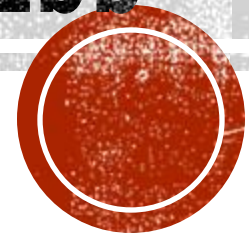# PREDICTING SUICIDAL EXPERIENCES AMONG YOUTH EXPERIENCING HOMELESSNESS (YEH): STRIVING FOR STABILITY, ACCURACY, AND FAIRNESS

Anthony Fulginiti, PhD, MSW

Associate Professor, Graduate School of Social Work

University of Denver

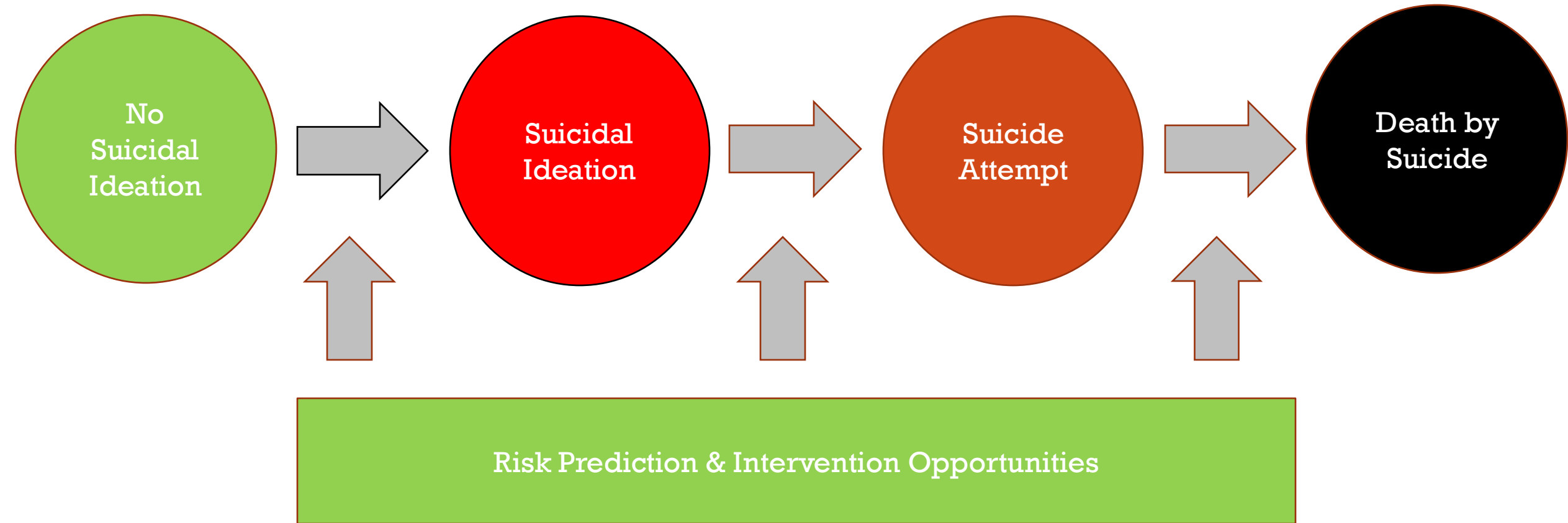# HOMELESSNESS & SUICIDAL EXPERIENCES

- ~4.2 million youth experience homelessness each year in the U.S.

- Lifetime rates of suicidal ideation: ~40%

- Lifetime rates of suicide attempts: ~35%

- Past-year rates of suicidal ideation: ~30%

- Past-year rates of suicide attempts: ~20%

http://www.nekys.org/homeless-youth/

SUICIDE PREDICTION IS HARD!!!

