# Digital Analytics and Causal Inference Project Report

**JEZI Co. Consulting Group:**
Jordan Meacham, Efe Comu, Zhizhi Wang, and Innocent Mukoki

## Executive Summary

ACME Manufacturing launched the Career 2030 training program to support employee career development and internal mobility. To assess its effectiveness, our consulting group, JEZI Co., conducted a rigorous causal inference analysis to evaluate the program's impact on employee promotions. Given concerns about selection bias due to managerial intervention and self-selection, we applied advanced statistical methods to ensure a robust and unbiased assessment. Our findings provide strong statistical evidence that training participation significantly increases promotion likelihood, reinforcing data-influenced decision-making for workforce development.

The primary objectives of this study were to determine whether participation in Career 2030 significantly increased the likelihood of promotion, account for confounding variables using statistical methods, provide data-influenced recommendations to optimize the training program and assess the robustness of findings through sensitivity analysis. By utilizing 1:1 Nearest Neighbor Matching (NNM), Propensity Score Matching (PSM), and Inverse Probability of Treatment Weighting (IPTW), we estimated the true effect of training while mitigating confounding variables.

Our findings provide strong statistical evidence that training significantly increases promotion odds. A baseline logistic regression model showed trained employees were 1.24 times (24%) more likely to be promoted, but significant covariate imbalances required adjustment. NNM improved balance but reduced the sample to 105 matched pairs, estimating a 157% increase in promotion odds. PSM retained 1,503 pairs (OR = 2.34, 134%), while a stricter PSM caliper (0.002) confirmed a 157% increase. IPTW provided the most robust estimate, retaining the full sample and finding a 287% increase (OR = 3.87, $p < 0.001$). A Wilcoxon Signed-Rank Test ($p < 2.2e-16$) reinforced statistical significance, and Leave-One-Out Sensitivity Analysis confirmed result stability.

To optimize Career 2030, we recommend enhancing data collection and experiment design by reinforcing randomization, tracking managerial overrides, and collecting additional data on employee motivation and performance. Longitudinal tracking should assess the program's long-term impact, while stratified randomization and survey-based motivation indicators can help mitigate self-selection bias.

From a business perspective, ACME should expand access to Career 2030, address participation barriers, and incorporate training completion into promotion criteria. Post-training assessments and a real-time analytics dashboard can enhance program monitoring and continuous improvement. By implementing these recommendations, ACME can maximize Career 2030's impact, ensure equitable access, and drive data-informed workforce development decisions.

# 1. Introduction

ACME Manufacturing, a company with over 60,000 employees, launched the Career 2030 training program as a strategic initiative to foster career development and promote internal mobility. The program aims to equip employees with the necessary skills to improve their chances of promotion and long-term retention within the organization. ACME's leadership team is interested in evaluating the effectiveness of this training program, particularly its impact on promotion rates. However, anecdotal evidence suggests that external factors, such as managerial intervention and employee self-selection, may have influenced participation. Some managers reportedly advocated for their direct reports to join the program, and certain highly motivated employees actively pursued training opportunities. These potential sources of selection bias complicate a straightforward comparison of promotion rates between trained and untrained employees.

To address these challenges, our consulting group, JEZI Co., conducted a rigorous causal inference analysis to evaluate the relationship between Career 2030 participation and employee promotions. The key objectives of this project were to determine whether training significantly increases the likelihood of promotion, control for confounding variables using statistical methods, and offer data-influenced recommendations for improving the training program. We employed various advanced causal inference techniques to account for biases introduced by managerial influence and motivated self-selection, allowing for a more accurate estimate of Career 2030's effect. Additionally, sensitivity analyses were conducted to test our findings' robustness and ensure our conclusions' reliability.

This report is structured as follows. The Background section provides an overview of the Career 2030 program, its intended goals, and the challenges in evaluating its effectiveness. The Data Cleaning and Exploratory Data Analysis (EDA) section presents a detailed examination of the dataset, including summary statistics, data visualizations, and identifying trends or biases that may impact the analysis. The Data Preprocessing section outlines the steps to prepare the data for causal inference techniques, ensuring that variables are properly formatted and potential biases are addressed. The Methodology section details the causal inference techniques employed, such as 1:1 Nearest Neighbors Matching (NNM), Propensity Score Matching (PSM), and Inverse Probability of Treatment Weighting (IPTW), explaining their role in mitigating selection bias and improving the validity of the study. The Results and Findings section presents the outcomes of the statistical models, providing estimates of the training program's impact on promotion rates and assessing the robustness of these findings through sensitivity analyses. Finally, the Discussion and Interpretation section contextualizes the results, explores potential limitations, and offers data-influenced recommendations for optimizing Career 2030.

# 2. Background

Measuring the true causal impact of Career 2030 on promotion is critical for ACME's long-term workforce planning. The results of this study will help ACME quantify the program's effect, identify biases in training participation, and provide actionable insights for optimizing future training programs. ACME can make data-driven decisions regarding its talent development strategies by understanding whether the program effectively contributes to career advancement and whether access to training is equitable. Furthermore, as ACME considers expanding and refining Career 2030, ensuring that the program benefits employees fairly and efficiently will be essential.

The dataset used for this analysis consists of 6,000 employee records, capturing a range of demographic, employment, and social factors. The key variables include whether an employee participated in the Career 2030 program (training), received a promotion within a year (promoted), and various attributes such as height, weight, age, salary, education level, distance from the training facility, and test scores. While the dataset originates from a randomized controlled trial (RCT) where 5% of employees were randomly selected for training and another 5% were excluded, external managerial influence and self-selection factors required additional statistical adjustments to obtain an unbiased estimate of the training program's impact.

# 3. Data Cleaning & Exploratory Data Analysis (EDA)

To gain an initial understanding of the dataset and identify potential confounders, we conducted an extensive exploratory data analysis (EDA). A key focus of the analysis was to compare trained and untrained employees to assess potential differences that could impact promotion rates.

*Data Cleaning*

Before conducting the analysis, the dataset was carefully examined for inconsistencies and missing values. A completeness check revealed no missing values across variables, allowing for a straightforward analysis without imputation. However, an issue was identified in the education variable, where some records contained a value of -1, likely representing an error or a placeholder for missing data. To correct this, all occurrences of education below zero were adjusted to 0, ensuring a more accurate representation of employees with no formal education. Apart from this correction, the dataset was already clean and structured.

*Summary Statistics of Key Variables*

The summary statistics computed for the variables help give us more insight into the distribution and characteristics of the employee population.

- **Age**: The dataset includes employees ranging from 18 to 71 years old, with a mean age of approximately 43 years.
- **Education**: Employee's education levels range from 0 to 30 years, with a median of 12 years. The majority of employees have between 10 and 13 years of education.
- **Salary Brackets**: Employees are distributed across four salary groups (Under $20k, $20k-$40k, $40k-$80k, and Above $80k), with a significant number earning below $20k.
- **Test Scores**: The test scores range from 17 to 100, with a median of 63, serving as a

potential indicator of cognitive ability or job readiness.

- **Distance from Training Facility:** Employees live at varying distances from the training facility, with a median distance of approximately 22 miles.
- **Children:** Employees have between 0 and 5 children, with an average of 1-2.
- **Training Participation**: Approximately 38.2% of employees participated in the Career 2030 training program, while 61.8% did not.
- **Promotion Rate:** Overall, 65.1% of employees received a promotion within a year, while 34.9% did not.

