

# **Maestría en Gestión de Información**

## **Caracterización de Aspirantes para el Planteamiento de Estrategias de Captación de las Instituciones de Educación Superior**

**Angela Daniela Prieto Parra**

**Bogotá D.C., 8 de febrero de 2019**



**Caracterización de Aspirantes para el Planteamiento de  
Estrategias de Captación de las Instituciones de Educación  
Superior**

**Trabajo de grado para optar al título de Magister en Gestión de  
Información**

**PhD. Oswaldo Castillo Navetty  
Director trabajo de grado**

**Jurado:  
Olga Lucia Giraldo  
Sofía López Ruiz**

**Bogotá D.C., 30 de noviembre de 2018**

El trabajo de grado de Maestría titulada "Caracterización de Aspirantes para el Planteamiento de Estrategias de Captación de las Instituciones de Educación Superior", presentada por Angela Daniela Prieto Parra, cumple con los requisitos establecidos para optar al título de Magister en Gestión de Información.

Olga Lucia Giraldo

Sofía López Ruiz

PhD. Oswaldo Castillo Navetty

## **Agradecimientos**

Doy gracias a Dios por permitirme culminar este proyecto en el que él me guío y me embarcó sin estar en mis planes, y a mi familia por su apoyo incondicional.

## Tabla de contenido

RESUMEN .....	7
1. INTRODUCCIÓN .....	9
1.1. CONTEXTO PROBLEMÁTICA .....	9
1.2. PROBLEMA DE INVESTIGACIÓN .....	10
2. OBJETIVOS Y PREGUNTA DE INVESTIGACIÓN .....	12
2.1. OBJETIVO GENERAL.....	12
2.2. OBJETIVOS ESPECÍFICOS.....	12
2.3. PREGUNTA.....	13
3. FUNDAMENTACIÓN TEÓRICA .....	13
4. DESARROLLO .....	20
4.1. OBJETIVO 1 .....	21
4.2. OBJETIVO 2 .....	24
4.3. OBJETIVO 3 .....	35
4.4. OBJETIVO 4.....	43
4.4.1. VISUALIZACIÓN.....	46
5. CONCLUSIONES .....	47
6. TRABAJOS FUTUROS.....	48
7. PRESUPUESTO DE INVERSIÓN .....	48
8. EQUIPO INVESTIGADOR .....	49
REFERENCIAS.....	50

## **Tabla de ilustraciones**

<i>Ilustración 1 Pirámide de Inteligencia de Negocios (Davenport &amp; Prusak, 1998)</i>	14
<i>Ilustración 2 Fases KDD</i>	18
<i>Ilustración 3. Fases KDD</i>	19
<i>Ilustración 4. Técnicas Minería de datos</i>	20
<i>Ilustración 5. Metodología Propuesta</i>	21
<i>Ilustración 6. Resumen SNIES</i>	27
<i>Ilustración 7. Datos SNIES</i>	28
<i>Ilustración 8. Resumen SPADIES</i>	29
<i>Ilustración 9. Datos SPADIES</i>	30
<i>Ilustración 10. Resumen ICFES</i>	31
<i>Ilustración 11. Datos ICFES</i>	32
<i>Ilustración 12. Datos Agregados</i>	33
<i>Ilustración 13. Cruce información Saber11</i>	35
<i>Ilustración 14. Ciclo de los Datos</i>	36
<i>Ilustración 15. Maqueta Inicio</i>	37
<i>Ilustración 16. Maqueta SNIES</i>	37
<i>Ilustración 17. Maqueta SPADIES</i>	38
<i>Ilustración 18. Maqueta DANE</i>	39
<i>Ilustración 19. Maqueta OLE</i>	39
<i>Ilustración 20. Maqueta ICFES1</i>	40
<i>Ilustración 21. Maqueta ICFES2</i>	40
<i>Ilustración 22. Maqueta PERFIL</i>	41
<i>Ilustración 23. Definición perfil</i>	41
<i>Ilustración 24. Ejemplo Perfil</i>	43
<i>Ilustración 25 Fases KDD</i>	44
<i>Ilustración 26. Modelos de aprendizaje</i>	44

## RESUMEN

Actualmente las empresas requieren tomar decisiones de forma ágil y las Instituciones de Educación Superior (IES) no son la excepción, por esto mismo se propone el aprovechamiento de la información abierta, pública y privada, donde se puedan tomar decisiones más acertadas sobre el comportamiento de sus futuros estudiantes y las características de los mismos.

Dando respuesta a esta necesidad un insumo esencial son los exámenes gubernamentales que miden el conocimiento y las habilidades de los estudiantes que culminan un ciclo de formación, permitiendo determinar comportamientos de los futuros estudiantes, actualmente existen fuentes de información pública que permite perfilar a los graduados de la educación media y ayuda a las Instituciones de Educación Superior a definir el perfil de los estudiantes a los cuales desea llegar, esta información se puede trabajar por medio de la minería de datos y la inteligencia de negocios, permitiendo a los interesados acceder a cuadros de mando (DashBoards) dinámicos con la información necesaria para cada tomador de decisiones o su equipo de trabajo logrando generar estrategias más efectivas que se alineen con la misión de la Institución.

Tener una herramienta que permita conocer las características sociodemográficas y académicas de los graduados de la educación media dándole valor a la información, gestionarla y generar sistemas de evaluación, aplicación de herramientas y buenas prácticas para medir la gestión, es la respuesta para ubicar a la demanda objetivo de las Instituciones de Educación Superior, permitiendo seleccionar las características de estos estudiantes y conocer la ubicación y los nombres de los planteles donde se encuentran los estudiantes con estas características.

**Palabras clave:** gestión de información, estrategia, sociodemográfico, minería de datos.

## **Abstract**

Currently, companies must make decisions in an agile way and Higher Education Institutions (IES) are not the exception, therefore, the use of open, public and private information is proposed, where the most accurate decisions can be made about the behavior of their future students and the characteristics of them.

Answering this need an essential input are the exams that the half knowledge and skills of students that culminate in a training cycle, which allow the behavior of future students, that currently exist sources of information that allow to obtain an excellent performance. middle education and help the IES a profile of the students you want to reach, this information can be worked in the medium of data mining and business intelligence, to those interested accessing the tables of dynamic control (Dash Boards) with the necessary information for each decision-maker or his work team managing to generate more effective strategies that align with the mission of the Institution.

Having a tool that allows to know the sociodemographic and academic characteristics of the graduates of education means giving value to the information, managing and generating the evaluation systems, the application of the tools and the good practices to measure the management, is the answer for locate the objective demand of the IES, select the characteristics of these students and know the location and the names of the plants where the students with these characteristics are located.

**Key words:** information management, strategy, sociodemographic, data mining.



## 1. INTRODUCCIÓN

### 1.1. CONTEXTO PROBLEMÁTICA

Ponce, 2011, afirma que “... La **información** se entiende como una dimensión relevante de la **calidad**, ya que permite un **control** efectivo de los procesos institucionales para **tomar decisiones** sobre los ajustes y mejoras requeridos, y también porque la información sobre las universidades y sus resultados normalmente se revela espontáneamente sólo cuando se ajusta a los **parámetros de calidad reconocidos**.” (Feet & Rodrigez - Ponce, 2011, pág. 570)

Las Instituciones de Educación Superior plantean sus estrategias de captación en su mayoría por la experiencia para poder llegar a los futuros estudiantes, sin mirar información específica, solo por la percepción que se tiene de la población, asistiendo a ferias universitarias y actividades con los colegios a la que son invitadas las universidades en muchas ocasiones no se evalúa que tan viable o no es este colegio para mi objetivo como universidad.

Entonces los recursos se invierten a todo lo que alcance, la toma de decisiones basada únicamente en la experiencia es buena pero no es suficiente, se necesita información para enfocar mejor los esfuerzos.

*Así como Quiroga nos enfoca que “Sin información de calidad, sin los datos o los conocimientos necesarios, es imposible gestionar calidad.” (Aja Quiroga, 2002), se debe tener la unión entre la información y la experiencia para el planteamiento de las estrategias, viendo la fotografía del momento.*

*Las Instituciones de Educación Superior afrontan retos, y uno relevante es realizar su gestión comercial a través de actividades de reclutamiento que muestren atributos de valor, tangibles para captar potenciales estudiantes y gestionar la matrícula y retención del mismo*

*(Bejou D., Bejou A, 2012; Anctil, 2008)*

*Las estrategias de captación deben evolucionar al ritmo cómo evolucionan las generaciones, no puedo plantear una estrategia de captación como Institución de Educación Superior para personas que nacieron entre el 80 y el 2000 a la generación actual, porque son diferentes.*

*Esto lo han entendido las Instituciones de Educación Superior, pero en su gran mayoría no basan sus decisiones en información.*

*“El panorama del marketing educativo cambia de manera constante e impredecible, es por eso que las Instituciones de Educación Superior deben reaccionar de manera rápida y estratégica para implementar las buenas prácticas, como la planeación estratégica, gestión de marca, atraer y retener los potenciales estudiantes y los cambios demográficos y culturales.” (Bejou D. & Bejou A, 2012)*

Entonces es necesario hablar de Marketing Educativo siempre cambiante, las instituciones deben marcar la diferencia, tener reconocimiento en los lugares a los que van, buscan estar presentes en la mente de los futuros estudiantes, colegios y sus familias como una opción.

Las actividades de reclutamiento o de captación son un reto, atraer a la persona para estudiar en la Institución, en este caso hablaremos específicamente de actividades de captación.

No se busca plantear las estrategias de captación en esta metodología, se busca brindar una solución para que las Instituciones de Educación Superior tomen mejores decisiones y planteen mejores estrategias de captación inicialmente.

## **1.2. PROBLEMA DE INVESTIGACIÓN**

Para las Instituciones de Educación Superior – IES, es un reto conocer las características de sus potenciales estudiantes, con el fin de proponer actividades innovadoras para atraer candidatos y mejorar las estrategias de captación.

La caracterización de los aspirantes sin información válida, no permite enfocar las estrategias con mayor objetividad.

Las generaciones están en constante cambio y las estrategias de captación de las Instituciones de Educación Superior deben evolucionar al mismo ritmo, para poder plantear estas estrategias es necesario conocer el público objetivo a fondo, no solo basarse en la experiencia que se tiene al trabajar con este medio y los datos internos, sino buscar información que permita determinar el perfil de la población actual, por medio de herramientas que transformen los datos en información y sean de ayuda en la toma de decisiones, integrando los datos externos, internos y el conocimiento del medio educativo de los tomadores de decisión.

Durante los últimos años la gestión de la información ha adquirido gran valor debido a su capacidad para generar y para mejorar el proceso de toma de decisiones. La mejora en el desempeño de las organizaciones es gracias a un análisis óptimo de la información que permite mejorar la toma de decisiones y obtener ventajas competitivas. Las utilidades incrementan a medida que se va alineado la gestión de

la información a la estrategia de la empresa. Los sectores relacionados con las tecnologías de la información y la comunicación desempeñan un papel particularmente importante dentro de este esquema. En este contexto, en la sociedad de la información se propone que las tecnologías de información y comunicaciones constituyen el motor del progreso y del desarrollo. (Rodríguez - Ponce, 2009)

Se busca dar a conocer la importancia y el impacto que tiene para las Instituciones de Educación Superior, la implementación de una metodología que permita orientar el análisis de información como medio para mejorar la toma de decisiones respecto a las estrategias de captación, apropiando de manera estratégica el perfil de los aspirantes.

En la Conferencia Mundial sobre la Educación Superior "La educación superior en el siglo XXI: Visión y acción", organizada en París por la UNESCO en octubre de 1998 se expresó el deber de generalizar en la mayor medida posible la utilización de las nuevas tecnologías para que ayuden a los establecimientos de educación superior a reforzar el desarrollo académico, a ampliar el acceso, a lograr una difusión universal y extender el saber, y a facilitar la educación durante toda la vida. Los gobiernos, los establecimientos de enseñanza y el sector privado deberán procurar, enfatizar el informe, en que se faciliten en un nivel suficiente infraestructuras de informática y de redes de comunicaciones, servicios informáticos y formación de recursos humanos (UNESCO, 1998)

Este planteamiento de la UNESCO deja un reto para las Instituciones de llegar a la mayor población posible, que se identifique con la oferta académica de la misma, entonces es necesario conocer donde encontrar este público y solo se logra con información de calidad; sin información de calidad, sin los datos o los conocimientos necesarios, es imposible gestionar (Aja Quiroga, 2002), las TI son parte central de la mejora de las Instituciones de Educación Superior (Lau & Yen, 2014).

Tableau plantea que, (Tableau, 2018)

*“... las personas comienzan a trabajar con datos de una manera más natural y afirma que la tecnología ha reemplazado el scripting y las tablas dinámicas con interfaces intuitivas de arrastrar y soltar, además asegura que, a partir del 2017, la interfaz de acceso a los datos comenzará a percibirse de un modo aún más natural por las mejoras en la generación y el procesamiento del lenguaje natural. Las interfaces de lenguaje natural son una nueva incorporación a los recursos de la BI. Pueden hacer que los datos, los gráficos y los dashboards sean aún más accesibles. Permiten a las personas*

*interactuar con ellos utilizando texto y lenguaje natural. Según Gartner, esta es la “próxima fase en la evolución de la creación de informes estándar hacia la narración de historias”.*

¿Cuáles características son relevantes en la definición del perfil de los aspirantes a ingresar a las Instituciones de Educación Superior, en qué lugares del país y en qué colegios se encuentran los futuros estudiantes que cumplen con el perfil buscado?

La gestión de información es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios. (Sinnexus, 2017), poder apoyar la toma de decisiones organizando los datos que permita mejorar y aplicar buenas prácticas es una buena opción para apoyar los procesos y robustecerlos.

