

**NOMBRE COMPLETO:** Ing. Hilaría Adima Vásquez Durán  
**NAO ID:** 3033  
**FECHA:** 20 de septiembre de 2024  
**NOMBRE DE LA TRAYECTORIA**  
**EN LA QUE ESTÁS ENROLADO:** DATA ANALYST CORE  
**Título del Reto:**  
**Manejo eficiente de datos al programar aplicaciones distribuidas**

**BACKLOG´S**

Backlog		
-lista de requerimientos-		
Backlog		
HISTORIA DE USUARIO	REQUERIMIENTOS	ESTIMACIÓN DE TIEMPO
Averiguar framework Hadoop	Buscar como se desarrolla dentro de Framework Hadoop	3 DIAS
Conocer el entorno	Crear un programa que construya una tabla con el cambio porcentual del crecimiento de ventas de cada país a nivel mensual (con respecto al mes anterior).	3 DIAS
Desarrollo de datos de del mes de marzo de 2024	Diseñar un programa que construya la agrupación de datos de marzo de 2024	3 DIAS
Desarrolla un script en Bash ( <b>extract_hdfs.sh</b> )	<ul style="list-style-type: none"><li>• Cree en el HDFS a la ruta <b>/data/</b>,</li><li>• Coloque los datos de vuelos correspondientes a <b>Marzo de 2024</b> en la ruta anterior,</li><li>• Renombre los datos anteriores en el mismo directorio de hdfs pero con la denominación "2024_3.csv"</li></ul>	3 DIAS
Paradigma de programación que permite una escalabilidad masiva en cientos o miles de servidores en un clúster de Hadoop.	<ul style="list-style-type: none"><li>• Permita visualizar los primeros 10 renglones del archivo,</li><li>• Liste el contenido de todos los directorios de cluster.</li><li>• El resultado de la ejecución a consola deberá ser salvada en un archivo de texto (<b>extract_hdfs.txt</b>) y adjuntar imágenes de la consola en cada paso (usar el formato <b>extract_hdfs_1.png, extract_hdfs_2.png</b>..., etcétera).</li></ul>	3 DIAS

