**Actividad 1**

**Nombre**: Carlos Saldivia Susperreguy

1. Principales competencias técnicas y habilidades personales que debe tener un ingeniero de datos.

**Competencias técnicas**

* **Programación y lenguajes de programación** como Python, SQL, R, Java, Scala.
* **Manejo y Gestión de Bases de Datos**
  + Bases de datos relacionales: PostgreSQL, MySQL, SQL Server.
  + Bases de datos no relacionales: MongoDB, Cassandra.
  + Optimización de consultas y manejo de integridad de datos.
* Procesamiento y Transformación de Datos
  + Uso de herramientas como Apache Spark, Hadoop, Airflow, etc.
  + Conocimiento de técnicas ETL (Extract, Transform, Load).
* **Machine Learning y Modelado Predictivo**
  + Aplicación de modelos supervisados (regresión, clasificación) y no supervisados (clustering).
* **Arquitectura de Datos**
  + Diseño y mantenimiento de pipelines de datos robustos y escalables.
* **Cloud Computing**
  + Experiencia con servicios en la nube: AWS, Azure, Google Cloud.

**Habilidades Personales**

* **Trabajo en equipo y la colaboración interdisciplinaria**: ya que interactúan a menudo con científicos de datos, ingenieros de software y otros profesionales.
* **El pensamiento analítico:** es clave para evaluar grandes volúmenes de datos, descubrir patrones y proponer soluciones.
* **Resolución de problemas complejos**: dada la naturaleza de los desafíos técnicos en los proyectos de ingeniería de datos.
* **La comunicación efectiva:** es fundamental para explicar conceptos técnicos complejos a audiencias no técnicas y crear documentación clara.

1. **Ejemplo de una empresa o industria donde los ingenieros de datos sean clave y explicar su función en dicho contexto:**

Los ingenieros de datos son fundamentales para el análisis de grandes volúmenes de datos en diferentes rubros como:

* **Finanzas y Banca**: Creación de sistemas para la optimización de datos financieros, la detección de fraudes y el análisis de datos de mercado.
* **Salud**: Creación y gestión de infraestructuras de datos para el análisis de datos médicos y de pacientes, con aplicaciones en medicina predictiva.
* **Retail y Comercio Electrónico**: Manejo de datos de clientes, comportamiento de compra y análisis predictivos para mejorar la experiencia del consumidor.
* **Energía y Utilities**: Optimización de la gestión de recursos energéticos mediante el análisis y procesamiento de datos generados por redes inteligentes.
* **Gobierno y Organizaciones Públicas**: Gestión de grandes volúmenes de datos para la mejora de servicios públicos, análisis de políticas públicas y desarrollo de infraestructuras tecnológicas.

1. **Comparación de herramientas de bases de datos**

|  |  |  |
| --- | --- | --- |
| Tipo de base de datos | Relacional | No relacional |
| Sistema | PostgreSQL, MySQL | MongoDB |
| Estructura de datos | Tablas con esquemas definidos | Documentos flexibles |
| Lenguaje de consulta | SQL estándar | MongoDB Query Language (MQL) |
| Utilidad | Administración eficiente y optimización de consultas. | Manejo de datos no estructurados y alta escalabilidad. |