Journal of Research High School

Volume 2023 (1) 199 - 206
ISSN 2688-3651 [online]

On the Normality of the Distribution of Colors of m&m Candy

Ankita S. Achanta[1 *]

[1]George School, Newtown, PA USA

Received August 31, 2022; Revised April 2, 2023; Accepted, May 6, 2023

Abstract

This investigation explores the shape of the distributions of colors of m&m candy. Statistical tests: kurtosis, skewness, and Shapiro-Wilk were used to assess normality of the distribution of the colors. Kurtosis measures "peakedness" and skewness measures distortion of a distribution. The Shapiro-Wilk test outputs a categorical yes/no, to indicate normality. Distributions of colors were also visually assessed for normality through histograms. This investigation also considers how normality varies based on sample size and different samples of the same size. It was hypothesized that if m&m candy colors are randomly allocated to packets, with an adequately large sample size, they would follow a normal distribution (skewness =0, kurtosis = 3, Shapiro-Wilk test result = yes). The results support that increasing sample size does not make kurtosis and skewness closer to normally distributed with small samples (n<50). Also, there is evidence that normality varies within samples of the same size, based on skewness/kurtosis. Finally, approximate probability density functions (PDFs) were created for each color distribution. With the functions, probabilities of getting a certain number of candies of a given color were calculated. Based on the PDFs, it was found that the probability of getting a packet with only blue m&ms is higher than that of a packet with only red m&ms, though both are highly unlikely. There are certain limitations related to this investigation. It is possible due to the small group sizes that the kurtosis/skewness values did not approach normality. Experimentation with larger sample sizes (group size > 100 each) would facilitate more accurate estimation of probabilities.

Keywords: Normality, Distribution, Kurtosis, Skewness, Shapiro-Wilk test

1. Introduction

M&ms are candy-coated chocolates distributed by Mars, Inc. They come in six colors: blue, orange, green, yellow, red, and brown. Usually, m&ms are distributed in small "fun-sized" packets with, on average, 15.9 candies per packet. This investigation explores the color distribution of the m&m candies. The intent is to evaluate if Mars, Inc. considers consumer color preferences and controls the distribution of each color to appeal to customers. Assuming no deliberate measures taken by Mars, Inc., to control the colors of candy, the distributions should be random and thus normal. The null hypothesis is that Mars, Inc. did not intervene and the colors are normally distributed. The alternate hypothesis is that there was intervention and the distributions are not normal.

The primary aim of this study is to investigate the distribution of each of the six colors of m&ms. It employs statistical techniques to check normality and estimate the probability of obtaining any given number and color of m&m candies in a fun-sized packet.

* Corresponding Author
ankitaachanta@gmail.com

Advisor: Autumn Thayer
athayer@georgeschool.org

According to LaSalle University, m&m color blends were selected via consumer preference tests, which indicate the most attractive assortment of colors to people. Further, LaSalle reports that according to Mars, Inc., on average each packet of m&ms contains 24% blue, 13% brown, 16% green, 20% orange, 13% red, and 14% yellow m&m's.

1.1 Assumptions

Since Mars, Inc. produces all m&m packets, one wholesale bag was purchased instead of many small bags from different locations. It was assumed that the variability in the colors of m&ms purchased across wholesale bags or purchase points is trivial.

To calculate probabilities, a probability density function (PDF) for the distribution must be written. A PDF is "a function of a continuous random variable, whose integral across an interval gives the probability that the value of the variable lies within the same interval" (Oxford Languages, 2022). An assumption was made for the probability calculations. In order to approximate the probability of getting a single number calculated via integration, a correction is needed. For example, to calculate probability that proportion of blue m&ms is exactly 1, the boundaries of the integral were changed to (0.95,1.05).

Also, it was possible that the changes to kurtosis/skewness were not observable due to the small sample sizes of some groups.

2. Materials and Methods

A wholesale bag of fun-sized m&ms, labeled to contain 50 fun-sized packets and distributed by Mars, Inc. was used. An ideal statistical sample surveys 10% of the population as long as it does not exceed 1000 (MV Organizing 2021 para. 2). Since m&m distributes over 10 million packs per day, it would not be reasonable to sample 10% of all candies (about 1 million packs), so a sample size of 50 fun-sized bags was obtained for the study. While the label on the bag stated that there were 50 packets, the bag had 48 packets in reality, so a sample size of 48 packets was used. Microsoft (MS) Excel was used to create data tables and histograms with the frequency of each color in a packet.

