Internship Project Report on

Enhancing Workforce Management System using Data Analytics

By

Devangnaba Jadeja
92100374014

Guided by
Dr. Prasanta Chatterjee Biswas
Associate Professor, FMS

A Thesis Submitted to

Marwadi University in Partial Fulfillment of the Requirements for the MBA in the Faculty of Business Management

September 2022

MARWADI UNIVERSITY
Rajkot-Morbi Road,
At & Po. Gauridad, Rajkot-360003, Gujarat, India

STUDENT DECLARATION

I, Devangnaba Jadeja, (92100374014) hereby, declare that the Summer Internship Project Work entitled “Enhancing Workforce Management System using Data Analytics”, submitted in partial fulfillment of the requirements for the award of the degree of **Master of Business Administartion (MBA)** of Marwadi University, Rajkot, India is my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission.

This authentic work has been carried out by me under the supervision of Dr./Prof Prasanta Chatterjee Biswas. I also declare that the content of this project report does not form a basis for the award of any previous degree to anyone else.

I understand that any violation of the above will be cause for disciplinary action by the university and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Date: DD/MM/YYYY

Place: Rajkot

Signature of the Student

Devangnaba

Jadeja

92100374014

UNIVERSITY CERTIFICATE

This is to certify that Summer Internship Project work embodied in this dissertation titled "Enhancing Workforce Management System using Data Analytics" was carried out by **Devangnaba Jadeja** at **Marwadi University** for partial fulfillment of **MBA** in **Faculty of Management Studies** to be awarded by Marwadi University. This research/project work has been carried out under my guidance and supervision and it is up to my satisfaction.

[MU Logo Watermark]

Date:

Place:

Table of Contents

Part A

Background of IT Industry

Growth and evolution:

There's a full-fledged history that led to the development of the field of Information and technology. But we can date back the definition of Information and Technology to the 1980s, which stated, "Over the last decade new technology has begun to take hold in American business, one so new that its significance is still difficult to evaluate ... The new technology does not yet have a single established name. We shall call it information technology." This appeared in a Harward Business Review article that predicted its upcoming effects on the world.

If we simply define what is Information and Technology, then we can define it as solving complex issues regarding computer systems or controlling who can access a particular system. So it is not just about installing hardware and software, it is much more than that.

There is so much more demand for IT professionals today and the demand will be on the more increasing day by day. IT Professionals today help in various different tasks in the organization. Their work and task is related to creating various policies that ensure that the systems run effectively, the strategic goals are met, business efficiency is maintained and improved with automated processes, longevity and continuity of the business are maintained, etc. Thus, we cannot even stand without the role of Information Technology in business. The business might collapse with this much amount of data without Information Technology.

We will discuss the slow evolution of this field in the following paragraphs.

Many complicated activities had to be carried out, but as there were no computers in use at the time, precursors were used to complete them. The Abacus was the earliest known tool for calculation and has been in use ever since. It was made up of rows of movable beads that represented numbers on a rod and dates back to 2400 BC. However, the concept of programming gadgets didn't fully take off until the 1800s. The Jacquard loom was created at this period, allowing weavers to create fabrics with complex woven patterns. To regulate weaving patterns, this device fed punched cards into the loom. People employed the loom's automatic machine instruction system in the 20th century. But this approach was subsequently supplanted by electronic devices.

The Difference Engine was created in the 1820s by Charles Babbage, an English mechanical engineer who is regarded as the creator of the computer. This was considered to be the first mechanical computer.

He then published the blueprints for his Analytical Engine in the 1830s. A punch card system would have been used to run the analytical engine. Ada Lovelace, Babbage's student, developed these ideas. She developed a set of operational instructions for the machine—now known as a computer program—going beyond simple mathematical computations. The Analytical Engine was going to be the first general-purpose computer ever built. However, it was never finished, and the directives were never followed.

The early work of Jacquard, Babbage, and Lovelace served as the foundation for many of the data processing and execution capabilities of modern IT, such as conditional branches (if statements) and loops.

Punch cards were also utilised in the 1890s by American statistician and inventor Herman Hollerith to input data into his census tabulating machine. This served as a significant forerunner of the contemporary electronic computer. By automatically reading and sorting cards that were numerically encoded by perforation location, Hollerith's machine recorded statistics. In order to produce these machines, Hollerith established the Tabulating Machine Company in 1911. In 1924, it adopted the name International Business Machines Corp. (IBM).

Konrad Zuse, a German engineer, created the Z2, one of the first electromechanical relay computers, in 1940. It operated at extremely slow speeds that are unthinkable today. Colossus computers, created by British codebreakers during World War II, arrived later in the 1940s. German encryption machines known as "Tunny" were used to encrypt messages, which these computers intercepted and decrypted. The Bombe was created at around the same period by British mathematician Alan Turing. Using this device, German Enigma machine messages were deciphered.

The modern computer was initially envisioned by Turing, whose work is immortalised by the Turing Test, in his 1936 paper "On Computable Numbers." In this essay, Turing proposed that a machine's memory could be used to hold programmable instructions that could be used to carry out specific tasks. The essential foundation of contemporary computing technology is this idea.

The world's first general-purpose commercial digital computer, the Ferranti Mark 1, was created in 1951 by the British electrical engineering firm Ferranti Ltd. The Manchester Mark 1 device, created at Victoria University of Manchester, served as the inspiration for this device.

Then the IT revolution picked up a really good pace.

The first commercial application for the LEO I computer was launched by J. Lyons & Co. in 1951. One of the first digital computers capable of real-time operation was the 1951-released Whirlwind from MIT. It was also the first computer that allow users to enter commands via a keyboard in 1956.

The factors that finally gave rise to the discipline of IT grew along with computers. The creation of the following gadgets starting in the 1960s paved the way for an IT revolution:
Screens
editors for text
Mouse
Disc drives
Optical fibre
Integrated circuits
Programming languages such as FORTRAN and COBOL.

Mathematicians are no longer the only experts in the IT industry nowadays. Network engineers, programmers, business analysts, project managers, and cybersecurity experts are just a few of the people it employs.

Then came the invention of the Internet.

Governments, defence organisations, and universities dominated the computing IT industry throughout the 1940s, 1950s, and 1960s. With the creation of office programmes like spreadsheets and word processing software, it also filtered into the business sphere. Specialists who could design, develop, modify, and maintain the hardware and software needed to support corporate processes became necessary as a result.

There were many different computer languages developed, and specialists in various languages also emerged. Databases were controlled by Oracle and SAP programmers, and networking software was created and updated by C programmers. These were in great demand, which is still the case today, particularly in the fields of compliance, AI, and cybersecurity.

Email's creation in the 1970s changed communications and IT. Email was first developed as a test to check if two computers could communicate, but it has now developed into a quick and simple tool for people to communicate. Even though the term "email" wasn't invented until much later, many of its early conventions, such the use of @, are still in use today.

The internet and the world wide web are responsible for the development of many IT technologies. But the U.S. government-funded network known as ARPANET, which was first conceived of as an interstellar computer network by MIT researchers in the 1960s, is regarded as the forerunner of the current internet. From just four computers, ARPANET expanded into a network of networks. Transmission Control Protocol (TCP) and Internet Protocol eventually resulted from it (IP). This made it possible for remote computers to virtually converse with one another. Machine-to-machine communication became a reality because to packet switching, which transmits data from one computer to another.

In 1991, Tim Berners-Lee created the World Wide Web, a web of information that anybody could access. The Nokia 9000 Communicator, which debuted in 1996, was the first mobile handset with internet access. The first laptop, the first search engine for domains, and the first search engine in the globe were all already available at this point. Google, a dominant search engine, was founded in the late 1990s.

An open-source web content management system called WordPress was created at the turn of the century. This made it possible for people to go from passive online users to active contributors who posted their own material.

The world of information technology has grown significantly since the World Wide Web was created. Today's IT includes things like tablets, smartphones, speech recognition software, quantum computers, and nanometer computer chips. Since its inception in the 1960s, cloud computing has been integral to many firms' IT plans. The idea of time-sharing—sharing computing resources with numerous users at once—was developed in the 1960s and 1970s. In addition, by 1994, the cloud metaphor was used to depict virtual services and devices that function like actual computer systems.

But cloud computing didn't really take off until 2006, when Amazon Web Services (AWS) was founded. AWS currently controls the greatest portion of the cloud computing business, followed by its three main rivals: Google Cloud Platform, Microsoft Azure, and Alibaba Cloud. In the first quarter of 2021, the top three providers—AWS, Google, and Azure—were responsible for 58% of all cloud spending.

Over the past decade, other technological advancements have really affected positively the field of Information and Technology. This includes developments in:

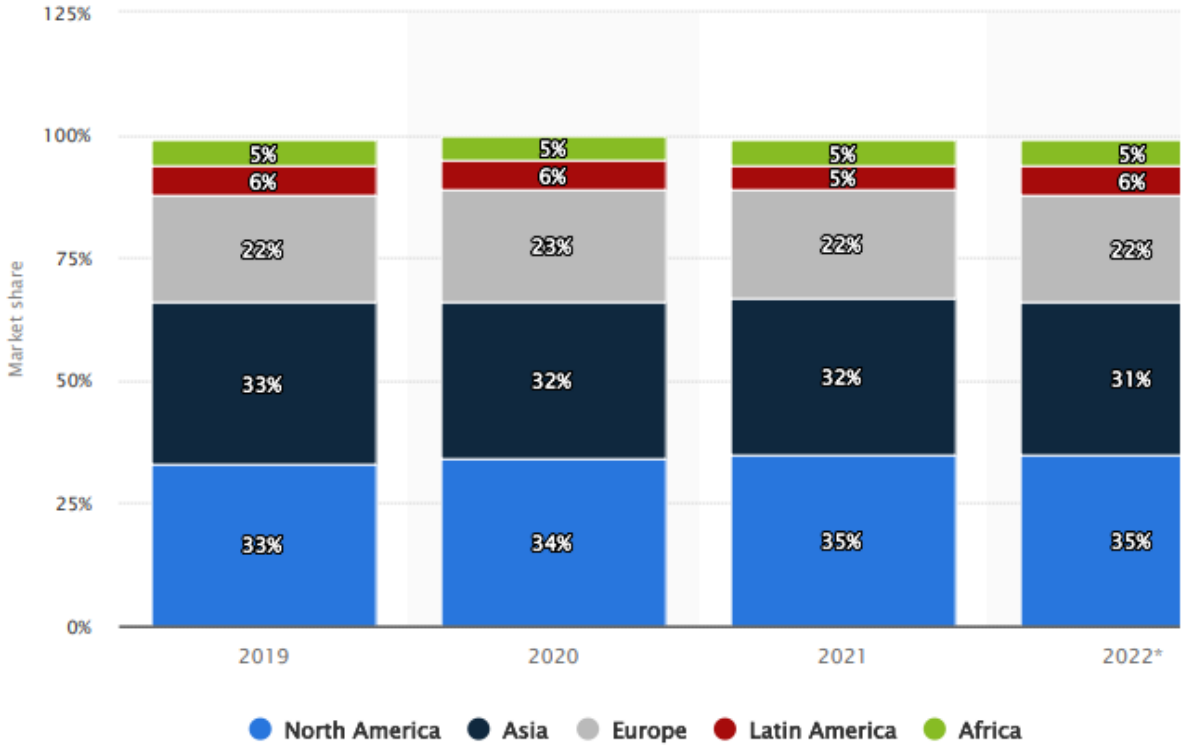
- Social media
- Internet of things
- Artificial intelligence
- Machine learning
- Robotic process automation
- Big data
- Mobile computing -- in both devices and communications technologies such as 4G and 5G

Network and system connectivity is also increasing. According to Cisco research, there will likely be 500 billion internet-connected gadgets by the year 2030.

Key indicators and future trends of the Information Technology industry

World Market:

Distribution of the information technology (IT) industry worldwide from 2019 to 2022, by region:



Source: Statista

Our access to technology is at the heart of most aspects of our professional and personal life. Technology underpins a lot of what we do on a daily basis, from personal web access to commercial platforms, third-party suppliers, and apps, and there is only future growth. Here are some essential technological and statistical data that you should be aware of.

Information technology is one of the fastest growing industries in the world because technology is unquestionably still progressing at a steady clip. By 2030, some segments of the workforce should increase at a rate that is 4-5 times that of now, which is predicted to double the rate of other industries. According to projections, the value of the technology sector will surpass $5.3 trillion in 2022. Through 2024, the U.S. IT sector is anticipated to grow at a CAGR of 5%. By 2025, cloud computing is anticipated to increase at a CAGR of 17.5%.

Technology now rules the roost in the commercial world, far beyond the compartmentalized department of old. Here are some statistics about technology in general. According to Zippia, the technology sector accounts for 35% of the global market. 90% of the data in existence today was produced between 2019 and the present. In the US, there are more than 585,000 tech businesses. The combined value of the Big Four—Amazon, Apple, Google, and Facebook—is $4 trillion. According to 84% of firms, a top priority for 2022 is cybersecurity and resilience.

The role of technology in our corporate environment is only going to increase. The US economy is driven more by technology than by any other industry besides healthcare. Jobs in the ICT sector are constantly increasing and will reach a new high in Q4 2021.
The top five technological categories are:
Software Hardware
Telecom
Business and IT Services
Latest Tech

By 2030, it is anticipated that artificial intelligence will have made $15.7 trillion in economic contributions. At a CAGR of 38.6%, the market for machine learning is expected to reach $152.24 billion in 2028. According to 99Firms, 8.4 billion voice assistants will be powered by machine learning by 2024. According to Genpact, companies that have adopted AI will be 10 times more efficient and hold double the market share by 2025 compared to those whthatave not. In terms of innovation, 60% of business owners think AI is now the most promising technology. 91.6% of Fortune 1000 organizations want to increase their expenditures on big data and AI projects.

New innovations are always being developed in order to support effective operations and ongoing developments in technology, which is constantly changing. According to Statista, between 2018 and 2023, the growth rate for new technology will be 104%. In 2022, it is anticipated that investments in the Internet of Things would increase by around 14%. By 2025, quantum computing is anticipated to generate $780 million. By 2025, Statista projects that there will be 30.9 billion IoT devices.
With budgets always expanding to devote more money to tech initiatives, technology is a pervasive and powerful force for organizations. Finding resources for innovations can be challenging due to persistent skills deficits, but there are still many prospects for technological development and expansion.

