15.459 Financial Data Science and Computing - Project E

Devin Connolly

November 30, 2019

1. **Question 1 - Data Modelling**

(a) The entity definitions for the database designed to model WTI Futures are as follows:

- Oil Futures - This is the primary entity in the table, and it should capture all information specific that describes a single futures contract. That should include a description of the security, its FIGI number, key dates associated with it, tick sizes and values, as well as where it is traded and who underwrites it.
- Daily Prices - This entity captures the historical price data of the security. For each day, it should store the OHLC data for the security.
- Trades - This entity captures all the trades made on a specific contract. It should contain the volume traded, the time and date, and the price at which the trade occured.
- Exchange - This should describe the exchange on which each security is traded.
- Underwriter - This should the describe the underwriters of the various securities.

(b) The logical entity diagram displays the entities in table form, showing the data points required for each entity (PK/FK designations are included in the physical diagram).

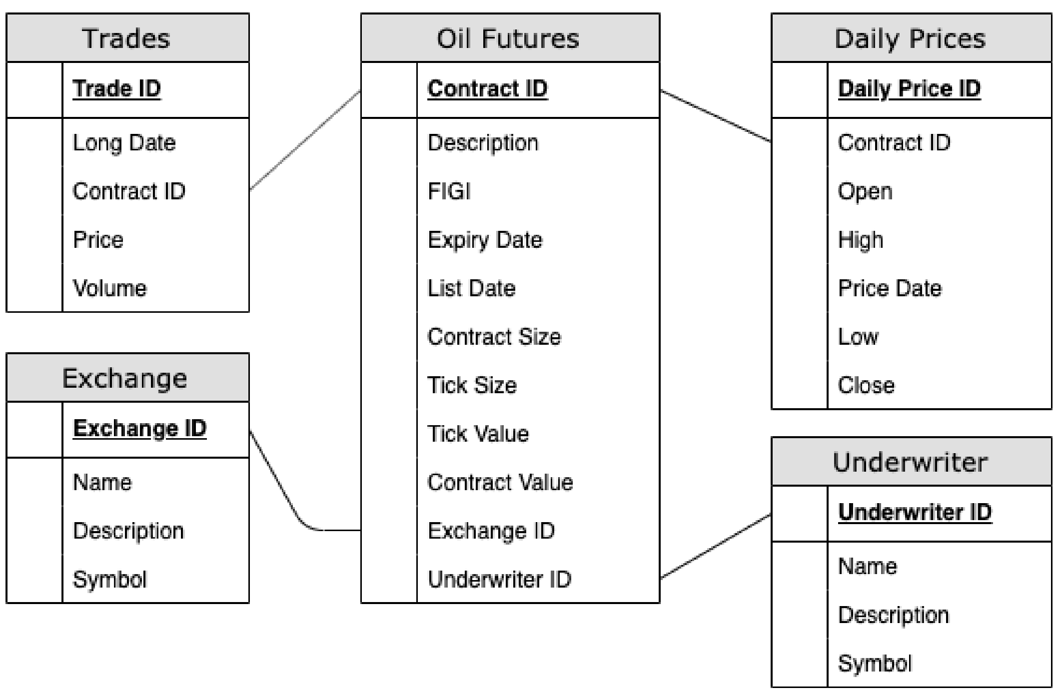

Figure 1: Logical Entity Diagram for WTI Futures

(c) The physical entity diagram features not only the entities and corresponding data points (as before), but also includes relationship links, primary and foreign keys, and data types.