*Training Status and Promotion Insights*

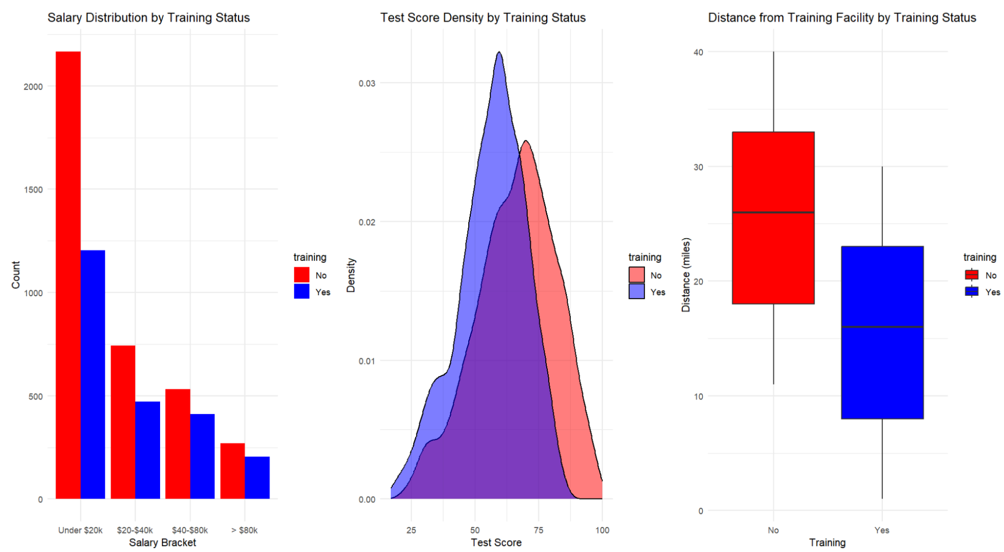



**Figure 1:** *EDA Insights on potentially confounding variables in relation to Training Status.*

A key focus of the exploratory analysis was understanding how training participation and promotion rates are distributed. Visualization of salary by training status highlighted a stark difference: untrained employees were more likely to earn below $20k, while trained employees had a more balanced representation across salary levels (**Figure 1**). Additionally, test score density plots revealed that untrained employees generally had higher test scores than their trained counterparts. This suggests that those not participating in the program may already have stronger initial qualifications or preparedness for career advancement (**Figure 1**). A boxplot comparing the distance from their home to the training facility between trained and untrained employees indicated that trained employees tended to live closer, implying that logistical barriers may have influenced participation (**Figure 1**).

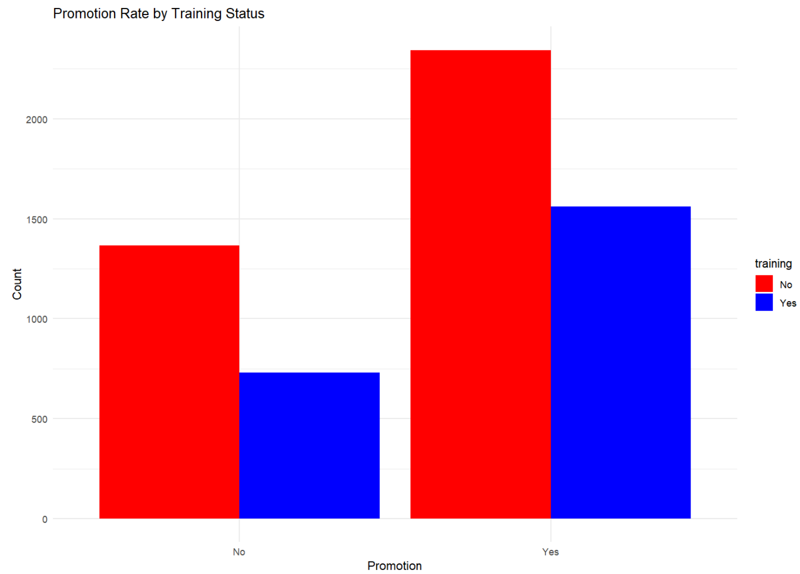



**Figure 2:** *Bar chart comparison between promotion and training status of employees.*

Promotion rates varied notably between trained and untrained employees. **Figure 2**, which compares promotion outcomes, indicates that a larger proportion of untrained employees were promoted compared to trained employees. This contradicts the expectation that training would lead to higher promotion rates and suggests potential confounding factors influencing promotions. Given the observed disparities in salary, test scores, and geographic proximity between the two groups, it is unclear whether training participation directly impacted promotion likelihood or if pre-existing advantages among untrained employees contributed to the difference.

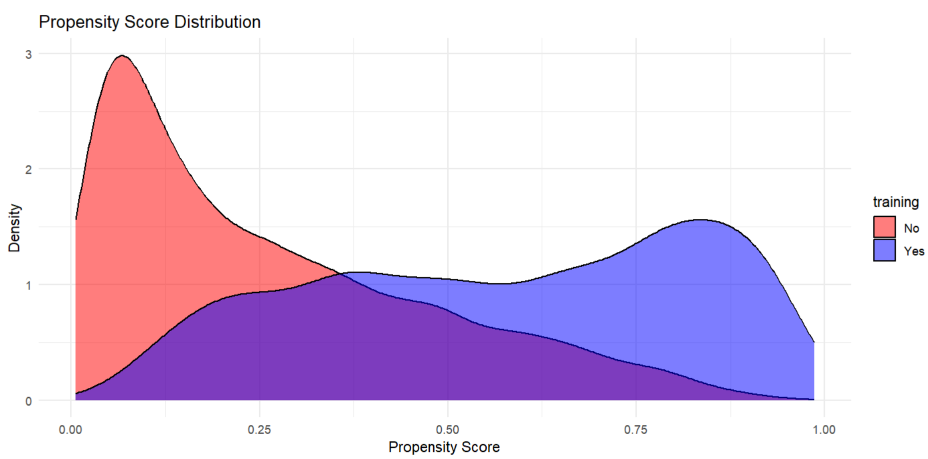



**Figure 3:** *Distribution of propensity scores based on untrained vs. trained employees.*

Additionally, in **Figure 3**, the distribution of propensity scores, which estimate the likelihood of training participation based on observed characteristics, reinforces concerns that the two groups were not perfectly comparable, necessitating causal inference techniques to account for selection bias.

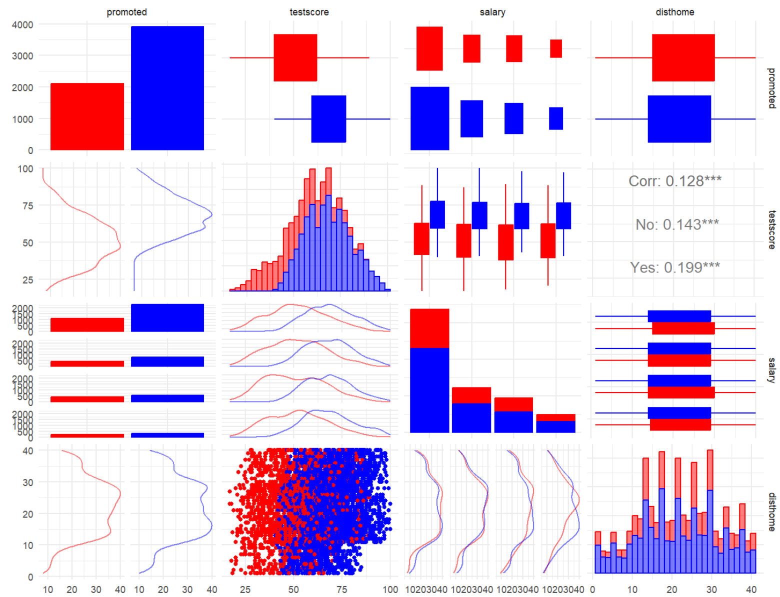



**Figure 4:** *Pairwise plots for potential confounding variables and promoted vs. not promoted employees. Red = Not Promoted, Blue = Promoted*

The pairwise visualizations of key variables in **Figure 4** provide an in-depth comparison between promoted employees and those who were not, highlighting disparities that may impact promotion outcomes. The bar chart of promotion rates shows that a greater proportion of employees who were promoted had not participated in the training program. This suggests that factors beyond training may have played a significant role in determining career advancement.

