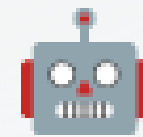
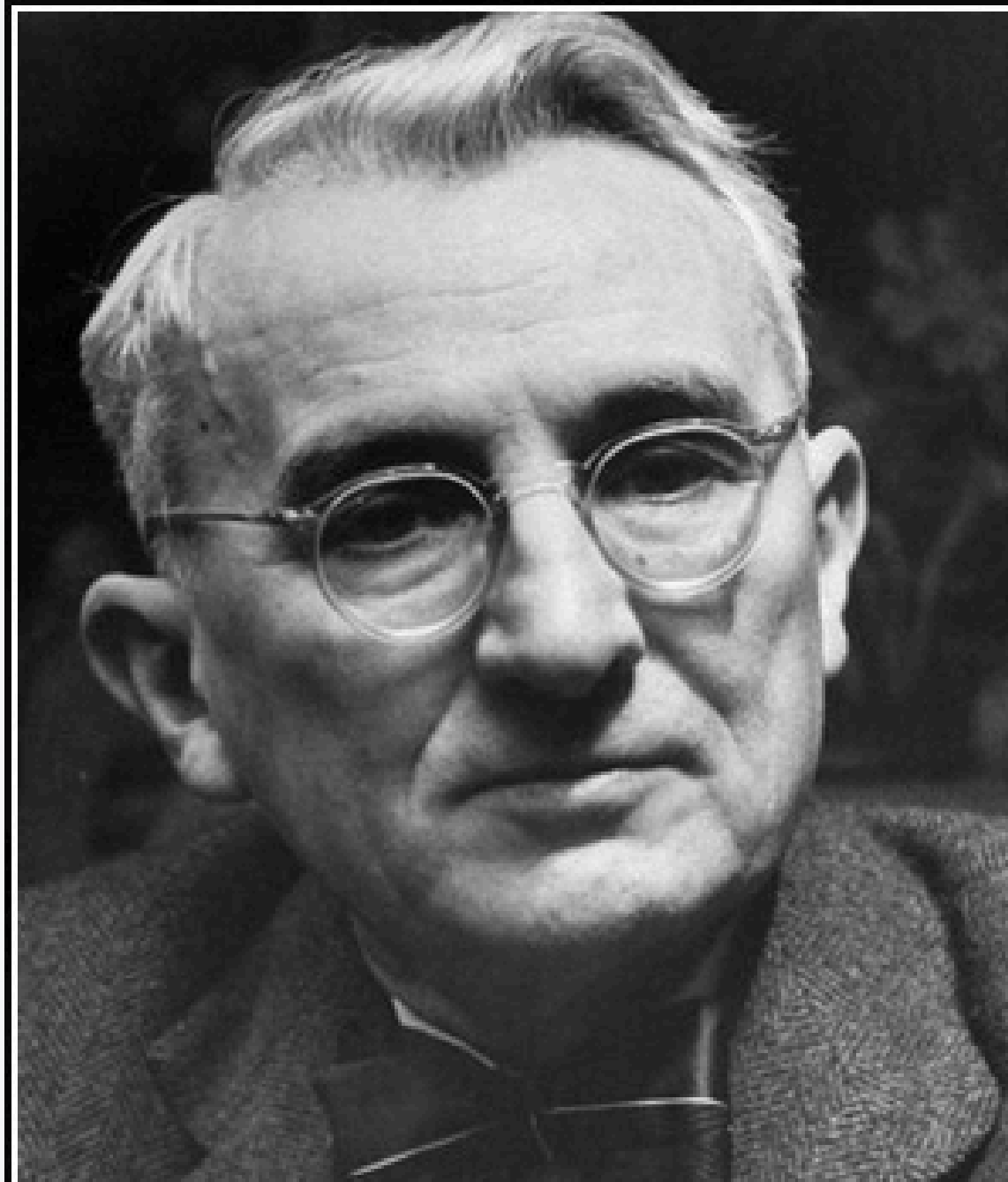


**GRADO Y MÁSTER EN INGENIERÍA Y GESTIÓN DE SISTEMAS DE
INFORMACIÓN - ÁREA DE ESPECIALIZACIÓN EN INGENIERÍA DE
SISTEMAS DE INFORMACIÓN**



