

TESIS DE MAESTRÍA

ADMINISTRACIÓN DE NEGOCIOS

**Título:**

**“Desarrollo de estrategias de atención a clientes en la industria de las telecomunicaciones soportada por herramientas de procesamiento y análisis de datos. Caso de estudio”**

Autor: Julián David Quintero

Director de Tesis: Dr. Martín Larrea

Buenos Aires - 2020



## **TESIS**

### **Maestría en Administración de Negocios MBA**

**Desarrollo de estrategias de atención a clientes en la industria de las telecomunicaciones soportada por herramientas de procesamiento y análisis de datos. Caso de estudio.**

**Alumno: Julián David Quintero**

**Director: Dr. Martín Larrea**

**Ciudad Autónoma de Buenos Aires, Junio 2020**



## Dedicatoria

Esta tesis está dedicada principalmente a mi esposa Tatiana Duque y a mi hija Luciana Quintero quienes me acompañaron en este proyecto y me dieron el aliento para sacar adelante esta investigación con su apoyo, cariño y comprensión, y también porque hoy son la principal motivación para alcanzar mis objetivos e iniciativas. A mis padres Clara Inés Ocampo y Gustavo Quintero que me dieron las bases y las herramientas para adelantar todos mis proyectos y a quienes debo todo lo que soy.

Finalmente quiero dedicar esta tesis a todos los profesores y compañeros que tuve durante la Maestría quienes fueron de gran soporte durante esta etapa, especialmente a la Doctora Sandra Patricia Fernández y al Doctor Martín Larrea por orientarme y estar siempre dispuestos a apoyar en la construcción de esta investigación, a ellos muchas gracias.

Julián David Quintero

## ÍNDICE

ÍNDICE DE TABLAS .....	6
ÍNDICE DE ILUSTRACIONES.....	6
1. Título .....	10
2. Introducción .....	10
2.1 Marco de referencia eTOM .....	14
2.1 Planteo del problema: contexto global .....	17
2.2 Contexto específico empresa CableTelco .....	20
2.3 Casos de estudio en la industria.....	23
2.4 Justificación .....	23
3. Objetivos. ....	26
3.1 General.....	26
3.2 Específicos .....	26
4. Estrategia y toma de decisiones en la industria.....	28
4.1 Administración estratégica .....	28
4.2 Integración de intuición y análisis en la toma de decisiones .....	30
4.2.1 Proceso de toma de decisiones.....	31
4.2.2 La tarea de los managers.....	33
4.2.3 Toma de decisiones apoyada en el análisis de datos.....	35
4.3 Ciencia de datos para los negocios. ....	38
4.3.1 Datos y estrategia de negocio .....	40
4.3.2 Definición de preguntas claves .....	42
4.3.3 Dónde encontrar las respuestas a esas preguntas.....	43
4.3.4 Análisis de datos.....	44
4.3.5 Minería de datos .....	47
4.3.6 Presentar y distribuir los hallazgos.....	52
5. Método y entrevistas semiestructuradas. ....	54
6. Desarrollo del caso .....	54
6.1 Conocimiento de negocio.....	54
6.1 Comprensión y preparación de los datos.....	55
6.2 Modelado y evaluación .....	58
6.3 Implementación .....	72
6.3.1 Establecer objetivos .....	74

6.3.2	Políticas de datos.....	76
6.3.3	Creación de una cultura de apoyo a la estrategia de datos.....	76
6.3.4	Estructura y estrategia de Datos.....	78
6.3.5	Plan de acción.....	82
6.4	Seguimiento .....	84
7.	Discusión .....	86
8.	Conclusiones.....	96
9.	Trabajo a futuro .....	98
10.	Bibliografía .....	99
11.	Anexos.....	100
11.1	Anexo A Cuestionario entrevista.....	100
11.2	Anexo B Script de evaluación de vecinos Cercanos Python.....	101

## ÍNDICE DE TABLAS

Tabla 1	Atributos del Set de Datos.....	57
Tabla 2	Tasa de deserción por barrio se listan los primeros 12 .....	63

## ÍNDICE DE ILUSTRACIONES

Ilustración 1	-- Servicios Telecomunicaciones.....	12
Ilustración 2	-- Modelo de Arquitectura empresarial eTOM .....	15
Ilustración 3	-- Ciclo De negocio Telecomunicaciones. ....	17
Ilustración 4	-- Previsión de Crecimiento 2015 - 2021 Mercado de Servicios de Telecomunicaciones.....	18
Ilustración 5	-- Tasa de Churn por Industria EEUU.....	18
Ilustración 6	-- Tasa de Churn por Operador de Telefonía Móvil Brasil.....	19
Ilustración 7	-- Churn AT&T Telefonía Móvil 2007 - 2017 EEUU.....	19
Ilustración 8	-- Churn empresa Sprint Telefonía Móvil 2008 - 2017 EEUU .....	20
Ilustración 9	-- Tablero Comercial Televisión CableTelco Diciembre 2017 .....	21
Ilustración 10	-- Tablero Comercial Internet CableTelco Diciembre 2017 .....	21
Ilustración 11	-- Diagrama de Administración estratégica.....	29
Ilustración 12	-- Modelo de Toma de Decisiones.....	34
Ilustración 13	-- Diagrama de Toma de decisiones Donald Waters .....	36
Ilustración 14	-- Espectro del Análisis de Datos .....	44
Ilustración 15	-- Modelo de predicción de vecindad en Telcos.....	46
Ilustración 16	-- CRISP DM – Cross Industry Estándar Proccess Data Mining .....	48
Ilustración 17	-- Distribución de los clientes por barrio se listan los primeros 20 barrios.....	59
Ilustración 18	-- Gráfico de torta participación de clientes en el mercado por barrio .....	59
Ilustración 19	-- Clientes por género y rango de edad.....	60

Ilustración 20-- Mapa de Calor de servicios contratados.....	60
Ilustración 21 -- Cuadro de Control.....	61
Ilustración 22 -- Diagrama de Barras de cliente en retiro se listan primeros 20 barrios .....	61
Ilustración 23 -- Comparativo de Productos vs Retiros por barrio.....	62
Ilustración 24 -- Mapa de Calor de retiros con filtro constitución.....	64
Ilustración 25 -- cantidad de clientes en Retiro por género y rango de edades.....	65
Ilustración 26 -- Cantidad de retiros por llamadas y Scoring .....	66
Ilustración 27 – K-NN Evaluación de Exactitud dependiendo del número de vecinos .....	69
Ilustración 28 -- Tablero comercial Cabletelco Diciembre 2017 .....	74
Ilustración 29 -- Diagrama de causa y efecto entre factores críticos del negocio .....	75
Ilustración 30 -- Organigrama funcional hasta el nivel de directorio y las gerencias .....	79
Ilustración 31 -- Diagrama de la oficina de Informática y Operaciones.....	80
Ilustración 32 -- Diagramas del área de Sistemas de información.....	81
Ilustración 33 -- Documento Estrategia de Negocio propuesto por Bernard Mars en el libro Data Strategy.....	83
Ilustración 34 -- Tablero de control comercial .....	85
Ilustración 35 -- Marco de referencia Administración estratégica .....	88
Ilustración 36 -- CRISP Data Mining Process .....	88
Ilustración 37 -- Propuesta de Estrategias soportadas por análisis de datos en el marco de Administración estratégica. ....	90

## **Resumen**

Uno de los grandes retos a los que se enfrenta la administración, entendiendo la administración como el manejo óptimo de los recursos en entidades privadas, públicas y organizaciones no gubernamentales es la adopción de nuevas tecnologías y con ellas la modificación o adecuación de los procesos administrativos a éstas.

Este trabajo de tesis busca a través de un caso de estudio encarar el reto de transformación siguiendo una metodología que permita la integración de nuevas herramientas en la toma de decisiones enfocadas en la Retención de clientes en la industria de telecomunicaciones, esto mediante la predicción de clientes con la posibilidad de dar por terminado el contrato, permitiendo generar acciones preventivas e incluso en línea que contribuyan a evitar la cancelación de cuentas.

Durante décadas las compañías y organizaciones estuvieron almacenando información relevante de la compañía; comportamiento de los clientes y la operación, la transformación digital crea la necesidad de analizar esta información para mejorar procesos tanto en tiempo como efectividad la toma de decisiones.

Para esto se centró el desarrollo del trabajo en el estudio de la toma de decisiones bajo el marco de la administración estratégica unido con técnicas de Data Sciences como la metodología Crisp-DM, permitiendo hallar puntos de encuentro entre las dos metodologías y desarrollar las modificaciones necesarias que se deben dar en una compañía para la adopción de nuevas tecnologías y delegar cada vez más la toma de decisiones a procesos en línea y automáticos.

## **Abstrac**

One of the great challenges that the management faces, understanding the administration as the optimal management of resources in private entities, public and non-governmental organizations, is the adoption of new technologies and with them the modification or adaptation of the administrative processes to.

This thesis work seeks through a case study to face the transformation challenge following a methodology that allows the integration of new



tools in decision making focused on customer retention in the telecommunications industry, this through customer prediction with the possibility of terminating the contract, allowing the generation of preventive and even online actions that help avoid account cancellation.

For decades companies and organizations have been storing relevant company information; customer behavior and operation, digital transformation creates the need to analyze this information to improve processes both in time and decision-making effectiveness.

For this, the development of the work was focused on the study of decision-making under the framework of strategic management combined with Data Sciences techniques such as the Crisp-DM methodology, allowing finding meeting points between the two methodologies and developing the necessary modifications that must be given in a company for the adoption of new technologies and increasingly delegate decision-making to online and automatic processes.

### **Palabras Claves**

Churn, Data Sciences, CRISP-DM, Triple Play/Quad Play, PQR , CRM, Framework ,SLAs, Ciencia de Datos, Atención a cliente, Estrategia, Data Strategy, Administración.

## 1. Título

Desarrollo de estrategias de atención a clientes en la industria de las telecomunicaciones soportada por herramientas de procesamiento y análisis de datos. Caso de Estudio.

## 2. Introducción

Los servicios de telecomunicaciones son de gran importancia para las sociedades y son motor del desarrollo económico de los países. La era digital en la que vivimos ha sido modelada por la mano del avance en las telecomunicaciones y el crecimiento de la oferta de servicios digitales. Hoy es posible contar con servicios en la nube; consumir contenidos personalizados, establecer enlaces digitales entre diferentes puntos, ofrecer almacenamiento de información, recibir y transmitir en vivo y otro sin fin deservicios sobre internet.

Con el auge de los sistemas de información y su implementación en los diferentes sectores económicos ha aumentado la necesidad de implementar metodologías que permitan rápidamente analizar gran cantidad de información, y usarla de forma adecuada para la correcta toma de decisiones. Esto exige a las organizaciones de diferentes sectores; industria, gobierno, educación , entre otras contar con soluciones en IT (Informática y Telecomunicaciones), y específicamente herramientas para el procesamiento y análisis de datos, que recolecten, ordenen y administren información para soportar la toma de decisiones y permita el desarrollo de estrategias para mejorar el desempeño y la consecución de los objetivos administrativos.

Para que un proceso de toma de decisiones en la era digital sea efectivo se requiere integrar y apoyar el proceso de dirección sobre una base tecnológica que permita tener disponibilidad, capacidad y velocidad para analizar datos y sobre todo metodologías orientadas a la creación de valor, y de esta forma encaminar la dirección en las compañías hacia los desafíos del siglo XXI.

En mercados saturados como el de los servicios de telecomunicaciones donde existen pocos competidores y el producto ofrecido hace parte de los principales bienes de consumo como telefonía, Internet o televisión, se hace necesario implementar estrategias en la toma de decisiones que permitan la captura de nuevos clientes y evitar la pérdida de los ya existentes.

Uno de los desafíos que se encuentran en esta industria es evitar la pérdida de clientes y que estos sean capturados por sus competidores. A este escenario donde los clientes cambian de un operador a otro se le conoce como Churn y se estima que anualmente mueve el 10% del mercado.

Actualmente las compañías empeñan esfuerzos por evitar la pérdida de clientes y en la consecución de nuevos. Sin embargo, capturar nuevos clientes en este tipo de mercados suele ser más costoso que evitar la retirada de los clientes que se tienen, es por eso que una gran apuesta de las empresas en el sector de las telecomunicaciones es evitar el Churn, diseñando campañas de mercadeo y ofertas especiales enfocadas en la retención.

Una de las prioridades para las áreas de Servicio al Cliente es evitar que el cliente dé por terminado su contrato. En la actualidad se diseñan planes y ofertas especiales de retención con el fin de mantener al cliente y el producto actual. Pese a eso, en la práctica esta acción es reactiva, una vez que el cliente solicita la cancelación de su servicio, el personal de servicio al cliente tiene como tarea evitar que éste se vaya, las alternativas pueden ser ofrecer descuentos por periodos u otorgar servicios adicionales sin cargo por un tiempo limitado, incluso, sacrificando ingresos por la retención.

Permitir que un cliente solicite el retiro del servicio puede significar la pérdida de éste para siempre. Si se quiere tener ventajas sobre la competencia y evitar el Churn, uno de los principales retos es saber quién va a dar por finalizado su contrato tiempo antes de que esto suceda. Anticiparse a esta decisión del cliente y predecir los casos de retiro puede convertirse en valor para la compañía, evitando pérdida de clientes y asegurando ingresos.

Los operadores de servicios de telecomunicaciones conocidos como Telcos, son empresas del sector público o privado que ofrecen los servicios de interconexión entre terceros. Entre su oferta comercial se encuentran los servicios de Telefonía fija y móvil, Internet de banda ancha, televisión y dependiendo del sector del mercado al que se atiende corporativo o residencial existen otras ofertas de servicios especializadas, como almacenamiento de datos, enlaces punto a punto, PBX y otras más.

Para ofrecer estos productos las compañías de Telecomunicaciones cuentan con infraestructura sobre la cual se soportan cada uno de los servicios y para poder acceder a éstos los usuarios cuentan con dispositivos que permitan conectarse a la red del operador, permitiendo de esta manera consumir los servicios ofrecidos.

Adicional a la red física, plataformas tecnológicas, redes de fibra o cobre, inalámbricas o satelitales el operador debe contar con sistemas de información que permitan gestionar la información asociada tanto a la infraestructura como a la atención y gestión de los clientes y los servicio que se prestan.



Ilustración 1 -- Servicios Telecomunicaciones

La demanda creciente de servicios y la dinámica del mercado ha hecho que las compañías de telecomunicaciones integren sus plataformas y puedan ofrecer servicios agrupados o paquetes de servicios. Comercialmente a este tipo de ofertas se les conoce como paquetes o Plays (En Inglés); Triple Play/Quad Play dependiendo de la cantidad de servicios que se agrupan y estos pueden ser Telefonía fija, Telefonía Móvil, Internet y Televisión.

Las compañías de telecomunicaciones gestionan la información de la operación modelando los procesos de negocio con aplicaciones especializadas

conocidas como sistemas de información, que permiten entre otras ofrecer los servicios mencionados y administrar los recursos.

Estos sistemas especializados van desde aplicativos comerciales CRMs (Customer Relation Manager) donde se modela la oferta comercial, se registra la información asociada a; los clientes y los servicios contratados, PQR (Petición, Quejas y Reclamos), sistemas que permiten administrar la fuerza de trabajo, es decir; materiales, equipos y personas que realizan las tareas de campo como instalaciones, mantenimiento y construcción de las redes, hasta aplicativos de gestión de las redes donde se modela la infraestructura y las novedades en ellas, daños, cambios en la infraestructura y nuevas configuraciones.

“La industria de las Telecomunicaciones y de los servicios digitales se está enfrentando a una sin precedente pérdida de clientes (Churn) y a la lucha por seguir siendo rentables de cara a la fuerte competencia, las altas exigencias de los clientes, a la caída del mercado compartido y la creciente presión de precios” (TM Forum , 2015).

La industria de las telecomunicaciones tiene una necesidad de definir y entender los procesos de negocio con el objetivo de enfrentar estas problemáticas y para esto es necesario el consenso de todos los actores del proceso; proveedores de equipos, constructores de aplicaciones e integradores de sistemas de administración, es decir una arquitectura empresarial que permita acoplar dinámicamente estas partes.

La arquitectura empresarial es la organización lógica de procesos de negocio y de infraestructura tecnológica que reflejan el modelo de operación de una organización. Existen marcos de referencia para apoyar a las organizaciones a definir esta arquitectura empresarial, entre estos marcos se encuentra el FrameworkX desarrollado con la colaboración de más de 850 compañías del sector de telecomunicaciones que hacen parte del TeleManagement Forum conocido como TMForum; una asociación global que fomenta la colaboración y la solución de problemas en conjunto entre empresas del sector Telecomunicaciones con el fin de maximizar el éxito de los negocios de operadores de servicios, su ecosistema y la relación con sus proveedores.

El Framework de procesos de negocio (FrameWorkx) desarrollado por el TM Forum define una visión orientada hacia los procesos de las empresas Telcos. Este Framework se puede descomponer en dos caras; una de ellas orientada a negocio, cliente y productos, y otra orientada al interior de la compañía, sistemas y soluciones que permiten soportar los servicios. Este Framework provee una referencia común, consensuada entre compañías, y a través del ITU-T (International Telecommunication Union) un estándar oficial para las compañías del sector.

## **2.1 Marco de referencia eTOM**

El modelo FrameWorkX (eTOM) ofrece un marco de referencia para estudiar los procesos en las compañías de telecomunicaciones y será usado en esta tesis para definir el modelo de negocio de la compañía CableTelco, empresa latinoamericana del sector telecomunicaciones que ofrece los servicios de Telefonía, Televisión e Internet y que cuenta con más de 6.800.000 de servicios activos y alrededor de 5 millones de contratos en gestión. (Fuente interna).

Según el modelo planteado por el TM-Forum una empresa de Telecomunicaciones se puede ver como la interacción entre tres áreas principales de procesos en siete dominios.

Áreas de procesos:

1. Estrategia, Infraestructura y producto: La cual cubre la planeación y el ciclo de vida de administración.
2. Operación: Enfocándose en la administración de la operación de los procesos “Core” de la compañía.
3. Administración de la empresa: Comprendiendo la administración corporativa y la administración de los procesos internos.

Dominios:

Dominio de Ventas y Mercado.

Dominio de Producto.

Dominio de Cliente.

Dominio de Servicio.

Dominio de recursos.

Dominio de Proveedores y Partners (Aleados estratégicos).

Dominio empresarial.

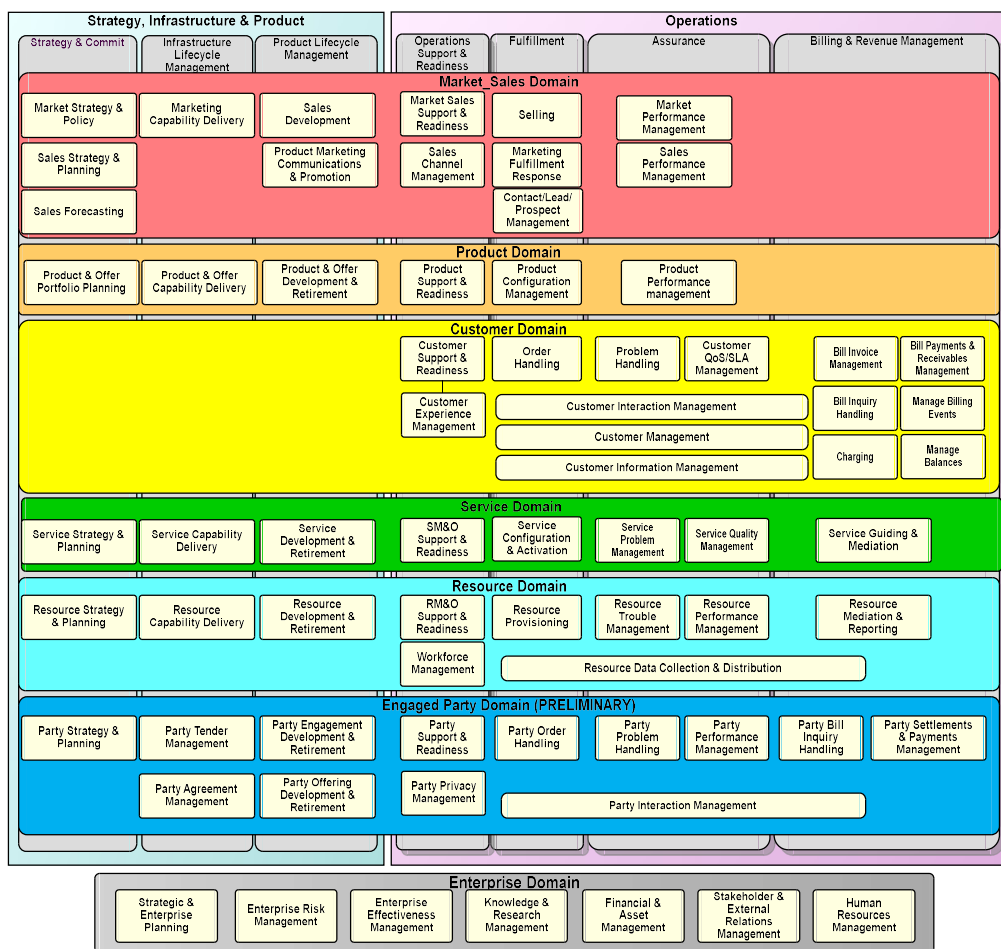


Ilustración 2 -- Modelo de Arquitectura empresarial eTOM

Teniendo en cuenta el modelo (eTom) podemos relacionar las áreas que intervienen en el ciclo de negocios de cualquier empresa de Telecomunicaciones y definir su ubicación dentro del Framework.

El ciclo de negocio de la empresa CableTelco inicia por el diseño de la oferta comercial y el soporte de ésta se desarrolla dentro del modelo eTOM en el

dominio de producto y área de proceso Estrategia, Infraestructura y Producto, y operación. Para desarrollar esta oferta comercial la compañía debe contar con redes e infraestructura que permitan soportar los productos ofrecidos, sistemas de información y aplicaciones de sistemas que permitan diseñar la oferta y gestionarla.

Una vez con los productos a comercializar (Oferta Comercial) la compañía requiere diseñar estrategias de mercadeo y de ventas para que este producto sea ofrecido a los clientes según la segmentación, estimando las capacidades de venta y que permita ser contratado, este proceso se puede enmarcar en el Framework dentro del Dominio de Mercado y ventas.

Para que el servicio contratado por el cliente sea activado es necesario asignar recursos de personal como de infraestructura, para que el usuario cuente con la oferta contratada es necesario desplegar unidades de trabajo al lugar para realizar la instalación, materiales, como del espacio físico en las diferentes plataformas de infraestructura además reservar espacio de almacenamiento, canales y redes. Esta parte del proceso se enmarca dentro del Dominio de Recursos.

Cuando el servicio ya ha sido entregado al cliente hay que poder soportarlo, tener diferentes canales de atención para realizar una gestión adecuada de los servicios prestados, comunicar los cambios, soportar el servicio contratado, poder definir acuerdos de servicio (SLAs), hacer seguimiento a estos acuerdos, facturar el servicio y atender las inquietudes de los usuarios sobre su facturación. En el modelo eTom podemos encontrar estos procesos en el dominio de Cliente.

Los procesos de impresión y distribución de facturas, recaudo y gestión de cobro (Cobranza) se dejan normalmente a proveedores ya que no corresponden a los procesos Core (Procesos centrales o estratégicos) de la empresa, es decir no es la especialidad de una empresa de telecomunicaciones, los proveedores de estos servicios como bancos, empresas de correo, y la forma como se relacionan con el negocio se encuentra en el Dominio de Engaged Party descrito en el modelo eTom.

En términos generales el modelo de negocios de telecomunicaciones enmarca los procesos descritos anteriormente desde el diseño de la oferta, el proceso de venta y entrega del servicio, el soporte y servicio al cliente, facturación y recaudo. Y para esto las compañías de telecomunicaciones acoplan las



áreas con sistemas de información que a través de flujos de trabajo permiten la integración y comunicación de las diferentes áreas.



Ilustración 3 -- Ciclo De negocio Telecomunicaciones.

## 2.1 Planteo del problema: contexto global

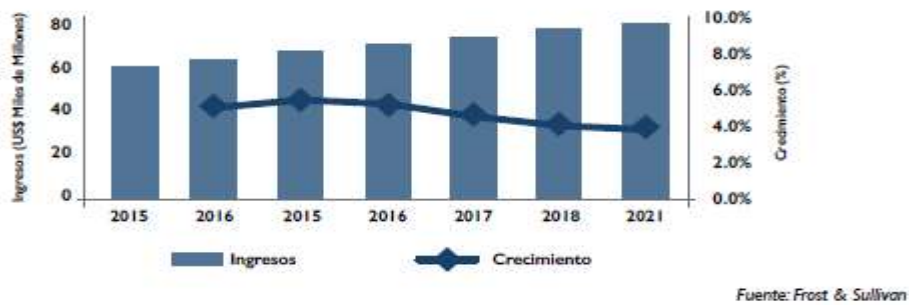
Los operadores de servicios de telecomunicaciones se enfrentan a una larga lista de retos en el mundo actual. La saturación de los principales mercados en los que participan ha provocado el estancamiento comercial tanto a nivel global como en Latinoamérica.

Según un estudio que recoge información sobre la evolución del mercado de las telecomunicaciones elaborado por Frost & Sullivan, las líneas de telefonía fija continuarán relativamente estables en los próximos años en los principales países latinoamericanos, con una tasa de crecimiento anual compuesto (TCAC) que oscila entre el 0 % y el 2 % para el periodo 2014/2020. Los accesos de banda tienen mejores tasas de crecimiento pero muestran una clara tendencia decreciente para los siguientes años (Frost & Sullivan, 2016).

El mercado de las telecomunicaciones solo presenta una TCAC del 4,7 % en el período 2015-2021 y el promedio de ingresos por usuario (ARPU) de todos los servicios mencionados está disminuyendo y continuará haciéndolo. Por consiguiente, el crecimiento de los ingresos del sector y los márgenes de beneficios están cada vez más amenazados (Frost & Sullivan, 2016).

