

Data Analytics
AN INTERNSHIP REPORT

Submitted by

Swar S. Parikh

200050131124

In Partial fulfillment for the award of the degree Of

BACHELOR OF ENGINEERING

in

Department of Computer Science &

Engineering Babaria Institute of Technology,

Vadodara



Gujarat Technological University,
Ahmedabad

Academic year (2023-2024)



Babaria Institute of Technology

Vadodara-Mumbai National Highway

#8, Varnama, Vadodara,

Gujarat 391240

CERTIFICATE

Date: 01/11/2023

*This is to certify that the Internship Work entitled “**Data Analytics**” has been carried out by **Swar Sanjaykumar Parikh** under my guidance in fulfillment of the degree of Bachelor of Engineering in Information Technology (8th Semester) of Gujarat Technological University, Ahmedabad during the academic year 2023-24.*

Professor HENIL D. SUTHAR

(Internal Guide)

Dr. Rocky S. Upadhyay

(Head of the Department)

OFFER LETTER



V-Ex Tech Solution
Software Solution

301,Dhun Complex,Nizampura
Above Riya Bridal Studio,
vadodara-390002

Internship Offer Letter

Dear Swar Parikh
we are pleased to offer you internship at V-Ex Tech Solution.
we feel that your skills and background will be valuable assets
to our team.

As per discussion,the position is Data Analysis Intern,Your
Starting Date Will Be 16-01-2024,Time Period Will Be 3Months
as per college.

we look forward to welcoming as a new Intern at V-Ex Tech
Solution,we wish you a Successful Carrer With us.



HIMANSHU
Director & CEO.
Director & CEO

v-extechsolution.in

CERTIFICATE OF COMPLETION



301, 3rd Floor, Dhun Complex,
A-22 Santoknagar Society,
Old Chhani Road, Nava Yard,
Vadodara-390024, GUJARAT
96647 68292

INTERNSHIP COMPLETION LETTER

Date:- 16-03-2024

To,

Head of Department,
Babaria Institute of Technology,
Varnama,
Vadodara. GUJARAT

As per the requirement from the college regarding the Internship, we have provided an internship to:-

Name	<i>Swar Parikh</i>
Enr. No.	<i>200050131124</i>
Student of	<i>Bachelor in Computer Science and Engineering</i>
College	<i>Babaria Institute of Technology, Varnama,</i>
Place	<i>Vadodara. GUJARAT</i>
From	<i>16-01-2024</i>
To	<i>16-04-2024</i>

This is to certify that Swar Parikh, a student of Babaria Institute of Technology is pursuing his internship in the field of Data Analysis from 16th January 2024 to be ongoing under the guidance of Mr. Veer.

During the internship, he Learned SQL, Python, Excel, Power BI, Advance Excel, & many more related concepts which require in Data Analysis.

Additionally, he is contributing to the construction of secure authentication mechanisms and the development of efficient payment systems.

During the period of his internship program with us, he had been exposed to different processes and was found diligent, hardworking, and inquisitive.

We wish him all the best in his future endeavor.



Authorized Signature

ACKNOWLEDGEMENT

First of all, I am grateful to the Almighty, ho graces us, and without the blessing from whom, we cannot think of breathing to learn.

This is My Internship Training Report after Completion of Internship Training at **V-EX Tech Solution**. I wish to express My Sincere Thanks to **Prof. Rocky S. Upadhyay**, Head of the Department of Information Technology. And also, to **Prof. HENIL D. SUTHAR** the Supervisor for the Internship Training.

I express My Gratitude to **V-EX Tech Solution** and **Mr. Himanshu Aggarwal Sir**, HR Manager, for allowing me to work with them and make the best out of My Internship.

I Especially Thank My Supervisor, **Mr. Veer Agraval Sir**, for constantly guiding and supporting me throughout the Training. My Heartfelt Gratitude also goes out to all the staff and employees **V-EX Tech Solution** for Cooperating with me and guiding me throughout My Internship.

Last but not least Important, I take this Opportunity to Thank My Parents and Friends who have been with me and offered emotional Strength and Moral Support.

ABSTRACT

This Internship Report outlines the activities, achievements, and experiences gained during a Data Analytics Internship. The Report begins with an Introduction to the Company and the Internship Program. The Primary focus of the Internship was to apply Data Analytics Techniques to solve Business Problems and Provide Insights to Stakeholders. The Report highlights the key activities undertaken during the Internship, such as Data Collection, Data Cleaning, Data Analysis, and Data Visualization. The Report also covers the Techniques used for Data Analysis, including Statistical Analysis, Machine Learning, and Predictive Modeling. Additionally, the report discusses the tools and technologies used, such as Python, SQL, Excel, and Power BI. The Report concludes with a Summary of the achievements and lessons learned during the Internship, including the Development of Technical Skills, Communication Skills, and Problem-Solving Abilities. Overall, the Internship provided Valuable Experience and Exposure to Real-World Data Analytics Projects, and the Skills developed are Transferable to a wide range of Industries and Roles.

LIST OF FIGURES

Sr. No.	Particulars	Page No.
1	Fig 1.1 V-EX Tech Solution	2
2	Fig 5.1.1. E-Commerce Sales Dash Board	10
3	Fig 5.1.2 IPL Analysis Dash Board	11
4	Fig 5.2.1 E-Commerce Excel Measures-1	12
5	Fig 5.2.2 E-Commerce Excel Measures-2	13
6	Fig 5.2.3 E-Commerce Excel Measures-3	14
7	Fig 5.2.2.1 IPL Analysis Excel Measures-1	15
8	Fig 5.2.2.2 IPL Analysis Excel Measures-2	16
9	Fig 5.2.2.3 IPL Analysis Excel Measures-3	17
10	Fig 5.3.1.1 SQL Project on E-Commerce Task 1	18
11	Fig 5.3.1.2 SQL Project on E-Commerce Task 2	19
12	Fig 5.3.1.3 SQL Project on E-Commerce Task 3	20
13	Fig 5.3.2.1 SQL Project on IPL Task 1	21
14	Fig 5.3.2.2 SQL Project on IPL Task 2	22
15	Fig 5.3.2.3 SQL Project on IPL Task 3	23
16	Fig 5.3.2.4 IPL SQL Project on IPL Task 4	24
17	Fig 5.3.2.5 IPL SQL Project on IPL Task 5	25
18	Fig 5.3.2.6 SQL Project on IPL Task 6	26

INDEX

CERTIFICATE OF COMPLETION	I
ACKNOWLEDGEMENT	II
ABSTRACT.....	III
LIST OF FIGURES.....	IV
Chapter 1. Introduction	1
1.1 Introduction to company	2
1.1.1 Organization Products/ Services	2
1.1.2 Contact.....	2
Chapter 2. Overview of Different Department in Company.....	7
2.1 Python Programming... ..	8
2.2 Web Designing and Development	9
2.3 PHP Laravel.....	9
2.4 React JS and Node JS	9
2.5 Full Stack Developers	9
Chapter 3. Introduction to Project/Internship	8
3.1 Internship Summary	9
3.2 Objectives	9
3.3 Scope	9
3.4 Technology Used.....	10
3.5 Internship Planning.....	10
Chapter 4. System Analysis	8
4.1 Study of Current Tools.....	9
4.2 Problem and Weakness of Current Tools.....	9
Chapter 5. Project Screenshot	8
5.1 Power BI Project	9
5.2 Advanced Excel Project	9
5.3 Python Project	9
5.4 SQL Project.....	9

Chapter 6. Implementation.....	8
6.1 Process/Program/Technology/Modules Specification	9
6.2 Finding/Result/Outcomes	9
Chapter 7. Conclusion and Discussion.....	8
6.1 Overall Analysis of Internship Viabilities.....	9
6.2 Limitation and Future Enhancement	9
Appendices	8
References.....	8

Chapter 1. Introduction

INTRODUCTION

1.1 Introduction to Company



[Figure 1.1 V-EX Tech Solution]

V-Ex Tech is an ISO Certified software consulting & service Company. V-Ex tech is Having Strong Experience of 16+ Years in designing software & create dynamic web pages, creating admin panel with back-end.

It is a part of V-Ex tech, which is in existence since 2001. it is having software company in Vadodara (Gujarat).

They have Specialties on Time Punctuality, Easy to Use, Best Management, Good Concept, Web Development, Front-End / Back-End and Full Stack and Data Analytics.

V-Ex Tech is delivering software solutions across industry verticals like banking, finance, spanning from large multinational corporation to small, medium & large enterprises located in USA, Canada, UK, Europe, Africa and Australia.

