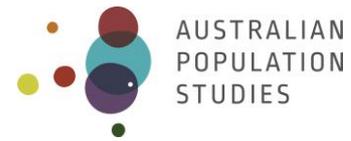

Introductory Guide



How to create an interactive dashboard using R: the example of the Queensland COVID-19 tracker

Julia Loginova* The University of Queensland

Pia Wohland The University of Queensland

* Corresponding author. Email: j.loginova@uq.edu.au. Queensland Centre for Population Research, School of Earth and Environmental Sciences, Chamberlain Building, The University of Queensland, St Lucia, Qld 4072, Australia

Paper received 15 September 2020; accepted 2 November 2020; published 16 November 2020

Abstract

Background

Interactive tools like data dashboards enable users both to view and interact with data. In today's data-driven environment it is a priority for researchers and practitioners alike to be able to develop interactive data visualisation tools easily and where possible at a low cost.

Aims

Here, we provide a guide on how to develop and create an interactive online data dashboard in R, using the COVID-19 tracker for Health and Hospital Regions in Queensland, Australia as an example. We detail a series of steps and explain choices made to design, develop, and easily maintain the dashboard and publish it online.

Data and methods

The dashboard visualises publicly available data from the Queensland Health web page. We used the programming language R and its free software environment. The dashboard webpage is hosted publicly on GitHub Pages updated via GitHub Desktop.

Results

Our interactive dashboard is available at <https://qcpr.github.io/>.

Conclusions

Interactive dashboards have many applications such as dissemination of research and other data. This guide and the supplementary material can be adjusted to develop a new dashboard for a different set of data and needs.

Key words

Interactive dashboards; R; data visualisation; COVID-19; Queensland

1. Introduction

Researchers and practitioners increasingly rely on interactive visualisation tools to draw insights from their data, improve strategic and operational decision-making and to disseminate information to stakeholders and the wider public. Effective interaction with data can substantially enhance the understanding of underlying trends and patterns. One way to achieve this is by presenting data in a dashboard.

Dashboards are “a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance.” (Few 2006, p. 12). Dashboards are attractive to users as they amplify cognition and capitalise on human perceptual capabilities. They can be broadly categorised as interactive or static.

Interactive dashboards represent a new trend of data intelligence that enable not only viewing but also interacting with data through tables, charts, maps and text. These are different from static dashboards that are “read-only” (Yigitbasioglu and Velcu 2012). Interactivity is especially useful in presenting and interpreting large complex data, as it assists in summarising data and alleviating information overload.

Different techniques and software can be used to present information and provide interactivity (Smith 2013). The possible options are ArcGIS Dashboards, Tableau, as well as numerous tools developed for the purposes of data intelligence. Although providing numerous advantages and appealing aesthetics, they might not be well integrated with the process and outputs of research which is performed in R, a widely used software environment for statistical computing, data science, and graphics. R is often used in conjunction with the integrated development environment RStudio. Both are available as free and open-source editions with a range of packages and exemplary support. The following section explains the process of development of the Queensland COVID-19 tracker in R.

2. Queensland COVID-19 tracker in R

As the COVID-19 pandemic started to spread around the world and Australia, dashboards that visualise a large amount of rapidly changing information have become useful for the general public and government institutions. Examples include the dashboard by the World Health Organisation (WHO 2020) and COVID-19 Map coordinated by Johns Hopkins Coronavirus Resource Center (Dong et al. 2020). Numerous interactive web-dashboards have been developed to track the spread of COVID-19 in Australia, for example, COVID-19 dashboard developed by Government of South Australia (2020), New South Wales COVID-19 Cases and Community Profile by the University of Sydney (2020), Charting the COVID-19 spread in Australia by the Australian Broadcasting Corporation (ABC 2020), and COVID LIVE (2020) developed by volunteers. Covering data at the level of states and territories, these dashboards initially did not provide information related to the spread of COVID-19 on a smaller geography level. We aimed at filling this gap through the development of the COVID-19 tracker focused on Queensland regions. To this end, we wanted to effectively utilise publically available information from Queensland Health (2020a), a ministerial department of the Queensland Government, and provide interactive analytical features to track the spread of COVID-19 throughout time and space.

