



Universidad Don Bosco

PROYECTO

Alumno: Marco Rodrigo Funes Bonilla

Carné: FB200456

Docente: Karens Medrano

Datawarehouse y Minería de datos Grupo 01T

OBJETIVOS

- Analizar la información almacenada y generada por los informes del gobierno central para el año 2018
- Aplicar técnicas de recolección de la información para elaborar el análisis de los informes “Esquelas” y “Parque vehicular”
- Realizar la documentación concerniente de los datos, como lo puede ser el cubo de información y su diseño, la creación de interfaces de consulta mediante herramientas de software como es power Bi, documentos PDF y/o archivos Excel con tablas dinámicas o estáticas.
- Diseñar, limpiar y estructurar la base de datos para establecer la información que se maneja en los documentos resultantes de los análisis realizados.

MARCO TEÓRICO

Inteligencia de negocios

La inteligencia de negocios (BI) combina análisis de negocios, minería de datos, visualización de datos, herramientas e infraestructura de datos, y las prácticas recomendadas para ayudar a las organizaciones a tomar decisiones más basadas en los datos.

La inteligencia de negocios debe que ser parte de la estrategia empresarial, esta le permite optimizar la utilización de recursos, monitorear el cumplimiento de los objetivos de la empresa y la capacidad de tomar buenas decisiones para así obtener mejores resultados

Datawarehouse

Un Data Warehouse (depósito de datos) es una plataforma utilizada para recolectar y analizar datos provenientes de múltiples fuentes heterogéneas. Ocupa un lugar central dentro de un sistema de Business Intelligence. Esta plataforma reúne diversas tecnologías y componentes que permiten explotar los datos.

Se basan en estructuras multidimensionales en las que se almacena la información calculando previamente todas las combinaciones de todos los niveles de todas las aperturas de análisis.

Data mart

Un data mart es una forma sencilla de un almacén de datos que se centra en un único tema o línea de negocio, como ventas, finanzas o marketing. Dado su enfoque, los data marts obtienen datos de menos orígenes que los almacenes de datos.

Constituye una parte de un data warehouse. Si un data warehouse está formado por todos los procesos de la organización, un Data mart constituye un determinado proceso.

Data mining

La minería de datos o data mining es un proceso técnico, automático o semiautomático, que analiza grandes cantidades de información dispersa para darle sentido y convertirla en conocimiento.

Asociado al nivel organizacional estratégico y tiene por objetivo eliminar los errores cometidos por las personas al analizar los datos debido a prejuicios y dejar que sean los datos que demuestren hechos reales.

Tecnología OLAP (online analytical process)

OLAP es la respuesta más próxima al análisis de consultas de naturaleza dimensional. Es una parte de la amplia categoría de inteligencia de negocios, que también incluye ETL, reportes relacionales y minería de datos

Esta tecnología permite aprovechar como está estructurada la información de un Data mart o un Data warehouse. Fundamentalmente es una tecnología que permite analizar información dinámicamente a los niveles tácticos y estratégicos basados en Cubos que contienen las medidas y las dimensiones.

ETL (Extract, Transform and Load)

Extraer, transformar y cargar (ETL por sus siglas en inglés), es un proceso en data warehouse que se encarga de extraer la información de las fuentes de datos externas, transformarlas para que se adapten a las necesidades del negocio y por último cargarla en el data warehouse

Dashboard

El nombre Dashboard se refiere al tablero de un automóvil, el cual ofrece al conductor información permanente sobre el estado del vehículo. El mundo de los negocios toma la palabra con un sentido similar, pero en lugar de aplicarlo a los automóviles lo refiere a la empresa.

Entonces un dashboard es una herramienta utilizada para seguir, analizar y mostrar datos, por lo general para obtener información del bienestar general de una organización, un departamento o un proceso específico.

Modelado de datos dimensional

La informática empresarial moderna usa la tecnología de almacenamiento de datos para guardar grandes cantidades de estos para su análisis. Puede usar proyectos de modelado de datos dimensionales para el almacenamiento y la recuperación de datos a alta velocidad desde un almacén de datos. Los modelos dimensionales usan datos duplicados o redundantes y priorizan el rendimiento sobre el uso de menos espacio para el almacenamiento de datos.

Las tablas de dimensiones describen entidades empresariales (las cosas que se modelan). Las entidades pueden incluir productos, personas, lugares y conceptos, incluido el propio tiempo.

Las tablas de hechos pueden almacenar observaciones o eventos, y pueden ser pedidos de ventas, existencias, tasas de cambio, temperaturas, etc. Una tabla de hechos contiene columnas de clave de dimensiones relacionadas con las tablas de dimensiones y columnas de medida numéricas. Las columnas de clave de dimensiones determinan la dimensionalidad de una tabla de hechos, mientras que los valores de clave de dimensiones determinan la granularidad de una tabla de hechos.

ANTECEDENTES

La modernización de la infraestructura de transporte que comenzó con los ferrocarriles, también se pudo apreciar en las principales ciudades de San Salvador y Santa Ana. Las carretas y carruajes que llevaban a las personas de un punto de la ciudad a otro fueron reemplazados primero por tranvías de tracción animal y después por tranvías eléctricos. Ya en la década de 1920 fueron asfaltadas las principales calles de San Salvador, y la mejoría de las calles obedecía también a otra consideración fundamental: la llegada del automóvil allá por 1915 y, pocos años más tarde, del camión y del autobús.

A partir de entonces el crecimiento de la infraestructura vial urbana e interurbana ha ido incrementándose aceleradamente, de acuerdo a la expansión de centros industriales, de producción, de servicios así como de los habitacionales, prueba de ello es la ampliación de la “mancha urbana” en la ciudad de San Salvador, la cual siempre se ha considerado la principal fuente generadora de crecimiento económico del país, esto genera una demanda de servicios, especialmente de comunicación y transporte, ya que sin ellos no se puede lograr la movilidad de productos para su comercialización, además de que influyen directamente en los costos de los artículos a través de los importes en concepto de producción.

Fue hasta en 1917, que se emite un Decreto Legislativo de creación del Ministerio de Fomento y Obras Públicas, la cual posteriormente asumiría todas las funciones encomendadas a las anteriores oficinas de regulación vial.

En 1920, la Dirección General de Obras Públicas dentro del ramo de Fomento contaba con una Sección de Caminos, así como una Sección de Arquitectura, Saneamiento y Aguas y una Sección de Caminos, Puentes y Calzadas.

En 1936, la Dirección de Obras Públicas estaba integrada por el Departamento de Hidráulica y Mantenimiento del Servicio de Aguas y de la Pavimentación de la Capital y por el Departamento de Urbanización y Arquitectura.

En 1948, El Ministerio de Fomento y Obras Públicas contaba con la Dirección General de Carreteras.

En 1949, El Ramo de Fomento y Obras Públicas estaba formado por:

- Secretaría de Estado
- Comisión Nacional de Electricidad
- Oficina de Cartografía y Geografía
- Bodega
- Dirección General de Obras Públicas
- Dirección General de Carreteras.

En 1951, el Ramo de Fomento y Obras Públicas estaba formado por:

- Secretaría de Estado
- Dirección de Bodegas, Talleres y Canteras
- Dirección de Caminos
- Dirección de Urbanismo y Arquitectura
- Dirección de Obras Hidráulicas, y
- Dirección de Cartografía

EN LA ACTUALIDAD

Actualmente el Ministerio de Obras Públicas, dentro de su organización cuenta con dos Viceministerios: de Transporte, el cual se encarga de la reglamentación del tráfico, tanto rural como urbano, así como de los transportes aéreos, terrestre y marítimos; y de Obras Públicas, que es el encargado de dirigir la planificación, construcción, rehabilitación, reconstrucción, ampliación, expansión y mantenimiento de la infraestructura vial del país.

