

# UEMC R

*Héctor Lozada*

*19 de noviembre de 2018*

## INFORME A SUAVECITO.ES

El siguiente estudio está basado en los datos publicados por el usuario Miguel Gil en la red social kaggle.

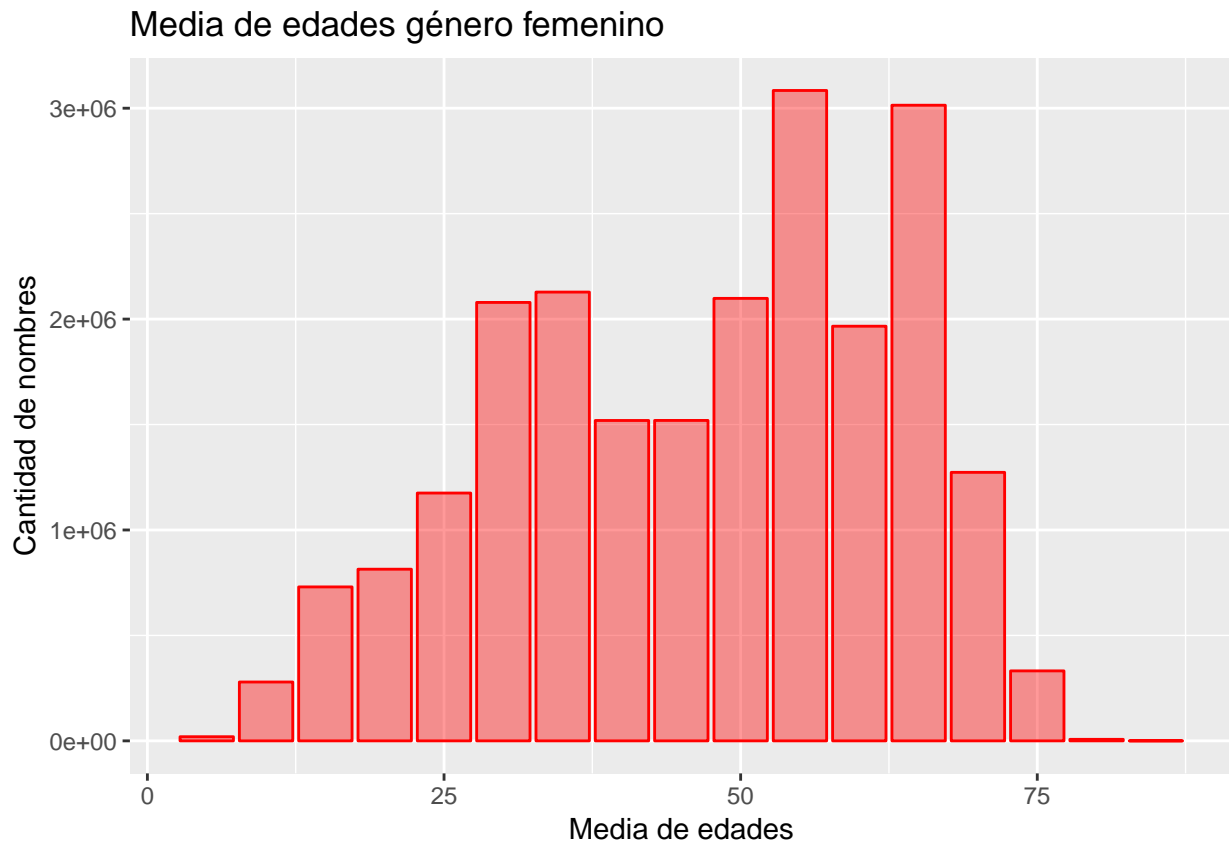
### Estudio y análisis de datos

Haciendo uso de técnicas de análisis de datos en el lenguaje de programación **R**, se ha determinado:

