

Data Analytics
AN INTERNSHIP REPORT

Submitted by

Dev Satishkumar Panchal

In Partial fulfillment for the award of the degree Of

BACHELOR OF ENGINEERING

in
Information Technology



K.J. Institute of Engineering & Technology

Opp. ITI, Javla, Savli- Halol Road, Savli, Vadodara.

**Gujarat Technological University,
Ahmedabad**

Academic year (2023-2024)

K.J. Institute of Engineering & Technology

Opp. ITI, Javla, Savli- Halol Road, Savli, Vadodara.

Information Technology

2023-24



CERTIFICATE

Date:01/11/2023

This is to certify that the Internship Work entitled "**Data Analytics**" has been carried out by **Dev Satishkumar Panchal** under my guidance in fulfillment of the degree of Bachelor of Engineering in Information Technology (7th Semester) of Gujarat Technological University, Ahmedabad during the academic year 2023-24.

Internal Guide

I/C Head of the Department

CERTIFICATE OF COMPLETION



CERTIFICATE

of Completion

DATA ANALYSIS INTERN

DEV PANCHAL

THIS CERTIFICATE IS GIVEN TO YOU BECAUSE YOU HAD DONE
INTERNSHIP OF DATA ANALYSIS INTERN AT V-EX TECH SOLUTION.



Himanshu A.
CEO & Director

ACKNOWLEDGEMENT

First of all, I am grateful to the Almighty, who 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.**, Head of the Department of Information Technology. And also, to **Prof.** 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 Aggarwal 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. Atliq HR Analysis Dash Board	10
3	Fig 5.1.2 Meet E-Commerce Sales Dash Board	11
4	Fig 5.1.3 IPL Analysis Dash Board	11
5	Fig 5.2.1 Shree Ram Store Annual Excel Report	12
6	Fig 5.2.2 Road Accident Excel Report	12
7	Fig 5.3.1.1 Diwali Sales Report of Gender	13
8	Fig 5.3.1.2 Diwali Sales Report of Gender Number	13
9	Fig 5.3.1.3 Diwali Sales Report of Age Group	14
10	Fig 5.3.1.4 Diwali Sales Report of Top 10 States	14
11	Fig 5.3.1.5 Diwali Sales Report of Marital Status	15
12	Fig 5.3.1.6 Diwali Sales Report of Marital Status with Gender and Amount	
13	Fig 5.3.1.7 Diwali Sales Report of Occupation	
14	Fig 5.3.1.8 Diwali Sales Report of Occupation with Amount	
15	Fig 5.3.1.9 Diwali Sales Report of Product Category	
16	Fig 5.3.1.10 Diwali Sales Report of Occupation with Amount	
17	Fig 5.3.2.1 IPL Analysis Report of Team Won Batting First	
18	Fig 5.3.2.2 IPL Analysis Report of Top 5 Players of Man Of The Match Awards	
19	Fig 5.3.2.3 IPL Analysis Report of Top 3 Team Won Batting First	

19	Fig 5.3.2.4 IPL Analysis Report of Team Won Batting First in Pie-Chart	
20	Fig 5.3.2.5 IPL Analysis Report of Team Frequency of Wins w.r.t Number of Wickets	
21	Fig 5.3.2.6 IPL Analysis Report of Most Win after Batting Second	
22	Fig 5.3.2.7 IPL Analysis Report of Most Win after Batting Second in Pie-Chart	
23	Fig 5.4.1.1 SQL Project on Music Store Task 1 (Section-1)	
24	Fig 5.4.1.2 SQL Project on Music Store Task 2 (Section-1)	
25	Fig 5.4.1.3 SQL Project on Music Store Task 3 (Section-1)	
26	Fig 5.4.1.4 SQL Project on Music Store Task 4 (Section-1)	
27	Fig 5.4.1.5 SQL Project on Music Store Task 5 (Section-1)	
28	Fig 5.4.1.6 SQL Project on Music Store Task 1 (Section-2)	
29	Fig 5.4.1.7 SQL Project on Music Store Task 2 (Section-2)	
30	Fig 5.4.1.8 SQL Project on Music Store Task 3 (Section-2)	
31	Fig 5.4.2.1 SQL Project on IPL Database Task 1-2	
32	Fig 5.4.2.2 SQL Project on IPL Database Task 3	
33	Fig 5.4.2.3 SQL Project on IPL Database Task 4	
34	Fig 5.4.2.4 SQL Project on IPL Database Task 5	
35	Fig 5.4.2.5 SQL Project on IPL Database Task 6	
36	Fig 5.4.2.6 SQL Project on IPL Database Task 7	
37	Fig 5.4.2.7 SQL Project on IPL Database Task 8	
38	Fig 5.4.2.8 SQL Project on IPL Database Task 9	

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
6.3 Date of Continuous Evaluation.....	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 penal 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-Ex Tech software solution>
- **Email:** himanshu0409agraval@gmail.com
- **Address:** 101, First Floor, Kunal Complex, Delux Char Rasta, Opposite Passport Office, 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 must 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 and resources.
- Insufficient training in data privacy and security.
- Lack of diversity and inclusion.
- Limited feedback and evaluation.

Chapter 5. Project Screenshot



HR Analytics Dash Board

Apr 22

May 22

Jun 22

170.3% 11.15% 2.01%

Presence %

WFH %

SL %

Name	WFH %	SL %	Presence %
Aditya Walls	7.14%	0.00%	200.00%
Adriel Pace	5.45%	0.00%	177.42%
Adyson Moyer	10.91%	0.00%	177.42%
Alexander Davenport	100.00%	0.00%	157.14%
Alyson Huber	0.00%	0.00%	260.00%
Ana Little	0.00%	3.33%	140.00%
Andrew Cummings	40.00%	0.00%	166.67%
April Ayers	0.00%	3.23%	174.19%
Athena Rios	7.41%	0.00%	192.86%
Ayanna Atkins	9.09%	19.35%	106.45%
Bo Cordova	33.33%	2.63%	142.11%
Boston Morse	0.00%	0.00%	174.19%
Briley Orr	5.77%	0.00%	167.74%
Total	11.15%	2.01%	170.33%

Day of week	Presence %
Tue	92.69%
Mon	92.66%
Wed	91.89%
Thu	90.54%
Fri	90.08%
Total	170.33%

SL %

Presence % by Date



Name	01 April 2022	04 April 2022
Aditya Walls	P	P
Adriel Pace	P	P
Adyson Moyer	P	P
Alexander Davenport		
Alyson Huber	P	P
Ana Little	P	P
Andrew Cummings	P	P
April Ayers	P	p
Total	HML	HPL

