displays units with higher processing time getting concentrated in the assembly station 2.

The third and final improvement action was related to the “Time to Last Station” variable and a rule in the production line managing software. For every display unit that took 15 minutes (900 seconds) or more from the display inspection unit until an assembly station, it would be automatically rejected by the software. The Process Engineering team, along with the Product Industrialization team, performed some studies that found out that after the filling LOCA was dispensed over the display unit, these could be put on hold inside the production for a period of time as nearly as 45 minutes without affecting assembly quality. By changing that rule from 15 to 45 minutes, this means that for example, on shift breaks, the production staff would no longer need to stop the production line to avoid automatic rejection completely.

### 3.5 Control phase - Results Summary and Discussion

When the improvement project team decided to wrap up the project in August 2021, the dispensing station manufacturer did not have the chance to perform the required changes on the station. Therefore the new parameters were not being monitored yet. However, by that time it was possible to confirm that after changes made to the software's rules that manage the entire production line, The “Time to Last Station” variable average had dropped nearly 10 seconds per unit, increasing the station output. Also, production was more evenly distributed between the three assembly stations. These improvements impacted the entire production line output. An increase in output of about 32 pcs/shift was recorded, each meant an increase of 7.3% in shift production capacity.

In addition, and although this was not defined as one of the project objectives, the development and implementation of an ETL process using a data science tool to transform an unstructured machine log into a structured, organized and tabular log allowed the company to identify new parameters for improving control, bugs and outdated data that were being collected and store unnecessarily. Moreover, it is the improvement project team belief that this is not the only station on the production line that collects, records and stores information in unstructured logs, suggesting that the developed ETL process for the improvement project can be used as a baseline for projects that involve handling similar unstructured machine logs.

## Final Conclusions

Undoubtedly, the combination of data science tools with traditional process improvement tools and methodologies is very prolific. Although it would be possible to develop such an improvement project with other softwares, it would not be easy to develop this project within the desirable timeframe without the help and inclusion of some data science and data mining tools. The inclusion of such helped not only to extract several information from unstructured source files, but also to transform and organize it in a structured manner.
Furthermore, this was done considering large amounts of data to process (or at least large enough for traditional improvement projects). However, it was also shown that a considerable amount of the information extracted from the logs was of no-good use. It was due to the experience and knowledge of the process experts that “the dots were connected” and value insights were taken from the available data, converting these insights into improvement actions that contributed to increasing the production output.

## Acknowledgments

This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the R&D Units Project Scope: UIDB/00319/2020 and the doctoral scholarship grant: PD/BDE/142911/2018.

## References

Abrahamson, J. T., Beagi, H. Z., Salmon, F., & Campbell, C. J., Optically Clear Adhesives for OLED. In *Luminescence - OLED Technology and Applications. IntechOpen,2019*. https://doi.org/https://doi.org/10.5772/intechopen.88659

Antony, J., Sony, M., Dempsey, M., Brennan, A., Farrington, T., & Cudney, E. A., An evaluation into the limitations and emerging trends of Six Sigma: an empirical study. *The TQM Journal*, *31*(2), 205–221, 2019. https://doi.org/10.1108/TQM-12-2018-0191

Bag, S., Yadav, G., Wood, L. C., Dhamija, P., & Joshi, S. , Industry 4.0 and the circular economy: Resource melioration in logistics. *Resources Policy*, *68*, 101776,2020. https://doi.org/10.1016/j.resourpol.2020.101776

Bahadur, B., Sampica, J. D., Tchon, J. L., & Butterfield, A., Direct dry film optical bonding - A low-cost, robust, and scalable display lamination technology. *Journal of the Society for Information Display*, *19*(11), 732,2011. https://doi.org/https://doi.org/10.1889/JSID19.11.732

Butt, J. (2020). A Strategic Roadmap for the Manufacturing Industry to Implement Industry 4.0. *Designs*, *4*(2), 11. https://doi.org/10.3390/designs4020011

