



Tribhuvan University

Faculty of Humanities and Social Science

An Internship Report

On

Data Analyst

At

Sanima Bank PVT.LTD

Under the supervision of

Roshan Chaudhary

Submitted to

Department of Computer Application

Orchid International College

In partial fulfillment of the requirements for the Bachelors in Computer Application

Submitted by

Sourav Rajbanshi (93902085)

April 2025



Tribhuvan University
Faculty of Humanities and Social Science
Orchid International College

SUPERVISOR'S RECOMMENDATION

I hereby recommend that this report be prepared under my supervision by Sourav Rajbanshi (TU Roll No (93902085) in partial fulfillment of the requirements for the degree of Bachelor of Computer Application (BCA) and be processed for evaluation.

.....

Roshan Chaudhary

Internship Supervisor, Department of BCA

Orchid International College

Bijaychowk, Gaushala



Tribhuvan University
Faculty of Humanities and Social Science
Orchid International College

CERTIFICATE OF APPROVAL

The undersigned certify that they have read and recommended to the Department of Humanities and Social Science the acceptance of an internship report submitted by Mr. Sourav Rajbanshi in fulfillment of the degree requirements for the Bachelor in Computer Application (BCA).

<hr style="border-top: 1px dashed black;"/> Roshan Chaudhary Internal Examiner/Project Coordinator, Department of IT Orchid International College Bijaychowk, Gaushala	<hr style="border-top: 1px dashed black;"/> External Examiner Department Of Humanities and Social Science Tribhuvan University Kiritipur, Kathmandu
---	--

ACKNOWLEDGEMENT

Sincere gratitude is extended to all those who contributed to the successful completion of this internship and the preparation of this report. Immense thanks are owed to Sanima Bank PVT. LTD. for providing me with the opportunity to work as a Data Analyst Intern. The experience significantly enriched my practical knowledge and allowed me to apply analytical skills in a real-world banking environment.

I would like to express my special appreciation to the Card & Digital Service Department for their constant support, expert guidance, and valuable insights. Their mentorship played a crucial role in helping me understand data-driven decision-making processes and develop a deeper understanding of the financial sector.

Appreciation is also extended to all the team members at Sanima Bank PVT. LTD. for their warm welcome, encouragement, and for creating a collaborative work environment that fostered learning and professional growth.

Heartfelt thanks go to my family and friends for their unwavering support, motivation, and belief in my abilities throughout this internship journey.

Lastly, I express my gratitude to all individuals, institutions, and resources whose work has been referenced in this report. Their contributions provided valuable information that supported the successful completion of this document.

With sincere regards,

Sourav Rajbanshi

(Reg No: 6-2-939-78-2020)

TU Roll No: 93902085

ABSTRACT

This internship report presents an in-depth account of the professional experience acquired during my internship as a Data Analyst at Sanima Bank PVT. LTD. The primary aim of the internship was to bridge the gap between academic learning and practical application by immersing in a real-world work environment within the banking industry. Over the course of the internship, I was exposed to a variety of data-related tasks and analytical processes that are critical to the bank's daily operations and strategic decision-making.

The report details my involvement in several key activities, including data extraction, cleaning, preprocessing, and analysis of large datasets using tools such as Microsoft Excel, SQL, and Python. I assisted in generating dashboards and reports, identifying patterns and trends, and contributing to data-driven insights that supported business objectives across departments. Through this hands-on experience, I gained a deeper understanding of how financial institutions use data analytics to monitor performance, manage risks, improve customer services, and ensure regulatory compliance.

In addition to technical skills, the internship enhanced my soft skills, such as teamwork, communication, time management, and problem-solving, by allowing me to collaborate with experienced professionals in a structured corporate setting. This report covers the objectives of the internship, organizational structure, roles and responsibilities assigned, challenges faced, and the valuable lessons learned throughout the period.

Overall, the internship was a valuable step in my academic and professional growth, providing practical exposure to data analytics in the banking sector and strengthening my interest and skills for a future career in this field.

Keywords: *Data Analyst, Internship, Sanima Bank, Data Analytics, Banking Sector, SQL, Python, Microsoft Excel, Data Visualization, Business Intelligence, Financial Data, Professional Development, Real-world Experience*

TABLE OF CONTENT

SUPERVISOR’S RECOMMENDATION	i
CERTIFICATE OF APPROVAL	ii
ACKNOWLEDGEMENT	iii
ABSTRACT.....	iv
TABLE OF CONTENT	v
LIST OF FIGURES	vii
LIST OF TABLES	viii
LIST OF ABBREVIATION.....	ix
Chapter 1: Introduction.....	1
1.1 Introduction.....	1
1.2 Problem Statement	2
1.3 Objectives	2
1.4 Scope and limitation	2
1.4.1 Scope.....	2
1.4.2 Limitation.....	3
1.5 Report Organization.....	3
Chapter 2: Introduction to Organization.....	4
2.1 Organization Details	4
2.2 Organization Hierarchy.....	5
2.3 Working Domains of the Organization	6
2.4 Description of Intern Department Unit	7
Chapter 3: Background Study and Literature Review	9
3.1 Background Study.....	9
3.2 Core Responsibilities	10
3.3 Literature Review.....	11

Chapter 4:	Internship Activities	15
4.1	Roles and Responsibilities	15
4.2	Weekly log	17
4.3	Description of the Project Involved During Internship.....	19
4.4	Tasks and Activities Performed	21
4.4.1	Data Cleaning and Preparation	21
4.4.2	Exploratory Data Analysis (EDA)	21
4.4.3	Data Visualization	22
4.4.4	Drafting report	23
4.4.5	Final Report and Presentation	23
4.4.6	Tools Used	23
Chapter 5:	Conclusion and Learning Outcomes	28
5.1	Conclusion	28
5.2	Learning Outcomes	28
References.....		30
Appendix.....		32

LIST OF FIGURES

Figure 1: Organizational Hierarchy	5
Figure 2: Sample Dataset in excel	21
Figure 3: statistical summary of dataset.....	22
Figure 4: Data Visualization using matplotlib	22
Figure 5: Correlation of the dataset	23

LIST OF TABLES

Table 1: Organizational Details.....	5
Table 2: Internship Details	8
Table 3: Weekly log with Task Performed and Details	17

LIST OF ABBREVIATION

AI	: Artificial Intelligence
API	: Application Program Interface
BI	: Business Intelligence
CSV	: Comma Separated Value
ETL	: Extract, Transform, Load
ML	: Machine Learning
IDE	: Integrated Development Environment

Chapter 1: Introduction

1.1 Introduction

This internship report encapsulates the enriching experience and professional growth achieved during a data analyst internship at Sanima Bank PVT. LTD., with a particular emphasis on the Bank Data Analysis project. The internship was designed to cultivate practical expertise in key data analytics processes, including data collection, cleaning, analysis, and visualization, leveraging tools such as Excel, SQL, and Python. The primary objective was to bridge the gap between theoretical knowledge acquired in academic settings and its practical application in a real-world business environment, fostering advanced problem-solving, data interpretation, and analytical thinking skills. These skills are increasingly critical in the banking sector, where data-driven insights drive strategic decisions, optimize operations, and enhance customer engagement.

The cornerstone of the internship was the Bank Data Analysis project, which focused on analyzing a comprehensive Bank Marketing Dataset to predict whether clients would subscribe to a term deposit based on various demographic and behavioral attributes. The dataset, comprising 45,211 records and 17 features, served as a robust platform to apply machine learning techniques, specifically Decision Tree, Random Forest, Naive Bayes, and K-Nearest Neighbors models, to build and evaluate predictive models. The project involved preprocessing complex, initially unstructured data, developing and comparing multiple classification models, and visualizing performance metrics to identify the most effective approach for predicting deposit outcomes. This hands-on experience not only reinforced technical proficiency in Python-based data analysis and visualization but also underscored the importance of translating raw data into actionable business insights.

The internship provided a unique opportunity to explore the pivotal role of data analytics in the financial industry, where accurate predictions and informed strategies can significantly impact customer acquisition and retention. By working on real-world data challenges, the intern developed a deeper understanding of how analytical tools and methodologies support organizational growth, streamline marketing efforts, and enhance operational efficiency in a competitive and rapidly evolving business landscape. This report details the objectives, methodologies, challenges, and outcomes of the Bank Data Analysis

project, highlighting its contributions to the intern's skill set and professional readiness. Ultimately, the internship experience has laid a solid foundation for a career in data analytics, equipping the intern with the technical expertise, analytical acumen, and practical insights needed to thrive in data-driven decision-making environments.