Como lo expresó en su momento el ministro TIC de Colombia, David Luna:

*“No importa si su empresa es de lo digital o de lo físico, no importa si usted hace páginas web o zapatos, o hace aplicaciones o pantalones, hay que digitalizarse. ¿Para qué? Para prestarle un mejor servicio a su cliente (...)”*  
(EL TIEMPO, 2017)

## **2. OBJETIVOS Y PREGUNTA DE INVESTIGACIÓN**

### **2.1. OBJETIVO GENERAL**

Definir la metodología y las técnicas que se deben aplicar en la definición del perfil de los aspirantes a ingresar a las Instituciones de Educación Superior, dando una herramienta para definir mejor las estrategias de captación utilizadas por estas instituciones, respondiendo a la dinámica generacional.

### **2.2. OBJETIVOS ESPECÍFICOS**

1. Definir las fuentes de información abiertas, públicas y privadas, que aportan valor para la toma de decisiones en la construcción del perfil adecuado de los aspirantes de cada Instituciones de Educación Superior basados en información.
2. Determinar las características que conforman el perfil de un aspirante a la Educación Superior.

3. Detallar la metodología que permita construir el perfil de los futuros estudiantes de las Instituciones de Educación Superior.
4. Seleccionar las técnicas que permiten llevar a cabo la metodología propuesta.

### **2.3. PREGUNTA**

La evaluación de los candidatos y el público objetivo es fundamental para las Instituciones de Educación Superior, entonces se busca responder:

¿Cuáles características son relevantes en la definición del perfil de los aspirantes a ingresar a las Instituciones de Educación Superior, en qué lugares del país y en qué colegios se encuentran los futuros estudiantes que cumplen con el perfil buscado?

### **3. FUNDAMENTACIÓN TEÓRICA**

La gestión de la información ocupa, cada vez más, un espacio mayor en la economía de los países a escala mundial. (Aja Quiroga, 2002), desde el año 95 Fernández-Molina afirma que los profesionales de la información y sus instituciones son un factor indispensable para la permanencia de la organización en el mercado (Fernandez - Molina, 1995), esto significa que la información se convirtió en el factor de producción primaria (Ikram, Mesut, & Abdullah, 2011).

La información es la materia prima para la inteligencia de Negocios (Business intelligence – BI), esta es una alternativa que permite trabajar los datos y obtener información con potencial, es una solución que integra todas las “actividades relacionadas a la organización y entrega de información, de análisis del negocio, por medio de la minería de datos, administración del conocimiento, aplicaciones analíticas, sistemas de reportes y principalmente Data Warehousing y modelado multidimensional” (Reyes M., 2004), esta es una herramienta que permite apoyar la toma de decisiones, apoyándose en DSS (Sistemas de soporte a la toma de decisiones), Modelado Multidimensional, Data Warehouse, Data Marts y Data Mining, OLAP (On line analitic procesing), herramientas de consulta y reporte de datos y portales de información empresarial.

La inteligencia de negocios entonces permite a las empresas definir mejor sus estrategias por medio de información que aporta valor, actualmente existen herramientas de inteligencia de negocios que permiten analizar los datos estructurados y no estructurados, como lo son Power BI, SAS, Oracle ODI, sap, Tableau, Microsoft SSIS, Sisense, Thoughtspot, Qlik, Informatica, Information

Builders, Ibm Datastage, Microstrategy, Talend, entre otros como puede ser la construcción propia de cada empresa.

Toda universidad cuenta con diversos sistemas que soportan la información para la toma de decisiones, algunos son digitales otros son documentos físicos pero tienen la información para mejorar sus estrategias, específicamente hablando de la promoción de los programas, si se logra identificar la efectividad de las comunicaciones en redes sociales, home page, CRM y sistemas de información en general, unificar estos datos analizarlos y determinar el comportamiento de los aspirantes actualmente y a futuro se garantizaría aún más la aceptación y matriculas efectivas que no generen deserción (Ikram, Mesut, & Abdullah, 2011).

Para esto se proponen sistemas como lo son la inteligencia de negocios que consiste en, (Jones, 2017):

- Integrar la Información
- Permitir la disponibilidad de la información en tiempo real
- Mejorar los procesos
- Realizar un análisis de la información

En resumen, BI es un conjunto de tecnologías y procesos que permiten a las personas de la organización, tener acceso a datos y su análisis. Sin personas para interpretar información y actuar con base a ella, la inteligencia de negocios nada logra. BI es menos sobre tecnología que creatividad, cultura e individuos considerando la información como un activo crucial. La tecnología posibilita la BI, pero a veces, demasiada importancia puesta en la tecnología puede sabotear iniciativas de inteligencia de negocios. Son las personas quienes harán de los esfuerzos de BI un enorme éxito o un profundo fracaso. (Howson, 2009)



Ilustración 1 Pirámide de Inteligencia de Negocios (Davenport & Prusak, 1998)

La inteligencia de negocios como se ve en el grafico está compuesta por diferentes pasos que permiten tener la explotación de la información en un Sistema de Soporte a la Decisión – DSS, que es una de las herramientas más emblemáticas del Business Intelligence (BI), que permiten resolver gran parte de las limitaciones de los programas de gestión, por medio de informes dinámicos, flexibles e interactivos, no requiere conocimientos técnicos, tiempo de respuesta, integración de sistemas, perfilamiento de la información y la información histórica. (Sinnexus, 2017)

La extracción, transformación y carga en sus siglas en ingles ETL, permite organizar los datos, estandarizarlos y darles la posibilidad de ser transformados en información valiosa, en esta labor la minería de datos es una técnica y tecnología que permite explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto. (Molina Félix, 2002)

Robert Brunet (Brunet, 2016), expresa en su artículo que la minería de datos utiliza métodos como son la inteligencia artificial, el aprendizaje automático, la estadística y sistemas de base de datos. Este identifica que el proceso de minería de datos tiene 4 etapas, determinación de los objetivos, trata el tipo de información que el cliente desea extraer de la base de datos; procesamiento de los datos, consiste en seleccionar, limpiar, enriquecer, reducir y transformar la base de datos; determinación del modelo y análisis de los resultados.

Una solución de BI está basada también en la exploración de los datos que el big data genera por tal razón es importante saber integrar los datos de estas fuentes externas con las fuentes internas de las empresas.

*“Actualmente no hemos puesto aún el pie en Marte, pero tecnologías como Big Data, BI y la analítica hacen del Gobierno IT un proceso empresarial que, añadiendo la ventaja de la automatización y el cálculo masivo de datos en muy poco tiempo, cada vez se parece más a la forma de analizar del cerebro humano”* (haycanal.com, 2017)

Los Sistemas de Información (SI) son fundamentales para la gestión, fundamentalmente en la toma de decisiones y también para informar las acciones de las universidades a la sociedad. (Barcos, 2008). Los SI en educación deben proporcionar una amplia gama de datos, a los estudiantes, docentes, investigadores y a la administración y gestión para la solución de diversos problemas y para evaluar el efecto de las acciones educativas internas y sobre la sociedad (Barcos, 2008)

Según (Barcos, 2008) la información se entiende de varias formas, pero generalmente no es solo datos estructurados también texto libre, gráficos, sonidos, fotografías, videos, films, etc. La información de una Instituciones de Educación

Superior es toda información que se relaciona con los hechos de formación y es muy diversa, citando los siguientes casos:

- Datos demográficos e información sobre determinantes sociales, culturales, económicos y ambientales de la educación;
- Datos de alumnos atendidos, ingresantes, ingresantes efectivos, regulares, perfiles de ingreso y graduación, movilidad, retención, deserción y similares;
- Resultados provenientes de evaluaciones y estadísticas de todo tipo,
- Datos del personal docente, de investigación, extensión y transferencia y datos del personal no docente en cantidad, calidad, dedicación, niveles de capacitación, edad, sexo, antigüedad, etc.;
- Características de los planes y programas de estudio de pre grado, grado, postgrado y doctorados;
- Poblaciones atendidas y no atendidas;
- Presupuestos, costos, estructuras patrimoniales y proyecciones económicas y financieras;
- Edificios y recursos materiales de diverso tipo y, en su caso, con sus respectivos inventarios;
- Acciones complementarias a la función docencia
- Acciones de extensión, investigación, bienestar estudiantil, transferencia, etc.;
- Registros de asistencia, de progreso en los estudios y similares
- Archivo de docentes y alumnos;
- Responsables estructurales, funcionales o no, por la información y los sistemas de información
- Todo tipo de información cruzada y de relación inherente a la actividad educativa.

En su artículo (Barcos, 2008, pág. 5) es muy claro al expresar que hay evaluaciones de calidad y acreditación que tienen expresiones como:

*“Entre las principales falencias encontradas se verifica el hecho de que los registros y la información para las decisiones se encuentra atrasada, desordenada, repetida y no es confiable ni completa. Se recomienda perfeccionar los sistemas de información de modo que sirvan para las decisiones operativas de la institución y como apoyo al proyecto institucional”.*

*“Entre las debilidades más importantes aparecen los sistemas de información implementados en la universidad. Entre ellos, la información contable no es fácilmente comparable con el presupuesto. El control presupuestario es malo.*



*No existe plan de capacitación ni detección de necesidades en los procesos de trabajo y, en particular, con relación a los sistemas logísticos o de apoyo y de información.*

*La institución carece de indicadores y de sistemas de información adecuados para evaluar calidad y productividad de los procesos sustantivos. En algunas áreas, la información parece suficiente y se cuentan con sistemas de información bastante consolidados, pero existen problemas para satisfacer nuevas necesidades de información, para planificar y evaluar a la institución como un conjunto.*

*Ausencia generalizada de un análisis sistemático de los procesos de gestión y servicios centrales y de sus resultados y escasa atención a las opiniones de los usuarios que requieren información para el trabajo cotidiano y para sus decisiones.*

*Escasa coordinación vertical y horizontal en los servicios y en los sistemas de información para la gestión y existencia de información fragmentada e incompleta.” (CONEAU, 1997/2006)*

Esto afirma la importancia de tener información que apoye la toma de decisiones.

Viviana Duro y Carlos Pérez (2016), expresan en su artículo que las universidades reivindican su papel de protagonismo en la sociedad y el ejercicio de su libertad académica y de investigación, como contrapartida se comprometen a rendir cuenta de sus actuaciones, aumentando la transparencia en su gestión, e identificando las demandas de los distintos colectivos sociales, internos y externos, para darles respuesta, concluyen de su investigación que el enorme caudal de información generado a diario tanto por las universidades internamente como por su entorno, ha elevado aún más la importancia de la gestión de la información y de la aplicación de las herramientas de Inteligencia de Negocios como lo son los Sistemas de Soporte a las Decisiones.

*“Lograr la gestión online de todos los procesos de apoyo de la universidad tendría, sin duda, un impacto trascendental, supondría inmediatez en la información, así como niveles de respuesta que nada tienen que ver con los actuales. Eliminar la espera, lograr el aprovechamiento múltiple del tiempo, tener la posibilidad de saber en cada momento el estado de cada gestión, pondría a la Universidad en un escalón superior a cualquier resultado en la gestión actual de la administración pública en Cuba.” (Duro & Perez, 2016, pág. 41)*

La economía del conocimiento describe una etapa particular del desarrollo del sistema capitalista, basada en el conocimiento, que sucede a una fase de acumulación del capital físico. Tal como Marx había predicho a mediados del siglo

XIX, parece ser que el conocimiento está sustituyendo a la fuerza de trabajo y que la riqueza creada se mide cada vez menos por el trabajo en su forma inmediata, mensurable y cuantificable, y depende cada vez más del nivel general de la ciencia y del progreso de la tecnología. (UNESCO, 2005)

Gestión estratégica es otra rama muy importante de la Gestión de Información que permite orientar la estrategia mejorando la productividad, enfocándose en la misión y visión de la institución permite la mejor determinación de objetivos y no solo en general como empresa sino como áreas de trabajo también permite organizar y priorizar la estrategia enfocados en la misión; es una necesidad la planeación prospectiva estratégica, para evitar luchar contra el azar que deja muchas acciones al tiempo, al desperdicio de recursos, a los gastos excesivos, cuando algo que es urgente se ataca demasiado tarde. Se precisan las características de una planeación prospectiva estratégica en tanto pensar escenarios y planear escenarios, así como sobre algunas guías de aplicación. (Arroyo,e.t, 2015)

#### Descubrimiento del Conocimiento en Bases de Datos (KDD)

Camacho trae en su artículo Análisis del desempeño del profesorado universitario en el uso de MOODLE a través de técnicas de minería de datos: propuestas de necesidades formativas la referencia a Fayyad donde expresa, el Descubrimiento del Conocimiento en Bases de Datos (Knowledge Discovery in Databases, KDD), es “el proceso no trivial de identificar patrones válidos, novedosos y potencialmente útiles y en última instancia, comprensible a partir de los datos” (Camacho Chacón, Zapata Gonzalez, Menéndez Domínguez, & Canto Herrera , 2018)

También referencia a Hernández quien presenta el KDD en 5 fases:



*Ilustración 2 Fases KDD  
Fuente: Construcción propia*

Camacho adicionalmente menciona en su artículo, expresa que la experimentación con las técnicas de minería de datos es una de las partes fundamentales para la extracción de conocimiento de la investigación y se resume en la siguiente ilustración: (Camacho Chacón, Zapata Gonzalez, Menéndez Domínguez, & Canto Herrera , 2018)

### Integración y recopilación

- Determinar las fuentes de información útiles y donde conseguirlas. Luego se transforman los datos a un formato común.

### Selección, limpieza y transformación

- Eliminar o corregir los datos incorrectos y decidir el paso a seguir con los datos incompletos, identificar los datos, aquellas variables o atributos que son importantes y hacer más fácil la tarea de minería y los resultados sean más útiles.

### Minería de datos

- Se toma la decisión (clasificar, agrupar, entre otras) y elige el método a utilizar.

