Development and Implementation of WISE (Workshop to Inspire Statistical Excellence) to Recruit Future Statisticians

Michelle L. DePoy Smith, Amanda Ellis, Shane P. Redmond

Eastern Kentucky University, 521 Lancaster Ave, Richmond, KY 40475

Abstract

Faculty at Eastern Kentucky University hosted a one-day workshop in April for current high school students who were enrolled in an advanced placement statistics course. The pilot program consisted of 53 students from two local high schools to participate in the day-long activities.

The goals of the workshop were to use group learning and competition to 1) prepare students for an end of course exam, 2) perform data analysis and exploration, and 3) recruit top high school students to major in mathematics and/or statistics. Two sessions consisted of using active learning to cover sampling distributions and one-sample hypothesis tests. Two additional sessions consisted of reviewing and mock exam questions for two-sample inference and regression. Prizes were given to the students with the best solutions for each session and overall best solutions. In addition, faculty and current undergraduate students discussed careers and opportunities in statistics.

We hope to expand this workshop to all the high schools within the Eastern Kentucky University service region next year and other surrounding counties. Also, we hope that this workshop will serve as a model for a future high school calculus workshop.

The goals of this paper are to present the workshop curriculum, follow-up survey results, and lessons learned from the workshop.

Key Words: Recruitment, Active Learning, High School Students, Introductory Statistics, Competition, Exam Preparation

1. Introduction

The number students enrolling in AP Statistics in our service region has increased in recent years. As a consequence, fewer students are taking our College Algebra based Applied Statistics Course. Historically, this is the course where our department recruits most of our statistics majors. Due to more students being exposed to statistics for the first time in high school, we decided to focus some of our recruiting efforts within the high schools.

In the spring of 2019, we invited teachers to bring their students enrolled in AP statistics from two local high schools to participate in a workshop designed to help them prepare for the upcoming end of year exam. Our aim in offering this workshop was two-fold we wanted to support students and teachers in their statistical education as well as introduce them to our university, department and faculty to aid in recruitment efforts.

2. Workshop Details

In order to make the workshop more accessible for students and faculty, the workshop was designed to be completed within a school day and allowed for one hour of travel time before and after the workshop. In addition, funding was secured from the College of Science to provide lunch for all students and teachers.

One of the main goals of the workshop was recruitment of students to Eastern Kentucky University (EKU) in general and specifically to the Department of Mathematics and Statistics. Each student was given an EKU t-shirt and cinch bag, which were obtained free of charge from the Director of Brand Management at EKU. In addition, we designated one of the sessions for a presentation by an EKU admission specialist.

The workshop consisted of four forty-minute sessions designed for curriculum. Two of the sessions consisted of mock exam problems. One of the co-authors, who has 11 years of experience as an AP Statistics Reader, hosted a 25-minute session providing tips for taking the AP Exam. In addition, we designated a session for careers in statistics and invited some undergraduate statistics students to share their experiences as a statistics major. The day ended with a student feedback survey.

The mock exam problems were graded by EKU faculty using an AP statistics exam style rubric. An EKU statistics faculty member visited each school to return the graded problems, give feedback, present awards for best solutions, and to address any additional questions or concerns.

2.1 Workshop Schedule

The schedule we used for our workshop is given below.

WISE Workshop - April 12, 2019

- 9:00 AM – Students arrived in Science Lobby – pick up name tag, t-shirt, & bag
- 9:10 AM – 9:50 AM – Session 1 – Sampling Distributions – Gettysburg Activity – Active Learning
- 10:00 AM – 10:40 AM – Session 2 – Inference – Psychic Activity - Active Learning
- 10:50 AM – 11:30 AM – Session 3 – Inference Review/Inference Mock Problem
- 11:30 AM – 12:30 PM - Lunch at Case Dining (No cost to students)
- 12:40 PM – 1:20 PM – Session 4 – Regression Review/Regression Mock Problem
- 1:20 PM – Tips for AP Statistics Exam
- 1:25 PM – Careers in Statistics
- 1:35 PM – Statistics Majors share experiences
- 1:40 PM – EKU Admissions
- 1:50 PM – Follow-Up Survey

2.2 Workshop Curriculum

The curriculum consisted of both active learning components and mock testing environments.

