# M11-intro

Learning Spoons R

2018-12-22

# 시작

#### "Hello World"

## [1] "Hello World"

- 0. WIFI + R/Rstudio 설치
- 1. 강사 소개
- 2. Final Project로 보는 강의 목적 (M41)
- 3. 수강생 Project 소개

## R로 시작하는 데이터 분석 및 시각화

- 부제: 마케팅/영업/기획/매출관리 직무가 무엇이든 데이터 분석은 필수
- · 강의페이지 https://github.com/LearningSpoonsR/LS-DS
- FB그룹(가입신청) https://www.facebook.com/groups/440130033101009/
- 일정
  - · 2018.12.23 ~ 2019.01.27 (5회)
  - · 2018.12.30 휴강
  - 매주 일요일 10:30 ~ 13:30
  - 2018.12.23, 2019.1.6, 2019.1.13, 2019.1.20, 2019.1.27
- 코스매니저: 김형종 (출결등 기타)
- 강사이메일: learningSpoonsR@gmail.com

## 데이터 사이언스 vs 프로그래밍

- 데이터 사이언스
  - · 데이터를 과학적인 방법으로 접근하여
  - 이해, 분석, 의사결정의 도구로 삼는 것
- 프로그래밍
  - · 컴퓨터가 특정한 목적을 가진 작업을 수행하게 하는 것
- 관계
  - 1. 프로그래밍은 데이터 사이언스의 도구로 사용됨
  - 2. 데이터 규모가 커질 수록, 이해와 분석의 깊이가 깊어질 수록 프로그래밍 의존도 상승

## DS 프로젝트 과정과 요구 Skill Set

- DS?
  - · Mathematics is to find patterns, Dr, Feynman
  - DS is to find patterns from data.
- DS 프로젝트 과정
  - 1. 결론이 될 "검증하고 싶은 가설"을 정의
  - 2. "가설의 검증"을 프로그램의 "목표"로 설정
  - 3. "목표"를 이루기 위한 과정을 작업 단위로 분리하여 설계
  - 4. 프로그램 작성
  - 5. 가설 검증
  - 6. 결론 도출
  - 7. 결론의 공유
- Skill Set (현재 자신의 Skill을 별5개 만점으로 표시해보세요)
  - 1. 컴퓨터 기술 (SW/HW/Programming)
  - 2. 수학/통계학
  - 3. 커뮤니케이션
  - 4. 논리적/과학적 접근
  - 5. 영어

# 수업 목표 & 진행 방식

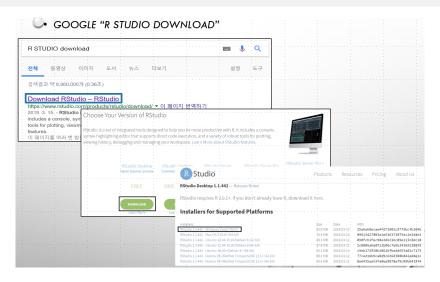
#### • 목표

- 1. DS를 위한 프로그래밍을 배우고 프로젝트의 과정을 체득합니다.
- 2. Final Project의 결과물을 낼 수 있을 정도로 R 프로그램을 다룰 수 있습니다.
- 3. 수업 Final Project에 준하는 결과물을 강의 마지막에 만들어 냅니다.
- 4. R 프로그래밍을 스스로 프로그래밍할 수 있게 됩니다.
- 5. 앞의 별표 갯수를 늘립니다!
- 6. Skill set을 일터로 가져갑니다.
- 진행 방식
  - 1. 데이터 분석 과정을 포함한
  - 2. Mini-Project를 반복하며
  - 3. 프로젝트 과정을 체득합니다.

