

The session will begin at 12pm EST.

Please submit questions to the Q&A box.

This session is being recorded.

See ZOOM Help Center for connection issues: https://support.zoom.us/hc/en-us

If issues persist and solutions cannot be found through Zoom, please contact Andres Arroyo at aa17@cornell.edu.

NDACAN SUMMER TRAINING SERIES

National Data Archive on Child Abuse and Neglect

Duke University, Cornell University, UC San Francisco, & Mathematica





LIFE CYCLE OF AN NDACAN RESEARCH PROJECT

NDACAN SUMMER TRAINING SERIES SCHEDULE

- July 2nd, 2025
 - Developing a research question & exploring the data
- July 9th, 2025
 - Data management
- July 16th, 2025
 - Linking data
- July 23rd, 2025
 - Exploratory Analysis
- July 30th, 2025
 - Visualization and finalizing the analysis

SESSION AGENDA

• Developing a research question

Exploring the data

DEVELOPING A RESEARCH QUESTION

FIRST PRINCIPLES FOR GOOD QUESTIONS

- Clarity
- Focus
- Answerability

IMPROVING RESEARCH QUESTIONS

- How unequal is child welfare system contact?
 - What is the ratio of annual incidence of maltreatment investigation across ethnoracial groups?
- How do youth experience instability after aging out of foster care?
 - What is the risk of homelessness three years after aging out of foster care?
- Why are children placed into different types of out-of-home care?
 - What is the effect of maltreatment type on placement setting?

NESTING RESEARCH QUESTIONS

- Consider organizing your research question as a set of nested questions with different levels of specificity
- This can help identify both theoretical implications (higher levels) and empirical details (lower levels) of your question(s)
- What is the relationship between family stability and child well-being?
 - What is the relationship between foster placement and health?
 - What is the association between placement type and health in the transition to adulthood?
 - What is the effect of institutional placement (relative to foster family placement) on the risk of substance abuse referral among children three years after aging out of foster care?

WHERE TO START

- Interesting topic
- Influential theory
- Valuable data source
- Innovative method
- Compelling study

QUESTIONS AND "FRAMING"

- Puzzle: tension between theory and observation that your question can resolve
 - Caution: perhaps there's a better-suited theory
- Gap: absence of prior observation that your question can fill in
 - Caution: perhaps there's a good reason for the gap

Further reading: "On Genre: A Few More Tips to Article-Writers" by Ezra W. Zuckerman

https://www.dropbox.com/scl/fi/sv1u2txx9djc74gfywj9v/On-Genre.pdf?rlkey=6urcf87h57m4s75wktds4qnqe&e=1

QUESTION DEVELOPMENT AND EXPLORATORY RESEARCH

- 1. Develop research question
- 2. Identify/explore potential data sources
 - a. Data documentation
 - b. Data
- 3. Revise/refine question
- 4. Repeat 2-3 as necessary
 - a. Avoid data dredging
 - b. Consider pre-registration

Further reading:

"Pre-Registering Studies – What Is It, How Do You Do It, and Why?" by Kriti Jain, Marie Lawrence, and Jenessa Malin

https://acf.gov/opre/blog/2022/08/pre-registering-studies-what-it-how-do-you-do-it-and-why

NOT REINVENTING THE WHEEL

- What sources and methods have prior researchers used to answer similar research questions?
 - Standard library research methods
- What research questions have prior researchers asked using potential data sources?
 - Child Abuse and Neglect Digital Library (canDL)

IDENTIFYING RELEVANT RESEARCH

U.S. Department of Health & Human Services Administration for Children & Families





ABOUT DATASETS FAQ USER SUPPORT CMRLE-LIST EVENTS CANDL BIBLIOGRAPHY MEASURES INDEX CONTRIBUTE DATA SEARCH

DATASET DETAILS

National Youth in Transition Database (NYTD) - Outcomes File, Cohort Age 17 in FY2020, Waves 1-3 (Complete)

Current Data Version: 1

Investigator(s)

Dataset Number: 297

Children's Bureau, Department of Health and Human Services

Dataset 297 Datasets List

Dataset 297

Inquire About

Abstract

The John H. Chafee Foster Care Independence Program (CFCIP) was initiated in an effort to improve outcomes for youth in foster care who are likely to reach their 18th birthday without having found a permanent home. The program provides funding to states to develop and administer programs designed to help ease the transition from foster care to independence.

