Paige Leeds

PREDICTING EMPLOYEE ATTRITION

PACE STRATEGY

Google Advanced Data Analytics Professional Certificate

Project Documentation

2026

Table of Contents

00

Overview

Objective

Develop a predictive model to identify employees at risk of leaving Salifort Motors and translate analytical findings into actionable retention strategies.

Business Question

Which workplace factors are most strongly associated with employee attrition, and how accurately can employee turnover be predicted using historical workforce data?

Dataset

- 14,999 employee records
- 10 workplace variables
- Employee demographics, satisfaction, performance, workload, tenure, department, salary, and attrition status

PACE Workflow

Plan → Analyze → Construct → Execute

01

Plan

Approach

Establish a structured workflow to define the business problem, understand the workforce dataset, and prepare the data for predictive modeling.

Business Context

Employee turnover creates costs through recruitment, onboarding, lost productivity, and knowledge loss. The objective is to identify the drivers of attrition and inform data-driven retention strategies.

Preparation

- Review project objectives and stakeholder needs
- Inspect the dataset and data dictionary
- Validate data quality and variable definitions
- Identify potential outliers and inconsistencies
- Define the target variable and modeling approach

Resources

Python • pandas • NumPy • scikit-learn

02

Analyze

Dataset Assessment

- 14,999 employee records • 10 variables
- Reviewed data types, missing values, duplicates, and distributions

Key Observations

- Verified overall data quality
- Explored relationships between workplace factors and attrition
- Identified patterns, potential outliers, and class balance

Exploratory Data Analysis

- Generated descriptive statistics
- Examined correlations and feature relationships
- Prepared the dataset for logistic regression

Deliverables

- Cleaned dataset
- Exploratory analysis notebook
- Feature summary and EDA visualizations

03

Construct

Model Development

Developed a logistic regression model to predict employee attrition using workplace and employee characteristics.

Feature Engineering

- Selected predictor variables
- Created training and testing datasets
- Encoded categorical variables
- Standardized numerical features

Model Evaluation

- Evaluated accuracy, precision, recall, and F1-score
- Interpreted model coefficients and predictive performance

Deliverables

- Trained logistic regression model
- Model evaluation report
- Attrition predictions

04

Execute

Key Findings

- Long working hours were strongly associated with attrition.
- Satisfaction, workload, and evaluation influenced turnover.
- Logistic regression identified significant attrition predictors.

Business Recommendations

- Improve workload balance and employee well-being.
- Monitor high-risk employees using predictive analytics.
- Support retention with targeted HR interventions.

Project Outcome

Developed a predictive model that identified employees at risk of attrition and translated findings into actionable retention strategies.

Deliverables

- Attrition risk predictions
- Retention recommendations
- Final project documentation