Dynamic Documents with R and knitr

Leonel Muñoz Cedano leoneling@gmail.com

Universidad Distrital Francisco José de Caldas Maestría en Ciencias de la información y las comunicaciones Tendencias en Ingeniería de Software Bogotá DC, 16 de Abril 2016





Agenda

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

taken di saktala

Albout the Autho

Inotallatio

Example Knitr - co

Example Knitr - res

Reproducib

Literature

Good and Bad Practice

Barrie

A First Lo

Setu

Quick Reportin

Feedback

Questions, Comments a

Contact Information

Bibliograph

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC. 16 de Abril Introduction

About the Author

Installation

Example Knitr - code

Example Knitr - result

Reproducible Research

Literature

Good and Bad Practices

Barriers

A First Look

Setup

Minimal Examples

Quick Reporting

Extracting R Code

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliography



Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Ar

Inetallation

Example Knitr - cod

Reproducib

Research

Literature

Good and Bad Practic

Barrie

A First Look

Setup

Minimal Examp

Extracting B (

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliograph

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC. 16 de Abril ► The basic idea behind dynamic documents stems from literate programming, a programming paradigm conceived by Donald Knuth (Knuth, 1984).



Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Ai

Inetallation

Example Knitr - co

Reproducib

Research

Literature

D .

A First Look

Setup

Outsto December

Extracting B.C.

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliograph

- ► The basic idea behind dynamic documents stems from literate programming, a programming paradigm conceived by Donald Knuth (Knuth, 1984).
- Web scripting language such as PHP (which can embed program code into HTML documents).



Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the A

Installation

Example Knitr - co

Reproducib

11030410

Good and Bad Practic

Barriers

A First Loc

Setup

Quick Reportin

Extracting R C

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliograph

- ► The basic idea behind dynamic documents stems from literate programming, a programming paradigm conceived by Donald Knuth (Knuth, 1984).
- Web scripting language such as PHP (which can embed program code into HTML documents).
- Easy manage and maintain for the author.



Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the A

Installation

Example Knitr - co

Reproducit

Research

Literature

Barriers

A First Loc

Setup

Outst. Demonstr

Extracting R Co

Feedback Questions, Comments ar

Questions, Comments and Suggestions

Ribliograph

- ► The basic idea behind dynamic documents stems from literate programming, a programming paradigm conceived by Donald Knuth (Knuth, 1984).
- Web scripting language such as PHP (which can embed program code into HTML documents).
- Easy manage and maintain for the author.
- ► That is why we need the literate programming paradigm. In this paradigm, an author has two tasks.



Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the A

Inetallatio

Example Knitr - cod

Reproducib

nesearci

Cood and Rad Broatia

Rarrie

A First Loc

Setup

Quick Reportin

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliograph

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril Technically, literate programming involves three steps:

- Parse the source document and separate code from narratives.
- Execute source code and return results.
- Mix results from the source code with the original narratives.



Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introductio

About the Author

Installatio

Example Knitr - coo

Reproducible

Research

Literature

Barrie

A First Look

Setup

Minimal Examp

Quick Reporti

Feedback

Questions, Comments and

Contact Information

Bibliography

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril ▶ Who is Yihui Xie.



Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Author

Installatio

Example Knitr - coo

Reproducib

Research

Literature

Good and Bad Practic

Barrier

A First Look

Setup

Minimal Examp

Quick Reporting

Feedback

Questions, Comments and

Contact Information

Bibliography

- Who is Yihui Xie.
- Won the 2009 John M. Chambers Statistical Software Award (ASA).



Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Author

Installatio

Example Knitr - coo

Reproducib

Research

Literature

Barrie

A First Look

Setup

Ouisk Reporting

Extracting B (

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliograph

- ▶ Who is Yihui Xie.
- Won the 2009 John M. Chambers Statistical Software Award (ASA).
- ► Honorable Mention prize in the Applications of R in Business Contest 2012.



Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Author

Inetallatio

Example Knitr - co

Reproducib

Research

Literature

Rarrie

A First Loc

Setup

Quick Reporting

Feedback

Questions, Comments and

Contact Information

Ribliograph

- Who is Yihui Xie.
 - Won the 2009 John M. Chambers Statistical Software Award (ASA).
 - ► Honorable Mention prize in the Applications of R in Business Contest 2012.
- ► He founded the "Capital of Statistics".



