

SESUG 2023 Paper 235
Teaching Business Statistics with JMP Software

Shanzhen Gao
Department of Computer Information Systems
Reginald F Lewis College of Business
Virginia State University
sgao@vsu.edu

Abstract

JMP is a suite of computer programs for statistical analysis developed by JMP, a subsidiary of SAS Institute. We will list topics that can be introduced with JMP software in an undergraduate business statistics course. We will present strategies for engaging students in active learning and enhancing students' confidence in learning business statistics with JMP, such as implementing ChatGPT in education.

Keywords

Business Statistics, JMP Software, ChatGPT, active learning, engagement, collaborative learning

Introduction

In the next section, you will see a long list of undergraduate business statistics topics that we plan to cover with the help of JMP software. There are various challenges in a business statistics course. We will discuss several approaches to conquer or reduce the difficulties.

Active learning is any instructional method that provokes thinking about an activity, which engages students to carry out meaningful learning activities in the learning process [1,2]. Active Learning Environments boost students' creative thinking, motivation for learning outcomes and performance, communication, and interaction, leading students to learn more and achieve better examination grades [3,4,5]. However, most computer course instructors in higher education do not use active learning methods, and traditional teaching methods such as simple face-to-face-lecturing remain the standard practice [6].

We at this moment present our different and varied experiences in engaging students in active learning in our Business Statistics course, from setting up the first day of class to introducing techniques such as collaborative learning to help engage students and improve course participation overall.

Business Statistics Topics with JMP Software

We plan to introduce the following topics in an undergraduate business statistics course with the help of the JMP software.

Descriptive Statistics

Measures of Central Tendency

Mean, Geometric Mean, Median, Mode

Measures of Dispersion

Variance, Standard Deviation, Range, Quantile, Minimum, Maximum, Summation, Five-Number-Summary

Dependency

Covariance, Correlation

Visualizing Data

Frequency Distribution, Histograms

Bar Chart, Pie Chart, Box Plot, Dot Plot, Bubble Plot

Scatter Diagram and Trendline, Scatter Plot Matrix

Line Graph, Multiple Line Graph

Stem-and-leaf Plot

Outlier Analysis

Cluster Analysis

Hierarchical Clustering, K-Means Clustering

Sampling

Subset, Random Sampling

Regression Analysis with JMP

Simple Linear Regression, Multiple Linear Regression, Logistic Regression

Time Series Analysis and Forecasting

Time Series Plot, Forecasting

Hypothesis Testing

Chi-Square Test, t-Test

Probability Distributions

Discrete Probability Distributions, Continuous Probability

Hypergeometric Probability Distributions, Exponential Distribution

Poisson Probability

Principal Component Analysis

Data Reduction, Principal Component Analysis

Association Rules

Generating Association Rules, Interpreting Association Rules, Visualizing Association Rules

Student's Questions Before and After Using JMP Software

Student's question when introducing JMP software

We can use a calculator and Excel. Why do we learn JMP?

Student's questions after using JMP software

JMP is much easier and more powerful. Why do we need to use Excel? Why do we need to calculate by hand/calculators?

Challenges in Business Statistics with JPM Course

There are various challenges in business statistics with the JMP course, such as

- the installation of software
- diversity in mathematics, logic, and reasoning
- classroom participation, attendance, and engagement
- complexity of concepts and topics
- time-consuming and time-constraint
- designing and creating proper and effective assessments
- diverse student backgrounds
- post Covid pandemic effects

In the following sections, we will discuss overcoming these uncertainties and difficulties and engaging students in active learning.

Student Engagement: Before Class

To prepare students for their engagement in our Business Statistics with JPM course, we introduce the class through the following resources:

Syllabus

At the very beginning, the course syllabus is posted online through the course learning management system.

Textbook

An e-version of the textbook, including its references such as title, author(s), and some brief descriptions of essential elements of the course are posted in the learning management system. This enables students to preview textbook content and facilitates resolving possible access issues once in class.