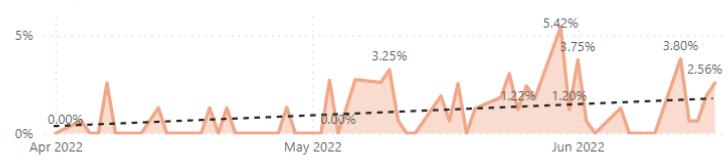
Day of week SL %

Day of week	SL %
Fri	0.69%
Mon	1.59%
Thu	1.03%
Tue	1.20%
Wed	0.92%
Total	2.01%

WFH % by Date



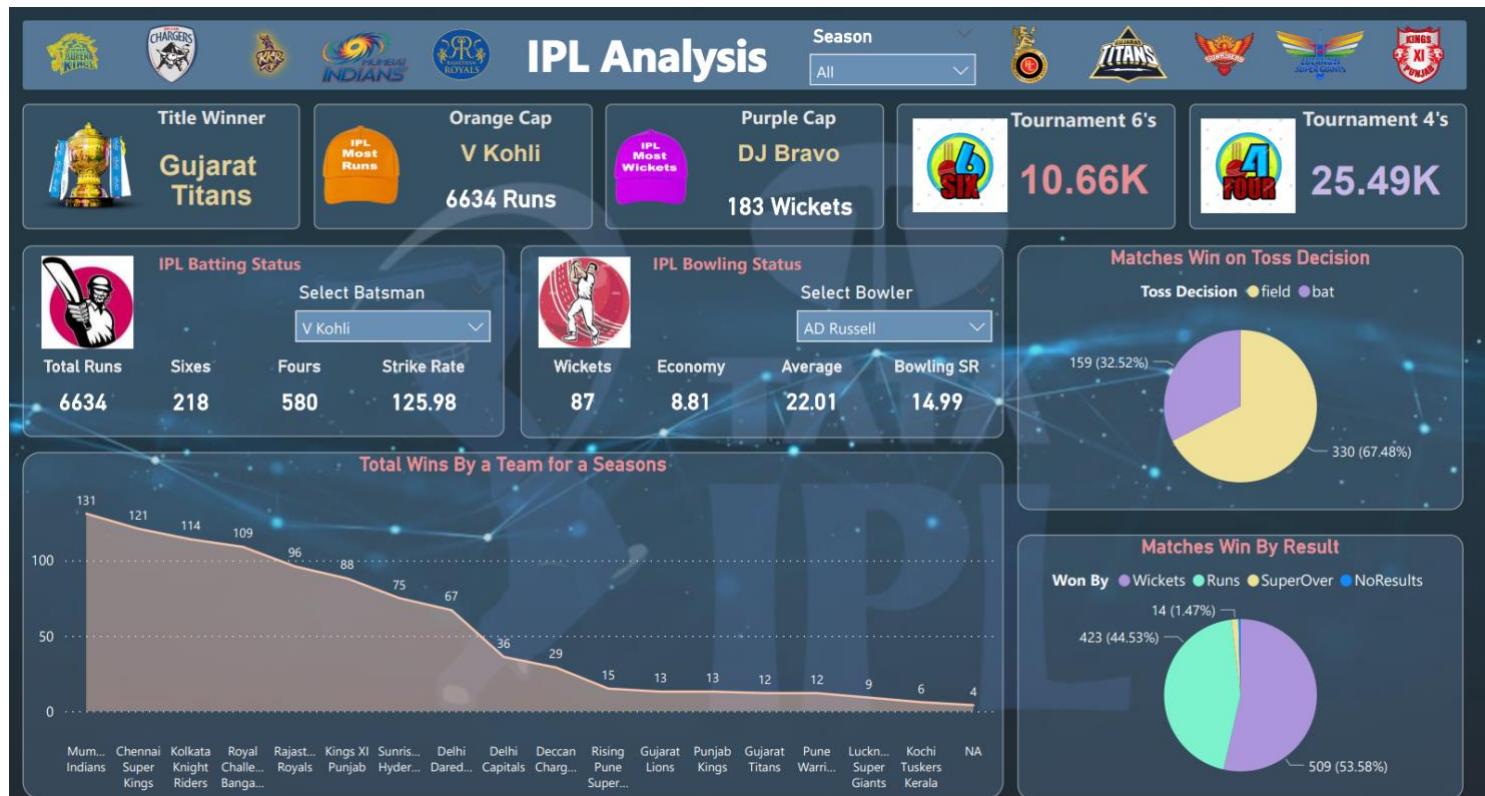
SL % by Date



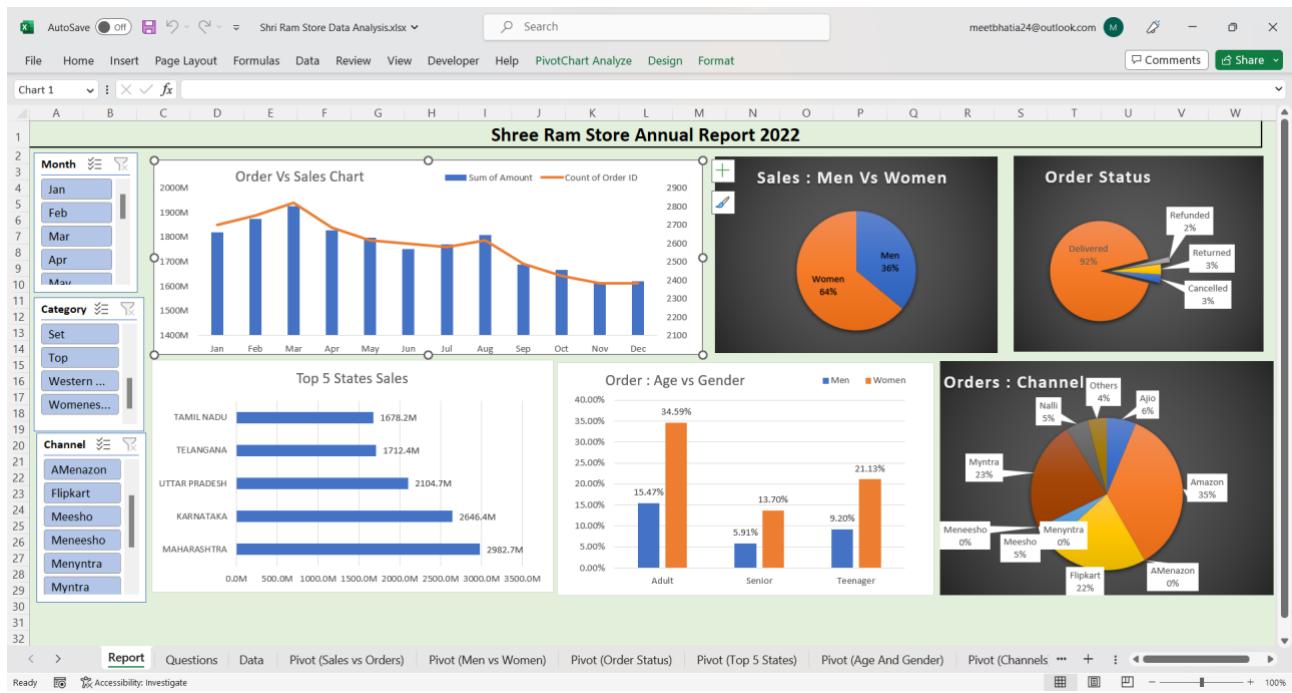
[Fig 5.1.1. Atliq HR Analysis Dash Board]



[Fig 5.1.2 Meet E-Commerce Sales Dash Board]



[Fig 5.1.3 IPL Analysis Dash Board]

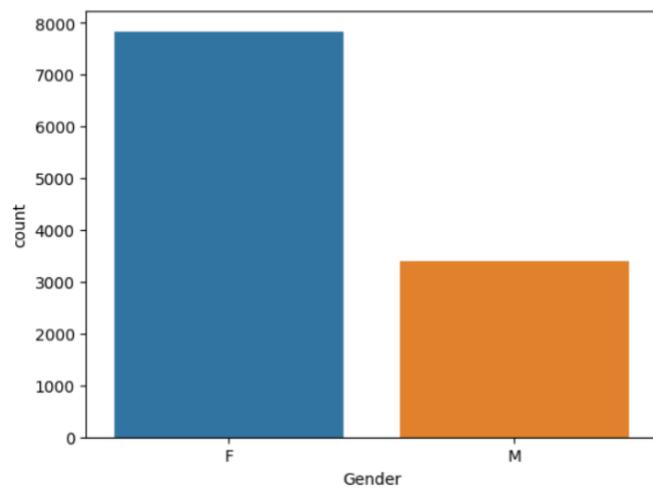


[Fig 5.2.1 Shree Ram Store Annual Excel Report]



