

# Assignment Brief and Front Sheet PGT

This front sheet for assignments is designed to contain the brief, the submission instructions, and the actual student submission for any WMG assignment. As a result the sheet is completed by several people over time, and is therefore split up into sections explaining who completes what information and when. Yellow highlighted text indicates examples or further explanation of what is requested, and the highlight and instructions should be removed as you populate 'your' section.

This sheet is only to be used for components of assessment worth more than 3 CATS (e.g. for a 15 credit module, weighted more than 20%; or for a 10 credit module, weighted more than 30%).

## **To be completed by the student(s) prior to final submission:**

Your actual submission should be written at the end of this cover sheet file, or attached with the cover sheet at the front if drafted in a separate file, program or application.

Student ID or IDs for group work	5569029
----------------------------------	---------

**To be completed (highlighted parts only) by the programme administration after approval and prior to issuing of the assessment; to be consulted by the student(s) so that you know how and when to submit:**

Date set	03/02/2025
Submission date (excluding extensions)	<b>3 March 2025 by 12 pm, UK time</b>
Submission guidance	Tabula link
Marks return date (excluding extensions)	31/03/2025
Late submission policy	<p>If work is submitted late, penalties will be applied at the rate of <b>5 marks per University working day</b> after the due date, up to a <b>maximum of 10 working days</b> late. After this period the mark for the work will be reduced to 0 (which is the maximum penalty). "Late" means <b>after the submission deadline time as well as the date</b> – work submitted after the given time even on the same day is counted as 1 day late.</p> <p>For <b>Postgraduate</b> students only, who started their <b>current course before 1 August 2019</b>, the daily penalty is <b>3 marks</b> rather than 5.</p>
Resit policy	<p>If you fail this module and/or component, the University allows students to remedy failure (within certain limits). Decisions to authorise resits are made by Exam Boards. These will be issued at specific times of the year, depending on your programme of study. More information can be found from your programme office if you are concerned.</p>

	If this is <b>already a resit</b> attempt, this means you will not be eligible for an additional attempt. The University allows as standard a maximum of two attempts on any assessment (i.e. only one resit). Students can only have a third attempt under exceptional circumstances via a Mitigating Circumstances Panel decision.
--	--

To be **completed** by the **module leader/tutor** prior to approval and issuing of the assessment; to be **consulted** by the **student(s)** so that you understand the assignment brief, its context within the module, and any specific criteria and advice from the tutor:

<b>Module title &amp; code</b>	WM9G1-15 Big Data Analytics for Industry
<b>Module leader</b>	Dr Leonardo Alves Dias
<b>Module tutor</b>	Liping Zheng
<b>Assessment type</b>	Essay
<b>Weighting of mark</b>	70%

<b>Assignment brief</b>
<p>You have been approached by a US Systems Automation company called SensorMesh. The company installs sensor networks in manufacturing production lines. It is known for its professionalism and reliability. The company now wants to expand its portfolio by using big data technologies to provide data analytics services for its sensor networks. However, its analytics and data management capabilities are quite limited, so it is asking your consultancy company to assist with the expansion.</p> <p>You have been hired as a consultant to provide insight into the challenges and opportunities of utilising big data technologies to manage sensor data from the manufacturing process of an industry. For this purpose, you must select a specific industry segment that manufactures their products (e.g., cars, pharmaceuticals, aeroplanes, cement, etc.) and write a business report about the data engineering and data analytics transition from a traditional to a big data methodology and how it can assist in the manufacturing process. Your report should address the following key areas:</p> <ol style="list-style-type: none"> <li>1. Analyse and critically evaluate the traditional approaches used in the chosen manufacturing industry segment and opportunities for utilising big data approaches.</li> <li>2. Identify and discuss the best practices of adopting a big data pipeline in a manufacturing process and provide recommendations for adopting these practices. Specifically, you need to choose two stages of the big data pipeline to give recommendations together with a real-world example/case study based on the best practices and provide.</li> <li>3. Discuss potential challenges associated with the implementation of your proposed recommendations, as well as mitigation solutions.</li> </ol>