Country Market:

India's IT sector is thriving, and the country has secured a well-deserved place in the global market thanks to strong governmental support, an excellent English-based educational system, a vast talent pool, and the free market. The IT sector, which will contribute 7.4% of GDP in the fiscal year 2022, has significantly aided India's socioeconomic development and may eventually serve as the engine of modern India.

Over 19% of worldwide IT spending was made in India. However, the nation's achievements in this area were limited at first. A new software policy was developed by the government since it lacked the necessary infrastructure for software development, allowing for greater involvement and simpler trade processes. As a result, the sector expanded rapidly, particularly after the 1990s' economic liberalization.

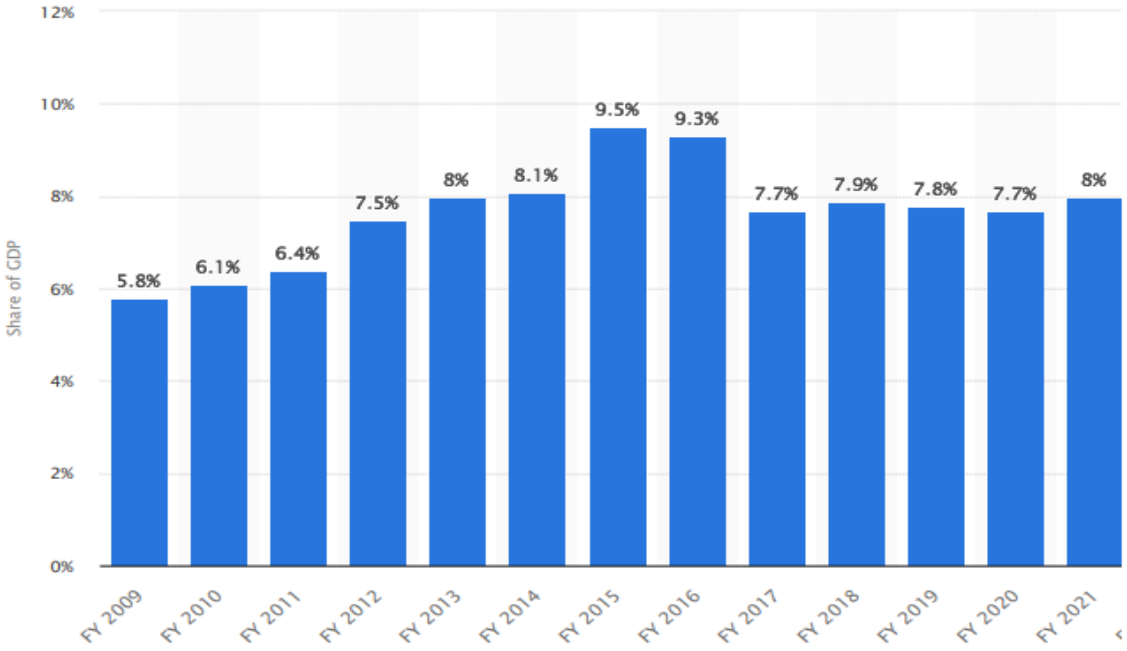
India is well-known throughout the international business sector for its top-notch IT outsourcing services. India is aptly referred to as the IT outsourcing hub because it serves the requirements of the big companies in North America and Europe. India's export revenue was approximately 150 billion dollars at the end of the fiscal year 2021, five times the size of its domestic market. The IT services and IT-enabled services were the most in-demand on the global market among the many sub-sectors. Moreover, half of the export revenue made by the information technology sector was from IT services. As a result, the sector has helped the nation add millions of new employments.

Even if India's IT industry guarantees development in both employment and money, there are still plenty of obstacles to overcome. The effects of the coronavirus epidemic, fierce global rivalry, and evolving technology have put it to the test while opening up new prospects for success in the future.

The market size of the domestic information technology industry across India from the financial year 2014 to 2020, with an estimate for 2021: (in US billion dollars)

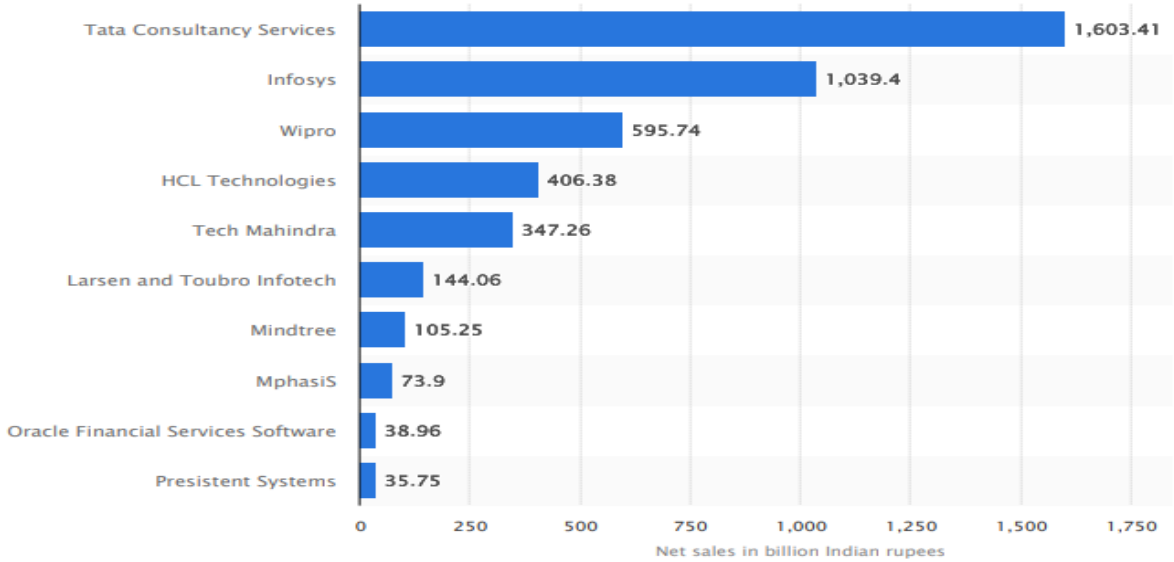
Source: Statista

Share of Information technology/business process management sector in the GDP of India from the financial year 2009 to 2022:



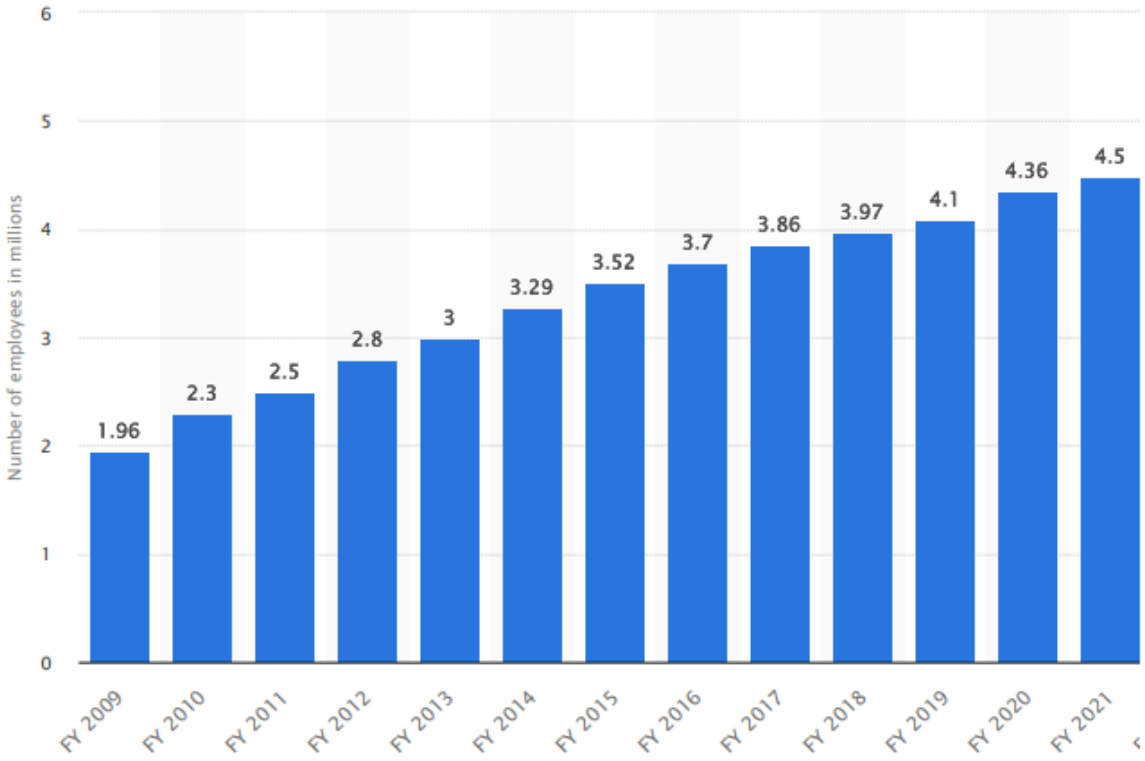
Source: Statista

Leading IT services and consulting companies in India as of June 2022, based on net sales (in billion Indian rupees)



Source: Statista

Employment of the IT-BPM industry in India from the financial year 2009 to 2021, with an estimate for 2022 (in millions)



Source: Statista

State Market:

Gujarat has traditionally generated the greatest conversation because it is home to the most successful businesspeople, including Mafatlal, the Ambanis (Reliance), and others. India's IT industry is expanding quickly, thus it is seeking new areas to expand into. This state's potential to become the next rapidly developing hub of India is supported by a number of factors. Many factors that support Gujarat include an increase in per capita income, business investment, reliable power supply to villages and industry, computerization of the electricity board, Prime Minister Narendra Modi's Digital India initiative, a rise in literacy rates, and good rail and air connectivity. The key benefits include the state's low crime rate, safety, and the prohibition of alcohol use, which makes it an ideal place for people to live and work**.**

According to a representative of a regional IT association from Ahmedabad, IT companies in Gujarat are planning on growing by 20% this fiscal year and will concentrate on product development and worldwide access. Although the state's official output data for IT and IT-enabled services (ITeS) has not yet been released, ST PI's local office exports have been pegged at Rs 1,500 crore.

Gujarat's capital city, Gandhinagar, constructed the Infocity-IT park.
According to reports, this city is built around the idea of "A city within a city," giving its residents access to a clubhouse, a 24/7 work culture, educational resources, and recreational amenities. Gujarat's Infocity is a center for international IT outsourcing. TCS, paramount web solutions, Cybage Software, i-serve Systems, Avaya Global Connect, Infosys, and many other IT businesses have their headquarters in Gandhinagar. In numerous more Gujarati cities, including Ahmedabad, Surat, and Rajkot, there are numerous additional software technology parks and IT businesses operating.

According to NASSCOM, Gujarat is expected to lead IT growth in the ensuing ten years. A better IT infrastructure and education are needed According to Som Mittal, president of the IT industry group, 90% of the overall IT business is currently concentrated in six cities—Mumbai, Delhi, Chennai, Hyderabad, Bangalore, and Pune—newer towns like Ahmedabad, Jaipur, Mysore, and Chandigarh will dominate it within the next ten years.

Together, the State government and educational institutions are genuinely working to develop talent. The Gujarat Electronics and Software Association of India, or GESIA, is working with academic institutions to generate trained human resources while also improving market access to support the growth of the IT sector. Gujarat has a tremendous amount of untapped potential that could eventually establish it as a top IT center in India.

Background of the company

Company Details:

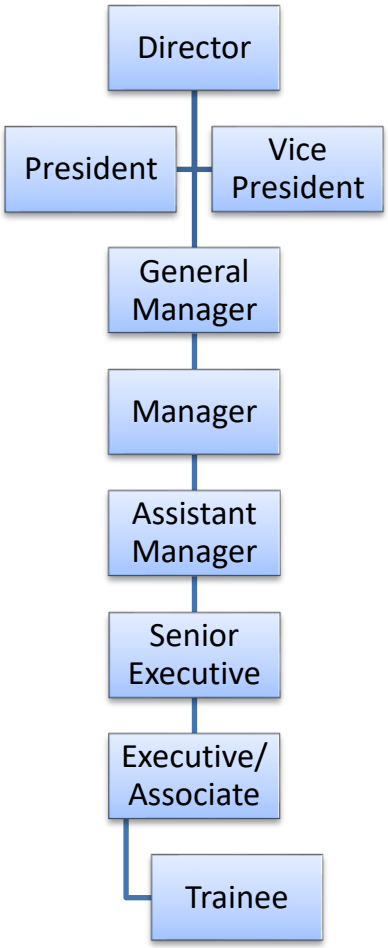
Since 2000, they have remained dedicated to its vision— by providing technology-enabled business solutions to keep its clients ahead of the competition.
Their work features a bouquet of Fortune 500 corporations, mid-size businesses, fast-paced start-ups, and government bodies.
Their solutions take you on a journey of digital transformation through unique products, services, and superior product engineering. From traditional technologies to emerging technologies powering innovation, their teams have the expertise to help transform your organization's digital journey by identifying the right technology platforms that deliver the right solutions for the clients.
They do this through their array of products and solutions, from an automation suite to their products aimed specifically at digitizing the clients' tax journey, complete with a blockchain-based digital signature solution. Its aim is to provide end-to-end solutions for its client's most pressing business needs.

The company has 6 directors and 2 reported key management personnel.
The longest-serving director currently on board is Niraj Karnik Hutheesing who was appointed on 22 August 2000. Niraj Karnik Hutheesing has been on the board for more than 22 years. The most recently appointed director is Keval Niraj Hutheesing, who was appointed on 20 November 2020.

The functional goals of the company are as follows:

- Financial objectives – Budgeting process along with documented assumptions, reduce payroll processing error by 20%
- Customer objectives – Improvising existing processes and experiences, to see appreciation in the eyes of internal customers.
- Technology objectives – Leveraging Cygnature and Hrinnova, Exploring document management systems, Process and tool enhancement towards Automation.
- People objectives – Create a culture of trust within the team, Periodic knowledge and feedback sharing sessions, Identify key areas of friction and solve the problems.