Ilustración 4 -- Previsión de Crecimiento 2015 - 2021 Mercado de Servicios de Telecomunicaciones

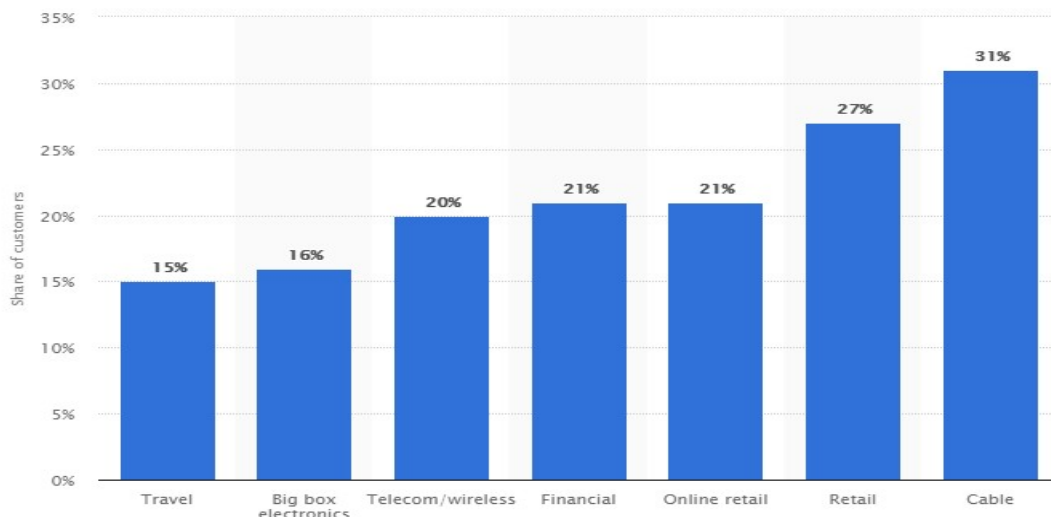
Figure 1: Mercado de servicios de telecomunicaciones móviles- Previsión de crecimiento, Latinoamérica, 2015-2021



Como resultado de la saturación de los principales segmentos del mercado de las telecomunicaciones el aumento de los volúmenes de negocio está relacionado con adquirir clientes de la competencia, y hacer esfuerzos por retener a los clientes existentes, ya que cuesta de seis a siete veces más atraer nuevos clientes que retener los ya existentes.

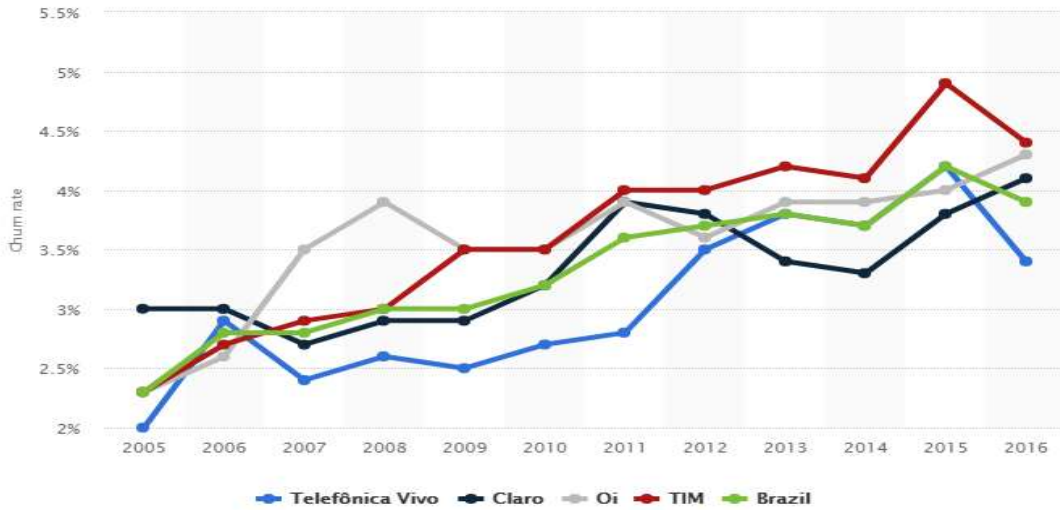
La siguiente gráfica muestra la tasa de Churn por industria en los Estados Unidos en el año 2017 (Estatista, 2018).

Ilustración 5 -- Tasa de Churn por Industria EEUU.



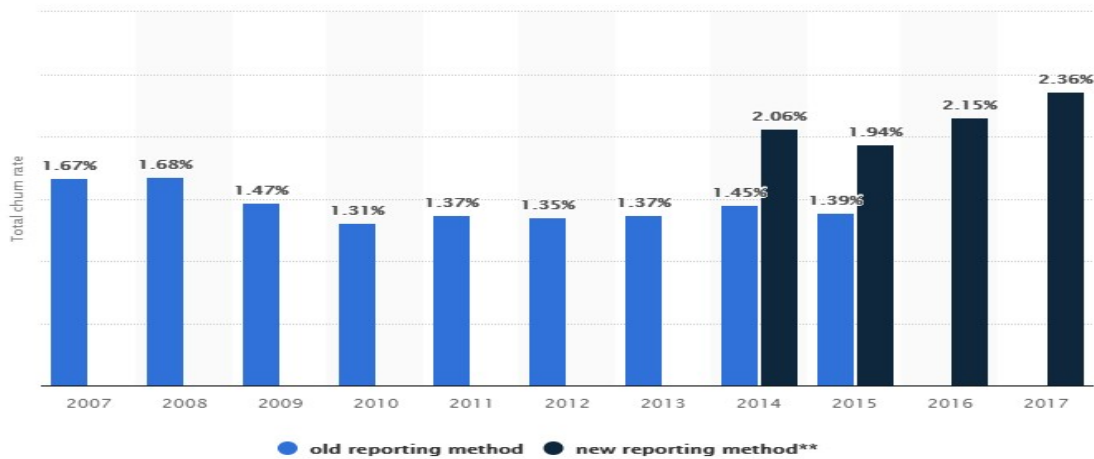
La siguiente estadística muestra la tasa de Churn de operadores de telefonía móvil en Brasil desde 2005 hasta el 2016 (Statista, 2017).

Ilustración 6 -- Tasa de Churn por Operador de Telefonía Móvil Brasil



La siguiente imagen muestra el Churn de AT&T en el mercado de telefonía Móvil entre 2007 y 2017 en los Estados Unidos.

Ilustración 7-- Churn AT&T Telefonía Móvil 2007 - 2017 EEUU.

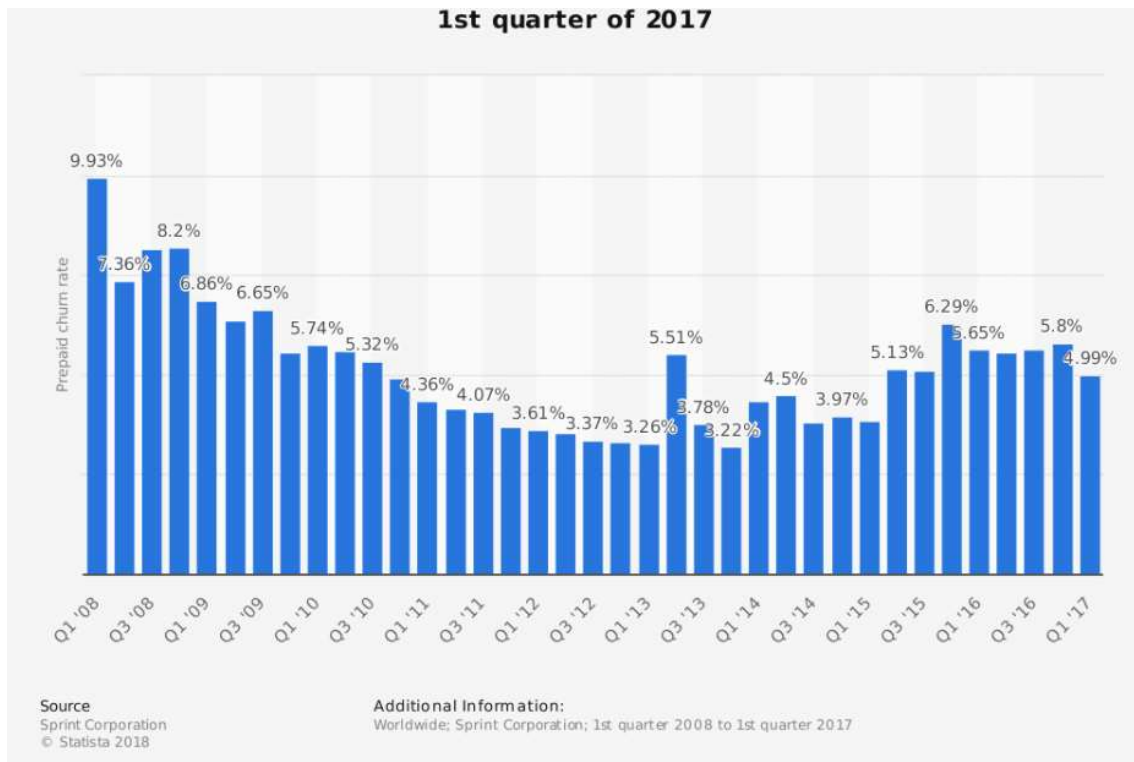


Old Reporting Method: Total de clientes que finalizaron contrato en el periodo sobre el total de cuentas a inicio del mismo periodo.

New Reporting Method: Se excluye las migraciones hacia otras compañías AT&T.

La siguiente imagen muestra la evolución del Churn para la compañía Sprint de telefonía Móvil en los Estados Unidos desde el 2008 hasta el 2017.

Ilustración 8 -- Churn empresa Sprint Telefonía Móvil 2008 - 2017 EEUU



Analizando las imágenes se puede deducir que existe una porción de mercado de alrededor del 4-5% que está trasladándose entre operadores, poder anticipar quienes dejarán el contrato con la compañía puede convertirse un factor diferenciador que permita generar valor. Esto nos indica que es necesario diseñar estrategias de retención anticipada que impidan el retiro de dichos clientes.

## 2.2 Contexto específico empresa CableTelco

La empresa CableTelco es una compañía del sector telecomunicaciones con alrededor de 6'724.800 productos activos al mes de febrero del 2018. En el mes anterior, diciembre, se retiraron alrededor de 146.600 productos lo que corresponde a un 2.18% del total de clientes. Es necesario tener en cuenta que los clientes que nos interesan en el estudio son los clientes que se dan de baja voluntariamente (Churn Voluntario), este 2.18% corresponde al total retiros entre voluntarios y retiros por mora.

Según el Tablero de control comercial para la empresa CableTelco el Churn voluntario para el mes de diciembre de 2018 y del producto Televisión fue de 1.16% con una variación con respecto al mes anterior de 0.0% lo que indica que se mantiene estable el número de bajas voluntarias, lo mismo para los meses anteriores.

Ilustración 9 -- Tablero Comercial Televisión CableTelco Diciembre 2017

	nov-16	dic-16	oct-17	nov-17	dic-17		Var. vs Mes pasado	Var. vs año pasado
	25.0	19.5	23.0	24.5	21.0			
Vtas Ingresadas	2,887	2,875	2,679	2,584	2,637	●	2.0%	● -8.7%
<b>*Vtas Pendientes</b>	<b>19,455</b>	<b>18,948</b>	<b>17,020</b>	<b>20,913</b>	<b>17,279</b>	●	-17.4%	● -11.2%
Altas	2,229	2,281	2,234	1,989	2,321	●	16.7%	● 4.1%
Demanda de Bajas	-4,127	-4,464	-3,849	-3,858	-4,443	●	15.2%	● 7.7%
Bajas Voluntarias	-1,912	-2,192	-2,007	-2,051	-2,419	●	17.9%	● 26.6%
Recuperos Voluntarios	431	472	472	484	530	●	9.5%	● 22.9%
Bajas Netas Voluntarias	-1,480	-1,720	-1,535	-1,567	-1,889	●	20.6%	● 27.6%
Churn voluntario	1.03%	0.97%	1.03%	1.12%	1.16%	●	0.0%	● 0.2%
<b>*Bajas Netas Mora</b>	<b>-11,492</b>	<b>-20,911</b>	<b>-9,610</b>	<b>-11,142</b>	<b>-19,082</b>	●	71.3%	● -8.7%
Churn Mora	0.32%	0.61%	0.28%	-0.33%	-0.56%	●	-0.2%	● -1.2%
<b>*Captura Gestión</b>	<b>7,229</b>	<b>-9,974</b>	<b>6,474</b>	<b>-792</b>	<b>-10,006</b>	●	1163.4%	● 0.3%
*Ajuste	473	-147	88	-145	-323		-	-
<b>*Captura Total</b>	<b>7,678</b>	<b>-10,121</b>	<b>6,562</b>	<b>-937</b>	<b>-10,329</b>	●	1002.3%	● 2.1%

Si se observan los datos para otros productos (Internet) se evidencia el mismo comportamiento.

Ilustración 10 -- Tablero Comercial Internet CableTelco Diciembre 2017

Vtas Ingresadas	3,032	3,146	3,269	3,140	3,218	●	2.5%	● 6.1%
<b>*Vtas Pendientes</b>	<b>19,004</b>	<b>19,553</b>	<b>20,012</b>	<b>24,518</b>	<b>19,998</b>	●	-18.4%	● 5.2%
Altas	2,368	2,483	2,768	2,373	2,823	●	19.0%	● 19.2%
Demanda de Bajas	-3,441	-3,920	-3,601	-3,582	-4,289	●	19.7%	● 24.7%
Bajas Voluntarias	-1,882	-2,222	-2,210	-2,134	-2,514	●	17.8%	● 33.6%
Recuperos Voluntarios	412	497	526	494	554	●	12.3%	● 34.5%
Bajas Netas Voluntarias	-1,470	-1,724	-1,684	-1,640	-1,960	●	19.5%	● 33.3%
Churn voluntario	1.72%	1.55%	1.70%	1.75%	1.78%	●	0.0%	● 0.2%
<b>*Bajas Netas Mora</b>	<b>-4,227</b>	<b>-9,550</b>	<b>-6,053</b>	<b>-5,838</b>	<b>-9,893</b>	●	69.5%	● 3.6%
Churn Mora	0.20%	0.44%	0.27%	-0.25%	-0.43%	●	-0.2%	● -0.9%
<b>*Captura Gestión</b>	<b>18,218</b>	<b>5,253</b>	<b>18,866</b>	<b>12,114</b>	<b>8,225</b>	●	-32.1%	● 56.6%
*Ajuste	-898	-1,754	-977	-350	-2,833		-	-
<b>*Captura Total</b>	<b>17,408</b>	<b>3,499</b>	<b>17,889</b>	<b>11,784</b>	<b>5,392</b>	●	-54.2%	● 54.1%

Open Systems International, empresa de soluciones TICs que atiende el mercado de las telecomunicaciones y Utilities (Servicios domiciliarios) en América Latina, como aliado estratégico de las compañías en las que está implementada solución Open Smartflex, sabe que a partir del conocimiento del negocio que atiende sus problemáticas, el modelo de datos de su solución es posible generar estrategias

preventivas de retención de cliente soportadas por el análisis y de los datos, evitando que sus clientes pierdan suscriptores e ingresos y de esta forma generar valor.

Open Smartflex es una plataforma que permite modelar procesos y registrar información asociada a las operaciones en todo el ciclo de negocio de las empresas de Servicios de telecomunicaciones, soportando desde la atención a clientes CRM (Customer Relation Manager), procesos operativos y administración de la fuerza de trabajo OSS (Operations Support Systems) hasta procesos relacionados a la generación de cobros, facturación y recaudo BSS (Billing Support Systems).

El Churn, como objeto de estudio, es una problemática de todas las empresas del sector Telecomunicaciones y lo que se pretende demostrar es que al analizar el modelo de datos de Open Smartflex (Minería de datos) se puede encontrar características similares entre los clientes que dan por terminado el contrato y de esta forma con análisis predictivo tomar decisiones de forma anticipada y evitar el retiro de estos. Esta tesis se soportará por investigaciones realizadas, técnicas, herramientas y metodologías adoptadas por la industria con este fin.

La mayor parte de la pérdida de clientes en el negocio de las telecomunicaciones está asociada a la insatisfacción debido a la calidad del servicio y la deficiente respuesta del servicio al cliente. Para dar solución a esta problemática se propone adoptar estrategias de retención apoyadas en análisis de datos, enfocándose en el análisis del segmento de suscriptores con mayores niveles de insatisfacción y con la mayor probabilidad de dar de baja sus contratos.

Para este fin se propone analizar el modelo de datos de Open Smartflex asociado a historial comercial, historial de problemas técnicos (Daños), la experiencia con el servicio al cliente y el soporte técnico para de esta forma detectar los clientes con probabilidad de retiro y así generar posibles acciones preventivas como: predecir la oferta comercial de retención, promociones anticipadas, mejorar la calidad del servicio por zonas geográficas o localidad, mejorar la percepción en la atención a clientes (Levantamiento de alarmas de atención especial), mejorar la experiencia con el servicio técnico, entre otras.

## **2.3 Casos de estudio en la industria**

Movistar, líder del mercado de las telecomunicaciones en América Latina, después de implementar un sistema de análisis de datos para la retención de clientes obtuvo una mejora concreta sobre satisfacción en canales de venta de 2%, disminución de la tasa de abandono en 29%, y una mejora del 3% de la satisfacción general (Predictive Analytics Experience, 2016).

Después de 15 días de haber implementado estrategias preventivas de retención mediante Data Analytics en Cablevisión se logró revertir la opinión del 70% de los clientes insatisfechos. Con respecto al total de clientes se logró mejorar la calificación del 50%, lo que representa 1.15 millones de suscriptores satisfechos, dando como resultado un aumento de clientes que recomendarían el servicio, reducción del Churn e incrementando la lealtad (IBM Offering Information, 2015).

La retención de clientes y la detección de posibles desertores es un problema para resolver en cualquier tipo de industria, en el caso de la editorial El Tiempo (Casa editorial Colombiana) por medio del uso de herramientas para el análisis de datos y modelos predictivos se logró la reducción de la deserción de los clientes entre 3 y 10 puntos porcentuales, y se ha conseguido mantener a clientes fieles y generando así crecimiento de lectores (Predictive Analytics Experience, 2016).

## **2.4 Justificación**

En el mundo se generan diariamente 2,5 quintillones de Bytes y cada dos días se genera más datos de los que existían antes del 2003. El 90% de los datos generados se ha producido en los últimos 3 años. Se estima que para el 2018 se generarán 50,000 Giga Bytes por segundo y para el 2020 la cantidad de información disponible habrá crecido de 5 Zetabytes a 50 Zetabytes (Bernard Marr, 2017 ).

“Tras el inicio del tercer milenio, las compañías tienen acceso a la información de una forma sin precedentes. Los datos, que se generan a través de un sin número de aplicaciones y de sistemas que sirven como soporte a los procesos de negocio, las transacciones comerciales y las redes sociales, entre otros, se encuentran no sólo en formatos estructurados y en bases de datos tradicionales, sino también en forma de imágenes, voz, posicionamiento geográfico, etc. La información está disponible en

tiempo real, lo que significa que las organizaciones pueden acceder a ella tan pronto como se genera” (Salvador, 2014).

Las fuentes son cada vez más diversas; redes sociales, correos personales y corporativos, datos de navegación en internet, búsquedas, se calcula que cada segundo se envían 2,9 millones de mails en el mundo, se obtienen datos de sensores, sistemas de posicionamiento, equipos de navegación, estaciones meteorológicas, dispositivos médicos, consumo electrónico. Todo esto a gran velocidad. (Ben Walker, 2013).

La inversión en infraestructura TICs, en los últimos años, ha desarrollado e incrementado el uso de sistemas transaccionales de información como: ERPs, CRMs, Supply Chain Management, Marketing Campaign, Workflow Procedures Frameworks, Work Force Management, entre otros. Estas herramientas permiten soportar la operación de las diferentes industrias, y registran todos los movimientos y novedades de los diferentes procesos de negocio ya sean ventas, recaudo, logística e inventarios, cadena de producción, facturación, atención a clientes entre otros.

Todo este volumen de información hace que las empresas estén ricas en datos, sin embargo, se tienen deficiencias en la generación de conocimiento valioso a partir del análisis de ellos. Generar conocimiento del negocio y construir valor a partir de la información registrada es una necesidad y objeto de esfuerzo para las empresas pero teniendo en cuenta que los datos carecen de valor por sí solos, éstos necesitan ser ordenados, procesados y analizados para que cumplan con el fin de generar valor.

Con el fin de mejorar el rendimiento de las organizaciones, hacer productiva esta información y generar valor es necesaria la implementación de procesos de administración estratégica apoyado sobre herramientas de análisis de datos que permitan construir conocimiento útil del negocio, que soporte la toma de decisiones, permita aumentar ganancias y disminuir costos, mejorar el performance de la compañía.

“Para la implementación de un modelo de administración estratégica soportada por el análisis de datos y herramientas de Data Analytics Las compañías deben dejar de lado la prevalente toma de decisiones basada en la intuición,



aprender a analizar la información de forma confiable y crear estructuras y roles nuevos que posiblemente afecten el balance de poder actual. En pocas palabras, deben superar las poderosas fuentes de inercia que impiden la innovación” (Salvador, 2014).

Hoy se dice que los datos son el oro del siglo XXI y esto genera una necesidad en las compañías, la cual se convierte en el objeto de estudio de esta Tesis, donde se pretende mostrar cómo el uso de herramientas para el procesamiento y análisis de datos ayuda a impulsar estrategias de negocio y generar valor dentro de las compañías.

Para soportar este estudio se desarrollará el caso de generación de estrategias de retención a partir del análisis de Datos, analizando el modelo de datos de Open Smartflex, con los que se soportan el negocio de telecomunicaciones y se demostrará qué al integrar herramientas de análisis de datos sobre el comportamiento de sus clientes es posible generar estrategias de retención.

En consecuencia con el objetivo de esta tesis se pretende identificar los retos asociados a la implementación de un modelo de toma de decisiones soportado por herramientas de análisis de datos, identificar patrones que permitan definir una metodología de implementación de una estrategia de datos aplicable a diferentes sectores económicos, y que adicional, permitan la generación de valor.

### **3. Objetivos.**

#### **3.1 General**

Mostrar la relación entre la implementación de herramientas para el procesamiento y análisis de datos y las estrategias de retención de clientes en la industria de las telecomunicaciones, sector residencial.

#### **3.2 Específicos**

- Identificar los retos organizacionales asociados a la implementación de un modelo de Data Science orientado a la generación de estrategias de retención de clientes en el sector Telecomunicaciones.
- Identificar patrones que permitan definir una metodología de implementación de Data Science como desarrollador de estrategias de negocio aplicable a la retención de clientes.



## **4. Estrategia y toma de decisiones en la industria**

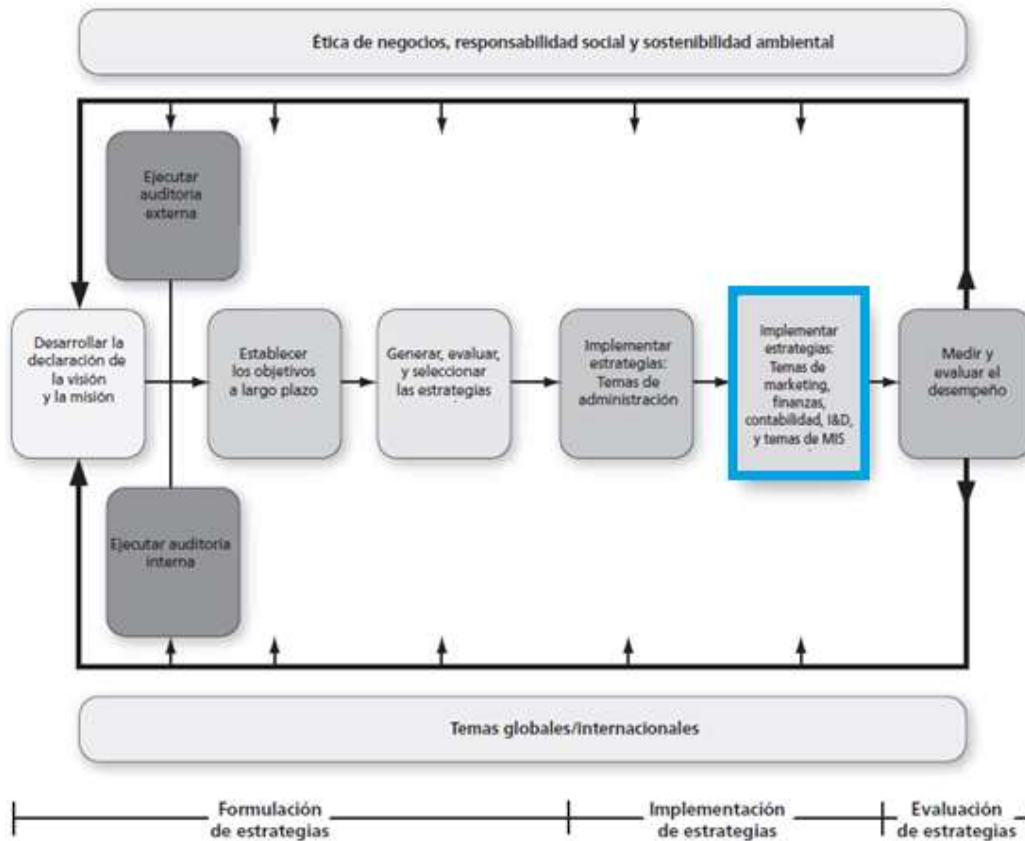
### **4.1 Administración estratégica**

El proceso de administración estratégica se define como el método objetivo, lógico y sistemático para la efectiva toma de decisiones. Su finalidad es organizar información cuantitativa y cualitativa para la elaboración, implementación y control de estrategias enfocadas a la consecución de metas u objetivos de las organizaciones (David, 2013).

La administración estratégica es una tarea multidisciplinaria que requiere la intervención de todas las áreas de la compañía, por lo que en éste se debe integrar todos los procesos que conforman el ciclo de negocio; desde la dirección, el marketing, las finanzas y la contabilidad, la producción, las operaciones, la investigación y el desarrollo, hasta los sistemas de información, con el fin de lograr mejores rendimientos y así el éxito de las empresas y negocios.

