**Modelling business clusters in regions**

# Introduction /Business Problem:

UK high streets are turning into ghost centres, number of automobile manufacturing business including the big brands such as Honda and Land Jaguar are shutting down their operations and financial capital is losing its shine because of uncertain Brexit where some of the banks have already started exiting London (BBC, 2019). In such a volatile business environment, it becomes increasingly important for government and regional developmental organisations to plan and design more effective and sustainable business models to boost the economy and one of the alternatives could be the creation of innovative business clusters or science and technology parks - STPs in right locations.

The process of creation of successful business clusters has always drawn a great attention in the management literature, since the seminal work of Porter (1990) on ‘industrial cluster theory.’ The development of successful business clusters has been a major driver in efforts around the world to improve the economic growth and competitiveness of locals, regions, and nations. It is keenly understood that the right combination of factors and policies can guide in unleashing the tremendous entrepreneurship benefiting individuals and societies. (Engel, 2015)

Business clusters are defined as a ‘socioeconomic entity characterised by a social community of people and a population of economic agents localised in close proximity in a specific geographic region’ (Morosini, P 2004). The companies within the clusters thrive by competing with each other while able to network and knowledge share with each other. It is “accepted wisdom” that successful examples of science parks etc. promote regional development and thus many governments and regional authorities have tried to construct imitations, but actually only a small proportion are successful (BBC, 2010 for a UK perspective and Wadhwa, 2013, for an international perspective).

The increasingly transglobal and knowledge based nature of the world business is motivating the organisations to find out innovative ways of collaborating with each other to gain a competitive advantage in their operations and strategic business objectives. The phenomenon of clusters, deﬁned by Porter (2000; p 15) as ‘geographic concentrations of interconnected companies’, is an approach which really explains how ﬁrms gain because of proximity to each other. Some of the well known and most frequently studied clusters are - Central London’s financial cluster (Keeble and Nachum, 2002), the biotech cluster in Malmö–Copenhagen region (Moodysson et al, 2008), and the Chilean wine cluster (Giuliani and Bell, 2005). Klofsten, et al. (2015)

The number of studies have shown an evidence of the additional benefits of clusters to the society through their value enhancement which includes learning and knowledge repository; advantages of economies of scale; synergies through collaboration, networking and social relations, more information flow and construction of infrastructure (Porter, 1998; Johannisson and Lindholm Dahlstrand, 2009; Sölvell, 2009; Smith et al, 2013). Klofsten, et al. (2015)

The cluster presence within related businesses in a location will foster entrepreneurship by reducing the start-up costs of a business, enabling better access to a more diverse range of inputs and complementary products and enhancing opportunities for innovations. (Saxenian, 1994; Porter, 1998a; Feldman, Francis, and Bercovitz, 2005; Glaeser and Kerr 2009). The co-location of businesses, suppliers, customers and other institutions increases the perception of innovation opportunities while amplifying the pressure to innovate. (Porter, 2000). (Mercedes et.al. 2010)

The firms generally tend to locate in regions that help to improve their economic activity - the reason may be to access the resources or to find ways to enter new markets. Presence of positive externalities determine the onset of clustering process and the competition between the rival firms is the main driver that influences the further development of cluster which drives firms to be more creative, innovative and to keep pace with advancement in technology. Hence, stimulating the research and development activity where employees becoming more specialised by gaining higher professional qualification and presence of higher value added services sector. The freedom of movement of employees from one firm to another within the cluster and proximity allow easy and fast transfer of knowledge to new firms enabling increased competitiveness and growth and therefore leading to onset of new activities which will, in turn, result in the emergence of new firms both upstream and downstream. In brief, cluster becoming part of an organic development process over time. (Porter 1990)

I will investigate business clusters in a landscape. Questions include how many potential clients (and at what input and churn rate amongst inhabitants) does a cluster need? Methods for identifying the number of client companies will initially be taken from company house.

Unfortunately many of the clusters fail because of number of reason sometimes probably because the cluster or science park was started in a wrong location? Study will aim to find out a right location for the next UKSPA to start up using regression analysis and Foursquare application.

Literature Review:

Though the emergence of SPs goes far back in early 1950s with the establishment of Stanford Research Park, it was not until 1980s the topic of SP concept became popular and truly flourished becoming one of the most appealing initiative for regional development (Link and Link 2003; Anttiroiko 2004 in (Lecluyse et.al., 2019). Interestingly, early success story of some of some of the science parks such as Research Triangle motivated all countries over the world to establish SPs (Castells and Hall 1994 in (Lecluyse et.al., 2019). SP activity worldwide has almost doubled over last 15 years Recent statistics report has revealed that there are approximate 400 SPs in Europe, employing more than 750,000 staff. (Rowe 2014). As a result, the discussion of this worldwide SP phenomenon has become main focus academic debate.

Monck et al. (1988), in their seminal work, aimed to find out whether SPs contribute to their tenants and economy. This soon was followed by number of studies trying to understand the SPs contribution to numerous stakeholders such as private sector, local and national economy.

The main purpose of providing this literature review is to take an overview of the current state of science parks contribution to economy and society and to find out ways to enhance the knowledge in this area to improve the overall contribution of the SPs. The importance of the contribution of science park and understanding of this phenomenon is particularly more relevant in the current times than ever before and matter of vibrant debate for academician. (Lecluyse et.al., 2019)

The literature contribution of SP is very theoretical in nature (Leyden et al. 2008; Ratinho and Henriques 2010) and henceforth there are minimal theoretical contributions based on these studies. There had been some efforts to relate the SP contribution to theory but most of the studies draw prior empirical findings to build hypotheses and mainly focus on providing the empirical implications.

