

# Once Upon A Time, When Epidemiology Met Data Science

Ahmend Hifaz	Alisa Ishikawa	Pirneeta Bhambhani
@hifaz.ahmed	@ishikawa.alisa	@bhambhani.pirneeta
0151 660 77919	0174 320 6012	0157 332 68112
OH: 11:00-16:00	OH: 10:30-14:30	OH: 10:00-15:00

January 5, 2020

## Abstract

R is a free, open-source programming language that is a powerful tool used by epidemiologists to combat viruses such as the one we are currently battling. It is useful to analyze the degree of the outbreak, visualize it, and estimate key statistics that play a vital role in public health, in terms of how and when we react to the outbreak. Through the utilization of data sets and related tools, the analysis on rate of growth, expected rate of growth, and the reproduction number(R number) could be estimated to construct ongoing reports, and aid in major decision making. The key argument here is that, “R helps support public health authorities in their “life-or-death” battle against the viruses.

## Learning Objectives

During this course you should:

- Understand the importance of R in large-scale data analysis
- Accumulate data from desired authorised sources to produce a customized analysis
- Explore alternative resources for COVID-19 data acquisition
- Utilising R as a tool to visualize COVID-19 with graphs and maps

## Topics Covered in the Course

**1 Why R is a significant tool of choice for outbreak epidemiologists**

**2 Data Acquisition**

2.1 Gathering all Covid related data through authentic sources

## 2.2 Strength of R as a statistical application for forecasting/predictions

### 3 Visualization Of COVID-19 Using R

#### 3.1 Overview of plots and ggplots

#### 3.2 Visualization using maps and line graphs

#### 3.3 DIY: Visualization Exercise

### Literature Resources

<https://arxiv.org/pdf/2009.01091.pdf>

<https://www.emeraldgrouppublishing.com/journal/idd/using-data-science-understand-coronavirus-pandemic>

<https://news.harvard.edu/gazette/story/2020/09/harvard-journal-keeps-data-scientists-connected-during-covid/>

<https://www.repidemicsconsortium.org/projects/> <https://medium.com/swlh/a-simple-way-to-gather-all-coronavirus-related-data-with-r-b1e7ecb74346>

<https://rviews.rstudio.com/2020/03/05/covid-19-epidemiology-with-r/>

<https://towardsdatascience.com/visualize-the-pandemic-with-r-covid-19-c3443de3b4e4>

<https://infographics.channelnewsasia.com/covid-19/map.html>

### GitHub Resources

<https://github.com/aqualisa/HSF-MIBM2-COVID19.git>

<https://github.com/CSSEGISandData/COVID-19>

<https://geodacenter.github.io/covid/>