

STATS 3N03/3J04

2004-09-20

3-1

DROP-IN COMPUTER HELP

RSB-241

MON-THURS 7:30 - 8:30 PM

THIS WEEK & NEXT WEEK

WORK ON EX 1, ASS. 1.

ASSIGNMENT 1

ON WEB BY THE END OF  
THE DAY

TEST 1

STATS 3N03

MON OCT 4 7-9 PM

RSB-241, 242, 244, 245

STATS 3J04

TUE OCT 5 7-9 PM

RSB-241, 244, 245

MAKEUP

WED OCT 6 8:30 - 10:30 AM

RSB-241

3-2

IF YOU CAN'T COME ON MON  
- COME ON TUE.

IF YOU CAN'T COME ON TUE  
- COME ON WED, NOT MON.

---

DATA EDITOR (WINDOWS ONLY)  
edit (prices)

ONLINE HELP

help (boxplot)

? boxplot

help.start()

OPENS HTML HELP  
IN YOUR BROWSER

WRITING & EDITING YOUR FUNCTIONS

options (keep.source = F)

- AT START OF SESSION

- FUNCTION GETS "PRETTIED UP"  
WHEN YOU EXIT THE  
FUNCTION EDITOR

3-3

```

> options(keep.source=F)
> 3*6
[1] 18
> aa <- 3*6
> aa
[1] 18
> by3 <- function(x) 3*x
> by3(5)
[1] 15
> by3(aa)
[1] 54
> by3()
Error in by3() : Argument "x" is missing, with no default
> by3
function (x)
  3 * x
> fix(by3)
> by3
function (x = 1)
  3 * x
> by3()
[1] 3
> fitbricks <- lm(density~temp, data=bricks)
> summary(fitbricks)

```

← CREATE A SIMPLE 1-LINE FUNCTION

← OPENS TEXT EDITOR; EDIT, SAVE, QUIT

← REVISED FUNCTION

← NOW PROVIDES DEFAULT TO USE WHEN ARGUMENT NOT GIVEN

Call:

lm(formula = density ~ temp, data = bricks)

Residuals:

	Min	1Q	Median	3Q	Max
	-0.28092	-0.11936	0.03006	0.09711	0.24104

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	2.157e+01	1.482e-01	145.571	<2e-16 ***
temp	8.786e-04	1.061e-03	0.828	0.418

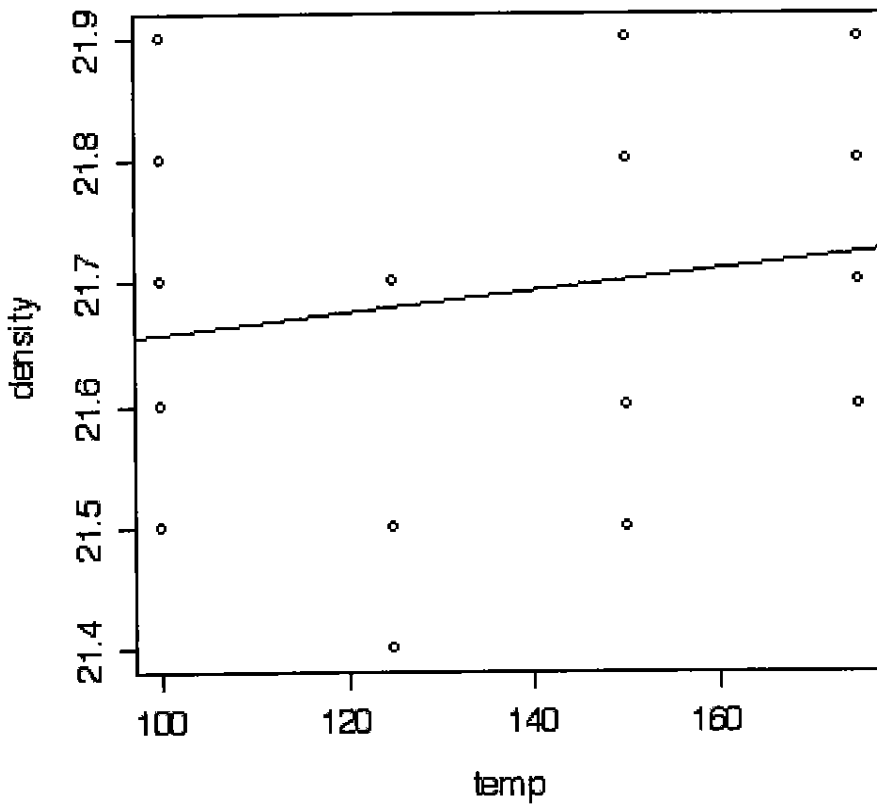
---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1488 on 20 degrees of freedom  
 Multiple R-Squared: 0.03312, Adjusted R-squared: -0.01522  
 F-statistic: 0.6851 on 1 and 20 DF, p-value: 0.4176

