

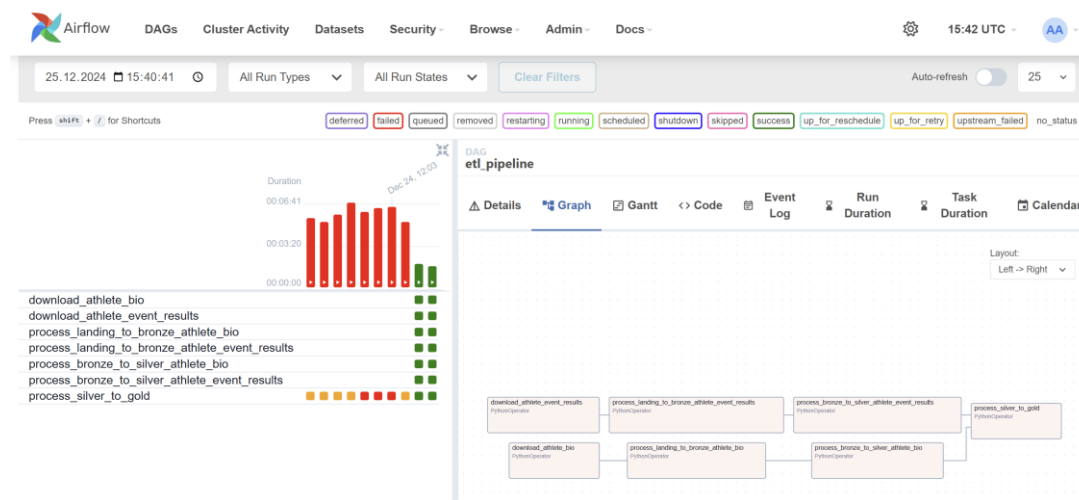
Завдання 1.

