Final Exam

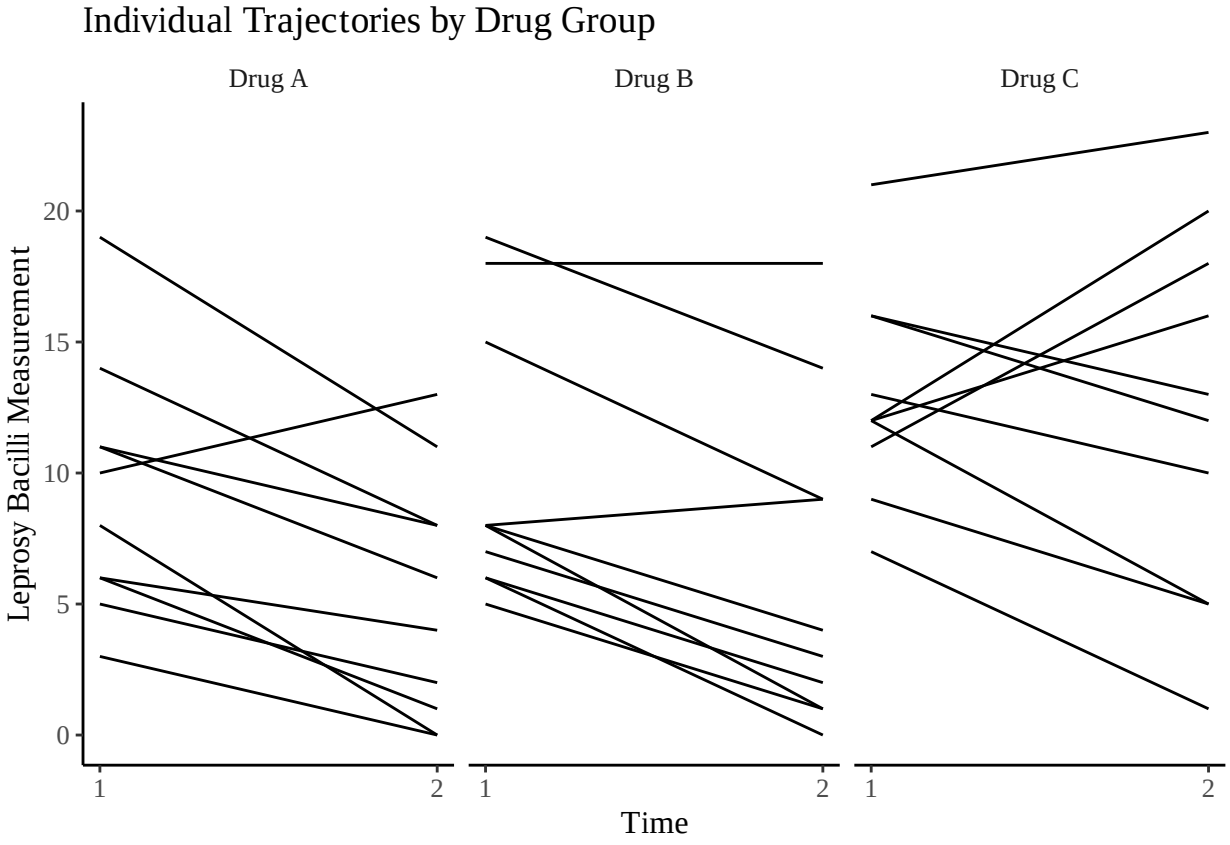
Blain Morin

October 27, 2018

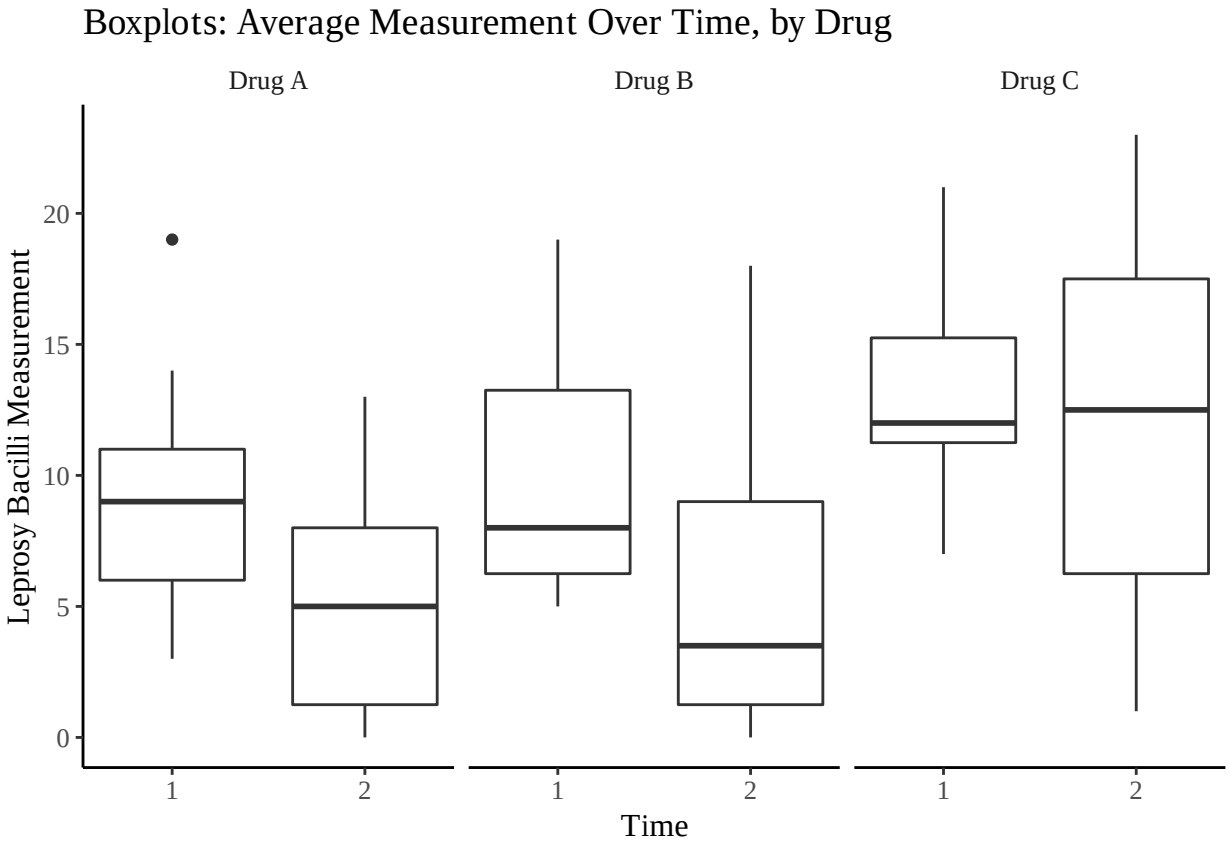
Part 1:

Question 1: Create the following plots:

i. Line ("spaghetti") plots to present the individual trajectories of the primary outcome by treatment group



ii. Plots to present the overall change in the mean primary outcome by treatment group



Question 2: Choose an appropriate regression model to describe the effect of treatment on the change in the primary outcome over time. Fit the model to the data and answer the following questions:

Table 1: Mixed Model Regressions

	Dependent variable:		
	Leprosy Bacilli Measurement		
	Random Intercept Model	Random Slope Model	Fixed Effects Model
	(1)	(2)	(3)
Drug B	0.600	0.600	0.600
	(3.473)	(2.756)	(3.473)
Drug C	0.200	0.200	0.200
	(3.473)	(2.756)	(3.473)
Time (Continuous)	-4.000^{***}	-4.000^{***}	-4.000^{***}
	(1.243)	(1.547)	(1.243)
Drug B : Time	0.100	0.100	0.100
	(1.758)	(2.187)	(1.758)
Drug C : Time	3.400^{*}	3.400	3.400^{*}
	(1.758)	(2.187)	(1.758)
Constant	13.300^{***}	13.300^{***}	13.300^{***}
	(2.456)	(1.949)	(2.456)
Observations	60	60	60
Log Likelihood	-164.239	-164.042	-164.239
Akaike Inf. Crit.	344.478	344.084	344.478
Bayesian Inf. Crit.	361.233	360.839	360.390

Note: *p<0.1; **p<0.05; ***p<0.01
Reference Group is Drug A

i. Explain why you chose the particular model.

ii. Is the change in the primary outcome the same for all three treatments according to this model?

