

# Programming Assignments

- General goal:
  - How to use a regular programming language (e.g. Java) and SQL together to *take a decision about a real-life situation*
- Specific goal:
  - *Data Mining* on available stocks of different companies

# Assignment 1

- Analyze **256K+ data points**
- Find when stock split occurs (stock split)
- Find when stock price is behaving in a strange way (crazy day)
- Assignment 1 helps:
  - To understand the stock data very well
  - To understand how to process stock data in a regular text file using Java
- You will require “stock split” part of Assignment 1 to complete Assignment 2

# Assignment 2

- Analyze ~**22.57 Million** data points
- Assignment 2 helps
  - To learn how to use Java and SQL together
    - How to connect to remote database server
    - How to write query to get data from the remote database server
  - To design a stock investment strategy
    - When to buy a stock
    - When to sell a stock
    - When to do nothing
- You will require “stock split” and “split adjustment” parts of Assignment 2

# **Guidelines for Assignment 2**

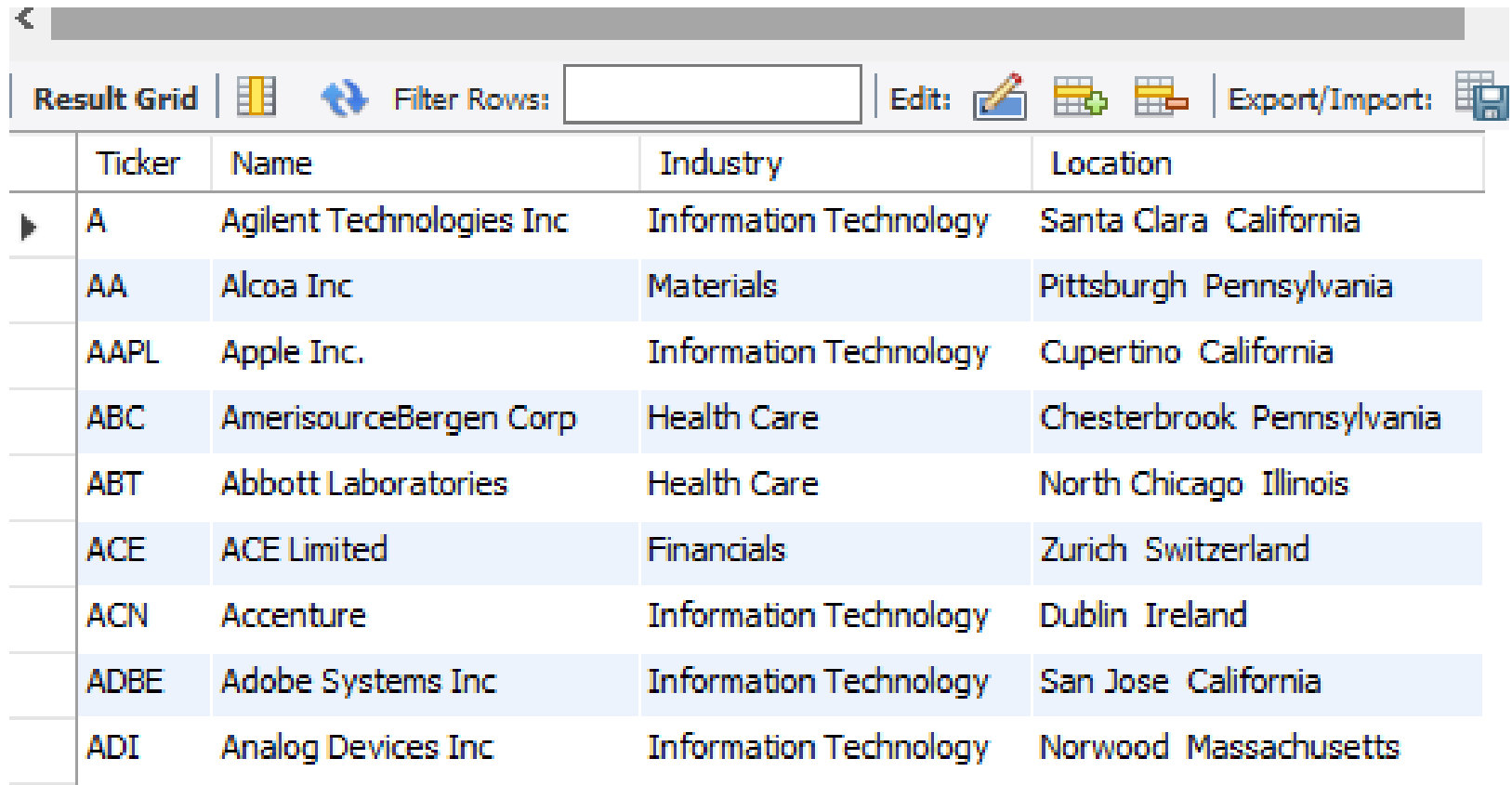
# Connect to Remote MySQL Sever

- If you are working remotely (e.g., using your laptop from home), you must connect to remote MySQL server first
- Please look at the separate file available on canvas for the steps

# Database Schema

Entity	Attributes	Primary Key	Foreign keys
company	Ticker Name Industry Location	Ticker	
pricevolume	Ticker TransDate OpenPrice HighPrice LowPrice ClosePrice Volume AdjustedClose	Ticker TransDate	Ticker
dividend	Ticker DivDate Amount	Ticker DivDate	Ticker






select \* from company limit 10;



The image shows a screenshot of a database query result grid. The interface includes a toolbar with icons for 'Result Grid', 'Filter Rows' (with a search box), 'Edit', and 'Export/Import'. The data is presented in a table with four columns: Ticker, Name, Industry, and Location. The first row is highlighted with a blue background.

