Characterize a Given OMOP CDM Database

Initial Set-Up

Activating Project Environment

```
knitr::opts_chunk$set(echo = TRUE, message = F, results = 'hold', cache = T)
options(scipen=999)
devtools::load_all()
```

```
## i Loading MentalHealthEquity
```

```
library(MentalHealthEquity)
```

Packages Used in Analysis

```
library(DatabaseConnector)
```

`DatabaseConnector`

Package Description:

This package will be used to create connections to the database and execute queries to obtain characterizations about the cohorts in the database.

Introduction

Background

The intent of this report is to investigate the characterization statement:

> **Characterization Statement 1:** Characterize the individuals being seen for mental health care services (related to depression, bipolar disorder, and suicidality) at least one time – including hospitalization events.

This characterization statement is founded on the central research topic for this study:

> **Research Topic:** Based on [CLAIMS], we see X% of all persons with at least one claim indicating [DEPRESSION/BIPOLAR DISORDER/SUICIDALITY] are not seen again.

By which the phrases "...all persons..." refer to those seen by patience care provider, etc. and "...are not seen again." implies lack of adherence to care.

Data Analysis Preparation

Creating Initial Connection

Defining connection details for connecting to a given database:

```
connectionDetails <- createConnectionDetails(
  dbms = "postgresql",
```

```
  server = "data.hdap.gatech.edu/synpuf_v5",
  user = keyring::key_list("synPUF")[1,2],
  password = keyring::key_get("synPUF", keyring::key_list("synPUF")[1,2]),
  port = 5434,
  pathToDriver = "../utils"
)
```

Creating Initial OMOP Tables

In this case, the schema follows the OMOP CDM v5 schema:

For this analysis, we will work with the following tables from the schema:

Per the [Book of OHDSI], here are the break downs for these tables:

- `PERSON` -
- `LOCATION` -
- `OBSERVATION_PERIOD` -
- `CONDITION_OCCURRENCE` -

General Characterization of Data

```
sql <- "
  SELECT {@x} ? {@table.@concept_name, count(DISTINCT PERSON_ID)} : {*}
  FROM person
      LEFT JOIN (SELECT LOCATION_ID, STATE FROM location) AS location ON
        person.location_id = location.location_id
      LEFT JOIN (SELECT CONCEPT_ID,
                    CONCEPT_NAME as RACE_CONCEPT_NAME
                  FROM concept WHERE DOMAIN_ID = 'Race'
                  AND STANDARD_CONCEPT = 'S') AS race ON
        person.race_concept_id = race.concept_id
      LEFT JOIN ( SELECT CONCEPT_ID,
                    CONCEPT_NAME as ETHNICITY_CONCEPT_NAME
                  FROM concept WHERE DOMAIN_ID = 'Ethnicity'
                    AND STANDARD_CONCEPT = 'S') AS ethnicity ON
        person.ethnicity_concept_id = ethnicity.concept_id
      LEFT JOIN (SELECT CONCEPT_ID,
                    CONCEPT_NAME as GENDER_CONCEPT_NAME
                  FROM concept WHERE DOMAIN_ID = 'Gender'
                  AND STANDARD_CONCEPT = 'S') AS gender ON
        person.gender_concept_id = gender.concept_id
  {@x}? {GROUP BY @table.@concept_name
      ORDER BY COUNT DESC} :
      {LIMIT 10}
"
```

Cohort Populations

General Cohort

Overall Trends

```
conn <- connect(connectionDetails)
date_sql <- "
  SELECT  MIN(visit.visit_end_date) as first_visit,
          MAX(visit.visit_end_date) as last_visit
  FROM visit_occurrence visit
"
dates <- renderTranslateQuerySql(conn, date_sql)
disconnect(conn)
```

The dates in the dataset range from 2008-01-01 and 2010-12-31

Patient Spread   This calculates the number of patients in the dataset:

```
conn <- connect(connectionDetails)
count <- querySql(conn, sql = "SELECT COUNT(DISTINCT PERSON_ID) FROM PERSON")
disconnect(conn)
```

In this dataset, there are 2326856 patients. To further examine this data, we can break them down across the following axes:

- State
- Age
- Race
- Gender
- Care setting

```
conn <- connect(connectionDetails)
state <- renderTranslateQuerySql(conn, sql, table = "location",
                                 concept_name = "state", x = T)
disconnect(conn)
```

