BU Athletics: Team-B

What have I worked on?

We have begun work on the extension project. We have been most focused on the extension project proposed where we are using the dataset with the polling values from each team. We have made a new dataset with all the polling features. These polling features are the averages for each team. We mapped these features to the athletes dataset. Although this is not individual data, this can give us a good way of identifying features that are affecting athletes GPA on a team-by-team basis.
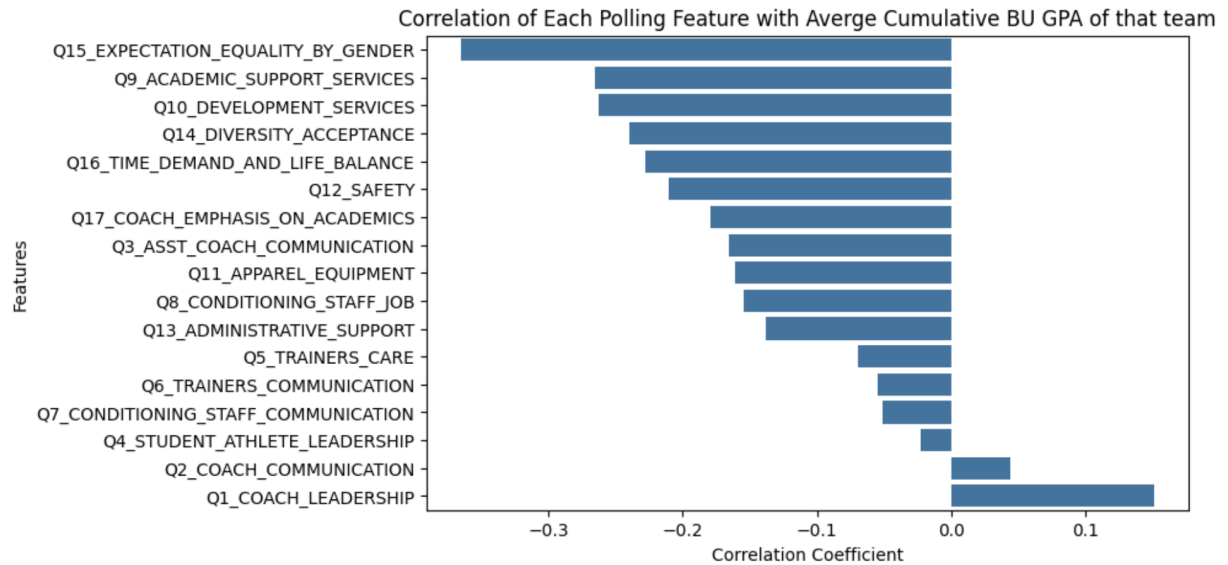
```python
for i, row in df.iterrows():
    if row['Data Input Year'] == '2021':
        year = 2021
        Q1_COACH_LEADERSHIP.append(df2021.iloc[0][row['Athletic Code']])
        Q2_COACH_COMMUNICATION.append(df2021.iloc[1][row['Athletic Code']])
        Q3_ASST_COACH_COMMUNICATION.append(df2021.iloc[2][row['Athletic Code']])
        Q4_STUDENT_ATHLETE_LEADERSHIP.append(df2021.iloc[3][row['Athletic Code']])
        Q5_TRAINERS_CARE.append(df2021.iloc[4][row['Athletic Code']])
        Q6_TRAINERS_COMMUNICATION.append(df2021.iloc[5][row['Athletic Code']])
        Q7_CONDITIONING_STAFF_COMMUNICATION.append(df2021.iloc[6][row['Athletic Code']])
        Q8_CONDITIONING_STAFF_JOB.append(df2021.iloc[7][row['Athletic Code']])
        Q9_ACADEMIC_SUPPORT_SERVICES.append(df2021.iloc[8][row['Athletic Code']])
        Q10_DEVELOPMENT_SERVICES.append(df2021.iloc[9][row['Athletic Code']])
        Q11_APPAREL_EQUIPMENT.append(df2021.iloc[10][row['Athletic Code']])
        Q12_SAFETY.append(df2021.iloc[11][row['Athletic Code']])
        Q13_ADMINISTRATIVE_SUPPORT.append(df2021.iloc[13][row['Athletic Code']])
        Q14_DIVERSITY_ACCEPTANCE.append(df2021.iloc[14][row['Athletic Code']])
        Q15_EXPECTATION_EQUALITY_BY_GENDER.append(df2021.iloc[15][row['Athletic Code']])
        Q16_TIME_DEMAND_AND_LIFE_BALANCE.append(df2021.iloc[16][row['Athletic Code']])
        Q17_COACH_EMPHASIS_ON_ACADEMICS.append(df2021.iloc[17][row
```

```
Effectiveness of head coach's leadership and d...
Effectiveness of head coach's communication sk...
Effectiveness of assistant coach's communicati...
Effectiveness of student-athlete leadership wi...
Confidence in the quality of care provided to ...
When injured, the extent to which Athletic Tra...
Appropriate communication, monitoring and indi...
Extent to which the strength and conditioning ...
Accessibility, knowledge, and quality of resou...
Accessibility, effectiveness, and knowledge of...
Quality and quantity of necessary apparel and ...
Condition of locker room and competition venue...
Adequate attention to safety while traveling f...
Travel accomodations were well organized and m...
Accessibility and effectiveness of athletics a...
Extent to which our team spoke about competiti...
Accepting, welcoming environment for diversity...
Similar performance expectations, behavioral s...
Ability to successfully manage time demands an...
Extent to which your coaches emphasized the im...
```

