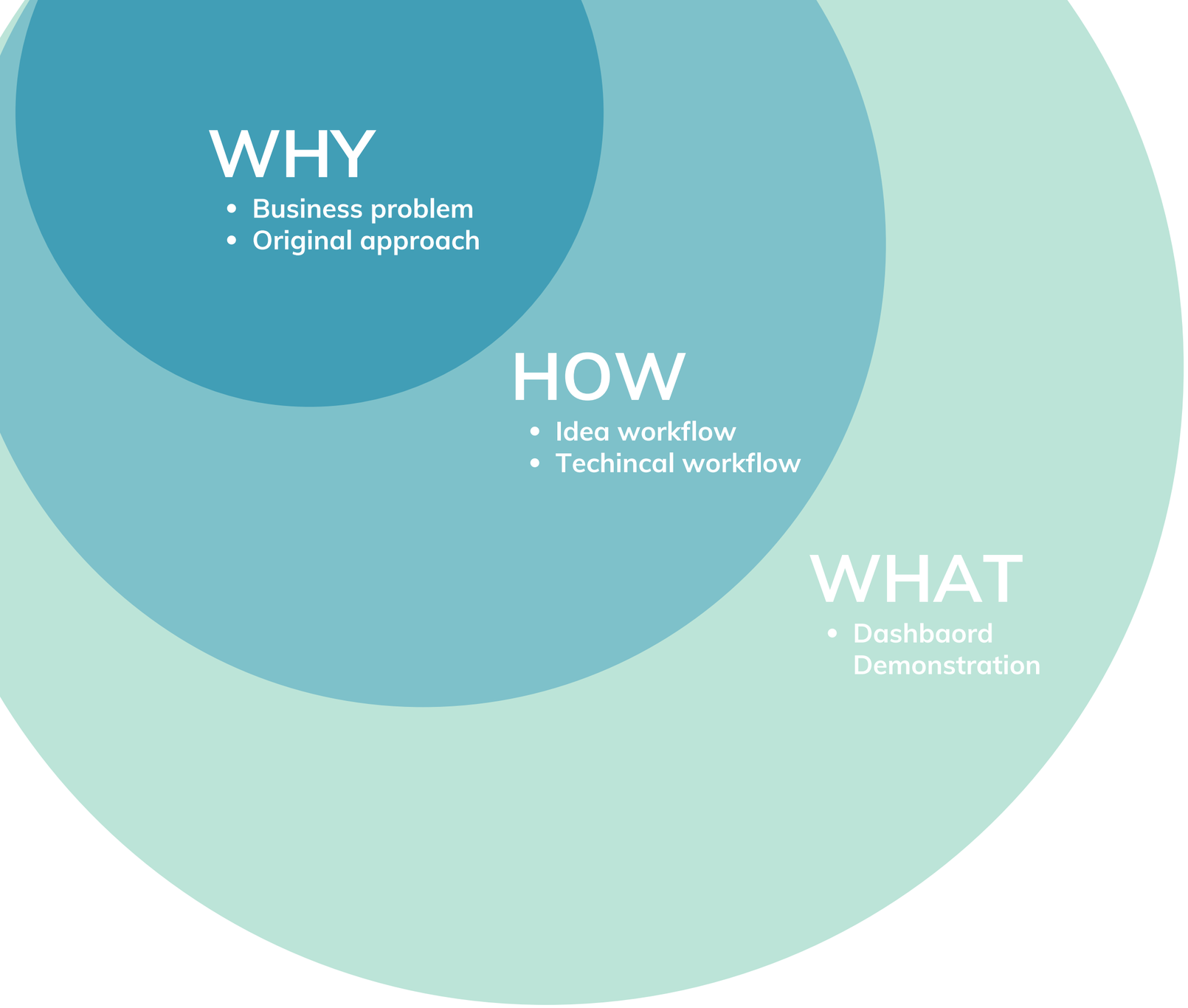
Matt Cheng

Data Project Presentation

Data Analytic with AI -
Comparable analysis in Corporate Finance



In this Presentation

Here’s what we’ll cover:

WHY?

- Personal Background
- Business problem
- Original approach

Personal Background

Investment banking analyst

- A major part of my job is **creating pitch books** for potential IPO candidates
- Involves **identifying comparable** listed companies to project potential IPO earnings.

