



## Actividad | 3 | Cotizaciones

### Minería de datos

Ingeniería en Desarrollo de Software

**TUTOR: Félix Acosta Hernández**  
**ALUMNO: Hazaél Atlai Hervert Martínez**  
**FECHA: 24/10/2024**



# Índice

<b>Portada-----</b>	<b>1</b>
<b>Índice -----</b>	<b>2</b>
<b>introducción -----</b>	<b>3</b>
<b>Descripción -----</b>	<b>3</b>
<b>Justificación -----</b>	<b>4</b>
<b>Etapa 1:Software de Data Mining – Perfiles y roles – Proceso del proyecto -----</b>	<b>5-10</b>
<b>Etapa 2:Tabla de cotizaciones – Preguntas – Propuesta-----</b>	<b>11-17</b>
<b>Conclusión -----</b>	<b>18</b>
<b>Referencias -----</b>	<b>19</b>

# **Introducción**

En el presente trabajo se desarrolla una propuesta para la implementación de un proyecto de minería de datos en el área de Inteligencia de Negocios de una organización. Este tipo de proyectos resulta fundamental en un entorno empresarial cada vez más orientado al análisis de grandes volúmenes de datos para la toma de decisiones. La minería de datos permite identificar patrones ocultos en los datos que pueden ofrecer valiosas oportunidades de optimización y mejoras en la eficiencia operativa. A lo largo de este proyecto, se analizarán los recursos necesarios, incluyendo software, hardware y personal especializado, con el fin de generar una tabla de cotización detallada que permita tener una visión clara de los costos y el tiempo estimado para la ejecución del proyecto. El objetivo es presentar una solución técnica y económicamente viable que responda a las necesidades actuales del área, aportando un valor significativo en el análisis de datos y facilitando la toma de decisiones.

# **Descripción**

El proyecto propuesto se centra en el diseño e implementación de un sistema de minería de datos aplicado al área de Inteligencia de Negocios. La minería de datos es una técnica que permite extraer información útil a partir de grandes volúmenes de datos, lo cual es esencial en entornos empresariales donde la toma de decisiones estratégicas depende del análisis preciso de la información disponible. Este proyecto implica la selección de las herramientas de software adecuadas, la identificación de la infraestructura de hardware necesaria y la definición de los perfiles profesionales que participarán en el desarrollo. Además, se considera el mantenimiento de los equipos y sistemas a lo largo del tiempo. A través de este enfoque integral, se busca ofrecer una solución que optimice los procesos de análisis y permita a la organización manejar de manera más eficiente sus recursos de información. La propuesta también incluye una tabla de cotización detallada que refleja los costos estimados para la adquisición e implementación de los recursos necesarios.

## **Justificación**

La razón por la que este proyecto de minería de datos es importante es porque, hoy en día, las empresas manejan grandes cantidades de información y necesitan herramientas que les ayuden a entenderla mejor. La minería de datos es una forma muy útil de analizar esa información, ya que permite encontrar patrones y datos importantes que no se ven fácilmente. Con esto, las empresas pueden tomar mejores decisiones, ser más eficientes y ahorrar recursos. Además, el uso de tecnología avanzada, como software especializado y equipos adecuados, ayuda a hacer el trabajo más rápido y preciso. Invertir en este tipo de proyectos es clave para que las empresas sean más competitivas y puedan adaptarse a los cambios del mercado.