Do you want to install th

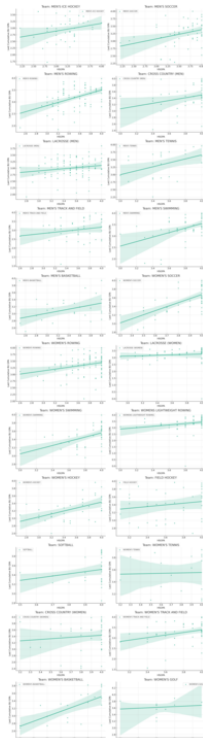| Q1_COACH_L	Q2_COACH_	Q3_ASST_CC	Q4_STUDEN	Q5_TRAINER	Q6_TRAINER	Q7_CONDITI	Q8_CONDITI	Q9_ACADEM	Q10_DEVELC	Q11_APPARI	Q12_SAFETY	Q13_ADMIN	Q14_DIVERS	Q15_EXPECT	Q16_TIME_C	Q17_COACH_EMPHASIS_ON_ACADEMICS
3.88	3.42	3.58	3.79	3.79	3.96	4	3.92	3.83	3.67	3.13	3.75	3.74	3.79	3.74	3.67	3.71
3.78	3.37	3.7	3.37	3.7	3.85	3.37	3.26	3.81	3.43	3.33	3.96	3.7	3.67	3.74	3.37	3.63
2.46	2.31	2.68	3.41	3.78	3.87	3.55	3.29	3	3.19	2.79	3.84	3.46	3.81	3.63	3.13	3.51
3.78	3.37	3.7	3.37	3.7	3.85	3.37	3.26	3.81	3.43	3.33	3.96	3.7	3.67	3.74	3.37	3.63
2.46	2.31	2.68	3.41	3.78	3.87	3.55	3.29	3	3.19	2.79	3.84	3.46	3.81	3.63	3.13	3.51
3.88	3.42	3.58	3.79	3.79	3.96	4	3.92	3.83	3.67	3.13	3.75	3.74	3.79	3.74	3.67	3.71
3.88	3.42	3.58	3.79	3.79	3.96	4	3.92	3.83	3.67	3.13	3.75	3.74	3.79	3.74	3.67	3.71
2.46	2.31	2.68	3.41	3.78	3.87	3.55	3.29	3	3.19	2.79	3.84	3.46	3.81	3.63	3.13	3.51
3.78	3.37	3.7	3.37	3.7	3.85	3.37	3.26	3.81	3.43	3.33	3.96	3.7	3.67	3.74	3.37	3.63
3.88	3.42	3.58	3.79	3.79	3.96	4	3.92	3.83	3.67	3.13	3.75	3.74	3.79	3.74	3.67	3.71
3.73	3.19	3.39	3.44	2.95	3	3.6	3.48	3.24	3.44	2.73	3.75	3.5	3.73	3.69	3.46	3.31
3.78	3.37	3.7	3.37	3.7	3.85	3.37	3.26	3.81	3.43	3.33	3.96	3.7	3.67	3.74	3.37	3.63
3.78	3.37	3.7	3.37	3.7	3.85	3.37	3.26	3.81	3.43	3.33	3.96	3.7	3.67	3.74	3.37	3.63
3.65	3.52	3.68	3.68	3.39	3.67	3.68	3.68	3.39	3.52	3.35	3.74	3.67	3.7	3.73	3.43	3.72
3.65	3.52	3.68	3.68	3.39	3.67	3.68	3.68	3.39	3.52	3.35	3.74	3.67	3.7	3.73	3.43	3.72
3.65	3.52	3.68	3.68	3.39	3.67	3.68	3.68	3.39	3.52	3.35	3.74	3.67	3.7	3.73	3.43	3.72
3.78	3.37	3.7	3.37	3.7	3.85	3.37	3.26	3.81	3.43	3.33	3.96	3.7	3.67	3.74	3.37	3.63
4	3.91	3.91	4	3.73	3.91	2	1.91	3.9	3.45	3.27	3.91	3.5	4	3.64	3.45	3.91
2.46	2.31	2.68	3.41	3.78	3.87	3.55	3.29	3	3.19	2.79	3.84	3.46	3.81	3.63	3.13	3.51
2.91	2.82	3.48	2.91	2.45	2.44	3.53	3.74	3.08	3.15	2.36	3.27	3.07	3.63	3.7	3.45	3.55
3.73	3.19	3.39	3.44	2.95	3	3.6	3.48	3.24	3.44	2.73	3.75	3.5	3.73	3.69	3.46	3.31
3.78	3.37	3.7	3.37	3.7	3.85	3.37	3.26	3.81	3.43	3.33	3.96	3.7	3.67	3.74	3.37	3.63
3.78	3.37	3.7	3.37	3.7	3.85	3.37	3.26	3.81	3.43	3.33	3.96	3.7	3.67	3.74	3.37	3.63
4	3.91	3.91	4	3.73	3.91	2	1.91	3.9	3.45	3.27	3.91	3.5	4	3.64	3.45	3.91
3.88	3.42	3.58	3.79	3.79	3.96	4	3.92	3.83	3.67	3.13	3.75	3.74	3.79	3.74	3.67	3.71
3.78	3.37	3.7	3.37	3.7	3.85	3.37	3.26	3.81	3.43	3.33	3.96	3.7	3.67	3.74	3.37	3.63
2.46	2.31	2.68	3.41	3.78	3.87	3.55	3.29	3	3.19	2.79	3.84	3.46	3.81	3.63	3.13	3.51
3.88	3.42	3.58	3.79	3.79	3.96	4	3.92	3.83	3.67	3.13	3.75	3.74	3.79	3.74	3.67	3.71
3.65	3.52	3.68	3.68	3.39	3.67	3.68	3.68	3.39	3.52	3.35	3.74	3.67	3.7	3.73	3.43	3.72
3.7	3.48	3.83	3.78	3.71	3.95	3.82	3.73	3.39	3.47	2.5	3.73	3.45	3.86	3.55	3.5	3.64
2.46	2.31	2.68	3.41	3.78	3.87	3.55	3.29	3	3.19	2.79	3.84	3.46	3.81	3.63	3.13	3.51
3.7	3.48	3.83	3.78	3.71	3.95	3.82	3.73	3.39	3.47	2.5	3.73	3.45	3.86	3.55	3.5	3.64
3.7	3.48	3.83	3.78	3.71	3.95	3.82	3.73	3.39	3.47	2.5	3.73	3.45	3.86	3.55	3.5	3.64
2.73	2.87	3.33	2.71	3.4	3.73	3.93	4	3.38	3.55	2.13	3.93	3.21	3.47	3.73	3	3.8

