

“Modelación de flujo de turistas en Parque Nacional Torres del Paine”

Autores Derek Corcoran y Rodrigo Gomez

Last update: 2015-09-15

Estudio de distribución del flujo de turistas

Distribución horaria de las visitas

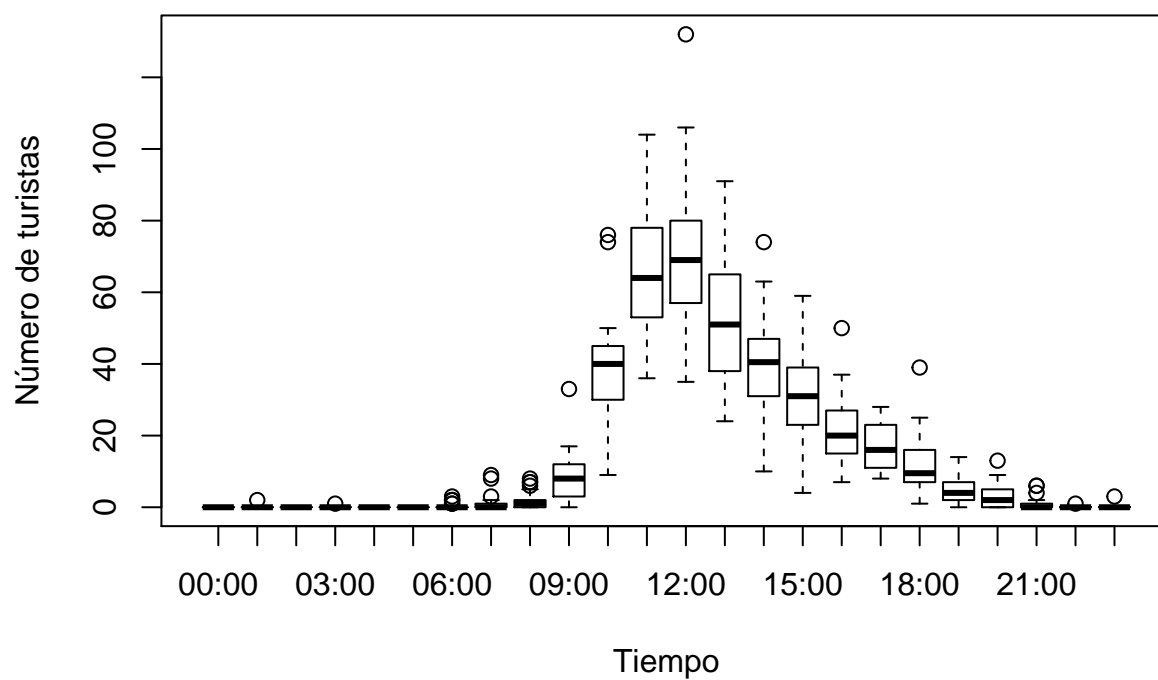
Inicialmente se realizará un estudio de la distribución de turistas para el parque Torres del Paine. En forma de ejemplo se utilizarán los datos disponibles actualmente para el contador de los Cuernos.

Usaremos las horas y fechas tanto como factores como tiempos por lo que agregaremos algunas variables

```
## 'data.frame':   622 obs. of  9 variables:
## $ date.as.factor : Factor w/ 27 levels "01/02/2015","02/02/2015",...: 24 24 24 24 24 24 24 24 24 24
## $ X.1            : Factor w/ 24 levels "00:00","01:00",...: 13 14 15 16 17 18 19 20 21 22 ...
## $ Eco.Cuernos    : int  97 36 16 17 24 24 16 5 1 0 ...
## $ Eco.Cuernos_IN : int  40 20 6 13 7 8 1 0 0 0 ...
## $ Eco.Cuernos_OUT: int  57 16 10 4 17 16 15 5 1 0 ...
## $ date.time      : POSIXct, format: "2015-01-28 12:00:00" "2015-01-28 13:00:00" ...
## $ date.as.date   : POSIXct, format: "2015-01-28" "2015-01-28" ...
## $ dia.semana     : chr  "miércoles" "miércoles" "miércoles" "miércoles" ...
## $ Hora           : int   12 13 14 15 16 17 18 19 20 21 ...
```

Primero que nada observemos lo que ocurre al graficar la cantidad de turistas totales agrupados por horas en que se mueven.

