

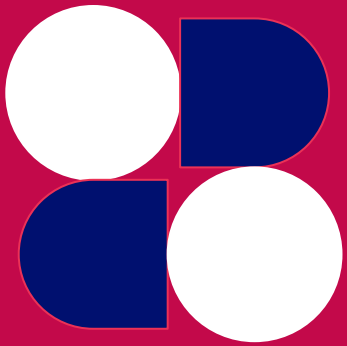


Programa PUCP DATA LITERACY

Formación tecnológica

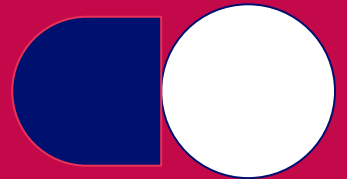
Dirección Académica de Planeamiento y Evaluación
Dirección de Gestión del Talento Humano





SESIÓN 6:

Planeación de proyectos de analítica





Contenido:



1

Ciclo de vida de un proyecto de analítica

2

Roles en los proyectos de analítica de datos

3

ChatGPT en proyectos de analítica

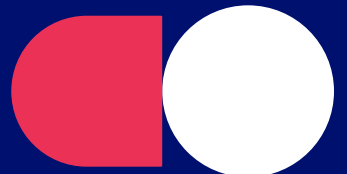


Al finalizar la sesión, lograrás:

- Identificar los componentes clave involucrados en la estructuración de un proyecto de analítica de datos.
- Comprender los diferentes roles involucrados en un proyecto de analítica, como analista de datos, científico de datos y analista de negocios.
- Adquirir comprensión de las tecnologías y herramientas utilizadas en proyectos de analítica y cómo ChatGPT puede asistir en la definición de proyectos.



Ciclo de vida





"Una buena estrategia de datos no se determina por los datos que están fácilmente disponibles o potencialmente disponibles; se trata de lo que tu organización quiere lograr y cómo los datos pueden ayudarte a llegar allí."

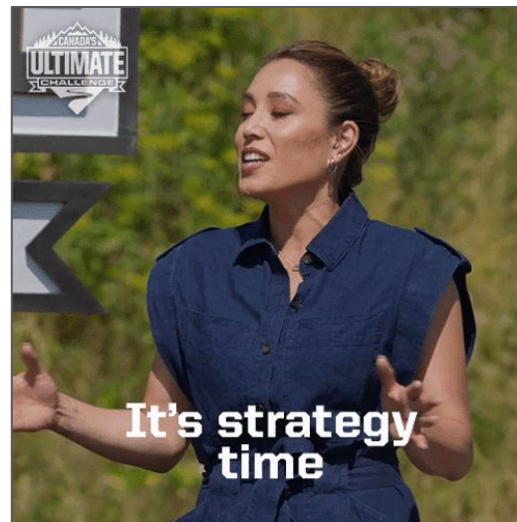
Marr (2017)



La estrategia



Un concepto central e integrado
que articula cómo los datos
habilitarán e inspirarán la
estrategia organizacional





¿A qué debe responder la estrategia de datos?



- ¿Cuál es la visión de la organización para los datos?
- ¿Cómo avanzará la organización en la ejecución de las actividades necesarias para conseguir los datos deseados?
- ¿Cómo impulsará la organización la adopción y el uso de datos?
- ¿Cuándo la organización ejecutará las actividades propuestas?
- ¿Cuál es la lógica económica de la organización para monetizar sus datos?



¿Por qué es importante?