### Evaluación e interpretación

- Expertos evalúan y analizan los patrones, y de ser necesario se vuelve a fases anteriores para una nueva iteración.

### Difusión y uso

- Hacer uso del nuevo conocimiento y comunicarlo a todos los posibles usuarios. Este conocimiento se suele utilizar en procesos de toma de decisiones.

*Ilustración 3. Fases KDD  
Fuente: Construcción propia*

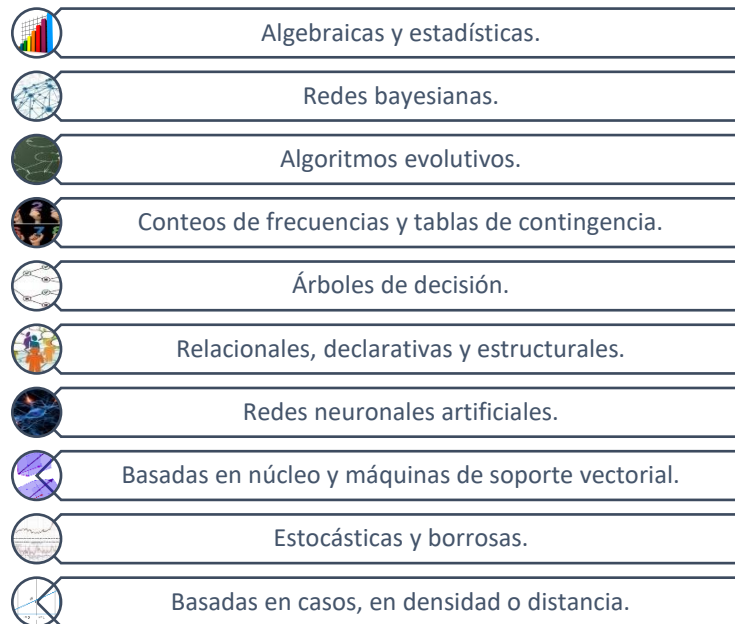
Jiawei Han (Han, Kamber, & Pei, 2012), define que la minería de datos se traduce a extracción de conocimiento a partir de datos, extracción de conocimiento, análisis de datos / patrones, datos arqueología y dragado de datos, la extracción de conocimiento a partir de datos.

Se definen los siguientes pasos para trabajar con los datos:

1. Limpieza de datos (para eliminar ruidos y datos inconsistentes).
2. Integración de datos (donde se pueden combinar múltiples fuentes de datos)
3. Selección de datos (donde los datos relevantes para la tarea de análisis se recuperan de la base de datos)

4. Transformación de datos (donde los datos se transforman y se consolidan en formularios apropiado para la minería al realizar operaciones de resumen o agregación)
5. Minería de datos (un proceso esencial en el que se aplican métodos inteligentes para extraer patrones de datos)
6. Evaluación de patrones (para identificar los patrones realmente interesantes que representan el conocimiento basado en medidas de interés)
7. Presentación del conocimiento (donde las técnicas de visualización y representación del conocimiento. Se utilizan para presentar conocimiento minado a los usuarios)

Camacho (Camacho Chacón, Zapata Gonzalez, Menéndez Domínguez, & Canto Herrera , 2018) hace un resumen de las técnicas más populares:



*Ilustración 4. Técnicas Minería de datos  
Fuente: Construcción propia*

#### **4. DESARROLLO**

La información de los futuros estudiantes en su gran mayoría es pública, es abierta a todo el público, cualquier persona la puede trabajar.

Actualmente no hay una solución que tenga estas fuentes analizadas en un solo lugar, se quiere entonces entregar la información actualizada en un solo lugar.

Luego de la revisión de la fundamentación teórica se revisaron temas como lo son la Inteligencia de Negocios, es la gran base de este proyecto; la planeación estratégica, para la alineación de la información a la estrategia de las Instituciones de Educación Superior; la gestión de la información, que es la base de todo lo revisado; Marketing, para conocer cómo se define un perfil y como se llega al público objetivo; Gestión del conocimiento, como aprendo de la historia, de las experiencias de los tomadores de decisión; la Minería de Datos, que sin sus técnicas no se puede trabajar los datos y presentar la información.

Buscando alinear los objetivos del proyecto con la metodología se propone la siguiente metodología de trabajo:

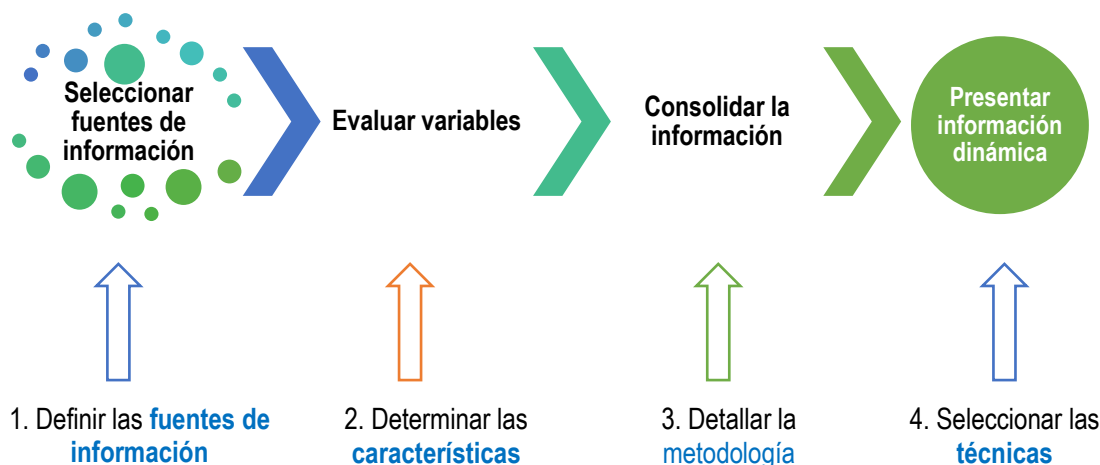


Ilustración 5. Metodología Propuesta  
Fuente: Construcción propia

Esta es la manera de abordar los cuatro objetivos, con el fin de llegar a dar respuesta a la problemática planteada.

#### 4.1. OBJETIVO 1

*Definir las fuentes de información abiertas, públicas y privadas, que aportan valor para la toma de decisiones en la construcción del perfil adecuado de los aspirantes de cada Institución de Educación Superior basados en información.*

En la búsqueda de la información que permita definir las características de un perfil de un aspirante que se ajuste a lo buscado por la Institución de Educación Superior, se encuentran múltiples fuentes públicas y privadas, como lo son datos abiertos, información de cada institución y la información que el Ministerio de Educación pública en sus diferentes Sistemas como lo son ICFES, SNIES, SPADIES y otras fuentes distintas al Ministerio de Educación como el DANE.

Como características de estos sistemas de información tenemos que:

En el Sistema de aseguramiento de la calidad de la educación superior se encuentra la información donde las Instituciones de Educación Superior rinden cuentas ante la sociedad y el Estado sobre el servicio educativo que prestan, proveyendo información confiable a los usuarios del servicio educativo y se propicia el auto examen permanente de instituciones y programas académicos en el contexto de una cultura de la evaluación, el sistema lo definen la información, evaluación y fomento. (MEN, 2018)

Este sistema lo componen los siguientes sistemas de información:

- Sistema Nacional de Información de Educación Superior (**SNIES**), el cual ofrece datos confiables sobre las instituciones de educación superior en Colombia y los programas que ofrecen.
- Observatorio Laboral para la Educación - **OLE**, que ofrece un seguimiento permanente de los graduados de la Educación Superior en Colombia.
- Sistema de Información para el Aseguramiento de la Calidad (**SACES**), el cual contiene información para el proceso de Registro Calificado de programas académicos.
- Sistema de Prevención y Análisis de la Deserción en las Instituciones de Educación Superior (**SPADIES**), que permite el seguimiento a cada estudiante para calcular el riesgo de la deserción y prevenirlo.

El Instituto Colombiano para la Evaluación de la Educación (**ICFES**) , aplica diferentes evaluaciones que miden la calidad de la Educación en Colombia como los es para los bachilleres el Examen Saber11, el Examen Saber Pro para los profesionales, también se aplican pruebas de calidad durante el proceso de formación de los estudiantes de colegio como son el Saber 3, 5 y 9 que permiten identificar las competencias y habilidades que todos los estudiantes colombianos desarrollan durante su trayectoria escolar (ICFES, 2018).

Los colegios, así como las universidades tienen medidas de calidad de la educación lideradas por el Ministerio de Educación por medio del ICFES, quienes generan una clasificación de los colegios de acuerdo a los resultados del examen Saber 11 presentado por los estudiantes de estos colegios.

Cada Instituciones de Educación Superior tiene su sistema de información que contiene datos importantes para la toma de decisiones con respecto a su calidad de la educación y el perfil de los estudiantes actuales y futuros, para definir un objetivo de los estudiantes que se desea tener es importante lograr las mayores fuentes de información interna y externa que permitan plantear estrategias para alcanzar los objetivos.

Específicamente la información de interés del ICFES son los resultados agregados de los Colegios y las instituciones que certifican el título de Bachiller Colombiano que permiten tener una visión general de la Educación en Colombia por medio de los resultados del Examen Saber11 a partir del año 2011 al 2018.

El ICFES permite conocer la información del Examen Saber11 por secretarías, es material para trabajar la estadística a nivel geográfico y compararlos con el resultado a nivel nacional en las áreas de evaluación del examen.

El ICFES también presenta la información agregada de la Clasificación de Planteles, que presenta la información de cada una de los establecimientos educativos dando una calificación de A+, siendo esta la mayor, A, B, C y D siendo esta la más baja.

Actualmente en Colombia la página Web del ICFES permite consultar la información de interés de los establecimientos educativos por medio de búsquedas en el espacio de resultados, permite seleccionar el periodo de interés; el examen saber 11 cuenta con dos versiones, antes del 2014-2 y 2014-2 y posterior, esto se dio a partir del segundo periodo de 2014 donde se realizaron cambios en la evaluación del examen.

En esta búsqueda de información se encuentra dos tipos de bachillerato Grado 11 o de Grado 26 (ciclo para adultos), los establecimientos de educación media los identifica el ministerio de educación por medio de un código, el código Dane anteriormente se identificaban por el código ICFES.

La herramienta que publica los resultados del examen Saber 11 permite conocer la información de un establecimiento o descargar todos los establecimientos de un municipio o departamento.

Estos datos en el caso de no estar públicos en la página web del ICFES se pueden solicitar de forma gratuita al ICFES por medio de una solicitud escrita, esta entidad dará respuesta a la solicitud y entregará la información por medio de correo electrónico o medio magnético dependiendo del tamaño de los archivos. Para las

Universidades esta solicitud de información es sencilla de realizar dada la naturaleza de sus instituciones, el acceso a esta información no tiene limitaciones por parte del ICFES o el Ministerio de Educación Nacional - MEN.

**SINEB** es una herramienta que permite conocer la información de los colegios en Colombia, actualmente se encuentra fuera de servicio por esta razón no será tenida en cuenta como fuente de información, esta fuente de información es importante pero no es vital puesto que hay otras fuentes que permiten tener acceso a información similar, brinda datos sobre los colegios de Colombia permitiendo medir costos de matrícula, ubicación, tipo de institución (público o privado), Nombre del colegio, Dirección, Teléfono, Departamento, Municipio, Estado, Tipo, Calendario, Sector, Zona EE, Rector, Género, Carácter, Especialidad, Niveles, Grados, Etnias, Discapacidades, Capacidades Excepcionales, Modelos, Idiomas, Matrícula Contratada, Jornada, Cantidad de alumnos, Tarifas y Sedes.

El **DANE**, presenta información que ayuda a definir información sociodemográfica a nivel educativo, permite verificar información por medio de las estadísticas que esta organización ya ha tratado y permite tener un mejor contexto de la realidad colombiana para la toma de decisiones.

La información de cada institución es importante para conocer la población actual y poder identificar que quiere mantener o modificar del perfil de ingreso de los estudiantes de pregrado.

Redes sociales, fuente muy importante, se propone usarla a futuro, por lo robustas y porque los datos tienen un costo, si la institución que acoge la solución lo desea hacer en su momento se hará la evaluación pertinente.

Las fuentes de información pueden variar, cambiar las plataformas donde se encuentra la información y se debe acomodar a las disposiciones de información del momento de la implementación y la actualización de la información cuando el proyecto esté en marcha.

Luego de identificar las fuentes de información es necesario determinar que se usara y que no, para este fin se da paso al siguiente objetivo de investigación:

#### **4.2. OBJETIVO 2**

*Determinar las características que conforman el perfil de un aspirante a la Educación Superior.*

Esta información la reportan las Instituciones semestre a semestre, actualmente se cuenta con información desde el 2000-1 al 2017-2. En el momento que se



implemente la solución estos periodos de información se deben evaluar para determinar de qué año a qué año se tendrá en cuenta la información a utilizar.

Se propone tener información a partir del año 2014-2, cuando el Saber 11 cambio a la fecha para tener los datos más homogéneos.

Para definir las características del perfil, es necesario determinar la información sociodemográfica del público objetivo, según (Vicente Molina & Echevarria Gonzalez de Durana, 2003), los factores sociodemográficos aluden a características objetivas del consumidor y, por tanto, son perfectamente identificables y medibles; se trata de variables especialmente útiles para la segmentación del mercado, según esta definición es todo lo que nos permita conocer el comportamiento de la demanda, queriendo decir que se debe tener un conocimiento previo del medio de la educación para definir estas características.

Siguiendo esta definición es importante identificar las siguientes variables resultado de la evaluación de las fuentes de información:

¿Cuáles son los Resultados de evaluación de la educación media en Colombia?,  
¿Cuáles son las variables de evaluación y los tipos de exámenes que existen?