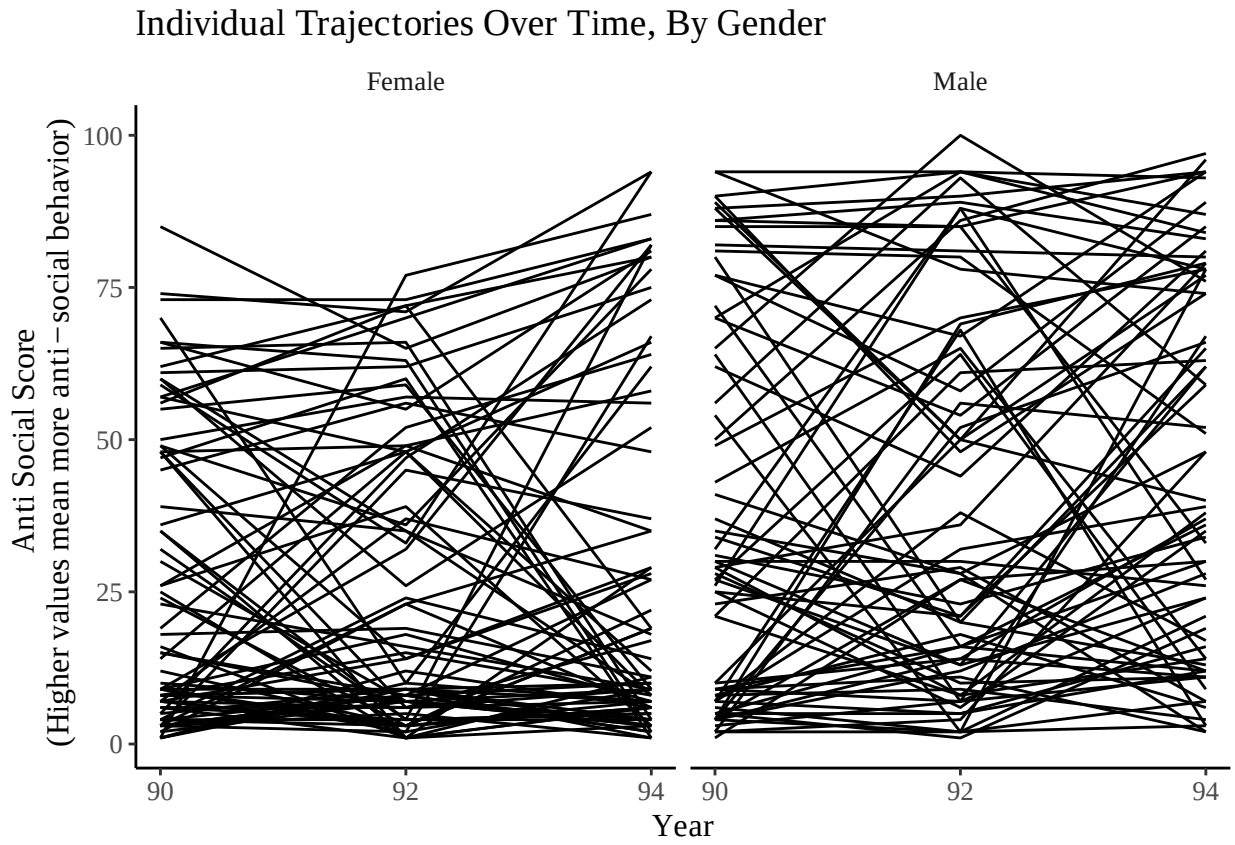
iii. Interpret the regression coefficients.

