Codebook

Overview of Data

This steak data was created in hopes of assessing the various factors affecting individual steak preferences. There are 5 different temperatures to cook steak at; rare, medium rare, medium, medium well, and well. The temperature preference for each is different but it is often believed that individuals avoid rare steak is often because they want to avoid food born illnesses, which can be associated with risk. Therefore, what this data hope to achieve is a correlation between risk and steak preferences. Accordingly, it mostly contains questions to assess risk. Given that it is a poll, most of the data are character variables with a few factor variables.

Sources and methodology

The dataset was extracted from SurveyMonkey Audience polls and published by FiveThirtyEight. The polling methodology used allowed researchers to gather information directly from the source. Both websites are trusted sources and well-fitted to provide quality data needed to carry out analyses. Furthermore, FiveThirtyEight also uses hard data and statistical analysis to tell stories about politics, sports, economics, and culture. With this data, they also wrote an article to asses their findings. The poll used to gather the data is composed of 14 questions, with 7 questions meant to measure the participants' willingness to take a risk.

Note on missing values

Missing values are questions that were left unanswered by the participants. Given that it is a poll, there is a higher chance that there are missing answers. In this data set, there

are missing values in several of the variables for unknown reasons. However, we can assume that the participant overlooked or chose not to answer.

Itemized presentation of variables

Respondent ID

- Variable name: RespondentID
- Variable type: Numeric
- Recorded ID number of the respondent.

Consider the following hypothetical situations: In Lottery A, you have a 50% chance of success, with a payout of $100. In Lottery B, you have a 90% chance of success, with a payout of $20. Assuming you have $10 to bet, would you play Lottery A or Lottery B?

- Variable name: lottery
- Variable type: character
- Respondent's lottery risk taking willingness. Coded as “Lottery A” or “Lottery B”

Do you ever smoke cigarettes?

- Variable name: smoke
- Variable type: character
- Respondent's smoking status coded as “Yes” or “No”.

Do you ever drink alcohol?

- Variable name: drink
- Variable type: character
- Respondent's drinking status coded as “Yes” or “No”.

Do you ever gamble?

- Variable name: gamble
- Variable type: character
- Respondent's gambling status coded as “Yes” or “No”.

Have you ever been skydiving?

- Variable name: skydive
- Variable type: character
- Respondent's skydiving past coded as “Yes” or “No”.

Do you ever drive above the speed limit?

- Variable name: above_speed
- Variable type: character
- Respondent's driving habits coded as “Yes” or “No”.

Have you ever cheated on your significant other?

- Variable name: cheat
- Variable type: character
- Respondent's cheating habit coded as “Yes” or “No”.

How do you like your steak prepared?

- Variable name: steak_cook
- Variable type: factor
- Respondent's steak preference

steak_cook	n	prop
Well	36	0.06
Medium well	75	0.14
Medium	132	0.24
Medium rare	66	0.3
Rare	23	0.04
NA	118	0.21

How do you like your steak prepared? (new mutated factor variable)

- Variable name: cook
- Variable type: factor
- Respondent's steak preference coded as numbers 1 throught 5, 1 for rare and 5 for well.

steak_cook	n	prop
5	36	0.06
4	75	0.14
3	132	0.24
2	66	0.3
1	23	0.04
-9	118	0.21

Do you eat steak?

- Variable name: steak
- Variable type: factor
- Respondent's willingness to eat steak, coded as “Yes” and “No”.

Steak	n	prop
No	109	0
Yes	430	0.78
NA	11	0.02

Do you eat steak? (new mutated factor variable)

- Variable name: steak1
- Variable type: factor
- Respondent's willingness to eat steak coded as numbers. 0 for no, 1 for yes and -9 for NA.

Steak	n	prop
0	109	0
1	430	0.78

-9	11	0.02

Gender

- Variable name: gender
- Variable type: character
- Respondent's gender coded as “Male” and “Female”.

Respondent age

- Variable name: age
- Variable type: Numeric
- Respondent's age in years at time of interview.

Min	**Medium**	**Mean**	**Standard Dev**	**Max**
18	45	39.2	15.4	60

Household Income

- Variable name: income
- Variable type: character
- Respondent's household income

Min	**Medium**	**Mean**	**Standard Dev**	**Max**
$0	$50,000	$60,988	$44,609	$150,000

Location (Census Region)

- Variable name: location
- Variable type: character
- Respondent's location

Reference

F. (2014, January 1). *Steak Survey*. Github. https://github.com/fivethirtyeight/data/tree/master/steak-survey