1.2 Problem Statement

In today's data-driven world, organizations face challenges in collecting, processing, and interpreting large volumes of data to make informed decisions. Sanima Bank PVT. LTD., like many financial institutions, generates vast amounts of data daily, which requires effective analysis to extract meaningful insights. However, transforming raw data into actionable information is often hindered by issues such as data inconsistency, lack of automation, and limited analytical tools. The objective of this internship was to address these challenges by assisting in data cleaning, analysis, and visualization processes, thereby supporting the bank's efforts in data-driven decision-making and operational efficiency.

1.3 Objectives

The main purpose of the internship at Sanima bank is to achieve following objectives:

- To apply theoretical knowledge of data analysis in a real-world setting.
- To enhance skills in data cleaning, processing, and visualization.
- To gain hands-on experience with tools like Excel, SQL, and Python
- To assist in identifying trends and patterns from large datasets.
- To support data-driven decision-making within the organization.

1.4 Scope and limitation

1.4.1 Scope

Some significant scopes include:

- Gained practical experience in handling and analyzing real-world banking data.
- Worked with tools like Excel, SQL, and Python for data-related tasks.
- Assisted in creating data visualizations and reports for internal use.
- Contributed to identifying patterns and trends to support business insights

- Applied data cleaning and preprocessing techniques to improve data quality.

1.4.2 Limitation

Some significant limitations include:

- Limited access to sensitive or confidential data due to privacy policies.
- Time constraints restricted the depth of analysis for complex projects.
- Exposure was confined to specific departments and data sources.
- Some advanced tools and techniques were not explored due to internship duration.
- Limited decision-making authority in actual business operations.

1.5 Report Organization

The report is divided into four sections:

Chapter 1:

The first chapter details the project's introduction, including the problem statement, objectives, and scope.

Chapter 2:

The second chapter contains organization details.

Chapter 3:

The third chapter contains background study and Literature review.

Chapter 4:

The fourth chapter dives deep into the analysis of activities completed and problems resolved.

Chapter 5:

The fifth chapter contains the project's conclusion and learning outcome

Chapter 2: Introduction to Organization

2.1 Organization Details

Sanima Bank Limited, established in 2004, is a prominent "A" Class Commercial Bank in Nepal, initially founded as a national-level development bank by a group of dynamic Non-Resident Nepalese (NRN) businesspersons. With its motto of being "strong and reliable," the bank transitioned into a commercial bank, receiving its operating license from Nepal Rastra Bank (NRB), and is now recognized as the 32nd commercial bank in Nepal. Headquartered in Naxal, Kathmandu, Sanima Bank operates with a robust network of over 27 full-fledged branches across the country, catering to both urban and rural markets.

The bank's mission focuses on delivering superior customer service, ensuring employee welfare, and upholding corporate social responsibility. Sanima Bank offers a wide range of financial products and services, including savings accounts like Prime Saving Account and Sanima Bishesh Bachat Khata, fixed deposits, loans, and modern banking facilities such as Any Branch Banking System (ABBS), Electronic Cheque Clearing (ECC), and fee-free ATM services across various bank terminals. Its financial structure comprises a paid-up capital of NPR 2.02 billion, with a shareholding pattern of 70% promoters and 30% general public.

Sanima Bank emphasizes technological innovation and operational efficiency to enhance customer experience, making it a key player in Nepal's banking sector. As a data analyst intern at Sanima Bank, the opportunity to engage with its data-driven processes, such as analyzing financial metrics, customer transactions, and operational performance, provides valuable insights into the practical application of analytical tools in a dynamic financial environment.

Sanima Bank is also committed to sustainable growth and financial inclusion, actively supporting microfinance initiatives and community development programs to uplift underserved populations in Nepal. By leveraging advanced digital banking platforms, such as mobile banking and internet banking, the bank ensures seamless access to financial services, fostering customer convenience and trust. Its strategic focus on risk management and compliance with regulatory standards further strengthens its reputation as a reliable institution. Working as a data analyst intern in such a forward-thinking organization offers exposure to cutting-edge data analytics practices, contributing to informed decision-making and the bank's continued success in a competitive market.

Table 1: Organizational Details

Organization	Sanima Bank PVT.LTD.
Area	Kathmandu Province No.3
Address	Alakapuri, Naxal, Kathmandu
Telephone	+977-1-5970033
Email	sanima@sanimabank.com
Official Website	https://www.sanimabank.com/know-us
Opening Hours	9:30 AM to 5:30 PM

2.2 Organization Hierarchy

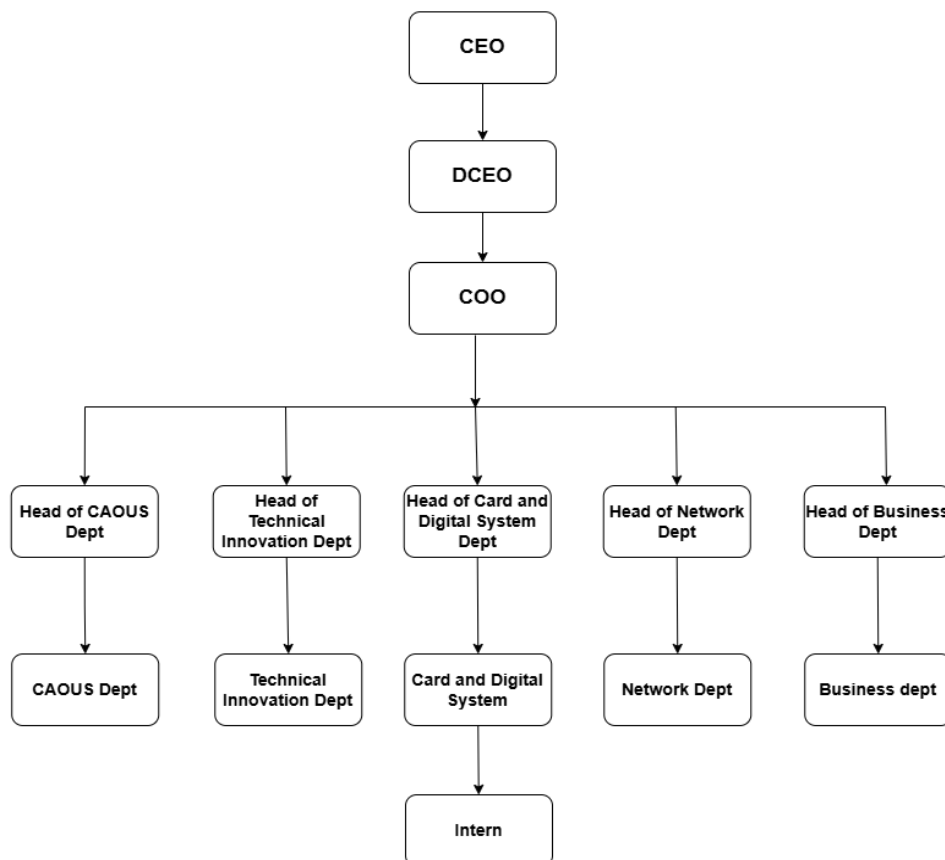


Figure 1: Organizational Hierarchy

The above diagram represents the organizational hierarchy of Sanima Bank Pvt. Ltd., illustrating the chain of command and departmental structure within the bank. At the top of the hierarchy is the Chief Executive Officer (CEO), who is responsible for the overall leadership and strategic direction of the bank. The Deputy Chief Executive Officer (DCEO) supports the CEO and oversees various operational functions.

Below the DCEO is the Chief Operating Officer (COO), who manages the bank's day-to-day operations and supervises multiple departmental heads. Reporting directly to the COO are the heads of several key departments, including:

- Head of CAOUS Department
- Head of Technical Innovation Department
- Head of Card and Digital System Department
- Head of Network Department
- Head of Business Department

Each department head manages their respective units. The Card and Digital System Department is further extended to include interns, showing the practical involvement of trainees within the organizational framework. My position as a Data Analyst Intern falls under this department, highlighting my contribution and the learning experience gained in the digital services area of the bank.