Organizational Structure:



Cygnet Infotech follows Functional Organisational Structure. The departments of Cygnet Infotech are as follow:

- Accounting and Finance
- Learning and Development
- Marketing
- Purchase
- Sales
- MIS
- Research and Development
- Analytics
- Legal
- Human Resources
- IT Support
- Operations
- Product Management

Company Products/Services:

Products:

- Tax technology
- Fintech
- digital publishing platform
- Automation – Test and Intelligent Automation
- Digital signature
- Ordering apps – Web and mobile commerce, Field sales management
- Online Publishing – Print and ebook solution

Services:

- Product engineering
- Cloud and DevOps
- Application development
- Testing
- UI/UX development
- Artificial Intelligence
- Business Intelligence
- SAP services
- Implementation services – ERP, CRM, CMS
- Infrastructure management
- Staff augmentation

What Architecture they deliver:

- Micro Services – Spring boot, .NET, Node.js, Go
- Serverless – AWS lambda, Google Cloud Functions, Azure functions

Principles that the Technology team follows:

- Reusability – Modular approach, be it code, designs, test cases, HTML components
- Testing
- Scalability
- Telemetry
- API first development
- Open source first – Tools and Libraries

Principles that the Accounting team follows:

- Keeping books for accessing the real-time situation of the company
- Following the accounting standards
- Internally auditing the books to check the compliances with various laws like TDS, GST, etc
- Timely reimbursement of expenses done on behalf of the company
- Preparing and generating invoices as per the law to prevent any revenue leakages
- Compiling data from various heterogeneous sources and presenting it in a reportable format

Principles that Finance and Taxation team follows:

- Preparing the budget for various departments as per need of the hour
- Managing the finances available for the business operations

- Managing the day-to-day cash flow of the company for smooth payment processes
- Dealing with the intricate laws of different countries where we operate to safeguard from any litigation and penalty
- Handling the dynamic laws of Direct Tax and Indirect Tax to reduce the risk exposure and be a responsible compliant corporate citizen
- Complying with the FEMA laws for cross-border transactions such as exports, imports, acquisitions, etc
- Processing the payrolls on time

Principles that the Purchase and Travel desk follows:

- Understanding the technical requirement for the procurement
- Raising the requisition to the vendors in the form of RFP
- Negotiating to get the best quality product at the best price
- Managing vendors for effective cost planning
- Helping to successfully plan the outstation/foreign business meetings and making our business travel a memorable experience
- Arranging for the resources required for seminars, exhibitions, etc

Principles that the Legal and Compliance team follows:

- For management to avoid direct dealings on legal matters
- Governance and compliance
- Ethics and gatekeeper of information
- Risk mitigation for providing timely counsel, liaison with external legal counsel, contract review management to prevent litigation exposure
- Secretarial compliance

Financial Information:

Here is a summary of financial information of CYGNET INFOTECH PRIVATE LIMITED for the financial year ending on 31st March 2021.

- Revenue/turnover of CYGNET INFOTECH PRIVATE LIMITED is INR 100 cr - 500 cr
- Net worth of the company has increased by 21.82 %
- EBITDA of the company has increased by 111.67 %
- Total assets of the company have increased by 14.47 %
- Liabilities of the company have increased by 22.72 %

Operating Revenue	INR 100 Cr - 500 Cr
EBITDA	111.67 %
Net worth	21.82 %
Debt/Equity Ratio	0.15
Return on Equity	17.91 %
Total Assets	14.47 %
Fixed Assets	-3.15 %
Current Assets	44.70 %
Current Liabilities	22.72 %
Trade Receivables	12.17 %
Trade Payables	-13.39 %
Current Ratio	5.05

Part B

Enhancing Workforce Management System using Data Analytics

Introduction:

Workforce management entails identifying an organization's key priorities and foreseeing human capital issues so that appropriate actions may be made to reduce liabilities and uphold productivity. Every employer has distinct priorities, but regardless of size or industry, the majority of businesses aspire to be more productive and efficient than their rivals. Productivity improvement without increasing labor costs is one of workforce management's main goals. It typically takes the appropriate people and resources to be in the right location at the right time to achieve this aim.

Among other things, employers utilize workforce management to complete a lot of the following tasks:

- Predicting labor demand
- Scheduling employees
- Recording attendance and time
- Dealing with absences
- Analyzing data
- Assisting with compliance
- Decreasing Attrition rates

Data scientists use exploratory data analysis (EDA), which frequently makes use of data visualization techniques, to examine and analyze data sets and summarise their key properties. It makes it simpler for data scientists to find patterns, identify anomalies, test hypotheses, or verify assumptions by determining how to modify data sources to achieve the answers they need. EDA helps with a better understanding of the variables in the data collection and their relationships and is usually used to investigate what data might disclose beyond the formal modeling hypothesis testing assignment. It can also assist in determining the suitability of the statistical methods you are considering using for data analysis. EDA approaches, were initially created by American mathematician John Tukey in the 1970s. These methods have been in use since then and contribute a lot in the field of Analytics.

This research cum analysis hovers around the fictitious data provided by Cygnet Infotech which has been fetched from their Hrinnova portal.

Research Problem:

The greatest competitive advantage sought after but not completely achieved by firms coping with the turbulence and uncertainty driving major changes in today's workplace is employee and worker insights. This is accompanied by the rising demand for the human resource (HR) function to comprehend how workforce analytics informs the company and supports success. Data collection and management for the workforce consume a lot of time and resources for organizations. But simply possessing this information is insufficient. The HR manager should be able to examine the data to recognize patterns and unearth important insights. It's simpler to say than to do. Timesheets are very much important in any organization to calculate the number of hours an employee spends on projects for payroll processing, absenteeism management, and much more. Project managers rely on timesheet data to fulfill clients' project completion requirements. There's an ample amount of data that is stored in the timesheet software and to fetch that data and present it to the higher authorities is a task that has to be achieved. These insights can be presented in a very good way with the help of Power BI dashboards. Any organization's human resources are its most valuable assets. In order to find and develop their personnel, businesses invest a significant amount of time and money. Employee turnover costs businesses a lot of money, especially when it comes to important personnel. Attrition can occur for a variety of reasons, including low pay, few or no career advancement opportunities, poor employee supervision, a desire to work for multinational corporations, a lack of recognition, a lack of freedom of speech within the workplace, and underutilization of employees' talents and skills. As a result, the attrition rate is increasing as more and more people leave the company. So, if with the enhancement of workforce management and using exploratory data analysis on Cygnet Infotech's fictitious data, the rates can be decreased then the techniques of data analytics are of much use. Moreover, dashboards are also very much helpful in presenting insights in a very creative way. Along with the exploratory data analysis of our employee attrition data, timesheet data of Cygnet Infotech has been fetched from the portal to create summary reports in Power BI. The problem can be easily tackled with the help of visualizations as visualizations can help us see more clearly which variable is affecting attrition and which is not.

Literature Review:

Workforce Management:

An organization uses a comprehensive set of procedures called "workforce management" to maximize staff productivity. Workforce Management entails accurately estimating labor needs and developing and overseeing staff schedules to carry out specific work on a day-to-day and hour-to-hour basis. It takes into account business metrics as any employee's daily performance can be known, and analyses can be done easily with the help of historical data. It is typically a technical system that uses employee data and plans and strategize according to the human resources planning. Effective workforce management with modernized tools and techniques can lead to the success of an organization.

It's critical to remember that Workforce Management is not a novel idea. Organizations have been automating workforce management activities for many years in order to produce better results; time and attendance monitoring was one of the first processes to be automated and then came all the machine learning and artificial intelligence techniques that can be used with the automation additionally. The idea has changed dramatically over time, though, and is now utilized by technological businesses and HR divisions to track and boost labor productivity and efficiency and decrease one of the important goals of any organization which is decreasing attrition rates.

Workforce Analytics:

Managers can learn more about various abilities through workforce analytics, and they can then plan how to use these talents for new company initiatives and leadership development. Analytics is an emerging field presently and with the advancement of machine learning and artificial intelligence techniques, things have become a little easier to make effective and fast decisions regarding any topic. As well as conventional ratios like time to fill, cost per hire, accession rate, retention rate, add rate, replacement rate, time to start, and offer acceptance rate, it examines hiring, staffing, training and development, personnel, remuneration, and benefits. A new corporate philosophy is workforce analytics. It may promote a vibrant, new psychological outlook and an excellence-focused workplace culture. Successful HR analytics also require human intervention and intuition to comprehend and extract the knowledge from the data to better influence corporate choices.

Operational and analytical dashboards for data exploration and visualization are provided by workforce analytics. Complex data can be displayed on dashboards in simple-to-understand ways including bar charts, infographics, maps, and scatter plots. These dashboards are available to HR managers as needed. Dashboards provide information on trends in hiring and firing over time, promotions across organizations, and employee distribution by geography, demographic diversity across locations, and compensation and benefits. Additionally, dashboards make it simpler to see connections between various corporate processes and employee performance. They can assist the HR managers in preventing minor issues from developing into significant ones.

Companies will have 30 or 40 criteria to analyze worker data instead of just 3 or 4, thanks to AI, ML, and data analytics. This will enable them to decide what they must do as a company based on data-driven, fact-based decisions. They will also be able to determine which solutions are effective and whether workers are reacting favorably.

Attrition:

A company's workforce shrinks through attrition when employees depart and are not replaced. Attrition, also referred to as a hiring freeze, is seen to be a less disruptive method of reducing the staff and payroll than layoffs. A client base's decline can also be referred to as attrition, frequently as a result of consumers leaving and fewer people choosing to become new customers. Layoffs, which happen when a corporation lets workers go without replacing them, are distinct from attrition caused by voluntarily leaving employees. When employees leave their positions, either freely or involuntarily, in a short period of time, they are replaced with fresh talent. The attrition rate is the

frequency with which employees depart an organisation over a specific time frame. A company can determine if departures are rising or falling by tracking attrition rates over time. When the attrition rate changes, management may be made aware of internal issues that could be contributing to employee departures.

Timesheet management and dashboarding:
An employer can track the hours a certain employee worked over a specific time period using a timesheet, which is a data table. Timesheets are used by businesses to keep track of the time spent on tasks, projects, or clients. Timesheets have been recorded using a variety of ways, including paper, spreadsheet software, and internet time-tracking tools. Digital timesheet formats have now replaced paper-based ones. Timesheets reduce wastage of resources, bills are automated and recorded efficiently, and can help project managers achieve goals in time. You can see exactly what your staff are working on thanks to timesheets. By contrast the number of time employees spend working with the results they create, they may monitor employee productivity. Timesheets make it easier to identify workers who are ineffective and cost your business money. Timesheets help your business save a lot of money by showing you who is and is not productive.

Enterprise dashboards and business intelligence dashboards are both analytical tools for visualizing Big Data in a variety of businesses. These dashboards are essential to business performance management since they offer vital reporting and measurement data. Similar to the dashboard in your car, dashboards show important data and performance indicators in real-time to help you make decisions and better navigate your environment. Power BI is dashboarding tool that can summarize and present timesheet data effectively.

Exploratory Data Analysis and Machine Learning Classification method:
Exploratory data analysis is a strategy for learning from the data rather than using it for formal modeling or hypothesis testing. EDA aids in the analysis of the data sets to highlight their statistical properties, concentrating on four essential factors: the distribution's shape, the presence of outliers, and measures of central tendency (which include the mean, mode, and median). Data exploration, data cleaning, model building, and presenting the results are the steps in the process of machine learning. EDA is an important part of machine learning and analysis. It can help with the following:

- Using EDA, errors and anomalies can be found
- We can learn new things about different kinds of data
- Data outliers can be identified
- Utilizing EDA, we can test our assumptions
- It can be used to identify significant elements
- We can comprehend the connections between distinct data
- Utilizing visualization techniques, data can speak for itself.

Python is an interpreted, dynamically semantic object-oriented programming language. It is particularly desirable for rapid application creation as well as for usage as a scripting or glue language to tie existing components together due to its high-level, built-in data structures, dynamic typing, and dynamic binding. In order to select how to treat missing values for machine learning, it is critical to be able to discover missing values in a data set using Python and EDA combined.

A non-parametric supervised learning technique for classification and regression is called a decision tree (DT). The objective is to learn straightforward decision rules derived from the data features in order to build a model that predicts the value of a target variable. A piecewise constant approximation of a tree can be thought of. Decision trees are very easy and simple to understand and formulate. It is capable to handle both numerical as well as categorical variables. Visualizations can also be done in a very simple way with little data exploration and codes.

Research Objectives:

The objective of this report is to tackle the problems that are arising in workforce management in the organization. We are dealing with two kinds of data here, one is related to the attrition rates of employees in the organization. The attrition data is fictitious data that was provided by Cygnet Infotech and the Timesheet data is a Staging application data on which queries have been run and then reporting has been done in Power BI.