1.1.1 Organization Products/ Services

- Data Analytics
- Python Programming
- Web Design and Development
- PHP Laravel, React JS or Node JS Development
- Full-Stack Development

1.1.2 Contact:

- **Website:** <https://v-extechsolution.in/>
- **Email:** email-veeragraval@v-extechsolution.in
- **Address:** 301, Third Floor, Dhun Complex Complex, Above Riya Bridal Studio, Nizampura, Vadodara, Gujarat 390002

Chapter 2. Overview of Different Department in Company

2.1 Python Programming

V-EX Tech Solutions is a software development company that provides Python programming services to clients worldwide. The company has a team of experienced Python developers who specialize in building high-quality, scalable, and robust software solutions.

V-EX Tech Solutions offers a range of Python programming services, including: Web Development, Data Science, Automation, Custom Software Development, Support and Maintenance.

V-EX Tech Solutions has a proven track record of delivering high-quality Python-based solutions to clients in various industries, including finance, healthcare, education, and e-commerce. Their team of Python developers is skilled in the latest technologies and frameworks, ensuring that clients receive solutions that are up-to-date and efficient.

2.2 Web Designing and Development

V-EX Tech Solutions is a web design and development service provider, offering a range of services to help businesses establish their online presence. With a team of skilled designers, developers, and digital marketers, the company provides customized solutions to meet the unique needs of each client.

The company's web development services include front-end and back-end development, e-commerce development, and CMS development. They have experience working with a variety of content management systems, such as WordPress, Drupal, and Magento, and can create custom solutions to meet the unique needs of each client.

Overall, the company is a reliable and experienced web design and development service provider, offering a range of services to help businesses establish a strong online presence and achieve their digital marketing goals.

2.3 PHP Laravel

V-EX Tech Solution is a leading software development company that specializes in providing PHP Laravel services to clients around the world. With a team of experienced developers, designers, and project managers, V-EX Tech Solution is committed to delivering high-quality, cost-effective solutions that meet the unique needs of each client.

V-EX Tech Solution offers a wide range of PHP Laravel services, including custom web application development, e-commerce development, CMS development, and API development. They use the latest technologies and best practices to create scalable and secure solutions that are tailored to each client's specific requirements.

Overall, V-EX Tech Solution is a reliable and experienced provider of PHP Laravel services that delivers high-quality, cost-effective solutions to clients around the world. Their expertise with PHP Laravel and other related technologies, combined with their commitment to customer satisfaction, makes them a trusted partner for any organization looking to develop a custom web application or e-commerce solution.

2.4 React JS and Node JS

V-EX Tech Solution is a company that provides ReactJS and NodeJS services is likely a web development company that specializes in building web applications using these technologies. ReactJS is a JavaScript library that is used for building user interfaces, while NodeJS is a runtime environment for executing JavaScript code outside of a web browser

Overall, V-EX Tech Solution may have a team of experienced developers who work closely with clients to understand their requirements and deliver high-quality projects on time and within budget.

2.5 Full-Stack Developers

V-EX Tech Solution is a leading provider of full stack developer services, offering clients end-to-end development solutions for web and mobile applications. The company is headquartered in Vadodara, with a team of experienced developers and designers who specialize in building custom applications using the latest technologies and frameworks.

The company offers a wide range of full stack development services, including front-end development, back-end development, database design, API development, and quality assurance. Their team of full stack developers are experienced in a range of programming languages such as JavaScript, Python, Ruby on Rails, and more.

Overall, V-EX Tech Solution is a reliable and experienced provider of full stack developer services, offering clients end-to-end solutions for web and mobile application development. Their team of developers are skilled in a range of programming languages and frameworks, and are dedicated to delivering high-quality, scalable, and secure applications that meet the needs of their clients.

Chapter 3. Introduction to Project/Internship

3.1 Internship Summary

Data analytics involves using statistical methods and tools to analyze data and extract insights.

The process involves collecting, cleaning, and transforming data, as well as modeling and visualizing data to identify patterns and trends.

The goal of data analytics is to provide actionable insights that can inform decision-making and drive business outcomes.

Data analytics is used across a range of industries and applications, including finance, healthcare, marketing, and more.

The field is constantly evolving, with new technologies and techniques emerging to help analysts work more efficiently and effectively.

3.2 Objectives

Identifying patterns and trends: Data analytics can help identify patterns and trends in large data sets that might not be apparent through traditional analysis methods.

Improving decision-making: Data analytics can help organizations make more informed decisions based on data-driven insights, rather than relying on gut instincts or intuition.

Optimizing operations: Data analytics can help organizations optimize their operations by identifying inefficiencies, streamlining processes, and improving overall performance.

Enhancing customer experiences: Data analytics can help organizations understand their customers better, identify their needs and preferences, and provide personalized experiences that drive loyalty and satisfaction.

Innovating products and services: Data analytics can help organizations identify new opportunities for innovation, by uncovering unmet customer needs or identifying emerging trends.

3.3 Scope

The scope of data analytics is constantly expanding, as more organizations realize the value of data-driven decision-making. With the growth of big data and advancements in technology, the opportunities for data analytics are only expected to increase in the future.

- **Business:** Data analytics is used in business to analyze customer behavior, optimize marketing strategies, identify new revenue opportunities, and improve operational efficiency.
- **Healthcare:** Data analytics is used in healthcare to improve patient outcomes, identify disease trends, and reduce healthcare costs.
- **Finance:** Data analytics is used in finance for fraud detection, risk assessment, credit scoring, and investment analysis.
- **Education:** Data analytics is used in education to improve student performance, identify learning gaps, and measure the effectiveness of educational programs.
- **Government:** Data analytics is used in government to improve public safety, optimize resource allocation, and measure the impact of government policies.

3.3 Technology Used

3.3.1 Python

Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation via the off-side rule.

Python is a popular high-level programming language used for a wide variety of applications, including Web Development, Data Analysis, Artificial Intelligence, and Scientific Computing. It was First released in 1991 and has since become one of the most widely used programming languages in the world.

Python is known for its simple, clear syntax that is easy to read and write. It uses indentation to indicate blocks of code, which makes it highly readable and helps maintain consistency in code formatting

Python is dynamically typed and garbage-collected. It supports multiple programming paradigms, including structured (particularly procedural), object-oriented and functional programming. It is often described as a "batteries included" language due to its comprehensive standard library.

Guido van Rossum began working on Python in the late 1980s as a successor to the ABC programming language and first released it in 1991 as Python 0.9.0. Python 2.0 was released in 2000. Python 3.0, released in 2008, was a major revision not completely backward-compatible with earlier versions.

Python 2.7.18, released in 2020, was the last release of Python 2.

Python consistently ranks as one of the most popular programming languages.

3.3.2 Advanced Excel

Microsoft Excel is a spreadsheet application developed by Microsoft Corporation that allows users to organize, analyze, and manipulate data using rows and columns of cells. Excel is a powerful tool for creating and managing complex spreadsheets, as well as for performing calculations and generating graphs and charts.

Excel is used in a variety of industries and professions, including finance, accounting, engineering, and data analysis. It offers a range of features and tools that can help users to automate tasks, create custom formulas, and analyze large sets of data quickly and easily.

Some of the key features of Excel include the ability to create and format spreadsheets, perform calculations and functions, use charts and graphs to visualize data, filter and sort data, use conditional formatting to highlight important information, and collaborate with others in real-time.

Excel can be used for a wide range of tasks, including creating budgets, tracking expenses, managing inventory, analyzing sales data, and creating financial statements. Its flexibility and versatility make it a valuable tool for individuals and businesses alike.

3.3.3 SQL

MySQL is an open-source relational database management system (RDBMS). Its name is a combination of "My", the name of co-founder Michael Widenius's daughter, and "SQL", the abbreviation for Structured Query Language.

A relational database organizes data into one or more data tables in which data may be related to each other; these relations help structure the data. SQL is a language programmers use to create, modify and extract data from the relational database, as well as control user access to the database.

In addition to relational databases and SQL, an RDBMS like MySQL works with an operating system to implement a relational database in a computer's storage system, manages users, allows for network access and facilitates testing database integrity and creation of backups.

and facilitates testing database integrity and creation of backups.

3.3.4 Power BI

Power BI is a business intelligence and data visualization tool developed by Microsoft. It allows users to connect, analyze, and visualize data from various sources, including Excel spreadsheets, databases, cloud-based and on-premises data sources, and third-party applications.

With Power BI, users can create interactive dashboards, reports, and data visualizations that enable data-driven insights and decision-making. The platform provides a range of features, including data modeling, data preparation, custom calculations, and advanced analytics.