Both active learning components were adapted from activities located in Lock, 2016. The psychic activity was used to reinforce and review one–sample inference procedures (Figure 1). The Gettysburg Address Activity was used to review and reinforce the concepts of sampling distributions (Figure 2).

The authors reviewed many years of past AP statistics free response exam questions in order to design similar mock questions. We chose two free response questions covering two-sample statistical inference (Figure 3) and regression analysis (Figure 4) due to the frequency of these types of problems on the exam. The session began with a review question that the students and faculty member completed together. Then the students were given 20 minutes to complete a similar mock exam problem on their own. The students were permitted to use their calculator and the official AP Statistics Exam Formula Sheet.

Exercises from the mock testing environment were graded utilizing a grading rubric (Figures 5 & 6) developed by an EKU faculty member with 11 years experience as an AP Statistics exam reader. The faculty were trained using rubric before grading the exams.

A summary of the scored mocked problems are listed in Table 1. The results were very encouraging with 68% (36/53) students scoring a 2 or higher on the mock inference question and 69% (35/51) students scoring a 2 or higher on the mock regression question.

Table 1: Mock Testing Results.

	Number of Scores				
	4	3	2	1	0
Inference	0	13	23	5	12
Regression	0	3	32	10	6

Do you have ESP?

Today we are going to try to answer the question: "Is there such a thing as extrasensory perception (ESP) or a "sixth sense"?

One way to test for ESP is with Zener cards:

Subjects draw a card at random and telepathically communicate this to someone who then guesses the symbol.

We have developed our own set of cards and are going to perform a study to see if anyone in this room has ESP.[1] We will be using the following set of cards ...

Let p denote the true percentage of correct guesses a person can make. There are three cards with three different symbols, so if there is no such thing as ESP, $p = 1/3$ or $p=0.333$ of all guesses are expected to be correct. We are going to conduct trials to see if we can find evidence that a person with ESP exist in this room i.e. someone has a greater than 1/3 chance of getting the card right.

Set up your hypothesis:

Before collecting data you need to set up your hypothesis that you want to test. For the ESP experiment define an appropriate null and alternative hypothesis.

Collect your Data:

Break into pairs. If there is an odd number of people in the room, one group will have 3 people. Each pair of people will need a set of cards and a score sheet. Select a person in the group to be the guesser and someone to read the card. If you have a third person they will be the scorekeeper, otherwise the reader is the scorekeeper. Using the Zener cards have the guesser guess 36 times in a row recording whether they are correct each time. Once you are done switch roles.

Calculate your Statistic: Using the data, give the symbol and the value for your statistic for the proportion of times you guessed correctly.

Conduct the appropriate hypothesis test at the α=0.05 to determine if you have ESP:

Potential Issues

- Based on your conclusion could a Type I or Type II error have potentially occurred and why?
- In this room, this same hypothesis test was performed approximately 70 times, using what we know about multiple testing issues how many type 1 errors would you expect to occur?

---

[1] This exercise was adapted from Lock, R. H., Lock, P. F., Morgan, K. L., Lock, E. F., & Lock, D. F. (2016). *Statistics: Unlocking the power of data*. Wiley Global Education.

Figure 1: Example of active learning worksheet (hypothesis testing)

Constructing Sampling Distributions Activity

For this activity, we will consider the words in the Gettysburg Address as our population.[1] The address is as follows.

"Four score and seven years ago our fathers brought forth, on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation or any nation so conceived and so dedicated, can long endure. We are met on a great battlefield of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this. But, in a larger sense, we cannot dedicate—we cannot consecrate—we cannot hallow—this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us—that from these honored dead we take increased devotion to that cause for which they here gave the last full measure of devotion—that we here highly resolve that these dead shall not have died in vain—that this nation, under God, shall have a new birth of freedom—and that government of the people, by the people, for the people, shall not perish from the earth."

Part 1: Your goal is to find to the average word length of all of the words in the address by taking samples 10 words. In the table below write down 2 samples of 10 words that you feel resemble the overall address.

1st Sample of Words	Length

2nd Sample of Words	Length

Consider your 2 samples and answer the following:

- Calculate the mean of your first sample. Use appropriate notation.
- Calculate the mean of your 2nd sample. Use appropriate notation.
- Come to the front of the room and plot both of your mean values on the board. Pick up the additional handout and continue to part 2 until everyone else has entered their data.