State Breakdown   The top states are: CA, FL, TX, NY, PA, IL, OH, MI, NC, NJ

STATE	COUNT
CA	201651
FL	155040
TX	133809
NY	132059
PA	102191
IL	86557
OH	86325
MI	79556
NC	78912
NJ	63475

```
conn <- connect(connectionDetails)
age_sql <- "
```

```
  SELECT  MIN(YEAR(visit.visit_end_date) - person.year_of_birth) as min_age,
          MAX(YEAR(visit.visit_end_date) - person.year_of_birth) as max_age
  FROM person
    JOIN visit_occurrence visit ON
      person.person_id = visit.person_id
"
age <- renderTranslateQuerySql(conn, age_sql)
disconnect(conn)
```

Age Breakdown The age of the patient at the time of visit ranged from 25 years of age to 101 years of age.

```
conn <- connect(connectionDetails)
race <- renderTranslateQuerySql(conn, sql, table = "race",
                                concept_name= "RACE_CONCEPT_NAME", x = T)
disconnect(conn)
```

Racial Breakdown In the dataset, there are 1926708 White persons, 247723 Black or African American persons, 152425 NA persons

RACE_CONCEPT_NAME	COUNT
White	1926708
Black or African American	247723
NA	152425

```
conn <- connect(connectionDetails)
gender <- renderTranslateQuerySql(conn, sql, table = "gender",
                                  concept_name = "GENDER_CONCEPT_NAME", x = T)
disconnect(conn)
```

Gender Breakdown In the dataset, there are 1292861 FEMALE persons, 1033995 MALE persons

GENDER_CONCEPT_NAME	COUNT
FEMALE	1292861
MALE	1033995

```
conn <- connect(connectionDetails)
caresite_sql <- "
  SELECT concept.concept_name, count(*) VISIT_COUNT,
    count(DISTINCT visit.person_id) PERSON_COUNT
  FROM visit_occurrence visit
    JOIN (SELECT CONCEPT_NAME , CONCEPT_ID
          FROM concept
          WHERE DOMAIN_ID = 'Visit'
              AND STANDARD_CONCEPT = 'S') concept ON
      visit.visit_concept_id = concept.concept_id
  GROUP BY concept.concept_name
  ORDER BY count(*) DESC
"
```

```
care_site <- renderTranslateQuerySql(conn, caresite_sql)
disconnect(conn)
```

Care Site Breakdown   In the dataset, there are 1984582 unique patients with an indicated care site. The care site setting Inpatient Visit had 111637582 visits

CONCEPT_NAME	VISIT_COUNT	PERSON_COUNT
Inpatient Visit	111637582	1984582

Inpatient Cohort

Overall Trends

Patient Spread   This calculates the number of patients in the dataset:

```
conn <- connect(connectionDetails)
inpatient_sql <- "
  SELECT @query
  FROM (SELECT *
        FROM visit_occurrence visit
        JOIN
          ( SELECT CONCEPT_NAME , CONCEPT_ID
            FROM concept
            WHERE DOMAIN_ID = 'Visit'
            AND STANDARD_CONCEPT = 'S') concept
        ON visit.VISIT_CONCEPT_ID = concept.CONCEPT_ID
        WHERE LOWER(concept.CONCEPT_NAME) LIKE '%@a%'
        ) VISIT_INFO
  JOIN
    ( SELECT *
      FROM person
        LEFT JOIN (SELECT LOCATION_ID, STATE FROM location) AS location ON
          person.location_id = location.location_id
        LEFT JOIN (SELECT CONCEPT_ID,
                      CONCEPT_NAME as RACE_CONCEPT_NAME
                    FROM concept WHERE DOMAIN_ID = 'Race'
                    AND STANDARD_CONCEPT = 'S') AS race ON
          person.race_concept_id = race.concept_id
        LEFT JOIN ( SELECT CONCEPT_ID,
                      CONCEPT_NAME as ETHNICITY_CONCEPT_NAME
                    FROM concept WHERE DOMAIN_ID = 'Ethnicity'
                      AND STANDARD_CONCEPT = 'S') AS ethnicity ON
          person.ethnicity_concept_id = ethnicity.concept_id
        LEFT JOIN (SELECT CONCEPT_ID,
                      CONCEPT_NAME as GENDER_CONCEPT_NAME
                    FROM concept WHERE DOMAIN_ID = 'Gender'
                    AND STANDARD_CONCEPT = 'S') AS gender ON
          person.gender_concept_id = gender.concept_id) DEMO_INFO
  USING(PERSON_ID)
  {@x} ? {GROUP BY DEMO_INFO.@concept_name
          ORDER BY COUNT DESC}
"
inpatient <- renderTranslateQuerySql(conn,
                                     inpatient_sql,
                                     x = F, a = "inpatient",
                                     query = "COUNT(*) VISIT_COUNT,
                                     COUNT(DISTINCT PERSON_ID) PERSON_COUNT
                                     ")
disconnect(conn)
```

