

Proyecto fase # 2: Implementación de la solución BI

Elaborado por Nelson Rivas

abril 2020

Habiendo definido en la fase I el modelo dimensional del *Datawarehouse*, en esta fase se realizará la implementación de la solución BI. Esto involucraría:

- Modelo relacional de la BD transaccional.
- Datawarehouse (almacén de datos).
- Procesos de extracción, transformación y carga (ETL's).
- *Dashboard*.

Las herramientas que se utilizaron son expuestas a continuación:

- **MySQL** (versión 8) como Sistema Manejador de Base de Datos.
- **Pentaho Data Integration** (PDI 8.3) community version, para los procesos ETL's.
- **Tableau Desktop** (versión 2019.4), para el *look and feel* (visualización).

En lo que sigue se dejará en evidencia la implementación.

Creación de tablas de la BD transaccional:

id_cat	nomb_cat
1	Children Bicycles
2	Comfort Bicycles
3	Cruisers Bicycles
4	Cydocross Bicycles
5	Electric Bikes
6	Mountain Bikes
7	Road Bikes
NULL	NULL

Fig. 1: Tabla categoria y su data.

id_marca	nomb_marca
1	Electra
2	Haro
3	Heller
4	Pure Cycles
5	Ritchey
6	Strider
7	Sun Bicycles
8	Surly
9	Trek
NULL	NULL

Fig. 2: Tabla marca

id_est	nomb_est	siglas
1	Anzoategui	ANZ
2	Aragua	ARG
3	Distrito Capital	DTC
4	Merida	MER
5	Zulia	ZLA
NULL	NULL	NULL

Fig. 3: Tabla estado

id_ciud	nomb_ciud	siglas	id_est
1	Barcelona	BRC	1
2	Caracas	CRC	3
3	Maracay	MCY	2
4	Maracaibo	MCB	5
5	Merida	MER	4
NULL	NULL	NULL	NULL

Fig. 4: Tabla ciudad

id_sub	nomb_subc	id_cat
10	SUBC Children Bicycles	1
20	SUBC Comfort Bicycles	2
30	SUBC Cruisers Bicycles	3
40	SUBC Cyclocross Bicycles	4
50	SUBC Electric Bikes	5
60	SUBC Mountain Bikes	6
70	SUBC Road Bikes	7
NULL	NULL	NULL

Fig. 5: Tabla subcategoria

id_prod	nomb_prod	id_marca	pvp	id_sub
1	Trek 820 - 2016	9	379.99	70
2	Ritchey Timberwolf Frameset - 2016	5	749.99	70
3	Surly Wednesday Frameset - 2016	8	999.99	50
4	Trek Fuel EX 8 29 - 2016	9	2899.99	50
5	Heller Shagamaw Frame - 2016	3	1320.99	50
6	Surly Ice Cream Truck Frameset - 2016	8	469.99	50
7	Trek Slash 8 27.5 - 2016	9	3999.99	50
8	Trek Remedy 29 Carbon Frameset - 2016	9	1799.99	40
9	Trek Conduit+ - 2016	9	2999.99	40
10	Surly Straggler - 2016	8	1549	10
11	Surly Straggler 650b - 2016	8	1680.99	10
12	Electra Townie Original 21D - 2016	1	549.99	20
13	Electra Cruiser 1 (24-Inch) - 2016	1	269.99	10

Fig. 6: Tabla producto

id_tienda	nomb_tienda	rif	telf	fax	capacidad_prod	id_ciudad
1	Santa Cruz Bikes	G20007749	(831) 476-4321	(024) 476-4321	10000	4
2	Baldwin Bikes	J41253680	(516) 379-8888	(023) 379-8888	15000	3
3	Rowlett Bikes	J30776801	(972) 530-5555	(024) 530-5555	6000	2
4	Start Sport	J9572525	(621) 432-8830	(021) 432-8830	4000	1
5	Tamanaco Sport	J40112923	(958) 305-1298	(025) 305-1298	30000	5
NULL	NULL	NULL	NULL	NULL	NULL	NULL

Fig. 7: Tabla tienda

id_alm	nomb_alm	rif	telf	dimensiones	capacidad_prod	id_ciudad
1	ALM Santa Cruz Bikes	J31183476	(631) 254-8497	40mts2	12000	4
2	ALM Baldwin Bikes	J08635241	(845) 707-6088	42mts2	12100	3
3	ALM Rowlett Bikes	J30716727	(507) 186-2004	50mts2	12000	3
4	ALM Santa Baldwin Bikes	J40423168	NULL		20000	2
5	ALM Santa Rowlett Bikes	J29644397		28mts2	2550	4
6	ALM Cruz Baldwin Bikes	J40832284	NULL	35mts2	3000	1
7	ALM Cruz Rowlett Bikes	J40460391			6900	3
8	ALM Baldwin Rowlett Bikes	J19306404	(408) 807-6157	50mts2	10000	5
9	ALM Tam Start Sport	J40423183	(621) 886-2496	60mts2	15000	5
10	ALM CORPORACION DEPORTIVA HORIZONTE	J40439183	(631) 807-2088	60mts2	15000	5
NULL	NULL	NULL	NULL	NULL	NULL	NULL