The justification of the establishment of SPs is often referred to the theory of agglomeration economies (Shearmur and Doloreux 2000; Koçak and Can 2014; Ramirez et al. 2013), cluster theory (Phillips and Yeung 2003; Hu et al. 2005; Link and Scott 2007), regional development theory (Goldstein and Luger 1990, 1992), structural theory (Chan and Lau 2005) and transaction cost theory (Cabral 1998). (Lecluyse et.al., 2019)

A SP or cluster is the critical mass of firms in a specific field in a particular location, whether it may be a country, a state, a region or a city. Clusters can take different forms depending upon their sophistication and depth, but most of the clusters include group of companies, suppliers, machinery, components and firm related industries. Clusters often include firms in downstream industries such e.g. channels and customers, providers of specialist infrastructure, complementary products producers, information, education and other specialised training institutions, technical support and research such as universities, think tanks, standard setting agencies and vocational training providers. (Porter, 1998)

It is a paradoxical to mention the importance of economic geography in an era of ultracompetitive business environment which boasts to be dependent heavily on state of art high technology communication, rapid transportation and easily accessible markets – it is obvious for someone to expect location to diminish its importance. However, the opposite is true. The competitive advantage of today’s global economy is often very local, arising from conglomeration of specialised knowledge and skills, competitors, enterprises, related business, and highly sophisticated customers. Geographical, institutional and cultural proximity leads to special access, better information, strong closer relationship, powerful incentives and other related advantages in innovation and productivity that are really difficult to tap from a distance. The more this is true as the world economy becomes more knowledge based, complex and dynamic. (Porter 1998: 90) (Piperopoulos, Panos G. 2012)

Providing resources such as material, human capital and financial support creates an environment that will attract new members strengthening the competitive position of companies within the cluster which will help to diversify the cluster, bring synergies and create new opportunities for collaboration. As a result, the transformation of new resources over time will lead to creating more value, increase in investment, more wages & salary and prosperity in the cluster. Cojocaru and Ionescu (2016)

Day et al (2000) states that co-operation and trust amongst firms in the industrial cluster are central to its existence. The firms must share their information and knowledge in order to live in this innovative and creative environment. And as a result this mix of trust and co-operation helps on creating an atmosphere where competitors also becomes collaborators allowing specialist firms to offer its expertise to create products that benefit from a pool of expert knowledge. (Piperopoulos, Panos G. 2012)

# Aims and Objectives:

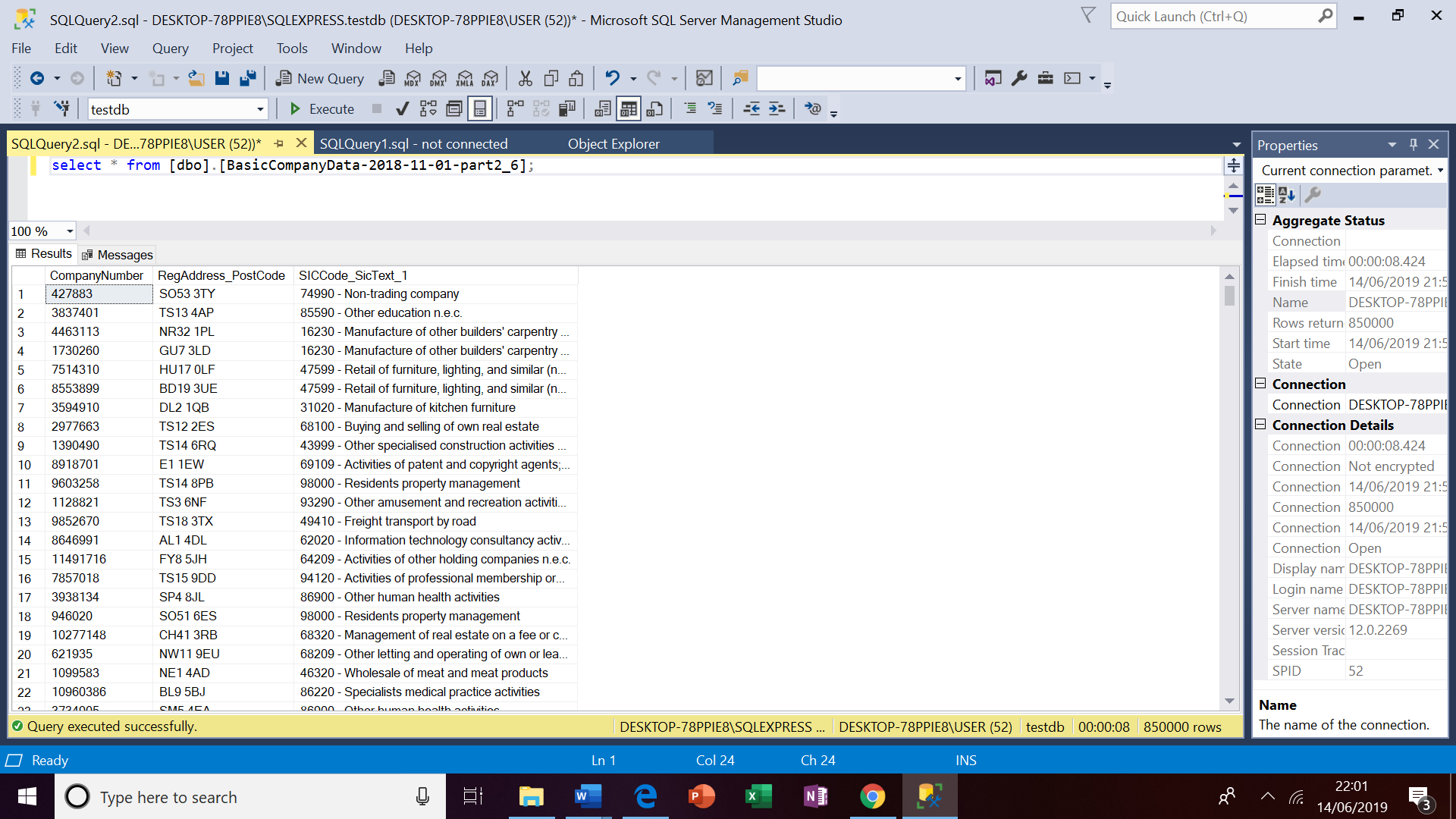
The research aims to find out what is the optimal size of science and technology park for it to grow and flourish successfully, and to find out the right location for the success of next STP. Research will focus to identify the right combination of firms in a cluster in terms of their specialisation using Standard Industrial Classification (SIC) codes. A balanced cluster structure is not only about sharing facilities, knowledge and information but also about rejecting bad projects while implementing good ones. A successful business cluster contains a combination of value creation processes such as specialised facilities, technology, knowledge spill over and cluster’s/firm’s capability to handle the risks. Though there are numerous factors responsible for the success of a cluster, the main objective of the research is to make a better understanding about right size of the cluster and types of companies that need to coexist in the cluster for its success.

