Mobilité & Geography France
Interactive Chatbot Report

Generated on July 14, 2025

1. Introduction

This document presents the scope, structure, and content of the "Mobilité & Geography France" dataset powering our interactive chatbot. The dataset combines:

- **French geographical data**: communes, departments, regions, life-zones, arrondissements, and municipal fusions/scissions through 2024.
- **Low-Emission Zone (ZFE) mappings**: with dates of vignette application.
- **Vehicle fleet composition**: passenger cars (VP) and light commercial vehicles (VUL) from 2011 to 2022, at both national and commune levels, broken down by age bracket, fuel type, sector, and vignette category.

The chatbot allows users to query this rich, multi-table resource in natural language—whether for local code lookups, fleet statistics, or ZFE compliance dates.

2. Objectives

- Support Renault's new mobility initiatives by providing instant access to spatial and temporal transport patterns.
- Enable non-technical users (planners, policy makers, researchers) to explore the data without writing SQL.
- Offer decision-support for local authorities and business units planning vehicle fleets, emission strategies, or infrastructure.

3. Dataset Overview

- **Number of Tables Indexed**: 19
- **Total Rows**: ∼3.5 million records
- **Naming Convention**: `France_Table_<topic>`
- **Main Categories**:
  - *Geography*: communes, departments, regions, life-zones, arrondissements, fusions/scissions.
  - *ZFE*: municipal zones subject to low-emission regulations, with application dates.
  - *Vehicle Fleet*: VP and VUL statistics by age, fuel, sector, vignette—both nationally and per commune.

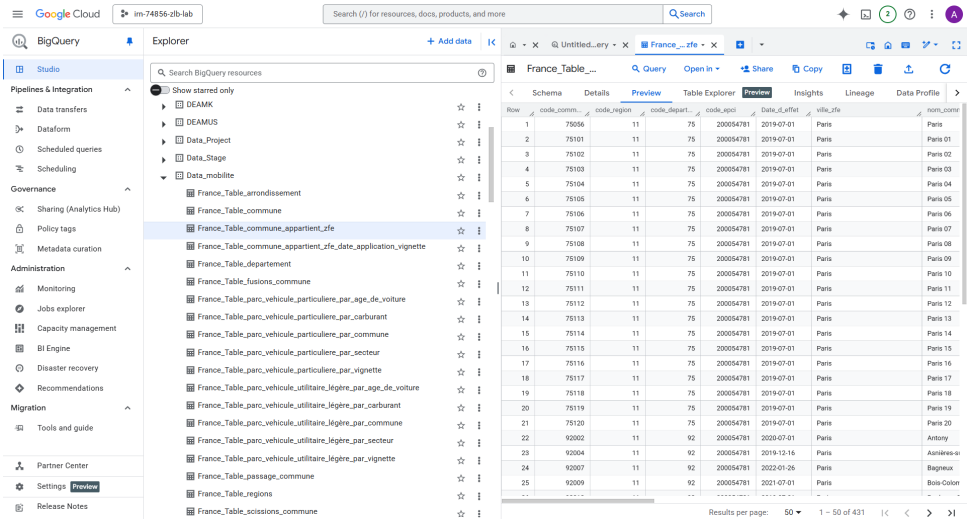

Figure 1: BigQuery dataset tree with all `France_Table_*` entries

4. Detailed Table Descriptions

4.1 Communes

Table: `France_Table_commune`
Columns: `code_commune` (INSEE), `nom_commune`, `code_departement`, `code_region`, plus ZFE and fleet keys.
Rows: 34 935 (all French communes, including overseas).
Purpose: Core lookup for municipality codes and names.

4.2 Departments

Table: `France_Table_departement`
Columns: `code_departement`, `nom_departement`
Rows: 101 (metropolitan + overseas).

4.3 Regions

Table: `France_Table_regions`
Columns: `code_region`, `nom_region`
Rows: 18 (13 metropolitan, 5 overseas).

4.4 Life-Zones (Bassin de vie)

Table: `France_Table_bassin_de_vie_2022`
Columns: `code_bassin_de_vie`, `nom_bassin_de_vie`
Use: Functional urban area delineation for socio-economic studies.

4.5 Arrondissements

Table: `France_Table_arrondissement`
Columns: `code_arrondissement`, `nom_arrondissement`
Rows: ∼330 in metropolitan France.

4.6 Fusions & Scissions

Tables:

- `France_Table_fusions_commune`
- `France_Table_scissions_commune`

Columns: Old/new commune codes, event dates.
Use: Historical tracking of municipality boundary changes.

4.7 Low-Emission Zones (ZFE)

Tables:

- `France_Table_commune_appartient_zfe`
- `France_Table_commune_appartient_zfe_date_application_vignette`

Columns: Commune code → ZFE zone ID, vignette application date.

4.8 Vehicle Fleet Composition

4.8.1 National Level **Tables** (VP & VUL):

- `_parc_vehicule_particuliere_par_age_de_voiture`
- `_parc_vehicule_particuliere_par_carburant`
- `_parc_vehicule_particuliere_par_secteur`
- `_parc_vehicule_particuliere_par_vignette`

Years: 2011–2022.

4.8.2 Commune Level Same tables partitioned by `code_commune`.

5. Indexing & Retrieval Methodology

1. **Metadata Extraction**: Schema + sample unique values formatted into descriptive text.
2. **Embedding Generation**: Vertex AI text embeddings (768-dim) on descriptions.
3. **FAISS Index**: Inner-product flat index for semantic retrieval (RAG).
4. **SQL Generation**: LLM produces parameterized SQL based on user query + retrieved contexts.

6. Chatbot Use Cases & Examples

- "What is the INSEE code for Nice?"
- "How many diesel passenger cars in Île-de-France in 2020?"
- "Which communes joined a ZFE on July 1, 2021?"

7. Access & Download

- **Streamlit Interface**: `https://your-app-url`
- **Download Full PDF Report**:
  Download PDF Context Report

8. Sources & References

- **INSEE**: `https://www.insee.fr/fr/statistiques`
- **Data.gouv.fr**: National open data portal
- **Renault New Mobility**: internal documentation