Individual Coursework Submission Form

Specialist Masters Programme

Surname: Parashar	**First Name: Tridev**
MSc in: Business Analytics	**Student ID number: 180032806**
Module Code: SMM638	
Module Title: Network Analytics	
Lecturer: Dr. Simone Santoni	**Submission Date: 14 Dec 2022**

Declaration:
By submitting this work, I declare that this work is entirely my own except those parts duly identified and referenced in my submission. It complies with any specified word limits and the requirements and regulations detailed in the coursework instructions and any other relevant programme and module documentation. In submitting this work, I acknowledge that I have read and understood the regulations and code regarding academic misconduct, including that relating to plagiarism, as specified in the Programme Handbook. I also acknowledge that this work will be subject to a variety of checks for academic misconduct.

We acknowledge that work submitted late without a granted extension will be subject to penalties, as outlined in the Programme Handbook. Penalties will be applied for a maximum of five days lateness, after which a mark of zero will be awarded.

Marker's Comments (if not being marked on-line):

Deduction for Late Submission: | | **Final Mark:** | %

SMM638 — Network Analytics

Final course project submission template

The target business analytics problem of your choice (500 words)

As a quant analyst at one of the renowned hedge funds in the UK, I have been asked to assess the factors impacting UK focused investment strategies during times of global financial crisis in order to mitigate losses faced by the hedge fund during such events in the past.

With reference to my previous experiences as a quant analyst, the primary factors usually responsible for spread of financial crisis phenomenon are similar investor behaviour, interconnectedness of markets etc. (also stated by *Hansen (2021)* in his research on problems of connectivity in financial markets). Therefore, it would be useful to analyse the global financial markets network structure and further assess contagion characteristics during times of financial crisis.

Regarding the same, I propose to move ahead with the analysis using the following approach:

- Use concepts of network contagion (in this case similar investor behaviour in financial markets) to understand the spread of financial crisis in the UK markets.
- For the purpose of understanding financial crisis contagion, I propose to use the US subprime crisis 2007 - 08 as a basis to understand the spread of the crisis from the US (origination of the crisis) to the UK financial markets.
- To deliver an objective outcome, I propose to divide and compare the analysis between three different time periods i.e., prior to the crisis, during the crisis and post the crisis.

The aim of this analysis would be to devise business analytics-oriented UK focused investment strategies that would enable the hedge fund to maximise certainty in investment returns and minimise losses to the farthest extent possible.

The justification for the choice of the problem (300 words)

The existence of financial crisis contagion has been covered by multiple research papers and has been an area of caution for hedge funds during investment decisions. *Hansen (2021)* in his research on problems of connectivity in financial markets clearly covers the contagious strategies adopted by investment firms/hedge funds, keeping in mind the severity of a brewing financial crisis.

With regards to the same, we firstly ought to analyse contagion characteristics during times of financial crisis since majority of reasons for spread of crisis phenomenon listed through my previous experiences as well as noted by *Hansen (2021)* and *Zhu, Yang and Ye (2018)* in their research on financial contagion behaviour are related to similar behaviour (homogeneity) portrayed by investors during such crisis events.

Hansen (2021) also notes that similar investor behaviour is due to the increased interconnectedness of the markets which allows seamless flow of capital, eventually allowing market actors to trade across different markets at the same time.

Therefore, to devise effective UK focused investment strategies, we will need to analyse the contagion characteristics arising from interconnectedness between US and UK markets. Such interconnectedness will help us answer the spread of crisis between these two countries during the US subprime crisis 2007-08.

Further, we propose to divide the analysis into three different time periods to allow us to compare the periods in terms of changes in investor behaviour observed and devise particular investment strategies that might be helpful especially at the onset and during the financial crisis.

The network dataset suited to address the chosen problem (500 words)

For the purpose of analysing financial crisis contagion behaviour of the US subprime crisis 2007-08, we propose to employ the following data points for our analysis –

- Logarithmic daily returns of the S&P 500 index (equity benchmark for the USA wherein the subprime crisis originated) and the FTSE 100 index (equity benchmark for the UK). Log returns will be considered to ensure that the dataset follows a normal distribution.
- Based on the logarithmic returns, we shall observe three correlation/co-movement scenarios between the two indices i.e., positive corr., negative corr. or no change in the movements of both the indices. Therefore, based on the direction of movement of returns, we shall use the following co-movement symbols for daily returns i.e.,

$$\begin{cases} P, \textit{ if there is a positive correlation} \\ N, \textit{ if there is a negative correlation} \\ O, \textit{ if no change} \end{cases}$$

- After defining the co-movement symbols, we need to focus on defining the nodes, which in this case pertains to investor behaviour. The movement in prices/returns from indices can be taken as a representation of investor behaviour, and hence the co-movement symbols can be used as nodes. Since co-movement symbols are limited in number, we can use a 'coarse graining' process to increase the number of nodes. As stated by *Zhu, Yang and Ye (2018),* the process can be carried out by grouping the co-movement symbols on a particular day along with the preceding 4 days (since markets usually operate for 5 days in a week). A snapshot has been attached below for better understanding:

Date	S&P Price	FTSE Price	S&P Log Returns	FTSE Log Returns	Co-movement symbol	Co-movement groupings
11-01-2006	1294.18	5731.5	0.35%	0.75%	P	NPNPP
12-01-2006	1286.06	5735.1	-0.63%	0.06%	N	PNPPN
13-01-2006	1287.61	5711	0.12%	-0.42%	N	NPPNN
17-01-2006	1282.93	5699	-0.36%	-0.21%	P	PPNNP
18-01-2006	1277.93	5663.7	-0.39%	-0.62%	P	PNNPP

For example – the co-movement grouping on 18-01-2006 will be a grouping of the co-movement symbols on that day as well as the 4 preceding days

- After forming the co-movement groups (nodes), we should be able to assign edges and its attributes. In this case, the transition from one group to another forms an edge, and number of times a transition is made between two similar groups can be assigned as weights (attribute) for the edges.
- Some additional things to note –
    i. The log returns will be collected for three different time periods (similar to the ones used by *Zhu, Yang and Ye (2018))* i.e., pre-crisis (Jan 2006 to Mar 2007), during crisis (Apr 2007 to Mar 2009) and post crisis (Apr 2009 to May 2013).
    ii. The network structure formed between the nodes will be a directed weighted network since we are trying to assess whether the FTSE 100 returns are affected by S&P 500 returns. In other words, we will assess the spread of the crisis from the USA to the UK.

Main steps of the analysis (300 words)

After gathering the variables discussed above, we will firstly draw the networks for each time period using NetworkX (can be found in the Python Notebook). As discussed, the networks will be weighted directed graphs for each of the three periods wherein –

- The co-movement groups are the nodes.
- The transition between the co-movement groups are the edges.
- The number of transitions between two similar co-movement groups will be the weights of the edges.

Thereafter, our analysis shall consist of the calculation of the following network measures:

- Firstly, we will check the Pearson correlation coefficient measures between returns of the S&P 500 and FTSE 100 in order to confirm whether UK index returns are affected by the US index returns.

- We will further take a dive in to the outdegree measures of the nodes along with accounting for weights of the edges to check the strength of nodes and its distribution in the network.

- Next, we will check the grouping characteristics of the nodes through clustering coefficient measures to check the possibility of grouping of the nodes.

- We will further check the transformational pattern of the nodes through betweenness centrality measure to check whether nodes occupy any central hub positions in the network.

- We will also check whether any of the nodes occupy positions with shorter distances to other nodes through closeness centrality measure.

- Some more things to note –
    i. All of the above steps will be performed on the three different time periods, as well as on all the time periods combined.

The justification for each step (600 words)

- **Pearson Correlation Coefficient** –

This will be calculated to check whether both markets have any correlation in terms of index return/price movements. Our preliminary findings indicate the following:

	S&P Price	FTSE Price
S&P Price	1.000000	0.950422
FTSE Price	0.950422	1.000000

The Pearson correlation coefficient indicate that both the indices have 95% correlation, which means that any event in the US markets will have a significant effect in the UK markets.

- **Degree and its Distribution** –

The degree of a node represents the number of connections it has with other nodes. We have considered the out-degree measure here, along with weights of the edges to assess its connectivity with other nodes. The higher the out-degree connectivity, the greater the strength of the node and its distribution, and eventual strength to transform to other nodes. Our preliminary findings indicate the following (top 5 nodes with highest degrees):

1. All Data

Nodes	Degree	Distribution of Degree
PPPPP	272	0.03125
PPNPP	145	0.03125
PNPPP	140	0.03125
PPPNP	138	0.03125
NPPPP	134	0.06250

2. Pre Crisis

Nodes	Degree	Distribution of Degree
PPPPP	34	0.03125
PNPPP	19	0.03125
PPPNP	18	0.03125
NPPPP	17	0.09375
PPPPN	17	0.09375

3. During Crisis

Nodes	Degree	Distribution of Degree
PPPPP	65	0.03125
PPNPP	44	0.03125
NPPPP	39	0.06250
PPPPN	39	0.06250
PPPNP	38	0.06250

4. Post Crisis

Nodes	Degree	Distribution of Degree
PPPPP	173	0.032258
PPNPP	84	0.032258
PNPPP	83	0.032258
PPPNP	82	0.032258
PPPPN	78	0.032258

As observed from our preliminary findings, the node *PPPPP* has the highest out-degree connectivity across all the time periods. For the following node, the strength was greatest during the *post crisis* phase followed by *during crisis* and *pre crisis* time periods. Also, the symbol '*P*' is dominant across all the three periods.

