# Paso 2.  Recopilación de Información

Con el propósito de realizar una correcta recopilación de información, se propone abarcar esta fase por medio de 3 factores principales:

## 2.1 Marco Teórico

Se hace una búsqueda de los términos relacionados con el problema planteado. Esta búsqueda se realiza en fuentes reconocidas y confiables para reconocer que conceptos hacen parte del problema.

* Data mining:

*Data Mining* (minería de datos) es el proceso de extracción de información significativa de grandes bases de datos, información que revela inteligencia del negocio, a través de factores ocultos, tendencias y correlaciones para permitir al usuario realizar predicciones que resuelven problemas del negocio proporcionando una ventaja competitiva.

* Big Data:

Big data (en español, grandes datos o grandes volúmenes de datos) es un término evolutivo que describe cualquier cantidad voluminosa de datos estructurados, semiestructurados y no estructurados [que tienen el potencial de ser extraídos para obtener información](https://searchdatacenter.techtarget.com/es/cronica/Big-data-una-espada-de-doble-filo-en-la-ciberestrategia-de-este-CIO).

* Windows Forms

Las características principales de la programación de Windows Forms y cómo puede usar Windows Forms para compilar smart clients que satisfagan las necesidades actuales de las empresas y usuarios.



* C#

es un lenguaje de programación orientado a objetos desarrollado y estandarizado por Microsoft como parte de su plataforma .NET, que después fue aprobado como un estándar por la ECMA (ECMA-334) e ISO (ISO/IEC 23270). C# es uno de los lenguajes de programación diseñados para la infraestructura de lenguaje común

* Visual Studio

Microsoft Visual Studio es un [entorno de desarrollo integrado](https://es.wikipedia.org/wiki/Entorno_de_desarrollo_integrado) (IDE, por sus siglas en inglés) para sistemas operativos [Windows](https://es.wikipedia.org/wiki/Windows). Soporta múltiples lenguajes de programación, tales como [C++](https://es.wikipedia.org/wiki/Visual_C%2B%2B), [C#](https://es.wikipedia.org/wiki/C_sharp), [Visual Basic .NET](https://es.wikipedia.org/wiki/Visual_Basic_.NET), [F#](https://es.wikipedia.org/wiki/F_Sharp), [Java](https://es.wikipedia.org/wiki/Java_(lenguaje_de_programaci%C3%B3n)), [Python](https://es.wikipedia.org/wiki/Python), [Ruby](https://es.wikipedia.org/wiki/Ruby) y [PHP](https://es.wikipedia.org/wiki/PHP), al igual que entornos de desarrollo web, como [ASP.NET](https://es.wikipedia.org/wiki/ASP.NET) MVC, Django, etc., a lo cual hay que sumarle las nuevas capacidades online bajo Windows Azure en forma del editor Monaco.

* Base de datos

Una base de datos de un SI (Sistema de Información) es la representación integrada de los conjuntos de entidades instancia correspondientes a las diferentes entidades tipo del SI y de sus interrelaciones. Esta representación informática (o conjunto estructurado de datos) debe poder ser utilizada de forma compartida por muchos usuarios de distintos tipos.

En otras palabras, una base de datos es un conjunto estructurado de datos que representa entidades y sus interrelaciones. La representación será única e integrada, a pesar de que debe permitir utilizaciones varias y simultáneas.



* Microsoft.NET

Es un [framework](https://es.wikipedia.org/wiki/Framework) de [Microsoft](https://es.wikipedia.org/wiki/Microsoft) que hace un énfasis en la transparencia de [redes](https://es.wikipedia.org/wiki/Red_de_computadoras), con independencia de plataforma de [hardware](https://es.wikipedia.org/wiki/Hardware) y que permite un rápido desarrollo de [aplicaciones](https://es.wikipedia.org/wiki/Aplicaci%C3%B3n_inform%C3%A1tica).