MARCO LEGAL Y NORMATIVO

- Ley de carreteras y caminos vecinales.
- Reglamento Interno y de Funcionamiento MOPT.

MISIÓN

Institución rectora de la planificación y ejecución de la obra pública y de la gestión inteligente del transporte, que promueve el desarrollo económico y social del país.

VISIÓN

Ser una institución moderna, innovadora y con altos estándares de calidad y de servicio, que, a través de la ejecución de infraestructura habilitante y la gestión inteligente del transporte, contribuya al desarrollo sostenible del país.

VALORES

- Honestidad
- Servicio
- Excelencia
- Lealtad

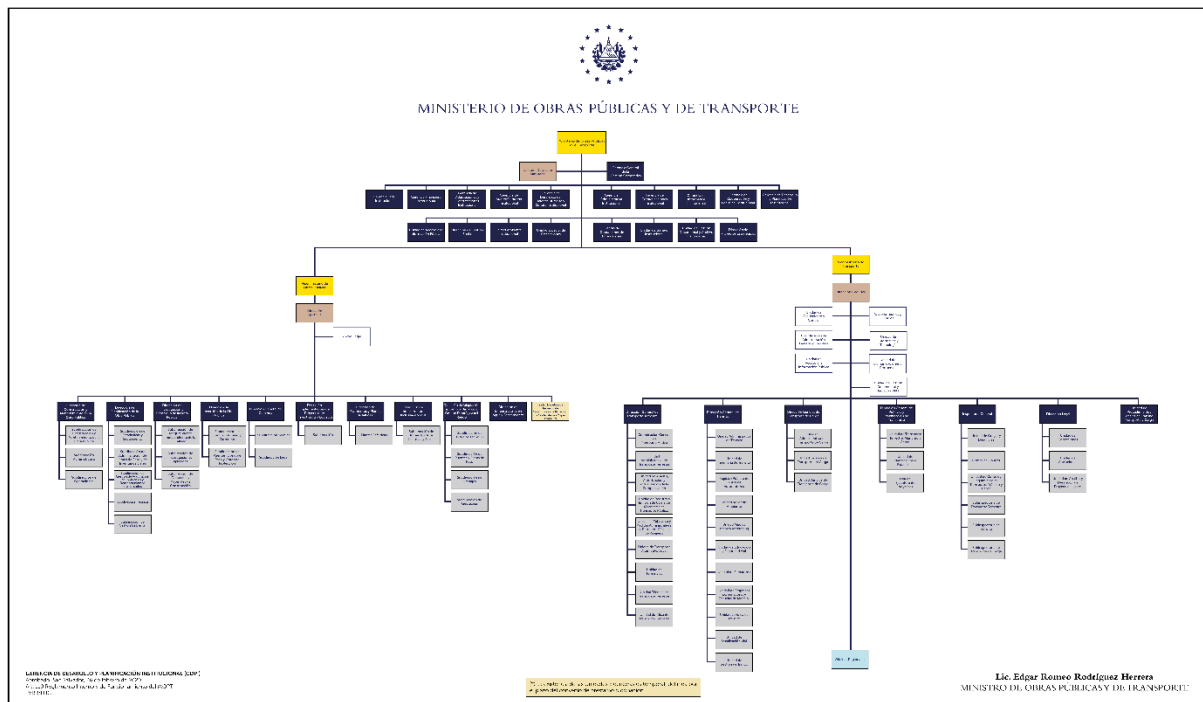
VICEMINISTERIOS

- Viceministerio de Transporte
- Viceministerio de Obras Públicas

SUBDIRECCIONES

- Dirección de Construcción y Mantenimiento de la Obra Pública.
- Dirección de Planificación de la Obra Pública.
- Dirección de Investigación y Desarrollo de la Obra Pública.
- Dirección de Inversión de la Obra Pública.
- Dirección General de Caminos.

Organigrama



METODOLOGIA

La metodología seguida para la investigación de la situación actual está basada en herramientas de recolección de información. La utilización de estas herramientas por medio de softwares como lo es power BI, reportería, archivos excel y el uso de gráficos, tablas y el conteo de datos.

Recolección de datos

En este caso, los datos son proporcionados por medio de un archivo csv que contiene los datos.

Manejo de los datos

Se revisa la estructura de los datos del archivo csv, para mayor facilidad convertimos el formato de datos separado por comas a una estructura de tablas que facilita la lectura de los datos, se revisan y limpian los datos para seguir con el modelado que se utilizara para facilitar el análisis y presentación de la reportería de los datos.

Utilizamos el modelado de datos dimensional, consiste en asignarle variables a atributos como el departamento, el estado y el tipo de falta cometida por el infractor. En este caso se utiliza una tabla de hechos donde dejaremos la base de una esquila como el valor, la fecha, descripciones y luego agregamos las dimensiones, en este caso se crean tres dimensiones con los atributos mencionados anteriormente que serían departamento, estado y el tipo de falta cometida.

Metodología del proyecto

Se define de forma clara y específica el propósito de saber de qué manera se ejecutará el proyecto. Es necesario tener claro cómo se desarrollarán cada una de las etapas del proyecto como lo puede ser el manejo de extracción, limpieza, carga y transformación de los datos, para evitar pérdidas de tiempo y además utilizar las herramienta o métodos específicos de implementación para llevarlo a cabo.

Informes de prueba

Se han realizado informes de prueba en distintos formatos como lo es excel o PDF, pero en este documento será presentado por medio de imágenes, sin embargo, en el repositorio estará anexado estos informes de prueba.

ENUNCIADO DEL PROBLEMA

El ministerio de obras públicas (MOP), ha recolectado mucha información sobre dos temas de suma importancia, pero no sabe cómo presentar los informes al gobierno central para el cierre del año 2018. Los informes que deben entregar serían “Esquelas de infracción de tránsito” y “Parque vehicular”. Por tal motivo a contratado un grupo de consultores para que analicen la información y proporcione un documento con el análisis que realizaron, solo tiene una restricción el análisis debe de ocupar dos estrategias por tema de minería de datos, puede ser cualquier de las descritas a continuación:

- Cubos OLAP
- Reporting Services
- Power BI
- Reglas de asociación
- Agrupamiento con k-means
- Arbol de decisión

Los recursos son proporcionados en archivos CSV, es decir que los datos estarán separados por comas, en este caso los datos fueron separados por punto y coma.

FACTIBILIDAD

Existen distintos tipos de factibilidad como lo puede ser la económica, operativa, técnica, etc. Para este trabajo tendremos un enfoque ligeramente superior al lado operativo. Esto debido a que, para el desarrollo del proyecto y la creación del sistema conformado por los datos, la reportería, el modelado y las distintas estrategias utilizadas, están los usuarios finales (gobierno central) que emplearan el producto generado.

Los beneficios del manejo de datos mediante el proceso ETL, modelado y creación de la data warehouse para hacer el data mining y crear la reportería, presenta una serie de beneficios como lo puede ser:

- Ayudar a comprender una enorme cantidad de datos para que estos puedan ser utilizados para la toma de decisiones, creación de conclusiones para contribuir en la mejora y el crecimiento de regulaciones para los transeúntes.
- Influye también en la parte económica como lo puede ser el ahorro de costes con respecto a la forma en que se analizan los datos, abriendo nuevas oportunidades para invertir en otros sectores del ministerio de obras públicas.
- Con respecto al uso de cubos OLAP nos otorgan beneficios como:
 - La velocidad que permiten a la hora de realizar consultas.
 - El volumen de datos que se puede analizar.
 - La posibilidad que ofrecen al usuario de orientar los informes o análisis según diferentes perspectivas.
 - Buenas capacidades analíticas.
- Para el caso de uso de algunos informes de prueba o plataformas como lo pueden ser Power BI nos otorgan otras ventajas tales como:
 - Optimizar, Limpiar, transformar y combinar datos de múltiples orígenes. Analizar en profundidad los datos y encontrar patrones
 - Innovación. Podrás crear informes sorprendentes con visualizaciones de datos interactivas.
 - Personalizar. Diseñar tu informe mediante las herramientas de creación de temas, formato y diseño.
 - Multiplataforma. Crear informes optimizados para dispositivos móviles.
 - Flexibilidad. Te permite extraer información importante para una amplia gama de escenarios.

