Response Bias Project

AP Statistics 2018-2019
**TASK:** In this project, you (and a partner) will design and conduct an experiment to investigate the effects of response bias in surveys. You may choose the topic for your surveys, but you must design your experiment so that it can answer one of the following questions:

- Can the wording of a question create response bias?
- Do the characteristics of the interviewer create response bias?
- Does anonymity change the responses to sensitive questions?
- Does providing additional information (facts/quote/photos) change the response?
Type a proposal describing the design of your experiment. Be sure to include:

a) Your chosen topic and why you chose this topic. Why does this topic interest you? Why should we “care”?

b) Which of the above general response bias questions (bullets 1 – 4) you will try to answer with your project.

c) What your questions will be and how they will be asked. Be very clear in demonstrating how your chosen surveys pertain to one of the four bulleted response bias questions above.

d) Make a hypothesis on what results you expect to get (which direction the bias will be).

e) A detailed description of how you will obtain your subjects (minimum of 50 subjects, but more is better). Your plan must be practical and follow the statistical protocol we have practiced in Section 4.1. Provide a clear description of your sampling method (i.e. simple random sample, cluster sample, stratified random sample… Hat method? Random number generator? Table D?).

f) A clear explanation of how you will implement your design once you reach your subjects (will you be handing out surveys? Are you interviewing people individually? When? Where? How?).

g) Precautions you will take to collect data ethically. (Section 4.3)
Please note: if I believe your project will not induce bias, you must redo your proposal on your topic/questions and resubmit for approval.
The whole point of this project is to induce bias, so make sure your questions and project design will get bias.
Proposal Submission

- Proposal Due Date: **Wednesday, October 31st, 2018 by 11:59pm.**

- Share to avoinea@sjusd.org using **Google Docs.**

- Use your **school email account**, not your chickentacofreak123@yahoo.com email...

- Allow me the ability to **comment.**

- You may begin data collection as soon as you receive an approval from me on your proposal (next week sometime).

- Proposal is worth **15 points of your total project grade** – take it seriously and submit quality work.

- Treat this like your chance as a researcher to request funding from a potential research investor, or to seek approval from the Institutional Review Board (IRB).
Once you’re approved...

**Step 2. Carry out the experiment** once your teacher has approved your design. Take pictures during the data collection process for evidence for your presentation. **Record** your data in a table.

*Please note: if you do not get the bias you expected once you carry out the experiment, you must redo your proposal on your topic/questions and resubmit for approval! And no, you cannot “fake it”. Don’t even try. 😜*

**Step 3. Analyze your data.** What conclusion do you draw? Provide appropriate graphical and numerical evidence to support your answer.

**Step 4. Prepare a final report** that includes the data you collected, your analysis from Step 3, and a discussion of any problems you encountered and how you dealt with them.
Grading Breakdown

Proposal Plan: 10 – 15 points
Final Report: 50 points
Presentation Poster: 10 – 15 points

Total: 75 points in Test/Quiz category

Due Friday November 30th, 2018
# AP Statistics – Project 1 (Response Bias Project)