References from both academic and commercial sources should be included to support your analysis and recommendations. Ensure your report is clear, objective, and well-structured, adhering to academic standards.	
<b>Word count</b>	The suggested word count is approximately 2800 (a plus or minus 10%-word limit is acceptable). The word limit applies only to the main body of the report. Therefore, it excludes the executive summary, table of contents, list of figures and tables, reference list, and appendices.
<b>Module learning outcomes (numbered)</b>	<ul style="list-style-type: none"> <li>L1. Critically evaluate the key differences between Big Data technologies and analysis methods and traditional approaches in engineering business management.</li> <li>L2. Critically evaluate real-world engineering scenarios/case studies and devise appropriate analytical solutions.</li> <li>L3. Demonstrate a comprehensive understanding of the core concepts of visual communication and data visualisation.</li> <li>L4. Collaboratively analyse engineering business requirements and practically implement analytics and optimisation techniques in real-world settings.</li> </ul>
<b>Learning outcomes assessed in this assessment (numbered)</b>	L1, L2, L3.
<b>Marking guidelines</b>	<i>See below.</i>
<b>Academic guidance resources</b>	Contact the e-business management tutors

#### Where to get help:

1. Talk to your module tutor if you don't understand the question or are unsure as to exactly what is required.
2. There are also numerous online courses provided by the University library to help in academic referencing, writing, avoiding plagiarism and a number of other useful resources.  
<https://warwick.ac.uk/services/library/students/your-library-online/>
3. If you have a problem with your wellbeing, it is important that you contact your personal tutor or wellbeing support services <https://warwick.ac.uk/services/wss>

## Contents

Introduction .....	5
Traditional Data Management Approaches.....	6
Big Data .....	10
Data Ingestion .....	11
Data Visualisation .....	13
Recommendations on Data Visualisation for SensorMesh based on above case study – .....	15
Challenges and their solutions for the Big Data Pipeline Stages .....	16

## List of Figures

Fig. 1 Automotive Data.....	5
Fig. 2 Visual Representation of Manual Data collection.....	8
Fig. 3 Limitations of Traditional Data Management.....	9
Fig.4 The Flow used for Data Ingestion in Farplas.....	13
Fig. 5 Failure Data Visualisation Page Volvo case Study.....	15

# Introduction

As the auto industry is facing challenges like globalization, it has now become more important to maintain the important data which possesses customer expectation along with environmental changes. However, this huge amount of data creates challenges, and to tackle those challenges Automotive manufacturers need to efficiently manage their data. Ultimately, implementing Big Data techniques is essential in order to improve product quality, reliability and sustainability along with being competitive in market (Johanson, Belenki et al. 2014). SensorMesh, a US-based automation company, is expanding its services by introducing big data analytics for sensor networks in auto manufacturing. This report focuses on the automotive manufacturing industry, examining the shift from traditional data management to big data approaches. It highlights best practices, offers recommendations, and presents real-world examples to help SensorMesh to improve data analytics capabilities and improve efficiency and product quality for its clients.



Fig. 1 Automotive Data (Williams 2024).

## Traditional Data Management Approaches

1. Based on the (KPMG 2020) report, below are some traditional data management tools and methods that have been used in Automotive industry.

- **OEM Managed Proprietary Systems** (KPMG 2020) – In the past car manufacturers used to collect and store vehicle data on their own dedicated servers and private databases. This meant that the data was kept in a separate group and had limited access to it.
- **In-House Legacy Data Collection** (KPMG 2020) - Most of the data was gathered using older manual methods or very basic automated systems. Since these systems were not connected to a larger digital network, the data could not be accessed in real time or used for advanced analysis.
- **Limited Use of Open APIs** – Older systems relied on private, closed interfaces to access and manage data instead of using modern, open interfaces which allowed easy sharing of data across different departments or with external partners.

