

## Anexo VII. PCA Por décadas

A continuación se muestra el proceso llevado a cabo para obtener los diferentes PCAs por décadas

### Carga del conjunto de datos

```
load("C:/DANIEL/UNIVERSIDAD/CD2/PROY II/Proyecto_R/Datos/ncdata.RData")
load("C:/DANIEL/UNIVERSIDAD/CD2/PROY II/Proyecto_R/Datos/desc_data.RData")
```

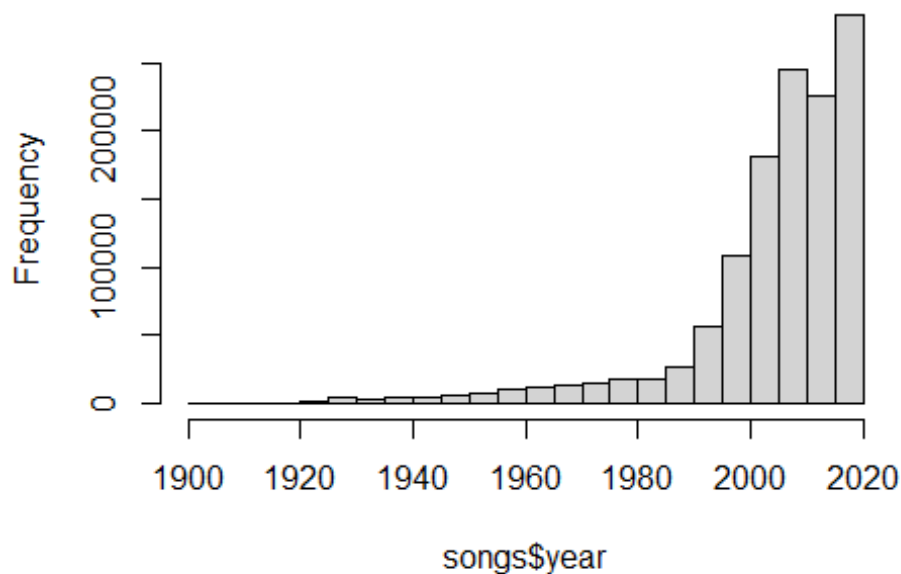
### Obtención de los intervalos de años

```
songs = ncdata[ncdata$year != 0, -1]
summary(songs$year)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	1900	2001	2008	2005	2015	2020

```
hist(songs$year, main = "Histograma de frecuencia de años")
```

**Histograma de frecuencia de años**



```
categories = c("1900-1909", "1910-1919", "1920-1929", "1930-1939",  
"1940-1949",
```

```

"1950-1959", "1960-1969", "1970-1979", "1980-1989",
"1990-1999",
      "2000-2009", "2010-2020")
songs$year_c = cut(songs$year, breaks = seq(1900, 2020, 10), labels =
categories, include.lowest = T)

```

Ahora que tenemos los años con intervalos, vamos a ver como queda el histograma de años (que ahora es un diagrama de barras):

```

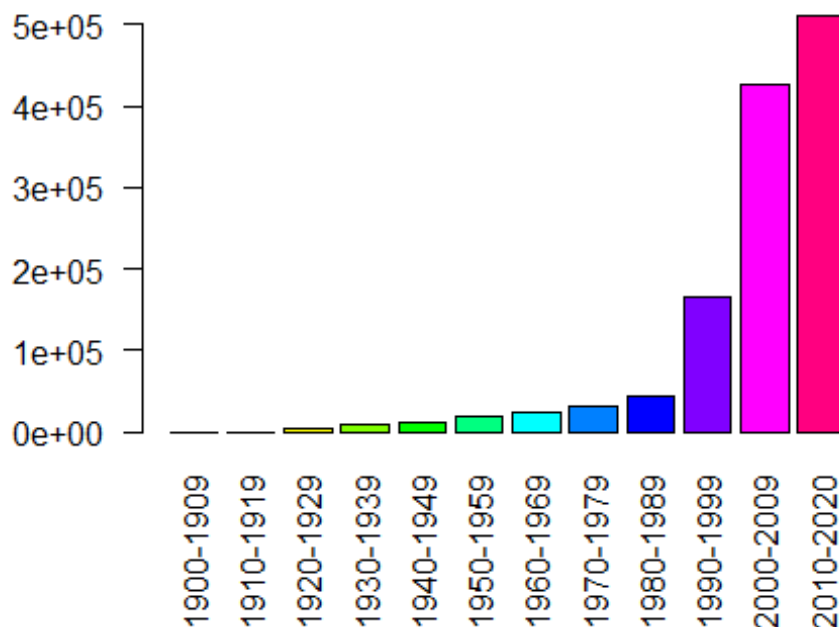
summary(songs$year_c)

## 1900-1909 1910-1919 1920-1929 1930-1939 1940-1949 1950-1959 1960-1969
1970-1979
##          58          63        4901        7949        10255        18051        24157
32299
## 1980-1989 1990-1999 2000-2009 2010-2020
##        44746        164287        427144        510500

barplot(table(songs$year_c), main = "Histograma de frecuencia de años (por
intervalos)", las = 2, col = rainbow(12))

```

## Histograma de frecuencia de años (por intervalos)



Vamos a empezar ahora a realizar los distintos PCAs.

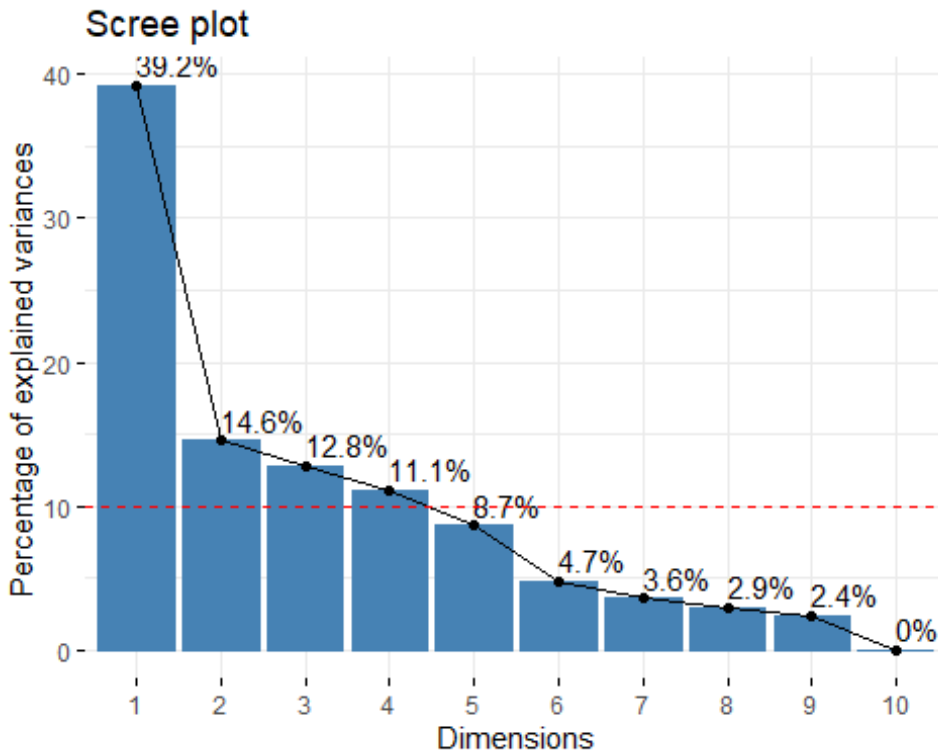
## PCA 1900-1909

```

songs1900_1909 = songs[songs$year_c ==
"1900-1909", c(2,3,4,5,7,9,10,11,15,16,17)]

```

```
songs1900_1909[, -4] = scale(songs1900_1909[, -4])
pca1900_1909 = PCA(songs1900_1909, graph = F, scale.unit = F, ncp = 10,
quali.sup = c(9))
eig1900_1909 = get_eig(pca1900_1909)
fviz_eig(pca1900_1909, addlabels = TRUE) +
geom_hline(yintercept=100*(1/nrow(eig1900_1909)), linetype=2, color="red")
```



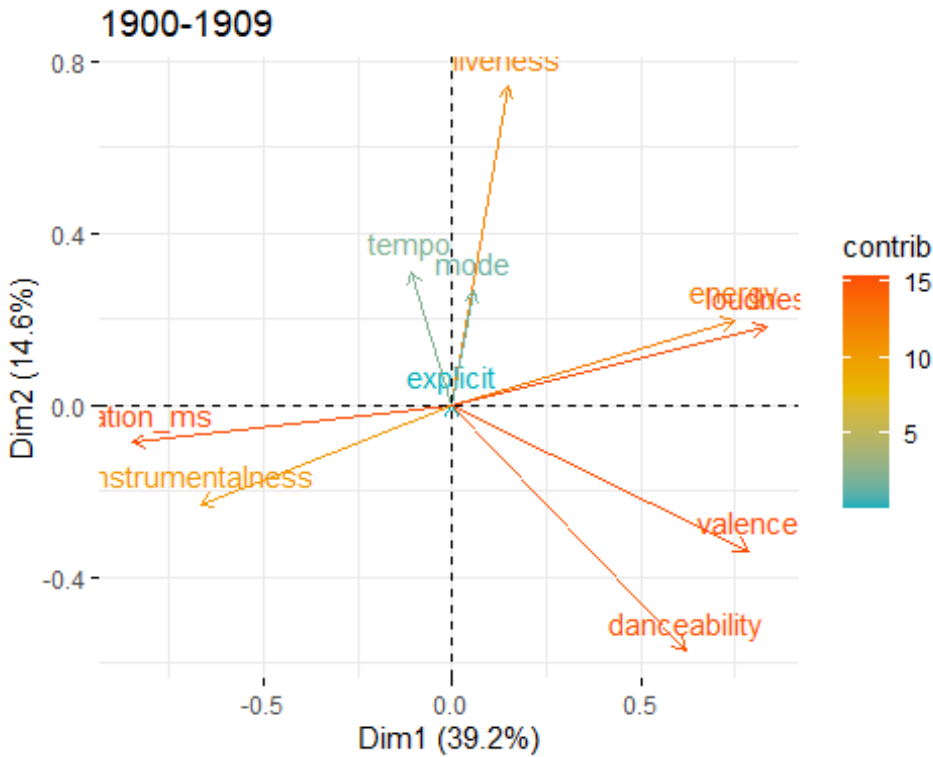
```
eig1900_1909
```

##	eigenvalue	variance.percent	cumulative.variance.percent
## Dim.1	3.465110e+00	3.917668e+01	39.17668
## Dim.2	1.288071e+00	1.456299e+01	53.73967
## Dim.3	1.131080e+00	1.278804e+01	66.52770
## Dim.4	9.859519e-01	1.114721e+01	77.67492
## Dim.5	7.654828e-01	8.654582e+00	86.32950
## Dim.6	4.188851e-01	4.735932e+00	91.06543
## Dim.7	3.172347e-01	3.586669e+00	94.65210
## Dim.8	2.608945e-01	2.949685e+00	97.60179
## Dim.9	2.121178e-01	2.398213e+00	100.00000
## Dim.10	2.665192e-33	3.013277e-32	100.00000

```
#Escogemos 4
```

```
pca1900_1909 = PCA(songs1900_1909, graph = F, scale.unit = F, ncp = 4,
quali.sup = c(9))
```

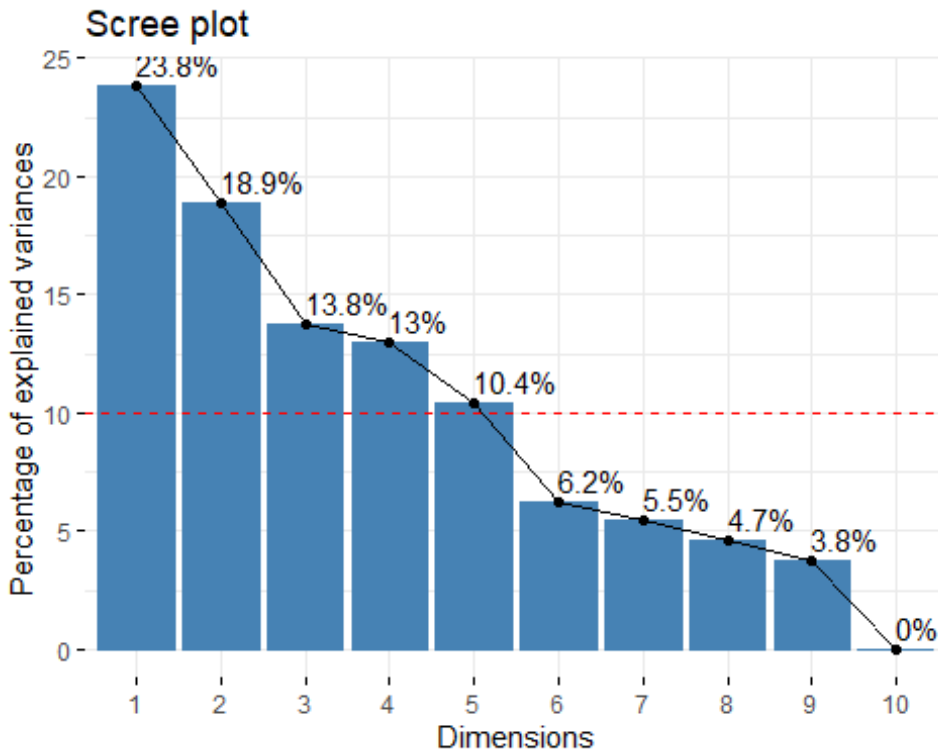
```
fviz_pca_var(pca1900_1909, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1900-1909")
```



