

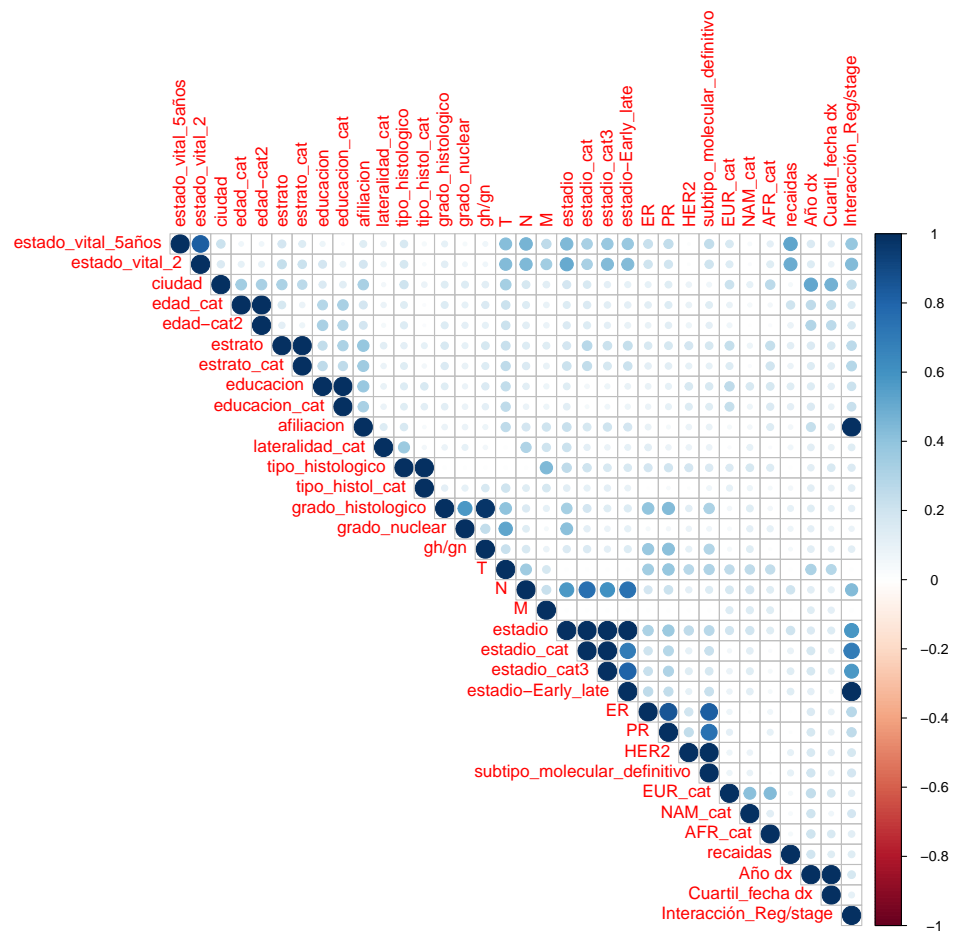
# **Informe Consultoria**

Luis Hernández y Juan Carvajal

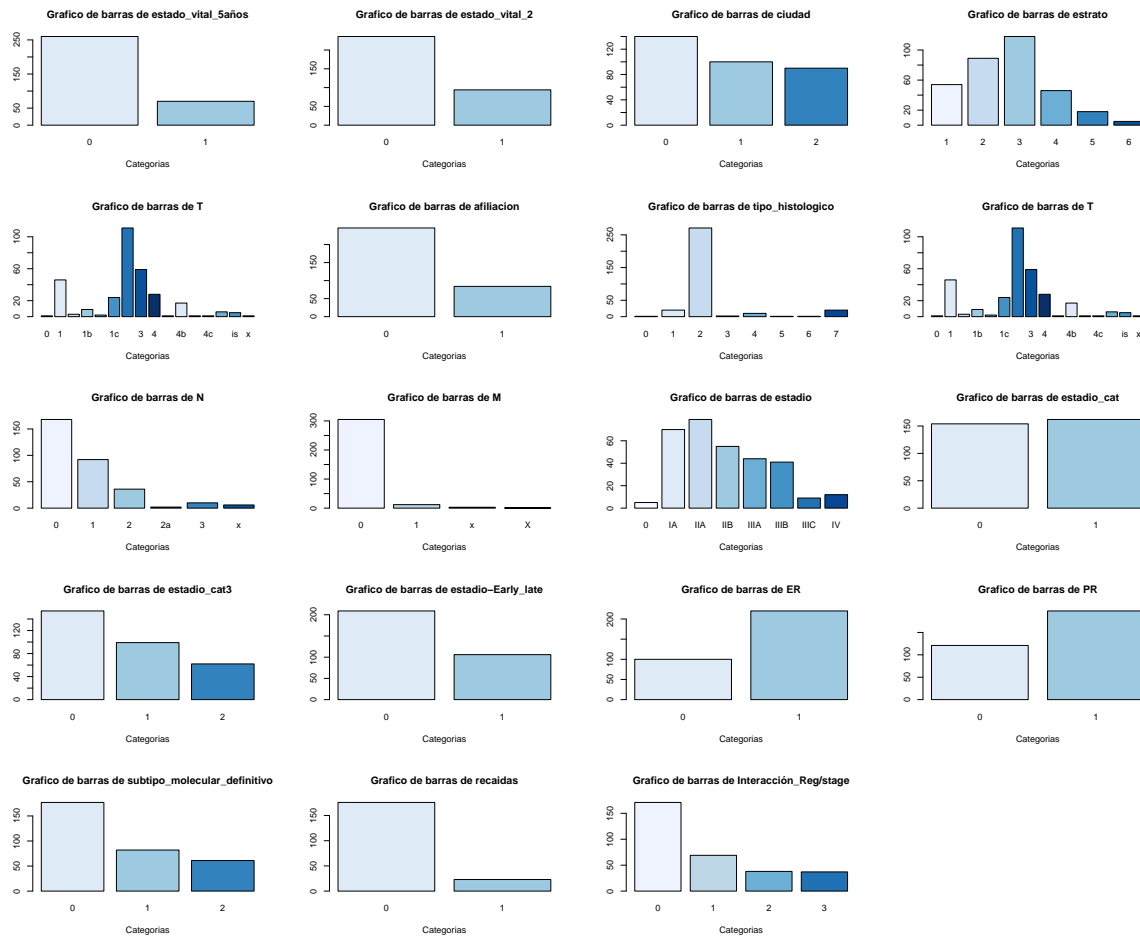