This structured hierarchy ensures smooth coordination, clear reporting lines, and effective execution of tasks across departments, supporting the bank's goal of operational excellence and innovation.

2.3 Working Domains of the Organization

Sanima Bank Pvt. Ltd. operates across a diverse range of domains to ensure comprehensive financial services and digital innovation. The key working domains of the organization include:

- **Retail and Corporate Banking**

Providing banking services such as savings accounts, loans, fixed deposits, and financial solutions for individual and corporate clients.

- **Digital Banking and Card Services**
Managing mobile banking, internet banking, debit/credit card operations, and other digital financial services to promote cashless transactions.
- **Technical Innovation and IT Infrastructure**
Implementing and maintaining advanced technological solutions, software systems, and network security to ensure reliable and efficient operations.
- **Strategic and Research Development (SARD)**
Conducting data-driven research, performance analysis, and strategic planning to support informed decision-making and long-term growth.
- **Compliance, Audit, and Operational Support (CAOUS)**
Ensuring regulatory compliance, risk management, internal audit operations, and overall procedural integrity of the bank.
- **Business and Market Development**
Expanding the bank's market presence, developing new financial products, and building relationships with customers and business partners.

2.4 Description of Intern Department Unit

During my internship at Sanima Bank Pvt. Ltd., I was placed in the **Card and Digital System Department**. This department is instrumental in steering the bank's digital transformation and managing its electronic banking services.

Key Responsibilities of the Department:

- **Digital Banking Services:** Overseeing and enhancing platforms such as mobile banking (Sanima Sajilo e-Banking), internet banking, and QR code payments to ensure seamless customer experiences.

- **Card Services Management:** Administering the issuance and operation of debit and credit cards, including the VISA domestic debit card and the Sanima USD eCard, facilitating both local and international transactions.
- **Technical Coordination:** Collaborating with external vendors and partners to integrate new digital tools and services, ensuring the bank remains at the forefront of technological advancements.
- **System Security and Performance:** Ensuring the security, reliability, and optimal performance of all digital banking systems to protect customer data and maintain trust.
- **Innovation Projects:** Supporting digital innovation initiatives aimed at enhancing customer experience and operational efficiency.

My Role as a Data Analyst Intern:

As a Data Analyst Intern, I contributed to various tasks, including data analysis, report generation, and monitoring digital service performance. This role provided me with valuable exposure to real-time banking data, customer behavior patterns, and the practical applications of data analytics in a financial institution. Working alongside experienced professionals, I gained insights into the operational aspects of digital banking and the strategic importance of data-driven decision-making.

Table 2: Internship Details

Start Date	January 22, 2025
End Date	April 20, 2025
Position	Data Analyst Intern
Work Hours	8 hours a day
Office Days	Sunday – Friday (6 days)
Internship Period	3 months

Chapter 3: Background Study and Literature Review

3.1 Background Study

The role of a data analyst has become increasingly vital in today's data-driven world, particularly within the banking sector, where organizations like Sanima Bank Pvt. Ltd. rely on data to make informed strategic decisions, enhance operational efficiency, and improve customer experiences. A data analyst is responsible for collecting, processing, and analyzing large volumes of data to uncover meaningful insights, trends, and patterns that guide business strategies. In the context of a commercial bank like Sanima Bank, which operates in a competitive and highly regulated environment, the data analyst plays a pivotal role in transforming raw financial and operational data into actionable intelligence. This background study explores the significance of the data analyst role, its relevance to the banking industry, and its specific application within Sanima Bank, providing a foundation for understanding the scope of responsibilities undertaken during an internship.

The banking sector in Nepal, including institutions like Sanima Bank, has undergone significant transformation due to technological advancements and the growing adoption of digital banking platforms. With an expansive network of branches and services such as mobile banking, internet banking, and Any Branch Banking System (ABBS), Sanima Bank generates vast amounts of data daily, including customer transactions, loan portfolios, savings account activities, and operational metrics. Analyzing this data is critical for optimizing processes, identifying market opportunities, managing risks, and ensuring regulatory compliance. Data analysts employ statistical tools, data visualization techniques, and programming languages like Python, R, and SQL to extract insights from complex datasets. For instance, they may analyze customer behavior to tailor personalized financial products or assess credit risk to minimize loan defaults, directly contributing to the bank's profitability and customer satisfaction.

The role of a data analyst in a bank extends beyond technical expertise to encompass business acumen and communication skills. Analysts must translate complex data findings into clear, actionable recommendations for stakeholders, such as bank managers or executives, who may not have a technical background. At Sanima Bank, which emphasizes innovation and customer-centric services, data analysts are instrumental in supporting strategic initiatives like financial inclusion and microfinance programs. They may evaluate the performance of initiatives like the Sanima Bishesh Bachat Khata or assess the impact

of community development programs, ensuring alignment with the bank's mission of sustainable growth. Furthermore, data analysts contribute to regulatory reporting by ensuring accurate and timely submission of financial data to Nepal Rastra Bank (NRB), reinforcing the bank's compliance with national standards.

In the context of an internship at Sanima Bank, the data analyst role offers hands-on exposure to real-world applications of data analytics in a financial institution. Interns are typically tasked with assisting in data cleaning, performing exploratory data analysis, creating dashboards, and generating reports that support decision-making. Tools such as Microsoft Excel, Tableau, or Power BI are commonly used for visualization, while SQL is leveraged for querying large databases. Interns may also work on predictive modeling or trend analysis to forecast key metrics like deposit growth or loan repayment rates. This practical experience not only hones technical skills but also fosters an understanding of the banking sector's operational dynamics, risk management frameworks, and customer engagement strategies. The internship provides a unique opportunity to bridge theoretical knowledge with industry practices, preparing interns for future roles in data analytics or related fields.

The growing importance of data analytics in banking is driven by global trends such as big data, artificial intelligence, and machine learning, which are reshaping how financial institutions operate. Sanima Bank, with its commitment to technological innovation, is well-positioned to leverage these advancements, making the data analyst role a cornerstone of its strategic vision. By analyzing data from diverse sources such as ATM transactions, digital banking interactions, and customer feedback analysts help the bank stay competitive in a rapidly evolving market. For an intern, contributing to these efforts offers valuable insights into the intersection of technology and finance, equipping them with skills that are highly sought after in the global job market. This background underscores the critical role of data analysts in driving Sanima Bank's success and highlights the learning opportunities available during an internship in this dynamic field.

3.2 Core Responsibilities

- **Data Collection and Cleaning:** Gathering data from various sources, such as customer transaction records, loan portfolios, and operational databases, and ensuring its accuracy, completeness, and consistency for analysis.

- **Data Analysis:** Using statistical tools and techniques to interpret complex datasets, identifying trends, patterns, and correlations that inform business strategies and operational improvements.
- **Data Visualization:** Creating visual representations, such as charts, graphs, and dashboards, using tools like Tableau or Power BI to make complex data findings accessible and understandable to non-technical stakeholders.
- **Reporting:** Preparing comprehensive reports that summarize data analysis results, highlight key insights, and provide actionable recommendations to support decision-making processes.
- **Collaboration:** Working closely with departments like risk management, marketing, and finance to understand their data needs and provide tailored analytical support to achieve organizational goals.

3.3 Literature Review

The application of data analytics in the banking sector has garnered significant attention in academic and industry literature, particularly as financial institutions like Sanima Bank Pvt. Ltd. increasingly rely on data-driven decision-making to enhance operational efficiency, customer satisfaction, and regulatory compliance. This literature review synthesizes key findings from existing studies and publications to contextualize the role of a data analyst in a commercial bank, highlighting theoretical frameworks, practical applications, and emerging trends relevant to an internship at Sanima Bank. The review incorporates insights from recent works to provide a comprehensive backdrop for the Bank Data Analysis project undertaken during the internship, which focused on predicting client deposit behavior using machine learning models.

- **Data Analytics in Banking Operations:**
According to Laudon and Laudon (2020), data analytics transforms raw data into actionable insights, a process critical for banks managing vast datasets from customer transactions, loan applications, and digital banking platforms. Similarly,

Challa (2023) emphasizes the transformative impact of data analytics on the future of banking, highlighting its role in optimizing operational processes through predictive modeling and real-time analytics. In the context of Nepal's banking sector, Shrestha (2019) notes that commercial banks, such as Sanima Bank, leverage data analytics for credit risk assessment and fraud detection. By employing statistical models and machine learning algorithms, banks can predict loan default probabilities, reducing financial losses. This aligns with the responsibilities of a data analyst intern, who may assist in cleaning transactional data and developing predictive models, as demonstrated in the Bank Data Analysis project, where models like Random Forest and Decision Tree were used to predict deposit outcomes.