## Scoring Rubric

<table>
<thead>
<tr>
<th>Category</th>
<th>Complete</th>
<th>Substantial</th>
<th>Developing</th>
<th>Minimal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proposal</strong></td>
<td>* Chosen topic and type of bias is clearly stated</td>
<td>* Chosen topic and type of bias is stated</td>
<td>* Chosen topic and type of bias is stated</td>
<td>* Chosen topic and/or type of bias is stated</td>
</tr>
<tr>
<td></td>
<td>* Detailed description of how subjects will be obtained is included</td>
<td>* Description of how subjects will be obtained is included</td>
<td>* Description of how subjects will be obtained is included</td>
<td>* What questions will be and/or how they will be asked is stated</td>
</tr>
<tr>
<td></td>
<td>* What questions will be and how they will be asked is clearly stated</td>
<td>* What questions will be and how they will be asked is clearly stated</td>
<td>* What questions will be and/or how they will be asked is clearly stated</td>
<td>* What questions will be and/or how they will be asked is stated</td>
</tr>
<tr>
<td></td>
<td>* A clear explanation of how design will be implement is included</td>
<td>* An explanation of how design will be implement is included</td>
<td>* An explanation of how design will be implement is included</td>
<td>* An explanation of design is included</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td>* Describes the context of the research</td>
<td>* Introduces the context of the research and has a specific question of interest</td>
<td>* Introduces the context of the research and has a specific question of interest</td>
<td>* Briefly describes the context of the research</td>
</tr>
<tr>
<td></td>
<td>* Has a clearly stated question of interest</td>
<td>* Suggests hypothesis or has appropriate difficulty</td>
<td>* Suggests hypothesis or has appropriate difficulty</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Provides a hypothesis about the answer to the question of interest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Question of interest is of appropriate difficulty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data Collection</strong></td>
<td>* Method of data collection is clearly described</td>
<td>* Method of data collection is clearly described</td>
<td>* Method of data collection is clearly described</td>
<td>* Some evidence of data collection</td>
</tr>
<tr>
<td></td>
<td>* Includes appropriate randomization</td>
<td>* Some effort is made to incorporate principles of good data collection</td>
<td>* Some effort is made to incorporate principles of good data collection</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Describes efforts to reduce bias, variability, confounding</td>
<td>* Quantity of data collected is appropriate</td>
<td>* Quantity of data collected is appropriate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Quantity of data collected is appropriate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Graphs and Summary Statistics</strong></td>
<td>* Raw data is included in a two-way table (categorical) or in lists (quantitative)</td>
<td>* Appropriate graphs are included</td>
<td>* Appropriate graphs are included</td>
<td>* Graphs or summary statistics are included</td>
</tr>
<tr>
<td></td>
<td>* Appropriate graphs are included</td>
<td>* Graphs are neat, clearly labeled, and easy to compare</td>
<td>* Graphs are neat, clearly labeled, and easy to compare</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Graphs are neat, easy to compare, and clearly labeled, including clear identification of treatments</td>
<td>* Appropriate summary statistics are included</td>
<td>* Appropriate summary statistics or raw data are included</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Appropriate summary statistics are included</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Conclusions</strong></td>
<td>* Uses the results of the study to correctly answer question of interest</td>
<td>* Makes a correct conclusion</td>
<td>* Makes a partially correct conclusion</td>
<td>* Makes a conclusion</td>
</tr>
<tr>
<td></td>
<td>* Discusses what inferences are appropriate based on study design</td>
<td>* Discusses what inferences are appropriate or shows good evidence of critical reflection</td>
<td>* Discusses what inferences are appropriate or shows good evidence of critical reflection</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* Shows good evidence of critical reflection (discusses possible errors, limitations, etc.)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sample Questions

Please do not copy.
Seeimg Red

Question 1: “Do you find women more attractive when they are wearing red” (interviewer was wearing red)

Question 2: “Do you find females who wear red more attractive?” (interviewer not wearing red)

More males preferred females who wear red when asked the question by the interviewer wearing red.

All questions asked to females in this case.
“Do you prefer males with blue eyes or brown eyes?”

All females were asked the same question (whether they prefer blue eyes or brown eyes), but they were asked either by a male interviewer with blue eyes or a male interviewer with brown eyes. When asked by the blue-eyed interviewer, more females preferred blue eyes when asked than by the brown-eyed interviewer.
Females were asked how much they weigh anonymously and in person. The average weight was higher for the anonymous responses.
Aid for the Addicted?

“Should there be more rehabilitation funding for drug-addicted people?”

All subjects were asked the same question, but half of the people were randomly shown a picture of a book cover with a distressed mother finding her apparently overdosed son. When the photo was shown with a verbal rendition of the mother’s story, more people supported increased funding for drug addicts.
Students were asked a series of two questions, in random order. The questions were, “How hungry are you on a scale from 1 to 10?”, and “How long has it been since your last meal?” When students were asked about their last meal first, their average hunger rating was higher.
Sample Posters

(Not necessarily in any apparent order or the best/worst quality)
ANONYMOUS!

DATA COLLECTION

HAVE YOU SMOKED?

HYPOTHESIS

When I began this experiment I predicted that more people would respond yes, that they have smoked when they were asked anonymously.