Test scores show a notable disparity between promoted and non-promoted employees. Promoted employees had a higher distribution of test scores, reinforcing the idea that initial competency plays a role in career progression. Similarly, salary distributions reveal that promoted employees were more likely to fall into higher salary brackets. In contrast, those who were not promoted had a larger share in the lower-income categories. This suggests that compensation level may be indicative of career progression opportunities.

The correlation matrix further highlights key trends. Higher test scores and salaries positively correlate with promotion, confirming that employees with stronger qualifications and financial stability had higher chances of moving up in the organization. Distance from the training facility does not have a strong direct relationship with promotion, but its influence on training participation could still make it an indirect factor.

*EDA Conclusion*

The exploratory data analysis reveals substantial differences in key employee attributes between promoted and non-promoted individuals. The data suggests that factors such as test scores, salary, and training participation are associated with promotion outcomes, though the direction of causality remains unclear. Since training participation was not entirely random, and imbalances exist in multiple covariates, further statistical techniques must be applied to measure the program's effectiveness accurately. To address these concerns, causal inference methods such as 1:1 Nearest Neighbors Matching, Propensity Score Matching (PSM), and Inverse Probability of Treatment Weighting (IPTW) will be implemented to adjust for selection bias and ensure a more reliable assessment of Career 2030's impact on promotion likelihood.

# 4. Data Pre-processing

We performed the necessary preprocessing steps before conducting causal inference techniques to prepare the dataset for analysis. This included examining the balance between trained and untrained employees, converting categorical variables for analysis, and developing a matching approach to mitigate selection bias.

*Handling Missing or Incorrect Values*

As mentioned in the previous section, a preliminary assessment of missing data showed no significant missingness across the dataset, eliminating the need for imputation. However, an inconsistency was identified in the education variable, where some employees had an education value of -1, likely indicating an error or a missing value placeholder. To correct this, we replaced all negative instances with 0, ensuring the data accurately reflected employees with no reported education beyond the required minimum.

*Transformations and Variable Adjustments*

Several variables required transformation to improve performance and standardization. To improve the accuracy of our matching process and ensure the best possible results, we applied transformations to variables with high or moderate skew. Specifically, we adjusted children, weight, and vacation days taken in the previous year. Given their high skew, we transformed children and weight using a log(x + 1) transformation to normalize their distributions. We applied a square root transformation for vacation days, which exhibited moderate skew to smooth the data while preserving its interpretability.

*Encoding Categorical Variables*

Since the dataset contained multiple categorical variables, we applied dummy encoding to facilitate numerical analysis while preserving meaningful group distinctions. Variables were converted into binary indicators or dummy variables, ensuring they could be appropriately utilized in propensity score modeling.

*Selection and Justification of Key Variables*

The key variables in the analysis include demographic factors such as age, sex, race, and marital status, as well as job-related factors like manager status, raise, salary, insurance, flexible spending, retirement contributions, distance from home to training facility, and test scores. Additionally, due to skewness, transformations were applied to children (log), weight (log), and vacation days (square root). These variables were used in matching and propensity score estimation to balance the dataset and assess the causal effect of training on promotion.

By implementing these pre-processing steps, we ensured the dataset was clean, structured, and ready for causal inference techniques such as Propensity Score Matching (PSM) and Inverse Probability of Treatment Weighting (IPTW). This rigorous preparation was critical to addressing selection bias and producing a valid estimate of Career 2030's impact on employee promotions.

# 5. Methodology

To evaluate the causal impact of the Career 2030 training program on employee promotions, we employed multiple causal inference techniques designed to address selection bias and ensure comparability between trained and untrained employees. Given the evidence of confounding factors, we applied 1:1 Nearest Neighbors Matching (NNM), Propensity Score Matching (PSM), and Inverse Probability of Treatment Weighting (IPTW) to balance the groups and estimate the true effect of training on promotion likelihood.

**5.1 Assumption Testing: Evaluating Covariate Balance and Confounding Variables**

A key assumption of causal inference is that treatment and control groups should be comparable except for their treatment assignment. To test this assumption, we ran an initial logistic regression model predicting promotion based solely on training participation to establish a baseline effect:

$$log\ odds\ (promotion) = \beta_0 + \beta_1(training) + \epsilon$$

The results indicated that trained employees were 1.24 times (24%) more likely to be promoted than untrained employees ($p < 0.001$). However, the residual deviance was close to the null deviance, suggesting that unobserved confounders influenced promotion likelihood.

Additionally, we conducted a pre-matching balance check using standardized mean differences (SMDs) across key covariates. The results confirmed that several variables—including raise, salary, marital status, weight, insurance, participation in flexible spending account, retirement contribution, distance from home to the training facility, and test score—were significantly imbalanced between trained and untrained employees (SMD > 0.1). This imbalance highlighted the need for statistical adjustments before making causal claims. Given these findings, we proceeded with matching and weighting techniques to address selection bias.

**5.2 Matching Techniques**

*One-to-One Nearest Neighbor Matching (NNM)*

To create a more balanced dataset, we first applied 1:1 Nearest Neighbor Matching without replacement. Each trained employee was matched to an untrained employee with the most similar covariate values, minimizing differences in pre-treatment characteristics. A caliper (maximum allowed distance for matching) was tuned to improve balance. However, this method alone did not fully address imbalances.

*Propensity Score Matching (PSM)*

Given the remaining imbalances, we implemented Propensity Score Matching to ensure that trained and untrained employees were comparable based on observable characteristics. A logistic regression model was used to estimate the probability of training participation based on covariates:

$$P(Training) = \frac{e^{\beta X}}{1 + e^{\beta X}}$$

Employees with similar propensity scores were matched, reducing selection bias. We tested multiple caliper values (ranging from 0.2 to 0.001) to optimize balance. The final matched sample significantly improved covariate balance throughout all variables but did have a violation of positivity, which we will address in the next section.

*Inverse Probability of Treatment Weighting (IPTW)*

For our next approach, we applied Inverse Probability of Treatment Weighting, which assigns weights based on the inverse probability of receiving the observed treatment. This method reweights the dataset so that trained and untrained employees resemble a randomized experiment. The weight formula is:

$$w_i = \frac{1}{P(T_i | X_i)}$$

Where the denominator represents the estimated propensity score, after weighting, we reassessed covariate balance and observed significant improvements, particularly in salary, test scores, and distance to training. We also trimmed extreme weights (top 1%) to stabilize estimates and reduce variance.

