

**SURANA COLLEGE-AUTONOMOUS**

Affiliated to Bangalore University

Recognized under 2(f) and 12 B of UGC, Certified by ISO, Re-accredited by NAAC with A+

No 16, South End Road, Bengaluru-560 004

# **INTERNSHIP REPORT**

**PYTHON AND DATA ANALYTICS INTERNSHIP – EMPLOYEE DATA ANALYSIS**

**Submitted to**

**SURANA COLLEGE-Autonomous**

In partial fulfilment for the curriculum requirement for

# **BACHELOR OF COMPUTER APPLICATIONS**

**By**

**ABHISHEK. V**

**U03ME23S0037**

**Under the Guidance of**

**Prof. ASHA H**

Assistant Professor, Dept of Computer Science SURANACOLLEGE-Autonomous, Bengaluru- 560 004

**2025-2026**

# **DECLARATION**

**ABHISHEK. V AUGUST 2025**

**BCA**

I**, Abhishek V** hereby declare that the internship report, entitled **“PYTHON AND DATA ANALYTICS – EMPLOYEE DATA ANALYSIS”** submitted to **SURANA COLLEGE-Autonomous, i**n partial fulfilment for the curriculum requirement for **BACHELOR OF COMPUTER APPLICATIONS done** by me during 2025-2026 under the guidance of **Prof ASHA.H,** Assistant Professor, Department of Computer Science, SURANA COLLEGE-Autonomous, Bengaluru. It has not formed the basis for the award any Degree/Diploma/Associateship/Fellowship or other similar title to any candidate in any College/University.

**ABHISHEK .V**

# **ACKNOWLEDGEMENT**

I am extremely delighted and grateful to thank every single person who helped me in completing my dissertation successfully. I thank the almighty for bestowing me with his blessings. I’m very much indebted and extend my deep sense of gratitude to **Prof ASHA H,** Assistant Professor, Department of Computer Science, internship guide, for her consistent words of motivation, encouragement and filling me completely with insight regarding the topic throughout the period of my study. She was present whenever I needed her the most during the completion of my internship.

My sincere thanks are to the faculty members of Department of Computer Science, for their encouragement in my academic endeavors.

I thank my parents and Ms. **SHAMBHAVI,** External Guide**,** Dyashin Technosoft pvt ltd for being very liberal towards me and for having complete trust in me during the course of my internship.

**ABHISHEK V**

|  |  |  |
| --- | --- | --- |
| Sl. No. | Content | Page No |
| 1 | **Chapter 1: Introduction** | 01 |
| 1.1 | Industry Overview | 01 |
| 2 | **Chapter 2: Organisation Profile** | 02-06 |
| 2.1 | History | 02-03 |
| 2.2 | About | 04 |
| 2.3 | Vision and Mission | 04 |
| 2.4 | Structure of the Organisation | 05-06 |
| 3 | **Chapter 3: Technology used** | 07-09 |
| 3.1 | Day/Week wise report | 07-08 |
| 3.2 | Application used | 09 |
| 4. | **Chapter 4: Conclusion** | 10-25 |
| 4.1 | Learning | 10 |
| 4.2 | Outcome | 11 |
| 4.3 | Snapshots of Project Work and Analysis | 12-23 |
| 4.4 | Summary of the Project | 24 |
| 4.5 | Recommendations to Improve Employee Performance | 25 |

**TABLE CONTENTS**

**CHAPTER 1: INTRODUCTION**

**1.1 INDUSTRY OVERVIEW**

The data analytics industry is experiencing rapid expansion, driven by the exponential growth of digital data and the increasing demand for data-driven decision-making across sectors. As of 2024, the global market is valued at approximately USD 69.54 billion and is projected to reach USD 302.01 billion by 2030, with a compound annual growth rate (CAGR) of 28.7%. This surge is fuelled by advancements in artificial intelligence (AI), machine learning (ML), and cloud computing, which have made analytics tools more powerful, scalable, and accessible. Major players such as IBM, Microsoft, Oracle, SAP, AWS, and Tableau are leading the charge, offering sophisticated platforms that cater to diverse business needs. Key trends shaping the industry include predictive analytics for forecasting, real-time analytics for instant insights, and data democratization, which empowers non-technical users to interpret data through intuitive interfaces. Industries like healthcare, retail, finance, manufacturing, and human resources are leveraging analytics to improve efficiency, personalize services, and mitigate risks.

Despite its promising trajectory, the data analytics industry faces several challenges. Data privacy and security are critical concerns, especially with stringent regulations like GDPR and CCPA requiring responsible data handling. Organizations must invest in secure infrastructure and ethical practices to maintain trust and compliance. Another major hurdle is the shortage of skilled professionals, as demand for data scientists and analysts continues to outpace supply.

**2.ORGANIZATION PROFILE**

**2.1 HISTORY**

Dyashin Technosoft, while officially incorporated as a private limited company in India on September 12, 2023, has a deeper story rooted in the experience of its founding team. The company describes itself as being "led by a group of industry veterans," suggesting that its leaders brought significant knowledge and expertise to the new venture. The name "Dyashin" itself combines the concepts of "dynamic growth" and "shining excellence", reflecting the company's aspirations.

**The early stages:**

* September 12, 2023: Dyashin Technosoft was officially incorporated in India. Some sources also list September 11, 2023
* Initial focus: The company initially focused on providing custom software development, managed IT services, and other related services.
* Building a foundation: During its early years, Dyashin Technosoft likely concentrated on establishing credibility, building a client base, and developing its expertise in core competencies such as software development and enterprise integration.

**Key milestones and growth:**

* Early recognition: By 2012, Dyashin had secured its first international client, marking a turning point that propelled the company to develop a global delivery model and enhance its project management capabilities.
* Infrastructure expansion: In 2013, the company moved into its first fully owned office in Whitefield, Bangalore, signifying its transition from a startup to a more established organization.
* Service diversification: From 2015 to 2018, Dyashin Technosoft expanded its service offerings to include cloud computing, DevOps, mobile app development, enterprise integration services, and data warehousing.
* Global expansion: The company established offices in the United States and the United Kingdom, along with offshore support hubs in other countries, to strengthen its global presence.
* Brand consolidation: In 2020, the company undertook a brand identity revamp, launching a
* new visual identity, website, and corporate messaging focused on "Accelerating Digital Futures".

**Innovation and sustainability:**

* Dyashin Labs: The company launched Dyashin Labs, its dedicated innovation hub, to focus on emerging technologies like IoT, machine learning, and Robotic Process Automation (RPA).
* Proprietary platforms: Dyashin has invested in and launched its own platforms like Dyashin\_ View (a business intelligence tool), Dyashin Edge (a cloud-native DevOps management platform), and DyasChain (a blockchain integration layer).
* Focus on CSR: Dyashin is committed to green IT practices, carbon footprint reduction, and corporate social responsibility (CSR) through its "Dyashin Cares" initiative.

**Present-day operations:**

* Global presence: Dyashin Technosoft operates globally with offices in the USA, India, UK, and UAE.
* Structured operations: The company's operations are organized into three key business units: the Digital Solutions Group (DSG), Enterprise Services Group (ESG), and Innovation & Products Group (IPG).
* Industry verticals: They serve diverse industries such as FinTech, healthcare, retail, logistics, and education.

In essence, Dyashin Technosoft's history is one of continuous growth, diversification, and a strong commitment to innovation and customer success. The company's focus on industry expertise, global expansion, and strategic partnerships has allowed it to establish itself as a recognized player in the software development and engineering services industry.