Respondiendo a las anteriores preguntas se conocen dos exámenes que se aplicaron a los bachilleres en Colombia, A partir del 2014-2 (corresponde al segundo semestre del año 2014) se viene aplicando el examen Saber11 con 5 áreas de evaluación, Matemáticas, Lectura Crítica, Ciencias Naturales, Sociales y ciudadanas e Inglés.

Y 2014- 1 y anteriores donde se evaluaban las áreas Matemáticas, Lenguaje, Física, Biología, Química, Inglés y Sociales.

Para esta aplicación se limita la información del examen Saber 11 a partir del periodo 20142 al 20172 como los periodos con información más relevante de la población actual, bien sea para definir comportamientos futuros o para identificar desempeño académico.

Del ICFES se toman los informes donde se publica la clasificación de los planteles de educación media y los reportes de los resultados del examen por planteles donde se resume los resultados en las áreas de evaluación del examen.

En la definición de la población objetivo se requiere tener un panorama de su círculo más cercano, para este efecto los informes de la población como los del DANE permiten conocer el panorama nacional.

El SNIES brinda la información del comportamiento de las Instituciones de Educación Superior en los procesos de admisión y formación, identificando el comportamiento actual y las tendencias que se han presentado versus el perfil de

la población. Con características como edad, sexo, estrato, colegio de bachillerato, valor de pensión del colegio, desempeño de estos colegios en el Saber11, conformación del núcleo familiar y nivel de estudios de los padres del estudiante.

La información abierta que tienen los entes estatales que miden la calidad de la educación en Colombia como actualmente lo hace el SNIES y SPADIES, permite determinar cuál es el camino a tomar según el comportamiento de las poblaciones que ingresan a la educación superior, la tasa de deserción y el éxito de quienes logran el título profesional o técnico si este es el interés.

Para definir con mayor detalle la información de los futuros estudiantes hoy es sumamente importante tener información de las redes sociales, comportamientos en redes e internet, como lo es Facebook, Twitter, Instagram, Youtube, NETFLIX, entre otras que nos permitan conocer más al detalle los gustos de los futuros estudiantes y llegar a ellos como la respuesta a su necesidad de educación, como lo menciona Caminos en su libro Dirección de Marketing. Fundamentos y aplicaciones (Camino & De Garcillán, 2012) donde menciona que con la realización de un buen marketing, no hay necesidad de vender, pues la gente comprara la oferta que satisface sus necesidades.

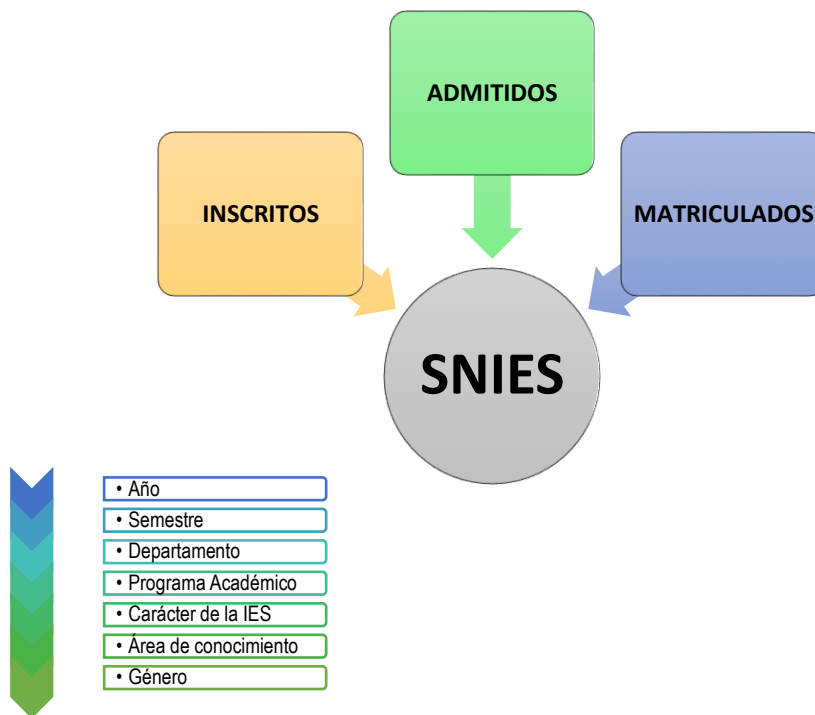
Esto se puede determinar cómo lo menciona el Doctor Carlos Adolfo Piña en la entrevista realizada por Cacho por medio de su estudio en Minería Social que permite conocer la información de los medios digitales actuales y realizar análisis del comportamiento de las poblaciones (Cacho, 2016), cuando la información es demasiado grande y cumple con características correspondientes del Big Data entonces aplicar diferentes algoritmos de inteligencia artificial o machine learning para determinar el tratamiento de los datos y como resultado obtener características de inclinación a ciertas áreas del conocimiento y lograr llegar de forma más asertiva a los futuros estudiantes de la Institución de Educación Superior.

Estos son los datos de las fuentes de información que se identificaron anteriormente:

El Sistema Nacional de Información de Educación Superior (SNIES) entrega la información de los Inscritos, Admitidos y Matriculados a los programas de pregrado con las siguientes variables a partir del periodo 2000-1 al 2017-2.

SNIES- que presenta la información de los Inscritos, Admitidos y Matriculados en Instituciones de Educación Superior y presenta la información del año, semestre, departamento, programa académico, carácter de la Institución de Educación Superior (Técnica, Tecnología, Profesional...), Área del conocimiento y el Género del estudiante.

Ilustración 6. Resumen SNIES  
Fuente: Elaboración propia



INSCRITOS		ADMITIDOS		MATRICULADOS 1ER CURSO	
PK	CODIGO INST	PK	CODIGO INST	PK	CODIGO INST
	AÑO PERIODO IES PADRE Institución de Educación Superior (IES) PRINCIPAL Sector IES Carácter IES CODIGO DEPARTAMENTO DEPARTAMENTO CODIGO MUNICIPIO MUNICIPIO SNIES PROGRAM NIVEL DE FORMACION NIVEL ACADEMICO METODOLOGIA AREA DE CONOCIMIENTO NUCLEO BASICO CODIGO DEPARTAMENTO DEPARTAMENTO CODIGO MUNICIPIO MUNICIPIO GENERO CANTIDAD		AÑO PERIODO IES PADRE Institución de Educación Superior (IES) PRINCIPAL Sector IES Carácter IES CODIGO DEPARTAMENTO DEPARTAMENTO CODIGO MUNICIPIO MUNICIPIO SNIES PROGRAM NIVEL DE FORMACION NIVEL ACADEMICO METODOLOGIA AREA DE CONOCIMIENTO NUCLEO BASICO CODIGO DEPARTAMENTO DEPARTAMENTO CODIGO MUNICIPIO MUNICIPIO GENERO CANTIDAD		AÑO PERIODO IES PADRE Institución de Educación Superior (IES) PRINCIPAL Sector IES Carácter IES CODIGO DEPARTAMENTO DEPARTAMENTO CODIGO MUNICIPIO MUNICIPIO SNIES PROGRAM NIVEL DE FORMACION NIVEL ACADEMICO METODOLOGIA AREA DE CONOCIMIENTO NUCLEO BASICO CODIGO DEPARTAMENTO DEPARTAMENTO CODIGO MUNICIPIO MUNICIPIO GENERO CANTIDAD
MATRICULADOS		GRADUADOS			
PK	CODIGO INST	PK	CODIGO INST		
	AÑO PERIODO IES PADRE Institución de Educación Superior (IES) PRINCIPAL Sector IES Carácter IES CODIGO DEPARTAMENTO DEPARTAMENTO CODIGO MUNICIPIO MUNICIPIO SNIES PROGRAM NIVEL DE FORMACION NIVEL ACADEMICO METODOLOGIA AREA DE CONOCIMIENTO NUCLEO BASICO CODIGO DEPARTAMENTO DEPARTAMENTO CODIGO MUNICIPIO MUNICIPIO GENERO CANTIDAD		AÑO PERIODO IES PADRE Institución de Educación Superior (IES) PRINCIPAL Sector IES Carácter IES CODIGO DEPARTAMENTO DEPARTAMENTO CODIGO MUNICIPIO MUNICIPIO SNIES PROGRAM NIVEL DE FORMACION NIVEL ACADEMICO METODOLOGIA AREA DE CONOCIMIENTO NUCLEO BASICO CODIGO DEPARTAMENTO DEPARTAMENTO CODIGO MUNICIPIO MUNICIPIO GENERO CANTIDAD		

Ilustración 7. Datos SNIES  
Fuente: Elaboración propia

SPADIES - permite conocer un contexto más claro del núcleo familiar y económico del estudiante.



Ilustración 8. Resumen SPADIES  
Fuente: Elaboración propia

APOYO ICETEX		APOYO U		CLASIFICACION ICFES		EDAD		NUMERO DE HERMANOS	
PK	PERIODO	PK	PERIODO	PK	PERIODO	PK	PERIODO	PK	PERIODO
	RECIBIO NO RECIBIO		RECIBIO NO RECIBIO		SIN INFORMACION BAJO MEDIO ALTO		15 ó menos años 16 a 20 años SIN INFORMACION 21 a 25 años 26 ó más años	1 2 3 4 > 4 SIN INFORMACION NINGUNO	
POSICION HERMANOS		TIPO APOYO ICETEX		TRABAJABA		SEXO			
PK	PERIODO	PK	PERIODO	PK	PERIODO	PK	PERIODO		
	SIN INFORMACION PRIMERO SEGUNDO TERCERO CUARTO QUINTO POSTERIOR AL QUINTO		NO RECIBIO CREDITO LARGO PLAZO MEDIANO PLAZO ACCES OTRO		TRABAJABA NO TRABAJABA SIN INFORMACION		SIN INFORMACION MUJER HOMBRE		

ESTRATO		AREA DE CONOCIMIENTO		NUMERO DE PERSONAS	
PK	PERIODO	PK	PERIODO	PK	PERIODO
	ESTRATO 1 ESTRATO 2 ESTRATO 3 ESTRATO 4 ESTRATO 5 ESTRATO 6 HOGARES NO CLASIFICADOS		Sin información Agronomía veterinaria y afines Bellas artes Ciencias de la educación Ciencias de la salud Ciencias sociales y humanas Economía, administración, contaduría y afines Ingeniería, arquitectura, urbanismo y afines Matemáticas y ciencias naturales		1 2 3 4 5 6 7 8 9 10 11 12 13 SIN INFORMACION

INGRESO DE LA FAMILIAR		NIVEL EDUCATIVO MADRE		INGRESO DE LA FAMILIA - NUEVO	
PK	PERIODO	PK	PERIODO	PK	PERIODO
	< 1 SALARIO MINIMO Entre 1 y menos de 2 salarios mínimos Entre 2 y menos de 3 salarios mínimos Entre 3 y menos de 5 salarios mínimos Entre 5 y menos de 7 salarios mínimos Entre 7 y menos de 9 salarios mínimos >=15 SALARIOS SIN INFORMACION Entre 9 y menos de 11 salarios mínimos Entre 11 y menos de 13 salarios mínimos Entre 13 y menos de 15 salarios mínimos		SIN INFORMACION PRIMARIA SECUNDARIA TECNICO O TECNOLOGICO SUPERIOR Y POSGRADO Entre 7 y menos de 9 salarios mínimos >=15 SALARIOS SIN INFORMACION Entre 9 y menos de 11 salarios mínimos Entre 11 y menos de 13 salarios mínimos Entre 13 y menos de 15 salarios mínimos		< 1 SALARIO MINIMO Entre 1 y menos de 2 salarios mínimos Entre 2 y menos de 3 salarios mínimos Entre 3 y menos de 5 salarios mínimos Entre 5 y menos de 7 salarios mínimos Entre 7 y menos de 10 salarios mínimos >=10 SALARIOS SIN INFORMACION

SISBEN		NUMERO DE HERMANOS	
PK	PERIODO	PK	PERIODO
	NIVEL 1 NIVEL 2 NIVEL 3 CLASIFICA EN OTRO NIVEL NO CLASIFICADO POR SISBEN		1 2 3 4 > 4 SIN INFORMACION NINGUNO

Ilustración 9. Datos SPADIES  
Fuente: Elaboración propia

ICFES – Presenta la información en tres informes, la clasificación de planteles, resultados agregados y resultados por secretaria. La clasificación de planteles permite conocer la información de los establecimientos de educación media (colegios), brindando datos de la ubicación, municipio, departamento, sector (Oficial, Privado), Calificación (A+, A, B, C y D) y los resultados en promedio obtenidos por cada establecimiento en las áreas del conocimiento evaluadas.

El informe de Resultados Secretarías, permite conocer la información de las secretarías de Colombia, como es la secretaría de educación de Bogotá, la secretaria de educación de Cundinamarca, entre otras, cual fue el comportamiento

de la población en las áreas del conocimiento, comparando con los resultados obtenidos a nivel nacional.

Por último están los resultados agregados que son las estadísticas por área del conocimiento en el examen saber11 por las regiones del país.