Fig. 8: Tabla almacen

id_prov	nombre	rif	telf	id_ciudad
1	Debra Burks	J40793867	NULL	1
2	Kasha Todd	J31020432	NULL	2
3	Tameka Fisher	J09510587	NULL	3
4	Daryl Spence	J30825425	NULL	4
5	Charolette Rice	J31412867	(916) 381-6003	5
6	Lyndsey Bean	J31501632	NULL	5
7	Latasha Hays	J29450221	(716) 986-3359	4
8	Jacqueline Duncan	J29677169	NULL	3
9	Genoveva Baldwin	J31183476	NULL	2
10	Pamela Newman	J31299455	NULL	1
11	Deshawn Mendoza	J30558369	NULL	2
12	Robby Sykes	J31217572	(516) 583-7761	3
13	Lashawn Ortiz	J40003164	NULL	4

Fig. 9: Tabla proveedor

fecha	id_prod	id_tienda	cant_vend	cant_exist	nopa	nmrp
2010-01-01	9	2	17	23	34	25
2010-01-02	10	2	13	10	35	99
2010-01-03	11	2	21	27	12	20
2010-01-10	1	2	14	34	50	10
2010-01-11	2	2	16	20	23	20
2010-01-12	3	2	28	24	29	10
2010-01-13	4	2	2	50	56	100
2010-01-24	5	2	1	99	27	23
2010-01-29	6	2	11	35	30	50
2010-01-30	7	2	8	87	30	40
2010-02-04	1	3	14	6	15	25
2010-02-05	2	3	24	20	20	20
2010-02-06	3	3	0	100	100	100

Fig. 10: Tabla relación tiene

fecha	id_alm	id_prod	cant_desp	cant_exist	nopal	nmrs
2010-01-03	1	41	2	123	234	56
2010-01-04	1	42	18	23	45	87
2010-01-05	3	43	17	34	43	35
2010-01-06	5	44	11	10	54	98
2010-01-10	1	45	11	23	6	6
2010-01-14	1	47	1	43	8	2
2010-01-14	6	46	15	23	87	9
2010-01-20	8	48	1	1	56	100
2010-01-22	9	49	22	34	54	89
2010-01-26	9	40	19	54	43	98
2010-02-06	1	41	30	59	37	32
2010-02-07	1	42	21	12	100	76
2010-02-08	1	44	30	12	15	234

Fig. 11: Tabla relación almacena

fecha_rec	fecha_env	id_prov	id_alm	id_prod	cant	costo_prod	costo_env	costo_total
2010-01-01	2010-01-03	12	2	35	45	832.99	37484.6	31224300
2010-01-01	2010-01-03	29	1	49	4	3499.99	14000	48999700
2010-01-02	2010-01-03	23	2	42	4	2299.99	9199.96	21159800
2010-01-03	2010-01-05	17	1	37	34	379.99	12919.7	4909340
2010-01-13	2010-01-16	24	2	38	5	549.99	2749.95	1512440
2010-01-14	2010-01-15	24	2	44	20	539.99	10782	5822160
2010-01-16	2010-01-17	24	2	44	10	539.99	5390.99	2911080
2010-01-24	2010-01-25	20	2	48	23	1499.99	34499.7	51749200
2010-01-25	2010-01-26	22	2	35	10	832.99	8320.99	6938800
2010-01-25	2010-01-28	10	1	45	4	869.99	3479.96	3027530
2010-02-05	2010-02-07	26	2	35	5	832.99	4164.95	3496360
2010-02-05	2010-02-07	26	8	35	5	832.99	4164.95	3496360
2010-02-06	2010-02-09	11	2	36	8	832.99	6663.92	5597630

Fig. 12: Tabla relación provee

fecha_rec	fecha_desc	id_alm	id_tienda	id_prod
2010-02-01	2010-02-02	1	2	45
2010-02-03	2010-02-04	1	1	30
2010-02-03	2010-02-04	1	2	45
2010-02-04	2010-02-07	1	2	45
2010-06-06	2010-06-08	1	3	39
2010-06-17	2010-06-18	1	4	38
2010-06-17	2010-06-19	1	2	38
2010-06-28	2010-06-29	1	2	35
2010-06-29	2010-06-30	1	3	34
2010-07-01	2010-07-02	1	2	45
2010-07-03	2010-07-04	1	1	30
2010-11-06	2010-11-08	1	1	39
2010-11-17	2010-11-19	1	2	37

Fig. 13: Tabla relación abastece

Encontrándose ya llena la base de datos fuente, se mostrará a continuación el *Datawarehouse*:

sk_dim_prod	id_cat	desc_cat	id_subc	desc_subc	id_marca	nomb_marca	id_prod	desc_prod	pvp	codigo_concatenado
1	1	Children Bicycles	10	SUBC Children Bicycles	8	Surly	10	Surly Straggler - 2016	1549	110810
2	1	Children Bicycles	10	SUBC Children Bicycles	8	Surly	11	Surly Straggler 650b - 2016	1680.99	110811
3	1	Children Bicycles	10	SUBC Children Bicycles	1	Electra	13	Electra Cruiser 1 (24-Inch) - 2016	269.99	110113
4	1	Children Bicycles	10	SUBC Children Bicycles	8	Surly	33	Surly Wednesday Frameset - 2017	469.99	110833
5	1	Children Bicycles	10	SUBC Children Bicycles	9	Trek	34	Trek Session DH 27.5 Carbon Frameset - 2017	469.99	110934
6	1	Children Bicycles	10	SUBC Children Bicycles	7	Sun Bicycles	35	Sun Bicycles Spider 3i - 2017	832.99	110735
7	1	Children Bicycles	10	SUBC Children Bicycles	8	Surly	36	Surly Troll Frameset - 2017	832.99	110836
8	1	Children Bicycles	10	SUBC Children Bicycles	2	Haro	37	Haro Flightline One ST - 2017	379.99	110237
9	1	Children Bicycles	10	SUBC Children Bicycles	2	Haro	38	Haro Flightline Two 26 Plus - 2017	549.99	110238
10	1	Children Bicycles	10	SUBC Children Bicycles	9	Trek	39	Trek Stache 5 - 2017	1499.99	110939
11	1	Children Bicycles	10	SUBC Children Bicycles	9	Trek	40	Trek Fuel EX 9.8 29 - 2017	4999.99	110940
12	2	Comfort Bicycles	20	SUBC Comfort Bicycles	1	Electra	12	Electra Townie Original 21D - 2016	549.99	220112
13	2	Comfort Bicycles	20	SUBC Comfort Bicycles	1	Electra	22	Electra Girl's Hawaii 1 (16-inch) - 2015/2016	269.99	220122
14	2	Comfort Bicycles	20	SUBC Comfort Bicycles	1	Electra	23	Electra Girl's Hawaii 1 (20-inch) - 2015/2016	299.99	220123