Based on the data input year and the 'Athlete Code' in both the polling and athlete dataset we were able to map these values.

Grouped by Teams and took Cumulative GPA's mean for each team. Then retrieved the correlation between these mean GPAs and each of the polling scores, to understand the polling features' importance.

Correlation of Each Polling Feature with Averge Cumulative BU GPA of that team

Features such as Gender Equality, Academic Support, and Development Services are observed to be inversely correlated with the average BU GPA of teams.

Features such as Coaches Leadership and Communication are observed to be correlated with the average BU GPA of teams.

We have also done a regression analysis for HSGPA and BU GPA between different sports teams.

What will I be working on next?

Planning to perform more statistical analysis with polling data collected. We may start with Regression Analysis.

We are looking to test out different features and develop a regression model for predicting GPA based on certain features.

We are hoping to test out a number of different features to give the athletic department the sentiments behind their highest and lowest performing athletes from an academic standpoint.

Have I run into any issues? Do I need help?

We ran into issues trying to map the data between the polling dataset and the athletics dataset. There are abbreviations below that are pretty ambiguous for the sports teams. We did not get a direct answer from Spark or Athletics on what the abbreviations were, so we went to the athletic department and just spoke to enough people till we were able to get an answer. We requested individual performance data from athletics and were not able to receive that information for a different extension project.

Team Lead: Have I talked to the client recently? When are we meeting with them next?

We spoke to Matthew, our TPM, at our lab and over Slack over the past few weeks. He has been helpful in facilitating our communication with Spark for additional information requests. We are just working on our main extension now so will not have additional questions since we are scoping this extension on our own.