

# Analyzing Susceptibility **Health Issues**

# **Meet the Team**

Data Divas







**Steph Yap** Practitioner **Cat Greenman** Practitioner

Margot Herman Practitioner

Whit Brooks Executive







### **Michell Li**

Practitioner

# **Problem Overview**



### The goal of this project and analysis is to **identify** factors that make individuals more susceptible to mental health issues.

# Content Outline

01	Datasets Overview	
02	Data Cleaning Methodology	
03	Data Insights	
04	Data Visualization	
05	Statisticaly Analysis & Findings	
		7
06	Recommendation and Limitation	



# Datasets

### National Survey on Drug Use and Health (NSDUH)

Survey level data on drug use and health (2015 - 2019)

Dataset shape: (214505, 95)

### Health Professional Shortage Area (HPSA)

Clinic level data (inclusive of HPSA ID and Date of Designation/Withdrawal)

### Health Resources and Services Administration (HRSA)

Grants awarded to qualified non-federal entities



### National Survey of Drug Use and Health (NSDUH)

- Race/ethnicity
- Health insurance status and type
- Adult age groups
- Susceptibility to mental illness by demographic subset

### Health Professional Shortage Area (HPSA)

- Metropolitan indicator by state
- The average number of withdrawn days by state
- HPSA status by county



#### ANALYZING SUSCEPTIBILITY TO MENTAL HEALTH ISSUES

In 2019, national spending on mental health services totaled \$225.1 billion and accounted for 5.5% of all health spending. (1) Furthermore, approximately 40% of Americans live in a designated mental health provider shortage area, which exacerbates the problem. Across the US, each state has discretionary funding allocated specifically for mental health. Sufficient funds and effective resource allocation are necessary for the diagnosis and treatment of mental health issues. Mental health issues are pervasive and, now more than ever, need to be better understood to address their causes and impacts in a meaningful way.

The goal of this project was to identify factors that make individuals more susceptible to mental health issues, based on self-administered substance use, demographics, and geographic information from the National Survey on Drug Use and Health (NSDUH).





Num Weeks Mental Health Difficulties



#### High HPSA Scored Areas The Brighter It Is The Higher the HPSA score) vs. Received Mental Health Assistance



# Data Cleaning Methodology

### NSDUH

- Determine relevant columns out of 2000 existing columns
- Refactor column names
- Consolidate non-committal answers

### **HPSA**

- Remove variables with more than 30% data missing
- Remove variables without any variation





#### **Statistical Analysis & Methods**

To analyze the relationship between a participant's serious psychological distress indicator (0=NO/1=YES) in the past month and exposure variables for adults using the NSDUH survey.

- Bivariate Analysis (Chi-Square) test
- SMOTE (Synthetic Minority Oversampling Technique)
- Feature Selection Forward selection, RFE
- Multivariate Analysis/Logistic Regression

County's Population Density Type in NSDUH Survey (2015-2019)

	High susceptibility to mental health issues (n = 16,078); n(%)	Low susceptibility to mental health issues (n= 181,505), n(%)	χ²	p-valu e
Population Density Type			46.83	< 0.01
>=1 Million Persons	7,017 (44)	84,317 (46)	-	-
<1 Million Persons	9,061 (56)	97,188 (54)	-	-

### Table 2. Bivariate Associations Between Susceptibility to Mental Health Issues with

# **Final Model**

**R^2 Score:** 0.5797

Features: 20

**Accuracy:** 0.94 (2.398% above baseline) (2015-2019)

**Outcome: Susceptibility to** Overall\_Health\_Fair\_Poor Education\_Category\_Less\_than Perceived\_Unmet\_Need Adult\_Employment\_Status\_Un Education\_Category\_HS\_Grad Overall\_Health\_Good Treatment\_Type\_Past\_Year\_In Worst\_Psychological\_Distress Adult\_Employment\_Status\_Ot Education\_Category\_Some\_Co Race\_Ethnicity\_Black Gender\_Male Overall\_Health\_Very\_Good Num\_Days\_Skipped\_Work\_Pas Year PDEN10\_Less\_than\_1\_Mil Total\_Income\_Family\_Recode\_ Age\_Category\_Six\_Levels\_35-4 Age\_Category\_Six\_Levels\_26-3 Age\_Category\_Six\_Levels\_50-6 Age\_Category\_Six\_Levels\_65\_A

#### Table 4. Multivariate Logistic Regression Analysis Between Susceptibility to Mental Health Issues with Population Density Type, and Covariates in NSDUH Survey Participants

Mental Health Issues	Coefficient	Adjusted Odds Ratio
	0.9477***	2.58
n_HS	0.6343***	1.89
	0.5933***	1.81
nemployed	0.5616***	1.75
	0.5268***	1.69
	0.4836***	1.62
patient_Only	0.4594***	1.58
_Level	0.3719***	1.45
her	0.3012***	1.35
ollege_Assoc	0.2479***	1.28
	0.2352***	1.27
	0.1904***	1.21
	0.1414***	1.15
st_30_Days	0.0656***	1.07
	-0.0038***	1.00
	-0.0552**	0.95
_75000orMore	-0.1595****	0.85
49	-0.228***	0.80
34	-0.2291***	0.80
64	-0.2628***	0.77
And_Above	-0.3817***	0.68
	-	-



### Likelihood of Having Experienced Serious Psychological Distress In the Past Month Using Adjusted Odds Ratio





Below high school: 1.8x more likely **n Level raduate** High school: 1.6x more likely

College level: 1.2x more likely

to 25	Ages 35 to 49: 20% less likely
	Ages 26 to 34: 20% less likely
	Ages 50 to 64: 23% less likely
	Ages 65+: 32% less likely

# Recommendation



- Combat stigma around mental health
- Targeted distribution of how to get help
- Target younger age groups and people with lower education levels

# Limitations

- Inability to relate geographic HPSA and Grant data to individuals
- Methods should be performed to adjust for confounding

# **Future Work**



Breaking down models by age group and education level

# Thank you!

Questions and comments?