JUSTIFICACIÓN

El contenido de estos reportes es utilizado en dos escenarios, uno como análisis interno para los departamentos inmediatos dependiendo del área que lo solicita, dicha información puede ser vital para el nivel estratégico y táctico para toma de decisiones, así como para análisis del trabajo que se desarrolla como institución. El otro uso que se le puede dar a la información de los reportes es para brindar a la población en general un estado de los casos como tal, muchas veces en alguna conferencia o entrevistas la Gerencia General solicita datos específicos de casos a los que se les da seguimiento de la institución, para lo que se obtienen cifras significativas de las esquelas generadas por los infractores.

Es importante mencionar que tanto los ciudadanos como los empleadores consideran que el acceso a la información es de gran importancia aportando al hecho de que generar reportes por ellos mismos (los empleadores) lo consideran como una mejora al rendimiento de su trabajo diario. Se menciona el acceso a los datos debido a que muchos jefes a pesar de tener mayores niveles de control sobre la información requieren siempre de un equipo tecnológico o informático, encargados de manejar la base de datos de forma más técnica por lo que el crear un sistema como lo pueden ser los dashboards les permite crear un enlace en los cuales se pueden generar reportes sin necesidad de solicitarlo a los departamentos de informática.

IMPORTANCIA

Los siguientes elementos resaltan la importancia que conlleva el realizar el proyecto:

- Un mejor análisis interno de la información al poseer un acceso rápido y fácil a los contenidos de información de Esquelas y parque vehicular.
- Creación de reportes a la medida de acuerdo a las necesidades del usuario final (gobierno central), contemplando en todo momento el grado de libertad para filtrar la información de manera sencilla.
- Todos los procesos de la esquila, la adquisición, el estado, el lugar de procedencia y los valores de la información estarán incluidas dentro de la reportería.
- Crear una cultura en el ministerio de obras públicas enfocada a la utilización de herramientas informáticas para la agilización de los procesos operativos internos, lo que aportará un valor agregado al desarrollo profesional de los recursos en general.

Por lo antes expuesto podemos asegurar que la trascendencia que tendrá el desarrollo y finalización del proyecto, tendrá efectos y consecuencias directamente sensibles para la institución en cuanto a la agilización de los procesos internos mediante la automatización de la gestión de información y para la sociedad a través del mejoramiento de la forma en que se presenta la información y la generación de futuras entregas al gobierno central.

ALCANCES

1. La creación de procesos ETL (son las siglas en inglés de Extraer, Transformar y Cargar) que permitan la extracción de datos contenidos en los datos de Esquelas logrando cumplir con las 3 operaciones básicas del ETL las cuales son: extracción, transformación y carga.
2. Integración de informes de prueba que permita generar información estática para exportarse en documentos PDF o tablas de excel que resuma algunos datos como lo pueden ser “Esquelas provenientes del departamento de San Salvador”, “Total de esquelas”, “Total del valor monetario de todas las esquelas”, “Esquelas de tipo transito”, y muchas más oportunidades de filtrado para la presentación de estos datos.
3. Creación de dashboards interactivos como lo es Power BI que mejora la interacción con el analista de información ya que tiene funciones como la facilidad de filtrar datos, el conteo de datos representado en porcentajes y el manejo de la lectura de los datos por medio de gráficos.

LIMITANTES

1. Para este proyecto existen pocas limitantes con respecto al apartado técnico (hardware), no se necesita de un equipo gama alta porque como se menciona anteriormente se ha realizado la integración de informes de prueba y la reportería multiplataforma como lo es Power BI.
2. Respecto al apartado de software power BI es gratuito por lo que no se requieren licencias para acceder al dashboard interactivo con los datos.
3. Datos vacíos, a pesar de haberse hecho una limpieza y transformación de los datos nos hemos percatado de que algunos datos no han sido registrados durante la creación de la esquila.

RECOMENDACIONES

- Recomendamos el requerir de forma obligatoria llenar todos los campos de la esquila para no tener este inconveniente de tener datos vacíos en el futuro. Ejemplo: En el atributo “Departamento” que nos dice la procedencia en la que fue generada la Esquila podemos ver que en lugar de ser 14 departamentos son 15, esto debido a que se tuvo que agregar el concepto de “Desconocido” dando a entender que se desconoce el lugar en el que fue realizada la falta.
- Se recomienda dejar a un encargado ya sea una persona capacitada o a un departamento de IT de dar mantenimiento a la data warehouse como lo puede ser la actualización de los datos, o el ingreso de nuevos datos para mejorar la comunicación y la comprensión de todos aquellos elementos involucrados con la generación de informes y reportería.

INFORMES DE ESQUELA

Power BI

Presentación de tabla en Power BI con todos los datos de las esuelas.

Count of IDEsuela	Year	Quarter	Month	Day	Falta	DescripcionFalta
1	1968	Qtr 1	March	23	TRANSITO	CONducir con licencia inadecuada al tipo de vehiculo
1	1994	Qtr 1	March	10	TRANSITO	NO PORTAR LA LICENCIA DE CONDUCIR
6	1995	Qtr 1	January	1	TRANSITO	CIRCULAR A MAYOR VELOCIDAD QUE LA REGLAMENTARIA
1	1995	Qtr 1	January	1	TRANSITO	CIRCULAR A MAYOR VELOCIDAD QUE LA REGLAMENTARIA
1	1995	Qtr 1	January	1	TRANSITO	NO PORTAR LA LICENCIA DE CONDUCIR
1	2000	Qtr 3	August	21	TRANSITO	NO RESPETAR LA LUZ ROJA DEL SEMAFORO
1	2000	Qtr 3	August	21	TRANSITO	NO RESPETAR SEÑALES DE PRECAUCION
5	2000	Qtr 3	August	21	TRANSITO	NO RESPETAR SEÑALES DE PRECAUCION
1	2000	Qtr 4	October	13	TRANSITO	NO PORTAR LA LICENCIA DE CONDUCIR
2	2000	Qtr 4	October	28	TRANSITO	REBASAR LA DOBLE LINEA AMARILLA CENTRAL QUE DIVIDE LA VIA EN DOBLE SENTIDO DE CIRCULACION
1	2000	Qtr 4	December	31	TRANSITO	CARECER PARCIALMENTE DE LUZ DELANTERA O TRASERA
1	2000	Qtr 4	December	31	TRANSITO	CIRCULAR VEHICULOS DE TRACCION HUMANA O ANIMAL CON RUEDAS METALICAS POR CARRETERAS PAVIMENTADAS, ASI COMO CONDUC AGRICOLAS VEHICULOS PARA CONSTRUCCION Y MONTACARGAS
1	2000	Qtr 4	December	31	TRANSITO	CIRCULAR VEHICULOS DE TRACCION HUMANA O ANIMAL CON RUEDAS METALICAS POR CARRETERAS PAVIMENTADAS, ASI COMO CONDUC AGRICOLAS VEHICULOS PARA CONSTRUCCION Y MONTACARGAS
1	2000	Qtr 4	December	31	TRANSITO	CIRCULAR VEHICULOS DE TRACCION HUMANA O ANIMAL CON RUEDAS METALICAS POR CARRETERAS PAVIMENTADAS, ASI COMO CONDUC AGRICOLAS VEHICULOS PARA CONSTRUCCION Y MONTACARGAS
1	2000	Qtr 4	December	31	TRANSITO	CONducir con licencia inadecuada al tipo de vehiculo
1	2000	Qtr 4	December	31	TRANSITO	CONducir en sentido contrario salvo cuando se sobrepase a otro vehiculo en las zonas y momentos permitidos
1	2000	Qtr 4	December	31	TRANSITO	CONducir en sentido contrario salvo cuando se sobrepase a otro vehiculo en las zonas y momentos permitidos
1	2000	Qtr 4	December	31	TRANSITO	CONducir en sentido contrario salvo cuando se sobrepase a otro vehiculo en las zonas y momentos permitidos
1	2000	Qtr 4	December	31	TRANSITO	CONducir en sentido contrario salvo cuando se sobrepase a otro vehiculo en las zonas y momentos permitidos
1	2000	Qtr 4	December	31	TRANSITO	CONducir motocicleta sin el casco protector
1	2000	Qtr 4	December	31	TRANSITO	CONducir sin estar autorizado para ello
1	2000	Qtr 4	December	31	TRANSITO	CONducir sin estar autorizado para ello
1	2000	Qtr 4	December	31	TRANSITO	CONducir sin estar autorizado para ello
1	2000	Qtr 4	December	31	TRANSITO	CONducir sin estar autorizado para ello
2	2000	Qtr 4	December	31	TRANSITO	CONducir sin estar autorizado para ello
1	2000	Qtr 4	December	31	TRANSITO	CONducir sin estar autorizado para ello
1	2000	Qtr 4	December	31	TRANSITO	CONducir sin estar autorizado para ello
2	2000	Qtr 4	December	31	TRANSITO	CONducir sin estar autorizado para ello
1	2000	Qtr 4	December	31	TRANSITO	ESTACIONAR EL VEHICULO EN AREA EXCLUSIVA PARA MINUSVALIDOS O MUJERES EMBARAZADAS
1	2000	Qtr 4	December	31	TRANSITO	ESTACIONARSE EN PARADAS DE BUSES

