Background and Motivation

- Social media is important to Purdue Athletics because it allows each sports account to grow and engage with their fan bases, create brands for themselves, and interact with other companies and sports teams via sponsorships and reposting.
- The 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 accounts we scope are for Volleyball, Women's Basketball, Men's Basketball, Men's Golf, Wrestling, Soccer, Men's Swim/Dive, Women's Tennis, Baseball, Track and Field, Purdue Sports, Softball, Football, Men's Tennis, Women's Swim/Dive, and Women's Golf.
- The social media platforms we scrape for are Facebook, Instagram, and Twitter.
- Purdue Athletics was previously collecting data manually, and presenting data manually, so help automate this process as much as possible and conduct analysis on the data using tools such as R, Python, and Tableau that we found useful to implement.
- Our data analysis involves scraping the social media data (date/time, type of post, likes, and comments) from respective platforms, calculating growth metrics (engagement rate, follower growth) and summarizing the data in reports through visualization.

Scraping and Automation

- Rather than scraping the data by hand every week or two weeks, we decided to speed up the process and make it more efficient.
- We are using a CRON job to automatically scrape the data from all the accounts and input it into an Excel file.
- A CRON job is a job scheduler that runs a script, the scraper in this case, at a given time every day.

Categorizing the Data

- After meeting with Dr. David Pierce of Indiana University-Purdue University Indianapolis about his work in categorizing social media posts, the team added categories as a part of our analytics.
- Originally the categories were based on Dr. Pierce’s categories.
- Met with our team mentor and the Social Media Team of Purdue Athletics and altered the categories to best fit Purdue Athletics’ Social Media needs.
- The categories used are Offseason General, Pro Players/Alumni, Statistics, Purdue History, Mic’d Up, Fans & Community, Game Day, Team Series/Behind the Scenes, Player Features, Special Recognition & Initiatives, Team Features, Recruiting/Signees, Facilities, Fundraising Content, Sponsored Content, In-Season General, Game Week, Ticket Information/Sales, and Other.

Future Goals

- Continue to work on the Shiny App in R
- Digital engagement: using our tool with the Purdue Sports app and website
- Have more consistent scraping with Facebook and not getting our accounts banned
- Continue to automate data scraping