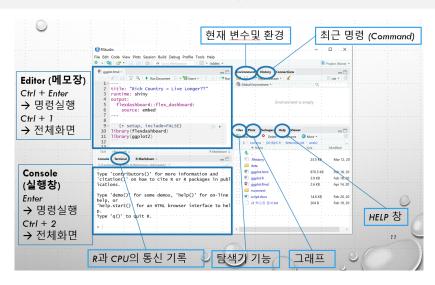
- GOOGLE "R DOWNLOAD"
- LINK → BASE → DOWNLOAD → INSTALL → 아이콘 2개



## RStudio (1)



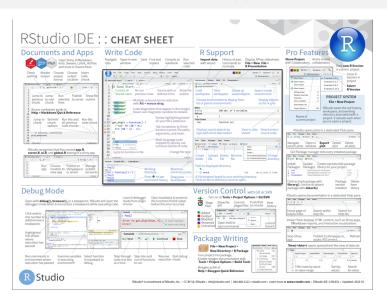
## RStudio (2)



## RStudio (3)

- 1. Tool -> Global Option
- Appearance: 폰트와 색상 등을 조정 (어두운 바탕에 밝은 글씨 vs 밝은 바탕에 어두운 글씨)
- PANE Layout: 모니터 크기와 작업 목적에 따라 조정 (좌우 vs 상하로 Editor와 Console창 조정)
- CODE -> Saving -> Text Encoding in -> UTF-8
- 2. 윈도우 사용자 계정과 R파일 폴더 이름은 영어로
- 제어판 -> 사용자 계정 -> 계정 이름 변경 -> 영어로 입력
- R파일, 불러올 데이터 파일 등은 가급적이면 영어로
- R Studio는 관리자 권한으로 실행해야 함 (default로 관리자 권한)
- 3. 한영 전환
- 한글로 전환: Sys.Setlocale("LC\_ALL", "ko\_KR.UTF-8")
- 영어로 전환: Sys.Setlocale("LC\_ALL", "en\_KR.UTF-8")

