Questions.

1. Calculate the annual and daily growth and volatility estimates for the DJIA and 4 stocks comprising the ISCO index. Note that the daily growth values will be pretty small.

mu_daily	0.00019	-0.00047	0.000626	0.000681	0.000031
sigma_daily	0.031127	0.037504	0.016556	0.016544	0.012306
mu_annual	**0.047874**	**-0.11839**	**0.157711**	**0.171523**	**0.00786**
sigma_annual	**0.494119**	**0.595354**	**0.262822**	**0.262628**	**0.195345**
r_bar	-0.00029	-0.00117	0.000489	0.000544	-4.5E-05
sigma	0.031127	0.037504	0.016556	0.016544	0.012306

	CSC	CSCO	SYY	SMJ	DJIA
8/30/00	76.19	66.56	20.97	19.78	11103

For convenience and reference, we also calculate parameters for the generated indexes:

|		
mu_daily	-0.000060	0.000042
sigma_daily	0.017877	0.017377
mu_annual	**-0.01502**	**0.010646**
sigma_annual	**0.283796**	**0.275853**
r_bar	-0.00022	-0.00011
sigma	0.017877	0.017377

	D-ISCO	G-ISCO
8/30/00	11103	11103
First Value	45.875	38.08325
Multiplier	242.0272	291.5455

2. Plot price charts for the DJIA and each component stock.

        First we show the Dow as compared with the ISCO indexes.
Note that the scale plotted is the same for all graphs, enabling
one to immediately see that the Dow held up better than ISCO. A
little harder to see is that the G-Isco got higher than the Dow
ever did (around day 70).
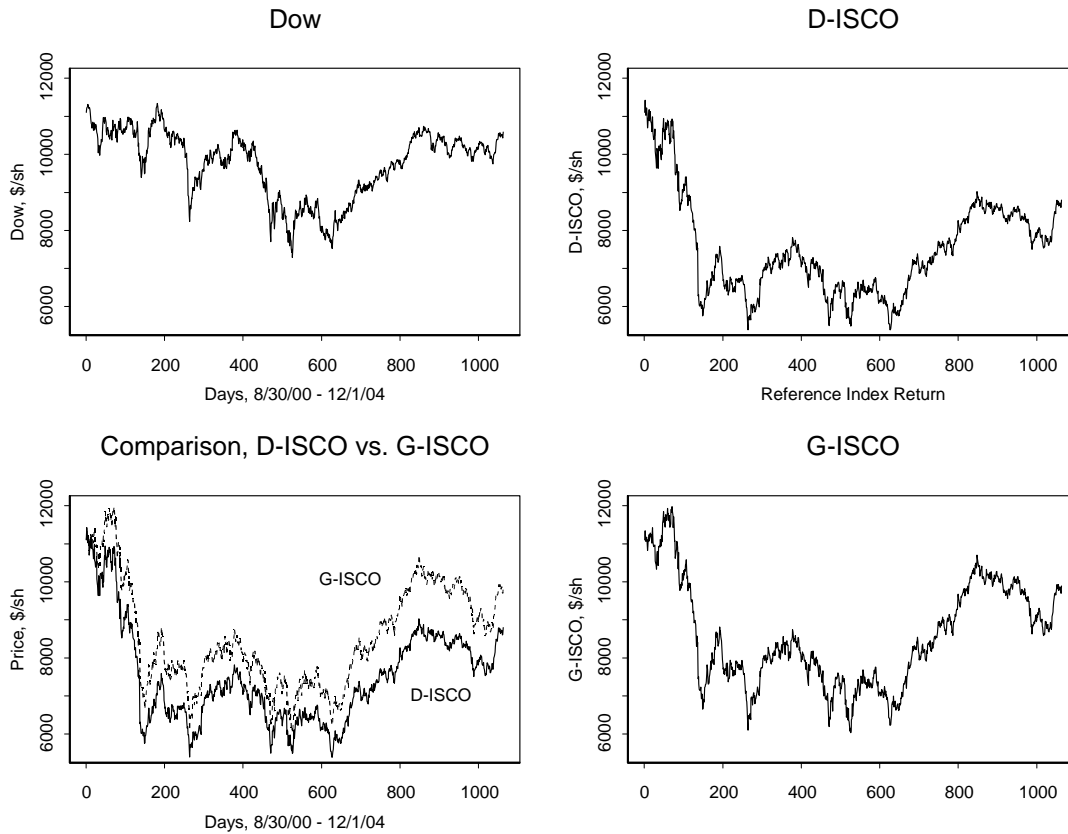
The ISCO stock components are shown below:

