live02

January 23, 2025

1 Live Coding Session Module 02

1.1 Gallup Poll # 1936-0053

- Teachers' Oath/Government Loans for Farmers/Employers Insurance Contributions/Presidential Candidates [Roper #31087039]

[4]:
```
import pandas as pd
import numpy as np
```

[5]:
```
gallup = pd.read_csv('data/USAIPO1936-0053.csv')
```

[7]:
```
gallup.head(3).T
```

[7]:

	0	1	\
form	NaN	NaN	
state	Indiana	Illinois	
region	East Central	East Central	
female	Male	Male	
age	NaN	NaN	
class	Av+	Av+	
OCCUPATION1	Skilled workers	Skilled workers	
OCCUPATION2	NaN	NaN	
OCCUPATION3	NaN	NaN	
black	NaN	NaN	
size	Urban	Urban	
education	NaN	NaN	
AGE_3WAY	NaN	NaN	
AGE40	NaN	NaN	
OCC8	Labor	Labor	
prof	Not Professional	Not Professional	
REGION4	Midwest	Midwest	
EDU_RECODE	NaN	NaN	
VOTE_PRO	Landon	Landon	
VOTE_RETRO	Hoover	Hoover	
PHONE_RECODE	NaN	NaN	
CAR_RECODE	NaN	NaN	
ballot	53	53	

Q1	Yes	Yes
Q2	Yes	NaN
Q3	Yes	Yes
Q4A	Roosevelt	Landon
Q4B	Roosevelt	Landon
Q4C	Landon	Landon
Q5A	Landon	Landon
Q5B	Yes, voted for Hoover	Yes, voted for Hoover
farm	Non-Farm	Non-Farm
SIZE3	Urban	Urban
urban	Urban	Urban
StPOAbrv	in	il
SOUTH11	Non-South	Non-South
SOUTH11xBLACK	NaN	NaN
SOUTH12	Non-South	Non-South
SOUTH12xBLACK	NaN	NaN
south	Non-South	Non-South
SOUTHxBLACK	NaN	NaN
year	1936	1936
WtPubFeas	NaN	NaN
WtVotFeas	NaN	NaN

	2
form	NaN
state	Michigan
region	East Central
female	Male
age	NaN
class	Av
OCCUPATION1	Business
OCCUPATION2	NaN
OCCUPATION3	NaN
black	NaN
size	Urban
education	NaN
AGE_3WAY	NaN
AGE40	NaN
OCC8	Professional
prof	Professional
REGION4	Midwest
EDU_RECODE	NaN
VOTE_PRO	Landon
VOTE_RETRO	Hoover
PHONE_RECODE	NaN
CAR_RECODE	NaN
ballot	53
Q1	No

```
Q2                                 No
Q3                                Yes
Q4A                            Landon
Q4B                            Landon
Q4C                            Landon
Q5A                            Landon
Q5B             Yes, voted for Hoover
farm                         Non-Farm
SIZE3                           Urban
urban                           Urban
StPOAbrv                           mi
SOUTH11                     Non-South
SOUTH11xBLACK                     NaN
SOUTH12                     Non-South
SOUTH12xBLACK                     NaN
south                       Non-South
SOUTHxBLACK                       NaN
year                             1936
WtPubFeas                         NaN
WtVotFeas                         NaN
```

[9]:
```
pd.crosstab(gallup['Q3'],gallup['Q5A'])
```

[9]:
```
Q5A  Landon  Lemke  Other party  Roosevelt  Thomas
Q3
No       921     41            6        442      10
Yes      820    109            6       2048      41
```

1.2 Dataset 2

[11]:
```
russia = pd.read_csv('data/31121191.csv',encoding='latin')
```

[15]:
```
#pd.set_option('display.max_rows'=63)
russia.head(3).T
```

[15]:
```
                                    0  \
ID                     0791e2f21e4cea8f
SEX                                Male
AGE                                  63
EDU      Incomplete secondary education
REGION       The Trans-Baikal Territory
…                                    …
ss6
ss7_1
ss7_2
ss8
weight1                       1.466481
```

```
                                                        1  \
ID                                         0791e3674d8ced7d
SEX                                                   Male
AGE                                                     47
EDU       Secondary special education (technical school)
REGION                       The Primorsky Krai/ Territory
…                                                       …
ss6
ss7_1
ss7_2
ss8
weight1                                          1.249258

                                                        2
ID                                         0791e2f9facceb98
SEX                                                 Female
AGE                                                     30
EDU       Secondary special education (technical school)
REGION                       The Primorsky Krai/ Territory
…                                                       …
ss6
ss7_1
ss7_2
ss8
weight1                                          1.658539

[62 rows x 3 columns]
```

[17]:
```
russia['Q2_1'].value_counts()
```

[17]:
```
Q2_1
Approve        37597
Disapprove      6818
Do not know     3768
Name: count, dtype: int64
```

1.3 Third Dataset: Pakistan

[28]:
```
widthsd = pd.read_csv('data/USIA data widths.csv')
widthsd
```

[28]:
```
     Field  Width
0        1      4
1        2      3
2        3      1
3        4      1
```

```
4        5      1
..     …    …
131    132      2
132    133      1
133    134      1
134    135      1
135    136      1

[136 rows x 2 columns]
```

[29]:
```
usia = pd.read_fwf('data/i20068.dat',widths=widthsd['Width'],header=None)
usia.head(3).T
```

[29]:
```
          0      1      2
0         1      2      3
1     155.0  155.0  155.0
2       2.0    2.0    2.0
3       1.0    1.0    1.0
4       2.0    2.0    2.0
..      …    …    …
131     4.0    NaN    NaN
132     NaN    NaN    NaN
133     6.0    NaN    NaN
134     4.0    3.0    5.0
135     NaN    NaN    NaN

[136 rows x 3 columns]
```

[ ]: