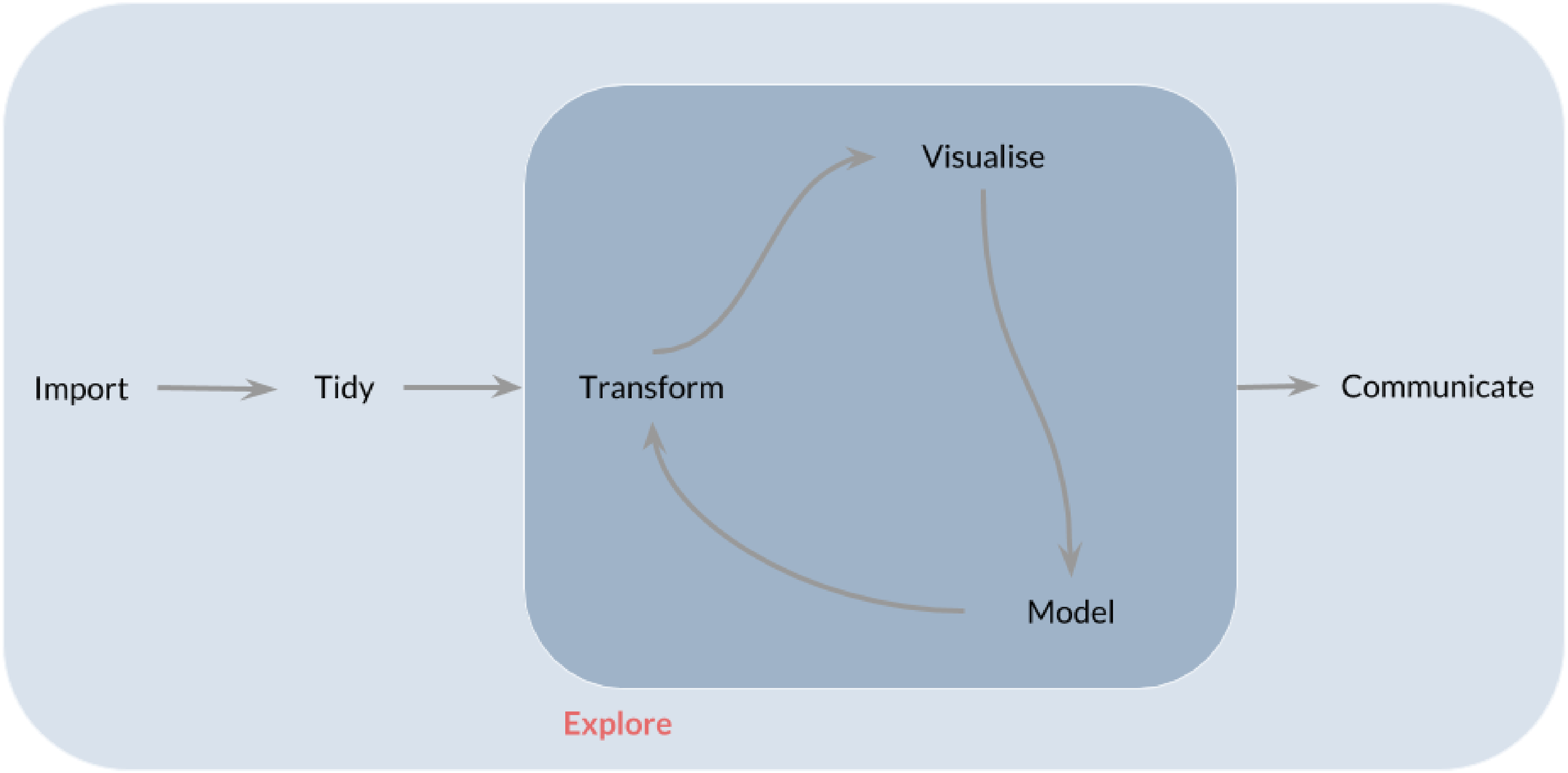