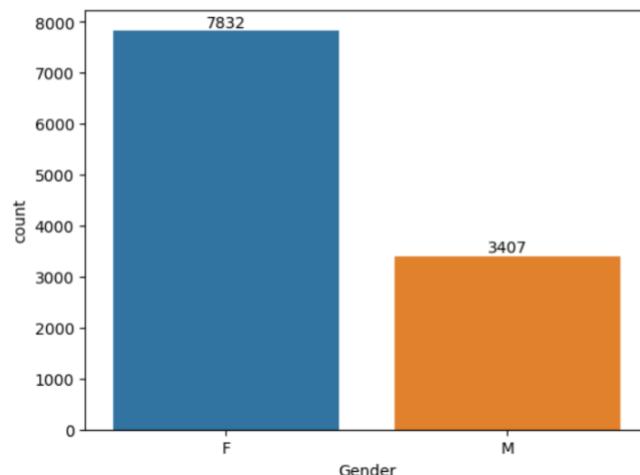
[Fig 5.2.2 Road Accident Excel Report]

```
In [35]: sns.countplot(x = 'Gender',data = df)
Out[35]: <Axes: xlabel='Gender', ylabel='count'>
```



[Fig 5.3.1.1 Diwali Sales Report of Gender]

```
In [34]: ax = sns.countplot(x = 'Gender',data = df)
for bars in ax.containers:
    ax.bar_label(bars) #using this we got a value in graph
```

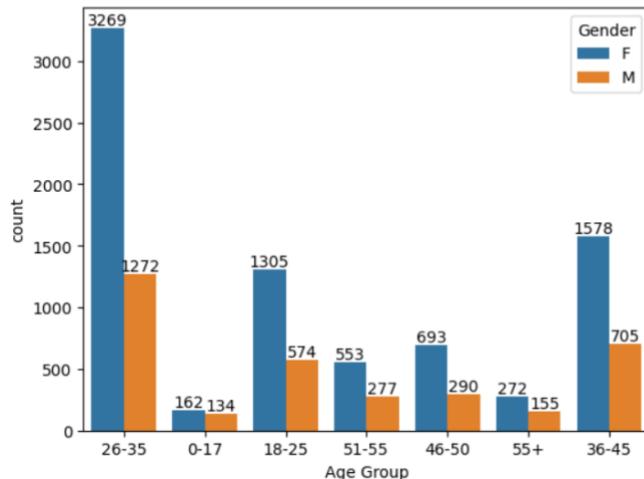


[Fig

5.3.1.2
Diwali Sales Report of Gender Number]

```
In [43]: ax = sns.countplot(data = df, x = 'Age Group', hue = 'Gender')

for bars in ax.containers:
    ax.bar_label(bars)
```



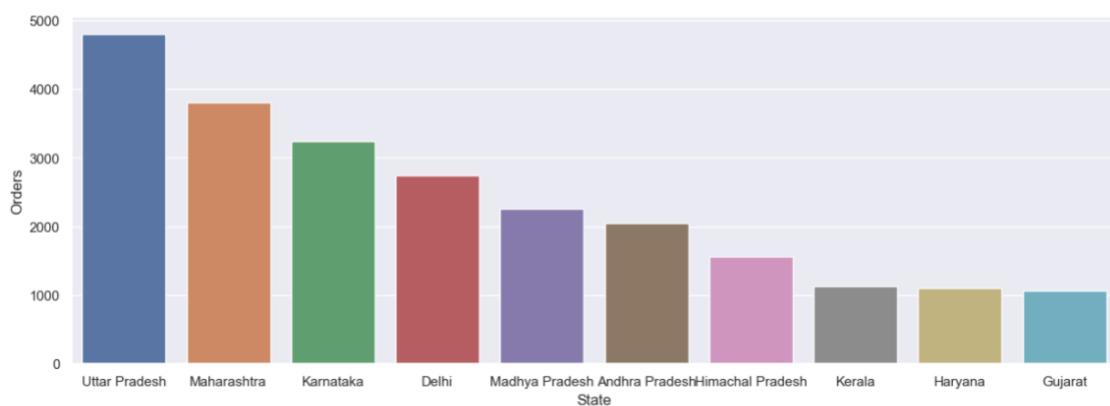
[Fig 5.3.1.3 Diwali Sales Report of Age Group]

```
In [45]: # total number of orders from top 10 states

sales_state = df.groupby(['State'], as_index=False)[['Orders']].sum().sort_values(by='Orders', ascending=False).head(10)

sns.set(rc={'figure.figsize':(15,5)})
sns.barplot(data = sales_state, x = 'State',y= 'Orders')
```

```
Out[45]: <Axes: xlabel='State', ylabel='Orders'>
```

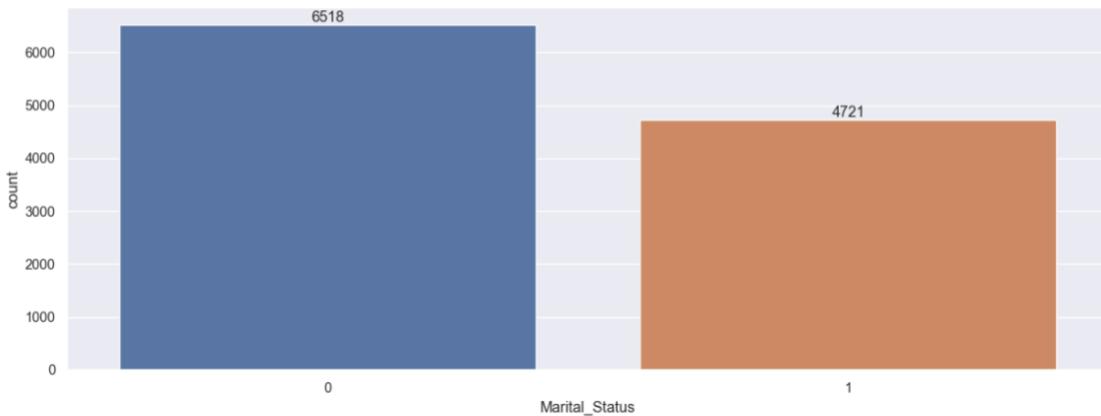


[Fig

5.3.1.4 Diwali Sales Report of Top 10 States]

```
In [47]: ax = sns.countplot(data = df, x = 'Marital_Status')

sns.set(rc={'figure.figsize':(7,5)})
for bars in ax.containers:
    ax.bar_label(bars)
```

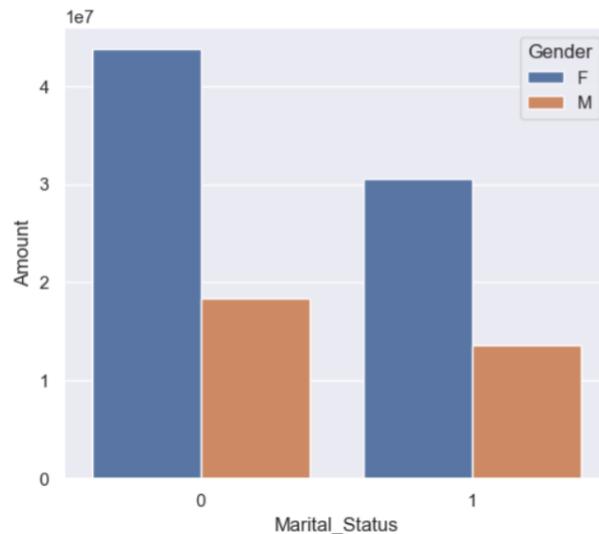


[Fig 5.3.1.5 Diwali Sales Report of Marital Status]

```
In [48]: sales_state = df.groupby(['Marital_Status', 'Gender'], as_index=False)[['Amount']].sum().sort_values(by='Amount', ascending=False)

sns.set(rc={'figure.figsize':(6,5)})
sns.barplot(data = sales_state, x = 'Marital_Status',y= 'Amount', hue='Gender')

out[48]: <Axes: xlabel='Marital_Status', ylabel='Amount'>
```

