Final Presentation

Team-Foodies

Datasets and Question

- Source: Substance Abuse and Mental Health Data Archives (SAMHDA)
- Primary Dataset: National Mental Health Services Survey
  - 2014, 2015, 2016, 2017, 2018
- Secondary Dataset: National Survey of Substance Abuse Treatment Services
  - 2014, 2015, 2016, 2017, 2018
- Our Goal:
  - Trends among different states and types of facilities
  - How they differ in the services that they offer

TREATMT: Facility offers substance abuse treatment (Q.A1)

Value	Label	Frequency	%
0	No	5,009	42.9%
1	Yes	6,673	57.1%
	Total	*11,682*	*100%*

Variable Type: numeric

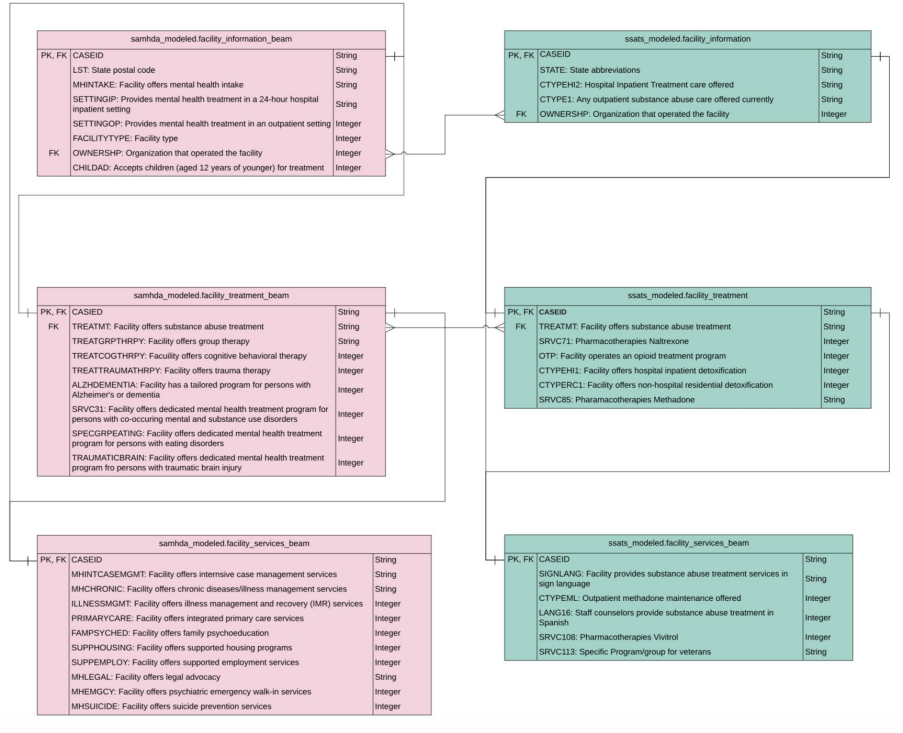
FACILITYTYPE: Facility type (Q.A4)

Value	Label	Frequency	%
1	Psychiatric hospital	692	5.9%
2	Separate inpatient psychiatric unit of a general hospital	1,066	9.1%
3	Residential treatment center for children	580	5.0%
4	Residential treatment center for adults	840	7.2%
5	Other type of residential treatment facility	72	0.6%
6	Veterans Administration medical center (VAMC)	459	3.9%
7	Community mental health center (CMHC)	2,553	21.9%
8	Partial hospitalization/day treatment facility	360	3.1%
9	Outpatient mental health facility	4,665	39.9%
10	Multi-setting mental health facility	382	3.3%
11	Other	13	0.1%
	Total	*11,682*	*100%*