3-4

```
> coef(fitbricks)
(Intercept)      temp
2.157110e+01 8.786127e-04
> plot(density~temp, data=bricks)
> abline(fitbricks)  - ADD TO EXISTING PLOT
abline(fitbricks$coef[1], fitbricks$coef[2])
```



```
fitbricks
```

```
print(fitbricks)
```

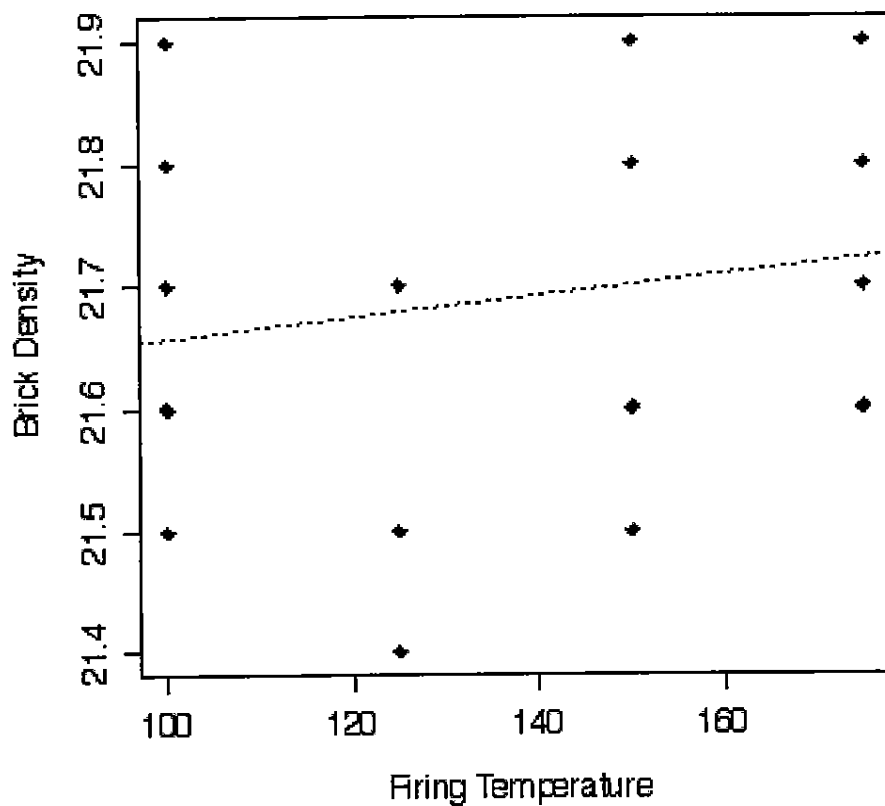
```
as.list(fitbricks)
```

```
names(fitbricks)
```

```
abline(coef(fitbrick)[1], coef(fitbrick)[2])
```