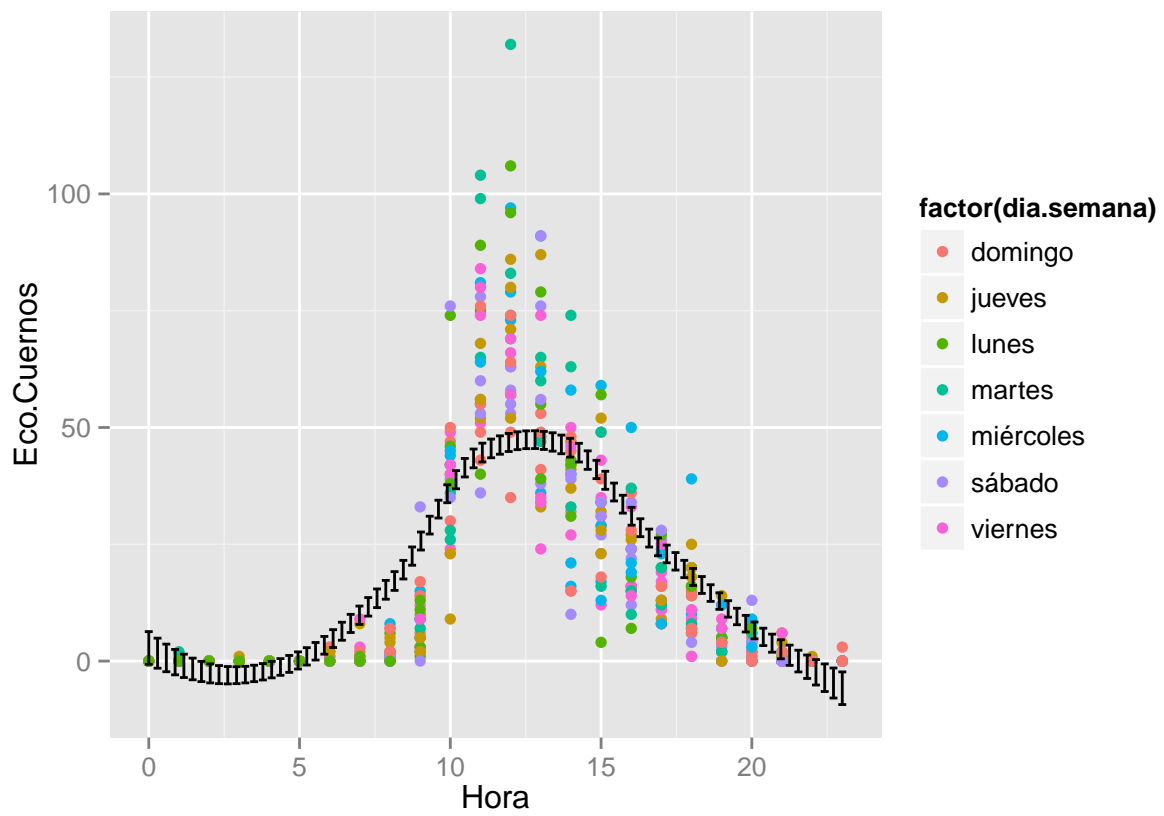
Contador Cuernos



```
p <- ggplot(Eco_Cuernos, aes(Hora, Eco.Cuernos))  
p + geom_point(aes(color = factor(dia.semana)))+ stat_smooth(geom = "errorbar")
```

