

Tarea 1. Big Data

Manual de ejecución del programa

Gabriel Valentine Fonseca

27/11/2023

El archivo comprimido consta de:

1. Archivos necesarios para el correr el programa. Dentro de este podemos encontrar:
 - `build_image.sh`: Este script se encarga de compilar la imagen del contenedor Docker con la configuración necesaria
 - `run_image.sh`: Proporciona las instrucciones para iniciar el contenedor utilizando la imagen previamente construida.
 - `run_program.sh`: Texto para ejecutar el programa dentro del contenedor.
 - `run_test.sh`: Instrucciones para realizar las pruebas unitarias para verificar el correcto funcionamiento del programa
2. Tres archivos CSV: `ciclista.csv`, `ruta.csv` y `actividad.csv`. Los archivos no contienen encabezados y tienen las columnas en el orden mencionado en las instrucciones. Corresponden a los datos para el programa.
3. `main.py`: punto de entrada principal para ejecutar el programa.
4. `functions.py`: contiene funciones utilitarias y operaciones relacionadas con el manejo de datos y consultas en Spark según las instrucciones de la tarea.
5. `controller.py`: coordina la ejecución del programa, utiliza funciones del módulo `functions` para realizar operaciones específicas en datos y se encarga de interactuar con Spark para procesar y analizar datos.

6. `test_unit_tests.py`: contiene pruebas unitarias para las funciones clave en el archivo `functions.py`. Estas pruebas aseguran que las funciones realicen las operaciones previstas y devuelvan los resultados esperados.
7. `conftest.py`: entorno de prueba de `pytest` se utiliza para definir fixtures que pueden ser compartidas entre múltiples archivos de prueba.

Procedimiento para ejecutar el Programa:

1. Construir la Imagen del Contenedor:

Ejecutar el script `build_image.sh` para compilar la imagen del contenedor Docker.

2. Iniciar el Contenedor:

Ejecutar el script `run_image.sh` para iniciar el contenedor con la imagen construida.

3. Ejecutar el Programa:

Acceder al contenedor.

Ejecutar el script `run_program.sh` para ejecutar el programa dentro del contenedor.

Procedimiento para ejecutar Pruebas Unitarias:

1. Construir la Imagen del Contenedor:

Ejecutar el script `build_image.sh` para compilar la imagen del contenedor Docker.

2. Iniciar el Contenedor:

Ejecutar el script `run_image.sh` para iniciar el contenedor con la imagen construida.

3. Ejecutar Pruebas Unitarias:

Ejecutar el script `run_test.sh` para realizar pruebas unitarias y verificar el correcto funcionamiento del programa.

Nota: Asegurarse de seguir el orden correcto de los pasos y verificar que los scripts estén configurados correctamente.