Problem 1.

In your own words describe how RM ANOVA can be used to test for differences between groups, differences between repeated observations, and differences in observations as a function of group membership. These tests have specific names we discussed in lecture.

(a) Groups Hypothesis (Test of Levels)

    (i) Do groups score similarly on the collected set of measures?

    (ii) This test checks if the mean of the groups over all measurement times is the same. This does not mean that all the groups performed the same at each time point, just that their average performance is the same over all time points. Mathematically, this is simple measuring the relative contributions of between-group and within-group contributions to the total sum of squared errors. If the group 'levels' are significantly different, then the equal levels null hypothesis is rejected.

(b) Flatness Hypothesis (Test of Flatness)

    (i) Is the dependent variable similar across assessments?

    (ii) This test checks if the difference in the response between assessments (i.e., the slope) is nonzero. The flatness null hypothesis is that the slope between all assessments is zero, and that the results are flat. This is checked for each group, and is effectively a within-subjects test. If the slope is statistically significantly nonzero, then there is a within groups main effect of time.

(c) Parallelism Hypothesis (Test of Parallelism)

    (i) Do different groups have parallel profiles?

    (ii) This test checks if the slope between each set of assessments is the same across all groups. This is equivalent to a one-way ANOVA on the slopes, we are checking that there is no interaction of the within-subjects factor with the between-subjects factor.

Problem 2.

Why do we say that in RM ANOVA time is treated as a categorical variable?

    RM ANOVA treats time as a categorical factor. Time is considered a set of discrete, fixed conditions in which each subject contributes one outcome variable per time condition. The ANOVA model for describing mean change over time uses as many parameters as there are discrete occasions. This is also known as a saturated means model—the model for time is saturated by using all possible degrees of freedom for differences across conditions of the time variable. The goal of the ANOVA model is not to predict or summarize the pattern of means, but simply to reproduce the observed mean per occasion using fixed effects equal to the number of occasions minus 1 (given that the fixed intercept is already included).

Problem 3.

What is the difference between polynomial contrasts and comparisons of means between different observation periods?

Polynomial contrast can be used to test which level of polynomial (linear, quadratic, cubic) explains the relation under study. The comparison of means between different observation periods doesn't lead to information about how the dependent variable changes with time, just that there may be some difference between the observations (e.g. categorical differences). In that respect, polynomial contrasts can give the statistician more information about the underlying variable.

Problem 4.

Researchers were interested in how confidence in students fluctuates from the end of sophomore year through the end of senior year in college. Further, researchers were interested in whether students attended a private (i.e., Harvard, Stanford, and Yale) or public (i.e., UMass—Boston, San Francisco State University, and Southern Connecticut State University) university would influence end-of-year student confidence ratings, and possible trajectories.

Use the data set confidence.csv to answer the following questions. For your information, t1-t3 indicate time of observation 1-3; i.e., end of sophomore year, end of junior year, and end of senior year, respectively. The variable public indicates whether the student attended a public (public = 1) or private (public = 0) university.

1. Convert the wide format data set to a long format data set. Remember that you will need to create an ID variable, and an Observation/Time variable. Show syntax and the header and footer of the long format data set.

```
# Load data
gwide <- read.csv("confidence.csv")

# Convert from wide to long format
public <- rep(c(0, 1), each = 150)
obs <- rep(c("t1", "t2", "t3"), 100)
time <- rep(c(0.5, 1.5, 2.5), 100)
id <- paste("id", 1:100, sep = "")
observation <- unlist(lapply(1:100, function(x) {gwide[x, 1:3]}))
glong <- data.frame(id, obs, time, public, observation)
> head(glong)
   id obs time public observation
1 id1  t1  0.5      0           4
2 id2  t2  1.5      0           6
3 id3  t3  2.5      0           6
4 id4  t1  0.5      0           3
5 id5  t2  1.5      0           5
6 id6  t3  2.5      0           5
> tail(glong)
        id obs time public observation
295  id95  t1  0.5      1           3
296  id96  t2  1.5      1           4
```

```
23  297   id97   t3   2.5        1            4
24  298   id98   t1   0.5        1            4
25  299   id99   t2   1.5        1            5
26  300  id100   t3   2.5        1            4
```