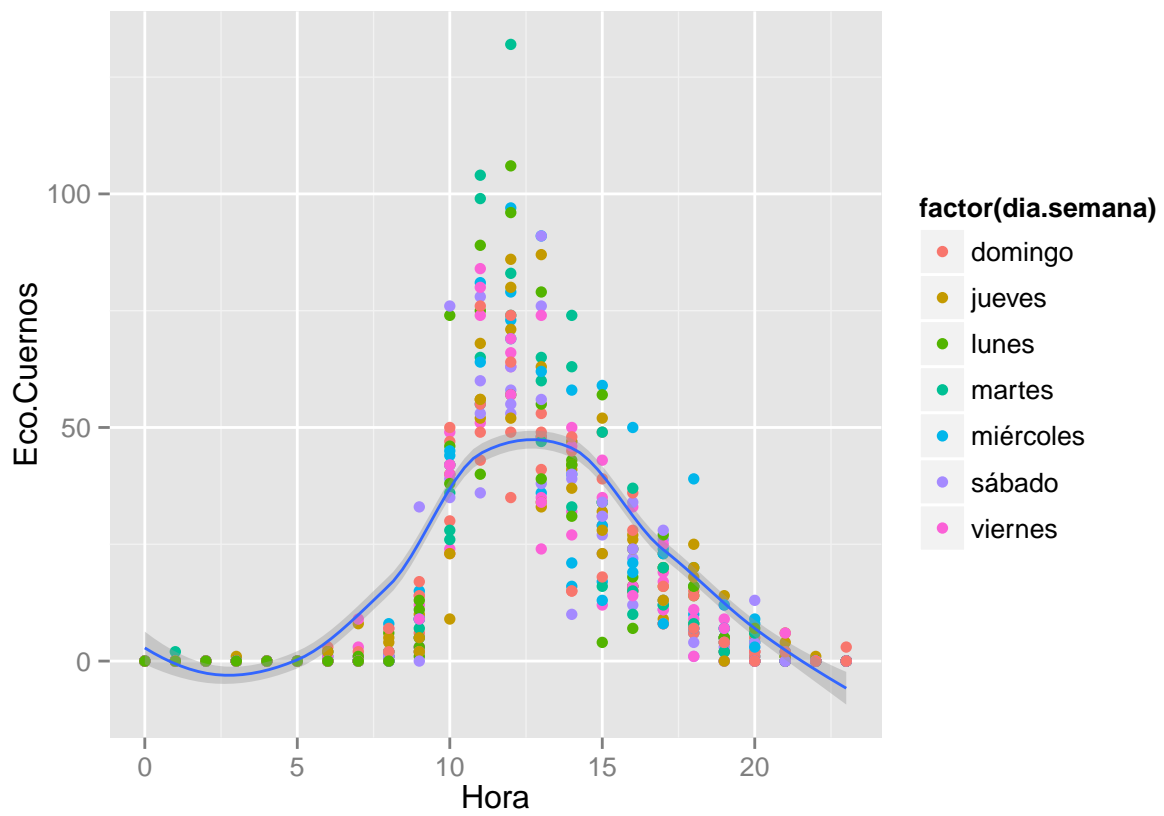
opcion grafica 2

geom_smooth: method="auto" and size of largest group is <1000, so using loess. Use 'method = x' to c