**5.3 Justification for Method Selection**

While NNM improved balance across most covariates, PSM was more effective in achieving balance across all variables. IPTW provided the most robust solution, ensuring that trained and untrained employees were statistically comparable across all key confounders. We still did have an imbalance within two covariates, but we accounted for them within the regression model, which we will address in the next section. Given ACME's goal of obtaining an unbiased estimate of Career 2030's impact, IPTW was chosen as the primary method for final analysis.

Through this structured approach—starting with balance testing, implementing multiple matching techniques, and selecting IPTW as the final method based on the adjusted odds ratio—we ensured that our analysis accurately measured the effect of training on promotion while minimizing the influence of confounding factors.

# 6. Results and Findings

Our analysis aimed to estimate the causal effect of the Career 2030 training program on employee promotion rates by addressing selection bias through NNM, PSM, and IPTW. Below, we summarize the key findings from each method, balance assessments, and the overall treatment effect estimation.

**Balance Assessment and Initial Findings**

Before applying matching or weighting techniques and to understand our starting point with the data, we conducted a pre-matching balance assessment to evaluate differences between the treatment group (training = Yes) and the control group (training = No). The standardized mean difference (SMD) was used to measure the covariate imbalance, with an SMD threshold of 0.1 requiring further investigation and 0.2 indicating a meaningful imbalance. The results revealed significant imbalances across several covariates, including raise, salary, marital status, weight, insurance, flexible spending account participation, retirement contributions, distance from home to the training facility, and test score, with many exceeding an SMD of 0.2 or higher (**see appendix**). This suggests that the assignment to training was not random and that confounding variables could bias the estimated effect of training on promotion.

To establish a baseline estimate, we fit an unadjusted logistic regression model predicting promotion as a function of the training. The results showed a statistically significant effect, where employees who underwent training at 1.24 times (24%) had greater odds of being promoted than those who did not. However, given the observed imbalances in the key covariates, this estimate is likely biased due to confounding. This underscores the necessity of employing matching or weighting techniques to ensure a more accurate estimate of the causal effect of training on promotion.

**1:1 Nearest Neighbor Matching (NNM)**

To improve the covariate balance between the treatment (training = Yes) and control groups, we implemented 1:1 Nearest Neighbors Matching (NNM). By utilizing a loop to test all caliper levels within a given range, we found that the caliper of 1.10 resulted in the fewest imbalances (**see appendix**). After implementing this caliper, we observed a significant improvement in covariate balance, reducing the number of imbalanced covariates to only two. However, this improvement came at the cost of a drastic reduction in sample size, with only 105 matched pairs retained. While most covariates have near-perfect balance, our distance to home from the training facility variable (disthome) remained highly imbalanced (SMD = 0.249), suggesting that even with various calipers, full balance was not achieved for all variables. This highlights the trade-off between improving balance and reducing statistical power, and stricter matching criteria eliminated a large portion of the sample.

The logistic regression model on the matched data set produced a statistically significant result for training, indicating that participation in the training program significantly increases the likelihood of being promoted; specifically, trained individuals are 2.57 times (157%) more likely to get promoted in comparison to their untrained counterparts. The intercept is not statistically significant, indicating that the baseline probability of promotion (without training) does not differ significantly from zero. Additionally, the model's lower AIC suggests that we have a reasonable model fit given the reduced sample size.

**Propensity Score Matching (PSM)**

When using Propensity Score Matching (PSM), we aimed to create a balanced comparison between the treatment (training = Yes) and control groups by matching participants based on their propensity scores. We gathered insights from PSM using both the largest optimal propensity score, 0.31, and the smallest, 0.002 (**see appendix**).

*PSM with 0.31 Caliper*

When applying the largest optimal caliper of 0.31, we successfully balanced all covariates. While this is good news, unfortunately, we had a positivity violation within the Medicare & Medicaid insurance category. Although, because of the nature of the insurance variable, being the subgroups are not mutually exclusive, this positivity violation could potentially not exist if we were to combine the insurances into a single metric or be insignificant. Despite this, the estimated effect of training on promotion remained statistically significant, with an odds ratio of 2.34, indicating that participants are 134% more likely to be promoted than their untrained counterparts. This model retained 1,503 matched pairs, ensuring a relatively large sample size while achieving a strong covariate balance.

*PSM with 0.002 Caliper*

Additionally, we applied a stricter 0.002 caliper to minimize SMDs further. This approach resulted in 1,297 matched participants, improving balance further while still maintaining a significant effect of training. The corresponding odds ratio was 2.57, suggesting that training increased promotion odds by 157%. We did still have the same occurrence of a positivity violation, but as stated previously, this could not be a significant issue. Overall, PSM effectively balanced covariates and confirmed the strong positive impact of training on promotion, with results consistent across different calipers.

**Inverse Probability of Treatment Weighting (IPTW)**

When using IPTW, we aimed to create a pseudo-population where treatment assignment is independent of baseline characteristics, improving the validity of our causal estimate. This weighting approach helped reduce bias and allowed us to compare trained and non-trained employees more fairly in terms of their likelihood of promotion.

*Unadjusted IPTW Model*

Applying IPTW improved covariate balance substantially, and a logistic regression model without additional adjustments estimated $\beta = 0.731$ ($p < 0.001$), corresponding to an odds ratio of 2.08, suggesting trained employees were about twice as likely to be promoted as their non-trained counterparts. Despite this, for the unadjusted model, we still had two unbalanced covariates.

*Adjusted IPTW Model*

When using IPTW, we adjusted for confounding and estimated the causal effect of training on promotion. The final weighted logistic regression model included distance from home and test score variables, as both had SMDs exceeding 0.2 in the balance assessment (**see appendix**). Although raise and flexible spending account participation were other potential imbalanced covariates, they were not controlled for due to their insignificant impact on the model's results.

The regression analysis confirmed that training had a statistically significant effect on promotion with an estimated odds ratio of 3.87, meaning individuals who participated in the program were 287% more likely to be promoted than those who did not. While the test score variable was also highly significant, indicating a strong relationship between performance and promotion, the distance between home and training facility was not, suggesting that distance from home had no meaningful impact on promotion likelihood.

Overall, IPTW effectively reduced imbalance and provided a strong estimate of training's impact on promotion, reinforcing findings from other matching methods. The consistent significance across the different approaches supports the conclusion that training substantially increases promotion likelihood.

# 7. Discussion and Interpretation

Our analysis provides strong statistical evidence that participation in the Career 2030 training program significantly increases the likelihood of promotion. However, causal inference relies on the assumption that treatment and control groups are comparable, except for their treatment assignment. To test this assumption, we first established a baseline logistic regression model predicting promotion solely as a function of training. The results showed that trained employees were 1.24 times (24%) more likely to be promoted ($p < 0.001$). Still, the residual deviance was close to the null deviance, suggesting that unobserved confounders likely influenced promotion outcomes. Further, a pre-matching balance check using standardized mean differences (SMDs) revealed significant imbalances across several key covariates, including raise, salary, marital status, weight, insurance, flexible spending account participation, retirement contributions, distance to the training facility, and test score (SMD > 0.1). These findings confirmed the presence of selection bias, necessitating statistical adjustments before making causal claims.