## RStudio (4)



# RStudio (5)

				4 WRITE CODE	Windows /Linux	Mac	WHY RSTUDIO SERVER PRO?		
LAYOUT		Windows/Linu		Attempt completion	Tab or Ctrl+Space	Tab or Cmd+Space	RSP extends the the open source:	- 44	
Move focus to Source Editor Ctrl+1			Ctrl+1	Navigate candidates	↑/↓	↑/↓			
Move focus to Console Ctrl+2			Ctrl+2	Accept candidate	Enter, Tab, or >	Enter, Tab, or >	commercial license, support, and more:		
Move focus to Help Ctrl+3			Ctrl+3	Dismiss candidates	Esc.	Ender, 180, or 7	open and run multiple R sessions at once		
Show History Ctrl+4			Ctrl+4	Undo	Ctrl+Z	Cmd+7	tune your resources to improve performance		
		Ctrl+5	Ctrl+5	Redo	Ctrl+Shift+Z	Cmd+Shift+Z	edit the same project at the same time as others		
		Ctrl+6	Ctrl+6	Cut	Ctrl+X	Cmd+X	see what you and others are doing on your server		
		Ctrl+7	Ctrl+7	Copy	Ctrl+C	Cmd+C	<ul> <li>switch easily from one version of R to a different version</li> </ul>		orsion
		Ctrl+8	Ctrl+8	Paste	Ctrl+V	Cmd+V	<ul> <li>Integrate with your authentication, authorization, and audit practic</li> </ul>		
		Ctrl+9	Ctrl+9	Select All	Ctrl+A	Cmd+A			nu auun practices
show Build		Ctrl+0	Ctrl+0	Delete Line	Ctrl+D	Cmd+D	Download a free 45 day evaluation		
RUN CODE		Windows/Linu	x Mac	Soloct	Shift+(Arrow)	Shift+(Arrow)	www.rstudio.com/products/rst	udio-server-pro	
Search command his	ton	Ctrl+↑	Cmd+♠	Select Word	Shift+(Arrow) Ctrl+Shift+ ◆/→	Option+Shift+ +/+	S DEBUG CODE	Windows/Li	non Mar
lavigate command his		1/4	1/4					Shift+F9	Shift+F9
fove cursor to start of		Home	Cmd+4	Select to Line Start	Alt+Shift+	Cmd+Shift+◆	Toggle Breakpoint		
Nove cursor to start or Nove cursor to end of		End	Cmd+	Select to Line End	Alt+Shift+	Cmd+Shift+	Execute Next Line	F10	F10
hange working direct		Ctrl+Shift+H	Ctrl+Shift+H	Select Page Up/Down	Shift+PageUp/Down	Shift+PageUp/Down	Step Into Function	Shift+F4	Shift+F4
nange working direct nterrupt current con		Esc	Esc.	Select to Start/End	Shift+Alt+↑/↓	Cmd+Shift+♠/◆	Finish Function/Loop	Shift+F6	Shift+F6
lear console	minano	Ctrl+L	Ctrl+L	Delete Word Left	Ctrl+Backspace	Ctrl+Opt+Backspace	Continue	Shift+F5	Shift+F5
uit Session (desktop)		Ctrl+O	Cmd+O	Delete Word Right		Option+Delete	Stop Debugging	Shift+F8	Shift+F8
guit Session (desktop)	OHIN	Ctrl+Shift+F10	Cmd+Shift+F10	Delete to Line End		Ctrl+K			
Restart K Session Run current line/sele		Ctrl+Enter	Cmd+Snift+10	Delete to Line Start		Option+Backspace	6 VERSION CONTROL	Windows/Lin	
		Alt+Enter	Option+Enter	Indent	Tab (at start of line)	Tab (at start of line)	Show diff	Ctrl+Alt+D	Ctrl+Option+I
Run current (retain cur Run from current to en		Alt+Enter Ctrl+Alt+E	Option+Enter Cmd+Option+E	Outdent	Shift+Tab	Shift+Tab	Commit changes	Ctrl+Alt+M	Ctrl+Option+
				Yank line up to cursor	Ctrl+U	Ctrl+U	Scroll diff view	Ctrl+◆/◆	Ctrl+◆/◆
Run the current function	on	Ctrl+Alt+F	Cmd+Option+F	Yank line after cursor	Ctrl+K	Ctrl+K	Stage/Unstage (Git)	Spacebar	Spacebar
Source a file		Ctrl+Alt+G	Cmd+Option+G	Insert yanked text	Ctrl+Y	Ctrl+Y	Stage/Unstage and move to next	Enter	Enter
iource the current fil	ie	Ctrl+Shift+Ente	Cmd+Shift+S	Insert <-	Alt+-	Option+-			
