UBS Capstone Project:
Portfolio Optimization

Ken Chen, Lin Hui, Yu Li, Lukuan Wang, Jingzhao Zhang

Motivation

- Objective: Use quantitative techniques to manage portfolio
- Senario: PMs want to find an alternative stock under current trend
- Focus: find similar stocks to replace the one in portfolio
- Features: User-friendly interface, interactive explanations

Methodology

- Similar company finder with Barra risk factors
- Similar company finder with revenue and sector data
    - Same company finder by path code
    - Similar company finder by revenue segments
- General recommender system research
- Optimize the portfolio

1 Similar company finder with Barra risk factors

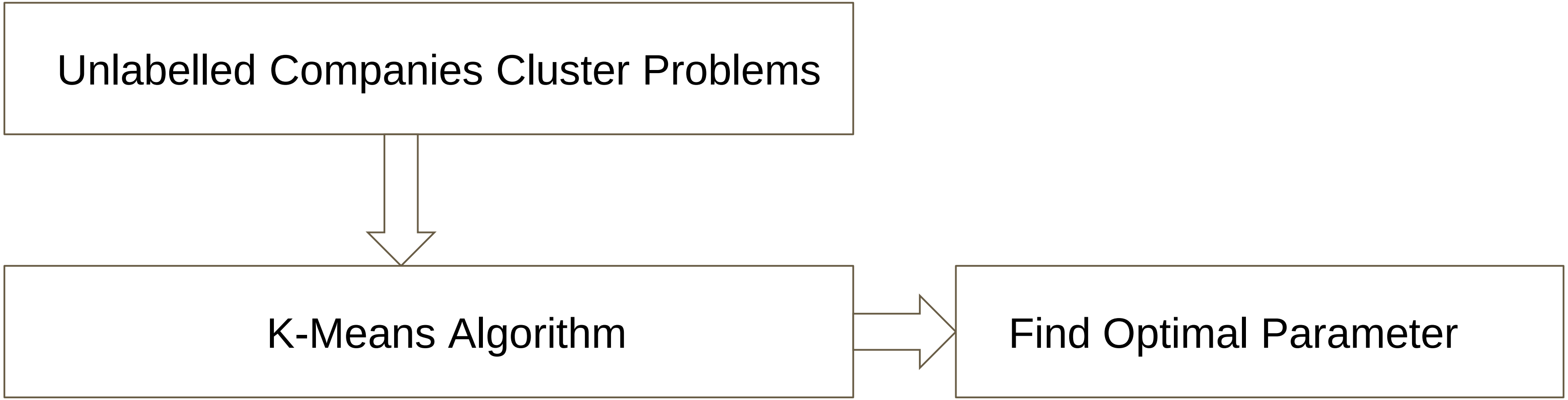
1.1 Barra risk factors data:

	INSTRUMENT_NM	LABEL_NM	MODEL_ID	IMNT_CD	LABEL_ID	VAL	FIRST_DAY	LAST_DAY
85	ACCENTURE PLC CLASS A	MOMENTUM	5	USA4JB1	7205	-0.004	2021-01-01	2021-01-31
235	ACCENTURE PLC CLASS A	GROWTH	5	USA4JB1	7209	-0.276	2021-01-01	2021-01-31
340	ACCENTURE PLC CLASS A	LEVERAGE	5	USA4JB1	7212	-0.751	2021-01-01	2021-01-31
270	ACCENTURE PLC CLASS A	DIVYILD	5	USA4JB1	7210	0.008	2021-01-01	2021-01-31
515	ACCENTURE PLC CLASS A	USAC	5	USA4JB1	7372	1.000	2021-01-01	2021-01-31

LABEL_NM	BANKS	BETA	BIOTECH	BTOP	CAPGOODS	CHE	CHEC	CHEMICAL	COMMSVCS	COMMUNIC	...	RETAIL	SEMICOND	SIZE	SIZENL
INSTRUMENT_NM															
ACCENTURE PLC CLASS A	0.0	0.325	0.0	-0.409	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.579	-0.364
AMERICAN ELECTRIC POWER	0.0	-0.576	0.0	0.656	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	-0.383	0.513
AMGEN CORPORATION	0.0	-0.457	1.0	-0.465	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.407	-0.223

1 Similar company finder with Barra risk factors

1.2 Research Procedures



1 Similar company finder with Barra risk factors

1.3 Find optimal Parameters(k)-------Efficient Turning Point

Loss Function Inertia(*I*)

Fixed K

Unique loss (*I*_k)

Inertia(*I*):

$$\sum_{S_i}\sum_{x\in S_i}|x-u_i|^2$$

1 Similar company finder with Barra risk factors

1.3 Efficient Turning Point(Continued)

Minimize Inertia ?