El proceso de Administración Estratégica está dividido en tres etapas; planeación, etapa en la cual se formula el plan estratégico, implementación, etapa en la que se ejecutan los puntos del plan y finalmente evaluación, donde se mide y se comparan los resultados obtenidos después de la implementación (David Fred R, 2013).

Ilustración 11 -- Diagrama de Administración estratégica.



Con el auge de los sistemas de información y su implementación en los diferentes sectores de la economía: Industria, gobierno, educación, ha aumentado la necesidad de implementar estrategias de analítica de negocios y metodologías que permitan rápidamente analizar gran cantidad de información, y usarla de forma adecuada para la correcta toma de decisiones.

Esto exige a las compañías contar con soluciones en IT (Informática y telecomunicaciones), específicamente herramientas para el procesamiento y análisis de datos que recolecten, ordenen y administren información valiosa y que permitan soportar el diseño de estrategias para mejorar el desempeño de los negocios y la consecución de los objetivos administrativos.

Para que un proceso de toma de decisiones en la era digital sea efectivo se requiere integrar y apoyar el conocimiento de negocio (Business Knowledge) sobre una base tecnológica que permita tener disponibilidad, capacidad y velocidad para analizar datos y sobre todo metodologías orientadas a la creación de valor, y de esta forma encaminar la dirección en las compañías hacia los desafíos del siglo XXI.

Las empresas capaces de obtener, analizar y evaluar con mayor eficiencia la información externa e interna son las que están generando ventajas competitivas. Como estrategia corporativa soportar decisiones con sistemas de información, herramientas de data Analytics puede ser un factor diferencial que permita la supervivencia de las compañías.

El proceso de la administración estratégica es más efectivo si las empresas cuentan con sistemas de información y análisis de datos que permitan; Obtener, almacenar y analizar información de diferentes fuentes, seguir las ventas cruzadas a los clientes, monitorear los proveedores, mantener informados a gerentes y empleados, coordinar actividades entre divisiones y administrar fondos.

Al igual que el inventario y los recursos humanos, ahora se reconoce que la información es un recurso organizacional valioso que se puede controlar y administrar. Las empresas que implementan estrategias y hacen uso de la mejor información serán las que logren ventajas competitivas en el siglo XXI (David Fred R, 2013).

#### **4.2 Integración de intuición y análisis en la toma de decisiones**

El proceso de la administración estratégica como un método objetivo, lógico y sistemático para tomar decisiones hace necesario organizar la información cualitativa y cuantitativa de manera que sea posible tomar decisiones efectivas en condiciones de incertidumbre. Sin embargo, la administración estratégica no es una ciencia pura que se adapte a un modelo sencillo o sistemático.

Para tomar buenas decisiones la mayoría de las personas reconocen que la intuición es esencial. La intuición se considera la suma de experiencias pasadas, juicios y sentimientos, que permite tomar decisiones de forma inmediata sin intervención aparente de la razón, es muy útil para tomar decisiones, y en particular en situaciones de gran incertidumbre o con pocos precedentes.

Los acelerados cambios actuales y el volumen de información que se produce están generando un mundo de negocios en el que los hábitos administrativos de las empresas resultan cada vez más inadecuados. La expertís del director era una guía adecuada cuando los cambios eran paulatinos. Sin embargo, las filosofías de administración intuitiva y basada en la experiencia resultan obsoletas cuando los cambios tecnológicos y el flujo de información no alcanzan a ser percibido.

Hoy más que nunca en el entorno empresarial la única constante es el cambio. Las empresas exitosas sortean los cambios con gran efectividad, para esto, adaptan continuamente sus estructuras, estrategias, sistemas, productos y culturas con el fin de sobrevivir a los impactos y prosperar.

#### **4.2.1 Proceso de toma de decisiones**

La primera teoría general sobre el proceso de toma de decisión fue enunciada por el filósofo francés Condorcet (1743-1794). Él explica el proceso de decisión en tres etapas.

En la primera etapa se contemplan opiniones propias sobre los principios que servirán de base para la toma de la decisión, se examinan varios aspectos del problema y las consecuencias de tomar alguna de las posibles decisiones. En la segunda etapa se hace claridad sobre el problema o situación a resolver, se exteriorizan las opiniones y se ponderan frente a diferentes opiniones, al final de esta etapa quedará un grupo reducido de opciones con el fin de tener un pequeño grupo de alternativas. En la tercera etapa consiste en evaluar y elegir entre las posibles alternativas resultantes de las dos etapas anteriores.

El punto de partida en la discusión moderna sobre el proceso de toma de decisiones es basado en las teorías expuesta por John Dewey's (1910) donde define etapas en la solución de problemas. El modelo de Dewey consiste en 5 etapas consecutivas: Una dificultad evidenciada, la definición de esa dificultad, sugerir posibles soluciones, evaluar las posibles soluciones y posterior a la observación y experimentación, aceptar o rechazar las sugerencias.

Hebert Simon (1964) modifica el planteamiento de Dewey para asimilarlo con los procesos organizacionales. Simon plantea que el proceso de toma decisiones consiste tres fases principales, donde encontrar la ocasión para tomar acciones,

determinar posibles cursos de acción, elegir entre una cantidad de cursos de acción es la primera fase, la cual llama inteligencia, la segunda etapa la llama diseño y la última etapa selección (Hansson, 2005).

Otra teoría influyente fue propuesta por Brim y Hebert Simon (1962), modificando el anterior planteamiento, dividen el proceso de decisión en los siguientes 5 pasos:

1. Identificar el problema.
2. Obtener la información necesaria.
3. Producción de posibles soluciones.
4. Evaluar las posibles soluciones.
5. Seleccionar la estrategia para implementación.

Más adelante, el psicólogo Leon Mann, Ros Harmoni y Colin Power (1983) desarrollaron el proceso de toma de decisiones llamado GOFER, éste ha servido de marco teórico de referencia para enseñar en escuelas y universidades alrededor del mundo, su nombre se debe al acrónimo en inglés de lo que en conjunto pensaron eran los 5 pasos en la toma de decisiones:

1. Clasificación de objetivos.
2. Considerar un amplio rango de cursos de acción.
3. Buscar información que apoye la toma de decisión.
4. Considerar los posibles efectos de las decisiones tomadas.
5. Planear como implementar la decisión.

Mann y su equipo plantean que la actividad de toma de decisiones involucra procesos cognitivos como búsqueda de información y evaluación, juicios y solución de problemas, adicional plantean que éste proceso también responde a un conjunto de fuerzas motivacionales que determinan la forma en que se toma la decisión.



Las componentes emocionales que podrían intervenir en la actividad de toma de decisiones pueden ser de resistencia a tomar la decisión, presiones sociales, estrés psicológico, tendencias a no tomar decisiones o procrastinación, idealizar algunas alternativas y delegar o no tomar otras decisiones. En el estudio de la toma de decisiones no se puede prescindir de los factores emocionales que hacen parte del proceso. (Jonathan Baron, Rex V. Brown, 1991).

#### **4.2.2 La tarea de los managers**

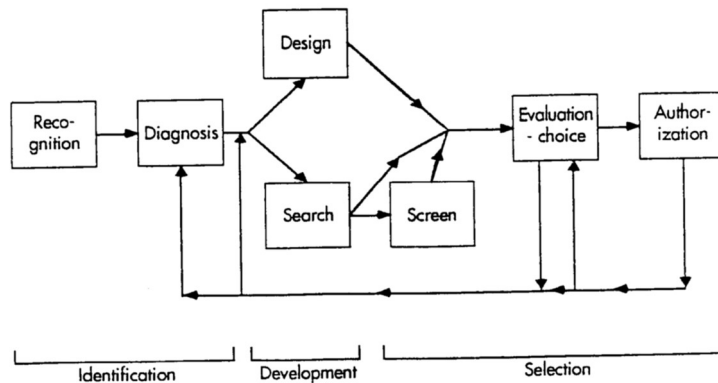
Los directores tienen muchas responsabilidades pero estas se pueden generalizar en analizar y dar solución a los problemas, lo que implica una constante toma de decisiones, un manager debe ser capaz de tomar una decisión y de apoyarse en todos los recursos que tenga a su disposición para hacerlo.

El proceso de toma de decisiones no es un proceso exclusivo de la intuición o de la expertís del gerente o director del negocio. Los responsables de esta tarea deben también soportar sus acciones apoyado en herramientas de procesamiento de información, equipos de comunicación, abogados y otros especialistas técnicos, análisis financiero y presupuestal, encuestas, análisis de mercados, información externa e interna de la compañía, de tal forma que permita reducir los riesgos asociados a la decisión y ésta sea más efectiva.

Uno de los modelos de toma de decisiones más aceptado en la industria fue propuesto por Henry Mintzberg en compañía de Raisnghani y Théoret (1976). Para estos autores el proceso de toma de decisiones sucede en distintas etapas, pero al contrario de los modelos mencionados anteriormente, estas etapas no tienen una relación secuencial simple.

En el modelo intervienen las fases generales propuestas por Simon comentadas anteriormente pero con los siguientes nombres: Identificación, Desarrollo y selección.

Ilustración 12 -- Modelo de Toma de Decisiones



La etapa de identificación se encuentra dos rutinas Reconocimiento y diagnóstico. En el reconocimiento los problemas y oportunidades son identificados mediante información recibida de los agentes del problema. En la rutina de diagnóstico se consulta más información, cerrando los canales de comunicación ya consultados en la primera fase y se abren nuevos canales o espacios con el fin de dar claridad y definir el problema o situación a resolver.

En la etapa de desarrollo se definen y clarifican las opciones. Esta etapa también está constituida por dos rutinas: Diseño y Búsqueda. En esta etapa se identifican posibles soluciones al problema planteado; en la rutina de búsqueda se investiga sobre soluciones ya implementadas que puedan resolver la problemática, y en la rutina de diseño se buscan nuevas soluciones.

La fase de selección, la última del proceso descrito, consiste en tres rutinas. La primera Screen es utilizada cuando a partir de la búsqueda se han encontrado diferentes modelos que pueden ayudar a conseguir el objetivo planteado y necesiten ser evaluados. En esta etapa las soluciones ineficientes ya han sido descartadas.

La rutina de evaluación es la etapa donde se evalúan las alternativas propuestas, esta evaluación se puede dar de tres modos: Juicios, Negociación y análisis. En la tercera y última rutina, Autorización, se requiere una aprobación para ser implementado. Normalmente este tipo de autorización depende del nivel de decisión que se tenga para realizarlo.

El modelo planteado por Minbtzberg y colegas plantea una relación cíclica entre las etapas del proceso y quien toma la decisión (Decission Maker) puede circular dentro de la Identificación para encontrar situaciones o parámetros para ser tenidos en cuenta durante el diseño.

Se puede iterar dentro de los procesos de Búsqueda y Diseño para encontrar mejores soluciones o diseñar una solución combinada al problema planteado, o iterar entre el desarrollo e investigación para mejorar el conocimiento del problema a resolver u objetivo a conseguir. Se puede iterar entre Selección y Desarrollo para conciliar objetivos con alternativas.

La visión del proceso de toma de decisiones como un modelo cíclico permite tener control sobre las diferentes etapas del proceso y adicional permite la mejora continua de la toma decisiones y de los procesos en las compañías.

#### **4.2.3 Toma de decisiones apoyada en el análisis de datos**

El siguiente modelo es el propuesto por Donald Waters en el libro *Quantitative Methods for Business* (Waters, 2011). En él se presentan diferentes etapas del proceso de toma decisiones soportadas por análisis de datos.

El proceso inicia con la identificación del problema, en esta etapa se define el problema en términos de negocio, en la segunda etapa, Análisis, se realiza una valoración tanto cuantitativa como cualitativa, modelamiento y evaluación del modelo. Una vez comprobado y validado el modelo, se toma la decisión contando con la información necesaria y suficiente para soportarla, y por último la implementación.

Posterior a la implementación se requiere también ejecutar control sobre la decisión tomada con el fin de hacer correcciones y mejoras a los modelos. Este proceso es cíclico y está enfocado a ser mejorado continuamente.

Ilustración 13 -- Diagrama de Toma de decisiones Donald Waters

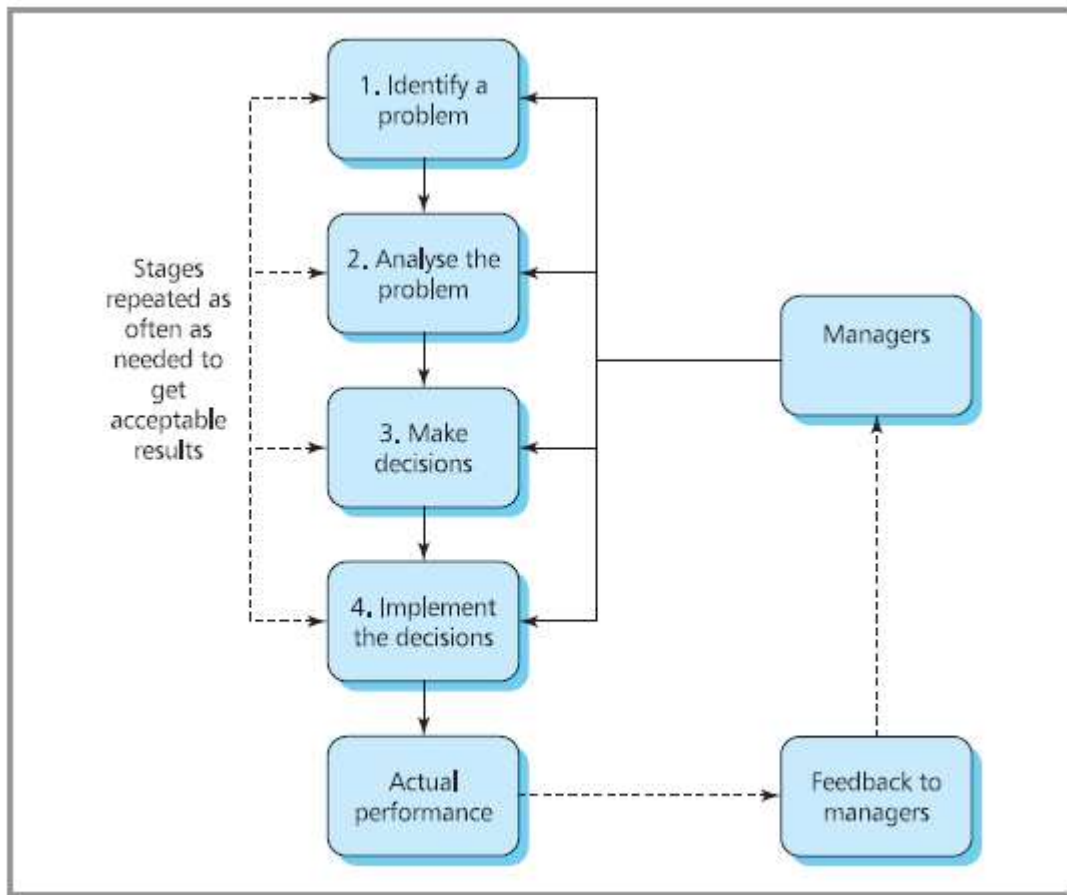


Figure 1.2 Stages in decision-making

Etapa 1: Identificar el problema.

En esta etapa se requiere conseguir el conocimiento necesario sobre el negocio, el contexto e identificar requerimientos o necesidades para resolver el problema al cual se están enfrentando.

- a. Realizar una investigación inicial, conocer los detalles de la operación, identificar dificultades, y reconocer que efectivamente existe un problema a resolver.
- b. Definir el problema, adicionar detalles a la investigación inicial, definir exactamente el problema y no los síntomas, el contexto, alcances y limitaciones.

- c. Establecer objetivos, identificar los responsables en la toma de la decisión, sus objetivos, mejoras que se quieren obtener, identificar los efectos en la organización, y establecer medidas de éxito.
- d. Identificar variables, posibles alternativas y cursos de acción.
- e. Planear el trabajo de tal forma que se muestre como enfrentar el problema, programar actividades, diseñar el proyecto estableciendo tiempos y recursos.

#### Etapa 2: Analizar el problema

Se busca que en esta etapa quien toma la decisión consiga un entendimiento claro del problema, las opciones para resolverlo y las consecuencias.

- a. Considerar diferentes enfoques para la solución del problema.
- b. Investigar trabajos realizados para solucionar problemas similares.
- c. Estudiar a fondo el problema y ajustar los detalles.
- d. Identificar variables claves y la relación entre ellas.
- e. Construir un modelo del problema y realizar pruebas del modelo.
- f. Obtener información necesaria para el modelo.
- g. Realizar más pruebas del modelo con más datos para asegurar que el modelo describe las condiciones reales y está trabajando correctamente.
- h. Experimentar con el modelo para encontrar resultados en diferentes circunstancias y bajo condiciones diferentes.
- i. Analizar los resultados y asegurarse de que el modelo esté funcionando de manera precisa y óptima.

#### Etapa 3: Toma de decisiones.

Este es la etapa en la que se toma la decisión teniendo en cuenta el resultado de la etapa de análisis y de pruebas.

- a. Se comparan todas las soluciones y todos los aspectos de su performance.
- b. Seleccionar las soluciones que más se ajusten a los objetivos corporativos o de las especificaciones de la solución.

- c. Identificar y acordar la mejor solución sobre todas las soluciones.

Etapa 4: Implementación de la solución.

En esta etapa se llevan las ideas del modelo a la práctica.

- a. Verificar que la solución propuesta efectivamente funcione y que exista una mejora en el desempeño de la compañía.
- b. Planear detalles de la implementación.
- c. Cambiar la operación para introducir nuevas formas de hacer las cosas.
- d. Monitorear la operación y comparar el actual desempeño de la solución frente al modelo y si el rendimiento no es el esperado se pasa a realizar nuevamente ajustes a la operación.

Tal como se muestra en los modelos de Henry Mintzberg y Donald Waters respectivamente, el proceso de solución de problemas y toma de decisiones en la industria es un proceso cíclico y requiere de un constante monitoreo y retroalimentación con el fin de evidenciar; situaciones en la implementación, posibles mejoras y determinar otros posibles cursos de acción.

### **4.3 Ciencia de datos para los negocios.**

Un factor diferenciador de las industrias es la toma de decisiones, no importa cuál sea el sector en el que compitan tomar decisiones es la principal actividad, es necesario enfatizar que cada decisión que se toma difiere una de la otra y que son éstas una combinación de diferentes factores, desde objetivos corporativos, deseos, limitaciones hasta motivaciones personales.

Industrias de diferentes sectores han implementado sistemas automáticos para la toma de decisiones. El sector financiero; bancos y aseguradoras, y telecomunicaciones fueron los primeros en adoptar sistemas de análisis de datos y equipos de Data Science para toma de decisiones gracias a su infraestructura y personal técnico especializado en sistemas, esto permitió la exploración de datos a gran escala, creación de aplicaciones y así el modelamiento de procesos, llegando hasta automatizar la toma de decisiones.

Entre los casos más conocidos del uso de Análisis de datos para toma de decisiones y diseño de estrategias comerciales es el sistema de recompensas de Harra's Casinos, en el cual por medio del análisis de los datos de la base de clientes pudieron segmentar y crear planes personalizados dependiendo del nivel de fidelidad del cliente. Este modelo de análisis es conocido como el sistema de puntos o millas y es usado ampliamente en diferentes sectores; Retail, hoteles, aerolíneas, el cual otorga puntos de lealtad y posibilidad de descuentos y mejores condiciones a cambio de fidelidad como cliente. Con el desarrollo de este modelo la línea Harra's Casino pasó de ser un competidor pequeño en los años 90s a ser la cadena de entretenimiento más importante del mundo hoy conocida como Caesar's Entertainment.

Otro caso de éxito en la industria que ejemplifica el caso de análisis de datos para la toma de decisiones fue el desarrollado por Walmart, quienes mediante el análisis del historial de ventas de fechas o eventos, como el paso de un huracán, pudieron predecir los productos a ofrecer y su ubicación dentro de las góndolas, dado el evento.

Target, competidor de Walmart en los Estados Unidos también tiene un caso de éxito. Target basado en el estudio de los hábitos de los clientes, específicamente estudiando la inercia en el momento de la compra, se dieron cuenta que estos hábitos se mantienen a menos de que algo cambie, así desarrollaron un sistema predictivo que permitiera relacionar a clientes con algún atributo que se quiera predecir como padres en espera (Padres en embarazo).

En este caso los padres que están a la espera de un hijo cambian sus hábitos de consumo, modifican la dieta, la ropa, consumen suplementos vitamínicos, reducen el consumo de alcohol, aumentan los productos para cuidado personal entre otros. Estos históricos pueden ser extraídos de las bases de datos para entrenar modelos de predicción del indicador, una vez se logre tener resultados que permitan predecir la variable esperada, o permitan dar respuesta al problema planteado el modelo es implementado con la base de clientes permitiendo diseñar campañas de mercadeo con ofertas personalizadas, atención especial, capturando al cliente y de esta forma estar un paso adelante de la competencia.

La toma de decisiones apoyándose en el uso de análisis predictivo es una forma de abstraer la complejidad de los problemas a resolver dentro de las compañías, permitiendo enfocarse en un grupo de indicadores para correlacionarlos

con cantidad o grado de interés y permitiendo resolver preguntas como ¿quién comprará X producto? ¿Cuál es nivel de lealtad de mis clientes? ¿Quién dejará el contrato con mi compañía? (Provost, Foster; Fawcett, Tom, 2013).

La forma en la que se abordan problemáticas en las que se requiere conocimiento de negocio, manipulación y procesamiento de grandes cantidades de datos y diseño de estrategia es a través de equipos de Data Science que permitan modelar el negocio, obtener y procesar estos datos, e implementar soluciones que se ajusten a esas necesidades.

La ciencia de datos es la práctica de obtener información valiosa a partir de los datos, esto incluye la combinación de herramientas de procesamiento con técnicas que incluyen estadística, bases de datos y almacenamiento, programación, entre otros que permitan dar solución a la problemática.

Una habilidad importante en Data Science (Ciencia de datos) es la habilidad de descomponer un problema en problemas más pequeños o que estos se relacionen y ajusten a soluciones y herramientas conocidas así evitar perder tiempo y recursos y permitiendo dedicar la atención a otras partes del proceso, las partes no automatizadas donde se requiere de la creatividad y la intervención humana.

El caso de estudio diseño de estrategias de retención y prevención del Churn es un problema específico de CableTelco, y puede diferir del caso Churn de otro operador de servicios, pero la problemática de identificar o determinar la probabilidad de que uno de los clientes dé por terminado el contrato a partir del análisis de una gran cantidad de factores es general a todas las compañías de este sector, así que muchas de ellas conforman equipos de Data Science con el fin de atender esta situación.

#### **4.3.1 Datos y estrategia de negocio**

Los Datos se han convertido en un activo clave de los negocios y eje central del éxito de las compañías; entre más avances y digitalización exista mayor es el nivel de atención que debemos poner en los datos para obtener ventajas competitivas y generar valor.

“De acuerdo con el Instituto Internacional de Analítica las compañías que usen analítica de datos para el desarrollo de estrategias y dirección del



negocio para el 2020 tendrán US \$430 billones de dólares en ganancias más que las compañías que no los usen” (Bernard Marr, 2017 ).

En los negocios tener información es poder y las posibilidades que ofrece hoy el Big Data, el acceso a diferentes fuentes de información, la capacidad de almacenamiento y velocidad de análisis era impensable hace unos años, hoy es un factor diferencial entre las empresas y hace que las que no estén en línea con estos avances tecnológicos y no se adapten en la era digital estén destinadas a ser desplazadas por otros competidores y finalmente desaparecer.

Los datos por sí solos no son útiles en términos de negocio, éstos se convierten en útiles cuando existe una necesidad y ayudan a satisfacer esa necesidad, alcanzar los objetivos estratégicos o se convierten en un activo con el cual se genera ingresos.

Es necesario tener en cuenta que para que los datos sean verdaderamente útiles no es necesario la cantidad de datos que se tienen o la cantidad de fuentes, o cuál es el volumen de datos analizado o disponibles para analizar, lo realmente importante en términos de negocio es saber qué es lo que se quiere, hacia dónde se está dirigiendo la compañía, cuáles son los objetivos y cómo esos datos me pueden ayudar a conseguirlos. En síntesis cuando los datos tienen un enfoque desde la estrategia.

“Hay incontables formas en las que los datos pueden ayudar en el éxito de los negocios, pero en términos generales éstos se pueden agrupar en tres categorías: Uso de datos para mejorar la toma de decisiones, uso de datos para mejorar la operación, y el uso de datos como activo de la compañía” (Bernard Marr, 2017 ).

Tomar mejores decisiones es uno de los objetivos de los negocios; en casos en los que se quiere entender el mercado, desarrollar un nuevo producto, incrementar las ganancias, o alcanzar nuevos clientes es necesario tomar más y mejores decisiones y los datos son la principal entrada para este proceso, ya que ayudan a diagnosticar, analizar y así tener los elementos necesarios para tomar decisiones mejor informadas y sustentadas.

Para desarrollar una estrategia basada en datos es necesario identificar las prioridades de la organización y definir cuáles son las preguntas sin resolver por ejemplo: ¿cómo puedo alcanzar un determinado segmento de mercado? ¿Cómo incrementar las ventas en un 5%? ¿Cuál segmento de mercado tiene más

probabilidades de dejar la compañía? Saber cuáles son estas preguntas define las fuentes y los datos correctos a analizar.

Tener una estrategia de datos permite definir las preguntas claves de negocio, priorizarlas, y utilizar eficientemente los recursos y el tiempo.

#### **4.3.2 Definición de preguntas claves**

La definición de preguntas es el proceso inicial de la estrategia y se nutre del conocimiento de negocio de los usuarios claves o power users, normalmente los patrocinadores de los proyectos de estrategia en análisis de datos y responsables de tomar decisiones.

Los encargados de tomar decisiones (Decisions makers) comprenden que un entendimiento profundo del entorno habilita a la compañía para tomar mejores y más inteligentes, logrando acercar estas al conocimiento y separarlas de la intuición.