Business Problems

Company Name	Stock Code	Core Business
EXTRA COMPARABLES		
CHOW SANG SANG	00116.HK	The manufacture and retail of **jewellery**, distribution of **watches**...
LUK FOOK GOLD	00590.HK	Retailing of a variety of **gold** and platinum **jewellery** and gem-set jewellery products...
CHOW TAI FOOK	01929.HK	Manufacturing and selling of **jewellery** products...
PRADA	01913.HK	Producing and distributing **leather** goods, **footwear** and **apparel**...
ORIGINAL COMPARABLES		
TOKYO CHUO	01939.HK	**Auction** and **artwork sales** in Hong Kong and Japan...

Example of real comparable search for art trading platform

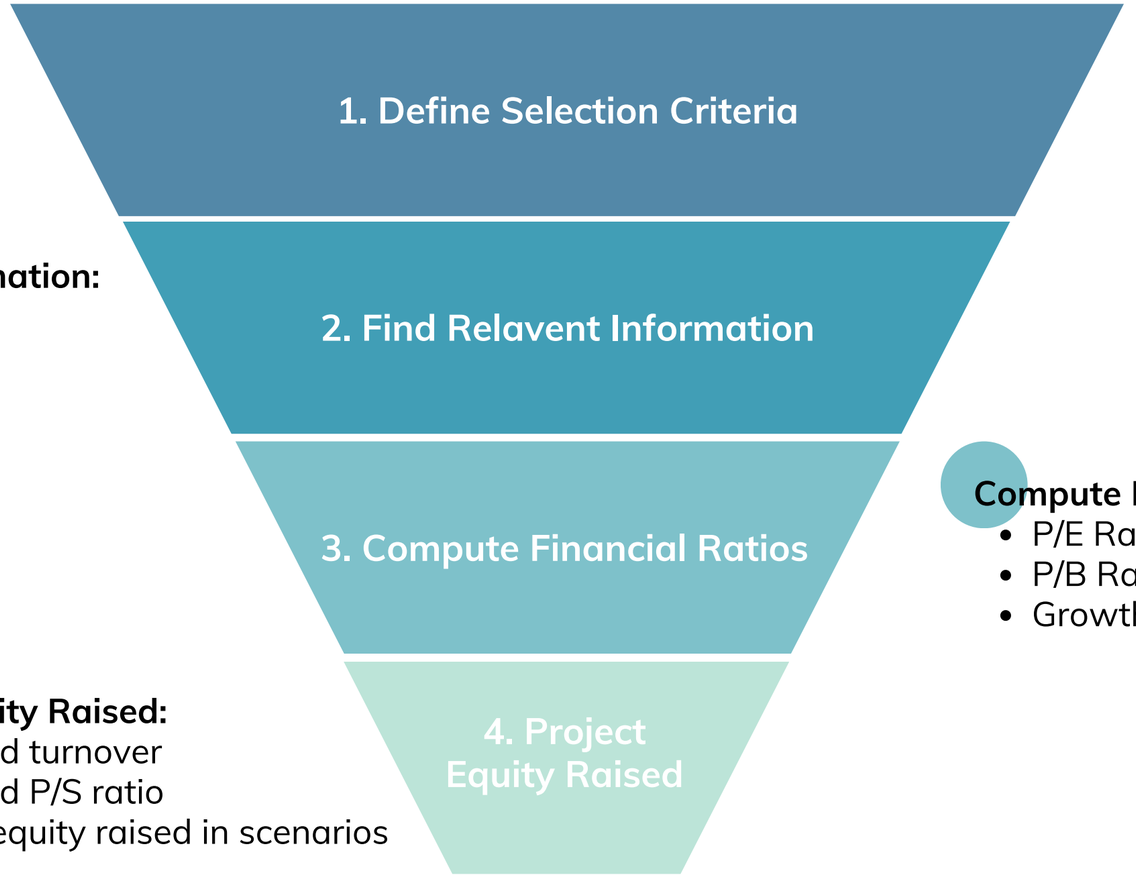
- Process of finding comparables is imprecise due to **varying judgment criteria** among parties.
- For instance, an **art trading platform** seeking comparables in the HK market faced a challenge as there were no direct competitors.
- We have to use unrelated companies like “Chow Sang Sang” (00116.hk) and “Prada” (01913.hk) as comparables.

So there is an question in my mind:

“

Can I find a legit comparable company for investment analysis in 30 seconds?

