

**A REPORT
ON
DATA ANALYSIS IN THE PHARMACEUTICAL INDUSTRY**

Submitted by,

Ms. MANYA A J - 20211CSE0571

Under the guidance of,

Dr. Joseph Michael Jerard V

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

At



PRESIDENCY UNIVERSITY

BENGALURU

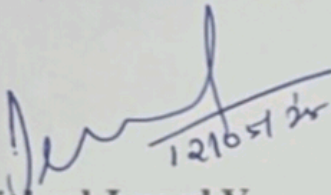
MAY 2025

PRESIDENCY UNIVERSITY

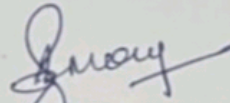
PRESIDENCY SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the Internship report "DATA ANALYSIS IN THE PHARMACEUTICAL INDUSTRY" being submitted by "Ms. Manya A J" bearing roll number "20211CSE0571" in partial fulfillment of the requirement for the award of the degree of Bachelor of Technology in Computer Science and Engineering is a Bonafide work carried out under my supervision.



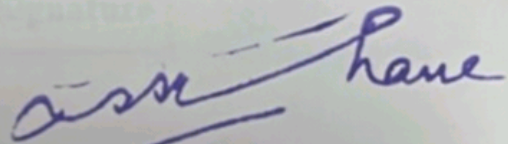
Dr. Joseph Michael Jerard V
Professor
PSCS
Presidency University



Dr. Asif Mohammed
HoD, CSE
PSCS
Presidency University



Dr. MYDHILI NAIR
Associate Dean
PSCS
Presidency University



**Dr. SAMEERUDDIN
KHAN**
Pro-Vice Chancellor -
Engineering
Dean -PSCS / PSIS
Presidency University

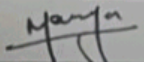
PRESIDENCY UNIVERSITY

PRESIDENCY SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

DECLARATION

I hereby declare that the work, which is being presented in the report entitled "DATA ANALYSIS IN THE PHARMACEUTICAL INDUSTRY" in partial fulfillment for the award of Degree of Bachelor of Technology in Computer Science and Engineering, is a record of my own investigations carried under the guidance of **Dr. Joseph Michael Jerard V**, Presidency School of Computer Science and Engineering, Presidency University, Bengaluru.

I have not submitted the matter presented in this report anywhere for the award of any other Degree.

Name	MANYA A J
Roll No	20211C5E0571
Signature	

ABSTRACT

This internship report details the work carried out during my tenure at Datazymes, where I contributed to data analysis within the pharmaceutical sector. The primary objective of my role was to work with real-world healthcare datasets, analyze emerging trends, and design interactive dashboards using Power BI and Tableau. Throughout my internship, I was involved in the end-to-end process of data analysis, including ETL (Extract, Transform, Load) processing, data modeling, and generating actionable insights to support business decisions. I also worked on enhancing the accessibility and visual appeal of reports to ensure effective communication of findings to stakeholders. This experience allowed me to gain hands-on expertise in handling large-scale data, working with complex datasets, and navigating the various stages of reporting workflows. Furthermore, I was exposed to industry-specific challenges, such as understanding pharmaceutical terminology and ensuring data accuracy in the context of regulatory requirements. Ultimately, this internship offered valuable insights into the intricacies of the healthcare industry, data-driven decision-making, and the importance of clean, well-structured data for meaningful analysis.

ACKNOWLEDGEMENTS

Firstly, I am indebted to the **GOD ALMIGHTY** for giving me an opportunity to excel in our efforts to complete this project on time.

I express my sincere thanks to our respected dean **Dr. Md. Sameeruddin Khan**, Pro-VC - Engineering and Dean, Presidency School of Computer Science and Engineering & Presidency School of Information Science, Presidency University for getting me permission to undergo the Internship.

I express my heartfelt gratitude to our beloved Associate Dean Dr. Mydhili Nair, Presidency School of Computer Science and Engineering, Presidency University, and Dr. Asif Mohammed, Head of the Department, Presidency School of Computer Science and Engineering, Presidency University, for rendering timely help in this internship successfully.

I am greatly indebted to our guide **Dr. Joseph Michael Jerard V, Professor**, Presidency School of Computer Science and Engineering, Presidency University, for his inspirational guidance, and valuable suggestions and for providing me with a chance to express my technical capabilities in every respect for the completion of the internship work.

I would like to convey my gratitude and heartfelt thanks to the PIP4004 Internship Coordinator **Mr. Md Ziaur Rahman** and **Dr. Sampath A K**, and Git hub coordinator **Mr. Muthuraj**.

I thank my family and friends for the strong support and inspiration they have provided me in this internship.

Manya A J

TABLE OF CONTENTS

Chapter No.	Title	Page No.
	ABSTRACT	iv
	LIST OF FIGURES	vii
	LIST OF ABBREVIATIONS	viii
Chapter 1	INTRODUCTION	1
Chapter 2	WEEK 1 EXCEL	2
Chapter 3	WEEK 2 PHARMA	4
Chapter 4	WEEK 3 DATA WAREHOUSING & MODELLING	6
Chapter 5	WEEK 4 & 5 POWER BI	10
Chapter 6	WEEK 6 & 7 TABLEAU	13
Chapter 7	WEEK 8-11 DASHBOARD REPLICATION IN TABLEAU AND POWER BI	17
Chapter 8	WEEK 12-16 LIVE PROJECT	21
Chapter 9	WEEK 17,18 LIVE PROJECT	23
Chapter 10	TIMELINE FOR EXECUTION	24
Chapter 11	OUTCOMES	
Chapter 12	CONCLUSION	25

LIST OF FIGURES

FIGURE NO	NAME OF THE FIGURE
2.1	Excel
3.1	Pharma
4.1	Data Warehousing
4.2	SQL Query Execution
5.1	Flow Process in Power BI
6.1	Tableau Flow Process
7.1	Sample Dashboard

LIST OF ABBREVIATIONS

Sl No.	ABBREVIATION	EXPANSION
1	BI	Business Intelligence
2	KPI	Key Performance Indicator
3	SQL	Structured Query Language
4	ETL	Extract, Transform, Load
5	HIPPA	Health Insurance Portability and Accountability Act
6	NCFB	Non-Cystic Fibrosis Bronchiectasis
7	SSC	Systemic Sclerosis
8	IPF	Idiopathic Pulmonary Fibrosis
9	GPP	Generalized Pustular Psoriasis

Chapter 1

INTRODUCTION

My internship learning journey has been a comprehensive and rewarding experience, offering deep insights into pharmaceutical analytics, data visualization, and advanced data manipulation. It significantly enhanced both my technical expertise and business communication skills. I began with foundational training in Excel and a strong emphasis on information security, particularly within the pharma context—which established a solid understanding of the importance of data protection in this industry. As the internship progressed, I explored medical terminologies and advanced data modeling techniques, equipping myself to manage complex datasets and extract meaningful insights. I also gained hands-on experience with tools like Power BI and Tableau, where I focused on building interactive dashboards, defining key performance indicators (KPIs), and mastering advanced visualizations to effectively showcase data trends. Additionally, I was introduced to core aspects of pharmaceutical brand analysis, prescription trend monitoring, and patient analytics, which provided a well-rounded view of data interpretation and its impact on business decisions. My ability to validate, report, and communicate data-driven findings to stakeholders was further strengthened throughout this journey. Overall, this internship not only sharpened my skills in working with large datasets but also empowered me to convert data into actionable insights—preparing me to contribute effectively to the dynamic field of pharma analytics.

