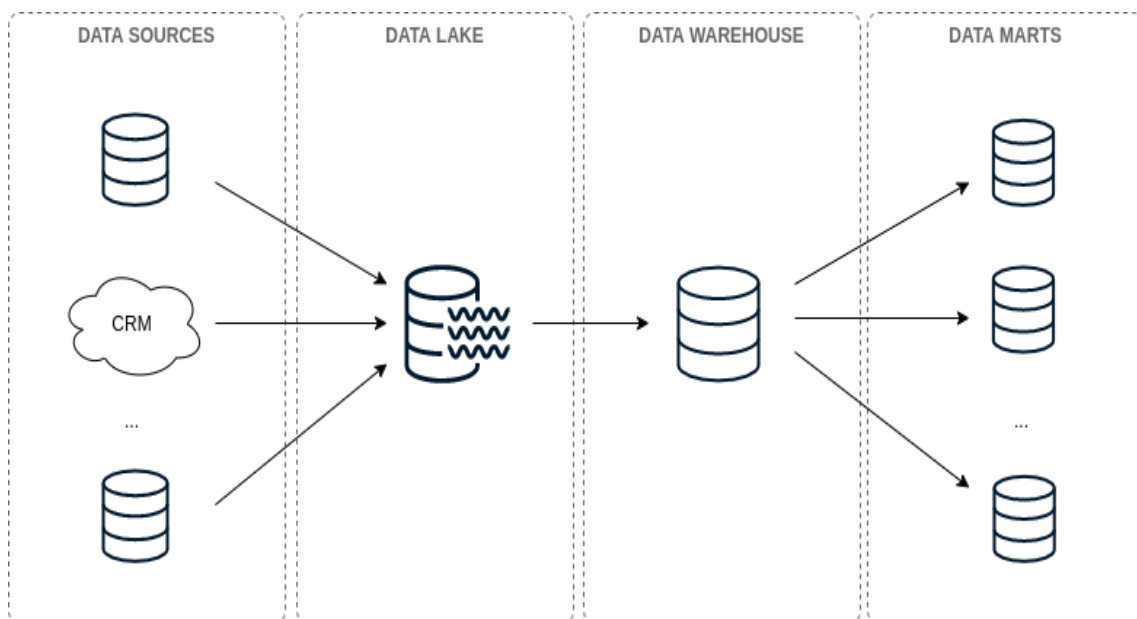


DATALAKE TO DATAMART



DESARROLLO DE APLICACIONES PARA LA CIENCIA DE DATOS
EDIFICIO DE INGENIERÍA INFORMÁTICA Y MATEMÁTICAS
UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
CIENCIA E INGENIERÍA DE DATOS
SEGUNDO CURSO

ADONAI OJEDA MARTÍN

14/01/2023

VERSIÓN I

ÍNDICE

1. Resumen.....	4
2. Recursos Utilizados.....	5
3. Diseño.....	6
4. Conclusiones.....	7
5. Líneas Futuras.....	8
6. Bibliografía.....	9

RESUMEN

En primer lugar, la AEMET, agencia estatal de meteorología, es un organismo público español cuyo objetivo básico es la prestación de servicios meteorológicos que sean del Estado.

En segundo lugar, esta práctica se basa en la descarga de datos de todas las estaciones meteorológicas de la web service de AEMET y quedarnos con los de Gran Canaria para posteriormente trabajar con ellos.

A continuación, se crea una base de datos SQLite donde se crean dos tablas una para los valores máximos y otra para los valores mínimos, donde insertamos la fecha, la estación, el lugar y su temperatura correspondiente según fuera máxima o mínima.

Finalmente, se procedió a crear una api rest donde podríamos ver los datos almacenados a través de distintos get.

RECURSOS UTILIZADOS

La práctica ha sido realizada en el entorno de desarrollo IntelliJ IDEA CE, con la versión java 11 jdk.