Introduction

I chose to ask this question because it is a sensitive subject. Most people do not want to admit to underestimating smoking if they have to put their names on it. In my survey, I asked 50 students if they had ever smoked a cigarette before. 25 of the students were asked to put their real names on the survey, and 25 were asked to keep their name anonymous.

RESULTS

At the end of the survey, I concluded that the people who were surveyed anonymously were more likely to admit to smoking than those who had to write their names. There was a 10% increase in yes responses when they were anonymous.
Does Adding a Fact Affect the Answer?

In this survey, I asked 50 people how many hours they spend on their cell phones texting & calling in a day. 25 people got only the question, and the other 25, I added a fact along with it. I chose this because I was interested in what the results would be and because pretty much everyone has some sort of cell phone.

Hypothesis

By adding the fact to the original question, it will result in them picking a lower number.

Data Collection

I made 2 different survey sheets: one with the question, and one with the question and fact. I made 25 of each, 50 in all, and questioned 50 people, tech and DHS combined. The same person was never asked both questions.

Conclusion

It seems that even with the fact, it doesn't make a difference. The fact surprisingly increases their answers. The mistake and what to do different teenagers aren't pressed by the idea of harm coming to them in their later years. Perhaps it would be different to ask about something they aren't so dependent on.
Does anonymity change the response to sensitive questions? Do you drink alcohol, yes or no?

**Anonymous**

Introduction

Austin and I asked people whether they drink alcohol or not. We first asked the question anonymously. After we gathered those results, we asked 30 face to face. We picked this question because it was the perfect sensitive subject to go with the question we picked.

**Not Anonymous**

Conclusion

At the end we realized that at Drumright High School people aren’t afraid of saying whether they drink or not. Both the anonymous and not anonymous results are very close to each other.
Do students support the decision to terminate DACA?

<table>
<thead>
<tr>
<th></th>
<th>Non-Bias</th>
<th>Bias</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>No</td>
<td>19</td>
<td>12</td>
<td>31</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>13</td>
<td>38</td>
</tr>
</tbody>
</table>

With the recent controversy around the Trump administration's actions, we decided to conduct a survey to determine the opinions of students on the matter. The survey was conducted among students at ABC High School, with a total of 50 participants. The results are as follows:

- 7 students support the decision to terminate DACA.
- 31 students do not support the decision.

This survey was conducted anonymously, and we do not share the personal details of the participants. The results are based on the opinions of the students surveyed. We believe that further research is needed to understand the implications of these decisions.
How Many Hours a day are Teens REALLY on their phones?

Initial Question: Does wording of the question affect response bias?

Hypothesis: The wording of the question in question 6 is showing that the average teenager spends 9 hours 30 minutes on their phones. The wording of the question in question 2, asking how much time students say they spend on their phones, will vary.

Procedure:
1. Randomly choose 4 history classes to sample, clustering classes by grade level.
2. Sample one freshman class, one sophomore class, one junior class, and one senior class.
3. Send the test email to sample students in each class by having equal amounts of G1 and G2.
4. Have every student within each class randomly choose 1 question from the list and answer it.
5. After all the surveys are returned, count the responses and graph the data for G1 and G2 by grade level and level students.

Data Analysis:
- Question 1
- Question 2

Conclusion:
Based on this survey, I can conclude that wording of the question does have a response bias. In this case, I believe that the addition of the last sentence of time that students say they spend on their phones is biased to the students who responded G2. The average time that students with response G2 spent on their phones was 3.5 hours more than the students who responded G1. Despite surveying 20 students, these surveys would need to be conducted before making a definitive conclusion that the wording of question 6 really affected the results.
Does the Wording of a Question Lead to Bias?

Introduction

We are trying to answer the question of whether or not the wording of a question will lead to bias. The topic we chose to work on because we were curious to see how many people would actually admit to not being healthy. Our two questions are: do you stay healthy by exercising frequently? and do you work out frequently? We expected that the majority of students who were asked the first question would choose yes, and the majority of the students who were asked the second question would choose no. We believed this because the students who answered the first question would feel bad about themselves and their health if they said no because it's implying that they aren't healthy. For the second question, we believed that the majority of students would say no because it's not mentioning anything about their health.

