

UNIVERSIDAD DE EL SALVADOR  
FACULTAD DE INGENIERIA Y ARQUITECTURA  
ESCUELA DE INGENIERIA DE SISTEMAS INFORMÁTICOS  
CICLO I - 2017



**ANTEPROYECTO:**

PROPUESTA DE FORMACIÓN ACADÉMICA Y PROFESIONAL PARA EL PERFIL  
DE ESPECIALIZACIÓN EN EL ÁREA DE INTELIGENCIA DE NEGOCIOS Y  
ANALÍTICA DE DATOS PARA LA CARRERA DE PREGRADO DE INGENIERÍA DE  
SISTEMAS INFORMÁTICOS DE LA UNIVERSIDAD DE EL SALVADOR

**PRESENTADO POR:**

MÓNICA GUADALUPE FLORES BONILLA  
MARÍA JOSÉ FLORES RAMOS  
MELVIN IVÁN REYES ORTIZ  
JUAN DIEGO ZAVALA MEJÍA

**DOCENTE ASESOR:**

ING. ELMER ARTURO CARBALLO RUIZ

CIUDAD UNIVERSITARIA, 08 DE MAYO DE 2017

## Contenido

INTRODUCCIÓN .....	1
CAPÍTULO I: EL PROBLEMA DE INVESTIGACIÓN .....	4
1.1 Situación problemática .....	4
1.2 Planteamiento del problema .....	5
1.2.1 Variables de estudio.....	5
1.3 Situación Deseada .....	5
1.4 Preguntas de Investigación .....	5
1.5 Objetivos de Investigación.....	6
1.5.1 Objetivo General.....	6
1.5.2 Objetivo Específicos .....	6
1.6 Alcances y Limitaciones .....	7
1.6.1 Limitaciones de la investigación .....	7
1.6.2 Alcances .....	7
1.7 Justificación del estudio.....	8
1.7.1 Importancia .....	10
1.7.2 Resultados Esperados .....	12
CAPÍTULO II: MARCO TEÓRICO.....	13
2.1 Antecedentes del estudio .....	13
2.2 Bases teóricas.....	17
2.2.1 Inteligencia de negocios.....	17
2.2.2 Analítica de Big Data.....	32
2.3 Marco Conceptual .....	42
CAPÍTULO III: METODOLOGÍA.....	45
3.1 Tipo de investigación.....	45
3.2 Metodología para la recolección de datos.....	47
3.2.2 Población y muestra.....	53
3.3 Metodología del proyecto .....	56
3.3.1 Creación de perfil profesional.....	56
3.3.2 Creación de curso académico .....	57
CAPÍTULO IV: ASPECTOS ADMINISTRATIVOS .....	63
4.1 Cronograma de Actividades.....	63
4.2 Presupuesto .....	65
CONCLUSIONES .....	67
RECOMENDACIONES .....	68

REFERENCIAS .....	69
ANEXOS.....	70
Anexo 1: Lluvia de Ideas para el análisis y determinación de la situación problemática .	70
Anexo 2: Análisis de Resultado de Encuesta “Investigación de la Situación Actual de la inteligencia de negocios y analítica de datos” .....	71
Anexo 3: Cálculo de muestra para poblaciones infinitas o desconocidas .....	86

## INTRODUCCIÓN

---

El término inteligencia de negocios fue utilizado por primera vez en 1865 por Richard Miller Devens en su libro “Cyclopaedia of Commercial and Business Anecdotes”, para describir cómo Sir Henry Furnese, un banquero, obtuvo una ventaja sobre sus competidores utilizando y actuando sobre la información que lo rodeaba, lo que mejoró su capacidad para tomar decisiones empresariales.<sup>1</sup>

En los años 1990 y 2000 el mercado de inteligencia de negocios mostraba signos de aceleración, el desarrollo de herramientas de inteligencia de negocios también aumentaba y los objetivos principales de producir informes y visualizar percepciones seguían siendo los mismos.

Se entiende por inteligencia de negocios a la combinación de tecnologías, herramientas y procesos que permiten transformar datos almacenados en información, esta información en conocimiento y este conocimiento dirigido a un plan o una estrategia comercial.<sup>2</sup>

Cuando se aborda el tema de inteligencia de negocios, es imposible dejar de hablar de analítica de datos, el cual se define como la ciencia que examina datos en bruto con el propósito de sacar conclusiones sobre la información.<sup>3</sup>

La analítica de datos comenzó a mediados de los años 50, principalmente con el análisis descriptivo basado en datos estructurados. Hace aproximadamente una década atrás, se empezó a hablar de la segunda versión de la analítica de datos en donde se habla del análisis de datos no estructurados.<sup>4</sup>

Actualmente ya se habla de la tercera versión de la analítica de datos, que es una combinación del análisis tradicional y Big Data.

El término de Big Data se utilizó por primera vez para referirse a volúmenes de datos en aumento, a mediados de los años noventa. En 2001, Doug Laney, entonces analista de la consultora Meta Group Inc., amplió la noción de grandes datos para incluir también aumentos en la variedad de datos generados por las organizaciones y la velocidad a la que se creaban y actualizaban esos datos. Estos tres factores volumen, velocidad y variedad, se conocieron como las 3Vs del Big Data, un concepto que Gartner popularizó después de adquirir Meta Group y contratar a Laney en 2005.

---

<sup>1</sup>Davies N.. (2015). A history of Business Intelligence. marzo,14,2016, de LinkedIn Sitio web: <https://www.linkedin.com/pulse/history-business-intelligence-natalie-davies>

<sup>2</sup> Oracle. (2008). Qué es inteligencia de Negocios. Febrero 26, 2017, de Oracle Sitio web: [http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529\\_esa.pdf](http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529_esa.pdf)

<sup>3</sup> Margaret Rouse. (2012). Análisis de datos. Febrero 20 ,2017, de TechTarget Sitio web: <http://searchdatacenter.techtarget.com/es/definicion/Analisis-de-Datos>

<sup>4</sup> Handfield R.. (2013). A brief story for big data analytics. marzo 15, 2017, de International Institute for Analytics Sitio web: <http://iianalytics.com/research/a-brief-history-of-big-data-analytics>

La analítica de Big Data es el proceso de examinar volúmenes crecientes y variados conjuntos de datos para descubrir patrones, correlaciones, tendencias de mercado y preferencias para ayudar a las organizaciones en la toma de decisiones basadas en información de valor.<sup>5</sup>

La inteligencia de negocios y analítica de datos son áreas que han empezado a abrirse paso en las organizaciones y empresas de El Salvador, como evidencia se tienen las ofertas de plazas en estos rubros. Además, en el ámbito académico se han empezado a impartir cursos, talleres y conferencias en estas áreas en diferentes universidades y otras instituciones.

La Facultad de Ingeniería y Arquitectura como parte de La Universidad de El Salvador, a través de la formación académica superior deberá satisfacer las necesidades por parte de las organizaciones y empresas que demandan profesionales en esta área.

Se ha identificado que en el presente año 2017 no se tiene un perfil académico de pregrado completo en el área de inteligencia de negocios y analítica de datos que satisfaga las necesidades de profesionales que demandan las organizaciones y empresas en El Salvador. Por ende, a través del desarrollo de una investigación aplicada, empleando como metodología la estrategia mixta, se determinarán los conocimientos y las habilidades técnicas que forman parte del perfil académico y profesional en el área de inteligencia de negocios y analítica de Big Data para la creación de un curso de especialización de pregrado para la carrera de Ingeniería de Sistemas Informáticos de la Facultad de Ingeniería y Arquitectura de la Universidad de El Salvador.

Para el proyecto de investigación propuesto se presenta, como punto de partida, el anteproyecto, para dar el panorama general que permitirá describir lo que se llevará a cabo durante todo el desarrollo de la investigación. El anteproyecto consta de tres capítulos, los cuales se describen a continuación.

En el planteamiento del problema se describen varios puntos críticos para el desarrollo de la investigación como es la formulación del problema, las preguntas de investigación, así como el establecimiento de objetivos. Además, en este capítulo se justifica la importancia que tendrá llevar a cabo la investigación y los alcances que se pretenden lograr.

El siguiente capítulo es el marco teórico en el cual se presentan las bases teóricas que sustentarán el desarrollo de la investigación. Contiene además los antecedentes, que plasma la historia del crecimiento de la inteligencia de negocios y el análisis de los datos, desde un contexto internacional y en El Salvador. Adicionalmente, la base teórica presentada, se enfoca en los temas de estudio a desarrollar en el proyecto. Finalmente, y no menos importante, se define el marco conceptual, que consta de un conjunto de conceptos cuyo entendimiento son claves para el desarrollo del proyecto.

---

<sup>5</sup> Rouse, M., (2017). Big Data Analytics. abril 10,2017, de Tech Target Sitio web: <http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics>

En la metodología se plantea la forma en que se llevará a cabo la investigación, las técnicas de recolección de datos, la planificación de estas, además la forma en cómo se proseguirá para construir el perfil y el curso planteado. Se incluye en esta parte el cronograma de todo el ciclo de vida del proyecto, así como la estimación de costos a los que se incurre con este proyecto.

Como último capítulo se establecen los aspectos administrativos que incluyen el cronograma de actividades que se realizarán durante todo el desarrollo de la investigación.

## **CAPÍTULO I: EL PROBLEMA DE INVESTIGACIÓN**

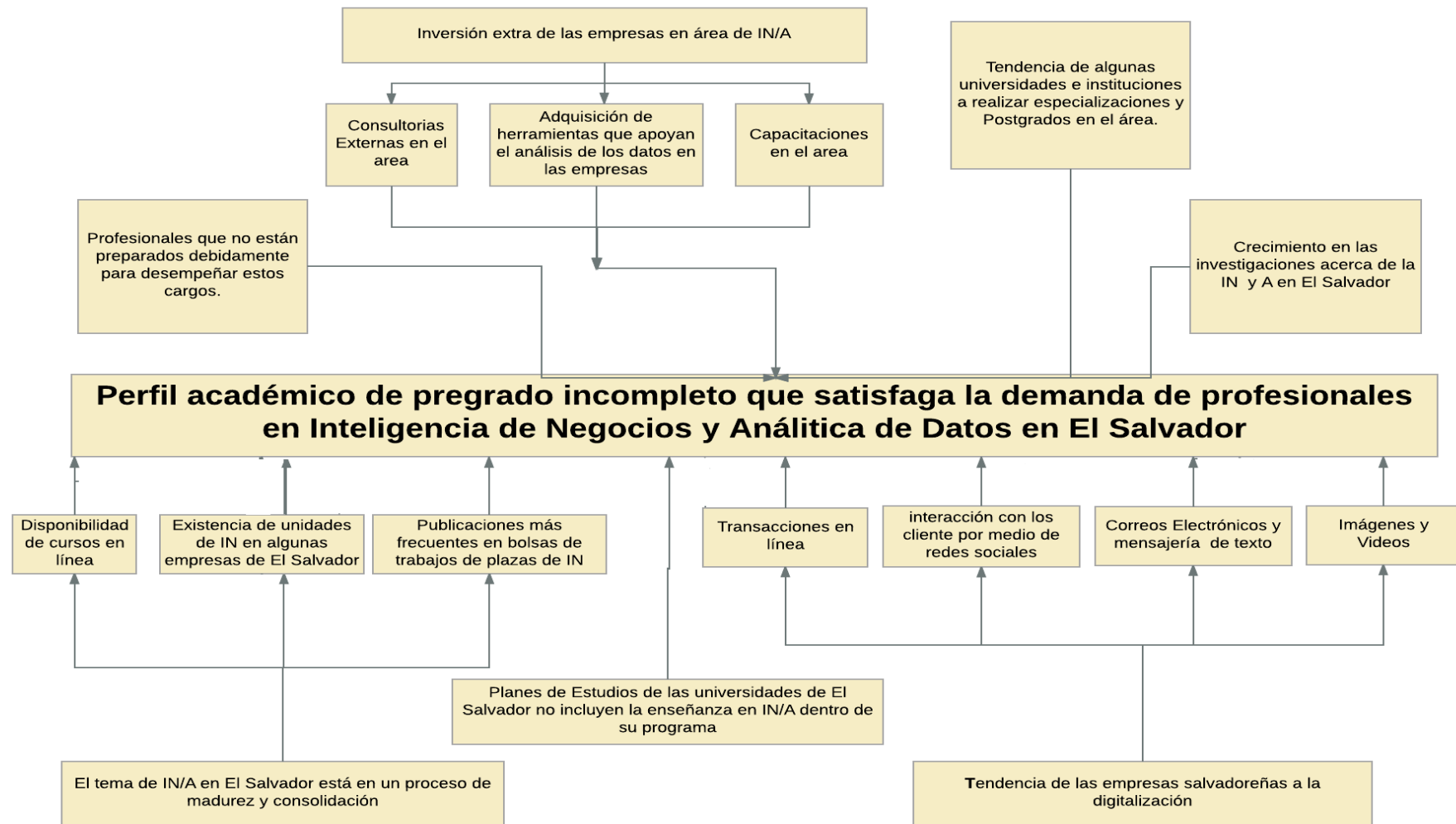
---

### **1.1 Situación problemática**

La inteligencia de negocios y analítica de datos se ha posicionado en las empresas de El Salvador como una necesidad emergente ya que aproximadamente desde hace más de 10 años se inició la creación de unidades de inteligencia de negocios en las empresas salvadoreñas y por ende la oferta de plazas. En el año 2013 se empezaron a impartir cursos de especialización y posgrados en el área por parte de instituciones educativas superiores y organizaciones. Estos hechos hacen que los estudiantes de sistemas informáticos interesados en el tema, requieran como parte de su aprendizaje, recibir los conocimientos, las habilidades técnicas y destrezas en el manejo de herramientas de inteligencia de negocios y analítica de datos, lo cual al nivel de pregrado no se satisface en la mayoría de universidades del país.

La Escuela de Ingeniería de Sistemas Informáticos de la Facultad de Ingeniería y Arquitectura de La Universidad de El Salvador, es conocedora de dicha situación y le inquieta generar soluciones a las necesidades actuales y futuras para satisfacer las nuevas tendencias en formación profesional.

Para identificar la situación problemática de la investigación se utilizó la técnica denominada lluvia de ideas (Ver anexo 1), en donde los investigadores colocaron sus criterios en base a la experiencia en el tema, a la búsqueda bibliográfica y a los resultados obtenidos de un estudio de sondeo realizado, titulado “Investigación sobre la situación actual de la inteligencia de negocios y analítica de datos” cuyos resultados también se anexan (Ver anexo 2). Posteriormente se analizaron y consolidaron las ideas en un “árbol del problema” con el cual se describe el problema de estudio, así como sus causas y efectos. Dicha técnica se asemeja en su representación a un árbol; en la parte de en medio, simulando el tronco, se encuentra el problema; en la parte inferior, simulando las raíces, se han colocado las causas y en la parte superior simulando las hojas se han colocado los efectos.





## 1.2 Planteamiento del problema

Tomando como base la situación problemática, surge la siguiente interrogante: **¿Cuáles son los conocimientos, las habilidades técnicas y las herramientas que deben formar parte del perfil académico y profesional en el área de inteligencia de negocios y analítica de Big Data para poder formar en esta área a los estudiantes de pregrado de la carrera de Ingeniería de Sistemas Informáticos de la Universidad de El Salvador?**

### 1.2.1 Variables de estudio

- La necesidad de profesionales especializados en inteligencia de negocios y analítica de Big Data en El Salvador.
- La oferta de cursos de especialización y posgrados de parte de instituciones educativas superiores y otras organizaciones.
- Los conocimientos que forman parte del perfil de especialización de inteligencia de negocios y analítica de Big Data.
- Las habilidades técnicas y herramientas de BI y analítica de Big Data fundamentales para un perfil académico y profesional en el área.

## 1.3 Situación Deseada

Que esté definido un perfil de inteligencia de negocios en donde se definan los conocimientos que se requieren para desarrollarse en esta área.

Que esté definido un perfil de analítica de datos en donde se definan los conocimientos que se requieren para desarrollarse en esta área.

La Universidad de El Salvador en su afán de preparar profesionales que satisfagan las necesidades del campo laboral tendrá actualizada su currícula de acuerdo a las tendencias de la carrera de Ingeniería de Sistemas Informáticos.

## 1.4 Preguntas de Investigación

- ¿Qué conocimientos debe tener un profesional que se dedique a la inteligencia de negocios y analítica de Big Data?
- ¿Cuáles son los requerimientos técnicos más demandados en el área laboral de inteligencia de negocios y analítica de Big Data?
- ¿Cuáles son las técnicas y herramientas que ayudan a los futuros profesionales a aplicar inteligencia de negocios y analítica de Big Data?
- ¿De qué manera se formará al profesional para que se desempeñe en esas áreas?

## **1.5 Objetivos de Investigación**

### **1.5.1 Objetivo General**

- Crear un curso de especialización de pregrado en el área de inteligencia de negocios y analítica de Big Data para la carrera de Ingeniería de Sistemas Informáticos de la Facultad de Ingeniería y Arquitectura de la Universidad de El Salvador a través de la determinación de los conocimientos, las habilidades técnicas y herramientas que forman parte del perfil académico y profesional del área

### **1.5.2 Objetivo Específicos**

- Identificar el perfil demandado de los profesionales en el área de estudio de inteligencia de negocios y Big Data mediante la comparación de los requerimientos técnicos de las diferentes plazas publicadas en las bolsas de trabajo y periódicos nacionales y otras fuentes de información para determinar los contenidos que conformarán el curso de especialización
- Desarrollar el perfil académico y profesional de inteligencia de negocios y analítica de Big Data mediante las fuentes primarias y secundarias para consolidar los conocimientos que deben adquirir los profesionales en su formación
- Identificar las técnicas, estándares y herramientas que se utilizan en el proceso de hacer inteligencia de negocios y analítica de Big Data mediante la recolección de datos por medio de entrevistas, encuestas, investigación bibliográfica con el propósito de formar una base de conocimientos sólida en el área.
- Diseñar un curso académico y profesional a partir del perfil de los conocimientos obtenidos del área de inteligencia de negocios y analítica de Big Data, para formar profesionales competitivos en el campo laboral.

## **1.6 Alcances y Limitaciones**

### **1.6.1 Limitaciones de la investigación**

- El estudio se limitará solo a las empresas e instituciones del área de San Salvador.
- Madurez en el desarrollo de las áreas de inteligencia de negocio y analítica de datos a nivel internacional con respecto a nivel nacional.
- La selección de la(s) herramienta(s) para el desarrollo del curso académico y profesional estará condicionado por el acceso gratuito o bien por licencias estudiantiles.
- El tema de machine learning no se considera como parte del desarrollo del proyecto de investigación, dado que su complejidad de contenido sobrepasa el alcance de una solución para pregrado.

### **1.6.2 Alcances**

- Identificación de las necesidades del país sobre el perfil de formación en el área de inteligencia de negocios y analítica de Big Data.
- Investigación sobre el cuerpo de conocimiento necesario para un perfil de un especialista en inteligencia de negocios y analítica de Big Data de acuerdo a las necesidades del país.
- Propuesta de áreas de conocimientos, contenidos, habilidades y valores que el profesional en el área de inteligencia de negocios y analítica de Big Data que deberá poseer en su formación profesional.
- Creación de los contenidos para el área de inteligencia de negocios y Big Data de datos en la formación del profesional.
- Propuesta de técnicas, estándares y herramientas para el uso y explotación de los datos para hacer inteligencia de negocios y analítica de Big Data.
- Creación de un curso de pregrado de formación académica especializado en inteligencia de negocios y analítica de Big Data.

## 1.7 Justificación del estudio

La Universidad de El Salvador, crítica de la realidad y con capacidad de proponer soluciones a los problemas nacionales se visualiza ser un transformador y líder de la Educación Superior. Dado que la situación problemática identificada tiene sus alcances en el área de la informática le corresponde a La Escuela de Ingeniería de Sistemas Informáticos, de la Facultad de Ingeniería y Arquitectura investigar, innovar y desarrollar la educación de tal manera que los profesionales que forja tengan las capacidades que satisfagan las necesidades presentes del sector laboral salvadoreño.

A partir del año 2000 la inteligencia de negocios y la analítica de datos vinieron a impactar a las empresas del país, reconociéndose la necesidad de profesionales especializados en el área. Desde el año 2013 hasta la fecha en que se está realizando este estudio, existen universidades e instituciones que ofrecen especializaciones y postgrados en la materia:

- La Universidad Tecnológica ofrece la pre especialización de graduación “Desarrollo de soluciones en inteligencia de negocios para entornos empresariales” para sus estudiantes de pregrado <sup>6</sup>
- La Universidad Francisco Gavidia posee una especialización de graduación en Administración y Desarrollo de Base de datos e Inteligencia de Negocios<sup>7</sup>.
- Fepade ofreció un postgrado internacional en Business Intelligence en el año 2015<sup>8</sup>.

El contenido que se ofrece en los cursos de especialización de graduación mencionados no se ha actualizado desde su creación considerando que esta área evoluciona aceleradamente. Además, están pensados para llevarse a cabo en un periodo de seis a ocho meses lo que limita la profundidad del aprendizaje y el desarrollo de las competencias que se demandan. Estos dos hechos hacen que surja la oportunidad de crear una propuesta de formación académica y profesional para el perfil de especialización en el área de inteligencia de negocios y analítica de datos para la carrera de pregrado de esta alma mater.

Se realizó un sondeo preliminar para conocer la situación actual de la inteligencia de negocios y analítica de datos visto desde la perspectiva de los profesionales que consumen la información, es decir profesionales que ejercen cargos de jefaturas, gerencias y altos mandos en instituciones públicas y empresas privadas.

---

<sup>6</sup> Universidad Tecnológica de El Salvador. (2013). Programación Preespecialidad. Marzo 14, 2017, de Universidad Tecnológica de El Salvador Sitio web: <http://graduadoutec.edu.sv/banner/Programacion-Preespecialidad-Prueba01-2013.pdf>

<sup>7</sup> Universidad Francisco Gavidia. (2015). Curso de Especialización Administración y Desarrollo de Base de datos e Inteligencia de Negocios. Marzo 14, 2017, de Universidad Francisco Gavidia Sitio web: <http://registro.ufg.edu.sv/InformacionEgresados/Especializacion.aspx>

<sup>8</sup> Isdae Fepade. (2015). Postgrado Internacional en Business Intelligence - Septiembre 2015. Marzo 14, 2017, de Fepade Sitio web: <https://www.iseade.edu.sv/index.php/postgrados-especializados/30-postgrados-especializados/historico-de-postgrados/259-postgrado-internacional-en-business-intelligence-septiembre-2015>

De los resultados obtenidos se analiza lo siguiente:

1. Se evaluó el conocimiento de las personas que consideraron estar familiarizadas con el tema de estudio, la mayoría identificó correctamente el concepto de inteligencia de negocios y el de analítica de datos, sin embargo, para el segundo caso, sólo el 4.76% separa a las personas que si saben de las que no. Una proporción tan pequeña demanda interés y esfuerzos por especializar a los profesionales en las áreas de su competencia, para garantizar que sean expertos en lo que se desenvuelven y no se conviertan en ingenieros genéricos de conocimiento que tratan saber de todo.
2. Todas las personas encuestadas que desempeñan sus labores en un nivel estratégico consideran de manera unánime que los datos que se generan en sus empresas deben analizarse con un mayor énfasis; lo que se traduce en un reconocimiento de la necesidad de aplicar inteligencia de negocios y analítica de datos en los procesos de sus negocios.
3. El 52% de las personas que participaron en la encuesta conocen la existencia de la inteligencia de negocios porque en su empresa existe un área como tal y un 33% sabe que por lo menos en una empresa del país existe esta área. Este 85% en conjunto representa el posicionamiento de la inteligencia de negocios en las empresas salvadoreñas y por ende el crecimiento de la demanda de profesionales que las trabajen.
4. Existe una disconformidad de parte de los directivos con respecto a la información que actualmente poseen para tomar decisiones, reconociendo que con la inteligencia de negocios y analítica de datos la toma de decisiones es más oportuna y precisa. La realización de proyecciones, el descubrimiento de nuevos datos y el conocimiento de comportamientos futuros de los clientes fueron los beneficios de la inteligencia de negocios y analítica de datos que los directivos identificaron con más frecuencia.

Por tanto, existe un pronunciamiento de inteligencia de negocios y analítica de datos en las empresas salvadoreñas y por ende de profesionales dedicados al área, es preciso entonces, que la Universidad de El Salvador cumpliendo a su misión educativa, se enfoque en darle solución a las necesidades actuales desde el área de su competencia: la educación.

### 1.7.1 Importancia

La creación de un curso académico y profesional de especialización de pregrado en la carrera de Ingeniería de Sistemas Informáticos es la propuesta de solución planteada; dicho curso contendrá el material didáctico necesario para que los estudiantes que lo cursen adquieran los conocimientos y habilidades técnicas solicitados en el perfil de inteligencia de negocios y analítica de Big Data. Esto impactaría a tres diferentes sectores: los estudiantes y futuros ingenieros de sistemas informáticos, la Universidad de El Salvador y la Sociedad Salvadoreña.

#### 1. Profesionales Especializados en inteligencia de negocios y analítica de datos

Dado que la inteligencia de negocios y la analítica de datos tiende al crecimiento a nivel internacional, y como se ha demostrado, también a nivel nacional, se espera que la necesidad de profesionales en el área crezca y por ende que haya profesionales que la satisfagan. El Foro Económico Mundial (The World Economic Forum) en colaboración de Business Insider publicó en mayo del 2016 “Los 8 trabajos que las compañías demandarán para el 2020”.<sup>9</sup> En éste, los analistas de datos serán uno de las profesiones más demandadas dado a la necesidad de darle sentido a la mezcla de datos no estructurados y estructurados que está generando la tecnología. Por tanto, esto beneficia a los futuros profesionales que decidan especializarse en el área, al tener mayor oportunidad de encontrar trabajo y mayores posibilidades de selectividad gracias a los conocimientos técnicos específicos y las habilidades que se pretende inculcar en los estudiantes.