	Ticker	Name	Industry	Location
▶	A	Agilent Technologies Inc	Information Technology	Santa Clara California
	AA	Alcoa Inc	Materials	Pittsburgh Pennsylvania
	AAPL	Apple Inc.	Information Technology	Cupertino California
	ABC	AmerisourceBergen Corp	Health Care	Chesterbrook Pennsylvania
	ABT	Abbott Laboratories	Health Care	North Chicago Illinois
	ACE	ACE Limited	Financials	Zurich Switzerland
	ACN	Accenture	Information Technology	Dublin Ireland
	ADBE	Adobe Systems Inc	Information Technology	San Jose California
	ADI	Analog Devices Inc	Information Technology	Norwood Massachusetts

select \* from pricevolume  
where Ticker = 'INTC' limit 10;

<								
Result Grid								
Filter Rows: <input type="text"/>								
Edit:   								
Export/Import:  								
Wrap C								
	Ticker	TransDate	OpenPrice	HighPrice	LowPrice	ClosePrice	Volume	AdjustedClose
▶	INTC	1985.01.02	28.00	28.25	27.25	27.50	27257600	0.42
	INTC	1985.01.03	27.50	28.50	27.50	28.00	31075200	0.43
	INTC	1985.01.04	28.00	28.75	28.00	28.50	11686400	0.43
	INTC	1985.01.07	28.50	29.25	28.25	29.25	12464000	0.45
	INTC	1985.01.08	29.25	29.75	27.75	28.25	33734400	0.43
	INTC	1985.01.09	28.25	29.25	27.75	28.75	16880000	0.44
	INTC	1985.01.10	28.75	30.25	28.75	30.00	31644800	0.46
	INTC	1985.01.11	30.00	30.50	29.75	30.50	33379200	0.46
	INTC	1985.01.14	30.50	31.25	29.75	31.00	42153600	0.47



# Sample Outputs

Database connection is established.

Enter a ticker symbol [start/end dates]: INTC

Intel Corp.

2:1 split on 2000.07.28 129.12 --> 65.44

2:1 split on 1999.04.09 130.81 --> 61.62

2:1 split on 1997.07.11 153.81 --> 77.25

2:1 split on 1995.06.16 116.12 --> 58.50

2:1 split on 1993.06.04 112.75 --> 60.12

3:2 split on 1987.10.28 31.75 --> 21.75

6 splits in 7470 trading days

Transaction executed

= How many buy or sales you have done in this time period

Executing investment strategy

Transactions executed: 690

Net cash: 14717.72

Enter ticker symbol [start/end dates]: INTC 1980.01.01 1999.12.31

Intel Corp.

