

# Social Interactions at a Public High School

*Lauren Treiman*

*8/3/2018*

Over the summer, I worked on a project at an Epidemiology lab at the University of Utah that investigated how social interactions in elementary schools, middle schools, and high schools affected the transmission of viruses. This project identified the social networks of a highschool in Salt Lake City, Utah. Students wore sensors that recorded the length and time of each interaction, enabling me to analyze these variables. This data will help epidemiologists better determine the spread of viruses through schools.

There were four datasets downloaded in the project. The first two were the node attributes for Day 1 and Day 2. The third file was the highschool schedule enabling determination of the class period the interaction occurred. The last interaction was the edgelist for both Day 1 and Day 2.

I used R packages `igraph`, `dplyr`, `chron`, and `lubridate`.

I created two datasets containing interactions occurring on either Day 1 or Day 2. I recoded the node attributes to match the edgelist in each dataset.

I then created a “Period” edge attribute to identify when each interaction occurred. There were 4 periods, 3 hallway Passings, and 2 lunches. Although some students had different lunchtimes, I coded students who interacted during lunchtime but had opposite lunches in “Lunch.”

## Summary of Network

There were two networks for each day of the experiment. I created smaller networks for each period, hallway passing, and lunch (lunch 1 and lunch 2 were grouped together).

Day 1

```
## IGRAPH 2ac9b73 UN-- 1425 153999 --
## + attr: name (v/c), GradeCat (v/n), GenderCat (v/n), Lunch (v/n),
## | TotalSeconds (e/n), BeginDateTime (e/c), EndDateTime (e/c),
## | Duration (e/n), period (e/c)
```

For day 1, there were 1425 students (vertices) who participated and 153999 interactions (edges).

Day 2

```
## IGRAPH 9421841 UN-- 1220 556935 --
## + attr: name (v/c), GradeCat (v/n), GenderCat (v/n), Lunch (v/n),
## | TotalSeconds (e/n), BeginDateTime (e/c), EndDateTime (e/c),
## | Duration (e/n), period (e/c)
```

For day 2, there were 1220 students who participated and 556935 interactions.

## Distribution

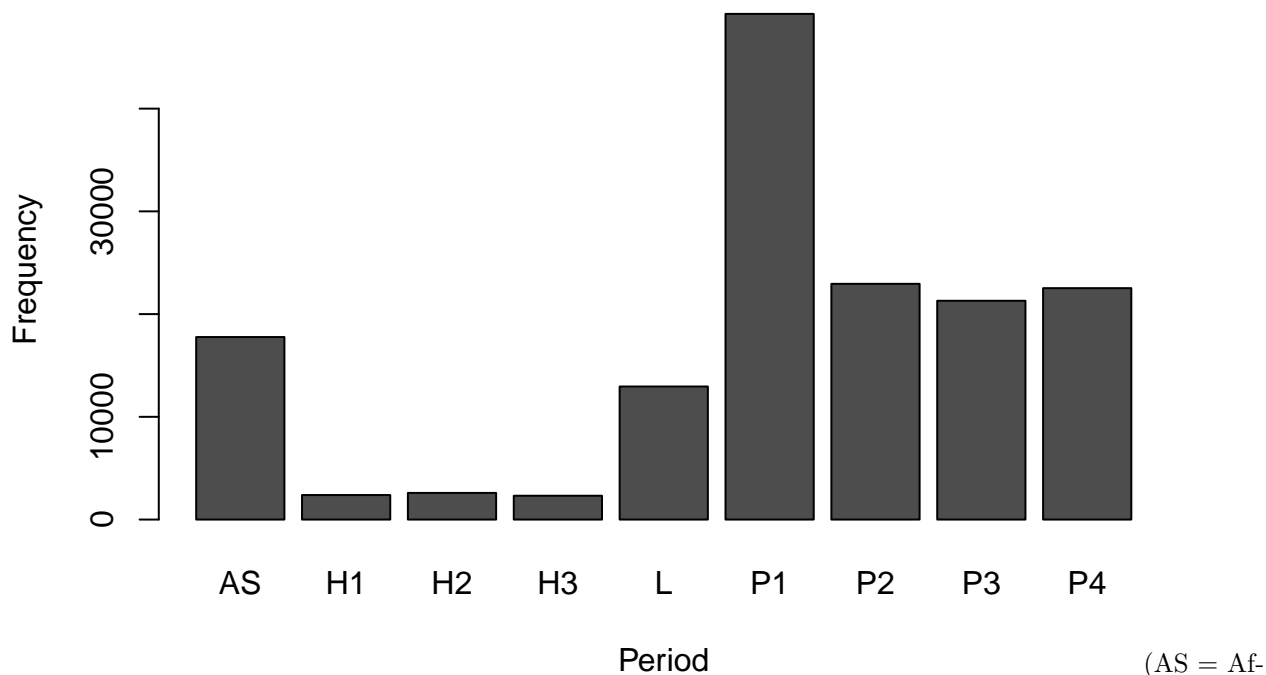
I graphed the distribution for the number of interactions during each hallway passing, period, lunch, and after school. Since sensors were handed out to students in period 1, there was an overestimation of the number of interactions for this period on Day 1. Since sensors were collected in period 4 on Day 2, there was an overestimation of the number of interactions for this period.

For all graphs, P = Period, H = Hallway Passing, and AS = After school. Since there were four periods, P1, P2, P3, and P4 denoted periods 1-4. H1, H2, and H3 denoted hallway passings 1-3.

Day 1

```
## After School Hallway1 Hallway2 Hallway3 Lunch Period1 Period2 Period3 Period4
##
##      17765      2385      2592      2317 12951  49223  22944  21297  22525
```

**Distribution of contacts (Day 1)**

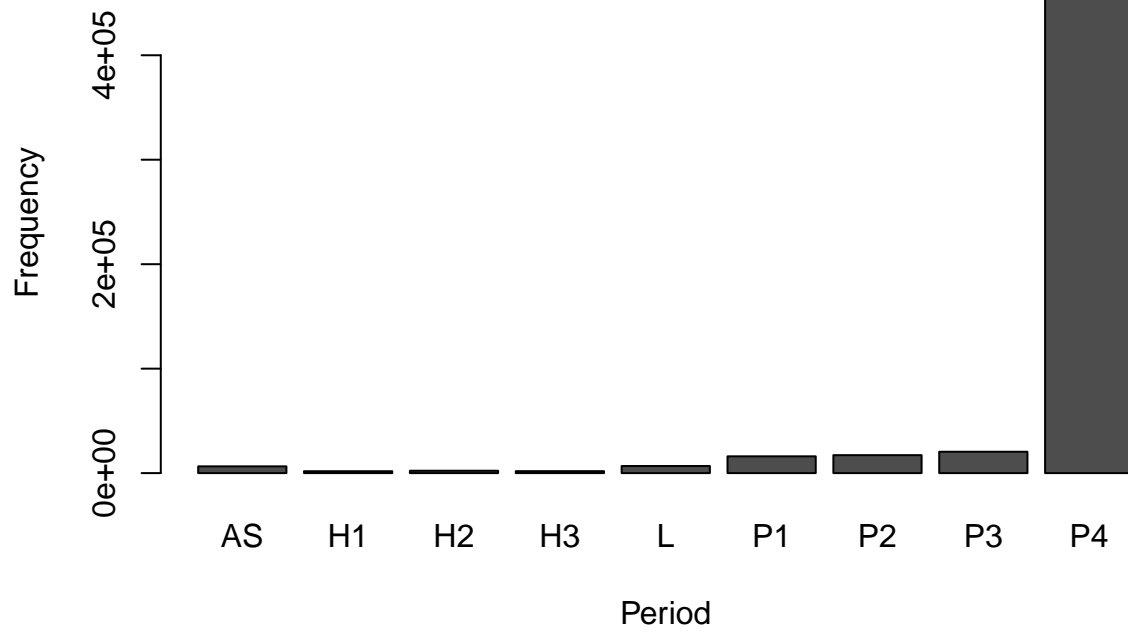


ter School)

