RESEARCH PROPOSAL

Module: Research Methods and professional Issues INM373
Name: Mussa Yousef
Course: MSc Data Science
Email: mussa.yousef.1@city.ac.uk

Evaluating the accessibility and equity of vaccination centre locations

1. Introduction

The purpose of the proposed project is to evaluate the geospatial distribution of vaccination centres in England, hence justifying the level of parity of accessibility for vaccination centres in England. The recent pandemic was completely unprecedented, the reconstructions of policies and healthcare have caused huge disruptions across all sectors. This project is relevant and is aimed to withstand time, throughout western history, there have been recorded pandemics that each shaped our history, society, and inclusive of shaping the fundamental principles of modern health sciences.[1] Globally, since the 1960s we have only been able to eradicate smallpox, other diseases since have been hard to dislodge and could easily outbreak again.[2]

Therefore, the recent Covid-19 merged with our developments of infrastructures and technology this project aims to evaluate the distribution of vaccination centres, assess the level of access citizens have to them. Henceforth, our underlying question is "How could the geospatial allocation network location of vaccination centre be improved?". Through our extensive analysis we will be building metrics on the level of accessibility, in the following perspectives:

- Transportation modes available in each geographical location
- The demographic surrounding each vaccination centre thus using statistical analysis such as clustering
- Infrastructure surrounding vaccination centres

We will further develop our analysis by challenging [3] from the perspectives above, equipped with the geographic and spatial-statistical strategies introduced in [4] and [5]. Our research generated will be from a critical social perspective hence our beneficiaries of this research will be aimed at:

- The national health care service providers (NHS) - providing social perspectives to review current strategies and provide decision-making metrics for future pandemic strategies.

- Political leaders – An equitable research allows social responsibility to be exposed therefore provides accountability to be taken for reactive decisions from a pandemic. Therefore, minimising discrepancies is essential deliverance in healthcare.

- Businesses – Due to the lockdown caused by the pandemic and the stagnated rollout of the vaccinations, businesses across all sectors were adversely affected. By optimising centres the wider community will be able to be provided tailored health care i.e., Vaccination, therefore minimising the impact on businesses respectively.

- Education – Schools, universities, and education institutes alike would need mobilisation; therefore, community mobilisation is targeted here bringing together all resources available and tailoring this to our research perspectives, we aim to provide a contingency to limit the disruption to the education sector which we have seen with the Covid-19 pandemic.

2. Critical context

Geography is one of few disciplines that reasons to allow a synthetic approach to the interaction between biophysical and human variables [4]. Twenty years ago, evidence was scarce, certainly in published literature, that Geographical Information Systems (GIS) was being used in strategic decision making in the healthcare sector.[6] The recent, pandemic has highlighted the need for GIS, Hospitals, and healthcare providers utilised online and Real-time geospatial analysis of communities to monitor coronavirus cases [Figure 1], which enhanced forecasts of the disease spread. However geospatial analysis today remains a technology for static data and generally lacks the representation of dynamics, which is a major impediment to its use in spatial modelling.[8] Thus, focusing on decision making, our fundamental ideas area of research is rooted in "Tobler's" first law [7], we have therefore focused on our perspectives to find appropriate literature.

