1. 데이터 시각화를 위한 ggplot2

This is an [R Markdown](http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Ctrl+Shift+Enter*.

plot(cars)



Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Ctrl+Alt+I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Ctrl+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.

import pandas as pd  
  
from dfply import \*  
import seaborn as sns  
import matplotlib.pyplot as plt  
import matplotlib.font\_manager as fm  
font\_list = [font.name for font in fm.fontManager.ttflist]  
font\_list

## ['STIXNonUnicode', 'DejaVu Serif', 'DejaVu Serif', 'STIXGeneral', 'cmtt10', 'STIXGeneral', 'STIXSizeTwoSym', 'DejaVu Sans', 'cmr10', 'cmss10', 'DejaVu Sans Mono', 'cmex10', 'DejaVu Sans', 'STIXSizeFourSym', 'STIXGeneral', 'DejaVu Serif', 'STIXSizeFiveSym', 'cmsy10', 'STIXSizeOneSym', 'DejaVu Sans', 'DejaVu Sans Mono', 'STIXNonUnicode', 'STIXSizeTwoSym', 'DejaVu Sans Mono', 'DejaVu Serif', 'STIXSizeOneSym', 'STIXSizeFourSym', 'DejaVu Sans Display', 'STIXNonUnicode', 'STIXSizeThreeSym', 'cmb10', 'STIXSizeThreeSym', 'STIXGeneral', 'DejaVu Serif Display', 'cmmi10', 'DejaVu Sans', 'STIXNonUnicode', 'DejaVu Sans Mono', 'KBIZgo M', 'Mongolian Baiti', 'Constantia', 'MGungHeulim', 'Segoe UI', 'HCR Batang ExtB', 'Trebuchet MS', 'Wingdings 2', 'KBIZgo R', 'MHunmin', 'Bell MT', 'Gill Sans MT', 'Book Antiqua', 'Calisto MT', 'Palatino Linotype', 'Gill Sans MT Ext Condensed Bold', 'Stencil', 'Eras Demi ITC', 'Palatino Linotype', 'HYPMokGak-Bold', 'Century Schoolbook', 'HyhwpEQ', 'Gill Sans MT', 'Ebrima', 'Tw Cen MT Condensed', 'HYGothic-Medium', 'Segoe UI', 'Cambria', 'Lucida Bright', 'HCR Batang Ext', 'MoeumT R', 'Castellar', 'HCR Batang', 'Batang', 'Tahoma', 'Arial', 'Segoe Print', 'Lucida Fax', 'Edwardian Script ITC', 'Calibri', 'Colonna MT', 'Calisto MT', 'Vladimir Script', 'Perpetua', 'OCR A Extended', 'Sitka Small', 'MSugiHeulim', 'Bahnschrift', 'Times New Roman', 'Franklin Gothic Medium Cond', 'Georgia', 'Footlight MT Light', 'KBIZmjo B', 'MingLiU-ExtB', 'Tw Cen MT', 'Corbel', 'Nirmala UI', 'Verdana', 'Arial Rounded MT Bold', 'Bodoni MT', 'Ami R', 'Nirmala UI', 'Bodoni MT', 'French Script MT', 'Magic R', 'Gulim', 'KBIZgo B', 'LG PC', 'Magneto', 'Eras Bold ITC', 'Wide Latin', 'Goudy Stout', 'Consolas', 'Franklin Gothic Heavy', 'Cooper Black', 'Franklin Gothic Demi', 'Consolas', 'MGungJeong', 'Leelawadee UI', 'Lucida Sans Typewriter', 'Microsoft JhengHei', 'Microsoft Uighur', 'Georgia', 