
# Introduction to Pair Trading

# Learning Objectives

- Understand the advantages of trading a pair of stocks versus a single stock
- Identify the patterns of price movements that cause a pair trade to be profitable
- Construct a beta-hedged pair strategy

# Agenda

Advantages of trading a pair of stocks versus a single stock

Price movements that cause a pair trade to be profitable

Beta Hedging a Stock Pair

# Single Stock Trading

- Trading a single stock is betting on returns moving in a single direction
- If you are bullish you go long and hope for the price to rise
- If you are bearish you sell short and hope for the price to decrease

# Short Selling Rules

- Selling a stock you don’t currently own
- Legal requirement to borrow stock (usually from a stock lender) and then sell it
- Close trade by buying the stock back and returning it to the lender

# Single Stock vs Pair Trading

- Trading a single stock is betting on returns moving in a single direction
- Pairs trading requires you to focus on relative performance rather than absolute performance

# Agenda

Advantages of trading a pair of stocks versus a single stock

Price movements that cause a pair trade to be profitable

Beta Hedging a Stock Pair

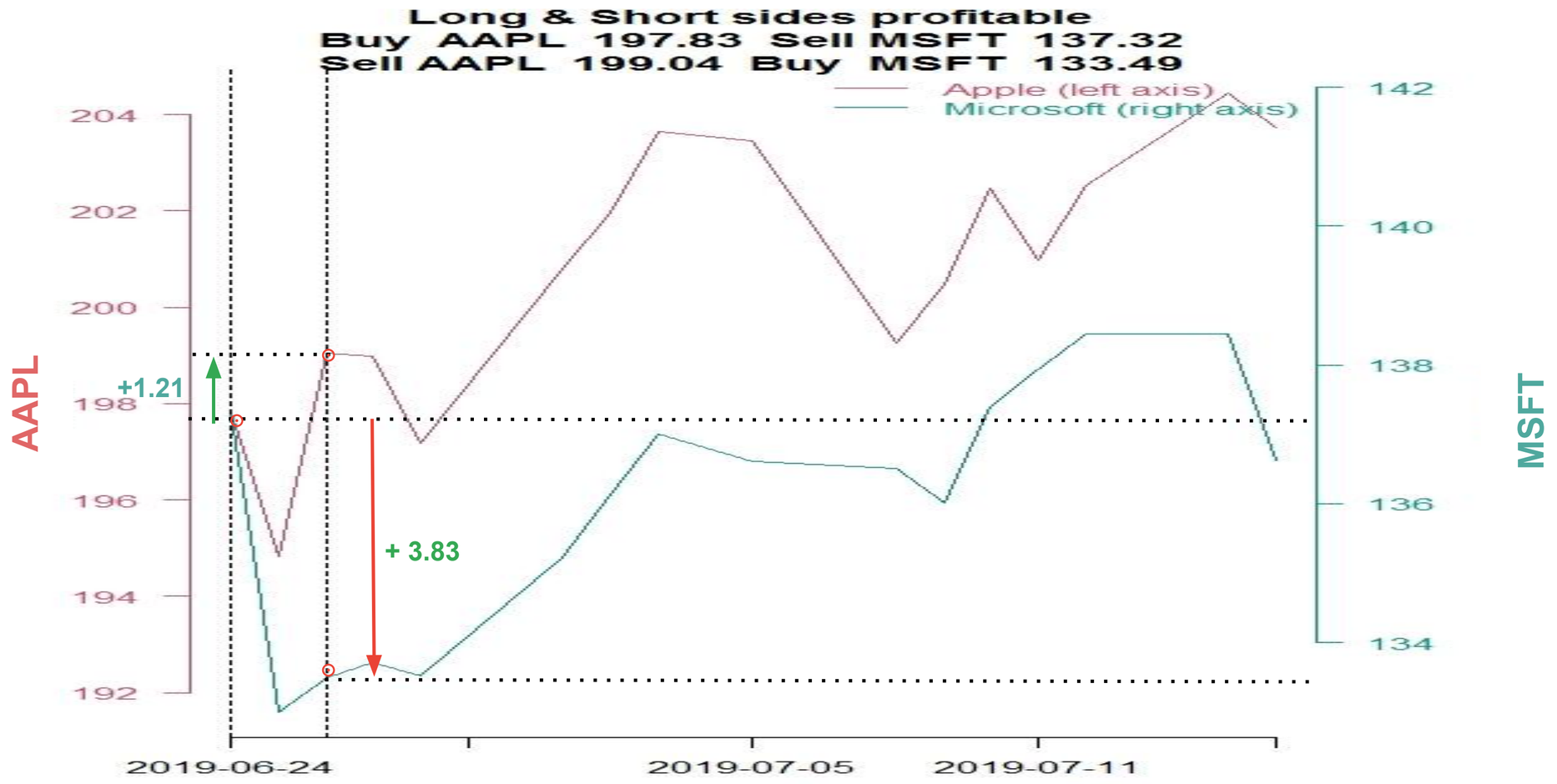
# Profitable Scenarios for Pair Trading

1. AAPL ↑ 5% MSFT ↓ 5%
2. AAPL ↑ 15% MSFT ↑ 5%
3. AAPL ↓ 5% MSFT ↓ 15%

➔ Long side outperforms

➔ Profit is +10% for all 3 scenarios

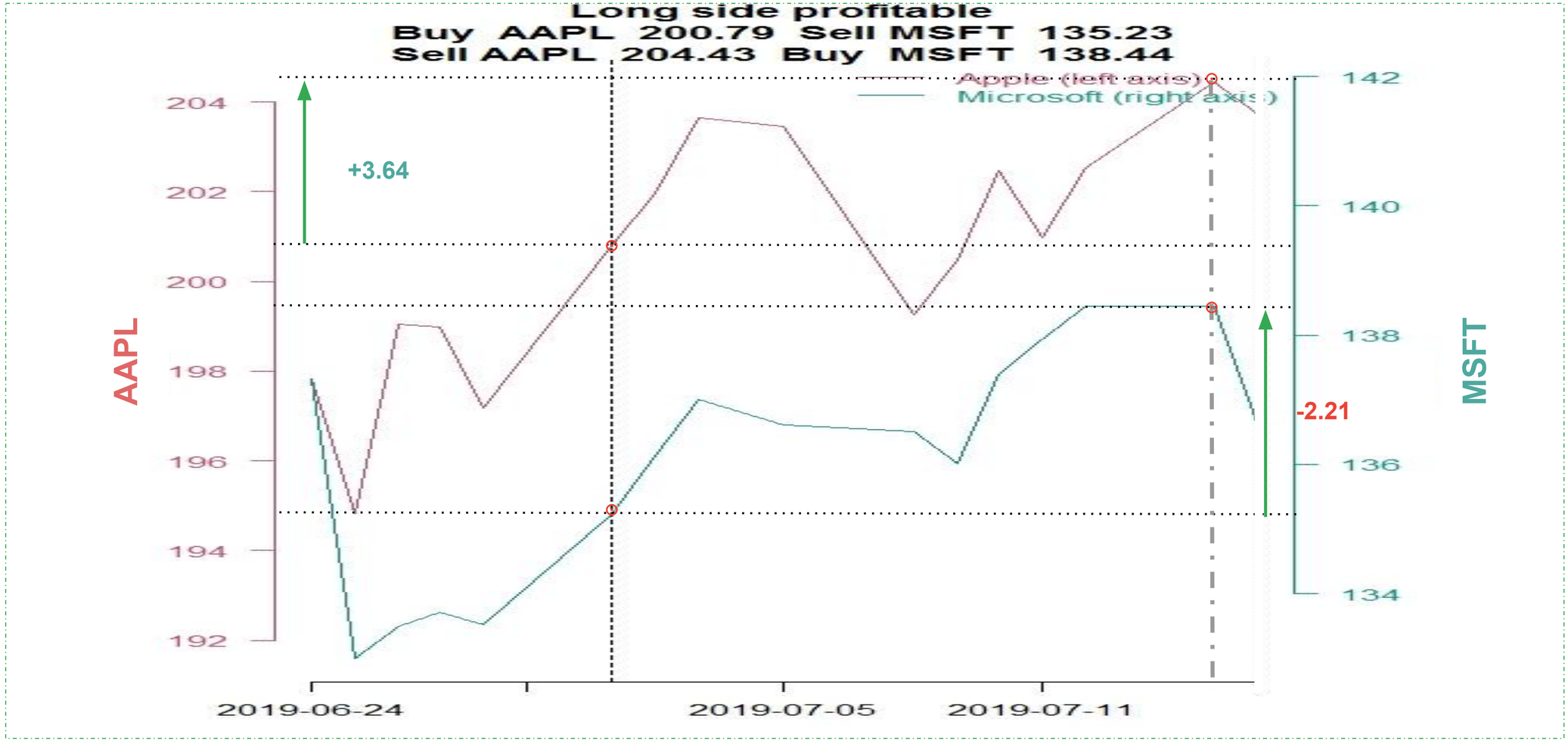
# Profitable Pair Trade: AAPL up 1.21 & MSFT down 3.83




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Profitable Pair Trade: AAPL up 3.64 & MSFT up 3.21




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Agenda

Advantages of trading a pair of stocks versus a single stock

Price movements that cause a pair trade to be profitable

Beta Hedging a Stock Pair

# Beta Hedging a Stock Pair

Beta is a stock's covariance with the overall market per unit of market risk

$$\beta_{AAPL} = \frac{COV(AAPL, MKT)}{VAR(MKT)}$$

Beta is a measure of a stock's systematic risk

$$\beta_{AAPL} = 1.10$$

⇒ AAPL has 10% more systematic volatility than the market

# Beta Hedging a Stock Pair

For Large Cap Stocks:

**⇒ MKT ≈ S&P 500**

If S&P 500 is up 20%

**⇒ AAPL expected to be up 22%**

If S&P 500 is down 20%

**⇒ AAPL expected to be down 22%**

# CAPM and Beta

**Capital Asset Pricing Model (CAPM):**

$$E(R_{stock}) - R_{rf} = \beta_{stock} * [E(R_{mkt}) - R_{rf}]$$

⇒ a stock's expected excess return is equal to its $\beta$ times the expected market risk premium $[E(R_{mkt}) - R_{rf}]$

**3-month $\beta_{AAPL} = 1.10$ $\beta_{MSFT} = 0.96$**

⇒ AAPL has 1.10/0.96 = 14.6% more exposure to market risk than MSFT

# CAPM and Beta

**Capital Asset Pricing Model (CAPM):**

$$E(R_{stock}) - R_{rf} = \beta_{stock} * [E(R_{mkt}) - R_{rf}]$$

⇒ a stock's expected excess return is equal to its $\beta$ times the expected market risk premium $[E(R_{mkt}) - R_{rf}]$

**3-month $\beta_{AAPL} = 1.10$ $\beta_{MSFT} = 0.96$**

⇒ AAPL has 1.10/0.96 = 14.6% more exposure to market risk than MSFT

# CAPM and Beta

**Capital Asset Pricing Model (CAPM):**

$$E(R_{stock}) - R_{rf} = \beta_{stock} * [E(R_{mkt}) - R_{rf}]$$

⇒ a stock's expected excess return is equal to its $\beta$ times the expected market risk premium $[E(R_{mkt}) - R_{rf}]$

**3-month $\beta_{AAPL} = 1.10$ $\beta_{MSFT} = 0.96$**

⇒ AAPL has 1.10/0.96 = 14.6% more exposure to market risk than MSFT

# CAPM and Beta

**Capital Asset Pricing Model (CAPM):**

$$E(R_{stock}) - R_{rf} = \beta_{stock} * [E(R_{mkt}) - R_{rf}]$$

⇒ a stock's expected excess return is equal to its $\beta$ times the expected market risk premium $[E(R_{mkt}) - R_{rf}]$

**3-month $\beta_{AAPL} = 1.10$ $\beta_{MSFT} = 0.96$**

⇒ AAPL has 1.10/0.96 = 14.6% more exposure to market risk than MSFT

# CAPM and Beta

**Capital Asset Pricing Model (CAPM):**

$$E(R_{stock}) - R_{rf} = \beta_{stock} * [E(R_{mkt}) - R_{rf}]$$

⇒ a stock's expected excess return is equal to its $\beta$ times the expected market risk premium $[E(R_{mkt}) - R_{rf}]$

**3-month $\beta_{AAPL} = 1.10$ $\beta_{MSFT} = 0.96$**

⇒ AAPL has 1.10/0.96 = 14.6% more exposure to market risk than MSFT

# Weighting Long and Short Components of Pair Trade

- Pair will not contain equal amounts of AAPL and MSFT
- Weighting must be **Beta Neutral** and also factor in relative stock prices.

$$\beta_{AAPL} * P_{AAPL} * Q_{AAPL} = \beta_{MSFT} * P_{MSFT} * Q_{MSFT}$$

# Weighting Long and Short Components of Pair Trade

- Pair will not contain equal amounts of AAPL and MSFT
- Weighting must be **Beta Neutral** and also factor in relative stock prices.

$$\beta_{AAPL} * P_{AAPL} * Q_{AAPL} = \beta_{MSFT} * P_{MSFT} * Q_{MSFT}$$

# Weighting Long and Short Components of Pair Trade

- Pair will not contain equal amounts of AAPL and MSFT
- Weighting must be **Beta Neutral** and also factor in relative stock prices.

$$\beta_{AAPL} * P_{AAPL} * Q_{AAPL} = \beta_{MSFT} * P_{MSFT} * Q_{MSFT}$$

# Weighting Long and Short Components of Pair Trade

- Pair will not contain equal amounts of AAPL and MSFT
- Weighting must be **Beta Neutral** and also factor in relative stock prices.

$$\beta_{AAPL} * P_{AAPL} * Q_{AAPL} = \beta_{MSFT} * P_{MSFT} * Q_{MSFT}$$

# Weighting Long and Short Components of Pair Trade

- Pair will not contain equal amounts of AAPL and MSFT
- Weighting must be **Beta Neutral** and also factor in relative stock prices.

$$\beta_{AAPL} * P_{AAPL} * Q_{AAPL} = \beta_{MSFT} * P_{MSFT} * Q_{MSFT}$$

# Weighting Long and Short Components of Pair Trade

Suppose **MSFT is $135** & **AAPL is $200**

You wish to sell short 1000 shares of MSFT. How many shares of AAPL do you need to buy?

$$1.10 * \$200 * Q_{AAPL} = 0.96 * \$135 * 1000$$

$$Q_{AAPL} = 589 \text{ shares}$$

# Weighting Long and Short Components of Pair Trade

Suppose **MSFT is $135** & **AAPL is $200**

You wish to sell short 1000 shares of MSFT. How many shares of AAPL do you need to buy?

