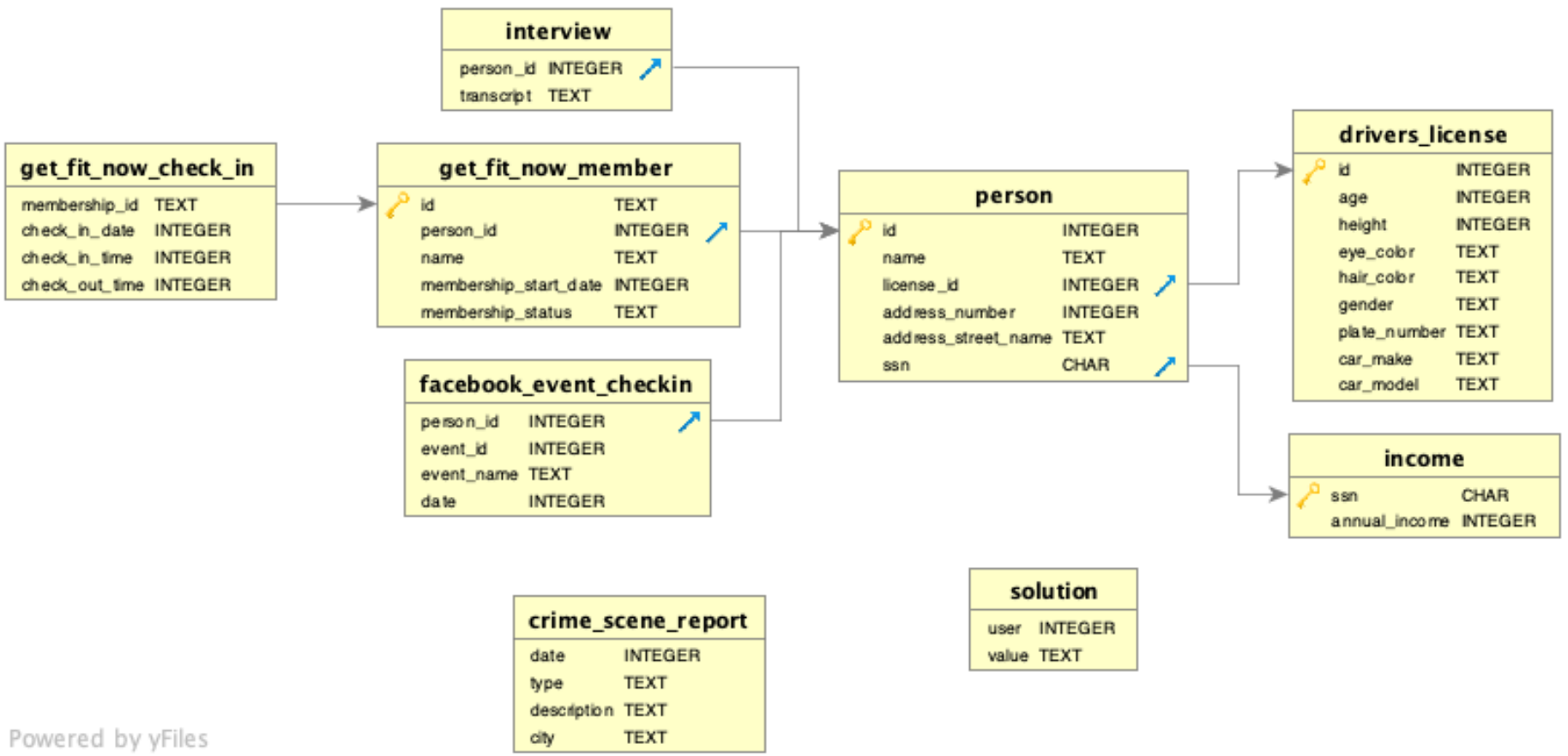
Database structure

ERD

An *entity-relationship diagram* describes the tables in a database and the relationships between them