Original Approach



Define Selection Criteria:
- Product details
- Geographical locations
- Accounting policies

Extract Relavent Information:
- Prospectus
- Annual reports
- Financial websites

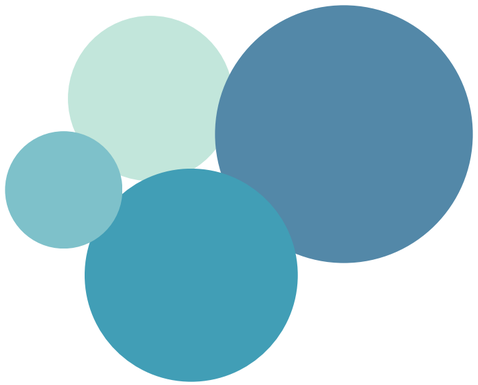
Compute Financial Ratios:
- P/E Ratio
- P/B Ratio
- Growth Margin

Project Equity Raised:
- Predicted turnover
- Predicted P/S ratio
- Project equity raised in scenarios

HOW?

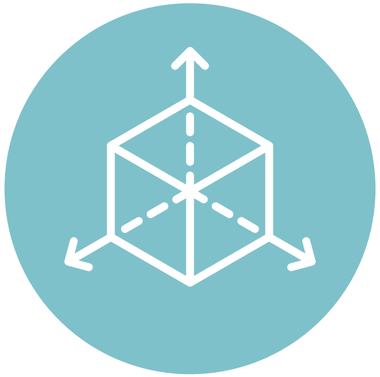
- Idea Workflow
- Data processing
- Vector Similarity Search Explanation
- Technical Workflow

Idea Workflow

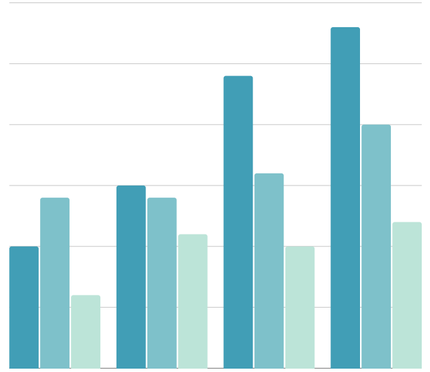


Web scrape
pre-processed Data

Extract product features
by LLM



Perform vector similarity
search from text embeddings



Dashboard visualisations

Data Processing

Investment related Data:
- Market Capitalisation
- Current stock price
- P/E, P/B ratios

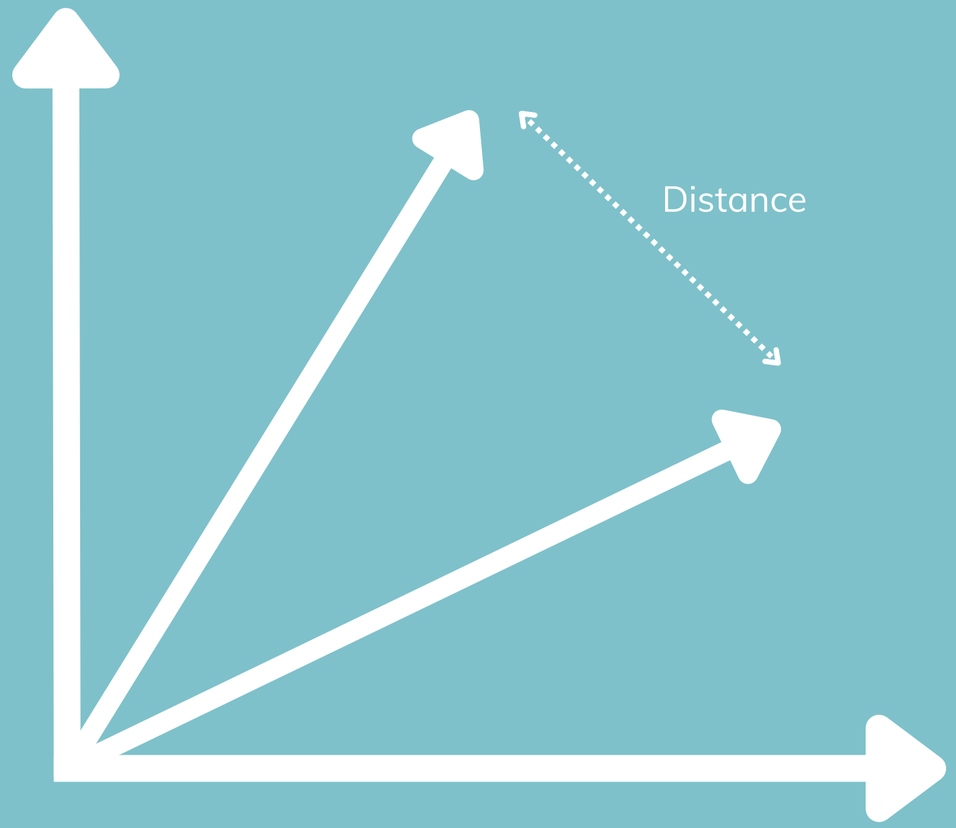
Document-type data:
- Annual Reports
- Prospectus