To address these imbalances, we applied 1:1 Nearest Neighbor Matching (NNM), Propensity Score Matching (PSM), and Inverse Probability of Treatment Weighting (IPTW). When comparing methodologies, NNM with a caliper of 1.10 significantly improved covariate balance, reducing imbalances to only two variables but drastically reduced the sample size to 105 matched pairs. This trade-off between balance and statistical power was also observed in PSM, where a larger caliper (0.31) retained 1,503 pairs and achieved strong balance while still showing a minor positivity violation in the Medicare & Medicaid insurance category. A stricter PSM caliper (0.002) further minimized SMDs, retaining 1,297 matched pairs, confirming the training program's impact with an odds ratio of 2.57 (157% increase in promotion odds).

IPTW, in contrast, allowed us to retain the full sample while addressing covariate imbalance, producing the highest estimated treatment effect (OR = 3.87, or a 287% increase in promotion odds). Although IPTW effectively balanced most variables, minor imbalances persisted in test scores and distance from the training facility (disthome), which ultimately had no significant effect on promotion ($p = 0.806$).

Despite the strength of these findings, potential biases and limitations must be considered. Selection bias remains a concern, as employees who opted into training may inherently differ from those who did not in ways that are not fully captured by our observed covariates. Additionally, while positivity violations were noted in some models, they were mainly confined to non-mutually exclusive insurance subcategories, making them unlikely to impact results significantly. Another limitation is unobserved confounding, as factors such as managerial favoritism or external skill development could influence promotion likelihood but were not explicitly included in the dataset.

To further validate our findings, we conducted two robustness checks: a Wilcoxon Signed-Rank Test and a Leave-One-Out Sensitivity Analysis (**see appendix**). The Wilcoxon Signed-Rank Test produced an extremely small p-value ($p < 2.2e-16$), reinforcing that the observed difference in promotion rates between trained and untrained employees was statistically significant. The Leave-One-Out Sensitivity Analysis also confirmed that no single observation significantly influenced the results, as the estimated treatment effect remained stable across iterations. This indicates that our findings are not driven by outliers and are robust to minor data perturbations.

In practical terms, these findings underscore the substantial benefits of the Career 2030 training program in facilitating career advancement. The statistical significance and consistency of our results across multiple methodologies provide strong empirical support for the program's effectiveness. Organizations seeking to enhance employee development and retention may benefit from expanding similar training initiatives, ensuring broader access to structured skill-building opportunities that promote career growth. However, continued monitoring of selection processes and eligibility criteria will be essential to maximize equity and effectiveness in promotion outcomes.

# 8. Recommendations

Given the findings from our analysis, we recommend several strategic actions to optimize the Career 2030 training program, enhance data collection and experimental design, and mitigate bias in future studies. Since ACME Manufacturing's Chief People Officer prioritizes data-informed decision-making, our recommendations focus on ensuring that training participation translates into meaningful career growth while refining future program evaluations to provide stronger causal insights. Additionally, as ACME evaluates potential long-term analytics partners, our approach highlights how rigorous causal inference methods can support continuous workforce development and decision-making.

To improve future evaluations of Career 2030, ACME should enhance its data collection and experiment design by reinforcing the randomized controlled trial (RCT) framework initially used to assign employees to training. While randomization was employed, anecdotal evidence suggests that managerial intervention and employee self-selection influenced participation, introducing potential selection bias. To maintain the integrity of the randomization process, we recommend that ACME limit managerial overrides and track cases where managers directly place employees into training to adjust for this confounding factor in future analyses. Furthermore, collecting additional data on employee motivation, performance before and after training, and managerial promotion recommendations would provide a more comprehensive picture of the program's impact. Expanding data collection to include longitudinal tracking of employees over multiple years would also help assess whether the effects of training on promotion persist over time.

Bias mitigation is also critical for ensuring that Career 2030 achieves its intended goals fairly and equitably. Our analysis revealed that employees who voluntarily enrolled in training may already be more motivated and career-driven, leading to an overestimation of training's true effect if this motivation is not properly accounted for. To address this, ACME could stratify randomization by department, tenure, and past performance to ensure that training assignments are balanced across employee groups. Additionally, incorporating survey-based motivation indicators into the dataset would allow future analyses to better control for self-selection bias. Our findings also identified positivity violations, particularly in insurance subcategories, suggesting a need for clearer categorization of employee benefits and work-related factors in future datasets. Finally, continuous refinement of propensity score weighting and inverse probability treatment weighting (IPTW) methods can help further reduce imbalances in observational analyses.

From a business and policy perspective, ACME should use these findings to expand and institutionalize Career 2030 as a core component of its workforce development strategy. Given the strong statistical evidence that training increases promotion likelihood, making training more widely accessible will help broaden career advancement opportunities across the organization. However, participation barriers such as time constraints, lack of managerial support, and uncertainty about the program's benefits may prevent some employees from engaging in training. To address this, ACME should implement targeted outreach initiatives for underrepresented employee groups, offer flexible training schedules, and provide incentives for training completion, such as leadership development opportunities or performance-based recognition.

Beyond expanding access, Career 2030 should be more deeply embedded into ACME's career progression framework. Training completion could be incorporated into promotion eligibility criteria, ensuring that employees recognize its value in advancing their careers. Additionally, post-training assessments and skill evaluations should be implemented to measure knowledge retention and fine-tune the training curriculum based on data-driven insights. ACME could also develop a real-time analytics dashboard to continuously track training participation, promotion outcomes, and retention trends across different departments and demographics. This would allow leadership to identify which aspects of the training program have the greatest impact and make iterative improvements based on ongoing data analysis rather than one-time evaluations.

By integrating these recommendations, ACME can enhance the effectiveness, accessibility, and long-term impact of Career 2030 while ensuring that future program evaluations are more rigorous and reliable. Our findings confirm that training significantly increases the likelihood of promotion, but by strengthening randomization controls, improving data collection, and refining evaluation methodologies, ACME can further optimize the program's effectiveness and equity. Additionally, as ACME considers long-term partnerships for program monitoring, our approach demonstrates how advanced causal inference techniques can provide actionable insights that drive meaningful workforce development decisions.

# 9. Conclusion

Our analysis provides compelling evidence that the Career 2030 training program has a significant impact on employee promotion rates. Through rigorous causal inference methods—including 1:1 Nearest Neighbor Matching (NNM), Propensity Score Matching (PSM), and Inverse Probability of Treatment Weighting (IPTW)—we addressed selection bias and improved covariate balance to estimate the true effect of training on promotion likelihood. The results across multiple methodologies consistently indicate that employees who participated in training were significantly more likely to be promoted, with our most robust model (IPTW) estimating a 287% increase in promotion odds. These findings provide strong empirical support for the effectiveness of Career 2030 in fostering career advancement at ACME Manufacturing.

Despite the strength of these findings, our analysis highlights the importance of continued refinement in both data collection and program evaluation. The presence of selection effects, where motivated employees or managerial influence may have impacted training participation, suggests that future iterations of Career 2030 should implement stricter randomization controls and collect additional data on employee motivation, managerial interventions, and long-term career progression. Furthermore, while our propensity score and weighting techniques effectively reduced bias, minor imbalances in specific covariates—such as distance from the training facility—highlight the need for further methodological enhancements in future analyses.

Moving forward, future research should explore several key areas to enhance both the evaluation and implementation of Career 2030. First, longitudinal tracking of trained employees over multiple years would allow ACME to measure the sustained impact of training on promotion and retention. Second, alternative weighting techniques such as entropy balancing or doubly robust estimation could further improve balance and refine causal estimates. Third, survey-based measures of motivation and career ambition could be integrated to better control for self-selection bias, ensuring that training effects are not overstated due to pre-existing employee characteristics.