$$1.10 * \$200 * Q_{AAPL} = 0.96 * \$135 * 1000$$

$$Q_{AAPL} = 589 \text{ shares}$$

# Weighting Long and Short Components of Pair Trade

Suppose **MSFT is $135** & **AAPL is $200**

You wish to sell short 1000 shares of MSFT. How many shares of AAPL do you need to buy?

$$1.10 * \$200 * Q_{AAPL} = 0.96 * \$135 * 1000$$

$$Q_{AAPL} = 589 \text{ shares}$$

# Weighting Long and Short Components of Pair Trade

Suppose **MSFT is $135** & **AAPL is $200**

You wish to sell short 1000 shares of MSFT. How many shares of AAPL do you need to buy?

$$1.10 * \$200 * Q_{AAPL} = 0.96 * \$135 * 1000$$

$$Q_{AAPL} = 589 \text{ shares}$$

# Weighting Long and Short Components of Pair Trade

Suppose **MSFT is $135** & **AAPL is $200**

You wish to sell short 1000 shares of MSFT. How many shares of AAPL do you need to buy?

$$1.10 * \$200 * Q_{AAPL} = 0.96 * \$135 * 1000$$

$$Q_{AAPL} = 589 \text{ shares}$$

# Weighting Long and Short Components of Pair Trade

Suppose **MSFT is $135** & **AAPL is $200**

You wish to sell short 1000 shares of MSFT. How many shares of AAPL do you need to buy?

$$1.10 * \$200 * Q_{AAPL} = 0.96 * \$135 * 1000$$

$$Q_{AAPL} = 589 \text{ shares}$$

# Weighting Long and Short Components of Pair Trade

Suppose **MSFT is $135** & **AAPL is $200**

You wish to sell short 1000 shares of MSFT. How many shares of AAPL do you need to buy?

$$1.10 * \$200 * Q_{AAPL} = 0.96 * \$135 * 1000$$

$$Q_{AAPL} = 589 \text{ shares}$$

# Weighting Long and Short Components of Pair Trade

Suppose **MSFT is $135** & **AAPL is $200**

You wish to sell short 1000 shares of MSFT. How many shares of AAPL do you need to buy?

$$1.10 * \$200 * Q_{AAPL} = 0.96 * \$135 * 1000$$

$$Q_{AAPL} = 589 \text{ shares}$$

# Checking Market Value for Each Side of Trade

Mkt Val. AAPL is $200*589 = $117,800

Beta = 1.10 ⇒ adjusted market value is:

**$117,800*1.10 = $129,580**

Mkt Val. MSFT is $135*1000 = $135,000

Beta = 0.96 ⇒ adjusted market value is:

**$135,000*0.96 = $129,600**

# Checking Market Value for Each Side of Trade

Mkt Val. AAPL is $200*589 = **$117,800**

Beta = 1.10 ⇒ adjusted market value is:

**$117,800*1.10 = $129,580**

# Checking Market Value for Each Side of Trade

Mkt Val. AAPL is $200*589 = **$117,800**

Beta = 1.10 ⇒ adjusted market value is:

**$117,800*1.10 = $129,580**

Mkt Val. MSFT is $135*1000 = **$135,000**

Beta = **0.96** ⇒ adjusted market value is:

**$135,000*0.96 = $129,600**

# Checking Market Value for Each Side of Trade

Mkt Val. AAPL is $200*589 = $117,800

Beta = 1.10 ⇒ adjusted market value is:

**$117,800*1.10 = $129,580**

Mkt Val. MSFT is $135*1000 = $135,000

Beta = 0.96 ⇒ adjusted market value is:

**$135,000*0.96 = $129,600**

**With rounding you are beta neutral!**

# Beta Hedging ✓
# Choosing Good Pair Candidates ?

- Constructing a Beta-hedged strategy is important for managing trading risk

- We still need a strategy to identify stocks that would make good pair candidates

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Picking Pairs

# Learning Objectives

- Understand the role of correlation in selecting pairs within an industry sector
- Create a hierarchical clustering using a daily return matrix
- Construct a scree plot using principal components

# Learning Objectives

- Understand the role of correlation in selecting pairs within an industry sector
- Create a hierarchical clustering using a daily return matrix
- Construct a scree plot using principal components

# Learning Objectives

- Understand the role of correlation in selecting pairs within an industry sector
- Create a hierarchical clustering using a daily return matrix
- Construct a scree plot using principal components

# Agenda

Use correlation to select pairs

Identify hierarchical clusters of stocks and create a cluster plot

Construct a scree plot for stocks within an industry sector

# Correlation between Stocks in Pair Trade

- Predicted correlation is negative
- Profit still depends on going long the the stock that will outperform...
- And going short the stock that will underperform

Long AAPL

Short MSFT

# Measuring Correlation

- Long-term negative correlation is unusual as almost all stocks have positive betas
- Correlations between pairs can be unstable and should be re-estimated regularly

# S&P 500 Sectors

1. Financials
2. Utilities
3. Consumer Discretionary
4. Consumer Staples
5. Energy
6. Health Care
7. Industrials
8. Technology
9. Telecom
10. Materials
11. Real Estate

# XLK Technology Sector Exchange-Traded Fund (ETF)

- Market-cap weighted exposure to entire sector
- **XLK** is the Technology sector fund
- **AAPL** and **MSFT** account for over **38%** of **XLK**

# Technology Sector: Top Sub-sectors

Sub-sectors form logical groups which can be useful for selecting pair candidates:

| | |
|---|---|
| Software & IT Services | 52.0% |
| Computers & Phones | 18.6% |
| Semiconductors | 17.4% |
| Communications & Network | 5.4% |

# DataFrame for XLK

- Daily returns matrix (**DataFrame**) for each of the 68 stocks in XLK
- Each row is a trading day (**Index**)
- Each Column is one of the stocks
- Each cell (the **Data**) consists of a single stock's return on a particular day

# Agenda

Use correlation to select pairs

Identify hierarchical clusters of stocks and create a cluster plot

Construct a scree plot for stocks within an industry sector

# What is Clustering?

**Cluster Analysis is a method of grouping similar objects:**

1. Decide metric to to measure differences between objects
2. Decide number of clusters
3. Assign objects to clusters so as to minimize differences

# Hierarchical Clustering Matrix

**Use the matrix containing daily returns for each stock in XLK**

1. **Columns** contain data for a **particular stock**
2. **Rows** contain data for a particular **trading day**

# Computing Distance between Returns

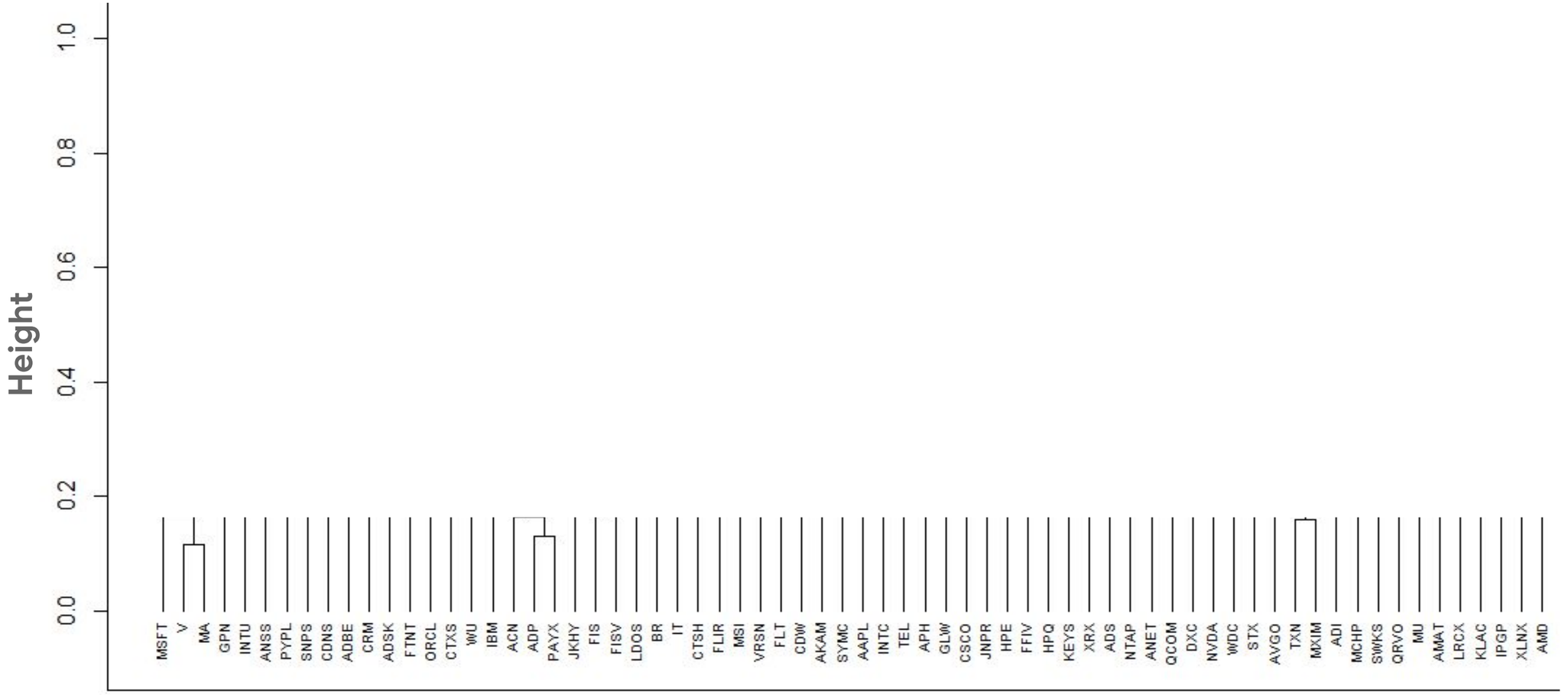
Objective is to group stocks that have similar returns (minimize distance between returns).

1. Calculate distance using the Euclidean method

$$\mathbf{d} = \sqrt{\sum_{\mathbf{i=1}}^{\mathbf{n}} (\mathbf{x_i} - \mathbf{y_i})^2}$$

2. Iteratively cluster stocks based on minimizing distance

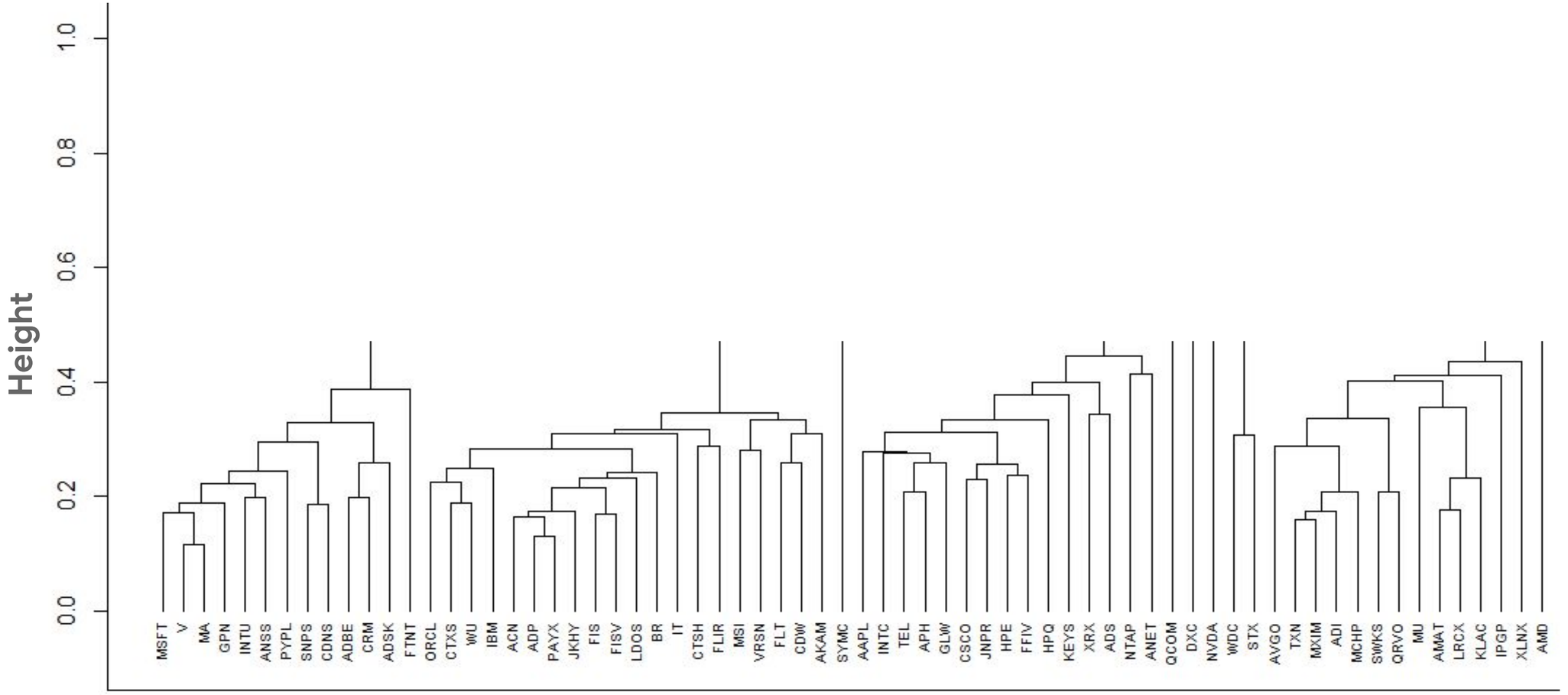
# Hierarchical Cluster Plot of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

