

Statistics For Data Science, 2024

Topic Schedule and Suggested Readings

Suggested Resources

- [OS] OpenIntro Statistics, 4th Edition: <https://www.openintro.org/book/os/>
- [OMS] OpenIntro Introduction to Modern Statistics: <https://www.openintro.org/book/ims/>
- [TS] Think Stats: Probability and Statistics for Programmers: <https://greenteapress.com/thinkstats/>
- [MD] Modern Dive: Statistical Inference via Data Science <https://moderndive.com/index.html>

Weekly Schedule and References

Week 1: Introduction to Statistics, Descriptive Statistics, and Single-Variable EDA

- OS, Sections 1.1-1.3, 2.1 - 2.2
- TS, Sections 2.1 - 2.9

Week 2: Multi-variable EDA

- OS, Sections 2.1 - 2.2
- TS, Sections 9.2 - 9.3

Week 3: Probability and Random Variables (Binomial and Normal Distributions)

- OS, Sections 3.1 - 3.2; 3.4 - 3.5; 4.1, 4.3
- TS, Chapter 5

Week 4: Estimation, Sampling Distributions, Confidence Intervals, and the (Empirical) Bootstrap

- OS, Section 5.1, 5.2, 7.1
- OMS, Chapter 12
- MD, Chapters 7 and 8

Week 5: Hypothesis Testing - Testing for means and proportions; permutation testing

- OS, Section 5.3, 6.2, 7.3
- OMS, Chapter 11
- MD, Chapter 9

Week 6: Linear Regression

- OS, Sections 8.1, 8.2, 8.4, 9.1 - 9.3

Week 7: Logistic Regression

- OS, Section 9.5