



Agencia Estatal de Meteorología

AEMET DATALAKE TO DATAMART

Desarrollo de Aplicaciones para la Ciencia
de Datos.

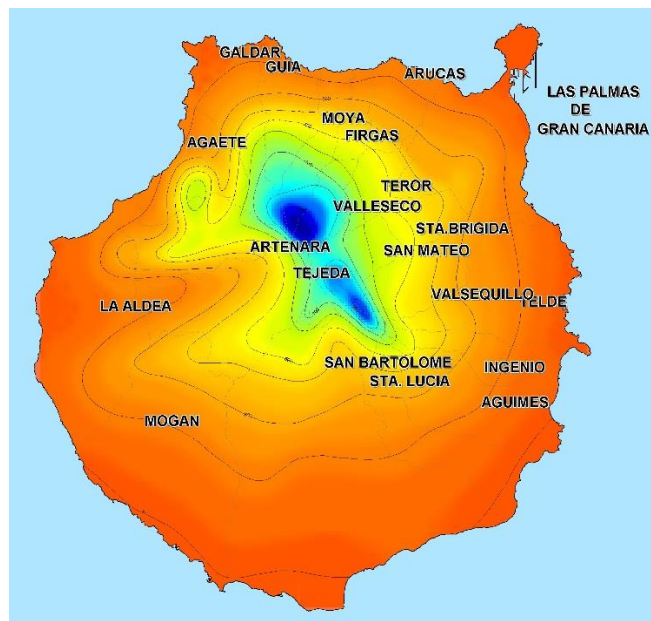
- **CURSO:** Segundo.
- **TITULACIÓN:** Grado en Ciencia e Ingeniería de Datos.
- **ESCUELA:** Escuela de Ingeniería Informática.
- **UNIVERSIDAD:** Universidad de Las Palmas de Gran Canaria.

RESUMEN

En este proyecto realizado en Java, he realizado un datalake donde guardamos eventos y un datamart para realizar consultas. Estos eventos que almacenamos en el datalake son datos sobre el clima de diferentes zonas de Gran Canaria sacados mediante llamadas a la API de Aemet. Usé un Timer Task para que estas llamadas se realizan cada hora. Para recoger únicamente datos de nuestra isla, he tenido que filtrar los datos que nos aporta la API mediante longitud y latitud. Además, también he filtrado la información para quedarnos solo con las variables que nos interesan. Esta información se almacena en el datalake separada por días.

Luego, usando SQL se crea el datamart donde almacenaremos información que nos interesa del datalake. Este datamart se basa en dos tablas, una para las temperaturas máximas de cada día y otra para las temperaturas mínimas. Estas tablas muestran la fecha, la hora, la estación que hizo el registro, el lugar y la temperatura máxima o mínima de cada día. Estos almacenamientos de la información al datamart se realizarán una vez al día mediante el uso de un Timer Task.

Finalmente creé otro programa que ofrece una API Rest para realizar consultas al datamart creado anteriormente.



ÍNDICE

1. Recursos utilizados. -----	3
2. Diseño. -----	4
3. Conclusiones. -----	5
4. Líneas futuras. -----	5
5. Bibliografía. -----	5