PERSON-CENTRIC VIEW OF SUICIDE: EXAMPLE CORRELATES OF SI & SA

- Age (Younger = higher risk)
- Gender (Female = higher risk)
- Depression
- Substance Use
- History of Trauma
- Post-traumatic Stress

Individual · Relationship · Community · Societal

https://www.cdc.gov/violenceprevention/about/social-ecologicalmodel.html

https://college.harvard.edu/about/news-announcements/ai-can-help-reduce-risk-hiv-high-risk-communities
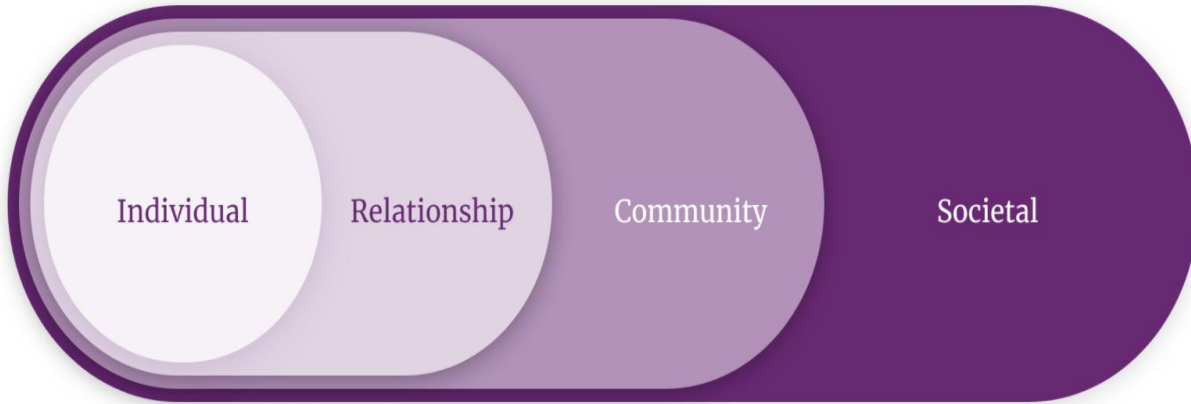
A SOCIOECOLOGICAL VIEW OF SUICIDE: THE ENVIRONMENT MATTERS

SOCIOECOLOGICAL VIEW OF SUICIDE: EXAMPLE CORRELATES OF SI & SA

RELATIONSHIP FACTORS

- Social support
- Contagion
- Relational Conflict/Victimization
- Coping Norms

COMMUNITY/SOCIETAL FACTORS
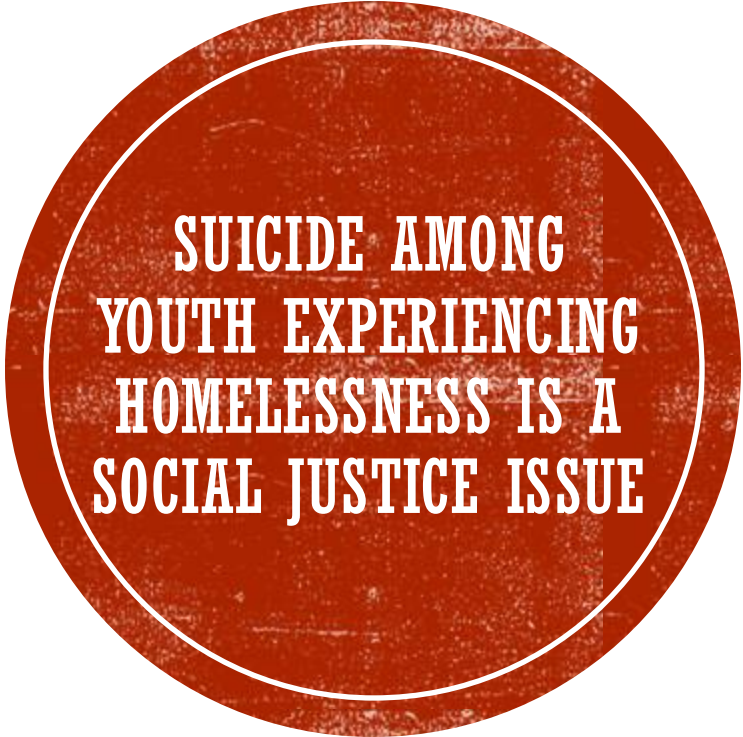
- Socioeconomic Disadvantage
- Inadequate/Inaccessible Healthcare
- Discrimination & Social Exclusion
- Historical Trauma