In the inpatient cohort, there are a total of 1984582 distinct patients, and there are a total of 111637582 entries. Compared to the original unique person count in the overall dataset, which is 2326856, 342274 are not accounted for in the inpatient count.

To further examine this data, we can break them down across the following axes:

- State
- Age
- Race
- Gender
- Care setting

```
conn <- connect(connectionDetails)
inpatient_state <- renderTranslateQuerySql(conn, inpatient_sql, x = T, a = "inpatient",
                    query = "DEMO_INFO.state, COUNT(DISTINCT PERSON_ID)",
                    concept_name = "state")
disconnect(conn)
```

State Breakdown   In the inpatient cohort, the top 5 states by person count are

STATE	COUNT
CA	178254
FL	140385
TX	123530
NY	119443
PA	90893
IL	79577
OH	77830
MI	70896
NC	65822
NJ	56971

```
conn <- connect(connectionDetails)
inpatient_age <- renderTranslateQuerySql(conn, inpatient_sql,
                              x = F, a = "inpatient",
                              query = "
            MIN(YEAR(VISIT_INFO.visit_end_date) -
              DEMO_INFO.year_of_birth) as min_age,
            MAX(YEAR(VISIT_INFO.visit_end_date) - DEMO_INFO.year_of_birth)
              as max_age,
            SUM(CASE WHEN YEAR(VISIT_INFO.visit_end_date) -
              DEMO_INFO.year_of_birth < 18 THEN 1 END) AS less_than_18,
            SUM(CASE WHEN YEAR(VISIT_INFO.visit_end_date) -
              DEMO_INFO.year_of_birth BETWEEN 18 AND 24 THEN 1 END) AS age_18_to_24,
            SUM(CASE WHEN YEAR(VISIT_INFO.visit_end_date) -
              DEMO_INFO.year_of_birth BETWEEN 25 AND 34 THEN 1 END) AS age_25_to_35,
            SUM(CASE WHEN YEAR(VISIT_INFO.visit_end_date) -
              DEMO_INFO.year_of_birth BETWEEN 35 AND 44 THEN 1 ELSE 0 END) AS age_35_44,
            SUM(CASE WHEN YEAR(VISIT_INFO.visit_end_date) -
              DEMO_INFO.year_of_birth BETWEEN 45 AND 54 THEN 1 ELSE 0 END) AS age_45_54,
            SUM(CASE WHEN YEAR(VISIT_INFO.visit_end_date) -
              DEMO_INFO.year_of_birth BETWEEN 55 AND 64 THEN 1 ELSE 0 END) AS age_55_64,
            SUM(CASE WHEN YEAR(VISIT_INFO.visit_end_date) -
              DEMO_INFO.year_of_birth >= 65 THEN 1 ELSE 0 END) AS age_65_and_more")
disconnect(conn)
```

Age Breakdown   In the inpatient dataset, it takes in the visit encounters, and the ages for the individuals were found according to the time of the inpatient visit. The youngest individual was of age 25, and the

oldest individual was of age 101

This table shows the number of occurences of the ages in the data set.

	COUNT
MIN_AGE	25
MAX_AGE	101
LESS_THAN_18	NA
AGE_18_TO_24	NA
AGE_25_TO_35	1297697
AGE_35_44	2581211
AGE_45_54	5357835
AGE_55_64	7806974
AGE_65_AND_MORE	94593865

```
conn <- connect(connectionDetails)
inpatient_race <- renderTranslateQuerySql(conn, inpatient_sql,
                                x = T,
                                a = "inpatient",
                                concept_name = "RACE_CONCEPT_NAME",
                                query = "
                                DEMO_INFO.RACE_CONCEPT_NAME,
                                COUNT(DISTINCT PERSON_ID)")
disconnect(conn)
```

Racial Breakdown   In the inpatient cohort, there are 1664596 White persons, 198723 Black or African American persons, 121263 NA persons