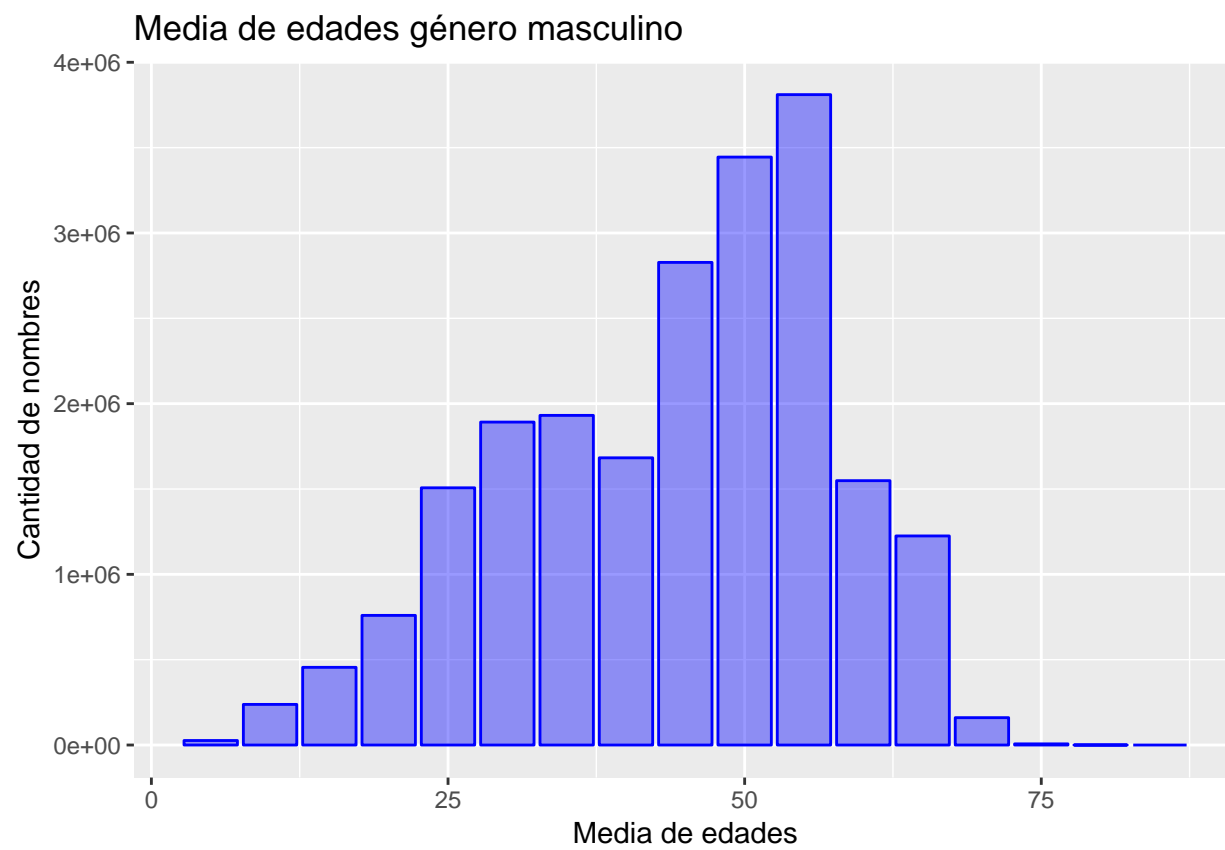
#### Cantidad de nombres femeninos y masculinos con mayor frecuencia

El nombre femenino mas repetido ha sido **MARIA CARMEN** con una aparición de **668 639**, por otra parte el nombre masculino mas repetido ha sido **ANTONIO** con una aparición de **715 215**

