

Penske Project

Project 1 - Warehouse Associate Productivity

Henry Yang, Alua Bekbossynova,
Jeffy Jong, Kevin Chen, Nick Moshos



Table of contents



01

Introduction

02

Analysis

03

Recommendation



01

Introduction





Data

Purpose: Monitors training progress and regulatory compliance

Takeaway: Introduces a detailed employee training overview

Highlight: Training strategy that guides employees through essential orientations, safety protocols, and job-specific skills, ensuring thorough preparedness for their roles within the company

Our goal

- Assess training duration (60-90 days) for optimal productivity and lower turnover.
- Evaluate training timeline variations across locations for standardization or customization.
- Analyze the link between training length and turnover to minimize attrition.
- Identify training delays by location to enhance performance and efficiency.



Data Summary

Start_date: The Date when an employee started.

Employee_type: The Type of employment, such as full-time or part-time.

Job_title: The Title of the job held by the employee.

Hr_orient: Indicates if human resources orientation was completed.

Safety_orient: Indicates if safety orientation was completed.

Pit_training_start: Start date of PIT (Powered Industrial Trucks) training.

Pit_training_end: End date of PIT training.

Days_pit: Number of days spent in PIT training.

Dept_training_start: Start date of department-specific training.

Dept_training_end: End date of department-specific training.

Days_dept: Number of days spent in department training.

Days_total: Total number of training days.

Term_before_cert: Indicates if the employee terminated before certification.

Location: Location of employment.

Data Cleaning

- Missing values Assumption: term_before_cert ->"Y", days_pit and days_dept -> 0
- Corrected days_total: days_pit + days_dept
- Data segmentation: Created subsets for employees who left before certification and those who stayed
- Added numeric conversion for analysis: Transformed term_before_cert into a numeric column for easier analysis.



02

Analysis



Analysis 1: Training time



Method



Step 1

Used linear regression analysis to find the correlation between total training days and Location/Job_Title



Step 2

Used logistic regression analysis to find the correlation between termination and training time



Step 3

Used Tableau to visualize the relationship between total training days and Location/Job_Title



Step 4

Finally, we used prediction analysis to predict the training time for each location