2:1 split on 1999.04.09 130.81 --> 61.62

2:1 split on 1997.07.11 153.81 --> 77.25

2:1 split on 1995.06.16 116.12 --> 58.50

2:1 split on 1993.06.04 112.75 --> 60.12

3:2 split on 1987.10.28 31.75 --> 21.75

5 splits in 3791 trading days

Enter ticker symbol [start/end dates]: T 2000.01.01 2014.08.18

AT&T Inc

0 splits in 3679 trading days

Executing investment strategy

Transactions executed: 148

Net cash: -1568.00

Executing investment strategy

Transactions executed: 358

Net cash: 44953.95

# How your program should proceed

Database connection is established.

Enter a ticker symbol [start/end dates]: INTC

Intel Corp.

2:1 split on 2000.07.28 129.12 --> 65.44

2:1 split on 1999.04.09 130.81 --> 61.62

2:1 split on 1997.07.11 153.81 --> 77.25

2:1 split on 1995.06.16 116.12 --> 58.50

2:1 split on 1993.06.04 112.75 --> 60.12

3:2 split on 1987.10.28 31.75 --> 21.75

6 splits in 7470 trading days

Step 1

Step 2.1 to 2.4

Executing investment strategy

Transactions executed: 690

Net cash: 14717.72

Step 2.5 to 2.10

# Two Queries

Entity	Attributes	Primary Key	Foreign keys
company	Ticker Name Industry Location	Ticker	
pricevolume	Ticker TransDate OpenPrice HighPrice LowPrice ClosePrice Volume AdjustedClose	Ticker TransDate	Ticker
<del>Dividend</del>	<del>Ticker</del> <del>Amount</del>	<del>Ticker</del>	<del>Ticker</del>

```
select Name  
from company  
where Ticker = 'INTC';
```

```
select TransDate, OpenPrice,  
HighPrice, LowPrice, ClosePrice  
from priceVolume  
where Ticker = 'INTC'  
order by TransDate DESC;
```

# Split Adjustment for INTC (Step 2.5): A Simple Way

Result Grid | Filter Rows: | Export:

TransDate	OpenPrice	HighPrice	LowPrice	ClosePrice
1999.04.19	57.12	59.00	55.00	55.50
1999.04.16	57.88	58.38	56.45	57.25
1999.04.15	58.50	58.62	56.50	58.44
1999.04.14	61.38	61.50	56.25	57.00
1999.04.13	62.94	63.00	59.81	60.50
1999.04.12	61.62	62.47	60.00	61.25
1999.04.09	132.38	132.69	130.06	130.81
1999.04.08	132.00	133.38	128.44	131.06
1999.04.07	131.44	133.50	128.50	132.12
1999.04.06	126.75	131.19	126.00	130.44
1999.04.05	121.88	127.50	121.75	127.50
1999.04.01	119.94	121.38	118.88	120.88

.....

TransDate	OpenPrice	HighPrice	LowPrice	ClosePrice
1997.07.21	86.19	86.50	84.56	85.69
1997.07.18	86.75	87.62	84.75	86.44
1997.07.17	88.25	89.69	86.12	87.81
1997.07.16	86.38	88.88	84.75	88.38
1997.07.15	80.75	81.94	79.69	80.91
1997.07.14	77.25	78.98	76.75	78.75
1997.07.11	150.50	154.00	149.88	153.81
1997.07.10	152.50	153.50	149.75	150.12
1997.07.09	151.00	154.56	151.00	152.62
1997.07.08	147.81	150.38	146.25	149.62
1997.07.07	145.69	149.00	145.38	147.38
1997.07.03	144.62	145.31	143.00	144.94

.....

Divide by 2 for split adjustment

Divide by 2 again  
for split adjustment

# Split Adjustment for INTC (Step 2.5): A Smart Way

Result Grid | Filter Rows: | Export:

TransDate	OpenPrice	HighPrice	LowPrice	ClosePrice
1999.04.19	57.12	59.00	55.00	55.50
1999.04.16	57.88	58.38	56.45	57.25
1999.04.15	58.50	58.62	56.50	58.44
1999.04.14	61.38	61.50	56.25	57.00
1999.04.13	62.94	63.00	59.81	60.50
1999.04.12	61.62	62.47	60.00	61.25
1999.04.09	132.38	132.69	130.06	130.81
1999.04.08	132.00	133.38	128.44	131.06
1999.04.07	131.44	133.50	128.50	132.12
1999.04.06	126.75	131.19	126.00	130.44
1999.04.05	121.88	127.50	121.75	127.50
1999.04.01	119.94	121.38	118.88	120.88

.....

