

Nombre: Jesús Alejandro Valencia Valadez		Matrícula: 3013480
<ul style="list-style-type: none">Nombre del curso: Infraestructura para Big Data		Nombre del profesor: Miguel de Jesús Martínez Felipe
Módulo: 1		Actividad: Avance 1 de Evidencia 1

Lee con atención plena el siguiente caso:

La empresa *Bratt & Houston* pertenece al giro aeroespacial en los Estados Unidos, teniendo operaciones a lo largo de los cinco continentes, siendo un referente a nivel mundial en el diseño, servicio y fabricación de motores aéreos. Su sede principal se localiza en Montgomery, Alabama (EE. UU). La empresa cuenta con más de 20,000 colaboradores para dar servicio a más 5,000 clientes localizados en 100 países.

Vanessa, quien es egresada, desde hace un año, de la carrera de Ingeniería en Computación Administrativa por parte de la Universidad Tecmilenio, no tuvo problemas para colocarse profesionalmente en la empresa *Innova Analytics*, dedicada al desarrollo de soluciones basadas en datos. Gracias a su desempeño y el desarrollo de sus competencias profesionales, ha logrado una plaza en la empresa *Bratt & Houston*, para trabajar como analista de datos en su sede central.

Uno de sus primeros proyectos asignados, consiste en implementar un ecosistema Hadoop utilizando la nube de Microsoft Azure, relacionado con el servicio de verificación del estado de las turbinas aéreas fabricadas en su sede central, mediante este servicio se logran recopilar una gran cantidad de datos de diferentes fuentes, con un énfasis especial en las rutas aéreas comerciales de las aeronaves que utilizan sus turbinas, además de conocer si estas rutas son cumplidas en tiempo y forma.

La importancia del proyecto radica, que, al día de hoy, miles de empresas están migrando sus infraestructuras a la nube, lo cual les permite, entre otras cosas, disponer de una mayor flexibilidad, agilizar sus procesos y reducir costos. Aunado a esto, la masiva generación de datos y su procesamiento vuelve primordial adquirir conocimientos en el uso, configuración y mantenimiento de estas herramientas.

Como resultado, se espera hacer una presentación con la implementación de la infraestructura para el procesamiento de datos masivos en la nube.

Para llevar a cabo este proyecto, imagina que tomarás el rol de Vanessa y deberás realizar cada una de las siguientes indicaciones, toma en cuenta que esta evidencia la entregarás en tres partes diferentes (avances).

Avance 1

1. En este primer avance, como paso previo al desarrollo del proyecto, se realizará una reflexión de arranque (con una extensión de al menos dos cuartillas), en la cual se estará dando respuesta a cada una de las siguientes preguntas:
 - ¿Cuál es la importancia de la infraestructura en proyectos de Big Data?

La infraestructura es fundamental en proyectos de Big Data debido a su papel crucial en el almacenamiento, procesamiento y análisis de grandes volúmenes de datos. Algunos puntos clave son:

- **Escalabilidad:** Una infraestructura bien diseñada permite manejar crecimiento exponencial de datos sin afectar el rendimiento.
 - **Disponibilidad:** Garantiza que los datos estén disponibles cuando se necesiten, evitando tiempos de inactividad.
 - **Rendimiento:** Una infraestructura optimizada acelera el procesamiento y análisis de datos.
 - **Seguridad:** Protege los datos sensibles y cumple con regulaciones.
- ¿Cuáles son las claves para construir una infraestructura adecuada, sólida y efectiva?
 - **Planificación:** Comprender los requisitos del proyecto y diseñar una infraestructura escalable y flexible.
 - **Selección de Tecnologías:** Elegir herramientas y plataformas que se adapten a las necesidades específicas de Big Data.
 - **Arquitectura:** Diseñar una arquitectura robusta que incluya almacenamiento, procesamiento y visualización.
 - **Monitoreo y Mantenimiento:** Establecer prácticas para supervisar y mantener la infraestructura.

- ¿Cuáles son las mejores prácticas en la creación de una infraestructura de Big Data?
 - **Cloud First:** Considerar soluciones en la nube para escalabilidad y flexibilidad.
 - **Contenedores:** Utilizar contenedores (como Docker) para facilitar la implementación y gestión.
 - **Automatización:** Automatizar tareas de aprovisionamiento, configuración y escalado.
 - **Seguridad:** Implementar medidas de seguridad como cifrado y autenticación.

- ¿Qué habilidades iniciales son las necesarias para instalar, configurar y poner en marcha una infraestructura hardware y software para Big Data?
 - **Conocimientos en Cloud:** Comprender servicios en la nube como Azure, AWS o Google Cloud.
 - **Administración de Sistemas:** Saber instalar y configurar servidores, redes y sistemas operativos.
 - **Bases de Datos y Almacenamiento:** Conocer bases de datos NoSQL, Hadoop, almacenamiento distribuido.
 - **Programación y Scripting:** Lenguajes como Python, Java o Scala.

- ¿Qué puntos débiles existen al momento de construir una infraestructura, además de propuestas para lograr resolverlos?
 - **Costos:** Las soluciones en la nube pueden ser costosas. Se debe optimizar el uso de recursos.
 - **Escalabilidad Compleja:** Planificar para escalabilidad desde el inicio.
 - **Seguridad:** Implementar medidas de seguridad sólidas y mantenerse actualizado.
 - **Falta de Experiencia:** Capacitación continua y colaboración con expertos.

2. Haz entrega de los archivos electrónicos con las actividades realizadas de los temas 1, 2, 3, 4 y 5.

Las Actividades se han subido en su respectivo espacio.