Finding

Call:

```
lm(formula = days_total ~ location, data = data)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-44.700	-10.570	0.071	2.923	85.300

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	18.111	8.050	2.250	0.0269	*
locationLOC_2	-8.034	9.340	-0.860	0.3919	
locationLOC_3	-4.049	10.063	-0.402	0.6884	
locationLOC_4	21.472	10.650	2.016	0.0467	*
locationLOC_5	4.817	10.318	0.467	0.6417	
locationLOC_6	27.589	11.097	2.486	0.0147	*
locationLOC_7	-5.611	10.650	-0.527	0.5995	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 24.15 on 92 degrees of freedom

Multiple R-squared: 0.2197, Adjusted R-squared: 0.1688

F-statistic: 4.316 on 6 and 92 DF, p-value: 0.000702

Call:
glm(formula = term_before_cert_numeric ~ days_total, family = binomial(),
data = data)

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	0.66927	0.31812	2.104	0.035396	*
days_total	-0.06353	0.01916	-3.316	0.000915	***

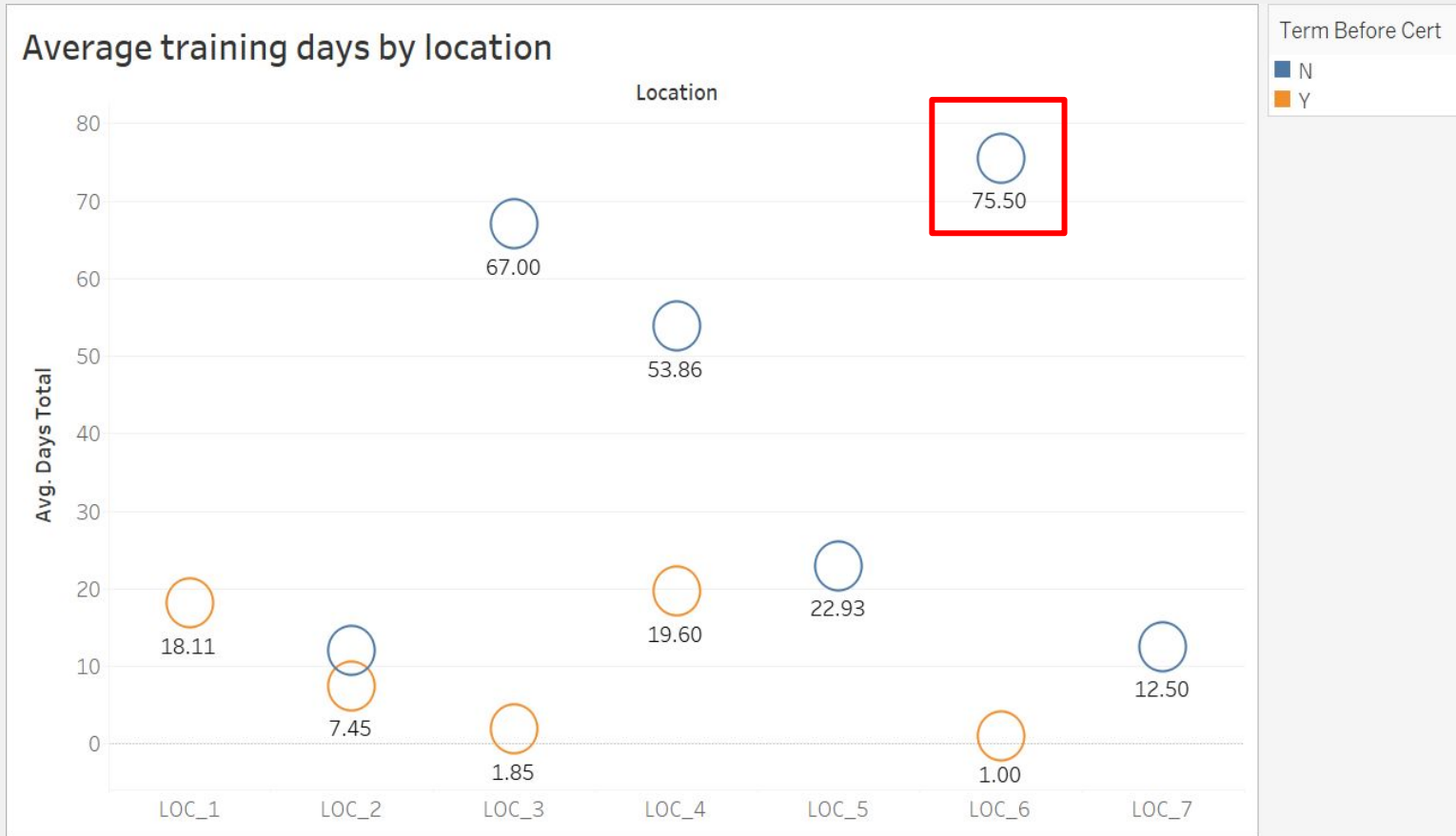
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

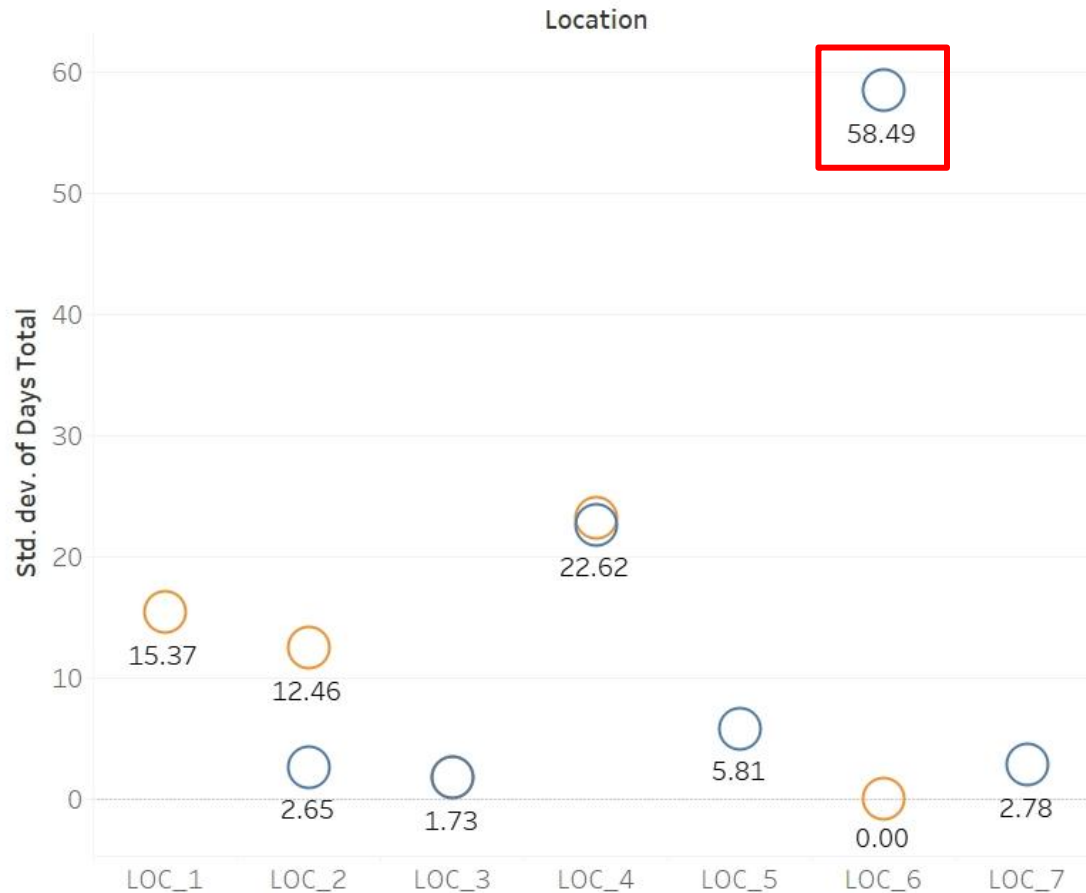
Null deviance: 134.96 on 98 degrees of freedom
