

# Фінальний проєкт

## Частина 1. Building an End-to-End Streaming Pipelin

roman\_enriched\_athlete\_avg X

ВластивостіДаніER Діаграма

roman\_enriched\_athlete\_avg Введіть вираз SQL для фільтрації результатів (використовуйте Ctrl+Пробіл)

СіткаТекстІм'я

	A-Z sport	A-Z medal	A-Z sex	A-Z country_noc	123 avg_height	123 avg_weight	timestamp
1	Judo	nan	Male	ARU	182	81	2025-06-10 17:55:00
2	Swimming	nan	Male	ARU	177	70	2025-06-10 17:55:10
3	Judo	nan	Male	ARU	177	74	2025-06-10 17:55:10
4	Swimming	nan	Male	ARU	184	75	2025-06-10 17:55:17
5	Judo	nan	Male	ARU	177	74	2025-06-10 17:55:17
6	Swimming	nan	Female	ARU	165	61	2025-06-10 17:55:17
7	Athletics	nan	Male	BAN	173	74	2025-06-10 17:55:25
8	Swimming	nan	Male	ARU	184	75	2025-06-10 17:55:25
9	Judo	nan	Male	ARU	177	74	2025-06-10 17:55:25
10	Swimming	nan	Female	ARU	165	61	2025-06-10 17:55:25
11	Sailing	nan	Female	ARU	171	65	2025-06-10 17:55:25
12	Athletics	nan	Male	BAN	173	74	2025-06-10 17:55:31
13	Weightlifting	nan	Male	ARU	178	150	2025-06-10 17:55:31
14	Swimming	nan	Male	ARU	184	75	2025-06-10 17:55:31
15	Judo	nan	Male	ARU	177	74	2025-06-10 17:55:31
16	Swimming	nan	Female	ARU	165	61	2025-06-10 17:55:31
17	Sailing	nan	Female	ARU	171	65	2025-06-10 17:55:31
18	Athletics	nan	Male	BAN	173	74	2025-06-10 17:55:39
19	Weightlifting	nan	Male	ARU	178	150	2025-06-10 17:55:39
20	Swimming	nan	Female	BAN	158	53	2025-06-10 17:55:39
21	Athletics	nan	Female	BAN	157	56	2025-06-10 17:55:39
22	Swimming	nan	Male	ARU	184	75	2025-06-10 17:55:39
23	Judo	nan	Male	ARU	177	74	2025-06-10 17:55:39
24	Swimming	nan	Female	ARU	165	61	2025-06-10 17:55:39
25	Sailing	nan	Female	ARU	171	65	2025-06-10 17:55:39
26	Athletics	nan	Male	BAN	173	74	2025-06-10 17:55:46
27	Weightlifting	nan	Male	ARU	178	150	2025-06-10 17:55:46
28	Swimming	nan	Female	BAN	158	53	2025-06-10 17:55:46
29	Athletics	nan	Female	BAN	157	56	2025-06-10 17:55:46
30	Swimming	nan	Male	ARU	184	75	2025-06-10 17:55:46
31	Judo	nan	Male	ARU	175.3333333333	71.6666666667	2025-06-10 17:55:46
32	Swimming	nan	Female	ARU	165	61	2025-06-10 17:55:46
33	Sailing	nan	Female	ARU	171	65	2025-06-10 17:55:46
34	Athletics	nan	Male	BAN	173	74	2025-06-10 17:55:52
35	Weightlifting	nan	Male	ARU	178	150	2025-06-10 17:55:52
36	Athletics	nan	Female	BAN	157	56	2025-06-10 17:55:52
37	Swimming	nan	Female	BAN	158	53	2025-06-10 17:55:52
38	Swimming	nan	Male	ARU	184	75	2025-06-10 17:55:52
39	Judo	nan	Male	ARU	175.3333333333	71.6666666667	2025-06-10 17:55:52
40	Athletics	nan	Male	ANZ	167	62	2025-06-10 17:55:52
41	Swimming	nan	Female	ARU	165	60.5	2025-06-10 17:55:52
42	Sailing	nan	Female	ARU	171	65	2025-06-10 17:55:52
43	Athletics	nan	Male	BAN	173	74	2025-06-10 17:55:59