```
#fviz_pca_biplot(pca1900_1909)
```

## PCA 1910-1919

```
songs1910_1919 = songs[songs$year_c ==
"1910-1919", c(2,3,4,5,7,9,10,11,15,16,17)]
songs1910_1919[, -4] = scale(songs1910_1919[, -4])
pca1910_1919 = PCA(songs1910_1919, graph = F, scale.unit = F, ncp = 10,
quali.sup = c(9))
eig1910_1919 = get_eig(pca1910_1919)
fviz_eig(pca1910_1919, addlabels = TRUE) +
geom_hline(yintercept=100*(1/nrow(eig1910_1919)), linetype=2, color="red")
```



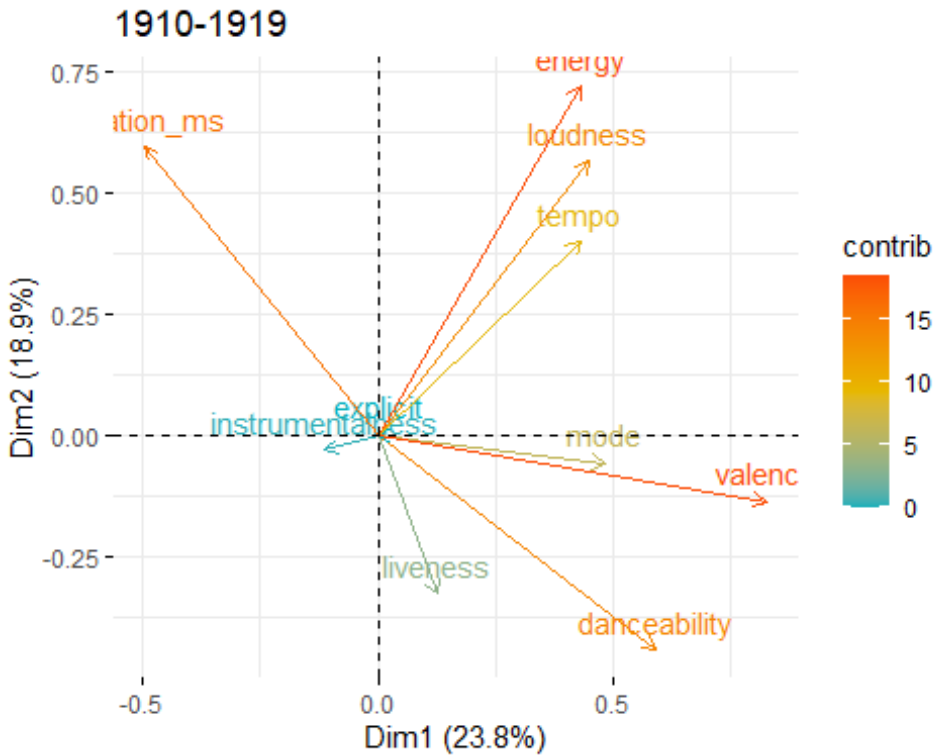
```
eig1910_1919
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1  2.111102e+00      2.383502e+01                23.83502
## Dim.2  1.674542e+00      1.890612e+01                42.74114
## Dim.3  1.217896e+00      1.375044e+01                56.49159
## Dim.4  1.149834e+00      1.298199e+01                69.47358
## Dim.5  9.213635e-01      1.040249e+01                79.87607
## Dim.6  5.527026e-01      6.240191e+00                86.11626
## Dim.7  4.841741e-01      5.466481e+00                91.58274
## Dim.8  4.122529e-01      4.654468e+00                96.23721
## Dim.9  3.332757e-01      3.762790e+00               100.00000
## Dim.10 6.898889e-35      7.789069e-34               100.00000
```

```
#Escogemos 5
```

```
pca1910_1919 = PCA(songs1910_1919, graph = F, scale.unit = F, ncp = 5,
quali.sup = c(9))
```

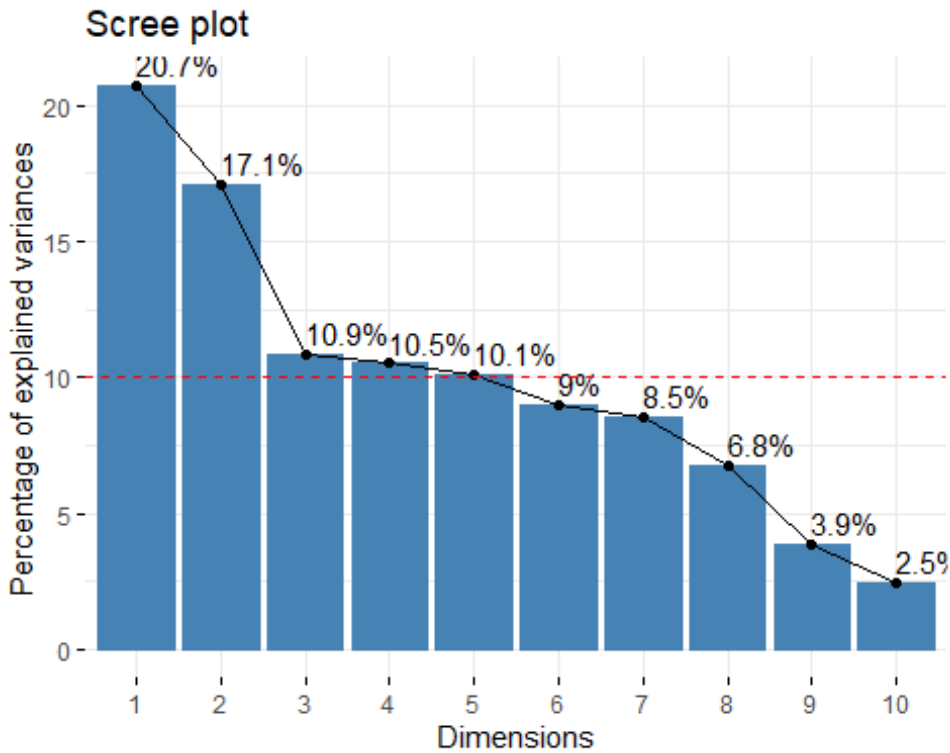
```
fviz_pca_var(pca1910_1919, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1910-1919")
```



```
#fviz_pca_biplot(pca1910_1919)
```

## PCA 1920-1929

```
songs1920_1929 = songs[songs$year_c ==
"1920-1929",c(2,3,4,5,7,9,10,11,15,16,17)]
songs1920_1929 = scale(songs1920_1929)
pca1920_1929 = PCA(songs1920_1929, graph = F, scale.unit = F, ncp = 10,
quali.sup = c(9))
eig1920_1929 = get_eig(pca1920_1929)
fviz_eig(pca1920_1929, addlabels = TRUE) +
geom_hline(yintercept=100*(1/nrow(eig1920_1929)), linetype=2, color="red")
```



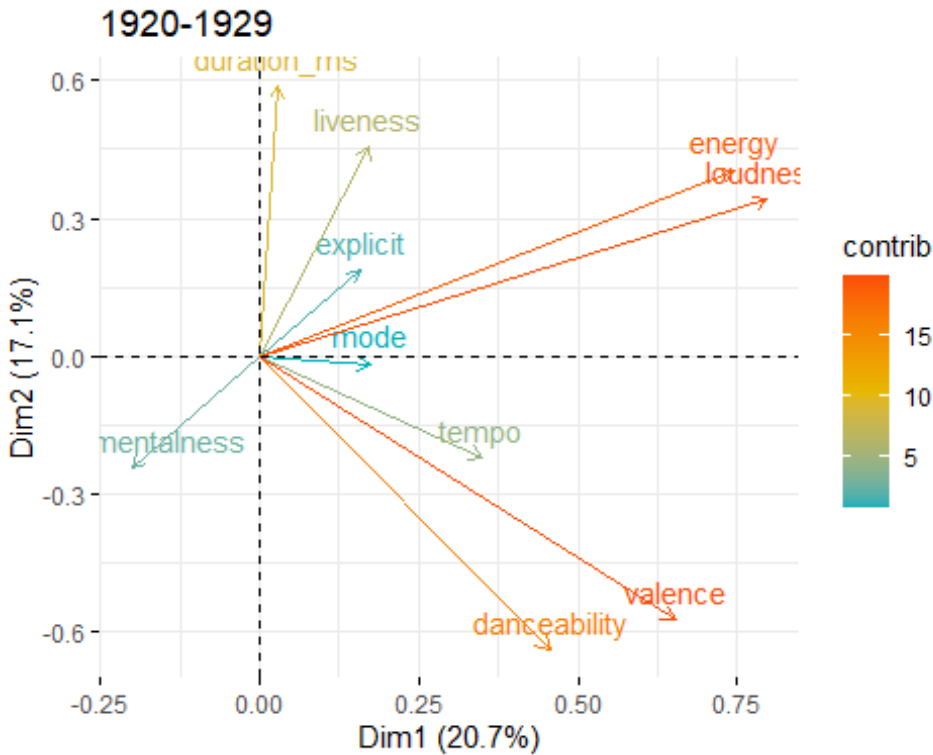
```
eig1920_1929
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1    2.0721526         20.725755          20.72576
## Dim.2    1.7087405         17.090892          37.81665
## Dim.3    1.0867289         10.869506          48.68615
## Dim.4    1.0524787         10.526935          59.21309
## Dim.5    1.0086009         10.088067          69.30116
## Dim.6    0.9012644          9.014483          78.31564
## Dim.7    0.8532538          8.534280          86.84992
## Dim.8    0.6781253          6.782637          93.63256
## Dim.9    0.3880609          3.881401          97.51396
## Dim.10   0.2485537          2.486044         100.00000
```

```
#Escogemos 6
```

```
pca1920_1929 = PCA(songs1920_1929, graph = F, scale.unit = F, ncp = 6,
quali.sup = c(9))
```

```
fviz_pca_var(pca1920_1929, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1920-1929")
```

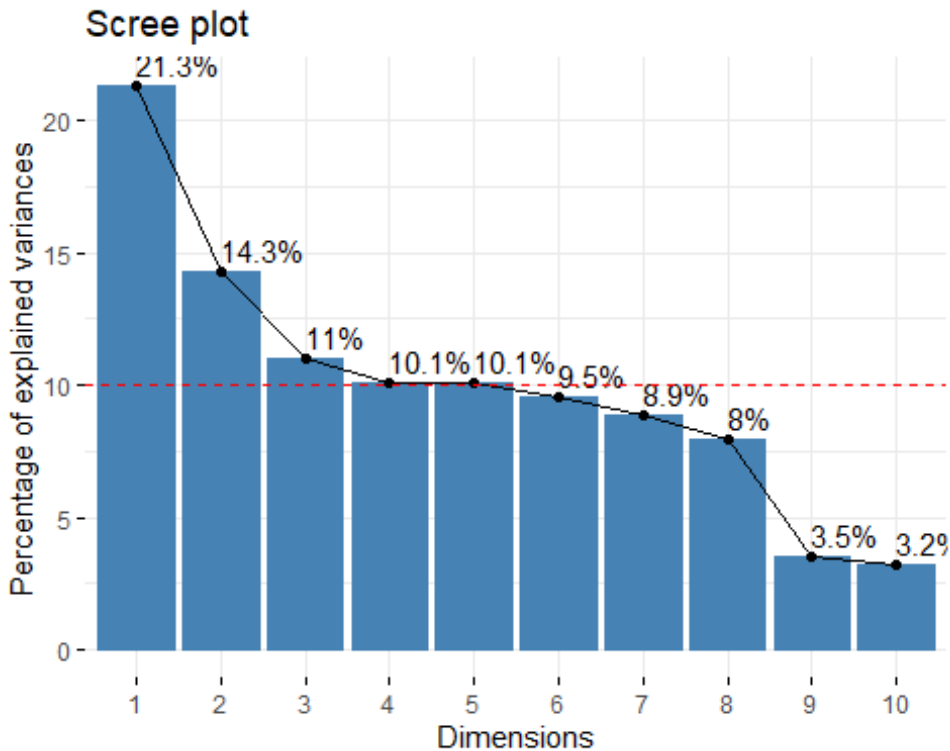


```
#fviz_pca_biplot(pca1920_1929)
```

## PCA 1930-1939

```
songs1930_1939 = songs[songs$year_c ==
"1930-1939",c(2,3,4,5,7,9,10,11,15,16,17)]
songs1930_1939 = scale(songs1930_1939)
pca1930_1939 = PCA(songs1930_1939, graph = F, scale.unit = F, ncp = 10,
quali.sup = c(9))
eig1930_1939 = get_eig(pca1930_1939)
fviz_eig(pca1930_1939, addlabels = TRUE) +
geom_hline(yintercept=100*(1/nrow(eig1930_1939)), linetype=2, color="red")
```





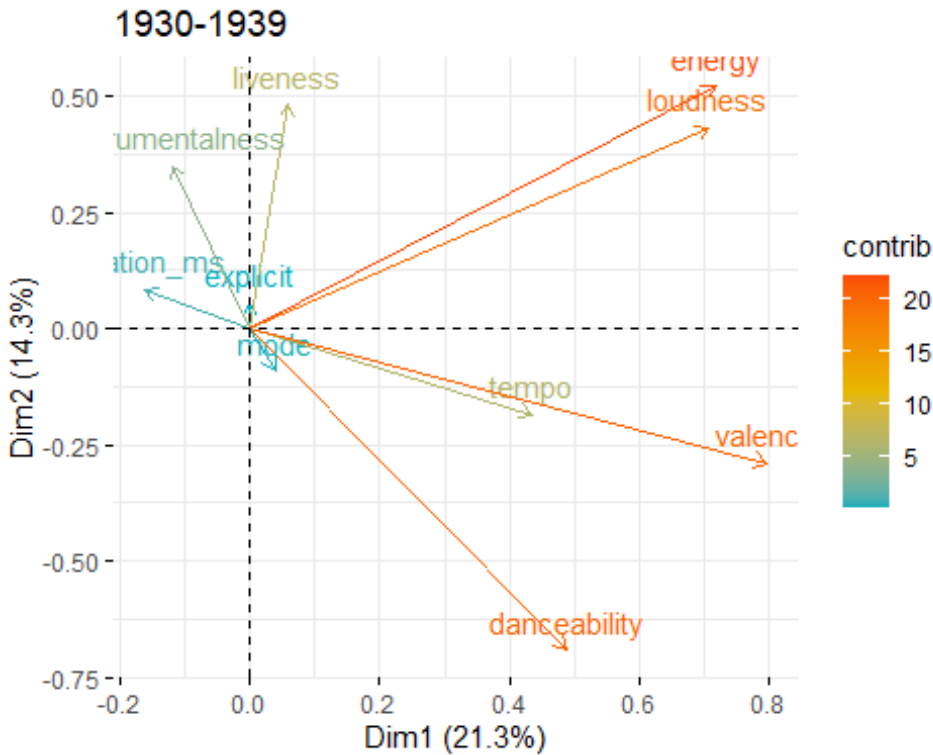
```
eig1930_1939
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1    2.1314266         21.316948          21.31695
## Dim.2    1.4293589         14.295388          35.61234
## Dim.3    1.1011847         11.013233          46.62557
## Dim.4    1.0110158         10.111430          56.73700
## Dim.5    1.0070023         10.071290          66.80829
## Dim.6    0.9545863          9.547064          76.35535
## Dim.7    0.8894296          8.895415          85.25077
## Dim.8    0.7973086          7.974089          93.22486
## Dim.9    0.3542231          3.542677          96.76753
## Dim.10   0.3232060          3.232467         100.00000
```

