# Interpreting Cell Types and Biological Context in a Single-Cell Dataset

The answers provided in this file were derived based on the analysis performed in `stage_two_task.ipynb`.

1. What cell types did you identify?
List each annotated cluster. No need for a novel classification system. Just show you understand what you're looking at.

Output:

| Cluster | Annotated Cell Type |
|---|---|
| 0 | Monocytes |
| 1 | B cells naive |
| 2 | T cells |
| 3 | Plasma cells / Plasmablasts |

2. Explain the biological role of each cell type
For every annotated label, give a short explanation of what that cell type actually does in bone marrow or peripheral immunity.

Output:

- Monocytes: Innate immune cells that phagocytose pathogens, present antigens, and produce inflammatory cytokines. Key players in early immune responses.

- B cells naive: Adaptive immune cells that recognize antigens and differentiate into plasma cells upon activation. Responsible for producing antibodies.

- T cells: Adaptive immune cells that coordinate immune responses. Include helper T cells (activate B cells and other immune cells) and cytotoxic T cells (kill infected or abnormal cells).

- Plasma cells / Plasmablasts: Differentiated B cells that actively produce and secrete antibodies. Expansion is a hallmark of recent antigen exposure (infection or vaccination).

3. Is the tissue source really bone marrow? Justify your answer. Your job is to reason your way toward (or away from) that conclusion using:

- expected vs. missing lineage populations

- typical frequency distributions
- presence or absence of progenitors

If you claim bone marrow, explain the flaws in your logic. Otherwise, justify it biologically. Hand-waving is a fail.

Output:
Evidence for bone marrow:

- Expected presence of hematopoietic progenitors (HSPCs), erythroid, megakaryocytes, and neutrophils.

Observation: The dataset shows no HSPCs, erythroid, or megakaryocytes. Only mature peripheral immune cells (T cells, B cells, monocytes, plasma cells) are present.

Interpretation: The sample behaves like peripheral blood mononuclear cells (PBMCs) rather than bone marrow. The absence of progenitors and neutrophils suggests this is not a true bone marrow sample.

4. Based on the relative abundance of cell types, is the patient healthy or infected? Use the cluster proportions to make a call. Your job: defend your conclusion using deviations in:

- neutrophils
- monocytes
- NK cell activation states
- lymphocyte depletion or expansion

Do not just guess. Interpret the landscape like a scientist.

Output:

1. Cluster proportions: T cells ~73%, Monocytes ~13%, B cells naive ~7%, Plasma cells ~6%.
2. Interpretation: Elevated plasma cells (6%) and slightly elevated monocytes (~13%) indicate recent immune activation, likely due to infection or recent vaccination. Also, Neutrophils are absent (expected for PBMC prep) and do not inform infection status, and NK cells were not explicitly analyzed; their activation could provide further confirmation.

Conclusion: Patient shows signs of an active immune response, not a completely healthy baseline.