

2. Estudie la siguiente URL:

<https://sagemaker-examples.readthedocs.io/en/latest/intro.html>

a. ¿Cuáles son las diferentes opciones que tiene para usar SageMaker?

Las diferentes opciones que se tienen para usar SageMaker son la “IDE: SageMaker Studio”, “Console: SageMaker Notebook Instances”, “Command line & SDK: AWS CLI, boto3, & SageMaker Python SDK” y “3rd Party integrations: Kubeflow & Kubernetes operators”.

b. Haga un resumen de los vídeos que se muestran en esta página web.

- Video 1 (Onboard Quickly to Amazon SageMaker Studio):

El primer video es en esencia la guía para crear un dominio, un usuario y un cuaderno dentro de AWS SageMaker. La primera parte se enfoca en definir que es “onboarding”, que se entiende como el proceso de crear un dominio asociado a una cuenta personal en una región determinada. La creación de un dominio conlleva muchas ventajas, tales como el acceso a la creación de cuaderno y la capacidad de compartirlos con otros usuarios de una forma muy sencilla, permitiendo un mejor y más fácil ambiente de trabajo sin dejar atrás la seguridad de los datos e información a tratar. Por otro lado, al crear un dominio, le permite al usuario tener “amazon elastic file systems (efs)”, en donde cada usuario podrá tener su directorio, garantizando un ambiente de desarrollo (IDE) privado.

La siguiente parte del video se centra en el paso a paso de como crear el dominio. Esta sección empieza mostrando la existencia de dos opciones para crear un dominio, la primera siendo una forma automática y rápida, mientras que la segunda siendo una alternativa estándar (standar set up), la cual permite tener todavía más control en lo concerniente a permisos, configuración para compartir cuadernos y formas de encriptación de la información. Por tal motivo, el video propone realizar la configuración estándar. Inmediatamente seleccionada esta opción se despliega un menú de configuraciones de por ejemplo el método de autorización que depende del tipo de cuenta que se tenga, los permisos a los buckets de AWS a los que el usuario podrá acceder, la opción de compartir o no los cuadernos, la forma de encriptación de los cuadernos y la forma de almacenamiento si será en la nube o no (private subnet si no tiene acceso a internet y public subnet si lo tiene). Cuando se haya terminado la configuración basta con oprimir el botón de submit y continuar con la guía para la obtención de usuario.

En este punto, la página deberá estar situada en el panel frontal, en donde no se podrá avanzar mucho más si no se tiene un usuario ya creado con un rol estipulado dentro del espacio de trabajo. Lo único posible en este punto por parte de la empresa es la posibilidad de eliminar usuarios y quitarles los permisos a los cuadernos que tenga acceso.

Una vez se haya creado el usuario, el video toma como objetivo el desglosar la creación de cuadernos y contextualización con el ambiente de desarrollo. Se propone como primer paso la elección del kernel/núcleo asociado, que por fines prácticos se elige comúnmente a Python 3. Luego de esto se explican algunas opciones de la barra de herramientas, tales como la posibilidad de clonar y/o anclar repositorios de git al directorio privado del usuario. Finalmente, el video acaba explicando algunas de las opciones que puede ofrecer el studio. Así como cualquier studio, en este se pueden crear carpetas y visualizar los datos o archivos a los que se tiene acceso, se pueden subir archivos, se puede acceder a la sesión de administrador que permite ver los kernels, terminales e imágenes que se estén corriendo actualmente, se tiene la paleta de funciones, las herramientas de cuaderno y se tiene la opción de abrir una terminal para ejecutar muchas de las funciones del studio partiendo de comandos como en Linux (file, new y terminal).

El video finaliza en este punto, dejando como recado la conclusión de que, si bien parece que el proceso de creación de dominio es difícil, su consecución termina siendo bastante satisfactorio para el orden y funcionamiento del usuario, sin mencionar que tal proceso solo debe ser llevado a cabo en una ocasión.

- **Video 2 (Quickly Launch Amazon SageMaker Notebooks):**

La primera parte de este video se centra en la configuración de los cuadernos de Jupyter, como la capacidad de instalar paquetes, correr scripts y configurar la seguridad de lo que se conoce como lifecycles configurations. Cuando se realizan configuraciones como estas se permite la personalización de los cuadernos, procesos que pueden ser supremamente importantes si se desea por ejemplo correr scripts siempre que se inicie un cuaderno o únicamente al punto de crearlos. Tal ejercicio se ve visualizado a la hora de enlazar los cuadernos con Github, pues de esta manera se tiene control de las versiones previas del código.