Day 2

```
## Before School Hallway1 Hallway2 Hallway3 Lunch Period1 Period2 Period3 Period4
##
##      6446      1821      2269      1863  6795  16050  17182  20441  484068
```

## Distribution of contacts (Day 2)

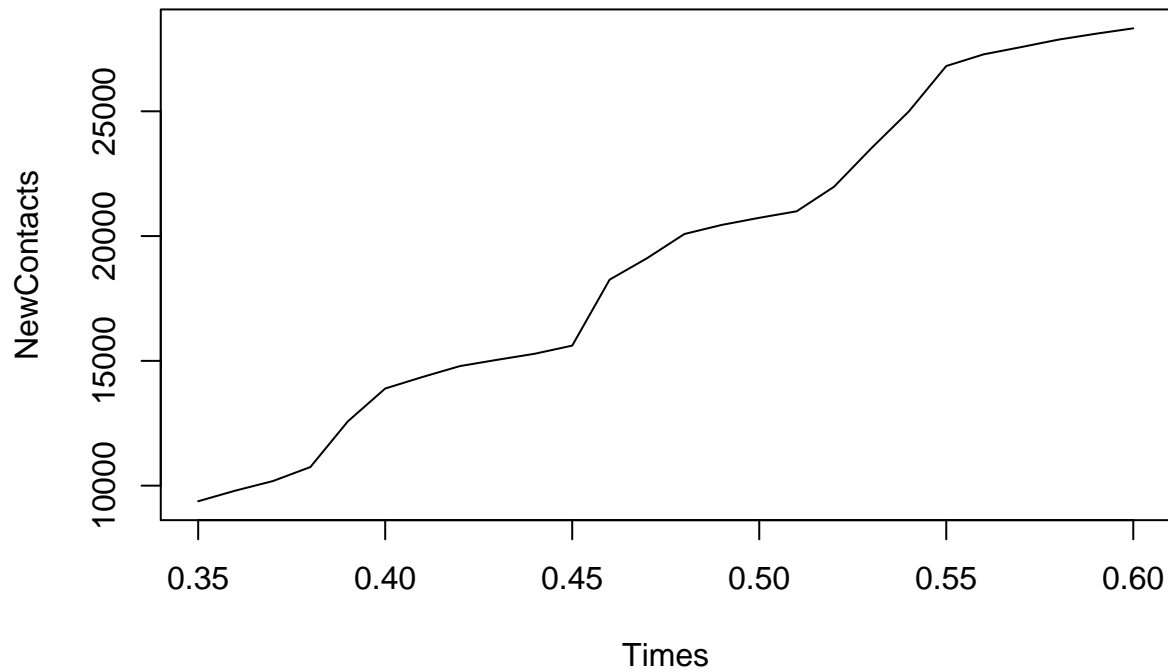


## Interactions Throughout the Day

The two plots showed the new interactions as the day progressed.

Day 1

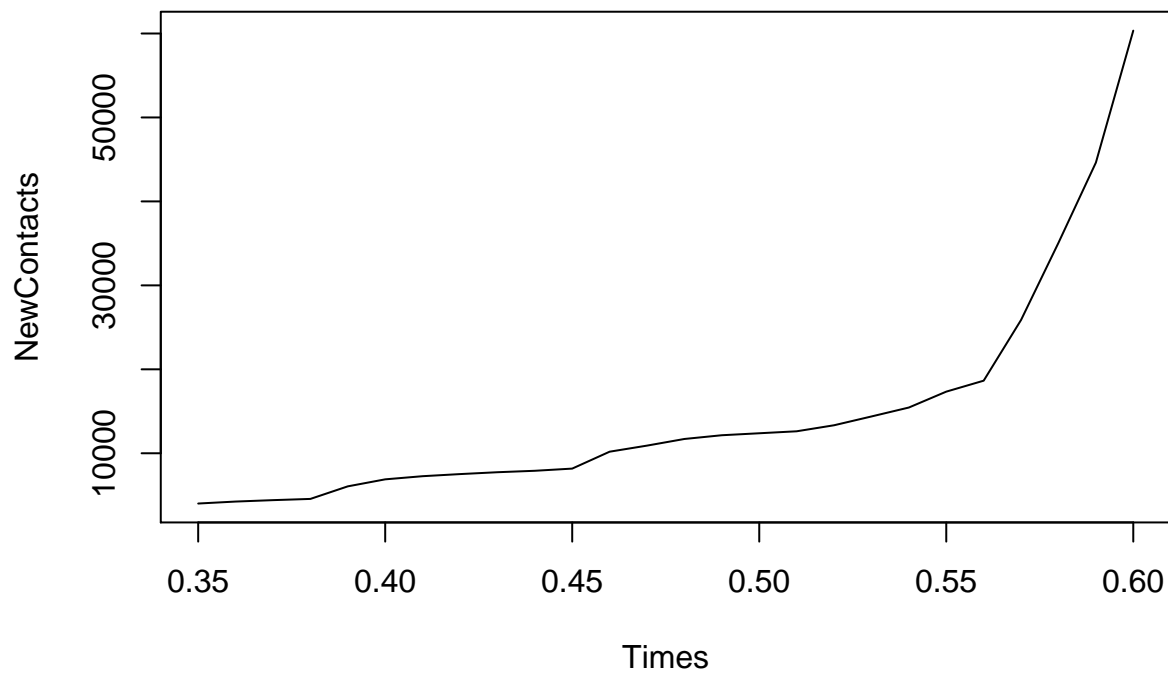
### New Contacts (Day 1)



There were three step-ups in the graph representing the start of a new period or hallway passing. This occurred since students were more likely to interact with other students they had not seen yet.

Day 2

### New Contacts (Day 2)



I examined different social networking attributes including degree, betweenness, closeness, transitivity, density

and assortativity. For degree, betweenness, and closeness, transitivity, and density I used the same technique to code the graphs and tables.