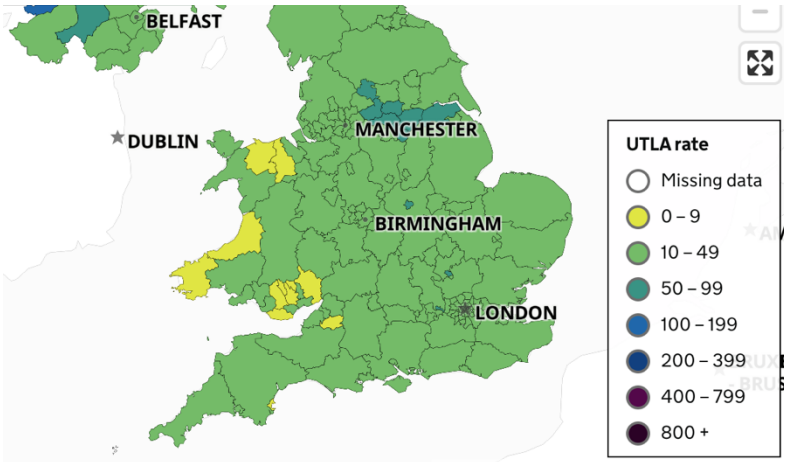

Figure 1: Geospatial mapping representation of covid-19 cases as of 26th April

Transportation and Infrastructure

Firstly, attempts have been made to highlight the significance of Transportation and Infrastructure on the accessibility of health care services. Tijs Neurens [9] developed the strategy to use the geographical study of gravity models. Fundamentally, gravity models, in this case, are where two vectors represent the health demand and supply, both share a relationship where the distance or time travel represents a time decay function. Therefore, the model characterises the relationship between distance and time travel which expresses the generalised cost for interaction with health centres, in our case vaccination centres.

Tailoring this strategy to our research, an advantage to this is that given the sheer metrics we would be able to calculate. We would therefore be able to visualise transportation flow, size, and demand to vaccination centres. Thus, optimising location, community distances, and healthcare services.

The drawbacks of using this methodology are that several factors have been omitted, for instance, the willingness of patients to travel, opportunities to travel, and usage of the aggregate district or community data which would lead to statistical bias when using point-based measures of geospatial analysis.
To develop and work around our disadvantages here, we will consider McGrail's [10], suggestion to employ a two-step floating catchment area (2SFCA) method to improve spatial

accessibility. This method allows us to involve key information which may affect vaccination centres accessibility, scrutinising distance, time, transportation available, and population. McGrail has highlighted that minimising the usage of aggregated data, although will enhance computational time will maximise sensitivity to small districts, consequently, limiting statistical bias. McGrail carried out this research looking at the access to a general practitioner (GP) in Australia. One interesting finding in this paper is that rural services are more dispersed, so that rural populations commonly access services beyond their immediate community [10].

Linking this back to our research, considering Tobler's first law that everything is relevant. We will dissect all relevant metrics, external factors and further develope sophisticated methodologies to justify the analysis on transportation and infrastructure on accessibility to healthcare. However, the nature of our review isn't for a GP or general health services, therefore the novelty of our research is the significance of a pandemic. Both pieces of literature reviewed here, provide a sophisticated approach to geospatial analysis and optimisation that are relevant. Providing insight on techniques that will be adopted when evaluating vaccination centres accessibility. This also allows a systematic plan to optimise accessibility, providing proactive solution planning for our beneficiaries.

Demographic

Diverse efforts have been made on the demographic study on accessibility to fundamental public services. Like the literature examined for geospatial analysis of infrastructure and transport. Most geospatial researchers equate people's accessibility with distance, however, most fail to account for acceptability or willingness for the service.[11] Hawthorne [11], has tailored this in his research on demographics and geographical accessibility to healthcare. The idea of satisfaction-adjusted distance (SAD) adjusts street network distance to account for the added perceived which residents feel accessible. This would require engagement and community research in the form of questionnaires, considering the scale of our research this would not be possible.

Kim [12] also utilised 2SFCA to investigate accessibility, further regression analysis is added introduced multiple factors of demographic. Hertel, Karen [13], researched the accessibility of healthcare from a demographic perspective in the US. The methodology introduced was the utilisation of census data and rather than scrutinising the accessibility of each region/city they analysed the concentration/scatter of different ethnic/education/gender/social status/crime rate/education in regions and built facets to their analysis. The advantage here is that the results provide a rich body of evaluative information tailored to the demographics.

3. Approaches: Design, Analysis and Evaluation

This section will describe the workflow of our project, breaking down our methods for design, analysis, and evaluation of the result. For each step of the workflow, we intend to present and justify what we aim to achieve in our outcomes.

