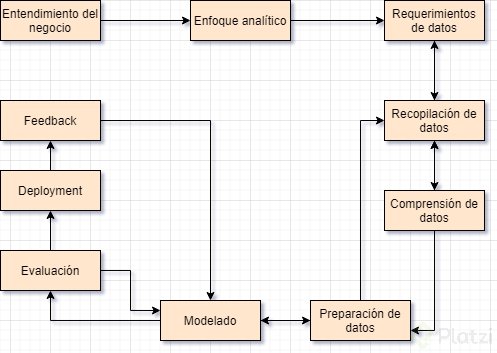
Asímismo, se requiere tomar en cuenta un “checklist” o serie de etapas por así decirlo



**1- Entendimiento del Negocio**  
Hace referencia a la pregunta ¿Cuál es el problema que estás tratando de resolver?, asimismo debemos conocer el campo de estudio donde vamos a trabajar con el análisis para definir una pregunta que nos dará las primeras directrices.

**2- Enfoque analítico**  
Relacionado con la pregunta ¿Cómo puedes usar los datos para responder la pregunta?, una vez establecida la directriz en esta etapa utilizaremos herramientas estadísticas y de machine learning(identifica relaciones y tendencias en datos que de otra manera no serían accesibles o identificados), asimismo la aplicación de un enfoque analítico:

* Descriptivo -> estado actual.
* Diagnostico (análisis estadístico) -> ¿Qué pasó? y ¿Por qué está pasando esto?.
* Predictivo (previsión) -> ¿Qué pasa si estas tendencias continúan? y ¿Qué pasará después?.
* Prescriptivo -> ¿Cómo lo resolvemos?.

**Si la pregunta requiere determinar las probabilidades de una acción:**

* Deberías usar un modelo predictivo.

**Si la pregunta requiere mostrar relaciones:**

* Deberías usar un modelo descriptivo.

**Si la pregunta requiere una respuesta sí / no:**

* Deberías usar un modelo de clasificación.

**3- Requerimiento de datos**  
Dependiendo del enfoque analítico vamos a saber qué tipos de datos, formatos y presentaciones vamos a necesitar (¿Qué datos necesitamos para responder la pregunta? y ¿Cuáles son los requisitos?).

Es un punto importante, por ejemplo si quieres cocinar una carne mongoliana y no conoces los ingredientes(datos) y la receta, no vas a poder determinar qué vas a necesitar para cocinarla solamente con tu criterio, posiblemente puedas acercarte, sin embargo, ¿Vamos a obtener el resultado esperado?.

**4- Recopilación de datos**  
Relacionada con la pregunta: ¿De dónde provienen los datos y cómo los obtendrá?.  
En esta etapa se recopilan los datos para el dominio del problema, esta etapa tiene revisión continua.

Es necesario ser muy meticuloso en esta etapa, asimismo esta etapa da a conocer si necesitamos o no más datos.

**5- Compresión de datos**  
Relacionada con la pregunta: ¿Los datos que recopiló son representativos del problema a resolver? y ¿Qué significa preparar o limpiar datos?.  
En esta etapa se utiliza la estadística descriptiva como también visualización de datos, en el caso de encontrar incongruencias es necesario regresar al paso anterior.

**6- Preparación de datos**  
Relacionada con la pregunta: ¿Qué trabajo adicional se requiere para manipular y trabajar los datos? y ¿De qué forma se preparan los datos?.  
En esta etapa vamos a construir el conjunto de datos a utilizar, en pocas palabras vamos a preparar los datos para el modelado.

**7. Modelado**  
Relacionada con la pregunta: ¿De qué manera se pueden visualizar los datos para llegar a la respuesta?.  
En esta etapa se utilizan los conjuntos de datos preparados en la anterior etapa, asimismo nos dedicamos a desarrollar modelo predictivos y descriptivos(hay que tomar en cuenta el enfoque analítico que se tomó en el punto 2).

**8.- Evaluación**  
Relacionada con la pregunta: ¿El modelo utilizado realmente responde a la pregunta inicial o necesita ser ajustado?.  
En esta etapa se evalúa el modelado para comprender su calidad y garantizar que el mismo nos ayude a responder la pregunta inicial, a través de esta etapa podemos saber si necesitamos ajustar el modelado para que este cumpla su objetivo.

**9- Deployment**  
Relacionada con la pregunta: ¿Puedes poner el modelo en práctica?.  
En esta etapa se realiza la implementación del modelado, se realiza de forma gradual hasta que el rendimiento del mismo sea evaluado en su totalidad, se podría decir que es un proceso netamente operativo.

**10- Feedback**  
Relacionada con la pregunta: ¿Puedes obtener retroalimentación constructiva para responder la pregunta?.  
En esta etapa se recibe toda la retroalimentación posible con fines de ajustar el modelo para mejorar su precisión y utilidad. La idea es obtener mejores resultados.

Cabe oportuno mencionar que no es la única metodología existente.

Sin nada más que agregar, un saludo cordial a todos.