ource with echo		Ctrl+Shitt+Ente	Cmd+Shift+Enter	Insert %>%	Ctrl+Shift+M	Cmd+Shift+M	7 MAKE PACKAGES	Windows/Lin	my Mac
NAVIGATE CODE	Windox	vs /Linux 1	fac	Show help for function	F1	F1	Build and Reload	Ctrl+Shift+B	Cmd+Shift+E
Soto File/Function	Ctrl+.		trl+.	Show source code	F2	F2	Load All (devtools)	Ctrl+Shift+L	Cmd+Shift+I
Fold Selected	Alt+L		md+Option+L	New document	Ctrl+Shift+N	Cmd+Shift+N			Cmd+Shift+
Infold Selected	Shift+Al		md+Shift+Option+L	New document (Chrome)	Ctrl+Alt+Shift+N	Cmd+Shift+Opt+N	Test Package (Desktop)	Ctrl+Shift+T	
old All	Alt+O			Open document	Ctrl+O	Cmd+O	Test Package (Web)	Ctrl+Alt+F7	Cmd+Opt+Fi
			md+Option+O	Save document	Ctrl+S	Cmd+S	Check Package	Ctrl+Shift+E	Cmd+Shift+E
Jnfold All	Shift+Al		md+Shift+Option+O	Close document	Ctrl+W	Cmd+W	Document Package	Ctrl+Shift+D	Cmd+Shift+
io to line	Shift+Al		md+Shift+Option+G	Close document (Chrome)	Ctrl+Alt+W	Cmd+Option+W			
ump to	Shift+Al		md+Shift+Option+J	Close all documents	Ctrl+Shift+W	Cmd+Shift+W	8 DOCUMENTS AND APPS	Windows/Linu	
witch to tab	Ctrl+Shi		trl+Shift+.	Extract function	Ctrl+Alt+X	Cmd+Option+X	Preview HTML (Markdown, etc.)	Ctrl+Shift+K	Cmd+Shift+K
revious tab	Ctrl+F11		tri+F11	Extract variable	Ctrl+Alt+V	Cmd+Option+V	Knit Document (knitr)	Ctrl+Shift+K	Cmd+Shift+K
lext tab	Ctrl+F12		trl+F12	Reindent lines	Ctrl+I	Cmd+I	Compile Notebook	Ctrl+Shift+K	Cmd+Shift+K
irst tab	Ctrl+Shi		trl+Shift+F11	(Un)Comment lines	Ctrl+Shift+C	Cmd+Shift+C	Compile PDF (TeX and Sweave)	Ctrl+Shift+K	Cmd+Shift+K
ast tab	Ctrl+Shi	R+F12 (	trl+Shift+F12	Reflow Comment	Ctrl+Shift+/	Cmd+Shift+/	Insert chunk (Sweave and Knitr)	Ctrl+Alt+I	Cmd+Option+I
lavigate back	Ctrl+F9		md+F9	Reformat Selection	Ctrl+Shift+A	Cmd+Shift+A	Insert code section	Ctrl+Shift+R	Cmd+Shift+R
lavigate forward	Ctrl+F10		md+F10	Select within braces	Ctrl+Shift+E	Ctrl+Shift+E	Re-run previous region	Ctrl+Shift+P	Cmd+Shift+P
lump to Brace	Ctrl+P		trl+P	Show Diagnostics	Ctrl+Shift+Alt+P	Cmd+Shift+Opt+P	Run current document	Ctrl+Alt+R	
Select within Braces	Ctrl+Shi	ft+Alt+E (	trl+Shift+Option+E	Transpose Letters		Ctrl+T			Cmd+Option+F
ise Selection for Find	Ctrl+F3		md+E	Move Lines Up/Down	Alt+ <b>↑/</b> ↓	Option+1/4	Run from start to current line	Ctrl+Alt+B	Cmd+Option+
ind in Files	Ctrl+Shi		md+Shift+F	Copy Lines Up/Down	Shift+Alt+↑/↓	Cmd+Option+↑/◆	Run the current code section	Ctrl+Alt+T	Cmd+Option+
ind Next	Win: F3	Linux: Ctrl+G (	md+G	Add New Cursor Above	Ctrl+Alt+Up	Ctrl+Option+Up	Run previous Sweave/Rmd code	Ctrl+Alt+P	Cmd+Option+I
ind Previous	W: Shift		md+Shift+G	Add New Cursor Below	Ctrl+Alt+Down	Ctrl+Option+Down	Run the current chunk	Ctrl+Alt+C	Cmd+Option+0
lump to Word	Ctrl+ 4		option+ •/•	Move Active Cursor Up	Ctrl+Alt+Shift+Up	Ctrl+Option+Shift+Up	Run the next chunk	Ctrl+Alt+N	Cmd+Option+1
ump to Word ump to Start/End	Ctrl++/		imd+ <b>↑/</b> ↓	Move Active Cursor Down		Ctrl+Option+Shift+Down	Sync Editor & PDF Preview	Ctrl+F8	Cmd+F8
oggle Outline	Ctrl+Shi		md+Shift+O	Find and Replace	Ctrl+F	Cmd+F			Cmd+Option+F11
	Coltail		and same of	Use Selection for Find	Ctrl+F3	Cmd+E	Next plot		Cmd+Option+F12
-00									
Studi				Replace and Find	Ctrl+Shift+J	Cmd+Shift+J	Show Keyboard Shortcuts	Alt+Shift+K	Option+Shift+

