**MSBC 5490**

**BUAN Experiential Project**

**Primoris**

**Agile and Data Report**

**Submitted by Group 4:**

| **Name** | **Email Address** |
| --- | --- |
| Rafael Cintron | rafael.cintron@colorado.edu |
| Supria Deka | supria.deka@colorado.edu |
| Pranathi Manthri | pranathi.manthri@colorado.edu |
| Murali Prateek Manthri | murali.manthri@colorado.edu |
| Jinal Mehta | jinal.mehta@colorado.edu |

**TABLE OF CONTENTS**

[**1. Introduction 3**](#_div8zoi7rbll)

[Project Vision Statement 3](#_ymauglelbgcu)

[**2. Project Management Elements 3**](#_g8gb1tc5m8lk)

[2.1 Business Understanding 3](#_ng6ekv3v9d0u)

[2.2 Problem Statement 3](#_q2y0eekwqcgo)

[**3. Agile/Planning Elements 4**](#_7jkfc3jd7a1u)

[3.1 Agile Software Choice 4](#_7pvcy6av7s0f)

[3.2 Team Members 4](#_c6wyddkqxkfr)

[3.3 Key Performance Indicators (KPIs) 5](#_qmys37ql9bnb)

[**4. Data Report 7**](#_58ka2xol2hl1)

[4.1 Data Description 7](#_voeu16bwyxgu)

[4.2 Dataset Structure 7](#_eljyx68v1m8v)

[4.3 Metadata 8](#_i0xunvcq444)

[4.4 Data Analysis Tools/Software 9](#_cifqusnbcz2b)

[4.5 Data Exploration 10](#_lxi7jmuu2mdp)

[4.6 Data Visualization 11](#_pouoeb4w8yof)

[4.5 Data Issues 16](#_h9wwdnkgwx93)

[**5. Data Challenges and Considerations 17**](#_30qe6t34loz)

[**6. Future Enhancements & Data Refinement 17**](#_rnta5ddb4lav)

[**7. Appendix 18**](#_44u9phw3ub1p)

[Generative AI Usage 18](#_z3yz0j6dvfv9)

# **1. Introduction**

## **Project Vision Statement**

“At Primoris Renewable Energy, we believe that people are the foundation of sustainable growth. As we continue to lead in utility-scale solar, energy storage, and renewable fuels, we recognize the need to cultivate a thriving workforce that is engaged, empowered, and prepared for the future. Our vision is to harness data-driven workforce insights to optimize employee retention, improve professional development, and refine workforce planning—ensuring that every team member has the opportunity to grow alongside our expanding business.”

This project will leverage predictive analytics to understand workforce trends, employee satisfaction, and proactively address talent management challenges. Through an agile approach, we will refine workforce strategies that drive engagement and create an environment that fosters long-term career success and organizational stability.

# **2. Project Management Elements**

## **2.1 Business Understanding**

Primoris Renewable Energy is a division of Primoris Services Corporation, specializing in large-scale solar energy, battery storage, and renewable fuel projects. As a leading engineering, procurement, and construction (EPC) provider, Primoris is committed to advancing clean energy solutions. With a rapidly growing workforce and expanding operations, the company recognizes the need to enhance workforce planning and employee retention strategies. Through this project, we aim to use data-driven insights to optimize workforce management, improve employee engagement, and drive long-term sustainability.

## **2.2 Problem Statement**

As Primoris Renewable Energy continues its rapid expansion in the renewable energy sector, the company faces significant workforce management challenges. The need to attract, develop, and retain top talent is critical to sustaining long-term growth and operational excellence. However, with a growing workforce, the company struggles to effectively analyze employee movement patterns, predict future workforce needs, and implement proactive retention strategies.

Key workforce challenges include:

* **Workforce Trend Analysis:** Understanding employee career progression, job transitions, and factors influencing long-term retention.
* **Predictive Workforce Planning:** Developing data-driven models to anticipate hiring demands, identify turnover risks, and optimize workforce distribution.
* **Employee Engagement & Satisfaction:** Enhancing onboarding, training, and professional development programs to improve job satisfaction and reduce attrition.

# **3. Agile/Planning Elements**

## **3.1 Agile Software Choice**

**Tool Selection: JIRA**

**Justification:** Jira provides a structured and robust platform for agile project management. It is well-suited for managing workforce analytics projects due to its:

* Sprint planning and backlog tracking capabilities, ensuring efficient progress monitoring.
* Kanban and Scrum boards for streamlined task management.
* Advanced reporting features that allow the team to visualize project metrics and insights.
* Seamless integration with data analytics tools for improved workflow automation.

By utilizing Jira, the team can enhance collaboration, track project milestones effectively, and ensure an agile approach to workforce optimization.

## **3.2 Team Members**

* **Jinal Mehta — Product Owner:** Responsible for defining the project vision, maintaining the product backlog, and ensuring alignment with business objectives. Jinal acts as the bridge between stakeholders and the development team, prioritizing tasks and refining user stories to deliver maximum value.
* **Rafael Cintron** **— Scrum Master:** Facilitates agile processes, ensuring smooth sprint execution, removing roadblocks, and promoting collaboration within the team. Rafael helps maintain team productivity by implementing agile best practices and tracking project progress.
* **Supria Deka** **— Developer:** Focuses on data integration, report automation, model building and implementing workforce analytics tools. Supria collaborates with the team to ensure that technical solutions align with project objectives.
* **Pranathi Manthri** **— Developer:** Works on backend processing, data transformation, and the development of predictive models. Pranathi ensures that analytics models provide meaningful workforce insights that drive decision-making.
* **Murali Prateek Manthri** **— Validation Analyst:** Ensures data accuracy, validates workforce predictions, and improves overall reporting reliability. Murali also works closely with the development team to test automated analytics systems and refine data visualization techniques.

## **3.3 Key Performance Indicators (KPIs)**

1. **Employee Retention Rate :** Employee retention rate is a critical measure of workforce stability and engagement. A higher retention rate suggests that employees are satisfied with their work environment, career growth opportunities, and overall job experience. By tracking retention trends across different departments, Primoris Renewable Energy can identify patterns that lead to longer employee tenure. Reducing turnover also minimizes hiring and training costs while fostering a more experienced workforce. The insights gained from retention analysis will guide strategic decisions in employee engagement and workforce planning.
2. **Workforce Satisfaction Index:** The workforce satisfaction index evaluates employee morale, motivation, and overall workplace engagement. Using surveys, feedback mechanisms, and sentiment analysis, this metric provides insights into how employees perceive job satisfaction, work-life balance, and career advancement opportunities. A high satisfaction index reflects a positive and productive workplace, while a lower score may indicate issues requiring immediate attention. By consistently monitoring workforce sentiment, Primoris can implement targeted initiatives to enhance employee well-being and workplace culture.
3. **Talent Acquisition Efficiency:** Efficient talent acquisition ensures that the company can quickly attract and onboard top talent without compromising quality. This metric measures the time taken to fill vacant positions, the success rate of recruitment efforts, and the retention of new hires. An optimized hiring process helps maintain workforce continuity and reduces disruptions caused by long vacancy periods. By improving recruitment efficiency, Primoris Renewable Energy can better align talent acquisition strategies with long-term business needs.
4. **Training Program Impact:** Training programs play a vital role in workforce development and career growth. This KPI assesses how effectively employees apply newly acquired skills in their roles after training. Performance improvements, productivity gains, and career progression post-training are analyzed to determine the program’s success. Well-structured training enhances job proficiency, reduces knowledge gaps, and prepares employees for leadership roles. Tracking these outcomes allows Primoris to continuously refine its learning and development initiatives.
5. **Turnover Prediction Accuracy:** Predictive analytics can help forecast employee turnover and identify individuals at risk of leaving the company. This KPI measures how accurately predictive models anticipate attrition trends based on various factors such as job satisfaction, career progression, and workload balance. By leveraging data insights, Primoris Renewable Energy can take proactive measures to retain employees, improve engagement strategies, and mitigate risks associated with high turnover. Accurate turnover predictions allow the company to build a more stable and committed workforce.
6. **Tableau Dashboard Implementation:** A Tableau dashboard will be developed to visualize workforce trends, key metrics, and actionable insights. This dashboard will provide a dynamic and interactive way for stakeholders to analyze real-time data related to employee retention, workforce satisfaction, and hiring efficiency. With customizable filters and graphical representations, decision-makers can quickly identify trends, outliers, and performance indicators to drive strategic workforce planning.
7. **Workable Predictive Model:** Two predictive models will be developed to enhance workforce analysis. The first model will focus on predicting employee attrition based on historical trends and engagement factors. The second model will analyze training effectiveness, evaluating how training participation impacts performance and career progression. These models will provide actionable insights, helping Primoris make data-driven decisions to optimize workforce stability and development.
8. **Feature for Workforce Recommendation System:** An important feature for workforce recommendations will be integrated into the analysis. This feature will provide tailored suggestions for employee training, career advancement opportunities, and retention strategies based on individual performance and engagement metrics. By leveraging AI-driven insights, Primoris Renewable Energy can personalize employee development plans, ensuring long-term job satisfaction and alignment with business goals.