A comparison of *pre crisis* and *during crisis* periods, the node degree/strength enhances significantly in the *during crisis* period, indicating significant financial crisis contagion behaviour during the US subprime crisis 2007-08.

	All Data	Pre Crisis	During Crisis	Post Crisis
Node(s) with highest clustering coefficient	PPPPP	PPPPP	PPPPP	PPPPP
Values of highest clustering coefficients	0.202	0.209	0.271	0.175
Nodes with '0' clustering coefficients	20	23	23	22
Node(s) with maximum betweenness centrality	NNNNP PNNNN NNNNN	PPNNN	PPNNN	PNNNN NNNNP
Values of maximum betweenness centrality	0.627 0.627 0.627	0.382	0.315	0.664 0.664
Node(s) with minimum closeness centrality	PPPPP	PPPPP	PPPPP	PPPPP
Values of minimum closeness centrality	0.0059	0.032	0.0183	0.0116

Fig – Some preliminary findings on network measures across the three different time periods

- **Clustering Coefficient** –

Larger the clustering coefficient of the nodes, the greater is the possibility for grouping of the nodes. In the figure given above, we have included the nodes with largest coefficients across the three periods and their values along with mentioning the number of nodes with '0' clustering coefficient.

As observed in our preliminary findings above, there are significant number of nodes with no clustering coefficient indicating that changes between different nodes happen only in case of a few nodes. The node with the highest clustering coefficient across the three time periods is *PPPPP,* wherein its highest coefficient lies in the *during crisis* period indicating that the co-movement node between UK and USA weakens and changes frequently during this period. But upon observing the nodes with the second largest clustering coefficients in the *during crisis* period i.e., *NPPPP* and *PPPPN* with coefficients of 0.04519, we can conclude that the co-movements strengthen (mainly concentrated in the vicinity of *PPPPP)* during this period and illustrates the effect of the financial crisis contagion.

- **Betweenness Centrality -**

This measure enables us to understand whether nodes act as a medium between two nodes i.e., occupying a centrally important position. As observed from our preliminary findings given above, four nodes with the highest betweenness centrality in the *during crisis* phase are *PPNNN* (0.315), *PNNNP* (0.257), *NNNPP* (0.246) and *NNPNP* (0.212), which all follow a certain pattern i.e., they consist primarily of negative co-movements.

- **Closeness Centrality –**

The importance of a node can also be determined by its distance to other nodes (preferably shortest), indicated by the closeness centrality measure. As observed from our preliminary findings, *PPPPP* is the node with the least closeness centrality across all the three time periods. In comparison to the *pre-crisis* period, the closeness centrality measure in the *during crisis* period is lesser implying that the nodes are relatively less scattered compared to before the financial crisis (also implying that the transformation between nodes happen faster). Also, the four least closeness centrality measures in the *during crisis* phase are *PPPPP* (0.0183), *PPPPN* (0.0256), *PPNPP* (0.0313) and *NPPPP* (0.0338), all of which indicate primarily positive co-movements.

Set of possible actionable business analytics emerging from the project (300 words)

As suggested by our preliminary findings gathered through network measures, we can possibly state that UK is affected by the US subprime crisis. The contagion effect is confirmed by out degree measures and clustering coefficients. Further, betweenness centrality and closeness centrality determine the co-movement direction of the nodes with highest betweenness centrality and least closeness centrality measures during the financial crisis.

With reference to our findings, we can devise the following possible business analytics-oriented recommendations for building UK focused investment strategies –

- **Avoid building correlated strategies with significant exposure to the USA –** Since the nodes with highest betweenness centrality indicate negative co-movements, nodes with least closeness centrality indicate positive co-movements and so does clustering coefficients (node with highest clustering coefficient is *PPPPP*), the investments in asset classes in the UK which have a similar exposure to the USA (either through operational reliance or investments etc.) will have either a significant upside or an unlimited downside risk due to higher correlation. Hence, it is better for the hedge fund to avoid correlated strategies with significant exposure to the USA.

- **Optimum risk management to obtain better mix of investments –** Since degree measures and clustering coefficients confirm significant contagion behaviour during times of crisis, it would be useful to develop risk management controls that shall help us monitor the composition of investments in multiple asset classes and enact changes whenever the probability of a crisis significantly increases.

- **Assessment of investor sentiments regularly –** In continuation to the previous recommendation, investor sentiments shall also be observed regularly to discover patterns indicating the onset of a financial crisis etc.

References

- Data collection source – Bloomberg
- Research papers used as references –

i. Zhu, Y., Yang, F. and Ye, W. (2018) 'Financial contagion behavior analysis based on complex network approach', *Annals of Operations Research*, 268(1/2), pp. 93–111. doi:10.1007/s10479-016-2362-6

ii. Kristian Bondo Hansen (2021) Financial contagion: problems of proximity and connectivity in financial markets, Journal of Cultural Economy, 14:4, 388-402, DOI: 10.1080/17530350.2021.1879211