Fig. 14: Dimensión producto.

sk_dim_local	id_estado	desc_estado	id_ciudad	desc_ciudad	codigo_concatenado
1	1	Anzoategui	1	Barcelona	11
2	3	Distrito Capital	2	Caracas	32
3	2	Aragua	3	Maracay	23
4	5	Zulia	4	Maracaibo	54
5	4	Merida	5	Merida	45
NULL	NULL	NULL	NULL	NULL	NULL

Fig. 15: Dimensión localidad.

sk_dim_tienda	id_tienda	desc_tienda	rif	telf	fax	capacidad
1	1	Santa Cruz Bikes	G20007749	(831) 476-4321	(024) 476-4321	10000
2	2	Baldwin Bikes	J41253680	(516) 379-8888	(023) 379-8888	15000
3	3	Rowlett Bikes	J30776801	(972) 530-5555	(024) 530-5555	6000
4	4	Start Sport	J9572525	(621) 432-8830	(021) 432-8830	4000
5	5	Tamanaco Sport	J40112923	(958) 305-1298	(025) 305-1298	30000
NULL	NULL	NULL	NULL	NULL	NULL	NULL

Fig. 16: Dimensión tienda.

sk_dim_almacen	id_almacen	desc_almacen	rif	telf	dimensiones	capacidad
1	1	ALM Santa Cruz Bikes	J31183476	(631) 254-8497	40mts2	12000
2	2	ALM Baldwin Bikes	J08635241	(845) 707-6088	42mts2	12100
3	3	ALM Rowlett Bikes	J30716727	(507) 186-2004	50mts2	12000
4	4	ALM Santa Baldwin Bikes	J40423168	NULL	NULL	20000
5	5	ALM Santa Rowlett Bikes	J29644397	NULL	28mts2	2550
6	6	ALM Cruz Baldwin Bikes	J40832284	NULL	35mts2	3000
7	7	ALM Cruz Rowlett Bikes	J40460391	NULL	NULL	6900
8	8	ALM Baldwin Rowlett Bikes	J19306404	(408) 807-6157	50mts2	10000
9	9	ALM Tam Start Sport	J40423183	(621) 886-2496	60mts2	15000
10	10	ALM CORPORACION DEPORTIVA HORIZONTE	J40439183	(631) 807-2088	60mts2	15000
NULL	NULL	NULL	NULL	NULL	NULL	NULL

Fig. 17: Dimensión almacen.

sk_dim_prov	id_prov	desc_prov	rif	telf
1	1	Debra Burks	J40793867	NULL
2	2	Kasha Todd	J31020432	NULL
3	3	Tameka Fisher	J09510587	NULL
4	4	Daryl Spence	J30825425	NULL
5	5	Charolette Rice	J31412867	(916) 381-6003
6	6	Lyndsey Bean	J31501632	NULL
7	7	Latasha Hays	J29450221	(716) 986-3359
8	8	Jacqueline Duncan	J29677169	NULL
9	9	Genoveva Baldwin	J31183476	NULL
10	10	Pamelia Newman	J31299455	NULL
11	11	Deshawn Mendoza	J30558369	NULL
12	12	Robby Sykes	J31217572	(516) 583-7761
13	13	Lashawn Ortiz	J40003164	NULL

Fig. 18: Dimensión proveedor.

sk_dim_tiempo	id_anio	desc_anio	id_semestre	desc_semestre	id_trimestre	desc_trimestre	id_mes	desc_mes	id_dia	desc_dia	fecha
20100101	2010	NULL	NULL	1	NULL	1	1	Enero	1	NULL	2010-01-01
20100102	2010	NULL	NULL	1	NULL	1	1	Enero	2	NULL	2010-01-02
20100103	2010	NULL	NULL	1	NULL	1	1	Enero	3	NULL	2010-01-03
20100104	2010	NULL	NULL	1	NULL	1	1	Enero	4	NULL	2010-01-04
20100105	2010	NULL	NULL	1	NULL	1	1	Enero	5	NULL	2010-01-05
20100106	2010	NULL	NULL	1	NULL	1	1	Enero	6	NULL	2010-01-06
20100107	2010	NULL	NULL	1	NULL	1	1	Enero	7	NULL	2010-01-07
20100108	2010	NULL	NULL	1	NULL	1	1	Enero	8	NULL	2010-01-08
20100109	2010	NULL	NULL	1	NULL	1	1	Enero	9	NULL	2010-01-09
20100110	2010	NULL	NULL	1	NULL	1	1	Enero	10	NULL	2010-01-10
20100111	2010	NULL	NULL	1	NULL	1	1	Enero	11	NULL	2010-01-11
20100112	2010	NULL	NULL	1	NULL	1	1	Enero	12	NULL	2010-01-12
20100113	2010	NULL	NULL	1	NULL	1	1	Enero	13	NULL	2010-01-13

