# Introduction(week 1, part 3)

6 Jan 2023

## Table of contents



### books

- ESL and ADA cover very similar material
- both compare linear regression and nearest-neighbour methods as opposite ends of a specrum

### Fisher's irises

- Canadian content (irises of the Gaspé peninsula) (Fisher 1936)
- Fisher was a eugenist (Bodmer et al. 2021)
- multiple versions/errors! (Bezdek et al. 1999)
- alternative: Palmer penguins dataset




Fisher, R. A. 1936. "The Use of Multiple Measurements in Taxonomic Problems." *Annals of Eugenics* 7 (2): 179–88. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x.

Bodmer, Walter, R. A. Bailey, Brian Charlesworth, Adam Eyre-Walker, Vernon Farewell, Andrew Mead, and Stephen Senn. 2021. "The Outstanding Scientist, R.A. Fisher: His Views on Eugenics and Race." *Heredity* 126 (4): 565–76. https://doi.org/10.1038/s41437-020-00394-6.

Bezdek, J. C., J. M. Keller, R. Krishnapuram, L. I. Kuncheva, and N. R. Pal. 1999. "Will the Real Iris Data Please Stand Up?" *IEEE Transactions on Fuzzy Systems* 7 (3): 368–69. https://doi.org/10.1109/91.771092.

## linear models

- can write out as $\hat{Y} = \hat{\beta}_0 + \sum X_j \hat{\beta}_j$
- go almost immediately to $\hat{Y} = X^\top \hat{\beta}$ or $\langle X, \beta \rangle$ or $\mathbf{X}\beta$

## bias-variance expansion

## linear regression

## nearest-neighbor

## dimensionality

## references