RACE_CONCEPT_NAME	COUNT
White	1664596
Black or African American	198723
NA	121263

```
conn <- connect(connectionDetails)
inpatient_gender <- renderTranslateQuerySql(conn, inpatient_sql,
                    x = T, a = "inpatient",
                    concept_name = "GENDER_CONCEPT_NAME",
                    query = "
                    DEMO_INFO.GENDER_CONCEPT_NAME,
                    COUNT(DISTINCT PERSON_ID)
                    ")
disconnect(conn)
```

Gender Breakdown   In the inpatient cohort, there are 1131807 FEMALE persons, 852775 MALE persons

GENDER_CONCEPT_NAME	COUNT
FEMALE	1131807
MALE	852775

General Trends for Bipolar Disorder

```r
bipolar_id <- 436665
bipolar_df <- get_atlas_concept_set(get_atlas_concept(bipolar_id))
bipolar_concept <- paste0('(', paste(bipolar_df$CONCEPT_ID, collapse = ", "), ')')
condition_sql <- "
  SELECT {@x} ?
      { person.PERSON_ID,
        person.YEAR_OF_BIRTH,
        person.MONTH_OF_BIRTH,
        person.DAY_OF_BIRTH,
        location.STATE,
        race.RACE_CONCEPT_NAME,
        ethnicity.ETHNICITY_CONCEPT_NAME,
        gender.GENDER_CONCEPT_NAME,
        condition.CONDITION_OCCURRENCE_ID,
        condition.CONDITION_CONCEPT_ID,
        condition.CONDITION_CONCEPT_NAME,
        condition.CONDITION_START_DATE,
        condition.CONDITION_END_DATE,
        condition.VISIT_OCCURRENCE_ID,
        visit_info.VISIT_CONCEPT_NAME} :
      {@query}
  FROM person
      LEFT JOIN (SELECT LOCATION_ID, STATE FROM location) AS location ON
        person.location_id = location.location_id
      LEFT JOIN (SELECT CONCEPT_ID,
                    CONCEPT_NAME as RACE_CONCEPT_NAME
                  FROM concept WHERE DOMAIN_ID = 'Race'
                  AND STANDARD_CONCEPT = 'S') AS race ON
        person.race_concept_id = race.concept_id
      LEFT JOIN ( SELECT CONCEPT_ID,
                    CONCEPT_NAME as ETHNICITY_CONCEPT_NAME
                  FROM concept WHERE DOMAIN_ID = 'Ethnicity'
                    AND STANDARD_CONCEPT = 'S') AS ethnicity ON
        person.ethnicity_concept_id = ethnicity.concept_id
      LEFT JOIN ( SELECT CONCEPT_ID,
                  CONCEPT_NAME as GENDER_CONCEPT_NAME
                  FROM concept WHERE DOMAIN_ID = 'Gender'
                      AND STANDARD_CONCEPT = 'S') AS gender ON
        person.gender_concept_id = gender.concept_id
      LEFT JOIN ( SELECT *
                  FROM CONDITION_OCCURRENCE
                  LEFT JOIN
                    ( SELECT CONCEPT_ID,
                      CONCEPT_NAME as CONDITION_CONCEPT_NAME
                      FROM concept WHERE DOMAIN_ID = 'Condition'
                          AND STANDARD_CONCEPT = 'S') concept ON
                  CONDITION_OCCURRENCE.CONDITION_CONCEPT_ID = concept.CONCEPT_ID) condition ON
        person.PERSON_ID = condition.PERSON_ID
      LEFT JOIN ( SELECT *
                  FROM VISIT_OCCURRENCE visit
                  JOIN
                        ( SELECT concept.CONCEPT_NAME AS VISIT_CONCEPT_NAME,
```

```
                         concept.CONCEPT_ID
                       FROM concept
                       WHERE DOMAIN_ID = 'Visit'
                       AND STANDARD_CONCEPT = 'S') concept
                 ON visit.VISIT_CONCEPT_ID = concept.CONCEPT_ID) as visit_info ON
                 person.PERSON_ID = visit_info.PERSON_ID AND
                 condition.VISIT_OCCURRENCE_ID = visit_info.VISIT_OCCURRENCE_ID
  WHERE condition.CONDITION_CONCEPT_ID in @ids
  {@y}? {@query2}
"
conn <- connect(connectionDetails)
bipolar_counts <- renderTranslateQuerySql(conn, condition_sql,
                        ids = bipolar_concept, x = F, y = F,
                        query = "COUNT(*) AS RECORDS, COUNT(DISTINCT person.PERSON_ID) AS COUNT")
disconnect(conn)
```