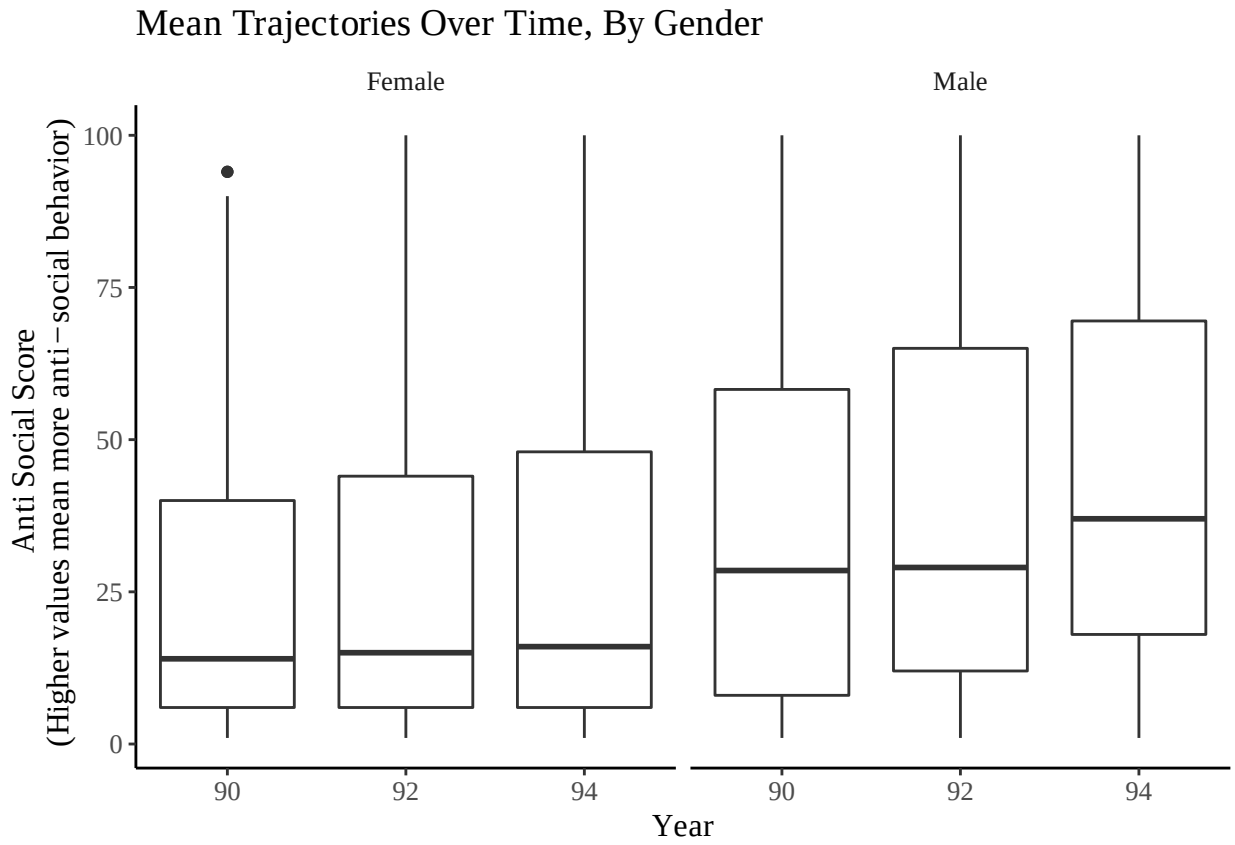
Part 2:

Question 3: Explore the data. Provide descriptive statistics and graphs to present the information included in this dataset in a meaningful and comprehensive way to your collaborators.

Table 2: Summary Statistics

Statistic	Mean	St. Dev.	Min	Max	N
Anti Score	33.271	28.964	1	100	1,743
Female	0.504	0.500	0	1	1,743
Poverty	0.329	0.470	0	1	1,743
Mom's Age	20.656	2.188	16	25	1,743
Child's Age	8.944	0.601	8	10	1,743
Hispanic	0.244	0.430	0	1	1,743
Black	0.363	0.481	0	1	1,743
Mom Works	0.336	0.472	0	1	1,743
Mom Married	0.236	0.425	0	1	1,743


Individual Trajectories Over Time, By Gender
Female
Male
Anti Social Score
(Higher values mean more anti−social behavior)
100
75
50
25
0
90
92
94
Year



Question 4: : One of the doctors (Doctor 1) is only interested in overall differences in the antisocial behavior scores between baseline and the end of the follow-up period (1994). He is not interested in describing any particular time trend. He also wants to compare these differences between boys and girls.

i. What methodology would you apply in order to answer this doctor's research question? State the form (write the mathematical formula) of the model (M1) that you will fit to your data.

$$M1 : E[Y_i|X_i] = \beta_0 + \beta_1 Female + \beta_3 Time94 + \beta_4 Female * Time94$$

ii. Based on the model results what is your conclusion about the changes in the mean response over time between boys and girls?

Table 3: Regression Results M1

	Dependent variable:
	Anti Social Score
Female	-10.552^{***}
	(2.336)
Year = 94	8.396^{***}
	(1.680)
Female * Time94	-5.416^{**}
	(2.365)
Constant	35.934^{***}
	(1.659)
Observations	1,162
Log Likelihood	−5,440.820
Akaike Inf. Crit.	10,893.640
Bayesian Inf. Crit.	10,923.970
Note:	*p<0.1; **p<0.05; ***p<0.01
	Year is categorical, reference year is 1990

iii. What is the model estimate of the difference in the average antisocial behavior scores between:

a. boys and girls at baseline?

b. boys and girls in 1994?

c. 1990 and 1994 for boys?

d. 1990 and 1994 for girls?

e. boys in 1994 and girls in 1990?

f. boys in 1990 and girls in 1994?