3-5

```
> plot(density~temp, data=bricks, col="blue",
pch=18,xlab="Firing Temperature",ylab="Brick Density")
> abline(fitbricks,col="red",lty="2")
```



LOG SCALES:

log = "x"  
 "y"  
 "xy"

plot() STARTS NEW PLOT

points() ADDS POINTS TO EXISTING PLOT

lines() " LINES " " " "

type = "l" LINES (l NOT 1)

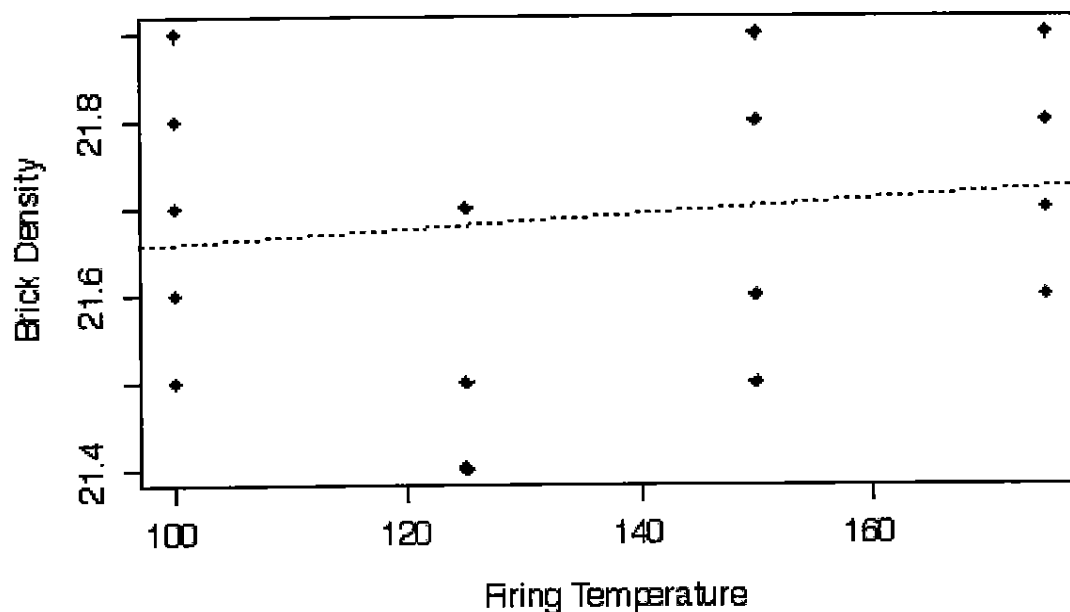
"p" POINTS

"b" BOTH

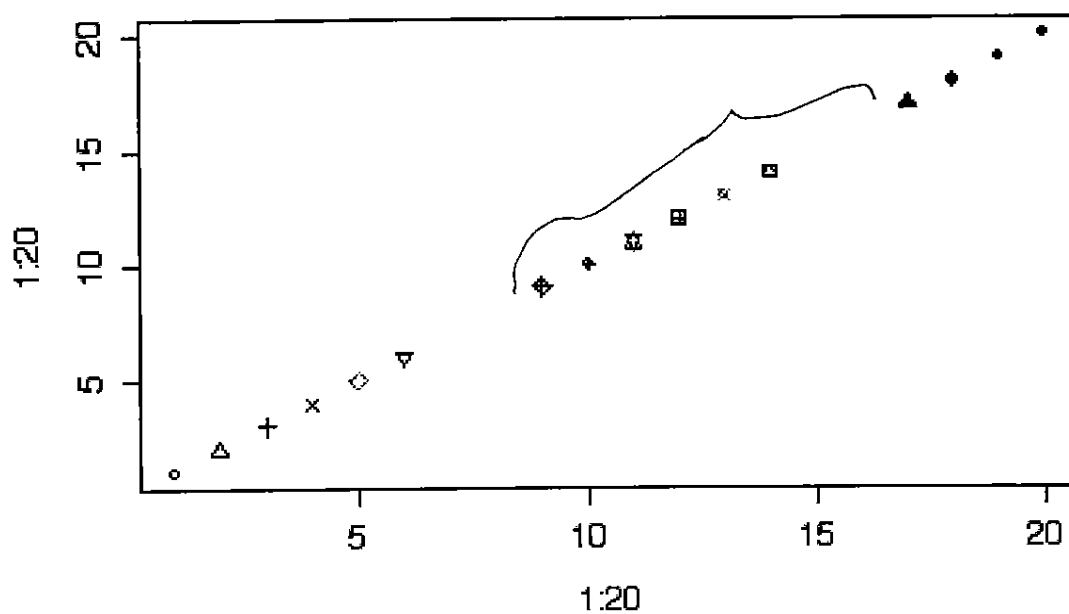
"n" NOTHING

3-6

```
> macintosh(h=4,w=6)
> plot(density~temp, data=bricks, col="blue",
pch=18,xlab="Firing Temperature",ylab="Brick Density")
> abline(fitbricks,col="red",lty="2")
```



```
> plot(1:20,1:20,pch=1:20,col=1:20)
```



3-7

## IMPORTING DATA INTO A DATA FRAME:

- ENTER DATA IN EXCEL OR IMPORT INTO EXCEL
- MAKE SURE THERE IS ONLY DATA AND COLUMN HEADERS IN WORKSHEET
- MISSING VALUES CAN BE EMPTY CELLS
- SAVE AS TAB-DELIMITED TEXT bricks.txt
- OPEN IN NOTEPAD TO CHECK FOR BAD CHARACTERS, CORRECT FIRST AND LAST LINES

file =

```
bricks ← read.table("bricks.txt",  
header = T, sep = "\t", fill = T)
```

J-8

sep = "\t"

COLS. SEPARATED BY  
TABS

sep = " "

SEPARATED BY ANY  
WHITE SPACE  
(WON'T WORK IF MISSING  
VALUES ARE BLANK COLS)