Power BI nos ofrece muchas opciones de filtrado con respecto a todos los datos, a continuación, se presenta un ejemplo de la interfaz del filtrado y los resultados de la misma

Fecha - Month
is (All)

Filter type
Basic filtering

Search

☐ January
☐ February
☐ March
☐ April
☐ May
☐ June
☐ July

☐ Require single selection

Count of IDEsuela	Year	Quarter	Month	Day	Falta	DescripcionFalta
6	1995	Qtr 1	January	1	TRANSITO	CIRCULAR A MAYOR VELOCIDAD QUE LA REGLAMENTARIA
1	1995	Qtr 1	January	1	TRANSITO	CIRCULAR A MAYOR VELOCIDAD QUE LA REGLAMENTARIA
1	1995	Qtr 1	January	1	TRANSITO	NO PORTAR LA LICENCIA DE CONDUCIR
1	2002	Qtr 1	January	20	TRANSITO	ESTACIONARSE EN ZONA PROHIBIDA O DE PREFERENCIAL
1	2003	Qtr 1	January	1	TRANSITO	NO PORTAR LA LICENCIA DE CONDUCIR
1	2003	Qtr 1	January	4	TRANSITO	CIRCULAR VEHICULOS DE TRACCION HUMANA O ANIMAL CON RUEDAS METALICAS POR CARRETERAS PAVIMENTADAS, ASI COMO CONDUC AGRICOLAS VEHICULOS PARA CONSTRUCCION Y MONTACARGAS
1	2003	Qtr 1	January	15	TRANSITO	CONducir con licencia inadecuada
1	2003	Qtr 1	January	21	TRANSITO	NO PORTAR TABETA DE CIRCULACION
1	2003	Qtr 1	January	22	TRANSITO	NO PORTAR LA LICENCIA DE CONDUCIR
1	2003	Qtr 1	January	26	TRANSITO	REBASAR LA DOBLE LINEA AMARILLA CENTRAL QUE DIVIDE LA VIA EN DOBLE SENTIDO DE CIRCULACION
1	2004	Qtr 1	January	1	TRANSITO	CARECER DE ESPEJO RETROVISOR Y ESPEJO LATERAL, IZQUIERDO
5	2004	Qtr 1	January	1	TRANSITO	CARECER DE LUZ DE FRENO
3	2004	Qtr 1	January	1	TRANSITO	CARECER DE LUZ DE FRENO
2	2004	Qtr 1	January	1	TRANSITO	CARECER DE LUZ DE PLACA
1	2004	Qtr 1	January	1	TRANSITO	CARECER DE LUZ DE PLACA
1	2004	Qtr 1	January	1	TRANSITO	CARECER DE LUZ DE PLACA
16	2004	Qtr 1	January	1	TRANSITO	CARECER PARCIALMENTE DE LUZ DELANTERA O TRASERA
9	2004	Qtr 1	January	1	TRANSITO	CARECER PARCIALMENTE DE LUZ DELANTERA O TRASERA
5	2004	Qtr 1	January	1	TRANSITO	CARECER PARCIALMENTE DE LUZ DELANTERA O TRASERA
13	2004	Qtr 1	January	1	TRANSITO	CIRCULAR A MAYOR VELOCIDAD QUE LA REGLAMENTARIA
12	2004	Qtr 1	January	1	TRANSITO	CIRCULAR A MAYOR VELOCIDAD QUE LA REGLAMENTARIA
2	2004	Qtr 1	January	1	TRANSITO	CIRCULAR A MAYOR VELOCIDAD QUE LA REGLAMENTARIA
2	2004	Qtr 1	January	1	TRANSITO	CIRCULAR A MAYOR VELOCIDAD QUE LA REGLAMENTARIA
1	2004	Qtr 1	January	1	TRANSITO	CIRCULAR A MAYOR VELOCIDAD QUE LA REGLAMENTARIA
1	2004	Qtr 1	January	1	TRANSITO	CIRCULAR A MAYOR VELOCIDAD QUE LA REGLAMENTARIA
1	2004	Qtr 1	January	1	TRANSITO	CIRCULAR A MAYOR VELOCIDAD QUE LA REGLAMENTARIA
2	2004	Qtr 1	January	1	TRANSITO	CIRCULAR CON LAS LUZES ARRAJADAS EN HORAS NOCTURNAS
1	2004	Qtr 1	January	1	TRANSITO	CIRCULAR CON LAS LUZES ARRAJADAS EN HORAS NOCTURNAS
1	2004	Qtr 1	January	1	TRANSITO	CIRCULAR CON LAS PLACAS NO VISIBLES
1	2004	Qtr 1	January	1	TRANSITO	CIRCULAR CON LAS PLACAS NO VISIBLES
1	2004	Qtr 1	January	1	TRANSITO	CIRCULAR CON UNA PLACA
3	2004	Qtr 1	January	1	TRANSITO	CIRCULAR CON VEHICULOS NO MATRICULADOS

Filters

Search

LA PAZ 6618

☐ Require single selection

DescripcionEstado
is (All)

DescripcionFalta
is (All)

Estado
is (All)

Falta
is (All)

Fecha - Day
is (All)

