PURDUE UNIVERSITY | The Data Mine

Social Media Marketing Analytics

Youngstown State University

Central Insurance Companies

Project Overview

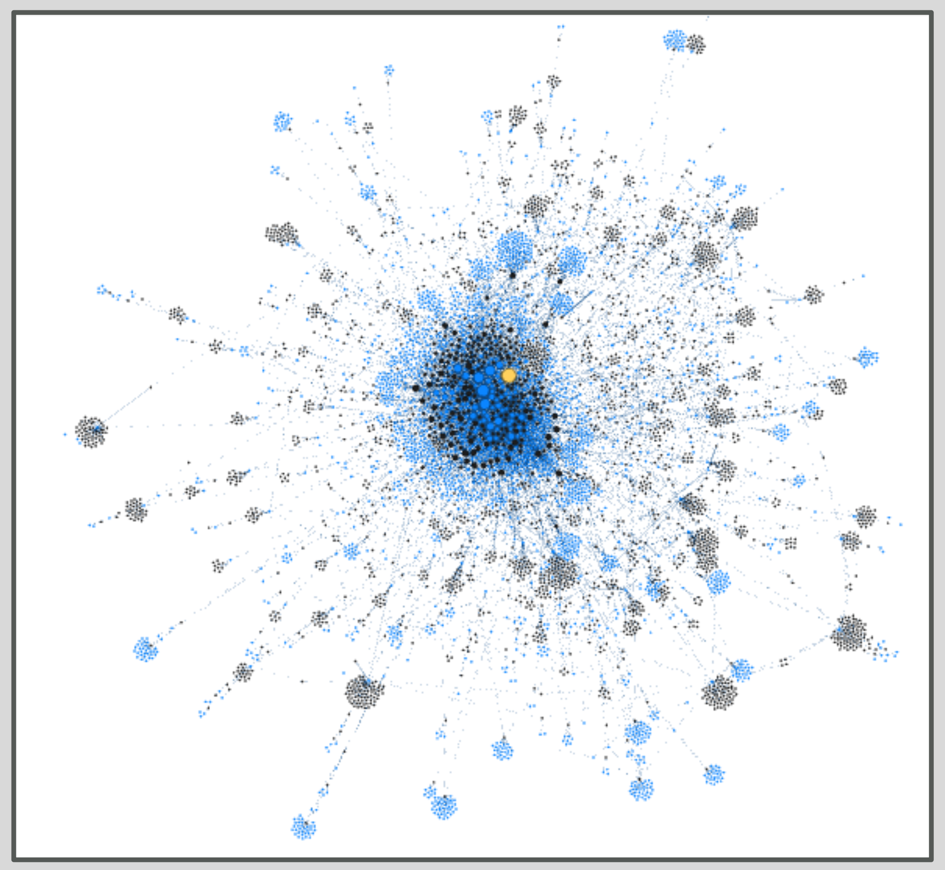
We collected social media data using web scraping to find insights in the social media campaigns of Central Insurance Company (CI). We analyzed comment and reaction rates of competitors on Facebook, Instagram, LinkedIn, and Twitter, and developed networks to evaluate Central's outreach.

Data Acquisition	
Web Scraping	Automated API Pipeline
Data Preprocessing	
Exploratory Data Analysis	Python Scripts
Data Analysis	
Competitor Analysis	Graph Network Analysis

Definitions

- **Web Scraping** – The process of extracting data from web pages.
- **Network Science –** Extension of graph theory that studies networks (graphs) that represent biological, physical, and social relations.
- **Eigenvector Centrality** – The ranking of each node of a network based on the eigenvector centrality scores of its neighbors.
- **Modularity** – Measures the strength of divisions in a graph.

LinkedIn



Legend
Yellow Node – Central Insurance
Black Nodes – Posts
Blue Nodes – Users

Graph Statistics

- 11,000 Edges
  - 7,400 Likes
  - 1,500 Comments
  - 400 Empathies
- 5,000 Nodes
  - 3,000 Posts
  - 1,500 Users

