```
---
title: "French Elite analysis"
author: "Gary Sztajnman"
date: "February 22, 2016"
output: pdf_document
---
```

In this homework, we are going to study the "FRENCH FINANCIAL ELITE" dataset
downloaded from http://moreno.ss.uci.edu/.
This dataset is composed of three 28 by 28 matrices on the influence, the
belonging of an elite and the friendship of 28 top French leaders in Finance.
We also have information about their background such as religion, education or
age...

\section*{Hypothesis}
I want to study the characteristics of powerfull member of an elite.
At first, I wanted to see the impact of higher education on the influence of
leaders in an elite group. However, education level may not be a differentiable
factor as all leaders when through inpressive school (in particular in France
where people are considered based on the school they went to.
Another factor that is more easy to anticipate is the level of religion defined as
the belief in god and the frequency of praying.

We want to test the relation between religion and position of power in an elite
group. In particular, we hypothesis that the more powerful a leader is the less he
prays and believe in god.

\section*{Explanation}
In France, the concept of "laïcité" (secularism) is particularly important in
higher education school and in high business circle and elite people come from
these school. Hence, even if they may come from religious family, they will
probably get away from religion after and during their higher education.
Furthermore, the financial elite is busy and we could think that don't have to
much time to pray and think about religion and god. Finally, financial leaders
tends to prove their skills by their understanding of financial markets and
business and not by showing their faith in god.

So I would think that French financial leaders are less linked to religion when
they are more influencial.

Let's use the dataset to analyze that

\section*{Retrieveing and cleaning the data}
```{r}

### Libraries----
library(data.table)
library(dplyr)
library(plyr)
library(ggplot2)
library(igraph)

### import data----
netw <- read.csv("http://moreno.ss.uci.edu/ffe.dat", header=T, sep=";", skip= 36)
netw = setDT(netw)[, tstrsplit(DATA., ' ')]
netw = netw[,V1:=NULL]
```

```
#splitting the network data to advice/friendship/report
influence <- as.matrix(slice(netw, 1:28))
elite <- as.matrix(slice(netw, 29:56))
friend <- as.matrix(slice(netw, 57:84))
```

```{r}
#importing attributes data
setwd("/Users/garyair/Desktop/Dropbox/Columbia/SNA/Labs/SNA-EgoNetwork-1/French
elite/")
attributes <- read.csv("attributes.csv", header=TRUE)

#Clean missing data in attributes
attributes$igyear <- NULL
attributes$birthplace <- NULL
attributes$polyyear <- NULL
attributes$enayear <- NULL
attributes$zipcode <- NULL


#We decide to restrict our analysis to 3 major topics: age, education and
religion. Hence, we delete other attributes
attributes$fathers.lev <- NULL
attributes$masons <- NULL
attributes$socialreg <- NULL
attributes$eliteprom <- NULL
attributes$prestige <- NULL
attributes$clubs <- NULL
attributes$topboards <- NULL
attributes$inspec.gen <- NULL
attributes$cabinet <- NULL
attributes$finance.min <- NULL
attributes$party <- NULL
attributes$elitevote <- NULL


#Cleaning and polishing data
names(attributes)
#"sciencepoly" "polytech"    "university"  "normal.sch"  "ena" "religion"
attributes$sciencepoly[attributes$sciencepoly == 2] <- FALSE
attributes$sciencepoly[attributes$sciencepoly == 1] <- TRUE

attributes$polytech[attributes$polytech == 2] <- FALSE
attributes$polytech[attributes$polytech == 1] <- TRUE

attributes$university[attributes$university == 2] <- FALSE
attributes$university[attributes$university == 1] <- TRUE

attributes$normal.sch[attributes$normal.sch == 2] <- FALSE
attributes$normal.sch[attributes$normal.sch == 1] <- TRUE

attributes$ena[attributes$ena == 2] <- FALSE
attributes$ena[attributes$ena == 1] <- TRUE


```
```

\section*{Variables}

For this study, 3 types of variables have been selected.
 1."birthdate"
 2."sciencepoly", "polytech", "university", "normal.sch", "ena"
 3. "religion"

Birthdate is an independent variable, it gives an idea of which elite member is
part of which generation
Education variables are binary coded (true or false)
Religion is a class variable coded from 2 to 4 (2 is low religion level and 4 is
high religion level)

To this attribute, we add the ego network measure: transity and network size for
elite and influence

Influence network: "WHO INFLUENCES YOU"
```{r}
influenceg=graph.adjacency(influence,mode="directed",weighted=NULL)

### calculate ego network influence size ###
attributes$influence.degree <- degree(influenceg)
ggplot(attributes, aes(x= religion, y =influence.degree )) +
    geom_point(shape=1) +    # Use hollow circles
    geom_smooth()

### calculate transitivity of elite ego network ###
attributes$Influence.transitivity =transitivity(influenceg, type="local")
attributes$Influence.transitivity[attributes$Influence.transitivity == "NA"] <-
mean(attributes$Influence.transitivity)

ggplot(attributes, aes(x= religion, y =Influence.transitivity )) +
    geom_point(shape=1) +    # Use hollow circles
    geom_smooth()

```

Elite network: "WHO IS IN THE ELITE"
```{r}
eliteg=graph.adjacency(elite,mode="undirected",weighted=NULL)
### calculate ego network elite size ###
attributes$elite.degree <- degree(eliteg)
ggplot(attributes, aes(x= religion, y =elite.degree )) +
    geom_point(shape=1) +    # Use hollow circles
    geom_smooth()

### calculate transitivity of elite ego network ###
attributes$elite.transitivity =transitivity(eliteg, type="local")
attributes$elite.transitivity[attributes$elite.transitivity == "NaN"] <-
mean(attributes$elite.transitivity)

ggplot(attributes, aes(x= religion, y =elite.transitivity )) +
    geom_point(shape=1) +    # Use hollow circles
    geom_smooth()
```

```
```

\section*{Initial results}

If we look at the elite network and the influence network, we see in both case a clear decreasing factor between religion and the place in the elite.
Elite member tend to have a more influential place when they spend less time to pray.
If we just plot the average amount of excitement people claim as a function of the average educational attainment of their social circle, we see that as one's social circle is more educated, the higher excitement ego's report.

However, it's interesting to see that transitivity is positively correlated with religion in this dataset. Religious people may be less influential but more connected to their religious peers.

We will now use regression model to further study these ideas.

\section*{Model 1}
We first start by studying the impact on influence:
```{r}
summary(lm(influence.degree~ religion, attributes))
summary(lm(Influence.transitivity~ religion, attributes))
```

The regression on influence degree validates the idea of negative correlation between religion and influence with a p-value lower than 5% and a beta estimate of approximately -3.
Clearly we note that more religion means less influence in our dataset.

However, we can not confirm our hypothesis on religion and the transitivity of influence because the p-value is too high aroung 30%.

\section*{Model 2}
To be sure of our first idea on education we will the education variables in our model:
```{r}
summary(lm(influence.degree ~ ena + sciencepoly + polytech + university + normal.sch, attributes))
summary(lm(Influence.transitivity~ ena + sciencepoly + polytech + university + normal.sch, attributes))
```

As expected, we note that we cannot infer any correlation between education and influence because p-values are too high. The dataset is too small to detects the impact of these binary variables on the influence levels.

\section*{Model 3}
To confirm our analysis, we will reproduce our linear regression on the elite data.
This will help us determine if being part of the elite is linked to being influential.

```{r}
summary(lm(elite.degree~ religion, attributes))
summary(lm(elite.transitivity~ religion, attributes))
summary(lm(elite.degree ~ ena + sciencepoly + polytech + university + normal.sch,
attributes))
summary(lm(elite.transitivity~ ena + sciencepoly + polytech + university +
normal.sch, attributes))
```

The results are similar than with the influence network. We have relatively
similar level of p-value and the only thing that we can infer is that religion is
negatively correlated with being part of the elite.

\section*{Conclusion}

The answer to the main problem seems to be clear.
In the case of the French financial elite, one's relation to god does influence
its position in his network. We can propose some interpretation about the role of
religion in elite networks but more data and more analysis would help us confirm
our ideas.