Software

Information on the JMP, Anaconda Jupyter Notebook, Google Cloud Platform: Colab, and other software, their installations, setup, troubleshooting tips, and help links, along with a Frequently Asked Questions (FAQs) page, are available well in advance. Students are encouraged well in advance to review provided software resources; they can download the software and its allied resources before arriving in class. We encourage students to help each other with software installations which can offer multiple benefits, such as enhancing the learning experience, equipping students with essential skills, and cultivating a supportive community.

First Day of Class: Survey

On the first day of class, students must complete a survey to gauge their expectations and level of preparedness for the class experience. The survey includes ten questions, and we list five random questions.

What are your goals and expectations from the Introduction to Computing Class?

How do you plan to achieve your stated goals for the course?

You may face some learning challenges at some point in the course. How do you plan to meet such challenges?

What are your grades' expectations for the course?

What do you consider to be an appropriate and inappropriate usage of technology during class time?

First Day of Class: Collaborative Learning

To help students engage in a collaborative learning environment, we require them to perform the following tasks on Day 1:

Task 1: Form groups of 2-3 and introduce each other within your group.

Task 2: Help your group members with any technology-related issues, such as software installation and how to find and submit materials in the learning management system.

Task 3: Take turns introducing a member of your group to the class.

Task 4: Exchange information with some classmates to form a possible study partnership or workgroup.

We anticipate that collaboration can be challenging in the first two or three weeks of the semester or quarter. We teach students cooperative skills, such as active listening, asking thoughtful questions, and evaluating group processes, such as oral and written communication, consultation, and negotiation.

Using Anaconda Jupyter Notebook and Google Cloud Platform: Colab

Jupyter Notebook and Google Cloud Platform offer dynamic and interactive platforms that can significantly enhance the teaching and learning experiences. We use them for data visualization. We hope to prepare students for the modern, data-centric business world.

Business Statistics with ChatGPT-4

We will observe ChatGPT's responses to a few questions in the following.

Question to ChatGPT

When Tom was born, his father was 30 years old. Now Tom is 30 years old. What is the average age of Tom and Tom's father now? Why?

Response from ChatGPT

Let's break this down step by step:

1. When Tom was born, his father was 30 years old.
2. Now Tom is 30 years old.
3. This means 30 years have passed since Tom was born.
4. Thus, Tom's father's current age = 30 years (his age when Tom was born) + 30 years (that have passed since then) = 60 years.

Now, let's calculate the average age of Tom and his father:

Average age = (Tom's age + Tom's father's age) / 2 = (30 years + 60 years) / 2 = 90 years / 2 = 45 years.

The average age of Tom and his father now is 45 years.

To explain why: The average is calculated by summing up the ages of the two individuals and dividing by 2. In this scenario, Tom is half the age of his father (30 and 60 respectively), so the average is simply halfway between their two ages, which is 45.

Question to ChatGPT

Please write down your age and your mother's age. What is the average age of you and your mother? Why?

Response from ChatGPT

I am a machine learning model, so I don't have personal experiences, emotions, or a concept of age. Similarly, I don't have a mother in the traditional or biological sense. Therefore, I cannot provide an average age for a non-existent entity and myself. However, the concept behind calculating an average age is simple: you add the ages together and then divide by the number of ages you added. This gives the mean, or average, of the set of numbers. If you have any other questions or topics you'd like to explore, feel free to ask!

Same Question to ChatGPT

Please write down your age and your mother's age. What is the average age of you and your mother? Why?

Response from ChatGPT

I apologize for any confusion, but as a digital entity, I don't have an age or a mother. Consequently, I can't provide an average age for the two of us.

However, I can explain the concept: To find the average age of two individuals, you would add their ages together and then divide by 2 (since there are two people). The reason we average in this manner is to find the midpoint or central value of a set of numbers.

Remarks

ChatGPT can assist students in answering questions. Instructors may create questions that ChatGPT is not able to provide a solution although it could be a challenge.

Question to ChatGPT

Given seven numbers: 20, 18, 20, 19, 18, 18 and 19. Please use JMP software to find the mean, median, mode, range, maximum, minimum, summation, variance, and standard deviation. Please show me your work in detail.

