

# NYTD Modeling: Visual Data Report



A Statistical Analysis of Foster Care Youth Outcomes in Tennessee

## Table of Contents

Introduction	.....
Summary Dashboard	.....
Executive Summary	.....
Data Curation & Methodology	.....
Data Overview	.....
Significant Variables	.....
Significant Model Results	.....
Descriptive Statistics	.....
Outcome Distributions	.....
Variable Importance & Effect Sizes	.....
Appendix: Variable Definitions	.....
Appendix: All Tables	.....
Appendix: Binary Classification Model Matrix	.....
Appendix: Model Details	.....
Appendix: Charts and Visualizations	.....
Appendix: Data Breakdown	.....
Appendix: Source Files, Model Creation & References	.....

## Note on Navigating This Report

Click on any section in the **Table of Contents** to jump directly to that analysis

Each section contains **Key Findings** highlighted in blue boxes for easy scanning

All charts and visualizations are referenced in the **Appendix** for detailed viewing

Statistical tables provide the complete numerical details behind each finding



## Introduction

This report presents a **comprehensive statistical analysis** of National Youth in Transition Data (NYTD), focusing on outcomes for youth in foster care. The analyses include:

- ▶ Identification of significant predictors for key outcomes
- ▶ Visualization of outcome and predictor distributions
- ▶ Summary statistics and data quality assessments
- ▶ Model effect sizes and variable importance

### Things You Need To Know:

- ▶ **Outcome variables** reflect key areas of life, such as education, housing, and incarceration, for youth (ages 18-21) who have exited foster care.
  - ▶ *Note: all data used is from participants 1-4 years after leaving foster care.*
- ▶ **Predictor variables** are the factors that may influence those outcomes (like placement history or service access).
  - ▶ *Note: all data is from participants while in care or directly before leaving care.*

### Data Limitations:

- ▶ **Small sample size** of 720 youth; final dataset reduced to 599 rows due to AFCARS data merging issues.
- ▶ Reliance on voluntary participation introduces **selection bias**.
- ▶ Inconsistent longitudinal follow-up across waves

See more about these variables in:

[Variable Names and Descriptions](#) | [Variable Explanations](#)

[Return to Table of Contents](#)



## Summary Dashboard

Supporting Data: See [Appendix: All Tables](#) for complete statistical results, [Appendix: Variable Definitions](#) for variable descriptions, and [Appendix: Charts and Visualizations](#) for supporting graphs.

### Critical Risk Factors

RISK FACTOR	IMPACT / STAT	OUTCOME VARIABLES
Placement Instability	Higher Incarceration, Higher levels of Homelessness, Higher levels of Substance Abuse, Lower levels of Education enrollment	Negative: Incarc_w23, Homeless_w23, SubAbuse_w23 Positive: Connected_Youth
Group Home Placement	2 times more vulnerable to incarceration	Negative: Incarc_w23
Disability	Higher level of Homelessness, Less Education/Employment	Negative: Homeless_w23, Connected_Youth
Sex (Male)	30% more likely to be incarcerated	Negative: Incarc_w23

**1 in 3** youth incarcerated

**1 in 3** youth homeless

**1 in 5** substance abuse referrals

### Positive: Protective Factors & Success Stories

PROTECTIVE FACTOR	IMPACT / STAT	OUTCOME VARIABLES
Early Adult Connections	94.3% maintain supportive adults	Positive: CnctAdult_w23
Foster Home Placement	Higher outcomes compared to any other placement	Positive: Multiple outcomes
Longer Time in Care	Lower Incarceration, Higher Positive Outcomes	Positive: Incarc_w23 (reduced), Connected_Youth
Service Access	More Education, More Higher Education Enrollment	Positive: Connected_Youth, CurrenRoll_w23

**94.3%** maintain adult connections

**64.8%** in education/employment

**81.5%** avoid substance abuse

[Return to Table of Contents](#)



# Executive Summary

## Supporting Data:

See [Appendix: All Tables](#) for complete statistical results, [Appendix: Variable Definitions](#) for variable descriptions, and [Appendix: Charts and Visualizations](#) for supporting graphs.

## Key Findings

- Within the Tennessee Dataset, 19 important connections were identified between what happens to youth in foster care and their life as adults.

## Key Analytical Decisions

- Outcome Focus:** Concentrated on post-transition outcomes (ages 18-21) to capture long-term effects
- Predictor Selection:** Based on prior academic research and factors predicted to impact outcomes
- Statistical Significance:** Applied  $p < 0.05$  threshold for identifying meaningful relationships

## Limitations

- Analysis is limited by a sample size of 720
- The final dataset used for modeling was consisted of only 3 cohorts and 599 rows due to AFCARS data
  - Note:** While the NYTD Outcomes dataset had data for four cohorts (2011-2024), AFCARS data from the 2011 cohort could not be merged into the final dataset due to inconsistent participant IDs. The AFCARS dataset has 599 participants compared to the 720 in the NYTD Outcomes dataset.
- Reliance on voluntarily participating youth introduces selection bias
- Inconsistent longitudinal follow-up across waves, affects the generalizability and completeness of the findings.

[Return to Table of Contents](#)



## Most Impactful Predictors

### 1. Placement Stability

- ▶ **Key Finding:** Frequent moves between homes is the most universal predictor of poor outcomes
- ▶ **Impact:** Affects education, criminal justice, housing, and substance use outcomes

### 2. Sex Differences

- ▶ **Key Finding:** Females are significantly less likely to be incarcerated
- ▶ **Impact:** Sex has strong protective effects across multiple domains

### 3. Service Access

- ▶ **Key Finding:** Number of services received strongly predicts positive outcomes
- ▶ **Impact:** More services increase educational aid access and current enrollment

### 4. Disability Status

- ▶ **Key Finding:** Youth with disabilities face multiple barriers but have greater service needs
- ▶ **Impact:** Affects homelessness, service access, enrollment, and eligibility

[Return to Table of Contents](#)



## Most Impactful Findings

### Critical Risk Factors:

- ▶ Placement instability is the most universal predictor of poor outcomes
- ▶ Youth with disabilities are more likely to experience homelessness and less likely to be enrolled in education or employment
- ▶ Youth placed in foster homes show better outcomes than group homes
- ▶ Early supportive relationships have lasting positive effects

### Demographics Most in Need of Support:

- ▶ Youth with disabilities are in multiple high risk domains
- ▶ Males are ~30% more likely to be incarcerated than females.
- ▶ Youth in Group home placements are in multiple high risk domains

### Positive Post-Transition Outcomes

- ✓ 94.3% of youth maintain supportive adult connections
- ✓ 64.8% of youth pursue higher education or full time employment
- ✓ Early supportive relationships have the strongest positive effects

### Critical Challenges

- ⚠ 1 in 3 youth experience incarceration or homelessness
- ⚠ Placement instability affects nearly every outcome measured
- ⚠ Homelessness and lack of education are high, particularly for youth with disabilities

[Return to Table of Contents](#)

# Data Curation & Methodology



 **Detailed Methods:** See [Appendix: Binary Classification Model Matrix](#) for model coverage details and [Appendix: Source Files](#) for methodology references.

