Universidad Pontificia Bolivariana
Tópicos Avanzados en Bases de Datos
Luisa María Flórez Múnera ID 000449529 – Samuel Pérez Hurtado ID 000459067

Mejoras de desempeño – Consultas
Examen 2 - Sistema de gestión de recursos humanos y empleos

IMPLEMENTACIÓN DE LAS CONSULTAS

CONSULTA 1

Esta consulta inicialmente no tenía como mínimo 4 tablas, entonces se cambió un poco la estructura de la consulta.
Antes: ¿Cuál es el promedio de los salarios por nivel de experiencia y año de trabajo?

Ahora:
```sql
--¿Cuál es el promedio de los salarios por nivel de experiencia, año de trabajo y ubicación de la empresa?
select
    el.experience_level AS "Nivel de Experiencia",
    e.work_year AS "Año de Trabajo",
    cl.company_location AS "Ubicación de la Empresa",
    ROUND(AVG(e.salary_usd)::numeric, 2) AS "Promedio de Salario"
from inicial.employees e
join inicial.experience_levels el ON e.experience_level_id = el.id
join inicial.companies c ON e.company_id = c.id
join inicial.company_locations cl ON c.company_location_id = cl.id
group by el.experience_level, e.work_year, cl.company_location
order by el.experience_level, e.work_year, cl.company_location;
```

	"Nivel de Experiencia"	"Año de Trabajo"	"Ubicación de la Empresa"	"Promedio de Salario"
1	Entry-level	2020	AE	295989.13
2	Entry-level	2020	AR	302944.74
3	Entry-level	2020	AT	292728.97
4	Entry-level	2020	AU	292753.19
5	Entry-level	2020	BE	299673.58
6	Entry-level	2020	BG	349357.95
7	Entry-level	2020	BO	294580.2
8	Entry-level	2020	BR	301648.9
9	Entry-level	2020	CA	299613.84
10	Entry-level	2020	CH	291068.48
11	Entry-level	2020	CL	305734.91
12	Entry-level	2020	CN	300995.94
13	Entry-level	2020	CO	310142.32
14	Entry-level	2020	CZ	304320.09
15	Entry-level	2020	DE	289150.49
16	Entry-level	2020	DK	297553.04
17	Entry-level	2020	DZ	308147.84
18	Entry-level	2020	EE	298009.14
19	Entry-level	2020	ES	297163.1
20	Entry-level	2020	GB	287819.86

CONSULTA 2

```
¿Cuál es la cantidad de empleados en estado "On Leave" por ubicación y tamaño de Empresa?
select
    cl.company_location AS "Ubicación de la Empresa",
    cs.company_size AS "Tamaño de Empresa",
    COUNT(e.id) AS "Cantidad de Empleados en Estado On Leave"
from inicial.employees e
join inicial.companies c ON e.company_id = c.id
join inicial.company_locations cl ON c.company_location_id = cl.id
join inicial.company_sizes cs ON c.company_size_id = cs.id
join inicial.employment_statuses es ON e.employment_status_id = es.id
where es.employment_status = 'On Leave'
group by cl.company_location, cs.company_size
order by cl.company_location, cs.company_size;
```

	"Ubicación de la Empresa"	"Tamaño de Empresa"	"Cantidad de Empleados en Estado On Leave"
1	AE	L	252
2	AE	M	567
3	AR	L	265
4	AR	S	501
5	AT	L	265
6	AT	S	250
7	AU	M	263
8	AU	S	760
9	BE	M	237
10	BE	S	240
11	BG	S	251
12	BO	L	237
13	BO	S	552
14	BR	S	255
15	CA	L	252
16	CA	S	520
17	CH	L	776
18	CL	L	281
19	CL	M	280
20	CN	M	537
21	CO	M	729

CONSULTA 3

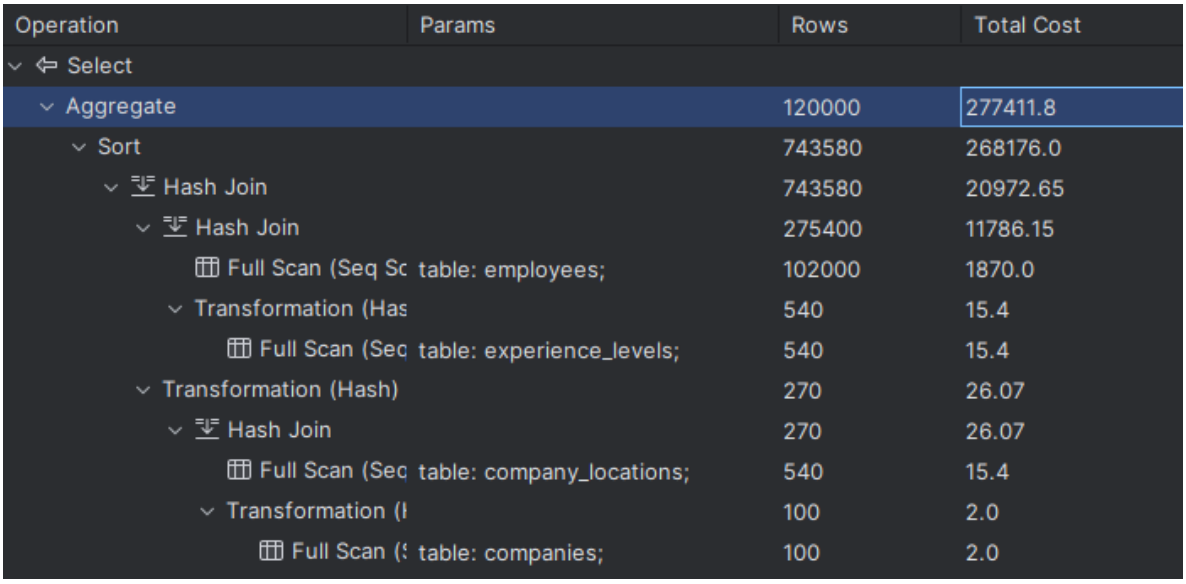
```
¿Cuál es el análisis de empleados remotos (100%) por título de trabajo y tamaño de empresa?
```

```sql
select
    jt.job_title AS "Título de Trabajo",
    cs.company_size AS "Tamaño de Empresa",
    COUNT(e.id) AS "Cantidad de Empleados Remotos"
from inicial.employees e
join inicial.companies c ON e.company_id = c.id
join inicial.company_sizes cs ON c.company_size_id = cs.id
join inicial.job_titles jt ON e.job_title_id = jt.id
join inicial.remote_ratios rr ON e.remote_ratio_id = rr.id
where rr.remote_ratio = '100'
group by jt.job_title, cs.company_size
order by jt.job_title, cs.company_size;
```