No está demás decir, que la aplicación de los conocimientos adquiridos no tendrá frontera física para los profesionales especializados. El Índice Global de Habilidades 2015-2016, elaborado por Hays en colaboración con Oxford Economics, a petición de Forbes México, expresa la demanda de profesionales en las industrias de alta especialización: tecnológica, telecomunicaciones, energética, farmacéutica y manufactura. Gerente en inteligencia de negocios, en el sector TI es una de las diez tendencias laborales del 2016 publicadas por Forbes México<sup>10</sup>. Según éste, la tendencia indica que cada vez más compañías alrededor del mundo necesitan profesionales con un perfil de Business Intelligence, que incorporen habilidades de comunicación y gestión de proyectos.

#### 2. Universidad de El Salvador

La Universidad de El Salvador ha innovado recientemente en la manera de educar, inaugurando en el año 2016 la nueva modalidad de educación en línea a distancia, con la cual se pretende la inclusión de los jóvenes salvadoreños a la educación superior. Los desafíos por seguir transformado la educación superior y

---

<sup>9</sup>Thompson, C.. (2016). 8 jobs every company will be hiring for by 2020. abril 09,2017, de World Economic Forum Sitio web: <https://www.weforum.org/agenda/2016/01/8-jobs-every-company-will-be-hiring-for-by-2020/>

<sup>10</sup> Forbes Staff. (2016). Las 10 profesiones más demandadas en 2016. Febrero,14,2016, de Forbes México Sitio web: <http://www.forbes.com.mx/las-10-profesiones-mas-demandadas-en-2016/#gs.fzevR8o>

mantenerse en un panorama protagónico de entre todas las universidades del país se encuentran latentes. La creación de una especialización de pregrado en la carrera de Ingeniería de Sistemas Informáticos, contribuirá a la ya empezada labor de la universidad y seguramente servirá de ejemplo para otras escuelas de las diferentes carreras de la universidad para que se preocupen por investigar la situación actual del campo de su competencia y se contraste con los conocimientos que actualmente se están impartiendo en sus programas de educación.

Los resultados de una investigación con fines educativos y propositivos pueden incluir el fortalecimiento del sistema educativo universitario actual (incluyendo el contenido y la pedagogía), la promoción de políticas y estrategias de estado en pro de la educación superior que involucren actores gubernamentales y/o no gubernamentales, apertura de nuevas alianzas educativas que se reflejen en beneficios para los estudiantes como becas, trabajos en el exterior o intercambios por mencionar algunos.

Tener en mente las funciones básicas de la Universidad de El Salvador: la docencia, la investigación y la proyección social, contribuirá con la innovación educativa y excelencia académica, funciones que se pretenden se alcanzarán con el tema propuesto.

### 3. Sociedad Salvadoreña

En un contexto del crecimiento de la economía, y según el PNUD en El Salvador, la generación de trabajo decente sigue siendo un gran desafío. De cada cien personas que forman parte de la población económicamente activa (PEA), 7 están desempleadas y 44 subempleadas (con ingresos inferiores al salario mínimo vigente en el sector económico en el que laboran). Únicamente la quinta parte de la fuerza laboral cuenta con trabajo decente y solamente el 28% de la PEA es cotizante para recibir una pensión.<sup>11</sup>

En septiembre de 2015, los Estados Miembros de la ONU aprobaron la Agenda 2030 para el Desarrollo Sostenible, que incluye un conjunto de 17 Objetivos de Desarrollo Sostenible (ODS)<sup>12</sup> que deberán ser alcanzados por sus países miembros (El Salvador es uno de ellos). Dos de ellos están enfocados a la educación de calidad y al trabajo decente y crecimiento económico:

“...Para 2030, aumentar sustancialmente el número de jóvenes y adultos que tienen las competencias necesarias, en particular técnicas y profesionales, para acceder al empleo, el trabajo decente y el emprendimiento”...

“...Lograr niveles más elevados de productividad económica mediante la diversificación, la modernización tecnológica y la innovación, entre otras cosas

---

<sup>11</sup> PNUD. (2017). El Salvador en breve. abril 09,2017, de Programas de las Naciones Unidas para el Desarrollo Sitio web: [http://www.sv.undp.org/content/el\\_salvador/es/home/countryinfo.html](http://www.sv.undp.org/content/el_salvador/es/home/countryinfo.html)

<sup>12</sup> PNUD. (2017). Agenda de desarrollo post 2015. abril 09,2017, de Programas de las Naciones Unidas para el Desarrollo Sitio web; [http://www.sv.undp.org/content/el\\_salvador/es/home/post-2015/sdg-overview.html](http://www.sv.undp.org/content/el_salvador/es/home/post-2015/sdg-overview.html)

centrando la atención en sectores de mayor valor añadido y uso intensivo de mano de obra”...

Dicho esto, se espera que al tener profesionales con competencias que satisfagan las demandas laborales actuales y futuras, se está contribuyendo al desarrollo económico del país, al reducir el empleo informal, mover la actividad económica de otros sectores y aumentar el ingreso salarial en las familias salvadoreñas.

### **1.7.2 Resultados Esperados**

- Perfil Académico de Pregrado que satisfaga la demanda profesional del Área de Inteligencia de Negocios.
- Perfil Académico de Pregrado que satisfaga la demanda profesional del Área de Analítica de Datos.
- Curso Académico de Inteligencia de Negocios y Analítica de Datos a partir de los perfiles investigados.



## CAPÍTULO II: MARCO TEÓRICO

---

### 2.1 Antecedentes del estudio

La inteligencia de negocios y analítica de datos y el campo intrínseco a estos dos: Big Data, han ido ganando importancia y protagonismo en el área de los negocios debido a las ventajas que se obtienen de los datos y del análisis de los mismos, lo que ha hecho que las personas generen especial interés en la inteligencia de negocios y analítica de datos, de ahora en adelante IN&A.

IN&A es generalmente entendida como las técnicas, tecnologías, sistemas, prácticas, metodologías y aplicaciones que analizan datos críticos de las empresas en función del mejor entendimiento del negocio y del mercado y la toma de decisiones en tiempo real.<sup>13</sup>

Las tecnologías y aplicaciones de IN&A que adoptaron las empresas entre la década de los 90 y el año 2000 se conoce como IN&A 1.0. En sus inicios, los datos de las industrias eran en su mayoría estructurados, los cual se obtenían de sistemas heredados y se almacenan en bases de datos relacionales (RDBMS).

Se considera que el data management y data warehousing son la base de esta primera versión de IN&A:

- El diseño de data marts y las herramientas de extracción, transformación y carga (ETL) son esenciales para convertir e integrar datos,
- Consultas a la base de datos, el procesamiento analítico en línea (OLAP) y las herramientas de reportería basados en gráficos simples pero intuitivos utilizados para la exploración de datos,
- La gestión del rendimiento empresarial (Business Performance Management o BPM) a través del uso de dashboards y scorecards para ayudar a visualizar el desempeño de métricas.

Adicional a estas tres funciones de la IN&A 1.0 se suman los análisis estadísticos desarrollados en los años 70 y las técnicas de minería de datos desarrolladas en los 80, a saber: análisis de asociación, segmentación de datos y clustering, análisis de clasificación y regresión, modelos predictivos y detección de anomalías. La mayoría de estos técnicas de procesamiento y análisis de datos han sido incorporados a las herramientas de inteligencia de negocios incluyendo Microsoft, IBM, Oracle y SAP (Sallam et al. 2011).

A partir del año 2000, el internet y la web empezaron a ofrecer una manera diferente de recolección datos y búsquedas analíticas:

---

<sup>13</sup> Chen, H., Chiang, R. & Storey, V.. (2012). Business Intelligence and Analytics: From Big Data to Big Data. En Special Issue: Business Intelligence Research (1188). Estados Unidos: Special Issue Associate Editors.

- Los motores de búsqueda como Google y Yahoo, así como la tendencia del e-commerce como Amazon e eBay, permitieron a las empresas presentar sus negocios en línea e interactuar con sus clientes directamente.
- Los datos que se recogen a través de las cookies y logs de los servidores se convirtieron en una mina de oro para el entendimiento de las necesidades de los consumidores.

La inteligencia en la web, el análisis web y todo el contenido de los usuarios generados a través de la web 2.0 inauguró la nueva era de la IN&A 2.0, centrada en el contenido web desestructurado. Las fuentes de datos de la web 2.0, por mencionar algunos son: foros, grupos en línea, blogs, redes sociales, sitios multimedia y juegos sociales.

Con el análisis web se logra determinar aspectos como el diseño de sitios web, la optimización de ubicar un producto, analizar las transacciones de un consumidor, analizar la estructura del mercado y hacer recomendaciones de productos.

En el 2011 The Economist<sup>14</sup> publicó un artículo donde se reportaba que los teléfonos móviles y las tabletas (más de 480 unidades) sobre pasaban el número de laptops y PCs (cerca de 380 unidades). El estudio comparó la tendencia conocida y la estimada del crecimiento de los dispositivos móviles desde el año 2005-2013, proyectando que para el 2020 la cantidad de dispositivos móviles conectados alcanzarían los 10 billones.

Los dispositivos móviles como iPads, iPhones, teléfonos inteligentes además de otros dispositivos con sensores y con conexión a internet, dispositivos con RFID y códigos de barra, son los que dan inicio la versión 3.0 de IN&A y es aquí el punto de inflexión en donde el IN&A y Big Data se encuentran.

El término de Big Data se utilizó por primera vez para referirse a volúmenes de datos en aumento a mediados de los años noventa. En 2001, Doug Laney, entonces analista de la consultora Meta Group Inc., amplió la noción de grandes datos para incluir también aumentos en la variedad de datos generados por las organizaciones y la velocidad a la que se creaban y actualizaban esos datos. Estos tres factores -volumen, velocidad y variedad- se conocieron como las 3Vs del Big Data, un concepto que Gartner popularizó después de adquirir Meta Group y contratar a Laney en 2005.

La analítica de Big Data es el proceso de examinar volúmenes crecientes y variados conjuntos de datos para descubrir patrones, correlaciones, tendencias de mercado y preferencias para ayudar a las organizaciones en la toma de decisiones basadas en información de valor.<sup>15</sup>

El análisis de Big Data incluye datos estructurados, además de otras formas de datos que a menudo se dejan sin explotar por los programas convencionales de IN&A. Esto

---

<sup>14</sup> Special Report on Personal Technology. (2011). Beyond PC. abril 10,2017, de The Economist Sitio web: <http://www.economist.com/node/21531109>

<sup>15</sup> Rouse, M., (2017). Big Data Analytics. abril 10,2017, de Tech Target Sitio web: <http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics>

incluye una mezcla de datos semi estructurados y no estructurados, por ejemplo: datos de clics de Internet, registros de servidores web, contenido de redes sociales, texto de correos electrónicos de clientes y respuestas de encuestas, registros de llamadas de teléfonos móviles y datos de máquinas capturados por sensores conectados (Internet de las cosas).

Los tipos de datos no estructurados y semi estructurados no encajan bien en los data warehouse tradicionales que se basan en bases de datos relacionales orientadas a conjuntos de datos estructurados. Además, estos pueden no ser capaces de manejar las demandas de procesamiento que plantean los conjuntos de Big Data que necesitan actualizarse con frecuencia, o incluso continuamente. Como resultado, muchas organizaciones que recopilan, procesan y analizan Big Data recurren a Hadoop y sus herramientas complementarias, como YARN, MapReduce, Spark, HBase, Hive, Kafka y Pig, así como bases de datos NoSQL.

## **Inteligencia de negocios y analítica de Big Data en El Salvador.**

### **Educación**

- El plan de estudios 2009 de la carrera de Ingeniería en Ciencias de la Computación la Universidad Don Bosco contempla la materia “Inteligencia de negocios” como una opción electiva técnica.<sup>16</sup>
- En el 2010, estudiantes de la Universidad de El Salvador realizaron el primer trabajo de graduación sobre inteligencia de negocios en la Escuela de Ingeniería de Sistemas Informáticos denominado “Desarrollo de un Data Warehouse para el proceso de Denuncias de la Defensoría del Consumidor”.
- En el 2012, estudiantes de la Universidad Matías Delgado llevaron a cabo el seminario de trabajo de investigación “Uso de la Inteligencia de Negocios en las PYMES en El Salvador”
- La Universidad Tecnológica ofrece desde el año 2013 la pre especialización de graduación “Desarrollo de soluciones en inteligencia de negocios para entornos empresariales” para sus estudiantes de pregrado <sup>17</sup>
- El 26 de julio de 2014 la Escuela Europea de Negocios impartió un seminario gratuito en San Salvador abordando el tema “Las nuevas tecnologías en los modelos de negocio (BIG DATA)” <sup>18</sup>
- Fepade ofreció un postgrado internacional en Business Intelligence en el año 2015<sup>19</sup>.

---

<sup>16</sup> Universidad Don Bosco. (2014). Ingeniería en Ciencias de la Computación Plan 2009. abril 11,2017, de Universidad Don Bosco Sitio web: [http://www.udb.edu.sv/udb/archivo/pensum/ing\\_compu.pdf](http://www.udb.edu.sv/udb/archivo/pensum/ing_compu.pdf)

<sup>17</sup> Universidad Tecnológica de El Salvador. (2013). Programación Preespecialidad. Marzo 14, 2017, de Universidad Tecnológica de El Salvador Sitio web: <http://graduadoutec.edu.sv/banner/Programacion-Preespecialidad-Prueba01-2013.pdf>

<sup>18</sup> Escuela Europea de Negocios. (2014). Las nuevas tecnologías en los modelos de negocio (BIG DATA). abril 11,2017, de Escuela Europea de Negocios Sitio web: <http://www.escuelaeuropea.eu/SV/san-salvador-1/las-nuevas-tecnologias-en-los-modelos-de-negocio-big-data/evento/es>

- La Universidad Francisco Gavidia posee desde el 2015 una especialización de graduación en Administración y Desarrollo de Base de datos e Inteligencia de Negocios.<sup>20</sup>
- En el 2016, estudiantes de la Universidad de El Salvador llevaron a cabo el trabajo de graduación “Business Intelligence como soporte de las decisiones estratégicas, tácticas y operacionales de las organizaciones”
- En el 2016, la Universidad Tecnológica publicó en su revista Entorno los resultados de su investigación “Big Data, análisis de datos en la nube” con el objetivo de conocer sobre el auge de Big Data, y cómo se podrían utilizar algunas herramientas para el procesamiento, análisis y visualización de los datos <sup>21</sup>
- El 29 de marzo de 2017 la Universidad Matías Delgado abordó el tema de Big Data en la actualidad y los Millennials en el área financiera, como parte de su evento Global Money Week 2017<sup>22</sup>
- La Universidad Don Bosco ha ofrecido a estudiantes y profesionales, por medio de su Centro de Tecnología de la información y las telecomunicaciones CTIC las siguientes conferencias:
  - Conferencia: “Analítica de Big Data”, 30 de agosto de 2016.<sup>23</sup>
  - Taller Internacional de Big Data: " Procesamiento de Big Data a gran escala usando Spark", 22 y 23 de septiembre de 2016 <sup>24</sup>
  - Conferencia Internacional Big Data, 14 de marzo de 2017 <sup>25</sup>
  - Taller Internacional: Analítica y Gestión de Big Data, 16 y 17 de marzo de 2017<sup>26</sup>

---

<sup>19</sup> Isdae Fepade. (2015). Postgrado Internacional en Business Intelligence - Septiembre 2015. Marzo 14, 2017, de Fepade Sitio web: <https://www.iseade.edu.sv/index.php/postgrados-especializados/30-postgrados-especializados/historico-de-postgrados/259-postgrado-internacional-en-business-intelligence-septiembre-2015>

<sup>20</sup> Universidad Francisco Gavidia. (2015). Curso de Especialización Administración y Desarrollo de Base de datos e Inteligencia de Negocios. Marzo 14, 2017, de Universidad Francisco Gavidia Sitio web: <http://registro.ufg.edu.sv/InformacionEgresados/Especializacion.aspx>

<sup>21</sup> Rosa, V. & Rivera, J.. (2016, junio 01). Big Data, análisis de datos en la nube. Revista entorno, Universidad Tecnológica de El Salvador, 61, 24.

<sup>22</sup> Reyes, G., (2017). Segundo día del Global Money Week 2017. abril 11, 2017, de Universidad Dr. José Matías Delgado Sitio web: <http://www.ujmd.edu.sv/contenidoslideshow/29-noticias/1354-segundo-dia-del-global-money-week-2017>

<sup>23</sup> CTIC-UDB. (2016, agosto, 18). La Universidad Don Bosco, a través del CTIC (Centro de Tecnología de la Información y las Comunicaciones) le invitan a la Conferencia: “Analítica de Big Data”. Recuperado de: <https://www.facebook.com/cticudb/photos/a.104339129684803.4866.104214889697227/1004593346326039/?type=3&theater>

<sup>24</sup> CTIC-UDB. (2016, septiembre, 1). La Universidad Don Bosco a través del CTIC le invita a su 3er Taller Internacional de Big Data: " Procesamiento de Big Data a gran escala usando Spark" impartido por el especialista Magister Willis Polanco de República Dominicana. Recuperado de:

<https://www.facebook.com/cticudb/photos/a.104339129684803.4866.104214889697227/1016800408438666/?type=3&theater>

<sup>25</sup> CTIC-UDB. (2017b, marzo, 3). Quieres saber más acerca de BIG DATA? Te invitamos a participar de nuestra Conferencia Internacional, entrada gratis, entérate de toda la información necesaria, danos like y comparte esta publicación. Recuperado de: <https://www.facebook.com/cticudb/photos/a.104339129684803.4866.104214889697227/1188136064638432/?type=3&theater>

<sup>26</sup> CTIC-UDB. (2017a, marzo, 3). CTIC le invita a participar en el Taller Internacional: “Analítica y Gestión de Big Data” impartido por el especialista Magister Willis Polanco de República Dominicana. Recuperado de: <https://www.facebook.com/cticudb/photos/a.104339129684803.4866.104214889697227/1188022041316501/?type=3&theater>

- La Escuela Superior de Economía y Negocios ESEN ofrecerá a partir de mayo 2017 culminando en el año 2018 los siguientes programas:
  - Postgrado en Analítica para Ciencia de Datos
  - Certificado superior en Minería de Texto y Analítica de Datos
  - Postgrado en Estadística para Ciencias Sociales

## **2.2 Bases teóricas**

### **2.2.1 Inteligencia de negocios**

#### **Definición**

Inteligencia de negocios puede ser definida como un conjunto de modelos matemáticos y metodologías de análisis que explotan los datos disponibles para generar información y el conocimiento útil para procesos de toma de decisiones complejos.<sup>27</sup>

Proceso de tomar grandes cantidades de datos, analizarlos y presentar un conjunto de reportes de alto nivel que condensan la esencia de esos datos en las bases de las actividades de los negocios, permitiendo a la gerencia tomar decisiones fundamentales de negocio diarias.<sup>28</sup>

Conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de información que permite tomar mejores decisiones a los usuarios de una organización.<sup>29</sup>

Se observan elementos comunes en cada una de las definiciones de inteligencia de negocios entre ellos están: el tratamiento o análisis de los datos para obtener información por medio de metodologías, el proceso de inteligencia de negocios da soporte a la toma de decisiones dentro del área que se aplica

#### **Áreas o categorías en las que se aplica Inteligencia de negocios**

1. AQL - Lógica de consulta asociativa
2. Scorecarding
3. Gestión del rendimiento empresarial y Medición del desempeño
4. Planificación empresarial
5. Reingeniería de procesos de negocio
6. Análisis competitivo

---

<sup>27</sup> Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. Reino Unido: Wiley.

<sup>28</sup> Rajan, J.. (2005). *Business Intelligence: Concepts, Components, Techniques And Benefits*. Journal of Theoretical and Applied Information Technology, Vol. 9, pp 60-70.

<sup>29</sup> Curto, J. & Conesa, J. . (2010). Capítulo I: Introducción al Business Intelligence. En *Introducción al Business Intelligence*(p. 18). Barcelona: UOC.

7. Gestión de relaciones con los clientes (CRM) y Marketing
8. Sistemas de Apoyo a la Decisión (DSS) y Predicción
9. Almacenes de documentos y administración de documentos
10. Sistemas de gestión empresarial
11. Sistemas de Información Ejecutiva (EIS)
12. Finanzas y Presupuesto
13. Recursos humanos
14. Gestión del Conocimiento
15. Mapeo, visualización de información y Tableros (Dashboards)
16. Sistemas de Información Gerencial (MIS)
17. Sistemas de Información Geográfica (SIG)
18. Inteligencia de negocios en tiempo real
19. Estadísticas y análisis de datos técnicos
20. Gestión de la Cadena de Suministro / Gestión de la cadena de Demandas
21. Inteligencia de sistemas
22. Análisis de tendencia
23. Consulta del usuario / usuario final y generación de informes
24. Personalización Web y Minería Web
25. Extracción de textos

## **Componentes de la Inteligencia de Negocios**

### **Data Warehouse**

Un data warehouse es un repositorio de datos que proporciona una visión global, común e integrada de los datos de la organización –independientemente de cómo se vayan a utilizar posteriormente por los consumidores o usuarios–, con las propiedades siguientes: estable, coherente, fiable y con información histórica. Al abarcar un ámbito global de la organización y con un amplio alcance histórico, el volumen de datos puede ser muy grande (centenas de terabytes). Las bases de datos relacionales son el soporte técnico más comúnmente usado para almacenar las estructuras de estos datos y sus grandes volúmenes. Resumiendo, presenta las siguientes características:

- Temático: los datos están vinculados al negocio y organizados por las diferentes funciones que lleva a cabo el mismo.
- Integrado: incluye datos de múltiples orígenes y presenta consistencia de datos.
- Variable en el tiempo: se realizan fotos de los datos basadas en fechas o hechos.
- No volátil: sólo de lectura para los usuarios finales.

Frecuentemente el data warehouse está constituido por una base de datos relacional, pero no es la única opción factible. Debemos tener en cuenta que existen otros elementos en el contexto de un data warehouse:

- **Data Warehousing:** es el proceso de extraer y filtrar datos de las operaciones comunes de la organización, procedentes de los distintos sistemas de información operacionales y/o sistemas externos, para transformarlos, integrarlos y almacenarlos en un almacén de datos con el fin de acceder a ellos para dar soporte en el proceso de toma de decisiones de una organización.
- **Data Mart:** es un subconjunto de los datos del data warehouse cuyo objetivo es responder a un determinado análisis, función o necesidad, con una población de usuarios específica. Al igual que en un data warehouse, los datos están estructurados en modelos de estrella o copo de nieve, y un data mart puede ser dependiente o independiente de un data warehouse.  
Por ejemplo, un posible uso sería para la minería de datos o para la información de marketing. El data mart está pensado para cubrir las necesidades de un grupo de trabajo o de un determinado departamento dentro de la organización.
- **Operational Data Store:** es un tipo de almacén de datos que proporciona sólo los últimos valores de los datos y no su historial; además, generalmente admite un pequeño desfase o retraso sobre los datos operacionales.
- **Staging Area:** es el sistema que permanece entre las fuentes de datos y el data warehouse con el objetivo de:
  - Facilitar la extracción de datos desde fuentes de origen con una heterogeneidad y complejidad grande.
  - Mejorar la calidad de datos.
  - Ser usado como caché de datos operacionales con el que posteriormente se realiza el proceso de data warehousing.
  - Uso de la misma para acceder en detalle a información no contenida en el data warehouse.
- **Procesos ETL:** tecnología de integración de datos basada en la consolidación de datos que se usa tradicionalmente para alimentar data warehouse, data mart, staging área y ODS. Usualmente se combina con otras técnicas de consolidación de datos.
- **Metadatos:** Son datos acerca de los datos, para un data warehouse los metadatos almacenan la ruta del mapa del almacén de datos, funcionan como un directorio para localizar los objetos y el contenido que conforma el data warehouse.

### Modelado de Data warehouse

Es el proceso para la creación de un Data warehouse, existen varios enfoques para implementarlos:

## Enfoque de Inmon

Este enfoque es fundado por Bill Inmon en los años 90 para satisfacer los requisitos de las empresas y permitirles desarrollar sistemas para la toma de decisiones. Permite almacenar todos los eventos de una empresa; este enfoque se basa en los diagramas entidad-relación de los sistemas transaccionales. Los datos de la empresa se cargan sin conocer a priori los requisitos del usuario; esta es la razón por la que se llama enfoque basado en datos.

La arquitectura del almacén de datos Inmon (véase la figura 1) incluye todos los sistemas de información de la empresa en lugar de considerar sólo fragmentos de información.

Inmon divide el entorno de las bases de datos de la empresa en cuatro niveles, incluyendo: operacional, atómico (data warehouse), departamentales (data marts) y niveles individuales.

Los tres últimos niveles se refieren al data warehouse mientras el primer nivel (operacional) apoya las operaciones diarias y contiene los datos transaccionales de la empresa. Estos datos se transforman y cargan en el atómico data warehouse utilizando ETL (Extract, Transform and Load).

De acuerdo con los requerimientos de los departamentos de la empresa, los datos del nivel departamental están orientados a un tema, siempre consistentes porque provienen de un mismo data warehouse. Inmon considera que un data warehouse y un data mart están físicamente separados.