Overall, this study confirms that Career 2030 is a valuable program with demonstrable benefits for employee development. By expanding access to training, refining evaluation methods, and ensuring continuous program monitoring, ACME can maximize the program's effectiveness and equity. The findings from this analysis serve as a foundation for data-driven decision-making, positioning Career 2030 as a critical driver of workforce advancement in ACME's long-term talent strategy.

# Appendix

## 1. Base Case Balance Table and Logistic Regression Results

```
                          Stratified by training
                           No             Yes            SMD
  n                         3709           2291
  manager = Yes (%)          485 (13.1)     379 (16.5)    0.098
  raise = Yes (%)           1540 (41.5)     664 (29.0)    0.265
  salary (%)                                              0.134
     $20-$40k                743 (20.0)     471 (20.6)
     $40-$80k                531 (14.3)     412 (18.0)
     > $80k                  269 ( 7.3)     204 ( 8.9)
     Under $20k             2166 (58.4)    1204 (52.6)
  children (mean (SD))      0.74 (0.49)    0.77 (0.50)    0.068
  mstatus (%)                                             0.133
     divorced                910 (24.5)     467 (20.4)
     married                 724 (19.5)     554 (24.2)
     single                 2075 (55.9)    1270 (55.4)
  age (mean (SD))          43.26 (11.96) 42.56 (10.80)   0.061
  sex = Male (%)            1998 (53.9)    1340 (58.5)    0.093
  edu (mean (SD))          11.51 (3.15)   11.80 (3.17)    0.093
  vacation (mean (SD))      3.18 (0.43)    3.22 (0.40)    0.081
  weight (mean (SD))        5.03 (0.27)    5.11 (0.25)    0.283
  height (mean (SD))       64.48 (6.65)   64.91 (6.85)    0.063
  hrfriend = Yes (%)        2001 (53.9)    1215 (53.0)    0.018
  cxofriend = Yes (%)       1985 (53.5)    1318 (57.5)    0.081
  insurance (%)                                           0.231
     Covered                1236 (33.3)     940 (41.0)
     Covered & Medicaid      640 (17.3)     393 (17.2)
     Covered & Medicare       27 ( 0.7)      10 ( 0.4)
     Medicaid               1517 (40.9)     789 (34.4)
     Medicare                 85 ( 2.3)      16 ( 0.7)
     Medicare & Medicaid      14 ( 0.4)       1 ( 0.0)
     Other                   190 ( 5.1)     142 ( 6.2)
  flexspend = Yes (%)       1057 (28.5)     957 (41.8)    0.281
  retcont = Yes (%)          594 (16.0)     131 ( 5.7)    0.335
  race (%)                                                0.041
     black                   621 (16.7)     352 (15.4)
     other                   221 ( 6.0)     148 ( 6.5)
     white                  2867 (77.3)    1791 (78.2)
  disthome (mean (SD))     25.62 (8.67)   15.69 (8.68)    1.145
  testscore (mean (SD))    66.21 (15.57) 56.57 (13.36)   0.665

Call:
glm(formula = (promoted == "Yes") ~ training, family = binomial(link = "logit"),
    data = data)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.54070    0.03405  15.881  < 2e-16 ***
trainingYes  0.21732    0.05629   3.861 0.000113 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 7764.4  on 5999  degrees of freedom
Residual deviance: 7749.4  on 5998  degrees of freedom
AIC: 7753.4

Number of Fisher Scoring iterations: 4

            exp(coef) [confint] p
(Intercept) "1.72 [1.61, 1.84]" "<0.001"
trainingYes "1.24 [1.11, 1.39]" "<0.001"
```

## 2. 1:1 Nearest Neighbors Matching (1.10 Caliper) Balance Table and Logistic Regression Results

```
                                         Stratified by training.Yes
                                          0             1             SMD
n                                           105           105
manager.Yes (mean (SD))                    0.08 (0.27)   0.08 (0.27) <0.001
raise.Yes (mean (SD))                      0.29 (0.45)   0.29 (0.45) <0.001
salary.$40-$80k (mean (SD))                0.08 (0.27)   0.08 (0.27) <0.001
salary.> $80k (mean (SD))                  0.04 (0.19)   0.04 (0.19) <0.001
salary.Under $20k (mean (SD))              0.82 (0.39)   0.82 (0.39) <0.001
children (mean (SD))                       0.82 (0.48)   0.86 (0.48)  0.070
mstatus.married (mean (SD))                0.20 (0.40)   0.20 (0.40) <0.001
mstatus.single (mean (SD))                 0.66 (0.48)   0.66 (0.48) <0.001
age (mean (SD))                           46.08 (9.51)  45.54 (9.32)  0.057
sex.Male (mean (SD))                       0.56 (0.50)   0.56 (0.50) <0.001
edu (mean (SD))                           11.08 (2.20)  11.13 (2.28)  0.026
vacation (mean (SD))                       3.13 (0.31)   3.13 (0.32)  0.030
weight (mean (SD))                         5.06 (0.21)   5.05 (0.19)  0.026
height (mean (SD))                        64.11 (5.79)  64.45 (6.12)  0.057
hrfriend.Yes (mean (SD))                   0.52 (0.50)   0.52 (0.50) <0.001
cxofriend.Yes (mean (SD))                  0.54 (0.50)   0.54 (0.50) <0.001
insurance.Covered & Medicaid (mean (SD))   0.22 (0.42)   0.22 (0.42) <0.001
insurance.Covered & Medicare (mean (SD))   0.00 (0.00)   0.00 (0.00) <0.001
insurance.Medicaid (mean (SD))             0.61 (0.49)   0.61 (0.49) <0.001
insurance.Medicare (mean (SD))             0.00 (0.00)   0.00 (0.00) <0.001
insurance.Medicare & Medicaid (mean (SD))  0.00 (0.00)   0.00 (0.00) <0.001
insurance.Other (mean (SD))                0.01 (0.10)   0.01 (0.10) <0.001
flexspend.Yes (mean (SD))                  0.44 (0.50)   0.44 (0.50) <0.001
retcont.Yes (mean (SD))                    0.04 (0.19)   0.04 (0.19) <0.001
race.other (mean (SD))                     0.00 (0.00)   0.00 (0.00) <0.001
race.white (mean (SD))                     0.92 (0.27)   0.92 (0.27) <0.001
disthome (mean (SD))                      20.93 (7.34)  19.01 (8.11)  0.249
testscore (mean (SD))                     62.23 (11.98) 60.61 (9.25)  0.151
```

```
Call:
glm(formula = (promoted.Yes == 1) ~ training.Yes, family = binomial(link = "logit"),
    data = matching_results$matched_data)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)    0.3267     0.1978   1.652  0.09860 .
training.Yes   0.9445     0.3079   3.068  0.00216 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 262.98  on 209  degrees of freedom
Residual deviance: 253.19  on 208  degrees of freedom
AIC: 257.19

Number of Fisher Scoring iterations: 4

             exp(coef) [confint] p
(Intercept)  "1.39 [0.94, 2.05]" " 0.099"
training.Yes "2.57 [1.42, 4.76]" " 0.002"
```

