# MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR

(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)


**Final Year Internship Report**

**on**

## BUSINESS INTELLIGENCE

**Submitted By:**

**Aman**

**0901CS181006**

**Faculty Mentor:**

**Dr. Manish Dixit**

**Professor and Head,**

**Department of Computer Science and Engineering**

## DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE
GWALIOR - 474005 (MP) est. 1957

MAY-JUNE 2022

# MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR

(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)


# BUSINESS INTELLIGENCE

A final year internship report submitted in partial fulfillment of the requirement for the degree of

**BACHELOR OF TECHNOLOGY**

in

**COMPUTER SCIENCE AND ENGINEERING**

Submitted by:

**Aman**

**0901CS181006**

Internship Faculty Mentor:

**Dr. Manish Dixit**

**Professor and Head,**

**Department of Computer Science and Engineering**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE

GWALIOR - 474005 (MP) est. 1957

MAY-JUNE 2022

# Internship Certificate

Infosys | Education, Training and Assessment

CERTIFICATE OF COMPLETION
OF
INTERNSHIP

This is to certify that

Aman

Of

Madhav Institute of Technology & Science, Gwalior,
Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal

has completed the internship program at Infosys Limited

from

January 2022 – April 2022

Satheesha B N

Satheesha B Nanjappa
Vice President and Head, Global Education Center

**MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR**

(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)

# CERTIFICATE

This is certified that **Aman** (0901CS181006) has submitted the Internship report titled **Business Intelligence** of the work he has done under the mentorship of **Dr. Manish Dixit**, in partial fulfilment of the requirement for the award of degree of Bachelor of Technology in Computer Science and Engineering from Madhav Institute of Technology and Science, Gwalior.

**Dr. Manish Dixit**
Faculty Mentor
Professor and Head,
Computer Science and Engineering

27/05/202

**Dr. Manish Dixit**
Professor and Head,
Computer Science and Engineering

Dr. Manish Dixit
Professor & HOD
Department of CSE
M.I.T.S. Gwalior

**MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR**

(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)

# DECLARATION

I hereby declare that the work being presented in this Internship report, for the partial fulfilment of requirement for the award of the degree of Bachelor of Technology in CSE at Madhav Institute of Technology & Science, Gwalior is an authenticated and original record of my work under the mentorship of **Dr. Manish Dixit**, **Professor, and Head**, Department of CSE.

I declare that I have not submitted the matter embodied in this report for the award of any degree or diploma anywhere else.

Aman

Aman

0901CS181006

IV Year,

Computer Science and Engineering

MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE, GWALIOR

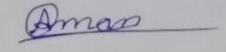
(A Govt. Aided UGC Autonomous & NAAC Accredited Institute Affiliated to RGPV, Bhopal)

# ACKNOWLEDGEMENT

The full semester internship has proved to be pivotal to my career. I am thankful to my institute, **Madhav Institute of Technology and Science** to allow me to continue my disciplinary/interdisciplinary internship as a curriculum requirement, under the provisions of the Flexible Curriculum Scheme (based on the AICTE Model Curriculum 2018), approved by the Academic Council of the institute. I extend my gratitude to the Director of the institute, **Dr. R. K. Pandit,** and Dean Academics, **Dr. Manjaree Pandit** for this.

I would sincerely like to thank my department, **Department of Computer Science and Engineering, for allowing** me to explore this internship. I humbly thank **Dr. Manish Dixit**, Professor, and Head, Department of Computer Science and Engineering, for his continued support during the course of this engagement, which eased the process and formalities involved.

I am sincerely thankful to my faculty mentors. I am grateful to the guidance of **Dr. Manish Dixit**, Professor and Head, Department of Computer Science and Engineering, for his continued support and close mentoring throughout the internship. I am also very thankful to the faculty and staff of the department.



Aman

0901CS181006

IV Year,

Computer Science and Engineering

# ABSTRACT

In terms of academic and non-academic viewpoints, the internship report outlines the significant work completed by me. The purpose of this document is to identify and summarise the research conducted, projects accomplished, and experience acquired, with a focus on the intern's accomplishments. I have completed my internship as a Business Intelligence Intern at Infosys in which I've learned and used BI tools for report generation.

Business intelligence (BI) is a word that refers to a combination of technologies, structures, procedures, applications, and processes for collecting, integrating, analyzing, and presenting business data in order to drive profitable business decisions. Business intelligence is thought to have a big impact on companies. The basic goal of business intelligence is to help people make better decisions. BI tools aid in data analysis and the creation of summaries, dashboards, maps, graphs, charts, and reports that give users detailed intelligence about the business. A well-implemented Extract, Transform, and Load (ETL) process is an important component of BI systems.

There are four chapters in this report. The first chapter of this report covers the topic, objectives, and scope of the internship, as well as an overview of the industry. The second chapter of this report deals with the literature review and some theoretical concepts involved during my internship. The third chapter is about project analysis which contains a detailed description of my project. The fourth chapter is about the project overview which contains results, limitations, conclusion, etc.

**Keywords:** Business Intelligence, BI, Extract, transform, and Load, ETL, SSAS, SSRS, SSMS, SSIS

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS

| Abbreviation | Description |
|---|---|
| **BI** | Business Intelligence |
| **ETL** | Extract, transform, and load |
| **OLAP** | Online Analytic Processing |
| **API** | Application Programming Interface |
| **CRM** | Customer Relationship Management |
| **ERP** | Enterprise Resource Planning |
| **SSMS** | SQL Server Management Studio |
| **SSIS** | SQL Server Integration Services |
| **SSAS** | SQL Server Analysis Services |
| **SSRS** | SQL Server Reporting Services |
| **OI** | Operational Intelligence |

# Chapter 1: Introduction

The technology that enables data preparation, data mining, data management, and data visualization is referred to as business intelligence (BI). End-users can identify relevant information from raw data using business intelligence tools and methods, making data-driven decision-making easier in a variety of industries. Dashboarding and data visualization are anticipated in business intelligence tools. Machine learning and AI are being integrated into corporate intelligence platforms to stay competitive. Data warehouses, ETL, and OLAP are the foundations of their systems.

It's worth noting that this is a relatively new definition of BI, and the term has a tangled history as a buzzword. Traditional Business Intelligence, complete with capital letters, first appeared in the 1960s as a framework for exchanging data across enterprises. It evolved with computer models for decision-making and converting data into insights in the 1980s before becoming a particular product from BI teams with IT-reliant service solutions. Flexible self-service analysis, controlled data on trustworthy platforms, empowered business users, and speed to insight are all priorities in modern BI solutions.