Toma de
decisiones

Optimizar
procesos

Monetizar los datos

Reducir riesgos

Mantenerse
competitiva



Elementos de una estrategia de datos



•Adquisición, estandarizar y limpieza de datos

•Gobernanza y calidad de Datos

•Visualización e integración de los datos

•Acceso, distribución, privacidad y seguridad a los datos

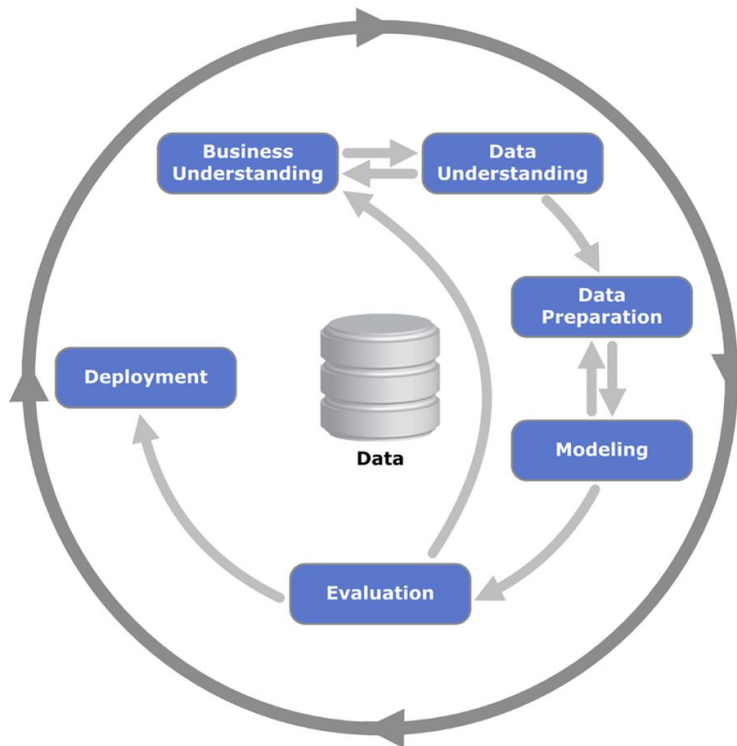
•Acuerdos de nivel de servicio

•Analítica de datos



Proyecto de analítica

Metodología CRISP - DM





Etapas de un proyecto de analítica



Entendimiento del
negocio



Características

- Enmarcar el problema
- Identificar a los interesados clave
- Identificar al líder del proceso de analítica
- Desarrollar hipótesis iniciales

ChatGPT

Actúa como consultor de analítica. Te daré un problema de negocio. Ayúdame a clarificar los objetivos, métricas de éxito y posibles preguntas analíticas para saber qué tipo de modelo analítico es más apropiado. Problema: [aquí va el caso]



Retos

- Mayoría de las empresas no realizan planeación estratégica.
- Recursos (tiempo, competencias técnicas y analíticas)
- Redactar el objetivo de negocio con claridad.





Etapas de un proyecto de analítica



Entendimiento de
los datos



Características

- Identificar fuentes de datos potenciales y cómo puede resolver la pregunta de investigación
- Requerimientos de los datos
- Recolectar los datos
- Entender los datos (diccionario de variables, cómo se recolectaron, cuándo, quién, dónde, para qué)

ChatGPT

- Con base en este problema y posible aproximación analítica, sugiere qué fuentes de datos serían necesarias, qué variables clave recolectar, y qué problemas de calidad de datos podrían surgir.
- O pueden indicar las fuentes de datos y variables con la que cuentan.



Retos

- Datos no existan.
- Costos de los datos (internos y externos)
- Captura y almacenamiento de datos no estructurados y costos.





Etapas de un proyecto de analítica



Preparación de los
datos



Características

- Explorar, preprocesar y limpiar datos.
- Resultados:
- Determinar qué datos están disponibles, que no están disponible y fuentes externas a la organización que podrían ser útiles.
- Reporte de Calidad de Datos

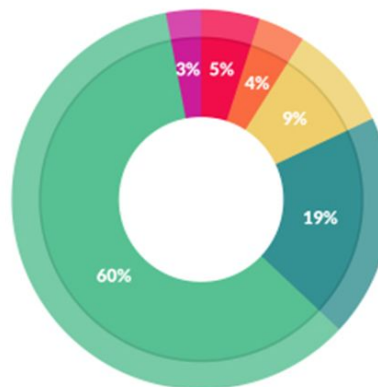
ChatGPT

Diseña un plan de preparación de datos: qué transformaciones, limpieza o integración se deben hacer para tener una base analítica lista.



Retos

- Datos no existan.
- Costos de los datos (internos y externos)
- Captura y almacenamiento de datos no estructurados y costos.