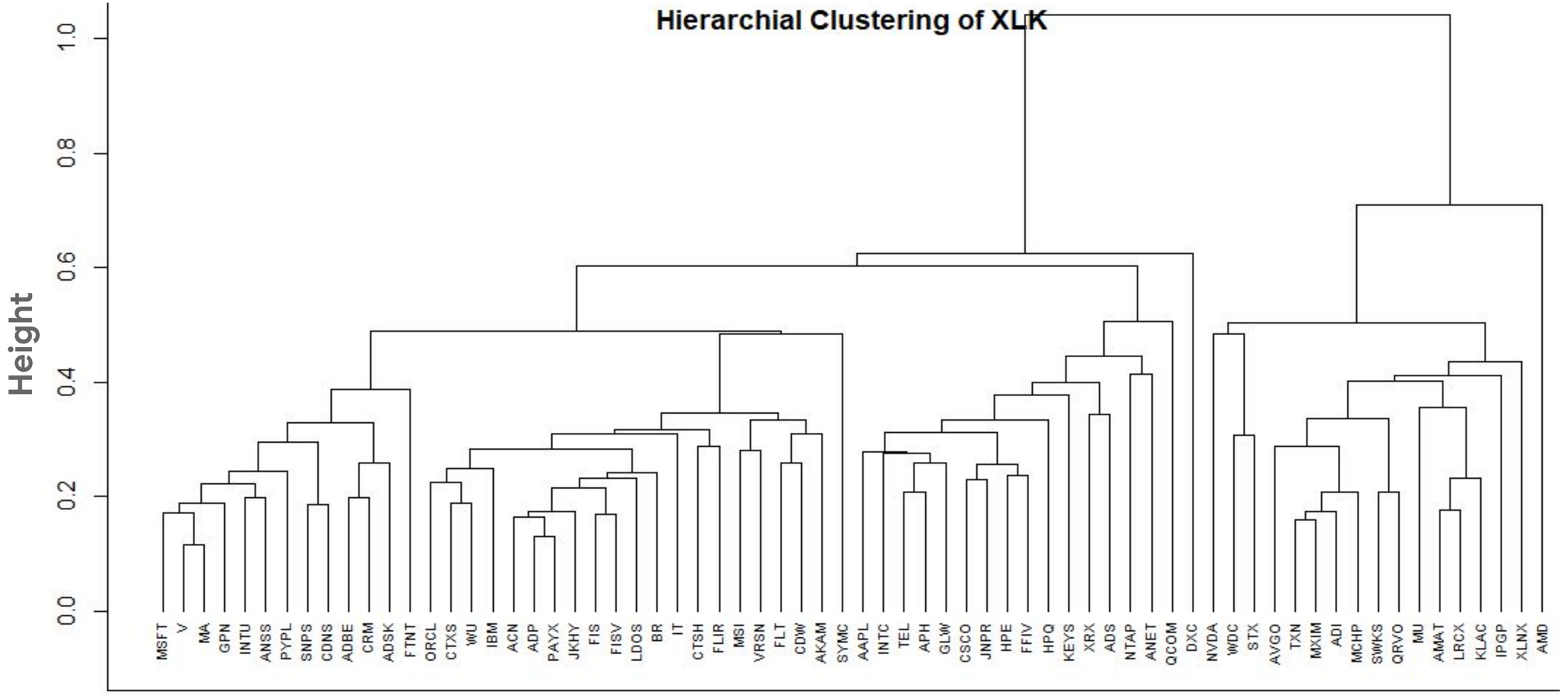
# Hierarchical Cluster Plot of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

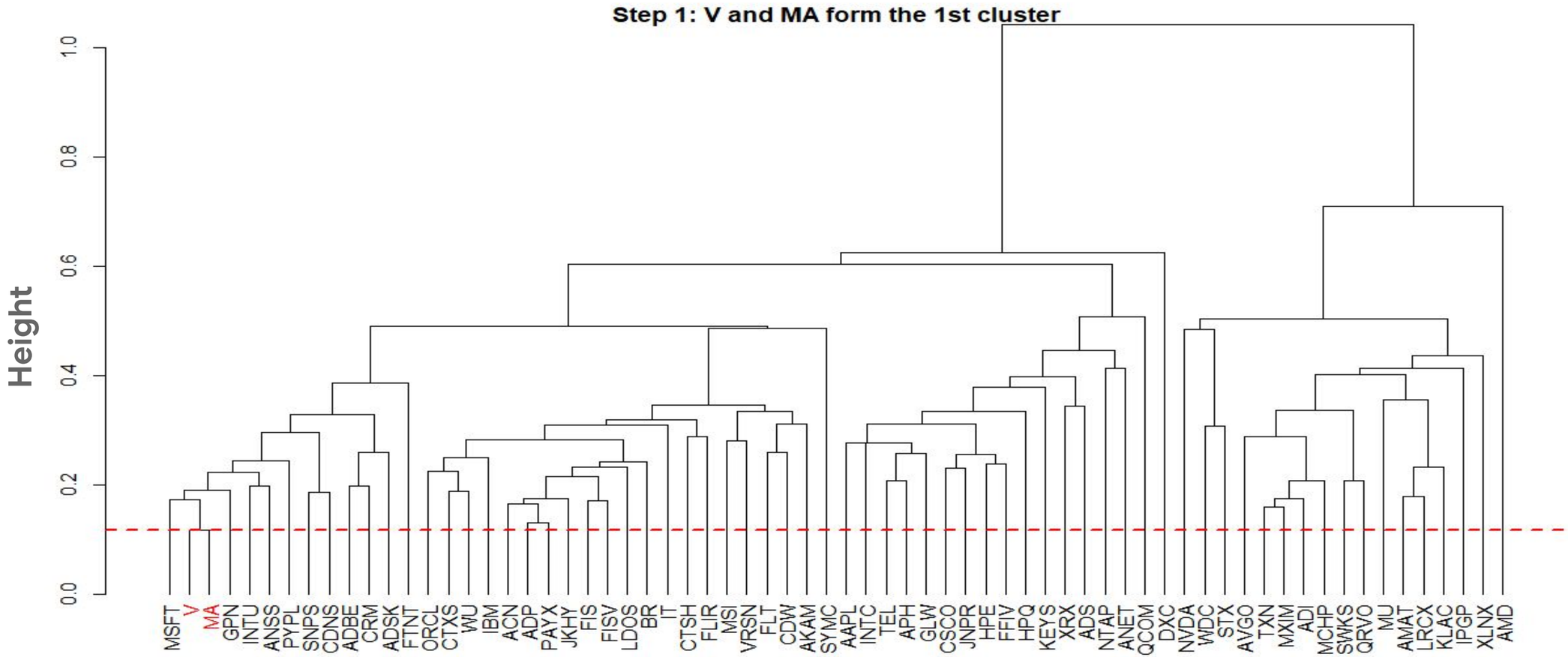
# Hierarchical Cluster Plot of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

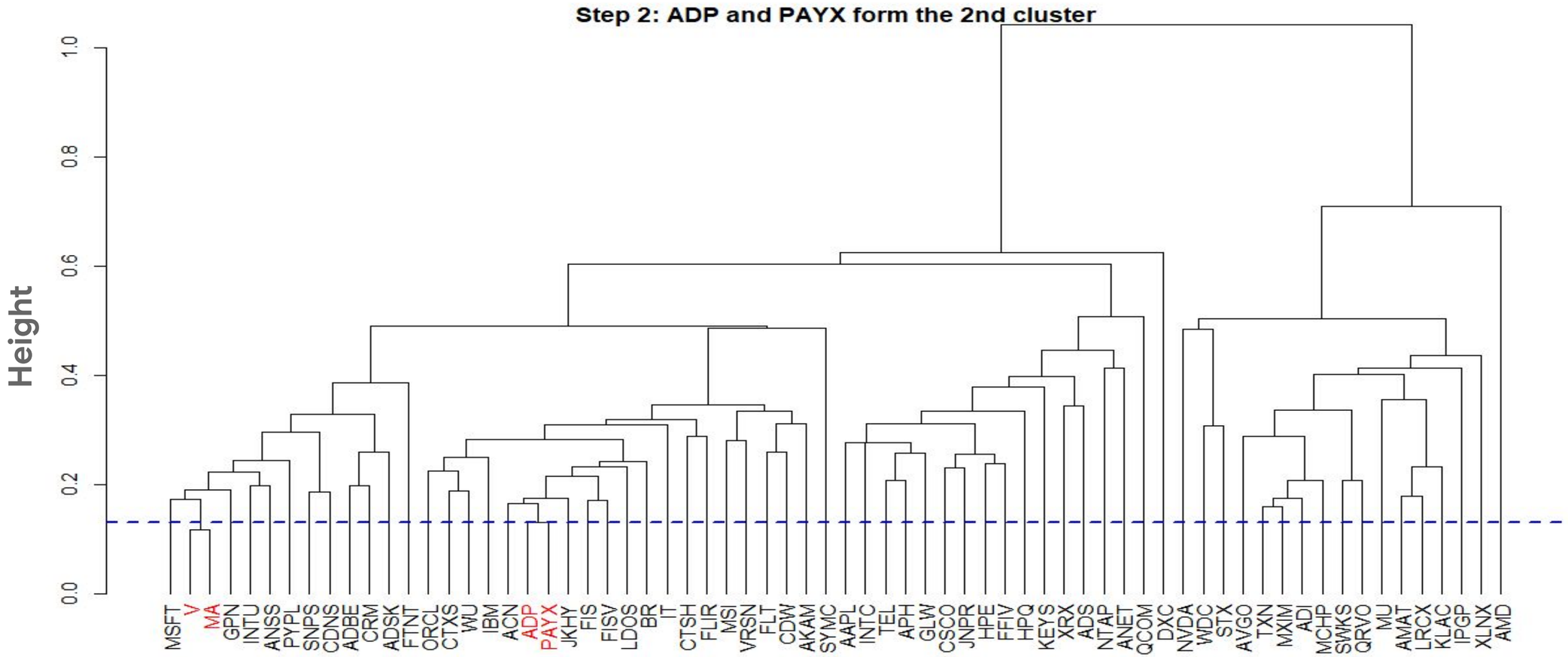
# Hierarchical Cluster Plot of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

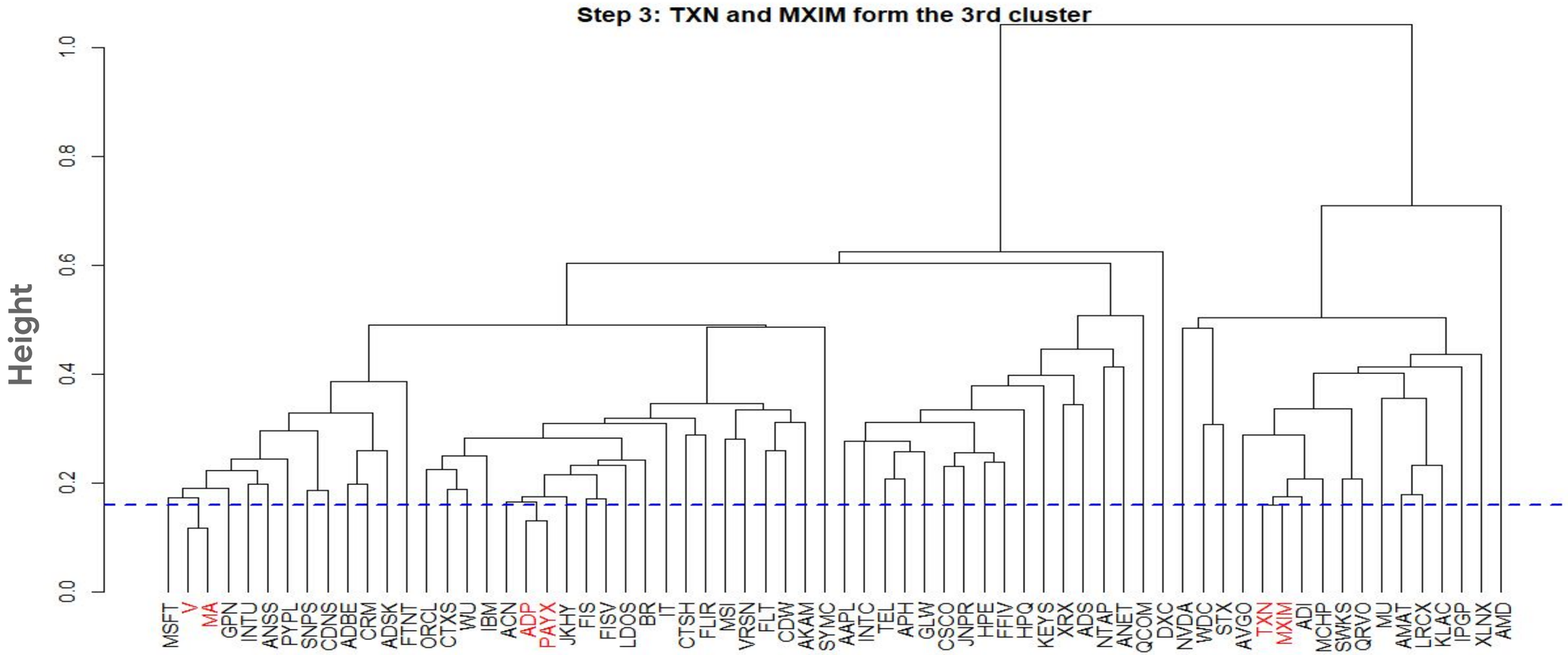
# Hierarchical Cluster Plot of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

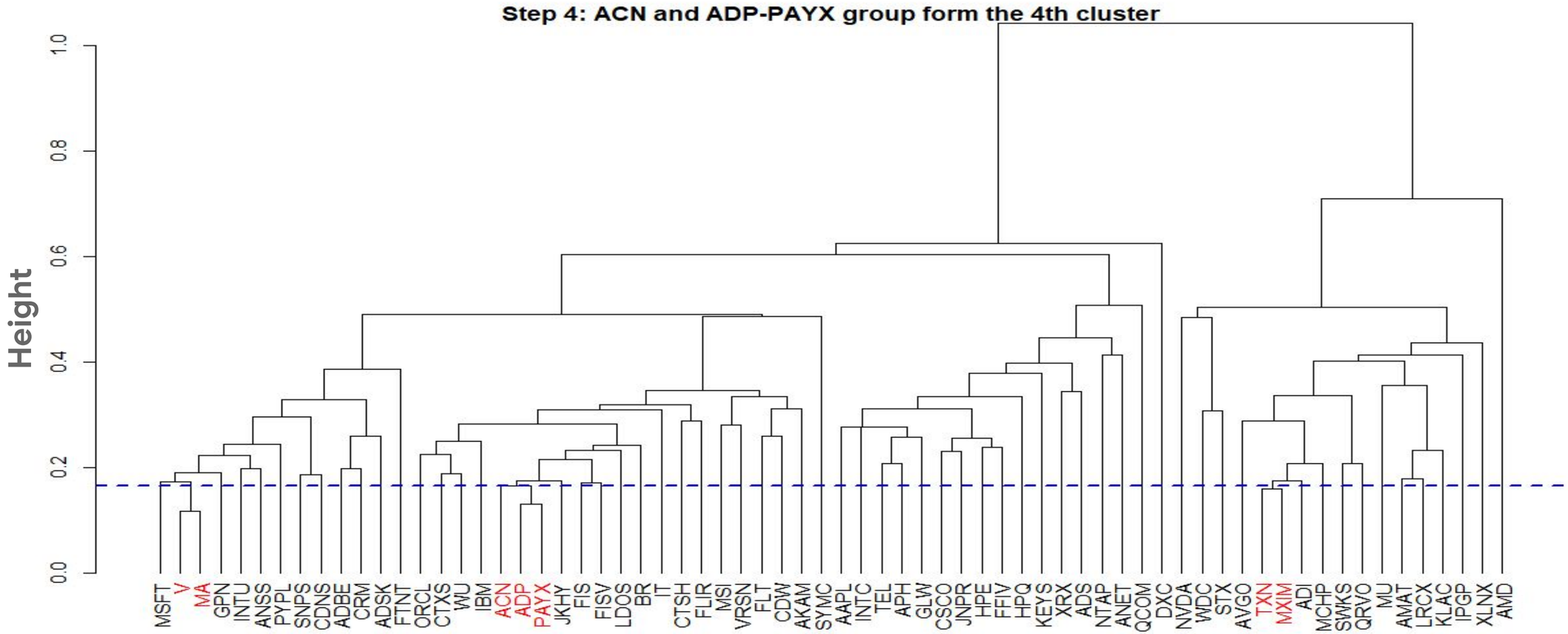
# Hierarchical Cluster Plot of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