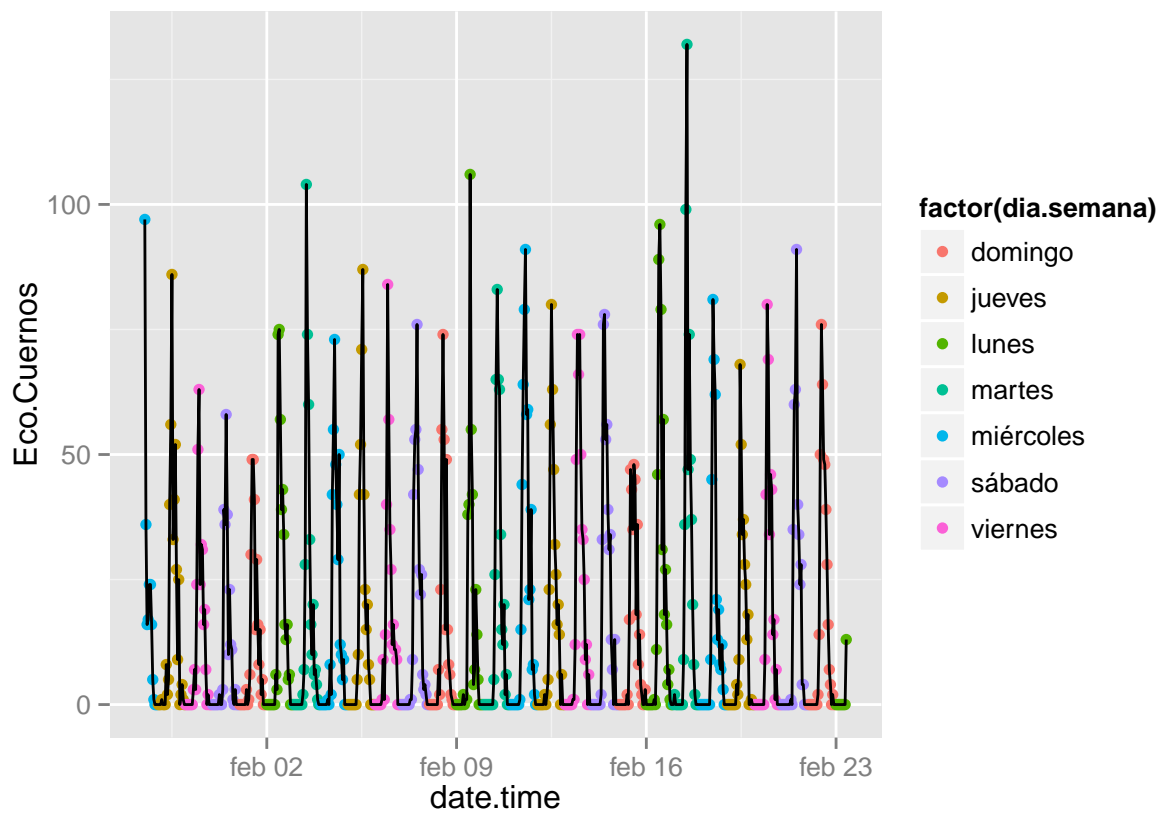


```
p + geom_point(aes(color = factor(dia.semana)))+ stat_smooth()
```

```
## geom_smooth: method="auto" and size of largest group is <1000, so using loess. Use 'method = x' to c
```

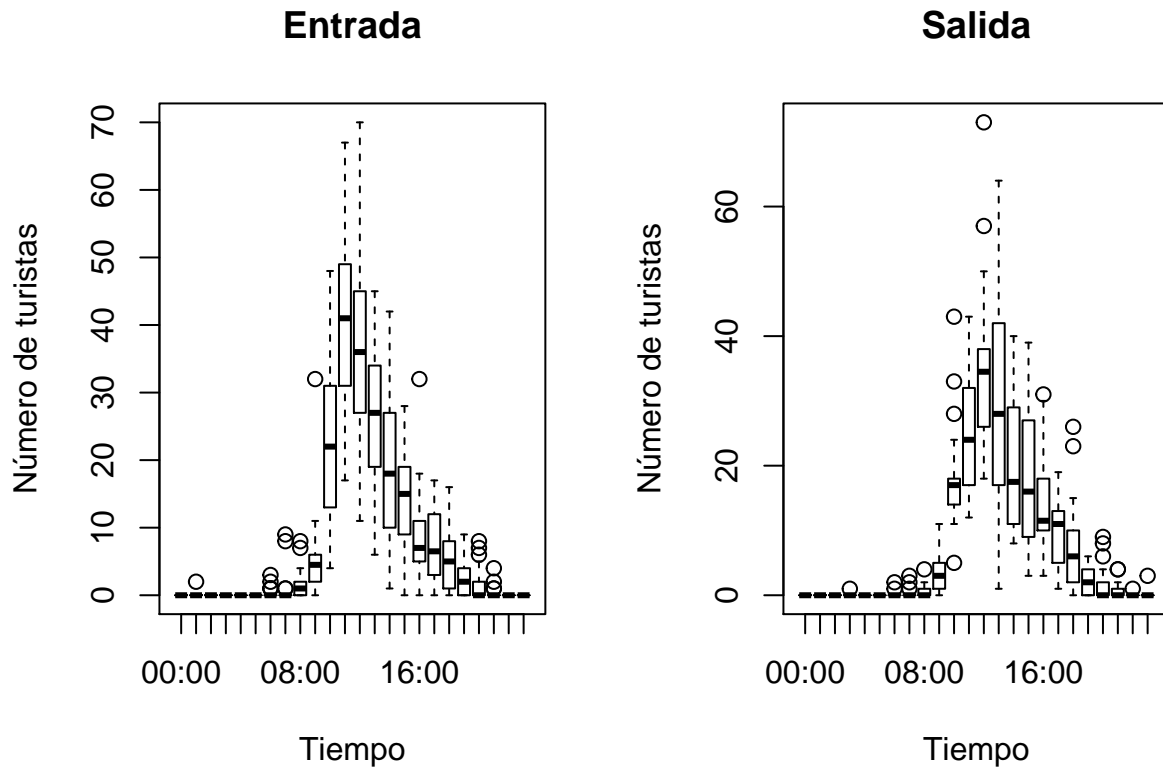


```
ts <- ggplot(Eco_Cuernos, aes(date.time, Eco.Cuernos))
ts + geom_point(aes(color = factor(dia.semana)))+ geom_line()
```