Una estrategia enfocada en datos debe ser transversal a la compañía permitiendo resolver preguntas de todo el ecosistema, financiero, operacional y mercado. Algunas preguntas pueden ser ¿Son nuestros partners (Terceros) idóneos para conseguir nuestros objetivos? ¿Cuáles son los cuellos de botella? ¿Cómo optimizar la cadena de abastecimiento? ¿Cuál área de negocio está más expuesta al fraude?

Por ejemplo, si se requiere tener mejor entendimiento de los clientes y el mercado en el que se compite es necesario preguntarse: ¿Quiénes son nuestros clientes? ¿Dónde están? ¿Cuáles son sus expectativas y preferencias? Conociendo esto identificar tendencias, entender la competencia y construir una foto completa del mercado permitiendo resolver preguntas como ¿Por qué compran? ¿Cómo prefieren comprar? ¿Qué comprarán próximamente? ¿Qué hace que ellos definan comprar a una compañía que a otra? ¿Cuáles son las tendencias en el mercado en el que competimos? ¿Cómo se comporta la tendencia de demanda del producto y/o servicio que se ofrece? ¿Cuál mercado se debe abandonar y cuándo? ¿Por qué algunos clientes no están comprando? ¿Cómo están cambiando las expectativas y patrones de compra de los clientes? ¿Qué tan bien segmentado tengo el mercado? ¿Quiénes son nuestros competidores? ¿Cuál canal de venta es más eficiente y por qué? ¿Tenemos el mejor precio del mercado?

Para el caso de estudio Churn sería bueno contestar preguntas como ¿Cuántos clientes tienen probabilidad de dejar la compañía? ¿Cómo es percibida la marca en el mercado comparada frente los competidores? ¿Qué tan satisfechos están los clientes con el producto y/o servicio ofrecido? ¿Qué tan buena es la relación comercial con el cliente? ¿Cuál es el historial de compra de servicios? ¿Cuál es el historial de pagos? Responder estas preguntas requiere de un análisis de grandes volúmenes de datos para esto es necesario saber dónde y cómo obtener la información.

### **4.3.3 Dónde encontrar las respuestas a esas preguntas**

Habiendo identificado y ordenado por prioridad las preguntas se debe pasar a la siguiente etapa del proceso que es identificar las fuentes y la recolección de los datos. Existen muchas formas de acceder a datos y estos pueden ser externos o internos, pueden ser datos libres o con costo, pueden estar siendo recolectados o puede ser necesario empezar a recolectarlos, pueden ser estructurados como las bases de datos, hojas de cálculo o no estructurados como los contenidos digitales, audio, chats, imágenes. No existe un mejor tipo de datos que el otro, la selección de la fuente y el formato depende de la pregunta que el negocio quiera resolver.

Desde una mirada estratégica la selección de las fuentes de datos y el conjunto de datos analizar deben ser evaluados desde tres perspectivas; qué tanto pueden ayudar a conseguir los objetivos corporativos, qué tan fácil es el acceso a estos datos y la relación costo efectividad de los datos seleccionados frente al objetivo que se persigue.

De forma general estos datos pueden ser internos y estructurados como la información de los sistemas de información ERPs y CRMs, estructurados y externos como datos demográficos, internos y no estructurados como chats y canales de comunicación mails, y no estructurados y externos como redes sociales entre otros.

“El tipo de dato más usado es el estructurado y representa solo el 20% del total de datos disponibles en el mundo, el 80% restante se encuentra de forma desestructurada” (Bernard Marr, 2017).

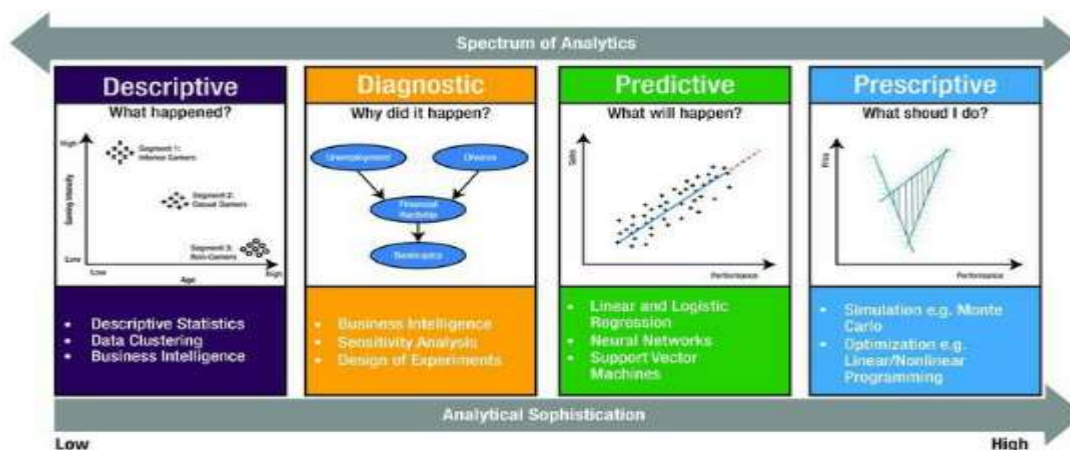
La recolección de datos no es algo nuevo existen diversas fuentes y formas de recolección de datos pero con el paso del tiempo este fenómeno se ha ido incrementando de forma exponencial; ahora generamos información más confiable, con mayor variedad de fuentes y cantidad de datos a una velocidad mucho mayor. A este fenómeno se le conoce como Big Data y a la forma de encontrar hallazgos dentro de estos datos se le conoce como ciencia de datos.

#### 4.3.4 Análisis de datos

Teniendo los datos para resolver las preguntas de negocio el siguiente paso en la estrategia de datos es convertir estos datos en hallazgos que ayuden a conseguir los objetivos planteados, la forma en la que se hace esta transformación es mediante el análisis de los datos lo cual es una combinación de conocimiento del negocio y algoritmos y herramientas que permiten transformar los datos en valor.

La ciencia de datos utilizada en los negocios puede ser clasificada en cuatro categorías de análisis; Descriptivo, Diagnóstico, Predictivo, Prescriptivo. La siguiente gráfica muestra el espectro del Data Science para toma de decisiones y su nivel de sofisticación (Barga, 2014).

Ilustración 14 -- Espectro del Análisis de Datos



Análisis Descriptivo: El análisis Descriptivo es usado para explicar, o como el nombre sugiere, describir que está pasando en una situación dada. Este tipo de análisis involucra intervención humana y puede ser usado para resolver preguntas como ¿Qué pasó? ¿Quiénes son mis clientes? ¿Cuántos tipos de clientes se

tienen? Este tipo de análisis descriptivo involucra Tableros de control, gráficas, histogramas entre otros.

**Análisis Diagnóstico:** El análisis diagnóstico sirve para ayudar a entender que pasó y por qué pasó, y cuáles son los componentes claves que originaron cierta situación. Ayuda a entender fenómenos y las preguntas se centran en analizar el pasado, por ejemplo: Por qué aumentó el número de llamadas al Call Center? ¿Por qué disminuyó el número de ventas?

**Análisis Predictivo:** El análisis predictivo permite anticipar y predecir qué pasará en el futuro. Es usado para predecir la probabilidad de un resultado desconocido evaluado parámetros de entrada sobre condiciones dadas. Este tipo de análisis puede ayudar a predecir una transacción de tarjeta de crédito fraudulenta, si un cliente tiene la probabilidad de dar por finalizado su contrato o si éste podría adquirir o mejorar su plan.

**Análisis prescriptivo:** Este tipo de análisis ayuda a recomendar lo que se debiera hacer ante una entrada dada, este tipo de análisis combina varios tipos de análisis como el predictivo y reglas de negocio con el fin de optimizar un proceso. En el caso del Churn este tipo de análisis ayuda a clasificar los clientes y partir de esa clasificación determinar cuál es la mejor oferta a ofrecer, o el nivel de atención que deba dársele a un cliente, o en procesos operativos como instalaciones, servicios de domicilio o servicios postales, cuál sería la mejor ruta a tomar según las condiciones dadas, por ejemplo, el tráfico.

El nivel de sofisticación incrementa desde análisis descriptivo a análisis prescriptivo, como se muestra en la ilustración 14, cuando el problema incluye mayor nivel de complejidad mayor es el nivel de sofisticación de la solución.

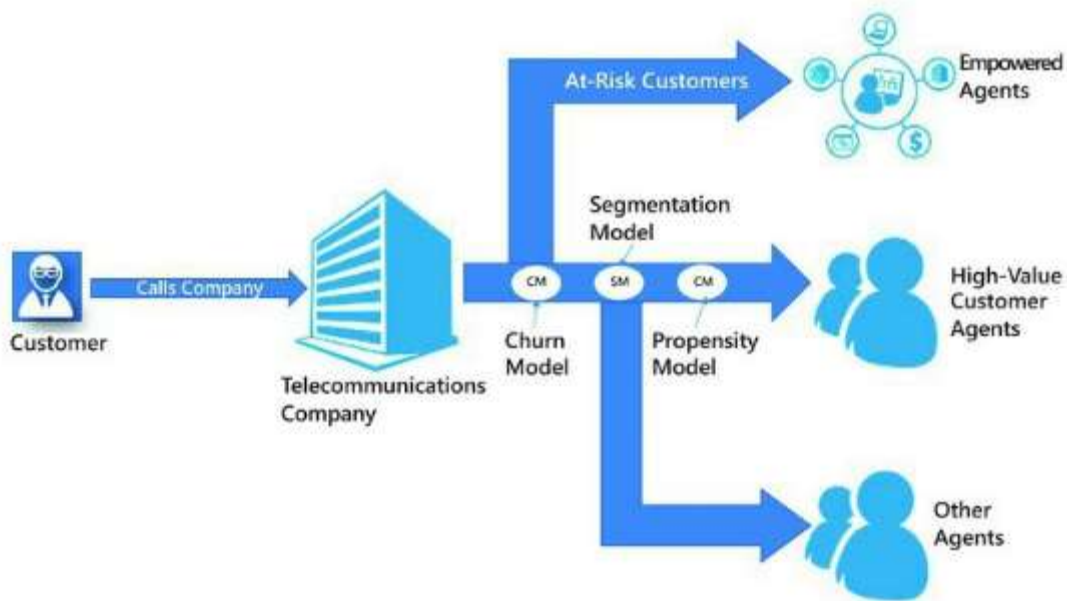
Un ejemplo de un modelo Prescriptivo puede encontrarse en el proceso de atención a clientes en una empresa de telecomunicaciones en este se combina modelos de predicción de vecindad y toma de decisiones, donde dependiendo de la cercanía en probabilidad de pertenecer a ciertos grupos definidos (Vecindad) se direcciona y personaliza la atención del cliente (Barga, 2014).

Las vecindades definidas son: vecindad Churn, clientes con alta probabilidad de dar por terminado el contrato, Vecindad comercial en la cual se

clasifica a los clientes en diferentes segmentos con fines comerciales, Vecindad Clientes de alto Valor es el grupo de clientes con la probabilidad de mejorar sus condiciones en el plan contratado.

En este caso al entrar la llamada el sistema clasifica al cliente en una vecindad según el modelo de predicción de vecindad, una vez clasificado el sistema direcciona la llamada a diferentes grupos de atención; de Retención, de Ofertas especiales o de ventas de nuevos planes. Es la combinación del modelo de predicción junto con la decisión de direccionar la atención del cliente lo que hace que un sistema sea prescriptivo y la toma de decisiones sea automática, asegurando ingresos y mejorando el rendimiento de la compañía.

Ilustración 15 -- Modelo de predicción de vecindad en Telcos



### 4.3.5 Minería de datos

La minería de datos es esencialmente una combinación entre inteligencia artificial, estadística, sistemas de bases de datos, investigación y Machine Learning (Aprendizaje de máquinas), el proceso incluye el análisis automático o semiautomático de grandes sets de datos para encontrar patrones, anomalías y dependencia entre diferentes factores o variables que pueden ser utilizadas para generar valor y predecir el comportamiento del negocio bajo circunstancias antes desconocidas.

De forma general el proceso tiene tres etapas: exploración de datos, construcción de modelos y validación y por último implementación y modificación o inclusión de los hallazgos en los procesos de negocio con el fin de modificar y realizar cambios que contribuyan a mejorar el rendimiento del negocio.

Predecir el comportamiento ayuda no solo a reducir costos, mejorar la operación y generar ganancias también ayuda a conseguir los objetivos corporativos, mejorar la toma de decisiones y disminuir los riesgos.

El proceso de extraer información de los datos para resolver problemas en la industria puede ser tratado como un proceso sistemático y estructurado, con etapas bien definidas. Existe actualmente una metodología estandarizada a través de la industria para desarrollar proyectos de Data Mining (Minería de Datos), es conocido como CRISP-DM (CRISP-DM Project, The Cross Industry Standard Process for Data Mining). Este estándar provee un Framework para estructurar el pensamiento y la solución de problemas que requieran analítica de datos.

El ciclo de vida de un proceso de Data Mining y sus etapas se muestra en la siguiente ilustración:

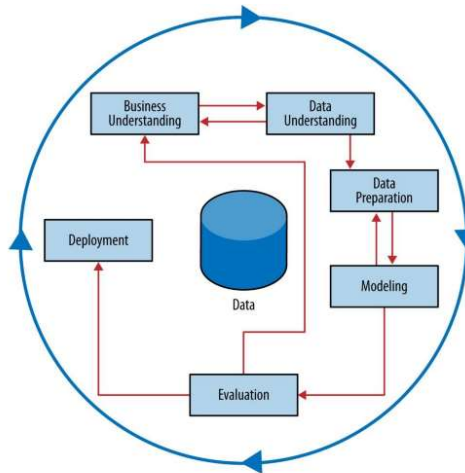


Figure 2-2. The CRISP data mining process.

Ilustración 16 -- CRISP DM – Cross Industry Estándar Procces Data Mining

Etapa I. Conocimiento de Negocio: Definición de necesidades del negocio. Esta etapa inicial se enfoca en la comprensión de los objetivos de proyecto y en las preguntas sin resolver. Después se convierte este conocimiento del negocio en la definición de un problema diseñado para alcanzar los objetivos.

Etapa II. Comprensión de los datos: En la etapa de comprensión de los datos se identifican las fuentes de datos y se define la necesidad de generarlos, recolectarlos, descubrir subconjuntos que permitan generar hipótesis.

Etapa III. Preparación de los datos: La etapa de preparación de datos incluye las actividades necesarias para construir el conjunto final de datos a partir de los datos iniciales. Las tareas de este proceso incluyen la selección de tablas, registros y atributos, así como la transformación y la limpieza para las herramientas que modelan.

Etapa IV Modelado: En esta etapa, se seleccionan y aplican las técnicas de modelado que sean necesarias para resolver el problema y se ajustan los parámetros. En esta etapa se incluye el análisis estadístico y dependiendo del problema se pueden implementar algunas de las siguientes técnicas.

Correlación: La correlación es una técnica estadística que ayuda a determinar la relación existente entre dos variables y qué tan fuerte es esa relación. Este tipo de análisis es usado en datos categorizados como la información demográfica y sirve para resolver preguntas como ¿Qué tan relacionado es que un cliente de por



terminado su contrato dependiendo del sector socioeconómico o por localidad geográfica?

Este análisis resulta en un número entre +1 y -1 si el valor es positivo las dos variables están relacionadas de forma positiva, es decir que cuando aparece una también aparece la otra, si el resultado es negativo quiere decir que cuando aparece la una no aparece la otra, si el resultado tiende a 0 quiere decir que son variables independientes la una de la otra.

Regresión: La regresión es una herramienta estadística para investigar la relación entre variables, por ejemplo cuando existe una relación causal entre precio, producto y demanda. Regresión y Correlación son similares y muchas veces confundidos, la principal diferencia es que la regresión puede mostrar el comportamiento hacia futuro de la relación de las variables mientras que la correlación muestra que tan relacionadas están dos variables. Esta técnica se puede usar con cuando se considera que una variable puede afectar a otra y se quiere comprobar esa hipótesis.

Análisis de escenarios: Análisis de escenarios también conocido como análisis horizontal es un proceso analítico que permite analizar una variedad de eventos futuros considerando escenarios futuros. Este tipo de análisis permite implementar decisiones en el proceso de análisis y ver el comportamiento, medir las consecuencias y estimar riesgos sin implementarlos realmente.

Análisis de Series de tiempo: El análisis de series tiempo es una medición estadística de información que ha sido recolectada en espacios uniformes de tiempo y permite identificar patrones, medir acciones implementadas y predecir el futuro dependiendo de lo que ha pasado en el pasado.

Simulación Monte Carlo: Simulación de Montecarlo es un método de solución de problemas matemáticos y aseguramiento de riesgos que aproxima la probabilidad y el riesgo de ciertos escenarios usando simulación computarizada con el ingreso de valores aleatorios en las variables de entrada. Esta técnica muestra la posibilidad de que algo ocurra.

Esta técnica se usa cuando se quiere un mejor entendimiento de las implicaciones y ramificaciones de un curso de acción, estrategia, plan o decisión y especialmente cuando existe un alto grado de incertidumbre sobre las suposiciones o hipótesis que se quieren comprobar.

**Programación Linear:** También conocida como optimización lineal es un método que permite identificar la mejor salida de un proceso basado en un set de condiciones usando un modelo matemático lineal. Este método permite solucionar problemas en los que se requiere minimizar y maximizar condiciones como ¿cómo maximizar los ingresos y minimizar costos? Este método es ampliamente usado en la cadena de producción donde se quiere determinar los mejores niveles de producción.

**Análisis de Cohorte:** Este tipo de análisis permite estudiar el comportamiento de un grupo segmentado, el grupo o cohorte en este contexto es un sub-set de datos de diferentes frentes de negocio y esta información puede venir de plataformas de comercio electrónico, aplicaciones web, redes sociales, o información de los sistemas de información de ventas (CRM). Los grupos son analizados relacionando características comunes permitiendo extraer información y de esta manera encontrar hallazgos.

**Análisis de factores:** El análisis de factores es el nombre dado a un grupo de técnicas estadísticas que son usadas para reducir la información con la que se trabaja y específicamente el número de variables. En la actualidad las compañías tienen a su disposición un gran volumen de datos tanto propios como externos, existen muchos de estos datos que no ayudan a alcanzar los objetivos corporativos y debilitan o desvían el análisis a realizar. Reducir la cantidad de variables a analizar facilita la detección de estructuras y generación de hallazgos, la identificación de relaciones causales que afectan de forma positiva la estrategia, y mejorando la toma de decisiones.

**Redes neuronales:** Una red neuronal es un programa computacional modelado como el cerebro humano que permite analizar grandes cantidades de información encontrando patrones, aprender y mejora continuamente las habilidades de reconocimiento de la misma manera en como nosotros lo hacemos, ayudando a hacer predicciones que pueden ser probadas e implementadas ayudando a los procesos de toma de decisiones y el desempeño de la compañía.

**Analítica avanzada:** La analítica avanzada es un conjunto de herramientas desarrolladas para el análisis de datos y aprendizaje automático con el fin

de generar información valiosa para el negocio integrándose con los procesos de tomas de decisión mejorando la operación y el desempeño de los negocios.

Este conjunto de herramientas; Deep Learning (Aprendizaje profundo), computación cognitiva, Machine Learning (Aprendizaje de máquina) cada vez son más implementadas en el mundo de los negocios con el fin de mejorar la toma de decisiones. Estas tecnologías por medio de algoritmos de auto aprendizaje aprenden de los datos suministrados y aprenden continuamente, mejorándose. Donde aplicarlas y que tanto delegar responsabilidad en la toma de decisión en las máquinas depende de lo que se quiere conseguir como objetivo

Árboles de clasificación: El proceso de árboles de clasificación es un proceso recursivo de división de la información a clasificar donde el objetivo en cada etapa es seleccionar el atributo que cree particiones del grupo completo dejando subgrupos más puros con respecto a la variable objetivo.

El concepto de pureza es un concepto asociado al análisis estadístico relacionado con la homogeneidad del conjunto de registros a estudiar, entre el grupo sea más homogéneo con respecto a la variable de clasificación, por ejemplo clientes que dieron por cancelado su suscripción, el grupo es más puro. Es decir lo que se busca es tener un grupo que al cumplir con el mismo atributo la variable objetivo sea la misma.

Este tipo de método es recursivo ya que la idea es aplicar la metodología al grupo completo, después a los subgrupos restantes hasta tener una clasificación óptima.

Etapa V Evaluación (obtención de resultados):

En esta etapa en el proyecto, se tienen construidos uno o varios modelos que alcanzan la calidad suficiente desde la perspectiva de análisis de datos. Antes de la implementación es necesario probar y revisar los pasos ejecutados para conseguir el modelo, evaluarlo con respecto a los objetivos de negocio y dependiendo de esta evaluación pasar a implementar estos cambios.

Etapa VI Implementación (puesta en producción):

En la etapa de implementación se debe introducir los cambios o modificaciones en la dinámica del negocio. La implementación no es el final del proceso de estrategia enfocada en análisis de datos, este proceso es continuo y requiere evaluar los cambios, mejorar el modelo y realizar nuevas implementaciones con el fin de que se ajuste cada vez más a las necesidades del negocio.

#### **4.3.6 Presentar y distribuir los hallazgos**

Los datos pueden ser analizados de diferentes formas y una de las formas más simples es creando información visual o gráfica que permita identificar patrones, tendencias y comunicarlas de manera efectiva.

Este proceso integra las ventajas que combinan el análisis de datos, visualización e interacción humana. Esta herramienta es usada cuando se requiere encontrar sentido a gran cantidad de información, encontrar patrones y especialmente cuando se requiere dar información entendible sobre las evidencias encontradas.



## **5. Método y entrevistas semiestructuradas.**

Se adjunta como Anexo el cuestionario de preguntas que se realizó a expertos relacionados con la problemática de Churn y la toma de decisiones basadas en analítica de datos. Para la obtención de la información se utilizará la metodología de entrevistas semiestructuradas que permite realizar preguntas con un objetivo, comprender el proceso de toma de decisiones soportada por analítica de datos y los retos que se tienen al implementarlo, adicionalmente también permitirá conocer detalles o profundizar sobre ciertos temas como metodologías, herramientas y métodos estadísticos, así como tips basados en su experiencia.

## **6. Desarrollo del caso**

### **6.1 Conocimiento de negocio**

Para el caso de estudio de detección de Churn se requiere construir un modelo de probabilidad representado como una función de los atributos de la cuenta del cliente tales como años con la compañía, número de llamadas al servicio técnico, características demográficas, localidad, edad si existe, género, estado civil entre otros.

Como se planteó en el punto 2.3 de este documento: La empresa CableTelco es una compañía del sector telecomunicaciones con alrededor de 6'724.800 productos activos al mes de febrero del 2018. En el mes de enero, se retiraron 20229 productos lo que corresponde a un 2.18% del total de clientes.

Según el Tablero de control comercial (Ver ilustración 9) para la empresa CableTelco el Churn voluntario para el mes de diciembre de 2018 y del producto Televisión fue de 1.16% con una variación con respecto al mes anterior de 0.0% lo que indica que se mantiene estable el número de bajas voluntarias, lo mismo para los meses anteriores. Si se observan los datos para otros productos (Internet) (Ver Ilustración 10) se evidencia este mismo comportamiento.

Una vez comprendida la problemática del negocio la cual indica que existe una parte del conjunto de clientes con alta probabilidad de dejar el contrato y que esto representa pérdidas y disminución de ingresos mes a mes entonces vale la pena seguir adelante y preguntarse quiénes son estos clientes y cómo los identifico.

La definición de preguntas clave para el caso de retención de clientes y poder tomar decisiones informadas deberíamos poder saber:

- ¿Quiénes son estos clientes? ¿Se pueden caracterizar?
- ¿Cuáles son los clientes que dejan mayor ganancias a la compañía?
- ¿Qué diferencia un cliente rentable de un cliente común?
- ¿Cómo se puede segmentar la población de clientes con respecto a la variable que se quiere predecir?
- ¿En el caso específico cuáles son los clientes con alta probabilidad de retirarse?
- ¿Cuál de las variables de caracterización de un cliente proporciona la mayor cantidad de información sobre si un cliente se dará por terminado: ¿La edad? ¿La ocupación? ¿Lugar de residencia? ¿Ingresos? ¿Cantidad de llamados al servicio al cliente? ¿Cantidad de registros de quejas?

## **6.1 Comprensión y preparación de los datos**

La comprensión y preparación de los datos es la etapa del proceso en el que apoyado en el conocimiento de la información y su estructura se prepara para responder las preguntas planteadas por el negocio en la primera etapa y así ayudar en la toma de decisiones.

En el caso de estudio Churn las preguntas planteadas generan los siguientes problemas o situaciones a resolver: ¿Tengo disponible toda la información necesaria? ¿Cuáles son las posibles fuentes de información a usar? ¿Qué tanto de esa información es externa o interna? ¿El costo de conseguirla? ¿Cuál es el formato en que está disponible? ¿Es necesario algún proceso adicional para acondicionarla para el análisis?

Como se mencionó en el capítulo 2 de esta tesis Open Smartflex es una plataforma que permite modelar procesos y registrar información asociada a las operaciones en todo el ciclo de negocio de las empresas de Servicios de telecomunicaciones, soportando la atención a clientes CRM (Customer Relation

Manager), procesos operativos y administración de la fuerza de trabajo OSS (Operations Support Sytems) y procesos de facturación y recaudo BSS (Billing Support Systems).

La fuente de información para analizar el caso será tomada del modelo de datos de Open Smartflex el cual es un aplicativo que permite gestionar información de los clientes y registra esta información en bases de datos relacionales y de forma estructurada. Es decir la información está registrada en Tablas y esas tablas están relacionadas con otras que complementan la información.

Para este caso se analizará el modelo de datos de Smartflex CRM donde se encuentra información comercial del cliente. Para un análisis más profundo se podrían incluir otras fuentes de datos como información de facturación e infraestructura, o fuentes externas como redes sociales o información disponible de centrales de riesgo y comportamientos de crédito.

Es necesario aclarar que la no inclusión de estas fuentes no afecta el objetivo final de esta investigación que es el diseño de estrategias a partir del análisis de datos, definir retos asociados y patrones que permitan definir una metodología para el desarrollo de estrategias a partir del análisis de datos.

