

Task 1 Presentation - ASE 2024/25

Robert Richter 26.11.2024

Reflection Task 1 - Implementing the Query

Own Implementation:

- + The performance was better compared to ChatGPT's variant (30s vs. 90s)
- + More correct variant (⇒ ChatGPT did not realize a person is *not* foreign to themself)
- It took quite a while to develop the code
- Bad readability

ChatGPT's Implementation:

- + Quick working solution (⇒ Copy-pasting task was nearly sufficient)
- + Perfectly Readable
- Worse Performance (not relevant for smaller one-time queries..)

Reflection Task 2 - Enhancing the Query

Own Implementation:

- + Still better performance (45s vs. 120s)
- High effort (≈ an hour of work)
- Code became messy

ChatGPT's Implementation:

- + No knowledge required (again simply copy-pasting of new task)
- + Highly readable
- Less performance

Reflection Task 3 - DBeaver Query Planning

Sort 46770350.11 - 46770350.10 154	Knotentyp	Entität	Kosten	Reihen	Zeit	Bedingung
Y Hash Join 4715204.77 - 46770295.71 154 (foreignlikes.l_messageid) = m_1.m_messageid) Y Subquery Scan 4703936.19 - 46759027.12 4476 (foreignlikes.r.m = 1) Y WindowAgg 4703936.19 - 46759027.05 57521 (foreignlikes.r.m = 1) Y WindowAgg 4703936.19 - 46759027.05 57521 (k2.k_person1id is null) Y Nested Loop 4703936.19 - 467461813.85 57521 (k2.k_person1id is null) Y Nested Loop 4702936.17 - 46745178.47 19174 19174 Y Nested Loop 31575.95 - 46745178.02 19174 Y Nested Loop 31575.95 - 46745178.02 19174 Y Nested Loop 31575.95 - 34674012.02 36480 Y Nested Loop 31575.95 - 346740.02 36480 Y Nested Loop 31575.95 - 346740.02 36480 Y Sort 18370.49 - 18669.18 95553 Parallel Seq Sca message 0.00 - 6250.77 95581 Y Materialize 1305.41 - 13752.61 101849 Y Seq Scan knows 0.00 - 461.46 18188 Y Materialize 0.00 - 602.19 28146 Y Hash 1160.04 - 11241.44 345 Y Aggregate 11160.04 - 11241.44 345 <td>∨ Sort</td> <td></td> <td>46770350.11 - 46770350.1</td> <td>154</td> <td></td> <td></td>	∨ Sort		46770350.11 - 46770350.1	154		
Subquery Scan 4703936.19 - 46759027.12 4476 (foreignlikes.rn = 1) ✓ WindowAgg 4703936.19 - 46759027.11 4476 ((row_number() over (?)) <= 1)	∨ Nested Loop		4715204.77 - 46770350.10	154		
✓ WindowAgg 4703936.19 - 46759027.11 4476 ((row_number() over (?)) <= 1)	→ Hash Join		4715204.77 - 46770295.71	154		(foreignlikes.l_messageid = m_1.m_messageid)
V WindowAgg 4703936.19 - 46759027.0E 57521 ((k2.k_person1id is null) V Nested Loop 4703936.19 - 46746181.38 57521 ((k2.k_person1id is null) V Gather Merge 4703936.17 - 46745178.47 19174 (k2.k_person1id is null) V Incremental Sort 4702936.17 - 46745178.47 19174 V Nested Loop 31575.95 - 346745178.02 19174 V Merge Join 31575.95 - 33450.02 36480 V Sort 18370.49 - 18669.18 95553 Parallel Seq Sca message 0.00 - 6250.77 95581 V Materialize 13205.41 - 13752.61 101849 V Sort 13205.41 - 13479.01 101849 Seq Scan kines 0.00 - 1798.40 109440 Seq Scan knows 0.00 - 461.46 18188 V Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 V Hash 11241.44 - 11241.44 345 V Hash 1160.04 - 11241.44 345 (count([-1.L_personid) >= 20) V Hash Join 8853.65 - 11127.48 21623 ([-1.L_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 V Hash 8640.30 - 8640.30 17445	 Subquery Scan 		4703936.19 - 46759027.12	4476		(foreignlikes.rn = 1)