Observamos un peak en actividad a las 11 y 12 del día e inicios de actividad cercano a las 9 am finalizando al rededor de las 21 horas

Esto es para el total de turistas, si solo consideramos la entrada o salida de turistas vemos que esto cambia.



El estudiar estos flujos permitiría el estudiar los horarios con mayor número de encuentros en sentido contrario, lo cual no necesariamente debe coincidir con el numero total de turistas.

Análisis estadístico de factores (Día de semana y hora)

para estudiar esta relacion se realizara un analisis de dos vias, primero veamos un resumen de el numero de turistas por hora para cada dia

| lunes | martes | miércoles | jueves | viernes | sábado | domingo |
|-------|--------|-----------|--------|---------|--------|---------|
| 14.89 | 17.82 | 17.62 | 15.15 | 14.77 | 14.89 | 13.54 |

y el numero de turistas por hora

| 00:00 | 01:00 | 02:00 | 03:00 | 04:00 | 05:00 | 06:00 | 07:00 | 08:00 | 09:00 | 10:00 | 11:00 | 12:00 | 13:00 | 14:00 | 15:00 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0 | 0.08 | 0 | 0.04 | 0 | 0 | 0.42 | 0.96 | 2.08 | 8.42 | 39.6 | 65.72 | 71.58 | 54.54 | 38.54 | 31.46 |

Ahora realizamos un ANOVA de dos vias

```
attach(Eco_Cuernos)
```

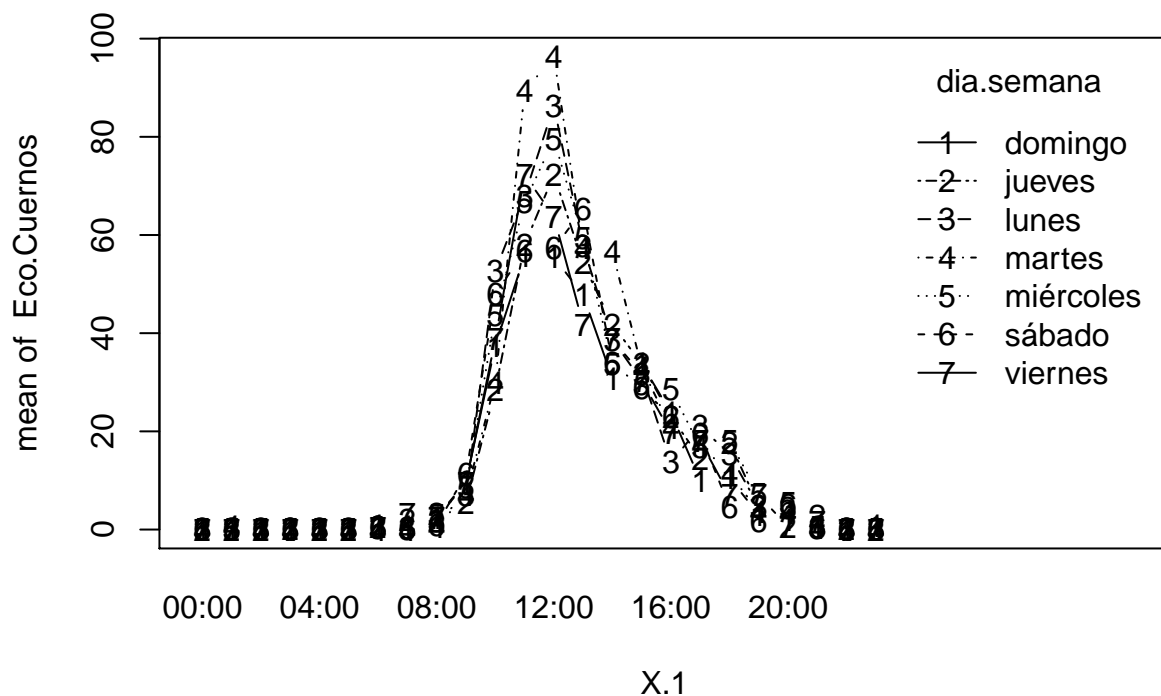
```
## The following objects are masked _by_ .GlobalEnv:
```

```
##
##      date.as.date, date.time
```

```
ANOVA.Cuernos <- aov(Eco.Cuernos ~ X.1*dia.semana)
summary(ANOVA.Cuernos)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## X.1           23 300703   13074 163.049 <2e-16 ***
## dia.semana     6   1102     184   2.290 0.0345 *
## X.1:dia.semana 138  12904     94   1.166 0.1237
## Residuals     454  36404     80
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
interaction.plot(X.1, dia.semana, Eco.Cuernos, type="b")
```



