

PHP 1510 Midterm Exam 2019

Due: October 29, 2019 at 11:59pm

Name: _____

Instructions

- Write your code and answers under each of the following question.
- You may use course notes, labs and homeworks for this exam.
- You may not ask any person for any help.
- If you are found copying code or responses from anyone, you will receive a 0 for the entire exam.
- **Explain your reasoning.** The final answer is not as important as the process.
- All interpretations must be in context to the original problem including units.
- **All answers must be in complete sentences**
- Download the Rmd document to run with your code such.
- Please submit to: ***PHP 1510: Canvas***

Scoring

Problem	Point Value	Problem Grade
1	3 ‘	‘ _____ ‘
2	3 ‘	‘ _____ ‘
3	4 ‘	‘ _____ ‘
4	4 ‘	‘ _____ ‘
5	6 ‘	‘ _____ ‘
6	6 ‘	‘ _____ ‘
7	6 ‘	‘ _____ ‘
8	6 ‘	‘ _____ ‘
9	12	----- ‘
10	20 ‘	_____ ‘
		‘ ‘
	Total	70

The Data

The Behavioral Risk Factor Surveillance System (BRFSS) is a collaborative project between all of the states in the United States (US) and participating US territories and the Centers for Disease Control and Prevention (CDC). The BRFSS is administered and supported by CDC's Population Health Surveillance Branch, under the Division of Population Health at the National Center for Chronic Disease Prevention and Health Promotion. The BRFSS is an ongoing surveillance system designed to measure behavioral risk factors for the noninstitutionalized adult population (aged 18 years of age and older) residing in the United States. The BRFSS was initiated in 1984, with 15 states collecting surveillance data on risk behaviors through monthly telephone interviews. Over time, the number of states participating in the survey increased, and by 2001, 50 states, the District of Columbia, Puerto Rico, Guam, and the US Virgin Islands were participating in the BRFSS. Today, all 50 states, the District of Columbia, Puerto Rico, and Guam collect data annually; American Samoa, the Federated States of Micronesia, and Palau collect survey data over a limited point-in-time (usually 1 to 3 months). In this document, the term state is used to refer to all areas participating in the BRFSS, including the District of Columbia, Guam, and the Commonwealth of Puerto Rico.

Factors assessed by the BRFSS in 2014 include tobacco use, HIV/AIDS knowledge and prevention, exercise, immunization, health status, healthy days health-related quality of life, health care access, inadequate sleep, chronic health conditions, alcohol consumption, oral health, falls, drinking and driving, cancer screenings (including breast, cervical, prostate, colorectal cancers), and seatbelt use. Since 2011, the BRFSS conducts both landline telephone- and cellular telephone-based surveys. In conducting the BRFSS landline telephone survey, interviewers collect data from a randomly selected adult in a household. In conducting the cellular telephone version of the BRFSS questionnaire, interviewers collect data from an adult who participates by using a cellular telephone and resides in a private residence or college housing.

Variable Desc	Description
genhlth	Excellent Very Good Good Fair Poor
genhlth_bin	Excellent/Very Good/Good Fair/Poor
Unhealthy.days	0-30
menthlth	0-30
poorhlth	0-30
imrace	White, Non-Hispanic Black, Non-Hispanic Asian, Non-Hispanic AIAN, Non-Hispanic Hispanic Other Race, Non-Hispanic
insurance	yes No
trnsghdr	Yes, mtf Yes, ftm Yes, non-conforming No
trnsghdr_bin	Yes No
sexorient	Heterosexual Homosexual Bisexual Other
sexorient_bin	Heterosexual Other
lstisfy	Very Satisfied Satisfied Disatisfied Very Disatisfied
lsatisfy_bin	Satisfied Disatisfied
emtsuprt	Always Usually Sometimes Rarely Never
emtsuprt_bin	Always/Usually Sometimes/Rarely/Never
medcost	Yes No

Conceptual Questions

1. **(3 points)** What does the expectation tell you?
2. **(3 points)** What does the variance tell you?
3. **(4 points)** What is the story behind Bernoulli Data?
4. **(4 points)** What are some traits of the normal distribution?

Data Questions

Our goal here will to be consider the days that stress, depression, and problems with emotions caused days where mental health was not good. We will be addressing this specifically in context to the transgender experience. The reason for this being that we do see different rates of depression, suicide and other outcomes associated with the transgender experience. It is important to begin to understand what the cause might be.

- Download the file ***brfss.rda***

Click the link or go to: <https://drive.google.com/file/d/188YvZMXQxegZY5oXDoeTa8Lkh8hGbzx4/view?usp=sharing>

5. **(6 points)** Plot and Describe the Distribution of poor mental health days.

6. **(6 points)** Plot and Describe the Distribution of Race.

7. **(6 points)** Plot and Describe the Distribution of Transgender Binary.

8. **(6 points)** Plot and Describe the Distribution of General Health (not binary).

9. **(12 points)** Display graphs of variables that have relationships with poor mental health days. (**Note:** Look at the notes with general health to see how to combine multiple plots in a larger grid image. (lect 13 slide 58))

10. **20 points** We have been exploring the outcome of days off due to mental health. Papers have been published suggesting that there is a relationship between the transgender experience and mental health outcomes.

Health related to the transgender experience in general has not been well researched. This is an attempt to get everyone to be thinking about the number of days in which stress, depression and other emotions may have led to poor mental health days.

Do there appear to be relationships between the transgender experience and the number of days in which stress, depression and emotions led to poor mental health?

Remember: we have 3 things to consider with a hypothesis.

Note: This problem is much more complex than we have data to consider, this is one starting place with data which is available to us.

Note: You will need to use transgender binary to make this work like the other examples.