# Certificate in Financial Data Analytics

## Syllabus

**Module 1: The Data-Driven Financial Services Landscape**

- Role of data and analytics in financial services.
- Data across banking, capital markets, asset management, wealth, insurance and fintech.
- Evolution from manual reporting to automated, real-time analytics.
- Descriptive, diagnostic, predictive and prescriptive analytics.
- Data roles: analyst, data scientist, data engineer, quant and analytics translator.
- The financial data value chain: sourcing, processing, analysis and decision-making.
- Indian financial data ecosystem: regulators, exchanges, depositories, bureaus and data vendors.
- Digital public infrastructure: Aadhaar, UPI, Account Aggregator and DigiLocker.
- Global data and analytics trends in finance.
- The value proposition of a financial data strategist.

**Module 2: Financial Markets and Market Data**

- Structure of equity, debt, currency and commodity markets.
- Exchanges, clearing corporations and depositories.
- Market data types: prices, volumes, order books and corporate actions.
- Tick, intraday and end-of-day data.
- Index construction and index data.
- Mutual fund, PMS and AIF data.

- Fixed income data: yields, curves, spreads and ratings.
- Reference data, security master and identifiers: ISIN, LEI and ticker.
- Data vendors and terminals.
- Exchange data feeds and APIs.
- Data licensing, cost and usage rights.
- Survivorship bias and look-ahead bias in market data.

**Module 3: Financial Statements and Corporate Data**

- Balance sheet, income statement and cash-flow statement.
- Ind AS and IFRS reporting basics.
- Notes to accounts and related-party disclosures.
- Ratio analysis: profitability, liquidity, leverage and efficiency.
- Segment and consolidated reporting.
- Corporate filings: MCA, stock-exchange disclosures and annual reports.
- XBRL and machine-readable financial filings.
- Shareholding patterns, promoter pledges and insider trading data.
- Earnings quality and red flags.
- Standardising company data for comparison.
- Building company and sector databases.

**Module 4: Data Governance, Quality and Management**

- Data governance principles and frameworks.
- Data ownership, stewardship and accountability.
- Data quality dimensions: accuracy, completeness, timeliness, consistency and validity.
- Data profiling and cleansing.

- Master data and reference data management.
- Metadata, data catalogues and data lineage.
- Data classification and access control.
- Data retention and archival.
- Data quality metrics and issue management.
- Data governance operating models.
- Data governance case studies in banks and asset managers.

**Module 5: Data Architecture and Engineering**

- Data architecture for financial institutions.
- Relational databases and data modelling.
- Data warehouses, data lakes and lakehouses.
- Batch and real-time data processing.
- ETL and ELT pipelines.
- APIs, web services and data integration.
- Cloud data platforms.
- Big data technologies.
- Data pipelines and workflow orchestration.
- Data storage costs and performance.
- Core banking, trading and portfolio systems as data sources.
- Building a single customer view.

**Module 6: Spreadsheet Analytics and Financial Modelling**

- Advanced spreadsheet functions for finance.
- Lookup, logical, text and date functions.

- Pivot tables and data summarisation.
- Data validation and error checking.
- Financial functions: NPV, IRR, XIRR, PMT and yield.
- Scenario, sensitivity and goal-seek analysis.
- Three-statement financial models.
- Model structure, documentation and best practices.
- Spreadsheet automation.
- Spreadsheet risk and controls.
- When to move from spreadsheets to code.

**Module 7: SQL and Database Querying**

- Relational database concepts.
- SELECT, filtering and sorting.
- Aggregations and grouping.
- Joins across financial datasets.
- Subqueries and common table expressions.
- Window functions for time series and rankings.
- Date and time handling.
- Query performance and indexing.
- Building analytical views and reporting tables.
- SQL for portfolio, transaction and customer data.
- Data quality checks in SQL.

**Module 8: Python for Finance**

- Python fundamentals for analysts.