The law that created CFCIP also required states to develop 1) a system for tracking the services provided through CFCIP, and 2) a method for collecting outcome measures so that the effectiveness of the program can be assessed.

These two components together form the National Youth in Transition Database (NYTD). The files contain data from all 50 states, as well as from the District of Columbia and Puerto

The Outcomes component (this dataset) contains the results of surveys conducted with youth to examine certain well-being, financial, and educational outcomes as they get older. Data from a survey on outcomes for a cohort of youth is provided every other year beginning with federal fiscal year 2011 (Oct 1, 2010 through Sep 30, 2011). There are three phases of outcome data for a given cohort -- a baseline survey during the year of the youth's 17 birthday, and two follow-ups: one at age 19, one at age 21. New cohorts will be established every three years (2014, 2017, 2020...) after the initial one in FY2011.

The Services component contains cross-sectional information on the services provided by states under CFCIP and the youth who receive those services. Data are submitted by the states every 6 months on a continuing basis.

Bibliographic Citation

Children's Bureau, Administration on Children, Youth and Families, Administration for Children and Families, U.S. Department of Health and Human Services. (2025). National Youth in Transition Database (NYTD) - Outcomes Survey, Cohort Age 17 in FY2020 Waves 1-3 [Data set]. National Data Archive on Child Abuse and Neglect (NDACAN). https://doi.org/10.34681/021Q-V394

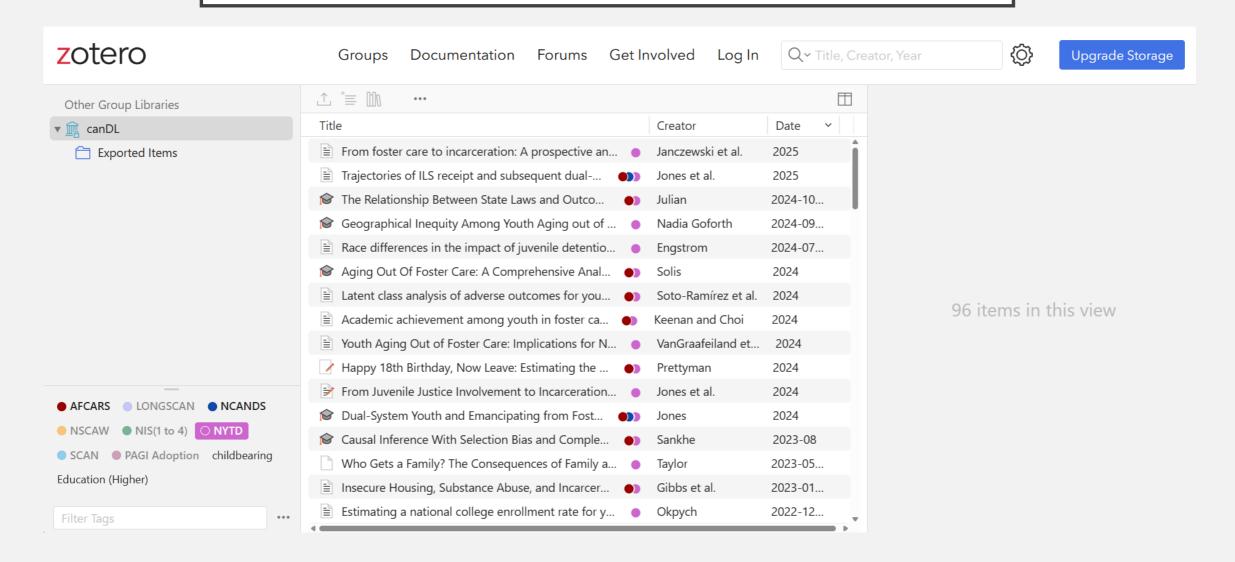
Data Documentation

- · NYTD Outcomes User's Guide
- NYTD Outcomes Codebook

Publications from this Dataset

NYTD Publications

CANDL



ADDITIONAL CONSIDERATIONS

- What are your goals for your research?
- What will sustain your interest?
- What is your comparative advantage?

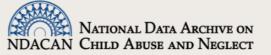
PRELIMINARY RESEARCH QUESTION

What is the relationship between placement history and unemployment among youth aging out of foster care?