.....

TransDate	OpenPrice	HighPrice	LowPrice	ClosePrice
1997.07.21	86.19	86.50	84.56	85.69
1997.07.18	86.75	87.62	84.75	86.44
1997.07.17	88.25	89.69	86.12	87.81
1997.07.16	86.38	88.88	84.75	88.38
1997.07.15	80.75	81.94	79.69	80.91
1997.07.14	77.25	78.98	76.75	78.75
1997.07.11	150.50	154.00	149.88	153.81
1997.07.10	152.50	153.50	149.75	150.12
1997.07.09	151.00	154.56	151.00	152.62
1997.07.08	147.81	150.38	146.25	149.62
1997.07.07	145.69	149.00	145.38	147.38
1997.07.03	144.62	145.31	143.00	144.94




.....

.....

Divide by 2 for split adjustment

Divide by 4 for split adjustment

# Investment Strategy for INTC (Now Ascending order): Step 2.6 to Forward

Result Grid		 Filter Rows:			Export: 
	TransDate	OpenPrice	HighPrice	LowPrice	ClosePrice
▶	1985.01.02	28.00	28.25	27.25	27.50
	1985.01.03	27.50	28.50	27.50	28.00
	1985.01.04	28.00	28.75	28.00	28.50
	1985.01.07	28.50	29.25	28.25	29.25
	1985.01.08	29.25	29.75	27.75	28.25
	1985.01.09	28.25	29.25	27.75	28.75
	1985.01.10	28.75	30.25	28.75	30.00
	1985.01.11	30.00	30.50	29.75	30.50
	1985.01.14	30.50	31.25	29.75	31.00
	1985.01.15	31.00	31.75	31.00	31.25
	1985.01.16	31.25	32.00	30.75	31.00
	1985.01.17	31.00	31.25	29.75	30.50
	1985.01.18	30.50	31.25	30.25	31.00
	1985.01.21	31.00	31.75	30.75	31.75
	1985.01.22	31.75	32.50	31.75	32.00
	1985.01.23	32.00	32.00	31.25	31.50
	1985.01.24	31.50	32.50	31.00	31.75
	1985.01.25	31.75	31.87	31.25	31.75
	1985.01.28	31.75	31.75	30.75	30.75
	1985.01.29	30.75	31.25	30.25	31.25
	1985.01.30	31.25	32.25	31.13	31.25
	1985.01.31	31.25	31.50	30.50	30.75




# The Deque Interface in Java

- Pronounced as deck
- A double-ended-queue.
- Implements both stacks and queues at the same time.

Deque Methods		
Type of Operation	First Element (Beginning of the Deque instance)	Last Element (End of the Deque instance)
Insert	addFirst(e)	addLast(e)
Remove	removeFirst()	removeLast()
Examine	getFirst()	getLast()

# Investment Strategy for INTC

## (Now Ascending order): Step 2.6 to Forward

Result Grid		 Filter Rows:		Export:	
	TransDate	OpenPrice	HighPrice	LowPrice	ClosePrice
▶	1985.01.02	28.00	28.25	27.25	27.50
	1985.01.03	27.50	28.50	27.50	28.00
	1985.01.04	28.00	28.75	28.00	28.50
	1985.01.07	28.50	29.25	28.25	29.25
	1985.01.08	29.25	29.75	27.75	28.25
	1985.01.09	28.25	29.25	27.75	28.75
	1985.01.10	28.75	30.25	28.75	30.00
	1985.01.11	30.00	30.50	29.75	30.50
	1985.01.14	30.50	31.25	29.75	31.00
	1985.01.15	31.00	31.75	31.00	31.25
	1985.01.16	31.25	32.00	30.75	31.00
	1985.01.17	31.00	31.25	29.75	30.50
	1985.01.18	30.50	31.25	30.25	31.00
	1985.01.21	31.00	31.75	30.75	31.75
	1985.01.22	31.75	32.50	31.75	32.00
	1985.01.23	32.00	32.00	31.25	31.50
	1985.01.24	31.50	32.50	31.00	31.75
	1985.01.25	31.75	31.87	31.25	31.75
	1985.01.28	31.75	31.75	30.75	30.75
	1985.01.29	30.75	31.25	30.25	31.25
	1985.01.30	31.25	32.25	31.13	31.25
	1985.01.31	31.25	31.50	30.50	30.75

