

Universidad Don Bosco
Facultad de Ingeniería
Escuela de Electrónica
Datawarehouse y minería de datos

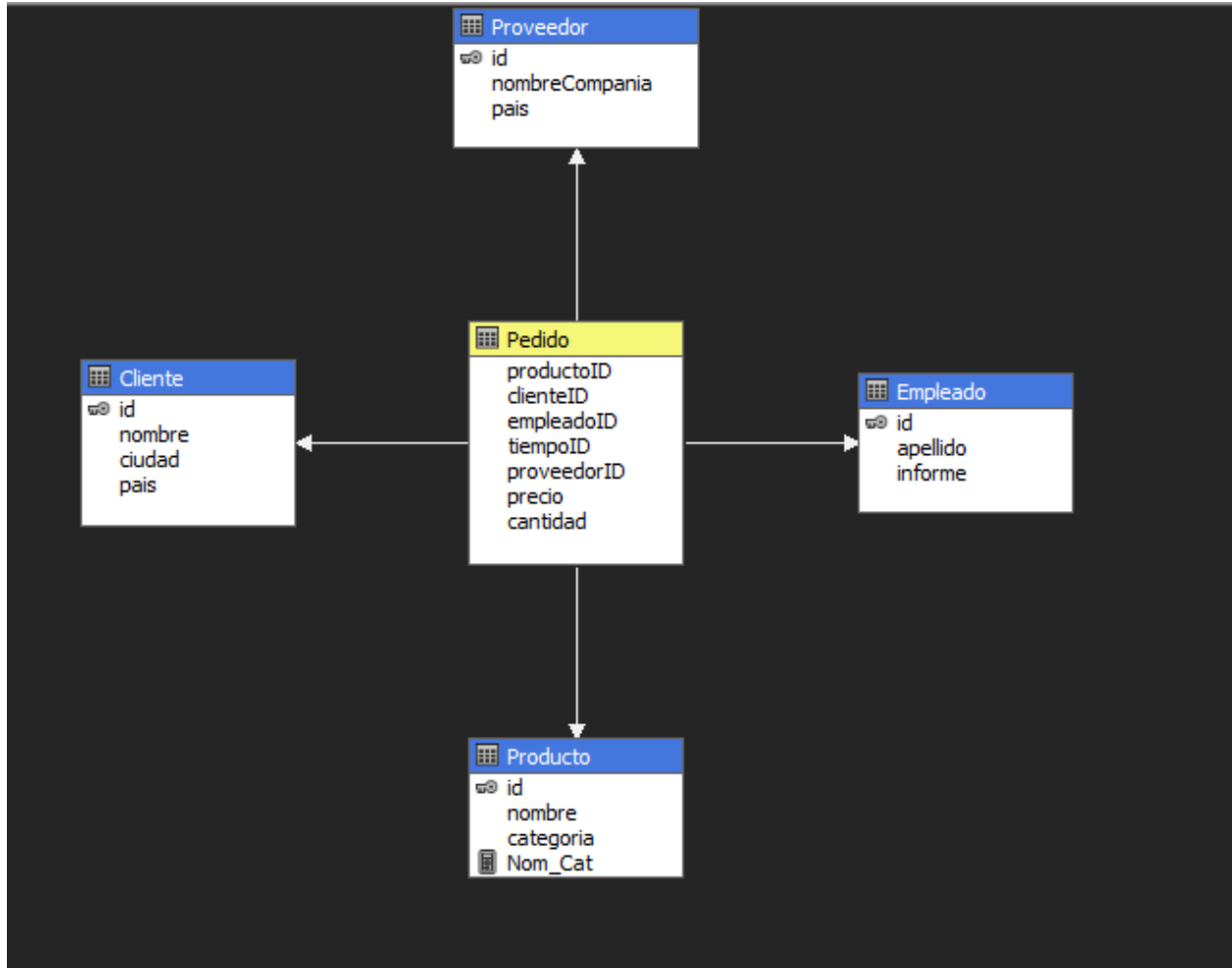


Desafío Practico 2 DMD
Grupo 01L
Integrantes:
Rivas González, Cesar Josué RG180141
Docente:
Alexander Sigüenza

Fecha:
4 de octubre de 2020

Parte 1

Para la primera parte, se creo un cubo con las especificaciones que se pedían para poder visualizar los datos de mejor manera, ademas, se creo un nuevo campo calculado Nom_Cat en donde se concatena el nombre del producto y la catergoría del producto.