El almacén de datos tiene su propia existencia física y se orienta al almacenamiento, trazabilidad y escalabilidad en respuesta a nuevos requerimientos. Además, los data marts tienen su propia existencia física y ofrecen restitución, estructura orientada al rendimiento en respuesta a los requerimientos del usuario.

Los usuarios crean el último nivel cuando analizan y explotan los datos cargados en los data marts (OLAP Analysis, Informes, Dashboards,... etc.).

Inmon propone una arquitectura denominada Corporate Information Factory (o Enterprise Data Warehouse), donde existe un data warehouse corporativo y unos data marts (o incluso cubos OLAP) dependientes del mismo. El acceso a datos se realiza a los data marts o a la ODS en caso de existir, pero nunca al propio data warehouse. Puede existir en el caso de ser necesaria una staging área.



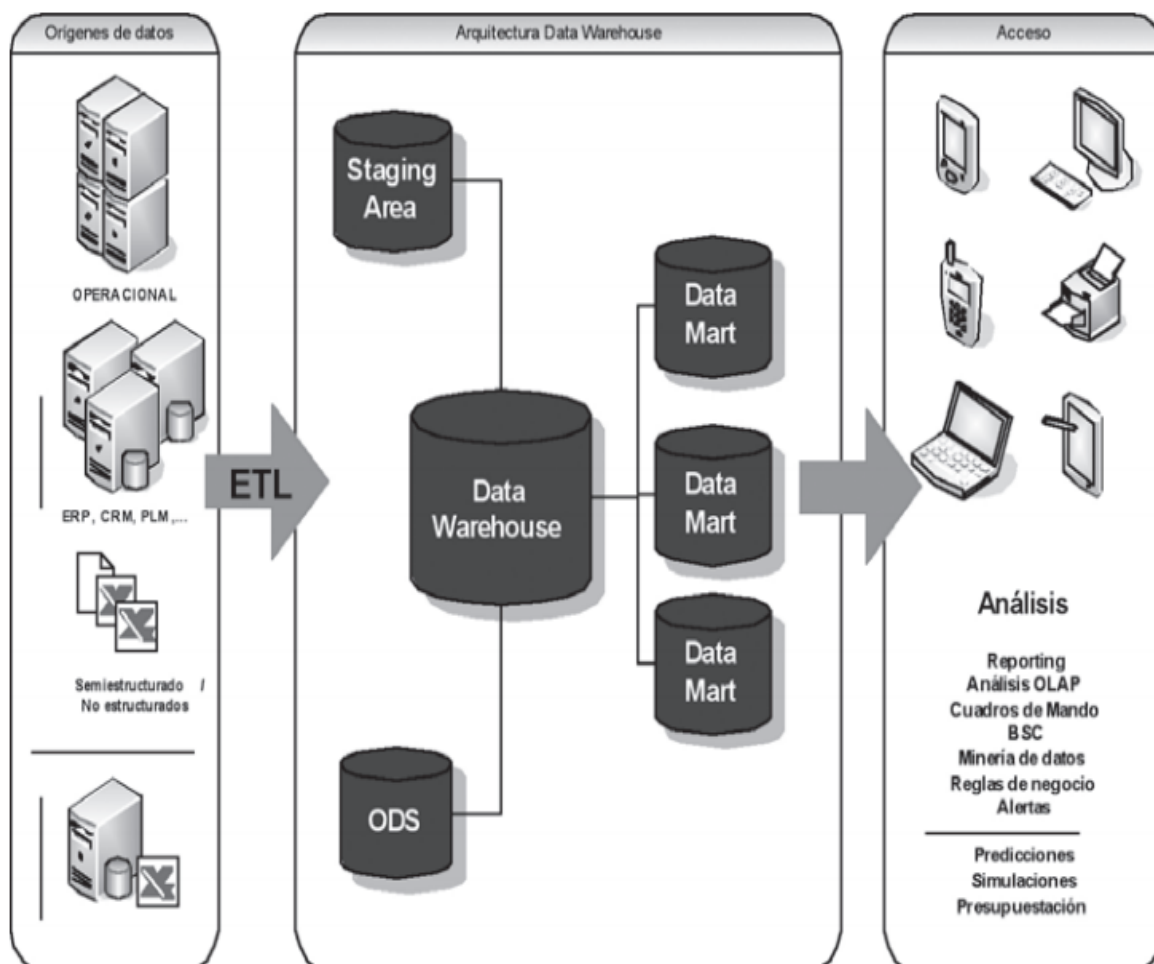


Figura 1. Corporate Information Factory. *Introducción al business Intelligence* (p.40), por Josep Curto Díaz y Jordi Conesa Caralt, 2010, Barcelona: UOC, Copyright [2010] por Josep Curto Díaz y Jordi Conesa i Caral. Reimpresión autorizada

## Enfoque de Kimball

Kimball creó su enfoque en los años 90 ofreciendo una nueva arquitectura, una nueva visión y un modelo innovador de data warehouse.

Este enfoque se basa en el concepto de modelado dimensional. Este enfoque se opone al principio de aislamiento de los usuarios finales enunciado por Inmon. De hecho, en esta propuesta se involucra a los usuarios finales en las primeras etapas del proyecto, por eso se le denomina enfoque basado en los requerimientos del usuario. Se presenta una visión diferente de los data warehouse pues considera que estos pueden ser vistos como un conjunto de data marts de datos consistentes y basados en dimensiones compartidas.

Kimball propone una arquitectura denominada Enterprise Bus Architecture (o Data Warehouse Virtual/Federado) la cual permite un desarrollo incremental del data warehouse y de inteligencia de negocios (DW / BI). Descompone el proceso de planificación DW / BI en partes manejables, centrándose en los procesos de negocio de la organización, junto con las dimensiones conformadas asociadas.

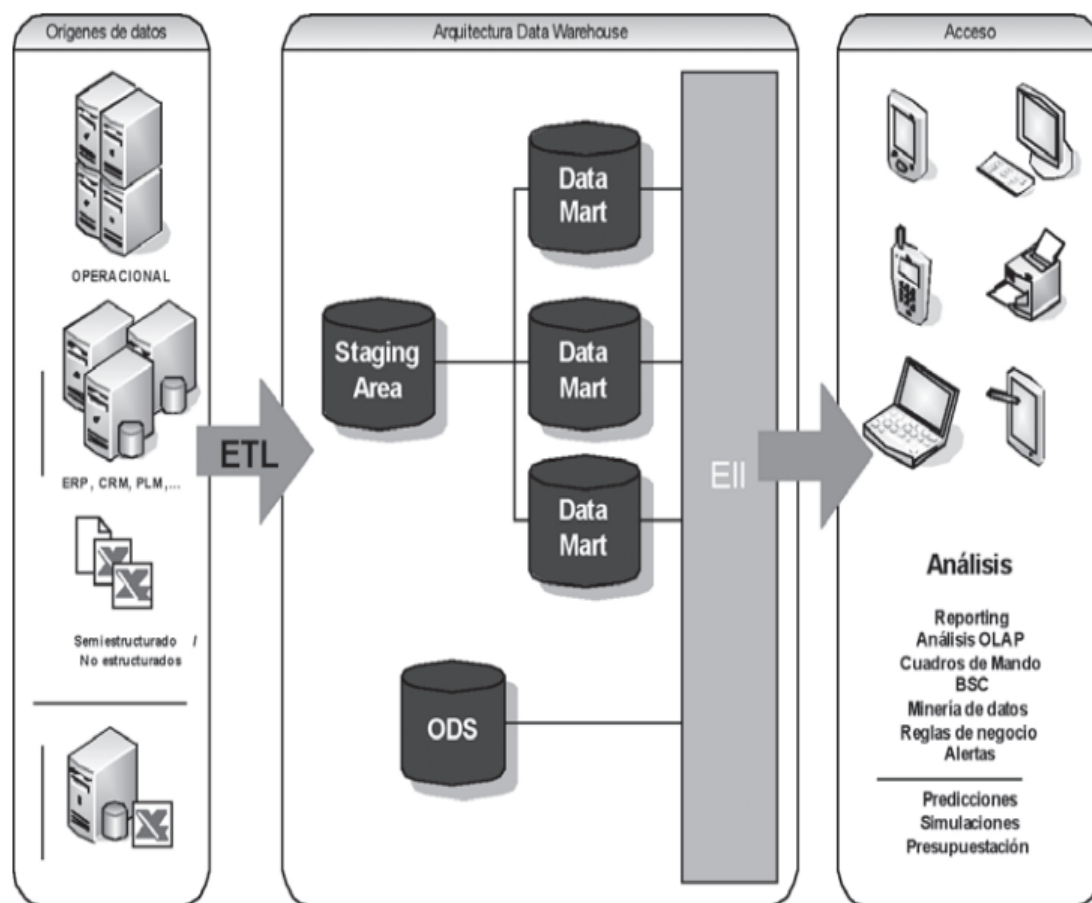


Figura 2. Enterprise bus. *Introducción al business Intelligence* (p.39), por Curto, J. & Conesa, J., 2010, Barcelona: UOC, Copyright [2010] por Josep Curto Díaz y Jordi Conesa i Caral. Reimpresión autorizada

## Elementos de un Datawarehouse

La idea principal es que la información sea presentada desnormalizada para optimizar las consultas. Para ello se debe identificar, en el núcleo de nuestra organización, los procesos de negocio, las vistas para el proceso de negocio y las medidas cuantificables asociadas a los mismos. De esta manera hablaremos de:

- **Tabla de hecho:** es la representación en el data warehouse de los procesos de negocio de la organización. Por ejemplo, una venta puede identificarse como un proceso de negocio de manera que es factible, si corresponde en la organización, considerar la tabla de hecho ventas.
- **Dimensión:** es la representación en el data warehouse de una vista para un cierto proceso de negocio.
- **Métrica (Medida):** son los indicadores de negocio de un proceso del mismo. Aquellos conceptos cuantificables que permiten medir nuestro proceso de negocio. Por ejemplo, en una venta se tiene el importe de la misma.

## Esquemas

Para estructurar los datos en el data warehouse este proceso se realiza mediante un esquema (en inglés, Database Schema) el cual describe la estructura de una Base de datos, en un lenguaje formal soportado por un Sistema administrador de Base de datos (DBMS). En una Base de datos Relacional, el Esquema define sus tablas, sus campos en cada tabla y las relaciones entre cada campo y cada tabla. En la práctica el término esquema de la base de datos se refiere al diseño físico de la base de datos.

Existen tres tipos de esquemas:

- **Esquema en estrella:** consiste en estructurar la información en procesos, vistas y métricas recordando a una estrella (por ello el nombre). A nivel de diseño, consiste en una tabla de hechos (lo que en los libros encontraremos como fact table) en el centro para el hecho objeto de análisis y una o varias tablas de dimensión por cada punto de vista de análisis que participa de la descripción de ese hecho. En la tabla de hecho encontramos los atributos destinados a medir (cuantificar): sus métricas. La tabla de hechos sólo presenta uniones con dimensiones.
- **Esquema en copo de nieve:** es un esquema de representación derivado del esquema en estrella, en el que las tablas de dimensión se normalizan en múltiples tablas. Por esta razón, la tabla de hechos deja de ser la única tabla del esquema que se relaciona con otras tablas, y aparecen nuevas uniones. Es posible distinguir dos tipos de esquemas en copo de nieve:
  - Completo: en el que todas las tablas de dimensión en el esquema en estrella aparecen ahora normalizadas.
  - Parcial: sólo se lleva a cabo la normalización de algunas de ellas.

Una tabla de hecho es una representación de un proceso de negocio. A nivel de diseño es una tabla que permite guardar dos tipos de atributos diferenciados:

- Medidas del proceso/actividad/flujo de trabajo/evento que se pretende modelar.
- Claves foráneas hacia registros en una tabla de dimensión (o en otras palabras, como ya sabemos, hacia una vista de negocio).

Las dimensiones recogen los puntos de análisis de un hecho. Por ejemplo, una venta se puede analizar en función del día de venta, producto, cliente, vendedor o canal de venta, entre otros.

- **Esquema Constelación:** este modelo está compuesto por una serie de esquemas en estrella, está formado por una tabla de hechos principal y por una o más tablas de hechos auxiliares, las cuales pueden ser sumalizaciones de la principal. Dichas tablas yacen en el centro del modelo y están relacionadas con sus respectivas tablas de dimensiones.

No es necesario que las diferentes tablas de hechos compartan las mismas tablas de dimensiones, ya que, las tablas de hechos auxiliares pueden vincularse con solo algunas de las tablas de dimensiones asignadas a la tabla de hechos principal, y también pueden hacerlo con nuevas tablas de dimensiones.

## OLAP

El concepto de Business Intelligence engloba múltiples conceptos. Uno de los más importantes es el concepto OLAP que es una tecnología que permite un análisis multidimensional<sup>30</sup> a través de tablas matriciales o pivotantes.

Se entiende por OLAP, o proceso analítico en línea, al método ágil y flexible para organizar datos, especialmente metadatos, sobre un objeto o jerarquía de objetos como en un sistema u organización multidimensional, y cuyo objetivo es recuperar y manipular datos y combinaciones de los mismos a través de consultas o incluso informes.

OLAP forma parte de lo que se conoce como sistemas analíticos, que permiten responder preguntas como: ¿por qué pasó? estos sistemas pueden encontrarse tanto integrados en suites de Business Intelligence o ser simplemente una aplicación independiente.

Una herramienta OLAP está formada por un motor y un visor. El motor es, en realidad, justo el concepto que acabamos de definir. El visor OLAP es una interfaz que permite consultar, manipular, reordenar y filtrar datos existentes en una estructura OLAP mediante una interfaz gráfica de usuario que dispone funciones de consulta MDX (Multidimensional Expressions) y otras. Las estructuras OLAP permiten realizar preguntas que serían sumamente complejas mediante SQL.

### Tipos de OLAP

Existen diferentes tipos de OLAP, que principalmente difieren en cómo se guardan los datos:

- **MOLAP (Multidimensional OLAP):** es la forma clásica de OLAP y frecuentemente es referida con dicho acrónimo. MOLAP utiliza estructuras de bases de datos generalmente optimizadas para la recuperación de los mismos. Es lo que se conoce como bases de datos multidimensionales (o, más coloquialmente, cubos). En definitiva, se crea un fichero que contiene todas las posibles consultas pre calculadas. A diferencia de las bases de datos relacionales, estas formas de almacenaje están optimizadas para la velocidad de cálculo. También se optimizan a menudo para la recuperación a lo largo de patrones jerárquicos de acceso. Las dimensiones de cada cubo son típicamente atributos tales como periodo, localización, producto o código de la cuenta. La forma en la que cada dimensión será agregada se define por adelantado.

---

<sup>30</sup> En 1962, se introduce el análisis multidimensional en el libro de Ken Iverson A Programming Language.

- **ROLAP (Relational OLAP):** trabaja directamente con las bases de datos relacionales, que almacenan los datos base y las tablas dimensionales como tablas relacionales mientras se crean nuevas tablas para guardar la información agregada.
- **HOLAP (Hybrid OLAP):** no hay acuerdo claro en la industria en cuanto a qué constituye el OLAP híbrido, exceptuando el hecho de que es una base de datos en la que los datos se dividen en almacenaje relacional y multidimensional. Por ejemplo, para algunos vendedores, HOLAP consiste en utilizar las tablas relacionales para guardar las cantidades más grandes de datos detallados, y utiliza el almacenaje multidimensional para algunos aspectos de cantidades más pequeñas de datos menos detallados o agregados.
- **DOLAP (Desktop OLAP):** es un caso particular de OLAP ya que está orientado a equipos de escritorio. Consiste en obtener la información necesaria desde la base de datos relacional y guardarla en el escritorio. Las consultas y los análisis son realizados contra los datos guardados en el escritorio.
- **In-memory OLAP:** es un enfoque por el que muchos nuevos fabricantes están optando. Consiste en que la estructura dimensional se genera sólo a nivel de memoria y se guarda el dato original en algún formato que potencia su despliegue de esta forma (por ejemplo, comprimido o mediante una base de datos de lógica asociativa).

#### **Reglas para evaluar soluciones OLAP de E. F. Codd**

- Vista conceptual multidimensional: se trabaja a partir de métricas de negocio y sus dimensiones, el negocio debe verse por naturaleza de forma multidimensional.
- Transparencia: el sistema OLAP debe formar parte de un sistema abierto que soporta fuentes de datos heterogéneas (lo que llamamos actualmente arquitectura orientada a servicios).
- Accesibilidad: se debe presentar el servicio OLAP al usuario con un único esquema lógico de datos (lo que, en definitiva, nos indica que debe presentarse respecto una capa de abstracción directa con el modelo de negocio).
- Rendimiento de informes consistente: el rendimiento de los informes no debería degradarse cuando el número de dimensiones del modelo se incrementa.
- Arquitectura cliente/servidor: basado en sistemas modulares y abiertos que permitan la interacción y la colaboración.
- Dimensionalidad genérica: capacidad de crear todo tipo de dimensiones y con funcionalidades aplicables de una dimensión a otra.
- Dynamic sparse-matrix handling: la manipulación de datos en los sistemas OLAP debe poder diferenciar valores vacíos de valores nulos y además poder ignorar las celdas sin datos.

- Operaciones cruzadas entre dimensiones sin restricciones: todas las dimensiones son creadas igual y las operaciones entre dimensiones no deben restringir las relaciones entre celdas.
- Manipulación de datos intuitiva: dado que los usuarios a los que se destinan este tipo de sistemas son frecuentemente analistas y altos ejecutivos, la interacción debe considerarse desde el prisma de la máxima usabilidad de los usuarios.
- Reporting flexible: los usuarios deben ser capaces de manipular los resultados que se ajusten a sus necesidades conformando informes. Además, los cambios en el modelo de datos deben reflejarse automáticamente en esos informes.
- Niveles de dimensiones y de agregación ilimitados: no deben existir restricciones para construir cubos OLAP con dimensiones y con niveles de agregación ilimitados.

## ETL

Los procesos de ETL en sus siglas en inglés Extract, Transform and Load que significa <<extraer, transformar y cargar>> están íntimamente ligados con la creación de data warehouse. Para que los datos contenidos en un data warehouse sean precisos y tengan relevancia, tienen que pasar por un proceso que permita ser transformado de acuerdo a como la estructura donde se almacenarán lo requieran. El objetivo principal de los procesos ETL es facilitar la migración y la transformación de los datos.

Los procesos ETL se pueden definir como la tecnología de integración de datos basada en la consolidación de datos que se usa tradicionalmente para alimentar data warehouse, data mart, staging área y ODS. Usualmente se combina con otras técnicas de consolidación de datos.

Los procesos ETL son una colección de aplicaciones o programas que aportan valor añadido a los datos. Esto va más allá de solo depurar los datos que serán almacenados en un datawarehouse. Los sistemas ETL realizan varias tareas:

- Remueven datos incorrectos o corrigen datos faltantes.
- Proveen de indicadores de confiabilidad en los datos.
- Capturan el flujo de los datos transaccionales de manera segura.
- Ajustan los datos provenientes de múltiples fuentes para ser usados de manera conjunta.
- Reestructuran los datos para ajustarse a las herramientas usadas por el usuario final.

## **Técnicas de modelado**

La construcción del ETL puede dividirse en tres subprocesos o componentes: componente de extracción, componente de transformación y componente de carga.

La tarea de diseñar procesos ETL involucra: (1) analizar las fuentes de datos existentes para encontrar la semántica oculta en ellas y (2) diseñar el flujo de trabajo que extraiga los datos desde las fuentes, repare sus inconsistencias, los transforme en un formato deseado, y, finalmente, los inserte en la bodega de datos. Con el propósito de facilitar la tarea de modelar los procesos ETL, se han desarrollado diferentes técnicas, dos categorías que sobresalen son las siguientes.

### **Enfoque orientado al modelado de procesos.**

Con este enfoque existen dos tipos de modelado: el conceptual y el lógico. Para el modelado conceptual se plantea identificar cuáles son las fuentes, los atributos y las transformaciones que se necesitan realizar. Para el modelado lógico se plantea considerar además otros elementos como el conjunto de valores que el atributo admite, cuales son los parámetros necesarios para que las funciones se realicen de manera correcta.

### **Enfoque orientado a objetos.**

En este enfoque se plantea usar diagramas de UML. Al emplear la adopción del diagrama de clases de UML, se logra cubrir las especificaciones de conceptos como clases, atributos y métodos. Además de ello se logra acortar la curva de aprendizaje al ser una buena práctica del desarrollo de software, al asimilar una nueva tecnología.

## **Técnicas de integración de datos**

Las herramientas de ETL han sido la opción usual para alimentar el data warehouse. La funcionalidad básica de estas herramientas consiste básicamente en la extracción, transformación y carga de datos.

La integración de datos se conoce como al conjunto de aplicaciones, productos, técnicas y tecnologías que permiten tener una visión única y consistente de los datos del negocio. Existen diferentes tecnologías basadas en estas técnicas.

### **ETL de generación de código**

Constan de un entorno gráfico en donde se diseñan y se especifican los datos de origen, sus transformaciones y sus entornos de destino. EL resultado es un programa de tercera generación, usualmente basado en Cobol que permite realizar las transformaciones de datos.

### **ETL basados en motor**

Permite crear flujos de trabajo en tiempo de ejecución definidos mediante herramientas gráficas. El entorno gráfico permite hacer un mapping de los entornos de datos de origen y destino, las transformaciones de datos necesarios, el flujo de procesos y los procesos por

lotes necesarios. Los motores generalmente usados para realizar estos procesos se pueden componer por diversos motores como adaptadores ODBC, JDBC, JNDI, etc.

### **ETL integrado en la base de datos**

Algunos fabricantes incluyen capacidades de ETL dentro del motor de la base de datos. En general, presentan menos funcionalidades y complejidad, y son solución menos completa que los ETL comerciales basados en motor o generación de código.

### **Diseño de reportes / Visualizaciones**

La intención de los reportes es que el usuario del negocio pueda recibir información pertinente que pueda contribuir a la toma de decisiones. Existen diferentes tipos de informe en función de la interacción ofrecida al usuario.

- **Estáticos:** tienen un formato preestablecido inamovible.
- **Paramétricos:** presentan parámetros de entrada y permiten múltiples consultas.
- **Ad-hoc:** son creados por el usuario final a partir de la capa de metadatos que permite usar el lenguaje de negocio propio.

### **Elementos de un informe**

Es necesario destacar cuáles son los componentes que tiene un informe, puede estar conformado por:

- **Texto:** que describe el estado del proceso de negocio o proporciona las descripciones necesarias para entender el resto de elementos del informe.
- **Tablas:** este elemento tiene forma de matriz y permite presentar una gran cantidad de información.
- **Gráficos:** este elemento persigue el objetivo de mostrar información con un alto impacto visual que sirva para obtener información agregada o resumizada con mucha más rapidez que a través de tablas.
- **Mapas:** este elemento permite mostrar información geolocalizada.
- **Métricas:** que permiten conocer cuantitativamente el estado de un proceso de negocio.
- **Alertas visuales y automáticas:** consiste en avisos del cambio de estado de información que pueden estar formadas por elementos gráficos como fechas o colores resultados y que deben estar automatizadas en función de reglas de negocio encapsuladas en el cuadro de mando.



## **Tipos de métricas**

Los informes poseen métricas que son propias del negocio, existen diferentes tipos de medida que se clasifican según el tipo de información que se requiere.

- Métricas: valores que recogen el proceso de una actividad o los resultados de la misma. Estas medidas proceden del resultado de la actividad de negocio.
  - Métricas de realización de actividad (leading): miden la realización de una actividad. Por ejemplo, la participación de una persona en un evento.
  - Métricas de resultado de una actividad (lagging): recogen los resultados de una actividad. Por ejemplo, la cantidad de puntos de un jugador en un partido.
- Indicadores clave: entendemos por este concepto, valores correspondientes que hay que alcanzar, y que suponen el grado de asunción de los objetivos. Estas medidas proporcionan información sobre el rendimiento de una actividad o sobre la consecución de una meta.
  - Key Performance Indicator (KPI): indicadores clave de rendimiento. Más allá de la eficacia, se definen unos valores que nos explican en qué rango óptimo de rendimiento nos deberíamos situar al alcanzar los objetivos. Son métricas del proceso.
  - Key Goal Indicator (KGI): indicadores de metas. Definen mediciones para informar a la dirección general si un proceso TIC ha alcanzado sus requisitos de negocio, y se expresan por lo general en términos de criterios de información.

## **Tipos de gráficos**

Existen diferentes tipos de gráficos que ayudan a una mejor visualización y dependiendo de ciertos criterios es mejor utilizar un gráfico por sobre otro, a continuación el siguiente diagrama reconoce los criterios necesarios para la elección de un gráfico adecuado de acuerdo a la información que se desea presentar.

