MSBA 6450

Please attach all code.

Name: **Danny Moncada**

1. The COO of US Airlines is trying to solve a staffing problem of pilots on its international route network. The table below lists the needs of pilots needed to fly the Boing 737 aircrafts for the next six months. Due to seasonality of its route network, a pilot's salary varies by the month as shown in the table. At the beginning of Month 1 the airline has 120 pilots, but this can be adjusted each month.

Month	1	2	3	4	5	6
Needed Pilots	100	110	115	140	110	200
Salary (in '000)	$9	$7	$8	$6	$6	$9

Pilots can be hired and fired at the beginning of each month. Newly hired pilots can start working the same month, and fired pilots stop working the day they are fired. The severance cost of firing a pilot is $10k, and the hiring cost is $5k per pilot. If it is convenient, the airline can keep a staff of pilots larger than their actual need.

By an existing contract between the airlines and the pilot's union the airline cannot hire pilots if it has fired pilots in either of the previous two months. No plot has been fired in either of the previous two months prior to month 1.

[i] Develop a model and solve with R to prescribe the (integer) number of pilots to hire and fire each month to minimize the total staffing cost (cost of salary, hiring and firing).

Optimized cost: $6710

Month	0	1	2	3	4	5	6
Needed Pilots	-	100	110	115	140	110	200
Pilots Hired	-	0	0	5	25	0	60
Pilots Fired	-	10	0	0	0	0	0
Total Pilots (Contracted)	120	110	110	115	140	140	200

[ii] After negotiations following a financial loss by the airline, the pilot's union has expressed a willingness to temporarily alter its agreement with the airline to prevent the latter from being able to hire if it had fired pilots in either of the previous two months.

Instead, the union has proposed the following clauses:

- The airline will be able to hire and fire a will for the next 6 months.
- However, none of the existing original 120 pilots can be fired in the first 3 months, but they can be fired starting in the beginning of month 4.
- All pilot salaries (existing and new hires) need to be fixed at $7k per month.
- The severance cost incurred by the airline for firing during these 6 months will be reduced to$5k.

Develop a model and solve with R to prescribe the (integer) number of pilots to hire and fire each month to minimize the total staffing cost (costs of salary plus hiring and firing) for the proposed agreement. Advise the COO if this is a better or worse deal for the airline over this six months planning horizon.

[ii]

Optimized cost: $6280

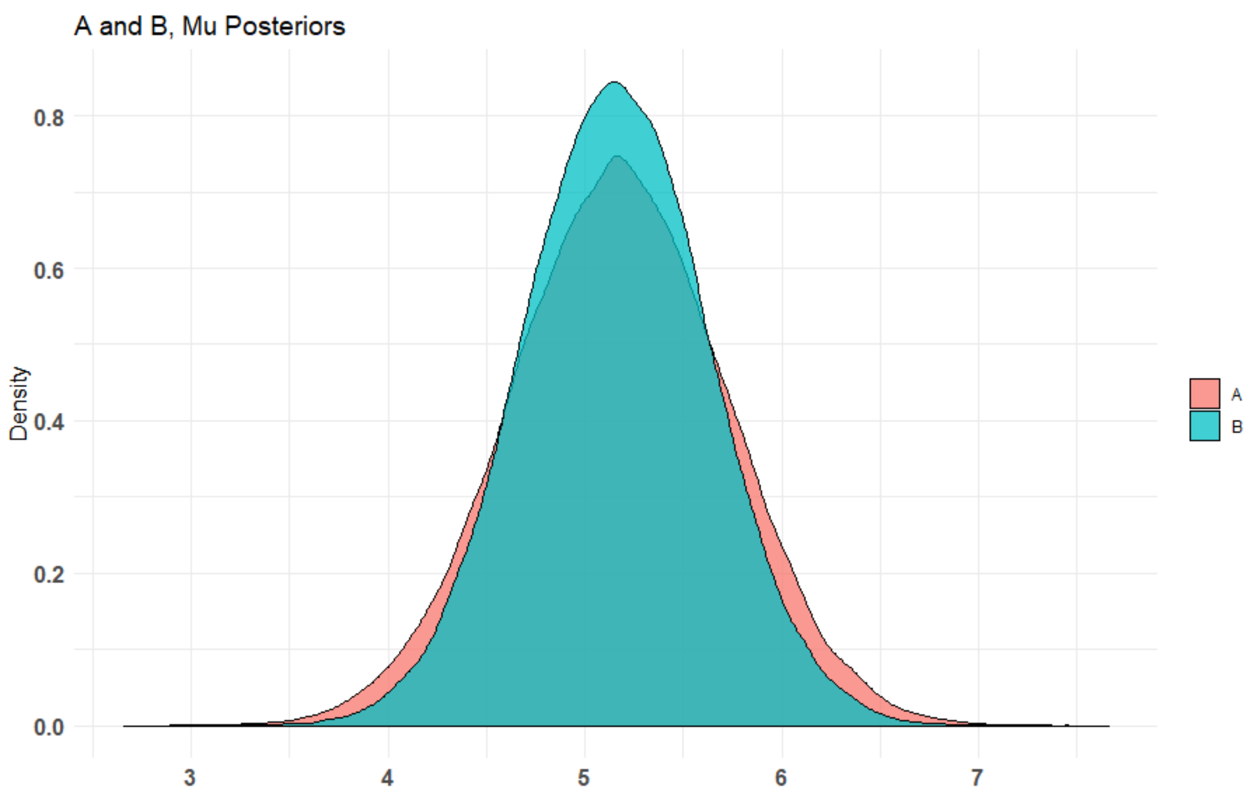
Month	0	1	2	3	4	5	6
Needed Pilots	120	100	110	115	140	110	200
Pilots Hired	0	0	0	0	20	0	60
Pilots Fired	0	0	0	0	0	30	0
Total Pilots (contracted)	120	120	120	120	140	110	200