 **Detailed Model:** See [Appendix: Model Details](#) for technical specifications, [Appendix: Binary Classification Matrix](#) for model coverage tables and [Appendix: Variable Definitions](#) for variable descriptions

## Data Processing & Integration

The datasets were carefully merged and processed to create a unified analysis framework:

### Data Integration Process:

- ▶ **Matching Strategy:** Used unique identifiers to link youth across all three datasets
- ▶ **Temporal Alignment:** Aligned data collection periods to create consistent timeframes
- ▶ **Variable Creation:** Developed composite measures (e.g., placement instability, service counts)
- ▶ **Final Dataset:** Created comprehensive dataset [NYTD outcomes, NYTD Services, AFCARS] with 599 youth and complete outcome/predictor information.

### Data Sources and Cohorts:

**Note:** See the [univariate.pdf](#) and [bivariate.pdf](#) for more information. You can also see [Appendix: Binary Classification Model Matrix](#)

- ▶ **NYTD Outcome Survey:** Four cohort groups (2011-2015, 2014-2018, 2017-2021, 2020-2024) with longitudinal follow-ups at ages 17, 19, and 21
- ▶ **NYTD Services:** Services that participants utilized under the age of 18 were summed and matched using a unique ID
- ▶ **AFCARS:** Three matchable Cohort groups (2014-2018, 2017-2021, 2020-2024) with case level information about youth adopted & in foster care

[Return to Table of Contents](#)

## 🔧 Data Cleaning and Preparation:



### ▶ Dataset Creation:

- ▶ Only youth identified as foster care children in Tennessee were included.
- ▶ Outcome, Services and AFCARS data was merged into a single dataset based on a unique personal identifier and cohort year.
- ▶ For each variable, if the participant experienced it, they are assigned a 1; if they did not, they are assigned a 0. If the participant chose not to answer or could not be found, the response is left blank.

### ▶ Variable Standardization:

- ▶ The sum of NYTD Services used by each individual was summed into a column [services\_amount] due to lack of data for each individual service
  - ▶ **Note:** services captured include: [Special Education, Independent Living Needs Assessment, Academic Support, Post-Secondary Educational Support, Career Preparation, Employment Programs or Vocational Training, Budget and Financial Management, Housing Education and Home Management Training, Health Education and Risk Prevention, Family Support and Healthy Marriage Education, Mentoring, Supervised Independent Living, Room and Board Financial Assistance, Education Financial Assistance, Other Financial Assistance]

- ▶ The NYTD outcomes data set was organized into pre-transition and post-transition
  - ▶ Pre-transition: Youth answered the survey while in fostercare.
  - ▶ Post-transition: If youth responded in waves 2 and/or 3, they were included.

### ▶ Missing Data Handling:

- ▶ Applied appropriate imputation strategies for key variables.

### ▶ Data Validation:

- ▶ Unique identifiers were cross-checked across the final dataset and sources to ensure accuracy.
- ▶ The accuracy of final calculations was hand-checked by sampling.

[Return to Table of Contents](#)

## Modeling Approach



### Logistic Regression & Binary Classification Framework

#### Model Type:

- ▶ **Logistic Regression:** A statistical method used to model the probability of the outcome occurring *(for each variable, 1 means it did happen and 0 means it did not)*
- ▶ **Function:** Uses the logistic function to transform linear combinations of predictors into probabilities between 0 and 1 *(this is the average of all participants' answers)*
- ▶ **Output:** Provides odds ratios and coefficients that indicate the strength and direction of relationships between predictors and outcomes
- ▶ **Advantages:** Interpretable results, handles multiple predictors simultaneously, and provides confidence intervals for effect estimates

#### Outcome Classification:

- ▶ **Binary Classification:** All outcome variables were treated as binary classifications of whether an outcome occurred or not ([Click here to see tables](#))
- ▶ **Interpretation:** This approach allows us to model the probability of each outcome occurring based on predictor variables

## Variable Selection

#### Predictor Variables:

[Click here to see more information](#)

- ▶ Selected based on responses provided by youth prior to exiting foster care
- ▶ All variables were grouped into specific domains (demographics, placement, services, etc.)
- ▶ Selection aimed to include variables that reflect known risk and protective factors
- ▶ Only variables that were consistently available across all cohort years and waves were considered

#### Outcome Variables:

[Click here to see more information](#)

- ▶ Selected from post-transition waves of the NYTD (ages 18-21)
- ▶ Represent a combination of positive and negative outcomes
- ▶ Outcomes were modeled separately to enable analysis of both successful transitions and areas of vulnerability
- ▶ Only variables that had sufficient response rates and were consistently reported across all cohorts were included

#### Control Variables:

- ▶ Included to account for baseline differences among youth that could influence outcomes
- ▶ Controls help ensure that observed effects are not confounded by core demographic characteristics

[Return to Table of Contents](#)



## Data Overview

Primary Data Sources:

- ▶ **NYTD Outcome Survey**: Longitudinal data on youth outcomes at ages 17, 19, and 21
- ▶ **NYTD Services**: Detailed service utilization records for youth in care
- ▶ **AFCARS** (Adoption and Foster Care Analysis and Reporting System): Administrative data on placement history and demographics
- ▶  **Final Analysis Dataset:** `wave_service_afcars_final.csv` (720 youth with complete data)
- ▶  **Model Results:** `significant_model_results.csv`
- ▶  **Statistical Summaries:** `outcome_stats_detailed_df.csv` and `predictor_stats_detailed_df.csv`

[Return to Table of Contents](#)



## Significant Variables: Key Findings Summary

- 🔍 **Big Picture:** Our analysis identified **19** statistically significant relationships between foster care experiences and young adult outcomes.
- 📋 **Detailed Tables:** See [Appendix: All Tables](#) for complete statistical results and [Appendix: Variable Definitions](#) for variable descriptions.

### Most Important Predictors

PREDICTOR	IMPACT	OUTCOMES AFFECTED
Placement Instability	Frequent moves between homes	Connected_Youth, Homeless_w23, Incarc_w23, SubAbuse_w23
Sex	Strong effects on incarceration rates	Connected_Youth, Incarc_w23
Time in Care (MonthsInCare)	Months spent in foster care	Connected_Youth, Incarc_w23
Having a Disability (DiagDis)	Increases risk for negative outcomes	Connected_Youth, Homeless_w23

### Statistically Significant Predictors

#### 🎓 Education & Employment Success:

##### Success

- ▶ Youth with more stable placements
- ▶ Females
- ▶ Those with longer time in care

##### Struggle

- ▶ Having a disability

[Return to Table of Contents](#)



## Criminal Justice Involvement:

### **Success**

- ▶ Females are significantly less likely to be incarcerated

### **Struggle**

- ▶ Youth in group homes have higher incarceration rates
- ▶ More placement moves increase incarceration risk

## Housing & Basic Needs:

### **Struggle**

- ▶ Youth with disabilities are more likely to experience homelessness
- ▶ More placement instability increases homelessness risk

## Substance Use:

### **Success**

- ▶ Longer time in care slightly reduces this risk

### **Struggle**

- ▶ White youth are more likely to receive substance abuse referrals
- ▶ Those with more placement moves are more likely to receive substance abuse referrals

## Adult Connections:

### **Success**

- ▶ Youth who had supportive adults while in care are much more likely to maintain these connections as young adults

## What This Means

**Key Insight:** These findings suggest that **placement stability** and **early supportive relationships** are crucial for positive outcomes. Youth with disabilities need extra support, and the type of placement (foster home vs. group home) matters significantly for future success.