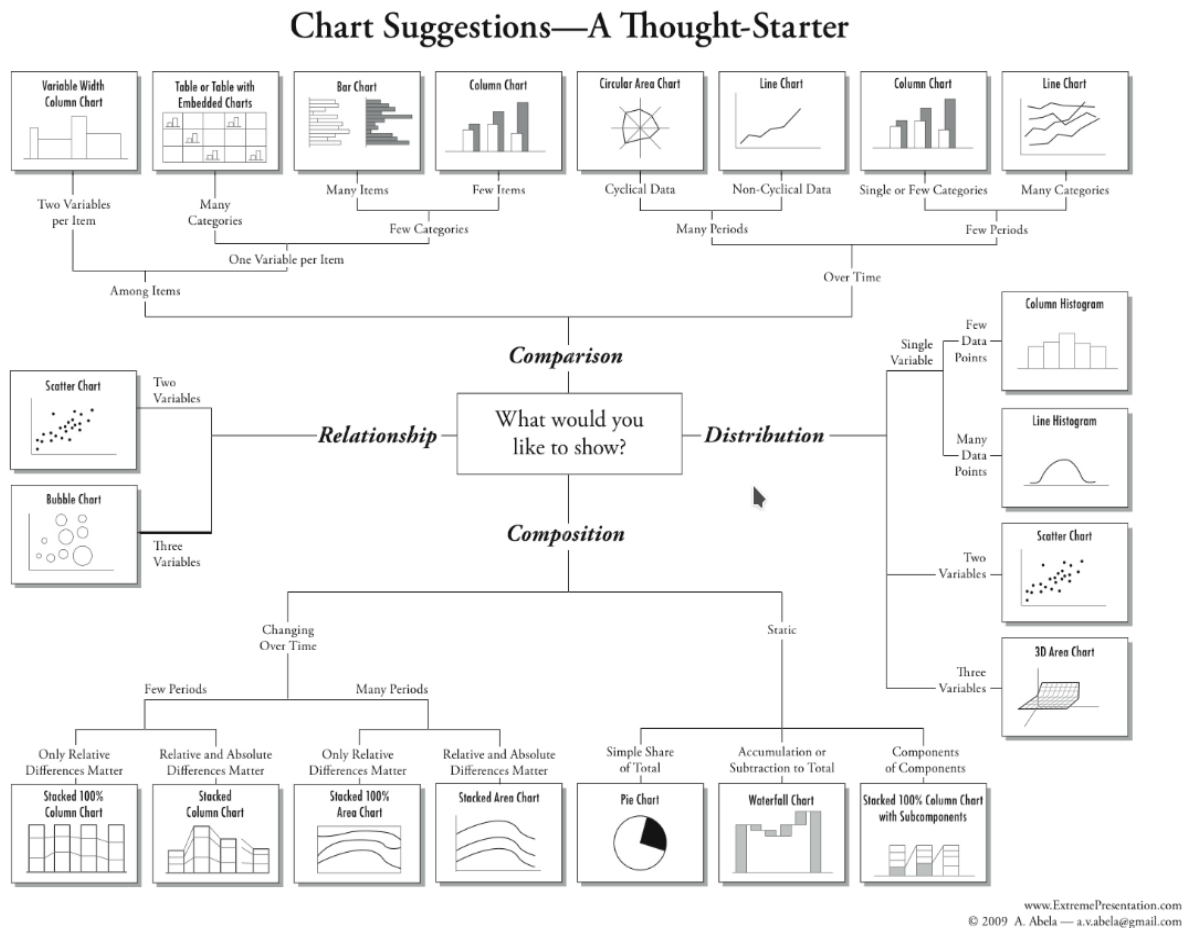


Figura 3. *Chart Suggestions - A Thought-Starter*. En “Extreme presentation tools”, por A. Abela, 2009, <http://extremepresentation.com>. Copyright [2009]. The Extreme Presentation Method. Reimpresión autorizada.

## Selección de herramientas

El cuadro mágico de Gartner es una gráfica donde se colocan los proveedores de tecnología que ofrecen las mejores plataformas para realizar inteligencia de negocios y analítica. Dentro de los mejores 3 se encuentran posicionados los siguientes.



Figura 4. *Gartner Magic Quadrant*. En “Microsoft breaks through in the Gartner Magic Quadrant for Business Intelligence and Analytics Platforms”, por K. Hathi, 2017, <https://blogs.microsoft.com/blog/2017/02/16/microsoft-breaks-gartner-magic-quadrant-business-intelligence-analytics-platforms/#sm.0000n49zhpbs4e41tvk14qnbii4p> . Copyright [2017]. Official Microsoft blog. Reimpresión autorizada.

## Tableau

Es una herramienta que está enfocada a la visualización de datos de manera interactiva, con Tableau Desktop se puede conectar a datos provenientes de distintas fuentes como base de datos SQL, Big Data, hojas de cálculo, etc. Además, esta herramienta incorpora formas para realizar cálculos complejos, estadísticas, pronósticos, etc.

## Microsoft BI

Dentro de las herramientas que proporciona Microsoft, Power BI es un conjunto de aplicaciones de análisis de negocios que permite analizar de datos y además compartir información. Esta herramienta consta de varias funcionalidades como la combinación de

datos provenientes de múltiples fuentes como bases de datos, archivos, servicios web y provee de herramientas visuales para la construcción de dashboard personalizados.

## **Qlik**

Es un set de herramientas que permite la exploración de datos y la visualización de información. Permite analizar varias fuentes de datos de manera dinámica e interactiva. Ofrece visualizaciones atractivas e inteligentes además que permite guardar de manera segura toda información sensible de la organización.

### **2.2.2 Analítica de Big Data**

Hasta este punto ya se ha hablado sobre inteligencia de negocios, pero para poder abordar el tema de analítica de datos y su enfoque en Big Data, es necesario abordar la diferencia entre estos tres tópicos.

#### **Analítica de datos**

Es preciso empezar diciendo que la analítica de datos es el medio para llegar a la inteligencia de negocios y Big Data. Para hacer inteligencia de negocios nos basamos en el análisis de los datos, para hacer analítica de Big Data nos basamos igualmente del mismo proceso. Entonces, el análisis o la analítica de datos es un término general para cualquier tipo de procesamiento, que implica la recopilación, integración y preparación de los datos (data profiling y data cleansing). Además, si el análisis es más avanzado se habla del desarrollo, prueba y revisión de modelos analíticos para asegurar que se produzcan resultados precisos.

La analítica de datos tiene diferentes metodologías que corresponden a diferentes niveles de aplicación. A groso modo, las metodologías de análisis de datos incluyen el *análisis exploratorio de datos (EDA)*, cuyo objetivo es encontrar patrones y relaciones en los datos y el *análisis de datos confirmatorios (CDA)*, que aplica técnicas estadísticas para determinar si las hipótesis sobre un conjunto de datos son verdaderas o falsas.

El análisis de datos también puede separarse en *análisis de datos cuantitativos* y *análisis de datos cualitativos*. La primera implica el análisis de datos numéricos con variables cuantificables que pueden ser comparadas o medidas estadísticamente. El enfoque cualitativo es más interpretativo - se centra en la comprensión del contenido de datos no numéricos como texto, imágenes, audio y video, incluyendo frases comunes, temas y puntos de vista.

Los tipos de análisis de datos más avanzados incluyen la *minería de datos*, que implica clasificar grandes conjuntos de datos para identificar tendencias, patrones y relaciones; *analítica predictiva*, que busca predecir el comportamiento del cliente, fallas de equipos y otros eventos futuros; y *machine learning*, una técnica de inteligencia artificial que utiliza algoritmos automatizados para manipular conjuntos de datos más rápidamente de lo que puede hacerse a través del modelado analítico convencional.

## **Análisis Exploratorio de Datos**

Es un modo de análisis de los datos que utiliza el resumen numérico y visual para explorar datos en busca de patrones no anticipados. En general, el análisis exploratorio de datos se caracteriza por el uso de herramientas o técnicas con mucha carga visual o gráfica, con énfasis en revelar información vital sobre la data examinada. El arsenal correspondiente está compuesto, entre otros, por instrumentos como:

- Diagrama de caja y bigotes (Box-and-whisker): un resumen visual de la distribución (comportamiento) de una variable que provee detalles acerca de si uno o ambos extremos de la distribución contienen valores inusualmente grandes o pequeños.
- Diagrama de tallo y hojas (Stem-and-leaf): es una muestra visual de la distribución de una variable. Se asemeja a un histograma y "... se usa mucho para mostrar tanto el orden de rangos como la forma de un conjunto de datos en forma simultánea". (Anderson y otros, 1999: 40). Cada caso es representado por uno o más dígitos colocados a la derecha de una línea vertical y en la fila correspondiente al primer dígito del valor observado.
- Diagrama de dispersión (Scatter plot): gráfico que muestra la relación entre variables. Es útil para examinar la dirección, fuerza y forma de la relación.

## **Análisis Confirmatorio de Datos**

Es un modo de análisis de datos que utiliza estadísticos numéricos de resumen generados a partir del empleo de un modelo, definido a priori, para confirmar o no una hipótesis. Se caracteriza por el empleo de indicadores como la media, la varianza y los coeficientes de correlación y regresión, así como las pruebas de hipótesis.

Así mismo, puede decirse que este es el modelo de análisis de datos cuya rutina es más sencilla de enseñar y de estandarizar mediante el uso de computadores, como se puede constatar por el desarrollo de programas informáticos como el SPSS, SAS, SPAD y STATGRAPHICS, entre otros.

## **Análisis Cuantitativo (Análisis descriptivo)**

### **Análisis Univariante**

Los métodos de análisis univariable se utilizan para estudiar el comportamiento de las variables de forma individual.

1. Distribuciones de frecuencias:  
Permiten obtener una primera aproximación de la tendencia de los resultados, ya que indican el número de individuos que tanto en valores absolutos como en valores porcentuales han mencionado cada uno de los códigos posibles (respuestas) que puede tomar una variable determinada.

## 2. Medidas de tendencia central

Los tipos estadísticos que miden la tendencia central permiten apreciar cuál es el comportamiento medio de cada variable. Los tres más utilizados son la moda, la mediana y la media.

## 3. Medidas de dispersión

Permiten analizar la variabilidad de la distribución, es decir, determinar si las respuestas que han dado las personas entrevistadas se han concentrado sólo en unos cuantos valores o si, por el contrario, han sido muy variadas. La dispersión se mide respecto del comportamiento medio de la variable, por lo que la elección de la medida de dispersión que hay que utilizar también depende de la escala en que esté medida la variable que se analiza.

La medida que permite evaluar la dispersión de las respuestas respecto de la media cuando la escala de medida es cuantitativa es la varianza. Otras medidas de dispersión que permiten completar la información suministrada por la varianza son las siguientes:

- El coeficiente de simetría (skewness): indica el grado de simetría o asimetría de la distribución.
- El coeficiente de apuntamiento (curtosis): valora si las respuestas están concentradas en pocos valores o están repartidas.

## 4. Inferencia estadística

Una vez obtenidos los resultados es importante validarlos, es decir, ver si hay diferencias entre los resultados obtenidos en la investigación y unos valores determinados conocidos a priori o teóricos; en caso de que las haya, hay que comprobar si estas diferencias son estadísticamente significativas o si, por el contrario, se deben al azar. El proceso que hay que seguir para validar los resultados se denomina test de inferencia estadística.

## **Análisis Multivariante**

Según Rosana Ferrero, científica de datos y profesora del Máster en Estadística Aplicada con R Software, en su artículo “Análisis Multivariado para Big Data” clasifica las técnicas de análisis multivariante de la siguiente manera:

### 1. Por grupos

- a. Encontrar grupos (Clúster)
- b. Comparar grupos (análisis de varianza multivariado MANOVA, el test de permutación multirespuesta MRPP, el análisis de similitudes de grupos ANOSIM, y el test de mantel MANTEL)
- c. Discriminar entre grupos (análisis discriminante lineal LDA, árbol de regresión o clasificación CART/MRT)

## 2. Gradientes

- a. Ordenación no restringida (análisis de componentes principales PCA, análisis de correspondencia CA, análisis de correspondencias sin tendencia DCA, escalamiento multidimensional métrico MDS y no métrico NMDS)
- b. Ordenación restringida (análisis de redundancia RDA, análisis de correspondencia canónico CCA)

### **Análisis cualitativo (Análisis predictivo)**

Hoy en día, se utilizan cada vez más software informáticos para análisis de datos cualitativos (QDA). Un software QDA se utiliza en muchos campos académicos, como la sociología, la psicología, la ciencia política, la medicina y la ciencia educativa, y es también una herramienta popular para las empresas y los investigadores de mercado.

Un software QDA sirve para proporcionar información sobre conjuntos de datos cualitativos sin sugerir interpretaciones. Basándose en un análisis de contenido, se puede sacar conclusiones sobre el objeto que se está estudiando (por ejemplo, datos de entrevistas). Las herramientas de software para el análisis cualitativo de datos y textos permiten una fácil clasificación, estructuración y análisis de grandes cantidades de texto u otros datos y facilitan la gestión de las interpretaciones y evaluaciones resultantes.

Existe una tendencia hacia la inclusión y el análisis de los archivos de imágenes, así como de los materiales de audio y video. Por lo tanto, un software de análisis de datos cualitativo de última generación apoya idealmente el análisis de textos, así como de archivos multimedia (imágenes, videos, grabaciones de audio) y otros tipos de documentos.

Ejemplo de herramientas que apoyan al análisis cualitativo son: QDA Miner, Atlas.ti, HyperRESEARCH, MAXQDA, NVivo, entre otras.

Aclarado este punto, corresponde tratar la diferencia entre la inteligencia de negocios y Big Data. La primera ayuda a encontrar respuestas a preguntas conocidas (puesto que define qué información es útil y relevante para la toma de decisiones en base a los datos existentes) mientras que Big Data ayuda a encontrar las cuestiones que no se sabe encontrar (se requiere de la integración de los datos de las diferentes fuentes para sacar conclusiones y aprovechar oportunidades futuras).

### **Análisis predictivo**

La analítica predictiva es una forma de análisis avanzado que utiliza datos nuevos e históricos para predecir la actividad futura, el comportamiento y las tendencias. Implica la aplicación de técnicas de análisis estadístico, consultas analíticas y algoritmos automáticos de machine learning sobre conjuntos de datos para crear modelos predictivos que sitúen un valor numérico o puntuación en la probabilidad de que ocurra un suceso particular.

Una amplia gama de herramientas y técnicas se utiliza en los modelos y análisis predictivos. IBM, Microsoft, el SAS Institute y muchos otros proveedores de software ofrecen herramientas analíticas predictivas, incluyendo software de machine learning y tecnologías relacionadas que soportan aplicaciones de aprendizaje profundo. Además, el software de código abierto juega un papel importante en el mercado de análisis predictivo. El lenguaje de código abierto R se utiliza comúnmente en aplicaciones de análisis predictivo, al igual que los lenguajes de programación Python y Scala.

## Data mining

La minería de datos o data mining es el proceso de clasificar grandes conjuntos de datos para identificar patrones y establecer relaciones para resolver problemas a través del análisis de datos. Las herramientas de minería de datos permiten a las empresas predecir tendencias futuras.

En la minería de datos, las reglas de asociación se crean analizando datos para patrones frecuentes si/entonces (if/then), usando los criterios de soporte y confianza para localizar las relaciones más importantes dentro de los datos. El soporte es la frecuencia con la que aparecen los elementos en la base de datos, mientras que la confianza es el número de veces que las sentencias if/then son exactas.

Otros parámetros de minería de datos incluyen Análisis de secuencia o trayectoria (Sequence or Path Analysis), Clasificación y Agrupamiento (clustering).

- Los parámetros de Análisis de secuencia o trayectoria buscan patrones en los que un evento conduce a otro evento posterior. Una secuencia es una lista ordenada de conjuntos de elementos y es un tipo común de estructura de datos que se encuentra en muchas bases de datos.
- Un parámetro de clasificación busca nuevos patrones y puede resultar en un cambio en la forma en que se organizan los datos. Los algoritmos de clasificación predicen variables basadas en otros factores dentro de la base de datos.
- Los parámetros de agrupación encuentran y documentan visualmente grupos de hechos previamente desconocidos. Clustering agrupa un conjunto de objetos y los agrega basados en lo similares que son entre sí.

## Big Data

No existe unanimidad en la definición de Big Data, aunque sí un cierto consenso en la fuerza disruptiva que suponen los grandes volúmenes de datos. En general, casi todas las definiciones están de acuerdo en la esencia de Big Data: crecimiento exponencial de la creación de grandes volúmenes de datos y la necesidad de su captura, almacenamiento y análisis para conseguir el mayor beneficio para organizaciones y empresas.

La empresa multinacional de auditoría Deloitte lo define como: “El término que se aplica a conjuntos de datos cuyo volumen supera la capacidad de las herramientas informáticas de uso común, para capturar, gestionar y procesar datos en un lapso de tiempo razonable. Los volúmenes.”

Una de las definiciones más acreditadas por venir de la consultora Gartner es: “Big Data son los grandes conjuntos de datos que tiene tres características principales: volumen (cantidad), velocidad (velocidad de creación y utilización) y variedad (tipos de fuentes de datos no estructurados, tales como la interacción social, video, audio)”.



## Características de Big Data

Big Data abarca diversas tecnologías. Los datos de entrada a los sistemas de Big Data pueden proceder de redes sociales, logs, registros de servidores web, sensores de flujos de tráfico, imágenes de satélites, flujos de audio y de radio, transacciones bancarias, datos de mercados financieros. ¿Todos estos datos son realmente los mismos?

Uno de los mayores desafíos ha sido identificar las características de dichos datos. Diferentes investigadores han tenido puntos de vistas distintos hacia estas características; como algunos mencionan las 3Vs [*Volumen, Velocidad y Variedad*] de datos (por ejemplo Shah, Rabhi, & Ray, 2015). Otros informaron 4Vs [*Volumen, Velocidad, Variedad y Variabilidad*] de datos (por ejemplo Liao, Yin, Huang, & Sheng, 2014) y 6Vs [*Volumen, Velocidad, Veracidad, Variedad, Variabilidad y Valor*] de datos (Gandomi y Haider, 2015).

Según una Crítica de la Revista de Investigación Empresarial<sup>31</sup> se identificaron 7 Vs, siete características de los datos que son: Volumen, Variedad, Velocidad, Variabilidad, Veracidad, Visualización y Valor.

### 1. Volumen:

Son datos a gran escala y volumen, lo cual conlleva a un gran desafío para la recuperación, procesamiento, integración e inferencia de los mismos. Podemos estar hablando de terabytes, petabytes o zettabytes generados en un momento. Facebook genera diariamente más de 500 terabytes de datos y Wal-Mart recolecta cada hora más de 2.5 petabytes como resultado de las transacciones con sus clientes.

### 2. Variedad:

Se refiere a las diferentes formas y cualidades de los datos, los datos son heterogéneos y no siguen un formato específico, debido a que son capturados de diversas formas y de múltiples fuentes. Por ejemplo: mensajes (texto, correo electrónico, tweets, blogs), contenido transaccional (registros web, transacciones comerciales), datos científicos (datos procedentes de experimentos, del área de la salud, bases de datos biológicas) y datos web (imágenes publicadas en medios de comunicación y lecturas de sensores).

### 3. Veracidad:

IBM presentó esta característica de Big Data, asociándola a la falta de confianza inherente en muchas fuentes de datos estructurados, así como no estructurados. Akerkar (2014) y Zicari (2014) refieren la veracidad a cómo hacer frente a los sesgos, dudas, imprecisiones, ambigüedad y evidencia extraviada en los datos.

### 4. Velocidad:

Esto se aplica principalmente a los conjuntos de datos que se generan a través de grandes redes complejas, incluidos los datos generados por la proliferación

---

<sup>31</sup> Uthayasankar, S., Muhammad, M., Zahir, I., & Vishanth, W.. (2016, agosto 10). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 263–286, p.7.

dispositivos digitales. Los datos procedentes de dispositivos móviles y que fluyen a través de aplicaciones móviles o el uso de tarjetas de tiendas (por ejemplo, la tarjeta de Sainsbury para recoger puntos de néctar) genera información que se puede utilizar al tiempo en que se produce, como para ofertar productos personalizados según comportamientos, patrones de compra o la ubicación espacial. (Gandomi & Haider, 2015).

5. Variabilidad:

La variabilidad no debe confundirse con la variedad, ya que variabilidad hace referencia a los diferentes significados que puede tomar un dato en el tiempo y no a los distintos tipos de datos de los que describe la variedad. Variabilidad también está relacionado con el análisis de sentimientos (sentiment analysis). Por ejemplo, en tweets una misma palabra puede tener un significado totalmente diferente. Para realizar un análisis de sentimiento adecuado, los defensores afirman que los algoritmos necesitan ser capaces de entender el contexto y ser capaces de descifrar el significado exacto de una palabra en ese contexto (Zhang, Hu et al., 2015).

6. Visualización:

Consiste en representar información clave y el conocimiento de una manera más intuitiva y efectiva mediante el uso de diferentes formatos visuales tal como en una representación pictórica o gráfica (Taheri, Zomaya, Siegel, & Tari, 2014).

7. Valor:

Extraer conocimiento de la cantidad de datos estructurados y no estructurados para los usuarios finales. Los investigadores de Big Data consideran el valor de los datos como una característica esencial, ya que en algún lugar dentro de estos existe información valiosa (datos de alto valor), aun si la mayoría de los datos pudiera parecer insignificantes (Zaslavsky, Perera, & Georgakopoulos, 2012).

## **Tipos de datos**

Los Big Data son diferentes de las fuentes de datos tradicionales que almacenan datos estructurados en las bases de datos relacionales. Es frecuente dividir las categorías de datos en dos grandes tipos: estructurados (datos tradicionales) y no estructurados (datos Big Data). Sin embargo, las múltiples fuentes que están generando datos propician a la creación de nuevas categorías dentro de los tipos de datos no estructurados: datos semiestructurados y datos no estructurados propiamente dichos.

- Datos Estructurados:

La mayoría de las fuentes de datos tradicionales son datos estructurados. En estas fuentes, los datos vienen en un formato bien definido y conforman las bases de datos relacionales, las hojas de cálculo y los archivos, fundamentalmente. Los datos estructurados se componen de piezas de información que se conocen de antemano, es decir, vienen en un formato específico, y se producen en un orden específico, facilitando el trabajo con dichos datos. Formatos típicos son: fecha de nacimiento (DD, MM, AA); documento nacional de identidad o pasaporte (por ejemplo, 8 dígitos y una letra); número de la cuenta corriente en un banco (20 dígitos), entre otros.

- **Datos Semiestructurados:**

Los datos semiestructurados tienen un flujo lógico y un formato que puede ser definido, pero no es fácil su comprensión para el usuario. Son datos que no tienen formatos fijos, pero contienen etiquetas y otros marcadores que permiten separar los elementos. La lectura de datos semiestructurados requiere el uso de reglas complejas que determinan cómo proceder después de la lectura de cada pieza de información. Ejemplos típicos de datos semiestructurados son: el texto de etiquetas de lenguajes XML y HTML y los registros de web logs de las conexiones a Internet. Un web log se compone de diferentes piezas de información, cada una de las cuales sirve para un propósito específico.

- **Datos No Estructurados:**

Los datos no estructurados son datos sin tipos predefinidos. Se almacenan como “documentos” u “objetos” sin estructura uniforme, y se tiene poco o ningún control sobre ellos. Por ejemplo, las imágenes se clasifican por su resolución en píxeles y no tienen campos fijos. Ejemplos típicos de datos no estructurados son: audio, video, fotografías, documentos impresos, cartas, hojas electrónicas, imágenes digitales, formularios especiales, mensajes de correo electrónico y texto, mensajes instantáneos SMS, artículos, libros, mensajes de mensajería instantánea tipo WhatsApp, Une, Joyn, Viber, Une, Wechat, Spotbros. Al menos, el 80% de la información de las organizaciones no reside en las bases de datos relacionales o archivos de datos, sino que se encuentran esparcidos a lo largo y ancho de la organización.

## **Plataformas**

Sin duda, los datos más difíciles de dominar por los analistas son los datos no estructurados, pero su continuo crecimiento ha provocado el nacimiento de frameworks y plataformas para su manipulación como es el caso de Hadoop, Spark y NoSQL.

### **1. Hadoop**

Es un framework de software que soporta aplicaciones distribuidas bajo una licencia libre. Permite a las aplicaciones trabajar con miles de nodos y petabytes de datos. Hadoop se inspiró en los documentos Google para MapReduce y Google File System (GFS). Hadoop es un proyecto de alto nivel Apache que está siendo construido y usado por una comunidad global de contribuyentes, mediante el lenguaje de programación Java. Yahoo ha sido el mayor contribuyente al proyecto, y usa Hadoop extensivamente en su negocio.

Al hablar de Hadoop es inevitable hablar del tema de MapReduce, que es considerado el corazón de este, al ser su entorno de programación.

MapReduce es un modelo de programación para dar soporte a la computación paralela sobre grandes colecciones de datos en grupos de computadoras. Su nombre está inspirado en los nombres de dos importantes métodos, macros o funciones en programación funcional: Map y Reduce. Su desarrollo fue liderado inicialmente por Yahoo y actualmente lo realiza el proyecto Apache. Se han escrito

implementaciones de bibliotecas de MapReduce en diversos lenguajes de programación como C++, Java y Python.

2. Spark

Apache Spark es un motor de procesamiento rápido y en memoria que permite ejecutar eficientemente grandes cargas de datos que requieran procesamiento continuo de datos (data streaming), machine learning o procesamiento SQL.

Spark puede integrarse con Hadoop y añadir nuevas funcionalidades, puede funcionar con muchos otros productos de Big Data como: CassandraDB, Google Big Query, almacenamiento de Amazon S3, Elastic Search, etc. O puede correr sobre cluster independiente.

3. Sistema de bases de datos NoSQL: también llamadas No Solo SQL, son un enfoque hacia la gestión de datos y el diseño de base de datos que es útil para grandes conjuntos de datos distribuidos. Abarca una amplia gama de tecnologías y arquitecturas, busca resolver los problemas de escalabilidad y rendimiento de Big Data que las bases de datos relacionales no fueron diseñadas para abordar. Es especialmente útil cuando una empresa necesita acceder y analizar grandes cantidades de datos no estructurados o datos que se almacenan de forma remota en varios servidores virtuales en la nube.