El Set de datos a analizar contiene información demográfica del cliente como edad, género y ubicación, así como el comportamiento con respecto al servicio de los últimos meses, esta base de información o muestra se toma específicamente para la ciudad de Buenos Aires Capital Federal y para un solo tipo de producto.

A partir de la información en el CRM se busca establecer un conjunto de variables o asociadas al cliente que permita caracterizar al cliente y responder preguntas asociadas. La información básica del suscriptor está registrada en la Entidad (Tabla) Ge\_Subscriber y a partir de esta tabla y otras tablas relacionadas en la base de datos se construirá el Set de datos a analizar.

El dato único que identificará el cliente es el id de suscriptor, este permitirá tener información asociada al tipo de cliente, la información adicional como Edad y Género está asociada a la entidad ge\_subs\_general\_data, de la tabla ab\_address se obtiene la dirección, la localidad, de la tabla suscripción se obtiene información como



el plan de facturación, el segmento de mercado, y el scoring, que es una puntuación dada por reglas de negocio dependiendo del movimiento de la cuenta en cuanto a consumos, servicios adquiridos, dinámica de pago, y asociada a la entidad de solicitudes se puede encontrar el historial de peticiones registradas; llamadas, daños reportados al servicio, reclamos entre otras.

Según la información recopilada en las entrevistas con los expertos se escogen los datos para armar el Data Set.

Tabla 1 Atributos del Set de Datos

CAMPO	DESCRIPCIÓN
SUBSCRIBER_ID	Identificador interno del Subscriptor o Cliente
EDAD	Edad del cliente
GENERO	Genero del cliente
TIPO_SERVICIO	Tipo de Servicio contratado
PLAN DE FACTURACIÓN	Plan de facturación
TIPO DE SUBSCRIPTOR	Tipo de suscripción
SEGMENTO DE MERCADO	Segmento de mercado al que pertenece, residencial, corporativo, salud, gobierno otros.
TIPO_DE_COBRO	Forma de pago, en casillas, Débito automático
SHAPE (GEO-LOCALIZACIÓN)	Ubicación geográfica, coordenadas (x,y)
BARRIO	Barrio donde se encuentra el cliente
SCORING	Calificación de cliente
ESTADO DE CORTE	Esta variable permite conocer el estado de corte del producto, suspendido, retirado, activo, en proceso de activación o proceso de baja.
CALLS	Llamadas al servicio al cliente en los últimos tres meses.
SERVICES	Daños registrados en los últimos 3 meses
RECLAMOS	Reclamos registrados en los últimos 3 meses
CHURN	Variable objetivo, indica si el cliente dejó el servicio (1) o permanece activo (0)

Se analizará una base de información (Data Set) de la compañía con 70.939 registros de suscriptores del producto Internet localizados en la ciudad de Buenos Aires. Para el análisis y extracción de la información detallada anteriormente se

usaron herramientas de extracción de datos como PL/SQL ya que la base de datos está sobre un motor de Oracle, base relacional de datos en formato SQL, para su transformación se utilizó Python e inserción en la base de datos no relacionada de análisis Elastic Search, base de datos no relacional que ofrece un mejor performance en el análisis de datos.

## **6.2 Modelado y evaluación**

Para lograr una mejor comprensión de los datos y obtener información valiosa de estos es necesario explorarlos con herramientas de visualización que acerquen el objeto de estudio a quien toma decisiones, una cantidad de variables mayor a tres dimensiones dificulta el entendimiento y hace que se tomen decisiones sin contemplar el universo total de variables y sus dependencias.

Para iniciar el análisis exploratorio de los gráficos responderemos a preguntas planteadas como ¿Quiénes son nuestros clientes?, ¿Cuántos y dónde están ubicados? ¿Su género y edad? ¿Su comportamiento con el servicio al cliente? ¿Ha reportado daños o reclamos?

Del conjunto de datos a analizar, especialmente para el caso de estudio, Diseño de estrategias de retención, es necesario saber cuántos de los clientes que se dieron de alta en el último año permanece y cuántos de ellos dieron por terminada la suscripción y así saber en cuál porción del mercado capturado se debe enfocar las estrategias.

Según la información analizada tenemos que un 3% da por retirado su contrato y es en esta porción del mercado que se enfocará el análisis.

¿Cómo están distribuidos demográficamente todos los clientes?

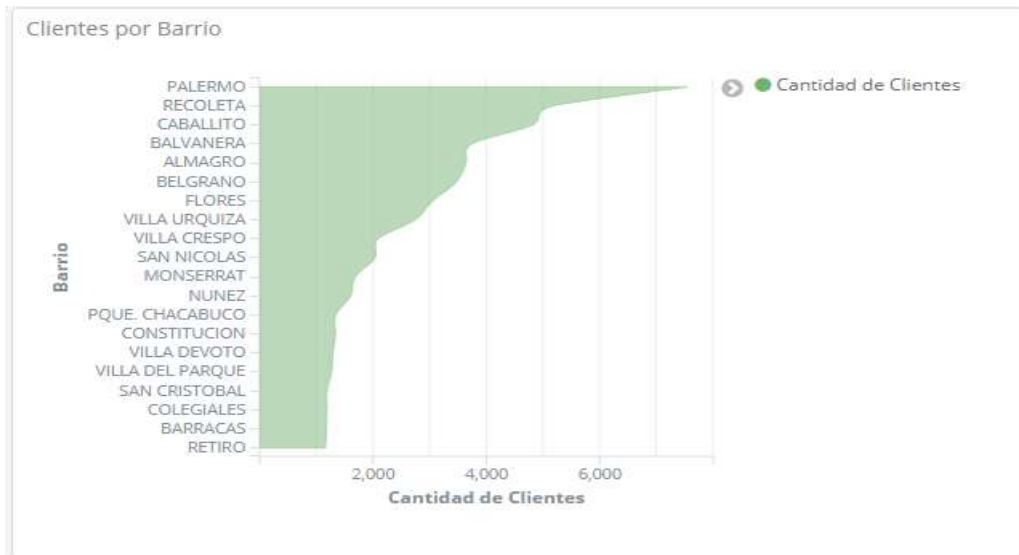


Ilustración 177 -- Distribución de los clientes por barrio se listan los primeros 20 barrios.

¿Y cuánto de la porción del mercado aporta cada ubicación al universo de clientes?

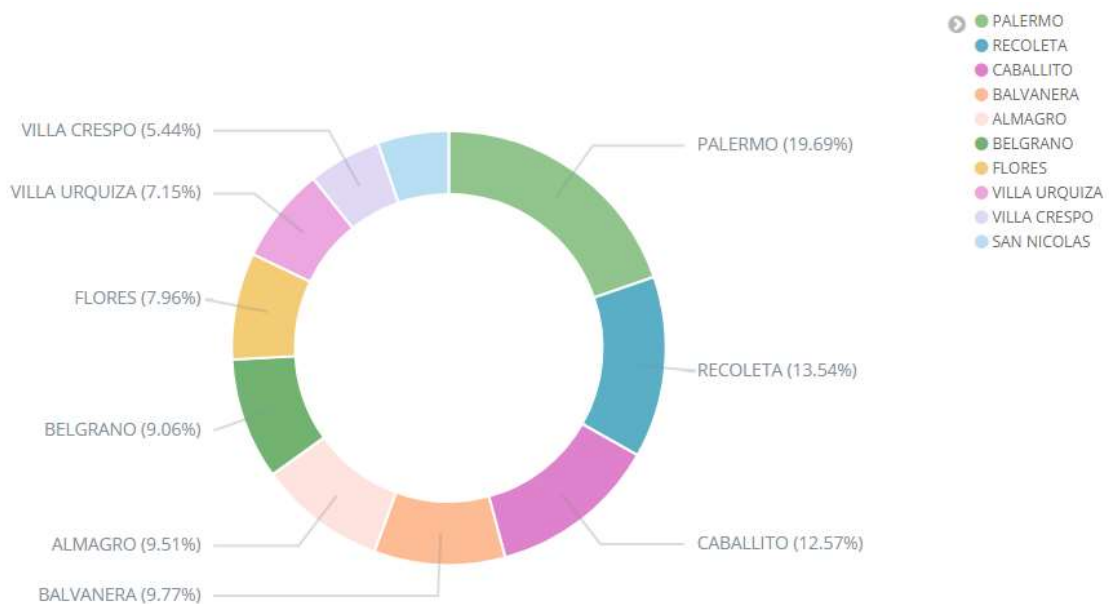


Ilustración 18 -- Gráfico de torta participación de clientes en el mercado por barrio

¿Cuántos de estos son hombres o mujeres y en qué rangos de edades? Esta información puede ser importante en el momento de diseñar una campaña de retención para saber a cuál segmento de mercado direccionarla.

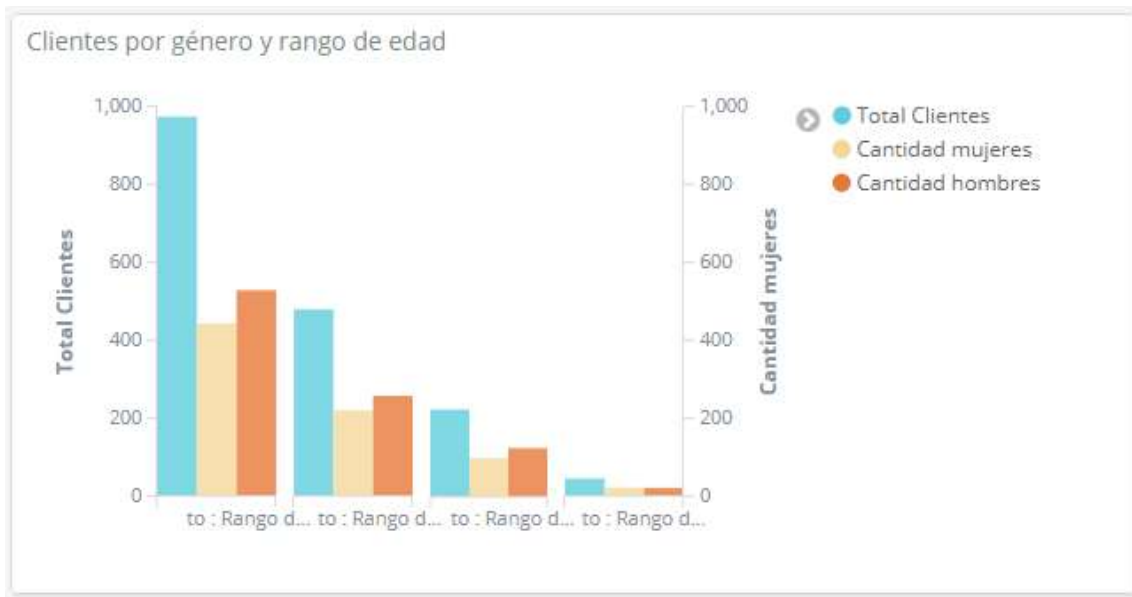


Ilustración 199 -- Clientes por género y rango de edad

También es necesario saber dónde están localizados y cuáles son las áreas de calor, es decir dónde más productos estén contratados.

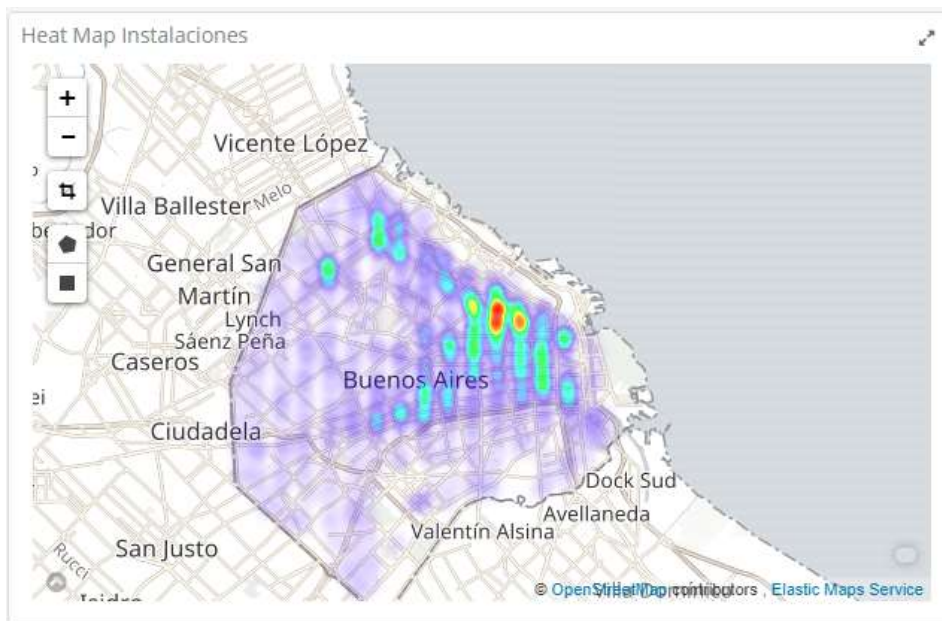


Ilustración 20-- Mapa de Calor de servicios contratados

La siguiente imagen muestra un Dashboard o Tablero de Control en el cual se puede profundizar en la visualización de los datos y explorar cada variable individualmente y cómo ésta afecta a las otras dimensiones:

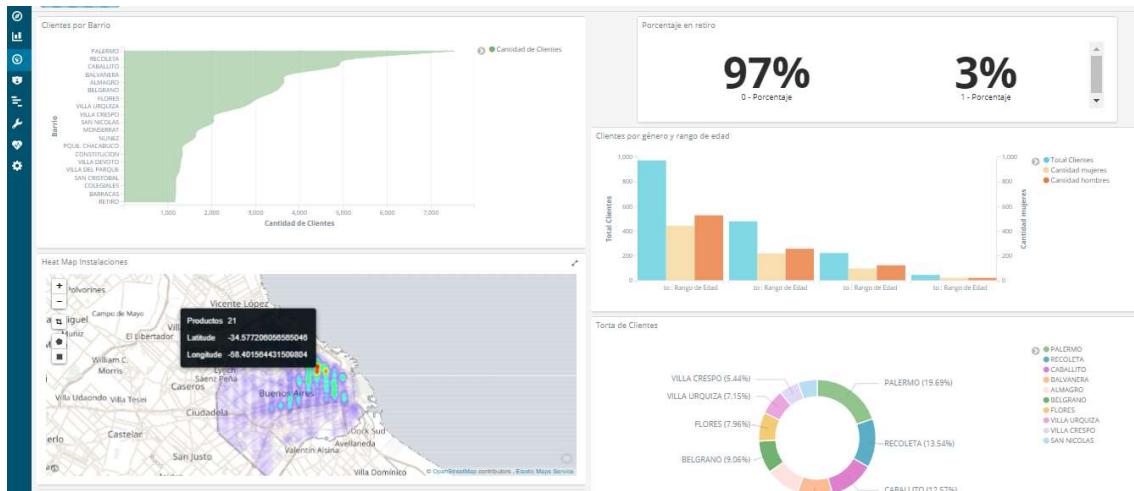


Ilustración 21 -- Cuadro de Control

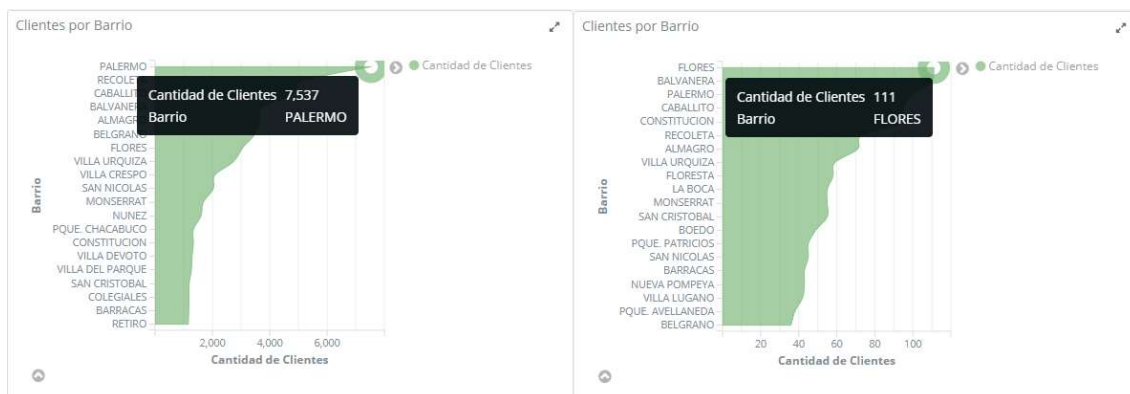
En el caso Retención de Clientes la investigación se centra en quienes dieron por terminado el contrato y se excluyen los que continúan activos y así encontrar patrones que permitan conocer mejor esta porción del mercado.



Ilustración 22 -- Diagrama de Barras de cliente en retiro se listan primeros 20 barrios

La imagen anterior muestra la distribución por barrio de los clientes que se dieron de baja, al contrario que en imagen inicial donde se mostraba todo el universo de clientes y el barrio con más productos era Palermo en esta imagen el barrio con más productos en retiro es el barrio Flores. A primera vista podemos

concluir que el retiro no responde a densidad de productos instalados por barrio pues el barrio con más productos activos no fue el primero con mayor número de retiros.



Barrio	Count
PALERMO	7,438
RECOLETA	5,110
CABALLITO	4,715
BALVANERA	3,630
ALMAGRO	3,567
BELGRANO	3,432
FLORES	2,937
VILLA URQUIZA	2,676
VILLA CRESPO	2,049
SAN NICOLAS	1,984

Barrio	Count
FLORES	111
BALVANERA	110
PALERMO	99
CABALLITO	95
CONSTITUCION	94
RECOLETA	73
ALMAGRO	71
VILLA URQUIZA	59
FLORESTA	58
LA BOCA	55

Ilustración 23 -- Comparativo de Productos vs Retiros por barrio

De la comparación entre estas dos gráficas podemos concluir que existe más probabilidad de que una persona dé por terminado su contrato en alguno de los barrios listados en la gráfica de retiros si evaluamos el conjunto de clientes como un universo completo pero como el objetivo es segmentar y encontrar patrones de comportamiento es necesario evaluar esta conclusión barrio por barrio ya que hay cosas que el gráfico no alcanza a decir y es cuánto es esa probabilidad, se sabe de las gráficas que hay barrios con pocos servicios contratados comparativamente con otros barrios pero con una tasa de deserción es mayor.

Tabla 2 Tasa de deserción por barrio se listan los primeros 12

Barrio	Cantidad Activos	Cantidad Churn	Tasa de deserción %
VERSAILLES	221	20	9.05
CONSTITUCION	1,239	94	7.59
FLORESTA	851	58	6.82
PQUE. PATRICIOS	731	45	6.16
LA BOCA	1,045	55	5.26
NUEVA POMPEYA	848	43	5.07
PQUE. AVELLANEDA	758	38	5.01
SAN CRISTOBAL	1,132	55	4.86
BOEDO	1,043	49	4.7
MONTE CASTRO	712	33	4.63
VILLA RIACHUELO	307	14	4.56
LA PATERNAL	302	13	4.3

El cuadro anterior nos indica valores conocidos como la cantidad de productos activos, la cantidad de productos en retiro y adicional, ordenado por este campo calculado, muestra la tasa de deserción o en términos matemáticos la frecuencia base, o frecuencia de ocurrencia del evento, calculada como  $(n/n+m) * 100$  donde N es la cantidad de eventos del tipo que se está estudiando y m es la cantidad del otro tipo, en el caso de estudio; n= cantidad de productos en retiro y m= cantidad de productos que permanecieron activos.

A diferencia de los gráficos anteriores y si se ordena esta vez por la nueva variable, tasa de deserción, se encuentran listados los barrios que no aparecían en las primeras gráficas donde se mostraba la cantidad de ocurrencia del evento, esto quiere decir que hemos ganado conocimiento o información respecto del conjunto de datos o población y de la variable que se busca predecir.

Si se requiere diseñar un proceso que permita predecir quién se dará de baja es necesario construirlo a partir de modelos matemáticos evaluando la cantidad de información que aporta cada una de las variables del set a la predicción del objetivo, en este caso la variable Churn.

Cómo podemos segmentar la población con respecto de la variable que se quiere predecir, cuál modelo usar para este propósito. En el párrafo anterior se introduce un concepto fundamental en la ciencia de datos y la minería de datos y es la

cantidad de información, esta es la cantidad que reduce la incertidumbre sobre algo entre más certeza se tenga sobre el objetivo mejor es la información ganada.

Entendiendo el concepto de ganancia de información y teniendo como objetivo reducir el churn y tomar mejores decisiones con respecto al objetivo, el siguiente paso es preguntarse por si existen más variables que podamos correlacionar con la variable objetivo y que permitan aumentar nuestro conocimiento sobre la población estudiada y que permita la reducción de la incertidumbre.

Para continuar con el análisis tomaremos como muestra uno de los barrios con mayor tasa de deserción: Constitución.

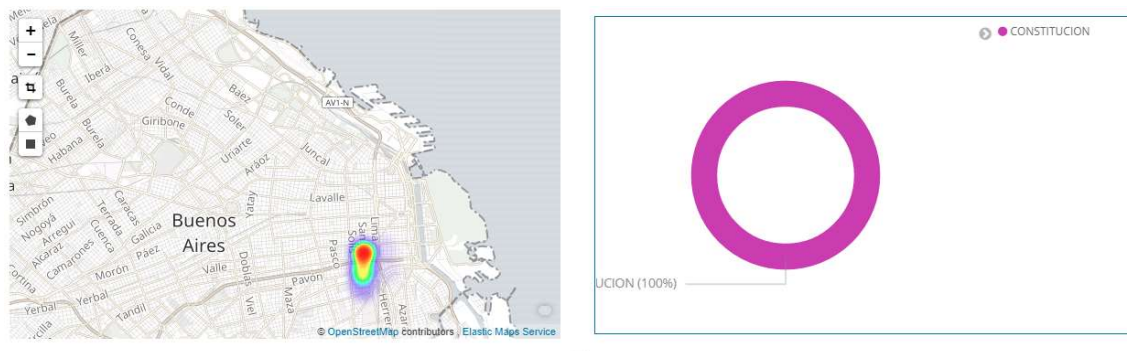


Ilustración 24 25-- Mapa de Calor de retiros con filtro constitución

En la siguiente gráfica se muestra la cantidad de clientes en Retiro por género y rango de edades. Como se observa la cantidad de mujeres y hombres no aporta gran cantidad de información y pero el rango de edad sí caracteriza mejor un posible caso de retiro. Finalmente podemos concluir que la cantidad de retiros no depende del género pero está fuertemente relacionada con la edad del suscriptor.



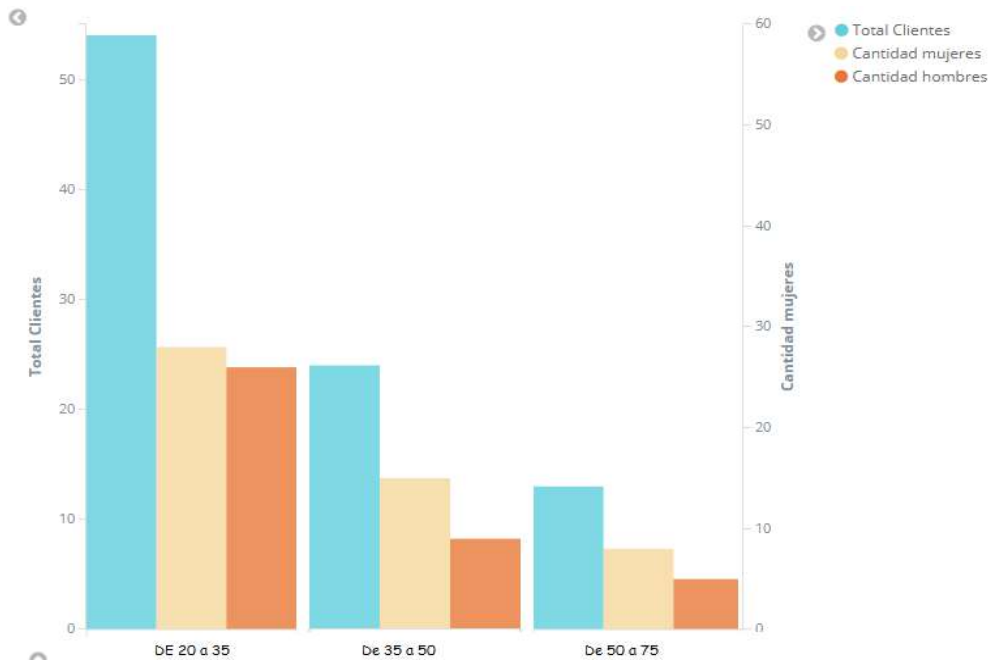


Ilustración 26 -- cantidad de clientes en Retiro por género y rango de edades.

Las otras variables de análisis seleccionadas en el estudio es el Scoring, que es una calificación del cliente dada por la empresa dependiendo del uso del producto, dinámica de pago, cantidad de llamadas al servicio al cliente. Para mostrar las variables en un mismo cuadro, se selecciona un cuadrante de calor y ver el comportamiento dependiendo de la variación de cada una de ellas. El eje X da información sobre la calificación y el eje y sobre la cantidad de llamadas.

