Problem Set 0: Bias in Self-reported Turnout

Shuai

2022-05-14

```
turnout <- read.csv("turnout.csv")

t_dim <- dim(turnout)
year_range <- range(turnout$year)
```

Q1: How many observations are there? What is the range of years covered in this data set?

A1: There are 9 observations. The range of years covered in this data set is from 1980 to 2008.

```
turnout$VAPtr <- turnout$total / (turnout$VAP + turnout$overseas) * 100
turnout$VEPtr <- turnout$total / turnout$VEP * 100

knitr::kable(turnout[,c("year", "VAPtr", "VEPtr")])
```

year	VAPtr	VEPtr
1980	52.03972	54.19551
1982	40.24522	42.13701
1984	52.53748	55.24860
1986	36.07845	38.14115
1988	49.72260	52.76848
1990	35.93884	38.41895
1992	54.04097	58.11384
1994	38.03086	41.12625
1996	47.53376	51.65793
1998	34.83169	38.09316
2000	49.34211	54.22449
2002	35.82850	39.51064
2004	54.54777	60.10084
2008	55.67409	61.55433

Q2: What difference do you observe between the different measures of turnout?

A2: 2.1557847, 1.8917891, 2.7111147, 2.0627026, 3.0458778, 2.4801053, 4.0728658, 3.0953969, 4.1241656, 3.2614698, 4.882388, 3.6821454, 5.553078, 5.8802387

```
turnout$diffVAP <- turnout$ANES - turnout$VAPtr
turnout$diffVEP <- turnout$ANES - turnout$VEPtr

turnout[,c("diffVAP", "diffVEP")]
```

```
##     diffVAP   diffVEP
## 1  18.96028 16.804491
## 2  19.75478 17.862987
```

```
## 3  21.46252 18.751404
## 4  16.92155 14.858846
## 5  20.27740 17.231520
## 6  11.06116  8.581054
## 7  20.95903 16.886160
## 8  17.96914 14.873745
## 9  25.46624 21.342072
## 10 17.16831 13.906838
## 11 23.65789 18.775507
## 12 26.17150 22.489359
## 13 22.45223 16.899156
## 14 22.32591 16.445672
```

Q3: How big are each of these differences on average? What is the range of the differences?

A3:

- average of diffVAP: 20.3291382, average of diffVEP: 16.8363437
- range of diffVAP: (11.0611594, 26.1715043)
- range of diffVEP: (8.5810541, 22.4893589)