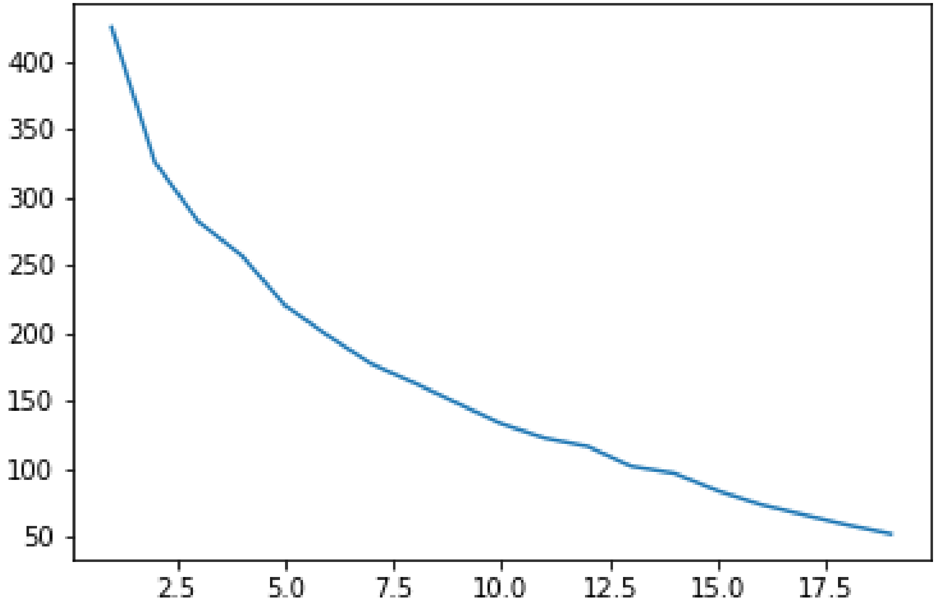
Wrong !

$$argmin_k \sum_{S_i} \sum_{x \in S_i} |x - u_i|^2$$

Find the K with greatest marginal contribution！ √

1 Similar company finder with Barra risk factors

- 1.4 Inertia Graph



Best K candidates: 3,4,5

1 Similar company finder with Barra risk factors

1.5 Example:Find similar stocks using Euclidean distance, for Microsoft

```
0  :  ['ACCENTURE PLC CLASS A', 'AMGEN CORPORATION', 'CISCO SYSTEMS INC',
'COCA COLA', 'DIAGEO PLC ADR', 'HOME DEPOT', 'JOHNSON & JOHNSON', 'LOCKHE
ED MARTIN CORP', 'MCDONALDS CORP', 'MICROSOFT CORP.', 'NEXTERA ENERGY IN
C', 'NOVARTIS AG ADRS', 'PROCTER & GAMBLE CO', 'TEXAS INSTRS INC', 'UNION
PAC CORP', 'UPS']

2  :  ['AMERICAN ELECTRIC POWER', 'ANALOG DEVICES INC', 'AUTOMATIC DATA P
ROCESSING INC', 'MARSH & MC LENNAN', 'REPUBLIC SERVICES', 'ROCKWELL AUTOM
ATION COM U$1', 'V F CORP']

3  :  ['BLACKROCK INC', 'CHUBB LTD', 'COMCAST CORP COM CLASS A', 'INTEL C
ORP', 'JPMORGAN CHASE & COMPANY', 'LINDE PLC', 'MEDTRONIC PLC']

1  :  ['CRANE CO', 'EOG RESOURCES', 'PHILLIPS 66', 'RAYTHEON TECHNOLOGIES
CORP', 'TRUIST FINANCIAL CORP']
```

→ 'ACCENTURE PLC CLASS A'

1 Similar company finder with Barra risk factors

1.5 Example: Factor filter and weights

```
Weight of first factor: .4
Weight of second factor: .3
Weight of thrid factor: .2
Weight of forth factor: .1
```