```
#Escogemos 6
```

```
pca1930_1939 = PCA(songs1930_1939, graph = F, scale.unit = F, ncp = 6,
quali.sup = c(9))
```

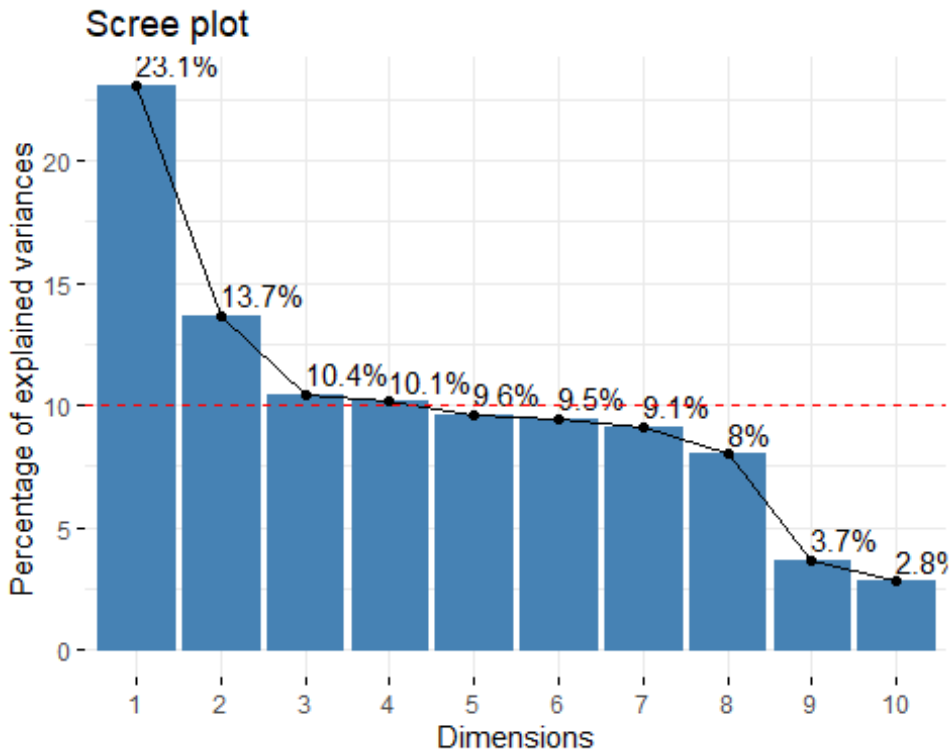
```
fviz_pca_var(pca1930_1939, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1930-1939")
```



```
#fviz_pca_biplot(pca1930_1939)
```

## PCA 1940-1949

```
songs1940_1949 = songs[songs$year_c ==  
"1940-1949",c(2,3,4,5,7,9,10,11,15,16,17)]  
songs1940_1949 = scale(songs1940_1949)  
pca1940_1949 = PCA(songs1940_1949, graph = F, scale.unit = F, ncp = 10,  
quali.sup = c(9))  
eig1940_1949 = get_eig(pca1940_1949)  
fviz_eig(pca1940_1949, addlabels = TRUE) +  
geom_hline(yintercept=100*(1/nrow(eig1940_1949)), linetype=2, color="red")
```



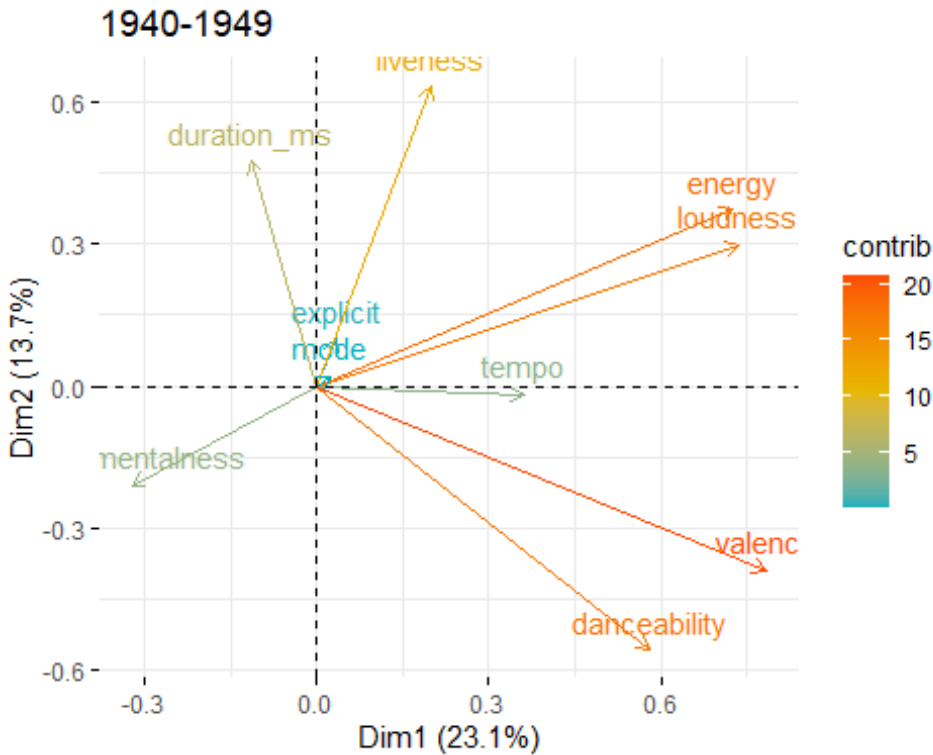
```
eig1940_1949
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1    2.3075087         23.077337          23.07734
## Dim.2    1.3686018         13.687353          36.76469
## Dim.3    1.0427107         10.428124          47.19281
## Dim.4    1.0142266         10.143255          57.33607
## Dim.5    0.9632285          9.633224          66.96929
## Dim.6    0.9453189          9.454111          76.42340
## Dim.7    0.9135395          9.136286          85.55969
## Dim.8    0.7992680          7.993459          93.55315
## Dim.9    0.3672356          3.672714          97.22586
## Dim.10   0.2773865          2.774136         100.00000
```

```
#Escogemos 6
```

```
pca1940_1949 = PCA(songs1940_1949, graph = F, scale.unit = F, ncp = 6,
quali.sup = c(9))
```

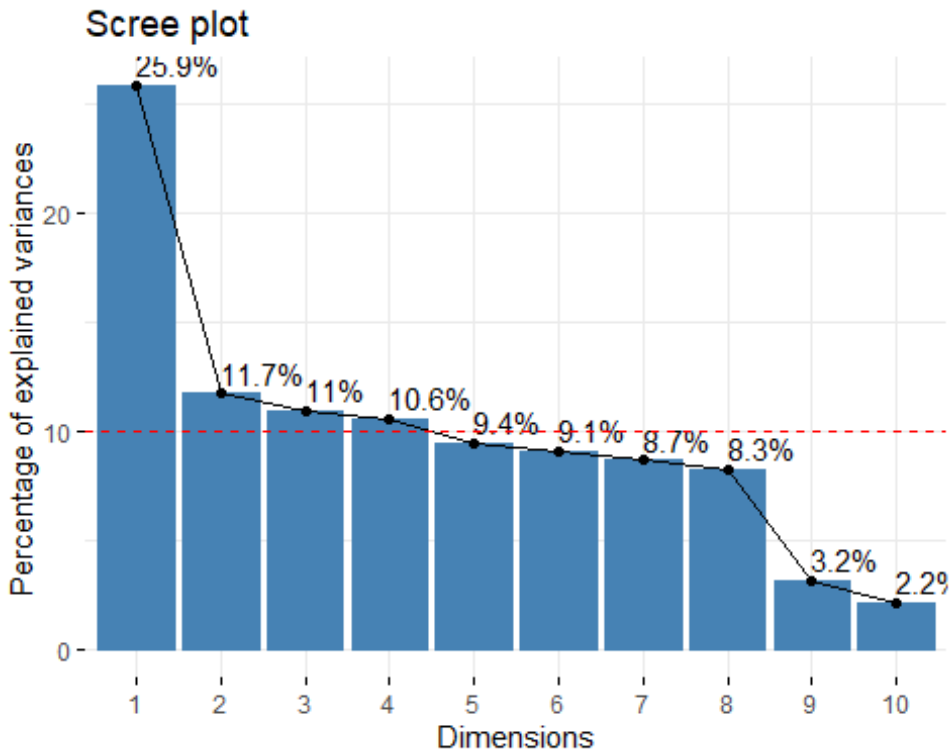
```
fviz_pca_var(pca1940_1949, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1940-1949")
```



```
#fviz_pca_biplot(pca1940_1949)
```

## PCA 1950-1959

```
songs1950_1959 = songs[songs$year_c ==  
"1950-1959",c(2,3,4,5,7,9,10,11,15,16,17)]  
songs1950_1959 = scale(songs1950_1959)  
pca1950_1959 = PCA(songs1950_1959, graph = F, scale.unit = F, ncp = 10,  
quali.sup = c(9))  
eig1950_1959 = get_eig(pca1950_1959)  
fviz_eig(pca1950_1959, addlabels = TRUE) +  
geom_hline(yintercept=100*(1/nrow(eig1950_1959)), linetype=2, color="red")
```



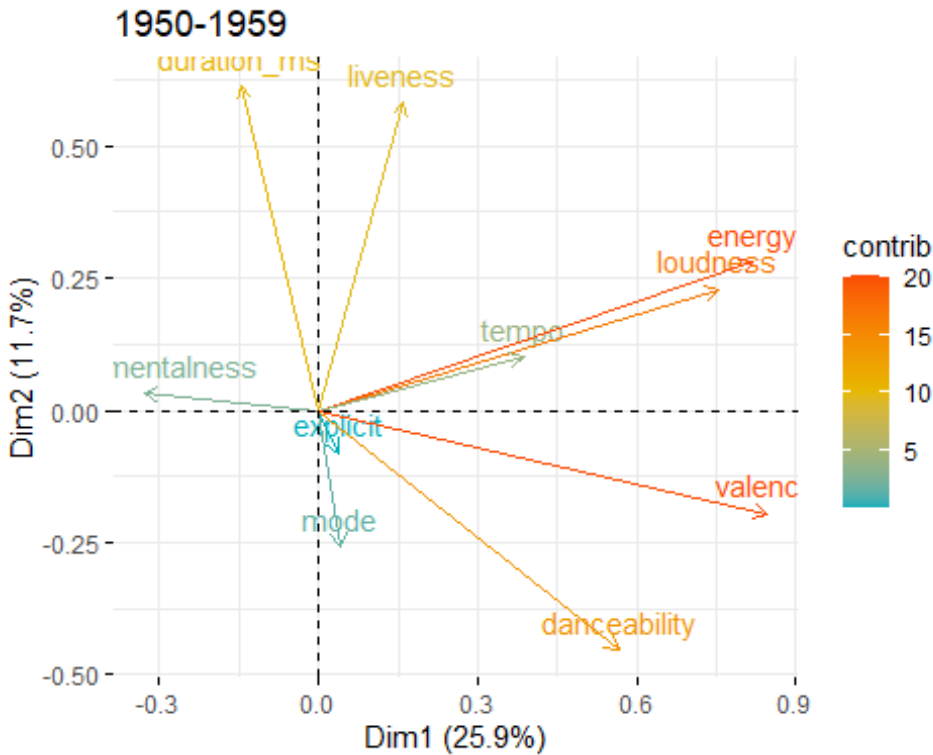
```
eig1950_1959
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1    2.5871604         25.873037         25.87304
## Dim.2    1.1740501         11.741151         37.61419
## Dim.3    1.0968762         10.969369         48.58356
## Dim.4    1.0608565         10.609153         59.19271
## Dim.5    0.9447967          9.448491         68.64120
## Dim.6    0.9090150          9.090653         77.73185
## Dim.7    0.8686421          8.686902         86.41876
## Dim.8    0.8260268          8.260726         94.67948
## Dim.9    0.3158009          3.158184         97.83767
## Dim.10   0.2162214          2.162333        100.00000
```

```
#Escogemos 6
```

```
pca1950_1959 = PCA(songs1950_1959, graph = F, scale.unit = F, ncp = 6,
quali.sup = c(9))
```

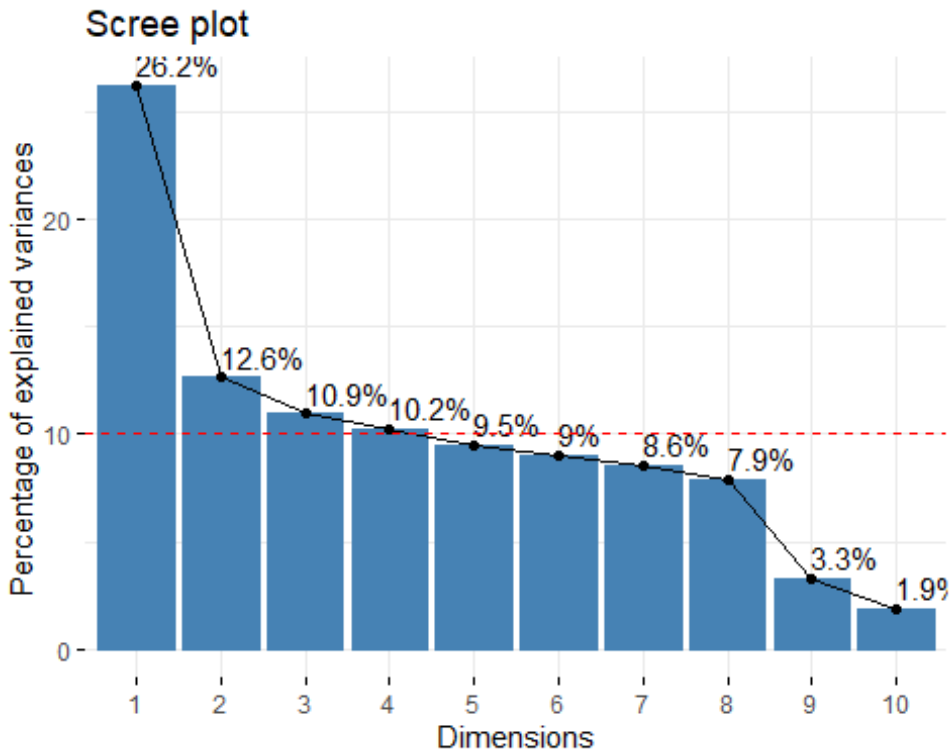
```
fviz_pca_var(pca1950_1959, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1950-1959")
```