**2.2 ABOUT THE DYASHIN TECHNOSOFT**

Dyashin Technosoft is a renowned [Software Development](https://dyashin.com/solutions/software-development) and [Engineering services](https://dyashin.com/services) company led by a group of industry veterans. With our passion, knowledge, and commitment to excellence, we believe in transforming the arena of software development and empowering customer businesses

to embrace the future and achieve amazing success. Our comprehensive range of custom software services & Solutions are designed to assist our customers at all stages of the SDLC, assuring project success from start to end and beyond.

**2.3 VISION AND MISSION**

**Vision**

To be a global standard company to provide Value Driven IT services, Innovative Engineering Solutions & Reliable NexGen products for a better tomorrow.

**Mission**

Fostering long lasting relationships and driving mutual growth with our customers, employees & vendor partners in the ever-evolving landscape of technology**.**

**2.4 STRUCTURE OF ORGANISATION**

Dyashin Technosoft is a privately held technology company in India, specializing in information technology and computer service activities. Founded and led by CEO Madan Shamachar, the organization operates with a well-defined hierarchical management structure that ensures smooth coordination across all functions. The leadership team includes key figures such as Swathi Rao, who serves as the Account Manager, overseeing client relationships and project delivery, and Sushma Malali, the Senior Business Analyst responsible for bridging business needs with technical solutions.

The company is organized into specialized departments—Sales, Operations, and Product & Engineering—each with its own leadership and dedicated teams. This clear departmental division allows for focused execution of responsibilities, from generating business opportunities to delivering high-quality technology solutions. The Product & Engineering wing is guided by Chief Technology Officer Praveen Dyamappa, who spearheads software development, product innovation, and engineering excellence.

Dyashin Technosoft primarily delivers software development and engineering services, positioning itself as a solutions-oriented partner for clients in various industries. Its expertise spans building robust software products, optimizing business processes, and offering tailored IT services.

From a legal standpoint, the company is registered as a Private Limited Company under Indian corporate law, which offers operational flexibility while ensuring compliance with regulations. Interestingly, Dyashin Technosoft is an unfunded enterprise, meaning it grows through internally generated resources rather than external investments. This approach reflects a self-sustaining business model, emphasizing steady, organic growth and operational independence, which in turn strengthens its long-term stability and credibility in the competitive IT sector.

**Services provided by Dyashin technosoft company are:**

**1. Software Development:**

• Dyashin offers custom software development services tailored to various business needs.

• They provide support throughout the entire Software Development Life Cycle (SDLC) or Product Development Life Cycle (PDLC).

**2. Software QA & Testing:**

• Dyashin provides quality assurance and testing services to ensure software meets high standards.

• This includes various testing methods like manual, automation, performance, security, and API testing.

**3. Engineering Services:**

• Dyashin offers engineering services, including those related to semiconductor and embedded systems.

**4. Consulting Services:**

• Permanent Hire: They assist with permanent placements, finding the right talent for specific roles.

• Contract Staffing: They provide dedicated account managers for sourcing talent for contract positions.

• Contract-to-Hire Staffing: They offer a pathway for clients to transition contract employees to permanent roles.

• Build, Operate & Transfer (BOT): Dyashin helps clients build, operate, and then transfer teams to the client's payroll.

• Just in Time Hiring: They focus on providing resources as needed, ensuring timely support.

**5. Learning Services:**

• Corporate Learning Programs: They offer skill development initiatives for corporate clients.

• Campus Programs: They provide comprehensive tech-enabled learning experiences for students.

Off-Campus Programs: They cater to the retail segment with technology-focused training.

**CHAPTER 3: TECHNOLOGY USED**

**3.1 WEEKLY REPORT**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Date** | **Task Description** | **Time spent** | **Task Category** | **Status** |
| 01-07-25 | Introduction, History and  Architecture of python. | 1 hour 30 minutes | Theory class | Completed |
| 02-07-25 | Creating project in  PyCharm,  In-built functions and  Comments. | 1 hour 30 minutes | Practical + Theory class | Completed |
| 03-07-25 | Variables, Data Type and Type casting. | 1 hour 30 minutes | Practical + Theory class | Completed |
| 04-07-25 | Operators,  Conditional statement | 1 hour 30 minutes | Practical + Theory class | Completed |
| 07-07-25 | Basic programs,  Control Statement | 1 hour 30 minutes | Practical + Theory class | Completed |
| 08-07-25 | Strings  String functions | 1 hour 30 minutes | Practical +  Theory class | Completed |
| 09-07-25 | User defined functions,  Collections-list and  Tuple | 1 hour 30 minutes | Practical + Theory class | Completed |
| 10-07-25 | Set and dictionary,  Class and object | 1 hour 30 minutes | Practical + Theory class | Completed |
| 11-07-25 | Oops concept | 1 hour 30 minutes | Practical + Theory class | Completed |
| 14-07-25 | NumPy library in  python, Arrays | 1 hour 30 minutes | Practical + Theory class | Completed |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 15-07-25 | Basic Functions in  NumPy | 1 hour 30 minutes | Practical + Theory class | Completed |
| 16-07-25 | Indexing and slicing  of array | 1 hour 30 minutes | Practical + Theory class | Completed |
| 17-07-25 | Arrays access, Import.  image using array | 1 hour 30 minutes | Practical + Theory class | Completed |
| 21-07-25 | Pandas’ library Analysing data from datasets | 1 hour 30 minutes | Practical + Theory class | Completed |
| 22-07-25 | Merge, Concatenation of  Data | 1 hour 30 minutes | Practical + Theory class | Completed |
| 23-07-25 | Filtering data | 1 hour 30 minutes | Practical + Theory class | Completed |
| 24-07-25 | Visualization, EDA | 1 hour 30 minutes | Practical + Theory class | Completed |
| 25-07-25 | SQL Installation and  Basic | 1 hour 30 minutes | Practical + Theory class | Completed |
| 28-07-25 | Introduction to DBMS  Introduction to SQL | 1 hour 30 minutes | Theory class | Completed |
| 29-07-25 | DDL commands  DML commands  DQL commands | 1 hour 30 minutes | Practical + Theory class | Completed |
| 30-07-25 | Data types  Constraints  Primary keys | 1 hour 30 minutes | Practical + Theory class | Completed |
| 31-07-25 | Operators | 1 hour 30 minutes | Practical + Theory class | Completed |
| 01-08-25 | Joins  Functions | 1 hour 30 minutes | Practical + Theory class | Completed |

**3.2 APPLICATION USED**

**Visual Studio Code (VS Code)**

* Purpose: Executed Python scripts for data preprocessing, cleaning, and transformation.
* Why VS Code: It offers an efficient coding environment with features like IntelliSense, debugging tools, and seamless integration with Python extensions.
* Usage: I wrote and tested core Python functions, including data manipulation using libraries such as pandas, NumPy, and custom logic for handling missing values and formatting.

**Google Colab**

* Purpose: Performed exploratory data analysis (EDA) and visualizations.
* Why Google Colab: It provides a cloud-based, collaborative environment with access to GPU acceleration and pre-installed libraries, making it ideal for data analysis.
* Usage: I imported the dataset into Colab to generate visual insights using libraries like matplotlib and seaborn. Colab’s interactive cells allowed for iterative exploration and real-time adjustments to the analysis.

**SQL Compiler**

* Used to practice and execute SQL queries in a structured environment.
* Facilitated learning of key SQL concepts such as joins, constraints, functions, and primary keys.
* Enabled interaction with sample databases to simulate real-world relational data scenarios.

**CHAPTER 4: CONCLUSION**

**4.1 LEARNING**

During my internship, I developed a solid understanding of several essential tools and technologies for programming and data analysis. I began with the basics of Python, learning core concepts such as variables, data types, loops, conditional statements, and functions. I also explored Python’s built-in data structures like lists, tuples, dictionaries, and sets, which helped me write efficient and organized code.