## Etapa 1

# Software de data mining

Tecnología de Minería de datos	¿Por qué propondrías ese software?	¿Qué procesos de minería de datos puede realizar el software?
Oracle SQL	<p>como jefe del proyecto yo usaría Oracle sql por estas razones:</p> <p><b>Integración directa:</b> Acceso y manipulación de grandes volúmenes de datos sin moverlos a otras plataformas.</p> <p><b>Funciones avanzadas:</b> Herramientas para análisis complejos como agregaciones y análisis de series temporales.</p> <p><b>Escalabilidad:</b> Capacidad para manejar millones o miles de millones de registros.</p> <p><b>Modelos de Data Mining:</b> Algoritmos de clasificación, regresión, clustering, entre otros, integrados en Oracle Database.</p> <p><b>Optimización del rendimiento:</b> Uso de Oracle Optimizer para ejecutar consultas y análisis de manera eficiente.</p> <p><b>Seguridad:</b> Controles de acceso robustos y características de seguridad para proteger los datos.</p> <p><b>Visualización y reporting:</b> Integración con herramientas para representar los resultados gráficamente.</p>	<p><b>Clasificación:</b> Asigna categorías basadas en patrones.</p> <p><b>Regresión:</b> Predice valores continuos como ingresos o demanda.</p> <p><b>Clustering:</b> Agrupa datos según similitudes.</p> <p><b>Asociación:</b> Identifica relaciones entre elementos, como productos comprados juntos.</p> <p><b>Ánálisis de series temporales:</b> Estudia patrones a lo largo del tiempo.</p> <p><b>Detección de anomalías:</b> Encuentra datos fuera de lo esperado, útil para detectar fraudes.</p> <p><b>Reducción de dimensionalidad:</b> Simplifica conjuntos de datos eliminando atributos irrelevantes.</p>
Amazon web services	<p>Como jefe de proyecto, usaría Amazon Web Services (AWS) para minería de datos por las siguientes razones:</p> <p><b>Escalabilidad infinita:</b> Permite ajustar recursos bajo demanda según las necesidades del proyecto.</p> <p><b>Amplia gama de servicios:</b> Ofrece herramientas como Amazon SageMaker, Amazon Redshift y Amazon EMR para análisis de datos.</p> <p><b>Costo eficiente:</b> Modelo de pago por uso que optimiza los costos según el consumo.</p> <p><b>Seguridad y cumplimiento:</b> Garantiza altos niveles de seguridad y cumplimiento de normativas internacionales.</p> <p><b>Disponibilidad global:</b> Infraestructura en múltiples regiones que asegura alta disponibilidad y tiempos de respuesta rápidos.</p> <p><b>Integración con IA:</b> Servicios de inteligencia artificial que facilitan el análisis y la detección de patrones.</p> <p><b>Facilidad de uso:</b> Ecosistema amigable y bien documentado que acelera el desarrollo del proyecto.</p>	<p><b>Clasificación:</b> Asignar categorías a datos.</p> <p><b>Regresión:</b> Predecir valores continuos.</p> <p><b>Clustering:</b> Agrupar datos similares.</p> <p><b>Asociación:</b> Identificar relaciones entre variables.</p> <p><b>Ánálisis de series temporales:</b> Estudiar datos a lo largo del tiempo para detectar tendencias.</p> <p><b>Detección de anomalías:</b> Encontrar datos que se desvian de patrones esperados.</p> <p><b>Reducción de dimensionalidad:</b> Simplificar conjuntos de datos eliminando atributos irrelevantes.</p> <p><b>Optimización:</b> Encontrar soluciones óptimas en problemas complejos.</p>
Azure	<p>Como jefe del proyecto yo usaría azure por estas razones:</p> <p><b>Integración en la nube:</b> Facilita el acceso a datos y servicios de análisis.</p> <p><b>Herramientas de machine learning:</b> Proporciona algoritmos para construir y desplegar modelos fácilmente.</p> <p><b>Escalabilidad:</b> Permite ajustar recursos según las necesidades del proyecto.</p> <p><b>Análisis en tiempo real:</b> Permite decisiones rápidas con Azure Stream Analytics.</p> <p><b>Seguridad y cumplimiento:</b> Garantiza la protección de datos sensibles y el cumplimiento de regulaciones.</p> <p><b>Integración con otras herramientas:</b> Se conecta con servicios de Microsoft y terceros para un flujo de trabajo ágil.</p> <p><b>Facilidad de uso:</b> Interfaz intuitiva y plantillas que simplifican la creación de modelos.</p> <p><b>Capacidades de visualización:</b> Herramientas como Power BI para presentar resultados de forma clara.</p>	<p><b>Clasificación:</b> Asignar etiquetas a datos.</p> <p><b>Regresión:</b> Predecir valores continuos.</p> <p><b>Clustering:</b> Agrupar datos similares.</p> <p><b>Ánálisis de asociación:</b> Identificar relaciones entre variables.</p> <p><b>Ánálisis de series temporales:</b> Estudiar datos a lo largo del tiempo para detectar tendencias.</p> <p><b>Detección de anomalías:</b> Encontrar datos que se desvian de patrones esperados.</p> <p><b>Reducción de dimensionalidad:</b> Simplificar conjuntos de datos eliminando atributos irrelevantes.</p> <p><b>Optimización:</b> Encontrar soluciones óptimas en problemas complejos.</p>

### ¿Cuál es el mejor gestor de base de datos para este proyecto?

Yo como jefe usaría Amazon Web Services (AWS) para el proyecto de minería de datos debido a su amplia gama de servicios diseñados para análisis y machine learning, como Amazon SageMaker y Redshift. La escalabilidad de AWS permite ajustar recursos fácilmente según las necesidades del proyecto, lo que optimiza costos con su modelo de pago por uso. Además, ofrece alta disponibilidad y confiabilidad, garantizando el acceso continuo a nuestros datos. Su robusta seguridad protege información sensible y cumple con regulaciones. La capacidad de realizar análisis en tiempo real con herramientas como Kinesis facilita la toma de decisiones rápida y fundamentada, haciendo de AWS una opción ideal para nuestras necesidades. Aunque claro ay algunos puntos a tomar a cuenta y que debemos de fortalecer para poder obtener buenos resultados.