Chapter 2

WEEK 1

EXCEL

The Excel training session was aimed at enhancing skills in data refinement, manipulation, and analysis to improve business intelligence and reporting efficiency.

Key Learnings

1. Data Formatting & Operations

- Gained proficiency in cell formatting (CTRL + 1), managing columns (hide/unhide/freeze), and applying filters (ALT + A + T).
- Enhanced productivity through quick selection techniques (CTRL + Shift + Down) and the use of absolute references (\$) for maintaining structured formulas.
- Learned to extract components from date values using functions like MONTH(), YEAR(), and DATE().

2. Array & Aggregate Functions

- Applied array formulas for efficient bulk calculations (e.g., =SUM(B3:B7 * C3:C7)).
- Utilized key aggregation functions such as COUNT, SUM, AVERAGE, MAX, and MIN to summarize and analyze datasets.

3. Logical & Conditional Functions

- Implemented logic-based functions including IF, SUMIF, SUMIFS, COUNTIF, and AVERAGEIF to handle conditional scenarios and evaluate data dynamically.

4. Text Functions for Data Cleaning

- Improved data quality using functions like CONCATENATE, UPPER, LOWER, LEFT, RIGHT, MID, FIND, and LEN for text extraction and formatting.

5. Advanced Lookup Functions

- Developed skills in dynamic data retrieval through advanced lookup tools like VLOOKUP, INDEX, and MATCH.

6. Advanced Excel Features

- Enabled and utilized the Developer tab for automation via Macros and basic VBA scripting.
- Created and analyzed PivotTables to summarize large datasets effectively.
- Performed scenario-based analysis and optimization using tools such as Solver and Goal Seek.

7. Additional Functions & Productivity Enhancements

- Applied functions like COUNTA, COUNTBLANK, RAND, RANDBETWEEN, and UNIQUE for advanced data tracking and manipulation.
- Enhanced worksheet navigation and usability through features like Freeze Panes (ALT + W + F).

8. Power Query (Excel 201)

- Explored advanced data transformation techniques similar to Power BI's Query Editor.
- Gained hands-on experience in merging, appending tables, and preprocessing datasets for structured analysis and reporting.

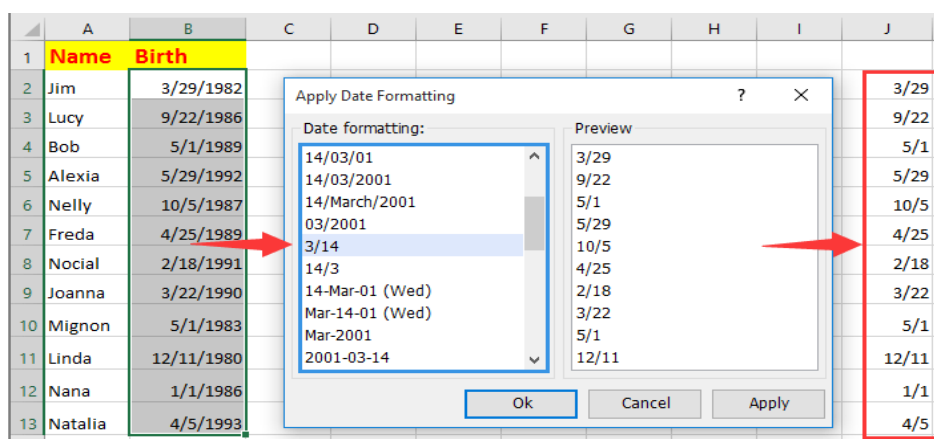


Fig 2.1. Excel

Chapter 3

WEEK 2

PHARMA TRAINING

The Pharma training provided in-depth exposure to the pharmaceutical industry, healthcare ecosystem, and broader life sciences sector. It emphasized the end-to-end drug development lifecycle, sales and marketing strategies, regulatory requirements, and the pivotal role of data in driving informed decision-making.

Key Learnings

1. Pharmaceutical Industry & Drug Lifecycle

- Gained a comprehensive understanding of the drug development process, from initial discovery through preclinical and clinical trials, regulatory approval (e.g., FDA), market launch, patent expiration, and eventual generic competition.
- Recognized the importance of accurately identifying target patient populations to maximize treatment outcomes and minimize risks.

2. Sales & Marketing in Pharmaceuticals

- Learned to distinguish between the primary customers (physicians) and end users (patients), a key factor in shaping effective sales strategies.
- Identified major stakeholders in the healthcare supply chain, including prescribers, pharmacies, PBMs (Pharmacy Benefit Managers), insurance providers, and government programs such as Medicare and Medicaid.
- Explored lifecycle management strategies like evergreening, product flanking, and Rx-to-OTC conversions to extend market exclusivity and maintain product relevance.

3. Regulatory Environment & Compliance

- Developed a clear understanding of regulatory processes including FDA approvals, drug indications, and the concept of off-label use.
- Gained familiarity with essential coding systems such as the National Drug Code (NDC) for drug identification and the International Classification of Diseases (ICD) for clinical data standardization.

4. Healthcare Models & Insurance Systems

- Examined global healthcare models, including the Beveridge (publicly funded), Bismarck (insurance-based), out-of-pocket, and national insurance frameworks.

-
- Learned the purpose and structure of U.S. healthcare programs like Medicare (for the elderly and disabled) and Medicaid (for low-income populations).
- 5. Data-Driven Insights in Pharma**
- Explored the growing role of analytics and AI in identifying high-value customers, forecasting market dynamics, and supporting strategic business decisions.
 - Studied tier-based drug pricing and formulary placements as tools to manage healthcare costs while ensuring patient access to medications.

Conclusion

This Pharma training offered a well-rounded understanding of the pharmaceutical industry's operational and strategic dimensions. By integrating knowledge of regulatory systems, market dynamics, and patient-centered approaches with data analytics, the program enhanced my ability to interpret healthcare trends and contribute effectively to data-driven initiatives in the life sciences sector. I now feel more equipped to support the transformation of healthcare through informed analytics and strategic insight.