- **Customer-Centric Analytics:**

Customer relationship management (CRM) is a pivotal area where data analytics drives value. Kotler and Keller (2016) argue that analyzing customer data enables businesses to segment markets and personalize offerings, a strategy widely adopted in banking to enhance customer retention. Kiran Singh and Ramaswamy (2016) further elaborate on the importance of data analysis and visualization in understanding customer behavior, such as purchase patterns in sales data, which can be extended to banking contexts like deposit subscriptions. In Nepal, Adhikari and Sharma (2021) highlight how banks use visualization tools like Tableau to analyze customer behavior, enabling targeted marketing campaigns for products like savings accounts. For example, Sanima Bank's Sanima Bishesh Bachat Khata may rely on such analytics to identify eligible customer segments. The Bank Data Analysis project mirrors this by using a dataset to predict deposit behavior, with visualization techniques like bar charts and ROC curves to communicate model performance. Interns contribute to creating such visualizations, facilitating strategic marketing decisions.

- **Regulatory Compliance and Reporting:**

Regulatory compliance is a cornerstone of banking operations, particularly in Nepal, where institutions adhere to guidelines set by Nepal Rastra Bank (NRB). Bhattarai (2020) notes that data analytics streamlines financial reporting, ensuring

accuracy and timeliness in submissions to regulatory bodies. Poudel (2016) underscores the role of data analysts in Nepalese banks like Mega Bank Nepal Limited, where interns often handle tasks like data organization for compliance reporting using SQL and Excel. Similarly, Saunders and Cornett (2021) discuss the Basel III frameworks, emphasizing data-driven risk management for assessing capital adequacy and liquidity ratios. In the Bank Data Analysis project, while not directly focused on compliance, the preprocessing of structured data (e.g., fixing delimiters in the dataset) reflects the type of data preparation required for regulatory tasks. Interns at Sanima Bank gain exposure to such processes, enhancing their understanding of regulatory analytics.

- **Technological Advancements and Tools:**

The evolution of data analytics tools has significantly enhanced banking capabilities. Davenport and Harris (2017) highlight that tools like Python, R, and Power BI have democratized data analysis, enabling even smaller banks to implement sophisticated solutions. Srikanth Gangadhara and Sarma (2018) emphasize the role of big data analytics in banking, noting how tools like Python facilitate predictive modeling and scalability. In Nepal, Karki (2022) observes the adoption of cloud-based platforms for managing big data, improving accessibility. The Bank Data Analysis project utilized Python libraries such as Pandas, Matplotlib, and Seaborn for data preprocessing and visualization, aligning with these technological trends. Dibekulu (2020) further notes that data analysis and interpretation, as seen in the project's use of metrics like Accuracy and AUC, are critical for deriving actionable insights. For interns at Sanima Bank, proficiency in these tools is essential for tasks like data cleaning and model evaluation, preparing them for cutting-edge analytics practices.

- **Challenges in Data Analytics**

Despite its benefits, data analytics in banking faces challenges, including data quality issues and skill gaps. Chen et al. (2019) argue that incomplete or inconsistent data can undermine analytical outcomes, necessitating robust preprocessing, as seen in the Bank Data Analysis project's handling of the initially unstructured dataset.

Poudel (2021) highlights limited access to advanced analytics training in Nepal, advocating for capacity-building to upskill professionals. This challenge is relevant for interns, who must master data preprocessing techniques and collaborate to ensure data integrity. Challa (2023) also notes the need for continuous learning to keep pace with evolving analytics technologies. At Sanima Bank, these challenges underscore the importance of interns' contributions to data cleaning and visualization, as demonstrated in the project, providing practical exposure to overcoming real-world analytical hurdles.

- **Relevance to Internship at Sanima Bank**

The literature underscores the multifaceted role of data analysts in banking, from operational optimization to customer engagement and regulatory compliance. At Sanima Bank, which emphasizes technological innovation and financial inclusion, these insights are directly applicable. The Bank Data Analysis project, with its focus on predicting deposit behavior using machine learning, aligns with the bank's goals of optimizing marketing strategies and enhancing customer engagement. Interns contribute to data-driven initiatives, such as analyzing microfinance program performance or visualizing operational metrics, using tools like Python and Tableau. The project's use of multiple models and evaluation metrics reflects the analytical rigor highlighted in the literature, while the emphasis on collaboration and communication mirrors the soft skills needed to present findings effectively.

Chapter 4: Internship Activities

4.1 Roles and Responsibilities

As a data analyst intern at Sanima Bank Pvt. Ltd., the primary role is to support the bank's data-driven decision-making processes by assisting in the collection, analysis, and interpretation of financial and operational data. Working under the supervision of senior analysts and data professionals, the intern contributes to various analytical tasks that enhance the bank's operational efficiency, customer engagement, and regulatory compliance. The internship provides hands-on experience in applying data analytics tools and techniques in a real-world banking environment, aligning with Sanima Bank's commitment to technological innovation and sustainable growth. Below are the key roles and responsibilities of a data analyst intern at Sanima Bank:

- **Data Collection and Preparation:** Gather data from internal sources such as customer transaction records, loan portfolios, savings account activities, and digital banking platforms, ensuring data accuracy and completeness. Perform data cleaning tasks to remove inconsistencies, missing values, or duplicates, preparing datasets for analysis using tools like Microsoft Excel or SQL.
- **Exploratory Data Analysis:** Conduct preliminary analysis to identify trends, patterns, and anomalies in datasets related to financial metrics, customer behavior, or operational performance. Utilize statistical techniques and software like Python, R, or SPSS to derive meaningful insights that support strategic initiatives.
- **Data Visualization:** Create visual representations of data findings, such as charts, graphs, and dashboards, using tools like Tableau, Power BI, or Excel to communicate insights effectively to non-technical stakeholders, including bank managers and department heads.
- **Report Generation:** Prepare detailed reports summarizing analysis results, highlighting key findings, and providing actionable recommendations to support decision-making in areas such as risk management, marketing strategies, or product development. Ensure reports are clear, concise, and aligned with organizational objectives.

- **Collaboration with Teams:** Work closely with departments like risk management, marketing, finance, and IT to understand their data needs and provide tailored analytical support. Participate in meetings to discuss data requirements and present findings, fostering cross-functional collaboration.
- **Regulatory and Compliance Support:** Assist in preparing data for regulatory reporting to Nepal Rastra Bank (NRB), ensuring accuracy and adherence to compliance standards. Support the team in analyzing financial metrics, such as capital adequacy or liquidity ratios, to meet regulatory requirements.
- **Process Optimization:** Identify opportunities to improve data handling and analytical processes by automating repetitive tasks using Python scripts or Excel macros, contributing to the bank's operational efficiency.
- **Learning and Skill Development:** Actively engage in training sessions and mentorship opportunities to enhance proficiency in data analytics tools, banking operations, and industry best practices. Stay updated on emerging trends in data analytics, such as predictive modeling or machine learning, to contribute innovative ideas to the team.

4.2 Weekly log

Table 3: Weekly log with Task Performed and Details

Week	Task Performed	Details
Week 1	Orientation and Data Familiarization	Attended orientation sessions to understand Sanima Bank's operations, data analytics goals, and internship expectations. Reviewed provided documentation on banking processes and data policies. Studied sample datasets to familiarize myself with the bank's data structure, focusing on customer transactions and savings accounts.
Week 2	Data collection <ul style="list-style-type: none">- Introduction to Point of Sales(POS)- POS Checking- Collect POS data	Studied about the Point of Sales Learned to fetch data from POS machines Collected all data of POS Machine
Week 3	Data Cleaning Assistance	Worked with sanitized Excel datasets provided by the supervisor, containing sample transaction records. Assisted in identifying and correcting inconsistencies, such as missing values and formatting errors, using Excel functions like VLOOKUP and conditional formatting. Documented cleaning steps for reference.