**2. Traditional Data Ware Housing** (Ponnusamy 2023) - Traditional data warehousing started in the 1980s when companies needed better ways to analyse their data, beyond what regular transaction systems (OLTP) could do. These warehouses gathered historical data from different sources, stored it in an organised way, and allowed employees to access it using SQL queries. This helped companies get useful insights from their data. However, traditional data warehouses had several problems. Loading data into the system was slow, and companies needed large and expensive infrastructure to handle times when lots of data came in at once. Storage would also fill up quickly, and running and maintaining these systems required highly skilled technical experts.

- 3 In the automotive Industry, **Relational Database Management System** (RDBMS) has been used widely to manage the data sets. Relational databases like Oracle, SQL Server and PostgreSQL have been used for many years as the core system to manage structured data (Luckow, Kennedy et al. 2015).

#### **Traditional Role of Relational Database Management System**

1. **Transactional Data Management** (Luckow, Kennedy et al. 2015) – RDBMS platforms stored structured data generated from enterprise system like ERP systems. These data bases held critical information related to inventory, production schedules, procurement transactions and supply operations.
2. **Internal Reporting and Analysis** (Luckow, Kennedy et al. 2015) – Automotive companies traditionally relied on RDBMS for internal reporting such as monthly production reports, defect tracking and supplier performance evaluation.

**4. Manual Data Collection and Extraction** – It primarily involved gathering information from various systems like spreadsheets, offline databases and other internal databases. These sources usually lacked standardized structures which made it harder to collect the data collection processes. MS Excel was a main storage manipulation platform for this data. The initial steps involved manually extracting the raw data which described the shop floor behavior such as cycle times and downtimes, and then this data was moved from offline production databases into performance analysis tools using manual methods like copy and paste (Skoogh, Johansson et al. 2012). After collecting the data, the users had to manually filter out the errors and irrelevant points based on their knowledge to clean the data. Afterwards, they performed additional calculations to convert the raw data into simulation ready values like Time to Failure (TTF) with the help of spreadsheets. Tools like ExpertFit, were used to create statistical distributions to improve the accuracy of the data. Finally, the data was formatted and adjusted Discrete Event Simulation (DES) models (Skoogh, Johansson et al. 2012).



Fig. 2 Visual Representation of Manual Data collection (Data 2017)

#### **Supporting Case Study on Traditional Data Management (Liu, Rouse et al. 2015) –**

The article discusses about the failure of several automobile brands. The traditional approach for data management contributed to their downfall. These companies relied on the outdated, siloed system that collected and processed data separately, which made it difficult to see the full picture of the market trends and customer preferences. For example, many companies failed to notice shifts like growing demand for fuel efficient vehicles or the impact of the economic downturn, because the entire data management and forecasting was poor, they struggled to adapt the rapid change of the market behaviors.

If these brands had used modern and accurate data collection, data management tools and techniques which integrate and analyse the real time data, they could have made better decisions. For example, tracking the production trends, data of consumer behavior, or data of economic indicators in real time would have allowed for the faster response to the changing market (Liu, Rouse et al. 2015).



**Limitations in Traditional Data Management Approaches** – Above discussed traditional approaches often involve separate system across different departments to collect the data. This broken setup of data collection causes several problems which include **limited Real Time collaboration** which makes it difficult for teams to work together and work efficiently which further leads to delays and increases the chance of outdated information being used. Second, **version control issues** arise when different system store different versions of the same data. Moreover, **Data Inconsistency** between linked departments can result in conflicting or outdated information transfer which can lead to error, rework and production delays. **Manual process inefficiency** makes data entry and updates slow, also it **increases the chances of errors** as well. It also consumes more time which further leads to additional administration work (Sharma 2024). Finally, these combined limitations make the overall cost go high, lowered profitability and less productivity.



Fig. 3 Limitations of Traditional Data Management (Sharma 2024).

## Big Data

Big Data refers to large volume of digital data that cannot be effectively processed or managed by traditional data processing techniques due to its size, complexity, velocity and variety. It includes diverse data types and sources that require specialised technologies for storage, analysis and management (Khine and Shun 2017). A big data pipeline is a structured process that involves data collection which is often called as data ingestion, data storage, data processing, data analysis and finally data visualisation. All of the processes are aimed at transforming the raw data into actionable insights for informed decision making.