d = Trading day,

d+1 = the next trading day,

d-1 = the prior trading day

close(d) = closing price for day d

open(d) = opening price for day d

### Step 2.7

- Maintain a moving average of closing prices over a 50-day window.
- For day d, the 50-day average is the average closing price for the 50 previous trading days (days d-50 to d-1).

### Step 2.8

- If less than 51 days of data, no trading
- Net gain of zero
- Repeat from beginning (step 2.1) to get next user input.

CP1	CP2	CP3	..	..	..	CP50	CP51	CP52	...
-----	-----	-----	----	----	----	------	------	------	-----

CP means closing Price

- Ignore the average closing price from day 1 to day 50
- Avg of closing price for day 51=(CP1+ CP2 +.....+ CP50)/50
- Avg of closing price for day 52=(CP2 + CP3 +..... + CP50 +CP51)/50
- Avg of closing price for day 53=(CP3 + CP4 +..... + CP50 +CP51+CP52)/50
- .....

.....

.....



# Investment Strategy for INTC

## (Now Ascending order): Step 2.6 to Forward

```

43 1985.03.04 open: 1.1666667 high: 1.1666667 low: 1.1200000 close: 1.1354167
44 1985.03.05 open: 1.1354167 high: 1.1354167 low: 1.0729167 close: 1.0937500
45 1985.03.06 open: 1.0937500 high: 1.1041667 low: 1.0729167 close: 1.0833333
46 1985.03.07 open: 1.0833333 high: 1.0833333 low: 1.0520833 close: 1.0729167
47 1985.03.08 open: 1.0729167 high: 1.0937500 low: 1.0520833 close: 1.0625000
48 1985.03.11 open: 1.0625000 high: 1.0833333 low: 1.0520833 close: 1.0625000
49 1985.03.12 open: 1.0625000 high: 1.1041667 low: 1.0575000 close: 1.0937500
50 1985.03.13 open: 1.0937500 high: 1.1041667 low: 1.0312500 close: 1.0416667
51 Day fifty: 1985.03.13, average = 1.2229167
52 1985.03.14 open: 1.0416667 high: 1.0520833 low: 1.0104167 close: 1.0208333 (average 1.2229167)
53 1985.03.15 open: 1.0208333 high: 1.0625000 low: 1.0104167 close: 1.0312500 (average 1.2204167)
54 1985.03.18 open: 1.0416667 high: 1.0833333 low: 1.0416667 close: 1.0729167 (average 1.2177083)
55 1985.03.19 open: 1.0729167 high: 1.1666667 low: 1.0729167 close: 1.1562500 (average 1.2154167)
56 1985.03.20 open: 1.1562500 high: 1.1666667 low: 1.1250000 close: 1.1458333 (average 1.2141667)
57 1985.03.21 open: 1.1458333 high: 1.1666667 low: 1.1145833 close: 1.1250000 (average 1.2135417)
58 1985.03.22 open: 1.1250000 high: 1.1666667 low: 1.1250000 close: 1.1250000 (average 1.2120833)
59 1985.03.25 open: 1.1250000 high: 1.1354167 low: 1.0575000 close: 1.0675000 (average 1.2095833)
60 1985.03.26 open: 1.0675000 high: 1.0833333 low: 1.0575000 close: 1.0729167 (average 1.2055167)
61 Buy: 1985.03.26 100 shares @ 1.0675000, total shares = 100, cash = -114.7500000
62 1985.03.27 open: 1.0729167 high: 1.1300000 low: 1.0625000 close: 1.1041667 (average 1.2011417)
63 1985.03.28 open: 1.1041667 high: 1.1562500 low: 1.1041667 close: 1.1300000 (average 1.1971833)

```

### Step 2.9

- If more than 51 days of data, compute 50-day average for the first fifty days.
- From day 51 through the second-to-last trading day, execute following strategy

2.9.1 Initially; cash = 0.0 and number of stock=0;

2.9.2 Buy criterion (here, d = 1985.03.25)

- If  $\text{close}(d) < 50\text{-day average}$  and  $(\text{close}(d) / \text{open}(d) \leq 0.97)$  buy 100 shares at price  $\text{open}(d+1)$ .