	"Título de Trabajo"	"Tamaño de Empresa"	"Cantidad de Empleados Remotos"
1	3D Computer Vision Researcher	L	226
2	3D Computer Vision Researcher	M	203
3	3D Computer Vision Researcher	S	272
4	AI Scientist	L	233
5	AI Scientist	M	180
6	AI Scientist	S	250
7	Analytics Engineer	L	257
8	Analytics Engineer	M	206
9	Analytics Engineer	S	221
10	Applied Data Scientist	L	233
11	Applied Data Scientist	M	181
12	Applied Data Scientist	S	225
13	Applied Machine Learning Scientist	L	243
14	Applied Machine Learning Scientist	M	200
15	Applied Machine Learning Scientist	S	246
16	BI Data Analyst	L	235
17	BI Data Analyst	M	191
18	BI Data Analyst	S	261
19	Big Data Architect	L	211
20	Big Data Architect	M	197
21	Big Data Architect	S	208

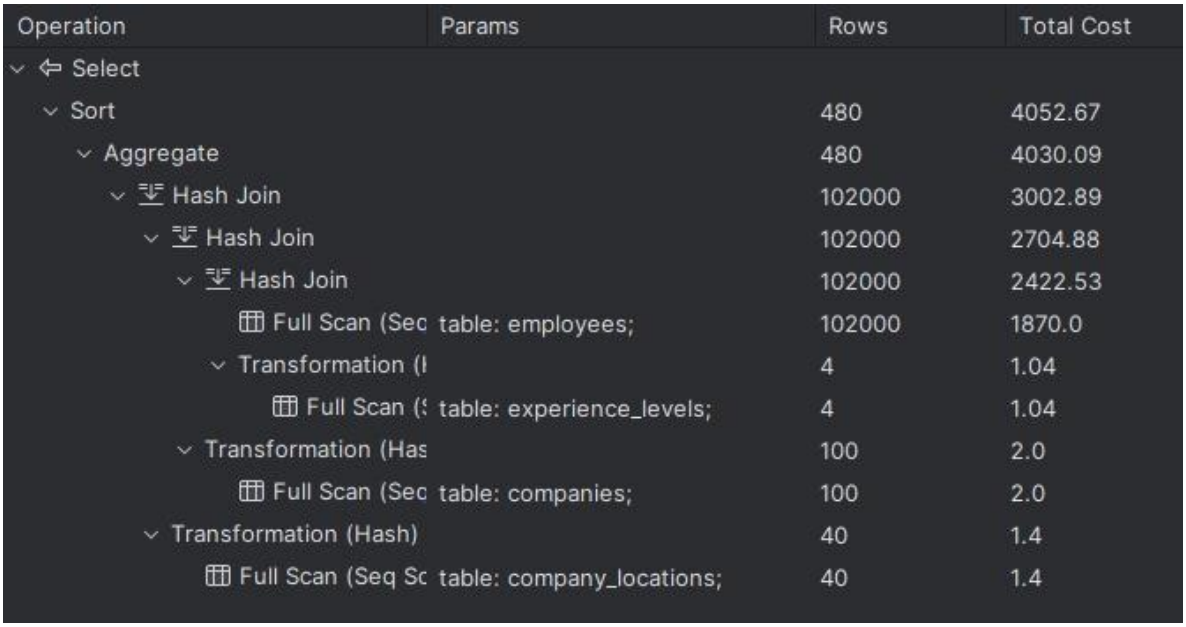
POSTGRES

CONSULTA 1
PLAN DE MEJORA EN EL ESQUEMA INICIAL



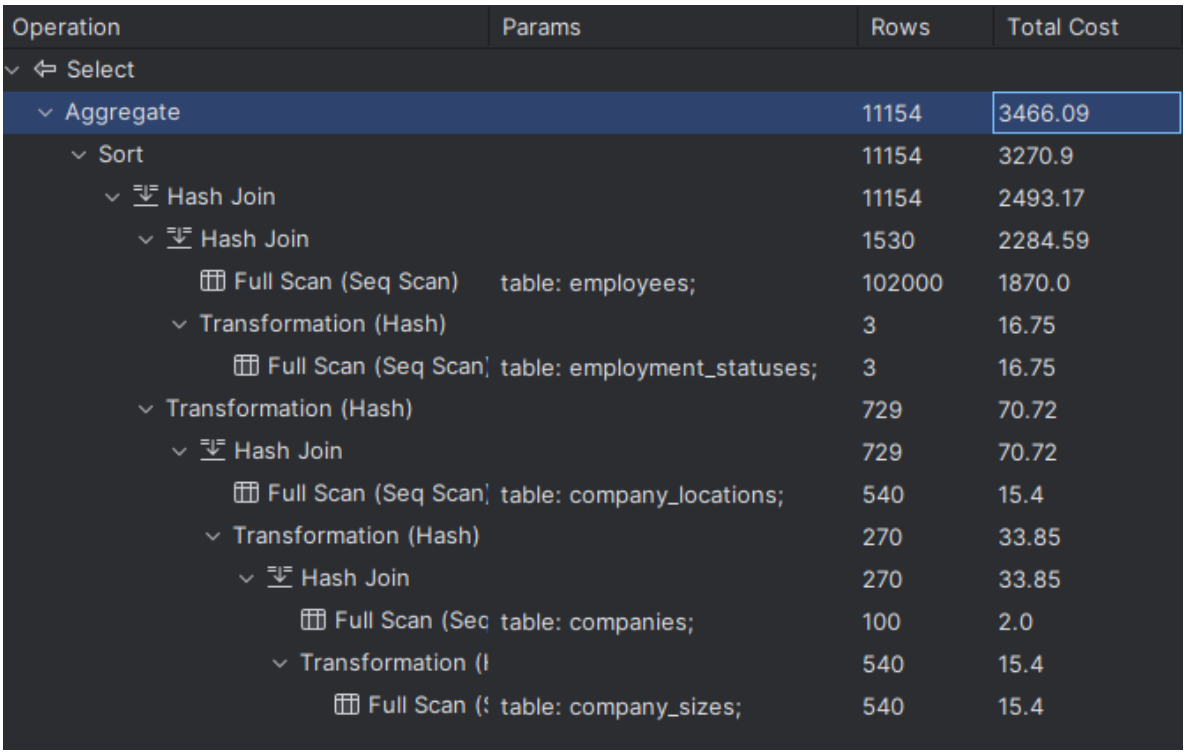
Operation	Params	Rows	Total Cost
Select			