EXPLORING THE DATA

FINDING DATA AND DATA DOCUMENTATION







ABOUT DATASETS FAQ USER SUPPORT CMRLE-LIST EVENTS CANDL BIBLIOGRAPHY MEASURES INDEX CONTRIBUTE DATA SEARCH

DATASETS

The National Data Archive on Child Abuse and Neglect (NDACAN) distributes the data for each study listed below. All datasets are distributed free of charge. Dataset ordering requirements vary: click on a dataset name to read its details and to access its ordering instructions.

All Datasets NIS NSCAW NYTD AFCARS: Foster Care Adoption 6-month File NCANDS: Child File Agency File

Dataset Number	Study Title		
MULTIPLE	Adoption and Foster Care Analysis and Reporting System (AFCARS) <u>Foster Care Files</u> , <u>Adoption Files</u> , <u>6-month Files</u>		
MULTIPLE	National Youth in Transition Database (NYTD)		
MULTIPLE	National Child Abuse and Neglect Data System (NCANDS) State Level Data Years available: 1990-1999 (SDC), 2000-2008 (CAF), 2009-2023 (NCANDS Agency File)		
MULTIPLE	National Child Abuse and Neglect Data System (NCANDS) Child Level Data Years available: 1995-1999 (DCDC), 2000-2023 (Child File)		
MULTIPLE	National Incidence Study of Child Abuse and Neglect (NIS) Years Available: 1980 (NIS-1), 1987 (NIS-2), 1996 (NIS -3), 2006 (NIS-4)		
290	Regional Partnership Grants National Cross-Site Evaluation (RPG-5 and RPG-6)		
296	State Child Abuse and Neglect (SCAN) Policies Database 2019-2023		
295	State Child Abuse and Neglect (SCAN) Policies Database 2023		
287	State Child Welfare (SCW) Data Linkages Descriptive Study		
283	Florida Study of Professionals for Safe Families (FSPSF), Waves 1-8		
282	Neglect and Prevention Policies (NAPP) Data		
272	Child and Caregiver Outcomes Using Linked Data (CCOULD)		
271	Rockwool-Duke Global Child Welfare Database		

ACCESSING DATA AND DATA DOCUMENTATION

U.S. Department of Health & Human Services Administration for Children & Families





ABOUT DATASETS FAQ USER SUPPORT CMRLE-LIST EVENTS CANDL BIBLIOGRAPHY MEASURES INDEX CONTRIBUTE DATA SEARCH

Dataset 297

Inquire About Dataset 297

Datasets List

DATASET DETAILS

National Youth in Transition Database (NYTD) - Outcomes File, Cohort Age 17 in FY2020, Waves 1-3 (Complete)

Dataset Number: 297

Current Data Version: 1

Investigator(s)

Children's Bureau, Department of Health and Human Services

Abstract

The John H. Chafee Foster Care Independence Program (CFCIP) was initiated in an effort to improve outcomes for youth in foster care who are likely to reach their 18th birthday without having found a permanent home. The program provides funding to states to develop and administer programs designed to help ease the transition from foster care to independence.

The law that created CFCIP also required states to develop 1) a system for tracking the services provided through CFCIP, and 2) a method for collecting outcome measures so that the effectiveness of the program can be assessed.

These two components together form the National Youth in Transition Database (NYTD). The files contain data from all 50 states, as well as from the District of Columbia and Puerto

The Outcomes component (this dataset) contains the results of surveys conducted with youth to examine certain well-being, financial, and educational outcomes as they get older. Data from a survey on outcomes for a cohort of youth is provided every other year beginning with federal fiscal year 2011 (Oct 1, 2010 through Sep 30, 2011). There are three phases of outcome data for a given cohort -- a baseline survey during the year of the youth's 17 birthday, and two follow-ups: one at age 19, one at age 21. New cohorts will be established every three years (2014, 2017, 2020...) after the initial one in FY2011.

The Services component contains cross-sectional information on the services provided by states under CFCIP and the youth who receive those services. Data are submitted by the states every 6 months on a continuing basis.