Power BI also offers a wide range of connectors to various data sources, including Azure, Salesforce, Google Analytics, and many more. The platform also supports collaboration and sharing, allowing users to share their dashboards and reports with colleagues or external stakeholders in real-time.

Power BI is available in different versions, including Power BI Desktop (for creating reports and data models), Power BI Service (for sharing and collaborating on reports and dashboards), and Power BI Mobile (for accessing reports and dashboards on mobile devices).

3.4 Internship Planning

3.4.1 Internship Development Approach and Justification

- Before The Starting of Internship, Interns should have the knowledge about His/Her's Choosed Technology.
- A Successful Internship requires not only a good deal of effort on the side of The Intern, but managers and supervisors most also put in some work to ensure that the Intern gets a meaningful experience.
- The Key is to Accept Responsibility when things don't go as planned. taking ownership and articulating possible solutions will result in faster resolution and enable others to see us as a Leader.
- On behalf of students, we are always working with them on things they can do to become a better

3.4.2 Internship Effort and Time

- As an Intern, I have gained the experience of Live Project & How to work in Hectic Situation.
- I conduct Myself in a professional manner at all times and this is the role and responsibility of an Intern.
- Develop a respectful and co-operative relationship with the company mentors and the other interns at the working place.
- Always be punctual to work and always behave in an ethical manner.
- Completed My Internship in Data Analytics in 12 weeks of time period.

Chapter 4. System Analysis

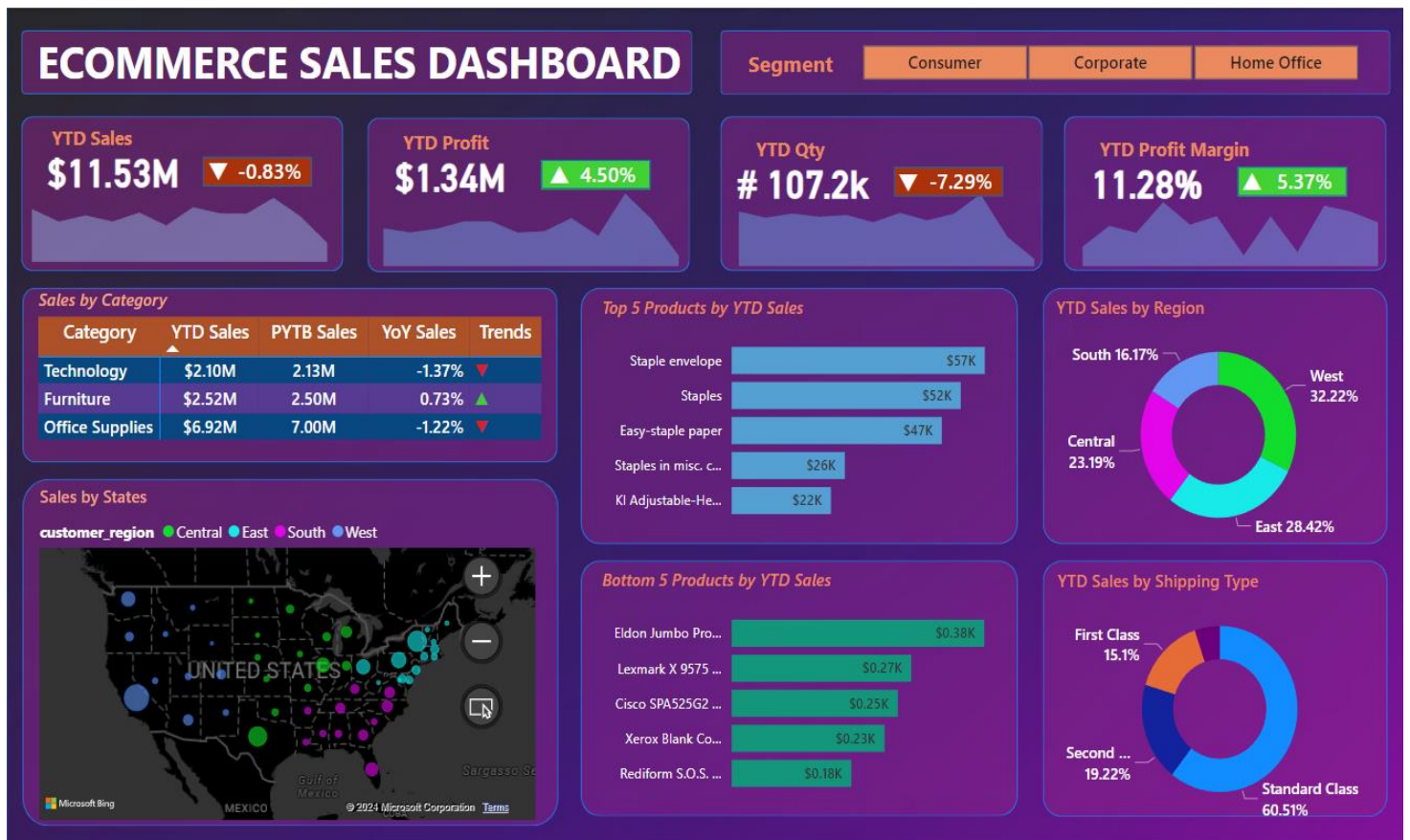
4.1 Study of Current Tools

- Using Power BI, SQL, Advanced Excel and Python to Create a project on IPL, Road Accident, Music Store Report, Diwali Sales Report, and E-Commerce Sales Report.
- Using Different tools, I Create a Dash board on Meet-Ecommerce Sales, Road Accident and IPL Analysis.
- Also using Python to Create a Exploratory Analysis Report on Diwali Sales Data and IPL 2008-2020 Data.
- Using SQL, we store Music Store Database and Solve the insights and also Store the IPL Database and Solve the insights using Query.
- Also Created a Excel report on Road Accident and Shri Ram Store Report.

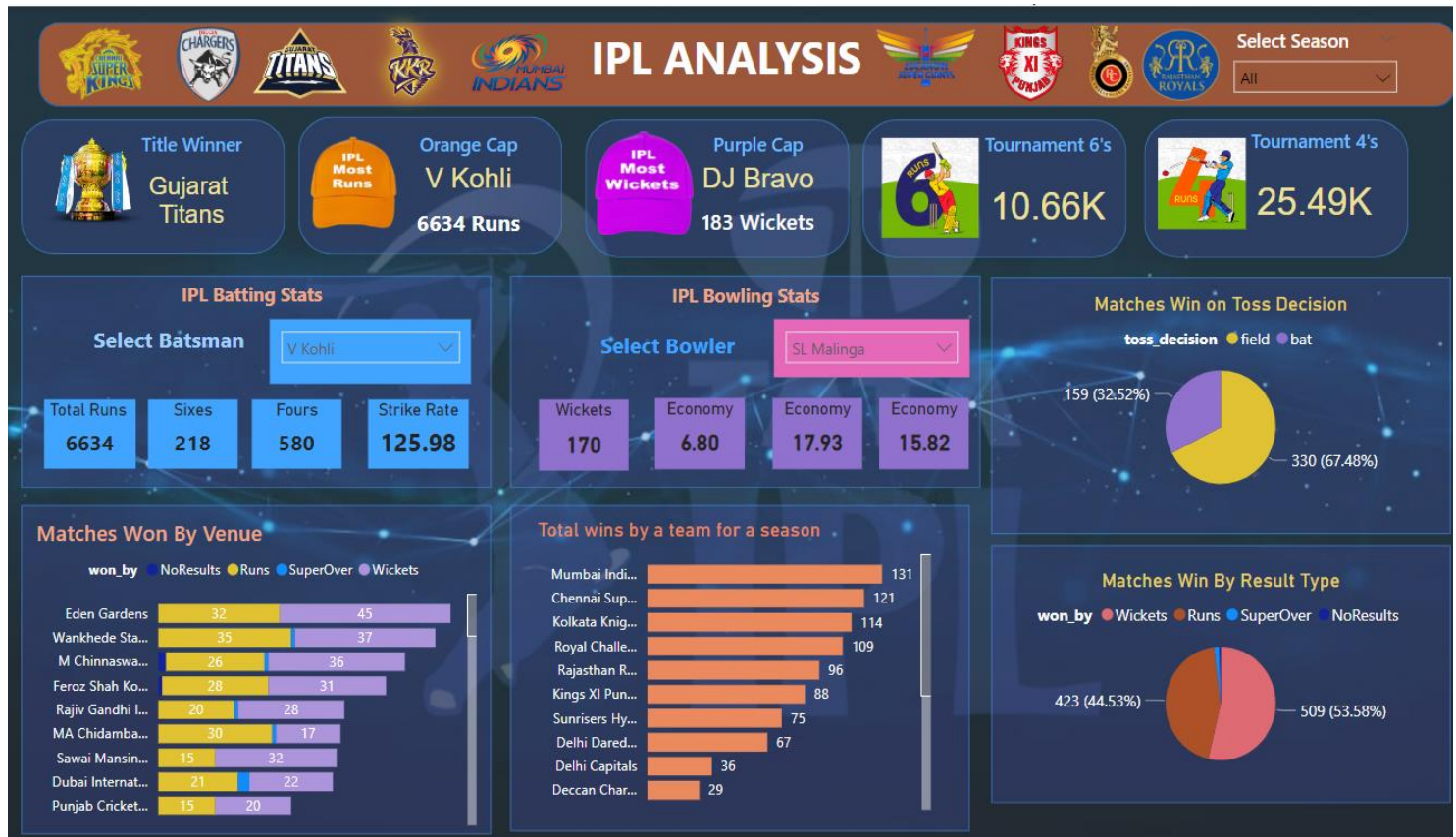
4.2 Problem and Weakness of Current Tools