[Return to Table of Contents](#)



## 🔍 Significant Model Results: Key Findings

**Analysis Summary:** Our comprehensive analysis revealed key insights about the relationships between foster care experiences and outcomes:

**Total Significant Effects:** 19 statistically significant relationships found between predictors and outcomes

**Outcomes Analyzed:** 6 different outcome measures studied across multiple domains

**Unique Variables:** 9 distinct predictors with significant effects on youth outcomes

**Detailed Results:** See [Appendix: All Tables](#) for complete statistical tables and [Appendix: Model Details](#) for methodology.

[Return to Table of Contents](#)



## Strongest Effects (by absolute coefficient)

**Key Insight:** The strongest effects in our analysis show which factors have the biggest impact on outcomes:

### Top 5 Most Important Findings

**Understanding Odds Ratios:** An odds ratio tells us how much more or less likely an outcome is to occur. An odds ratio of 2.0 means the outcome is twice as likely, while 0.5 means it's half as likely. Values above 1.0 indicate increased odds, while values below 1.0 indicate decreased odds.

RANK	WHAT WE FOUND	ODDS RATIO	WHAT THIS MEANS
1st	Placement Instability and Incarceration	13.9	More placement moves dramatically increase incarceration risk
2nd	Placement Instability and Substance Abuse	12.2	More placement moves significantly increase substance abuse risk
3rd	Placement Instability and Homelessness	4.4	More placement moves increase homelessness risk
4th	Placement Instability and Education/Employment	0.14	More placement moves decrease education/employment success
5th	Substance Abuse Prevention	0.04	Youth in foster care are much less likely to have substance abuse problems

**Key Insight:** The strongest effects are overwhelmingly related to **placement instability**, suggesting that the number of placement moves is the most critical factor affecting youth outcomes. This has major implications for policy and practice.

[Return to Table of Contents](#)



## ⌚ Most Significant Effects (by p-value)

**Statistical Confidence:** Within this dataset, these effects have the strongest statistical evidence, meaning we're most confident these relationships are real:

## ☒ Top 5 Most Reliable Predictions

RANK	WHAT WE CAN PREDICT	HOW SURE WE ARE	WHAT THIS MEANS
1st	Sex and Incarceration Risk	99.9996% sure	Females are much less likely to be incarcerated than males
2nd	Group Home and Incarceration Risk	99.99% sure	WYouth in group homes are more likely to be incarcerated
3rd	Other Placements and Incarceration Risk	99.93% sure	Youth in other placement types are more likely to be incarcerated
4th	Frequent Moves and Incarceration Risk	99.89% sure	More placement moves increase incarceration risk
5th	Time in Care and Incarceration Risk	99.85% sure	Longer time in care reduces incarceration risk

💡 **Key Insight:** All the most statistically significant effects relate to **incarceration risk**, with **sex and placement factors** being the most reliable predictors. This suggests that criminal justice involvement is the outcome we can predict most confidently.

[Return to Table of Contents](#)



## 📋 Descriptive Statistics

### 📋 Detailed Data:

See [Appendix: Data Breakdown](#) for comprehensive data quality metrics and [Appendix: Charts and Visualizations](#) for supporting graphs

## Outcome Distributions

### Frequency Breakdown:

#### 🏡 Housing & Family Outcomes:

OUTCOME	PERCENTAGE	COUNT	INTERPRETATION
Having Children Without Marriage [OutcmFCS_w23]	22.4%	159 out of 709	About 1 in 5 youth had children while unmarried
Experiencing Homelessness [Homeless_w23]	29.3%	211 out of 719	Nearly 1 in 3 youth experienced homelessness

#### 🎓 Education & Employment:

OUTCOME	PERCENTAGE	COUNT	INTERPRETATION
Connected to Education/Employment [Connected_Youth]	64.8%	464 out of 716	Nearly two-thirds pursuing education or employment
Supportive Adult Connections [CnctAdult_w23]	94.3%	678 out of 719	Most youth maintain adult connections

[Return to Table of Contents](#)

**Criminal Justice & Health:**

OUTCOME	PERCENTAGE	COUNT	INTERPRETATION
Incarceration [Incarc_w23]	35.7%	256 out of 718	About 1 in 3 youth were incarcerated
Substance Abuse Referrals [SubAbuse_w23]	18.5%	133 out of 717	About 1 in 6 youth received substance abuse referrals

**💡 Key Insights:**

**Most youth maintain adult connections** (94.3%) - a positive finding

**Nearly two-thirds are connected to education/employment** (64.8%)

**Significant challenges remain:** About 1 in 3 youth experience incarceration or homelessness

**Data quality is excellent:** Missing data is very low (under 2% for all outcomes)

[Return to Table of Contents](#)

## Appendix



### All Data, Visuals and Model Outcomes

## Variable Definitions

### Related Findings:

These variable definitions support the analysis in [Significant Variables](#), [Descriptive Statistics](#), and [Summary Dashboard](#).

### Outcome Variables Overview:

OUTCOME DOMAINS	OUTCOME VARIABLES
Education & Employment	Pursuing higher education or full-time employment [Connected_Youth]
Social Connections	Having a supportive adult connection [CnctAdult_w23]
Housing Stability	Experiencing homelessness [Homeless_w23]
Health & Wellness	Receiving a substance abuse referral [SubAbuse_w23]
Justice System	Being incarcerated [Incarc_w23]
Family Formation	Having children without being married [Children_NoMarriage_w23]

### Preictor Variables Overview:

PREDICTOR DOMAIN	PREDICTOR VARIABLES
Core Demographics	Being either male or female [Sex]
Disability Status	Being diagnosed with a disability [DiagDis]
Removal Reasons	Why they were removed from their home. Grouped neglect, abuse or any other reason [RR_neglect], [RR_abuse], [RR_other]
Placement Factors	Where they were placed (group home, foster home or other), how many different places they were in and how long they were in the foster care system [PlacementType], [PlacementInstability], [MonthsInCare]
Early Indicators	If they were referred for substance abuse help. If they have a supportive adult relationship [SubAbuse_w1], [CnctAdult_w1]

## Appendix: All Tables



**Related Findings:** These tables support the findings in [Significant Variables](#), [Significant Model Results](#), and [Summary Dashboard](#).

**Note:** In a logistic regression 'Coefficient' represents the change in the log-odds of the outcome for a one-unit increase in the predictor variable.

**Understanding Log-Odds:** Log-odds is the natural logarithm of the odds ratio. It's used in logistic regression because it can range from negative infinity to positive infinity, making it suitable for linear modeling. The odds ratio can be calculated as  $e^{\text{coefficient}}$ .