Дані в топіку athlete_event_results

```
C:\Windows\system32\cmd.exe
aggregated_results
athlete_event_results

C:\Users\Asus>kafka-console-consumer --bootstrap-server localhost:9092 --topic athlete_event_results --from-beginning --max-messages 10
{"result_id":1626,"event_title":"GiantsSlalom_Women","edition":"1998WinterOlympics","edition_id":46,"sport":"Snowboarding","sport_url":"editions46sportsSBD","result_date":"9February1998","result_location":"Mt. Yakebitai, ShigaKogen, Yamanouchi","result_participants":"31from14countries","result_format":"Tworuns,totaledeterminedplacement.","result_detail":"Gates3836length936mStartAltitude1968mVerticalDrop290m","result_description":"The women's giant slalom was postponed one day because of a snowstorm that interrupted the competition. The heavy favorite was Frances Karine Ruby, who had won a silver medal in the 1994 FIS World Championships. On the first run, she overcame the field, leading by 1.95 seconds over her teammate, Isabelle Blanc. On the second run, Ruby boarded carefully to protect her lead but still won the gold medal by a margin of two seconds over Germanys Heidi Renoth. Blanchad the silver medal in her grasp when she missed the final gate on her second run, and fell, sliding across the finish line. Ruby would compete again at the 2002 and 2006 Winter Olympics, winning a silver medal in 2002 in parallel giant slalom, and switching to bobsledding for the 2006 Winter Olympics. She then began training to become a mountainguide and was killed in late May 2009 when she fell into a crevasse on Mont Blanc. Isabelle Blanc returned in 2002 to win the gold medal in the parallel giant slalom."}
{"result_id":76,"event_title":"Singles_Men","edition":"1976WinterOlympics","edition_id":40,"sport":"Luge","sport_url":"editions40sportsLUG","result_date":"47February1976","result_location":"Kunsteis Bobund Rodelbahn, Igls","result_participants":"43from15countries","result_format":"Fourruns,totaledeterminedplacement.","result_detail":"Curves14Length1220mStartAltitudeVerticalDrop103m","result_description":"Once more, the competitors from East and West Germany divided the medals among them. The strength of the East German team was shown from the fact that the reigning World Champion and 1972 bronze medalist, Wolfgang Fiedler, did not make the cut. This left 1973 World Champion Hans Rinnas as the clear favorite. But he made a serious error in the first run, putting him 0.6 seconds behind the leader, who had been the 1975 European Champion, but had only managed 7th spot at the World Championships. On the second run, giving him a comfortable lead. Only West German Josef Fend could seriously challenge him, but even the former World Champion 1978, 1973 had to concede half a second in the final standings. The unlucky skier in Innsbruck was without a doubt Huang Liu Chong. He was actually the best non-European, as shown by his performances in the first and fourth run, but he finished last because of two crashes in the other two runs. Remarkably, he did finish both times, but ended up more than one minute behind the next to last luger, a huge gap in luge terms."}
{"result_id":962,"event_title":"1,500metres,Men","edition":"1928WinterOlympics","edition_id":30,"sport":"SpeedSkating","sport_url":"editions30sportsSSK","result_date":"14February1928","result_location":"Olympia Eisstadion Badrutts Park, St. Moritz","result_participants":"30from14countries","result_format":"na","result_detail":"na","result_description":"There was a little doubt that the Olympic 500m title would be retained by Clas Thunberg. He had clocked a 21.8 in the day earlier in Davos, and all of the other season best times were at least 1.8 seconds slower. In fact, Thunberg had not lost a 500m race since winning this distance at the Finnish, European and World Championships. Behind him, the Norwegians were expected to divide the medals, with the Americans as outsiders. The Americans opened, with Eddie Murphy 225.9 and Irving Jaffee 226.7 setting the pace. The third American, John O'Neil Farrell, was beaten by Roald Larsen of Norway, who also moved in front. Two more Norwegians followed, Olympic 500m champion Ivar Ballangrud and Voller Nygren. The latter, a substitute for the sick Staksrud, had done well in Davos, but disappeared in St. Moritz. Not so for Ballangrud, who took the lead with a very decent 222.6. He didn't have to wait long to see what his time was worth. The very next race paired Thunberg and Bernt Evensen, the fastest of the Norwegians in the metric mile. The two fought a great race, and were still level coming out of the final bend. Thunberg then emerged as the better sprinter, winning in 221.1, eight tenths ahead of Evensen. The twelfth place of Charlie Gorman is somewhat contentious. He showed up late for the start, but was allowed to skate against Sieme de Groot in the final pair, but out of competition. The Official Report mentions that his time would be noted, but would not count. Yet, he is listed as 12th in several newspapers, while the other 1928 sources do list him as out of competition."}
{"result_id":901297,"event_title":"800metresWheelchair_Women","edition":"2004SummerOlympics","edition_id":26,"sport":"Athletics","sport_url":"editions26sportsATH","result_date":"22August2004","result_location":"OlympiakoStadio, OlympiakoAthletikoKentroAthinaSpirosLoues, Maroussi","result_participants":"77from4countries","result_format":"na","result_detail":"na","result_description":"Wheelchair racing was held as a demonstration event on 2004."}
```

Завдання 2.



download_athlete_bio:

Task Instance: download_athlete_bio at 2024-12-25, 14:59:59

[Grid](#) [Task Instance Details](#) [<> Rendered Template](#) [Log](#) [XCom](#)

Log by attempts

1

```
*** Found local files:
*** * /opt/airflow/logs/dag_id=etl_pipeline/run_id>manual__2024-12-25T14:59:59.896378+00:00/task_id=download_athlete_bio/attempt=1.1
[2024-12-25, 15:00:02 UTC] {local_task_job_runner.py:123} ► Pre task execution logs
[2024-12-25, 15:00:45 UTC] {logging_mixin.py:190} INFO - File athlete_bio.csv downloaded successfully.
[2024-12-25, 15:00:45 UTC] {python.py:240} INFO - Done. Returned value was: None
[2024-12-25, 15:00:45 UTC] {taskinstance.py:340} ► Post task execution logs
```

download_athlete_event_results:

Task Instance: download_athlete_event_results at 2024-12-25, 14:59:59

[Grid](#) [Task Instance Details](#) [<> Rendered Template](#) [Log](#) [XCom](#)


Log by attempts

1

[Jump To End](#)

```
*** Found local files:
*** * /opt/airflow/logs/dag_id=etl_pipeline/run_id>manual__2024-12-25T14:59:59.896378+00:00/task_id=download_athlete_event_results/attempt=1.log
[2024-12-25, 15:00:03 UTC] {local_task_job_runner.py:123} ► Pre task execution logs
[2024-12-25, 15:00:34 UTC] {logging_mixin.py:190} INFO - File athlete_event_results.csv downloaded successfully.
[2024-12-25, 15:00:34 UTC] {python.py:240} INFO - Done. Returned value was: None
[2024-12-25, 15:00:34 UTC] {taskinstance.py:340} ► Post task execution logs
```

landing_to_bronze_athlete_bio:

 Airflow DAGs Cluster Activity Datasets Security Browse Admin Docs

Log by example


1 [Jump To End](#)

```
*** Found local files:
*** * /opt/airflow/logs/dag_id=etl_pipeline/run_id>manual_2024-12-25T14:59:59.896378+00:00/task_id=process_landing_to_bronze_athlete_bio/attempt=1
[2024-12-25, 15:00:48 UTC] {local_task_job_runner.py:123} ▶ Pre task execution logs
[2024-12-25, 15:00:58 UTC] {logging_mixin.py:190} INFO -
```

athlete_id	name	sex	born	height	weight	country	country_noc	description	special_notes
65649	Ivanka Bonova	Female	4 April 1949	166	55	Bulgaria	BUL	Personal Best: 40...	null
112510	Nataliya Uryadova	Female	15 March 1977	184	70	Russian Federation	RUS		Listed in Olympia...
114973	Essa Ismail Rashed	Male	14 December 1986	165	55	Qatar	QAT	Personal Best: 10...	Listed in Olympia...
30359	Péter Boros	Male	12 January 1908	null	null	Hungary	HUN	Between 1927 and ...	null
50557	Rudolf Piowaty	Male	28 April 1900	null	null	Czechoslovakia	TCH	Rudolf Piowaty jo...	null
146111	Svetlana Kholomina	Female	9 November 1997	null	null	ROC	ROC		null
133041	Vincent Riendeau	Male	13 December 1996	178	68	Canada	CAN		Listed in Olympia...
110425	Tanja Morel	Female	4 October 1975	164	58	Switzerland	SUI		Listed in Olympia...
110705	Maksim Shabalín	Male	25 January 1982	183	76	Russian Federation	RUS		Listed in Olympia...
54541	Gé Regter	Male	6 March 1916	null	null	Netherlands	NED		null
22721	Aristide Pontenani	Male	null	null	null	Italy	ITA		null
56266	Go Yeong-Chang	Male	21 March 1926	167	75	Republic of Korea	KOR		null
82227	Marlies Rostock	Female	20 April 1960	null	null	East Germany	GDR	Marlies Rostock w...	Wife (divorced) o...
93334	Craig Hutchison	Male	26 May 1975	198	97	Canada	CAN		Listed in Olympia...
146013	Raquel Queirós	Female	4 March 2000	167	56	Portugal	POR		null
109912	Vyacheslav Kurginyan	Male	22 December 1986	170	65	Russian Federation	RUS		Listed in Olympia...
37019	Philippe Lot	Male	13 June 1967	189	85	France	FRA		Listed in Olympia...
22885	Rudy Kugeler	Male	11 August 1928	187	86	Luxembourg	LUX		null
95497	Yoshihiro Miyazaki	Male	10 May 1930	173	70	Japan	JPN		null
76	Roper Barrett	Male	24 November 1873	null	null	Great Britain	GBR	Roper Barrett lea...	null

only showing top 20 rows

landing_to_bronze_athlete_event_results:

 Airflow DAGs Cluster Activity Datasets Security Browse Admin Docs

Log by example


1 [Jump To End](#)

```
*** Found local files:
*** * /opt/airflow/logs/dag_id=etl_pipeline/run_id>manual_2024-12-25T14:59:59.896378+00:00/task_id=process_landing_to_bronze_athlete_event_results/attempt=1
[2024-12-25, 15:00:36 UTC] {local_task_job_runner.py:123} ▶ Pre task execution logs
[2024-12-25, 15:00:45 UTC] {logging_mixin.py:190} INFO -
```

