**Empirical Study of Tail Distributions in Equity Returns**

**1. Objective**

This project aims to empirically analyze the tail behaviour of daily log returns of major global equity indices using a 10-year dataset. The study quantifies tail risk through Value-at-Risk (VaR) and Expected Shortfall (ES) under three statistical assumptions:

* **Normal distribution** (baseline)
* **Student’s t-distribution** (heavy-tailed, symmetric)
* **Generalized Pareto Distribution (GPD)** via Extreme Value Theory (EVT) (non-parametric)

**2. Dataset**

* **Indices:** NSEI (India), GSPC (US S&P 500), FTSE (UK), N225 (Japan)
* **Time Horizon:** Sep 2015 – Aug 2025 (10 years)
* **Data Source:** YFinance, FRED exchange rates for currency normalization
* **Note:** Returns were converted to USD to ensure comparability across indices

**3. Methodology**

**3.1 Log Returns & Distributions**

* Computed daily log returns for each index
* Compared return distributions against:
  + Normal PDFs and CDFs
  + Student-t fits using Q-Q plots
  + GPD fits via Peaks Over Threshold (POT)

**3.2 EVT & Tail Modeling**

* Fitted GPD on losses exceeding thresholds
* Used Mean Residual Life (MRL) plots and ξ (shape parameter) stabilization to identify appropriate thresholds
* Observed that for NSEI, GSPC, FTSE: threshold **u ≈ 0.010** gave stable results
* For N225, a slightly higher threshold was required

**3.3 Heavy-Tail Evidence**

* Student-t distribution outperformed normal, particularly in the tails
* EVT shape parameters (ξ) showed:
  + **ξ > 0** across all indices → power-law (fat) tails
  + High excess kurtosis and negative skewness observed

**4. Value-at-Risk (VaR) & Expected Shortfall (ES)**

**4.1 VaR Methodologies**

Computed 1-day VaR at confidence levels **α ∈ [95%, 99.9%]** using:

* **Normal VaR:** −(μ + σz\_α)
* **Student-t VaR:** −(μ + σt\_α)
* **EVT VaR:** GPD-based tail modeling beyond threshold

**4.2 Expected Shortfall (ES)**

* **Normal ES:** Uses PDF-based tail formula
* **Student-t ES:** Adjusted for degrees of freedom (ν)
* **EVT ES:** Based on conditional GPD expectations

**4.3 Observations**

* EVT-based VaR/ES estimates were **significantly higher** than Gaussian estimates, especially at 99.9% CI
* Confirmed that Gaussian models **underestimate tail risk** in extreme market scenarios

**5. Key Insights**

* **Normal distribution is inadequate** for tail modeling of equity returns
* Student-t improves fit but **EVT-GPD gives the most realistic tail characterization**
* EVT-based VaR/ES provides more conservative and actionable risk metrics
* Tail risk was **highest for N225**, **lowest for GSPC** over the 10-year period
* GPD fits aligned well with **historical market shocks** (e.g., COVID-19 crash, 2022 rate hikes)

**6. Tools & Libraries Used**

* **Languages:** Python
* **Libraries:** NumPy, Pandas, Matplotlib, SciPy, statsmodels, sklearn