A continuación se muestran los resultados en el cubo después de procesado:

Nombre_Cliente	Pais_Cliente	Apellido_Empleado	Pais_Proveedor	Nom Cat	Precio	Cantidad
Alejandra Camino	Spain	Davolio	Australia	Perth Pasties-Meat/Poultry	262	10
Alejandra Camino	Spain	Davolio	Canada	Tourtière-Meat/Poultry	29.5	5
Alejandra Camino	Spain	Davolio	Germany	Nord-Ost Matjeshering-Seafood	207	10
Alejandra Camino	Spain	Peacock	Brazil	Guaraná Fantástica-Beverages	21.6	6
Alejandra Camino	Spain	Peacock	France	Camembert Pierrot-Dairy Products	340	10
Alejandra Camino	Spain	Peacock	Germany	Nord-Ost Matjeshering-Seafood	150.09	7
Alejandra Camino	Spain	Peacock	Italy	Ravioli Angelo-Grains/Cereals	31.2	2
Alejandra Camino	Spain	Peacock	UK	Teatime Chocolate Biscuits-Confections	7.3	1
Alejandra Camino	Spain	Peacock	USA	Steeleye Stout-Beverages	57.6	4
Alexander Feuer	Germany	Buchanan	Australia	Vegie-spread-Condiments	263.4	6
Alexander Feuer	Germany	Buchanan	Finland	Lakkalikööri-Beverages	180	10
Alexander Feuer	Germany	Buchanan	France	Raclette Courdavault-Dairy Products	660	12
Alexander Feuer	Germany	Buchanan	Italy	Mozzarella di Giovanni-Dairy Products	1044	30
Alexander Feuer	Germany	Leverling	Australia	Pavlova-Confections	349	20
Alexander Feuer	Germany	Leverling	Canada	Tarte au sucre-Confections	986	20
Alexander Feuer	Germany	Leverling	Netherlands	Zaanse koeken-Confections	114	12
Alexander Feuer	Germany	Peacock	Italy	Gorgonzola Telino-Dairy Products	125	10
Alexander Feuer	Germany	Peacock	Japan	Konbu-Seafood	120	20
Ana Trujillo	Mexico	King	Australia	Outback Lager-Beverages	60	5
Ana Trujillo	Mexico	King	Norway	Cudbrandalekost-Dairy Products	28.8	1

Parte 2

Para migrar los datos del documento .csv a una base de datos se utilizó el siguiente ETL, en el que se enviaron los datos con campos nulos hacia un documento de texto externo, y se cambió el nombre de departamento en el componente de script para que sea más fácil su comprensión, el segundo componente de script sirve para que la tabla timeline espere a las demás tablas antes de llenarse para no tener error en la referencia a la llave foránea.