En este punto entramos a la creación de los cuadernos. La configuración de este es el inicio de este segmento, pues define el precio de este y la capacidad computacional que tendrá, aclarando que en cualquier instante se pueden modificar cualquiera de estos dos atributos, siempre y cuando se mantenga dentro del rango permitido por AWS (5GB – 16TB). Luego de establecer estas propiedades, el video apunta a algo llamado IAM role, que es donde la persona debe elegir un usuario capaz de acceder a los múltiples servicios de AWS SageMaker, como bien pueden ser los S3 Buckets.

Este video, al igual que el primero, deja un apartado para hablar de la privatización de los cuadernos, pues se explica la posibilidad de aislarlos de la nube y del servidor en general (VPC). A su vez, mencionan la opción de etiquetas, que permite nombrar los recursos disponibles de forma más organizada.

Por último, el video deja como advertencia los tiempos de ejecución de un cuaderno, pues recalca que se debe estipular una condición de parada si se desea evitar costos de

ejecución en la nube. Esto es posible desde el menú principal, donde se encuentra la opción de Stop o la opción de borrar en definitivo el cuaderno.

3. Lea el siguiente artículo:

<https://earthclipse.com/environment/pollution/wastewater-types-sources-effects-solutions.html>

Haga un resumen sobre las características de las diferentes aguas residuales.

El tema principal de este artículo es sobre todo lo que rodea a las aguas residuales, las preocupaciones que están generando y su caracterización.

Se define el agua residual como los desechos que se acumulan dentro de los líquidos residuales que fluyen alrededor de las zonas urbanas y rurales, sin discriminar si su procedencia y/o tratamiento se da en propiedades individuales de cada ciudadano o si se hace en las plantas de tratamiento específico. El riesgo de no tratar estas aguas residuales es altísimo, porque por lo general estas pueden terminar por transmitir virus, bacterias, etc. A su vez, también pueden transportar residuos supremamente nocivos, tales como contaminantes biológicos, físicos, químicos, materiales pesados, macro sólidos, toxinas, etc.

Por lo general, el tratamiento de estas aguas va en función de su procedencia y contenido. Por tal motivo, se hace necesario la discriminación en sectores de esta.

- **Industriales:** La diferencia con el agua residual doméstica recae esencialmente en el contenido alto de sustancias nocivas dentro de las industriales, causando que su tratamiento sea más engorroso y generando que algunas entidades no realizan tal procedimiento de la forma más idónea. Algunos de los contenidos dentro de las aguas industriales son los medicamentos, químicos tóxicos, arena, productos químicos, arena y demás. Usualmente su origen recae en las fábricas o en los grandes comercios.
- **Proveniente de hogares (domésticas):** Las aguas clasificadas dentro de este sector pueden ser las que su fuente son las descargas líquidas de uso regular. Como por ejemplo de la cocina, el lavaplatos, hospitales, algunos comercios y pequeñas industrias. Si bien estas aguas contienen altas impurezas, por lo general tienden a ser tratadas y purificadas correctamente.

Discriminar las aguas residuales por procedencia no es suficiente, pues a partir de su contenido también es posible encontrar todo tipo de sustancias y elementos. Dicho eso, se presenta la categorización a partir de sus sustancias.

- **Aguas amarillas:** Este tipo de agua residual se refiere esencialmente a los ductos por donde circula la orina.

- **Aguas Negras:** Cuando se trata de aguas negras, se refieren a las aguas altamente contaminadas con desechos humanos, tales como las provenientes de los baños, cocina y lava platos. Tal combinación puede llevar consigo bacterias con un tiempo de vida tan largo que tienden a durar incluso después de el agua ser tratada y puesta en circulación nuevamente.
- **Aguas grises:** Se entiende como el agua que circula por el sistema de alcantarillado pero que no lleva consigo materia fecal ni orina, permitiendo un tratamiento de purificación mucho más fácil. Eso sí, aclarando que de no tratarse bien no es posible que circulen en otra ocasión pues pueden causar enfermedades de ser ingeridas.