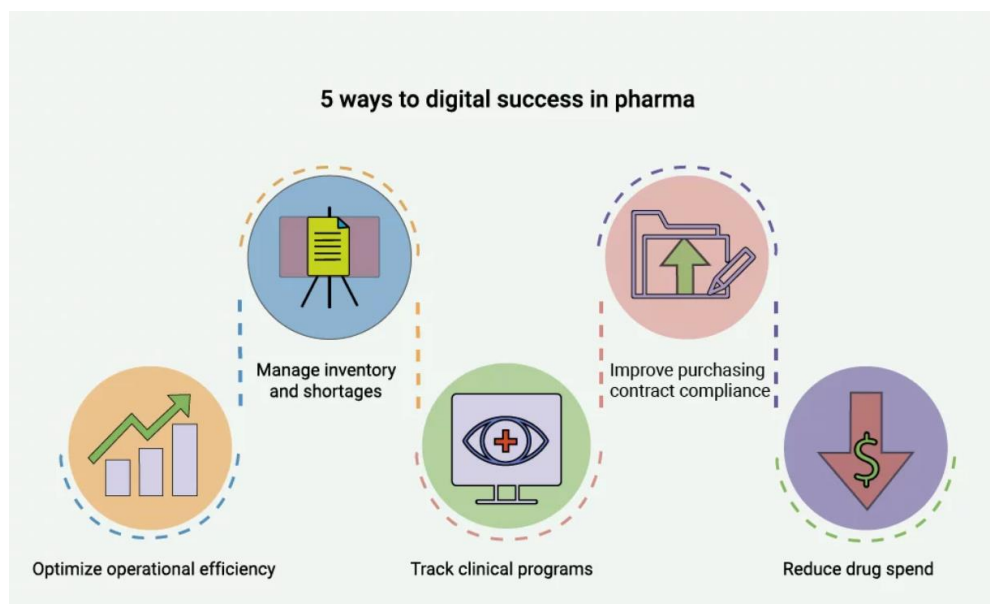


Fig 3.1. Pharma

Chapter 4

WEEK 3

4.1 DATA WAREHOUSING & MODELLING

This week's training offered a detailed exploration of data warehousing, data modeling, and relational database management systems—key components in supporting efficient data storage, organization, and analysis for business intelligence and reporting solutions.

Key Learnings

1. Data Warehousing & Storage Systems

- Gained insights into a variety of data storage architectures, including Data Lakes, Data Warehouses, Data Marts, Traditional Databases, and Operational Data Stores (ODS).
- Understood how each system is utilized to manage structured and unstructured data depending on specific business requirements and analytical use cases.

2. Fact and Dimension Tables

- Learned the critical role of fact and dimension tables in organizing data for effective reporting and analytics.
- Studied common schema designs such as the **Star Schema** and **Snowflake Schema**, which are used to enhance query efficiency and improve performance.
- Explored various dimension types, including **Slowly Changing Dimensions (SCDs)**, **Role-Playing**, **Conformed**, and **Junk Dimensions**, and their specific applications in modeling.

3. Data Modeling Concepts

- Built a solid foundation in data modeling fundamentals, including entities, attributes, constraints, and key relationship types (one-to-one, one-to-many, many-to-many).
- Examined different data modeling methodologies such as **Hierarchical**, **Network**, and **Relational Models**.
- Gained clarity on the three levels of data modeling: **Conceptual**, **Logical**, and **Physical**, which provide a structured approach to designing and implementing robust database systems.

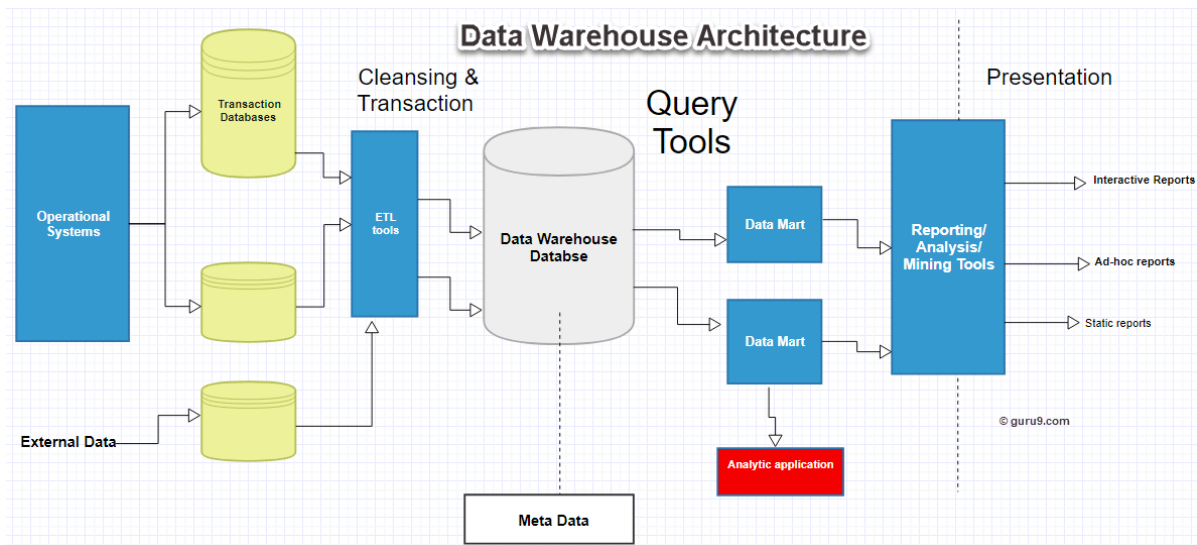


Fig 4.1..Datawarehouse Architecture

4.2. SQL TRAINING

The SQL training sessions provided a solid foundation in relational database management systems (RDBMS) and advanced data querying techniques using Structured Query Language (SQL). Emphasizing both conceptual understanding and hands-on experience, the program was designed to develop proficiency in managing, manipulating, and retrieving data effectively in real-world business scenarios.

Key Learnings

1. SQL Command Categories

- **Data Definition Language (DDL):**
 - **CREATE TABLE:** Defines new tables, including columns and constraints.
 - **DROP TABLE:** Removes an entire table and its data.
 - **ALTER TABLE:** Updates a table's structure by adding, modifying, or deleting columns.
- **Data Manipulation Language (DML):**
 - **INSERT INTO:** Adds new rows to a table.
 - **UPDATE:** Alters existing data based on conditions.
 - **DELETE:** Removes data from a table under specified criteria.
- **Data Query Language (DQL):**
 - **SELECT:** Retrieves specific data based on filtering and sorting criteria.

-
- **Data Transaction Language (DTL):**
 - `COMMIT`: Saves all operations in a transaction.
 - `ROLLBACK`: Reverts changes made within a transaction to maintain data consistency.
 - **Data Control Language (DCL):**
 - `GRANT`: Provides user access rights.
 - `REVOKE`: Removes previously granted permissions.