Fig. 19: Dimensión tiempo.

sk_dim_almacen	sk_dim_local	sk_dim_prov	sk_dim_prod	fecha_original	nmrs	nopal	fecha_pedido	fecha_recibido	cant_desp	cant_exist	pvp_prod	costo_productos	costo_envio	costo_total	inventario_dd
1	4	10	26	20100110	6	6	20100125	20100128	11	23	869.99	869.99	3479.96	3027530	1
1	4	10	26	20100210	21	34	20100125	20100128	10	32	869.99	869.99	3479.96	3027530	2
1	4	2	11	20100222	2	3	20100219	20100225	3	2	4999.99	4999.99	24999.9	125000000	3
10	5	8	22	20100306	32	37	20100208	20100211	30	59	1469.99	1469.99	14700	21608800	4
1	4	10	26	20100308	89	34	20100125	20100128	4	34	869.99	869.99	3479.96	3027530	5
2	3	23	23	20100402	76	100	20100102	20100103	21	12	2299.99	2299.99	9199.96	21159800	6
1	4	2	8	20100518	4	65	20100209	20100212	16	98	379.99	832.99	4164.95	3469360	7
1	4	17	8	20100518	4	65	20100103	20100105	16	98	379.99	379.99	12919.7	4909340	8
2	3	8	13	20100607	2	50	20100412	20100413	1	78	269.99	269.99	7559.72	2041050	9
3	3	25	46	20100701	20	20	20110103	20110104	8	20	379.99	379.99	3799.99	1443960	10
1	4	13	47	20100702	60	45	20100322	20100323	0	30	749.99	749.99	11249.8	8437280	11
2	3	7	33	20100703	22	10	20100310	20100312	22	12	999.99	999.99	5999.94	5999880	12
1	4	5	34	20100709	21	15	20100224	20100225	6	6	2899.99	2899.99	57999.8	168199000	13

Fig. 20: Fact almacen.

fecha_original	sk_dim_tend	sk_dim_local	sk_dim_almacen	sk_dim_prod	fecha_pedido	fecha_recibido	pvp_prod	cant_exist	cant_vendida	nmrp	nopa	inventario_dd
20100101	2	3	4	28	20100309	20100311	2999.99	23	17	25	34	2010010123428
20100101	2	3	5	28	20100310	20100311	2999.99	23	17	25	34	2010010123528
20100101	2	3	7	28	20110620	20110621	2999.99	23	17	25	34	2010010123728
20100101	2	3	10	28	20111125	20111127	2999.99	23	17	25	34	20100101231028
20100102	2	3	1	1	20110923	20110924	1549	10	13	99	35	201001022311
20100102	2	3	7	1	20110717	20100719	1549	10	13	99	35	201001022371
20100110	2	3	5	46	20100204	20100206	379.99	34	14	10	50	2010011023546
20100110	2	3	9	46	20100308	20100309	379.99	34	14	10	50	2010011023946
20100110	2	3	10	46	20101110	20100111	379.99	34	14	10	50	20100110231046
20100111	2	3	2	47	20100414	20100415	749.99	20	16	20	23	2010011123247
20100111	2	3	9	47	20100306	20100309	749.99	20	16	20	23	2010011123947
20100111	2	3	10	47	20100916	20100916	749.99	20	16	20	23	20100111231047
20100113	2	3	5	34	20111202	20111205	2899.99	50	2	100	56	2010011323534

Fig. 21: Fact tienda.

Ahora prosigue las transformaciones y el *job*:

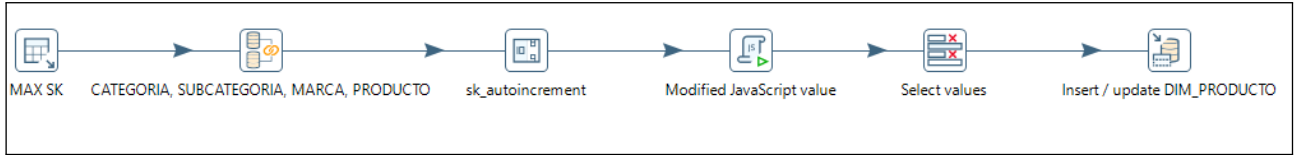


Fig. 22: Transformación de la dimensión producto.

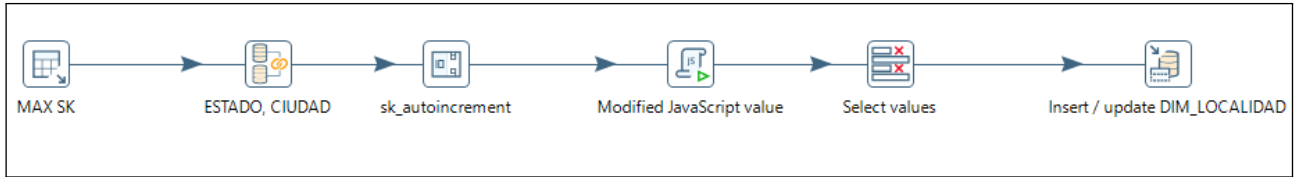


Fig. 23: Transformación de la dimensión localidad.