The process of hiring in any company is a relatively easy task but when it comes to retaining those hired employees, it becomes a little tough. But when the company has past or historical data like who has left the company, when, how much is their satisfaction level, the level of their work-life balance, their education levels, etc, then proper analyses could be done on this data and ways can be found out to retain the existing employees and find the reason behind the employees who have left the organization. Successful businesses frequently have a plan in place for maximizing employee retention, which not only helps them save money and resources by reducing turnover rates but also aids in the development of a network of skilled and self-driven workers. As a result, these workers are driven to take on complex and extensive tasks and to adhere to the corporate culture. Some of the very important objectives to achieve with this analysis are mentioned hereby:

- To know the attrition rates department wise
- To reduce it in order to better workforce management system
- To assist project managers with the summaries of Timesheet data
- To fulfill the client's objectives and targets with a better understanding of Timesheet data
- To see what are the major factors that are influencing the attrition levels in the organization with the help of the classification model that we have built using Python as our coding language
- To build efficient strategies for a better workforce management system for the organization's achievement of goals

Part C

Findings:

We have carried out exploratory data analysis in Jupyter Notebook with Python as our coding language to deal with the Attrition data. Next, we have also summarized Timesheet data into a BI report. The Findings are as follows:

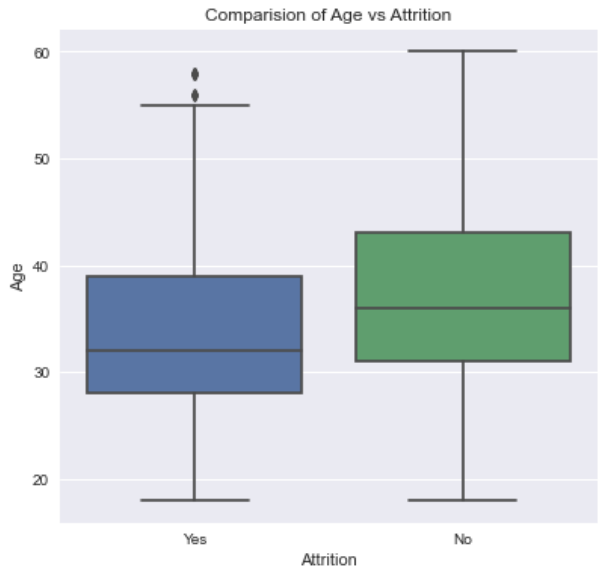
There are 35 columns and 1470 rows. Where the number of employees is 1469 bifurcated into various departments. The name of the columns or variables that we are dealing with are as follows:

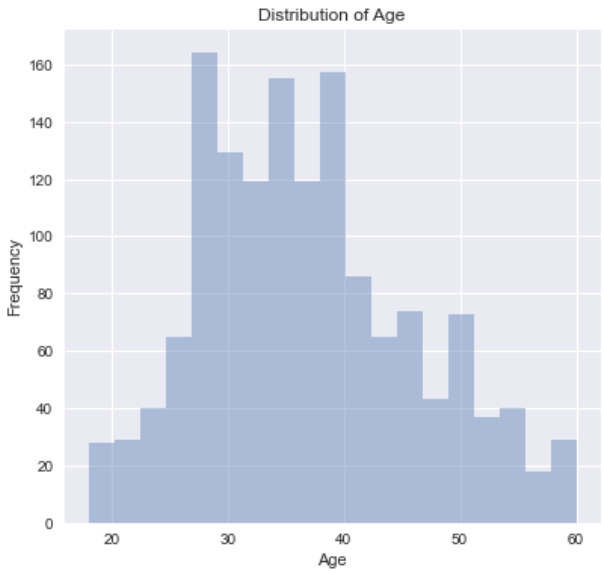
1. Emp Code
2. Age
3. Attrition
4. Daily rate
5. Department
6. Distance from home
7. Education
8. Education field
9. Emp count
10. Emp No.
11. Emp satisfaction
12. Gender
13. Hourly rate
14. Job Involvement
15. Job level
16. Job rate
17. Job satisfaction
18. Marital status
19. Monthly income
20. Monthly rate
21. No. of companies worked
22. Over 18
23. Overtime
24. % salary hike
25. Performance rating
26. Relationship satisfaction
27. Standard hours
28. Total working years
29. Training time last year
30. Work-life balance
31. Years at the company
32. Years since the last promotion
33. Years with current manager
34. Years in current role
35. Stock option level

Analyzing numerical as well as categorical variables:
First of all, we are checking whether Attrition rates are related with numerical variables or not.

Numerical variables:

1. Age



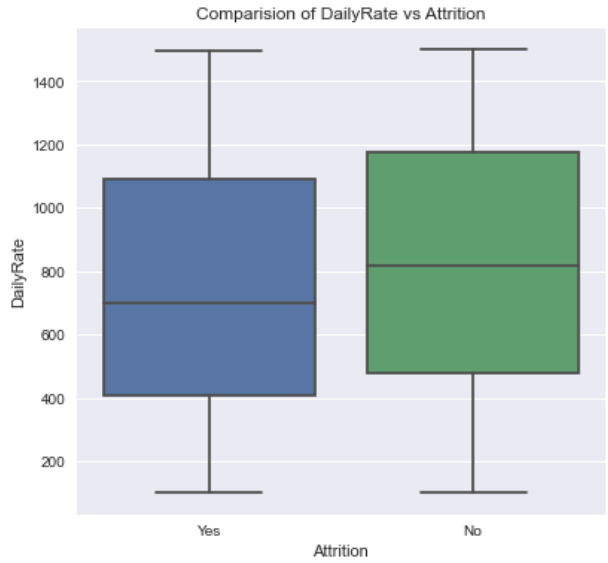


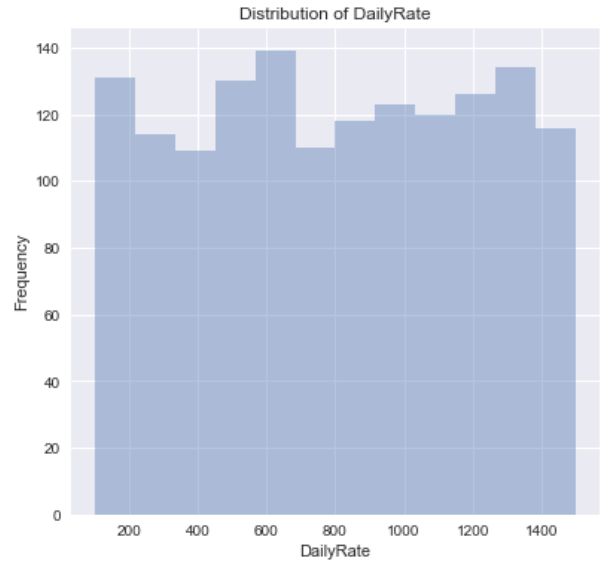

```
In [14]: NumericalVariables_targetPlots(employee_data,segment_by="Age")
```

We can explore the following things from this visualization:

- Minimum age of an employee is 18 years and the maximum is 60 years.
- Employees leaving the organization are in their late 20s and mid-30s.
- This can be due to many reasons like getting better opportunities in other places, switching their career, higher education etc.
- Thus we can say that Age significantly defines Attrition rate.

2. Daily Rate

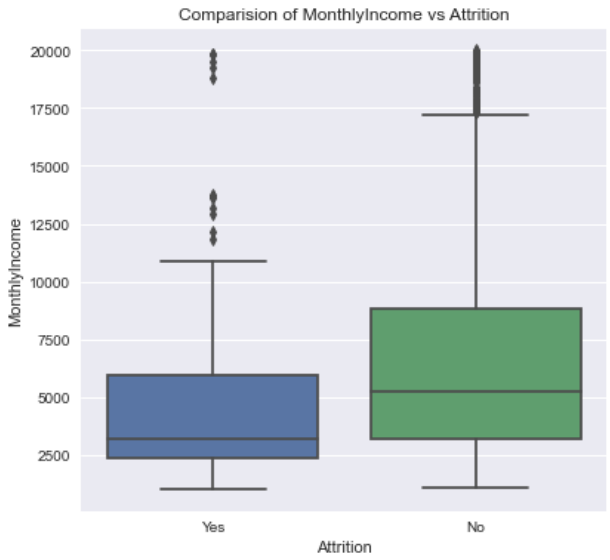


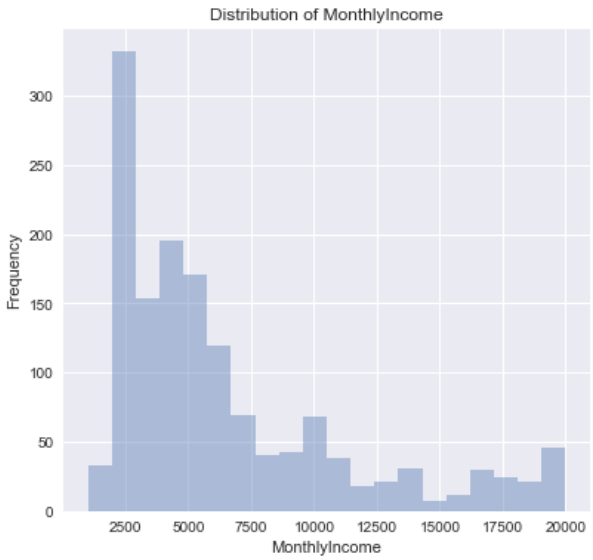



```
In [15]: NumericalVariables_targetPlots(employee_data,"DailyRate")
```

- Employees who are getting lower daily rates are more likely to leave the company as they might not feel secure with their salaries.

3. Monthly Income

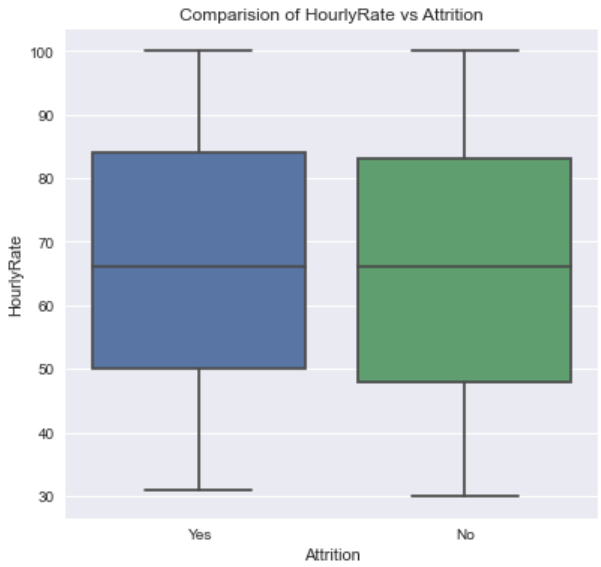


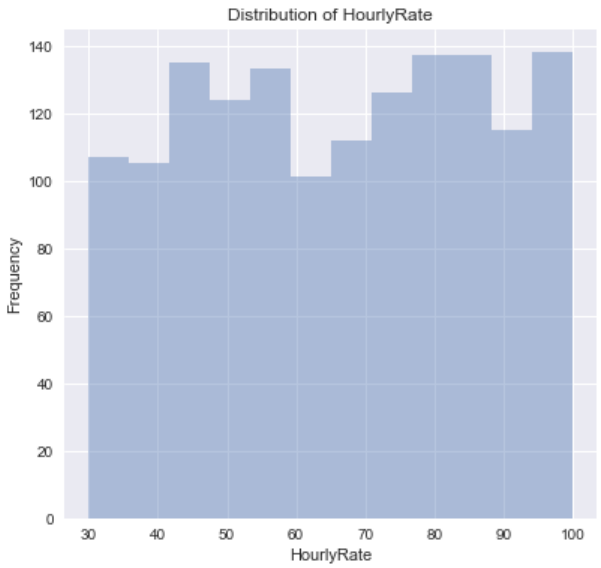



```
In [16]: NumericalVariables_targetPlots(employee_data,"MonthlyIncome")
```

- Here too, employees getting lower monthly income are leaving the company and the frequency too is increasing.

4. Hourly Rate

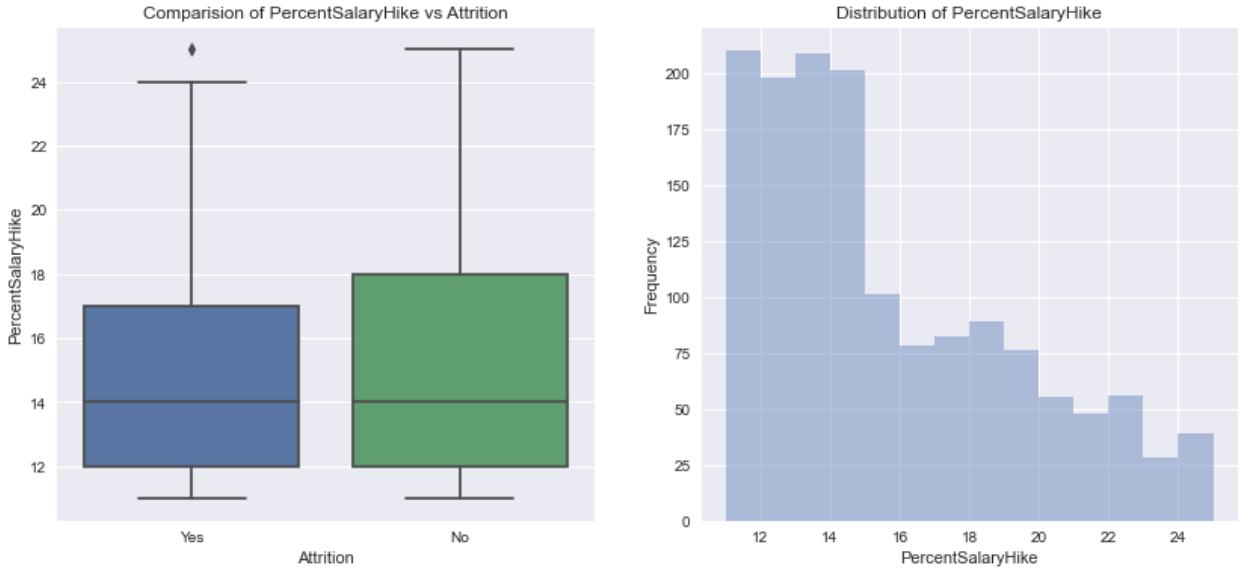




```
In [17]: NumericalVariables_targetPlots(employee_data,"HourlyRate")
```

- But here we can see that there is no significant difference between the employees leaving due to lower hourly rates.
- The frequency is almost the same.

