



# Teaching Principles of One-Way Analysis of Variance Using M&M's Candy

[Todd A. Schwartz](#)

University of North Carolina at Chapel Hill

*Journal of Statistics Education* Volume 21, Number 1 (2013),  
[www.amstat.org/publications/jse/v21n1/schwartz.pdf](http://www.amstat.org/publications/jse/v21n1/schwartz.pdf)

Copyright © 2013 by Todd A. Schwartz all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the author and advance notification of the editor.

---

**Key Words:** Teaching statistics; Active learning; Activities; ANOVA; M&M's.

## Abstract

I present an active learning classroom exercise illustrating essential principles of one-way analysis of variance (ANOVA) methods. The exercise is easily conducted by the instructor and is instructive (as well as enjoyable) for the students. This is conducive for demonstrating many theoretical and practical issues related to ANOVA and lends itself to multiple possible configurations of ANOVA results, leading to rich classroom discussion and deeper student understanding of real-world applications of the methods.

## 1. Background

M&M's candies have successfully been introduced to the statistics classroom to illustrate a variety of topics, including frequencies and proportions, probability functions ([Magel 1998](#)), sampling distributions ([Dyck and Gee 1998](#)), design of experiments ([Lin and Sanders 2006](#)), chi-square goodness-of-fit statistics ([Magel 1996](#)), and correlation/linear regression ([Froelich and Stephenson 2013](#)). This candy's popularity with students may stem from its colorful shells, as well as its propensity to serve as an edible reward at the conclusion of the experiment. The candies also provide an intuitive basis on which instructors can build, as students tend to have a reasonable degree of what to expect based on their previous experiences in dealing with seeing and eating M&M's.

This exercise is consistent with the advantages gained by incorporating fun activities and active learning into the classroom ([Lesser and Pearl 2008](#)), such as reducing students' anxiety or stress as well as enhancing student engagement and learning. Furthermore, it aligns with the first five

recommendations of the American Statistical Association's GAISE College Report ([ASA 2010](#)) through its use of student-generated data to develop statistical problem-solving, promotion of deeper conceptual understanding, and utilization of active learning techniques and technology.

I have had success in utilizing M&M's to provide a straightforward illustration of the methods of one-way analysis of variance (ANOVA), including fodder for discussion of many of the relevant considerations that accompany these methods. Although an instructor following this approach may not know *a priori* what the results from a given classroom experiment will yield, the exercise provides several alternative response variables for analysis; thus, it is likely to provide a variety of scenarios from which the instructor may select those that will be useful in classroom demonstrations, homework exercises, or examination questions.

## 2. Materials

The setting in which I have conducted this demonstration is a university course comprised of 12 to 15 students in a doctoral-level nursing program, though it is easily adaptable to classes comprised of undergraduate students in various disciplines. The classroom is equipped with a computer loaded with spreadsheet and statistical software, which can be projected onto a screen in the classroom. The classroom has individual desks which are not fixed and may be rearranged to suit the needs of the class, or is otherwise configured to allow the students to align themselves into pre-assigned groups. Students do not need access to individual computers during this exercise, as they will be observing me and my use of the classroom computer.

These students have had previous, introductory exposure to one-way ANOVA, though this demonstration should also be suitable to motivate the topic for novice students who may not have had any prior ANOVA experience.

In terms of the sequencing of course material, this exercise occurs after I have discussed two-independent group *t*-tests, and have motivated the need for an extension to those methods for comparing more than 2 groups. I make introductory comments to motivate further the utility of ANOVA, including discussing the types of research questions that ANOVA can address, as well as the issues of within- and between-group variability and the influence of sample size. At that point, I am ready to begin this activity.

Prior to the in-class demonstration, I need a minimal amount of preparation time. First, I create groupings of the students. Each group is approximately one-third of the total class size (typically 4-6 students in my case). Further, each group will receive only one type of M&M's (e.g., peanut M&M's), though each member of the group receives his or her own packet of that type of M&M's.

Also, I prepare an unpopulated spreadsheet to record the data from the students in real-time (see [Figure 1](#)). This step gives the students experience in how to structure their primary data collection to facilitate analysis after the data are collected. In the first column, I list each student's name in a separate row, grouped according to their assignment.