The number of each color of m&ms in a fun-sized pack (blue, green, brown, orange, red, and yellow) was manually counted and recorded in MS Excel. Since the total number of candies per packet is variable, percentages (the percentage of the total number of candies in that packet that are the specific color of interest) were calculated to show frequencies of colors.

Normal distributions are perfectly symmetrical curves where the mean, median, and mode are equal. The total area under the curve of a standard normal distribution equals 1. Such distributions are important because they can be used to calculate the probability of a random variable taking a certain value or range of values. For example, the probability that a bag of fun-sized m&ms will contain 4 or more blue candies can be calculated using the probability density function (equation for the standard normal curve) for the random variable of the number of blue candies per packet. The probability density functions (PDF) were estimated for each color distribution to predict such probabilities.

Normality was tested in three different ways: kurtosis, skewness, and the Shapiro-Wilk test. An MS Excel add-in called Xrealstats (Real Statistics, 2022) was used to calculate these statistics and conduct the test.

"Kurtosis is a statistical measure that defines how heavily the tails of a distribution differ from the tails of a normal distribution" (Corporate Finance Institute,). In other words, kurtosis measures the "peakedness" of a distribution. The kurtosis of a normal distribution is always equal to 3.

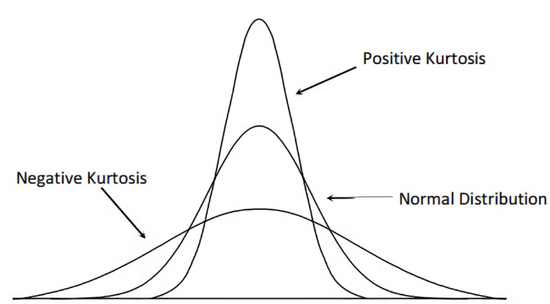

Figure 1: Examples of distributions with varying kurtosis (Glenn, 2015)

Positive kurtosis (leptokurtic) refers to a distribution with a higher kurtosis than a normal distribution (kurtosis > 3). Negative kurtosis refers to a distribution with lower kurtosis than a normal distribution (kurtosis < 3). As shown in *Figure* 1, the distribution with a positive kurtosis is far more peaked than a normal distribution, and the normal distribution is more peaked than the one with a negative kurtosis. Also, data sets with high kurtosis have heavy tails, or outliers.

In MS Excel, the kurtosis formula is modified so that the kurtosis of a normal distribution is zero. To calculate kurtosis, standard deviation must be calculated first. Then, the following formula may be used:

$$\left\{\frac{n(n+1)}{(n-1)(n-2)(n-3)}\Sigma\left(\frac{x_j-\bar{x}}{s}\right)^4\right\}-\frac{3(n-1)^3}{(n-2)(n-3)}$$

Note: n = number of terms in data set; x_i= observed data point; x-bar = sample mean; s = sample standard deviation

"Skewness refers to a distortion or asymmetry that deviates from the…normal distribution." The skewness for a normal distribution is equal to zero. Negative values for skewness indicate data that are skewed left, and positive values indicate data that are skewed right (Investopedia, 2022).

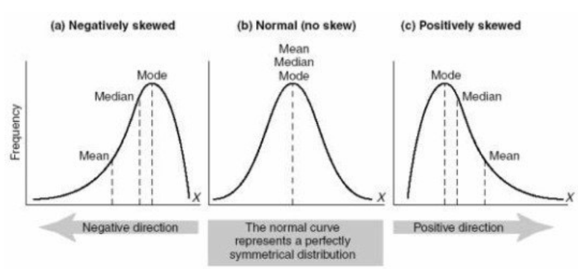

Figure 2: Examples of distributions with varying skews (Hogan, 2022)

Figure 2 shows how skewed distributions differ from normal distributions. Skewed data is not symmetric, and the mean, median, and mode are not equal. MS Excel's skewness formula:

$$\frac{n}{(n-1)(n-2)}\Sigma\left(\frac{x_j-\bar{x}}{s}\right)^3$$

Note: n = number of terms in data set; x_i= observed data point; x-bar = sample mean; s = sample standard deviation

Finally, the Shapiro-Wilk test for normality outputs a categorical "yes/no" to indicate if the data is normal. The null hypothesis is that the variable is normally distributed, and the alternative hypothesis is that it is not normally distributed. The test produces a W statistic, along with a p-value. Testing at the 5% level of significance, if the p-value is less than 0.05, the null hypothesis is rejected and there is sufficient evidence to support that variable is not normally distributed (SPSS Tutorials, 2022).