CSC

SYY

CSCO

SMJ
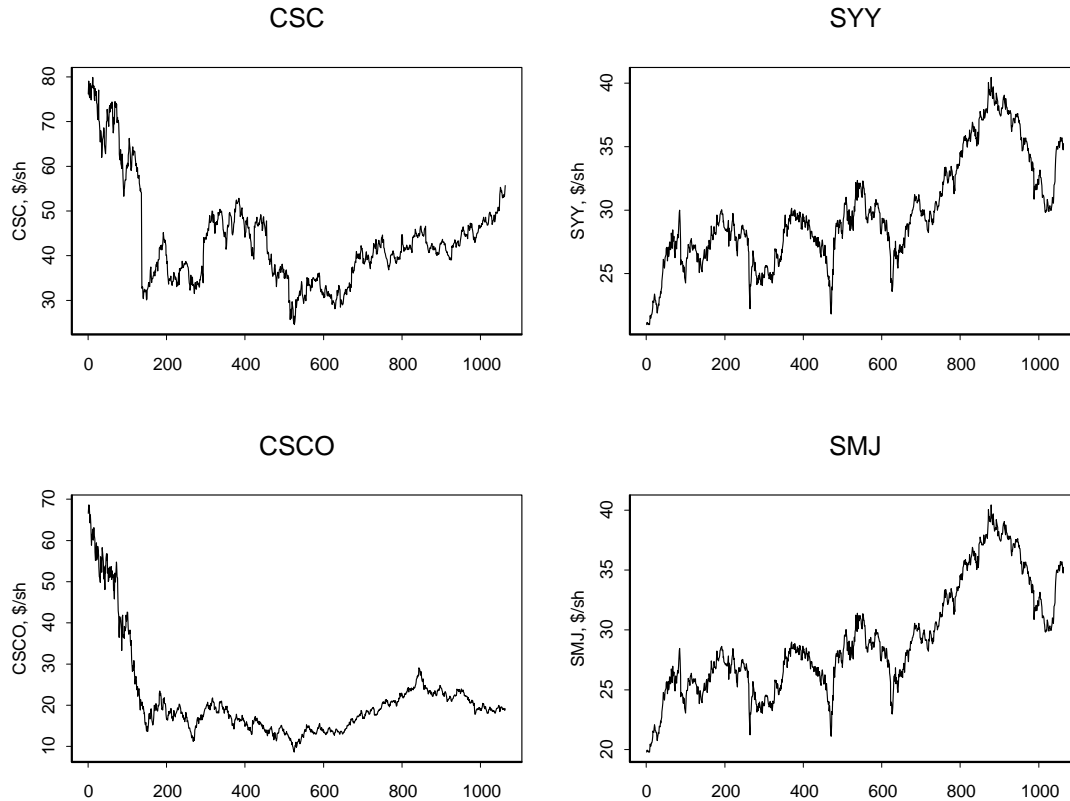
3. Perform a correlation and regression analysis for each ISCO stock and the DJIA. Try to make statements on what the correlations mean. Plot your regression lines.

```
We provide what is known as the "correlation matrix", which is
symmetric, with Corr(Xᵢⱼ) = Corr(Xⱼᵢ).  On the diagonals are the
correlation of each index/stock with itself (1.00); the
intersection of any row/column (Xᵢⱼ)is the correlation between
that pair.
```