- Limited exposure to real-world data.
- Lack of mentorship and guidance.
- Limited access to tools a resource.
- Insufficient training in data privacy and security.
- Lack of diversity and inclusion.
- Limited feedback and evaluation.

Chapter 5. Project Screenshot



[Fig 5.1.1. E-Commerce Sales Dash Board]



[Fig 5.1.2 IPL Analysis Dash Board]

1 PYTB Sales = CALCULATE(SUM(ecommerce_data[sales_per_order]),DATESYTD(SAMEPERIODLASTYEAR(Calendar[Date])))							
customer_id	customer_first_name	customer_last_name	Category	product_name	customer_segment	customer_city	customer_state
C_ID_29162	Mary	Jacobs	Office Supplies	Recycled Eldon Regeneration Jumbo File	Consumer	Los Angeles	California
C_ID_53661	Mary	Miller	Office Supplies	Portable Personal File Box	Consumer	Los Angeles	California
C_ID_66378	Mary	Bell	Office Supplies	Acco 6 Outlet Guardian Premium Plus Surge Suppressor	Consumer	Los Angeles	California
C_ID_29254	Mary	Olsen	Office Supplies	Ibico Recycled Linen-Style Covers	Consumer	Los Angeles	California
C_ID_33239	Mary	Swindell	Office Supplies	Fellowes Superior 10 Outlet Split Surge Protector	Consumer	Los Angeles	California
C_ID_42136	Mary	Cousins	Office Supplies	Accohide Poly Flexible Ring Binders	Consumer	Los Angeles	California
C_ID_43861	Mary	Lee	Office Supplies	TOPS Voice Message Log Book, Flash Format	Consumer	Los Angeles	California
C_ID_39004	Mary	Kastensmidt	Office Supplies	Strathmore Photo Frame Cards	Consumer	Los Angeles	California
C_ID_31297	Mary	Trafton	Office Supplies	Avery 52	Consumer	Los Angeles	California
C_ID_57963	Mary	DeCherney	Office Supplies	Dixon Ticonderoga Core-Lock Colored Pencils	Consumer	Los Angeles	California
C_ID_50743	Mary	Phan	Office Supplies	BIC Brite Liner Highlighters	Consumer	Los Angeles	California
C_ID_43653	Mary	Collins	Office Supplies	Xerox 1971	Consumer	Los Angeles	California
C_ID_52308	Mary	Coram	Office Supplies	Xerox 1893	Consumer	Los Angeles	California
C_ID_46373	Mary	Swindell	Office Supplies	Avery Reinforcements for Hole-Punch Pages	Consumer	Los Angeles	California
C_ID_49806	Mary	Triggs	Office Supplies	Tuf-Vin Binders	Consumer	Los Angeles	California
C_ID_74201	Mary	Hansen	Office Supplies	Staples	Consumer	Los Angeles	California
C_ID_55256	Mary	Ducich	Office Supplies	Stur-D-Stor Shelving, Vertical 5-Shelf: 72"H x 36"W x 18 1/2"D	Consumer	Los Angeles	California
C_ID_47712	Mary	Vittorini	Office Supplies	Universal Ultra Bright White Copier/Laser Paper, 8 1/2" x 11", Ream	Consumer	Los Angeles	California
C_ID_28782	Mary	Drucker	Office Supplies	SAFCO Commercial Wire Shelving, Black	Consumer	Los Angeles	California
C_ID_28548	Mary	Blackwell	Office Supplies	Belkin Premiere Surge Master II 8-outlet surge protector	Consumer	Los Angeles	California
C_ID_48943	Mary	Triggs	Office Supplies	Hoover Commercial SteamVac	Consumer	Los Angeles	California
C_ID_32531	Mary	Thornton	Office Supplies	Xerox 1899	Consumer	Los Angeles	California
C_ID_36871	Mary	Stewart	Office Supplies	Gould Plastics 18-Pocket Panel Bin, 34w x 5-1/4d x 20-1/2h	Consumer	Los Angeles	California
C_ID_34578	Mary	Glocke	Office Supplies	Xerox 1889	Consumer	Los Angeles	California
C_ID_66325	Mary	Beltran	Office Supplies	Newell 319	Consumer	Los Angeles	California
C_ID_53233	Mary	Bell	Office Supplies	Avery Premier Heavy-Duty Binder with Round Locking Rings	Consumer	Los Angeles	California
C_ID_33316	Mary	Swindell	Office Supplies	Turnerise Lead Holder with Pocket Clin	Consumer	Los Angeles	California

[Fig 5.2.1 E-Commerce Excel Measures-1]

1 YoY Sales = ([YTD Sales]-[PYTB Sales])/[PYTB Sales]								
customer_id	customer_first_name	customer_last_name	Category	product_name	customer_segment	customer_city	customer_state	
C_ID_29162	Mary	Jacobs	Office Supplies	Recycled Eldon Regeneration Jumbo File	Consumer	Los Angeles	California	
C_ID_53661	Mary	Miller	Office Supplies	Portable Personal File Box	Consumer	Los Angeles	California	
C_ID_66378	Mary	Bell	Office Supplies	Acco 6 Outlet Guardian Premium Plus Surge Suppressor	Consumer	Los Angeles	California	
C_ID_29254	Mary	Olsen	Office Supplies	Ibico Recycled Linen-Style Covers	Consumer	Los Angeles	California	
C_ID_33239	Mary	Swindell	Office Supplies	Fellowes Superior 10 Outlet Split Surge Protector	Consumer	Los Angeles	California	
C_ID_42136	Mary	Cousins	Office Supplies	Accohide Poly Flexible Ring Binders	Consumer	Los Angeles	California	
C_ID_43861	Mary	Lee	Office Supplies	TOPS Voice Message Log Book, Flash Format	Consumer	Los Angeles	California	
C_ID_39004	Mary	Kastensmidt	Office Supplies	Strathmore Photo Frame Cards	Consumer	Los Angeles	California	
C_ID_31297	Mary	Trafton	Office Supplies	Avery 52	Consumer	Los Angeles	California	
C_ID_57963	Mary	DeCherney	Office Supplies	Dixon Ticonderoga Core-Lock Colored Pencils	Consumer	Los Angeles	California	
C_ID_50743	Mary	Phan	Office Supplies	BIC Brite Liner Highlighters	Consumer	Los Angeles	California	
C_ID_43653	Mary	Collins	Office Supplies	Xerox 1971	Consumer	Los Angeles	California	
C_ID_52308	Mary	Coram	Office Supplies	Xerox 1893	Consumer	Los Angeles	California	
C_ID_46373	Mary	Swindell	Office Supplies	Avery Reinforcements for Hole-Punch Pages	Consumer	Los Angeles	California	
C_ID_49806	Mary	Triggs	Office Supplies	Tuf-Vin Binders	Consumer	Los Angeles	California	
C_ID_74201	Mary	Hansen	Office Supplies	Staples	Consumer	Los Angeles	California	
C_ID_55256	Mary	Ducich	Office Supplies	Stur-D-Stor Shelving, Vertical 5-Shelf: 72"H x 36"W x 18 1/2"D	Consumer	Los Angeles	California	
C_ID_47712	Mary	Vittorini	Office Supplies	Universal Ultra Bright White Copier/Laser Paper, 8 1/2" x 11", Ream	Consumer	Los Angeles	California	
C_ID_28782	Mary	Drucker	Office Supplies	SAFCO Commercial Wire Shelving, Black	Consumer	Los Angeles	California	
C_ID_28548	Mary	Blackwell	Office Supplies	Belkin Premiere Surge Master II 8-outlet surge protector	Consumer	Los Angeles	California	
C_ID_48943	Mary	Triggs	Office Supplies	Hoover Commercial SteamVac	Consumer	Los Angeles	California	
C_ID_32531	Mary	Thornton	Office Supplies	Xerox 1899	Consumer	Los Angeles	California	
C_ID_36871	Mary	Stewart	Office Supplies	Gould Plastics 18-Pocket Panel Bin, 34w x 5-1/4d x 20-1/2h	Consumer	Los Angeles	California	
C_ID_34578	Mary	Glocke	Office Supplies	Xerox 1889	Consumer	Los Angeles	California	
C_ID_66325	Mary	Beltran	Office Supplies	Newell 319	Consumer	Los Angeles	California	
C_ID_53233	Mary	Bell	Office Supplies	Avery Premier Heavy-Duty Binder with Round Locking Rings	Consumer	Los Angeles	California	
C_ID_32216	Mary	Swindell	Office Supplies	Tuxton Lead Holder with Pocket Clip	Consumer	Los Angeles	California	