# **4. Data Report**

## **4.1 Data Description**

This dataset contains various employee-related records for ***Primoris Renewable Energy***, covering aspects such as new hires, promotions, demotions, rehires, terminations, and job title changes. The data spans multiple sheets, each representing a different employment event.

## **4.2 Dataset Structure**

File Format: The dataset is provided in Excel format(.xlsx) consists of 9 sheets, each representing different HR activities.

Total Number of Rows and Columns:

| **Sheet Name** | **Rows** | **Columns** |
| --- | --- | --- |
| ALL Employee | 12364 | 11 |
| New Hires | 11585 | 13 |
| Promotions | 1840 | 13 |
| Demotions | 24 | 13 |
| Rehires | 2338 | 13 |
| TransferIns | 7475 | 13 |
| TransferOuts | 99 | 13 |
| Terms(Terminations) | 18180 | 10 |
| ALL Job Title Changes | 7278 | 12 |

Table 1: Number of Records/Column in .xlsx file.

File Size: The size of the dataset provided is approximately 3.8 MB and the memory usage in colab notebook is 4.1 MB.

General Structure: The dataset captures employee movements and job changes within the company, including *new hires, promotions, demotions, rehires, transfers, and terminations*. It consists of categorical fields such as *Employee ID, Job Title, Project Name, Reason for Change*, along with *date-based* attributes like *Hire Date, Termination Date, and Effective Date*. Each row represents an individual employee's record for a specific HR event.

## **4.3 Metadata**

Definitions of Key Columns: The column names and its role are mentioned as follows, most sheets share the following key attributes:

* *Co #*: Company identification number.
* *Co Name*: Name of the company.
* *EE # (Employee Number)*: Unique identifier for each employee.
* *Job Title:* Employee’s job title.
* *Project & Project Code:* Project associated with the employee.
* *Employment Status:* Active (T) or Terminated.
* *Hire and Termination Dates:* Key employment dates.
* *Reason:* Explanation for employee status changes (e.g., resignation, promotion, demotion).

Column Types:

| **Categorical Columns (Qualitative attributes)** | **Numerical(Continuous) Columns (Measurable Quantities)** | **Date Columns** |
| --- | --- | --- |
| Employee ID | Length of Service (Days) | Original Hire Date |
| Job Title | Project Code | Last hire Date |
| Employment Status |  | Termination Date |
| Reason for Termination |  | Effective Date |
| * Categorical columns categorize employees based on job title, employment status, and termination reasons. * Numerical columns provide measurable information, such as length of service and project codes. * Date columns track key events like hire date, termination date, and effective date. * Possible data inconsistencies include missing termination dates for active employees and variations in job titles that may require standardization. | | |

Table 2: Detailed Columns Type at glance.

#### 

#### 

## **4.4 Data Analysis Tools/Software**

**Tools Used/Prefered :** Python (Google Colab Notebook & Pandas), Excel, and Tableau.

To do the analysis, we utilized ***Python, Excel*,** and ***Tableau*** to efficiently process, analyze, and visualize the dataset. Each tool played a crucial role in different stages of the workflow, from data loading to exploratory analysis and visualization.

**Python (Google Colab) :**

Python Google Colab Notebook is used as it is cloud based, allowing our team to perform collaboration and real -time coding. The following libraries were used:

* *pandas:* For loading, cleaning, and manipulating structured data.
* *numpy:* For numerical computations, including handling missing values.
* *matplotlib & seaborn:* For data visualization to identify trends, distributions, and correlations in employee movement.
* *scikit-learn (if predictive modeling is needed)*: For machine learning applications like employee retention prediction.

We were able to execute code interactively, document insights, and visualize results efficiently.

**Excel :**

Microsoft Excel was used as a preliminary analysis tool for as follows :

* Data exploration and quick inspections (checking missing values, inconsistencies, and duplicates).
* Pivot tables & filtering to analyze employee movements across departments, projects, and job roles.
* Initial data cleaning, including formatting date fields and identifying outliers.

Excel’s familiar interface made it an excellent tool for quick insights before moving into Python for deeper analysis.

**Tableau :**

For enhanced data visualization, we will be using Tableau to create interactive dashboards and charts which will help stakeholders to better understand trends in the data. This included:

* Attrition and retention analysis to visualize termination reasons over time.
* Employee movements (transfers, promotions, demotions, rehires) across different departments.
* Project-based employment trends, tracking workforce distribution across various projects.

Tableau enabled dynamic filtering for detailed analysis by time periods, job titles, and termination causes.

## **4.5 Data Exploration**

We performed the Exploratory Data Analysis (EDA) of the given dataset and we found the following *unique values* and *counts* for categorical columns.

| **Sheet Name** | **Key Categorical Columns** | **Unique Values Count** |
| --- | --- | --- |
| **ALL Employees** | Job Title, Employment Status, Project, Reason Code | Job Title: 151, Status: 3, Project: 109, Reason: 23 |
| **New Hires** | FLSA Category, PR Dept Desc., Craft Class Desc., Project | FLSA: 2, PR Dept: 31, Craft Desc: 384, Project: 107 |
| **Promotions** | FLSA Category, PR Dept Desc., Craft Class Desc., Project | FLSA: 2, PR Dept: 26, Craft Desc: 222, Project: 62 |
| **Demotions** | FLSA Category, PR Dept Desc., Craft Class Desc., Project | FLSA: 2, PR Dept: 6, Craft Desc: 20, Project: 16 |
| **Rehires** | FLSA Category, PR Dept Desc., Craft Class Desc., Project | FLSA: 2, PR Dept: 14, Craft Desc: 208, Project: 62 |
| **TransferIns** | FLSA Category, PR Dept Desc., Craft Class Desc., Project, Reason | FLSA: 2, PR Dept: 21, Craft Desc: 224, Project: 65, Reason: 2 |
| **TransferOuts** | FLSA Category, PR Dept Desc., Craft Class Desc., Project, Reason | FLSA: 2, PR Dept: 20, Craft Desc: 49, Project: 15, Reason: 1 |
| **Terms** | Job Title, Employment Status, Project, Termination Reason | Job Title: 126, Status: 3, Project: 98, Termination Reason: 24 |
| **ALL Job Title Changes** | Field Updated, Previous Job Title, Current Job Title, EE Status, Segment | Field Updated: 2, Previous Job Title: 444, Current Job Title: 390, EE Status: 3 |

Table 3: Unique Values and Counts for Categorical Columns

The table presents an overview of key categorical columns and their unique value counts across various employee datasets. The "ALL Employees" sheet has the most diverse job titles (151) and projects (109), while "New Hires" and "Rehires" highlight FLSA categories, department descriptions, and project allocations with varying levels of categorization. "Promotions," "Demotions," and "TransferIns" show similar categorical distributions, focusing on PR Department descriptions and craft classifications, with "Promotions" having the most diversity in job roles.

Additionally, "TransferOuts" has relatively lower categorical diversity, with only 15 unique project assignments. The "Terms" dataset captures employment status, job titles, and termination reasons, while "ALL Job Title Changes" tracks modifications in job titles and employee statuses, revealing 444 unique previous job titles and 390 current job titles. These insights help in understanding workforce movement, hiring trends, and job classification variations across the organization.