TECHNOLOGY CASE

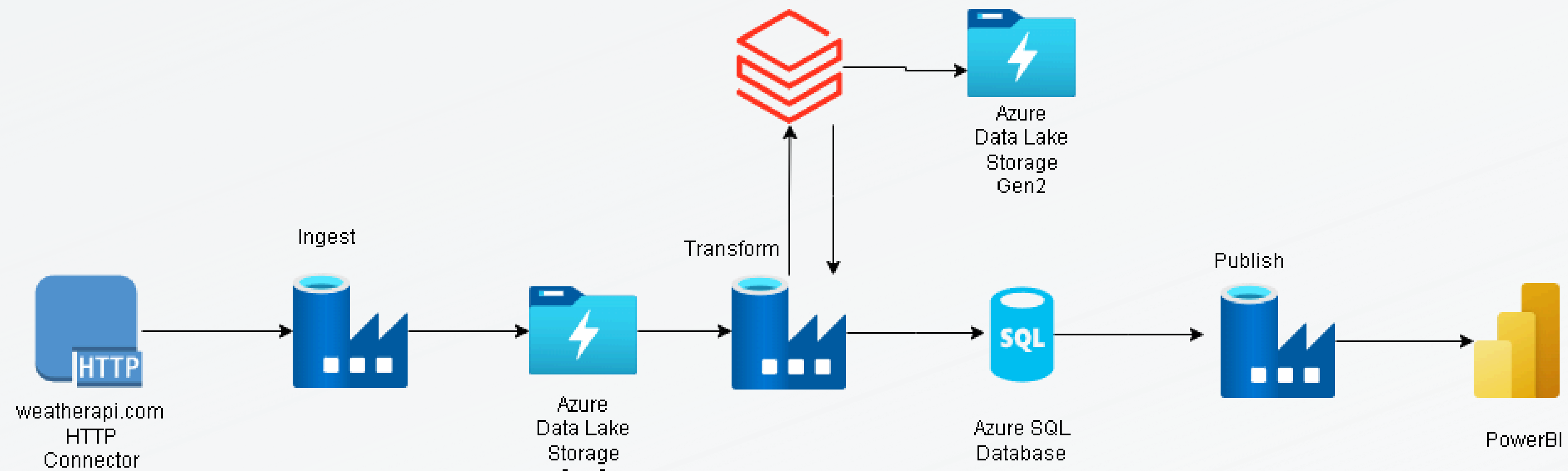
GUSTAVO JOÃO GONÇALVES CALDAS



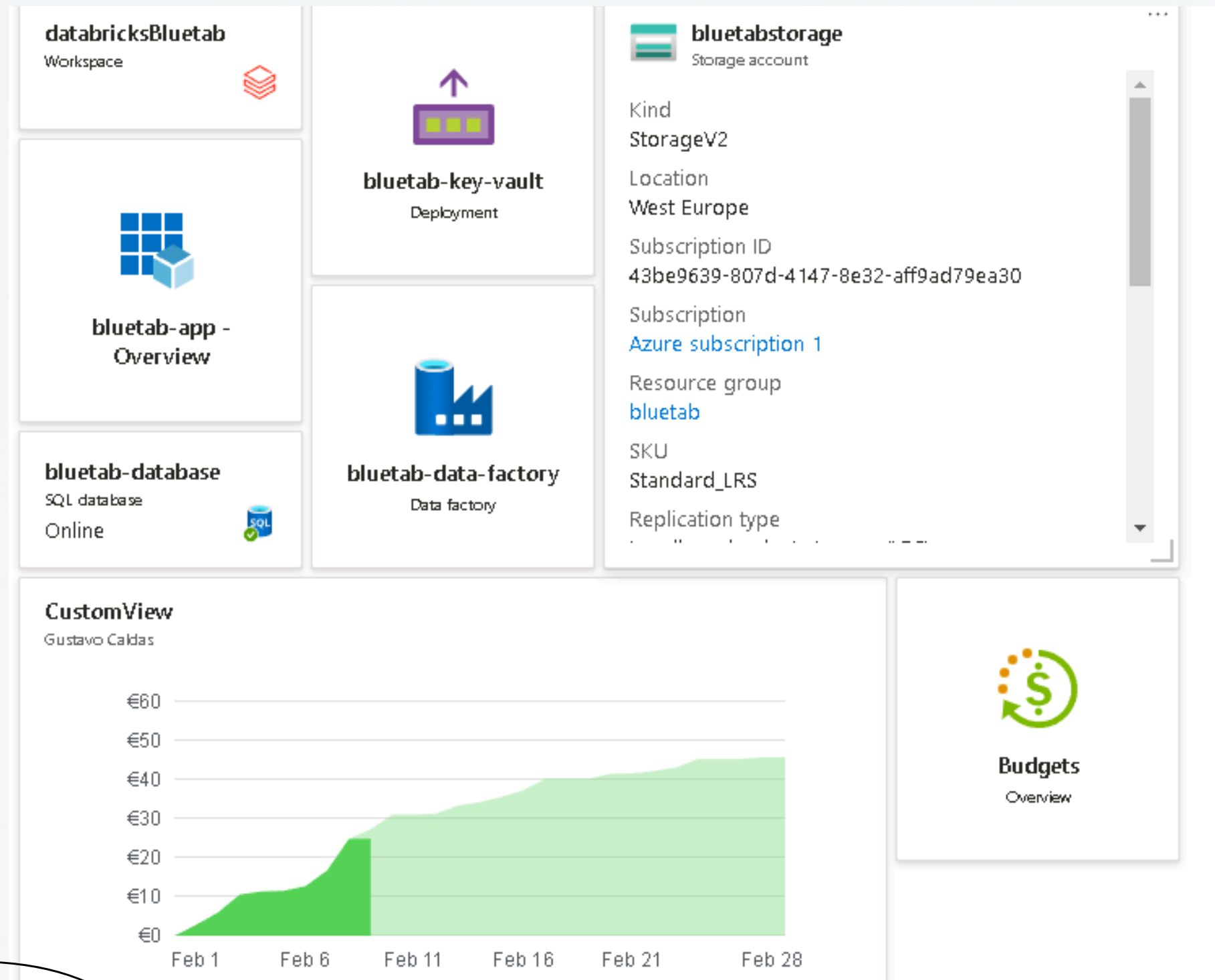
Learning is an active process. We learn by doing.. Only knowledge that is used sticks in your mind.

