Reproducible Research
What?
Why?
001-motivating-example
002-tables

Reproducible Research

An Introduction to knitr

Sahir Bhatnagar[1]

May 28, 2014

[1]sahir.bhatnagar@mail.mcgill.ca

Acknowledgements

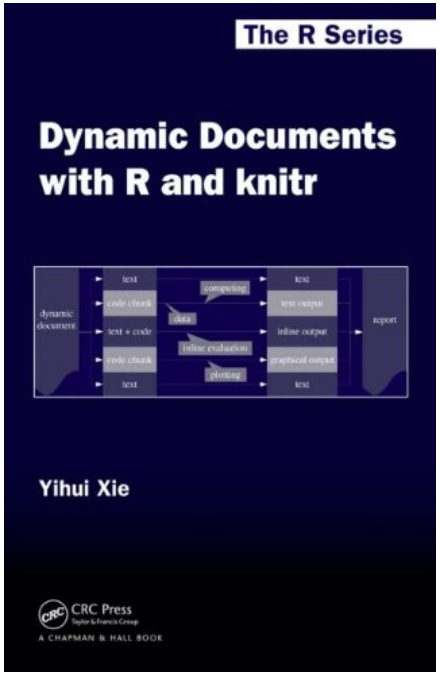
- Dr. Erica Moodie
- Maxime Turgeon, Kevin McGregor
- You

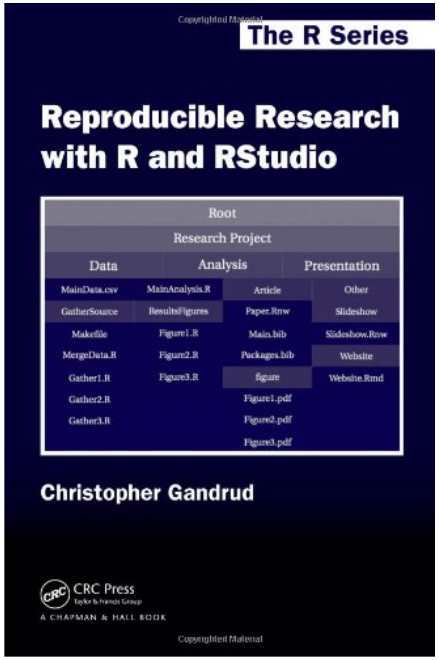
Disclaimer

I don't work for, nor am I an author of any of these packages. I'm just a messenger.

Disclaimer

- Material for this tutorial comes from many sources. For a complete list see: https://github.com/sahirbhatnagar/knitr-tutorial
- Alot of the content in these slides are based on these two books





Reproducible Research
What?
Why?
001-motivating-example
002-tables

Eat Your Own Dog Food

- These slides are reproducible
- Source code: https://github.com/sahirbhatnagar/knitr-tutorial

Reproducible Research
What?
Why?
001-motivating-example
002-tables

What is Science Anyway?

What is Science Anyway?

According to the American Physical Society:

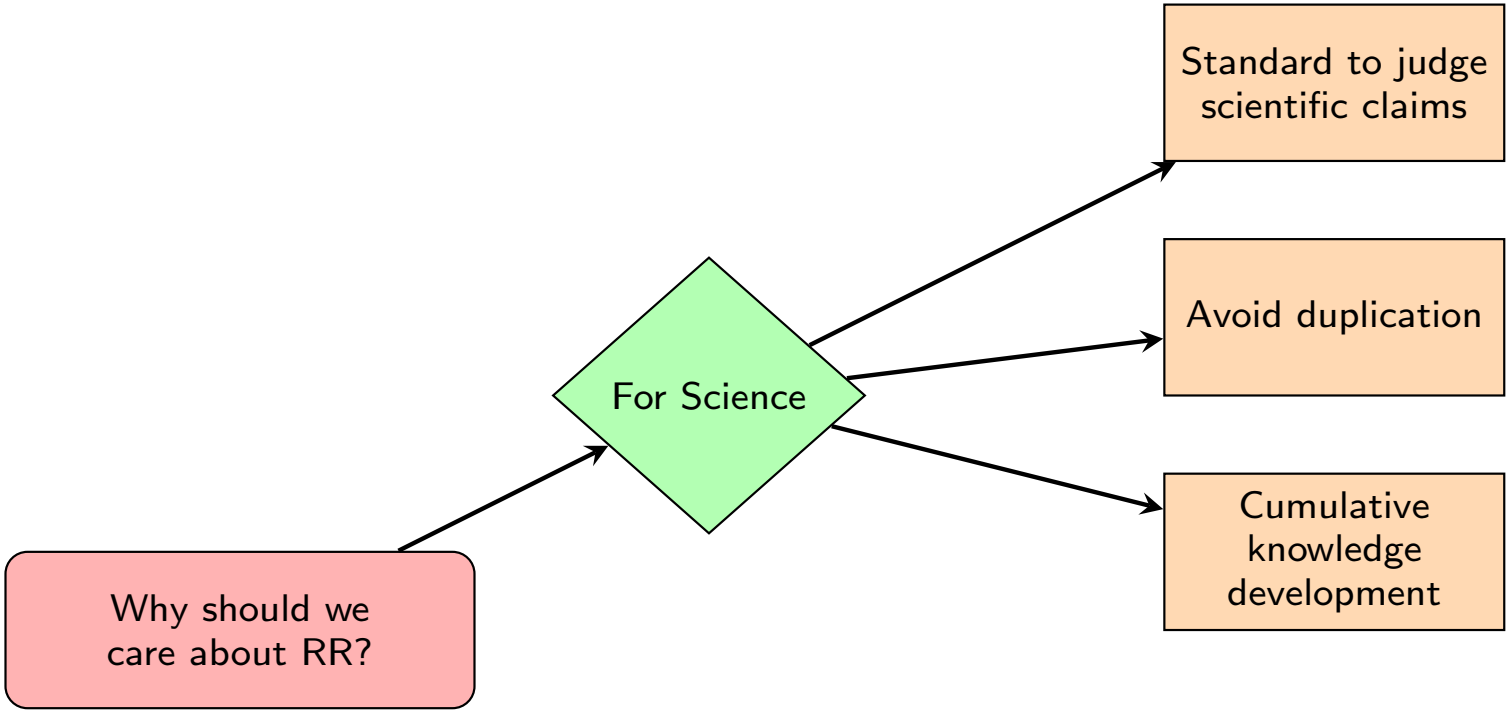
Science is the systematic enterprise of gathering knowledge about the universe and organizing and condensing that knowledge into **testable** *laws and theories. The* **success and credibility of science** *are anchored in the* **willingness** *of scientists to* **expose their ideas** *and results to* **independent testing** *and* **replication** *by other scientists*

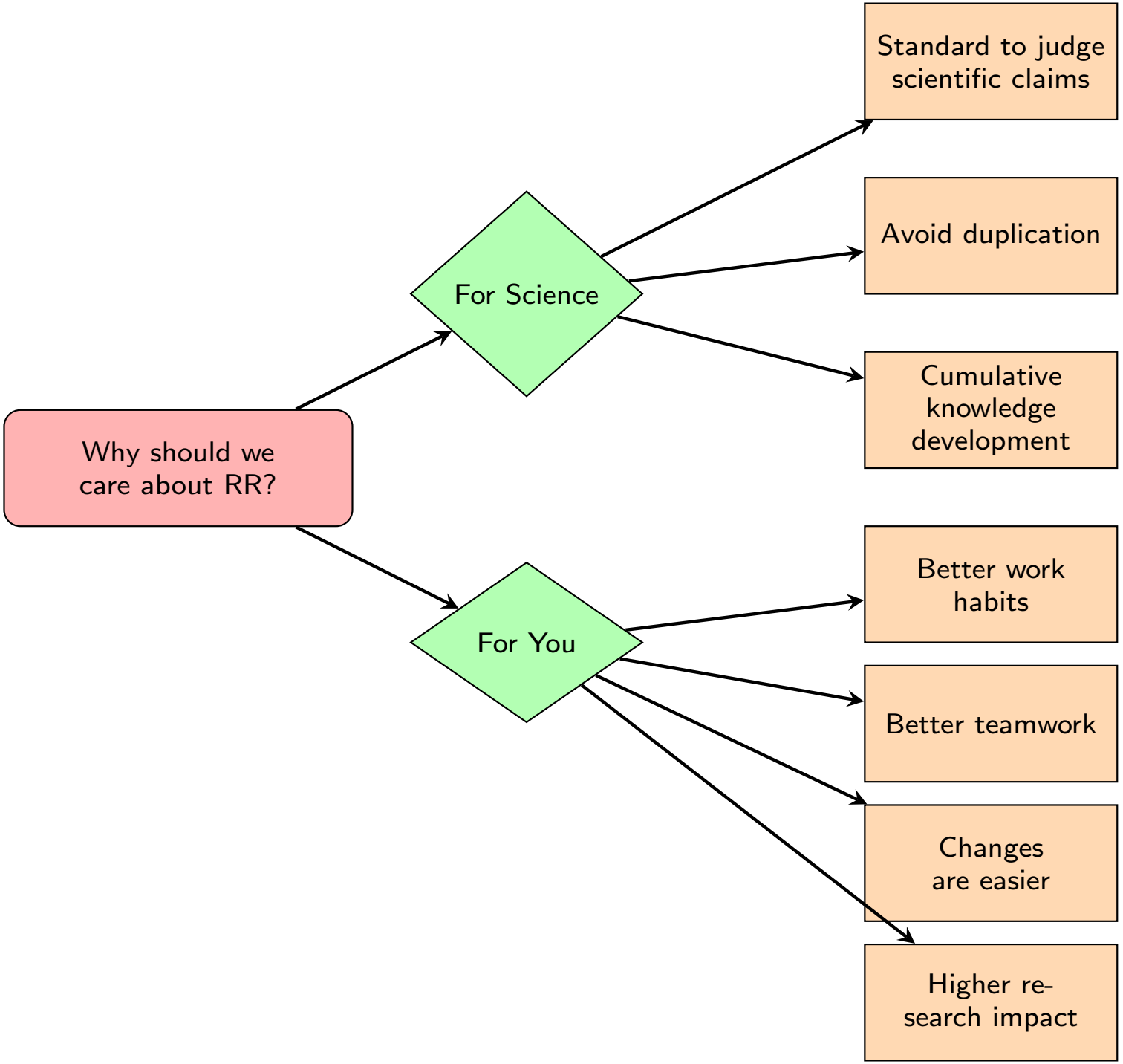
RR: A Minimum Standard to Verify Scientific Findings