Ilustración 10. Resumen ICFES  
Fuente: Elaboración propia

Resultados Secretarías		LECTURA CRITICA		MATEMÁTICAS	
PK	Código DANE	PK	Código DANE	PK	Código DANE
	Fecha de presentación del examen Fecha de corte del SIMAT Fecha de actualización de datos Entidad territorial Tipo de entidad Número de municipios adscritos Número de establecimientos Evaluados (últimos 3 años) País/ET/Municipios Establecimientos educativos Matriculados Registrados Presentes Publicados Promedio Puntaje Global Comparación PG Desviación Comparación DPG		País/ET/Municipios Establecimientos educativos Promedio Comparación PLC Desviación Comparación DLC %1 %2 %3 %4 Aprendizaje 1 Aprendizaje 2 Aprendizaje 3		País/ET/Municipios Establecimientos educativos Promedio Comparación PMAT Desviación Comparación DMAT % 1 %2 %3 %4 Aprendizaje 1 Aprendizaje 2 Aprendizaje 3

SOCIALES Y CIUDADANAS		INGLES		CIENCIAS NATURALES	
PK	Código DANE	PK	Código DANE	PK	Código DANE
	País/ET/Municipios Establecimientos educativos Promedio Comparación PSYC Desviación Comparación DSYC % 1 %2 %3 %4 Aprendizaje 1 Aprendizaje 2 Aprendizaje 3 Aprendizaje 4 Aprendizaje 5 Aprendizaje 6		País/ET/Municipios Establecimientos educativos Promedio Comparación PING Desviación Comparación DING A- A1 A2 B1 B+		País/ET/Municipios Establecimientos educativos Promedio Comparación PCN Desviación Comparación DCN % 1 %2 %3 %4 Aprendizaje 1 Aprendizaje 2 Aprendizaje 3 Aprendizaje 4 Aprendizaje 5 Aprendizaje 6 Aprendizaje 7 Aprendizaje 8 Aprendizaje 9 Aprendizaje 10 Aprendizaje 11 Aprendizaje 12 Aprendizaje 13 Aprendizaje 14 Aprendizaje 15 Aprendizaje 16 Aprendizaje 17 Aprendizaje 18 Aprendizaje 19 Aprendizaje 20 Aprendizaje 21 Aprendizaje 22 Aprendizaje 23 Aprendizaje 24

Ilustración 11. Datos ICFES  
Fuente: Elaboración propia

Los datos agregados de la Clasificación de Planteles:

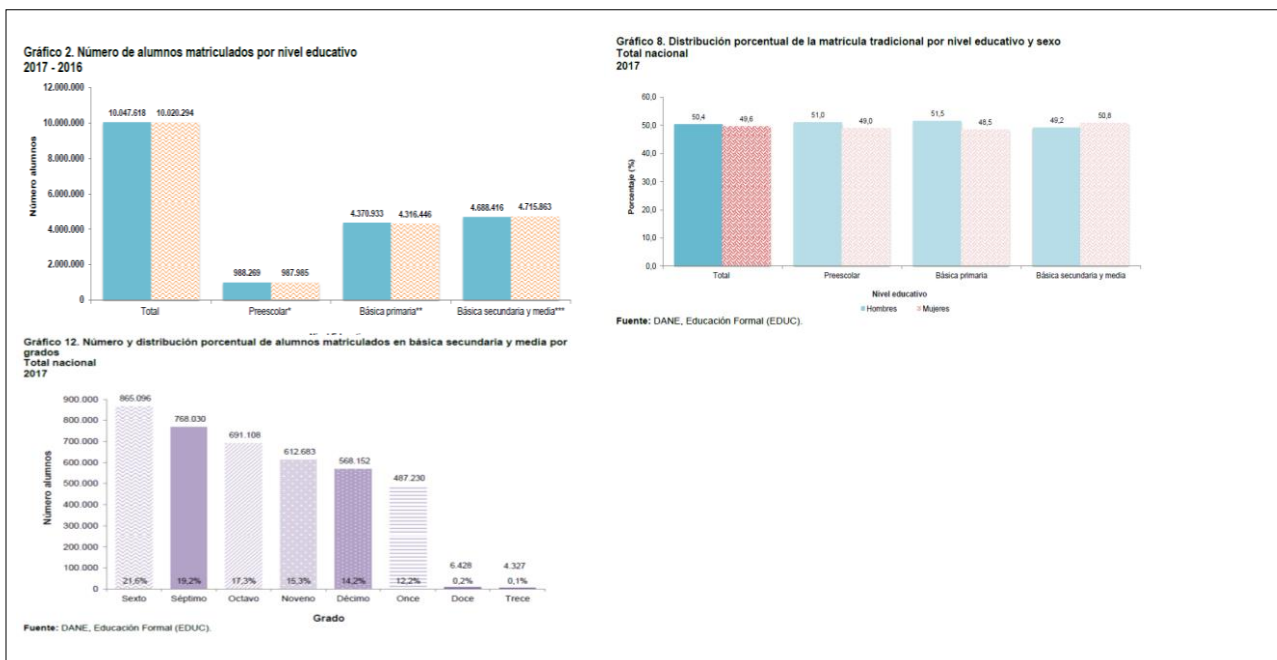


POR PLANTEL	
<b>PK</b>	<b>Código Dane</b>
	Nombre del Establecimiento
	Tipo
	Municipio
	Sector
	Clasificación
	Matriculados (últimos 3 años)
	Evaluados (últimos 3 años)
	Índice de Matemática
	Índice de Ciencias Naturales
	Índice de Sociales y Ciudadanas
	Índice de Lectura Crítica
	Índice de Inglés
	Índice Total

Ilustración 12. Datos Agregados  
Fuente: Elaboración propia

La información del DANE permite evaluar comparativos de la información del censo y la información del MEN como se puede ver en el informe:

Información DANE:



Para esta propuesta de investigación a manera de ejemplo se definen un perfil académico, que permitan conocer el desempeño según el examen Saber11 de los

estudiantes de colegio; esta definición se debe hacer en compañía del tomador de decisión para tener mayor asertividad en la definición de la información que muestra la herramienta, en esta ocasión se plantea como investigador con algo de experiencia en el ámbito educativo.

Para este ejercicio, se definen los siguientes datos y sus fuentes que permiten definir el perfil del futuro estudiante de una Institución de Educación Superior:

La información que se requiere para este ejemplo es, el colegio de procedencia (plantel), puntaje saber11 en cada una de las 5 áreas (Matemáticas, Lectura Crítica, Ciencias Naturales, Sociales y Ciudadanas e Inglés), Departamento, Ciudad/Municipio, naturaleza del plantel (oficial/no oficial), resultados nacionales en cada una de las áreas de conocimiento. Estos datos los entregan los informes del ICFES clasificación de planteles y el informe de resultados agregados.

Los datos que permitan conocer el porcentaje de deserción de la educación superior, Áreas de conocimiento en la educación superior, Institución de Educación Superior, ciudad, departamento, inscritos, admitidos, matriculados y sexo. Esta información la provee el informe de SNIES.

Incluir el comportamiento de la población según encuesta del DANE a nivel educación, para obtener una visión más global de la realidad del país y tener mayores herramientas para el planteamiento de las estrategias.

Luego de la evaluación de la información más coherente y con facilidad de cruzar para el ejemplo se determinan las siguientes características:

- Los resultados del examen Saber 11, el promedio que obtuvo cada establecimiento de educación superior en a partir del 2014 2 hasta el 2017 2, recordar que son 5 áreas de conocimiento.
- La ubicación del establecimiento de educación departamento y municipio, a futuro se puede definir con la dirección exacta para visualizar mejor los datos.
- La clasificación del establecimiento de educación siendo A+ la clasificación más alta, seguida de A, B, C y D, esta clasificación la define el ICFES.
- Luego se define el grado 11 y grado 26 (ciclo para adultos), el calendario académico del establecimiento de educación, el sector oficial y no oficial, la población que presenta el examen saber en los últimos tres años de cada establecimiento educativo.

Con esta información se plantearán los datos que se visualizaran en la maqueta de ejemplo, estos datos se modelaran en la herramienta de Power BI, por el convenio

que se tienen como institución académica y el conocimiento en la misma se seleccionó para este ejemplo.

En la ilustración 13. se observa el cruce de información de este ejercicio.

[00]



Ilustración 13. Cruce información Saber11  
Fuente: Creación propia

### 4.3. OBJETIVO 3

*Detallar la metodología que permita construir el perfil de los futuros estudiantes de las Instituciones de Educación Superior*

Se plantea un ciclo para poder trabajar la información, siempre alineado con los objetivos del proyecto.

Lo primero realizar una revisión de las fuentes de información, el siguiente paso a seguir es determinar los datos sociodemográficos que se van a utilizar, donde se plantea generar un listado de la información que le interesa conocer a la Institución de Educación Superior, luego determinar cuál de las fuentes de información me permite conocer la información de interés.

No se propone un modelo predictivo en esta fase, porque es necesario determinar un modelo económico o evaluar con un experto en el tema, cómo se realizaría esta propuesta y por su extensión puede ser el inicio de otro proyecto que fortalezca esta propuesta.

Como último paso se presenta la información por medio de la inteligencia de negocios.

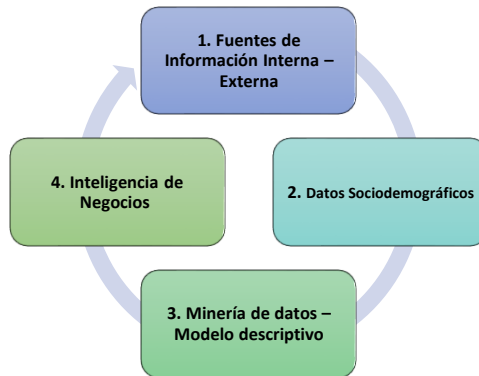


Ilustración 14. Ciclo de los Datos  
Fuente: Creación propia

En la revisión de la fundamentación teórica se habla de los DSS – Sistema de Soporte a la Decisión donde Sinnexus, expresa que estos sistemas permiten resolver gran parte de las limitaciones de los programas de gestión, informes dinámicos, flexibles e interactivos, no requiere conocimientos técnicos, tiempo de respuesta, integración de sistemas, perfilamiento de la información y la información histórica. (Sinnexus, 2017)

*“...Las interfaces de lenguaje natural son una nueva incorporación a los recursos de la BI. Pueden hacer que los datos, los gráficos y los dashboards sean aún más accesibles. Permiten a las personas interactuar con ellos utilizando texto y lenguaje natural. Según Gartner, esta es la “próxima fase en la evolución de la creación de informes estándar hacia la narración de historias”.*” (Tableau, 2018)

Si el futuro de la presentación de informes es la narración de historias, es necesario presentar la información lo más natural que se pueda para el tomador de decisiones y que el lenguaje sea de fácil acceso.

Se propone una maqueta del proyecto como la visual que tiene el tomador de decisiones de la información que se le brinda para contextualizarlo y para tomar la decisión.

La página inicial entonces se propone que sea un menú donde el usuario pueda escoger la información que desea conocer de las diferentes fuentes de información, como se ve en la *Ilustración 16.*, cabe anotar que estos diseños se deben trabajar en conjunto con la Institución que lo decida implementar y pueden cambiar totalmente.

En este caso se requiere información dinámica donde el usuario pueda interactuar con los datos y seleccionar el perfil de su interés y además que puedan ser

actualizados semestre a semestre con la información del ICFES y los otros Sistemas de Información Pública, se definen entonces marcos que contienen la información de contexto como son:

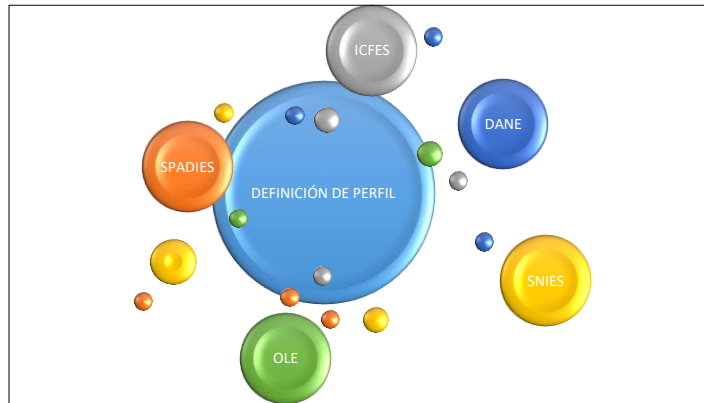


Ilustración 15. Maqueta Inicia  
Fuente: Elaboración propia

En esta ilustración, se propone tener 6 botones, donde se puede escoger cual es la información de interés del usuario o ir directamente a la definición del perfil.

Los círculos pequeños permiten tener un contexto de la población de estudiantes. Debe ser interactivo, entonces si el usuario da clic en el botón SNIES, le permite tener la siguiente visual (también es una propuesta que puede ser modificada).

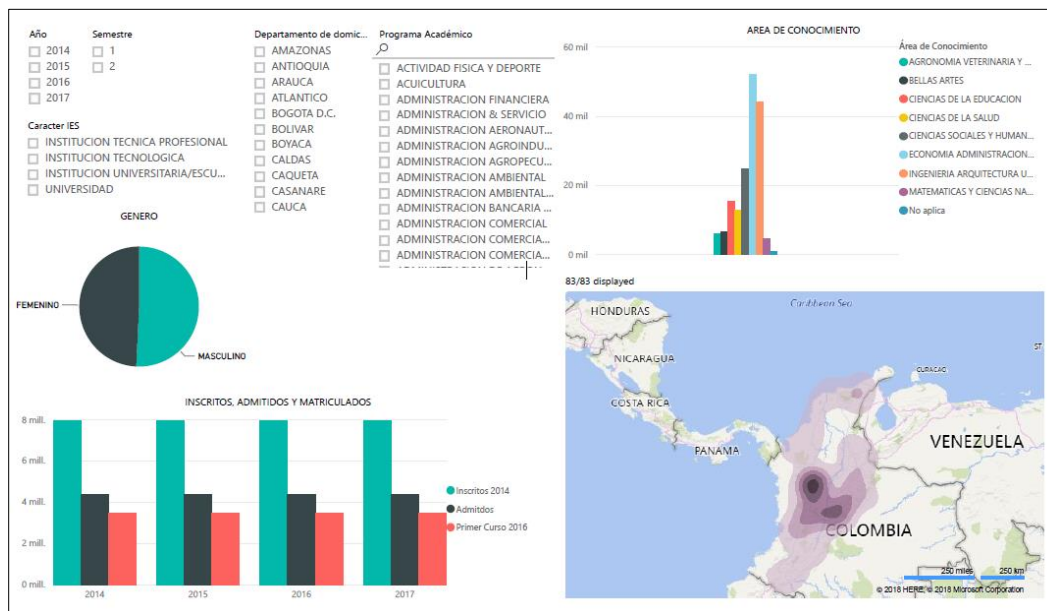


Ilustración 16. Maqueta SNIES  
Fuente: Elaboración propia

Esta información debe ser interactiva como anteriormente, se ha hecho énfasis.