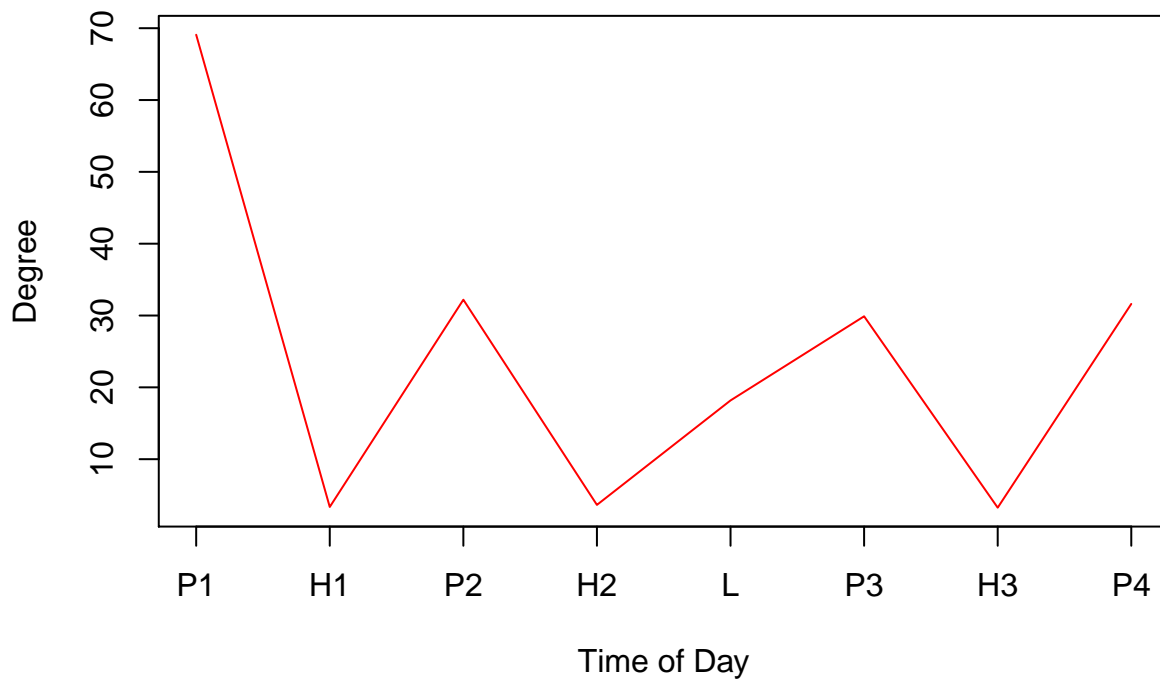
## Degree

Degree represented the number of edges for each node.

Day 1

##	Measure	Min	Q1	Median	Mean	Q3	Max
## 1	All day	2	136	207	216.138947	282	704
## 2	All periods	1	98	154	162.791579	215	500
## 3	Period 1	0	35	61	69.084912	97	307
## 4	Hallway 1	0	1	3	3.347368	5	30
## 5	Period 2	0	11	26	32.202105	46	185
## 6	Hallway 2	0	1	3	3.637895	5	32
## 7	Lunch	0	7	16	18.176842	26	98
## 8	Period 3	0	8	23	29.890526	45	159
## 9	Hallway 3	0	1	2	3.251930	5	20
## 10	Period 4	0	8	24	31.614035	46	182

**Degree (Mean) (Day 1)**

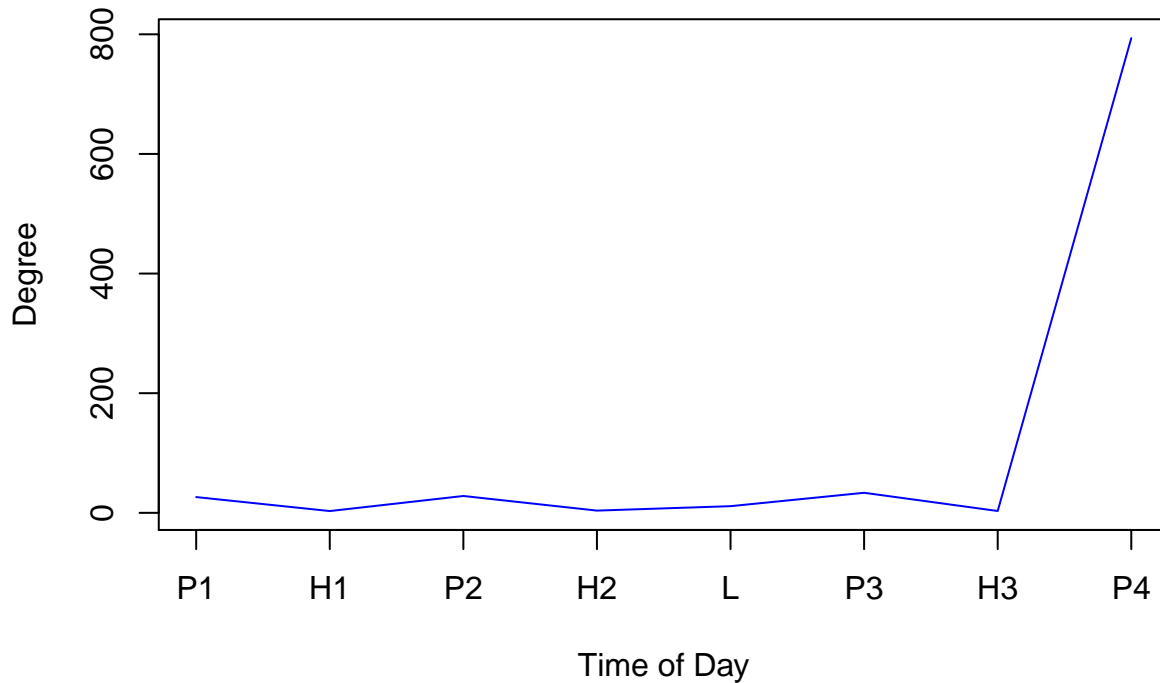


Day 2

##	Measure	Min	Q1	Median	Mean	Q3	Max
## 1	All day	1	117	208	913.008197	523.50	5223
## 2	All periods	0	102	184	892.109836	481.75	5209
## 3	Period 1	0	5	18	26.311475	39.00	152
## 4	Hallway 1	0	1	2	2.985246	4.00	18
## 5	Period 2	0	7	22	28.167213	40.25	206
## 6	Hallway 2	0	1	3	3.719672	5.00	41

## 7	Lunch	0	3	8	11.139344	16.00	74
## 8	Period 3	0	13	27	33.509836	48.00	158
## 9	Hallway 3	0	1	2	3.054098	5.00	16
## 10	Period 4	0	19	50	793.554098	344.00	5151

### Degree (Mean) (Day 2)



## Betweenness

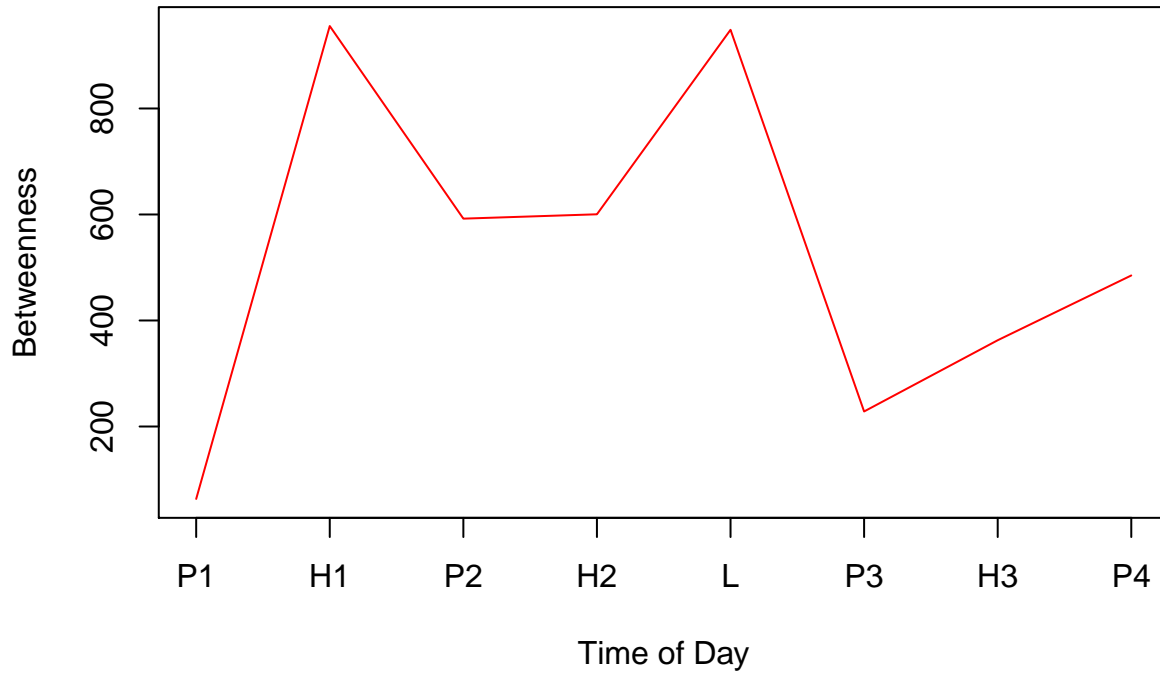
Betweenness represented the nodes that control the “communication flow” in a network. Students with higher betweenness helped connect other students with one another.