edition	edition_id	country_noc	sport	event	result_id	athlete	athlete_id	pos	medal	isTeamSport
1908 Summer Olympics	5	ANZ	Athletics	100 metres, Men	56265	Ernest Hutcheon	64710	DNS	null	false
1908 Summer Olympics	5	ANZ	Athletics	400 metres, Men	56313	Henry Murray	64756	DNS	null	false
1908 Summer Olympics	5	ANZ	Athletics	800 metres, Men	56338	Harvey Sutton	64808	3 h8 r1/2	null	false
1908 Summer Olympics	5	ANZ	Athletics	800 metres, Men	56338	Guy Haskins	922519	DNS	null	false
1908 Summer Olympics	5	ANZ	Athletics	800 metres, Men	56338	Joseph Lynch	64735	DNS	null	false
1908 Summer Olympics	5	ANZ	Athletics	800 metres, Men	56338	Henry Murray	64756	DNS	null	false
1908 Summer Olympics	5	ANZ	Athletics	1,500 metres, Men	56349	Joseph Lynch	64735	5 h2 r1/2	null	false
1908 Summer Olympics	5	ANZ	Athletics	1,500 metres, Men	56349	Charles Swain	79576	AC h3 r1/2	null	false
1908 Summer Olympics	5	ANZ	Athletics	1,500 metres, Men	56349	Guy Haskins	922519	DNS	null	false
1908 Summer Olympics	5	ANZ	Athletics	1,500 metres, Men	56349	George Blake	64619	DNS	null	false
1908 Summer Olympics	5	ANZ	Athletics	5 miles, Men	56360	George Blake	64619	3 h1 r1/2	null	false
1908 Summer Olympics	5	ANZ	Athletics	5 miles, Men	56360	Joseph Lynch	64735	AC h5 r1/2	null	false
1908 Summer Olympics	5	ANZ	Athletics	5 miles, Men	56360	Guy Haskins	922519	DNS	null	false
1908 Summer Olympics	5	ANZ	Athletics	5 miles, Men	56360	Victor Aitken	64594	DNS	null	false
1908 Summer Olympics	5	ANZ	Athletics	Marathon, Men	56369	Victor Aitken	64594	DNF	null	false
1908 Summer Olympics	5	ANZ	Athletics	Marathon, Men	56369	George Blake	64619	DNF	null	false
1908 Summer Olympics	5	ANZ	Athletics	Marathon, Men	56369	Joseph Lynch	64735	DNF	null	false
1908 Summer Olympics	5	ANZ	Athletics	110 metres Hurdle...	56370	Henry Murray	64756	2 h1 r1/3	null	false
1908 Summer Olympics	5	ANZ	Athletics	400 metres Hurdle...	56392	Henry Murray	64756	2 h3 r1/3	null	false
1908 Summer Olympics	5	ANZ	Athletics	3,200 metres Stee...	56412	Guy Haskins	922519	DNS	null	false

only showing top 20 rows



bronze_to_silver_athlete_bio:

 Airflow DAGs Cluster Activity Datasets Security Browse Admin Docs

1 Jump

```
*** Found local files:
*** * /opt/airflow/logs/dag_id=etl_pipeline/run_id=manual__2024-12-25T14:59:59.896378+00:00/task_id=process_bronze_to_silver_athlete_bio/a
[2024-12-25, 15:01:11 UTC] {local_task_job_runner.py:123} ▶ Pre task execution logs
[2024-12-25, 15:01:21 UTC] {logging_mixin.py:190} INFO - +-----+-----+-----+-----+-----+-----+-----+-----+
|athlete_id|      name|  sex|   born|height|weight|   country|country_noc|   description|   special_notes|
+-----+-----+-----+-----+-----+-----+-----+-----+
| 58415| Franz Schmitt| Male| 5 September 1937| 170| 72| Germany| GER|Greco-Roman light...| null|
| 57779| Mahmoud Hassan| Male| 15 December 1918| null| null| Egypt| EGY|Mahmoud Hassan to...|Listed in Olympia...|
| 74884| Kazimierz Zimny| Male| 4 June 1935| 172| 60| Poland| POL|Kazimierz Zimny w...|Listed in Olympia...|
| 58947| Riccardo Pizzocaro| Male| 26 March 1901| null| null| Italy| ITA| null| null|
| 28093| Edeltraud Schramm| Female| 16 December 1923| null| null| Austria| AUT|Edeltraud Schramm...| null|
| 117651| Samantha Peszek| Female| 14 December 1991| 152| 50| United States| USA| null|Listed in Olympia...|
| 112730| Levan Zhorzholiani| Male| 20 January 1988| 192| 100| Georgia| GEO| null| null|
| 67311| Ignacio Sola| Male| 1 February 1944| 179| 75| Spain| ESP|Personal Best: PV...|Listed in Olympia...|
| 2091| Gloria Rosa| Female| 4 February 1958| null| null| Puerto Rico| PUR| null|Listed in Olympia...|
| 103286| Cathrin Carlzon| Female| 25 April 1983| 173| 73| Sweden| SWE| null|Listed in Olympia...|
| 127630| Gary di Silvestri| Male| 3 February 1967| 183| 82| Dominica| DMA|Gary di Silvestri...|Husband of Angeli...|
| 106476| Barbara Madejczyk| Female| 30 September 1976| 180| 81| Poland| POL|Personal Best: JT...|Listed in Olympia...|
| 106653| Alexandre Lee| Male| 2 September 1977| 170| 60| Brazil| BRA| null|Listed in Olympia...|
| 93821| Aytekin Mindan| Male| 31 August 1981| 193| 70| Türkiye| TUR| null| null|
| 125119| Mária Czaková| Female| 2 October 1988| 166| 56| Slovakia| SVK| null| null|
| 32880| Vladimír Belov| Male| 26 April 1958| 195| 92| Soviet Union| URS|A 1.95 cm tall ce...| null|
| 43517| John H. Larsen, Jr.| Male| 18 August 1943| 170| 66| Norway| NOR| null|Son of John H. La...|
| 50816| Nikolay Yevseyev| Male| 16 April 1966| 190| 81| Soviet Union| URS|Nikolay Yevseyev ...|Husband of Angela...|
| 143367| Robert Farken| Male| 20 September 1997| 186| 79| Germany| GER| null| null|
| 50214| Therèse Alshammar| Female| 26 August 1977| 180| 64| Sweden| SWE| null|Daughter of Britt...|
+-----+-----+-----+-----+-----+-----+-----+-----+
only showing top 20 rows
```