# Investment Strategy for INTC (Now Ascending order): Step 2.6 to Forward

443	1986.08.18	open: 0.8333333	high: 0.8333333	low: 0.8125000	close: 0.8229167	(average 0.8710250)
444	1986.08.19	open: 0.8229167	high: 0.8541667	low: 0.8125000	close: 0.8437500	(average 0.8660250)
445	1986.08.20	open: 0.8437500	high: 0.8958333	low: 0.8437500	close: 0.8854167	(average 0.8620667)
446	1986.08.21	open: 0.8958333	high: 0.9062500	low: 0.8854167	close: 0.8958333	(average 0.8589417)
447	Sell: 1986.08.21 100 shares @ 0.8958333, total shares = 3000, cash = -3171.0833333					
448	1986.08.22	open: 0.8958333	high: 0.9479167	low: 0.8854167	close: 0.9375000	(average 0.8562333)
449	1986.08.25	open: 0.9375000	high: 0.9479167	low: 0.9062500	close: 0.9062500	(average 0.8546750)
450	1986.08.26	open: 0.9062500	high: 0.9375000	low: 0.9062500	close: 0.9375000	(average 0.8522833)
451	1986.08.27	open: 0.9375000	high: 0.9479167	low: 0.9166667	close: 0.9375000	(average 0.8512417)
452	1986.08.28	open: 0.9270833	high: 0.9479167	low: 0.9270833	close: 0.9270833	(average 0.8512417)
453	1986.08.29	open: 0.9270833	high: 1.0104167	low: 0.9270833	close: 0.9583333	(average 0.8508250)
454	1986.09.02	open: 0.9583333	high: 0.9687500	low: 0.8958333	close: 0.9062500	(average 0.8502000)
455	1986.09.03	open: 0.8958333	high: 0.9166667	low: 0.8854167	close: 0.9166667	(average 0.8483250)
456	1986.09.04	open: 0.9166667	high: 0.9687500	low: 0.9166667	close: 0.9479167	(average 0.8470750)
457	1986.09.05	open: 0.9479167	high: 0.9687500	low: 0.9166667	close: 0.9375000	(average 0.8460333)
458	1986.09.08	open: 0.9375000	high: 0.9375000	low: 0.9166667	close: 0.9270833	(average 0.8447833)
459	1986.09.09	open: 0.9270833	high: 0.9479167	low: 0.9062500	close: 0.9062500	(average 0.8441583)
460	1986.09.10	open: 0.8958333	high: 0.9166667	low: 0.8645833	close: 0.9062500	(average 0.8431167)
461	1986.09.11	open: 0.8958333	high: 0.9062500	low: 0.8541667	close: 0.8541667	(average 0.8420750)
462	1986.09.12	open: 0.8541667	high: 0.8645833	low: 0.8333333	close: 0.8333333	(average 0.8402000)
463	1986.09.15	open: 0.8437500	high: 0.8645833	low: 0.7812500	close: 0.8020833	(average 0.8381167)
464	1986.09.16	open: 0.8125000	high: 0.8437500	low: 0.7916667	close: 0.8333333	(average 0.8356167)
465	Buy: 1986.09.16 100 shares @ 0.8125000, total shares = 3100, cash = -3260.3333333					
466	1986.09.17	open: 0.8333333	high: 0.8645833	low: 0.8125000	close: 0.8541667	(average 0.8343667)
467	1986.09.18	open: 0.8437500	high: 0.9062500	low: 0.8437500	close: 0.8854167	(average 0.8333250)
468	1986.09.19	open: 0.8854167	high: 0.8854167	low: 0.8333333	close: 0.8541667	(average 0.8331167)
469	1986.09.22	open: 0.8645833	high: 0.8854167	low: 0.8333333	close: 0.8750000	(average 0.8327000)
470	Sell: 1986.09.22 100 shares @ 0.8697917, total shares = 3000, cash = -3181.3541667					
471	1986.09.23	open: 0.8958333	high: 0.9166667	low: 0.8437500	close: 0.8595833	(average 0.8341583)
472	Sell: 1986.09.23 100 shares @ 0.8777083, total shares = 2900, cash = -3101.5833333					
473	1986.09.24	open: 0.8645833	high: 0.8750000	low: 0.8125000	close: 0.8333333	(average 0.8353083)