Part 2: Your goal is to find to the average word length by taking random samples 10 words. Generate a set of 10 random numbers from 1 to 268 using your calculator. Write down the numbers in the first table. Repeat to get another set of random numbers for the second table. Find the words, which correspond, to the random numbers you generated.

Random Number	1st Sample of Words	Length

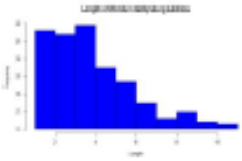
Random Number	2nd Sample of Words	Length

Consider your 2 samples and answer the following:

- Identify the mean of your first sample. Use appropriate notation.
- Identify the mean of your 2nd sample. Use appropriate notation.
- Come to the front of the room and plot both of your mean values on the board.

Part 3: There are 268 total words in the Gettysburg address with a mean length of 4.29 and standard deviation of 2.12.

1. What is the mean and standard deviation of the sampling distribution of the sample mean if we take samples of size 10.
2. Now that you know the population mean what can you conclude from the sampling distributions in parts 1 and 2?
3. Below is a histogram of the data. Explain why this histogram does not have the same shape as the plot in part 2?



[1] *This exercise was adapted from Lock, R. H., Lock, P. F., Morgan, K. L., Lock, E. F., & Lock, D. F. (2016). *Statistics: Unlocking the power of data*. Wiley Global Education.

Figure 2: Example of active learning worksheet (sampling distributions)

Two-Sample Inference Review Question

A survey of students at a large university asked, "How many tattoos do you currently have on your body?" Of the 560 males randomly selected, 15% responded that they had at least one tattoo. Of the separate 605 females randomly selected, 13% responded that they had at least one tattoo. Do the data provide convincing evidence that the proportions for males and females with at least one tattoo at this college is significantly different using $\alpha = 0.05$?

Mock Exam Two-Sample Inference Question

The Fear of Negative Evaluation (FNE) psychological test measures a person's fear about other's evaluations, distress over negative evaluations by others, and the expectation that others would evaluate one negatively. People who have a high FNE score are highly concerned with seeking social approval or avoiding disapproval by others, and may tend to avoid situations where they have to undergo evaluations.

A psychology experiment was conducted to compare FNE scores of females with bulimia nervosa to females with normal eating habits. One sample consisted of 11 students who were randomly selected from students who suffer from the eating disorder bulimia nervosa. The other sample of 11 female students were randomly selected from students with normal eating habits. Each student completed a questionnaire from which a "fear of negative evaluation" (FNE) score was produced.

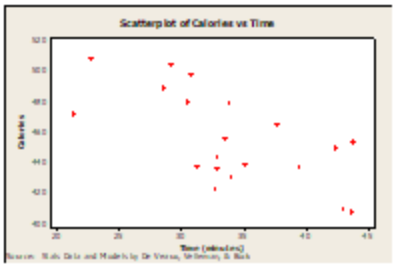
Bulimic Nervosa	21	13	10	20	25	19	16	21	24	17	17
Normal Eating Habits	13	12	16	13	8	19	23	18	11	19	7

a) At the significance level $\alpha = 0.05$, do the data provide convincing evidence that the mean FNE score is higher for bulimic nervosa females?
b) Each inference test can result in either a type I or type II error. Based on the conclusion to the hypothesis test in part a) which error is possible? Explain your reasoning.

Figure 3: Example of review questions (2 sample inference)

Regression Review Problem

Does how long children remain at the lunch table help predict how much they eat? The following scatterplot and MINITAB results are based on data on 20 toddlers observed over several months at a nursery school. "Time" is the average number of minutes a child spent at the table when lunch was served. "Calories" is the average number of calories the child consumed during lunch, calculated from careful observation on what the child ate each day.



Regression Analysis: Calories versus Time

Predictor	Coef	SE Coef	T	P
Constant	560.65	29.37	19.09	0.000
Time	-3.0771	0.8498	-3.62	0.002

S = 23.3980 R-Sq = 42.1% R-Sq(adj) = 38.9%

a) Identify and interpret in context the estimate of the slope for the least-squares regression line.
b) Identify and interpret in context the estimate of the intercept for the least squares regression line.
c) Predict the average amount of calories consumed during lunch for a child who spends on average 0.5 hours eating. Use the coefficient of determination to discuss the usefulness of this prediction?