**Figure 1.** Unpopulated spreadsheet

|    | A              | B            | C            | D             | E          | F           | G             | H            | I            | J            |
|----|----------------|--------------|--------------|---------------|------------|-------------|---------------|--------------|--------------|--------------|
| 1  | <b>Student</b> | <b>Group</b> | <b>Total</b> | <b>Yellow</b> | <b>Red</b> | <b>Blue</b> | <b>Orange</b> | <b>Brown</b> | <b>Green</b> | <b>Check</b> |
| 2  | A              | Plain        |              |               |            |             |               |              |              | 0            |
| 3  | B              | Plain        |              |               |            |             |               |              |              | 0            |
| 4  | C              | Plain        |              |               |            |             |               |              |              | 0            |
| 5  | D              | Plain        |              |               |            |             |               |              |              | 0            |
| 6  | E              | Peanut       |              |               |            |             |               |              |              | 0            |
| 7  | F              | Peanut       |              |               |            |             |               |              |              | 0            |
| 8  | G              | Peanut       |              |               |            |             |               |              |              | 0            |
| 9  | H              | Peanut       |              |               |            |             |               |              |              | 0            |
| 10 | I              | Pretzel      |              |               |            |             |               |              |              | 0            |
| 11 | J              | Pretzel      |              |               |            |             |               |              |              | 0            |
| 12 | K              | Pretzel      |              |               |            |             |               |              |              | 0            |
| 13 | L              | Pretzel      |              |               |            |             |               |              |              | 0            |

The second column contains a code for the group to which they have been assigned. This code could be generic (e.g., ‘A,’ ‘B,’ or ‘C’), or it could be a descriptor of the type of M&M’s with which they will be dealing (e.g., ‘Plain’ or ‘Peanut’). I have found it better not to use a numeric label to avoid confusion, as this variable will be treated as a qualitative variable (i.e., a ‘factor’) rather than a quantitative one where the numbers would have intrinsic meaning.

Subsequent columns will contain the observed M&M counts as reported by the students, including the total count and counts by color, as described more fully below. I choose to include a final column in the spreadsheet for which I program a formula to sum across the colors and subtract that sum from the total; this allows a built-in “error check” to compare that sum to the student-reported total and can assist in verifying the students’ accuracy.

Finally, I prepare the M&M’s for distribution to the students in their groups. I choose three different varieties of M&M’s, though this approach could easily extend to more than three types. In past years, the two major types have been Plain (milk chocolate in a brown wrapper) and Peanut (yellow wrapper). Today, one may find many more varieties on the store shelves, including Peanut Butter, Coconut, Dark Chocolate, Pretzel, and Almond. While packages with themed colors are available seasonally (e.g., green and red at Christmastime), the instructor will want to choose varieties carefully, as the goal is to have roughly the same colors represented across all types of M&M’s involved in the demonstration.

I supply each student with his or her own individually-sized packet of M&M’s. The cost of each packet will vary. In my local grocery store, the retail price per packet is approximately \$1.20, but instructors should be able to achieve cost savings through store promotions or by ordering

these in bulk. Ideally, the packets for all types should all be approximately the same size or weight; this will invoke variability among the types due to the varying sizes of the individual M&Ms, as a single milk chocolate M&M is smaller (and weighs less) than its peanut M&M counterpart.

I tend to choose two of the types to have individual candies of similar size, with the third being substantially different. For my example, individual peanut and pretzel M&M's are similarly sized, with both being substantially larger than the milk chocolate (or plain) M&M's. This will facilitate illustrations of varying patterns of statistically significant findings that can be intuitively understood, despite potentially small sample sizes due to the number of students in the class assigned to each group.

### 3. Methods

I start the in-class demonstration by announcing the groupings of the students, and having them rearrange into their groups. I distribute the corresponding packets of M&M's to each group, cautioning them not to "eat the data" until the conclusion of the demonstration. This step is usually accompanied by a rising noise level from the students as their interest is piqued and their anticipation rises, thus "breaking the ice" and lowering their guard (and, for some, their anxiety) to pave the way for learning to occur.

The research question can be posed in natural and easily understandable language as "*Do different types of M&M's have a different number of candies inside similarly-sized packages?*" This question may be repeated for each color under consideration. I can frame this question in the context of wanting to know which type to purchase at the store in order to maximize either the total number of M&M's or the number of a certain favorite color. Also, while each packet is of a similar weight, the students tend to gravitate toward wanting more M&M's than their classmates; high (or low) counts often elicit audible responses from the class. This framework can serve to further pique the curiosity of the students as well as encouraging their full engagement in the activity.