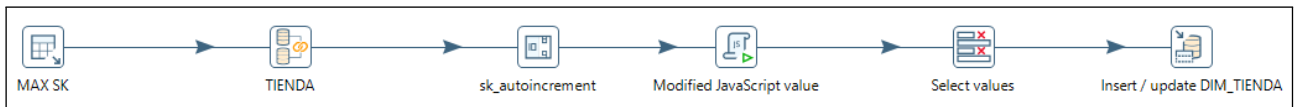


Fig. 24: Transformación de la dimensión tienda.

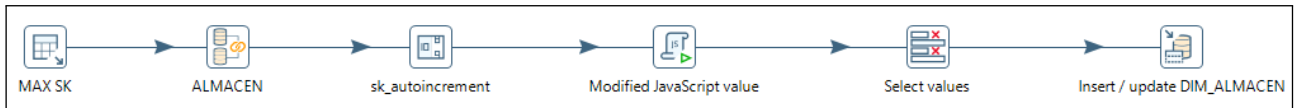


Fig. 25: Transformación de la dimensión almacen.

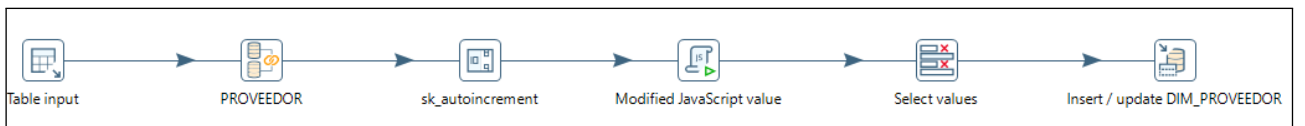


Fig. 26: Transformación de la dimensión proveedor.

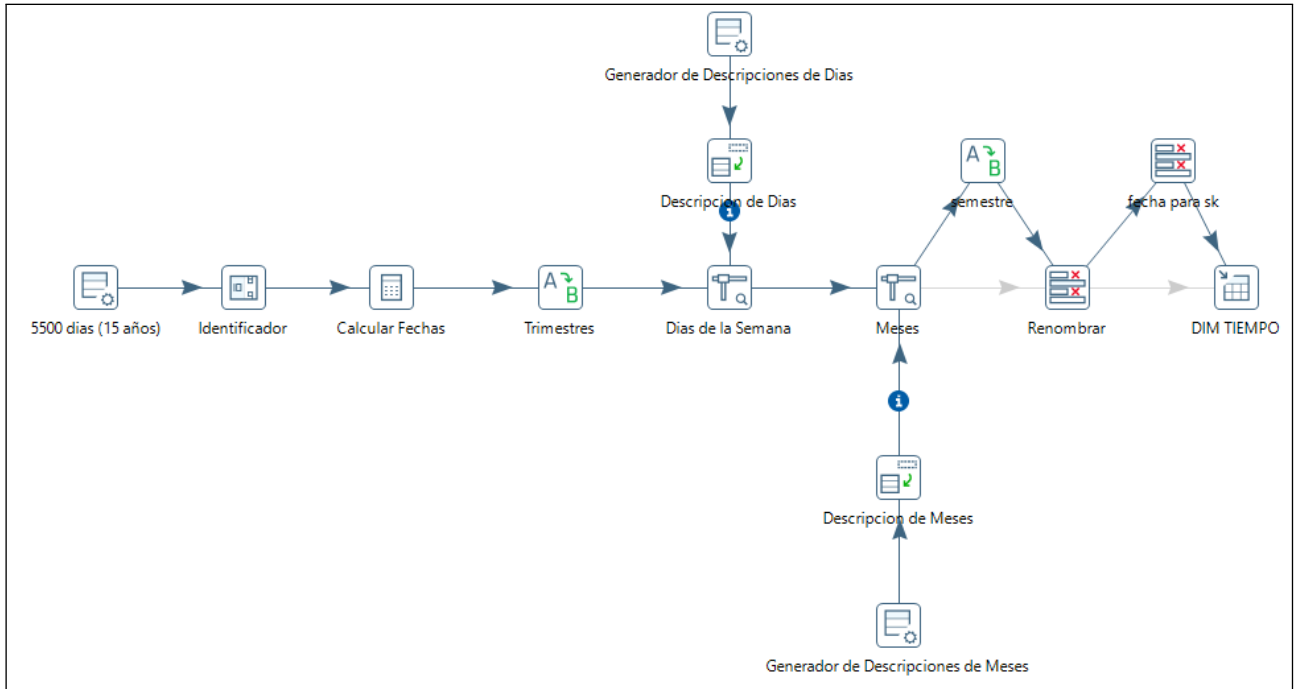


Fig. 27: Transformación de la dimensión tiempo.

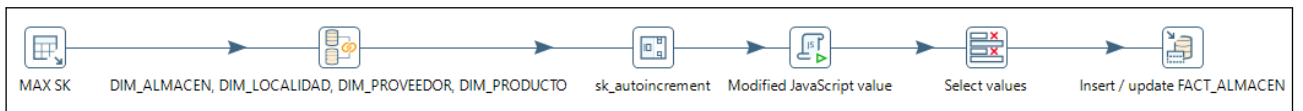


Fig. 28: Transformación de la fact almacen.