[Fig 5.2.2 E-Commerce Excel Measures-2]

1 YTD Sales = TOTALYTD(SUM(ecommerce_data[sales_per_order]),Calendar[Date])							
customer_id	customer_first_name	customer_last_name	Category	product_name	customer_segment	customer_city	customer_state
C_ID_29162	Mary	Jacobs	Office Supplies	Recycled Eldon Regeneration Jumbo File	Consumer	Los Angeles	California
C_ID_53661	Mary	Miller	Office Supplies	Portable Personal File Box	Consumer	Los Angeles	California
C_ID_66378	Mary	Bell	Office Supplies	Acco 6 Outlet Guardian Premium Plus Surge Suppressor	Consumer	Los Angeles	California
C_ID_29254	Mary	Olsen	Office Supplies	Ibico Recycled Linen-Style Covers	Consumer	Los Angeles	California
C_ID_33239	Mary	Swindell	Office Supplies	Fellowes Superior 10 Outlet Split Surge Protector	Consumer	Los Angeles	California
C_ID_42136	Mary	Cousins	Office Supplies	Accohide Poly Flexible Ring Binders	Consumer	Los Angeles	California
C_ID_43861	Mary	Lee	Office Supplies	TOPS Voice Message Log Book, Flash Format	Consumer	Los Angeles	California
C_ID_39004	Mary	Kastensmidt	Office Supplies	Strathmore Photo Frame Cards	Consumer	Los Angeles	California
C_ID_31297	Mary	Trafton	Office Supplies	Avery 52	Consumer	Los Angeles	California
C_ID_57963	Mary	DeCherney	Office Supplies	Dixon Ticonderoga Core-Lock Colored Pencils	Consumer	Los Angeles	California
C_ID_50743	Mary	Phan	Office Supplies	BIC Brite Liner Highlighters	Consumer	Los Angeles	California
C_ID_43653	Mary	Collins	Office Supplies	Xerox 1971	Consumer	Los Angeles	California
C_ID_52308	Mary	Coram	Office Supplies	Xerox 1893	Consumer	Los Angeles	California
C_ID_46373	Mary	Swindell	Office Supplies	Avery Reinforcements for Hole-Punch Pages	Consumer	Los Angeles	California
C_ID_49806	Mary	Triggs	Office Supplies	Tuf-Vin Binders	Consumer	Los Angeles	California
C_ID_74201	Mary	Hansen	Office Supplies	Staples	Consumer	Los Angeles	California
C_ID_55256	Mary	Ducich	Office Supplies	Stur-D-Stor Shelving, Vertical 5-Shelf: 72"H x 36"W x 18 1/2"D	Consumer	Los Angeles	California
C_ID_47712	Mary	Vittorini	Office Supplies	Universal Ultra Bright White Copier/Laser Paper, 8 1/2" x 11", Ream	Consumer	Los Angeles	California
C_ID_28782	Mary	Drucker	Office Supplies	SAFCO Commercial Wire Shelving, Black	Consumer	Los Angeles	California
C_ID_28548	Mary	Blackwell	Office Supplies	Belkin Premiere Surge Master II 8-outlet surge protector	Consumer	Los Angeles	California
C_ID_48943	Mary	Triggs	Office Supplies	Hoover Commercial SteamVac	Consumer	Los Angeles	California
C_ID_32531	Mary	Thornton	Office Supplies	Xerox 1899	Consumer	Los Angeles	California
C_ID_36871	Mary	Stewart	Office Supplies	Gould Plastics 18-Pocket Panel Bin, 34w x 5-1/4d x 20-1/2h	Consumer	Los Angeles	California
C_ID_34578	Mary	Glocke	Office Supplies	Xerox 1889	Consumer	Los Angeles	California
C_ID_66325	Mary	Beltran	Office Supplies	Newell 319	Consumer	Los Angeles	California
C_ID_53233	Mary	Bell	Office Supplies	Avery Premier Heavy-Duty Binder with Round Locking Rings	Consumer	Los Angeles	California
C_ID_33316	Mary	Sunlev	Office Supplies	Turquoise Lead Holder with Pocket Clin	Consumer	Los Angeles	California

[Fig 5.2.3 E-Commerce Excel Measures-3]

1 Calendar Table = CALENDAR(MIN(ipl_matches_details[match_date]),MAX(ipl_matches_details[match_date]))	
Date	Year
18-04-2008 00:00:00	2008
19-04-2008 00:00:00	2008
20-04-2008 00:00:00	2008
21-04-2008 00:00:00	2008
22-04-2008 00:00:00	2008
23-04-2008 00:00:00	2008
24-04-2008 00:00:00	2008
25-04-2008 00:00:00	2008
26-04-2008 00:00:00	2008
27-04-2008 00:00:00	2008
28-04-2008 00:00:00	2008
29-04-2008 00:00:00	2008
30-04-2008 00:00:00	2008
01-05-2008 00:00:00	2008
02-05-2008 00:00:00	2008
03-05-2008 00:00:00	2008
04-05-2008 00:00:00	2008
05-05-2008 00:00:00	2008
06-05-2008 00:00:00	2008
07-05-2008 00:00:00	2008
08-05-2008 00:00:00	2008
09-05-2008 00:00:00	2008
10-05-2008 00:00:00	2008
11-05-2008 00:00:00	2008
12-05-2008 00:00:00	2008
13-05-2008 00:00:00	2008
14-05-2008 00:00:00	2008
15-05-2008 00:00:00	2008

[Fig 5.2.2.1 IPL Excel Measures-1]

1 Batters Runs = CONCATENATE(SUM(ipl_details[batsman_run]), " Runs")													
id	innings	overs	ball_number	batter	bowler	non_striker	extra_type	batsman_run	extras_run	total_run	non_boundary	iswicket_delivery	player_out
1304097	1	0	4	Ishan Kishan	Mohammed Shami	RG Sharma	NA	1	0	1	0	0	NA
1304097	1	6	4	Ishan Kishan	PJ Sangwan	RG Sharma	NA	1	0	1	0	0	NA
1304064	1	0	4	Ishan Kishan	DJ Willey	RG Sharma	NA	1	0	1	0	0	NA
1304064	1	1	4	Ishan Kishan	Mohammed Siraj	RG Sharma	NA	1	0	1	0	0	NA
1304064	1	4	4	Ishan Kishan	PWH de Silva	RG Sharma	NA	1	0	1	0	0	NA
1304060	1	1	4	Ishan Kishan	Rasikh Salam	RG Sharma	NA	1	0	1	0	0	NA
1304060	1	2	4	Ishan Kishan	UT Yadav	RG Sharma	NA	1	0	1	0	0	NA
1304048	1	7	4	Ishan Kishan	AR Patel	RG Sharma	NA	1	0	1	0	0	NA
1254096	1	7	4	Q de Kock	CV Varun	RG Sharma	NA	1	0	1	0	0	NA
1254088	1	0	4	Ishan Kishan	Mohammad Nabi	RG Sharma	NA	1	0	1	0	0	NA
1254088	1	2	4	Ishan Kishan	Mohammad Nabi	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	0	4	Q de Kock	MC Henriques	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	1	4	Q de Kock	DJ Hooda	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	4	4	Ishan Kishan	MC Henriques	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	8	4	SA Yadav	Ravi Bishnoi	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	14	4	SA Yadav	FA Allen	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	15	4	SA Yadav	Mohammed Shami	RG Sharma	NA	1	0	1	0	0	NA
1254062	1	3	4	SA Yadav	Shakib Al Hasan	RG Sharma	NA	1	0	1	0	0	NA
1254062	1	13	4	HH Pandya	M Prasidh Krishna	RG Sharma	NA	1	0	1	0	0	NA
1254062	1	14	4	HH Pandya	CV Varun	RG Sharma	NA	1	0	1	0	0	NA
1216511	1	0	4	Q de Kock	AS Rajpoot	RG Sharma	NA	1	0	1	0	0	NA
1216508	1	4	4	SA Yadav	PJ Cummins	RG Sharma	NA	1	0	1	0	0	NA
1216508	1	8	4	SA Yadav	SP Narine	RG Sharma	NA	1	0	1	0	0	NA
1216508	1	9	4	SA Yadav	Kuldeep Yadav	RG Sharma	NA	1	0	1	0	0	NA
1216508	1	10	4	SA Yadav	SP Narine	RG Sharma	NA	1	0	1	0	0	NA
1216503	1	6	4	Ishan Kishan	JDS Neesham	RG Sharma	NA	1	0	1	0	0	NA

