

Module 7: Assignment

Aaron Hum

Stephanie Moyerman, PhD

LSC 541: Statistics for Biological Data Science I

2024-08-08

```
install.packages(c("lme4", "ggplot2", "knitr", "rmarkdown",  
"reshape2"))
```

```
Installing packages into '/usr/local/lib/R/site-library'  
(as 'lib' is unspecified)
```

```
also installing the dependencies 'minqa', 'nloptr', 'RcppEigen',  
'plyr'
```

```
# Load required libraries
```

```
library(lme4)  
library(ggplot2)  
library(knitr)  
library(rmarkdown)  
library(reshape2)
```

```
Loading required package: Matrix
```

```
# Read the data
```

```
data <- read.table('tlc-data.txt', header = FALSE, col.names =  
c("SubjectID", "Group", "Baseline", "Week1", "Week4", "Week6"))  
data
```

	SubjectID	Group	Baseline	Week1	Week4	Week6
1	1	P	30.8	26.9	25.8	23.8
2	2	A	26.5	14.8	19.5	21.0
3	3	A	25.8	23.0	19.1	23.2
4	4	P	24.7	24.5	22.0	22.5
5	5	A	20.4	2.8	3.2	9.4
6	6	A	20.4	5.4	4.5	11.9
7	7	P	28.6	20.8	19.2	18.4
8	8	P	33.7	31.6	28.5	25.1
9	9	P	19.7	14.9	15.3	14.7
10	10	P	31.1	31.2	29.2	30.1
11	11	P	19.8	17.5	20.5	27.5

12	12	A	24.8	23.1	24.6	30.9
13	13	P	21.4	26.3	19.5	19.0
14	14	A	27.9	6.3	18.5	16.3
15	15	P	21.1	20.3	18.4	20.8
16	16	P	20.6	23.9	19.0	17.0
17	17	P	24.0	16.7	21.7	20.3
18	18	P	37.6	33.7	34.4	31.4
19	19	A	35.3	25.5	26.3	30.3
20	20	A	28.6	15.8	22.9	25.9
21	21	P	31.9	27.9	27.3	34.2
22	22	A	29.6	15.8	23.7	23.4
23	23	A	21.5	6.5	7.1	16.0
24	24	P	26.2	26.8	25.3	24.8
25	25	A	21.8	12.0	16.8	19.2
26	26	A	23.0	4.2	4.0	16.2
27	27	A	22.2	11.5	9.5	14.5
28	28	P	20.5	21.1	17.4	21.1
29	29	A	25.0	3.9	12.8	12.7
30	30	P	33.3	26.2	34.0	28.2
:	:	:	:	:	:	:
71	71	A	26.4	15.3	24.6	32.4
72	72	A	21.8	10.6	14.4	18.7
73	73	P	27.2	28.5	35.0	30.5
74	74	P	22.4	22.0	19.1	18.7
75	75	P	32.5	25.1	27.8	27.3
76	76	P	24.9	23.6	21.2	21.1
77	77	P	24.6	25.0	21.7	23.9
78	78	P	23.1	20.9	21.7	19.9
79	79	A	21.1	5.6	7.3	12.3
80	80	P	25.8	21.9	23.6	24.8
81	81	P	30.0	27.6	24.0	23.7
82	82	A	22.1	21.0	8.6	24.6
83	83	P	20.0	22.7	21.2	20.5
84	84	P	38.1	40.8	38.0	32.7
85	85	A	28.9	12.5	16.7	22.2
86	86	P	25.1	28.1	27.5	24.8
87	87	A	19.8	11.6	13.0	23.1
88	88	P	22.1	21.1	21.5	20.6
89	89	A	23.5	7.9	12.4	18.9
90	90	A	29.1	16.8	15.1	18.8
91	91	A	30.3	3.5	3.0	11.5
92	92	P	25.4	24.3	22.7	20.1
93	93	A	30.6	28.2	27.0	25.5
94	94	A	22.4	7.1	17.2	18.7
95	95	A	31.2	10.8	19.8	22.2
96	96	A	31.4	3.9	7.0	17.8
97	97	A	41.1	15.1	10.9	27.1
98	98	A	29.4	22.1	25.3	4.1

