



Informe 1: “Telco Costumers”

Integrantes:

Francisco Olivares 201473575-8 francisco.olivars.14@sansano.usm.cl
Andrés Huerta 201473544-8 andres.huerta.14@sansano.usm.cl
Gabriel Valenzuela 201473505-7 gabriel.valenzuel.14@sansano.usm.cl

1 Descripción

El caso asignado consiste en el estudio de clientes de una empresa de telecomunicaciones, cuya finalidad es la de detectar posibles usuarios que se quieran dar a la fuga y así poder tomar acciones a tiempo, como lo es ofrecerle algún tipo de promoción.

2 Requisitos

Requisito	Perfil de los clientes fugados
Descripción	Identificar comportamientos de clientes ya fugados de la compañía para, en un futuro, tomar medidas y así evitar que se den a la fuga.
KPI's	<ul style="list-style-type: none">• Cantidad promedio de cargos totales de clientes fugados por sexo, adulto_mayor, tipo_contrato, pago_mensual, antigüedad, plan.• Cantidad promedio de cargos mensuales de clientes fugados por sexo, adulto_mayor, tipo_contrato, pago_mensual, antigüedad, plan.

Requisito	Planes con mayor satisfacción
Descripción	Determinar cuales son los planes con mayor satisfacción entre los clientes con el fin de recomendar o generar promociones para incentivar la contratación éstos.
KPI's	<ul style="list-style-type: none">• Cantidad de clientes no fugados por sexo, adulto_mayor, plan.• Cantidad de clientes por sexo, adulto_mayor, plan.



Requisito	Tipos de clientes mas importantes para la compañía
Descripción	Reconocer cual grupo de clientes son los que aportan más ingresos a la compañía, de esta forma la empresa puede determinar que grupos son más importantes retener y cuales no.
KPI's	<ul style="list-style-type: none"> • Cantidad promedio de cargos mensuales de clientes que no se han fugado, por sexo, adulto_mayor, antigüedad, plan. • Cantidad promedio de cargos totales de clientes que no se han fugado, por sexo, adulto_mayor, antigüedad, plan.

Requisito	Perfiles de cliente con más riesgo de fugarse
Descripción	Identificar los perfiles de cliente que tiene más probabilidad de fugarse dado a la gran cantidad histórica de fugados con esas características y así poder enfocarse en disminuir la cantidad de fugados más fácilmente
KPI's	<ul style="list-style-type: none"> • Cantidad de fugados por sexo, adulto_mayor, tipo_contrato, pago_mensual, antigüedad, plan. • Cantidad de personas en ese perfil por sexo, adulto_mayor, tipo_contrato, pago_mensual, antigüedad, plan. • Porcentaje de fugados en relación al perfil por sexo, adulto_mayor, tipo_contrato, pago_mensual, antigüedad, plan.

Requisito	Medios de pagos mas populares
Descripción	Identificar los medios de pago más populares para así hacer descuentos asociados a éstos con la finalidad de que los clientes los prefieran mejorando su experiencia en la empresa y en consecuencia aumenten su probabilidades de quedarse.
KPI's	<ul style="list-style-type: none"> • Cantidad de personas no fugadas por método de pago. • Cantidad total de personas por método de pago.

