Project Report: Fama–French ML Quant Strategy (Amsterdam UCITS ETFs)

This document summarizes the discussion, steps, and work carried out in the chat session regarding building a machine learning trading strategy based on Fama–French factors, mapped to investable UCITS ETFs listed on the Amsterdam exchange. It also covers dataset preparation, visualization, and practical considerations for execution in Interactive Brokers (IBKR).

# 1. Data Acquisition

- We set up scripts to download both daily and monthly European Fama–French 5 Factors + Momentum from the Ken French Data Library.  
- Cleaned the datasets: standardized column names (MKT\_RF, SMB, HML, RMW, CMA, MOM, RF), converted returns from percent to decimal, and aligned indexes.  
- Merged FF5 with Momentum into `europe\_ff5\_plus\_mom\_monthly.csv` and `europe\_ff5\_plus\_mom\_daily.csv`.  
- Dataset coverage: 1990–2025 (monthly ~420 obs, daily ~9000 obs).

# 2. Data Visualization

- Built visualization pipeline:  
 • Time series of factor returns  
 • Growth of $1 (compounding)  
 • Rolling Sharpe ratios  
 • Correlation heatmaps  
 • Rolling correlations vs Market  
 • Drawdowns  
 • Distributions (histograms)  
- These plots provide intuition into factor performance and behavior across regimes (dot-com bubble, GFC, euro crisis, COVID, etc.).

# 3. Mapping Factors to UCITS ETFs (Amsterdam)

- Fama–French anomalies were mapped to Amsterdam-listed UCITS ETFs:  
  
 • Market (MKT\_RF): IWDA.AS, CSPX.AS, SXR8.AS  
 • Size (SMB): WSML.AS (iShares MSCI World Small Cap)  
 • Value (HML): IUVL.AS (iShares Edge MSCI World Value)  
 • Momentum (MOM): IWMO.AS (iShares Edge MSCI World Momentum)  
 • Profitability (RMW): IWQU.AS (iShares Edge MSCI World Quality)  
 • Investment (CMA): IWLV.AS (iShares Edge MSCI World Minimum Volatility)  
  
- Benchmark: IWDA.AS (iShares MSCI World UCITS).

# 4. Machine Learning Strategy

- ML model: Random Forest / XGBoost re-weighting engine.  
- Inputs: Fama–French factor returns (monthly or daily).  
- Targets: next-period ETF returns.  
- Rolling training window (e.g., 10 years monthly data = 120 samples).  
- Portfolio construction: long-only tilts (if ETF not shortable) or long–short replication (if shortable).  
- Benchmark comparison: Buy & Hold IWDA.AS.  
- Outputs: equity curves, CAGR, Sharpe, max drawdown, rolling alpha.

# 5. Shortability Considerations

- True Fama–French factors are long–short portfolios.  
- In practice, IBKR shortability depends on borrow availability.  
- Large liquid UCITS ETFs (IWDA.AS, CSPX.AS, SXR8.AS) are often shortable.  
- Specialty factor ETFs (WSML.AS, IWMO.AS, IUVL.AS, IWQU.AS, IWLV.AS) may or may not be shortable.  
- If not shortable: implement ML as a long-only tilting strategy vs benchmark.  
- If shortable: construct closer-to-theory anomaly replication portfolios.

# 6. Extensions

- Apply same ML pipeline to \*\*European sector ETFs\*\* (Amsterdam-listed, e.g., EXSA.AS, EXV6.AS).  
- Add macroeconomic features (EUR/USD, ECB rates, volatility indices).  
- Compare Europe vs US Fama–French factor behavior.  
- Deploy live via IBKR API in paper trading mode.

# 7. Career/Project Takeaways

- The repos (BTC GARCH, IWDA RSI signals, chess bot, Fama–French ML) demonstrate quant research capability.  
- Vibe coding with ChatGPT is valid: employers care about reproducible projects, not who typed every line.  
- These projects prove ability to connect theory → code → data → strategy.  
- Next step: polish GitHub READMEs, add plots and performance tables, and frame them as portfolio-ready quant research projects.