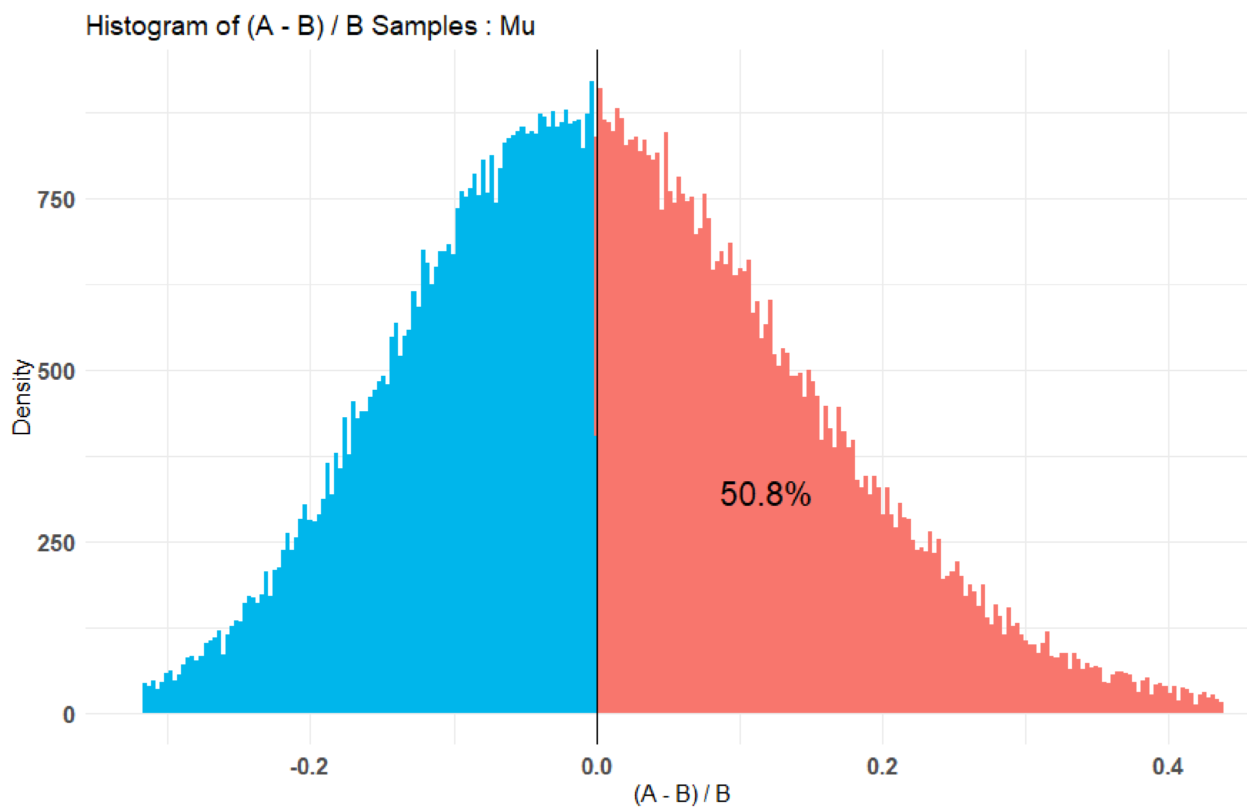
Per the two models, the second of the two deals is a better deal for the COO. The first model has an optimized cost of $6710, while the second model has an optimized cost of $6280. Also, fixing the salaries and reducing the severance cost for firing pilots may have contributed to a much better outcome.