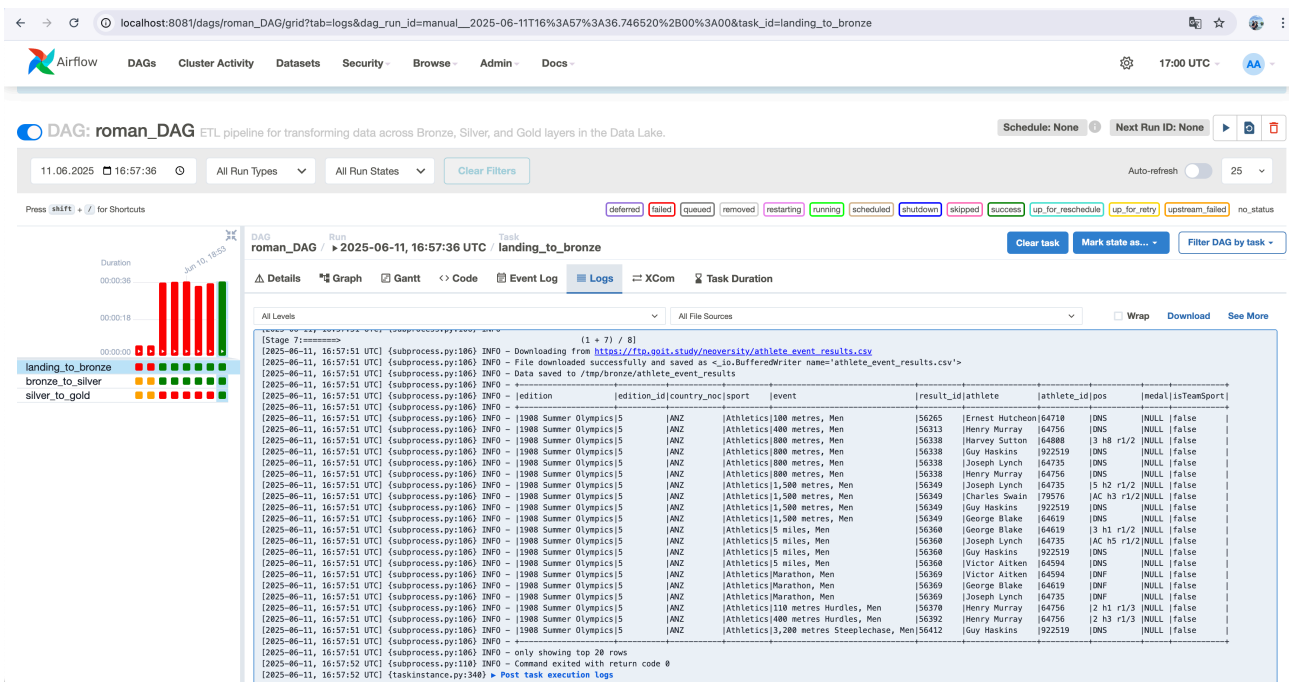
Скріншот бази даних MySQL який підтверджує, що створилась таблиця roman\_enriched\_athlete\_avg та в неї записались дані.

sport	medal	sex	noc_country	avg_height	avg_weight	timestamp
Sailing	nan	Female	NULL	169.5	63.5	2025-06-10T17:56:13.906+02:00
Athletics	nan	Male	NULL	173.0	74.0	2025-06-10T17:56:20.842+02:00
Sailing	nan	Male	NULL	184.0	88.0	2025-06-10T17:56:20.842+02:00
Weightlifting	nan	Male	NULL	178.0	150.0	2025-06-10T17:56:20.842+02:00
Rowing	nan	Male	NULL	189.0	85.0	2025-06-10T17:56:20.842+02:00
Taekwondo	nan	Female	NULL	171.0	49.0	2025-06-10T17:56:20.842+02:00
Athletics	nan	Female	NULL	157.0	56.0	2025-06-10T17:56:20.842+02:00
Swimming	nan	Female	NULL	158.0	53.0	2025-06-10T17:56:20.842+02:00
Swimming	nan	Male	NULL	168.0	62.0	2025-06-10T17:56:20.842+02:00
Shooting	nan	Male	NULL	169.0	50.0	2025-06-10T17:56:20.842+02:00
Swimming	nan	Male	NULL	184.0	75.0	2025-06-10T17:56:20.842+02:00
Judo	nan	Female	NULL	170.0	59.0	2025-06-10T17:56:20.842+02:00
Judo	nan	Male	NULL	175.33333333333334	71.66666666666667	2025-06-10T17:56:20.842+02:00
Athletics	nan	Male	NULL	167.0	62.0	2025-06-10T17:56:20.842+02:00
Swimming	nan	Female	NULL	165.0	60.5	2025-06-10T17:56:20.842+02:00
Sailing	nan	Female	NULL	169.5	63.5	2025-06-10T17:56:20.842+02:00
Athletics	nan	Male	NULL	173.0	74.0	2025-06-10T17:56:28.833+02:00
Sailing	nan	Male	NULL	184.0	88.0	2025-06-10T17:56:28.833+02:00
Weightlifting	nan	Male	NULL	178.0	150.0	2025-06-10T17:56:28.833+02:00
Rowing	nan	Male	NULL	189.0	85.0	2025-06-10T17:56:28.833+02:00