2.1. Setup

Several R packages provide opportunities for the development of interactive dashboards. The Queensland COVID-19 tracker has been developed using the R package ‘flexdashboard’ (Iannone et al. 2020). It provides an effective solution for developing and publishing a group of related data visualisations as a dashboard. Another package that could be used is ‘shinydashboard’ (Chang and Borges 2018). It is more suitable for developing dynamic dashboards where input interacts with the output requiring a server to execute R code on user input. ‘Flexdashboard’ operates as an R Markdown framework allowing the dashboard to be ‘knitted’ into a dynamic but independent HTML.

The complete and annotated code in R Markdown document “QueenslandCOVID19tracker.Rmd” and associated files to create the Queensland COVID-19 tracker are published online (Loginova and Wohland 2020). As demonstrated in this document, the first step in setting up the dashboard is to add the YAML code as the header of the dashboard, including title and type of output file which in this case is HTML (Box 1). The rest of the document contains code to specify the outline of the dashboard and numerous code chunks. A code chunk is a piece of R code that start and end with ``` and can be named. In the code chunk ‘setup’, we assign main directory (the folder where “QueenslandCOVID19tracker.Rmd” and data files were downloaded), install and load packages, and set global options that affect how R displays the results.

It is important to stress that versions of R, R Studio and R packages may affect the execution of the code. Therefore, control of R and packages is necessary to ensure the stability of the dashboard over time. To this end, we indicated that the code for the Queensland COVID-19 tracker was developed and works in R version 4.0.0 and versions of each package are specified in “QueenslandCOVID19tracker.Rmd”.

```
---  
title: "COVID-19 IN QUEENSLAND, AUSTRALIA"  
output: flexdashboard::flex_dashboard  
---
```

Box 1: The YAML code specifying the header of the dashboard

2.2. Data updates and calculations

Queensland Health has published summary data for 16 Queensland Health and Hospital regions on a daily basis since the first cases of COVID-19 were identified in Queensland on January 21st 2020 (Queensland Health 2020b). We started gathering data from March 16th 2020, at the point when the number of new cases had begun to increase rapidly. As this paper went to press in November 2020, there were 1,177 confirmed cases of COVID-19 in Queensland. The data for Confirmed, New Confirmed, Active and Recovered cases as well as Deaths were downloaded daily from the

Queensland Health website (Queensland Health 2020a) and combined in a .csv (comma-separated values) file. As the situation evolved, additional information has been included in the reporting, including the number of active self-quarantine notices.

Data updates, preparation, and calculations are performed in a single R Markdown file. In the code chunk 'updates', we load updated data into the R environment, specify values to update, and add updated daily data to existing data files. In 'data preparation', we format the data to be used in maps and graphs. In 'calculations', we obtain values to be displayed in the dashboard for Queensland totals and doubling time (the time it takes for confirmed cases to double in value). It is important to stress that these three code chunks contain 'include = FALSE' condition to prevent R from showing the code and its results in the dashboard

2.3. Layout, design and components

The layout of the dashboard is specified in the R Markdown document using 'flexdashboard' syntax, as shown in Box 2. The layout of the Queensland COVID-19 trackers includes four pages organised according to themes, including 'Situation Today', 'Change Over Time', 'Individual Regions', and 'About'.

The first three pages contain a collection of tabsets to visualise information relevant to each page (Figure 1). The layout of each page can be column-based or row-based, allowing control of the width of columns. The colour of the main panel (dark purple) has been assigned at the beginning of the R Markdown document using CSS (Cascading Style Sheets) easily understood by HTML (see code for <style> in "QueenslandCOVID19tracker.Rmd").

