A Comparative Text Analysis of DSM Volumes 1-5

Initial Insights and Observations

Analise Bottinger, Ryan Costa, Justin Guthrie, Tyler Nguyen, Josh Yu

Wordcloud

- Diagnostic and Statistical Manual of Mental Disorders (DSM)
  - 1952
  - 1968
  - 1980
  - 1994
  - 2013
- Published by American Psychiatric Association (APA)
  - Originally a standardized classification system
  - Now used to guide diagnoses and treatments
- Language/Classification changes throughout its history
  - Homosexuality (sociopathic personality disturbance)
  - Autism Spectrum Disorder (childhood schizophrenic reaction)
  - Classification of disorder severity

DSM-I

DSM-II

DSM-III

DSM-IV

DSM-5*

Sankey Visualization
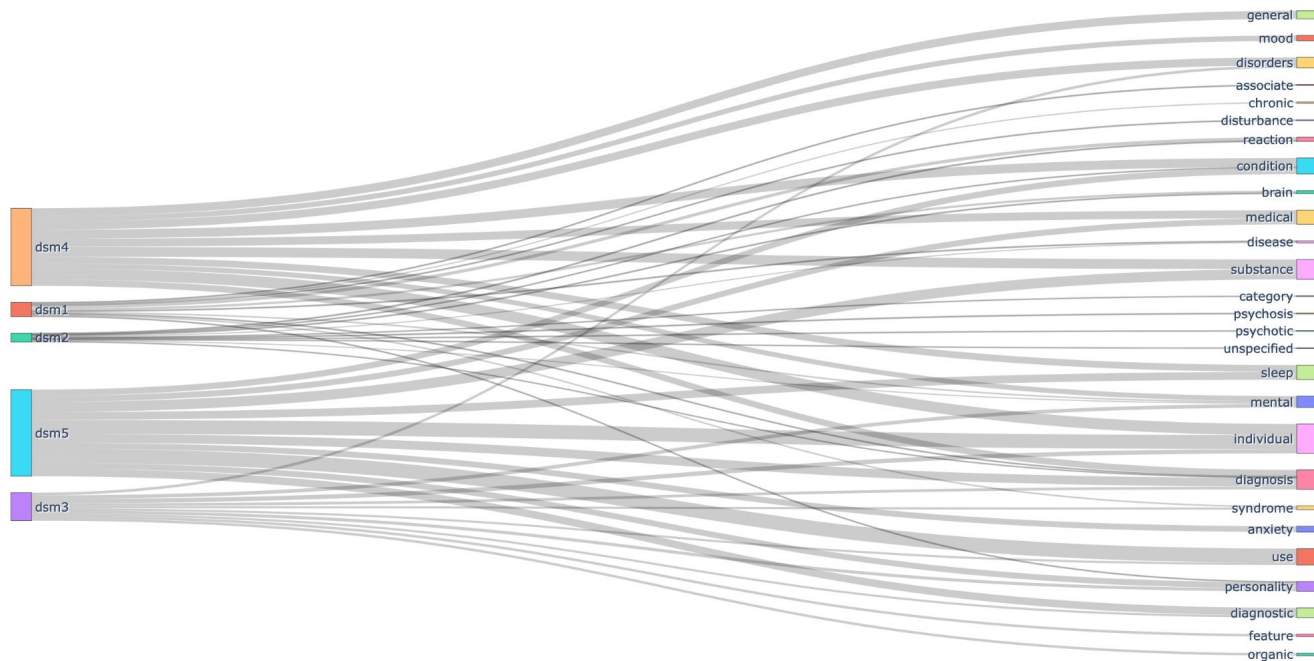
Purpose: Sankey visualization shows relationships between different DSM versions as whole texts

Important notes: "individual", "diagnosis", "substance" and "sleep" are some of the most common words

"individual": as shared mental disorders are not common, it makes sense individual is used frequently to describe individual cases (especially since they vary between people)

"diagnosis": this is expected as these disorders are diagnosed, and the novels are made for this purpose. This is also the only word shared among all five versions

"substance": this is interesting because this word is most common in the later versions of the DSM (4 and 5), which is likely due to the increased use of pharmaceuticals in helping resolve these disorders

"sleep": this is interesting because this word is only used in DSMs 4 and 5, which could indicate the discovery of sleep-related disorders like narcolepsy and sleep apnea

