

به نام خدا

پوریا کلاتتری ۹۷۳۱۱۲۷

گزارش قسمت عملی تمرین ۱ درس داده کاوی

برای این قسمت، من از لایبرری pandas استفاده کردم و در خط ۳ مشخص است که به چه صورت ورودی را دریافت میکنم:

```
1 import pandas as pd
2
3 data = pd.read_csv("players.csv")
4
```

سپس برای قسمت a با استفاده از تابع iloc که درواقع متدی برای متغیر data است، عنصر اول و آخر لیست بازیکنان را نمایش میدهم:

```
print("\n\n\n\n-----PART A-----\n\n")
print(f'first and last index of file is:\n{data.iloc[[0, -1]]}')
```

!!! خروجی ها را در انتهای فایل گزارش قرار داده ام!!!!

قسمت b را نتوانستم به دست بیاورم ولی برای قسمت c مقدار مینیمم، ماکسیمم و یا حتی مقدار میانگین را با استفاده از توابع (به ترتیب) min و max و mean به دست می آورم که قبل از استفاده از این توابع باید مشخص کنیم که مقدار کدام ستون مدنظر است که مطابق تصویر زیر مشخص است که چگونه این کار را انجام دادم:

```
print("\n\n\n\n-----PART C-----\n\n")
print(f'minimum weight is {data["Weight"].min()} and maximum is {data["Weight"].max()}')
print(f'average of weight is {data["Weight"].mean()}')
```

برای قسمت d نیز ابتدا همه ملیت ها را group میکنم سپس تعدادشان را با استفاده از size به دست می آورم سپس به کمک تابع reset index در یک ستون جدید به نام players\_count قرار میدم. در خط بعدی نیز مشخص میکنم از دیتاست جدیدمان (a) آن بازیکنایی رو نشون بده که مقدار players\_count شون مساوی min و max همون ستون باشه.

```
print("\n\n\n\n-----PART D-----\n\n")
a = data.groupby(['Nationality']).size().reset_index(name='players_count')
print(f'minimum country players is:\n {a[a["players_count"]==a["players_count"].min()]} \nand maximum is:\n\n
```

برای قسمت e و f نیز از دستون query استفاده کردم تا کسانی که مقادیر رشد و پتانسیلشان به ترتیب کمتر از ۳ و ۸۴ باشد، با این تفاوت که برای قسمت f گفتم تنها ستون های زیر را نشان دهد:

"Name", "Growth", "Potential", "Positions"

کد:

```
print("\n\n\n\n-----PART E-----\n\n")
print(f'players with growth < 3 and potential < 84 are:\n{data.query("Growth < 3 and Potential < 84")}')

print("\n\n\n\n-----PART F-----\n\n")
print(f'past report with position:\n{data.query("Growth < 3 and Potential < 84")}[["Name", "Growth", "Potentia
```

قسمت H: برای این قسمت نیز دقیقاً مانند قسمت قبلی ابتدا یک query نوشتم تا کسانی که قراردادشان تا ۲۰۲۱ تمام میشود و در تیم ملی نیستند (مقدار NationalTeamشان برابر با Not in team باشد) را جدا کند و در متغیر a بریزد سپس در دو خط بعدی هم تعدادشان و هم خودشان را نمایش میدهم:

```
print("\n\n\n\n-----PART H-----\n\n")
a = data.query("ContractUntil < 2021 and NationalTeam == 'Not in team'")
print(f'number of players contract until 2021 and not in national team: {len(a)}')
print(f'those players are:\n{a}')
```

برای قسمت I نیز مانند قسمت های قبلی یک query نوشتم تا نیاز های این قسمت را ارضا کند سپس تنها ستون های اسم و سن و ValueEURشان را نمایش میدهم:

```
print("\n\n\n\n-----PART I-----\n\n")
a = data.query("Age < 24 and Potential < 84 and Club == 'Chelsea')[["Name", "Age", "ValueEUR"]]
print(f'chelsea under 24 years old\'s ARZESH:\n{a}')
```

و در انتها برای موارد مورد نیاز هازارد ابتدا سطرهای که فیلد Nameش برابر با e. Hazard بود جدا کردم سپس ستون های مورد نظر را جدا کردم:

```
print("\n\n\n\n-----PART J-----\n\n")
a = data.query("Name == 'E. Hazard')[["Name", "Club", "Positions", "WageEUR"]]
print(f'hazard\'s info is:\n{a}')
```

در انتها تمام خروجی های بدست آمده را screen shot گرفتم و نمایش میدهم:

خروجی قسمت های a و c:

```
-----PART A-----

first and last index of file is:
   ID      Name      FullName  Age  Height  Weight  ...  CDMRating  RWBRating  LBRating  CBRating  RBRating  GKRating
0   158023  L. Messi    Lionel Messi  33    170     72    ...      68         69         65         55         65         22
19019 241493  S. Cartwright Samuel Cartwright  19    185     75    ...      45         46         47         51         47         15

[2 rows x 90 columns]

-----PART C-----

minimum weight is 50 and maximum is 110
average of weight is 75.05241850683491
```

خروجی قسمت d:

-----PART D-----

minumum country players is:

	Nationality	players	count
3	Andorra		1
8	Aruba		1
12	Barbados		1
17	Bermuda		1
28	Chad		1
62	Guam		1
63	Guatemala		1
69	Hong Kong		1
73	Indonesia		1
84	Korea DPR		1
95	Malawi		1
96	Malaysia		1
98	Malta		1
108	New Caledonia		1
117	Papua New Guinea		1
123	Puerto Rico		1
127	Rwanda		1
129	Saint Lucia		1
135	Singapore		1
139	South Sudan		1
142	Suriname		1
146	São Tomé & Príncipe		1
147	Tanzania		1

and maximum is:

	Nationality	players	count
47	England		1706

خروجی قسمت e:

```

-----PART E-----

players with growth < 3 and potential < 84 are:

```

ID	Name	FullName	Age	Height	...	RWBRating	LBRating	CBRating	RBRating	GKRating
139	Renato Augusto	Renato Augusto	32	186	...	79	77	75	77	21
140	C. Vela	Carlos Vela	31	177	...	62	58	50	58	22
143	D. Zapata	Duván Zapata	29	189	...	56	54	56	54	18
144	Rafa	Rafael A. Ferreira Silva	27	172	...	72	68	56	68	20
147	M. Dúbravka	Martin Dúbravka	31	190	...	33	32	32	32	83
...	...	...	...	...	...	...	...	...	...	...
18684	Xu Wu	Wu Xu	27	184	...	52	52	52	52	16
18742	O. Alici	Ömer Alici	34	185	...	17	17	22	17	51
18888	J. Russell	John Russell	35	172	...	49	49	50	49	16
18968	Teng Shangkun	Shangkun Teng	29	186	...	23	23	24	23	51
18986	Li Hao	Hao Li	28	173	...	51	51	51	51	16

[8261 rows x 90 columns]

خروجی قسمت f:

```

-----PART F-----

past report with position:

```

ID	Name	Growth	Potential	Positions
139	Renato Augusto	0	83	CM,CAM,LM
140	C. Vela	0	83	RW,LW,CAM
143	D. Zapata	0	83	ST
144	Rafa	0	83	LM,RM
147	M. Dúbravka	0	83	GK
...	...	...	...	...
18684	Xu Wu	2	54	CDM
18742	O. Alici	0	51	GK
18888	J. Russell	0	50	CM,CDM
18968	Teng Shangkun	1	51	GK
18986	Li Hao	1	51	CDM,RM

[8261 rows x 4 columns]

خروجی قسمت h و i:

```

-----PART H-----

number of players contract until 2021 and not in national team: 1223
those players are:

```

ID	Name	FullName	Age	Height	...	RWBRating	LBRating	CBRating	RBRating	GKRating
441	Hulk	Givanildo Vieira de Souza	33	180	...	68	65	63	65	20
477	Pozuelo	Alejandro Pozuelo Melero	28	170	...	69	65	58	65	19
622	C. Pavón	Cristian Pavón	24	174	...	69	66	58	66	20
658	Alex Teixeira	Alex Teixeira dos Santos	30	174	...	68	64	56	64	20
694	I. Piatti	Ignacio Piatti	35	180	...	66	62	61	62	21
...	...	...	...	...	...	...	...	...	...	...
19004	J. O'Brien	Jake O'Brien	19	190	...	46	48	51	48	16
19008	Wen Da	Da Wen	20	185	...	39	37	32	37	16
19011	S. Doherty	Stephen Doherty	20	176	...	45	44	39	44	17
19015	M. Nzongong	Mike Nzongong	19	179	...	44	42	40	42	18
19018	R. Dinanga	Ricardo Dinanga	18	174	...	36	34	30	34	16

[1223 rows x 90 columns]

```

-----PART I-----

chelsea under 24 years old's value:

```

ID	Name	Age	ValueEUR
2647	C. Musonda	23	4800000
7888	D. Sterling	20	2300000
16763	J. Wakely	19	500000

خروجی قسمت l:

-----PART J-----

hazard's info is:

	Name	Club	Positions	WageEUR
22	E. Hazard	Real Madrid	LW	350000