Graphs and Tables

<table>
<thead>
<tr>
<th>Question</th>
<th>Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>25</td>
</tr>
<tr>
<td>No</td>
<td>15</td>
</tr>
</tbody>
</table>

Data Collection

- Place 30 slips of each question into equal size cups of paper making a total of 60 slips.
- Pick 15 slips out of the hat and mix it in well.
- Go to the first class (unclassified) and have each student peek out of the hat to answer. Repeat this same step for the second class (art). Use the color of the slips.
- This idea chance decides which student answers which question.
- Collect the slips and write down the results of each question.

Conclusion

According to our data, our hypothesis was correct. The majority of students that were asked question 1 answered "yes" and the majority that were asked question 2 answered "no." This proves that the wording of a question can lead to bias. The fact that the first question seemed healthy makes them feel guilty or embarrassed to say no. Meanwhile, the students who answered the second question didn't feel so inclined to say yes because it wasn't questioning anything about their health.

If you look at Table 1 and Graph 1, you'll see the student's answers for both questions by count. It illustrates that for question 1, 25 students answered yes and 15 answered no. That means 66.67% more students answered yes to question 1. For question 2, 14 students answered yes and 16 answered no. This means 30% more students said no. If you look at Table 2 and Graph 2 you'll also see the same results but as percents. 63.33% of 30 students said yes to the first question while 33.33% of another 30 students said no to the second question.

Both of the questions received the results that were expected, and this data, therefore, shows that the wording of a question can lead to bias. We were able to secretly "manipulate" the students into answering how we wanted them to just by wording our question a certain way.
Teen Driving
How does wording of the question create response bias?

Introduction
Driving is a top concern among senior high school students. In our project, we wanted to discover if adding a piece of hard evidence about teen driving would make teenagers think of themselves as worse drivers. In addition, we wanted to create more awareness of teen driving in order to see if there would be a change in teen perceptions. We chose this topic because we knew many senior students who have been in car accidents and are very aware of the problem. Because of this, we thought it would be important to create more awareness amongst students who want to be safe drivers. As a result, we wanted to see if the students would take action on the issue. In addition to the main piece of evidence, we also gathered data from the same students who had been in car accidents and were aware of the problem. We wanted to see if the students would take action on the issue. In addition to the main piece of evidence, we also gathered data from the same students who had been in car accidents and were aware of the problem. We wanted to see if the students would take action on the issue.

Data & Analysis
- Test Collection Method
  - Students were divided into two groups: Group A and Group B.
  - Group A was given the test with wording that made it clear that the question was about teen driving.
  - Group B was given the test without any indication that the question was about teen driving.

- Results
  - Group A had a significantly higher rate of correct answers than Group B.
  - This suggests that wording of the question can create response bias.

Conclusion
Based on our results, we can conclude that wording of the question can create response bias. Students who are aware of the problem of teen driving are more likely to give the correct answer. However, students who are not aware of the problem are more likely to give the wrong answer. We recommend that educators make a clear distinction between teen driving and other types of driving in the future.
To Kneel or not to Kneel?

Introduction

The main idea behind this project was to investigate if the wording of a question could create response bias. We considered the questions: “Do you think that NFL players should be allowed to kneel in protest during the National Anthem?” and “The 1st Amendment of the United States Constitution guarantees citizens the right to peacefully protest and petition the government, as well as the freedom of speech. Do you think that NFL players should be allowed to kneel in protest during the National Anthem?” The first question being the one without any additional information, and the second question being the one with additional information. Our intention was for the first question to provide a baseline response to compare the responses to the second question and determine if response bias was present. We predicted that more people would claim that they would support the athletes if they were given the additional information because the mention of the United States Constitution and the 1st Amendment will most likely result in the survey takers thinking of the situation in a legal sense, rather than personal belief.

Data Collection

For the second portion of the project, we decided to split up. Using a cluster sample, I surveyed Mr. Milewski’s 5th and 6th period classes. Then I surveyed 40 students total between the classes. We utilized the box method to ensure randomness by putting strips of paper in a bag with either the first question or the question with additional information printed on them. Making sure there was an equal amount of both questions within the bag for each class. Before allowing the students to randomly select a question, we gave the students the option of what the survey was about and thoroughly explained the strips of paper. Then we walked around the classrooms, holding the bag and having each student select one question from the bag without looking. Where all of the students were able to choose the strips and completed the data. None of the students that already knew the answers were allowed to do the survey. We then distributed the survey, allowing the students to choose the survey number. As a result, we would collect data on 50% of high school students. Overall, the majority of our students found the survey very interesting and informative.