Conocería la ubicación geográfica, áreas del conocimiento, como está creciendo o decreciendo la inscripción, admisión y matrícula, bien sea por toda la población o por un programa o programas específicos, cuantas personas son de género femenino, masculino.

La siguiente vista sería de SPADIES, donde conocer la información del estrato, el ingreso familiar, integrantes de la familia y la clasificación ICFES, todo esto sigue siendo parte del contexto.

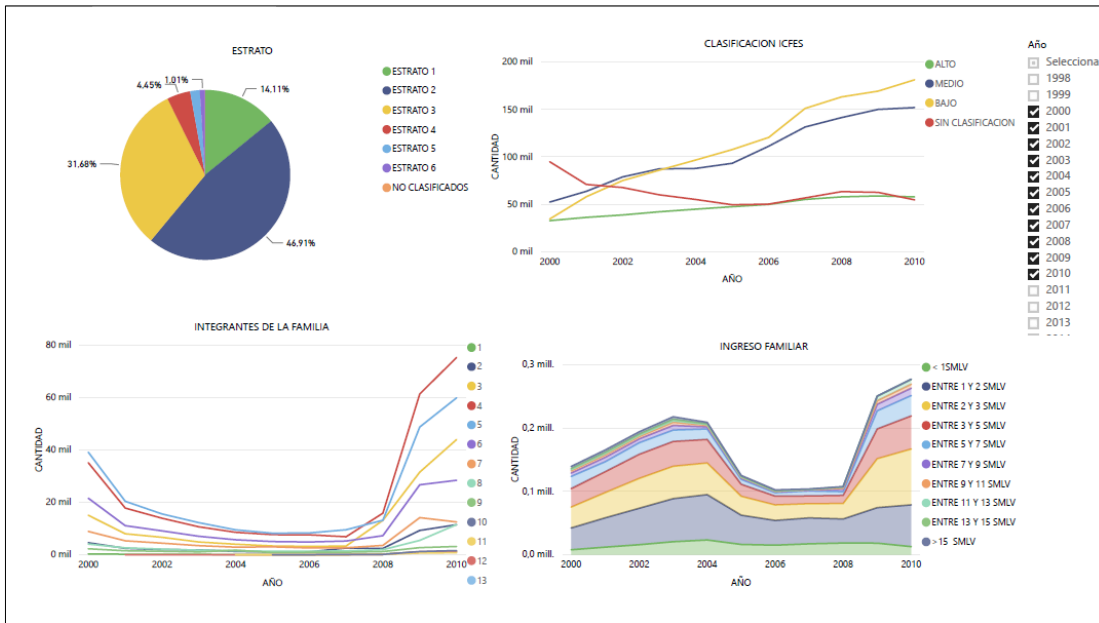
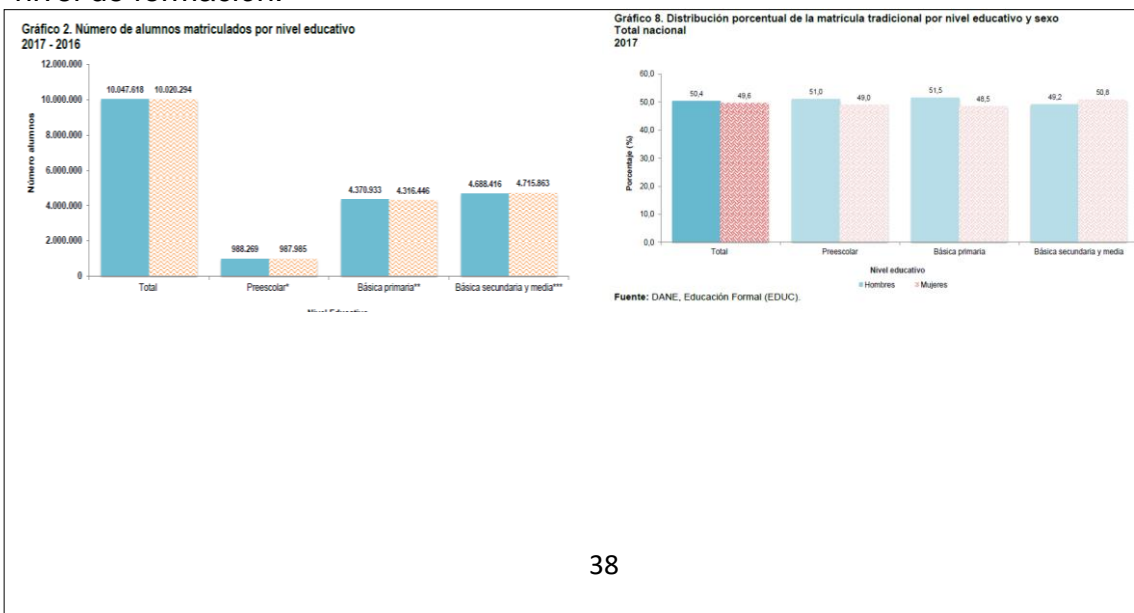


Ilustración 17. Maqueta SPADIES  
Fuente: Elaboración propia

La tercera vista es la información del DANE, contextualizando a nivel Colombia por nivel de formación.



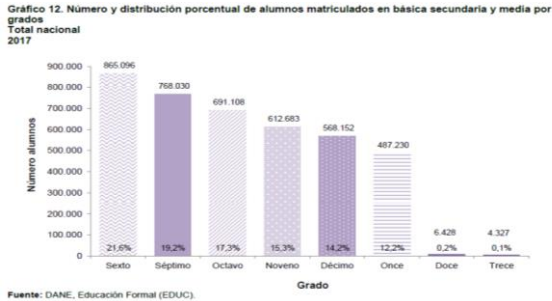


Ilustración 18. Maqueta DANE  
Fuente: Elaboración propia

La cuarta vista es del OLE, donde se identifica el comportamiento de que programas tienen más graduados, de que Regiones del país, el nivel académico (Doctorado, Maestría, Especialización, Pregrado, ...)

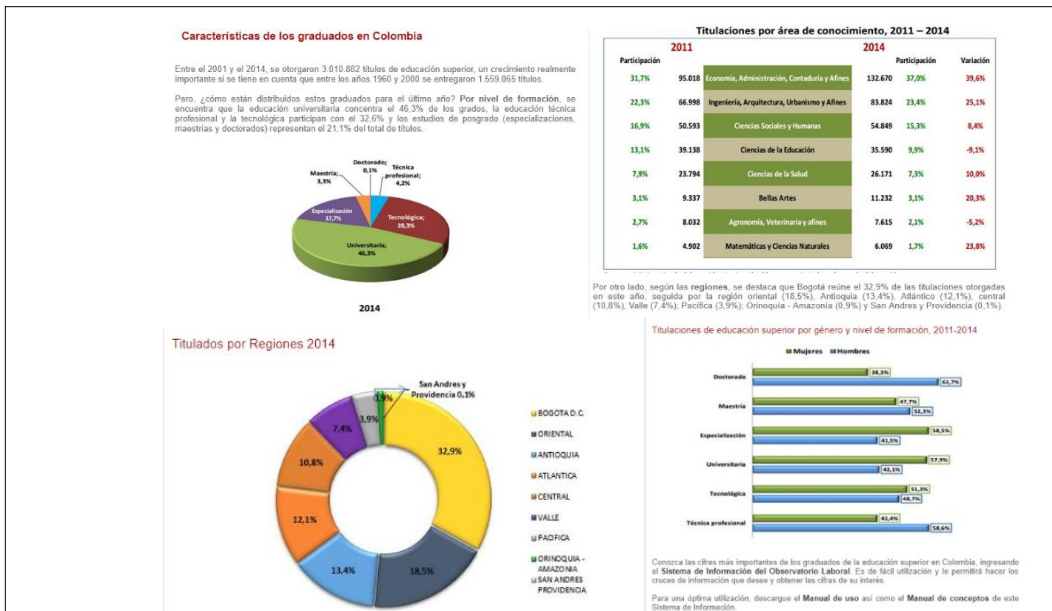


Ilustración 19. Maqueta OLE  
Fuente: Elaboración propia

La quinta y sexta vista es la información del ICFES, periodo, jornada, la naturaleza del establecimiento, departamento, si quiere conocer un establecimiento (colegio) específico, cuál es el desempeño académico.



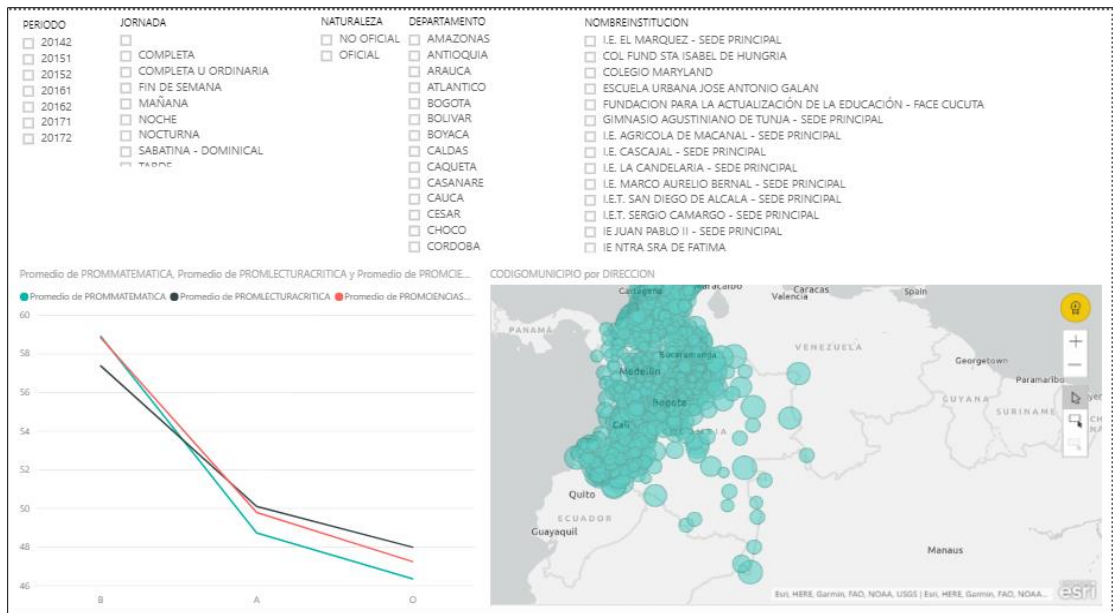


Ilustración 20. Maqueta ICFES1  
Fuente: Elaboración propia

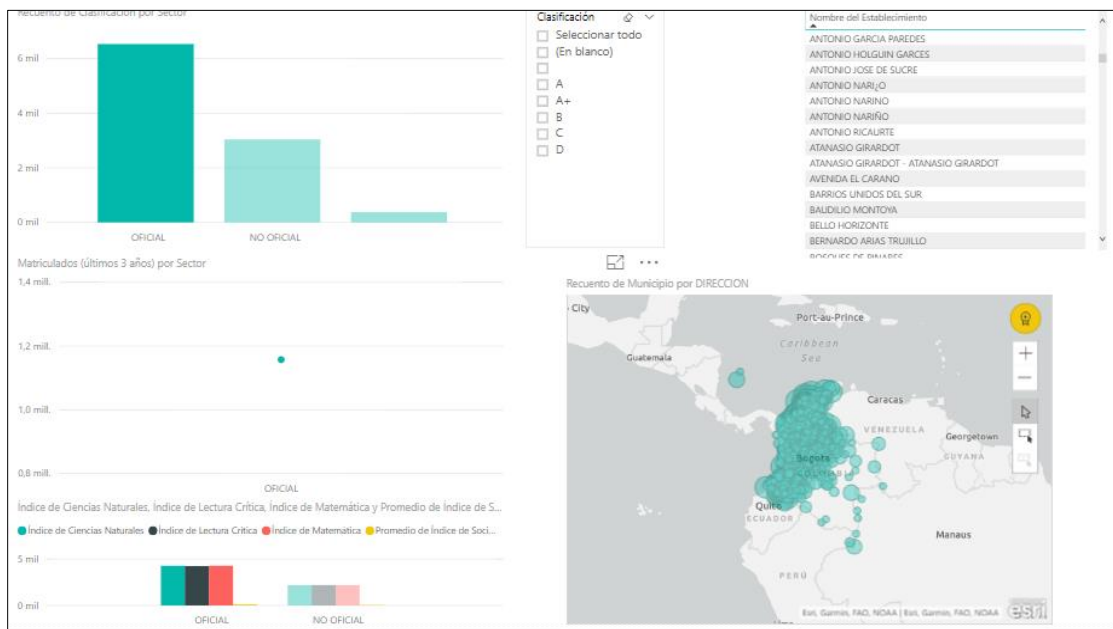


Ilustración 21. Maqueta ICFES2  
Fuente: Elaboración propia

Esta séptima vista que es una opción que se plantea para definir el perfil del estudiante basándose únicamente en la información del ICFES, definiendo el calendario del colegio, sector, departamento, Municipio, Grado, Clasificación, Puntaje en el examen saber 11 en un área específica; le mostrará entonces la cantidad de estudiantes a la que puede llegar y el listado de los colegios que tienen el perfil definido para la ilustración de la estrategia de captación.



