#### **1. Project Overview**

The primary objective of my thesis is to deepen my understanding of financial markets, specifically focusing on the factors that influence volatility. To achieve this, I aim to apply machine learning techniques learned throughout the year to analyze these factors. Through this project, I hope to gain expertise in both finance and data analysis, while also identifying career paths in either or both of these domains.

The end goal is to deliver a comprehensive **100-page paper** that includes not only the required theoretical content, but also a practical, quantitative study of the phenomenon. This study will focus on volatility in financial markets, leveraging machine learning models to extract insights.

#### **2. Work Completed So Far**

* **Initial Research**: The literature review has been completed, providing a foundational understanding of financial volatility and machine learning applications in finance.
* **Data Collection**: Data has been gathered from **yfinance** and **FRED** (Federal Reserve Economic Data), covering key financial variables for the analysis of market volatility.
* **Data Preprocessing**: The preprocessing steps have been successfully completed, including cleaning the data, handling missing values, and transforming features to make them suitable for machine learning models.
* **Model Development**:  
  + Three machine learning models have been developed and tested:  
    - **K-Means Clustering**: Applied for unsupervised learning of volatility regimes.
    - **Gaussian Mixture Model (GMM)**: Used for modeling probability distributions of volatility states.
    - **DBSCAN (Density-Based Spatial Clustering of Applications with Noise)**: Implemented to identify volatility regimes in a non-parametric way.
    - **XGBoost**: Employed as a supervised learning model for volatility prediction.
  + I also plan to add a **Hidden Markov Model (HMM)** to further enhance the analysis of volatility regimes and improve prediction accuracy.
* **Thesis Writing**:  
  + Most sections of the thesis have been written, including the introduction, methodology, and analysis.
  + The **discussion** section requires the most refinement and completion, as I aim to integrate the findings from the models with financial theory and existing literature.

#### **3. Challenges Faced**

At this stage, no significant challenges have arisen that hinder progress. However, I anticipate that fine-tuning the HMM model and integrating it into the existing analysis may present some challenges in terms of model complexity and data fitting.

#### **4. Upcoming Tasks**

* **Add the HMM model**: I plan to incorporate the **Hidden Markov Model** into the analysis to enhance the study of volatility regimes.
* **Complete Thesis Writing**: The **discussion** and **conclusion** sections need to be completed and refined. The final paper will integrate the results from the machine learning models with broader financial theory and provide clear conclusions on the findings.
* **UBS Interview**: As part of my career exploration, I have an interview with the **Quantitative Strategies Executive Director at UBS**, which will also provide valuable insights for the career section of my thesis.

#### **5. Goals for the Next Period**

* **Finalize the Thesis Writing by the End of April**: My goal is to complete the thesis by the end of April, focusing on refining the discussion and concluding sections.
* **Prepare for Thesis Defense by May**: After completing the thesis, I will begin preparing for the defense, aiming to schedule it before the end of May.