I then instruct the students (individually) to count the total number of M&M's in their packets, as well as to sort them by color and to count the number of each color. The six traditional colors of M&Ms are red, green, blue, yellow, orange, and brown.

After giving the students a few minutes to complete this task and to record their numbers, I project the unpopulated spreadsheet onto the classroom screen. This provides the opportunity to discuss the structure of how we will record the data and organize it for analysis (e.g., rows for observations, columns for variables). This often provides insights to the students as far as pre-planning data entry to facilitate analysis. If they understand how the data must be structured for their statistical software to conduct ANOVA, they can directly enter data to match that structure, rather than needing an additional step later in the process to rearrange the data into a different format.

At this point, the instructor can present a brief review of key concepts of one-way ANOVA, including the overall null hypothesis of testing all group population means equal to one another,

and how that specifically relates to our data. We can specify a null hypothesis for the mean total counts of each type, as well as separate hypotheses for the type means of each color subtotal. Based on their intuition, a few students can be called upon to guess whether the null hypotheses will be supported or rejected for our data, as well as to provide the rationale for their responses.

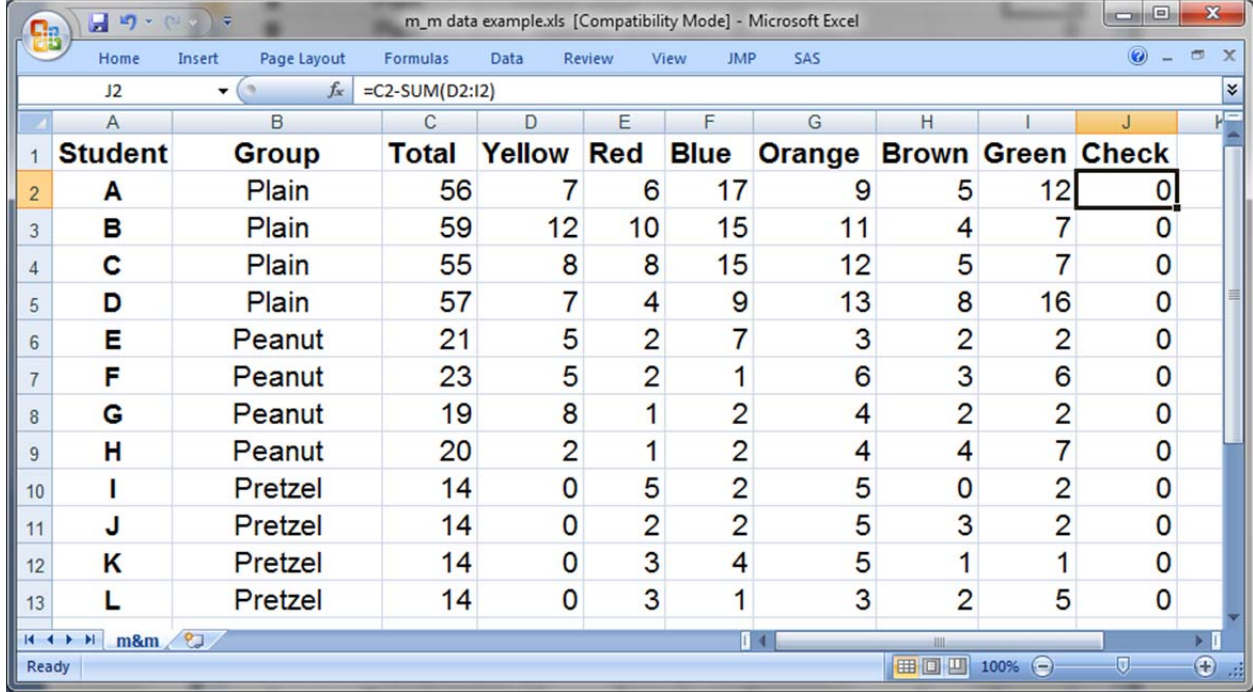
The instructor can now tie in important concepts to the dataset, by determining such specifics as identifying the dependent and independent variables and how they are represented in the spreadsheet (i.e., the columns for the [sub]totals and the column for the grouping variable, respectively).

Now that we are ready to populate the spreadsheet with the data, I call on the students individually, in the order their names appear in the spreadsheet. Each student will then orally recite his or her data values, in the order of the columns of the spreadsheet. In this example, he or she would first give the total number of M&M's in her packet, then the number of yellow M&M's, red M&M's, etc.