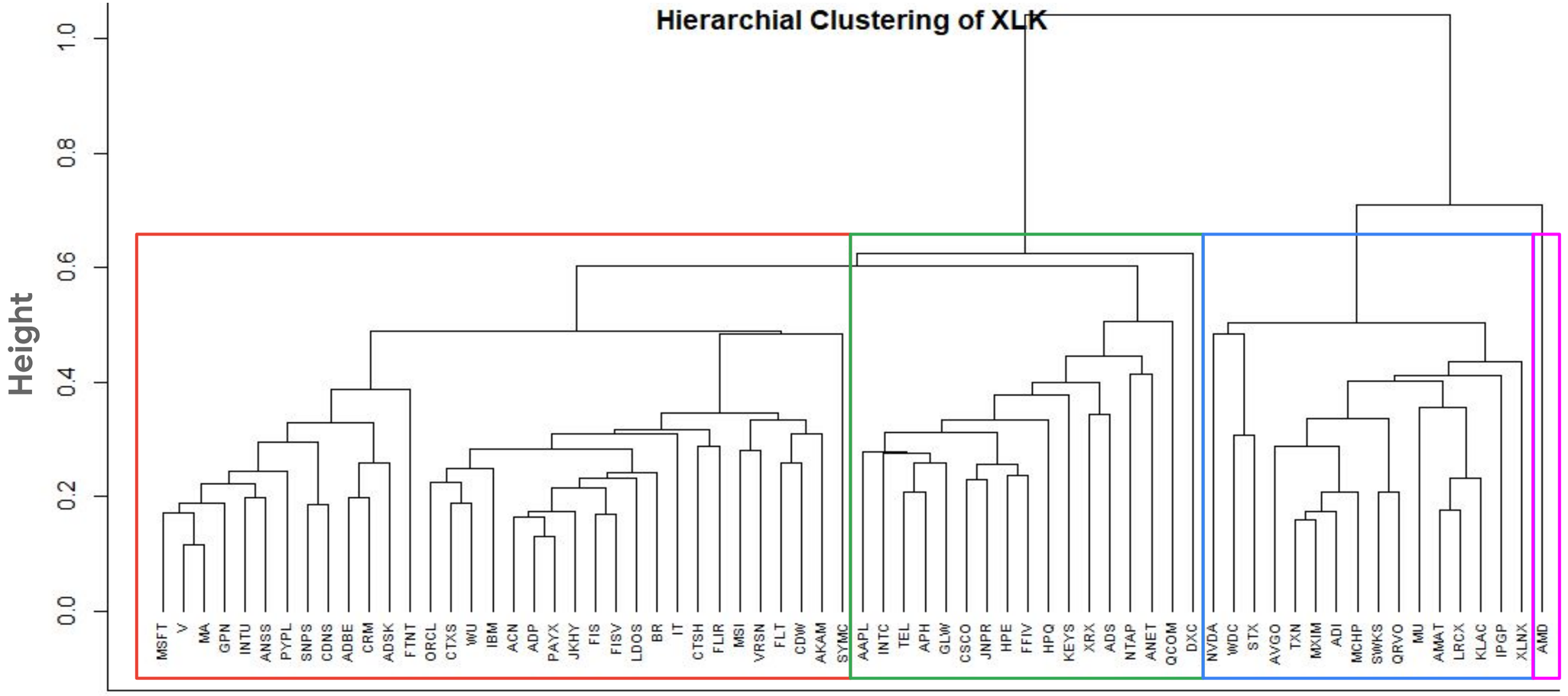
# Hierarchical Cluster Plot of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

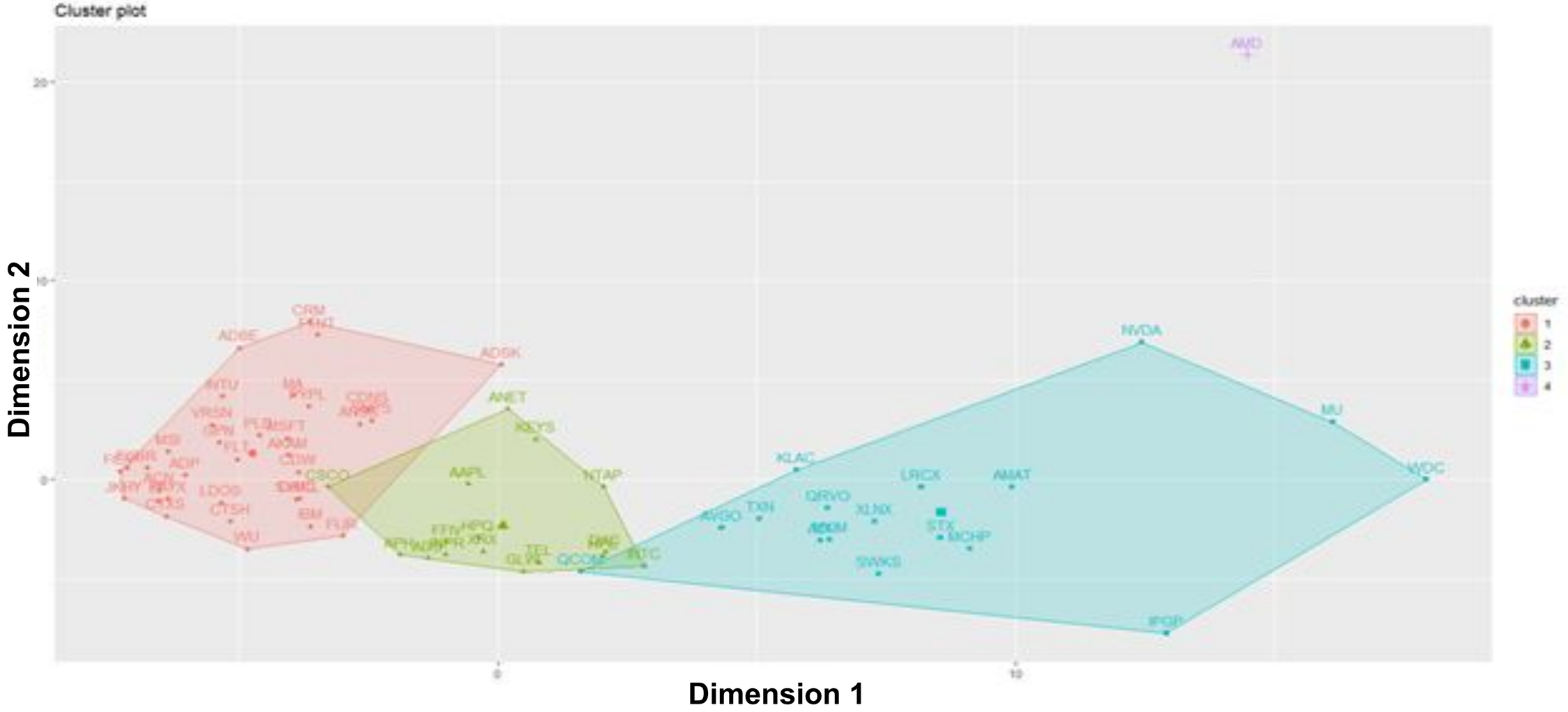
# Hierarchical Cluster Plot of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

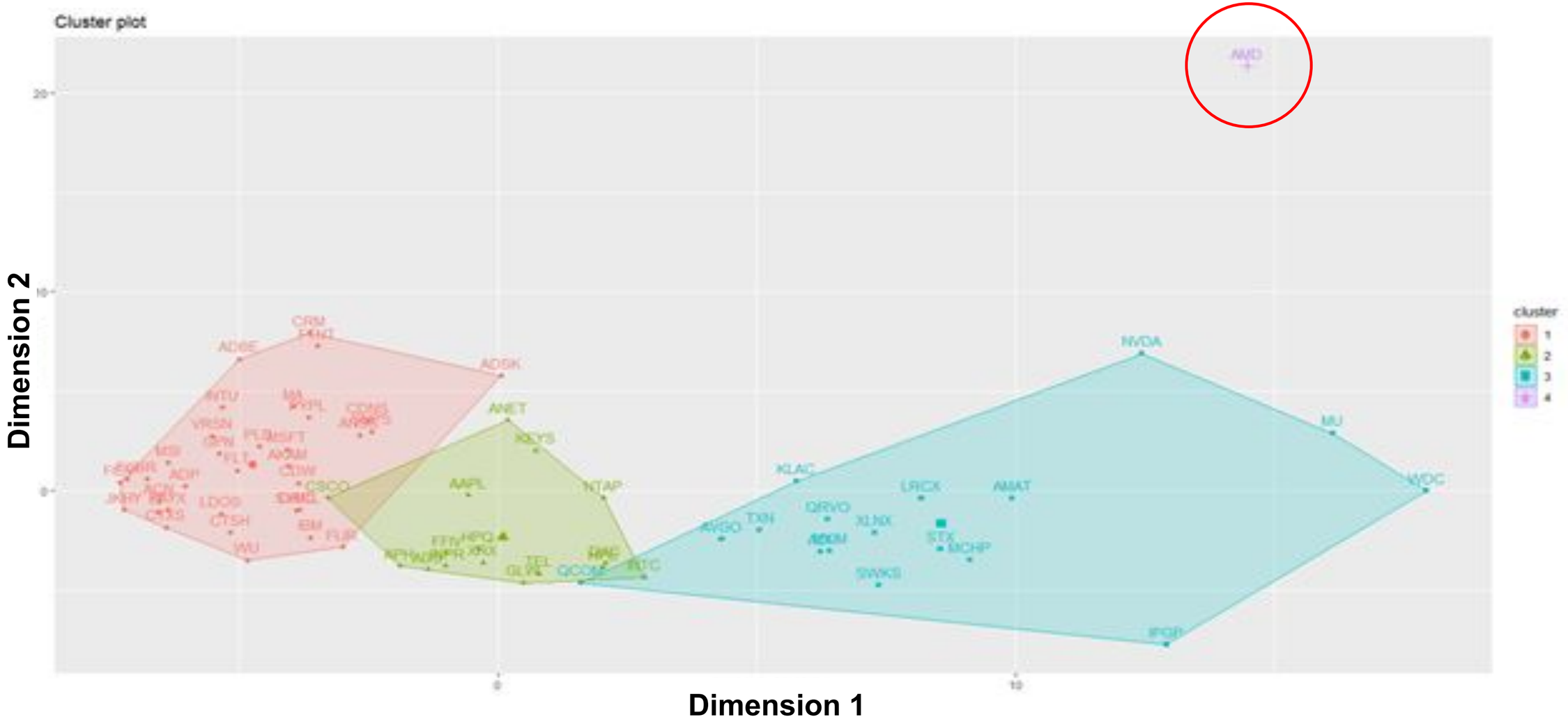
# Cluster Plot of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Cluster Plot of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