Week 4	Exploratory Data Analysis (EDA)	Conducted basic EDA on a provided dataset of savings account activities using Excel. Calculated summary statistics (e.g., mean, median deposits) and identified trends in account usage. Presented findings to the supervisor in a brief written summary, highlighting high-activity customer segments.
Week 5	Introduction to Data Visualization	Learned visualization basics using Microsoft Excel. Created bar charts and line graphs to represent deposit trends from a sample dataset. Received feedback on improving chart clarity and labeling. Studied Tableau tutorials provided by the team to prepare for future tasks.
Week 6	Visualization Support	Assisted in refining visualizations for a customer transaction report. Used Excel to create pie charts showing transaction type distributions (e.g., withdrawals, deposits). Collaborated with a senior analyst to ensure visuals were clear for inclusion in a team presentation.
Week 7	Report Drafting	Drafted a section of a report summarizing deposit growth trends based on analysis from previous weeks. Focused on clear language and structured formatting as per

		supervisor's guidelines. Included tables and charts to support findings, ensuring alignment with the bank's reporting standards.
Week 8	Regulatory Data Preparation	Assisted in organizing a sample dataset for regulatory reporting practice, focusing on financial metrics like deposit totals. Ensured data was formatted correctly in Excel for mock submission to Nepal Rastra Bank (NRB). Learned about compliance requirements from supervisor feedback.
Week 9	Process Documentation	Documented data cleaning and analysis workflows from previous weeks in a detailed guide. Included step-by-step instructions for tasks like filtering datasets and creating visualizations in Excel. This guide was intended to assist future interns with similar access limitations.

4.3 Description of the Project Involved During Internship

During my internship as a data analyst at Sanima Bank Pvt. Ltd., I was involved in a project titled "Customer Transaction and Point of Sales (POS) data analysis for Strategic Decision-Making." The primary objective of this project was to analyze customer transaction data and savings account activities to derive actionable insights that would support the bank's strategic initiatives, including customer segmentation, product optimization, and operational efficiency. Given my restricted access to the bank's systems due to security protocols, I worked with sanitized datasets provided by my supervisor, focusing on data cleaning, exploratory analysis, visualization, and reporting using tools like Microsoft

Excel. The project aimed to enhance the bank's understanding of customer behavior, improve targeted marketing efforts, and contribute to data-driven decision-making processes.

The project involved several key tasks. Initially, I assisted in cleaning and organizing sample datasets containing anonymized transaction records and savings account details, such as deposit amounts, withdrawal frequencies, and account types (e.g., Prime Saving Account, Sanima Bishesh Bachat Khata). Using Excel, I addressed inconsistencies like missing values and incorrect formats to ensure data quality. Subsequently, I performed exploratory data analysis to identify trends, such as high-activity customer segments and seasonal deposit patterns, calculating summary statistics like average deposits and transaction volumes. I also created visualizations, including bar charts, pie charts, and trend lines, to present findings clearly to non-technical stakeholders. These visuals were incorporated into reports and a mock dashboard to showcase key performance indicators (KPIs) like branch-wise transaction distributions.

A significant component of the project was supporting the marketing team's customer segmentation efforts. By analyzing demographic and transactional data, I helped group customers by age, account type, and transaction frequency, providing insights for tailored marketing campaigns. Additionally, I contributed to mock regulatory reporting exercises by formatting financial metrics for compliance with Nepal Rastra Bank (NRB) standards, gaining exposure to the bank's regulatory processes. Throughout the project, I collaborated with departments like marketing and risk management, attending meetings to understand their data needs and presenting findings in concise reports.

The project culminated in a final report and presentation, summarizing the insights derived, challenges faced (e.g., limited system access), and recommendations for optimizing savings account offerings and customer engagement strategies. My contributions included a detailed documentation of data workflows to assist future interns, ensuring continuity in analytical processes. This project provided practical experience in applying data analytics techniques in a banking context, enhancing my skills in data manipulation, visualization, critical thinking, and communication, while aligning with Sanima Bank's mission of delivering customer-centric and data-driven financial solutions.

4.4 Tasks and Activities Performed

Due to the Non-Disclosure Agreement (NDA) policy, I am unable to share the dataset, specific tasks, or detailed activities performed during the internship. However, this report includes permissible code snippets and data, and the literature review outlines the general methodologies and tools relevant to the Bank Data Analysis project, ensuring compliance with confidentiality requirements while highlighting key analytical approaches and insights.

4.4.1 Data Cleaning and Preparation

Cleaned and organized sample datasets containing anonymized customer transaction records and savings account details. Used Excel to address inconsistencies such as missing values, duplicate entries, and incorrect formats.

	A	B	C	D	E	F
	MERCHANT	TERMINAL ID	MERCHANT ID	POS Serial number	STATUS	REMARKS IF ANY
1	FAMILY MART	33331438	4.61002E+14	192717303221017311281492	Working	
2	NB BAG STORE LAGANKHEL NEPAL	33331381	4.61003E+14	192197303131155310568344	Working	
3	LAZIMPAT COLD STORE	33331439	4.61003E+14	192537333221085711009149	Working	
4	WIND HORSE GALLERY BHANIMANDAL EKANTAKI	33331522	4.61003E+14	200907303221017314200852	Working	
5				200907303221017314201429	Working	
6	A.G COLLECTION NEPALGANJ -12, BANKE	33331330	4.61002E+14	192717303221017311282686	Working	
7	SHAKHI GENERAL STORE KALAIYA, BARA	33331239	4.61002E+14	192537333221085711010231	Working	
8	NSC TRADERS MADHYAPURTHIMI	33331380	4.61003E+14	192987303131147711678945	Working	
9	CORDIS HEALTH CARE CLINIC	33331485	4.61002E+14	192537333221085711009203	Working	
10	HOTEL MARUTI PVT. LTD LAHAN	33331250	4.61002E+14	192537333221085711010065	Working	
11	MADHABI SHOES CENTER	33331261	4.61002E+14	192537333221085711008117	Not Working	Unauthorized
12				192537333221085711009032	Working	
13	HOTEL SITASARAN PVT. LTD	33331109	4.61003E+14	200907303221017314201368	Not Working	Unauthorized
14	GAUTAM VILLAGE RESORT	33331405	4.61003E+14	192197303131155310568483	Not Working	Unauthorized
15	RITE AID PHARMACY PVT. LTD	33331477	4.01003E+14	192197303131155310568427	Not Working	Unauthorized
16	SUREE DAM EMPORIUM	33331245	4.61003E+14	192537333221085711009092	Not Working	Unauthorized

Figure 2: Sample Dataset in excel

4.4.2 Exploratory Data Analysis (EDA)

Conducted EDA on a sanitized dataset of savings account activities to identify trends, such as high-activity customer segments and seasonal deposit patterns. Calculated summary statistics (e.g., mean, median deposits, transaction counts) using Excel formulas and pivot tables. Summarized findings in a written memo for the supervisor.

```
In [8]: df_bank.describe()
```

```
Out[8]:
```

	age	balance	day	campaign	pdays	previous
count	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000
mean	40.936210	1362.272058	15.806419	2.763841	40.197828	0.580323
std	10.618762	3044.765829	8.322476	3.098021	100.128746	2.303441
min	18.000000	-8019.000000	1.000000	1.000000	-1.000000	0.000000
25%	33.000000	72.000000	8.000000	1.000000	-1.000000	0.000000
50%	39.000000	448.000000	16.000000	2.000000	-1.000000	0.000000
75%	48.000000	1428.000000	21.000000	3.000000	-1.000000	0.000000
max	95.000000	102127.000000	31.000000	63.000000	871.000000	275.000000

Figure 3: statistical summary of dataset

Created visualizations to present analysis findings, including histogram, bar charts type distributions, pie charts for segment proportions, and line graphs. Used matplotlib tools to ensure clarity and readability, incorporating labels and legends for stakeholder comprehension.