As noted previously, I have the final column pre-programmed to sum across the color subtotals. After the student is finished reciting his or her values, this will verify that the student's total reported does indeed match the sum of the color subtotals. If it does not, the student will need to re-count M&M's and provide any necessary corrections to the values. This provides an important lesson on data integrity: it is much easier to take precautions to detect data errors as they occur or are entered, rather than trying to reconstruct those values in a post hoc fashion.

When all students have reported their values and the spreadsheet is completely populated (see [Figure 2](#)), a few more students can be called upon to update their guesses regarding the validity of the null hypotheses, based on the evidence they now see on the screen. Finally, I save the spreadsheet file and make it available to the class (e.g., via a course management system or email attachment). It is not necessary to situate this activity in a setting where all students have access to a computer. However, if they do, instructors could take advantage of that fact to ask students to replicate the in-class findings as the activity unfolds, individually or in groups. At this point, I have a primary source dataset with which the students have been involved from beginning to end, providing seven different potential one-way ANOVAs, one for the mean total number and one for the mean of each color's subtotal. In addition to the in-class demonstration, this dataset is also useful for subsequent assignments, projects, or examinations.

**Figure 2.** Populated spreadsheet



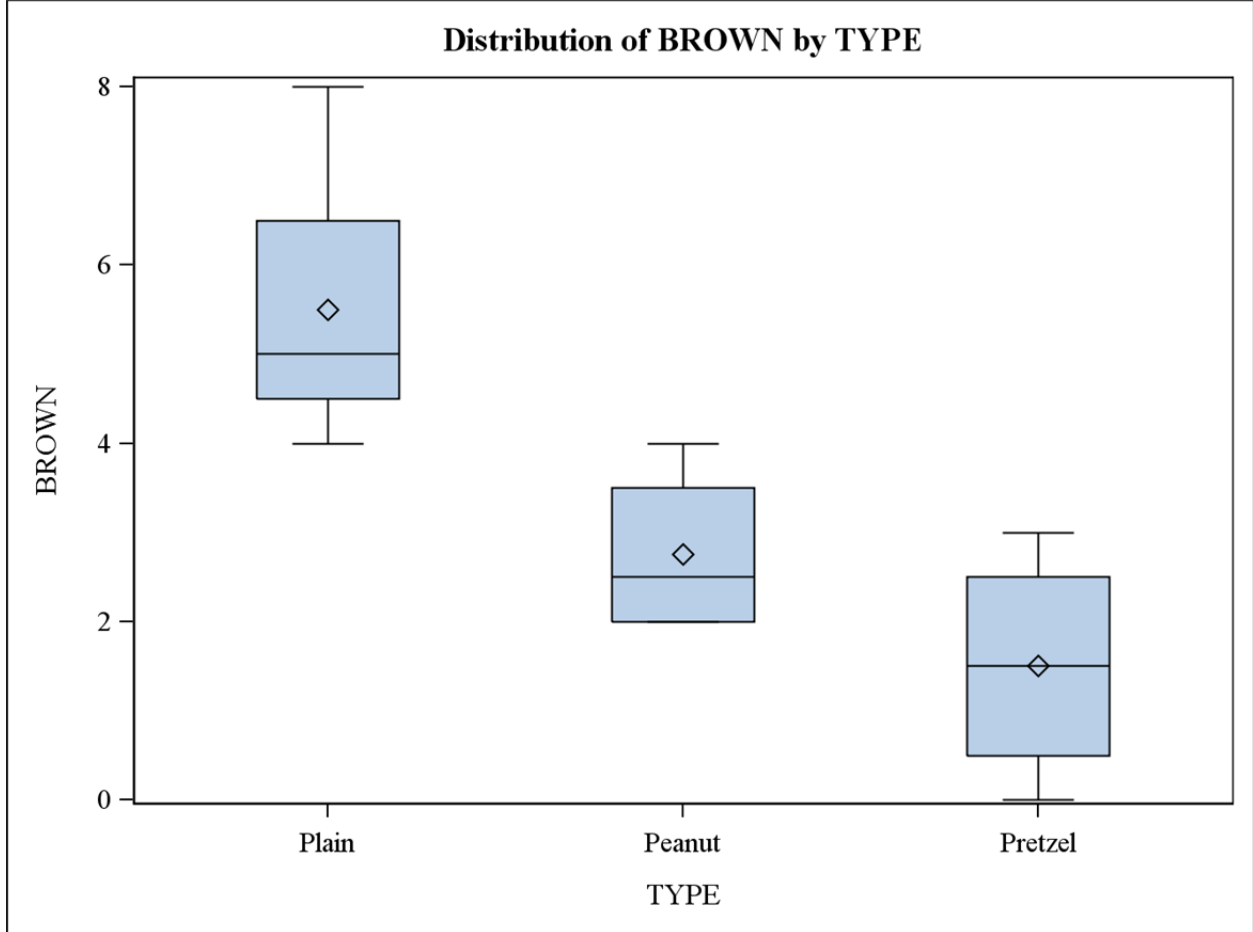
## 4. Results

In this section, I illustrate some of the results generated from a replication of this exercise. As I cover each section, I am teaching the students how to use statistical software to produce the necessary output. The output shown in this article is from SAS Software ([SAS Institute Inc. 2011](#)), but other software packages will provide corresponding results.

Preliminary analyses, including computation of univariate statistics, such as means and standard deviations ([Table 1](#)), and information on the distribution (e.g., through boxplots) can be generated separately for each type of M&M's ([Figure 3](#)). We can have a rich classroom discussion as we review these descriptive statistics, and I can tie together these numerical results with the visual impact of the plots.

**Table 1.** Selected descriptive statistics for the subtotal of brown M&M's

| GROUP   | N | Mean | Std Dev | Minimum | Maximum |
|---------|---|------|---------|---------|---------|
| Plain   | 4 | 5.5  | 1.7     | 4.0     | 8.0     |
| Peanut  | 4 | 2.8  | 1.0     | 2.0     | 4.0     |
| Pretzel | 4 | 1.5  | 1.3     | 0.0     | 3.0     |