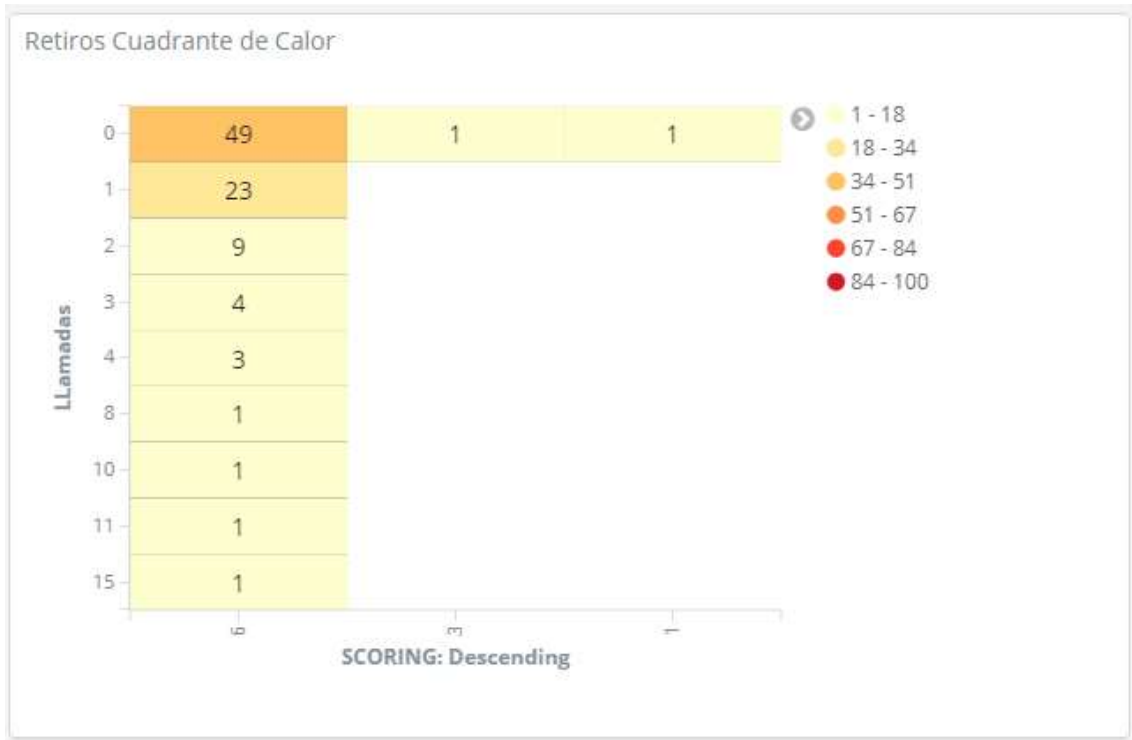


Ilustración 27 -- Cantidad de retiros por llamadas y Scoring

De esta podemos deducir que no importa el número de llamadas la mayor deserción se da en clientes con Scoring 6 incluso sin haber llamado al servicio en los últimos meses.

Con este análisis visual preliminar se puede concluir que existen variables que aportan a la predicción de la variable objetivo y también que no todas las variables aportan información y en consecuencia no es necesario tenerlas en cuenta dentro del modelado y evaluación.

A medida que vamos avanzando en el análisis preliminar de la información se gana conocimiento sobre la población estudiada, encontrando atributos que aportan información es la base de un método de predicción conocido como arboles de inducción.

Un modelo es una representación de la realidad creado para servir a un propósito, en ciencia de datos un modelo predictivo es una representación matemática o lógica para la estimación de un valor desconocido de una variable de interés, en este caso predecir cuáles clientes darán por retirado su contrato es el propósito.

Para la predicción del Churn se requiere construir un modelo que permita determinar la posibilidad de retiro de un cliente como una función de atributos descriptivos de la cuenta del cliente como número de llamadas del servicio al cliente, scoring; función del uso y de la dinámica de pago del cliente, e información demográfica como el barrio, edad y género.

Al seleccionar las variables que aportan más información el objetivo es llegar a que el grupo resultante después de la segmentación sea tan homogéneo como sea posible, es decir que sea puro. Si cada miembro del grupo tiene el mismo valor para la variable objetivo se puede decir que el grupo es puro, en casos reales como el caso de estudio llegar a un grupo homogéneo es complejo teniendo en cuenta la cantidad de variables analizadas, pero se busca optimizar la pureza del grupo cuanto sea posible y así tomar decisiones más acertadas.

Para problemas en los que se requiere clasificar se puede resolver creando una fórmula matemática que evalúe que tan bien cada atributo parte el set de datos en segmentos más puros. En la industria y en la ciencia actualmente para resolver este tipo de situaciones se usa el Machine Learning (aprendizaje de máquina) el cual consiste en enseñar a las máquinas a tomar decisiones a través de un algoritmo o conjunto fórmulas lógicas y matemáticas que describen el modelo de comportamiento.

Uno de los métodos para abordar problemas en los que se necesita segmentar un elemento dado un conjunto de elementos con diferentes atributos es la clasificación, y la metodología de aprendizaje de máquina es supervisada. Se dice que es supervisado porque para un conjunto dado se conoce la variable objetivo y es dependiendo de esta variable y la proximidad con otros elementos que se realiza la clasificación del nuevo elemento con variable objetivo desconocida.

La similitud entre elementos del conjunto permite describir y solucionar muchos de los problemas en los negocios que son abordados desde la ciencia de datos. Si dos elementos, clientes, compañías, productos comparten características y de alguna forma son similares también compartirán otras características desconocidas, es decir que pertenecen a una misma vecindad. Muchos de los algoritmos están basados en este concepto agrupando y ordenando por el atributo correcto.

El procedimiento para usar el concepto de vecindad o similitud como modelo predictivo consiste en dado un nuevo elemento donde queremos predecir la variable objetivo se requiere evaluar o entrenar en términos de minería de datos, un set de datos de muestra y seleccionar los elementos que comparten mayor similitud con el nuevo elemento y de esta forma, basado en el algoritmo de vecinos cercanos (k-Nearest Neighbors) predecir la variable objetivo para este elemento.

El método de los k vecinos más cercanos (en inglés, k-nearest neighbors, es un método de clasificación supervisada que sirve para estimar la función de densidad. Este método estima el valor de la función de densidad de probabilidad o directamente la probabilidad de que un elemento  $X$  pertenezca a la clase  $C$  a partir de la información proporcionada por el conjunto de elementos de muestra o set de entrenamiento.

En el reconocimiento de patrones, el algoritmo k-nn es usado como método de clasificación de elementos basado en un entrenamiento mediante ejemplos cercanos en el espacio de los elementos.

El procedimiento para la evaluación analítica consistió en separar los datos de trabajo en data de entrenamiento y datos de evaluación para que a través del algoritmo K-nn de la librería de analítica “sklearn” de Python, lenguaje de programación y de uso libre y con librerías disponibles en internet con código abierto.

Entrenamiento del set de datos:

Para lograr esto, se crea un objeto Dataframe (Similar a una tabla Excel) el cual se limpia, esto quiere decir: se eliminan valores nulos y valores que puedan no ser reales distorsionando la data, en este caso el DataSet original previo al proceso de limpieza contenía valores de edades superiores a 100 años, y en otros casos entre los 0 y 10 años, para el caso de una compañía de Telco estos valores no son posibles. Finalmente se eliminan las columnas que no van a ser procesadas por el algoritmo de clasificación.

el DataSet que queda después de la limpieza se divide en data de Testing y en data de entrenamiento, esto con el fin de tener data para realizar el modelo y data para testear el modelo, al procesar este Data set con el algoritmo Knn se clasifica

a cada uno de los elementos del set de datos de entrenamiento, creando una instancia de entrenamiento de los elementos.

Posterior a esto se procesan los elementos que no fueron utilizados para el entrenamiento del algoritmo con la función de evaluación de la librería K-nn y esta arroja la probabilidad de que el elemento evaluado pertenezca a la clase de la variable objetivo comparándolo con su valor real.

En el caso de estudio se tomó la muestra y se evaluó el 30% de los elementos del conjunto, este algoritmo clasificó a cada uno de los elementos ingresados en la muestra en una vecindad de 7, esto quiere decir que toma el valor dependiendo de los 7 vecinos más cercanos, y con la muestra define una función de distribución la cual se evalúa con los nuevos elementos entrantes, prediciendo en este caso de estudio con una precisión del 97.2%.

Para definir el número de vecinos se debe simular el modelo con los diferentes valores de vecinos cercanos. Esto quiere decir medir la exactitud (Accuracy) de la predicción evaluando el modelo con los diferentes valores posibles.

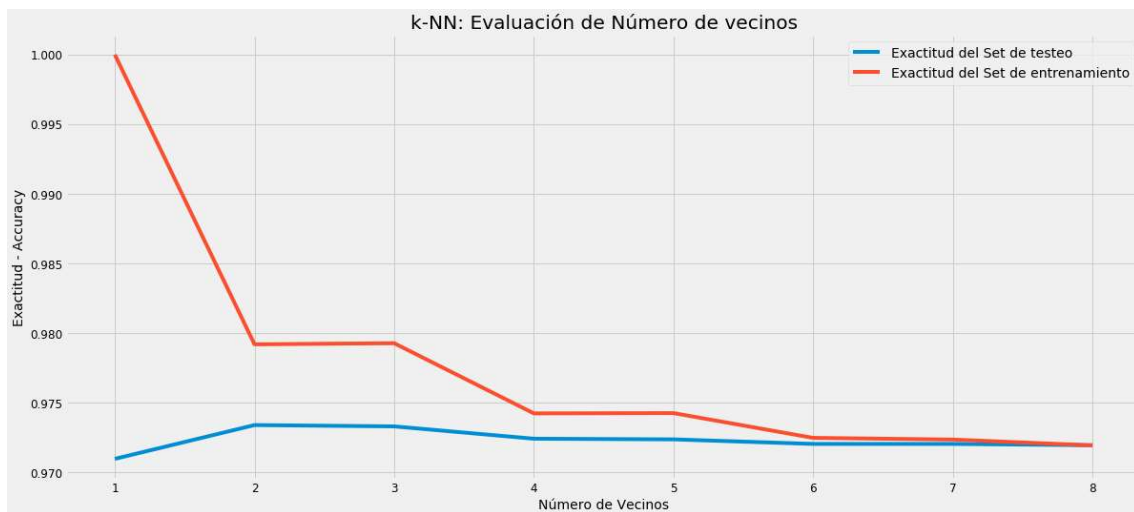


Ilustración 28 – K-NN Evaluación de Exactitud dependiendo del número de vecinos

Del modelamiento nos queda que dadas las condiciones de entrada (Variables conocidas) valores de cada atributo, el modelo es capaz de predecir con un 97.2% la probabilidad de que un cliente dé por terminado su contrato, de implementarse este algoritmo en la base de clientes permite tomar acciones en el momento y así evitar pérdidas de clientes y por lo tanto pérdida de ingresos.

El uso de análisis de datos para obtener información valiosa de negocio y la incorporación de herramientas automáticas de decisión es cada vez más usado y necesario en la industria. Las oportunidades de implementar un sistema aprendizaje automático (Machine Learning) son amplios: predecir negociaciones comerciales, predecir tendencias de consumo, mejoramiento de la producción, medicina personalizada, educación, desarrollos de infraestructura, eco-movilidad, agricultura, consumo energético, apoyo judicial, análisis macroeconómicos, desafíos éticos, lucha contra la corrupción y equidad social, esta tecnología está modificando el desarrollo de las actividades del hombre, no existe un campo de aplicación asociado a la ciencia donde no sea aplicable.

“Estamos en los albores de una tecnología que se constituye en un nuevo factor de producción. La inteligencia artificial, guiada por una sabia y renovada visión humanista puede contribuir a consolidar una integración regional predictiva e inclusiva para todos los latinoamericanos”. Revista BID (Banco Interamericano de Desarrollo (BID, 2018).

A continuación se presentan ejemplos de casos de uso en los que soluciones de tipo Data Analytics pueden ser implementadas en las empresas operadoras de servicios de telecomunicaciones con el fin de obtener información valiosa a partir de los sistemas de información de la compañía e integración con fuentes externas de datos. Estos casos buscan mejorar el desempeño en la industria de las telecomunicaciones desde una mirada transversal en los procesos de negocio.

Los casos están relacionados con los procesos de negocio definidos en el modelo ETOM que se abordaron en el capítulo 2.1 de este documento:

Mercadeo y ventas:

- Monitoreo del mercado y tendencias de consumo.
- Empaquetamiento de productos para generar ofertas atractivas.
- Ofertas personalizadas en tiempo real durante navegación, por tipo de dispositivo, basado en la localidad, durante interacción en vivo.
- Publicidad inteligente basada en el historial de navegación.

#### Producto:

- Definición y desarrollo de producto.
- Análisis de introducción de productos.
- Optimización del desempeño de los productos.
- Análisis de probabilidad de compra.
- Reducción de errores en el proceso de órdenes de trabajo.

#### Administración de Clientes y Servicio:

- Planes de atención personalizados.
- Interacción personalizada en tiempo Real.
- Predicción de la motivación de Churn.
- Cálculo de probabilidad de aceptación de ofertas en tiempo real.
- Atención al cliente proactiva.
- Optimización de auto instalación del cliente.

#### Recursos e infraestructura:

- Análisis predictivo de fallas en la red.
- Detección de fallas por localidad y auto recuperación.
- Optimización de la asignación de órdenes técnicas.
- Optimización de rutas de servicios técnicos.

#### Operación de la compañía:

- Detección de fraudes.
- Informe de tráfico analizando el tráfico en la red.
- Análisis de perfiles.

Para poder explotar el potencial de información que hay en las compañías y empezar a generar valor a partir de los datos se necesita de una estrategia orientada a datos que permita acondicionar a la compañía con este nuevo modelo de administración, diseñar políticas, destinar recursos, capacitar personal y generar una cultura de datos que permita explotar la información que existe tanto en los sistemas internos como en el entorno, en términos generales implementar una estrategia de datos.

### **6.3 Implementación**

Crear un modelo que defina el comportamiento de la variable de estudio, Churn o probabilidad de dar por terminado el contrato en nuestro caso, es una de las tareas del proceso de toma de decisiones pero para que este modelo se convierta en parte de la estrategia de la compañía se tiene que articular el análisis de las variables con las acciones a tomar, y que estas sean formuladas y ejecutadas de manera apropiada y que sean transversal a toda la organización.

La ejecución de una estrategia exitosa debe involucrar a todos los niveles de la compañía y cada uno de estos niveles debe entender la importancia de poner los datos como el eje en la toma de decisiones. “El éxito de la formulación de la estrategia no garantiza una implementación exitosa de la estrategia. Siempre es más difícil hacer algo (implementar la estrategia) que decirlo” (David, 2013).

Los líderes de los negocios (Administradores, Managers) en esta nueva era deben fortalecer la cultura de los datos en sus organizaciones y comprender que una clave de éxito de las compañías en el siglo XXI es el entendimiento de ellos, especialmente cuando se pone en consideración la velocidad de generación de información y la evolución de las herramientas para el análisis. Entendiendo estos dos factores como amenaza y oportunidad respectivamente, debe desarrollarse la fortaleza del análisis y toma de decisiones a partir de la generación de valor en los datos, es decir una estrategia con enfoque en los datos (Data Strategy).

Según el Tablero de control comercial el Churn voluntario para el mes de diciembre de 2017 y del producto Televisión fue de 1.16% y el valor estimado de cada cliente (Fuente interna Análisis Customer Care) está calculado en USD1000 según estas dos premisas la suma de las pérdidas al año por este concepto representa alrededor de USD 22.668.000 es por esto que predecir quien va a dar por retirado su



contrato es un objetivo estratégico para la compañía y se constituye en clave de éxito para el alcance de la estrategia corporativa.

En un mercado saturado y altamente competitivo como el de las telecomunicaciones donde el incremento de una tarifa, una mala experiencia con el servicio al cliente o un error en la facturación pueden derivar en la pérdida de un contrato comercial, mantener una buena relación con el cliente y mantenerlos satisfechos no es una tarea fácil.

Es por esto que cada vez los operadores deben enfocarse en introducir ofertas personalizadas para retener al cliente y atender a cada cliente de manera diferenciada, es en esta dimensión donde se conjuga la necesidad de involucrar técnicas de Data Science (Big Data, Machine Learning, Minería de datos) que permita analizar la cantidad de clientes y la cantidad de información que ellos producen a una velocidad que permita tomar decisiones de forma adecuada.

Para el caso de estudio en el que se propone modificar el proceso de toma de decisiones y llevará a la compañía en otra dirección, los cambios desde la administración estratégica son extensos ya que implica modificar procesos, cultura y soportarla.

Entre las responsabilidades administrativas centrales para la implementación de las estrategias están: establecer objetivos anuales, diseñar políticas, asignar recursos, modificar la estructura organizacional existente, reingeniería y/o reestructuración, desarrollar una cultura de apoyo a la estrategia, adaptar procesos de producción y operaciones, desarrollar una función efectiva de recursos.

El principio de la ejecución de una estrategia enfocada en Datos es el mismo que el de cualquier estrategia. La estrategia actúa como una hoja de ruta que ayuda a estructurar acciones y a dimensionar todo lo que se necesita para que se consigan las metas planteadas, esto involucra establecer objetivos, asignar recursos para realizar inversión en infraestructura, diseñar políticas con enfoque en datos, trabajar en la cultura organizacional que apoye la estrategia de datos, modificar los procesos de producción y operaciones; recolectar información, invertir en herramientas de análisis de datos, generar acciones que afecten la operación a partir de los hallazgos

en los datos, emplear nuevos talentos, capacitar al personal actual, entre otras actividades que garanticen el éxito en la implementación de la estrategia.

### 6.3.1 Establecer objetivos

Los objetivos son el soporte para la implementación de estrategias ya que representan la base para la asignación de recursos, y son el mecanismo para la evaluación gerencial, adicional son el principal instrumento para monitorear el progreso y establecer las prioridades dentro de la organización.

Normalmente los objetivos están definidos en términos financieros, participación de mercado, procesos internos de negocio y recursos humanos. En términos de una estrategia de datos los objetivos además deben ser consecuentes o estar alineados con las preguntas que se quieren resolver tal como se mencionó en el punto 4.3.2 Definición de preguntas claves de este documento.

Estos objetivos deben ser de corto plazo (9 a 12 meses), deben involucrar a más de un área de la compañía, tener un responsable; área y gerente, deben ser medibles y tener un tiempo límite de consecución y que genere una ventaja competitiva o financiera.

	nov-16	dic-16	oct-17	nov-17	dic-17	Var. vs Mes pasado	Var. vs año pasado
	25.0	19.5	23.0	24.5	21.0		
Vtas Ingresadas	2,887	2,875	2,679	2,584	2,637	2.0%	-8.7%
*Vtas Pendientes	19,455	18,948	17,020	20,913	17,279	-17.4%	-11.2%
Altas	2,229	2,281	2,234	1,989	2,321	16.7%	4.1%
Demanda de Bajas	-4,127	-4,464	-3,849	-3,858	-4,443	15.2%	7.7%
Bajas Voluntarias	-1,912	-2,192	-2,007	-2,051	-2,419	17.9%	26.6%
Recuperos Voluntarios	431	472	472	484	530	9.5%	22.9%
Bajas Netas Voluntarias	-1,480	-1,720	-1,535	-1,567	-1,889	20.6%	27.6%
Churn voluntario	1.03%	0.97%	1.03%	1.12%	1.16%	0.0%	0.2%
*Bajas Netas Mora	-11,492	-20,911	-9,610	-11,142	-19,082	71.3%	-8.7%
Churn Mora	0.32%	0.61%	0.28%	-0.33%	-0.56%	-0.2%	-1.2%
*Captura Gestión	7,229	-9,974	6,474	-792	-10,006	1163.4%	0.3%
*Ajuste	473	-147	88	-145	-323	-	-
*Captura Total	7,678	-10,121	6,562	-937	-10,329	1002.3%	2.1%

Ilustración 29 -- Tablero comercial Cabletelco Diciembre 2017

Las respuestas a las preguntas claves de negocio deben mostrarnos el estado actual de la compañía, sirven para priorizar las problemática,

asegurar el uso de tiempo y recursos de la forma más efectiva y enfocarnos en las soluciones.

El tablero comercial de la compañía de telecomunicaciones muestra el mes de diciembre 2017 comparado con los dos meses anteriores y el valor para el mismo mes en el año anterior. Si vemos el mes de diciembre las bajas voluntarias fueron 1889 para uno de los tipo de producto ofrecido por la compañía, pero en total se registraron 2419 retiros (Intenciones de retiro) de las cuales se evitaron 530 bajas, evitar estas bajas se conoce como recupero y es un medidor de la efectividad de la estrategia de Churn.

Teniendo en cuenta que el objetivo del caso de estudio es el apoyo en las estrategias de retención de cliente mediante análisis de datos se toma como objetivo estratégico disminuir la tasa de Churn en la compañía en un 20% en los siguientes 12 meses.

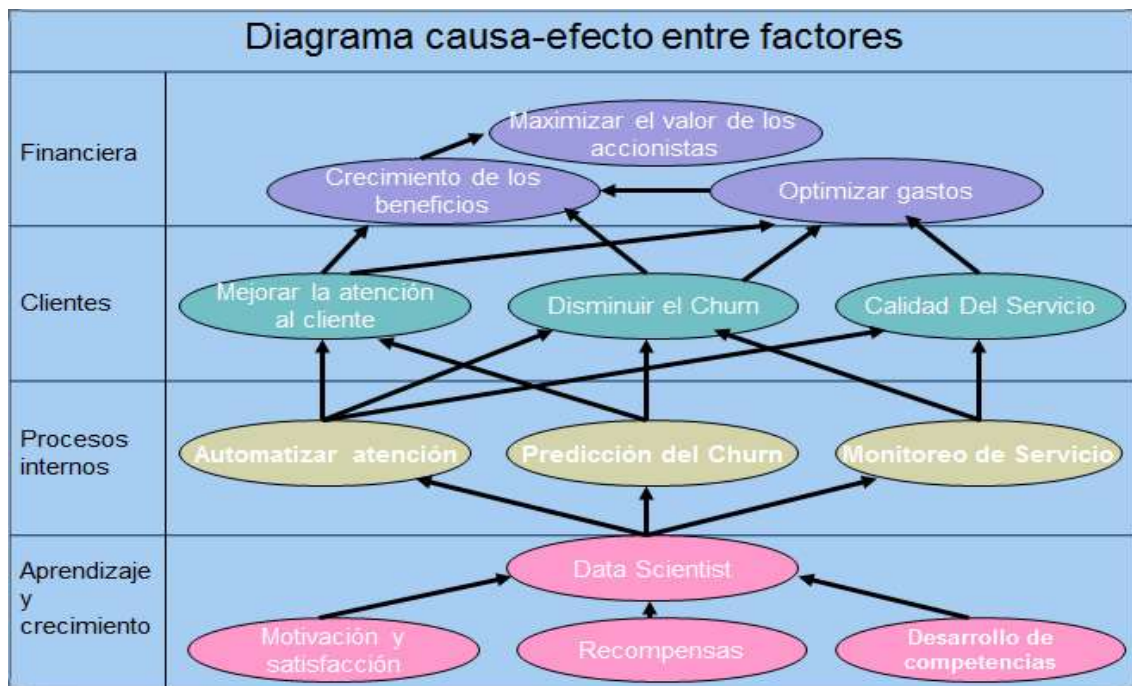


Ilustración 30 -- Diagrama de causa y efecto entre factores críticos del negocio

El diagrama anterior muestra la relación entre el objetivo estratégico o iniciativa de negocio con los demás factores organizacionales. Se puede deducir que una disminución del Churn generaría ingresos a la compañía, teniendo en cuenta que cada cliente está valorado en 1000USD reducir la tasa de Churn al 20% evitaría una pérdida de ingresos relacionada directamente con el objetivo.

Es así como un objetivo estratégico relacionado con los procesos a través de toda la organización aporta en la consecución de los objetivos financieros y responde a los intereses del directorio o grupo de accionistas de la compañía.

Adicional también se muestra la relación entre el objetivo y los procesos internos de análisis de datos, como predicción del churn, la automatización de la atención al cliente, o el monitoreo de la calidad del servicio, lo cual son procesos que afectan directamente la tasa de retiros y para los cuales se requiere tener condiciones que permitan alcanzar el objetivo; como personal capacitado, una cultura orientada a la estrategia de datos, políticas que permitan desplegar la estrategia y una estructura organizacional que soporte esta forma de orientar la estrategia.

### **6.3.2 Políticas de datos**

Los cambios en la dirección estratégica de una empresa no son automáticos. Las políticas cotidianas son necesarias para hacer que una estrategia funcione.

Las políticas facilitan la solución de problemas recurrentes y guían la implementación de una estrategia. En términos generales, la política se refiere a los lineamientos específicos, métodos, procedimientos, reglas, formas y prácticas administrativas establecidas para respaldar y fomentar el trabajo hacia el logro de metas conocidas. Las políticas son instrumentos para la implementación de estrategias.

Las políticas establecen límites y restricciones sobre el tipo de acciones administrativas que se pueden aplicar para recompensar o sancionar conductas; aclaran lo que se puede y no se puede hacer en el logro de los objetivos de la organización.

### **6.3.3 Creación de una cultura de apoyo a la estrategia de datos**

La cultura organizacional se puede definir como el conjunto de percepciones, sentimientos, actitudes, hábitos, creencias, valores, tradiciones y formas de interacción entre los grupos existentes en las organizaciones.

Para la implementación de una estrategia exitosa los líderes deben esforzarse en conservar, reforzar y utilizar aspectos de la cultura existente que apoyen las nuevas estrategias propuestas, saber reconocer y reemplazar cuales aspectos de la cultura actual no se alinean con la nueva estrategia y tener la habilidad de crear nuevos conjuntos de valores, mitos, creencias y demás aspectos que permitan su despliegue o implementación.

Una cultura que soporte la estrategia de datos entiende estos como un activo de la compañía, y se busca que esto sean aprovechados de manera efectiva en todos los niveles de la organización con el objetivo de hacer mejoras que generen; mejores decisiones de negocio como un mejor entendimiento de los clientes permitiendo mejorar los esfuerzos de marketing, hacer más eficiente la cadena de abastecimiento y producción, soportando los procesos de calidad reduciendo la cantidad de reclamos, diseñando nuevas opciones de generar ingresos entre otros.

El cambio hacia una estrategia de datos debe ser responsabilidad del directorio de la compañía y debe bajar el compromiso a todos los niveles corporativos e involucrar al resto del personal, esta es la forma como se construyen los mitos y se crean héroes o heroínas, muchas de las iniciativas en análisis de datos vienen de la base operativa quien es la que se enfrenta día a día a las situaciones del negocio.

La responsabilidad de los directores es crucial para que la implementación de sistemas de análisis de datos como apoyo a la estrategia y toma de decisiones sea exitoso, se le dé la importancia que requiere y no sea considerado como otro proyecto del área de tecnología sino como el soporte de la estrategia.

“Una cultura de datos es sobre todas las personas a través de la compañía entendiendo el valor de los datos y cómo pueden estos ayudar al éxito del negocio” (Marr, 2017).