EFFORTS TO IMPROVE PREDICTIONS: ITERATION 1
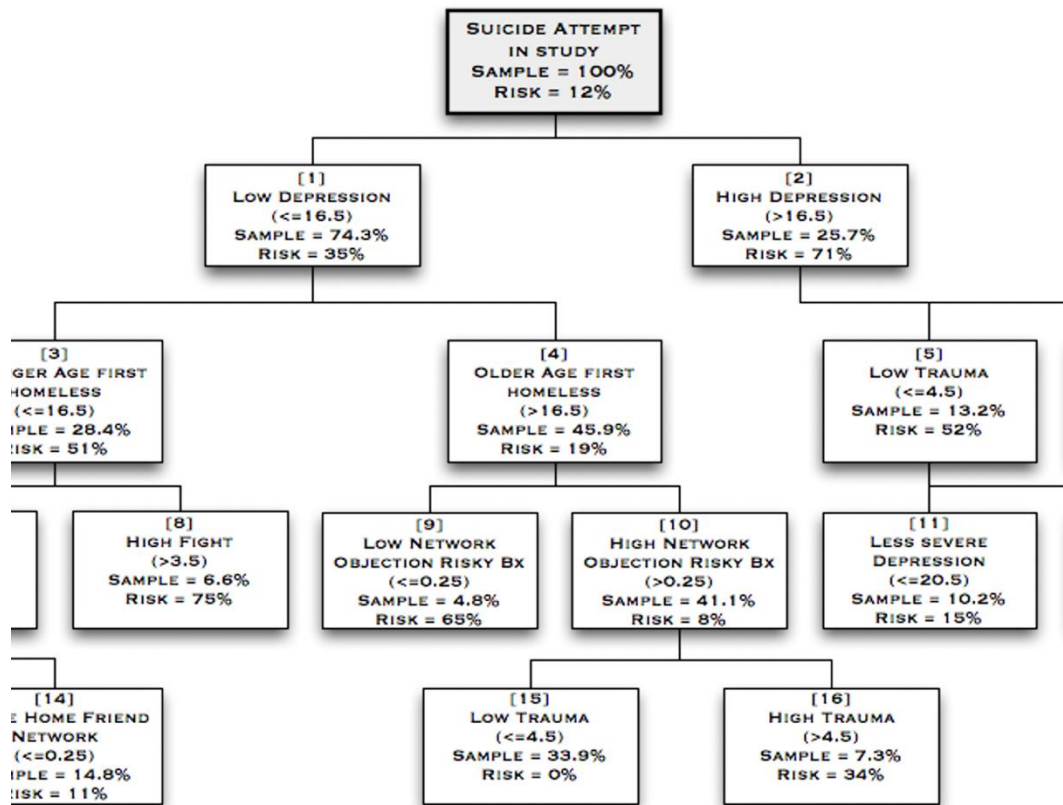
- We included both personal and social environment characteristics into our prediction models

- We used decision tree analyses (specifically CART models) to account for complex interactions and produce results that are more interpretable to guide service delivery

WHAT WERE THE RESULTS OF OUR DT MODELS?
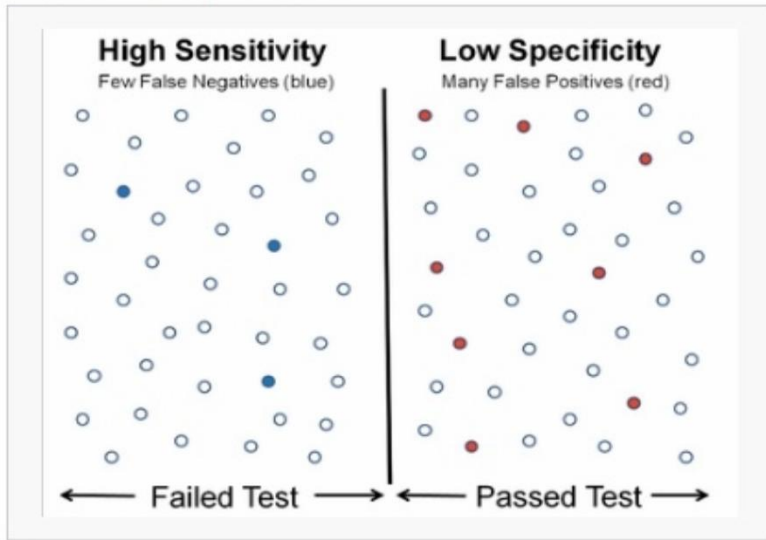
- Suicidal Ideation
  - AUC: .79
  - Sensitivity: 48%
  - Specificity: 73%

- Suicide Attempts
  - AUC: .86
  - Sensitivity: 68%
  - Specificity: 72%

- Literature Comparison
  - Median/Mean AUC = .76/.78

ITERATION 1: LIMITATIONS

- We used a train-split approach: a single split was used to create the training and testing datasets (80:20) for our decision-tree models. This raised questions about the robustness of the model performance.