[Return to Table of Contents](#)

## Significant Variables From Paper Full Table



OUTCOME	VARIABLE	COEFFICIENT	P-VALUE	CI LOWER	CI UPPER
Children_NoMarriage_w23	const	-2.408	0.0113	-4.272	-0.544
Children_NoMarriage_w23	White_Only	-0.543	0.0268	-1.023	-0.062
CnctAdult_w23	CnctAdult_w1	1.336	0.0324	0.112	2.560
CnctAdult_w23	const	2.407	0.0477	0.024	4.790
Connected_Youth	PlacementInstability	-1.980	0.0082	-3.448	-0.511
Connected_Youth	Sex	0.518	0.0221	0.074	0.962
Connected_Youth	DiagDis	-0.490	0.0228	-0.911	-0.068
Connected_Youth	MonthsInCare	0.019	0.0312	0.002	0.036
Homeless_w23	DiagDis	0.546	0.0116	0.122	0.970
Homeless_w23	PlacementInstability	1.489	0.0433	0.045	2.933
Incarc_w23	Sex	-1.157	0.0000	-1.650	-0.664
Incarc_w23	PlacementType_3	1.116	0.0001	0.553	1.680
Incarc_w23	PlacementType_2	1.074	0.0007	0.451	1.697
Incarc_w23	PlacementInstability	2.634	0.0011	1.054	4.214
Incarc_w23	MonthsInCare	-0.031	0.0015	-0.051	-0.012
SubAbuse_w23	const	-3.237	0.0039	-5.435	-1.039
SubAbuse_w23	White_Only	1.006	0.0045	0.312	1.700
SubAbuse_w23	PlacementInstability	2.498	0.0108	0.577	4.419
SubAbuse_w23	MonthsInCare	-0.034	0.0109	-0.061	-0.008

## ALL Significant Variables Full Table

OUTCOME	VARIABLE	COEFFICIENT	P-VALUE	CI LOWER	CI UPPER
OutcmFCS_w23	DiagDis	-1.16	0.0010	-1.85	-0.47
OutcmFCS_w23	Service_Count	0.83	0.0023	0.30	1.37
EducAid_w23	Service_Count	0.54	0.0206	0.08	0.99
PubFoodAs_w23	const	-3.57	0.0124	-6.36	-0.77
PubFoodAs_w23	Sex	1.77	0.0002	0.84	2.71
CurrenRoll_w23	DiagDis	-0.92	0.0083	-1.60	-0.24
CurrenRoll_w23	Service_Count	0.69	0.0130	0.15	1.24
Homeless_w23	DiagDis	0.85	0.0254	0.10	1.60
Homeless_w23	RR_other	-1.72	0.0162	-3.13	-0.32
OthrHlthln_w23	const	-3.55	0.0298	-6.76	-0.35
Elig19_w23	const	3.78	0.0256	0.46	7.10
Elig19_w23	DiagDis	1.34	0.0048	0.41	2.27
Elig19_w23	RR_other	-2.26	0.0484	-4.51	-0.02
Responded_w23	const	3.78	0.0256	0.46	7.10
Responded_w23	DiagDis	1.34	0.0048	0.41	2.27
Responded_w23	RR_other	-2.26	0.0484	-4.51	-0.02
HS_or_GED_w23	Sex	0.75	0.0284	0.08	1.41

[Return to Table of Contents](#)

## Strongest Effects (by absolute coefficient) Full Table



VARIABLE	OUTCOME	COEFFICIENT	P-VALUE	EFFECT DESCRIPTION
PlacementInstability	Incarc_w23	2.634	0.0011	Increases Incarc_w23 by 2.63 units
PlacementInstability	SubAbuse_w23	2.498	0.0108	Increases SubAbuse_w23 by 2.50 units
PlacementInstability	Homeless_w23	1.489	0.0433	Increases Homeless_w23 by 1.49 units
PlacementInstability	Connected_Youth	-1.980	0.0082	Decreases Connected_Youth by 1.98 units
const	SubAbuse_w23	-3.237	0.0039	Decreases SubAbuse_w23 by 3.24 units

## Most Significant Effects (by p-value) Full Table

VARIABLE	OUTCOME	COEFFICIENT	95% CI	P-VALUE
Sex	Incarc_w23	-1.157	[-1.650, -0.664]	0.000004
PlacementType_3	Incarc_w23	1.116	[0.553, 1.680]	0.000104
PlacementType_2	Incarc_w23	1.074	[0.451, 1.697]	0.000730
PlacementInstability	Incarc_w23	0.098	[0.039, 0.156]	0.001085
MonthsInCare	Incarc_w23	-0.031	[-0.051, -0.012]	0.001542

## Frequency Breakdown Full Table

OUTCOME	0 (NO)	1 (YES)	VALID N	MISSING
Children_NoMarriage_w23	550	159	709	11
CnctAdult_w23	41	678	719	1
Connected_Youth	252	464	716	4
Homeless_w23	508	211	719	1
Incarc_w23	462	256	718	2
SubAbuse_w23	584	133	717	3

[Return to Table of Contents](#)

## Summary Statistics for All Outcomes Full Table



OUTCOME	COUNT	MEAN	STD	MISSING_PCT
Children_NoMarriage_w23	709.0	0.224	0.417	1.53
CnctAdult_w23	719.0	0.943	0.232	0.14
Connected_Youth	716.0	0.648	0.478	0.56
Homeless_w23	719.0	0.293	0.456	0.14
Incarc_w23	718.0	0.357	0.479	0.28
SubAbuse_w23	717.0	0.185	0.389	0.42

## Variable Importance Table Full Table

VARIABLE	FREQUENCY	MEAN COEFF	MEAN ABS COEFF	MEAN P-VALUE
PlacementInstability	4	1.16	1.90	0.0159
const	3	-1.08	2.68	0.0205
CnctAdult_w1	1	1.34	1.34	0.0324
PlacementType_3	1	1.12	1.12	0.0001
PlacementType_2	1	1.07	1.07	0.0007
Sex	2	-0.32	0.84	0.0111
White_Only	2	0.23	0.77	0.0156
DiagDis	2	0.03	0.52	0.0172
MonthsInCare	3	-0.02	0.03	0.0146

[Return to Table of Contents](#)

## Appendix: Binary Classification Model Matrix



### Binary Classification Model Matrix

**Model Coverage:** Significant predictor→outcome relationships (1=significant, 0=not significant, p<0.05)< p>

**Related Findings:** See [Significant Model Results](#) for detailed effect sizes and [Summary Dashboard](#) for key risk and protective factors.