These three tests are used to determine the normality of the groups of bags because they are readily available through Xrealstats.

For visual analysis, graphs were created with the data through MS Excel. Box plots were made for groups of bags. Also, histograms with normal curve overlays were created to check normality and write the approximate probability density function of each color.

The study investigated the following questions of interest:

<u>*Are the distributions of each color across packets normal?*</u>

Xrealstats was used to create a histogram with a normal curve overlay to visually assess normality, in addition to the kurtosis, skewness, and Shapiro-Wilk calculations. Further areas of inquiry that arise from this question are the following:

<u>*How does normality vary based on different samples of size 5 and 10 bags in each group?*</u>

Kurtosis, skewness, and Shapiro-Wilk test results of the first 10 groups of 5 bags for each color (bags 1-5, 6-10, etc., the last group was bags 44-48) will be compared to assess how normality varies based on different samples of the same size. The same tests will be used to compare variation in the normality of the 5 groups of 10 bags for each color.

<u>*How does normality vary based on the sample size (number of fun-sized bags)?*</u>

The first group of 5, 15, 25, 35, and 45 bags for each color will be taken from the data. The kurtosis, skewness, and Shapiro-Wilk test results for each color distribution will be compared to understand how normality changes.

Can the probabilities of obtaining certain numbers of certain colors be calculated with the estimated probability density functions?

The probability density functions of the color distributions will be estimated with sample data, and probabilities will be calculated using integration.

3. Results and Discussion

The raw data and percentages for the counts of m&m colors in each packet are available upon request from the first author. Further, kurtosis, skewness, box plots, and Shapiro-Wilk results for groups of 5/10 bags may also be requested.

3.1 Normality of the distribution of each color

Xrealstats was used to create the following histograms. The total area of the bars in the graph was used to create a normal curve overlay. Note that due to the bin sizing in the histograms below, the graphs show a bin less than zero; this is an artifact of the software used.

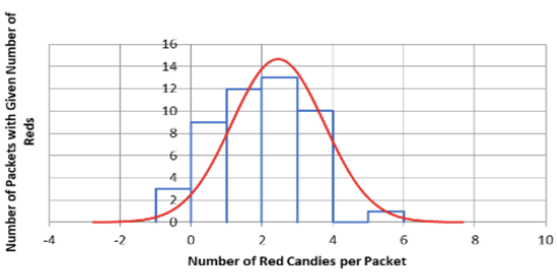

Figure 3: Red Color Distribution in 48 Packets

Visually, red is a near-normal distribution. The histogram is approximately symmetric, and the mode is around three red m&ms in a packet. If there were two or three packets with five red m&ms in them, the histogram would be even more symmetrical.

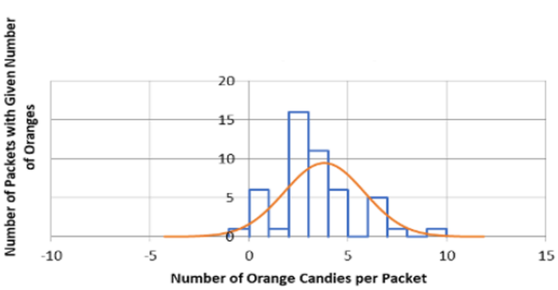

Figure 4: Orange Color Distribution in 48 Packets

The orange histogram looks slightly right skewed and is not as symmetrical as red. This is because the mode occurs at 2 orange candies per bag, and the distribution has larger values that are potential outliers (9 and 10). Also, the curve underestimates the mode of the histogram. The orange distribution is not normal.

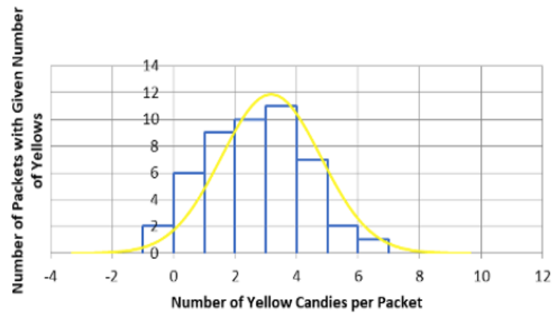

Figure 5: Yellow Color Distribution in 48 Packets

Yellow is the most near-normal distribution of the six colors. The curve seems to accurately represent the histogram, which is close to symmetrical and does not have any visible outliers. The mode is around three yellow candies in a packet.