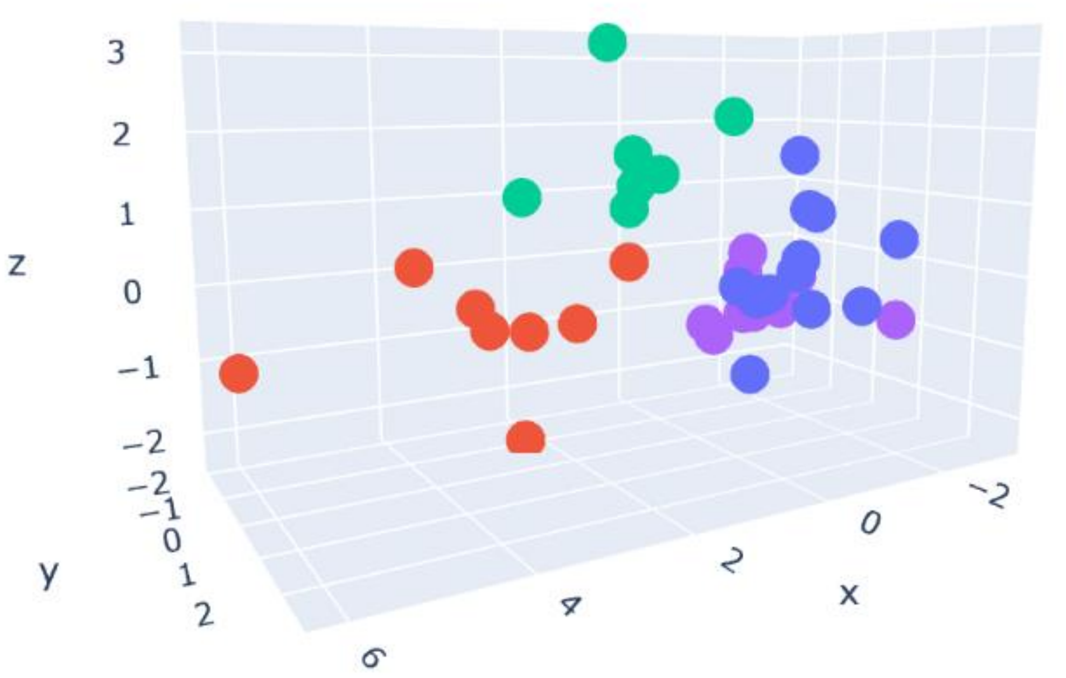
INSTRUMENT_NM	BANKS	BETA	BTOP	SIZE
ACCENTURE PLC CLASS A	0.0	-0.010161	-0.128951	0.050604
AMERICAN ELECTRIC POWER	0.0	-0.370197	0.189940	-0.101675
AMGEN CORPORATION	0.0	-0.322645	-0.145719	0.023378
ANALOG DEVICES INC	0.0	0.060567	-0.033433	-0.072074
AUTOMATIC DATA PROCESSING INC	0.0	-0.089681	-0.149611	-0.037566
BLACKROCK INC	0.0	0.146480	0.038130	0.002483

```
0  :  ['ACCENTURE PLC CLASS A', 'AMGEN CORPORATION', 'BLACKROCK INC', 'CI
SCO SYSTEMS INC', 'COCA COLA', 'COMCAST CORP COM CLASS A', 'DIAGEO PLC AD
R', 'HOME DEPOT', 'INTEL CORP', 'JOHNSON & JOHNSON', 'JPMORGAN CHASE & CO
MPANY', 'LINDE PLC', 'LOCKHEED MARTIN CORP', 'MCDONALDS CORP', 'MEDTRONIC
PLC', 'MICROSOFT CORP.', 'NEXTERA ENERGY INC', 'NOVARTIS AG ADRS', 'PROCT
ER & GAMBLE CO', 'TEXAS INSTRS INC', 'UNION PAC CORP', 'UPS']

1  :  ['AMERICAN ELECTRIC POWER', 'ANALOG DEVICES INC', 'AUTOMATIC DATA P
ROCESSING INC', 'MARSH & MC LENNAN', 'REPUBLIC SERVICES', 'ROCKWELL AUTOM
ATION COM U$1', 'V F CORP']

3  :  ['CHUBB LTD', 'EOG RESOURCES', 'PHILLIPS 66', 'RAYTHEON TECHNOLOGIE
S CORP', 'TRUIST FINANCIAL CORP']

2  :  ['CRANE CO']
```

```
'COCA COLA'
```

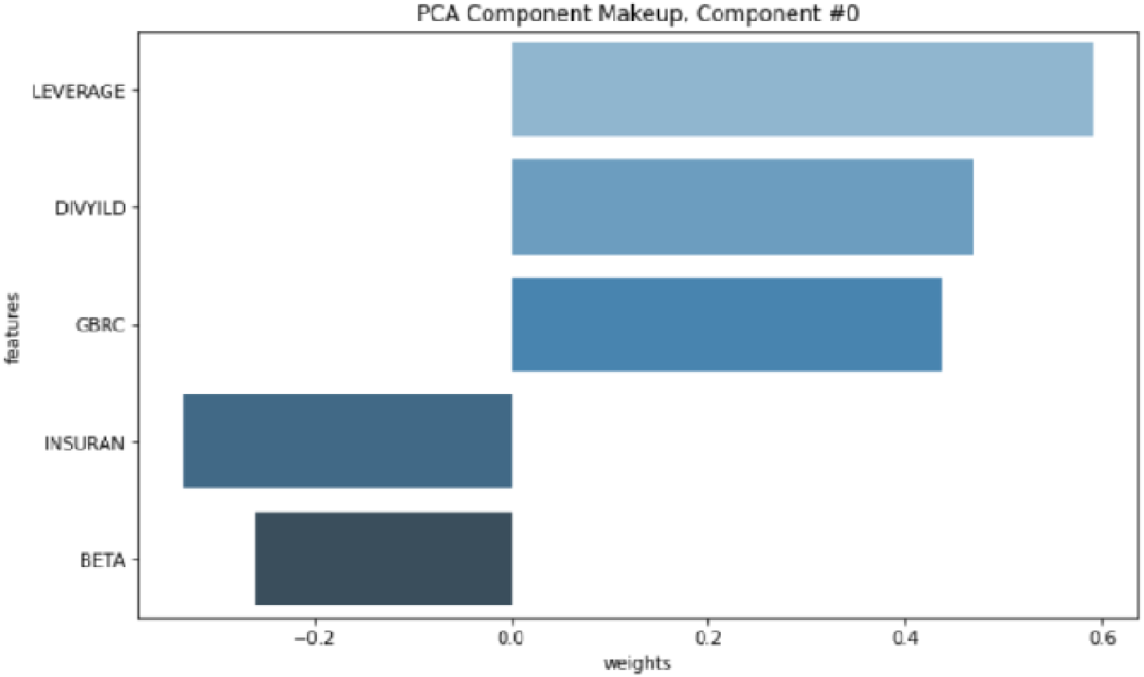
1 Similar company finder with Barra risk factors