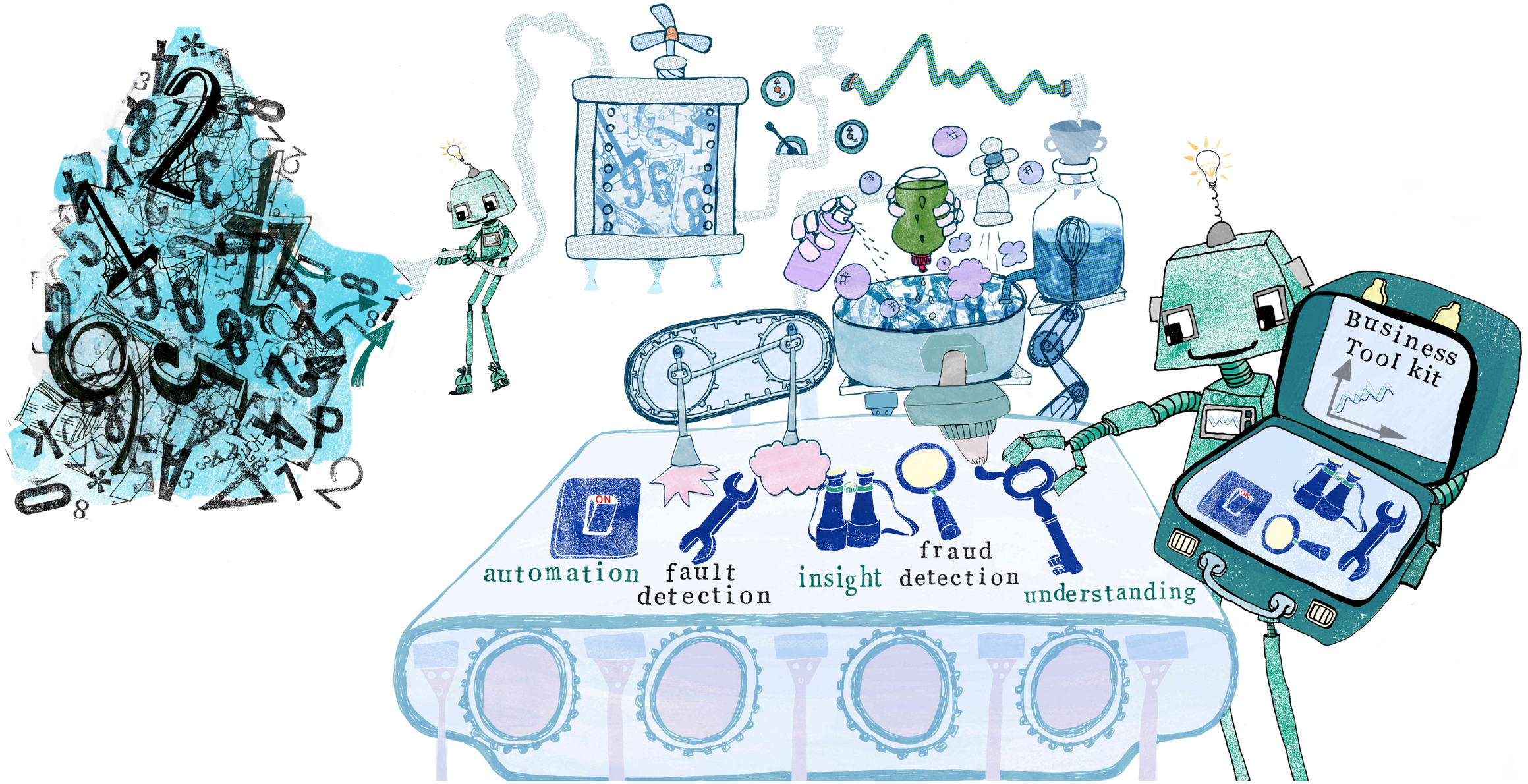
Defensive R Programming