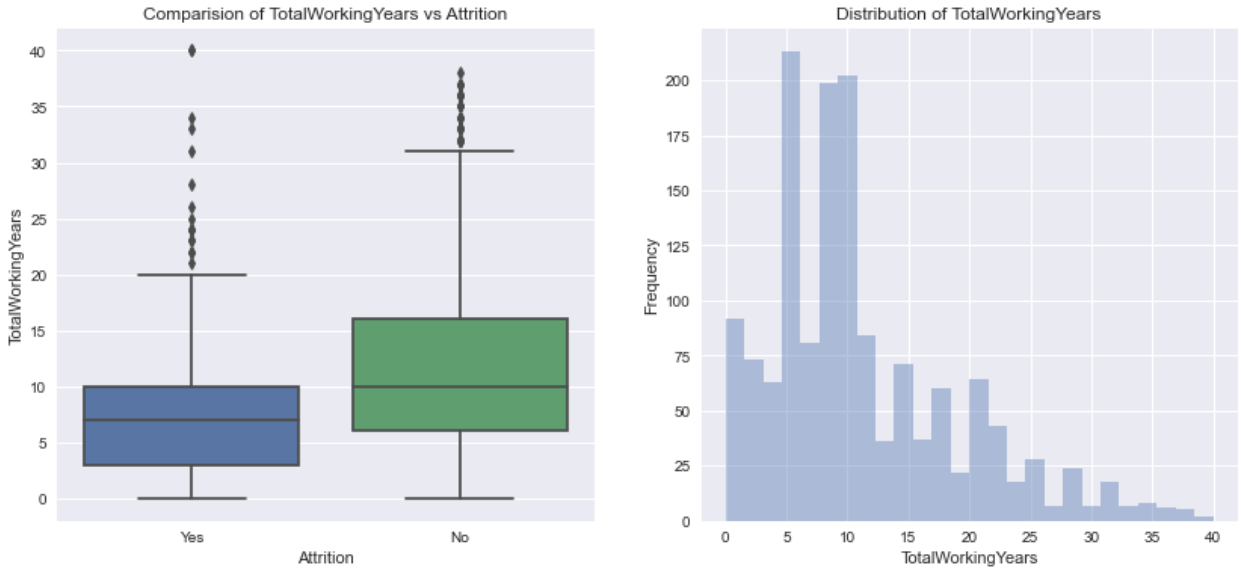
5. % Salary hike




```
In [18]: NumericalVariables_targetPlots(employee_data,"PercentSalaryHike")
```

- Here, employees getting lower salary hike that is less than 16 %, their frequency is increasing.
- Thus, salary hike is also a significant factor in determining Attrition rates.

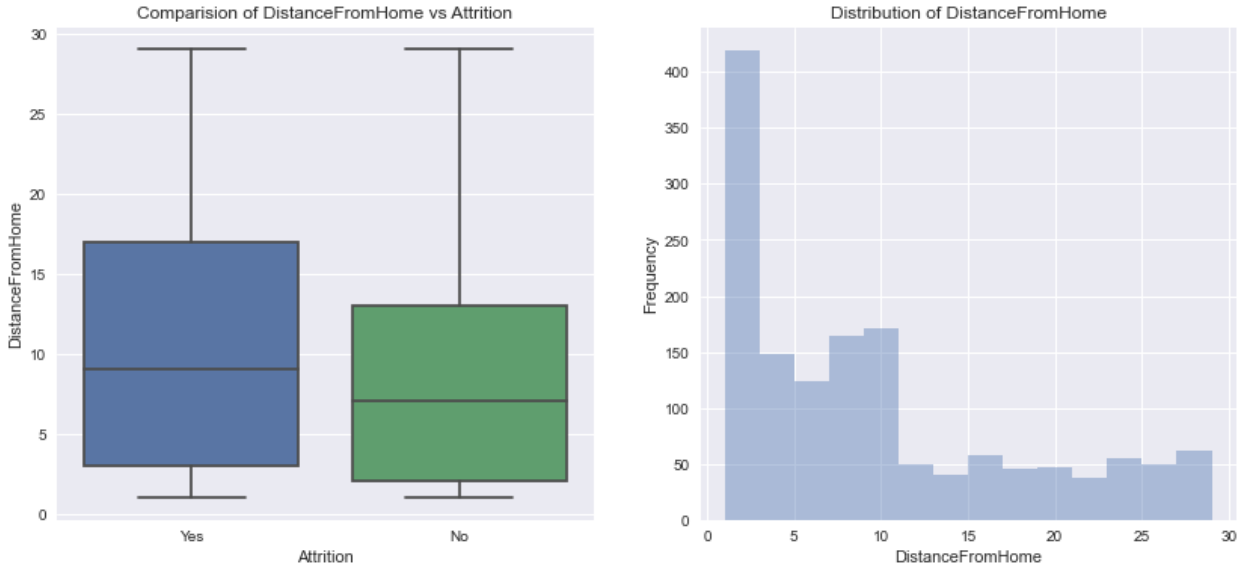
6. Total Working Hours



```
In [19]: NumericalVariables_targetPlots(employee_data,"TotalWorkingYears")
```

- Here we can see, a totally different scenario.
- With the increase in total working hours of an employee, the frequency of employees leaving the company decreases.
- Hence, total working hours are not determining the attrition number.
- Employees with total working hours less than 10 are also prone to leave the company than the employees working for more than 10.

7. Distance from Home

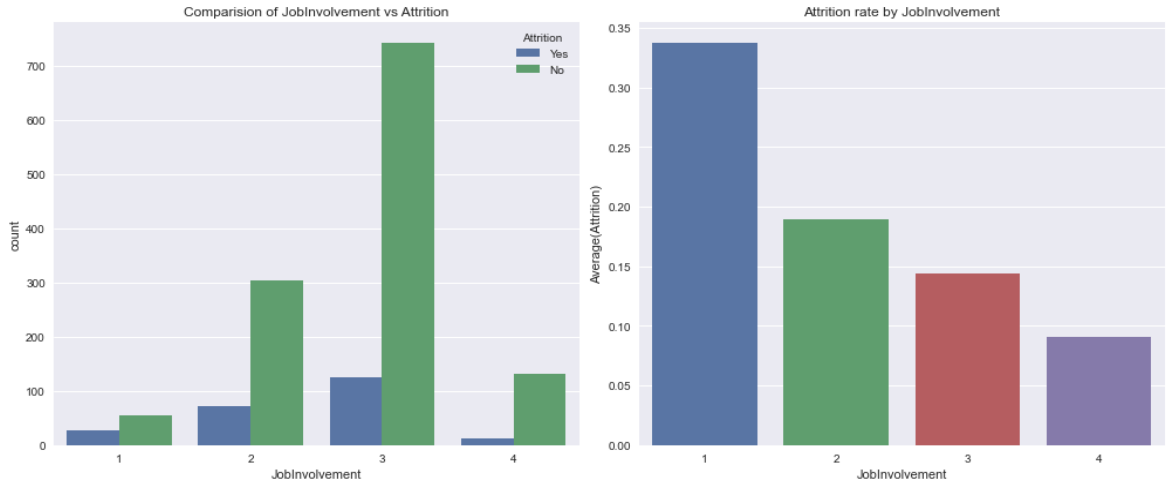



```
In [22]: NumericalVariables_targetPlots(employee_data,"DistanceFromHome")
```

- With the increase in distance from home, the rates of employees leaving the company also increasing as people might find travelling costs higher and might leave the organization because of this.

Categorical Variables:

1. Job Involvement

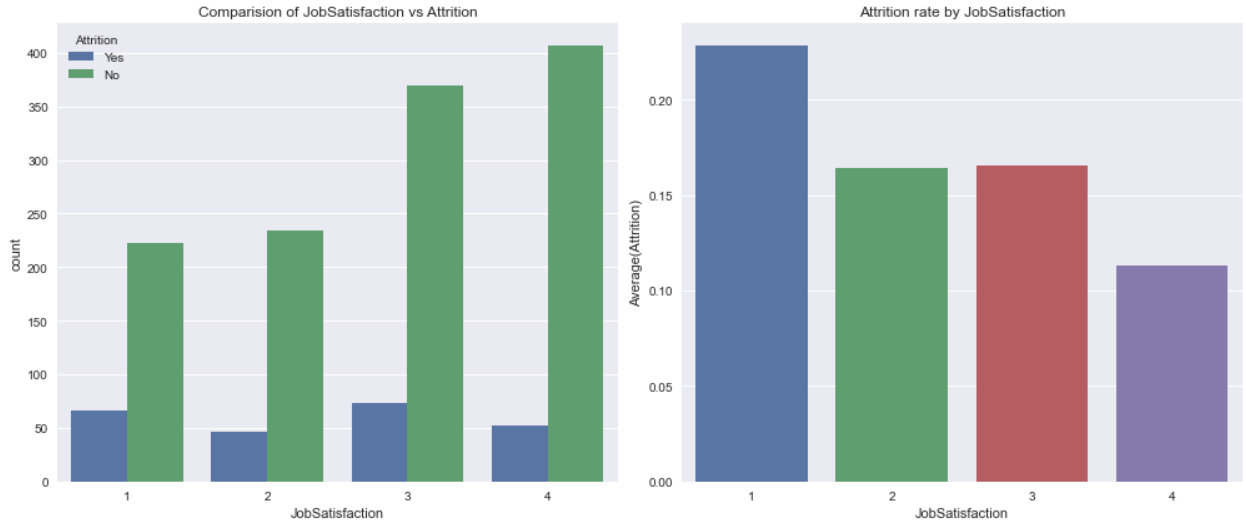



```
In [25]: CategoricalVariables_targetPlots(employee_data,"JobInvolvement")
```

Here's what we get to know while we explore the categorical variables:

- 1,2,3,4 states Highest, High, Medium, Low levels of Job involvement respectively.
- We can see from here that employees with high Job involvement level have left the organization than with medium Job involvement level.
- Average attrition rate is 35% and the frequency is also more of employees having high Job involvement level.

2. Job Satisfaction

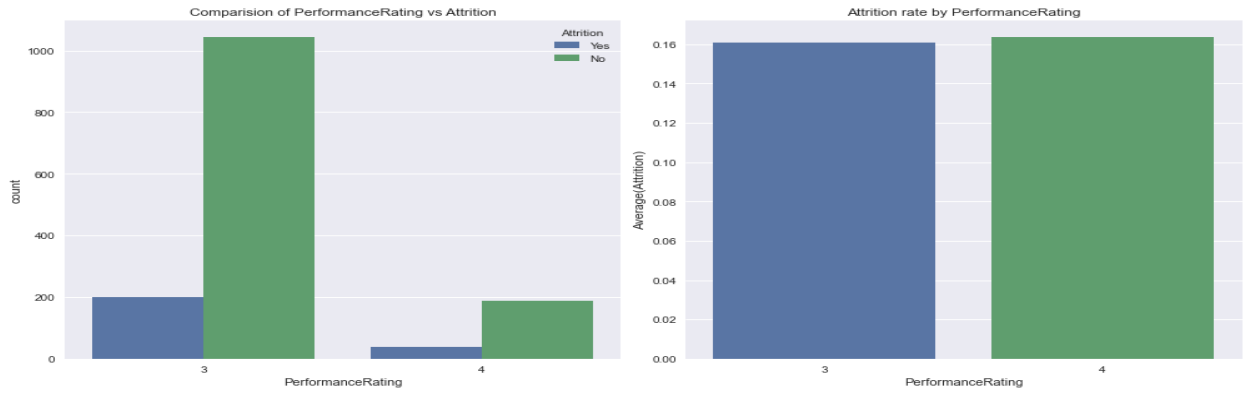



```
In [26]: CategoricalVariables_targetPlots(employee_data,"JobSatisfaction")
```

- Employees having high Job satisfaction level are too leaving the organization with the employees having low job satisfaction levels.

- Hence there is no significant difference between the two.
- Average rate of attrition is somewhere around 25% when the Job satisfaction is low.

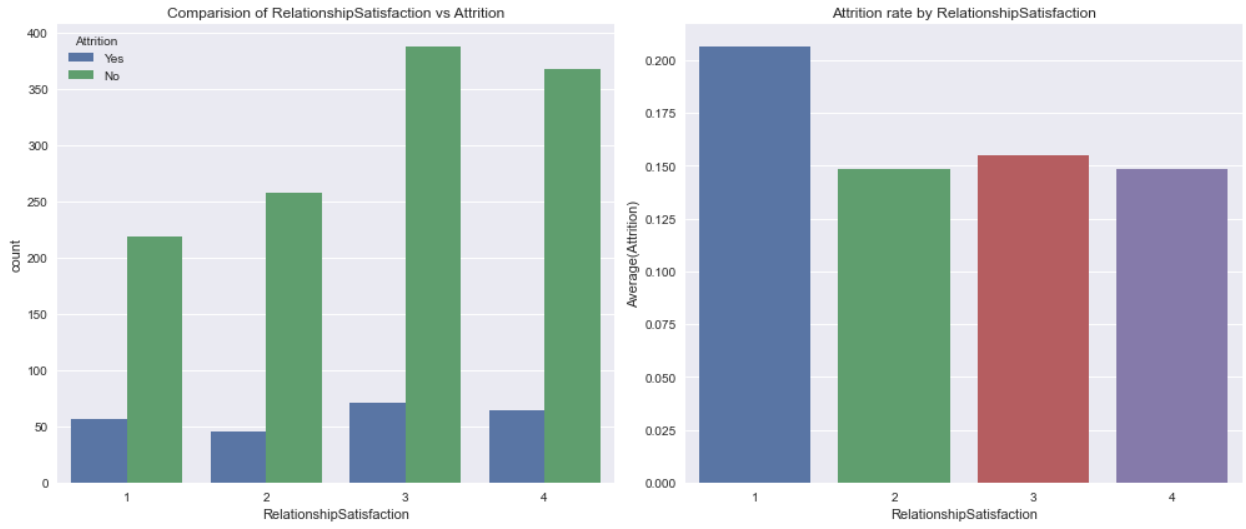
3. Performance rating




```
In [29]: CategoricalVariables_targetPlots(employee_data,"PerformanceRating")
```

- We can see a count of around 200 of employees leaving the organization with high performance rating.
- Average attrition rate is 16% with high performance rating.
- So we can conclude that even though employees are being rated higher in their performance level, there are still chances of them leaving the organization.

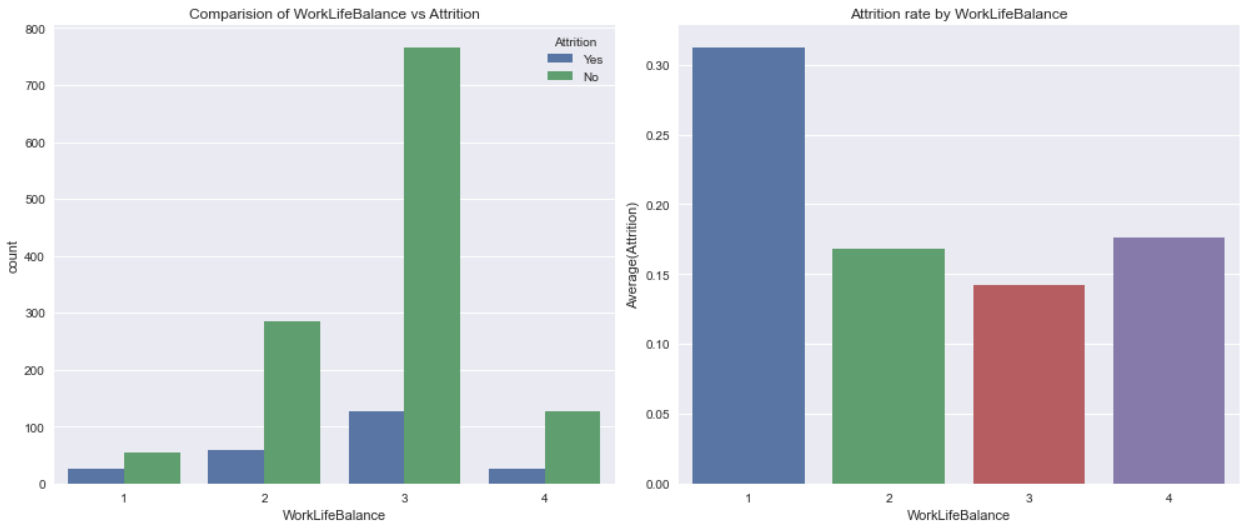
4. Relationship satisfaction




```
In [31]: CategoricalVariables_targetPlots(employee_data,"RelationshipSatisfaction")
```

- The attrition level is same at all the levels of relationship satisfaction.
- The average attrition rate is around 200%.