## 2.2 Entrevista con el cliente

Se realiza una visita por parte del personal de la empresa Allers, en donde se logra recolectar información específica del estado en el que opera la empresa actualmente:

* La empresa Allers cuenta con una base de datos en SQL Server.
* La empresa cuenta con equipos de buena capacidad para el manejo de grandes cantidades de datos.
* La solución a plantear va dirigida a los directivos de Allers para la toma de decisiones.
* La empresa cuenta con un inventario activo de aproximadamente 15000 productos activos.
* Se enfocan en la distribución de insumos hospitalarios, teniendo como objetivo cubrir la mayor parte del territorio nacional.
* Realizan sus ventas basándose en promedios históricos.
* Cuenta con personal que realiza análisis según las estadísticas que arrojan los datos que tiene la empresa

## 2.3 Antecedentes

Se realiza una investigación de situaciones y proyectos los cuales sean similares al problema planteado. Esta investigación se realiza teniendo en cuenta proyectos en donde se abarque una solución por medio de Big Data o Data mining.

* Proyecto (Big data para acabar con los atascos en Nueva York):

La ciudad de Nueva York, como muchas ciudades en el mundo, tienen presente una gran problemática en cuanto al tráfico urbano, muchas de las soluciones que se han tratado de abordar son el uso de algoritmos para sincronizar los semáforos, de tal manera que dé prioridad a vías principales, aún así solo se ha logrado disminuir un poco los atascos. Según la investigadora del MIT en congestión de redes urbanas, Carolina Osorio, plantea una solución a esta situación para un futuro muy cercano, esta solución se trata de utilizar la big data, principalmente extrayendo información de las agendas de las personas, con el propósito de crear algoritmos que permita realizar una organización en las vías, recomendando cuando realizar una actividad utilizando claramente factores de prioridad.

* Proyecto (Big data para salvar animales):

El mundo está sufriendo una [sexta ola de extinción en masa](http://wwf.panda.org/about_our_earth/biodiversity/biodiversity/) en la que grandes porcentajes de las especies del planeta están desapareciendo. Un equipo de científicos del Snow Leopard Trust y de la National Conservation Foundation están [usando los datos](http://www.huffingtonpost.com/2015/01/12/india-wild-tigers-big-data_n_6458386.html) para saber cuándo, dónde y cómo los [cazadores furtivos intentarán cazar los tigres salvajes en la India.](http://dataconomy.com/2015/01/how-big-data-can-help-in-the-fight-to-save-indias-tigers/) La fauna de la India alberga la mayoría de la población de tigres salvajes del planeta. Por esta razón están analizando cantidades increíbles de datos (Big data) para intentar interceptar los cazadores furtivos antes de que puedan llegar a su objetivo.

* Proyecto (Big data para detectar y prevenir ciberataques):

El análisis de big data se está empleando de forma intensiva en la mejora de la seguridad informática, como lo está solucionando el sistema de IA creado en el MIT el cual hasta la fecha a logrado predecir el 85% de los ciberataques.

* Proyecto (Big data para la recomendación de libros - Amazon):

La empresa Amazon realiza un análisis de big data sobre los datos que recolectan por sus ventas de libros, gracias a este análisis logran realizar sugerencias de otros libros  por medio de un algoritmo. Un análisis en su base de datos permite identificar que un cliente que compra un libro, por lo general su comportamiento está basado en comprar más libros de este mismo tipo, razón por la cual el sistema sugiere otros libros similares basándose en los datos recolectados de miles de clientes.

Fuentes:

* Rafael Camps Paré, Luis Alberto Casillas Santillán, Dolors Costal Costa, Marc Gibert Ginestà, Carme Martín Escofet, Oscar Pérez Mora.”Bases de datos”.Fundació per a la Universitat Oberta de Catalunya.
* https://docs.microsoft.com/es-es/dotnet/framework/winforms/windows-forms-overview
* <https://searchdatacenter.techtarget.com/es/definicion/Big-data>

* <https://www.gestiopolis.com/que-es-data-mining/>
* Vodafone, [El futuro es aspirante].(2017,09,17).Big data para acabar con los atascos en Nueva York. Recuperado de https://www.youtube.com/watch?v=Vljm5JRxHKU&t=90s
* Big data para detectar y predecir ciberataques. Recuperado de <https://www.baoss.es/10-ejemplos-usos-reales-big-data/>
* Big data para salvar animales en extinción. Recuperado de <https://bismart.com/es>