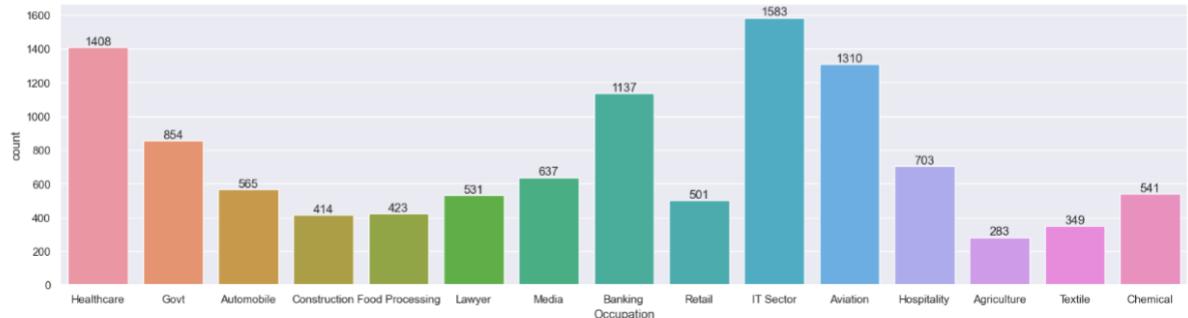


[Fig 5.3.1.6 Diwali Sales Report of Marital Status with Gender and Amount]

Occupation

```
In [49]: sns.set(rc={'figure.figsize':(20,5)})
ax = sns.countplot(data = df, x = 'Occupation')

for bars in ax.containers:
    ax.bar_label(bars)
```

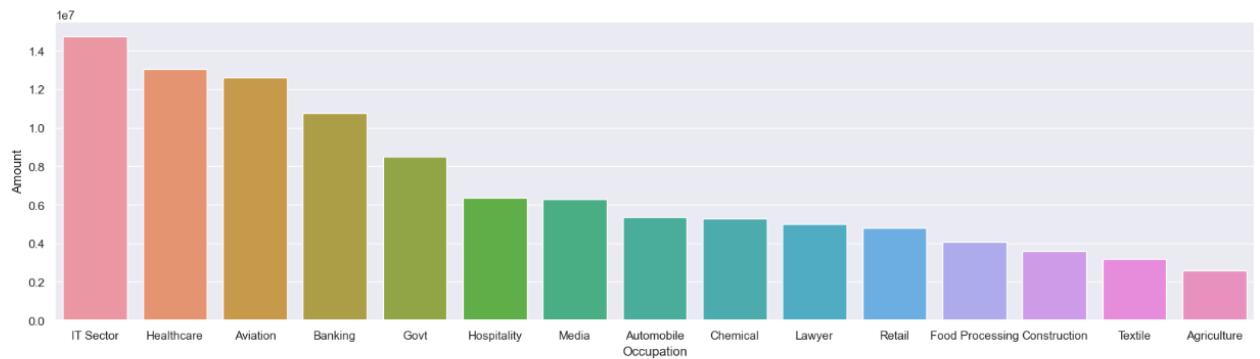


[Fig 5.3.1.7 Diwali Sales Report of Occupation]

```
In [50]: sales_state = df.groupby(['Occupation'], as_index=False)[['Amount']].sum().sort_values(by='Amount', ascending=False)

sns.set(rc={'figure.figsize':(20,5)})
sns.barplot(data = sales_state, x = 'Occupation',y= 'Amount')

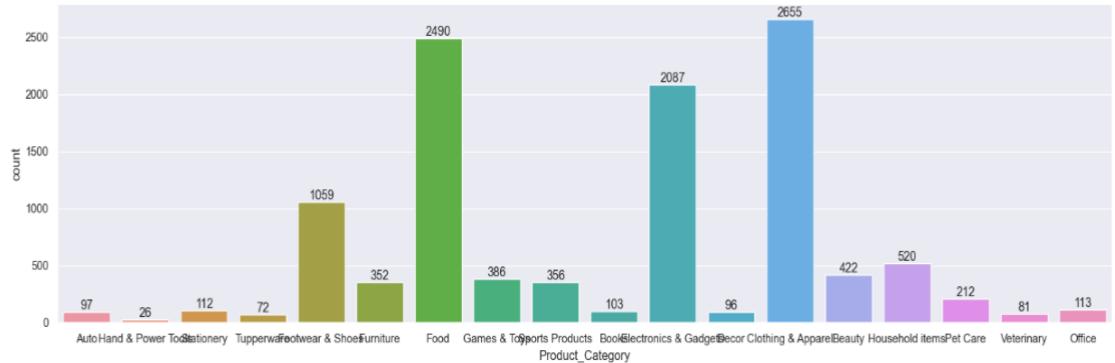
Out[50]: <Axes: xlabel='Occupation', ylabel='Amount'>
```



[Fig 5.3.1.8 Diwali Sales Report of Occupation with Amount]

```
In [51]: sns.set(rc={'figure.figsize':(20,5)})
ax = sns.countplot(data = df, x = 'Product_Category')

for bars in ax.containers:
    ax.bar_label(bars)
```

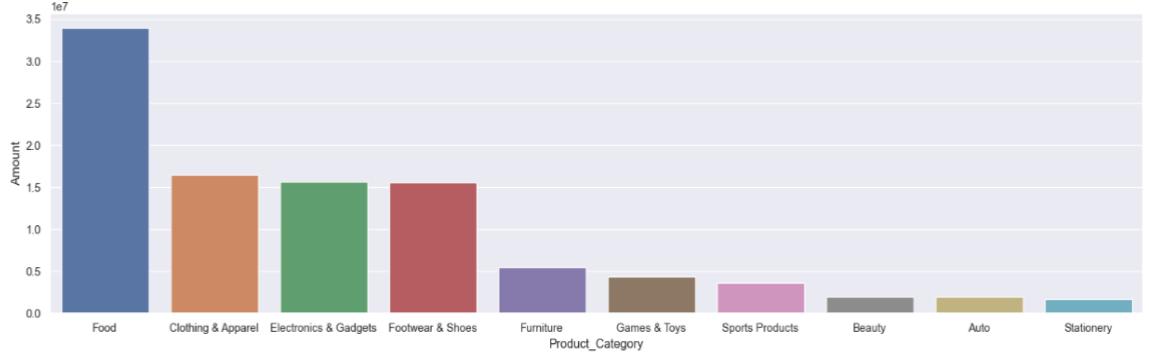


[Fig 5.3.1.9 Diwali Sales Report of Product Category]

```
In [52]: sales_state = df.groupby(['Product_Category'], as_index=False)[['Amount']].sum().sort_values(by='Amount', ascending=False).head(10)

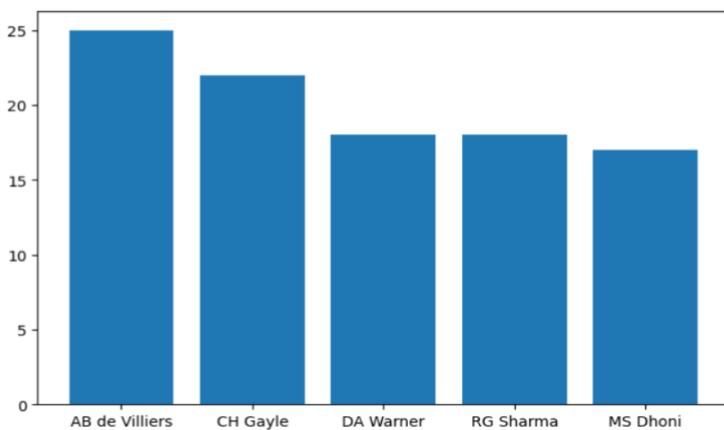
sns.set(rc={'figure.figsize':(20,5)})
sns.barplot(data = sales_state, x = 'Product_Category',y= 'Amount')
```

```
Out[52]: <Axes: xlabel='Product_Category', ylabel='Amount'>
```