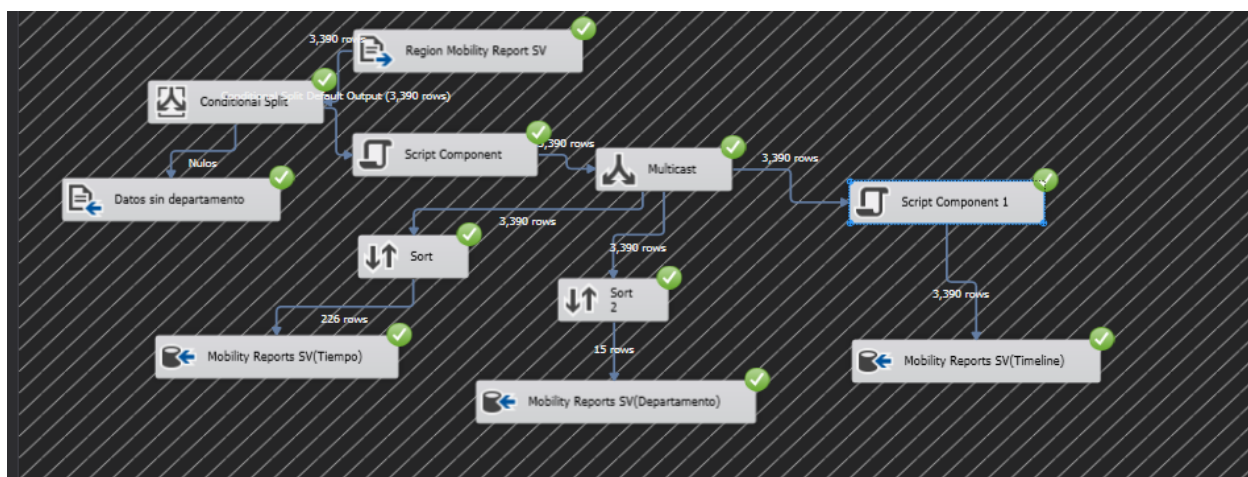
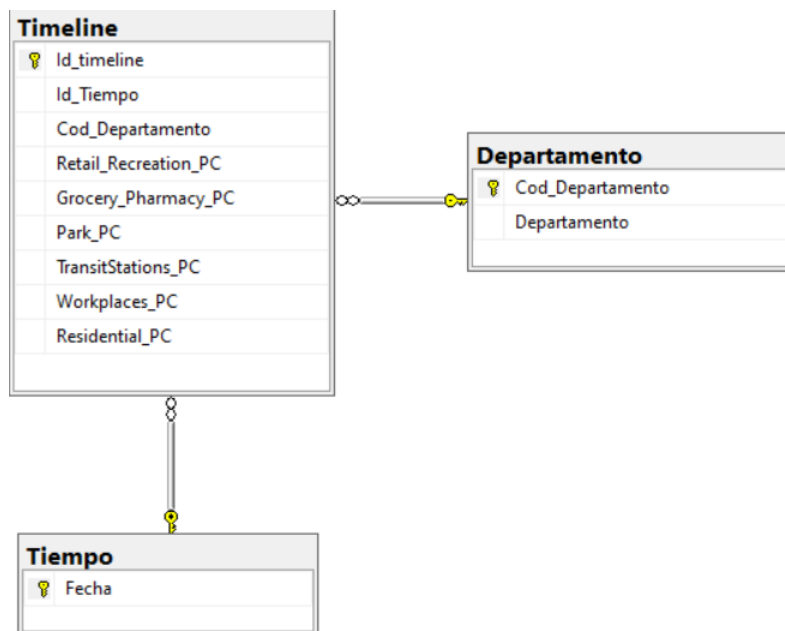


Diagrama de la base de datos:



ANÁLISIS DE INFORMACIÓN

Para realizar el análisis de los datos proporcionados, se crearon dos nuevas columnas calculadas para poder tener una visión mas concreta de como fueron evolucionando los cambios de movilidad en el país, para ello se añadieron dos columnas nuevas: mes y trimestre, en el que nos será posible ir viendo como ha ido cambiando la pandemia a la manera en la que la gente se moviliza, los lugares que frecuentan, etc. A continuación se muestra el cálculo de las dos nuevas columnas y del cubo.

Edit Named Calculation

Column name:

Description:

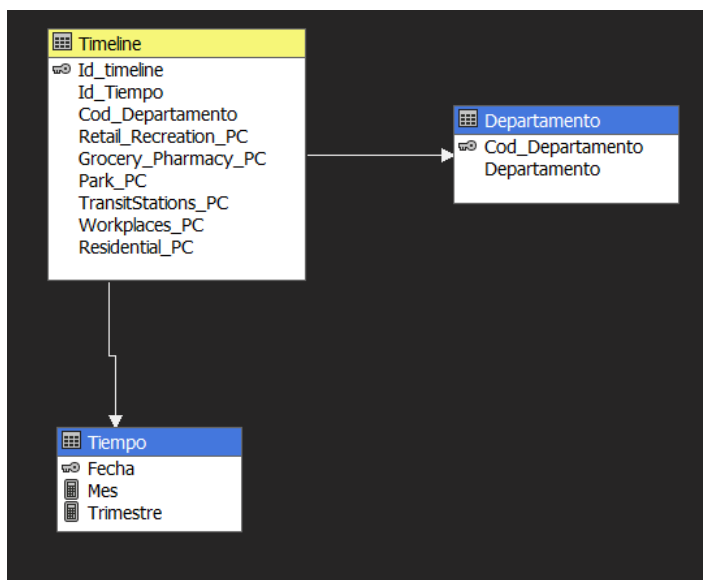
Expression:

Edit Named Calculation

Column name:

Description:

Expression:



1. Momento en que inicia la pandemia

Lo primero que se notó inmediatamente al empezar el análisis, es el cambio brusco que hubo en la cantidad de personas que asistían a las zonas públicas del país alrededor del momento en que se declaró la cuarentena en El Salvador, que fue exactamente el 11 de marzo de 2020 y se puede notar perfectamente el cambio en ese día como podemos observar:

Fecha	Retail Recreation PC	Grocery Pharmacy PC	Park PC	Transit Stations PC	Workplaces PC	Residential PC
2020-03-07	-24	-21	-75	-49	76	1
2020-03-08	-21	-18	-47	-57	37	5
2020-03-09	-26	-4	-66	-25	107	-4
2020-03-10	-26	-4	-44	-28	96	-12
2020-03-11	-13	84	-28	-25	87	-11
2020-03-12	-64	44	-87	-99	-116	36
2020-03-13	-82	66	-123	-110	-19	34
2020-03-14	-126	39	-217	-144	-1	29
2020-03-15	-280	3	-314	-182	-38	32

A partir del 12 de marzo todas las zonas publicas observaron un gran decremento en la cantidad de personas que asistían, sobre todo los lugares de trabajo, mientras que las personas que se quedaban en su casa aumentaron en gran medida.

2. Auge de la pandemia

Aunque no se aprecien todos, la tendencia que se nota en la gran mayoría de departamentos del país es que en el segundo trimestre del año es donde la gente pasaba la mayor parte de su tiempo en cuarentena en su casa, y donde los lugares públicos pasaban vacíos, ya que la gran parte de ellos estaban cerrados a excepción de las farmacias y supermercados.

Departamento	Trimestre	Retail Recreati...	Grocery Pharmac...	Park PC	Transit Stations PC	Workplaces PC	Residential PC
Ahuachapan	1	-875	-201	-694	-455	-489	-11
Ahuachapan	2	-5354	-2126	-5598	-4598	-4166	0
Ahuachapan	3	-1910	-956	-2267	-906	-2897	49
Cabañas	1	0	0	-621	0	-510	0
Cabañas	2	0	0	-4237	0	-4440	0
Cabañas	3	0	0	-1735	0	-3318	0
Chalatenango	1	-918	0	-928	-72	-517	0
Chalatenango	2	-5372	0	-4797	-1311	-4200	0
Chalatenango	3	-1964	0	-1891	0	-3055	0
Cuscatlan	1	-332	0	-898	-560	-538	-6
Cuscatlan	2	-1735	0	-5219	-4164	-4605	0
Cuscatlan	3	-166	0	-1946	-1002	-3049	41
La Libertad	1	-1157	-539	-1013	-798	-728	472
La Libertad	2	-6854	-4725	-6231	-5192	-5511	3067
La Libertad	3	-2751	-1597	-2370	-1969	-3814	1942
La Paz	1	-1071	-735	-659	-1884	-504	0
La Paz	2	-6036	-4639	-3822	-8453	-4395	0
La Paz	3	-2239	-1327	-1295	-4287	-3066	166