— Dale Carnegie —

ARCHITECTURE



SERVICES



SET-UP

DATA FACTORY LINKED SERVICES

Linked services

Linked service defines the connection information to a data store or compute. [Learn more](#)

+ New

Filter by name Annotations : **Any**

Showing 1 - 7 of 7 items

Name ↑↓	Type ↑↓	Rel
AzureDatabricks1	Azure Databricks	4
AzureDataLakeStorage1	Azure Data Lake Storage Gen2	2
Is_ablob_bluetabsimplestorage	Azure Blob Storage	2
Is_adls_bluetabstorage	Azure Data Lake Storage Gen2	14
Is_http_opendata_ecdc_europa_eu	HTTP	1
Is_http_weather_api	HTTP	1
Is_sql_covid_db	Azure SQL Database	4

[Learn more](#)

Name *

Is_http_weather_api

Description

Connect via integration runtime * ⓘ

AutoResolveIntegrationRuntime

Base URL *

https://api.weatherapi.com

Information will be sent to the URL specified. Please ensure you trust the URL entered.

Server certificate validation ⓘ

☒ Enable ☐ Disable

Authentication type * ⓘ

Anonymous

SET-UP

DATA FACTORY FILE LIST

▲ Pipelines11

▶ Covid Project7

▶ F1 Project3

▲ Real-Time Weather Project1

pl_ingest_weather_madrid

▶ Change Data Capture (preview)0

▲ Datasets23

▶ Covid Project14

▶ F1 Project2

▲ Weather Project7

ds_sql_weather

ds_sql_weather_forecast

ds_weather_filelist

ds_weather_forecast_processed

ds_weather_processed

ds_weather_raw_json_dl

ds_weather_raw_json_http

▶ Data flows2

▶ Power Query0

JSON
ds_weather_filelist

Preview data

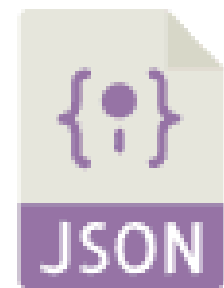
Linked service: ls_adls_bluetabstorage

Object: weather_file_list.json

```
[
  {
    "sourceRelativeURL": "v1/forecast.json?key=1058c00ce8b54ad2939100319250802&q=Madrid&days=7",
    "sinkFileName": "raw/weather_madrid.json"
  },
  {
    "sourceRelativeURL": "v1/forecast.json?key=1058c00ce8b54ad2939100319250802&q=Alicante, Spain&days=7",
    "sinkFileName": "raw/weather_alicante.json"
  },
  {
    "sourceRelativeURL": "v1/forecast.json?key=1058c00ce8b54ad2939100319250802&q=Barcelona, Spain&days=7",
    "sinkFileName": "raw/weather_barcelona.json"
  },
  {
    "sourceRelativeURL": "v1/forecast.json?key=1058c00ce8b54ad2939100319250802&q=Bilbao, Spain&days=7",
    "sinkFileName": "raw/weather_bilbao.json"
  },
  {
    "sourceRelativeURL": "v1/forecast.json?key=1058c00ce8b54ad2939100319250802&q=Distrito%20Federal, Mexico&days=7",
    "sinkFileName": "raw/weather_cidadedomexico.json"
  },
  {
    "sourceRelativeURL": "v1/forecast.json?key=1058c00ce8b54ad2939100319250802&q=Lima, Peru&days=7",
    "sinkFileName": "raw/weather_limaperu.json"
  },
  {
    "sourceRelativeURL": "v1/forecast.json?key=1058c00ce8b54ad2939100319250802&q=bogota&days=7",
    "sinkFileName": "raw/weather_bogota.json"
  }
]
```


SET-UP

DATA FACTORY DS



JSON

ds_weather_raw_json_http

Connection

Schema

Parameters

Linked service *

Is_http_weather_api

Test connection

Connection successful

Base URL

https://api.weatherapi.com

Relative URL ①

@dataset().relativeURL

Preview data

Compression type

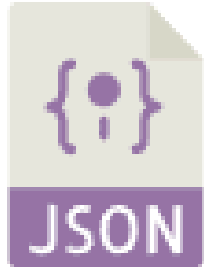
No compression

Encoding

Default(UTF-8)

