

**UNIVERSIDAD DE CHILE
FACULTAD DE ECONOMÍA Y NEGOCIOS
ESCUELA DE ECONOMÍA Y ADMINISTRACIÓN**

**“MODELOS DE NEGOCIOS BASADOS EN DATOS:
DESAFÍOS DEL *BIG DATA* EN LATINOAMÉRICA”**

**Seminario para optar al título de
Ingeniero Comercial, Mención Administración**

Participantes:

Myrla Alcaíno Ruiz

Valeska Arenas Miranda

Felipe Gutiérrez Baeza

Director:

Claudio Bravo-Ortega

Profesor Guía:

David Díaz Solís, Ph. D.

Santiago, Noviembre 2015

Agradecimientos

Agradecemos al profesor David Díaz, quién confió en nosotros desde el primer momento y que, con su paciencia, apoyo absoluto y exigencia guió este trabajo, permitiéndonos lograr todos los objetivos propuestos al momento de decidir el comienzo de esta tesis.

Por supuesto, nuestras familias tienen un lugar especial en estos reconocimientos, han sido actores importantes desde nuestros primeros días en el colegio y durante los últimos días en la universidad. Este trabajo refleja las enseñanzas, el cariño y el sustento que nos han otorgado y, si no fuera por ustedes, no seríamos las personas que somos y, esta tesis sin duda no se hubiera concretado.

Queremos agradecer también, a los emprendimientos y emprendedores que, con su amable y entusiasta disposición a atendernos, contribuyeron a confeccionar este estudio. De igual manera, agradecemos a Start-Up Chile quienes sin aprensión ni desconfianza nos facilitaron su información.

Finalmente, es importante considerar en estas palabras a quienes nos ayudaron durante todo este proceso, desde nuestros amigos y compañeros hasta los funcionarios de la facultad que, con su disposición y simpatía, nos dieron energías para continuar día a día con nuestro trabajo.

Esperamos que este trabajo que con tanta dedicación se ha llevado a cabo, genere conocimiento valioso para las personas que hagan uso de éste y fomente la motivación a investigar más sobre este tema que hoy en día está tomando importancia en el mundo de los negocios.

Contenido

Resumen Ejecutivo	4
1. Introducción	5
2. Capítulo I: Revisión Literaria.....	8
2.1. <i>Big Data</i>	8
2.1.1. El aumento exponencial de los datos: ¿Cuál es su relevancia?	8
2.1.2. ¿Qué se entiende por <i>Big Data</i> ?	9
2.1.3. <i>Big Data</i> como una nueva industria y fuente de negocio	10
2.2. Modelos de Negocios.....	12
2.2.1. Definición de Modelo de Negocios	12
2.2.2. Modelos de Negocios Basados en Datos	15
2.3. <i>Startup</i>	18
2.3.1. Definiciones de <i>Startup</i>	18
2.3.2. <i>Lean Startup</i>	19
2.3.3. <i>Startup</i> en Latinoamérica	24
2.4. Principales conclusiones y críticas de la revisión literaria.....	26
2.5. Posicionamiento de la investigación	27
3. Capítulo II: Propuesta de Investigación y Objetivos.....	29
3.1. Propuesta de Investigación y Objetivo General.....	29
3.2. Objetivos Específicos.....	30
4. Capítulo III: Metodología.....	31
4.1. Entrevistas Semiestructuradas.....	33
4.2. Análisis de Páginas Web.....	34
5. Capítulo IV: Resultados	37
5.1. Etapa Exploratoria: entrevistas y análisis de contenido	37
5.2. Etapa Cuantitativa	40
5.2.1. Estadística Descriptiva.....	40
5.2.2. Análisis <i>Clustering</i>	46
5.2.3. Árboles de Decisión.....	48
5.2.4. Mapa de Posicionamiento.....	57
6. Capítulo V: Discusión.....	59

7.	Capítulo VI: Conclusiones	66
8.	Referencias	70
9.	Anexos	74
9.1.	Anexo N°1: Listado de empresas entrevistadas	74
9.2.	Anexo N°2: Boceto Entrevista semiestructurada <i>Startups</i>	75
9.3.	Anexo N°3: Material de apoyo Entrevista semiestructurada <i>Startups</i>	80
9.4.	Anexo N°4: Listado de empresas analizadas	88

Resumen Ejecutivo

El presente estudio tiene por objetivo identificar cuáles son los principales Modelos de Negocios basados en Datos desarrollados por las *Startups* en Latinoamérica, y determinar qué características definen cada uno de los modelos encontrados. La muestra estudiada incluyó un universo de 82 empresas, pertenecientes a 7 países listadas en Angellist y Start-Up Chile. Dentro de los objetivos específicos, se busca establecer qué patrones de innovación están desarrollando estas empresas y la forma en que podrían evolucionar. La metodología utilizada aborda la técnica CRISP-DM. Entre las principales etapas que se presentan destaca el Entendimiento de los Datos, en donde se explican las líneas de investigación que este trabajo desarrolló, específicamente, aquellas relacionadas al análisis del tipo cualitativo, a través de entrevistas semiestructuradas y, al del tipo cuantitativo, mediante el análisis de páginas web. Los resultados de entrevistas se examinaron mediante un análisis de contenido, y para el caso de los datos recolectados del análisis web se utilizaron tres tipos de algoritmos, análisis de *Clusters K-means*, *X-means* y *árboles de decisión*. Los resultados fueron representados en un mapa de posicionamiento.

En relación a los resultados encontrados, las entrevistas arrojaron que, respecto de los patrones de innovación, estas empresas no se familiarizan con ninguno de los patrones previamente identificados en el trabajo de (Parmar, Mackenzie, Cohn, & Gann, 2014), sin embargo, la mayoría destaca la importancia de recolectar información desde distintas industrias, la relevancia de asociarse con otras empresas y la posibilidad de estandarizar alguno de sus procesos para venderlos como un producto independiente. Respecto del análisis web, se identificaron seis tipos de modelos de negocios diferentes, los cuales se caracterizaron por realizar más de una actividad clave, enfocarse en análisis más complejos que los reportados en el trabajo previo de (Hartmann et al., 2014) y en realizar procesos de monitoreo. Tanto las fuentes de datos internas, como externas son utilizadas y, la principal propuesta de valor, tiene que ver con la entrega de información y conocimiento. La gran mayoría de las empresas examinadas se dirige al segmento B2B. Adicionalmente, se identificó una nueva actividad clave, relacionada al asesoramiento de las empresas una vez que se entregan los resultados del procesamiento de datos.

1. Introducción

En la actualidad nos enfrentamos a una explosión significativa en la cantidad de datos que se generan y son almacenados, basta con pensar en los datos que diariamente subimos a redes sociales, los datos recabados por los sensores de nuestros teléfonos inteligentes o por cualquier otra plataforma tecnológica. Lo anterior, no sólo impacta a la persona común y corriente, sino que también a las empresas, quienes ahora están insertas en un mundo altamente cambiante y competitivo, el cual es cada vez más dependiente de las Tecnologías de la Información y las Comunicaciones (TICs) para sus procesos comerciales, sociales y políticos; estas tecnologías han facilitado la interconexión e integración de mercados, creando nuevos canales de distribución y transformando profundamente la economía y la sociedad en general y, por ende, la manera en que los intercambios de bienes y servicios son realizados (Díaz, 2015). Por lo tanto, existe un desafío latente en la forma en que las compañías -tanto las actuales como las próximas a formarse- pueden agregar valor a partir de los datos, teniendo en cuenta que, deben adaptar sus modelos de negocios a estas nuevas tendencias. Este último punto es especialmente importante, ya que la adaptación de un modelo de negocios es especialmente complicado en empresas que ya están formadas, pero aún más desafiante para las empresas que se constituirán en el corto plazo. En este sentido, toma relevancia el concepto de *Startup*, debido a que, lo que caracteriza a este tipo de empresa es su flexibilidad y las condiciones de incertidumbre a la cual se debe enfrentar (Ries, 2012). Además, otro punto relevante, estrechamente relacionado con lo anterior es que, el contexto Latinoamericano se ha convertido en un ecosistema atractivo para la proliferación de este tipo de modelo de negocios, destacando las políticas públicas de Brasil y Chile, donde éstas juegan un rol predominante en la generación de incentivos y facilidades para la gestación de nuevas empresas innovadoras (OECD, 2013). Entonces, dadas estas condiciones y, teniendo en cuenta que, este es un tema reciente y en desarrollo, es que nace la motivación de estudiarlo y de aportar al conocimiento de este tipo de empresas y su relación con la gestión y el manejo de los datos.

El presente estudio pretende conocer más profundamente cuáles son los modelos de negocios basados en datos que se están gestando o que ya están presentes en esta región, así como también determinar en qué ámbitos específicos se está innovando -en relación a los datos- y tener

un panorama general sobre esta industria emergente. Para llevar a cabo lo anterior, el capítulo 1 comienza con una exhaustiva revisión de la literatura sobre los temas presentados anteriormente. En primer lugar, se define qué es *Big Data*, su relevancia y cómo éste se ha transformado en una fuente de ventaja competitiva para las empresas. En segundo lugar, se presenta el tópico de modelos de negocios, cómo se define, ahondando en sus principales componentes y en cómo se relacionan con el *Big Data*. Posteriormente, se aborda el tema de las *Startups*, específicamente, señalando cómo está definido este concepto en la literatura, así como también, se describe el Método *Lean Startup* (Ries, 2012), para luego enfocar la temática al contexto Latinoamericano. Este capítulo culmina con las principales conclusiones, brechas y elementos encontrados en la literatura que aún no han sido objeto de estudio, con el fin de posicionar el trabajo y determinar qué elementos serán considerados para el posterior análisis.

En el capítulo 2, se expone la propuesta de investigación mediante la pregunta de investigación, junto con los objetivos, tanto generales como específicos del estudio. En el capítulo 3, se presenta la metodología a utilizar para llevar a cabo el estudio, la cual está basada en la técnica CRISP-DM (Pete et al., 2000), la cual está compuesta, principalmente, por 5 etapas de desarrollo. La primera, corresponde al Entendimiento del Negocio; este paso tiene por objetivo dar a conocer el objeto de estudio, señalando qué se realizará a través de este estudio. La segunda etapa es el Entendimiento de los Datos, en donde se presentan las líneas de investigación que esta tesis abordará, específicamente, a través de dos procesos, entrevistas semiestructuradas y análisis de páginas web. Las siguientes etapas, Procesamiento de Datos y Modelamiento, presentan los pasos que se realizarán para analizar los datos, especificando, los métodos que se utilizarán, los cuales serán: modelos de *clustering*, árboles de decisión y mapa de posicionamiento. Por último, la etapa de Evaluación, presenta las diferentes métricas que permiten medir el desempeño o *performance* de cada una de las técnicas implementadas durante el análisis.

Posterior al desarrollo de la metodología, en el capítulo 4 se exponen los principales resultados obtenidos a través de los análisis realizados para ambas líneas de investigación, enfatizando en que las medidas de desempeño sean lo suficientemente aceptables para validar los modelos. Además, en esta sección se detallan las características principales de los clústeres de empresas encontrados, así como también, se grafican en un mapa de posicionamiento con el fin de

compararlo con el estudio similar realizado en Cambridge por (Hartmann, Zaki, Feldmann, & Neely, 2014). Tales puntos similares o diferentes son exhibidos y desarrollados de manera crítica en el siguiente capítulo 5 de discusión. Para finalizar con esta investigación, el capítulo 6 condensa las principales conclusiones y limitaciones que se experimentaron en su desarrollo, junto con las potenciales aristas de investigaciones futuras en torno al tema.

2. Capítulo I: Revisión Literaria

2.1. *Big Data*

2.1.1. El aumento exponencial de los datos: ¿Cuál es su relevancia?

La globalización, los avances tecnológicos y en especial, el aumento en cobertura e impacto del internet han propiciado condiciones óptimas para el desarrollo masivo de las Tecnologías de la Información y las Comunicaciones (TICs), provocando una caída en los costos, que ha facilitado el acceso a la información a todo tipo de empresas y organizaciones (democratización de la información), permitiéndoles mejorar sus modelos de negocios y las posibilidades de subsistir en el largo plazo (Díaz, 2015). Esta tendencia ha generado un incremento significativo en la cantidad de datos que diariamente son creados y almacenados por estas tecnologías, llegando incluso a estimar, por ejemplo, que para 2020, la humanidad creará, replicará y consumirá cerca de 40000 exabytes de datos en todo el mundo (Gantz, Reinsel, & Shadows, 2012). Sin ir más lejos, y para mejor comprensión de lo mencionado, es pertinente citar las palabras del CEO de Google, Erich Schmidt, quien en 2010 dijo que la humanidad *“Cada dos días de hoy, crea tanta información como la que creamos desde el comienzo de la civilización hasta 2003”*.

Ahora bien, explorando más allá de la cantidad de datos generados, es posible apreciar que el impacto de este nuevo escenario afecta, cada día más, a una mayor cantidad de industrias y ha cambiado el foco en las habilidades requeridas de los futuros profesionales dedicados a las áreas de inteligencia de negocio. Es así como un reporte de McKinsey Global Institute (Manyika et al., 2011) determina por ejemplo que, para el sector del retail, existen un 60% de aumento potencial en sus márgenes de operaciones con el uso del *Big Data* y, en relación a las nuevas capacidades requeridas, la investigación resuelve que, para el caso de Estados Unidos, sacar el máximo provecho a todos los datos producidos exigirá alrededor de 1.5 millones de nuevos administradores y managers especializados en datos.

Finalmente, respecto del gran aumento de las tasas de generación de datos, la evidencia empírica muestra que aquellas empresas, sin importar su rubro, que soportan sus decisiones de negocios en el análisis de datos tienen un mayor desempeño en términos de rentabilidad y productividad (McAfee & Brynjolfsson, 2012). A continuación, se presenta un cuadro resumen de lo expuesto anteriormente, presentado en el paper *Big Data: A Survey* (Chen, Mao, & Liu, 2014)

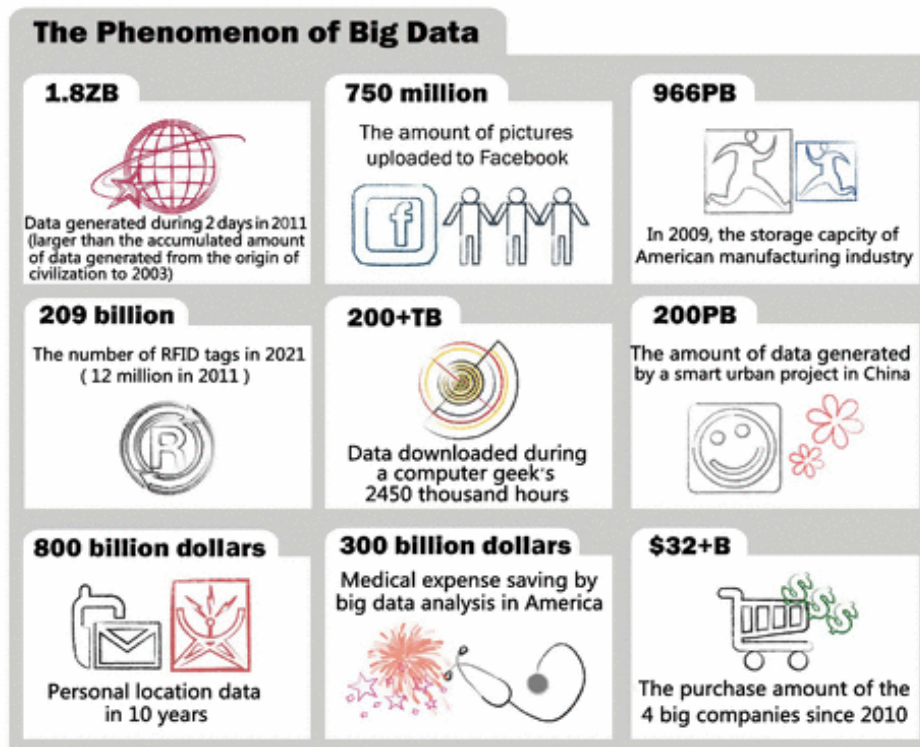


Imagen N°1: El crecimiento continuo de *Big Data*
 Fuente: A Survey (Chen et al., 2014)

2.1.2. ¿Qué se entiende por *Big Data*?

Para comprender este concepto, es recomendable partir por las definiciones utilizadas por (Manyika et al., 2011) quienes precisan *Big Data* como un “set de datos cuyo tamaño está más allá de las habilidades típicas de los software de datos tradicionales para capturarla, almacenarla, administrarla y analizarla”. De esta definición se puede desprender que, en un principio, la forma para describir este concepto contempla, principalmente, las características relacionadas a la cantidad y el volumen de datos que se está generando, sin considerar otras propiedades importantes del *Big Data*. Es por esto que, un año después de lo presentado anteriormente, (Gartner, 2012) agrega otros atributos a esta definición, los cuales hablan de la capacidad de los datos de ser un activos de información de alto volumen, alta velocidad y alta variedad y, siguiendo esta disposición, los autores (Sagiroglu & Sinanc, 2013) concluyen que este concepto es un término para denominar a aquellos datos masivos, que se caracterizan por su cantidad, variedad y complejidad, dificultando, consecuentemente, su almacenamiento, análisis y visualización para

consecución de resultados. Lo anterior va de la mano con lo expuesto por (Vesset et al., 2015) quien menciona que, dada la complejidad de la nueva generación de datos, el tratamiento de la información requiere de grandes esfuerzos en el desarrollo de tecnologías y arquitecturas TICs para asegurar una correcta interpretación y análisis de la gran cantidad de información que hoy se almacena.

Finalmente (Díaz & Zaki, 2015) postulan que el *Big Data* debe ser apreciado, principalmente, por el volumen de datos, su variedad, la velocidad o flujo de creación y procesamiento, la veracidad o calidad de la información y sobre todo, por el **valor** que entregan a las organizaciones, transformándose en un activo estratégico en sí mismo. También, los autores en su trabajo, se refieren al tipo de datos que actualmente se generan, señalando que en su mayoría estos son del tipo no estructurado, lo cual, impacta directamente a la forma en que las empresas los interpretan y analizan.

2.1.3. *Big Data* como una nueva industria y fuente de negocio

Si bien ya se ha hablado acerca de la importancia del *Big Data*, corresponde referirse a su aplicabilidad como una nueva fuente de ventaja competitiva y/o como una nueva fuente de innovación para empresas, ya sean existentes o próximas a formarse. Es así como el (World Economic Forum, 2012) enumera y describe el impacto del *Big Data* en el desarrollo de políticas internacionales enfocadas en distintas áreas de interés mundial, tales como la educación, la salud, la agricultura y los servicios financieros, mencionando cómo las organizaciones relacionadas a estos temas pueden mejorar sus sistemas analíticos, predictivos y de monitoreo. Otro trabajo importante de mencionar es el realizado por (Hartmann et al., 2014) donde sus autores analizan diferentes empresas *Startups*, con el objetivo de generar un marco teórico que identifique los principales modelos de negocios y sus características que se basan en datos, para ser exitosos. Los autores, de acuerdo a su investigación, determinan la existencia de 6 clústeres o grupos de empresas diferentes, los cuales se resumen en la siguiente tabla:

	Propuesta de Valor	Segmentos	Modelo de ingreso	Recursos Clave	Actividades Clave	Ejemplo
Tipo A	Recopilación y agregación de datos	B2C	Publicidad, suscripción y tarifas	Datos disponibles gratuitamente	Distribución de datos	Gnip
Tipo B	Análisis de datos, principalmente descriptivos	B2B	Suscripción o tasa de uso	Datos facilitados por el cliente	Distribuyen datos y visualización de resultados	Sendify y Granify
Tipo C	Generación y análisis de datos	B2B y B2C	Publicidad, suscripción y venta de productos físicos	Datos generados por ellos mismos	Recolección y análisis de datos	Swarmly, GoSquarey Automatic
Tipo D	Análisis de datos disponibles gratuitamente	B2B y B2C	Suscripción, uso, tarifa y publicidad	Datos disponibles gratuitamente	Recolección de datos	Gild y Olery
Tipo E	Agregación de datos desde múltiples fuentes internas	B2B	Suscripción y por el tipo de negocio del cliente	Datos facilitados por el cliente	Agregación y visualización de los datos	Bluenose y Who@
Tipo F	Agregación de datos facilitados desde distintas fuentes	B2B	Suscripción	Datos facilitados por el cliente y datos externos disponibles gratuitamente	Recolección y análisis de datos	FarmLogs y Next Big Sound

Tabla N°1: Resumen Tipos de Modelos de Negocios basados en datos

Fuente: Elaboración propia en base *Big Data for Big Business?* (Hartmann et al., 2014)

Ahora bien, el trabajo anterior nos muestra la realidad única de empresas *Startups*, por ende, es interesante además analizar que está ocurriendo en empresas ya establecidas y de mayor tamaño. En particular, en el estudio de (Parmar, Mackenzie, Cohn, & Gann, 2014) se definen los principales patrones o tendencias de innovación que las grandes empresas de Estados Unidos y el mundo están utilizando para incorporar el *Big Data* dentro de sus negocios. Los patrones encontrados por los autores son: el uso de datos o información de los productos para mejorar a los mismos o crear nuevas formas para agregar valor, la digitalización de activos físicos para la mejor comprensión, tratamiento y presentación de los productos, la combinación de datos e información desde diversas industrias con el objetivo de proporcionar ofertas de negocios integradas al cliente, la generación y condensación de información para que sea utilizada por otros negocios y la

estandarización de los procesos de negocios claves basados en datos para ser vendidos a otras compañías como un producto nuevo. Adicionalmente, los autores concluyen que, si bien estos patrones se pueden desarrollar de forma individual, la tendencia muestra que las grandes empresas combinan estos patrones buscando la sinergia adecuada entre ellos que les permitan generar un mayor valor para sus clientes.

Finalmente, (Díaz & Zaki, 2015) en su estudio muestran estadística descriptiva relacionada a la adopción de las tecnologías de *Big Data* y su situación en Latinoamérica señalando, entre otras cosas que, nuestra región aún se encuentra en etapa temprana de adopción de estas tecnologías y que los países que toman ventaja son Chile, Colombia y Perú. Por otro lado, los autores destacan que en Latinoamérica aún hay mucho espacio por crecer dado que, si bien la generación de datos aumenta en el territorio, la mayoría de las empresas existentes no consideran la *Big Data* como un activo estratégico (TechTarget, 2013) y cerca de un tercio de las empresas de la región ni siquiera tiene planes de incluir esta tecnología en su modelo de negocio (TechTarget, 2014).

