

# Aggregating Mobility Surveys to Understand Intermodality: Evidence from 68 French Surveys

**Lucas Javaudin**

LVMT, Energy4Climate, Institut Polytechnique de Paris

**WCTR 2026 Toulouse**

**July 9, 2026**

## Outline

Standardization and aggregation of surveys with MobiSurvStd

Intermodality analysis

## Motivations

Mobility surveys (or *household travel surveys*) are essential data sources for examining individual and household travel behaviors

- Household structure and persons' characteristics
- Trips / activities of the day previous to the interview (mode, duration, purpose, origin, destination, etc.)
- In France: national and local surveys

## Motivations

However, they are difficult to exploit:

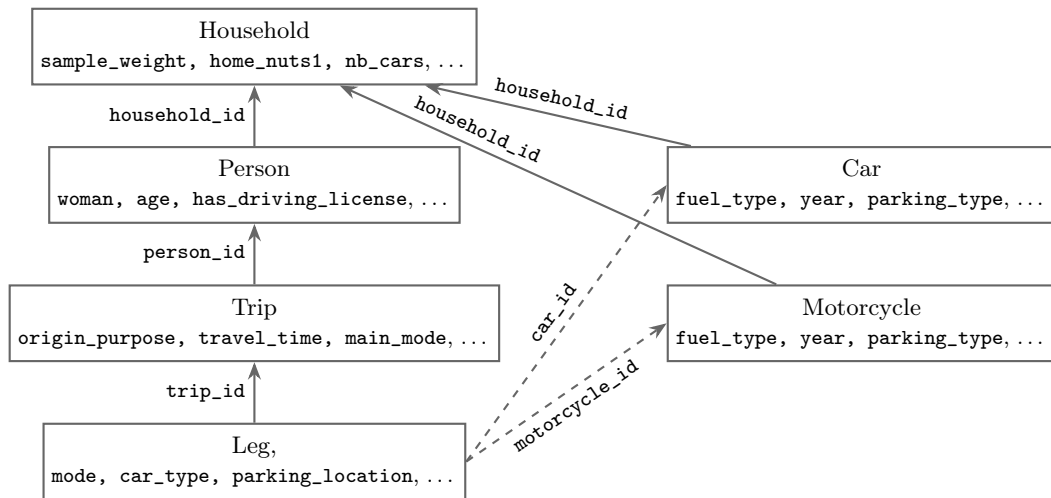
- Diverse formats:
  - ▶ Paris surveys (IDFM: *EGT 2010*, *EGT H2020*; IPR: *EMG*)
  - ▶ local surveys outside Paris (CEREMA: *EMC<sup>2</sup>*, *EDGT*, *EDVM*, *EMD*)
  - ▶ national surveys (SDS: *ENTD 2008*, *EMP 2019*)
- CSV files (+ GIS files) are not easy to read (separator? encoding? datatype?)
- Unclear variable names and modalities (e.g., “P2” for gender with modality 1 for men and 2 for women)
- Complex joint operations between households, persons, trips, legs, cars, etc.

## Solution: MobiSurvStd

- Open-source Python library
- Command-line tool to convert French mobility surveys in a standardized format
- Reading and converting households, persons, trips, legs, cars, motorcycles, and zones
- 286 variables
- Standardized format as Parquet files with a rich and clean documentation
- Currently 7 supported formats (EMC<sup>2</sup>, EDGT, EDVM, EMD, EMP, EGT-2010, EGT-H2020) for about 85 surveys

<https://mobisurvstd.github.io/MobiSurvStd/>

## Entities and their relationships



## Documentation

### main\_mode\_group

Mode group of the main mode of transportation used for the trip.

#### Modalities:

- "walking": walking, excluding wheelchair and personal transporter
- "bicycle": bicycle (traditional or electric, driver or passenger, personal or shared)
- "motorcycle": motorcycle (small and large, driver or passenger)
- "car\_driver": car as a driver
- "car\_passenger": car as a passenger (including taxi and VTC)
- "public\_transit": any form of public transit (excluding employer transport)
- "other": all other modes

#### Guarantees:

- The values are consistent with `main_mode`.

### intermodality

Whether the trip involved using two different modes of transportation.

Note that combining walking with another transportation mode is not classified as intermodality and combining various public-transit modes (e.g., bus then metro) is not either.

#### Type: Boolean

#### Guarantees:

- If the value is `true`, then there is at least two legs with different `mode_group` (excluding walking legs).