bronze_to_silver_athlete_event_results:

 Airflow DAGs Cluster Activity Datasets Security Browse Admin Docs  15:10

1 Jump To End Toggle

```
*** Found local files:
*** * /opt/airflow/logs/dag_id=etl_pipeline/run_id=manual__2024-12-25T14:59:59.896378+00:00/task_id=process_bronze_to_silver_athlete_event_results/attempte
[2024-12-25, 15:00:49 UTC] {local_task_job_runner.py:123} ▶ Pre task execution logs
[2024-12-25, 15:01:05 UTC] {logging_mixin.py:190} INFO - +-----+-----+-----+-----+-----+-----+-----+-----+
|      edition|edition_id|country_noc|      sport|      event|result_id| athlete|athlete_id|   pos|medal|isTeamSport|
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|1908 Summer Olympics| 5| BOH| Athletics|Javelin Throw, Fr...| 56522| Miroslav Šuster| 65445| DNS| null| false|
|1908 Summer Olympics| 5| BOH| Tennis| Doubles, Men| 44260| Károly Vitos| 79929| DNS| null| true|
|1988 Summer Olympics| 22| BRN| Athletics|3,000 metres Stee...| 63403| Abdullah Al-Dosari| 65635|11 h3 r1/3| null| false|
|1988 Summer Olympics| 22| BRN| Modern Pentathlon| Individual, Men| 50225|Saleh Sultan Farh...| 21360| 47| null| false|
|1992 Summer Olympics| 23| BRN| Athletics| 100 metres, Men| 63952|Khaled Ibrahim Jouma| 65640| 6 h1 r1/4| null| false|
|2016 Summer Olympics| 59| BAR| Athletics| 200 metres, Men| 358579| Ramon Gittens| 121363| 3 h5 r1/3| null| false|
|2008 Summer Olympics| 53| ANG| Handball| Handball, Women| 260050| Odeth Tavares| 91605| 12| null| true|
|2012 Summer Olympics| 54| ANG| Basketball| Basketball, Women| 303046| Felizarda Jorge| 124457| 12| null| true|
|1936 Summer Olympics| 11| AFG| Hockey| Hockey, Men| 35547| Zahir Shah Al-Zadah| 17345| =5| null| true|
|1956 Summer Olympics| 14| AFG| Hockey| Hockey, Men| 35641|Mohammad Anis She...| 17338| =11| null| true|
|1960 Summer Olympics| 15| AFG| Athletics|4 x 100 metres Re...| 59753| Ali Zaid| 64265| 4 h4 r1/3| null| true|
|2012 Summer Olympics| 54| AZE| Wrestling|Heavyweight, Grec...| 332230| Salva Qadabadze| 117359| 10| null| false|
|1996 Summer Olympics| 24| IVB| Athletics| High Jump, Women| 65154| Coralea Cline| 923789| DNS| null| false|
|2012 Summer Olympics| 54| BUR| Athletics|100 metres Hurdle...| 302235| Marthe Koala| 121488| 7 h3 r1/3| null| false|
|1994 Winter Olympics| 45| AND| Alpine Skiing| Giant Slalom, Men| 1410| Victor Gómez| 86127| DNF| null| false|
|2016 Summer Olympics| 59| AND| Shooting|Air Rifle, 10 met...| 395754| Esther Barrugués| 128861| 51| null| false|
|1948 Summer Olympics| 12| BER| Athletics| 400 metres, Men| 58528| Hazzard Dill| 65364| 3 h6 r1/4| null| false|
|1960 Summer Olympics| 15| BER| Sailing|Three Person Keel...| 9707| Dick Masters| 62928| 15| null| true|
|2000 Summer Olympics| 25| BER| Equestrian Eventing| Individual, Open| 185467| MJ Tumbridge| 11485| DNF| null| false|
|1984 Summer Olympics| 21| ALG| Boxing| Welterweight, Men| 25395| Kamel Aboud| 7423| =9| null| false|
+-----+-----+-----+-----+-----+-----+-----+-----+
only showing top 20 rows
```