Specific Disorder Analysis

- Performed Topic Modeling (not shown here) and created Co-Occurrence Network Graphs for depressive disorders, anxiety disorders, and schizophrenia specific disorders
- Network graphs show the evolution of co-occurring words over time with each new edition of DSM
    - Do the networks become more or less complex?
    - Are there new symptoms introduced? Are certain symptoms/disorders less prevalent in newer editions?
    - How does the co-occurrence of common words evolve over time?
- Edges in network graph were filtered using arbitrary edge-weight threshold (k=0.00095), can be changed to make graph more or less comprehensive but could lead to cluttering
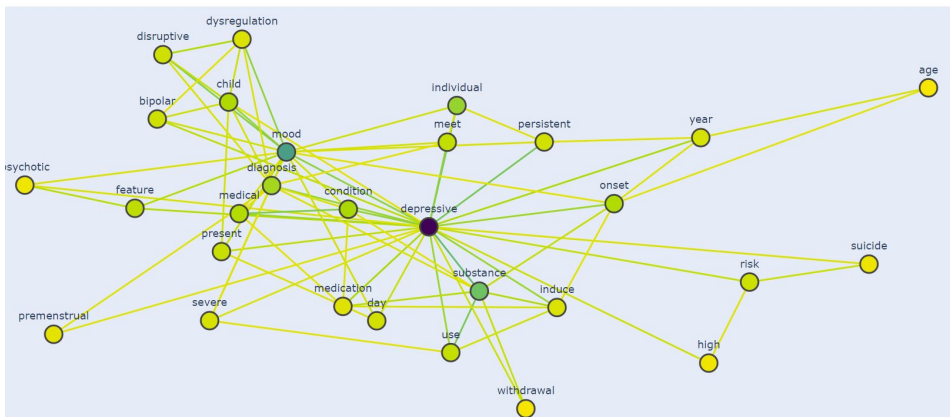
Depression Specific Disorders

Co-occurrence Network Graph for DSM-3 Depression
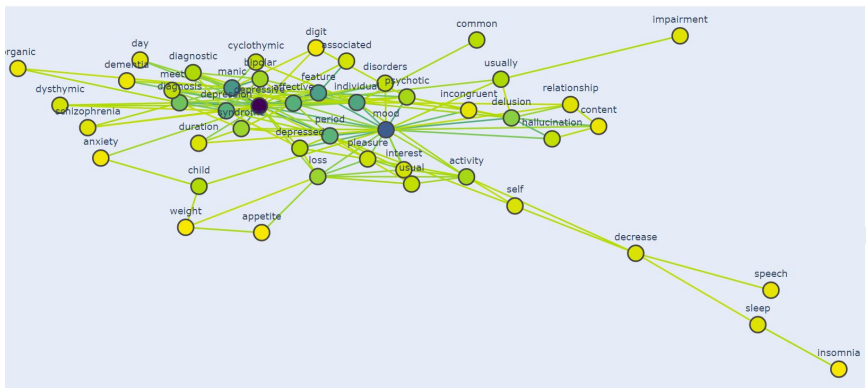
Co-occurrence Network Graph for DSM-4 Depression
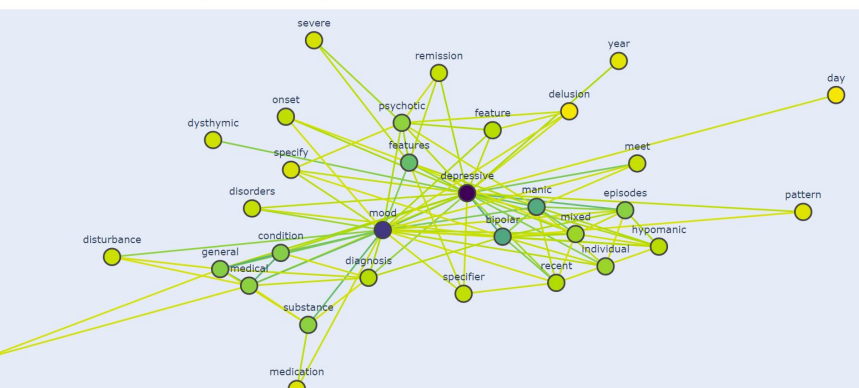
Co-occurrence Network Graph for DSM-5 Depression

