

## **Supplemental material: GEPHI VISUALISATIONS:**

We followed the following steps to create these visualizations:

1. Imported the subreddit\_data\_info\_nw\_insights file into Gephi.
2. Applied Fruchterman Reingold layout and Yin Hu layout.
3. Changed the size of the nodes based on modularity class.

**There are 2 Gephi files provided as supplemental material.** They are:

### **1. SMA-Project-Vis1: This file contains a visualization between Subreddit and Total Average Subscribers.**

The following are our insights from this interesting visualization:

- The visualization suggests the presence of varying community structures within the network. If larger nodes represent subreddits with a higher modularity class, it indicates that these subreddits serve as hubs or central points within their respective communities.
- The visualization shows that while there are several connections within groups of nodes, the connections between these groups (inter-community links) are less frequent. This pattern supports the idea of distinct communities or topics within the network that interact more within themselves than with other communities.
- There are several nodes with fewer or no connections, isolated from the main groups. These could represent niche or less popular subreddits with fewer common links to the main or larger subreddits.

### **2.SMA-Project-Vis2: This file contains a visualization between Subreddit and Average Sentiment scores.**

The following are our insights from this interesting visualization:

- The graph shows a range of sentiment scores across different subreddits. The scores are clustered around certain values, which indicate predominant positive or negative sentiments in specific communities.
- By examining the names of subreddits and their sentiment scores, we could analyze how topics might influence sentiment. For instance, subreddits focused on animals being happy namely AnimalsBeingHappy, cute\_animals, and so on show higher positive sentiment compared to the other subreddits.
- Subreddits with extremely positive or negative scores could be significant. These outliers might represent niche communities with very specific types of engagement, whether highly supportive or contentious.

- Larger nodes might represent subreddits with higher engagement or more central roles in disseminating culture or sentiments within their clusters. These could be influential communities where opinions form and spread, potentially affecting neighboring nodes (subreddits) within the same modularity class.