Next, I learned Pandas, a powerful Python library for data manipulation and analysis. I practiced working with Data Frames and Series, importing datasets from various sources, cleaning data, handling missing values, sorting, filtering, grouping, and performing aggregations. This allowed me to transform raw data into a structured format suitable for analysis.

To visually represent data, I used Matplotlib, creating plots such as bar charts, line graphs, histograms, and scatter plots. I learned how to customize visuals with labels, legends, and colours, making the presentation of results clearer and more engaging.

In addition, I worked with Pillow, a Python imaging library, to process and manipulate images. I learned how to open, resize, crop, and save images in different formats, as well as apply filters and enhancements. This skill added versatility to my Python knowledge, especially for projects involving image-based data.

I also gained knowledge of SQL (Structured Query Language) for database management. I wrote queries to create, retrieve, update, and delete records, and used advanced operations like joins, grouping, and ordering to extract meaningful information from relational databases.

Finally, I applied these skills to dataset analysis, performing data cleaning, exploratory analysis, visualization, and interpretation of results. The internship enhanced my problem-solving abilities, improved my technical expertise, and gave me firsthand experience in applying programming, visualization, image processing, and database skills to real-world scenarios.

**4.2 OUTCOME**

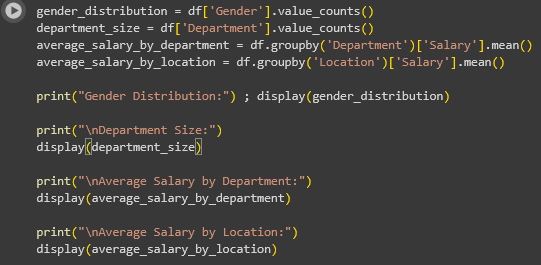
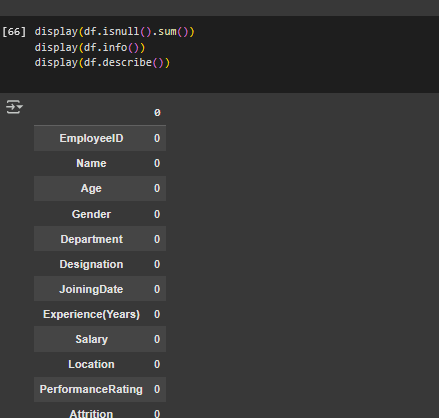
As part of my internship, I worked on a Student Performance Analysis project using Python. The main objective of the project was to identify the key factors affecting students’ exam performance and provide insights to help improve their results. I used a student dataset containing information such as attendance, study hours, parental education level, extracurricular activities, internet usage, sleep patterns, and other demographic and academic details.

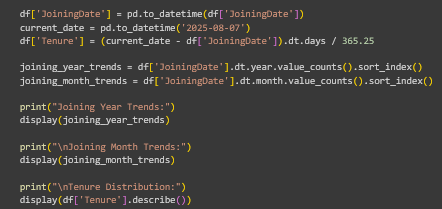
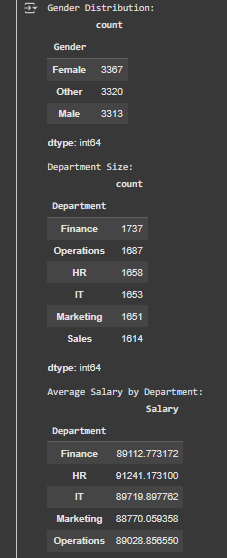
Using Pandas, I cleaned and organized the dataset, managed missing values, and grouped the data for meaningful comparisons. I applied Matplotlib to create visualizations like bar charts, histograms, and scatter plots, making it easier to identify patterns and relationships between distinct factors and exam scores. I analysed various aspects such as the impact of regular attendance, study habits, teacher quality, and motivation levels on student performance. The findings revealed that consistent attendance, adequate sleep, balanced extracurricular involvement, and strong teacher support significantly contribute to higher exam scores.

This project provided actionable insights that could help educators and institutions design better teaching strategies, improve student engagement, and create personalized learning plans, aiming to enhance overall academic performance.

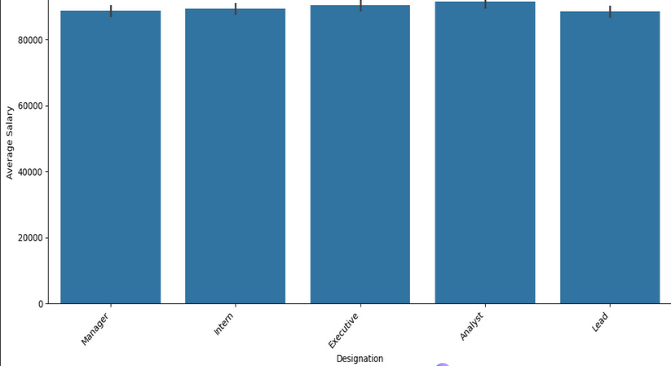
* 1. **SNAPSHOTS OF PROJECT WORK AND ANALYSIS**
     1. **Employee Data Analysis**

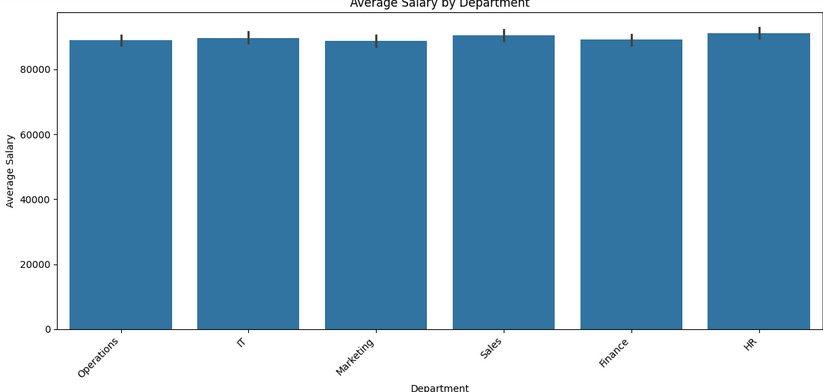
**4.3.1.1 Clean and preprocess the dataset 4.3.1.2 Analyze Gender distribution.**



**4.3.1.3 Identify trends in joining dates and tenure** 4.3.1.2**DESCRIPTION:**

**Clean and preprocess the dataset: Handle missing values, incorrect data types, and any inconsistencies in the df Data-Frame. There are no missing values and data types.** **The overall attrition rate was not explicitly calculated, but analysis of attrition employees suggests that salary and location are not strong individual predictors of attrition.**

**4.3.1.4 Visualize salary trends by Designation and Department**

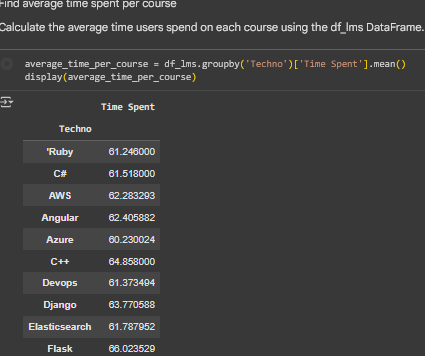
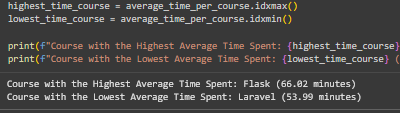