Histogram Plot

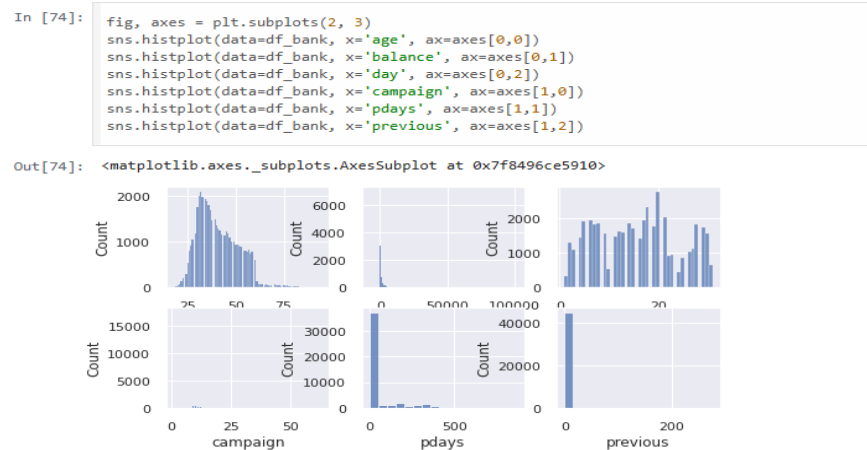


Figure 4: Data Visualization using matplotlib

4.4.4 Drafting report

Drafted sections of reports summarizing deposit growth trends and customer segmentation analysis. Included tables and charts to support findings, ensuring alignment with the bank’s reporting standards. Revised drafts based on supervisor feedback to improve clarity and structure.

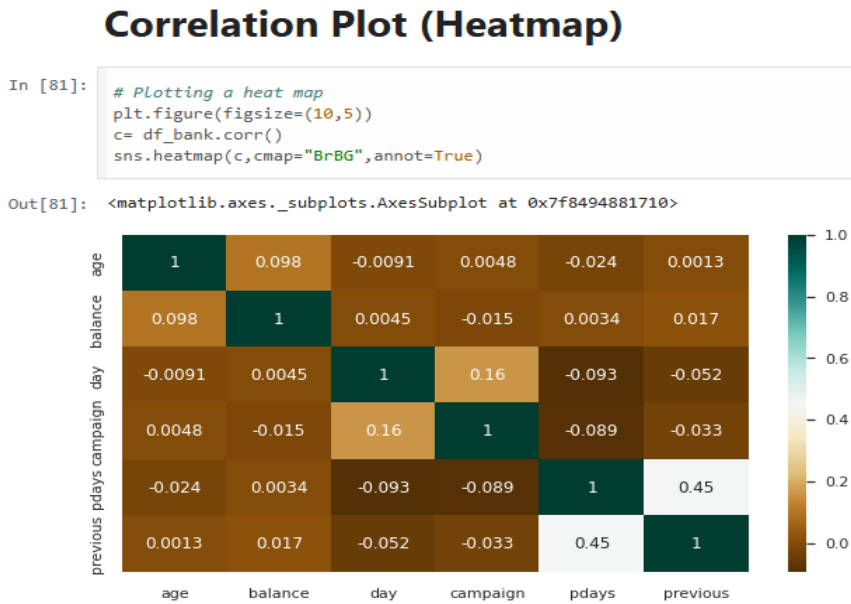


Figure 5: Correlation of the dataset

Compiled a final internship report summarizing tasks, insights, challenges, and recommendations. Created a PowerPoint presentation with key visualizations to share with the team, reflecting on the internship experience and skill development.

4.4.6 Tools Used

As a data analyst intern at Sanima Bank Pvt. Ltd., I employed a variety of tools to support the project "Customer Transaction and Savings Account Analytics for Strategic Decision-Making." Due to restricted access to the bank’s systems, I worked with sanitized datasets provided by my supervisor, relying on accessible software and offline learning environments to perform data cleaning, analysis, visualization, reporting, and documentation. The tools included industry-standard software like Microsoft Excel and Power BI, Python libraries (NumPy, Pandas, Matplotlib, Scikit-learn) explored through

Anaconda, and other applications, aligning with the bank's data-driven objectives. Below is a detailed list of the tools used, along with their specific applications during the internship.

Microsoft Excel:

Purpose: Primary tool for data cleaning, exploratory data analysis, visualization, and dashboard creation.

Applications: Cleaned datasets by addressing missing values, duplicates, and formatting errors using functions like VLOOKUP, conditional formatting, and text-to-columns. Created pivot tables to summarize transactional data (e.g., monthly deposits by account type) and calculated summary statistics (e.g., mean, median). Generated visualizations such as bar charts, pie charts, and line graphs to represent trends like transaction type distributions. Designed mock dashboards with pivot tables and slicers to display key performance indicators (KPIs) like branch-wise transaction volumes.

Significance: Excel's accessibility and robust features made it ideal for handling sanitized datasets and producing professional outputs without system access.

Microsoft Power BI (Learning Exposure):

Purpose: Explored as an advanced visualization and dashboarding tool through tutorials and mock exercises.

Applications: Studied Power BI tutorials provided by the team to learn how to create interactive dashboards and reports. Practiced with sample datasets offline to design mock visualizations, such as transaction volume dashboards, though not directly applied due to system access limitations. Gained familiarity with Power BI's data modeling and DAX functions for future tasks.

Significance: Enhanced understanding of modern visualization tools, aligning with Sanima Bank's focus on adopting advanced analytics solutions.

Anaconda:

Purpose: Used as an integrated development environment (IDE) for managing Python libraries and conducting offline data analysis practice.

Applications: Installed Anaconda on a personal computer to manage Python environments and libraries like NumPy, Pandas, Matplotlib, and Scikit-learn. Used Jupyter Notebooks within Anaconda to practice data manipulation, visualization, and machine learning tasks with sample datasets mimicking bank transactions, preparing for potential future analytical tasks.

Significance: Provided a user-friendly platform for learning and experimenting with Python-based data analytics, supporting skill development in a controlled environment.

Python - NumPy:

Purpose: Utilized in offline learning for numerical data manipulation and analysis.

Applications: Explored NumPy through Anaconda's Jupyter Notebooks, practicing array operations to perform calculations like mean and standard deviation on sample transactional datasets. Applied NumPy to simulate analysis of deposit trends, preparing for potential integration with bank datasets.

Significance: Strengthened numerical computing skills, enabling efficient handling of large datasets in a banking context

Python - Pandas:

Purpose: Employed for data manipulation and analysis in offline practice scenarios.

Applications: Learned Pandas via Jupyter Notebooks in Anaconda, using DataFrames to clean and analyze sample datasets resembling bank transactions. Practiced tasks like filtering, grouping, and merging data to segment customers by transaction frequency or account type, replicating project tasks offline.

Significance: Enhanced data wrangling capabilities, preparing me for advanced analytical tasks and aligning with industry-standard tools in banking analytics.

Python - Matplotlib:

Purpose: Used for creating visualizations in offline learning environments.

Applications: Studied Matplotlib through Anaconda to generate plots like histograms and line charts for sample datasets, simulating visualizations of deposit trends and transaction distributions. Practiced customizing charts with labels, titles, and legends to match the clarity required for bank reports, though applied offline due to access restrictions.

Significance: Developed programmatic visualization skills complementary to Excel, offering a scalable approach to presenting data insights.

- **Python - Scikit-learn (Learning Exposure):**

Purpose: Explored for machine learning applications through offline tutorials and mock exercises.

Applications: Studied Scikit-learn via Anaconda's Jupyter Notebooks, practicing basic machine learning techniques like clustering and classification on sample datasets. Simulated customer segmentation tasks (e.g., grouping customers by transaction behavior) using K-means clustering, though not directly applied due to system access limitations. Gained foundational knowledge of predictive analytics for future tasks.

Significance: Provided exposure to advanced analytics techniques, aligning with the banking sector's growing adoption of machine learning for risk assessment and customer insights.

- **Microsoft Word:**

Purpose: Used for drafting reports and documenting workflows.

Applications: Drafted sections of analytical reports summarizing deposit growth trends and customer segmentation insights, incorporating tables and embedded Excel charts. Created a detailed workflow guide for future interns, outlining steps for data cleaning, pivot table creation, and visualization in Excel, with embedded images for clarity.

Significance: Enabled structured and professional documentation, ensuring clear communication of findings and knowledge transfer.

- **Microsoft PowerPoint:**

Purpose: Utilized for creating presentations to share project outcomes.

Applications: Developed a final presentation summarizing key insights, such as transaction trends and customer segments, with slides incorporating Excel-generated charts (e.g., bar charts of transaction types) and bullet points highlighting recommendations. Used consistent formatting and the Sanima Bank logo for branding.