**Table 1: Predictors vs Outcomes 1-5**

PREDICTOR	HAVING CHILDREN OUTSIDE OF MARRIAGE	HAVING A SUPPORTIVE ADULT RELATIONSHIP	BEING HOMELESS	BEING INCARCERATED	HAVING A SUBSTANCE ABUSE REFERRAL
<b>Constant</b>	1	1	0	0	1
<b>Race: White (compared to Black/Other)</b>	1	0	0	0	1
<b>Relationship with a Supportive Adult</b>	0	1	0	0	0
<b>Sex (male)</b>	0	0	0	1	0
<b>Foster Home</b>	0	0	0	1	0
<b>Group Home</b>	0	0	0	1	0
<b>Other Placement Types</b>	0	0	0	0	0
<b>Placement Instability</b>	0	0	1	1	1
<b>Months In Care</b>	0	0	0	1	1
<b>Diagnosed Disability</b>	0	0	1	0	0
<b>Service Received</b>	0	0	0	0	0
<b>Removed from family for reasons other than abuse or neglect</b>	0	0	1	0	0

[Return to Table of Contents](#)

**Table 2: Predictors vs Outcomes 6-8**

PREDICTOR	RECEIVING PUBLIC FOOD ASSISTANCE	BEING CURRENTLY ENROLLED IN HIGHER EDUCATION	RECEIVING HEALTH INSURANCE	HAS A HIGH SCHOOL DEGREE OR GED
<b>Constant</b>	1	0	1	0
<b>Race: White (compared to Black/Other)</b>	0	0	0	0
<b>Relationship with a Supportive Adult</b>	0	0	0	0
<b>Sex (Male)</b>	1	0	0	1
<b>Foster Home</b>	0	0	0	0
<b>Group Home</b>	0	0	0	0
<b>Other Placement Types</b>	0	0	0	0
<b>Placement Instability</b>	0	0	0	0
<b>Months In Care</b>	0	0	0	0
<b>Diagnosed Disability</b>	0	1	0	0
<b>Services Received</b>	0	1	0	0
<b>Removed from family for reasons other than abuse or neglect</b>	0	0	0	0

**Key Patterns:** Placement Instability (3 outcomes), Sex (3), Disability (4), Services (3)

[Return to Table of Contents](#)



## ⌚ Appendix: Model Details

📋 **Related Findings:** These model details support the statistical analysis in **Significant Model Results** and **Summary Dashboard**.

- ▶ 📈 **Modeling approach:** Logistic regression for each outcome
- ▶ 📈 **Variance Inflation Factors (VIF):** Checked for multicollinearity
- ▶ 🗞 **Significance threshold:**  $p < 0.05$

### Model Selection

| **Logistic Regression Chosen:**

- ▶ Appropriate for binary outcome variables
- ▶ Provides interpretable metrics such as odds ratios and significance levels
- ▶ Allows for examination of multiple predictors simultaneously

### Effect Size Calculation

- ▶ Significant predictors were extracted for each model and compiled into summary tables for interpretation
- ▶ The magnitude and direction of coefficients were used to determine whether predictors had positive or negative associations with the outcome

| **Assessed using:**

- ▶ Coefficients from the logistic regression
- ▶ P-values to determine statistical significance
- ▶ 95% confidence intervals to evaluate the precision of estimates

### Validation

- ▶ No formal out-of-sample validation (e.g., test/train split or cross-validation) was implemented in this version because the focus was on significant relationships and not 'predictability'. The entire dataset was used for the analysis of relationships.
- ▶ Manual checks were performed to verify the accuracy of model outputs and ensure correct interpretation of key coefficients and p-values

| **Robustness checks:**

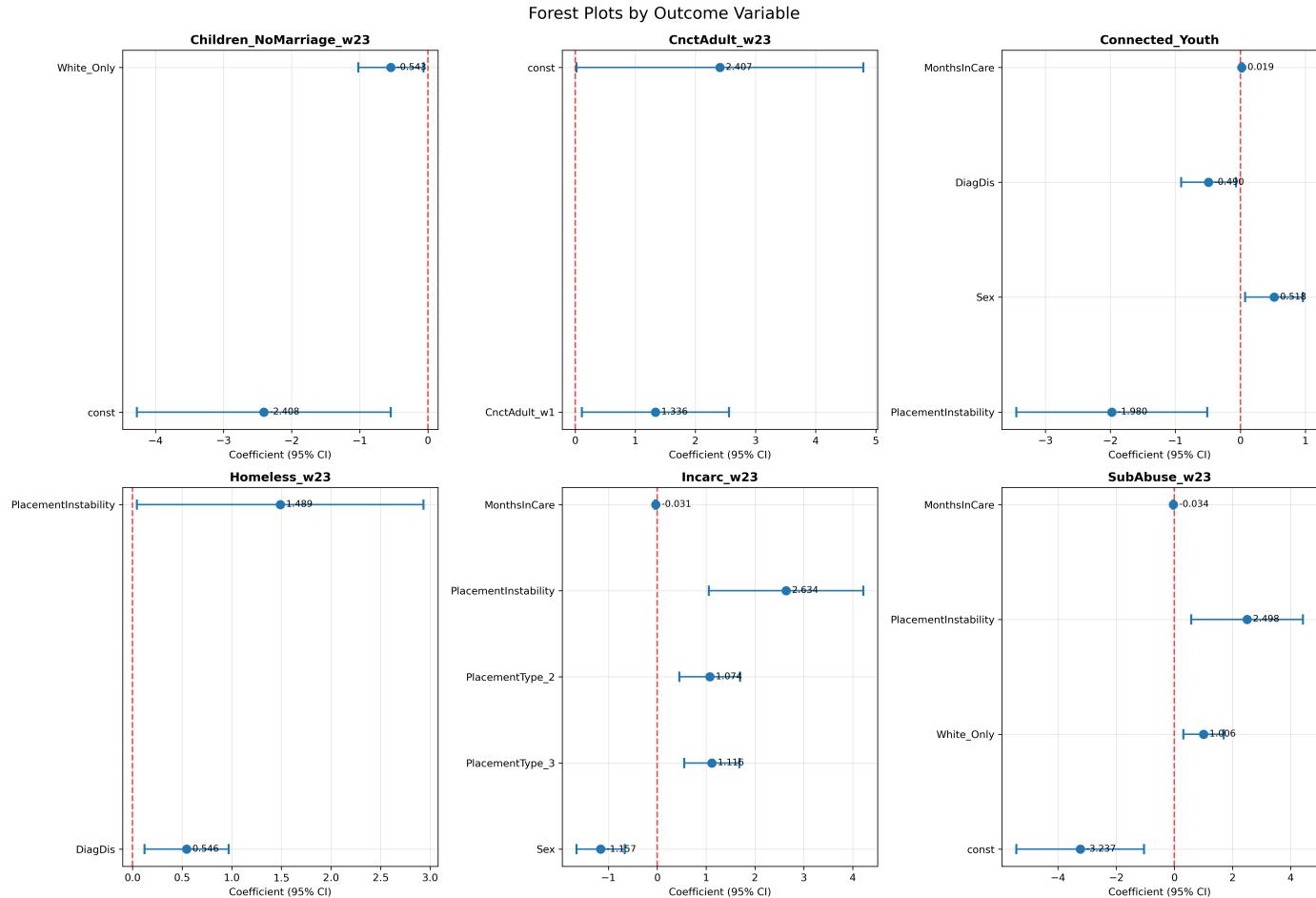
- ▶ Multiple outcomes analyzed independently
- ▶ Use of statistical thresholds ( $p < 0.05$ ) for interpreting effects
- ▶ Examination of confidence intervals for effect precision



## \_APPENDIX: Charts and Visualizations

**Related Findings:** These visualizations support the findings in **Descriptive Statistics**, **Significant Model Results**, and **Summary Dashboard**.