2.2. Modelos de Negocios

2.2.1. Definición de Modelo de Negocios

Una parte importante en la cual se sustenta la presente investigación está centrada en la comprensión de los modelos de negocios. Es por esto que, es conveniente precisar y hablar de esta base fundamental para cualquier empresa, la cual explica, entre otras cosas, cómo las organizaciones generan valor, se relacionan con el cliente y compiten en las distintas industrias. Es en este punto donde la definición de modelo de negocios cobra real importancia, ya que este ejercicio puede marcar la diferencia en el desempeño de la organización y su viabilidad en el largo plazo. Ahora bien, ¿Cómo se define e identifica un modelo negocios? ¿Cuáles son sus componentes y cómo se relacionan? A lo largo de la literatura, es posible encontrar numerosas referencias que responden este tipo de preguntas, la mayoría desde distintos enfoques. En el último tiempo, han surgido diversos estudios que han sido capaces de integrar todos estos trabajos y entregar lineamientos comunes acerca de la definición de un modelo de negocios. En términos básicos se aprecian definiciones relacionadas principalmente al modelo económico de la firma, tal como lo señala (Stewart & Zhao, 2000) que enfocan un modelo de negocios como una declaración de cómo la empresa gana dinero y sustenta los flujos de ingresos a lo largo del tiempo.

Junto con lo anterior, también se precisa a un modelo de negocios como *“el diseño de sistemas interdependientes clave que crean y sustentan un negocio competitivo”* (Mayo & Brown, 1999). Pero en términos de un análisis más acabado sobre lo que es un modelo de negocios y las partes relevantes que lo componen, es preciso citar a (Morris, Schindehutte, & Allen, 2005) quienes proponen, luego de analizar más de 15 definiciones - y sus componentes - un marco teórico *“que consiste de tres niveles cada vez más específicos de tomas de decisiones, denominados niveles de “fundación”, “propiedad” y “reglas”* “. En este trabajo, los autores indican que el nivel de fundación es aquel en donde se realizan la toma de decisiones más genéricas relacionadas al negocio y que incluyen *“la propuesta de valor, los consumidores, procesos internos y competencias y cómo la firma genera dinero”*. El siguiente nivel, de propiedad, hace referencia a las decisiones que tienen relación a cómo los componentes se interrelacionan y combinan de forma única para generar ventajas competitivas, es decir, en este nivel es en donde se toman las decisiones estratégicas. Por último el nivel de reglas, alude a los principios básicos que rigen al negocio en los otros dos niveles y que constituye un factor diferenciador entre dos modelos de negocios similares. Finalmente, los autores, utilizan para su estudio seis preguntas guía que ayudan a los emprendedores a identificar los siguientes componentes de un modelo de negocios: propuesta de valor, clientes, capacidades internas, estrategia competitiva, modelo de ingresos y objetivos del inversionista. Adicionalmente al marco teórico propuesto, no se debe obviar el hecho que un negocio posee etapas de crecimiento y de aprendizaje, donde los componentes de estos niveles pueden evolucionar y profundizarse en términos de su complejidad y, por lo tanto, generar mayores ventajas competitivas para la empresa.

Otro estudio importante en la materia y que ha tenido relevancia en el último tiempo, es el trabajo de (Osterwalder & Pigneur, 2010) quienes, en su libro *Generación de Modelos de Negocios* definen que *“un modelo de negocios describe las bases sobre las que una empresa crea, proporciona y capta valor”*. En este libro, los autores introducen el denominado modelo Canvas, el cual, utiliza nueve módulos que ayudan a describir y analizar el modelo de negocios de una empresa y cómo esta funciona para conseguir ingresos. Estos bloques abordan distintas dimensiones en las que se desenvuelve el negocio, tanto a nivel de los clientes, la propuesta de valor u oferta, la infraestructura y la viabilidad económica. Los nueve módulos definidos en este trabajo son: segmentos de mercado, propuesta de valor, canales, relaciones con clientes, fuentes

de ingresos, recursos clave, actividades clave, asociaciones clave y estructura de costos, los cuales se pueden apreciar en la imagen N°2 que se presenta a continuación.

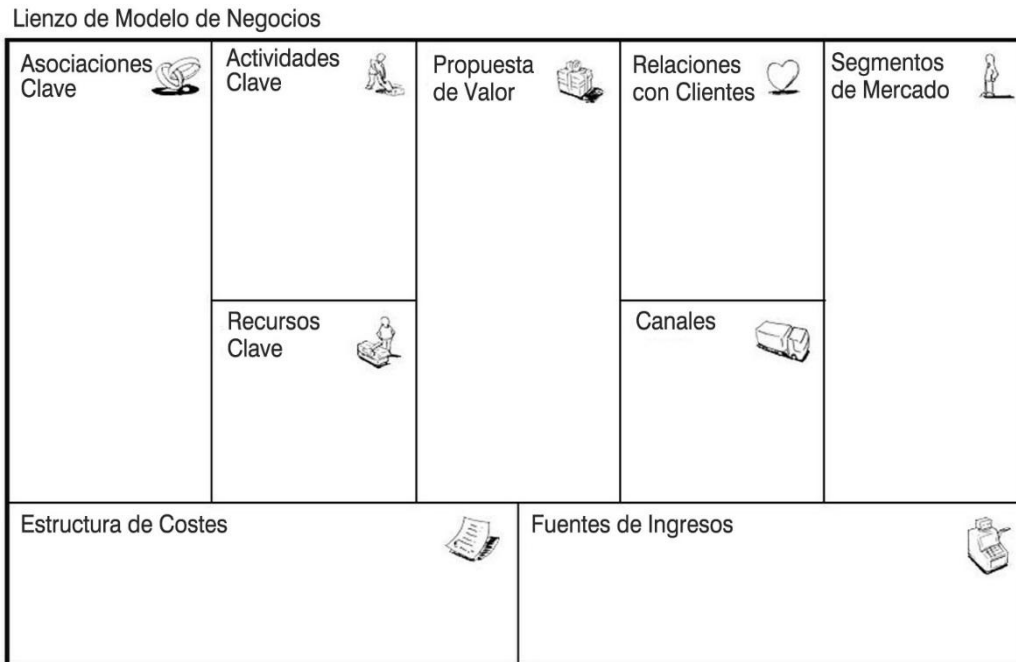


Imagen N°2: Lienzo Modelo de Negocios Canvas

Fuente: Business Model Canvas (Osterwalder & Pigneur, 2010)

Finalmente, los autores abordan diversos tipos de modelos de negocios explicados a través de su metodología Canvas. Además de lo anterior, el libro plantea los nuevos modelos de negocios que están siendo exitosos hoy en día en el mercado, dando una señal que motive y guíe a emprendedores para que se atrevan a comenzar nuevos negocios y a innovar en cada uno de los nueve módulos, por lo tanto, la base teórica de este trabajo permite no sólo comprender un modelo de negocio explicado en base a múltiples dimensiones, sino que también, entrega una guía para diseñar modelos de negocios, a través de diversas técnicas, tales como la ideación, pensamiento visual, aportaciones de clientes, entre otras.

En relación a esta última idea expuesta, un punto relevante a destacar, es el cambio que pueden tener los modelos de negocios cuando se llevan a cabo innovaciones. De acuerdo con (Afuah & Tucci, 2001) destacan que compañías con modelos de negocios innovadores bien formulados ganan ventajas competitivas que se traducen en ingresos más altos que los competidores. Algo

similar es lo encontrado por los autores (Pohle & Chapman, 2006), quienes entrevistaron a 765 CEOs alrededor del mundo preguntándoles con respecto al tópico de la innovación. En su investigación, los resultados muestran que las compañías con enfoque de innovación en sus modelos de negocios tenían un significativo aumento en su margen operativo, en contraste con aquellas compañías que sólo se enfocaban en innovaciones de su producto o servicio. Lo anterior, deja en evidencia la clara ventaja que otorga a las compañías la innovación en los modelos de negocios si es que éstas trabajan orientando sus esfuerzos fuertemente hacia ella.

Finalmente, con el fin de condensar todas las definiciones mencionadas anteriormente, es relevante citar el meta análisis realizado por (Burkhart, Krumeich, Werth, & Loos, 2011), quienes analizaron más de 100 diferentes investigaciones y documentos vinculados a este tópico, con el objetivo de consolidar los puntos comunes que cada trabajo menciona y, de esta manera, establecer una visión general del concepto de modelo de negocios en la última década. Los autores, luego de realizar un filtro de acuerdo a la relevancia de las publicaciones, en base a diferentes criterios entre los cuales se encuentra la cantidad de veces que ha sido utilizado en otras investigaciones y el ranking *Australian Research Council*, seleccionaron los 30 artículos más relevantes y los clasificaron en cinco categorías: *Classification of the underlying literature, Comprehension of business models, Usage of business model, Focus of business models y Representation and evaluation of business models*. De acuerdo a la categorización realizada, el trabajo identifica ciertos patrones que permiten expresar una comprensión común del concepto de modelo de negocios. Entre estos patrones se destaca que el modelo de negocios puede ser aplicado en cualquier tipo de organización, tiene aspectos tanto estáticos, como dinámicos, posee diferentes componentes como la oferta, el mercado, modelo de ingresos y competencias internas, y estos componentes son interdependientes.

2.2.2. Modelos de Negocios Basados en Datos

En esta sección se entrará de lleno al principal tópico de este trabajo, es decir, se darán a conocer las principales características de un modelo de negocios basado en datos, que es la relación de los temas anteriormente expuestos. (Hartmann et al., 2014), basándose en el *Business Model Canvas* propuesto por (Osterwalder & Pigneur, 2010), concluyen que los modelos de negocios basados en datos, poseen seis dimensiones principales, las cuales son: propuesta de valor, segmento de

clientes, modelo de generación de ingresos, recursos clave, actividades clave y estructura de costos. Además, estos autores crearon un esquema en el cual se pueden apreciar todos los componentes dentro de cada una de las dimensiones, el cual se muestra en la siguiente imagen:

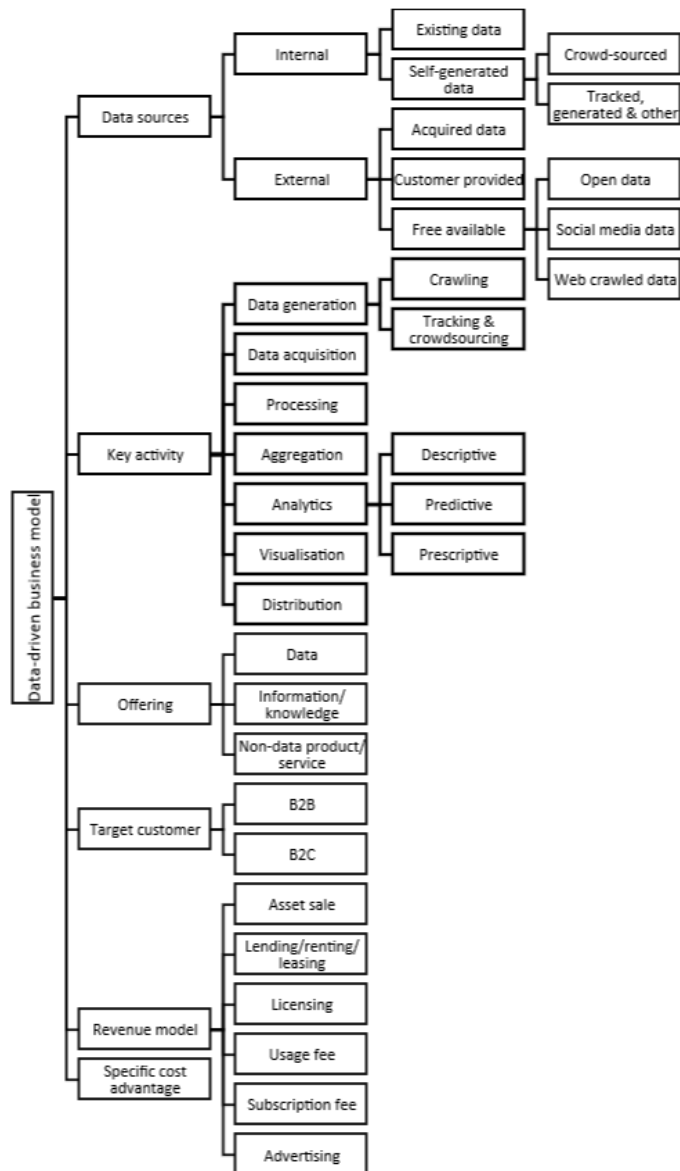


Imagen N°3: Framework de Modelos de Negocios Basados en Datos

Fuente: *Big Data for Big Business* (Hartmann et al., 2014)

A continuación, se detallarán los componentes de cada una de las dimensiones:

a. Recursos Clave.

Dado que son modelos de negocios basados en datos, este recurso es el que se debe considerar como clave en la organización, sin embargo, pueden existir otros recursos que sean el soporte de los datos. Los datos, pueden ser clasificados en dos tipos de fuentes: internas o externas. Dentro de las fuentes internas de la empresa, se pueden encontrar datos ya existentes o que se desarrollan mediante TICs y/o datos generados por medio de monitoreo, entre las cuales se encuentran el web tracking, los recolectados por sensores o los creados por otros agentes denominado *Crowdsourcing*. Por otro lado, dentro de las fuentes externas se pueden encontrar datos que son proporcionados por el mismo cliente o por socios estratégicos que tenga la organización, así como por la adquisición de datos, que son en su mayoría de uso privado; a esto también se suman los datos que se pueden encontrar libremente disponibles en internet. En relación a estos últimos, poseen la siguiente categorización: Open Data, que son datos libres y disponibles principalmente en bases de datos; *Social Media Data*, aquellos que provienen de los sitios o redes sociales como Twitter y Facebook y *Web Crawled Data*, datos recolectados de diferentes páginas web por medio de un software especializado.

b. Actividades Clave.

Dentro de las actividades clave para un modelo de negocios basado en datos se encuentran: Recolección, Generación, Adquisición, Procesamiento, Agregación, Analítica, Visualización y Distribución de datos. El procesamiento y la agregación de datos, incluye el hecho de transformar, limpiar, sintetizar y dar reportes de lo que contiene un base de datos. Por otro lado, el análisis de datos entrega una comprensión de los datos que permite tomar decisiones para la organización. Existen tres tipos de análisis, estos son del tipo descriptivo, predictivo y prescriptivo.

c. Propuesta de Valor.

Las ofertas que pueden generar este tipo de empresas, son de tres tipos: datos, es decir, la entrega de bases de datos o datos particulares sin un procesamiento previo; información y conocimiento, los cuales surgen del procesamiento y comprensión de los datos y la venta de

productos y/o servicios no digitalizados, que son artefactos o artículos materiales que se venden de forma complementaria a los datos y/o información y conocimiento.

d. Segmentos de clientes.

Los tipos de clientes se clasifican en clientes B2B o clientes tipo empresa y clientes B2C o clientes individuales, donde las organizaciones pueden servir a uno de ellos en forma particular o a ambos tipos de clientes en conjunto.

e. Modelo de generación de ingresos.

Todas las organizaciones, independiente de su modelo de negocios deben poseer al menos una fuente de ingresos, entre las cuales se encuentran seis categorías, pudiendo optar a una o más al mismo tiempo. Las categorías son: venta directa de activos (tangibles o intangibles), arriendo o préstamo, licenciamiento, cobro por uso, suscripción, comisión y publicidad.

f. Estructura de costos.

Esta dimensión tiene relación con los costos que incurre la compañía en el tratamiento u otra actividad relacionada a los datos y si esto le permite poseer algún tipo de ventaja relativa con respecto al costo de usar, recolectar y analizar datos.

2.3. Startup

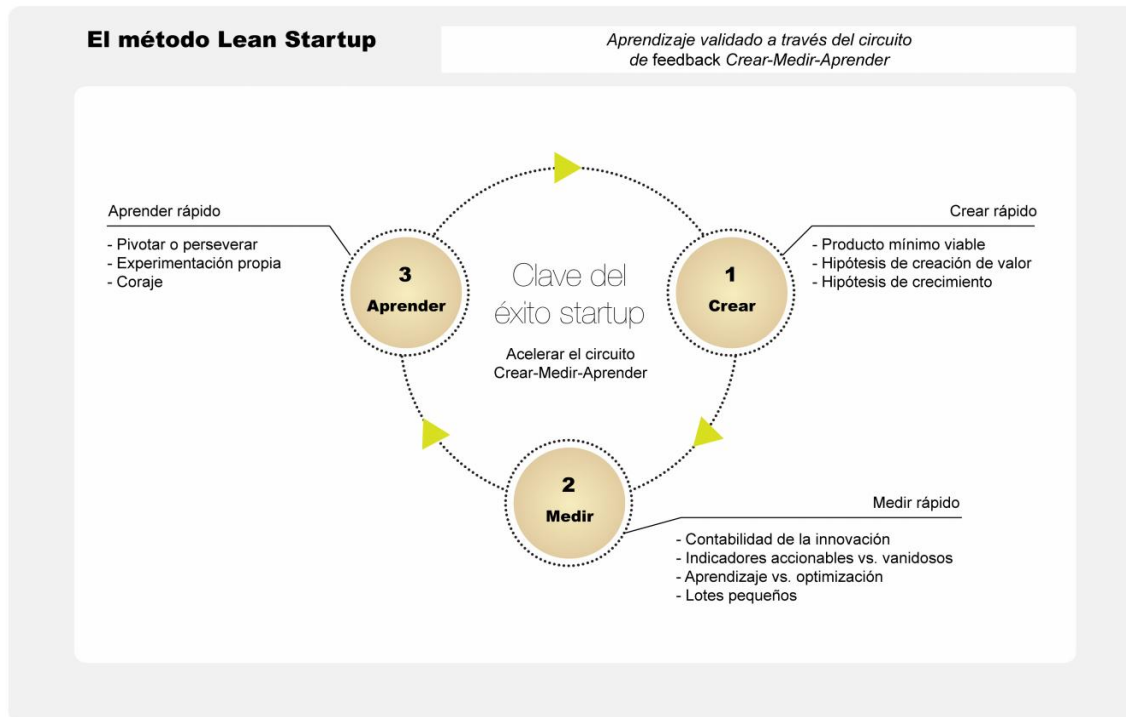
2.3.1. Definiciones de Startup

En primer lugar, es pertinente dejar en claro qué se entenderá por *Startup*, ya que la literatura muestra diversas perspectivas sobre este concepto. Al pensar en esta palabra, sin duda lo primero con lo cual se relaciona es con emprendimiento. Esta palabra deriva del concepto francés *entrepreneur*, cuyo significado es estar listo a tomar decisiones o a iniciar algo y ha evolucionado a través del tiempo, siendo definido por múltiples autores, destacando entre ellos a Joseph Alois Schumpeter, quien en la Teoría del Desarrollo Económico describe la teoría del “espíritu emprendedor”, la cual deriva del hecho que empresarios que crean innovaciones en un medio competitivo deben asumir riesgos, obteniendo beneficios que podrían cambiar a través del tiempo (Schumpeter, 1934). Por otro lado, (Amit & Zott, 2001) describe el emprendimiento o el espíritu

emprendedor como innovador, dinámico, flexible, creativo, orientado al crecimiento y capaz de asumir diversos riesgos. Más concretamente relacionado al concepto de *Startup* se encuentra lo propuesto por Eric Ries en su obra *El método Lean Startup* (Ries, 2012) quien la define como “una institución humana diseñada para crear nuevos productos en condiciones de incertidumbre extrema”. Lo anterior plantea a las *Startups* como instituciones y no sólo como una simple idea o un producto. Finalmente, otro punto importante que señala este autor es que las *Startups* están diseñadas para aprender a crear negocios, lo cual está directamente relacionado con el siguiente apartado donde se presenta un marco teórico orientado a guiar a este tipo de negocios en su constante experimentación.

2.3.2. Lean Startup

Así como Eric Ries en su obra *El método Lean Startup* (Ries, 2012) entrega definiciones concretas sobre lo que es una *Startup*, también ofrece una mirada sobre el hecho que, dados los constantes procesos de cambio, éstas no están libres del fracaso, por lo que necesitan desarrollar aprendizajes continuamente para ir evaluando sus soluciones, asimismo poseer flexibilidad, creatividad y dinamismo en sus decisiones, lo cual es el centro del espíritu emprendedor. Junto con lo anterior, este autor plantea que, dada la naturaleza de este tipo de organización, no pueden evaluarse en base a los márgenes estándares de las empresas consolidadas, es por esto que, define un método para conducir el desarrollo de las *Startups* por medio de los ensayos y errores que viven. Este método llamado “*Lean Startup*” está basado en un circuito de retroalimentación constante de tres etapas: crear, medir y aprender, por el cual todas las *Startup* debería pasar, ya que les servirá para iterar y comprender de manera más profunda a los consumidores. El método en detalle se aprecia en el siguiente esquema:



Esquema N°1: El Método Lean Startup

Fuente: El Método Lean Startup (Ries, 2012)

Del esquema anterior se puede ver que este método comienza con la creación del llamado *Producto Mínimo Viable* (PVM), el cual es una forma de aclarar las concepciones que se tienen sobre lo que los consumidores quieren y cuál es su problema a resolver. Este PVM permite a los emprendedores empezar con el proceso de aprendizaje rápidamente y así entrar en un circuito de *feedback* y de mejoramiento del producto. Luego, el segundo paso es medir a través de lo que el autor define como *contabilidad de la innovación*, dado que las *Startups* poseen elementos espontáneos que no son medibles bajo normas financieras precisas. Finalmente, el tercer paso es el aprendizaje, donde se debe hacer un cuestionamiento de lo que actualmente se está realizando y determinar si es conveniente realizar un cambio importante; este cambio es definido en la literatura como pivote, el cual tiene por objetivo probar una nueva hipótesis del producto y su estrategia. Lo anterior requiere de gran valor, ya que puede ir desde un pequeño reposicionamiento del producto existente, hasta la creación de uno totalmente nuevo, pero es

fundamental que se conserve la visión general que se tiene desde la fundación de la *Startup*, es decir, que idealmente se conserve la idea principal generando algunos cambios para ver florecer la solución que el emprendedor deseó desarrollar desde un comienzo.

