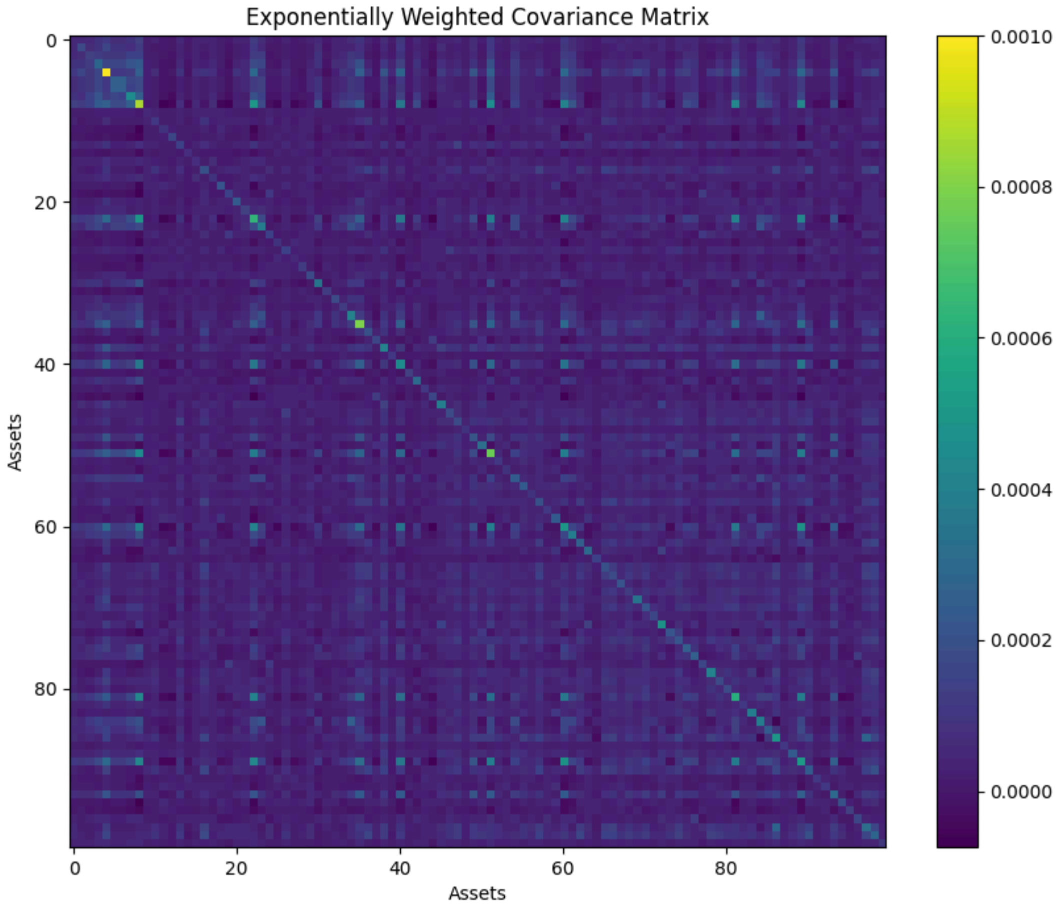
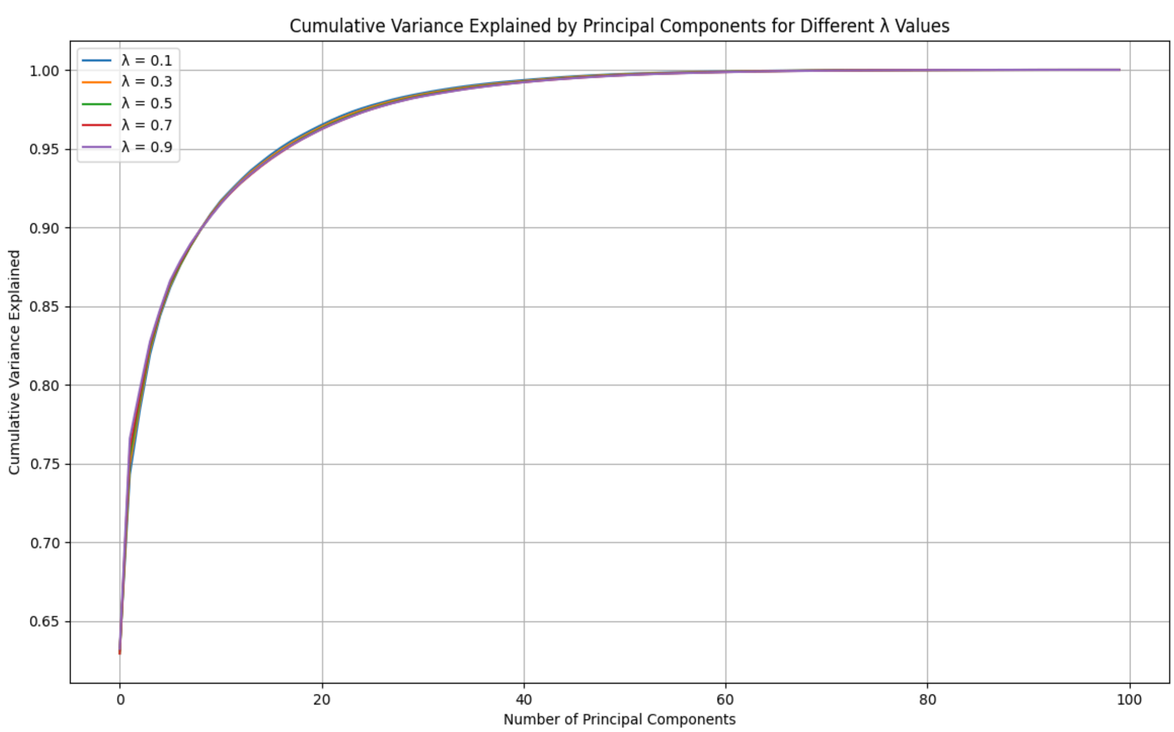
1. Problem 1