# Análisis Descriptivo

## Correlaciones entre las variables



## Analisis descriptivos de variables individuales

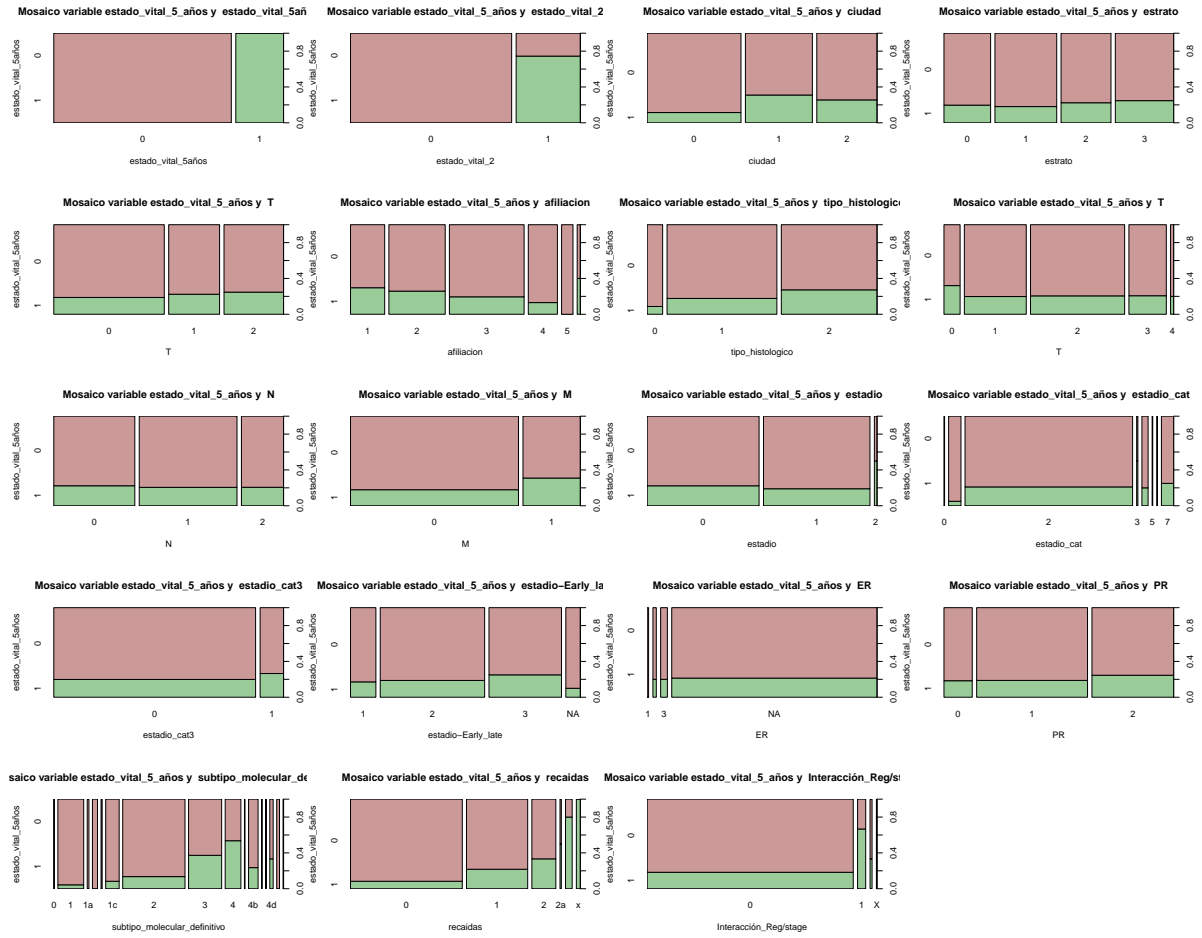


## Analisis descriptivo de