Aggregate		120000	277411.8
Sort		743580	268176.0
Hash Join		743580	20972.65
Hash Join		275400	11786.15
Full Scan (Seq Sc	table: employees;	102000	1870.0
Transformation (Has		540	15.4
Full Scan (Sec	table: experience_levels;	540	15.4
Transformation (Hash)		270	26.07
Hash Join		270	26.07
Full Scan (Sec	table: company_locations;	540	15.4
Transformation (I		100	2.0
Full Scan (S	table: companies;	100	2.0

PLAN DE MEJORA EN EL ESQUEMA OPTIMIZADO



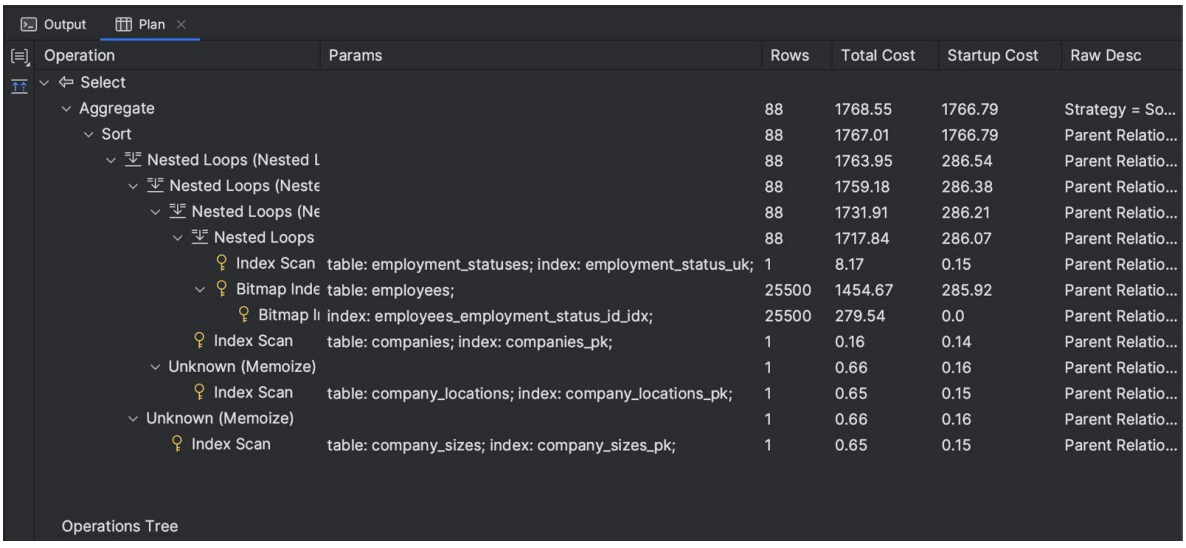
Operation	Params	Rows	Total Cost
Select			
Sort		480	4052.67
Aggregate		480	4030.09
Hash Join		102000	3002.89
Hash Join		102000	2704.88
Hash Join		102000	2422.53
Full Scan (Sec	table: employees;	102000	1870.0
Transformation (I		4	1.04
Full Scan (S	table: experience_levels;	4	1.04
Transformation (Has		100	2.0
Full Scan (Sec	table: companies;	100	2.0
Transformation (Hash)		40	1.4
Full Scan (Seq Sc	table: company_locations;	40	1.4