Mock Exam Regression Problem

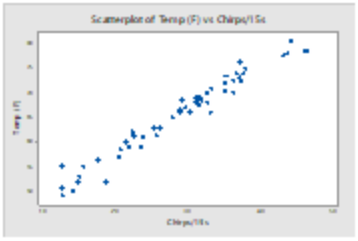
Can you use cricket chirps to predict the temperature? Data collected in 2007 from Boulder, Colorado was used to answer this question. The number of chirps per 15 second time period and temperature measured in degrees Fahrenheit at 55 different times was recorded. Use the MINITAB output below to answer the following questions. The data was obtained from the following website: https://www.gl obe.gov/explore-science/scientists-blog/archived-posts/sciblog/index.html_p=45.html

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.60975	96.09%	96.02%	95.69%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	40.025	0.744	53.79	0.000	
Chirps/15s	0.8918	0.0247	36.09	0.000	1.00



a) State the least squares regression model used to predict temperature using chirps per 15 seconds. Comment on the usefulness of this model to predict temperature.
b) Identify and interpret in context the estimate of the slope for the least-squares regression line.
c) Predict the temperature in degrees Fahrenheit if there are 40 cricket chirps in 15 seconds. How about if there are 8 chips in 15 seconds?
d) Dolbear's Law says to convert cricket chirps to degrees Fahrenheit, just count the number of chirps in 14 seconds, then add 40 to get the temperature. Do the data above support or refute Dolbear's Law? Explain.

Figure 4: Example of review questions (regression)

AP Review Scoring Rubric: Inference

Section 1 (stating the correct hypotheses, identifying the appropriate test procedure, and checking the technical conditions) is scored as follows:

Essentially correct (E) if the response correctly includes the following three components:

1. States both hypotheses correctly with proper symbols/notation
2. Identifies a two-proportion z-test as the procedure either by name or by formula
3. Verifies appropriate conditions that minimally include (a) independence and (b) checking the number of successes and failures in both groups are at least 5 and (c) does not include any incorrect conditions (such as normality)

Partially correct (P) if the response correctly completes two of the components required for an E.
Incorrect (I) if the response correctly completes one or fewer components required for an E.

Notes:

1. Stating numbers of successes and failures alone is not sufficient. The condition must at least show the smallest count of successes or failures is 5 or more.
2. The random samples condition was stated in the stem and so need not be explicitly checked.
3. Statements like "independence – yes" or "independence – check" will count as having verified that condition.

Section 2 (mechanics of calculating the test statistic and *p*-value) is scored as follows:

Essentially correct (E) if the response correctly includes the following two components:

1. The value of the test statistic
2. *p*-value or critical value

Partially correct (P) if the response correctly calculates one of the components required for an E.
Incorrect (I) if the response calculates neither component correctly.

Notes:

1. If the response makes an error in one calculation, future calculations are considered correct if they follow correctly from the initial miscalculation.
2. Work does not have to be shown for calculations of the test statistic or p-value. However, if incorrect work (other than minor arithmetic/transcription errors) is shown, it will be considered to be an incorrect calculation of the respective component, even if the correct value is given.
3. If a response provides a test statistic that is not for a two-proportion z-test, then that response is scored as an I on Section 2.

Section 3 (stating a correct conclusion with justification) is scored as follows:

Essentially correct (E) if the response correctly includes the following three components:

1. Provides a correct conclusion about the alternative hypothesis
2. Justification of the conclusion based on linkage between the p-value and alpha (or linkage between the test statistic and critical value)
3. The conclusion is stated in context, which must mention gender and tattoos.

Partially correct (P) if the response correctly completes two of the components required for an E.
Incorrect (I) if the response correctly completes one or fewer components required for an E.

Notes:

1. Incorrect statistical statements will be considered to be incorrect conclusions for the hypothesis test and therefore do not meet the criteria for component 1.
2. Statements in the conclusion like "we conclude …", "this proves …", or "this guarantees …" do not meet the criteria for component 1. Statements like "this is sufficient evidence …" are acceptable.
3. If the response provides a correct decision, in context, with linkage to the p-value, but the decision is stated in terms of the null hypothesis only, so no conclusion is made about the alternative hypothesis, it does not meet the criteria for component 1.
4. If the conclusion is consistent with the p-value from Section 2, and also in context with justification based on linkage to the p-value, then Section 3 is scored as E.
5. The justification must compare the p-value to the alpha, not merely list them next to each other. If alpha is not explicitly referred to, the solution must be explicit about the linkage by giving a correct interpretation of the p-value or explaining how the conclusion follows from the p-value (for example, the following are acceptable "because the p-value is small, we reject the null hypothesis" or "because the p-value is large, we do not reject the null hypothesis").
6. A decision about the null hypothesis (reject H_0 or fail to reject H_0) is not required. However, if one is given, then scoring of that decision is considered in component 2.
7. If a graph is given but does not include any statistically wrong statements, it is counted as extra information and ignored. If a graph includes incorrect statements or is incorrectly used as a basis for their inference, they cannot get credit for component 1.