V Nested Loop 4703936.19 - 46759027.0E 57521 ((k2.k_person1id is null) C Gather Merge 4703936.19 - 46746181.38 57521 ((k2.k_person1id is null) V Incremental Sort 4702936.17 - 46745181.78 47 19174 Nested Loop 31575.95 - 46745178.02 19174 Nested Loop 31575.95 - 34450.02 36480 Sort 18370.49 - 18669.18 95553 Parallel Seq Sca message 0.00 - 6250.77 95581 Naterialize 13205.41 - 13479.01 101849 Seq Scan likes 0.00 - 1798.40 109440 Seq Scan knows 0.00 - 461.46 18188 Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 Hash 11241.44 - 11241.44 345 (count([_1.i_personid) >= 20) Hash Join 8853.56 - 11127.48 21623 ([_1.i_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 Hash 109440 109440 109440	WindowAgg		4703936.19 - 46759027.11	4476		((row_number() over (?)) <= 1)
✓ Gather Merge 4703936.19 - 46746181.38 57521 ✓ Incremental Sort 4702936.17 - 46746178.47 19174 ✓ Nested Loop 31575.95 - 646746178.02 19174 ✓ Merge Join 31575.95 - 33450.02 36480 ✓ Sort 18370.49 - 18669.18 95553 Parallel Seq Sca message 0.00 - 6250.77 95581 ✓ Materialize 13205.41 - 13752.61 101849 ✓ Sort 13205.41 - 13479.01 101849 ✓ Seq Scan likes 0.00 - 1798.40 109440 Seq Scan knows 0.00 - 481.46 18188 ✓ Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 ✓ Hash 11241.44 - 11241.44 345 (count([-1.]_personid) >= 20) ✓ Hash Join 8853.65 - 11127.48 21623 ([-1.]_messageid = m_1.m_messageid) ✓ Hash 1886 0.00 - 1798.40 109440 109440 109440 109440 109440 109440 109440 109440 109440 109440 109440 109440 109440 109440 109440 109440 109440 109440	WindowAgg		4703936.19 - 46759027.09	57521		
V Incremental Sort 4702936.17 - 46745178.47 19174 V Nested Loop 31575.95 - 33450.02 36480 V Merge Join 31575.95 - 33450.02 36480 V Sort 18370.49 - 18669.18 95553 Parallel Seq Sca message 0.00 - 6250.77 95581 V Materialize 13205.41 - 13752.61 101849 Seq Scan likes 0.00 - 1798.40 109440 Seq Scan knows 0.00 - 461.46 18188 V Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 Y Hash 11241.44 - 11241.44 345 (count([-1.]_personid) >= 20) Y Hash Join 8853.65 - 11127.48 21623 ([-1.]_messageid = m_1.m_messageid) Y Hash 8640.30 - 8640.30 179445	 Nested Loop 		4703936.19 - 46759027.08	57521		((k2.k_person1id is null)
Nested Loop 31575.95 - 46745178.02 19174 Nerge Join 31575.95 - 33450.02 36480 Sort 18370.49 - 18669.18 95553 Parallel Seq Sca message 0.00 - 6250.77 95581 Naterialize 13205.41 - 13752.61 101849 Sort 13205.41 - 13479.01 101849 Seq Scan likes 0.00 - 1798.40 109440 Seq Scan knows 0.00 - 461.46 18188 Adaterialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 Hash 11241.44 - 11241.44 345 (count([.1.L_personid] >= 20) Hash 1160.04 - 11241.44 345 (count([.1.L_personid] >= 20) Hash 8853.65 - 11127.48 21623 ([.1.L_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 Hash 8640.30 - 8640.30 17445	 Gather Merge 		4703936.19 - 46746181.38	57521		
