Practical Summer Internship Report

On

Decoding Data

Submitted in Partial Fulfillment for the Requirements for the Award of the

Degree of

BACHELOR OF TECHNOLOGY

In

CSE –DATA SCIENCE

By
Birender Kumar
(Roll No: 23DS005)

Department of DATA SCIENCE

St. Andrews Institute of Technology & ManagementGurugram–122506

Affiliated to

MAHARISHI DAYANAND UNIVERSITY, ROHTAK

INTERNSHIP CERTIFICATION

CERTIFICATE

This is to certify that the Internship Report submitted by **Birender Kumar (Roll No. 23DS005)**

is a bona fide record of work done by him and submitted during the 2025-2026 academic year, in partial fulfilment of the requirements for the award of the degree of BACHELOR OF TECHNOLOGY in COMPUTER SCIENCE & ENGINEERING (Emerging Technologies)

Internship Co-Ordinator
Dept. Internship Co-Ordinator

Head of the Department
Associate Professor, AIML/DS
SAITM, Gurugram, India

CANDIDATE'S DECLARATION

I, **Birender Kumar (Roll No. 23DS005)**, hereby declare that the winter internship report titled **"Decoding Data: Insights and Impact through Analytics"** is a bona fide work carried out by me under the guidance of the project team at **IBM Skills Build Project Based Learning Program in collaboration with CSRBOX**,.

Further, I declare that this report has not previously formed the basis for the award of any degree or diploma and has not been submitted elsewhere.

Date:17/11/2025
Place: Gurugram

Candidate's Signature
Name: Birender Kumar
Roll no: 23DS005

ACKNOWLEDGEMENTS

I am highly indebted to Director Prof. Rakesh Rajpal , for the facilities provided to accomplish this internship. I would like to thank our Head of the Department Dr. Sandeep Sharma and Program Leader Ms.Preetishree for her constructive criticism throughout my internship. I would like to thank Dr. Dinesh Yadav, College internship coordinator and Dr. Shweta , Department internship coordinator for their support and advices to get a complete internship in above said organization. I am extremely grateful to my department staff members and friends.

Date: November 2025 **Name: Birender**

ABSTRACT

The internship titled **"Decoding Data: Insights and Impact through Analytics"**, conducted under the **IBM Skills Build Project Based Learning Program in collaboration with CSRBOX**, provided an in-depth understanding of data analysis, visualization, and decision-making using modern analytical tools. The focus of the internship was to bridge theoretical concepts with practical industry-oriented skills in data analytics and machine learning.

Throughout the internship, I gained hands-on experience in **data cleaning, preprocessing, exploratory data analysis (EDA), dashboard creation, and interpretation of analytical insights**. Tools such as **Python, Pandas, NumPy, Matplotlib, Seaborn,Power BI,Sql and IBM Skills Build analytics modules** were extensively used to understand the impact of data in real-world decision-making.

The internship emphasized understanding **problem statements, extracting insights, designing data workflows, building analytical solutions, and evaluating results**. The project contributed significantly to enhancing my technical proficiency, analytical thinking, and problem-solving ability.

The report summarizes the purpose of the internship, organizational structure, learning outcomes, analytical tasks, findings, and recommendations for future improvement.

TABLE OF CONTENTS

1. **Title Page** 1
2. **Internship Certificate (Received from Organization** 2
3. **Candidate's Declaration** 3
4. **Acknowledgements** 4
5. **Abstract** 5
6. **Table of Contents** 6
7. **List of Figures** 7
8. **List of Tables** 8
9. **List of Abbreviations** 9
10. **Introduction** 10
11. **Organization Profile** 11
12. **Objectives of the Internship** 12
13. **Description of Work / Project Details** 13
14. **Tools and Technologies Used** 14
15. **Roles, Responsibilities, and Learnings** 15
16. **Outcomes and Skills Gained** 16
17. **Conclusion and Future Scope** 17
18. **Annexure / Supporting Documents** 18