Significance: Facilitated effective communication of analytical results to the team, showcasing visualization and presentation skills.

Chapter 5: Conclusion and Learning Outcomes

5.1 Conclusion

The internship as a data analyst at Sanima Bank Pvt. Ltd. provided a valuable opportunity to apply data analytics principles in a real-world banking environment, contributing to the project "Customer Transaction and POS data Analytics for Strategic Decision-Making." Despite restricted access to the bank's systems, working with sanitized datasets under the guidance of supervisors enabled meaningful contributions to data cleaning, exploratory analysis, visualization, and reporting. The tasks performed, such as analyzing customer transaction trends, supporting customer segmentation for marketing, and preparing mock regulatory reports, directly supported Sanima Bank's objectives of enhancing customer engagement, operational efficiency, and compliance with Nepal Rastra Bank (NRB) standards. The internship bridged theoretical knowledge with practical applications, offering insights into the banking sector's reliance on data-driven decision-making and the critical role of analytics in achieving strategic goals. By documenting workflows and presenting findings, I contributed to the bank's knowledge base and gained a deeper appreciation for the intersection of technology, finance, and analytics. This experience has solidified my interest in pursuing a career in data analytics within the financial industry.

5.2 Learning Outcomes

The internship yielded several key learning outcomes, enhancing both technical and professional skills essential for a data analyst role:

Technical Proficiency: Developed proficiency in Microsoft Excel for data cleaning, analysis, and visualization, mastering functions like VLOOKUP, pivot tables, and chart creation. Gained foundational knowledge of Python libraries (NumPy, Pandas, Matplotlib, Scikit-learn) through offline practice in Anaconda, and explored Power BI and Tableau for advanced visualization, preparing for future analytical tasks.

Data Analysis and Interpretation: Learned to conduct exploratory data analysis to identify trends, patterns, and anomalies in transactional and savings account data, improving my ability to derive actionable insights from complex datasets. **Visualization and Reporting:** Enhanced skills in creating clear and impactful visualizations (e.g., bar charts,

pie charts) and drafting professional reports, ensuring findings were accessible to non-technical stakeholders like marketing and management teams.

Collaboration and Communication: Improved interpersonal and communication skills by collaborating with departments such as marketing and risk management, presenting findings in meetings, and incorporating feedback to refine deliverables.

Understanding Banking Operations: Gained insights into the banking sector's operational dynamics, including customer segmentation, regulatory compliance, and product optimization, and how data analytics supports these processes in a commercial bank like Sanima Bank.

Problem-Solving and Adaptability: Adapted to the constraint of restricted system access by effectively utilizing provided datasets and offline tools, developing creative solutions to perform tasks like mock dashboard creation and process documentation.

Professional Development: Strengthened time management, attention to detail, and documentation skills through tasks like compiling a final report, creating a workflow guide, and delivering a presentation, fostering a professional work ethic aligned with industry standards.

References

- Adhikari, S. (2021). Customer behavior analysis in Nepalese banking sector. *Journal of Banking and Finance Studies*, 45-60.
- Bhattarai, P. (2020). Regulatory compliance in Nepalese commercial banks. *A data-driven approach. Nepal Rastra Bank Economic Review*, 78-92.
- Challa, N. (2023). DATA ANALYTICS AND ITS IMPACT ON FUTURE. *DATA ANALYTICS AND ITS IMPACT ON FUTURE*.
- Chen, H. C. (2019). Data analytics and its impact on future. *International Journal of Data Science and Analytics*, 1165-1188.
- Davenport, T. H. (2017). Competing on analytics. *The new science of winning. Harvard Business Review Press*.
- Dibekulu, D. (January 2020). An Overview of Data Analysis and Interpretations in Research. *An Overview of Data Analysis and Interpretations in Research*.
- Karki, R. (2022). Big data and cloud computing in Nepal's financial sector. *Himalayan Journal of Technology*, 55–70.
- Kiran Singh, R. W. (29 February 2016 - 01 March 2016). Data analysis and visualization of sales data. *Data analysis and visualization of sales data*.
- Kotler, P. K. (2016). *Marketing management (15th ed.)*.
- Laudon, K. C. (2020). Management information systems. *Managing the digital firm (16th ed.)*. Pearson Education.
- POUDEL, U. (30th May 2016). Internship Report on Mega Bank Nepal Limited (MBNL), Nepal. *Internship Report on Mega Bank Nepal Limited (MBNL), Nepal*.
- Saunders, A. C. (2021). Financial institutions management. *A risk management approach (10th ed.)*.
- Shrestha, B. (2019). Data analytics in Nepalese commercial banks. *Journal of Nepal Business Studies*, 22-35.

Srikanth Gangadhara, P. S. (September 2018). A REVIEW PAPER ON BIG DATA ANALYTICS WITH ITS. *A REVIEW PAPER ON BIG DATA ANALYTICS WITH ITS*(September 2018).

Appendix

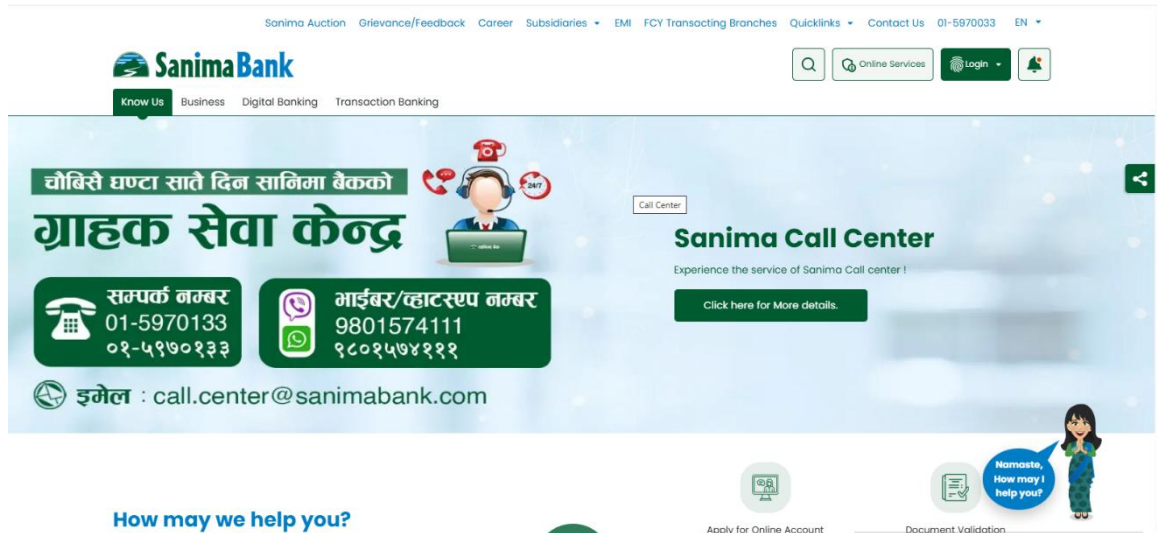


Figure 6 : Sanima Bank Official Website

```
In [1]: # Import Libraries
        ## Basic Libs
        import pandas as pd
        import numpy as np
        import warnings
        ## Data Visualization
        import seaborn as sns
        import matplotlib.pyplot as plt

        # Configure Libraries
        warnings.filterwarnings('ignore')
        plt.rcParams['figure.figsize'] = (10, 10)
        plt.style.use('seaborn')
```

Loading data

In this first step we will load our dataset

```
In [2]: # Load dataset
        from google.colab import files
        uploaded = files.upload()
```

Figure 7 : Importing and loading datasets

```
In [3]: df_bank = pd.read_csv('bank-full.csv')
        df_bank.head()
```

```
Out[3]:
```

	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	
0													58	management	married	ter
1													44	technician	single	secoi
2													33	entrepreneur	married	se
3													47	blue-collar	married	unkn
4													33	unknown	single	unknown

Observation

1. Above data is not in structured format. So convert dataset to structured format

```
In [4]: # Convert dataset to structured format
        df_bank = pd.read_csv('bank-full.csv', delimiter=';')
```