SET-UP

DATA FACTORY DS

 JSON
ds_weather_raw_json_dl

Connection

Schema

Parameters

Linked service *

Is_adls_bluetabstorage

Test connection

Edit

+ New

Learn more

File path

weather

/

Directory

/

@dataset().fileName

Compression type

No compression

Encoding

Default(UTF-8)

SET-UP DATABRICKS

Microsoft Azure | databricks

Search

+ New

Workspace

Recents

Catalog

Workflows

Compute

Marketplace

SQL >

Data Engineering

Job Runs

Data Ingestion

Delta Live Tables

Machine Learning >

HomePage / Create Secret Scope

Create Secret Scope

Cancel

Create

A store for secrets that is identified by a name and backed by a specific store type. [Learn more](#)

Scope Name ⓘ
bluetab-scope

Manage Principal ⓘ
All workspace users

Azure Key Vault ⓘ
DNS Name
https://bluetab-key-vault.vault.azure.net/
Resource ID
/subscriptions/43be9639-807d-4147-8e32-aff9ad79ea30/resourceGroups/Bluetab/

SET-UP

DATABRICKS

 **bluetab-key-vault** | Secrets

Key vault

Search

◊

«

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Access policies

Events

+ Generate/Import

↺ Refresh

↑ Restore Backup

</> View sample code

🔑 Manage deleted secrets

Name	Type	Status
bluetab-account-key	Storage Account Key	✓ Enabled
bluetab-client-id1		✓ Enabled
bluetab-client-secret1		✓ Enabled
bluetab-demo-sas-token1	sas-token	✓ Enabled
bluetab-tenant-id1		✓ Enabled

SET-UP DATABRICKS

Mount Azure Data Lake Containers for the Project

```
def mount_adls(storage_account_name, container_name):

    client_id = dbutils.secrets.get(scope = 'bluetab-scope', key = 'bluetab-client-id1')
    tenant_id = dbutils.secrets.get(scope = 'bluetab-scope', key = 'bluetab-tenant-id1')
    client_secret = dbutils.secrets.get(scope = 'bluetab-scope', key = 'bluetab-client-secret1')

    configs = {"fs.azure.account.auth.type": "OAuth",
               "fs.azure.account.oauth.provider.type": "org.apache.hadoop.fs.azurebfs.oauth2.ClientCredsTokenProvider",
               "fs.azure.account.oauth2.client.id": client_id,
               "fs.azure.account.oauth2.client.secret": client_secret,
               "fs.azure.account.oauth2.client.endpoint": f"https://login.microsoftonline.com/{tenant_id}/oauth2/token"}

    if any(mount.mountPoint == f"/mnt/{storage_account_name}/{container_name}" for mount in dbutils.fs.mounts()):
        dbutils.fs.unmount(f"/mnt/{storage_account_name}/{container_name}")

    dbutils.fs.mount(
        source = f"abfss://{container_name}@{storage_account_name}.dfs.core.windows.net/",
        mount_point = f"/mnt/{storage_account_name}/{container_name}",
        extra_configs = configs)

    display(dbutils.fs.mounts())
```

```
mount_adls('bluetabstorage', 'weather')
```

```
display(dbutils.fs.ls("/mnt/bluetabstorage/weather"), truncate=False)
```

(2) Spark Jobs

	path	name	size	modificationTime
1	dbfs:/mnt/bluetabstorage/weather...	raw/	0	1739012921000

