

Bài tập Assignment

Nguyễn Duy Hoàng – B22DCCN334 – Nhóm 11

(Chú thích và số liệu có thể chỉnh sửa được ghi trong file ‘.py’)

1: *Viết chương trình Python:*

- File: ‘btl1.py’; ‘results.csv’
- Chú ý: Do url của trang web <https://fbref.com/en/> đôi lúc sẽ gặp lỗi quá tải khi tra nhiều url liên tục (Rate Limited Request (429 error)), dẫn đến không lấy được data đầy đủ (KeyError: ‘Player’)
- Dữ liệu khi được tra đầy đủ ở trong file ‘results.csv’ có sẵn.

2:

a, Tìm top 3 cầu thủ có điểm cao nhất và thấp nhất ở mỗi chỉ số.

- File: ‘btl2.1.py’

b, Tìm trung vị của mỗi chỉ số. Tìm trung bình và độ lệch chuẩn của mỗi chỉ số cho các cầu thủ trong toàn giải và của mỗi đội. Ghi kết quả ra file results2.csv.

- File: ‘btl2.2.py’; ‘results2.csv’

c, Vẽ histogram phân bố của mỗi chỉ số của các cầu thủ trong toàn giải và mỗi đội.

- File: ‘btl2.3.py’

d, Tìm đội bóng có chỉ số điểm số cao nhất ở mỗi chỉ số. Theo bạn đội nào có phong độ tốt nhất giải ngoại Hạng Anh mùa 2023-2024.

- File: ‘btl2.4.py’

3:

a, Sử dụng thuật toán K-means để phân loại các cầu thủ thành các nhóm có chỉ số giống nhau.

- File: ‘btl3.py’

c, Sử dụng thuật toán PCA, giảm số chiều dữ liệu xuống 2 chiều, vẽ hình phân cụm các điểm dữ liệu trên mặt 2D.

- File: 'btl3.1.py'

d, Viết chương trình python vẽ biểu đồ rada (radar chart) so sánh cầu thủ:

- File: 'radarChartPlot.py'

- Attribute: không nên chọn cột có chỉ số chênh lệch quá lớn, dẫn đến radar chart không thể hiện rõ chỉ số cần xét.

4:

a, Thu thập giá chuyển nhượng của các cầu thủ trong mùa 2023-2024 từ trang web <https://www.footballtransfers.com>.

- File: 'btl4.py'; 'results3.csv'