Public Data (Not included):
- Website information

Financial Data:
- Turnover, net profit
- Financial Ratio (e.g. Profit Margin, Asset Ratio)

Put all different sources of data into
Large Language Model (LLM)
to process and generate output

Vector **Similarity**



The only way that we can compare words are **turning them into numbers (Vectors)**

Vectors similarity search compares the words' corresponding numbers (embeddings) to find out which terms are more similar

e.g. Apple vs Orange / Apple vs Airplane

Technical Workflow

Annual Reports &
Stock price

Financial Data

Web Scrape
through
Selenium

Extract product descriptions
through
Open AI API

Use RDS Service
to store relational data

Use Chroma DB
to store and perform
vector similarity search

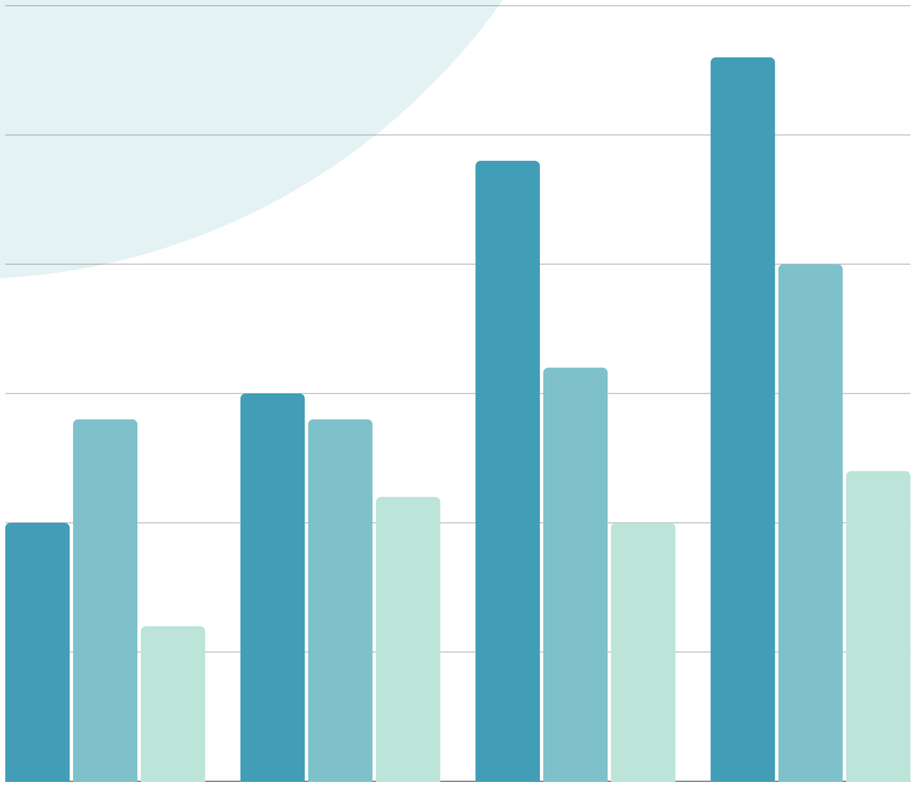
Use PostgresSQL
to query relational data

Use Tableau API (Tabpy)
to connect ChromaDB
and Python

Use Tableau
to visualise
insights

WHAT?