1.5 Example: Results Visualization(3D)



1 Similar company finder with Barra risk factors

1.5 Example:PCA Decomposition(First dimension Visualization)



2 Similar company finder with revenue and sector data

Goal: For a given large cap company (source dataset), find the similar small cap company (target dataset) which operate in the same/similar category

ticker	issuername	SEGMENT_PERCENT	PATH	DEPTH	PATH_CODE
AAPL-US	Apple Inc.	54.725299	Technology > Hardware > Communications Equipment > Wireless Equipment	4	20111115
AAPL-US	Apple Inc.	54.725299	Technology > Hardware > Communications Equipment > Wireless Equipment > Mobile > Cellular Phones > Smartphones	7	20111115111012
AAPL-US	Apple Inc.	54.725299	Technology > Hardware > Communications Equipment	3	201111
AAPL-US	Apple Inc.	54.725299	Technology > Hardware > Communications Equipment > Wireless Equipment > Mobile > Cellular Phones	6	201111151110
AAPL-US	Apple Inc.	54.725299	Technology > Hardware > Communications Equipment > Wireless Equipment > Mobile	5	2011111511
005930-KR	Samsung Electronics Co., Ltd.	41.561134	Technology > Hardware > Communications Equipment > Wireless Equipment	4	20111115
005930-KR	Samsung Electronics Co., Ltd.	41.561134	Technology > Hardware > Communications Equipment > Wireless Equipment > Mobile > Cellular Phones > Smartphones	7	20111115111012
005930-KR	Samsung Electronics Co., Ltd.	41.561134	Technology > Hardware > Communications Equipment	3	201111
005930-KR	Samsung Electronics Co., Ltd.	41.561134	Technology > Hardware > Communications Equipment > Wireless Equipment > Mobile > Cellular Phones	6	201111151110

- Segment percent: combined revenue percentage of a certain path
- Path: the revenue category which the sector operates in
- Depth: how deep the path goes
- Path code: represents the path in a unique way

2.1 Same company by path code

Search in target companies using path code to find the exact match

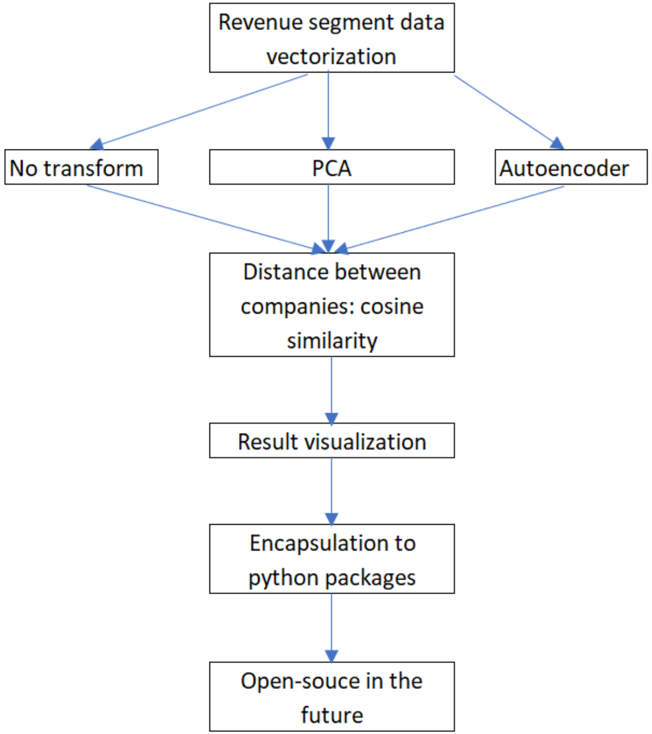
CSCO-US	Technology > Hardware > Communications Equipment
CSCO-US	Technology > Hardware > Communications Equipment > Wide Area Networking (WAN)

↓

```
2498-TW  :  Communications Equipment
088800-KR  :  Communications Equipment
2342-HK  :  Communications Equipment
050890-KR  :  Communications Equipment
ASCN-CH  :  Communications Equipment
CEL-IL  :  Communications Equipment
NTGR-US  :  Communications Equipment
6703-JP  :  Communications Equipment
EXTR-US  :  Wide Area Networking (WAN)
HLIT-US  :  Communications Equipment
5388-TW  :  Communications Equipment
ADTN-US  :  Wide Area Networking (WAN)
3681-JP  :  Communications Equipment
DGII-US  :  Communications Equipment
3596-TW  :  Wide Area Networking (WAN)
ATEN-US  :  Wide Area Networking (WAN)
3380-TW  :  Communications Equipment
6676-JP  :  Communications Equipment
```

2.2 Similar company by revenue segments

Search in target companies using <u>path, depth and segment percent</u> to find the similar match

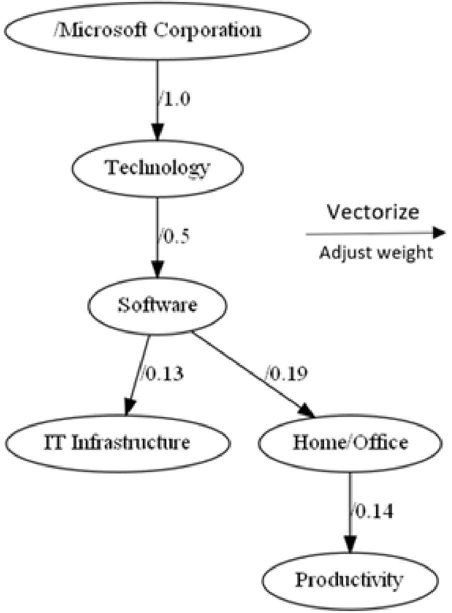


2.2 Similar company by revenue segments

Vectorization

$$\text{revenue percent of depth i adjusted for weights} = \frac{\text{revenue percent of depth i}}{1 + \text{i}}$$

