**TASK 5**

Research Sample

A research sample is a group of individuals, things, or products chosen for evaluation from a wider population (Education Centre, 2006). To guarantee that the findings from the study sample can be applied to the entire population, the sample should be representative of the population.

For my research, I collected Satellite data from the Sentinel-2 mission, which was applied for its relatively high 10 m spatial resolution, and for having radiometry with three vegetation red edge bands (Abdi, 2019). These two characteristics make the Sentinel-2 data appealing for LCLU mapping. The said datasets were downloaded from the satellite imagery platform, the Copernicus Hub. These are cloudless images, spread over the span of 4 years: 2 years between 2015 to 2017, and 2 years between 2020 to 2022. Each set of years has 11 images, for a total of 22 images for both groups, which are roughly in alternate months. It was crucial for these images not to be affected by distortion, such as clouds. As a result, care has been taken in determining which images to obtain for the winter months when cloud-free images are scarce.

Another dataset gathered was from ARPA, regarding agriculture activity detected within a whole year from their on-the-spot land inspections. This dataset was split into two classes, being NMM i.e., no-minimum agricultural activity, and Arable. These classes were observed as “no Ploughing” i.e., the declaration is confirmed, and another as “bare Soil” i.e., the declaration is not correct. For every parcel, there were 1, 2 or 3 confirmed inspection dates. The total number of parcels within the study area includes 85 ‘arable’ parcels, 48 ‘non-cultivated’ parcels and 34 ‘other’ parcels. After running both algorithms, data extraction was done, and a custom dataset in Geo Tiff format was created. The files were utilized on QGIS to view and extract the outputted marked land listed under their respective classes (arable, non-cultivated or other) and to see the changes in the land from the first set of images from 2015 to 2017 to the second set from 2020 to 2022.

Ethics

* *Prior to beginning the study*

For my study, I first had to research how my paper’s objective would be beneficial and what was missing in the field of remote sensing. After deciding, I then submitted a research proposal form, also known as the Statement of Intent (SOI) for the Institute Research Sub-Committee (IRC) formally established at every institute to verify that the proposal sufficiently follows the examination bodies' standards as well as the ethical guidelines. It then got approved for me to carry out the research intended. Considering that my study is a quantitative one, which did not consist of any participants, no ethical measures were needed. This is scientific research on proof of concept with an environmental significance. Hence, no personal and confidential data will be breached and no identities to divulge. There is also no consequence of possible physical harm during the implementation of this research and no consequence of possible moral harm during the implementation of this research. This study is not tied to any commercial interest, hence there will be no consequences on businesses during the implementation of this research.

* *Beginning the study*

In Kuss et al’s study (2022), the participants were chosen after a prior study was completed. Those that were interested were encouraged to give their contact information at the end of the survey. All participants who provided contact information were contacted and invited to participate in an interview. As a thank you for their time, participants were given a GBP 20 voucher. From this, 20% expressed an interest in participating. Participants in both studies had to be able to play video games, converse in English, and identify as female. Informed consent was sought from all subjects involved in the study, along with the individual's preference for how they wanted to be interviewed and their availability. For convenience, participants could choose from a range of interview techniques, including in-person, Skype audio, or Skype written, which was offered in case they didn't have access to a microphone or sound equipment, or if they were apprehensive about speaking out loud. All interviews were conducted in private, with the consent of the participants to record the session. In-depth and personalized comments were encouraged in private interviews.

* *Collecting data*

The data collected for my study included an on-the-spot land inspection dataset from ARPA, in which land surveyors inspect which lands are being cultivated and which are not. Prior to this, I contacted them to ask me to supply this information where a meeting was set up to discuss further. The manager I spoke to later contacted his superior to get permission for them to supply this Land Parcel Identification System (LPIS) GIS data. He obtained approval on the 22nd of November 2021 and sent me a non-disclosure agreement (NDA) for me to sign before forwarding the data to me and asked for my ID card number. The NDA covers the data specifications, the intended use of data, confidentiality and publication, indemnity, and the sharing of data. The document was duly filled and signed by both parties. The data was later sent via SharePoint as shapefile layer to be transferred to QGIS.

In Yang and Cornelius’s study (2004), data collection was done by interviews, observations, and documents. Each participant in this study underwent two interviews: an unstructured interview and a structured interview. To ensure ethical considerations, the interviews took place wherever the participants were most comfortable, such as in their dorm room, office, or the school library. Following the participants' consent, the first observation was arranged to see how he/she worked for the online class. Two observations were conducted to collect more qualitative data, where each session lasted approximately 60 minutes. Participants were given the option of scheduling their observations at their leisure and in conducive environments, in the same setting where they completed their online coursework, including the participants' homes, dorm rooms, offices, and the campus library, among other places. Documents were requested from all parties which contained printouts and other data to substantiate the information acquired and witnessed during the observation and interview process.

* *Analysing data*

When it came to analysing data from my study, I went for an unbiased approach with testing two common algorithms used in remote sensing for land cover change, the random forest classification and K-means clustering. According to past papers I read for my literature review which utilized such algorithms, they all conclude that the Random Forest algorithm gave better accuracy results. I applied both algorithms to my research to confirm the results of such past research papers.

Kuss et al. (2022) mention that their thematic analysis for all interviews follows full anonymization, meaning that their data processing technique removes or modifies personally identifiable information to prevent it from being linked to a specific person. It's also a significant factor in Google's privacy commitment.

* *Reporting, Sharing and Storing data*

According to the contractual agreement between me and ARPA, I am obligated to send them a final draft of the entire dissertation before it is published, with the citations and conclusions generated from ARPA's data being vetted by the agency. It states that the agency reserves the right to request amendments to any text that it believes, in its sole discretion, misrepresents any data it has provided or conclusions drawn from it, and I agree to make every effort to comply with such request and, in any case, not to publish any reference to which the agency has not expressly consented in writing, bearing sole responsibility for any failure to do so.

References:

Education Research, 2006, *Sampling in Research*, Available at: <https://www.thh.nhs.uk/documents/_Departments/Research/InfoSheets/16_sampling_research.pdf> (Accessed: 30 March 2022)

Kuss, D.J.; Kristensen, A.M.; Williams, A.J.; Lopez-Fernandez, O., 2022, To Be or Not to Be a Female Gamer: A Qualitative Exploration of Female Gamer Identity. *Int. J. Environ. Res. Public Health 2022*, 19(3):1169, DOI: 10.3390/ijerph19031169, <https://www.researchgate.net/deref/https%3A%2F%2Fdoi.org%2F10.3390%2Fijerph19031169>

Yang, Y, Cornelius, L, F, 2004, *Students’ Perceptions towards the Quality of Online Education: A Qualitative Approach*, Available at: <https://files.eric.ed.gov/fulltext/ED485012.pdf> (Accessed: 5 April 2022)