CONSULTA 2
PLAN DE MEJORA EN EL ESQUEMA INICIAL



Operation	Params	Rows	Total Cost
Select			
Aggregate		11154	3466.09
Sort		11154	3270.9
Hash Join		11154	2493.17
Hash Join		1530	2284.59
Full Scan (Seq Scan)	table: employees;	102000	1870.0
Transformation (Hash)		3	16.75
Full Scan (Seq Scan)	table: employment_statuses;	3	16.75
Transformation (Hash)		729	70.72
Hash Join		729	70.72
Full Scan (Seq Scan)	table: company_locations;	540	15.4
Transformation (Hash)		270	33.85
Hash Join		270	33.85
Full Scan (Seq	table: companies;	100	2.0
Transformation (H		540	15.4
Full Scan (S	table: company_sizes;	540	15.4

PLAN DE MEJORA EN EL ESQUEMA OPTIMIZADO



Output | Plan

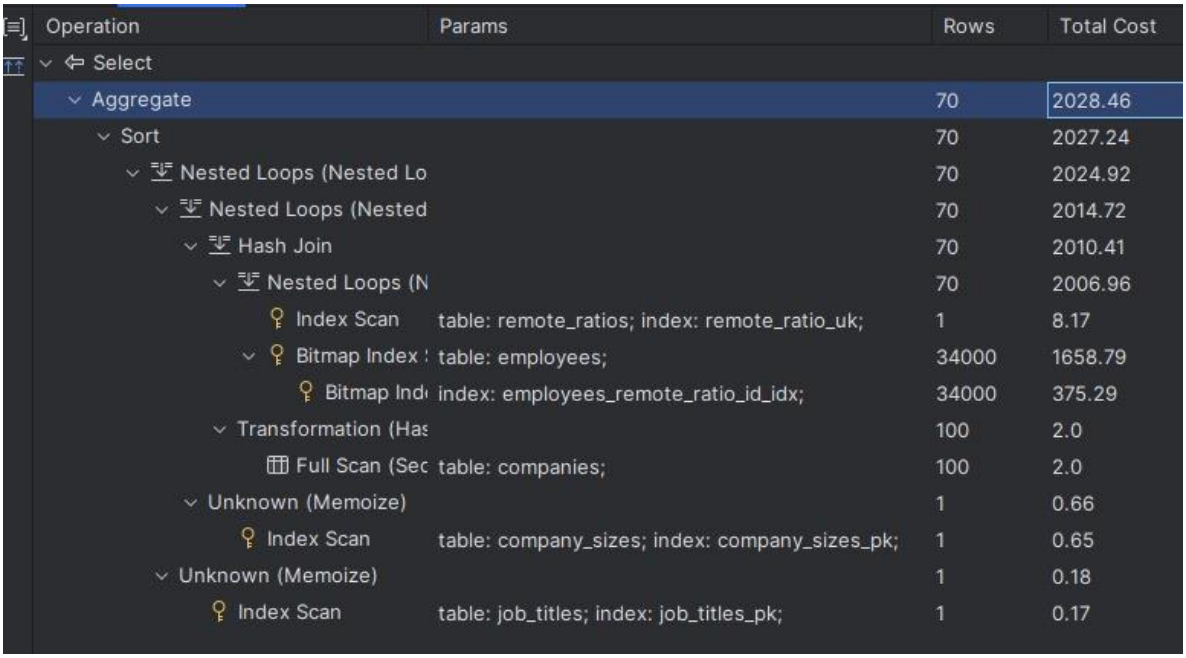
Operation	Params	Rows	Total Cost	Startup Cost	Raw Desc
Select					
Aggregate		88	1768.55	1766.79	Strategy = So...
Sort		88	1767.01	1766.79	Parent Relatio...
Nested Loops (Nested L		88	1763.95	286.54	Parent Relatio...
Nested Loops (Neste		88	1759.18	286.38	Parent Relatio...
Nested Loops (Ne		88	1731.91	286.21	Parent Relatio...
Nested Loops		88	1717.84	286.07	Parent Relatio...
Index Scan	table: employment_statuses; index: employment_status_uk;	1	8.17	0.15	Parent Relatio...
Bitmap Inde	table: employees;	25500	1454.67	285.92	Parent Relatio...
Bitmap I	index: employees_employment_status_id_idx;	25500	279.54	0.0	Parent Relatio...
Index Scan	table: companies; index: companies_pk;	1	0.16	0.14	Parent Relatio...
Unknown (Memoize)		1	0.66	0.16	Parent Relatio...
Index Scan	table: company_locations; index: company_locations_pk;	1	0.65	0.15	Parent Relatio...
Unknown (Memoize)		1	0.66	0.16	Parent Relatio...
Index Scan	table: company_sizes; index: company_sizes_pk;	1	0.65	0.15	Parent Relatio...