```
#fviz_pca_biplot(pca1950_1959)
```

## PCA 1960-1969

```
songs1960_1969 = songs[songs$year_c ==
"1960-1969",c(2,3,4,5,7,9,10,11,15,16,17)]
songs1960_1969= scale(songs1960_1969)
pca1960_1969 = PCA(songs1960_1969, graph = F, scale.unit = F, ncp = 10,
quali.sup = c(9))
eig1960_1969 = get_eig(pca1960_1969)
fviz_eig(pca1960_1969, addlabels = TRUE) +
geom_hline(yintercept=100*(1/nrow(eig1960_1969)), linetype=2, color="red")
```



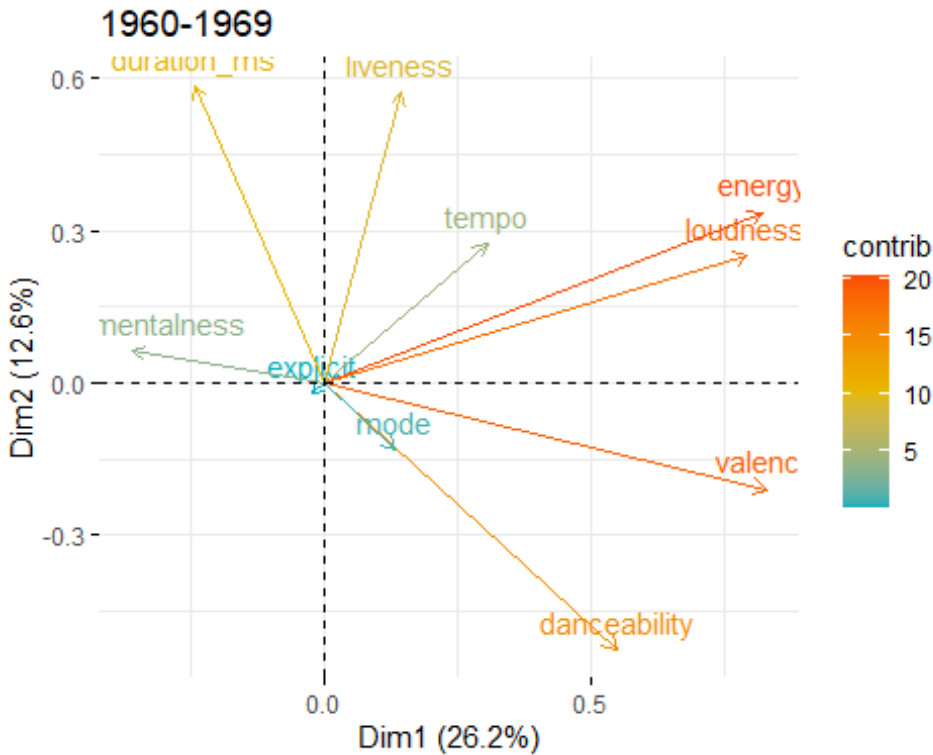
```
eig1960_1969
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1      2.6174068         26.175151          26.17515
## Dim.2      1.2637245         12.637768          38.81292
## Dim.3      1.0937238         10.937691          49.75061
## Dim.4      1.0218649         10.219072          59.96968
## Dim.5      0.9487303          9.487695          69.45738
## Dim.6      0.8983212          8.983584          78.44096
## Dim.7      0.8551817          8.552171          86.99313
## Dim.8      0.7858548          7.858874          94.85201
## Dim.9      0.3284722          3.284858          98.13686
## Dim.10     0.1863058          1.863135         100.00000
```

```
#Escogemos 6
```

```
pca1960_1969 = PCA(songs1960_1969, graph = F, scale.unit = F, ncp = 6,
quali.sup = c(9))
```

```
fviz_pca_var(pca1960_1969, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1960-1969")
```

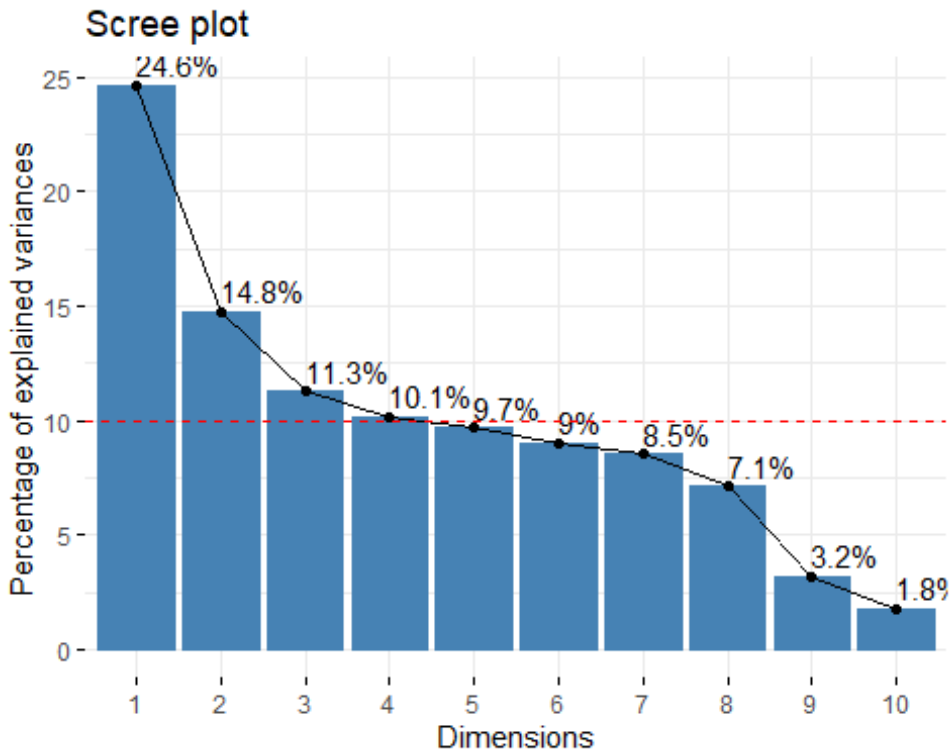


```
#fviz_pca_biplot(pca1960_1969)
```

## PCA 1970-1979

```
songs1970_1979 = songs[songs$year_c ==  
"1970-1979",c(2,3,4,5,7,9,10,11,15,16,17)]  
songs1970_1979 = scale(songs1970_1979)  
pca1970_1979 = PCA(songs1970_1979, graph = F, scale.unit = F, ncp = 10,  
quali.sup = c(9))  
eig1970_1979 = get_eig(pca1970_1979)  
fviz_eig(pca1970_1979, addlabels = TRUE) +  
geom_hline(yintercept=100*(1/nrow(eig1970_1979)), linetype=2, color="red")
```





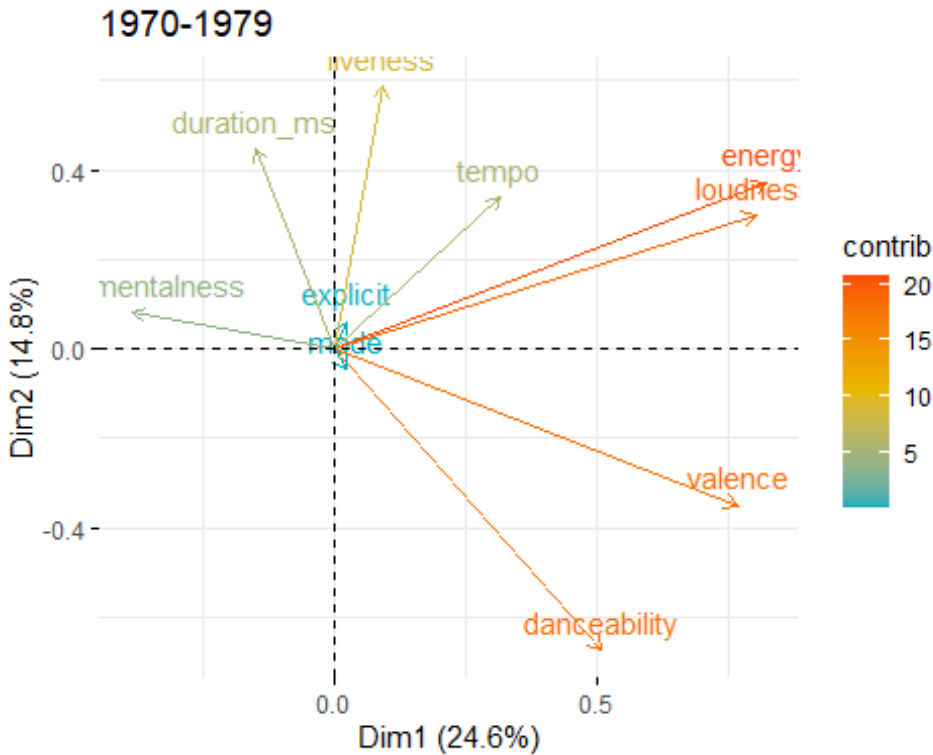
```
eig1970_1979
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1      2.4608868         24.609629             24.60963
## Dim.2      1.4770976         14.771433             39.38106
## Dim.3      1.1257258         11.257606             50.63867
## Dim.4      1.0140755         10.141069             60.77974
## Dim.5      0.9669428          9.669727             70.44947
## Dim.6      0.8976804          8.977082             79.42655
## Dim.7      0.8520427          8.520691             87.94724
## Dim.8      0.7116089          7.116309             95.06355
## Dim.9      0.3168069          3.168167             98.23172
## Dim.10     0.1768230          1.768285             100.00000
```

```
#Escogemos 6
```

```
pca1970_1979 = PCA(songs1970_1979, graph = F, scale.unit = F, ncp = 6,
quali.sup = c(9))
```

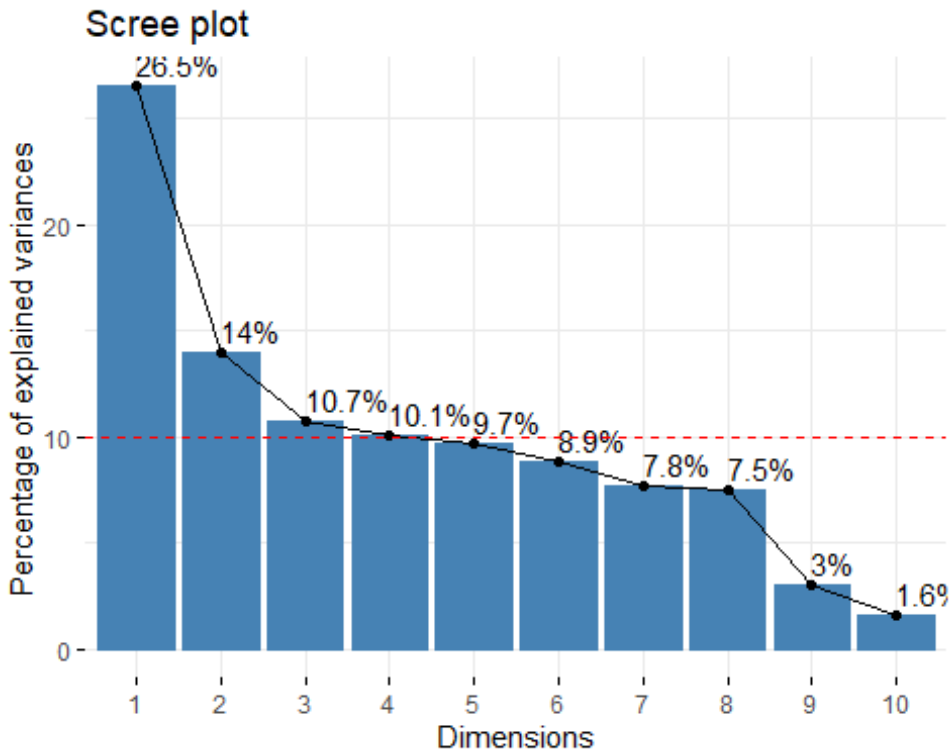
```
fviz_pca_var(pca1970_1979, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1970-1979")
```