Campbell, C. J. (2016). Optically Clear Adhesives. In *Handbook of Visual Display Technology* (pp. 1501–1514). Springer International Publishing. https://doi.org/https://doi.org/10.1007/978-3-319-14346-0_197

Cattaneo, L., Fumagalli, L., Macchi, M., & Negri, E., Clarifying Data Analytics Concepts for Industrial Engineering. *IFAC-PapersOnLine*, *51*(11), 820–825, 2018. https://doi.org/10.1016/j.ifacol.2018.08.440

Clancy, R., O’Sullivan, D., & Bruton, K., Data-driven quality improvement approach to reducing waste in manufacturing. *TQM Journal*, *35*(1), 51–72, 2023. https://doi.org/10.1108/TQM-02-2021-0061

Coccia, R., Corsini, A., Bonacina, F., Lombardi, M., Celsi, L. R., Pellicanò, N. M., Zangheri, N., Savini, N., Santucci, F., Salini, S., & Malavisi, M. (2021). An Application of Data-Driven Analysis in Road Tunnels Monitoring. In F. M., L. M., T. M., D. P., S. M.M., C. F., D. G. G., & P. R. (Eds.), *Proceedings of the International Conference on Industrial Engineering and Operations Management* (pp. 1445 – 1455). IEOM Society. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85126212351&partnerID=40&md5=9393deecf0d1dabfa9faa1e0bb75002c

Cruz, S., Sousa, A., Viana, J. C., & Martins, T. , Analysis of the bonding process and materials optimization for mitigating the Yellow Border defect on optically bonded automotive display panels. *Displays*, *48*, 21–28, 2017. https://doi.org/https://doi.org/10.1016/j.displa.2017.02.003

Dahlström, T. (2022). *Facilitating quality management through data mining* [Master Thesis Dissertation, Mälardalen University]. http://urn.kb.se/resolve?urn=urn:nbn:se:mdh:diva-59180

Dhamija, P., & Bag, S. (2020). Role of artificial intelligence in operations environment: a review and bibliometric analysis. *The TQM Journal*, *32*(4), 869–896. https://doi.org/10.1108/TQM-10-2019-0243

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

George, M. L. (2003). *Lean six sigma for service*. McGraw-Hill.

Hoerl, R. W., Snee, R. D., & De Veaux, R. D., Applying statistical thinking to 'Big Data' problems. *Wiley Interdisciplinary Reviews: Computational Statistics*, *6*(4), 222–232, 2014. https://doi.org/10.1002/wics.1306

Koppel, S., & Chang, S. (2021). MDAIC – a Six Sigma implementation strategy in big data environments. *International Journal of Lean Six Sigma*, *12*(2), 432–449. https://doi.org/10.1108/IJLSS-12-2019-0123

Kregel, I., Stemann, D., Koch, J., & Coners, A. (2021). Process Mining for Six Sigma: Utilising Digital Traces. *Computers and Industrial Engineering*, *153*(December 2020), 107083. https://doi.org/10.1016/j.cie.2020.107083

Mozdzyn, L., & Rudolph, M. , Optical bonding makes its mark with touch panels and other displays. *Information Display*, *27*(5–6), 26–29, 2011. https://doi.org/https://doi.org/10.1002/j.2637-496X.2011.tb00392.x

Nader, J. (2022). Advanced Analytics Tools for Process Improvement: A Case Study in a Brewery. *2022 International Conference on Interdisciplinary Research in Technology and Management, IRTM 2022 - Proceedings*. https://doi.org/10.1109/IRTM54583.2022.9791633

Oliveira, R., Viana, J., & Sampaio, P. (2023). Optical bonding process of flat panel displays and their critical-to-quality factors. *The International Journal of Advanced Manufacturing Technology*, *125*(11–12), 5631–5650. https://doi.org/10.1007/s00170-023-10810-w