'Yu Gothic', 'Constantia', 'Californian FB', 'Arial', 'Bookman Old Style', 'Rockwell', 'Trebuchet MS', 'Yu Gothic', 'Trebuchet MS', 'Poor Richard', 'Impact', 'Californian FB', 'Bodoni MT', 'Yet R', 'Tw Cen MT', 'HCR Batang ExtB', 'Perpetua Titling MT', 'Garamond', 'Cambria', 'Microsoft Yi Baiti', 'Segoe MDL2 Assets', 'Gadugi', 'Comic Sans MS', 'Agency FB', 'Webdings', 'HYMyeongJo-Extra', 'Cambria', 'Gabriola', 'Century Schoolbook', 'Kristen ITC', 'Wingdings 3', 'Myanmar Text', 'Calibri', 'Calisto MT', 'Century Gothic', 'Dubai', 'Centaur', 'Arial', 'Perpetua', 'Segoe UI Historic', 'MT Extra', 'Bookman Old Style', 'Script MT Bold', 'Corbel', 'Comic Sans MS', 'Constantia', 'Lucida Sans', 'Comic Sans MS', 'High Tower Text', 'Tw Cen MT Condensed Extra Bold', 'Bookshelf Symbol 7', 'Microsoft YaHei', 'Juice ITC', 'Trebuchet MS', 'Bell MT', 'Arial', 'Maiandra GD', 'Segoe UI', 'MS Outlook', 'Dubai', 'Algerian', 'Blackadder ITC', 'Lucida Console', 'Pyunji R', 'Bodoni MT', 'Rockwell Condensed', 'Franklin Gothic Book', 'Leelawadee UI', 'Courier New', 'Imprint MT Shadow', 'Playbill', 'Leelawadee', 'Segoe Script', 'Perpetua', 'Symbol', 'Lucida Bright', 'Snap ITC', 'Constantia', 'Calibri', 'Bodoni MT', 'Consolas', 'Tw Cen MT Condensed', 'Bodoni MT', 'Matura MT Script Capitals', 'SimSun', 'Book Antiqua', 'Goudy Old Style', 'Century Schoolbook', 'Century Gothic', 'Nirmala UI', 'Bodoni MT', 'Courier New', 'Kunstler Script', 'Microsoft JhengHei', 'New Gulim', 'Elephant', 'Microsoft JhengHei', 'Candara', 'Brush Script MT', 'Rage Italic', 'Corbel', 'LG Smart UI', 'Franklin Gothic Book', 'Segoe Script', 'Calibri', 'Niagara Solid', 'Franklin Gothic Demi', 'Britannic Bold', 'KBIZmjo M', 'LG Smart UI', 'Century Gothic', 'Sitka Small', 'Courier New', 'Headline R', 'Copperplate Gothic Bold', 'Franklin Gothic Medium', 'Cambria', 'HCR Dotum Ext', 'Franklin Gothic Medium', 'HYHeadLine-Medium', 'Lucida Handwriting', 'Microsoft Uighur', 'Goudy Old Style', 'MT Extra', 'Dubai', 'Gloucester MT Extra Condensed', 'Elephant', 'Bauhaus 93', 'Felix Titling', 'Segoe UI', 'Pristina', 'Gill Sans MT', 'Microsoft YaHei', 'Segoe UI', 'Sylfaen', 'Malgun Gothic', 'Gill Sans MT', 'Lucida Sans Unicode', 'Lucida Bright', 'Chiller', 'Californian FB', 'Segoe UI', 'Tw Cen MT', 'HCR Dotum', 'Haettenschweiler', 'Agency FB', 'Harrington', 'MJemokGothic', 'Lucida Bright', 'Franklin Gothic Heavy', 'High Tower Text', 'Book Antiqua', 'Dubai', 'Garamond', 'Lucida Sans