



dbt + python = dbt-fal

python within dbt models



Who are we?

Som Energia is a **consumer cooperative**

- Renewable green energy
- not-for-profit
- Solidarity-based community
- Present through Local Groups
- **More than ~83k members**
- **~120k contracts**
- **65GWh/year @ 17 plants**
- <https://github.com/Som-Energia>
- **since 2010**



Producció +

Transport  Distribució 

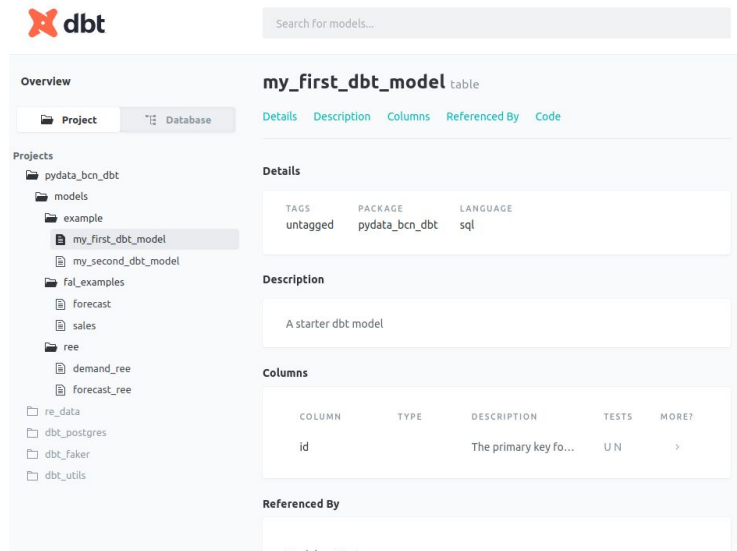
Comercialització +

[More details](#)



What is dbt?

- SQL-first **transformation** framework
- Data **modeling, testing** and **documentation**
- configure with **YAML**
- Relies on Jinja2 **SQL** templates
- Compatible with **CI/CD** and **Git**
- [Great community](#)
- Great [documentation](#), [examples](#) and [discussions](#)



dbt is SQL only

```
-- my_first_dbt_model.sql
{{ config(materialized="table") }}
with
    source_data as (
        select 1 as id
        union all
        select null as id
    )
select *
from source_data
where id is not null
```

```
-- Use the `ref` function to select from other
models
select * from {{ ref("my_first_dbt_model") }}
where id = 1
```

Python support is limited (for now)

- Targets snowflake, databricks and BigQuery only
- More **flexible** than *just* SQL
 - ML models
 - Notifications
 - etc
- pandas recommendation is still not clear. Some alternatives are ibis and polars
- Not the best debugging experience, some workarounds exist

```
# model.py
def model(dbt, session):
    # config
    dbt.config(
        materialized="table")
    df = dbt.ref("my_model")
    ...
    # some transformations taking place
    return df
```

dbt + fal = dbt-fal

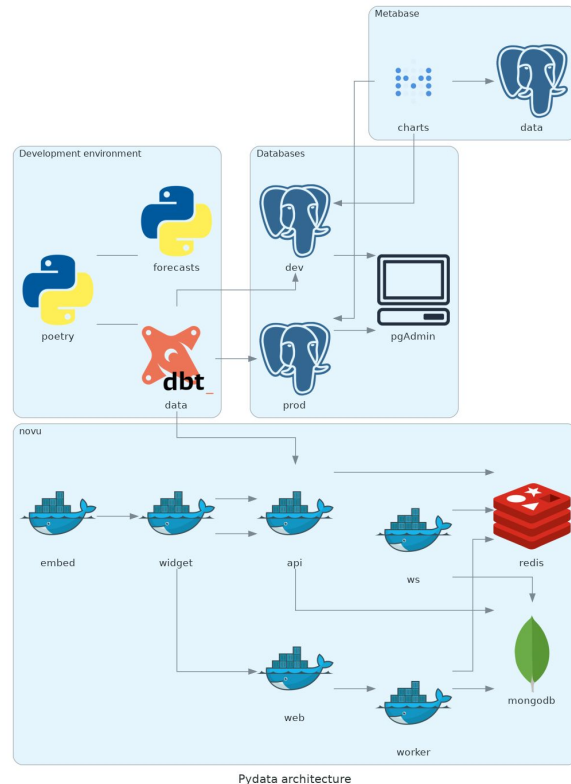
- Python **adapter**
- Brings in support to **other solutions** e.g. **Postgres**
- **Virtual environment** support

```
# profiles.yml
pydata_bcn_dbt:
  outputs:
    dev:
      type: postgres
  dev_with_fal: # new
    type: fal # new
  db_profile: dev # new
  target: dev_with_fal # new
```

```
# schema.yml
---
version: 2
models:
  - name: my_sales_model
    description: A model using
fal-dbt
    config:
      materialized: table
  meta:
    fal:
      scripts:
        - model.py
```