Bibliographic Citation

Children's Bureau, Administration on Children, Youth and Families, Administration for Children and Families, U.S. Department of Health and Human Services. (2025). National Youth in Transition Database (NYTD) - Outcomes Survey, Cohort Age 17 in FY2020 Waves 1-3 [Data set]. National Data Archive on Child Abuse and Neglect (NDACAN). https://doi.org/10.34681/021Q-V394

Data Documentation

- · NYTD Outcomes User's Guide
- NYTD Outcomes Codebook

Publications from this Dataset

NYTD Publications

National Youth in Transition Database (NYTD)

OUTCOMES FILE USER'S GUIDE

NYTD OUTCOMES SURVEY FY2020 COHORT WAVE 1

NDACAN DATASET #256

Revised October 1, 2024



Duke University / Cornell University Phone: 607-255-7799 Email: ndacan@cornell.edu Website: www.ndacan.acf.hhs.gov

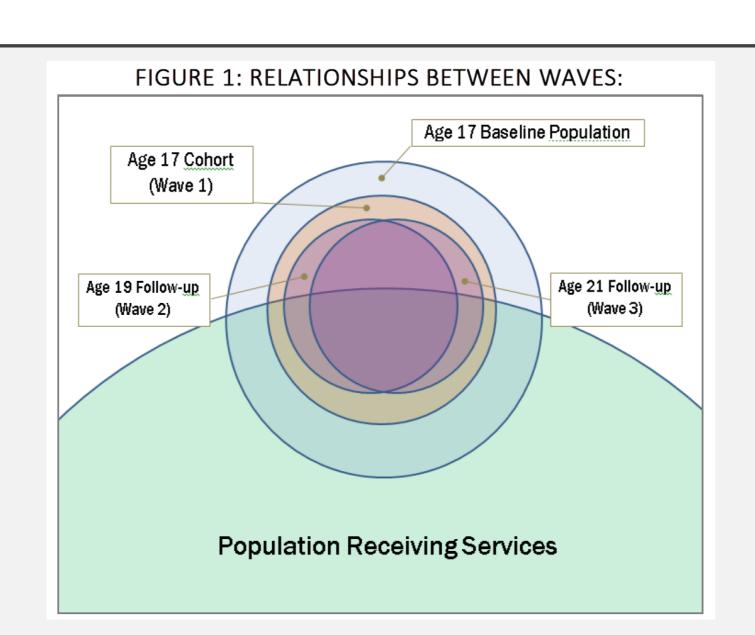
NATIONAL YOUTH IN TRANSITION DATABASE (NYTD)

NYTD OUTCOMES FILE CODE BOOK

Revised September 30, 2024



Duke University / Cornell University Phone: 607-255-7799 Email: ndacan@cornell.edu Website: www.ndacan.acf.hhs.gov



UNDERSTANDING AVAILABLE DATA: NYTD

NYTD OUTCOMES VARIABLES – BY POSITION

Page	Element	Variable Name	Variable Label	Туре
6	D	Wave	1 = Baseline (age 17); 2 = Age 19; 3 = Age 21	tinyint
7	D	StFCID	St and RecNumbr (for linking)	char(14)
8	#1	StFIPS	#1: State FIPS Code	tinyint
9	D	St	Two-letter USPS code for state	char(2)
10	#3	Recnumbr	#3: AFCARS Unique Child ID (Encrypted)	char(12)
11	#2	Repdate	#2: Report Date	char(6)
12	#4	DOB	#4: Date of Birth	date
13	#5	Sex	#5: Sex	tinyint
14	#6	Amiakn	#6: Race – American Indian Or Alaskan Native	tinyint
15	#7	Asian	#7: Race – Asian	tinyint
16	#8	Blkafram	#8: Race – Black or African American	tinyint
17	#9	Hawaiipi	#9: Race – Native Hawaiian or other Pacific Islander	tinyint
18	#10	White	#10: Race – White	tinyint
19	#11	Raceunkn	#11: Race – Unknown	tinyint
20	#12	Racedcln	#12: Race – Declined	tinyint
21	#13	Hisorgin	#13: Hispanic Or Latino Ethnicity	tinyint
22	#34	OutcmRpt	#34: Outcomes Reporting Status	tinyint
24	#35	OutcmDte	#35: Date of Outcome Data Collection	date
25	#36	OutcmFCS	#36: Foster Care Status at Outcomes Collection	tinyint
26	#37	CurrFTE	#37: Current Full Time Employment	tinyint
27	#38	CurrPTE	#38: Current Part-Time Employment	tinyint
28	#39	EmplySklls	#39: Employment Related Skills	tinyint
29	#40	SocSecrty	#40: Social Security	tinyint
30	#41	EducAid	#41: Educational Aid	tinyint
31	#42	PubFinAs	#42: Public Financial Assistance	tinyint
32	#43	PubFoodAs	#43: Public Food Assistance	tinyint
33	#44	PubHousAs	#44: Public Housing Assistance	tinyint
34	#45	OthrFinAs	#45: Other Financial Support	tinyint
35	#46	HighEdCert	#46: Highest Educational Certification Received	tinyint
36	#47	CurrEnroll	#47: Current Enrollment and Attendance	tinyint
37	#48	CnctAdult	#48: Connection to Adult	tinyint
38	#49	Homeless	#49: Homelessness	tinyint
39	#50	SubAbuse	#50: Substance Abuse Referral	tinyint
40	#51	Incarc	#51: Incarceration	tinyint
41	#52	Children	#52: Children	tinyint
42	#53	Marriage	#53: Marriage at Child's Birth	tinyint
43	#54	Medicaid	#54: Medicaid	tinyint

