Python for Genomic Data Science Training Course

Professional Development Training Course Outline

Category	Research and Data Analysis	**Duration**	10 Days
Base Fee	$3,000 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

The explosion of genomic data in modern biology has revolutionized healthcare, biotechnology, and personalized medicine. Harnessing this wealth of information requires robust computational skills, with Python programming emerging as the cornerstone of genomic data analysis. This training program equips participants with practical expertise in bioinformatics, genomic data processing, statistical modeling, machine learning, and visualization using Python. Python for Genomic Data Science Training Course is Designed for both beginners and experienced professionals, this course transforms raw genomic sequences into actionable insights, empowering participants to tackle complex datasets efficiently.

By integrating hands-on coding, real-world case studies, and advanced computational tools, this course ensures learners master the entire pipeline of genomic data analysis—from preprocessing and quality control to variant calling and multi-omics integration. Participants will gain proficiency in Python libraries such as Pandas, NumPy, Biopython, SciPy, Scikit-learn, and Matplotlib, enabling them to extract meaningful patterns from large-scale genomic datasets. Upon completion, learners will possess the skills to accelerate precision medicine initiatives, genomic research, and data-driven biological discoveries, positioning them at the forefront of the genomic data science revolution.

Programme Curriculum

Python for Genomic Data Science Training Course

Introduction

The explosion of genomic data in modern biology has revolutionized healthcare, biotechnology, and personalized medicine. Harnessing this wealth of information requires robust computational skills, with Python programming emerging as the cornerstone of genomic data analysis. This training program equips participants with practical expertise in bioinformatics, genomic data processing, statistical modeling, machine learning, and

visualization using Python. Python for Genomic Data Science Training Course is Designed for both beginners and experienced professionals, this course transforms raw genomic sequences into actionable insights, empowering participants to tackle complex datasets efficiently.

By integrating hands-on coding, real-world case studies, and advanced computational tools, this course ensures learners master the entire pipeline of genomic data analysis—from preprocessing and quality control to variant calling and multi-omics integration. Participants will gain proficiency in Python libraries such as Pandas, NumPy, Biopython, SciPy, Scikit-learn, and Matplotlib, enabling them to extract meaningful patterns from large-scale genomic datasets. Upon completion, learners will possess the skills to accelerate precision medicine initiatives, genomic research, and data-driven biological discoveries, positioning them at the forefront of the genomic data science revolution.

Course Duration

10 days

Course Objectives

1. Master Python programming for bioinformatics and genomics.
2. Analyze large-scale genomic datasets using Python libraries.
3. Perform data preprocessing and quality control on sequencing data.
4. Implement statistical analysis for genomic variation studies.
5. Develop machine learning models for genomics and predictive biology.
6. Visualize complex genomic data using Matplotlib, Seaborn, and Plotly.
7. Apply Biopython for sequence alignment and genome annotation.
8. Conduct variant calling and SNP analysis on next-generation sequencing data.
9. Integrate multi-omics datasets for comprehensive biological insights.
10. Explore genomic databases like NCBI, Ensembl, and UCSC Genome Browser.
11. Automate bioinformatics workflows with Python scripting.
12. Implement case studies in precision medicine and cancer genomics.
13. Prepare for real-world genomic data challenges in research and industry.

Target Audience

1. Bioinformaticians
2. Genomic researchers
3. Data scientists
4. Computational biologists
5. Healthcare and biotech professionals
6. Graduate students in genomics or biology
7. Python programmers seeking bioinformatics applications
8. Pharmaceutical researchers

Course Modules

Module 1: Introduction to Python for Genomics

- Python fundamentals: syntax, variables, data types
- Control structures and functions
- Jupyter Notebook for genomic data analysis
- Introduction to Python libraries for bioinformatics

- **Case Study:** Simple DNA sequence manipulation

Module 2: Genomic Data Handling and Preprocessing

- Reading and writing genomic files
- Data cleaning and preprocessing
- Handling missing and noisy data
- Sequence filtering and trimming
- **Case Study:** Preprocessing raw sequencing reads

Module 3: Biopython for Genomic Analysis

- Sequence manipulation
- Reading and writing sequence files
- Feature extraction from genomes
- GenBank and FASTA integration
- **Case Study:** Annotating a bacterial genome