Day 1

##	Measure	Min	Q1	Median	Mean	Q3
## 1	All day	0.1567588	334.93022869	772.89874	956.1979	1354.5604
## 2	All periods	0.0000000	437.95946877	916.32280	1203.0035	1708.0332
## 3	Period 1	0.0000000	1.86578144	63.36503	3461.3235	690.5111
## 4	Hallway 1	0.0000000	0.00000000	955.51833	2447.9186	3315.8083
## 5	Period 2	0.0000000	0.62557498	592.16790	4045.6393	3819.8502
## 6	Hallway 2	0.0000000	0.00000000	600.36704	2119.7972	2719.8638
## 7	Lunch	0.0000000	154.81162615	948.43728	1914.1972	2483.0749
## 8	Period 3	0.0000000	0.01320622	228.33839	4554.3060	2941.0047
## 9	Hallway 3	0.0000000	0.00000000	362.73761	1868.2112	2352.8285
## 10	Period 4	0.0000000	0.08803701	484.93794	3256.3832	3397.2198
##	Max					
## 1		6683.899				
## 2		9635.527				
## 3		220261.761				
## 4		39157.746				
## 5		97654.775				

```
## 6 43118.184
## 7 48022.364
## 8 130600.730
## 9 36580.794
## 10 99725.299
```

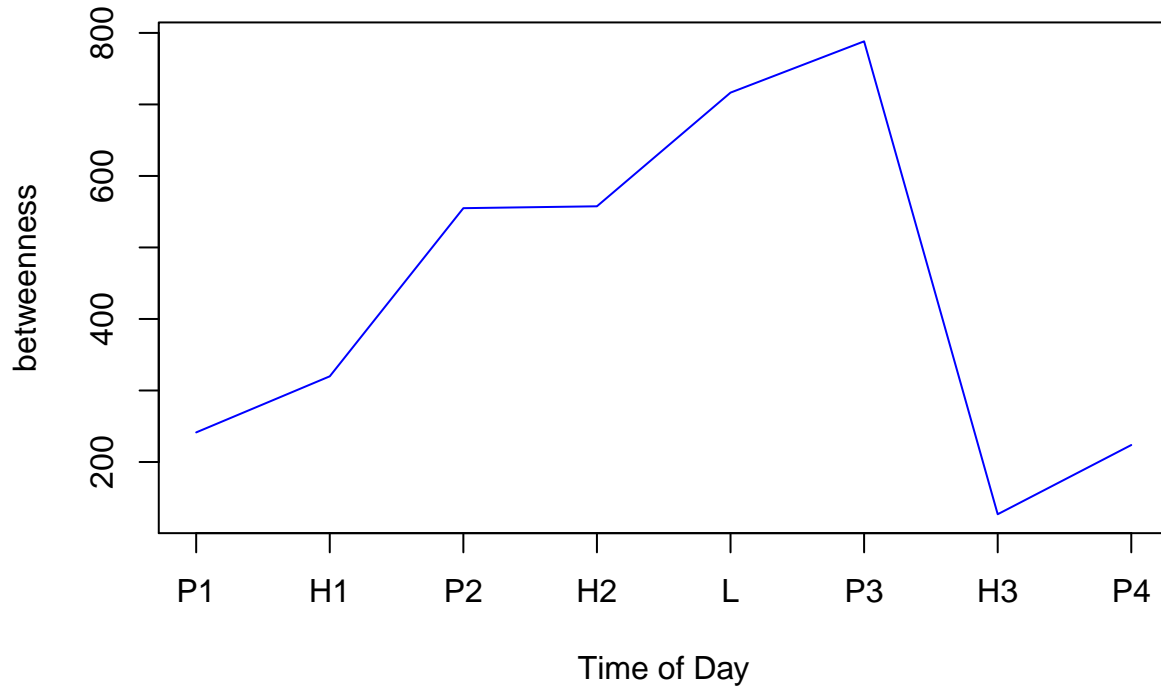
### Betweenness (Median) (Day 1)



Day 2

##	Measure	Min	Q1	Median	Mean	Q3	Max
## 1	All day	0	145.19109	369.3038	698.2984	909.9495	6178.156
## 2	All periods	0	142.07282	380.7294	736.4984	934.5057	6283.603
## 3	Period 1	0	0.00000	241.4009	2539.9492	2621.2555	94858.865
## 4	Hallway 1	0	0.00000	319.7882	1686.9025	2210.1744	32879.639
## 5	Period 2	0	0.00000	554.9565	2893.9148	3587.7324	55082.512
## 6	Hallway 2	0	0.00000	557.5593	1895.7697	2539.1902	26615.949
## 7	Lunch	0	0.00000	716.4939	2312.4434	2747.3453	44071.933
## 8	Period 3	0	152.47484	788.2054	1838.9770	2260.1258	45176.072
## 9	Hallway 3	0	0.00000	126.9584	1456.1705	1898.8175	30874.107
## 10	Period 4	0	26.13226	223.7442	1372.7361	1331.8410	22802.214

## Betweenness (Median) (Day 2)



## Closeness

Closeness measured how close a node was to all other nodes in the network.

Day 1

```
## Warning in closeness(westPeriod1): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

## Warning in closeness(westPeriod2): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

## Warning in closeness(westPeriod3): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

## Warning in closeness(westPeriod4): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

## Warning in closeness(westLunch): At centrality.c:2784 :closeness centrality
## is not well-defined for disconnected graphs

## Warning in closeness(westHallway1): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

## Warning in closeness(westHallway2): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

## Warning in closeness(westHallway3): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

##      Measure      Min      Q1      Median      Mean
## 1      All day 2.134016e-04 2.866972e-04 3.020236e-04 3.011528e-04
```