Typewriter', 'MS Reference Specialty', 'Gadugi', 'Eras Medium ITC', 'MS Reference Sans Serif', 'Segoe UI Symbol', 'Garamond', 'Comic Sans MS', 'Calisto MT', 'Jokerman', 'Microsoft Himalaya', 'Vivaldi', 'Berlin Sans FB Demi', 'Gigi', 'Calibri', 'Engravers MT', 'HYGungSo-Bold', 'Lucida Sans Typewriter', 'Papyrus', 'Perpetua Titling MT', 'Corbel', 'Ink Free', 'Verdana', 'Tahoma', 'Tw Cen MT', 'Informal Roman', 'Candara', 'Leelawadee', 'Candara', 'Lucida Fax', 'Century Schoolbook', 'Bookman Old Style', 'Georgia', 'Gill Sans Ultra Bold', 'HCR Batang', 'Old English Text MT', 'Ebrima', 'Lucida Sans', 'KBIZmjo R', 'HYSinMyeongJo-Medium', 'Candara', 'Times New Roman', 'Berlin Sans FB', 'Forte', 'Corbel', 'Modern No. 20', 'Arial', 'Microsoft Tai Le', 'Segoe UI', 'HYGothic-Extra', 'Bell MT', 'Palatino Linotype', 'Palace Script MT', 'Arial', 'Bodoni MT', 'Corbel', 'Candara', 'Microsoft YaHei', 'Lucida Fax', 'Gill Sans MT Condensed', 'HYPost-Light', 'Consolas', 'Onyx', 'Microsoft Sans Serif', 'Lucida Sans', 'Segoe UI', 'Berlin Sans FB', 'Copperplate Gothic Light', 'Bookman Old Style', 'HYGraphic-Medium', 'Segoe UI', 'Broadway', 'Century', 'Calibri', 'Segoe UI', 'MSugiJeong', 'Curlz MT', 'Viner Hand ITC', 'Malgun Gothic', 'Courier New', 'Bodoni MT', 'MS Gothic', 'MBatang', 'Arial', 'Rockwell Extra Bold', 'Bernard MT Condensed', 'Microsoft New Tai Lue', 'Microsoft PhagsPa', 'Palatino Linotype', 'Georgia', 'Century Gothic', 'Book Antiqua', 'Gill Sans Ultra Bold Condensed', 'Marlett', 'Malgun Gothic', 'LG Smart UI', 'Javanese Text', 'Bodoni MT', 'Rockwell Condensed', 'Rockwell', 'Microsoft PhagsPa', 'Arial', 'Lucida Sans Typewriter', 'HCR Dotum', 'Showcard Gothic', 'Sitka Small', 'Freestyle Script', 'HoloLens MDL2 Assets', 'Myanmar Text', 'Times New Roman', 'MV Boli', 'HYShortSamul-Medium', 'Segoe Print', 'HYPost-Medium', 'Sitka Small', 'Perpetua', 'Lucida Fax', 'Ravie', 'NewJumja', 'MJemokBatang', 'Candara', 'Rockwell', 'Times New Roman', 'Parchment', 'Yu Gothic', 'Mistral', 'Rockwell', 'Microsoft New Tai Lue', 'Harlow Solid Italic', 'Baskerville Old Face', 'Leelawadee UI', 'Yu Gothic', 'Segoe UI', 'LG Smart UI', 'Arial', 'Tempus Sans ITC', 'Goudy Old Style', 'SimSun-ExtB', 'Niagara Engraved', 'Wingdings', 'Segoe UI Emoji', 'Verdana', 'Microsoft Tai Le', 'Segoe UI', 'MDotum', 'Bradley Hand ITC', 'Lucida Sans', 'Monotype Corsiva', 'Bodoni MT', 'Eras Light ITC', 'Franklin Gothic Demi Cond', 'Verdana', 'Lucida Calligraphy']