We provide what is known as the "correlation matrix", which is symmetric, with $Corr(X_{ij}) = Corr(X_{ji})$. On the diagonals are the correlation of each index/stock with itself (1.00); the intersection of any row/column (X_{ij}) is the correlation between that pair.
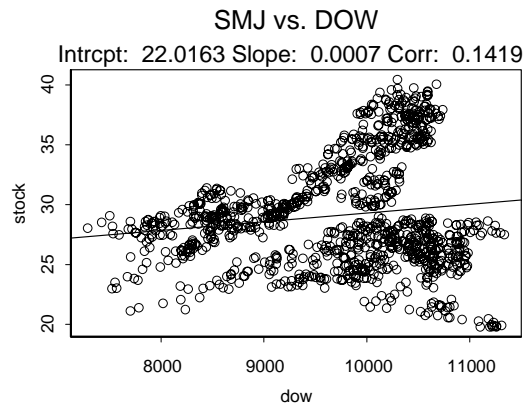
```
        Correlation Analysis
        --------------------

        > round(cor(Matrix), 4)

                 dow     csc     csco      syy      smj
        dow   1.0000   0.6644   0.5716   0.1657   0.1419
        csc   0.6644   1.0000   0.8661  -0.2066  -0.2118
        csco  0.5716   0.8661   1.0000  -0.1998  -0.2044
        syy   0.1657  -0.2066  -0.1998   1.0000   0.9972
        smj   0.1419  -0.2118  -0.2044   0.9972   1.0000
```
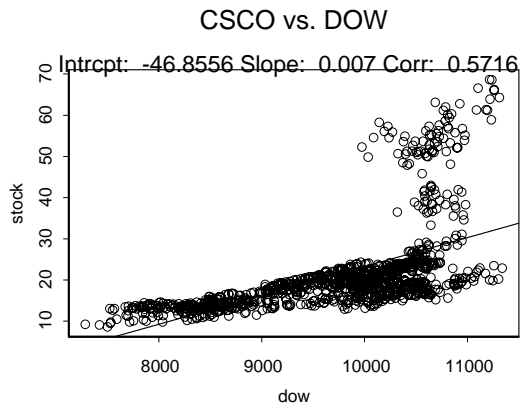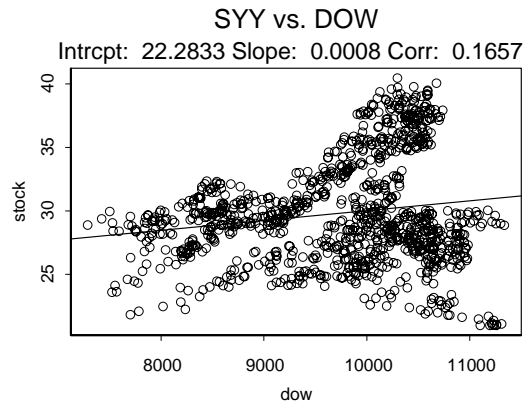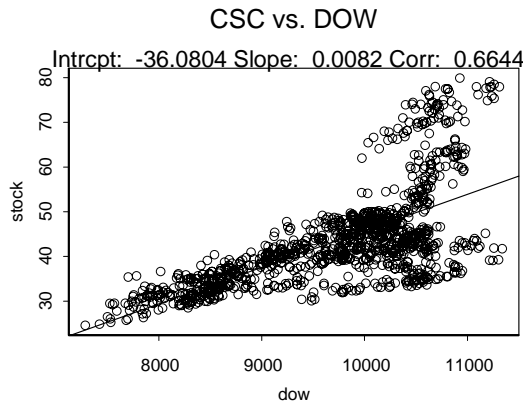
Some discussion of the correlations in the above table follows.
We note that all the stocks are positively correlated with the
DOW; CSC and CSCO moderately-high correlated (~.60), and SYY/SMJ
are faintly correlated. The correlation between CSC and CSCO is
high, almost 87%, but not as high as SYY and SMJ, at 97%. Both
pairs of stocks are in the same sector (tech and food service),
so this is not surprising. Both SYY and SMJ are negatively
correlated with CSCO/CSC, which is good in portfolio theory.

That SYY/SMJ are so close, as indicated by the price charts also,
would lead one to more closely inspect the data, which in fact
reveals that the prices are exactly the same for these stocks
from 7/1/03 thru 12/1/04 (about 300 days). If one recalculates,
the correlation reduces to .88, which still shows the sector
correlation, and does not change the other analysis except that
the annual parameters are SYY~GBM(.16, .26); SMJ~GBM(.26, .30).

Notice that the regression line slopes are very small, but that
is only because the Indexes themselves are so large. If the
index changes by +1 point, the stock changes by +.008 point.