### Forest Plots by Outcome

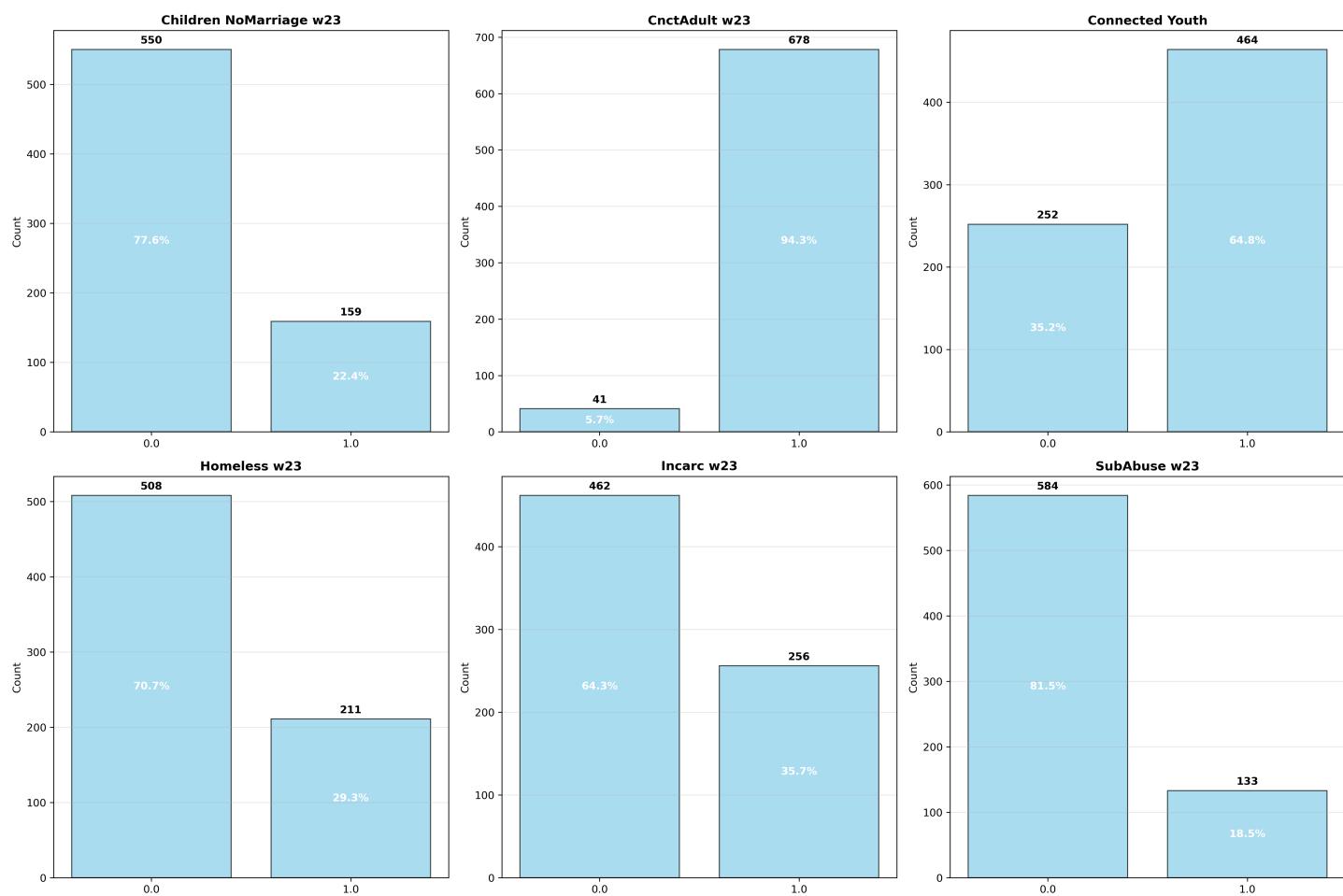


[Return to Table of Contents](#)