IssuerName	Microsoft Corporation	Microsoft Corporation	Microsoft Corporation
Segment Percent	28.93333	24.69391	40.28249
Segment Name	IT Infrastructure	Productivity	Home/Office
Path	Technology > Software > IT Infrastructure	Technology > Software > Home/Office > Productivity	Technology > Software > Home/Office



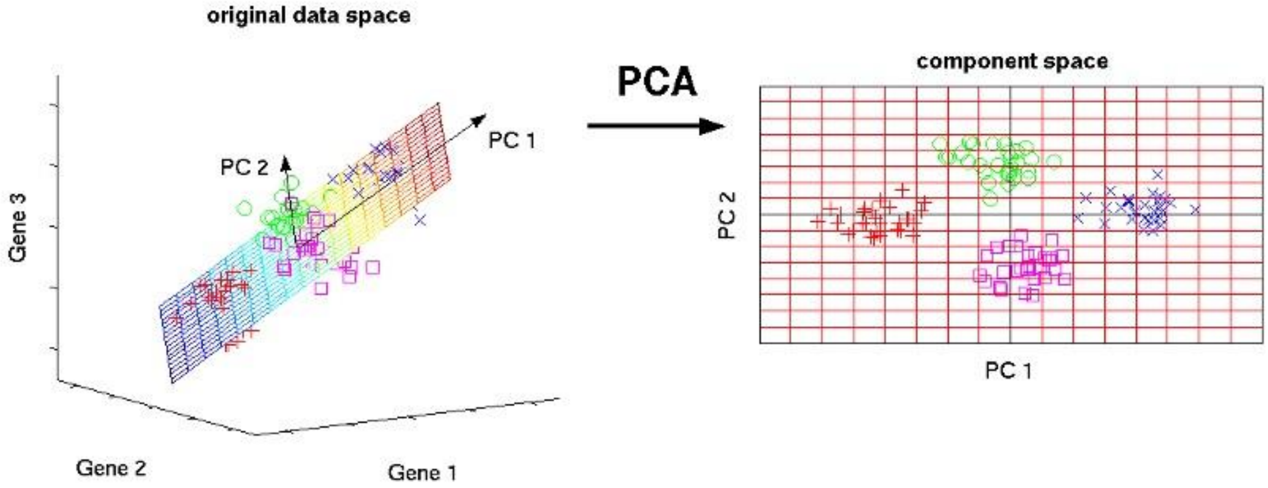
depth	name	Microsoft Corporation
0	Technology	1.00
1	Hardware	0.00
	Software	0.50
2	Communications Equipment	0.00
	Home/Office	0.19
	IT Infrastructure	0.14
3	Business Communications Equipment	0.00
	Games	0.00
	Handheld/Smart Phone	0.00
	Network	0.00
	Productivity	0.15
	System Software	0.00
4	Applications	0.00
	Communications	0.00
	Conferencing Equipment	0.00
	Desktop Security	0.00
	Network Security	0.00
	Online Games	0.00
	PC Games	0.00
5	Conferencing/Collaboration Tools	0.00
	Games	0.00
	Puzzle	0.00
	Video	0.00
6	Online Conference/Meeting	0.00
	Server/Infrastructure Equipment	0.00

2.2 Similar company by revenue segments

PCA for dimension reduction:

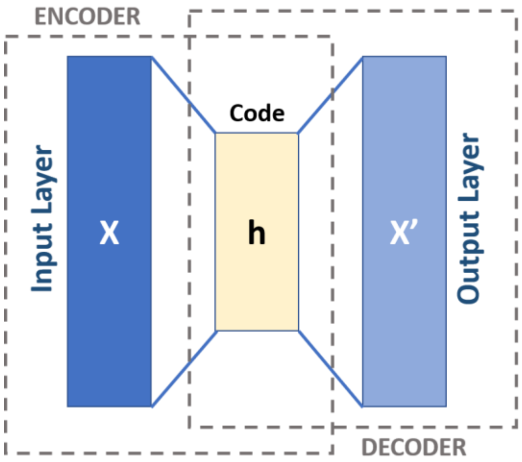
Automatically projects each data point onto only the first few components to obtain lower-dimensional data.