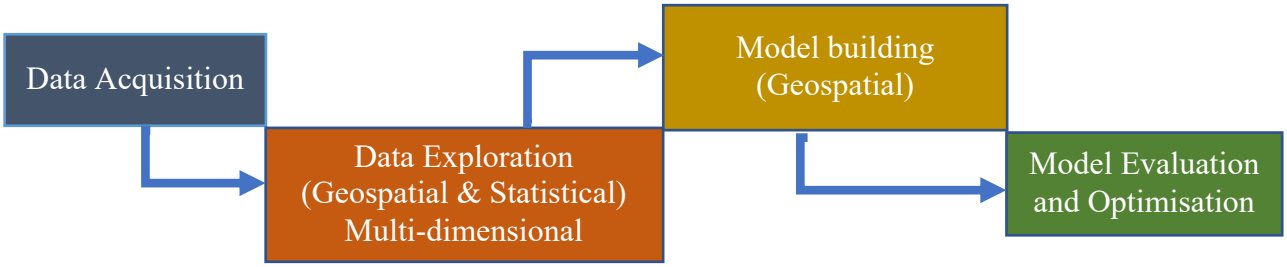

Figure 2: Workflow of project – Main pivotal stages

Before going into the core steps of our workflow, we will first need to cherry-pick the tools we will use. My most comfortable, programming language to use is python, mainly due to its ability to be used across multiple disciplines with paired libraries. For this project, I will relish using NumPy and pandas for data manipulation and scientific computers to assist with statistical analysis. We will do some visualisations of our statistical analysis to be able to communicate and highlight early finds to our beneficiaries, this will be done using Matplotlib and Seaborn. We will also use python for some of our geospatial analysis hence libraries such as Geopandas, Folium, pyProj and Arcpy. Further, into our project for our clustering and regression analysis, we will utilise the Machine learning module scikit-learn. As we aim to build extensive visualisations of both data exploratory and geospatial models, we will utilise Vega-Lite, a high-level visualization grammar. Tableau, an interactive data visualization software. QGIS, open-source geographical information systems application.

Our first phase begins with data acquisition, this project itself has multiple layers hence to be able to make sure we are correctly finding tailored sources of data. We must evaluate our phases of data explorations, model building, and model evaluation; each phase would relatively increase the depth of analysis thus build on our evaluation on accessibility.

Data Exploration

We have broken down our exploration facets into four phases. In phase one, we scrutinise the locations of vaccination centres and calculate the distance between centres and Towns/Districts/Cities. Compare this to the ratio of engagement/population. We explore using methods such as Multidimensional scaling (MDS), Principal component analysis (PCA), and Euclidian distance. In phase two we analyse landscape and infrastructure which may affect accessibility within these Towns/Districts/Cities, examine patterns with the engagement of vaccination centres scrutinised in phase one. In phase three, we now study transport links, car ownership, and disability accessible features surround vaccination centres. We, therefore, link this to our finding of phases 1 & 2, consequently, examine if further patterns could be identified. In phase four, we introduce the Two-step floating catchment area method (2SFCA), with our findings from phase 1-3, inspired by [9& 10] we calculate the distance in ratio hence incorporate distance decay between vaccination centres location. This will be translated onto Geospatial analysis, therefore, examining if further patterns could be found.

Model Building

We have similarly broken down our model building facets into three phases. Phase one, we introduce demographics of each consensus, building thematic choropleth maps, we use this to identify further patterns between different indexes of demographics such as ethnicity, crime, age, education, and social status. We overlay our positions of vaccination centres and scrutinise the equity in service and different spectrums of demographics. In phase two, we focus on building regressions to calculate the statistical differences between demographics and accessibility scrutinised in phase 4 in our data exploration. We couple this with cluster analysis, we build geospatial clusters with our 2SFCA analysis and demographic regression analysis. Phase three, inspired by [5] we, utilise radian and area composition with the research uncovered to assess where optimisation is needed.