99	99	A	21.9	7.6	10.8	13.0
100	100	A	20.7	8.1	25.7	12.3

Melting the data for easy plotting

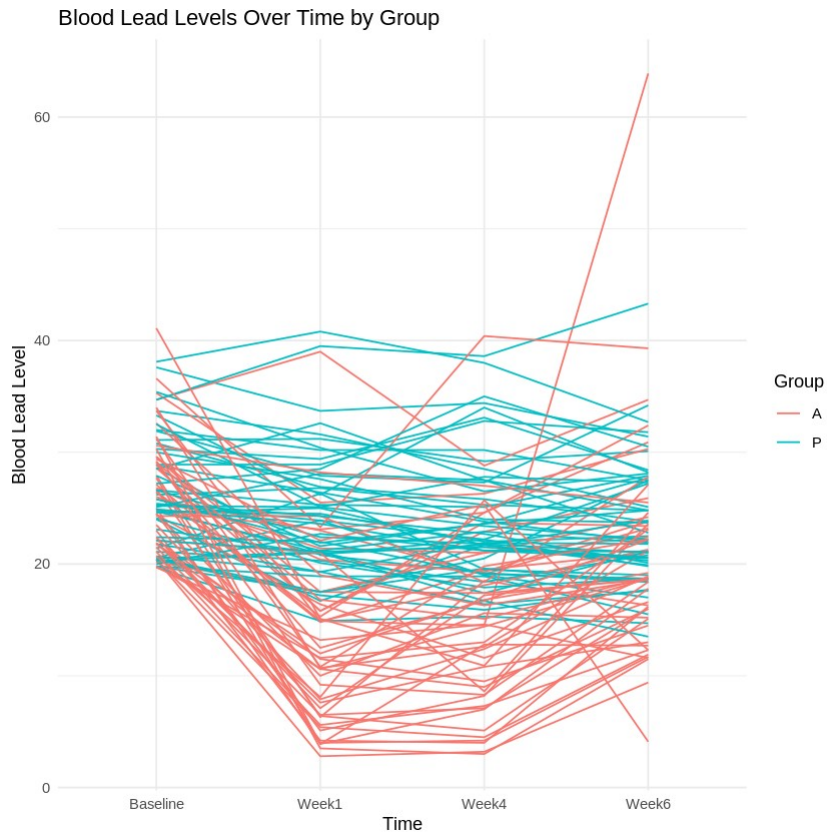
```
data_melted <- reshape2::melt(data, id.vars = c("SubjectID", "Group"),
                             variable.name = "Time", value.name =
"BloodLeadLevel")
data_melted
```

	SubjectID	Group	Time	BloodLeadLevel
1	1	P	Baseline	30.8
2	2	A	Baseline	26.5
3	3	A	Baseline	25.8
4	4	P	Baseline	24.7
5	5	A	Baseline	20.4
6	6	A	Baseline	20.4
7	7	P	Baseline	28.6
8	8	P	Baseline	33.7
9	9	P	Baseline	19.7
10	10	P	Baseline	31.1
11	11	P	Baseline	19.8
12	12	A	Baseline	24.8
13	13	P	Baseline	21.4
14	14	A	Baseline	27.9
15	15	P	Baseline	21.1
16	16	P	Baseline	20.6
17	17	P	Baseline	24.0
18	18	P	Baseline	37.6
19	19	A	Baseline	35.3
20	20	A	Baseline	28.6
21	21	P	Baseline	31.9
22	22	A	Baseline	29.6
23	23	A	Baseline	21.5
24	24	P	Baseline	26.2
25	25	A	Baseline	21.8
26	26	A	Baseline	23.0
27	27	A	Baseline	22.2
28	28	P	Baseline	20.5
29	29	A	Baseline	25.0
30	30	P	Baseline	33.3
:	:	:	:	:
371	71	A	Week6	32.4
372	72	A	Week6	18.7
373	73	P	Week6	30.5
374	74	P	Week6	18.7
375	75	P	Week6	27.3
376	76	P	Week6	21.1
377	77	P	Week6	23.9
378	78	P	Week6	19.9
379	79	A	Week6	12.3

380	80	P	Week6	24.8
381	81	P	Week6	23.7
382	82	A	Week6	24.6
383	83	P	Week6	20.5
384	84	P	Week6	32.7
385	85	A	Week6	22.2
386	86	P	Week6	24.8
387	87	A	Week6	23.1
388	88	P	Week6	20.6
389	89	A	Week6	18.9
390	90	A	Week6	18.8
391	91	A	Week6	11.5
392	92	P	Week6	20.1
393	93	A	Week6	25.5
394	94	A	Week6	18.7
395	95	A	Week6	22.2
396	96	A	Week6	17.8
397	97	A	Week6	27.1
398	98	A	Week6	4.1
399	99	A	Week6	13.0
400	100	A	Week6	12.3

Plot trajectories of blood lead levels over time for each subject

```
p <- ggplot(data_melted, aes(x = Time, y = BloodLeadLevel, group =
SubjectID, color = Group)) +
  geom_line() +
  labs(title = "Blood Lead Levels Over Time by Group", x = "Time",
y = "Blood Lead Level") +
  theme_minimal()
p
```



```
# Mixed effects model with random intercept
modell <- lmer(BloodLeadLevel ~ Time + Group + (1|SubjectID), data =
data_melted)
summary_modell <- summary(modell)
summary_modell
```

Linear mixed model fit by REML ['lmerMod']
Formula: BloodLeadLevel ~ Time + Group + (1 | SubjectID)
Data: data_melted

REML criterion at convergence: 2564

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.2841	-0.5432	0.0038	0.4504	6.9510

Random effects:

Groups	Name	Variance	Std.Dev.
SubjectID	(Intercept)	24.48	4.947
Residual		24.42	4.941

Number of obs: 400, groups: SubjectID, 100

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	23.6173	0.8915	26.492

TimeWeek1	-7.3150	0.6988	-10.468
TimeWeek4	-6.6140	0.6988	-9.465
TimeWeek6	-4.2020	0.6988	-6.013
GroupP	5.5775	1.1060	5.043

Correlation of Fixed Effects:

	(Intr)	TimWk1	TimWk4	TimWk6
TimeWeek1	-0.392			
TimeWeek4	-0.392	0.500		
TimeWeek6	-0.392	0.500	0.500	
GroupP	-0.620	0.000	0.000	0.000