[Fig 5.2.2.2 IPL Excel Measures-2]

1 Bowlers Wickets = CONCATENATE(SUM(ipl_details[iswicket_delivery]), " Wickets")													
id	innings	overs	ball_number	batter	bowler	non_striker	extra_type	batsman_run	extras_run	total_run	non_boundary	iswicket_delivery	player_out
1304097	1	0	4	Ishan Kishan	Mohammed Shami	RG Sharma	NA	1	0	1	0	0	NA
1304097	1	6	4	Ishan Kishan	PJ Sangwan	RG Sharma	NA	1	0	1	0	0	NA
1304064	1	0	4	Ishan Kishan	DJ Willey	RG Sharma	NA	1	0	1	0	0	NA
1304064	1	1	4	Ishan Kishan	Mohammed Siraj	RG Sharma	NA	1	0	1	0	0	NA
1304064	1	4	4	Ishan Kishan	PWH de Silva	RG Sharma	NA	1	0	1	0	0	NA
1304060	1	1	4	Ishan Kishan	Rasikh Salam	RG Sharma	NA	1	0	1	0	0	NA
1304060	1	2	4	Ishan Kishan	UT Yadav	RG Sharma	NA	1	0	1	0	0	NA
1304048	1	7	4	Ishan Kishan	AR Patel	RG Sharma	NA	1	0	1	0	0	NA
1254096	1	7	4	Q de Kock	CV Varun	RG Sharma	NA	1	0	1	0	0	NA
1254088	1	0	4	Ishan Kishan	Mohammad Nabi	RG Sharma	NA	1	0	1	0	0	NA
1254088	1	2	4	Ishan Kishan	Mohammad Nabi	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	0	4	Q de Kock	MC Henriques	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	1	4	Q de Kock	DJ Hooda	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	4	4	Ishan Kishan	MC Henriques	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	8	4	SA Yadav	Ravi Bishnoi	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	14	4	SA Yadav	FA Allen	RG Sharma	NA	1	0	1	0	0	NA
1254074	1	15	4	SA Yadav	Mohammed Shami	RG Sharma	NA	1	0	1	0	0	NA
1254062	1	3	4	SA Yadav	Shakib Al Hasan	RG Sharma	NA	1	0	1	0	0	NA
1254062	1	13	4	HH Pandya	M Prasidh Krishna	RG Sharma	NA	1	0	1	0	0	NA
1254062	1	14	4	HH Pandya	CV Varun	RG Sharma	NA	1	0	1	0	0	NA
1216511	1	0	4	Q de Kock	AS Rajpoot	RG Sharma	NA	1	0	1	0	0	NA
1216508	1	4	4	SA Yadav	PJ Cummins	RG Sharma	NA	1	0	1	0	0	NA
1216508	1	8	4	SA Yadav	SP Narine	RG Sharma	NA	1	0	1	0	0	NA
1216508	1	9	4	SA Yadav	Kuldeep Yadav	RG Sharma	NA	1	0	1	0	0	NA
1216508	1	10	4	SA Yadav	SP Narine	RG Sharma	NA	1	0	1	0	0	NA
1216503	1	6	4	Ishan Kishan	JDS Neesham	RG Sharma	NA	1	0	1	0	0	NA
1216503	1	0	4	Ishan Kishan	K Gaubhan	RG Sharma	NA	1	0	1	0	0	NA

[Fig 5.2.2.3 IPL Excel Measures-3]

The screenshot displays the Microsoft SQL Server Management Studio (SSMS) interface. The title bar indicates the connection to 'SWAR\SQLEXPRESS.E-commerce2 (SWAR\swarp (68))'. The left-hand 'Object Explorer' pane shows the database structure, including 'Databases', 'System Databases', 'Database Snapshots', 'Books', 'e_commerne', 'E-commerce', 'E-commerce2', 'Database Diagrams', 'Tables', 'Views', 'External Resources', 'Synonyms', 'Programmability', 'Query Store', 'Service Broker', 'Storage', 'Security', 'employee', 'IPL', 'SQL_Keys', 'SQL_Project', and 'studentdatabase'. The main query editor window shows a SQL script with the following content:

```
select * from dbo.ecommerce_data
select * from dbo.us_state_long_lat_codes

/* Question-1:*/

-- Average order placed by customers ?
-- Monthly Sales --

SELECT AVG(order_quantity) AS avg_orders_per_customer
FROM (
    SELECT customer_id, COUNT(*) AS order_quantity
    FROM ecommerce_data
    GROUP BY customer_id
) AS order_counts
```

The 'Results' pane at the bottom shows the output of the query, which is a single row with the value 2 for the column 'avg_orders_per_customer'.

avg_orders_per_customer
2

The status bar at the bottom indicates 'Query executed successfully.' and provides performance metrics: 'SWAR\SQLEXPRESS (16.0 RTM) | SWAR\swarp (68) | E-commerce2 | 00:00:00 | 1 rows'.

[Fig 5.3.1.1 SQL Project on E-Commerce]

The screenshot displays the Microsoft SQL Server Management Studio (SSMS) interface. The left pane shows the 'Object Explorer' with the 'E-commerce2' database selected. The central pane contains a SQL query titled 'SQLQuery.sql - SWA...(SWAR\swarp (68))'. The query is as follows:

```

/* Question-2:*/
-- What are our top-selling products?
SELECT product_name, SUM(order_quantity) as total_quantity_sold
FROM ecommerce_data
GROUP BY product_name
ORDER BY total_quantity_sold DESC;

```

The bottom pane shows the 'Results' tab with a table of 16 rows. The status bar at the bottom indicates 'Query executed successfully.' and '1,849 rows'.

product_name	total_quantity_sold
1 Staples	1132
2 Staple envelope	1065
3 Easy-staple paper	997
4 KI Adjustable-Height Table	439
5 Avery Non-Stick Binders	428
6 Staples in misc. colors	414
7 Staple remover	373
8 Storex Dura Pro Binders	361
9 Staple-based wall hangings	351
10 Brefford Rectangular Conference Table Tops	345
11 Situations Contoured Folding Chairs, 4/Set	324
12 Eldon Wave Desk Accessories	315
13 Wilson Jones Ledger-Size, Piano-Hinge Binder, 2", Blue	314
14 Logitech 910-002974 M325 Wireless Mouse for Web ...	313
15 Flat Face Poster Frame	310
16 Global High-Back Leather Tilter, Burgundy	309
17 Tennessee Double Time Lockers	306

[Fig 5.3.1.2 SQL Project on E-Commerce]

The screenshot displays the Microsoft SQL Server Management Studio (SSMS) interface. The title bar indicates the connection to 'SWAR\SQLEXPRESS.E-commerce2 (SWAR\swarp (68))'. The left-hand 'Object Explorer' pane shows a tree view of the database structure, including 'Databases', 'System Databases', 'Database Snapshots', 'Books', 'e_commerce', 'E-commerce', 'E-commerce2', 'Database Diagrams', 'Tables', 'System Tables', 'FileTables', 'External Tables', 'Graph Tables', 'dbo.us_state_long_lat_cod', 'Views', 'External Resources', 'Synonyms', 'Programmability', 'Query Store', 'Service Broker', 'Storage', 'Security', 'employee', 'IPL', 'SQL_Keys', 'SQL_Project', 'studentdatabase', 'Security', 'Server Objects', 'Replication', 'Management', and 'Event Viewer'. The central 'Query Editor' pane contains a SQL query titled 'Question-3:*/' with the following text:
--What is our customer retention rate?