Eigenvector Centrality

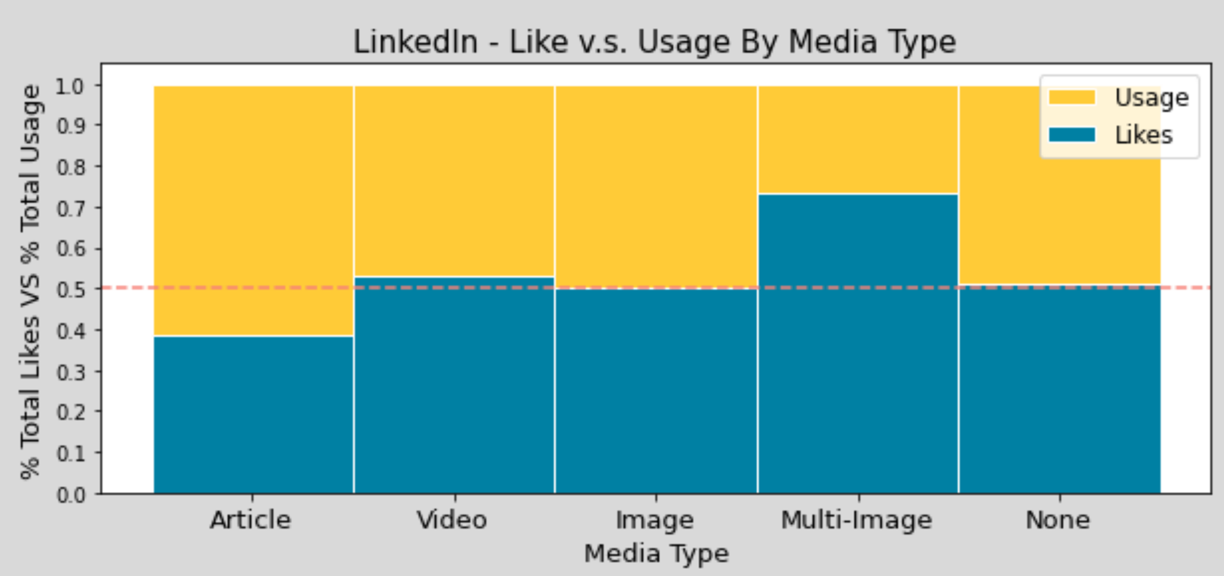
- Central Insurance has Centrality 1
- Next 11 Highest are Users
  - Centralities from 0.812 to 0.315

Modularity

- At Resolution 1, 111 Groups form.
  - Center split between groups 0 and 1.
    - Group 0 is largest; CI is in Group 1.
    - Most groups formed in branches on outside of graph model.
- At higher resolutions, center group unifies, and branches remain independent.

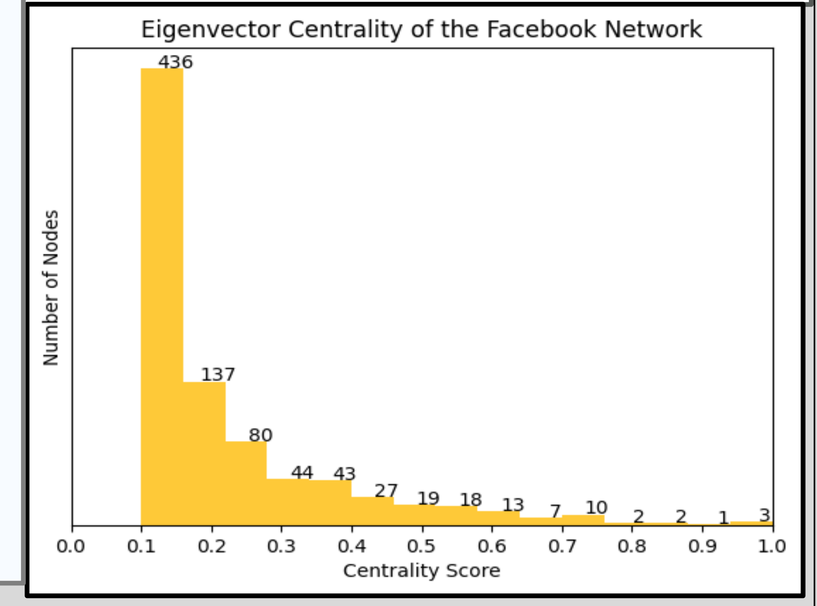
Note: Resolution determines group size.
- Higher Resolution = Fewer Groups
- Lower Resolution = More Groups

- Data collected from 5 companies' pages.
- Articles perform worst
  - Overused
- Multi-Image perform best
  - Underused



Facebook

Legend
Yellow – People of Influence in Ohio (VW)
Green – Retired Central
Blue – Outside Ohio
Magenta – Not CMI Employee