Comparación post-HOC via test de Tukey

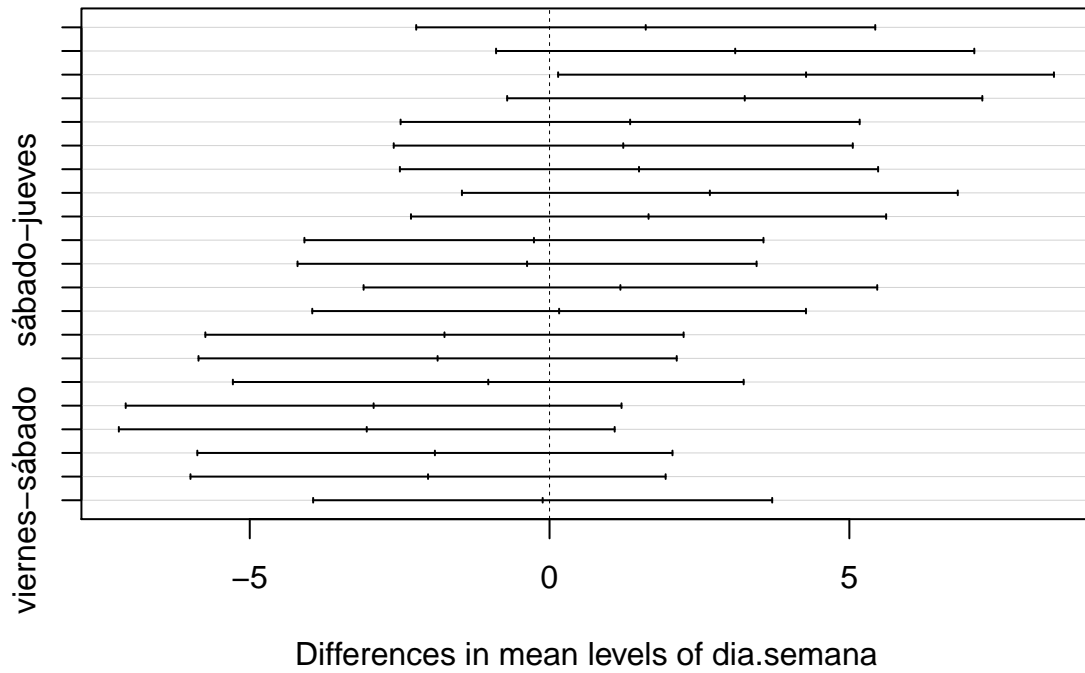
```
TukeyHSD(ANOVA.Cuernos, which="dia.semana", ordered=TRUE)
```

```
##      Tukey multiple comparisons of means
##      95% family-wise confidence level
##      factor levels have been ordered
##
```

```
## Fit: aov(formula = Eco.Cuernos ~ X.1 * dia.semana)
##
## $dia.semana
##          diff          lwr          upr          p adj
## viernes-domingo  1.2291667 -2.5986742  5.057008  0.9638370
## sábado-domingo   1.3437500 -2.4840909  5.171591  0.9445284
## jueves-domingo   1.6041667 -2.2236742  5.432008  0.8776575
## lunes-domingo    3.0960600 -0.8918188  7.083939  0.2465250
## miércoles-domingo 3.2555678 -0.7066240  7.217760  0.1870952
## martes-domingo   4.2777778  0.1432371  8.412318  0.0371476
## sábado-viernes    0.1145833 -3.7132576  3.942424  1.0000000
## jueves-viernes    0.3750000 -3.4528409  4.202841  0.9999512
## lunes-viernes     1.8668934 -2.1209855  5.854772  0.8089862
## miércoles-viernes 2.0264011 -1.9357907  5.988593  0.7359715
## martes-viernes    3.0486111 -1.0859296  7.183152  0.3063706
## jueves-sábado     0.2604167 -3.5674242  4.088258  0.9999944
## lunes-sábado      1.7523100 -2.2355688  5.740189  0.8512044
## miércoles-sábado  1.9118178 -2.0503740  5.874010  0.7858997
## martes-sábado     2.9340278 -1.2005129  7.068568  0.3533798
## lunes-jueves      1.4918934 -2.4959855  5.479772  0.9255740
## miércoles-jueves  1.6514011 -2.3107907  5.613593  0.8804509
## martes-jueves     2.6736111 -1.4609296  6.808152  0.4712810
## miércoles-lunes   0.1595077 -3.9575027  4.276518  0.9999998
## martes-lunes      1.1817177 -3.1014165  5.464852  0.9831695
## martes-miércoles  1.0222100 -3.2370181  5.281438  0.9919297
```