Graphs and Summary

Statistics

[Graphs showing data]

Conclusion

Through this project, we were able to discover whether wording bias affects the responses of students in a yes or no question. According to our data, students who answered the question without additional information regarding the question were more likely to respond in the affirmative. 50% of students who answered the question without additional information regarding the question never responded. However, when additional information was added into the question, the percent of students who responded was 80%. This decrease in students who believe that 50% of players should be able to kneel is huge, as it significantly cuts into the number of students who were previously comfortable with the question. 80% of students would be less likely to change their opinion if additional wording was added to the question.
Response Bias Project

Introduction

We want to determine if wording of the question causes response bias. Two questions were asked: “How much time do you spend watching TV per day?” and “How much time do you watch TV per day?” Participants who answered the “watch” question were asked to watch less TV because they didn’t want to admit to watching too much. Therefore, we expect the mean and median time spent watching TV to be lower than the mean and median time spent watching TV answered by those who answered the “watch” question.

Does the wording of the question cause bias?

Data Collection Process

1. Using a cluster sample, 2 of Mr. Wind’s regular classes were chosen. Every student in each class was surveyed, creating 36 participants.
2. Each question was printed on 30 slips of paper. The 90 total slips were identical. They were given to participants randomly using the hat method.
3. 20 students answered the question with the word “watch” and 27 students answered the question with the word “watch.”


Graphs and Data

<table>
<thead>
<tr>
<th>Time Spent Watching TV (hours)</th>
<th>Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

Number of hours spent watching TV

Conclusion

Subjects who answered the “watch” question watched a mean of 1.775 hours of TV per day with a median of 2 hours per day and a standard deviation of 1.04 hours per day. Subjects who answered the “watch” question watched a mean of 2.66 hours of TV per day with a median of 3 hours per day and a standard deviation of 1.12 hours. The means for the subjects who answered the “watch” question is lower by 1.37 hours and the median is one hour lower. This large of a difference per day is a statistically significant amount, meaning the difference cannot be accounted for by random chance, and therefore is likely due to the wording of the question, since we made no effort to control all variables for “watch” or “watch.” The hypothesis that the wording of the question would influence people answering the “watch” question to report watching less TV was correct.
How many people play sports competitively?

Introduction:
We decided to ask the question, “Do the number of competitive sports lead to increased stress?” Due to our interest in various sports, we wanted to conduct a study involving the types of sports that people play. We decided to ask the question “Do you play sports?” (33) and “Do you play sports competitively?” (27) with the goal of figuring out how many people actually play sports competitively rather than for fun or for the enjoyment.

Data Collection:
We chose to sample the Advanced Multimedia and Multimedia classes in Periods 2 and 3 because there is a mix of all the grades in these classes. We printed out copies of papers with one question per slip. Next, we made sure we had an equal amount of questions for the classes before we began sampling. The survey was done in a way that each person had an equal opportunity of being surveyed. Finally, we used the bar method to collect our data once we began the sampling process. We put the slips of paper in a box, mixed them around thoroughly, and walked around the classroom, having each person draw from the bar to select a question.

Data:

<table>
<thead>
<tr>
<th>Play Sports</th>
<th>Play Sports Competitively (Outside of School)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>63%</td>
<td>42.3%</td>
</tr>
<tr>
<td>27%</td>
<td>51.7%</td>
</tr>
</tbody>
</table>

Sports vs Competitive Sports

Questions:

Amount of People

Conclusion:
After analyzing the results from sampling, we determined that our hypothesis that more people would claim that they play sports rather than claim that they play sports competitively was correct. Only 63% of the people we surveyed claimed to not play sports competitively, whereas 42.3% of the people who responded claimed to play sports competitively. These percentages may be skewed because we did not account for those who answered “Yes” to the latter. For the purpose of this study, we should have asked the students both questions in order to get the most accurate data we could have possibly received.
Please email avoinea@sjusd.org with questions.

Good luck, team!