DEFENSIVE R PROGRAMMING

Dr. Colin Gillespie
Jumping Rivers


Visualise
Import
Tidy
Transform
Communicate
Model
Explore


Business
Tool kit
ON
automation
fault
detection
insight
fraud
detection
understanding

Goals

- Make our pipeline robust
- Avoid problems where possible
- Be defensive where ever possible

> The supreme art of war is to subdue the enemy without fighting
>
> Sun Tzu, **The Art of War**

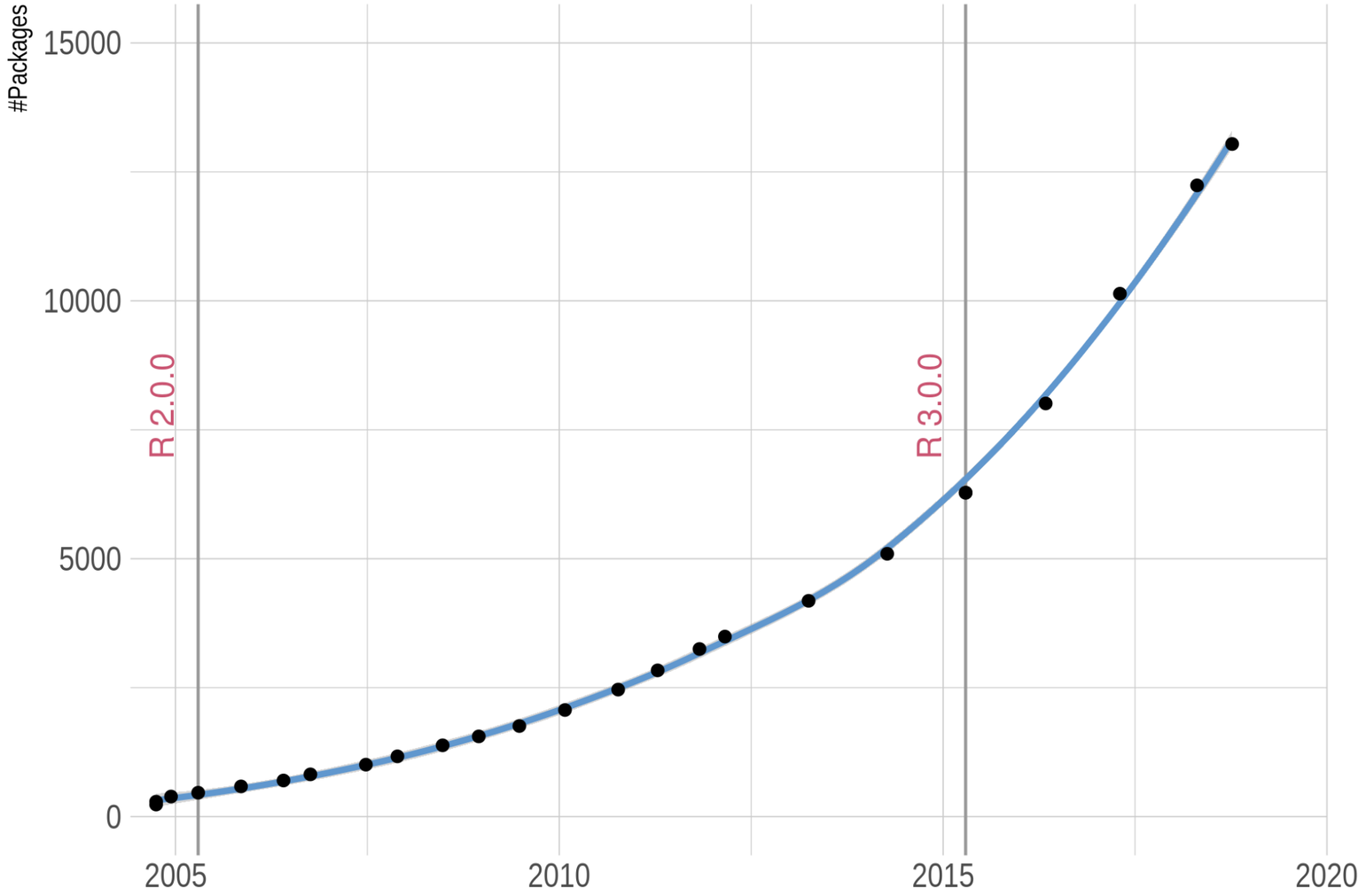
Let's go

DEFENSIVE R PROGRAMMING

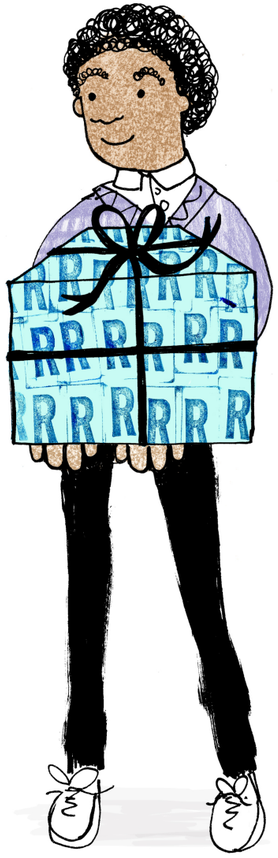
Don't reinvent the wheel/package

DEFENSIVE R PROGRAMMING

Dr. Colin Gillespie
Jumping Rivers


#Packages
15000
10000
5000
0
R 2.0.0
R 3.0.0
2005
2010
2015
2020

Choosing an R package



Is the package

- mature?
- actively developed?
- well documented?
- well used?

Context matters

Where are you using the package?

- Messing about on a toy project
- Production server

CRAN Task Views

CRAN task views aim to provide some guidance which packages on CRAN are relevant for tasks related to a certain topic. They give a brief overview of the included packages and can be automatically installed using the ctv package. The views are intended to have a sharp focus so that it is sufficiently clear which packages should be included (or excluded) - and they are *not* meant to endorse the "best" packages for a given task.

- To automatically install the views, the ctv package needs to be installed, e.g., via
  `install.packages("ctv")`
  and then the views can be installed via `install.views` or `update.views` (where the latter only installs those packages are not installed and up-to-date), e.g.,
  `ctv::install.views("Econometrics")`
  `ctv::update.views("Econometrics")`
- The task views are maintained by volunteers. You can help them by suggesting packages that should be included in their task views. The contact e-mail addresses are listed on the individual task view pages.
- For general concerns regarding task views contact the ctv package maintainer.

Topics

Bayesian	Bayesian Inference
ChemPhys	Chemometrics and Computational Physics