CSC vs. DOW
Intrcpt: -36.0804 Slope: 0.0082 Corr: 0.6644

SYY vs. DOW
Intrcpt: 22.2833 Slope: 0.0008 Corr: 0.1657

CSCO vs. DOW
Intrcpt: -46.8556 Slope: 0.007 Corr: 0.5716

SMJ vs. DOW
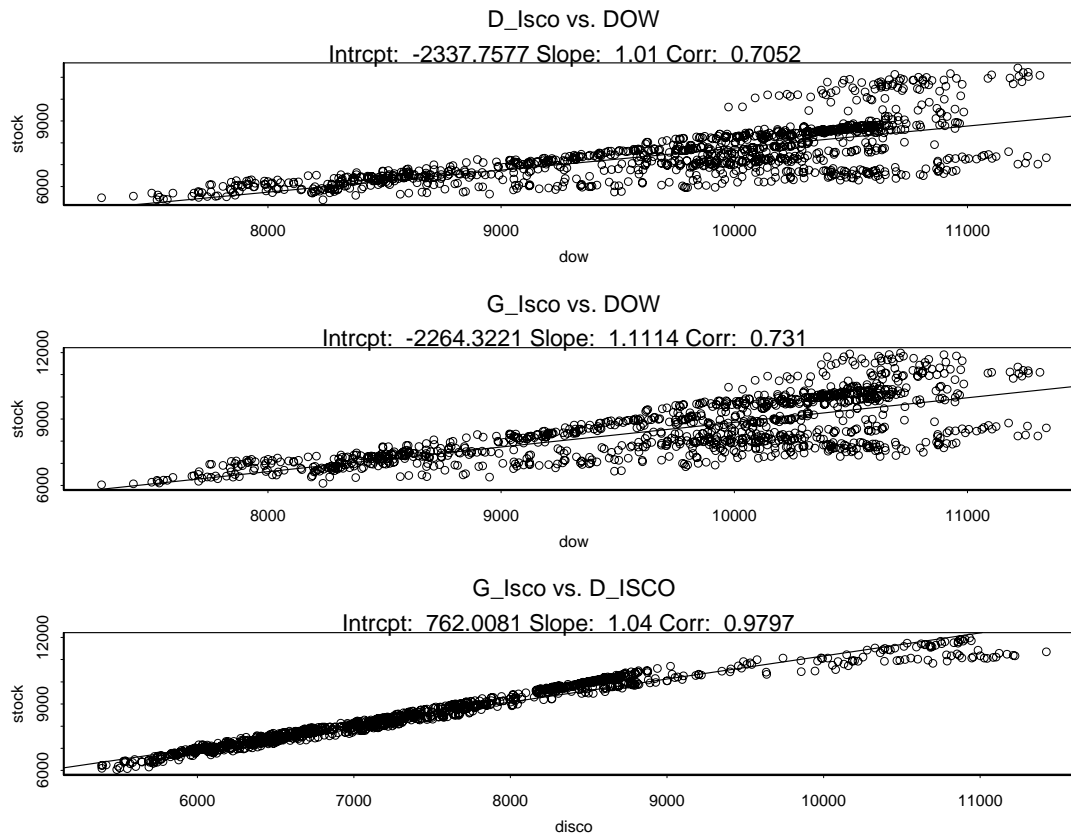Intrcpt: 22.0163 Slope: 0.0007 Corr: 0.1419

These are correlations between the indexes. Note that the ISCO
indexes are only about 70% correlated with the DOW, indicating a
possible portfolio diversification benefit.

```
> round(cor(Matrix), 4)

            dow   disco  gisco
  dow    1.0000  0.7052 0.7310
disco    0.7052  1.0000 0.9797
gisco    0.7310  0.9797 1.0000
```

The DISCO/GISCO correlation is almost 98%, which is not
surprising since they are composed of the exact same stocks.

WE can also look at the regressions among the indexes (below).
Note here that the regression line slope is much larger and
easier to get a feel for.

D_Isco vs. DOW
Intrcpt: -2337.7577 Slope: 1.01 Corr: 0.7052

G_Isco vs. DOW
Intrcpt: -2264.3221 Slope: 1.1114 Corr: 0.731

G_Isco vs. D_ISCO
Intrcpt: 762.0081 Slope: 1.04 Corr: 0.9797

4. Calculate and plot the D-ISCO and G-ISCO indexes, scaling to the value of the DJIA
on 8/30/2000.

```
        See plots for problem 1
```

5. Find the correlation between the 4 component stocks and each of the ISCO indexes.
Comment on this correlation with respect to those for the DJIA.

```
Correlation Analysis
--------------------
> round(cor(Matrix), 4)

       disco  gisco      csc     csco      syy      smj
disco 1.0000 0.9797   0.8822   0.8849   0.2023   0.1973
gisco 0.9797 1.0000   0.7976   0.8107   0.3523   0.3474
  csc 0.8822 0.7976   1.0000   0.8661  -0.2066  -0.2118
 csco 0.8849 0.8107   0.8661   1.0000  -0.1998  -0.2044
  syy 0.2023 0.3523  -0.2066  -0.1998   1.0000   0.9972
  smj 0.1973 0.3474  -0.2118  -0.2044   0.9972   1.0000
```

6. Would it be better to invest in the ISCO index or the DJIA? Why?

        There are many ways to answer this question. They would
fall into an analysis of the "buy-and-hold (BAH)," "portfolio re-
balancing", or "trading" strategies. We can measure their
performance over a horizon of 4 years, assuming an initial
investment of say, $100,000.