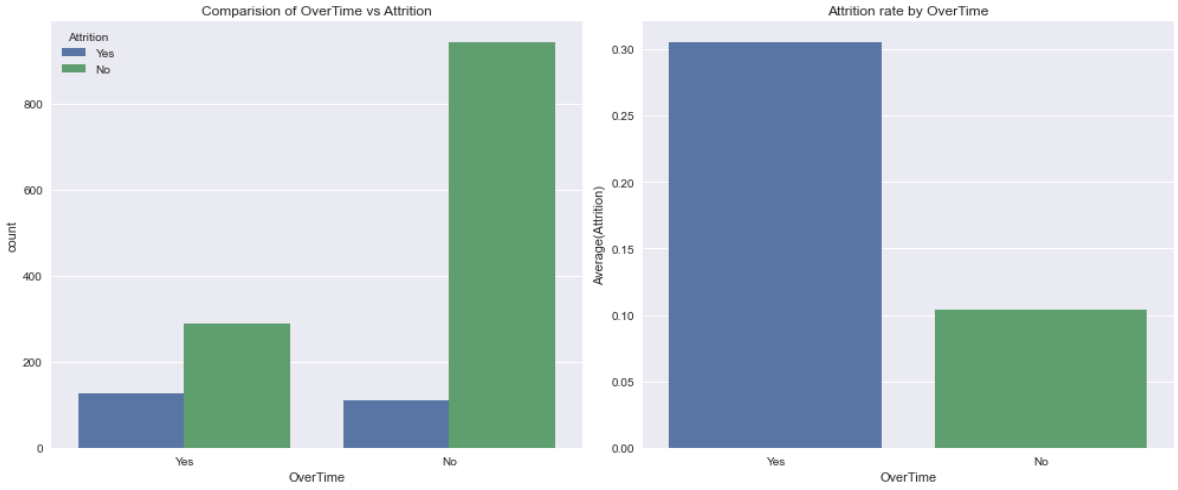
5. Work-life balance




```
In [33]: CategoricalVariables_targetPlots(employee_data,"WorkLifeBalance")
```

- Employees having high work-life balance level are still leaving the organization.
- Hence, it is not right to say that work-life balance of an employee increases the retention level.
- The average attrition level here is around 32%.

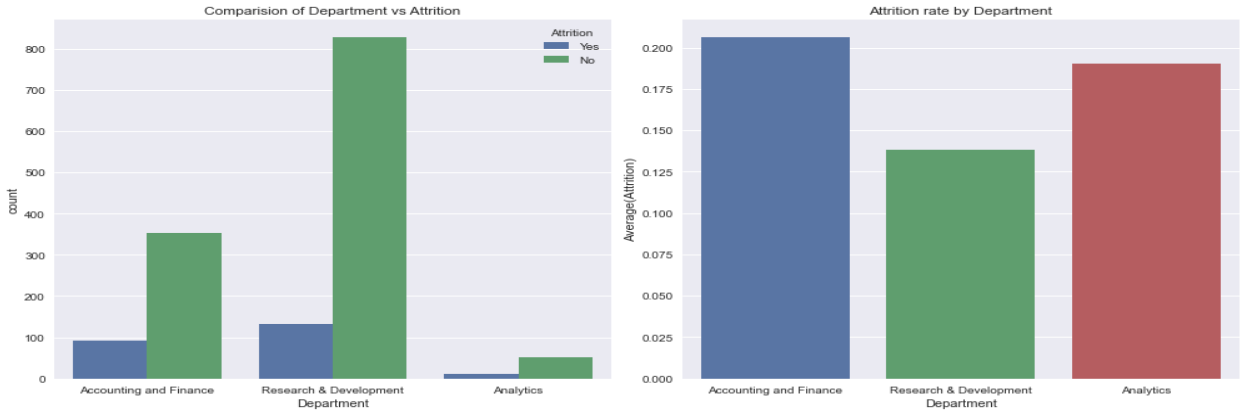
6. Overtime




```
In [34]: CategoricalVariables_targetPlots(employee_data,"OverTime")
```

- Overtime plays an important role generally, when the employees are not treated well or not paid though working overtime.
- But when it comes to attrition level here, the count is almost the same that is around 150 employees still leave the organization.
- The average rate here is around 30%.

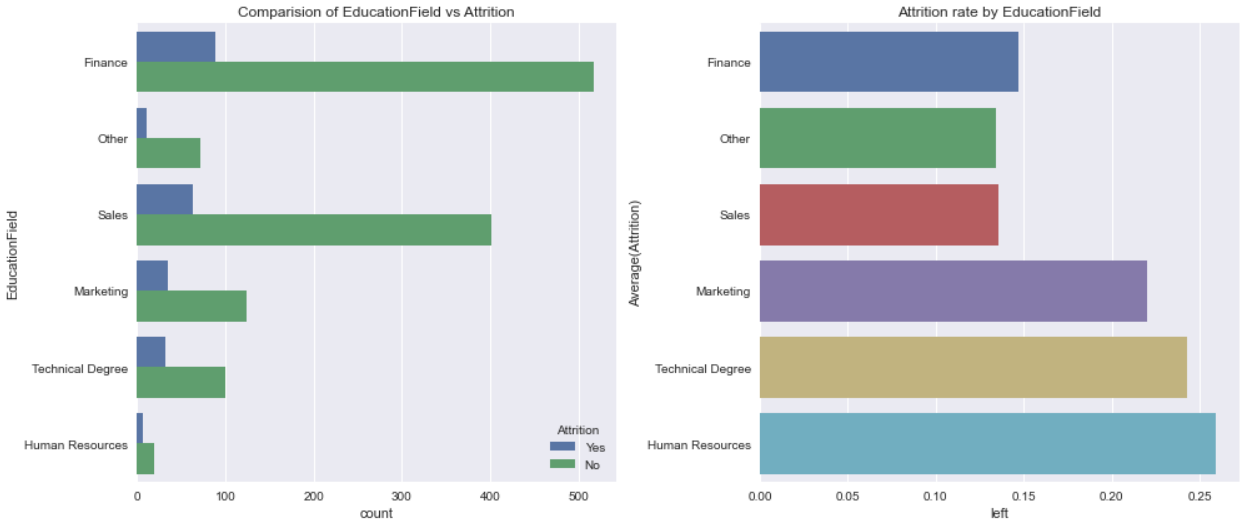
7. Department




```
In [36]: CategoricalVariables_targetPlots(employee_data,segment_by="Department")
```

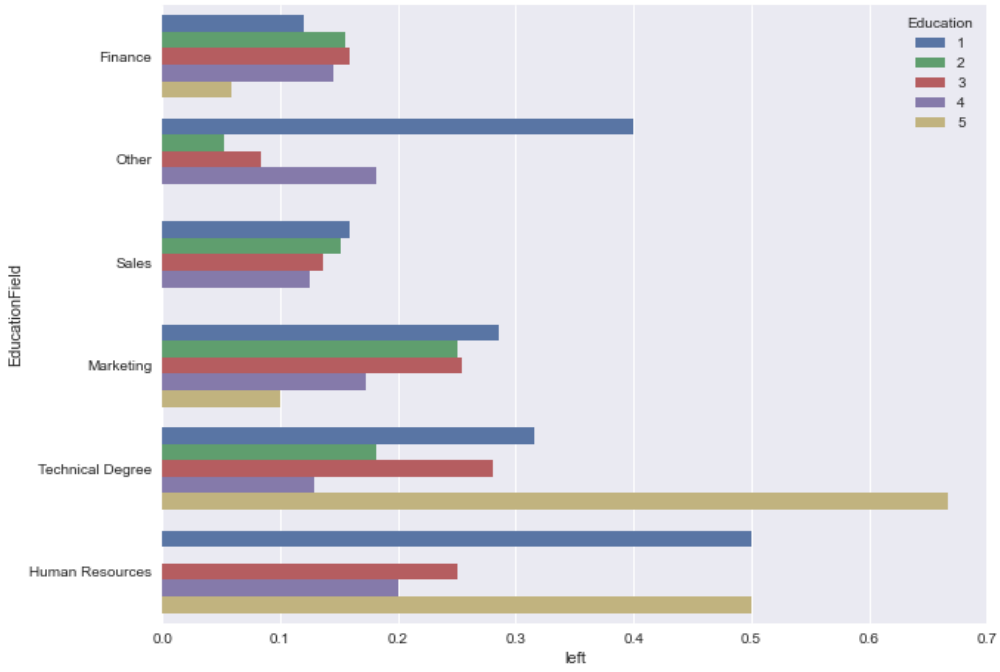
- Employees from analytics department have stayed more in the organization as we can see from here.
- But Research and Development department has seen more turnover rates, where the count has been around 150 or so. Followed by Accounting and Finance.

8. Education Field




```
In [38]: CategoricalVariables_targetPlots(employee_data,"EducationField",invert_axis=True)
```

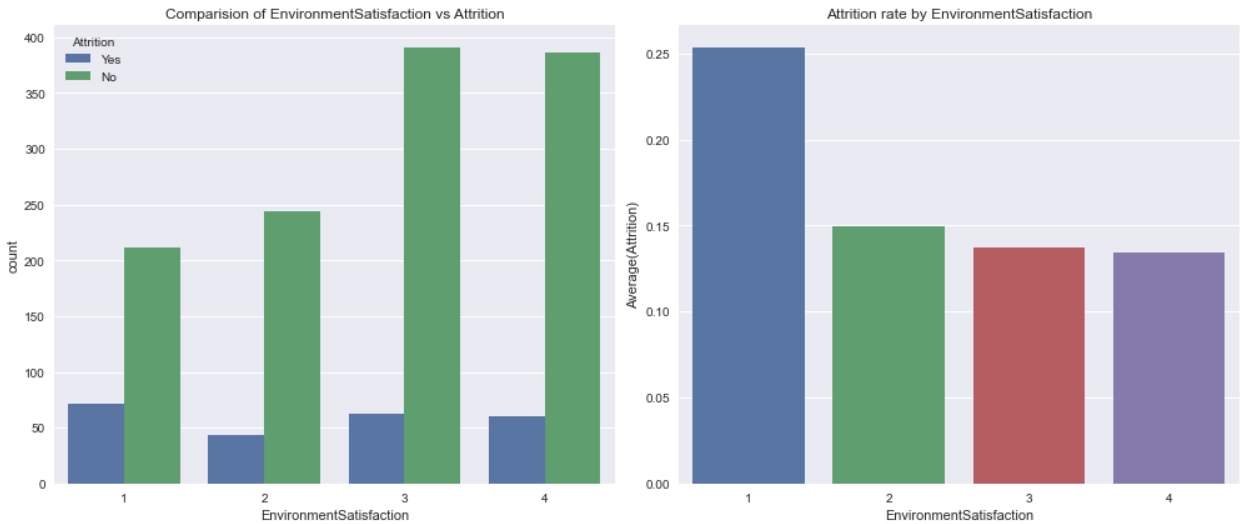
9. Education




```
In [39]: plt.figure(figsize=(10,8))
         sns.barplot(y = "EducationField", x = "left", hue="Education", data=employee_data,ci=None)
         plt.show()
```

- There are more number of employees having their education field as Human Resources or some kind of a Technical degree.
- But we can also see that the employees from these fields have left more too.
- The attrition levels are higher with these education fields.

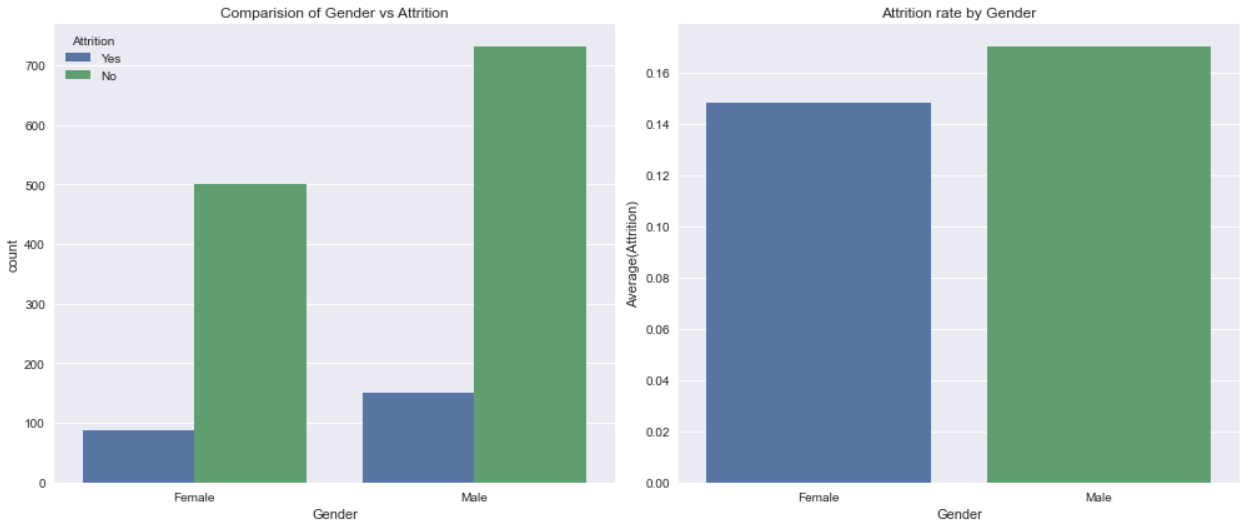
10. Environment Satisfaction




```
In [40]: CategoricalVariables_targetPlots(employee_data,"EnvironmentSatisfaction")
```

- Satisfaction of employees with the environment of their organization is an important factor.
- Here we can see that the average attrition rate is around 25% and the frequency of employees with low level of environment satisfaction is higher too.
- But still, there is no significant difference between all the levels, as employees are still leaving with higher level of environment satisfaction.

