

# Making a Case

HN

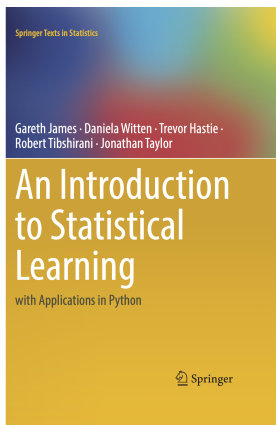
Washington State University

January 26, 2025



- Anything worth making needs patience and practice
- If you want mastery, you need to immerse yourself in it till it becomes your second nature, just like walking
- Top(?) two take aways from graduate school
  - 1 Try not to have a bias and be playful 🔍
  - 2 When you are wrong, you are wrong ⚠️

# Source of This Notes



## Notes from **An Introduction to Statistical Learning** with Applications in Python

- It's an easy read and as the name suggests, it's just Introduction. Good for intuition building
- Its PDF is available for free from the Authors' website.
- Sign-up & receive coupons (30%-50% off)

# Last thing first (There is no magic wand)

- There is no unique tool that outperforms other methods (there is no free lunch)

# Last thing first (There is no magic wand)

- There is no unique tool that outperforms other methods (there is no free lunch)
- There is no tool that is both accurate (in predicting) AND interpretable (unless we get lucky?!)

# Last thing first (There is no magic wand)

- There is no unique tool that outperforms other methods (there is no free lunch)
- There is no tool that is both accurate (in predicting) AND interpretable (unless we get lucky?!)

# Last thing first (There is no magic wand)

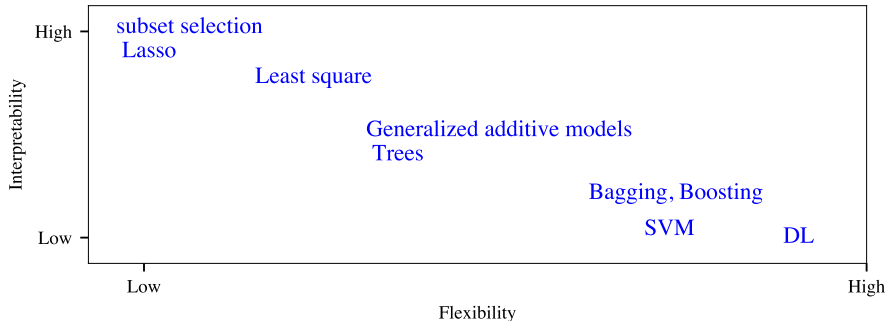
- There is no unique tool that outperforms other methods (there is no free lunch)
- There is no tool that is both accurate (in predicting) AND interpretable (unless we get lucky?!)

**define your goal and deal with it**

# Last thing first (There is no magic wand)

- There is no unique tool that outperforms other methods (there is no free lunch)
- There is no tool that is both accurate (in predicting) AND interpretable (unless we get lucky?!)

**define your goal and deal with it**





- True model  $y = f(x) + \varepsilon$

- True model  $y = f(x) + \varepsilon$
- Our model  $y = \beta_0 + \beta_1 x + \varepsilon$

- True model  $y = f(x) + \varepsilon$
- Our model  $y = \beta_0 + \beta_1 x + \varepsilon$
- The error term ( $\sim N(0, \sigma^2)$ ) is a catch-all for what we miss with this simple model: measurement errors, false modeling, omitted variables

- True model  $y = f(x) + \varepsilon$
- Our model  $y = \beta_0 + \beta_1 x + \varepsilon$
- The error term ( $\sim N(0, \sigma^2)$ ) is a catch-all for what we miss with this simple model: measurement errors, false modeling, omitted variables
- RSS (or  $RSE = \sqrt{RSS/(n-2)}$ ) depends on Y-units.

- True model  $y = f(x) + \varepsilon$
- Our model  $y = \beta_0 + \beta_1 x + \varepsilon$
- The error term ( $\sim N(0, \sigma^2)$ ) is a catch-all for what we miss with this simple model: measurement errors, false modeling, omitted variables
- RSS (or  $\text{RSE} = \sqrt{\text{RSS}/(n-2)}$ ) depends on Y-units.
- $R^2$  is unitless.

- True model  $y = f(x) + \varepsilon$
- Our model  $y = \beta_0 + \beta_1 x + \varepsilon$
- The error term ( $\sim N(0, \sigma^2)$ ) is a catch-all for what we miss with this simple model: measurement errors, false modeling, omitted variables
- RSS (or  $RSE = \sqrt{RSS/(n-2)}$ ) depends on Y-units.
- $R^2$  is unitless.
- Small  $R^2$  “might occur because the linear model is wrong, or the error variance  $\sigma^2$  is high, or both.”

- True model  $y = f(x) + \varepsilon$
- Our model  $y = \beta_0 + \beta_1x + \varepsilon$
- The error term ( $\sim N(0, \sigma^2)$ ) is a catch-all for what we miss with this simple model: measurement errors, false modeling, omitted variables
- RSS (or  $RSE = \sqrt{RSS/(n-2)}$ ) depends on Y-units.
- $R^2$  is unitless.
- Small  $R^2$  “might occur because the linear model is wrong, or the error variance  $\sigma^2$  is high, or both.”
- “...and an  $R^2$  value well below 0.1 might be more realistic!”  
(page 79)

# More Resources

- Guesstimation: (Weinstein and Adam, [2008](#))
- First Course in Probability (Ross et al., [1976](#))
- Introduction to Linear Regression Analysis (Montgomery, Peck, and Vining, [2021](#))

## Non-technical

- Excellent Sheep: The Miseducation of the American Elite and the Way to a Meaningful Life (Deresiewicz, [2014](#))
- The Culture Code: The Secrets of Highly Successful Groups (Coyle, [2018](#))

## Tools

- Software Carpentry has lots of tutorials including [GitHub](#)!








# Some Definition I

## Definition (Variance of a method)

*Variance refers to the amount by which  $\hat{f}$  would change if we estimated it using a different training data set.*

# Bibliography I

-  Coyle, D. (2018). *The Culture Code: The Secrets of Highly Successful Groups*. Bantam Books, an imprint of Random House, a division of Penguin Random House LLC New York.
-  Deresiewicz, W. (2014). *Excellent Sheep: The Miseducation of the American Elite and the Way to a Meaningful Life*. Free Press.
-  Montgomery, D. C., Peck, E. A., and Vining, G. G. (2021). *Introduction to Linear Regression Analysis*. John Wiley & Sons.
-  Ross, S. M. et al. (1976). *A first course in probability*. Vol. 2. Macmillan New York.
-  Weinstein, L. and Adam, J. A. (2008). *Guesstimation: Solving the world's problems on the back of a cocktail napkin*. Princeton University Press.