Each tabset consists of components including charts, maps, tables, text, images, and value boxes. Our dashboard contains the following components that summarise COVID-19 cases in Queensland, their spatial distribution and temporal change:

- 8 value boxes;
- geographical map with 6 layers using 'tmap' and 'leaflet' packages and 1 animated map using 'tmap' and 'ImageMagick';
- 3 summary tables using 'DT' package;
- 18 graphs using 'ggplot2' and 'plotly' packages;
- 2 texts tabs.

Each of these components was produced in individual code chunks, as outlined in "QueenslandCOVID19tracker.Rmd".

```

`r Sit.text` {} # Page 1
===== # PAGE 1

Column {data-width=175} # column for value boxes
-----

Column {.tabset} # tabset for Geographical Map, Summary Table, and Instructions
-----

Change Over Time {} # Page 2
===== # PAGE 2

Column {.tabset}
----- # tabset for Queensland, HHS Regions, Active cases, Recovered cases

Column {data-width=250}
----- # column for animated map

Individual Regions {} # Page 3
===== # PAGE 3

Column {.tabset}
----- # tabset for 15 individual regions

About {} # Page 4
===== # PAGE 4

```

Box 2: The layout of the components of the Queensland COVID-19 tracker

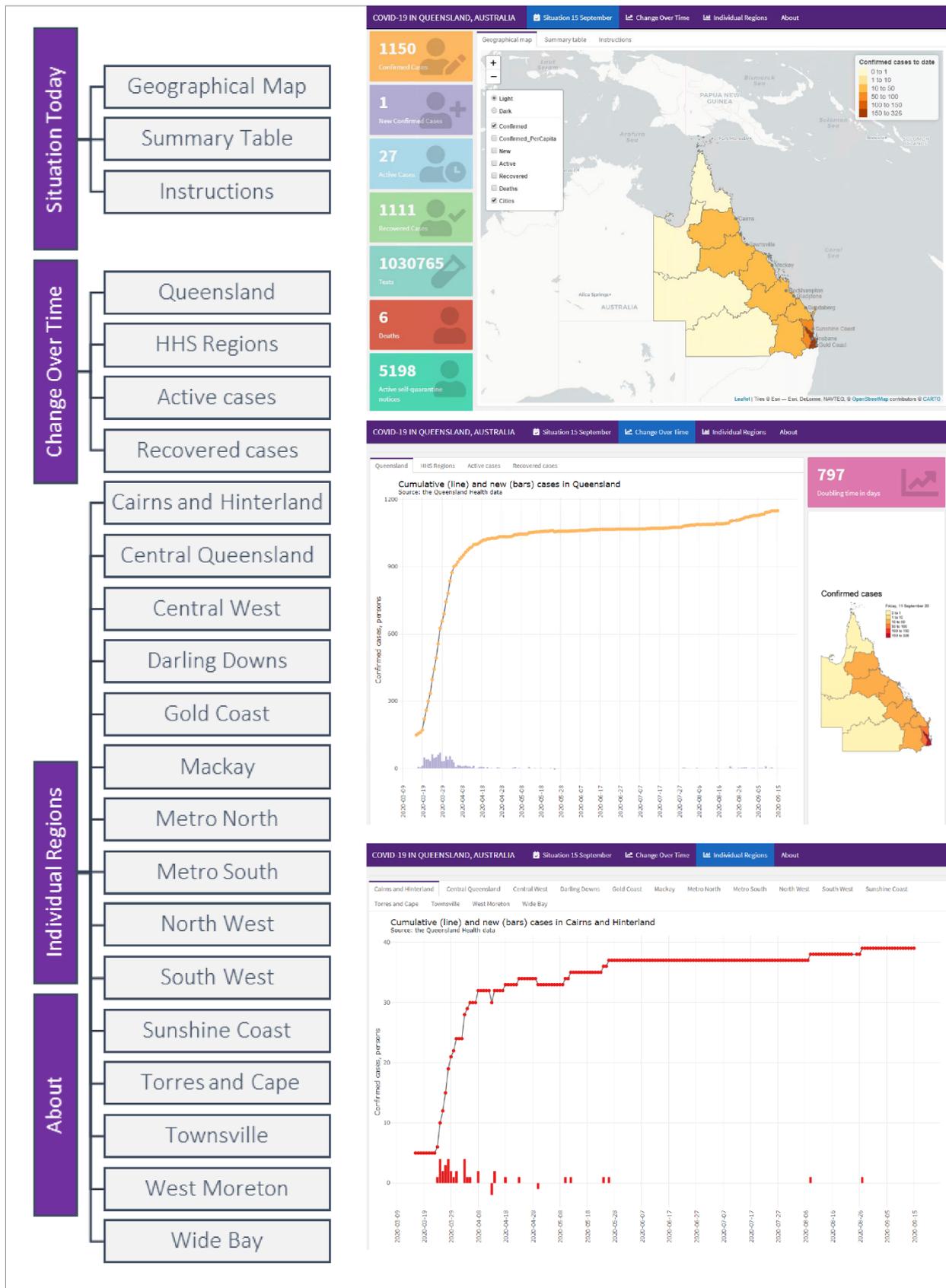


Figure 1: The layout of the dashboard and snapshots of three main pages

Source: <https://qcpr.github.io/>

2.4. Knitting and hosting