2. Test whether public and private universities differed in their confidence scores. Conduct any pairwise comparison necessary. Report your conclusions.

```
1  > summary(aov(observation ~ public + Error(id), data = glong))
2
3  Error: id
4             Df Sum Sq Mean Sq F value Pr(>F)
5  public      1   2.43   2.430   1.689  0.197
6  Residuals  98 140.97   1.438
7
8  Error: Within
9              Df Sum Sq Mean Sq F value   Pr(>F)
10 public       1  47.04   47.04    30.5 1.04e-07 ***
11 Residuals  199 306.96    1.54
12 ---
```

A one-way between subjects ANOVA was conducted to compare the effect of different types of universities on student confidence for students at public and private universities. There was not a significant effect of university type on student confidence at the $p = 0.05$ level for the two conditions $[F(1, 98) = 1.7, p = 0.197]$. Comparisons of the mean indicated that the mean confidence scores for the private university students ($M = 3.77, SD = 1.30$) was larger than the public university students ($M = 2.96, SD = 1.15$).

3. Test whether confidence differs over time. Do this by comparing all observations to each other. Also do this by testing the maximum allowable number of polynomial contrasts. Report your conclusions for both tests.

```
1  > # Test for effect of observation
2  > summary(aov(observation ~ obs + Error(id), data = glong))
3
4  Error: id
5             Df Sum Sq Mean Sq F value Pr(>F)
6  Residuals  99  143.4   1.448
7
8  Error: Within
9              Df Sum Sq Mean Sq F value Pr(>F)
10 obs          2   13.4   6.703   3.897 0.0219 *
11 Residuals  198  340.6   1.720
12 ---
13 >
14 > pairwise.t.test(glong$observation, glong$obs,paired = TRUE, p.adjust = c("bonf"))
15
16    Pairwise comparisons using paired t tests
17 data:  glong$observation and glong$obs
18      t1        t2
```

```
19 t2 1.7e-06 -
20 t3 0.024    1.000
21 P value adjustment method: bonferroni
```

A one-way within subjects ANOVA was conducted to compare the effect of time on student confidence. There was a significant effect of time on student confidence at the $p = .05$ level for the three conditions $[F(2, 198) = 3.9, p = 0.02]$. Post-hoc t-tests at the $p = .05$ level with the Bonferroni correction indicated that there was a significant difference between measurements at time 1 and 2, and time 1 and 3, but not between time 2 and 3.

```
1 > # Add Polynomial Contrasts to describe trajectories
2 > glong[, 6] <- contr.poly(3)[, 1]
3 > glong[, 7] <- contr.poly(3)[, 2]
4 >
5 > colnames(glong) <- c(colnames(glong[, 1:5]), c("l", "q"))
6 >
7 > # Test to see linear/quadratic contrasts describe the observations
8 > summary(aov(observation ~ l + q + Error(id), data = glong))
9
10 Error: id
11          Df Sum Sq Mean Sq F value Pr(>F)
12 Residuals 99  143.4   1.448
13
14 Error: Within
15           Df Sum Sq Mean Sq F value Pr(>F)
16 l          1    7.6   7.605   4.421 0.0368 *
17 q          1    5.8   5.802   3.373 0.0678 .
18 Residuals 198  340.6   1.720
19 ---
```

Polynomial contrasts were also investigated to explain the relation between time and student confidence scores at the $p = 0.05$ level. There was a significant linear trend, $[F(1, 198) = 4.4, p = 0.04]$, indicating that as time increased, the student confidence scores increased proportionally. A quadratic trend was also investigated, but was found to be non-significant, $[F(1, 198) = 3.4, p = 0.07]$.