## **4.6 Data Visualization**

1. Voluntary and involuntary Churn in Primoris

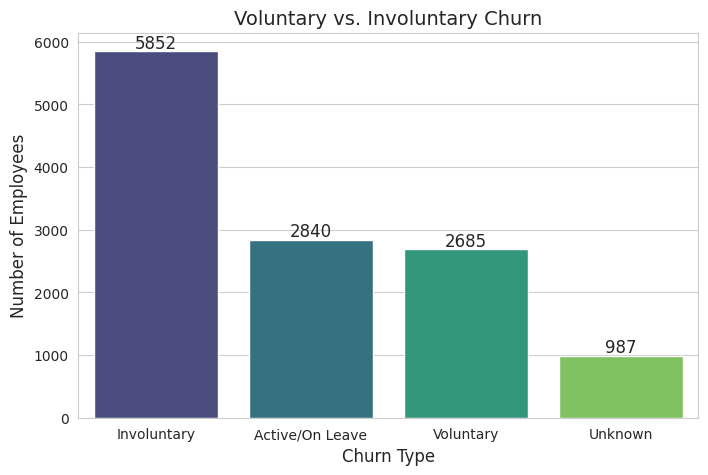


Fig. 1: Voluntary and Involuntary Churn

The data reveals that involuntary churn is the highest (5,852 employees), indicating potential workforce instability due to layoffs, performance issues, or restructuring. Voluntary churn, at 2,685 employees, suggests a considerable number leaving for better opportunities, retirement, or personal reasons, highlighting potential concerns around employee satisfaction and retention. Additionally, 2,840 employees remain active or on leave, necessitating closer monitoring to assess return rates and long-term workforce engagement. The involuntary reasons taken into account are : "Death", "Discharge", "Employment Eligibility", "Excessive Absenteeism", "Policy Viol-Safety", "Policy Viol-Other", "Lack of Work", "Layoff Restructuring", "No Call-No Show", "Performance", "Policy Viol-Drug/Alcohol", "Position Eliminated", "Seasonal Layoff".

A notable 987 employees have an unknown churn reason, emphasizing the need for improved HR data collection and classification. Addressing the high involuntary churn through better workforce planning and retention strategies, while also understanding voluntary departures, will be crucial for organizational stability. Cleaning up missing churn data can further enhance decision-making and workforce forecasting. The Voluntary reasons taken into consideration are: “Moving”, "Other Job", "Personal/family", "Resignation","<None>", "Return to School", "Resignation/Consultant", "Retirement/Early Retire", "Retirement/Consultant".

1. Attrition Rate Analysis based on *terms* sheet:

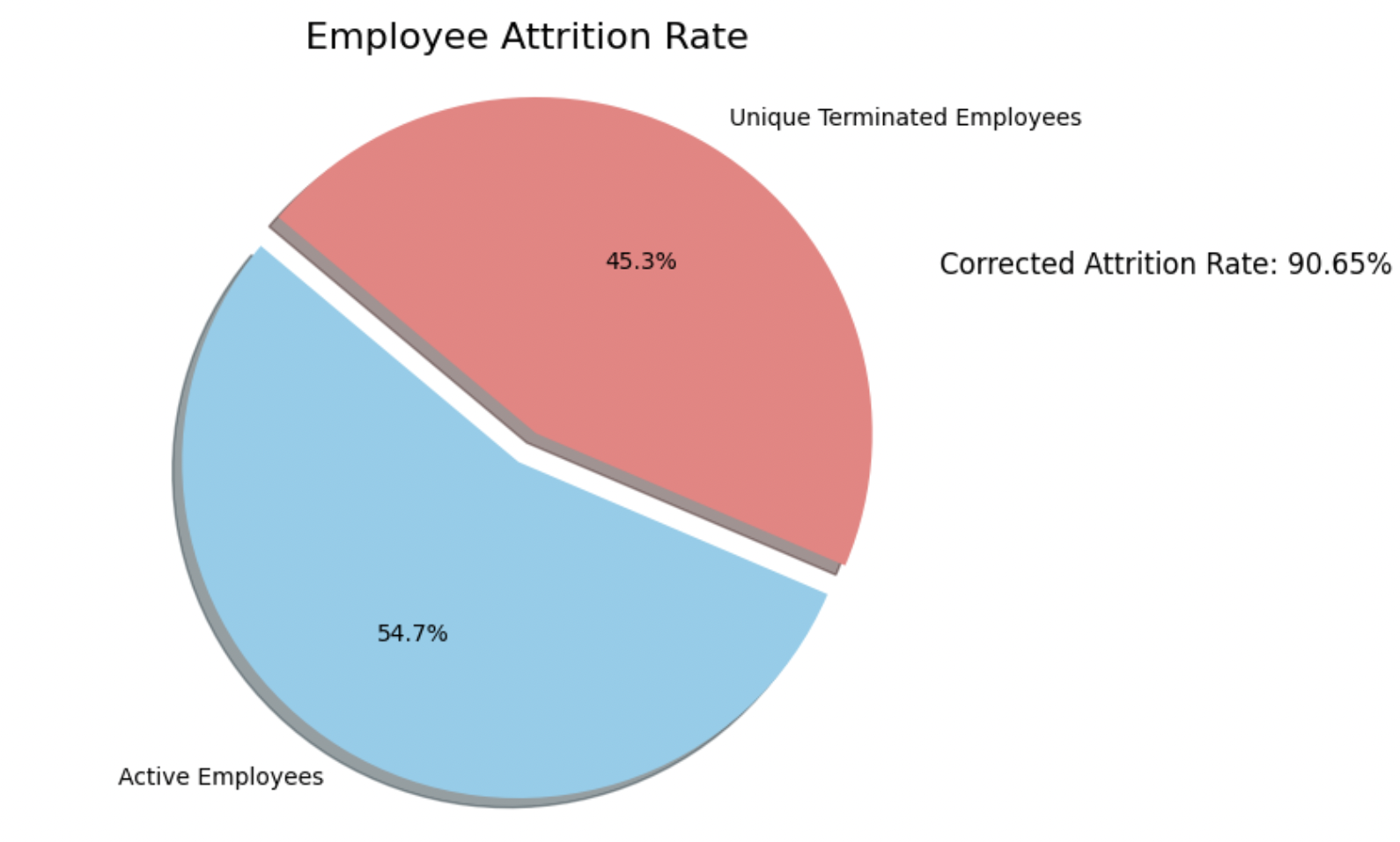


Fig. 2: Employee Attrition Rate

The pie chart illustrates the Employee Attrition Rate, where 45.3% of the workforce has been terminated, and 54.7% remain active. The attrition rate of 90.65% highlights a high turnover issue, suggesting potential instability in workforce retention. This could be due to voluntary resignations, layoffs, or other organizational factors.

