Social Media Engagement Analysis for Purdue Athletics

The Team: Brandon Allen, Nate Garrison, Noam Katzenelson, Gabbi Kraus, Trent Marchand, Nick Molinaro, Luke Nitschke, Jonah Rockey, Jordan Sabel, Emma Sartore, Cyrus Swafford, Maggie Wade

Background and Motivation

• Social media is important to Purdue Athletics because it allows each sports team, and Purdue Sports as a whole, to grow and engage with their fan bases, create brands for themselves, and interact with other companies and sports teams via sponsorships and reposting.

• This purpose of this project is to collect social media data and conduct data analysis to accomplish the previously outlined relevant social media goals of Purdue Athletics.

• The data we worked with is from the @PurdueSports, @BoilerBall (basketball), and @BoilerFootball accounts, each on Twitter, Instagram, and Facebook (so 9 accounts total).

• Purdue Athletics was previously collecting and presenting data manually, so we help automate this process as much as possible and conduct analysis on the data using tools such as R, Python, Shiny, and Tableau that we found useful to implement.

• Our data analysis involves scraping the social media data (date/time, type of post, likes, comments) from the respective platforms, calculating growth metrics (engagement rate, follower rate), and summarizing the data in weekly and monthly reports through visualizations.
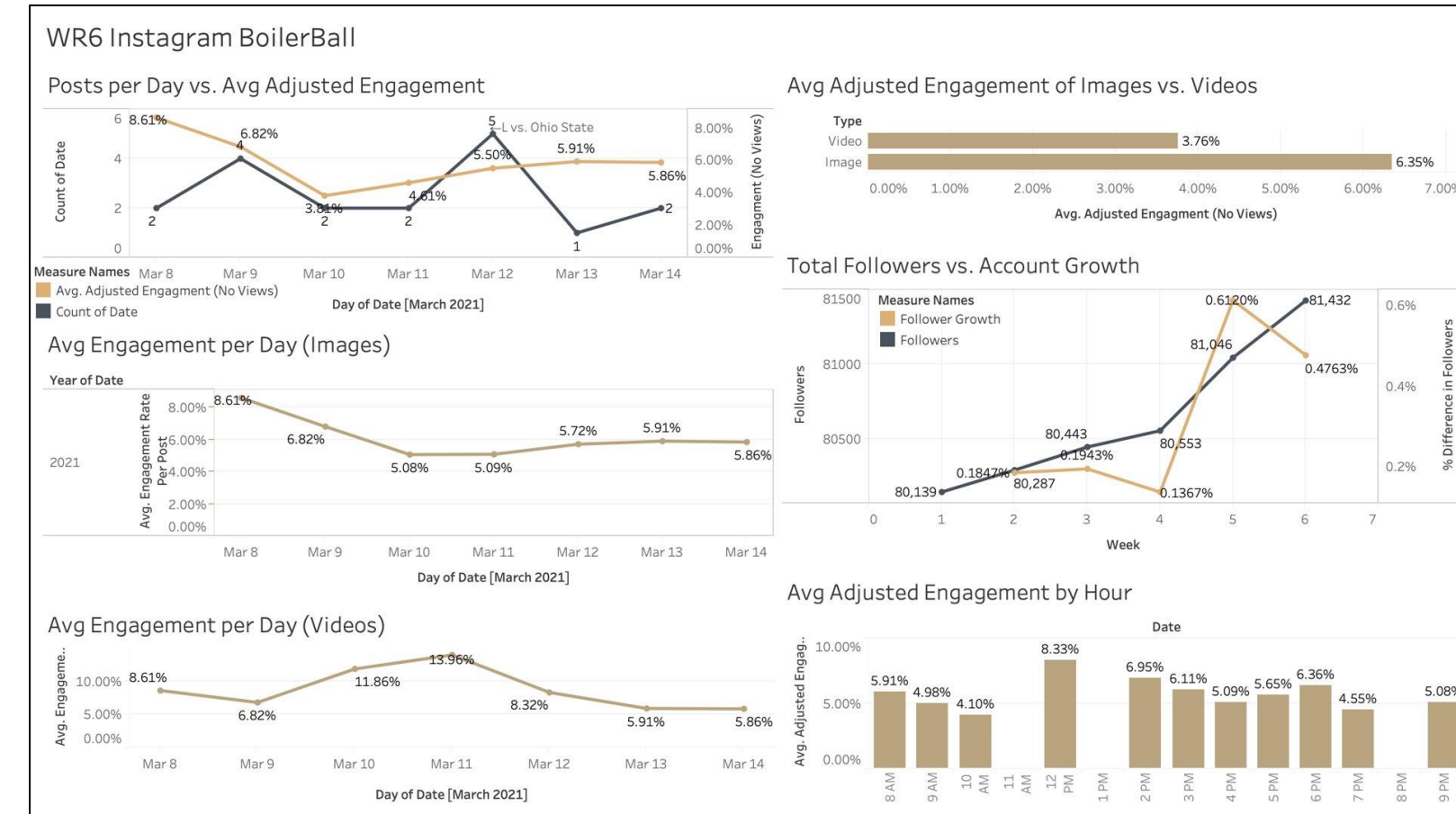
@BoilerBall Instagram account's Tableau dashboard data from our sixth weekly report (weekly reports discussed in the "Visualizing the Data" section).
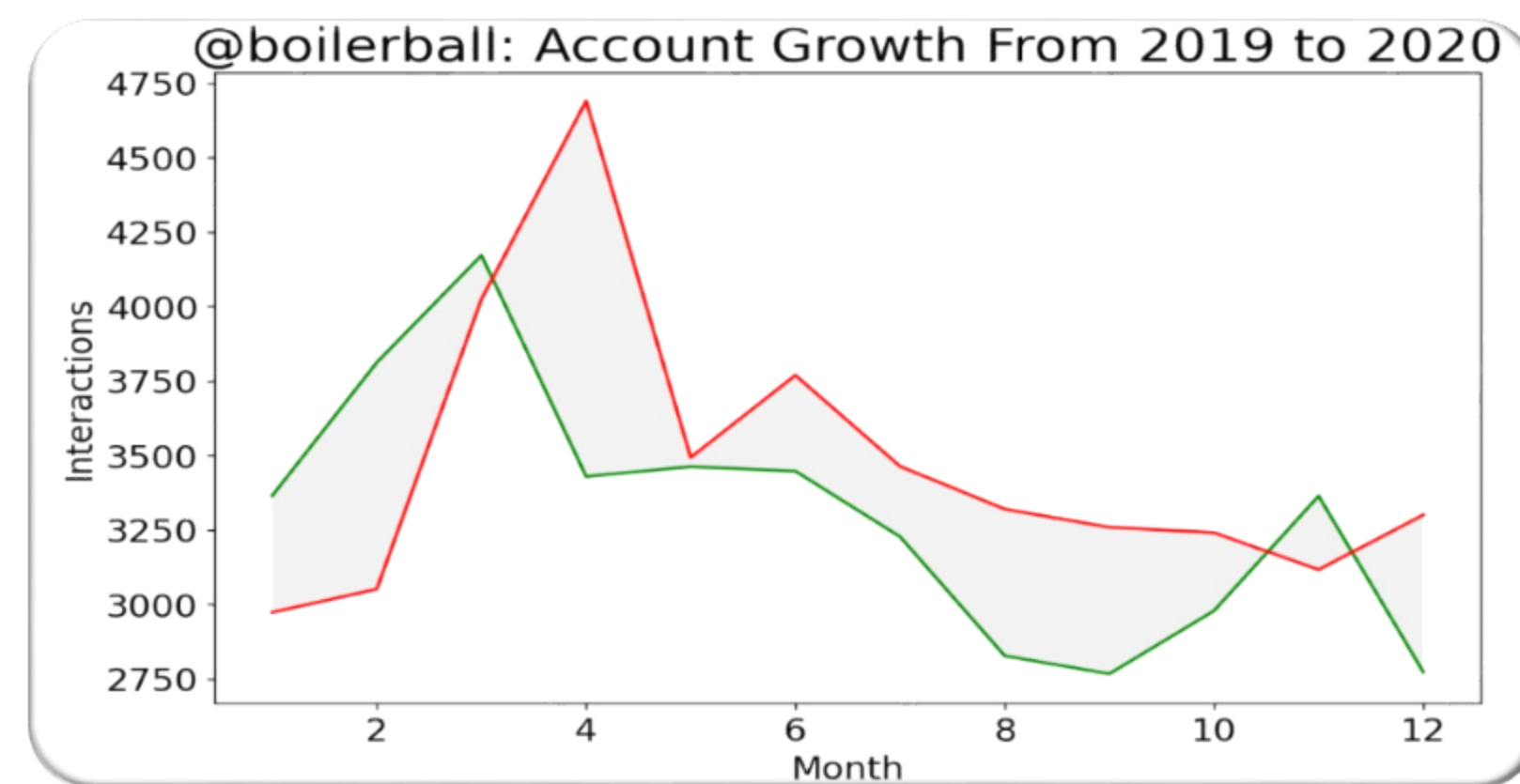
Interactions differential graph for the @BoilerBall Instagram account as a part of our "Creative Work" outlined in the "Visualizing the Data" section.