# Data collection and analysis:

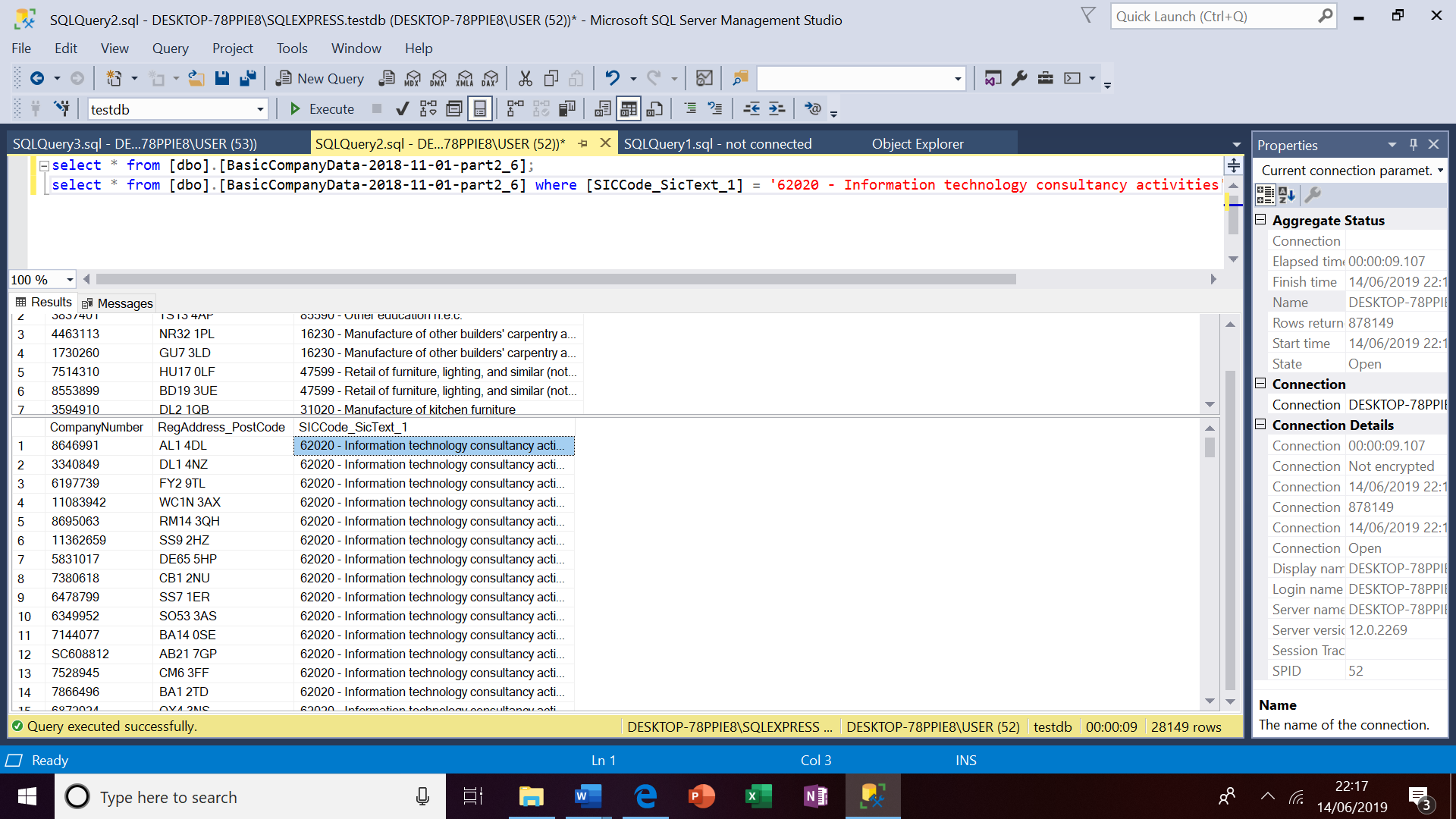
Data is available from the open Gov.UK company House, UK Office of National Statistics data of companies and data from UK Science Park Association (UKSPA), the International Association of Science Parks (IASP) and Catapult technology centres (Innovate UK). Data will be extracted cleaned then both by serial iteration through (outliers separated for individual analysis).

At a micro-level, Geographical Information Systems will be used with real data obtained from the UK Office of National Statistics and data.gov.uk to create GIS mashups of the actual (post code) locations of ‘prey’ firms onto maps and then superimpose real "predators", using data available pertaining to mature clusters (obtained from UKSPA, IASP etc.) as well as younger nascent initiatives (from innovate UK).