La finalidad de este proceso, es rescatar los aprendizajes que se logran para ayudar a los emprendedores a tomar decisiones relacionadas al futuro de la *Startup*, dando como opciones el pivoteo, es decir, cambiar o variar tanto la oferta como otros aspectos del modelo de negocios para progresar, o preservar los lineamientos actuales que presenta la organización. Por lo tanto, el aprendizaje es esencial para evaluar el progreso de este tipo de empresas.

Este último punto sobre el pivoteo ha sido de cierta forma abordado por diversos autores, entre ellos Steve Blank, quien en su artículo *“Why the Lean Startup Changes Everything”* (Blank, 2013) señala que las *Startups* van de fracaso en fracaso, adaptándose e iterando las ideas iniciales a partir del aprendizaje continuo de los clientes, por lo tanto, al crear este tipo de empresas se debe enfatizar en la búsqueda de un modelo de negocios, a diferencia de las empresas ya establecidas, quienes simplemente los ejecutan. Otra arista relevante que introduce este estudio son los factores que influyen en el crecimiento y que restringen el funcionamiento de las *Startups*, entre los cuales se encuentran: los altos costos de obtener a los primeros clientes y los altos costos de crear un producto equivocado; los largos ciclos de desarrollo de la tecnología; número limitado de personas que están dispuestas a asumir riesgos en la fundación y el trabajo en una *Startup* y, por último, la estructura de capital de la industria. Además, en este artículo se destaca la tendencia relacionada al fácil acceso a diferentes fuentes de financiamiento, como por ejemplo, los fondos ángeles, aceleradoras, capitales semillas y *crowdfunding*, que fomentan a la creación de nuevos emprendimientos.

Relacionando el concepto de modelos de negocios al ámbito de las *Startup*, es conveniente mencionar la propuesta que Ash Maurya ha desarrollado en su libro *“Running Lean”* (Maurya, 2012) llamada *“Lean Canvas”*, el cual es una adaptación del modelo Canvas propuesto por Alex Osterwalder (2010). Este concepto de *running* está basado en el hecho que dos tercios de las *Startups* que tienen éxito cambiaron sus planes drásticamente a lo largo del camino (Mullins & Komisar, 2009), por lo tanto, no necesariamente los emprendimientos comienzan con un excelente plan, sino que van iterando y cambiando de planes a lo largo del tiempo lo cual también

está en sintonía con lo propuesto en el método *Lean Canvas*. Dado lo anterior, el análisis del modelo de negocios que se presenta en este libro tiene por objetivo ayudar a los emprendedores en este proceso de cambio, de tal manera de aterrizar el problema que han identificado y la solución que han desarrollado para éste, además, de resaltar la ventaja competitiva y las métricas claves para evaluar el desempeño de la organización. A continuación, en la imagen N°4 se presenta el lienzo de este modelo, donde se destacan los principales cambios respecto al modelo Canvas original.

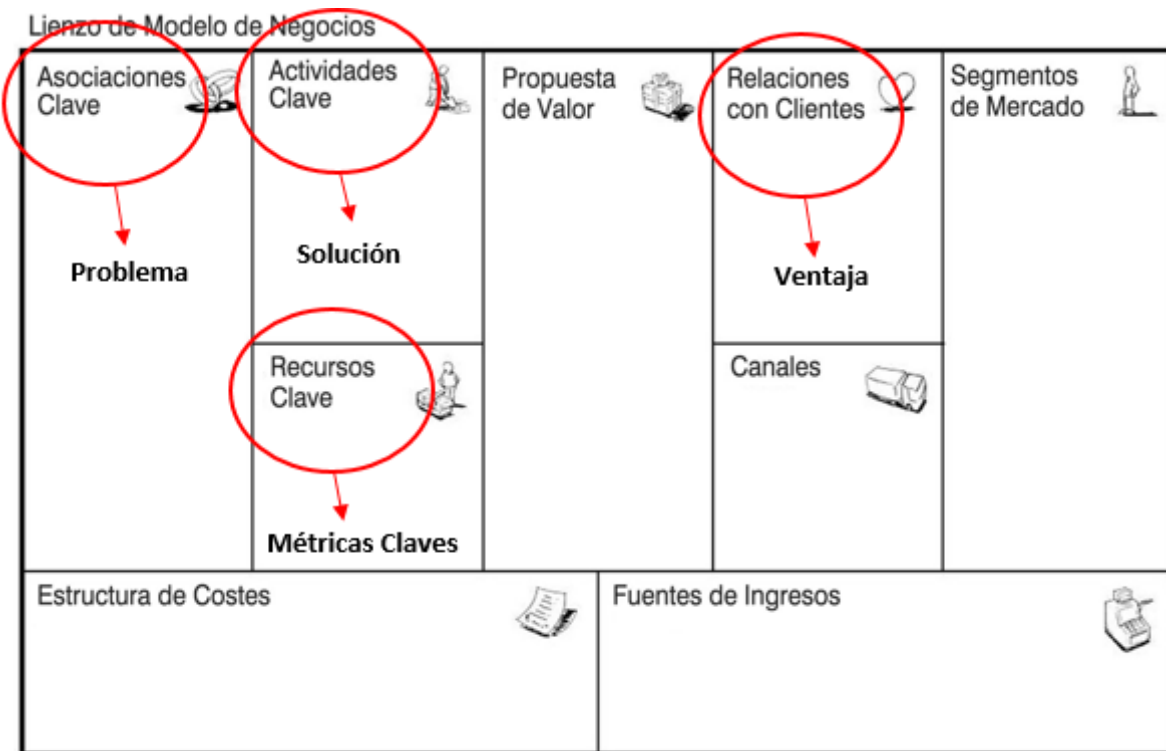


Imagen N°4: *Lean Canvas*

Fuente: *Running Lean* (Maurya, 2012)

Una vez presentado este tema, es pertinente tener una mirada sobre lo que está ocurriendo en el ámbito Latinoamericano en torno al desarrollo de este tipo de empresas, con el fin de presentar aquellos países líderes en este ámbito, que poseen políticas favorables para la formación y crecimiento de las *Startups*. En el siguiente apartado se presenta este tópico.

2.3.3. *Startup* en Latinoamérica

En las secciones anteriores se ha hablado básicamente sobre cómo las *Startups* nacen, se desarrollan y crecen, teniendo en cuenta que éstas están sujetas a rápidos cambios y deben adaptarse. Pero, además de lo anterior, un elemento clave que se considera este trabajo es la importancia y el aporte que realizan este tipo de empresas en el crecimiento de los países. Diversas investigaciones anteriores han abordado este punto y, para el contexto latinoamericano, el documento más destacado es el Estudio del Centro de Desarrollo de la (OECD, 2013), el cual enfatiza, entre otras cosas, que las *Startups* son un agente de cambio importante para generar empleo de calidad, mayor crecimiento e innovación. Asimismo, destaca el rol de éstas en los cambios estructurales que ocurren en la economía de los países al contribuir con la introducción de nuevos productos y servicios intensivos en conocimiento, dinamizando el crecimiento del país y ayudando a sostener la competitividad de las industrias nacionales (OECD, 2013). Por lo tanto, se aprecia el rol central que hoy en día están comenzando a tener las *Startups*, dentro de lo cual surgen disyuntivas sobre cómo apoyarlas y darles incentivos para que sigan desarrollándose. Este último punto ha sido desarrollado por los países latinoamericanos y, en el mismo estudio citado anteriormente se aborda y analiza el rol de las políticas públicas de los países de la región respecto de la creación y expansión de los *Startups*. Específicamente, presenta una visión comparativa de las iniciativas recientes de 6 países - Argentina, Brasil, Chile, Colombia México y Perú - con el objetivo de identificar y promover las buenas prácticas y la implementación de políticas dirigidas hacia la innovación. El trabajo se refiere a cómo los países pertenecientes a la OCDE han adoptado dentro de sus estrategias nacionales el apoyo a las *Startups* como fuente de competitividad e innovación, cuestión que aún no se ha implementado en Latinoamérica, principalmente porque para la OCDE, el desarrollo de *Startups* costa de innumerables barreras a la creación y su expansión, como lo son las relacionadas a la disponibilidad de financiamiento, escaso dinamismo de los sistemas de innovación y la prevalencia de marcos regulatorios que, en muchos casos, dificulta la creación de empresas. Por lo tanto, el trabajo postula que, las políticas públicas, en especial en Latinoamérica, juegan un rol predominante en la generación de incentivos y facilidades para la gestación de nuevas empresas innovadoras. Ahora bien, entre los países que han estado introduciendo a las *Startups* como un eje de innovación y desarrollo productivo, el documento resalta el trabajo realizado por Chile y Brasil y, en el último tiempo, se han sumado países como

Colombia y Perú. A continuación, se presenta un cuadro comparativo entre los países en estudio, que refleja el apoyo directo a las *Startups*:

Categoría	Instrumento	Argentina	Brasil	Chile	Colombia	México	Perú
Financiamiento	Capital semilla	●	●	●	●	○	●
	Inversionistas ángeles	○	●	●	●	○	○
	Capital de riesgo	○	●	●	●	●	○
Servicios de apoyo y capacitación empresarial	Incubadoras	●	●	●	●	●	●
	Aceleradoras	●	●	●	○	●	○
	<i>Spin-offs</i> corporativas	○	○	○	○	○	○
	Transferencia tecnológica y <i>Spin-offs</i> universitarias	●	●	●	○	●	○
	Capacitación empresarial	●	●	●	●	●	●
Marco regulatorio	Facilidad para crear/cerrar empresas	○	○	●	●	●	●
	Tributación y legislación especial	○	●	●	○	●	○

● En operación
● En fase de desarrollo
● Recién creados
○ Necesitan ser creados/reformados

Cuadro N°1: Apoyo directo a las *Startups* en América Latina: Comparación entre países, 2012.

Fuente: *Startup América Latina* (OCDE, 2013)

Finalmente, el trabajo concluye (y presenta propuestas) que, para la realidad de América Latina, el apoyo de las *Startups* requiere de un adecuado sistema de incentivos que tome en cuenta las necesidades de financiamiento, capacitación y regulación en cada una de las etapas de desarrollo de estas empresas (gestación, nacimiento, desarrollo y expansión).

2.4. Principales conclusiones y críticas de la revisión literaria

En relación al primer tópico presentado, se desprende la importancia hoy en día del *Big Data* para todo tipo de compañías, debido al aumento exponencial en la generación de datos, lo cual plantea desafíos en torno a la capacidad de gestión de las empresas para procesar y comprender esta información, a modo de generar valor que les permita crecer y proyectarse en el largo plazo. Esta información se caracteriza por ser variada, por generarse con gran rapidez y por la veracidad, de la cual muchas veces carece. Además, cabe mencionar que, ninguno de los autores revisados se refiere al potencial problema del exceso de datos y cómo las empresas pueden enfrentarlo convirtiéndose en una posible área de investigación.

Por otro lado, respecto al desarrollo del *Big Data* en el escenario empresarial latinoamericano, no existe suficiente evidencia de cómo se está adoptando esta tendencia en el desarrollo de modelos de negocios en la zona, lo que deja en evidencia la falta de una exploración más profunda en este ámbito y, probablemente, la pérdida de terreno que tienen las empresas de la región respecto de su implementación.

Con respecto a los modelos de negocios, si bien existen múltiples definiciones, que algunos autores han tratado de condensar, aún no existe un consenso sobre la definición exacta, sin embargo, todas incluyen ciertos puntos en común relacionados a las características que un modelo de negocios posee, tales como propuesta de valor u oferta, segmentos de clientes, actividades y recursos necesarios primordiales para su funcionamiento. Ahora bien, de acuerdo a lo presentado, las propiedades que distinguen a un modelo de negocios, para una empresa ya establecida no funcionan de la misma manera para los emprendimientos o *Startups*, ya que estos últimos, según lo mencionado, deben adaptar, en muchas ocasiones, la forma en que se desarrolla el negocio debido a que tienen mayor capacidad de innovación, pero a la vez deben enfrentar mayor incertidumbre y riesgos al momento de tomar decisiones. Lo anterior, de cierta manera es abordado por el tercer punto de esta sección que se plantea posteriormente.

En relación al *Big Data* y los modelos de negocios, el constante crecimiento de la información ha facilitado el surgimiento de diversas firmas dedicadas y especializadas en datos, entre las cuales se encuentran empresas expertas en análisis de información, agregación de datos, generación de datos, visualización y distribución, entre otras. Cabe señalar que, las investigaciones relacionadas a

este tipo de modelo de negocios se basan principalmente en la realidad de los países desarrollados dejando fuera del ámbito de estudio la situación en Latinoamérica.

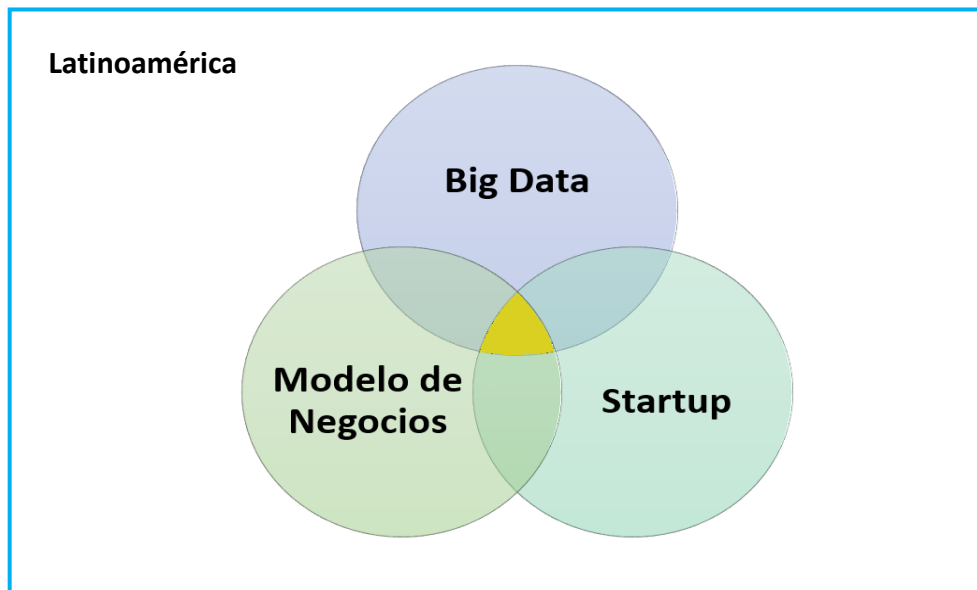
Finalmente, es importante mencionar que para el caso de las empresas especializada en datos, se aprecia la falta de distinción en cómo se definen ciertos elementos del modelo de negocios, como lo es la indiferenciación entre el cliente y el usuario, es decir, la diferencia de roles que puede existir entre quien compra y quien finalmente hace uso del producto o servicio. Además, para el caso de las formas de generar ingresos, se puede ver que no se incluyen modelos de ingresos como el sistema *Freemium*, el cual es ampliamente utilizado hoy en día.

Con respecto a la situación de las *Startups* en América Latina, las investigaciones encontradas distan de una clara perspectiva de la realidad actual del emprendimiento y de cómo se están gestando nuevos tipos de modelos de negocios relacionados al *Big Data*, luces de lo anterior sólo se encuentran en el trabajo realizado por (Díaz & Zaki, 2015). A pesar de esta situación, es posible mencionar que, esta región posee un ecosistema promotor del emprendimiento, ya que países como Chile, Brasil, Colombia y Perú han llevado a cabo políticas públicas destinadas a potenciar la innovación y la generación de nuevas ideas de negocios, lo cual ha provocado una proliferación de las *Startups* y los emprendimiento en Latinoamérica. Por lo tanto, toma importancia el hecho que en la actualidad y en el mediano-largo plazo se generen estudios para analizar cómo es la situación en la región. Junto con lo anterior, se observa que toda la literatura presentada se basa en estudios estáticos de las *Startups*, es decir, solo corresponden a una foto del momento actual de lo que está ocurriendo con este tipo de empresas, por lo tanto, existe una evidente carencia de estudios longitudinales que permitan conocer la evolución de éstas a través del tiempo, teniendo en cuenta que ese es un proceso natural de ellas y que, esa foto del momento, puede en pocos meses o años quedar obsoleta.

2.5. Posicionamiento de la investigación

De acuerdo a todo lo anteriormente expuesto en este capítulo, es relevante mencionar que, esta investigación pretende situarse en las brechas encontradas luego de la revisión literaria, las cuales están relacionadas con entender y comprender qué tipo de modelos de negocios basados en datos existen en Latinoamérica, centrando la mirada en las *Startups* de dicha región. Para realizar este estudio se utilizará como base para el análisis el Framework de (Hartmann et al., 2014) y los

patrones de innovación de (Parmar et al., 2014), el cual se adaptará a la realidad latinoamericana, incluyendo otros aspectos para profundizar más en la investigación. Como por ejemplo, se incluirá el modelo de ingresos *Freemium* y la distinción entre clientes y usuarios que la literatura revisada no ha abordado. Por lo tanto, el foco del estudio es la intersección de los tres tópicos desarrollados, lo cual se puede apreciar en el siguiente esquema:



Esquema N°2: Posicionamiento de la investigación

Fuente: Elaboración propia en base al posicionamiento de la investigación

En el siguiente capítulo, se detallarán los objetivos tanto generales como específicos de esta investigación, junto con la pregunta que guiará el trabajo.

3. Capítulo II: Propuesta de Investigación y Objetivos

3.1. Propuesta de Investigación y Objetivo General

El principal objetivo de esta investigación se relaciona con la motivación del equipo de trabajo por identificar y/o encontrar los principales modelos de negocios basados en datos, específicamente para las *Startups* creadas en Latinoamérica. Es importante mencionar que, esta investigación se enfoca en las *Startups*, puesto que, tal y como lo mencionan (Hartmann et al., 2014), este tipo de negocios, en primer lugar, tiene la capacidad de “comenzar de cero”, es decir, tienen la oportunidad de crear y llevar a cabo sus ideas sin limitaciones más allá de las económicas, no tienen barreras para la creatividad y, además tienen menor aversión al riesgo de fracasar. Adicionalmente, a lo largo de su desarrollo, las *Startups* deben enfrentarse continuamente a ambientes de gran incertidumbre, por lo cual, en muchas ocasiones, tendrán que cambiar sus modelos de negocios para sobrevivir, convirtiéndose en grandes focos de innovación y de desarrollo de nuevas ideas. Lo anterior está estrechamente relacionado con el *Big Data*, debido a que se ha convertido en una nueva industria y fuente de negocios innovadores, donde es pertinente preguntarse si las *Startups* están siendo protagonistas o cuál es el rol que están teniendo, dada la flexibilidad que por naturaleza poseen y, por lo tanto, la capacidad que tienen para enfrentar los desafíos de este nuevo panorama enfocado en la explosión exponencial de los datos y cómo están sobrellevando esto de tal manera de convertirlo en una ventaja competitiva que les permita crecer y adaptarse. Cabe mencionar que, de acuerdo a lo abarcado en la revisión literaria, actualmente, no existen trabajos que relacionen los tres ámbitos presentados (Ver Capítulo Revisión Literaria: Posicionamiento de la investigación) para otro contexto que no sea el de países desarrollados. Es por esto que, el equipo de trabajo ha planteado la siguiente pregunta de investigación:

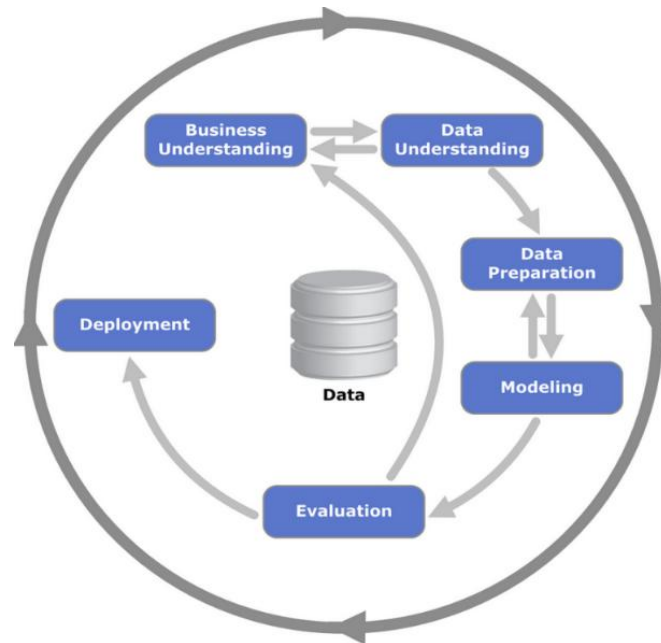
- 1. ¿Qué Modelos de Negocios basados en Datos están desarrollando las *Startups* en Latinoamérica?**
- 2. ¿Cuáles son las principales características de los Modelos de Negocios basados en Datos Latinoamericanos?**

3.2. Objetivos Específicos

1. Examinar y conocer la realidad de las *Startups* latinoamericanas dedicadas al tratamiento y gestión de datos.
2. Identificar qué patrones de innovación están desarrollando las *Startups* chilenas especializadas en la administración de datos.
3. Comparar e identificar factores comunes y diferentes de los modelos de negocios encontrados con los presentados por (Hartmann et al., 2014) en *Big Data* for Big Business?.
4. Identificar áreas o cuadrantes con potencial de explotación para futuros emprendimientos en Latinoamérica.
5. Identificar los principales factores o elementos de estudio que debiesen ser considerados para las futuras investigaciones relacionadas a este tema.

4. Capítulo III: Metodología

Para llevar a cabo este estudio, se utilizará la metodología CRISP-DM (*Cross-Industry Standard Process for Data Mining*), creado por tres empresas líderes de la industria del *Data Mining* en 1996: DaimierBenz, SPSS y NCR y descrito por (Pete et al., 2000). Este método ha sido diseñado en base a las actividades y experiencias que las personas realizan al momento de crear proyectos sobre minería de datos. En el esquema N°3 que se presenta a continuación, se aprecia el ciclo o secuencia de trabajo que plantea esta metodología, el cual está compuesto por seis etapas: (1) Entendimiento del negocio, (2) Entendimiento de los datos, (3) Preparación de los datos, (4) Modelamiento, (5) Evaluación e (6) Implementación.



Esquema N°3: Ciclo de Vida Metodología CRISP-DM

Fuente: CRISP-DM (Pete et al., 2000)

A continuación, se explicarán cada una de las etapas planteadas por la técnica CRIPS-DM, utilizada para comprender los modelos de negocios basados en datos de *Startups* latinoamericanas.