Installation

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

Installation

Example Knitr - co

Reproducib

Research

Literature

Barrier

A First Look

Setup

Quick Reportin

Extracting R Co

Feedback

Questions, Comments and Suggestions

Contact Informatio

Bibliograph

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril Installation requirements and configurations needid:

- 1. install R x64 (version 3.2.3 or top)
- 2. install RStudio (version 0.96.331 or top)
- 3. install packages "knitr" in R commander
- 4. install TexStudio with MiKTeX
- 5. Integration RStudio with TexStudio



Installation Example Knitr - code

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Example Knitr - code

Feedback

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril

Example document *.Rnw in TexStudio

\documentclass{article}

\usepackage[utf8]{inputenc}

\begin{document}

Código de LATEX normal, puede utilizar cualquier plantilla.

«»=

Crea una secuencia de números

= 2:50

Muestra las medidas estadisticas basicas

summary(X)

\end{document}



Installation Example Knitr - result

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Author

Installation

Example Knitr - cod

Example Knitr - result

Reproducible

Research

Literature

Good and Rad Practice

Barrie

A First Look

Setup

Minimal Examples

Extracting B Co

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliography

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril

Result of the example above

Código de La rormal, puede utilizar cualquier plantilla.

```
# Crea una secuencia de números
X = 2:50
```

A - 2.50

Muestra las medidas estadisticas basicas

summary(X)

##	${\tt Min.}$	1st	Qu.	Median	Mean	3rd	Qu.	Max
##	2		14	26	26		38	50



Reproducible Research

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Auti

Inetallation

Example Knitr - cod

Example Knitr - resi

Reproducible Research

Literature

Good and Bad Practices

A First Look

Setup

Minimal Examp

Extracting B (

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliograph

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril Characteristics document of Reproducible research:

Results from scientific research have to be reproducible to be trustworthy.



Reproducible Research

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Auti

Inetallatio

Example Knitr - cod

Example Knitr - res

Reproducible Research

Literature
Good and Bad Practice

A First Look

Setup

Quick Reportin

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliograph

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril

Characteristics document of Reproducible research:

- Results from scientific research have to be reproducible to be trustworthy.
- Reproducible research (RR) is one possible by-product of dynamic documents.



Reproducible Research Literature

Dynamic Documents with R and knitr Leonel Muñoz Cedano

Literature

A First Look

Feedback

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril

 Proposed by Jon Claerbout at Stanford University (Fomel and Claerbout, 2009).



Reproducible Research

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Auti

Installation

Example Knitr - co

Reproducib

Literature

Good and Bad Practice

Barriers

A First Lo

Setup

Minimai Exampi

Extracting B.C.

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliograph

- ▶ Proposed by Jon Claerbout at Stanford University (Fomel and Claerbout, 2009).
- Sweave (Leisch, 2002) was among the first implementations for dealing with dynamic documents in R.



Reproducible Research

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Auth

Installation

Example Knitr - co

Example Killi - res

Research

Literature

D :

Δ Firet Lo

AT IISE LO

Minimal Examp

Quick Reporting

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliograph

- Proposed by Jon Claerbout at Stanford University (Fomel and Claerbout, 2009).
- Sweave (Leisch, 2002) was among the first implementations for dealing with dynamic documents in R.
- ► The knitr package (Xie, 2015).



Reproducible Research Good and Bad Practices

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Aut

Installatio

Example Knitr - c

Example Knitr - re

Research

Good and Rad Practices

GOOD AND DAD I TACILO

Δ First Loc

ATTION LOC

Minimal F

Quick Reportin

Feedback

Questions, Comments and Suggestions

Contact Information

Bibliograph

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril The key to keep in mind for RR is that other people should be able to reproduce our results. Some good practices for RR follow:

- Manage all source files under the same directory and use relative paths whenever possible.
- ▶ Do not change the working directory after the computing has started: function setwd().
- ► Compile the documents in a clean R session.
- Avoid the commands that require human interaction.
- ► Avoid environment variables for data analysis.
- Attach session Info and instructions on how to compile this document.



Reproducible Research Barriers

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Barriers

Feedback

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril

Despite all the advantages of RR, there are some practical barriers, and here is a non-exhaustive list:

- ► The data can be huge.
- Confidentiality of data.
- Software version and configuration.
- Competition.



A First Look

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Auti

Installation

Example Knitr - coo

Reproducible

Research

Literature

Barrier

A First Loo

Setup

Quick Reporting

Feedback

Questions, Comments and

Contact Information

Ribliography

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril ➤ Since knitr is an R package, it can be installed from CRAN in the usual way in R:
install.packages("knitr", dependencies = TRUE)

► If you have any problems with knitr, the first thing to check is its version:

packageVersion("knitr")
if not the latest version, run
update.packages()

➤ Once we have knitr installed library(knitr) knit("your-file.Rnw")



A First Look Minimal Examples

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introductio

About the Au

Installatio

Example Knitr - co

Poproducib

Reproducio

Literature

Good and Bad Practic

Barrie

A First Look

Setup Minimal Examples

Quick Reporting

Feedback

Questions, Comments and Suggestions

Bibliography

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC. 16 de Abril

```
\documentclass{article}
```

\usepackage[utf8]{inputenc}

\begin{document}

\title{A Minimal Example}

\author{Yihui Xie}

\maketitle

We examine the relationship between speed and stopping distance using a linear regression model: $Y = \beta_0 + \beta_1 x + \epsilon$.