Exponentially Weighted Covariance Matrix
Assets
Assets
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000



Cumulative Variance Explained by Principal Components for Different λ Values
λ = 0.1
λ = 0.3
λ = 0.5
λ = 0.7
λ = 0.9
Cumulative Variance Explained
Number of Principal Components

1. λ Near 0 (e.g., λ = 0.01):

- **Greater Emphasis on Recent Data**: A lower λ places much more weight on recent observations, causing older data to quickly lose influence in the calculation.
- **Higher Reactivity, Increased Volatility**: The covariance matrix becomes highly sensitive to recent changes, capturing short-term fluctuations in the market. However, this heightened responsiveness can introduce volatility as it reacts more to short-term movements.
- **Less Stability**: With minimal influence from older data, the covariance matrix reflects immediate market conditions but may lack the smoothness associated with longer-term trends.

2. λ Near 1 (e.g., λ = 0.99):

- **Greater Emphasis on Historical Data**: A higher λ causes weights to decay more gradually, allowing the covariance matrix to incorporate a larger window of historical data. Older data points maintain significance, contributing to a more comprehensive analysis.
- **Smoother, More Stable Matrix**: The matrix becomes less responsive to short-term market fluctuations, as it places more weight on long-term trends. While this leads to greater stability, it may be slower to adapt to sudden changes in recent market dynamics.
- **Reduced Volatility**: The matrix reflects more stable, long-term correlations, making it less susceptible to the noise of short-term fluctuations.

3. Effect on PCA (Principal Component Analysis):

- **For Smaller λ**: PCA results will emphasize short-term variations. The first principal component (or leading eigenvalue) is likely to explain a large proportion of the variance, with the cumulative variance quickly diminishing for subsequent components.
- **For Larger λ**: The variance is distributed more evenly across multiple components. This results in a slower accumulation of explained variance, as more eigenvalues contribute meaningfully to the overall variance, capturing long-term relationships.

Conclusion:

The value of **λ** offers a way to balance reactivity and stability in the EWMA covariance matrix. A smaller λ focuses on short-term movements, resulting in greater responsiveness but increased volatility, while a larger λ emphasizes long-term trends, producing a more stable but less reactive matrix. The choice of λ depends on whether the analysis prioritizes recent market movements or seeks to capture broader, long-term patterns.