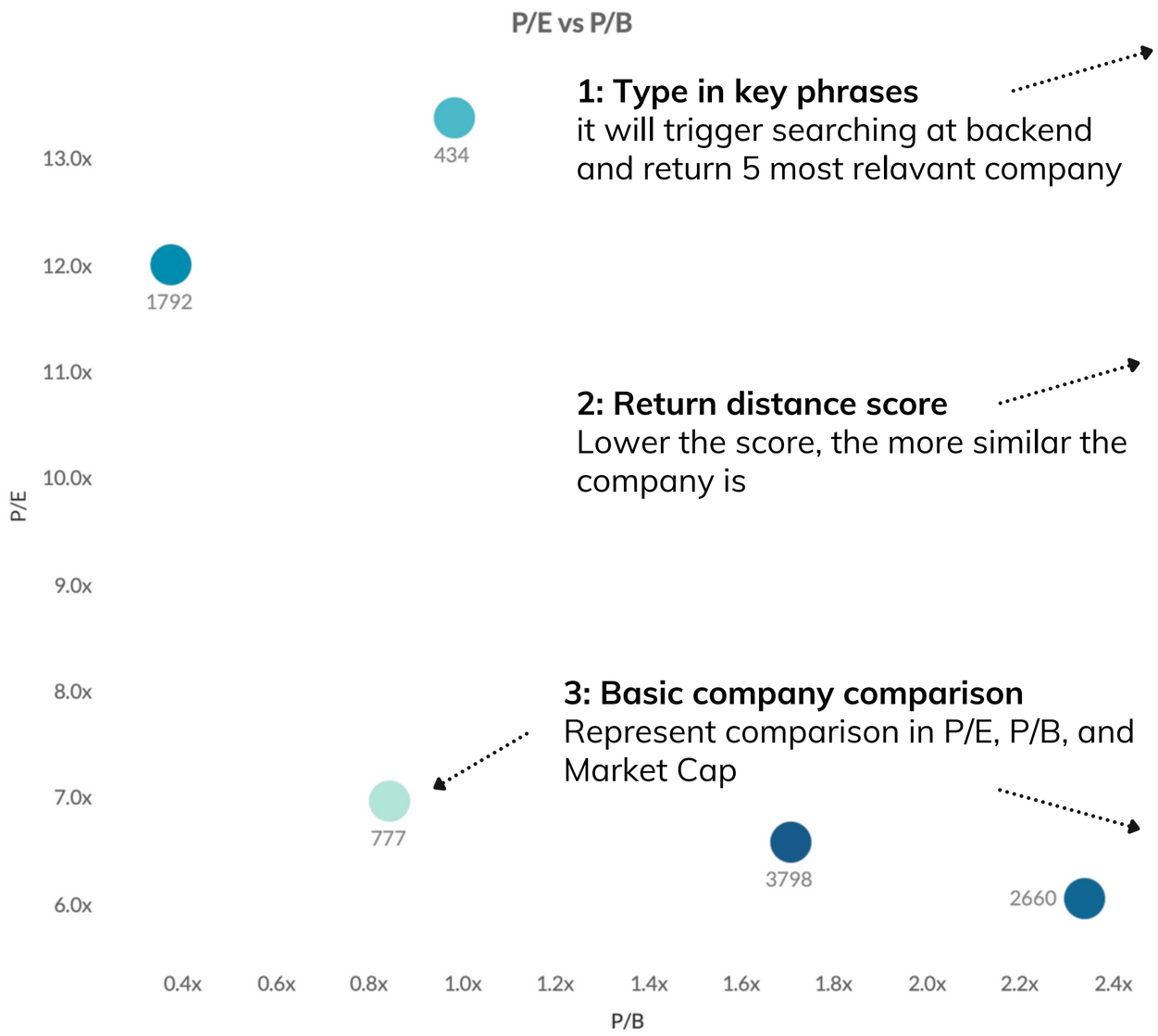
- Dashboard Demonstration
- Future Advancement
- Before & After



Dashboard Demonstration

PEER EVALUATION FINANCIAL ANALYSIS BY AI

Using LLM to extract features from more than 2000+ annual reports, and compare performance to provide insights on potential IPO investments



1: Type in key phrases
it will trigger searching at backend and return 5 most relavant company