## 3. Propensity Score Matching Balance Table and Logistic Regression Results (0.002 Caliper)

```
                                          Stratified by training.Yes
                                           0             1             SMD
n                                           1297          1297
manager.Yes (mean (SD))                     0.15 (0.35)   0.15 (0.36)   0.017
raise.Yes (mean (SD))                       0.35 (0.48)   0.34 (0.47)   0.015
salary.$40-$80k (mean (SD))                 0.16 (0.37)   0.17 (0.37)   0.023
salary.> $80k (mean (SD))                   0.08 (0.27)   0.08 (0.27)  <0.001
salary.Under $20k (mean (SD))               0.56 (0.50)   0.55 (0.50)   0.033
children (mean (SD))                        0.76 (0.49)   0.76 (0.50)   0.019
mstatus.married (mean (SD))                 0.21 (0.41)   0.23 (0.42)   0.041
mstatus.single (mean (SD))                  0.57 (0.50)   0.55 (0.50)   0.033
age (mean (SD))                            42.66 (11.74) 42.45 (11.08)  0.018
sex.Male (mean (SD))                        0.56 (0.50)   0.57 (0.50)   0.008
edu (mean (SD))                            11.57 (3.13)  11.69 (3.16)   0.038
vacation (mean (SD))                        3.20 (0.45)   3.20 (0.40)   0.014
weight (mean (SD))                          5.07 (0.27)   5.08 (0.25)   0.033
height (mean (SD))                         64.75 (6.63)  64.82 (6.81)   0.010
hrfriend.Yes (mean (SD))                    0.52 (0.50)   0.53 (0.50)   0.011
cxofriend.Yes (mean (SD))                   0.57 (0.50)   0.57 (0.49)   0.006
insurance.Covered & Medicaid (mean (SD))    0.18 (0.38)   0.17 (0.38)   0.004
insurance.Covered & Medicare (mean (SD))    0.01 (0.09)   0.01 (0.07)   0.037
insurance.Medicaid (mean (SD))              0.36 (0.48)   0.37 (0.48)   0.014
insurance.Medicare (mean (SD))              0.01 (0.11)   0.01 (0.10)   0.015
insurance.Medicare & Medicaid (mean (SD))   0.00 (0.03)   0.00 (0.03)  <0.001
insurance.Other (mean (SD))                 0.06 (0.23)   0.06 (0.24)   0.013
flexspend.Yes (mean (SD))                   0.37 (0.48)   0.36 (0.48)   0.011
retcont.Yes (mean (SD))                     0.09 (0.28)   0.09 (0.28)   0.008
race.other (mean (SD))                      0.06 (0.24)   0.06 (0.24)   0.013
race.white (mean (SD))                      0.79 (0.41)   0.77 (0.42)   0.032
disthome (mean (SD))                       19.84 (6.89)  20.16 (7.23)   0.046
testscore (mean (SD))                      60.33 (15.43) 59.87 (12.14)  0.033
propensity_score (mean (SD))                0.44 (0.19)   0.44 (0.19)  <0.001
```

```
Call:
glm(formula = promoted.Yes == 1 ~ training.Yes, family = binomial(link = "logit"),
    data = matching_results$matched_data)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)   0.09722    0.05560   1.749   0.0804 .
training.Yes  0.94562    0.08422  11.228   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3413.5  on 2593  degrees of freedom
Residual deviance: 3283.1  on 2592  degrees of freedom
AIC: 3287.1

Number of Fisher Scoring iterations: 4

             exp(coef) [confint] p
(Intercept)  "1.10 [0.99, 1.23]" " 0.080"
training.Yes "2.57 [2.18, 3.04]" "<0.001"
```

## 4. Propensity Score Matching Balance Table and Logistic Regression Results (0.31 Caliper)

| Stratified by training.Yes | 0 | 1 | SMD |
|---|---|---|---|
| n | 1503 | 1503 | |
| manager.Yes (mean (SD)) | 0.15 (0.35) | 0.15 (0.36) | 0.011 |
| raise.Yes (mean (SD)) | 0.33 (0.47) | 0.33 (0.47) | 0.007 |
| salary.$40-$80k (mean (SD)) | 0.16 (0.37) | 0.16 (0.37) | 0.004 |
| salary.> $80k (mean (SD)) | 0.08 (0.27) | 0.09 (0.28) | 0.027 |
| salary.Under $20k (mean (SD)) | 0.55 (0.50) | 0.55 (0.50) | 0.009 |
| children (mean (SD)) | 0.76 (0.49) | 0.77 (0.50) | 0.012 |
| mstatus.married (mean (SD)) | 0.22 (0.41) | 0.23 (0.42) | 0.026 |
| mstatus.single (mean (SD)) | 0.55 (0.50) | 0.55 (0.50) | 0.001 |
| age (mean (SD)) | 42.63 (11.72) | 42.53 (11.05) | 0.009 |
| sex.Male (mean (SD)) | 0.57 (0.50) | 0.56 (0.50) | 0.015 |
| edu (mean (SD)) | 11.69 (3.11) | 11.78 (3.16) | 0.028 |
| vacation (mean (SD)) | 3.20 (0.45) | 3.20 (0.40) | 0.007 |
| weight (mean (SD)) | 5.07 (0.27) | 5.08 (0.25) | 0.018 |
| height (mean (SD)) | 64.79 (6.73) | 64.86 (6.93) | 0.009 |
| hrfriend.Yes (mean (SD)) | 0.54 (0.50) | 0.52 (0.50) | 0.035 |
| cxofriend.Yes (mean (SD)) | 0.56 (0.50) | 0.57 (0.49) | 0.019 |
| insurance.Covered & Medicaid (mean (SD)) | 0.18 (0.39) | 0.18 (0.39) | 0.002 |
| insurance.Covered & Medicare (mean (SD)) | 0.01 (0.08) | 0.00 (0.07) | 0.027 |
| insurance.Medicaid (mean (SD)) | 0.38 (0.49) | 0.36 (0.48) | 0.033 |
| insurance.Medicare (mean (SD)) | 0.01 (0.10) | 0.01 (0.10) | 0.007 |
| insurance.Medicare & Medicaid (mean (SD)) | 0.00 (0.00) | 0.00 (0.03) | 0.036 |
| insurance.Other (mean (SD)) | 0.05 (0.22) | 0.05 (0.23) | 0.006 |
| flexspend.Yes (mean (SD)) | 0.36 (0.48) | 0.38 (0.48) | 0.025 |
| retcont.Yes (mean (SD)) | 0.09 (0.29) | 0.08 (0.26) | 0.053 |
| race.other (mean (SD)) | 0.07 (0.25) | 0.06 (0.24) | 0.016 |
| race.white (mean (SD)) | 0.78 (0.42) | 0.78 (0.42) | 0.002 |
| disthome (mean (SD)) | 19.56 (6.68) | 19.52 (7.42) | 0.005 |
| testscore (mean (SD)) | 60.70 (15.43) | 59.49 (12.29) | 0.086 |
| propensity_score (mean (SD)) | 0.45 (0.19) | 0.47 (0.20) | 0.097 |

```
Call:
glm(formula = promoted.Yes == 1 ~ training.Yes, family = binomial(link = "logit"),
    data = matching_results$matched_data)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.16672    0.05177   3.221  0.00128 **
training.Yes 0.85101    0.07805  10.903  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3935.0  on 3005  degrees of freedom
Residual deviance: 3812.8  on 3004  degrees of freedom
AIC: 3816.8

Number of Fisher Scoring iterations: 4

             exp(coef) [confint] p
(Intercept)  "1.18 [1.07, 1.31]" " 0.001"
training.Yes "2.34 [2.01, 2.73]" "<0.001"
```