### DESVENTAJAS

**Complejidad de la plataforma:** La variedad de servicios puede ser abrumadora, aumentando la curva de aprendizaje para los usuarios sin experiencia.

**Costos acumulativos:** Aunque es flexible, el modelo de pago por uso puede resultar en costos inesperados si no se gestionan adecuadamente.

**Dependencia de la conectividad a Internet:** Requiere una conexión constante, lo que puede ser problemático en áreas con acceso limitado.

**Gestión de seguridad y cumplimiento:** La responsabilidad compartida en seguridad implica que la organización debe mantenerse vigilante para cumplir con normativas.

**Curva de aprendizaje:** La capacitación del personal puede llevar tiempo y recursos, retrasando la implementación efectiva del proyecto.

# Perfiles y Roles

## ¿Qué roles o perfiles escogiste para el desarrollo del proyecto?

- Director del Departamento
- Project Manager (IT, Negocios)
- Product Manager
- Arquitecto de Soluciones (Ingenieros en Software)
- Ingeniero de Redes (DevOps)
- Científico de Datos (Matemáticos, Físicos)
- Analistas de Datos (x2)
- Desarrollador Frontend
- Desarrollador Backend
- Full Stack Developpe

## ¿Por qué son fundamentales?

- **Director del Departamento:** Es fundamental para asegurar que el proyecto esté alineado con los objetivos generales de la empresa y para tomar decisiones estratégicas.
- **Project Manager:** Coordina todos los esfuerzos del equipo y gestiona los recursos para que el proyecto sea ejecutado en tiempo y forma.
- **Product Manager:** Asegura que el producto resultante de la minería de datos sea útil y cumpla con las expectativas del cliente o usuario final.
- **Arquitecto de Soluciones:** Garantiza que la estructura técnica sea adecuada y escalable para manejar grandes volúmenes de datos.
- **Ingeniero de Redes:** Es esencial para garantizar que la infraestructura necesaria para el análisis de datos funcione de manera óptima.
- **Científico de Datos:** Responsable de extraer valor a partir de los datos a través de técnicas matemáticas y estadísticas avanzadas.
- **Analistas de Datos:** Son cruciales para transformar datos sin procesar en conjuntos que puedan ser usados por el científico de datos.
- **Desarrolladores (Frontend, Backend, Full Stack):** Son indispensables para crear la infraestructura digital que permitirá interactuar con los resultados de la minería de datos.

## ¿Cuánto personal se va a contratar?

- **Director del Departamento:** 1
- **Project Manager:** 1
- **Product Manager:** 1
- **Arquitecto de Soluciones:** 1
- **Ingeniero de Redes (DevOps):** 1
- **Científico de Datos:** 1
- **Analistas de Datos:** 2
- **Desarrollador Frontend:** 1
- **Desarrollador Backend:** 1
- **Full Stack Developer:** 1

## Propuesta de Software: Amazon Web Services (AWS)

### **Motivo de selección:**

AWS es una de las plataformas de servicios en la nube más robustas y ampliamente utilizadas en el mercado. Su selección para este proyecto de minería de datos se basa en los siguientes puntos:

1. **Escalabilidad:** AWS ofrece infraestructura elástica que permite escalar los recursos según el crecimiento del proyecto.
2. **Seguridad:** Proporciona un amplio conjunto de herramientas de seguridad y cifrado de datos, cumpliendo con estándares internacionales.
3. **Variedad de Servicios:** Incluye una amplia gama de servicios relacionados con almacenamiento de datos, análisis y machine learning, tales como Amazon S3, Redshift, AWS Glue y SageMaker.
4. **Costo Eficiente:** Su modelo de pago por uso permite optimizar el presupuesto del proyecto, ya que solo se paga por lo que se consume.
5. **Compatibilidad con Herramientas de Minería de Datos:** AWS se integra fácilmente con herramientas de minería de datos como Apache Spark, Hadoop, Jupyter Notebooks y más.
6. **Disponibilidad Global:** Gracias a su infraestructura distribuida, AWS garantiza alta disponibilidad y baja latencia para la ejecución del proyecto.