1. Promotion and Demotion Trends

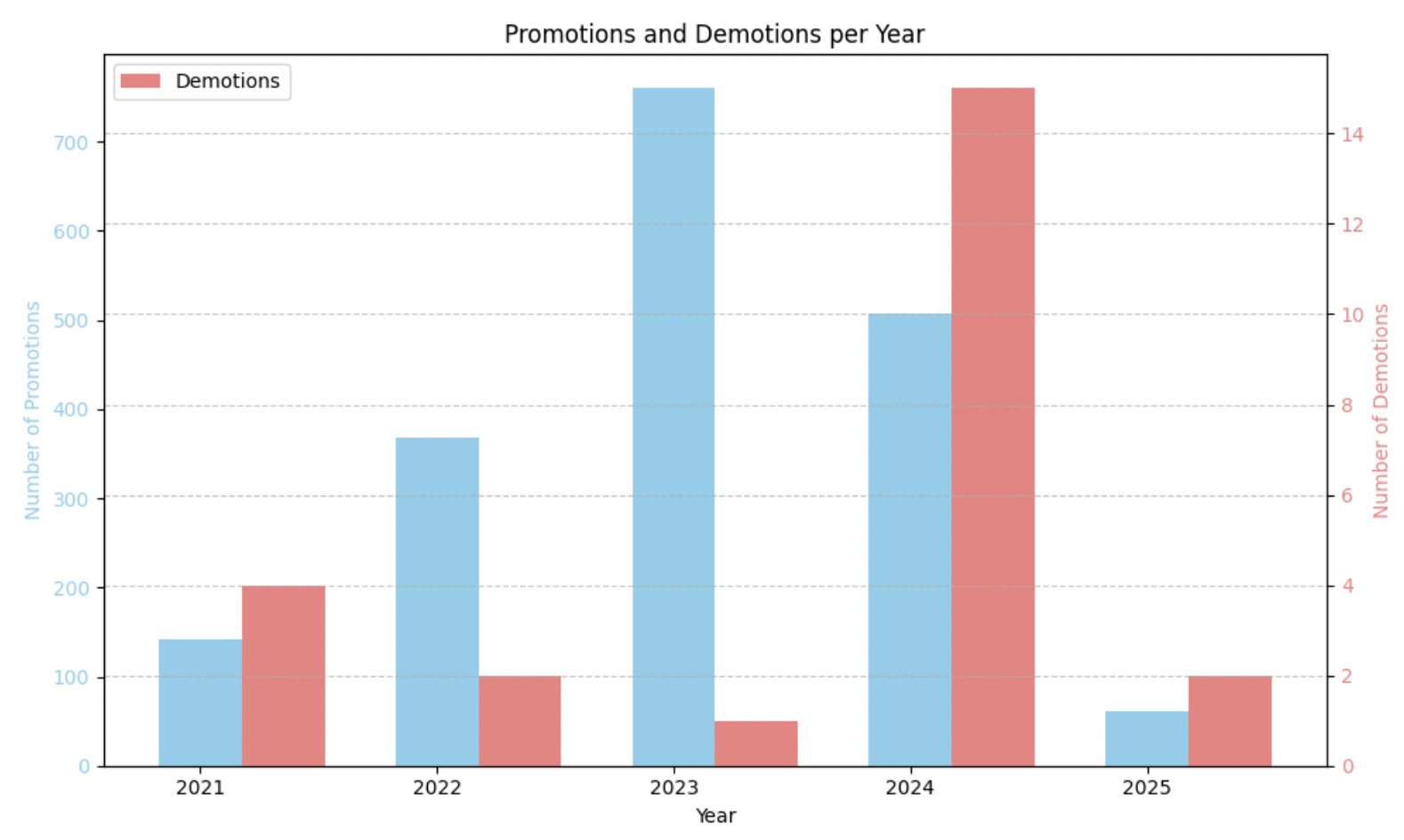


Fig. 3: Promotions and Demotions per year

The chart shows a steady increase in promotions from 2021 to 2023, peaking in 2023, followed by a decline in 2024 and 2025, while demotions remained low until 2024, where they sharply increased, possibly indicating organizational restructuring or performance-related adjustments

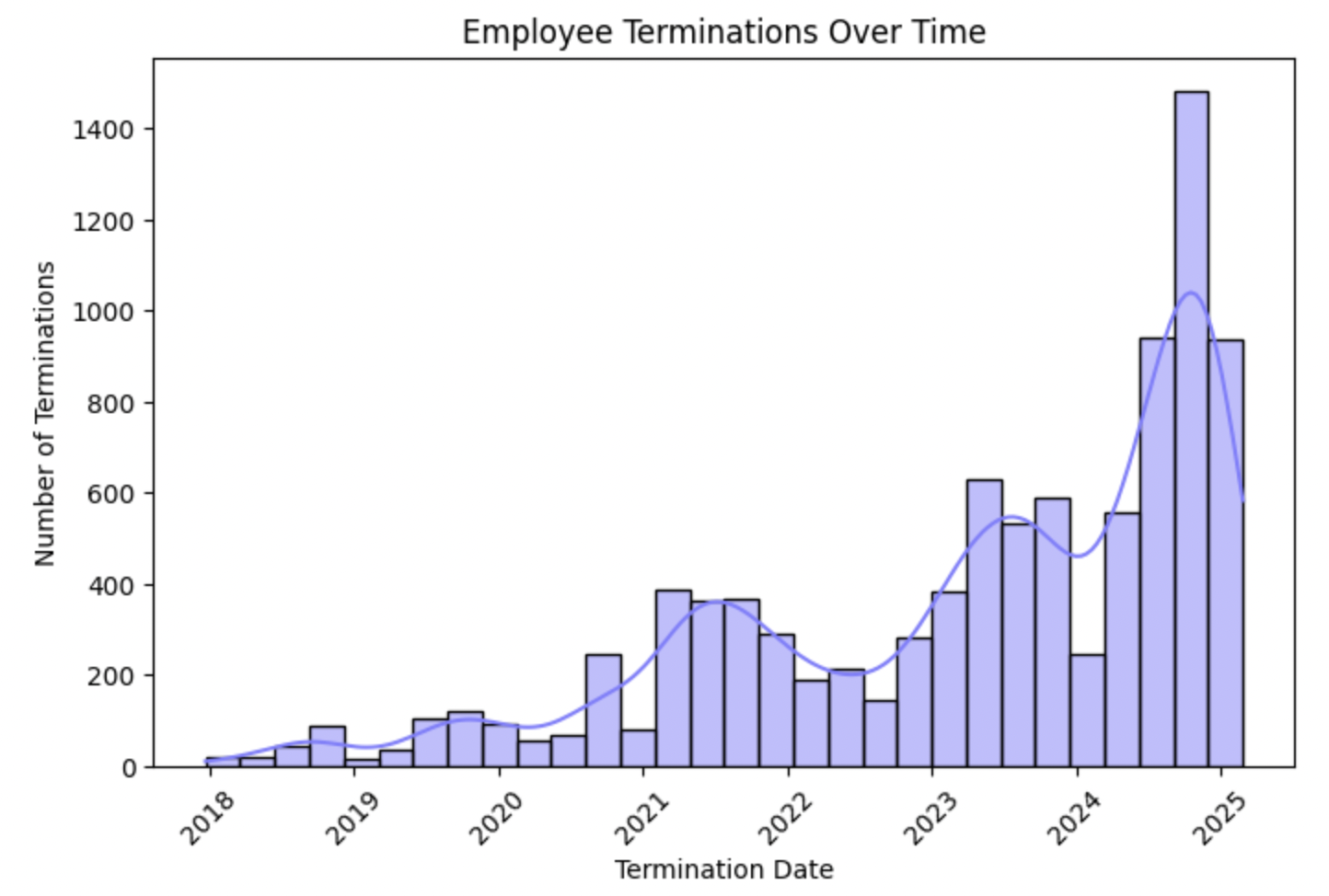
1. Employee Terminations Over Time

Fig. 4: Number of Termination over Time

The termination trend shows a gradual rise from 2018 to 2023, followed by a sharp spike in 2024, likely due to mass layoffs, restructuring, or economic factors. After peaking at nearly 8,000 terminations, numbers declined but remained higher than pre-2024 levels, suggesting a stabilization phase. Further analysis of departments, job roles, and termination reasons is needed to understand the impact and long-term workforce trends.