### public\_transit\_access\_mode

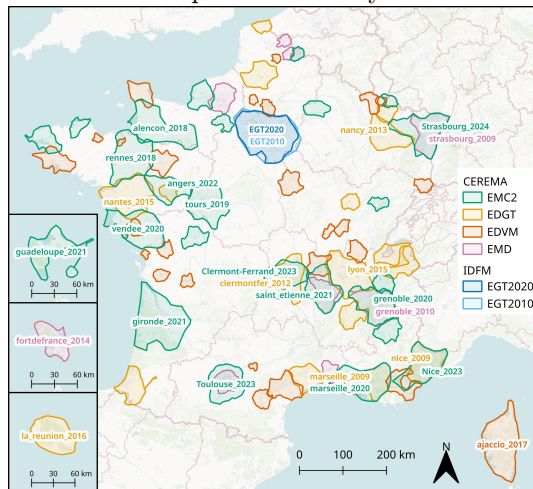
Mode of transportation used for the access part of the trip (for public-transit trip).

#### Modalities:

- Same modalities as `main_mode` (excluding all public-transit modes).

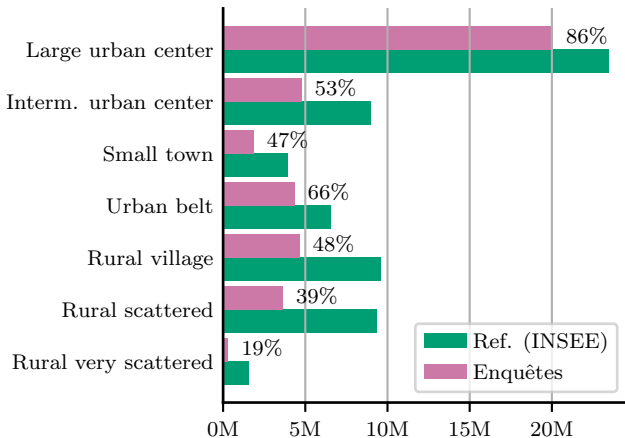
#### Guarantees:

## Map of local surveys



## Aggregation and representativeness

- Aggregation of 68 surveys (local only, most recent version in case of duplicates), covering 62 % of France population
- Better coverage in urban areas than rural areas
- Sample weights are re-calibrated on marginals to get a representative dataset
- Totals: 1.8 M trips, 0.5 M persons



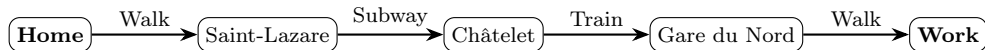
# Outline

Standardization and aggregation of surveys with MobiSurvStd

Intermodality analysis

## Definition of trips in surveys

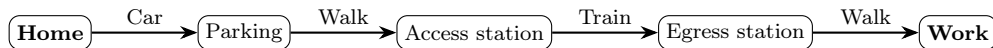
- **Trip:** movement from an origin to a destination for a specific purpose (e.g., work, education, shopping)
- **Leg:** continuous segment with a unique transportation mode
- A trip is composed of one or more leg



## Definition of intermodality

A trip is **intermodal** if it has at least two legs with different modes

- Walking legs are not considered
- All public-transit modes (train, subway, bus, etc.) are grouped in the same category
- Considered modes: car driver, car passenger, public transit, bicycle, motorcycle



Note. Non-local trips ( $> 80$  km) are excluded.

## Stylized fact 1: A small share of trips

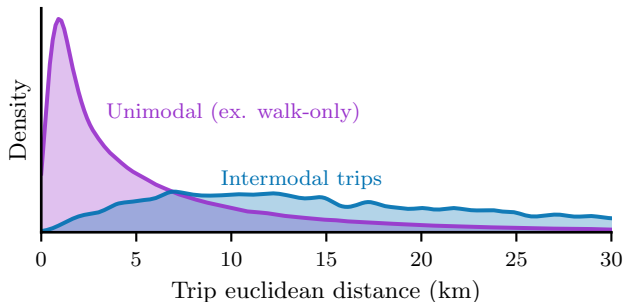
Over the aggregation of 68 surveys

- 16 982 intermodal trips (sample weight of 2.7 millions)
- 1.2 % of trips
- 4.6 % of (euclidean) distance

## Stylized fact 2: Long trips

Mean distance:

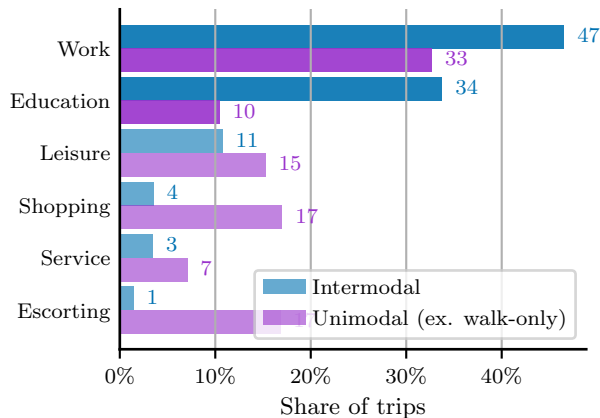
- Intermodal trips: 21.0 km
- Unimodal trips (ex. walk): 7.0 km



## Stylized fact 3: Trips for work or education purposes

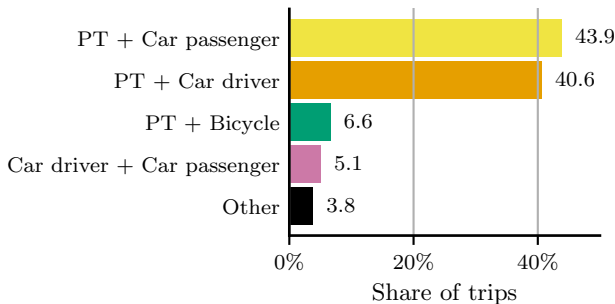
Share of work and education purposes:

- Intermodal trips: 81 %
- Unimodal trips (ex. walk): 43 %



## Stylized fact 4: A majority of car + public transit combinations

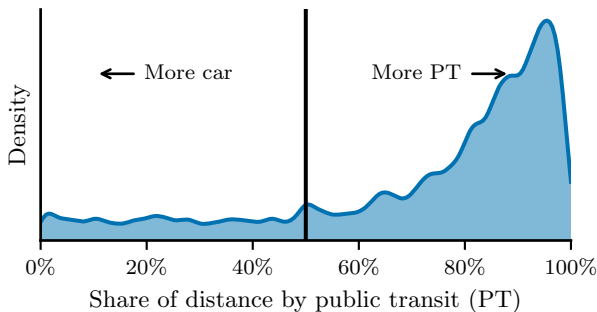
Combining car (driver or passenger) and public transit represents 84.5 % of intermodal trips



## Stylized fact 5: Most of the distance is traveled by public transit

Over the car + public transit trips, the mean euclidean distance is:

- By public transit: 18.8 km
- By car: 5.3 km



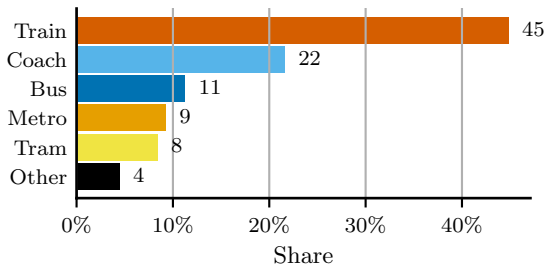
## Stylized fact 6: Train and coaches as main transfer modes

### Train use:

- Transfer car  $\leftrightarrow$  public transit: 45 %
- Overall: 17 %

### Coach use

- Transfer car  $\leftrightarrow$  public transit: 22 %
- Overall: 11 %

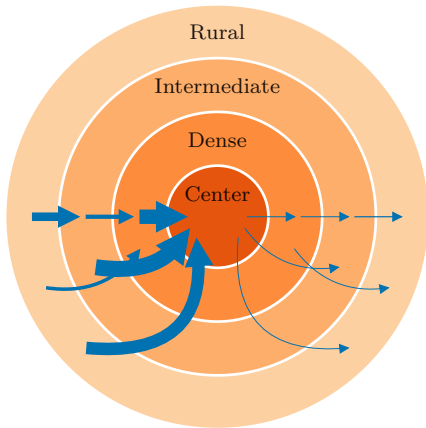


In 62 % of cases, the public-transit part has a single leg.