**Best Practices of Adopting Big Data Pipeline** – Implementing an effective big data pipeline is really crucial for manufacturing organisations which aiming to tackle the data driven insights (Labs 2024). Focusing on best practices can significantly improve the efficiency and reliability of the product.

1. **Define clear goals and requirements** (Labs 2024) - Identify the specific goal of your data pipeline, such as monitoring equipment performance or optimizing production schedules. Determine the frequency of data collection and their source like, machine sensors, operational database.
2. **Choose the right tools and technologies** – Select the appropriate data storage solutions like data warehouse or data lakes and processing tools like Apache Spark or Flink on the basis of requirement.
3. **Design for scalability** (Labs 2024) – Create modular pipeline in which components can be easily added, removed or modified without disrupting the entire system. This type of flexibility is really crucial as manufacturing processes evolve.
4. **Data Quality Checks** (Labs 2024)- Implement data quality checks at the point of data entry to identify the issues like missing values or duplicates. Early detection and correction prevent downstream problems and helps to improve reliable analytics.
5. **Automate Pipeline Deployment and Testing** – Utilise tools and scripts to automate the deployment and testing of the data pipeline. Automations ensures consistency, reduces manual efforts and enables faster repetition.

By following these practices, manufacturing organisations can develop data pipelines that are efficient, reliable, and scalable. This will help to improve the operational efficiency and support data driven decision making.

Integrating big data technologies into manufacturing can significantly boost its efficiency, product quality and decision making. Two important and major stages from the big data pipeline are data ingestion and data visualisation. Data Ingestion involves collecting and preparing data from various sources for analysis, while the data visualisation presents this data in an easily understandable format to aid decision making. Implementing best practices in these stages is really crucial for successful transition to big data.

## Data Ingestion

Data ingestion is the process of collecting and bringing data from different sources into a system, where it can be stored, analysed and used for making decisions. In simple terms data ingestion is the first step in turning raw data into useful information. In automotive manufacturing this involves collecting the data from different equipment, sensors and machines on production floor. Then the data is later cleaned, organised and converted to make sure it's accurate and consistent before being stored in a system or data base.

### Recommendations for Data Ingestion –

1. **Adopt Hybrid Data Ingestion Strategy** (Kotwal 2024) - Use both batch and real-time data ingestion methods to meet different data needs. For example, use real-time ingestion to capture continuous data from IoT sensors on machinery and employees, while batch ingestion can be used to analyse historical data from inventory systems and production logs. This approach makes sure that you get timely insights and maintain a full history of data records. Apache Kafka can be used for real time monitoring streaming and Apache NiFi for batch processing. However, Apache Flink can be used for both stream and batch data processing.
2. **Implement strict Data Validation** (Santos, Machado et al. 2007) - Set up automatic checks to monitor the quality of data throughout the ingestion process. These checks should ensure that the data is accurate, properly formatted, and consistent across different sources, like supply chains, production equipment, and quality assurance systems. Tools like Apache NiFi can automate these validation processes, which can help to reduce manual errors and ensuring the data is reliable.
3. **Focus on Scalable Architecture** (Kotwal 2024) - Choose a data ingestion system that can grow and adapt with your production needs. Since vehicle production can change depending on the market, make sure the system can handle large amounts of data, especially during busy production periods, without slowing down or losing data. Tools like Amazon Web Services (AWS) for cloud scalability, Google Cloud Performance (GCP) or Apache Hadoop for distributed storage and processing can be used.
4. **Maintain Flexibility Testing for Schema Evolution** (Santos, Machado et al. 2007) - Build systems that can easily handle changes in data structure. As new parts, technologies, or

production processes are introduced, having a flexible system makes sure that you can integrate them without major changes to the overall system. Apache Avro can be used for schema evolution or Google BigQuery can be implemented for handling dynamic schemas.