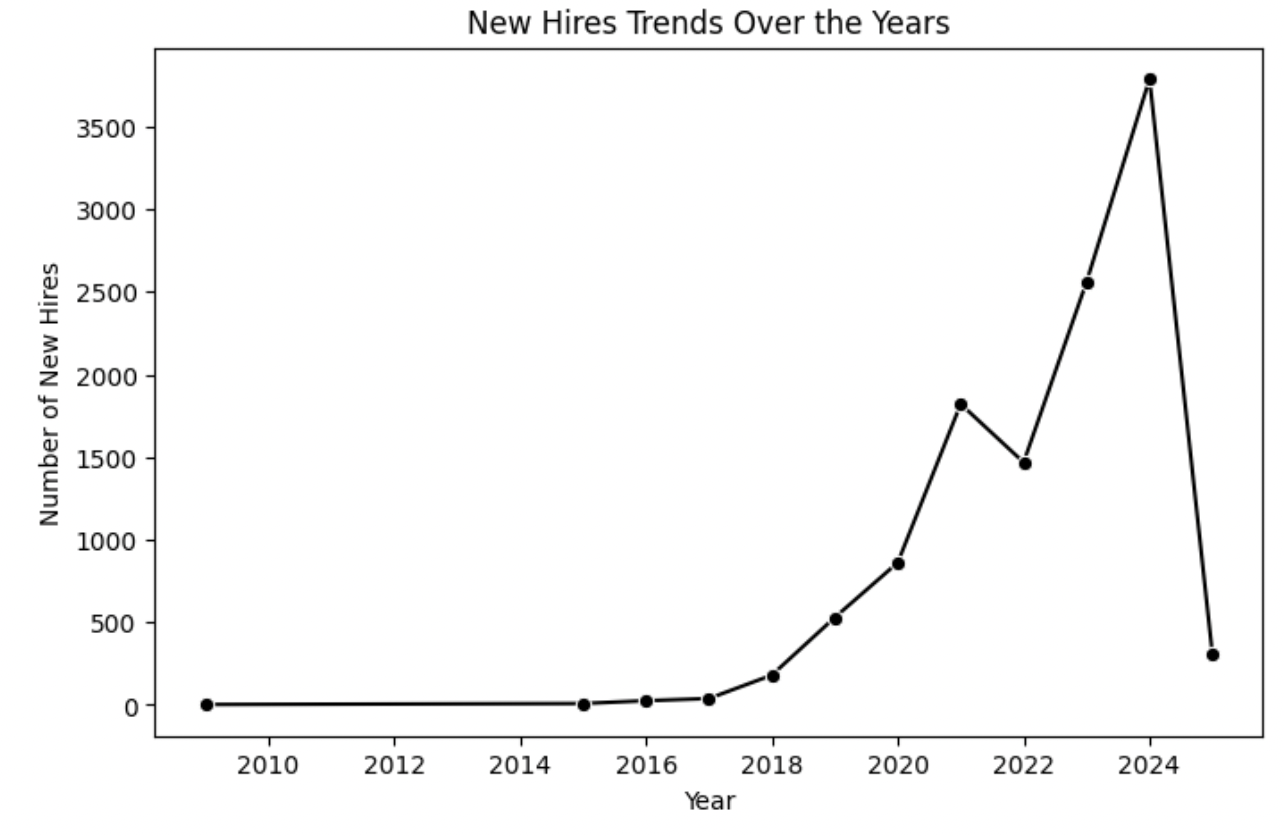
1. New hires trends over the year

Fig. 5: Line chart showing new hires trends over the years

The chart shows a steady rise in new hires from 2016, peaking in 2024, followed by a sharp decline, possibly indicating hiring freezes, layoffs, or shifting workforce strategies.

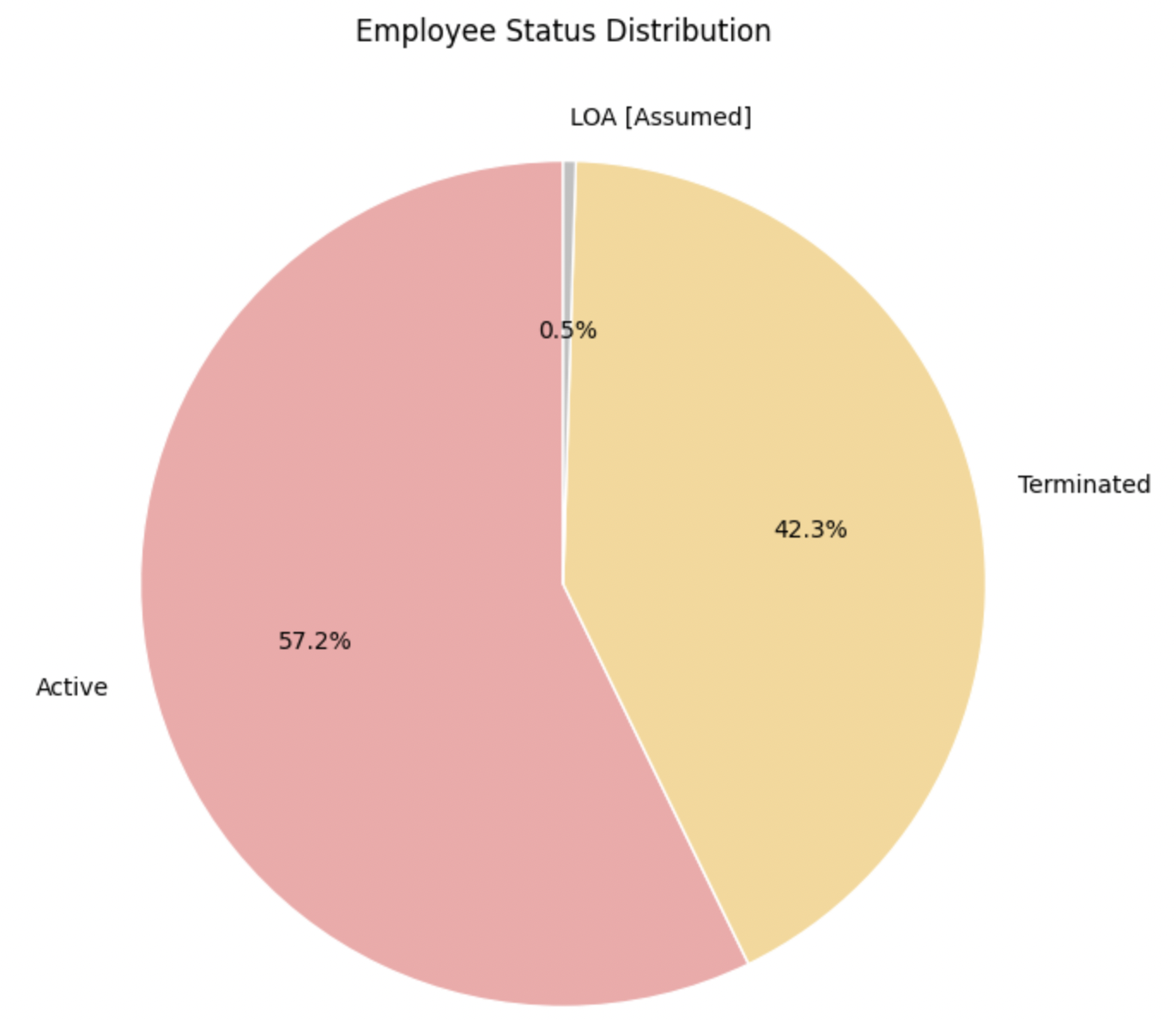
1. Employee Status Distribution 

Fig. 6: Employee Status

The Employee Status Distribution shows that 57.2% of employees are Active, 42.3% are Terminated, and 0.5% are on LOA, indicating a high termination rate that may impact workforce stability.