**Figure 3.** Boxplots for the subtotal of brown M&M's

As I move to inferential statistics, I first review the two-sample  $t$ -test, as these students would have previous familiarity with this method. I then generate a two-sample  $t$ -test for each pair of M&M's types. This can provide a tie-in to illustrate the equivalency of those results with a one-way ANOVA having  $k=2$  groups. This provides an ideal opportunity to compare and contrast the output from those two approaches. I now can generalize the one-way ANOVA from  $k=2$  groups to  $k=3$  groups, and I can motivate the advantages of analyzing all of the groups in a single analysis, rather than as pairs through separate  $t$ -tests.

Tables 2, 3 and 4 show selected sample output for the one-way ANOVA on the mean subtotal of brown M&M's in a comparison of Plain, Peanut, and Pretzel M&M's. For my students, I focus on the ANOVA (sums of squares) table (Table 2), the parameter estimates (especially their interpretations; data not shown), the least squares group means provided in Table 3 and how these values are exactly the same as the descriptive statistics (means) from Table 1), and the  $p$ -values for each of the pairwise comparisons (Table 4). Depending on the nature of the course, different aspects of the ANOVA may be emphasized.



**Table 2.** ANOVA table for the mean subtotal of brown M&M's

| Source                 | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|------------------------|----|----------------|-------------|---------|--------|
| <b>Model</b>           | 2  | 33.50          | 16.75       | 9.00    | 0.0071 |
| <b>Error</b>           | 9  | 16.75          | 1.86        |         |        |
| <b>Corrected Total</b> | 11 | 50.25          |             |         |        |

**Table 3.** Model-based least-squares group means for the subtotal of brown M&M's

| GROUP          | BROWN LSMEAN |
|----------------|--------------|
| <b>Plain</b>   | 5.50         |
| <b>Peanut</b>  | 2.75         |
| <b>Pretzel</b> | 1.50         |

**Table 4.** *P*-values for pairwise comparisons (unadjusted for multiple comparisons) testing row versus column means for the subtotal of brown M&M's

| Row/Column     | Plain  | Peanut | Pretzel |
|----------------|--------|--------|---------|
| <b>Plain</b>   |        | 0.0191 | 0.0025  |
| <b>Peanut</b>  | 0.0191 |        | 0.2273  |
| <b>Pretzel</b> | 0.0025 | 0.2273 |         |

For these data, despite small sample sizes ( $n=4$  for each type), the overall null hypothesis of equal type means would be rejected ( $p=.0071$ ). I can then lead a discussion on the implications of this result. In this example, the unadjusted *p*-values suggest that the mean number of brown plain M&M's is significantly higher than either brown peanut ( $p=.0191$ ) or brown pretzel M&M's ( $p=.0025$ ). However, the mean number of brown peanut M&M's does not differ significantly from the mean number of brown pretzel M&M's ( $p=.2273$ ). This provides a counter-example to some students' thinking that a statistically significant overall test implies that all means are different from one another, rather than not all means are equal.