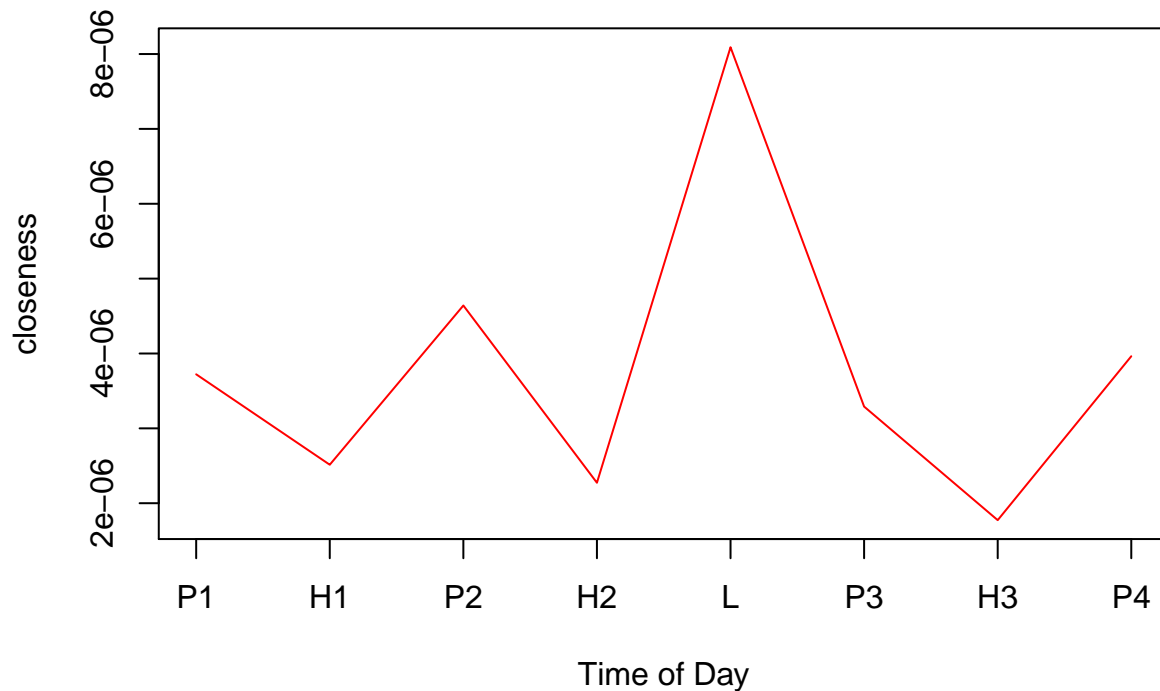


```

## 2 All periods 1.476233e-04 2.535497e-04 2.629503e-04 2.619763e-04
## 3 Period 1 4.928050e-07 3.702127e-06 3.722163e-06 3.312085e-06
## 4 Hallway 1 4.928050e-07 2.504182e-06 2.514199e-06 2.125963e-06
## 5 Period 2 4.928050e-07 4.610122e-06 4.641082e-06 4.222076e-06
## 6 Hallway 2 4.928050e-07 2.264559e-06 2.273203e-06 1.894226e-06
## 7 Lunch 4.928050e-07 8.053540e-06 8.090615e-06 7.643305e-06
## 8 Period 3 4.928050e-07 3.264667e-06 3.290903e-06 2.887783e-06
## 9 Hallway 3 4.928050e-07 4.931511e-07 1.774075e-06 1.423272e-06
## 10 Period 4 4.928050e-07 3.943482e-06 3.963221e-06 3.546797e-06
##      Q3      Max
## 1 3.165559e-04 3.526093e-04
## 2 2.718130e-04 3.082614e-04
## 3 3.736041e-06 3.765386e-06
## 4 2.518892e-06 2.526899e-06
## 5 4.661440e-06 4.705882e-06
## 6 2.277074e-06 2.285046e-06
## 7 8.115500e-06 8.179825e-06
## 8 3.303263e-06 3.326835e-06
## 9 1.776483e-06 1.780877e-06
## 10 3.976649e-06 4.004934e-06

```

### closeness (Median) (Day 1)



Day 2

```

## Warning in closeness(west2Period): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs
## Warning in closeness(west2Period1): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs
## Warning in closeness(west2Period2): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

```

```

## Warning in closeness(west2Period3): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

## Warning in closeness(west2Period4): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

## Warning in closeness(west2Lunch): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

## Warning in closeness(west2Hallway1): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

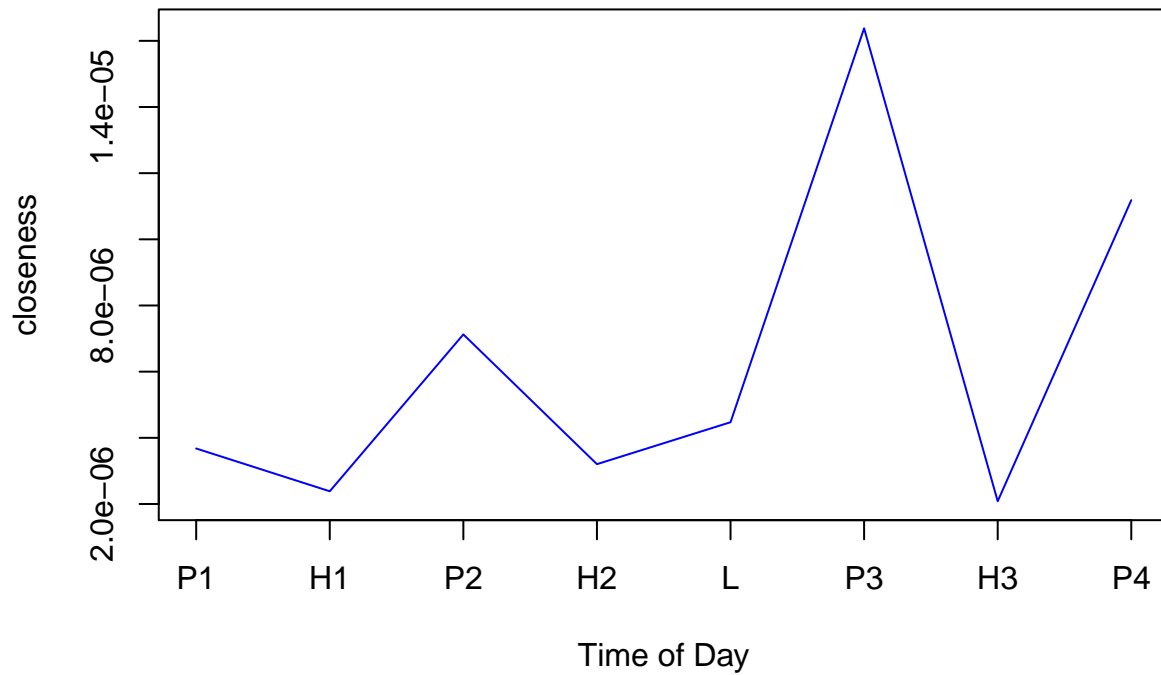
## Warning in closeness(west2Hallway2): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

## Warning in closeness(west2Hallway3): At centrality.c:2784 :closeness
## centrality is not well-defined for disconnected graphs

##      Measure      Min      Q1      Median      Mean
## 1      All day 2.391772e-04 3.524850e-04 3.757280e-04 3.886689e-04
## 2 All periods 6.724136e-07 2.403846e-04 2.510040e-04 2.577183e-04
## 3   Period 1 6.724136e-07 3.654420e-06 3.677174e-06 3.143374e-06
## 4   Hallway 1 6.724136e-07 6.729652e-07 2.384381e-06 1.909823e-06
## 5   Period 2 6.724136e-07 7.073874e-06 7.125679e-06 6.544543e-06
## 6   Hallway 2 6.724136e-07 3.188793e-06 3.204353e-06 2.684412e-06
## 7     Lunch 6.724136e-07 4.451265e-06 4.471292e-06 3.917690e-06
## 8   Period 3 6.724136e-07 1.626076e-05 1.637774e-05 1.577242e-05
## 9   Hallway 3 6.724136e-07 6.729652e-07 2.083062e-06 1.634518e-06
## 10  Period 4 6.724136e-07 1.106960e-05 1.118374e-05 1.057664e-05
##      Q3      Max
## 1 4.314529e-04 4.791567e-04
## 2 2.797398e-04 2.998501e-04
## 3 3.686133e-06 3.709446e-06
## 4 2.388104e-06 2.394768e-06
## 5 7.156506e-06 7.235419e-06
## 6 3.210832e-06 3.221847e-06
## 7 4.482948e-06 4.510335e-06
## 8 1.647202e-05 1.669533e-05
## 9 2.086366e-06 2.090594e-06
## 10 1.130135e-05 1.135525e-05

