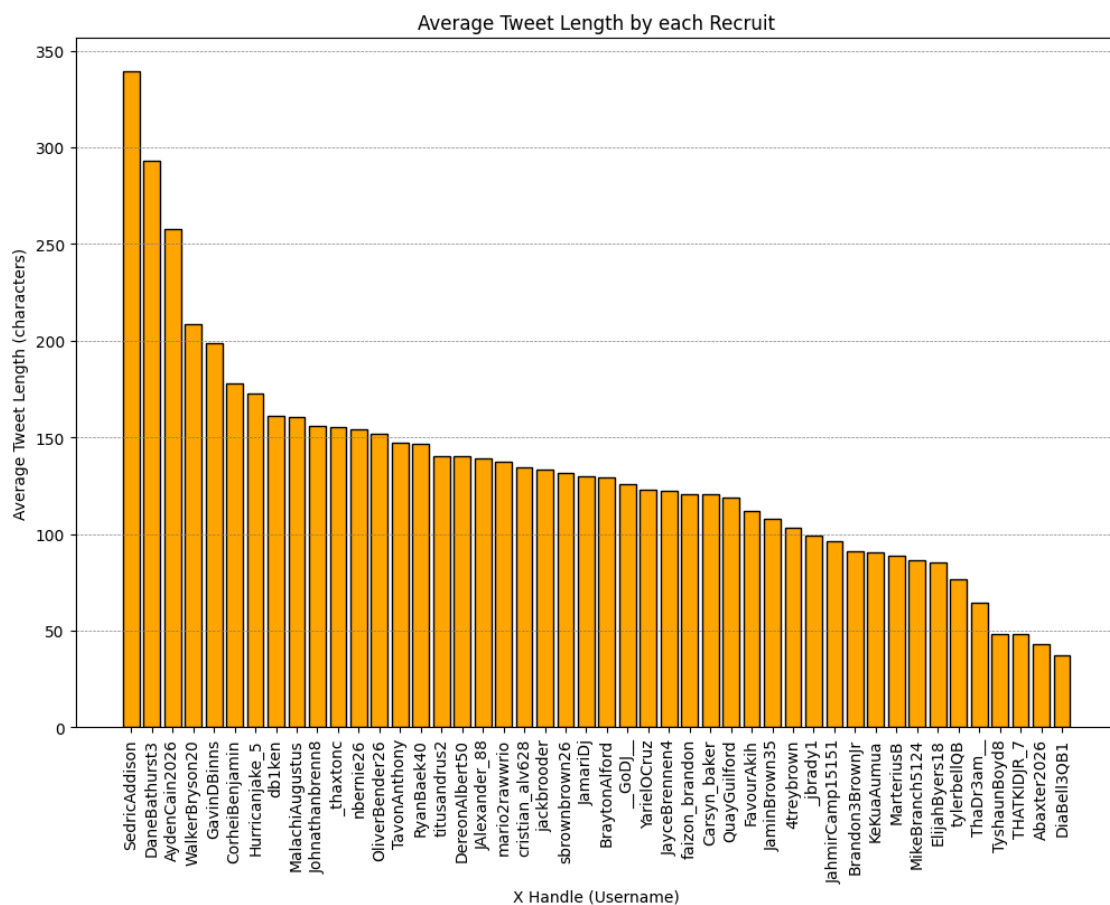
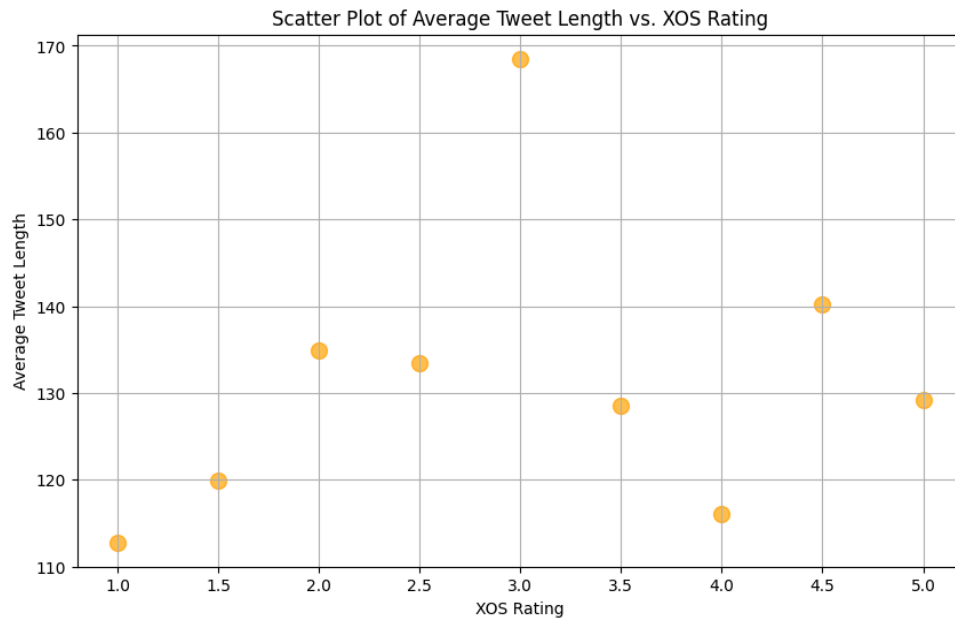