4. Test for a possible interaction between variables tested in question 2 (university type) and 3 (mean differences between observations and significant trends in confidence). Conduct any follow-up analyses necessary if there is a significant interactions. Report your conclusions for all tests.

```
1 > # Test for interaction effect
2 > summary(aov(observation ~ public:obs + Error(id), data = glong))
3 Error: id
4            Df Sum Sq Mean Sq F value   Pr(>F)
5 public:obs  3  32.87  10.958   9.518 1.45e-05 ***
6 Residuals  96 110.52   1.151
7 ---
8
9 Error: Within
10             Df Sum Sq Mean Sq F value   Pr(>F)
```

```
11  public:obs   3  52.39   17.464   11.41 6.26e-07 ***
12  Residuals  197 301.61    1.531
13
14  # Output truncated for follow-up tests
15
16  # Follow-up options Simple Effects split on interaction: e.g., public vs. private
17  summary(aov(observation ~ obs + Error(id), glong[which(glong$public == 0), ]))
18  summary(aov(observation ~ obs + Error(id), glong[which(glong$public == 1), ]))
19
20  # Test where differences exist
21  pairwise.t.test(glong[which(glong$public == 1), "observation"],
22                  glong[which(glong$public == 1), "obs"],
23                  paired = TRUE, p.adjust = c("bonf"))
24  pairwise.t.test(glong[which(glong$public == 0), "observation"],
25                  glong[which(glong$public == 0), "obs"],
26                  paired = TRUE, p.adjust = c("bonf"))
27
28  # Follow-up options Simple Effects split on Observation
29  summary(aov(observation ~ public + Error(id), glong[which(glong$obs == 't1'), ]))
30  summary(aov(observation ~ public + Error(id), glong[which(glong$obs == 't2'), ]))
31  summary(aov(observation ~ public + Error(id), glong[which(glong$obs == 't3'), ]))
32
33  # Find mean and sd
34  t3 = glong[glong$time == 2.5,]
35  aggregate(t3$observation, by=list(t3$public), FUN=mean)
36  aggregate(t3$observation, by=list(t3$public), FUN=sd)
```

A one-way between subjects ANOVA was conducted to test for an interaction between university type and observation on student confidence. There was a significant interaction detected at the $p = 0.05$ level $[F(3, 96) = 9.5, p < .001]$. Follow-up simple effects analyses indicated that there was a significant effect of observation for both private $[F(2, 97) = 29.7, p < .001]$ and public $[F(2, 97) = 5.2, p = .001]$ schools. Pairwise t-tests for both private and public schools indicated significant differences among all combinations of observation periods. Simple effects analyses also show a simple effect for the third observation $[F(1, 98) = 44.1, p < .001]$. Student confidence scores in the third observation were significantly higher for the students in private schools $(M = 4.38, SD = 1.50)$ than public school $(M = 2.54, SD = 1.27)$.

Source	SS	df	F	PR(> F)
School Type				
Private	29.66	2	11.97	< .001
Public	13.69	2	5.18	.001
Time				
1	0.81	1	1.1	.295
2	4.00	1	2.9	.091
3	84.64	1	44.1	< .001

Table 1: Results of Simple Effects Analysis

5. Create a plot of the data that represents the trajectory for public and private students. Make sure the trajectories have some indication of variability around the average trajectory. Sufficiently label and describe your figure.

```python
# Python!
import pandas as pd
import seaborn as sns
sns.set_style('ticks')

df = pd.read_csv('glong.csv', index_col=0)
p = sns.lmplot(x='time', y='observation',
               hue='public',
               data=df, order=2, ci=95,
               x_jitter=.3, y_jitter=.3,
               markers=["o", "x"],
               legend=False)
p.axes[0][0].set_xticks([0.5, 1.5, 2.5])
p.axes[0][0].set_xticklabels(['Sophomore', 'Junior', 'Senior'])
plt.xlabel('Observation')
plt.ylabel('End-of-Year Student Confidence Rating')
plt.savefig('prob5-quad.pdf')
plt.close('all')
```