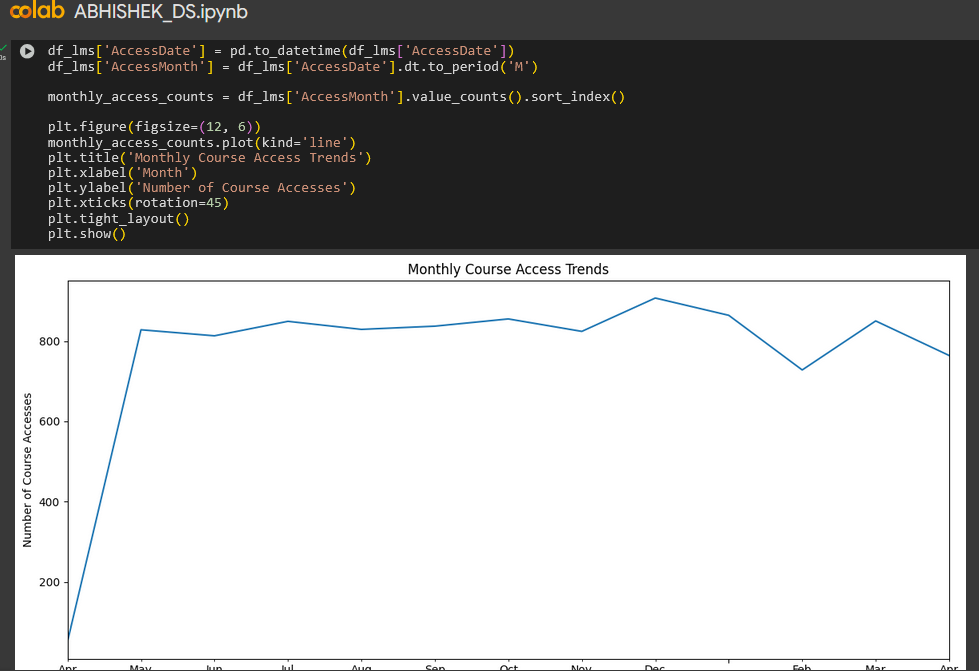
**DESCRIPTION:**

* Analyze gender distribution, department size, and average salaries: Analyze gender distribution, department size, and Average.
* Identify trends in joining dates and tenure: Identify trends in joining dates and tenure.
* Visualize salary distributions by designation and department for a better understanding.
* Employee joining trends show relatively consistent hiring year-over-year, with an even distribution across months. The average employee tenure is approximately 5.75 years.
  + 1. **LMS USAGE DETAILS**

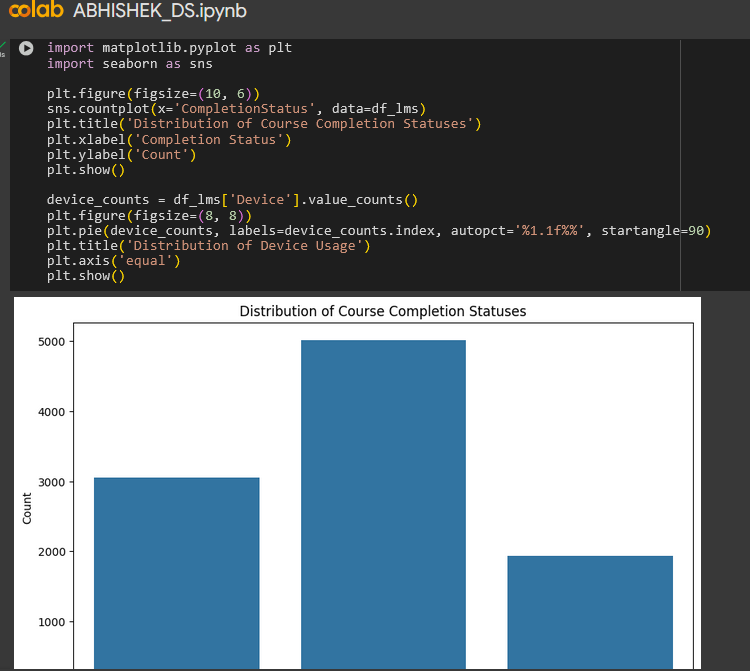
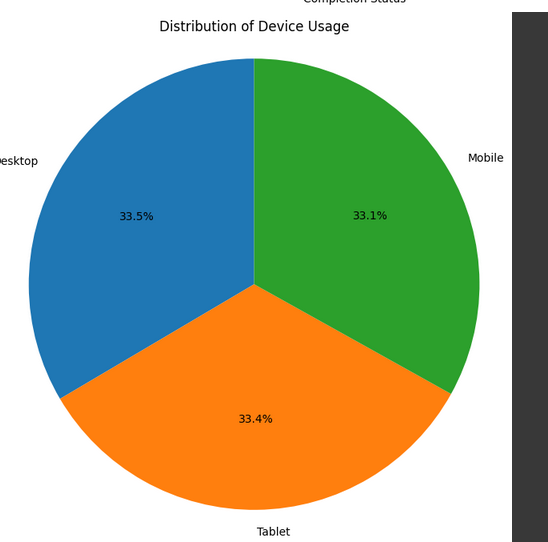
**4.3.2.1 Analyze most popular courses 4.3.2.2 Highlight courses where learners spend the most**

**and least time.**

  **4.3.2.3 Create time-series plots showing trends in course access over months**

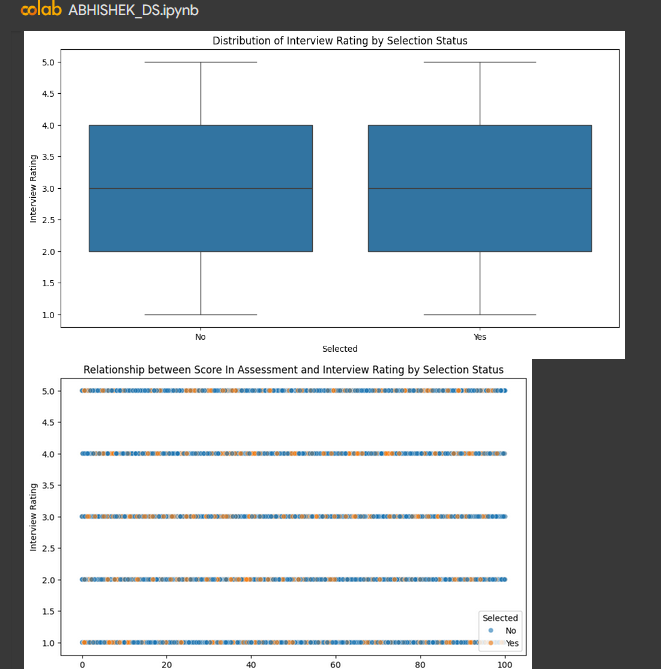
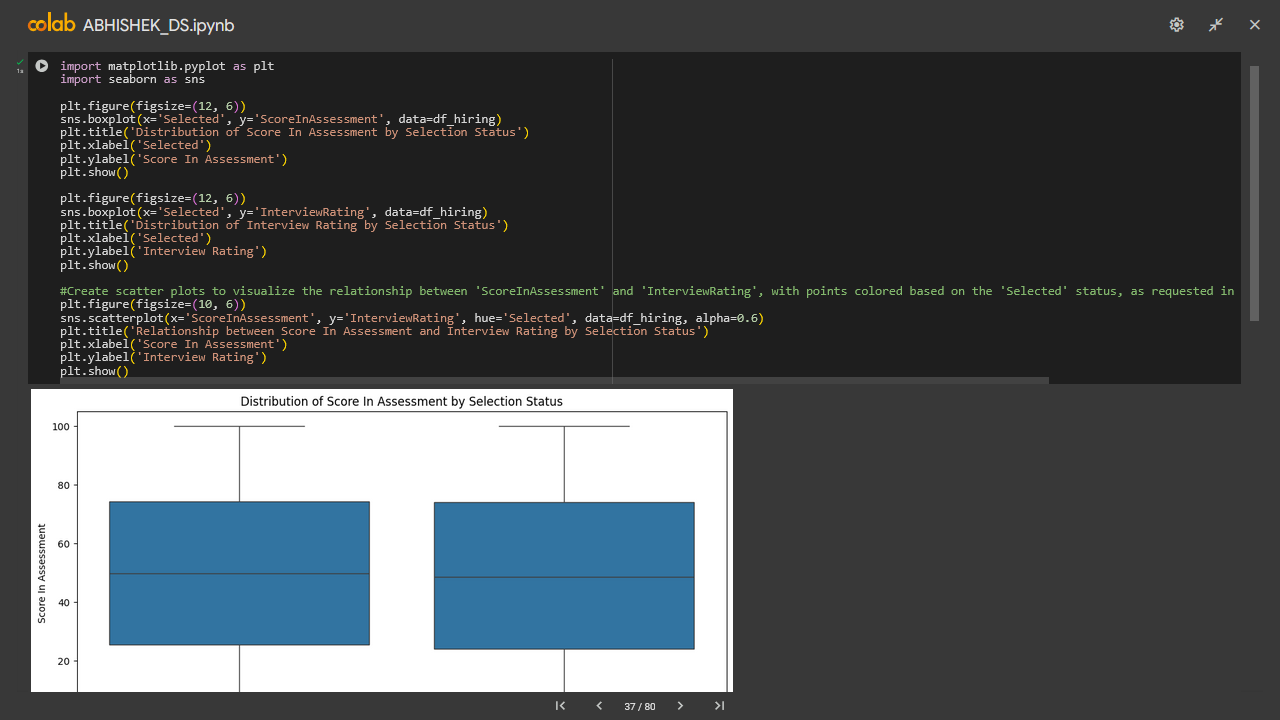
****

