Parth Naik

**Fair assignment**

1. 1. Are you aware that a data(set) should be assigned a globally unique persistent and resolvable identifier when deposited with a data repository?

Ans: The GEO dataset has been assigned a unique and persistent identifier, GSE276609, fulfilling the requirement for it to be ‘findable’.

1. Are you aware that when you deposit a data(set) in a data repository, you will need to provide discovery metadata in order to make the data(set) findable, understandable and reusable to others?

Ans: Additional ‘data’ within the GEO dataset are referenced to explain its overall composition such as sample title and project.

1. Are you aware that the data repository providing access to your data(set) should make the metadata describing your data(set) available in a format readable by machines as well as humans?

Ans: The processed dataset is provided in TSV and MTX format.

1. Are you aware that access to your data(set) may need to be controlled and that metadata should include licence information under which the data(set) can be reused?

Ans: Since the dataset hasn't been published yet, it remains private, and no licensing information for reuse has been observed.

1. Are you aware that metadata should remain available over time, even if the data(set) is no longer accessible?

Ans: Currently, the GEO dataset is unavailable to the public, making it inaccessible. If published metadata will be available.

1. Are you aware that the metadata describing your data(set) should use controlled vocabularies?

Ans: The dataset employs controlled vocabularies, such as the term Homo sapiens, breast cancer, expression profiling .

1. Are you aware that provenance information about the collection and/or generation of data should be included in the metadata?

Ans: This dataset references previously analyzed datasets, and it includes the platforms used for generating or collecting the data.

1. Are you aware that metadata describing your data(set) should follow the specifications of a community-endorsed standard?

Ans: I am unfamiliar with the community standards associated with this dataset, so I cannot determine if it complies with them.

1. Are you aware that your data(set) should be deposited preferably in a file format that is open and supported by the data repository for long-term preservation?

Ans: The dataset is presented in fastq, run.zq, sra, which is an open format supported by the repository for long-term preservation.

1. Are you aware that keeping your data(set) FAIR over time requires professional data curation and digital preservation?

Ans: Data integrity, Metadata curation are few points of many that I can see in this experiment.