CASA0009: Spatial Data Capture and Analysis
Assessment Guidelines

You are required, as part of a group, to develop an interactive website that conveys deep analysis and insights drawn from one or more datasets. The website should be designed with contemporary web development tools, should introduce advanced data analysis processes, and allow the user to gain a greater understanding of the subject being presented. The analysis should focus on a spatial phenomenon (e.g. a geographic or urban process), and the website should be easy to use and well designed.

The assessment projects will be completed in groups of four students. Groups will be required to submit a website describing the analysis and visualisation of a dataset chosen by the group. Each group will be required to also submit a report about the rationale and execution of the data analysis and visualisation project. Individuals are required to submit a short justification of their contributions to the group work, and the relative contributions of fellow group members.

Final grades will be awarded individually for this coursework. Where groups are observed to have contributed equally, each member will receive the same mark for the website and report. Where it is clear that a group member has contributed less than their colleagues (as reported by team members), that individual's mark will be reduced accordingly.

This coursework represents 100% of the overall module assessment.

Stages of Assessment

Group project pitch (unassessed but required)	
Group website - Interactive visualisation and analysis of dataset	**60%**
Group presentation of project	**10%**
Group report - 5,000 words (+/- 10%)	**30%**
Individual personal reflection (unassessed but required)	

Key Dates for 2020/2021

Group pitches of project idea – Thursday 25th / Friday 26th March
Group presentation of website – Monday 17th May
Individual report submission via Moodle – Wednesday 19th May at 5pm

Assessment Guidelines

Group Pitches – Unassessed but required - Lecture Week 10
5 - 7 min presentation with 3 minutes for questions

This presentation is rough overview of the group work, the identification of a relevant dataset, how they attempt to analyse the data and what the group aims to achieve at the end of the project. It is important to note that groups will not be held to what they present, and amendments can be made later on. The aim of this informal presentation is to ensure sure students have an achievable goal within the timeframe and to source valuable feedback from tutors and the rest of the class.

This pitch is not assessed but is a required element of the coursework. Due to time constraints, we expect at least 2 people from each group to present but we expect all members of the team to contribute to the presentation (i.e. be present at the presentation and able to answer question from the group).

Group Website – 60%

This website is the main component of the assessed group work for this course. Groups are required to create a website, with multiple pages that convey a story about a dataset that has been sourced by the group. As a bare minimum, this website should contain an interactive visualisation, a graph of data values, a description of the analysis, a link to the source data, and a short explanation of the project.

The website should be interactive and pull data from a database linked to a server-side component. Analysis results may be pre-processed, and results recorded in the database, then visualised on the website in the form of graphs, maps etc. You are expected to demonstrate the full extent of advanced data analysis and interactive visualisation skills you have been taught during this course. You are furthermore expected to make use of those tools taught on the course in completing the analysis and visualisation work (e.g. MySQL, Python, JavaScript).

Students are encouraged to use HTML/JavaScript frameworks (such as Bootstrap, Leaflet, Google Maps, etc.) to improve the functionality of the website. Students are reminded that the course is assessed on the skills and techniques learnt during the course, as such, we therefore expect all analysis to be performed using Python, extending to libraries beyond those taught during the course if necessary.

The website will be assessed by the following criteria:

- Use of advanced data analysis methods to explore dataset 30%
- Use of compelling and appropriate visualisation methods 30%
- Exploratory, 'storytelling' nature of analysis description 20%
- Design and aesthetics of website design 20%

Group Presentation – 10%

15-minute presentation (by all members of the group) with 5 minutes for questions.

Each group is expected to produce a presentation showing off the website created, how the group carried out the analysis of the dataset, the limitations of the analysis and how the interactive tool works behind the scenes. It is important to note that lecturers and researchers from CASA will be invited to the presentation sessions, so do not assume any prior knowledge of your project or its purpose.

Groups should detail what works, and potentially what doesn't work, and any future modifications that could be made given more time. Questions and feedback may be provided by the tutors or the class and can be referenced in the individual report.

The presentation will be judged on the following criteria:

- Scope and ambition of project 30%
- Execution of the project 30%
- Demonstration of teamwork 20%
- Clarity of presentation 20%

We expect all members of the team to contribute (i.e. talk!) and present their individual contribution to the group project (as a rough guide, each person should allow 5 minutes for their part of the main presentation).

Group Report – 30%
5000 word written report (+/- 10%)

Each group must write a report about the design, development, and execution of the project. The reports should include a case of where the project fits in the wider context of research (with relevant references given) and data visualisation available on the web. The report should describe how the entire website was put together, including the collection and handling of data, development of analysis, and design of the data visualisation and user interaction. A conclusion should detail how, if given more time, they would improve the project.