5. **Integrate Predictive Analytics** (Santos, Machado et al. 2007) - Use machine learning models to analyse the incoming data for predictive maintenance and quality control. For example, analyzing data from manufacturing equipment sensors can help predict when machines might fail, reducing downtime and keeping production on schedule.
6. **Implement Security Protocol** (Kotwal 2024) - Since manufacturing data, like designs, processes, and customer information, can be sensitive, it is really important to have strong security measures. Use encryption to protect data during transfer and set up role-based access controls to limit who can access critical data. Regular security checks and audits should also be done to comply with industry standards and protect data integrity. SSL can be used for encryption during data transmission.

#### **A Case Study of Successful Implementation of Data Ingestion – The Farplas Case Study .** (Kahraman, Onar et al. 2020)

Farplas, a leading automotive supplier in Turkey, implemented a successful data ingestion system through a pilot project designed for intelligent production. They created a pilot injection cell to test and develop scalable data collection solutions for their injection molding machines. The setup included a plastic injection molding machine, 6-axis robots, energy analysers, water collectors, precision scales, humidity sensors, and temperature sensors. These devices communicated using various protocols, such as Euromap63, which is specific to injection molding machines . (Kahraman, Onar et al. 2020). The data from Euromap63 sensors is sent to a central Programmable Logic Controller (PLC) via Ethernet. Other components, like energy analysers and sensors, are also connected to the PLC. The parameters are standardized using OPC software (Kepware). After processing, the data is transferred via MQTT protocol to Farplas's big data platform, called Flatform. At Flatform, Apache Ni-Fi is used to create data pipelines, process the data, and transform it into JSON format. The transformed data is then sent to Apache Kafka for distributed streaming. Kafka stores the data in HDFS for scalable storage and Elasticsearch for fast querying and analysis. Kafka also distributes the data to other platforms for further processing (Kahraman, Onar et al. 2020). This approach allowed Farplas to manage large volumes of data in real time, overcoming challenges like data variety, size, and speed. The data ingestion pipeline improved operational efficiency, provided real-time analytics, and enabled better decision-making.

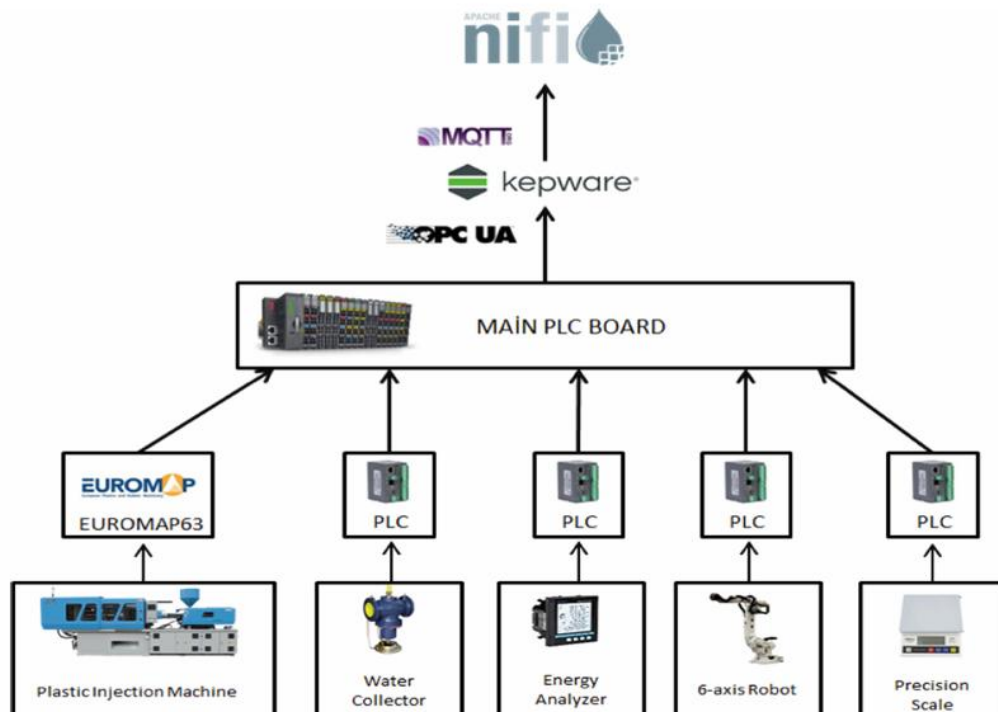


Fig.4 The Flow used for Data Ingestion in Farplas (Kahraman, Onar et al. 2020).