ERD

- **Primary key:** a unique identifier for rows of a table
  - A primary key can be built from one column or from multiple columns
  - Every row must have a value for the primary key
  - No two rows can have the same value for the primary key
  - Each table can have at most one primary key
- Example: `id` in the `person` table, `id` in the `drivers_license` table

ERD

- **Foreign key:** references a primary key in another table
  - Foreign keys describe relationships between tables
  - Values of a foreign key can be repeated in the table
- Example: `person_id` in `facebook_event_checkin` table references `id` in `person` table

Maintaining referential integrity

- A foreign key cannot have values which do not appear in the primary key for the referenced table
- Example: the `person_id` column in the `facebook_event_checkin` table cannot take values which do not appear in the `id` column of the `person` table
- Deleting a row from a table with a primary key is not permitted if that value appears in a referenced foreign key

Class activity

Work on the class activity (handout).

Class activity

What columns could be primary keys in the `nycflights13` tables?

airlines: carrier

airports: faa

planes: tailnum

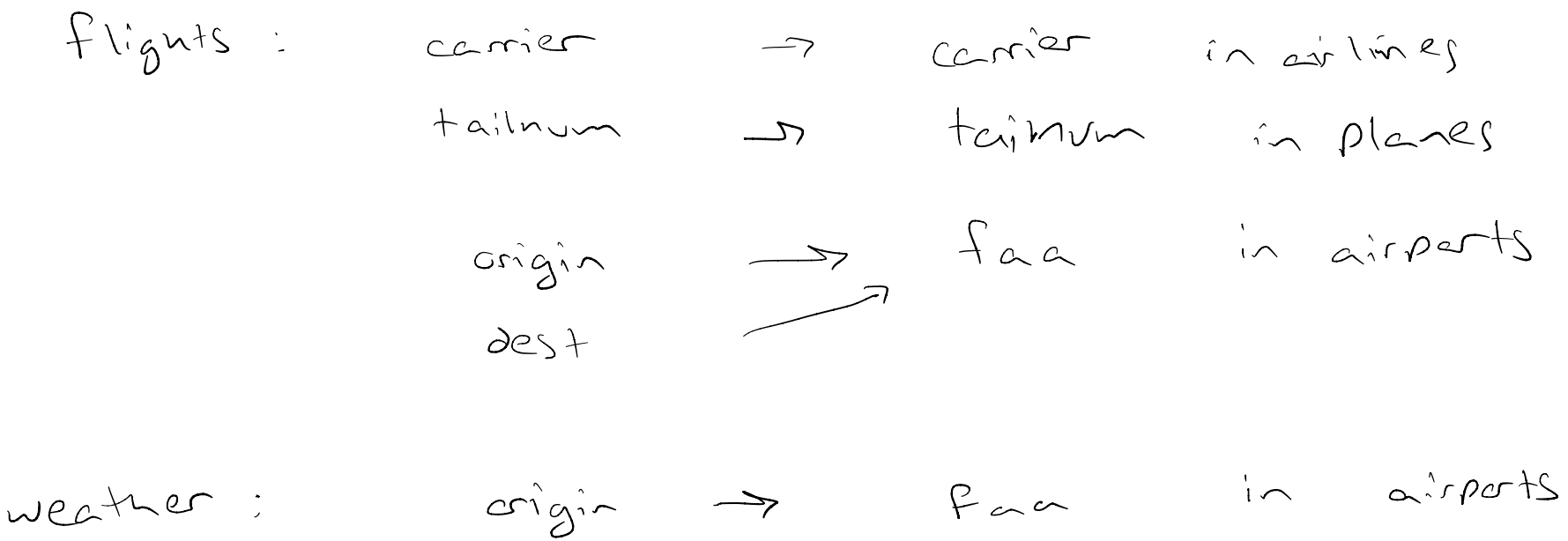
weather: origin, time_hour

flights: carrier, time_hour, flight

```
weather |>
  count(origin, time_hour) |>
  filter(n > 1) |>
  nrow()
```