Variable Type: numeric

ERD splits/modeled tables

samhda_staging.samhda_<year>		
PK	CASEID	Integer
	LST: State postal code	String
	MHINTAKE: Facility offers mental health intake	Integer
	TREATMT: Facility offers substance abuse treatment	Integer
	SETTINGIP: Provides mental health treatment in a 24-hour hospital inpatient setting	Integer
	SETTINGOP: Provides mental health treatment in an outpatient setting	Integer
	FACILITYTYPE: Facility type	Integer
	FOCUS: Primary treatment focus of facility	Integer
	OWNERSHP: Organization that operated the facility	Integer
	PUBLICAGENCY: Public agency or department that operated facility	Integer
	TREATPSYCHOTHRPY: Facility offers individual psychotherapy	Integer
	TREATGRPTHRPY: Facility offers group therapy	Integer
	TREATCOGTHRPY: Facuility offers cognitive behavioral therapy	Integer
	TREATTRAUMATHRPY: Facility offers trauma therapy	Integer
	MHINTCASEMGMT: Facility offers internsive case management services	Integer
	MHCHRONIC: Facility offers chronic diseases/illness management servcies	Integer
	ILLNESSMGMT: Facility offers illness management and recovery (IMR) services	Integer
	PRIMARYCARE: Facility offers integrated primary care services	Integer
	FAMPSYCHED: Facility offers family psychoeducation	Integer
	SUPPHOUSING: Facility offers supported housing programs	Integer
	SUPPEMPLOY: Facility offers supported employment services	Integer
	MHLEGAL: Facility offers legal advocacy	Integer
	MHEMGCY: Facility offers psychiatric emergency walk-in services	Integer
	MHSUICIDE: Facility offers suicide prevention services	Integer
	CHILDAD: Accepts children (aged 12 years of younger) for treatment	Integer
	SENIORS: Accepts seniors (ages 65 years and older) for treatment	Integer
	ALZHDEMENTIA: Facility has a tailored program for persons with Alzheimer's or dementia	Integer
	SRVC31: Facility offers dedicated mental health treatment program for persons with co-occuring mental and substance use disorders	Integer
	SPECGRPEATING: Facility offers dedicated mental health treatment program for persons with eating disorders	Integer
	TRAUMATICBRAIN: Facility offers dedicated mental health treatment program fro persons with traumatic brain injury	Integer



Beam Pipeline Transformations

Primary Dataset:

- Facility_information

```
if settingip == 0:
    settingip = 'No'
if settingip == 1:
    settingip = 'Yes'
```

- Facility_services

```
if mhlegal == 0:
    mhlegal = 'No'
if mhlegal == 1:
    mhlegal = 'Yes'
if mhlegal == -1:
    mhlegal = 'Missing'
```

- Facility_treatment

```
if treatgrpthrpy == 0:
    treatgrpthrpy = 'No'
if treatgrpthrpy == 1:
    treatgrpthrpy = 'Yes'
if treatgrpthrpy == -1:
    treatgrpthrpy = 'Missing'
if treatgrpthrpy == -5:
    treatgrpthrpy= 'Refused'
```

Secondary Dataset:

- Facility_information

```
if ctype1 == 0:
    ctype1 = 'No'
if ctype1 == 1:
    ctype1 = 'Yes'
```