Operations Tree

CONSULTA 3
PLAN DE MEJORA EN EL ESQUEMA INICIAL

Operation	Params	Rows	Total Cost	Startup Cost
Select				
Aggregate		4131	2738.51	2655.89
Sort		4131	2666.21	2655.89
Hash Join		4131	2407.77	68.06
Hash Join		1530	2319.68	30.84
Hash Join		1530	2284.59	16.79
Full Scan (S	table: employees;	102000	1870.0	0.0
Transformation		3	16.75	16.75
Full Scan	table: remote_ratios;	3	16.75	0.0
Transformation (H		180	11.8	11.8
Full Scan (S	table: job_titles;	180	11.8	0.0
Transformation (Hash		270	33.85	33.85
Hash Join		270	33.85	22.15
Full Scan (S	table: companies;	100	2.0	0.0
Transformation		540	15.4	15.4
Full Scan	table: company_sizes;	540	15.4	0.0

PLAN DE MEJORA EN EL ESQUEMA OPTIMIZADO

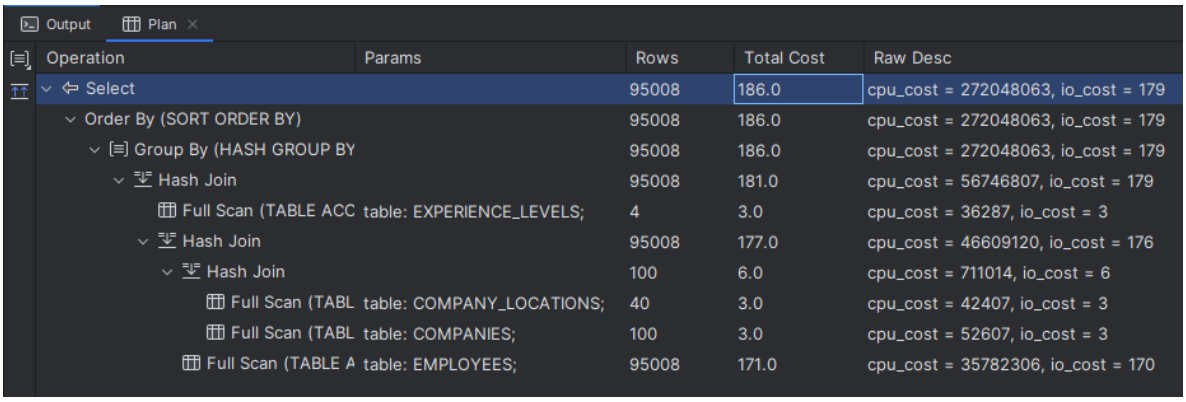


Operation	Params	Rows	Total Cost
Select			
Aggregate		70	2028.46
Sort		70	2027.24
Nested Loops (Nested Lo		70	2024.92
Nested Loops (Nested		70	2014.72
Hash Join		70	2010.41
Nested Loops (N		70	2006.96
Index Scan	table: remote_ratios; index: remote_ratio_uk;	1	8.17
Bitmap Index S	table: employees;	34000	1658.79
Bitmap Indi	index: employees_remote_ratio_id_idx;	34000	375.29
Transformation (Has		100	2.0
Full Scan (Sec	table: companies;	100	2.0
Unknown (Memoize)		1	0.66
Index Scan	table: company_sizes; index: company_sizes_pk;	1	0.65
Unknown (Memoize)		1	0.18
Index Scan	table: job_titles; index: job_titles_pk;	1	0.17

