Dashboards project

Richard White

2018-10-18

Contents

List of Tables

List of Figures

1 Introduction

1.1 Executive summary

The dashboards project is a project at FHI concerned with running automated analyses on data.

In principle, the dashboards project is split up into three parts:

1. The overarching infrastructure (i.e. Docker containers, continuous integration, chron jobs, etc.)
2. The R package for each automated analysis
3. The R executable for each automated analysis

1.2 What is an automated analysis?

An automated analysis is any analysis that:

1. Will be repeated multiple times in the future
2. Always has an input dataset with consistent file structure
3. Always has the same expected output (e.g. tables, graphs, reports)

1.3 Why not have one project for each automated analysis?

Automated analyses have a lot of code and infrastructure in common.

Automated analyses:

1. Need their code to be tested via unit testing to ensure the results are correct
2. Need their code to be tested via integration testing to ensure everything runs
3. Need to be run at certain times
4. Need to be able to send emails notifying people that the analyses have finished running
5. Need to make their results accessible to the relevant people

By combining them all in one umbrella project we can force everyone to use the same infrastructure, so we:

1. Only need to solve a problem once
2. Only need to maintain one system
3. Can easily work on multiple projects, as we all speak the same language

1.4 Important repositories

1.4.1 Infrastructure

https://github.com/raubreywhite/dashboards_control/ (private)

This contains the Docker files, cronfiles, all bash scripts, etc.

https://folkehelseinstituttet.github.io/dashboards/ (this one)

This contains the R executable for each automated analysis.

https://folkehelseinstituttet.github.io/fhi/

This is an R package that contains helper functions.

1.4.2 Automated analyses

https://folkehelseinstituttet.github.io/dashboards_sykdomspuls/

https://folkehelseinstituttet.github.io/dashboards_normomo/

https://folkehelseinstituttet.github.io/dashboards_sykdomspuls_pdf/

https://folkehelseinstituttet.github.io/dashboards_sykdomspuls_log/

2 Executable for each automated analysis

Each automated analysis has its own R package (e.g. sykdomspuls).

Each R package should contain 99%

3 R packages

Each automated analysis has its own R package (e.g. sykdomspuls).

Each R package should contain 99%

4 Introduction

4.1 Executive summary

The dashboards project is a project at FHI concerned with running automated analyses on data.

In principle, the dashboards project is split up into three parts:

1. The overarching infrastructure (i.e. Docker containers, continuous integration, chron jobs, etc.)
2. The R package for each automated analysis
3. The executable for each automated analysis

4.2 What is an automated analysis?

An automated analysis is any analysis that:

1. Will be repeated multiple times in the future
2. Always has an input dataset with consistent file structure
3. Always has the same expected output (e.g. tables, graphs, reports)

4.3 Why not have one project for each automated analysis?

Automated analyses have a lot of code and infrastructure in common.

Automated analyses:

1. Need their code to be tested via unit testing to ensure the results are correct
2. Need their code to be tested via integration testing to ensure everything runs
3. Need to be run at certain times
4. Need to be able to send emails notifying people that the analyses have finished running
5. Need to make their results accessible to the relevant people

By combining them all in one umbrella project we can force everyone to use the same infrastructure, so we:

1. Only need to solve a problem once
2. Only need to maintain one system
3. Can easily work on multiple projects, as we all speak the same language