

## Ejercicio 2: Sensor de Archivos en Supermercado

(Video relacionado: 11.2 Sensores. Contenido: Sensors)

Te has unido como Data Engineer a una empresa famosa de supermercados. Recientemente la compañía ha implementado nueva publicidad y quieren analizar si tiene efecto en las ventas de una compañía en específico (compania1) es por esto que te piden que generes un proceso que permita saber todos los días cuando esta compañía ha subido el resumen de sus ventas aplicando esta publicidad. Para esto deberás hacer uso del concepto de sensores y de Airflow. La data se debe subir a una carpeta dentro del cluster donde estas trabajando en la ruta `data/compania1/_SUCCESS/` y en `data/compania1/` si ambas condiciones se cumplen y los archivos empiezan por el patrón `data-*.csv` se debería tener un mecanismo que informe que los datos llegaron (puede ser un simple print)

### Solución

1. Creamos una carpeta que se llama `Sensor_Archivos`
2. Dentro de esta carpeta creamos 4 carpetas: `dags`, `logs`, `plugins` y `data`
3. En el terminal ejecutamos el siguiente código para obtener el `docker-compose` actualizado de Airflow  
`Invoke-WebRequest -Uri 'https://airflow.apache.org/docs/apache-airflow/2.6.3/docker-compose.yaml' -OutFile 'docker-compose.yaml'`
4. Asegurarse que la sección de `volumes` dentro del `docker-compose` sea como este:

```
volumes:
  - ./dags:/opt/airflow/dags
  - ./logs:/opt/airflow/logs
  - ./plugins:/opt/airflow/plugins
  - ./data:/opt/airflow/data
  - ./config:/opt/airflow
```

Y en esta sección false para evitar que se carguen ejemplos innecesarios:

```
AIRFLOW__CORE__LOAD_EXAMPLES: 'false'
```

La versión final debería ser como [esta](#)

5. Dentro de la carpeta `dags`, creamos un archivo llamado `dag_sensores.py` y lo llenamos con el siguiente código
6. Ejecutamos en el terminal `docker-compose up`
7. Vamos a `localhost:8080` y nos logueamos en el UI de airflow

8. Buscamos nuestro dag y lo ejecutamos
9. Ahora si colocamos dentro de la carpeta data/compania1/\_SUCCESS/ el archivo data-compania1.csv el cual lo puedes obtener de [aquí](#)
10. Cuando ejecutemos el Dag debería aparecer el proceso que estabamos buscando

La carpeta con todos los archivos necesarios la encuentras [aquí](#)