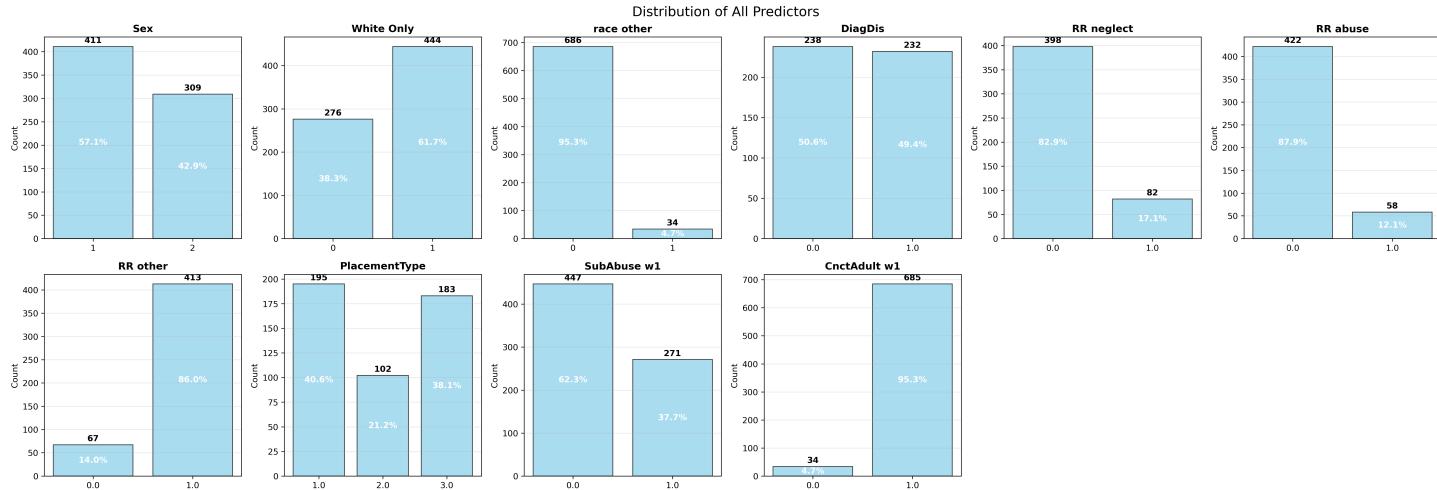
## Outcome Distribution Bar Charts



Distribution of All Outcomes

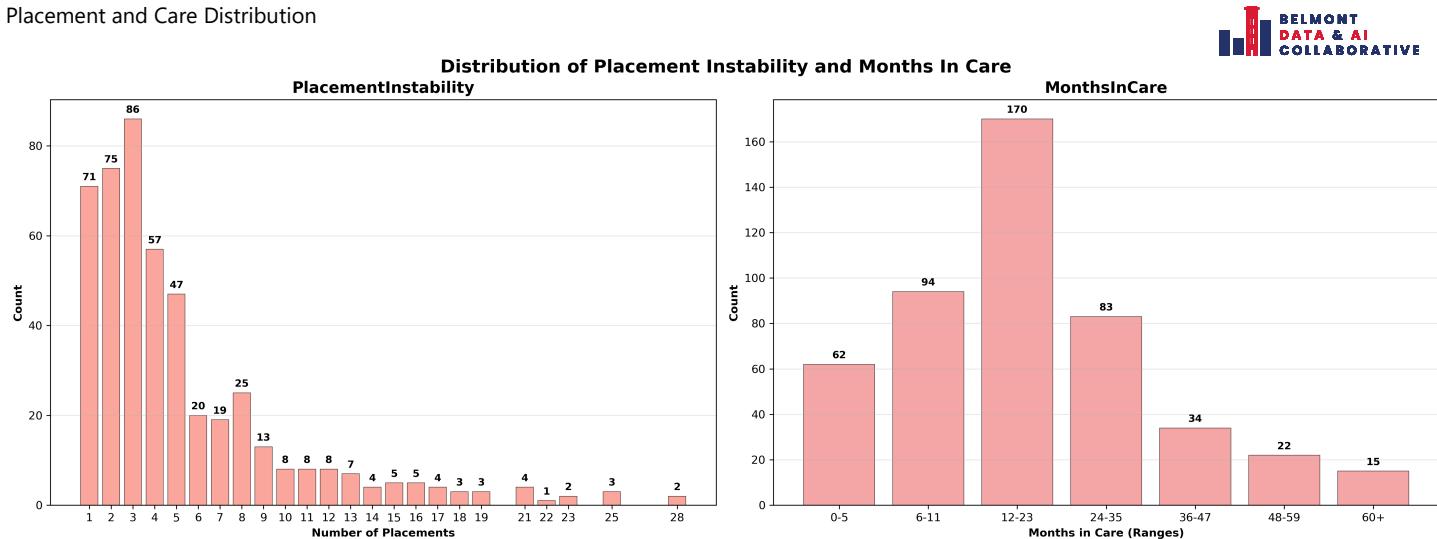


## Predictor Distribution Bar Charts

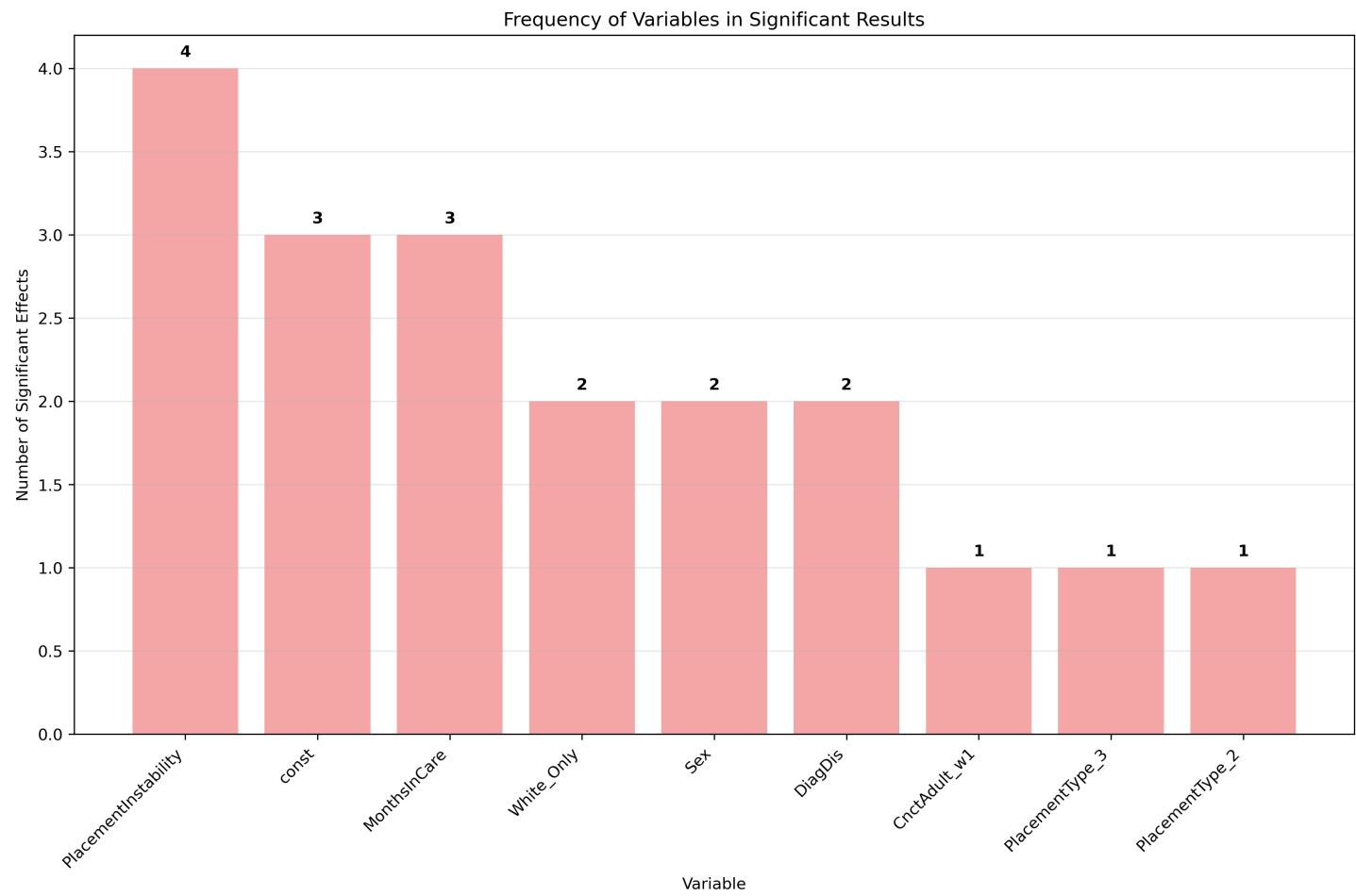


[Return to Table of Contents](#)

## Placement and Care Distribution



## Coefficient Magnitude Plot



[Return to Table of Contents](#)

## Statistical Model Results (ranked by Coefficient [log-odds])



**Understanding Log-Odds:** Log-odds is the natural logarithm of the odds ratio. It's used in logistic regression because it can range from negative infinity to positive infinity, making it suitable for linear modeling. The odds ratio can be calculated as  $e^{(\text{coefficient})}$ .