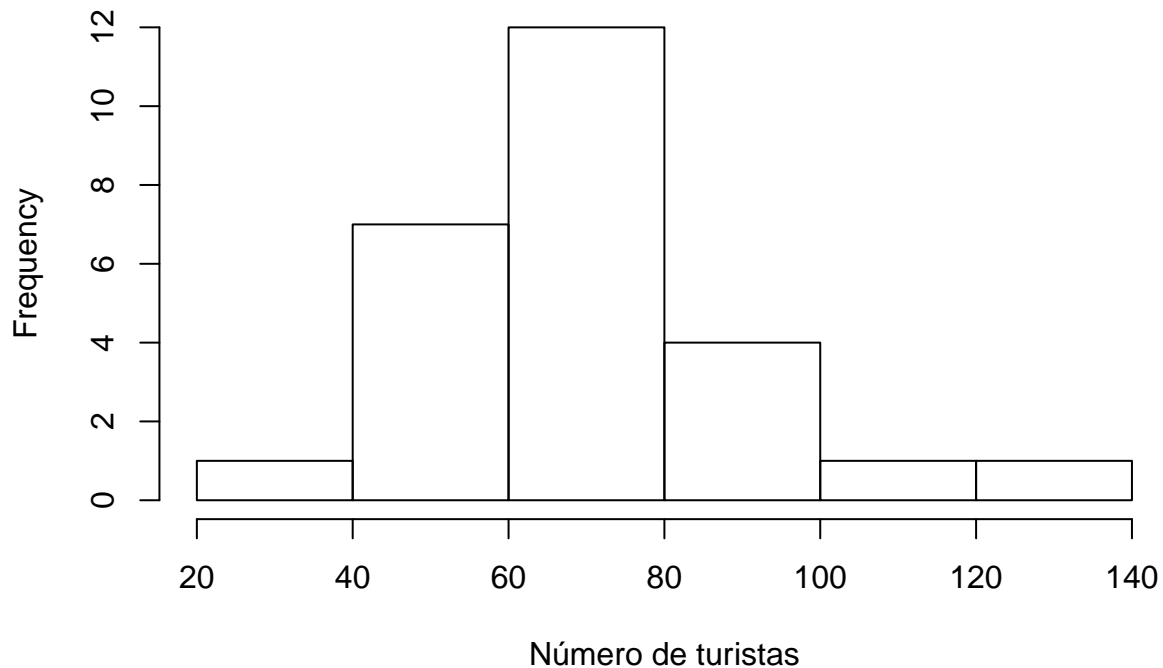
```
plot(TukeyHSD(ANOVA.Cuernos, which="dia.semana"))
```


95% family-wise confidence level



simulación

Para cada hora, se realizará un estudio de distribución del número de turistas que llegarían. Por ejemplo, para el mes de Febrero, a la hora peak (12:00), encontramos la siguiente histograma correspondiente a la distribución de turistas.