Model Evaluation
Here we evaluate our finding particularly all our facets of geospatial analysis linked with modelling building. We assess if we have adequately scrutinised the accessibility and equity of vaccination centres. Where optimisation is needed, therefore we can utilise our findings and provide a novel approach to justify our evaluation. A scoring system that takes into consideration both equity and accessibility of each vaccination centre. This could be calculated or evaluated using PCA or 2SFCA which maintains both variance and weighting on optimising the location of vaccination centres. Therefore, where discrepancies are found for equity and accessibility, we aim to optimise and provide centroid locations where both equity and accessibility meet utility maximisation.

Data Acquisition
Now after evaluating our workflow we can now understand the depth in which our research will be, we would need extensive data to be able to adequately meet our research objective. We would need the following primary data:

- List of vaccination sites
- Covid daily announced vaccinations
- Geospatial boundaries of England & London
- Major Road & Public transport network

Following supplementary data:

- Census Data (Most recent)
- Indices of deprivation 2019
- Ethnic population projection
- Data on disabled population per city/ disabled accessibility per city

To conclude, this section, in term of privacy and ethics, this project doesn't use any information which isn't made public hence is compliant with the regulations. Our data is anonymised hence there isn't any data which jeopardises any privacy regulations. Furthermore, no external participants will be involved in this research project, hence our project is not a risk to anyone.

4. Workplan

This research project will begin on the 1st of June 2021 it is expected to be completed on the 30th of September 2021. Figure 3 visually outlines the estimated time scales and tasks for this project. our next sector evaluates risk, which may cause a slight delay in task completion, these are unforeseen which we will make provision and make sure the impact is minimised.

