*Andy Brown*

*Data VISUALISATION*

*Assignment 1*

Author:

*Andy Brown*

16-05-2020

University of Buckingham

MSc Applied Data Science

L7 Apprenticeship

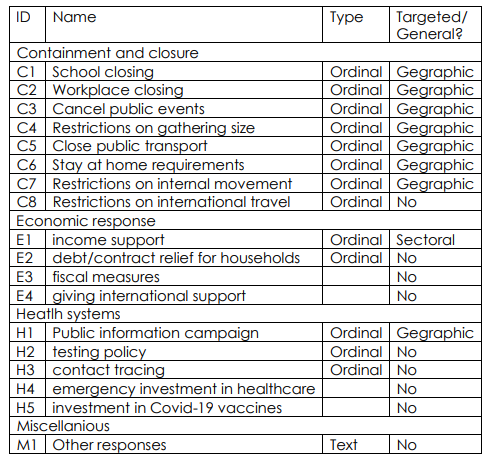
## Introduction

## Part 1 – Question1

Please see oxcgrt\_v9.py and associated output OxCGRT\_summary.xlsx

## Part 1- Question 2 – a

In the document <https://www.bsg.ox.ac.uk/sites/default/files/2020-05/BSG-WP-2020-032-v5.0_0.pdf>. The paper discusses the approach the approach to collecting stringency data an index providing a calculation/indication of the level of government response. The index is based on the metrics documented below and identifies general government policies in response to the Covid-19 virus pandemic. The metrics are a series of ordinal measures, judged/scored by the OxCGRT team using publicly available data. An algorithm is used to generate an overall index based on ordinal measures C1-8, E-14, H1-H5 and M1. It is possible that the required publicly available is missing and therefore it is impossible to make a judgement on one or more measures may be missing. To handle this the team assume that absence of a value for one metric e.g. C1 sets that ordinal value to zero and then take the mean of all the values.



This seems like a good scheme because the index should increase as more policies are put in place and therefore more ‘stringent’ government measures are put in place. In contrast if the mean based one the policies data could be scored of existing values were taken it would be possible to have one policy and get a score of 100. It would be hard to argue that one policy is a stringent regime however an index of 100 imply the opposite. Whilst this scheme is best of the two options documented. There are many other possibilities for example the mean or the last reported value. The mean would not seem like a good measure because this index is designed to show variation as new measures are introduced to curb the virus outbreak. An overall mean does not describe this instant in time. I would suggest a better method would be to use the last previously recorded value. This would seem the most likely score for current government policy because absence of data probably indicates lack of change in policy rather than a relaxing of policy.

To consider what measure may be most appropriate it’s worth considering what is being measured, publicly available of government policy response to the virus pandemic. I suggest it’s likely that when a new policy measure is added this has to be publicised to citizens to make sure this policy is followed. There may then be times when there is no published data about the policy however it seems unlikely that a policy would be switched on and off daily

## Part 1- Question 2 – b