3 Modelo de datos

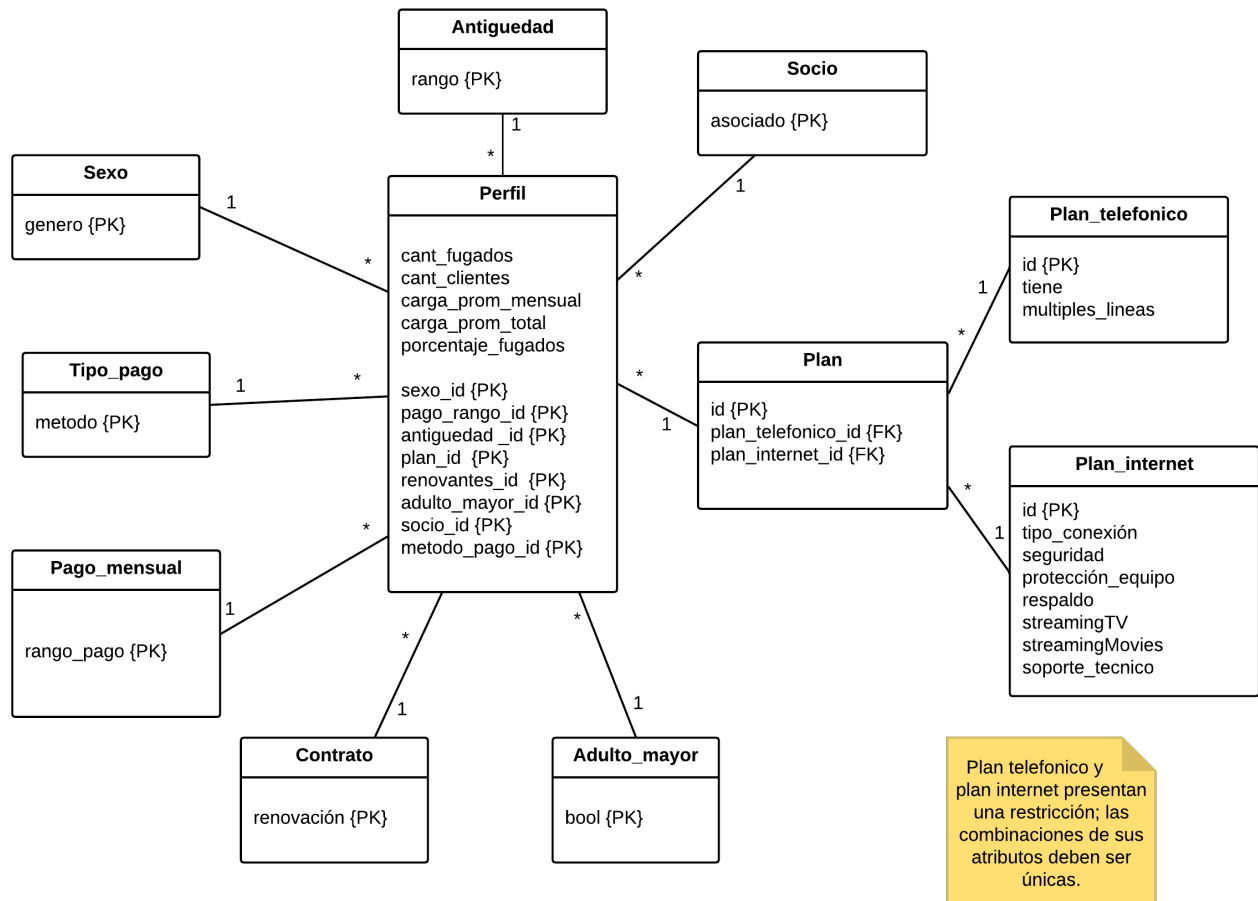


Figure 1: Modelo ROLAP.



4 Tratamiento de datos

4.1 Perfilado

Atributo	Cantidad de Nulos	Valores Distintos	Media	Desviación Estándar	Moda	Valor Mínimo	Valor Máximo
customerID	0	7043	–	–	–	0002-ORFBO	9995-HOTOH
gender	0	2	–	–	Male	Female	Male
SeniorCitizen	0	2	0,16	0,36	0	0	1
Partner	0	2	–	–	No	No	Yes
Dependents	0	2	–	–	No	No	Yes
tenure	0	73	32,37	24,55	1	0	72
PhoneService	0	2	–	–	Yes	No	Yes
MultipleLines	0	3	–	–	No	No	Yes
InternetService	0	3	–	–	Fiber Optic	DSL	No
OnlineSecurity	0	3	–	–	No	No	Yes
OnlineBackup	0	3	–	–	No	No	Yes
DeviceProtection	0	3	–	–	No	No	Yes
TechSupport	0	3	–	–	No	No	Yes
StreamingTV	0	3	–	–	No	No	Yes
StreamingMovies	0	3	–	–	No	No	Yes
Contract	0	3	–	–	Month-to-month	Month-to-month	Two year
PaperlessBilling	0	2	–	–	Yes	No	Yes
PaymentMethod	0	4	–	–	Electronic check	Bank transfer (automatic)	Mailed check
MonthlyCharges	0	1585	64,76	30,09	20,05	18,25	118,75
TotalCharges	0	6531	–	–	20,20	' '	'999.9'
Churn	0	2	–	–	No	No	Yes

Table 1: Perfilado de datos

4.2 ETL

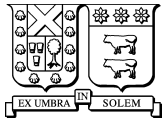
4.2.1 Limpieza

Atributo	Reglas de limpieza
costumerID	Se elimina la columna puesto que no entrega información para el análisis.
Dependents	Se elimina la columna ya que no sé sabe exactamente qué significa, buscando en internet sobre este concepto en el contexto de las telco no se llegó a una respuesta.