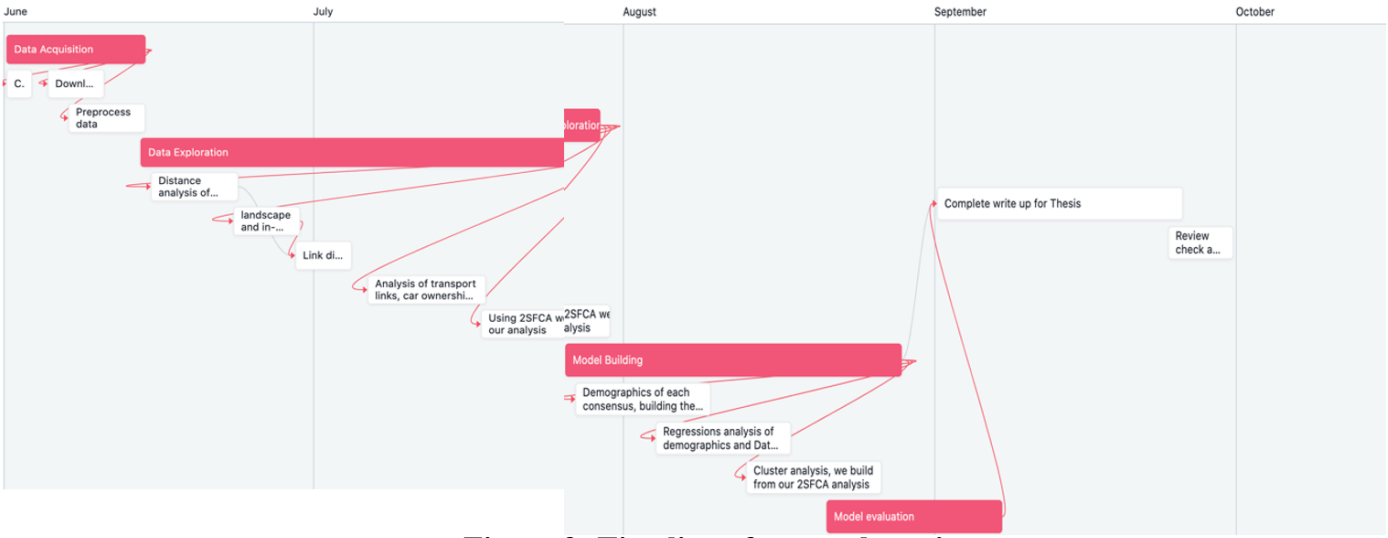

Figure 3: Timeline of research project

Project Stages	Description of tasks	Expected outcomes
Data Acquisition, Preparation and Tools	• Collect full spectrum of Data • Pre-process data • Download and familiarise with all software needed for our product	Prepare data and tools for project. Organise data and familiarise with tools.
Data Exploration	• Distance analysis of Vaccination Centres • Landscape and infrastructure analysis via geospatial analysis • Link distance and landscape analysis and evaluate comparison • Analysis of transport links, car ownership, and disability accessible features surround vaccination centres • Using 2SFCA we merge our analysis	Build graphical representations of multiple perspectives of accessibility. Produce intricate geospatial visualisation.
Model Building	• Demographics of each consensus, building thematic choropleth maps • Regressions analysis of demographics and Data exploration • Cluster analysis, we build from our 2SFCA analysis	Deepen our finding with models built with Demographics of respective cities and town. Contextualise our evaluation using visual.
Model evaluation	• Review, evaluate, interpretate findings • Suggest optimised vaccination centre	Suggest optimised accessibility and equity locations Complete report and visualise models

Table 1: Description and expected outcomes of the main project stages

5. Risks

Concluding our proposal, we scrutinise the risk involved in this project. For each risk we find we analyse and evaluate the impact this would have on our project. We first describe the type of risk this is, between it being a technical or non-technical type of risk. We then evaluate the likelihood of this risk happening, grading this between low, medium, and high. Similarly, we evaluate the consequence scrutinising this in the following five levels, very low to very high. Lastly, we provide a mitigating strategy, classifying these between deflection, avoidance, and contingency therefore briefly explaining our strategy.

Risk	Type	Likelihood	Consequence	Mitigating Strategy
Unforeseen illness	Non-Technical	Low	Very High	Contingency: we have overstated our time scales by a day to allow any flexibility
Geospatial knowledge difficulty	Technical	Low	High	Deflection: Seeking assistance from supervisor/training courses when familiarising with software
Computer/hard-disk failure	Technical	Low	Very High	Avoidance: Constant back-ups of work making sure an updated project is available
Researcher Bias	Non-Technical	Medium	Medium	Avoidance: Making sure statistics reinforces each assumption keeping evaluating factual
Milestones delayed	Non-Technical	Medium	High	Avoidance: Overestimate timings in which a milestone is reached, hence if anything is delayed, we have our contingency in place
Difficulty acquiring data	Non-Technical	Low	High	Avoidance: Using tools which I am already familiar with (Python)
Unavailability of Supervisor	Non-Technical	Medium	Medium	Contingency: Using other tools such as Slack & Moodle to contact colleagues/other professors. Contact supervisor to confirm progress when back available

Reference so far

[1] Huremović D. (2019). Brief History of Pandemics (Pandemics Throughout History). Psychiatry of Pandemics: A Mental Health Response to Infection Outbreak, 7–35. https://doi.org/10.1007/978-3-030-15346-5_2
[2] Laura Spinney (Oct 2019). How pandemics shape social evolution. Nature 574, 324-326 (2019), Security threat , p1. https://doi.org/10.1038/d41586-019-03048-8
[3] Barabiuk C (2021). "Covid-19: How the UK Vaccine rollout delivered success, so far", BMJ; 372 doi: https://doi.org/10.1136/bmj.n421 (Published 18 February 2021)

[4] Franch-Pardo et al(2020). "Spatial analysis and GIS in the study of COVID-19". A review, Science of The Total Environment, Volume 739,2020,140033, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2020.140033