```
#fviz_pca_biplot(pca1970_1979)
```

## PCA 1980-1989

```
songs1980_1989 = songs[songs$year_c ==  
"1980-1989",c(2,3,4,5,7,9,10,11,15,16,17)]  
songs1980_1989 = scale(songs1980_1989)  
pca1980_1989 = PCA(songs1980_1989, graph = F, scale.unit = F, ncp = 10,  
quali.sup = c(9))  
eig1980_1989 = get_eig(pca1980_1989)  
fviz_eig(pca1980_1989, addlabels = TRUE) +  
geom_hline(yintercept=100*(1/nrow(eig1980_1989)), linetype=2, color="red")
```



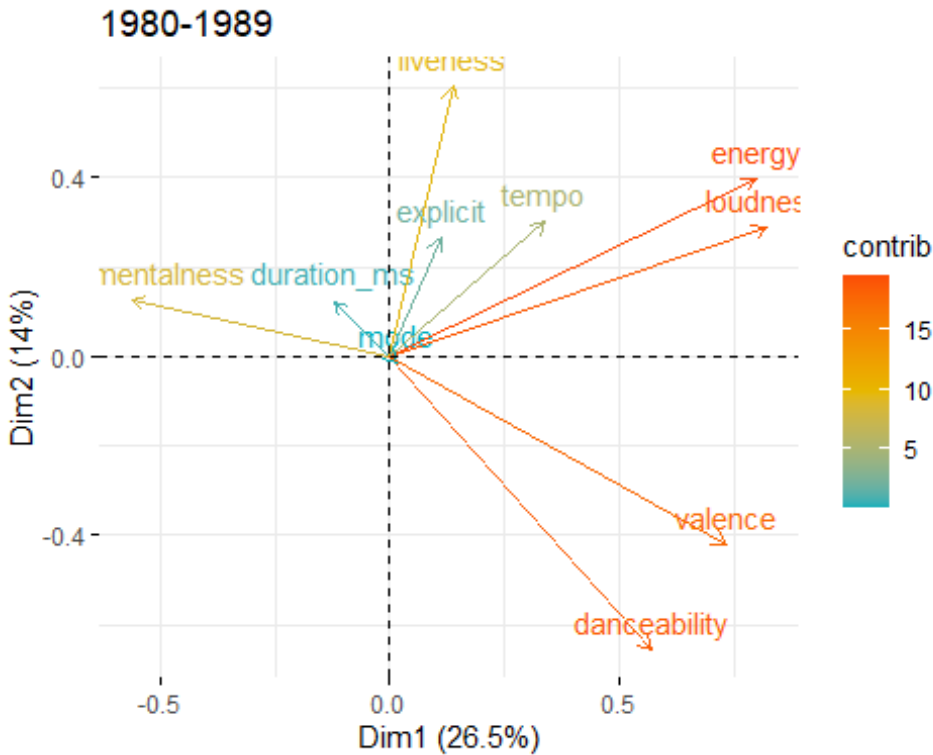
```
eig1980_1989
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1    2.6543955         26.544548             26.54455
## Dim.2    1.4032985         14.033299             40.57785
## Dim.3    1.0738888         10.739128             51.31697
## Dim.4    1.0085735         10.085960             61.40294
## Dim.5    0.9743879          9.744097             71.14703
## Dim.6    0.8893542          8.893741             80.04077
## Dim.7    0.7756101          7.756274             87.79705
## Dim.8    0.7507458          7.507626             95.30467
## Dim.9    0.3047193          3.047261             98.35193
## Dim.10   0.1648029          1.648066            100.00000
```

```
#Escogemos 6
```

```
pca1980_1989 = PCA(songs1980_1989, graph = F, scale.unit = F, ncp = 6,
quali.sup = c(9))
```

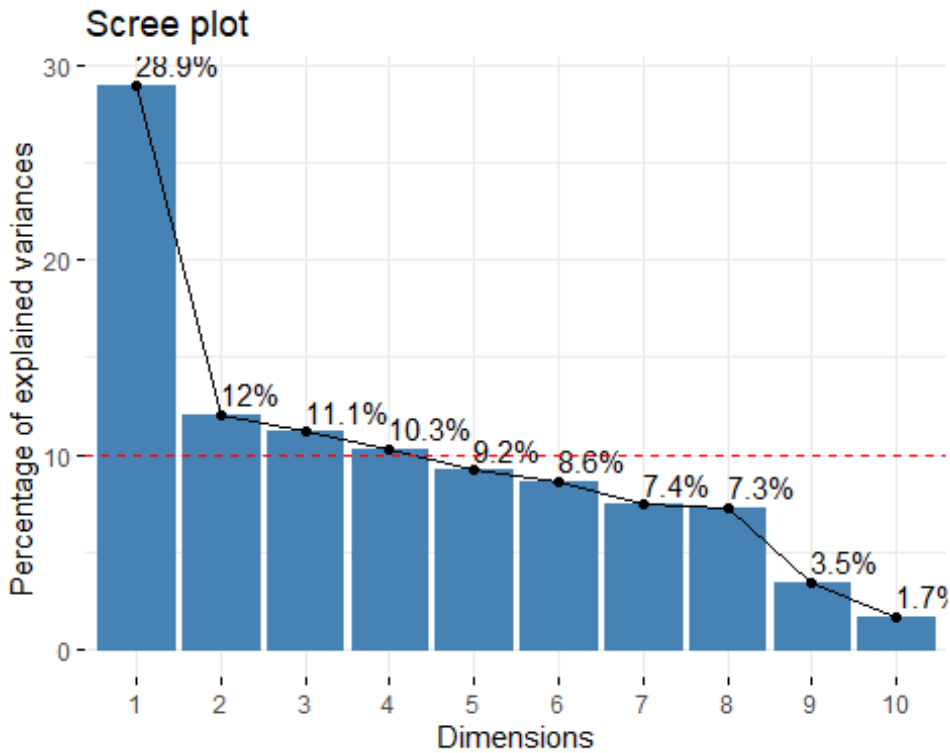
```
fviz_pca_var(pca1980_1989, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1980-1989")
```



```
#fviz_pca_biplot(pca1980_1989)
```

## PCA 1990-1999

```
songs1990_1999 = songs[songs$year_c ==
"1990-1999",c(2,3,4,5,7,9,10,11,15,16,17)]
songs1990_1999 = scale(songs1990_1999)
pca1990_1999 = PCA(songs1990_1999, graph = F, scale.unit = F, ncp = 10,
quali.sup = c(9))
eig1990_1999 = get_eig(pca1990_1999)
fviz_eig(pca1990_1999, addlabels = TRUE) +
geom_hline(yintercept=100*(1/nrow(eig1990_1999)), linetype=2, color="red")
```



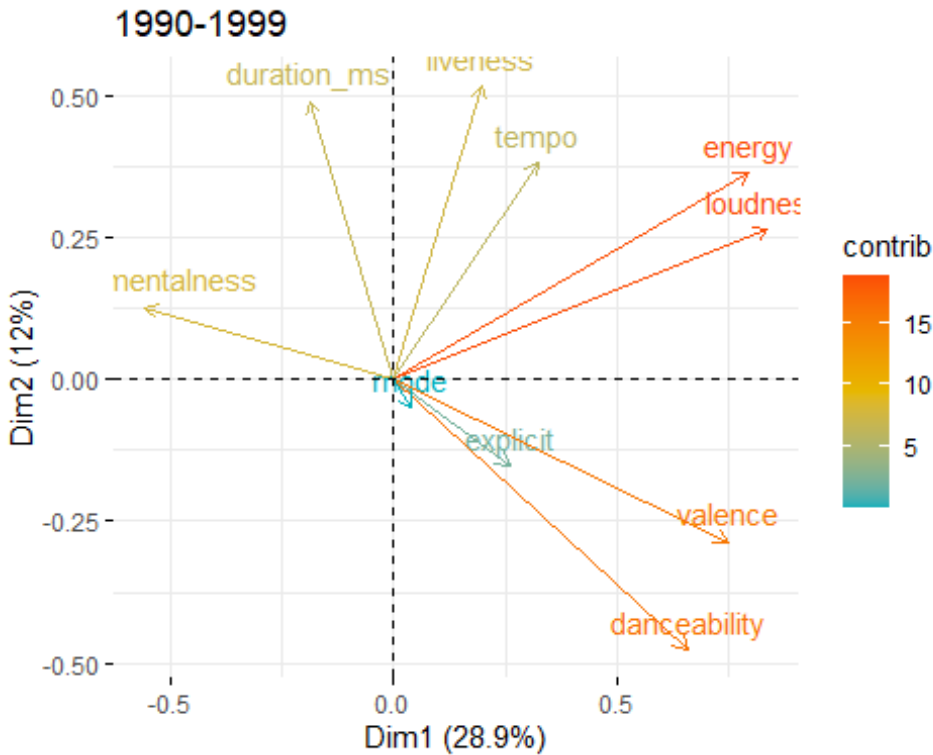
```
eig1990_1999
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1    2.8914704         28.914880          28.91488
## Dim.2    1.2016498         12.016571          40.93145
## Dim.3    1.1147388         11.147456          52.07891
## Dim.4    1.0270222         10.270285          62.34919
## Dim.5    0.9186825          9.186881          71.53607
## Dim.6    0.8630992          8.631044          80.16712
## Dim.7    0.7429407          7.429452          87.59657
## Dim.8    0.7255440          7.255485          94.85205
## Dim.9    0.3459630          3.459651          98.31171
## Dim.10   0.1688285          1.688295         100.00000
```

```
#Escogemos 6
```

```
pca1990_1999 = PCA(songs1990_1999, graph = F, scale.unit = F, ncp = 6,
quali.sup = c(9))
```

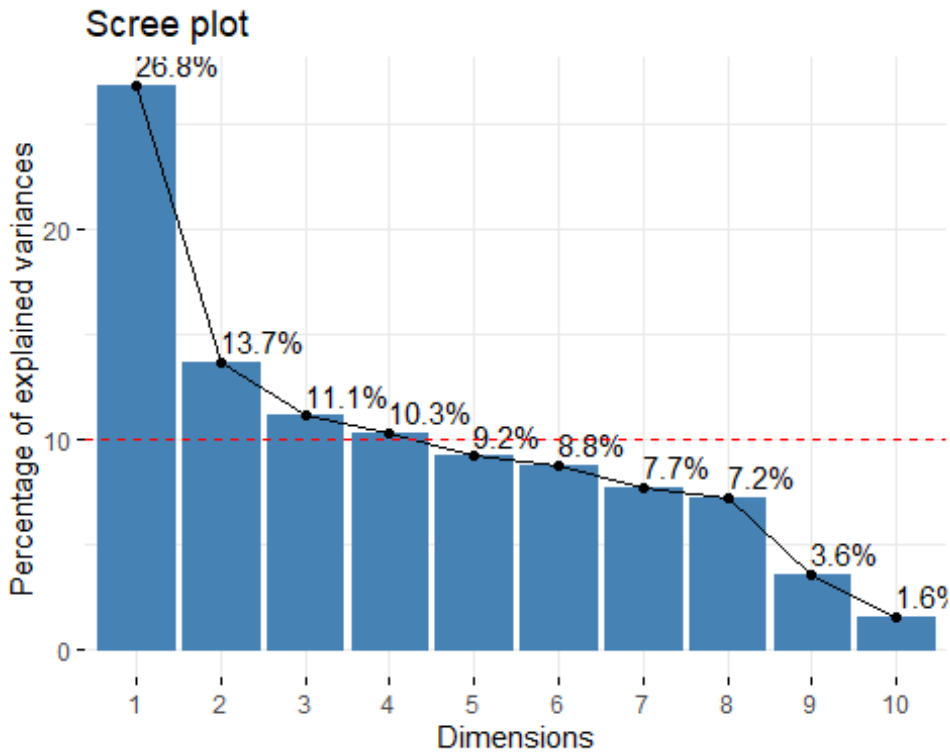
```
fviz_pca_var(pca1990_1999, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1990-1999")
```



```
#fviz_pca_biplot(pca1990_1999)
```

## PCA 2000-2009

```
songs2000_2009 = songs[songs$year_c == "2000-2009" & songs$popularity >= 10,
                        c(2,3,4,5,7,9,10,11,15,16,17)]
songs2000_2009 = scale(songs2000_2009)
pca2000_2009 = PCA(songs2000_2009, graph = F, scale.unit = F, ncp = 10,
                    quali.sup = c(9))
eig2000_2009 = get_eig(pca2000_2009)
fviz_eig(pca2000_2009, addlabels = TRUE) +
  geom_hline(yintercept=100*(1/nrow(eig2000_2009)), linetype=2, color="red")
```



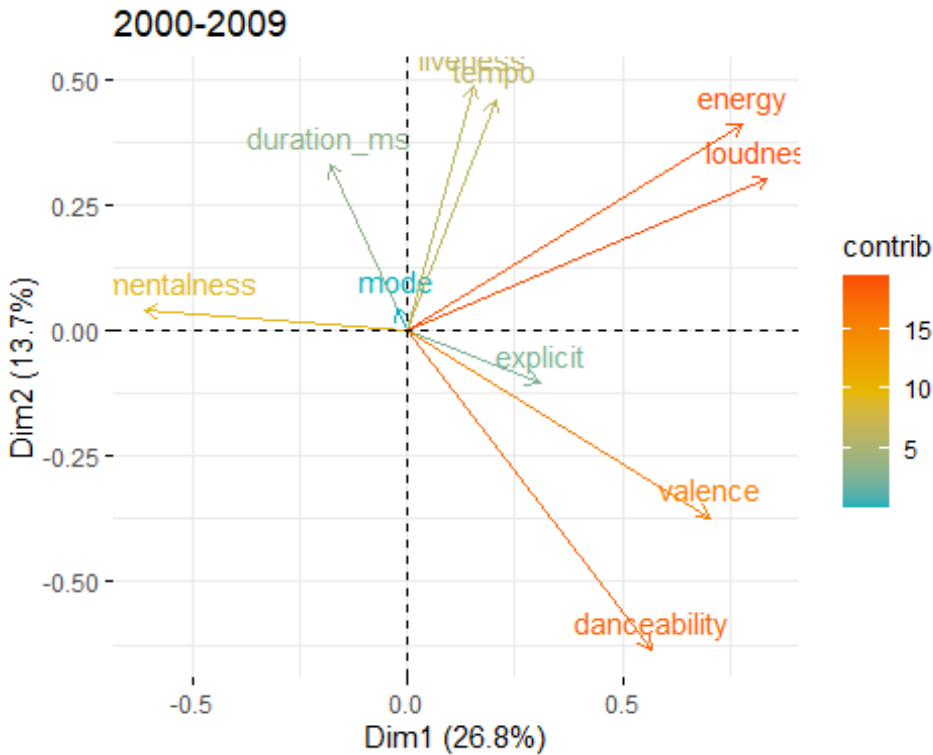
```
eig2000_2009
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1      2.6829556         26.829884          26.82988
## Dim.2      1.3685563         13.685730          40.51561
## Dim.3      1.1113196         11.113332          51.62895
## Dim.4      1.0308818         10.308944          61.93789
## Dim.5      0.9193739          9.193851          71.13174
## Dim.6      0.8771041          8.771149          79.90289
## Dim.7      0.7734384          7.734478          87.63737
## Dim.8      0.7197425          7.197513          94.83488
## Dim.9      0.3583857          3.583900          98.41878
## Dim.10     0.1581198          1.581218         100.00000
```

```
#Escogemos 6
```

```
pca2000_2009 = PCA(songs2000_2009, graph = F, scale.unit = F, ncp = 6,
quali.sup = c(9))
```