First of all the raw company data for preys was downloaded from open source – gov.uk/ company house in the excel format and then it was loaded on the MSSQL server where the simple sql query was executed to fetch total number of 850,000 rows with 3 fields named as Company Number, Post Code and SIC Code as shown below:



In the next step, companies with SIC code '62020 - Information technology consultancy activities' were identified making another sql query which fetched 28149 records as shown below:

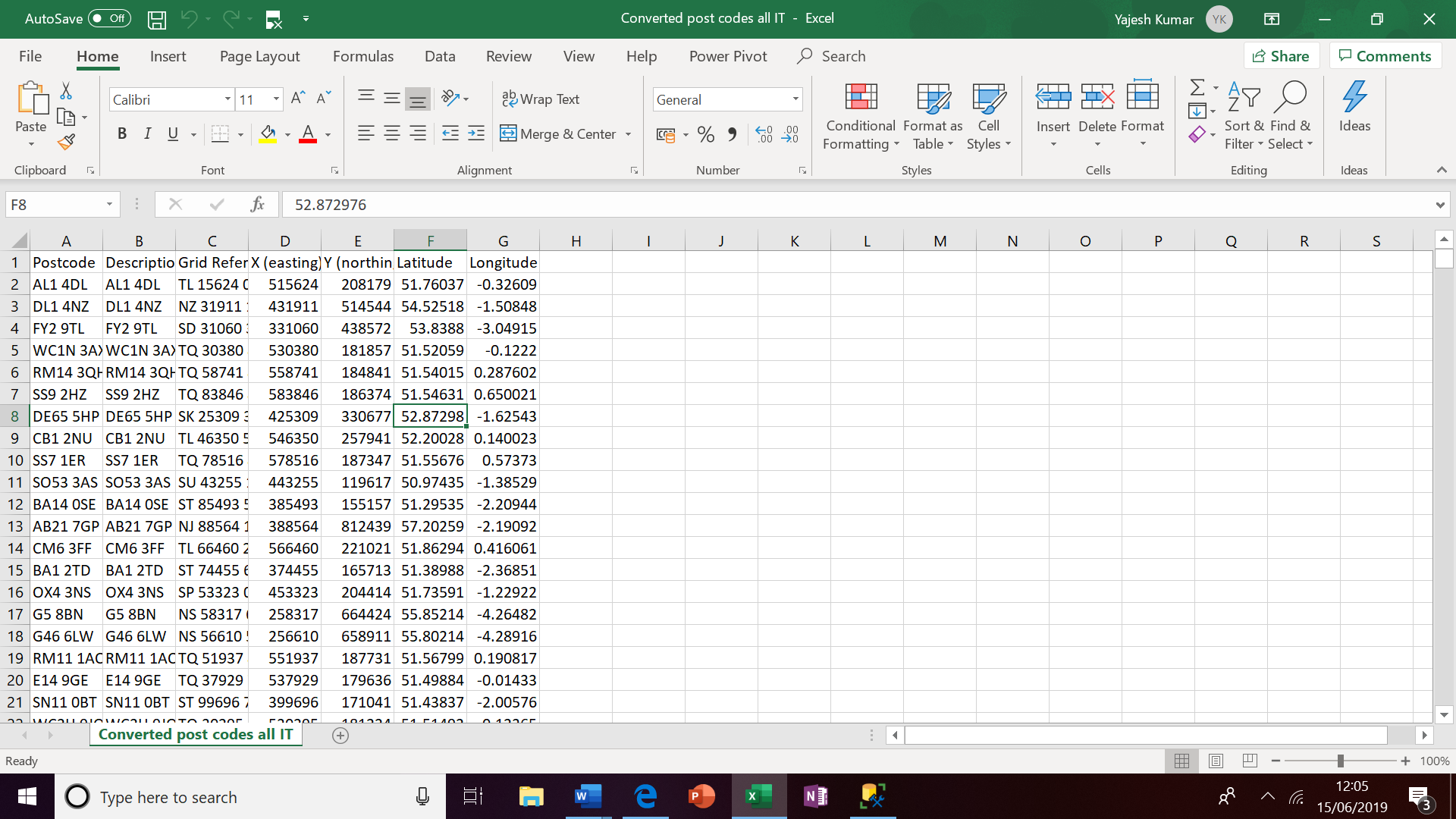


In the next step, these companies data with SIC code '62020 - Information technology consultancy activities' will be cleaned, any NaN values rows deleted and made ready to be mashed up using Geographical Information System – GIS’s application *ESRI’s Arc GIS.*