LIST OF FIGURES

1. **Figure 1.1 – Organization Hierarchy Chart** 7
2. **Figure 2.1 – Department Workflow Diagram** 8
3. **Figure 3.1 – Project Architecture / System Design** 9
4. **Figure 4.1 – Data Flow Diagram (DFD)** 10
5. **Figure 5.1 – User Interface / Dashboard Screenshot** 11
6. **Figure 6.1 – Database Schema or ER Diagram** 12
7. **Figure 7.1 – Output / Result Graph** 13
8. **Figure 8.1 – Summary Chart / Analysis Visualization** 14

LIST OF TABLES

1. **Table 1.1 – Internship Schedule / Work Plan** 15
2. **Table 2.1 – Departmental Structure / Team Allocation** 16
3. **Table 3.1 – Project Modules and Descriptions** 17
4. **Table 4.1 – Tools and Technologies Used** 18
5. **Table 5.1 – Dataset / Input Details (if applicable)** 19
6. **Table 6.1 – Task Summary and Daily Work Log** 20
7. **Table 7.1 – Results / Performance Analysis** 21
8. **Table 8.1 – Comparison or Evaluation Table** 22

LIST OF ABBREVIATIONS

1. **AI** – Artificial Intelligence 23
2. **ML** – Machine Learning 24
3. **CNN** – Convolutional Neural Network 25
4. **DFD** – Data Flow Diagram 26
5. **ER** – Entity Relationship 27
6. **GUI** – Graphical User Interface 28
7. **API** – Application Programming Interface 29
8. **DBMS** – Database Management System 30
9. **IoT** – Internet of Things 31
10. **SQL** – Structured Query Language 32
11. **HTML** – HyperText Markup Language 33
12. **CSS** – Cascading Style Sheets 34
13. **IDE** – Integrated Development Environment 35
14. **KPI** – Key Performance Indicator 36
15. **UI/UX** – User Interface / User Experience 37

CHAPTER 1: INTRODUCTION OF THE REPORT

1.1 Background

Data plays a key role in decision-making across all industries today. Organizations use data analytics to understand trends, improve performance, and make informed choices. To help students gain practical skills, the **IBM SkillsBuild Project-Based Learning Program**, in collaboration with **CSRBOX**, offered the internship **"Decoding Data: Insights and Impact Through Analytics."**

This report summarizes the internship work completed by **Birender Kumar** during **July–August 2025**, where I learned the basics of data collection, preprocessing, analysis, and visualization.

1.2 Purpose of the Report

The purpose of this report is to present the tasks, tools, and learning outcomes of the internship. It includes:

- Activities performed during the program
- Tools used like Python and visualization methods
- Insights gained from data analysis
- Skills developed through project-based learning

1.3 Objectives of the Internship

The main objectives were:

1. To understand real-world use of data analytics.
2. To learn data cleaning and preprocessing.
3. To perform exploratory analysis using Python.
4. To generate insights and visualizations.
5. To improve analytical and problem-solving skills.

1.4 Scope of the Report

The report focuses on the analytics tasks completed in the 8-week internship, including:

- Working with datasets
- Cleaning and analyzing data
- Creating visual charts
- Summarizing insights

The internship provided practical exposure to essential data analytics concepts.

CHAPTER 2: OVERVIEW OF THE ORGANIZATION

The internship was conducted under the **IBM SkillsBuild Project-Based Learning Program**, in collaboration with **CSRBOX**, during July–August 2025. This chapter gives a brief overview of the organizations involved and their role in the internship.

2.1 Introduction to IBM

IBM (International Business Machines Corporation) is a global technology company known for its work in **cloud computing, AI, data analytics, cybersecurity, and enterprise solutions**. With over a century of innovation, IBM focuses heavily on research and skill development. IBM also offers learning platforms like **IBM SkillsBuild** to support students and professionals with industry-ready skills.

2.2 IBM SkillsBuild

IBM SkillsBuild is a free learning platform that provides training in areas such as:

- Data Analytics
- AI & Machine Learning
- Cloud Computing
- Cybersecurity
- Professional Skills

The platform includes industry-aligned courses, hands-on labs, project work, and digital badges. It prepares learners for modern digital careers through structured modules and practical assignments.