En primer lugar, la etapa de **Entendimiento del Negocio**, permite vislumbrar cuál es el objetivo de esta investigación -detallado en el capítulo anterior- el cual es identificar los modelos de negocios basados en datos que se están desarrollando actualmente por las *Startups*. Como base para

realizar este trabajo y como fuente principal de comparación con los resultados que se obtengan de éste, se utilizará el Framework (Hartmann et al., 2014), que plantea la existencia de al menos seis tipos de modelos de negocios basados en datos.

En segundo lugar, la etapa de **Entendimiento de los Datos**, está compuesta de dos líneas investigativas de forma paralela, que ayudarán al equipo de trabajo a interpretar de mejor forma la información que se está recabando y tener una mirada más amplia de los resultados y, de esta manera, lograr conclusiones más robustas. En relación a las etapas, la primera consta de una investigación exploratoria, a través de entrevistas presenciales semiestructuradas con el fin de identificar los actuales modelos de negocios basados en datos y los patrones de innovación en los cuales pueden identificarse las empresas. Para la otra fase, se realizó una investigación basada en el análisis de páginas web pertenecientes a los emprendimientos de la región. La lista de empresas que fueron objeto de estudio se obtuvieron de dos fuentes de información: Angellist y Start-Up Chile. En relación a Angellist, ésta es una plataforma web que agrupa perfiles de *Startups* describiendo su nombre, ubicación, mercado al cual pertenece, página web, entre otras (Angel List, 2015). Junto con lo anterior, la plataforma posee un gran número de registros de inversionistas, por lo tanto, sirve como conexión entre ellos y los emprendimientos listados. Pero esto no está enfocado sólo en personas que quieran financiar emprendimientos, sino que también la página posee un buscador que permite localizar *Startups* según diversos criterios, tales como ubicación, industria a la cual pertenece, mercados que atienden, entre otras. Por otro lado, Start-Up Chile es un programa de la CORFO (Corporación de Fomento de la Producción) enfocado en emprendedores en etapa temprana, tanto nacionales, como extranjeros otorgándoles apoyo financiero, llamado capital semilla, para que puedan abrirse al mundo (Start-Up Chile, 2015). A cambio, los emprendedores deben permanecer en el país durante los 6 meses que dura el programa y participar en actividades de impacto social. Cabe señalar que, este programa busca convertir a Chile en un polo de innovación y emprendimiento de América Latina. Es por lo anterior que, esta organización gubernamental posee una amplia base de datos con todos los emprendimientos que han participado de sus programas, la cual fue proporcionada como punto de partida para seleccionar empresas para la investigación. Esta base contiene en detalle el nombre, rubro, página web, procedencia y fundadores, de aquellas empresas aceptadas en el

programa Start-Up Chile entre los años 2010 y 2015. Las etiquetas utilizadas para filtrar las empresas *ad hoc* a la investigación, se presentan a continuación:

AngelList	Start-Up Chile
South America	BigData
Central America	BigData, Motion Picture, Video & Photography,
Mexico	BigData, Investment Banking,
Analytics	BigData, Crowd (Funding, Sourcing),
Business Intelligence	BigData, Gaming,
Big Data	BigData, Recruitment,
Cowdsourcing	BigData, Entertainment & Leisure, Publishing, Sports,
Data Mining	BigData, Energy,
Cloud Data Services	BigData, Hospitals & Medical Practice, Investment Banking, Legal,
Big Data Analytics	BigData, Health Care, Hospitals & Medical Practice,
Data	BigData, Publishing,
	BigData, Call Centers,
	BigData, Human Resources
	BigData, Computer, Investment Banking
	BigData, Entertainment & Leisure, Government, Motion Picture & Video, Travel
	BigData, Biotechnology, Health Care, Hospitals & Medical Practice
	BigData, Cargo & Shipping Services
	BigData, Publishing, Telecommunications
	BigData, Food & Beverage, Restaurants
	BigData, Construction, Computer, Engineering, Smart Cities, Transportation
	BigData, Computer, Design
	IT & Enterprise Software
	Latinoamérica

Tabla N°2: Etiquetas en listas de *Startups*

Fuente: Elaboración Propia en base a las bases de datos de AngelList y Start-Up Chile

A continuación se describirán en mayor detalle cada una de las líneas de investigación.

4.1. Entrevistas Semiestructuradas

Se realizaron 7 entrevistas presenciales semiestructuradas a *Startups* pertenecientes a Chile, específicamente a la Región Metropolitana, las cuales respondieron de manera exitosa al contactarlas vía email o a través de su página web (Ver anexo N°1). La entrevista consta de dos

partes, la primera enfocada en identificar los aspectos más importantes de los modelos de negocios en cuestión, basados en el Framework de (Hartmann et al., 2014), donde las preguntas están orientadas a explorar cómo los emprendimientos gestionan y administran los datos dentro de sus procesos de negocio y la importancia de éstos en cada una de las dimensiones del marco teórico. La segunda parte busca identificar si las *Startups* están desarrollando o utilizando algunos de los patrones de innovación planteados por (Parmar et al., 2014). Ambas aristas de la entrevista, tienen como fin encontrar las principales tendencias que estos emprendimientos están o desean implementar en sus negocios. Las empresas entrevistadas fueron: WholeMeaning, Noisegrasp, Semant.io, RutaPro, Mi Entorno, Desert Point y Agronometrics. Para más detalles acerca del cuestionario realizado y el material de apoyo, ir a anexo N°2 y N°3, respectivamente.

4.2. Análisis de Páginas Web

En el desarrollo de esta etapa, se observaron 164 páginas web de *Startups* latinoamericanas obtenidas de las listas previamente mencionadas, las cuales fueron examinadas mediante un análisis de palabras y conceptos claves que tienen directa relación con las dimensiones presentadas en el Framework; éstas son: estructura de costos, recursos clave, actividades clave, propuesta de valor, modelo de ingreso y segmentos de clientes. A continuación, se presentan las palabras y conceptos claves que se utilizaron para examinar las páginas web:

Dimensión	Palabras Claves
Estructura de costos	Costos, Márgenes, Rentabilidad, Egresos, Inversión
Recursos Clave	Compra de datos, información de clientes o socios, información de mercado, redes sociales, información desde distintas páginas web, información interna o propia, información co-creada, datos generados, seguimiento o monitoreo.
Actividades Clave	Consolidación de distintas fuentes, análisis descriptivo, análisis predictivo, optimización, seguimiento o monitoreo, procesamiento, informes o reportes, gráficas o <i>dashboard</i> , tablero de comando.
Propuesta de Valor	Venta de datos, información o conocimiento, dispositivos o sensores.
Modelo de Ingresos	Publicidad, venta de producto, comisión, arriendo, licencia, <i>freemium</i> , suscripción, uso, prueba gratuita, plan, tarifa, pago mensual, cobro.
Segmentos	Empresas, personas, usuarios, clientes, negocios.

Tabla N°3: Palabras claves para análisis de páginas web

Fuente: Elaboración propia en base a palabras claves para el análisis de páginas web

Para recopilar la información de cada una de las empresas, se crea una base de datos con cada una de las dimensiones y sus respectivas subdivisiones, donde cada celda toma valor 1 o 0 dependiendo si cumple o no con la característica, respectivamente. Con el objetivo de disminuir el sesgo en el análisis, cada uno de los investigadores (3) realizará de manera independiente la observación de cada página web, para luego contrastar y consolidar los resultados.

Luego, en las etapas de **Procesamiento de Datos y Modelamiento**, se realizaron diferentes análisis para cada una de las líneas de investigación. Con respecto a los resultados obtenidos de las entrevistas semiestructuradas, se desarrolla un análisis de contenido, basado en los autores Neuendorf (Neuendorf, 2002) y Krippendorff (Krippendorff, 2004). El método planteado por el primer autor, enumera una lista con nueve pasos que se deben implementar durante la investigación, los cuales se detallan a continuación:

1. Teoría y racionalidad. Se refiere a la pregunta de investigación e hipótesis del estudio.
2. Conceptualización. Definición de las variables
3. Operacionalización. Definición de la unidad de medida, para medir y validar.
4. Codificar. Examinar el contenido, a través de palabras claves.
5. Muestreo. Determinar la muestra.
6. Entrenamiento y Prueba Piloto. Revisar la metodología utilizada para codificar
7. Codificación de los Datos. Establecer la validez de la codificación
8. Medida Final de Validez. Medir la validez de cada variable
9. Tabulación y Reportes. Dar a conocer los resultados por variable y a través del tiempo

Por otro lado, (Krippendorff, 2004), plantea que los datos no son solo representaciones físicas de eventos, como en el ámbito financiero, sino que también pueden ser extraídos de textos, imágenes y expresiones que se generan para ser vistas, leídas, interpretadas o vividas, de acuerdo a sus sentidos, y deben ser analizadas desde todos estos puntos de vista, es por esto que, los textos deben ser estudiados en base al contexto en el que fueron creados. El análisis de contenido de esta investigación, se realizará según el último autor descrito, permitiendo inferir cuáles son los diferentes lineamientos y relaciones que pudiesen existir entre conceptos descriptores de los modelos de negocios de estas *Startups* chilenas vinculados a los modelos de negocios *Big Data* y a los patrones de innovación. Entonces, se pretende encontrar tendencias y elementos comunes

que describan lo que está ocurriendo en el ecosistema de las *Startups* entorno a los modelos de negocios *Big Data*.

A partir de la información recabada, a través del análisis web, se presentarán los principales resultados del estudio, mediante diferentes tipos de análisis cuantitativo. En primer lugar, se presentará una estadística descriptiva que busca proporcionar un análisis más específico y detallado de los datos. En segundo lugar, se realizará un análisis de *clustering* mediante dos procedimientos: *K-means* y *X-means* para agrupar a las *Startups* en base a sus características comunes. En relación al algoritmo *K-means*, éste clasifica a un conjunto de objetos en un determinado K de clústeres, definiendo a priori el K (García & Gómez, 2006), donde los grupos son representados por la media de sus centroides u observaciones representativas. A diferencia del anterior, *X-means* sugiere o detecta el número óptimo de grupos que es posible encontrar en base a los datos que se están analizando. En tercer lugar, se utilizarán árboles de decisión con el objetivo de simplificar el entendimiento de la composición de cada clúster, es decir, a través de este mecanismo se visualizará de forma más apropiada las variables o atributos que posee cada uno de los segmentos encontrados. Finalmente, se realizará un mapa de posicionamiento, para organizar los clústeres de los modelos de negocios basados en datos de Latinoamérica, de tal forma de comparar los resultados obtenidos con el Framework (Hartmann et al., 2014), y así identificar aspectos nuevos, comunes y potencialmente explotables.

Para la etapa de **Evaluación**, los diferentes algoritmos serán evaluados con respecto a sus medidas de desempeño. En el caso de los clústeres, ambos algoritmos serán evaluados en relación al tamaño de los grupos, al desempeño obtenido en la distancia promedio de los centroides y al indicador *Davies Bouldin*. Por otro lado, el desempeño del árbol de decisión, se evaluará en base a tres criterios: *Accuracy*, que refleja el porcentaje de casos reales que fueron clasificados correctamente por el modelo; *Class recall*, que representa el nivel de aciertos del modelo de cada una de las clases existentes en la base de datos y *Class precision*, que indica el porcentaje de las predicciones que fueron acertadas.

5. Capítulo IV: Resultados

En el siguiente capítulo se presentarán los resultados obtenidos a partir de la etapa exploratoria, así como también, en la etapa cuantitativa.

5.1. Etapa Exploratoria: entrevistas y análisis de contenido

Con el fin de tener una mirada de primera fuente sobre cómo se están configurando los modelos de negocios basados en datos y encontrar posibles tendencias que sean significativas para la investigación, como se detalló anteriormente, se llevaron a cabo 7 entrevistas presenciales semiestructuradas a las siguientes empresas: WholeMeaning, NoiseGrasp, Semant.io, RutaPro, Mi Entorno, Desert Point y Agronometrics; donde las personas entrevistadas eran parte de la gerencia de la empresa, entre ellas CEOs y/o co-fundadores, área de operaciones, área comercial, entre otras. Las entrevistas se realizaron en las dependencias de cada una de las empresas en un horario convenido previamente. El equipo investigativo poseía un cuestionario realizado con anterioridad, así como también, material de apoyo visual para los entrevistados.

Entrando de lleno en los aspectos relevantes encontrados, en términos generales es posible notar que, la mayoría de las empresas tenía entre 1 año y medio y 2 años de funcionamiento, donde el número de integrantes que participan en ellas, fluctúa entre las 2 y 5 personas. Además, al mercado al cual apuntan estas empresas es de tipo industrial, es decir, está dirigido al segmento de clientes B2B. Junto con lo anterior, todas coincidieron que el negocio que están desarrollando actualmente no es totalmente igual a la idea inicial, pero según lo mencionado por Nicolás Bravo, socio y Gerente de Ventas en Chile de la empresa Semat.io -dedicada a realizar seguimiento en tiempo real de las expresiones y menciones en redes sociales asociadas a una marca en particular- *“El espíritu central se respeta, pero han habido muchos cambios”*, dejando en evidencia que, si bien ocurren cambios con el transcurso del tiempo, el foco siempre es la idea original, lo cual coincide con lo planteado por los autores (Maurya, 2012) y (Ries, 2012).

Otro punto importante a destacar, relacionado a los costos, es que una parte importante de ellos (cerca del 90%) son atribuibles a RRHH y, que los costos correspondientes a las actividades con los datos son prácticamente nulos; en palabras de Andrés Groisman -CEO y co-fundador de la empresa NoiseGrasp, quienes se definen como una plataforma que optimiza el mix de medios de

una empresa, recomendándoles en qué medios les conviene invertir- quien señala lo siguiente acerca de la estructura de costos “...mi estructura de costos está distribuida prácticamente en un 90% en RRHH... Los gastos relacionados a los datos no son tan relevantes”. Por otro lado, con respecto a los equipos que componen estas empresas, se evidenció que los profesionales son diversos y de múltiples disciplinas, más que sólo aquellas enfocadas en la gestión o el manejo de datos, ya que se encontraron lingüistas, ingenieros informáticos, industriales, agrónomos, entre otros.

Con respecto a las actividades, todas las empresas coincidieron que el foco del negocio está en el análisis, principalmente el predictivo. Pero, uno de los reiterados elementos que surgió, fue que el servicio de gestionar los datos va mucho más allá de eso, porque se debe acompañar al cliente durante el uso de las plataformas, generando relaciones más cercanas con él, lo cual se refuerza con lo expresado por José Rafael Campino CEO de Desert Point -empresa que, a través de modelos predictivos de accidentes, evalúan riesgos operacionales y proponen mecanismos de control a través de indicadores en tiempo real- quien definió su propuesta de valor en torno a la gestión de los datos y al acompañamiento del cliente a través de la consultoría; en sus propias palabras indica que “Los clientes no saben qué hacer con los datos hay que darles tareas, recomendarles cosas que tienen que hacer”. Por lo tanto, se aprecia que parte vital del servicio está centrado en el cliente y en facilitarle la manera en que se relacionan con los datos, entendiendo que no son expertos en el tema y que el acompañamiento es clave.

Por otro lado, una parte importante de la entrevista fue ahondar en los patrones de innovación planteados por (Parmar et al., 2014) y determinar si las empresas se identificaban con alguno de ellos o si eventualmente podrían realizarlo. En relación al primer Patrón, “Mejorar productos para generar datos”, en términos generales las empresas señalaron que, todos los datos que recaban son importantes para ellos y que no existen datos que no se estén utilizando o que no tengan valor, por lo tanto, no se identificaron con este concepto.

Con respecto al Patrón 2: “Digitalización de los activos físicos”, 5 de las 7 Startups señalaron que no aplicaba en su organización, destacando la frase “La información nace digital”, expresada por Juan Besa, Co-Fundador y CTO de Wholemeaning, empresa orientada en el entendimiento de los mensajes de clientes con dos principales focos: analítica básica (call center, reclamos), entregando

un servicio permeado a la organización y el “ruteo”, en donde se optimizan los procesos de interacción con terceros.

En relación al Patrón 3 corresponde a “Combinación de datos dentro y a través de las industrias”, hubo un amplio consenso entre todas las empresas sobre la importancia de compartir los datos a través de los mercados, ya que se genera una relación de ganar-ganar para ambas partes. En este sentido, Cristián Araneda -Gerente de Estudios y Cartografía de Mi Entorno, empresa que actúa como un repositorio de datos públicos, privados y socios basados en geografía- dice que *“las alianzas son clave”*, así como también, Juan Besa, indica que *“es valioso compartir datos”*.

En el penúltimo Patrón “Comercio de datos”, no existe un acuerdo general entre las empresas entrevistadas, ya que, por un lado, la mayoría manifiesta que sí tiene datos que podrían interesarle a otras empresas pero que no es su foco actual explotarlos. En cambio, hay otras empresas que no tienen claridad si poseen o no información valiosa para otros negocios.

Finalmente, referente al Patrón 5 “Codificar una capacidad de servicio distintivo” las empresas indicaron que existe la posibilidad de convertir en un nuevo producto algún servicio o software que tienen, pero no han decidido si hacerlo o no, es decir, ha sido tema de análisis pero no se ha tomado una decisión concreta respecto a qué camino continuar.

De acuerdo a lo expuesto anteriormente, es posible mencionar que, estas *Startups* chilenas se encuentran en pleno proceso de crecimiento, teniendo que ir perfeccionando sus métodos, adaptándose a los nuevos requerimientos de los clientes y mejorando sus propuestas valor. En este sentido, se aprecia que, la actividad que se debe explotar o en donde se debe poner el foco, es en la gestión de los datos y la asesoría continua con los clientes, velando por un proceso de ayuda permanente en la toma de decisiones y en cómo manejar los datos que poseen y sacarles el mayor provecho posible. Esta situación implica que, aquellos procedimientos que hoy les permiten posicionarse como *Startups* emergentes en la industria, deberán fortalecerse para lograr convertirse en empresas sostenibles en el largo plazo.

5.2. Etapa Cuantitativa

En primer lugar, cabe señalar que, de las 164 empresas analizadas, sólo un porcentaje de ellas se encontraban disponibles para incluirlas dentro de la investigación, principalmente, debido a que, algunas empresas tenían sus páginas web en funcionamiento y contaban con la suficiente información -ya sea de sus mismas páginas o de otras plataformas en la web- que permitiera identificar sus características y, por ende, su capacidad para entrar en el estudio. Por lo tanto, la base de datos utilizada para la estadística descriptiva y para los análisis siguientes, está compuesta de 82 empresas, quienes cumplían con todos los requerimientos necesarios para desarrollar la investigación (para más detalle de las empresas ver anexo N°4). Es relevante mencionar que, durante la consolidación de la base final, el equipo de trabajo obtuvo un 87% de acuerdo respecto de análisis de las empresas que llevaron a cabo de forma individual, este porcentaje, demuestra un alto nivel de coincidencia del equipo investigativo al momento de definir las principales características de cada empresa, lo cual es un input fundamental para desarrollar el análisis y entregar resultados con mayor nivel de confiabilidad. Finalmente, en una segunda ronda, los investigadores dirimieron los casos en que había diferencia en la clasificación, hasta lograr un consenso el cual fue validado por el profesor guía de la Tesis. Dichos resultados se presentan a continuación:

5.2.1. Estadística Descriptiva

En primer lugar, en términos de cantidad de empresas analizadas, se aprecia un amplio liderazgo de aquellas pertenecientes a Chile, las cuales representan el 47,6% de la muestra. Luego, en segundo lugar destaca Brasil con 15 empresas, lo cual representa un 18,3% del total; más atrás se queda Argentina con un 12,2% y México con un 11%. Los países que poseen empresas con modelos de negocios basados en datos y que, representan menos del 10% de las observaciones son, en orden decreciente, Colombia, Perú y Uruguay. Esta información se resume en la siguiente tabla:

País	Cantidad	%
Argentina	10	12,20%
Brasil	15	18,29%
Chile	39	47,56%
Colombia	6	7,32%
México	9	10,98%
Perú	2	2,44%
Uruguay	1	1,22%
Total general	82	100%

Tabla N°4: Distribución de empresas según País

Fuente: Elaboración propia en base a la estadística descriptiva

La sobre representación de Chile en la muestra se explica en gran manera por el éxito y antigüedad del programa de apoyo al emprendimiento Start-Up Chile, el cual ha otorgado capital semilla a más de 1000 emprendimientos desde su lanzamiento en el año 2010, especialmente considerando que otros países como Colombia y Perú están recién comenzando a generar iniciativas similares (ver Cuadro N°1: Apoyo directo a las *Startups* en América Latina: Comparación entre países, 2012)

Adicionalmente, respecto a la dimensión recursos clave, de los 82 modelos de negocios analizados, aproximadamente un 90% utiliza como fuente de información datos externos, lo que corresponde a 74 *Startup*. Estas empresas obtienen los datos, principalmente, de sus clientes o socios y de fuentes de datos libres, tales como datos públicos. Asimismo, no se observa una tendencia clara en relación a la utilización de datos pertenecientes a redes sociales, ya que de aquellas empresas que utilizan datos libres, un 49% si utiliza datos de medios sociales, mientras que 51% no lo hace. Por otro lado, en relación a los datos internos, sólo el 40% de las *Startups* los utiliza, los cuales se obtienen, primordialmente, a través de procesos de monitoreo. El detalle de la distribución y porcentaje de cada uno de los aspectos relacionados a la dimensión de recursos clave, se presentan en las gráficas que se muestran a continuación:

Recursos Clave		
Tipo de Fuente	Si Utiliza	% (del total de empresas)
Interna	33	40,24%
Externa	74	90,24%
Externa e Interna	25	30,49%

Tabla N°5: Distribución de empresas según Fuente de Datos que utilizan

Fuente: Elaboración propia en base a la estadística descriptiva

En relación a los datos externos:

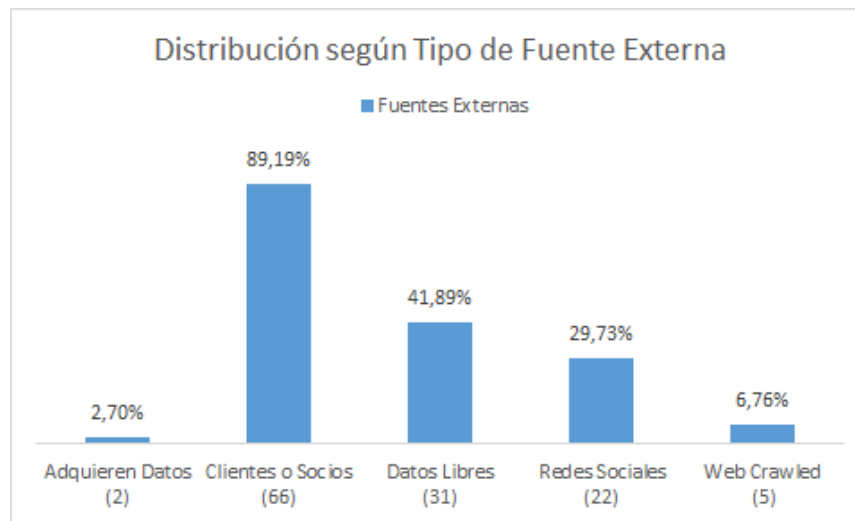


Gráfico N°1: Distribución de empresas según tipo de Fuente Externa

Fuente: Elaboración propia en base a la estadística descriptiva

Y a los datos internos:

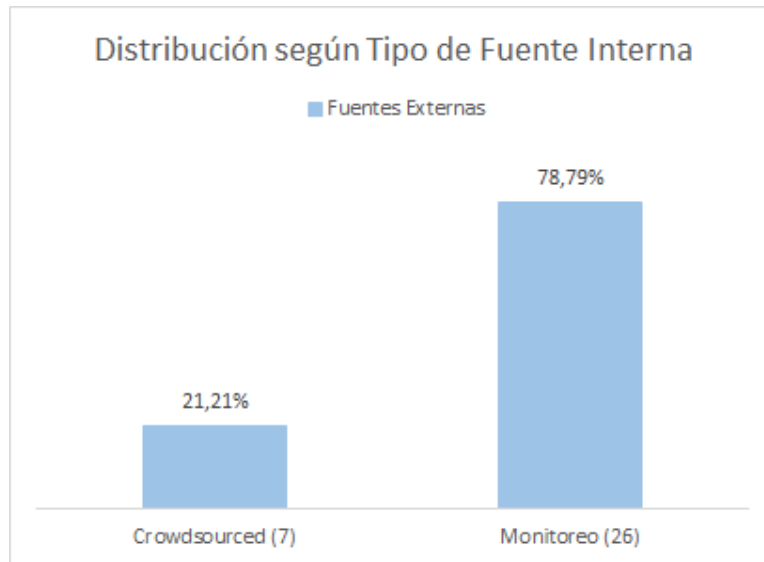


Gráfico N°2: Distribución de empresas según tipo de Fuente Interna

Fuente: Elaboración propia en base a la estadística descriptiva

Con respecto a la dimensión actividades clave, el 96,34% de las empresas realiza la agregación de datos desde diferentes fuentes y, en relación al análisis de datos, el 64% de las *Startups* realiza alguno de ellos, de las cuales el 97% desarrolla análisis del tipo descriptivo, el 45% del tipo predictivo y un 23% desarrolla un análisis prescriptivo u optimización. La adquisición de datos no es una actividad clave para ninguna de las *Startups* en estudio. Para el caso de la generación de datos, un 63% de las empresas realiza algún tipo de recolección o monitoreo de ellos. Sólo un 8% de aquellas empresas tiene como actividad clave el generar datos mediante *Crawling*, mientras que un 96% lo hace a través de Monitoreo y *Crowdsourcing*. Por otro lado, para un 82% de las empresas realiza procesamiento de datos y, un 92% desarrolla distribución de información o datos a través de informes. Finalmente, la visualización es clave para el 99% de las organizaciones estudiadas. Las siguientes gráficas muestran lo planteado en el párrafo anterior:

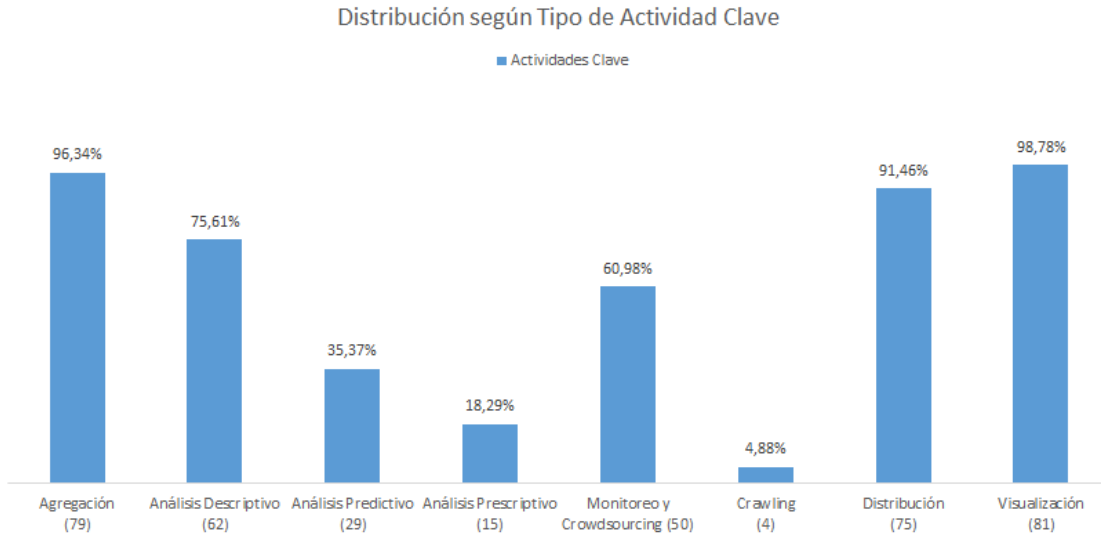


Gráfico N°3: Distribución de empresas según tipo de Actividad Clave

Fuente: Elaboración propia en base a la estadística descriptiva

En relación a la propuesta de valor, destaca que el 99% de las empresas está enfocado en entregar información o conocimiento, mientras que el 11% vende algún producto no digital (también denominado *Non-Data*) y sólo el 5% de ellas comercializa datos. La información anterior se detalla en el siguiente gráfico:

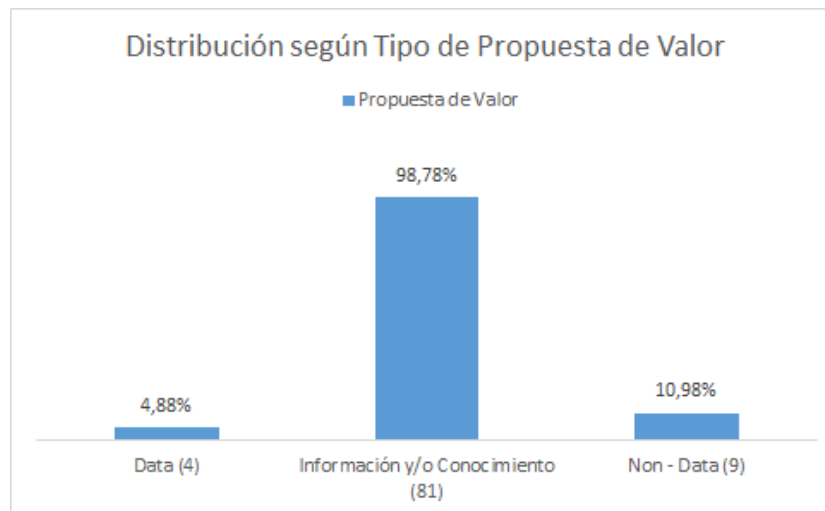


Gráfico N°4: Distribución de empresas según tipo de Propuesta de Valor

Fuente: Elaboración propia en base a la estadística descriptiva

Si nos referimos al segmento de clientes objetivo, el 93% de las *Startups* se dirige a empresas - segmento B2B- el 22% a clientes personas -segmento B2C- y el 26% además posee usuarios de sus productos/servicios.

Finalmente, entre los modelos de ingresos encontrados, en la mayoría no se encontraba disponible esta información (38%). Asimismo, el 26% emplea el modelo por Suscripción, el 9% el modelo *Freemium*, otro 9% el modelo de Licencia. Un menor porcentaje (2%) recibe ingresos por la Venta de Activos, el 8% por Cobro por Uso y el 8% final por Comisión. Lo anterior se visualiza en los siguientes gráficos:

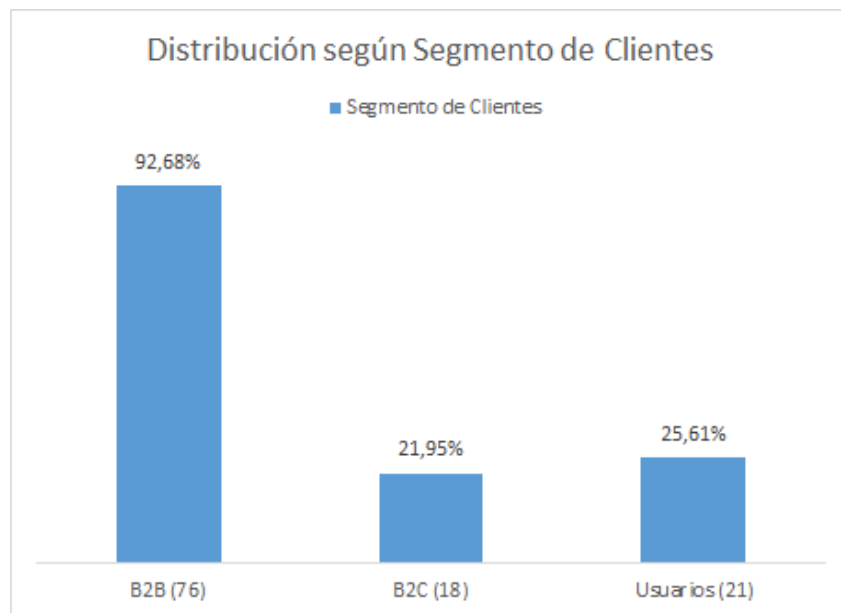


Gráfico N°5: Distribución de empresas según Segmento de Clientes

Fuente: Elaboración propia en base a la estadística descriptiva

Y en relación al modelo de ingresos:

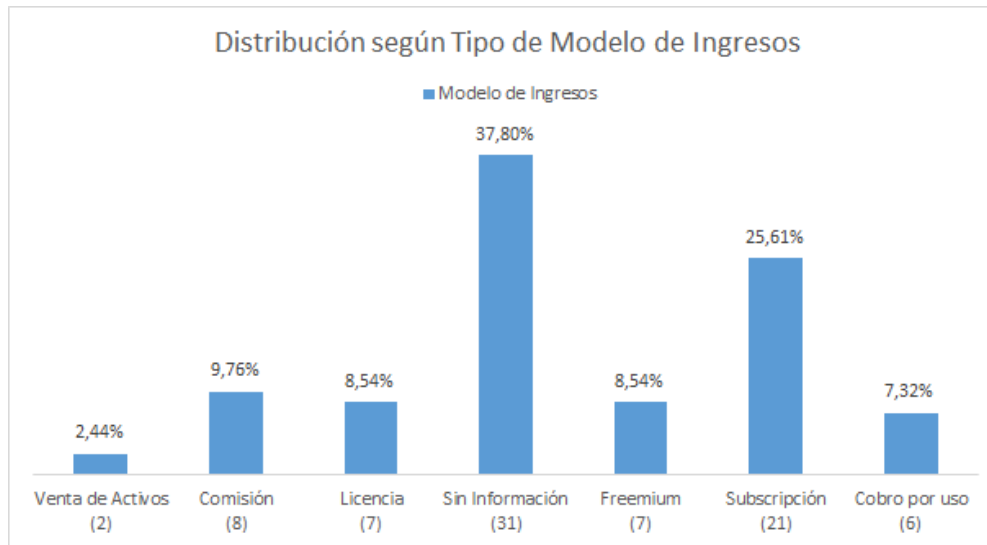


Gráfico N°6: Distribución de empresas según tipo de Modelo de Ingresos

Fuente: Elaboración propia en base a la estadística descriptiva

5.2.2. Análisis *Clustering*

En primer lugar, se realizó una selección de los atributos que serían relevantes para realizar la segmentación. Luego, de forma simultánea se llevaron a cabo los procesos de *clustering* por medio de los métodos de *X-means* y *K-means* -indicando a este último como mínimo 6 clústeres- lo cual dio como resultado 3 y 6 grupos, respectivamente. El detalle de la cantidad de empresas por cada clúster se detalla en las siguientes tablas:

K-Means	
Clúster	Nº Empresas
0	20
1	21
2	6
3	10
4	6
5	19
Total	82

Tabla N°6: Clústeres *K-Means*

Elaboración propia en base a resultados del Análisis *Clustering*

X-Means	
Clúster	N° Empresas
0	34
1	27
2	21
Total	82

Tabla N°7: Clústeres *X-Means*

Elaboración propia en base a resultados del Análisis *Clustering*

Luego, con el fin de evaluar el desempeño de los algoritmos de agrupación en clústeres en ambos métodos y determinar si eran lo suficientemente efectivos en la segmentación realizada, se llevó a cabo el procedimiento para obtener el *Cluster Distance Performance*, cuyos resultados se muestran a continuación:

K-Means	
Avg. within centroid distance	3,264
Davies Bouldin	2,051

Tabla N°8: Resultados Cluster Distance Performance

Fuente: Elaboración propia en base a resultados del Análisis *Clustering*

X-Means	
Avg. within centroid distance	4,439
Davies Bouldin	2,207

Tabla N°9: Resultados Cluster Distance Performance

Fuente: Elaboración propia en base a resultados del Análisis *Clustering*

De las tablas anteriores, se puede apreciar que, el modelo *K-means* presenta, en promedio, mayor homogeneidad entre las empresas clasificadas dentro de cada clúster con un promedio de 3,264, mientras que, para el caso del modelo *X-means*, su distancia promedio a cada centroide es de un valor igual a 4,439, presentando mayores niveles de heterogeneidad dentro de los clústeres que ha definido. En relación al indicador *Davies Bouldin*, nuevamente el modelo *K-means* presenta un mejor resultado, ya que presenta un indicador más cercano a cero, lo cual implica un mejor resultado en términos de homogeneidad intra-clúster y una mayor heterogeneidad inter-clúster.

Finalmente, se puede concluir que en términos de desempeño, el método que destaca por clasificar de manera más precisa las empresas es el procedimiento *K-means*.

5.2.3. Árboles de Decisión

Luego de tener una visión sobre la cantidad de clústeres que se pudieron encontrar en base al análisis de los datos, es preciso determinar qué tipo de empresas pertenecen a cada clúster, con el fin de caracterizar y determinar el tamaño de cada uno de los grupos. Para llevar a cabo lo anterior, se realizó el procedimiento de Árboles de Decisión, basados en los modelo de *clustering K-means* y *X-means*. Los resultados obtenidos se muestran a continuación:

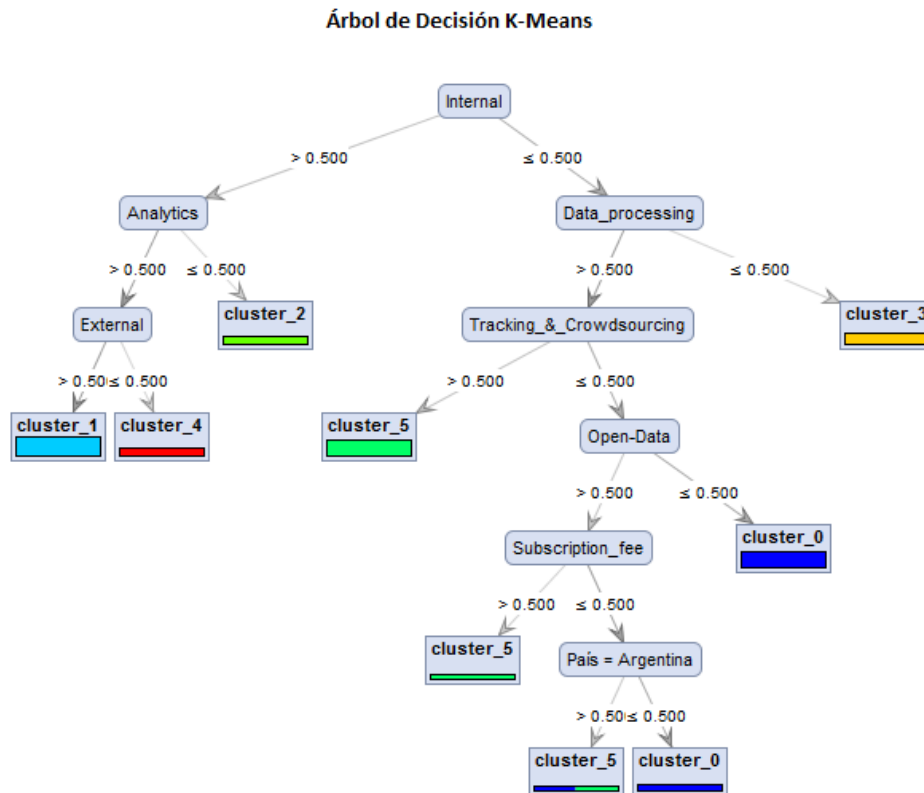


Imagen N°5: Árbol de Decisión Modelo de *Clustering K-means*
Elaboración propia en base a resultados del Análisis *Clustering*

El árbol del modelo *K-means*, para determinar los clústeres utiliza como características principales aspectos como: sí existe utilización o no de datos internos como recurso clave y si realizan o no análisis o procesamiento de datos. Como variables para continuar con la clasificación de las empresas, considera si utilizan o no datos externos y de qué tipo, y si realizan o no monitoreo de los datos. Una última variable para distinguir entre dos clústeres, es si poseen o no un modelo de ingresos del tipo suscripción.

Por otro lado, el árbol del modelo *X-means*, presenta como primera característica diferenciadora, si las empresas generan o no datos a partir de un proceso de monitoreo. La siguiente variable que utiliza, está relacionada a si las *Startups* realizan o no análisis del tipo descriptivo. Adicionalmente, para definir al clúster 0 ocupa variables como el país de origen de la empresa y si recolectan datos externos a través de un mecanismo del tipo *Web Crawled*. A continuación, se presenta el árbol de decisión del modelo *X-means*:

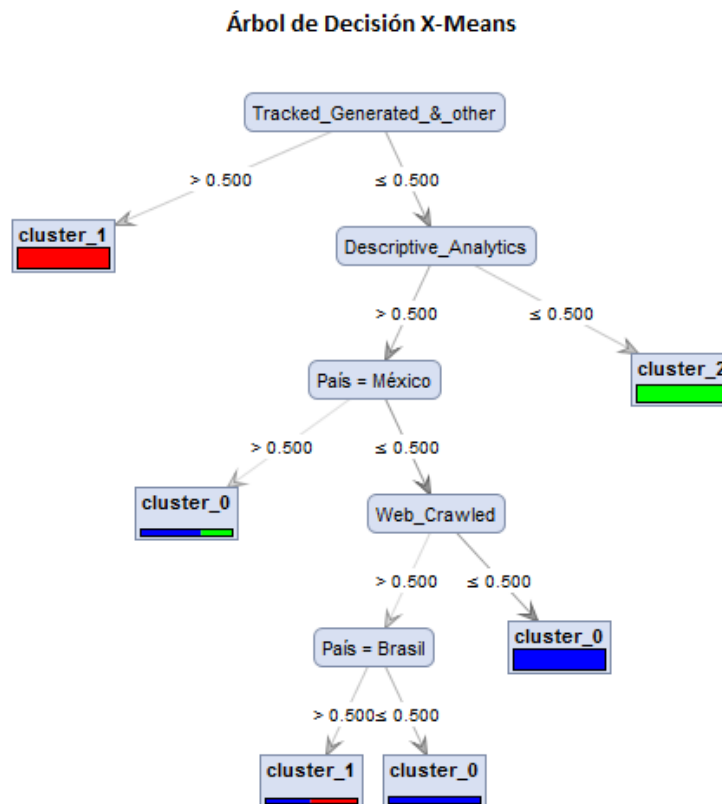


Imagen N°6: Árbol de Decisión Modelo de *Clustering X-means*
Elaboración propia en base a resultados del Análisis *Clustering*

Al igual que en la etapa anterior, es fundamental medir el desempeño de esta técnica, con el fin de evaluar qué nivel de acierto hay en los árboles encontrados. Los resultados del *Performance* tanto del árbol de *K-means*, como del árbol a partir de *X-means* se presentan a continuación:

Performance X-Means				
Accuracy: 97,56%				
	true cluster_0	true cluster_2	true cluster_1	class precision
pred. cluster_0	33	1	0	97.06%
pred. cluster_2	0	20	0	100.00%
pred. cluster_1	1	0	27	96.43%
class recall	97.06%	95.24%	100.00%	

Tabla N°10: Resultados Performance Árbol de decisión *X-Means*

Fuente: Elaboración propia en base a resultados del Análisis *Clustering*

Performance K-Means							
Accuracy: 98,78%							
	true cluster_0	true cluster_1	true cluster_5	true cluster_2	true cluster_3	true cluster_4	class precision
pred. cluster_0	19	0	0	0	0	0	100.00%
pred. cluster_1	0	21	0	0	0	0	100.00%
pred. cluster_5	1	0	19	0	0	0	95.00%
pred. cluster_2	0	0	0	6	0	0	100.00%
pred. cluster_3	0	0	0	0	10	0	100.00%
pred. cluster_4	0	0	0	0	0	6	100.00%
class recall	95.00%	100.00%	100.00%	100.00%	100.00%	100.00%	

Tabla N°11: Resultados Performance Árbol de decisión K-Means

Elaboración propia en base a resultados del Análisis *Clustering*

Con respecto a los resultados del análisis del *Performance* presentados en las tablas anteriores, se puede concluir que, ambos modelos presentan niveles similares de eficiencia y confianza, que se traducen en resultados equivalentes de *Accuracy*, *Class recall* y *Class precision*. Esta situación

demuestra que ambos modelos son capaces predecir casi de forma perfecta cómo están clasificadas las empresas y sus características representativas.