**Data Visualisation** – Data Visualisation is the graphical representation of information and data using visual elements like charts, graphs and maps. This technique provides an accessible way to identify and understand patterns, trends and outliers within the datasets (Tableau). By translating complex data into visual context, data visualisation makes information more comprehensible, improving better decision making along with better communication.

**Supportive Case Study for Data Visualisation** – Using Data Visualisation to find faults at Volvo car corporation (Jones and Da Silva Martins 2017).

The Research and Development team at Volvo Car Corporation (VCC) faced major challenges when trying to find and fix faults in their vehicle’s complex electronic system. Each Volvo vehicle contains more than 70 Electric Control Units (ECU) that work together and when something went wrong during testing, it was very hard to figure out the root cause. The process took a lot of time

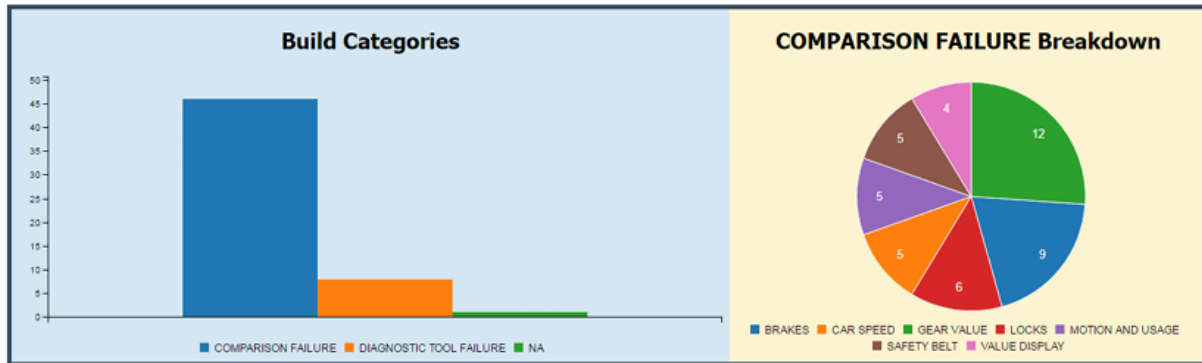
depended heavily on the knowledge of experienced experts. Engineers had to manually read through long, text-based logs, which was slow and confusing. To overcome this problem, Volvo worked closely with researchers to create a special data visualisation tool that would make it easier to find and understand the test failures. This tool collect historical data from past failures and grouped the failures in different categories (Jones and Da Silva Martins 2017). These categories were then displayed using interactive visual dashboards. Instead of searching through pages of logs, engineers could see clear charts and graphs that highlighted the most common types of failures and helped them go deeper with the specific issues.

The tool was linked directly with Volvo's automated testing system, where every new software update was tested on real hardware through hardware-in-the-loop or HIL testing. The visual dashboard included different types of charts and graphs. Bar charts were used to show often each type of failure occurred, Pie chart was used to breakdown failures into smaller sub groups (Jones and Da Silva Martins 2017). However, the Trend Analysis Graphs used to display failures occurring over time.

By using these visual tools, Volvo's engineers spent much less time searching for faults. They were also more accurate at finding the real causes of problems, which led to better testing processes overall. With both short-term and long-term failure trends available in visual form, Volvo's teams could spot recurring problems, improve their test coverage, and increase the overall quality of their vehicle software as well as hardware.

This case study shows how using data visualisation tools alongside automated data collection can turn a slow, manual process into a faster, smarter, and more data-driven system. Volvo's success proves that companies like SensorMesh can use similar visualisation tools to help their manufacturing clients monitor sensor data, spot problems faster, and improve product quality.