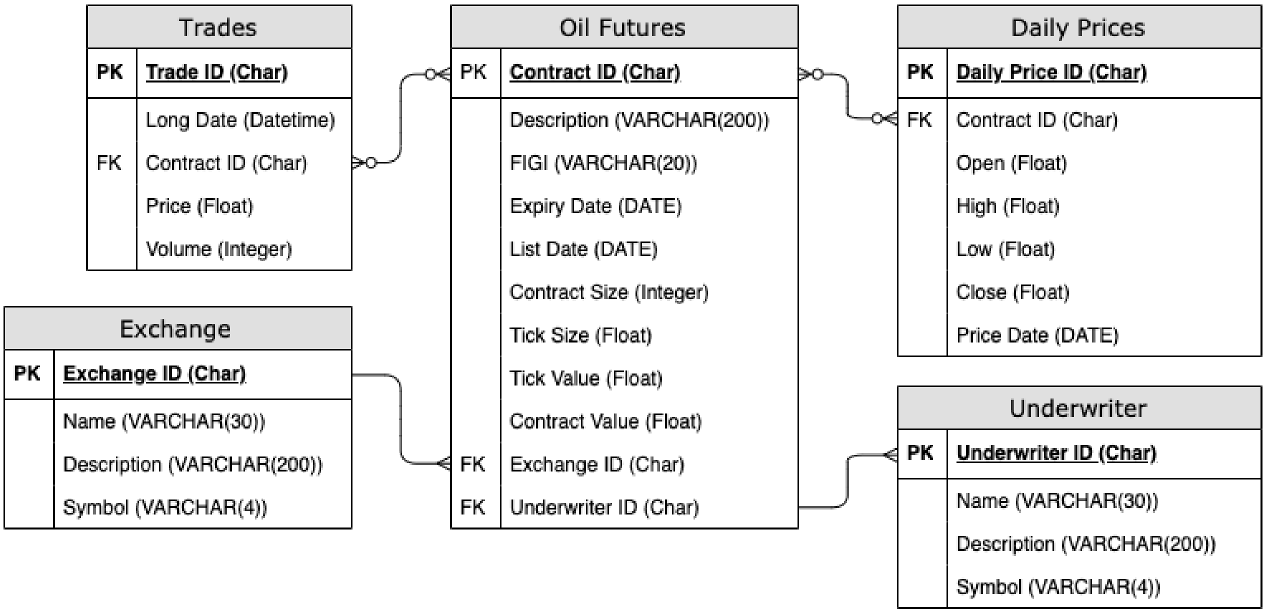

Figure 2: Physical Entity Diagram for WTI Futures

2. **Question 2 - Indices and Tradable Proxies**

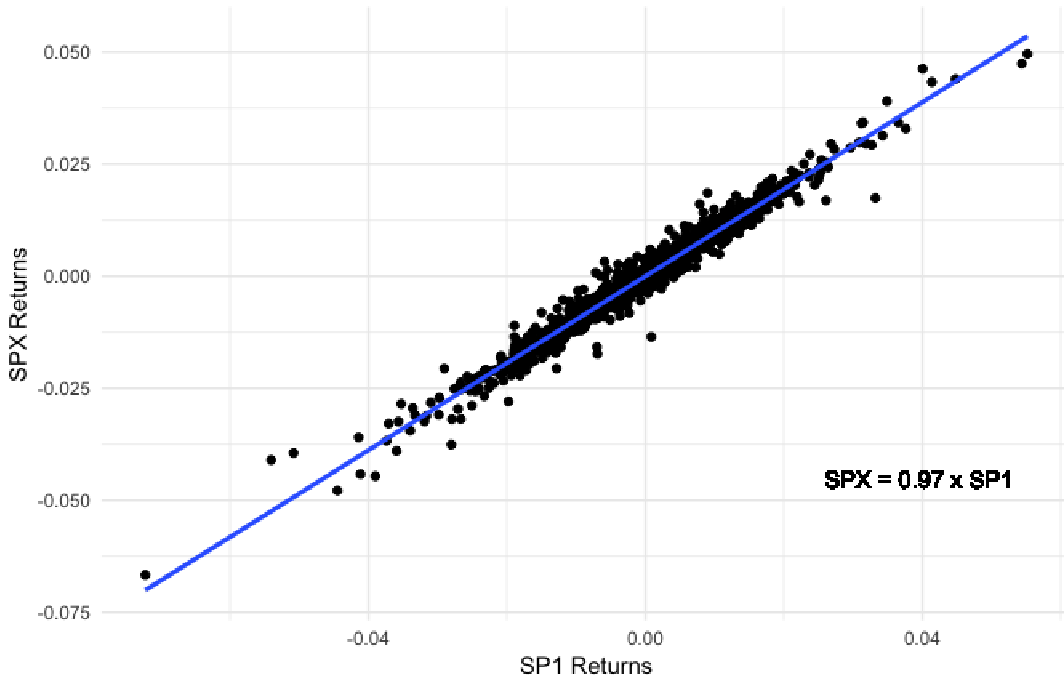

Figure 3: Scatterplot of SPX Spot Returns vs SP1 Index Returns

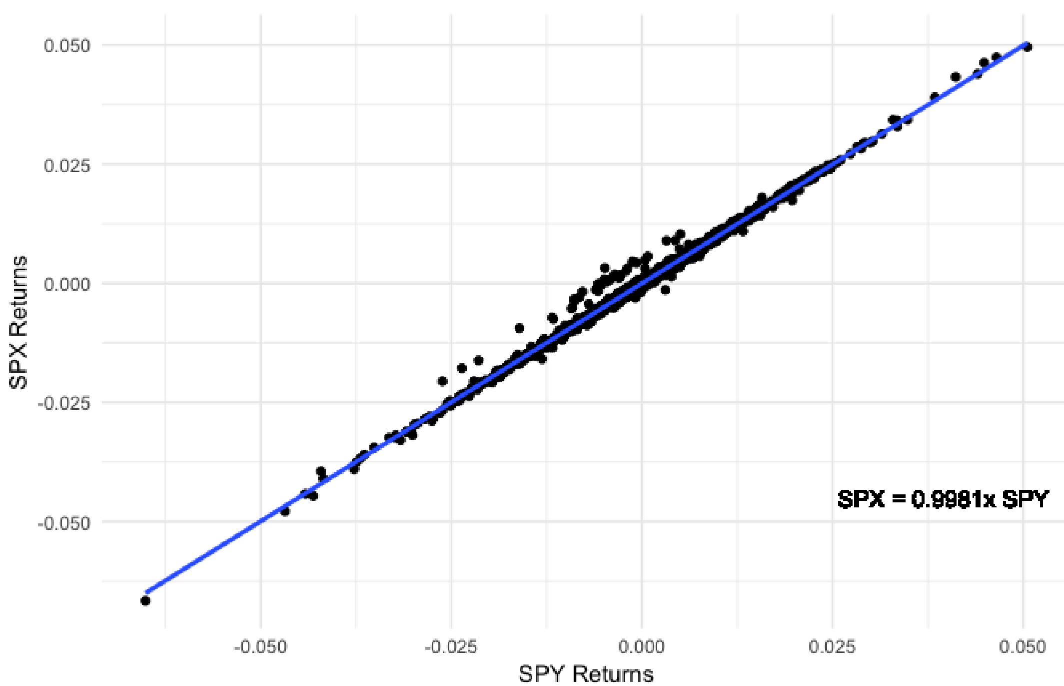

Figure 4: Scatterplot of SPX Spot Returns vs SPY Index Returns

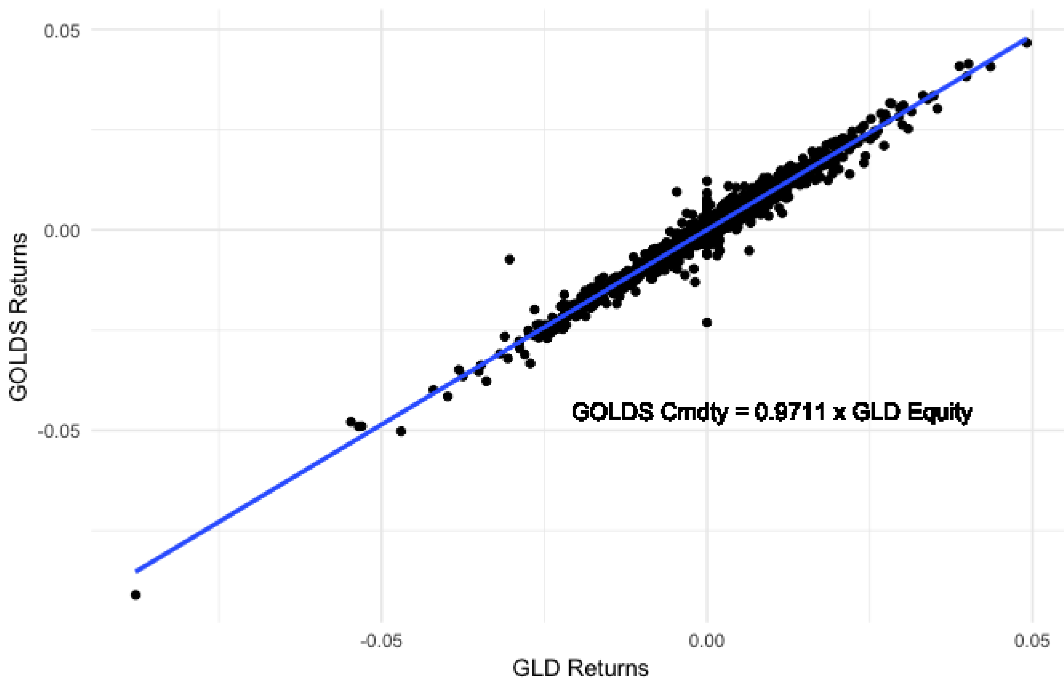

Figure 5: Scatterplot of Gold Spot Returns vs Gold Index Returns

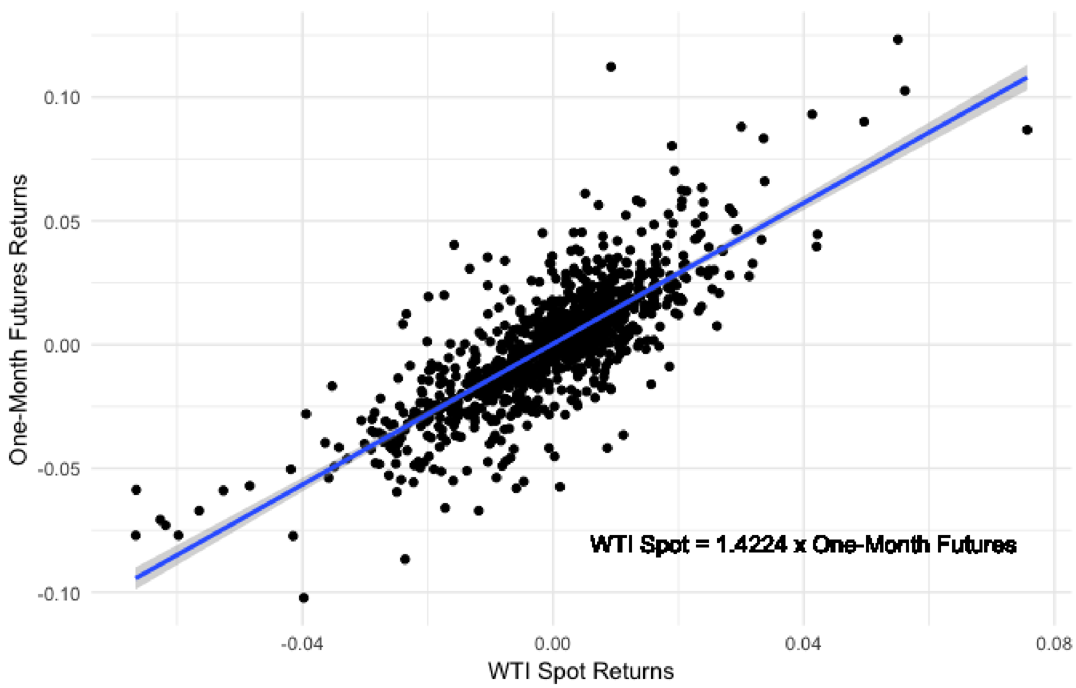

Figure 6: Scatterplot of WTI Spot Price Returns vs WTI 1-Month Future Returns

Comments on the Graphs:

- We can see the relationships are almost perfectly unitary for every pair except oil. In the case of oil, the WTI spot rate return tends to be higher than the one-month future return (by a factor of 1.4).

- The tradable security will always have a later start date than the index as the tradable security is based off the index itself. The index needs to be constructed and balanced before a security can be constructed and begin tracking it.
- **SPX/SP1** track each other well, with slight deviations consistent throughout (mostly no larger than 25 BP). There is a single negative outlier, which is Black Monday (8th of August 2011), a stock-market crash fuelled by the S&P's downgrade of US Debt from AAA to AA+.
