

Preterm birth data set

Degree of preterm birth:

mild	32-36 weeks' gestation
moderate	28-31
extreme	< 28

Many causes, most unknown. But three key categories:

iatrogenic (medically induced. E.g. to save mom or baby)

PPROM (water breaks – often due to infection)

spontaneous (stress, lifestyle, unknown)

Are these “indications” independent of degree?

Recall what the HairColor data set looked like:

An example in chapter 2 tabulated hair and eye color:

```
HairColor = HairEyeColor[, , 2]
```

```
HairColor
```

Hair	Eye			
	Brown	Blue	Hazel	Green
Black	36	9	5	2
Brown	66	34	29	14
Red	16	7	7	7
Blond	4	64	5	8

Are these two variables independent of each other?

Code for HairColor analysis

```
library("vcd"); library("ade4"); library("factoextra")  
  
mosaicplot(HairColor,shade=TRUE,las=1,type="pearson",  
  cex.axis=0.7,main="")  
HC=as.data.frame.matrix(HairColor)  
coaHC=dudi.coa(HC, scannf=FALSE,nf=2) # package ade4  
fviz_ca_biplot(coaHC,repel=TRUE,col.col="brown",  
  col.row="purple") + ggtitle("") + ylim(c(-0.5,0.5))
```

Try this on preterms.txt.

Which, if any, indications for and degrees of preterm birth are related (and in which direction?)