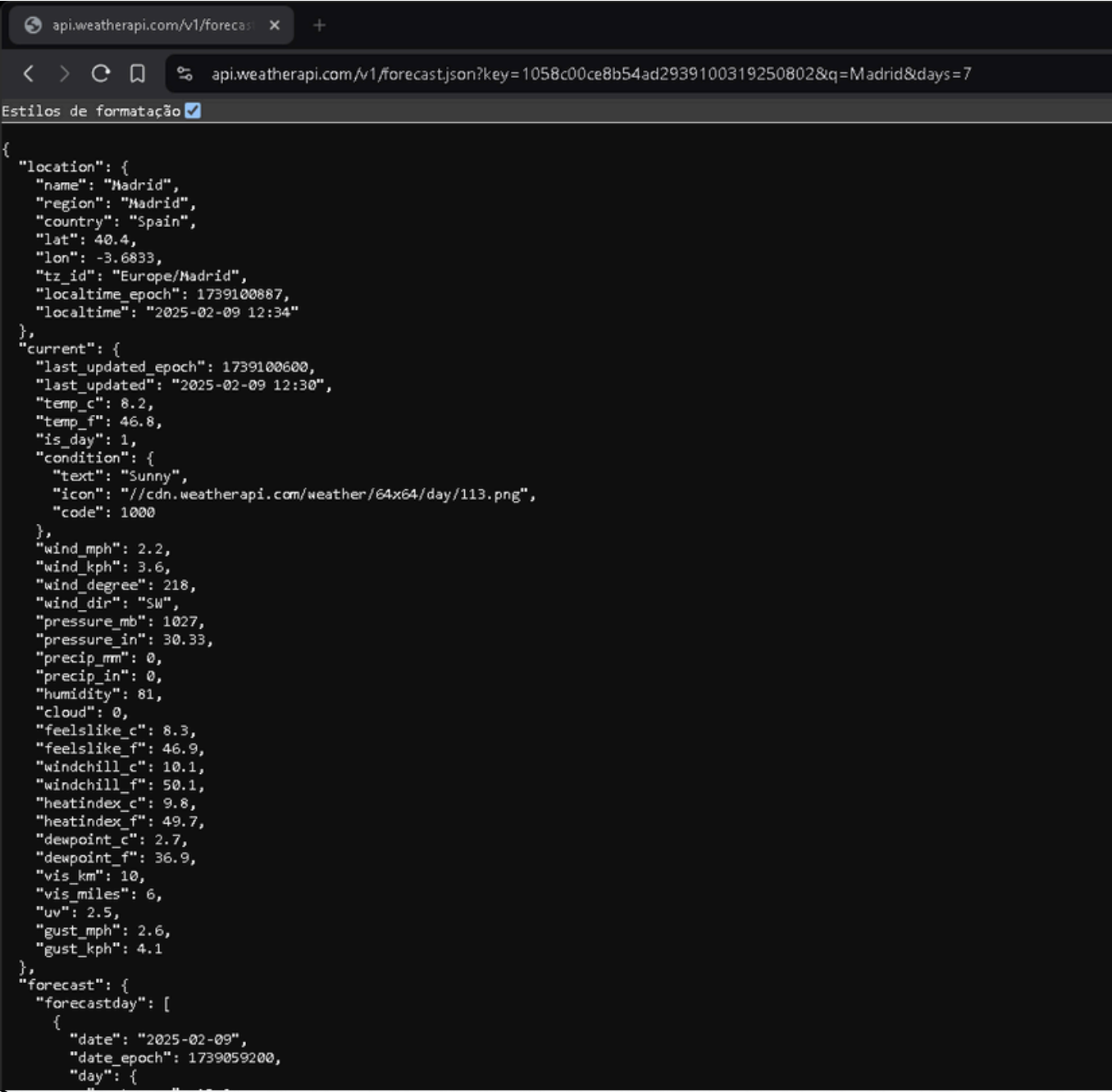
SET-UP

AZURE SQL

The screenshot displays the Azure SQL Query editor interface. The top bar shows the connection name "bluetab-database (bluetab-server/bluetab-database)" and the title "Query editor (preview)". Below the top bar is a navigation menu with options like Search, Login, New Query, Open query, Feedback, and Getting started. The left sidebar contains a list of navigation items: Overview, Activity log, Tags, Diagnose and solve problems, Query editor (preview), Mirror database in Fabric (preview), Settings, Data management, Integrations, Power Platform, Security, Intelligent performance, Monitoring, Automation, and Help. The main area is divided into two panes. The left pane shows the object explorer for "bluetab-database (gj.goncalvescal...)" with a message indicating limited functionality and a list of objects: Tables, Views, and Stored Procedures. The right pane, titled "Query 1", contains a SQL script for creating a schema and two tables.

```
1 CREATE SCHEMA weather;
2 GO
3
4 CREATE TABLE weather.forecast_weather (
5     ciudad NVARCHAR(100) NULL,
6     país NVARCHAR(100) NULL,
7     fecha DATE NULL,
8     temperatura_c FLOAT NULL,
9     humedad_porcentaje INT NULL,
10    indice_uv FLOAT NULL,
11    visibilidad_km FLOAT NULL,
12    precipitación_mm FLOAT NULL,
13    temperatura_max_c FLOAT NULL,
14    temperatura_min_c FLOAT NULL,
15    sunrise NVARCHAR(50) NULL,
16    sunset NVARCHAR(50) NULL
17 );
18 GO
19
20 CREATE TABLE weather.real_time_weather (
21     ciudad NVARCHAR(100) NULL,
```