- **SPX/SPY** track each other almost perfectly (offer the same returns essentially every day). The coefficient in the line of best fit is 0.9961, which clearly indicates this 1-to-1 relationship. There still is a slight trend for the SPX (spot) to outperform the SPY index on days where return is close to 0 (for the SPY). Black Monday is clearly visible on this plot too.
- **Oil** is much more volatile than the other securities (the y axis, returns for oil futures, spreads from -10% to + 10%). It is noteworthy that there are not any particular outliers in the oil graph, as there are multiple extreme values in each direction.
- **Gold** returns remain relatively in line with gold index returns. There is one specific outlier visible in returns, but the relationship was not broken in this case. There is a day where both gold and the gold index lost approximately 9%. This occurred on the 15th of April 2013, and seems to have been related to slowing growth in the Chinese economy, and a mass exodus from commodities securities worldwide.

3. **Question 3 - Options and Hedging**

(a) The mark-to-market value of the delta-hedged portfolio was calculated using the formulas outlined in the lecture slides. The initialisation at time zero followed the below formulae:

$$\boxed{\begin{aligned} Q_0 &\leftarrow -100(\text{number of contracts}) \\ q_0 &\leftarrow -Q_0\Delta_0 \\ \text{Cash}_0 &\leftarrow -Q_0V_0^{(\text{ask})} - q_0S_0 \\ MV_0 &\leftarrow Q_0V_0 + q_0S_0 + \text{Cash}_0 = Q_0(V_0 - V_0^{(\text{ask})}) \end{aligned}}$$

Q, the number of options held, remained constant throughout the period. The number of shares held, q_t, was calculated using: $Q_t * \Delta_t$.

