GANTT CHART

PROJECT TITLE	Team Arch Software Development Project	**MODULE**	BIO727P
MEMBERS	Roxana Bosnea, Farzard Hamzawe, Cheau Lee & Siril Sunil	**DATE**	23/01/24

STAGE	TASK TITLE	TASK OWNERS	START DATE	DUE DATE	DURATION (DAYS)	PROGRESS (%)
1	**Data Acquisition & Parsing**					
1.1	Download Data: Retrieve the sample genetic data from the VCF file .	ALL	24/01/24	24/01/24	1	100
1.2	Annotation Data: Retrieve data for human genes and SNPs. Find extra information on the samples from 1000 Genomes Project and clinical revelance of SNPs.	CL	24/01/24	28/01/24	1	100
1.3	Data Parsing and Cleaning: Process the VCF file, extract relevant data, and verify the integrity of the data to ensure compatbility with the project requirements.	CL + FZ	24/01/24	27/01/24	1	100
1.4	Data Loading, Transformation and Design Schema: Load the cleaned data into the SQL database. Design pre-computational tools for data access, filtering, and data storage for downstream analyses.	RB + CL + FZ	29/01/24	30/01/24	1	100
2	**Core Analysis Functions**					
2.1	Choose Clustering Algorithm(s): Select suitable clustering algorithms like Principal Component Analysis (PCA), Multidimensional Scaling (MDS), or Uniform Manifold Approximation and Projection (UMAP) based on data characteristics and desired output. Justify your choices in the documentation.	RB	02/02/2024	09/02/24	5	100
2.2	Implement Clustering Functionality: Develop code to perform the chosen clustering algorithm(s) on the selected population data. Allow user input to choose which populations or superpopulations to include in the analysis.	RB	02/02/2024	09/02/24	5	100
2.3	Visualise Clustering Results: Generate interactive visualisations to represent the clustered data and relationships between populations.	RB	02/02/2024	09/02/24	5	100
2.4	Choose Admixture Algorithm(s): Select suitable admixture algorithms like ADMIXTURE or STRUCTURE based on data characteristics and desired outputs. You can compare different algorithms and justify your choice based on accuracy and interpretability.	FZ	06/02/24	15/02/24	10	100
2.5	Implement Admixture Functionality: Develop code to perform the chosen admixture algorithm(s) on the selected population data. Allow user input to choose which populations or superpopulations to include in the analysis.	FZ	06/02/24	15/02/24	5	100
2.6	Visualize Admixture Results: Generate interactive visualisations of individual and population admixture proportions.	FZ	06/02/24	15/02/24	1	100
3	**Frontend Development**					
3.1	App Set-Up: Create an application using a front-end framework.	RB + CL	09/02/24	16/02/24	5	100
3.2	Project structure: Develop components for clustering and admixture analysis.	RB + FZ	09/02/24	16/02/24	5	100
3.3	Backend: Integrate these components with backend AP for data retrieval and analysis	RB + FZ	09/02/24	16/02/24	5	100
3.4	Develop User Interface: Create a user-friendly interface for users to input SNP IDs, genomic coordinates, or gene names for retrieving specific data.	CL	09/02/24	16/02/24	5	100
3.5	Data Retrieval Functionality: Implement code to efficiently query the database and extract relevant information (allele frequencies, genotypes, clinical annotations) based on user input.	RB + CL	09/02/24	16/02/24	5	100
3.6	Population Genetic Differentiation: Implement calculation component for pairwise population genetic differentiation.	RB + CL + SS	09/02/24	16/02/24	5	100
3.7	Data Visualisation: Present the retrieved differentiation matrix data in an interactive manner, and provide download options.	RB + CL + SS	09/02/24	16/02/24	5	100
4	**Web Testing, Deployment, and Documentation**					
4.1	Internal code: Add comprehensive comments to explain logic and functionalities.	ALL	22/02/24	29/02/24	4	100
4.2	Documentation: Create a detailed document explaining the code structure, API usage, and algorithms implemented.	ALL	22/02/24	29/02/24	2	100
4.3	Unit and Integration Testing: Conduct testing of each functionality, addressing any bugs or issues before deployment.	ALL	22/02/24	29/02/24	1	100
4.4	Final Refinements: Track user feedback and monitor application performance.	ALL	22/02/24	29/02/24	1	100
4.5	GitHub: Keep track and finalise upload of the repositories for code, data and documentation.	ALL	26/02/24	29/02/24	4	100