Module 4: Statistical Analysis in Genomics

- Introduction to NumPy and SciPy
- Descriptive statistics for genomic data
- Hypothesis testing in sequence data
- Correlation and regression analysis
- **Case Study:** Gene expression correlation study

Module 5: Data Visualization for Genomics

- Plotting genomic data with Matplotlib and Seaborn
- Visualizing variant frequencies and distributions
- Heatmaps and cluster visualization
- Interactive plots with Plotly
- **Case Study:** Visualizing RNA-seq expression profiles

Module 6: Next-Generation Sequencing (NGS) Data Analysis

- NGS overview: RNA-seq, DNA-seq, ChIP-seq
- Sequence alignment with Python
- Quality control using Python scripts
- Coverage and depth analysis
- **Case Study:** RNA-seq data preprocessing and QC

Module 7: Variant Calling and SNP Analysis

- Introduction to variant calling
- SNP identification using Python
- Functional annotation of variants
- Filtering and prioritization
- **Case Study:** SNP discovery in human exome sequencing

Module 8: Genomic Databases and APIs

- Accessing NCBI, Ensembl, UCSC Genome Browser

- Using Python to query databases
- Retrieving genomic annotations
- Integrating multiple databases
- **Case Study:** Fetching gene annotations via Ensembl API

Module 9: Machine Learning for Genomics

- Supervised and unsupervised learning
- Feature selection for genomic datasets
- Predicting gene function using ML
- Model evaluation metrics
- **Case Study:** Predicting cancer-related genes using ML

Module 10: Multi-Omics Data Integration

- Introduction to transcriptomics, proteomics, epigenomics
- Data normalization and scaling
- Integrating omics datasets with Python
- Correlation and network analysis
- **Case Study:** Multi-omics integration in cancer research

Module 11: Functional Genomics and Pathway Analysis

- Gene ontology enrichment
- Pathway mapping and visualization
- Network analysis of genes and proteins
- Functional annotation using Python tools
- **Case Study:** Pathway analysis in metabolic disorders

Module 12: Automation and Workflow Management

- Python scripting for repetitive tasks
- Creating reproducible pipelines
- Workflow automation with Snakemake
- Logging and debugging
- **Case Study:** Automating variant calling pipeline

Module 13: Genomic Data Security and Ethics

- Data privacy and ethical considerations
- Secure handling of genomic datasets
- GDPR and HIPAA compliance
- Sharing data responsibly
- **Case Study:** Ethical handling of patient genomic data

Module 14: Advanced Visualization and Dashboards

- Interactive genomic dashboards with Dash
- Real-time data visualization
- Custom plotting functions for genomic data
- Reporting results effectively
- **Case Study:** Interactive cancer genomics dashboard

Module 15: Capstone Project

- Real-world genomic data analysis project
- End-to-end pipeline: preprocessing to visualization
- Machine learning or multi-omics integration
- Presentation and interpretation of results
- **Case Study:** Personalized medicine genomics pipeline

Training Methodology

This course employs a participatory and hands-on approach to ensure practical learning, including:

- Interactive lectures and presentations.
- Group discussions and brainstorming sessions.
- Hands-on exercises using real-world datasets.
- Role-playing and scenario-based simulations.
- Analysis of case studies to bridge theory and practice.
- Peer-to-peer learning and networking.
- Expert-led Q&A sessions.
- Continuous feedback and personalized guidance.

Register as a group from 3 participants for a Discount

Send us an email: info@fineskilltrainingcenter.org or call *+254769199797*

Certification

Upon successful completion of this training, participants will be issued with a globally- recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

a. The participant must be conversant with English.

b. Upon completion of training the participant will be issued with an Authorized Training Certificate

c. Course duration is flexible and the contents can be modified to fit any number of days.

d. The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

e. One-year post-training support Consultation and Coaching provided after the course.

f. Payment should be done at least a week before commence of the training, to Fineskill Training Center account, as indicated in the invoice so as to enable us prepare better for you.

Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	04 Dec 2026	Online	$2,000
23 Nov 2026	04 Dec 2026	Physical	$3,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$9,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$12,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com