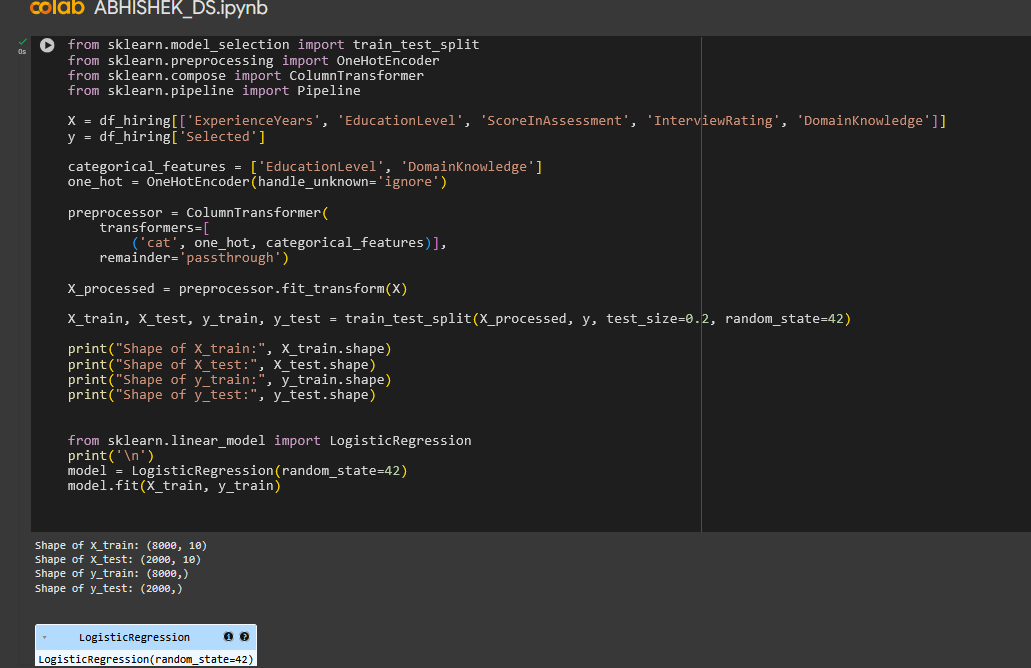
**4.3.2.4 Use bar charts and pie charts to depict completion distribution and device usage**

 ****

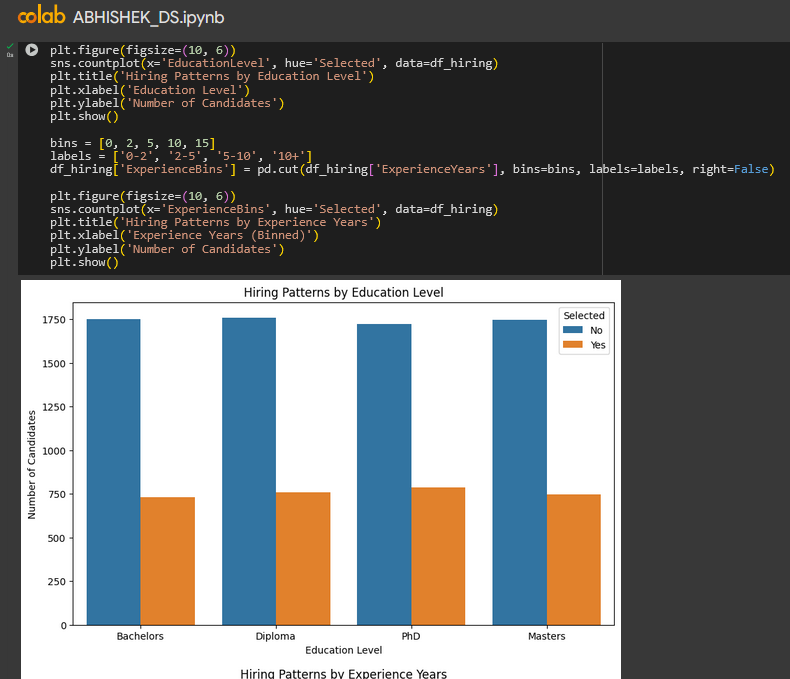
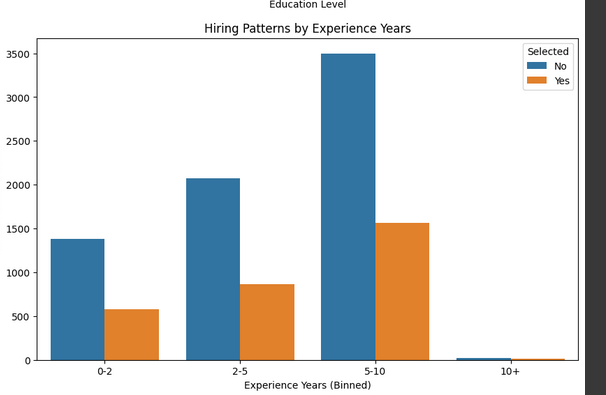
**DESCRIPTION:**

* Determine which course Techno column has the highest total number of users or accesses**.**
* Find average time spent per course. Calculate the mean Time Spent for each uniquecourse Techno
* Calculate course-wise Completion Status percentages. For each course, calculate the percentage of users who are Completed, In Progress, and Not Started.
* Aggregate data by month from the Access Date column and plot the total number users.
* . Visualize the distribution of the three Completion Status categories pie chart and the number of users accessing the LMS by Device bar chart
  + 1. **PREDICTIVE DATA ANALYTICS**
       1. **Perform Eda to understand relationships between test scores and hiring**