df\_입학생 = pd.read\_excel('C:/R/git/datavisualization/chap3/2021\_연도별\_입학자수.xlsx', sheet\_name='Sheet0', header=(1, 2), nrows=400)

## C:\ANACON~1\lib\site-packages\openpyxl\styles\stylesheet.py:226: UserWarning: Workbook contains no default style, apply openpyxl's default  
## warn("Workbook contains no default style, apply openpyxl's default")

df\_입학생.head()

## 연도 시도 입학자수 ... 석사 박사   
## 연도 시도 계 여 ... 계 여 계 여  
## 0 1999 전체 795998.0 356583.0 ... 73826.0 26647.0 10447.0 2536.0  
## 1 2000 전체 819779.0 381957.0 ... 82374.0 32487.0 11705.0 3099.0  
## 2 2001 전체 839516.0 393342.0 ... 86992.0 35937.0 12570.0 3547.0  
## 3 2002 전체 819052.0 391921.0 ... 89557.0 38640.0 13227.0 4034.0  
## 4 2003 전체 776291.0 374483.0 ... 91178.0 40813.0 13310.0 4328.0  
##   
## [5 rows x 32 columns]

df\_입학생.info()

## <class 'pandas.core.frame.DataFrame'>  
## RangeIndex: 400 entries, 0 to 399  
## Data columns (total 32 columns):  
## # Column Non-Null Count Dtype   
## --- ------ -------------- -----   
## 0 (연도, 연도) 400 non-null int64   
## 1 (시도, 시도) 400 non-null object   
## 2 (입학자수, 계) 400 non-null float64  
## 3 (입학자수, 여) 400 non-null float64  
## 4 (전문대학, 계) 400 non-null float64  
## 5 (전문대학, 여) 400 non-null float64  
## 6 (교육대학, 계) 400 non-null float64  
## 7 (교육대학, 여) 400 non-null float64  
## 8 (일반대학, 계) 400 non-null float64  
## 9 (일반대학, 여) 400 non-null float64  
## 10 (방송통신대학, 계) 400 non-null float64  
## 11 (방송통신대학, 여) 400 non-null float64  
## 12 (산업대학, 계) 400 non-null float64  
## 13 (산업대학, 여) 400 non-null float64  
## 14 (기술대학, 계) 400 non-null float64  
## 15 (기술대학, 여) 400 non-null float64  
## 16 (각종학교, 계) 400 non-null float64  
## 17 (각종학교, 여) 400 non-null float64  
## 18 (원격및사이버대학, 계) 400 non-null float64  
## 19 (원격및사이버대학, 여) 400 non-null float64  
## 20 (사내대학, 계) 400 non-null float64  
## 21 (사내대학, 여) 400 non-null float64  
## 22 (전공대학, 계) 400 non-null float64  
## 23 (전공대학, 여) 400 non-null float64  
## 24 (기능대학, 계) 400 non-null float64  
## 25 (기능대학, 여) 400 non-null float64  
## 26 (계, 계) 400 non-null float64  
## 27 (계, 여) 400 non-null float64  
## 28 (석사, 계) 400 non-null float64  
## 29 (석사, 여) 400 non-null float64  
## 30 (박사, 계) 400 non-null float64  
## 31 (박사, 여) 400 non-null float64  
## dtypes: float64(30), int64(1), object(1)  
## memory usage: 100.1+ KB

df\_입학생.describe()

## 연도 입학자수 ... 박사   
## 연도 계 ... 계 여  
## count 400.000000 400.00000 ... 400.000000 400.000000  
## mean 2010.157500 89694.30500 ... 2354.175000 907.470000  
## std 6.659472 177082.59102 ... 5175.558116 2102.481449  
## min 1999.000000 5106.00000 ... 39.000000 5.000000  
## 25% 2004.000000 25671.50000 ... 378.000000 113.750000  
## 50% 2010.000000 33323.50000 ... 671.500000 240.500000  
## 75% 2016.000000 55345.00000 ... 1223.000000 478.750000  
## max 2021.000000 839516.00000 ... 29594.000000 13251.000000  
##   
## [8 rows x 31 columns]

df\_입학생.loc[1:5]

## 연도 시도 입학자수 ... 석사 박사   
## 연도 시도 계 여 ... 계 여 계 여  
## 1 2000 전체 819779.0 381957.0 ... 82374.0 32487.0 11705.0 3099.0  
## 2 2001 전체 839516.0 393342.0 ... 86992.0 35937.0 12570.0 3547.0  
## 3 2002 전체 819052.0 391921.0 ... 89557.0 38640.0 13227.0 4034.0  
## 4 2003 전체 776291.0 374483.0 ... 91178.0 40813.0 13310.0 4328.0  
## 5 2004 전체 773741.0 372952.0 ... 88763.0 41092.0 14494.0 4917.0  
##   
## [5 rows x 32 columns]

df\_입학생.shape[0]

## 400

df\_입학생.index

## RangeIndex(start=0, stop=400, step=1)

df\_입학생 = df\_입학생 >> select(0, 1, 2, 4, 8, 28, 30)  
df\_입학생.columns = ['연도', '시도', '전체', '전문대', '일반대', '석사', '박사']  
  
df\_입학생\_전체 = df\_입학생 >> filter\_by(X.시도 == '전체')  
df\_입학생\_전체.info()

## <class 'pandas.core.frame.DataFrame'>  
## Int64Index: 23 entries, 0 to 22  
## Data columns (total 7 columns):  
## # Column Non-Null Count Dtype   
## --- ------ -------------- -----   
## 0 연도 23 non-null int64   
## 1 시도 23 non-null object   
## 2 전체 23 non-null float64  
## 3 전문대 23 non-null float64  
## 4 일반대 23 non-null float64  
## 5 석사 23 non-null float64  
## 6 박사 23 non-null float64  
## dtypes: float64(5), int64(1), object(1)  
## memory usage: 1.4+ KB

plt.rcParams['axes.unicode\_minus'] = False  
plt.rcParams['font.family'] = 'Malgun Gothic'  
sns.lineplot(x='연도', y='전체', data=df\_입학생\_전체)  
plt.show()