## **Propuesta del Personal Necesario:**

El organigrama presentado establece los siguientes roles y personal para el proyecto:

1. **Director del Departamento:** Juan supervisará y garantizará que el proyecto esté alineado con los objetivos estratégicos.
2. **Project Manager (IT, Negocios):** Se encargará de la gestión integral del proyecto, supervisando el cumplimiento de plazos y objetivos.
3. **Product Manager:** Definirá las características del producto y priorizará las funcionalidades.
4. **Arquitecto de Soluciones (Ingeniero en Software):** Diseñará la arquitectura técnica del sistema, asegurando la integración de AWS con las soluciones internas.
5. **Ingeniero de Redes (DevOps):** Se encargará de configurar y mantener la infraestructura necesaria en AWS.
6. **Científico de Datos:** Desarrollará los algoritmos y modelos de minería de datos utilizando los servicios de AWS.
7. **Analistas de Datos (x2):** Ayudarán en la limpieza y preparación de los datos.
8. **Desarrolladores (Frontend, Backend, Full Stack):** Crearán las interfaces necesarias y configurarán las bases de datos para interactuar con los resultados.

## ***Requisitos del personal:***

- **Conocimiento en AWS:** El personal técnico debe tener experiencia con la plataforma de AWS y certificaciones como AWS Certified Solutions Architect o AWS Certified Big Data.
- **Experiencia en Minería de Datos y Ciencia de Datos:** Deben estar familiarizados con herramientas como Python, R, Spark y otros frameworks de análisis de datos.
- **Habilidades en Desarrollo Web y API:** Para garantizar la correcta interacción entre la plataforma de análisis y las interfaces de usuario.

# Proceso del proyecto

## Ejecución del proyecto

Para garantizar el éxito del proyecto, se seguirán los siguientes pasos:

1. **Fase de Planificación:**
  - a. **Definición de Objetivos y KPIs:** Juan y su equipo definirán los objetivos de negocio específicos que se deben alcanzar con la minería de datos.
  - b. **Ánálisis de Requisitos:** El Project Manager y el Arquitecto de Soluciones identificarán los requisitos técnicos y del negocio para la infraestructura de AWS.
  - c. **Estimación de Costos:** Se desarrollará un presupuesto estimado basado en el uso de AWS y el costo de personal necesario.
2. **Fase de Diseño:**
  - a. **Diseño de Arquitectura Técnica:** El Arquitecto de Soluciones diseñará la arquitectura en la nube, especificando los servicios de AWS que se utilizarán, como S3 para almacenamiento, Redshift para el análisis de datos y SageMaker para machine learning.
  - b. **Plan de Desarrollo:** Se definirá el cronograma de desarrollo, dividiendo las tareas entre los equipos de backend, frontend y ciencia de datos.
3. **Fase de Implementación:**
  - a. **Desarrollo e Implementación Técnica:** Los ingenieros y desarrolladores comenzarán a implementar la infraestructura en AWS, configurando bases de datos, integrando herramientas de análisis y desarrollando las interfaces necesarias.
  - b. **Despliegue de Modelos de Minería de Datos:** El Científico de Datos implementará modelos predictivos utilizando SageMaker y herramientas de machine learning.
  - c. **Pruebas Unitarias y de Integración:** Cada componente desarrollado será sometido a pruebas para asegurar su correcto funcionamiento.
4. **Fase de Evaluación:**
  - a. **Pruebas Piloto:** Se realizarán pruebas piloto con datos reales para verificar el rendimiento del sistema.
  - b. **Evaluación de Resultados:** El equipo analizará los resultados obtenidos para validar la efectividad de los modelos de minería de datos.
  - c. **Optimización:** Se harán ajustes y optimizaciones en los modelos y la infraestructura según sea necesario.
5. **Fase de Despliegue Final:**
  - a. **Lanzamiento a Producción:** Una vez probada la solución, se procederá al despliegue final, haciendo que el sistema esté disponible para su uso en producción.
  - b. **Capacitación del Personal:** Se capacitará al personal operativo sobre cómo usar las herramientas y sistemas desarrollados.
  - c. **Monitoreo y Soporte:** El Ingeniero de Redes y DevOps se encargará de monitorear el

rendimiento del sistema y de realizar tareas de mantenimiento cuando sea necesario.

#### 6. Fase de Mantenimiento y Optimización Continua:

- a. **Monitoreo de KPIs:** El Project Manager y el Product Manager revisarán periódicamente el desempeño del sistema basado en los KPIs definidos al inicio del proyecto.
- b. **Mejoras Incrementales:** Dependiendo del crecimiento de los datos y los resultados obtenidos, el equipo de desarrollo y los científicos de datos realizarán mejoras continuas.