2: Return distance score
Lower the score, the more similar the company is

3: Basic company comparison
Represent comparison in P/E, P/B, and Market Cap

Company Descriptions

Produce mobile games mainly for gamers in china

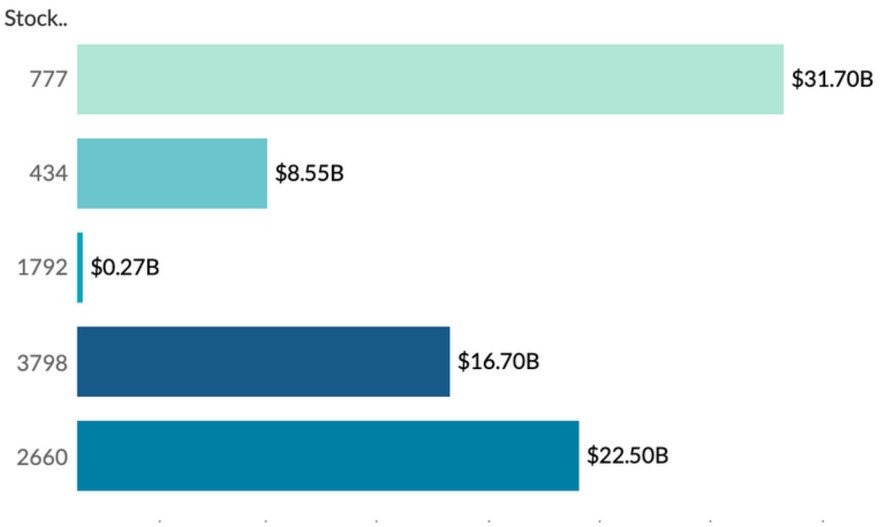
Year: 2023

Distance Score: 0.7973 – 1.0015

Distance Score

Stock Code	Stock Name	Distance Score
1792	CMON Limited	0.8661
434	Boyaa Interactive International Limited	0.9285
777	NetDragon Websoft Holdings Limited	1.0015
2660	Zengame Technology Holding Limited	0.8147
3798	Homeland Interactive Technology Ltd.	0.7973

Market Capitalisation

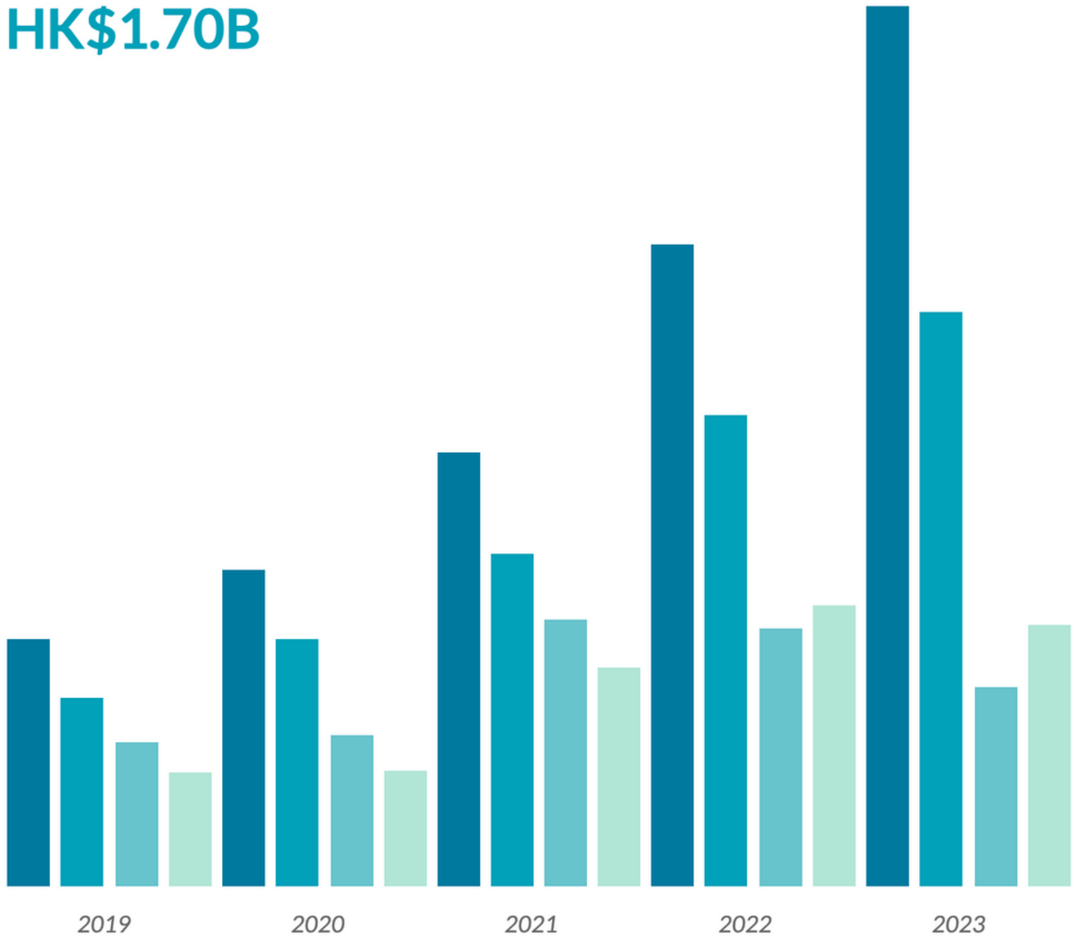


HOMELAND INTERACTIVE TECHNOLOGY LTD.