- The majority of youth did not have suicidal experiences, but we did not account for class imbalance issues when making predictions.

- We did not test whether predictions were fair across different gender, sexual orientation, and racial/ethnic groups.

https://portlandenglish.edu/pela-blog/have-or-be/question-mark-1421013_1280/

ITERATION 2: SUMMARY & CHALLENGES

- This problem was presented in this class several years ago and efforts were made to address the limitations of iteration 1.

- Although k-fold cross validation was used to better evaluate the robustness of the overall model performance, the individual surrogate models were substantially different from one another (e.g., different constituent features).

- Although it is common to have different feature/variable sets across individual trees when using ensemble methods (e.g., random forests), such differences can make it more challenging to interpret the results for applied purposes (e.g., treatment/program planning)

- Is it possible to develop a model that can produce more interpretable results (i.e., stability of features across individual trees) while also delivering accurate and fair predictions?

https://portlandenglish.edu/pela-blog/have-or-be/question-mark-1421013_1280/

OVERARCHING PROJECT AIMS

1. Increase the stability of feature/variable sets across individual surrogate trees

2. Increase the accuracy of our predictions related to suicidal experiences (e.g., sensitivity/specificity)

3. Improve the fairness of our predictions related to suicidal experiences across gender, sexual orientation, and racial/ethnic identities

IMPLICATIONS OF THE PROJECT

- Decision trees that produce accurate suicide-related predictions can lead to more effective suicide risk management

- Although ensemble methods are generally preferred to improve predictions, such methods can introduce challenges for interpretability. Finding a way to improve prediction accuracy while maintaining interpretability may help to inform suicide prevention programming, resource allocation, and perhaps even aid theory development

- Striving for fairness in our decision tree predictions is a social justice action that ensures we are not compounding the suicidogenic effects of inequity.

https://www.nuveen.com/en-us/insights/fixed-income-impact-investing/impact-framework-methodology
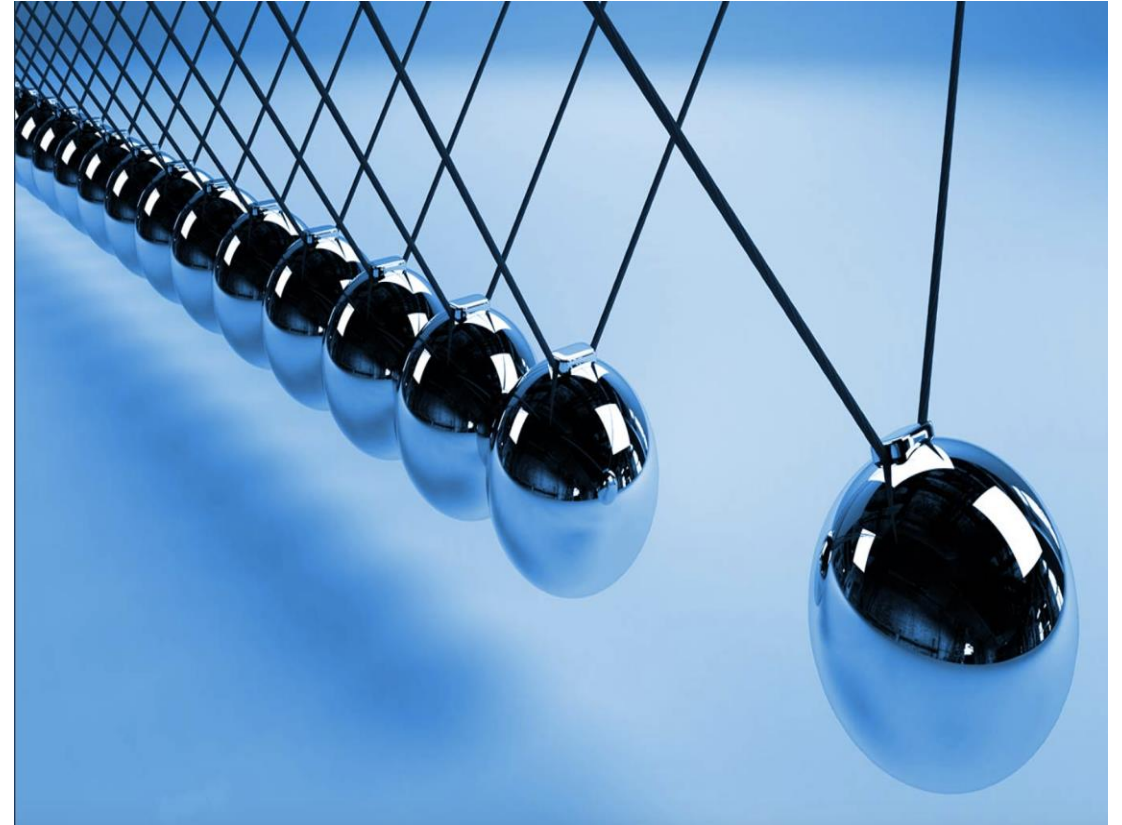
DESCRIPTION OF THE PROJECT DATASET

