

PUTTING UBCSECURE TO A TEST

Kratika Rathi 38763710

Philip Loewen 70359054

Abstract

Through this experiment, we attempted to test the upload speed of UBCSecure at different locations, time of day and devices. Going into this project, we expected the upload speeds to be fairly constant over the different factors, however, we were surprised to see the results.

Introduction

To measure the upload speed and collect our data, we used the website <https://speedtest.net>. We restricted ourselves to this website to minimize any noise in our data. Similarly, the same model of devices were used to further minimize instrumental error. Finally, every reading was taken from the same place in each location. These cautionary steps were taken to avoid variance from other random effects in the data. The rest of the report is as follows: Details of the experimental design will be described in Section 2, followed by Statistical Analysis in Section 3, Conclusion & Discussion in Section 4 and lastly Tables, Figures and Data Appendix in Section 5.

Experimental Design

Factor / Block	Description	Levels
Location	Where the speed was checked	Nest, Library, Residence
Time of Day	When the speed was checked	Morning, Afternoon, Evening
Device	On what device speed was checked	iPhone, iPad, Macbook
Day	On what day speed was checked	Wednesday, Thursday, Friday

Location, Time of Day are factors whereas Device and Day are blocking factors.

All counts of Upload Speed were measured in Mbps (megabits per second). All the other factors are unitless.

Under location, we chose 3 places that all serve different purposes to estimate which type of building would serve the best for uploading files. Library was chosen as an educational building requiring good internet to ease the uploading and downloading of assignments, Nest as a general purpose building without a need for very high internet speeds and Residence which we expect to fall in between.

Under Time of Day, we chose an interval of ~4 hours between every reading. This resulted in a reading taken every Morning, Afternoon and Evening.

Under devices, we chose 3 different devices that connect to the internet and are frequently used by students at UBC. The devices used work on the same operating system, to reduce instrumental bias. This was used as a blocking factor to incorporate many different devices. The experiment was set up over the course of 3 days, hence Day is also used as a blocking factor. Every time slot, the order in which the different locations were visited (to collect data) was randomized and for each location, the order in which the devices were used was also randomized. The actual order of the runs is included in the data file.

The experiment is a full factorial 3^4 design with 2 blocking factors. As such we have performed 81 different observations. Since we believe the 4 factor interaction effect to be negligible, this results in 16 degrees of freedom for the residuals, allowing us to properly test the significance of our estimated effects.

Statistical Analysis

Before we begin fitting a model, we look at a Boxcox plot to see if our model requires any transformations. From Figure 1, we can see that the confidence interval for log likelihood includes $\lambda = 1$, which indicates that our linear model does not require any transformation.

Continuing with our process, we attempt to fit a linear regression model to our data. This model takes the form of

$$Y_{ijkl} = \mu + \beta_i + \omega_j + \rho_k + \psi_l + (\beta\omega)_{ij} + (\beta\rho)_{ik} + (\beta\psi)_{il} + (\omega\rho)_{jk} + (\omega\psi)_{jl} + (\rho\psi)_{kl} + (\beta\omega\rho)_{ijk} + (\beta\rho\psi)_{ikl} + (\omega\rho\psi)_{jkl} + (\beta\omega\psi)_{ijl} + E_{ijkl}$$

Where,

Y_{ijkl} is the observed response for factor i , factor j , block k and block l ;

μ is the intercept or overall mean effect;

β_i is the effect of Location i ($i = 1, 2, 3$);

ω_j is the effect of Device j ($j = 1, 2, 3$);

ρ_k is the effect of Time k ($k = 1, 2, 3$);

ψ_l is the effect of Day l ($l = 1, 2, 3$);

$(\beta\omega)_{ij}$ is the interaction effect of level i of Location and level j of Device;

$(\beta\rho)_{ik}$ is the interaction effect of level i of Location and level k of Time;

$(\beta\psi)_{il}$ is the interaction effect of level i of Location and level l of Day;

$(\omega\rho)_{jk}$ is the interaction effect of level j of Device and level k of Time;