12/2

# Cheatsheet & English

#### · Cheatsheets?

- 모든 문법과 명령어를 외우는 것은 불가능합니다.
- 언제든지 참조하고 검색하는 것이 프로그래밍의 과정입니다.
- Cheat Sheet은 "컨닝페이퍼"로서 사용법이 정리되어 있는 문서입니다.
- 여러분은 수강후에 Cheat Sheet과 과거의 경험을 참조하고 검색하면서 프로그래밍을 하게 될 것입니다.

#### English?

- Cheat Sheet을 가득 채운 영어가 겁나시나요?
- 웹 검색 결과와 프로그램 설명서 등은 대부분 영어로 되어있습니다.
- 그러나 대부분 컴퓨터와 프로그래밍에 관련된 용어일 뿐입니다.
- 자주 사용되는 용어는 약 1000개 정도일 뿐이므로 경험과 함께 금방 익숙해 집니다.
- 프로그래밍에 관련된 설명은 어려운 문법을 사용하는 경우가 드물어 구글/네이버 번역기으로도 아주 잘 이해할 수 있습니다.

## Excel vs R

#### Excel

- GUI (화면에 자동으로 보임)
- 분석 과정이 기록되지 않으며 결과 위주
- 매우 쉬움
- 데이터가 커지면 느려서 버벅됨
- 눈에 보이기에 직관적
- 실수에 관대함

#### R

- Console Interface (명령어로 접근)
- Reproducibility (재현가능성)
- 분석 과정이 기록되어 결과를 재생산할 수 있음
- 초반의 학습곡선이 있음
- 훨씬 빠르게 대용량 데이터 처리
- 추상적, 논리적인 사고가 요구되고 길러짐
- 결과물에 실수 없음이 어느정도 보장

## R vs 다른 프로그래밍 언어

분석 (Scie	ntific Computing)	개발 (Development)						
R	가장 배우기 쉽고 빠르게 성장	C, C++	가장 빠르며 다른 언어의 기반					
MATLAB	수리 계산 (유료)	JAVA	웹 개발용으로 주로 쓰임					
SAS	슈퍼 대용량 Data (매우 유료)							
Python: R보다 좀 더 어려움, 분석 및 개발 양쪽을 충분히 지원 R보다 조금 더 빠르며 R을 배우고 나면 더 쉽게 배울 수 있음								

#### · R의 장점

- 무료
- · 가장 왕성한 open-source 확장 프로그램들
- · Rstudio 회사가 존재
  - 확장 프로그램들을 통합하여 관리
  - 매뉴얼 제공, 교육
  - 더 쉽고 빠르고 편리하고 강력한 프로그램 개발
- 예를 들어, 어려운 문법의 딥러닝 라이브러리인 Tensorflow도 현재는 Keras를 거쳐 Rstudio에서 거의 모든 기능을 사용할 수 있음.

## R은 왜 배워야 하나요?

- 1. 일 잘하는 사람
- · 암산 -〉 주판 -〉 엑셀 -〉 프로그래밍
- 계산기 -> 도스 -> 윈도우 -> 프로그래밍
- 반복되는 작업을 컴퓨터에게 맡겨서 처리
- 프로그래밍 언어 사용자 중 최고 연봉 (미국)
- 2. 빠르게 변하는 세상과 교육 수준
- 구구단 -> 19단 -> 코딩 교육 열풍
- 처리해야 할 데이터는 급격하게 증가
- 커리어의 발전을 위해 오퍼레이션의 반복이 아닌 **구조를 파악**하여 **자동으로 해결**할 수 있는 접근법을 익혀야 함

#### M1X: Introduction & Basics

- M11 intro
  - · 강의 계획서. Data Science와 R 프로그램에 대해 소개합니다.
- M12 base
  - base
  - Data Type과 Data Structure에 대해서 논의합니다.
- M13 Quiz 1
  - rmarkdown
  - M12 review (.docx)

## M2X: The First Project - Carseat Sales

- 사용하기 편한 간단한 패키지들을 학습합니다.
- Carseats Sales 데이터를 이용해서 rmarkdown으로 작성된 report를 만듭니다.
- M21 flow
  - · dplyr, ggplot2, rmarkdown, ISLR
  - ISLR의 Carseats 매출데이터로 자료 처리, 시각화, 문서화의 전 과정을 다룹니다.
- M22 CarseatsReport
  - rmarkdown (.html, .docx, .pdf)
  - 실제로 분석 결과를 보고서로 작성합니다.

#### M23 More on ggplot2

- ggplot2
- 시각화 기법에 대해서 좀 더 깊이있게 다룹니다.

#### M24 ggplot2 Gallery

- ggplot2
- 사용성이 높은 ggplot2를 이용한 아름다운 시각화 예제를 살펴봅니다.

#### M25.pdf

- rmarkdown, texlive
- 한글로 된 pdf문서를 작성할 수 있는 템플릿입니다. M22에 pdf 조판을 위한 textlive 프로그램 설치 과정을 진행해야 합니다.

#### M26 beamer

- · rmarkdown, texlive
- 한글로 된 프리젠테이션 문서(수업 강의노트)를 작성할 수 있는 템플릿입니다. M22에 pdf 조판을 위한 textlive 프로그램 설치 과정을 진행해야 합니다.

#### · M27 slidy

- rmarkdown
- rmarkdown을 사용하면 손쉽게 .html로 된 프리젠테이션 문서를 만들 수 있습니다. 이에 해당하는 slidy 포맷의 템플릿입니다.