```

## Closeness (Median) (Day 2)

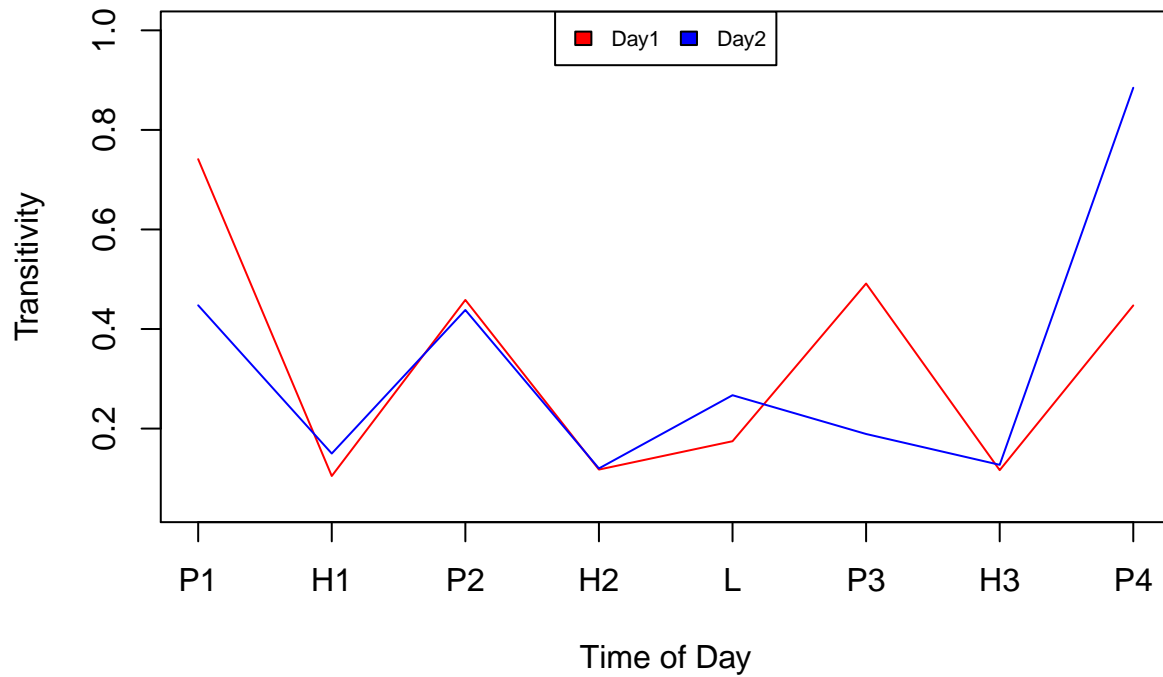


## Transitivity

Transitivity measured the likelihood nodes will cluster together.

##	Measure	Day1	Day2
## 1	All day	0.1823022	0.7506676
## 2	All periods	0.2978069	0.7779085
## 3	Period 1	0.7412191	0.4475851
## 4	Hallway 1	0.1046884	0.1498767
## 5	Period 2	0.4584243	0.4382355
## 6	Hallway 2	0.1176294	0.1198721
## 7	Lunch	0.1746072	0.2670097
## 8	Period 3	0.4913340	0.1891207
## 9	Hallway 3	0.1164544	0.1271959
## 10	Period 4	0.4473970	0.8846711

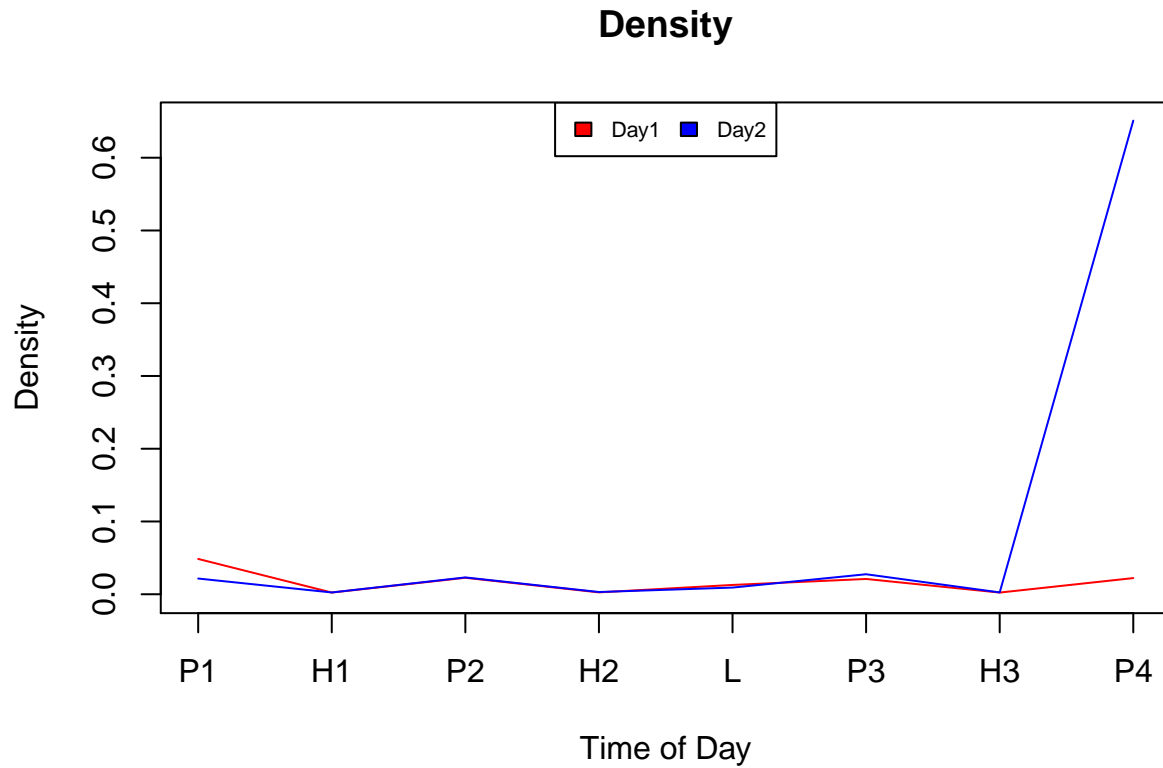
## Transitivity



## Density

Density was the proportion of actual edges divided by potential edges.

