Pandemic Anxiety & Depression

Chan, Natural

25 January, 2022

Abstract

The goal of this project is to identify groups that are most likely to experience symptoms of anxiety or depression. We will primarily observe demographics and reports of symptoms from the COVID-19 pandemic. From there, we will sample our data and analyze different demographics to identify groups with a significantly higher likelihood of suffering from symptoms. Once we have identified key groups, we will be able to create a plan to raise mental health awareness within those groups and therefore increase Prozac and Zoloft prescriptions.

1   Design

The Anxiety & Depression Association of America (ADAA) estimates that approximately 18% of adults suffer from anxiety and only 37% receive treatment. The Depression and Bipolar Support Alliance (DBSA) estimates that major depressive disorder affects about 7% of adults in the U.S.

Pfizer manufactures both anti-anxiety medication (Xanax) and anti-depressant medication (Zoloft). They would like to raise mental health awareness and in turn, increase prescriptions for both Xanax and Zoloft. To do this, they would like to know which groups they should focus their marketing on - more specifically, which groups would benefit the most from mental health awareness. To this end, we would like to answer two key questions:

- What groups of people are most likely to experience anxiety and/or depression?

- Are there any factors due to the COVID-19 pandemic that seems to be correlated with symptoms of anxiety or depression?

We will answer these questions in two ways. We will analyze data on individuals experiencing symptoms of anxiety or depression which will help us identify what demographics they belong to. Also, we can create a classification model to predict how likely an individual is to experience anxiety or depression. Ultimately, this will allow us to create an avenue to raise mental health awareness by potentially allowing people to take a survey and receive information on how likely they are to experience symptoms.

2   Data

The CDC released the Household Pulse Survey to gauge how the pandemic has affected households within the U.S. The survey has been taking place from April 2020 to the present, but due to the survey changing every few weeks, we were only able to make use of the weeks from April 2020 to July 2021. Also, since the survey kept changing, we had to limit the questions to ones that appeared in every weekly data set. Since each week contained anywhere between $30,000$ and $200,000$ observations, we had to take a random sample of 1% from each week to maintain a manageable number of rows while also maintaining any time related trends. Each observation contained data on an individual's demographic, pandemic related situations that they are facing, and frequency of anxiety/depression symptoms. The survey does not explicitly ask for a diagnosis of anxiety or depression disorders, so we must generate these labels ourselves using the frequency of symptoms as well as the statistics that we mentioned earlier. Note that these labels are not a diagnosis, but simply labels that we can use for our analysis.

3   Algorithms

3.1   Data Cleaning/Preprocessing

- Identifying approximately 18% of people that have experienced the most symptoms of anxiety to create labels

- Identifying approximately 7% of people that have experienced the most symptoms of depression to create labels

- Removal of observations with too many missing values

- Removal of any columns that we would not have when marketing

3.2   Analysis

Once our data was cleaned, we looked at various trends regarding anxiety and depression:

- The most significant trend when looking between each week was a spike in symptoms during December 2020 and January 2021

- Females are more likely than males to experience symptoms

- Lower income individuals see a significantly higher occurrence of anxiety and depression

- There is a strong relationship between symptoms and age, where the younger the individual, the more likely they are to experience symptoms

Other factors that seemed to be related to anxiety and depression symptoms include:

- Unemployment

- Uncertainty of paying rent/mortgage

- No home ownership

When seeking a solution, we need to keep in mind that there must be a large enough discrepancy between the groups we are targeting. While targeting certain demographics is important for marketing, we must also avoid as much bias as possible when raising mental health awareness. With this in mind, we ultimately look at two groups: Low income and younger individuals.

When looking at low income individuals, we also see that people between 25-64 make up a significant percentage of those likely to suffer from anxiety or depression. This means that Millennials and Gen Xers should be our primary targets for marketing. To target both groups while minimizing costs, we should consider behaviour from both groups. Social media is a key link between these two groups. To minimize negative perceptions towards Xanax and Zoloft, the marketing would need to be in the form of mental health awareness. First, Millennials tend to react better when there is a more personal connection so we should establish a partnership with social media influencers to raise mental health awareness. A vast majority of Gen Xers are users of Facebook so promoting mental health posts would be a cost effective solution.

To measure success, we can create a control group based on the demographics we identify. By using our proposed solutions, we can see whether or not the test group sees an increase in prescriptions of Xanax or Zoloft compared to the control group.

4 Tools

- Pandas to create a random sample from each week

- Google Sheets for data cleaning, feature engineering, and preliminary analysis

- Tableau to create visualizations for analysis

5 Communication

Presentation slides, Google Sheets, and Tableau visualizations are located in my github repository.