De acuerdo a lo expuesto anteriormente, el equipo de trabajo ha decidido continuar el estudio de las *Startups* a partir del resultado entregado por el modelo *K-means*, principalmente, por sus niveles de eficiencia, homogeneidad dentro de cada clúster y heterogeneidad que ha encontrado entre ellos. A continuación, se presenta una tabla resumen de la media y desviación estándar por clúster, en base a las variables relevantes arrojadas por el árbol de decisión:

	Clúster 0		Clúster 1		Clúster 2		Clúster 3		Clúster 4		Clúster 5	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
País = Argentina	0,100	0,308	0,048	0,218	0,000	0,000	0,200	0,422	0,000	0,000	0,263	0,452
External	1,000	0,000	1,000	0,000	0,667	0,516	1,000	0,000	0,000	0,000	1,000	0,000
Open-Data	0,250	0,444	0,619	0,498	0,333	0,516	0,100	0,316	0,000	0,000	0,579	0,507
Internal	0,000	0,000	1,000	0,000	1,000	0,000	0,000	0,000	1,000	0,000	0,000	0,000
Analytics	0,950	0,224	1,000	0,000	0,000	0,000	0,000	0,000	1,000	0,000	0,947	0,229
Tracking_ & Crowdsourcing	0,000	0,000	1,000	0,000	1,000	0,000	0,200	0,422	1,000	0,000	0,842	0,375
Data_processing	1,000	0,000	1,000	0,000	0,167	0,408	0,000	0,000	1,000	0,000	1,000	0,000
Subscription_fee	0,100	0,308	0,333	0,483	0,167	0,408	0,200	0,422	0,000	0,000	0,474	0,513
Tamaño (n)	20		21		6		10		6		19	

Tabla N°12: Tabla de media y desviación estándar por clúster
Elaboración propia en base a resultados del Análisis *Clustering*

De la tabla anterior, es posible notar que, los clústeres más homogéneos son el número 3 y el número 4, ya que evidencian desviación cercana a cero y medias cercanas a 1 o 0; lo anterior predice que la mayoría de las empresas que componen ese grupo poseen con certeza o no, las características presentes en la tabla. Entre los grupos más heterogéneos, se encuentran el clúster 2 y el clúster 5, dado que presentan medias entre 0,4 y 0,6, dejando en evidencia que el atributo no presenta una fuerte tendencia entre las empresas que los componen.

Para complementar los resultados encontrados, se utilizará esta información junto con la base de datos, cuestión de caracterizar de forma más específica los principales rasgos de cada clúster y cuáles son los modelos de negocios de las *Startups* que los componen. En base a sus

características más relevantes cada clúster ha recibido una etiqueta o nombre descriptivo. A continuación, se detallan los seis clústeres encontrados y sus etiquetas:

- **Clúster 0: Recolector de datos externos para hacer analítica avanzada**

Este clúster está compuesto por 20 empresas, las cuales provienen principalmente de Chile (35%), Brasil (20%) y México (15%). Además, la totalidad de estas *Startups* trabajan con datos proporcionados por fuentes externas, particularmente, por sus clientes o socios (85%) y datos disponibles libremente en la web (55%), específicamente, en las redes sociales. Otros tipos de fuentes externas son menos utilizadas, tales como datos públicos, ya que sólo el 25% de ellas lo utiliza, así como también, sólo el 10% extrae datos a través de *Web Crawled*. Cabe señalar que, ninguna de las empresas utiliza datos internos. Ahora bien, respecto de las actividades clave que desarrollan para crear valor, destaca la agregación de datos, el procesamiento de ellos, específicamente, a través del análisis descriptivo (85%), predictivo (45%) y prescriptivo (35%) (optimización), la entrega de informes (90%) y la visualización de los resultados (90%). Respecto de la propuesta de valor, entregan información y conocimiento pero se desconoce su modelo de ingresos. Finalmente, como segmento de clientes, estas empresas se dirigen a otros negocios, es decir, un modelo B2B.

Entre las empresas representativas de este clúster, se encuentran NoiseGrap y QualCanal. La primera es una *Startup* chilena que, a través de técnicas matemáticas avanzadas, modela cómo los consumidores reaccionan a diferentes actividades de marketing con el fin de optimizar el mix de medios de una compañía. Cabe destacar la alianza estratégica que esta compañía posee con el Centro de Modelamiento Matemático de la Universidad de Chile, quienes aportan con los algoritmos necesarios para realizar los procesos de optimización. Por otro lado, cabe señalar que, el modelo de ingresos que utiliza esta empresa es del tipo Licencia y se dirige al segmento B2B. Por otro lado, QualCanal es una empresa brasileña dedicada a monitorear el impacto de los programas de televisión en las redes sociales, específicamente Twitter, organizando en tiempo real los tópicos más hablados sobre una marca en particular. Esta *Startup* se dirige al segmento B2B y también al segmento usuarios.

- **Clúster 1: Recolector multi-fuente de datos y oferente de un servicio completo**

El segundo grupo, es levemente más grande que el anterior y está compuesto por 21 empresas, principalmente originarias de Chile (43%) y Brasil (29%). Entre sus características más importantes destaca que, son empresas que utilizan como fuente de datos aquellas provenientes de plataformas externas, específicamente de sus clientes o socios (95%) y de fuentes disponibles libre y gratuitamente en internet (62%). Adicionalmente, el 100% de ellas genera datos de fuentes internas a través de procesos de monitoreo (95%) y de *Crowdsourcing* (5%). Respecto de las actividades clave que llevan a cabo, es posible mencionar que, todas realizan agregación de datos, así como también análisis descriptivo y predictivo. Además de lo anterior, otra actividad clave también llevada a cabo por el 100% de las empresas que componen este grupo, es la generación de datos, la cual se realiza a través del seguimiento a las actividades que sus clientes realizan. Estos datos se traducen en información, la cual es entregada en reportes y a través de interfaces gráficas preparadas especialmente para los usuarios, es decir, efectúan tanto la distribución como la visualización de los datos, respectivamente. En relación al modelo de ingresos, no se aprecia una clara tendencia a poseer algún tipo de modelo en particular, ya que algunas de estas empresas cobran por venta de activos (5%), cobro por comisión (5%), *Freemium* (10%), Licencia (14%), Suscripción (33%) y al 29% de ellas no fue posible determinar su forma de generar ingresos. Finalmente, el segmento de negocio al cual apuntan corresponde al B2B.

Ejemplos de este tipo de empresas son Wivo, que analiza los puntos de venta a través de la integración de información de diversas fuentes de datos, dándole sentido en tiempo real, y además, ofrece visualización de los datos de la tienda desde cualquier lugar. Todo lo anterior, permite que las firmas entiendan el comportamiento de los clientes y descubran oportunidades de mejora para aumentar su rentabilidad. Otra compañía distintiva en este grupo es Bovcontrol, la cual pertenece al rubro agrícola; su principal actividad está centrada en recolectar información sobre animales a través de dispositivos de identificación, la cual es presentada en un tablero de comando o *dashboard* que permite tener un inventario de los animales, realizar control sanitario y nutricional, entre otros. El servicio funciona por medio de una aplicación móvil que permite a los clientes ver en cualquier momento y lugar los datos. El tipo de modelo de ingresos es *Freemium* y se dirige tanto al segmento B2B, como al B2C.

- **Clúster 2: Monitoreo como un servicio**

Este clúster agrupa sólo a 6 empresas, las cuales se distribuyen en partes iguales en los países de Colombia, México y Chile. Se aprecia que, se caracterizan por recabar datos, tanto de fuentes externas (67%) -mediante datos proporcionados por los clientes o socios- como interna (100%), a través de un proceso de *Crowdsourcing*. En cuanto a las actividades que identifican a este grupo, estas firmas realizan agregación de datos (67%), así como también seguimiento (100%). Otra actividad que realizan la totalidad de estas *Startups* es el proceso de visualización y, como propuesta de valor, entregan conocimiento e información. Como modelo de ingresos destaca el modelo de cobro por comisión en el 50% de ellas y, respecto del segmento de clientes, estas firmas distinguen 3 grupos, las empresas -B2B- los compradores individuales -B2C- y los usuarios, quienes utilizan sus plataformas.

Respecto de las empresas representativas de este grupo destacan MiCarga y Rutanet. La primera es una empresa colombiana perteneciente al rubro del transporte, que a través de una amplia base de datos de vehículos con capacidad ociosa, ayudan a las empresas a optimizar sus operaciones logísticas a través de un software especializado. El modelo de ingresos utilizado por esta *Startup* es cobro por comisión. La segunda, es una empresa originaria de México que, a partir de la información que publican sus usuarios genera un directorio de transportistas, donde se encuentran las personas que buscan transportar carga, así como también, personas con transporte disponible. La principal diferencia de ambas, es que la segunda sólo actúa como una plataforma para que coincidan las personas que desean transportar algo con las que tiene disponible transporte, en cambio MiCarga ofrece un servicio de optimización de las operaciones logísticas.

- **Clúster 3: Agregador de datos externos**

El siguiente clúster está compuesto por 10 empresas, provenientes de Chile (50%), Brasil (20%) y Argentina (20%). Para recolectar datos, utilizan sólo fuentes externas, principalmente, de los clientes o socios con quienes trabajan (90%) y, en menor medida, datos disponibles libremente (30%). En relación a las actividades que desarrollan, estas compañías realizan agregación (100%), distribución de datos a través de informes (90%) y visualización de los datos (100%). Como

propuesta de valor, entregan información y conocimiento. Los modelos de ingresos identificados para estas empresas son, en su mayoría, el modelo *Freemium* (30%), de suscripción (20%) y por comisión (20%). Finalmente, el segmento de clientes a quienes se dirigen son tanto B2B como B2C.

Algunos ejemplos de empresas que integran este grupo son Kurv y Packdocs. Por un lado, Kurv es una empresa que pertenece al rubro alimenticio y permite tener toda la información de la empresa integrada en una sola plataforma, de manera instantánea. El modelo de ingresos que posee esta empresa es del tipo *Freemium*. Con respecto a Packdocs, cabe destacar que, es una plataforma que permite gestionar la información dentro de grupos de trabajo. La plataforma permite invitar a amigos para que se unan a ella y, de esta manera, gestionar y compartir documentos. Las personas pueden participar en una red pública o crear su propia red particular. De acuerdo a la descripción de ambas empresas, es posible apreciar que las *Startups* de este clúster están enfocadas principalmente en agregar datos, pero no entregan ningún tipo de análisis a partir de ellos.

- **Clúster 4: Generador de datos internos y analítica básica**

Las 6 compañías que componen este clúster provienen principalmente de Chile (63%), y, en menor medida de Brasil (17%), las cuales obtienen sus datos de fuentes internas mediante un mecanismo de monitoreo. Cabe señalar que, ninguna de ellas recurre a fuentes de datos externas como input. Respecto de las actividades clave que llevan a cabo, se caracterizan porque el 100% de ellas lleva a cabo procesos de agregación, análisis descriptivo y procesamiento de datos, específicamente a través de un seguimiento de la información recolectada. Adicional a esto, entregan informes y visualizan los resultados obtenidos. Como propuesta de valor ofrecen información, conocimiento y productos tangibles, tales como sensores. No entregan información de su modelo de ingresos y se dirigen a empresas, es decir, su segmento de cliente es el B2B.

Una de las empresas representativas de este grupo es Cuboverde la cual monitorea de manera continua y en tiempo real diversas variables de la empresa tales como: consumos energéticos, procesos industriales, emisión de contaminantes, entre otros. Esta actividad la realiza por medio de sensores de bajo costo y permite que las empresas sean más conscientes del impacto que generan sus acciones sobre el medio ambiente, para poder minimizarlo, así como también, reducir

sus costos. Junto con la anterior, cabe destacar a Eye3, que al igual que la anterior, también tiene foco en el monitoreo y se dedica a entregar soluciones tecnológicas e industriales con énfasis en la detección, monitoreo y control por proximidad o cercanía para mejorar la seguridad y hacer más eficientes los procesos industriales.

- **Clúster 5: Recolector de datos externos, monitoreo y análisis**

El último clúster encontrado está compuesto por 19 firmas, provenientes de Chile (58%) y Argentina (26%). Estas compañías se caracterizan por recabar datos de fuentes externas (100%), principalmente, de clientes o socios y de información disponible libremente en internet, tal como aquellos datos originados en redes sociales. Como actividades clave, todas las empresas de este grupo realizan agregación y procesamiento de datos, específicamente, mediante análisis descriptivo (95%) y predictivo (32%). Adicionalmente, realizan monitoreo y seguimiento, entregan informes y presentan sus resultados mediante la actividad de visualización. Su propuesta de valor es conceder información y conocimiento, y la mayoría de ellas utiliza como modelo de ingresos el cobro por suscripción (47%) o licencia (16%) y están dirigidas a otros negocios, es decir, al segmento B2B.

Una de las empresas que destaca en este grupo es TopicFlower, la cual es una empresa de origen argentino enfocada en el segmento B2B, que mediante diversos análisis presenta los temas de conversación más relevantes en los medios sociales, con el fin de ayudar a las empresas a definir su estrategia de medios y luego, determinar mediante indicadores, dónde y cómo está teniendo impacto la marca. Por otro lado, RelCase es una *Startup* dedicada al rubro judicial, que facilita el proceso de gestión y seguimiento de los casos judiciales en Chile, a través de una aplicación web. Esta *Startup* utiliza un modelo de ingresos del tipo suscripción y se dirige tanto a los segmentos de B2B y B2C, como al segmento de usuarios.

A continuación, se presenta un cuadro resumen con las principales características de cada clúster:

	Propuesta de Valor	Segmentos	Modelo de ingreso	Recursos Clave	Actividades Clave	Ejemplo
Clúster 0: Recolector de datos externos para hacer analítica avanzada	Información - Conocimiento	B2B y usuarios	No entregan información	- Fuentes externas: Clientes o socios y redes sociales	Agregación, análisis descriptivo, predictivo y prescriptivo; distribución y visualización	Qual Canal, Rocket.la y NoiseGrasp
Clúster 1: Recolector multi-fuente de datos y oferente de un servicio completo	Información - Conocimiento	B2B	Subscripción o no entregan información	- Fuentes externas: Clientes o socios y datos disponibles libremente en web - Fuentes internas: Monitoreo	Agregación, análisis descriptivo y predictivo; monitoreo, distribución y visualización	Wivo, BovControl y Datastom
Clúster 2: Monitoreo como un servicio	Información - Conocimiento	B2B, B2C y usuarios	Comisión	- Fuentes externas: Clientes o socios. - Fuentes internas: <i>Crowdsourced</i>	Agregación, monitoreo y visualización	MiCarga, Bycler y Rutanet
Clúster 3: Agregador de datos externos	Información - Conocimiento	B2B y B2C	<i>Freemium</i> , subscripción y comisión	- Fuentes externas: Clientes o socios	Agregación, distribución y visualización	Kurv, Packdocs y Lynk
Clúster 4: Generador de datos internos y analítica básica	Información – Conocimiento, Non-Data	B2B	No entregan información	- Fuentes internas: Monitoreo	Agregación, análisis descriptivo, monitoreo, distribución y visualización	Cuboverde, Monitor y Eye3
Clúster 5: Recolector de datos externos, monitoreo y análisis	Información - Conocimiento	B2B	Subscripción	- Fuentes externas: Clientes o socios, información disponible libremente en web y redes sociales	Agregación, análisis descriptivo, monitoreo, distribución y visualización	Eland, RelCase y TopicFlower

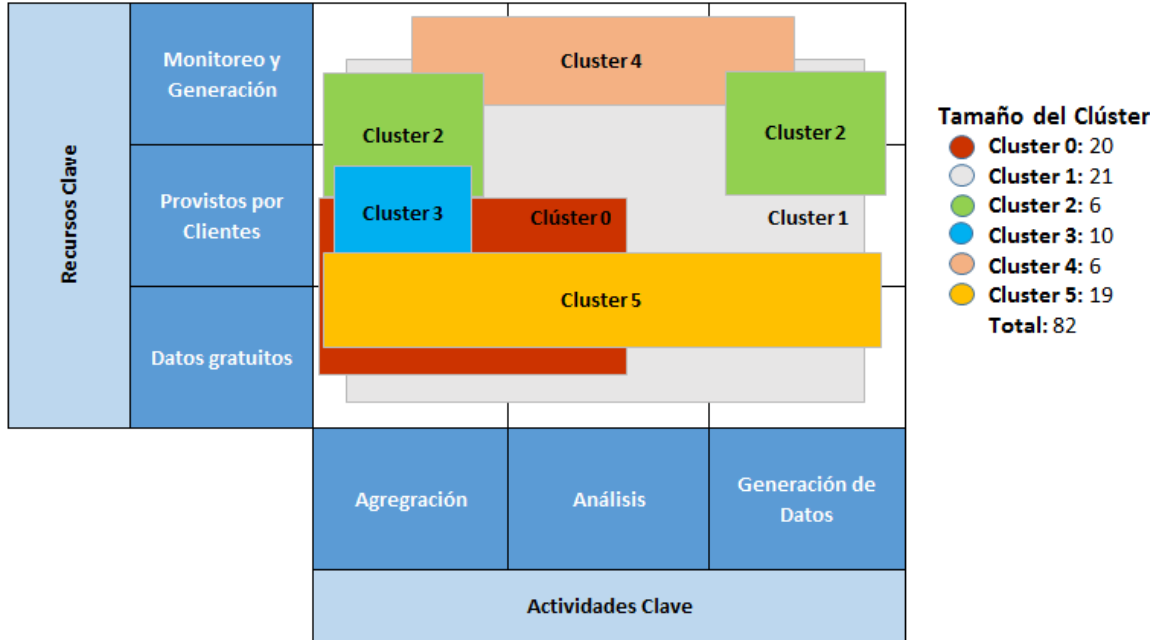
Tabla N°13: Resumen de clúster de Modelos de Negocios Basados en Datos

Fuente: Elaboración propia en base a resultados del Análisis *Clustering*

5.2.4. Mapa de Posicionamiento

Con el fin de representar gráficamente los diferentes clústeres de modelos de negocios basados en datos encontrados, con respecto al tipo de recurso clave que utilizan y a la actividad clave que llevan a cabo, se presenta el siguiente mapa de posicionamiento:

Mapa de Posicionamiento Modelos de Negocios *Startups* Latinoamericanas



CLÚSTER 0: RECOLECTOR DE DATOS EXTERNOS PARA HACER ANALÍTICA AVANZADA
CLÚSTER 1: RECOLECTOR MULTI-FUENTE DE DATOS Y OFERENTE DE UN SERVICIO COMPLETO
CLÚSTER 2: MONITOREO COMO UN SERVICIO
CLÚSTER 3: AGREGADOR DE DATOS EXTERNOS
CLÚSTER 4: GENERADOR DE DATOS INTERNOS Y ANALÍTICA BÁSICA
CLÚSTER 5: RECOLECTOR DE DATOS EXTERNOS, MONITOREO Y ANÁLISIS

Esquema N°4: Mapa de Posicionamiento Modelos de Negocios *Startups* Latinoamericanas

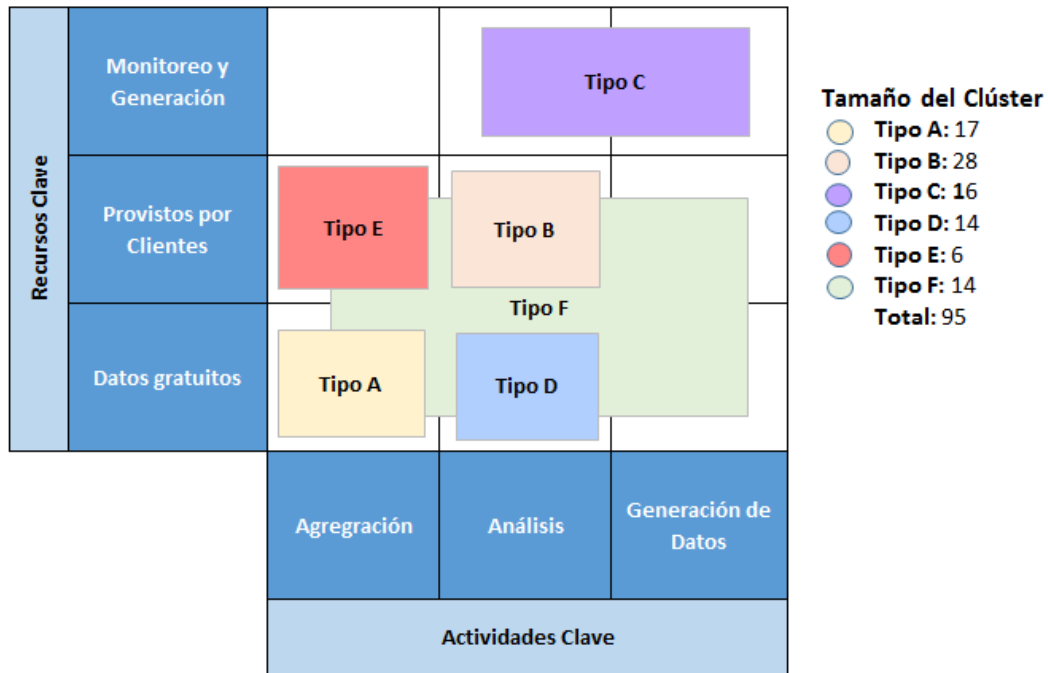
Fuente: Elaboración propia en base a resultados del Análisis *Clustering*

6. Capítulo V: Discusión

El presente estudio buscó encontrar los principales modelos de negocios basados en datos en el ámbito Latinoamericano, caracterizar sus principales componentes y agruparlos en clústeres o segmentos, con el fin de tener una visión sobre el panorama general de estos tipos de modelos y las tendencias en innovación que se están generando en la región en torno a este tema.

En primer lugar, y para comprender de mejor manera los resultados entregados en el capítulo anterior, se contrastará la información encontrada, con los resultados expuestos en el paper *Big Data for Big Business?* (Hartmann et al., 2014), el cual, como se ha mencionado a lo largo de este estudio, ha sido utilizado como guía para la investigación. En el esquema N° 5, se pueden apreciar los principales resultados de este estudio, que clasifica a los diferentes modelos de negocios encontrados -específicamente en Estados Unidos y parte Europa- de acuerdo a la fuente de obtención los datos y a la actividad clave que realizan con ellos, encontrando seis grupos o tipos de empresas.

Mapa Posicionamiento Estados Unidos y parte de Europa



TIPO A: FREE DATA COLLECTOR AND AGGREGATOR
TIPO B: ANALYTICS AS A SERVICE
TIPO C: DATA GENERATION AND ANALYSIS
TIPO D: FREE DATA KNOWLEDGE DISCOVERY
TIPO E: DATA AGGREGATION AS A SERVICE
TIPO F: MULTI SOURCE DATA MASH-UP AND ANALYSIS

Esquema N°5: Mapa Posicionamiento Estados Unidos y parte de Europa.