2. Advanced SQL Concepts

- **Aggregation & Grouping:**
 - Utilized `GROUP BY` and aggregation functions (`SUM()`, `AVG()`, `COUNT()`, etc.) to summarize data.
 - Applied `HAVING` to filter aggregated data based on conditions.
- **Sorting & Filtering:**
 - Used `ORDER BY` to sort query results.
 - Applied `WHERE` clauses for conditional filtering before aggregation.
- **Joins & Relationships:**
 - Explored multiple types of joins:
 - `INNER JOIN`, `LEFT JOIN`, `RIGHT JOIN`, `FULL OUTER JOIN`, and `CROSS JOIN` to combine data across related tables.
- **Subqueries & Common Table Expressions (CTEs):**
 - Implemented subqueries for nested filtering and aggregations.
 - Used `WITH` clause for defining CTEs to improve readability and modularity of complex queries.
- **Handling NULL Values:**
 - Managed missing data using `IS NULL` and `IS NOT NULL`.
- **Set Operators:**
 - Employed `UNION`, `UNION ALL`, `INTERSECT`, and `EXCEPT` to merge or differentiate result sets.
- **Window Functions:**
 - Used functions like `ROW_NUMBER()`, `RANK()`, `DENSE_RANK()`, and `NTILE()` to rank or number rows.
 - Leveraged the `OVER()` clause for advanced analytics across partitions.
- **Indexes & Performance Optimization:**
 - Understood how indexes enhance performance by reducing full table scans.
 - Distinguished between **clustered** and **non-clustered** indexes.
- **Views:**
 - Created simplified, reusable query structures using `CREATE VIEW`.
- **Normalization & Database Design:**
 - Learned principles of normalization to eliminate redundancy and ensure data integrity.

-
- Explored Normal Forms (1NF, 2NF, 3NF) for designing efficient database structures.

Assessments and Documentation

Each session was reinforced with assignments, practical labs, and assessments, allowing hands-on application of learned concepts. Tasks ranged from writing basic queries to modeling entire datasets and optimizing complex SQL statements. Comprehensive documentation of learnings, query examples, and debugging efforts was maintained throughout, ensuring systematic progress and deeper understanding.

Conclusion

This SQL training has equipped me with essential skills to manage relational data efficiently and apply complex querying techniques in business environments. With a strong grasp of database design, query optimization, and data handling best practices, I am now prepared to contribute to data analytics and business intelligence initiatives. The structured approach to learning has sharpened both my technical and analytical acumen, reinforcing the importance of precision in database management.

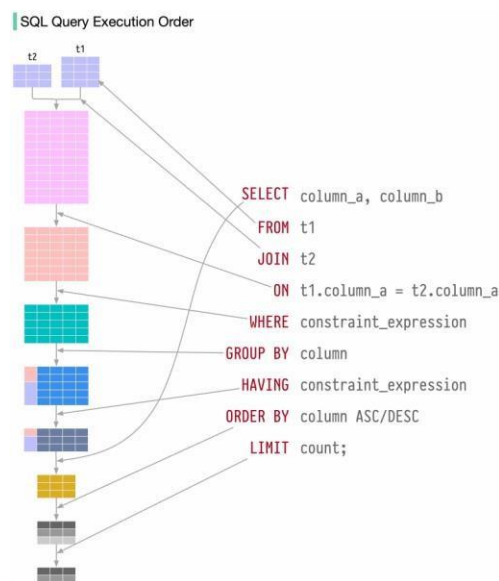


Fig 4.2.SQL Query Execution order

Chapter 5

WEEK 4 & 5

POWER BI TRAINING

The Power BI training sessions provided a thorough and practical understanding of how to transform raw data into actionable business insights through visually compelling and interactive dashboards. This hands-on experience covered the entire data analysis lifecycle—from data ingestion and modeling to visualization and report sharing—strengthening our ability to support data-driven decision-making in dynamic business environments.

Key Areas of Learning

1. Data Modeling

A strong foundation in data modeling was a major focus of the training. Key learnings included:

- Creating relationships between tables using primary and foreign keys for efficient model design.
- Using **DAX (Data Analysis Expressions)** to build calculated columns, measures, and tables that extract meaningful insights.
- Applying **Star Schema** modeling best practices to streamline performance and simplify analytics.
- Managing complex scenarios such as **many-to-many relationships** and **composite models** (combining DirectQuery and Import modes).

These modeling techniques helped improve performance, ensured accurate aggregations, and minimized report complexity.

2. Visualizations

Developing effective visuals was essential for conveying insights clearly and persuasively. We explored:

- Core visual types such as **bar charts, line graphs, pie charts, tree maps, waterfalls, and scatter plots**.
- Integration of **custom visuals** from Microsoft AppSource to enhance report versatility.

-
-
- Best practices in visual selection based on data types and audience needs.
 - Use of **color schemes**, **conditional formatting**, and **tooltips** to increase interpretability and visual appeal.

The training emphasized that impactful dashboards prioritize clarity and simplicity over volume.

3. Interactivity

To ensure reports were engaging and user-driven, we learned to implement:

- **Slicers** for dynamic filtering.
- **Drillthrough pages** for deeper insights from summary views.
- **Cross-filtering and cross-highlighting** to allow intuitive navigation across visuals.
- Use of **Bookmarks**, **Selection Pane**, and **Buttons** for guided storytelling and enhanced navigation.

These interactive features empowered end-users to explore data at their own pace, based on their unique information needs.

4. Dashboard Design Principles

Beyond functionality, we were taught to design dashboards with an emphasis on user experience and aesthetics:

- Use of **consistent layouts**, **fonts**, and **color themes** to create a professional look and feel.
- Logical organization using **grids**, **cards**, and **containers**.
- Techniques to optimize performance by minimizing heavy visuals and complex DAX.
- Responsive design practices to ensure compatibility across desktop, tablet, and mobile devices.

Designing with the user in mind helped us build dashboards that are not only informative but also easy to use and navigate.

5. Publishing, Sharing, and Collaboration

We also covered best practices for deployment and collaboration in a production environment:

- Publishing reports to the **Power BI Service**.
- Configuring **scheduled data refreshes** for real-time updates.
- Implementing **Row-Level Security (RLS)** to ensure appropriate data access.
- Sharing content with stakeholders via secure links, workspaces, and **Power BI Apps**.

-
- Leveraging features like **subscriptions**, **comments**, and **version control** to enhance collaboration and governance.

Understanding these practices was essential for building secure, scalable, and collaborative BI solutions.

Assignments and Assessments

Each topic was reinforced with practical assignments simulating real-world business use cases. These included dashboards for:

- **Sales performance**
- **Financial metrics**
- **Market segmentation**
- **Customer behavior analysis**

We documented each assignment meticulously—capturing the problem statement, data preparation steps, DAX logic, visualization choices, and key takeaways. This structured approach enhanced our problem-solving skills and analytical thinking.

