

5.1. Ciclo de Elaboración

5.1.1. Inicio

T1 Tabla 1 Ciclo de Elaboración, fase de Inicio.
Fuente: Elaboración Propia

Tarea	Evaluar la preparación de la entidad para la implementación. Es decir, evalúa que tan lista está la organización/entidad para iniciar el proyecto de analítica.	
Descripción	Se evalúa los diferentes factores de implementación / administrativos, factores de construcción y ubicación, factores de tecnología, fuentes de datos y factores de volúmenes que podrían afectar la implementación del proyecto e identifique soluciones para mitigar el impacto de esos factores.	
Actividades	Producir plan del proyecto	
Roles	RA01 Patrocinador del Proyecto RA13 Científico de Datos RA19 Arquitecto Empresarial RA04 Gerente de Datos y Analítica RA05 Analista de Negocio RA06 Arquitecto de Datos	
Salidas	Plan de Proyecto. Declaración de trabajo (documento o contrato en donde se especifica los objetivos y entregables para el proyecto a realizar).	

5.1.2. Entendimiento del negocio

T2 Tabla 2 Ciclo de Elaboración, fase de Entendimiento del negocio
Fuente: Elaboración Propia

Tarea	Entender los objetivos del proyecto y sus requerimientos desde la perspectiva del negocio, para convertir este conocimiento en un proceso de minería de datos y en un plan preliminar que logre cumplir con los objetivos.	
Descripción	Se determinan los objetivos del negocio y se evalúa su situación como por ejemplo conocer los riesgos/ contingencias, requerimientos. Además, determinar los objetivos de la minería de datos.	
Actividades	Determinación de los objetivos de negocio: <ul style="list-style-type: none"> Recopilación del contexto del negocio 	

	<ul style="list-style-type: none"> • Objetivos de negocio • Criterios de éxito de negocio <p>Evaluación de la situación:</p> <ul style="list-style-type: none"> • Inventario de recursos, requerimientos, suposiciones y restricciones y contingencias. • Definición de Roles • Terminología • Análisis de costo beneficio. <p>Determinar objetivos de minería de datos Criterios de éxito de minería de datos</p>
Roles	RA02 Líder de Negocio RA03 Usuarios de Negocio RA04 Gerente de Datos y Analítica RA05 Analista de Negocio RA06 Arquitecto de Datos RA11 Coordinador de Metadatos RA13 Científico de Datos RA15 Analista de Datos RA16 Analista de Datos Geográficos
Salidas	Documento de entendimiento de negocio.

5.1.3. Descubrimiento y Entendimiento de datos

T3

Tabla 3 Ciclo de Elaboración, fase de Entendimiento de datos.
Fuente: Elaboración propia

Tarea	Acceder y explorar los datos con el fin de determinar la calidad de la información.
Descripción	Se recolectan los datos iniciales con el fin de describirlos, explorarlos y verificar su calidad.
Actividades	<p>Recolectar datos iniciales: Conseguir y acceder a los datos (si estos son adquiridos de diferentes fuentes será necesario integrarlos). Los datos pueden provenir de fuentes como:</p> <ul style="list-style-type: none"> • Datos existentes: incluye una gran variedad de información, por ejemplo: datos transaccionales, datos de encuestas y otros. • Compra de datos: Se usa en caso de necesitar datos que no se

encuentren disponibles en las fuentes de información de la organización.

- Datos adicionales: En caso de que los tipos de datos anteriores no sean suficientes se debe iniciar una exploración adicional de otros posibles almacenes de datos disponibles en la organización.

Descripción los datos: Describir los datos teniendo en cuenta los siguientes criterios

- Fuente
- Tipos de datos: Numérico, categórico, booleano, fecha, geográfico. Para el caso de los datos geográficos tener en cuenta el tipo de geometría y sistema de referencia.
- Dominios: Ejm. Sexo, rango etario, tipo de producto, etc.
- Volumen / Cantidad de datos
- Velocidad de cambio de los datos
- Variedad: Estructurado, No estructurado, semi estructurado
- Veracidad: Fiabilidad de los datos (Algunas de las fuentes utilizadas pueden tener características variables de fiabilidad)
- Valor: Valor que puede ser obtenido del análisis de los datos

Exploración de datos: Aborda las preguntas de minería de datos utilizando técnicas de consulta, visualización y reportes.