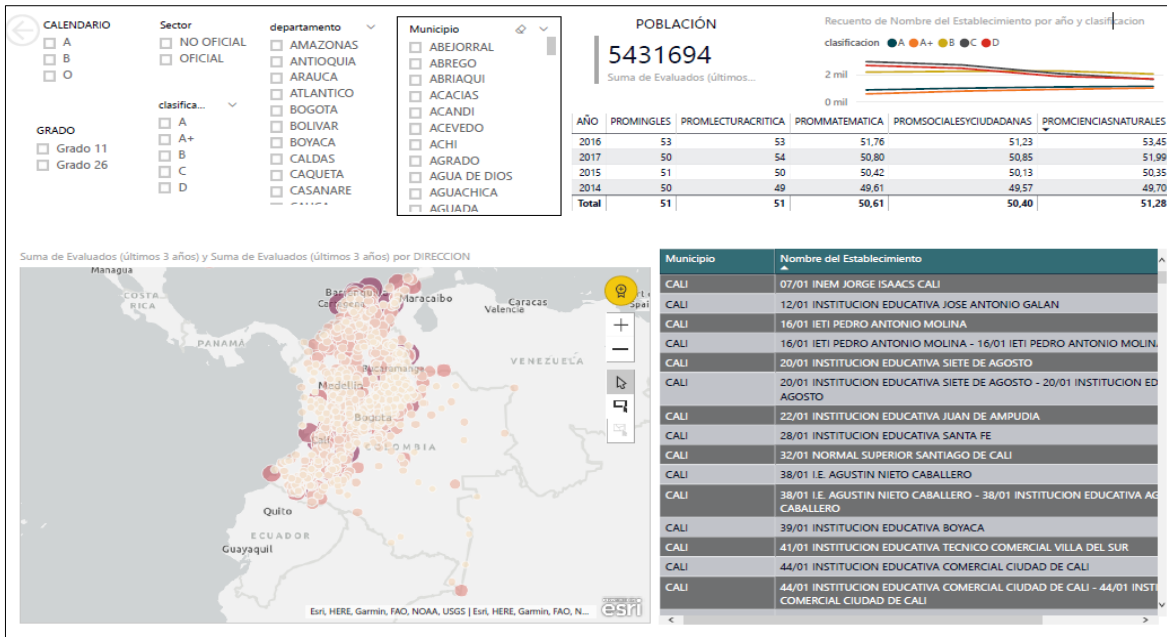


Ilustración 22. Maqueta PERFIL

Fuente: Elaboración propia

Es necesario que la información que se le brindó para contextualizar al tomador de decisiones permita integrar todas o la gran mayoría de fuentes de información como se muestra en la pantalla de definición del perfil a continuación:

SNIES		SPADIES		GENERAL																							
PROGRAMA ACADEMICO	<input type="text"/>	ESTRATO	<input type="text"/>	AÑO	<input type="text"/>																						
AREA DE CONOCIMIENTO	<input type="text"/>	CLASIFICACION ICFES	<input type="text"/>	SEMESTRE	1 <input type="checkbox"/> 2 <input type="checkbox"/>																						
<b>OLE</b>		INTEGRANTES FLIA	<input type="text"/>	GÉNERO	F <input type="checkbox"/> M <input type="checkbox"/>																						
TECNICO	<input type="checkbox"/>	INGRESO FAMILIAR	<input type="text"/>	DEPARTAMENTO	<input type="text"/>																						
TECNICO PROFESIONAL	<input type="checkbox"/>	<p>POBLACIÓN: XXXXX</p> <table border="1"> <thead> <tr> <th>COLEGIO</th> <th>DEPARTAMENTO</th> <th>CIUDAD</th> <th>CLASIFICACION</th> <th>PROM MAT</th> <th>PROM CI</th> <th>PROM LEC</th> <th>PROM SC</th> <th>PROM ING</th> <th>AÑO</th> <th>SEMESTRE</th> </tr> </thead> <tbody> <tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr> </tbody> </table> <p>DESCARGAR LISTADO</p>				COLEGIO	DEPARTAMENTO	CIUDAD	CLASIFICACION	PROM MAT	PROM CI	PROM LEC	PROM SC	PROM ING	AÑO	SEMESTRE											
COLEGIO	DEPARTAMENTO					CIUDAD	CLASIFICACION	PROM MAT	PROM CI	PROM LEC	PROM SC	PROM ING	AÑO	SEMESTRE													
UNIVERSITARIA	<input type="checkbox"/>																										
DECIMO	<input type="checkbox"/>																										
ESPECIALIZACIÓN	<input type="checkbox"/>																										
MAESTRIA	<input type="checkbox"/>																										
DOCTORADO	<input type="checkbox"/>																										
AREA DE CONOCIMIENTO	<input type="text"/>																										
REGION	<input type="text"/>																										
<b>ICFES</b>																											
JORNADA	<input type="text"/>																										
NATURALEZA	<input type="text"/>																										
PROM MAT	<input type="text"/>																										
PROM LECTURA	<input type="text"/>																										
PROM CIENCIAS NATUR.	<input type="text"/>																										
PROM SOC Y CIUDADA.	<input type="text"/>																										
PROM INGLES	<input type="text"/>																										
CLASIFICACIÓN PLANTEL	<input type="text"/>																										
CALENDARIO	<input type="text"/>																										
<b>DANE</b>																											
PRESCOLAR	<input type="checkbox"/>																										
BASICA PRIMARIA	<input type="checkbox"/>																										
BASICA SECUNDARIA Y MEDIA	<input type="checkbox"/>																										
SEXTO	<input type="checkbox"/>																										
SEPTIMO	<input type="checkbox"/>																										
OCTAVO	<input type="checkbox"/>																										
NOVENO	<input type="checkbox"/>																										
DECIMO	<input type="checkbox"/>																										
ONCE	<input type="checkbox"/>																										
DOCE	<input type="checkbox"/>																										
TRECE	<input type="checkbox"/>																										

Ilustración 23. Definición perfil

Fuente: Elaboración propia

Se plantea una vista para poder seleccionar las características de sus futuros estudiantes, en las áreas de conocimiento cabe recordar que el puntaje esta entre

o y 100, luego de ingresar los datos en cada característica de interés, este le presenta el listado de los colegios con las características definidas en la herramienta.

Para esto es necesario poder integrar la información histórica de la Institución de Educación Superior bien sea para contextualizar o para integrar en la definición del perfil.

La información histórica puede determinar el tamaño de la población objetivo, porque se conoce el porcentaje de absorción único de esta Institución de inscritos a admitidos y de admitidos a matriculados.

Entonces esta información puede ayudar a definir el tamaño de la población y la utilización de los recursos económicos y de personal que se tienen para llevar a cabo la estrategia de captación definida por la institución.

A manera de ejemplo se propone:

Se define el perfil del estudiante de colegio de **calendario A**, con un estrato socioeconómico **3**, nivel de ingresos familiar **mayor a 3 salarios mínimos**, con un desempeño en el examen saber 11 en el área de **Matemáticas de 50 a 70 puntos**, que provenga de un colegio con **clasificación A+**, no importa si es público o privado, del departamento de **Cundinamarca, del año 2018, periodo 2**.

Para este ejercicio se define un **porcentaje de absorción**, la selección de la población depende el objetivo planteado por la institución.

Se encuentra que se puede llegar a una población de aproximadamente 1.000.000 personas con este perfil se muestra la información de interés para determinar mejor la estrategia de promoción.

La cantidad de colegio a los que se puede llegar son 100.000.

Este ejemplo se puede observar en la siguiente ilustración:

SNIES	
PROGRAMA ACADEMICO	<input type="text"/>
AREA DE CONOCIMIENTO	<input type="text"/>


OLE	
TECNICO	<input type="checkbox"/>
TECNICO PROFESIONAL	<input type="checkbox"/>
UNIVERSITARIA	<input type="checkbox"/>
DECIMO	<input type="checkbox"/>
ESPECIALIZACIÓN	<input type="checkbox"/>
MAESTRIA	<input type="checkbox"/>
DOCTORADO	<input type="checkbox"/>
AREA DE CONOCIMIENTO	<input type="text"/>
REGION	<input type="text"/>

ICFES	
JORNADA	<input type="text"/>
NATURALEZA	<input type="text"/>
PROM MAT	50 <input type="checkbox"/> 70 <input type="checkbox"/>
PROM LECTURA	<input type="text"/>
PROM CIENCIAS NATUR.	<input type="text"/>
PROM SOC Y CIUDADA.	<input type="text"/>
PROM INGLES	<input type="text"/>
CLASIFICACIÓN PLANTEL	A+ <input type="checkbox"/>
CALENDARIO	A <input type="checkbox"/>

SPADIES	
ESTRATO	3 <input type="checkbox"/>
CLASIFICACIÓN ICFES	<input type="text"/>
INTEGRANTES FLIA	<input type="text"/>
INGRESO FAMILIAR	>3smly <input type="checkbox"/>

<b>POBLACIÓN:</b> 1.000.000	
--------------------------------	--

GENERAL	
AÑO	2018 <input type="text"/>
SEMESTRE	1 <input type="checkbox"/> 2 <input checked="" type="checkbox"/>
GÉNERO	F <input type="checkbox"/> M <input type="checkbox"/>
DEPARTAMENTO	CUND <input type="text"/>

DANE	
PRESCOLAR	<input type="text"/>
BASICA PRIMARIA	<input type="text"/>
BASICA SECUNDARIA Y MEDIA	<input type="text"/>
SEXTO	<input type="checkbox"/>
SEPTIMO	<input type="checkbox"/>
OCTAVO	<input type="checkbox"/>
NOVENO	<input type="checkbox"/>
DECIMO	<input type="checkbox"/>
ONCE	<input checked="" type="checkbox"/>
DOCE	<input type="checkbox"/>
TRECE	<input type="checkbox"/>

COLEGIO	DEPARTAMENTO	CIUDAD	CLASIFICACION	PROM MAT	PROM CN	PROM LEC	PROM SC	PROM ING	AÑO	SEMESTRE
NUEVO REINO	CUND	CHIA	A+	60	50	48	70	80	2018	2

[DESCARGAR LISTADO](#)

Ilustración 24. Ejemplo Perfil  
Fuente: Elaboración propia

El uso de la información queda entonces a disposición de la Institución de Educación Superior para definir el perfil de los estudiantes a lo que desea llegar y poder generar estrategias con información de calidad, la Institución de Educación Superior puede definir y utilizar la herramienta cuantas veces desee diferentes perfiles dado que no tiene limitantes, la muestra de la información se hace a nivel nacional para brindar un panorama completo y poder ir de lo macro al detalle.

#### 4.4. OBJETIVO 4

*Seleccionar las técnicas que permiten llevar a cabo la metodología propuesta.*

Esta maqueta anteriormente presentada es la propuesta, pero para poder lograr este resultado se requiere definir cuáles son las técnicas que permiten llevar a cabo la metodología propuesta una forma es la que propone Camacho Chacón y sus compañeros de investigación en el 2018:

También referencia a Hernández quien presenta el KDD en 5 fases:



Ilustración 25 Fases KDD  
Fuente: Construcción propia

Entonces se propone para tratar los datos dos formas, los modelos predictivos y los modelos descriptivos Camacho (Camacho Chacón, Zapata Gonzalez, Menéndez Domínguez, & Canto Herrera , 2018):

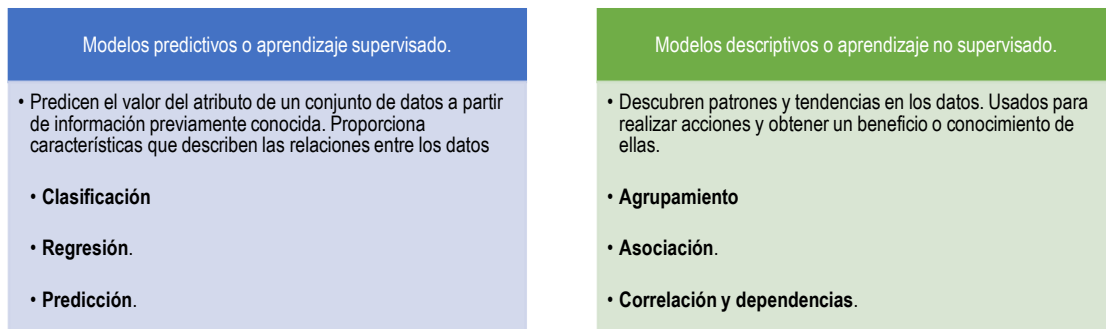


Ilustración 26. Modelos de aprendizaje  
Fuente: Elaboración propia

El siguiente paso a seguir es, consolidar la información por medio de la metodología KDD mencionada en el marco teórico donde se explican los pasos para hacer la limpieza y estandarización de los datos por medio de la minería de datos puede realizarse por medio de una herramienta para este fin o en un Excel dependiendo de la capacidad de las herramientas tecnológicas que se tienen a la mano.

Luego de la revisión de esta bibliografía se define tomar un modelo descriptivo de la información por medio de Agrupamiento y Asociación, el agrupamiento se hace de las características del perfil de futuro estudiante, realizando el cruce de datos de las diferentes fuentes de información hasta completar una base completa o un repositorio que permita analizar los datos y obtener información más completa. En

el momento de evaluar la calidad de la información se puede modificar el modelo de minería a utilizar eso dependerá de que defina el experto en minería.

Teniendo la información cruzada de las diversas bases y fuentes de información se define el uso de una herramienta de visualización para definir los diferentes tableros donde el tomador de decisión puede jugar con los datos y ver diferentes panoramas con la información, se debe tener varias visualizaciones una donde se realiza la introducción a la información con la que se enfrentará y otras tantas como se necesiten pensando siempre en acceso a la información de forma dinámica y gerencial.