2.3 About CSRBOX

CSRBOX is India's leading CSR implementation and training organization. It partners with companies and educational institutions to deliver large-scale skill development programs. As IBM's partner, CSRBOX handled:

- Training sessions
- Guidance and doubt support
- Project evaluation
- Internship coordination

2.4 IBM SkillsBuild + CSRBOX Collaboration

Together, IBM SkillsBuild and CSRBOX provided:

- Structured online learning
- Project-based tasks
- Real datasets

- Mentor support
- Certification

This collaboration ensured that students received both technical knowledge and practical exposure.

2.5 Internship Program Overview

The internship titled **"Decoding Data: Insights and Impact through Analytics"** focused on data analytics fundamentals. The key components included:

- Data collection & preprocessing
- Exploratory Data Analysis (EDA)
- Visualization and dashboards
- Insight generation
- Final project submission

2.6 Summary

IBM, IBM SkillsBuild, and CSRBOX collectively created a practical and structured learning experience. Their collaboration helped students develop essential data analytics skills and gain exposure to industry-standard tools and workflows.

CHAPTER 3: Internship Experience and Technical Learning

The IBM SkillsBuild internship (July–August 2025) provided me with practical exposure to data analytics through live sessions, self-paced modules, and project-based tasks. The program helped me understand how real datasets are cleaned, analyzed, and converted into meaningful insights.

3.1 Duration

The internship lasted **8 weeks**, during which I completed training in Python, data preprocessing, EDA, visualization, and report preparation.

3.2 Work Areas

My work mainly involved:

- Data cleaning and preprocessing
- Exploratory Data Analysis
- Visualization and dashboard creation
- Insight generation and documentation

3.3 Key Learnings

I learned:

- Python for data handling (Pandas, NumPy, Matplotlib)
- Techniques for cleaning, filtering, and preparing data
- Creating charts, dashboards, and KPI visuals
- Identifying trends, patterns, and business insights

3.4 Weekly Summary

The internship progressed from basics to advanced tasks:

- Week 1–2: Fundamentals of analytics & Python
- Week 3–4: Data cleaning and EDA
- Week 5–6: Visualizations and insight generation
- Week 7–8: Final project compilation and submission

3.5 Projects Completed

Project 1: Netflix Content Analysis
Analyzed global Netflix data to study content trends, top countries, genres, and yearly growth. Built an interactive dashboard using Python and Power BI.

Project 2: Agromech Sales Dashboard
Worked on sales and profit data to identify best-performing products, states, and customer behavior. Created KPIs and visuals using SQL, Power Query, and Power BI.

4. TOOLS AND TECHNOLOGIES USED

During the internship, a set of essential tools and technologies were used for data cleaning, analysis, and visualization.

4.1 Software Tools

- **Power BI:** Used to create dashboards and visualize insights for the project.
- **Power Query:** Helped in cleaning and transforming raw data before analysis.
- **DAX:** Used inside Power BI to create calculated metrics and measures.
- **Python (Pandas):** Used for advanced data cleaning and restructuring that required coding.
- **SQL:** Used to extract data, join tables, and run basic queries before analysis.
- **VS Code:** IDE used to write and run Python scripts.
- **Excel/CSV:** Used for storing and reviewing raw datasets.
- **Git/GitHub:** Used conceptually for version control and organizing project files.

4.2 Hardware Requirements

- A laptop/PC with minimum **8GB RAM** for running Power BI and Python smoothly.
- Stable **internet connection** for virtual sessions, dataset downloads, and submissions

- Collected and cleaned data using SQL, Python, and Power Query.
- Built data models and KPIs in Power BI using DAX.
- Created dashboards and generated insights for Netflix & Agromech projects.
- Documented scripts, measures, and project steps.

4.3 Challenges

- Handling messy and unstructured data.
- Writing complex DAX formulas.
- Keeping data consistent across SQL → Python → Power BI.
- Turning raw data into clear business insights.

4.4 Learnings Improved skills in SQL, Python, Power BI, DAX, and ETL.

- Learned data modeling, KPI creation, and dashboard storytelling.
- Gained soft skills: problem-solving, documentation, time management, adaptability

5. Roles, Responsibilities & Learnings

5.1 Roles & Responsibilities

- Extracted data using SQL.
- Cleaned and transformed data using Python and Power Query.
- Built data models and KPIs in Power BI (DAX).
- Designed dashboards for Netflix & Agromech projects.
- Analyzed trends and generated insights.
- Maintained documentation of scripts and metrics.