SELECT
COUNT(DISTINCT customer_id) AS repeat_customers,
COUNT(DISTINCT CASE WHEN order_count = 1 THEN customer_id ELSE NULL END) AS one_time_customers,
COUNT(DISTINCT customer_id) / COUNT(DISTINCT CASE WHEN order_count > 1 THEN customer_id ELSE NULL END) AS repeat_customer_rate
FROM (
SELECT customer_id, COUNT(*) as order_count
FROM ecommerce_data
GROUP BY customer_id
) AS customer_orders;
The bottom 'Results' pane shows a single row of data: repeat_customers: 42047, one_time_customers: 13430, repeat_customer_rate: 1. The status bar at the bottom indicates 'Query executed successfully.' and 'SWAR\SQLEXPRESS (16.0 RTM) | SWAR\swarp (68) | E-commerce2 | 00:00:00 | 1 rows'.

```
/* Question-3:*/  
  
--What is our customer retention rate?  
  
SELECT  
COUNT(DISTINCT customer_id) AS repeat_customers,  
COUNT(DISTINCT CASE WHEN order_count = 1 THEN customer_id ELSE NULL END) AS one_time_customers,  
COUNT(DISTINCT customer_id) / COUNT(DISTINCT CASE WHEN order_count > 1 THEN customer_id ELSE NULL END) AS repeat_customer_rate  
FROM (  
SELECT customer_id, COUNT(*) as order_count  
FROM ecommerce_data  
GROUP BY customer_id  
) AS customer_orders;
```

repeat_customers	one_time_customers	repeat_customer_rate
42047	13430	1

[Fig 5.3.1.3 SQL Project on E-Commerce]

The screenshot displays the Microsoft SQL Server Management Studio (SSMS) interface. The title bar indicates the connection to 'SQLQuery.sql - SWAR\SQLEXPRESS.IPL (SWAR\swarp (81)) - Microsoft SQL Server Management Studio (Administrator)'. The 'Object Explorer' on the left shows a tree view of the database structure, including 'Databases', 'System Databases', 'Database Snapshots', 'Books', 'e_commerce', 'E-commerce', 'E-commerce2', 'employee', 'IPL', 'Database Diagrams', 'Tables', 'System Tables', 'FileTables', 'External Tables', 'Graph Tables', 'dbo.ipl_details', 'dbo.ipl_matches_details', 'Views', 'External Resources', 'Synonyms', 'Programmability', 'Query Store', 'Service Broker', 'Storage', 'Security', 'SQL Keys', 'SQL Project', 'studentdatabase', 'Security', 'Server Objects', 'Replication', 'Management', and 'XEvent Profiler'. The 'IPL' database is selected, and its tables are expanded. The 'Query Editor' window shows a SQL query with the following text:

```
select * from dbo.ipl_details
select * from dbo.ipl_matches_details

/* Question-1:*/

--List out players with most man of the match?

select player_of_match,count(*) as topplayer
from ipl_matches_details
group by player_of_match
order by count(*)
desc;
```

The 'Results' pane at the bottom displays the output of the query, showing a list of players and their corresponding counts. The status bar at the bottom indicates 'Query executed successfully.' and 'SWAR\SQLEXPRESS (16.0 RTM) SWAR\swarp (81) IPL 00:00:00 263 rows'.

player_of_match	topplayer
AB de Villiers	25
CH Gayle	22
DA Warner	18
RG Sharma	18
MS Dhoni	17
SR Watson	16
YK Pathan	16
V Kohli	14
SK Raina	14
KA Pollard	14
G Gambhir	13
AD Russell	12
A Mishra	12
AM Rahane	12
KL Rahul	12
MEK Hussey	12
SD Mhambre	12

[Fig 5.3.2.1 SQL Project on IPL]

The screenshot displays the Microsoft SQL Server Management Studio (SSMS) interface. The title bar indicates the connection is to 'SQLQuery.sql - SWAR\SQLEXPRESS.IPL (SWAR\swarp (81)) - Microsoft SQL Server Management Studio (Administrator)'. The 'Object Explorer' on the left shows a tree view of the database structure, including 'Databases', 'System Databases', 'Database Snapshots', 'Books', 'e_commerne', 'Ecommerce', 'E-commerce2', 'employee', 'IPL', 'Database Diagrams', 'Tables', 'System Tables', 'FileTables', 'External Tables', 'Graph Tables', 'dbo.ipl_details', 'dbo.ipl_matches_details', 'Views', 'External Resources', 'Synonyms', 'Programmability', 'Query Store', 'Service Broker', 'Storage', 'Security', 'SQL Keys', 'SQL Project', 'studentdatabase', 'Security', 'Server Objects', 'Replication', 'Management', and 'XEvent Profiler'.

The main query editor shows the following SQL code:

```

/* Question-2:*/
--Find out the number of matches each team has won so far?

select winning_team, count(*) as win from ipl_matches_details
group by winning_team
order by count(*)
desc ;

```

The 'Results' pane at the bottom displays the output of the query as a table with two columns: 'winning_team' and 'win'. The results are sorted in descending order of the number of wins.

winning_team	win
1 Mumbai Indians	131
2 Chennai Super Kings	121
3 Kolkata Knight Riders	114
4 Royal Challengers Bangalore	109
5 Rajasthan Royals	96
6 Kings XI Punjab	88
7 Sunrisers Hyderabad	75
8 Delhi Daredevils	67
9 Delhi Capitals	36
10 Deccan Chargers	29
11 Gujarat Lions	13
12 Punjab Kings	13
13 Pune Warriors	12
14 Gujarat Titans	12
15 Rising Pune Supergiant	10
16 Lucknow Super Giants	9
17 Kolkata Knight Riders	8

[Fig 5.3.2.2 SQL Project on IPL]

The screenshot displays the Microsoft SQL Server Management Studio (SSMS) interface. The title bar indicates the connection to 'SQLQuery.sql - SWAR\SQLEXPRESS.IPL (SWAR\swarp (81)) - Microsoft SQL Server Management Studio (Administrator)'. The 'Object Explorer' on the left shows a tree view of the database structure, including 'Databases', 'System Databases', 'Database Snapshots', 'Books', 'e_commerne', 'E-commerce', 'E-commerce2', 'employee', 'IPL', 'Database Diagrams', 'Tables', 'System Tables', 'FileTables', 'External Tables', 'Graph Tables', 'dbo.ipl_matches_details', 'Views', 'External Resources', 'Synonyms', 'Programmability', 'Query Store', 'Service Broker', 'Storage', 'Security', 'SQL_Keys', 'SQL_Project', 'studentdatabase', 'Security', 'Server Objects', 'Replication', 'Management', and 'XEvent Profiler'. The 'IPL' database is expanded, showing 'Tables' and 'Views'. The 'Query Editor' window shows a SQL query:

```

/* Question-3:*/
-- List out all the ground with total number of matches played each of them?

select venue,count(*) as total_matches_held
from ipl_matches_details
group by venue
order by count(*) desc;

```

The 'Results' pane at the bottom shows the output of the query as a table with two columns: 'venue' and 'total_matches_held'. The data is sorted in descending order of the number of matches held. The status bar at the bottom indicates 'Query executed successfully.' and 'SWAR\SQLEXPRESS (16.0 RTM) | SWAR\swarp (81) | IPL | 00:00:00 | 49 rows'.

venue	total_matches_held
1 Eden Gardens	77
2 Wankhede Stadium	73
3 M Chinnaswamy Stadium	65
4 Feroz Shah Kotla	60
5 Rajiv Gandhi International Stadium, Uppal	49
6 MA Chidambaram Stadium, Chepauk	48
7 Saijal Mansingh Stadium	47
8 Dubai International Cricket Stadium	46
9 Punjab Cricket Association Stadium, Mohali	35
10 Wankhede Stadium, Mumbai	31
11 Sheikh Zayed Stadium	29
12 Sharjah Cricket Stadium	28
13 Maharashtra Cricket Association Stadium	22
14 Dr DY Patil Sports Academy, Mumbai	20
15 Dr DY Patil Sports Academy	17
16 Brabourne Stadium, Mumbai	17
17 Sahara Sports Stadium	16

[Fig 5.3.2.3 SQL Project on IPL]