La cual corresponde a una distribución normal de acuerdo al test de Shapiro-wilk

```
shapiro.test(eco.cuernos.12$Eco.Cuernos)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  eco.cuernos.12$Eco.Cuernos
## W = 0.93584, p-value = 0.1068
```

Con una media de 71.5769231 y una desviación estándar de 20.2645959.

Dado esto, si hicieramos simulaciones aleatorias del número de turistas que pasarían por los cuernos en 20 días consecutivos tendríamos como resultado los siguientes 20 numeros.

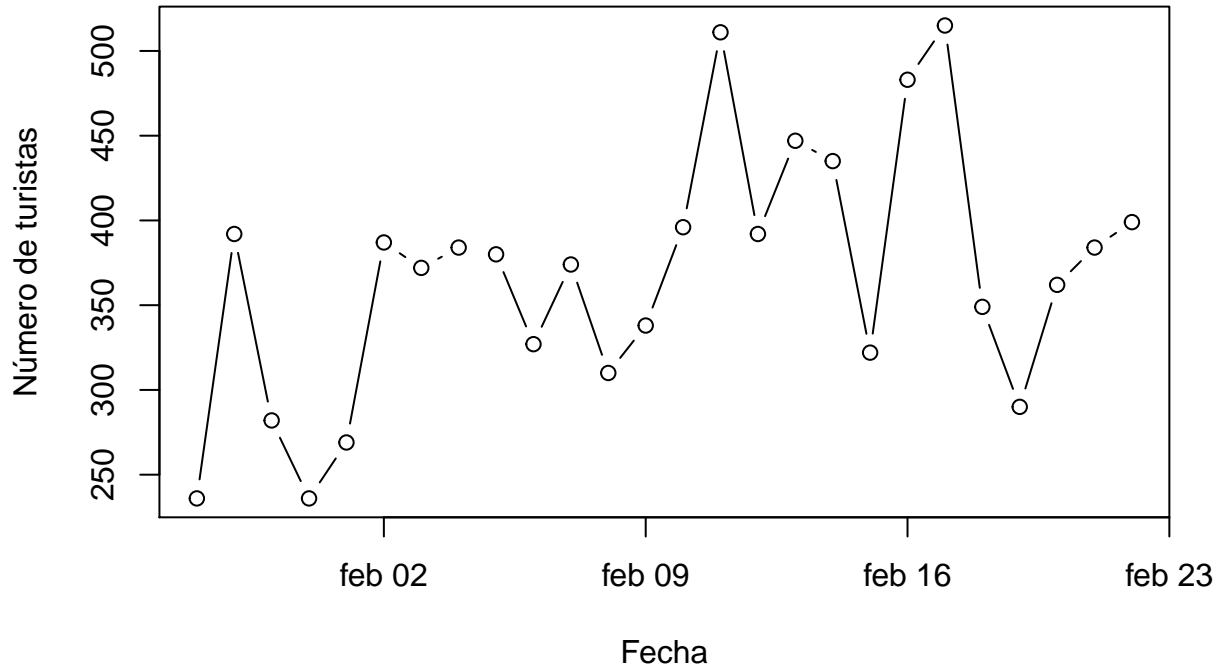
```
set.seed(100)
b<-rnorm(20, mean=mean(eco.cuernos.12$Eco.Cuernos), sd=sd(eco.cuernos.12$Eco.Cuernos))
print(b, digits=2)
```

```
## [1] 61 74 70 90 74 78 60 86 55 64 73 74 67 87 74 71 64
## [18] 82 53 118
```

Distribución diaria del uso por turistas

Además de esto se debe estudiar como varia por fecha la cantidad de turistas diarios, para esto sumaremos la cantidad de turistas al día y se estudiará el comportamiento de cada mes, así como los días de semana y fin

de semana.



Modelación espacialmente explícita

Estas caracterizaciones se utilizarán para una simulación espacialmente explícita utilizando modelaciones en base de agentes, apoyándose en la información generada por el resto de los grupos del CEQUA.