Quantitative Sociological Analysis

Descriptive Statistics

Exercise 4

February 18, 2025

Summarizing descriptive statistics: GSS

- descriptive statistics table
  - includes summary for all measures of interest

Descriptive Statistics Table: General Social Survey 1972-2022 (N = 64,555)

Variable	Mean (SD)	Median	Min.	Max.	Level of Measurement
Happiness			1	3	ordinal
Not too Happy	0.14				
Pretty Happy	0.56				
Very Happy	0.30				
Age	46.45 (17.63)	44.00	18	89+	interval-ratio
Female	0.56		0	1	nominal
White	0.80		0	1	nominal
Educational Attainment			0	4	ordinal
Less than HS	0.21				
HS	0.30				
Some College	0.24				
BA	0.17				
Graduate Deg.	0.08				
Married	0.53		0	1	nominal
Household Size	2.64 (1.51)	2.00	1	16	interval-ratio
Political Party Affiliation			1	3	nominal
Democrat	0.49				
Indep./Other	0.17				
Republican	0.34				

Especially this section of the RScript

```
176 ###
177 # Computing Descriptive Statistics for Summary Table
178 ##
179
180 # Happiness: ordinal variable
181 prop.table(table(GSS$happy))
182
183 # Age: interval-ratio variable
184 # see how the summary command provides many different statistics
185 summary(GSS$age)
186 sd(GSS$age)
187
188 # Female: nominal variable (binary, means only two categories)
189 # this is a special case where only need to report one category,
190 # because remainder is intuitive (sums to 100%, see Descriptive Table)
191 mean(GSS$female)
192
193 # White: nominal variable (binary)
194 mean(GSS$white)
195
196 # Educational Attainment: ordinal variable
197 prop.table(table(GSS$educ_deg))
198
199 # Married: nominal variable (binary)
200 mean(GSS$married)
201
202 # Household Size: interval-ratio variable
203 summary(GSS$hhsize)
204 sd(GSS$hhsize)
205
206 # Political Party Affiliation: nominal variable
207 prop.table(table(GSS$polit_party))
208
209 ### End Descriptive Example for Summary Table ###
```

Summarizing descriptive statistics: Netflix

Descriptive Statistics Table: Netflix Survey SOC303 Spring 2025 (N = 22)

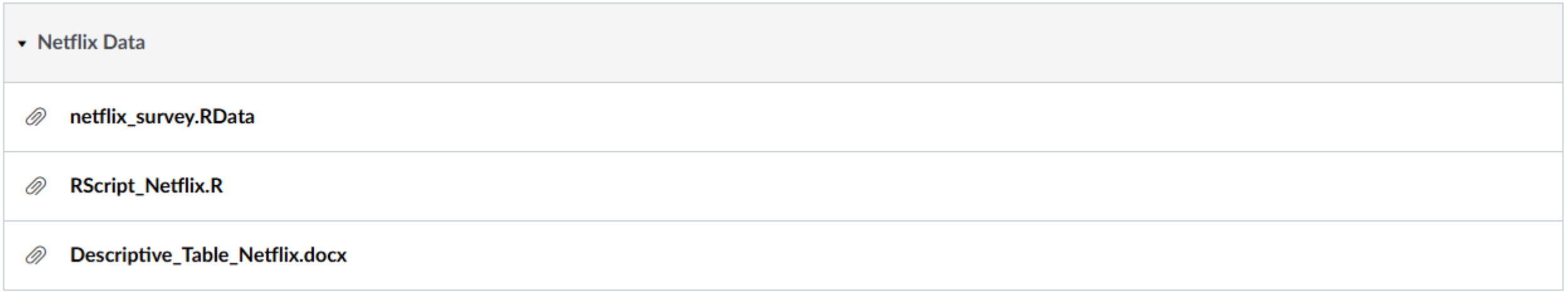
Variable	Mean (SD)	Median	Min.	Max.	Level of Measurement
Comedy Preference [a]			1	5	ordinal
Most Favorite	0.18				
Second Favorite	0.27				
Third Favorite	0.27				
Fourth Favorite	0.23				
Least Favorite	0.05				
Comedy Preference [b]	2.68	3.00	1	5	interval-ratio
Age					
Subjective Age					
Female					
Gender					
Region					
North					
East					
South					
West					
Rurality					
Urban					
Suburban					
Rural					
Subjective SES					
Political Party Affiliation					
Strongly Republican					
Republican					
Independent/Other					
Democrat					
Strongly Democrat					
Anxiety					
Minimal					
Mild					
Moderate					
Severe					
Anxiety Binary					
Depression					
Minimal					
Mild					
Moderate					
Moderately Severe					
Depression Binary					
Attention Binary					
Macro Experience?					

Notes: [a,b] we may want to treat Comedy Preference like an ordinal [a] or an interval-ratio [b] variable, so let's consider descriptive statistics for both levels of measurement.

- Now we are ready to start making a descriptive statistics table for our Netflix survey data
  - I made some executive decisions on variable construction, but we can revisit later if you want
- I started us off by providing a skeleton and summarizing our DV
  - Comedy Preference
- Let's see how this goes working in our groups

I'll check in toward the end of class today, around 1:30pm

Summarizing descriptive statistics: Netflix



- everything you need is in our Netflix Data module on Canvas
  - Try a variable or two as a team, then divide the remaining variables amongst your group as individuals or pairs/trios.
    - You're a team. Help each other when stuck. If all stuck, ask me.

It may be helpful to keep a working draft of your own RScript and descriptive statistics table. Then share programming and results with one another, maybe even check each others work. Make sure to provide kind and constructive feedback. Each team should have a completed descriptive statistics table when we are finished.