[5] Kuupiel, D.; Adu, K.M.; Bawontuo, V.; Adogboba, D.A.; Drain, P.K.; Moshabela, M.; Mashamba-Thompson, T.P. Geographical Accessibility to Glucose-6-Phosphate Dioxygenase Deficiency Point-of-Care Testing for Antenatal Care in Ghana. Diagnostics 2020, 10, 229. https://doi.org/10.3390/diagnostics10040229

[6] Gary Higgs, Myles Gould, "Is there a role for GIS in the 'new NHS'?,", Health & Place, Volume 7, Issue 3, 2001, Pages 247-259, ISSN 1353-8292, https://doi.org/10.1016/S1353-8292(01)00014-4

[7] Waters, Nigel. (2017). Tobler's First Law of Geography. 10.1002/9781118786352.wbieg1011

[8] T.-X. Yue, Z.-P. Du, Y.-J. Song, Spatial Models and Geographic Information Systems, Encyclopedia of Ecology, Academic Press, 2008, Pages 3315-3325, ISBN 9780080454054, https://doi.org/10.1016/B978-008045405-4.00233-0

[9] Tijs Neutens, Accessibility, equity and health care: review and research directions for transport geographers, Journal of Transport Geography, Volume 43, 2015, Pages 14-27, ISSN 0966-6923, https://doi.org/10.1016/j.jtrangeo.2014.12.006

[10] McGrail, Matthew. (2012). Spatial accessibility of primary health care utilising the two-step floating catchment area method: An assessment of recent improvements. International journal of health geographics. 11. 50. 10.1186/1476-072X-11-50

[11] T.L.Hawthorne (2011). Using GIS and perceived distance to understand the unequal geographies of healthcare in lower-income urban neighbourhoods. TheGeographicalJournal.Mei-po Kwan. https://doi.org/10.1111/j.1475-4959.2011.00411.x

[12] Y.Kim (2018). Enhancing healthcare accessibility measurements using GIS: A case Study in Seoul, Korea. Department of Civil and environment, institute of science and technology, Khalifa University, Abu Dhabi, UAE. DOI: 10.1371/journal.pone.0193013

[13] Hertel, Karen (2007). "GIS and census data: tools for library planning". Nancy Sprague University of Idaho Library, Moscow, Idaho, USA. https://www.emerald.com/insight/content/doi/10.1108/07378830710755009/full/html

RESEARCH PROPOSAL

Research Ethics Review Form: BSc, MSc and MA Projects

Computer Science Research Ethics Committee (CSREC)

http://www.city.ac.uk/department-computer-science/research-ethics

Undergraduate and postgraduate students undertaking their final project in the Department of Computer Science are required to consider the ethics of their project work and to ensure that it complies with research ethics guidelines. In some cases, a project will need approval from an ethics committee before it can proceed. Usually, but not always, this will be because the student is involving other people ("participants") in the project.
In order to ensure that appropriate consideration is given to ethical issues, all students must complete this form and attach it to their project proposal document. There are two parts:

PART A: Ethics Checklist. All students must complete this part. The checklist identifies whether the project requires ethical approval and, if so, where to apply for approval.

PART B: Ethics Proportionate Review Form. Students who have answered "no" to all questions in A1, A2 and A3 and "yes" to question 4 in A4 in the ethics checklist must complete this part. The project supervisor has delegated authority to provide approval in such cases that are considered to involve MINIMAL risk.
The approval may be ***provisional*** – *identifying the planned research as* likely to involve MINIMAL RISK. In such cases you must additionally seek ***full approval*** from the supervisor as the project progresses and details are established. ***Full approval*** must be acquired in writing, before beginning the planned research.