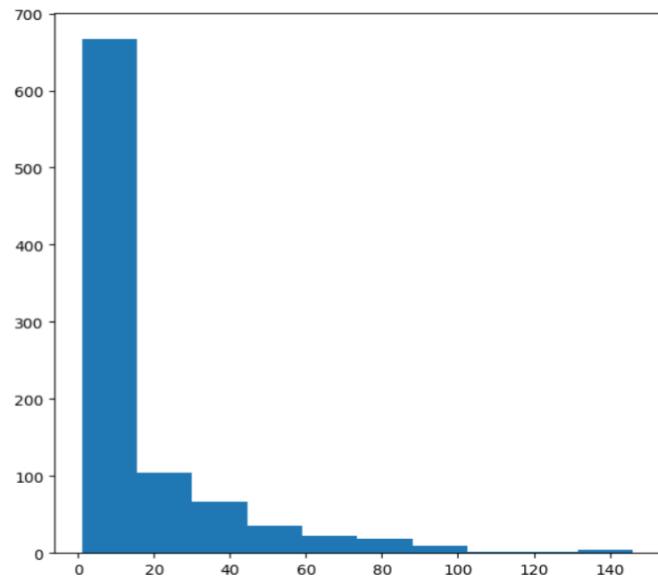


[Fig 5.3.1.10 Diwali Sales Report of Occupation with Amount]

```
In [9]: #making a bar-plot for the top 5 players with most man of the match awards
plt.figure(figsize=(8,5))
plt.bar(list(ipl['player_of_match'].value_counts()[0:5].keys()),list(ipl['player_of_match'].value_counts()[0:5]))
plt.show()
```

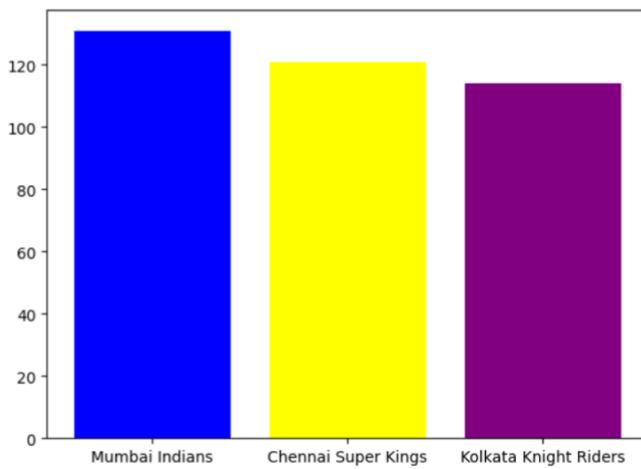


[Fig 5.3.2.1 IPL Analysis Report of Team Won Batting First]



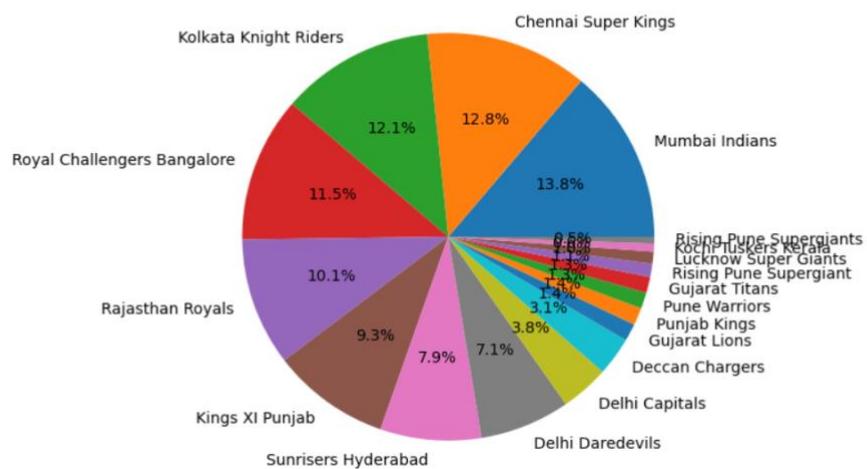
[Fig 5.3.2.2 IPL Analysis Report of Top 5 Players of Man Of The Match Awards]

```
In [16]: #Making a bar-plot for top 3 teams with most wins after batting first
plt.figure(figsize=(7,5))
plt.bar(list(batting_first['winning_team'].value_counts()[0:3].keys()),list(batting_first['winning_team'].value_counts()[0:3]),color=['blue','yellow','purple'])
plt.show()
```



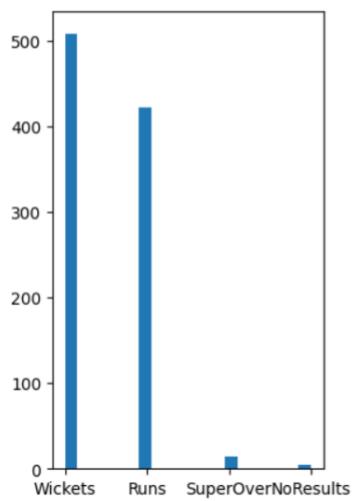
[Fig 5.3.2.3 IPL Analysis Report of Top 3 Team Won Batting First]

```
In [17]: #Making a pie chart
plt.figure(figsize=(13,6))
plt.pie(list(batting_first['winning_team'].value_counts()),labels=list(batting_first['winning_team'].value_counts().keys()),autopct='%.2f')
plt.show()
```



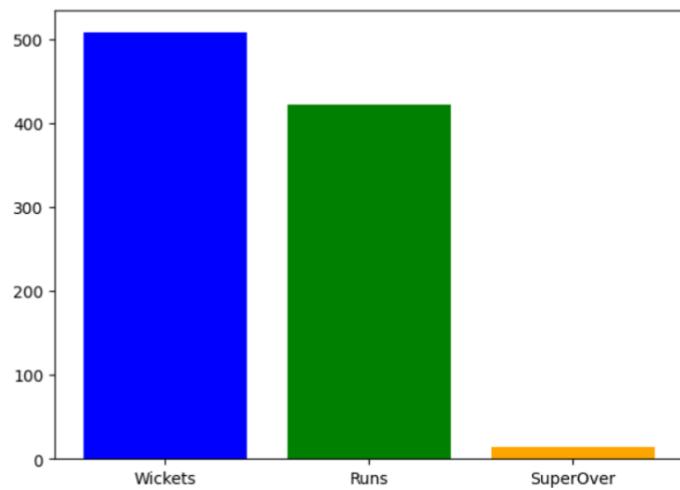
[Fig 5.3.2.4 IPL Analysis Report of Team Won Batting First in Pie-Chart]

```
In [20]: #Making a histogram for frequency of wins w.r.t number of wickets
plt.figure(figsize=(3,5))
plt.hist(batting_second['won_by'],bins=20)
plt.show()
```



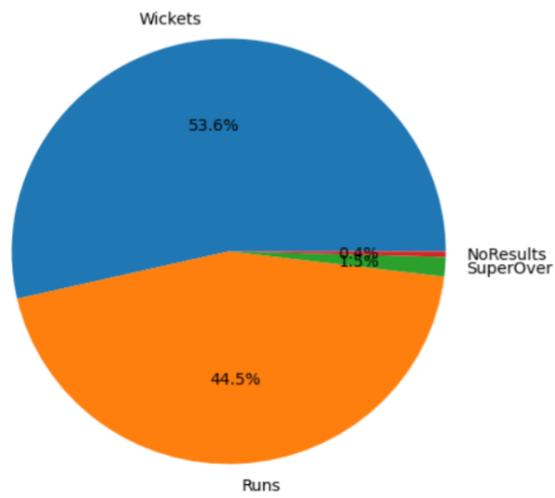
[Fig 5.3.2.5 IPL Analysis Report of Team Frequency of Wins w.r.t Number of Wickets]

```
In [23]: #Making a bar plot for most wins after batting second
plt.figure(figsize=(7,5))
top_3 = list(batting_second['won_by'].value_counts().keys()[:3])
plt.bar(top_3, batting_second['won_by'].value_counts()[top_3], color=["blue", "green", "orange"])
plt.show()
```



[Fig 5.3.2.6 IPL Analysis Report of Most Win after Batting Second]

```
In [76]: #Making a pie chart for distribution of most wins after batting second
plt.figure(figsize=(6,7))
plt.pie(list(batting_second['won_by'].value_counts()),labels=list(batting_second['won_by'].value_counts().keys()),autopct='%0.1f%%')
plt.show()
```



[Fig 5.3.2.7 IPL Analysis Report of Most Win after Batting Second in Pie-Chart]

Object Explorer

```

select * from PlaylistTrack
select * from Track

/*Question Set 1*/

-- Q.1 Who is the senior most employee based on job title?

select * from Employee
SELECT TOP 1 title, last_name, first_name
FROM employee
ORDER BY levels DESC;

```

Results

	title	last_name	first_name
1	Senior General Manager	Madan	Mohan

Query executed successfully.