For the BAH strategy, the ISCO indexes UNDERPERFORMED the DOW.

Initial value   100,000
Terminal Portfolio Value

	8/30/00	8/31/01	8/30/02	8/29/03	8/30/04	12/1/04	BAH
CSC	76.19	37.6	36.83	42.57	46.16	55.73	73,146
CSCO	66.56	16.33	13.82	19.14	19.03	19.15	28,771
SYY	20.97	28.02	28.36	30.89	32.1	35.24	168,050
SMJ	19.78	26.78	27.44	30.89	32.1	35.24	178,160
DJIA	11,103	9,950	8,664	9,416	10,123	10,590	95,382
DISCO	11,103	6,579	6,441	7,472	7,812	8,795	79,215
GISCO	11,103	7,596	7,314	8,657	8,956	9,892	89,094

Even with the DOW, however, you lose money over a really long
time (4 years). You lose 5% on the DOW, 21% on the DISCO, and
11% on the GISCO. What drives the game is noticing that "If I'd
invested in SYY I would have made 68%."

The portfolio re-balancing or trading strategy would show a
different story. Since the markets started a major decline
(based on technical performance) starting March 2000, you would
be careful when investing huge chunks of money. Suppose you had
a protection scheme in place which sold you out after your
portfolio showed a slight loss after 8/2000; you stayed out until
8/02, and then re-entered on 9/1/02 (again, based on your crystal
ball). Then we would have the following returns:

	Buy Here				Initial value 100,000 Terminal Portfolio Value	
	9/1/02	8/29/03	8/30/04	12/1/04	BAH	Buy in 02
CSC	36.83	42.57	46.16	55.73	73,146	151,317
CSCO	13.82	19.14	19.03	19.15	28,771	138,567
SYY	28.36	30.89	32.1	35.24	168,050	124,260
SMJ	27.44	30.89	32.1	35.24	178,160	128,426
DJIA	8,664	9,416	10,123	10,590	95,382	122,240
DISCO	6,441	7,472	7,812	8,795	79,215	136,552
GISCO	7,314	8,657	8,956	9,892	89,094	135,248

Here both ISCO indexes did much BETTER than the DOW. Even if you
suffered a 5% loss in 2000, you'd still be way ahead after
staying out for 2000 to mid-2002, and the ISCO investment would
have been much better.

Another way to compare relative performance would be to look at
the difference between the indexes (such as DISCO – DOW) each
day. In statistics and mathematics, this usually is interpreted
as a ratio, like GISCO/DOW, which begins at 1.00 on day 1, or
0.00 if you subtract 1.0. This appears as:

Comparison, G-ISCO vs. DOW

Days, 8/30/00 - 12/1/04
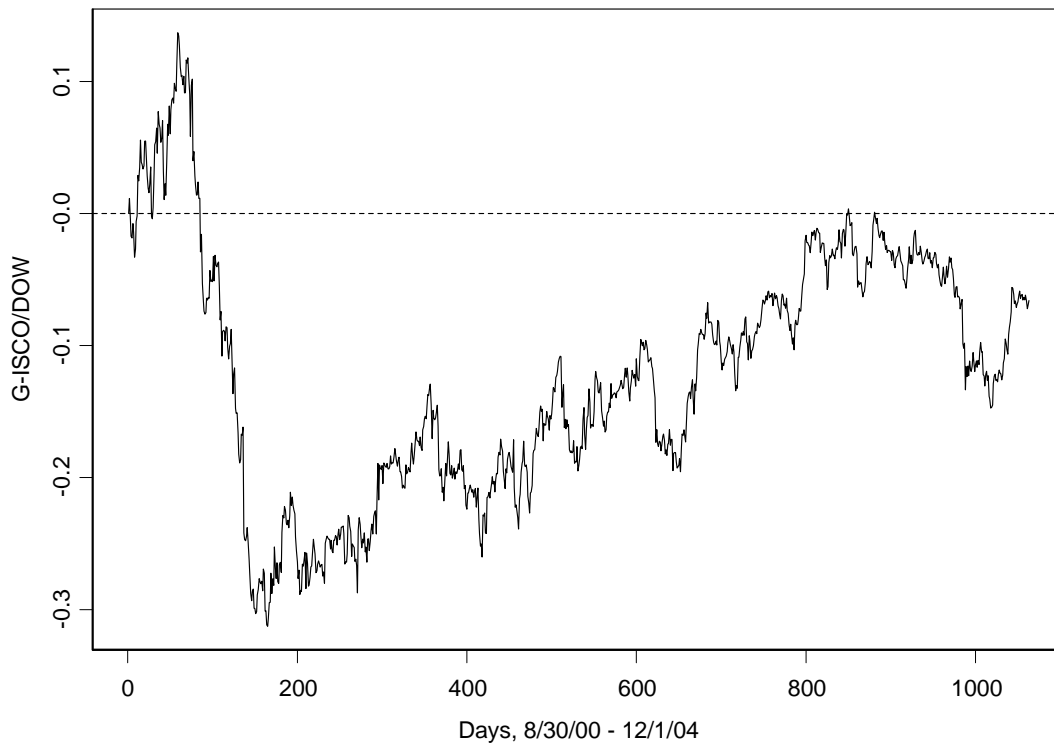
For the BAH strategy, if you could buy the ratio, it is apparent
that the Isco index immediately underperforms the Dow, and that
you would have to wait 820 days or so just to break even. And
then, you would immediately be disappointed again by the end of
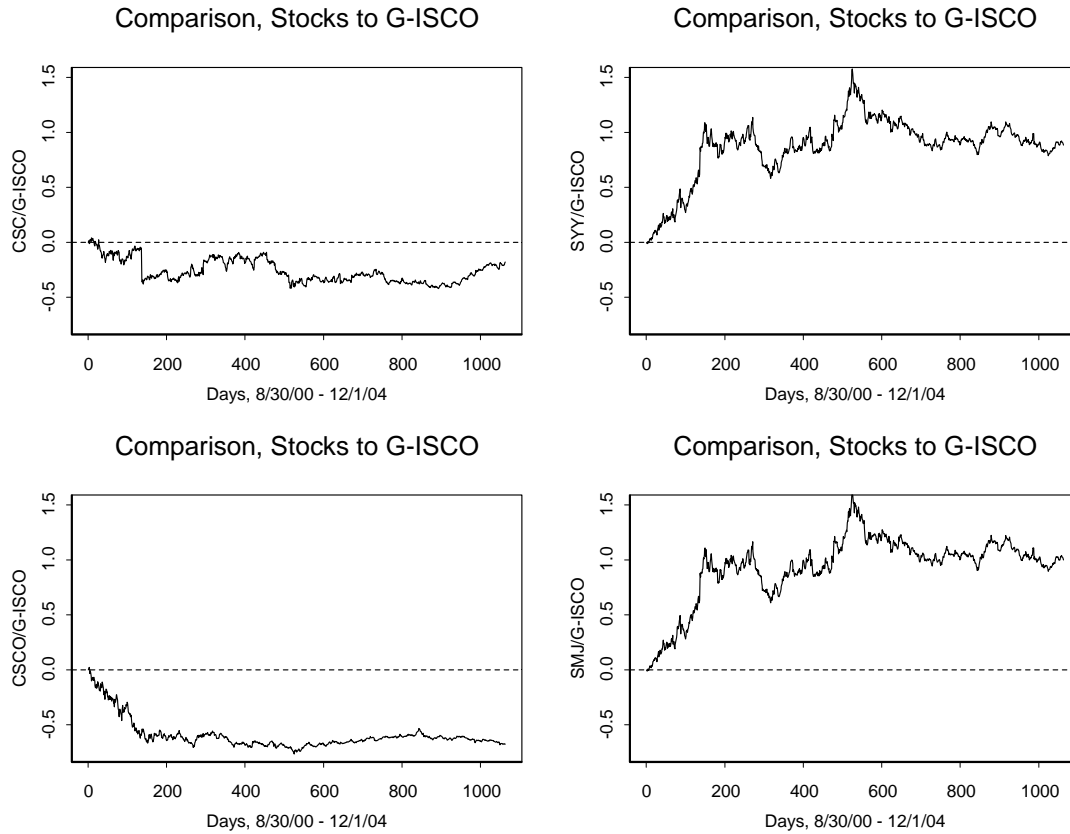the holding period when Isco had lost 15% by day 1000.