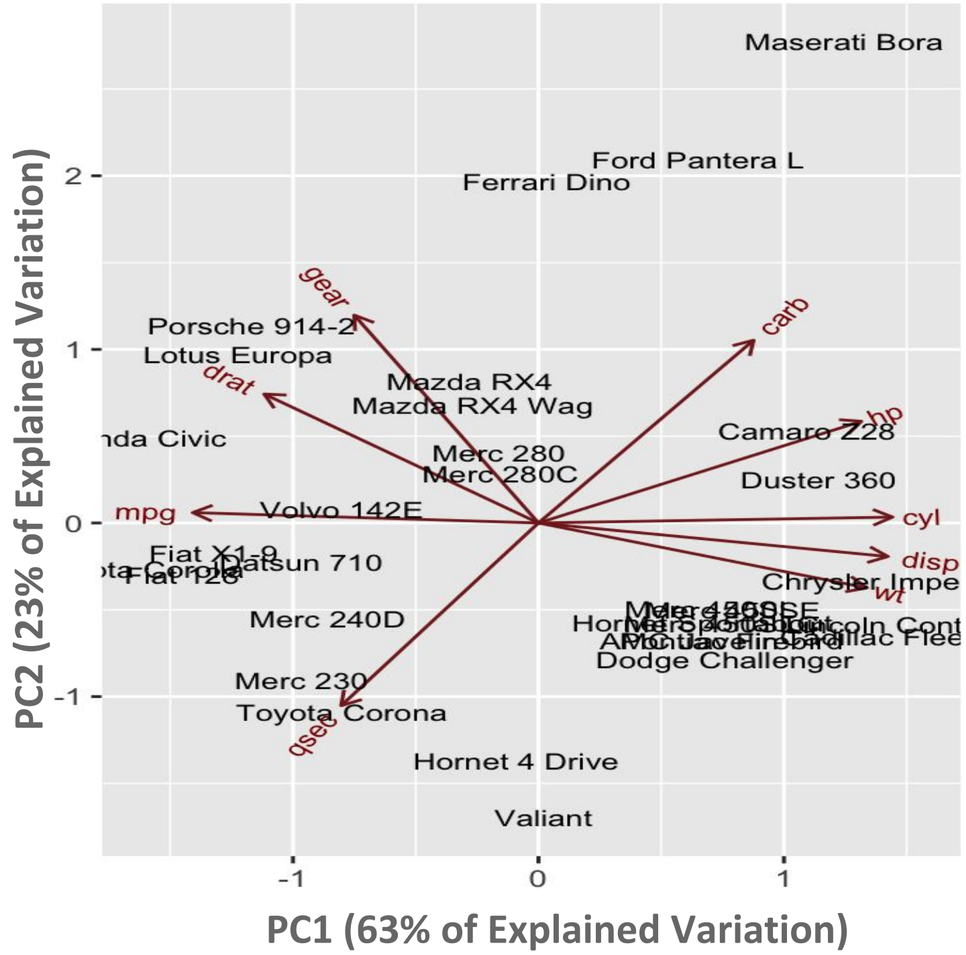
# Agenda

Use correlation to select pairs

Identify hierarchical clusters of stocks and create a cluster plot

Construct a scree plot for stocks within an industry sector

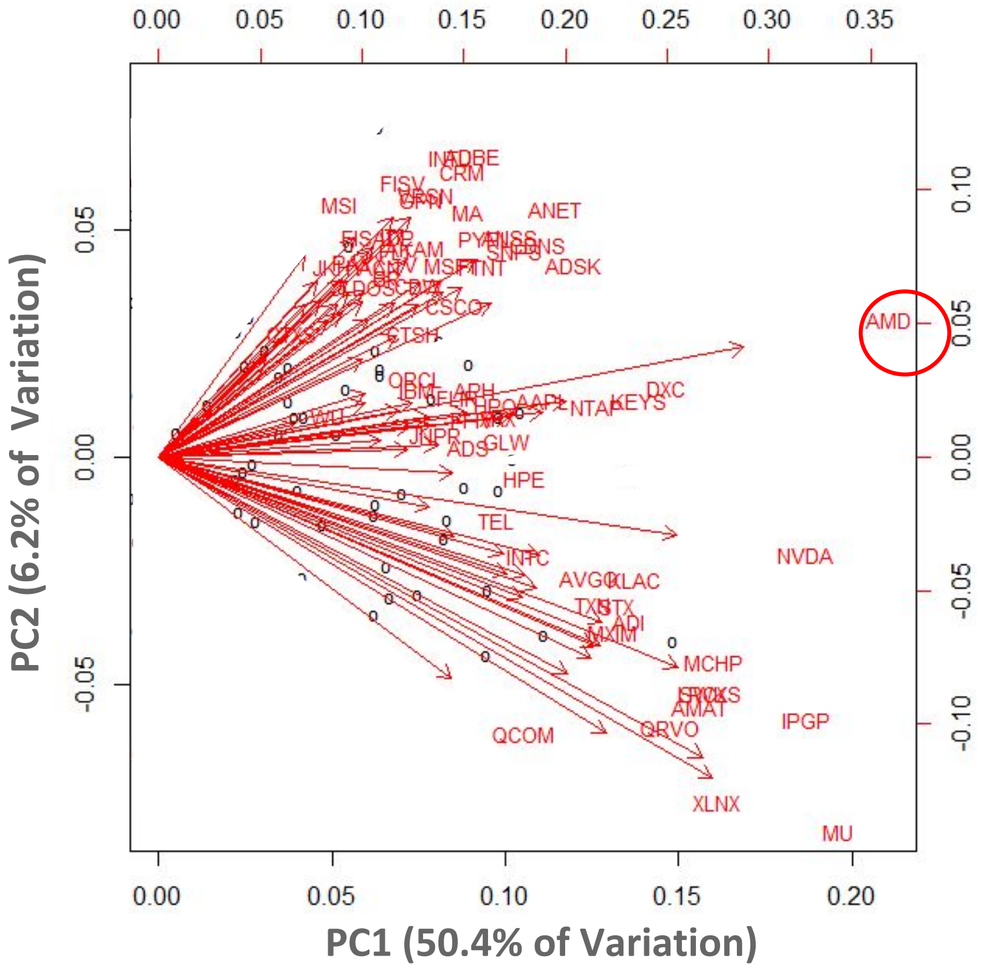
# Principal Components




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Principal Components




Google Cloud

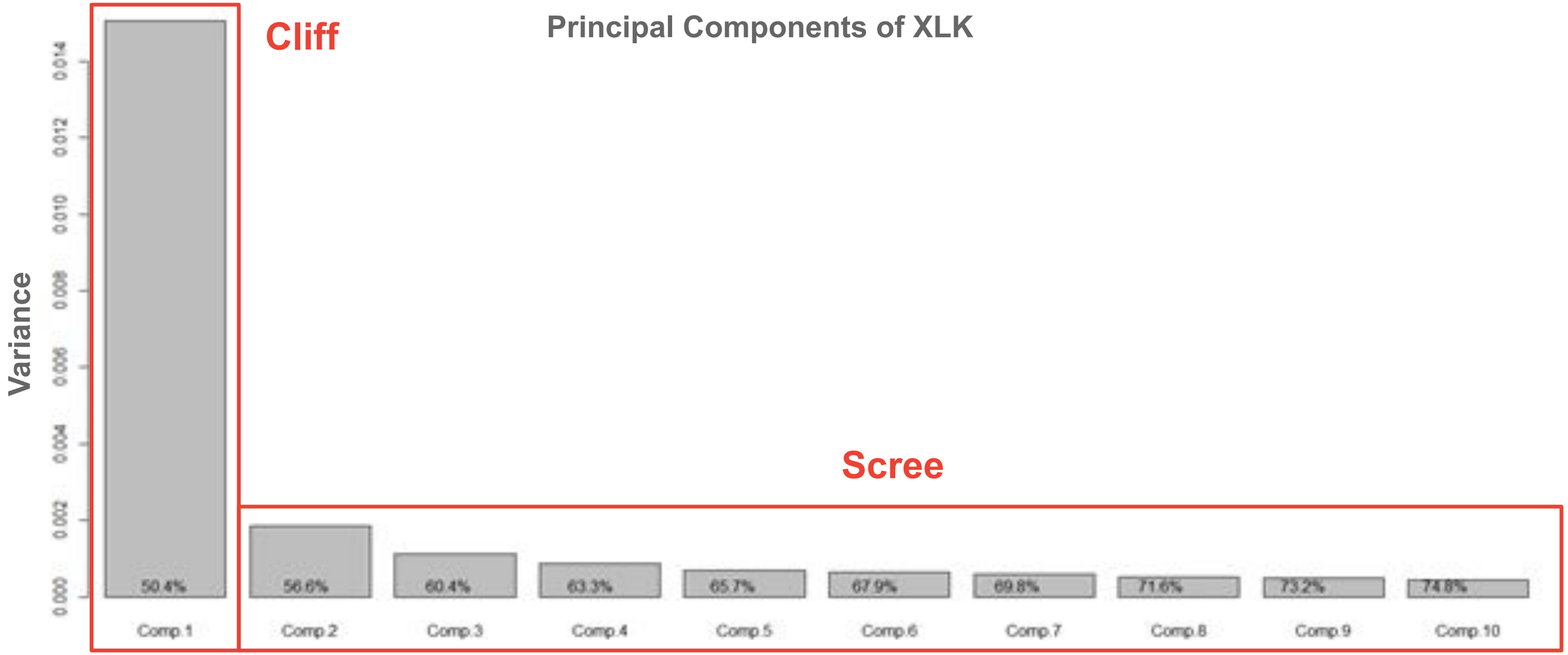
ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Scree Plots

Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

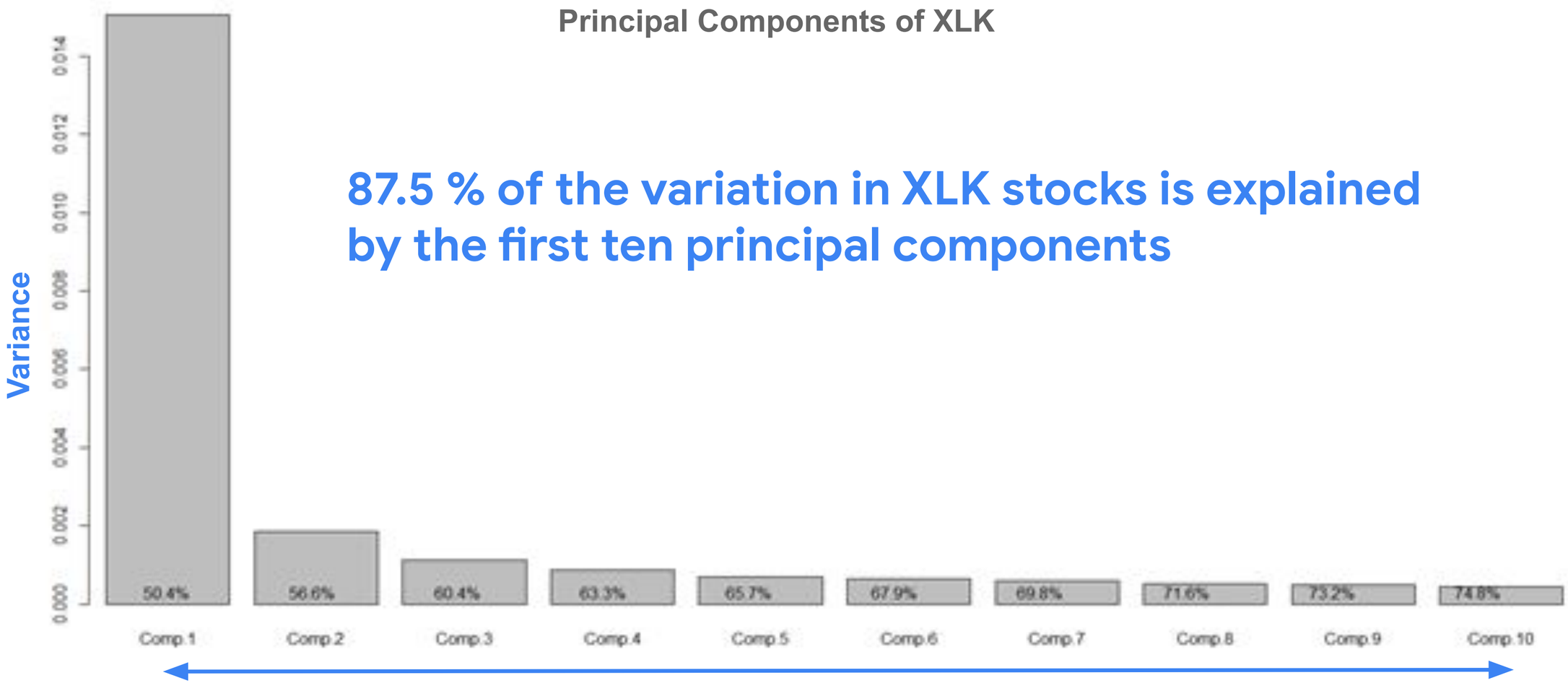
# Scree Plot for Principal Components of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Scree Plot for Principal Components of XLK




**87.5 % of the variation in XLK stocks is explained by the first ten principal components**

Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

ESTABLISHED 1922

NEW YORK INSTITUTE OF FINANCE

# How to Implement a Pair Strategy

# Learning Objectives

- Understand what the loadings in a PCA indicate
- Use PCA loadings for the first two components to select a pair
- Design a beta weighted pair strategy for two stock in SLK