This figure shows a scatter plot of each observation of student confidence ratings. The ratings of public school students are shown as green x's, and blue o's represent the private schools. As the dataset contains only three distinct measurement conditions and eight observation values, a small amount of jitter is added to each observation. This jitter allows the viewer to get a better idea of how many data points are at each location. The quadratic fit to the public and private schools is overlaid on the scatter plot along with 95% confidence regions. Private schools show an overwhelmingly linear fit, while public schools shown a quadratic fit which slopes downward in the final observation category, making the interaction evident.
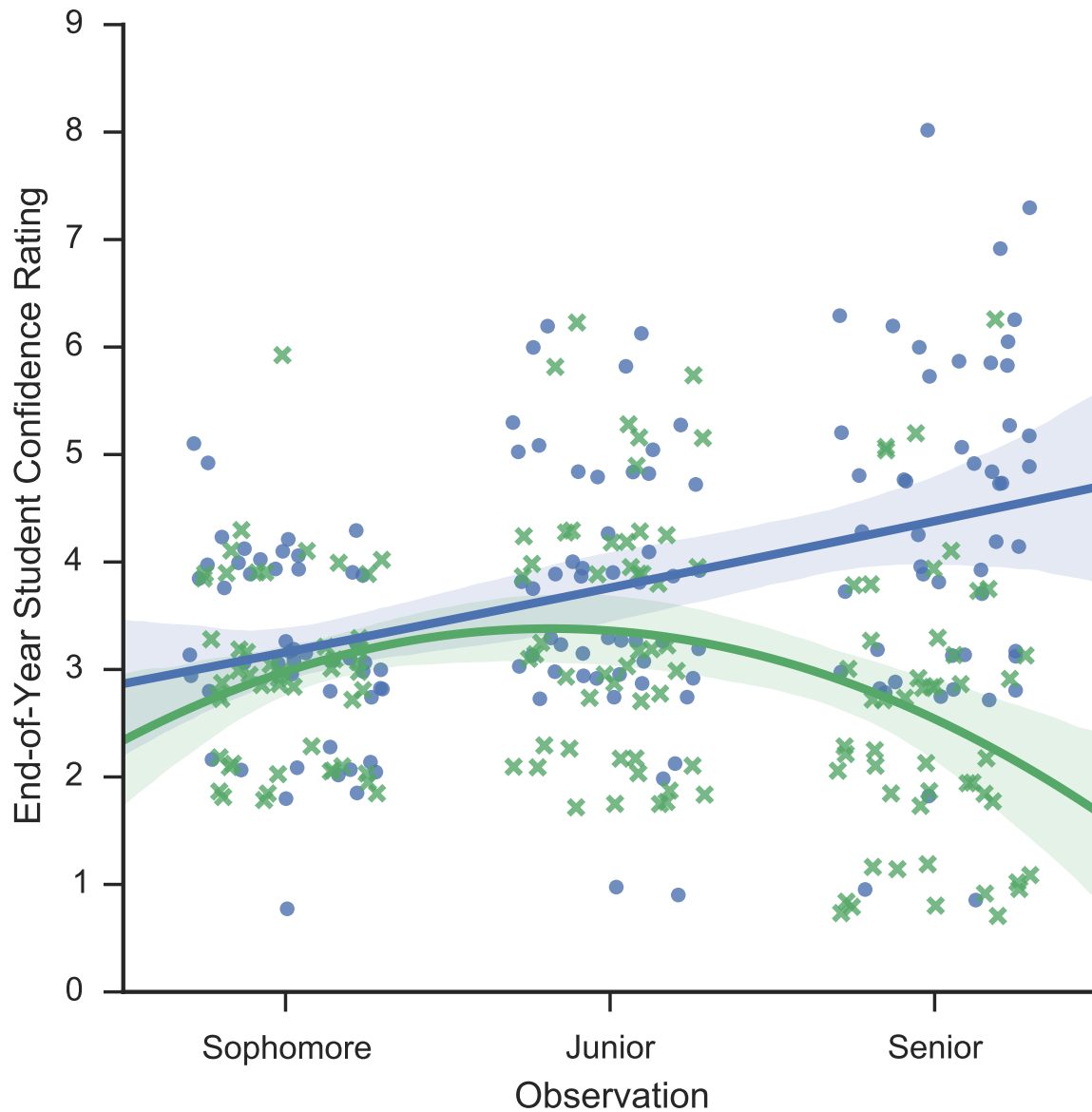
Figure 1: Quadratic trajectories for end-of-year confidence ratings for students in public schools (green) and private schools (blue). Shaded regions represent 95% confidence regions. (Some jitter introduced for visibility.)