Asimismo, para su desarrollo y obtener todas las versiones anteriores y posteriores del proyecto se ha subido a un repositorio de GitHub, siendo esta la primera versión, el scrapping a hoteles de Las Palmas de Gran Canaria de Booking.

Finalmente, Java Doc 11 ha sido la herramienta de documentación usada.

DISEÑO

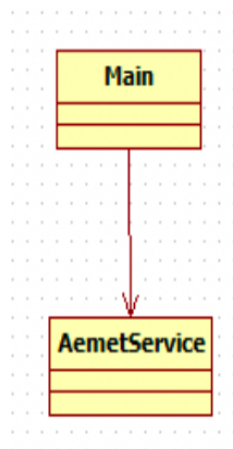
En primer lugar, la práctica está dividida en tres módulos el “feeder”, donde hacemos la conexión con el api de AEMET y recogemos todos los datos necesarios de Gran Canaria para esta práctica.

En segundo lugar, almacenamos todos estos datos en un directorio llamado “datalakeDir” donde registramos los eventos de cada día.

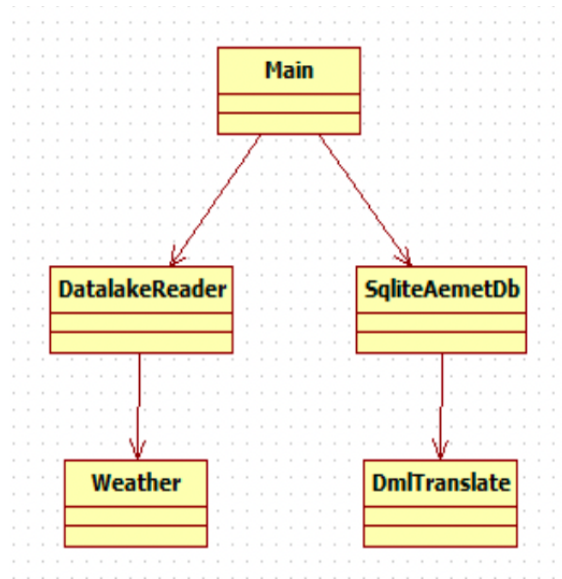
En tercer lugar, el módulo “datamart” donde se crea la base de datos con sus tablas correspondientes, almacenando los datos leídos del “datalakeDir” e insertándolos en su tabla correspondiente.

Finalmente, representamos estos datos en una api rest con dos get uno para la temperatura máxima y otra para la mínima filtrando por fechas.

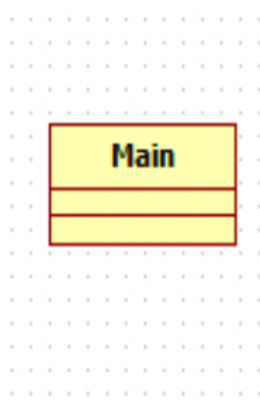
FEEDER:



DATAMART:



API:



CONCLUSIONES

En conclusión, esta práctica nos ha servido para recordar como recoger datos de una API, repasar la dependencia gson y todo lo que eso conlleva, poder poner en práctica como leer ficheros de algún directorio, repasar como crear conexiones, tablas e insertar datos en SQL, y finalmente volver a poder hacer una api la cuál hemos tenido que aprender como pasar parámetros por la query.

LÍNEAS FUTURAS

Nuestra práctica podría ser de gran utilidad a la hora de poder pronosticar temperaturas a corto plazo viendo las diferencias entre los últimos días, aunque esto no es del todo fiable.

Asimismo, podríamos observar los cambios de temperatura entre la misma época del año pero de distintos años y ver si ha cambiado algo o no, o si ha afectado al cambio climático o no, y con esto sacar conclusiones de si se debe este cambio a la vida humana o no, y poder prevenir acontecimientos.

BIBLIOGRAFÍA

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

<https://www.sqlitetutorial.net/sqlite-java/select/>

<https://www.jc-mouse.net/java/microservicio-rest-java-con-spark-parte-1>

<https://www.guru99.com/buffered-reader-in-java.html>