In our model: 1605 components → 500 components



2.2 Similar company by revenue segments

Autoencoder for dimension reduction:



depth	name	AB Dynamics plc	AB Science SA	ABC arbitrage SA
0	Business and Public Services	0	0	0
0	Consumer	0	0	0
0	Energy	0	0	0
0	Finance	0	0	1
0	Healthcare	0	1	0
0	Industrial and Materials	1	0	0
0	Technology	0	0	0
1	Banking	0	0	0
1	Biopharmaceuticals	0	0	0
1	Business Services	0	0	0
1	Consumer Products and Services	0	0	0
1	Downstream	0	0	0
1	Electronic Components	0	0	0
1	Hardware	0	0	0
1	Healthcare Services	0	0	0
1	Hospitality	0	0	0
1	Industrial Manufacturing	0.5	0	0

	AB Dynamics plc	AB Science SA	ABC arbitrage SA
0	-0.027	-0.028	-0.027
1	0.092	0.090	0.090
2	-0.098	-0.099	-0.099
3	0.069	0.069	0.069
4	0.015	0.016	0.016
5	-0.017	-0.017	-0.018
6	-0.009	-0.006	-0.007
7	-0.055	-0.055	-0.055
8	0.030	0.030	0.031
9	0.112	0.113	0.114
10	-0.064	-0.065	-0.066
11	0.092	0.093	0.094

2.2 Similar company by revenue segments

Compare PCA and Autoencoder (AE):	**No transform**	**PCA**	AE
PCA is a linear transformation vectors while AE can be non-linear.	√	√	
PCA is faster than AE.	√	√	
PCA projects data into dimensions that are orthogonal to each other; AE transformed data doesn't guarantee that.		√	
Rule of thumb: go for AE with larger data sets			√
AE with multiple layers and non-activation function is prone to overfitting.	√	√	

2.2 Similar company by revenue segments

Use Cosine similarity to get similar company:

After we get different vectors for different company, we use cosine similarity to measure the distance between companies.

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}},$$

2.2 Similar company by revenue segments

Example:

If our customer wants to find Microsoft's similar company, our model will calculate the cosine distance between Microsoft's vector and other companies' vectors. The result will be presented as a sorted company list according to cosine distance.

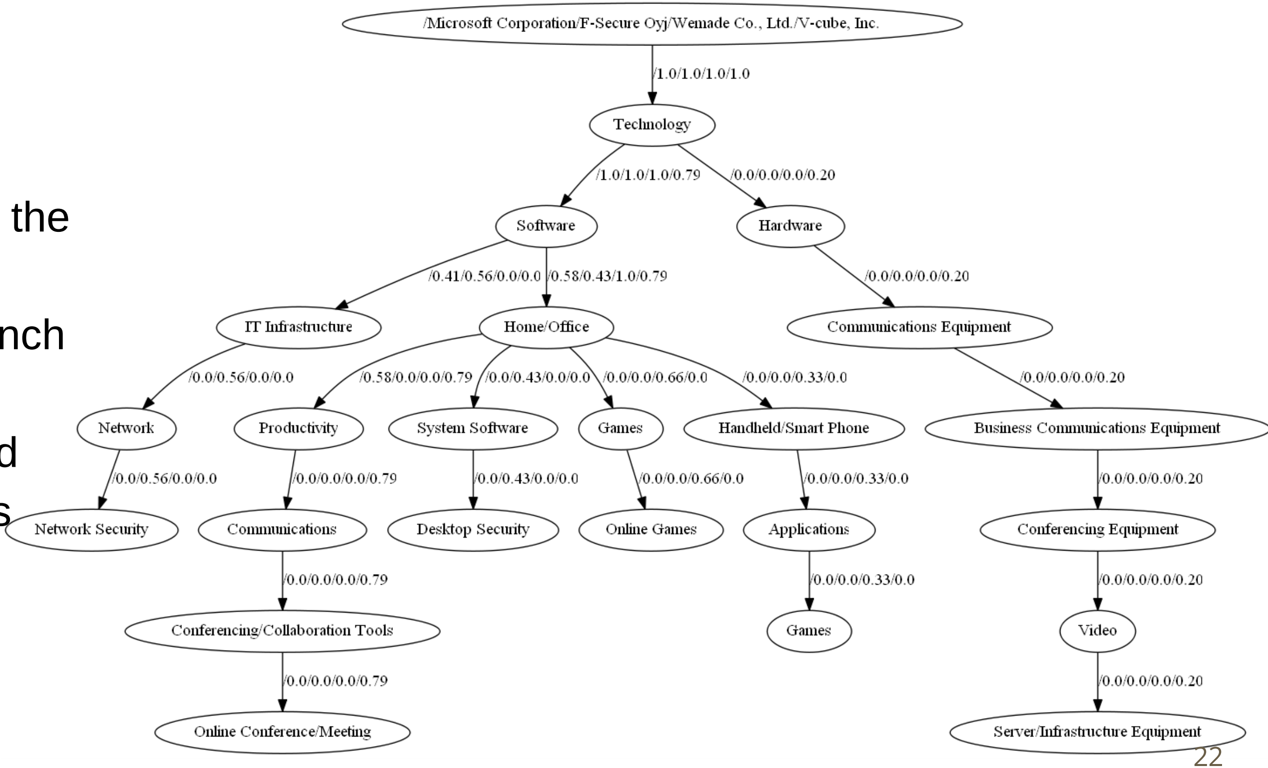
```
Source Company: Microsoft Corporation
Similar Companies:  ['F-Secure Oyj', 'Wemade Co., Ltd.', 'V-cube, Inc.', 'Pexip Holding ASA', 'DoubleUGames Co., Ltd.']
```

2.2 Similar company by revenue segments

Result presented:

2.2.1 Visual trees

- Similar company list is at the top of the tree.
- Numbers along each branch are segment percent.
- We can know the detailed area for these companies



2.2 Similar company by revenue segments

Result presented:

2.2.2 Table of calculated vectors

We present the result by using table to show the vectors that we calculated at the first step.