API



The screenshot shows a web browser window with the address bar displaying the URL: `api.weatherapi.com/v1/forecast.json?key=1058c00ce8b54ad2939100319250802&q=Madrid&days=7`. The page content displays a JSON response from the API, which includes location details, current weather conditions, and a forecast for the next seven days. The JSON is formatted with syntax highlighting.

```
{
  "location": {
    "name": "Madrid",
    "region": "Madrid",
    "country": "Spain",
    "lat": 40.4,
    "lon": -3.6833,
    "tz_id": "Europe/Madrid",
    "localtime_epoch": 1739100887,
    "localtime": "2025-02-09 12:34"
  },
  "current": {
    "last_updated_epoch": 1739100600,
    "last_updated": "2025-02-09 12:30",
    "temp_c": 8.2,
    "temp_f": 46.8,
    "is_day": 1,
    "condition": {
      "text": "Sunny",
      "icon": "//cdn.weatherapi.com/weather/64x64/day/113.png",
      "code": 1000
    },
    "wind_mph": 2.2,
    "wind_kph": 3.6,
    "wind_degree": 218,
    "wind_dir": "SW",
    "pressure_mb": 1027,
    "pressure_in": 30.33,
    "precip_mm": 0,
    "precip_in": 0,
    "humidity": 81,
    "cloud": 0,
    "feelslike_c": 8.3,
    "feelslike_f": 46.9,
    "windchill_c": 10.1,
    "windchill_f": 50.1,
    "heatindex_c": 9.8,
    "heatindex_f": 49.7,
    "dewpoint_c": 2.7,
    "dewpoint_f": 36.9,
    "vis_km": 10,
    "vis_miles": 6,
    "uv": 2.5,
    "gust_mph": 2.6,
    "gust_kph": 4.1
  },
  "forecast": {
    "forecastday": [
      {
        "date": "2025-02-09",
        "date_epoch": 1739059200,
        "day": {
          "date": "2025-02-09",
          "date_epoch": 1739059200,
          "day": {
            "temp_c": 8.2,
            "temp_f": 46.8,
            "is_day": 1,
            "condition": {
              "text": "Sunny",
              "icon": "//cdn.weatherapi.com/weather/64x64/day/113.png",
              "code": 1000
            },
            "wind_mph": 2.2,
            "wind_kph": 3.6,
            "wind_degree": 218,
            "wind_dir": "SW",
            "pressure_mb": 1027,
            "pressure_in": 30.33,
            "precip_mm": 0,
            "precip_in": 0,
            "humidity": 81,
            "cloud": 0,
            "feelslike_c": 8.3,
            "feelslike_f": 46.9,
            "windchill_c": 10.1,
            "windchill_f": 50.1,
            "heatindex_c": 9.8,
            "heatindex_f": 49.7,
            "dewpoint_c": 2.7,
            "dewpoint_f": 36.9,
            "vis_km": 10,
            "vis_miles": 6,
            "uv": 2.5,
            "gust_mph": 2.6,
            "gust_kph": 4.1
          }
        }
      }
    ]
  }
}
```

DATAFACTORY

PIPELINE

The screenshot displays the Azure Data Factory (ADF) interface, divided into three main sections: the left-hand navigation pane, the top center canvas, and the bottom configuration pane.

Left-hand navigation pane: This pane shows a hierarchical view of the workspace. The 'Real-Time Weather Project' is selected, containing a pipeline named 'pl_ingest_weather_madrid'. Below it, a list of datasets is shown, including 'ds_sql_weather', 'ds_sql_weather_forecast', 'ds_weather_filelist', 'ds_weather_forecast_processed', 'ds_weather_processed', 'ds_weather_raw_json_dl', and 'ds_weather_raw_json_http'. The 'Data flows' and 'Power Query' sections are also visible.

Top center canvas: This area shows the visual representation of the pipeline. It starts with a 'Lookup' activity (green checkmark) connected to a 'ForEach' activity (blue box with green checkmark). The 'ForEach' activity contains a 'Copy Weather...' activity. The pipeline then branches into two parallel paths. Each path consists of a 'Notebook' activity (green checkmark) followed by a 'Copy data' activity (green checkmark). The first path uses 'Transformation Data Current' and 'Sqlize'. The second path uses 'Transformation Data Forecast' and 'Sqlize_copy1'.

Bottom configuration pane: This pane shows the configuration for the selected 'Copy Weather...' activity. The 'Source' tab is active, displaying the 'Source dataset' as 'ds_weather_raw_json_http'. Below this, the 'Dataset properties' section is visible, showing the 'relativeURL' property set to '@item().sourceRelativeURL'.

DATAFACTORY

OUTPUT

▶ 15 hours ago (4s)

3

Python

```
df_madrid = extraer_datos_forecast(spark.read.json("dbfs:/mnt/bluetabstorage/weather/raw/weather_madrid.json"))
df_alicante = extraer_datos_forecast(spark.read.json("dbfs:/mnt/bluetabstorage/weather/raw/weather_alicante.json"))
df_barcelona = extraer_datos_forecast(spark.read.json("dbfs:/mnt/bluetabstorage/weather/raw/weather_barcelona.json"))
df_bilbao = extraer_datos_forecast(spark.read.json("dbfs:/mnt/bluetabstorage/weather/raw/weather_bilbao.json"))
df_ciudademexico = extraer_datos_forecast(spark.read.json("dbfs:/mnt/bluetabstorage/weather/raw/weather_ciudademexico.json"))
df_limaperu = extraer_datos_forecast(spark.read.json("dbfs:/mnt/bluetabstorage/weather/raw/weather_limaperu.json"))
df_bogota = extraer_datos_forecast(spark.read.json("dbfs:/mnt/bluetabstorage/weather/raw/weather_bogota.json"))
```