variables en conjunto

### Análisis de Supervivencia

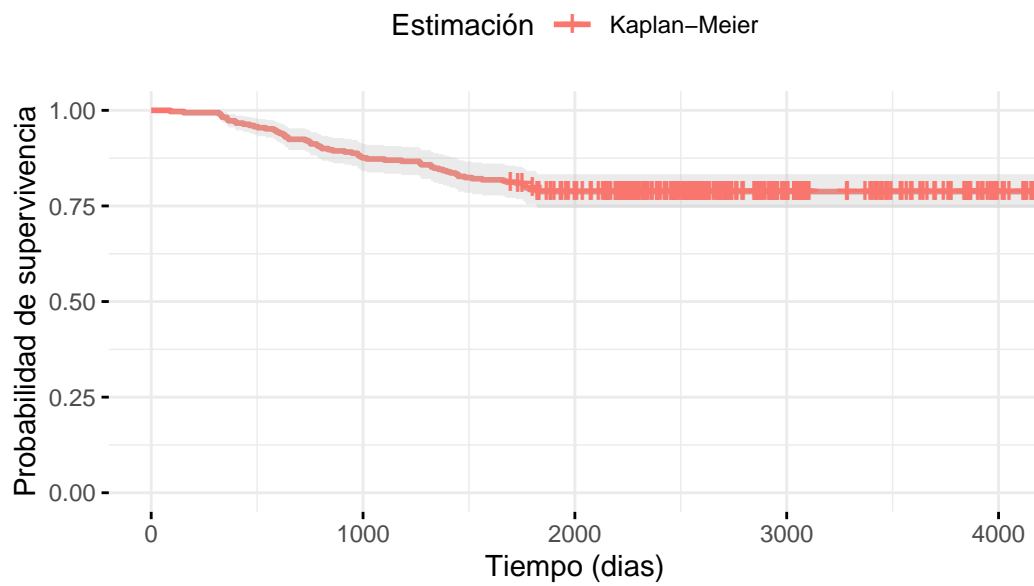
Para las curvas de supervivencia utilizaremos el estimador de Kaplan-Meier.