```
fviz_pca_var(pca2000_2009, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "2000-2009")
```

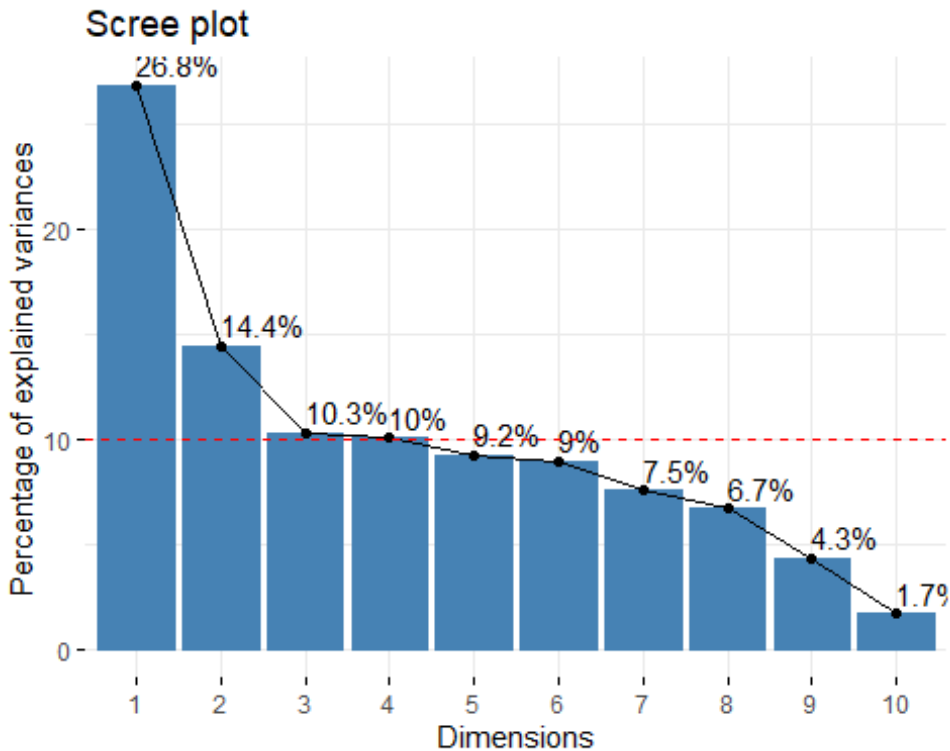


```
#fviz_pca_biplot(pca2000_2009)
```

## PCA 2010-2020

```
songs2010_2020 = songs[songs$year_c == "2010-2020" & songs$popularity >= 20,
                        c(2,3,4,5,7,9,10,11,15,16,17)]
songs2010_2020 = scale(songs2010_2020)
pca2010_2020 = PCA(songs2010_2020, graph = F, scale.unit = F, ncp = 10,
                  quali.sup = c(9))
eig2010_2020 = get_eig(pca2010_2020)
fviz_eig(pca2010_2020, addlabels = TRUE) +
  geom_hline(yintercept=100*(1/nrow(eig2010_2020)), linetype=2, color="red")
```





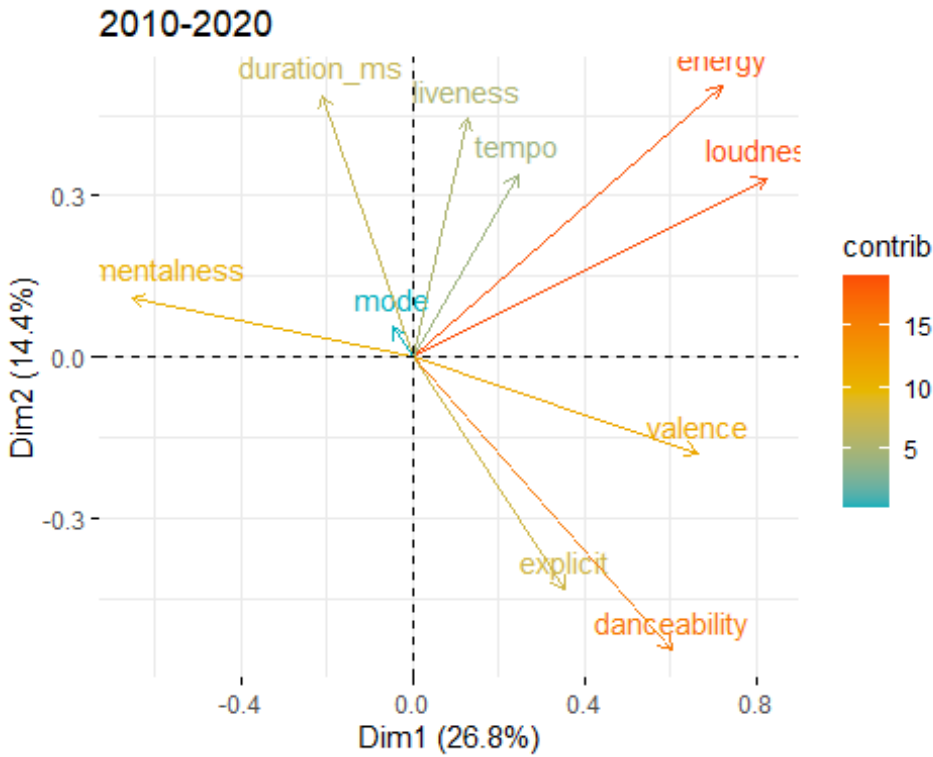
```
eig2010_2020
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1    2.6794617         26.794868          26.79487
## Dim.2    1.4448409         14.448545          41.24341
## Dim.3    1.0287005         10.287101          51.53051
## Dim.4    1.0048868         10.048962          61.57948
## Dim.5    0.9240717          9.240803          70.82028
## Dim.6    0.8949978          8.950062          79.77034
## Dim.7    0.7546701          7.546772          87.31711
## Dim.8    0.6679395          6.679458          93.99657
## Dim.9    0.4307670          4.307711          98.30428
## Dim.10   0.1695703          1.695719         100.00000
```