Residual deviance: 113.16 on 97 degrees of freedom
AIC: 117.16

Number of Fisher Scoring iterations: 5

Location_6 with the highest training days

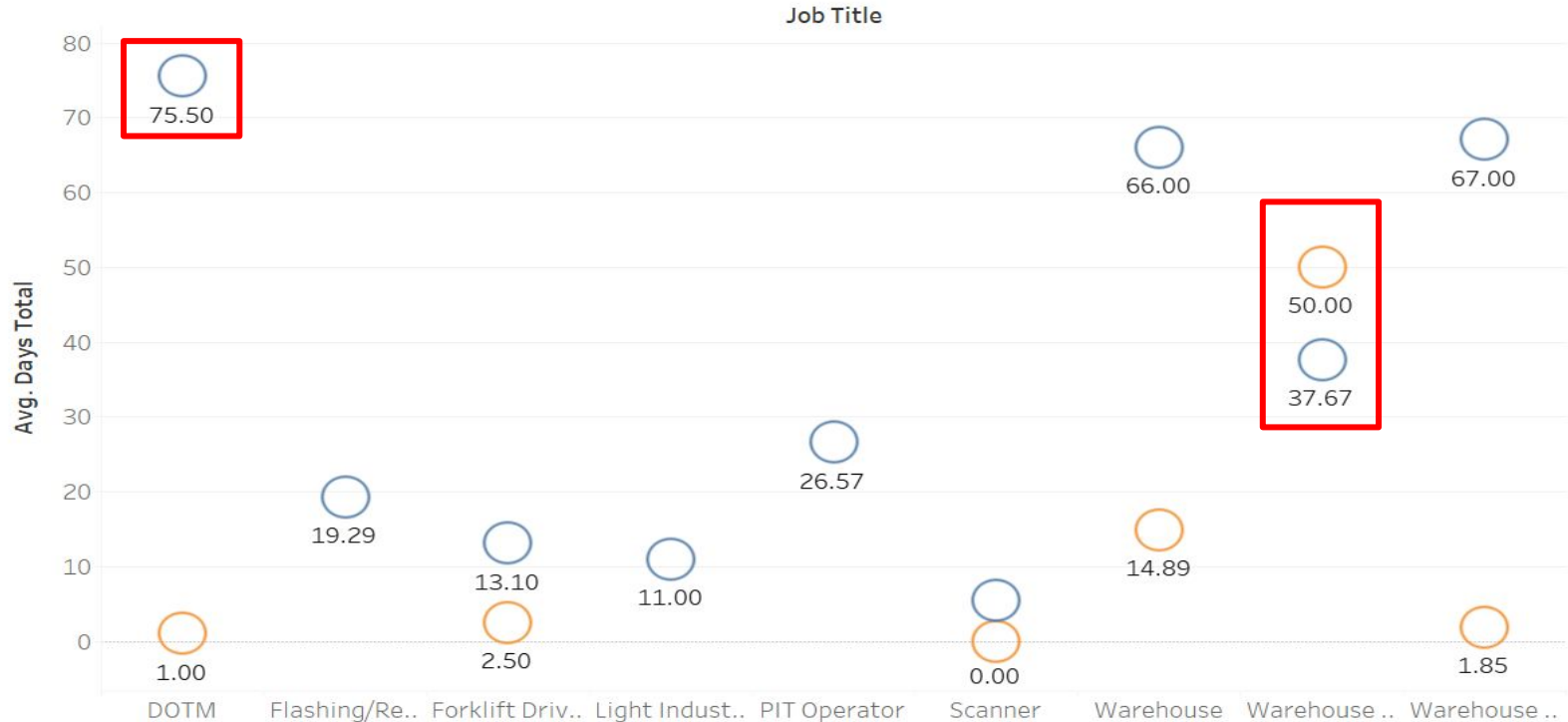


Standard Deviation of training days by location

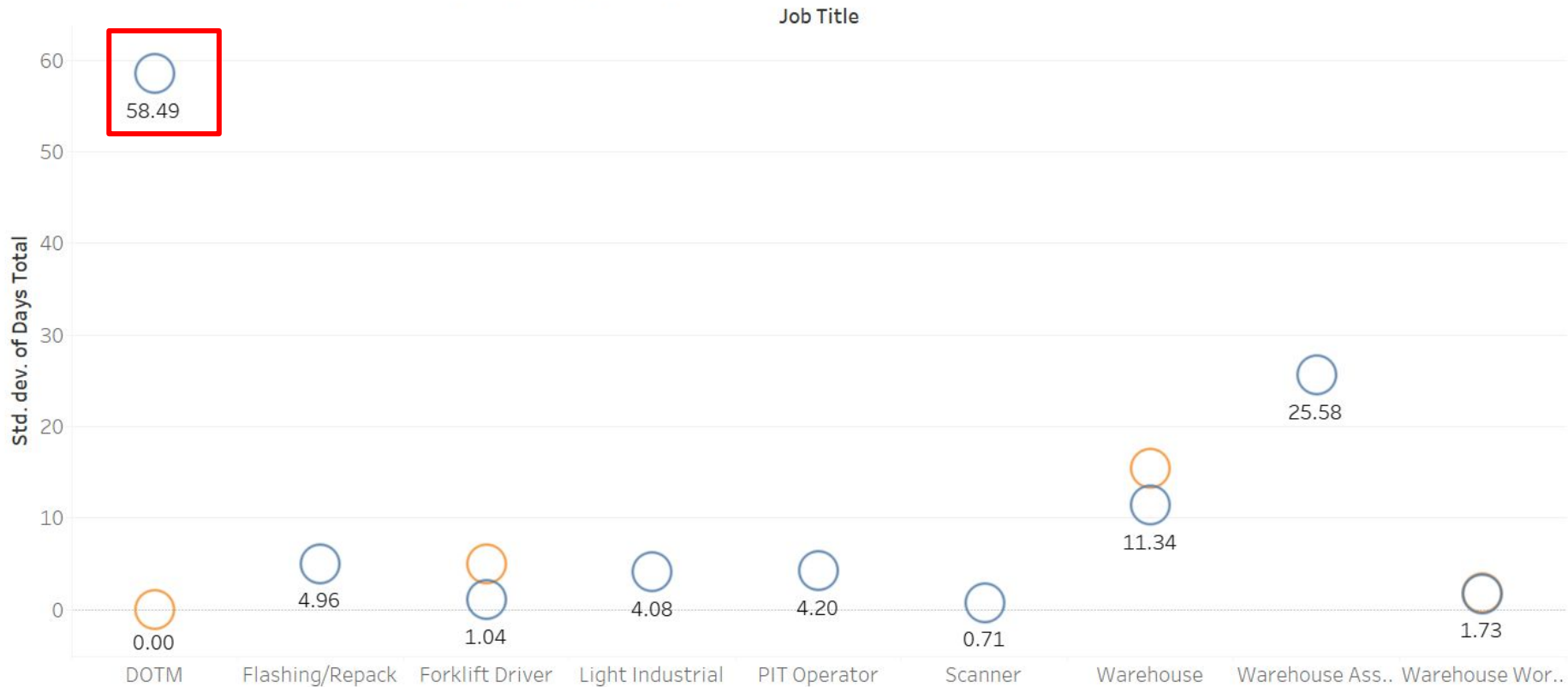


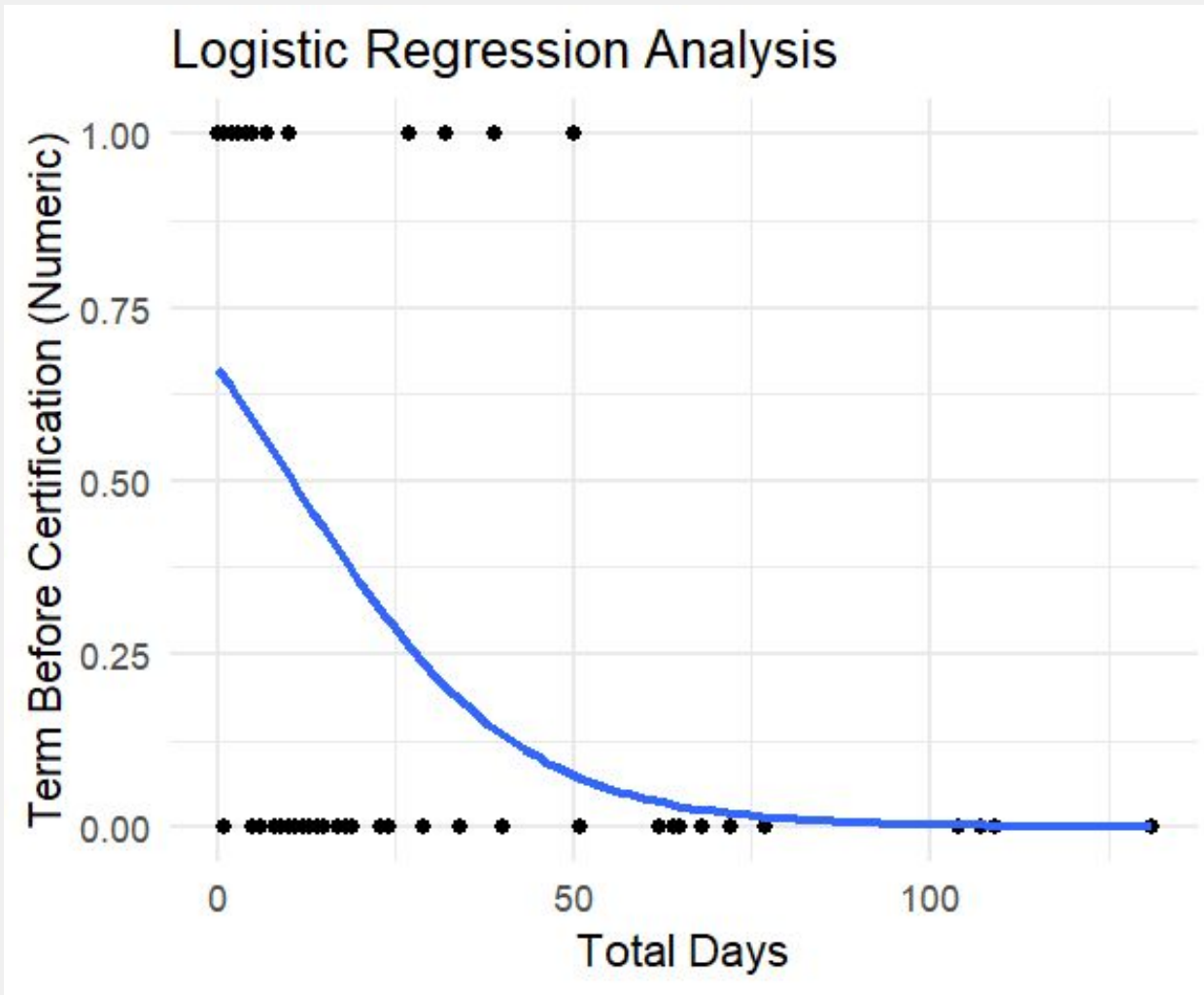
DOTM and Warehouse inconsistency

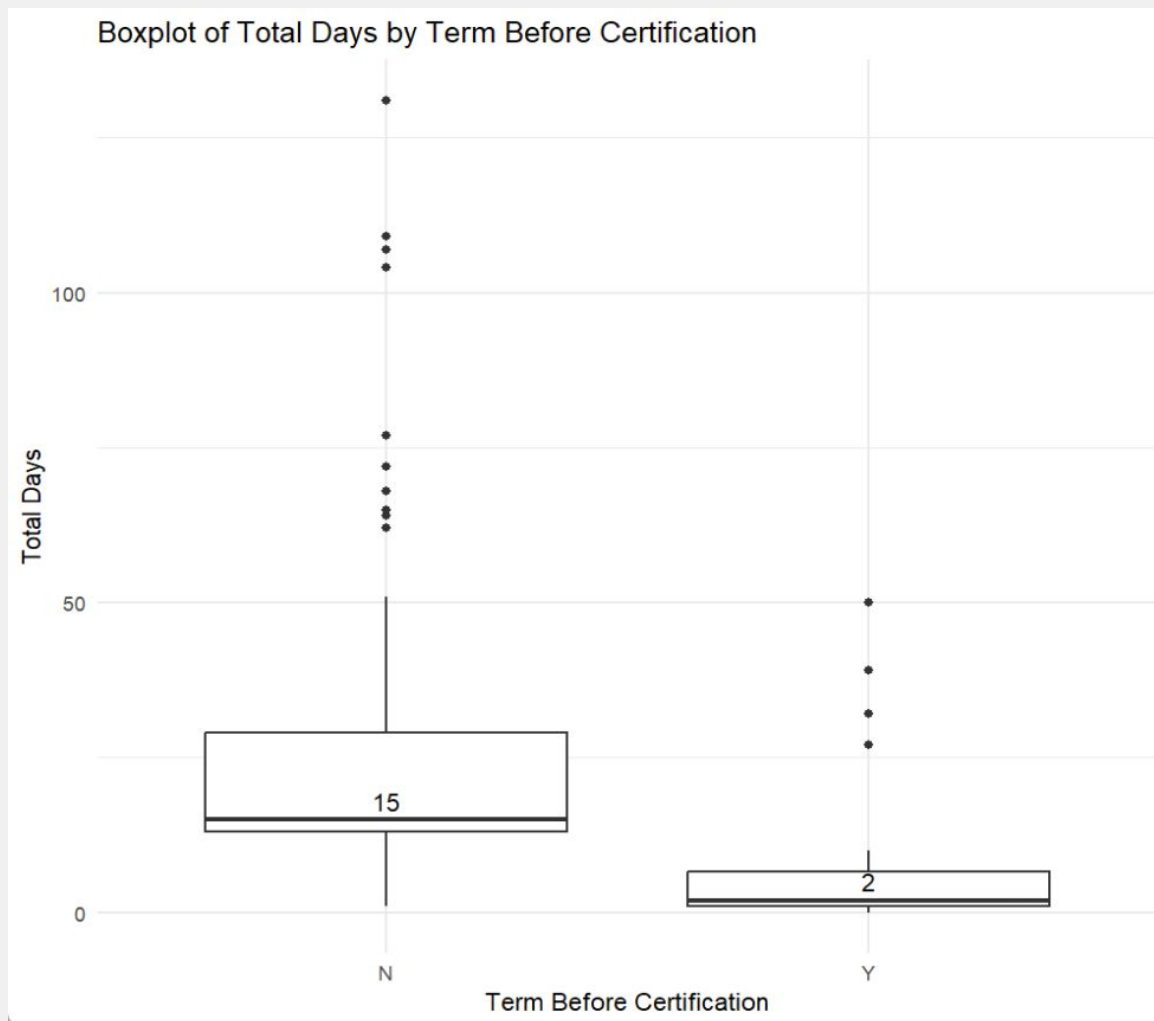
Average training days by job type



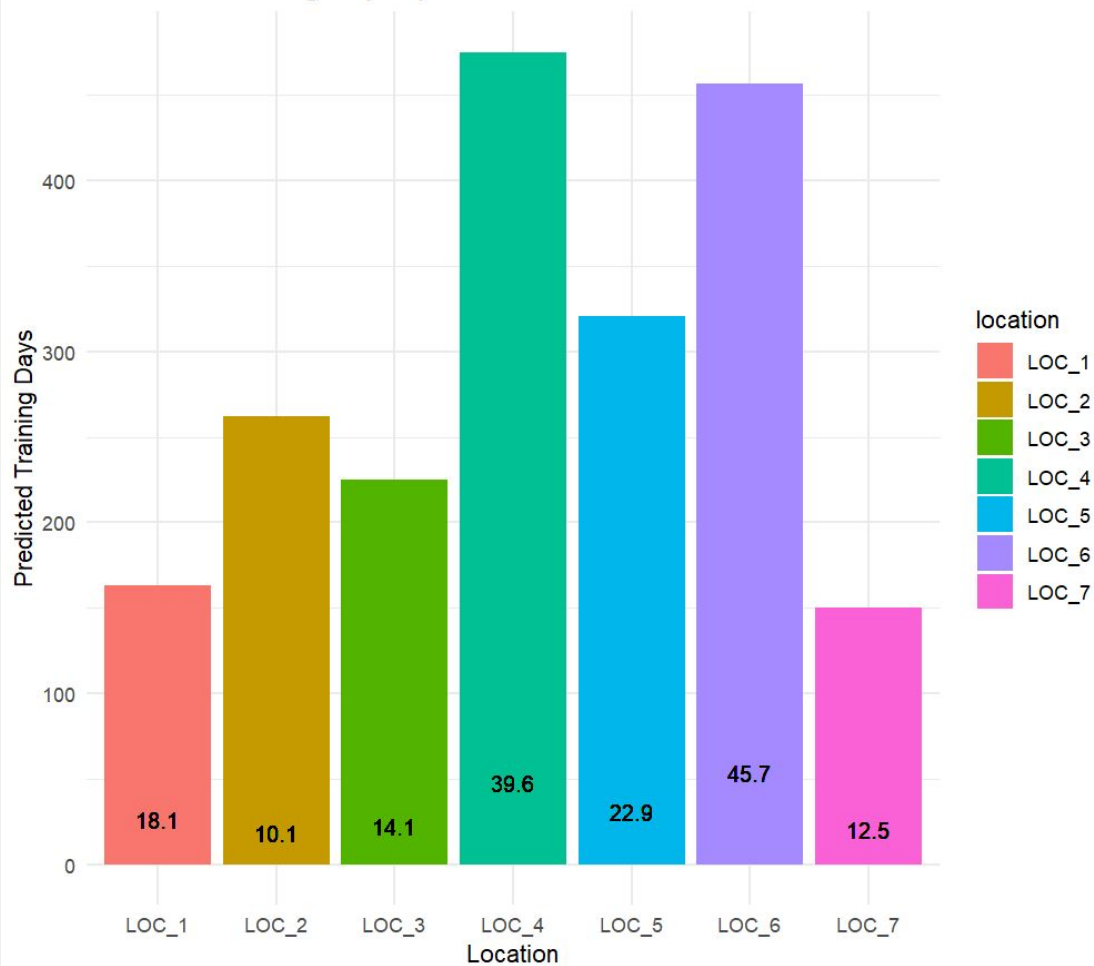
Standard Deviation of training days by job type







Predicted Training Days by Location