- Data structures, functions and control flow.
- NumPy for numerical computing.
- pandas for data manipulation.
- Working with time series in pandas.
- Importing data from files, databases and APIs.
- Cleaning, merging and reshaping data.
- Financial calculations in Python.
- Visualisation with Matplotlib and Plotly.
- Notebooks, scripts and reproducible analysis.
- Version control with Git.
- Writing clean, tested and documented code.

**Module 9: Statistics for Finance**

- Descriptive statistics.
- Probability concepts and distributions.
- Normal, lognormal and fat-tailed distributions.
- Sampling and the central limit theorem.
- Confidence intervals.
- Hypothesis testing.
- Correlation and covariance.
- Simple and multiple linear regression.
- Regression diagnostics.
- Outliers and robust statistics.
- Bootstrapping and resampling.
- Monte Carlo simulation.

- Common statistical pitfalls in finance.

**Module 10: Time Series and Econometrics**

- Characteristics of financial time series.
- Returns: simple, log and cumulative.
- Stationarity and unit-root tests.
- Autocorrelation and partial autocorrelation.
- ARIMA models.
- Volatility modelling: ARCH and GARCH.
- Cointegration and pairs relationships.
- Regime-switching models.
- Forecasting and forecast evaluation.
- Panel data methods.
- Macroeconomic data and nowcasting.
- Event studies.

**Module 11: Data Visualisation and Business Intelligence**

- Principles of effective data visualisation.
- Choosing the right chart.
- Visualising financial data: performance, risk, flows and exposures.
- Dashboard design.
- Business intelligence tools: Power BI, Tableau and others.
- Data models for dashboards.
- KPIs and metrics for financial institutions.
- Interactive and self-service analytics.

- Data storytelling.
- Presenting analytics to boards, investment committees and clients.
- Avoiding misleading visualisations.

**Module 12: Portfolio and Investment Analytics**

- Return calculations: time-weighted and money-weighted.
- Risk measures: volatility, beta, tracking error and drawdown.
- Risk-adjusted performance ratios.
- Performance attribution.
- Factor models and factor exposures.
- Style analysis.
- Portfolio optimisation: mean-variance and its alternatives.
- Backtesting investment strategies.
- Avoiding overfitting and data-snooping.
- Fund and manager analytics.
- Holdings-based and returns-based analysis.
- Investment reporting and client analytics.

**Module 13: Risk Analytics**

- Risk data and risk taxonomy.
- Market risk measurement.
- Value at risk: historical, parametric and Monte Carlo.
- Expected shortfall.
- Stress testing and scenario analysis.
- Liquidity risk metrics.

- Interest-rate risk in the banking book.
- Counterparty credit risk measurement.
- Operational risk data and loss analysis.
- Risk aggregation and risk data reporting.
- Risk dashboards and early-warning indicators.
- Climate and ESG risk data.

**Module 14: Credit Analytics and Scoring**

- Credit data: bureaus, bank statements, GST and alternative data.
- Credit bureaus in India.
- Application, behaviour and collection scorecards.
- Probability of default, loss given default and exposure at default.
- Logistic regression for credit scoring.
- Machine learning in credit decisions.
- Scorecard development and validation.
- Reject inference.
- Credit portfolio monitoring.
- Expected credit loss modelling under Ind AS 109.
- Early-warning systems.
- Cash-flow based lending and the Account Aggregator framework.
- Fairness and explainability in credit models.

**Module 15: Customer and Business Analytics**

- Customer data in banks, wealth firms and insurers.
- Customer segmentation.

- Customer lifetime value.
- Churn prediction and retention analytics.
- Cross-sell and next-best-action models.
- Campaign analytics and attribution.
- Pricing analytics.
- Branch, channel and distribution analytics.
- Relationship manager productivity analytics.
- A/B testing and experimentation.
- Product and digital analytics.
- Building a customer analytics roadmap.