The finished .Rmd file needs to be 'knitted' into an HTML document. Rendering executes all of the R code chunks and creates the interactive dashboard. The final step is deploying it online. The decision in relation to publishing a dashboard online is required early in the dashboard development process. GitHub Pages provides a suitable environment for continuous integration and serverless deployment of dashboards through GitHub, a repository hosting service (Utomo and Falahah 2020). Web hosting on GitHub Pages has several advantages as it is free, has collaborative support from the community of users, and is easy to use (see Hanson and Seeger 2017). The deployment requires an account and a repository on GitHub and an installed git component on a computer, such as GitHub Desktop. However, it is not necessary to have data and .Rmd file included the repository, they can all be kept locally as done in this case. Figure 2 summarises the process from the dashboard development to hosting it online.

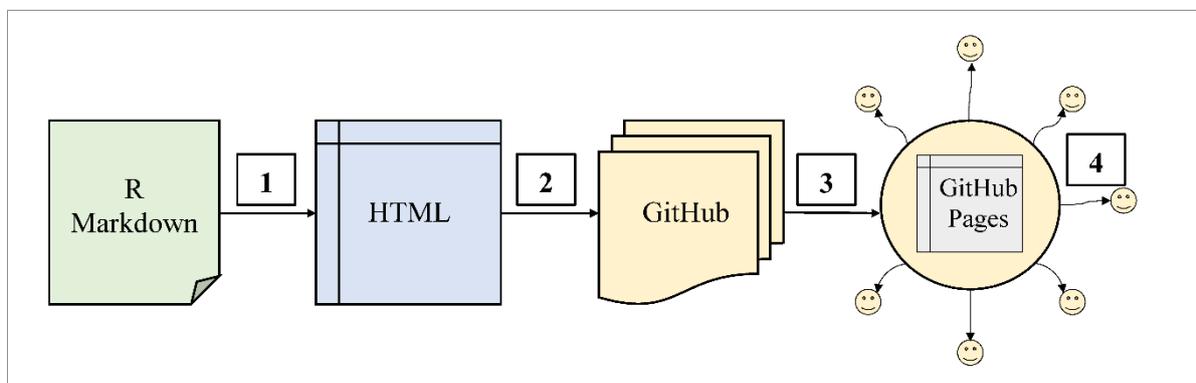


Figure 2: The process of building a dashboard in R and its deployment of GitHub Pages

Notes: 1. The HTML is generated using R Markdown framework in R. 2. The HTML website is uploaded into GitHub using GitHub Desktop. 3. The HTML is committed to GitHub Pages. 4. Users can access the interactive dashboard hosted in GitHub Pages.