Recommendations



Further Investigation

- No one in Location 1 completed their training
- Location_4 and Location_6 as they have high average days and high STDEV



Utilize Data-Driven Feedback

- Use data from the training outcomes and employee performance post-training to refine training programs.



Standardization

- Standardize the training process for warehouse job to lower the variability



Customized Training Time

- Different training “Sweet Spot” for different locations and job types
- Loc_4 & Loc_6: 40 days
- Loc_1,2,3,5: 10-20 days
- DOTM and Warehouse requires longer training



Analysis 2:

Turnover rate



Method



Step 1

Used Python to determine the number of "Y" in the the column "Termination_Before_Certification"



Step 2

Used Python to Determine the Number of "Y" in different Locations, Job Titles, and Employee Types



Step 3

Used Excel to determine the turnover rate for each category by calculating the number of "Y" divided by the number of "Y+N"



Step 4

Finally, we graphed turnover rate of each category to see the differences

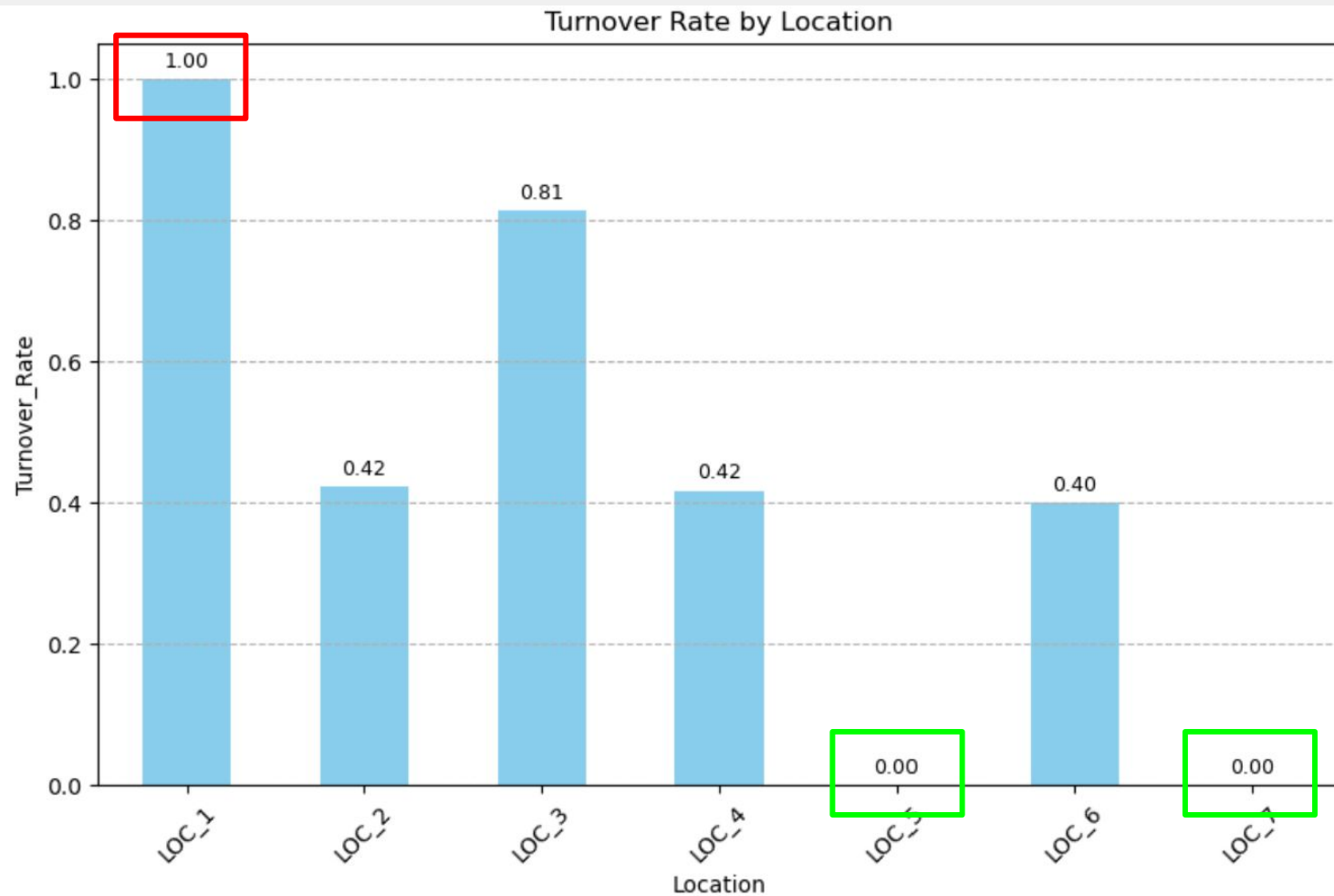
Finding

