To reproduce the findings in the **FY25 Supply and Demand Study** using **Census Local Employment Dynamics (LED)/LHD employment security filings**, you'll need to assemble data cohorts that map both **childcare demand** and **supply**, adjusted for **workforce commuting patterns**, **household characteristics**, and **provider capacity**. Here's a breakdown of the **required cohorts**:

🔹 1. Population Cohorts for Demand Estimation

A. Children Under Age 5 by Census Block Group

- Source: ACS 5-Year Estimates, Claritas, or Census Population Estimates
- Fields:
    - Total children under age 5
    - Split by age groups:
        - Infants & Toddlers (0–2 years)
        - Preschool Age (3–5 years)

B. Workforce Commuter Adjustment

- Source: Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES)
- Fields:
    - Inflow and outflow of workers by census block group
    - Adjusted number of children needing care near job locations (vs residence)

C. Household Employment Characteristics

- Source: ACS Public Use Microdata Sample (PUMS), LODES
- Fields:

- % of families where all parents work
- % working non-traditional hours (before 6:30 AM / after 6 PM)
- % working remotely ≥3 days/week

D. Language and Socioeconomic Indicators

- Source: ACS
- Fields:
  - % of children in households where no parent speaks English at home
  - % of children eligible for Child Care Subsidy (CCS) or WPA based on income
  - Income-to-poverty ratio, household size

🔹 2. Childcare Supply Cohorts

A. Licensed Provider Data

- Source: State/local licensing databases (MD License DB, MSDE, DHHS-Federal)
- Fields:
  - Location (geocoded)
  - Age range served (infant, toddler, preschool)
  - Licensed capacity by age group
  - Evening/weekend care offered
  - Language services available
  - EXCELS quality rating
  - Participation in subsidy programs (WPA, CCS)

B. School-Based and Community Pre-K Programs

- Source: MCPS, MSDE, Head Start data
- Fields:
    - Total seats by program
    - Program type (school-based, community, special ed)
    - Participation in CCS/WPA

🔹 3. Subsidized Care Utilization

A. Subsidy Program Data

- Source: State subsidy administrators (e.g., WPA, CCS)
- Fields:
    - Number of enrolled children by age and location
    - % utilization of eligible capacity
    - Program eligibility by EXCELS rating or other indicators

---

🔹 4. Relative Shortage Analysis

To conduct **spatial mismatch analysis**:

- Match demand (child count adjusted for commuting and demographics) vs supply (licensed slots) by **census block group**
- Compute:
    - Absolute Gap = Demand – Supply

- Supply Ratio = Supply / Demand
- Relative Shortage = Standardized index of supply-to-demand across the county

---

🔹 5. Survey and Qualitative Cohorts (optional if replicating full scope)

While not part of the LED/Census data, the original study included:

- Parent surveys (n=249)
- Provider surveys (n=185)
- Focus groups

You can optionally replicate this with new primary research or proxy ACS/PUMS variables (e.g., hours worked, income, race/ethnicity).

✅ Summary Table of Required Cohorts

Category	Data Elements	Primary Source
Population (by age)	Children 0–2, 3–5 by census block	ACS, Claritas
Employment & Commuting	Worker inflow/outflow, work hours, remote work	LODES, ACS PUMS
Language & Income	English spoken at home, income eligibility for CCS/WPA	ACS
Provider Capacity	Capacity by age, EXCELS rating, language, hours, subsidy participation	MD License DB, MSDE, DHHS
School & Pre-K Programs	Seats in school-based and community programs	MCPS, Head Start, State agencies
Subsidy Enrollment	Seats used vs available in WPA/CCS programs	State subsidy agencies

Relative Shortage Index	Spatial supply-to-demand comparison by block group	Derived from all above sources

---

Here's a **draft data schema and field mapping** for ingesting the necessary childcare supply and demand data into a database (e.g., PostgreSQL, BigQuery, or similar). This schema is designed to align with the components used in the FY25 Supply and Demand Study and will help you structure the data request and downstream analysis.

🧩 Database Schema Overview

You'll need to build tables in the following categories:

Table Name	Purpose
`demographics_block`	Census and household characteristics by geography
`employment_commute`	Worker inflow/outflow by census block
`childcare_demand`	Modeled demand estimates by child age group
`childcare_supply`	Licensed provider and capacity data
`subsidy_programs`	WPA/CCS eligibility and utilization
`pre_k_programs`	School-based and community Pre-K programs