Much more than a specific “thing,” business intelligence is rather an umbrella term that covers the processes and methods of collecting, storing, and analyzing data from business operations or activities to optimize performance. All of these things come together to create a comprehensive view of a business to help people make better, actionable decisions. Over the past few years, business intelligence has evolved to include more processes and activities to help improve performance.

In this internship, I’ve learned the methods involved in Business Intelligence. I’ve also implemented those methods to develop many reports from raw data.

## 1.1 Objectives

This internship has the following goals: -

- Learn and understand the methods involved in Business Intelligence.
- Learn how to gather, analyze, and evaluate data for relevant departments in order to make informed decisions in the face of ambiguity in order to achieve the corporate goal.
- Learn how to synthesize useful knowledge from collections of data.
- Develop user-friendly reports using BI tools.

## 1.2 Scope

This report summarises the activities I participated in during my internship at Infosys Limited's Business Intelligence department. The study of reporting and information processing is not the primary focus of this report. It also covers the working environment and working area of the industry. I have used data from many primary and secondary sources to summarize all the knowledge I've gained from the internship.

## 1.3 Feasibility

Many businesses are operating in very competitive economic conditions, where they must constantly lower expenses and implement the most intelligent business strategies in order to stay afloat.

If certain factors are taken into account, an investment in a corporate IT project, such as the deployment of a Business Intelligence approach or any other Enterprise Information Systems perspective, can be lucrative for the investor.

The value generated and supplied for the organization's shareholders will be measured using the established BI value chain, which will discover chances to enhance revenue, cut expenses, and improve asset utilisation. Process performance (budget, time schedule) and infrastructure performance are used to assess the success of BI system implementations (system quality, information quality, system use).

A Feasibility study is the first step in a business-driven approach to BI project deployment. The decision-making process for huge projects is quite complex, and this paper will not go into detail about it. A feasibility analysis using the Monte Carlo simulation approach will be undertaken with a middle-sized BI project in mind. Project management best practises, according to Gonzalez (2009), propose the most appropriate probabilistic, statistical, and simulation methods for project analysis.

## 1.4 System Requirement

**Hardware Requirements**

| | | |
|---|---|---|
| RAM | : | 4GB |
| Processor | : | Intel i3 Processor |
| Hard Disk | : | 500 GB HDD |
| Monitor | : | 12" minimum needed |

**Software Requirements**

| | | |
|---|---|---|
| Operating System | : | Windows 7/8/10 |
| Browser | : | Google Chrome/Firefox |
| Language | : | Python, MySQL, MongoDB |

**Extra Software Required**: Visual Studio Code, Anaconda, Jupytor Notebook

## 1.5 Organization Overview

### 1.5.1 Introduction

Infosys Limited is an Indian IT firm that specialises in worldwide business consulting and information technology services. Infosys assists clients in 45 countries in developing and implementing various digital transformation initiatives. Infosys assists organisations in renewing and improving existing circumstances so that they may achieve better efficiency and remain competitive in today's market. Infosys has over 200,000 workers and has developed to a US $10.9 billion (revenues FY18) company with a market valuation of US $39 billion because to their hard work and devotion.

**Mission**

"To achieve our objectives in an environment of fairness, honesty, and courtesy towards our clients, employees, vendors, and society"

**Products & Services Offered by Infosys**

Infosys Offers a Variety of Products and Services Infosys Offers a Variety of Products and Services Infosys works with a wide range of firms in a variety of industries, including insurance, finance, manufacturing, and others, to build and maintain software. • Mana, which is now known as NIA - Next Generation Integrated AI Platform - is one of the most essential work-related platforms utilised by Infosys.

• Finacle, a banking system with different modules relating to corporate and retail banking; Infosys Information Platform, the company's own analytics platform; and Infosys Consulting, a worldwide management consulting service.

### 1.5.2 Organization History

N. R. Narayana Murthy and a team of six other engineers founded Infosys Limited, formerly known as Infosys Technologies Limited, in Pune, India, in 1981 with an initial capital investment of only US $250. Infosys went public in 1993 and started an employee stock option programme. Infosys relocated its headquarters from Pune to Bangalore, India, in 1994. Infosys became the first Indian IT business to be listed on NASDAQ in 1999, making it the most expensive stock on the Indian market at the time. In 1999, Infosys was one of the NASDAQ's top 20 businesses by market capitalization. In 1999, Infosys' yearly revenue was US$100 million, and in 2004, it was US$1 billion.

### 1.5.3 Organization Heads

1. **N.R. Narayana Murthy** is the founder of Infosys.
2. **Salil S. Parekh** is the CEO and managing director of Infosys.
3. **Pravin Rao** is the chief operating officer of Infosys.

# Chapter 2: Literature Review

In this report with the aid of several papers and secondary sources of data collecting, I used secondary information from the prior literature study on Business Intelligence. This literature review contains Business Intelligence Architecture, its components, and the tools I've used during my internship period.

## 2.1 Business Intelligence

The quality of corporate decision-making is entirely based on high-quality business data. This is a truth in today's competitive business climate, which necessitates quick, accurate, and relevant data insights from a data storage warehouse arranged in such a way as to improve business performance. Data warehousing is the backbone of these procedures, and business intelligence architecture has arisen to suit those needs.

The rapid increase in unstructured data created by IoT devices, smart cities, sensors, cameras, and other devices is driving the expansion of the Business Intelligence market size.

Increasing need for data visualization dashboards to aid in making educated decisions that have a significant influence on corporate performance.

Increasing analytics spending and usage of cloud-based BI products.

## 2.2 Business Intelligence Tools

BI tools are specialized software that collects, processes, and analyses data, as well as generates reports and dashboards, to provide meaningful business insights. On-Line Analytical Processing (OLAP), predictive and enhanced analytics and much more are all possible with BI technologies.

BI tools used to be confined to querying and producing reports, which didn't help much with making quick choices. Modern BI solutions, such as Tableau and PowerBI, are more adaptable and versatile, allowing users to develop actionable insights, create reports and dashboards, and create visualizations and performance scorecards to present KPIs and business indicators.

**SQL Server Management Studio (SSMS)**

SQL Server Management Studio (SSMS) provides a unified management environment for any SQL system. All components of SQL Server, Azure SQL Database, Azure SQL Managed Instance, SQL Server on Azure VM, and Azure Synapse Analytics may be accessed, configured, managed, administered, and developed using SSMS. SSMS is a single, complete solution that combines a wide range of graphical tools with a number of

powerful script editors to provide developers and database administrators of all skill levels access to SQL Server.

**SQL Server Management Studio for Business Intelligence**

SQL Server Management Studio is used to access, setup, manage, and administer Analysis Services, Integration Services, and Reporting Services. Although SQL Server Management Studio is used by all three business intelligence solutions, the administrative chores associated with each are slightly different.