## M3X: Textmining & Web Applications

- M3X 부터는 dependency가 없습니다.
- 즉, M1X와 M2X까지를 익숙하게 다룰 수 있다면 M3X이후에는 모듈 각각에 대해서 독립적으로 이해하고 사용할 수 있습니다.
- Textmining 기법(워드클라우드, 크롤링, 번역)에 대한 예제를 제공합니다.
- Flexdashboard와 Shiny Application으로 구현해 봅니다.

#### M31 I have a dream

- tm, KONLP, wordcloud, function
- 마틴루터킹 목사의 "I have a dream"에 대해 Wordcloud 분석을 하고, 이를 다양한 텍스트 파일 포맷과 한/영 문서에 대해서도 일반화 시킵니다. |

#### M32 flexdashboard

- · flexdashboard, rmarkdown
- rmarkdown을 다룰수 있다면 손쉽게 이를 .html의 대시보드로 제작할 수 있습니다. 이는 반응형 대시보드(사용자의 마우스/키보드 동작에 반응)로 연장되는 shiny 패키지의 기본이 됩니다.

#### · M33 shiny on fd

- shiny flexdashboard rmarkdown
- flexdashboard와 rmarkdown을 다룰수 있다면 손쉽게 반응형 대시보드(사용자의 마우스/키보드 동작에 반응)또한 구현할 수 있습니다.
- shiny는 빠른 시간에 구현할 수 있는, elegant하고 professional한 프리젠테이션 기법이 됩니다.

## M34 Happy B-day, ggplot

- rvest, googleLanguageR
- ggplot2는 최근 10년에 걸쳐 데이터 시각화의 혁신을 가져왔습니다. 이에대한 영문 article을 web crawling하여 google 번역기를 이용해서 자동번역하는 프로그램을 제각합니다. |

### · M35 R and Python are joining

- · rvest, googleLanguageR
- 위의 M34와 같은 프로그램입니다. 데이터 사이언스에 있어서 R와 Python의 장단점과 둘의 융합이라는 트렌드에 대한 article을 역시 web crawling으로 얻어와 google 번역기를 자동으로 실행시켜 한글 문서로 제작합니다.

## M4X: Case Studies - Data Analysis

- M41 longevity
  - · shiny
- M42 subway

.

M43 retail

.

M44 SNS

.

M45 sales log

.

## M5X: Other Advanced Applications

- M51 tidyr
  - · tidyr, SQL
- M52 Time Series
  - xts, dygraph, lubridate
- M53 python
  - feather
- M54 KMA & Auto
  - jsonlite, mailR, taskscheduleR
- M55 Outtro

#### M6X: References

- Book
- Cheatsheets
- Dr. Hadley

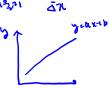
# M41-longevity (1/2)

상관관계 (correlation) vs 인과관기 (causation) .

- - 1. income -> 기대수명
  - 부모수명 -> 자녀수명
  - 쌍둥이 형 -> 쌍둥이 동생
- - 2104:219L

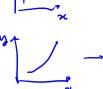


- y = ax + b



76







# M41-longevity (2/2)

# 수강생 Survey (1)

- 1. 이름, 이메일
- 2. 프로그래밍 경험이 있으신가요?
- 3. 수강 후에 어떤 능력을 배워가고 싶으신가요?

# 수강생 Survey (2)

- 1. 생각하고 있는 프로젝트에 대해서 적어주세요.
  - 주제 & 목적:
  - 데이터셋 설명:
  - 결과물 포맷 (보고서/대시보드):
- 2. 팀워크 진행 선호 여부에 대해서 아래 중에서 선택하고 적어주세요.
  - 2.1 해당 주제를 혼자 진행하기를 원함
  - 2.2 해당 주제 crew를 원함 (최대 인원을 적어주세요)
  - 2.3 다른 이의 팀에 join하기를 원함 (어떤 주제의 프로젝트에 join하기를 원하시는 지 적어주세요)
- 3. 팀워크 가이드라인
  - 호칭은 XX님으로; 상호 존칭 및 매너 엄수; 실례가 될 수 있는 질문 삼가
  - .Rmd파일 공유 및 소통
  - The best way to learn is to teach!
  - 소통, 질문, 대답하는 분위기를 조성해주세요.

## M11-intro