RECURSOS UTILIZADOS

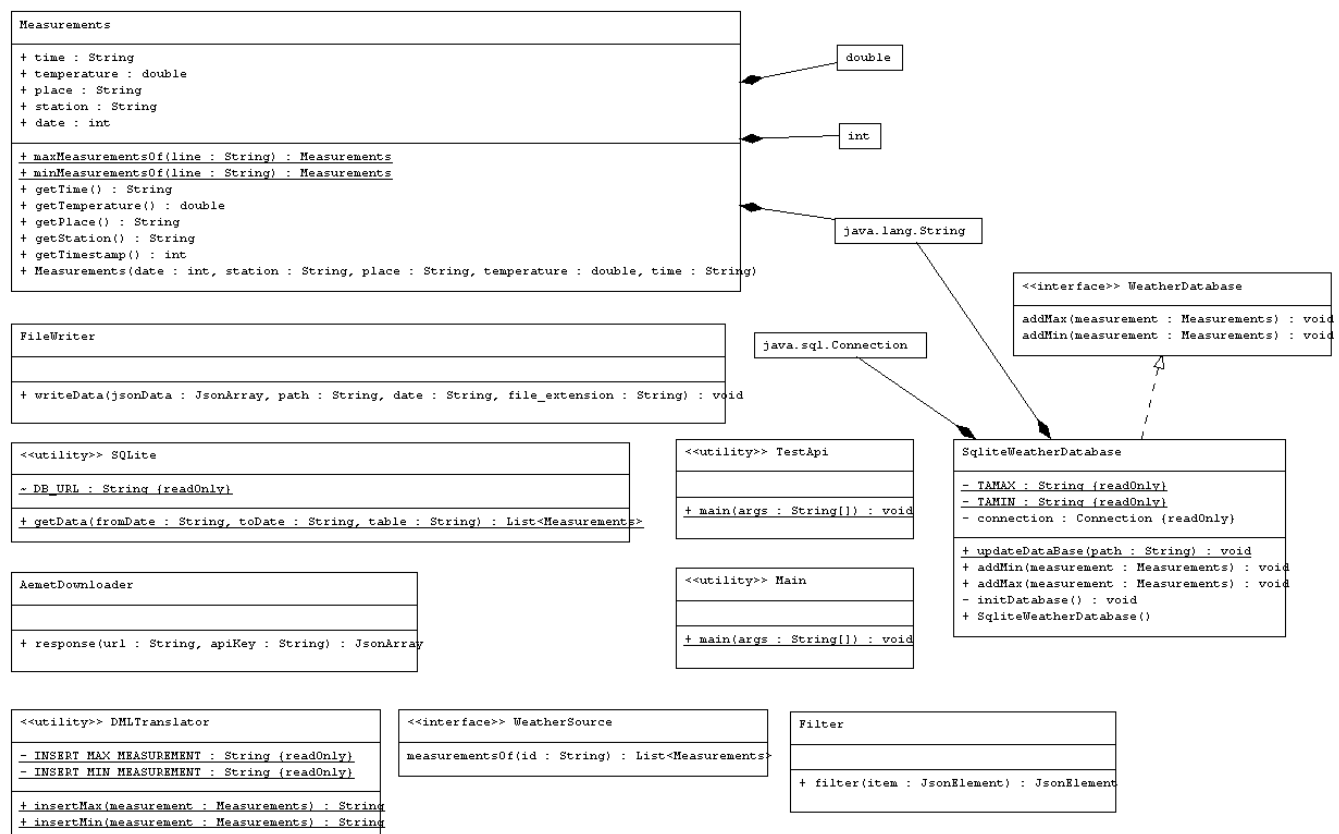
Para este proyecto he utilizado “IntelliJ IDEA Community Edition 2022.2.2” como entorno de desarrollo. Para realizar esta memoria, he utilizado el procesador de texto “Microsoft Word”. Además usé “Git”, que es un software de control de versiones diseñado por Linus Torvalds, pensando en la eficiencia, la confiabilidad y compatibilidad del mantenimiento de versiones de aplicaciones cuando estas tienen un gran número de archivos de código fuente. También el “GitHub” donde he creado un repositorio donde se encuentra este proyecto.



DISEÑO

Este proyecto se basa en tres módulos distintos: “feeder”, “datamart-builder” y “temperatura-api”. Cada uno de ellos tiene una funcionalidad. El módulo “feeder” se encarga de sacar datos mediante llamadas a la API de Aemet, y de guardar los eventos en el datalake. Por otro lado, “datamart-builder” crea el datamart con los datos de interés a partir del datalake. Finalmente, “temperatura-api” ofrece una API Rest para realizar consultas al datamart de los sitios con máxima y mínima temperatura.

Respecto al diseño, he seguido los principios SOLID como el principio de responsabilidad única, es decir, las clases tienen una sola responsabilidad. Además del principio de inversión de dependencias, que se basa en que los módulos de alto nivel no dependan de módulos de bajo nivel.



CONCLUSIONES

En definitiva, este proyecto me ha resultado muy interesante ya que hemos puesto en práctica muchas de las cosas que hemos ido dando durante el curso y además el tema del proyecto es muy curioso y cercano a la realidad. Además, me he sentido más cómodo con el entorno de desarrollo que en el primer proyecto. A pesar del trabajo que conlleva este proyecto, en general me ha gustado y la satisfacción de conseguir un buen resultado es enorme.

LÍNEAS FUTURAS

Para convertirlo en un producto comercializable, habría que automatizarlo todo para que el programa este continuamente guardando datos de las temperaturas máximas y mínimas de los lugares. También habría que expandirlo a más sitios para tener un público mayor, por lo que tendríamos que incluir datos de toda España. Finalmente habría que realizar una interfaz intuitiva y sencilla para que todos los usuarios puedan acceder a las temperaturas de cualquier lugar del país de una manera fácil.

BIBLIOGRAFÍA

<https://www.sqlitetutorial.net/>

<https://stackoverflow.com/>

<https://www.baeldung.com/>

<https://opendata.aemet.es/>

ALEJANDRO BOLAÑOS GARCÍA

12 de enero de 2023