Shiny App

Shiny app of hourly interaction breakdown. Details in the "Shiny App" section.

Our Shiny app, which can be seen to the left, breaks down Twitter data into an hourly breakdown. The information can be filtered by the following ways:

• start and end date
• with or without outliers
• only show sponsored posts
• only get posts from a specific day of the week
• which account you want the data from

Methodology

Scraping Social Media

Twitter
• Created a Twitter Developer account and used an R package called rtweet
• With rtweet we can –
  • scrape data such as favorites, retweets, quotes, and replies
  • pull data from the previous 30 days or a specific time frame
  • separate the data based on whether the tweet contains a video or not

Instagram
• Used a Python package called Instaloader
• With Instaloader, we can –
  • scrape data such as type of post, likes, comments, video views and duration, location, etc.
  • scrape data from profiles, posts, and hashtags

Facebook
• Used a Python package called facebook-scraper
• With facebook-scraper, we can –
  • scrape data such as likes, comments, shares, and the various ways you can interact with Facebook posts (love, wow, sad, ha-ha, etc.)
  • use the pandas within Python to save the data to a dataframe and export to Microsoft Excel

Visualizing the Data

Tableau
• Once the data has been scraped and growth metrics have been calculated, we use Tableau to create interactive visuals.

• Excel files can easily be imported into Tableau Desktop and from there, we create 9 different visual dashboards with the Tableau software, one for each account on each platform, to present in report form.

• Tableau allowed us to visualize engagement on a weekly basis as well as break things down and look at many different variables.
  • For example, we can look at time of day posted and if the day was a game day for the @BoilerBall accounts during the season.
  • We were also able to visualize account growth and engagement for the social media team at Purdue Athletics.

Creative Freedom
• Visualize the data without strict guidelines or expectations to allow for individual creativity and application of information learned in classes.

• Many Python libraries (pandas, requests, etc.) that can be used to create visualizations
  • For example, we analyzed year by year differentials in account interactions and focused on gameday interactions.

Future Goals

General Goals
• Compare performance of sponsored vs non-sponsored posts
• Further automating the process of data collection and visualization production
• Continue pursuing creative freedom with visualizing the social media data