A.1 If you answer YES to any of the questions in this block, you must apply to an appropriate external ethics committee for approval and log this approval as an External Application through Research Ethics Online - https://ethics.city.ac.uk/		*Delete as appropriate*
1.1	Does your research require approval from the National Research Ethics Service (NRES)? *e.g. because you are recruiting current NHS patients or staff?* *If you are unsure try - https://www.hra.nhs.uk/approvals-amendments/what-approvals-do-i-need/*	**NO**
1.2	Will you recruit participants who fall under the auspices of the Mental Capacity Act? *Such research needs to be approved by an external ethics committee such as NRES or the Social Care Research Ethics Committee - http://www.scie.org.uk/research/ethics-committee/*	**NO**
1.3	Will you recruit any participants who are currently under the auspices of the Criminal Justice System, for example, but not limited to, people on remand, prisoners and those on probation? *Such research needs to be authorised by the ethics approval system of the National Offender Management Service.*	**NO**
A.2 If you answer YES to any of the questions in this block, then unless you are applying to an external ethics committee, you must apply for approval from the Senate Research Ethics Committee (SREC) through Research Ethics Online - https://ethics.city.ac.uk/		*Delete as appropriate*
2.1	Does your research involve participants who are unable to give informed consent? *For example, but not limited to, people who may have a degree of learning disability or mental health problem, that means they are unable to make an informed decision on their own behalf.*	**NO**
2.2	Is there a risk that your research might lead to disclosures from participants concerning their involvement in illegal activities?	**NO**

2.3	Is there a risk that obscene and or illegal material may need to be accessed for your research study (including online content and other material)?	**NO**
2.4	Does your project involve participants disclosing information about special category or sensitive subjects? *For example, but not limited to: racial or ethnic origin; political opinions; religious beliefs; trade union membership; physical or mental health; sexual life; criminal offences and proceedings*	**NO**
2.5	Does your research involve you travelling to another country outside of the UK, where the Foreign & Commonwealth Office has issued a travel warning that affects the area in which you will study? *Please check the latest guidance from the FCO - http://www.fco.gov.uk/en/*	**NO**
2.6	Does your research involve invasive or intrusive procedures? *These may include, but are not limited to, electrical stimulation, heat, cold or bruising.*	**NO**
2.7	Does your research involve animals?	**NO**
2.8	Does your research involve the administration of drugs, placebos or other substances to study participants?	**NO**
A.3 If you answer YES to any of the questions in this block, then unless you are applying to an external ethics committee or the SREC, you must apply for approval from the Computer Science Research Ethics Committee (CSREC) through Research Ethics Online - https://ethics.city.ac.uk/ **Depending on the level of risk associated with your application, it may be referred to the Senate Research Ethics Committee.**		*Delete as appropriate*
3.1	Does your research involve participants who are under the age of 18?	**NO**
3.2	Does your research involve adults who are vulnerable because of their social, psychological or medical circumstances (vulnerable adults)? *This includes adults with cognitive and / or learning disabilities, adults with physical disabilities and older people.*	**NO**
3.3	Are participants recruited because they are staff or students of City, University of London? *For example, students studying on a particular course or module.* *If yes, then approval is also required from the Head of Department or Programme Director.*	**NO**
3.4	Does your research involve intentional deception of participants?	**NO**
3.5	Does your research involve participants taking part without their informed consent?	**NO**
3.5	Is the risk posed to participants greater than that in normal working life?	**NO**
3.7	Is the risk posed to you, the researcher(s), greater than that in normal working life?	**NO**
A.4 If you answer YES to the following question and your answers to all other questions in sections A1, A2 and A3 are NO, then your project is deemed to be of MINIMAL RISK. **If this is the case, then you can apply for approval through your supervisor under PROPORTIONATE REVIEW. You do so by completing PART B of this form.** **If you have answered NO to all questions on this form, then your project does not require ethical approval. You should submit and retain this form as evidence of this.**		*Delete as appropriate*
4	Does your project involve human participants or their identifiable personal data? *For example, as interviewees, respondents to a survey or participants in testing.*	**NO**