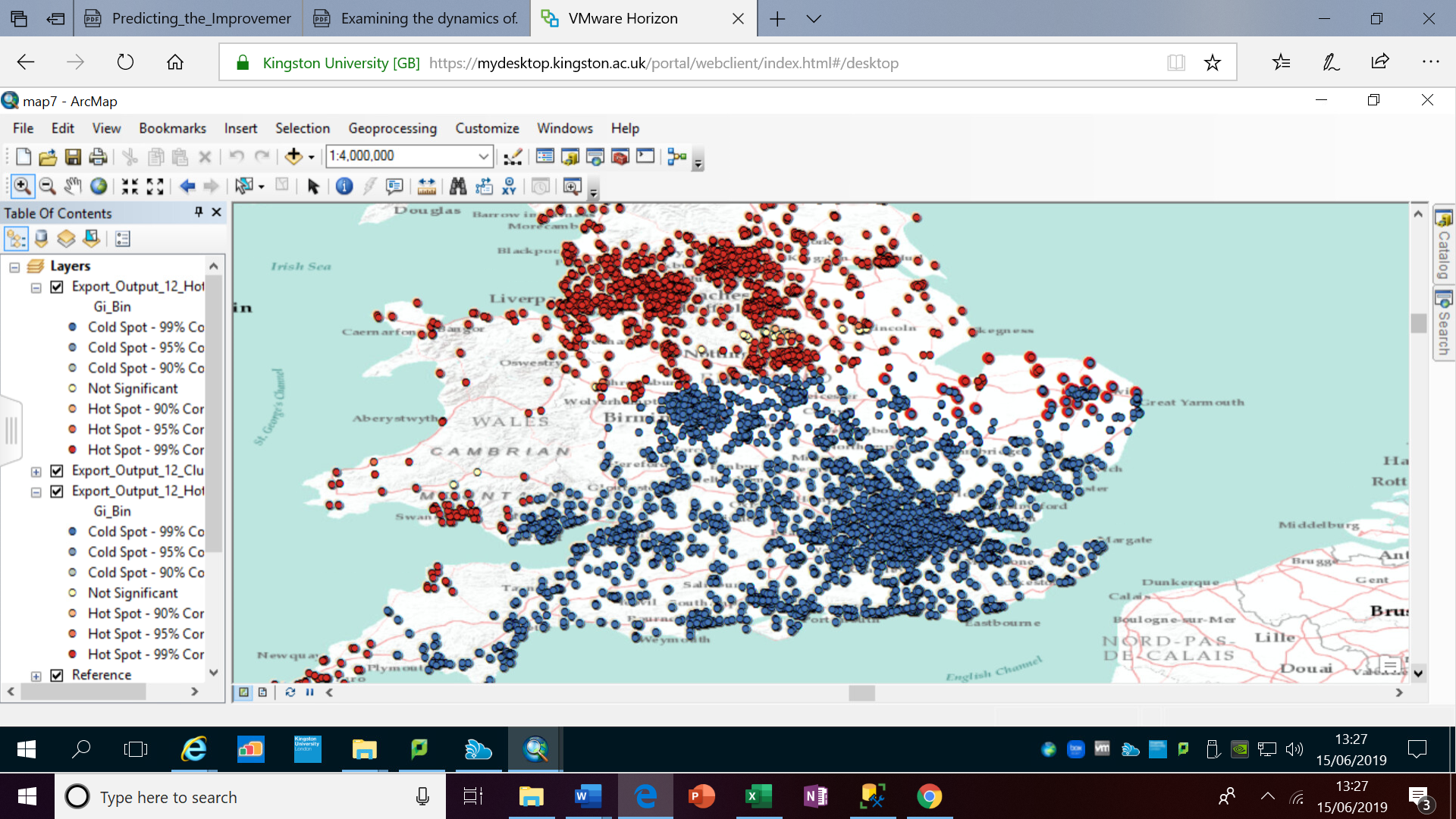
Methodology:

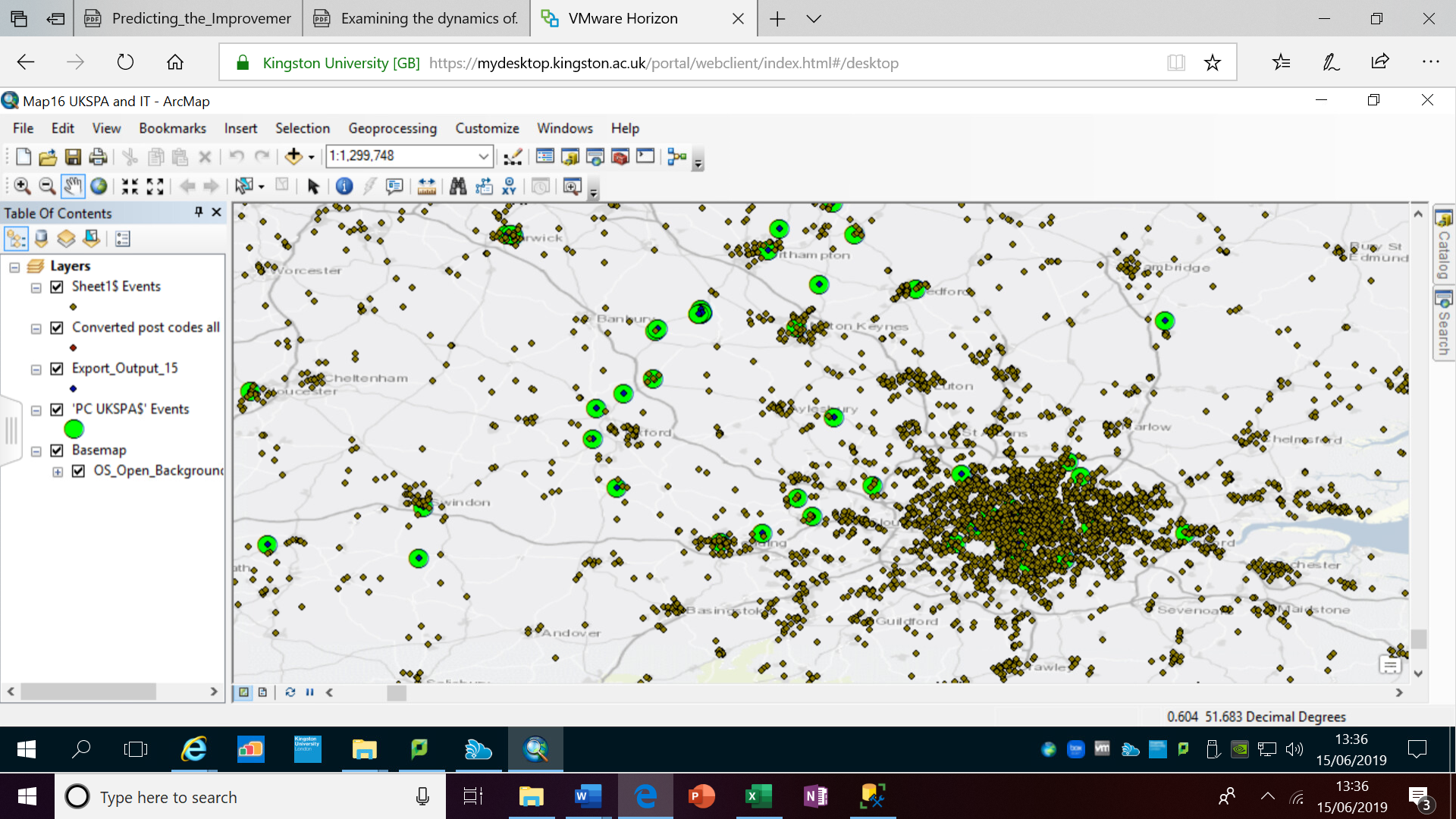
Now data for prey and predator firms will be mashed up will be mashed up using – GIS’s application *ESRI’s Arc GIS.*

