

## Workshop: Finding Objects, then classes, then model

### Problem

We need a **system** that compares analysts from **records** of the analyst's own **stock buy/sell transactions**, so we can **recommend investors** to invest with the best **analyst**.

To understand what the **program** needs to do, it is first **necessary** to understand a little bit about the **stocks, buy** and **sell transactions**, and their **records**. We'll keep things **simple**, though, and only deal with **basic stock purchases** and **sales**.

### Overview

In the **domain** of **financial investment**, **individual investors** may base **buy/sell decisions** on the **opinions** of **investment analysts**, who spend considerable time studying the **fundamentals** and **potential** of selected **companies**. When an investor relies on an analyst's **recommendations** to any **degree**, then that **investor** typically wants know something about the **analyst's track record**.

### Background

Think of a **share** of **stock** as a piece of a **company**, albeit typically a very small **piece**. The value of a **share** is loosely related to the **value** of the **company divide** by the **number** of **outstanding** shares (i.e., total **shares** owned by **shareholders**.) For example, if a **company** was worth \$10M and there was a total of 1M shares owned by **shareholders**, then each share would be worth \$10.

In a stock **market**, all **companies** and their stock are **identified** by **symbols**, which are short character **strings**. For example, Amazon's **symbol** is AMZN, Apple's is AAPL, and Microsoft's is MSFT.

An **investor** will **buy** some **number** of shares at a **specific purchase price** and pay a small fee to a **broker** to **complete** that **transaction**. Then, at late **time**, the **investor** will sell all or some of those **shares** at a sales **price** and **pay** another small fee to a **broker**. Obviously, if an **investor sells** at a **price** higher than the **purchase price** (adjusting for the fees), the **investor** will make **money**. For example, if an **investor buy** 10 **shares** of AMZN at \$800/share for a \$10 fee, then **sell** those 10 shares at \$850/share for \$10 fee, that **investor** will make a profit of \$480

$(850 \times 10 - 10 - 800 \times 10 - 10)$ .

For **taxes purpose** and asset **management**, **investors** correlate the sale of a **block** of stocks to **specific purchases**. For example, consider an **investor** who buy 100 shares of ABC at \$10/share and a little while later another 100 shares at \$11/share. Then, **image** that the **investor sell** 125 shares. In **making** that **sale**, the **investor** will say how many of the 125 shares come from the first **block** of shares **purchase** and how many come from the **second block**, because it will **impact** how many **profit** the **investor** makes.

An **investor** does not **actually** make or loss **money** until shares are sold. Shares that an investor is **holding** onto only **represent potential** profits or losses. So, to rank an **analyst** perform, we need to know about completed **purchase/sales** cycles. Therefore, a **history** will contain **records** with following data:

Stock **symbol** A short **string** that **identifies** the **company**  
Quantity (Q) A integer that represent the **number** of shares  
purchase and sold Purchase Date/time (PT) A number that represents time of purchase in **minutes** since Jan 1, 2016  
Purchase Price (PP) An integer **represents** the purchase price in **cents**. For example, 1234 would be a **purchase price** of \$12.34.  
**Purchase** Trans. Fee (PF) An integer represents the purchase **transaction** fee in **cents**. Sale Date/time (ST) A **number** that represents time of purchase in **minutes** since Jan 1, 2016  
Sale Price (SP) An **integer represents** the **purchase price** in **cents**. For example, 1234 would be a purchase **price** of \$12.34. Sale Trans. Fee (SF) An integer **represents** the **purchase transaction** fee in cents. Q, PT, PP, PF, ST, SP, SF  
See **Table 1** for an example of a small **history** for an **investor** who purchased and sold stock for four **companies**: Amazon (AMZN), Apple (AAPL), Microsoft (MSFT), and Google (GOOGL).

**Table 1 – Sample Purchase-Sale History**

<b>Symbol</b>	<b>Quantity (Q)</b>	<b>Purchase Date/Time (PT)</b>	<b>Purchase Price (PP)</b>	<b>Purchase Trans. Fee (PF)</b>	<b>Sale Date/Time (ST)</b>	<b>Sale Price (SP)</b>	<b>Sale Trans. Fee (SF)</b>
AMZN	10	5256000	\$510.00	\$9.95	5306400	\$630.00	\$9.95
AAPL	20	5258880	\$105.00	\$9.95	5303520	\$98.35	\$9.95
MSFT	50	5261760	\$48.20	\$9.95	5282160	\$49.50	\$9.95
GOOGL	10	5263200	\$712.43	\$9.95	5298240	\$740.23	\$9.95
AAPL	20	5263200	\$96.00	\$5.00	5442240	\$110.03	\$9.95

The **amount** of **money** invested (INV) in a purchase/sale is the cost of the shared adjust plus both the **purchase** fee and the

sales fees  $INV = Q*PP + PF + SF$

So, the **profit/loss** (PL) for each purchase/sale can be computed as  $PL = Q*SP - INV$

**Table 2** shows INV and PL for each of the **row**, **identified** by the **symbol** and ST, of **Table 1**.

**Table 2 – Computed Values for Purchase/Sales Records for Table 1**

<b>Symbol</b>	<b>Sale Date/Time (ST)</b>	<b>Investment INV = Q*PP + PF + SF</b>	<b>Profit/Loss PL = Q*SP - INV</b>
AMZN	5306400	\$5,119.90	\$1,180.10
AAPL	5303520	\$2,119.90	\$(152.90)
MSFT	5282160	\$2,429.90	\$45.10
GOOGL	5298240	\$7,144.20	\$258.10
AAPL	5442240	\$1,934.95	\$265.65
AAPL	5503800	\$3,774.95	\$433.85
AAPL	5347200	\$5,539.90	\$599.30
MSFT	5336760	\$4,859.90	\$50.10
GOOGL	5279160	\$7,340.10	\$(437.80)
GOOGL	5289240	\$14,219.90	\$207.10
MSFT	5355840	\$2,524.90	\$(64.90)
MSFT	5295240	\$2,529.90	\$(9.90)

## Analyst Comparison

To **compare analysts**, we give each analyst “play” **money** (called the seed amount) and let them make whatever **purchases** and **sales** they desire, using **simulated transaction**. We record the **purchase-sales** for each **investor** in a separate purchase-sale **histories file**. Each investor starts off with same seed **money**, but they can start and end their simulation at **different times**. The **number of days** between the start and end of an analyst’s **simulation** is called the **Simulation Days** (D).

One measure of overall **investor** performance is **total** of all PL’s computed for each **purchasesale** in that **investor’s history**. We’ll call this **Total Profit-Loss** (TPL). Another measure of overall investor performance is TPL divided by the D. We’ll call this Profit/Loss per **Day** (PLPD).

Another **measure** is an **investor’s** performance for an **individual** stock. This is the **sum** of PL for that stock, divided by the **total days invested** in that **stock**. See **Table 3** for an example.

**Table 3 – Example Performance Measures for an Analyst**

**TPL**                \$1,596.05  
**D**                    200    days  
**PLPD**            7.98    \$/day

<b>Symbol</b>	<b>PL</b>	<b>Min PT</b>	<b>Max ST</b>	<b>Stock Investment Days (SID) (Max ST – Min PT) / (24*60)</b>	<b>Stock's Profit/Loss Per Day (SPLPD) PL / SID</b>
AMZN	\$1,395.25	5256000	5507280	174.50	\$8.00
AAPL	\$1,145.90	5258880	5503800	170.08	\$6.74
MSFT	\$20.40	5261760	5355840	65.33	\$0.31
GOOGL	\$301.10	5263200	5335380	50.13	\$6.01