- Explorar distribución de los atributos clave, relaciones entre atributos, resultados de agregaciones simples, propiedades de subpoblaciones significativas y análisis estadísticos simples.
- Estos análisis pueden abordar directamente los objetivos de extracción de datos; también pueden contribuir o refinar la descripción de los datos y los informes de calidad, y alimentar la transformación y otros pasos de preparación de datos necesarios para un análisis más detallado⁷.
- Explorar los datos con las tablas, cuadros y otras herramientas de visualización.
- Formular hipótesis y dar forma a las tareas de transformación de datos que tienen lugar durante la preparación de los datos.

Roles	<p>Verificación de calidad de los datos</p> <ul style="list-style-type: none"> • ¿Están completos? • ¿Son correctos? • ¿Qué tipos de errores hay? • ¿Con qué frecuencia ocurren errores? • ¿Existen nulos? • ¿Hay datos duplicados? ¿Con qué frecuencia? ¿Cómo se representan? • ¿Hay errores en el formato de los datos? • ¿Hay inconsistencias en los dominios o códigos en los datos? • ¿Hay errores entre los metadatos del conjunto de datos y la información encontrada? • ¿Falta información clave para poder utilizar los datos? (Ejem. Sistema de referencia) 	
	RA06 Arquitecto de Datos RA07 Administrador de Datos RA08 Administrador de Base de Datos RA11 Coordinador de Metadatos RA13 Científico de Datos RA15 Analista de Datos RA16 Analista de Datos Geográficos RA03 Usuarios de Negocio RA05 Analista de Negocio	
Salidas	<p>Reporte de entendimiento de datos que debe contener:</p> <ul style="list-style-type: none"> • Reporte de recolección de datos. • Reporte de descripción de datos. • Reporte de exploración de datos. • Reporte de calidad de datos. 	

5.1.4. Preparación y gestión de datos

5.1.4.1. Diseñar y validar la infraestructura

T4

Tabla 4 Ciclo de Elaboración, fase de Preparación y gestión de datos:
Diseñar y validar la infraestructura.
Fuente: Elaboración Propia

Tarea	Diseñar estrategias de infraestructura técnica y seguridad
Descripción	Verificar que hardware, software y demás componentes técnicos se necesitan. También se diseñará la infraestructura de seguridad.
Actividades	<p>Diseño de infraestructura técnica:</p> <ul style="list-style-type: none"> • Diseñar la arquitectura de los entornos requeridos (Ejm: Analítico, Control de calidad y producción) en términos del hardware y software que se requieran • Validar la infraestructura técnica propuesta <p>Diseño de infraestructura de seguridad:</p> <ul style="list-style-type: none"> • Diseñar los mecanismos de autenticación y autorización requeridos • Documentar el modelo de seguridad y validarlo con los interesados (stakeholders) del proyecto
Roles	Arquitecto empresarial
Salidas	Documento de entendimiento de negocio.

5.1.4.2. Configurar entornos

T5

Tabla 5 Ciclo de Elaboración, fase de Preparación y gestión de datos:
Diseñar y validar la infraestructura.
Fuente: Elaboración Propia

Tarea	Configurar Entornos
Descripción	Configurar los diferentes entornos requeridos para las tareas de análisis ya sea en sitio o en la nube según el diseño y requerimientos establecidos.

Actividades	Configurar entorno analítico: Entorno donde se ejecutan las tareas relacionadas con la elaboración y pruebas de los modelos. Configurar entorno de control de calidad Configurar entorno de producción
Roles	RA19 Arquitecto Empresarial RA21 Administrador de tecnología RA20 Arquitecto de Soluciones
Salidas	Documento de entendimiento de negocio.

5.1.4.3. Preparar datos

T6

Tabla 6 . Ciclo de Elaboración, fase de Preparación y gestión de datos: Preparar datos.
Fuente: Elaboración Propia

Tarea	Preparar datos
Descripción	Preparación de los datos para las tareas de análisis. Por lo general esta es una de las actividades que más consume tiempo en el proceso.
Actividades	<ul style="list-style-type: none"> • Seleccionar datos: Decidir qué datos son útiles para el análisis. • Limpiar datos: Selección de subconjuntos, asignación de valores por defecto, etc. • Construir datos: Producción de atributos derivados o transformación de datos. • Integrar datos: Integración de datos de múltiples fuentes. • Formatear datos: Transformación de datos respecto a formatos sin cambiar su significado.
Roles	RA13 Científico de Datos RA14 Investigadores especializados RA15 Analista de Datos RA16 Analista de Datos Geográficos RA09 Arquitecto/Diseñador de ETL RA10 Desarrollador de ETL
Salidas	Documento detallado de preparación de datos