Conclusion

In summary, the Power BI training significantly elevated my ability to design, develop, and deliver meaningful analytics solutions. It enhanced my technical skills in data modeling and visualization, while also developing a keen eye for storytelling and user-centric design. With these capabilities, I am better equipped to drive business intelligence initiatives and support data-informed decision-making within any organization.

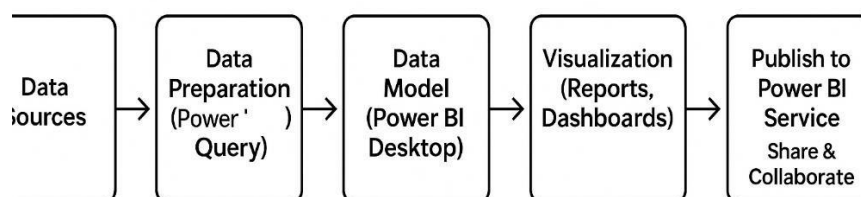


Fig 5.1.Flow process in Power BI

Chapter 6

WEEK 6 & 7

TABLEAU TRAINING

The Tableau training program offered a deep and structured understanding of designing interactive dashboards and data visualizations to extract actionable business insights. The sessions gradually advanced from basic data connectivity and cleaning to sophisticated analytical techniques, interactivity, and deployment strategies, ensuring a well-rounded learning experience.

Key Areas of Learning

1. Data Connection and Preparation

We began by learning to connect Tableau to various data sources, including:

- **Excel, SQL Server, Google Sheets**, and web-based services.
- Leveraging tools like **Data Interpreter** for automatic cleaning and restructuring of raw data.
- Creating **calculated fields** to derive custom metrics and handle missing or incorrect values.
- Validating **metadata and data types** before visualization to ensure consistency and reliability.

Effective data preparation proved essential for building accurate and trustworthy dashboards.

2. Data Blending and Joins

For scenarios involving multiple data sources, we explored:

- **Data Blending** to integrate datasets across platforms using common fields.
- Various **join types** (Inner, Left, Right, Full Outer) for merging tables from the same source.
- Addressing challenges such as **field mismatches** and **aggregation discrepancies** during blending.

These techniques allowed us to create unified datasets tailored to specific analytical needs.

3. Creating Visualizations

We practiced building a wide range of visualizations, such as:

- **Bar charts, line graphs, scatter plots, tree maps, heat maps, bullet charts, and area charts.**
- Choosing the right visual based on the **data type** (categorical, time-series, hierarchical) and intended insights.
- Enhancing visuals with **color gradients, labels, size controls, and interactive tooltips.**

This reinforced the importance of simplicity, clarity, and purposeful design in storytelling.

4. Calculations and Aggregations

To expand Tableau's analytical depth, we learned:

- **Calculated Fields** for row-level logic and KPI creation.
- Use of **aggregate functions** and **conditional logic** (e.g., IF, CASE, IIF).
- Applying **Table Calculations** (e.g., Running Total, Percent of Total, Rank).
- Introduction to **Level of Detail (LOD) Expressions** (FIXED, INCLUDE, EXCLUDE) to control aggregation granularity.

These concepts enabled more dynamic, flexible, and customized analytics.

5. Interactivity and Filters

To improve usability and engagement, we integrated interactive elements:

- **Filters** (dimension, measure, context) for data subsetting.
- **Parameters** for dynamic user input and flexible control over calculations or visuals.
- **Dashboard Actions** (Filter, Highlight, URL) for contextual interactivity.
- Advanced interactions using **Set Actions** and **Parameter Actions** for responsive dashboards.

These features empowered users to explore insights independently based on their needs.

6. Dashboard Building

We developed skills in designing dashboards that are both intuitive and performance-optimized:

-
- Organizing content with **grids**, **containers**, and logical flow.
 - Applying **consistent color themes**, **fonts**, and branding elements.
 - Emphasizing **data storytelling** to guide user navigation through key insights.
 - Creating **responsive designs** suitable for desktops, tablets, and mobile devices.

Performance optimization, including extract tuning and reducing complex calculations, was emphasized to ensure smooth user experiences.

7. Publishing, Sharing, and Collaboration

In the final phase, we learned about secure deployment and teamwork:

- Publishing to **Tableau Server** or **Tableau Online** environments.
- Scheduling **automatic data refreshes** for real-time updates.
- Configuring **user permissions** and **Row-Level Security (RLS)** to protect sensitive data.
- Using **project folders**, **dashboard subscriptions**, and **comments** for collaborative work.

These practices ensured that dashboards were not only insightful but also enterprise-ready and secure.

Assignments and Assessments

Each module included hands-on assignments based on real-world use cases like **sales trends**, **customer segmentation**, and **financial dashboards**. Assessments tested our ability to apply both theoretical and practical knowledge effectively.

We maintained detailed documentation for each task, covering:

- Objectives
- Data preparation steps
- Design decisions
- Challenges and resolutions
- Final outputs

This disciplined approach reinforced technical skills, analytical thinking, and solution-oriented design.

Conclusion

The Tableau training equipped me with strong competencies in visual analytics, interactive dashboard creation, and business storytelling. I am now confident in transforming complex data into clear, compelling visuals that drive informed decision-making. These skills position me to support data-driven initiatives and contribute meaningfully to analytical teams across industries.

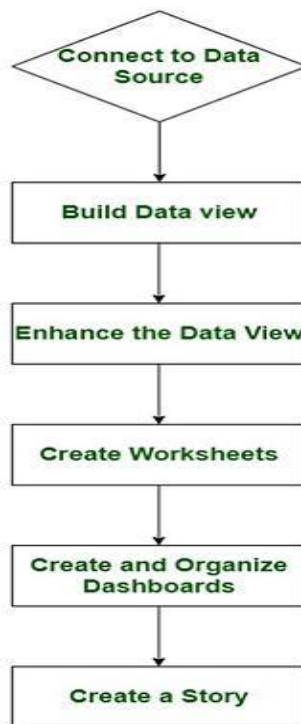


Fig.6.1.Tableau flow process

Chapter-7

Week 8-11

Dashboard Replication in Tableau and Power BI

Objective:

For this practical assessment, I was tasked with replicating four dashboards—Sales, Prescription, Promo Spend, and Patient—across both Tableau and Power BI. These dashboards incorporated key performance indicators (KPIs) such as Rx, Tx, and Market Share, along with a variety of charts and parameter controls to enhance data interactivity and filtering. The primary goal was to achieve a high level of consistency in design, functionality, and user experience across both platforms, testing my ability to interpret dashboard requirements and adapt my skills accordingly.

Key Challenges and Solutions:

1. Divergent Methods and Workflows:

- **Approach:** I adapted to the distinct data preparation and analysis methods in each tool. In Tableau, I utilized its intuitive drag-and-drop interface and calculated fields to quickly structure data. For Power BI, I relied on Power Query for data transformations and DAX for custom calculations, ensuring the underlying data models remained consistent.