Observations
- DSM 5 has a less complex network
  - Could indicate more focused language around depressive disorders (Originally called "mood disorders" in DSM 3)
- DSM 5 new words: "Premenstrual" could indicate increasing focus on resulting hormonal changes in women, increasing focus on "substance"/substance use
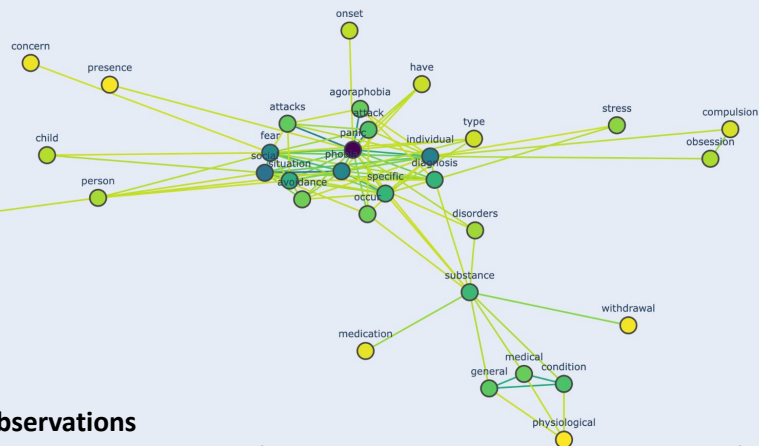
Anxiety Specific Disorders