- Sample size = 940

- Data were collected as part of project designed to understand HIV/AIDS risk-taking behavior among young people experiencing homelessness.

- Data collection occurred between October 2011 and February 2013

- Recruited from two drop-in centers that serve young people experiencing homelessness in Hollywood and Santa Monica, CA.

- Survey + Social Network Interview

https://www.linkedin.com/pulse/seven-key-characteristics-project-stanley-epstein/

PROSPECTIVE DEPENDENT VARIABLES

- **Active Suicidal Ideation (*label= suicidea*)**
  - "During the past 12 months, did you ever seriously consider attempting suicide?"
  - Coded: 0= No and 1= Yes

- **Suicide Plan (*label= suicplan*)**
  - During the past 12 months, did you make a plan about how you would attempt suicide?
  - Coded: 0= No and 1= Yes

- **Suicide Attempt (*label= suicattempt*)**
  - "During the past 12 months, how many times did you actually attempt suicide?"
  - Coded: 0= No Attempts and 1= ≥ 1 Attempt

- **Any Suicidal Experience (*label= suic_any*)**
  - Coded: 0= No and 1= Yes

https://www.bhsarkansas.org/about/news-events/national-suicide-prevention-month

INDEPENDENT VARIABLES: PERSONAL CHARACTERISTICS

- 47 variables that assess personal attributes

- Sociodemographic factors (e.g. education)

- Mental health factors (e.g., depression, trauma)

- Substance use (e.g., alcohol, marijuana, hard drug use),

- Sexual risk behaviors (e.g., # sex partners),

- Violence perpetration and victimization (e.g., weapon-carrying, IPV victim),

- Justice system involvement (e.g., spending time in jail; Harris et al., 2009),

- Online activity (e.g., time spent online)

- Homelessness indices (e.g., time spent homeless; age first homeless).

Variable	Obs	Mean	Std. dev.	Min	Max
pid	940	1351.826	1029.468	1	2618
site	940	1.541489	.4985409	1	2
wave	940	2.015957	.988087	1	4
age	921	21.34528	2.113672	14	29
gender	922	.2722343	.4453514	0	1
sexori	906	.2450331	.4303444	0	1
raceall	921	.6080347	.4884542	0	1
trauma_sum	886	2.965011	2.42583	0	8
ptsd_score	940	.2882979	.4532116	0	1
cesd_score	904	12.06259	6.017107	0	30
harddrug_l~e	940	6.3	4.364563	0	18
harddrug_30	940	3.953191	2.342354	0	18
otherdrug_~e	940	6.297872	3.8677	0	15
otherdrug_30	940	3.659574	1.804344	0	15
school	918	1.874728	.3312079	1	2
degree	940	1.701064	.5955094	0	3
job	909	1.869087	.337491	1	2
intertime	911	3.809001	1.920613	1	7
sntime	907	3.589857	1.850191	1	6
interlast	907	2.858875	1.785994	1	6