$(\omega\psi)_{jl}$ is the interaction effect of level j of Device and level l of Day;

$(\rho\psi)_{kl}$ is the interaction effect of level k of Time and level l of Day;
 $(\beta\omega\rho)_{ijk}$ is the interaction effect of level i of Location with level j of Device and level k of Time;
 $(\beta\omega\psi)_{ijl}$ is the interaction effect of level i of Location with level j of Device and level l of Day;
 $(\omega\rho\psi)_{jkl}$ is the interaction effect of level j of Device with level k of Time and level l of Day;
 $(\beta\omega\psi)_{ijl}$ is the interaction effect of level i of Location with level j of Device and level l of Day;
 E_{ijkl} are the random errors which are assumed to be i.i.d. Normal random variables with mean 0 and variance σ^2 .

After fitting this model in RStudio, we obtain the residuals of the model. It is necessary that the Normality assumptions of the residuals are checked, which may otherwise cast a doubt on our inference procedure.

Figure 2 shows the Q-Q Plot, which plots the quantiles of our data with the theoretical quantiles of the Normal Distribution. We see that the standardized residuals from our model form a $y = x$ line when plotted against Normal quantiles. Hence our assumption of Normality is justified. Figure 3 shows the Residuals vs Fitted Values plot. We see a random distribution of the residuals from our model with equal variance, indicating that the assumption that the errors are independent and identically distributed is justified and the linear model is appropriate.

Table 1 shows the Anova table. From this table we can see that 6 most important effects are Location, Device, Time, Day, the interaction between Location and Device and finally the interaction between Time and Day at the 1% significance level. The main effects of Device and Day are the most significant. As such we set up a combination of contrasts to further explore their levels. These contrasts are shown in Figure 4.

Before the experiment we believed that Friday would be the day with the greatest upload speed as it is the day which has the least number of students on campus and that MacBook would be the device with the greatest upload speed. As such we decided to set up contrasts to estimate the effects of these two versus the other levels. For our second contrasts we decided to then compare the two levels which had been grouped into 'others'. The analysis of these contrasts revealed them to be highly statistically significant as well as practically significant. For example, we are 95% confident that changing from a MacBook to either an iPad or an iPhone will on average increase upload speed by 29.5 Mbps with a confidence interval of [15.75, 43.28] Mbps.

Although we did not set up contrasts for the other factors, we have included in Figure 7 the mean effects plot for each factor. This figure shows that for the Time factor there is a large difference between Morning and the rest of the day, but there is very little difference between Afternoon and evening. Similarly there is a large difference in upload speed between the nest and the residence and library, but little difference between the two latter.

We also see that interaction between Location and Device and the interaction between Time and Day were significant. To further hone in on these effects, we plot their interaction plots, as shown in Figure 5 and 6.

According to Figure 5, the upload speed for an iPhone and iPad is higher in a library or in residence and lowest at the Nest, but the opposite is true for a Macbook. The most optimal conditions are Macbook in the Nest and iPad in residence.

A more interesting interaction plot, however, is in Figure 6. For Wednesday and Thursday, upload speeds see the biggest dip during the evenings. On the other hand, Friday evenings see the highest upload speeds. This plot indicates that for each day, the optimal levels are Wednesday and Thursday mornings, and Friday evenings.

Conclusions

After fitting a linear model on our data, and doing the required analysis, we conclude that the most important main effects are Device and Day at the 1% significance level. The important interaction effects include Location - Device and Day - Time.

Within the levels of Device, we see that Macbook has the highest upload speed. In fact, we are 95% confident that on average the upload speed increases by [15.75,43.28] Mbps when switching from any other device to a Macbook. Through the interaction plot we see that the Macbook works best in the Nest (while speed in the library and residence are fairly similar). Within the levels of Day, we initially expected Friday to be the best performing day. However, Friday saw the slowest upload speeds, with Wednesday being the best. We are 95% confident that the upload speed on Friday, compared to other days, is slower by [10.26,37.8] Mbps.