silver_to_gold (агрегуюча таблиця):

 Airflow

DAGs

Cluster Activity

Datasets

Security

Browse

Admin

Docs

```
*** * /opt/airflow/logs/dag_id=etl_pipeline/run_id>manual__2024-12-25T14:59:59.896378+00:00/task_id=process
[2024-12-25, 15:01:26 UTC] {local_task_job_runner.py:123} ▶ Pre task execution logs
[2024-12-25, 15:01:37 UTC] {logging_mixin.py:190} INFO - +-----+-----+-----+-----+-----+-----+-----+
| sport | medal | sex | bio_country_noc | avg_height | avg_weight | timestamp |
+-----+-----+-----+-----+-----+-----+-----+
| Sailing | Silver | Male | GDR | 178.66666666666666 | 82.83333333333333 | 2024-12-25 15:01:... |
| Equestrian Jumping | Gold | Male | USA | 182.11111111111111 | 71.66666666666667 | 2024-12-25 15:01:... |
| Canoe Marathon | Gold | Male | URS | null | null | 2024-12-25 15:01:... |
| Archery | Silver | Male | FRA | 180.0 | 83.0 | 2024-12-25 15:01:... |
| Figure Skating | Bronze | Male | HUN | 168.0 | 63.0 | 2024-12-25 15:01:... |
| Ice Hockey | Bronze | Male | CZE | 184.44736842105263 | 88.86842105263158 | 2024-12-25 15:01:... |
| Judo | Silver | Female | JPN | 160.6 | 62.55555555555556 | 2024-12-25 15:01:... |
| Swimming | null | Female | KGZ | 173.66666666666666 | 61.26666666666666 | 2024-12-25 15:01:... |
| Sailing | null | Male | ITA | 179.26351351351352 | 78.32432432432432 | 2024-12-25 15:01:... |
| Judo | null | Male | USA | 179.3684210526316 | 84.52631578947368 | 2024-12-25 15:01:... |
| Basketball | null | Male | PHI | 181.2972972972973 | 76.21621621621621 | 2024-12-25 15:01:... |
| Artistic Swimming | null | Female | AUS | 167.06349206349208 | 56.34920634920635 | 2024-12-25 15:01:... |
| Archery | null | Female | SWE | 169.03225806451613 | 61.83870967741935 | 2024-12-25 15:01:... |
| Athletics | null | Male | DOM | 174.0952380952381 | 68.03174603174604 | 2024-12-25 15:01:... |
| Fencing | null | Female | ITA | 169.3695652173913 | 60.70652173913044 | 2024-12-25 15:01:... |
| Football | null | Male | UAE | 175.5 | 70.04545454545455 | 2024-12-25 15:01:... |
| Table Tennis | null | Female | TPE | 163.5625 | 55.6875 | 2024-12-25 15:01:... |
| Luge | null | Male | UKR | 183.06060606060606 | 82.92307692307692 | 2024-12-25 15:01:... |
| Athletics | null | Male | MON | 182.4 | 86.8 | 2024-12-25 15:01:... |
| Athletics | null | Female | KGZ | 173.0 | 58.04761904761905 | 2024-12-25 15:01:... |
+-----+-----+-----+-----+-----+-----+-----+
only showing top 20 rows
```