2.2 Similar company by revenue segments

Handwritten Model cont.

(Detailed path for JPMorgan's similar company)

depth	name	JPMorgan Chase & Co.	Boston Private Financial Holdings, Inc.	Bryn Mawr Bank Corporation	First Bancshares, Inc.	Berkshire Hills Bancorp, Inc.	Great Western Bancorp, Inc.	Origin Bancorp, Inc.
0	Finance	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1	Banking	0.36	0.38	0.39	0.50	0.50	0.50	0.50
	Investment	0.14	0.12	0.11	0.00	0.00	0.00	0.00
2	Asset Management and Financial Advisory	0.00	0.08	0.07	0.00	0.00	0.00	0.00
	Securities Sales and Trading	0.10	0.00	0.00	0.00	0.00	0.00	0.00
	United States Banks	0.24	0.26	0.26	0.33	0.33	0.33	0.33
3	Commercial Banks	0.18	0.19	0.19	0.25	0.25	0.25	0.25
	Institutional/High-Net Advisory	0.00	0.06	0.06	0.00	0.00	0.00	0.00
4	Midwest	0.00	0.00	0.00	0.00	0.00	0.20	0.00
	Northeast	0.00	0.00	0.16	0.00	0.20	0.00	0.00
	Private Wealth Management	0.00	0.00	0.04	0.00	0.00	0.00	0.00
	South	0.00	0.00	0.00	0.20	0.00	0.00	0.20
	West	0.00	0.15	0.00	0.00	0.00	0.00	0.00
5	Middle Atlantic	0.00	0.00	0.13	0.00	0.00	0.00	0.00
	New England	0.00	0.00	0.00	0.00	0.17	0.00	0.00
	Pacific	0.00	0.13	0.00	0.00	0.00	0.00	0.00
	West North Central	0.00	0.00	0.00	0.00	0.00	0.17	0.00
	West South Central	0.00	0.00	0.00	0.00	0.00	0.00	0.17
6	California	0.00	0.11	0.00	0.00	0.00	0.00	0.00
	Pennsylvania	0.00	0.00	0.11	0.00	0.00	0.00	0.00

All in Finance sector

2.2 Similar company by revenue segments

Autoencoder cont.

(Detailed path for JPMorgan's similar company)

depth	name	JPMorgan Chase & Co.	Boston Private Financial Holdings, Inc.	Bryn Mawr Bank Corporation	North Pacific Bank, Ltd.	Bank Millennium SA	Liberbank SA	illimity Bank SpA	Credito Valtellinese SCARL	TBC Bank Group Plc
0	Finance	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	Banking	0.36	0.38	0.39	0.40	0.50	0.50	0.50	0.50	0.50
1	Investment	0.14	0.12	0.11	0.00	0.00	0.00	0.00	0.00	0.00
	Specialty Finance	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00
	Asset Management and Financial Advisory	0.00	0.08	0.07	0.00	0.00	0.00	0.00	0.00	0.00
	Commercial Finance	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00
2	International Banks (excluding United States)	0.00	0.00	0.00	0.27	0.33	0.33	0.33	0.33	0.33
	Securities Sales and Trading	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	United States Banks	0.24	0.26	0.26	0.00	0.00	0.00	0.00	0.00	0.00
	Asia/Pacific	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00
	Commercial Banks	0.18	0.19	0.19	0.00	0.00	0.00	0.00	0.00	0.00
3	EMEA (Europe, Middle East, Africa)	0.00	0.00	0.00	0.00	0.25	0.25	0.25	0.25	0.25
	Institutional/High-Net Advisory	0.00	0.06	0.06	0.00	0.00	0.00	0.00	0.00	0.00
	Leasing	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	Europe	0.00	0.00	0.00	0.00	0.20	0.20	0.20	0.20	0.20
	Northeast	0.00	0.00	0.16	0.00	0.00	0.00	0.00	0.00	0.00
4	Other Asia/Pacific	0.00	0.00	0.00	0.16	0.00	0.00	0.00	0.00	0.00
	Private Wealth Management	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00
	West	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Middle Atlantic	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00
	Pacific	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	Southern Europe	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17
	Western Europe	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00
	California	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Corporate	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.07
6	Pennsylvania	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00
	Retail	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.07

All in Finance sector

Some results are international banks

Some are not commercial banks

2.2 Similar company by revenue segments

Encapsulation to python package:

1. Easy to use.
2. Support fuzzy search for company name.
3. Clients can set flexible threshold.
4. Full visualization in Jupyter Notebook.

```python
from SimilarCompanyFinder import SimilarCompanyFinder
```

3 General Recommender System

- Help the user to select the right item by suggesting list of items
- Two most popular ways:
  - Content based filtering
    - based on user preferences for product features
  - Collaborative filtering
    - Mimics user-to-user recommendations (predicts users preferences as a weighted linear combination of other users' preferences)
    - mix the features of the item itself and the preferences of other users

3 General Recommender System -item based

- looks for similar items based on the items users have already liked or positively interacted with
- i.e. Movie recommendations:
  - search for movies that the user has watched/liked (A, B, C)
  - search for movies similar to A, B, C, recommend D
- Based on the similarity between items calculated using the rating users have given to items