I also use this opportunity to introduce the issue of multiple comparisons. At this point, I can now teach my students a few of the more commonly used methods for adjustment of pairwise testing in the context of multiple comparisons, such as the Bonferroni or Tukey corrections, as well as the advantages of each. Tables 5 and 6 illustrate these results for my dataset for the Bonferroni- and Tukey-adjusted results, respectively. Here, an interesting situation arises where the Bonferroni approach provides  $p>.05$  for the comparison of Plain vs. Peanut, while the Tukey approach yields  $p<.05$ . I can engage the students in a practical (it's no longer hypothetical!)



discussion of strategically selecting the multiple comparisons technique *a priori*, depending on their objectives for the study. I can also take advantage of this situation to delve into the rationale for using a significance level of .05, and how it is not a “magical” criterion that distinguishes important findings from unimportant ones.

**Table 5.** *P*-values for Bonferroni-adjusted pairwise comparisons for testing row versus column means for the subtotal of brown M&M’s

| Row/Column | Plain  | Peanut | Pretzel |
|------------|--------|--------|---------|
| Plain      |        | 0.0572 | 0.0075  |
| Peanut     | 0.0572 |        | 0.6819  |
| Pretzel    | 0.0075 | 0.6819 |         |

**Table 6.** *P*-values for Tukey-adjusted pairwise comparisons for testing row versus column means for the subtotal of brown M&M’s

| Row/Column | Plain  | Peanut | Pretzel |
|------------|--------|--------|---------|
| Plain      |        | 0.0456 | 0.0063  |
| Peanut     | 0.0456 |        | 0.4321  |
| Pretzel    | 0.0063 | 0.4321 |         |

As I examine each of the seven different possible ANOVAs, interesting configurations inevitably arise. As an example, for one color we might find a significant *p*-value when evaluating the overall null hypothesis, with a mixture of significant and nonsignificant pairwise differences; for another, we might find the counter-intuitive pattern of a significant *p*-value for the overall null hypothesis, with all pairwise differences being nonsignificant. For yet another, we might encounter the scenario where we report a nonsignificant *p*-value for the overall null hypothesis, but find one or more significant pairwise differences. These scenarios afford the opportunity to discuss the possibilities of such paradoxical findings. Of course, I am exerting influence on some of these configurations by choosing at least one of the types to be significantly smaller or larger than the others. This facilitates the existence of interesting findings in at least one of the ANOVAs, since I do not know *a priori* what the findings will be.

[Table 7](#) shows the pairwise comparisons that arise from the ANOVA for the total number of M&M’s among the three types. Clearly, all three types differ from one another with respect to their means. At this point, we are finally able to answer the research question posed at the beginning of the exercise. For this sample, we would conclude that each different type of M&M’s does indeed differ from the others in the average number of candies inside similarly-sized packages.

**Table 7.** *P*-values for pairwise comparisons (unadjusted for multiple comparisons) for testing row versus column means for the total M&M’s counts

| Row/Column | Plain  | Peanut | Pretzel |
|------------|--------|--------|---------|
| Plain      |        | <.0001 | <.0001  |
| Peanut     | <.0001 |        | <.0001  |
| Pretzel    | <.0001 | <.0001 |         |

## 5. Discussion

One advantage to this approach is that this demonstration need not be conducted in a computer lab, nor do the students need to have access to their own computers during the demonstration. Also, students are involved in the entire process, experiencing all stages from actual, real-time data collection and processing, through the data analysis and the appropriate follow-up steps.

This exercise provides an important lesson on how to structure data. While the proper formatting may come as second nature to experienced statisticians, students often do not have a natural sense of how to arrange the data to accommodate the structure necessary for analysis. For example, students may initially want to arrange separate columns to represent each group when the software requires a case-record format with separate observations in each row so the number of rows is equal to the overall sample size.