✓ Merge Join 31575.95 - 33450.02 36480 ✓ Sort 18370.49 - 18669.18 95553 Parallel Seq Sca message 0.00 - 6250.77 95581 ✓ Materialize 13205.41 - 13752.61 101849 ✓ Sort 13205.41 - 13479.01 101849 Seq Scan likes 0.00 - 1798.40 109440 Seq Scan knows 0.00 - 481.46 18188 ✓ Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 ✓ Hash 11241.44 - 11241.44 345 (count([-1personid] >= 20) ✓ Hash Join 8853.56 - 11127.48 21623 ([-1messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 ✓ Hash 8640.30 - 8640.30 17445	 Incremental Sort 		4702936.17 - 46745178.47	19174		
Sort 18370.49 - 18669.18 95553 Parallel Seq Sca message 0.00 - 6250.77 95581 ✓ Materialize 13205.41 - 13752.61 101849 ✓ Sort 13205.41 - 13479.01 101849 Seq Scan likes 0.00 - 1798.40 109440 Seq Scan knows 0.00 - 461.46 18188 ✓ Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 ✓ Hash 11241.44 - 11241.44 345 ✓ Aggregate 11160.04 - 11241.44 345 (count([-1.]_personid) >= 20) ✓ Hash Join 8853.56 - 11127.48 21623 ([-1.]_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 ✓ Hash 8640.30 - 8640.30 17445	Nested Loop		31575.95 - 46745178.02	19174		
Parallel Seq Sca message 0.00 - 6250.77 95581 ✓ Materialize 13205.41 - 13479.01 101849 ✓ Sort 13205.41 - 13479.01 101849 Seq Scan likes 0.00 - 1798.40 109440 Seq Scan knows 0.00 - 461.46 18188 ✓ Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 ✓ Hash 11241.44 - 11241.44 345 ✓ Aggregate 11160.04 - 11241.44 345 (count([.1.L.personid] >= 20) ✓ Hash Join 8853.56 - 611127.48 21623 ([.1.L.messageid = m_1.m.messageid) Seq Scan likes 0.00 - 1798.40 109440 ✓ Hash 8640.30 - 8640.30 17445	∨ Merge Join		31575.95 - 33450.02	36480		
✓ Materialize 13205.41 - 13752.61 101849 ✓ Sort 13205.41 - 13479.01 101849 Seq Scan likes 0.00 - 1798.40 109440 Seq Scan knows 0.00 - 461.46 18188 ✓ Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 ✓ Hash 11241.44 - 11241.44 345 ✓ Aggregate 11160.04 - 11241.44 345 (count([-1.]_personid) >= 20) ✓ Hash Join 8853.56 - 11127.48 21623 ([-1.]_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 ✓ Hash 8640.30 - 8640.30 17445	∨ Sort		18370.49 - 18669.18	95553		
Sort 13205.41 - 13479.01 101849 Seq Scan likes 0.00 - 1798.40 109440 Seq Scan knows 0.00 - 461.46 18188 Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 Hash 11241.44 - 11241.44 345 Aggregate 11160.04 - 11241.44 345 (count(I_1.I_personid) >= 20) Hash Join 8853.56 - 11127.48 21623 (I_1.I_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 Hash 8640.30 - 8640.30 17445	Parallel Seq Sc	a message	0.00 - 6250.77	95581		
Seq Scan likes 0.00 - 1798.40 109440 Seq Scan knows 0.00 - 461.46 18188 Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 Hash 11241.44 - 11241.44 345 (count([_1.L_personid] >= 20) Hash Join 8853.56 - 11127.48 21623 ([_1.L_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 Hash 8640.30 - 8640.30 17445	✓ Materialize		13205.41 - 13752.61	101849		