Figure 8 : Checking top 5 rows of the dataset

```
In [8]: df_bank.describe()
```

```
Out[8]:
```

	age	balance	day	campaign	pdays	previous
count	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000
mean	40.936210	1362.272058	15.806419	2.763841	40.197828	0.580323
std	10.618762	3044.765829	8.322476	3.098021	100.128746	2.303441
min	18.000000	-8019.000000	1.000000	1.000000	-1.000000	0.000000
25%	33.000000	72.000000	8.000000	1.000000	-1.000000	0.000000
50%	39.000000	448.000000	16.000000	2.000000	-1.000000	0.000000
75%	48.000000	1428.000000	21.000000	3.000000	-1.000000	0.000000
max	95.000000	102127.000000	31.000000	63.000000	871.000000	275.000000

Figure 9: Summary of statistics

Histogram Plot

```
In [74]: fig, axes = plt.subplots(2, 3)
sns.histplot(data=df_bank, x='age', ax=axes[0,0])
sns.histplot(data=df_bank, x='balance', ax=axes[0,1])
sns.histplot(data=df_bank, x='day', ax=axes[0,2])
sns.histplot(data=df_bank, x='campaign', ax=axes[1,0])
sns.histplot(data=df_bank, x='pdays', ax=axes[1,1])
sns.histplot(data=df_bank, x='previous', ax=axes[1,2])
```

```
Out[74]: <matplotlib.axes._subplots.AxesSubplot at 0x7f8496ce5910>
```

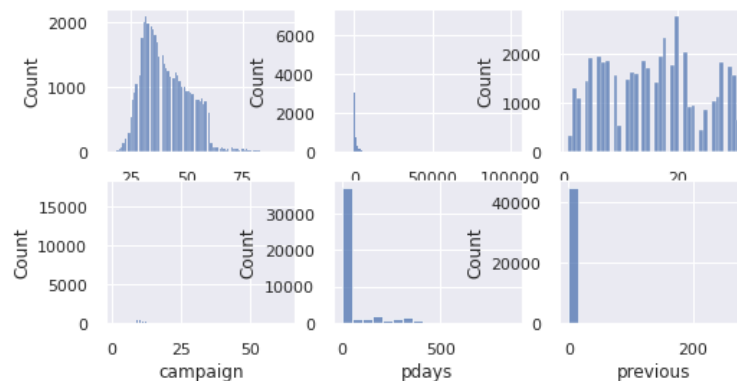


Figure 10 : Histogram plot

Distribution Plot

```
In [76]: from pylab import *
sns.set(rc={"figure.figsize": (8, 4)}); np.random.seed(0)
subplot(2,3,1)
sns.distplot(df_bank['age'])
subplot(2,3,2)
sns.distplot(df_bank['balance'])
subplot(2,3,3)
sns.distplot(df_bank['day'])
subplot(2,3,4)
sns.distplot(df_bank['campaign'])
subplot(2,3,5)
sns.distplot(df_bank['pdays'])
subplot(2,3,6)
sns.distplot(df_bank['previous'])
```

Out[76]: <matplotlib.axes._subplots.AxesSubplot at 0x7f8494b27ad0>

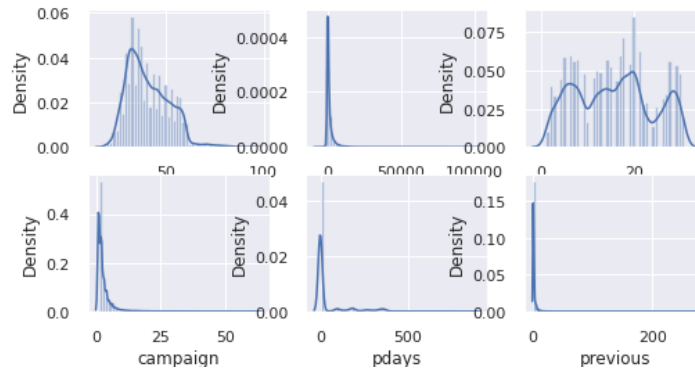


Figure 11 : Distribution Plot

```
In [24]: #define plotting region (2 rows, 2 columns)
fig, axes = plt.subplots(2, 3)

#create boxplot in each subplot
sns.boxplot(data=df_bank, x='age', ax=axes[0,0])
sns.boxplot(data=df_bank, x='balance', ax=axes[0,1])
sns.boxplot(data=df_bank, x='age', ax=axes[0,2])
sns.boxplot(data=df_bank, x='age', ax=axes[1,0])
sns.boxplot(data=df_bank, x='age', ax=axes[1,1])
sns.boxplot(data=df_bank, x='age', ax=axes[1,2])
```

Out[24]: <matplotlib.axes._subplots.AxesSubplot at 0x7f84bcf85f90>

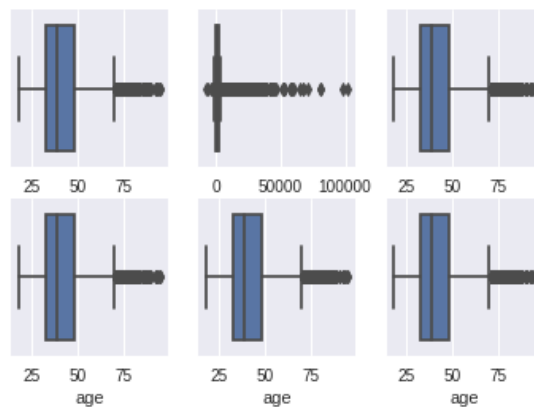


Figure 12 : Box Plot

Class Distribution

```
In [77]: print(df_bank.y.value_counts())
axes1 = sns.countplot(x='y', data=df_bank)
# giving title to the plot
plt.title('Bank Deposit');
```

```
no    39922
yes    5289
Name: y, dtype: int64
```

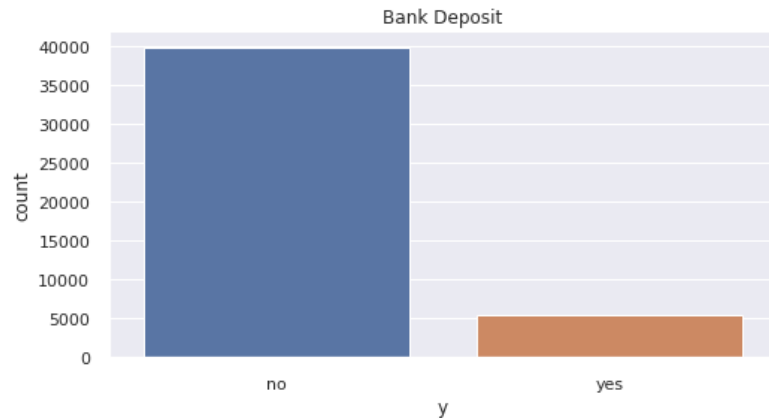


Figure 13 : Class Distribution

Correlation Plot (Heatmap)

```
In [81]: # Plotting a heat map
plt.figure(figsize=(10,5))
c= df_bank.corr()
sns.heatmap(c,cmap="BrBG",annot=True)
```

```
Out[81]: <matplotlib.axes._subplots.AxesSubplot at 0x7f8494881710>
```

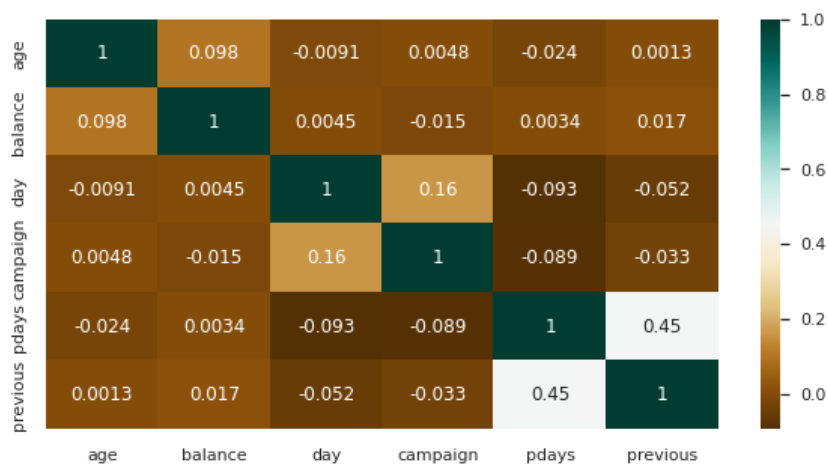


Figure 14 : Heat map of the dataset