Fecha - Month
is January

Filter type
Basic filtering

Search

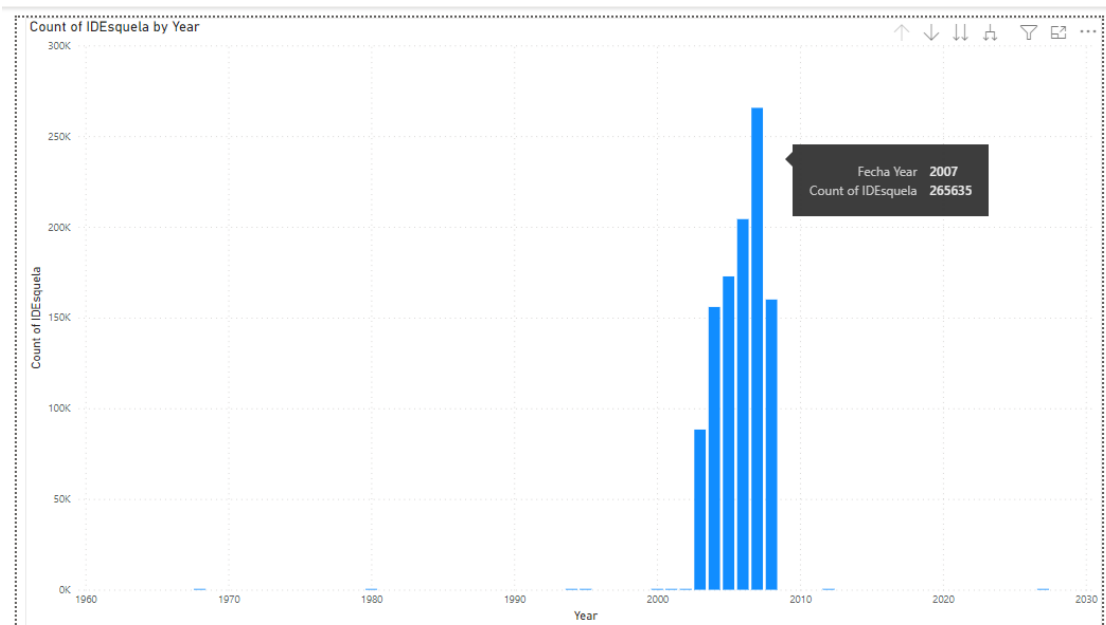
☒ January
☐ February
☐ March
☐ April
☐ May
☐ June
☐ July

☐ Require single selection

Fecha - Quarter
is (All)

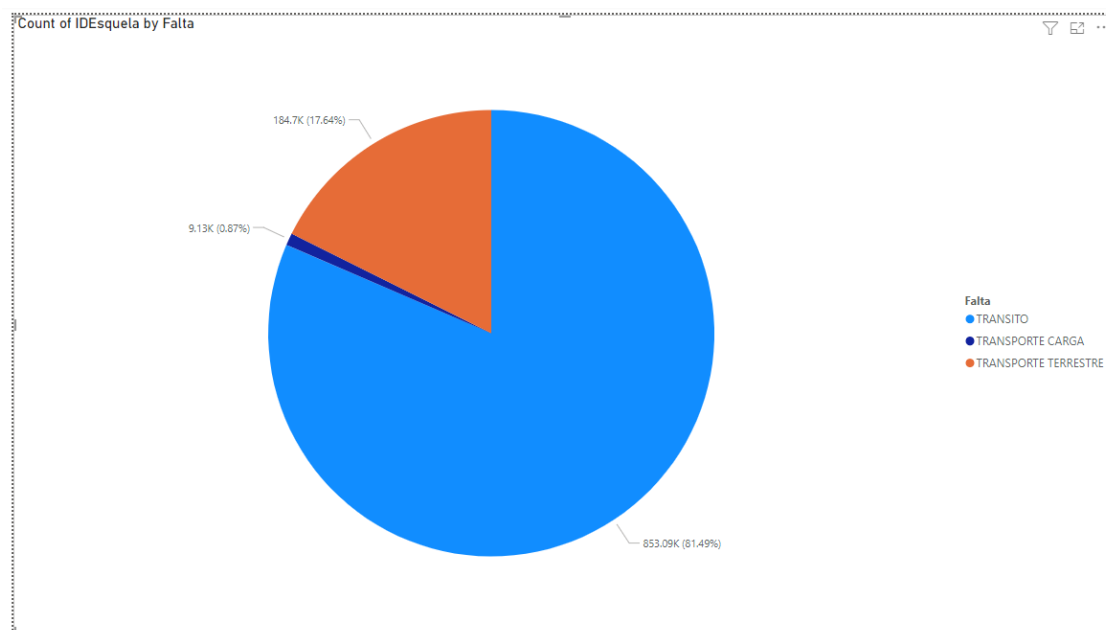
Fecha - Year
is (All)

Otra funcionalidad de Power BI es la generación de gráficas para poder resumir nuestros datos, en este caso queremos saber el conteo de esquelas de los distintos años.



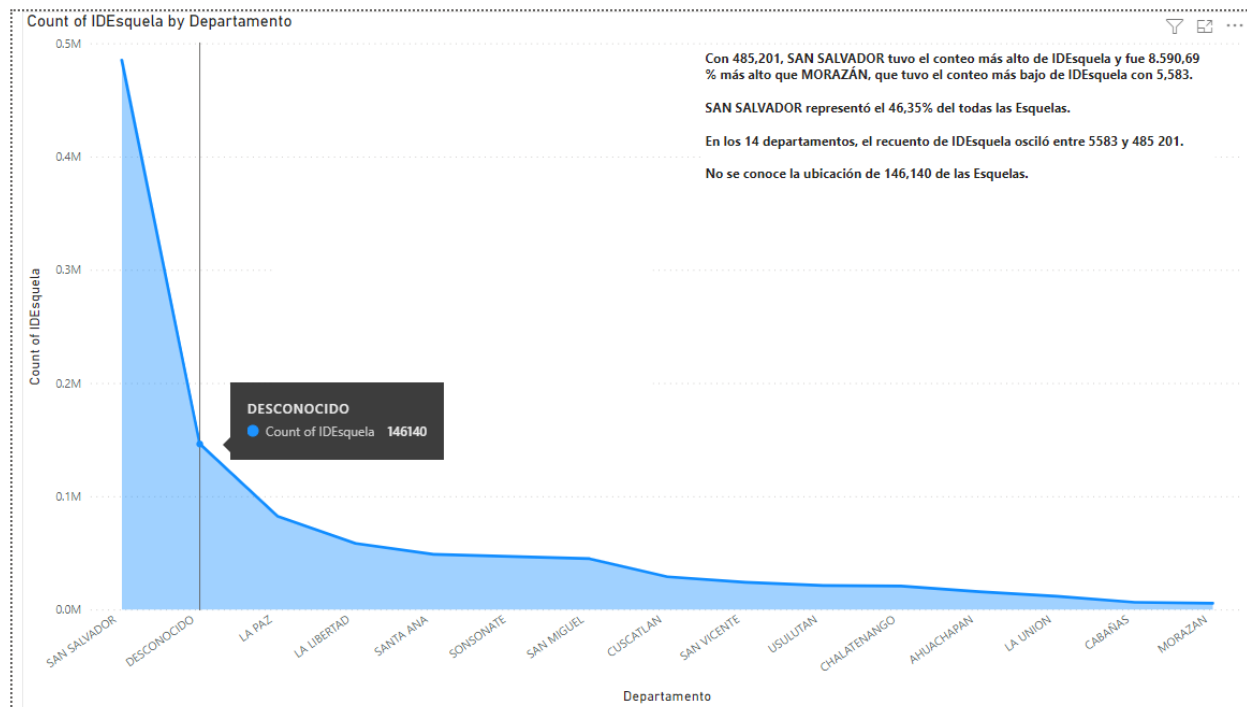
Observamos que el 2007 fue el año con más esquelas con un total de 265,635

Segundo ejemplo de gráficas en Power BI, en este caso queremos ver la distribución de los tipos de faltas cometidas



Mostrándonos que la más frecuente es la de tránsito con un 81.4% (853,088)