DSM 4

DSM 3
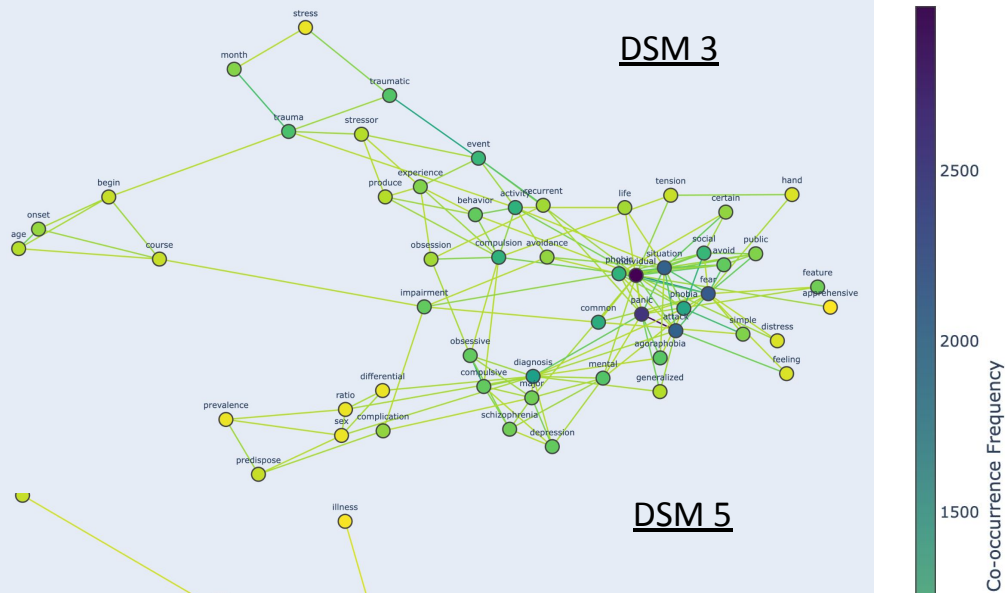
DSM 5

Observations
1. DSM 3 – A lot of co-occurrence, similar to what we saw for depression
    a. little to no dedicated language to describe anxiety, very general and misrepresentative terms
2. DSM 4 – complex network but smaller than DSM3, could indicate a widened understanding of anxiety and an increased effort to properly diagnose/define it
    a. Panic strongly co-occurs with specific disorders now, language is much less subjective
3. DSM 5 – Slightly less complex network, refining of language surrounding anxiety
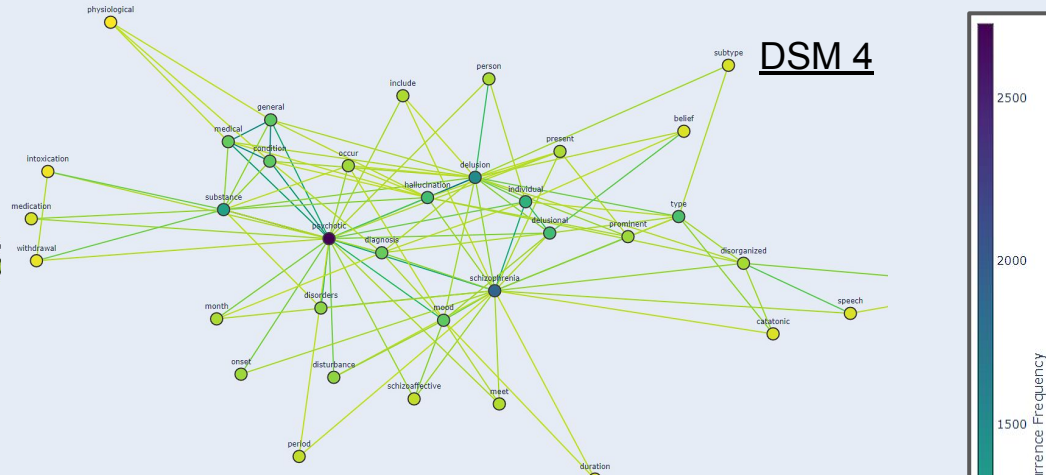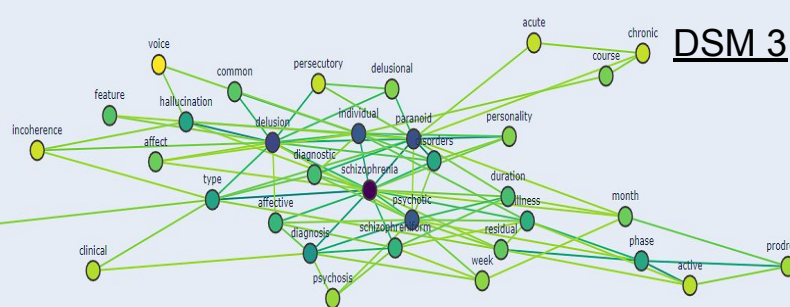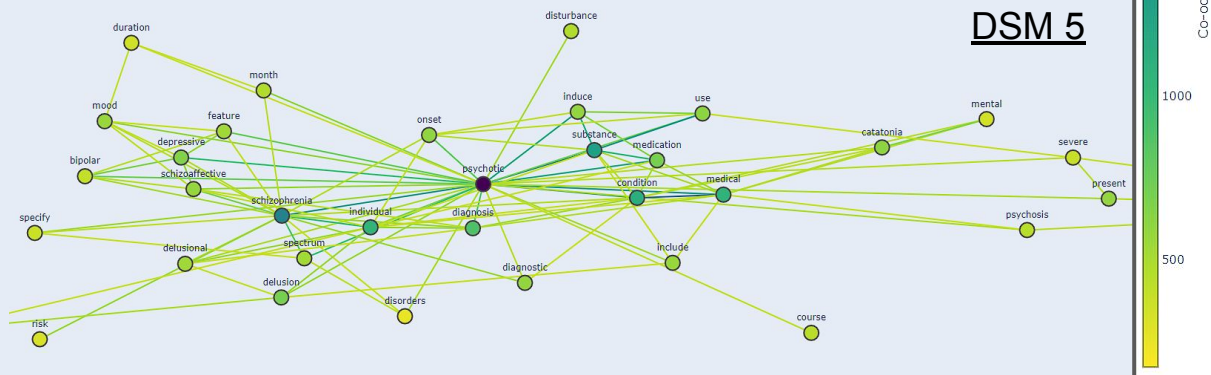   a. Strong connection of agoraphobia/diagnosis to panic
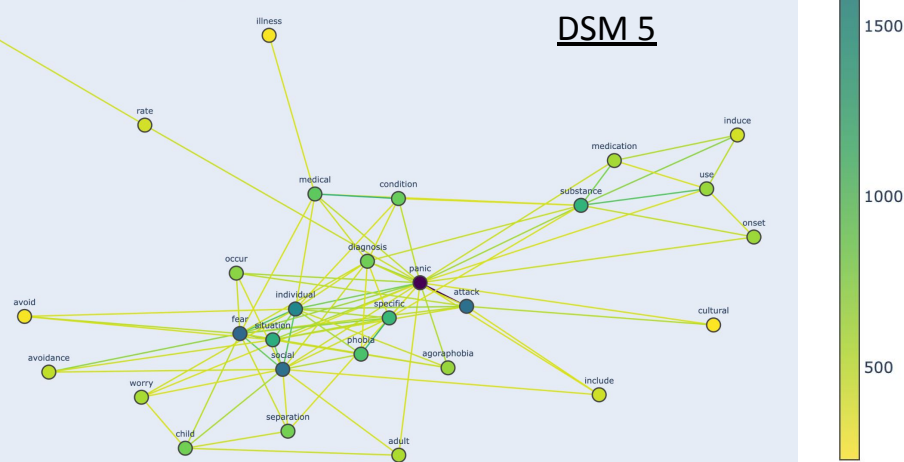
Schizophrenia Specific Disorders

DSM 3

DSM 4

DSM 5

Observations:

1. DSM 3, 4, and 5 display fairly similar webs featuring large focus on "psychotic" and "schizophrenia"
   a. 5 features a greater stress on substance induced psychosis
   b. "Spectrum" appears in 5
   c. 3 features paranoia and delusions in equal frequency that of psychotic
2. DSM 1 and 2 are significantly smaller than successors
   a. First installment appears to stress the "reaction"
   b. 2 elaborates on related depressive disorders (relocated in future editions)