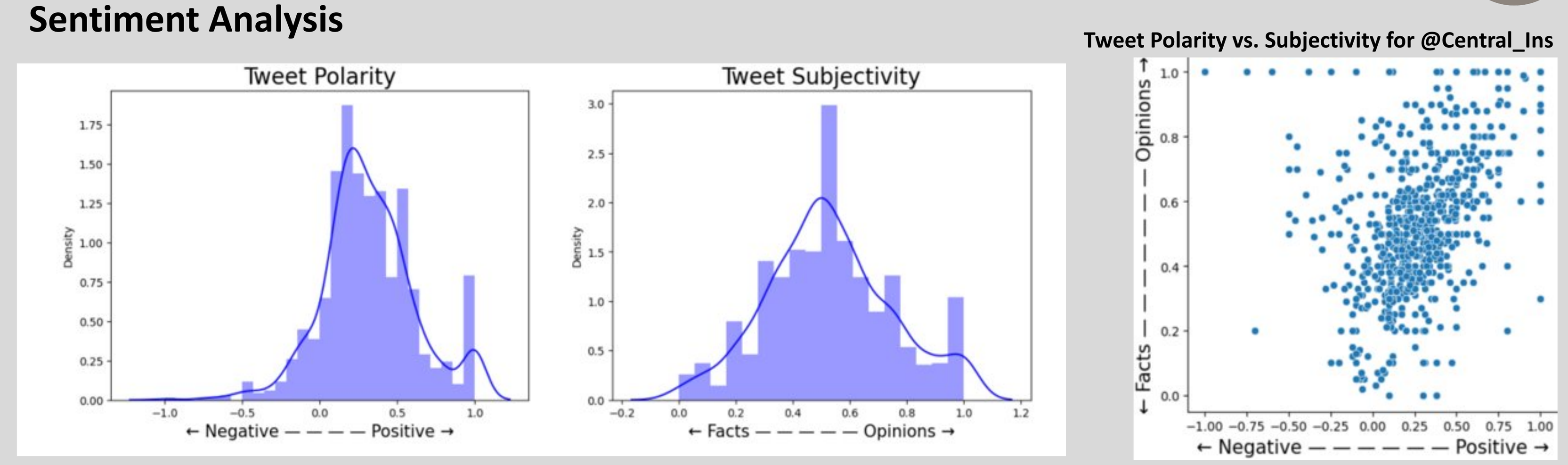
Page/Network Information

- 4.3K Followers
- 202 Post
- 1287 Unique Reactors
- 6170 Reactions
- 83K Nodes
- 116K Edges
- Colored by Community
- Sized by Degree Centrality
- 76k Communities
- Modularity 0.794

- An important reactor may not need lots of followers to impact a community.
- Central Mutual Insurance should encourage employees to react to post.

Twitter

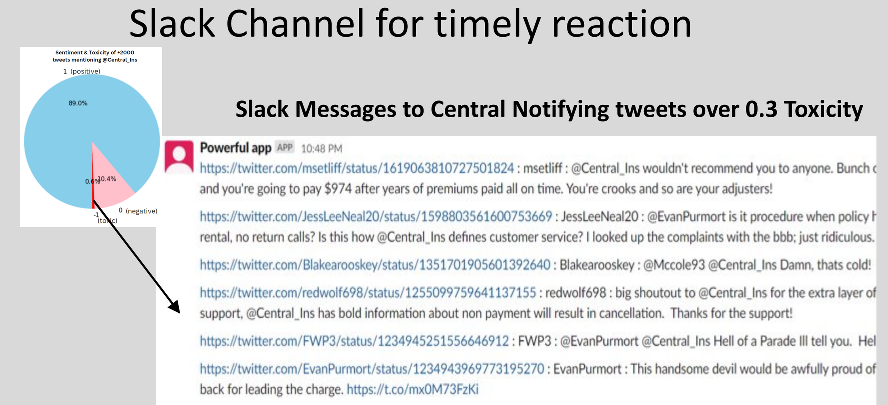
Sentiment Analysis



- Web scraped data of over 2000 tweets mentioning @Central_Ins over the past 3 years
- Polarity: Sentiment score of positive, neutral, or negative on a scale of 1 to -1
- Subjectivity: Sentiment score of opinion or fact on a scale of 1 to 0