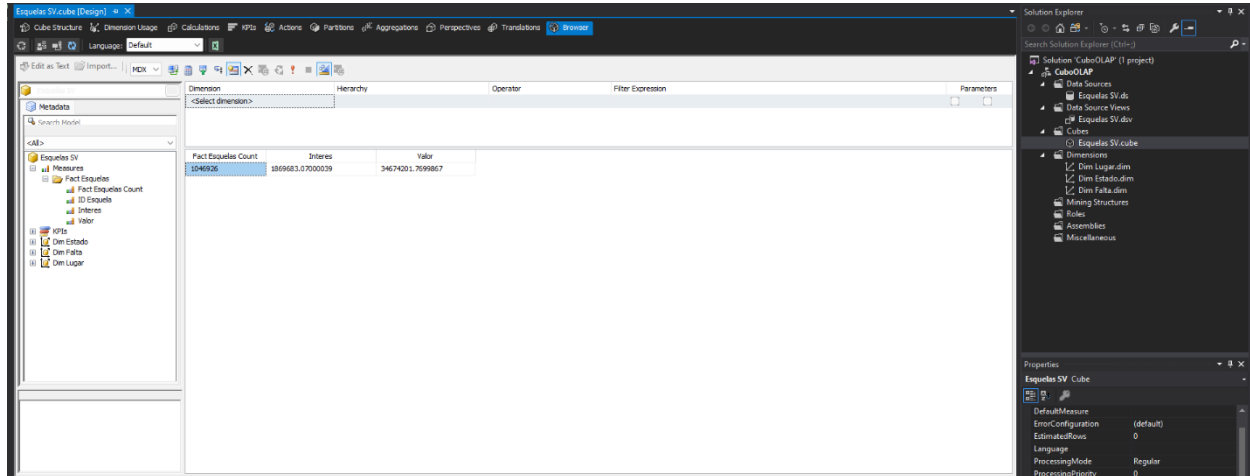
Un último ejemplo de análisis de datos, en este ejemplo analizaremos el departamento o el lugar de origen de la esquila realizada al transeúnte que cometió la falta.



Podemos ver que Power Bi nos permite realizar la sumarización de los datos, haciendo asi un pequeño resumen de la gráfica.

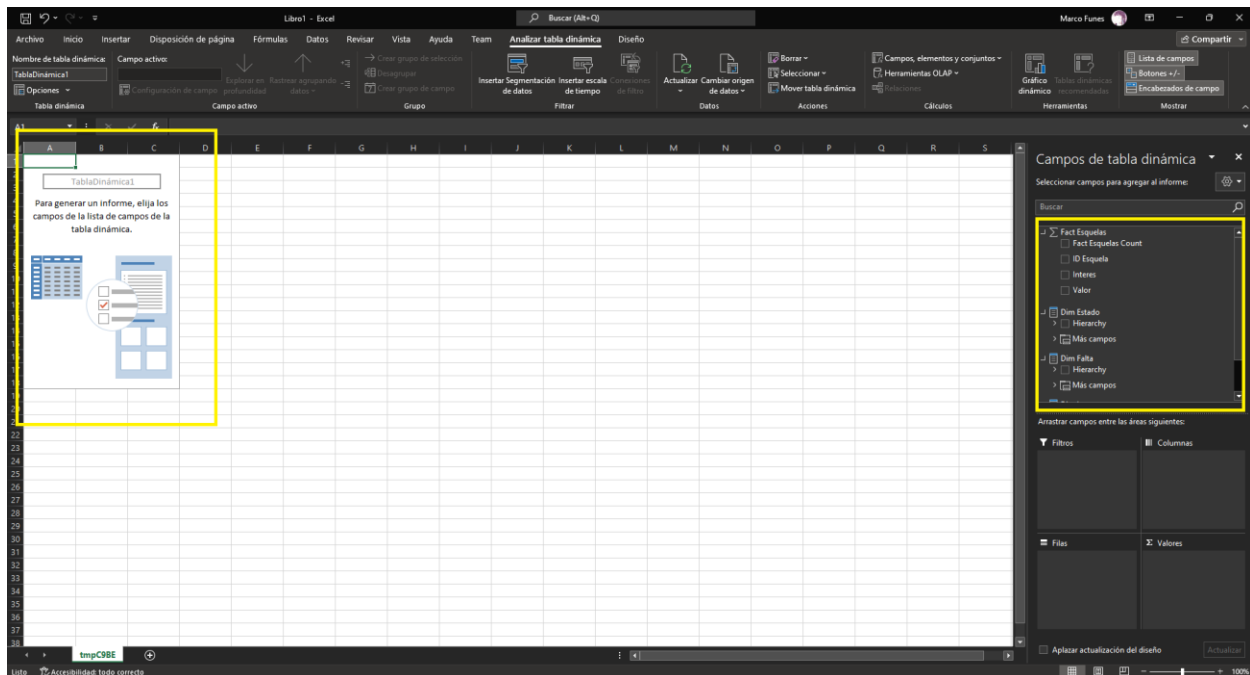
Cubo OLAP

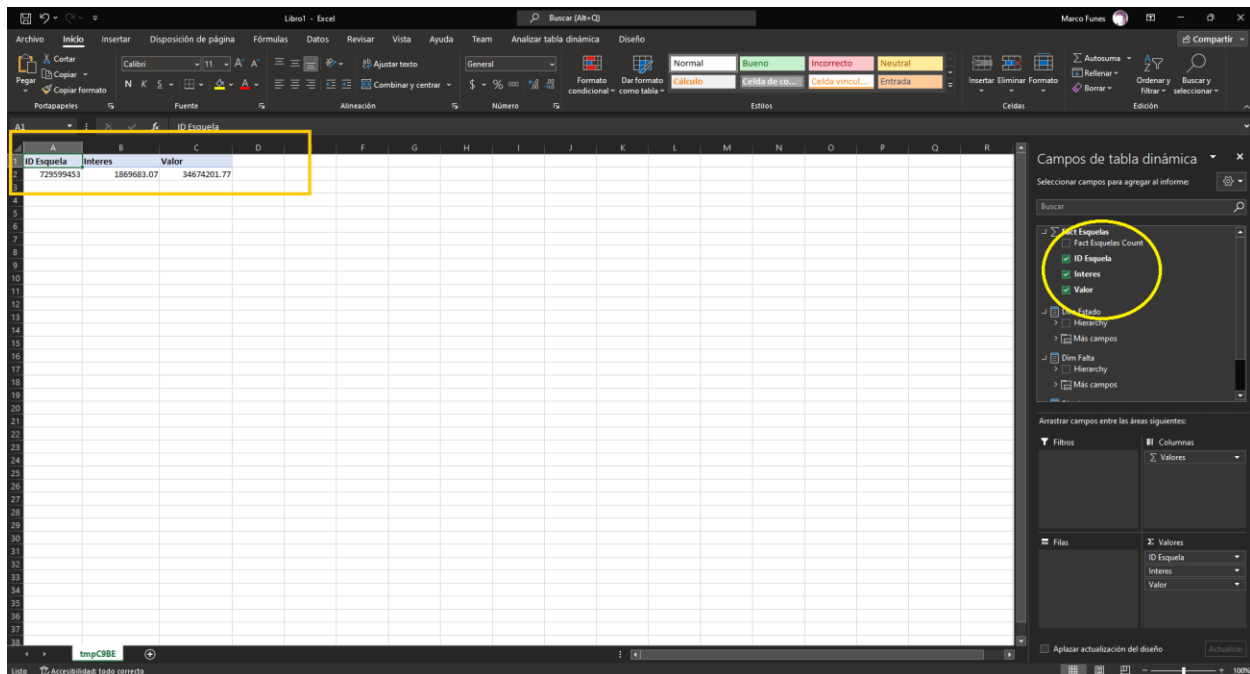
Otra estrategia de análisis que hemos utilizado es la creación de un cubo OLAP para el conteo de métricas, a continuación, dejamos un ejemplo:



Podemos ver un conteo de todas las esquelas, intereses y valores monetarios generados.

