
 <p>Universidad de los Andes Colombia</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Semestre: 2022-10</p>	
--	---	---

CASO A TRABAJAR: Infraestructura visible

El proyecto se realiza en la temática de Infraestructura visible¹, una **iniciativa de la Universidad de los Andes que permite el libre acceso a la información sobre la infraestructura colombiana**. El proyecto está enfocado en el análisis de indicadores socioeconómicos y dimensiones municipales de Colombia. Las dimensiones que se desean analizar son:

1. Demografía y Población
2. Vivienda y servicios públicos
3. Educación
4. Salud
5. Conflicto armado
6. Medición de desempleo

El proyecto busca aprovechar los datos abiertos para generar conocimiento útil que aporte a diferentes tipos de actores, desde el ciudadano normal hasta el experto en temáticas allí trabajadas, pasando por actores como periodistas, entre otros. De igual manera, quieren ampliar la gama de análisis que se pueden observar en la página web de la iniciativa.



Al final del proyecto, el cliente espera contar con tableros de control que le permitan analizar indicadores socioeconómicos y las dimensiones de demografía y población en conjunto con otras tres, que su grupo seleccione dentro de las 5 enumeradas previamente.

InfraestructuraVisible está muy interesada en conocer los estilos de análisis que pueden hacer y por supuesto, en ver propuestas que pueden ayudarles a definir un mapa de ruta para futuras consultorías. En este orden de ideas, junto con los expertos en inteligencia de negocio de la organización y algunos asesores, han planteado el siguiente alcance para esta consultoría.

Identificación e implementación de requerimientos de analítica basada en tableros de control, acompañados de *data marts* que incluyen el modelo dimensional, la carga de datos en una base de datos y el perfilamiento y preparación de los datos. Este desarrollo debe tener en cuenta las recomendaciones dadas por el cliente a los entregables en semestres anteriores para no repetir los mismos errores.

Este alcance se detalla a continuación:

¹ <https://www.infraestructuravisible.org/>

 <p>Universidad de los Andes Colombia</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Semestre: 2022-10</p>	
--	---	---

ENUNCIADO SEGUNDO PROYECTO

OBJETIVOS

Este trabajo tiene como objetivo principal

- Desarrollar un proyecto de inteligencia de negocios centrado en el uso de datos para monitorear métricas e indicadores de interés para una organización, incluida la sociedad.

A nivel de objetivos específicos



- Identificar requerimientos analíticos que puedan ser resueltos con análisis descriptivos a desplegar en tableros de control.
- Implementar requerimientos analíticos basados en tableros de control que son de interés para la organización.
- Desarrollar *data marts* que permitan resolver requerimientos analíticos de interés para el cliente.
- Proponer una arquitectura de solución de BI para los análisis a construir.
- Crear procesos de ETL.

DESCRIPCIÓN

Implementar dos requerimientos de analítica que consideren son de interés para el cliente del proyecto. En este proyecto se debe: (1) crear *data marts* con la o las tablas de hecho que considere necesarias, (2) cargar las tablas de hechos seleccionadas con sus respectivas dimensiones, y (3) diseñar e implementar los tableros de control. La descripción detallada de las actividades a realizar y de los elementos mínimos a incorporar en el documento se presenta a continuación:

1. **[10%] Identificar necesidades analíticas** siguiendo la metodología *Business Life Cycle* [3],[6]– Grupo Kimball (las referencias corresponden a las descritas en el cronograma del curso).

A partir de entrevistas realizadas a personas relacionadas con el proyecto que están trabajando, los datos compartidos y la retroalimentación a la consultoría realizada en semestres anteriores, identificar y documentar 5 requerimientos analíticos, utilizando la matriz de requerimientos de negocio, los temas analíticos, análisis requeridos, procesos de negocio, fuentes de datos y datos requeridos (en algunas referencias asociadas al curso, en la matriz de documentación, en cambio de “fuentes de datos y datos” aparece el título “comentario” o solo “datos”). Para este punto descargue el documento de Excel llamado Plantilla-Entrega1.xlsx, del enlace asociado a esta entrega. En ese documento de Excel debe llenar el formato de temas analíticos y análisis requeridos/inferidos que se encuentra en la primera hoja. Estos análisis son de analítica basados en tableros de control. Adicionalmente debe describir y justificar los análisis propuestos. La justificación debe incluir elementos de la entrevista realizada al cliente o de la revisión de proyectos similares a nivel nacional

 <p>Universidad de los Andes Colombia</p> <p>Accreditación institucional de alta calidad 10 años Protección de la Educación Resolución 182 del 9 de enero de 2023</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Semestre: 2022-10</p>	
--	---	---

o mundial. La invitación en este punto es a hacer una revisión del estado del arte que en este caso corresponde a revisar proyectos similares relacionados con **analítica sobre indicadores socioeconómicos**. Los requerimientos analíticos no seleccionados para implementar en esta entrega se consideran sugerencias para un futuro proyecto en la misma temática.