3 General Recommender System -item based

- Use cosine similarity between item i and j

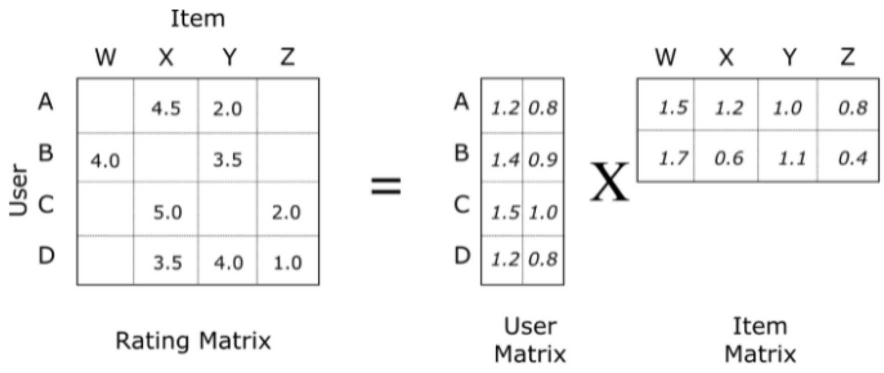
$$similarity(i,j) = \frac{\sum_u^U r_{(u,i)} r_{(u,j)}}{\sqrt{\sum_u^U r_{(u,i)}^2} \sqrt{\sum_u^U r_{(u,j)}^2}}$$

$$score(u,i) = \frac{\sum_j^I similarity(i,j)(r_{(u,j)} - \bar{r_j})}{\sum_j^I similarity(i,j)} + \bar{r_i}$$

- Other similarity measures:
    - Mean Squared Difference ; Pearson correlation coefficient; Pearson_baseline (shrunk)

3 General Recommender System -user based

- Identify neighboring users based on the similarity with the active user, then the scoring of the items is done based on neighbor’s ratings followed by a recommendation of an item based item’s scores by the recommendation system
- i.e. “Customers who bought this item also bought…”
- Use Matrix Factorization



3 General Recommender System -potential application

- For portfolio optimization:
  - Different Funds/Portfolio managers are users
  - Different companies/securities/products are items
  - Ratings could be the weights of each product in each fund which reflect different portfolio managers' preferences on different products/firms/securities
  - Apply the recommender system model to recommend alternative choices for portfolio managers based on their preferences

4 Portfolio Optimization

- Scenario

After we recommend a stock or find similar stocks for portfolio manager, we will optimize the portfolio again.

- Risk factors

We will build the similar portfolio based on barra risk factors.

- Example

Microsoft, US company, Technology/Software.

$$F\ (k,\ m) = \text{Covariance}\ [\tilde{f}\ (k),\ \tilde{f}\ (m)]$$

where $F\ (k,\ m)$ = factor covariance matrix, and
$k,\ m$ = common factors.

$$F = \left[\begin{array}{ccc:ccc:ccc}
F(1,1) & \ldots & F(1,48) & F(1,49) & \ldots & F(1,86) & F(1,87) & \ldots & F(1,90) \\
\vdots & \text{Local Markets} & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
F(48,1) & \ldots & F(48,48) & F(48,49) & \ldots & F(48,86) & F(48,87) & \ldots & F(48,90) \\
\hdashline
F(49,1) & \ldots & F(49,48) & F(49,49) & \ldots & F(49,86) & F(49,87) & \ldots & F(49,90) \\
\vdots & & \vdots & \vdots & \text{Industries} & \vdots & \vdots & & \vdots \\
F(86,1) & \ldots & F(86,48) & F(86,49) & \ldots & F(86,86) & F(86,87) & \ldots & F(86,90) \\
\hdashline
F(87,1) & \ldots & F(87,48) & F(87,49) & \ldots & F(87,86) & F(87,87) & \ldots & F(87,90) \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & \text{Risk Indices} & \vdots \\
F(90,1) & \ldots & F(90,48) & F(90,49) & \ldots & F(90,86) & F(90,87) & \ldots & F(90,90)
\end{array}\right]$$

4 Portfolio Optimization

- Problem

$$
\begin{array}{ll}
\text{maximize} & \mu^T w - \gamma\left(f^T \tilde{\Sigma} f + w^T D w\right) \\
\text{subject to} & \mathbf{1}^T w = 1, \quad f = F^T w \\
& w \in \mathcal{W}, \quad f \in \mathcal{F},
\end{array}
$$

Variables are the allocations $w \in R_n$ and factor exposures, $f \in R_k$ and F gives the factor exposure constraints.

- Solve with **cvxpy, OSQP** solver

```
# Factor model portfolio optimization.
ret = mu.T@w
risk = cp.quad_form(f, Sigma_tilde) +
        cp.sum_squares(np.sqrt(D) @ w)
prob_factor = cp.Problem(cp.Maximize(ret - gamma*risk),
                         [cp.sum(w) == 1,
                          f == F.T@w,
                          cp.norm(w, 1) <= Lmax])
```

Not enough data provided, we will run the code later.

Conclusion

- We could find similar companies based on their Barra risk factors or revenue and sector data
- Deep learning techniques such as autoencoders could be used for larger datasets.
- User based recommender system and the above similar company finder could be used by portfolio managers
- Packages such as cvxpy and OSQP could be utilised to solve optimization problems