```
#Escogemos 6
```

```
pca2010_2020 = PCA(songs2010_2020, graph = F, scale.unit = F, ncp = 6,
quali.sup = c(9))
```

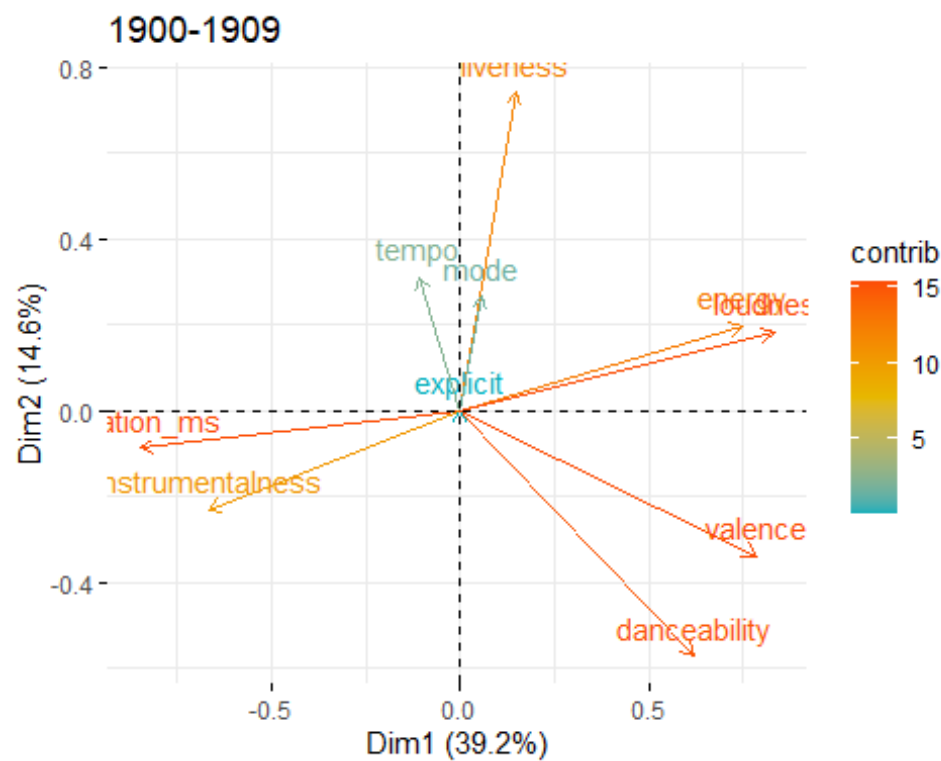
```
fviz_pca_var(pca2010_2020, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "2010-2020")
```



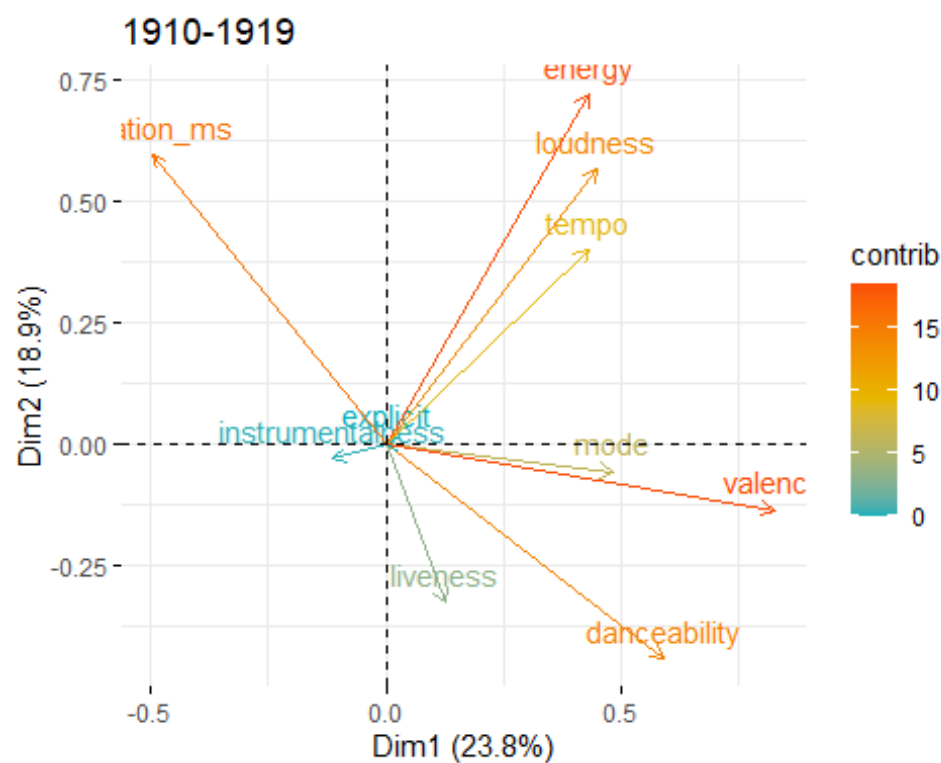
```
#fviz_pca_biplot(pca2010_2020)
```

## Resultados finales por décadas

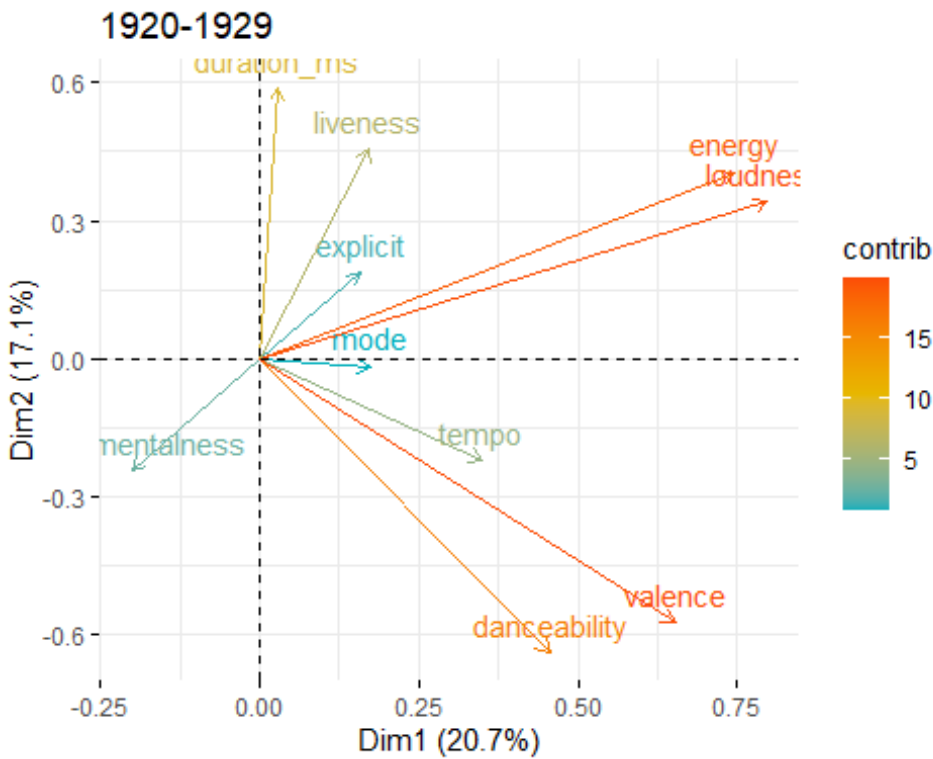
```
par(mfrow = c(3,5))
fviz_pca_var(pca1900_1909, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1900-1909")
```



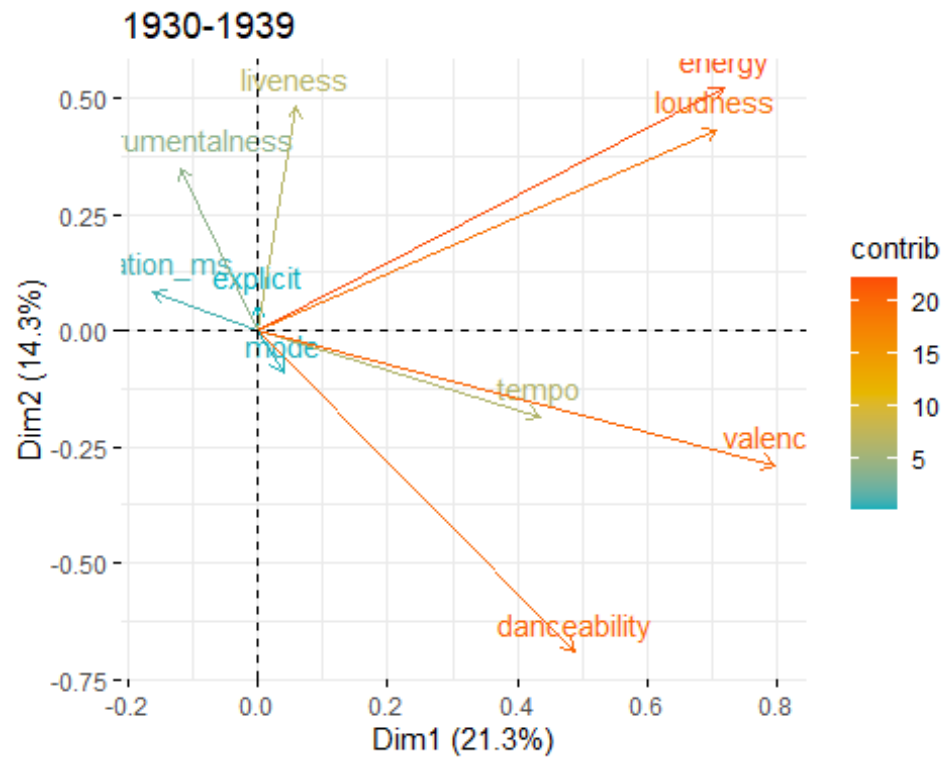
```
fviz_pca_var(pca1910_1919, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1910-1919")
```



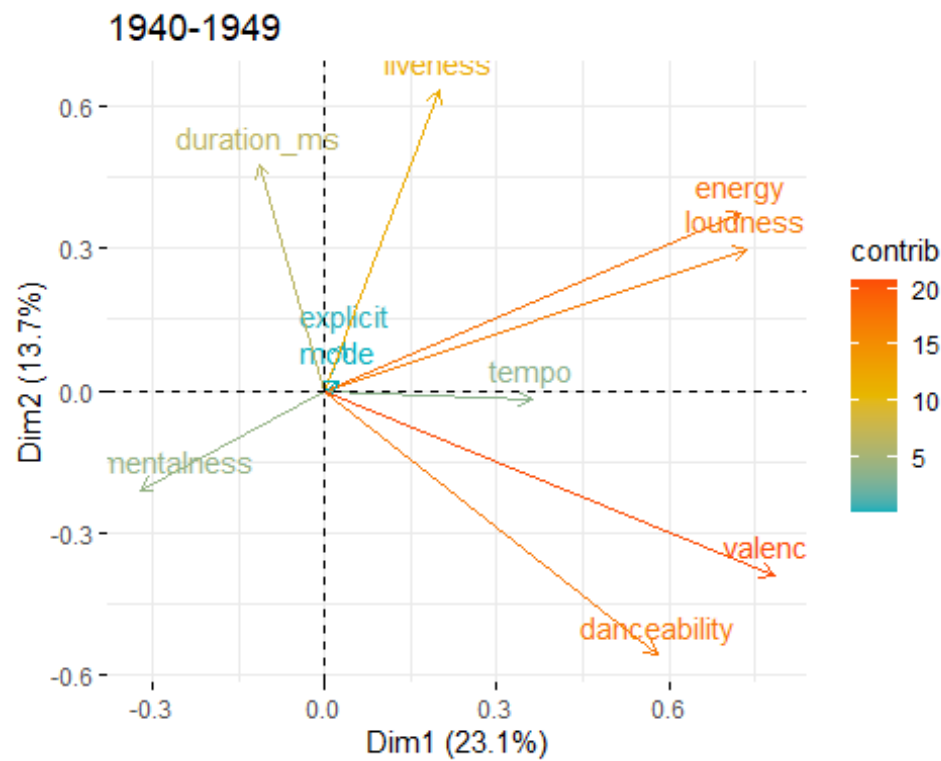
```
fviz_pca_var(pca1920_1929, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1920-1929")
```



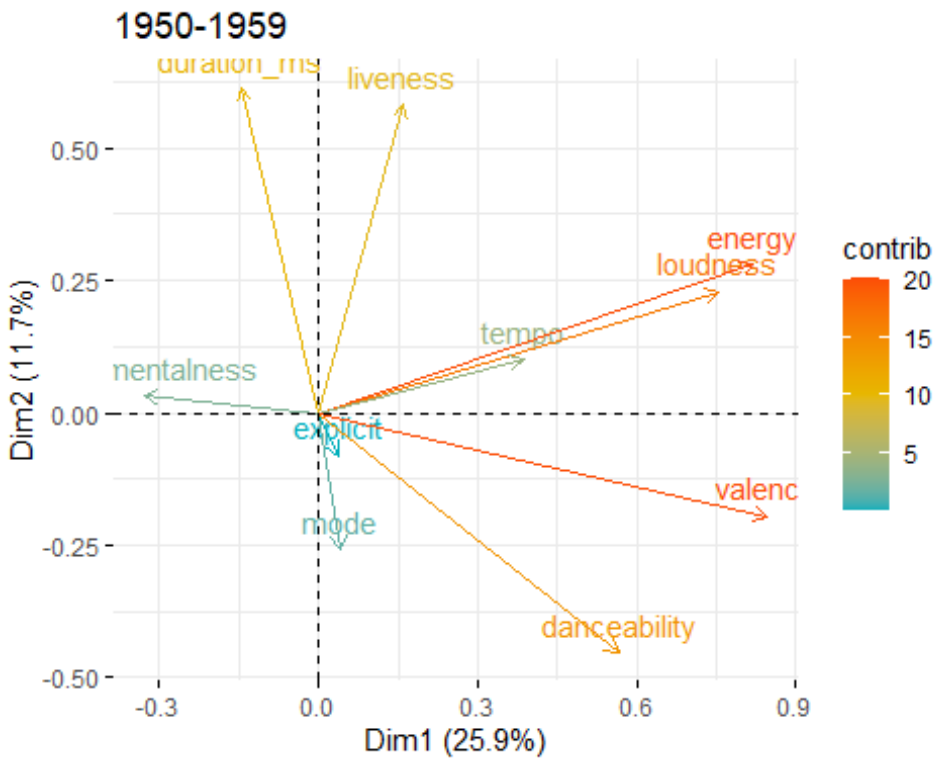
```
fviz_pca_var(pca1930_1939, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1930-1939")
```



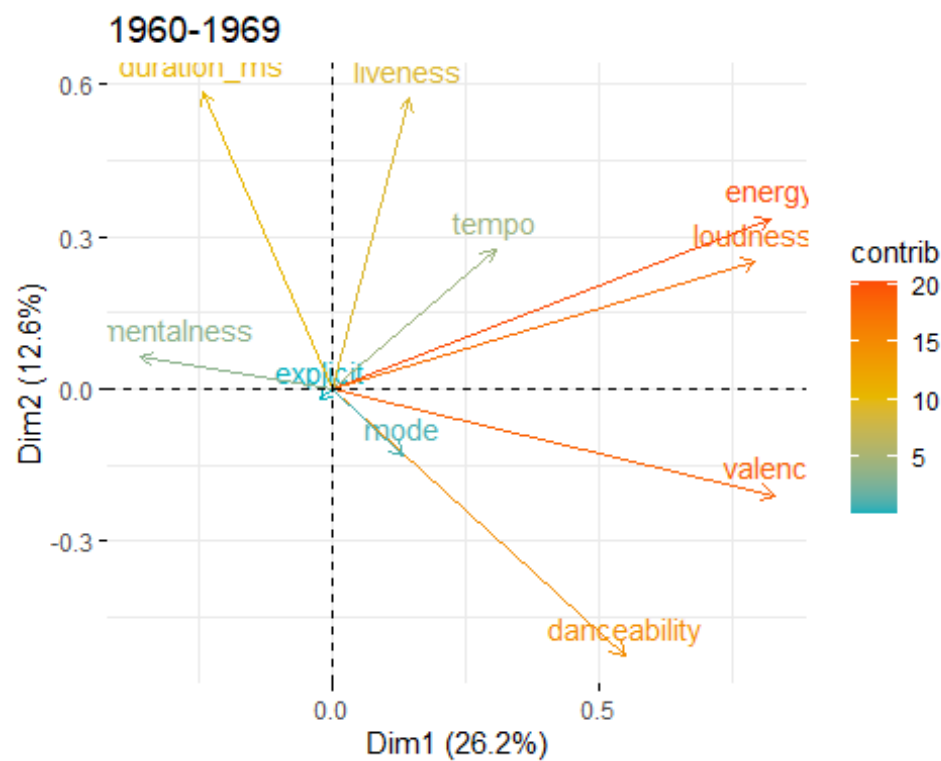
```
fviz_pca_var(pca1940_1949, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1940-1949")
```



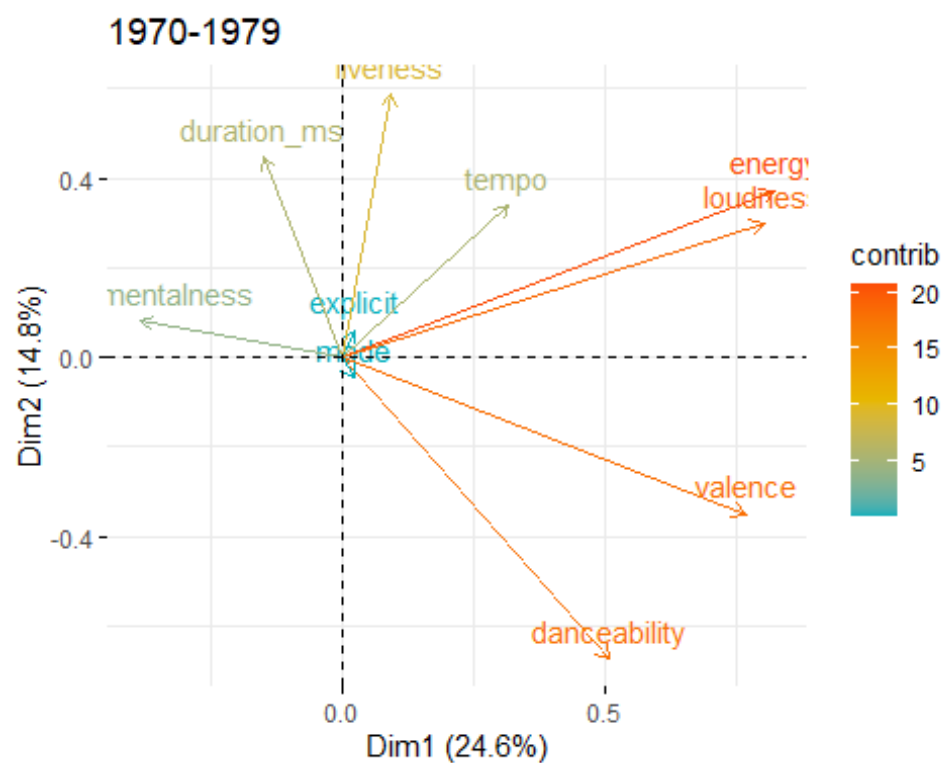
```
fviz_pca_var(pca1950_1959, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1950-1959")
```



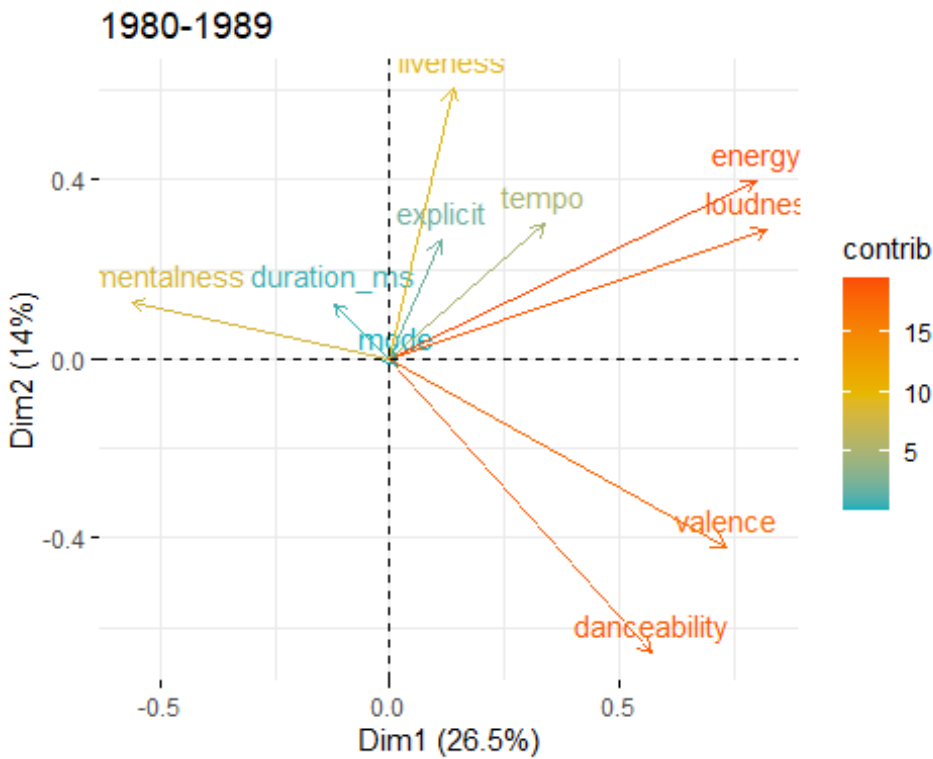
```
fviz_pca_var(pca1960_1969, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1960-1969")
```