**Module 16: Machine Learning Foundations**

- Machine learning concepts and workflow.
- Supervised and unsupervised learning.
- Feature engineering for financial data.
- Training, validation and test sets.
- Cross-validation for time series.
- Regression and classification algorithms.
- Decision trees, random forests and gradient boosting.
- Clustering and dimensionality reduction.
- Model evaluation metrics.
- Overfitting, regularisation and bias-variance trade-off.
- Imbalanced data.
- Machine learning libraries in Python.

**Module 17: Advanced Machine Learning and Deep Learning**

- Neural networks.
- Deep learning for time series.
- Recurrent networks, LSTMs and transformers.
- Anomaly detection.
- Reinforcement learning in finance.
- Graph analytics and network models.
- Ensemble methods.
- Hyperparameter tuning.
- Model deployment and MLOps.
- Monitoring models in production.
- Model drift and retraining.

**Module 18: Text Analytics, NLP and Generative AI**

- Unstructured data in finance.
- Text preprocessing and representation.
- Sentiment analysis of news, filings and earnings calls.
- Named entity recognition and information extraction.
- Topic modelling.
- Document intelligence and OCR.
- Large language models: concepts and capabilities.
- Prompt design for financial tasks.
- Retrieval-augmented generation.
- AI assistants and copilots for analysts and advisers.
- Hallucination, accuracy and verification.
- Generative AI governance.

**Module 19: Alternative Data**

- Definition and types of alternative data.
- Transaction, card and payment data.
- Web, app and social-media data.
- Satellite, geolocation and supply-chain data.
- GST, e-way bill and trade data.
- Web data collection: methods and legal boundaries.
- Evaluating alternative datasets.
- Integrating alternative data into investment and credit decisions.
- Privacy, consent and compliance considerations.
- Case studies in alternative data.

**Module 20: Fraud, AML and Financial Crime Analytics**

- Financial crime typologies.
- Transaction monitoring.
- Rules-based and model-based detection.
- Anomaly detection for fraud.
- Network analytics for fraud rings and money laundering.
- KYC and customer risk scoring.
- Sanctions and adverse-media screening.
- Digital payment and UPI fraud.
- False positives and alert optimisation.
- Reporting to FIU-IND.
- Case management and investigation analytics.

**Module 21: Regulation, Reporting and Data Protection**

- Regulatory reporting in banking, capital markets and insurance.
- RBI, SEBI, IRDAI and PFRDA data and reporting requirements.
- Supervisory data and automated reporting.
- RegTech and SupTech.
- Digital Personal Data Protection Act.
- Consent management and data principal rights.
- Data localisation requirements.
- Cybersecurity frameworks for regulated entities.
- Data breach response and notification.
- Cross-border data transfers.
- Global frameworks: GDPR and BCBS 239.

**Module 22: Model Risk Management and Responsible AI**

- Model risk and model risk management frameworks.
- Model inventory and tiering.
- Model development standards and documentation.
- Independent model validation.
- Backtesting and benchmarking of models.
- Explainability: SHAP, LIME and interpretable models.
- Bias, fairness and discrimination in models.
- Ethical use of data and AI.
- Human oversight and accountability.
- AI governance frameworks in India and globally.

- Responsible AI policies for financial institutions.

**Module 23: Data Strategy and Analytics Leadership**

- Building a data strategy.
- Aligning analytics with business objectives.
- Identifying and prioritising use cases.
- Measuring return on analytics investment.
- Build, buy and partner decisions.
- Building and leading analytics teams.
- Data culture and data literacy.
- Working with business, technology, risk and compliance teams.
- Change management and adoption.
- Vendor and technology selection.
- Communicating with senior management and boards.
- Case studies in data transformation.

**Module 24: Emerging Trends in Financial Data**

- Open finance and data sharing.
- Account Aggregator ecosystem growth.
- Real-time analytics and streaming data.
- Tokenisation and blockchain data.
- Synthetic data.
- Privacy-enhancing technologies.
- Agentic AI in financial services.
- Quantum computing in finance.

- ESG and climate data.
- Data marketplaces and data monetisation.
- Regulatory trends in data and AI.
- Scenarios for financial data in 2030 and beyond.