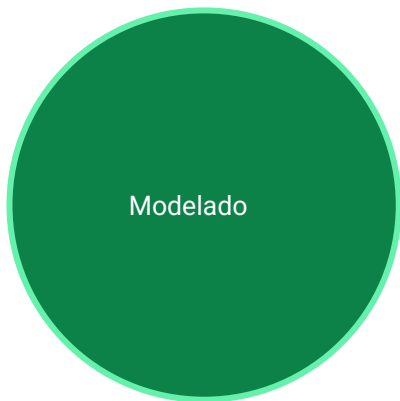


What data scientists spend the most time doing

- Building training sets: 3%
- Cleaning and organizing data: 60%
- Collecting data sets: 19%
- Mining data for patterns: 9%
- Refining algorithms: 4%
- Other: 5%



Etapas de un proyecto de analítica



ChatGPT

De acuerdo con el problema y los datos descritos, estable que métricas técnicas se usarían para evaluarlas y justifica por qué las incluirías



Etapas de un proyecto de analítica





Características

- ¿El modelo parece válido y exacto en los datos de prueba?
- ¿El resultado / comportamiento del modelo tiene sentido para los expertos en el dominio?
- ¿Los valores de los parámetros tienen sentido en el contexto del dominio?
- ¿Es el modelo suficientemente preciso para cumplir el objetivo?
- ¿Se necesitan más datos o insumos?

ChatGPT

Supongamos que ya tenemos un modelo entrenado con buena precisión técnica.

- ¿Cómo verifico que el modelo realmente responde al problema de negocio?
- ¿Qué criterios de éxito usaría la organización para adoptarlo?
- ¿Qué riesgos, costos u objeciones podría haber desde el negocio?
- ¿Cómo lo comparo con lo que la empresa hacía antes (baseline)?



Etapas de un proyecto de analítica



Despliegue



Características

- Comunicación de resultados
- Plan piloto – pequeña escala
- Controlar la precisión del modelo y volver a entrenar el modelo si es necesario
- Acceso a los resultados

ChatGPT

Describe cómo este modelo se puede llevar a la práctica en la organización.

- posibles formas de implementación (dashboard, alertas, informes, APIs),
- roles que deberían usarlo,
- acciones que podrían tomarse a partir de sus resultados



Despliegue entregables



Presentación
a tomadores
de decisión

+

Presentación
a usuarios

+

Código

+

Manual de
usuario y
manual
técnico

+

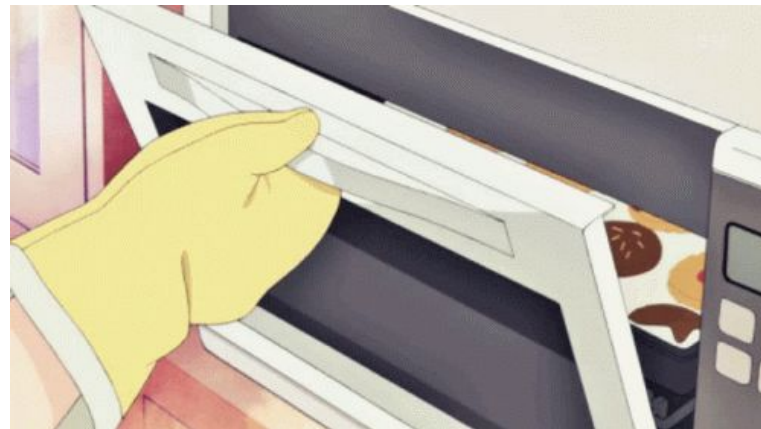
Soporte y
pruebas



¿Por qué es bueno iniciar con un piloto?