Based on our analysis, when needing to upload a large file, we would recommend uploading this file on a Macbook in residence on a Wednesday Morning. We are 90% confident that this combination would lead to an upload speed between 125.0959 and 318.8493 Mbps with an expected speed of 221.9726 Mbps.

If given the option to do this experiment again, we would include a blocking factor for UBCSecure vs UBCVisitor. While we expect UBCVisitor to perform poorly, it could be interesting to see how Location, Time or Day affects the upload speed in interaction with this choice. To do so we can perform a paired study measuring the internet speed on both networks using the same device. Additionally, since the age and background network usage of each Device is different, we could introduce random effects in our model by treating Device as a random effect instead of a fixed one. This could potentially lead to interesting conclusions.

Tables and Figures

Figure 1 : Boxcox Plot

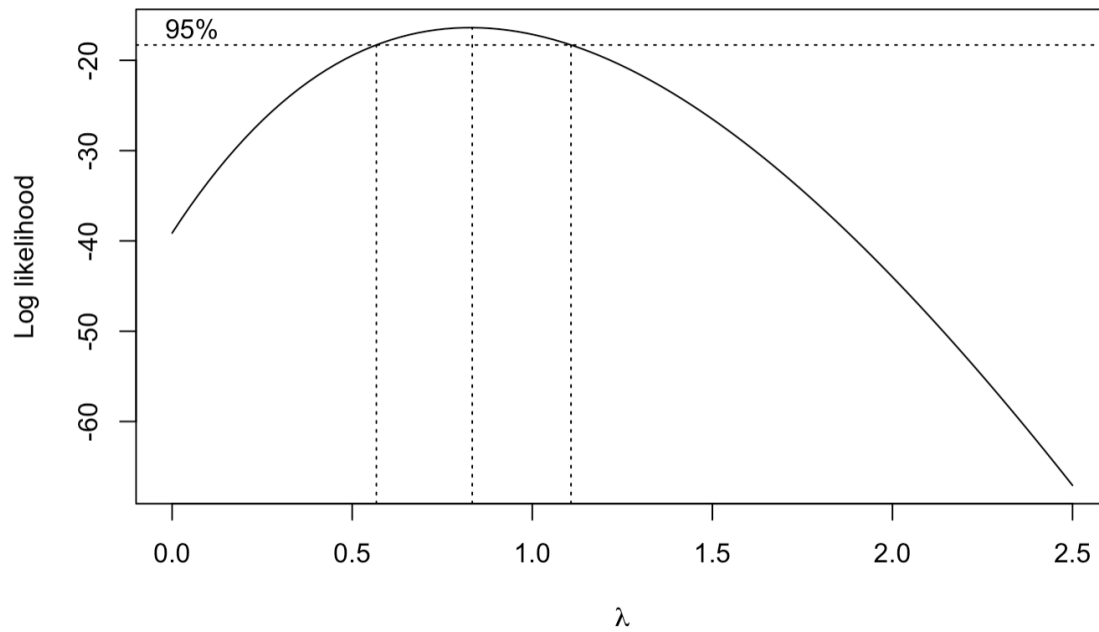


Figure 2 : Normal Q-Q Plot

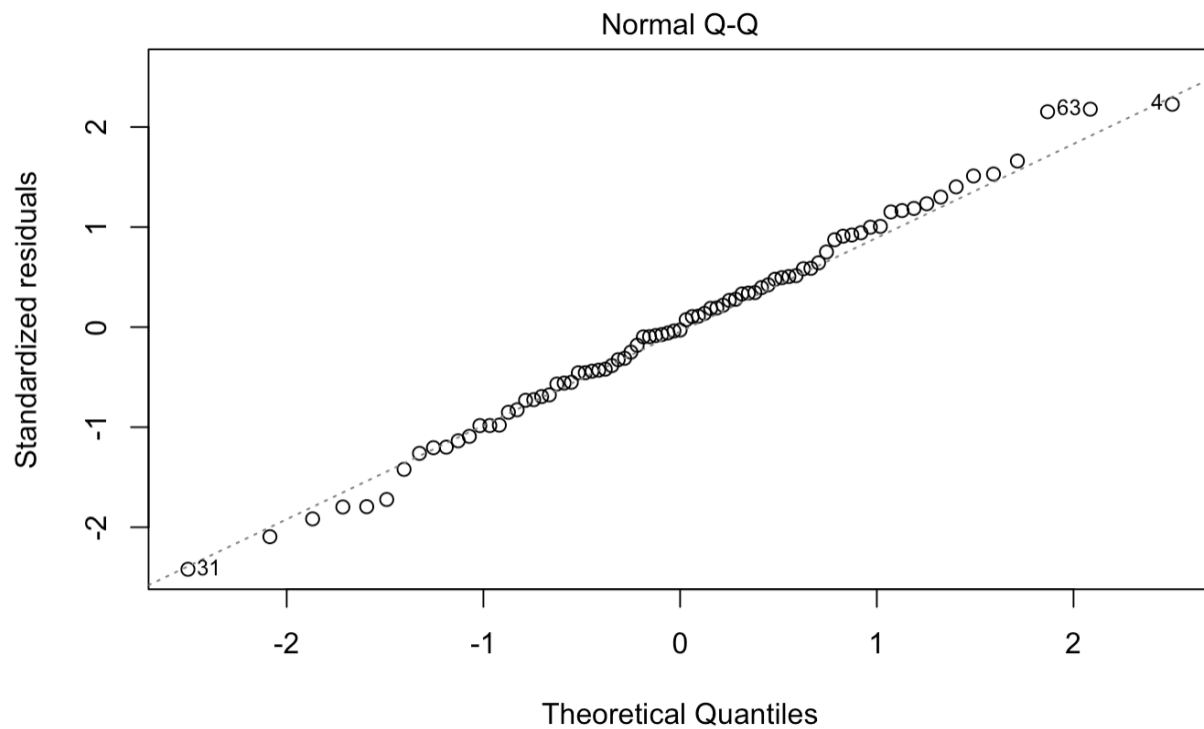


Figure 3 : Residuals vs Fitted Values

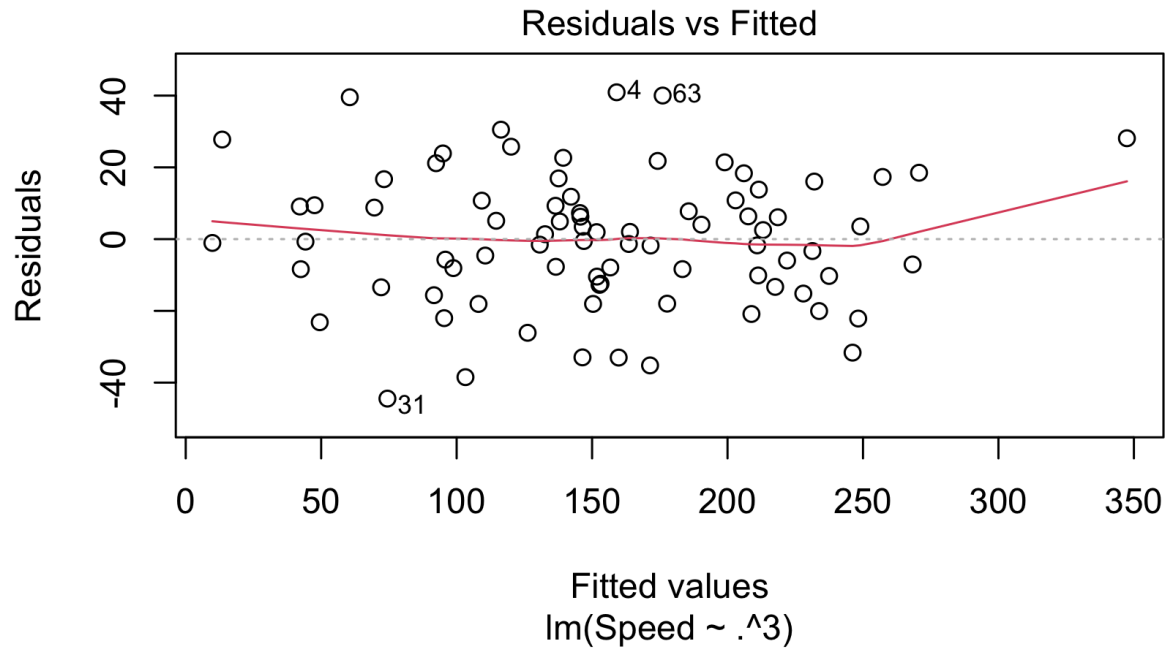


Figure 4 : Contrasts

MacbookVsOthers:	0.5	0.5	-1
iPhoneVsiPad:	1	-1	0
FridayVsOthers:	-1	0.5	0.5
WednesdayVsThursday:	0	1	-1