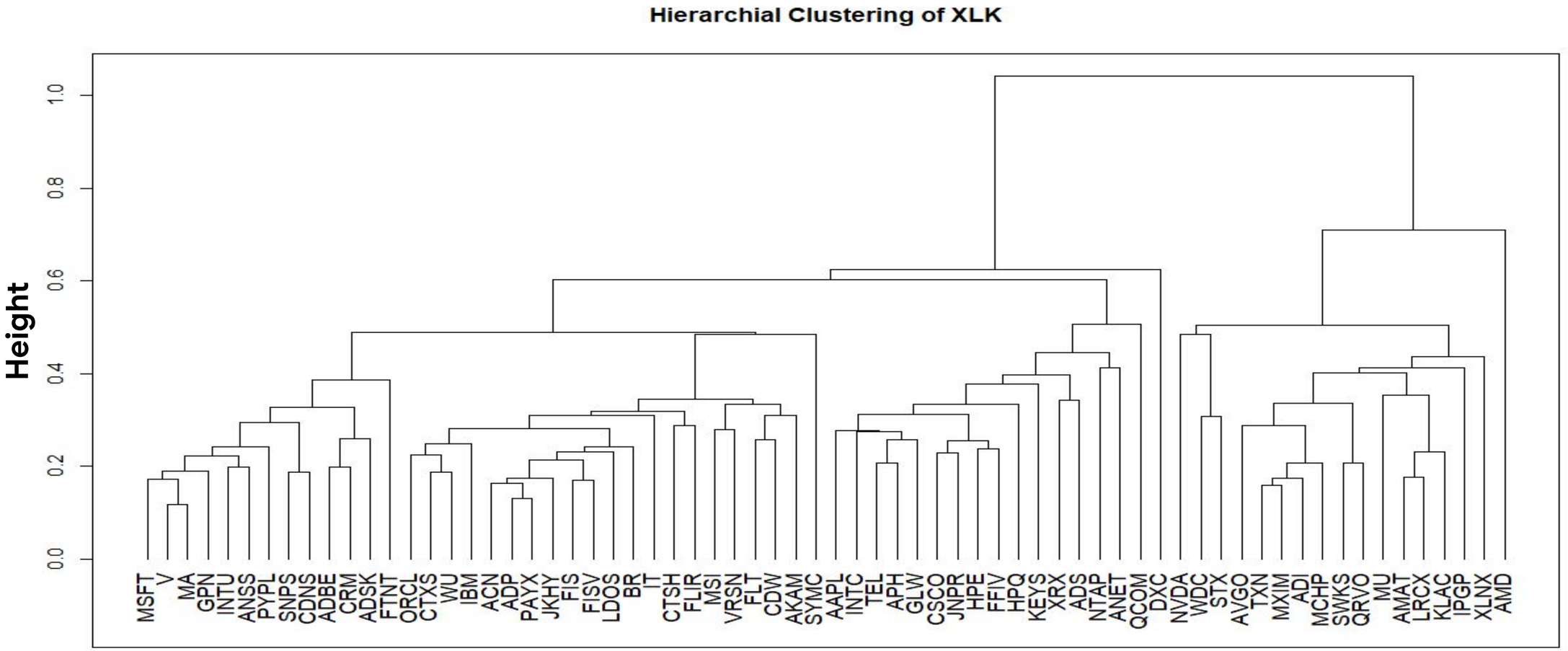
# Agenda

Understand the significance of the first two principal components

Use PCA component loadings to select pairs

Construct a beta-weighted pair

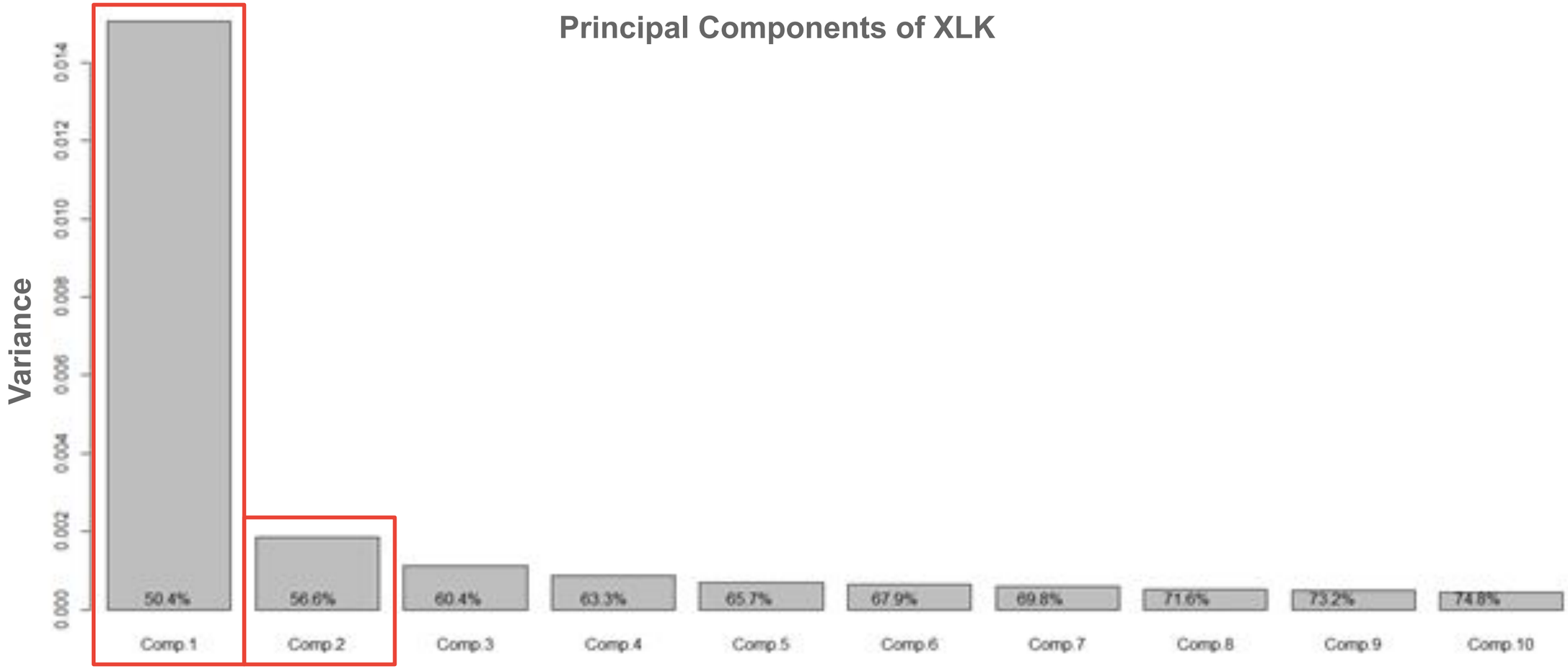
# Hierarchical Cluster Plot of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

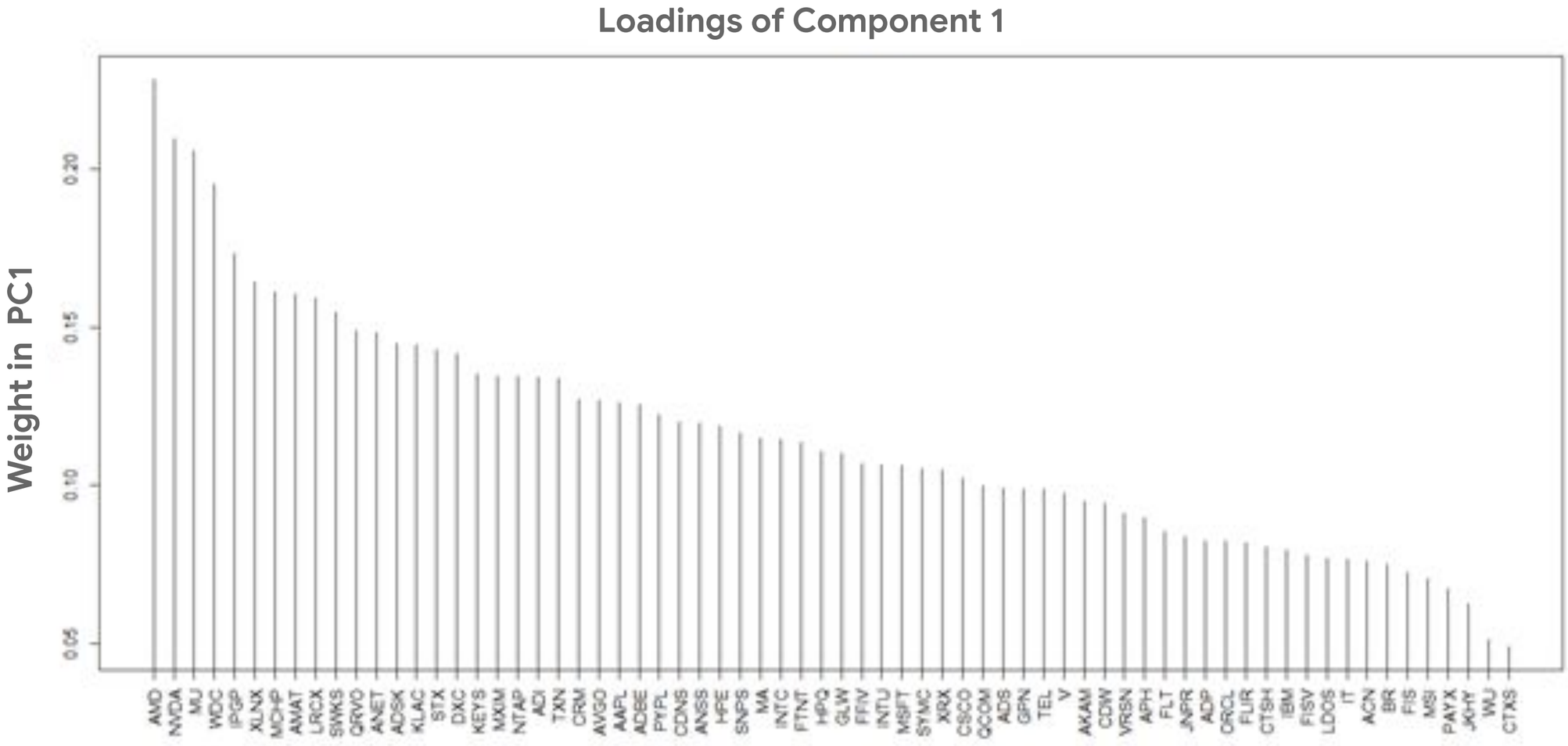
# Scree Plot for Principal Components of XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

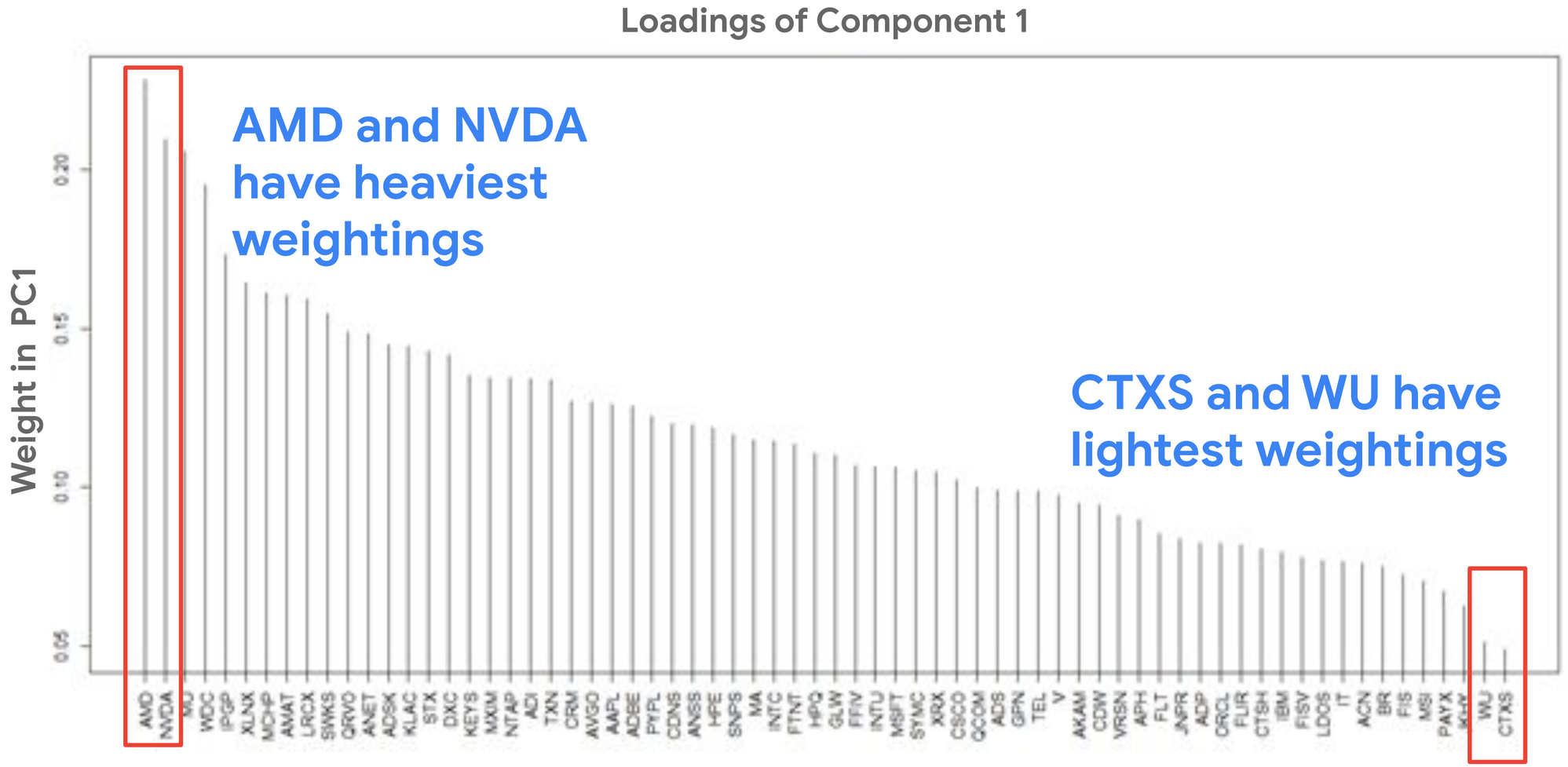
# Loadings of PCA Component 1 for XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

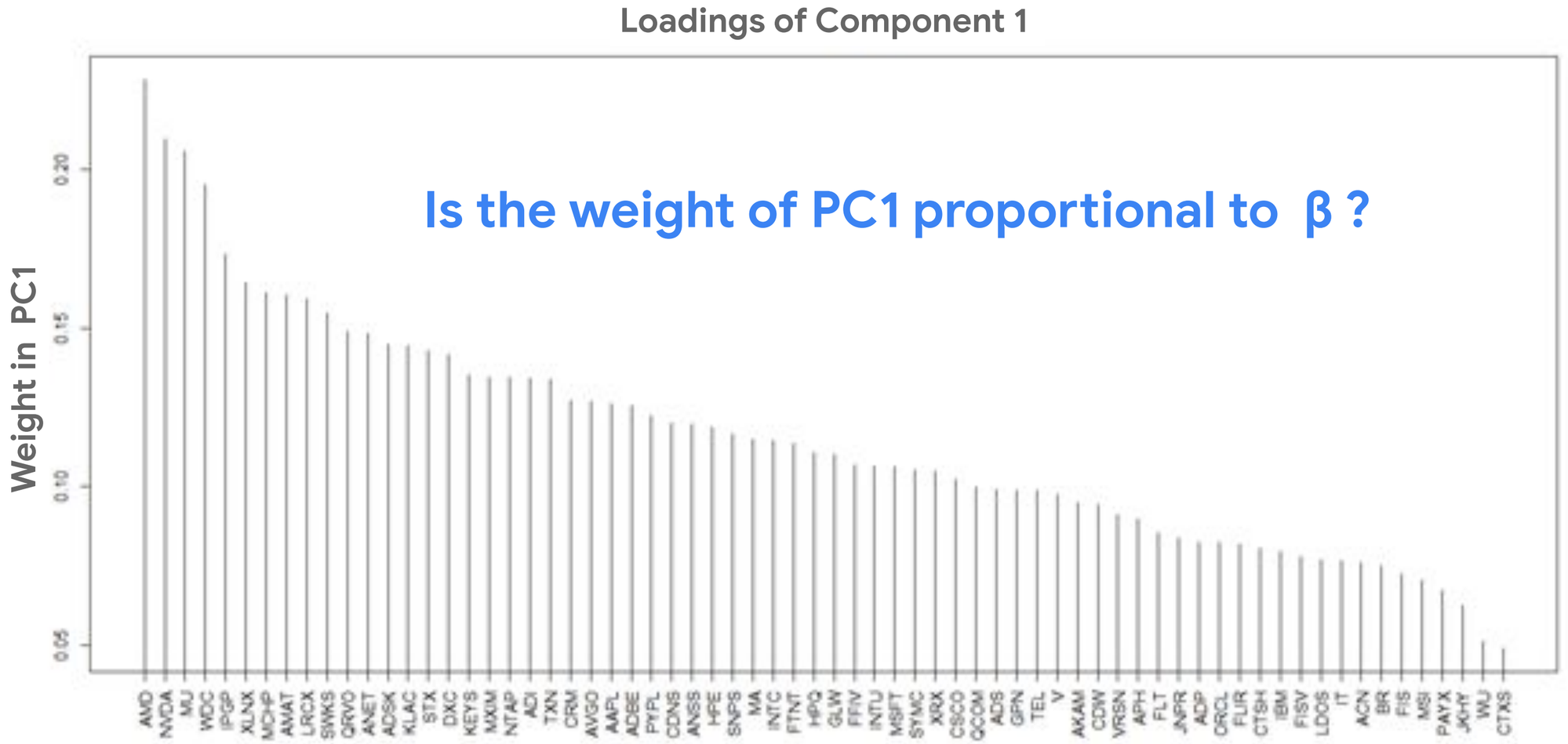
# Loadings of PCA Component 1 for XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Loadings of PCA Component 1 for XLK