2. Visual Consistency and Layout Replication:

- **Approach:** I focused on replicating the visual presentation of KPIs like Rx, Tx, and Market Share across both platforms. In Tableau, this required precise control over formatting and layout, while in Power BI, I leveraged custom visualizations and formatting options to achieve a similar design.

3. Interactivity and User Controls:

- **Approach:** I developed dynamic filters and parameter-driven controls for interactive data exploration. In Tableau, I implemented parameters and dashboard actions, while in Power BI, I used slicers and dynamic DAX measures to mirror the same interactivity, providing a seamless user experience.

4. Complex Calculations and Filtering Logic:

- **Approach:** I translated calculated fields from Tableau into equivalent DAX expressions in Power BI, ensuring that growth metrics and performance indicators remained consistent across both platforms, despite their differing calculation engines.

5. Parameter-Driven Interactivity:

- **Approach:** I implemented parameter controls to switch between different metrics like Rx and Tx. In Tableau, this involved linking parameters to filters and visual elements, while in Power BI, I used What-If Parameters and slicers to replicate this functionality, aligning the experience as closely as possible.

Outcome:

Despite the technical challenges, I successfully replicated all four dashboards across both platforms. This project significantly enhanced my problem-solving skills, deepened my understanding of both Tableau and Power BI, and improved my ability to translate visual and analytical concepts across different tools. I prioritized consistency in functionality, design, and interactivity, ultimately achieving a seamless user experience across both platforms.

Conclusion:

This practical assessment provided valuable insights into the nuances of dashboard design and data visualization, reinforcing my proficiency in both Tableau and Power BI. It strengthened my ability to create dynamic KPIs, parameter-driven interactions, and visually consistent dashboards, preparing me for more advanced data visualization tasks in the future.



Fig 7.1.Sample Dashboard

Chapter 8

Weeks 12–16

Live Project: Case Studies, Knowledge Transfer, and Ad-Hoc Reporting

During Weeks 14 to 15, I was deeply involved in the initial stages of a new live project where I focused on case studies, participated in knowledge transfer sessions, and handled ad-hoc report generation. This phase marked a critical point in my internship, where I transitioned from learning tools and techniques to applying them in real-world scenarios.

1. Case Study Analysis and Learning

In the beginning, I was assigned a series of case studies related to the pharmaceutical industry. These case studies were designed to deepen my understanding of industry-specific data analysis practices and the types of insights required by stakeholders.

- **Analyzing Pharmaceutical Data:** I worked with datasets that included patient demographics, treatment outcomes, and prescription patterns. I had to interpret this data to generate insights on how drugs were performing across various markets and how patient adherence affected treatment outcomes.
- **Scenario-Based Insights:** The case studies also required me to simulate different market conditions, such as the launch of a new drug or a change in patient behavior, to see how these factors could affect the performance of pharmaceutical products.

This phase helped me strengthen my analytical mindset by applying theoretical concepts to practical, real-world data scenarios.

2. Knowledge Transfer Sessions

As part of the internship, I attended multiple knowledge transfer (KT) sessions with senior analysts and project managers. These sessions were aimed at providing me with an in-depth understanding of the project's objectives and the tools used to achieve them.

-
- **Understanding Data Sources:** The KT sessions gave me a comprehensive overview of the different data sources involved in the project. I learned how clinical trial data, patient management systems, and pharmacy records contributed to the overall analysis.
 - **Tool Training:** I received hands-on training in the tools used for the project, such as SQL, Power BI, and Tableau, focusing on how to extract, process, and visualize the data.
 - **Project Workflow:** A key takeaway was learning about the workflow for data extraction, including how the data was retrieved from various systems and integrated into our analysis pipelines.

3. Ad-Hoc Report Generation and Data Refreshes

I was tasked with generating **ad-hoc reports** that provided key insights into the ongoing project. These reports were generated as required by the stakeholders and helped provide actionable insights into patient treatment patterns, market trends, and drug efficacy.

- **Report Generation:** Using Power BI and Excel, I created custom reports based on stakeholder requests. These reports ranged from simple descriptive analysis to more complex predictive analytics.
- **Data Refresh and Automation:** I also worked on refreshing the data for these reports, ensuring that the most recent data was included. This involved refreshing datasets, checking for data integrity, and ensuring that the reports reflected the latest patient data or clinical trial results.

By the end of Week 15, I had developed a solid understanding of how to generate actionable reports and ensure that data remained up-to-date for decision-making.

Chapter 9

WEEK 17 & 18

DASHBOARD MAINTENANCE, DOCUMENTATION, AND WORKBOOK CREATION

During weeks 17 and 18 of my internship, my role shifted to dashboard maintenance, documentation, and workbook creation. This transition allowed me to refine my understanding of dashboard operations while enhancing the overall user experience. I focused on ensuring that the dashboard was fully functional, well-documented, and easy to use by preparing supporting workbooks for future updates. Below is a summary of the activities I was involved in during this period:

1. Dashboard Maintenance

As part of my responsibilities, I maintained the official dashboard, ensuring that it was up-to-date and fully functional. My focus was on ensuring that the dashboard remained accurate, user-friendly, and optimized for use across various platforms.

- **Formatting and Layout Consistency:** I continued to format the dashboard using Virtual Desktop Infrastructure (VDI). Ensuring consistency across the dashboard was essential, so I spent time adjusting visual elements to align with the design guidelines of the organization. This included ensuring that all KPIs, charts, and visual components were clear and aligned.
- **Data Accuracy:** I worked closely with the team to ensure the accuracy of the data presented on the dashboard. While I was not responsible for refreshing the data, I reviewed it to ensure that all updates reflected in the dashboard were accurate and that no discrepancies occurred during data refresh cycles.

2. Dashboard Documentation

One of my key tasks during these weeks was to create comprehensive documentation for the dashboard. This was crucial for ensuring that the dashboard could be maintained and updated easily in the future, especially as the team grew or as new analysts took over. The documentation included:

- **Dashboard Overview:** I created a detailed summary of the dashboard's purpose, highlighting its key functionalities, the KPIs displayed, and the value it provided to end-users. This helped to provide a clear understanding of the dashboard's objectives for both technical and non-technical stakeholders.

-
- **Field and Metric Definitions:** I documented the various fields, metrics, and calculations used in the dashboard. This included providing clear definitions for each KPI, as well as explaining how they were derived (e.g., calculated fields or DAX formulas used in Power BI).
 - **Parameters and Filters:** I outlined the various parameters and filters used in the dashboard, explaining how users could interact with the dashboard and modify the view based on their requirements. This included a step-by-step guide on how to use filters to drill down into data and how to switch between different metrics or views.
 - **Version Control:** I included version control details in the documentation to track any changes made to the dashboard over time. This was important for keeping track of updates and ensuring that any modifications to the dashboard were logged properly.