Figure 5 : Interaction Plot Between Location and Device

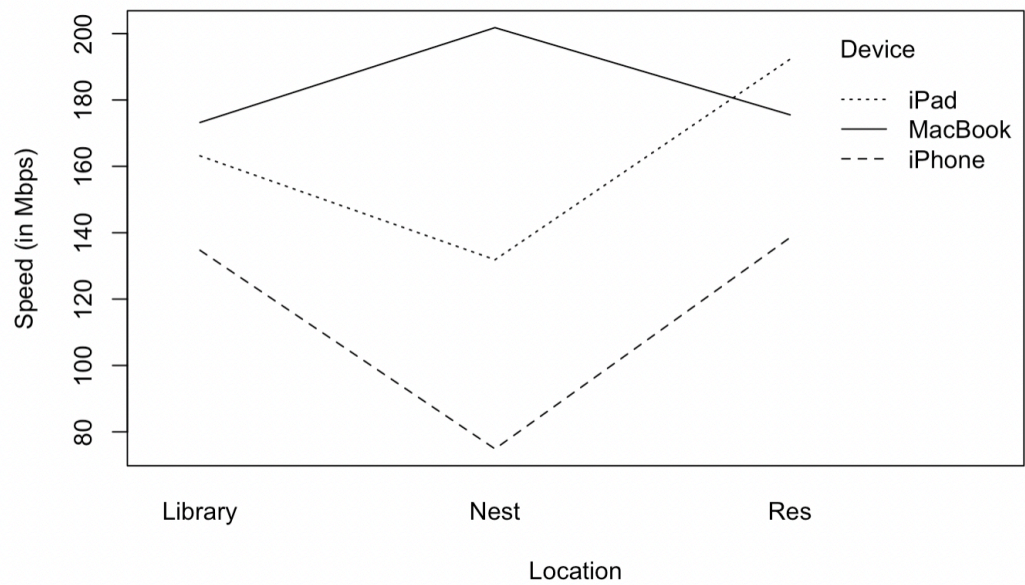


Figure 6 : Interaction Plot Between Time and Day

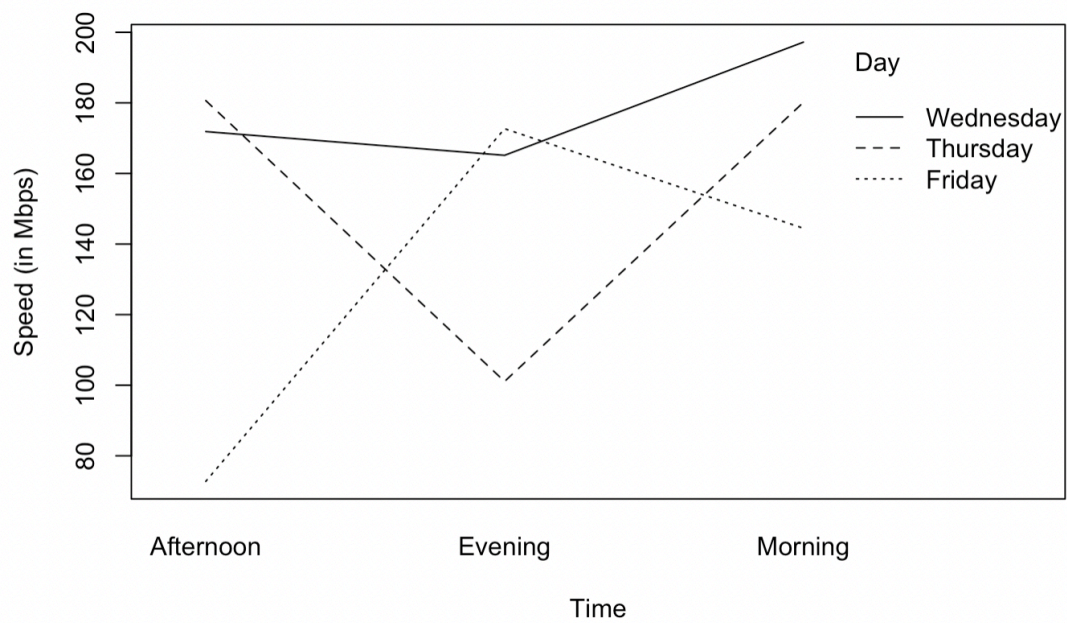


Figure 7 : Mean effects plot

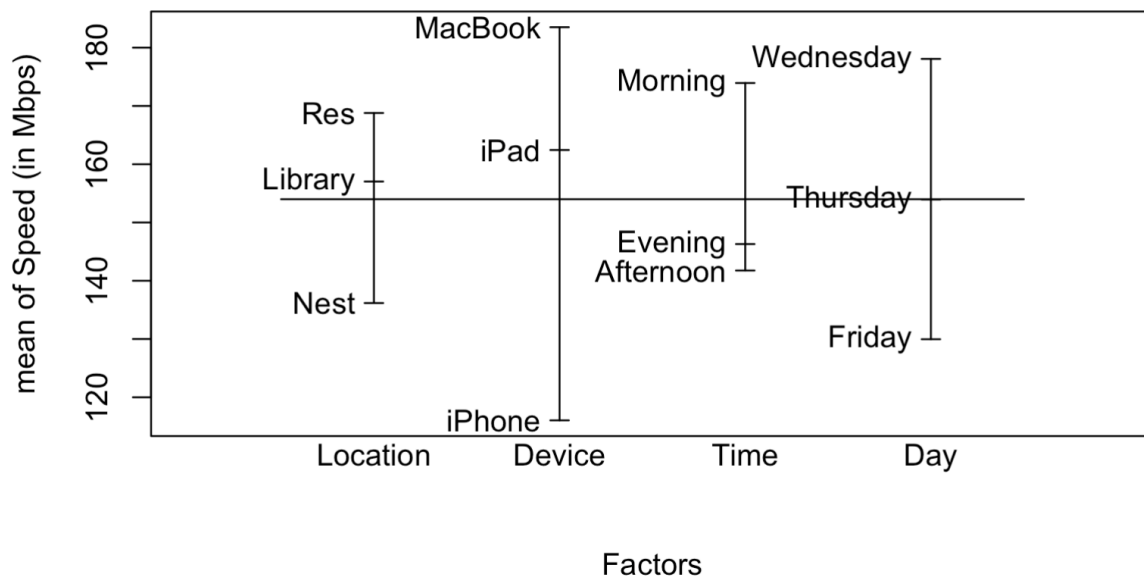


Table 1 : ANOVA Table

Analysis of Variance Table

Response: Speed

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Location	2	14732	7366	4.3120	0.0317750	*
Device	2	64343	32171	18.8334	6.242e-05	***
Time	2	16386	8193	4.7961	0.0233396	*
Day	2	31246	15623	9.1458	0.0022462	**
Location:Device	4	29212	7303	4.2753	0.0152962	*
Location:Time	4	18187	4547	2.6618	0.0709009	.
Location:Day	4	10900	2725	1.5953	0.2238201	
Device:Time	4	5865	1466	0.8584	0.5095629	
Device:Day	4	17706	4427	2.5913	0.0762503	.
Time:Day	4	74284	18571	10.8716	0.0001876	***
Location:Device:Time	8	12431	1554	0.9096	0.5325402	
Location:Device:Day	8	12668	1583	0.9270	0.5206392	
Location:Time:Day	8	20519	2565	1.5015	0.2323877	
Device:Time:Day	8	20760	2595	1.5191	0.2265347	
Residuals	16	27331	1708			