## Etapa 2

Tabla de Cotizaciones				
Amazon Web Services (AWS)				
Nombre	Precio individual	Cantidad	Descripción	Total
Amazon Simple Storage Service (S3)	48.87 USD	2T	es una herramienta poderosa y flexible que te permite almacenar y gestionar tus datos en la nube de forma segura, escalable y rentable.	Mensual 1097.38 USD
Amazon Athena	74.27 USD	50 consultas/día  10 GB	Un servicio de análisis interactivo que permite ejecutar consultas SQL directamente sobre datos almacenados en S3.	Anual 13,168.56 USD
Amazon Aurora MySQL-Compatible	309.84 USD	1 tb	Base de datos relacional compatible con MySQL y PostgreSQL que ofrece alto rendimiento y disponibilidad.	
Amazon QuickSight	268,00 USD	1 lector y 1 autor	es un servicio de inteligencia empresarial rápido y basado en la nube que facilita la entrega de información a todos los miembros de su organización.	
Amazon Virtual Private Cloud (VPC)	275,00 USD	24hrs /día	Amazon Virtual Private Cloud (Amazon VPC) le permite aprovisionar una sección de la nube de AWS aislada de forma lógica, en la que puede lanzar recursos de AWS en una red virtual que usted defina.	
AWS IAM Access Analyzer	120,40USD	2 analizadores/ cuenta	Gestiona el acceso a los recursos de AWS.	
Infraestructura (Hardware)				
Al hacerlo en Amazon web services tú no tienes que preocuparte por el hardware físico (servidores, discos duros, etc.), ya que el proveedor de la nube se encarga de todo eso. Tú simplemente alquilas los recursos de computación que necesitas				
Personal				
Nombre	sueldo	cantidad	descripción	total

Director del departamento	\$216,018	1	Responsable de supervisar las operaciones generales de un departamento, asegurando que estén alineadas estratégicamente con los objetivos de la organización y gestionando un equipo de gerentes y empleados.	\$1,227,783 USD anuales \$102,315 USD mensuales
Project manager	\$136,382	1	Gestiona proyectos específicos planificando, ejecutando y cerrando proyectos, asegurándose de que se completen a tiempo, dentro del alcance y del presupuesto.	
Product manager	\$123,660	1	Supervisa el desarrollo y la gestión de un producto a lo largo de su ciclo de vida, incluyendo la estrategia, la hoja de ruta y la definición de características.	
Arquitecto de soluciones	\$104,920.	1	Diseña y supervisa la implementación de soluciones tecnológicas, asegurándose de que cumplan con las necesidades empresariales y se integren correctamente con otros sistemas.	
Ingeniero en redes	\$95,571	1	Se enfoca en mantener y optimizar las infraestructuras de redes, implementando principios de DevOps como la automatización y la integración continua para optimizar las operaciones.	
Científico de datos	160,617	1	Analiza e interpreta datos complejos para proporcionar información valiosa, utilizando técnicas estadísticas, de aprendizaje automático y otras herramientas computacionales.	
Analista de datos	\$108,981	2	Interpreta los datos para apoyar la toma de decisiones, generalmente enfocándose en identificar patrones, tendencias y correlaciones dentro de los conjuntos de datos.	
Desarrollador fronted	76,566	1	Se especializa en el desarrollo de las partes visuales e interactivas de un sitio web o aplicación web, garantizando una buena experiencia de usuario y rendimiento	
Desarrollador backend	114,205	1	Se enfoca en la lógica del servidor, las bases de datos y las APIs que alimentan la funcionalidad de una aplicación web o software desde el "detrás de escena"	
Full stack developer	90,863	1	Trabaja tanto en el frontend como en el backend de aplicaciones web, siendo capaz de manejar todo el proceso de desarrollo, desde el diseño de la interfaz de usuario hasta la lógica del servidor y	

			la gestión de bases de datos.	
Mantenimiento				
mensual: \$500 USD				
Anual \$6,000 USD				
<b>TOTAL:</b>				
ANUAL			MENSUAL	
\$1.246.948,56 USD			\$103.912,38 USD	

## Preguntas

**¿Cuáles son los costos de las licencias de servicio de los softwares y gestor de base de datos que seleccionaste? ¿Es pago mensual, anual o de una sola compra?**

El costo es mensual

Amazon Simple Storage Service (S3)	48.87 USD
Amazon Athena	74.27 USD
Amazon Aurora MySQL-Compatible	309.84 USD
Amazon QuickSight	268,00 USD
Amazon Virtual Private Cloud (VPC)	275,00 USD
AWS IAM Access Analyzer	120,40USD

**¿Qué tipo de infraestructura vas a requerir para poder llevar a cabo el proyecto? ¿Qué tipo de equipos son mejores para este tipo de procesos y cuántos vas a utilizar? ¿Son necesarios servidores para almacenar la base de datos?**

Al hacerlo en Amazon web services tú no tienes que preocuparte por el hardware físico (servidores, discos duros, etc.), ya que el proveedor de la nube se encarga de todo eso. Tú simplemente alquilas los recursos de computación que necesitas.