Pongboonchai-Empl, T., Antony, J., Garza-Reyes, J. A., Komkowski, T., & Tortorella, G. L. , Integration of Industry 4.0 technologies into Lean Six Sigma DMAIC: a systematic review. *Production Planning and Control*, 2023. https://doi.org/10.1080/09537287.2023.2188496

Skalli, D., Charkaoui, A., Cherrafi, A., Garza-Reyes, J. A., Antony, J., & Shokri, A. , Industry 4.0 and Lean Six Sigma integration in manufacturing: A literature review, an integrated framework and proposed research perspectives. *Quality Management Journal*, *30*(1), 16–40,2023. https://doi.org/10.1080/10686967.2022.2144784

Sony, M., Antony, J., Park, S., & Mutingi, M., Key Criticisms of Six Sigma: A Systematic Literature Review. *IEEE Transactions on Engineering Management*, *67*(3), 950–962, 2018. https://doi.org/10.1109/TEM.2018.2889517

Titmarsh, R., Assad, F., & Harrison, R., Contributions of lean six sigma to sustainable manufacturing requirements: An industry 4.0 perspective. *Procedia CIRP*, *90*, 589–593,2020. https://doi.org/10.1016/j.procir.2020.02.044

Veile, J. W., Kiel, D., Müller, J. M., & Voigt, K.-I., Lessons learned from Industry 4.0 implementation in the German manufacturing industry. *Journal of Manufacturing Technology Management*, *31*(5), 977–997, 2019. https://doi.org/10.1108/JMTM-08-2018-0270

Yeh, G., Rudolph, M. L., & Chen, J. Y. C. (2013). Minimizing the Impa ct of Bonding Induced LCD Defects. *SID Symposium Digest of Technical Papers*, *44*(1), 951–954. https://doi.org/https://doi.org/10.1002/j.2168-0159.2013.tb06380.x

Zheng, L., Zeng, C., Li, L., Jiang, Y., Xue, W., Li, J., Shen, C., Zhou, W., Li, H., Tang, L., Li, T., Duan, B., Lei, M., & Wang, P., Applying data mining techniques to address critical process optimization needs in advanced manufacturing. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1739–1748,2014. https://doi.org/10.1145/2623330.2623347

Zwetsloot, I. M., Kuiper, A., Akkerhuis, T. S., & de Koning, H., Lean Six Sigma meets data science: Integrating two approaches based on three case studies. *Quality Engineering*, *30*(3), 419–431,2018. https://doi.org/10.1080/08982112.2018.1434892

## Biographies

**Rui Oliveira** is a PhD student of the Advanced Engineering Systems for Industry Doctoral Program, at the University of Minho. He completed his Master Degree in Industrial Engineering in 2007 at the University of Minho. His research activities are developed under the Industrial Engineering and Management Research Line of the ALGORITMI Research Centre, which is intergraded in the LASI (Intelligent Systems Associated Laboratory). He is a member of the Research Group on Quality and Organizational Excellence.

**Paulo Sampaio** is graduated in Industrial Engineering and Management at the University of Minho in 2002 (5-year degree). He completed his PhD in Industrial Engineering in 2008 at the University of Minho. His research activities are developed under the Industrial Engineering and Management Research Line of the ALGORITMI Research Centre. He is the Coordinator of the Research Group on Quality and Organizational Excellence. He has co-authored or authored more than 200 publications, 160 of them ISI/Scopus indexed papers.

**Júlio C. Viana** received the bachelor’s degree in industrial engineering–plastics branch from the University of Minho in 1991, the M.Sc. degree in computer aided engineering–mechanical design from Queen’s University Belfast, U.K., in 1993, and the Ph.D. degree in science and engineering of polymers and composites from the University of Minho in 2000.,In 1991, he joined the University of Minho as a Professor, where he is currently an Associate Professor with the Department of Polymer Engineering. He is also the Director of the Institute for Polymers and Composites (IPC) Research Unit, University of Minho. His research interests include advanced polymer systems and manufacturing technologies, flexible electronics, and smart polymeric products