ClinicalTrials	Clinical Trial Design, Monitoring, and Analysis
Cluster	Cluster Analysis & Finite Mixture Models
DifferentialEquations	Differential Equations
Distributions	Probability Distributions
Econometrics	Econometrics

Let's get to work

DEFENSIVE R PROGRAMMING

Packages and Namespaces

DEFENSIVE R PROGRAMMING

Dr. Colin Gillespie
Jumping Rivers

Your global environment

- The default environment is the `.GlobalEnv`
- Objects are stored in environments
- To view the contents, use `ls()`
- This environment quickly fills up
- We can start to mixing up objects

```
n
# or did I want
N
```

Packages and environments

- Packages use *namespaces* as spaces for names
  - You can think of a namespace as a box that contains the package
- A namespace helps keep things tidy
  - Sort of like folders

Exported functions

- `library()` gives you direct access to this package box
- `library("dplyr")` gives you direct access to the exported functions in **dplyr**

```
## 238 exported functions
getNamespaceExports("dplyr")
```

- Alternatively, we can use `::` to directly access a function

```
## The filter function from dplyr
dplyr::filter
```

Great minds think alike

- Sometimes package authors come with the same function name

```
## The filter() function
stats::filter()
dplyr::filter()
```

- If I type `filter()` which version do I get?
- It depends on the package load order!

Great minds think alike

- `search()` to the rescue

```
search()
```

```
# [1] ".GlobalEnv"        "package:dplyr"     "package:stats"
# <Other packages>
```

Be defensive

Namespace clashes are more than just a pain

- They can lead to hard to diagnose bugs

The **conflicted** package tries to make your function choice *explicit*

```
library("conflicted")
library("dplyr")
```

```
filter
```

```
#Error: [conflicted] `filter` found in 2 packages.
#Either pick the one you want with `::`
# * dplyr::filter
# * stats::filter
#Or declare a preference with `conflict_prefer()`
# * conflict_prefer("filter", "dplyr")
# * conflict_prefer("filter", "stats")
```

Time to wake up

DEFENSIVE R PROGRAMMING