Response from ChatGPT

While I can't directly operate JMP software here, I can walk you through how to calculate these statistical measures using the given data. After the explanation, you can replicate the steps in JMP or any other software. Given numbers: 20,18,20,19,18,18,19

- 1. Sorting the numbers:**
 - Sorted set: 18, 18, 18, 19, 19, 20, 20
- 2. Mean (Average):**
 - Mean = $\frac{\text{Sum of all numbers}}{\text{Total number of items}}$
 - Mean = $\frac{18+18+18+19+19+20+20}{7} = \frac{132}{7} = 18.857$
- 3. Median (Middle Value):**
 - Since there are 7 numbers (odd), the median is the fourth number.
 - Median: 19
- 4. Mode (Most Frequent Number):**
 - 18 appears 3 times, which is the most frequent, so the mode is 18.
- 5. Range:**
 - Range = Maximum – Minimum
 - Range: $20 - 18 = 2$
- 6. Maximum and Minimum:**
 - Maximum: 20
 - Minimum: 18
- 7. Summation:**
 - Total sum: 132
- 8. Variance:**
 - $s^2 = \frac{\sum(X_i - \text{mean})^2}{n - 1}$
 - $s^2 = \frac{(20 - 18.857)^2 + (18 - 18.857)^2 + (20 - 18.857)^2 + \dots}{7 - 1}$
 - After calculations, $s^2 = 0.9524$
- 9. Standard Deviation:**
 - $s = \sqrt{\text{Variance}}$
 - $s = \sqrt{0.9524} = 0.9759$

To use JMP for these calculations:

1. Input your data into a JMP data table.
2. Under the 'Analyze' menu, choose 'Distribution.'
3. Assign the column of numbers to the 'Y, Columns' box.
4. Click 'OK.'
5. JMP will provide a histogram and various statistics, including the Mean, Std Dev, Min, and Max. For median, you can right-click on the output statistics table and select 'Columns' > 'Median.' For Mode, JMP doesn't directly provide this, but you can infer it from the histogram or frequency table. For variance, right-click on the output statistics table and select 'Columns' > 'Variance.'

This will give you the desired statistical measures!

Question to ChatGPT

Assume that your 70-year-old grandma retired last week and saved one million dollars in her retirement account after 54 years of hard work. You just received an email from her with the following.

Dear Tom,

I hope you are having a good time at school, and everything is going well.

I am very excited to inform you that I will be wealthy in one day, next Friday. 10/20/2023.

I will use all the money in my retirement account to buy the "You will be rich in only one-day Lottery" lotteries next Friday. I will win 100 million dollars.

Then I will give you one million to buy the "You will be rich in only one day Lottery." The lottery is on Friday, 10/27/2023. You will win 100 million dollars.

Best wishes for your studies and good luck with your exams.

Love you!

Grandma Kate

Questions:

Are you happy? Do you worry about your grandma? Why?

You researched and found something about "You will be rich in only one-day Lottery".

One ticket is one dollar. The probability distributions are as follows.

The probability of winning 1000 is 0.00001.

The probability of winning 100 is 0.00002.

The probability of winning 0 is 0.99997.

Task 1: You write a friendly and respectful email to your grandma to explain your findings and analysis about the "You will be rich in only one-day Lottery". Please use the words and sentences that your grandma understands. For instance, you may use "average return" as Expected Value.

Task 2: Make a few slides (page) of PowerPoint for presenting your ideas.

Task 3: Record your presentation.

Response from ChatGPT

Task 1: Email to Grandma

Dear Grandma Kate,

I hope you're doing well. Thank you so much for thinking of me. Your email made me smile, and it's always heartwarming to know you have such big hopes and dreams for both of us! I took some time to look into the "You will be rich in only one-day Lottery" to better understand it. Here's what I found, explained in the simplest way:

1. For every ticket you buy, there's a very, very small chance (think of it as one in a hundred thousand) that you could win \$1000.
2. There's an equally small chance to win \$100.
3. However, almost all the time (like 99,997 times out of 100,000), the ticket doesn't win any money.