ORACLE

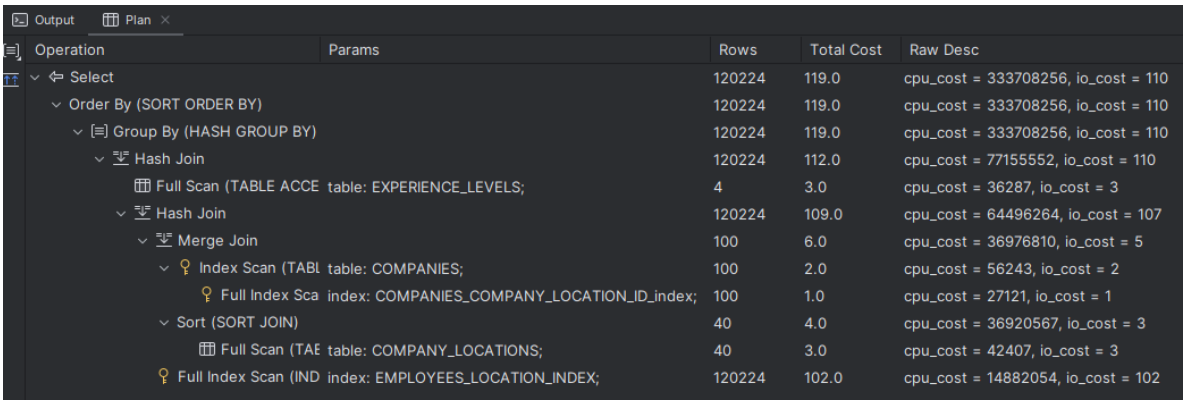
CONSULTA 1

PLAN DE MEJORAMIENTO EN EL ESQUEMA INICIAL



Operation	Params	Rows	Total Cost	Raw Desc
Select		95008	186.0	cpu_cost = 272048063, io_cost = 179
Order By (SORT ORDER BY)		95008	186.0	cpu_cost = 272048063, io_cost = 179
Group By (HASH GROUP BY		95008	186.0	cpu_cost = 272048063, io_cost = 179
Hash Join		95008	181.0	cpu_cost = 56746807, io_cost = 179
Full Scan (TABLE ACC	table: EXPERIENCE_LEVELS;	4	3.0	cpu_cost = 36287, io_cost = 3
Hash Join		95008	177.0	cpu_cost = 46609120, io_cost = 176
Hash Join		100	6.0	cpu_cost = 711014, io_cost = 6
Full Scan (TABL	table: COMPANY_LOCATIONS;	40	3.0	cpu_cost = 42407, io_cost = 3
Full Scan (TABL	table: COMPANIES;	100	3.0	cpu_cost = 52607, io_cost = 3
Full Scan (TABLE A	table: EMPLOYEES;	95008	171.0	cpu_cost = 35782306, io_cost = 170

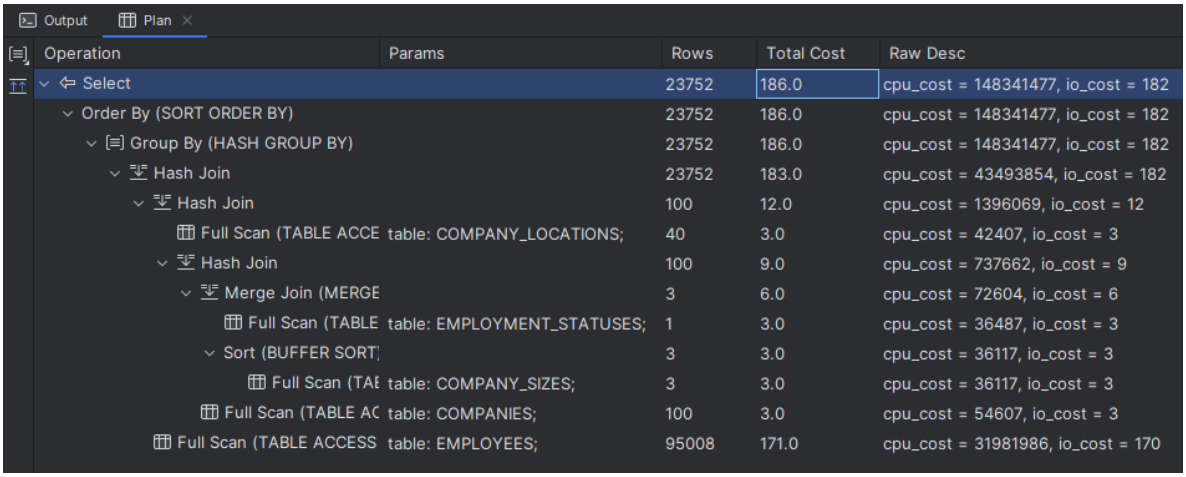
PLAN DE MEJORAMIENTO EN EL ESQUEMA OPTIMIZADO



Operation	Params	Rows	Total Cost	Raw Desc
Select		120224	119.0	cpu_cost = 333708256, io_cost = 110
Order By (SORT ORDER BY)		120224	119.0	cpu_cost = 333708256, io_cost = 110
Group By (HASH GROUP BY)		120224	119.0	cpu_cost = 333708256, io_cost = 110
Hash Join		120224	112.0	cpu_cost = 77155552, io_cost = 110
Full Scan (TABLE ACCE	table: EXPERIENCE_LEVELS;	4	3.0	cpu_cost = 36287, io_cost = 3
Hash Join		120224	109.0	cpu_cost = 64496264, io_cost = 107
Merge Join		100	6.0	cpu_cost = 36976810, io_cost = 5
Index Scan (TABL	table: COMPANIES;	100	2.0	cpu_cost = 56243, io_cost = 2
Full Index Sca	index: COMPANIES_COMPANY_LOCATION_ID_index;	100	1.0	cpu_cost = 27121, io_cost = 1
Sort (SORT JOIN)		40	4.0	cpu_cost = 36920567, io_cost = 3
Full Scan (TAE	table: COMPANY_LOCATIONS;	40	3.0	cpu_cost = 42407, io_cost = 3
Full Index Scan (IND	index: EMPLOYEES_LOCATION_INDEX;	120224	102.0	cpu_cost = 14882054, io_cost = 102

