

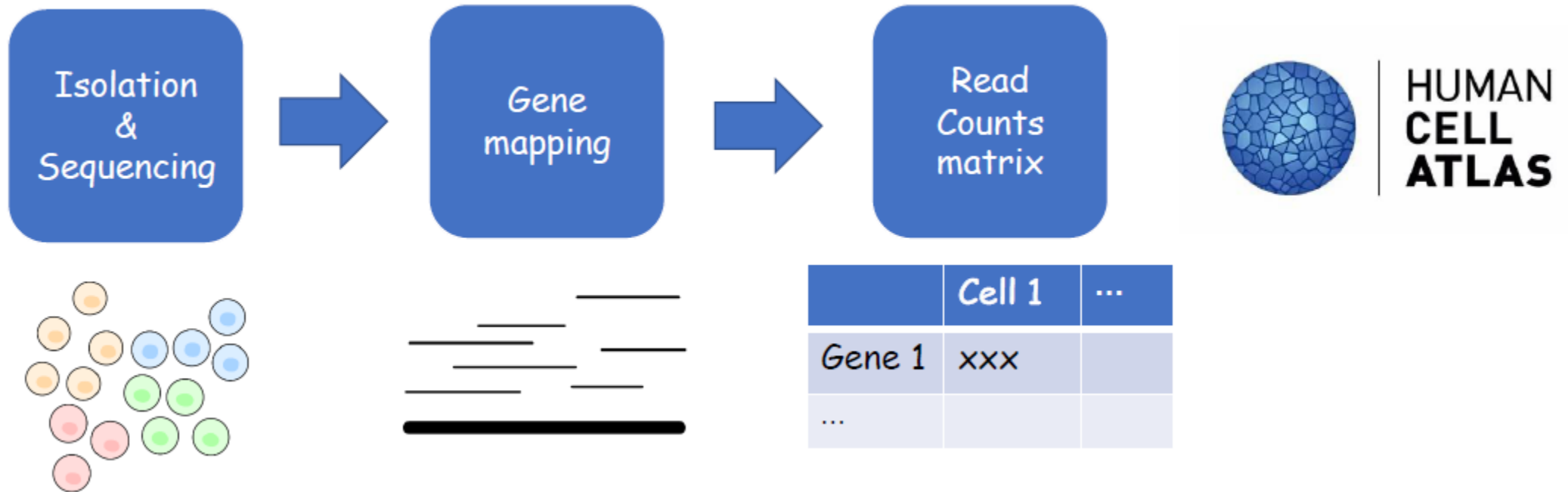


上海科技大学
ShanghaiTech University

SI100B Python final Project – Single-cell RNA sequencing

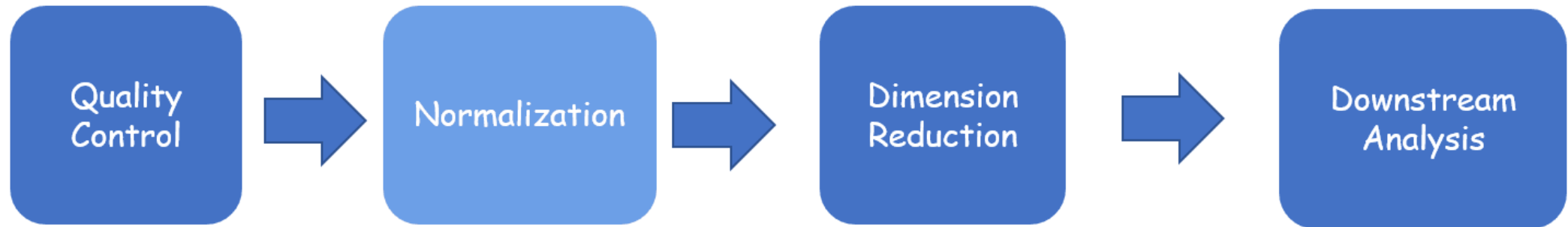
Background

- Single-cell RNA sequencing (scRNA-seq) is an emerging technique to assess individual cell function and cell-to-cell variability at the single cell level.

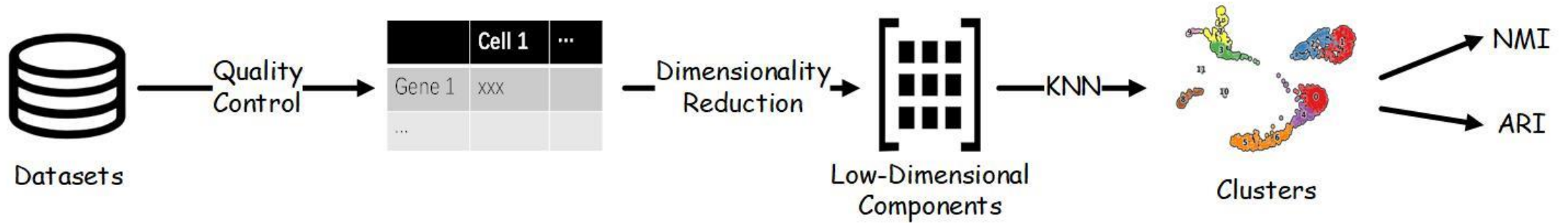


Background

- Widely used scRNA-seq protocols include Quartz-seq2, Chromium, Smart-seq2, CEL-seq2 Drop-seq, inDrop, MARS-seq and so on.
- scRNA data are always noisy
 - low expression counts
 - Sparse expression matrix due to drop-out event



Benchmark Methods



Tasks

- **Week 1 15pts**
 - Get familiar with git
- **Week 2 30pts**
 - Read the datasets (15pts)
 - Quality Control (15pts)
- **Week 3 30pts**
 - Dimensionality reduction (sklearn) PCA, tSNE (10pts)
 - Clustering + Visualization (sklearn) (10 pts)
 - Compare the performance of different dimensionality reduction methods using several metrics (NMI, ARI).(10pts)
- **Week 4 25pts**
 - Automate the whole pipeline (5pts)
 - Chase for better performance! (10pts)
 - Submit your final report! (10pts)

Brief reports, no more than 2 pages: What you have learned + How you implemented these tasks.

When & How

- We will distribute each part of the project by the end of week 12, 13, 14, 15.
 - We will distribute them as soon as we can, probably much earlier so that you can have enough time.
- Your work is collected at the end of week 16 (Sunday, 2nd Jan, 11:59pm).
- The offline check will be held afterwards.
- Your grade is **only** made of the offline check score, online judge does not count.

Work as a team

- You may team up as a 3-persons team.
 - You will get the same score if the contributions are fair.
 - Please mark your contributions in the final report.
- 17 groups in total
- Room 1D-107, SIST

Good Luck & Have Fun!

Special Thanks to Mr. **Qi QIN** for contributing the project!