Before mashing up the data on map using *ESRI’s Arc GIS*, data of companies with SIC code '62020 - Information technology consultancy activities' will be converted in the longitude and latitude format as shown below:



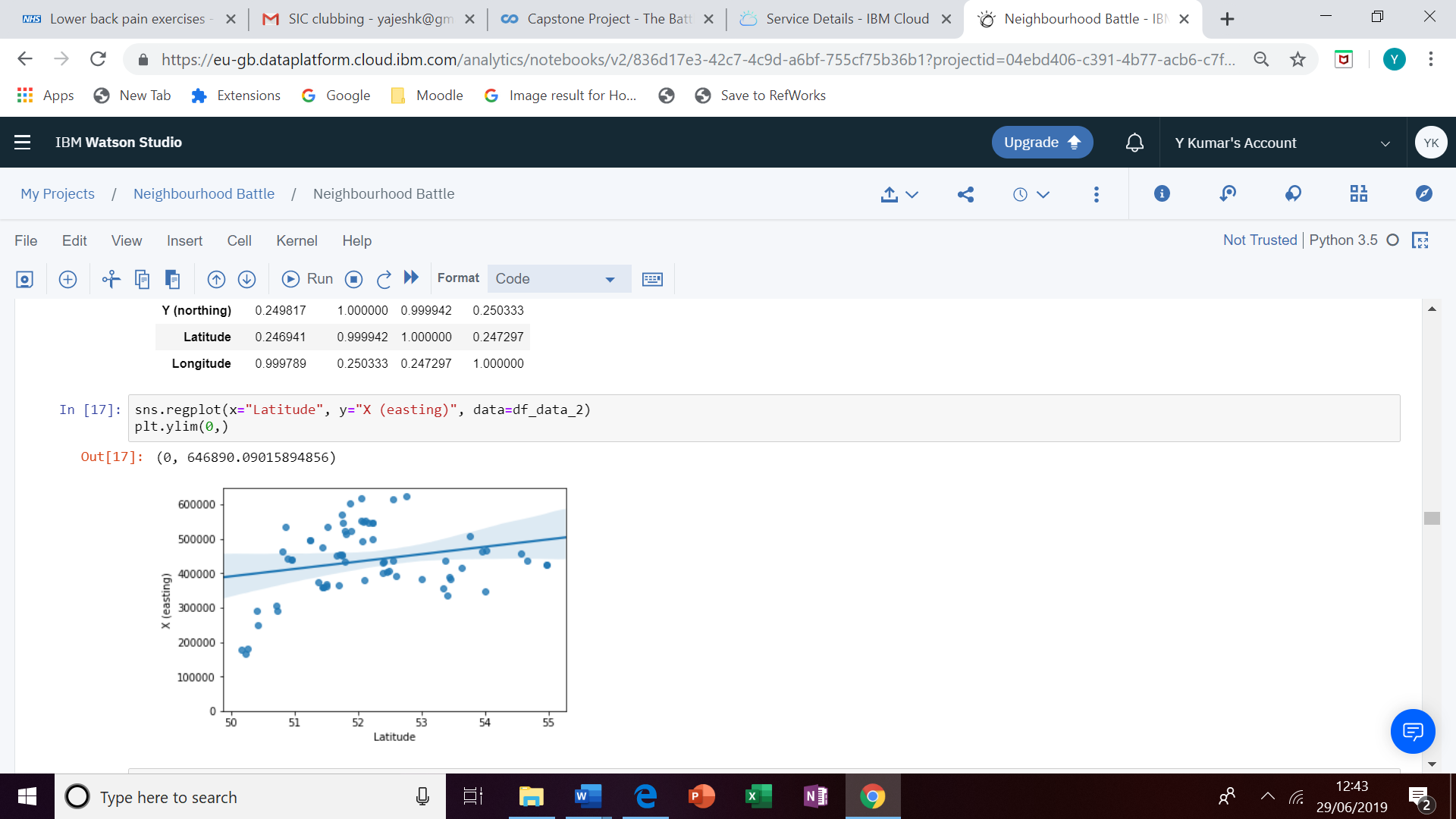
Now this data is ready to be mashed up using using *ESRI’s Arc GIS* application. On mashing up the data the map with related statistical information is presented below:





Converted post codes all IT and UKSPA mashed up together.

Also **regression analysis** will be performed using Jupiter notebook python to find out the best suitable location for the next UKSPA and it would be further supported by using Foursquare application.