Siguiendo la metodología propuesta se define el paso a paso del ejercicio de ejemplo:

a. Integración y recopilación

Continuando con el ejercicio planteado en los anteriores objetivos y a modo de ejemplo se determina que es necesario unificar la información definida en el primer y segundo objetivo, en este ejercicio se utiliza la información del ICFES<sup>2</sup>, la información de los resultados del examen Saber11 que es pública.

La información del SNIES, OLE, DANE, SPADIES y datos agregados como información que permite tener un contexto adecuado al tomador de decisiones. Revisando como insumo inicial la información del examen Saber11, Agregados por clasificación de planteles, por secretaria, por establecimiento educativo.

Por ejemplo, la clasificación de los planteles son varios archivos de acuerdo al periodo académico en este caso contamos 2014-2, 2015-1, 2015-2, 2016-1, 2016-2, 2017-1 y 2017-2, se reúnen en un solo lugar donde se pueden tener hasta 70.0000 registros, luego se debe cruzar los datos contra un archivo donde se encuentra la información del calendario académico de los planteles.

b. Selección, limpieza y transformación

Para este ejercicio en la limpieza de los datos se usó la herramienta de minería de datos llamada Open Refine (de uso libre), se usa esta herramienta por los conocimientos previos de los investigadores, cada proponente de la solución puede seleccionar la herramienta de su preferencia como lo puede ser Excel que también se utilizó en el proceso u otro programa.

c. Minería de datos

---

<sup>2</sup> [www.ICFESinteractivo.gov.co](http://www.ICFESinteractivo.gov.co)

Con este compilado se hace el cruce contra los resultados promedio de los planteles en las áreas de conocimiento donde el código ICFES es la llave y los demás archivos tienen en código DANE de los planteles entonces se debe hacer la transformación de los datos para poder continuar con la llave DANE, cada uno de los archivos de los resultados los va a encontrar con los resultados por periodo académico entonces se debe hacer la relación y organizar esta información.

#### d. Evaluación e interpretación

Con la base de datos definida se inicia con la agrupación de la información, es necesario contar con una herramienta de visualización que permita ver la agrupación de la información de forma integrada, para este caso a manera de ejemplo pero no como una herramienta definitiva se usa Power BI dado los conocimientos previos de la herramienta, donde se arman los diferentes tableros con la información para contextualizar al tomador de decisiones y el tablero que es el corazón del proyecto donde se pueden seleccionar las características de los futuros estudiantes de la institución de educación superior.

#### e. Difusión y uso

Luego de tener definida la información en los Dashboard se le da acceso al tomador de decisión para que conozca la información y pueda plantear estrategias de captación inicialmente.

En este punto es muy importante decir que la definición del perfil se puede usar en diferentes poblaciones objetivo, como lo es la asignación de una beca y la promoción de la misma, o la definición de la Institución de Educación Superior de llegar a una población de planteles con bajo rendimiento en el examen Saber 11 y apoyar la política de apoyo a la comunidad con un proyecto de mejoramiento de los contenidos programáticos o el apoyo académico de estas poblaciones, también como se ha mencionado desde el principio definir la población objetivo para realizar actividades que permitan volver estos estudiantes de colegio en estudiantes de la Institución de Educación Superior y enfocar los esfuerzos de mercadeo y promoción de las mismas tomando decisiones basados en información.

### **4.4.1. VISUALIZACIÓN**

Actualmente, existen varias herramientas de visualización, la selección de la herramienta depende del presupuesto que se destine al proyecto y las necesidades identificadas de la Institución, evaluando el conocimiento de la herramienta, la capacitación en la misma, los costos de mantenimiento, cantidad de usuarios que

brinda, entre otros que en su momento se debe definir antes de la compra o decisión por una herramienta.

Algunas herramientas a evaluar pueden ser: Power BI, SAS, Oracle ODI, sap, Tableau, Microsoft SSIS, Sisense, Thoughtspot, Qlik, Informatica, Information Builders, Ibm Datastage, Microstrategy, Talend, entre otros como puede ser la construcción propia de cada empresa.



## 5. CONCLUSIONES

Se entrega una propuesta útil para enfocar mejor las estrategias de captación y los recursos de una Institución de Educación Superior, actualmente conocen estas fuentes de información, pero no la saben manejar, entonces se le da valor a la información.

Se da a conocer la ubicación geográfica y los colegios donde se encuentran los aspirantes que cumplen con las características del perfil definidas por el tomador de decisiones.

Con mayor información se define mejor el perfil de las personas objetivo de las Instituciones de Educación Superior, si se trabaja con información de las redes sociales se puede determinar características del perfil como, gustos musicales, personas de interés, deportes, entre otras y se puede personalizar y llegar casi de forma individual.

## 6. TRABAJOS FUTUROS

Plantear modelos predictivos enfocados a determinar cuáles características tienen mayor probabilidad de grado.

Trabajar con información de las redes sociales.

Unificar la información histórica única de la Institución de Educación Superior que acoja el proyecto e incluirla como parte de la definición del perfil.

Aportar en la evaluación y mejora de la calidad de la educación como un proyecto a futuro.

Conectar con la realidad del país, con el ICETEX, las necesidades de acceder a créditos educativos por parte de la población objetivo, el plan de desarrollo como país, cuáles son las tendencias del mercado, las necesidades de tecnología.

Plantear una revisión constante de la metodología y las fuentes de información que están en constante cambio.

## 7. PRESUPUESTO DE INVERSIÓN

En la siguiente tabla se expresa el presupuesto de inversión para la investigación:

<b>Rubros</b>	<b>Valor</b>
Personal	\$15,000,000.00
Materiales e insumos	
Viajes	
Inscripciones a eventos	\$500,000.00
Publicaciones y difusión de resultados	\$800,000.00
Actividades de campo	
Equipos y software	\$3,700,000.00
<b>Total</b>	<b>\$20,000,000.00</b>



## **8. EQUIPO INVESTIGADOR**

Estudiante de Maestría: Ángela Daniela Prieto Parra, Ingeniera Industrial de la Escuela Colombiana de Ingeniería, Profesional de la Oficina de Admisiones adscrita a la Secretaría General de la Escuela Colombiana de Ingeniería.

Director: Ph. D. Oswaldo Castillo

## REFERENCIAS

- Aja Quiroga, L. (2002). Gestión de información, gestión del conocimiento y gestión de la calidad en las organizaciones. *ACIMED*, 10(5), 7-8. Recuperado en 06 de octubre de 2017, de [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1024-94352002000500004&lng=es&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352002000500004&lng=es&tlng=es).
- Arroyo, M., Guillermina, B. P., Baena Paz, G. L., Balbi, E. R., Bitar, S., Cabrera Jiménez, J. A., . . . Didriksson Takayanagi, A. (2015). EL ANTIÁZAR: LA PLANEACIÓN. En M. Arroyo, G. Baena Paz, G. L. Baena Paz, E. R. Balbi, S. Bitar, J. A. Cabrera Jiménez, . . . F. Ortiz Manbuscay, *Planeación prospectiva estratégica*. (pág. 644). Mexico: Metadata.
- Barcos, S. J. (2008). Reflexiones acerca de los Sistemas de Información Universitarios ante los desafíos y cambios generados por los procesos de evaluación y acreditación. *Avaliação (Campinas; Sorocaba)*, 209 - 244.
- Brunet, R. (2016). ¿Qué es la minería de datos? *Muy Interesante*.
- Cacho, Y. C. (26 de 07 de 2016). Twitter, área emergente para la investigación. Ciudad de Mexico, Ciudad de Mexico, Mexico.
- Camacho Chacón, P., Zapata Gonzalez, A., Menéndez Domínguez, V., & Canto Herrera , P. (2018). Análisis del desempeño del profesorado universitario en el uso de MOODLE a través de técnicas de minería de datos: propuestas de necesidades formativas. *Revista de Educación a Distancia*, 1-41.
- Camargo-Vega, J. J., Camargo-Ortega, J. F., & Joyanes-Aguilar, L. (2015). Conociendo Big Data. *Revista Facultad de Ingeniería*, 63 - 77.
- Camino, J. R., & De Garcillán, M. L. (2012). *Dirección de Marketing. Fundamentos y aplicaciones*. Madrid: ESIC.
- CONEAU, C. N. (1997/2006). *Informes sobre evaluaciones externas*. Buenos Aires.
- Davenport, T. H., & Prusak, L. (15 de 09 de 1998). *Working knowledge: How organizations manage what they know*. Boston: EUA: Harvard Business School Press.

- Duro, V., & Perez, C. (2016). BUSINESS INTELLIGENCE SYSTEM AND DECISION SUPPORT OF ECONOMIC MANAGEMENT AT THE UNIVERSITY OF HAVANA. *3C TIC*, 38 - 54.
- EL TIEMPO. (4 de OCTUBRE de 2017). Nuevas tecnologías, apuesta de empresas para transformar sus negocios. *EL TIEMPO*.
- Fayyad., P.-s. &. (1996). *From Data Mining to Knowledge Discovery in Databases*. Rhode Island: AAI.
- Feet, N., & Rodrigez - Ponce, E. (2011). Information management and quality of universities: Empirical study on Universities in Chile. *Interciencia*, 570 - 577.
- Fernandez - Molina, J. (1995). La responsabilidad de los profesionales de la documentacion en la prestacion de servicios de información. *Esp Doc Cient*, 320 - 32.
- Gobierno de España. (21 de 10 de 2017). *Gobierno de España*. Obtenido de Gobierno de España: [http://administracionelectronica.gob.es/pae\\_Home/pae\\_Documentacion/pae\\_Metodolog/pae\\_Metrica\\_v3.html#.WewC12iCzIV](http://administracionelectronica.gob.es/pae_Home/pae_Documentacion/pae_Metodolog/pae_Metrica_v3.html#.WewC12iCzIV)
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques*. USA: Morgan Kaufmann
- haycanal.com. (3 de 10 de 2017). *hay canal* . Obtenido de haycanal.com: <http://haycanal.com/noticias/6702/El-Big-Data-hace-que-la-inteligencia-artificial-ya-no-sea-ciencia-ficcion>
- Hernández, J. R., & M.J., F. C. (2004). *Introducción a la Minería de Datos*. Pearson Prentice Hall.
- Howson, C. (2009). *Business Intelligence - Estrategias para una implementación exitosa*. Mc Graw Hill.
- ICFES. (26 de 03 de 2018). *Instituto Colombiano para la Evaluación de la Educación*. Obtenido de Instituto Colombiano para la Evaluación de la Educación: <http://www.icfes.gov.co/quienes-somos>:
- Ikram, D., Mesut, C., & Abdullah, N. (2011). The Effects Of Information Technology Supported. *Procedia Social and Behavioral Sciences*, 1134 - 1142.
- Jones, V. (4 de 2017 de 2017). *YouTube*. Obtenido de YouTube: <https://www.youtube.com/watch?v=Yrp3Fw3Nd-M>
- Lau, W., & Yen, A. H. (2014). Developing and validating of a perceived ICT literacy scale for junior secondary school students: Pedagogical and educational contributions. *Computers & Education*, 1 - 9.

- MEN. (26 de 3 de 2018). *Ministerio de Educacion Nacional*. Obtenido de <https://www.mineducacion.gov.co/portal/Educacion-superior/Sistema-de-Educacion-Superior/235585:Sistema-de-aseguramiento-de-la-calidad-de-la-educacion-superior>
- MINTIC. (1 de 11 de 2017). *Ministerio de Tecnologias de Informacion y Comunicaciones*. Obtenido de Ministerio de Tecnologias de Informacion y Comunicaciones: <http://www.mintic.gov.co/arquitecturati/630/w3-propertyvalue-8078.html>
- Molina Félix, L. C. (2002). Data mining: torturando a los datos hasta que confiesen. *UOC*, 1-11.
- Reyes M., C. (2004). Una breve introducción a las tecnologías de la. *Intangible Capital*, 1-12.
- Rodriguez - Ponce, E. (2009). El rol de las universidades en la sociedad del conocimiento y en la era de la globalización: evidencia desde chile. *Interciencia*, 824-829.
- Sinnexus. (4 de Octubre de 2017). *Sinnexus*. Obtenido de Sinnexus: [http://www.sinnexus.com/business\\_intelligence/index.aspx](http://www.sinnexus.com/business_intelligence/index.aspx)
- Tableau. (25 de 06 de 2018). *Tableau*. Obtenido de Tableau: <https://itperforma.com/wp-content/uploads/2017/03/10-tendencias-de-inteligencia-de-negocios-para-2017.pdf>
- UNESCO. (1998). *La educación superior en el siglo XXI*. Paris: UNESCO.
- UNESCO. (2005). Hacia las sociedades del conocimiento. *Organización de las Naciones Unidas*, 1 - 240.
- Vicente Molina, M. A., & Echevarria Gonzalez de Durana, C. A. (8 de 09 de 2003). *Aproximación al perfil sociodemografico del consumidor ecologia a traves de la evidenciia empirica: propuesta para el desarrollo del mix de marketing*. Obtenido de BOLETIN ECONOMICO DE INFORMACION COMERCIAL ESPAÑOLA: [https://www.researchgate.net/publication/28064279\\_Aproximacion\\_al\\_perfil\\_sociodemo-grafico\\_del\\_consumidor\\_ecologico\\_a\\_traves\\_de\\_la\\_evidencia\\_empirica\\_propuestas\\_para\\_el\\_desarrollo\\_del\\_mix\\_de\\_marketing?enrichId=rgreq-57ffbe7f5fcce28dd631f8eb1a68e6ef-XXX&enr](https://www.researchgate.net/publication/28064279_Aproximacion_al_perfil_sociodemo-grafico_del_consumidor_ecologico_a_traves_de_la_evidencia_empirica_propuestas_para_el_desarrollo_del_mix_de_marketing?enrichId=rgreq-57ffbe7f5fcce28dd631f8eb1a68e6ef-XXX&enr)