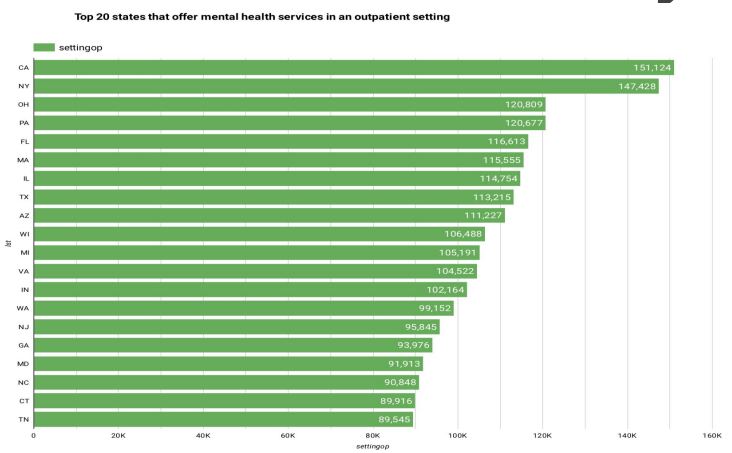
- Facility_services

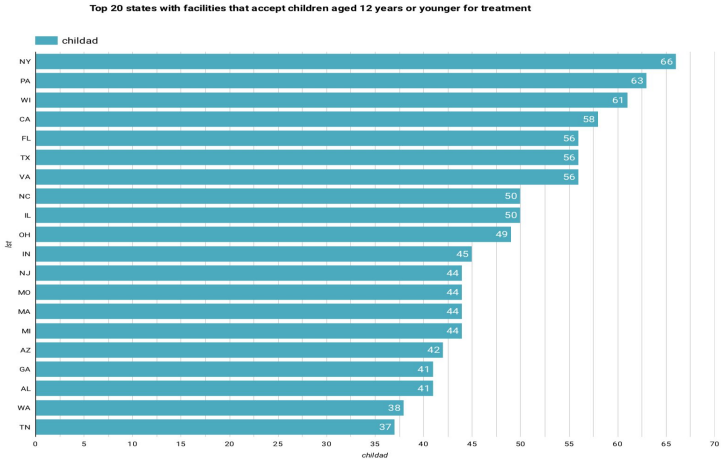
```
if srvc113 == 0:
    srvc113 = 'No'
if srvc113 == 1:
    srvc113 = 'Yes'
if srvc113 == -1:
    srvc113 = 'Missing'
if srvc113 == -2:
    srvc113 = 'Refused'
if srvc113 == -3:
    srvc113 = 'Do not know'
```

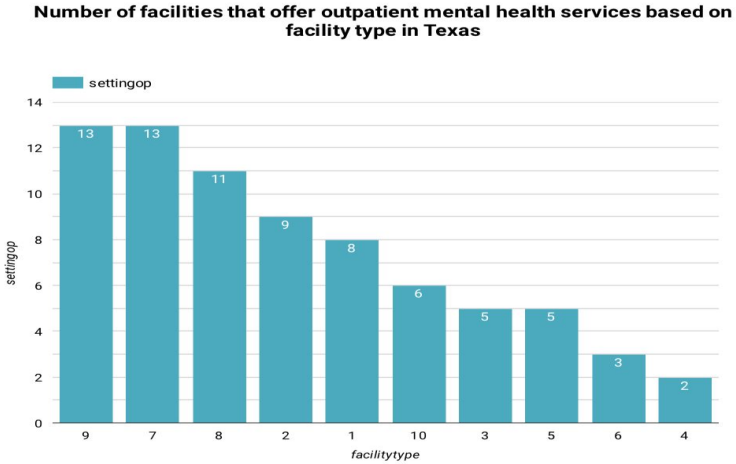
- Facility_treatment

```
if srvc85 == 0:
    srvc85 = 'No'
if srvc85 == 1:
    srvc85 = 'Yes'
if srvc85 == -1:
    srvc85 = 'Missing'
if srvc85 == -2:
    srvc85 = 'Refused'
if srvc85 == -3:
    srvc85 = 'Do not know'
```