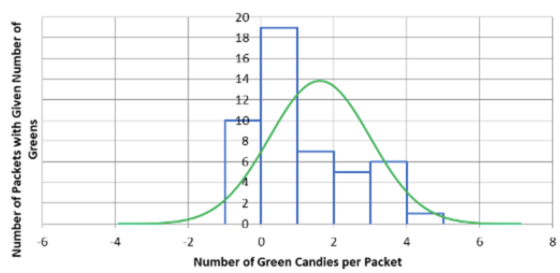

Figure 6: Green Color Distribution in 48 Packets

The green distribution is right-skewed. This may be because there were relatively fewer green candies per bag than the other colors, and with smaller sample size, the distribution showed a skew. The histogram shows the mode to be around 1, and the highest value of green m&ms per bag is 5, quite small compared to red, orange, and yellow. The curve does not seem to adequately represent the histogram, as the peak is far below the modal peak. Green is not normally distributed.

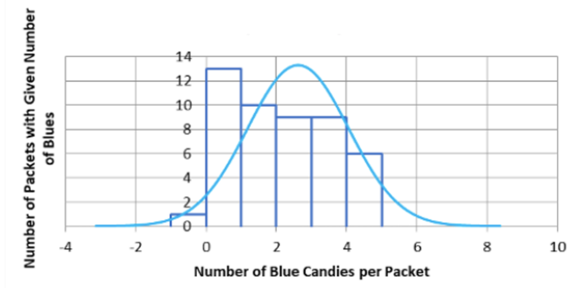

Figure 7: Blue Color Distribution in 48 Packets

The blue distribution has a mode around 1 and does not have more than 5 candies per bag. This histogram is uniform, but the curve does not represent the bars. The peak of the curve is around three, while the true mode is around one, so the blue distribution is not normal according to this data.

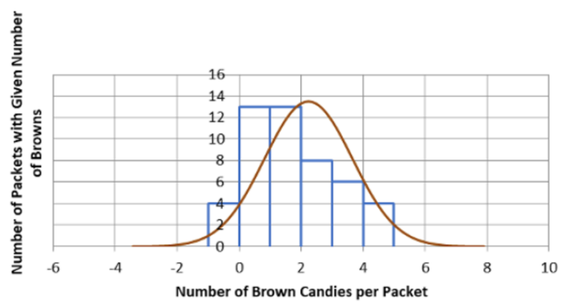

Figure 8: Brown Color Distribution in 48 Packets

The brown distribution is almost symmetrical. The modes are from 1-2, indicating that brown is a relatively less frequent color to occur in bags of m&ms. The curve almost matches the histogram, but because the bars are not perfectly symmetrical, brown is not normally distributed.

3.2 Variation in the normality of different samples of size 5 and 10 bags each

Observations from *Table* 3 include that the kurtosis of red m&ms in groups of 5 bags varies from close to zero to almost three. This means that some of the groups of red are close to normally distributed (0) while others are slightly peaked (3). The 10th group of red bags has the most symmetrical box plot, but most of the plots are not approximately normal.

In orange, the kurtosis scores vary from -2.79 to 4.30, a wide range. Since many of the orange scores were negative, the orange distribution likely has lighter tails than a normal distribution, meaning that most of the values are not extremely far from the mean. None of the orange box plots seem to be perfectly normal, and many are skewed or have outliers. These results may be discordant due to the small sample size of 5 bags (not large enough to be a representative sample).

The yellow kurtosis scores vary from -2.95 to 2.03 and are not near normal. Some of the box plots- 1,6, and 7 are symmetrical.

Green scores range from -3.04 to 2.06. Since there were very few green m&ms per packet, the green color distribution is expected to be the least normal of all of the colors due to the small sample size. This is reflected in the box plots as many are skewed or have wide ranges.

Blue kurtosis scores range from -2.69 to 1.21. There is one kurtosis score (0.16) that is close to normally distributed for the blue color. The box plots for the blue distributions have wide ranges and many are skewed.