[Fig 5.4.1.1 SQL Project on Music Store Task 1 (Section-1)]

Object Explorer

```

-- Q.2 Which Countries have the most Invoices?

select * from Invoice
SELECT COUNT(*) AS c, billing_country
FROM invoice
GROUP BY billing_country
ORDER BY c DESC

```

Results

c	billing_country
131	USA
76	Canada
61	Brazil
50	France
41	Germany
30	Czech Republic
29	Portugal
28	United Kingdom
21	India
13	Ireland
13	Chile
11	Finland
11	Spain
10	Sweden
10	Poland
10	Netherlands
10	U.S.A.

Query executed successfully.

[Fig 5.4.1.2 SQL Project on Music Store Task 2 (Section-1)]

Data_Analytics_Project.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.Music Store (LAPTOP-UU3PCMFC\MEET BHATIA (60)) - Microsoft SQL Server Management Studio

-- Q.3 What are top 3 values of total invoice?

```
select * from Invoice
SELECT TOP 3 total
FROM invoice
ORDER BY total DESC
```

Results

total
23.7599999999999
19.8
19.8

Query executed successfully.

[Fig 5.4.1.3 SQL Project on Music Store Task 3 (Section-1)]

Data_Analytics_Project.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.Music Store (LAPTOP-UU3PCMFC\MEET BHATIA (60)) - Microsoft SQL Server Management Studio

-- Q.4 Which city has the best customers? We would like to throw a promotional Music Festival in the city we made the most sales.
--Write a query that returns one city that has the highest sum of invoice totals.
--Return both the city name & sum of all invoice totals

```
select * from Invoice
SELECT TOP 5 billing_city, SUM(total) AS InvoiceTotal
FROM invoice
GROUP BY billing_city
ORDER BY InvoiceTotal DESC;
```

Results

billing_city	InvoiceTotal
Prague	273.24
Mountain View	169.29
London	166.32
Berlin	158.4
Paris	151.47

Query executed successfully.

[Fig 5.4.1.4 SQL Project on Music Store Task 4 (Section-1)]

Object Explorer

```
-- Q.5 Who is the best customer? The customer who has spent the most money will be declared the best customer.
-- Write a query that returns the person who has spent the most money.

select * from Customer
SELECT TOP 1 customer.customer_id, customer.first_name, customer.last_name, SUM(total) AS total_spending
FROM customer
JOIN invoice ON customer.customer_id = invoice.customer_id
GROUP BY customer.customer_id, customer.first_name, customer.last_name
ORDER BY total_spending DESC;
```

Results

customer_id	first_name	last_name	total_spending
5	František	Wichterlová	144.54

Query executed successfully.

[Fig 5.4.1.5 SQL Project on Music Store Task 5 (Section-1)]

Object Explorer

```
-- Q.1 Write query to return the email, first name, last name, & Genre of all Rock Music listeners.
-- Return your list ordered alphabetically by email starting with A.

select * from Genre
select * from Customer
SELECT DISTINCT email AS Email, first_name AS FirstName, last_name AS LastName, genre.name AS Name
FROM customer
JOIN invoice ON invoice.customer_id = customer.customer_id
JOIN invoiceline ON invoiceline.invoice_id = invoice.invoice_id
JOIN track ON track.track_id = invoiceline.track_id
JOIN genre ON genre.genre_id = track.genre_id
WHERE genre.name LIKE 'Rock'
ORDER BY email;
```

Results

Email	FirstName	LastName	Name
aaronmitchell@yahoo.ca	Aaron	Mitchell	Rock
alex@vol.com.br	Alexandre	Rocha	Rock
astrid.gruber@apple.at	Astrid	Gruber	Rock
bjorn.hansen@yahoo.no	Bjørn	Hansen	Rock
camille.bernard@yahoo.fr	Camille	Bernard	Rock
daan_peeters@apple.be	Daan	Peeters	Rock
diego.gutierrez@yahoo.ar	Diego	Gutiérrez	Rock
dmliller@comcast.com	Dan	Miller	Rock
dominiquelefevre@gmail.com	Dominique	Lefèvre	Rock
edfrancisco@yahoo.ca	Edward	Francis	Rock
eduardo@woodstock.com.br	Eduardo	Martins	Rock
ellie.sullivan@shaw.ca	Ellie	Sullivan	Rock
emma_jones@hotmail.com	Emma	Jones	Rock
enrique_munoz@yahoo.es	Enrique	Muñoz	Rock
femadaromas4@vol.com.br	Fernanda	Ramos	Rock
fharris@google.com	Frank	Harris	Rock
karolina.kowalczyk@wp.pl	Karolina	Kowalczyk	Rock

Query executed successfully.

[Fig 5.4.1.6 SQL Project on Music Store Task 1 (Section-2)]

Data_Analytics_Project.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.Music Store (LAPTOP-UU3PCMFC\MEET BHATIA (60)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Music Store Execute

Object Explorer

Connect Object Explorer

LAPTOP-UU3PCMFC\SQLEXPRESS (SQL Server) Databases System Databases Database Snapshots College Country-State-City CustomersBackup2017 EmployeeDb IPL Database Music Store Database Diagrams Tables System Tables FileTables External Tables Graph Tables dbo.Album dbo.Artist dbo.Customer dbo.Employee dbo.Genre dbo.Invoice dbo.InvoiceLine dbo.MediaType dbo.Playlist dbo.PlaylistTrack dbo.Track Views External Resources Synonyms Programmability Query Store Service Broker Storage Security

Query Editor

-- Q.2 Let's invite the artists who have written the most rock music in our dataset.
-- Write a query that returns the Artist name and total track count of the top 10 rock bands.

```
select * from Artist
select * from Track
SELECT artist.artist_id, artist.name, COUNT(artist.artist_id) AS number_of_songs
FROM track
JOIN album ON album.album_id = track.album_id
JOIN artist ON artist.artist_id = album.artist_id
JOIN genre ON genre.genre_id = track.genre_id
WHERE genre.name LIKE 'Rock'
GROUP BY artist.artist_id, artist.name
ORDER BY number_of_songs DESC
OFFSET 0 ROWS
FETCH NEXT 10 ROWS ONLY;
```

Results Messages

artist_id	name	number_of_songs
1	Led Zeppelin	114
2	U2	112
3	Deep Purple	92
4	Iron Maiden	81
5	Pearl Jam	54
6	Van Halen	52
7	Queen	45
8	The Rolling Stones	41
9	Creedence Clearwater Revival	40
10	Kiss	35

Query executed successfully.