## Métodos de análisis de Big Data

Dado que Big Data puede mejorar la toma de decisiones e incrementar el rendimiento en las organizaciones, es necesario recurrir a métodos analíticos usados para extraer sentido a los datos. La Universidad pública de London, Brunel University, clasifica los métodos analíticos para Big Data en función de la finalidad que se busca obtener del análisis, sea este: conocer la información que se tiene, tener mejor visión, tomar decisiones y tomar acciones.<sup>32</sup>

1. Analítica descriptiva

Hace un análisis de los datos e información para definir el estado actual de la situación del negocio en la que los acontecimientos y los patrones ya son evidentes, produciendo reportes estándares, reportes ad hoc y alertas (Joseph & Johnson, 2013);

2. Analítica inquisitiva

Consiste en probar datos para aprobar o rechazar propuestas de negocios. Por ejemplo, análisis drill downs, análisis estadísticos, análisis de factores (Bihani & Patil, 2014);

3. Analítica predictiva

Se ocupan de la predicción y el modelo estadístico para determinar las posibilidades futuras (Waller & Fawcett, 2013);

---

<sup>32</sup> Sivarajah, U., Kamal, M., Irani, Z. & Weerakkody, V.. (2016, Agosto 10). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 286. 2017, abril 14, De n.d Base de datos.

4. Analítica prescriptiva

Se trata de pruebas aleatorias y de optimización para evaluar cómo los negocios mejoran sus niveles de servicio mientras disminuyen sus gastos (Joseph & Johnson, 2013);

5. Analítica preventiva

Consiste en tener la capacidad de tomar acciones preventivas sobre eventos que puedan influir negativamente en el desempeño de las organizaciones. Por ejemplo: identificar posibles riesgos y recomendar estrategias de mitigación adelantadas en el tiempo (Szongott, Henne y von Voigt, 2012).

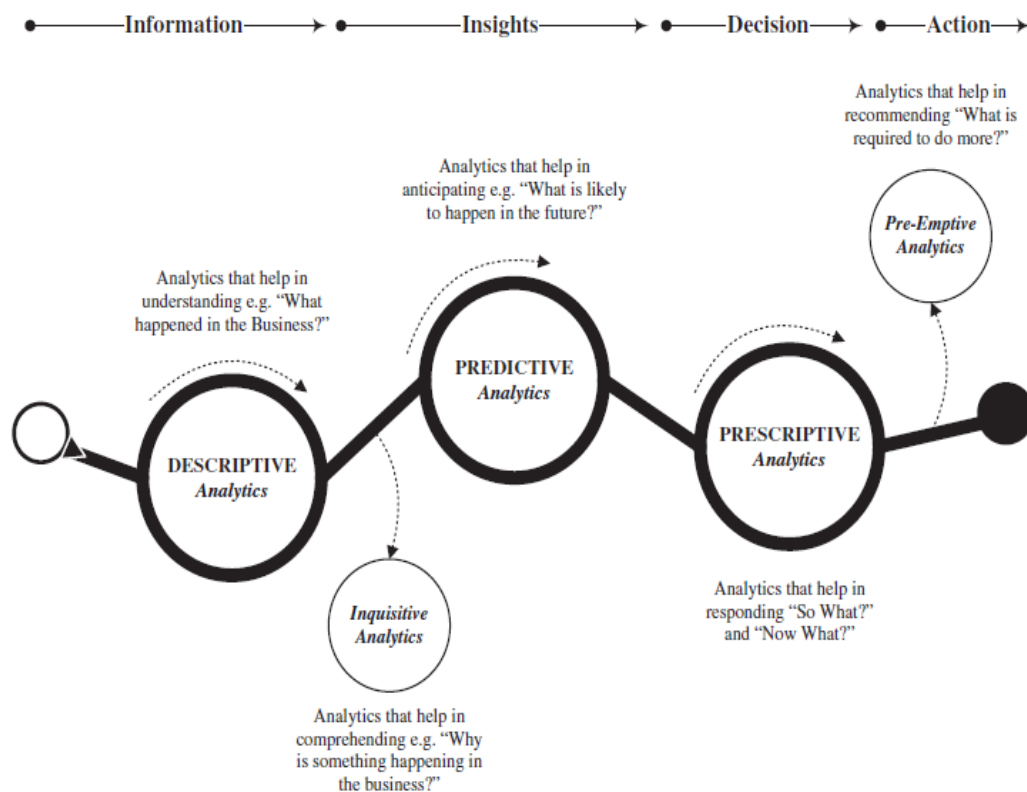


Fig. 2. Classification of types of big data analytical methods.

Figura 5. Classification of types of Big Data analytical methods. En *Critical analysis of Big Data challenges and analytical methods* (p.266). por Journal of Business Research, 2016

## ¿Quién usa Big Data en el mundo?

IBM publicó en el 2016 un artículo titulado "How 10 industries are using Big Data to win big" donde habla de las 10 industrias que están aplicando Big Data en sus actividades, a saber:

1. Telecomunicaciones,
2. Manufactura,
3. Comercio (retail),
4. Fitness,

5. Seguros,
6. Bancos,
7. Finanzas,
8. Gobiernos,
9. Salud y
10. Aerolíneas.

## DATA IS TRANSFORMING MULTIPLE INDUSTRIES



Figura 6. *Data is transforming multiples industries.* En *How 10 industries are using Big Data to win big.* por M.,Belfiore, 2016. Recuperado de <https://www.ibm.com/blogs/watson/2016/07/10-industries-using-big-data-win-big/>

### 2.3 Marco Conceptual

**Análisis de Datos:** Es la ciencia que examina datos en bruto con el propósito de sacar conclusiones sobre la información. El análisis de datos es usado en varias industrias para permitir que las compañías y las organizaciones tomen mejores decisiones empresariales y también es usado en las ciencias para verificar o reprobar modelos o teorías existentes. El análisis de datos se distingue de la extracción de datos por su alcance, su propósito y su enfoque sobre el análisis. Los extractores de datos clasifican inmensos conjuntos de datos usando software sofisticado para identificar patrones no descubiertos y establecer relaciones escondidas. El análisis de datos se centra en la inferencia, el proceso de derivar una conclusión basándose solamente en lo que conoce el investigador.<sup>33</sup>

<sup>33</sup> Rouse, M.. (2012). *Análisis de datos*. abril 14, 2017, de Data Profiling Sitio web: <http://searchdatacenter.techtarget.com/es/definicion/Analisis-de-Datos>

**Análisis de Trayectoria o Path Analysis:** El análisis de trayectoria es una forma de análisis estadístico de regresión múltiple utilizado para evaluar modelos causales examinando las relaciones entre una variable dependiente y dos o más variables independientes. Usando este método uno puede estimar tanto la magnitud como la importancia de las conexiones causales entre las variables<sup>34</sup>

**Big Data:** Activos de información de gran volumen, alta velocidad y / o de alta variedad que exigen formas innovadoras y rentables de procesamiento de información que permiten un mejor conocimiento, toma de decisiones y automatización de procesos.<sup>35</sup>

**Clustering:** Es un algoritmo de agrupamiento en donde datos con las mismas características son juntados para su posterior análisis.

**Data cleansing:** También conocido como data scrubbing (borrado de datos), es el proceso de modificar o eliminar de una base de datos aquellos datos incorrectos, incompletos, formateados incorrectamente o duplicados.<sup>36</sup>

**Data mart:** Es un subconjunto de los datos del data warehouse cuyo objetivo es responder a un determinado análisis, función o necesidad, con una población de usuarios específica.

**Data profiling:** También conocida como data archeology (arqueología de los datos). Es el análisis estadístico y la evaluación de los valores dentro de un conjunto de datos para determinar la coherencia, singularidad y lógica de los mismos.<sup>37</sup>

**Data Source:** Fuente de datos se define como los diferentes medios de los cuales se pueden obtener datos estas pueden ser bases de datos operacionales, datos externos o Información de los datos ya existentes en un data warehouse; así como también bases de datos relacionales o cualquier otra estructura de datos que soporté la línea de aplicaciones de negocio.

**Data streaming:** Proceso en el que los grandes flujos de datos se procesan en tiempo real con el objetivo de extraer conocimiento y tendencias útiles. Un flujo continuo de datos no estructurados se envía para su análisis en memoria antes de ser almacenarla en disco. La velocidad es lo más importante en el data streaming, el valor de los datos, si no se procesa rápidamente, disminuye con el tiempo.<sup>38</sup>

**Datawarehouse:** Repositorio de datos que proporciona una visión global, común e integrada de los datos de la organización, con las propiedades siguientes: estable, coherente, fiable y con información histórica.

---

<sup>34</sup> Crossman, A..(2017). *Understanding Path Analysis*. abril 15,2017, de Thought Co. Sitio web: <https://www.thoughtco.com/path-analysis-3026444>

<sup>35</sup> Big Data.(n.d).Gartner IT Glossary.Recuperado de <http://www.gartner.com/it-glossary/big-data/>

<sup>36</sup> Rouse, M..(2010). *data scrubbing (data cleansing)*. abril 14,2017, de TechTarget Sitio web: <http://searchdatamanagement.techtarget.com/definition/data-scrubbing>

<sup>37</sup> Rouse, M.. (2014). *Data Profiling*. abril 14,2017, de TechTarget Sitio web: <http://searchdatamanagement.techtarget.com/definition/data-profiling>

<sup>38</sup> Big data streaming.(n.d).Techopedia.Recuperado de <https://www.techopedia.com/definition/31752/big-data-streaming>

**ETL:** Tecnología de integración de datos basada en la consolidación de datos que se usa tradicionalmente para alimentar data warehouse, data mart, staging área y ODS. Usualmente se combina con otras técnicas de consolidación de datos.

**Inteligencia de negocios:** Conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de información que permite tomar mejores decisiones a los usuarios de una organización.

**Metadatos:** Datos estructurados y codificados que describen características de instancias; aportan informaciones para ayudar a identificar, descubrir, valorar y administrar las instancias descritas.

**ODS:** Es un tipo de almacén de datos que proporciona sólo los últimos valores de los datos y no su historial; además, generalmente admite un pequeño desfase o retraso sobre los datos operacionales.

**OLAP:** Proceso analítico en línea, método ágil y flexible para organizar datos, especialmente metadatos, sobre un objeto o jerarquía de objetos como en un sistema u organización multidimensional, y cuyo objetivo es recuperar y manipular datos y combinaciones de los mismos a través de consultas o incluso informes.

**Reportes ad hoc:** Reportes que son creados por usuarios finales en tiempo real, con los que se pretende dar respuesta a un objetivo o una pregunta específica. Se utilizan herramientas o asistentes que ofrecen la posibilidad a los usuarios de elegir las funciones de los informes y establecer el alcance, la variabilidad y otros aspectos. También pueden proporcionar una presentación visual personalizada con gráficos.<sup>39</sup>

**RFID (radio frequency identification):** Identificación por radiofrecuencia es una tecnología que incorpora el uso de acoplamiento electromagnético o electrostático en la porción de radiofrecuencia (RF) del espectro electromagnético para identificar de forma única un objeto, animal o persona. RFID está entrando en uso creciente en la industria como una alternativa al código de barras. La ventaja de la RFID es que no requiere el contacto directo.<sup>40</sup>

**Sentiment Analysis:** Uso de métodos estadísticas, la programación neuro-lingüística, o machine learning para extraer, identificar o caracterizar el contenido de sentimiento de una unidad de texto.<sup>41</sup>

**Sistemas transaccionales:** Es un tipo de sistema diseñado para soportar las operaciones del negocio, generalmente es usado para almacenar, recolectar y modificar cualquier tipo de información.

---

<sup>39</sup> Ad hoc reporting. (n.d).Techopedia.Recuperado de <https://www.techopedia.com/definition/30294/ad-hoc-reporting>

<sup>40</sup> Rouse, M.. (2007). *RFID (radio frequency identification)*. mayo 4,2017, de TechTarget Sitio web: <http://internetofthingsagenda.techtarget.com/definition/RFID-radio-frequency-identification>

<sup>41</sup> n.d. (n.d). *Introduction to Sentiment Analysis*. abril 15,2017, de European Masters Program in Language and Communication Technologies (LCT) i Sitio web: <https://lct-master.org/files/MullenSentimentCourseSlides.pdf>



## CAPÍTULO III: METODOLOGÍA

---

### 3.1 Tipo de investigación

El tipo de investigación se puede definir por medio de distintos criterios; uno de ellos es por el objetivo de la investigación es decir cuál es la finalidad que se busca al llevar a cabo la investigación; esta puede ser investigación básica o investigación aplicada.

#### INVESTIGACIÓN BÁSICA

También recibe el nombre de investigación pura, teórica, dogmática y fundamental. Se caracteriza porque parte de un marco teórico y permanece en él; la finalidad radica en formular nuevas teorías o modificar las existentes, en incrementar los conocimientos científicos o filosóficos, pero sin contrastarlos con ningún aspecto práctico.

#### INVESTIGACIÓN APLICADA

El proyecto o investigación busca la aplicación, el uso y generar posibles beneficios de los conocimientos resultantes de la misma, es decir, busca conocer para actuar, le interesa la aplicación sobre la realidad antes que el mero desarrollo de teorías generales, está dirigida hacia un fin u objetivo práctico determinado.<sup>42</sup>

Al analizar estos conceptos, se identifica que debido a la finalidad de la investigación de formar un curso de especialización a partir de la investigación de los conocimientos, habilidades y técnicas presentes en un profesional de inteligencia de negocios y analítica de Big Data para poder ser impartidos a estudiantes de pregrado de ingeniería de sistemas, se establece que la investigación es **aplicada**.

Otro aspecto importante de la clasificación de la investigación es según el enfoque de investigación que se realizará; este puede ser cuantitativo, cualitativo o mixto.

Los métodos mixtos representan un conjunto de procesos sistemáticos, empíricos y críticos de investigación e implican la recolección y el análisis de datos cuantitativos y cualitativos, así como su integración y discusión conjunta, para realizar inferencias producto de toda la información recabada (meta inferencias) y lograr un mayor entendimiento del fenómeno bajo estudio (Hernández Sampieri y Mendoza, 2008).

La definición anterior demuestra que, para ésta, investigación a realizar, el hecho de necesitar datos tanto cualitativos como cuantitativos hace que el método mixto sea el que mejor cumple los requerimientos; esto se fundamenta gracias a las siguientes preguntas que debe responder el investigador para elegir el diseño mixto:

---

<sup>42</sup> Enciclopedia de clasificaciones. (2016). Tipos de investigación. Marzo, 10, 2017, de Enciclopedia de clasificaciones Sitio web: <http://www.tiposde.org/general/484-tipos-de-investigacion/>

1. ¿Qué clase de datos tienen prioridad los cuantitativos, los cualitativos o ambos por igual?

#### **Cualitativos.**

El tema de inteligencia de negocios y analítica de Big Data se encuentra en una etapa de madurez a nivel internacional, y en nuestro ambiente, en proceso de crecimiento y posicionamiento en el área laboral, por tanto, prevalecerá toda la información disponible del avance de la inteligencia de negocios y analítica de Big Data, que es resultado de estudios realizados por expertos en el área. No siendo menos importante, la experiencia actual que ha tenido el tema en nuestro ambiente. Por tanto, los datos cuantitativos resultantes de la investigación a realizar, asistirán en la interpretación de lo cualitativo.

2. ¿Qué resulta más apropiado para el estudio en particular: recolectar los datos cuantitativos y cualitativos de manera simultánea (al mismo tiempo) o secuencial (un tipo de datos primero y luego el otro)?

#### **Concurrente**

Se aplican ambos métodos (CUAN y CUAL) de manera simultánea, los datos cuantitativos y cualitativos se recolectan y analizan más o menos en el mismo tiempo. Desde luego, los datos cualitativos requieren de mayor tiempo para su obtención y análisis.

Los diseños concurrentes implican cuatro condiciones (Onwuegbuzie y Johnson, 2008):

- i. Se recaban en paralelo y de forma separada datos cuantitativos y cualitativos.
- ii. Ni el análisis de los datos cuantitativos ni el análisis de los datos cualitativos se construye sobre la base del otro análisis.
- iii. Los resultados de ambos tipos de análisis no son consolidados en la fase de interpretación de los datos de cada método, sino hasta que ambos conjuntos de datos han sido recolectados y analizados de manera separada se lleva a cabo la consolidación.
- iv. Después de la recolección e interpretación de los datos de los componentes CUAN y CUAL, se efectúa una o varias “meta inferencias” que integran las inferencias y conclusiones de los datos y resultados cuantitativos y cualitativos realizadas de manera independiente.

Se propone el **Diseño anidado o incrustado concurrente de modelo dominante (DIAC)**

El diseño anidado concurrente colecta simultáneamente datos cuantitativos y cualitativos teniendo en consideración que un método predominante guía el proyecto (cualitativo). El método que posee menor prioridad es anidado o incrustado dentro del que se considera central. Tal incrustación puede significar que el método

secundario responda a diferentes preguntas de investigación respecto al método primario.

Los datos recolectados por ambos métodos son comparados y/o mezclados en la fase de análisis.

Este diseño suele proporcionar una visión más amplia del fenómeno estudiado que si usáramos un solo método.

3. ¿Cuál es el propósito central de la integración de los datos cuantitativos y cualitativos? Por ejemplo: triangulación, complementación, exploración o explicación.

### **Complementación**

Obtener una visión más comprensiva sobre el planteamiento si se emplean ambos métodos.

4. ¿En qué parte del proceso, fase o nivel es más conveniente que se inicie y desarrolle la estrategia mixta? Por ejemplo: desde y/o durante el planteamiento del problema, en el diseño de investigación, recolección de los datos, análisis de los datos, interpretación de resultados o elaboración del reporte de resultados.

Desde el diseño de la investigación ya que es desde esta etapa que se aprovechan las bondades del enfoque mixto pues se hará uso de instrumentos cuantitativos (encuestas), y cualitativos (entrevistas) para obtener datos que en la etapa de análisis ayudarán a profundizar en la construcción del perfil de conocimientos, habilidades y técnicas al proporcionar una visión más amplia del fenómeno estudiado que si se usará un solo método.<sup>43</sup>

## **3.2 Metodología para la recolección de datos**

### **3.2.1 Técnicas e instrumentos para la recolección de datos**

#### **Técnicas**

<b>Técnica</b>	<b>Objetivo</b>	<b>Personas dirigidas</b>	<b>¿Qué se obtiene?</b>
Encuesta	Conocer la experiencia laboral de profesionales o no en el desarrollo de actividades relacionadas con inteligencia de negocios y análisis de datos	Profesionales o no con experiencia laboral en el área de inteligencia de negocios y analítica de datos.	-Conocimientos técnicos -Herramientas de uso -Habilidades -Deficiencias técnicas

<sup>43</sup> Hernández Sampieri, R., Fernández Collado, C. and Baptista Lucio, P. (2010). *Metodología de la investigación*. 1st ed. México: McGraw-Hill.p572

Entrevista	Determinar cuáles son los conocimientos fundamentales que se imparten en el área de inteligencia de negocios y analítica de datos para establecer las áreas de formación necesarias del perfil académico.	Docentes que imparten cursos de pregrado o postgrado y especializaciones en los temas de inteligencia de negocios y analítica de datos.	-Conocimientos impartidos en los cursos. -Duración de los cursos. -Metodologías de enseñanza. -Valores
Análisis de documentos: bolsas de trabajos nacionales e internacionales.	Comparar y consolidar los requerimientos solicitados en las ofertas de trabajo para plazas relacionadas con inteligencia de negocios y analítica de datos	Bolsas de trabajo nacionales  Bolsas de trabajo internacionales	-Conocimientos técnicos requeridos  -Habilidades y valores.
Análisis de documentos: contenido de cursos en línea	Consolidar los contenidos que imparten universidades o entidades de renombre en el área de inteligencia de negocios y analítica de datos	Cursos en línea que ofrecen universidades Nacionales e internacionales.	-Conocimientos técnicos -Herramientas de uso -Metodologías de creación de cursos académicos.

## Instrumentos

### Diseño de instrumento: Encuesta

#### Estimado encuestado

Somos estudiantes de la Universidad de El Salvador optando al título de Ingeniero de Sistemas Informáticos, estamos trabajando en una investigación que servirá para elaborar el trabajo de graduación “Propuesta de formación profesional y académica para el perfil de especialización en el área de inteligencia de negocios y analítica de datos para la carrera de pregrado de Ingeniería de Sistemas informáticos de la Universidad de El Salvador”

Esta encuesta está destinada a profesionales o estudiantes de ingenierías o licenciaturas de sistemas informáticos, ciencias de la computación o carreras afines que tenga experiencia en el desarrollo de actividades en inteligencia de negocios, análisis de datos y/o Big Data. Solicitamos de su valioso tiempo para contestar unas preguntas.

Las opiniones de todos los encuestados serán de mucho aporte para los resultados de esta investigación, pero nunca se comunicarán datos individuales.

**Objetivo:** Conocer la experiencia de las personas que ejercen actividades relacionadas con la inteligencia de negocios, analítica de datos y/o Big Data.

**Indicaciones:** Conteste las siguientes interrogantes según el criterio que usted crea conveniente. Si alguna respuesta no se acomoda a su criterio, elija la que más se asemeje.

### Preguntas generales

**Objetivo:** Conocer el área de cada encuestado, para que a partir de esta segmentar las preguntas siguientes

1. ¿En qué área se desempeña actualmente?

Inteligencia de negocios

Análisis de datos

Big Data

**Objetivo:** Obtener una lista de las habilidades para realizar el perfil académico.

2. En base a su experiencia, mencione ¿cuáles son las habilidades que un profesional en su área debe poseer?

**Objetivo:** Obtener una lista de los valores profesionales para realizar el perfil académico

3. En base a su experiencia, menciona que valores profesionales debe poseer un profesional de su área

### Analítica de datos

**Objetivo:** Medir el porcentaje de participación de los diferentes tipos de análisis, para inferir las necesidades actuales y futuras en el área laboral en cuanto a análisis de datos se trate.

4. ¿Qué tipo de análisis realiza en su cargo?
  - a. Diagnóstico: Un vistazo al desempeño del pasado para determinar qué sucedió y por qué.
  - b. Descriptivo: Lo que está sucediendo ahora basado en los datos entrantes.
  - c. Prescriptivo: Este tipo de análisis revela qué acciones deben tomarse.
  - d. Predictivo: Un análisis de los posibles escenarios de lo que podría suceder.

**Objetivo:** Obtener una lista de conocimientos técnicos para realizar el perfil académico.

5. En base a su experiencia, mencione que conocimientos técnicos o base teórica ha requerido para llevar a cabo análisis de datos

**Objetivo:** Conocer qué herramientas tiene mayor uso, para considerarla en el desarrollo el curso académico.

6. ¿Qué herramientas utiliza/ha utilizado para el análisis?
- a) SAS
  - b) IBM
  - c) KNIME
  - d) RapidMiner
  - e) Otro

**Objetivo:** Conocer los cursos que los encuestados han recibido, para su posterior indagación y obtención de temas de Big Data a considerar en la creación del curso académico y profesional.

7. En cuanto al tema de Big Data, ¿ha recibido algún curso o capacitación?

Si la Respuesta es **Sí**, responder a lo siguiente:

- 7.1 Modalidad: ¿Presencial o en línea?
- 7.2 ¿Cuál fue la duración (horas, meses, etc.)?
- 7.3 Menciona que cursos recibió y quién lo impartió

Si la respuesta es **No**, responder a lo siguiente:

- 7.4 ¿En qué temas relacionadas con analítica de datos y Big Data considera debería recibir curso, capacitación o consultoría profesional en función de sus metas profesionales o los objetivos de la compañía a la que labora?

**Objetivo:** Obtener los conocimientos técnicos relacionados con Big Data para su incorporación en el curso académico a desarrollar.

8. En base a su experiencia o conocimiento teórico de Big Data, ¿Qué temas considera un analista de Big Data debe manejar?

**Objetivo:** Identificar los temas complejos relacionados al análisis de datos y Big Data para evaluar y determinar su incorporación al curso a crear en función de los alcances del estudio.

9. En base a su experiencia, ¿Qué temas relacionados a la analítica de datos y Big Data considera que requieren de mayor estudio y profundidad debido a su alto nivel de complejidad?

**Objetivo:** Determinar necesidades futuras en el área de análisis de datos y Big Data para su incorporación en el curso.

10. ¿Cuál considera son las nuevas tendencias en el área de análisis de datos y Big Data?

## Inteligencia de negocios

**Objetivo:** Obtener una lista de conocimientos técnicos para realizar el perfil académico

11. En base a su experiencia, mencione que conocimientos técnicos ha requerido para llevar a cabo inteligencia de negocios

**Objetivo:** Conocer qué herramientas tiene mayor uso, para considerarla en el desarrollo el curso académico.

12. ¿Qué herramientas utiliza/ha utilizado para realización de reportes?
- Tableau
  - Qlik
  - Microsoft Power BI
  - Pentaho
  - Microstrategy
  - Otro

**Objetivo:** Conocer los cursos que los encuestados han recibido, para su posterior indagación y obtención de temas de inteligencia de negocios a considerar en la creación del curso académico y profesional.

13. En cuanto al tema de inteligencia de negocios, ¿ha recibido algún curso o capacitación?

Si la Respuesta es **Sí**, responder a lo siguiente:

- 13.1 Modalidad: ¿Presencial o en línea?
- 13.2 ¿Cuál fue la duración (horas, meses, etc.)?
- 13.3 Menciona que cursos recibió y quién lo impartió

Si la respuesta es **No**, responder a lo siguiente:

- 13.4 ¿En qué temas relacionados con inteligencia de negocios considera debería recibir curso, capacitación o consultoría profesional en función de sus metas profesionales o los objetivos de la compañía?

**Objetivo:** Determinar necesidades futuras en el área de inteligencia de negocios data para su incorporación en el curso.

14. ¿Cuál considera son las nuevas tendencias en el área de inteligencia de negocios?