▶ (7) Spark Jobs

df_alicante: pyspark.sql.dataframe.DataFrame = [ciudad: string, país: string ... 10 more fields]

df_barcelona: pyspark.sql.dataframe.DataFrame = [ciudad: string, país: string ... 10 more fields]

df_bilbao: pyspark.sql.dataframe.DataFrame = [ciudad: string, país: string ... 10 more fields]

df_bogota: pyspark.sql.dataframe.DataFrame = [ciudad: string, país: string ... 10 more fields]

df_ciudademexico: pyspark.sql.dataframe.DataFrame = [ciudad: string, país: string ... 10 more fields]

df_limaperu: pyspark.sql.dataframe.DataFrame = [ciudad: string, país: string ... 10 more fields]

df_madrid: pyspark.sql.dataframe.DataFrame = [ciudad: string, país: string ... 10 more fields]

▶ 15 hours ago (1s)

4

Python

```
df_final = df_madrid.unionByName(df_alicante).unionByName(df_barcelona).unionByName(df_bilbao).unionByName(df_ciudademexico).unionByName(df_limaperu).unionByName(df_bogota)
df_final = df_final.orderBy("ciudad", "país", "fecha")
display(df_final)
df_final.printSchema()
```

▶ (1) Spark Jobs

df_final: pyspark.sql.dataframe.DataFrame = [ciudad: string, país: string ... 10 more fields]

Table

+

Q

S

	A ^B ciudad	A ^B país	A ^B fecha	1.2 temperatura_c	1.2 humedad_porcentaje	1.2 índice_uv	1.2 visibilidad_km	1.2 precipitación_mm	1.2 ten
36	Lima	Peru	2025-02-08	22.7	80	3.3	10	0.01	
37	Lima	Peru	2025-02-09	22.6	80	2.9	10	0.02	
38	Lima	Peru	2025-02-10	22.9	75	3.1	10	0.05	
39	Lima	Peru	2025-02-11	23.2	72	3.2	10	0.3	
40	Lima	Peru	2025-02-12	23.3	72	0.1	10	0.01	
41	Lima	Peru	2025-02-13	23.5	77	9	10	0.01	
42	Lima	Peru	2025-02-14	23	81	9	10	0.01	
43	Madrid	Spain	2025-02-08	6.7	73	0.6	10	0	
44	Madrid	Spain	2025-02-09	8.5	68	0.5	10	0	
45	Madrid	Spain	2025-02-10	10.5	67	0.4	10	0	
46	Madrid	Spain	2025-02-11	10.4	66	0.5	10	0.86	
47	Madrid	Spain	2025-02-12	10.4	73	0.3	8.8	11.12	
48	Madrid	Spain	2025-02-13	8.3	71	3	10	0	
49	Madrid	Spain	2025-02-14	10.2	60	3	10	0	

POWERBI

DATA

Consultas [2]

✕

✓

f_x

= Origem{[Schema="weather",Item="real_time_weather"]}[Data]

▼

weather real_time_weath...

weather forecast_weather

	A ^B _C ciudad	A ^B _C país	A ^B _C hora_local	A ^B _C última_atualización	1.2 temperatura_c	1.2 sensación_térmica_c	1.2 viento
1	Madrid	Spain	09/02/2025 10:51:00	09/02/2025 10:45:00	6,2	6	
2	Alicante	Spain	09/02/2025 10:51:00	09/02/2025 10:45:00	15,1	15,1	
3	Barcelona	Spain	09/02/2025 10:51:00	09/02/2025 10:45:00	8,3	7,8	
4	Bilbao	Spain	09/02/2025 10:52:00	09/02/2025 10:45:00	10	8,7	
5	Distrito Federal	Mexico	09/02/2025 03:52:00	09/02/2025 03:45:00	16,8	16,8	
6	Lima	Peru	09/02/2025 04:45:00	09/02/2025 04:45:00	21,8	21,8	
7	Bogota	Colombia	09/02/2025 04:59:00	09/02/2025 04:45:00	13,4	14	