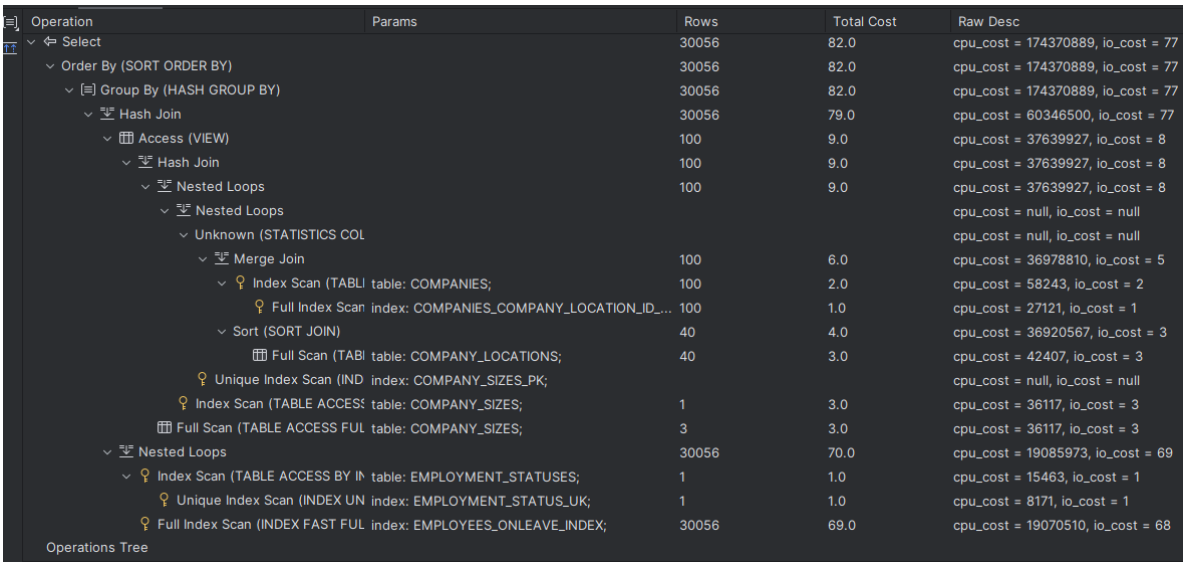
CONSULTA 2

PLAN DE MEJORAMIENTO EN EL ESQUEMA INICIAL



Operation	Params	Rows	Total Cost	Raw Desc
Select		23752	186.0	cpu_cost = 148341477, io_cost = 182
Order By (SORT ORDER BY)		23752	186.0	cpu_cost = 148341477, io_cost = 182
Group By (HASH GROUP BY)		23752	186.0	cpu_cost = 148341477, io_cost = 182
Hash Join		23752	183.0	cpu_cost = 43493854, io_cost = 182
Hash Join		100	12.0	cpu_cost = 1396069, io_cost = 12
Full Scan (TABLE ACCE	table: COMPANY_LOCATIONS;	40	3.0	cpu_cost = 42407, io_cost = 3
Hash Join		100	9.0	cpu_cost = 737662, io_cost = 9
Merge Join (MERGE		3	6.0	cpu_cost = 72604, io_cost = 6
Full Scan (TABLE	table: EMPLOYMENT_STATUSES;	1	3.0	cpu_cost = 36487, io_cost = 3
Sort (BUFFER SORT		3	3.0	cpu_cost = 36117, io_cost = 3
Full Scan (TAE	table: COMPANY_SIZES;	3	3.0	cpu_cost = 36117, io_cost = 3
Full Scan (TABLE AC	table: COMPANIES;	100	3.0	cpu_cost = 54607, io_cost = 3
Full Scan (TABLE ACCESS	table: EMPLOYEES;	95008	171.0	cpu_cost = 31981986, io_cost = 170

