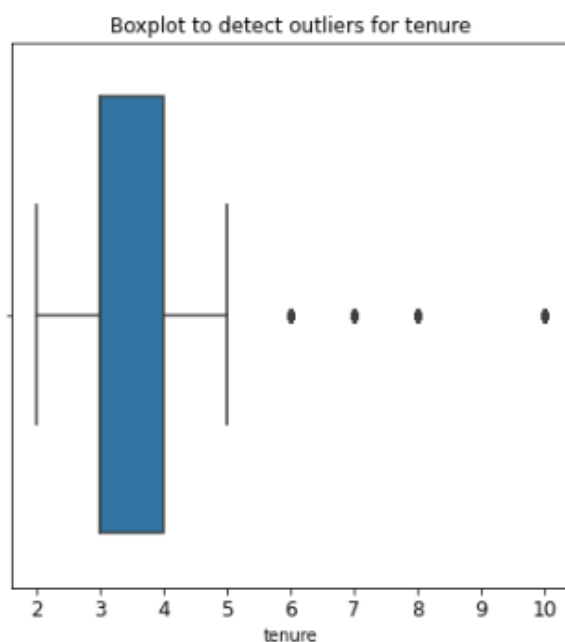


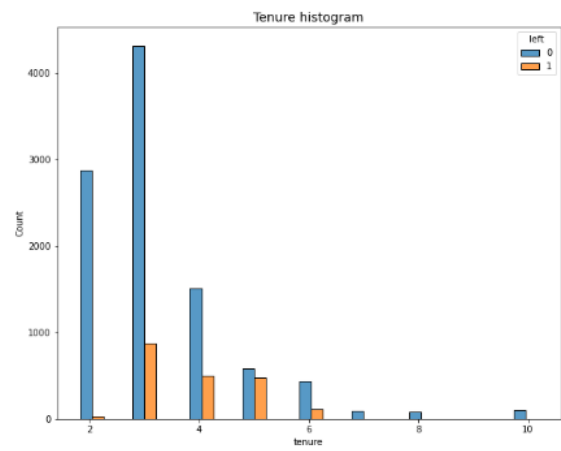
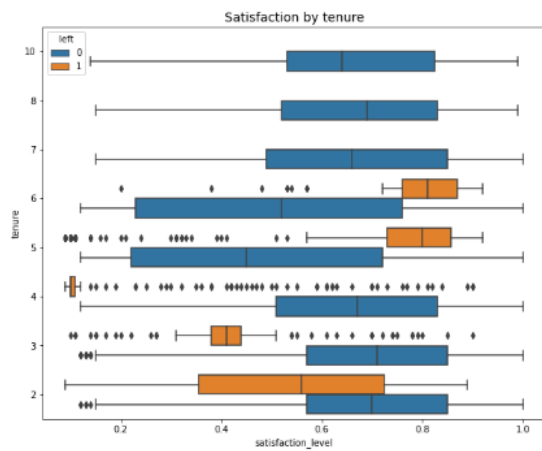
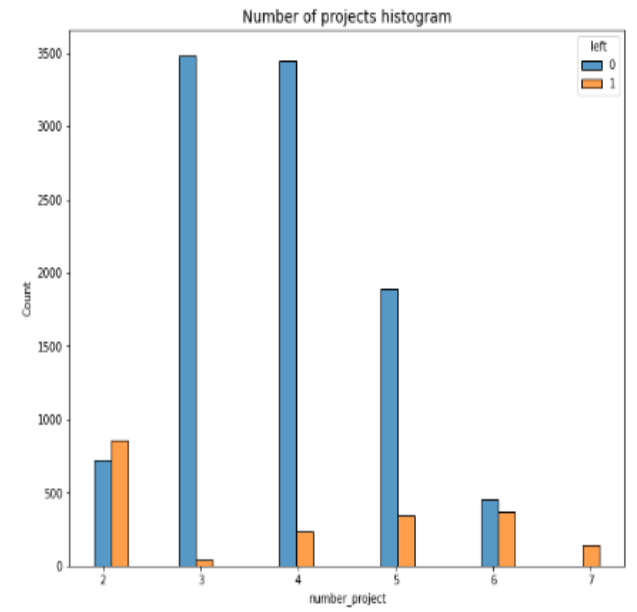
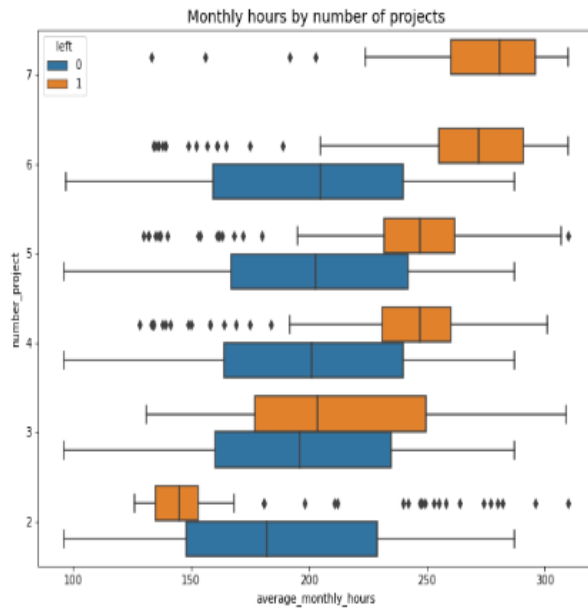
The visualizations from the Python lab in Jupyter Notebook provide a comprehensive overview of the dataset, highlighting key insights into employee turnover at Salifort Motors.