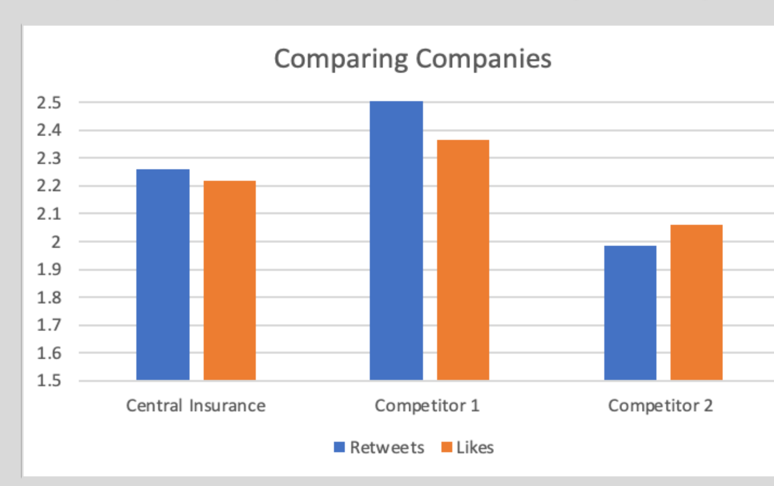
Toxicity Analysis

- Periodic toxicity scoring of tweets mentioning CI using Perspective API
- Tweets scoring over 0.3 are sent to CI's Slack Channel for timely reaction

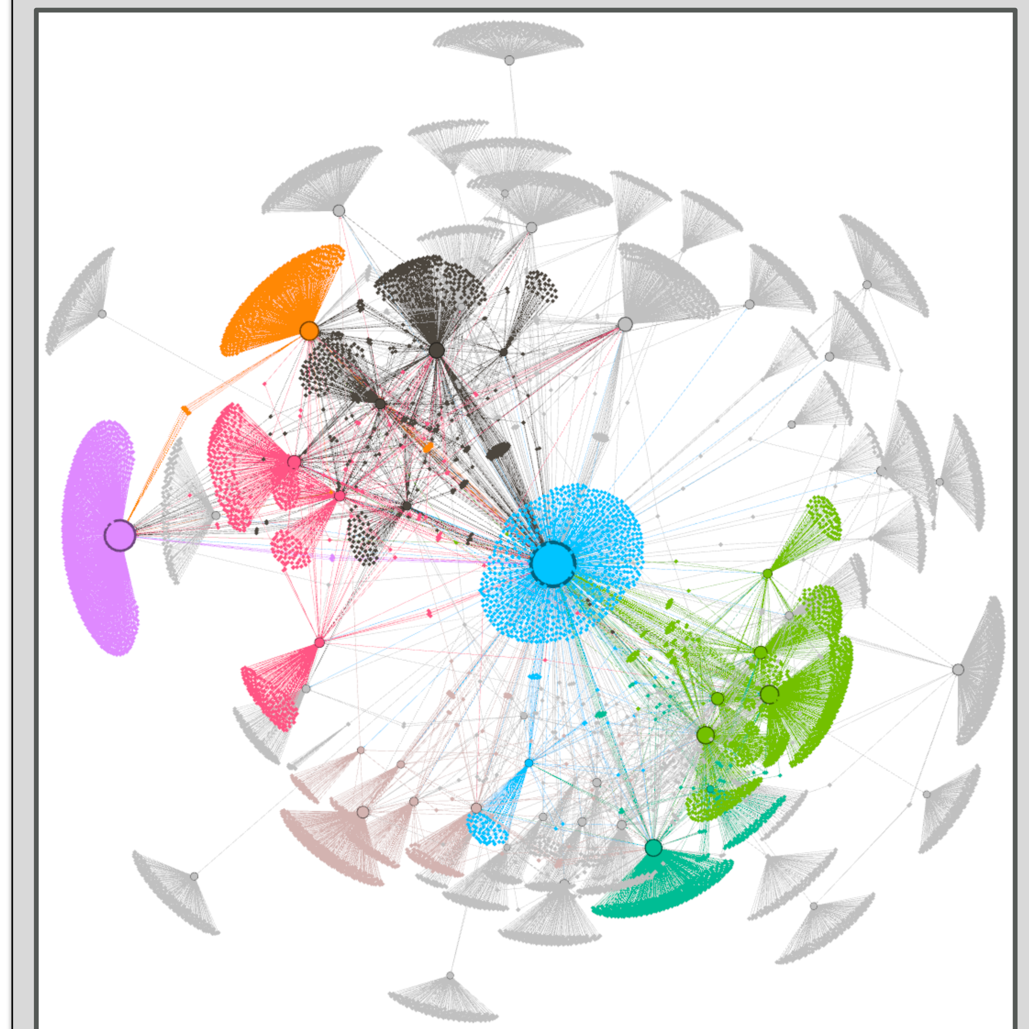


Competitor Analysis

- Determined comparative competitor analytics and presence boosting strategies



Instagram



Page/Network Information

- 1.4K Followers
- 9.8K Nodes
- 11.6K Edges
- Colored by Community
- Sized by Eigenvector Centrality

After scraping 62 accounts, we currently stand with 9.8K distinct users. From these, only 1.4K are directly following Central Insurance.