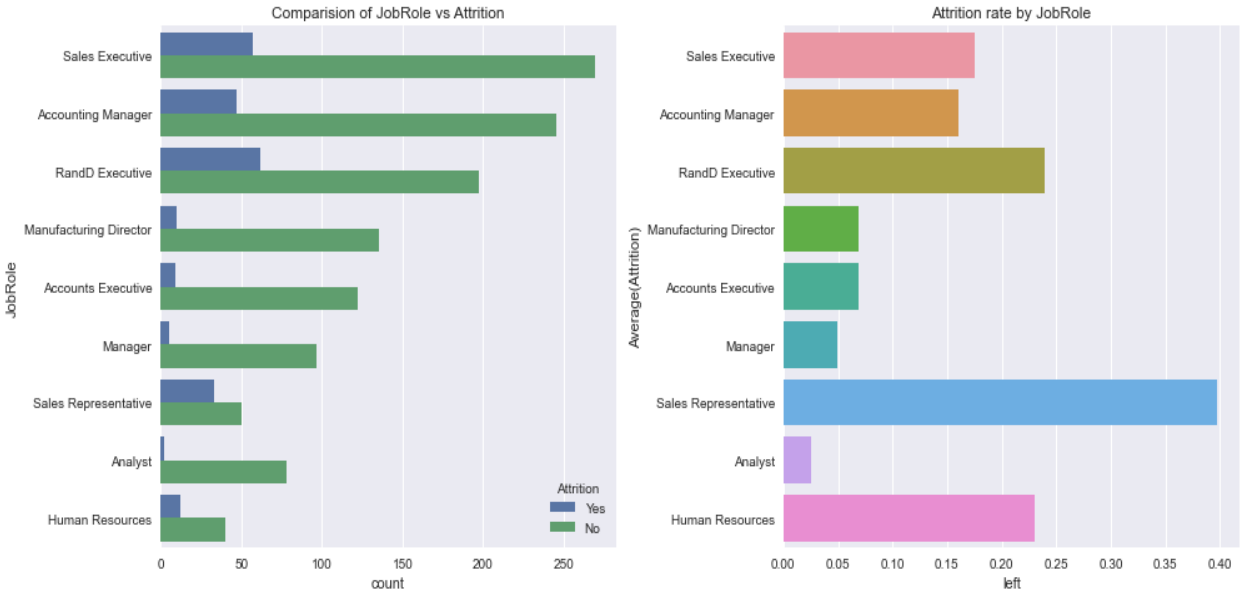
11. Gender




```
In [42]: CategoricalVariables_targetPlots(employee_data,"Gender")
```

- There is a minor difference between the attrition count of male and female.
- The average rate of attrition is also around 16%.
- This does not significantly state the reason since the variable gender is not showing any major difference.

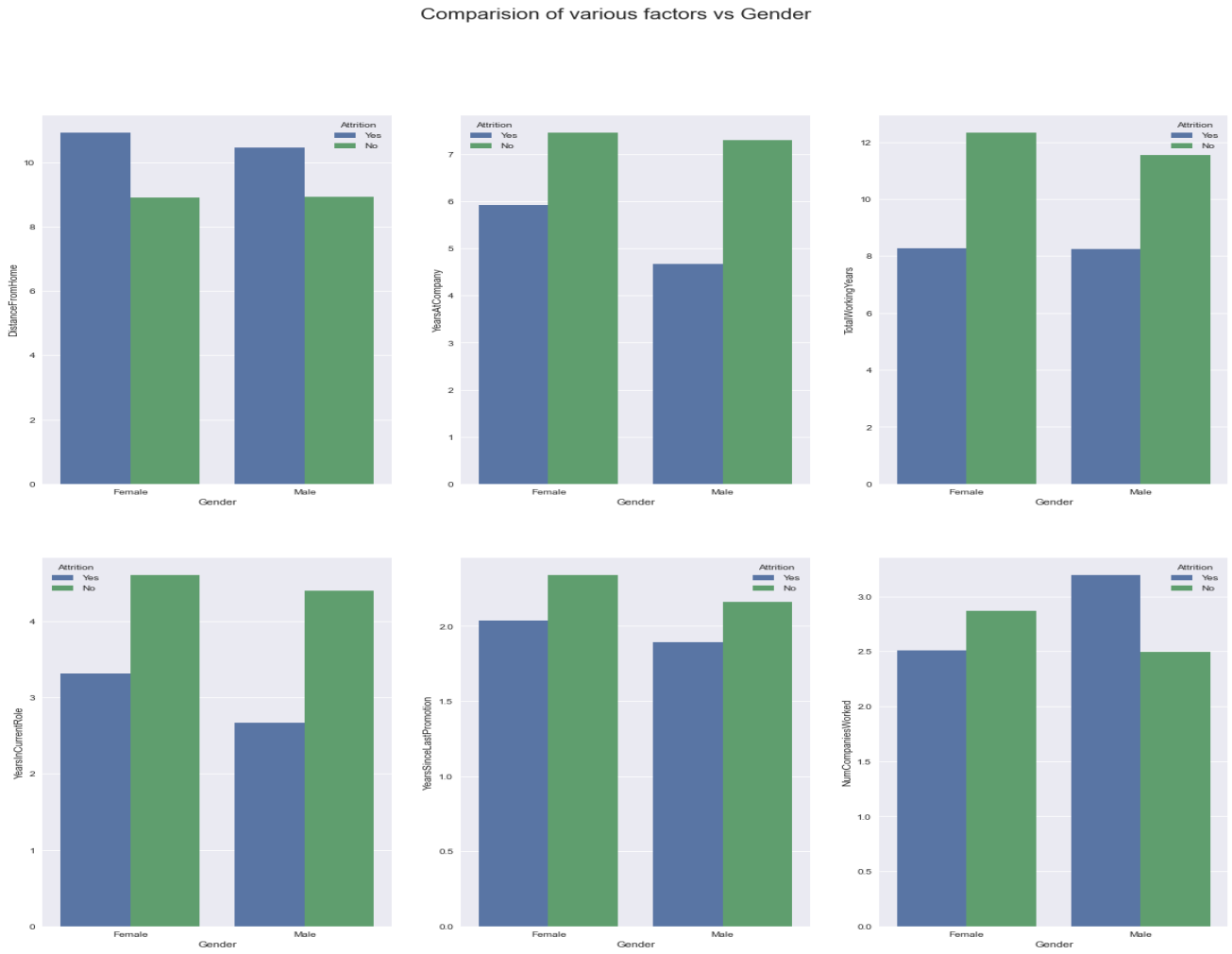
12. Job role



```
In [44]: CategoricalVariables_targetPlots(employee_data,"JobRole",invert_axis=True)
```

- From this we can see that, Sales executive are prone to leave the organization more likely followed by R and D executive and Accounting Manager and Sales representative.
- The attrition rate is higher in Sales representative designated employee, that is around 40%.

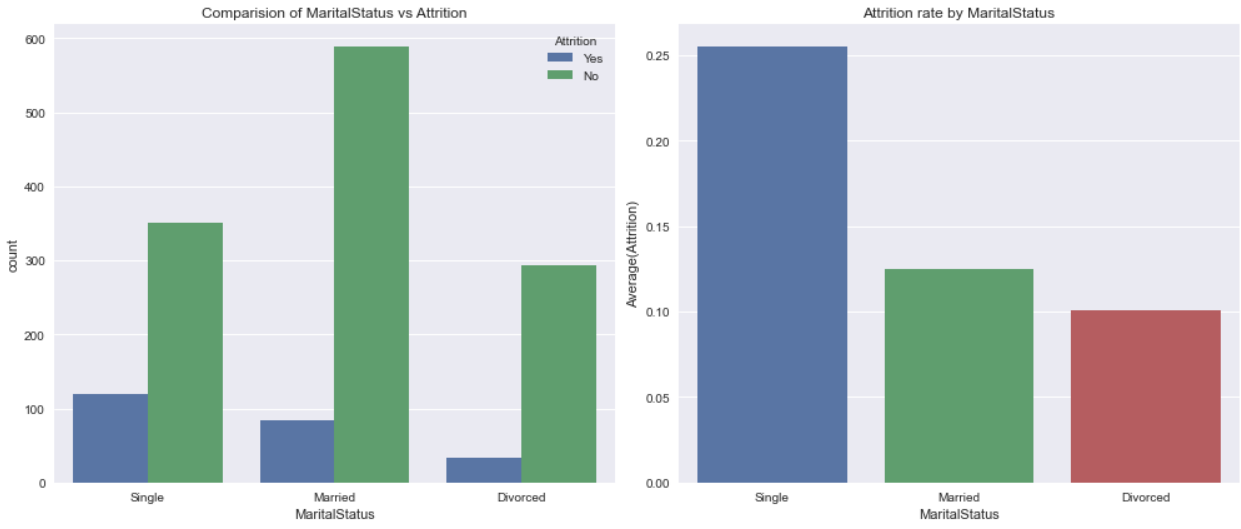
13. Comparing Gender with various other factors




```
In [43]: fig,ax = plt.subplots(2,3, figsize=(20,20))              # 'ax' has references to all the four axes
plt.suptitle("Comparision of various factors vs Gender", fontsize=20)
sns.barplot(employee_data['Gender'],employee_data['DistanceFromHome'],hue = employee_data['Attrition'], ax = ax[0,0],ci=None);
sns.barplot(employee_data['Gender'],employee_data['YearsAtCompany'],hue = employee_data['Attrition'], ax = ax[0,1],ci=None);
sns.barplot(employee_data['Gender'],employee_data['TotalWorkingYears'],hue = employee_data['Attrition'], ax = ax[0,2],ci=None);
sns.barplot(employee_data['Gender'],employee_data['YearsInCurrentRole'],hue = employee_data['Attrition'], ax = ax[1,0],ci=None);
sns.barplot(employee_data['Gender'],employee_data['YearsSinceLastPromotion'],hue = employee_data['Attrition'], ax = ax[1,1],ci=No
sns.barplot(employee_data['Gender'],employee_data['NumCompaniesWorked'],hue = employee_data['Attrition'], ax = ax[1,2],ci=None);
plt.show()
```

- Here we are comparing the factors with Gender, so that we can know how significantly it affects the level of attrition.
- Females are more likely to leave the organization when the distance from home is more.
- They are leaving the organization when the years they have spent is increasing.
- But the total working years does not affect the attrition rate because the average is same for both male as well as female.
- Here the only important factor that significantly determines the rates of attrition is distance from home, where in number of females are more than men.

14. Marital Status

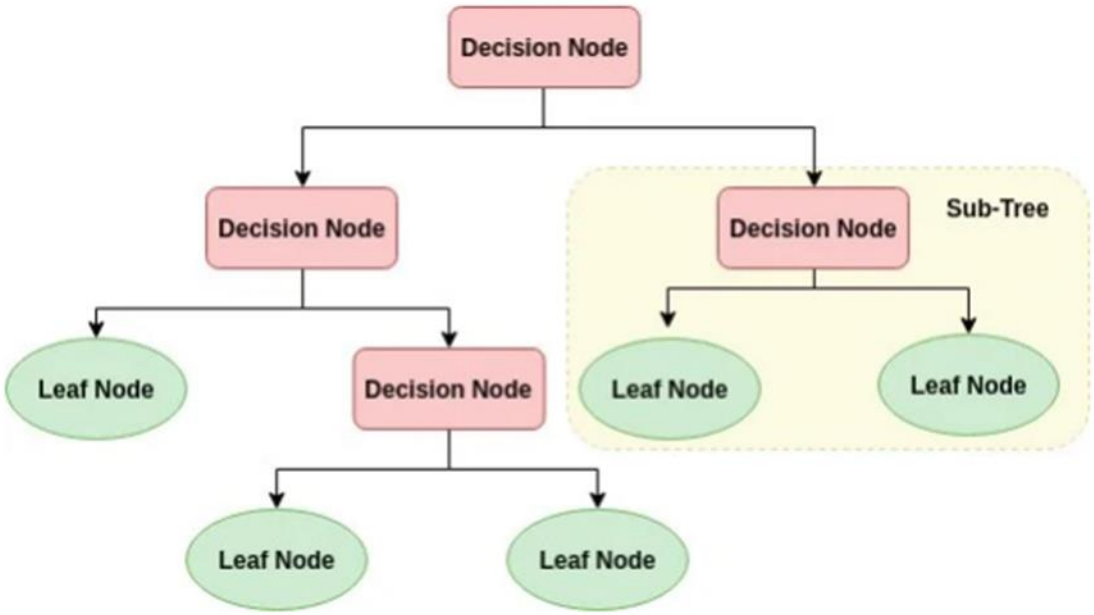



```
In [45]: CategoricalVariables_targetPlots(employee_data,"MaritalStatus")
```

- Marital status of an employee does not significantly determine the increase in attrition.
- Because here we can see that there is only a minor difference between single and married employees leaving the organization.
- The average rate of attrition shown here is around 25% which in it shows that single employees are more likely to leave the organization than the other two.
- This can be due to a number of reasons.
- They might be interested in some other good opportunity, or higher education or career change, anything.

Decision Tree:

An internal node represents a feature (or property), the branch represents a decision rule, and each leaf node indicates the conclusion in a decision tree, which resembles a flowchart. The root node in a decision tree is the first node from the top. It gains the ability to divide data according to attribute values. Recursive partitioning is the process of repeatedly dividing a tree. This framework, which resembles a flowchart, aids in decision-making. It is a flowchart-like representation that perfectly replicates how people think.



In comparison to the Neural Network Algorithm, Decision Tree training takes less time. The amount of records and number of attributes in the given data determine the temporal complexity of decision trees. The decision tree is a non-parametric or distribution-free strategy that does not rely on the assumptions of a probability distribution. High dimensional data can be handled by decision trees accurately.

