Exploration 4: Two Variables Relationships as Causal Relationships

Jake Bowers

September 16, 2016

```
download.file("http://jakebowers.org/Data/ANES/anes_pilot_2016_csv.zip",
    destfile = "anespilot2016.csv.zip")
unzip("anespilot2016.csv.zip")
anespi16 <- read.csv("anes_pilot_2016.csv", as.is = TRUE, strip.white = TRUE)
```

You friend calls back. "I really appreciated all of the work that you did with describing the relationships between age and Trump support. Of course, as soon as your results came in, the team began to bicker:"Older people like Trump because they are just inherently conservative. It is not age **per se** that causes this relationship, but the fact that ideology changes over time within the life span." Then the others argued, "First, there is no fact that ideology changes over time within the life span. Second, I think that this relationship just shows that people who are older have had more time to dislike Clinton — the young people don't know that much about the Clintons, and so they are not yet disillusioned." And other group argued, "So one of you is saying that age causes Trump support because ideology is related to age? And the other is that age causes Trump support because political knowledge and experience is related to age? But, both of you seem to be ignoring the nonlinear descriptions that we did. Didn't it seem like there were different age groups that responded to Trump differently? Don't you think that this is caused by the fact that the ethnic composition of the country has changed, and that Trump consistently offends non-white voters? That is, that the relationship is caused by the underlying relationship between age and ethnicity and Trump's offensive comments?"

"I started to try to calm the waters in my team, but I only go this far."

```
require(knitr)
kable(t(as.matrix(table(anespi16$fttrump, useNA = "ifany"))))
```

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
162	93	51	46	23	31	33	12	9	6	18	12	5	9	8	2	4	9	6	7	14	7	10	1	1

```
table(anespi16$birthyr, useNA = "ifany")
```

```
1921 1924 1925 1926 1927 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939
   1    1    2    1    1    1    2    3    4    2    1    9    7   10    6
1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954
  12    8   10   14    8   16   12   16    9   16   18   20   24   28   38
1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969
  31   34   27   23   26   26   29   21   17   22   17    9   14   18    8
1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984
  13   10   24   34   20   24   21   25   26   16   24   20   22   19   18
1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997
  17   21   22   24   27   26   20   22   19   20   13   17   14
```

```r
anespi16$fttrump[anespi16$fttrump == 998] <- NA
anespi16$age <- 2016 - anespi16$birthyr

summary(anespi16$age)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  19.00   34.00   48.00   48.06   61.25   95.00
```

```r
summary(anespi16$fttrump)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   0.00    2.00   30.00   38.38   72.00  100.00       3
```

"First, I was getting confused about the use of the word, 'cause'. I didn't quite know what they meant. And, I didn't know how to start using the data on hand to engage with these alternative causal explanations if I didn't know what they meant by cause. The problem is now that I'm back abroad, I can't ask them. What is your best guess? Can you explain it to me? For example, why would thinking about cause one way help me use data to engage with these kinds of arguments? And, of course, I'm relying on you for your advice about which explanation is correct. Please help! I'm also wondering about your own favorite explanation. One that is not one of those three. What is it? What is the evidence in favor of it or against it?"

References