```
y_n_counts_by_location = df.groupby(['location', 'term_before_cert']).size().unstack(fill_value=0)

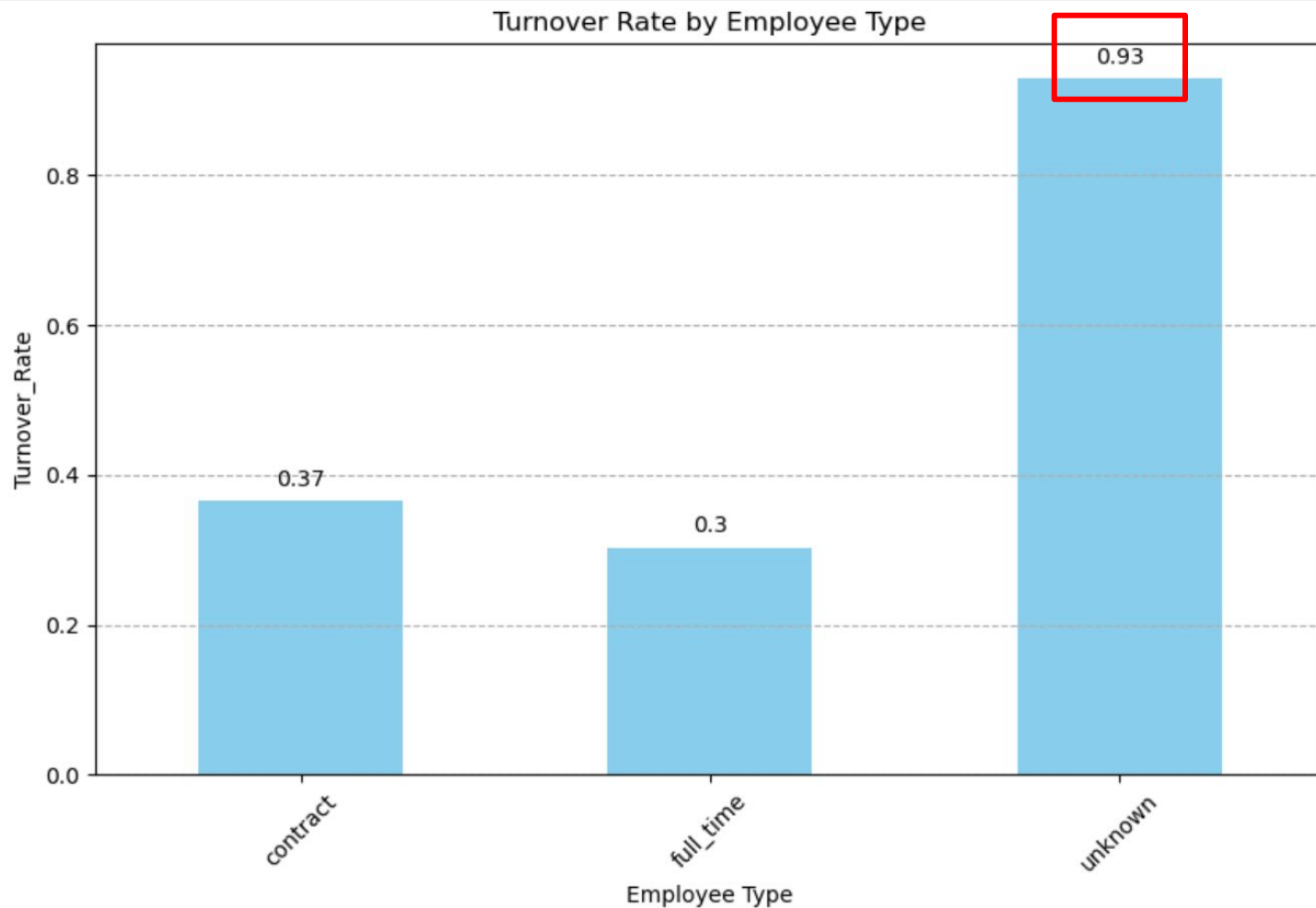
Turnover_Rate = y_n_counts_by_location['Y'] / (y_n_counts_by_location['Y'] + y_n_counts_by_location['N'])