## Failures Summary



## Test Case Failures

Test Suite	Test Case ID	Test Case Name	Category	Warn Message	Fail Message
HIL » HIL stability 94032 » BasicDrivingCycle	s1-s2-t5	Simple drive cycle - slow acceleration - soft braking	COMPARISON FAILURE	FAILED CHECK: car speed is greater than or equal to [10]=10.0 -- Actual value: 2.53368 [m/s]	Comparison failure
PKE » PKE Verified » PKE 017 S-17-04 GearPositionWarning	s1-s4-t1	S-17-04 Gear Position Warning	COMPARISON FAILURE	FAILED CHECK: usage mode is driving -- Actual value: UsgModInActv	Comparison failure
PKE » PKE Verified » PKE 017 S-17-04 GearPositionWarning	s1-s4-t2	Disengage EPB and drive off	COMPARISON FAILURE	FAILED CHECK: electrical parking brake status is released -- Actual value: applied	Comparison failure
PKE » PKE Verified » PKE 022 S-22-04 L1-RatioManagement	s1-s5-t1	S-22-04 L1-Ratio Management	COMPARISON FAILURE	FAILED CHECK: usage mode is driving -- Actual value: UsgModInActv	Comparison failure
PKE » PKE Verified » PKE 022 S-22-04 L1-	e1-c6-t2	Verify gears changing	COMPARISON	FAILED CHECK: actual gear is 1 --	Comparison failure

Fig. 5 Failure Data Visualisation Page Volvo case Study (Jones and Da Silva Martins 2017)

## Recommendations on Data Visualisation for SensorMesh based on above case study –

- Select the Appropriate Visualisation Tool** (Chapman 2024) - Choose tools that align with the organisation's needs and capabilities. Tools like Tableau offer various options, including desktop applications and online versions, allowing users to create interactive and shareable dashboards.
- Show Historical Trends** (Jones and Da Silva Martins 2017) – Along with live data, SensorMesh should show the historical data as well as mentioned in the Volvo case study, so that users can spot long term patterns and recurring issues.
- Create a customisable Dashboard** - SensorMesh should offer role-specific dashboards that are customisable by users themselves, allowing managers, operators, and engineers to personalise the data views as per their requirement (Jones and Da Silva Martins 2017). Managers might focus on high-level KPIs, while operators may need detailed machine data. Allowing user-defined filters, thresholds, and chart types ensures each team gets exactly the insights they need without information overload.

## Challenges and their solutions for the Big Data Pipeline Stages

1. **High Initial Investment Cost** (Brettel, Friederichsen et al. 2017) - Implementing real-time data ingestion pipelines, visualisation tools, and advanced analytics platforms requires significant upfront investment in hardware, software, and skilled personnel. This is especially challenging for small to mid-sized automotive manufacturers, who may lack the budget for comprehensive digital transformation.

**Solution** – SensorMesh can offer phased implementation starting with a smaller project first which has single production line before scaling it to the larger quantity.

2. **Complexity of Data** - Automotive manufacturing generates high-dimensional data from various sensor readings, machine downtimes, supply chain metrics which are difficult to represent coherently.

**Solution** - Tableau is a leading data visualization tool that enables users to transform complex datasets into intuitive and interactive visual representations(Batt, Grealis et al. 2020). To simplify the complex data sets SensorMesh can include software like Tableau to simplify the data sets.

3. **Resistance to change and skill gap** - Many manufacturing plants operate with long-serving staff who are used to traditional paper-based or spreadsheet-based processes. Transitioning to real-time dashboards and big data systems may face resistance, and staff may lack the necessary digital skills.