**Is the weight of PC1 proportional to β ?**

Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

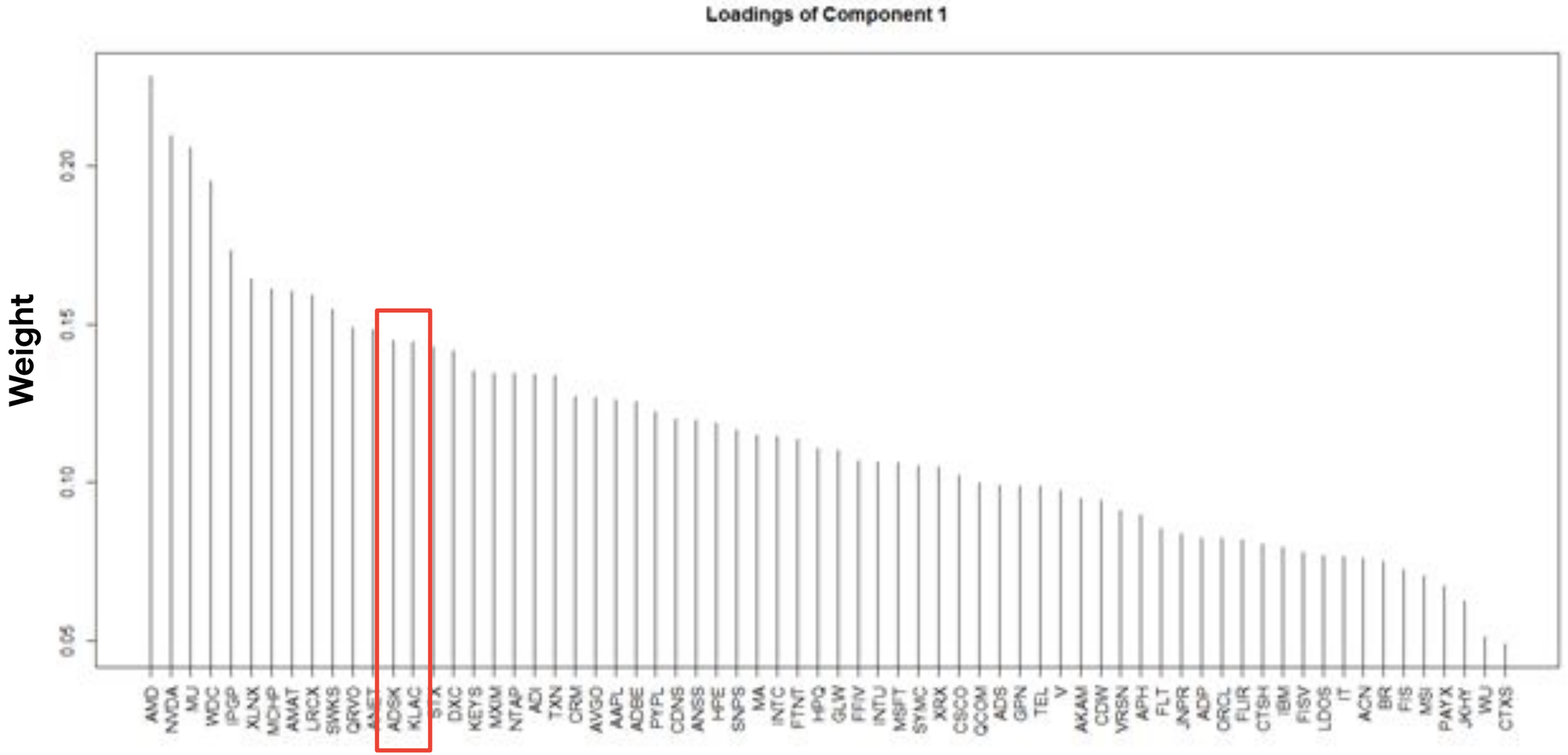
# Agenda

Understand the significance of the first two principal components

Use PCA component loadings to select pairs

Construct a beta-weighted pair

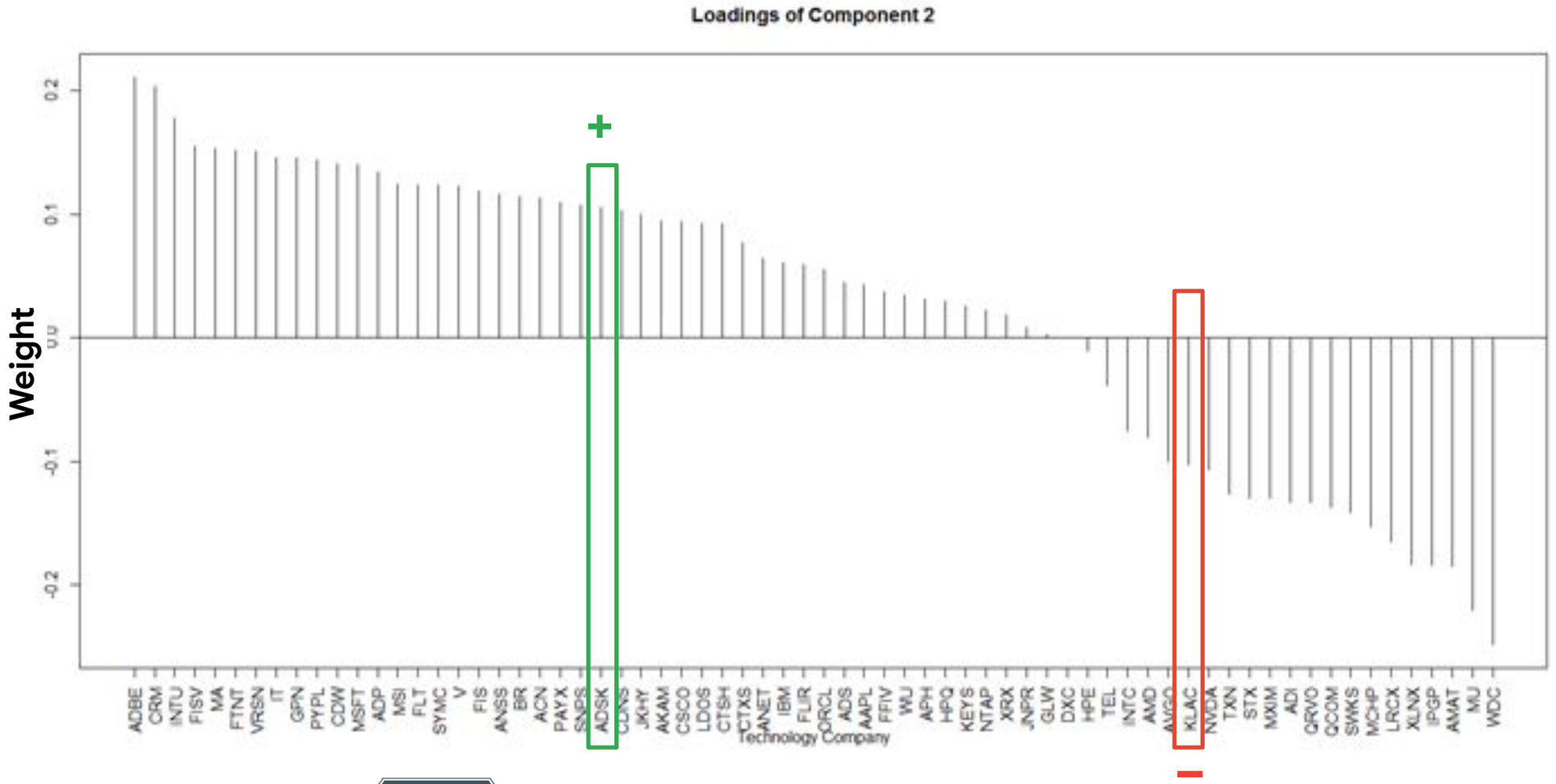
# Loadings of PCA Component 1 for XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