Turnover_Rate
```

```
location
LOC_1    1.000000
LOC_2    0.423077
LOC_3    0.812500
LOC_4    0.416667
LOC_5    0.000000
LOC_6    0.400000
LOC_7    0.000000
dtype: float64
```







Recommendations



Employee Surveys

- Location 1 has 100 % turnover rate
- Conduct root cause analysis
- Factors like poor working conditions, inadequate management, or non competitive wages.



Rationalize Workloads

- Ensure workloads are reasonable and well-distributed to prevent burnout.
- Use workload assessments to ensure that employees are not consistently overburdened.



Retention Programs

- Develop targeted retention programs, such as career development paths, better benefits packages, or work-life balance initiatives.



Analysis 3:

Excessive delays

DELA



Method



Step 1

First used Tableau to visualize locations, job type, and total days for training and how termination and completion time was affected



Step 2

Next we identified locations that we considered having excessive differences in total training days to certification



Step 3

Calculated percentage difference between max and min days needed



Step 4

Define what we believe qualifies as reasonable to classify as an excessive delay in training completion

Excessive Delays Overview



Excessive Delays Location 4



Job Title

- Warehouse
- Warehouse Associate

Term Before Cert

- N
- Y

Days Total

- 0
- 20
- 40
- 60
- 77

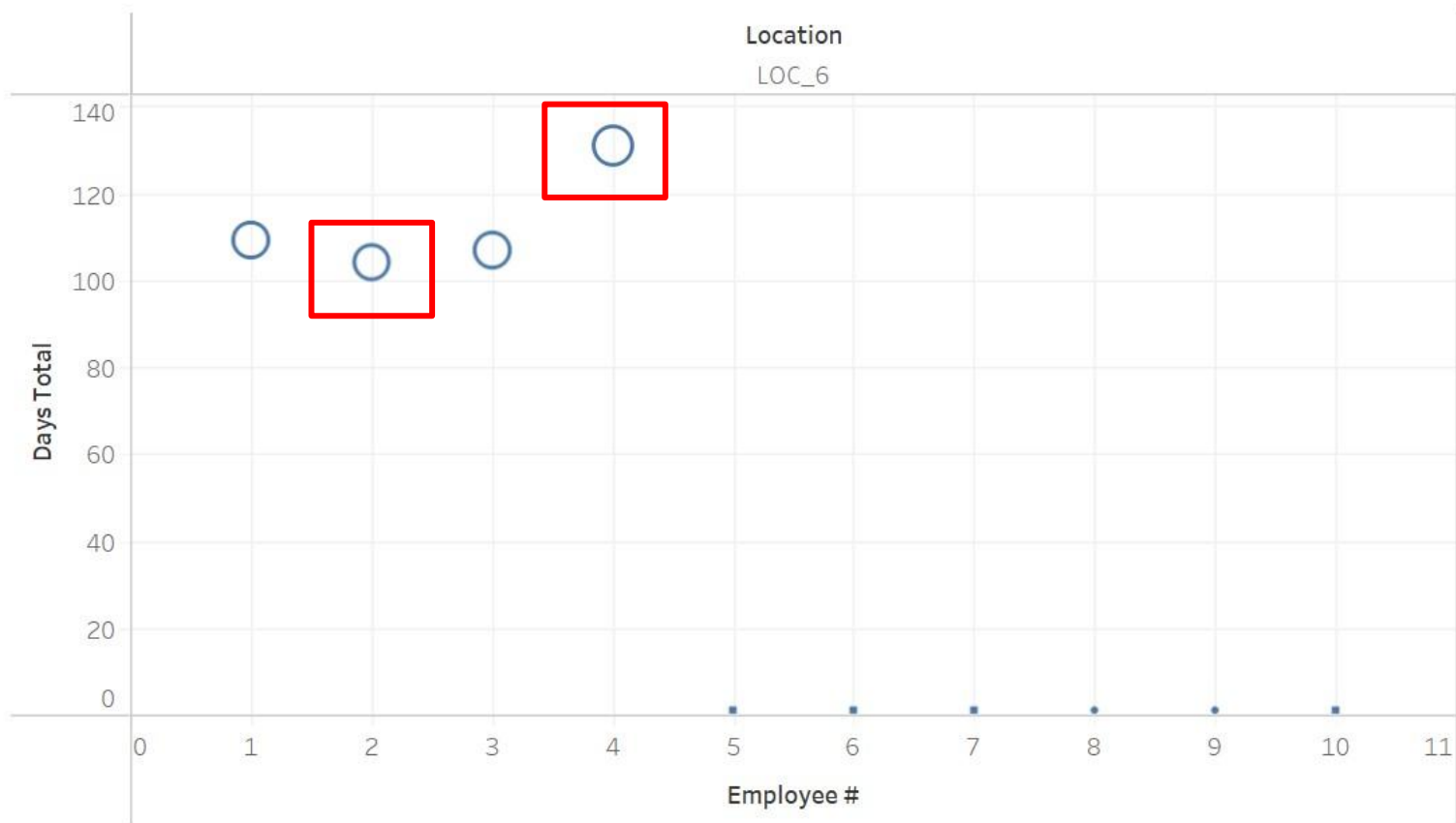


Excessive Delays: Location 4

- **Warehouse** training showed significant disparity in training time
 - Took as little as 51 days
 - As much as 77 days
 - 51 -> 77 is a 50.98% increase
 - 50% of warehouse applicants were terminated
- **Warehouse Associate** displayed an immense difference in training time
 - Took as little as 11 days
 - As much as 62 days
 - 11 -> 62 days is a 463.636% increase
 - 25% of warehouse associate applicants were terminated



Excessive Delays Location 6



Job Title

DOTM

Term Before Cert

N

Y

Days Total

1

50

100

131



Excessive Delays: Location 6

- **DOTM** training also had an excessive difference
 - As little as 104
 - As much as 131
 - Significant increase of 25.96% from 104 -> 131
- Many of the DOTM job applicants were terminated in the infancy of their training
 - 60% of applicants were terminated before more than 1 day had elapsed



Recommendations



Streamline Onboarding Processes

- Make onboarding procedures more efficient.
- Eliminate unnecessary steps/paperwork reduces burden of employee and those training them.
- More efficient processes reduce delays and terminations.



Just-in-Time Training

- Easy to overwhelm new employees with training.
- Providing training that is relevant to the current needs, the company can avoid overwhelming the employee
- Aids in retaining information and training success.



Feedback and Assessment

- Allows for correction and encouragement.
- Ultimately lead to shorter training times
 - This will motivate employees to develop and refine their skills more effectively.
- Would allow excessive delays to be flagged and necessary steps to be taken.





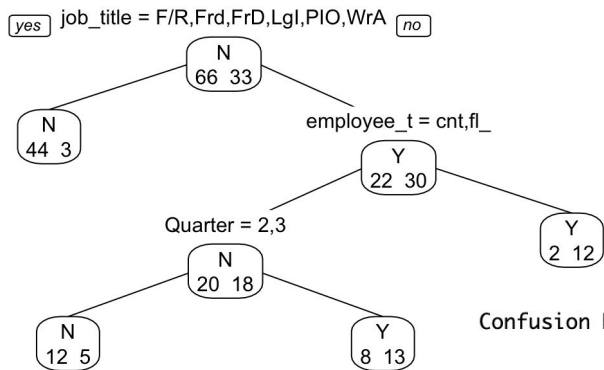
03

Predictive Modeling



Decision Tree Model

Decision Tree for Term Before Certification



Confusion Matrix and Statistics

	actual	
predictions	N	Y
N	56	8
Y	10	25

Accuracy : 0.8182
95% CI : (0.728, 0.8885)
No Information Rate : 0.6667
P-Value [Acc > NIR] : 0.0006096

Benefits:

- Simple & comprehensive
- Easy implementation
- Highlights important features

Risks:

- Sensitive
- Built using the data we cleaned
- Built on small dataset



04

Conclusion





Final Takeaways

- **Standardize training periods:** 60-90 day training window will not fit for every position.
- **Customize Training:** Revise DOTM and Warehouse training experience.
- **Standardize data collection:** key information such as “term before cert” was missing
 - Data was open to interpretation which can cause issues with seeing the bigger picture
- **Employee Engagement:** Conduct Exit surveys.





Thanks!

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**