Figure 5: Example of grading rubric (2 sample inference)

AP Review Scoring Rubric: Regression

Section A is scored as follows:

Essentially correct (E) if the response includes the following two components:

1. The correct regression equation is given.
2. The interpretation includes <u>all</u> the following: (a) references the correct r^2 value (r^2= 96.09% or r^2 = .9609), (b) argues the model is useful, (c) explains the reason the model is useful is because it measures the proportion (or percent) of variation in temperature due to variation in the number of chirps.

Partially correct (P) if the response includes one of the two components required for E.

Incomplete (I) otherwise.

Note: For Section A, the student does not need to use context (e.g. saying r^2 measures the variation in y due to x satisfies component 2c).

Section B is scored as follows:

Essentially correct (E) if the response includes the following 4 components:

1. A correct value for the slope*
2. Establishes context by mentioning chirps (or number of chirps)
3. Establishes context by mentioning temperature (mentioning "degrees" is sufficient)
4. Identifies that temperature <u>increases</u> with each additional chirps.

Partially correct (P) if the response includes 3 of the 4 components required for E.

Incomplete (I) otherwise.

Note: If the slope was incorrect in Section (a), the student will need to interpret that same incorrect value in Section (b) in order to earn the first component. (In other words, they words – they don't get deducted twice for using the wrong slope, but they must consistently use that wrong slope in Sections (a) and (b)).

Sections C and D are scored as follows:

Essentially correct (E) if the response includes the follow components:

1. A correct prediction (75.697) from the regression equation.
2. An decision not to predict for 8 chirps
3. A reason not to predict of 8 chirps that mentions extrapolation or that 8 is outside of the range of original data values
4. A decision to support or to refute Dolbear's Law
5. A justification of their decision as follows:
   (a) Support: compares slopes and intercepts and says they are <u>roughly equal</u> or compares predictions for a specific number of chirps within the range of the original data
   (b) Refute: compares slopes and says they are <u>not equal</u>

Partially correct (P) if the response includes either:

Component 1 and 2 out of 4 other components required for an E. Or, All of components 2 – 5 required for an E. Incorrect (I) otherwise.

Overall score for the problem is determined as follows:

4	**Complete Response** Three sections essentially correct (EEE)
3	**Substantial Response** Two sections essentially correct and one section partially correct (EEP, EPE, PEE)
2	**Developing Response** Two sections essentially correct and no sections partially correct (EEI, EIE, IEE) *Or* One section essentially correct and one or two sections partially correct (EPP, PEP, PPE, EPI, EIP, PIE, PEI, IEP, IPE) *Or* Three sections partially correct (PPP)
1	**Minimal Response** One section essentially correct and no sections partially correct (IIE, IEI, EII) *Or* No sections essentially correct and two sections partially correct (PPI, PIP, IPP)
0	**Other Response** No sections are essentially correct and at most one section is partially correct (III, IPI, PII, IIP)

Figure 6: Example of grading rubric (regression)

2.3 Follow-Up Survey

We designed and administered a follow-up student survey to obtain student feedback. The survey is given in Figure 7. There were an equal number of male (28) and female (28) participants. In addition, the participants by class rank consisted of 22 seniors, 26 juniors, and 4 sophomores.