2. **(15%) Modelar Data Marts:** A partir de las fuentes de datos anexas a esta entrega y de requerimientos analíticos identificados por ustedes como resultado de la entrevista a realizar al cliente, proponga el modelo dimensional que representa mejor los tipos de requerimientos que tiene en mente el cliente. En particular usted debe:
 - a. (7,5%) Elaborar los modelos dimensionales propuestos. Entregue una representación gráfica de los modelos multidimensionales. En el modelo se deben representar nombres de atributos, llaves primarias, llaves foráneas y roles. Evite diagramas ilegibles.
 - b. (7,5%) Justificación del modelo

Para cada tabla de hechos

 - i. (2%) Especificar y justificar la granularidad
 - ii. (3%) Definir los hechos/medidas que contiene. Para cada medida indicar y justificar el tipo de medida (aditiva, semi-aditiva y no aditiva).

Para cada dimensión:

 - iii. (2%) **Para cada atributo**, si se requiere, especificar el tipo de manejo de historia (1,2, 3,...) de variación lenta (*Slowly Changing Dimension*), y justificar la elección.
3. **(30%) Entendimiento de los datos, creación del Datamart y proceso ETL.** Perfilamiento de datos y análisis de calidad, creación de la base de datos y procesos ETL. Tomando como base las fuentes de datos proporcionadas y los ejemplos de requerimientos analíticos de la consultoría anterior, realice los siguientes pasos.
 - a. (15%) **Entender las fuentes de datos** recibidas y presentar el resultado del análisis (estadísticos de los datos y análisis de calidad de datos).
 - b. (15%) **Diseñar e implementar el proceso de ETL.** En este punto debe describir las actividades (o tareas del flujo de trabajo en la herramienta utilizada), realizadas en el proceso.
4. **(30%) Proponer la arquitectura de solución** de BI a nivel de componentes e **implementar** dos tableros de control. Recuerde que un tablero de control está compuesto por más de un indicador de interés para una organización.
 - a. (8%) **Proponer la arquitectura de solución** de BI para resolver los análisis realizados hasta este momento en el proyecto.
 - b. (22%) **Implementar los tableros de control**, utilizando un software especializado como PowerBI, Tableau, GoogleDataStudio, etc. y conectándose a la base de datos donde tiene los datos que representan el modelo multidimensional propuesto.

 <p>Universidad de los Andes Colombia</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Semestre: 2022-10</p>	
--	---	---

5. **(10%) Preparar** un video acompañado de una presentación para **la junta directiva de la empresa** donde venda la solución propuesta en esta etapa del proyecto de BI. Esta presentación debe estar en un video de máximo 8 minutos. Este video debe incluir en la primera parte una presentación más ejecutiva sin entrar en detalles técnicos, en la cual, muestre la **interacción** con los tableros de control contruidos y muestre hallazgos obtenidos gracias a su uso. En la segunda parte del video, incluya detalles técnicos de los modelos dimensionales, procesos ETL entre otros elementos que considere son de interés para el área técnica de la empresa.
6. **[5%] Describir las actividades realizadas**, el tiempo asignado y el responsable o responsables de cada punto de esta entrega. Adicionalmente, reparta 100 puntos entre los integrantes del grupo de acuerdo con el aporte al desarrollo del proyecto
7. Si tiene algún avance del bono de estadística de la etapa 1 proyecto 1, inclúyala en este entregable.

ENTREGA Y EVALUACIÓN

- El proyecto se realiza en grupos de máximo 3 estudiantes.
- El documento a entregar tienen máximo 8 páginas (sin incluir portada, tabla de contenido, ni referencias), a una columna y con letra arial tamaño 12.
- Debe entregar las fuentes de los análisis (e.g. si es tableau, los archivos twb) y todo lo necesario para poder ejecutar los tableros de control en una sustentación o en la presentación con el cliente.
- Los entregables del proyecto podrán ser utilizados por Infraestructura Visible para inspirarse en la definición y desarrollo de nuevos proyectos, dando los créditos respectivos, sin autorización de los autores.
- La fecha máxima de entrega es el **lunes 30 de mayo a las 22:00**.
- La nota del proyecto es individual a pesar de entregar un trabajo a nivel de grupo. Por eso es importante la respuesta que dan en el punto 5.

ANEXO 1. Descripción de datos

La fuente original que quiere analizar el cliente se encuentra en esta página del Departamento Nacional de Planeación, en el siguiente link: <https://terridata.dnp.gov.co/index-app.html#/descargas>

Los invitamos a revisar su descripción y contenido. Sin embargo, para evitar problemas por cambio en los datos, las fuentes a utilizar estan disponibles en la sección unificada, dentro del espacio del segundo proyecto del curso.