La Union	1	-627	-8	-1239	0	-466	0
La Union	2	-5280	-124	-5893	0	-4141	0
La Union	3	-1994	-443	-2726	0	-3267	25
Morazan	1	0	0	-1108	0	-514	0
Morazan	2	0	0	-4986	0	-4220	0
Morazan	3	0	0	-1944	0	-3318	0
San Miguel	1	-1040	-550	-875	-737	-587	311
San Miguel	2	-6109	-4260	-5204	-6434	-4668	1897
San Miguel	3	-2369	-1520	-2168	-3031	-3223	1180
San Salvador	1	-988	-441	-1082	-1012	-716	433
San Salvador	2	-6603	-4534	-6635	-6199	-5619	2929
San Salvador	3	-2510	-1327	-2567	-2713	-3724	1779
San Vicente	1	-190	0	-941	0	-470	0
San Vicente	2	-3901	0	-5273	0	-4361	0
San Vicente	3	-1368	0	-2131	0	-3209	0
Santa Ana	1	-984	-471	-1016	-1215	-705	338
Santa Ana	2	-6238	-4412	-5713	-6525	-4913	2204
Santa Ana	3	-2283	-1183	-2193	-2612	-3008	1431
Sonsonate	1	-1135	-633	-774	-878	-466	251
Sonsonate	2	-5819	-4132	-4682	-6198	-4337	1570
Sonsonate	3	-2139	-1273	-1721	-2540	-2865	1226
Usulután	1	-1172	-495	-523	0	-487	32
Usulután	2	-6081	-4244	-4711	0	-4535	0

3. Últimos meses que han transcurrido

Por otro lado, en los últimos meses que han transcurrido, se nota un gran aumento en la concurrencia de lugares públicos de nuevo, ya que desde hace agosto mas o menos, es que se empezaron a reabrir los negocios y la cuarentena dejo de ser obligatoria. Por lo que se puede ver, la reducción de personas es mucho menor que en el segundo trimestre del año, por lo que se puede esperar que desde el tercer trimestre del año los casos de covid-19 en El Salvador empiecen a crecer de una manera más acelerada que en los meses anteriores.

Departamento	Mes	Retail Recreation PC	Grocery Pharmacy PC	Park PC	Transit Stations PC	Workplaces PC	Resider
Ahuachapán	September	0	0	0	0	-836	49
Cabañas	August	0	0	-523	0	-1161	0
Cabañas	September	0	0	0	0	-910	0
Chalatenango	August	-658	0	-587	0	-1085	0
Chalatenango	September	0	0	0	0	-917	0
Cuscatlán	August	0	0	-631	-277	-1045	0
Cuscatlán	September	0	0	0	0	-818	41
La Libertad	August	-908	-521	-751	-662	-1341	672
La Libertad	September	0	0	0	0	-1006	453
La Paz	August	-724	-391	-342	-1468	-1044	36
La Paz	September	0	0	0	0	-883	130
La Union	August	-623	-37	-905	0	-1078	0
La Union	September	0	0	0	0	-968	25
Morazan	August	0	0	-634	0	-1116	0
Morazan	September	0	0	0	0	-1027	0
San Miguel	August	-771	-506	-713	-1012	-1106	399
San Miguel	September	0	0	0	0	-885	276
San Salvador	August	-832	-427	-822	-911	-1322	624
San Salvador	September	0	0	0	0	-947	402

4. Vista general del impacto en los diferentes lugares

De los datos mostrados, podemos concluir que los lugares que mas fueron afectados por la pandemia fueron los lugares de trabajo, los negocios de productos no esenciales y de

recreación (retail) sobre todo los centros comerciales y también los parques. Y los meses en los que estos tipos de lugares mas sufrieron fue en abril y en mayo que era cuando aun estaba la cuarentena obligatoria y empezaban los primeros casos del virus en el país.

Mes	Retail Recreation PC	Grocery Pharmacy PC	Park PC	Transit Stations PC	Workplaces PC	Residential PC
April	-24889	-13269	-27082	-18086	-25269	4985
August	-8600	-3783	-9716	-7104	-16745	3221
February	-141	56	-454	-155	1240	-80
July	-17757	-8516	-21232	-14677	-18254	3879
June	-20400	-9935	-23198	-16959	-19620	4285
March	-11354	-4574	-12886	-8457	-9600	2300
May	-26506	-14432	-28550	-20291	-24446	5039
September	0	0	0	0	-13528	2434