

DATA 140



Spring 2023

WEEK 12 STUDY GUIDE

The Big Picture

One of the rare weeks when we just cover one chapter:

The least squares predictor of one variable given another, and the error in it

- If you have the scatter diagram of simulated (X, Y) pairs, then Data 8 ideas say that given X , the best predictor of Y is the “center of the vertical strip at X .”
- Formally, “best” means “least squares,” and the “center of the vertical strip at X ” is the conditional expectation of Y given X .
- The error in this estimate, given X , is the conditional SD of Y given X .
- This allows us to decompose the variance of Y into two easier pieces, by conditioning on X .

Week At a Glance

Mon 4/10	Tue 4/11	Wed 4/12	Thu 4/13	Fri 4/14
	Lecture	Section	Lecture	Section
Lab 6B Due				
HW 11 Due HW 12 (Due Mon 4/17)				HW 12 Party 3 PM to 5 PM
Skim Sections 22.1-22.2	Work through Sections 22.1-22.2	Skim Sections 22.3-22.4	Work through Chapter 22	Work through Chapter 22; study for Quiz 3

Reading, Practice, and Class Meeting

Book	Topic	Lectures: Prof. A.	Sections: GSIs	Optional Additional Practice
Ch 22	An approach to prediction - 22.1 develops the main reason why conditional expectation is important for prediction - 22.2 shows that conditional expectation is a least squares predictor, and defines the error in the estimate - 22.3 decomposes variance into two pieces, by conditioning - 22.4 is a series of examples of varied uses of the method of 22.3	Tuesday 4/11 - The random variable equivalent of “dropping a perpendicular” - Least squares prediction, and a new variance	Wednesday 4/12 - Ch 22 Ex 6, 1, 7	Ch 22 - 2
		Thursday 4/13 - Variance by conditioning - Examples, including a look back at Section 9.2	Friday 4/14 - Ch 22 Ex 5, 3, 4	