INDEPENDENT VARIABLES: PERSONAL CHARACTERISTICS

sex	912	1.074561	.2628265	1	2
sexpart30	852	3.0223	1.399419	1	9
sexpartlife	854	5.772834	1.879715	1	7
concurrent	813	1.633456	.482157	1	2
exchange	852	1.826291	.3790815	1	2
pregnant	899	1.830923	1.245169	1	6
children	901	1.432852	.9160225	1	6
weapon	906	2.315673	1.758866	1	5
fight	906	2.626932	2.06146	1	8
fighthurt	896	1.350446	.8349622	1	5
ipv	904	1.800885	.3995558	1	2
alcfirst	905	4.01547	1.82224	1	7
alc30	903	2.976744	1.886617	1	7
alcbinge	893	1.583427	2.023138	0	6
potfirst	901	4.299667	1.882401	1	7
pot30	905	3.941436	2.070627	1	6
jail	881	1.358683	.4798864	1	2
staycurrent	854	.2611241	.4395048	0	1
homelage	824	16.79126	3.993935	0	30
time_homel~h	792	53.76768	196.3012	1	4381

traveler	899	1.452725	.4980372	1	2
jailkid	888	1.096847	1.422033	0	4
jailadult	884	2.46267	1.827737	0	5
gethelp	940	34.72234	12.24478	0	54
gettherapy	891	5.020202	1.658359	1	6
sum_na	940	3.175532	10.25594	0	61
suicidea	904	.170354	.3761516	0	1
suicplan	900	.1177778	.322524	0	1
suicattempt	906	.1192053	.3242089	0	1
suic_any	940	.2074468	.4056941	0	1

INDEPENDENT VARIABLES: SOCIAL NETWORK CHARACTERISTICS

- 74 variables that assess social network attributes

- Two types of variables---proportion and sum scores---were derived for each network attribute (i.e., 37 unique network attributes measured in two different ways)

- Relationship types (i.e., family, friends, intimate partners, home-based peers, street-based peers)

- Frequency of contact

- Substance use behavior (e.g., hard drug use)

- Sexual risk behavior (e.g., unprotected sex)

- Social support provision (e.g., emotional or tangible support)

- Encouragement/rejection of risky behavior

sum_alter	940	10.49468	7.211617	1	63
sum_family	940	2.602128	2.924538	0	20
sum_friends	940	6.55	5.844329	0	52
sum_home_f~s	940	2.497872	3.383383	0	30
sum_street~s	940	4.052128	4.300785	0	27
sum_unknow~r	940	7.021277	5.852971	0	45
prop_family	940	.2613085	.2523844	0	1
prop_friends	940	.6396915	.330578	0	1
prop_home_~s	940	.2388298	.263763	0	1
prop_stree~s	940	.4005426	.310988	0	1
prop_unkno~r	940	.6805638	.3007123	0	1
average_al~s	562	2159.5	1579.938	1	12045
sum_talk_o~k	940	5.698936	4.850678	0	51
prop_talk_~k	940	.5584362	.2974345	0	1
alter_part~s	940	.5234043	.4997178	0	1
sum_alter3~e	940	2.771277	.9513314	0	18
sum_all_comm	940	9.129787	6.524645	0	63
sum_sex30	940	.706383	.9542025	0	10
sum_sex	940	1.094681	1.312905	0	11
serious_se~r	940	.5531915	.4974273	0	1

