

Rattacher un dénominateur de population

D'où vient un dénominateur, comment le joindre sans apparier sur un nom, et pourquoi les populations de bassins de desserte ne totalisent pas un district. Plus le taux de couverture au-dessus de 100 % et ce qu'il dit réellement.

Pierre Robentz Cassion

Leçon 5 sur 8

Jointures et restructuration des données de programme — Faire coïncider les pièces

Ce que couvre cette leçon

- Le numérateur est la moitié facile
- D'où vient réellement un dénominateur
- Joignez sur le code, jamais sur le nom
- La projection, et l'année que personne n'énonce
- Les bassins de desserte ne s'additionnent pas
- Une couverture supérieure à 100 %
- Taux pour mille, et le piège de l'annualisation
- Écrivez le dénominateur
- La suite

Le numérateur est la moitié facile

- Compter ce qu'a fait votre programme relève de la comptabilité.

D'où vient réellement un dénominateur

- **Une projection de recensement.** L'institut national de statistique publie un recensement et un taux de croissance ; . . .
- **Une base administrative de population.** Les dénominateurs du ministère de la santé par bassin de desserte, ce qu'est . . .
- **Une cible de programme.** Le nombre de personnes que le programme prévoyait d'atteindre. Dénominateur légitime pour « . . .
- **Une enquête.** La plus défendable et la moins disponible, car une enquête donne une proportion avec un intervalle de . . .
- Nommez la source dans la colonne — `target_population_moh_2024` et `target_population_census_projection` sont deux . . .

Joignez sur le code, jamais sur le nom — En Python

```
population = pd.read_csv("admin-population-2024.csv")    # admin2_pcode, year, pop_total,
                                                         pop_under1

vax["admin2_pcode"] = vax["admin2_pcode"].astype("string").str.strip().str.upper()

with_denominator = vax.merge(
    population.query("year == 2024"),
    on="admin2_pcode", how="left", validate="many_to_one",
)

missing = with_denominator["pop_under1"].isna().sum()
assert missing == 0, f"{missing} rows have no population figure"
```

Joignez sur le code, jamais sur le nom — En R

```
population <- read_csv("admin-population-2024.csv")

with_denominator <- vax |>
  mutate(admin2_pcode = toupper(stringr::str_squish(admin2_pcode))) |>
  left_join(filter(population, year == 2024),
            by = "admin2_pcode", relationship = "many-to-one")

stopifnot(!any(is.na(with_denominator$pop_under1)))
```

La projection, et l'année que personne n'énonce — En Python

```
GROWTH = 0.024    # national annual growth rate, published with the census  
CENSUS_YEAR = 2015  
  
population["pop_2024"] = population["pop_census"] * (1 + GROWTH) ** (2024 - CENSUS_YEAR)
```

La projection, et l'année que personne n'énonce — En R

```
GROWTH <- 0.024  
CENSUS_YEAR <- 2015  
  
population <- population |>  
  mutate(pop_2024 = pop_census * (1 + GROWTH)^(2024 - CENSUS_YEAR))
```

La projection, et l'année que personne n'énonce

- Écrivez l'année du recensement, le taux de croissance et sa source à côté du nombre.
- Utilisez la **même** projection pour tous les indicateurs d'un rapport. Deux indicateurs sur deux projections ne sont. . .
- Là où des déplacements ont eu lieu, dites que la projection n'en tient pas compte, et donnez le sens de l'erreur au. . .

Les bassins de desserte ne s'additionnent pas — En Python

```
by_facility = vax.query("antigen == 'penta3' and period == '2024-01-01'")  
print("sum of facility targets:", by_facility["target_population"].sum())
```

Les bassins de desserte ne s'additionnent pas — En R

```
vax |>  
  filter(antigen == "penta3", period == "2024-01-01") |>  
  summarise(total = sum(target_population))
```

Les bassins de desserte ne s'additionnent pas — En Python

```
district = (  
    vax.query("antigen == 'penta3'")  
    .groupby("period")  
    .agg(doses=("doses_administered", "sum"),  
         target=("target_population", "sum"))  
)  
district["coverage"] = district["doses"] / district["target"]
```

Les bassins de desserte ne s'additionnent pas — En R

```
district <- vax |>
  filter(antigen == "penta3") |>
  summarise(doses = sum(doses_administered),
            target = sum(target_population), .by = period) |>
  mutate(coverage = doses / target)
```

Une couverture supérieure à 100 %

- **Le dénominateur est trop petit.** Une projection périmée, ou un coefficient de bassin qui suppose moins de moins d'un. . .
- **Le numérateur inclut des gens de l'extérieur.** Traversée de limites, ou une campagne qui a attiré des gens d'un. . .
- **La granularité est fausse.** Des doses comptées là où on voulait des enfants — un enfant recevant trois doses d'un. . .
- **Une jointure a multiplié quelque chose.** Tout ce qui est dans la leçon 1.

Une couverture supérieure à 100 % — En Python

```
over = district[district["coverage"] > 1]
print(over)
```

Une couverture supérieure à 100 % — En R

```
district |> filter(coverage > 1)
```

Une couverture supérieure à 100 %

- Rapportez-le plutôt que de l'écrêter — Un taux de couverture plafonné à 100 % a jeté la preuve que le dénominateur est. . .

Taux pour mille, et le piège de l'annualisation — En Python

```
district["per_1000"] = 1000 * district["doses"] / district["target"]
```

Taux pour mille, et le piège de l'annualisation — En R

```
district <- district |> mutate(per_1000 = 1000 * doses / target)
```

Écrivez le dénominateur

Champ	Valeur
Indicateur	Couverture penta3, district
Numérateur	Doses de penta3 administrées, formations ayant rapporté
Dénominateur	Nourrissons survivants, bassins du ministère sommés au district
Source	Estimations de population du ministère 2024, projetées du recensement 2015 à 2,4 %
Désagrégation	Mois, type de formation sanitaire
Réserve	Taux de complétude de 29 % en août ; couverture calculée sur les seules formations ayant rapporté

- Le dénominateur répond à « sur combien de personnes ».

Lire la leçon complète, avec le code exécutable

[https://data-analysis.cassion.dev/fr/cours/joining-and-reshaping/
joining-05-population-denominators](https://data-analysis.cassion.dev/fr/cours/joining-and-reshaping/joining-05-population-denominators)