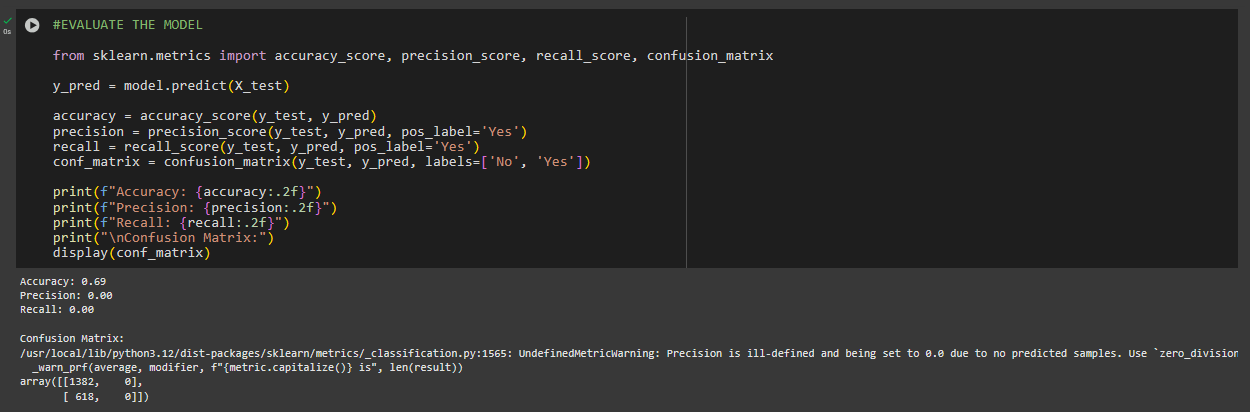
** 4.3.3.2 Build a binary classification model Logistic Regression or Decision Tree**

****

**4.3.3.3 Visualize hiring patterns by experience level and education**

** **

**4.3.3.4 Evaluate model using accuracy, precision, recall, and confusion matrix**

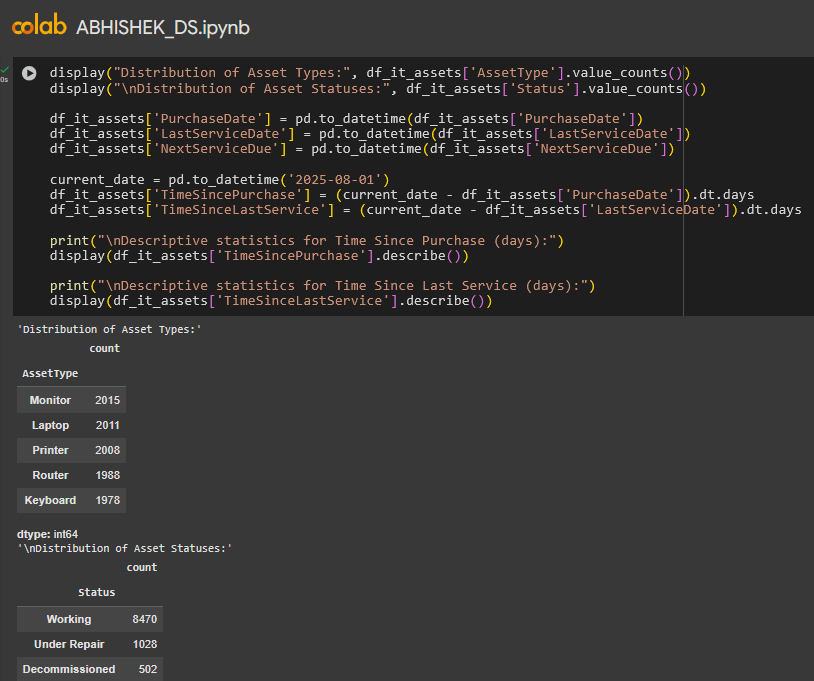
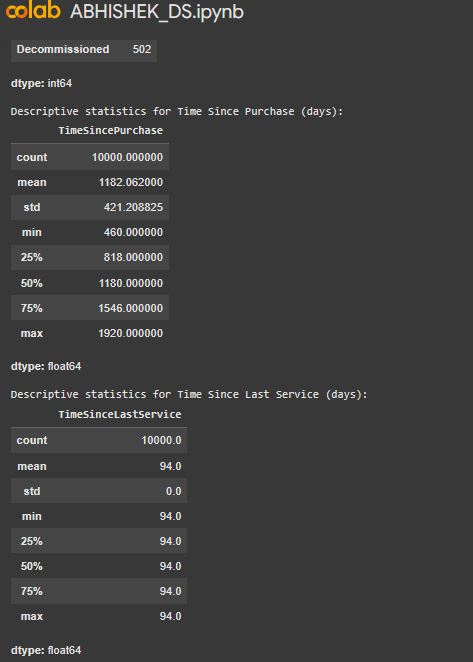
****

**DESCRIPTION:**

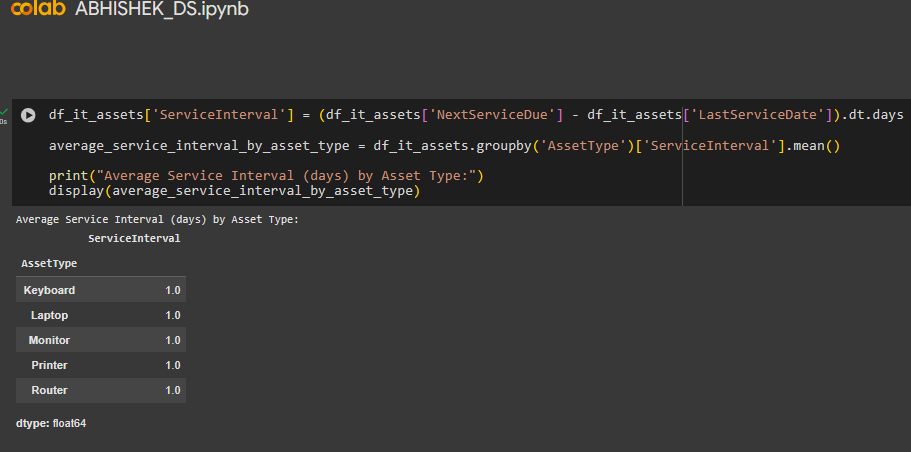
* Problem Statement: Quantify the immediate and long-term organizational deficiency in specifically required skills.
* Analytical and Categorical Analysis: Aggregate the count of Yes in the Gap column, grouped by Project Skill Required.
* Segmentation: Analyze the distribution of Proficiency Level for the Current Skill relative to the required skills.
* Actionable Insights: Provides a skills dashboard detailing where the company is most vulnerable to high demand for ML with a large 'Gap' count.
* This directly informs the L&D department on which training programs to launch and what expertise the recruitment team must source externally.
* The model predicts the probability of a candidate succeeding. This moves the selection process beyond subjective judgment, enabling the hiring team to focus interview efforts on candidates with the highest predictive chance of success, reducing hiring mistakes and improving new-hire quality

**4.3.4 IT ASSET MAINTENANCE FORECASTING**

**4.3.4.1 Analyze distribution of asset types and their service history**

** **

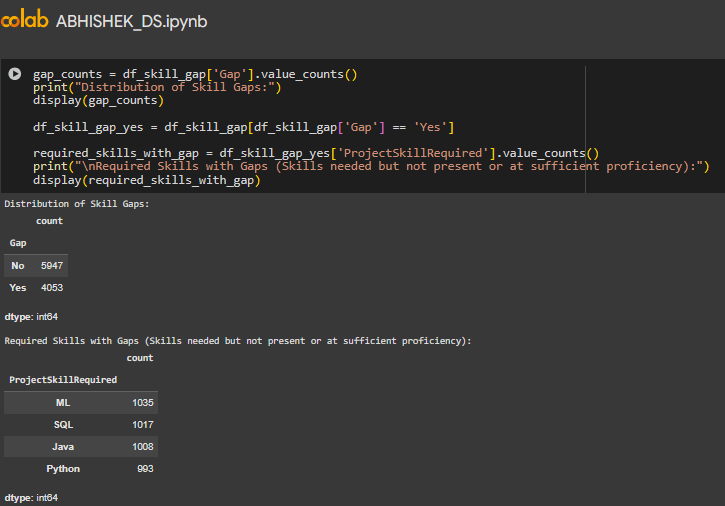
**4.3.4.2 Calculate average time between services**