3. Conclusions

The aim of this introductory guide was to make it easier for researchers and other practitioners to incorporate data dashboards as a visual data communication tool into their work. Today, tools are available in R to develop dashboards and deploy them online easily. Interactive dashboards allow presenting data in a way that is informative and visually engaging. Furthermore, dashboards can be used for data collections, validation of research results, interactive educational activities, and research collaborations. Underlying R codes can be easily shared supporting reproducibility and accountability in research and other areas. However, developing effective and stable dashboards requires building visualisation competence and further research is needed about benefits that justify their adoption.

4. Resources

The complete code and associated files to create the dashboard can be accessed at

<https://doi.org/10.6084/m9.figshare.12957065.v1>.

Key messages

- Interactive dashboards are great data visualisation and communication tools.
- Researchers and practitioners can easily develop interactive dashboards using R, a widely used software environment for statistical computing.
- GitHub Pages enable continuous integration and deployment of dashboards through GitHub, a repository hosting service.
- In this guide, we provide the complete R code that can be easily adjusted to suit specific needs.

Acknowledgements

We thank editors and reviewers for their valuable feedback. We also acknowledge input from Seraphim Alvanides, Alexander Smith, Paul Dennis and QCPR's colleagues in relation to the initial design of the dashboard.

References

- ABC (2020) Charting the COVID-19 spread in Australia. Available at: <https://www.abc.net.au/news/2020-03-17/coronavirus-cases-data-reveals-how-covid-19-spreads-in-australia/12060704?nw=0>. Accessed on 30 October 2020.
- Chang W and Borges B (2018) Shinydashboard: create dashboards with 'Shiny'. <https://rstudio.github.io/shinydashboard>. Accessed on 30 October 2020.
- COVID LIVE (2020) Available at: <https://covidlive.com.au/>. Accessed on 30 October 2020.
- Dong E, Du H and Gardner L (2020) An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* 20(5): 533-534.
- Few S (2006) *Information Dashboard Design: The Effective Visual Communication of Data*. Sebastopol, CA: O'Reilly Media.
- Government of South Australia (2020) COVID-19 dashboard. Available at: <https://www.covid-19.sa.gov.au/home/dashboard>. Accessed on 30 October 2020.
- Hanson B and Seeger C J (2017) Creating GitHub Pages. GIS and Mapping Task Sheets. 4. https://lib.dr.iastate.edu/gis_tasksheets/4. Accessed on 30 October 2020.
- Iannone R, Allaire J J, and Borges B (2020) Flexdashboard: R Markdown format for flexible dashboards. <http://rmarkdown.rstudio.com/flexdashboard>. Accessed on 30 October 2020.
- Loginova J and Wohland P (2020) Queensland COVID19 tracker. Figshare. <https://doi.org/10.6084/m9.figshare.12957065.v1>. Accessed on 30 October 2020.
- Pauwels K, Ambler T, Clark B H, LaPointe P, Reibstein D, Skiera B, Wierenga B, and Wiesel T (2009) Dashboards as a service: why, what, how, and what research is needed? *Journal of Service Research* 12(2): 175-189.
- Queensland Health (2020a) Queensland COVID-19 statistics. Available at: <https://www.covid19.qld.gov.au/>. Accessed on 30 October 2020.

- Queensland Health (2020b) Queensland coronavirus update. Available at: <https://www.health.qld.gov.au/news-events/doh-media-releases/releases/queensland-coronavirus-update-290120>. Accessed on 30 October 2020.
- Smith V S (2013) Data dashboard as evaluation and research communication tool. *New Directions for Evaluation* 140: 21-45.
- University of Sydney (2020) NSW COVID-19 Cases and Community Profile. Available at: <https://covid19-data.sydney.edu.au/>
- Utomo P and Falahah S (2020) Building Serverless Website on GitHub Pages. In IOP Conference Series: Materials Science and Engineering 879(1): 012077. IOP Publishing.
- WHO (2020) Coronavirus Disease (COVID-19) Dashboard. Available at: <https://covid19.who.int/>. Accessed on 30 October 2020.
- Yigitbasioglu O M and Velcu O (2012) A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems* 13(1): 41-59.