**¿Alguien alguna vez ha
horneado una torta o
una galleta?**





Razones



Validación rápida de conceptos

Menor inversión inicial

Experimentación y aprendizaje
iterativo

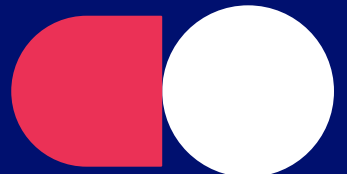
Ajustes a la realidad del negocio

Reducción de riesgos y medición
del ROI

Alineación con los objetivos



Roles en los proyectos de analítica





Roles



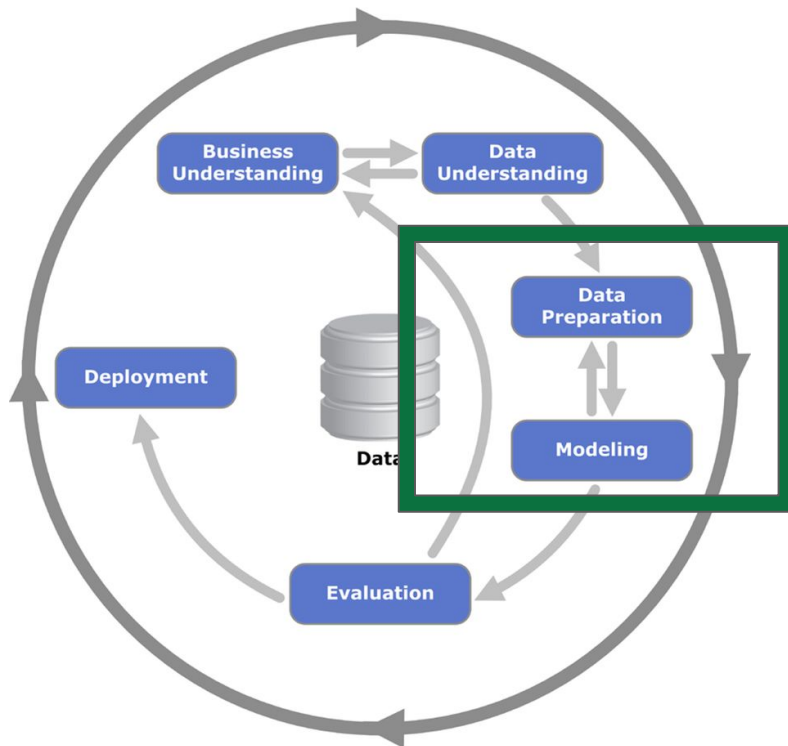
PUCP





Funciones y Responsabilidades:

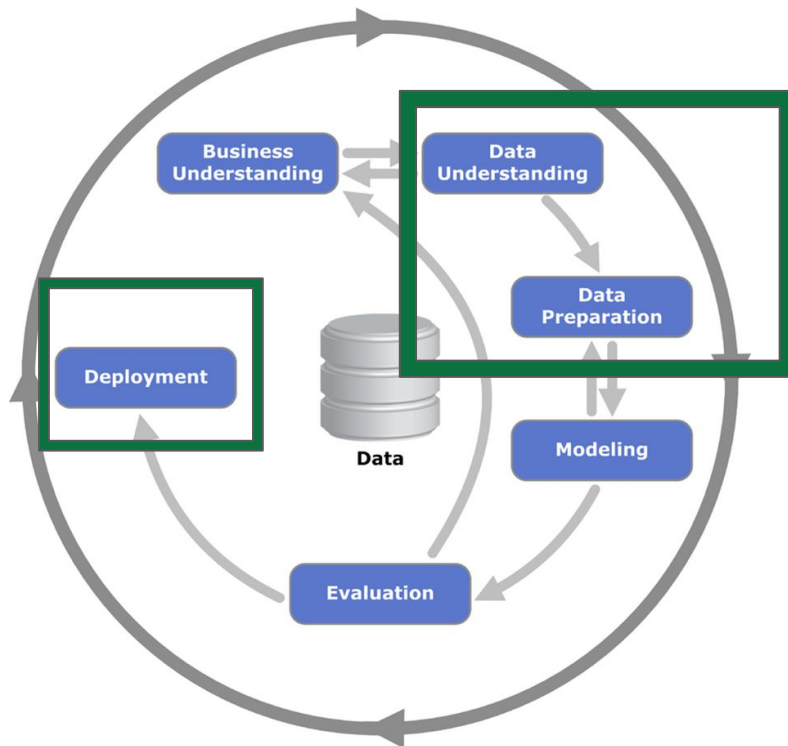
- **Análisis y Modelado:** El **científico de datos** es responsable de analizar los datos y aplicar técnicas estadísticas, algoritmos de machine learning y técnicas avanzadas para encontrar patrones, hacer predicciones y crear modelos.
- **Preparación de Datos:** Aunque el trabajo inicial de limpieza y transformación de datos puede ser realizado por un ingeniero de datos, el científico de datos también trabaja en la preparación de los datos para los modelos predictivos.
- **Creación de Modelos Predictivos:** Desarrolla modelos matemáticos o de machine learning que ayuden a predecir o clasificar eventos futuros basados en los datos.