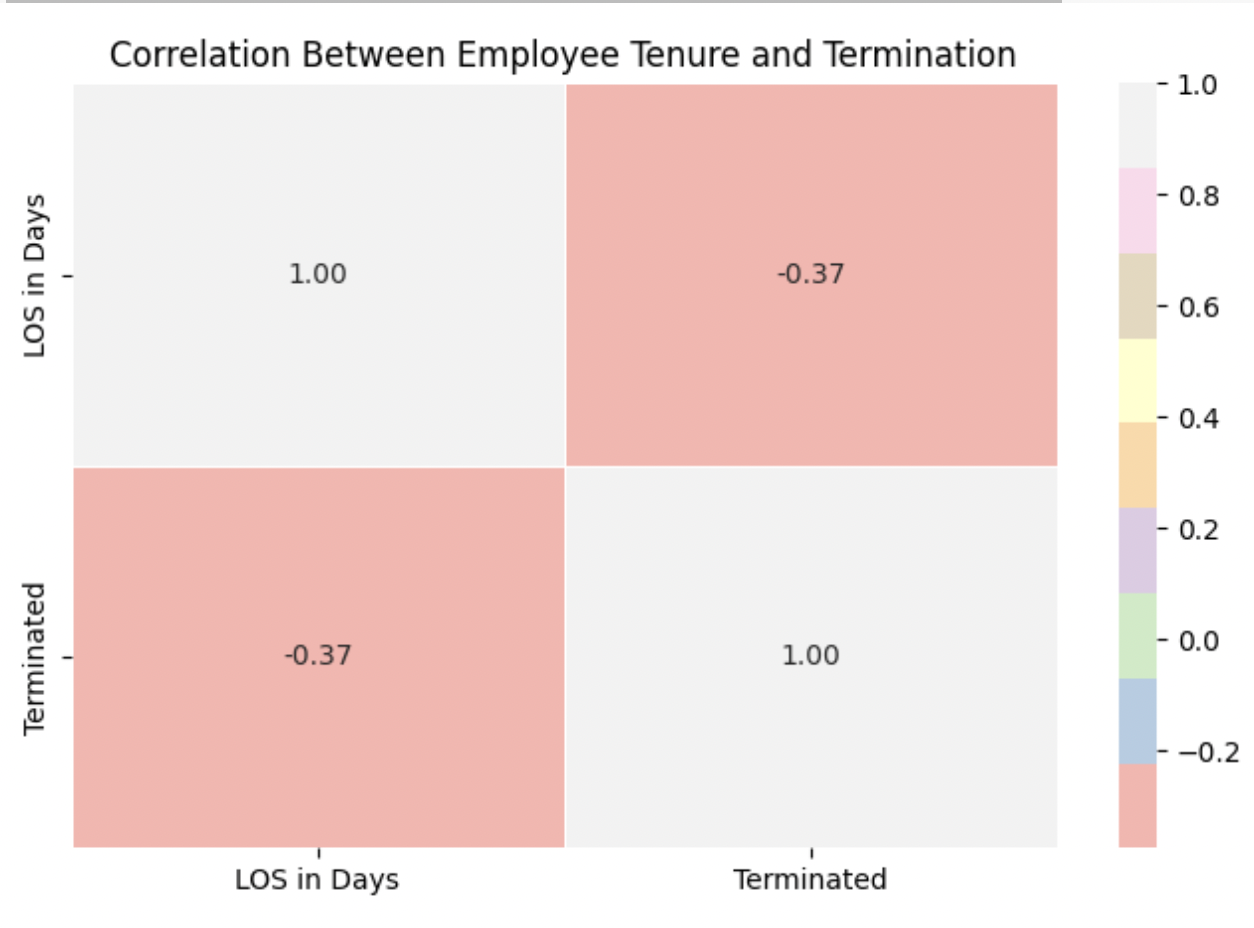
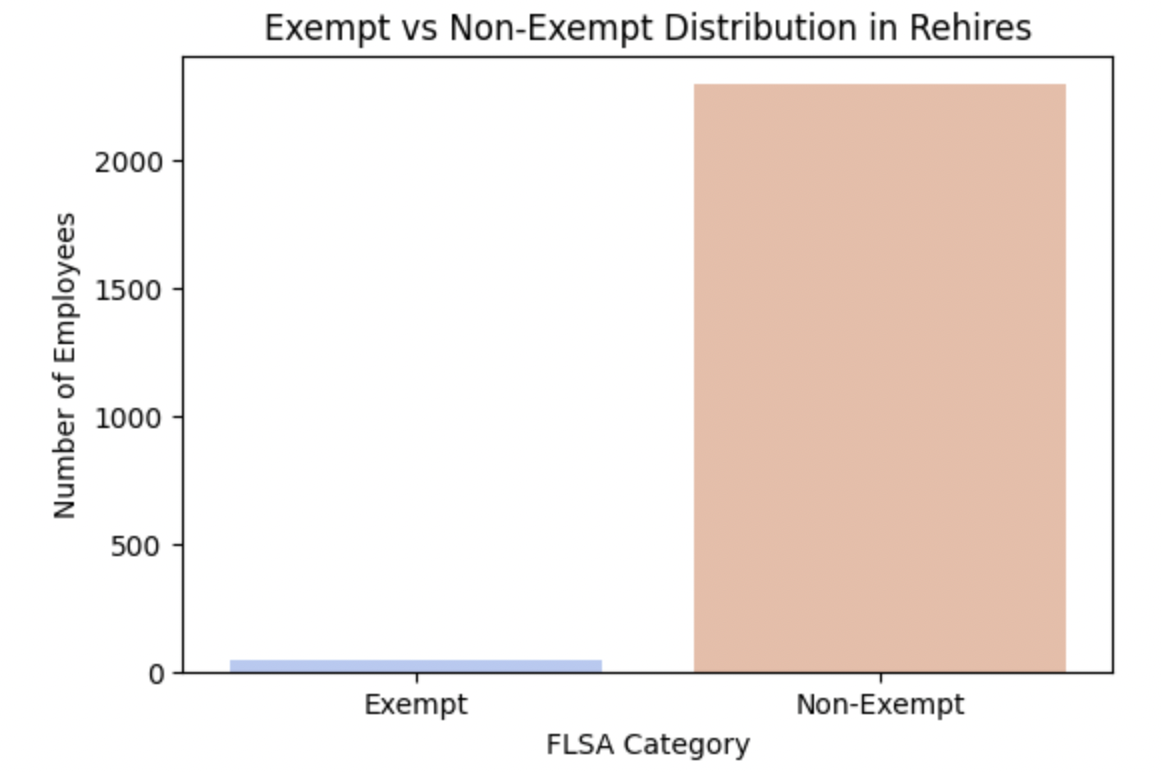
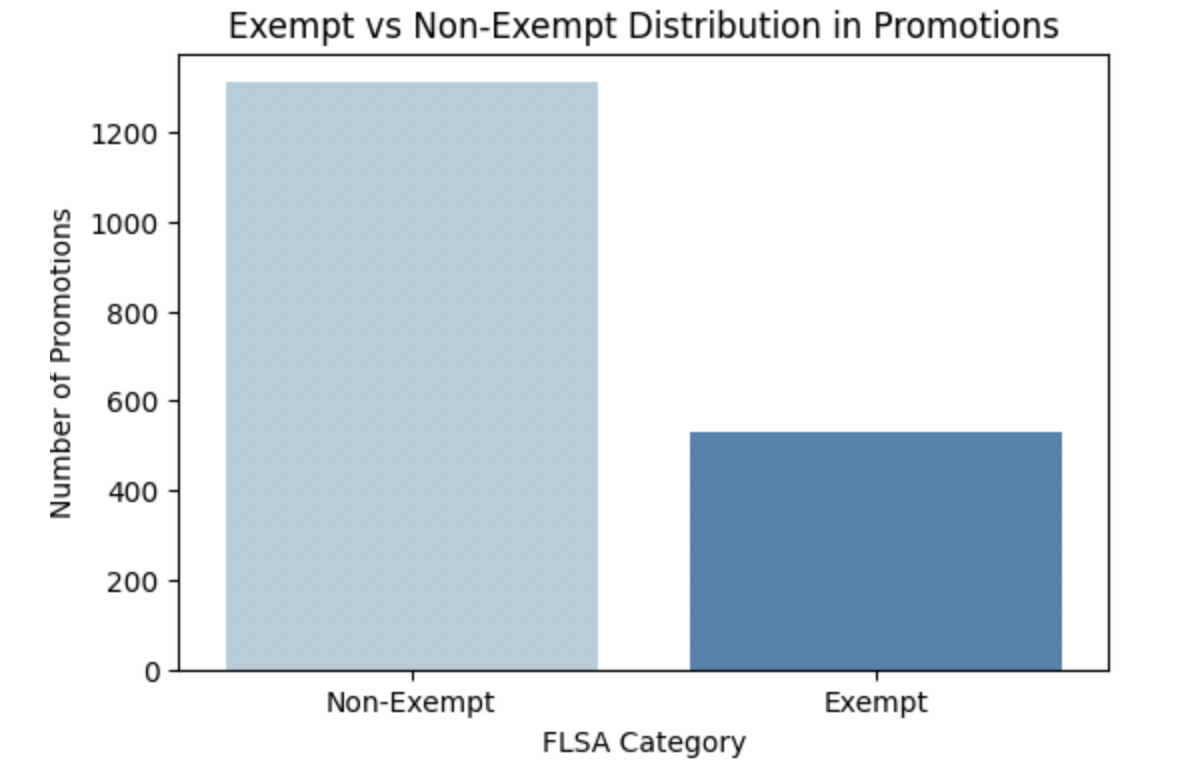
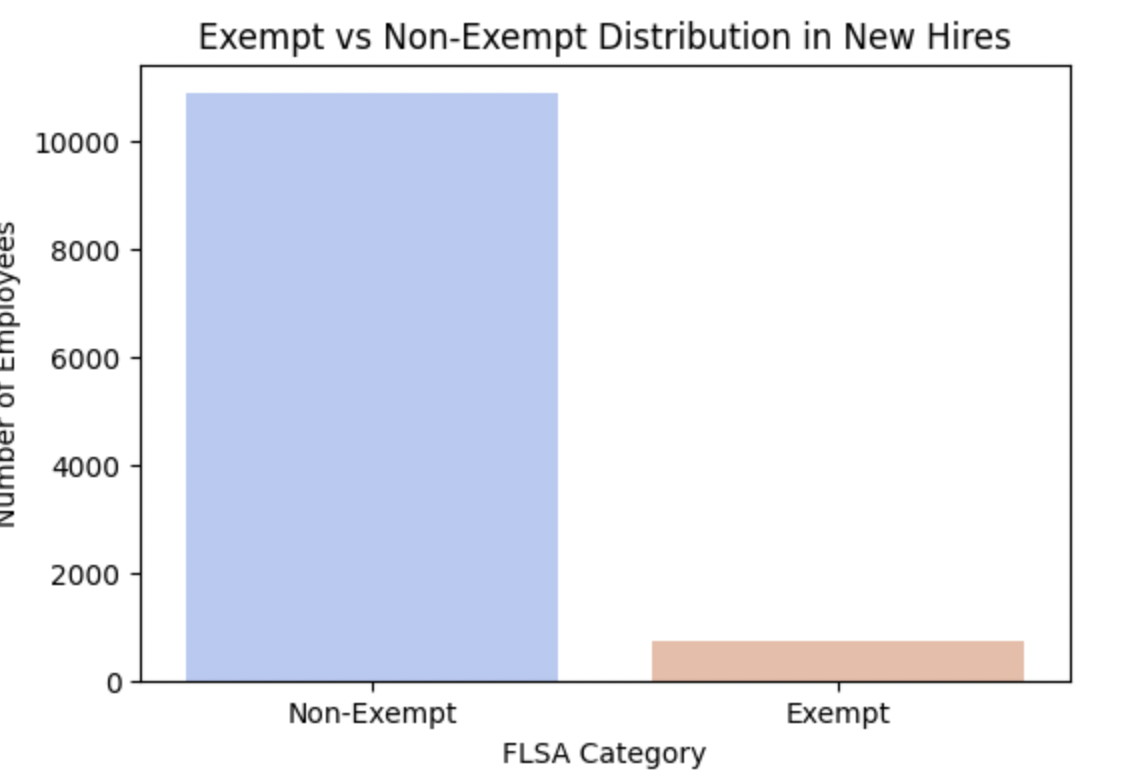
1. Correlation between Tenureship and Termination

Fig. 7: Correlation between employee tenureship and termination.