In our data collection process, there are inevitably discrepancies, reflecting a reporting problem either in the total or in one or more of the color subtotals. When this occurs, it serves as a natural segue to discuss data quality issues and the importance to take precautions to ensure accuracy of the data.

In addition to the above mentioned, we also discuss a myriad of considerations students may encounter in their subsequent research careers, including: issues of random assignment to groups, within- versus between-person factors, effect size, within- and between-group variability, equal versus unequal group sample sizes, sample size and power considerations, the importance of preliminary descriptive analysis to confirm or debunk investigators' intuition, the difference between the hypothesis of equality of all group means versus pairwise comparisons, multiple comparisons, understanding of the various components of the software output, proper interpretation of that output, and the appropriate reporting of findings, just to name a few.

The issue of variability is essential for the instructor to understand and, to some extent, control. Within-group variability is expected to be low due to the manufacturers' quality control. This can lead to instances of zero variability within a group. On one hand, increasing the size of the M&M packets should increase the variability in the total counts, but that will also increase the cost of the lesson. This creates an inherent tension between the risk of having less desirable results and the expense of using M&M's for this exercise.

Between-group variability is expected to be large by construction. While there is a chance that students will predict the eventual outcome of the exercise, I have found that this foreknowledge is beneficial in that it provides a familiar structure through which the students can see how the

quantitative approach of ANOVA can produce findings that corroborate their intuition. On the other hand, the patterns in the color-specific hypotheses are less clear to the students, so elements of curiosity, surprise, and discovery are present throughout the activity.

We may also find different patterns of significant pairwise differences, due to the choice of the multiple comparison procedure. This helps to illustrate the importance of weighing the choice of the most appropriate multiple comparison procedure and selecting the most appropriate one on an *a priori* basis.

This approach also yields a dataset that can be used to serve multiple purposes. After using one of the dependent variables from this dataset as a classroom illustration, I can later assign one or more of the other possibilities – chosen strategically by me to zero in on important distinctions I want my students to understand (such as a nonsignificant overall test) – as a homework assignment, or even as the basis for an exam problem. By doing so, the students will require little description of the dataset, as they will immediately identify the context.

At the conclusion of the exercise, I can generalize the one-way ANOVA from  $k=3$  groups to an unspecified number of groups. To help the students crystallize their thinking on this topic, I can ask students to describe how this experiment could be repeated with  $k=4$  or  $k=5$  (i.e., by adding additional types of M&M's).

Later in the course, I can return to the M&M's data to illustrate one-way analysis of covariance (ANCOVA) (e.g., examining one of the color subtotals, while covariate-adjusting for the total counts). At that point, we can ask a more focused research question for each color: “*Do different types of M&M's have a different number of candies of a specific color when the total number in the packages is the same?*” After covering two-way ANOVA, I can also use the demonstration to reinforce concepts by asking how we might have extended the M&M's experiment to include a second factor (e.g., repeating the process, but adding bags of M&M's of a different size, and doing so for each type).

A major advantage of this approach is that the students serve as stakeholders in the data – they have generated the data and have an intuitive sense of any obvious differences among groups. This intuition is based both on prior experience with M&M's as well as the process of reporting their data values and hearing obvious similarities or differences in the corresponding values from their colleagues. For example, when a total for a packet of plain M&M's is twice as high as the total for a packet of peanut M&M's, the students take note. This can lead to rich classroom discussion when statistical findings do or do not corroborate their intuitive hunches. These data can also be more relevant to the students than the textbook examples that are often taken from subject areas unfamiliar to the students.

One limitation to this approach is that the subject matter of M&M's does not naturally translate to the subject matter of interest to the students. For example, my classroom is comprised of students with varying interests from childbirth to gerontology to workplace bullying. This limitation can be offset by having the instructor explicitly bridge the gap from the foundational knowledge built using the M&M's to variables that would be relevant to the students' area of study.

As a concrete example, I could have the students imagine that the packet of M&M's represents a patient. The three types of M&M's represent three different medical conditions. If the total M&M's count for a packet represents the white blood cell count of that patient, I could ask how the results might be interpreted. I have also found that a useful examination question is to ask the students to contextualize the material: specifically, I ask them to express a dependent variable matched with an independent factor that would be relevant to their particular field of interest. In this way, I can ensure the students are engaging with the material at an appropriate level.