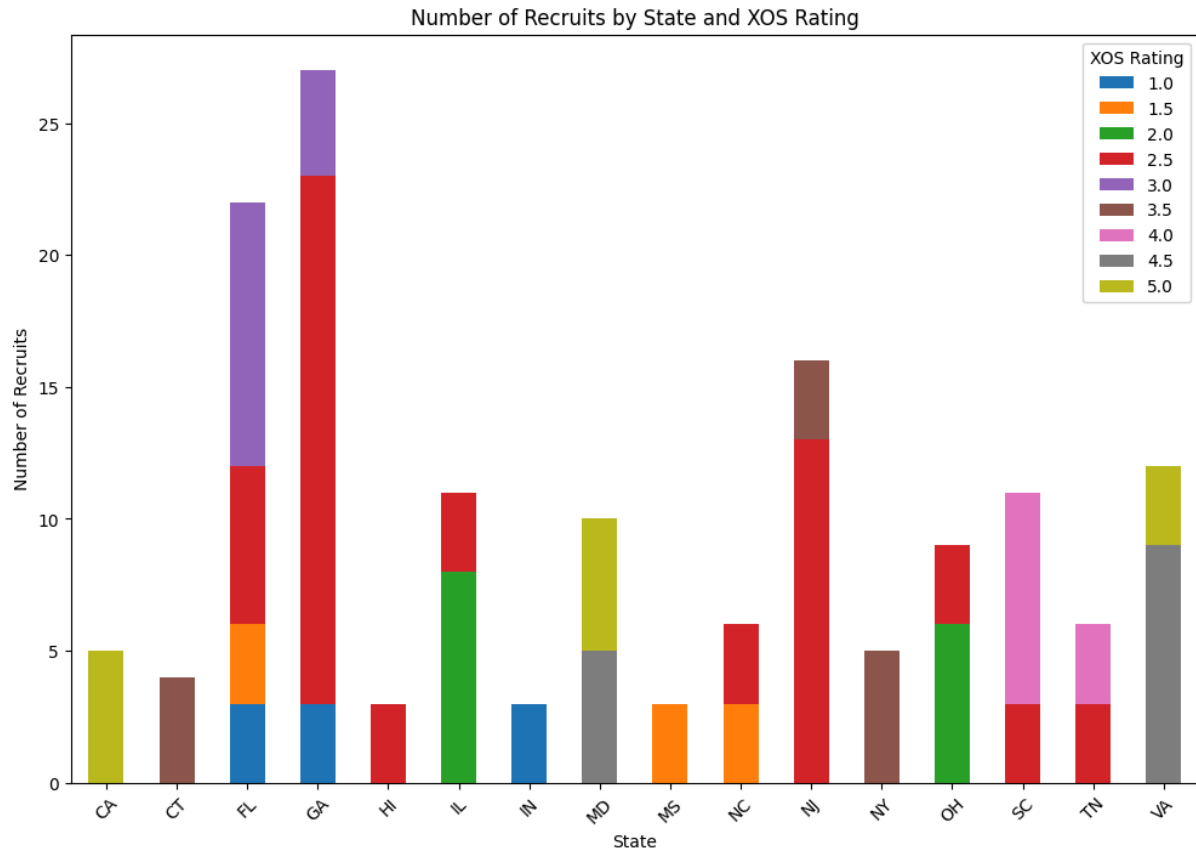
**Recruit\_Data:**

- Unit of observation is player-tweet (each row represents a player's unique tweet). Data set of tweets by 47 high school players considered for recruitment by UVA's football team.
- Tweet data pulled from X using the free X API plan and the Tweepy package, and then copy-pasted into rows to create processed data file.
- 227 rows and 5 columns
- 47 unique players

Column	Description	Example Data	Number of Observations
User	Categorical variable of the X username of the player	MarteriusB	168(0)
Full_Name	Categorical variable of the first and last name of the recruit	Boyd, Marterius	168(0)
XOS_Rating	Ordinal variable of the rating of the recruit that represents the football performance value of the recruit. Ranges from 1 to 5 in 0.5 increments, with 1 being the highest and 5 indicating the lowest skill level.	4.5	168(12)
High_School	Categorical variable of the high school attended by the player	Lausanne Collegiate School	168(0)
State	Categorical variable of player's home state	Virginia	168(0)
Tweet	Text data of tweet written by player pulled from X	Thankful to be recognized as a top prospect in my class!	168(0)

- Exploratory data analysis:





### Established\_names:

- Unit of observation: player-tweet (each row represents a unique tweet by player). Data set of tweets by 8 players established to be “gritty” by UVA football coaching staff collected using a survey.
- Description of survey and directions appearing to the participants:
  - I am completing a project where I am analyzing whether certain recruits exemplify grit. In order to find this, I am looking for players that you have coached or been associated within the past 10 years that exemplify grit. These players do not have to have played at UVA. Thank you for your answers and time, it is greatly appreciated. - Henry Duke
- Tweet data pulled from X using the free X API plan and the Tweepy package, and then copy-pasted into rows to create processed data file.
- Variables:
  - 80 rows and 3 columns
  - 8 unique players

Column	Description	Example	Number of Observations
--------	-------------	---------	------------------------

Full_Name	Categorical variable of the full name of the established player	Perry Jones	80(0)
User	Categorical variable of X username	@supe_jones	80(0)
Tweet	Text data of tweet written by player pulled from X	@supe_jones: Looking for Virginia's best! <a href="https://t.co/dLIaHxSxK9">https://t.co/dLIaHxSxK9</a>	80(0)

- Descriptive statistics:
  - Tweets per player (created using Excel pivot table)

Player	Count of Tweets
Ben Boulware	10
Deshaun Watson	14
Eli Wood	8
Jahmeer Carter	10
Jalen Pitre	
Perry Jones	8
Rodney McCloud	10
Wayne Gallman II	10
<b>Grand Total</b>	<b>80</b>

### Recruit\_VADER

- Unit of observation is player-tweet (each row represents a player's unique tweet). Data set of sentiment analysis using the VADER python package for each tweet from each possible recruit.
- Variables:

Column	Description	Example Data	Number of Observations
User	Categorical variable of the X username of the player	MarteriusB	168(0)
Full_Name	Categorical variable of the first and last name of the recruit	Boyd, Marterius	168(0)
XOS_Rating	Ordinal variable of the rating of the	4.5	168(12)

	recruit that represents the football performance value of the recruit. Ranges from 1 to 5 in 0.5 increments, with 1 being the highest and 5 indicating the lowest skill level.		
High_School	Categorical variable of the high school attended by the player	Lausanne Collegiate School	168(0)
State	Categorical variable of player's home state	Virginia	168(0)
Tweet	Text data of tweet written by player pulled from X	Thankful to be recognized as a top prospect in my class!	168(0)
sentiment_scores	List of sentiment scores of each tweet as calculated by the VADER python package.	{'neg': 0.0, 'neu': 0.5, 'pos': 0.5, 'compound': 0.7901}	168(0)
compound	Compound sentiment score calculated by VADER between -1 and 1, with -1 signifying a solely negative sentiment and 1 signifying a solely positive sentiment.	-0.79	168(0)
pos	Proportion of words within the tweet classified as "positive" by VADER. Between 0 and 1.	0.5	168(0)
neu	Proportion of words within the tweet classified as "neutral" by VADER. Between 0 and 1.	1	168(0)
neg	Proportion of words within the tweet classified as "negative" by VADER. Between 0 and 1.	0.619	168(0)

#### **Established\_VADER:**

- Unit of measurement is player-tweet.

- Database with VADER scores from established players achieved from sentiment analysis.
- Variables:

Column	Description	Example	Number of Observations
Full_Name	Categorical variable of the full name of the established player	Perry Jones	80(0)
User	Categorical variable of X username	@supe_jones	80(0)
Tweet	Text data of tweet written by player pulled from X	@supe_jones: Looking for Virginia's best! <a href="https://t.co/dLIaHxSxK9">https://t.co/dLIaHxSxK9</a>	80(0)
sentiment_scores	List of sentiment scores of each tweet as calculated by the VADER python package.	{'neg': 0.0, 'neu': 0.5, 'pos': 0.5, 'compound': 0.7901}	80(0)
compound	Compound sentiment score calculated by VADER between -1 and 1, with -1 signifying a solely negative sentiment and 1 signifying a solely positive sentiment.	-0.79	80(0)
pos	Proportion of words within the tweet classified as “positive” by VADER. Between 0 and 1.	0.5	80(0)
neu	Proportion of words within the tweet classified as “neutral” by VADER. Between 0 and 1.	1	80(0)
neg	Proportion of words within the tweet classified as “negative” by VADER. Between 0 and 1.	0.619	80(0)

- Exploratory data analysis:

