

STATS 3N03/3504

2004.09.22

4-1

PLOTS

ONE CONTINUOUS VARIABLE

DOT PLOT

BOX PLOT

STEM & LEAF PLOT

HISTOGRAM

TWO CONTINUOUS VARIABLES

SCATTER PLOT

MORE THAN TWO

SCATTER PLOT MATRIX

ONE CONTINUOUS, ONE CATEGORICAL

COMPARATIVE BOX PLOT

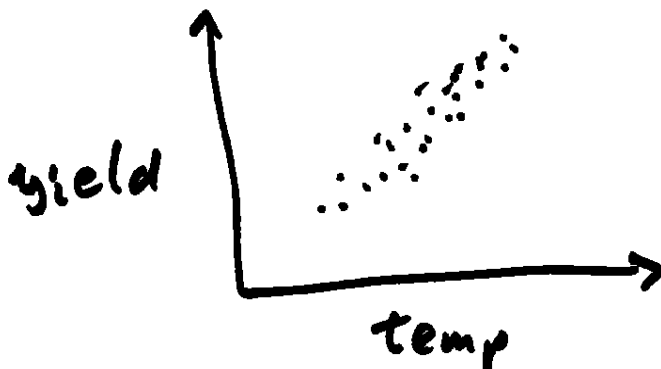
TIME SERIES

TIME SERIES PLOT

LAG PLOT

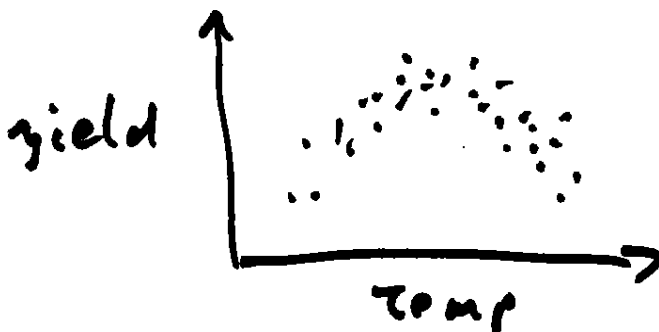
4-2

SCATTER PLOT

plot($x = \text{temp}$, $y = \text{yield}$)

"YIELD IS A LINEAR FUNCTION OF TEMPERATURE, OVER THE RANGE OF TEMPERATURES STUDIED"

⇒ COULD FIT A STRAIGHT LINE, USE IT TO PREDICT YIELD AT A GIVEN TEMPERATURE



"YIELD IS A NON-LINEAR FUNCTION OF TEMPERATURE OVER THE RANGE OF TEMPERATURE STUDIED, AND THERE IS A TEMPERATURE THAT OPTIMIZES YIELD"

4-3

⇒ FIND THE OPTIMAL TEMPERATURE



"THERE IS NO EVIDENCE FROM THESE DATA THAT TEMPERATURE AFFECTS YIELD, OVER THE RANGE OF TEMPERATURES STUDIED"

STEM & LEAF PLOTS

EX 1, 3, 4, 7, 8, 6, 7, 4, 3, 5, 2

UNORDERED:

	STEM	LEAVES	
01 02	00, 01	0	1
02 04	02, 03	0	3 3 2
03 06	04, 05	0	4 4 5
04 08	06, 07	0	6 6 7
05 09	08, 09	0	8

STEM: TENS
LEAF: UNITS

4-4

ORDERED →

F	f	STEM	LEAVES	F	f	STEM	LEAVES
1	1	0	1	6	6	0	123344
4	3	0	233	11	5	0	56778
7	3	0	445				
10	3	0	677				
11	1	0	8				

SCATTERPLOT MATRICES

SEE: TEST #1 SOLUTIONS
ON WEBS 2002, 2003

EX trees DATA FRAME CATEGORICAL
↓
ht dbh bands cci quality

pairs(trees) ⇒ 5x5 SPLOM

pairs(trees[, -5]) ⇒ OMIT quality

TO CODE POINTS BY quality:

col = as.numeric(trees\$quality)

pch = as.numeric(trees\$quality)