Data Studio: Primary Dataset







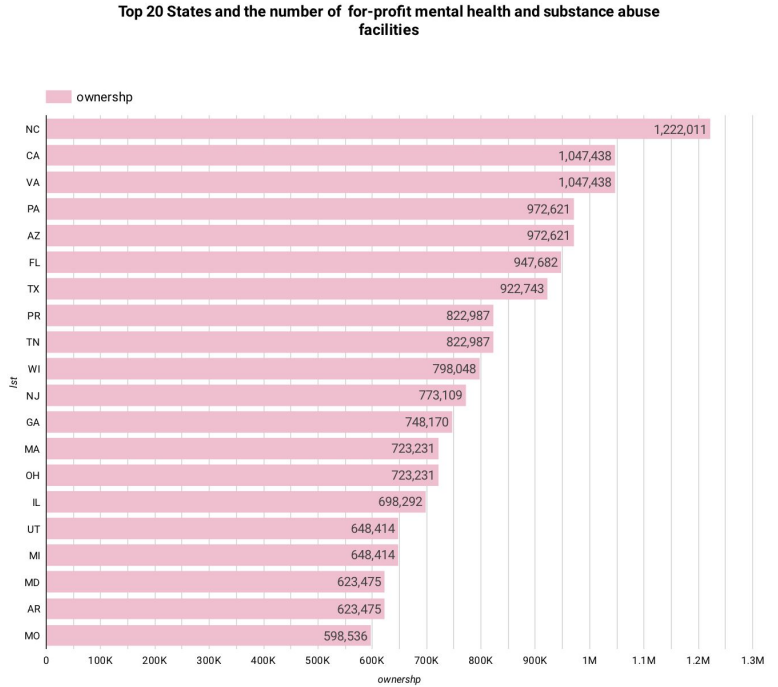
"facilitytype" Key:

1: Psychiatric Hospital
2: Separate inpatient psychiatric unit of gen. hosp.
3: Residential treatment center for children only
4: Residential treatment center for adults only
5: Other type of residential treatment facility
6: veterans Administration medical center (VAMC)
7: Community mental health center (CMHC)
8: Outpatient mental health facility
9: Multi-setting mental health facility
10: Other

Cross-dataset query #1

```
select s1.lst, sum(s1.ownershp) as ownershp,
from samhda_modeled.facility_information_Beam_DF as s1
inner join ssats_modeled.ssats_facility_information_Beam_DF as s2
on s1.ownershp = s2.ownershp where s1.ownershp = 1 and s2.ownershp = 1
group by lst
```

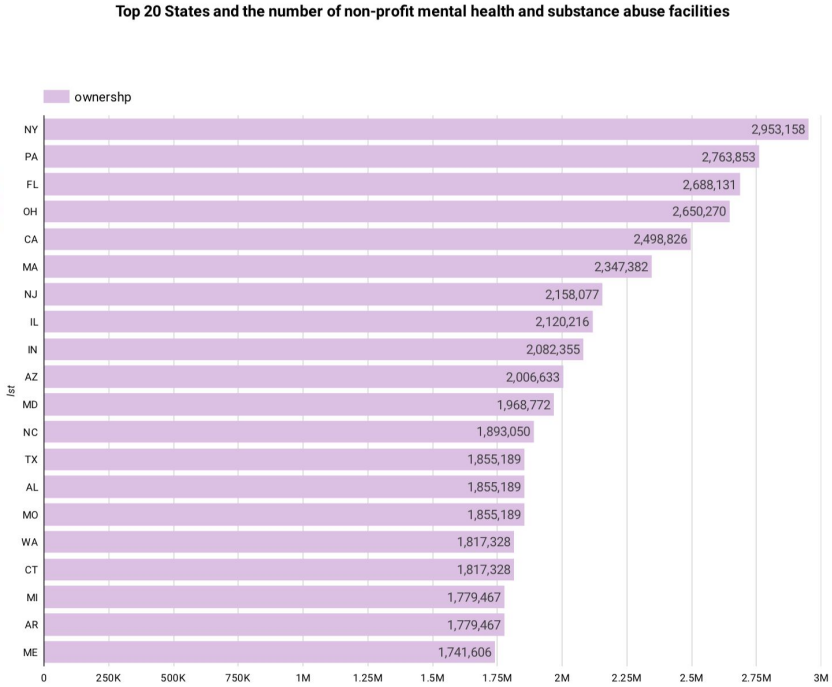
- Joined on ownership of facility
  - Ownership = 1
    - Private for profit
- Grouped by state
- Used sum because data was equal to 1



Cross-dataset query #2

```
select s1.lst, count(*) as ownershp,
from samhda_modeled.facility_information_Beam_DF as s1
inner join ssats_modeled.ssats_facility_information_Beam_DF as s2
on s1.ownershp = s2.ownershp where s1.ownershp = 2 and s2.ownershp = 2
group by lst
```