```
fviz_pca_var(pca1970_1979, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1970-1979")
```

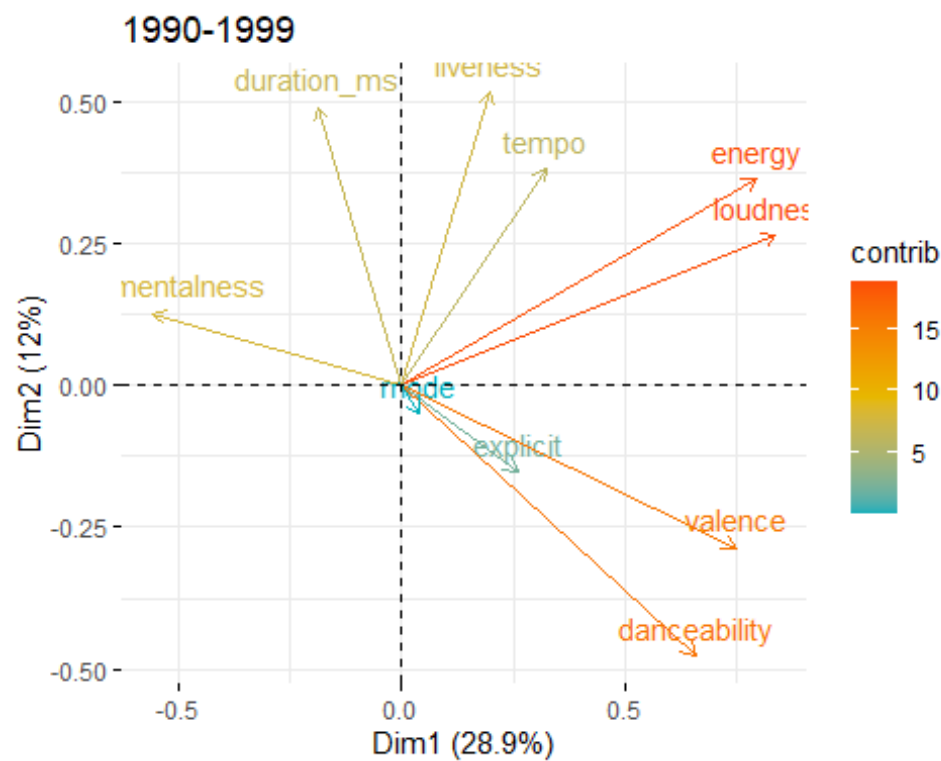


```
fviz_pca_var(pca1980_1989, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1980-1989")
```

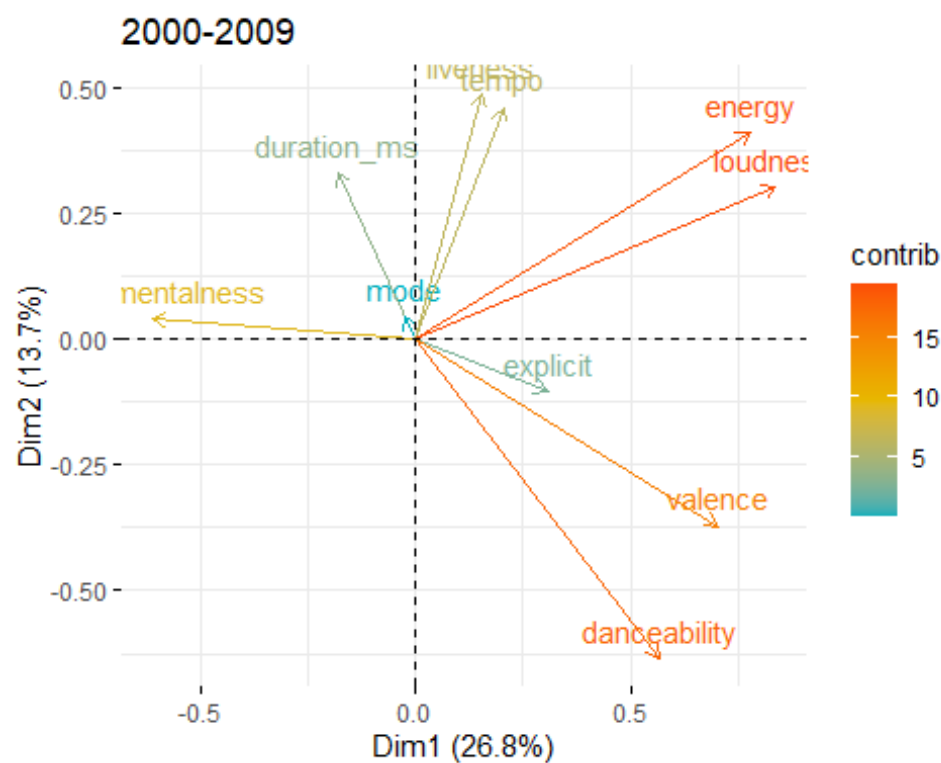


```
fviz_pca_var(pca1990_1999, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "1990-1999")
```

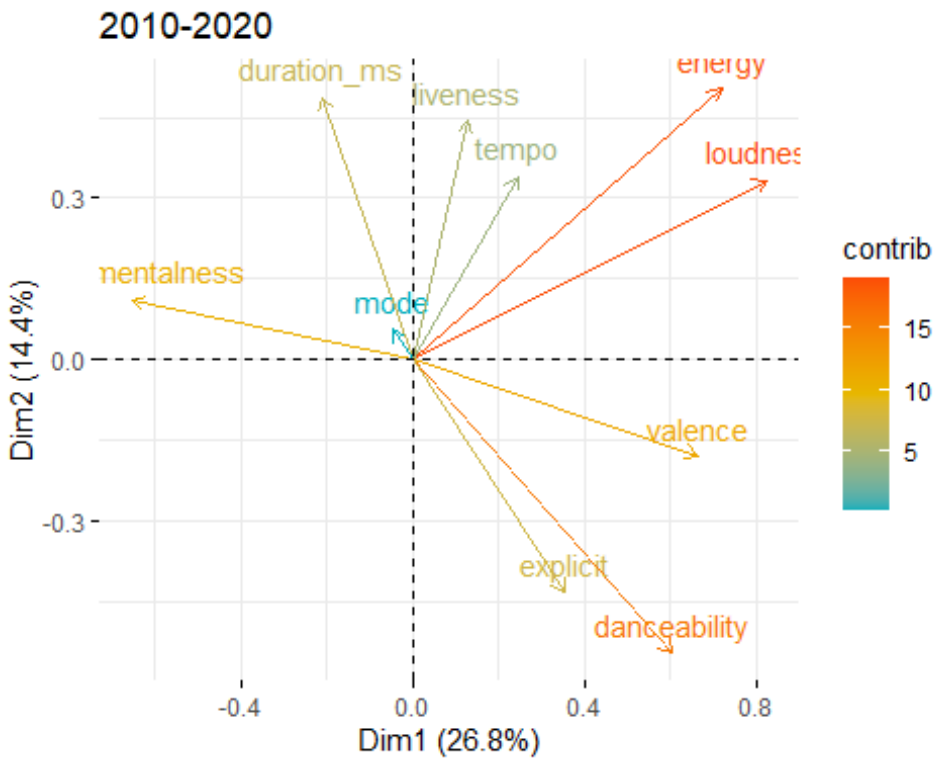




```
fviz_pca_var(pca2000_2009, col.var = "contrib", gradient.cols = c("#00AFBB",  
"#E7B800", "#FC4E07"), title = "2000-2009")
```



```
fviz_pca_var(pca2010_2020, col.var = "contrib", gradient.cols = c("#00AFBB",
"#E7B800", "#FC4E07"), title = "2010-2020")
```



Como se puede comprobar, existen cambios significativos en muchas variables, sobre todo en las primeras décadas:

### 1. En 1900-1910

- En la primera componente, las variables más importantes son la energía, el volumen, la valencia, la danceabilidad (positivamente), lo instrumental de la canción y la duración (negativamente). En esta década se favorecen las canciones cortas, con mucha letra (no instrumentales), muy altas en volumen y enérgicas, que transmiten una sensación positiva y por lo tanto son bailables.
- En la segunda componente, la variable más importante es el espectáculo en vivo, con lo que en esta década se favorecen canciones que han sido tocadas delante de un público (es decir, que no han sido pregrabadas).

### 2. En 1910-1920

Existen ciertas similitudes con la anterior década pero con sutiles diferencias:

- En la primera componente, las variables más importantes son parecidas a las de la década pasada (energía, volumen, valencia, danceabilidad), aunque su importancia es diferente: en esta década se le da mucha más importancia a la valencia (positividad de la canción), y variables como la energía o el sonido disminuyen su importancia. Cabe

destacar la aparición del tempo y de la clave (mayor o menor) como variables importantes. Por lo tanto, en esta década se favorecen las canciones alegres, rápidas, cortas y que transmiten una sensación muy positiva. Estas canciones tenderán a usar más el modo mayor.

- En la segunda componente, la variable más importante es la energía de las canciones, resultado que concuerda y respalda el tipo de canciones descritas gracias a la primera componente principal.

### 3. En 1920-1930

En esta década se pueden apreciar un par de cambios importantes:

- En la primera componente, la energía y volumen vuelven a cobrar más importancia, siendo de hecho más significativas que la valencia y la danceabilidad. No existe ninguna variable que impacte de forma significativa y negativamente a las canciones. Por lo tanto, en esta década se favorecen canciones con un volumen muy elevado, y en consecuencia muy energéticas, perfectas para ser bailadas y disfrutadas. A diferencia que la década pasada, la tonalidad de las canciones (mayor/menor) ya no es tan importante.
- En la segunda componente, las variables más significativas son la duración de la canción y el espectáculo en vivo: en esta década se favorecen también las canciones largas que se han cantado delante de un público. Este tipo de canciones podrían ser conciertos.
- Cabe destacar que esta es la primera década donde la variable explicit (si la canción contiene lenguaje explícito) empieza a tener una proyección no nula (o prácticamente no nula) en alguna de las componentes.

### 4. En 1930-1940

En esta década se “combinan” características de las dos anteriores:

- En la primera componente, la energía, la positividad, el volumen y la danceabilidad son las variables más significativas, teniendo las tres primeras un grado de importancia similar, mientras que la danceabilidad tiene un poco menos de relevancia. La duración de las canciones vuelve a tener un impacto negativo (aunque pequeño), y el tempo vuelve a cobrar relevancia de forma positiva (aunque no tanto como las otras cuatro variables mencionadas). En esta década se favorecen las canciones muy altas, rápidas, que transmitan mucha positividad y energía y que sean bailables (similar a la década 1900-1910).
- En la segunda componente, la única variable significativa es el espectáculo en vivo, que afecta de forma positiva: se favorecen las canciones no pregrabadas, como en las décadas anteriores.
- A diferencia de la década anterior, explicit tiene proyecciones muy bajas en ambas componentes.

## 5. En 1940-1950

Esta década es muy similar a la anterior:

- No existen cambios significativos en la primera componente salvo el impacto negativo de la ausencia de voces (vocals) en una canción. En esta década se favorecen el mismo tipo de canciones que en la década pasada, pero preferiblemente con letra.
- En la segunda componente, de igual manera que en la década de los 30, se favorecen las canciones largas y cantadas delante de un público.

Como se puede comprobar parece que aunque existan pequeñas diferencias entre cada década, hay una preferencia universal a lo largo del tiempo: canciones rápidas, de alto volumen, que puedan ser bailables y que hagan felices a quienes las oigan.

## 6. En 1950-1960

No existen muchos cambios de esta década a la anterior:

- En la primera componente todo se mantiene igual. No obstante, cabe destacar que la energía ahora es ligeramente más importante que el volumen (en la década anterior eran prácticamente igual de importantes). En esta década se favorecen el mismo tipo de canciones que en la década anterior.
- En la segunda componente, lo único destacable es que la duración de las canciones ahora tiene la misma importancia que el espectáculo en vivo (antes la duración era menos importante).

## 7. En 1960-1970

- La única diferencia considerable en esta década es que la tonalidad (mayor/menor) de la canción cobra menos relevancia en la segunda componente y gana un poco en la primera componente, aunque su importancia sigue siendo poca.

## 8. En 1970-1980

- De nuevo, se presentan similitudes con las dos décadas anteriores. La única diferencia a mencionar es que la importancia de la duración en la segunda componente vuelve a disminuir. El resto de variables importantes no cambian en magnitud.

Como se puede apreciar, a partir de la década de los 50 la tendencia parece estabilizarse, con pequeños cambios no tanto en la dirección de las variables, sino en la magnitud de estas.

## 9. En 1980-1990

A pesar de todo, en este década hay más cambios que en las décadas anteriores:

- La importancia de la duración de las canciones se ve enormemente reducida en ambas componentes, con lo que deja de ser una variable a tener en cuenta para determinar qué tipo de canciones se solían escuchar en este periodo.
- La variable explicit (si la canción contiene letra explícita) gana una importancia relevante (de casi la misma magnitud que el tempo) en la segunda componente. Por lo tanto, en esta década se favorecen canciones enérgicas, bailables, rápidas, cantadas en directo y con cierto lenguaje explícito. La duración de estas canciones puede variar, al no ser esta una variable tan importante.

#### 10. En 1990-2000

- La duración vuelve a cobrar importancia en la segunda componente, al igual que el tempo que aumenta su importancia. La variable explicit sufre un cambio brusco, ya que pasa de encontrarse en el cuadrante superior derecho al inferior derecho, siendo mayor su proyección sobre la primera componente: es decir, las canciones de esta década son similares a las de la década anterior, pero se vuelve a tener en cuenta la duración de estas.
- Como la variable explicit sigue siendo relevante pero en la otra componente, los motivos por los que se sigue prefiriendo canciones con contenido explícito son diferentes que los de la década pasada.

#### 11. En 2000-2010

- La duración de las canciones vuelve a perder relevancia. Parece que esta variable oscila entre décadas, aunque cabe destacar que a pesar de todo, y a diferencia de la década de los 80, aún sigue siendo una variable a considerar.
- Tras muchas décadas en las que el tempo de las canciones cobraba más importancia en la primera componente (su proyección era mayor en dicha componente), esta variable se consolida como importante pero en la segunda componente: es decir, se siguen favoreciendo canciones con un tempo rápido, pero por motivos diferentes.

#### 12. En 2010-2020

- Como hemos visto estas últimas décadas, la duración vuelve a ganar importancia esta década.
- El tempo es menos relevante al ser su proyección sobre la segunda componente menor que la década pasada. Ocurre lo mismo con la emisión en vivo de las canciones.
- Sorprendentemente, la variable explicit aumenta mucho su importancia en la segunda componente de forma negativa.
- El resto de variable se mantienen invariables,

- Por lo tanto, en la actualidad se favorecen las canciones con letra, preferiblemente sin contenido explícito, que sean bailables, rápidas, enérgicas, positivas y por lo general emitidas en directo, con una duración no muy corta, de unos 3 minutos y medio.

## Conclusión

A lo largo del último siglo el gusto por la música ha ido variando ligeramente: no obstante, a pesar del paso del tiempo, las canciones enérgicas, positivas y bailables parecen ser un punto en común para todas las décadas estudiadas. Las mayores diferencias parecen estar en variables más sutiles de las canciones, como la duración, la aparición de contenido explícito o el tempo de la canción. En conclusión: la diferencia está en los detalles.