0. Welcome (10 min)

Brief outline

Troubleshoot installing Unix

1. GEO Basics (20 min)
- Accession types: GSE/GSM (processed/metadata), SRR (raw reads)
- Common files: Series Matrix, SOFT, fastq.gz, sample TSV
- <u>Exercise 1.1</u>: Download a file
- FTP and Data Download (15 min)
    - What FTP is and why GEO use it
    - Command-line for downloading data
- <u>Exercise 1.2</u>: Download the dataset with FTP

2. UNIX Basics (30 min)

- File management and navigation (mkdir, rm, mv, pwd, ls, cd).
- File inspection (head, tail, less, wc).
- Pipes and redirects (|, >, >>).
- Text search (grep, cut, sort, uniq).
- Exercise 2: Reviewing material

Exercise 3.1 — Counts CSV (10 min)

Preview first/last lines → count rows (no header) → print header with indices → extract gene_id + two samples → sort descending by one sample (keep header) → simple unique group counts

Exercise 3.2 — FASTQ (15 min)

Browser download (small FASTQ) → count reads (^@) → 5′ 6-mers (top-10) → motif search (e.g., ACGTAC) → count reads with N → take first 100 reads safely (400 lines)

6. Conclusion (10 min)

Conclusion of today's session
Troubleshoot installing R