W.I.S.E. Workshop Survey

1. Please rate your overall satisfaction of the statistics workshop.

   Not satisfied 1 2 3 4 5 6 7 8 9 10 Highly Satisfied

   Comments:

2. Please rate your overall satisfaction with Session 1: Gettysburg Address Activity

   Not satisfied 1 2 3 4 5 6 7 8 9 10 Highly Satisfied

   Comments:

3. Please rate your overall satisfaction with Session 2: Psychic Cards

   Not satisfied 1 2 3 4 5 6 7 8 9 10 Highly Satisfied

   Comments:

4. Please rate your overall satisfaction with Session 3: Inference

   Not satisfied 1 2 3 4 5 6 7 8 9 10 Highly Satisfied

   Comments:

5. Please rate your overall satisfaction with Session 4: Regression

   Not satisfied 1 2 3 4 5 6 7 8 9 10 Highly Satisfied

   Comments:

6. Please rate your overall satisfaction with Session 5: Careers in Statistics

   Not satisfied 1 2 3 4 5 6 7 8 9 10 Highly Satisfied

   Comments:

7. How would you improve this workshop? (Check all that apply.)

___Make the workshop less difficult. ___Speed up the pace of the workshop.

___Make the workshop more difficult. ___Allot more time for the workshop.

___Slow down the pace of the workshop. ___Shorten the time for the workshop.

8. What other improvements would you recommend in this workshop?

9. What is least valuable about this workshop?

10. What is most valuable about this workshop?

Demographics

High School: Madison Central Model

Class Rank: Freshman Sophomore Junior Senior **Gender**: _____

Anticipated College Major: _____

Anticipated College: _____

Figure 7: Follow-up survey

The results for the first six questions are summarized in table 2.

Table 2: Survey results rating satisfaction of different workshop activities

Activity	Min	Q_1	Median	Q_3	Max	Mean	Mode
Overall	2	6	7	8	9	6.96	7
Gettysburg	3	7	9	10	10	8.23	9
Psychic	4	8	9	10	10	8.53	10
Inference	3	5	6	8	10	6.41	5
Regression	1	5	7	8	10	6.51	8
Careers	2	7	8	9	10	7.56	8

The responses to the question, "How would you improve the workshop?" are given in table 3.

Table 3: Survey results on how to improve the workshop

	Student Responses
Make Less Difficult	6
Make More Difficult	10
Slow Down Pace	34
Speed Up Pace	2
Allot More Time	28
Shorten The Time	1

The follow-up survey asked three opened ended questions. Below is each question and a brief summary of the responses.

1. What other improvements would you recommend in this workshop?
   - More hands-on activities
   - Smaller class size
   - Give more time for exam tips
   - Cover broader range of content
2. What is least valuable about this workshop?
   - Too fast paced
   - Admission to EKU information
   - Allow students to choose topics
   - If I didn't understand something there really was no time to explain it
3. What is most valuable about this workshop?
   - Getting feedback from AP grader
   - Practice/mock questions for the AP exam
   - The Stats review
   - The hands-on experiences taught a lot and put things better into perspective
   - Interesting careers

3. Conclusions

3.1 Lessons Learned

We found that students fell into two groups, those planning to take the exam and those who did not. For those students not planning to take the exam, we need to create optional sessions with less focus on taking an exam. In general, the content per session needs to be decreased and sessions need to be shorter. Additionally, more time needs to be spent on older topics they have already learned, including probability.

The students need more variety in the types of sessions offered. There was a split in preference for the active learning versus exam preparation. For future iterations, it would be better if students who prefer activities could just choose more of these activities, but students who prefer for exam prep material could chose more of those activities. In both types of sessions, prizes helped to motivate students. We also found that many students exhibited lack of teacher independence. Additionally, the workshop needs to include opportunities for students to work in groups

3.2 Future Endeavors

Based on experience and student feedback, there are several future endeavors that we would like to pursue. Instead of inviting teachers from only two schools, we plan to invite teachers of AP students from all schools in the EKU service region and surrounding counties. In order to promote the workshop, we plan to create brochures or an online presence. Additionally, we plan to seek external funding opportunities.

We would like to expand the scope of the workshop and the types of sessions offered. In the future we plan to open the workshop with a session utilizing Kahoot or other similar service to review important concepts with students. Building on the current sessions, we would like to run multiple concurrent sessions to offer a choice of sessions to participants. In order to facilitate multiple sessions, we plan to recruit more faculty to participate. We would also like to include sessions for the teachers that accompany the students and adapt the workshop to also offer a session for AP calculus students.

Further, we plan to collect more data, including a question asking if students plan to take the end of course national exam on the follow-up survey.

References

Lock, R. H., Lock, P. F., Morgan, K. L., Lock, E. F., & Lock, D. F. (2016). *Statistics: Unlocking the power of data*. Wiley Global Education.