## 2.9.3 Sell criterion (here, d= 1986.08.21)

- If shares  $\geq 100$  and  $\text{open}(d) > 50\text{-day average}$  and  $\text{open}(d) / \text{close}(d-1) \geq 1.01$ , sell 100 shares at price  $(\text{open}(d) + \text{close}(d))/2$ .

# Investment Strategy for INTC

Result Grid   Filter Rows: <input type="text"/> Export: 					
	TransDate	OpenPrice	HighPrice	LowPrice	ClosePrice
▶	1985.01.02	28.00	28.25	27.25	27.50
	1985.01.03	27.50	28.50	27.50	28.00
	1985.01.04	28.00	28.75	28.00	28.50
	1985.01.07	28.50	29.25	28.25	29.25
	1985.01.08	29.25	29.75	27.75	28.25
	1985.01.09	28.25	29.25	27.75	28.75
	1985.01.10	28.75	30.25	28.75	30.00
	1985.01.11	30.00	30.50	29.75	30.50
	1985.01.14	30.50	31.25	29.75	31.00
	1985.01.15	31.00	31.75	31.00	31.25
	1985.01.16	31.25	32.00	30.75	31.00
	1985.01.17	31.00	31.25	29.75	30.50
	1985.01.18	30.50	31.25	30.25	31.00
	1985.01.21	31.00	31.75	30.75	31.75
	1985.01.22	31.75	32.50	31.75	32.00
	1985.01.23	32.00	32.00	31.25	31.50
	1985.01.24	31.50	32.50	31.00	31.75
	1985.01.25	31.75	31.87	31.25	31.75
	1985.01.28	31.75	31.75	30.75	30.75
	1985.01.29	30.75	31.25	30.25	31.25
	1985.01.30	31.25	32.25	31.13	31.25
	1985.01.31	31.25	31.50	30.50	30.75

## •2.9.4 Transaction Fee

- For either a buy or sell transaction, cash is reduced by a transaction fee of \$8.00.

## •2.9.5 If neither the buy nor the sell criterion is met, do not trade on that day.

## •2.9.6 Regardless of trading activity, update 50-day average to reflect average over last 50 days, and continue with day d+1.

## Step 2.10

After processing data through second-to-last day, if there are any shares remaining, on the last day add ***open(d) \* shares remaining*** to cash to account for the value of those remaining shares (No transaction fee applies to this).

.....  
 .....

# Sample Outputs one more time

Database connection is established.

Enter a ticker symbol [start/end dates]: INTC

Intel Corp.

2:1 split on 2000.07.28 129.12 --> 65.44

2:1 split on 1999.04.09 130.81 --> 61.62

2:1 split on 1997.07.11 153.81 --> 77.25

2:1 split on 1995.06.16 116.12 --> 58.50

2:1 split on 1993.06.04 112.75 --> 60.12

3:2 split on 1987.10.28 31.75 --> 21.75

6 splits in 7470 trading days

Transaction executed

= How many buy or sales you have done in this time period

Executing investment strategy

Transactions executed: 690

Net cash: 14717.72

Enter ticker symbol [start/end dates]: INTC 1980.01.01 1999.12.31

Intel Corp.

2:1 split on 1999.04.09 130.81 --> 61.62

2:1 split on 1997.07.11 153.81 --> 77.25

2:1 split on 1995.06.16 116.12 --> 58.50

2:1 split on 1993.06.04 112.75 --> 60.12

3:2 split on 1987.10.28 31.75 --> 21.75

5 splits in 3791 trading days

Enter ticker symbol [start/end dates]: T 2000.01.01 2014.08.18

AT&T Inc

0 splits in 3679 trading days

Executing investment strategy

Transactions executed: 148

Net cash: -1568.00

Executing investment strategy

Transactions executed: 358

Net cash: 44953.95

The End