Table 2: Reglas de limpieza

4.2.2 Integración

No se aplica integración de datos debido a que la data utilizada para la actividad proviene desde un mismo archivo csv.



4.2.3 Transformación

Atributo	Reglas de transformación
tenure	Se distribuyen los valores en intervalos para obtener rangos de estadía con la compañía
MonthlyCharges	Se distribuyen los valores en intervalos para obtener rangos de ganancias mensuales para la compañía

Table 3: Reglas de transformación

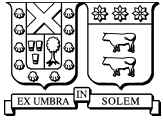
4.2.4 Reducción

No se aplica reducción de datos debido a que no se considera necesario.

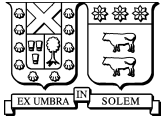
5 Modelo implementado en motor de bases de datos

A continuación se mostrarán los comandos creados para implementar el modelo en postgresSQL

```
CREATE TABLE IF NOT EXISTS sexo (  
    genero VARCHAR(6) PRIMARY KEY  
);  
  
CREATE TABLE IF NOT EXISTS tipo_pago (  
    metodo VARCHAR(30) PRIMARY KEY  
);  
  
CREATE TABLE IF NOT EXISTS pago_mensual (  
    rango_pago VARCHAR(15) PRIMARY KEY  
);  
  
CREATE TABLE IF NOT EXISTS contrato (  
    renovacion VARCHAR(15) PRIMARY KEY  
);  
  
CREATE TABLE IF NOT EXISTS adulto_mayor (  
    bool boolean PRIMARY KEY  
);
```



```
CREATE TABLE IF NOT EXISTS plan_internet (  
    id SERIAL PRIMARY KEY,  
    tipo_conexion VARCHAR(11) NOT NULL,  
    seguridad VARCHAR(19) NOT NULL,  
    proteccion_equipo VARCHAR(19) NOT NULL,  
    respaldo VARCHAR(19) NOT NULL,  
    streaming_tv VARCHAR(19) NOT NULL,  
    streaming_movies VARCHAR(19) NOT NULL,  
    soporte_tecnico VARCHAR(19) NOT NULL,  
    UNIQUE(tipo_conexion, seguridad, proteccion_equipo, respaldo, streaming_tv,  
    streaming_movies, soporte_tecnico)  
);  
  
CREATE TABLE IF NOT EXISTS plan_telefonico (  
    id SERIAL PRIMARY KEY,  
    tiene VARCHAR(3) NOT NULL,  
    multiples_lineas VARCHAR(16) NOT NULL,  
    UNIQUE(tiene, multiples_lineas)  
);  
  
CREATE TABLE IF NOT EXISTS socio (  
    asociado VARCHAR(3) PRIMARY KEY  
);  
  
CREATE TABLE IF NOT EXISTS antiguedad (  
    rango VARCHAR(15) PRIMARY KEY  
);  
  
CREATE TABLE IF NOT EXISTS plan (  
    id SERIAL PRIMARY KEY,  
    plan_telefonico_id int REFERENCES plan_telefonico (id) ON UPDATE CASCADE,  
    plan_internet_id int REFERENCES plan_internet (id) ON UPDATE CASCADE  
);
```



```
CREATE TABLE IF NOT EXISTS perfil (  
    cant_fugados int DEFAULT 0,  
    cant_clientes int DEFAULT 0,  
    carga_prom_mensual real DEFAULT 0.0,  
    carga_prom_total real DEFAULT 0.0,  
    porcentaje_fugados real DEFAULT 0.0,  
  
    sexo_id VARCHAR(6) REFERENCES sexo (genero) ON UPDATE CASCADE,  
    tipo_pago_id VARCHAR(30) REFERENCES tipo_pago (metodo) ON UPDATE CASCADE,  
    pago_mensual_id VARCHAR(15) REFERENCES pago_mensual (rango_pago) ON UPDATE CASCADE,  
    contrato_id VARCHAR(15) REFERENCES contrato (renovacion) ON UPDATE CASCADE,  
    adulto_mayor_id boolean REFERENCES adulto_mayor (bool) ON UPDATE CASCADE,  
    plan_id int REFERENCES plan (id) ON UPDATE CASCADE,  
    socio_id VARCHAR(3) REFERENCES socio (asociado) ON UPDATE CASCADE,  
    antiguedad_id VARCHAR(15) REFERENCES antiguedad (rango) ON UPDATE CASCADE,  
  
    PRIMARY KEY(sexo_id, tipo_pago_id, pago_mensual_id, contrato_id, adulto_mayor_id,  
    plan_id, socio_id, antiguedad_id)  
);
```