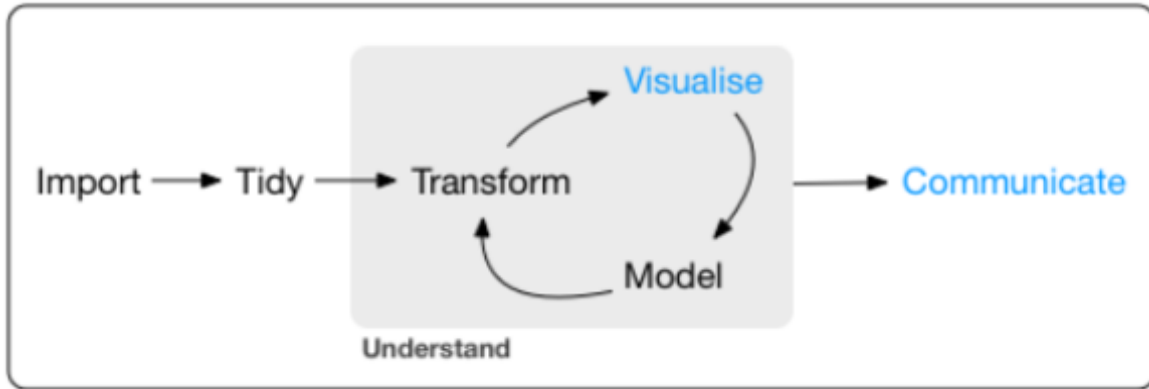


데이터 연동형 문서 만들기(1) - R Markdown 및 Knitr 패키지 소개

데이터 분석의 OSEMN 프로세스에서

Hadley Wickham 이 R for data Science 의 가장 마지막 부분에서 강조하고 있는
'공유(Communicate)' 부분을 살펴보자.



내가 아닌, 다른 사람이 생성해서 데이터가 변경 또는 삭제되고,
또 수식이 잔뜩 걸린 엑셀 파일을 받아서 다시 분석을 한 경험이 한 번이라도 있는 사람들은 알 것이다.
그게 얼마나 힘들고 짜증스러운 일인지..

R 은 txt 로 된 코드 형태라 ,
코드와 데이터만 받아서 실행하면 남이 분석한 것을 그대로 재현할 수 있기에 이 문제를 보완할 수 있는데,
R Markdown 와 Knitr 패키지를 활용하면
더 손쉽게 다른 사람들과 작업 내용을 효율적으로 공유할 수 있다.

Markdown

기존에 우리가 인터넷 세상에서 보는 웹 페이지(= 결국 문서지..)는
HTML 이라는 Markup(마크업) 언어인데,
HTML 을 제대로 쓰려면 태그 등을 알아야 하는 부담이 있다.

이러한 Markup 언어의 단점을 보완하기 위해
2004 년 John Grube 가 읽기도 쓰기도 쉬운 Markup 언어를 만들었는데,
이게 바로 Markdown 이며 확장자는 .md 을 사용.

R Markdown, knitr

R 에서도 이 Markdown 을 활용하여
코드, 데이터가 연동된 Dynamic 한 문서를 만들 수 있는데,
이게 바로 'R Markdown'이며
R 와 Markdown 을 체계적으로 연결시켜주는 도구가 'knitr'(니터'라고 읽음) 패키지'이다*.

RStudio 홈페이지에서는

R Markdown 을 '**Markdown 으로 작성되었으며 R 코드가 내장되어 있는 문서**'**라고 정의하고 있다.

Markdown 문서는 확장자가 .md 인데, R Markdown 문서는 .rmd 인게 차이점.

* R 과 knitr 를 활용한 데이터 연동형 문서 만들기(고석범)

** R Markdown document is written in markdown (an easy-to-write plain text format) and contains chunks of embedded R code, like the document below.

구 분	역 할
데이터 분석용 프로그래밍 언어	R
사람이 읽을 수 있는 문서를 만들어주는 언어(마크업 언어)	Markdown
두 언어를 체계 있게 연결시키는 도구	Knitr 패키지
이 모두를 통합한 개발 환경	Rstudio

주) 출처 : R 과 knitr 를 활용한 데이터 연동형 문서 만들기(고석범) 214p

RStudio 홈페이지의 Cheatsheet 자료 : 잘 정리된 요약자료

R 세계에서 유명한 Hadley Wickham 과 Garrett Grolemond 가 쓴

'R for Data Science'에는

R Markdown 과 관련된 요약자료 2 개를 소개하고 있는데,

이 cheatsheet 들은 ggplot2 패키지의 경우도 그랬지만 정리가 잘 되어 있어, 매우 유용하다.

1) R Markdown Cheat Sheet

2) R Markdown Reference Guide

이 자료들은 RStudio IDE 의 도움말(help) 아래에서 볼 수 있으며,

RStudio 홈페이지의 Cheatsheet 섹션에서도 볼 수 있다.

Cheatsheets - RStudio

RStudio Cheat Sheets The cheat sheets below make it easy to learn about and use some of our favorite packages.

From time to time, we will add new cheat sheets to the gallery. If you'd like us to drop you an email when we do, let us know by clicking the

www.rstudio.com

한편, 'R for Data Science'의 공동저자인 Garrett Grolemond 가

R Markdown 이 무엇인지에 대해서 간단하면서도 잘 소개한 자료도 Rstudio 에 올라와 있으니 참고.. (봐야할 건 정말 너무나 많다..)

Introduction to R Markdown

R Markdown is a file format for making dynamic documents with R. An R Markdown document is written in markdown (an easy-to-write plain text format) and contains chunks of embedded R code, like the document below. --- output: html_document --- This is an R Markdown document. Markdown is a simple for rmarkdown.rstudio.com

해외 도서

R Markdown 관련해서는

2018 년 7 월, knitr 패키지를 개발한 Yihui Xie 와 Garrett Grolemund 등이

"**R Markdown : The definitive guide(= 최종 안내서)**"라는 거창한 제목의 책을 쓰기도 하였다.