Dive into Data 2019

The aims of the data challenge are to develop novel and innovative ways to analyse and/or visualise Consumer Data Research Centre (CDRC) open or safeguarded data (https://data.cdrc.ac.uk) and to gain new insights into one (or more) of the UCL Grand Challenge priority themes below through data re-use.

Your data analysis and/or visualisation should be a clear, accessible, and engaging three-page case study that re-uses CDRC data to address one or more of the priority themes in a novel way. The results should benefit UCL Grand Challenge into gaining new insights, CDRC and your own research interests.

Data sets provided by CDRC cover a wide range of topics and therefore we invite UCL academics, researchers and master's students with interest in data from **ALL** disciplines to participate. Entrants will have the chance of winning a prize up to £500 in vouchers.

Sustainable Cities	Cultural Understanding	Human Wellbeing
PRIORITY THEMES		
Increasing and relentless urbanisation is highlighted as one of the most urgent issues facing the world.	Focus on relationships between people, communities and culture.	A wide range of disciplines, beyond traditional definitions of health, are examined to build up our understanding of wellbeing.
• Future Cities • Rethinking Housing • Resourcing Cities • Urban Mobilities	• Brexit • Cultural Engagement & Health • Growing up Multicultural • Immigration • Dynamics of Globalisation.	• Adolescent Wellbeing • Demographic Change • Disability • Environment • Achieving Wellbeing & Happiness.

Take part to:

- Increase visibility of your research
- Add value to your current research
- Expand your network
- Improve your skills
- Develop new skills
- Win a prize!
- Chance to present your work at the eResearch Domain conference (June 2019)

Who can participate?

- The challenge is open to anybody from master's level and above with interest in data
- It is open to teams consisting of one or more individuals, but ordinarily a maximum of three per team (remember, less team-mates, higher the value of the prize!)

For example (and inspiration), CDRC recently held a [data challenge](#) investigating "The immigration paradox explaining the Brexit vote". EU has also organised a [data challenge](#) to highlight the power of open data, developing an app related to a variety of EU themes.

Competition timeline: key dates

12 pm Friday 18th January 2019	Submission of Expression of Interest (EoI)
You should send the EoI form ([download here](#)) to **Dr Valerie Karaluka** (ovpr.beams@ucl.ac.uk) with the title "UCL Data Challenge 2018 Expression of Interest" by **12 pm 18th January 2019**. The form includes a brief project description, selected Grand Challenge and the themes they wish to address, and datasets you wish to use for the case study. **IMPORTANT:** If you plan to use **CDRC safeguarded data**, you will need to contact Sarah Sheppard (s.sheppard@ucl.ac.uk) to arrange access and ethical approval or submit an initial proposal at https://www.cdrc.ac.uk/data-services/using-our-data/ clearly stating that access is for the UCL Data Challenge. **Please note, those who would like to use safeguarded data should allow an additional 3 weeks for the request to be processed and we therefore advise you to do this well before the EoI submission deadline.**	
12 pm Wednesday 8th May 2019	Submission of the Case Study
Please submit a three-page case study suitable for academic but non-specialist audience to ovpr.beams@ucl.ac.uk by no later than 12 pm Wed 8 May 2019!	
June 2019	Conference – Winners announced
The panel will shortlist submissions and the finalists will be invited to present their work as a short oral presentation at the conference where the winner will be selected. In addition, participants may be invited to present their results as posters.	

Conditions & Rules:

- Participants must use open and/or safeguarded datasets obtained from CDRC. Other open data sets can be used in conjunction with the CDRC Data.
- Data that are not openly available to everyone cannot be used (except CDRC safeguarded)
- Reference all data sets used appropriately.
- PhD student's case studies must be complementary to their PhD research topic!
- You cannot submit an existing research project.

CDRC & Data

The CDRC creates, supplies, maintains, and delivers data to a wide range of users who wish to conduct research using big consumer data. Data comes in three level of access: open, safeguarded, and secure. For this challenge, you may use open and safeguarded data.

Open data: Data that are freely available to all for any purpose. A large variety of topics is covered and CDRC have in excess of 10,000 open datasets! The list of the most popular data is in **appendix**.

Safeguarded data: Data to which access is restricted due to license conditions. To apply please see Using CDRC Data Services and submit an initial application form clearly stating access is for UCL Data Challenge.

You can use any number of datasets provided by CDRC as well as any other open source datasets to complement the CDRC data. Examples of other open data sources are:

- The Government https://data.gov.uk/
- EU open data portal https://data.europa.eu/euodp/en/data/dataset?sort=views_total+desc
- Google https://toolbox.google.com/datasetsearch

Evaluation criteria (judged by representatives from eResearch Domain, CDRC, and Grand Challenges)

1) Relevance
   a. Relevance to the Grand Challenge priority themes

2) Novelty
   a. Does your study introduce new ideas to Grand Challenges?
   b. Does it introduce new methods for re-use of data?

3) Quality of Storytelling and Visualisation (communication)
   a. How engaging is the story and visual aids?
   b. Is it clear to be understood by non-specialist?

4) Impact
   a. Does your case study add new knowledge to the Grand Challenges?
   b. Does your case study create an opportunity for further research or funding?
   c. Does you study make optimal use of open or safeguarded data?

Help & further information

If you have any questions or would like assistance in selecting the theme and datasets, please contact Valerie Karaluka (v.karaluka@ucl.ac.uk) or Sarah Sheppard (s.sheppard@ucl.ac.uk), respectively.

Online tutorials are available on CDRC website: https://data.cdrc.ac.uk/tutorial and details of short course programme at https://www.cdrc.ac.uk/events/category/training-session/

Tutorial on how to use Python to manipulate and analyse data is available through Research IT Service Online training on Moodle: https://moodle-1819.ucl.ac.uk/course/view.php?id=11923

Further information on Research IT Service training can be found here:
https://www.ucl.ac.uk/isd/services/research-it/research-it-training

Lynda: A huge range of video tutorials supporting learning in software, creative and business skills – all free to UCL staff and currently enrolled students.

Appendix I: Popular CDRC Data Sets

Open Data sets:

2017 General Election Data

Business Census

CDRC 2011 Output Area Classification

CDRC 2011 London Output Area Classification

2011 Census Data Packs for England & Wales, Northern Ireland, Scotland

CDRC 2014 and 2018 Internet User Classification

CDRC Access to Healthy Assets and Hazards

CDRC Index of Multiple Deprivation Data Packs

CDRC Individual Income Estimates (PAYE)

CDRC Retail Centres and Catchments

Consumer Vulnerability

Electricity and Gas Domestic Consumption & Energy Performance Certificates

Workplace Zone Classifications

Safeguarded data sets

FCA Financial Lives Survey 2017

OOMap Bicycle sharing system docking stations

SmartStreet Sensor Footfall Atlas

Cluttons Real Estate Indices

High Street Retailer transactions

Domestic Energy Provider smart meter data

LDC Retail type or vacancy classification

LDC SmartStreetSensor footfall data

Otley BID Footfall Data

Registry Trust County Court Judgements

CGA 2017 Alcohol on Trade Outlet Index & Volume Estimations

Shop Direct transactions and products

Synthetic population data based on 2011 census

Appliances Online transactions

Callcredit CAMEO postcode directories

Youth Banking Card Provider – aggregated data

British Population Survey

Active Inspiration activity data

YouGov Survey Data

UK Women's Cohort Study

More data is available: https://data.cdrc.ac.uk