The cash position was calculated using the following formula:

$$Cash_t = (1 + RF_{t-1}) * dt * Cash_{t-1} - (Shares_t - Shares_{t-1}) * Spot_t$$

The market value was calculated as:

$$MV_t = (Q_t * P_t) + (q_t * Spot_t) + Cash_t + (Shares_t * Div_t)$$

Profit was calculated as:

$$\pi_t = MV_t - MV_{t-1}$$

The final position was closed out by selling all currently held shares (or buying back shares if it was a short position) and buying options to balance out the option position.

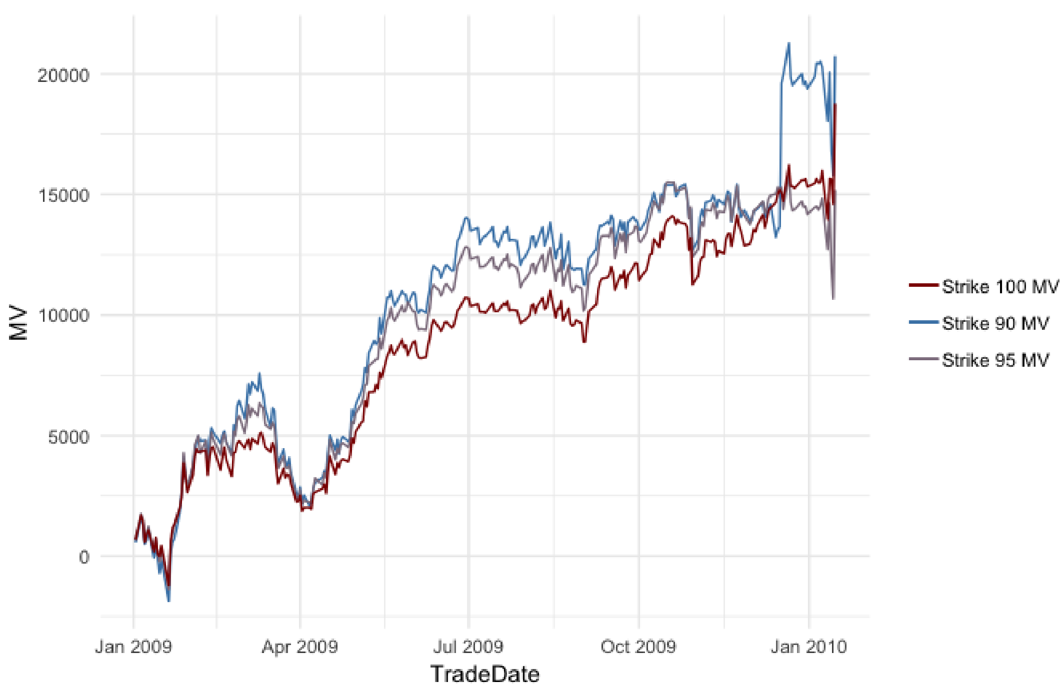

Figure 7: Market Value of K = 90,95,100 Trades Over Time

(b) The expected profit of the trade at the time of opening, as the market-maker, is equivalent to the half-spread on the option multiplied by the number of options sold (N):

$$E[\pi] = \frac{V_{ask} - V_{bid}}{2} * (N)$$

Calculating this for each of the three option strikes:

- K = 90
$$E[\pi_{90}] = \frac{11.15 - 10.9}{2} * 4500 = \$562.5$$
- K = 95
$$E[\pi_{95}] = \frac{8.8 - 8.5}{2} * 4500 = \$675$$
- K = 100
$$E[\pi_{100}] = \frac{6.75 - 6.45}{2} * 4500 = \$675$$

(c) To estimate the standard deviation of the final P&L, we must simulate multiple price paths, and use the same delta hedging trading strategy under those paths. Then, the terminal P&L can be calculated for each path. In perfect market conditions (no arbitrage, all assumptions hold), we expect this standard deviation to be 0, as the profit for the market-maker should be exactly equal the expected profit in all cases. Obviously, we don't expect this to be the case in practice.

In order to attempt to this question, I assumed the option prices at the terminal date were the same. I understand this is a wildly incorrect assumption, but the attempts I made to reprice the option at the terminal date were futile. My hope is that although the incorrect price is used, it is used in each case, and we are ultimately only interested in the terminal standard deviation (which is independent of the mean).

I simulated 1000 paths of the stock price, using a very simple, noise-based formula:

$$S = S_0 + 3 * \mathcal{N}(0, 0.4)$$

Where S_0 is the same starting price as the sample ($90.2) and the 3 and 0.4 were chosen arbitrarily. I then compute the terminal P&L using the same approach as before, and calculate the standard deviation of the 1000 results. This simulation, for each of the three options, gave the following:

$$\sigma_{P\&L,K=90} = \$14504.65$$

$$\sigma_{P\&L,K=95} = \$6027.994$$

$$\sigma_{P\&L,K=100} = \$9138.77$$

In each case, the standard deviation is significantly different from 0.

(d) The final realised P&L for each of the three trades are as follows:

Strike Price	**Final P&L**
K = 90	$20,747.02
K = 95	$15,192.55
K = 100	$18,775.71

(e/f) The annualised standard deviations and annualised Sharpe ratios for each of the trades are shown in the table below. They were calculated using the following formulas:

$$\sigma_{annual} = \sigma_{daily} * \sqrt{252}$$

$$SR_{annual} = \frac{\mu_{daily} * 252}{\sigma_{annual}}$$

Strike Price	σ	Sharpe Ratio
K = 90	$12,750.72	1.565
K = 95	$10,356.61	1.411
K = 100	$8,942.92	2.019

(f) The option deltas for the three strikes are shown below. There are two versions of the plot; one with zero values removed (for continuity), and the raw version.