The negative correlation (-0.374) between Length of Service (LOS) and Termination Status indicates that employees with longer tenure are less likely to be terminated, suggesting that newer hires face a higher risk of termination, potentially due to performance issues, probationary periods, or job fit challenges.

1. Fair Labor Standards Act (FLSA) classification: Understanding the *Exempt* vs. *Non-Exempt* employee classification helps in:

* Workforce Planning: It helps determine how many employees qualify for overtime, affecting budget allocation.
* Compensation Analysis: Exempt employees are salaried, while non-exempt employees receive hourly wages, impacting payroll costs.
* Turnover Insights: Identifying whether more exempt or non-exempt employees leave can help refine retention strategies.
* Career Progression Trends: Understanding how FLSA status impacts promotions, demotions, and transfers helps assess job mobility.
* Compliance & Legal Risks: Ensures the company is adhering to labor laws related to overtime pay and fair classification.

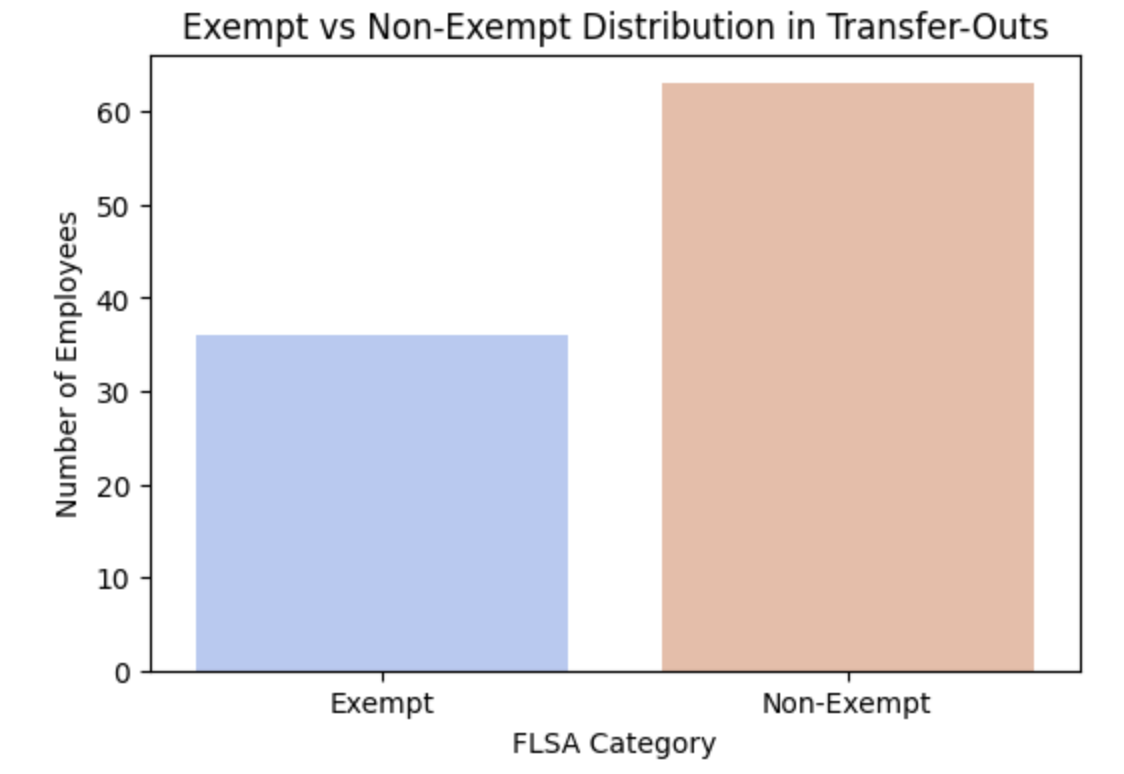
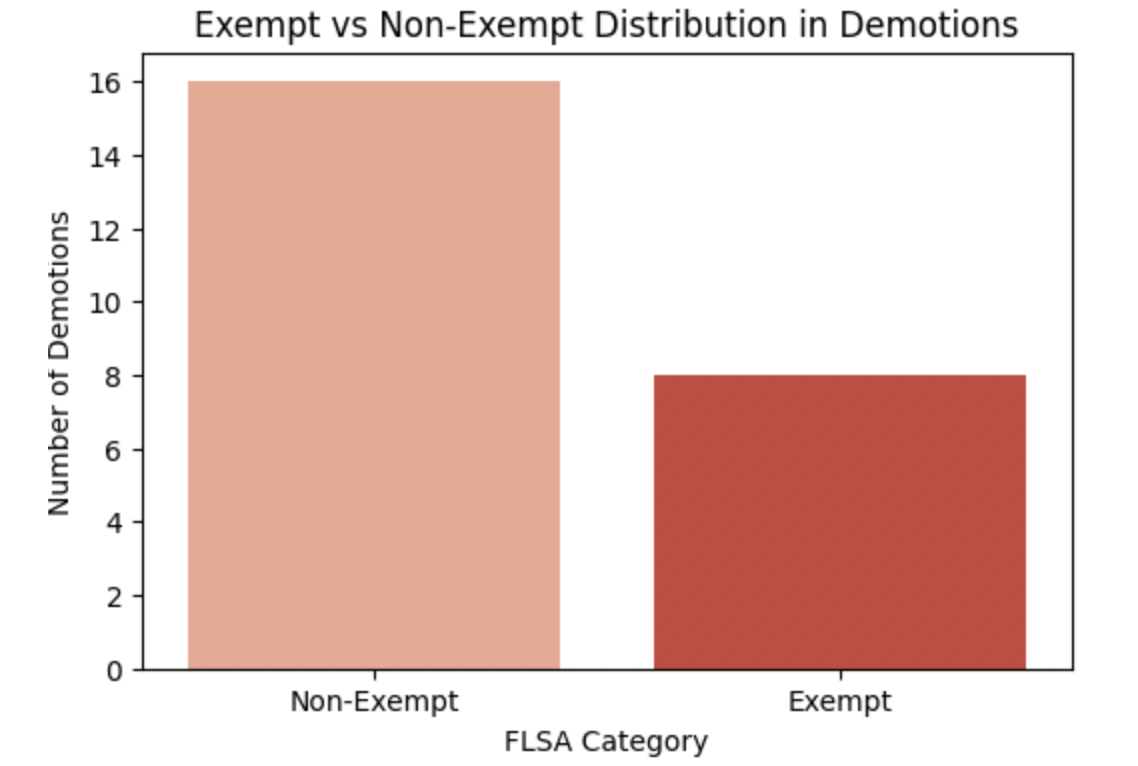
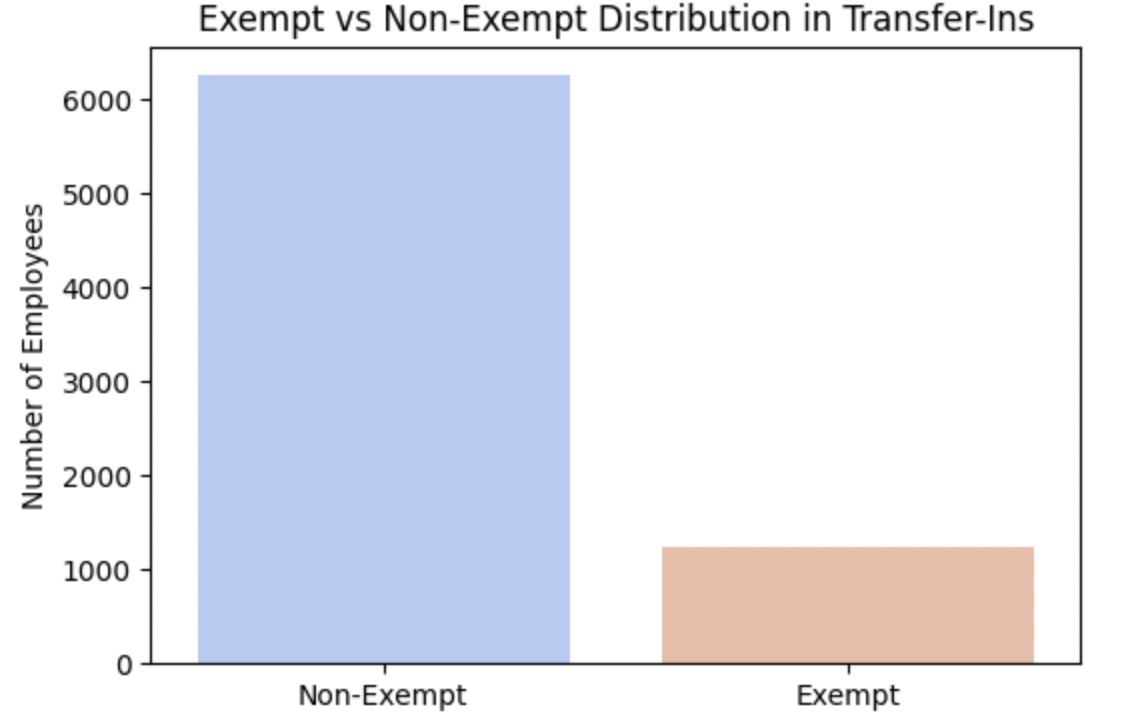