For a trading approach, the comparison chart clearly shows that
the BEST time to sell the DOW and go long the ISCO would be
around day 250, or 1 year after the 8/00 start date.

Unfortunately, in the past you don't know the future. Using
technical analysis, you would have waited for the pop up around
day 180, wait for it to settle back down (around day 200); when
it did not go below the previous low, you would have gone long
around day 300. Note that the market was aware of all this; do
you see how at day 300 the big jump over the top from day 199?
So "everyone" was getting in on this signal.

Sure enough, the ratio declines again to a low around day 400,
which would have made many people sell the ratio. If you stayed
in, however, buying on day 300 (November 12, 2001) would have
handsomely rewarded you.

7. Would it be better to invest in the ISCO index or for an individual stock in the ISCO
index? Why?

       This one is a little trickier, since you would have to know
which Isco stock would outperform the others. There is no way to
tell that SYY/SMJ would do better than CSC/CSCO, other than a
prediction that the "tech-wreck" which began March 2000 would
continue for 3 years. There are be divergence measures you could
employ to tell which sector was doing better, in which case it
would have been better to reduce your allocation of the tech
components and increase those for the food service stocks. This
is very hard to do, however, and most advisors fail.

Comparison, Stocks to G-ISCO

Comparison, Stocks to G-ISCO

Comparison, Stocks to G-ISCO

Comparison, Stocks to G-ISCO

8. Suppose after the Bush re-election each of the ISCO stocks gained 10% return on the day. What are the p-values for these returns how often should such a return be encountered in terms of years (or centuries, or millennia)?

Recall the daily parameters for our 4 stocks:

	CSC	CSCO	SYY	SMJ
mu_daily	0.00019	-0.00047	0.0006258	0.0006806
sigma_daily	0.031127	0.037504	0.0165562	0.016544
R	1.10	1.10	1.10	1.10
$r = \ln(R)$	0.095	0.095	0.095	0.095
$z = \dfrac{r - \mu_{daily}}{\sigma_{daily}}$	3.056	2.554	5.719	5.720
p-value	0.002244	0.010653	1.075E-08	1.069E-08
days	445.7	93.9	93,028,415	93,531,720
years	1.77	0.37	369,160	371,158

A 10% move on a stock with a daily volatility of 1% is a "10-sigma" event; a 10% move on a stock such as CSCO with an almost 4% daily volatility is only 2.5 sigma, which should occur about 3 times a year, based on the p-values and time-between-events indicated in the table.

9. Do your results in question 8 make you wonder about the Normality assumptions of the log of the daily returns?

Since we have so many greater than 6σ events happening so often, and since the Normal distribution just hardly ever gives values over 4σ , no matter how long you wait, it begs the question, "how can these returns be normal?" Much research has gone into this, and many solutions are attempted, mostly exotic, fat-tailed continuous distribution models, some of which do not even have finite moments such as the mean or standard deviation. Other models are employed which include a time-varying volatility concept.

Check out these "humdingers" (you can do this daily using http://fast2.quote.com/fq/stocktrak/group?mode=IwatchMenu1&group=IndustryWatch/ga&mode=ga&page=stocks)

Wednesday, December 15, 16:02 PM Eastern Time

Symbol	Name	Last	Change	Change%	Volume
LVS	NYSE:LVS (IPO)	46.56	17.56	60.55	26,260,300
CKCM	Click Commerce, Inc.	13.13	2.29	21.13	795,867
OPWV	Openwave Systems Inc	16.11	1.89	13.29	5,940,279
TRGL	Toreador Resources Corporation	18.73	2.16	13.03	451,379
LEN	Lennar Cp Cl A	56.35	5.52	10.85	9,374,600
LEN.B	Lennar Corp Cl B	51.85	4.97	10.6	580,000
MSO	Martha Stewart Liv	30.05	2.81	10.31	2,627,000

We see several 10% returns today; are they significant? I just checked three, and here are the parameters and the z-scores:

```
ckcm00-04    Mu_hat =   0.58021021    Sigma_hat =   1.27300990
opwv99-04    Mu_hat =   0.47903508    Sigma_hat =   1.21799147
len00-04     Mu_hat =   0.40983891    Sigma_hat =   0.41790226
```