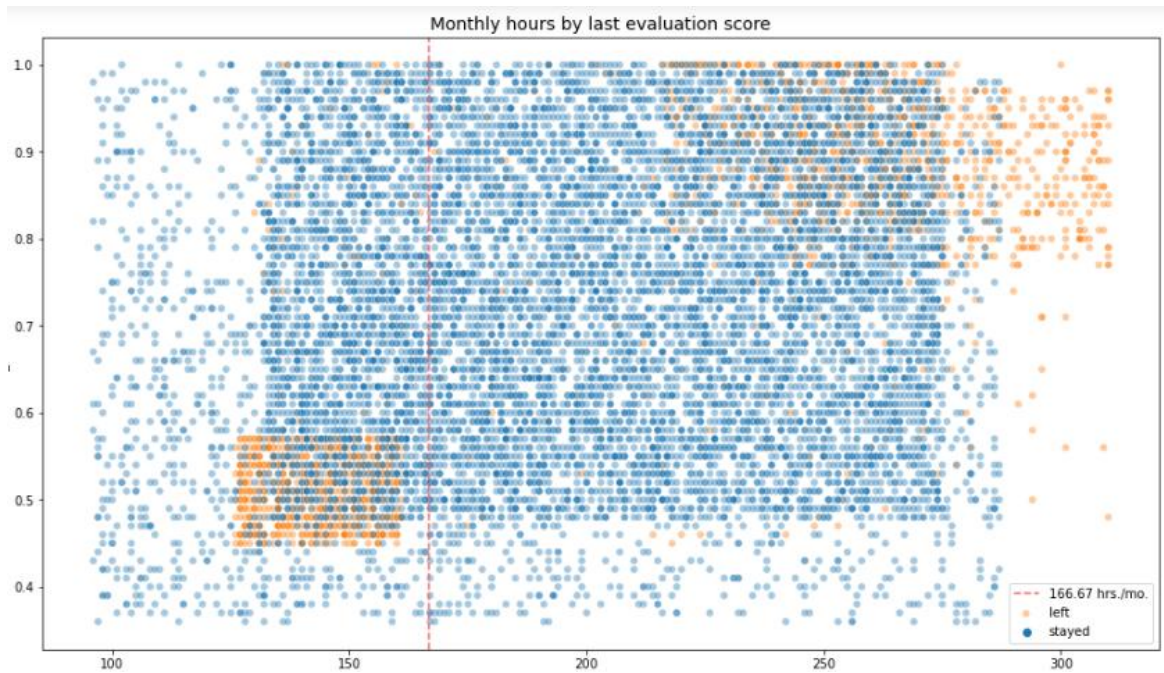
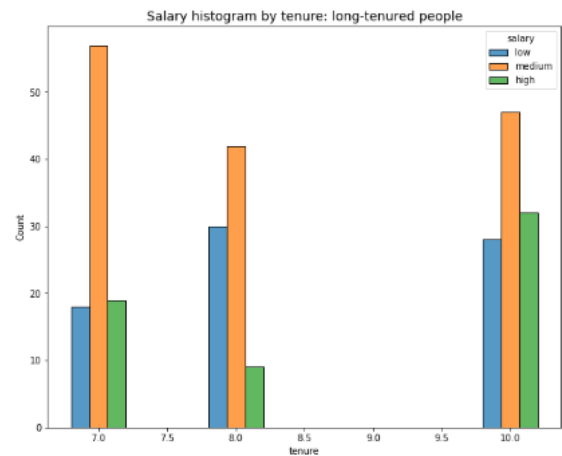
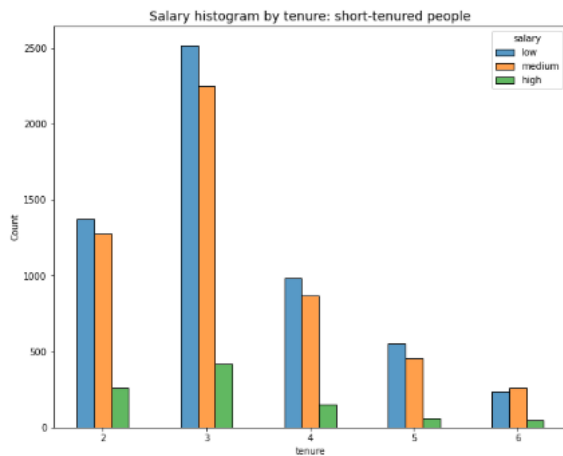
1. **Correlation Heatmap:** Displays relationships between numerical features, identifying satisfaction level and average monthly hours as significant predictors of turnover.