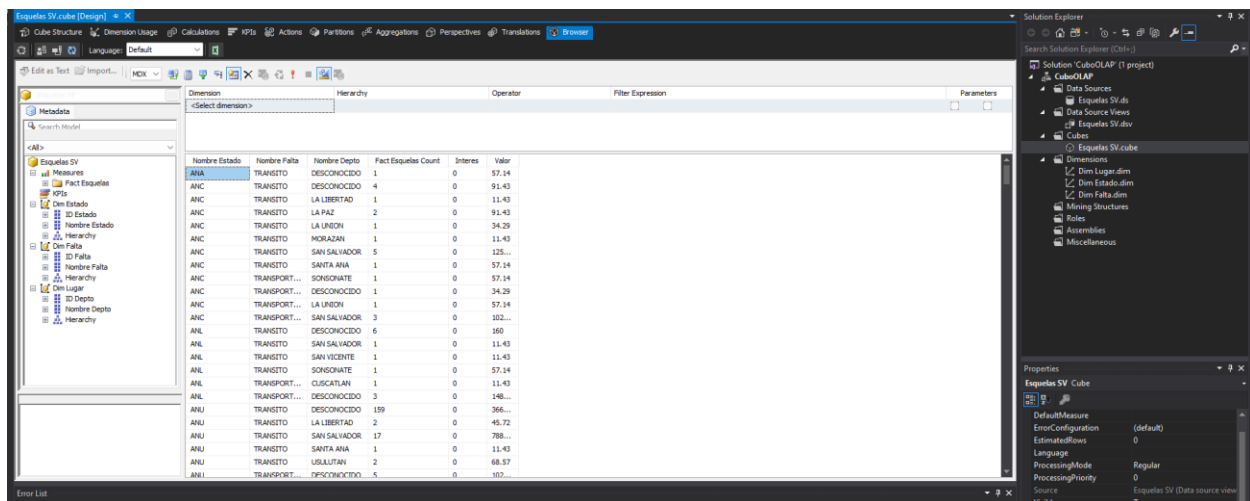
Importante mencionar que esta estrategia nos permite exportar nuestro análisis de datos a un archivo excel.





Podemos observar que al seleccionar nuestras métricas se van agregando de manera automática a la tabla con su respectivo formato.

Por último, vamos a generar una tabla completa de los datos con distintos atributos y filtrarlos para manejar la información.



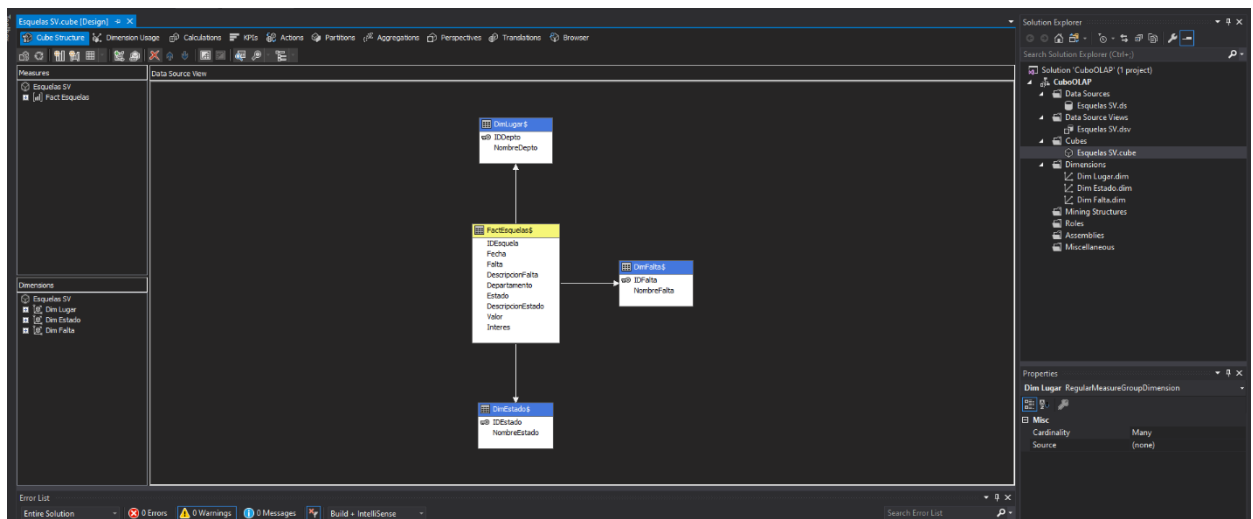
Agregando dimensiones a la tabla para poder observar el estado, departamento y el tipo de falta

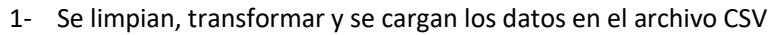
Ejemplo de reporteria CUBO OLAP utilizando filtro de estado y tipo de falta - Excel

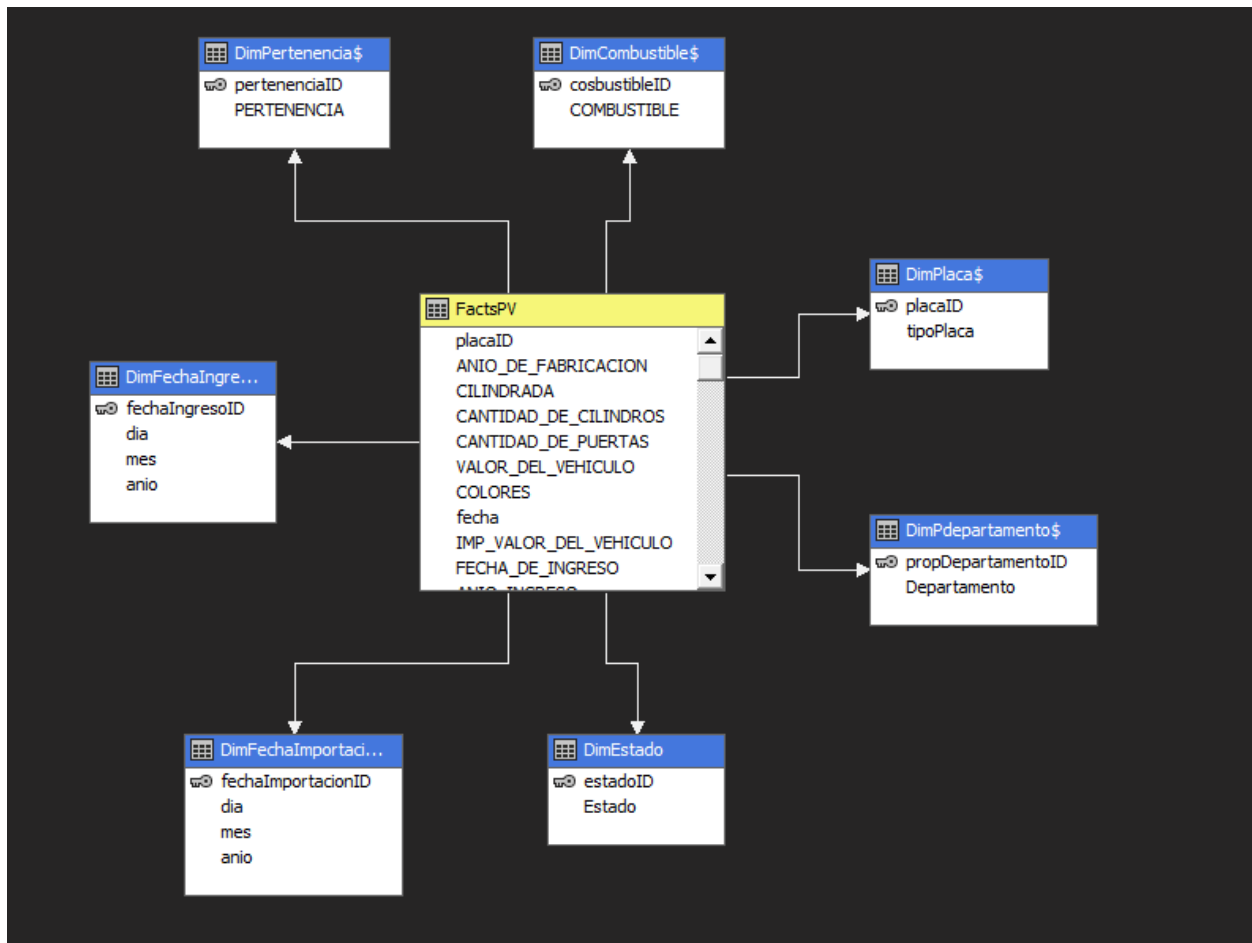
Etiquetas de fila	ID Esquela	Interes	Valor
ANA			
DESCONOCIDO	1725	0	57.14
ANC			
DESCONOCIDO	1990158	0	125.72
LA LIBERTAD	5	0	11.43
LA PAZ	761233	0	91.43
LA UNION	420139	0	91.43
MORAZAN	760373	0	11.43
SAN SALVADOR	4383072	0	238.58
SANTA ANA	673842	0	57.14
SONSONATE	671611	0	57.14
ANL			
CUSCATLAN	669632	0	11.43
DESCONOCIDO	801350	0	308.57
SAN SALVADOR	644384	0	11.43
SAN VICENTE	433342	0	11.43
SONSONATE	1150242	0	57.14
ANU			
DESCONOCIDO	1651014	0	4034.4
LA LIBERTAD	1738758	0	217.15
SAN SALVADOR	13531368	0	822.85
SANTA ANA	704926	0	11.43
SONSONATE	1541318	0	114.28
USulután	1724576	0	68.57
APL			
CHALATENANGO	1128778	0	57.14
CUSCATLAN	1755599	0	171.42
DESCONOCIDO	8226570	0	514.26
LA LIBERTAD	6292395	0	342.84
LA PAZ	4197064	0	228.56
SAN MIGUEL	1347644	0	57.14
SAN SALVADOR	379730122	0	18901.94
SANTA ANA	708795	0	57.14
SONSONATE	1087501	0	57.14
CBR			
AHUAHUAPAN	-1541501105	0	372841.05
CABAÑAS	-20005051241	0	130149.72