5.2 Challenges

- Messy data handling in Python (exploding lists).
- Difficulty in writing complex DAX measures.
- Managing smooth workflow from SQL → Python → Power BI.
- Converting charts into meaningful business insights.

5.3 Learnings & Soft Skills

Technical: SQL, Python (Pandas), Power Query, Power BI, DAX, data modeling, visualization.
Soft Skills: Problem-solving, documentation, time management, adaptability.

6. ANALYSIS

6.1 SWOT Analysis

Strengths: Real datasets, hands-on **Opportunities:** Move into ML, cloud databases, advanced analytics.
Threats: Fast-changing BI tools require constant upskilling.

6.2 Strategies

- Use dashboards & Python scripts as portfolio (SO).
- Learn cloud tools (WO).
- Regularly update BI/analytics skills (ST).
- Study ML and predictive analytics after internship (WT).

6.3 Overall Analysis

The internship gave practical end-to-end analytics experience using SQL, Python, and Power BI. It strengthened both technical skills and professional confidence, preparing me for real-world BI and data analytics roles.

7. **RECOMMENDATIONS AND CONCLUSION**

7.1 Recommendations

• Add a small module on predictive analytics using Python.
• Provide practice on cloud-based SQL environments.
• Include short advanced DAX sessions.
• Introduce team-based ETL work with Git.
• Add training on presenting insights in simple business language.

7.2 Conclusion

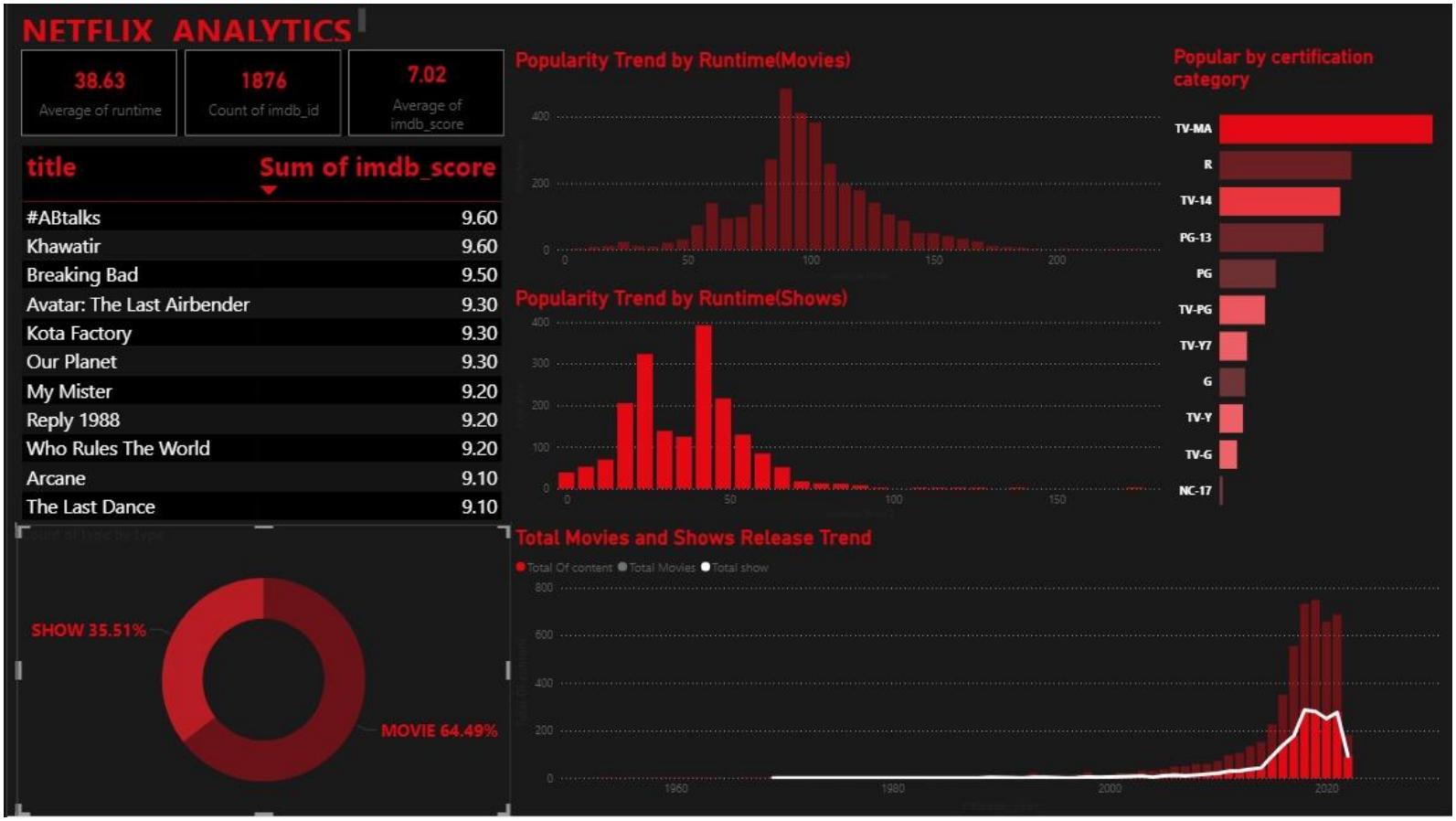
The internship at CODTECH IT SOLUTIONS helped me apply SQL, Python, Power Query, and Power BI to real datasets. Working on Netflix and Agromech projects strengthened my data cleaning, modeling, and dashboard creation skills. Overall, it gave me confidence to handle the full data pipeline and work as a BI/Data Analyst.