When we think of an "average return" from buying many tickets, it means that, in the long run, we might end up with much less than we spent.

I love your optimism, Grandma, but I'm a little worried about using all your hard-earned savings on this. Perhaps we can talk more about it and maybe consider other fun or meaningful ways to use or invest the money?

Sending you all my love and looking forward to our chat!

Warmly, Tom

Task 2: PowerPoint Slides

Slide 1: Title Slide

- Title: "The Real Odds of the Lottery"
- Image: Brightly colored lottery ticket

Slide 2: The Dream of Winning Big

- Image: Dollar signs and big jackpot numbers
- Text: "The hope of turning \$1 into millions!"

Slide 3: What Are The Chances?

- Bullet Point 1: "Winning \$1000 = 0.00001 chance" (Image: Stack of \$1000 bills)
- Bullet Point 2: "Winning \$100 = 0.00002 chance" (Image: \$100 bill)
- Bullet Point 3: "Not winning = 0.99997 chance" (Image: Empty piggy bank)

Slide 4: What Does "Average Return" Mean?

- Image: Scale tipping more towards cost than returns
- Text: "Over time, we might spend more than we win. It's like buying a candy bar every day but only getting a candy once a month."

Slide 5: There's More to Money Than Winning

- Image: Heartwarming photo of Grandma and Tom together
- Text: "The value of hard-earned savings and memories together."

Task 3: Recording the Presentation

Unfortunately, I can't directly record a presentation. However, I'd suggest using software like Zoom or PowerPoint's in-built recording feature to narrate the slides. Ensure you communicate your genuine concern for Grandma Kate and the value of the memories and hard work behind the savings.

Advantages and Disadvantages of Using ChatGPT in an Undergraduate Business Statistics Course

According to our limited experience in business statistics courses, ChatGPT is a learning tool that may help students know some concepts and stimulate critical thinking. It is a challenge for instructors and students to use ChatGPT wisely. ChatGPT can also be a barrier for students to really understand and master certain knowledge. Creating proper and effective assessments and preventing academic dishonesty is also challenging.

Advantages and Disadvantages of Using JMP Software in an Undergraduate Business Statistics Course

Advantages

Easy to use and user-friendly.

Efficient data visualization.

Free access to educational resources and real-world applications.

Do not need to spend time on a new programming language.

Disadvantages

Cost of licensing.

The JMP Scripting Language (JSL) is not as popular as Python or R.

Updates might not be as frequent as in Python.

Python, R, and SAS have been widely adopted in the analytics and data science industry. Students who are familiar with these tools are more marketable when entering the job world.

The operations and features of JMP might pose initial challenges to students.

The limitation of advanced analytical capabilities.

Conclusion/Remarks

Most students in our Business Statistics course demonstrate a positive attitude toward class participation, engagement, and completing assignments and quizzes.

Our approaches to active learning may apply to other subjects/fields such as Engineering or Mathematics. These fields have been known to have few interactions between instructors and students in the classrooms. It is important to promote active learning methods in any STEM field and encourage instructors to participate in active learning. This will help increase the stem workforce that is needed in this country.

References

1. Bonwell, C.C., and J. A. Eison, "Active Learning: Creating Excitement in the Classroom," ASHEERIC Higher Education Report No.1, George Washington University, Washington, DC, 1991.
2. Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, 93(3), 223–231.
3. Robert, A.P. (2016), *Fostering Creativity Through the Active Learning Model Kanban2C*, <https://doi.org/10.13021/G89C7V>
4. Bonwell C.C., Eison A.J., *Active Learning: Creating Excitement in the Classroom*. Washington, DC: George Washington University Press; 1991. ASHE-ERIC Higher Education Report No. 1.
5. Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410-8415., 8410-8415.
6. Deslauriers, L., McCarty, L.S., Miller, K., Callaghan, K., Kestin.G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences*, 1-7.