Patient Spread   This calculates the number of inpatient patients in the dataset:

In this dataset, there are a total of 1298679 records of bipolar disorder in the dataset, with 524063 patients

To further examine this data, we can break them down across the following axes:

- Condition
- State
- Age
- Race
- Gender
- Care setting

```
conn <- connect(connectionDetails)
bipolar_conditions <- renderTranslateQuerySql(conn, condition_sql, ids = bipolar_concept, x = F,
                                        query = " condition.CONDITION_CONCEPT_NAME, COUNT(*) RECO
disconnect(conn)
```

Condition Breakdown   Here is the breakdown of the top bipolar diagnoses:

CONDITION_CONCEPT_NAME	RECORD_COUNT	PERSON_COUNT
Bipolar disorder	274467	193040
Bipolar I disorder	88224	68356
Depressed bipolar I disorder	63863	49923
Bipolar affective disorder, current episode manic	51860	40876
Bipolar affective disorder, current episode mixed	51414	40581
Severe depressed bipolar I disorder without psychotic features	51398	38381

```
conn <- connect(connectionDetails)
bipolar_state <- renderTranslateQuerySql(conn, condition_sql, ids = bipolar_concept, x = F,
                    query = "location.STATE, COUNT(DISTINCT person.PERSON_ID)",
                    y = T, query2 = "GROUP BY location.STATE ORDER BY COUNT DESC")
disconnect(conn)
```

State Breakdown   The top ten states by person count with bipolar disorder are

STATE	COUNT
CA	44093
FL	36922
TX	34627
NY	31864
PA	23009
IL	22406
OH	20921
MI	19052
NC	18158
NJ	15820

```
conn <- connect(connectionDetails)
bipolar_age <- renderTranslateQuerySql(conn, condition_sql, ids = bipolar_concept, x = F,
query = "
                MIN(YEAR(condition.CONDITION_START_DATE)- person.YEAR_OF_BIRTH)
                  as min_age,
                MAX(YEAR(condition.CONDITION_START_DATE) - person.YEAR_OF_BIRTH)
                  as max_age,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH < 18 THEN 1 END) AS less_than_18,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH BETWEEN 18 AND 24 THEN 1 END)
                  AS age_18_to_24,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH BETWEEN 25 AND 34 THEN 1 END)
                  AS age_25_to_35,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -              pe
                  AS age_35_44,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH BETWEEN 45 AND 54 THEN 1 ELSE 0 END)
                  AS age_45_54,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -              pe
                  AS age_55_64,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH >= 65 THEN 1 ELSE 0 END)
                  AS age_65_and_more",
y = F)
disconnect(conn)
```

Age Breakdown

	COUNT
MIN_AGE	24
MAX_AGE	101
LESS_THAN_18	NA
AGE_18_TO_24	2
AGE_25_TO_35	30143
AGE_35_44	53704
AGE_45_54	98514
AGE_55_64	116770

	COUNT
AGE__65__AND__MORE	999546

```
conn <- connect(connectionDetails)
bipolar_race <- renderTranslateQuerySql(conn, condition_sql, ids = bipolar_concept,
                                x = F, query = "race.RACE_CONCEPT_NAME, COUNT(DISTINCT person.PE
disconnect(conn)
```

Racial Breakdown

RACE__CONCEPT__NAME	COUNT
White	437760
Black or African American	54827
NA	31476

```
conn <- connect(connectionDetails)
bipolar_gender <- renderTranslateQuerySql(conn, condition_sql, ids = bipolar_concept,
                                x = F, query = "gender.GENDER_CONCEPT_NAME, COUNT(DISTINCT perso
disconnect(conn)
```

Gender Breakdown

GENDER__CONCEPT__NAME	COUNT
FEMALE	309470
MALE	214593

```
conn <- connect(connectionDetails)
bipolar_caresite <- renderTranslateQuerySql(conn, condition_sql, ids = bipolar_concept, x= F, query = "
disconnect(conn)
```