3. Workbook Creation

In addition to maintaining the dashboard, I was tasked with creating supporting workbooks to help streamline future updates and analyses. These workbooks were essential for tracking changes, performing ad-hoc analysis, and providing additional context for the dashboard.

- **Workbook Design:** I created clean, organized workbooks that would serve as sources for the dashboard data. These workbooks included tables, pivot tables, and charts designed to assist with future updates, calculations, and analysis.
- **Data Analysis:** I performed basic data analysis within the workbooks to identify trends, outliers, and patterns that could inform future dashboard updates. I used these workbooks to track any data discrepancies or performance issues that arose during the data refresh process.
- **Ad-Hoc Reports:** Occasionally, I created ad-hoc reports within these workbooks to support the team's needs. These reports included quick summaries, detailed tables, or chart-based reports that could be referenced in discussions with stakeholders or in the planning of future dashboard updates.

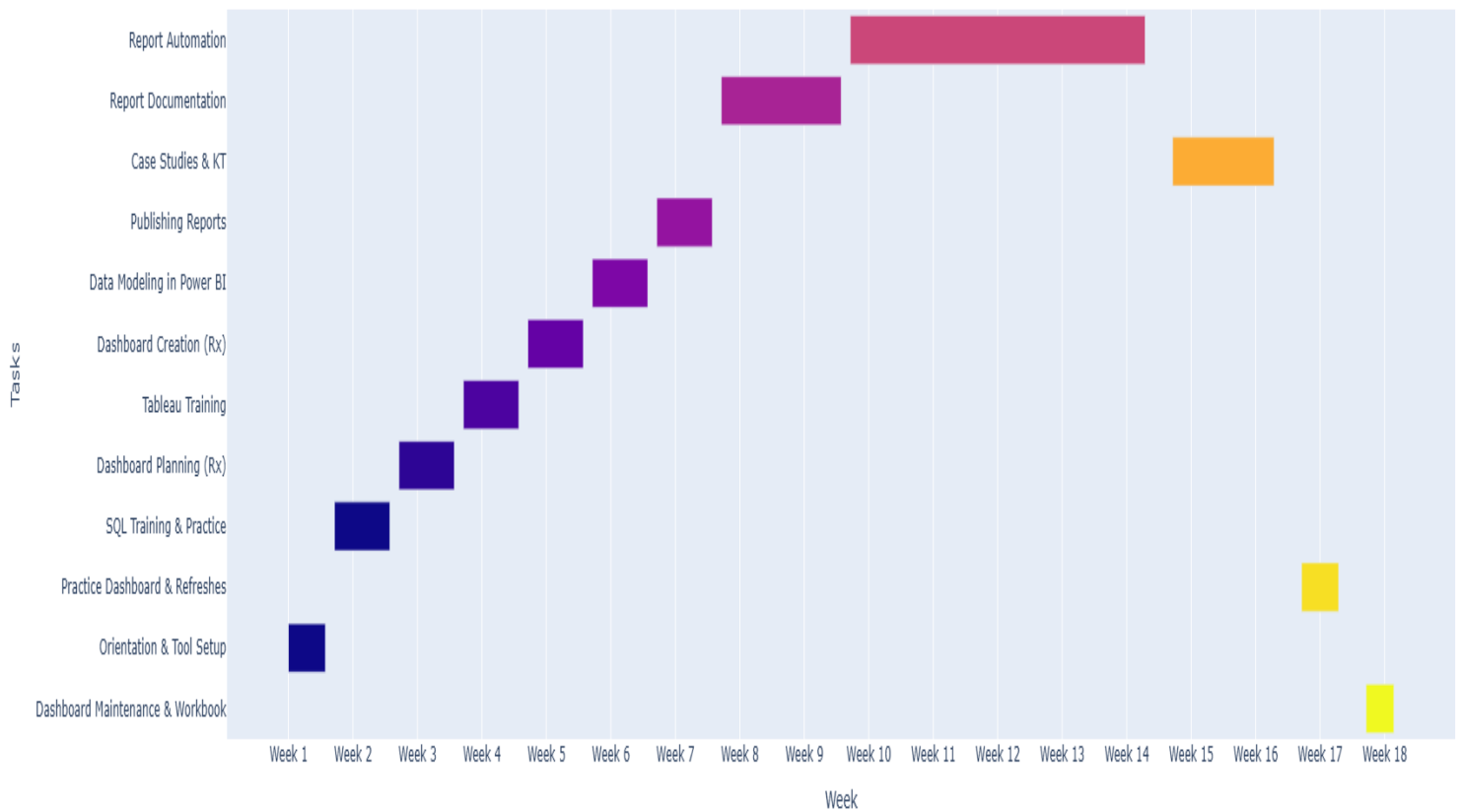
Conclusion

Weeks 17 and 18 allowed me to take a more hands-on approach to dashboard maintenance and documentation. I ensured that the dashboard remained functional, well-documented, and ready for future updates. The tasks I performed—ranging from formatting and maintaining consistency in the dashboard to creating detailed documentation and supporting workbooks—were critical for ensuring the continued success and usability of the dashboard. This experience strengthened my understanding of the dashboard lifecycle and the importance of clear documentation and preparation for future modifications.

Chapter 10

TIMELINE FOR EXECUTION (GANTT CHART)

Internship Timeline (Weeks 1-18)



Chapter 11

OUTCOMES

I am currently in the process of preparing to undertake the Power BI and Tableau certifications, which are pivotal milestones in my professional development. These certifications will serve to further solidify my expertise in data visualization and business intelligence (BI) tools, both of which are integral for translating complex and large datasets into meaningful, actionable insights that can drive business decisions.

Power BI and Tableau are two of the most widely used BI tools in the industry, known for their powerful features and flexibility in creating dynamic dashboards, reports, and visualizations. The certification exams will test my skills in various critical areas, such as designing and implementing interactive dashboards, conducting advanced data analysis, and leveraging both tools to present insights in an engaging and intuitive manner.

Chapter 12

CONCLUSION

Overall, my hands-on experience with dashboards has played a vital role in sharpening both my technical and analytical abilities. I gained practical knowledge in areas such as data refresh mechanisms, SQL querying, error resolution, and validation techniques—all of which are essential when working with large and complex healthcare datasets.

My exposure to SQL and the structure of clinical databases helped me understand the critical importance of data accuracy in medical analytics. I also became proficient in configuring parameters and filters to enable interactive, user-driven insights, and I applied this knowledge effectively across both Tableau and Power BI platforms. In addition, assisting with the dashboard refresh cycles and troubleshooting data discrepancies has strengthened my attention to detail and problem-solving skills.