Facebook Goals
• Be able to automate the process of collecting data and presenting it, all in one process
• Create an interactive tool using the Dash library within Python
• Eventually, create a web server to easily access data whenever necessary
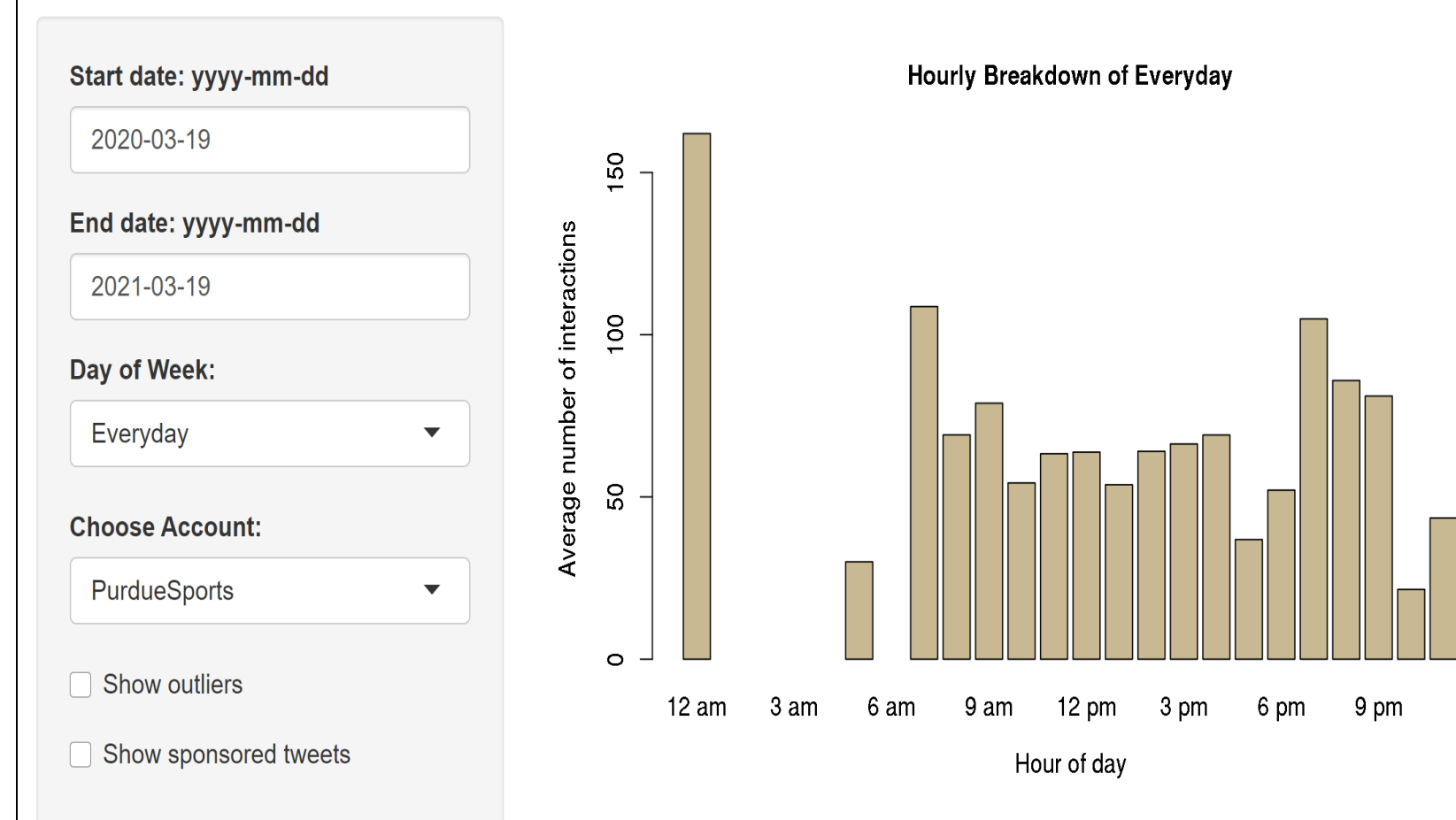
Conclusion

In conclusion, our project with Purdue Athletics is focused on helping them achieve their goals with social media: grow and engage with Purdue Sports fans, create brands, and establish sponsorship deals and relationships with other organizations. We are working to achieve these goals through scraping their social media data and conducting analysis by means of calculating growth metrics and creating insightful visuals using various software applications. While we have made decent progress with automating the process of data collection and presentation, we see furthering these efforts as a focus for this project in the future. Additional future goals include looking more at sponsored vs non-sponsored post performance, continuing to make new discoveries and learn new skills through creative visual freedom, and create an interactive Facebook tool similar to the Twitter Shiny app tool.

Acknowledgements