The Factors of Increased MIsinformation Belief and Spread

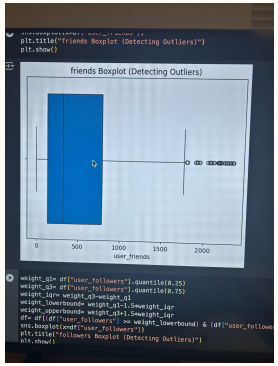
By Colin Thompson

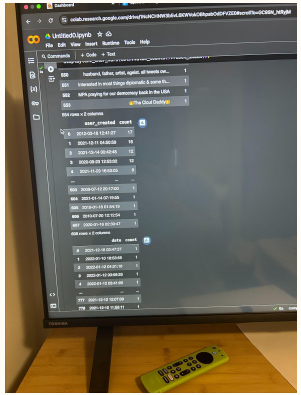
Project Topic and Motivation

The main topic for my project is misinformation, specifically investigated what factors lead to the a stronger belief or greater spread in misinformation. The main motivation behind me completing this project was due to the ever growing Presence of misinformation. Misinformation can be found for a multitude of topics and is present in most fields of study. Many also have trouble distinguishing between true infromationad false information. This drove me to be motivated enough to try and see why this was the case and what this could mean for how misinformation can be detected

Methods

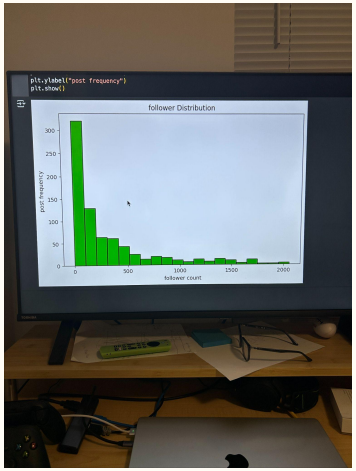
- The data used was from a Kaggle dataset contain information on accounts spreading conspiracy theories that birds do not exist
- For preprocessing the data was checked for NA values, outliers, and inconsistencies
- No NA Values or incosistiences were found in the data
- More extreme outliers were removed from the data using the IQR method
- The data was visualized in the project using a Histogram, regression plot, and a correlation heat map
- Basic descriptive statistics were collected for all numeric data
- The Mann Whitney U and point biserial correlation test were used to test data hypothesis
- A poisson regression was done to check the relationship between favorites and user verified status as a way of further investigating the relationship between verified factors and other data points

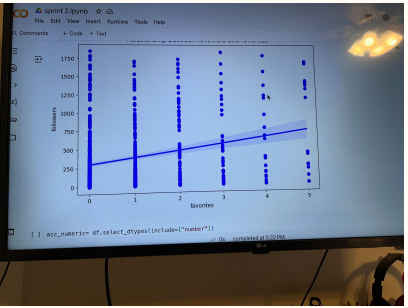


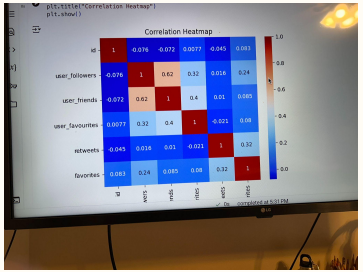


Results pt 1

- Descriptive statistics showed low mins and maxs for most numeric values
- Histogram showed distribution skewed to the left for follower distributions
- Regression plot showed a small positive correlation between follower count and favorites on a post
- Correlation Heatmap showed the previous correlation shown as well as a correlation between friends and followers and a slight correlation between retweets and favorites

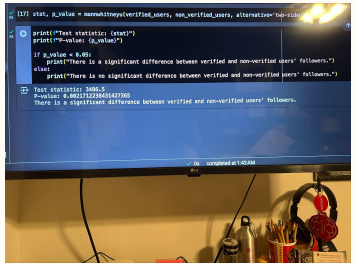


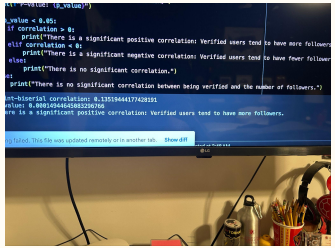


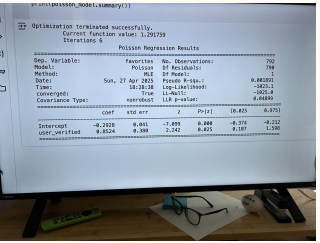


Results Pt 2

- Mann Whitney U test showed that there was a significant difference between the correlation of verified and unverified users
- Point Biserial Correlation Test showed that there was a significant positive correlation between follower counts and being a verified user
- The Poisson Regression analysis found no significant correlation between these two pieces of data, which means their is no relationship between these two factors that is of significant enough importance







Implications

- More smaller communities around accounts foster more belief and are more likely to spread misinformation
- Popularity does not seem to be important spreader for accounts
- People in these communities are somewhat closer knit with each other and may be somewhat close to the account poster
- Users may find verified accounts more popular due to perceived authenticity and sense of authority coming from the account

The End