LAPTOP-UU3PCMFC\SQLEXPRESS ... LAPTOP-UU3PCMFC\MEET B... Music Store | 00:00:00 | 10 rows

[Fig 5.4.1.7 SQL Project on Music Store Task 2 (Section-2)]

Data_Analytics_Project.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.Music Store (LAPTOP-UU3PCMFC\MEET BHATIA (60)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Music Store Execute

Object Explorer

LAPTOP-UU3PCMFC\SQLEXPRESS (SQL Server) Databases System Databases Database Snapshots College Country-State-City CustomersBackup2017 EmployeeDb IPL Database Music Store Database Diagrams Tables System Tables FileTables External Tables Graph Tables dbo.Album dbo.Artist dbo.Customer dbo.Employee dbo.Genre dbo.Invoice dbo.InvoiceLine dbo.MediaType dbo.Playlist dbo.PlaylistTrack dbo.Track Views External Resources Synonyms Programmability Query Store Service Broker Storage Security

Query Editor

--Q.3 Return all the track names that have a song length longer than the average song length.
-- Return the Name and Milliseconds for each track. Order by the song length with the longest songs listed first.

```
select * from Track
SELECT name, milliseconds
FROM track
WHERE milliseconds > (
    SELECT AVG(milliseconds) AS avg_track_length
    FROM track
)
ORDER BY milliseconds DESC;
```

Results Messages

name	milliseconds
Occupation / Precipice	5286953
Through a Looking Glass	5088838
Greetings from Earth, Pt. 1	2960293
The Man With Nine Lives	2956998
Battlestar Galactica, Pt. 2	2956081
Battlestar Galactica, Pt. 1	2952702
Murder On the Rising Star	2935394
Battlestar Galactica, Pt. 3	2927802
Take the Celeste	2927877
Fire In Space	2926593
The Long Patrol	2925008
The Magnificent Warriors	2924716
The Living Legend, Pt. 1	2924507
The Gun On Ice Planet Zero, Pt. 2	2924341
The Hand of God	2924007
Experiment In Terra	2923548
More of the Good, Pt. 3	2922581

Query executed successfully.

LAPTOP-UU3PCMFC\SQLEXPRESS ... LAPTOP-UU3PCMFC\MEET B... Music Store | 00:00:00 | 494 rows

[Fig 5.4.1.8 SQL Project on Music Store Task 3 (Section-2)]

SQLQuery1.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.IPL Database (LAPTOP-UU3PCMFC\MEET BHATIA (54)) - Microsoft SQL Server Management Studio

```
--Task 1
select * from ['IPL Matches 2008-2020']
select * from ['IPL Ball-by-Ball 2008-2020']

--Task 2
--SELECT THE TOP 20 ROWS FROM THE IPL MATCHES TABLE
SELECT TOP 20 *
FROM ['IPL Matches 2008-2020']
SELECT TOP 20 *
FROM ['IPL Ball-by-Ball 2008-2020']$
```

Results Messages

	id	city	date	player_of_match	venue	neutral_venue	team1	team2	toss_winner	toss_decision	winner
1	335982	Bangalore	2008-04-18 00:00:00.000	BB McCullum	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Kolkata Knight Riders	Royal Challengers Bangalore	field	Kolkata Knight
2	335983	Chandigarh	2008-04-19 00:00:00.000	MEK Hussey	Punjab Cricket Association Stadium, Mohali	0	Kings XI Punjab	Chennai Super Kings	Chennai Super Kings	bat	Chennai Super Kings
3	335984	Delhi	2008-04-19 00:00:00.000	MF Maharaj	Feroz Shah Kotla	0	Delhi Daredevils	Rajasthan Royals	Rajasthan Royals	bat	Delhi Daredevils
4	335985	Mumbai	2008-04-20 00:00:00.000	MV Boucher	Wankhede Stadium	0	Mumbai Indians	Royal Challengers Bangalore	Mumbai Indians	bat	Royal Challengers Bangalore
5	335986	Kolkata	2008-04-20 00:00:00.000	DJ Hussey	Eden Gardens	0	Kolkata Knight Riders	Deccan Chargers	Deccan Chargers	bat	Kolkata Knight Riders
6	335987	Jaipur	2008-04-21 00:00:00.000	SR Watson	Sawai Mansingh Stadium	0	Rajasthan Royals	Kings XI Punjab	Kings XI Punjab	bat	Rajasthan Royals
7	335988	Hyderabad	2008-04-22 00:00:00.000	V Sehwag	Rajiv Gandhi International Stadium, Uppal	0	Deccan Chargers	Delhi Daredevils	Deccan Chargers	bat	Delhi Daredevils
8	335989	Chennai	2008-04-23 00:00:00.000	ML Hayden	MA Chidambaram Stadium, Chepauk	0	Chennai Super Kings	Mumbai Indians	Mumbai Indians	field	Chennai Super Kings

Query executed successfully.

[Fig 5.4.2.1 SQL Project on IPL Database Task 1-2]

SQLQuery1.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.IPL Database (LAPTOP-UU3PCMFC\MEET BHATIA (54)) - Microsoft SQL Server Management Studio

```
--Task 3
--FETCH THE DATA OF ALL THE MATCHES PLAYED ON ANY DATE FROM THE IPL MATCHES TABLE
SELECT * FROM ['IPL Matches 2008-2020']
WHERE date = '2008-04-23 00:00:00.000'
```

Results Messages

	id	city	date	player_of_match	venue	neutral_venue	team1	team2	toss_winner	toss_decision	winner	result	result_margin	eliminator	method
1	335989	Chennai	2008-04-23 00:00:00.000	ML Hayden	MA Chidambaram Stadium, Chepauk	0	Chennai Super Kings	Mumbai Indians	Chennai Super Kings	field	Chennai Super Kings	6	N	NA	NA

Query executed successfully.

[Fig 5.4.2.2 SQL Project on IPL Database Task 3]

SQLQuery1.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.IPL Database (LAPTOP-UU3PCMFC\MEET BHATIA (54)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

IPL Database Quick Launch (Ctrl+Q)

Object Explorer

--Task 4
--FETCH DATA OF ALL THE MATCHES WHERE THE RESULT MODE IS "RUNS" AND "MARGIN" OF VICTORY IS MORE THEN 100 RUNS
SELECT * FROM ['IPL Matches 2008-2020']
WHERE result = 'runs' AND result_margin > 100;

	id	city	date	player_of_match	venue	neutral_venue	team1	team2	toss_winner	toss_decision	winner
1	335982	Bangalore	2008-04-18 00:00:00.000	BB McCullum	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Kolkata Knight Riders	Royal Challengers Bangalore	field	Kolkata Knight R
2	336038	Mumbai	2008-05-30 00:00:00.000	SR Watson	Wankhede Stadium	0	Deli Daredevils	Rajasthan Royals	Delhi Daredevils	field	Rajasthan Royal
3	501260	Dharamsala	2011-05-17 00:00:00.000	AC Gilchrist	Himachal Pradesh Cricket Association Stadium	0	Kings XI Punjab	Royal Challengers Bangalore	Kings XI Punjab	bat	Kings XI Punjab
4	598027	Bangalore	2013-04-23 00:00:00.000	CH Gayle	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Pune Warriors	Pune Warriors	field	Royal Challenge
5	829785	Bangalore	2015-05-06 00:00:00.000	CH Gayle	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Kings XI Punjab	Kings XI Punjab	field	Royal Challenge
6	980987	Bangalore	2016-05-14 00:00:00.000	AB de Villiers	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Gujarat Lions	Gujarat Lions	field	Royal Challenge
7	108235	Delhi	2017-05-06 00:00:00.000	LMP Simmons	Feroz Shah Kotla	0	Delhi Daredevils	Mumba Indians	Delhi Daredevils	field	Mumba Indians
8	1136601	Kolkata	2018-05-09 00:00:00.000	Ishan Kishan	Eden Gardens	0	Kolkata Knight Riders	Mumba Indians	Kolkata Knight Riders	field	Mumba Indians
9	1175366	Hyderabad	2019-03-31 00:00:00.000	JM Bairstow	Rajiv Gandhi International Stadium, Uppal	0	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	Sunrisers Hyder

Query executed successfully.