Si serán necesarios servidores para almacenar la base de datos ya que así evitamos el riesgo de perder los datos al tener hardware como discos.

**¿Cuánto cobra el personal que decidiste contratar para el desarrollo de este proyecto? Es importante recordar que se debe pagar por cada persona considerada.**

El sueldo es Anual y en USD

Nombre	sueldo
Director del departamento	\$216,018
Project manager	\$136,382
Product manager	\$123,660

Arquitecto de soluciones	\$104,920.
Ingeniero en redes	\$95,571
Científico de datos	160,617
Analista de datos	\$108,981
Desarrollador fronted	76,566
Desarrollador backend	114,205
Full stack developer	90,863

### ¿Cuánto cuesta mantener los equipos? ¿Cuánto cuesta mantener un servidor?

Al pagar las licencias no nos tenemos que preocupar por el costo de mantener el equipo sin embargo el costo de mantener optimo el servidor estimo un aproximado de

Mantenimiento
mensual: \$500 USD
Anual \$6,000 USD

### Con base en los costos de lo anterior, ¿cuánto tiempo va a tardar el desarrollo del proyecto?

Mes	Etapa del Proyecto	Detalles
1 - 2	Planificación	<ul style="list-style-type: none"> <li>- Definición de Objetivos y KPIs: Establecer objetivos claros (por ejemplo, análisis predictivo, clasificación).</li> <li>- Análisis de Requisitos: Determinar necesidades de almacenamiento, procesamiento, y permisos de seguridad.</li> </ul>
3	Estimación de Costos	<ul style="list-style-type: none"> <li>- Crear una estimación detallada de costos para los servicios de AWS necesarios.</li> <li>- Definir el presupuesto para almacenamiento (Amazon S3), procesamiento (Athena, Aurora) y visualización (QuickSight).</li> </ul>
4 - 5	Diseño	<ul style="list-style-type: none"> <li>- Diseño de Arquitectura Técnica: Crear la infraestructura en AWS utilizando Amazon VPC, IAM Access Analyzer para la seguridad, y diseño de bases de datos en Amazon Aurora.</li> <li>- Plan de Desarrollo: Dividir tareas entre los roles asignados: Científico de Datos, Ingenieros y Desarrolladores, detallando plazos y entregables para cada equipo.</li> </ul>
6 - 7	Implementación	<ul style="list-style-type: none"> <li>- Desarrollo e Implementación Técnica: Configuración de almacenamiento en S3, conexión con bases de datos en Aurora, y creación de roles IAM.</li> <li>- Despliegue de Modelos: Entrenar y desplegar modelos de minería de datos usando Amazon Athena.</li> </ul>
8	Pruebas Unitarias y de Integración	<ul style="list-style-type: none"> <li>- Pruebas Unitarias: Verificación de que cada componente funciona como se espera.</li> <li>- Pruebas de Integración: Asegurarse de que todos los servicios y modelos se integren correctamente y puedan comunicarse.</li> </ul>
9	Evaluación	<ul style="list-style-type: none"> <li>- Pruebas Piloto: Implementar una prueba piloto para analizar resultados en condiciones reales.</li> <li>- Evaluación de Resultados y Optimización: Validar precisión de modelos, realizar ajustes necesarios para mejorar rendimiento.</li> </ul>
10	Despliegue Final	<ul style="list-style-type: none"> <li>- Lanzamiento a Producción: Pasar los modelos y servicios a un entorno de producción.</li> <li>- Capacitación del Personal: Entrenamiento para el equipo sobre el uso de AWS y los modelos desplegados.</li> </ul>
11 - 12+	Mantenimiento y Optimización Continua	<ul style="list-style-type: none"> <li>- Monitoreo de KPIs: Revisar continuamente el rendimiento mediante dashboards en QuickSight y KPIs.</li> <li>- Mejoras Incrementales: Realizar ajustes de optimización según los datos recopilados, y asegurar la escalabilidad y precisión del sistema.</li> </ul>

### ¿Cuánto va a costar el desarrollo del proyecto en general?

TOTAL:	
ANUAL	MENSUAL

\$1.246.948,56 USD

\$103.912,38 USD

# Propuesta

## Propuesta de Proyecto: Minería de Datos en AWS

### 1. Resumen de la ejecución

El objetivo de este proyecto es desarrollar e implementar una infraestructura de minería de datos en Amazon Web Services (AWS). Se propone utilizar los servicios de AWS para crear un sistema eficiente y escalable que permita almacenar, analizar y visualizar grandes volúmenes de datos. La solución está diseñada para optimizar el procesamiento de datos, mejorando la toma de decisiones a partir de insights (**Un insight es una idea reveladora o descubrimiento que ayuda a resolver un problema**) obtenidos con técnicas de análisis predictivo y minería de datos. El proyecto se ejecutará en un período de 12 meses y cubrirá todas las fases críticas desde la planificación hasta el mantenimiento continuo.