2. (25 points) Here are 15 numbers generated from rnorm(mu, sigma) where mu and sigma were generated from runif ☺. Use your Monty Carlo/Bayesian Black Magic to determine what the actual **mu** was. **mu** +/- .03*mu is considered correct. For every extra +/- .03***mu** will cost you 5 points. I suggest multiple experiments. Incidentally, Frequentists will tell you that this cannot be done *LOL*.

4.653431, 4.446646, 5.087739, 5.052033, 6.920449, 6.968938, 4.452641, 4.948848, 5.054194
4.659720, 5.939746, 5.000139, 5.034643, 4.955450, 4.066340

mu is 5.213765 per my analysis. Based on the sample provided, and after running through numerous Bayes AB trails, I calculated that the actual mean was the sample_mean – sample_mean*.055. The results are in the code, but here are a few plots demonstrating my confidence intervals and the probability scores that lead me to that conclusion.





3. (25 points) Create a linear/quadratic/logarithmic/exponential model for Y in terms of X for the following data sets. For sure you're going to want to use everything in your tool box including Cross Validation. Hint: This one is tricky ☺ so be careful.

X = {0.52933466, 0.15552968, 0.36581071, 0.45692039, 0.91465466, 0.08033931, 0.76546048, 0.64505530, 0.62327284, 0.05966902}

Y = {2.629730, 1.281666, 1.958268, 2.284790, 3.578952, 1.409117, 3.663195, 2.749400, 2.540583, 2.576510}

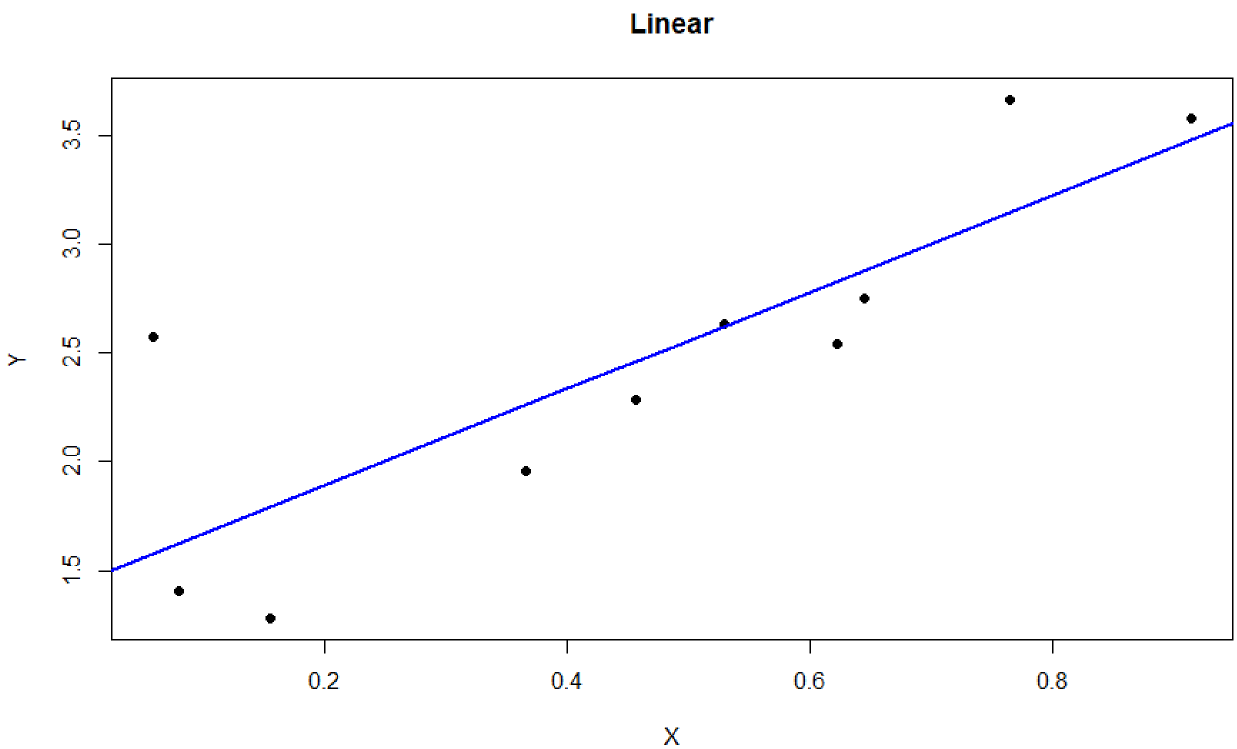
(a) What is your specific mathematical model?