****

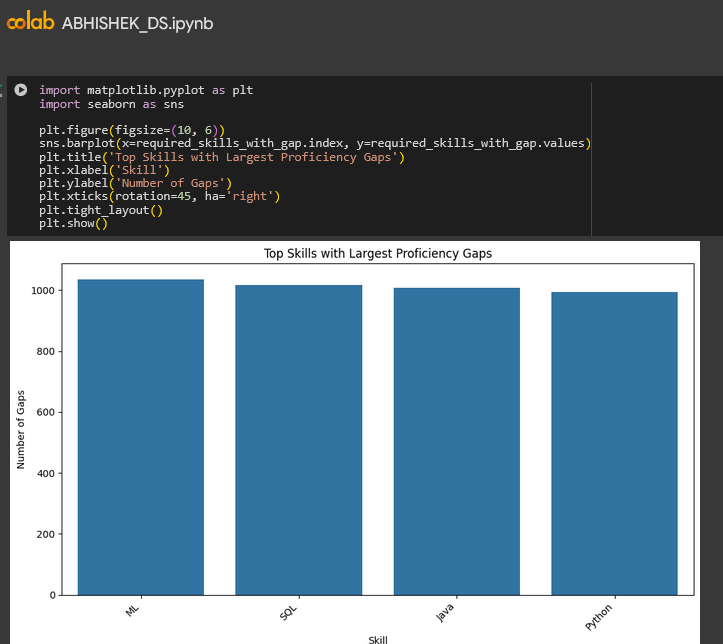
**DESCRIPTION:**

* Data Columns Used: Asset-ID, Asset-Type, Purchase-Date, LastService-Date, Next Service-Due, Status Working or Under Repair, and Location..
* Analytical Steps: Time Calculation: Calculate the Asset Age from Purchase Date and the Time Since Last Service from Last Service Date.
* Failure Rate Analysis: Group assets by AssetType and Location to identify types or locations with a high proportion of Under Repair in the Status column.
* Actionable Insights: Move from fixed-date servicing NextServiceDue to condition-based maintenance. Identify assets prone to failure before their designated service date to reduce interruptions and lower total cost of ownership.
  + 1. **SKILL GAP ANALYSIS**

**4.3.5.1 Compare current and required proficiency for each skill**



**4.3.5.2 Compare current and required proficiency for each skill**

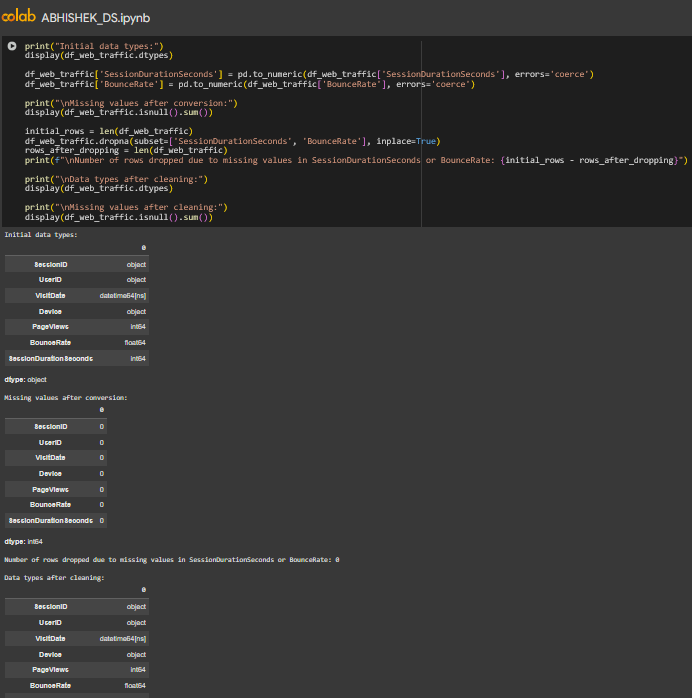
****

**DESCRIPTION:**

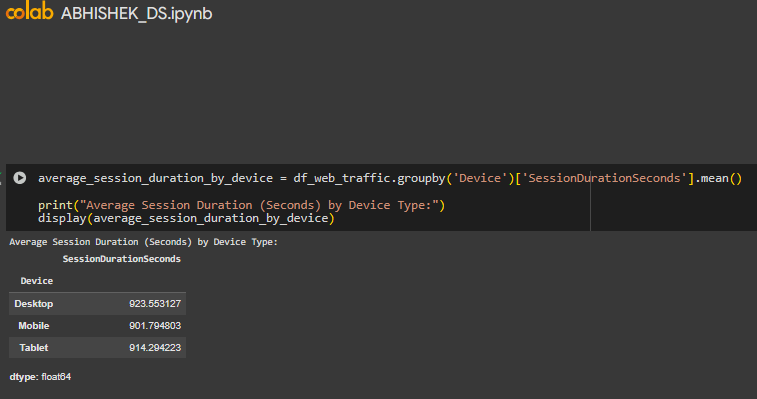
* The analysis reveals that 471 employees (35.3%) currently have a formal skill gap, with over 84% of all deficits concentrated in four high-value technical areas: Java, SQL, Python, and Machine Learning ML.
* These gaps primarily stem from employees who possess foundational analytical skills, such as Excel, but lack the core technical expertise needed for specific projects, pointing to a need for a transition into capabilities.
* The most significant challenge is the demand for ML, where the average current proficiency score is the lowest at 1.55 Beginner/Intermediate, suggesting a deep-seated deficit in advanced analytics roles requiring substantial training investment.
* When examining the entire workforce, the current proficiency levels are distributed relatively evenly: 35.4% are Advanced, 32.4% are Intermediate, and 32.1**%** areBeginners

**4.3.6 WEB TRAFFIC ANALYSIS**

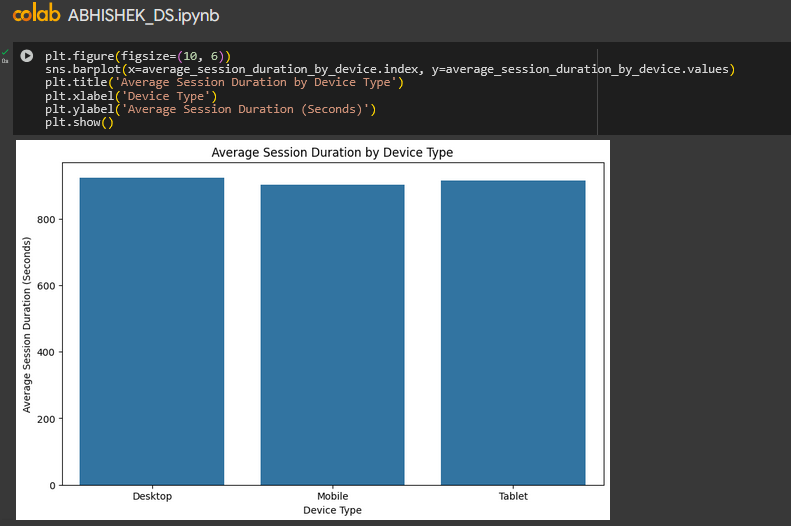
**4.3.6.1 Ensure Session Duration Seconds and Bounce Rate are in valid numerical formats**