INDEPENDENT VARIABLES:
SOCIAL NETWORK CHARACTERISTICS

sum_family~e	940	1.765957	3.211818	0	26		sum_enc_ba~e	940	10.04787	13.30365	0	126
sum~s_subuse	940	7.679787	9.609446	0	82		prop_enc_b~e	940	.1049149	.1062042	0	.67
sum~e_subuse	940	2.376596	4.567408	0	45		unprotecte~x	940	.5595745	.4967025	0	1
sum~t_subuse	940	5.303191	7.680147	0	70		sum_alter_~s	940	4.118085	4.632144	0	54
sum_alter_~e	940	11.03617	12.37972	0	117		prop_alter~s	940	.4070638	.3226625	0	1
prop_famil~e	940	.177883	.3368135	0	2.78		sum_family~p	940	3.806383	4.896888	0	40
pro~s_subuse	940	.7319362	.7150964	0	3.92		sum~s_emosup	940	9.419149	11.02543	0	88
pro~e_subuse	940	.2174362	.3730659	0	2.67		sum~e_emosup	940	4.407447	6.901633	0	56
pro~t_subuse	940	.5144468	.6235161	0	3.67		sum~t_emosup	940	5.011702	7.544746	0	69
prop_alter~e	940	1.011138	.7751352	0	4.1		sum_alter_~p	940	12.44043	12.30218	0	96
sum_family~g	940	.1797872	.8050797	0	11		sum_family~t	940	.4223404	1.031625	0	9
s~s_harddrug	940	.9010638	2.277781	0	21		sum_friend..	940	.2553191	.4362726	0	1
s~e_harddrug	940	.2351064	.8750728	0	7		sum_friend..	940	.3989362	1.071163	0	12
s~t_harddrug	940	.6659574	1.88122	0	20		sum_friend..	940	1.207447	2.351115	0	22
sum_alter_~g	940	1.393617	3.174103	0	32		sum_alter_~t	940	2.325532	3.275167	0	25
prop_famil~g	940	.0180638	.0826844	0	1.11		prop_famil~p	940	.4106383	.5347134	0	4
p~s_harddrug	940	.0858617	.1995896	0	1.5		pro~s_emosup	940	.9799468	.872193	0	4
p~e_harddrug	940	.0214255	.0807044	0	1		pro~e_emosup	940	.4481809	.6264129	0	4
p~t_harddrug	940	.0644149	.174216	0	1.5		pro~t_emosup	940	.531766	.6654454	0	4
prop_alter~g	940	.1275	.249768	0	1.5		prop_alter~p	940	1.300128	.9154396	0	4

INDEPENDENT VARIABLES:
SOCIAL NETWORK CHARACTERISTICS

prop_famil~t	940	.0527872	.1369067	0	1
prop_frien..	940	.0441915	.1077581	0	1
prop_frien..	940	.045	.1207718	0	1
prop_frien..	940	.132	.2258234	0	1
prop_alter~t	940	.2503298	.2900802	0	1
sum_alter_~f	940	.6117021	1.22508	0	11
sum_object~e	940	46.1383	37.75649	0	313
prop_objec~e	940	.5598191	.2505718	0	1
sum_enc_go~e	940	6.719149	7.106212	0	63
prop_enc_g~e	940	.643883	.4146918	0	1
sum_alter_~b	940	4.826596	4.934581	0	52
prop_alter~b	940	.4535426	.3024284	0	1
sum_alter_~w	940	2.73617	3.664958	0	32
prop_alter~w	940	.2942766	.3106964	0	1

PROSPECTIVE FAIRNESS VARIABLES

- **Gender Identity**
  - Male (coded 1)
  - Female (coded 2)
  - Transgender Female (coded 3)
  - Transgender Male (coded 4)

- **Sexual Orientation Identity**
  - Gay or Lesbian (coded 1)
  - Queer (coded 2)
  - Bisexual (coded 3)
  - Heterosexual (coded 4)
  - Questioning/Unsure (coded 5)

- **Race/Ethnic Group Identity**
  - American Indian or Alaska Native (coded 1)
  - Asian (coded 2)
  - Black or African American (coded 3)
  - Native Hawaiian or Other Pacific Islander (coded 4)
  - White (coded 5)
  - Latino/Hispanic (coded 6)
  - Mixed race (coded 7)

May need to combine categories to achieve adequate subgroup representation for analysis

MISSING DATA

Variable	Obs=.	Obs>.	Obs<.
age	19		921
gender	18		922
sexori	34		906
raceall	19		921
trauma_sum	54		886
cesd_score	36		904
school	22		918
job	31		909
intertime	29		911
sntime	33		907
interlast	33		907
sex	28		912
sexpart30	88		852
sexpartlife	86		854
concurrent	127		813
exchange	88		852
pregnant	41		899
children	39		901
weapon	34		906
fight	34		906

Variable	Obs=.	Obs>.	Obs<.
fighthurt	44		896
ipv	36		904
alcfirst	35		905
alc30	37		903
alcbinge	47		893
potfirst	39		901
pot30	35		905
jail	59		881
staycurrent	86		854
homelage	116		824
time_homel~h	148		792
traveler	41		899
jailkid	52		888
jailadult	56		884
gettherapy	49		891
suicidea	36		904
suicplan	40		900
suicattempt	34		906
average_al~s	378		562

QUESTIONS...I CAN'T GUARANTEE ANSWERS!

SUICIDE PREVENTION IS EVERYONE'S BUSINESS!

National Suicide Prevention Lifeline: 800-273-8255