Definições da Consulta

PROPRIEDADES

Nome

weather real_time_weather

Todas as Propriedades

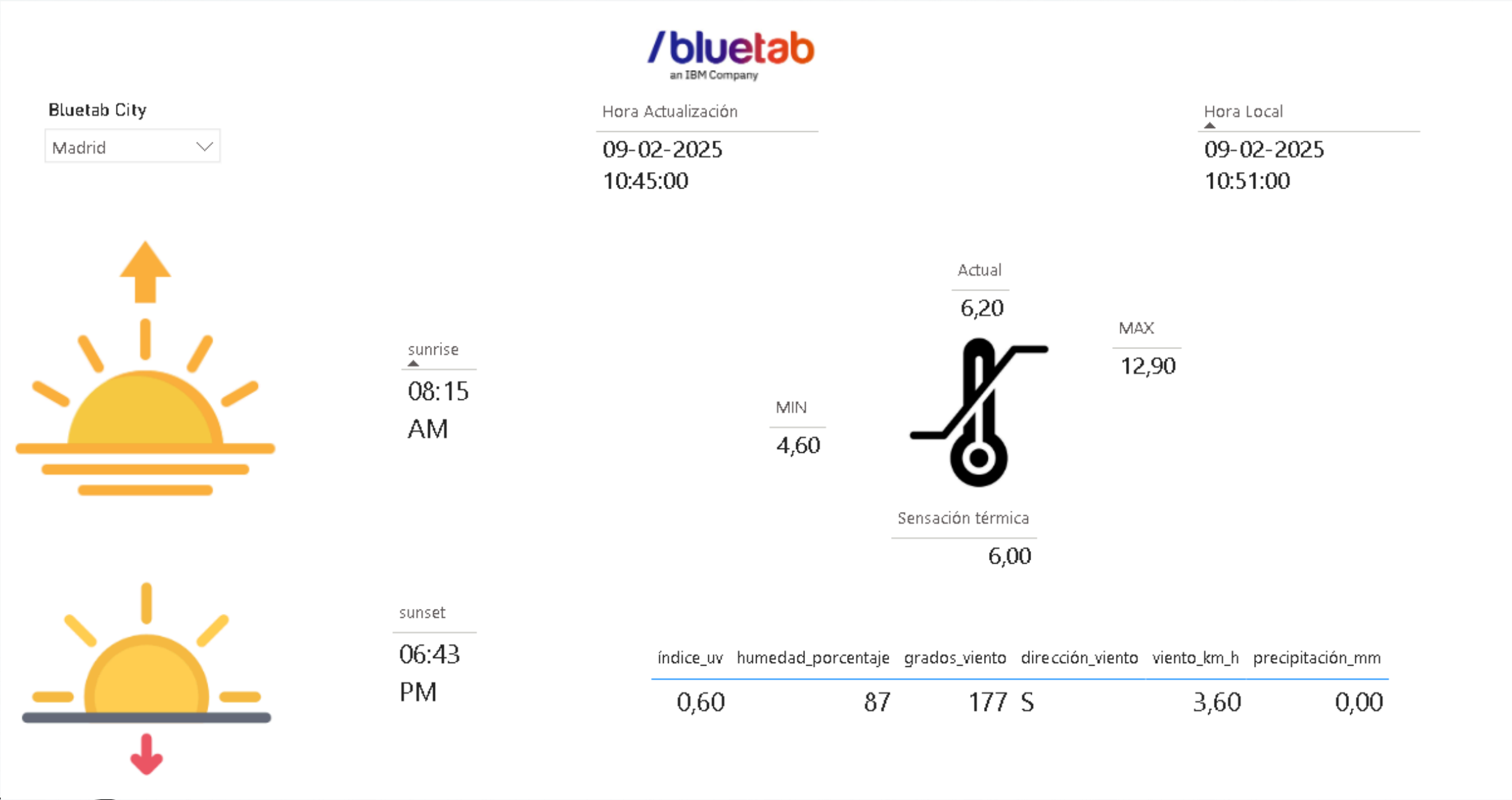
PASSOS APLICADOS

Origem

✕ Navegação

POWERBI

DASHBOARD



POWERBI

DASHBOARD

Bluetab City

Madrid

an IBM Company

país	ciudad	fecha	sunrise	sunset	temperatura_c	temperatura_max_c	temperatura_min_c	humedad_porcentaje	precipitación_mm	índice_uv	visibilidad_km
Spain	Madrid	09-02-25	08:15 AM	06:43 PM	8,20	12,90	4,60	69	0,00	0,60	10,00
Spain	Madrid	10-02-25	08:14 AM	06:44 PM	9,80	13,60	7,70	68	0,00	0,50	10,00
Spain	Madrid	11-02-25	08:13 AM	06:45 PM	9,80	14,50	7,20	67	0,31	0,50	9,70
Spain	Madrid	12-02-25	08:12 AM	06:46 PM	8,40	9,40	7,90	86	9,81	0,30	7,50
Spain	Madrid	13-02-25	08:10 AM	06:48 PM	9,60	13,30	7,80	70	0,02	0,50	10,00
Spain	Madrid	14-02-25	08:09 AM	06:49 PM	9,40	14,20	5,50	62	0,00	3,00	10,00
Spain	Madrid	15-02-25	08:08 AM	06:50 PM	9,90	14,90	5,70	59	0,00	3,00	10,00

LINKS



WEATHER IN BLUETAB CITIES

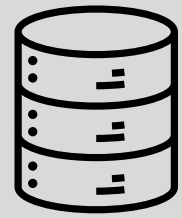
[LINK](#)



INTERACTIVE POWERBI

[LINK](#)

LINKS



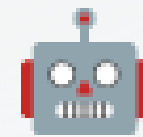
DATAFACTORY REPOSITORY

[LINK](#)



PROFILE
[LINK](#)

**GRADO Y MÁSTER EN INGENIERÍA Y GESTIÓN DE SISTEMAS DE
INFORMACIÓN - ÁREA DE ESPECIALIZACIÓN EN INGENIERÍA DE
SISTEMAS DE INFORMACIÓN**



TECHNOLOGY CASE

GUSTAVO JOÃO GONÇALVES CALDAS