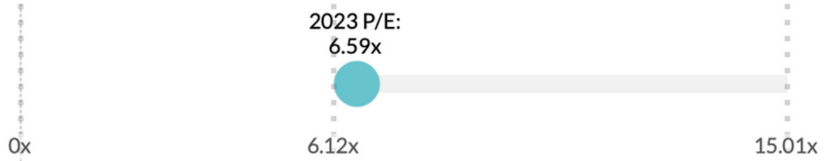
Input Company Descriptions: Produce mobile games mainly for gamers in china
Similarity Rank: 1

Sales & Profit Trend

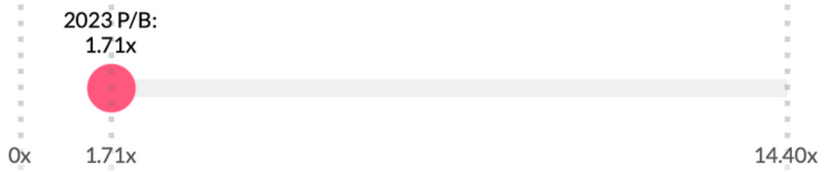
2023 Turnover:
HK$1.70B



P/E Range (2019 - 2023)



P/B Range (2019 - 2023)



Similarity Ranking
1

Year
2023

Profit Axis
- Turnover
- Gross Profit
- EBITDA
- Net Profit

Profit Ratios

Gross Profit Margin

EBITDA Margin

Net Profit Margin

Individual company performance
Select year or ranking, the dashboard will automatically refresh and show related information

Projection on equity raised
The ultimate purpose for the analysis is to know how much can they raise through this brokage firm

HOW MUCH CAN YOU RAISE?

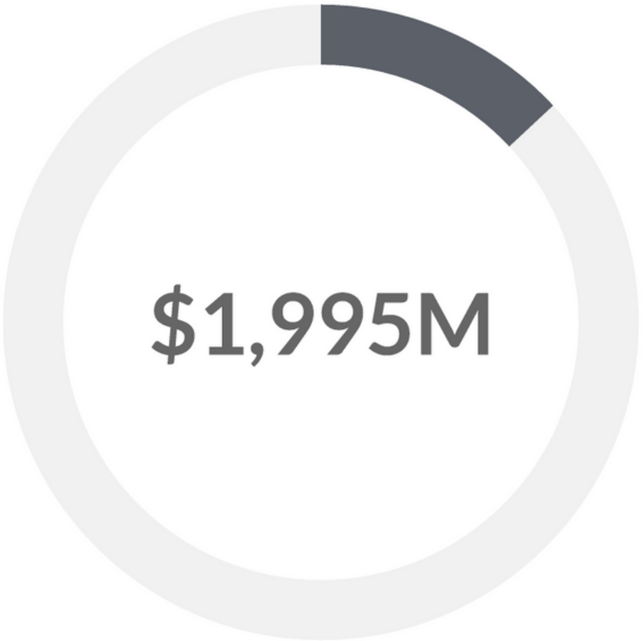
Predicted P/S

10.0x

Predicted Turnover

10.0M

Without Green Shoe Option



With Green Shoe Option

Green shoe option in HK IPO lets the banks selling the company's shares at a degree that up to **15% more** shares than their originally planned.
When lots of people want to buy the shares, this option is to **keep the share price stable** right after the company starts trading on the stock market.

Before project:

- Wasting time in reading annual reports and prospectus
- Lack standards for comparison
- Risky decisions from personal judgments and prejudices.

After project:

- Text inputs to find related comparables
- Multi-dimentional comparsion
- Informed investment decisions from data visualisations

Future Advancement

What to look out for and refine

Public Data

Website may include more organisational and marketing data

Accounting Policies

Accounting policies are one of the criteria in some industry-specifc comparable analysis

Vertical Analysis

Suppliers and customers data can be contributive to the comparable analysis

Thank You!