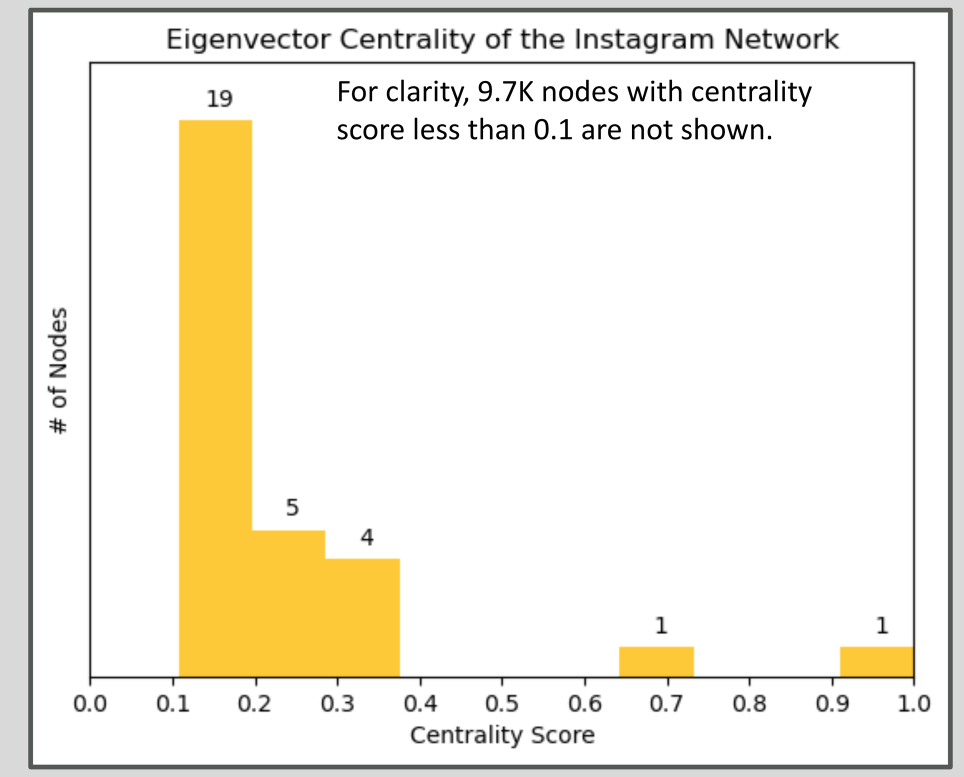
Legend
Blue – Central Insurance
Light Green – Career Opportunities
Grey – Individual Accounts
Other colors – Competitors

Community Detection

- 7 distinct communities
- Modularity of 0.842

Eigenvector Centrality

- All but 30 accounts < 0.1
- 10 influential accounts



Conclusions

Using various forms of web scraping, we created graph models that represent Central Insurance's presence on four social media platforms. Our insights included:

- **Community Detection** in user groups interacting with Central Insurance.
- **Influential Users** related to Central Insurance's social media activity.
- Full map of **Second-Degree Connections** for Central on Instagram, Facebook and LinkedIn.
- **Sentiment Scores** for user reviews on Twitter.
- **Media-Type Effectiveness** for posts on LinkedIn.

Our efforts have paved the way for further graph construction. Future research could define common factors in groups identified here and explore post effectiveness.

Acknowledgements

- Allie Coia, Dr. Jay Kerns, Dr. Yihao Deng
- John Aidoo, Nguyen Nguyen, and the Central Insurance Team
- David Glass and the Data Mine Staff
- Dr. Drake Olejniczak

Advising Team: Dr. G. Jay Kerns, Dr. Yihao Deng, Allie Coia. Students: Olamide Adekoya, Atharva Atre, Samuel Baah, Michael Bemus, Neal Birchfield, Jeric Cacal, Samantha Goodwill, James Lagese, Dariya Mamyrbek, Derek Miller, Shivi Pandey, Travis Schauer.

The Data Mine Corporate Partners Symposium 2023