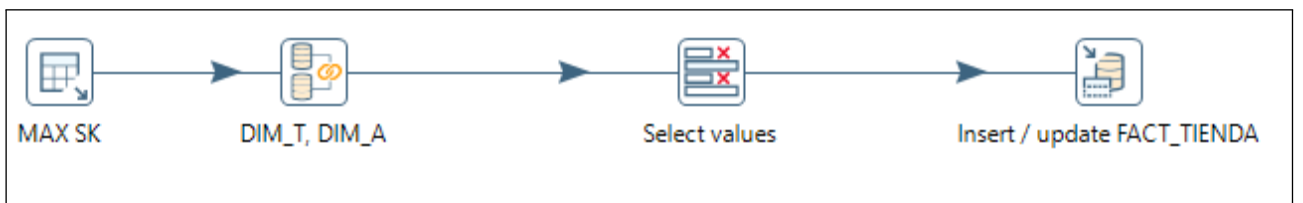


Fig. 29: Transformación de la fact tienda.

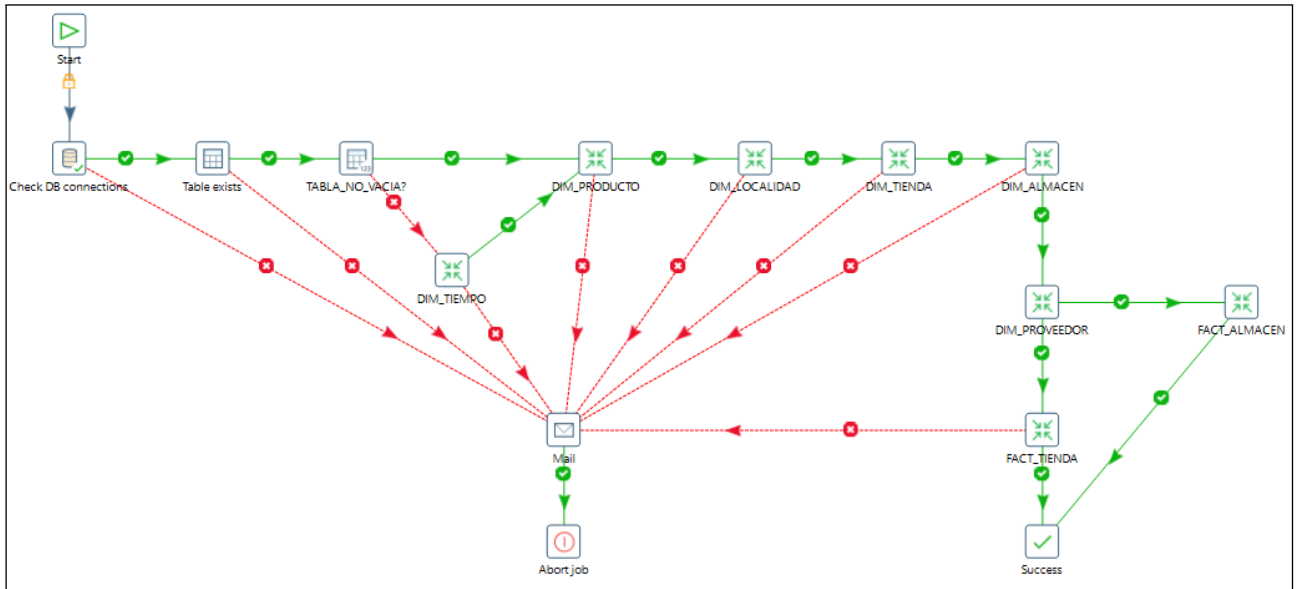


Fig. 30: Job para el almacén de datos.

Por último, se muestra el *Dashboard* principal y las consultas:

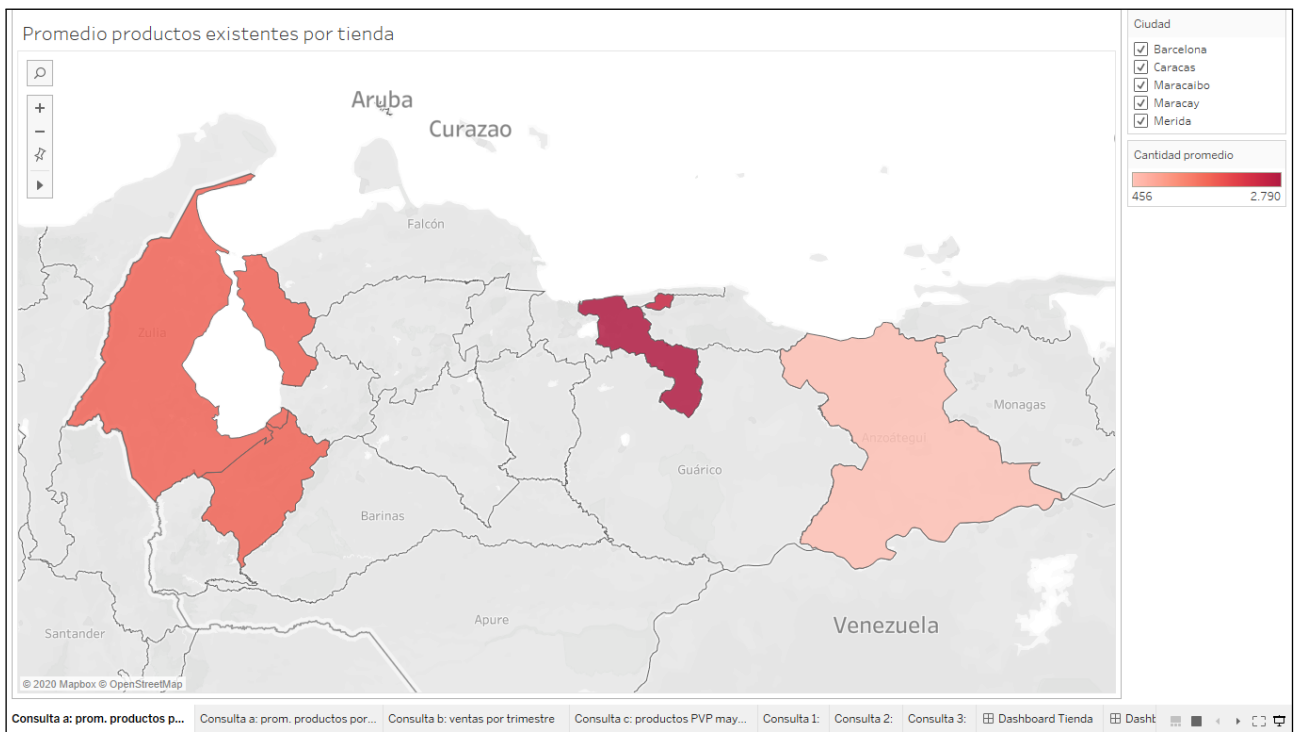


Fig. 31: Consulta a.

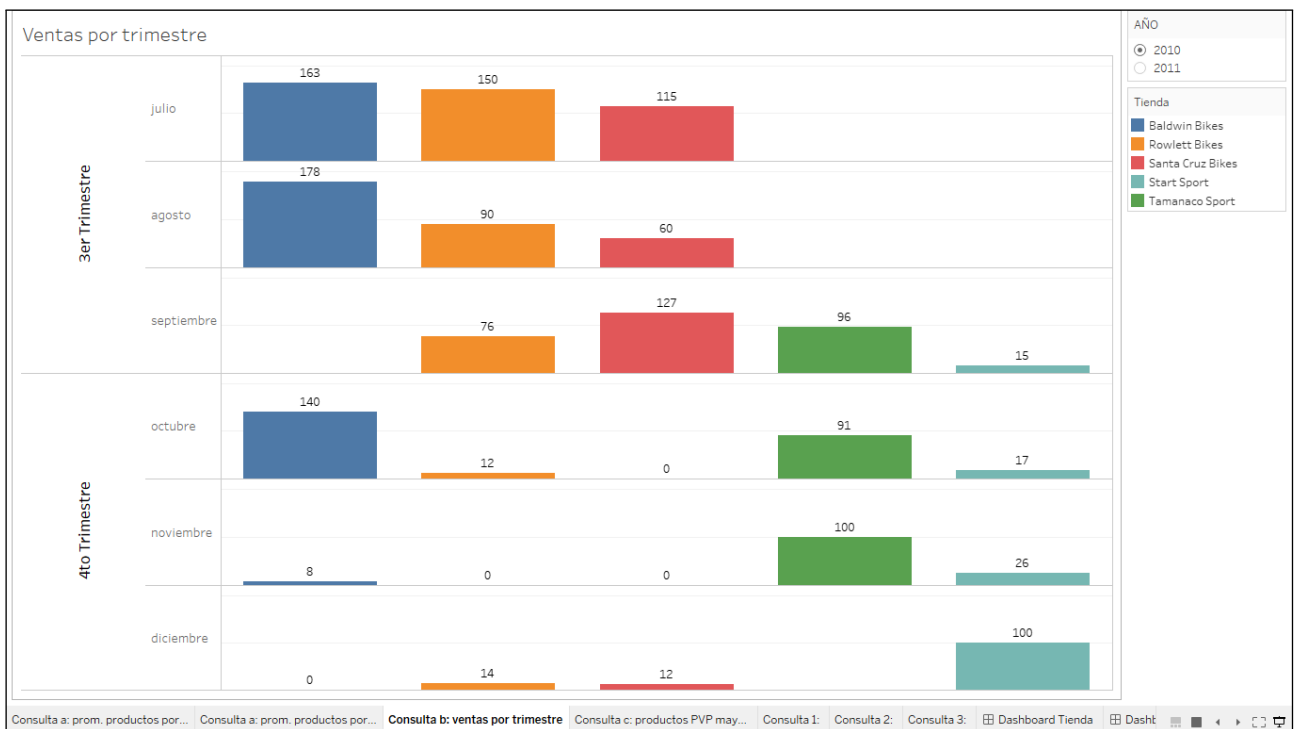


Fig. 32: Consulta b.