Decision tree algorithm CART (Classification and Regression Tree) uses the Gini method to create split points.

Gini index

$$\text{Gini}(D)=1-\sum_{i=1}^{m} Pi^2$$

Where, pi is the probability that a tuple in D belongs to class Ci.

The Gini Index considers a binary split for each attribute. You can compute a weighted sum of the impurity of each partition. If a binary split on attribute A partitions data D into D1 and D2, the Gini index of D is:

$$\text{Gini}_A(D)=\frac{|D1|}{|D|}\text{Gini}(D_1)+\frac{|D2|}{|D|}\text{Gini}(D_2)$$

When choosing a splitting attribute for a discrete-valued attribute, the subset that provides the lowest gini index for that attribute is picked. For continuous-valued characteristics, the technique is to choose the point with the smallest gini index as the splitting point after selecting each pair of neighboring values as a potential split-point.

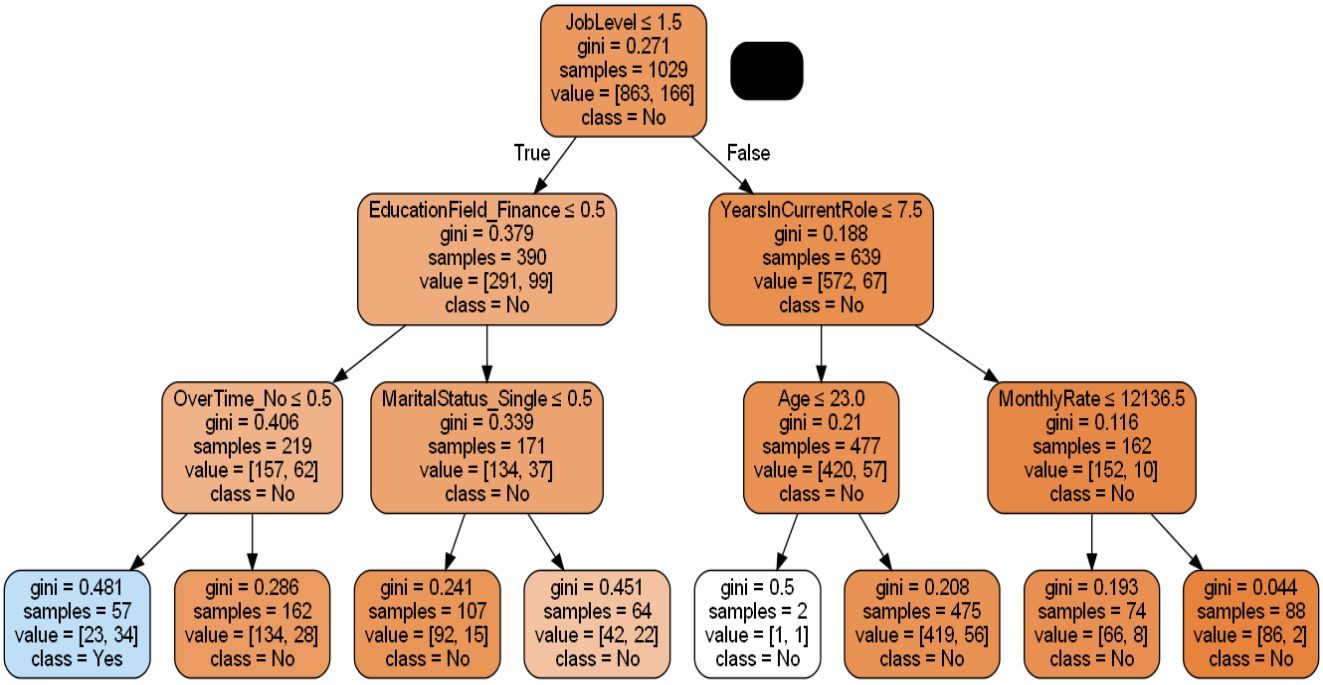
$$\Delta Gini(A) = Gini(D) - Gini_A(D).$$

The attribute with minimum Gini index is chosen as the splitting attribute.

We had to predict the level of attrition in the organization and had to classify the data into a model.

First of all we are doing feature selection that is dividing the columns into target variable and feature variables (Dependent variables and independent variables). Here, our target variable is Attrition and all other variables are independent variables. After feature selection we are splitting the data into training and test data to check the performance of our model.

The following is the Decision Tree model output that we got in Python:



In [91]:
```
from sklearn.tree import export_graphviz
```

In [93]:
```
import six
import sys
sys.modules['sklearn.externals.six'] = six
```

In [94]:
```
from sklearn.externals.six import StringIO
from IPython.display import Image
from sklearn.tree import export_graphviz
import pydotplus as pdot
```

In [95]:
```
#write the dot data
dot_data = StringIO()
```

In [102]:
```
#export the decision tree along with the feature names into a dot file format

export_graphviz(clf_best_model,out_file=dot_data,filled=True,
                rounded=True,special_characters=True,feature_names = X_train.columns.values,class_names = ["No","Yes"])

#make a graph from dot file
graph = pdot.graph_from_dot_data(dot_data.getvalue())
```

In [109]:
```
Image(graph.create_png())
```

- The dependent variable here is Job Level, which states the level of designation of an employee.
- The levels are 1,2,3,4 and 5. The first level is the lowest level whereas the last level is the highest.
- It has two decision criteria that is True or False/ Yes or No.
- If yes then, employees are from the education field finance where again we have decided between the criteria of marital status and overtime.
- So, if they have done overtime, then as shown by the gini index probability that there is 48% of chance of them leaving the organization.
- Since, the data is highly imbalanced due to its fictitious nature as we can also see from the confusion matrix here, it is only showing minor result of 48% and there are 1068 misclassifications in the data itself.

	actual	predicted	predicted_prob	actual_left	predicted_left
34	Yes	Yes	0.509804	1	1
1432	No	No	0.105263	0	0
334	No	No	0.055696	0	0
1068	Yes	No	0.075630	1	0
736	No	No	0.055696	0	0

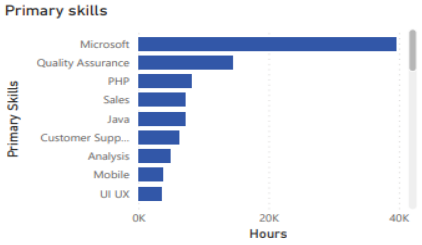
- So to conclude, we can say that employees who are working at lower level in their jobs and belongs to the finance field of education, doing overtime are likely to leave the organization.

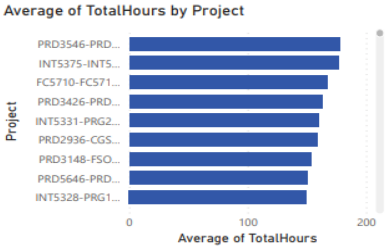
Now, if we understand the Power BI dashboarding, then we can see all the details clearly with the help of the report generated and the visualizations used in the reports.

- All the information regarding the project an employee is working upon, which company clients they are dealing with, how many total hours they are working, their primary skills, their department, reporting manager, present days etc.
- Here is the report that is generated and summarized in the form of a live dashboard, on the basis of which certain organizational decisions are to be taken.

Employee Timesheet Report
June 2021

Client	Project	Employee Name
Company 214	TST-HR194-UB001-Project - 772	Shweta M Name. LN.
Company 3747	TM5975-TM5975-Project - 4934	Abhinav M Name. LN.
Company 3747	TM5975-TM5975-Project - 4934	Ghanshyam M Name.
Company 3747	TM5975-TM5975-Project - 4934	Jigar M Name. LN.
Company 3747	TM5975-TM5975-Project - 4934	Mohammed Zahid M
Company 3747	TM5975-TM5975-Project - 4934	Nilay M Name. LN.
Company 3747	TM5975-TM5975-Project - 4934	Shalini M Name. LN.
Company 3730	-TM5945-Project - 4933	Jay M Name. LN.
Company 3730	-TM5945-Project - 4933	Monika M Name. LN.
Company 3730	-TM5945-Project - 4933	Shrishti M Name. LN.
Company 3729	TM5943-TM5943-Project - 4913	Jigar M Name. LN.





Designation	Reporting Manager	Account Cost Head	Employee Name
VP - Technology	Niraj M Name. LN.	Products	Gitesh M Name. LN.
VP - Technology	Niraj M Name. LN.	Products	Rajesh M Name. LN.
VP - Technology	Niraj M Name. LN.	Services	Gitesh M Name. LN.
VP - Technology	Niraj M Name. LN.	Support	Gitesh M Name. LN.
Visual Designer	Rachna M Name. LN.	Support	Gowtham M Name. LN.
Vice President - Customer Experience	Niraj M Name. LN.	Products	Maanish M Name. LN.
Vice President	Niraj M Name. LN.	Services	Premal M Name. LN.
Vice President	Niraj M Name. LN.	Support	Jigar M Name. LN.
Vice President	Niraj M Name. LN.	Support	Premal M Name. LN.
Trainee	Abhinavkumar M Name. LN.	Products	Harpreet Singh M Name. LN.
Trainee	Aditya M Name. LN.	Support	Kairvi M Name. LN.
Trainee	Amee M Name. LN.	Support	Riya M Name. LN.
Trainee	Digesh M Name. LN.	Support	Aakash M Name. LN.
Trainee	Gopinath M Name. LN.	Support	Kajal M Name. LN.
Trainee	Hardik M Name. LN.	Support	Ravi M Name. LN.
Trainee	Hiren M Name. LN.	Services	Mohammed M Name. LN.
Technical Project Manager	Deval M Name. LN.	Services	Sharvin M Name. LN.
Technical Project Manager	Deval M Name. LN.	Services	Sumit M Name. LN.
Technical Project Manager	Deval M Name. LN.	Support	Sharvin M Name. LN.

Average of Approved PM Hours	Project
0.00	BPO-2576-AA-25-Proj
2.20	BPO-HR112-Bpo12-Pr
0.29	BPO-HR216-BP216-Pr
1.00	FC3111-FC3111-Projec
2.50	FC3365-FC3365-Projec
3.57	FC3768-FC3768-Projec
8.60	FC5041-BIPS1-Project
1.00	FC5047-FC5047-Projec
1.86	FC5147-FC5147-Projec
4.99	FC5215-FC5215-Projec
0.00	FC5215-TM-GAH02-Pr
5.16	FC5223-FC5223-Projec
4.21	FC5329-FC5329-Projec
3.15	FC5415-FC5415-Roker
6.26	FC5458-FC5458-Projec
4.86	FC5467-FC5467-Projec
0.65	FC5547-FC5547-Projec
5.12	FC5548-FC5548-Projec
5.17	

128.96
Average of TotalHours

13.09
Average of PresentDays

Suggestions:

It is stated that employees account for 70% of a company's costs, so it is critical for businesses to comprehend to know the reason behind the leaving of employees. There are four types of employee turnover namely, Voluntary, Involuntary, Functional and Dysfunctional. Voluntary is when an employee leaves the job with his/her will. Involuntary is when any employee is dismissed by the organization. Functional is when any organization itself dismisses the poor or inefficient employees and Dysfunctional is when any organization loses its high-performing and efficient employees.

Organizations must look into the reasons or matters behind this attrition especially functional and dysfunctional. All the relevant factors significantly affecting the attrition rates must be taken into account while formulating a rigorous HR process. During the process of hiring, managers or assessors should look into a potential employee and whether he/she is a good fit for their organization or not to decrease the attrition rates. Though any employee leaving the organization is not in the hands of the manager but it can still be prevented with the right plan and strategies.

Cultural factor is also an important one that has to be looked upon. Understanding an organization's culture will help you choose individuals who will fit in well with the position, the team, and the larger organization. Understanding the company's culture as a whole and a candidate's compatibility with that culture are key components in evaluating organization fit. This could begin with the stated mission and values of the company. Unsatisfactory cultural fit must be clearly identified as a risk factor in a formal cultural interview and evaluation. There will undoubtedly be disparities in underlying values and working methods if a candidate moves from one industry to another or from a big organization to a small one. Hiring managers must be upfront with candidates about potential disparities and determine whether they are insurmountable. In this way, managers, recruiters, or assessors should plan and strategize while the hiring process. This can reduce the rates of attrition.

And now if we talk about timesheet reporting, then project managers should rely on effective dashboarding for the achievement of their goals. Reduction in hours, project time, timely completion of projects, leaves, everything can be summarized using dashboards or reports with Power BI.

Limitations:

This research cum analysis is only limited to attrition and timesheet data. Also, this data is fictitious as well as of staging application. Further, these methods or techniques of classification models of machine learning can be applied to the real data of any organization.

Contribution:

Timesheet reporting using Power BI has contributed greatly to Cygnet Infotech’s Project Management and Management Information Systems team as well as has helped other senior authorities in making effective decisions regarding Payroll, Project completion, leaves, etc.

Bibliography:

1. DiClaudio, (2019) "People analytics and the rise of HR: how data, analy,tics and emerging technology can transform human resources (HR) into a profit center", Strategic HR Review, https://doi.org/10.1108/SHR-11-2018-0096
2. LAL , (2015),"Transforming HR in the digital era", Human Resource Management International Digest, Vol. 23 Iss 3 pp. 1 – 4, http://dx.doi.org/10.1108/HRMID-03-2015-0051
3. http://www.htpub.org/Journal-Of-Social,-Management-And-Tourism-Letter/
4. Hota, (2013) "Workforce Analytics Approach: An Emerging Trend of Workforce Management
5. Marsden, T. (2016), "What is the true cost of attrition?", Strategic HR Review, Vol. 15 No. 4, pp. 189-190, https://doi.org/10.1108/SHR-05-2016-0039
6. https://www.statista.com/topics/2256/it-industry-in-india/
7. https://cygnet-infotech.com/
8. https://www.ibm.com/in-en/cloud/learn/exploratory-data-analysis
9. https://www.techtarget.com/searchhrsoftware/definition/workforce-management
10. https://www.investopedia.com/terms/a/attrition.asp
11. https://scikit-learn.org/stable/modules/tree.html

Annexures:

- Attrition Data of the organization (Fictitious)
- Timesheet Data (Staging Application Data)
- Power BI Report
- EDA codes Python file