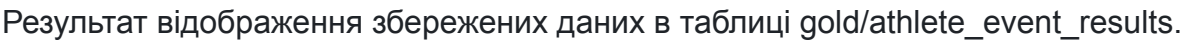
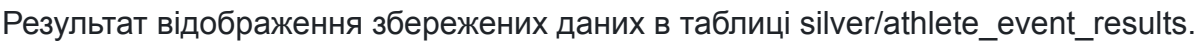
Скріншот підтверджує запис даних у kafka-топик.

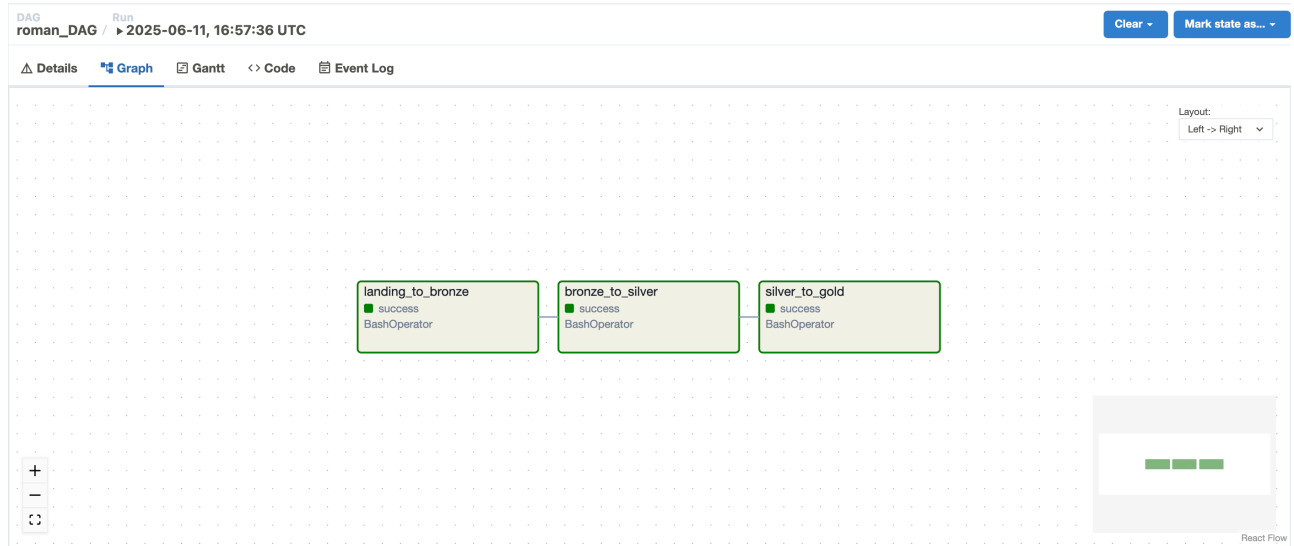
## Частина 2. Building an End-to-End Batch Data Lake

У цьому скріншоті відображено DAG з трьома успішно відпрацьованими батчами.



Результат відображення збережених даних в таблиці bronze/athlete\_event\_results.





Граф відпрацьованого DAGу.