Clearly CKCM and OPWV have volatilities over 120% per year! So these moves might not be that large. For Lennar homes, with 42% volatility, the 10% move might be unusual.

```
CKCM  mu_1d: 0.0023   sd_1d: 0.0802   z:  2.359  p_val:0.92E-02 Yrs: 0.433
OPWV  mu_1d: 0.0019   sd_1d: 0.0767   z:  1.602  p_val:0.55E-01 Yrs: 0.072
LEN   mu_1d: 0.0016   sd_1d: 0.0263   z:  3.854  p_val:0.58E-04 Yrs: 68.4
```

Stocks too low priced to really invest in; these show "non-normal" returns more often

Symbol	Name	Last	Change	Change%	Volume
SQNM	Sequenom, Inc.	1.63	0.52	46.84	16,066,866
NAVI	NaviSite, Inc.	3.02	0.77	34.22	117,569
CDIC	CardioDynamics International Corporation	5.37	1.16	27.55	1,729,825
EVST	Everlast Worldwide Inc.	4.55	0.92	25.34	122,113
BWNG	Broadwing Corporation	8.48	1.48	21.14	5,039,583
LOUD	Loudeye Corporation	2.49	0.4	19.13	16,350,960
CVSN	ChromaVision Medical Systems, Inc.	1.69	0.22	14.96	788,151
HANA	Hanaro Telecom, Inc.	3.44	0.44	14.66	153,295
ATCO	American Technology Corporation	7.50	0.88	13.29	1,080,026
WGRD	WatchGuard Technologies, Inc.	4.45	0.48	12.09	1,041,819
ARIA	ARIAD Pharmaceuticals, Inc.	7.17	0.77	12.03	2,894,132
PYX	Playtex Products Inc	7.70	0.66	9.37	963,700

10. What other questions does your analysis bring to mind?

Notice that we did not attempt to perform any hypothesis testing
in this lab. You could think of some tests. For example, are
CSC and CSCO really any different? The Null hypothesis would
assume the lognormal model, and that the parameters are the same.
Your test statistics would be the difference. Under the null,
you could clump together all the CSC and CSCO returns, draw two
samples, calculate the parameters, subtract them, and calculate
the probability that they exceed a certain value.

There are other tests too, such as distributional tests, non-
lognormal tests, etc. That would be a follow-on to this lab, but
would be questions you might have.

Other questions might be:

- How can anyone make money in these situations?

- What is stealth correlation, and what does correlation
  mean in the context of financial markets, especially
  since the correlation changes all the time?

- How could we model a trading scheme and test it in
  time?

- What other statistical tools are required to
  understand financial markets?

- Is there some sort of other model which could better
  explain the financial markets? We resort to
  stochastic models when a physical model is too

complicated. Perhaps there is a better "physical" paradigm that would account for (and hopefully predict) these markets.

- How is it that stock market analysts make 6-figures when they are as clueless as the average monkey throwing darts at a page of mutual funds?

These and many other questions come to mind after performing the simple tests we have studied in the course.

The following are some questions raised by students in the past:

o   Is there a political regime effect on the markets? I.e., is there a significant return during periods of "pro-business" vs. "anti-business" political regimes?

o   "A better question would be, what's the probability of a massive market withdrawal form American markets in favor of unstable, developing nations due to a Bush reelection?"

o   What sorts of analyses are needed to correlate market activity with business cycles?

o   How does this fit into an overall seasonality of stocks/sectors/markets?

o   What impacts do "terrorist attacks" have on the markets in general and which sectors or markets do well in these events? This is known as "structural change" in the time series literature.

o   General probabilistic question: What is the probability that an individual stock increases and how does this relate to its GBM growth parameters?

o   Is there fruit in coefficient of variation analysis for stocks (sigma/mu)?

o   How does one measure "consumer confidence," and does it attempt to discover individual psychology

o   Behavioral market efficiency: what causes investors to scare away from "undervalued" stocks when analysts are making strong buy recommendations?

o   Is there a statistically significant pattern on the scatterplots for SYY/DOW and SJM/DOW?

Sure looks like one.   "I noticed some funny patterns
in the scatterplots; it seemed to me it was likely due
to difference scenarios at different times."

   o   If one regresses the DOW on CSCO, one sees
a pattern which looks like could be brought linear via
a simple transform. Explore this.

   o   How develop a model which takes into
account the expectations of the market (without
modeling them as part of the EMH diffusion).

   o   What other type of regression might be
appropriate (besides linear)?

   o   How should one test for non-normality of
the market returns?

   o   Since stocks trade "continuously", how do
we justify a daily sampling given the Nyquist sampling
theory?

Please feel free to call or write if you have any questions,
former STAT 280 students always have unlimited technical support.