`shortage_index`	Derived indicators of relative and absolute shortage

📘 1. `demographics_block`

Field Name	Data Type	Description
`geo_id`	STRING	Census block group ID (e.g., GEOID)
`county_name`	STRING	County name (e.g., Montgomery County)
`children_under_2`	INTEGER	Number of children under age 2
`children_2_to_5`	INTEGER	Number of children age 2–5
`children_total_under_5`	INTEGER	Total children under age 5
`pct_non_english_home`	FLOAT	% children with parents not speaking English at home
`pct_all_parents_work`	FLOAT	% households where all parents work
`pct_parents_nontraditional_hours`	FLOAT	% parents working before 6:30am or after 6pm

`pct_remote_work`	FLOAT	% of families working remotely 3+ days/week
`pct_eligible_subsidy`	FLOAT	% of children eligible for CCS or WPA
`median_household_income`	INTEGER	Median household income
`poverty_level_ratio`	FLOAT	Income-to-poverty ratio
`data_year`	INTEGER	Year of the estimate (e.g., 2024)

📘 2. `employment_commute`

Field Name	Data Type	Description
`home_geo_id`	STRING	Census block group of residence
`work_geo_id`	STRING	Census block group of workplace
`commuter_count`	INTEGER	Number of commuters from home to work location
`net_commute_flow`	INTEGER	Net inflow or outflow of workers for given block
`data_year`	INTEGER	Year of data (e.g., 2024)

source	STRING	Data source (e.g., LEHD LODES)

📘 3. childcare_demand

Field Name	Data Type	Description
geo_id	STRING	Census block group ID
children_under_2_adj	INTEGER	Adjusted demand for children under 2 (with commute)
children_2_to_5_adj	INTEGER	Adjusted demand for children age 2–5
children_total_adj	INTEGER	Adjusted total demand for children under 5
children_non_traditional	INTEGER	Est. needing non-traditional hour care
children_non_english	INTEGER	Est. needing non-English language support
children_subsidy_eligible	INTEGER	Est. eligible for subsidy
data_year	INTEGER	Year of data

📘 4. `childcare_supply`

Field Name	Data Type	Description
`provider_id`	STRING	Unique provider ID
`provider_name`	STRING	Name of childcare provider
`geo_id`	STRING	Census block group or geolocation
`license_type`	STRING	Center-based, home-based, independent, etc.
`infant_capacity`	INTEGER	Licensed capacity for ages 0–2
`preschool_capacity`	INTEGER	Licensed capacity for ages 3–5
`total_capacity`	INTEGER	Combined licensed capacity
`excel_rating`	INTEGER	Quality rating (e.g., EXCELS 1–5)
`language_support`	BOOLEAN	Offers non-English language support
`evening_or_weekend`	BOOLEAN	Offers care after 8pm or weekends

`subsidy_acceptance`	BOOLEAN	Accepts WPA or eligible for CCS
`program_type`	STRING	Public, private, nonprofit, etc.
`data_year`	INTEGER	Year of record
`source`	STRING	Source database (MD License DB, DataAxle, etc.)

📘 5. `subsidy_programs`

Field Name	Data Type	Description
`provider_id`	STRING	Foreign key to `childcare_supply`
`program_name`	STRING	WPA or CCS
`enrolled_infants`	INTEGER	Number of infants enrolled under subsidy
`enrolled_preschool`	INTEGER	Number of preschoolers enrolled under subsidy
`total_enrolled`	INTEGER	Total enrollment
`capacity_available`	INTEGER	Seats available under subsidy

`program_eligible`	BOOLEAN	True if program is eligible but not currently enrolled
`data_year`	INTEGER	Year of data

📘 6. `pre_k_programs`

Field Name	Data Type	Description
`program_id`	STRING	Unique program ID
`program_type`	STRING	School-based, community-based, Head Start
`geo_id`	STRING	Census block group
`licensed_capacity`	INTEGER	Total seats available
`enrolled_students`	INTEGER	Students enrolled
`subsidy_eligible`	BOOLEAN	If students in program are subsidy-eligible
`data_year`	INTEGER	Year of record

📘 7. `shortage_index` (Derived Table)

Field Name	Data Type	Description
`geo_id`	STRING	Census block group
`demand_total`	INTEGER	Total adjusted demand
`supply_total`	INTEGER	Total available licensed capacity
`absolute_gap`	INTEGER	`demand_total - supply_total`
`supply_ratio`	FLOAT	`supply_total / demand_total`
`relative_shortage_score`	FLOAT	Standardized z-score or percentile relative to county average
`shortage_type`	STRING	Infant/toddler, preschool, subsidy, or quality
`data_year`	INTEGER	Year of estimate

🔧 Notes for Implementation

- Normalize values where possible (e.g., `geo_id`, `provider_id`) for joinability across tables.
- Use GIS data (point or polygon shapefiles) to visualize `geo_id` spatially.

- Build indexes on `geo_id` and `data_year` for efficient querying.
- This schema supports both point-in-time queries and longitudinal trend analysis.