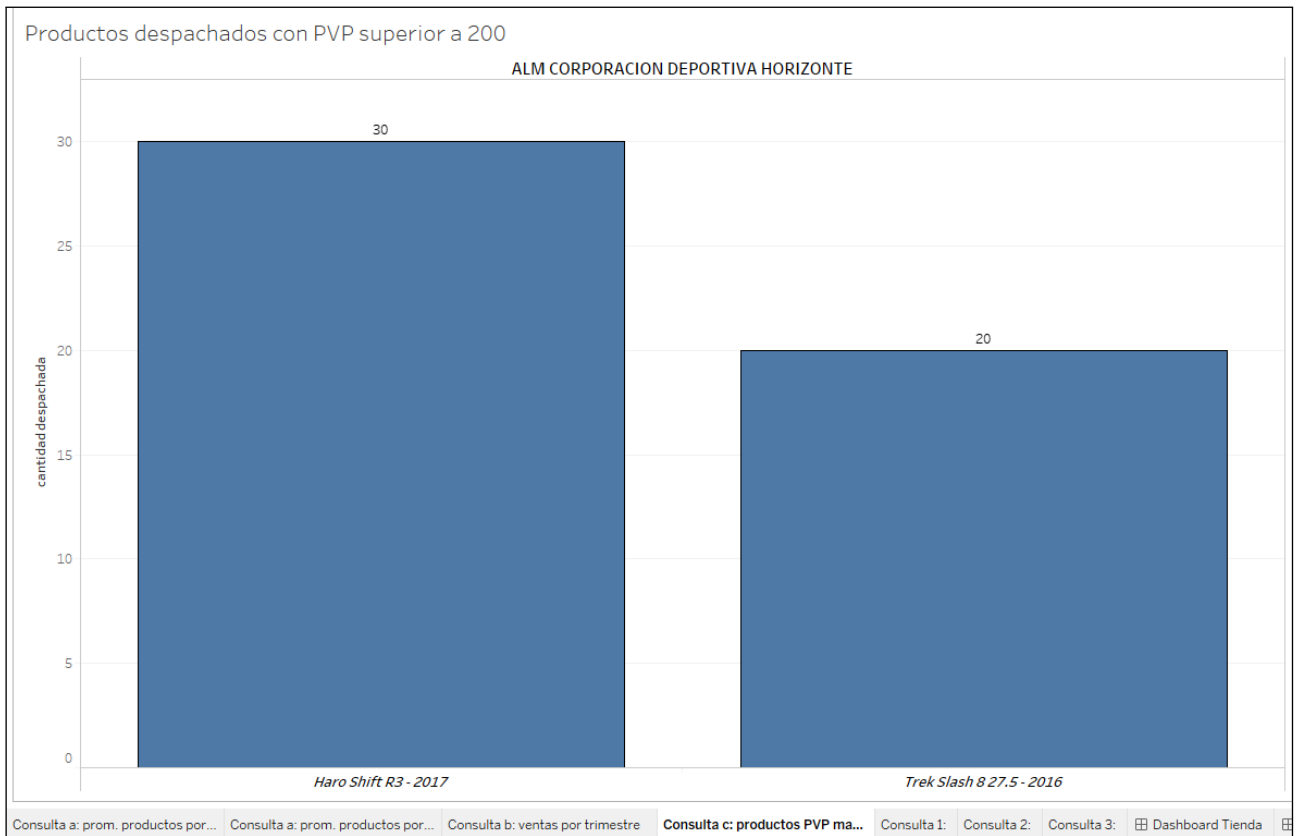


Fig. 33: Consulta c.

Fig. 34: Consulta 1.

Fig. 35: Consulta 2.

Fig. 36: Consulta 3.

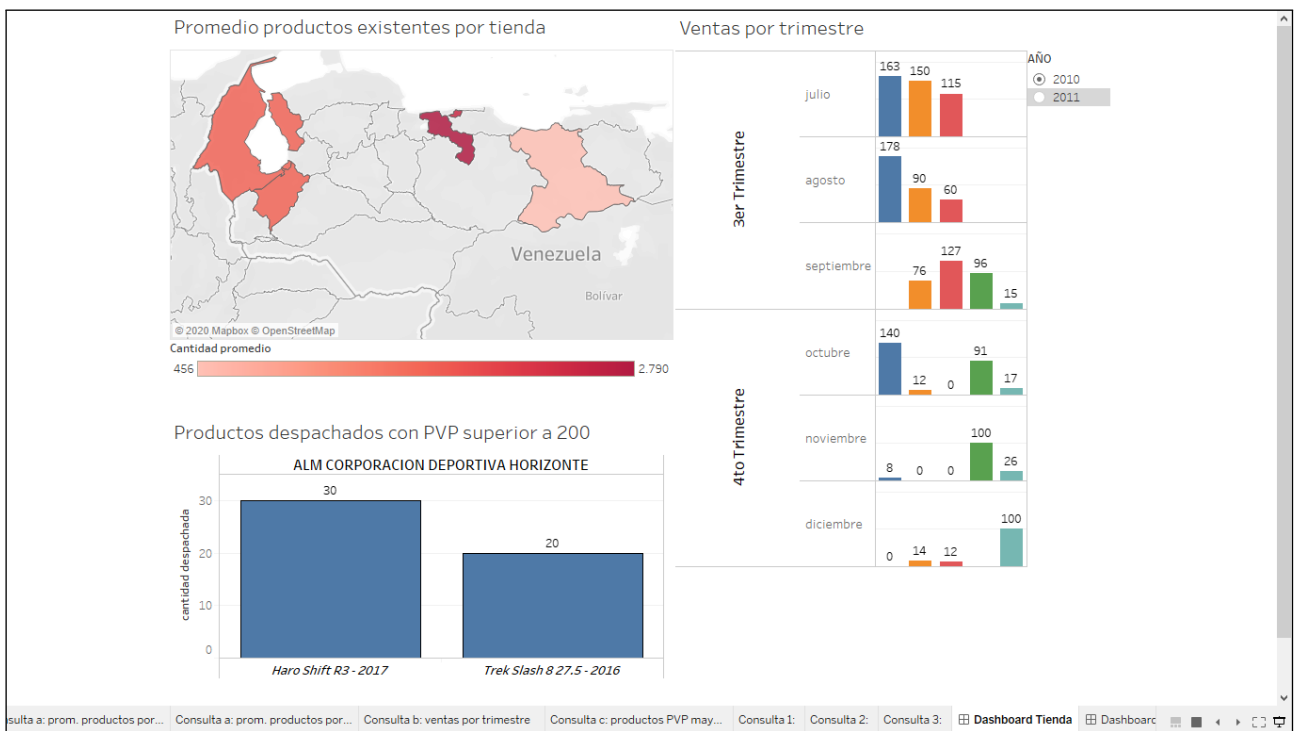


Fig. 37: Dashboard.