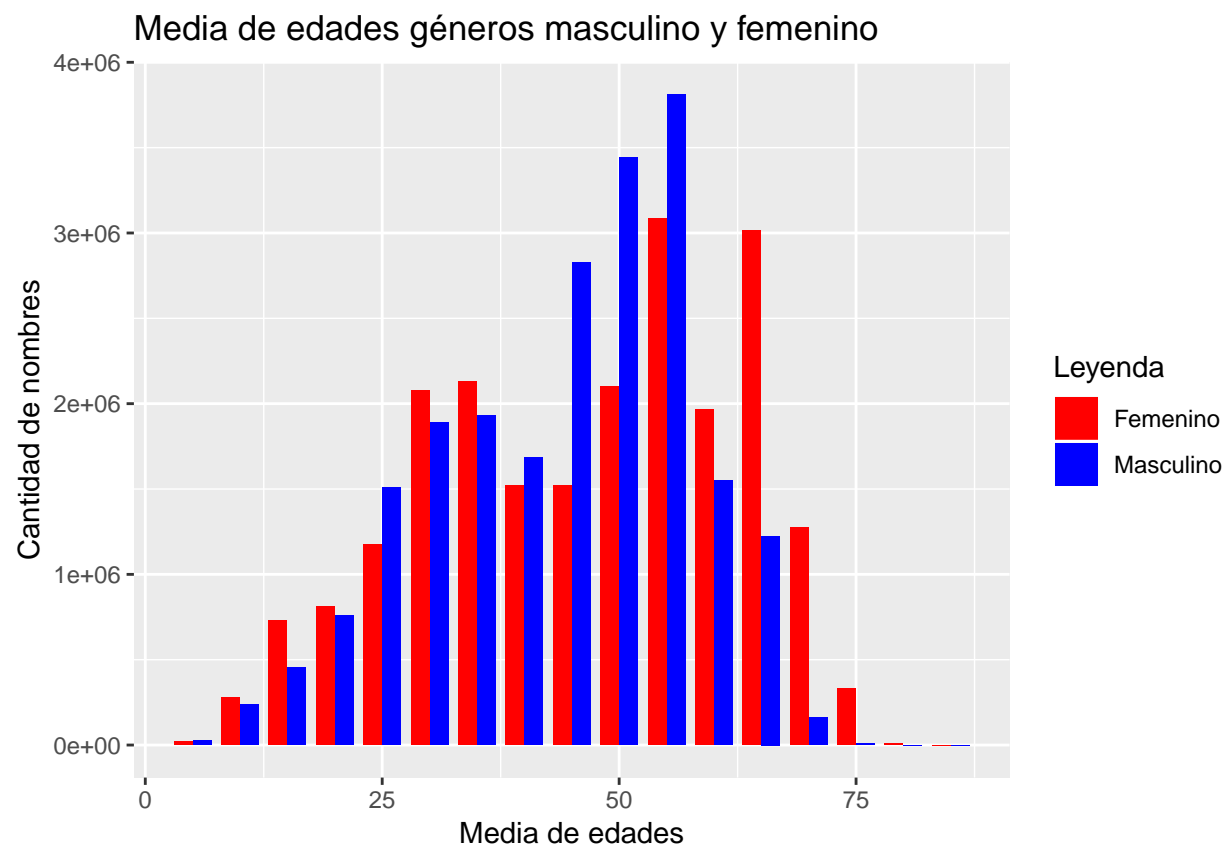
#### Gráfica de nombres femeninos



## Gráfica de nombres masculinos



## Gráfica combinada



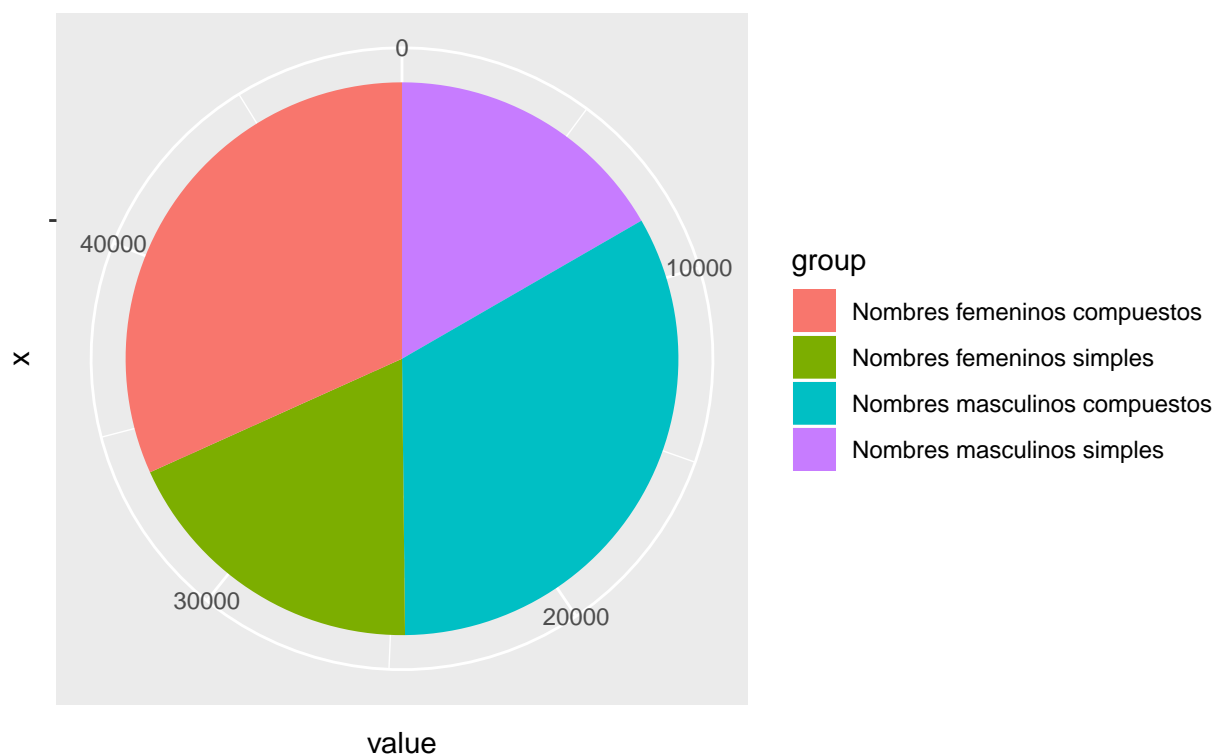
## Nombres compuestos

Los nombres compuestos son aquellos que están formados por dos o mas nombres, se evaluó el listado femenino y masculino de nombres llegando a la siguiente conclusión.

En el listado de nombres femeninos se ha determinado que la cantidad de nombres compuestos es de **15 659** de los cuales **15 359** son nombres compuestos por dos nombres y **300** son nombres compuestos por tres nombres

En el listado de nombres masculinos la cantidad de nombres compuestos son de **16 357**, **16 294** son nombres compuestos por dos nombres y **63** son nombres compuestos por tres nombres

## Total de nombres simples y compuestos.



## Total de nombres por letra.

```
total_letters <- data.frame(letter <- character(), total_female <- numeric(), total_male <- numeric())
for (letter in LETTERS) {
  total_letters <- rbind(total_letters, data.frame(
    letter = letter,
    total_female = count_frequency_letter(female_data, letter),
    total_male = count_frequency_letter(male_data, letter)))
}
```

letter	total_female	total_male
A	3019	3163
B	687	715
C	1697	1315
D	899	1232
E	1370	1301
F	629	1137
G	796	1039
H	369	684
I	924	840
J	1170	3084
K	481	350
L	1513	942
M	4192	2263
N	1056	596
O	282	397

letter	total_female	total_male
P	691	817
Q	38	35
R	1177	1407
S	1589	1276
T	447	388
U	50	69
V	581	558
W	112	312
X	179	161
Y	583	325
Z	224	178