Cineville Technical Assignment

Cineville members attend movies using their Cineville card. Each card has a unique barcode and belongs to a single member. When a member attends a movie, their card is scanned and a visit is created. A visit may be linked to a reservation, but reservations are optional. Members can also attend without a reservation (walk-ins).

To better understand how our members use their Cineville card, we would like to be able to see which visits each member made. Your task is to combine the visits data and members data to provide an overview of the visits per member.

Input files

- members.csv
- visits.csv

We exported two data sets in comma separated formatting.

members.csv
member_id,barcode

- Each row represents a Cineville member and their card
- Each barcode belongs to exactly one member
- Barcodes are unique

visits.csv
visit_id,barcode,reservation_id

- Each row represents one visit
- A visit is linked to a barcode
- reservation_id may be empty for walk-ins

Assignment

Write a program that reads the two files, members.csv and visits.csv, and generates an output CSV that groups visits per member and barcode.

The output file should have the following format:

member_id, barcode, [visit_id1, visit_id2, ...]

Each row represents a member, their Cineville card barcode, and all visits made using that card.

Validation Rules

- A visit must have a barcode
- A visit must reference a known barcode

Invalid data should be logged and not included in the output.

Bonus

- Print the top 5 members by number of visits. Each line should be printed to stdout in the following format: member_id, amount_of_visits.
- Print the total number of walk-ins (visits without a reservation).

Requirements

- Implement the solution in Python.
- Deliver the solution via GitHub or as a zip file.
- The solution should be production-ready, including:
    - Basic documentation
    - Instructions on how to run the code
    - Tests (if you think they add value)