5.1.5. Validación de datos

T7 Tabla 7 Ciclo de Elaboración, fase de Validación de datos.
Fuente: Elaboración Propia

Tarea	Validación de datos
Descripción	<p>La validación de los datos ocurre inmediatamente después de la preparación de los datos y antes del modelado. Este paso es necesario porque durante la preparación de datos existe una gran posibilidad de ocurrencia de errores, especialmente en escenarios complejos. La validación de datos garantiza que el modelado se realice en los datos correctos⁸.</p> <p>Causas comunes de errores:</p> <ul style="list-style-type: none"> • Falta de comprensión adecuada de los datos y por tanto aplicación de lógica errónea durante la preparación. • Errores en los programas ejecutados para la preparación de los datos que generan salidas defectuosas. • Formatos de datos especiales que pueden generar errores durante la preparación y que no generan alertas de error por los programas utilizados para la preparación de datos. <p>Ejemplo: en algunas ocasiones al aplicar herramientas de geoprocésamiento como Dice ⁹ puede suceder que el área de los polígonos se vea afectada.</p>
Actividades	Validar los datos resultantes de la etapa de preparación. En caso de encontrar errores, retornar a la etapa de preparación de datos.
Roles	RA13 Científico de Datos RA14 Investigadores especializados RA15 Analista de Datos RA16 Analista de Datos Geográficos RA18 Ingeniero de Datos RA09 Arquitecto / Diseñador de ETL RA10 Desarrollador de ETL
Salidas	Documento de validación de los datos generados en el proceso de preparación.

5.1.6. Análisis

T8

Tabla 8 . Ciclo de Elaboración, fase de Análisis.
Fuente: Elaboración Propia

Tarea	Análisis
Descripción	Construcción del modelo a través de múltiples iteraciones en donde se van refinando los diferentes parámetros o en ocasiones se determinan nuevos requerimientos previos para la preparación de los datos.
Actividades	<p>Seleccionar técnica del modelo según los datos disponibles, objetivos de análisis y requerimientos específicos de modelamiento. Consideraciones a tener en cuenta:</p> <ul style="list-style-type: none"> • ¿Se requiere separación de datos entre pruebas y entrenamiento? • ¿Existen datos suficientes para tener resultados confiables del modelo? • ¿El modelo requiere algún nivel especial de calidad de datos?, ¿Los datos actuales cumplen con esos niveles de calidad? • ¿Los datos existentes son apropiados o se necesitan transformaciones adicionales? <p>Generar diseño de la prueba: Generar mecanismo que permita evaluar la calidad y validez del modelo.</p> <p>Construir modelo: Ejecutar la herramienta del modelo seleccionado con los datos. Crear uno o más modelos. Para cada modelo determinar:</p> <ul style="list-style-type: none"> • ¿Se obtuvieron conclusiones significativas? • ¿Se encontraron hallazgos (insights) relevantes o patrones inusuales? • ¿Existieron problemas durante el procesamiento? ¿El tiempo de procesamiento es razonable? • ¿El modelo tuvo problemas con la calidad de datos? • ¿Se encontraron inconsistencias significativas en los cálculos? <p>Evaluar modelo: Determinar cuál o cuáles modelos son acertados o efectivos para ser considerados finales.</p> <ul style="list-style-type: none"> • Realizar evaluación integral de los modelos: <ul style="list-style-type: none"> ○ Efectividad ○ ¿Los resultados son lógicos? ○ ¿Los resultados son muy simples o triviales? • Llevar registro de los parámetros utilizados para el afinamiento del modelo

Roles	RA13 Científico de Datos RA14 Investigadores especializados
Salidas	Documento de análisis.

5.1.7. Prototipo

T9 Tabla 9 Ciclo de Elaboración, fase de Prototipo.
Fuente: Elaboración Propia

Tarea	Prototipo
Descripción	Construir un prototipo para los casos cuando los requerimientos impliquen la disposición de un prototipo de aplicación adicional a los modelos seleccionados.
Actividades	Definir arquitectura del prototipo: Puede ser aplicaciones basadas en software comercial o software desarrollado a la medida. Diseñar prototipo: Definir flujo de trabajo del prototipo. Integrar o implementar prototipo: Ejem. Visualizaciones. Evaluar prototipo: Determinar si la aplicación cumple con los criterios de éxito de negocio.
Roles	RA12 Desarrollador de BI RA13 Científico de Datos RA04 Gerente de Datos y Analítica RA05 Analista de Negocio RA06 Arquitecto de Datos RA20 Arquitecto de Soluciones
Salidas	Documento de arquitectura y diseño del prototipo. Prototipo.