Looking ahead, I am preparing for certification exams in Power BI and Tableau to further deepen my expertise in business intelligence and data visualization. These credentials will empower me to create impactful dashboards, perform advanced data analysis, and deliver insights that support informed decision-making. The skills I've gained through this journey will enable me to contribute meaningfully to the development, maintenance, and continuous improvement of robust BI solutions.



DataZymes Analytics Private Limited
(Subsidiary of DataZymes Inc.)

CIN: U74999KA2016PTC094712

December 19th, 2024

INTERNSHIP LETTER

Dear Manya Jagadish,

We are pleased to offer you an internship with our organization, for a period of 6 months from January 06th, 2025. Below are the terms of the internship:

- **Position held:** Associate
- **Stipend:** INR 20,000 (Twenty Thousand only) per month, subject to statutory deductions as per state & Income Laws (Responsibility of adherence with statutory provisions of tax lies with you)
- **Working hours:** As per company policies
- **Office Location:** Bengaluru

Terms & conditions:

- On successful completion of the internship, the company, if it wishes, can renew the internship, or offer you permanent employment, based on your performance or the work requirements.
- As an intern you will receive "temporary employment" status, and hence you will not be subject to any of the employee benefits that regular company employees receive, including, but not limited to, health insurance, gratuity, conveyance etc.
- In addition, you agree that, upon conclusion of your internship, you will immediately return to the company all its property, equipment, and documents, including electronically stored information.
- You will maintain utmost confidentiality of the proprietary information of the organization and client data/ information and will not disclose any such information to an outsider of the company/ third party without seeking the written permission of the company. You will also not get into malicious activities of sharing/ transferring /deleting any company/client data.
- The internship may be terminated at any time prior to the Cessation Date, at the Company's sole discretion. The termination of your internship without any notice period may be on grounds of: (i). Under performance (ii). Breach of confidentiality or IP related obligations; (iii). Violation of law; (iv). Gross Misconduct; (v). Material breach of Company policy.
- You are also required to adhere to and abide by all other policies, terms, and conditions as set by the company from time to time.

I hope that your association with the company will be successful and rewarding. Please indicate your acceptance of this offer

Regards,
For **DataZymes Analytics Private Limited.**


Anusuya Borkotoky
Associate Director- HR


12 / 26 / 2024
Manya Jagadish

Registered Office:
25, Primrose Rd, Craig Park Layout, Sivanchetti Gardens, Bengaluru, Karnataka 560025, Ph: 7406644226

Doc ID: 34a58ea1c9d6f0cd320e699a4e2dc2c165eee748



DataZymes Analytics Private Limited
(Subsidiary of DataZymes Inc.)

CIN: U74999KA2016PTC094712

Internship Confirmation Letter

TO WHOMSOEVER IT MAY CONCERN

This is to certify that Manya Jagadish is pursuing her internship with us at DataZymes Analytics Private Limited effective 8th January 2025.

This letter has been provided for the purpose of a confirmation to her college only.

For DataZymes Analytics Private Limited

A handwritten signature in black ink, appearing to read 'Anusuya Borkotoky'.

Anusuya Borkotoky
Associate Director– HR

APPENDIX-A

Compatibility with Sustainable Development Goals



1. SDG 3: Good Health and Well-being

a. Data-Driven Healthcare Improvements:

The internship focused on analyzing pharmaceutical data to support better treatment outcomes, prescription trends, and patient behavior analytics. This contributes to improving access to high-quality healthcare and promoting data-informed patient services.

b. Support for Public Health Monitoring:

The project involved dashboards and reporting tools that helped stakeholders monitor drug performance and patient adherence, supporting efforts to ensure healthy lives and promote well-being for all.

c. Evidence-Based Decisions:

By working with real-world healthcare datasets and using business intelligence tools, the internship contributed to creating systems that support evidence-based clinical and business decisions, thus aligning with health system strengthening.

2. SDG 9: Industry, Innovation and Infrastructure

a. Technological Advancement in Pharma Analytics:

The use of Power BI, Tableau, and Snowflake for scalable data solutions in the pharma sector reflects innovation in digital infrastructure and data-driven innovation in healthcare.

b. Capacity Building in Data Infrastructure:

The internship fostered skills in SQL, data modeling, and reporting, helping build human capital and digital capabilities for pharmaceutical firms—a direct contribution to resilient infrastructure and fostering innovation.

3. SDG 4: Quality Education

a. Industry-Aligned Skill Development:

The training involved hands-on sessions in data warehousing, pharma processes, business intelligence tools, and reporting workflows, contributing to inclusive, equitable, and quality education aligned with job market needs.

b. Practical Learning through Live Projects:

Engagement with real datasets and pharmaceutical dashboards promoted experiential learning and improved career readiness, especially in a specialized sector like healthcare analytics.

4. SDG 8: Decent Work and Economic Growth

a. Productivity through Automation:

The internship enhanced business reporting workflows, improved dashboard interactivity, and supported automation in analytics—helping increase productivity for business users.

b. Empowering Knowledge Workers:

By learning and applying automation and visualization techniques, the intern contributed to creating tools that help professionals focus on insights rather than manual work—fostering a smarter workforce.

5. SDG 12: Responsible Consumption and Production

a. Efficiency in Reporting and Data Use:

By streamlining dashboards and reports, the intern contributed to reducing reporting inefficiencies and promoted responsible data practices and efficient use of technological resources.

Joseph Michael Jerard_DataZymes_Internship_Report_final

ORIGINALITY REPORT

10%

SIMILARITY INDEX

8%

INTERNET SOURCES

6%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES

1

Matheen Fathima G, Shakkeera L. "Efficient task scheduling and computational offloading optimization with federated learning and blockchain in mobile cloud computing", Results in Control and Optimization, 2025

Publication

4%

2

Submitted to Presidency University

Student Paper

3%

3

www.swartzentrover.com

Internet Source

<1%

4

Submitted to Intercollege

Student Paper

<1%

5

www.jetir.org

Internet Source

<1%

6

rathinamcollege.ac.in

Internet Source

<1%

7

Submitted to eyglobal

Student Paper

<1%

8

www.chemanager-online.com

Internet Source

<1 %

9

www.globalmatrixsurvey.com

Internet Source

<1 %

10

careers.danone.com

Internet Source

<1 %

11

masudprogrammer.wordpress.com

Internet Source

<1 %

12

Haseeb Khan, Naiwrita Borah, Shaik Salma Begum, Afroj Alam, Mohammed Soudy.
"Transformer networks and autoencoders in genomics and genetic data interpretation: A case study", Elsevier BV, 2025

Publication

<1 %

13

commons.wikimedia.org

Internet Source

<1 %

14

coursetakers.com

Internet Source

<1 %

15

medium.com

Internet Source

<1 %

16

presidencyuniversity.in

Internet Source

<1 %

Exclude quotes Off

Exclude bibliography On

Exclude matches Off