Table 2 : Partition of the Anova Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Location	2	14732	7366	4.312	0.031775	*
Device	2	64343	32171	18.833	6.24e-05	***
Device: MacbookVsOthers	1	35281	35281	20.654	0.000332	***
Device: iPhoneVsiPad	1	29062	29062	17.013	0.000794	***
Time	2	16386	8193	4.796	0.023340	*
Day	2	31246	15623	9.146	0.002246	**
Day: FridayVsOthers	1	23388	23388	13.691	0.001942	**
Day: WednesdayVsThursday	1	7858	7858	4.600	0.047666	*
...						
Residuals	16	27331	1708			

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

Table 3 : Coefficient of Contrasts Table

DeviceMacbookVsOthers	-29.51481481	6.494455
DeviceiPhoneVsiPad	23.19888889	5.624363
DayFridayVsOthers	24.03074074	6.494455
DayWednesdayVsThursday	-12.06314815	5.624363

Table 4 : Summary of the Fitted Linear Model

Call:

```
lm(formula = Speed ~ .^3, data = clean)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-44.453	-12.438	-0.532	10.799	40.921

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	163.562	37.024	4.418	0.000431	***
LocationNest	-116.009	49.032	-2.366	0.030941	*
LocationRes	-64.778	49.032	-1.321	0.205041	
DeviceiPhone	-26.943	49.032	-0.550	0.590247	
DeviceMacBook	-121.433	49.032	-2.477	0.024813	*
TimeEvening	-21.362	49.032	-0.436	0.668891	
TimeMorning	-11.754	49.032	-0.240	0.813584	
DayThursday	47.977	49.032	0.978	0.342396	
DayWednesday	-10.444	49.032	-0.213	0.834013	
LocationNest:DeviceiPhone	-10.772	61.612	-0.175	0.863399	
LocationRes:DeviceiPhone	-2.218	61.612	-0.036	0.971731	
LocationNest:DeviceMacBook	116.349	61.612	1.888	0.077239	.
LocationRes:DeviceMacBook	66.851	61.612	1.085	0.293984	
LocationNest:TimeEvening	104.502	61.612	1.696	0.109226	
LocationRes:TimeEvening	160.064	61.612	2.598	0.019421	*
LocationNest:TimeMorning	120.914	61.612	1.963	0.067333	.
LocationRes:TimeMorning	118.999	61.612	1.931	0.071341	.
LocationNest:DayThursday	75.871	61.612	1.231	0.235953	
LocationRes:DayThursday	52.167	61.612	0.847	0.409649	
LocationNest:DayWednesday	82.986	61.612	1.347	0.196778	
LocationRes:DayWednesday	129.247	61.612	2.098	0.052163	.
DeviceiPhone:TimeEvening	-22.848	61.612	-0.371	0.715625	
DeviceMacBook:TimeEvening	125.684	61.612	2.040	0.058221	.
DeviceiPhone:TimeMorning	-52.722	61.612	-0.856	0.404788	
DeviceMacBook:TimeMorning	64.556	61.612	1.048	0.310311	
DeviceiPhone:DayThursday	-51.971	61.612	-0.844	0.411369	
DeviceMacBook:DayThursday	137.869	61.612	2.238	0.039818	*
DeviceiPhone:DayWednesday	81.481	61.612	1.322	0.204598	
DeviceMacBook:DayWednesday	186.881	61.612	3.033	0.007912	**
TimeEvening:DayThursday	-44.489	61.612	-0.722	0.480664	
TimeMorning:DayThursday	34.017	61.612	0.552	0.588501	
TimeEvening:DayWednesday	18.606	61.612	0.302	0.766557	
TimeMorning:DayWednesday	-24.993	61.612	-0.406	0.690367	

LocationNest:DeviceiPhone:TimeEvening	33.127	67.492	0.491	0.630217