##	Measure	Day1	Day2
## 1	All day	0.151782969	0.748981293
## 2	All periods	0.114319929	0.731837437
## 3	Period 1	0.048514686	0.021584475
## 4	Hallway 1	0.002350680	0.002448930
## 5	Period 2	0.022613838	0.023106820
## 6	Hallway 2	0.002554701	0.003051413
## 7	Lunch	0.012764636	0.009138100
## 8	Period 3	0.020990538	0.027489611
## 9	Hallway 3	0.002283659	0.002505413
## 10	Period 4	0.022200867	0.650987776



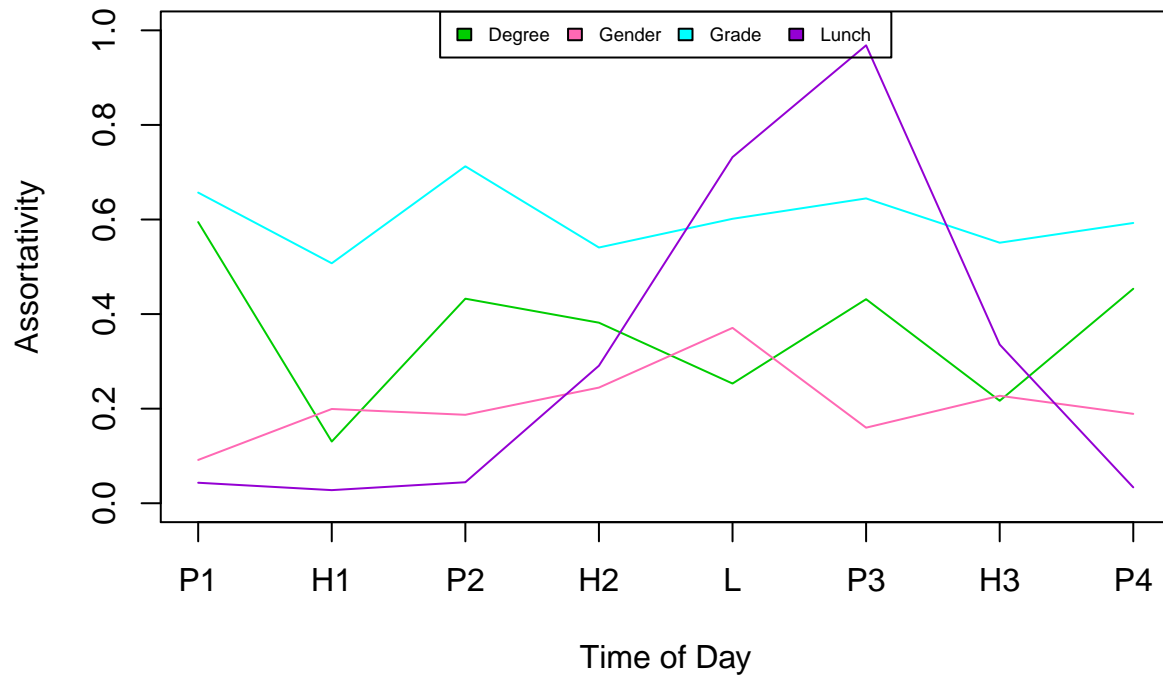
## Assortativity

Assortativity measured the probability that students interacted with other students with similar characteristics. The study evaluated four characteristics: gender, grade, lunch, and degree, and calculated the assortativity. Grade and Gender were coded as “GenderCat” and “GradeCat” to match other files.

Day 1

##	Measure	Degree	GenderCat	GradeCat	Lunch
## 1	All day	0.1729348	0.18532409	0.6176091	0.24153951
## 2	All periods	0.2143047	0.14201952	0.6555558	0.20689575
## 3	Period 1	0.5945900	0.09155421	0.6569461	0.04336572
## 4	Hallway 1	0.1304607	0.19934872	0.5074168	0.02772307
## 5	Period 2	0.4325414	0.18705206	0.7125158	0.04444116
## 6	Hallway 2	0.3817808	0.24454924	0.5407445	0.29084927
## 7	Lunch	0.2531446	0.37082585	0.6014776	0.73191453
## 8	Period 3	0.4314996	0.15979813	0.6446122	0.96825876
## 9	Hallway 3	0.2166636	0.22722294	0.5509142	0.33554356
## 10	Period 4	0.4535261	0.18897755	0.5926358	0.03361289

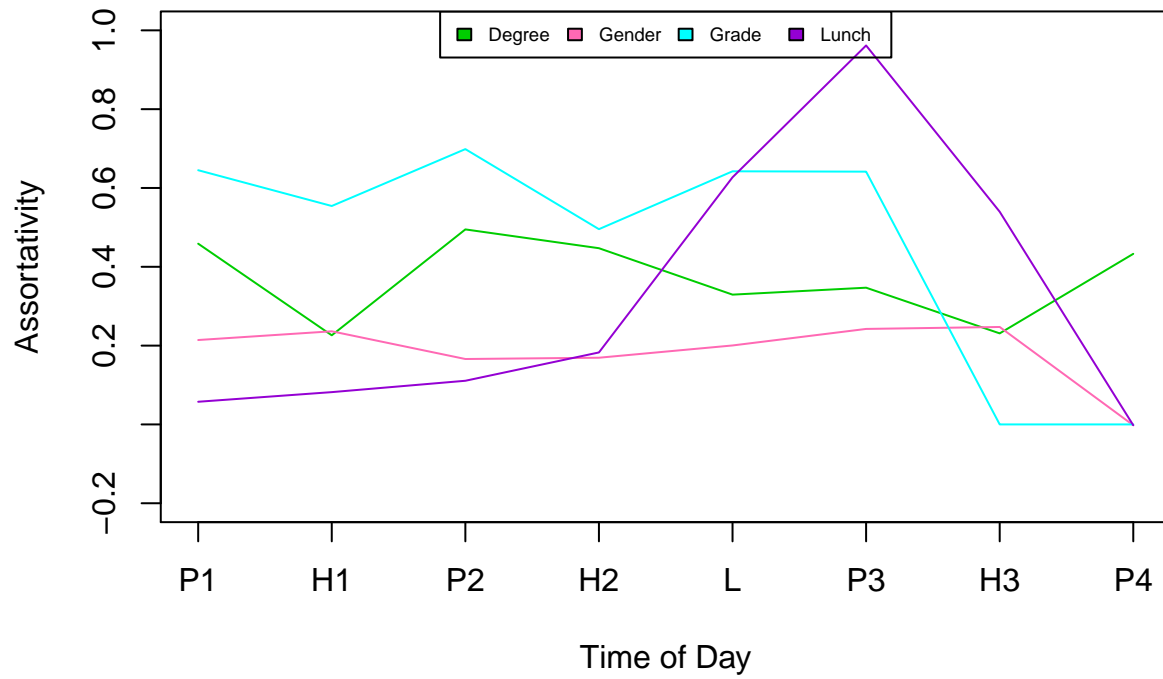
## Assortativity (Day1)