Demo

- [docker](#) + [docker compose](#)
- [novu](#) as notification structure
- [postgres](#) as dev/prod databases
- [pgadmin](#) for db admin
- [metabase](#) for BI
- [devcontainer](#) with [poetry](#)
 - dbt, dbt-fal, prophet, novu-client and other goodies
- **not a production setting**
- [github repo](#)



done with <https://diagrams.mingrammer.com/>

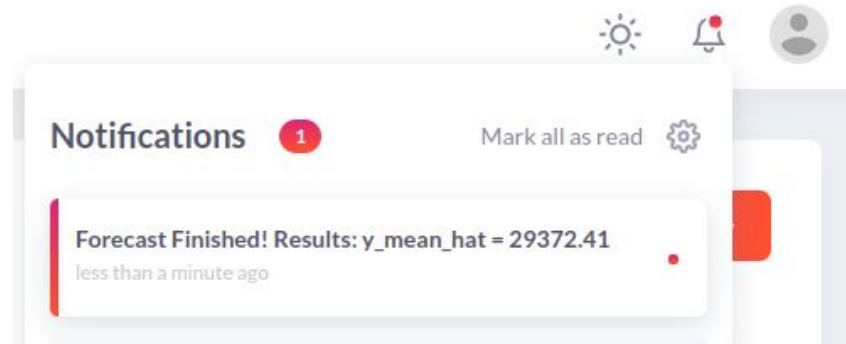


Demo

Demo Outline

- CSV data from [REE](#)
- Loaded to table using **dbt-fal**
- Forecast using [Prophet](#)
- Sends notification to **Novu** client

```
# report the forecast to novu
event_api = EventApi(
    "http://novu-api:3000",
    os.environ.get("NOVU_API_KEY"))
yhat_mean = forecast["yhat"].mean().item()
event_api.trigger(
    name=os.environ.get("NOVU_APP_NAME"),
    recipients=os.environ.get("NOVU_SUBSCRIBER_ID"),
    payload={"results": {"y_mean_hat": f"{yhat_mean:0.2f}"}})
```



Takeaways

- dbt makes pipelines easier to test, document and manage
- dbt relies on SQL and YAML for the most part
- Some tasks can't be / should not be implemented using SQL
- dbt-fal lets you run models using python
- Interfaces are a bit clunky at the moment, but support *will* increase

fal examples



Introducing fal-serverless

We are launching fal-serverless, a cloud based Python runtime that simplifies infrastructure for...

Mar 30, 2023 · 4 min read



Feature Selection for ML with dbt and fal

We explore how to use fal with dbt together for feature selection, starting with synthetic data...

Mar 22, 2023 · 6 min read



Build and Deploy Machine Learning Models from Jupyter Notebooks with fal and dbt

Introduction Machine Learning (ML) is increasingly important in data-driven decision...

Mar 9, 2023 · 8 min read



Composable OLAP

Separation of compute and storage was initially an architectural choice by database vendors to...

Mar 2, 2023 · 3 min read



Building language model powered pipelines with dbt

Yes, we know... Generative AI is all the hype right now. Whether you are an Analytics Engineer or...

Feb 23, 2023 · 5 min read



Populate dbt models with CSV data. Part 2: the power of dbt-fal

In this blog post, we will explore how dbt-fal and dbt can be combined to create a streamlined a...

Feb 8, 2023 · 3 min read

<https://blog.fal.ai/>



Thank you!

dades@somenergia.coop
diego.quintana@somenergia.coop

<https://github.com/Som-Energia/>
<https://github.com/diegoquintana/>