

Welcome

Prof. Eric A. Suess

August 21, 2019

Welcome

What is Data Science?

The best description of what Data Science is given by Wickham and Grolemund in their book R for Data Science.

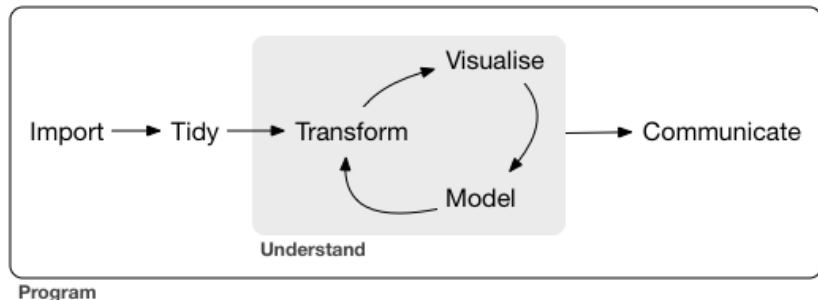


Figure 1: Data Science

Software

If you have not already done so you should

1. Download and install R from [r-project.org](https://www.r-project.org)
2. Download and install RStudio [rstudio.com](https://www.rstudio.com)
3. Get an RStudio Cloud account [RStudio Cloud](https://www.rstudio.com/cloud/)

Start reading the book

Read Chapter 1 What is data science?

Read Chapter 2 and 3 to get strart with ggplot2

Read Chapter 27 to become familiar with R Notebooks and R Markdown.

Install the packages

```
install.packages("tidyverse")
```

Reproducible Research

We will be practicing Reproducible Research using R Markdown and R Projects.

Or as Hadley Wickham describes it, start “Living a reproducible life.”

1. R Scripts
2. R Notebooks
3. R Projects

Start R and RStudio

Try all of the code in Chapter 3.

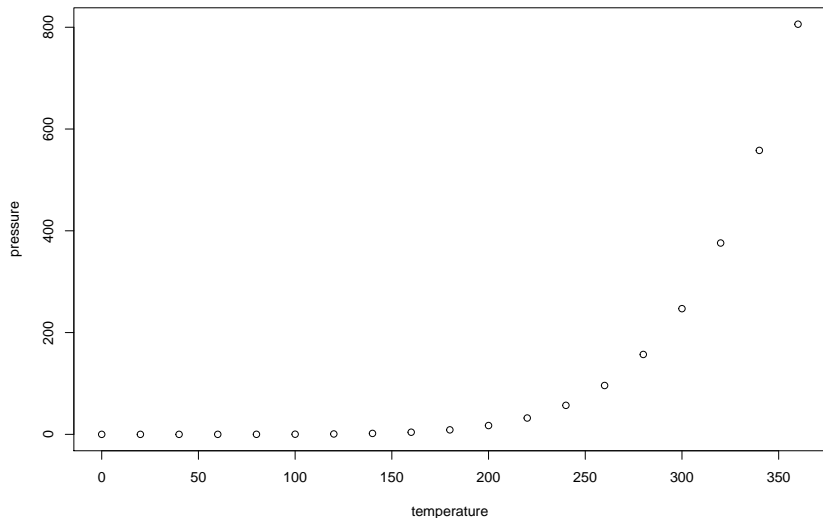
Slide With Code

```
summary(cars)
```

##	speed	dist
##	Min. : 4.0	Min. : 2.00
##	1st Qu.:12.0	1st Qu.: 26.00
##	Median :15.0	Median : 36.00
##	Mean :15.4	Mean : 42.98
##	3rd Qu.:19.0	3rd Qu.: 56.00
##	Max. :25.0	Max. :120.00

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.