Funciones y Responsabilidades:

- **Infraestructura y Almacenamiento de Datos:** El **ingeniero de datos** se encarga de diseñar, construir y mantener las infraestructuras necesarias para almacenar y procesar grandes volúmenes de datos (bases de datos, data lakes, pipelines de datos).
- **Integración de Datos:** Trabaja en la integración de datos provenientes de diversas fuentes, asegurando que los datos estén disponibles y sean accesibles para el análisis.
- **Automatización del Proceso de Datos:** Desarrolla pipelines de datos automatizados para transformar y cargar datos desde diferentes fuentes hasta las bases de datos.





Analista de Datos (Data Analyst)

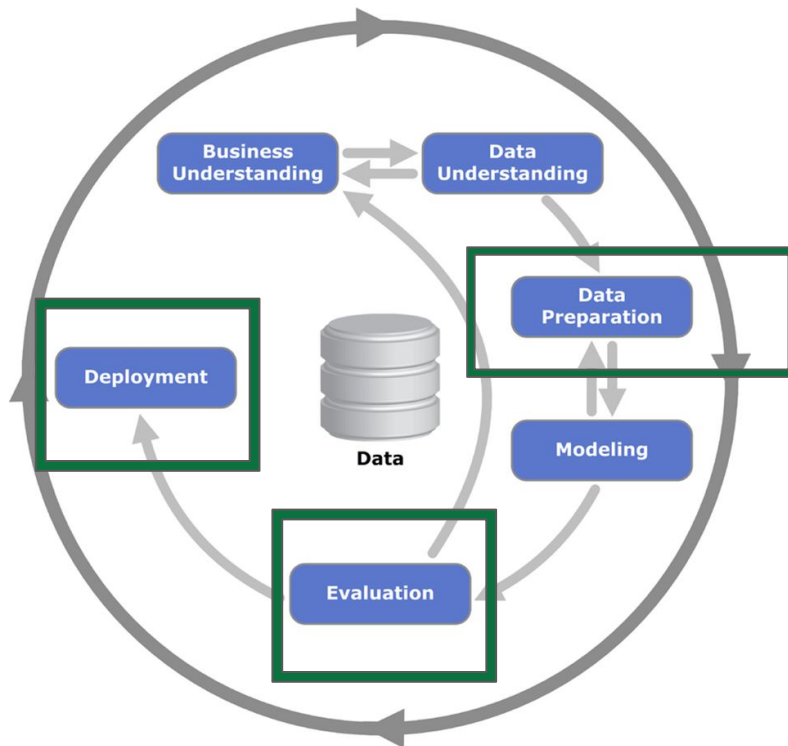


Funciones y Responsabilidades:

- **Exploración y Análisis de Datos:** El **analista de datos** se centra en analizar los datos para obtener insights valiosos. Utiliza herramientas de BI (Business Intelligence) como Power BI, Tableau, y Excel para crear informes y visualizaciones.
- **Visualización de Datos:** El analista crea gráficos y dashboards interactivos que ayudan a los stakeholders a tomar decisiones basadas en datos.
- **Interpretación de Resultados:** Su objetivo es presentar la información de manera clara y comprensible para los tomadores de decisiones.



Analista de Datos (Data Analyst)





Experto en Dominio (Business Analyst)