Figure 8: Delta of K = 90,95,100 Over Time, Zero Values Removed

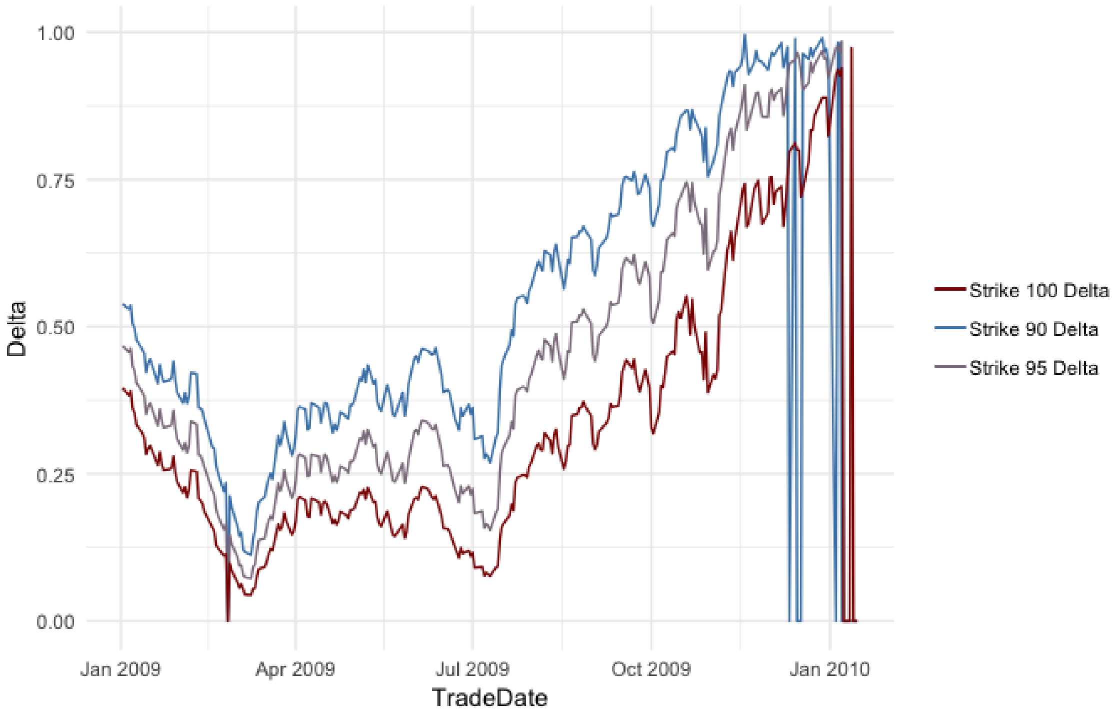

Figure 9: Delta of K = 90,95,100 Trades Over Time, Raw

(h) The realised 1-month volatility was calculated using a rolling regression of width 21 days (a trade month). As a result, there are no values until the 22nd day of the time series. The plot below shows the implied volatility for each of the three strikes and this realised 1-month volatility. As before, there is a version with zeros removed and the raw version.

Figure 10: IV of K = 90,95,100 Trades Over Time, Zeros Removed

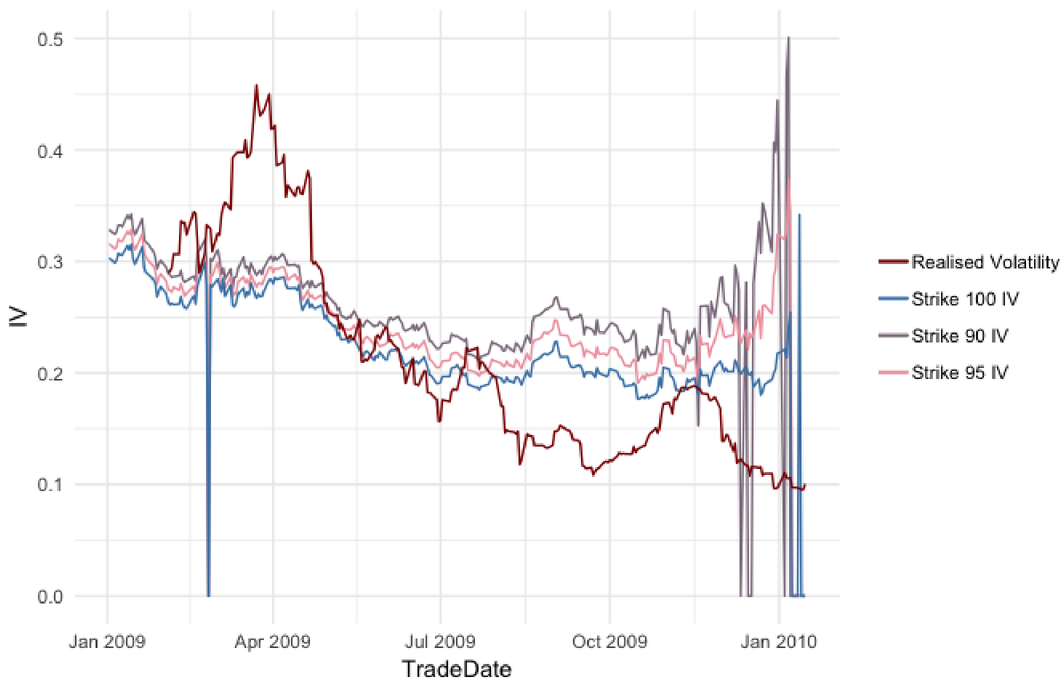

Figure 11: IV of K = 90,95,100 Trades Over Time, Raw

(i) The scatter-plots of daily P&L vs DIA return are show below:

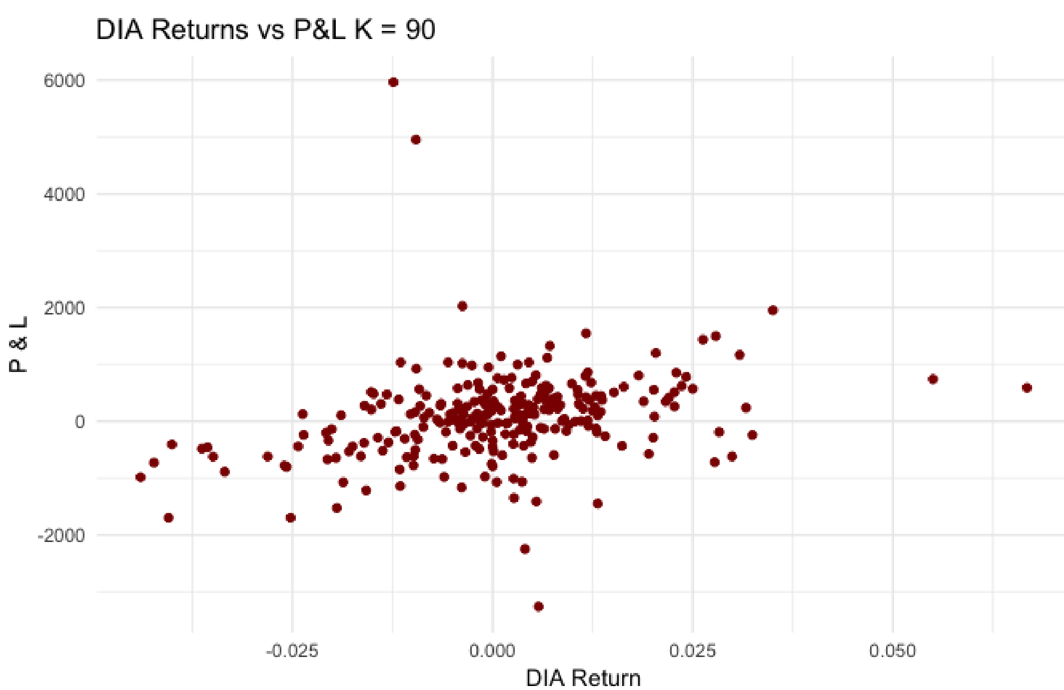

Figure 12: Scatter-plot of K = 90 P&L vs DIA Returns

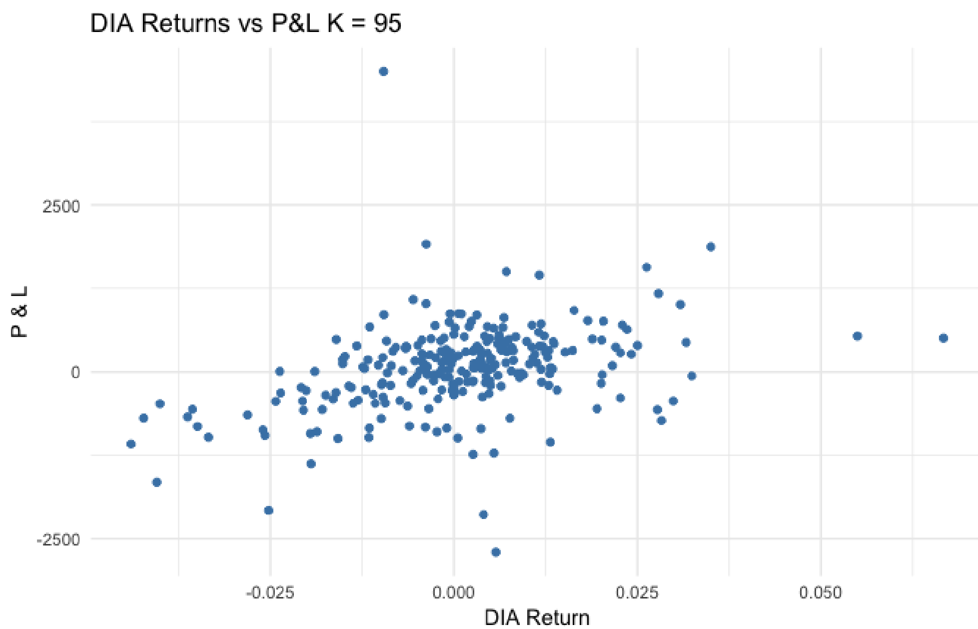

Figure 13: Scatter-plot of K = 95 P&L vs DIA Returns

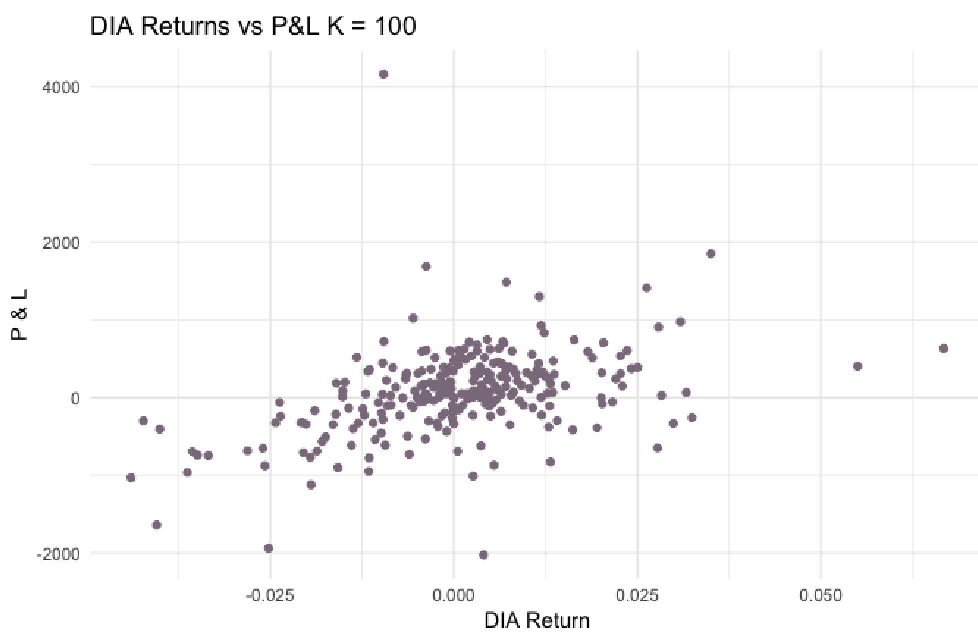

Figure 14: Scatter-plot of K = 100 P&L vs DIA Returns

In all three cases, the slight upward trend in the points indicates that none of the strategies are truly delta-neutral. If they are delta-neutral, the P&L of the trade should be independent of the underlying stock price returns (for small movements, at least). The upward trend is very gradual though, which indicates the strategies are at least close to being completely delta-neutral, which is a good result in practice (as opposed to in theory).