5.1.8. Validación y evaluación

T10

Tabla 10 Ciclo de Elaboración, fase de Validación y evaluación.
Fuente: Elaboración Propia

Tarea	Validación y evaluación
Descripción	Determinar cuál o cuáles modelos son acertados o efectivos para ser considerados finales acordes a los criterios de éxito del negocio.
Actividades	<p>Evaluar resultados: Evaluar los resultados según los criterios de éxito de negocio.</p> <ul style="list-style-type: none"> Entender los resultados del proceso de análisis Interpretar los resultados en términos de su aplicación Verificar el efecto en los objetivos planteados para el análisis Verificar los resultados del análisis en comparación al conocimiento base de los hechos con el fin de evaluar si la información descubierta es novedosa y útil Evaluar los resultados respecto a los criterios de éxito de negocio y los objetivos de negocio originales. Determinar si hay nuevos objetivos de negocios para abordar después en este u otros proyectos. Generar recomendaciones sobre futuros proyectos de análisis. Validar y aprobar los modelos según los resultados de la validación. <p>Revisión del proceso:</p> <ul style="list-style-type: none"> Proveer un resumen del proceso realizado Generar recomendaciones sobre el proceso realizado, fallas y sugerencias de mejora. <p>Determinar pasos siguientes</p> <ul style="list-style-type: none"> Opción 1: Pasar al ciclo de publicación Opción 2: Volver a ejecutar el ciclo de análisis para refinamiento o reemplazo de los modelos.
Roles	RA04 Gerente de Datos y Analítica RA13 Científico de Datos RA14 Investigadores especializados RA20 Arquitecto de Soluciones
Salidas	Documento de validación y evaluación.

5.1.9. Nueva adquisición de datos

T11

Tabla 11 . Ciclo de Elaboración, fase de Nueva adquisición de datos.
Fuente: Elaboración Propia

Tarea	Nueva adquisición de datos	
Descripción	<p>Identificar e incorporar nuevos datos que anteriormente no fueron tenidos en cuenta para que alimente el modelo y prototipo del ejercicio en cuestión.</p> <p>La Validación y evaluación del prototipo de analítica construido, puede arrojar la necesidad de incluir en el modelo del prototipo mismo, datos adicionales no tenidos en cuenta anteriormente, los cuales deberían ayudar a responder el requerimiento analítico del negocio.</p>	
Actividades	<p>Listar los hallazgos más significativos en la validación y evaluación de la fase preliminar.</p> <ul style="list-style-type: none"> • Consolidar el resultado de la validación • Consolidar el resultado de la evaluación • Identificar variables adicionales accesibles, que sean requeridas para el mejoramiento o precisión del modelo <p>Aplicar las demás fases del ciclo de elaboración</p> <ul style="list-style-type: none"> • Ejecutar la fase de “Entendimiento del negocio” y/o de “Descubrimiento y entendimiento de datos” en la medida de que sean requeridas para las nuevas variables incorporadas • Ejecutar la fase de “Preparación y gestión de datos”, junto con sus subfases • Ejecutar la fase de “Análisis” • Ejecutar la fase de “Prototipo” • Ejecutar la fase de “Validación y evaluación” <p>Revisión del proceso</p> <ul style="list-style-type: none"> • Proveer un resumen del proceso realizado • Generar recomendaciones sobre el proceso realizado, fallas y sugerencias de mejora. <p>Determinar pasos siguientes</p> <ul style="list-style-type: none"> • Opción 1: Pasar al ciclo de Entendimiento del negocio • Opción 2: Pasar al ciclo de Descubrimiento y entendimiento de datos 	
Roles	RA06 RA07 RA08 RA11 RA13 RA15	Arquitecto de Datos Administrador de Datos Administrador de Base de Datos Coordinador de Metadatos Científico de Datos Analista de Datos
	RA16 RA03 RA05	Analista de Datos Geográficos Usuarios de Negocio Analista de Negocio
Salidas	Documento de Nueva adquisición de datos.	