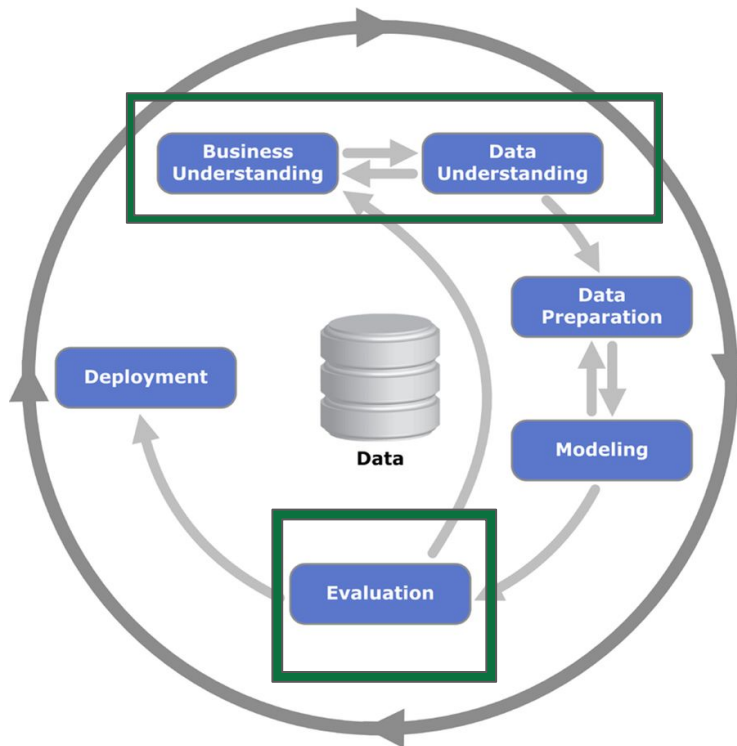


Funciones y Responsabilidades:

- **Definición de Requerimientos:** El **experto en dominio** o analista de negocios juega un papel crucial en la definición de los requisitos del proyecto. Trabaja estrechamente con los stakeholders para comprender sus necesidades y asegurarse de que el proyecto esté alineado con los objetivos de negocio.
- **Intermediario entre el negocio y el equipo técnico:** Su tarea principal es traducir las necesidades del negocio en términos que los técnicos puedan entender y viceversa.
- **Comunicación y Coordinación:** Es el punto de contacto entre el cliente y el equipo de analítica, asegurando que el proyecto esté siempre alineado con los intereses del negocio.



Experto en Dominio (Business Analyst)



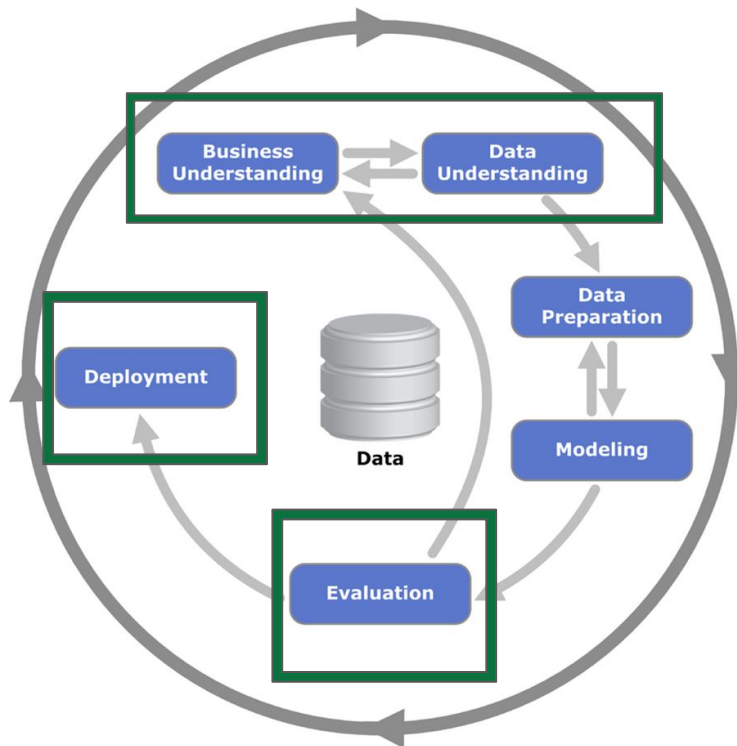


Funciones y Responsabilidades:

- **Definición de Necesidades:** Los stakeholders son los **usuarios finales** del proyecto. Pueden ser tanto del área académica como administrativa de la universidad. Ellos ayudan a definir las necesidades del proyecto desde la perspectiva del negocio.
- **Revisión y Validación:** Son responsables de revisar y validar los resultados del análisis para asegurarse de que se alinean con los objetivos del negocio.
- **Toma de Decisiones:** Finalmente, son los encargados de tomar decisiones basadas en los informes y modelos proporcionados por el equipo de analítica.



Stakeholders o Usuarios del Cliente





Gracias