fill = T

REQUIRED IF THERE  
ARE BLANK CELLS FOR  
MISSING VALUES AT  
THE ENDS OF LINES

MISSING VALUE CODES (NUMERIC)

EXCEL

BLANK

MINITAB

\*

SPSS, SAS

.

R, Splus

NA

IF YOU CODED MISSING VALUES  
WITH \* ADD:

na.strings = "\*"



3-9

```

> bricks
  density temp
1    21.8  100
2    21.9  100
3    21.7  100
4    21.6  100
5    21.7  100
6    21.5  100
7    21.8  100
8    21.7  125
9    21.4  125
10   21.5  125
11   21.5  125
12   21.9  150
13   21.8  150
14   21.8  150
15   21.6  150
16   21.5  150
17   21.9  175
18   21.7  175
19   21.8  175
20   21.7  175
21   21.6  175
22   21.8  175
> bricks$dens
 [1] 21.8 21.9 21.7 21.6 21.7 21.5 21.8 21.7 21.4 21.5 21.5 21.9 21.8 21.8
 [15] 21.6 21.5 21.9 21.7 21.8 21.7 21.6 21.8
> bricks$temp
 [1] 100 100 100 100 100 100 100 100 125 125 125 125 150 150 150 150 175
175
 [19] 175 175 175 175
> split(bricks$density, bricks$temp)
$"100"
 [1] 21.8 21.9 21.7 21.6 21.7 21.5 21.8

$"125"
 [1] 21.7 21.4 21.5 21.5

$"150"
 [1] 21.9 21.8 21.8 21.6 21.5

$"175"
 [1] 21.9 21.7 21.8 21.7 21.6 21.8

> boxplot(split(bricks$density, bricks$temp))
> boxplot(density~temp, data=bricks)

```

DATA  
FRAME

- CAN INTERPRET AS MATRIX LIST