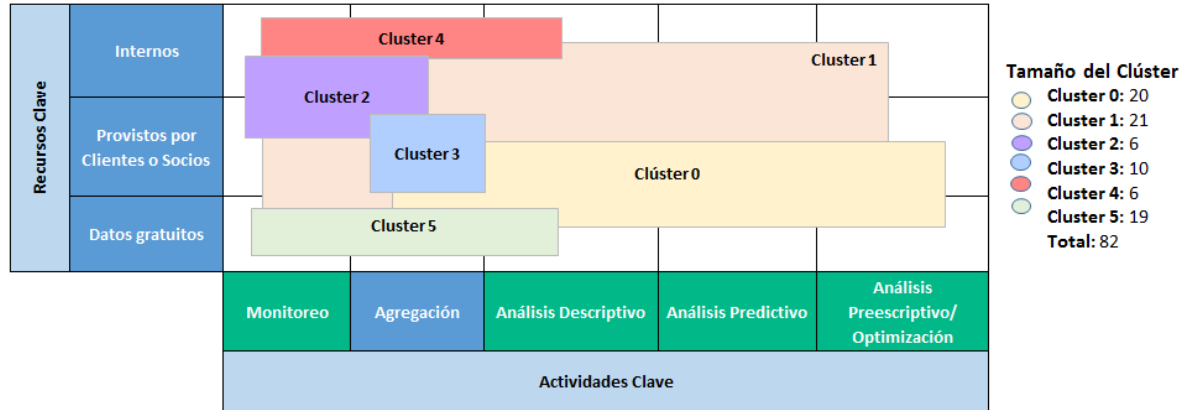
Fuente: Framework Modelos de Negocios Basados en Datos (Hartmann et al., 2014)

De acuerdo al mapa de posicionamiento presentado al inicio de este capítulo (Esquema N°5), correspondiente al trabajo de (Hartmann et al., 2014) y al mapa de posicionamiento presentado en el Capítulo de Resultados (Esquema N°4), que representa los hallazgos encontrados por el equipo de investigación, se puede observar que el clúster 3 sería muy similar al Tipo E, el cual contiene a empresas que realizan agregación de datos externos provistos por los clientes. Este grupo de empresas es relativamente más frecuente en las *Startups* analizadas en Latinoamérica. Adicionalmente, es posible identificar una similitud entre el clúster 4 y las empresas de Tipo C. En particular, las empresas del clúster 4 realizan agregación, análisis y generación de datos como actividad clave, sin embargo, las del Tipo C sólo realizan análisis y generación. Similarmente, en las empresas latinoamericanas es posible encontrar al clúster 1 que realiza la mayoría de las

actividades y utiliza todos los recursos claves. Éste clúster, es similar al Tipo F, pero no utiliza como recurso clave el monitoreo y la generación de datos. Respecto del clúster 2, es necesario notar que, dada las dimensiones establecidas en el mapa original de posicionamiento de (Hartmann et al., 2014), éste queda dividido en el espacio, dado que realiza actividades sobre recursos claves en cuadrantes que no han sido cubiertos por empresas de Estados Unidos y Europa.

Dado lo anterior, es posible notar que, las dimensiones escogidas en (Hartmann et al., 2014) no permiten distinguir claramente las características de los grupos encontrados en Latinoamérica dificultando su comparación y comprensión. Es así como, se ha decidido confeccionar un nuevo mapa de posicionamiento que incluya en la dimensión de actividades claves un mayor nivel de detalle respecto de las mismas. En particular, se han agregado y destacado con un color diferente las actividades Monitoreo, Análisis Descriptivo, Análisis Predictivo y Análisis Prescriptivo/Optimización que realizan las *Startups* Latinoamericanas, lo cual, permite identificar de mejor forma las actividades que llevan a cabo las empresas pertenecientes a los grupos encontrados y detallar cómo éstos se diferencian de la investigación presentada en el paper *Big Data for Big Business?* (Hartmann et al., 2014). A continuación, se presenta el nuevo mapa elaborado:

Nuevo Mapa de Posicionamiento Modelos de Negocios *Startups* Latinoamericanas



CLÚSTER 0: RECOLECTOR DE DATOS EXTERNOS PARA HACER ANALÍTICA AVANZADA
CLÚSTER 1: RECOLECTOR MULTI-FUENTE DE DATOS Y OFERENTE DE UN SERVICIO COMPLETO
CLÚSTER 2: MONITOREO COMO UN SERVICIO
CLÚSTER 3: AGREGADOR DE DATOS EXTERNOS
CLÚSTER 4: GENERADOR DE DATOS INTERNOS Y ANALÍTICA BÁSICA
CLÚSTER 5: RECOLECTOR DE DATOS EXTERNOS, MONITOREO Y ANÁLISIS

Esquema N°6: Nuevo Mapa de Posicionamiento Modelos de Negocios *Startups* Latinoamericanas

Fuente: Elaboración propia en base a resultados del Análisis *Clustering*

El esquema de (Hartmann et al., 2014) y la información que entregan en su estudio, se compara con la información expuesta por el equipo de trabajo en el capítulo de resultados junto con el nuevo mapa de posicionamiento, previamente presentado. Sus principales hallazgos se describen en la siguiente tabla resumen:

Hallazgo	<i>Big Data for Big Business?</i>	Tesis
Se encontraron empresas que utilizan como recurso clave datos internos y que realizan agregación	No	Sí
Se encontraron empresas que utilizan datos externos e internos y que realizan monitoreo, agregación y análisis	No	Sí
Se encontraron empresas que realicen análisis prescriptivo	No	Sí
Se encontraron empresas que adquieran (compre) datos	Sí	No
Se menciona como característica descriptora de los segmentos la estructura de costos	No	No
Cantidad de clústeres especializados en generar datos a través de un proceso de monitoreo	0	4
Cantidad de clústeres que sólo utilizan como recurso clave los datos internos	1	1
Cantidad de tipos de Modelos de Negocios especializados en solo una actividad clave	4	1

Tabla N°14: Tabla comparativa *Big Data for Big Business* – Tesis

Fuente: Elaboración propia en base a resultados del Análisis *Clustering*

Como se puede apreciar, los modelos de negocios encontrados en cada investigación varían, primordialmente, en los recursos y actividades clave que realizan, es así como la tendencia en *Big Data for Big Business?* (Hartmann et al., 2014) muestra que, en Estados Unidos y parte de Europa, las *Startups* que trabajan con datos se especializan en sólo una actividad clave, en realizar procesos de *Web Crawled* y en la existencia de empresas que compran datos. Mientras que, para el caso de Latinoamérica, los resultados encontrados hablan de una tendencia de los nuevos negocios a generar datos, además de obtenerlos de forma externa, a través del monitoreo y a dedicarse a más de una actividad clave, como es la agregación, el análisis y el propio monitoreo como servicio. Adicionalmente, es relevante mencionar que, para el contexto Latinoamericano, se encontraron *Startups* que realizan análisis prescriptivo o procesos de optimización situación no mencionada en el caso del estudio de países desarrollados. En relación a los cuadrantes

determinados como posibles modelos de negocios, para Latinoamérica todos estos están cubiertos o explorados por algún tipo de emprendimiento. Por otra parte, como puntos de similitud, es importante indicar que, en ambos estudios, la tendencia de las *Startups* Latinoamericanas es a recabar datos a través de fuentes externas, además, como actividad básica, la gran mayoría de las empresas realizan agregación. Finalmente, en las dos investigaciones existe un clúster dedicado sólo a recolectar información a través de sus clientes o socios y sólo a realizar como actividad clave, la agregación.

En relación a los aspectos que no fueron considerados en las líneas investigativas realizadas en el estudio o que faltó por desarrollar, en primer lugar, es relevante mencionar que, no se tomaron en cuenta aspectos más específicos de cada negocio y el entorno en que se desenvuelven, es así como no se incluyen características relacionadas a la cadena de suministro de las empresas, el nivel de competencia de los mercados en que participan, las capacidades de los equipos directivos, el potencial de sus recursos, entre otras particularidades, que permitan conocer en mayor profundidad cómo es el desempeño de las *Startups* en diversos escenarios o en distintos roles y, de esta manera, tener información más completa y concluyente. Además, como se mencionó en los resultados de las entrevistas, una arista que el equipo no consideró y que fue descrita por la mayoría de las *Startups*, que tiene que ver con los procesos de consultoría que realizan estas empresas y que podrían ser considerados como una nueva actividad clave dentro del marco teórico presentado en la revisión literaria. Asimismo, durante la investigación no se hizo distinción entre los datos proporcionados por clientes o socios, lo cual impidió determinar la importancia que tiene cada actor y el nivel de incidencia en el desempeño de las *Startups*.

Por otro lado, el equipo considera que, según los resultados de la investigación y el mapa de posicionamiento presentado anteriormente, existen potenciales oportunidades que podrían explotarse por algún nuevo modelo de negocios, siempre y cuando estos emprendimientos incluyan como actividad clave, los análisis predictivos y prescriptivos, los cuales son escasamente desarrollados hoy en día, sobre todo aquellos realizados a partir de datos internos como recurso clave, donde se observa que sólo dos segmentos están cubriendo esta parte del mapa. Otro cuadrante que podría aprovecharse es aquel que considera el monitoreo en base a datos

disponibles gratuitamente, ya que actualmente estos datos están siendo utilizado, principalmente, sólo como input para realizar análisis estáticos.

Respecto de la adquisición de datos, si bien es un proceso que no se realiza frecuentemente en Latinoamérica, el equipo consultor considera que, dada la antigüedad de las *Startups* de la región, esta actividad aún no ha sido necesaria de implementar pero, a medida que estos emprendimientos vayan creciendo y compitiendo cada vez de forma más global, la adquisición de datos se convertirá en un aspecto que toda empresa deberá tener cuenta, por ende, se recomienda a las nuevos negocios analizar este ámbito como una futura actividad clave que podrían desarrollar.

Finalmente, una de las principales diferencias que se aprecian con el estudio de (Hartmann et al., 2014) y que ha sido reflexionado por el equipo investigador, es aquella relacionada a que los hallazgos en Latinoamérica muestran una clara tendencia de las *Startups* a la realización de múltiples actividades, más que sólo especializarse en una actividad clave en particular como lo presentado en el trabajo de los autores mencionados anteriormente. En este sentido, sería interesante evaluar en un futuro cómo fue el desempeño de estas empresas o cuáles de ellas sobrevivieron, con el fin de determinar qué enfoque fue el más exitoso o qué lecciones se podrían extraer a partir de las diferencias en las orientaciones encontradas.

7. Capítulo VI: Conclusiones

El aumento exponencial de la generación de datos *-Big Data-* ha generado grandes cambios no sólo a nivel de cómo las empresas se relacionan con sus clientes, sino que también en la forma en que las empresas desarrollan sus modelos de negocios. El *Big Data* ha impactado en cada uno de los procesos de las organizaciones convirtiéndose en un recurso vital en la creación de ventajas competitivas sostenibles. Ahora bien, dado este nuevo escenario, marcado por la innovación y el aumento de la incertidumbre, es que nacen como principales agentes de cambio las *Startups*, las cuales se caracterizan por poseer mayor flexibilidad y capacidad de adaptación que una empresa ya constituida. Junto con lo anterior, Latinoamérica se ha convertido en un gran foco del desarrollo e implementación de nuevas ideas, principalmente, debido al crecimiento de las economías pertenecientes a la región. Este panorama convierte a esta zona en un ecosistema atractivo para el surgimiento y proliferación de nuevos negocios, generando así, mayor interés para realizar investigaciones enfocadas en los modelos de negocios que se están desarrollando en este territorio.

De esta manera, el equipo investigador se planteó identificar cuáles son los modelos de negocios basados en datos que las *Startups* Latinoamericanas han implementado y qué atributos y propiedades los caracterizan. Es así como este trabajo comenzó con una revisión de la literatura relacionada a los aspectos mencionados anteriormente, *Big Data*, Modelos de Negocios y *Startup*, analizando la situación actual de los estudios de estos conceptos e identificando qué autores y qué trabajos ayudarían a guiar la investigación. Sin ir más lejos, el trabajo realizado, se basó principalmente en los trabajos de (Hartmann et al., 2014), (Parmar et al., 2014) y (Díaz & Zaki, 2015). Posteriormente, se describió la propuesta de investigación y los objetivos del estudio (ya planteados anteriormente), tanto generales como específicos. Luego, se definió la metodología a utilizar, la cual, se sustenta en el modelo CRIPS-DM (Pete et al., 2000), donde se desarrollaron las diferentes etapas que ésta posee: Entendimiento del Negocios, Entendimiento de los Datos, Preparación de los Datos, Modelamiento, Evaluación e Implementación. Particularmente, en esta sección se destacan las dos líneas investigativas que se llevaron a cabo por el equipo de trabajo, las cuales se realizaron mediante entrevistas presenciales semiestructuradas y análisis de páginas web. Estas etapas se estudiaron a través de un proceso de análisis de contenido, para las

entrevistas, y de un proceso de análisis cuantitativo, para los resultados de las páginas web. Además, en este capítulo se definieron los indicadores para evaluar el desempeño de los resultados obtenidos. El siguiente capítulo, presentó precisamente los resultados encontrados a través del estudio, indicando, en primer lugar, las conclusiones de las entrevistas realizadas, luego la estadística descriptiva de la información recabada y los clústeres encontrados a través del análisis cuantitativo. Estos grupos o segmentos fueron definidos de acuerdo a sus características distintivas, además de presentarlos en un mapa de posicionamiento para facilitar su entendimiento. Finalmente, en el capítulo de discusión, se plantearon los principales hallazgos de la investigación, detallando las características nuevas encontradas, así como los aspectos comunes y distintos respecto del trabajo realizado por (Hartmann et al., 2014) en *Big Data for Big Business?*.

Dado lo anterior, entre las principales conclusiones que se desprenden, se puede mencionar que, los modelos de negocios que se están desarrollando en la región en su mayoría se diferencian en diversos aspectos de los presentes en Estados Unidos y parte de Europa, según lo expuesto por (Hartmann et al., 2014). La primera distinción importante es que en Latinoamérica se ha implementado una analítica más avanzada a través del análisis prescriptivo encontrado en algunas empresas en particular. Además, si bien la fuente de datos por excelencia en ambos estudios corresponde a datos externos, en el presente estudio, se aprecia una clara tendencia a la utilización de datos propios de las empresas, generados a partir de actividades de monitoreo u otras. De acuerdo a la dimensión de recursos clave, se aprecia que, ninguna de las *Startups* realiza la compra de datos; las razones de esto no fueron posibles determinarlas con certezas por el equipo investigador, pero se presume que esto ocurre debido a que, dado que son empresas en formación, aún no tienen recursos suficientes para destinarlos a este tipo de actividades, sino que priorizan en otras más relevantes para su operación. Por otro lado, en relación al análisis web, se encontró que ninguna de las empresas presentaba su estructura de costos, lo cual fue un impedimento, y una limitación del trabajo, para integrar este atributo en la conformación de los clústeres y no fue posible analizar su eventual o no trascendencia en ellos. Sin embargo, de los resultados de la etapa cualitativa se tiene una idea general respecto que en Latinoamérica, los principales costos estarían relacionados al gasto en RRHH, los cuales serían mucho mayores que los asociados a los componentes tecnológicos de desarrollo de software e implementación de hardware.

Con respecto a los Patrones de Innovación (Parmar et al., 2014), queda en evidencia que la mayoría de las *Startup* están en etapas de desarrollo y que, no están totalmente conectadas con estos patrones, pero que al plantearlos señalaron que eran puntos interesantes de explotar, sobre todo el Patrón 5, donde indicaron que sería positivo para sus negocios estandarizar sus productos y venderlos como un activo, dedicándose al asesoramiento del uso de estos, lo cual en muchas de ellas ya estaba en sus planes futuros o en discusión sobre si implementarlo o no.

De acuerdo a los objetivos generales planteados en este trabajo, a través de esta investigación, se lograron identificar qué modelos de negocios basados en datos están desarrollando las *Startups* en Latinoamérica, junto con identificar las principales características de éstos. Además –con respecto a los objetivos específicos- se indagó la realidad de las *Startups* latinoamericanas dedicadas al tratamiento y gestión de los datos, comparando los grupos de empresas encontrados con el estudio de (Hartmann et al., 2014), así como también, se identificaron las áreas potencialmente explotables para futuros emprendimientos. Para el caso de los patrones de innovación, se entrevistó a *Startups* chilenas que permitieron ejemplificar de qué manera utilizan los datos para innovar en sus modelos de negocios.

En relación a las nuevas aristas que esta investigación presenta, es importante recalcar los puntos inexplorados en la región (y planteados en el Capítulo de Discusión) que podrían convertirse en futuros modelos de negocios, es así como este trabajo enfatizó estas nuevas oportunidades en aquellas empresas que podrían dedicarse a análisis de mayor complejidad como lo sería los del tipo predictivo, prescriptivo u optimización, dado que, los clientes cada vez más necesitan tomar decisiones con la mayor efectividad posible, y hoy es un ámbito que aún está en desarrollo.

Adicionalmente, existen otros modelos de negocios que aún no se ha explotado en su totalidad, el cual tiene relación a las actividades de monitoreo a través de datos disponibles gratuitamente en internet. Respecto de la adquisición de datos, el estudio encontró, que es una actividad que no se realiza por las *Startups* de la región, lo cual sugiere dos interpretaciones, la primera, es que este tipo de actividad no agrega valor y ventajas competitivas y, la segunda, tiene relación a la capacidad económica de los emprendimientos latinoamericanos para acceder a datos pagados, la cual, dada la etapa inicial en la que se encuentran, no es lo suficientemente solvente. Por último, una situación digna de destacar, tiene que ver con el hallazgo de *Startups* o modelos de negocios

que realizan más de una actividad clave, lo cual, se diferencia de lo encontrado por (Hartmann et al., 2014), quienes identificaron varios modelos de negocios con mayor nivel de especialización. Esta situación presenta grandes desafíos para quienes continúen con investigaciones relacionadas a este ámbito dado que, sería relevante conocer, qué tipos de modelos de negocios han sido más exitosos en el tiempo y por qué razones. En concordancia con lo anterior, es importante mencionar que, los futuros trabajos que aborden la situación del *Big Data* en Latinoamérica deben considerar, entre otros aspectos, la inclusión de una nueva actividad clave relacionada a la gestión de datos y/o servicios de asesoría. Esta actividad, como se planteó en la mayoría de las entrevistas realizadas, aborda un tipo de prestación que se otorga post realización del trabajo directo con los datos y que se enfoca en guiar y acompañar a los clientes en el proceso de toma de decisiones cuando ya han accedido a los resultados. En este estudio, esta característica no se midió y, de acuerdo a lo investigado, es una función que podría volverse clave en los próximos años. En línea con lo anterior, para realizar esta investigación solo se entrevistaron a *Startups* que se han desarrollado en Chile, por lo cual, es un desafío relevante para los futuros trabajos, incluir entrevistados de toda Latinoamérica, ya que los resultados expuestos en este trabajo podrían presentar variaciones, sobre todo a nivel de los patrones de innovación que las *Startups* están desarrollando. Otro punto que es conveniente sopesar, está vinculado con la distinción entre Socio y Cliente que, tanto esta investigación como la de (Hartmann et al., 2014) no realizan, lo cual podría evidenciar posibles relaciones que estos trabajos no constataron.

Por otra parte, es necesario mencionar que, una de las limitaciones de este trabajo, y de los presentados en la revisión literaria, tiene relación a la temporalidad del análisis realizado de las *Startups*, ya que este tipo de investigaciones no presentan un seguimiento a la evolución y el desarrollo de estas empresas, por lo cual, al tratarse de estudios estáticos, no se podrá conocer cómo estos modelos de negocios cambiarán, se adaptarán y si sobrevivirán en el futuro. Finalmente, es relevante indicar que, esta investigación constituye un acercamiento de la situación actual que viven las *Startups* en Latinoamérica más no presenta resultados cien por ciento concluyentes, por ende, estudios similares podrían presentar resultados con cierto nivel de diferencia, refutando o confirmando de forma más acabada las conclusiones exploratorias de esta tesis.