Fig. 8: Exempt Vs Non-Exempt Distribution

The analysis reveals that *Non-Exempt* employees dominate workforce movements, including *new hires, rehires, transfers,* and *demotions*, indicating higher turnover and job instability in these roles. While *promotions* show career progression opportunities for both *Exempt* and *Nonexemp*t employees, the overall trend suggests that *Non-Exempt* roles are more volatile, with frequent transitions and exits. The significant number of *rehired* *Non-Exempt* employees further highlights potential retention challenges. Meanwhile, *Exempt* employees experience fewer *transfers* and *demotions*, implying greater stability in salaried positions. A department-wise breakdown could provide deeper insights into specific areas experiencing the most workforce movement.

## **4.5 Data Issues**

1. Checking Missing Value

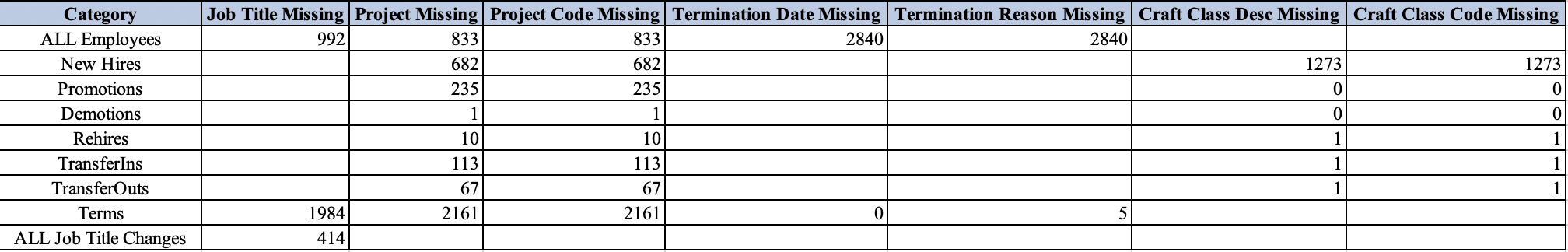
Upon checking for missing values on the provided dataset, it is observed that there are many records with missing values as shown below: 

Table 4: Missing Records

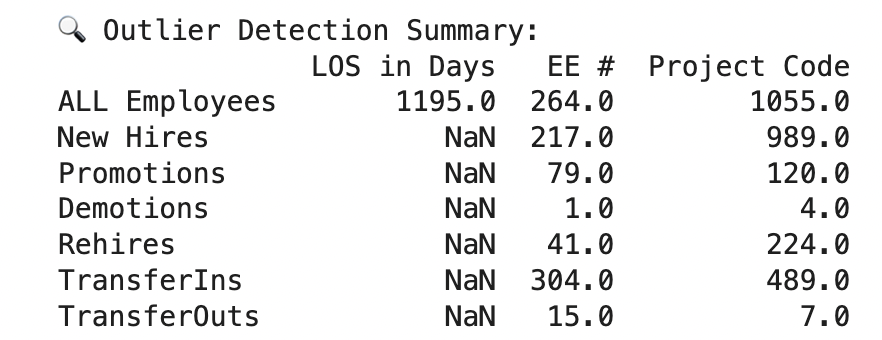
1. Outlier Detection: Relevant numeric columns for outlier detection include *Length of Service (LOS in Days)* to identify unusually long or short tenures, *Employee Number* (*EE #*) to detect anomalies in employee assignments, and *Project Code* to uncover irregular project allocations or transitions.

Fig 9: Outliers for the numerical column

From the output as shown above, the *Length of Service (LOS in Days)* outliers were detected only in the *ALL Employees* category indicating cases of extremely long or short tenures, which could be due to contract-based workers, rehired employees, or potential data inconsistencies. No significant LOS outliers were found in New Hires, Promotions, Demotions, Rehires, TransferIns, or TransferOuts.

For *Employee Number (EE #)*, outliers were present across all categories, with New Hires, Promotions, Demotion, Rehires, TransferIns, and TransferOuts. These anomalies might correspond to employees returning after a break, irregular hiring sequences, or leadership hires, with TransferIns reflecting frequent internal movements.

Regarding Project Code, major outliers were found in ALL Employees , New Hires, and TransferIns, while other categories had fewer irregularities. These outliers might indicate unusual project assignments, reassigned employees, or specific high-priority projects, and the 989 outliers in New Hires could suggest disproportionate hiring in certain projects or potential misclassification of project codes.

# **5. Data Challenges and Considerations**

During the analysis, several data-related challenges were identified that could impact the accuracy and reliability of insights:

Missing Data: Some records lack complete training history or exit reasons, making it difficult to determine the relationship between training participation and employee retention. This missing information could lead to gaps in understanding workforce movement and attrition trends.

Data Imbalance: Certain departments exhibit disproportionately high or low attrition rates, which may skew overall workforce analysis. This imbalance suggests the need for department-specific retention strategies rather than generalized company-wide policies.

Inconsistent Entries: Variations in job title formatting and training status labels were observed, leading to data inconsistencies that can affect classification and trend analysis. Standardizing these entries is essential to ensure accurate comparisons and predictive modeling.

# **6. Future Enhancements & Data Refinement**

As the dataset evolves with updated details and refined records, further enhancements will be implemented to ensure the accuracy and reliability of the data report. Further sanity checks will be conducted to ensure data accuracy and integrity. Missing details (e.g., training history, termination reasons) will be addressed for a more complete workforce analysis. Outlier detection models will be fine-tuned, and predictive analytics will be enhanced to anticipate attrition risks and promotion trends. Additionally, data privacy measures will be strengthened to ensure compliance with evolving regulations. These refinements will enable more accurate, actionable, and strategic workforce insights.

## 

# **7. Appendix**

## **Generative AI Usage**

The following prompts and queries were employed:

1. Summarize the operations and workforce expansion goals of Primoris Renewable Energy, emphasizing its role in the renewable energy sector.
2. Define the workforce challenges faced by a fast-growing renewable energy company and frame a problem statement.
3. Explain why Jira is a suitable project management tool for workforce planning and predictive analytics initiatives.
4. Suggest measurable KPIs for tracking workforce retention, engagement, training impact, and hiring efficiency.
5. What could be the important parameters to look after while preparing a preliminary data report?
6. Can you help me understand the code each line step by step to infer the graphs and its concept behind it.
7. What is the advantage of using Jira and colab notebook in terms of working collaboratively?
8. What are the necessary steps and format to begin with build the preliminary data report

### 

# 