UNDERSTANDING AVAILABLE DATA: AFCARS

NDACAN AFCARS FOSTER CARE VARIABLES - BY POSITION

Var#	Name	Variable Label	Regulation	Element	Page
1	FY	The Federal Fiscal Year of this Dataset	Derived	Licinciic	12
2	RepDatMo	Reporting Period End Date: Month	I.B-m	2	13
3	RepDatYr	Reporting Period End Date: Year	I.B-y	2	14
4	State	State FIPS Code	I.A	1	15
5	St	State Two-Character Code	Derived	-	16
6	FIPSCode	Local Agency FIPS Code	I.C	3	17
7	RecNumbr	Record Number (AFCARS ID)	I.D	4	18
8	PedRevDt	Periodic Review Date	I.E	5	19
9	DOB	Foster Child's Date of Birth	II.A	6	20
10	Sex	Child Sex	II.B	7	21
11	AmIAKN	Child Race American Indian or Alaska Native	II.C.1.a	8	22
12	Asian	Child Race Asian	II.C.1.b	8	23
13	BlkAfrAm	Child Race Black or African American	II.C.1.c	8	24
14	HawaiiPI	Child Race Native Hawaiian or Other Pacific Islander	II.C.1.d	8	25
15	White	Child Race White	II.C.1.e	8	26
16	UnToDetm	Child Race Unable to Determine	II.C.1.f	8	27
17	HisOrgin	Child Hispanic or Latino Ethnicity	II.C.2	9	28
18	ClinDis	Child Has Been Clinically Diagnosed with Disability	II.D	10	29
19	MR	Mental Retardation	II.D.a	11	30
20	VisHear	Visually or Hearing Impaired	II.D.b	12	31
21	PhyDis	Physically Disabled	II.D.c	13	32
22	EmotDist	Emotionally Disturbed	II.D.d	14	33
23	OtherMed	Other Medically Diagnosed Condition Requiring Special Care	II.D.e	15	35
24	EverAdpt	Child Has Previously Been Adopted	II.E.1	16	37
25	AgeAdopt	Age on Date of Legal Adoption	II.E.2	17	38
26	Rem1Dt	Date of First Removal	III.A.1	18	39
27	TotalRem	Total Number of Removals from Home	III.A.2	19	40
28	DLstFCDt	Discharge Date for Previous Removal	III.A.3	20	41

CurrFTE

Variable Label: #37: Current Full Time Employment

Definition:

A youth is employed full-time if employed at least 35 hours per week, in one or multiple jobs, as of the date of the outcome data collection.

"Yes" means the youth is employed fulltime.

"Declined" means the youth did not answer this question.

"Blank" means the youth did not participate in the survey.

Data Type: TinyInt

NYTD Element: #37

Value	Value Labe	
0	no	
1	yes	
2	declined	
77	blank	

REVISED RESEARCH QUESTION

What is the relationship between lifetime incidence of removal and full-time employment among youth three years after aging out of foster care?

QUESTIONS?

ALEX ROEHRKASSE AROEHRKASSE@BUTLER.EDU

NOAH WON DUKE.EDU

PAIGE LOGAN PRATER
PAIGE.LOGANPRATER@UCSF.EDU

NEXT WEEK...

Date: July 9th, 2025

Topic: Data Management

Instructor: Alex Roehrkasse