## Diseño de instrumento: Entrevista

### Entrevista para determinar los conocimientos que se imparten en los cursos de especialización y postgrado

**Objetivo:** Determinar los conocimientos fundamentales que se imparten en cursos de inteligencia de negocios y analítica de datos para establecer las áreas de formación necesarias del perfil académico.

La elaboración de la entrevista va orientada para conocer de las experiencias de las personas que se dedican a impartir cursos de inteligencia de negocios y analítica de Big Data, con el fin de poder crear un perfil académico que pueda englobar todos los conocimientos necesarios para desempeñarse en estos rubros.

**Objetivo:** Conocer el área en la que el entrevistado imparte el curso ya sea inteligencia de negocios, analítica de Big Data o ambas para identificar la más predominante.

1. ¿Cuál es el área en la que usted imparte?

**Objetivo:** Determinar cuál es la duración promedio de los módulos del curso para obtener estimaciones que aporten al diseño del curso.

2. ¿Cuánto tiempo es la duración de la materia o módulo?

**Objetivo:** Identificar el material de apoyo o las referencias que tomó el entrevistado para definir los temas del curso y así poder consultarlas.

3. ¿Cómo o en que se basó para determinar los temas del curso?

**Objetivo:** Determinar cuáles son los conocimientos que debe tener el estudiante para ser apto a recibir cursos.

4. ¿Cuáles son los prerrequisitos para una persona que se interese por llevar el curso o especialización?

**Objetivo:** Conocer la percepción del entrevistado con respecto al crecimiento del área.

5. ¿Ha crecido el número de personas que se decantan por llevar estos cursos?

**Objetivo:** Determinar qué temas presentan un mayor reto para el estudiante.

6. ¿Cuál es el contenido o tema en esta área que usted considera más complejo?

**Objetivo:** Conocer el grado de actualidad que tiene los contenidos del curso.

7. ¿El contenido del curso se mantiene desde la primera vez que lo impartió?



**Objetivo:** Determinar las competencias que tendrá el alumno al finalizar la realización del curso.

8. ¿Cuáles son las habilidades esperadas que el alumno tendrá al finalizar el curso?

**Objetivo:** Determinar aquellos temas más importantes identificados por los profesionales.

9. ¿Cuáles son los temas en los alumnos muestran más interés al momento de impartir su cátedra?

**Objetivo:** Descubrir cuáles son los temas que requiere más atención al tener más demanda en el campo laboral.

10. ¿Qué temas son los que más se ponen en práctica al momento de desempeñar un cargo en esta área?

**Objetivo:** Conocer la forma en que se imparten los cursos para determinar estrategias de enseñanza.

11. ¿Cuál es la metodología que utiliza al momento de impartir los cursos, depende ésta de la cantidad de participantes que asisten?

**Objetivo:** Conocer el grado de experiencia de la persona en el curso que imparte.

12. ¿Cuántas veces ha impartido el curso?

### 3.2.2 Población y muestra

**Técnica:** Encuesta

**Unidad de análisis:**

Profesionales o no que se desempeñan en cargos relacionados a inteligencia de negocios o análisis de datos.

**Población:**

Contenido: Profesionales o no que actualmente desempeñen los siguientes cargos: desarrolladores de inteligencia de negocios, analistas de inteligencia de negocios, consultores de inteligencia de negocios, coordinadores de proyecto de inteligencia de negocios, analistas de datos y analista de Big Data en empresas de cualquier naturaleza y tamaño.

**Lugar:** Área metropolitana

**Tiempo:** año 2017

### **Muestra:**

En este punto surge la siguiente interrogante ¿Cuál es el tamaño de la población definida?

1. El Instituto del Seguro social en su Anuario Estadístico 2016 totaliza los trabajadores cotizantes según actividad económica y sector público y privado, sin embargo no es posible conocer la cantidad de personas que se dediquen al área en estudio<sup>44</sup>.
2. El MINED en su informe Resultados de la información Estadística de Instituciones de Educación Superior 2015 publicó la cantidad de graduados de todas las universidades para todas las carreras, sin embargo no es posible determinar de estos datos la cantidad de profesionales que se dediquen al área de estudio.<sup>45</sup>
3. El cálculo probabilístico de muestras para poblaciones desconocidas o infinitas (Ver anexo 3) indica que la muestra es de 385 personas. Trabajar con muestras grandes o que no puedan ser alcanzadas afectan la confiabilidad de los resultados obtenidos. Pentti Routio, docente de la Universidad de Arte y Diseño en Helsinki dice lo siguiente:

“Sucedee a menudo que algunos casos en la muestra resultan infructuosos porque no pueden ser alcanzados, o las medidas fallan, o los entrevistados rechazan cooperar etc. El método normal es entonces sobredimensionar la muestra levemente, y después se olvida simplemente los casos que fallan.”<sup>46</sup>

Ante estas consideraciones, y teniendo en mente el objetivo del instrumento, la unidad de análisis y la calidad de los resultados, se determina que la muestra a utilizar para este instrumento se considera intencional, dirigido o por conveniencia.<sup>47</sup> En esta técnica de muestreo, el investigador selecciona a la muestra en función del conocimiento que tiene de ella, es decir en función de su representatividad. Por tanto, los investigadores encuestarán a las personas que sepan que cumplen con las características de la población definida, es decir, que estén desempeñando un cargo de los especificados anteriormente.

### **Técnica: Entrevista**

#### **Unidad de análisis:**

Docentes que imparten cursos de pregrado o postgrado y especializaciones en los temas de inteligencia de negocios y análisis de datos.

---

<sup>44</sup> Instituto Salvadoreño del Seguro Social. (2017, abril 01). *Estadísticas del ISSS 2016*. Anuario Estadístico, n.d, 96. 2017, abril 19, De n.d Base de datos.

<sup>45</sup> Ministerio de Educación República de El Salvador. (2016, diciembre 01). *Resultados de la Información Estadística de Instituciones de Educación Superior 2015*. Portal de Transparencia, n.d, 166. 2017, abril 19, De n.d Base de datos.

<sup>46</sup> Routio, P.. (2007). *Arteología, la ciencia de productos y profesiones: Muestreo*. mayo 1,2017, de Universidad de Arte y Diseño Sitio web: <http://www.uiah.fi/projects/metodi/252.htm>

<sup>47</sup> Gidalberto Bonilla. (1993). *Cómo hacer una tesis de investigación con técnicas estadísticas*. El Salvador: UCA Editores.

### **Población:**

Contenido: Docentes o profesionales que tengan experiencia en el campo académico. Que impartan o hayan impartido cursos de pregrado o postgrado en materia de inteligencia de negocios y/o analítica de datos, así como también cursos de especialización en estas áreas

**Lugar:** Área metropolitana.

**Tiempo:** Año 2017.

**Muestra:** Para estimar el número de personas a entrevistar, el modo más sencillo de construir el grupo de informantes es a través del uso de la técnica de bola de nieve<sup>48</sup>. Que consiste en conocer algunos informantes y lograr que ellos presenten a otros. De esta manera se espera que a través del desarrollo de las entrevistas, los informantes puedan referir a otras personas que colaboren con la investigación.

### **3.2.3 Plan de recolección y procesamiento de datos**

La recolección y procesamiento de datos se realizará para los datos obtenidos tanto para la encuesta como para la entrevista.

Con respecto a la entrevista se realizará la transcripción de la misma para luego identificar en un listado las ideas y datos aportadas de la misma, para luego realizar una matriz para vinculación de categorías; dicha matriz es una herramienta que ayudará a encontrar sentido y significado a las relaciones entre los temas<sup>49</sup>

Esta matriz se aplicará a cada pregunta de la entrevista y aquellas preguntas de la encuesta que sean abiertas.

Las categorías a evaluar serán las siguientes: áreas de conocimientos, herramientas más utilizadas, técnicas aplicadas y metodología de enseñanza-, esta última categoría sólo se evaluará en los datos recolectados de la entrevista

### **Modelo de Matriz para vinculación entre categorías**

Categoría o tema a comparar

Categoría/ Instrumento	Instrumento	Instrumento	Instrumento	Total
Categoría				
Categoría				
Categoría				

<sup>48</sup> P. Balcázar, N. González-Arratia, G. Gurrola & A. Moysén. (2013). Entrevista a profundidad. En Investigación Cualitativa (pp. 57-74). Toluca: UAEM.

<sup>49</sup> Hernández Sampieri, R., Fernández Collado, C. and Baptista Lucio, P. (2010). *Metodología de la investigación*. 1st ed. México: McGraw-Hill. p.464

Para las encuestas, específicamente para las preguntas cerradas se realizará un conteo de frecuencias, cálculo de media (promedio), se utilizará un diagrama de Pareto para encontrar una relación 80-20 y se representarán los datos por medio de gráficos representativos.

Se hará uso de triangulación de datos para comparar lo obtenido por cada uno de los instrumentos junto con los datos obtenidos a partir del análisis de documentos

### **3.3 Metodología del proyecto**

#### **3.3.1 Creación de perfil profesional**

En esta etapa se procede al análisis y consolidación de los datos recolectados en la etapa anterior para la creación del perfil profesional.

El análisis y consolidación de los datos recolectados implica identificar las variables a estudiar (conocimientos técnicos, herramientas, habilidades, valores profesionales) para que a partir de la aplicación de medidas estadísticas como la media y la moda se interprete y defina los aspectos que incorporará el perfil profesional a proponer tanto para inteligencia de negocios como analítica de Big Data.

La propuesta del perfil profesional pasará por un proceso de revisión de contenido, por lo cual, se recurrirá al método Delphi, con el afán de obtener un perfil depurado y definitivo para dar inicio a la creación del curso académico.

Para el perfil profesional se hará uso de una plantilla con la cual identificar las características esperadas en un estudiante al cursar el curso de especialización

### **Caracterización del Perfil del Egresado de la Especialización**

#### **1. Apoyo al perfil de Egreso a partir de la especialización**

En este apartado se pretende describir los aspectos y áreas las cuales el curso de especialización habrá formado y/o reforzado al finalizar con el mismo.

### **Ejes Transversales**

Son los temas que se pueden tratar desde otras disciplinas y en el proceso educativo sirven para dinamizar los procesos ya que ayudan a tener una visión en conjunto al recorrer diferentes asignaturas y temas

#### **1. Caracterización de la Especialización**

Se describirán las características específicas de esta especialización y cuál es el valor agregado de la misma para con el egresado

#### **2. Objeto de la Especialización**

Descripción de los temas o áreas generales en las que profundizará la especialización

- a. Esferas de actuación  
Consiste en los distintos escenarios en los que se podrá desempeñar un profesional con la especialización
3. Modelo del profesional con la especialización  
Descripción Características, habilidades y valores esperados en el profesional al momento de desempeñarse en el ámbito laboral.
4. Objetivos de la Especialización  
Los propósitos que se alcanzarán con el curso de especialización
5. Sistema de Conocimientos  
Cuadro comparativo en el cual se identificará las actividades que pondrán en práctica los conocimientos adquiridos

Acciones generalizadoras	Tareas terminales que ejecuta en la empresa	Conocimientos a trabajar

### 3.3.2 Creación de curso académico

A partir de la elaboración de los perfiles para el área de inteligencia de negocios y analítica de Big Data y con base a las metodologías de enseñanza propuestas por los expertos en la entrevista se procederá a la construcción del curso de especialización el cual pretende satisfacer los requerimientos expuestos en el perfil así como también los futuros requerimientos demandados

Como punto inicial, se utilizará una matriz de análisis de competencias para las áreas de estudio. Esta matriz resultante será el input para la creación de los programas de las asignaturas a desarrollar.

#### **Matriz de análisis de competencias para la carrera de Ingeniería de Sistemas Informáticos.**

La siguiente matriz será detallada para cada una de las siguientes áreas: inteligencia de negocios, análisis de datos y Big Data. En cada área se especificará las competencias, los prerrequisitos y se presentará un cuadro resumen.

#### **Competencias**

Se listan las capacidades que un profesional en el área debe emplear para responder a situaciones y resolver problemas aplicando para ello los conocimientos apropiados (saber) y las habilidades y destrezas (saber hacer). Las competencias podrán clasificarse en:

1. Metodológicas: corresponden a los niveles precisos de conocimientos y de información requeridos para desarrollar una o más tareas.
2. Técnicas: se refieren a las aplicaciones prácticas precisas para ejecutar una o más tareas.
3. Competencias Sociales: responden a la integración fluida y positiva de la persona a grupos de trabajo y a su respuesta al desafío social que ello implica, desde la perspectiva laboral.
4. Individuales: tiene relación con aspectos como la responsabilidad, la disciplina, la puntualidad, etc.

### **Prerrequisitos del área**

Se listan aquellos conocimientos que son la base y de carácter obligatorio para desarrollar las competencias de cada área. Dado que son prerrequisitos, no implicarán inversión de tiempo para su desarrollo, pues se espera que la persona los conozca de fuentes previas.

### **Resumen de Competencias, conocimientos, habilidades y valores**

Cuadro resumen en donde se detalla visualmente las habilidades, los conocimientos y valores que se requieren para desarrollar una competencia.

<b>Competencias</b>	<b>Habilidades</b>	<b>Conocimientos</b>	<b>Valores</b>

Gracias a este análisis, se identificará las temáticas que conformarán a cada una de las materias en las que estará dividido el curso de especialización; para cada materia se establecerá el programa de la asignatura en donde se detallaran todos los aspectos académicos de esta.

A continuación, se muestra un ejemplo de plantilla de programa de la asignatura con los elementos que la conforman.

UNIVERSIDAD DE EL SALVADOR  
FACULTAD DE INGENIERÍA Y ARQUITECTURA  
ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS

**PROGRAMA DE ASIGNATURA**

**I GENERALIDADES DE LA ASIGNATURA**

Asignatura	El nombre de la asignatura.
Código	El código por el que se identifica la asignatura.
Prerrequisito	Las materias que son prerrequisitos de la asignatura.
Número de horas	Las materias que son prerrequisitos de la asignatura.
Unidades valorativas	Las unidades valorativas de la asignatura.
Ciclo	El número de ciclo en el que se imparte.
Número correlativo	El número correlativo de la asignatura.
Horas teóricas semanales	Las horas teóricas destinadas a impartirse a lo largo de la semana.
Horas prácticas semanales	Las horas destinadas a laboratorios prácticos que refuercen la teoría vista en clase.
Duración del ciclo (semanas)	Duración del ciclo en que se impartirá la asignatura.
Duración horas clase (minutos)	Duración de la clase teórica expresada en minutos.
Plan de estudios	Año del plan de estudios donde se incluye la asignatura.

**II FUNDAMENTACIÓN DE LA ASIGNATURA**

Explica de manera general cuales son los conocimientos y habilidades que el estudiante logrará al completar la asignatura, además se señalan los temas principales a tratar durante todo el curso fundamentando la importancia de los contenidos como tal.

En este apartado también se explica los cuales son los temas que el estudiante ya debe dominar y estar familiarizado para sacar un mayor provecho del curso. Finalmente señala cual es la importancia que aportarán los conocimientos de la asignatura y de qué manera servirán al estudiante alcanzar tales competencias.

**III OBJETIVOS**

En este apartado se colocan los objetivos que se pretenden alcanzar durante el desarrollo de la asignatura por parte del estudiante.

**IV CONTENIDO**

En este apartado se colocan los temas que se verán desarrollados en el curso, estos en relación con los objetivos que se pretenden lograr por el estudiante. Además se da el detalle del temario de la asignatura divididos por unidades para una mejor organización del contenido.

Unidad	Contenido
.....	.....

## V METODOLOGÍA

La parte de metodología explica cómo se desarrollará el curso, cuáles son las estrategias para impartir el conocimiento al estudiante, además se hace mención de los recursos y herramientas que se utilizarán para en apoyo a la metodología escogida.

## VI EVALUACIÓN

En esta parte del programa se detallan todas las evaluaciones a las que se someterá el estudiante, especificando la descripción de la evaluación y su ponderación asignada.

Evaluación	Peso específico en calificación final
....	.....
....	.....
...	.....
TOTAL	100%

## VII BIBLIOGRAFÍA

Apartado utilizado para colocar todas las referencias utilizadas en la elaboración del programa de la asignatura.

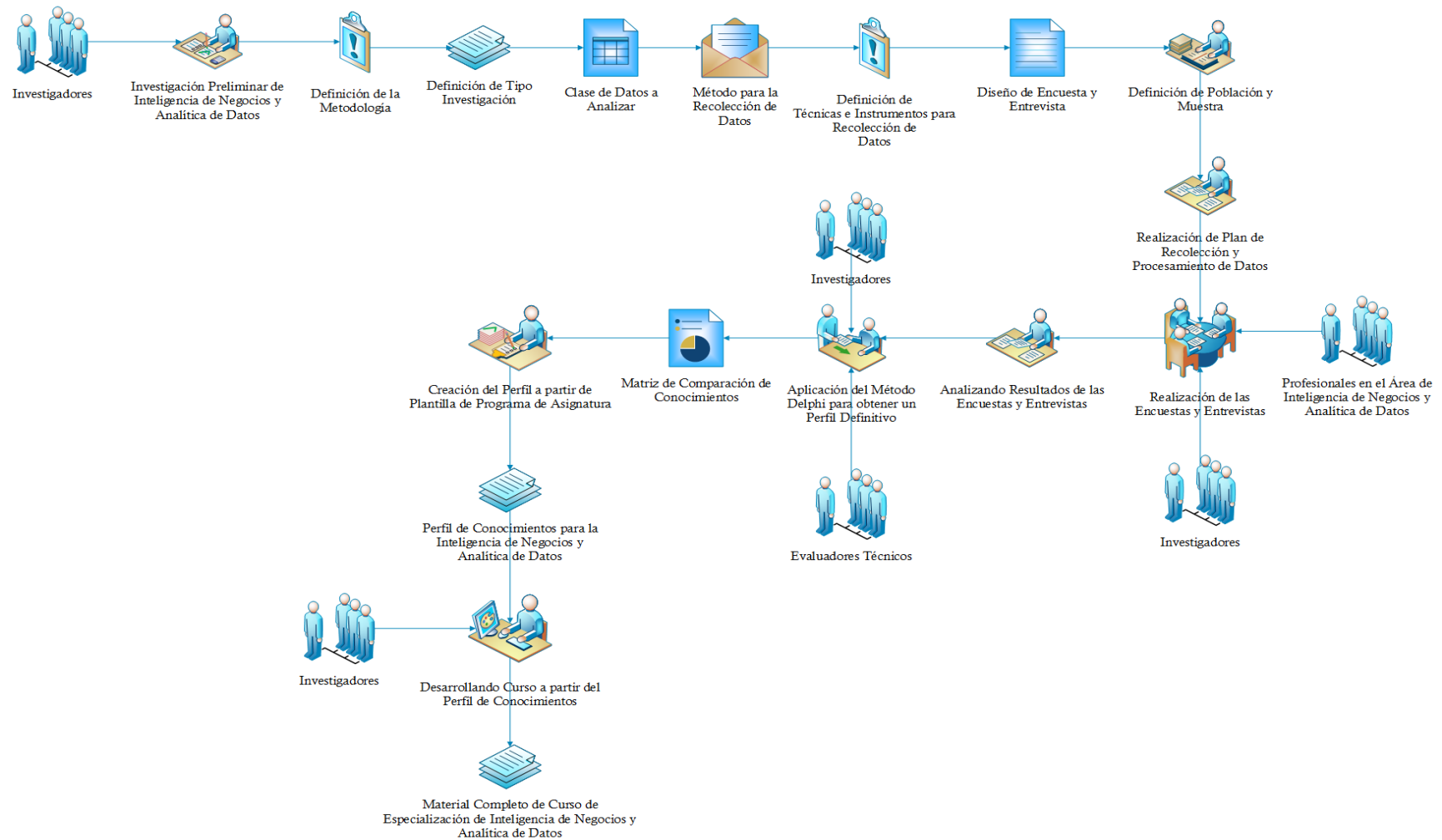
Con el fin de potenciar el aprendizaje de los candidatos al curso, el desarrollo de los contenidos tendrán una base teórica y el desarrollo de casos prácticos, abarcando desde lo fundamental hasta las partes más especializadas; para ello se construirá el material que brinde el apoyo para cumplir con dicho fin. Los insumos necesarios para desarrollar el material del curso de especialización se obtendrá a partir de investigación bibliográfica en libros, investigaciones académicas, cursos de universidades extranjeras.



Uno de los pilares del curso de especialización es la enseñanza de herramientas que ayudan a aplicar los temas relacionados a inteligencia de negocios así como de analítica de Big Data es por eso que dentro de los contenidos a desarrollar se realizarán guías de uso de las aquellas herramientas que mejor se adapten a las requerimientos del perfil. La selección de la(s) herramienta(s) a considerar en el curso será producto de un benchmarking de herramientas. Se presenta a continuación la matriz a utilizar.

Evaluación de Herramientas de inteligencia de negocios y análisis de datos																
Herramientas	Criterios															
	Plataforma			Tipo de licenciamiento				Disponibilidad de documentacion oficial		Ofrece Servicios de soporte		Requerimientos Técnicos minimos requeridos				
	Local	Web	Móvil	Pago	Open Source	Licencia Estudiantil	Trial	Si	No	Si	No	Sistema Operativo	Procesador	Disco Duro	Memoria RAM	Utros (tarjetas de video, etc)

A continuación, se presenta un diagrama que presenta todo el proceso de la metodología a seguir para el desarrollo del presente proyecto de investigación.



## CAPÍTULO IV: ASPECTOS ADMINISTRATIVOS

### 4.1 Cronograma de Actividades

Actividades	Fecha de Inicio	Fecha de Finalización
<b>Realización del Perfil</b>	<b>20/02/2017</b>	<b>20/03/2017</b>
Definir el Tema	20/02/2017	24/02/2017
Descripción del Tema	25/02/2017	26/02/2017
Objetivos	25/02/2017	26/02/2017
Búsqueda de Información en fuentes secundarias	27/02/2017	28/02/2017
Antecedentes	01/03/2017	03/03/2017
Justificación	04/03/2017	07/03/2017
Alcances	08/03/2017	09/03/2017
Diseño de Encuesta para Investigar la Situación Actual	10/03/2017	11/03/2017
Envío de Encuestas para recolectar datos	12/03/2017	17/03/2017
Análisis de Resultados	18/03/2017	19/03/2017
Entrega del Perfil	20/03/2017	20/03/2017
<b>Anteproyecto</b>	<b>21/03/2017</b>	<b>05/05/2017</b>
Definir la Estructura del Contenido del Proyecto	21/03/2017	23/03/2017
Búsqueda de Información en Bibliografías Relevantes	24/03/2017	28/03/2017
Marco Teórico	29/03/2017	03/04/2017
Definición de la Metodología	05/04/2017	26/04/2017
Planificación y Presupuesto	27/04/2017	30/04/2017

Entrega del Anteproyecto a evaluadores	05/05/2017	05/05/2017
Defensa del Anteproyecto	08/05/2017	12/05/2017
<b>Etapas 1</b>	<b>13/05/2017</b>	<b>03/07/2017</b>
Realizar Encuesta y Entrevista	13/05/2017	27/05/2017
Informe estadístico de las necesidades del perfil	28/05/2017	30/05/2017
Matriz de comparación de conocimientos	31/05/2017	01/06/2017
Informe descriptivo del perfil académico y profesional	02/06/2017	15/06/2017
Planificación de los Contenidos del Curso	16/06/2017	25/06/2017
Entrega de Etapas 1	26/06/2017	27/06/2017
Defensa de Etapas 1	28/06/2017	03/07/2017
<b>Etapas 2</b>	<b>04/07/2017</b>	<b>10/11/2017</b>
Desarrollo de los contenidos del Curso Guías del uso de las herramientas Preparación del material didáctico Guías de Laboratorio	01/07/2017	31/10/2017
Entrega de Etapas 2	02/11/2017	02/11/2017
Defensa de Etapas 2	06/11/2017	10/11/2017
<b>Defensa Final</b>	<b>13/11/2017</b>	<b>19/11/2017</b>

## 4.2 Presupuesto

Se presenta el presupuesto de este proyecto el cual tendrá una duración de 9 meses aproximadamente

### Recurso Humano

El equipo de investigadores invertirá 4 horas diarias por 6 días haciendo un total de 96 horas mensuales en el presente proyecto, se estima que el costo por hora de cada miembro del equipo será de \$5.625 en base al salario mensual de un Investigador de la Universidad de El Salvador <sup>50</sup>

Recurso	Costo Mensual	Tiempo (meses)	Costo Total
4 investigadores	\$2,160.00	9	\$19,440.00

### Servicios Básicos

De la misma manera para los costos de la energía eléctrica y el servicio de internet se tomó como base 96 horas de trabajo mensuales para cada miembro del equipo

Recurso	Costo Mensual	Tiempo (meses)	Costo Total
Energía Eléctrica	\$19.20	9	\$172.80
Internet Residencial 3 Mbps	\$11.52	9	\$103.68
Telefonía Móvil	\$20.00	9	\$180.00
Total			\$456.48

### Recursos consumibles

Recurso	Cantidad	Costo Unitario	Costo Total
Suscripción Survey Monkey	3	\$26.00/ mes	\$78.00
Ebook "Como hacer tu trabajo de investigación"	1	\$4.00	\$4.00

<sup>50</sup> Marco de Transparencia de la Universidad de El Salvador. (2016). Marco Presupuestario y Financiero. Mayo 01, 2017, de Universidad de El Salvador Sitio web: <http://www.transparencia.ues.edu.sv/sites/default/files/PDF/LEY%20DE%20SALARIOS%202016.pdf>

Papelería	2	\$3.50	\$7.00
Cartucho de tinta negra	4	\$22.00	\$88.00
Cartucho de tinta a color	3	\$25.00	\$75.00
Renta de proyector	8	\$7/hora	\$56.00
Consumo de gasolina mensual	9	\$30.00	\$270.00
Transporte público mensual	9	\$20.00	\$180.00
Cursos de apoyo de inteligencia de negocios y analítica de Big Data	1	\$200.00	\$200.00
Viáticos para la recolección de datos	3	30.00	\$90.00
<b>Total</b>			<b>\$1,048.00</b>

### Resumen de Gastos

<b>Recurso</b>	<b>Total</b>
Recurso Humano	\$19,440.00
Servicios Básicos	\$456.48
Recursos Consumibles	\$1,048.00
<b>Sub Total</b>	<b>\$20,944.48</b>
Imprevistos 10%	\$2,094.45
<b>Total</b>	<b>\$23,038.93</b>

## CONCLUSIONES

---

- Existe una necesidad actual en las empresas salvadoreñas de profesionales que hagan inteligencia de negocios y sean capaces de realizar análisis de los datos que poseen, los cuales se enriquecen al aprovechar de otros medios que la tecnología ofrece: redes sociales, geo localización, audiovisuales, comunicación por correo electrónico (por mencionar algunos). La incorporación de nuevas fuentes de datos hace del análisis de Big Data un reto y una necesidad latente en las empresas de El Salvador.
- La Universidad de El Salvador, como única institución pública de educación superior, tiene especial responsabilidad con la sociedad salvadoreña en cuanto a la calidad y el nivel de educación que ofrece. La Facultad de Ingeniería y Arquitectura como parte de ésta institución, a través de sus diferentes escuelas deberá proponer soluciones de acuerdo a la realidad laboral salvadoreña. El campo de la informática es rápidamente cambiante por lo que la Escuela de Ingeniería de Sistemas Informáticos debe procurar mantener en la vanguardia los programas de estudio que imparte en función de los requerimientos actuales y futuros.
- La especialización en un área de conocimiento para un profesional implica mayor calidad de trabajo realizado, dado que los esfuerzos y el tiempo invertido enfocados a un tema en específico aumenta la experiencia y la profundidad de conocimiento, los cuales no serán los mismos ni aporta el mismo valor de aquel profesional que trata hacer de todo. En el campo de la informática, las áreas de aplicación son extensas y múltiples (se habla de áreas de redes, administración de bases de datos, arquitectura de computadoras, por mencionar algunas) y cada una requiere de habilidades y conocimientos propios que no pueden ser desarrollados si se tiene un enfoque de educación holístico.