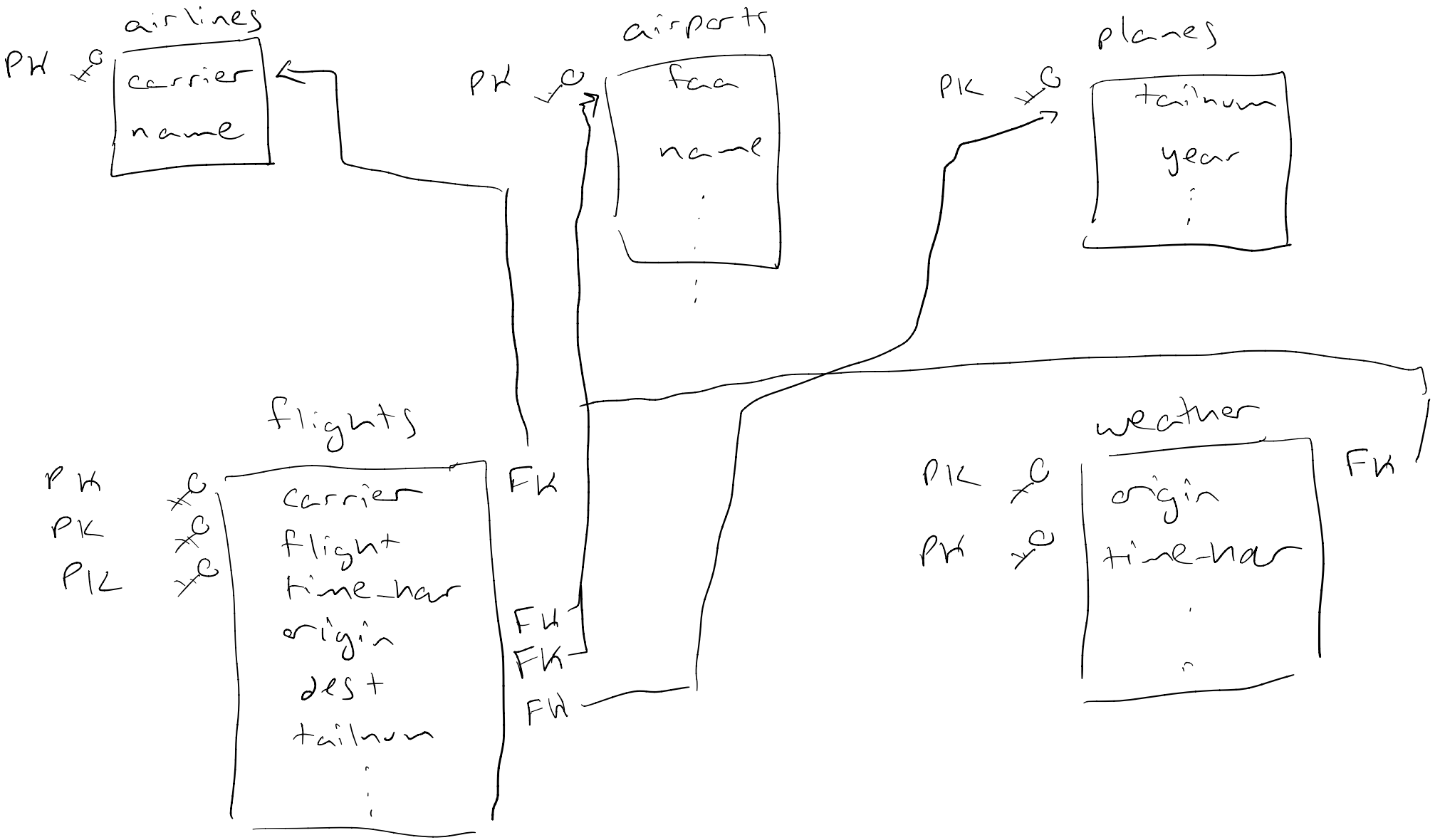
Class activity

What columns could be foreign keys in the `nycflights13` tables? What primary keys do they reference in other tables?



Class activity

ERD sketch for the `nycflights13` datasets:



Class activity, Part II

Now work on the second class activity (on course website). Render your work as an HTML and submit on Canvas when finished.