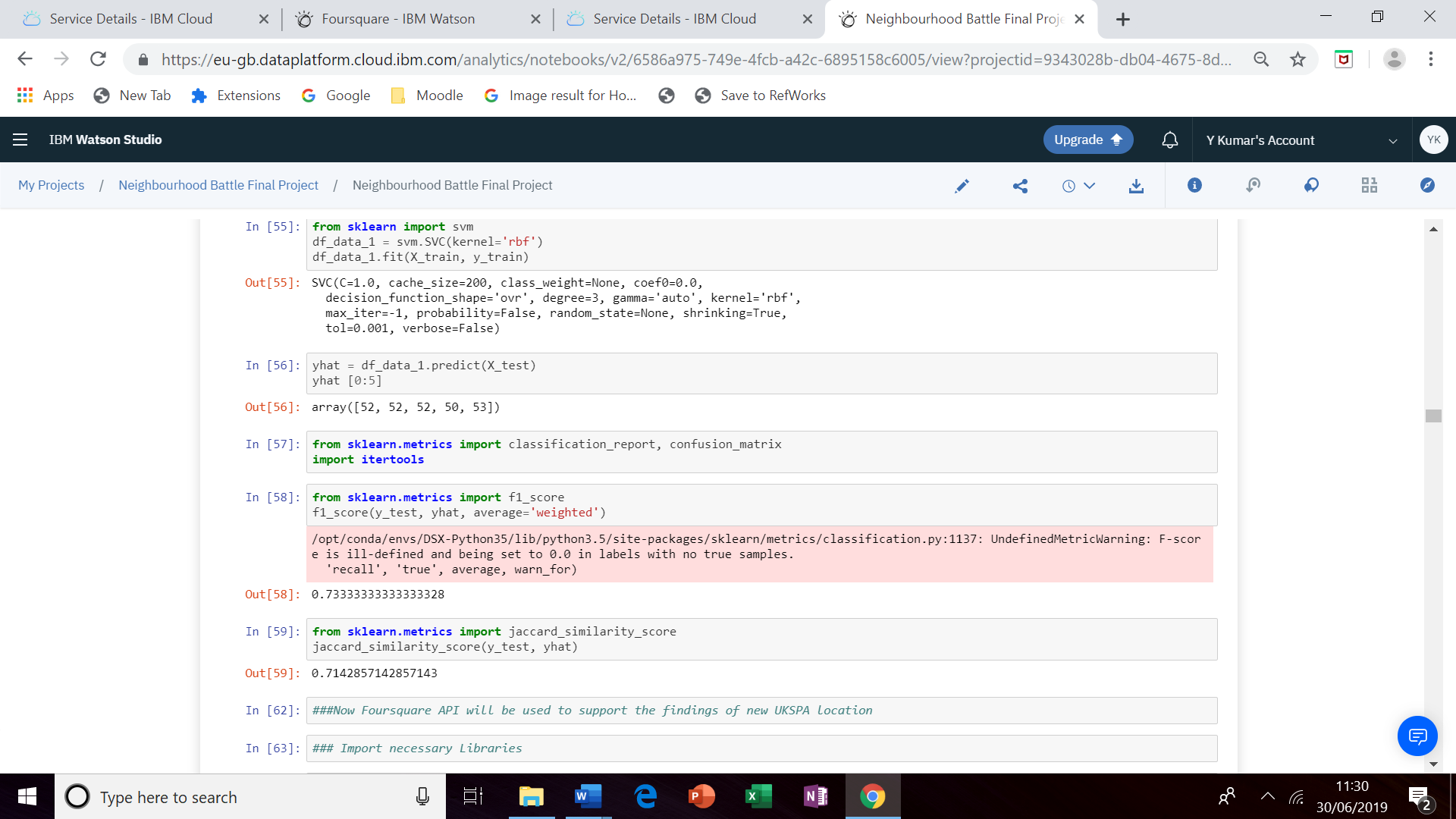
Discussion and Result:

This work clearly takes existing theories into account but it largely consists of experiments that are data-driven without preconceptions. This is in order to uncover fresh principles and construct a new and comprehensive theoretical framework(s) at macro, cluster and micro levels that will enable concrete recommendations to be made for a range of situations.