[Fig 5.4.2.3 SQL Project on IPL Database Task 4]

SQLQuery1.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.IPL Database (LAPTOP-UU3PCMFC\MEET BHATIA (54)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

IPL Database Quick Launch (Ctrl+Q)

Object Explorer

--Task 5
--FETCH THE DATA OF ALL THE MATCHES WHERE THE FINAL SCORES OF BOTH TEAMS TIED AND ORDER IT IN DESCENDING ORDER OF THE DATE
SELECT * FROM ['IPL Matches 2008-2020']
WHERE result = 'tie' OR result = 'NA'
ORDER BY date
DESC

	id	city	date	player_of_match	venue	neutral_venue	team1	team2	toss_winner	toss_decision	winner
1	1216512	Abu Dhabi	2020-10-18 00:00:00.000	LH Ferguson	Sheikh Zayed Stadium	0	Kolkata Knight Riders	Sunrisers Hyderabad	Sunrisers Hyderabad	field	Kolkata Knight Rid
2	1216517	Dubai	2020-10-18 00:00:00.000	KL Rahul	Dubai International Cricket Stadium	0	Mumba Indians	Kings XI Punjab	Mumba Indians	bat	Kings XI Punjab
3	1216547	Dubai	2020-09-28 00:00:00.000	AB de Villiers	Dubai International Cricket Stadium	0	Royal Challengers Bangalore	Mumba Indians	Mumba Indians	field	Royal Challengers
4	1216493	Dubai	2020-09-20 00:00:00.000	MP Stoinis	Dubai International Cricket Stadium	0	Delhi Capitals	Kings XI Punjab	Kings XI Punjab	field	Delhi Capitals
5	1178426	Mumbai	2019-05-02 00:00:00.000	JJ Bumrah	Wankhede Stadium	0	Mumba Indians	Sunrisers Hyderabad	Mumba Indians	bat	Mumba Indians
6	1178424	Bengaluru	2019-04-30 00:00:00.000	NA	M.Chinnaswamy Stadium	0	Royal Challengers Bangalore	Rajasthan Royals	Rajasthan Royals	field	NA
7	1175365	Delhi	2019-03-30 00:00:00.000	PP Shaw	Feroz Shah Kotla	0	Delhi Capitals	Kolkata Knight Riders	Delhi Capitals	field	Delhi Capitals
8	1082625	Rajkot	2017-04-29 00:00:00.000	KH Pandya	Saurashtra Cricket Association Stadium	0	Gujarat Lions	Mumba Indians	Gujarat Lions	bat	Mumba Indians
9	829813	Bangalore	2015-05-17 00:00:00.000	NA	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	field	NA
10	829763	Bangalore	2015-04-29 00:00:00.000	NA	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Rajasthan Royals	Rajasthan Royals	field	NA
11	829741	Ahmedabad	2015-04-21 00:00:00.000	SE Marsh	Sardar Patel Stadium, Motera	0	Rajasthan Royals	Kings XI Punjab	Kings XI Punjab	field	Kings XI Punjab
12	729315	Abu Dhabi	2014-04-29 00:00:00.000	JP Faulkner	Sheikh Zayed Stadium	1	Kolkata Knight Riders	Rajasthan Royals	Rajasthan Royals	bat	Rajasthan Royals
13	598017	Bangalore	2013-04-16 00:00:00.000	V Kohli	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	field	Royal Challengers
14	598004	Hyderabad	2013-04-07 00:00:00.000	GH Vihari	Rajiv Gandhi International Stadium, Uppal	0	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	bat	Sunrisers Hyder
15	501265	Delhi	2011-05-21 00:00:00.000	NA	Feroz Shah Kotla	0	Delhi Daredevils	Pune Warriors	Delhi Daredevils	bat	NA

Query executed successfully.

[Fig 5.4.2.4 SQL Project on IPL Database Task 5]

SQLQuery1.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.IPL Database (LAPTOP-UU3PCMFC\MEET BHATIA (54)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Object Explorer

SQLQuery1.sql - L...\MEET BHATIA (54)* Data_Analytics_Pr...\MEET BHATIA (60)*

```
--Task 6
--GET THE COUNT OF CITIES THAT HAVE HOSTED AN IPL MATCH
SELECT COUNT (DISTINCT city) FROM ['IPL Matches 2008-2020']
```

139 %

Results Messages

	(No column name)
1	33

Query executed successfully.

LAPTOP-UU3PCMFC\SQLEXPRESS ... LAPTOP-UU3PCMFC\MEET B... IPL Database 00:00:00 1 rows

[Fig 5.4.2.5 SQL Project on IPL Database Task 6]

SQLQuery1.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.IPL Database (LAPTOP-UU3PCMFC\MEET BHATIA (54)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Object Explorer

SQLQuery1.sql - L...\MEET BHATIA (54)* Data_Analytics_Pr...\MEET BHATIA (60)*

```
--Task 7
--WRITE A QUERY TO FETCH THE TOTAL NUMBER OF BOUNDARIES AND DOT BALLS FROM THE BALL-BY-BALL 2008-2020 TABLE
SELECT total_runs, COUNT(*)
FROM ['IPL Ball-by-Ball 2008-2020']
GROUP BY total_runs;
```

139 %

Results Messages

total_runs	(No column name)
1	23723
2	2640
3	257
4	16
5	7467
6	26883
7	170
8	4379

Query executed successfully.

LAPTOP-UU3PCMFC\SQLEXPRESS ... LAPTOP-UU3PCMFC\MEET B... IPL Database 00:00:00 8 rows

[Fig 5.4.2.6 SQL Project on IPL Database Task 7]

SQLQuery1.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.IPL Database (LAPTOP-UU3PCMFC\MEET BHATIA (54)) - Microsoft SQL Server Management Studio

```
--Task 8
--WRITE A QUERY TO FETCH THE TOTAL NUMBER OF BOUNDARIES SCORED BY EACH TEAM FROM THE BALL-BY-BALL 2008-2020 TABLE AND
SELECT * FROM ['IPL Ball-by-Ball 2008-2020']
SELECT batting_team, COUNT(*)
FROM ['IPL Ball-by-Ball 2008-2020']
WHERE total_runs IS NOT NULL
GROUP BY batting_team
ORDER BY COUNT(*) DESC
```

batting_team	(No column name)
Chennai Super Kings	8343
Royal Challengers Bangalore	7998
Deccan Chargers	7867
Mumbai Indians	7809
Kings XI Punjab	7682
Rajasthan Royals	7384
Delhi Daredevils	7320
Kolkata Knight Riders	7162
Pune Warriors	2388
Kochi Tuskers Kerala	1582

Query executed successfully.

[Fig 5.4.2.7 SQL Project on IPL Database Task 8]

SQLQuery1.sql - LAPTOP-UU3PCMFC\SQLEXPRESS.IPL Database (LAPTOP-UU3PCMFC\MEET BHATIA (54)) - Microsoft SQL Server Management Studio

```
--Task 9
--WRITE A QUERY TO FETCH THE TOTAL NUMBER OF DOT BALLS BOWLED BY EACH TEAM AND ORDER IT IN DESCENDING ORDER OF THE
SELECT * FROM ['IPL Ball-by-Ball 2008-2020']
SELECT bowling_team, COUNT(*)
FROM ['IPL Ball-by-Ball 2008-2020']
WHERE total_runs IS NOT NULL
GROUP BY bowling_team
ORDER BY COUNT(*) DESC
```

bowling_team	(No column name)
Royal Challengers Bangalore	8192
Chennai Super Kings	8164
Deccan Chargers	7853
Mumbai Indians	7725
Delhi Daredevils	7548
Kings XI Punjab	7460
Rajasthan Royals	7443
Kolkata Knight Riders	7063
Pune Warriors	2410
Kochi Tuskers Kerala	1614
NA	63

Query executed successfully.

[Fig 5.4.2.8 SQL Project on IPL Database Task 9]

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 analyse 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.

7.3 Date of Continuous Evaluation

7.3.1 Continuous Evaluation 1

- 1St Internship Evaluation Presentation was held on 28th February 2023. Students of 8th Semester 2022-2023 got the comments necessary changes and instruction was given to them.

7.3.2 Continuous Evaluation

- 2nd Internship Evaluation Presentation was held on 15th April 2023. Students of 8th Semester 2022-2023 and Reviewed there Internship work and ppt, gave remarks and some improvement on the projects. And gave information about the Report of the Internship.

7.3.3 Continuous Evaluation 3

- 3rd Internship Evaluation Presentation was held on 4th May 2023. Students of 8th Semester 2022-2023 and Reviewed there Internship work and ppt, gave final remarks and some improvement on the Reports. And gave Instruction about the Report Preparation of the Internship.

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