Seq Scan knows 0.00 - 461.46 18188 Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 Hash 11241.44 - 11241.44 345 Aggregate 11160.04 - 11241.44 345 (count(I_1.I_personid) >= 20) Hash Join 8853.56 - 11127.48 21623 (I_1.I_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 Hash 8640.30 - 8640.30 17445	∨ Sort		13205.41 - 13479.01	101849		
✓ Materialize 0.00 - 602.19 28146 Seq Scan knows 0.00 - 461.46 28146 ✓ Hash 11241.44 - 11241.44 345 (count(I_1.I_personid) >= 20) ✓ Hash Join 8853.65 - 11127.48 21623 (I_1.I_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 ✓ Hash 8640.30 - 8640.30 17445	Seq Scan	likes	0.00 - 1798.40	109440		
Seq Scan knows 0.00 - 461.46 28146 Y Hash 11241.44 - 11241.44 345 Y Aggregate 11160.04 - 11241.44 345 (count([_1]_personid) >= 20) Y Hash Join 8853.56 - 11127.48 21623 ([_1.l_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 Y Hash 8640.30 - 8640.30 17445	Seq Scan	knows	0.00 - 461.46	18188		
Y Hash 11241.44 - 11241.44 345 ✓ Aggregate 11160.04 - 11241.44 345 (count([_1.l_personid] >= 20) ✓ Hash Join 8853.56 - 11127.48 21623 (l_1.l_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 ✓ Hash 8640.30 - 8640.30 17445	→ Materialize		0.00 - 602.19	28146		
∨ Aggregate 11160.04 - 11241.44 345 (count([_1. _personid) >= 20) ∨ Hash Join 8853.65 - 11127.48 21623 ([_1. _messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 ∨ Hash 8640.30 - 8640.30 17445	Seq Scan	knows	0.00 - 461.46	28146		
 ✓ Hash Join ✓ Hash Join 8853.56 - 11127.48 21623 (L1.I_messageid = m_1.m_messageid) Seq Scan likes 0.00 - 1798.40 109440 ✓ Hash 8640.30 - 8640.30 17445 	→ Hash		11241.44 - 11241.44	345		
Seq Scan likes 0.00 − 1798.40 109440 ∨ Hash 8640.30 − 8640.30 17445	Aggregate		11160.04 - 11241.44	345		(count(I_1.I_personid) >= 20)
∨ Hash 8640.30 - 8640.30 17445	Hash Join		8853.56 - 11127.48	21623		(I_1.I_messageid = m_1.m_messageid)
	Seq Scan	likes	0.00 - 1798.40	109440		
00	→ Hash		8640.30 - 8640.30	17445		
Seq Scan message 0.00 - 8640.30 1/445 (m_length > 100)	Seq Scan	message	0.00 - 8640.30	17445		(m_length > 100)
Seq Scan person 0.00 - 35.28 1528	Seq Scan	person	0.00 - 35.28	1528		

Reflection Task 3 - ChatGPT's Query Planning

```
| Operation Table Index/Join Condition Filter Rows (Estimate)
| Sort ORDER BY (foreign_likes / total_likes) DESC 1000
| Nested Loop Join 1000
| Nested Loop Join 5000
| Filtered Aggregate message m_length >= 100 2000
| | Join likes m.m_messageid = 1.1_messageid 50000
| | Filtered Join person Exclude friends of m_creatorid Based on knows subquery 10000
| Hash Join total_likes 1.mid = r.mid total_likes >= 20 and foreign_likes * 2 >= total_likes 5000
| Nested Loop Join likes 12 r.mid = 12.1_messageid and r.min_date = 12.1_creationdate
| Hash Join person p2 r2.1_personid = p2.p_personid 1000
```

- Badly formatted
- Not clear what it means
- Precision? Less Operators, other results than in DBeaver ...
- We need knowledge to understand this ⇒ not simple anymore ...