**MongoDB**

MongoDB is an open-source document-oriented database that is the most popular NoSQL database. 'NoSQL' stands for 'non-relational'. It indicates that MongoDB doesn't use a table-like relational database structure to store and retrieve data, but instead uses a completely other technique. BSON is the name of the storage format (similar to JSON format).

**MongoDB Connector for BI**

Flat, tabular data is what traditional business intelligence tools operate with. These technologies aren't capable of comprehending three-dimensional data stored in MongoDB databases. The MongoDB Connector for Business Intelligence (BI) allows you to use relational business intelligence products like Tableau and Power BI to create SQL queries to display, graph, and report on your three-dimensional MongoDB data. The MongoDB Connector for BI works as a layer between a mongod or mongos instance and your reporting tool, translating queries and data. The BI Connector does not store any data; its sole purpose is to connect your MongoDB cluster to business intelligence tools.

**SQL Server Integration Services (SSIS)**

The data-warehousing arm of the SQL Server 2008 R2 package is SQL Server Integration Services. It comes with built-in ETL capabilities. This facilitates the transfer of data from one platform to another, as well as the modification of data if necessary.

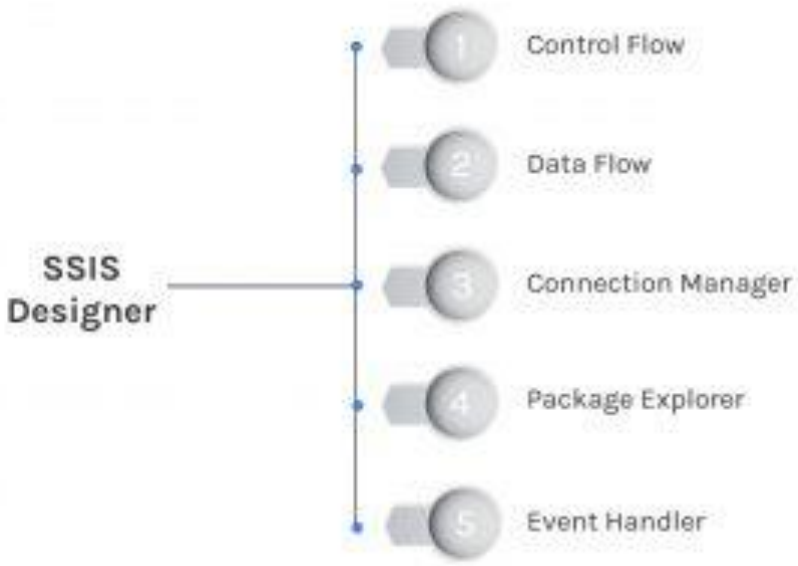



Fig 1: SSIS Designer

The SSIS Designer is a part of the Business Intelligence Development Studio that allows you to create and maintain integration service packages. The SSIS API Programming module enables you to write SSIS packages in any programming language. SSIS is a component that Microsoft SQL Server employs to perform a variety of data movement activities. This is a data warehousing tool for extracting, loading, and manipulating data that is exceptionally quick and adaptable. It also aids in data merging, cleansing, and aggregation.

**SQL Server Analysis Services (SSAS)**

The SSAS is a multidimensional analysis tool that does analytical processing in real time. The SSAS assists business information with a relational database by providing the most sophisticated data mining capabilities and deeper dimensions. Multidimensional analysis is an OLAP approach that allows you to examine enormous amounts of data by storing it in axes and cells rather than the typical relational 2D representation of rows and columns.

SSAS gives IT workers predictive analytic skills by allowing them to access to backend data using standard programs like Microsoft Excel and SharePoint for analysis, visual display, and collaboration.

SSAS is a technology that is used by a variety of companies to evaluate data that is dispersed across many databases. Microsoft has incorporated various business intelligence and data warehousing features in SQL Server.

**SQL Server Reporting Services (SSRS)**

The SQL Server Reporting Services (SSRS) is a framework of reporting mechanisms that operate together through a web interface to enable the construction of succinct interactive reporting solutions in print or web format.


Fig 2: SSRS Architecture

Two SSRS components for creating the most efficient reports are the report builder and report designer. Report builder is a basic tool that allows an information worker or business user to quickly construct reports without having to grasp the data's underlying structure. A report designer is a developer's tool since it complicates the creation of bespoke reports.

To utilise this tool, you must first grasp the business intelligence development studio and the data's underlying structure. The report server, according to Microsoft, is the primary process engine in SSRS that manages the processing and distribution of reports via processors.

The SSRS service connects to Microsoft Visual Studio, allowing developers and SQL administrators to quickly access to the SQL database and generate SQL reports using SSRS tools. It also has a "Report Builder" tool that allows technical users to construct SQL reports more easily.

# Chapter 3: Project Analysis

**Project: - Advanced ETL Pipeline**

For analysis, reporting, and data synchronization, an ETL pipeline is a series of operations that take data from an input source, convert it, and load it into an output destination such as a data mart, database, or data warehouse.

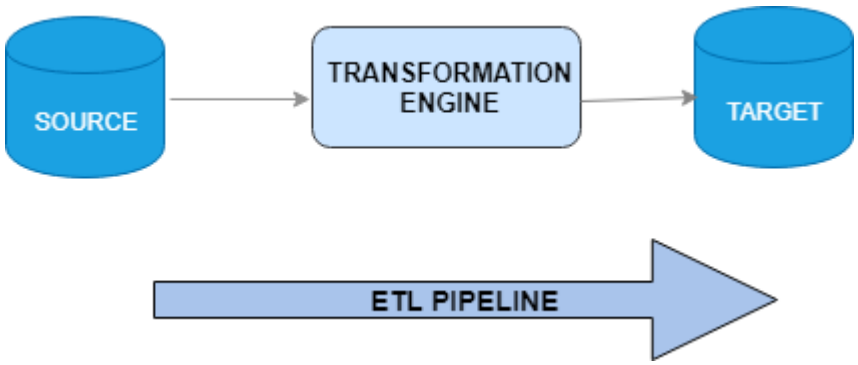



Fig 3: ETL Pipeline

ETL stands for Extract, Transform, and load.

**Extract**

The initial stage in this process is to extract data from heterogeneous sources such as business systems, APIs, sensor data, marketing tools, transaction databases, and other sources. As you can see, some of these data types are semi-structured JSON server logs, while others are structured outputs of frequently used systems.

**Transform**

The raw data gathered from the sources is then transformed into a format that can be used by various applications in the second stage. Data is cleaned, mapped, and converted at this step, frequently to a specified schema, in order to fulfill operational requirements. Several sorts of transformations are used in this process to assure data quality and integrity. Instead of loading data directly into the target data source, it is more convenient to upload information to a staging database. This step provides a speedy reversal in the event that something goes wrong. You can prepare audit reports for regulatory compliance during this stage, as well as identify and correct any data errors.