****

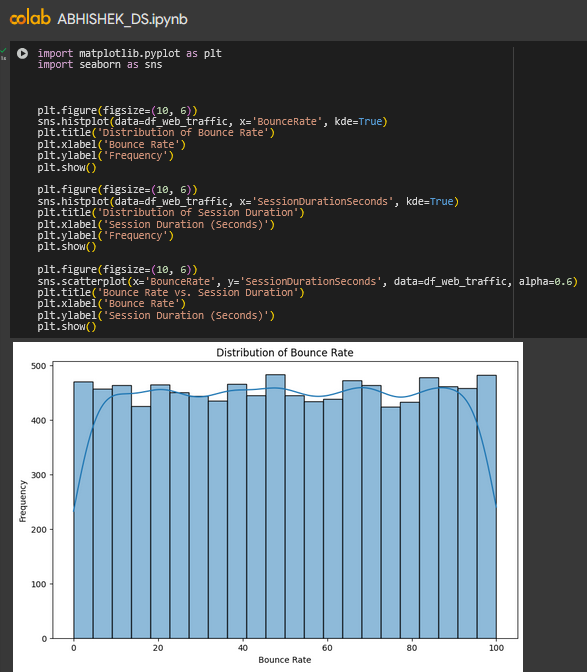
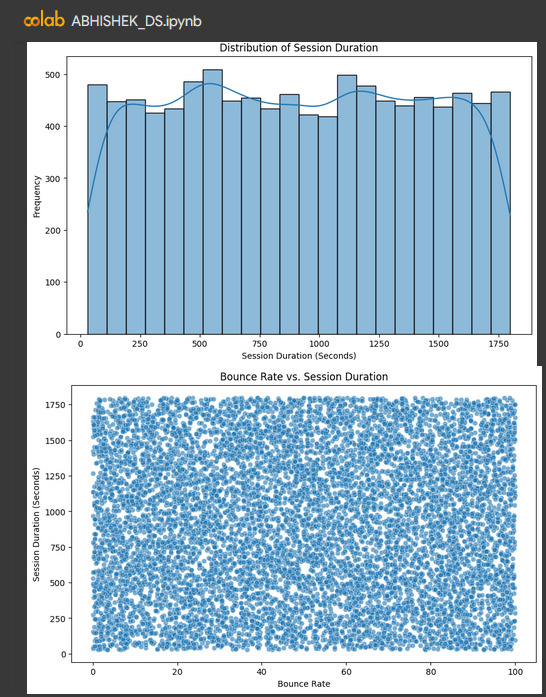
**4.3.6.2**  **Analyze average session duration seconds by device**

****

**4.3.6.3** **VISULALIZE DEVICE WISE AVERAGE SESSION DURATION**



**4.3.6.4** **CREATE HISTOGRAMS OR SCATTERPLOTS FOR BOUNCE\_RATE ,SESSION\_RATE**

**DESCRIPTION:**

* The Web Traffic Case Study is a Digital Marketing and Web Analytics analysis used to evaluate website performance and user behaviour.
* The primary goal is to diagnose performance issues and understand how users engage with the website across various devices.
* The key metrics analysed are Session Duration Seconds, which measures user engagement, and Bounce Rate, which indicates usability or relevance. Data preparation confirmed that both Session Duration Seconds and Bounce Rate were in valid numerical formats with no missing values, allowing for accurate analysis.
* Analysis of device-wise engagement revealed that Desktop users recorded the longest average session duration at approximately $923.55$ seconds.
* Tablet users followed with $Approx 914.29$ seconds, and Mobile users were the least engaged with $Approx 901.79$ seconds, which suggests that mobile optimization should be a priority to boost engagement for that segment.
* Finally, the scatterplot illustrates a clear inverse relationship between the two metrics sessions with a high Bounce Rate predictably correspond to a low Session Duration Seconds, which is a fundamental validation of user departure logic.
* The data provides actionable insights for the marketing and IT teams to focus optimization efforts on the most underperforming device.
* Finding**:** As expected, there is a general inverse relationship between the two metrics: sessions with a high Bounce Rate tend to have a low SessionDuration, while sessions with a low Bounce Rate are spread across all session durations, including the very long ones. This inverse pattern is a fundamental confirmation of digital engagement logic, where bouncing usually means immediate departure.

**4.5 SUMMARY OF PROJECT**

**Key Findings:**

* **Attrition Factors:** Salary and location were not strong individual predictors of attrition; however, a higher proportion of interns among attrited employees suggests role-specific factors may be at play.
* **Employee Trends:** Consistent hiring trends year-over-year and an even distribution of joining dates across months were observed, with an average employee tenure of around 5.75 years.
* **LMS Completion Rates:** The overall LMS course completion rate is 50%, with significant variation by course. Courses like PyTorch, Vue.js, AWS, PostgreSQL, and Elasticsearch show higher completion rates.
* **Workforce Characteristics:** The employee data revealed balanced gender distribution and variations in department sizes.
* **Hiring Model Improvement:** More complex models, addressing class imbalance, in-depth feature engineering, and understanding the business context are needed to improve hiring predictions.
* Explore user behavior patterns within sessions to understand engagement beyond simple metrics.
* **Predictive Data Analytics:** The hiring analysis underscores the challenge of accurately predicting successful hires based on the current data and model. The business implication is that the current hiring process, or the data collected from it, may not effectively identify the most suitable candidates.
* **IT Asset Maintenance Forecasting:** The issues identified in the IT asset data regarding service intervals have significant business implications for IT operations and budgeting. Inaccurate service data hinders effective maintenance planning, potentially leading to unexpected equipment failures, increased repair costs, and operational disruptions.
* **Web Traffic Analysis:** The web traffic analysis provides insights into user engagement with online platforms. While the current analysis shows weak correlations between basic metrics, understanding user behavior is essential for optimizing website design, content strategy, and marketing efforts.

.**4.6 RECOMMENDATIONS TO IMPROVE EMPLOYEE PERFORMANCE**

* **Analyze Relationship Between Tenure and Performance:** Explore if there's a correlation between employee tenure and performance rating. This could help identify if newer employees require more support or if long-tenured employees could benefit from new challenges or development opportunities
* **Link Performance to Skill Development:** The skill gap analysis is highly relevant here. Employees with identified skill gaps might have lower performance. Implement targeted training programs to address these gaps and potentially improve performance.
* **Offer Varied Training and Development Opportunities:** Provide access to a range of training programs, workshops, and online courses that cater to different skill levels and career aspirations, directly addressing the skill gaps identified **Offer Varied Training and Development Opportunities:** Provide access to a range of training programs, workshops, and online courses that cater to different skill levels and career aspirations, directly addressing the skill gaps identified.
* **Consider Department and Location Specific Initiatives:** While average salaries didn't show extreme variations, there might be nuances within departments or locations affecting performance. Further investigate performance by department and location to identify any specific issues or best practices.
* **Implement a Robust Mentorship Program:** Pair new employees, particularly in their first year, with experienced mentors to provide guidance, support, and facilitate their integration into the company.
* **Gather Employee Feedback Regularly:** Implement regular employee pulse surveys or feedback mechanisms to proactively identify areas of concern related to job satisfaction, work environment, and opportunities for improvement.
* **Foster a Positive Work Culture:** Focus on building a supportive and inclusive work environment that promotes collaboration, open communication, and employee well-being
* **Promote Internal Mobility and Career Progression:** Create clear pathways for career advancement within the company and encourage employees to explore different roles and departments to foster growth and reduce stagnation.