**Solution** – SensorMesh can arrange the training programs designed to different user levels like, operator, supervisor and manager



## List of References

- Batt, S., et al. (2020). "Learning Tableau: A data visualization tool." the Journal of economic education **51**(3-4): 317-328.
- Chapman, C. (2024). "A Complete Overview of the Best Data Visualization Tools." Retrieved 02/03/2025, 2025, from <https://www.toptal.com/designers/data-visualization/data-visualization-tools>.
- Data, H. o. (2017). 4 Data Collection Techniques: Which One's Right for You? Retrieved 02/03/2025, 2025, from <https://humansofdata.atlan.com/2017/08/4-data-collection-techniques-ones-right/>.
- Jones, M. and R. Da Silva Martins (2017). "Visualization of Test Failure Data to Support Fault Localisation in Distributed Embedded Systems within the Automotive Industry."
- Johanson, M., et al. (2014). Big automotive data: Leveraging large volumes of data for knowledge-driven product development. 2014 IEEE international conference on big data (Big Data), IEEE.
- Kahraman, C., et al. (2020). Intelligent and Fuzzy Techniques: Smart and Innovative Solutions: Proceedings of the INFUS 2020 Conference, Istanbul, Turkey, July 21-23, 2020, Springer Nature.
- Khine, P. P. and W. Z. Shun (2017). "Big data for organisations: a review." Journal of Computer and Communications **5**(3): 40-48.
- Kotwal, A. P. (2024). "DATA INGESTION PATTERNS IN BIG DATA SYSTEMS: AN ANALYSIS OF METHODS, ARCHITECTURES, AND IMPLEMENTATION STRATEGIES." INTERNATIONAL JOURNAL OF COMPUTER ENGINEERING AND TECHNOLOGY (IJCET) **15**(6): 1816-1826.
- KPMG (2020). Automotive Data Sharing. Automotive Data Sharing. Website: 2-11.
- Labs, D. (2024). "What is a Data Pipeline? Types, Architecture, Best Practices & Use Cases." Retrieved 1/03/2025, 2025, from <https://www.dqlabs.ai/blog/what-is-a-data-pipeline-types-architecture-components/>.
- Liu, C., et al. (2015). "When transformation fails: Twelve case studies in the american automobile industry." Journal of Enterprise Transformation **5**(2): 71-112.
- Lipovac, I. and M. B. Babac (2024). "Developing a data pipeline solution for big data processing." International Journal of Data Mining, Modelling and Management **16**(1): 1-22.
- Luckow, A., et al. (2015). Automotive big data: Applications, workloads and infrastructures. 2015 IEEE International Conference on Big Data (Big Data), IEEE.

Munappy, A. R., et al. (2020). Data pipeline management in practice: Challenges and opportunities. Product-Focused Software Process Improvement: 21st International Conference, PROFES 2020, Turin, Italy, November 25–27, 2020, Proceedings 21, Springer

Ponnusamy, S. (2023). "Evolution of Enterprise Data Warehouse: Past Trends and Future Prospects." International Journal of Computer Trends and Technology **71**(9): 1-6.

Raptis, T. P., et al. (2019). "Data management in industry 4.0: State of the art and open challenges." IEEE Access **7**: 97052-97093.

Santos, J. N. M. F., et al. (2007). "Progress in Artificial Intelligence."

Sharma, S. (2024). "Streamlining Automotive Product Data Management with Centralized Systems." Growing Pains of Traditional Data Management in Automotive Aftermarket. Retrieved 1/03/2025, 2025, from <https://www.credencys.com/streamlining-automotive-product-data-management/>.

Skoogh, A., et al. (2012). "Automated input data management: evaluation of a concept for reduced time consumption in discrete event simulation." Simulation **88**(11): 1279-1293.

Tableau. "Data visualisation beginner's guide: a definition, examples and learning resources." Data visualisation beginner's guide: a definition, examples and learning resources. Retrieved 03/03/2025, 2025, from <https://www.tableau.com/en-gb/learn/articles/data-visualization>.

Wedeniowski, S. (2015). The Mobility Revolution in the Automotive Industry, Springer.

Williams, J. (2024). "Unveiling Top Big Data Visualization Tools – A Guide for SMEs." Big Data Visualization Tools for Businesses. Retrieved 25/02/2025, 2025, from <https://www.promptcloud.com/blog/6-big-data-visualisation-tools-for-you/>.