**Social Media Activities: Mental Health**

**Question: Can mental illness be predicted by people's negative expressions/behaviors on social media?**

| **Problem**   * COVID's impact on social media use * COVID's impact on user’s mental status * Does topics of Reddit posts change over pre and post-COVID | **Solution**   * Time series for post trends and post engagements * Topic modeling for post text | **High level concept**   * ARIMA Time Series to compare topics, post trends and post engagements on mental health-related subreddits pre and post-COVID | **Assumptions**   * Post length from reddit dataset increases post-COVID * Upvote ratio and score increased post-COVID * Total comments increases in post-COVID * Topics of Post text of pre/post COVID will have differences | **Kill the assumptions**  How can we:   * use our design to prove the correlation between COVID and Reddit use * use our design to prove the correlation between Reddit use and mental health * perform topic modeling * prove the correlation between COVID-19 and mental health * reframe the problem if no correlation is found |
| --- | --- | --- | --- | --- |
| **Existing Alternatives**   * Existing studies on the effect of social media on users’ mental health | **Key metrics**   * Post trends * Post engagements * Topics * Time (Pre and Post-COVID)   (Can we use it for all 9 subreddit datasets?  Should we collect more data?) | **Unique value proposition**   * Using time series can mitigate history / maturation confound |