## Stylized fact 7: Connecting low-density areas to urban centers

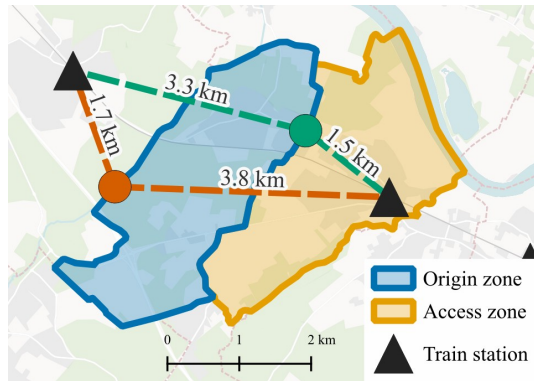
Among car→public transit trips:

- 38 % start from rural municipalities (vs 26 % for unimodal trips)
- 55 % arrive in urban centers (vs 18 % for unimodal trips)



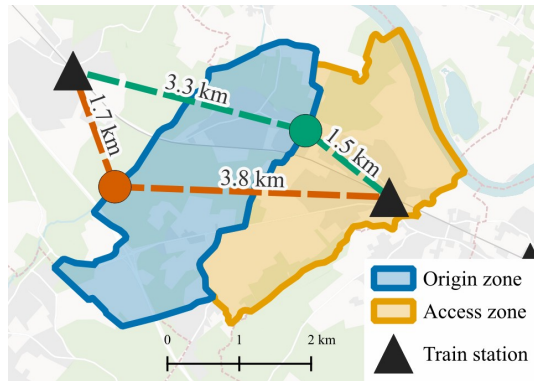
## Bonus: Choice of transfer station

- Surveys specify **origin** and **access** location at a detailed zone level (areas of  $6 \text{ km}^2$  on average)
- Exact location of **public-transit stops** in France are retrieved from GTFS files



## Bonus: Choice of transfer station

- The stop taken is the nearest from origin if, for all vertices of the **origin zone**, the nearest stop is within the **access zone**
- The stop taken is *not* the nearest from origin if, for all vertices of the **origin zone**, the nearest stop is *not* within the **access zone**
- In the other cases, we cannot conclude

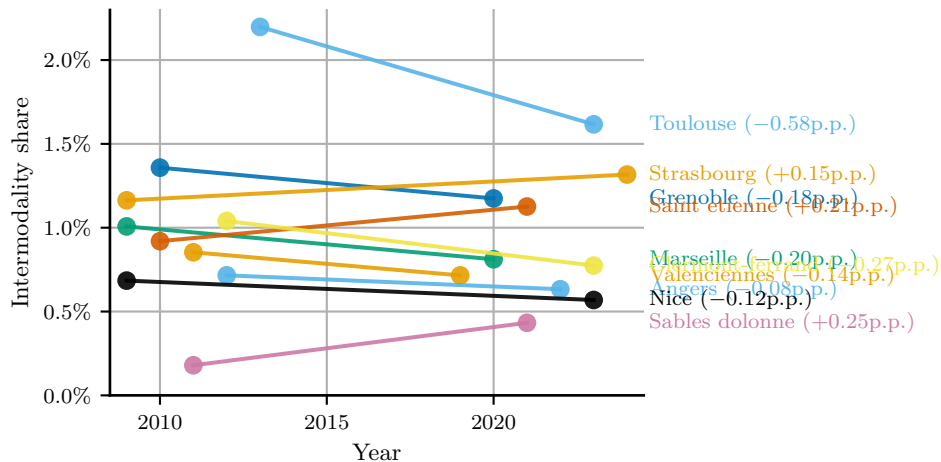


## Bonus: Choice of transfer station

- Train users usually take the nearest train station (between 35 % and 65 % of cases)
- Tramway and subway users seem to use a different logic (parking, line selection, global travel time)

|                            | Train | Tramway | Subway |
|----------------------------|-------|---------|--------|
| Take nearest stop          | 35 %  | 8 %     | 19 %   |
| Does not take nearest stop | 35 %  | 69 %    | 54 %   |
| Cannot know                | 30 %  | 23 %    | 27 %   |

## Bonus 2: Temporal variations



# Thank you

Contact: <https://lucasjavaudin.com> – [lucas.javaudin@enpc.fr](mailto:lucas.javaudin@enpc.fr)

Useful links:

- Survey standardization with MobiSurvStd:  
<https://mobisurvstd.github.io/MobiSurvStd>
- Source code of the intermodality analysis:  
<https://github.com/LucasJavaudin/intermodality-analysis>