



Institute  
and Faculty  
of Actuaries

# R modelling hackathon – Specification

April 2021

## **Objective**

The “R number” is a very popular metric used in connection with COVID-19 and other infectious diseases. The objective of this exercise is to perform analysis to illustrate the issues relating to the use of the R number. The specific output would be a demonstration of the potential pitfalls of using R numbers to draw conclusions and make policy decisions.

The judges will evaluate the submissions across 4 key areas:

1. Data collection, validation, adjustment
2. Analysis
3. Results
4. Communication

The exercise is not intended to be a specific forecast of R numbers, but rather a critical examination of the R number and the implications for policy decision-making relating to the COVID-19 pandemic. In particular, the judges are looking for a specific critical evaluation of the inherent assumptions of homogeneity in:

- Susceptibility
- Infection; and
- Virulence.

Examples of correct or incorrect evaluation of R numbers could be supplied to illustrate these particular issues.

## **Further detail**

The R number is the number of people that one person infected with COVID-19 infects. However, this R number:

- does not have a time period so does not give sufficient information to ascertain the number of concurrent infections
- does not technically differentiate between degrees of infection (asymptomatic, symptomatic, serious illness, critical illness, long illness, death).
- can vary by area, by time period, by human behaviour and by virus variant etc.
- is estimated from historical data and therefore by definition lags behind reality
- will have a variance around the true value in the normal statistical sense; and
- can be more volatile when the true value is low.

As a result of these disadvantages, policymakers need to be very cautious when making policy decisions based on estimated R numbers.

The aim of this hackathon is to demonstrate the variability of the underlying components of the R number and how these components can, in aggregate, lead to potentially misleading estimates of the R number.

The following links are to specific UK government data. The data is caveated by some of the discussion on R values and statistical considerations.

- <https://www.gov.uk/guidance/the-r-value-and-growth-rate>
- <https://www.gov.uk/guidance/coronavirus-covid-19-statistics-and-analysis>
- [https://coronavirus.data.gov.uk/?\\_ga=2.50738857.1094791912.1618849071-320608074.1618849071](https://coronavirus.data.gov.uk/?_ga=2.50738857.1094791912.1618849071-320608074.1618849071)

As an example, the following paper by the Actuaries Response Group illustrates one of the confounding aspects - namely super-spreaders.

- <https://www.covid-arg.com/post/overdispersion-how-superspreading-drives-the-pandemic>

Given the current events in India, studies of issues in that country would be of particular relevance. Some examples of data sources there would include the following free to use, and a reliable source for Confirmed, Deaths, Tests, Vaccinations (plus hospitalisations for some countries).

- <https://ourworldindata.org/covid-vaccinations>

Secondary data source, for variants, is outbreak.info:

- <https://outbreak.info/situation-reports>

This has a customisable analysis of the GISAID variant database. For India, last week it was straightforward to see the emergence of the B.1.617 variant. However, the data changed over the weekend; some counts of this variant was lost, and it now split over three variant reports: B.1.167, B.1.617.1 and B.1.167.2.

## Submission

Your answer should contain the following:

- references to any data used in the analysis (only publicly available data should be used)
- comments on the limitations of the data used
- analysis of the data to determine estimated R numbers
- theoretical models containing the main drivers of an R number
- illustration, by using these theoretical models, of the vast range of underlying circumstances possible to produce the following:
  - R number of 1
  - R number of 4.5
  - sudden change in R number due to new variant
  - impact of vaccines on the R number (across different age groups; susceptible groups; health groups etc)

- conclusions drawn from above analysis about the validity of publicly presented R numbers and the R number as a policy decision-making tool
- a communication of these findings aimed at an informed but non-expert, non-actuarial target audience. This communication will cover the conclusions derived from the analysis and the consequent concerns or otherwise of previously published conclusions in this area

Your submission should contain:

1. the data used
2. the models built and used
3. a report covering the above items, along with conclusions
4. a separate document of maximum 3 pages to cover the communication of the findings

The format of these should be Word or PDF for the documentation, Excel or CSV file for the data used, and the source code for the models actually built.

The judging panel will evaluate each submission on the following criteria:

1. The relevance of the data used and the appropriate caveats around the validity of the data used, including potential corrections required (10%)
2. The appropriateness of the modelling techniques used (15%)
3. The strength of support for the conclusions drawn based on the modelling of the data (25%)
4. The clarity and accuracy of the communication drafted to alert an interested audience to any material issues (50%)

Any decisions made by the judging panel will be final.

## FAQs

- 1. Does the data used need to be from the UK?*

No, any country or sub-population data can be used. We do however anticipate that many people will use UK data given the considerable amount of UK data available from reliable sources that has arisen over the past year. Due to current events we also believe that data for India would be of particular relevance.
- 2. Are there any restrictions on coding language/modelling style?*

No, although it is unlikely that Excel would be suitable for the task. The team will not be penalised according to the language or style used, but rather evaluated in terms of the appropriateness of the language/modelling for the analysis required.
- 3. Does the modelling need to be deterministic, scenario based or stochastic?*

There is no requirement around the nature of the analysis, but the analysis and its conclusion will be evaluated based on its completeness and robustness. For example, stochastic modelling can provide a significant amount of useful information, but the results can be relatively difficult to interpret and communicate. Similarly, deterministic methods would need to be carefully evaluated for sensitivity to key assumptions.
- 4. Does the output need to be in any particular format, e.g. Word, PDF etc?*

No specific format is required. However, the judges will be using standard computers across various platforms and thus PDF format files will be the most transportable across platforms and ensure limited ability for corruption.
- 5. Can additional questions and clarifications be answered?*

Yes, additional clarifications will be answered. We suggest you accumulate as many of these as you can at one time to facilitate comprehensive answering. All questions and the answers will be circulated to all entrants.



# Institute and Faculty of Actuaries

**DISCLAIMER** The views expressed in this publication are those of invited contributors and not necessarily those of the Institute and Faculty of Actuaries. The Institute and Faculty of Actuaries do not endorse any of the views stated, nor any claims or representations made in this publication and accept no responsibility or liability to any person for loss or damage suffered as a consequence of their placing reliance upon any view, claim or representation made in this publication. The information and expressions of opinion contained in this publication are not intended to be a comprehensive study, nor to provide actuarial advice or advice of any nature and should not be treated as a substitute for specific advice concerning individual situations. On no account may any part of this publication be reproduced without the written permission of the Institute and Faculty of Actuaries.

## **Beijing**

14F China World Office 1 · 1 Jianwai Avenue · Beijing · China 100004  
Tel: +86 (10) 6535 0248

## **Edinburgh**

Level 2 · Exchange Crescent · 7 Conference Square · Edinburgh · EH3 8RA  
Tel: +44 (0) 131 240 1300 · Fax: +44 (0) 131 240 1313

## **Hong Kong**

1803 Tower One · Lippo Centre · 89 Queensway · Hong Kong  
Tel: +852 2147 9418

## **London (registered office)**

7<sup>th</sup> Floor · Holborn Gate · 326-330 High Holborn · London · WC1V 7PP  
Tel: +44 (0) 20 7632 2100 · Fax: +44 (0) 20 7632 2111

## **Oxford**

1<sup>st</sup> Floor · Park Central · 40/41 Park End Street · Oxford · OX1 1JD  
Tel: +44 (0) 1865 268 200 · Fax: +44 (0) 1865 268 211

## **Singapore**

163 Tras Street · #07-05 Lian Huat Building · Singapore 079024  
Tel: +65 6717 2955

[www.actuaries.org.uk](http://www.actuaries.org.uk)

© 2017 Institute and Faculty of Actuaries