Care Site Breakdown

VISIT__CONCEPT__NAME	COUNT
Inpatient Visit	1298679

General Trends for Depression

Patient Spread   This calculates the number of patients in the dataset:

```
depression_id <- 440383
depression_df <- get_atlas_concept_set(get_atlas_concept(depression_id))
depression_concept <- paste0('(', paste(depression_df$CONCEPT_ID, collapse = ", "), ')')

conn <- connect(connectionDetails)
depression_counts <- renderTranslateQuerySql(conn, condition_sql,
                    ids = depression_concept, x = F, y = F,
                    query = "COUNT(*) AS RECORDS, COUNT(DISTINCT person.PERSON_ID) AS COUNT")
disconnect(conn)
```

In this dataset, there are a total of 1697749 records for 599108 patients. To further examine this data, we can break them down across the following axes:

- Condition
- State
- Age
- Race
- Gender
- Care setting

```
conn <- connect(connectionDetails)
depression_conditions <- renderTranslateQuerySql(conn, condition_sql, ids = depression_concept, x = F,
                                        query = " condition.CONDITION_CONCEPT_NAME, COUNT(*) RECO
disconnect(conn)
```

Condition Breakdown   Here is the breakdown of the top bipolar diagnoses:

CONDITION_CONCEPT_NAME	RECORD_COUNT	PERSON_COUNT
Recurrent major depressive episodes, moderate	490503	243514
Dysthymia	277263	205157
Recurrent major depressive episodes	154014	104917
Single major depressive episode	142648	110259
Recurrent major depressive episodes, severe, with psychosis	104429	69300
Recurrent major depressive episodes, in full remission	75501	63448

```
conn <- connect(connectionDetails)
depression_state <- renderTranslateQuerySql(conn, condition_sql, ids = depression_concept, x = F,
                        query = "location.STATE, COUNT(DISTINCT person.PERSON_ID)",
                        y = T, query2 = "GROUP BY location.STATE ORDER BY COUNT DESC")
disconnect(conn)
```

State Breakdown

STATE	COUNT
CA	50376
FL	42229
TX	39464
NY	36449
PA	26270
IL	25572
OH	23998
MI	22051
NC	20625
NJ	18381

```
conn <- connect(connectionDetails)
depression_age <- renderTranslateQuerySql(conn, condition_sql, ids = depression_concept, x = F,
query = "
            MIN(YEAR(condition.CONDITION_START_DATE)- person.YEAR_OF_BIRTH)
```

```
                  as min_age,
                MAX(YEAR(condition.CONDITION_START_DATE) - person.YEAR_OF_BIRTH)
                  as max_age,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH < 18 THEN 1 END) AS less_than_18,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH BETWEEN 18 AND 24 THEN 1 END)
                  AS age_18_to_24,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH BETWEEN 25 AND 34 THEN 1 END)
                  AS age_25_to_35,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -            per
                  AS age_35_44,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH BETWEEN 45 AND 54 THEN 1 ELSE 0 END)
                  AS age_45_54,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -            per
                  AS age_55_64,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH >= 65 THEN 1 ELSE 0 END)
                  AS age_65_and_more",
y = F)
disconnect(conn)
```

Age Breakdown

	COUNT
MIN_AGE	25
MAX_AGE	101
LESS_THAN_18	NA
AGE_18_TO_24	NA
AGE_25_TO_35	33096
AGE_35_44	62065
AGE_45_54	117171
AGE_55_64	145651
AGE_65_AND_MORE	1339766

```
conn <- connect(connectionDetails)
depression_race <- renderTranslateQuerySql(conn, condition_sql, ids = depression_concept,
                              x = F, query = "race.RACE_CONCEPT_NAME, COUNT(DISTINCT person.Pl
disconnect(conn)
```

Racial Breakdown

RACE_CONCEPT_NAME	COUNT
White	502259
Black or African American	61460
NA	35389

```
conn <- connect(connectionDetails)
depression_gender <- renderTranslateQuerySql(conn, condition_sql, ids = depression_concept,
                                             x = F, query = "gender.GENDER_CONCEPT_NAME, COUNT(DISTINCT perso
disconnect(conn)
```

Gender Breakdown

GENDER_CONCEPT_NAME	COUNT
FEMALE	352381
MALE	246727

```
conn <- connect(connectionDetails)
depression_caresite <- renderTranslateQuerySql(conn, condition_sql, ids = depression_concept, x= F, quer
disconnect(conn)
```