Uno de los factores claves en la creación de la cultura es la comunicación; líderes y administradores deben agendar tiempos de actividades con el

fin de comprometer a las personas con la estrategia de datos, reforzando sobre la importancia y los beneficios que ésta puede generar.

Una buena opción para construir una cultura de datos es exponer los resultados de acciones realizadas, o usando casos de usos de otras compañías donde se evidencie los impactos de haber adoptado estrategia de datos, construir una narrativa propia con casos de éxito que ya se han demostrado y probado en la organización.

Dar acceso a toda la organización de la información disponible, sentirse cómodos con el uso de datos para guiar las decisiones, generar propiedad intelectual de los desarrollos de analítica.

### **6.3.4 Estructura y estrategia de Datos**

El propósito de una organización es lograr un esfuerzo coordinado para la consecución de objetivos, definir las tareas y funciones, las relaciones de dependencia, reporte de actividades y alcance de objetivos, como las responsabilidades y niveles de responsabilidad permiten ubicar a los individuos dentro de la organización y dimensionar sus alcances y limitaciones.

“Combinar puestos para formar departamentos genera una estructura organizacional, un segmento de control y una cadena de mando. Los cambios de estrategia a menudo requieren cambios en la estructura porque se pueden crear, eliminar o fusionar los puestos. La estructura organizacional dicta cómo se asignan los recursos y cómo se establecen los objetivos en una empresa” (David Fred R, 2013).

Los cambios en la estrategia normalmente requieren cambios en la estructura organizacional ya que la estructura determina la forma en que se establecen los objetivos, se conciben las políticas y principalmente cómo se asignan los recursos. Si una organización está organizada por zonas geográficas de esta manera se establecerán los recursos si es por funciones entonces los recursos se asignan funcionalmente.

El rápido crecimiento en el volumen de datos, la velocidad de generación y la eficiencia de las herramientas de análisis está modelando el entorno y las empresas y organizaciones tienen la necesidad de encontrar valor y apoyar la estrategia y su crecimiento en esta nueva forma de mundo a la que se enfrentan.

Una de las formas en la que se evidencia esta transformación dentro de las organizaciones es la aparición del área de Big Data y Analytics, donde se requiere personal que entienda el negocio pero también cuente con habilidades de analítica y pueda operar o configurar herramientas para el análisis.

Se requiere un responsable de la inversión en estas herramientas y proyectos, ser capaz de valorar cuál iniciativa de negocio ofrece un mejor retorno de la inversión, es por esto que las organizaciones en sus grupos de tecnología o informática ha aparecido el líder o gerente de Data Analytics, dando paso a funciones y responsabilidades antes inexistentes en las compañías obligando a las organizaciones a redefinir la estructura permitiendo definir roles, funciones, dependencias y recursos disponibles.

En la organización del caso de estudio se tiene una organización funcional, es decir una dirección por funciones, en la cual hay un directorio de accionistas y debajo del directorio un CEO, encargado de la gerencia en general y que responde al directorio, y dependiendo de éste se encuentra, entre otras, la oficina de informática y operaciones.

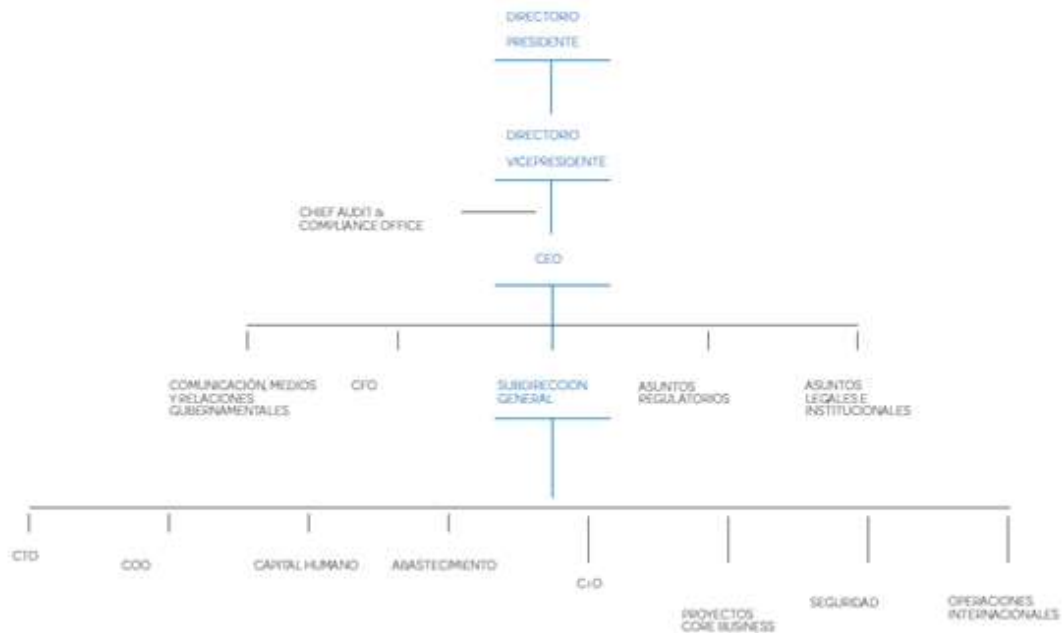


Ilustración 31 -- Organigrama funcional hasta el nivel de directorio y las gerencias

En la oficina de informática y operaciones, a cargo del CIO, se encuentra la generencia de desarrollo e implementaciones, la cuál se encarga de evaluar proyectos, desarrollarlos e implementarlos para que sean operativos a la compañía.

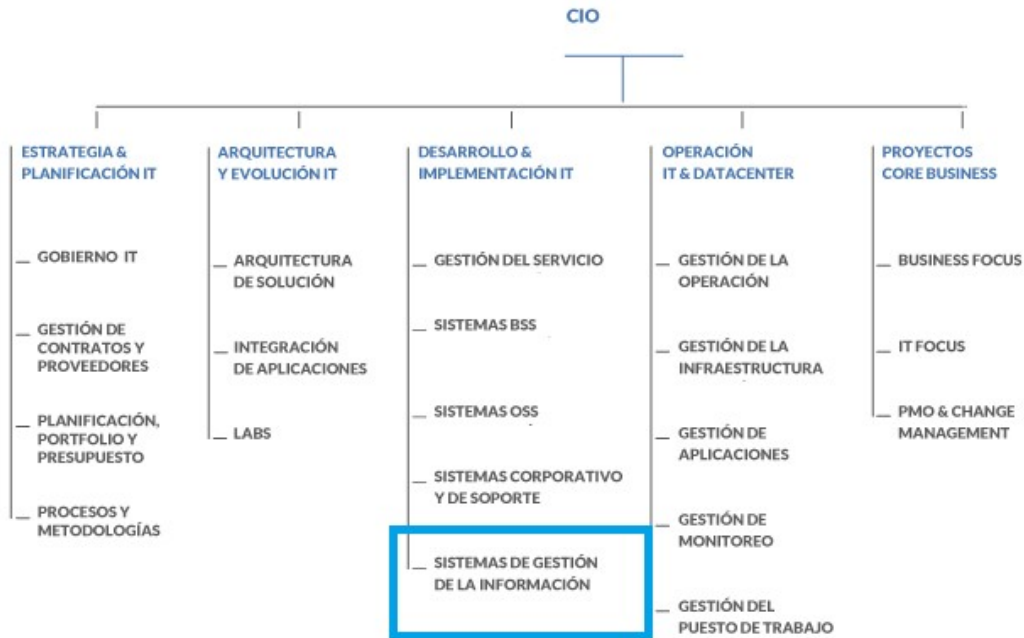


Ilustración 32 -- Diagrama de la oficina de Informática y Operaciones

Adicional a los sistemas de información, útiles a la operación y principal fuente de datos internos, empiezan a aparecer dentro de las compañías las gerencias relacionadas a otros sistemas como el de Big Data y analítica de datos. Dentro de la cual se debe contar con las habilidades y preparación necesaria para enfrentarse a la transformación.



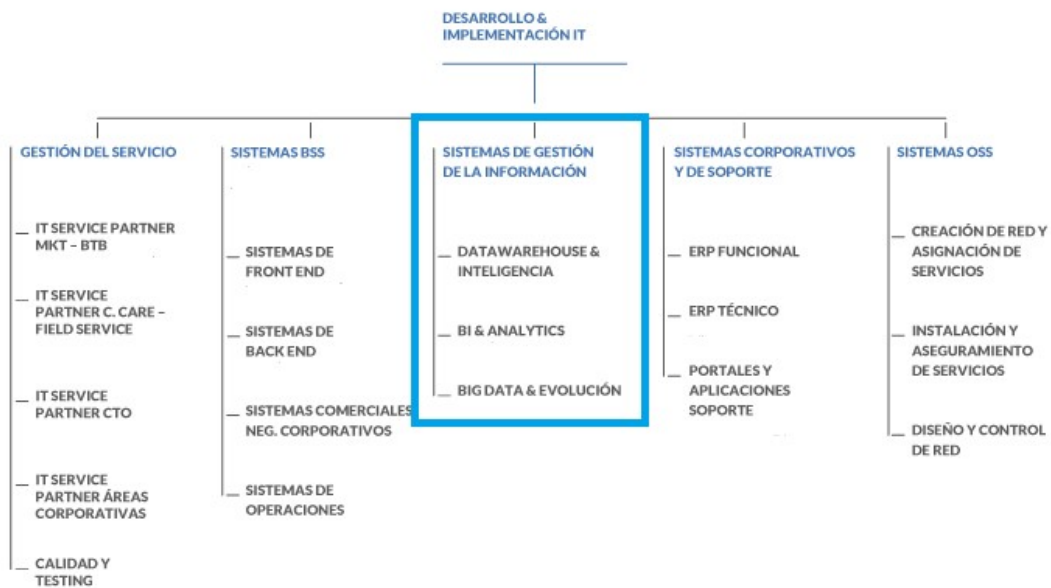


Ilustración 33 -- Diagramas del área de Sistemas de información

Es a cargo de esta oficina en la que se desarrollan los proyectos de Big Data y Analytics, y quienes lo lideran, se encargan de seleccionar las iniciativas de negocio que puedan ser implementadas y que representen una ganancia medible para la organización en términos de mejora de procesos y/o retorno de la inversión.

Desde la oficina de Sistemas de gestión de la información tiene como responsabilidad conducir el compromiso organizacional con la estrategia de datos y empoderar a las diferentes organismos dentro de la compañía con pensamiento creativo y analítico para que sea capaz de identificar donde y cuando puede aprovechar los datos para potencializar el modelo de negocio.

Esta área tiene el potencial para ayudar a la organización a analizar los datos que ya se recolectan e identificar nuevas fuentes de datos que valen la pena recolectar y pueden aportar al negocio a conseguir sus objetivos.

Entre las principales funciones de esta área son:

- Elaborar y proveer un proceso formal para identificar donde los datos y el análisis de estos pueden generar un impacto al negocio que apoye el alcance de las iniciativas de negocio en el corto plazo.
- Priorizar las iniciativas de negocio dependiendo de la relevancia de éstas y el impacto que generan en la obtención de objetivos.

- Empoderar a los líderes del negocio y a los usuarios clave para visionar el poder de soportar decisiones con el análisis de datos para que así nazcan las iniciativas y entiendan lo que es posible con el uso de datos tanto internos como externos.
- Conducir la alineación del negocio con IT alrededor de las mejores oportunidades en analítica con una hoja de ruta clara sobre lo que se debe hacer en los próximos 12 meses.
- Conformar equipos con las habilidades necesarias para enfrentar los retos organizaciones en analítica.

Los cambios en la estrategia generan cambios en la estructura organizacional. La estructura debe establecerse de forma que apoye las intenciones estratégicas de la empresa, y en consecuencia, facilite la implementación de la estrategia.

### **6.3.5 Plan de acción**

Durante esta etapa del proceso se formula el plan de acción para implementar de forma operativa el modelo de predicción desarrollado en el capítulo de modelamiento.

Estrategia de Negocio: Uso de herramientas de análisis predictivo (Data Analytics) como soporte para la toma de decisiones y apoyo a la consecución de objetivos en la organización. Objetivo primario: Disminuir los casos de Churn voluntario en un 20% durante el siguiente año en la ciudad de Buenos Aires para el producto Internet.

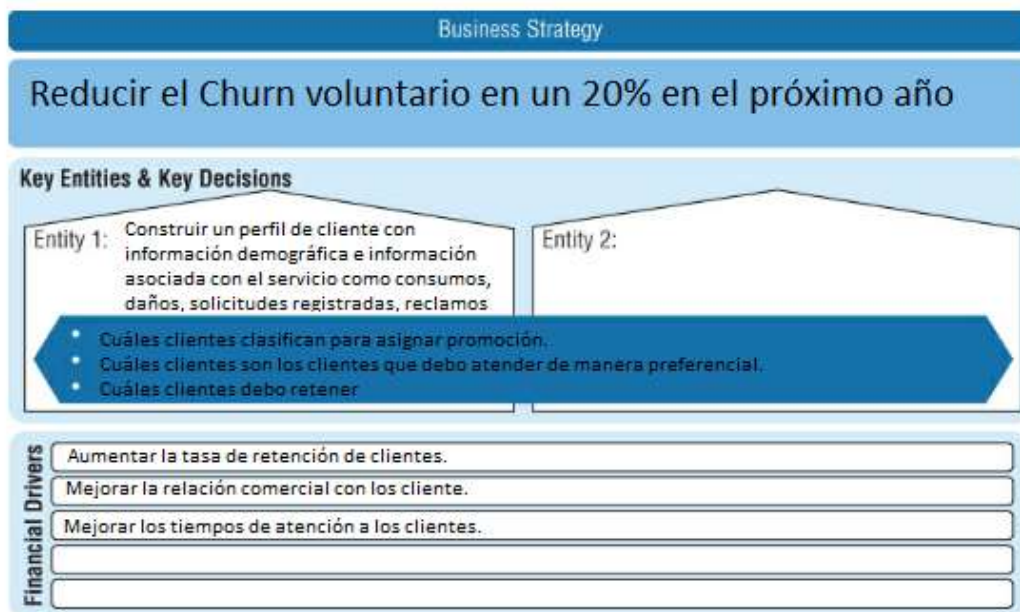


Ilustración 34 -- Documento Estrategia de Negocio propuesto por Bernard Mars en el libro Data Strategy.

Acciones a implementar:

- Ajustar la estructura de la compañía para el nuevo modelo de toma de decisiones y soporte a la estrategia.
- Desarrollar políticas de apoyo a la estrategia de datos.
- Capacitación del personal para potenciar el pensamiento analítico.
- Diseño de plan de incentivos para las propuestas de mejoras en procesos operativos o generación de valor a través de los datos.
- Asignar recursos humanos de las áreas involucradas: Atención al cliente, Comercial (Diseño de planes comerciales de retención), Marketing y sistemas para el desarrollo del modelo de análisis para predicción de Churn.
- Disponibilizar ambientes de desarrollo y pruebas de la solución.
- Diseñar Ofertas personalizadas basado en los consumos del cliente.
- Plan de capacitación al personal de atención al cliente para la integración con el sistema de analítica.
- Integrar el modelo de predicción de Churn para que en el momento del llamado del cliente se clasifique al cliente como posible caso de Churn y sea atendido por el departamento especializado en estos casos.

- Integrar el sistema de predicción de Churn para que el CRM permita asignarle planes de oferta de retención al cliente en Churn.
- Campañas Masivas enfocadas a medir el avance de la relación y el nivel de satisfacción de los clientes en posible Churn.
- Mejorar el modelo para que integre variables como las motivaciones del cliente a dejar la compañía a través del estudio del historial de comentarios de las llamadas y redes sociales.
- Campañas de atención al cliente de manera proactiva con el fin de mejorar la relación comercial.

La implementación de la estrategia o ejecución de las tareas planteadas es una guía de ejecución de actividades, no supone el éxito de la estrategia sino una hoja de ruta ya que para la implementación se debe ejecutar cada una de las partes propuestas en ella y crear las condiciones necesarias dentro de la compañía para que la estrategia funcione como modificar la estructura, crear políticas, desarrollar aspectos culturales, desarrollar las habilidades en las personas y asignar los recursos necesarios para la implementación.

#### **6.4 Seguimiento**

El seguimiento es la etapa del proceso en la que se evalúa la implementación de la estrategia, y se realiza el control de las acciones que se están ejecutando y revisa el alcance de los objetivos.

Las herramientas y equipos de Bussines Inteligence (BI) dentro de las organizaciones tienen las responsabilidad de entregar y habilitar, para uso de usuarios y gerentes, el acceso a reportes y cuadros de mando que permiten realizar un seguimiento a la estrategia y a las actividades permitiendo revisar el alcance de los objetivos y tomar las medidas correctivas.

El proceso de implementación de estrategias es un proceso cíclico que exige una retroalimentación sobre las acciones tomadas. Esta retroalimentación es proporcionada por las herramientas de BI, para el caso de estudio es el inicio del proceso. En el punto 2.3 contexto de la empresa de telecomunicaciones del caso de

estudio se hizo la introducción al tablero de comando de seguimiento y para nuestro caso será el mismo cuadro de evaluación.

Vtas Ingresadas	3,032	3,146	3,269	3,140	3,218	●	2.5%	●	6.1%
*Vtas Pendientes	19,004	19,553	20,012	24,518	19,998	●	-18.4%	●	5.2%
Altas	2,368	2,483	2,768	2,373	2,823	●	19.0%	●	19.2%
Demanda de Bajas	-3,441	-3,920	-3,601	-3,582	-4,289	●	19.7%	●	24.7%
Bajas Voluntarias	-1,882	-2,222	-2,210	-2,134	-2,514	●	17.8%	●	33.6%
Recuperos Voluntarios	412	497	526	494	554	●	12.3%	●	34.5%
Bajas Netas Voluntarias	-1,470	-1,724	-1,684	-1,640	-1,960	●	19.5%	●	33.3%
Churn voluntario	1.72%	1.55%	1.70%	1.75%	1.78%	●	0.0%	●	0.2%
*Bajas Netas Mora	-4,227	-9,550	-6,053	-5,838	-9,893	●	69.5%	●	3.6%
Churn Mora	0.20%	0.44%	0.27%	-0.25%	-0.43%	●	-0.2%	●	-0.9%
*Captura Gestión	18,218	5,253	18,866	12,114	8,225	●	-32.1%	●	56.6%
*Ajuste	-898	-1,754	-977	-330	-2,833		-		-
*Captura Total	17,408	3,499	17,889	11,784	5,392	●	-54.2%	●	54.1%

Ilustración 35 -- Tablero de control comercial

## 7. Discusión

Durante esta investigación se desarrolló el concepto de estrategias soportada por análisis de datos (Data Strategy), el cual consiste en la generación y desarrollo de estrategias, para el caso de estudio retención de clientes, soportadas por análisis de datos, y cómo este proceso de generación de estrategias se relaciona con la administración estratégica y la toma de decisiones en la industria, encontrando patrones que permiten definir una metodología y resaltar retos asociados a la implementación de una estrategia con enfoque en datos, y que ésta sea tomada por la compañía más allá que sólo un proyecto del grupo de TI (Tecnología e Informática).

Este concepto se encuentra en auge ya que en la actualidad se plantea la necesidad de la industria, y en general en todas las organizaciones, de incursionar en el proceso de transformación digital apoyada en nuevas tecnologías como el Big Data y Data Analytics.

Sim embargo, cuál es el esfuerzo que se hace desde las organizaciones para lograr esa transformación digital, cómo ir delegando mayor responsabilidad de decisión a herramientas como la inteligencia artificial y vencer la inercia de la toma de decisiones en cabeza de los Top Managers, cómo lograr la transformación digital y apoyar a las organizaciones en esta metamorfosis son parte de las preguntas que se responden en este documento.

El mayor reto de la transformación digital no se encuentra en la incorporación de la tecnología, los desarrollos que ésta conlleve, los costos de adquisición o encontrar el personal con las habilidades para operar e interactuar con estas herramientas, el mayor reto se encuentra en los cambios organizacionales que se deben asumir para conseguirlo.

Introducir a la organización en los retos de un cambio disruptivo como se plantea pone en desequilibrio el orden ya establecido y que ha tomado tiempo en construir, exponer a las personas y organizaciones a una nueva forma de interactuar con el entorno puede llegar a ser caótico y conducir, si no se toman las medidas necesarias, al fracaso de la organización.

Es por eso que para encontrar los patrones y retos asociados a la implementación de estrategias soportadas por análisis de datos, partimos desde el análisis estratégico y su relación con la toma de decisiones soportada por el análisis de datos.

La administración estratégica se define como el método objetivo, lógico y sistemático para la efectiva toma de decisiones. Su finalidad es organizar información cuantitativa y cualitativa para la elaboración, implementación y control de estrategias enfocadas a la consecución de metas u objetivos en las organizaciones.

Este método o framework propone una serie de pasos o puntos a tener en cuenta y está dividido en tres etapas; formulación de estrategias, implementación y evaluación. Este modelo propuesto se asemeja al proceso de toma de decisiones soportado por herramientas de análisis de datos en el cuál se necesita un entendimiento del negocio, un entendimiento de los datos, la creación de un modelo, evaluación del modelo, implementación y seguimiento.

En el primer conjunto de acciones en el marco de la administración estratégica se encuentra formulación de estrategias y éste está dividido a su vez por cuatro tareas principales; declaración de Misión y Visión, que definen el estado actual y futuro, qué somos y qué queremos ser respectivamente, una recolección de información, para lo cual se puede apoyar en auditorías internas y externas, el siguiente punto es definición de objetivos estratégicos, y finalmente la formulación de estrategias, que consiste en la enumeración de una serie de tareas a ejecutar.

Los siguientes puntos en el marco de referencia de la administración estratégica son implementación, el cual hace referencia a todas las consideraciones necesarias que se deben tener para ejecutar la estrategia, temas organizacionales como cultura, políticas, estructura y personas, y por último seguimiento a la estrategia en el cual se monitorea el avance del cumplimiento de los objetivos.

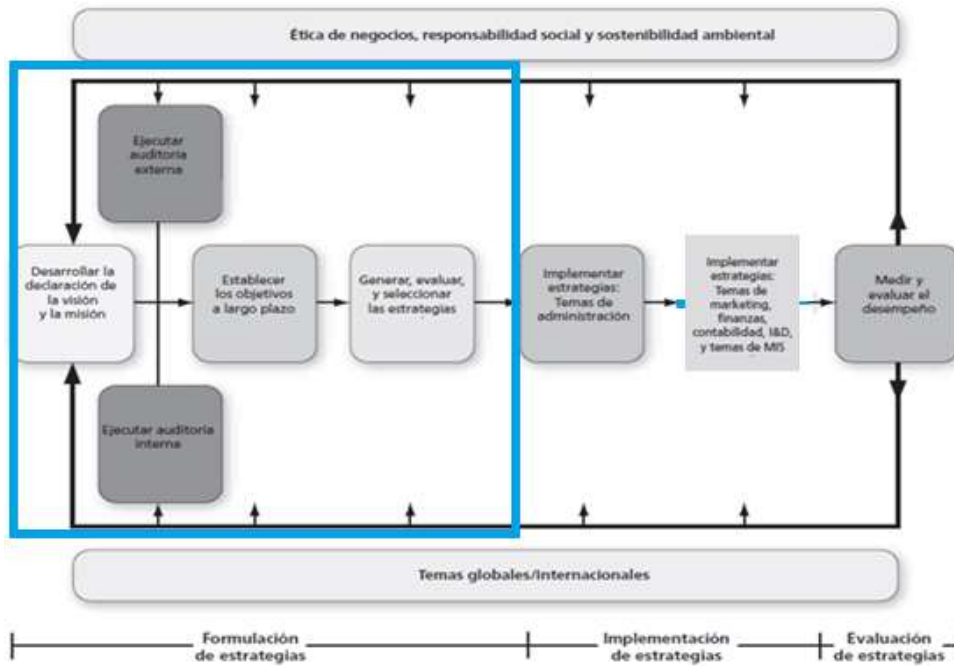


Ilustración 36 -- Marco de referencia Administración estratégica

Ahora si se analiza el proceso de toma de decisiones soportadas por análisis de datos, el modelo CRISP-DM, se plantea una etapa de conocimiento de negocio la cual parte de entender la razón de ser del negocio, su misión, y visión y a partir de éste comprender el estado actual tanto financiero como operativo y las particularidades de cada proceso.

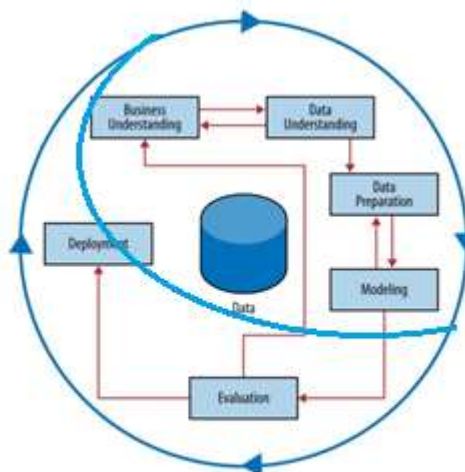


Figure 2-2. The CRISP data mining process.

Ilustración 37 -- CRISP Data Mining Process



También es necesario comprender cuál es la visión del negocio, qué se quiere ser es la base del entendimiento de negocio y de la formulación de preguntas, con estas preguntas se busca modelar las variables de negocio, entender su comportamiento y ayudar a descubrir hallazgos para la mejora del performance.

Por esto para desarrollar una estrategia de datos desde el marco de referencia de la administración estratégica se debe incluir como bloque dentro de la formulación de estrategias y después de la declaración de misión y visión el proceso CRISP ya que para orientar la decisión y el alcance de los objetivos establecidos se debe conocer el negocio, entender la información que existe disponible, modelar el comportamiento y finalmente a partir de este conocimiento plantear una serie de actividades para conseguir el objetivo, implementarlas y realizar seguimiento en línea o en tiempo real al cumplimiento de los objetivos.