Technical Outcomes

• Strong skills in SQL, Python (Pandas), Power Query, Power BI
• Built clean data models and automated ETL processes
• Created professional dashboards with KPIs
• Improved accuracy using DAX measures

Non-Technical Outcomes

• Better problem-solving and analytical thinking
• Improved documentation and reporting
• Strong time management
• Quick adaptability to multiple tools



Agriculture production management dashboard with forecast

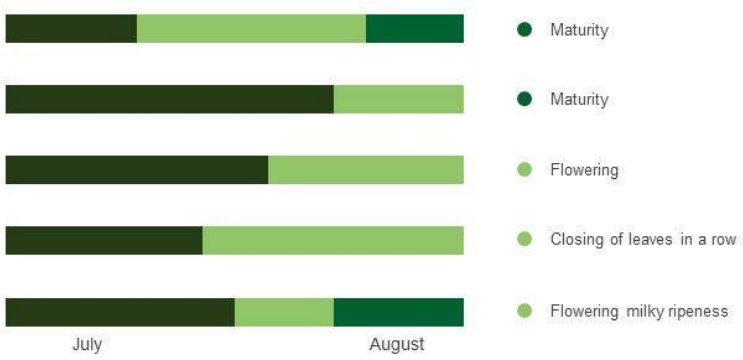
The slide showcases an agriculture management dashboard that helps monitor your field production state over the year. It covers aspects like total area of farmland, forecast production, main cultures and growth stages of various crops.

Total Area of Farmland

398000 / ha
Area under cultivation
+8%

22800 / ha
Black Land
-2%

Growth Stages



Forecast Production

1891 800 / t

+ 12%
vs . 2023

Main Culture

Winter Wheat	Sunflower	Corn	Sugar Beat	Add text here
121 800 / ha	69 200 / ha	28 500 / ha	13 000 / ha	90 500 / ha

++++++

This graph/chart is linked to excel, and changes automatically based on data. Just left click on it and select "edit data".

Conclusion and Future Scope

Conclusion

The internship under the IBM SkillsBuild Program gave me practical experience in the complete data analytics process—from cleaning data in Python to analyzing patterns and visualizing insights using Power BI. Working on the student performance dataset helped me understand how different academic and non-academic factors impact learning outcomes. Overall, the internship improved my technical skills, analytical thinking, and confidence in solving real-world data problems.

Future Scope

This project can be expanded by building predictive models to identify at-risk students, integrating behavioral or psychological data, creating real-time dashboards for schools, performing causal analysis, and applying advanced ML or deep learning techniques for deeper insights into student performance.