### Estimador de Kaplan-Meier.

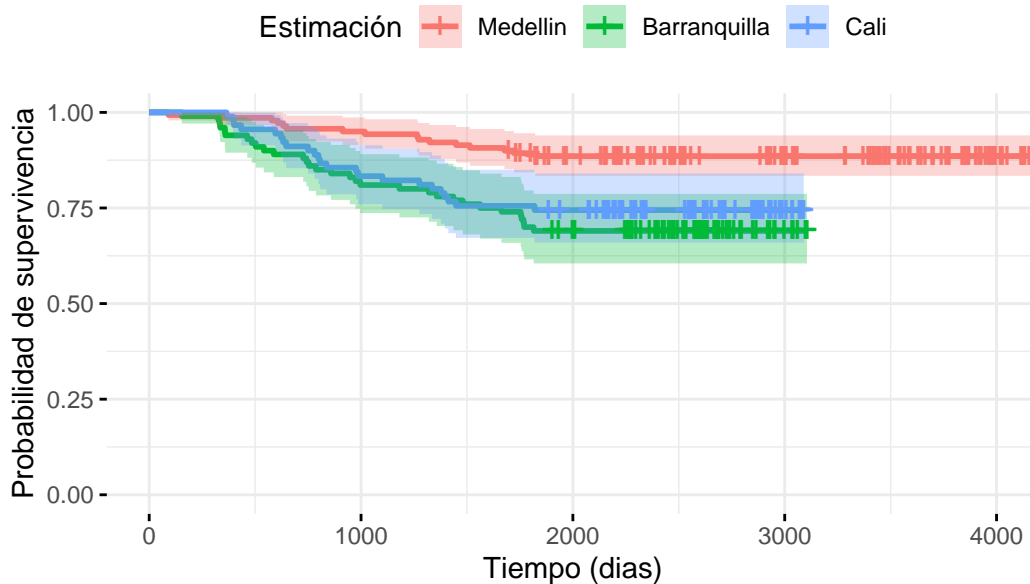


estado_vital_5anos	tiempo_supervivencia_dias
1	1271
0	4237
0	4323
0	4286
0	4293
1	1677

### Curva de Supervivencia



## Curva de Supervivencia por Ciudad



Call:

```
survdiffformula = Surv(tiempo_supervivencia_dias, estado_vital_5anos) ~
ciudad, data = bd, rho = 0)
```

	N	Observed	Expected	$(O-E)^2/E$	$(O-E)^2/V$
ciudad=0	140	16	31.5	7.62	13.86
ciudad=1	100	31	20.0	6.07	8.50
ciudad=2	90	23	18.5	1.08	1.47

Chisq= 14.8 on 2 degrees of freedom, p= 6e-04

## Modelo de Riesgos Proporcionales de Cox

En las situaciones experimentales en las que deseamos estudiar la supervivencia de un conjunto de sujetos en función de un conjunto  $X = (X_1, \dots, X_p)$  de variables predictoras, es decir, variables que pueden afectar o caracterizar su supervivencia, es necesario establecer modelos estadísticos capaces de analizar dichas relaciones. La construcción de este tipo de modelos que depende del tiempo y de las predictoras se hace a través del análisis de la función hazard asociada  $h(t; X)$ .

El modelo más habitual en esta situación es el **modelo hazard proporcional** que separa en dos componentes la función hazard, una correspondiente al tiempo de supervivencia y otra a

las variables predictoras. La finalidad de este modelo es para identificar factores que influyen en la supervivencia.

A manera de ejemplo se ajustara un modelo con algunas variables, las variables a considerar al modelo final, y se tranda en cuenta tambien el criterio de Akaike.

Call:

```
coxph(formula = Surv(tiempo_supervivencia_dias, estado_vital_5anos ==
  1) ~ ciudad + edad_cat + estrato_cat + educacion_cat + afiliacion,
  data = bd)
```

n= 330, number of events= 70

	coef	exp(coef)	se(coef)	z	Pr(> z )
ciudad1	1.0247	2.7862	0.3470	2.953	0.003147 **
ciudad2	1.2044	3.3347	0.3567	3.377	0.000733 ***
edad_cat1	-0.2012	0.8177	0.3809	-0.528	0.597242
edad_cat2	-0.3109	0.7328	0.4132	-0.753	0.451714
edad_cat3	-0.1475	0.8629	0.4144	-0.356	0.721891
estrato_cat1	0.8648	2.3746	0.7397	1.169	0.242356
estrato_cat2	1.3519	3.8648	0.7696	1.757	0.078980 .
educacion_cat1	0.1487	1.1603	0.3092	0.481	0.630482
educacion_cat2	0.5548	1.7415	0.4227	1.313	0.189335
afiliacion1	0.4964	1.6428	0.2795	1.776	0.075733 .

---

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ciudad1	2.7862	0.3589	1.4114	5.500
ciudad2	3.3347	0.2999	1.6575	6.709
edad_cat1	0.8177	1.2229	0.3876	1.725
edad_cat2	0.7328	1.3647	0.3261	1.647
edad_cat3	0.8629	1.1589	0.3830	1.944
estrato_cat1	2.3746	0.4211	0.5571	10.121
estrato_cat2	3.8648	0.2587	0.8551	17.466
educacion_cat1	1.1603	0.8618	0.6330	2.127
educacion_cat2	1.7415	0.5742	0.7606	3.987
afiliacion1	1.6428	0.6087	0.9499	2.841

Concordance= 0.678 (se = 0.031 )

Likelihood ratio test= 27.26 on 10 df, p=0.002

Wald test = 24.51 on 10 df, p=0.006

Score (logrank) test = 26.51 on 10 df, p=0.003