### 2. Objetivos del Proyecto

- **Implementar una infraestructura en la nube en AWS** para el almacenamiento y análisis de datos a gran escala.
- **Desarrollar y desplegar modelos de minería de datos** que permitan obtener insights predictivos y descriptivos.
- **Optimizar la toma de decisiones** mediante la visualización de datos y la generación de informes interactivos.
- **Monitorear y ajustar el rendimiento del sistema** continuamente para garantizar su eficacia y escalabilidad.

### 3. Alcance del Proyecto

- **Fase de Planificación y Diseño:** Incluye la definición de objetivos, análisis de requisitos y diseño de la arquitectura técnica.
- **Desarrollo y Despliegue:** Configuración y desarrollo de la infraestructura, creación de modelos y despliegue de herramientas de minería de datos.
- **Pruebas y Evaluación:** Pruebas unitarias, de integración y evaluación de resultados.

- **Despliegue Final y Capacitación:** Paso del sistema a producción, formación de personal y monitoreo.
- **Mantenimiento y Optimización Continua:** Monitoreo de KPIs y mejoras incrementales.

## 4. Metodología y Cronograma

Mes	Etapa del Proyecto	Detalles
1 - 2	Planificación	<ul style="list-style-type: none"> <li>- Definición de Objetivos y KPIs: Establecer objetivos claros (por ejemplo, análisis predictivo, clasificación).</li> <li>- Análisis de Requisitos: Determinar necesidades de almacenamiento, procesamiento, y permisos de seguridad.</li> </ul>
3	Estimación de Costos	<ul style="list-style-type: none"> <li>- Crear una estimación detallada de costos para los servicios de AWS necesarios.</li> <li>- Definir el presupuesto para almacenamiento (Amazon S3), procesamiento (Athena, Aurora) y visualización (QuickSight).</li> </ul>
4 - 5	Diseño	<ul style="list-style-type: none"> <li>- Diseño de Arquitectura Técnica: Crear la infraestructura en AWS utilizando Amazon VPC, IAM Access Analyzer para la seguridad, y diseño de bases de datos en Amazon Aurora.</li> <li>- Plan de Desarrollo: Dividir tareas entre los roles asignados: Científico de Datos, Ingenieros y Desarrolladores, detallando plazos y entregables para cada equipo.</li> </ul>
6 - 7	Implementación	<ul style="list-style-type: none"> <li>- Desarrollo e Implementación Técnica: Configuración de almacenamiento en S3, conexión con bases de datos en Aurora, y creación de roles IAM.</li> <li>- Despliegue de Modelos: Entrenar y desplegar modelos de minería de datos usando Amazon Athena.</li> </ul>
8	Pruebas Unitarias y de Integración	<ul style="list-style-type: none"> <li>- Pruebas Unitarias: Verificación de que cada componente funciona como se espera.</li> <li>- Pruebas de Integración: Asegurarse de que todos los servicios y modelos se integren correctamente y puedan comunicarse.</li> </ul>
9	Evaluación	<ul style="list-style-type: none"> <li>- Pruebas Piloto: Implementar una prueba piloto para analizar resultados en condiciones reales.</li> <li>- Evaluación de Resultados y Optimización: Validar precisión de modelos, realizar ajustes necesarios para mejorar rendimiento.</li> </ul>
10	Despliegue Final	<ul style="list-style-type: none"> <li>- Lanzamiento a Producción: Pasar los modelos y servicios a un entorno de producción.</li> <li>- Capacitación del Personal: Entrenamiento para el equipo sobre el uso de AWS y los modelos desplegados.</li> </ul>
11 - 12+	Mantenimiento y Optimización Continua	<ul style="list-style-type: none"> <li>- Monitoreo de KPIs: Revisar continuamente el rendimiento mediante dashboards en QuickSight y KPIs.</li> <li>- Mejoras Incrementales: Realizar ajustes de optimización según los datos recopilados, y asegurar la escalabilidad y precisión del sistema.</li> </ul>

## 5. Recursos y Roles Involucrados

Este proyecto se llevará a cabo con el equipo y roles siguientes:

- **Director del Departamento:** Supervisa el cumplimiento de los objetivos estratégicos.
- **Project Manager:** Coordina el equipo y asegura la entrega en tiempo y forma.
- **Product Manager:** Define los requisitos de producto y objetivos de negocio.
- **Arquitecto de Soluciones:** Diseña la arquitectura técnica y el flujo de trabajo.
- **Ingeniero de Redes:** Configura redes seguras en AWS.
- **Científico de Datos:** Desarrolla modelos de minería de datos.
- **Analistas de Datos:** Procesan y analizan datos para obtener insights.
- **Desarrollador Frontend y Backend:** Implementan interfaces y conexiones backend.
- **Desarrollador Full Stack:** Apoya en la construcción de funcionalidades adicionales.