Problem 2

```
Higham's method took 0.0266 seconds
Matrix multiplication successful
Is the corrected matrix PSD? True
Frobenius Norm - Original: 450.10494610587335, Near PSD: 450.10494160645254
```

Summary of Results

1. **Execution Time**:
    - Higham's method took **0.0266 seconds(varies every execution try)** to correct the non-PSD matrix. This indicates that the method is relatively efficient for the size of the matrix (500x500).
2. **Matrix PSD Check**:
    - The output confirms that the corrected matrix is **PSD** (True). This means that Higham’s method successfully adjusted the eigenvalues to ensure all are non-negative.
3. **Frobenius Norm Comparison**:
    - **Original Frobenius Norm**: 450.10494610587335
    - **Near PSD Frobenius Norm**: 450.10494160645254
    - The difference in Frobenius norms is very small, indicating that the adjustments made to the matrix were minor and that the structure of the matrix was largely preserved.
4. **Pros and Cons**:
    - **Higham's Method**:
        - **Pros**:

1. **Accuracy**: Higham's method ensures that the resulting matrix is positive semi-definite. The adjustment of eigenvalues maintains the overall structure, which is particularly important in statistical applications where matrix properties are crucial.

- **Cons**:
  1. **Complexity**: The method involves eigenvalue decomposition, which can become computationally expensive as matrix size increases (scales as O(N^3)). While the current execution time is acceptable, larger matrices may see increased runtimes.
  2. **Implementation**: If the data structure or matrix is poorly conditioned, Higham's adjustments may require more careful tuning, which can complicate implementation.
- Scenario: where maintaining the positive semi-definiteness of covariance matrices or correlation matrices is critical, such as in financial modeling or multivariate statistics.

- **near_psd Method**:
  - **Pros**:
    1. **Simplicity**: The near PSD method is straightforward to implement and typically involves fewer steps compared to more complex methods like Higham's. It primarily relies on eigenvalue adjustments, which are easier to understand conceptually.
    2. Can be faster due to fewer computations; often sufficient for practical applications.

- **Cons**:
  1. **Less Precision**: While it is effective for many cases, near PSD might not be as precise as Higham’s method in ensuring that the resulting matrix is truly PSD.
  2. **Potential for Larger Adjustments**: In some scenarios, the necessary adjustments to achieve PSD can be more significant than desired, potentially distorting the matrix properties more than a method like Higham's.
- Scenario: while having moderately non-PSD matrices and need a quick fix without extensive adjustments.

Problem 3

```
Frobenius Norm for Direct: 8.385599315149827e-08
Frobenius Norm for PCA 100.0%: 2.0327230522289617e-06
Frobenius Norm for PCA 75.0%: 5.205550563298977e-07
Frobenius Norm for PCA 50.0%: 1.108297589460235e-06
```

1. **Direct Simulation**

- Accuracy: Using the full covariance matrix for multivariate normal sampling ensures the highest accuracy, as the entire structure is preserved.
- Time: Direct simulation is computationally expensive and tedious since no dimensionality reduction is performed. Every sample drawn requires using the full covariance matrix.

2. **PCA-based Simulation:**

- Accuracy: Reducing the number of principal components compromises accuracy, as approximating the covariance matrix using fewer dimensions. However, the more we reduce the number of components, the more accuracy you sacrifice.
- Time: PCA-based simulations run much faster because of the dimensionality reduction. Fewer principal components mean fewer random variables are simulated, speeding up both sampling and matrix operations.

Trade-off: PCA introduces a trade-off between time and accuracy. Higher explained variance ratios (closer to 100%) mean greater accuracy but slower runtime. Lower variance ratios (like 50%) speed up computation but at the cost of potentially misrepresenting the true covariance structure.

2. Choosing the Right Trade-Off

- **For high-frequency trading or real-time analysis**, speed is critical; so, prioritizing faster methods like PCA with a lower explained variance threshold or simplified PSD corrections.
- **For risk management or portfolio simulations**, accuracy might be more important; so, full covariance matrices, high explained variance, and precise PSD corrections are preferred, even at the cost of longer runtimes.