Care Site Breakdown

VISIT_CONCEPT_NAME	COUNT
Inpatient Visit	1697749

General Trends for Suicidality

Patient Spread This calculates the number of patients in the dataset:

```
suicidality_id <- 4273391
suicidality_df <- get_atlas_concept_set(get_atlas_concept(suicidality_id))
suicidality_concept <- paste0('(', paste(suicidality_df$CONCEPT_ID, collapse = ", "), ')')

conn <- connect(connectionDetails)
suicidality_counts <- renderTranslateQuerySql(conn, condition_sql,
                    ids = suicidality_concept, x = F, y = F,
                    query = "COUNT(*) AS RECORDS, COUNT(DISTINCT person.PERSON_ID) AS COUNT")
disconnect(conn)
```

In this dataset, there are a total of 29342 records for 27914 patients.

To further examine this data, we can break them down across the following axes:

- Condition
- State
- Age
- Race
- Gender
- Care setting

```
conn <- connect(connectionDetails)
suicidality_conditions <- renderTranslateQuerySql(conn, condition_sql, ids = suicidality_concept, x = F
                                             query = " condition.CONDITION_CONCEPT_NAME, COUNT(*) RECOI
disconnect(conn)
```

Condition Breakdown Here is the breakdown of the top bipolar diagnoses:

CONDITION_CONCEPT_NAME	RECORD_COUNT	PERSON_COUNT
Suicidal thoughts	29342	27914

```r
conn <- connect(connectionDetails)
suicidality_state <- renderTranslateQuerySql(conn, condition_sql, ids = suicidality_concept, x = F,
                query = "location.STATE, COUNT(DISTINCT person.PERSON_ID)",
                y = T, query2 = "GROUP BY location.STATE ORDER BY COUNT DESC")
disconnect(conn)
```

State Breakdown

STATE	COUNT
CA	2331
FL	1934
TX	1869
NY	1649
IL	1233
PA	1185
OH	1170
MI	1011
NC	955
NJ	866

```r
conn <- connect(connectionDetails)
suicidality_age <- renderTranslateQuerySql(conn, condition_sql, ids = suicidality_concept, x = F,
query = "
                MIN(YEAR(condition.CONDITION_START_DATE)- person.YEAR_OF_BIRTH)
                  as min_age,
                MAX(YEAR(condition.CONDITION_START_DATE) - person.YEAR_OF_BIRTH)
                  as max_age,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH < 18 THEN 1 END) AS less_than_18,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH BETWEEN 18 AND 24 THEN 1 END)
                  AS age_18_to_24,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH BETWEEN 25 AND 34 THEN 1 END)
                  AS age_25_to_35,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -                       per
                  AS age_35_44,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH BETWEEN 45 AND 54 THEN 1 ELSE 0 END)
                  AS age_45_54,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -                       per
                  AS age_55_64,
                SUM(CASE WHEN YEAR(condition.CONDITION_START_DATE) -
                  person.YEAR_OF_BIRTH >= 65 THEN 1 ELSE 0 END)
                  AS age_65_and_more",
y = F)
disconnect(conn)
```

Age Breakdown

	COUNT
MIN_AGE	25
MAX_AGE	101
LESS_THAN_18	NA
AGE_18_TO_24	NA
AGE_25_TO_35	648
AGE_35_44	1207
AGE_45_54	2171
AGE_55_64	2676
AGE_65_AND_MORE	22640

```
conn <- connect(connectionDetails)
suicidality_race <- renderTranslateQuerySql(conn, condition_sql, ids = suicidality_concept,
                          x = F, query = "race.RACE_CONCEPT_NAME, COUNT(DISTINCT person.PI
disconnect(conn)
```

Racial Breakdown

RACE_CONCEPT_NAME	COUNT
White	23094
Black or African American	3129
NA	1691

```
conn <- connect(connectionDetails)
suicidality_gender <- renderTranslateQuerySql(conn, condition_sql, ids = suicidality_concept,
                          x = F, query = "gender.GENDER_CONCEPT_NAME, COUNT(DISTINCT perso
disconnect(conn)
```

Gender Breakdown

GENDER_CONCEPT_NAME	COUNT
FEMALE	16406
MALE	11508

```
conn <- connect(connectionDetails)
suicidality_caresite <- renderTranslateQuerySql(conn, condition_sql, ids = suicidality_concept, x= F, q
disconnect(conn)
```

Care Site Breakdown

VISIT_CONCEPT_NAME	COUNT
Inpatient Visit	29342