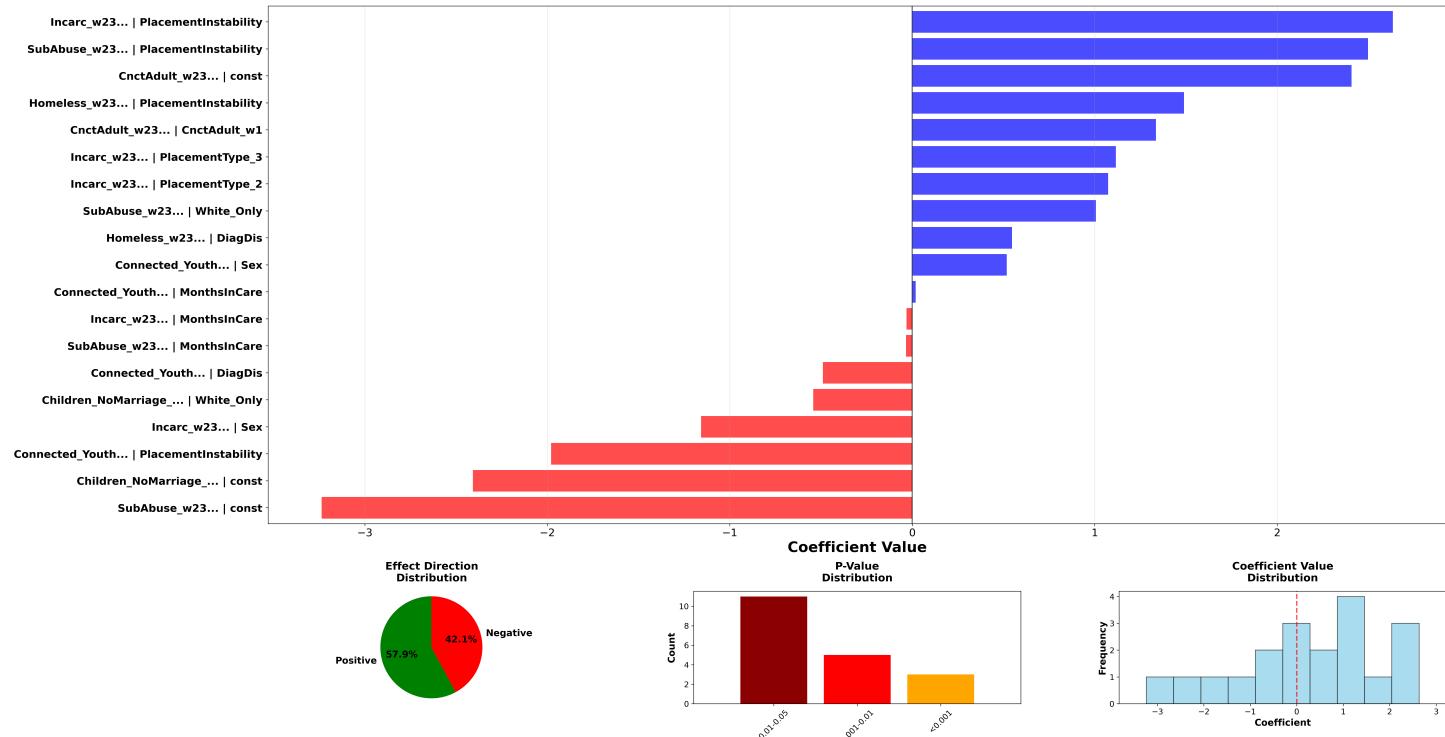
## Statistical Model Results Dashboard

**19**  
Total Significant Effects

**11**  
Positive Effects

**8**  
Negative Effects

## All Significant Effects Ranked by Coefficient



[Return to Table of Contents](#)



## 📋 Appendix: Data Breakdown

📋 **Related Findings:** This data breakdown supports the analysis in **Descriptive Statistics**, **Outcome Distributions**, and **Predictor Distributions**.

### Data Definition:

- ▶ **Valid Count:** Number of non-missing observations
- ▶ **Missing Count:** Number of missing observations
- ▶ **Completeness:** Percentage of complete data (Valid Count / Total Count \* 100)
- ▶ **Zero Values (%):** Percentage of observations with value 0
- ▶ **Mode Count:** Number of times the most frequent value appears
- ▶ **Mode (%):** Percentage of observations with the modal value
- ▶ **Quality Flags:** Data quality indicators (Good, Many Zeros, Low Completeness, High Outliers)

### Predictor Statistics

VARIABLE	VALID COUNT	MISSING COUNT	COMPLETENESS (%)	MEAN	MEDIAN	STD DEV	UNIQUE VALUES	ZERO VALUES (%)	MODE	M CO
Sex	720	0	100.0%	1.429	1.0	0.495	2	0.0%	1.0	
White_Only	720	0	100.0%	0.617	1.0	0.487	2	38.3%	1.0	
Black_Only	720	0	100.0%	0.0	0.0	0.0	1	100.0%	0.0	
race_other	720	0	100.0%	0.047	0.0	0.212	2	95.3%	0.0	
DiagDis	470	250	65.3%	0.494	0.0	0.5	2	50.6%	0.0	
PlacementType	480	240	66.7%	1.975	2.0	0.888	3	0.0%	1.0	
PlacementInstability	480	240	66.7%	5.292	4.0	4.867	24	0.0%	3.0	
MonthsInCare	480	240	66.7%	20.99	16.444	16.914	390	0.0%	17.774	
Service_Count	245	475	34.0%	0.449	0.0	1.291	8	73.9%	0.0	
RR_neglect	480	240	66.7%	0.171	0.0	0.377	2	82.9%	0.0	

VARIABLE	VALID COUNT	MISSING COUNT	COMPLETENESS (%)	MEAN	MEDIAN	STD DEV	UNIQUE VALUES	ZERO VALUES (%)	BELMONT DATA FAI MODE (%)	MISSING (%)
RR_abuse	480	240	66.7%	0.121	0.0	0.326	2	87.9%	0.0	
RR_other	480	240	66.7%	0.86	1.0	0.347	2	14.0%	1.0	
SubAbuse_w1	718	2	99.7%	0.377	0.0	0.485	2	62.3%	0.0	
CnctAdult_w1	719	1	99.9%	0.953	1.0	0.212	2	4.7%	1.0	

[Return to Table of Contents](#)

## Outcome Statistics



VARIABLE	VALID COUNT	MISSING COUNT	COMPLETENESS (%)	MEAN	MEDIAN	STD DEV	UNIQUE VALUES	ZERO VALUES (%)	MODE	MODE COUNT	MODE (%)	QUALITY FLAGS
Children_NoMarriage_w23	709	11	98.5%	0.224	0.0	0.417	2	77.6%	0.0	550	77.6%	High Outliers; Many Zeros
CnctAdult_w23	719	1	99.9%	0.943	1.0	0.232	2	5.7%	1.0	678	94.3%	High Outliers
Connected_Youth	716	4	99.4%	0.648	1.0	0.478	2	35.2%	1.0	464	64.8%	Many Zeros
Homeless_w23	719	1	99.9%	0.293	0.0	0.456	2	70.7%	0.0	508	70.7%	Many Zeros
Incarc_w23	718	2	99.7%	0.357	0.0	0.479	2	64.3%	0.0	462	64.3%	Many Zeros
SubAbuse_w23	717	3	99.6%	0.185	0.0	0.389	2	81.5%	0.0	584	81.5%	High Outliers; Many Zeros

[Return to Table of Contents](#)



## 📋 Appendix: Source Files, Modeling, and Method References

📋 Related Findings: These source files and methods support the analysis in Data Curation & Methodology and Significant Model Results.

Data used: [foster\\_care\\_univariate\\_analysis\\_report - v3.pdf](#), [foster\\_care\\_bivariate\\_analysis\\_report - v4.](#).

Models created by [all\\_visual\\_data.ipynb](#).

Variable Coding and Approach based on an academic article published in [The Journal Of Public Child Welfare](#) called Aging Out of Foster Care: Homelessness, Post-Secondary Education, and Employment