A breakdown of how the report will be marked is as follows:

- Project context and aims (incl. reference to relevant literature and projects) 20%
- Design Rationale 70%
    - Data collection, handling, cleaning and management
    - Data analysis, methods, and major findings
    - Data visualisation
    - Technical integration between elements
- Presentation and style of report 10%

Individual Personal Reflection – unassessed
500 word written report (+/- 10%)

Each student is required to submit a personal reflection of their contribution to the group work, and provide an indication of the relative contributions of their team members. This part of the work is unassessed, but will be used in judging the participation of all team members in executing the group work. Although this work is unassessed, it is a required submission from all students.

The personal reflection should cover the aspects of the group work that you have contributed, both with respect to the website (including technical contributions) and report. You should also describe the contributions of all other team members in similar detail.

Alongside the descriptive aspect, you should provide a score out of 10 for each group member, reflecting how much you think they have contributed to the project. If you feel each team member contributed equally then give all of your peers 10 out of 10. If you feel two members have contributed equally to yourself, and one less so, assign two peers 10, and the other a lower mark indicating their relative effort. These marks will contribute towards the assigning of marks to your team members.

The reflection should be personal and should not be a collaborative effort conducted with some or all other members of the group.

If you have any major issues within your team you should talk with the module leader (Steve).

Examples of Online Visualisations

Here are a few examples to get your minds working. We do not expect the level of quality or detail that some of these projects go into. However, this list should serve as a guide of the types of analysis and visualisation we are looking for.

- How Americans get to Work - http://flowingdata.com/2015/01/20/how-americans-get-to-work
- NYC Taxis: A day in the Life **-** http://nyctaxi.herokuapp.com
- Analysis of US microbreweries - https://pudding.cool/2017/04/beer/
- London Bike Share Map **-** http://bikes.oobrien.com/london
- A Month of Citi Bike- http://projects.newyorker.com/story/citi-bike.html

- Trace Media Flickr Cities - http://www.tracemedia.co.uk/luminous/flickr
- Global Slavery Index - http://www.globalslaveryindex.org/findings
- Ebola Timeline Map - http://simonbjohnson.github.io/Ebola-Timeline-Map
- Interactive NBA Followers Map - https://interactive.twitter.com/nba_followers
- Berlin Marathon analysis - https://interaktiv.morgenpost.de/berlin-marathon-2016/
- Airbnb Visualisation - http://www.dwarshuis.com/various/airbnb/newyork/
- Sanctions Tracker - https://labs.enigma.com/sanctions-tracker
- Live Cyber Attacks - http://map.norsecorp.com
- Global Historical Emissions Map - http://aureliensaussay.github.io/historicalemissions/interactive.html
- Description of a transport peak spotting website - http://www.nand.io/projects/clients/peak-spotting/

Examples Datasets

We'd encourage you to find an interesting dataset that you all want to work on. Here are a few examples in case you are struggling to find one.

- NYC GPS taxi data - http://chriswhong.com/open-data/foil_nyc_taxi
- Yelp dataset - https://www.yelp.com/dataset
- UK Land Registry house sales data - http://landregistry.data.gov.uk
- Stop and Search Data by US State - https://openpolicing.stanford.edu/data/
- Traffic Accident and Traffic Flow data for 16 years - https://www.kaggle.com/daveianhickey/2000-16- traffic-flow-england-scotland-wales/settings
- Real-time crime data in Seattle - https://data.seattle.gov/Public-Safety/Seattle-Police-Department- 911-Incident-Response/3k2p-39jp
- Various FOI data releases can be found on WhatDoTheyKnow - https://www.whatdotheyknow.com/list/successful
- Crime Data in Buenos Aires - https://github.com/ramadis/delitos-caba
- Lots of open data for Bahrain - https://datasource.kapsarc.org/pages/home/
- City Cellular Traffic Map - https://github.com/caesar0301/city-cellular-traffic-map
- Flight data (requires Google account) - https://bigquery.cloud.google.com/table/bigquery- samples:airline_ontime_data.flights
- Beijing GPS taxi data - http://research.microsoft.com/apps/pubs/?id=152883
- International Migration data - http://www.global-migration.info/
- Plant Diversity in American National Parks Biodiversity - https://www.kaggle.com/nationalparkservice/park-biodiversity/data
- Wildlife Trade Database - https://www.kaggle.com/residentmario/cites-wildlife-trade-database/data
- H1-B Visa Petitions - https://www.kaggle.com/nsharan/h-1b-visa/data
- Baltimore Crime Data - https://www.kaggle.com/sohier/crime-in-baltimore
- Chicago Crime Data - https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2
- AWS Honeypot Cyber Attack Data (with originating locations) - https://www.kaggle.com/casimian2000/aws-honeypot-attack-data/data
- Vancouver Crime Data - http://data.vancouver.ca/datacatalogue/crime-data.htm