Another drawback is that this exercise does not easily provide a rationale for the use of random assignment as might be expected in an experimental design context. However, it does allow for the discussion of that topic, as well as the fact that such randomization is not always possible. It is important to note that the context for my course is not primarily experimental design, but would also encompass quasi-experimental or observational designs that do not allow for the use of randomization as a basis for inference for ethical or other reasons. Nonetheless, these issues can lead to rich discussion about these topics, recognizing that in some instances, the understanding of one-way ANOVA is a means to an end, as it provides an essential bridge for the students to subsequent generalizations of the method, such as factorial ANOVA or ANCOVA.

One final limitation already mentioned is the cost involved in the exercise. While this exercise is adaptable to classes with larger enrollments, the costs will increase directly. Due to this constraint, I would recommend this activity for a class size ideally in the range of 12 to 25. Larger classes could also adversely impact the logistics, primary in data collection. This logistical drawback, however, could be reduced by segmenting a large class into smaller sections of the classroom. Every student within a given section could be given the same type of M&M's, or the exercise could be replicated in its entirety within each section. The section-specific spreadsheets could be concatenated to provide a single class-wide dataset. Of course, the benefit of a large class is the increased sample size.

I have anecdotal evidence that this exercise conveys the relevant concepts in a lasting fashion. A former student commented to me that during her subsequent dissertation research, she would return to our example and ask herself how elements of her dissertation dataset translated to the M&M's demonstration. She asked herself which variable in her dataset corresponded to the M&M counts, and what her grouping variable was that corresponded to the M&M type. She could then proceed through each step of the analysis as outlined in this exercise. I also continue to receive correspondence and have conversations with former students who continue to remember this example.

Overall, this exercise has proven to be a useful, fun, and memorable learning tool for my students. The in-class demonstration can easily be completed in one hour or less, or it can span more than one class period, depending on the depth of coverage desired by the instructor. Out-of-class preparation required on the part of the instructor is not time- or labor-intensive and is relatively inexpensive.

## Acknowledgement

A version of this material was presented as a poster at the Joint Statistical Meetings in Vancouver, BC in August 2010, as “A Motivating Example for Analysis of Variance Using Candies.”

---

## References

- Dyck, J.L., and Gee, N.R. (1998), “A Sweet Way to Teach Students about the Sampling Distribution of the Mean,” *Teaching of Psychology*, 25, 192-195.
- Froelich, A.G., and Stephenson, W.R. (2013), “How Much Do M&M’s Weigh?,” *Teaching Statistics*, 35, 14-20.
- Guidelines for Assessment and Instruction in Statistics Education: College Report (2010), American Statistical Association . Retrieved September 14, 2012 from [http://www.amstat.org/education/gaise/GaiseCollege\\_Full.pdf](http://www.amstat.org/education/gaise/GaiseCollege_Full.pdf).
- Lesser, L.M., and Pearl, D.K. (2008), “Functional Fun in Statistics Teaching: Resources, Research, and Recommendations,” *Journal of Statistics Education*, 16, <http://www.amstat.org/publications/jse/v16n3/lesser.html>.
- Lin, T., and Sanders, M.S. (2006), “A Sweet Way to Learn DoE,” *Quality Progress*, 39, 88.
- Magel, R.C. (1996), “Increasing Student Participation in Large Introductory Statistics Classes,” *The American Statistician*, 50, 51-56.
- Magel, R.C. (1998), “Using Cooperative Learning in a Large Introductory Statistics Class,” *Journal of Statistics Education*, 6, <http://www.amstat.org/publications/jse/v6n3/magel.html>.
- SAS Institute Inc. (2011), *SAS/STAT 9.3 User’s Guide*, Cary, NC: SAS Institute Inc.
- 

Todd A. Schwartz  
3106E McGavran-Greenberg Hall, Campus Box 7420  
Gillings School of Global Public Health  
The University of North Carolina at Chapel Hill  
Chapel Hill, NC 27599-7420  
Email: [Todd.Schwartz@unc.edu](mailto:Todd.Schwartz@unc.edu)  
Phone: (919) 966-7280  
Fax: (919) 966-7285

---

[Volume 21 \(2013\)](#) | [Archive](#) | [Index](#) | [Data Archive](#) | [Resources](#) | [Editorial Board](#) | [Guidelines for Authors](#) | [Guidelines for Data Contributors](#) | [Guidelines for Readers/Data Users](#) | [Home Page](#) | [Contact JSE](#) | [ASA Publications](#)