Question 4 - Butterfly

(a) The butterfly trade was tackled in the same way as the initial trades, but rather than incorporating a single security, we have to aggregate the positions in all three options (to calculate the shares needed). Furthermore, the market value is comprised of all three options:

$$MV_t = (Q_{90,t} * P_{90,t}) + (Q_{95,t} * P_{95,t}) + (Q_{100,t} * P_{100,t}) + (q_t * Spot_t) + Cash_t + (Shares_t * Div_t)$$

The market value of the butterfly trade over time is shown below:

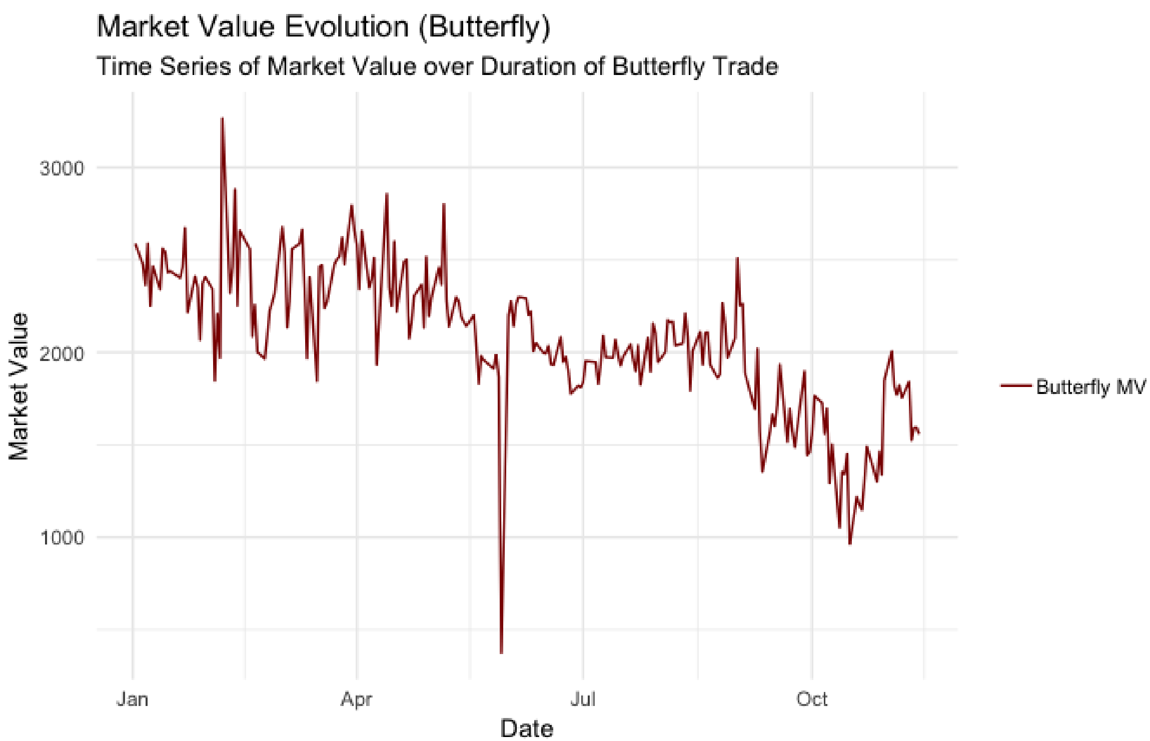

Figure 15: Butterfly Market Value over Duration of Trade

(b) The expected profit and loss can be calculated similarly to in Question 3. The formula is as follows:

$$\begin{aligned}
E[\pi] &= \sum_{n=1}^{3} \frac{V_{i,ask} - V_{i,bid}}{2} * (N_i) \\
&= (\frac{11.15 - 10.9}{2} * 4500) + (\frac{8.8 - 8.5}{2} * 9000) + (\frac{6.75 - 6.45}{2} * 4500) \\
&= 562.5 + 1350 + 675 \\
&= \$2,587.5
\end{aligned}$$

(c) This would be approached similar to the part in Question 3. We would need to simulate stock price paths independently, compute the option value at the terminal price (the price used to close out our position), and compute the terminal P&L for each of these paths using these option prices. The σ could then be calculated over the sample.

(d) The final realised P&L for the trade is equivalent to the cash value after closing out all positions, which is equal to the sum of the daily P&L over the duration of the trade.

$$\pi_{butterfly} = \$2,569.64$$

(e/f) The table below summarises the performance of the butterfly trade:

	σ	**Sharpe Ratio**
Butterfly	\$5,500.03	0.324

(g) The aggregate delta for the butterfly trade is calculated as follows:

$$\Delta_{butterfly} = 2 * \Delta_{95} - \Delta_{90} - \Delta_{100}$$

This gives the following delta over the duration of the trade:

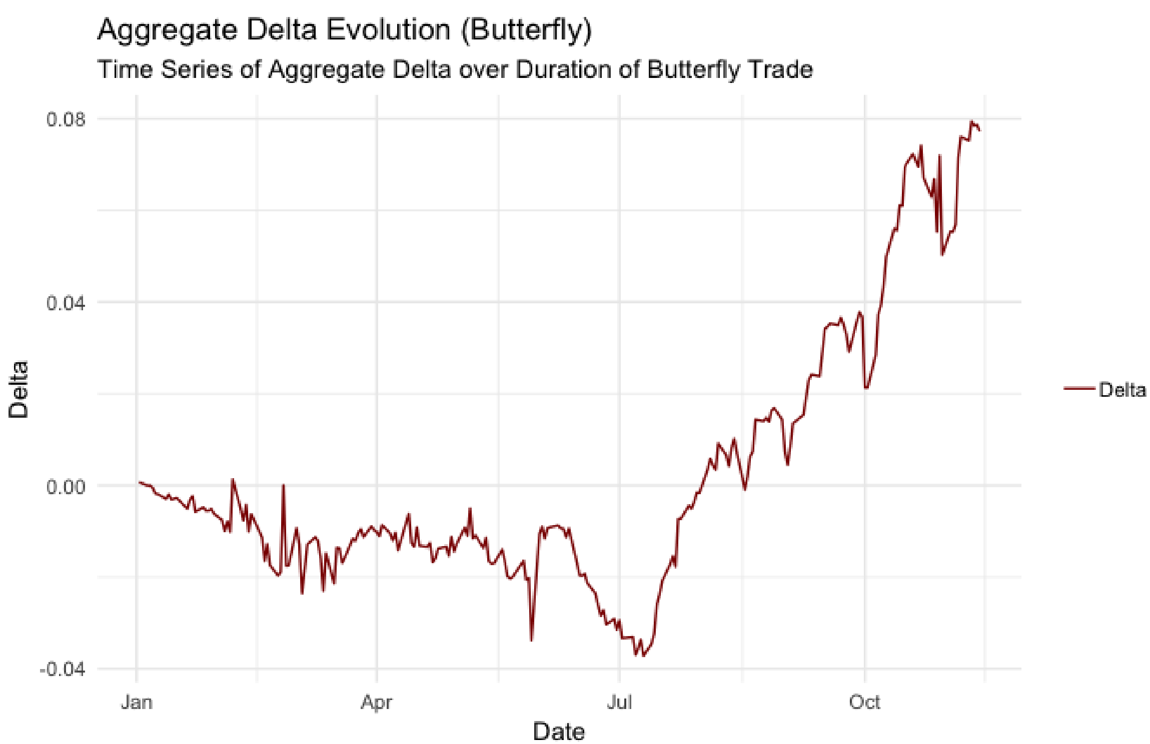

Figure 16: Butterfly Aggregate Delta over Duration of Trade

(h) The aggregate vega is calculated similarly to the delta:

$$\nu_{butterfly} = 2 * \nu_{95} - \nu_{90} - \nu_{100}$$

This gives the following vega over the duration of the trade:

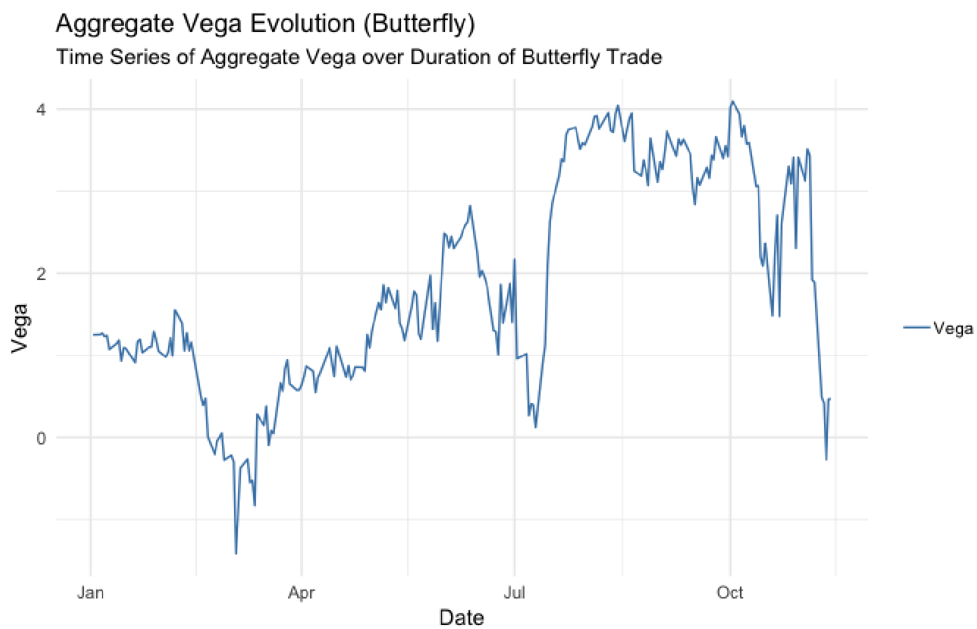

Figure 17: Butterfly Aggregate Vega over Duration of Trade

(i) The scatter-plot of daily Butterfly P&L vs the DIA return is shown below:

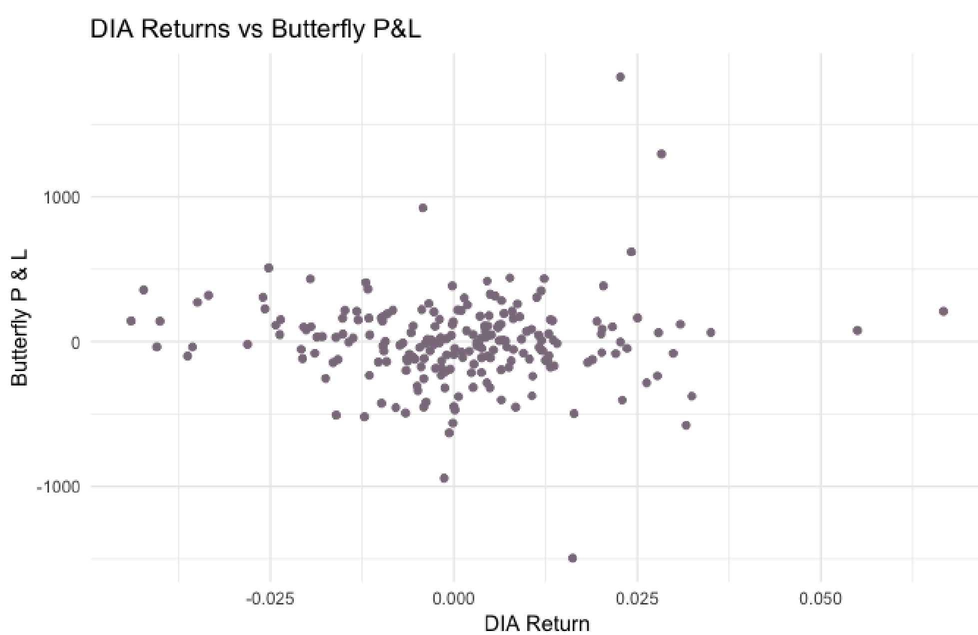

Figure 18: Daily Butterfly P&L vs DIA Returns

In this case, the trend line seems to be almost horizontal, right around zero. This indicates that the butterfly trade delta-hedge is more efficient, and much closer to delta-neutral. However, there still are some outliers (profitable days for the trade where the DIA returned about 2.5%), which indicates it is not yet perfect.