LocationRes:DeviceiPhone:TimeEvening	71.767	67.492	1.063	0.303418
LocationNest:DeviceMacBook:TimeEvening	-75.243	67.492	-1.115	0.281384
LocationRes:DeviceMacBook:TimeEvening	-40.280	67.492	-0.597	0.558986
LocationNest:DeviceiPhone:TimeMorning	-5.720	67.492	-0.085	0.933511
LocationRes:DeviceiPhone:TimeMorning	2.057	67.492	0.030	0.976067
LocationNest:DeviceMacBook:TimeMorning	29.967	67.492	0.444	0.662988
LocationRes:DeviceMacBook:TimeMorning	-30.333	67.492	-0.449	0.659147
LocationNest:DeviceiPhone:DayThursday	-8.517	67.492	-0.126	0.901156
LocationRes:DeviceiPhone:DayThursday	-9.720	67.492	-0.144	0.887286
LocationNest:DeviceMacBook:DayThursday	-33.497	67.492	-0.496	0.626429
LocationRes:DeviceMacBook:DayThursday	-50.890	67.492	-0.754	0.461796
LocationNest:DeviceiPhone:DayWednesday	-72.227	67.492	-1.070	0.300432
LocationRes:DeviceiPhone:DayWednesday	-133.397	67.492	-1.976	0.065602 .
LocationNest:DeviceMacBook:DayWednesday	-90.580	67.492	-1.342	0.198306
LocationRes:DeviceMacBook:DayWednesday	-159.523	67.492	-2.364	0.031089 *
LocationNest:TimeEvening:DayThursday	-114.213	67.492	-1.692	0.109979
LocationRes:TimeEvening:DayThursday	-146.180	67.492	-2.166	0.045770 *
LocationNest:TimeMorning:DayThursday	-155.520	67.492	-2.304	0.034951 *
LocationRes:TimeMorning:DayThursday	-131.370	67.492	-1.946	0.069380 .
LocationNest:TimeEvening:DayWednesday	-76.313	67.492	-1.131	0.274848
LocationRes:TimeEvening:DayWednesday	-171.923	67.492	-2.547	0.021520 *
LocationNest:TimeMorning:DayWednesday	-44.443	67.492	-0.658	0.519585
LocationRes:TimeMorning:DayWednesday	-86.727	67.492	-1.285	0.217089
DeviceiPhone:TimeEvening:DayThursday	5.543	67.492	0.082	0.935560
DeviceMacBook:TimeEvening:DayThursday	-150.167	67.492	-2.225	0.040820 *
DeviceiPhone:TimeMorning:DayThursday	72.040	67.492	1.067	0.301641
DeviceMacBook:TimeMorning:DayThursday	-103.850	67.492	-1.539	0.143420
DeviceiPhone:TimeEvening:DayWednesday	-18.070	67.492	-0.268	0.792325
DeviceMacBook:TimeEvening:DayWednesday	-109.477	67.492	-1.622	0.124327
DeviceiPhone:TimeMorning:DayWednesday	65.177	67.492	0.966	0.348570
DeviceMacBook:TimeMorning:DayWednesday	1.863	67.492	0.028	0.978316

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

Residual standard error: 41.33 on 16 degrees of freedom

Multiple R-squared: 0.9274, Adjusted R-squared: 0.6371

F-statistic: 3.194 on 64 and 16 DF, p-value: 0.006181

Data Appendix

Location : Place where the upload speeds were measured. Categorical variable with 3 levels - Nest, Library and Residence

Device : Device used to measure upload speeds. Categorical variable with 3 levels - Macbook, iPhone and iPad.

Time : When the upload speeds were measured. Categorical variable with 3 levels - Morning, Afternoon and Evening

Day : The day upload speeds were measured. Categorical variable with 3 levels - Wednesday, Thursday and Friday.