Day 2

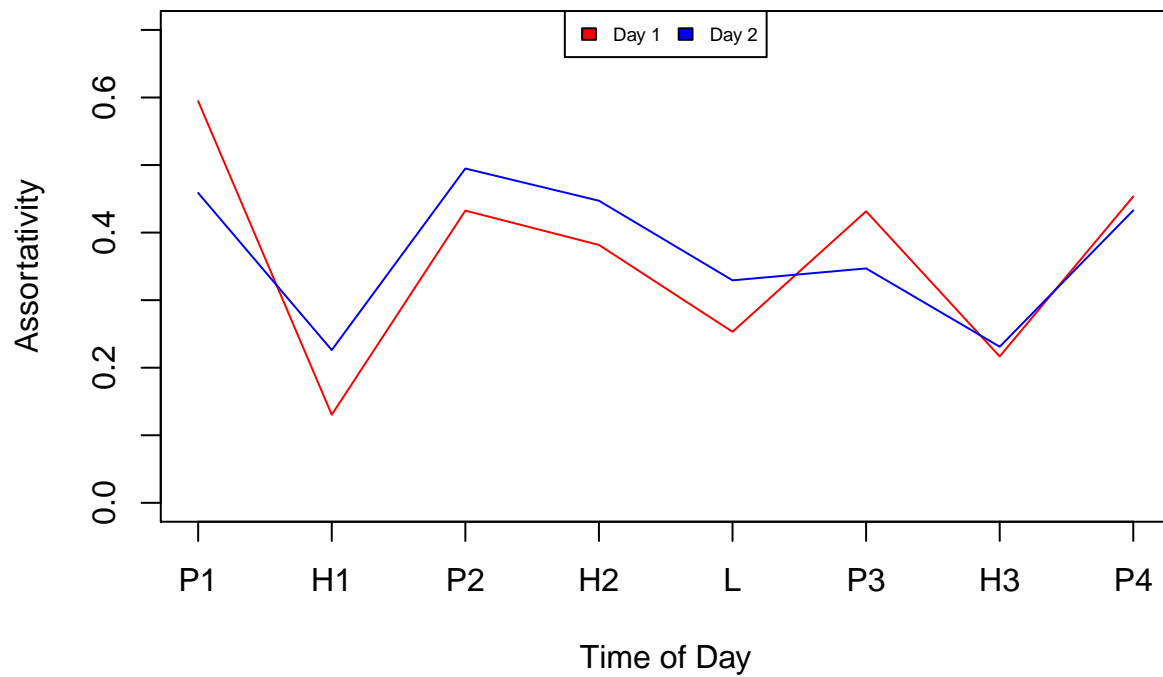
##	Measure	Degree	GenderCat	GradeCat	Lunch
## 1	All day	0.5618487	0.02708553	0.09573177	0.048471248
## 2	All periods	0.5514066	0.02267020	0.08264291	0.039003851
## 3	Period 1	0.4586588	0.21416905	0.64510699	0.057504029
## 4	Hallway 1	0.2261451	0.23631466	0.55427274	0.081951891
## 5	Period 2	0.4948827	0.16602239	0.69864308	0.110788210
## 6	Hallway 2	0.4471911	0.16923223	0.49546541	0.182620132
## 7	Lunch	0.3293667	0.20039163	0.64233551	0.627094628
## 8	Period 3	0.3469509	0.24230235	0.64143490	0.961614982
## 9	Hallway 3	0.2310550	0.24723797	0.00000000	0.539674517
## 10	Period 4	0.4328890	-0.00210218	0.00000000	-0.002105876

## Assortativity (Day2)

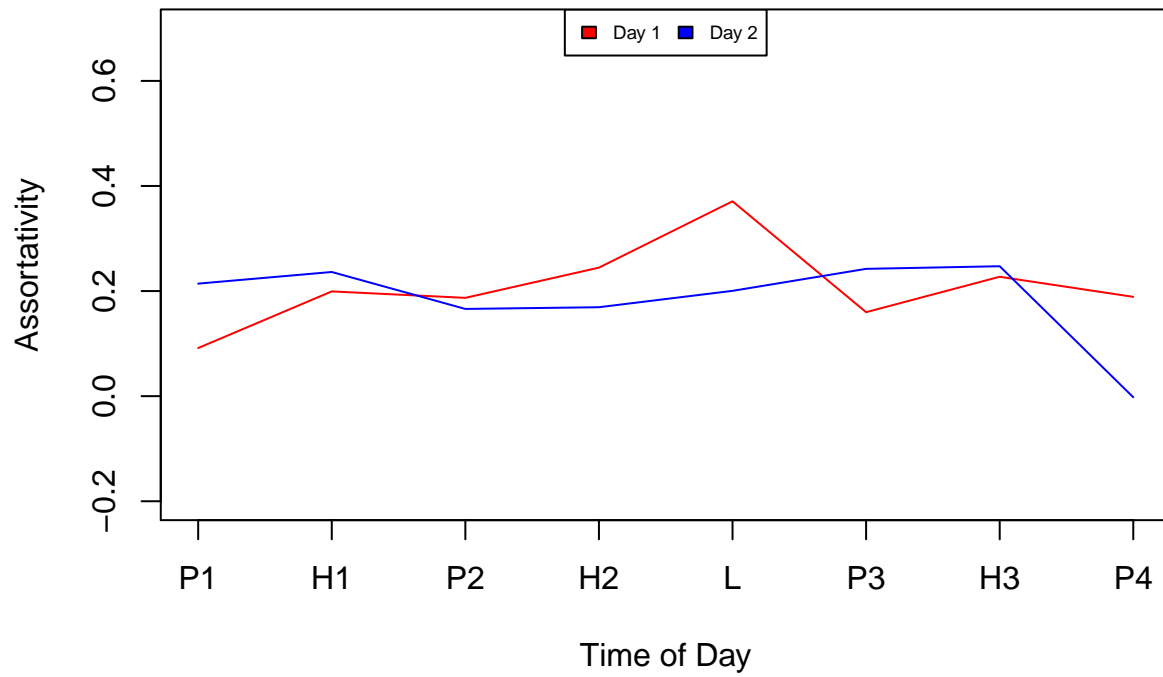


Comparing each characteristic between days:

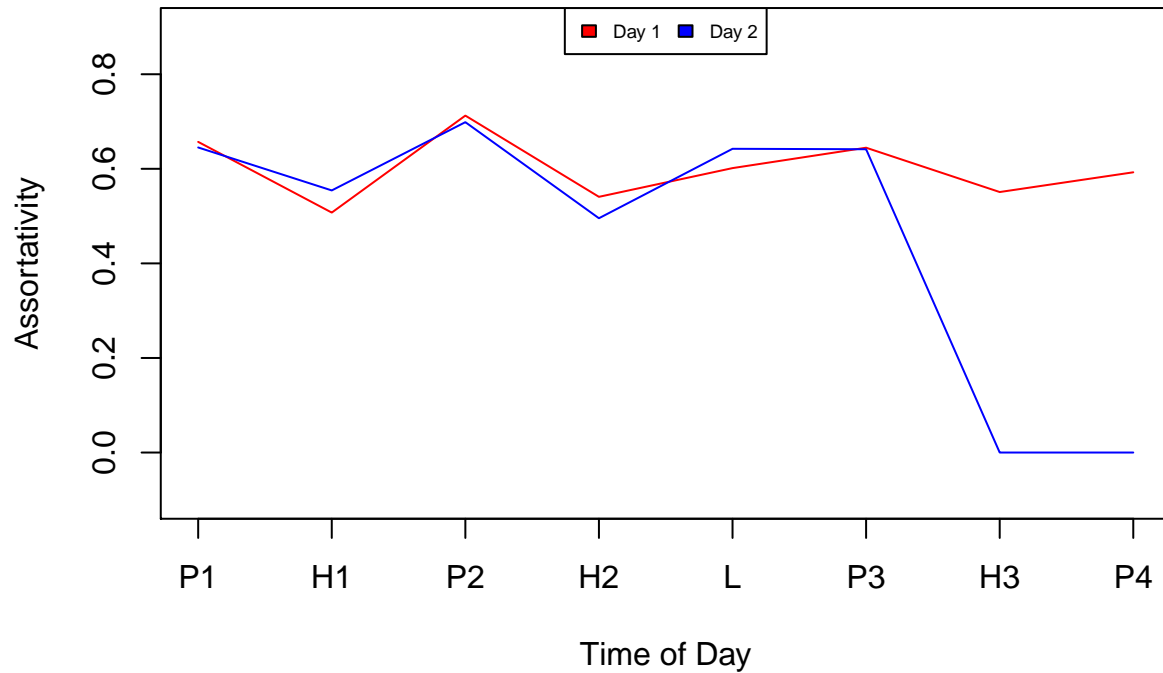
## Assortativity (Degree)



### Assortativity (Gender)



### Assortativity (Grade)





## Assortativity (Lunch)