**Load**

Finally, the load function is the process of transferring transformed data from a staging area to a destination database that may or may not have existed earlier. This method might be straightforward or complex, depending on the application's needs. ETL tools or custom code can be used to complete each of these processes.

# Chapter 4: Project Overview

During the internship, I was allocated to a team of 6 members. We were assigned to a project named "Advanced Etl Pipeline". The main features and aim of the project are: -

- ETL Pipeline refers to a set of processes that extract the data from an input source, transform the data, and loading into an output destination.

- ETL stands for Extract, Transform and Load.

- It is helpful in analysis, reporting, and data synchronization

- The main aim of this project is to analyze the given dataset and to represent the data in a proper way visually (data visualization) using Python and its packages.

**Task Performed**

- Created different user stories to determine the requirements of the project in an agile sheet.
- Extracted the raw data provided in the form of datasets.
- Performed different transformations using python based on user stories.
- Delivered the report in the form of data visualization for better understanding.

**Software Requirement**

- Anaconda
- Jupiter Notebook
- Python
- Agile Scrum

**Packages Used**

- Pandas: is an open-source Python package that is most widely used for data science/data analysis and machine learning tasks.
- NumPy: is a Python Library used for working with arrays. It is an open-source project and we can use it freely.
- Matplotlib: is a comprehensive library for creating static, animated, and interactive visualization in Python.

- Seaborn: is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

**Data Sources**

| match_id | inning | batting_te | bowling_t | over | ball | batsman | non_strik | bowler | is_super_ | wide_run | bye_runs | legbye_ru | noball_ru | penalty_r | batsman_ | extra_run | total_runs | player_dis | dismissal_ | fielder | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | Sunrisers | Royal Cha | 1 | 1 | DA Warne | S Dhawan | TS Mills | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 1 | 2 | DA Warne | S Dhawan | TS Mills | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 1 | 3 | DA Warne | S Dhawan | TS Mills | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 4 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 1 | 4 | DA Warne | S Dhawan | TS Mills | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 1 | 5 | DA Warne | S Dhawan | TS Mills | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 1 | 6 | S Dhawan | DA Warne | TS Mills | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 1 | 7 | S Dhawan | DA Warne | TS Mills | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 2 | 1 | S Dhawan | DA Warne | A Choudh | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 2 | 2 | DA Warne | S Dhawan | A Choudh | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 4 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 2 | 3 | DA Warne | S Dhawan | A Choudh | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 2 | 4 | DA Warne | S Dhawan | A Choudh | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 6 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 2 | 5 | DA Warne | S Dhawan | A Choudh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | DA Warne | caught | Mandeep Singh | |
| 1 | 1 | Sunrisers | Royal Cha | 2 | 6 | MC Henric | S Dhawan | A Choudh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 2 | 7 | MC Henric | S Dhawan | A Choudh | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 4 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 3 | 1 | S Dhawan | MC Henric | TS Mills | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 3 | 2 | MC Henric | S Dhawan | TS Mills | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 3 | 3 | MC Henric | S Dhawan | TS Mills | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 3 | 4 | MC Henric | S Dhawan | TS Mills | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 3 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 3 | 5 | S Dhawan | MC Henric | TS Mills | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 3 | 6 | MC Henric | S Dhawan | TS Mills | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 4 | 1 | MC Henric | S Dhawan | YS Chahal | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 4 | 2 | MC Henric | S Dhawan | YS Chahal | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 4 | 3 | S Dhawan | MC Henric | YS Chahal | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 4 | 4 | S Dhawan | MC Henric | YS Chahal | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | | | | |
| 1 | 1 | Sunrisers | Royal Cha | 4 | 5 | MC Henric | S Dhawan | YS Chahal | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | | | | |

Table 1: Deliveries dataset

| id | season | city | date | team1 | team2 | toss_winn | toss_decis | result | dl_applie | winner | win_by_r | win_by_w | player_of | venue | umpire1 | umpire2 | umpire3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2017 | Hyderaba | 4/5/2017 | Sunrisers | Royal Cha | Royal Cha | field | normal | 0 | Sunrisers | 35 | 0 | Yuvraj Sin | Rajiv Gan | AY Dandel | NJ Llong | |
| 2 | 2017 | Pune | 4/6/2017 | Mumbai In | Rising Pur | Rising Pur | field | normal | 0 | Rising Pur | 0 | 7 | SPD Smith | Maharash | A Nand Ki | S Ravi | |
| 3 | 2017 | Rajkot | 4/7/2017 | Gujarat Li | Kolkata Kr | Kolkata Kr | field | normal | 0 | Kolkata Kr | 0 | 10 | CA Lynn | Saurashtra | Nitin Men | CK Nandan | |
| 4 | 2017 | Indore | 4/8/2017 | Rising Pur | Kings XI P | Kings XI P | field | normal | 0 | Kings XI P | 0 | 6 | GJ Maxwe | Holkar Cri | AK Chaudl | C Shamshuddin | |
| 5 | 2017 | Bangalore | 4/8/2017 | Royal Cha | Delhi Dare | Royal Cha | bat | normal | 0 | Royal Cha | 15 | 0 | KM Jadhav | M Chinnaswamy Stadium | | | |
| 6 | 2017 | Hyderaba | 4/9/2017 | Gujarat Li | Sunrisers | Sunrisers | field | normal | 0 | Sunrisers | 0 | 9 | Rashid Kh | Rajiv Gan | A Deshmu | NJ Llong | |
| 7 | 2017 | Mumbai | 4/9/2017 | Kolkata Kr | Mumbai In | Mumbai In | field | normal | 0 | Mumbai In | 0 | 4 | N Rana | Wankhed | Nitin Men | CK Nandan | |
| 8 | 2017 | Indore | ######## | Royal Cha | Kings XI P | Royal Cha | bat | normal | 0 | Kings XI P | 0 | 8 | AR Patel | Holkar Cri | AK Chaudl | C Shamshuddin | |
| 9 | 2017 | Pune | ######## | Delhi Dare | Rising Pur | Rising Pur | field | normal | 0 | Delhi Dare | 97 | 0 | SV Samso | Maharash | AY Dandel | S Ravi | |
| 10 | 2017 | Mumbai | ######## | Sunrisers | Mumbai In | Mumbai In | field | normal | 0 | Mumbai In | 0 | 4 | JJ Bumrah | Wankhed | Nitin Men | CK Nandan | |
| 11 | 2017 | Kolkata | ######## | Kings XI P | Kolkata Kr | Kolkata Kr | field | normal | 0 | Kolkata Kr | 0 | 8 | SP Narine | Eden Gard | A Deshmu | NJ Llong | |
| 12 | 2017 | Bangalore | ######## | Royal Cha | Mumbai In | Mumbai In | field | normal | 0 | Mumbai In | 0 | 4 | KA Pollard | M Chinna | KN Ananth | AK Chaudhary | |
| 13 | 2017 | Rajkot | ######## | Rising Pur | Gujarat Li | Gujarat Li | field | normal | 0 | Gujarat Li | 0 | 7 | AJ Tye | Saurashtra | A Nand Ki | S Ravi | |
| 14 | 2017 | Kolkata | ######## | Kolkata Kr | Sunrisers | Sunrisers | field | normal | 0 | Kolkata Kr | 17 | 0 | RV Uthapp | Eden Gard | AY Dandel | NJ Llong | |
| 15 | 2017 | Delhi | ######## | Delhi Dare | Kings XI P | Delhi Dare | bat | normal | 0 | Delhi Dare | 51 | 0 | CJ Anders | Feroz Sha | YC Barde | Nitin Menon | |
| 16 | 2017 | Mumbai | ######## | Gujarat Li | Mumbai In | Mumbai In | field | normal | 0 | Mumbai In | 0 | 6 | N Rana | Wankhed | A Nand Ki | S Ravi | |
| 17 | 2017 | Bangalore | ######## | Rising Pur | Royal Cha | Royal Cha | field | normal | 0 | Rising Pur | 27 | 0 | BA Stokes | M Chinna | KN Ananth | C Shamshuddin | |
| 18 | 2017 | Delhi | ######## | Delhi Dare | Kolkata Kr | Delhi Dare | bat | normal | 0 | Kolkata Kr | 0 | 4 | NM Coulte | Feroz Sha | Nitin Men | CK Nandan | |
| 19 | 2017 | Hyderaba | ######## | Sunrisers | Kings XI P | Kings XI P | field | normal | 0 | Sunrisers | 5 | 0 | B Kumar | Rajiv Gan | AY Dandel | A Deshmukh | |
| 20 | 2017 | Rajkot | ######## | Royal Cha | Gujarat Li | Gujarat Li | field | normal | 0 | Royal Cha | 21 | 0 | CH Gayle | Saurashtra | S Ravi | VK Sharma | |
| 21 | 2017 | Hyderaba | ######## | Sunrisers | Delhi Dare | Sunrisers | bat | normal | 0 | Sunrisers | 15 | 0 | KS Willian | Rajiv Gan | CB Gaffan | NJ Llong | |
| 22 | 2017 | Indore | ######## | Kings XI P | Mumbai In | Mumbai In | field | normal | 0 | Mumbai In | 0 | 8 | JC Buttler | Holkar Cri | M Erasmu | C Shamshuddin | |
| 23 | 2017 | Kolkata | ######## | Kolkata Kr | Gujarat Li | Gujarat Li | field | normal | 0 | Gujarat Li | 0 | 4 | SK Raina | Eden Gard | CB Gaffan | Nitin Menon | |
| 24 | 2017 | Mumbai | ######## | Mumbai In | Delhi Dare | Delhi Dare | field | normal | 0 | Mumbai In | 14 | 0 | MJ McCler | Wankhed | A Nand Ki | S Ravi | |
| 25 | 2017 | Pune | ######## | Sunrisers | Rising Pur | Rising Pur | field | normal | 0 | Rising Pur | 0 | 6 | MS Dhoni | Maharash | AY Dandel | A Deshmukh | |

Table 2: Matches dataset

## 4.1 Results

After applying ETL to the datasets the result in the form of data visualization are:

**TASK: Visualization of the Dismissal kind of the Player**

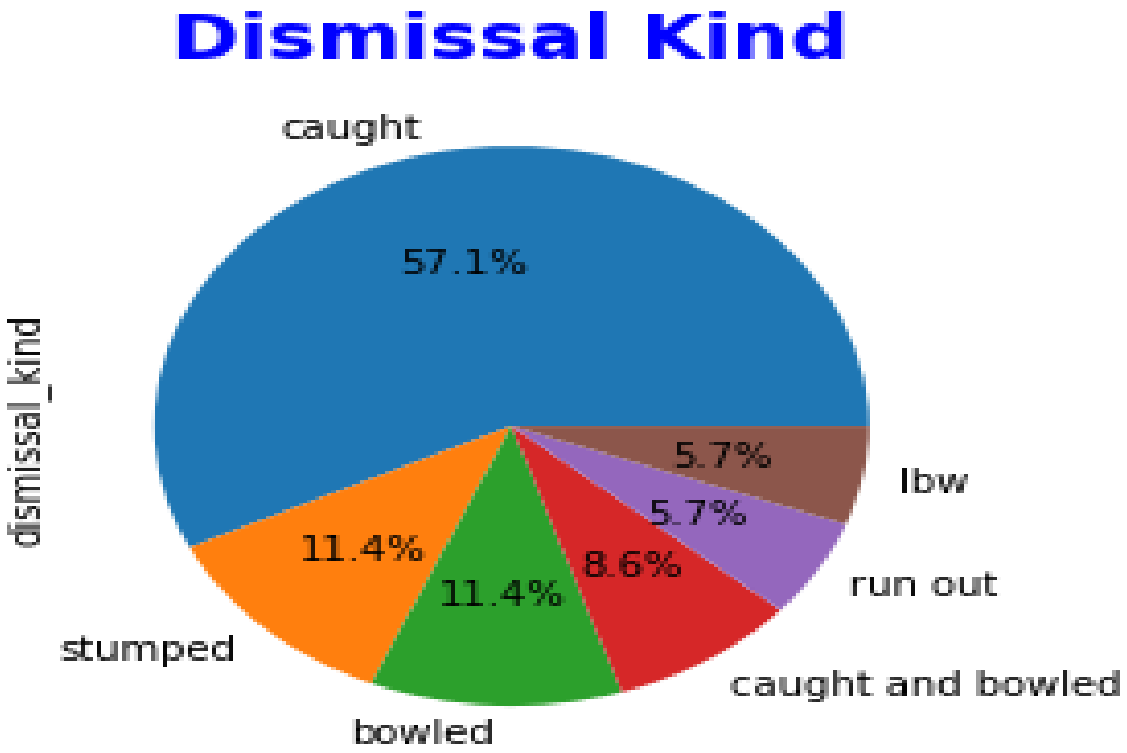



Fig 4: Dismissal Kind of the Player

**TASK: Box Plot Representation of the Runs Scored by each Team**

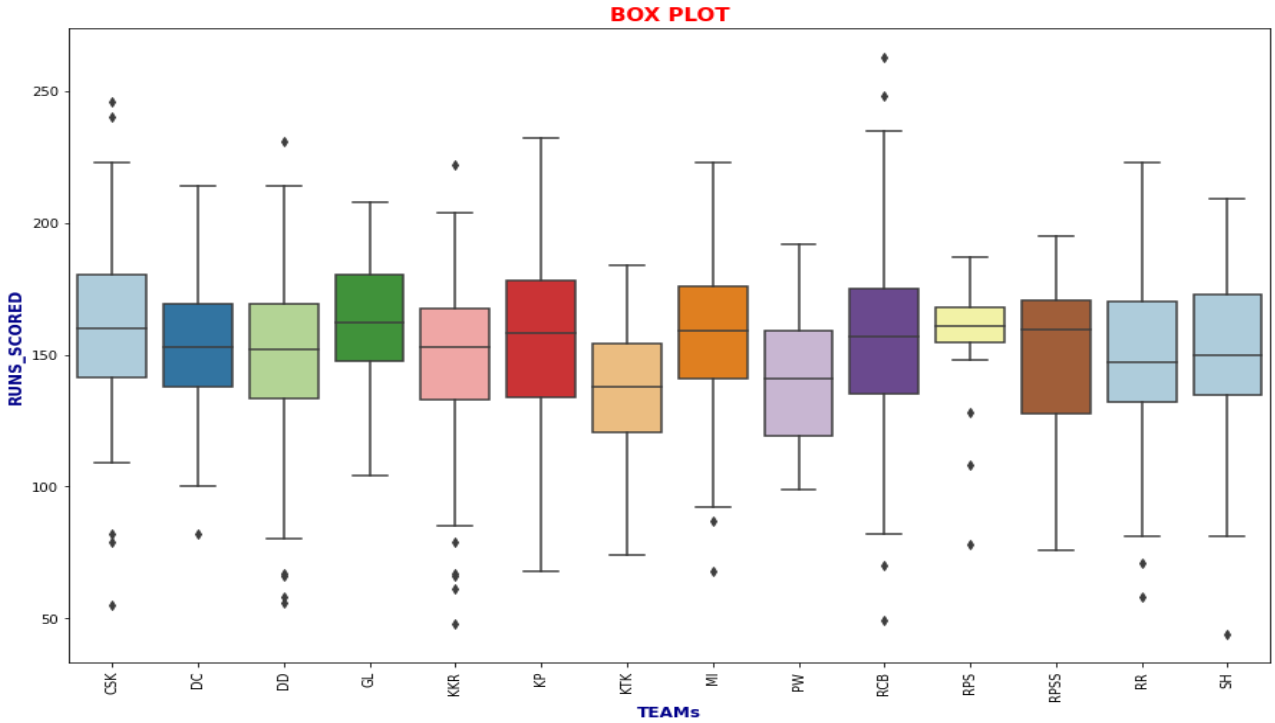



Fig 5: Runs scored by each team

**TASK: Top 10 Wicket Taking Bowlers**

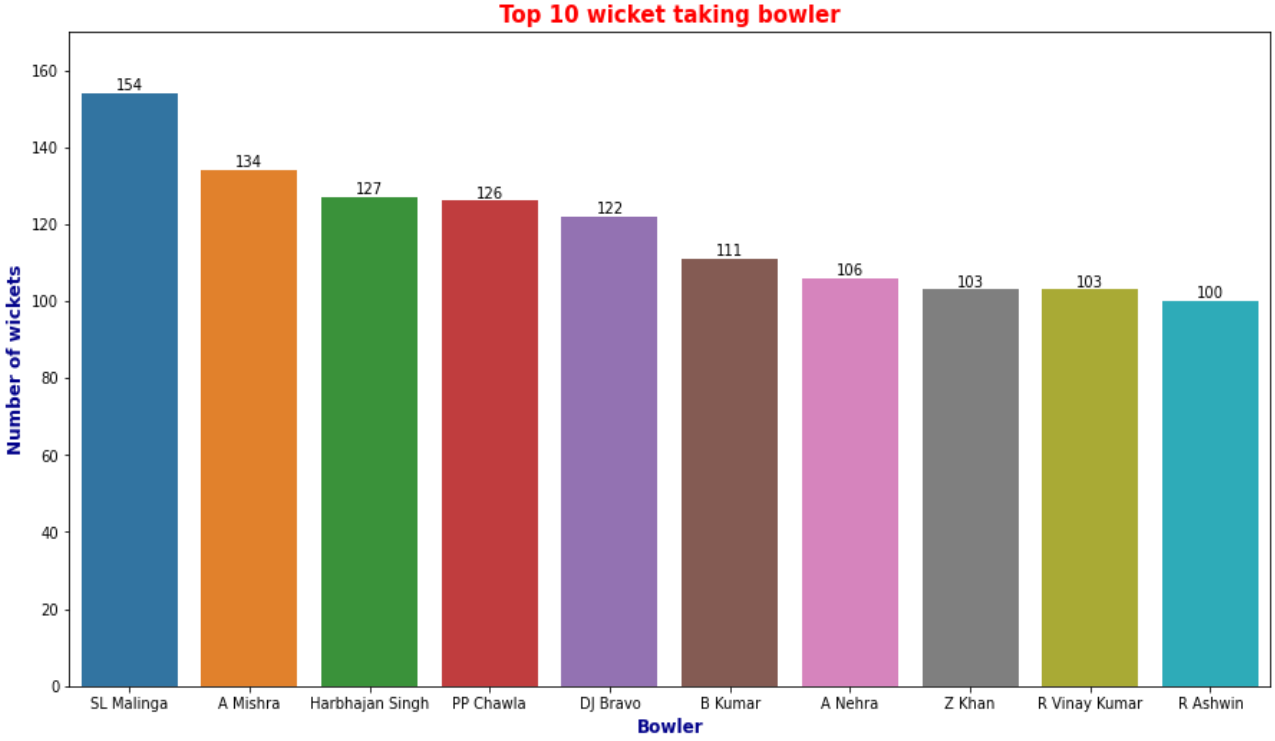



Fig 6: Top 10 wicket taking bowlers

**TASK: Double Hundreds Scored by Teams**

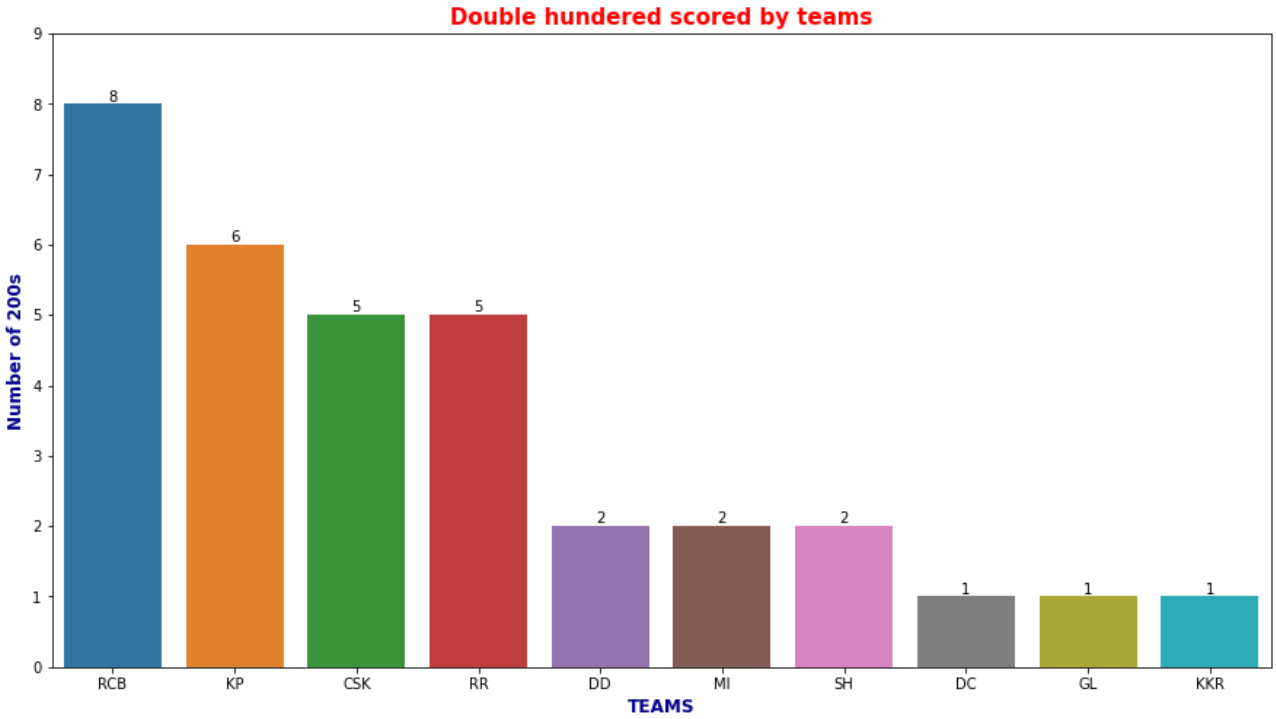



**Fig 7: Double Hundreds Scored by Teams**

**TASK: Top Bowlers with Low Economy Rate**

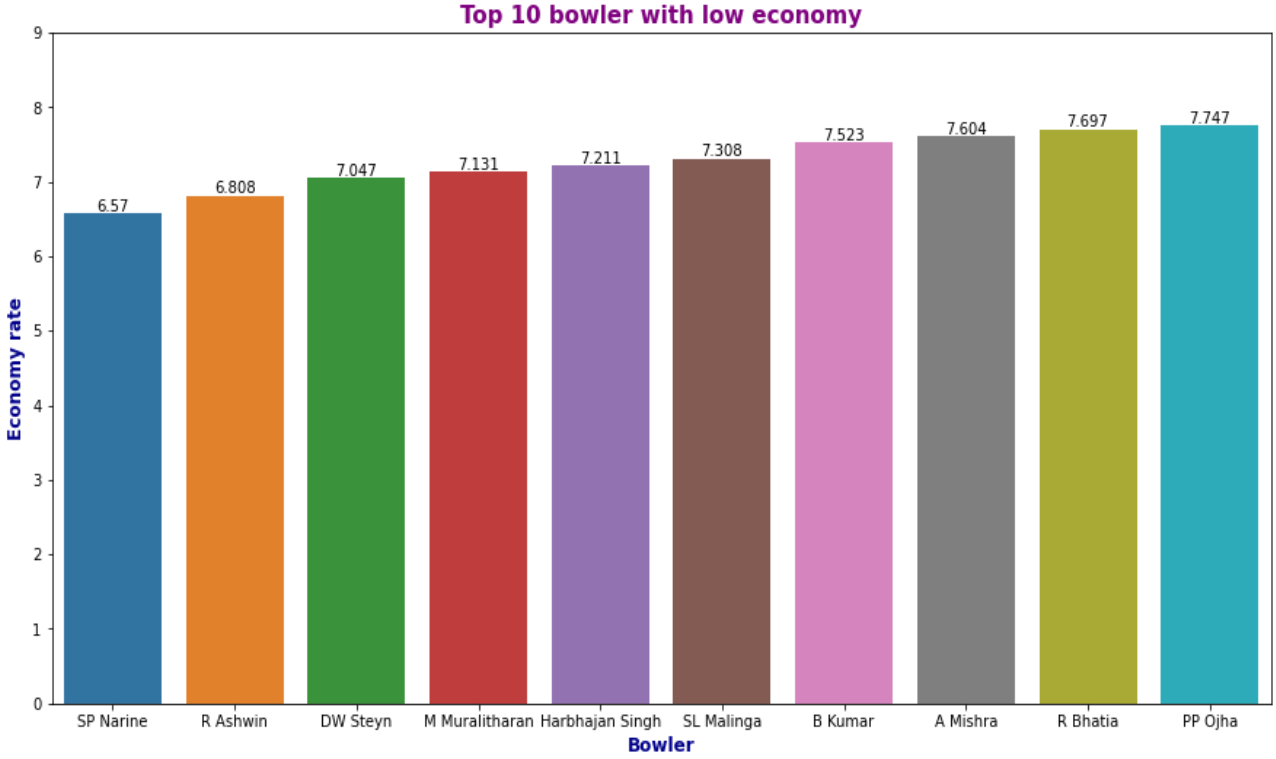



**Fig 8: Top Bowlers with Low Economy Rate**

## 4.2 Application

In BI, a wide range of applications are employed. Here are a few examples:

• Analyses ad hoc: Ad hoc querying is another name for it. It's a crucial component in today's business intelligence software. Queries are developed and performed in ad hoc analysis to investigate specific business-related challenges. The analytic findings are included in dashboards and reports.

• Mobile BI: It enables the use of BI apps on mobile devices. These are designed to be simple to use and are mostly used to see rather than analyze data.

• Real-time BI: This type of analysis examines data as it is being gathered, produced, and analyzed. This provides users with an up-to-date snapshot of all of their areas of interest.

## 4.3 Problems Faced

- Data loss due to VM maintenance.
- Develop the whole project a second time due to data loss.