## 6. Servicios de AWS Utilizados

- **Amazon Simple Storage Service (S3):** Almacenamiento de datos.

- **Amazon Athena:** Procesamiento de datos y consultas SQL.
- **Amazon Aurora (MySQL-Compatible):** Gestión de bases de datos relacionales.
- **Amazon QuickSight:** Visualización y análisis de datos.
- **Amazon Virtual Private Cloud (VPC):** Configuración de redes privadas.
- **AWS IAM Access Analyzer:** Gestión de accesos y seguridad.

## 7. Beneficios del Proyecto

- **Reducción de costos** mediante el uso de almacenamiento y procesamiento en la nube.
- **Escalabilidad y flexibilidad** para adaptarse al crecimiento de datos.
- **Mayor eficiencia en la toma de decisiones** gracias a los insights predictivos y descriptivos obtenidos a partir de los modelos.
- **Optimización continua** para mantener el rendimiento del sistema y mejorar la precisión de los modelos de minería de datos.

## 8. Conclusión

Este proyecto en AWS no solo permitirá una gestión avanzada de datos, sino que también fortalecerá las capacidades de análisis de la organización, aumentando la competitividad y promoviendo decisiones informadas. Con un enfoque en la optimización continua y la escalabilidad, la solución propuesta es sostenible y adaptable a largo plazo.

# Conclusión

A través de esta actividad, pude comprender de manera más profunda la importancia de la minería de datos en el ámbito empresarial, especialmente en el área de Inteligencia de Negocios. El cómo aprender la recolección, análisis y uso adecuado de grandes volúmenes de datos puede influir directamente en la toma de decisiones estratégicas, permitiendo que las empresas sean más eficientes y competitivas. Además, el proceso de estimar los costos y recursos necesarios para implementar un proyecto de esta envergadura me ha enseñado a valorar la planificación detallada y la gestión de los recursos tecnológicos y humanos. Entender cómo seleccionar el software adecuado, los equipos necesarios y el personal calificado, junto con la importancia del mantenimiento a largo plazo, ha sido fundamental para visualizar cómo un proyecto de minería de datos puede aportar valor real a una organización. Esta actividad me ha permitido aplicar conceptos teóricos a un escenario práctico, lo que sin duda enriquecerá mi aprendizaje y me preparará mejor para futuros desafíos en el futuro.

# Referencias

1. *LinkedIn Login, Sign in | LinkedIn.* (s. f.). LinkedIn. <https://www.linkedin.com/feed/>
2. *Precio de los servicios de la nube | AWS.* (s. f.). Amazon Web Services, Inc. [https://aws.amazon.com/es/pricing/?aws-products-pricing.sort-by=item.additionalFields.productNameLowercase&aws-products-pricing.sort-order=asc&awsf.Free%20Tier%20Type=\\*all&awsf.tech-category=\\*all](https://aws.amazon.com/es/pricing/?aws-products-pricing.sort-by=item.additionalFields.productNameLowercase&aws-products-pricing.sort-order=asc&awsf.Free%20Tier%20Type=*all&awsf.tech-category=*all)
3. Material de apoyo
4. Chat gpt
5. Gemini
6. Preguntas a Google ejemplo:

The screenshot shows a Google search results page. The search query "sueldo usd de desarrollador full stack" is entered in the search bar. Below the search bar, there are tabs for "Todo", "Imágenes", "Videos", "Noticias", "Shopping", "Web", "Libros", "Más", and "Herramientas". A blue diamond icon indicates that the results are "Visión general creada por IA". There is also a "Escuchar" button with a speaker icon.

The main content area displays a summary statement: "En Estados Unidos, el sueldo promedio de un desarrollador full stack es de USD 100,000 al año, y el pago total estimado es de USD 122,717 al año." Below this, a table titled "Sueldo" provides specific salary data:

Desarrollador full stack	USD 100,000 al año
Desarrollador full stack junior	USD 80,000 al año

Further down, a section titled "En cuanto a las industrias que pagan mejor a los desarrolladores full stack, las siguientes son las mejores:" lists several industries with their respective salaries:

- Energy, Mining & Utilities: USD 137,845
- Financial Services: USD 129,193
- Insurance: USD 124,704
- Healthcare: USD 122,283
- Aerospace & Defense: USD 121,441

At the bottom, a note states: "Las empresas que más pagan en Aerospace & Defense para un empleo de Full Stack Developer son Boeing, Maxar Technologies y Lockheed Martin."