The screenshot displays the Microsoft SQL Server Management Studio (SSMS) interface. The title bar indicates the connection is to 'SQLQuery.sql - SWAR\SQLEXPRESS\IPL (SWAR\swarp (81)) - Microsoft SQL Server Management Studio (Administrator)'. The Object Explorer on the left shows a tree view of the database structure, including 'Databases', 'System Databases', 'Database Snapshots', 'Books', 'e_commerne', 'Ecommerce', 'E-commerce', 'E-commerce2', 'employee', 'IPL', 'Database Diagrams', 'Tables', 'System Tables', 'FileTables', 'External Tables', 'Graph Tables', 'dbo.ipl_details', 'Views', 'External Resources', 'Synonyms', 'Programmability', 'Query Store', 'Service Broker', 'Storage', 'Security', 'SQL Keys', 'SQL Project', 'studentdatabase', 'Security', 'Server Objects', 'Replication', 'Management', and 'XEvent Profiler'. The central query editor shows the following SQL code:

```
/* Question-4:*/  
--List out top umpires who have stood the most umpire 1 ?  
  
SELECT umpire1, COUNT(*) AS ump_1  
FROM ipl_matches_details  
GROUP BY umpire1  
ORDER BY ump_1 DESC;
```

The Results pane at the bottom displays the output of the query, showing a list of umpires and their corresponding counts. The status bar at the bottom indicates 'Query executed successfully.' and 'SWAR\SQLEXPRESS (16.0 RTM) SWAR\swarp (81) IPL | 00:00:00 | 58 rows'.

umpire1	ump_1
AK Chaudhary	95
HDPK Dhamasena	76
Aaad Rauf	51
S Ravi	48
CB Gaffaney	47
KN Ananthapadmanabhan	44
M Erasmus	42
Aleem Dar	38
BF Bowden	36
BR Doctrove	34
Nitin Menon	31
AY Dandekar	30
BNJ Oxenford	30
C Shamshuddin	28
RE Koertzen	20
S Anani	19
M.L.Lee	17

[Fig 5.3.2.4 SQL Project on IPL]

The screenshot displays the Microsoft SQL Server Management Studio (SSMS) interface. The title bar indicates the connection to 'SQLQuery.sql - SWAR\SQLEXPRESS\IPL (SWAR\swarp (81)) - Microsoft SQL Server Management Studio (Administrator)'. The 'Object Explorer' on the left shows a tree view of the database structure, including 'Databases', 'System Databases', 'Database Snapshots', 'Books', 'e_commerne', 'E-commerce', 'E-commerce2', 'employee', 'IPL', 'Database Diagrams', 'Tables', 'System Tables', 'FileTables', 'External Tables', 'Graph Tables', 'dbo.ipl_details', 'Views', 'External Resources', 'Synonyms', 'Programmability', 'Query Store', 'Service Broker', 'Storage', 'Security', 'SQL Keys', 'SQL Project', 'studentdatabase', 'Security', 'Server Objects', 'Replication', 'Management', and 'XEvent Profiler'. The 'Query Editor' window shows a SQL query with the following text:

```
/* Question-5:*/  
--List out top umpires who have stood the most umpire 2 ?  
  
select umpire2,count(*) as upm2  
from ipl_matches_details  
group by umpire2  
order by count(*) desc ;
```

The 'Results' pane at the bottom displays the query output as a table with two columns: 'umpire2' and 'upm2'. The table contains 17 rows of data, with the first row being 'S Ravi' with a count of 83. The status bar at the bottom indicates 'Query executed successfully.' and 'SWAR\SQLEXPRESS (16.0 RTM) SWAR\swarp (81) IPL | 00:00:00 60 rows'.

umpire2	upm2
S Ravi	83
C Shamshuddin	60
SJA Taufel	54
Nitin Menon	52
RJ Tucker	46
CK Nandan	43
VK Sharma	42
VA Kulkarni	38
RB Tiffin	30
BNJ Oxenford	27
SK Tarapore	26
RK Illingworth	25
M Erasmus	25
YC Barde	22
CB Gaffaney	22
AM Saheba	22
DD Gaffaney	21

[Fig 5.3.2.5 SQL Project on IPL]

The screenshot displays the Microsoft SQL Server Management Studio (SSMS) interface. The title bar indicates the connection to 'SQLQuery.sql - SWAR\SQLEXPRESS.IPL (SWAR\swarp (81))'. The 'Object Explorer' on the left shows a tree view of the database structure, including 'Databases', 'System Databases', 'Database Snapshots', 'Books', 'e_commerce', 'E-commerce', 'E-commerce2', 'employee', 'IPL', 'Database Diagrams', 'Tables', 'Views', 'External Resources', 'Synonyms', 'Programmability', 'Query Store', 'Service Broker', 'Storage', 'Security', 'SQL Keys', 'SQL Project', 'studentdatabase', 'Security', 'Server Objects', 'Replication', 'Management', and 'XEvent Profiler'. The 'IPL' database is expanded, showing 'Tables' and 'Views'. The 'Tables' folder is expanded, showing 'System Tables', 'FileTables', 'External Tables', 'Graph Tables', 'dbo.ipl_details', and 'dbo.ipl_matches_details'. The 'Views' folder is expanded, showing 'External Resources', 'Synonyms', 'Programmability', 'Query Store', 'Service Broker', 'Storage', 'Security', 'SQL Keys', 'SQL Project', 'studentdatabase', 'Security', 'Server Objects', 'Replication', 'Management', and 'XEvent Profiler'. The 'Query Editor' in the center shows a SQL query:

```
/* Question-6:*/  
--Find out all the team names who have won matches with more than 50 matches?  
  
select  
winning_team, count(*) as won  
from ipl_matches_details  
group by winning_team  
having count(*) >50;
```

 The 'Results' pane at the bottom shows the output of the query, displaying a table with two columns: 'winning_team' and 'won'. The data is as follows:

winning_team	won
Mumbai Indians	131
Sunrisers Hyderabad	75
Royal Challengers Bangalore	109
Kolkata Knight Riders	114
Rajasthan Royals	96
Kings XI Punjab	88
Delhi Daredevils	67
Chennai Super Kings	121

 The status bar at the bottom indicates 'Query executed successfully.' and 'SWAR\SQLEXPRESS (16.0 RTM) | SWAR\swarp (81) | IPL | 00:00:00 | 8 rows'.

[Fig 5.3.2.6 SQL Project on IPL]

Chapter 6. Implementation

6.1 Process/Program/Technology/Modules Specification

In summary, data analytics involves a process of collecting, cleaning, exploring, analyzing, and visualizing data, using programming languages such as Python and R, technologies such as big data and cloud computing, and modules such as NumPy, Pandas, and Scikit-learn. By using these tools and techniques, data analysts can gain insights and make informed decisions based on data.

- Data Collection, Data Cleaning, Data Exploration, Data Analysis and Data Visualization.
- Python, SQL are used as a Program.
- Big Data, Cloud Computing and Machine Learning are used as Technologies.
- NumPy, Pandas, Matplotlib and Scikit-Learn are used as Modules.

6.2 Finding/Result/Outcomes

6.2.1 Technical Skills

- Data Manipulation
- Data Visualization
- Statistical Analysis
- Programming Languages
- Machine Learning

6.2.2 Soft & Management Skills

- Communication Skills
- Problem-Solving Skills
- Time Management Skills
- Project Management Skills
- Business Acumen

Chapter 7. Conclusion and Discussion

7.1 Overall Analysis of Internship Viabilities

- Due to Internship, I got the knowledge that how to analyze or work on live analysis project in Industrial and Professional Environment.
- Throughout My Internship, I could understand more about an IT Field and prepare Myself for Future.
- During My Training Period, I have received advice from mentors when mistakes were made. However, those advices are useful guidance for me to change myself and avoid myself making the same mistakes again.
- In sum, the activities that I had learned during Industrial Training really are useful for me in future to face challenges in a working environment.

7.2 Limitation and Future Enhancement

- Internship offer wonderful opportunities and are a great way to kickstart a career, blurring the line between education and employment in data analytics.
- I'm considering an Internship, I will be aware of the advantages it can offers me, and maximize these as much as I can.
- Similarly, I will take time to understand and gain more knowledge about my data analytics internship.
- Limitation of Data Analytics like Quality of Data, Bias and Interpretation, Data Security and Privacy and Technical Limitations.
- Future Enhancement in Data Analytics is Artificial Intelligence, Real-Time analytics, Data Visualization, Integration with other Technologies.

References:

- Website: <https://V-Ex Tech software solution>
- LinkedIn: <https://in.linkedin.com/company/v-ex-tech-software-company-in-vadodara>
- YouTube: (Codelike WE) https://www.youtube.com/@Veer_Agraval
- W3School: <https://www.w3schools.com/sql/default.asp>
- Kaggle: <https://www.kaggle.com/>
- GitHub: <https://github.com/codebasics>

THANK YOU