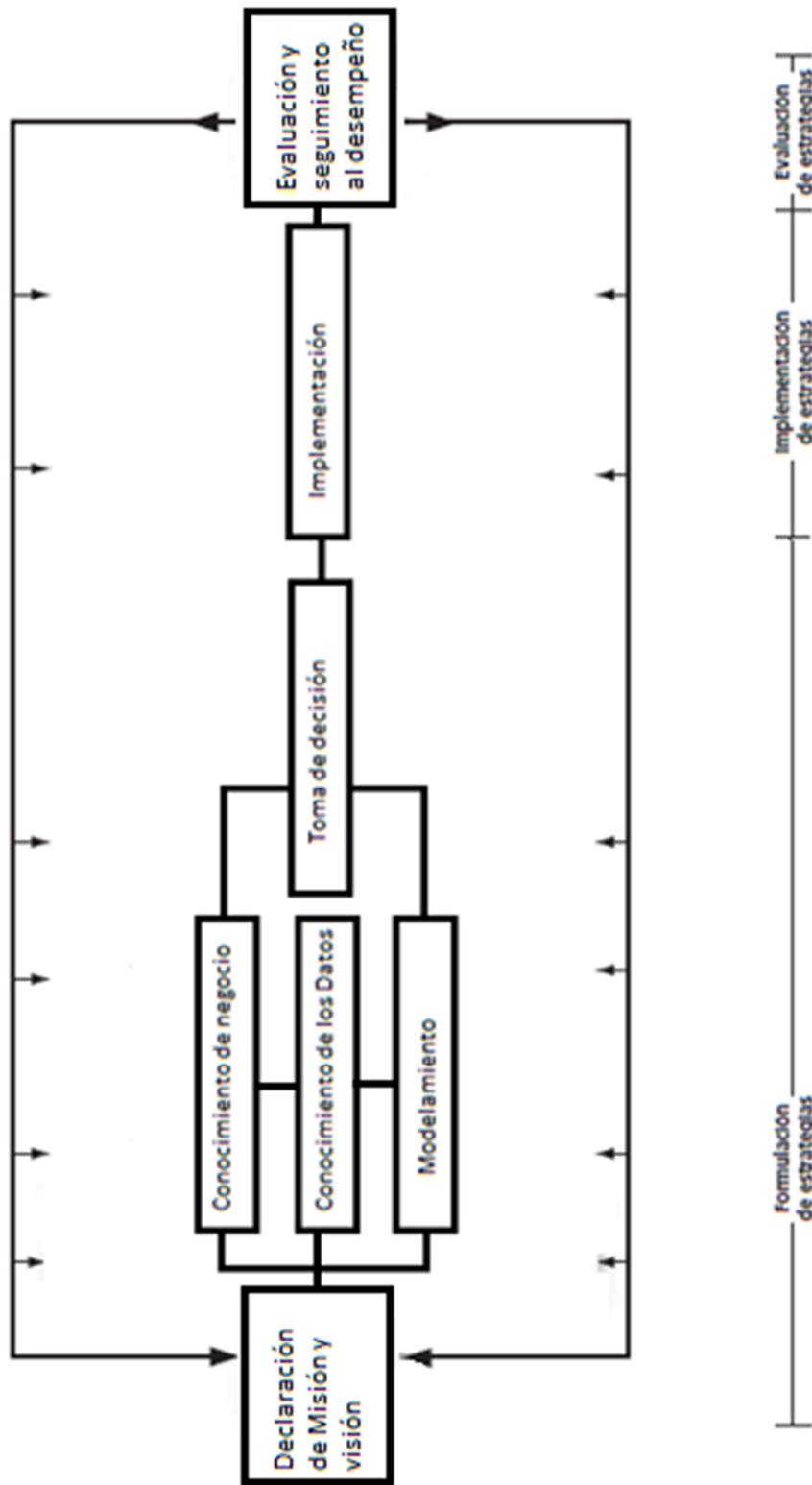


Ilustración 38 -- Propuesta de Estrategias soportadas por análisis de datos en el marco de Administración estratégica.

El conocimiento de negocio permite definir las iniciativas clave para el negocio, normalmente estas iniciativas son: los objetivos de negocio, metas financieras, indicadores, tiempos de respuesta, es decir lo que la organización plantea conseguir y/o mejorar en el corto plazo, normalmente entre 9 y 12 meses y que contribuya a la consecución de los objetivos financieros.

Para el caso de estudio el objetivo planteado es reducir el Churn voluntario en un 20% en los próximos 12 meses, pero el objetivo puede ser cualquiera que esté en consecuencia con los objetivos financieros y que aseguren el crecimiento de la organización y que esté dispuesta a alcanzar en los próximos meses, desde la reducción de reclamos o unidades defectuosas, hasta la reducción de delitos en una ciudad. Los anteriores son ejemplo de la iniciativa que motiva la estrategia de datos y el objetivo que se busca alcanzar en los próximos meses.

Una opción para identificar iniciativas clave de la organización es enfocarse en los objetivos relevantes de la organización enunciados normalmente en el marco de la administración estratégica y los cuales se pueden encontrar en: los reportes anuales, presentaciones del directorio, estados financieros, reportes de seguimiento a la operación, entre otros.

Para un proceso de incursión en estrategia con enfoque de datos es necesario tener en cuenta que las iniciativas de negocio deben ser de corto plazo de 9 a 12 meses, debe ser transversal al negocio, es decir afectar a varias áreas de proceso y contribuir a la estrategia general, por ejemplo: aumentar los ingresos, reducir costos, también debe tener un responsable desde el área de negocio que persigue la iniciativa, un responsable del área de TI y fuentes de datos disponibles.

La fase de entendimiento de los datos corresponde a identificar y comprender las entidades físicas que se requieren explorar, predecir, mejorar, o influenciar. Son las tablas y fuentes de datos relacionadas que se usan para responder a las preguntas claves. Si lo que se busca es comprender y/o modelar el comportamiento de los clientes la entidad clave de negocio sería clientes, si lo que se quiere es comprender o mejorar el funcionamiento o performance de un punto de venta la entidad de negocio sería Punto de venta y sus relaciones.

La preparación de los datos consiste en construir la entidad con la información necesaria para responder a las preguntas de negocio, el proceso de toma de decisiones soportado por datos incluye este punto porque al existir en la actualidad diversas fuentes de datos los formatos y presentación no se ajustan a las necesidades, para esto se requiere equipos de IT con las habilidades para realizar la extracción y transformación de estos datos, normalizar y/o estandarizar para que sean aptos para el procesamiento y análisis.

En la etapa de modelamiento se orienta en definir un prototipo que permita responder a las preguntas de negocio y que ayude a comprender el comportamiento de la variable de estudio, en nuestro caso del cliente, predecir su comportamiento y pueda sugerir decisiones a tomar dependiendo de la información explorada.

En esta etapa se discuten y definen los algoritmos que pueden ayudar a soportar una decisión, el algoritmo seleccionado requiere un proceso colaborativo entre el equipo de tecnología y los usuarios clave del proceso, este paso del proceso se conoce como Data Science (Ciencia de datos) el cual consiste en matemáticamente describir el comportamiento de la variable de estudio.

Tener una evaluación de las decisiones basada en el análisis matemático y/o estadísticos de los datos permite observar de manera lógica y objetiva un problema, analizar este problema y buscar soluciones prácticas, al tener un equipo y herramientas disponibles para la evaluación del modelo permite comparar soluciones alternativas y seleccionar la mejor.

Para lograr el compromiso de los sponsors es necesario poder soportar y defender una posición particular cuando existen opiniones subjetivas y sesgadas sobre un problema. Esta etapa permite comparar el desempeño de diferentes soluciones y poder explicar las diferentes opciones y alternativas generando confianza en el grupo que se encarga de asignar recursos y tomar la decisión.

Una vez definido el modelo se tiene la información suficiente para diseñar un plan de acción y realizar las mejoras en el performance que optimicen la operación y permitan la consecución de los objetivos o iniciativa de negocio planteada.

El planteamiento de una estrategia es insuficiente para que esta sea exitosa, se requiere el compromiso de toda la organización para la ejecución e implementación de la estrategia esto implica cambios culturales y estructurales que soporten la estrategia.

Para que las organizaciones puedan dar el paso hacia una estrategia de datos y asumir el reto de la transformación digital es necesario lograr cambios culturales u organizacionales como; dar acceso a toda la organización de la información disponible, las personas con posiciones de decisión dentro de la organización deben sentirse cómodos con el uso de datos para guiar la toma de decisiones y que los algoritmos permitan sugerir acciones a implementar, generar propiedad intelectual de los desarrollos de analítica, modificar la estructura organizacional de forma que aparezca una oficina que dirija la estrategia de datos (CDO Chief Data Officer), un equipo de científicos de datos y planes de capacitación y entrenamiento en pensamiento analítico de manera que todas las áreas comprendan la necesidad de soportar decisiones en los datos.

Lograr que el pensamiento analítico llegue a todas las áreas genera la sinergia y conexión necesaria para que aparezcan las iniciativas de negocio en los niveles operativos y con ellas mejoras en performance y factores que logren diferenciar a la organización.

Una vez diferentes roles entiendan los principios fundamentales de la ciencia de datos, las ideas y la creatividad para encontrar nuevas soluciones vendrán desde todas las direcciones de la organización, desde los ejecutivos para evaluar potenciales líneas de negocio, del directorio que es responsable de las ganancias y las pérdidas de la compañía, de los gerentes encontrando puntos críticos dentro del proceso de negocio, y desde el resto de empleados que se enfrentan con la operación y tienen conocimiento detallado del proceso.

La era digital impone nuevos retos para todas las organizaciones sean estas del sector gobierno o privadas. Los nuevos modelos de negocio están irrumpiendo en el ecosistema y desplazando a los negocios convencionales, caso UBER, Netflix, AirBnb. Este proceso de transformación exige de los líderes la responsabilidad de recorrer el camino de la transformación o el riesgo de desaparecer.

Es importante seguir una metodología que involucre a diferentes niveles de la organización tal como se planteó en este documento ya que ayuda a identificar las preguntas claves de negocio, priorizar la solución de estas, asegurar el uso de tiempo y recursos de la forma más efectiva y permite además que un proyecto de analítica de datos no sea un proyecto más del área de TI sino la forma de transformar el negocio y adecuarlo a las necesidades del siglo XXI.

En consecuencia con lo anterior no es suficiente tener herramientas para el análisis de datos y gente preparada para enfrentarse a estas herramientas, se requiere una visión estratégica del negocio que involucre el análisis de datos para apoyarse en ellos en la toma de decisiones y la búsqueda de objetivos.

Muchos de los limitantes para que una compañía inicie el proceso de transformación digital y fortalezca la estrategia con enfoque en los datos son basados en supuestos como el costo de las herramientas, en el caso de estudio Estrategias de Retención de clientes soportada por análisis de datos, objeto de esta tesis fue desarrollado en su totalidad con herramientas gratis y de código abierto, Python para el caso de procesamiento de datos y Elastic Search para el análisis gráfico exploratorio.

La gran mayoría de compañías actualmente son compañías que almacenan datos, en sus sistemas de información; CRMs, ERPs, sistemas de contabilidad, de recursos humanos, los datos están en todas partes, en las redes sociales, en internet, en aplicaciones como las del clima, el tráfico, mercado bursátil, bases abiertas de investigación y gobierno, entre otras. Las fuentes están disponibles solo falta la disposición de negocio para basarse en el análisis, ser creativo, integrar diversas fuentes y sobre todo una disposición de negocio para guiar la estrategia a partir del análisis de datos.

Otra barrera al iniciar es la falta de experiencia en sistemas, adicional porque el core de negocio posiblemente no sea la tecnología, y que existen competidores que ya lo están haciendo, este argumento con mayor razón debe concluir en un rápido accionar hacia la transformación, al menos para estar al mismo nivel de la industria en la que se participa.

Es claro que un cambio disruptivo como el que se propone debe involucrar aspectos como el plan de capacitación de los empleados y un cambio estructural que cree un departamento de data science y la figura del CIO como

responsable de la tecnología y los sistemas de información tenga mayor peso en la estrategia que permita visionar los cambios y soportar la puesta en marcha de la estrategia.

En el caso del plan de capacitación existen plataformas On line que ofrecen cursos de Data Science, analítica de datos y arquitectura de big data, que permita actualizar la fuerza laboral pero antes de tener un personal capacitado es necesario que los directores y responsables de las diferentes áreas también piensen de forma analítica y comprendan el valor que aporta al a compañía soportar sus decisiones en los datos de esta forma se garantiza que el cambio sea organizacional y no solo de formar o tener personal con las habilidades.

Ad portas de la cuarta revolución como se ha llamado a esta época en la que la tecnología y la información digital está impulsando cambios en la sociedad, la manera como nos comunicamos, como interactuamos, como nos relacionamos con los demás, vale la pena cuestionarse sobre el futuro de las organizaciones y que cambios se deben asumir hoy para mantenerse vigente y no desaparecer, la respuesta a esta pregunta está en la adopción no de la tecnología sino en la implementación de una estrategia con enfoque de datos que permita generar los cambios culturales y organizacionales necesarios para esta transformación.

## 8. Conclusiones

El objetivo principal de esta tesis fue mostrar la relación entre la implementación de herramientas para el procesamiento y análisis de datos y las estrategias de retención de clientes en la industria de las telecomunicaciones.

Como se evidenció durante el desarrollo de este caso de estudio e investigación, el análisis de datos juega un papel fundamental en el desarrollo de las estrategias de retención de cliente, ya que el conocimiento del negocio, en conjunto con la información de los clientes y el entorno, y esto; apoyado con el análisis gráfico y estadístico de esta información, le proporciona a las compañías de telecomunicaciones los inputs necesarios para actuar de forma anticipada generando valor.

El rápido cambio de la tecnología y la velocidad con la que se produce información ha generado la necesidad en las compañías de modificar sus procesos en consecuencia con estos cambios; adoptar nuevas tecnologías y herramientas que permitan mejorar el procesamiento de datos, adquirir personal capacitado para la operación de estas herramientas, entre otras.

Aunque estos puntos significan un reto para las compañías por temas económicos y la consecución de personal capacitado, estos no son los principales retos en las organizaciones en el momento de empezar a apoyar sus estrategias en los datos.

El principal reto es organizacional ya que implica realizar cambios estructurales, políticos y culturales dentro de las compañías. Una de las principales limitantes es ir dejando la responsabilidad en la toma de decisiones a las herramientas y esto tiene un costo desde la construcción de organización ya que implica cambiar la forma en la que se opera y sugiere eliminar responsabilidades a quienes la encabezan.

La implementación de una estrategia de datos exitosa exige a los líderes poder identificar, conservar, reforzar y utilizar aspectos de la cultura existente que apoyen las nuevas estrategias propuesta, pero también reemplazar aquellos aspectos de la cultura actual que no se alinean con la nueva estrategia.

Tener la habilidad de crear nuevos conjuntos de valores, mitos, creencias y demás aspectos que permitan su despliegue o implementación los cuales



requieren cambios en la estructura organizacional ya que la estructura determina la forma en que se establecen los objetivos, se conciben las políticas y principalmente cómo se asignan los recursos.

El desarrollo de este trabajo permitió identificar patrones que permitieron definir una metodología de implementación de Data Science como desarrollador de estrategias de negocio aplicable a la retención de clientes. En la cual se integraron conceptos de la administración estratégica para modelar un proceso de generación de estrategia soportada por análisis de datos.

## 9. Trabajo a futuro

El desarrollo de este trabajo deja un camino abierto para la investigación en procesos de cambio tecnológico, toma de decisiones y cómo se relacionan estos con la administración estratégica, involucrando a los procesos de cultura y estructura organizacional, desarrollo de políticas, manejo de recursos entre otros temas de interés organizacional.

Cada vez más las empresas empiezan a incluir dentro de su estructura el rol del líder de datos, pero al mismo tiempo se generan cuestiones sobre posicionamiento dentro de la estructura, responsabilidades, canales de comunicación y configuración de los equipos de trabajo.

Un cambio tecnológico como el que estamos evidenciando donde aumenta el uso de Big Data y herramientas de análisis de información para la toma de decisiones sugiere también grandes cambios culturales; ¿Cuál es el impacto en la cultura organizacional del cambio tecnológico propuesto por el Big Data? ¿Hasta qué punto el éxito de una organización en la adopción del Big Data depende de su capacidad de aceptar el cambio?

Otro de los temas en tendencia es el impacto de la transición hacia métodos automatizados de operación y toma de decisiones en la fuerza laboral. Muchas investigaciones sugieren que se reducirán los puestos de trabajo, otras al contrario, advierten sobre la necesidad de personal capacitado en estos temas indicando que esto no pondrá en riesgo la contratación, pero si exige de los gobiernos dirigir sus objetivos a atender las necesidades de esta transformación, como asumir retos educativos, políticas de usos de datos, bancos de datos abiertos, entre otros.

A partir de esta investigación apoyada en el caso de estudio: “Desarrollo de estrategias de atención a clientes en la industria de las telecomunicaciones soportada por herramientas de procesamiento y análisis de datos (Data Science)”, en el cual se muestra cómo con la ayuda de herramientas de análisis de datos se puede predecir cuáles clientes son propensos a cancelar su contrato y permitiendo diseñar estrategias para su retención, se dejan preguntas aún sin responder y que contribuirían en el debate sobre el uso del Data Science, incluso, en tareas en las que los seres humanos nos creíamos irremplazables como lo es la toma de decisiones.

## 10. Bibliografía

- Barga, R. (2014). Predictive Analytics with Microsoft Azure Machine Learning. En R. Barga, *Predictive Analytics with Microsoft Azure Machine Learning*. New York: Springer Science+Business Media New York, 233 Spring Street, 6th Floor, New York, NY 10013.
- Ben Walker, M. E. (2013). *VOUCHER CLOUD*. Obtenido de V CLOUDNEWS:  
<http://www.vcloudnews.com/every-day-big-data-statistics-2-5-quintillion-bytes-of-data-created-daily/>
- Bernard Marr, 2. (2017 ). *Data Strategy* . London United Kingdom : KoganPage.
- BID. (Julio de 2018). *Integración y Comercio BID (Banco Interamericano de Desarrollo)*. Recuperado el Septiembre de 2018, de El impacto de la Inteligencia Artificial en la Economía Digital : CLOUD.MAIL.IADB.ORG
- David Fred R. (2013). *Conceptos de Administración Estratégica 14E*. México: PEARSON EDUCACIÓN.
- David, F. R. (2013). *Conceptos de Administración Estratégica 14E*. México: PEARSON EDUCACIÓN.
- Estatista. (2018). *www.Estatista.com*. Recuperado el 23 de Abril de 2018, de [www.Estatista.com: https://www.statista.com/statistics/816735/customer-churn-rate-by-industry-us/](https://www.statista.com/statistics/816735/customer-churn-rate-by-industry-us/)
- Frost, S. &, & Sullivan, F. &. (2016). *TOTAL TELECOMMUNICATIONS SERVICES MARKET*. México: Frost & Sullivan.
- Hansson, S. O. (2005). *Decision Theory A Brief Introduction, Department of Philosophy and the History of Technology*. Stockholm: Department of Philosophy and the History of Technology Royal Institute of Technology (KTH).
- IBM Offering Information, I. (October de 2015). Obtenido de <https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=YTC04035USEN>.
- Jonathan Baron, Rex V. Brown. (1991). *Teaching Decision Making To Adolescents*. New Jersey : Lawrence Erlbaum associates.inc. Publishers ISBN 0-8058-0497-8.
- Laudon, K. C., & Jane P Laudon. (2012). *Sistemas de información gerencial. Decimosegunda edición*. Mexico: Pearson.
- Marr, B. (2017). Data Strategy. En B. Marr, *How to Profit from a World of Big Data, Analytics and the Internet of Things*. London: Printed and bound by CPI Group (UK) Ltd, Croydon CRO 4YY.

- Predictive Analytics Experience. (2016). *INFÓRMESE S.A.S.* Obtenido de INFÓRMESE S.A.S:  
[www.informese.co](http://www.informese.co)
- Provost, Foster; Fawcett, Tom. (2013). *Data Sciences For Bussiness* . En F. Provost, & T. Fawcett, *Data Sciences For Bussiness* (págs. Posición 390 - 393 Kindle). O'Reilly Media .
- Salvador, F. (2014). <https://www.ie.edu>. Recuperado el 2017, de <https://www.ie.edu>:  
[https://www.ie.edu/fundacion\\_ie/Comun/Publicaciones/Publicaciones/Big%20Data%20OESP%207.pdf](https://www.ie.edu/fundacion_ie/Comun/Publicaciones/Publicaciones/Big%20Data%20OESP%207.pdf)
- Statista. (Septiembre de 2017). <https://www.statista.com>. Recuperado el 23 de Abril de 2018, de <https://www.statista.com>: <https://www.statista.com/statistics/524074/wireless-providers-churn-rate-in-brazil/>
- Sven Ove Hansson. (2005). *Decision Theory A Brief Introduction*. Stockholm: Department of Philosophy and the History of Technology.
- TM Forum . (2015). *Business Process Framework (eTOM) R15.0.1*. Morristown, NJ : TM Forum .
- Waters, D. (2011). *Quantitative Methods For Business* (Fifth edition published 2011 ed.). Harlow, England: Pearson.

## 11. Anexos

### 11.1 Anexo A Cuestionario entrevista

#### Cuestionario

- ¿Cuál es el porcentaje de clientes que dan por terminado el contrato mensualmente?
- ¿Qué estrategia realizan para la retención de clientes?
- ¿En qué medida usa analítica de datos para definir estrategias de negocio o mejorar la productividad?
- ¿Ha mejorado el rendimiento de la compañía con el uso de estas tecnologías?
- ¿Tienen algún método para predicción de casos de Retiro?
- ¿En caso de ser positiva la anterior pregunta ha mejorado el performance de la compañía?

## 11.2 Anexo B Script de evaluación de vecinos Cercanos Python.

```
# -*- coding: utf-8 -*-  
  
Created on Tue Oct 22 15:45:22 2019  
  
@author: quintero  
  
#Se obtienen librerías de procesos Python  
  
import warnings  
  
import itertools  
  
import numpy as np  
  
from pylab import rcParams  
  
import matplotlib.pyplot as plt  
  
warnings.filterwarnings("ignore")  
  
plt.style.use('fivethirtyeight')  
  
import pandas as pd  
  
import statsmodels.api as sm  
  
import matplotlib  
  
import sklearn  
  
from sklearn import neighbors  
  
from sklearn.neighbors import KNeighborsClassifier  
  
from sklearn.model_selection import train_test_split  
  
from sklearn import datasets  
  
  
  
matplotlib.rcParams['axes.labelsize'] = 14  
  
matplotlib.rcParams['xtick.labelsize'] = 12
```

```
matplotlib.rcParams['ytick.labelsize'] = 12
```

```
matplotlib.rcParams['text.color'] = 'k'
```

Se realiza limpieza de los datos

```
def remove_outlier(df_in, col_name):
```

```
    q1 = df_in[col_name].quantile(0.25)
```

```
    q3 = df_in[col_name].quantile(0.75)
```

```
    iqr = q3-q1 #Interquartile range
```

```
    fence_low = q1-1.5*iqr
```

```
    fence_high = q3+1.5*iqr
```

```
    df_out = df_in.loc[(df_in[col_name] > fence_low) & (df_in[col_name] < fence_high)]
```

```
    return df_out
```

```
def quantil25 (series):
```

```
    q = series.quantile(0.25)
```

```
    return q
```

```
def quantil50 (series):
```

```
    q = series.quantile(0.5)
```

```
    return q
```

```
def quantil75 (series):
```

```
    q = series.quantile(0.75)
```

```
    return q
```

```
#Get DataSet from File
```

```
df = pd.read_csv('Churn_Data_19072018.csv', sep=';', encoding='latin-1',
error_bad_lines=False)
```

```

#Clean Data Frame

df.columns = df.columns.str.strip()

#saos = remove_outlier(saos,'Esfuerzo')

df = df.dropna(how='any',axis=0)

# Selecciona menores a 16

indexNames = df[df['EDAD'] <= 16].index

# Delete these row indexes from dataframe

df.drop(indexNames , inplace=True)

# Selecciona mayores de 90

indexNames = df[df['EDAD'] > 90].index

df.drop(indexNames , inplace=True)

#saos = saos.dropna(thresh=1000, axis='columns')

#print(saos.describe())

print(df)

print(df.describe())

print(df.info())

#print(df.describe())

# Definición de columnas a borrar Categorical DATA

cols = ['LOCALIDAD','TIPO_SERVICIO', 'LOCALIDAD', 'TIPO_DE_COBRO', 'SHAPE', 'DESCRIPTION',

        'TIPO_DE_COBRO','GENERO']

df.drop(cols, axis=1, inplace=True)

```

```
#Create feature and target arrays

#Cargar el archivo el campo data está definido como un arreglo con los datos

#es decir [att1,att2,att3...attn]

#El campo target solo es un arreglo con los resultados

y = df['ESTADO_SUSCRIPCION'].values

X = df.drop('ESTADO_SUSCRIPCION', axis=1).values

# Split into training and test set

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.3, random_state=42,
stratify=y)

# Create a k-NN classifier with 7 neighbors: knn

knn = KNeighborsClassifier(n_neighbors=7)

# Fit the classifier to the training data

knn.fit(X_train,y_train)

# Predict and print the label for the new data point X_test,

new_prediction = knn.predict(X_test)

print("Prediction: {}".format(new_prediction))

# Print the accuracy

print(knn.score(X_test, y_test))

# Setup arrays to store train and test accuracies

vecinos = np.arange(1, 9)

train_accuracy = np.empty(len(vecinos))

test_accuracy = np.empty(len(vecinos))
```



```

# Loop over different values of k
for i, k in enumerate(vecinos):

    # Setup a k-NN Classifier with k neighbors: knn
    knn = KNeighborsClassifier(n_neighbors=k)

    # Fit the classifier to the training data
    knn.fit(X_train,y_train)

    #Compute accuracy on the training set
    train_accuracy[i] = knn.score(X_train, y_train)

    #Compute accuracy on the testing set
    test_accuracy[i] = knn.score(X_test, y_test)

# Generate plot
plt.title('k-NN: Evaluación de Número de vecinos')
plt.plot(vecinos, test_accuracy, label = 'Exactitud del Set de testeo')
plt.plot(vecinos, train_accuracy, label = 'Exactitud del Set de entrenamiento')
plt.legend()
plt.xlabel('Número de Vecinos')
plt.ylabel('Exactitud - Accuracy')
plt.show()

```