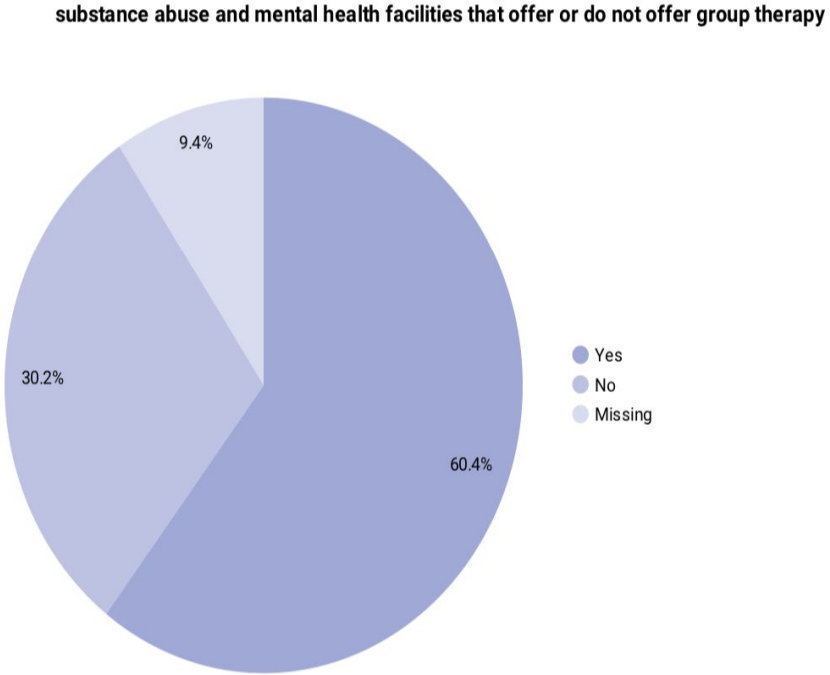
- Joined on ownership
  - Ownership = 2
    - Private non-profit facility
- Grouped by state



Cross-dataset query #3

```
select s1.treatgrpthrpy, count(*) as count,
from samhda_modeled.facility_treatment_Beam_DF as s1
inner join ssats_modeled.ssats_facility_treatment_Beam_DF as s2
on s1.treatmt = s2.treatmt where s1.treatmt = "Yes" and s2.treatmt = "Yes"
group by s1.treatgrpthrpy
```

- Joined on ‘treatmt’ variable which indicates whether or not mental health or substance abuse treatment is offered in that facility
  - treatmt = “yes”
- Grouped by treatgrpthrpy variable from s1
  - Treatgrpthrpy: whether or not the facility offers group therapy



Airflow DAG

```
create_staging >> create_modeled >> branch
load_ssats_2018 >> uuid_2018 >> join
load_ssats_2017 >> uuid_2017 >> join
load_ssats_2016 >> uuid_2016 >>join
load_ssats_2015 >> uuid_2015 >> join
load_ssats_2014 >> uuid_2014 >> join >> branch2
branch >> create_information >> information
branch >> create_services >> services
branch >> create_treatment >> treatment
```

Some issues we ran into while doing the Airflow milestone:

- Syntax error: when casting variables such as “CTYPEHI2” from 0 and 1 to “no” and “yes” we had to change syntax to \“yes\” and \“no\”
- Logic error: we had to use “generate_uuid” to generate the “CASEID” for each year

Future improvements to our solution

- Along with grouping by states, we could've grouped all the facilities by the different services they offer in order to look at any trends those facilities would have had
- Our dataset included all the facilities and services available by state and year, but if there was some way to also see how effective those services and treatments were on each state's population, we could have been able to get even more out of this data (ex: observe the different mental health trends in each state, observe what changes each state could make to accommodate to their population's mental health needs, observe how well each state was doing in terms of addressing its mental health needs, etc)