PLAN DE MEJORAMIENTO EN EL ESQUEMA OPTIMIZADO

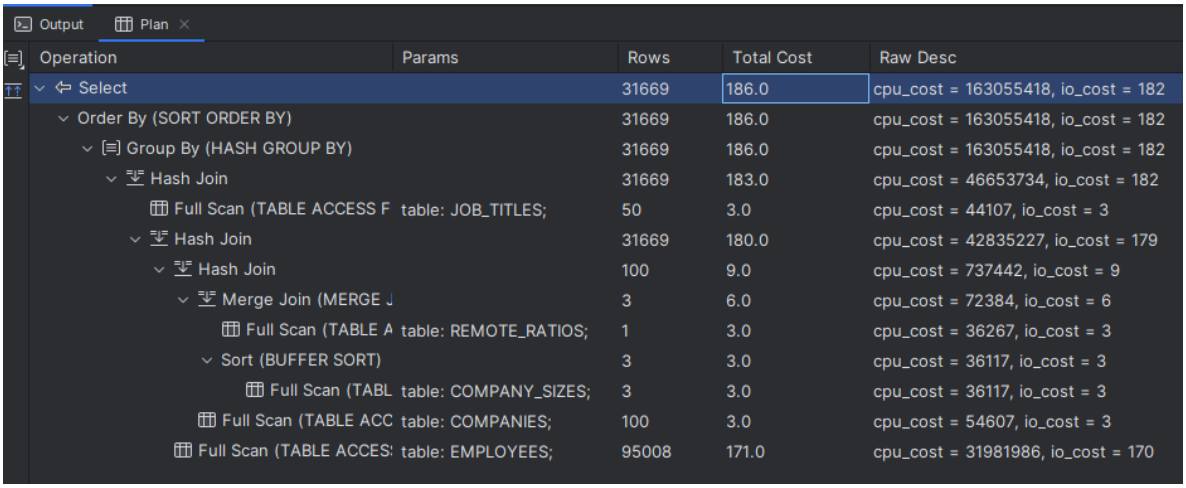


Operation	Params	Rows	Total Cost	Raw Desc
Select		30056	82.0	cpu_cost = 174370889, io_cost = 77
Order By (SORT ORDER BY)		30056	82.0	cpu_cost = 174370889, io_cost = 77
Group By (HASH GROUP BY)		30056	82.0	cpu_cost = 174370889, io_cost = 77
Hash Join		30056	79.0	cpu_cost = 60346500, io_cost = 77
Access (VIEW)		100	9.0	cpu_cost = 37639927, io_cost = 8
Hash Join		100	9.0	cpu_cost = 37639927, io_cost = 8
Nested Loops		100	9.0	cpu_cost = 37639927, io_cost = 8
Nested Loops				cpu_cost = null, io_cost = null
Unknown (STATISTICS COL				cpu_cost = null, io_cost = null
Merge Join		100	6.0	cpu_cost = 36978810, io_cost = 5
Index Scan (TABLI	table: COMPANIES;	100	2.0	cpu_cost = 58243, io_cost = 2
Full Index Scan	index: COMPANIES_COMPANY_LOCATION_ID_...	100	1.0	cpu_cost = 27121, io_cost = 1
Sort (SORT JOIN)		40	4.0	cpu_cost = 36920567, io_cost = 3
Full Scan (TABI	table: COMPANY_LOCATIONS;	40	3.0	cpu_cost = 42407, io_cost = 3
Unique Index Scan (IND	index: COMPANY_SIZES_PK;			cpu_cost = null, io_cost = null
Index Scan (TABLE ACCESS	table: COMPANY_SIZES;	1	3.0	cpu_cost = 36117, io_cost = 3
Full Scan (TABLE ACCESS FUL	table: COMPANY_SIZES;	3	3.0	cpu_cost = 36117, io_cost = 3
Nested Loops		30056	70.0	cpu_cost = 19085973, io_cost = 69
Index Scan (TABLE ACCESS BY IN	table: EMPLOYMENT_STATUSES;	1	1.0	cpu_cost = 15463, io_cost = 1
Unique Index Scan (INDEX UN	index: EMPLOYMENT_STATUS_UK;	1	1.0	cpu_cost = 8171, io_cost = 1
Full Index Scan (INDEX FAST FUL	index: EMPLOYEES_ONLEAVE_INDEX;	30056	69.0	cpu_cost = 19070510, io_cost = 68

Operations Tree

CONSULTA 3

PLAN DE MEJORAMIENTO EN EL ESQUEMA INICIAL



Output | Plan

Operation	Params	Rows	Total Cost	Raw Desc
Select		31669	186.0	cpu_cost = 163055418, io_cost = 182
Order By (SORT ORDER BY)		31669	186.0	cpu_cost = 163055418, io_cost = 182
Group By (HASH GROUP BY)		31669	186.0	cpu_cost = 163055418, io_cost = 182
Hash Join		31669	183.0	cpu_cost = 46653734, io_cost = 182
Full Scan (TABLE ACCESS F	table: JOB_TITLES;	50	3.0	cpu_cost = 44107, io_cost = 3
Hash Join		31669	180.0	cpu_cost = 42835227, io_cost = 179
Hash Join		100	9.0	cpu_cost = 737442, io_cost = 9
Merge Join (MERGE J		3	6.0	cpu_cost = 72384, io_cost = 6
Full Scan (TABLE A	table: REMOTE_RATIOS;	1	3.0	cpu_cost = 36267, io_cost = 3
Sort (BUFFER SORT)		3	3.0	cpu_cost = 36117, io_cost = 3
Full Scan (TABL	table: COMPANY_SIZES;	3	3.0	cpu_cost = 36117, io_cost = 3
Full Scan (TABLE ACC	table: COMPANIES;	100	3.0	cpu_cost = 54607, io_cost = 3
Full Scan (TABLE ACCES	table: EMPLOYEES;	95008	171.0	cpu_cost = 31981986, io_cost = 170

PLAN DE MEJORAMIENTO EN EL ESQUEMA OPTIMIZADO

Output | Plan

Operation	Params	Rows	Total Cost	Raw Desc
Select		40075	93.0	cpu_cost = 158327691, io_cost = 89
Order By (SORT ORDER BY)		40075	93.0	cpu_cost = 158327691, io_cost = 89
Group By (HASH GROUP BY)		40075	93.0	cpu_cost = 158327691, io_cost = 89
Hash Join		40075	90.0	cpu_cost = 29374976, io_cost = 89
Access (VIEW)		100	6.0	cpu_cost = 701174, io_cost = 6
Hash Join		100	6.0	cpu_cost = 701174, io_cost = 6
Full Scan (TABLE ACCES	table: COMPANY_SIZES;	3	3.0	cpu_cost = 36117, io_cost = 3
Full Scan (TABLE ACCES	table: COMPANIES;	100	3.0	cpu_cost = 54607, io_cost = 3
Hash Join		40075	84.0	cpu_cost = 24051302, io_cost = 83
Full Scan (TABLE ACCESS I	table: JOB_TITLES;	50	3.0	cpu_cost = 44107, io_cost = 3
Nested Loops		40075	81.0	cpu_cost = 19392195, io_cost = 80
Index Scan (TABLE ACC	table: REMOTE_RATIOS;	1	1.0	cpu_cost = 15463, io_cost = 1
Unique Index Scan (I	index: REMOTE_RATIO_UK;	1	1.0	cpu_cost = 8171, io_cost = 1
Full Index Scan (INDEX I	index: EMPLOYEES_REMOTE_INDEX;	40075	80.0	cpu_cost = 19376732, io_cost = 79