RR: A Minimum Standard to Verify Scientific Findings

Reproducible Research (RR) in Computational Sciences

The data and the code used to make a finding are available and they are sufficient for an independent researcher to recreate the finding


Why should we care about RR?
For Science
Standard to judge scientific claims
Avoid duplication
Cumulative knowledge development


Why should we care about RR?
For Science
For You
Standard to judge scientific claims
Avoid duplication
Cumulative knowledge development
Better work habits
Better teamwork
Changes are easier
Higher research impact

	GLM	GLMM	GL
Intercept	**2.64** [2.38; 2.90]*	**2.20** [0.98; 3.41]*	**3.55** [1.7
progabide	−0.02 [−0.23; 0.19]	−0.23 [−0.76; 0.30]	−0.15 [−
time	−0.04 [−0.10; 0.01]	−0.04 [−0.10; 0.01]	−0.05 [−
age	−**0.01** [−0.02; −0.01]*	−0.01 [−0.05; 0.03]	−0.05 [−
progabide:time	−0.03 [−0.11; 0.05]	−0.03 [−0.11; 0.05]	0.01 [−0
AIC	3268.84	1403.91	
BIC	3286.16	1424.69	
Log Likelihood	-1629.42	-695.95	
Deviance	2492.94		
Num. obs.	236	236	23
Num. groups: subject		59	
Variance: subject.(Intercept)		0.87	
Variance: Residual		1.00	
Num. clust.			5

* 0 outside the confidence interval

Table 1 : Comparing model estimates

Table 2

	Dependent variable:
	mpg
wt	-5.30^{***}
	(0.56)
Constant	37.00^{***}
	(1.90)
Observations	32
R^2	0.75
Adjusted R^2	0.74
Residual Std. Error	3.00 (df = 30)
F Statistic	91.00^{***} (df = 1; 30)
Note:	*p<0.1; **p<0.05; ***p<0.01

numeric(0)

	GLM	GLMM	GL
Intercept	**2.64** [2.38; 2.90]*	**2.20** [0.98; 3.41]*	**3.55** [1.7
progabide	−0.02 [−0.23; 0.19]	−0.23 [−0.76; 0.30]	−0.15 [−
time	−0.04 [−0.10; 0.01]	−0.04 [−0.10; 0.01]	−0.05 [−
age	−**0.01** [−0.02; −0.01]*	−0.01 [−0.05; 0.03]	−0.05 [−
progabide:time	−0.03 [−0.11; 0.05]	−0.03 [−0.11; 0.05]	0.01 [−0
AIC	3268.84	1403.91	
BIC	3286.16	1424.69	
Log Likelihood	-1629.42	-695.95	
Deviance	2492.94		
Num. obs.	236	236	23
Num. groups: subject		59	
Variance: subject.(Intercept)		0.87	
Variance: Residual		1.00	
Num. clust.			5

* 0 outside the confidence interval

Table 3 : Comparing model estimates