|  |
| --- |
| **1. 주제**  야구 선수 구종별 타율과 타구 분포의 시각화  **분반, 팀, 학번, 이름**  가반, 2팀, 20231782, 정민주 |

|  |  |
| --- | --- |
| **2. 요약**  목표는 주제와 유사하게 야구 선수의 구종별 타율과 타격 분포의 시각화입니다. 데이터를 받아 와서 리스트에 삽입하여 타율을 내고, 주로 안타를 치는 위치를 파악하여 그것을 시각적으로 한눈에 쉽게 볼 수 있게 하는 것이 핵심적인 내용입니다. 좋아하는 선수나 좋아하는 구단 내의 선수들이 어느 곳으로 안타를 많이 치는지, 어느 구종에 안타를 많이 치는지 알 수 있으며 좀 더 자세하게 야구에 대해 분석하며 볼 수 있다는 기대효과가 있습니다. | **3. 대표 그림**  **휴대 전화, 모바일 기기, 휴대용 통신 장치, 통신 장치이(가) 표시된 사진  자동 생성된 설명** |

|  |
| --- |
| **4. 서론**  야구를 보다 보면 좋아하는 선수, 좋아하는 구단 내의 선수들이 어느 곳으로 안타를 많이 치는지, 어느 구종일 때 안타를 많이 치는지 궁금하고 알고 싶어지게 됩니다. 방송사를 통해 중계를 볼 때 가끔 타구 분포도를 볼 수 있지 팬들이 접할 수 있는 사이트나 어플에서는 분포도를 시각화해 나타내 주지 않습니다.  KBO 사이트나 ‘KBO STATS’ 어플을 통해 타율과 안타 개수, 홈런 개수 등은 알 수 있지만 타구 분포나 구종별 타율은 알 수 없었습니다.  직접 데이터를 가져 와, 구종별 타율을 계산하고 타구 분포를 시각화할 수 있는 프로그램을 만들면 좀 더 자세하게 선수의 기록을 분석하며 볼 수 있게 될 것입니다. |

|  |
| --- |
| **5. 본론**  텍스트, 휴대 전화이(가) 표시된 사진  자동 생성된 설명  매 경기 이후에 KBO 사이트나 ‘KBO STATS’ 어플의 API 또는 데이터를 가져 와야 합니다. 안타를 쳤을 때 어느 방향(어느 수비수 앞이나 뒤로 타구가 떨어졌는지)으로 갔는지, 어느 구종(직구, 포크볼, 커브, 커터, 슬라이더 등)으로 안타를 만들었는지의 데이터를 가져 와야 합니다.  이제 이 데이터를 가지고 타자가 어느 방향으로 안타를 많이 쳤는지 시각화할 수 있어야 합니다. 아래는 예시 사진입니다.  텍스트, 스크린샷, 잔디, 야외이(가) 표시된 사진  자동 생성된 설명  받아 온 데이터에서 어느 수비수쪽인지, 타구가 앞쪽인지 뒷쪽인지를 특정 단어들로 판단합니다. 타율에 따라 색의 차이를 두어 한눈에 알아 볼 수 있게 시각화합니다.  선수 리스트를 미리 만들어 놓고 데이터를 받아 와서 안타를 친 구종을 리스트에 삽입합니다. 처음 데이터를 받아 올 땐 엑셀로 받아 와, 코랩의 오픈소스 pandas와 matplotlib를 통해서 미리 데이터를 정리해 놓습니다. 중복되는 변수도 각각 저장되게 합니다. 모든 구종에 대해 리스트 내에 그 구종의 개수를 리스트 내에 있는 변수 개수로 나눠주면 각 구종에 대한 타율이 계산됩니다. 모든 경기를 다 반영하진 않고 최근 열 경기만 반영할 것입니다. 선입선출 방식으로 데이터가 사라집니다.  타구 위치는 ‘좌익수’ ‘중견수’, ‘앞’ ‘옆’ 등의 단어들로 판단합니다. 미리 ‘우익수 앞’ ‘중견수 앞’ 등의 리스트를 만들어 놓고, 리스트에는 구종을 삽입합니다. 이 리스트에서는 구종이 유의미한 결과를 도출하는 데에는 관련이 없기 때문에 따로 변수를 만들지 않고 구종을 사용하여 단지 코스 별 안타 개수를 기록하면 됩니다. 각 코스의 변수 개수(구종 개수)를 전체 변수 개수(모든 리스트에 있는 구종 개수)로 나누면 각 코스로 타구가 가는 비율을 알 수 있게 됩니다.  추가로 좋아하는 팀을 설정할 수 있게 하여 하단에 구단 로고를 띄울 것입니다. |

|  |
| --- |
| **6. 결론**  구현하고자 하는 것은 야구 선수들의 구종별 타율과 타구 분포도를 시각화하는 것입니다. KBO 사이트 또는 ‘KBO STATS’ 어플을 통해 매 경기 이후 데이터를 가져 올 것이고 리스트에 넣어 구종별 타율을 계산하고 타구 위치별 타율을 계산할 것입니다.  데이터를 받아 오는 코드나 사이트 등을 검색해 봐야 하고 전반적인 코드 작성을 시작해야 합니다. 무엇보다 데이터를 가져 와서 조건에 따라 리스트에 삽입하는 것이 가장 중요하기 때문에 그 코드를 가장 공을 들여서 짜야 합니다. 데이터를 엑셀으로 정리할 것이기 때문에 코랩을 사용하여 데이터를 정리하는 것도 고려할 것입니다. |

**7. 출처**

[1] SPOTVNEWS, 2018, <https://sports.naver.com/news?oid=477&aid=0000140216>, 최정 타구 코스 별 분석.