«model, fig.width=4, fig.height=3,

fig.align='center'>=

par(mar = c(4, 4, 1, 1), mgp = c(2, 1, 0), cex = 0.8)

abline(fit, lwd = 2)

0

The slope of a simple linear regression is

\Sexpr{coef(fit)[2]}

\end{document}



A First Look Minimal Examples

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

A First Look

Minimal Examples

Feedback

Universidad Distrital Francisco José de Caldas Maestría en CIC

Tendencias en IS Bogotá DC, 16 de Abril

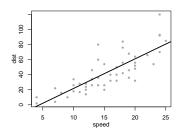
A Minimal Example

Yihni Xie

April 16, 2016

We examine the relationship between speed and stopping distance using a linear regression model: $Y = \beta_0 + \beta_1 x + \epsilon$.

```
par(mar = c(4, 4, 1, 1), mgp = c(2, 1, 0), cex = 0.8)
plot(cars, pch = 20, col = 'darkgray')
fit <- lm(dist " speed, data = cars)
abline(fit, 1wd = 2)
```



The slope of a simple linear regression is 3.9324088.



A First Look Quick Reporting

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

71000111107

Installation

Example Knitr - co

Reproducib

Literature

Good and Bad Practic

Barriers

A First Look

Setup

Quick Reporting

Feedback

Questions, Comments and Suggestions

Contact Informa

Bibliograph

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril If a user only has basic knowledge of R but knows nothing about knitr, or one does not want to write anything other than an R script, it is also possible to generate a quick report from this R script using the stitch() function.

The basic idea of stitch() is that knitr provides a template of the source document with some default settings, so that the user only needs to feed this template with an R script (as one code chunk); then knitr will compile the template to a report. Currently it has built-in templates for LATEX, HTML, and

Markdown.

The usage is like this: library(knitr) stitch("your-script.R")

stitch("your-script.K"



A First Look Extracting R Code

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Auti

Inetallation

Example Knitr - cod

Reproducib

nesearci

Good and Bad Practic

Barrier

A First Loo

Setup

Willillia Exali

Extracting R Code

Feedback

Questions, Comments and Suggestions

Ribliograph

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril For a literate programming document, we can either compile it to a report (run the code), or extract the program code in it. They are called "weaving" and "tangling," respectively. Apparently the function knit() is for weaving, and the corresponding tangling function is purl() in knitr. For example,

library(knitr)
purl("your-file.Rnw")
purl("your-file.Rmd")



Feedback Bugs, Comments and Suggestions

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Auti

Installation

Example Knitr - cod

D. . . . I . . I . I

Reproducib

Literature

Borriero

A First Look

Satur

Minimal Exampl

Quick Reporting Extracting R Code

Feedback

Questions, Comments and 17 Suggestions

Contact Informatio

Bibliography

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril ► Any Question, Comment or Suggestion.



Feedback Contact Information

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Aut

Installation

Example Knitr - coo

Reproducible

Research

Literature

Barrie

A First Look

Setup

Minimal Exampl

Extracting R C

Quartions Comments

Suggestions
Contact Information

Bibliography

Universidad Distrital Francisco José de Caldas Maestría en CIC Tendencias en IS Bogotá DC, 16 de Abril In case you have comments, suggestions or questions, please do not hesitate to contact me. You can find my contact details below.

Leonel Muñoz Cedano leoneling@gmail.com



Bibliography

Dynamic Documents with R and knitr

Leonel Muñoz Cedano

Introduction

About the Au

Installation

Example Knitr - coo

Reproducib

Research

Literature Good and Bad Practic

.

ATTION LOG

Minimal Exar

Quick Reporting Extracting R Cod

Feedback

Suggestions

Operated Information

Bibliography

....

- Yihui Xie, Dynamic Documents with R and knitr, Second Edition, A Chapman and Hall Book, Taylor and Francis Group, 2015.
- Borbón A. and Mora W., Edición de Textos Científicos LaTEX 2014, Second Edition, Escuela de Matemática, Instituto Tecnológico de Costa Rica, 2014.
- ► Pagina oficial R x64 (version 3.2.3) https://cran.r-project.org/bin/windows/base/.
- ▶ Jose Nelson Pérez Castillo, Documento para Integración de RStudio con TexStudio, 2016.

Thank you!