## 4.4 Limitations

- Not allowed to use a different software other than ones provided by the company or installed by the company.
- Not able to download files sent through emails during working hours.
- Couldn’t download the packages required for our project.
- Use of only limited resources provided by the company.

## 4.5 Conclusion

From my internship at Infosys, I was able to get a better understanding of how aspects of BI directly or indirectly influence the standard of managerial decision-making. It also taught me how Business Intelligence benefits many organizations by improving flexible reporting and analysis via maintaining higher data quality, and better decision making.

The unwavering involvement and devotion of the organization's business intelligence is critical to its success. In today's ever-changing business climate, BI managers' task has grown significantly more tough in order to compete in both worldwide and local marketplaces in bad scenarios.

Overall, I enjoyed my Business Intelligence internship, and I am certain that I will be able to apply the skills I acquired later in my career.

# Reference


- IBM Analytics
- The enterprise data warehouse of the future by Doug Dailey March 11, 2019 – Blog
- Literature Review of Business Intelligence Shubhi Jain, Anu Sharma M. Tech Student, Assistant Professor Teerthanker Mahaveer University, Moradabad, Volume 12, Issue 5, 2021
- https://www.datapine.com/blog/data-warehousing-and-business-intelligencearchitecture/
- https://www.prnewswire.com/news-releases/business-intelligence-bi-market-size-toreach-usd-27870-million-by-2026-at-cagr-5-4--valuates-reports301247353.html#:~:text=%2D%20Business%20Intelligence%20Platform%20market%20size,t he%20forecast%20period%202021%2D2026
- https://searchbusinessanalytics.techtarget.com/definition/business-intelligencearchitecture
- https://www.talend.com/resources/what-is-data-integration/
- https://olap.com/olapdefinition/#:~:text=OLAP%20(Online%20Analytical%20Processing)%20is,(budget%2C%20for ecast)%20planning.
- https://intellipaat.com/blog/what-is-business-intelligence/
- https://brightdata.com/blog/brightdata-in-practice/etl-pipeline?kw=&cpn=14745430544&cam=aw_all_products-all_geos-search_dsa_blog-kw_en-desktop_blog-brightdata-in-practice__592559539359&utm_term=&utm_campaign=all_products-all_geos-search_dsa_blog-kw_en-desktop&utm_source=adwords&utm_medium=ppc&utm_content=blog-brightdata-in-practice&hsa_acc=1393175403&hsa_cam=14745430544&hsa_grp=136943771713&hsa_ad=592559539359&hsa_src=g&hsa_tgt=dsa-1649388330904&hsa_kw=&hsa_mt=&hsa_net=adwords&hsa_ver=3&cq_src=google_ads&cq_cmp=14745430544&cq_term=&cq_plac=&cq_net=g&cq_plt=gp&gclid=Cj0KCQjwhLKUBhDiARIsAMaTLnFHQGJJ8f1ZO5lEWcTbm4UxOGi341ER0nWR6b80CyqP0AL_I3NwI8caAgH_EALw_wcB
- https://codestoresolutions.com/difference-between-microsoft-ssrs-ssas-and-ssis/

# Appendices

This section contains all the fortnightly progress reports(fpr) from the industry mentor.

**FORTNIGHTLY PROGRESS REPORT (FPR) FROM INDUSTRY MENTOR**

| | | | | | |
|---|---|---|---|---|---|
| Name of student | **Aman** | | Department | **Computer Science & Engineering** | |
| Industry/Organization | **Infosys Limited** | | Date/Duration | **March 2022** | |
| **Criterion** | **Poor** | **Average** | **Good** | **Very Good** | **Excellent** |
| Punctuality/Timely completion of assigned work | | | | | ✓ |
| Learning capacity/Knowledge up gradation | | | | | ✓ |
| Performance/Quality of work | | | | | ✓ |
| Behaviour/Discipline/Team work | | | | | ✓ |
| Sincerity/Hard work | | | | | ✓ |
| Comment on nature of work done/Area/Topic | Doing good work. | | | | |
| **OVERALL GRADE (Any one)** | **POOR/AVERAGE/GOOD/VERY GOOD/EXCELLENT** | | | | |
| **Name of Industry Mentor** | **Rajan Kushwaha** | | | | |
| **Signature of Industry Mentor** | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| **Receiving Date** | | **Name of Faculty Mentor** | | **Sign** | |

**FORTNIGHTLY PROGRESS REPORT (FPR) FROM INDUSTRY MENTOR**

| Name of student | **Aman** | | Department | **Computer Science & Engineering** | |
|---|---|---|---|---|---|
| Industry/Organization | **Infosys Limited** | | Date/Duration | **April 2022** | |
| **Criterion** | **Poor** | **Average** | **Good** | **Very Good** | **Excellent** |
| Punctuality/Timely completion of assigned work | | | | | ✓ |
| Learning capacity/Knowledge up gradation | | | | | ✓ |
| Performance/Quality of work | | | | | ✓ |
| Behaviour/Discipline/Team work | | | | | ✓ |
| Sincerity/Hard work | | | | | ✓ |
| Comment on nature of work done/Area/Topic | Doing good work. | | | | |
| **OVERALL GRADE (Any one)** | **POOR/AVERAGE/GOOD/VERY GOOD/EXCELLENT** | | | | |
| **Name of Industry Mentor** | **Rajan Kushwaha** | | | | |
| **Signature of Industry Mentor** | | | | | |

| **Receiving Date** | | **Name of Faculty Mentor** | | **Sign** | |
|---|---|---|---|---|---|

**FORTNIGHTLY PROGRESS REPORT (FPR) FROM INDUSTRY MENTOR**

| Name of student | **Aman** | | Department | **Computer Science & Engineering** | |
|---|---|---|---|---|---|
| Industry/Organization | **Infosys Limited** | | Date/Duration | **Feb22-Apr22** | |
| **Criterion** | **Poor** | **Average** | **Good** | **Very Good** | **Excellent** |
| Punctuality/Timely completion of assigned work | | | | | ✓ |
| Learning capacity/Knowledge up gradation | | | | | ✓ |
| Performance/Quality of work | | | | | ✓ |
| Behaviour/Discipline/Team work | | | | | ✓ |
| Sincerity/Hard work | | | | | ✓ |
| Comment on nature of work done/Area/Topic | Doing good work. | | | | |
| **OVERALL GRADE (Any one)** | **POOR/AVERAGE/GOOD/VERY GOOD/EXCELLENT** | | | | |
| **Name of Industry Mentor** | **Rajan Kushwaha** | | | | |
| **Signature of Industry Mentor** | | | | | |

| **Receiving Date** | | **Name of Faculty Mentor** | | **Sign** | |
|---|---|---|---|---|---|