

**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-**

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 (extract_hdfs.sh)	<ol style="list-style-type: none"><li>1. Cree en el HDFS a la ruta /data/,</li><li>2. Coloque los datos de vuelos correspondientes a <b>Marzo de 2024</b> en la ruta anterior,</li><li>3. Renombre los datos anteriores en el mismo directorio de hdfs pero con la denominación "2024_3.csv"</li></ol>	3 DIAS
Paradigma de programación que permite una escalabilidad masiva en cientos o miles de servidores en un clúster de Hadoop.	<ol style="list-style-type: none"><li>1. Permita visualizar los primeros 10 renglones del archivo,</li><li>2. Liste el contenido de todos los directorios de cluster.</li><li>3. El resultado de la ejecución a consola deberá ser salvada en un archivo de texto (extract_hdfs.txt) y adjuntar imágenes de la consola en cada paso (usar el formato <b>extract_hdfs_1.png, extract_hdfs_2.png,....</b>, etcétera).</li></ol>	3 DIAS

Crea una serie de programas para procesar y transformar los datos de vuelos en Estados Unidos usando el framework de Apache Hive en un sistema de datos distribuido basado en Hadoop.	Reglas de validación Valor máximo, mínimo, promedio y desviación estándar por columna,	3 DIAS
La migración es un proceso prioritario que debe gestionarse eficientemente y, para asegurarse de finalizarlo y lograr la operatividad necesaria de la base de datos, ha decidido sumar a Julia al proyecto	Valor máximo y mínimo Los 5 categorías más frecuentes de la columna Para poder los vuelos mas frecuentes que se tiene en la Aerolínea.	3 DIAS