## RECOMENDACIONES

---

- El tema de inteligencia de negocios y Big Data es una oportunidad de investigación que podría ser aprovechada por entidades enfocadas a la investigación, educación y desarrollo tecnológico en El Salvador. Actualmente la información que se posee del tema en el contexto nacional es escasa y dado el posicionamiento en el área laboral y las necesidades latentes se convierte en un tema de interés con futuro.
- La inteligencia de negocios y Big Data al impactar al sector educación, se convierte en una motivación y necesidad de constante actualización de conocimientos para el sector docente, por tanto, toda entidad que quiera abordar el tema desde el punto de vista académico deberá apostarle a la capacitación técnica de instructores, de ser posible, apoyándose de instituciones internacionales expertas en el área.



## REFERENCIAS

---

A. Simitsis y P. Vassiliadis, «*A method for the mapping of conceptual designs to logical blueprints for ETL processes*», *Decision Support Systems*, vol. 45, no. 1, págs. 22-40, Abr. 2008.

J. Conesa & J. Curto. (2010). *Introducción al Bussiness Intelligence*. Barcelona: UOC.

S. Dupor y V.Jovanovic(2014). *An approach to conceptual modelling of ETL processes*. MIPRO 2014.

¿Qué es Power BI? | Microsoft Power BI. (2017). Powerbi.microsoft.com. Recuperado 16 April 2017, de <https://powerbi.microsoft.com/es-es/what-is-power-bi/>

Tableau Desktop. (2017). *Tableau Software*. Recuperado 17 April 2017, a partir de <https://www.tableau.com/es-es/products/desktop>

Visualización de datos | Analítica integrada e informes | Qlik. (2017). Qlik.com. Recuperado 17 April 2017, a partir de <http://www.qlik.com/es-es/products>

Vercellis, C..(2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. Reino Unido: Wiley.

Rajan, J.. (2005). *Business Intelligence: Concepts, Components, Techniques And Benefits*. *Journal of Theoretical and Applied Information Technology*, Vol. 9, pp 60-70.

Olivares, J.. (2002, Marzo 01). *Análisis exploratorio y análisis confirmatorio de datos*. *Espacio Abierto*, 11, 124. n.d, De n.d Base de datos.

Ferrero, R.. (2017). *Análisis Multivariado para Big Data*. Abril 16,2017, de Maxima Información Sitio web: <https://www.maximaformacion.es/blog-r/41-big-data-an%C3%A1lisis-multivariado>

Rouse, M.. (2016). predictive analytics. Abril 15, 2017, de Tech Target Sitio web: <http://searchbusinessanalytics.techtarget.com/definition/predictive-analytics>

Rouse, M.. (2017). data mining. Abril 15, 2017, de Tech Target Sitio web: <http://searchsqlserver.techtarget.com/definition/data-mining>

Ramos, A.. (2012, Julio 01). Cálculo del tamaño óptimo de la muestra. *Estadística aplicada a la Investigación*, 1, 10. 2017, mayo 1, De n.d Base de datos.

## ANEXOS

---

### **Anexo 1: Lluvia de Ideas para el análisis y determinación de la situación problemática**

- El plan de estudio de la carrera de Ingeniería de Sistemas Informáticos no se ha actualizado desde 1998
- Existen otras tendencias para un ingeniero informático, el análisis y el desarrollo de sistemas puede ser realizado por técnicos, ahora se requiere que un ingeniero tenga un enfoque a los datos.
- El internet posibilita el acceso a la información, por este medio, es que se conocen las tendencias y los estudios que se están realizando en la carrera de Ingeniería de Sistemas Informáticos. Las bolsas de trabajo nacionales publican con más frecuencia plazas relacionadas con el data warehousing, elaboración de reportes, analistas.
- Existen universidades e instituciones que están innovando la enseñanza en ofrecer cursos de especialización y postgrados en inteligencia de negocios
- La generación de información por medio de sistemas transaccionales, mensajes de texto, redes sociales, transacciones en línea, correos se está incrementando y ahora es necesario no solo almacenar los datos, sino más bien, darles un tratamiento y generar conocimiento a partir de ellos.
- Hay empresas que ya contemplan unidades destinadas a la inteligencia de negocios
- El término Big Data se ha popularizado en El Salvador, la Escuela Superior de Economía y Negocios tiene a disposición a partir del 27 de mayo 2016 un Postgrado en Analítica para Ciencia de Datos
- En el país existen empresas que necesitan aplicar tecnologías que permitan analizar de mejor manera los datos que manejan.
- Cubrir conocimientos que el estudiante o profesional no cumplen en el campo laboral
- Identificar los conocimientos, las habilidades y técnicas actuales, que se requieren en un profesional especializado en inteligencia de negocios y analítica de datos en el salvador
- Existen sistemas transaccionales robustos que generan gran cantidad de datos de las diferentes operaciones de las grandes y medianas organizaciones; estos datos al ser detallados son necesarios que pasen por un análisis con el cual se genere un formación Con la cual se Puedan tomar decisiones, ese proceso de análisis se realiza por medio de la inteligencia de negocios y analítica de datos

## **Anexo 2: Análisis de Resultado de Encuesta “Investigación de la Situación Actual de la inteligencia de negocios y analítica de datos”**

Fecha de sondeo: 08 de marzo -12 de marzo de 2017

### **Pregunta 1**

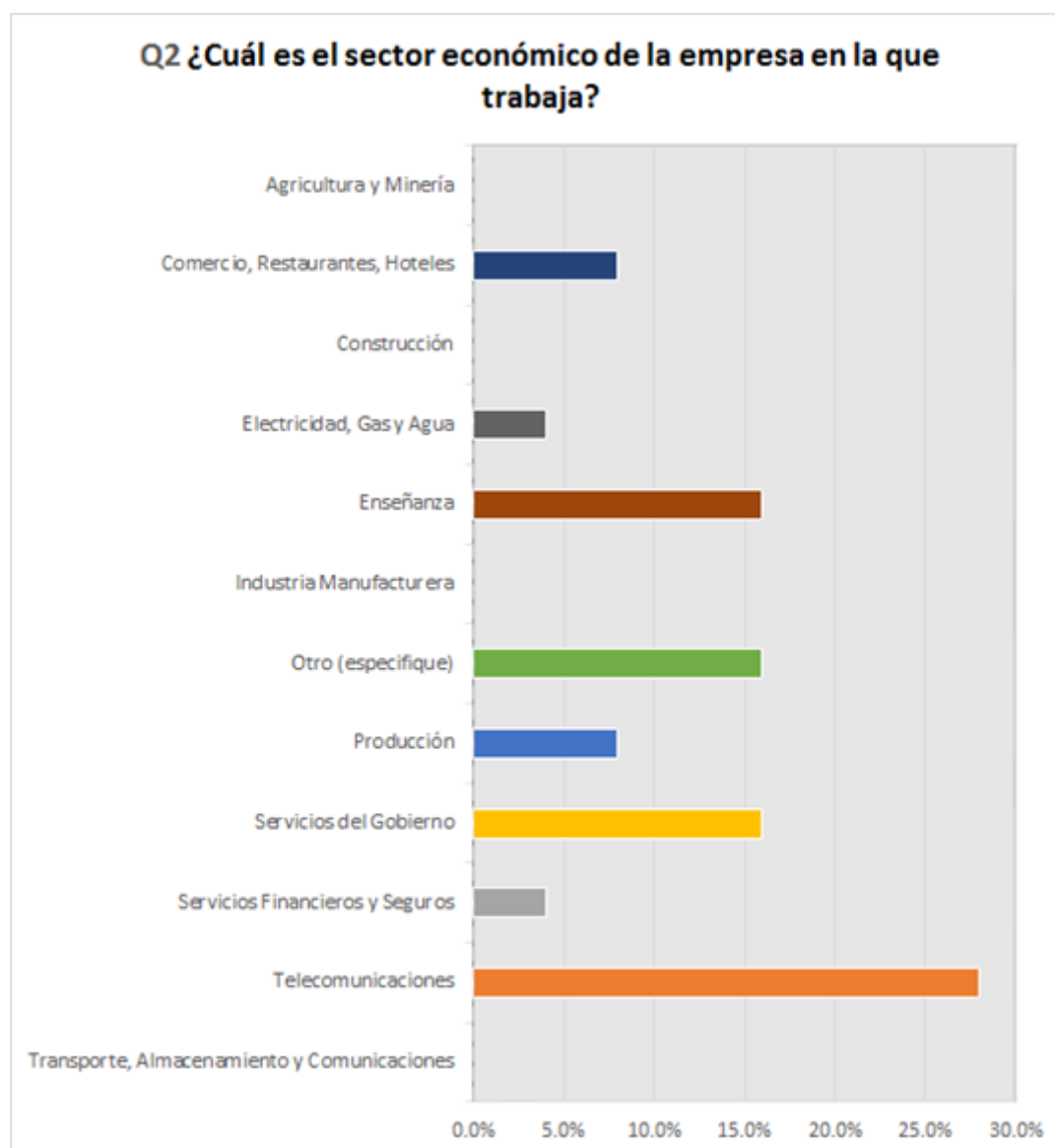
N/A: Correo electrónico y/o teléfono de los encuestados.

### **Pregunta 2**

#### **¿Cuál es el sector económico de la empresa en la que trabaja?**

Conocer el sector económico al que pertenece el encuestado hace reconocer que la aplicación de la inteligencia de negocios y analítica de datos puede aplicarse a cualquier actividad empresarial en donde se vea involucrada información.

Las respuestas a esta encuesta reafirman el amplio alcance de la inteligencia de negocios y analítica de datos al obtener resultados no sólo para sectores de tecnología e informática en sus varias expresiones(consultoría o desarrollo) sino también para sectores de Telecomunicaciones, Servicios de Gobierno, Enseñanza, Comercio/Restaurantes/Hoteles y Producción.



Opciones de Respuestas	Respuestas	
Transporte, Almacenamiento y Comunicaciones	0.0%	0
Telecomunicaciones	28.0%	7
Servicios Financieros y Seguros	4.0%	1
Servicios del Gobierno	16.0%	4
Producción	8.0%	2
Otro (especifique)	16.0%	4
Industria Manufacturera	0.0%	0
Enseñanza	16.0%	4
Electricidad, Gas y Agua	4.0%	1
Construcción	0.0%	0
Comercio, Restaurantes, Hoteles	8.0%	2
Agricultura y Minería	0.0%	0
<b>Total</b>		<b>25</b>

Otro (especifique)	
Servicios informáticos desarrollo de software	
Consultoría	
Software	
Tecnología	

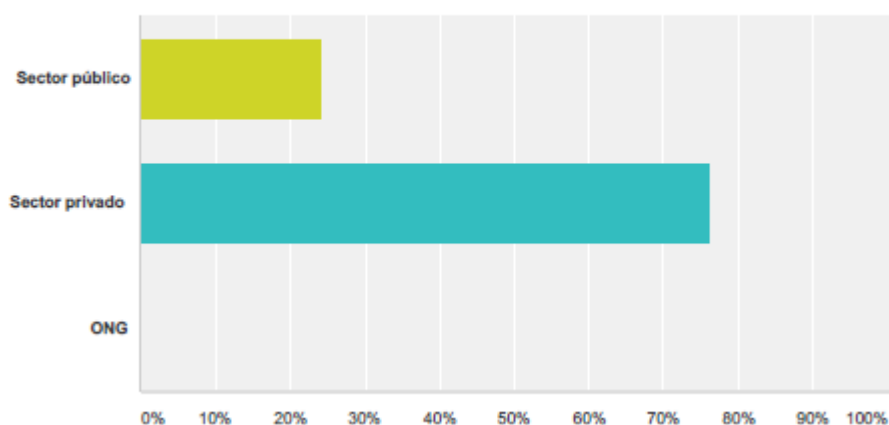
### Pregunta 3

#### ¿Tipo de empresa en la que trabaja?

Los resultados de esta encuesta identifican que el 76% de la muestra encuestada corresponde a profesionales que laboran en sector privado, siendo el porcentaje restante profesionales que laboran el sector público.

No se obtuvieron respuestas de profesionales pertenecientes al sector No gubernamental.

#### Q3 ¿Tipo de empresa en la que trabaja?



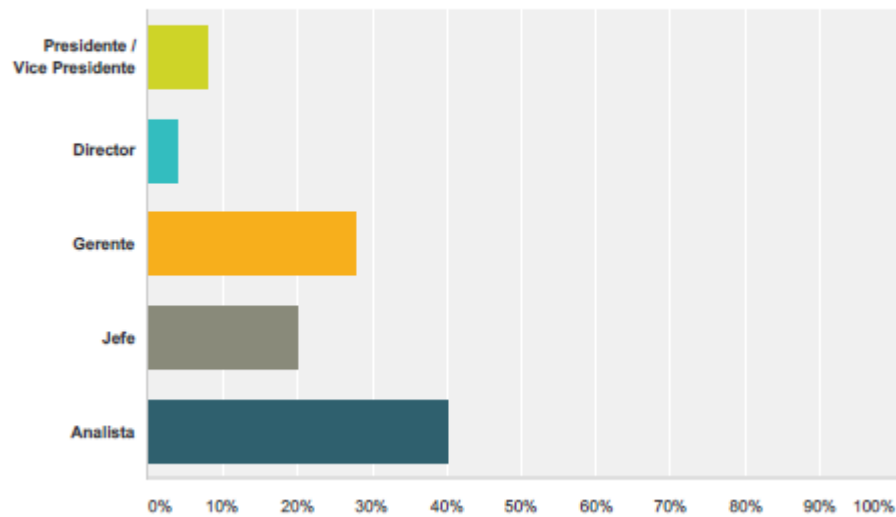
Opciones de respuesta	Respuestas
Sector público	24,00% 6
Sector privado	76,00% 19
ONG	0,00% 0
Total	25

### Pregunta 4

#### ¿Qué cargo desempeña?

El diseño muestral dirigido/intencional, utilizado en esta encuesta, deja a criterio del encuestador, elegir la muestra según la representatividad, en este caso, se determinó que se quería obtener la percepción de las personas que realizaban toma de decisiones y no quienes generaban la información, por tanto la encuesta debía de estar dirigida a personas de cargos de analistas, gerencias, jefaturas, vicepresidencias o presidencias.

#### Q4 ¿Qué cargo desempeña?

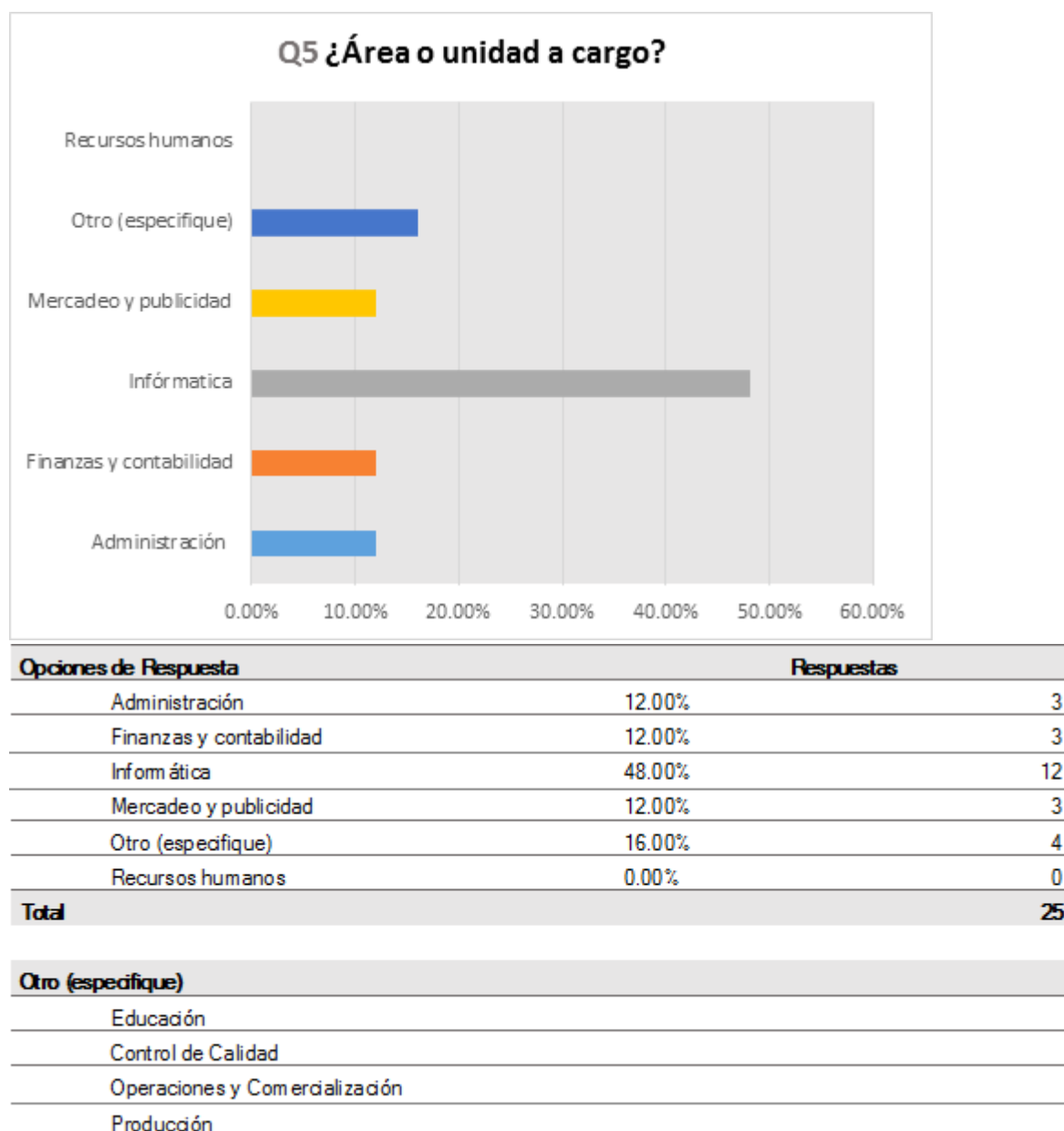


Opciones de respuesta	Respuestas
Presidente / Vice Presidente	8,00% 2
Director	4,00% 1
Gerente	28,00% 7
Jefe	20,00% 5
Analista	40,00% 10
Total	25

#### Pregunta 5

##### ¿Área o unidad a cargo?

Conocer el área a cargo del encuestado ayuda a hacer análisis de dos variables en el momento que se requiera

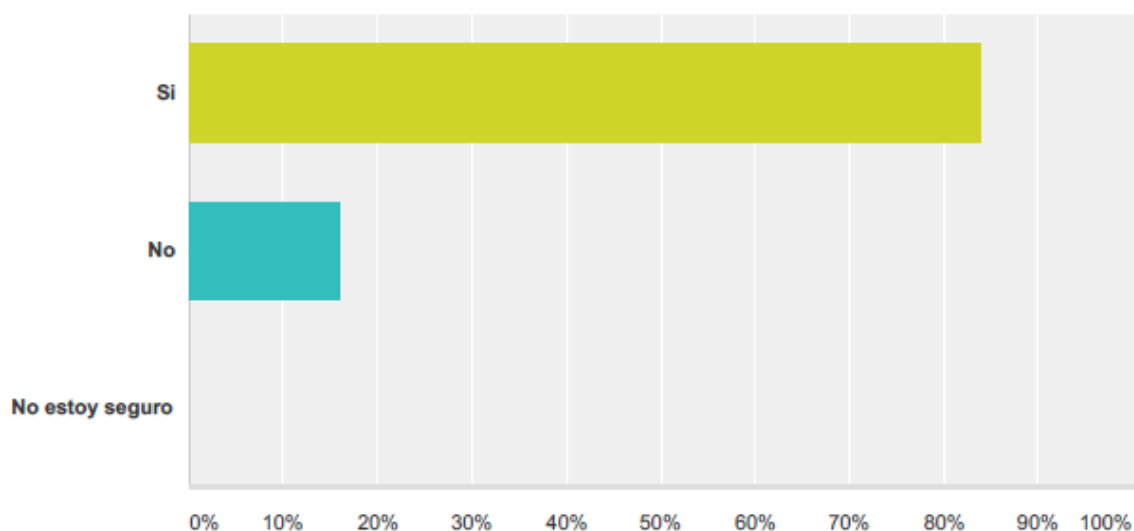


## Pregunta 6

### ¿Está familiarizado con el concepto de Inteligencia de negocios y analítica de datos?

A través de esta interrogante se midió si las personas de alguna forma estaban familiarizadas con estos conceptos, el 84% de los encuestados afirmó estar familiarizado y el restante (16%) no lo estaba; las personas encuestadas desempeñan cargos de jefatura, en donde la toma de decisiones está a la orden del día por tanto se esperaba que la mayoría estuviese familiarizado con el concepto.

### Q6 ¿Esta familiarizado con el concepto de Inteligencia de negocios y analítica de datos?



Opciones de respuesta	Respuestas
Si	84,00% 21
No	16,00% 4
No estoy seguro	0,00% 0
Total	25

### Pregunta 7

#### ¿Por medio de qué fuentes conoce de inteligencia de negocios y analítica de datos?

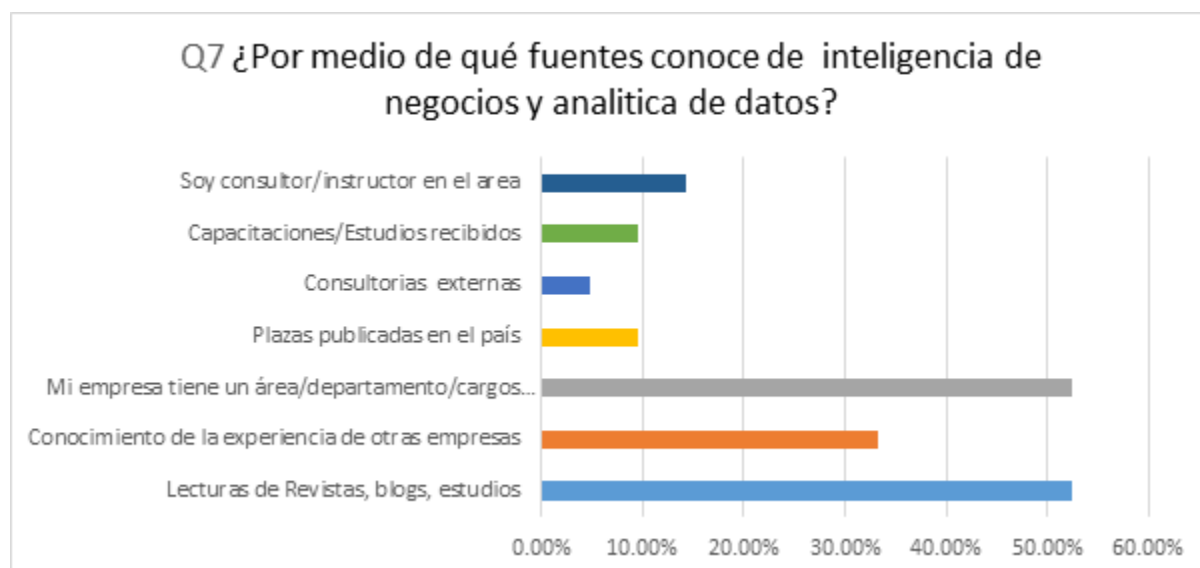
Con esta interrogante se determinaron las principales fuentes de información a través de las cuales los encuestados conocen los conceptos en estudio, las fuentes que comparten el mismo porcentaje de elección y el más representativo son dos: lecturas de revistas, blogs y estudios y la existencia áreas/departamento/cargos en las empresas donde laboran los encuestados.