Exponential model is going to be my selection, based on the tests generated in R.

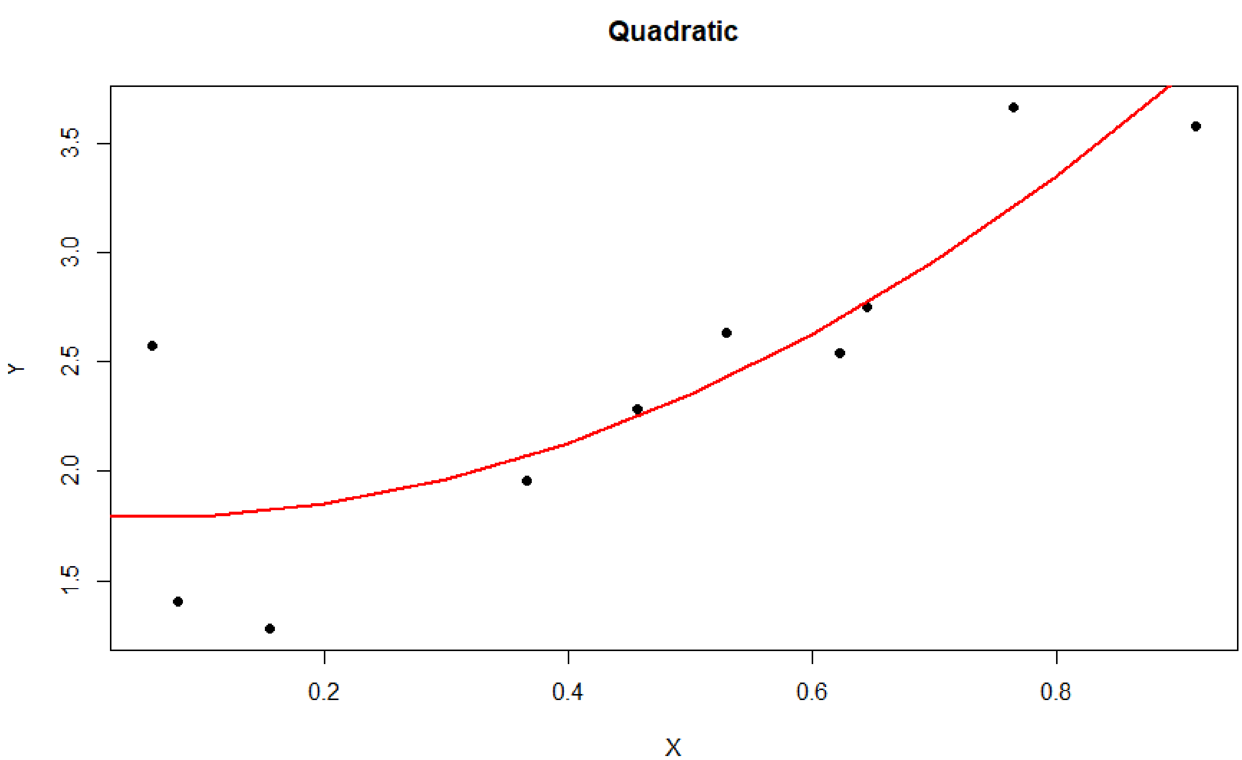
(b) Why specifically did you choose this model in light of the other options? What specifically in your analysis leads you to this conclusion?

At first glance, most of the models generally fit the pattern of the data provided – that's why a first eyeball test is not enough.

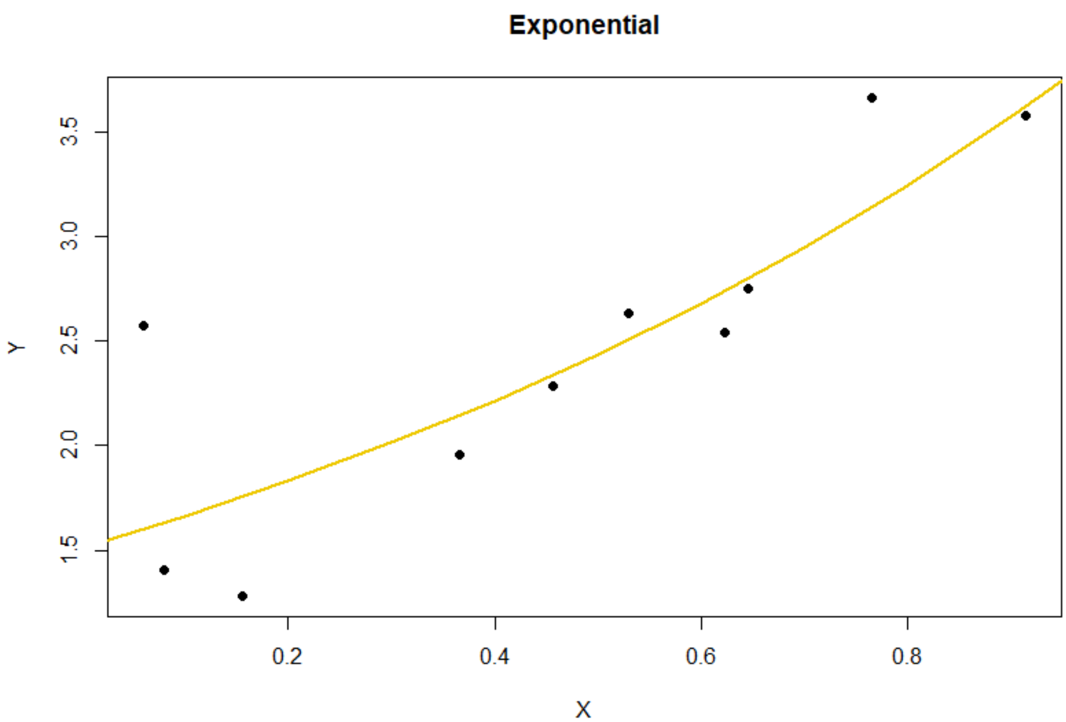
Linear model:



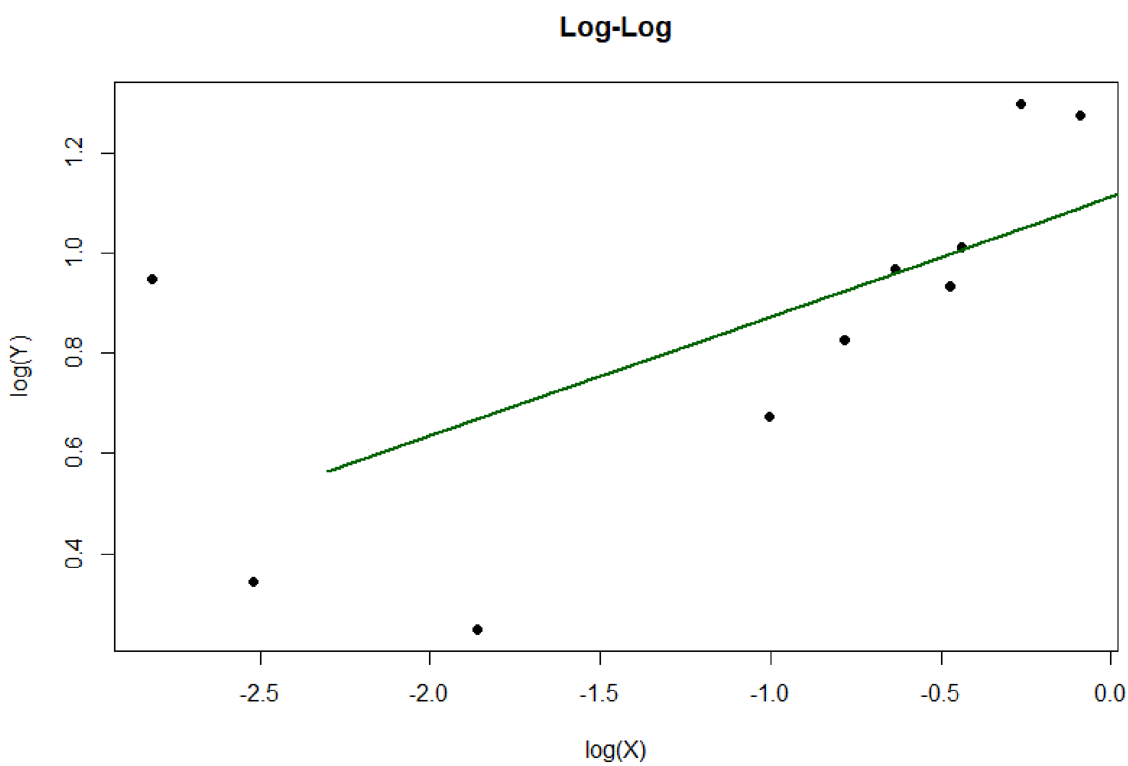
Quadratic model:



Exponential model:

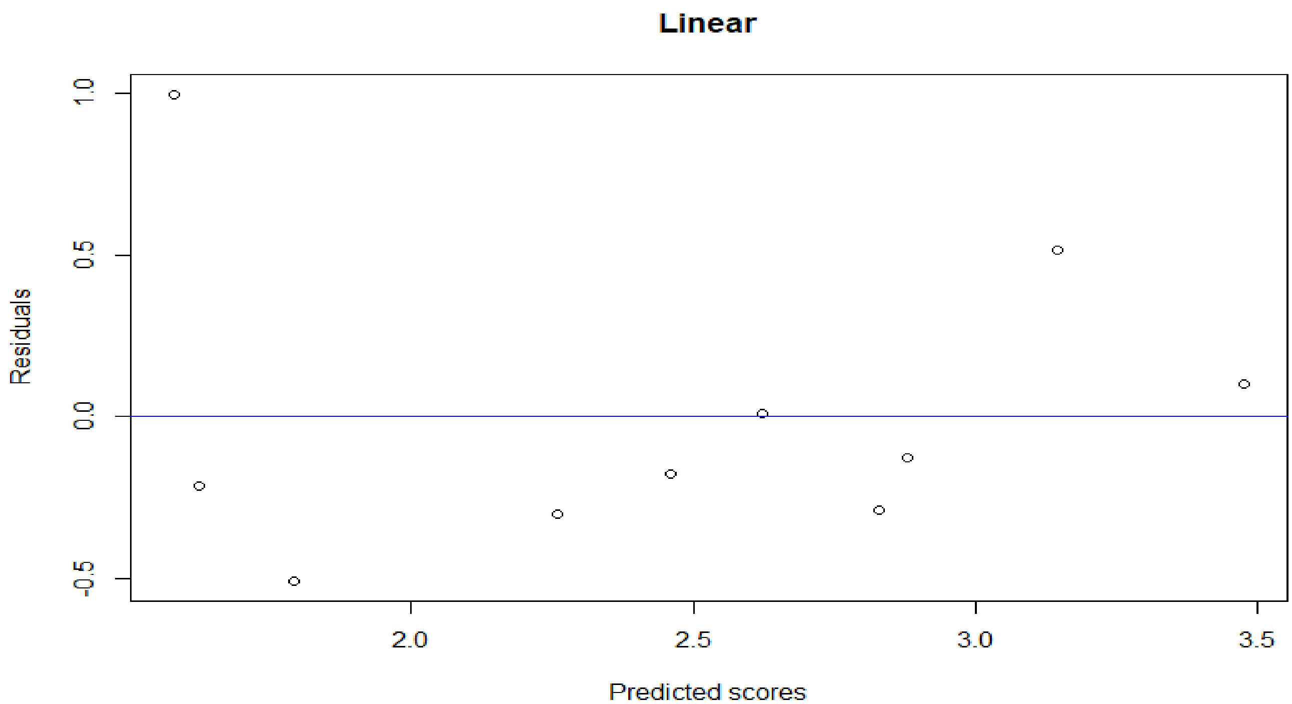


Logarithmic model:

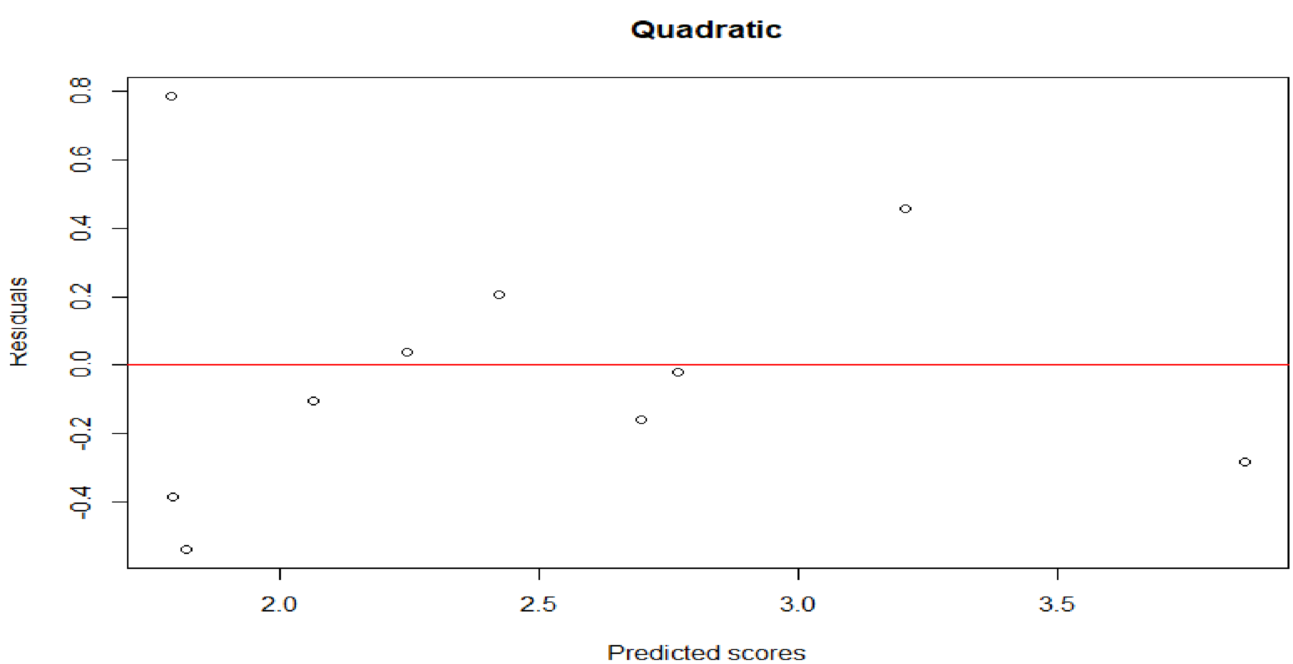


Following this, I decided to plot the standardized residuals for each one of my models; hopefully this helps me whittle down my options before I review the summary statistics to make my final decision.

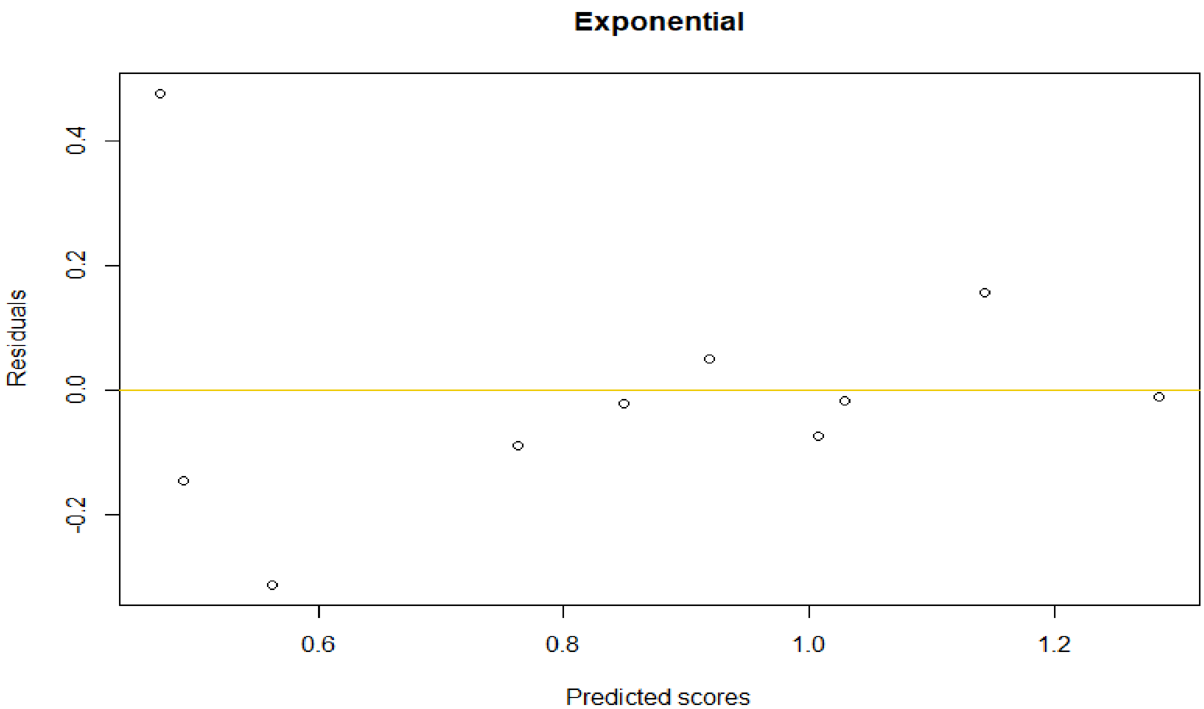
Linear Model:



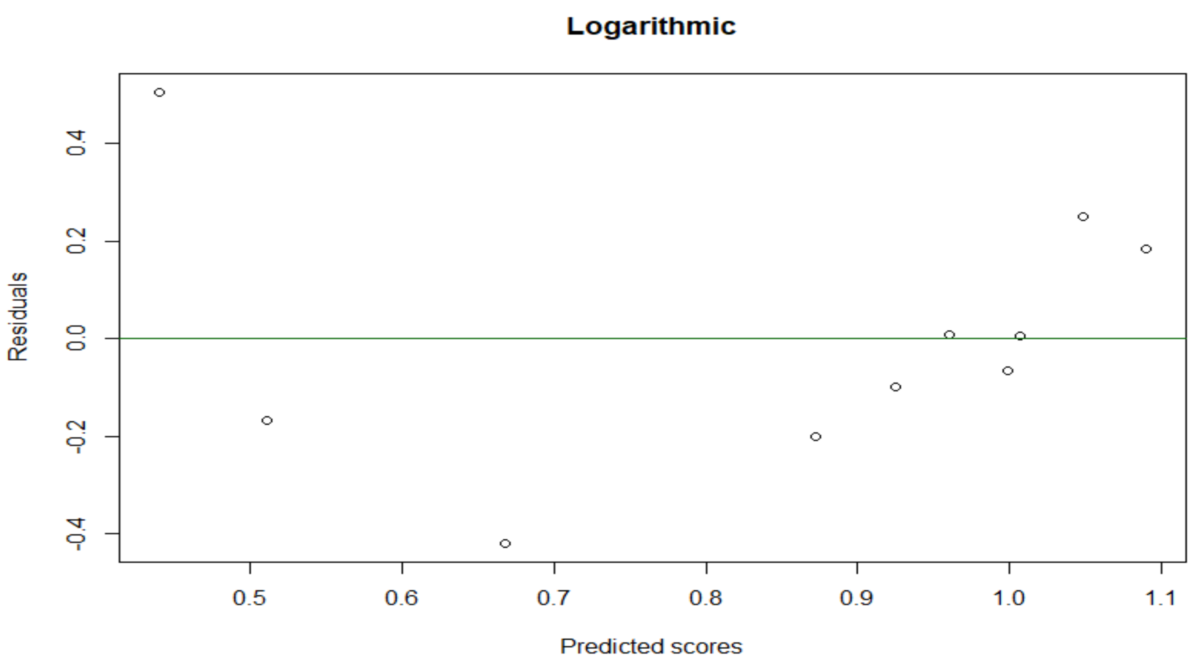
Quadratic Model:



Exponential Model:



Logarithmic Model:



After looking at the standardized residuals for each of my models, I am focused on two models: Linear and Exponential. For the last set of analysis, I will look at the summary statistics for each model.

Linear Model:

```
Residual standard error: 0.4746 on 8 degrees of freedom
Multiple R-squared:  0.6787,     Adjusted R-squared:  0.6385
F-statistic:  16.9 on 1 and 8 DF,  p-value: 0.003388
```

Quadratic Model:

```
Residual standard error: 0.4509 on 7 degrees of freedom
Multiple R-squared:  0.7462,     Adjusted R-squared:  0.6737
F-statistic: 10.29 on 2 and 7 DF,  p-value: 0.008234
```

Exponential Model:

```
Residual standard error: 0.2202 on 8 degrees of freedom
Multiple R-squared:  0.6439,     Adjusted R-squared:  0.5994
F-statistic: 14.47 on 1 and 8 DF,  p-value: 0.005211
```

Logarithmic Model:

```
Residual standard error: 0.2766 on 8 degrees of freedom
Multiple R-squared:  0.438,      Adjusted R-squared:  0.3678
F-statistic: 6.235 on 1 and 8 DF,  p-value: 0.03711
```

Out of the four models, the lowest p-value is for the Linear model at 0.003388, but close behind is the Exponential model at 0.005211. The best F-statistic again belongs to the Linear model at 16.9, followed again by Exponential at 14.47.

The standardized residual plot is what I will use to make my final decision. The Exponential model has the best out of the four; this is what I am basing my ultimate conclusion on.

Thanks for a great semester Professor Morley!

-DM