1. **Department Analysis:** Bar plots showcasing turnover rates across departments reveal that specific teams, such as "Technical" and "Support," experience higher turnover.
2. **Satisfaction Levels:** Histograms illustrate the distribution of satisfaction scores, revealing that employees with lower satisfaction are more likely to leave.
3. **Workload Distribution:** Scatter plots of average monthly hours and number of projects indicate that employees with excessive workloads exhibit higher turnover rates.
4. **Salary Insights:** Stacked bar charts show turnover trends by salary tiers, suggesting that employees with "low" salaries are disproportionately affected.
5. **Tenure Analysis:** Line graphs charting time spent at the company highlight tenure thresholds where dissatisfaction peaks.

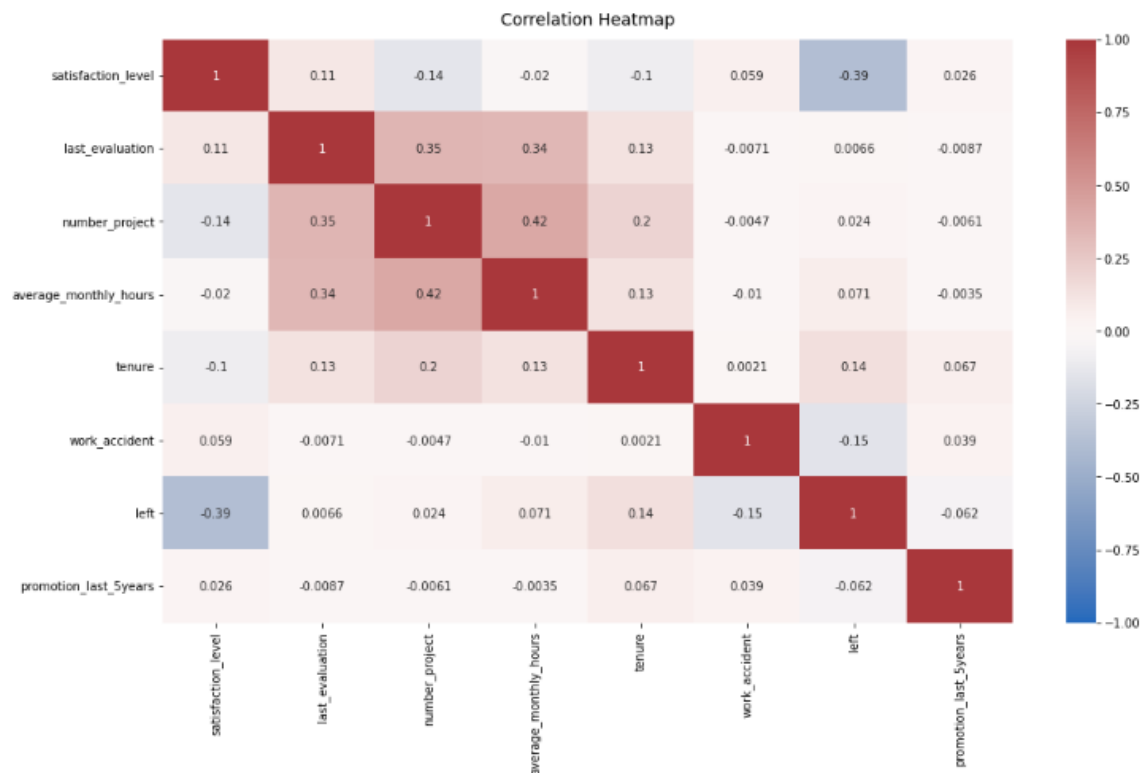
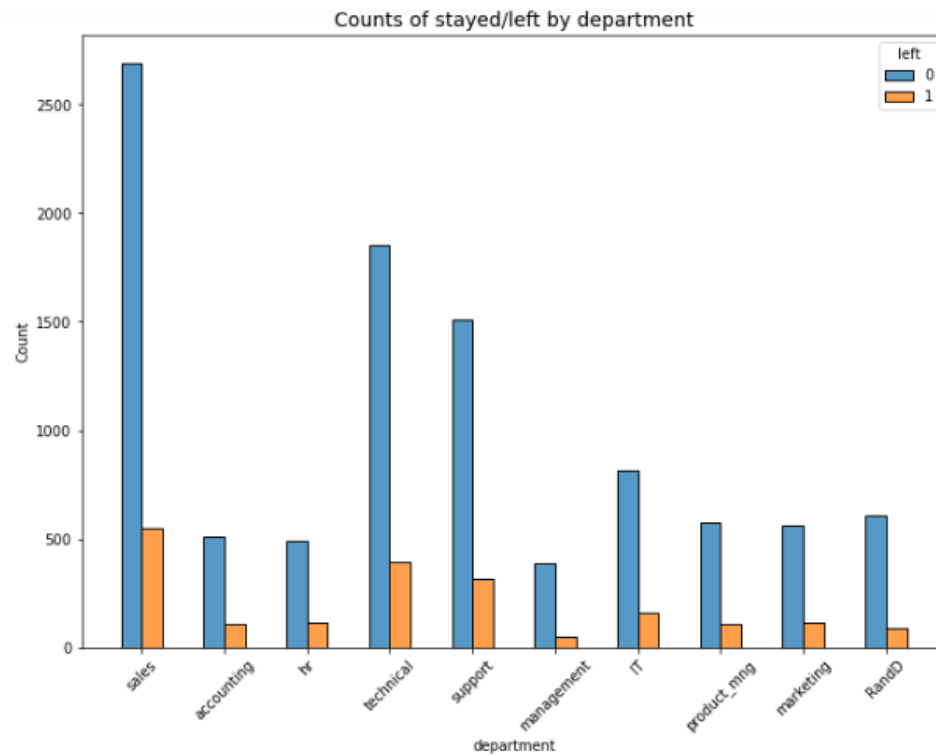


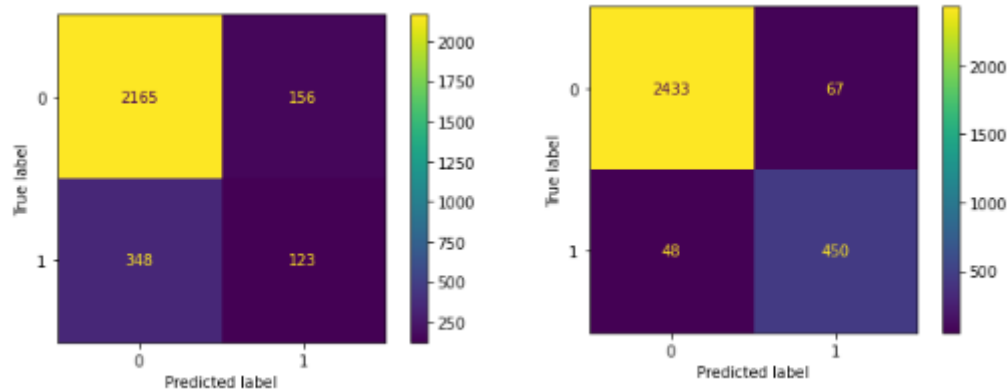
The boxplot above shows that there are outliers in the `tenure` variable.

It would be helpful to investigate how many rows in the data contain outliers in the `tenure` column.









Decision tree splits

```
In [78]: # Plot the tree
plt.figure(figsize=(85,20))
plot_tree(tree2.best_estimator_, max_depth=6, fontsize=14, feature_names=X.columns,
          class_names={0:'stayed', 1:'left'}, filled=True);
plt.show()
```