8. Referencias

- Afuah, A., & Tucci, C. L. (2001). *Internet Business Models and Strategies: Text and Cases*. McGrawHill international editions Management organization series (Vol. 2). McGraw-Hill Higher Education. Retrieved from <http://dl.acm.org/citation.cfm?id=579520>
- Amit, R., & Zott, C. (2001). Value creation in E-business. *Strategic Management Journal*, 22(6-7), 493–520. <http://doi.org/10.1002/smj.187>
- Angel List. (2015). AngelList. Retrieved from <https://angel.co/>
- Blank, S. (2013). Why the Lean Start-UP Changes Everything. Retrieved from <https://hbr.org/2013/05/why-the-lean-start-up-changes-everything>
- Burkhart, T., Krumeich, J., Werth, D., & Loos, P. (2011). Analyzing the Business Model Concept — A Comprehensive Classification of Literature. *ICIS 2011 Proceedings*. Retrieved from <http://aisel.aisnet.org/icis2011/proceedings/generaltopics/12>
- Chen, M., Mao, S., & Liu, Y. (2014). *Big Data: A survey*. *Mobile Networks and Applications*, 19(2), 171–209. <http://doi.org/10.1007/s11036-013-0489-0>
- Díaz, D. (2015). *Tecnologías de Información y Comunicaciones (TICs) y su rol en la Innovación*. Working Paper Series. Centro de Innovación para el Desarrollo. WP2015-06. Retrieved from <http://cid.uchile.cl/wp/WP-2015-06.pdf>
- Díaz, D., & Zaki, M. (2015). *Innovación en Modelos de Negocios Basados en Datos: Los “Big” y los “No Tanto.”* Working Paper Series. Centro de Innovación para el Desarrollo. WP2015-07. Retrieved from <http://cid.uchile.cl/wp/WP-2015-07.pdf>
- Gantz, B. J., Reinsel, D., & Shadows, B. D. (2012). *Big Data , Bigger Digital Shadow s , and Biggest Grow th in the Far East Executive Summary: A Universe of Opportunities and Challenges*. *Idc, 2007*(December 2012), 1–16.
- García, C., & Gómez, I. (2006). Algoritmos de aprendizaje: knn & kmeans. Retrieved from <http://www.it.uc3m.es/jvillena/irc/practicas/08-09/06.pdf>
- Gartner. (2012). *Big Data Analytics-Free* Gartner Research . Retrieved from <http://www.gartner.com/it-glossary/big-data/>
- Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2014). *Big Data for Big Business?* *Cambridge Service Alliance Blog*, 1–29. Retrieved from http://cambridgeservicealliance.blogspot.co.uk/2014/04/big-data-for-big-business_3.html

- Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology*. Facultyunlvedu (Vol. 15). <http://doi.org/10.2307/2288384>
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big Data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute, (June), 156. <http://doi.org/10.1080/01443610903114527>
- Maurya, A. (2012). *Running Lean: Iterate from Plan A to a Plan That Works*. Retrieved from https://books.google.cl/books?id=j4hXPn233UYC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false
- Mayo, M. C., & Brown, G. S. (1999). BUILDING A COMPETITIVE business model. *Ivey Business Journal*, 63(3), 18–23. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib&db=buh&AN=2026738&site=ehost-live>
- McAfee, A., & Brynjolfsson, E. (2012). *Big Data*. The management revolution. *Harvard Business Review*, 90(10), 61–68. <http://doi.org/10.1007/s12599-013-0249-5>
- Morris, M., Schindehutte, M., & Allen, J. (2005). The entrepreneur's business model: Toward a unified perspective. *Journal of Business Research*, 58(6), 726–735. <http://doi.org/10.1016/j.jbusres.2003.11.001>
- Mullins, J., & Komisar, R. (2009). Getting to Plan B: Breaking Through to a Better Business Model. *Harvard Business Review Press*, 7.2. Retrieved from <https://hbr.org/product/getting-to-plan-b-breaking-through-to-a-better-bus/an/2669-HBK-ENG>
- Neuendorf, K. a. (2002). A Flowchart for the Typical Process of Content Analysis Research from The Content Analysis Guidebook , by Kimberly A . Neuendorf. *Online*, 5–7. Retrieved from <http://academic.csuohio.edu/kneuendorf/content/resources/flowc.htm>
- OECD. (2013). *Startup América Latina*. Promoviendo la innovación en la región. *OECD Development Centre*, 228. <http://doi.org/10.1787/9789264202320-es>
- Osterwalder, A., & Pigneur, Y. (2010). *Business Model Generation*. self published. Retrieved from <http://www.consulteam.be/media/5985/businessmodelgenerationpreview.pdf>
- Parmar, R., Mackenzie, I., Cohn, D., & Gann, D. (2014). The New Patterns of Innovation. *Harvard Business Review*, 92(1-2), 86. Retrieved from

<http://search.ebscohost.com.ezproxy.library.wisc.edu/login.aspx?direct=true&AuthType=ip,uid&db=bth&AN=93302832&site=ehost-live&scope=site\nhttps://hbr.org/>

- Pete, C., Julian, C., Randy, K., Thomas, K., Thomas, R., Colin, S., & Wirth, R. (2000). Crisp-Dm 1.0. *CRISP-DM Consortium*, 76. Retrieved from <https://the-modeling-agency.com/crisp-dm.pdf>
- Pohle, G., & Chapman, M. (2006). Strategy & Leadership IBM's global CEO report 2006: business model innovation matters IBM's global CEO report 2006: business model innovation matters. *Strategy & Leadership Strategy & Leadership Journal of Business Strategy Iss Journal of Business Strategy*, 34(1), 34–40. Retrieved from <http://dx.doi.org/10.1108/10878570610701531>
- Ries, E. (2012). *El método Lean Startup*. Retrieved from <https://books.google.cl/books?id=tvfyz-4JlLwC&printsec=frontcover&dq=lean+Startup&hl=es-419&sa=X&ved=0CB4Q6AEwAGoVChMI1dC97t3myAlVxBKQCh1rpwnN#v=onepage&q=lean+Startup&f=false>
- Sagioglu, S., & Sinanc, D. (2013). *Big Data: A review. International Conference on Collaboration Technologies and Systems (CTS)*, 42–47. <http://doi.org/10.1109/CTS.2013.6567202>
- Schumpeter, J. A. (1934). The theory of economic development: an inquiry into profits, capital, credit, interest, and the business cycle. *Harvard Economic Studies*, 46(2), xii, 255 p. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1496199
- Start-Up Chile. (2015). Start-Up Chile. Retrieved from <http://www.Startupchile.org/>
- Stewart, D. W., & Zhao, Q. (2000). Internet Marketing, Business Models, and Public Policy. *Journal of Public Policy & Marketing*, 19(2), 287–296. <http://doi.org/10.1509/jppm.19.2.287.17125>
- TechTarget. (2013). *Big Data en América Latina avanza a pasos pequeños*. Retrieved from <http://searchdatacenter.techtarget.com/es/cronica/Big-Data-en-America-Latina-avanza-a-pasos-pequenos>

- TechTarget. (2014). Encuesta de Prioridades de TI en América Latina 2014. Retrieved from <http://searchdatacenter.techtarget.com/es/cronica/2014-Un-ano-de-proyectos-internos-de-TI>
- Vesset, D., Morris, H. D., Eastwood, M., Woo, B., Villars, R. L., Bozman, J. S., & Olofson, C. W. (2015). Worldwide *Big Data* Technology and Services 2012 – 2015 Forecast. *IDC Market Analysis*, (March 2012). Retrieved from <http://www.mendeley.com/catalog/worldwide-big-data-technology-services-2012-2015-forecast/>
- World Economic Forum. (2012). *Big Data*, Big Impact: New Possibilities for International Development. Retrieved from http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf

9. Anexos

9.1. Anexo N°1: Listado de empresas entrevistadas

Nº	Nombre	País	Contacto	Cargo	Página Web
1	WholeMeaning	Chile	Juan Besa	CTO	wholemeaning.com
2	NoiseGrasp	Chile	Andrés Groisman	CEO	noisegrasp.com
3	Semant.io	Chile	Nicolás Bravo	Gerente de Ventas	semant.io
4	RutaPro	Chile	Freddy Cea	Gente Comercial	rutapro.cl
5	Mi Entorno	Chile	Cristián Araneda	Gerente de Estudios y Cartografía	mientorno.cl
6	Desert Point	Chile	José Campino	CEO	gausscontrol.com
7	Agronometrics	Chile	Óscar Azocar	Gerente de Operaciones	agronometrics.com

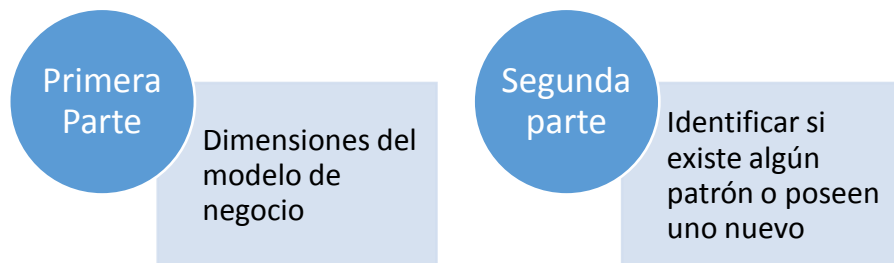
9.2. Anexo N°2: Boceto Entrevista semiestructurada *Startups*

Buenos días (tardes), mi nombre es _____. Somos tesistas de Ingeniería Comercial de la Facultad de Economía y Negocios de la Universidad de Chile. Actualmente estamos realizando una investigación que pretende identificar y caracterizar los modelos de negocios basados en datos (que generen y/o gestionen datos y los utilicen para distintos fines), específicamente enfocados en la Start-up de Chile y de Latinoamérica.

Hemos agendado esta reunión para conversar principalmente sobre el modelo de negocios de la organización a la cual usted pertenece y las preguntas estarán enfocadas a las diferentes dimensiones del modelo Canvas.

Preguntas Introductorias

- ✓ ¿De qué se trata su empresa?
- ✓ ¿Quiénes trabajan actualmente en ella?
- ✓ ¿Cómo nació el modelo de negocios?
- ✓ ¿Cómo ha evolucionado su negocio?
- ✓ ¿Cómo ha sido el camino recorrido para la realización de esta organización?



Primera Parte: Dimensiones del Modelo de Negocio

Dimensión: Recursos clave

1. ¿De qué tipo de fuente extrae los datos que utiliza su organización?
 - a. Si responde “Internos”, preguntar si son datos “Existentes” o “Generados por ellos mismos”
 - b. Si son “Generados por ellos mismos”, preguntar si son “Crowdsourced” o “Tracked, generated & other”.
 - c. Si responde “Externos”, preguntar si son datos “Adquiridos de otras empresas”, “Provisionadas por el cliente” o “Disponible gratuitamente”.
 - d. Si responde “Disponible gratuitamente”, preguntar si son “Datos descargables libremente desde internet” o “social-media data” o “Web Crawled Data”
 - e. Si responde “Otro”, preguntar a qué se refiere.

Dimensión: Actividades clave

2. ¿Cuál de estas alternativas considera usted que es una actividad clave respecto del uso de datos en su organización?
 - a. Generación de datos
 - b. Adquisición de datos
 - c. Procesamiento de datos
 - d. Análisis de datos
 - e. Visualización de datos
 - f. Distribución de datos
 - g. Agregación de datos
 - h. Otro
 - i. Si responde “Generación de datos”, preguntar mediante qué mecanismos los genera, si es “Crawling” o “Tracking & Crowdsourcing”?
 - j. Si responde “Análisis de datos”, preguntar qué tipo de análisis realiza, Descriptivo, Predictivo o prescriptivo?
 - k. Si responde “Otro”, preguntar a qué se refiere.

Dimensión: Propuesta de valor

3. ¿Qué tipo de propuesta de valor vinculada al *Big Data* poseen?
 - a. Vende datos
 - b. Procesa o interpreta la información para luego venderla
 - c. Ofrece un producto o servicio no virtual.
 - d. Otro

Dimensión: Segmento de clientes

4. ¿Quiénes son sus clientes?
 - a. Las empresas (B2B)
 - b. Consumidores individuales (B2C)
 - c. Otro

Dimensión: Modelo de ingresos

5. ¿Cuál es el modelo de ingresos que utiliza su organización?
 - a. Venta de activos físicos
 - b. Préstamo, arriendo o leasing
 - c. Licencia
 - d. Por comisión
 - e. Suscripción
 - f. A través de Publicidad
 - g. Pago por uso
 - h. *Freemium*
 - i. Otro

Dimensión: Estructura de costos

6. ¿Tiene identificada o definida su estructura de costos?
7. ¿Qué relevancia o porcentaje ocupan las actividades relacionadas al uso de datos dentro de su estructura de costos?
8. ¿Los datos que se generan en su organización son propios de la producción de su producto o servicio o debe producirlos de forma paralela a estos?

9. ¿Cuáles son los principales costos en los que incurren para recabar y/o generar datos?

Segunda Parte: Patrones de Innovación

Patrón 1: Uso de datos para mejorar productos

10. ¿Qué datos se generan a través de sus productos, servicios y/u operaciones y cuáles están utilizando o recopilando?
11. ¿Existen datos que podría empezar recopilar?
12. ¿Podrían agregar valor en la oferta a nuestros clientes?

Patrón 2: Digitalización de activos

13. ¿Han digitalizado activos físicos en su organización?
14. ¿Poseen activos que podrían digitalizarse?
15. ¿Qué tipo de activos han digitalizado dentro de la organización?
16. ¿Si posees activos digitales, estos podrían ser mejorados a través del fortalecimiento de su componente digital?

Patrón 3: Combinar datos desde distintas industrias

17. ¿Su organización utiliza datos provisionados por otras empresas?
18. ¿Tienen alianzas o asociaciones con otras empresas en donde comparten datos?
19. ¿Su organización tiene la capacidad de catalizar los datos de otras empresas y crear valor?
20. ¿Considera beneficioso para su organización compartir datos e información con otros negocios?

Patrón 4: Generar y condensar información

21. ¿Su organización posee datos o información que puede ser beneficiosa internamente, para potenciales consumidores o para otras empresas?

Patrón 5: Estandarizar procesos para luego venderlos

22. ¿Su organización posee datos o procesos relacionados a ellos que lo distinguen de la competencia?
23. ¿Esta actividad distintiva es apreciada por la competencia?
24. ¿Esta actividad distintiva puede estandarizarse y “venderse” como un nuevo producto/servicio?

Bueno, hemos terminado. Muchas gracias por su tiempo. Por último, queremos saber si desea recibir información sobre los resultados de nuestra investigación.

9.3. Anexo N°3: Material de apoyo Entrevista semiestructurada *Startups*



Material de apoyo

Entrevista semi-estructurada Start-ups

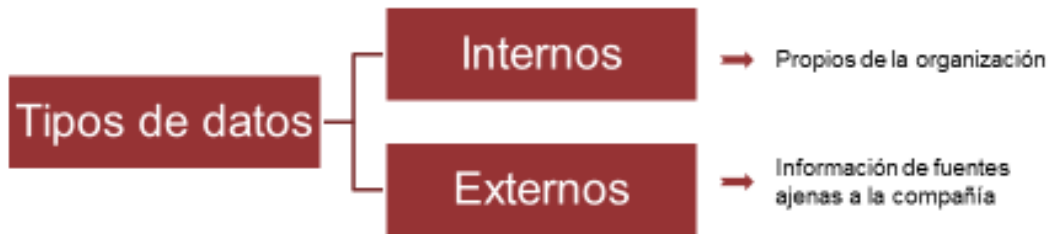
Profesor Guía: David Díaz S.

Myria Alcalino, Valeska Arenas y Felipe Gutiérrez

Dimensiones

Big Data for Big Business?

¿De qué tipo de fuente extrae los datos que utiliza su organización?



Facultad de Economía y Negocios Universidad de Chile

Fuente de datos internos



Facultad de Economía y Negocios Universidad de Chile

Fuente de datos externos



Facultad de Economía y Negocios Universidad de Chile

¿Cuál (es) de estas actividades clave realiza su organización?

Agregación	→	Condensar o integrar datos
Análisis	→	Estudio de los datos
Adquisición	→	Comprar datos
Procesamiento	→	Tratamiento de los datos
Generación	→	Crear nuevos datos
Distribución	→	Entregar los datos
Visualización	→	Presentación de datos

Facultad de Economía y Negocios Universidad de Chile

¿Qué tipo de propuesta de valor vinculada al Big data poseen?

Datos

- Hechos o mediciones de la realidad sin interpretación añadida.

Información y conocimiento

- Procesamiento de los datos dándoles contexto y sentido.

Non-data product/service

- Venta de productos no digitales pero que si están relacionado al tratamiento de datos.

Facultad de Economía y Negocios Universidad de Chile

¿Cuál es el modelo de ingresos que utiliza su organización?



¿Quiénes son sus clientes?



Facultad de Economía y Negocios Universidad de Chile

Patrones

The New Patterns of Innovation

Patrones

1. Mejorar productos para generar datos

- **Idea Principal:** Extraer información de los productos y servicios con el fin de mejorar su desempeño, desde el punto de vista de la productividad y la eficiencia.
- Actualmente se puede acceder a mayor información de los productos y servicios, esto es posible gracias a sensores, la comunicación inalámbrica, entre otros mecanismos.
- Ejemplo:



Facultad de Economía y Negocios Universidad de Chile

Patrones

2. Digitalización de los activos físicos

- **Idea Principal:** Digitalizar los productos o servicios con el objetivo de facilitar su tratamiento y funcionalidad, y así ofrecer mayor valor al cliente
- Esto ha permitido la transformación en la forma que las compañías operan en sus industrias, permitiendo reducir costos de distribución y administrar el inventario de forma más eficiente.
- Ejemplo:



Facultad de Economía y Negocios Universidad de Chile

Patrones

3. Combinación de datos dentro y a través de las industrias

- **Idea Principal:** Utilizar la información (datos) desde distintas industrias para generar nuevos modelos de negocios.
- Un ejemplo de esto, son las empresas que integran información a partir de su cadena de suministro.
- Existen negocios que recolectan la información, la estandarizan y "limpian" para luego entregarlas a las compañías que lo requieran.
- Ejemplo:

ZARA

Facultad de Economía y Negocios Universidad de Chile

Patrones

4. Comercio de datos

- **Idea Principal:** Condensar información de distintas fuentes (en especial la información que genera el propio negocio) para generar nuevas oportunidades de negocio. Por ejemplo, la venta de la información a otras empresas.
- El desafío es integrar y compartir la información disponible, la cual se presenta en diferentes formatos, generalmente no estructurados.
- Ejemplos:



Facultad de Economía y Negocios Universidad de Chile

Patrones

5. Codificar una capacidad de servicio distintivo

- **Idea Principal:** Generar y estandarizar procesos “únicos” y venderlos como un nuevo producto.
- Cualquier proceso relacionado a las TI que sea mejor en su clase -pero no fundamental para la ventaja competitiva de una empresa- puede, por lo tanto, ser convertido en un negocio rentable.
- No solo se refiere a la creación de nuevas alternativas de negocio sino que también existen empresas que han utilizado estos procesos para ampliar su oferta y el valor para los clientes.
- Ejemplo:



Facultad de Economía y Negocios Universidad de Chile

Material de apoyo

Entrevista semi-estructurada Start-ups

Profesor Guía: David Díaz S.

Myrta Alcaino, Valeska Arenas y Felipe Gutiérrez

9.4. Anexo N°4: Listado de empresas analizadas

N°	Nombre	País	Dirección web	Clúster
1	Adere.so	Chile	postcenter.io	Clúster 0
2	Bien.io	México	bien.io	Clúster 0
3	Comenta TV	Argentina	comenta.tv	Clúster 0
4	Comparabien.com	Perú	http://comparabien.com	Clúster 0
5	e-restó	Chile	e-resto.cl	Clúster 0
6	Eskalab	Colombia	www.eskalab.co	Clúster 0
7	Geniusly	Colombia	geniusly.co	Clúster 0
8	Innuy Powerful Software Products.	Uruguay	innuy.com	Clúster 0
9	Lagiar	Brasil	lagiar.com	Clúster 0
10	Mis Procesos	Chile	misprocesos.cl	Clúster 0
11	NoiseGrasp	Chile	noisegrasp.com	Clúster 0
12	Odysci	Brasil	odysci.com	Clúster 0
13	Percycle	Brasil	percycle.com	Clúster 0
14	Puerto Finanzas	Argentina	puertoфинanzas.com	Clúster 0
15	Qual Canal	Brasil	qualcanal.tv	Clúster 0
16	Rocket.la	México	rocket.la	Clúster 0
17	Sellywhere	Chile	www.sellywhere.com	Clúster 0
18	TimePoints	Chile	http://www.timepoints.cl	Clúster 0
19	WholeMeaning	Chile	wholemeaning.com	Clúster 0
20	YADA	México	yada.mx	Clúster 0
21	Agribots	Chile	www.agribots.com	Clúster 1
22	BovControl	Brasil	bovcontrol.com	Clúster 1
23	Colibrí	México	http://www.colibriteq.com/	Clúster 1
24	Crowdtask	Brasil	crowdtask.me	Clúster 1
25	Datastorm	Brasil	datastorm.com.br	Clúster 1
26	Desert Point Technologies	Chile	www.desertpointtech.com	Clúster 1
27	Eyso	Brasil	eyso.io	Clúster 1
28	Flipter	Chile	flipter.com	Clúster 1
29	FRACTTAL	Chile	www.fractal.com	Clúster 1
30	Hitmap	Chile	mientorno.com	Clúster 1
31	Mi entorno	Chile	mientorno.cl	Clúster 1
32	NAZAR	Brasil	http://nazar.io	Clúster 1
33	Nethub	Colombia	nethub.co	Clúster 1
34	RutaPro	Chile	rutapro.cl	Clúster 1
35	Sami	Chile	samimetrics.com	Clúster 1
36	SciCrop	Brasil	scicrop.com	Clúster 1

N°	Nombre	País	Dirección web	Clúster
38	Suruna	Perú	http://suruna.com	Clúster 1
39	Tapptus	Colombia	tapptus.com	Clúster 1
40	Wivo	Chile	wivo.cl	Clúster 1
41	Zolvers	Argentina	zolvers.com	Clúster 1
42	Bycler	Chile	www.bycler.com	Clúster 2
43	MiCarga	Colombia	micarga.com	Clúster 2
44	Rutanet	México	http://rutanet.com	Clúster 2
45	Tarefa	Colombia	tarefa.co	Clúster 2
46	Torneo de Ideas	México	torneodeideas.com	Clúster 2
47	Webcarga	Chile	www.webcarga.com	Clúster 2
48	Biletu.com	Chile	http://www.biletu.com	Clúster 3
49	DIVERZA	México	diverza.com	Clúster 3
50	Expensas Online	Argentina	http://expensasonline.pro/	Clúster 3
51	Kurv	Argentina	http://kurv.co	Clúster 3
52	Lynk	Chile	lynk.cl	Clúster 3
53	Meetcard	Chile	meetcard.cl	Clúster 3
54	NODO	Chile	http://www.nodoapp.com	Clúster 3
55	Packdocs	Brasil	http://www.packdocs.com	Clúster 3
56	Reqlut	Chile	www.reqlut.com	Clúster 3
57	SpreadsheetBooster	Brasil	http://spreadsheetbooster.com	Clúster 3
58	Cuboverde	Chile	cuboverde.cl	Clúster 4
59	Eye3	Chile	eye3.cl	Clúster 4
60	Hard Drones	Chile	harddrones.com	Clúster 4
61	Monitor	Chile	www.monitor.cl	Clúster 4
62	UP POINTS	Brasil	uppoints.com	Clúster 4
63	VisualProgress	Chile	http://www.visualprogress.cl	Clúster 4
64	Admetricks	Chile	admetricks.com	Clúster 5
65	Agronometrics	Chile	agronometrics.com	Clúster 5
66	Contenext	Brasil	contenext.com	Clúster 5
67	DeskMetrics	Brasil	http://deskmetrics.com	Clúster 5
68	Eddvantage	Chile	www.eddvantage.com	Clúster 5
69	Eland	Argentina	http://eland.es	Clúster 5
70	Insiders S.A.	Chile	www.insiders.cl	Clúster 5
71	InstaGIS	Chile	instagis.com	Clúster 5
72	JUNAR	Argentina	www.junar.com	Clúster 5
73	Meshh	Argentina	meshh.co	Clúster 5
74	PERAST	Chile	http://www.perast.cl/	Clúster 5
75	RelCase	Chile	relcase.cl	Clúster 5

N°	Nombre	País	Dirección web	Clúster
77	Safer List	Chile	saferlistglobal.com	Clúster 5
78	Sellpad	México	sellpad.com.mx	Clúster 5
79	Semant.io	Chile	semant.io	Clúster 5
80	Simplit Solutions	Chile	http://www.simplit.cl	Clúster 5
81	TopicFlower	Argentina	http://topicflower.com	Clúster 5
82	TWSocial	Chile	twsocial.com	Clúster 5