Ejemplo de exportación de datos utilizando filtros de departamento y estado de las esquelas.

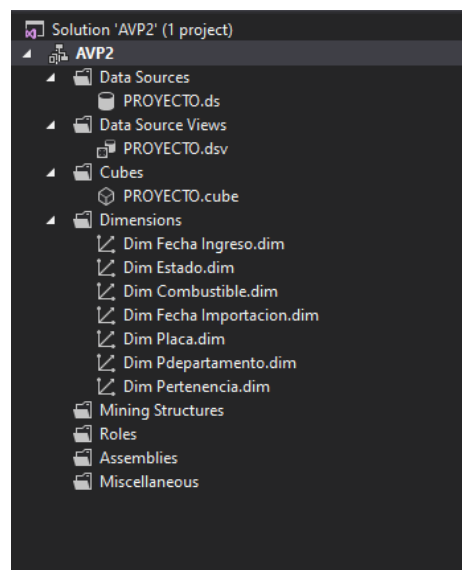
MODELADO CUBO OLAP





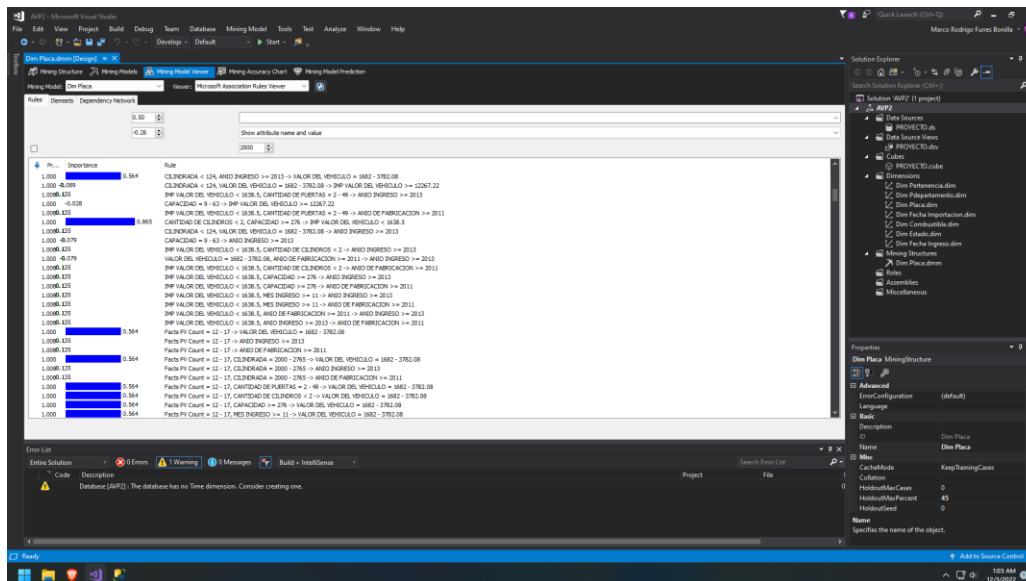
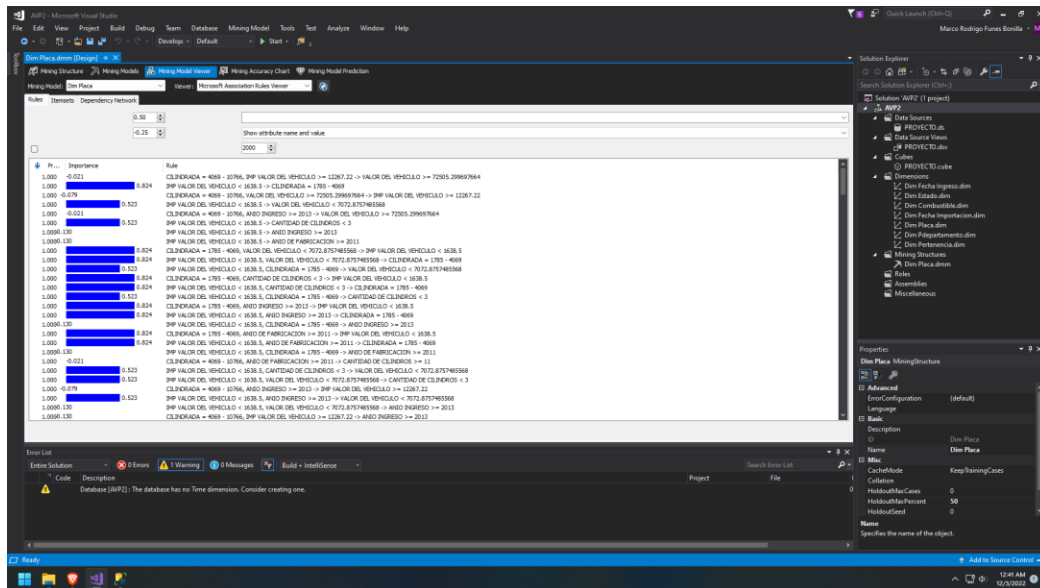


3- Integración de la base de datos en Visual Studio para la creación de reglas de asociación



4- Especificamos los datos que vamos a estudiar para las reglas de asociación

- 5- En este caso, como es recomendable trabajar con menos del 50% de los datos. Pero nosotros elegiremos exactamente la mitad para tener un análisis lo más completo posible.
- 6- Estos son los resultados de algunas métricas con respecto a los tipos de placa de vehículo



7- Ahora trabajaremos con un árbol de decisiones para poder ver otra forma de analizar nuestros datos

En este caso para poder ver la información necesitamos filtrar según la cantidad de niveles que queremos ver para este ejemplo solo estaremos usando 3 así podemos observar los datos de mejor manera.