## 5. Inverse Probability and Treatment Weighting Balance Table and Logistic Regression Results

| Stratified by training | No | Yes | SMD |
|---|---|---|---|
| n | 3465.1 | 1921.5 | |
| manager = Yes (%) | 486.1 (14.0) | 270.2 (14.1) | 0.001 |
| raise = Yes (%) | 1311.7 (37.9) | 631.1 (32.8) | 0.105 |
| salary (%) | | | 0.031 |
| $20-$40k | 698.6 (20.2) | 407.8 (21.2) | |
| $40-$80k | 534.6 (15.4) | 290.8 (15.1) | |
| > $80k | 265.3 ( 7.7) | 153.2 ( 8.0) | |
| Under $20k | 1966.7 (56.8) | 1069.7 (55.7) | |
| children (mean (SD)) | 0.75 (0.49) | 0.75 (0.50) | 0.011 |
| mstatus (%) | | | 0.035 |
| divorced | 812.2 (23.4) | 426.3 (22.2) | |
| married | 718.9 (20.7) | 418.8 (21.8) | |
| single | 1934.0 (55.8) | 1076.4 (56.0) | |
| age (mean (SD)) | 43.00 (11.80) | 42.84 (11.02) | 0.014 |
| sex = Male (%) | 1899.4 (54.8) | 1078.2 (56.1) | 0.026 |
| edu (mean (SD)) | 11.60 (3.13) | 11.62 (3.13) | 0.007 |
| vacation (mean (SD)) | 3.19 (0.44) | 3.20 (0.39) | 0.030 |
| weight (mean (SD)) | 5.05 (0.27) | 5.08 (0.25) | 0.100 |
| height (mean (SD)) | 64.67 (6.73) | 64.66 (6.81) | 0.001 |
| hrfriend = Yes (%) | 1834.3 (52.9) | 1012.4 (52.7) | 0.005 |
| cxofriend = Yes (%) | 1915.0 (55.3) | 1103.6 (57.4) | 0.044 |
| insurance (%) | | | 0.087 |
| Covered | 1235.9 (35.7) | 712.2 (37.1) | |
| Covered & Medicaid | 585.3 (16.9) | 344.0 (17.9) | |
| Covered & Medicare | 23.3 ( 0.7) | 11.2 ( 0.6) | |
| Medicaid | 1363.2 (39.3) | 723.5 (37.7) | |
| Medicare | 62.6 ( 1.8) | 20.0 ( 1.0) | |
| Medicare & Medicaid | 9.2 ( 0.3) | 2.1 ( 0.1) | |
| Other | 185.6 ( 5.4) | 108.4 ( 5.6) | |
| flexspend = Yes (%) | 1127.3 (32.5) | 710.9 (37.0) | 0.094 |
| retcont = Yes (%) | 443.8 (12.8) | 175.2 ( 9.1) | 0.118 |
| race (%) | | | 0.017 |
| black | 558.0 (16.1) | 302.9 (15.8) | |
| other | 214.0 ( 6.2) | 112.8 ( 5.9) | |
| white | 2693.0 (77.7) | 1505.8 (78.4) | |
| disthome (mean (SD)) | 23.18 (8.73) | 19.55 (8.35) | 0.425 |
| testscore (mean (SD)) | 63.39 (15.87) | 59.56 (12.49) | 0.268 |

```
Call:
svyglm(formula = promoted ~ training + disthome + testscore,
    design = tpdSvy, family = binomial(link = "logit"))

Survey design:
svydesign(ids = ~1, data = tpd, weights = ~iptw_weight)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -6.761032   0.233508 -28.954   <2e-16 ***
trainingYes  1.352638   0.091722  14.747   <2e-16 ***
disthome    -0.001059   0.004325  -0.245    0.806
testscore    0.114621   0.003201  35.804   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 0.9737424)

Number of Fisher Scoring iterations: 5

Odds Ratio for Training Impact on Promotion (IPTW): 3.867614
Confidence Interval for Training Impact on Promotion (IPTW): 1.172828 1.532447
```

**6. Wilcoxon Signed-Rank Test Results:**

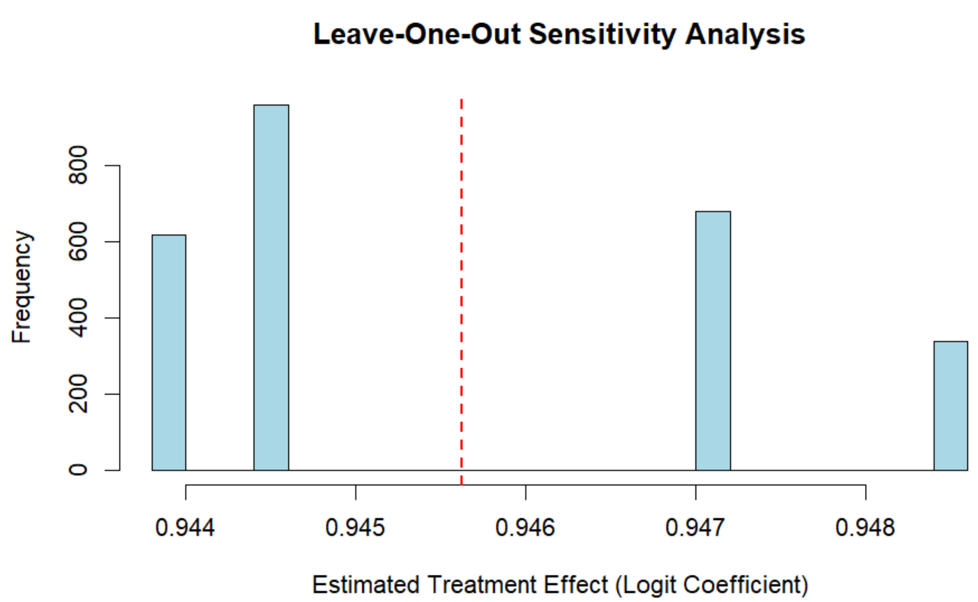
```
	Wilcoxon signed rank test with continuity correction

data:  treated and control
V = 139190, p-value < 2.2e-16
alternative hypothesis: true location shift is not equal to 0
```

V = the sum of differences between treated and control groups, which allows us to measure the difference in distributions between the groups.

Our p-value is extremely small, indicating strong evidence against the null hypothesis, meaning our observed difference is statistically significant. In other words, the treatment (training.Yes = 1) likely had a significant effect on the outcome (promoted.Yes).

**7. Leave-One-Out Sensitivity Analysis Results:**



# References


1. Laerd Statistics. (n.d.). Wilcoxon signed-rank test using SPSS Statistics. https://statistics.laerd.com/spss-tutorials/wilcoxon-signed-rank-test-using-spss-statistics.php
2. Jesussek, M. (n.d.). Wilcoxon signed-rank test. DATAtab. https://datatab.net/tutorial/wilcoxon-test
3. StataCorp LLC. (n.d.). Leave-one-out meta-analysis. Stata. https://www.stata.com/stata17/leave-one-out-meta-analysis/
4. Department of Statistics, The Pennsylvania State University. (n.d.). 16.8 - Random effects / sensitivity analysis. In STAT 509: Design and analysis of clinical trials. Retrieved March 9, 2025, from https://online.stat.psu.edu/stat509/lesson/16/16.8