La inteligencia de negocios y analítica de datos desde sus orígenes, ha sido objeto de innumerables estudios e investigaciones, por lo cual se espera que estas iniciativas internacionales y la facilidad de acceso a la información a través del internet sirvan de medio para que los profesionales salvadoreños conozcan acerca del tema.

Estas fuentes primarias disponibles y el impacto de la globalización hacen que las empresas salvadoreñas adopten las nuevas tendencias en función de mantenerse en sostenibilidad. No es extraño entonces, que empresas ya adopten dentro de su estructura organizacional áreas que ejerzan las funciones de la inteligencia de negocios y analítica de datos.



Estos macro factores (el conocimiento existente y la experiencia en las empresas) dan nacimiento a nuevas formas de generar conocimiento sobre la temática: las consultorías en el área se han vuelto más populares y por tanto muchas más personas y empresas están interesadas en recibir capacitaciones sobre el tema.



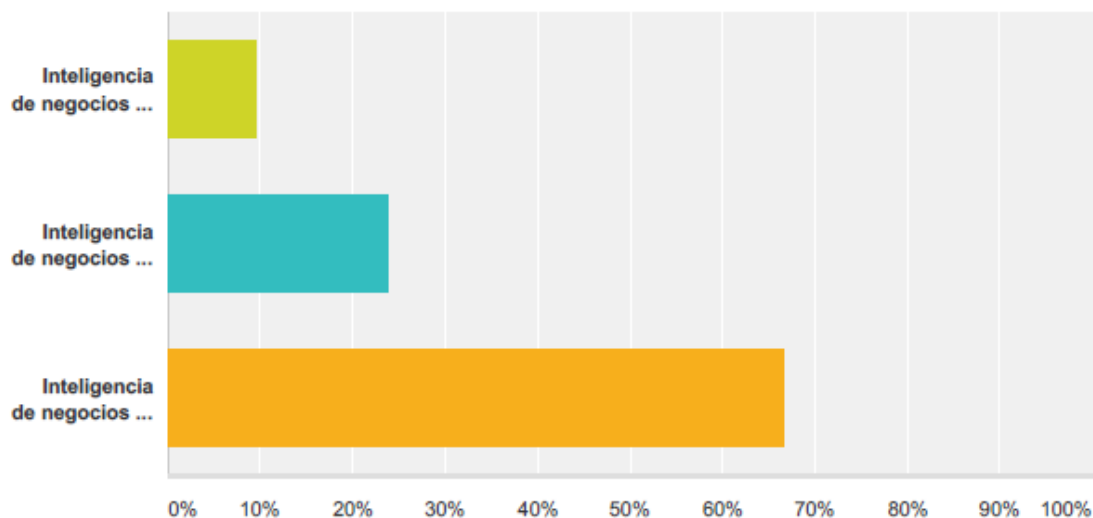
Opciones de Respuesta	Respuestas	
Lecturas de Revistas, blogs, estudios	52.38%	11
Conocimiento de la experiencia de otras empresas	33.33%	7
Mi empresa tiene un área/departamento/cargos para ello	52.38%	11
Plazas publicadas en el país	9.52%	2
Consultorias externas	4.76%	1
Capacitaciones/Estudios recibidos	9.52%	2
Soy consultor/instructor en el área	14.29%	3
<b>Total Encuestados: 21</b>		

## Pregunta 8

### ¿Qué entiende por Inteligencia de negocios?

Se quería saber si los encuestados conocían del concepto de inteligencia de negocios, para ellos se colocaron tres respuestas de las cuales solo una es cierta. El 66,67% de las personas eligió la respuesta correcta, mientras que el porcentaje restante eligió una de las dos respuestas erróneas. Con ello se concluye que a pesar de que las personas decían conocer el concepto de inteligencia de negocios, no todas tienen una comprensión correcta de lo que consiste. El 20.63% de los que contestaron que **Sí** a la pregunta “¿Está familiarizado con el concepto de Inteligencia de negocios y analítica de datos?” no pudo identificar el concepto de Inteligencia de negocios.

### Q8 ¿Que entiende por Inteligencia de negocios?



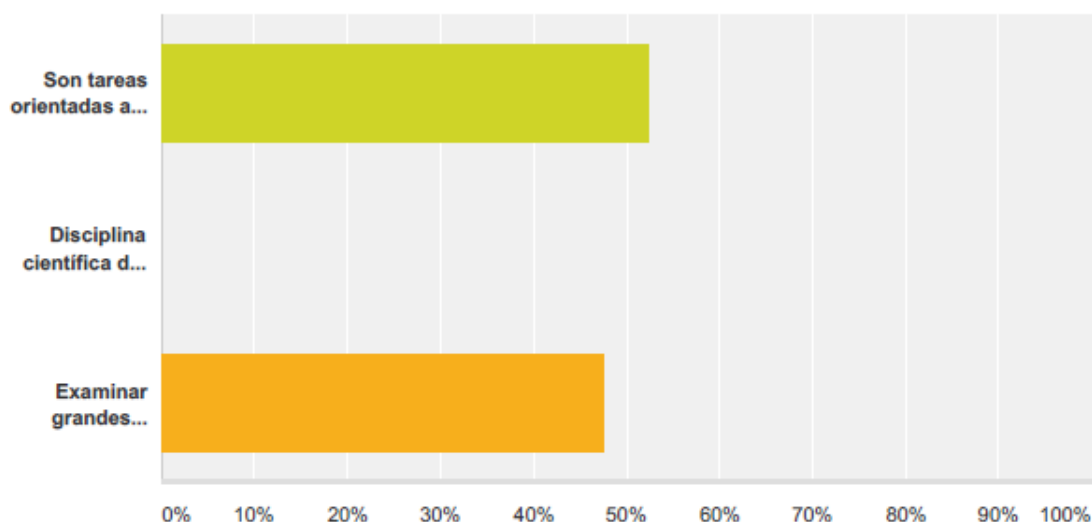
Opciones de respuesta	Respuestas
Inteligencia de negocios es CRM, Datawarehouse, ERP, DataMart y Cuadros de Mando	9,52%
Inteligencia de negocios es disponer en mi empresa de un repositorio de datos (Datawarehouse) en donde poder hacer consultas complejas para obtener valor añadido	23,81%
Inteligencia de negocios es disponer la información de la empresa orientada a conseguir su estrategia	66,67%

### Pregunta 9

#### ¿Qué entiende por analítica de datos?

Se siguió la misma lógica para medir el conocimiento sobre analítica de datos. Los resultados que se obtuvieron son muy parejos, lo cual no se traduce a resultados satisfactorios. Si bien fueron más las personas que acertaron a la respuesta (52.38%), el porcentaje de error es muy alto. El 37.64% de todos los que contestaron que *Sí* a la pregunta “¿Está familiarizado con el concepto de Inteligencia de negocios y analítica de datos?” no pudo identificar el concepto de analítica de datos. Por tanto, es un tema que deberá ser abordado con más profundidad.

### Q9 ¿Que entiende por analítica de datos?



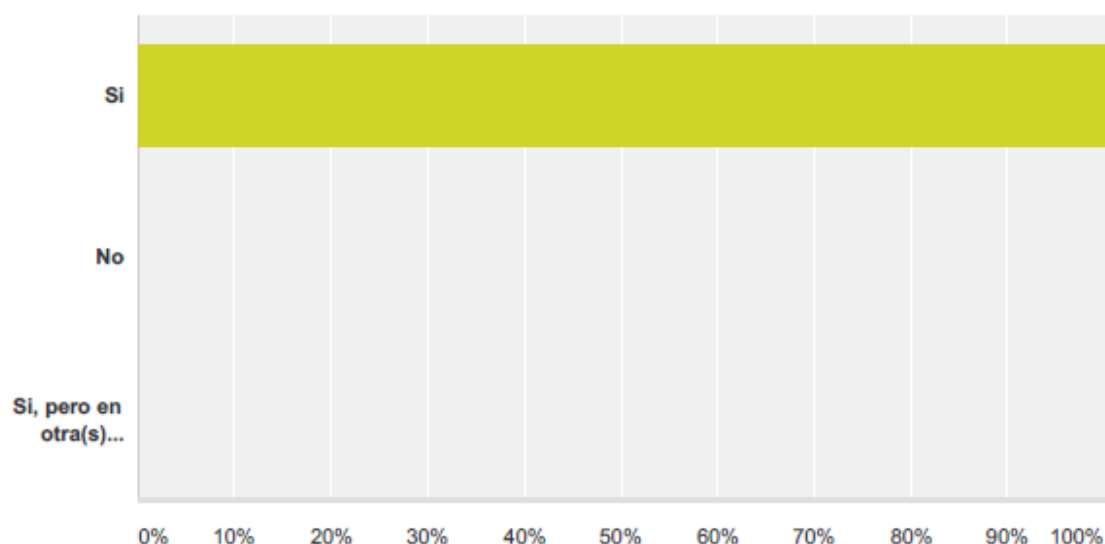
Opciones de respuesta	Respuestas
Son tareas orientadas a la exploración de los datos, con la intención de encontrar patrones o conocimiento útil, que permita optimizar o rentabilizar un proceso de negocio	52,38%
Disciplina científica del ámbito de la Inteligencia Artificial que crea sistemas que aprenden automáticamente	0,00%
Examinar grandes cantidades de datos de distintos tipos para descubrir patrones ocultos, correlaciones desconocidas y otra información útil	47,62%

### Pregunta 10

**¿Considera que en su empresa o en su gerencia/departamento es importante que los datos se analicen con mayor énfasis?**

El 100% de los encuestados reconoce la importancia de analizar los datos que se generan en sus empresas de una manera más profunda; uno de los propósitos de la inteligencia de negocio y analítica de datos es de hecho analizar los datos para poder generar información, por lo tanto se destaca la necesidad implícita de aplicar inteligencia de negocios y analítica de datos dentro de sus lugares de trabajo

### Q10 ¿Considera que en su empresa o en su gerencia/departamento es importante que los datos se analicen con mayor énfasis?



Opciones de respuesta	Respuestas	
Si	100,00%	25
No	0,00%	0
Si, pero en otra(s) gerencia(s) o departamento(s)	0,00%	0
<b>Total</b>		<b>25</b>

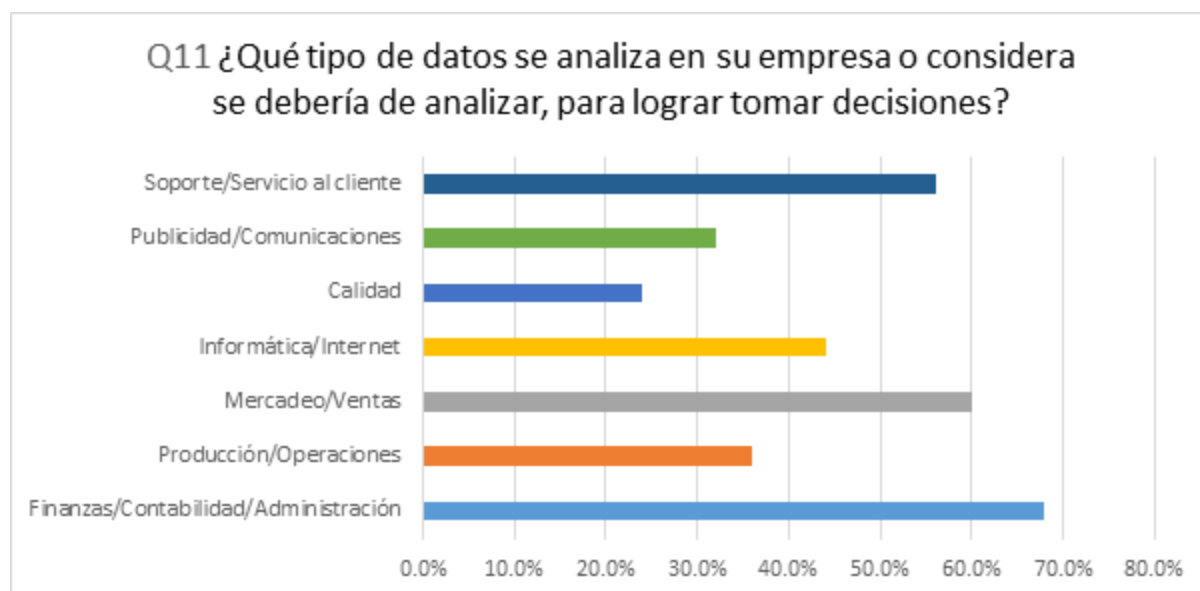
### Pregunta 11

#### ¿Qué tipo de datos se analiza en su empresa o considera se debería de analizar, para lograr tomar decisiones?

Los tres tipos de datos que resultaron más recurrentes para este estudio fueron: Finanzas/Contabilidad/Administración, Mercadeo/Ventas y Soporte/Servicio al cliente. Podría pensarse que las respuestas obtenidas tienen mucho que ver con el sector económico más representativo de los encuestados: las telecomunicaciones (28%). Sin embargo, ¿Qué empresa no lleva sus registros contables? ¿Cuántas empresas buscan ganar la fidelidad de los clientes a partir del servicio prestado o del producto entregado? ¿A caso la calidad de los servicios o de los productos, no debe ser un punto a analizar?

Todas las empresas, independiente del rubro económico, requieren un análisis de los distintos tipos de datos que manejan, estos se complementa: *“a mayor calidad de un servicio, mayores ventas, más ingresos”, “la optimización de la producción de un producto lleva a reducir costos”, etc.*

Por tanto, la inteligencia de negocios y analítica de datos puede aplicarse a cualquier tipo de dato que se genere en una empresa, lo cual se fundamenta con el hecho que ninguna de las opciones presentadas quedó sin respuesta.



Opciones de Respuesta	Respuesta
Finanzas/Contabilidad/Administración	68.0%
Producción/Operaciones	36.0%
Mercadeo/Ventas	60.0%
Informática/Internet	44.0%
Calidad	24.0%
Publicidad/Comunicaciones	32.0%
Soporte/Servicio al cliente	56.0%

## Pregunta 12

### ¿Qué hacen con los datos que generan/reciben en su gerencia/departamento?

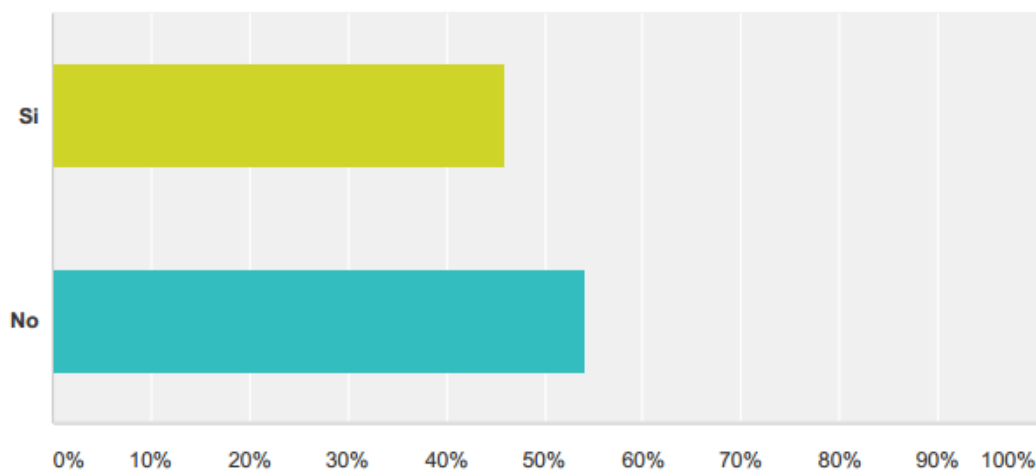
Esta pregunta no contiene respuestas ya que estaba condicionada a que solo los que respondieran *No* a “¿Considera que en su empresa o en su gerencia/departamento es importante que los datos se analicen con mayor énfasis?” serían redirigido a esta pregunta. El 100% de los encuestados consideró que sus datos deben pasar por un proceso de análisis debido al aporte que agrega esta información a las organizaciones.

## Pregunta 13

### ¿Considera que la información que actualmente posee es suficiente para tomar decisiones?

Un 54.17% de los encuestados consideran que la información que poseen no es suficiente para tomar decisiones, estos se traduce en una oportunidad para la inteligencia de negocios y analítica de datos para posicionarse en las empresas por su gran aporte al proceso de toma de decisiones.

### Q13 ¿Considera que la información que actualmente posee es suficiente para tomar decisiones?



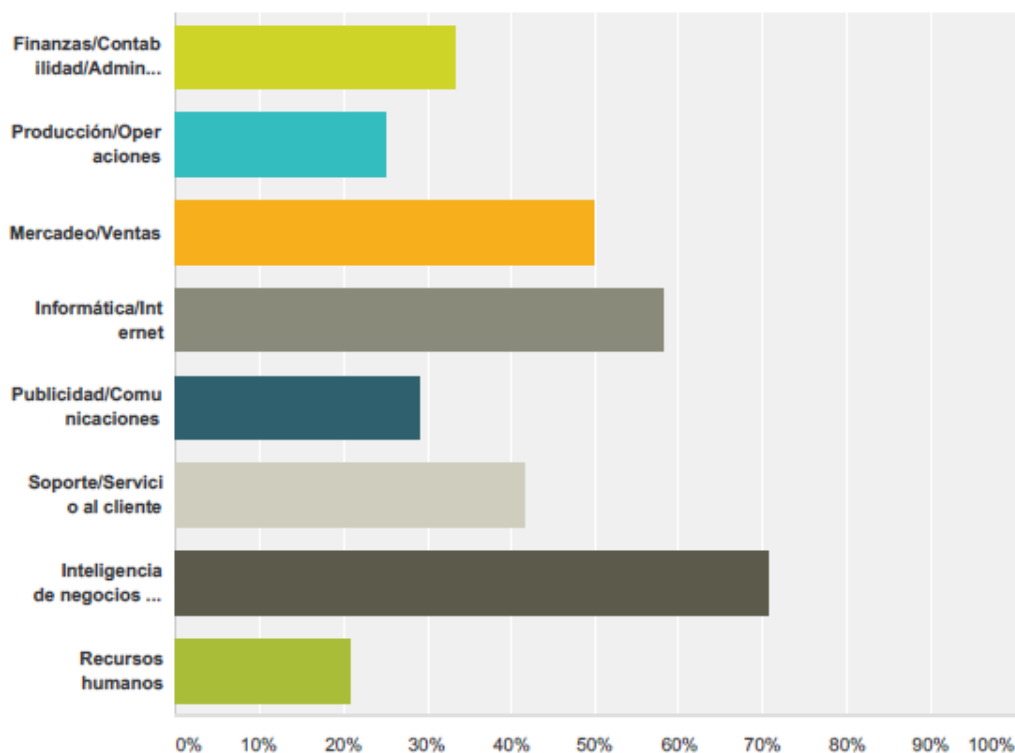
Opciones de respuesta	Respuestas	
Si	45,83%	11
No	54,17%	13
Total		24

### Pregunta 14

**¿En qué departamentos o áreas de la empresa considera que se debe invertir recursos?**

El objetivo de esta pregunta era conocer qué porcentaje de personas consideraba que de entre todas las áreas listadas, inteligencia de negocios y analítica de datos debía ser considerada también como un área en la cual las empresas deberían enfocarse. Ninguna opción fue elegida unánimemente, sin embargo fue inteligencia de negocios y analítica de datos la respuesta que más frecuencia de selección tuvo. Se reconoce el potencial de esta área para aportar valor agregado a las empresas.

### Q14 ¿En qué departamentos o áreas de la empresa considera que se debe invertir recursos ?



Opciones de respuesta	Respuestas	
Finanzas/Contabilidad/Administración	33,33%	8
Producción/Operaciones	25,00%	6
Mercadeo/Ventas	50,00%	12
Informática/Internet	58,33%	14
Publicidad/Comunicaciones	29,17%	7
Soporte/Servicio al cliente	41,67%	10
Inteligencia de negocios y análisis de datos	70,83%	17
Recursos humanos	20,83%	5
Total de encuestados: 24		

### Pregunta 15

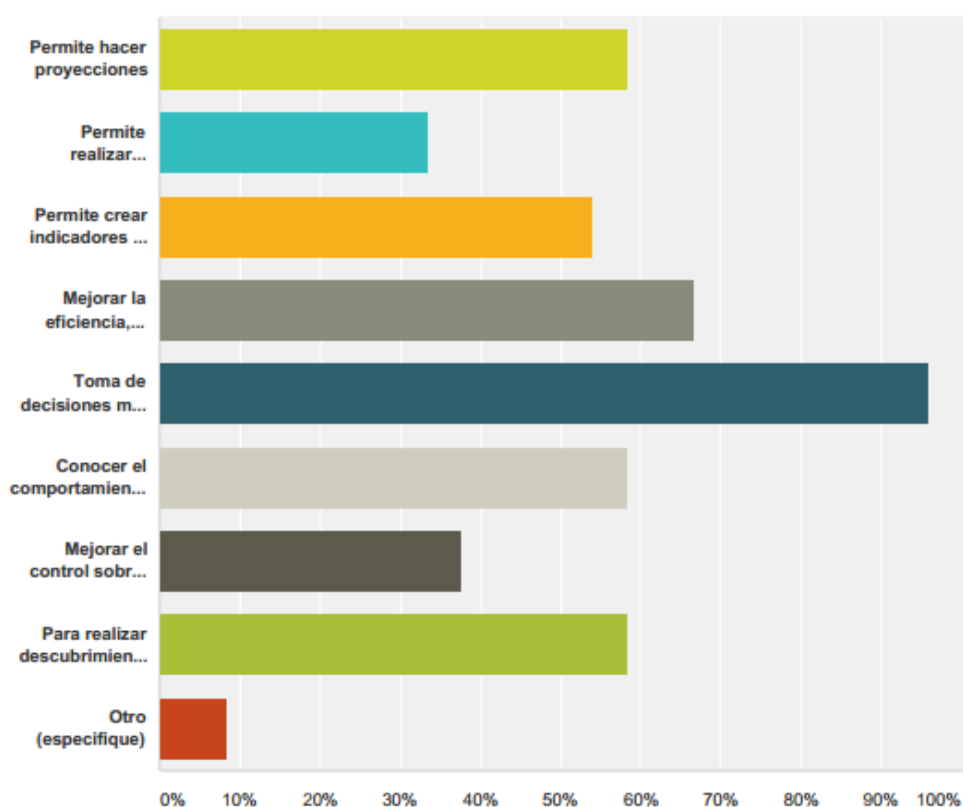
**Seleccione la(s) opción(es) con qué asocia la inteligencia de negocios y analítica de datos:**

La finalidad de esta pregunta es medir los beneficios que las personas consideran obtendrán de la inteligencia de negocios y analítica de datos. Se obtuvo como resultado que

las personas consideran que el mayor aporte de la inteligencia de negocios es para la toma de decisiones con un 95.83% de las respuestas.

Catorce personas de los encuestados consideraron unánimemente, que de la inteligencia de negocios y analítica de datos se obtienen proyecciones, descubrimientos de nuevos datos, conocimiento de los clientes y tomas de decisiones más oportunas. Estas opciones no representan lo mismo, más si están relacionadas. Hablamos entonces de la correlación de variables que en conjunto permiten mejorar el proceso de decisión.

**Q15 Seleccione la(s) opción(es) con qué  
asocia la inteligencia de negocios y  
analítica de datos:**





Opciones de respuesta	Respuestas	
Permite hacer proyecciones	58,33%	14
Permite realizar reportes ejecutivos	33,33%	8
Permite crear indicadores de desempeño	54,17%	13
Mejorar la eficiencia, reducir costos	66,67%	16
Toma de decisiones más precisas y oportunas	95,83%	23
Conocer el comportamiento de consumidores	58,33%	14
Mejorar el control sobre áreas funcionales de la empresa	37,50%	9
Para realizar descubrimiento con los datos	58,33%	14
Otro (especifique)	8,33%	2
Total de encuestados: 24		

### Anexo 3: Cálculo de muestra para poblaciones infinitas o desconocidas

La fórmula para calcular el tamaño de muestra cuando se desconoce el tamaño de la población es la siguiente:

$$n_{opt.} = \frac{Z^2 \times p \times q}{d^2}$$

Donde

Z = nivel de confianza,

p = probabilidad de éxito, o proporción esperada

q = probabilidad de fracaso

d = precisión (error máximo admisible en términos de proporción)

Consideraciones:

1. Cuando se desconoce la probabilidad de éxito esperada, se tiene que utilizar el criterio conservador ( $p = q = 0.5$ ), lo cual maximiza el tamaño de muestra.
2. Para el caso del nivel de confiabilidad de la muestra calculada (Z) existen algunos valores estandarizados en función del grado de confiabilidad asumido.  
 Para un: 99 % ----- z = 2,58 (Empleado con frecuencia)  
 95 % ----- z = 1,96 (El más empleado)  
 90 % ----- z = 1,64  
 Por tanto, se establece se asumirá un grado de confianza del 95%.
3. Toda expresión que se calcula contiene un error de cálculo debido a las aproximaciones decimales que surgen en la división por decimales, error en la selección de la muestra, entre otras, por lo que este error se puede asumir entre un 1 hasta un 10%; es decir, que se asume en valores de probabilidad correspondiente entre un 0.01 hasta un 0.1.  
 Por tanto, se establece un error máximo admisible del 5%

$$n_{opt} = \frac{1.96^2 \times 0.5 \times 0.5}{0.05^2} = 385$$