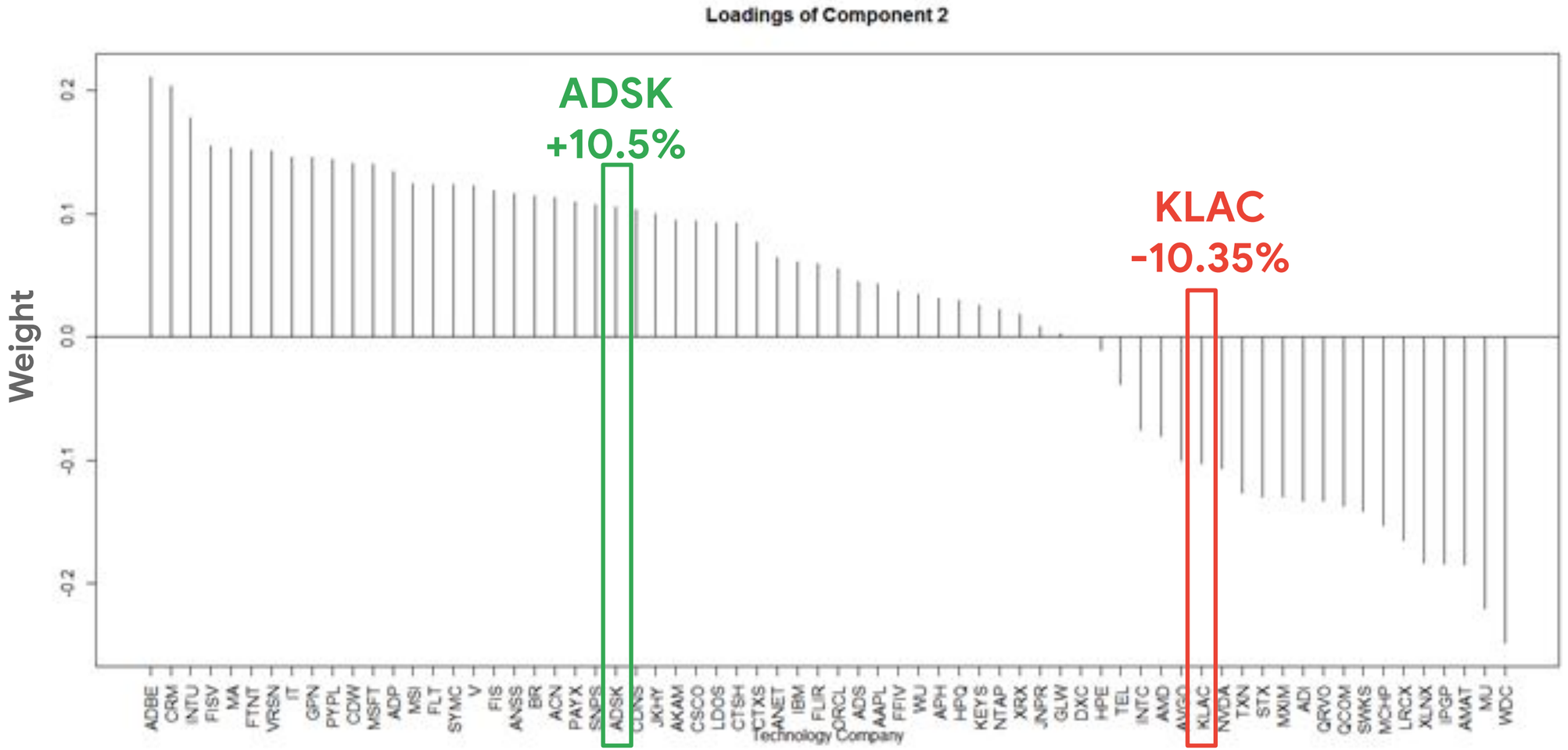
# Loadings of PCA Component 2 for XLK




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

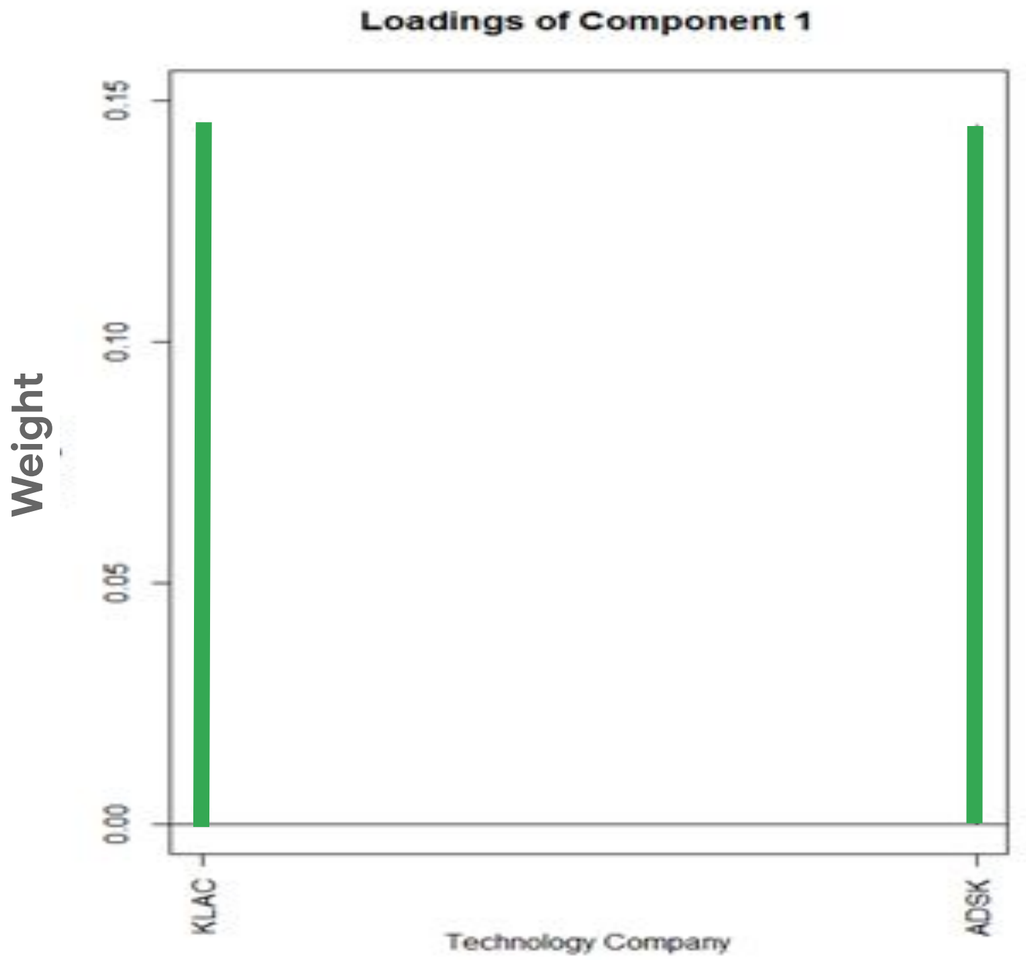
# Loadings of PCA Component 2 for XLK

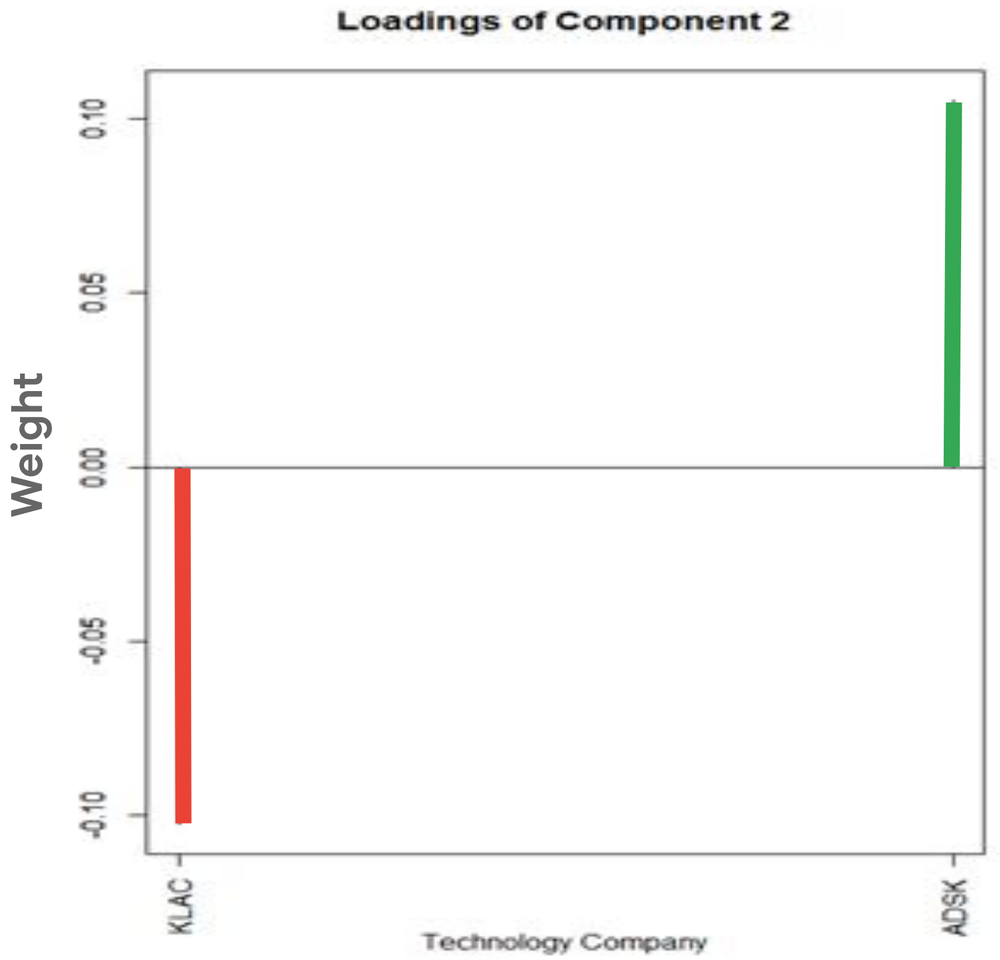



Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Loadings of KLAC and ADSK for Components 1 and 2




Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Agenda

Understand the significance of the first two principal components

Use PCA component loadings to select pairs

Construct a beta-weighted pair

# Weighting Long and Short Components of Pair Trade

- KLAC trades at $170 and has a beta of 1.56
- ADSK trades at $145 and has a beta of 1.41
- Weighting must be Beta Neutral and also factor in relative stock prices.

$$\beta_{KLAC} * P_{KLAC} * 1{,}000 = \beta_{ADSK} * P_{ADSK} * Q_{ADSK}$$

**1,000 shares of KLAC = $170,000**

# Weighting Long and Short Components of Pair Trade

$\beta_{KLAC} * \$170{,}000 / \beta_{ADSK} = P_{ADSK} * Q_{ADSK}$

We need 1.56 * $170,000 / 1.41 = $188,085 worth of ADSK

At a share price of $145, you would need about $188,085/145 = 1297 shares of ADSK.

Rounding, we'll trade 1,300 shares of ADSK

# Weighting Long and Short Components of Pair Trade

1.56 * 1000 sh * $170/sh = 265,200

1.41 * 1300 sh * $145/sh = $265,785

Long 1000 KLAC

Short 1300 ADSK

# Pair Strategy

**When do you get in?**

**How do you define your entry signal?**

Long 1000 KLAC

Short 1300 ADSK

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Evaluate Results of a Pair Trade

# Learning Objectives

- Construct a trade blotter to record entry and exit data
- Quantify and implement a trading rule
- Identify six key performance metrics for pairs trading

# Agenda

Construct a Trade Blotter

Quantify and Implement a Trading Rule

Evaluate Performance using Six Trading Metrics

# Trade Blotter Data

- The pair you trade
- The direction of the trade
- The entry date
- The entry level
- The exit level
- The exit category

# Trade Blotter Data

The pair you trade

**The direction of the trade**

The entry date

The entry level

The exit level

The exit category

# Trade Blotter Data

The pair you trade

The direction of the trade

The entry date

The entry level

The exit level

The exit category

# Trade Blotter Data

The pair you trade

The direction of the trade

The entry date

**The entry level**

The exit level

The exit category

# Trade Blotter Data

The pair you trade

The direction of the trade

The entry date

The entry level

**The exit level**

The exit category

# Trade Blotter Data

The pair you trade

The direction of the trade

The entry date

The entry level

The exit level

**The exit category**

# Agenda

Construct a Trade Blotter

**Quantify and Implement a Trading Rule**

Evaluate Performance using Six Trading Metrics

# Specifying Pairs using PCA

1. Use second component loadings to choose viable pairs for trading
2. Update PCA loading based on new data
3. Maintain that existing pair is still supported
4. Otherwise choose new pair based on current loadings

# Mean-Reverting Strategy

- Assume KLAC and ADSK will have similar long-term performance
- Sell pair member that has outperformed
- Buy pair member that has underperformed
- Wait for spread to converge

# Quantifying Your Trading Rule

- Typically, you will have an entry signal, a profit target, and a stop-loss
- If KLAC - ADSK > $30, enter the trade
- Your profit target is when KLAC - ADSK < $25
- Your stop loss is when KLAC - ADSK > $35
- Your time out will be one month

# Quantifying Your Trading Rule

- Typically, you will have an entry signal, a profit target, and a stop-loss
- If KLAC - ADSK > $30, enter the trade
- Your profit target is when KLAC - ADSK < $25
- Your stop loss is when KLAC - ADSK > $35
- Your time out will be one month

# Quantifying Your Trading Rule

- Typically, you will have an entry signal, a profit target, and a stop-loss
- If KLAC - ADSK > $30, enter the trade
- Your profit target is when KLAC - ADSK < $25
- Your stop loss is when KLAC - ADSK > $35
- Your time out will be one month

# Quantifying Your Trading Rule

- Typically, you will have an entry signal, a profit target, and a stop-loss
- If KLAC - ADSK > $30, enter the trade
- Your profit target is when KLAC - ADSK < $25
- Your stop loss is when KLAC - ADSK > $35
- Your time out will be one month

# Quantifying Your Trading Rule

- Typically, you will have an entry signal, a profit target, and a stop-loss
- If KLAC - ADSK > $30, enter the trade
- Your profit target is when KLAC - ADSK < $25
- Your stop loss is when KLAC - ADSK > $35
- Your time out will be one month

# Agenda

Construct a Trade Blotter

Quantify and Implement a Trading Rule

**Evaluate Performance using Six Trading Metrics**

# Backtest and Evaluate Trading Rule

- Apply trading rules to historical data set
- Create hypothetical trade blotter to record results of trading rule
- Measure hypothetical performance of trading rule

# Key Performance Metrics

1. Win Percentage
2. Total Return
3. Average Win/Loss Amounts
4. Extreme Win/Loss Amounts
5. Variance of Returns
6. Total Portfolio Value

# Performance Metrics: Win Percentage

- Hit Target = Win
- Hit Stop Loss - Loss

  **Win % = [Wins/Losses]*100**

- 75-80% puts odds in your favor and covers trading costs

# Performance Metrics: Total Return

- Need to calculate total invested capital and how much was gained or lost when the trade was closed
- Remember to subtract trading costs such as fees and commissions
- Market impact costs will be reflected in the prices at which you transact

# Performance Metrics: Average Win / Loss Amounts

- Average is for performance of both sides of pair combined
- PnL distribution can be asymmetric
- **Average win is $1** but .... **average loss is $10**
- You **exit trade profitably 90%** of the time **but still make a loss overall**

# Performance Metrics: Extreme Win and Loss Amounts

- PnL distribution can have have a **“long fat tails”**
- Significant probability of an **extreme gain or loss** that accounts for most of the expected gain/loss
- Implies **low confidence** in averages

# Performance Metrics: Variance of Returns

- High variance/volatility of returns also implies low confidence in expected returns
- Your alpha may have a low **Information Ratio (IR)**
- Skilled vs Lucky? **IR** penalizes luck

$$IR = alpha / volatility_{return}$$

# Performance Metrics: Track Total Portfolio Value

- Track linear trend of portfolio value against market value
- Calculate correlation
- Steady, stable trend preferred to wild oscillations even if less profitable

# Performance Metrics: Track Total Portfolio Value

- Prefer a return that is not “bumpy”
- Can measure bumpiness by calculating the second derivative (in discrete time) of portfolio value and then integrating it numerically

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

# Backtesting and Avoiding Overfitting

# Learning Objectives

- Identify the steps required to create a pairs trading model
- Describe a method to modify pair choices dynamically
- Identify the steps required to backtest a static pairs model

# Agenda

**Backtesting a Pairs Trading Model**

Modify model to Dynamically Update Pair Candidates

Detail Steps in a Static Model Backtest

# Backtesting Static Pairs

1. Backtest KLAC-ADSK pair using several months of data
2. Backtest other potential pairs that have been identified using PCA analysis
3. Rank pairs and perform an implementation backtest on top candidates

# Selecting Optimal Pairs

Selection based on a weighted average of the following criteria.

1. Minimize the distance between stock loadings in Component 1.
2. Minimize the distance between the absolute value of stock loadings in Component 2.
3. Choose two stocks with opposite sign loadings in Component 2.

# Agenda

Backtesting a Pairs Trading Model

**Modify model to Dynamically Update Pair Candidates**

Detail Steps in a Static Model Backtest

# Dynamically Update Principal Components

1. Update principal components periodically by incorporating more recent price data
2. Select optimal pairs using updated PCA analysis

# Dynamic Pair Updating: Practical Considerations

Decision to dynamically update pairs requires backtesting or qualitative analysis to determine optimal:

1. Update frequency
2. Selection criteria for new pairs

# Agenda

Backtesting a Pairs Trading Model

Modify model to Dynamically Update Pair Candidates

Detail Steps in a Static Model Backtest

# Backtest of a Static Example

Backtest incorporates varying three parameters:

1. Entry Signal
2. Profit-taking level
3. Stop Loss

Assume no time outs.

# Combinations of Backtest Parameters

- **50** Different **Entry Signals**
- **50** Different **Profit-taking levels**
- **50** Different **Stop Losses**

**50 x 50 x 50 ⇒ 125,000 potential combinations** requiring efficient software to:

- Produce a trade blotter
- Summarize each blotter with our key metrics

# Divide DataFrame into Training and Testing

Divide available historic price series into two unequal parts:

- Allocate **two thirds of data** for **training** our models
- Allocate **one third of data** for **testing** our models

# Backtesting Potential Parameter Combinations

- Back test each parameter combination/create a trade blotter
- Summarize key metrics:
  1. Total, average and percent win
  2. Expected loss
  3. Volatility
  4. Steadiness of the return
- Define optimal criterion as Sharpe ratio: **Total Return / Volatility**

# Measuring Performance of Selected Models

- Run selected models on test data
- Testing out-of-sample helps to identify overfitting model to training data
- **Poor performance on testing data indicates overfitting**

# Measuring Performance of Selected Models

- Run top 50% of models on test data
- Produce new trade blotter
- Choose best models based on Sharpe Ratio or other risk-adjusted performance criterion

# Compare Best Models from Testing Phase

- We have over 60,000 models to compare
- Select models with best overall performance based on your weighting of performance in the training and testing phase:

**Metric = 33% Training + 67% Testing**

ESTABLISHED 1922

NEW YORK INSTITUTE OF FINANCE

# Next Steps: Improvements to Your Pairs Strategy

# Learning Objectives

- Identify external factors that can affect asset prices
- Understand how external factors can impact trade profitability
- Learn how to incorporate external factors into trading models

# Agenda

Time and Day-of-the Week Effect

Related Markets

Economics and Company Specific Events

Updates to Market Indices

# Time-of-Day Effects

- Market Open
- Market Close
- Intraday drop in trading Volume

# Time-of-Day Effects

- Segregate trades by time of day
- Compare profit metrics for each time slice
- Filter your data and backtest most profitable parts of day
- Run a seperate backtest on excluded data
- Understand when and how your model works

# Time-of-Day Effects

- Segregate trades by time of day
- Compare profit metrics for each time slice
- Filter your data and backtest most profitable parts of day
- Run a seperate backtest on excluded data
- Understand when and how your model works

# Time-of-Day Effects

- Segregate trades by time of day
- Compare profit metrics for each time slice
- Filter your data and backtest most profitable parts of day
- Run a seperate backtest on excluded data
- Understand when and how your model works

# Time-of-Day Effects

- Segregate trades by time of day
- Compare profit metrics for each time slice
- Filter your data and backtest most profitable parts of day
- Run a separate backtest on excluded data
- Understand when and how your model works

# Time-of-Day Effects

- Segregate trades by time of day
- Compare profit metrics for each time slice
- Filter your data and backtest most profitable parts of day
- Run a seperate backtest on excluded data
- Understand **when** and **how** your model works

# Day-of-Week Effects

- Higher volume on some weekdays
- Higher volatility
- Avoid days with low volatility

# Agenda

Time and Day-of-the Week Effect

**Related Markets**

Economics and Company Specific Events

Updates to Market Indices

# Related Markets

- Options markets
- Expire on third Friday of each month
- Is there more volatility on Expiration days?
- Do market participants adjust their trading on Expiration days?

# Agenda

Time and Day-of-the Week Effect

Related Markets

Economics and Company Specific Events

Updates to Market Indices

# Economic Events

- Are economic events part of the rules you use to initiate or close trades?
- Are your results influenced by key economic reports
- Does your risk management model for the volatility that can occur after a “surprise” change in an economic indicator  (adjust stop-loss orders prior to announcement)

# Stock Specific Events

- Quarterly earnings announcements have a large impact on trading volume and volatility
- Announcements for either stock in a pair can trigger outsized gains and losses
- Adjusting risk management model is key

# Agenda

Time and Day-of-the Week Effect

Related Markets

Economics and Company Specific Events

**Updates to Market Indices**

# Updates to Market Indices

- Additions and subtractions to indices have enormous impact on affected assets
- Changes are partially anticipated by market
- Can still have volatility of announcement of change

# Improving Your Pairs Strategy

- Periodic events can be anticipated and modeled
- Infrequent events such as index additions and deletions are more difficult
- Can still have volatility of announcement of change