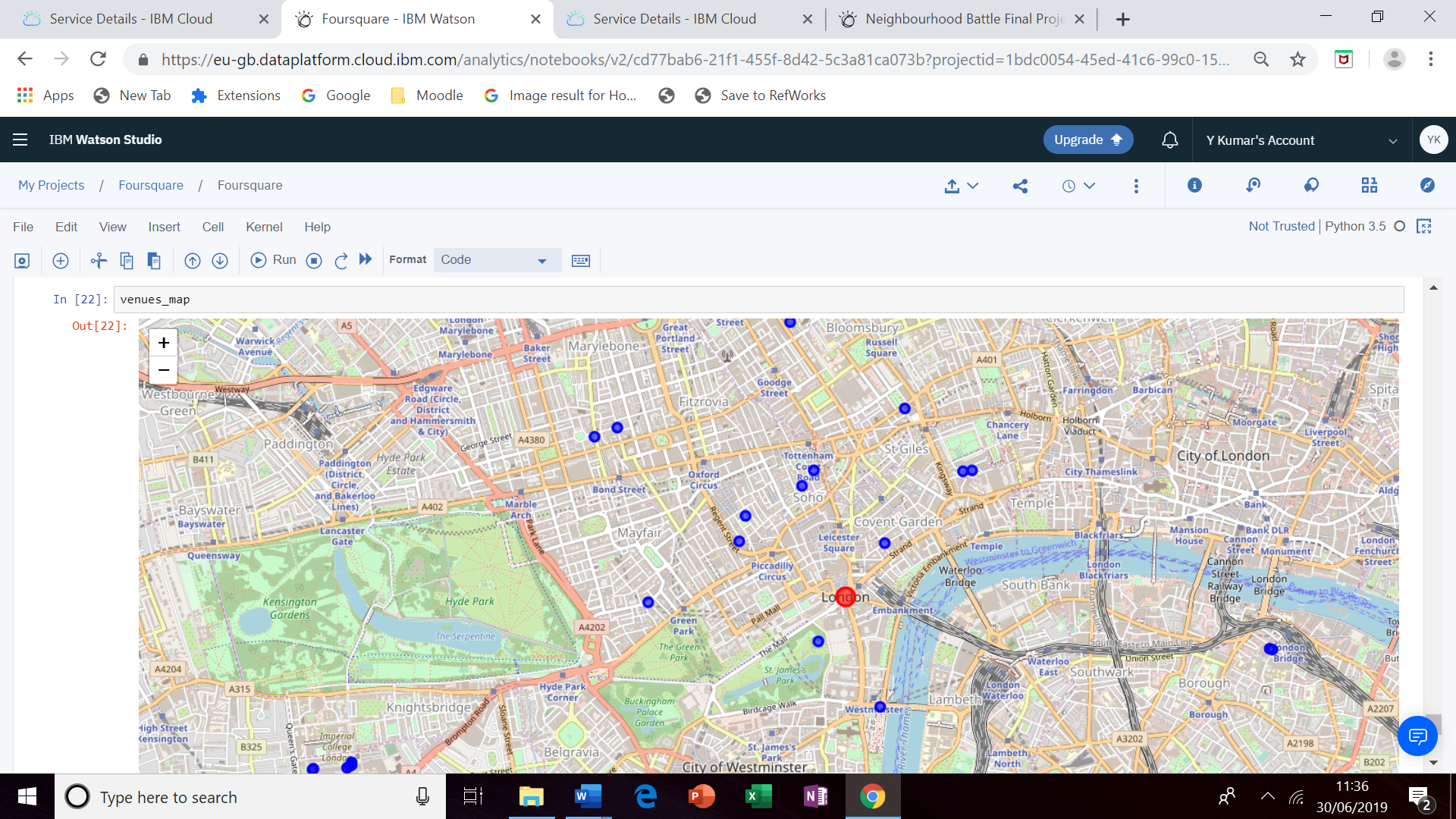
It can be easily observed that using sns regplot we can find out that what should be the location of our next UKSPA for its success.



We can see from above Decision tree accuracy to predict the location of next UKSPA is 0.905.



The analysis was also performed using ‘K Neighbors Classifier’ algorithm and SVM and it was observed that F1 Score and Jaccard Similarity values are in the range of 0.7 which is quite good for supporting the analysis.



Finally, I used the foursquare API to find out the most trending venues for the desired location as plotted in the map above.

# Conclusion:

Presently this project uses task-specific algorithms but in the future "Deep Learning" could be added, which in turn could establish principles for further studies, and in particular "Deep Learning" may help realize one ambitious future goal, a relatively real-time model. Also more attributes or features can be added for a more accurate prediction of the success factors for the initiation and development of business clusters in regions. This more accurate prediction of the success of business clusters will not only save the government funding (which currently goes waste because of the in the majority cases failure of such initiatives) but will also help to boost the economy and society at large.

References:

BBC. (2019): City centres 'could become ghost towns’ [https://www.bbc.co.uk/news/business-47307865 30.03.2016](https://www.bbc.co.uk/news/business-47307865%2030.03.2016) 21.02.2019

BBC. (2010): Six technium business centres closing in Wales. [http://www.bbc.com/news/uk-wales-11790468 30.03.2016](http://www.bbc.com/news/uk-wales-11790468%2030.03.2016)

Bahar, A & Mao, X (2004): Stochastic delay Lotka–Volterra model. J. Math. Anal. Appl. 292, 364–380

Cojocaru A.M.R. and Ionescu, S. (2016) 'The Advantages of Business Clusters', FAIMA Business & Management Journal, 4(2), pp. 31.

Engel, J.S. (2015) 'Global clusters of innovation', *California Management Review,*57(2), pp. 36-65.(Engel, 2015).

Klofsten, M. et al. (2015) 'Success Factors in Cluster Initiative Management', Industry and Higher Education, 29(1), pp. 65-77.

Lecluyse, L., Knockaert, M. and Spithoven, A. (2019) 'The contribution of science parks: a literature review and future research agenda', The Journal of Technology Transfer, 44(2), pp. 559-595.

Mercedes Delgado, Michael E. Porter and Scott Stern (2010) 'Clusters and entrepreneurship', *Journal of Economic Geography,*10(4), pp. 495-518.

Morosini, P. (2004): Industrial Clusters, Knowledge Integration and Performance, World Development, 32 (2) 305-326.

Porter, M. E. (1990) ‘The Adam Smith address, Location, clusters and the new…’ Business Economics Jan 1998, Vol. 33 Issue 1, p7 Palgrave Macmillan Ltd. accessed via Ebscohost

Porter, M. E. (1990) The Competitive Advantage of Nations. New York: The Free Press.

Piperopoulos, Panos G.. Entrepreneurship, Innovation and Business Clusters, Routledge, 2012. ProQuest Ebook Central,

Wadhwa, V. (2013): Silicon Valley can't be copied, Technology Review, 116(5), 87-88.

Wootton, K. L. & Stouffer, D.B (2016): Many weak interactions and few strong; food-web feasibility depends on the combination of the strength of species’ interactions and their correct arrangement. Theoretical Ecology, 9(2) 185–195.

Gov.uk

UKSPA

Company house

ESRI ARC GIS

Watson studio

MSSQL studio