Finally, the kurtosis scores for brown range from -3 to 4.7. Some distributions are very peaked for brown while others are not. Some of the box plots for the brown groups of 5 are symmetrical and look near normal. Kurtosis varied greatly among the distributions of colors for different groups of the same sample size. Also, skewness varied, but not as much as kurtosis. There was at least one group of bags with skewness of close to zero in each color, but there were more groups with skewness of around 1 overall.

For the distributions of groups of ten bags, the kurtosis varied less as the ranges were not as wide as the kurtosis of groups of 5 for each color. Across all colors, there were more negative values of kurtosis, so the distributions had lighter tails when the sample size was increased in each color. An increase in sample size decreases variability, so it makes sense that the distributions will have lighter tails with a larger sample size. However, the skewness seemed to be greater in groups of 10 bags, which was not expected. Overall, there were not as many values for skewness that were very close to zero. The box plots for each color were closer to symmetrical, especially the blue color distribution.

According to the Shapiro-Wilk test, the majority of the groups of 5 bags were normally distributed. However, results of this test were often discordant with those obtained using kurtosis and skewness. This may be due to the small sample sizes, and the experiment could be repeated with larger groups of bags for better results. This is a key consideration for analyzing the data as with smaller sample sizes, the Shapiro-Wilk test may not be accurate.

3.3 Variation in normality based on the sample size (number of fun-sized bags)

For the red distribution, the kurtosis gradually increases as the sample size increases. The skewness gets closer to zero, however, the group of five bags has a skewness of -0.02, which is closer to zero than the group of 45 bags, with a score of 0.06. Also, the Wilk-Shapiro test is not a good indicator of normality on its own, as all of the groups of bags were normally distributed according to the test.

The orange distribution's kurtosis gets closer to three (more peaked) as the sample size increases, with the exception of the first five bags as they start with a kurtosis score of close to three. Also, the skewness score gets closer to zero but does not decrease each time the sample size increases.

In the yellow distribution's results, kurtosis becomes closer to zero as sample size increases, getting closer to normal. Skewness is not very high and fluctuates slightly, it does not periodically decrease as sample size increases.

The green distribution's kurtosis scores get closer to normally distributed as the sample size increases. The skewness scores for green mostly approach zero as the sample size increases, and the closest score to zero is observed in the largest group of bags.

The blue kurtosis and skewness scores fluctuate as sample size increases, not showing a score that is close to normally distributed.

Finally, the brown group's kurtosis gets closer to normally distributed but is still far from the perfectly normal score of zero. The skewness for brown fluctuates.

3.4 Estimation of probabilities of obtaining certain numbers of certain colors candies

Estimated probability density functions for each color:

Probability Density Function (Duke Edu 2022):

$$F(x) = \frac{1}{\sigma * \sqrt{2\pi}} * e^{\frac{\left(-(x-\mu)^2\right)}{2\sigma^2}}$$

μ and σ for populations are unknown, so they will be estimated as shown below:

σ ≅ s (sample standard deviation)

μ ≅ x-bar (sample mean)

Estimated PDFs:

Red PDF: x-bar = 0.153; s= 0.076

$$F(x) = 5.001 * e^{\frac{\left(-(x-0.153)^2\right)}{0.0126}}$$

Orange PDF: x-bar = 0.238; s = 0.122

$$F(x) = 3.27 * e^{\frac{\left(-(x-0.238)^2\right)}{0.0297}}$$

Yellow PDF: x-bar = 0.198; s = 0.099

$$F(x) = 4.029 * e^{\frac{\left(-(x-0.198)^2\right)}{0.0196}}$$

Green PDF: x-bar = 0.101; s = 0.0897

$$F(x) = 4.447 * e^{\frac{\left(-(x-0.101)^2\right)}{0.016}}$$

Blue PDF: x-bar = 0.165; s = 0.0911

$$F(x) = 4.379 * e^{\frac{\left(-(x-0.165)^2\right)}{0.016}}$$

Brown PDF: x-bar = 0.141; s = 0.092

$$F(x) = 4.336 * e^{\frac{\left(-(x-0.141)^2\right)}{0.016}}$$

Using the PDFs above, some probabilities were calculated. First, the probability of getting one packet with all blue m&ms was calculated. Also, the probability of getting a packet full of red m&ms was calculated to be compared to the probability of getting all blues, since red candies were more frequent than blue ones.

Probability of getting a packet with all blue m&ms (Probability that proportion of blue m&ms in a randomly selected packet = 1)

$$\int_{0.95}^{1.05} 4.379 * e^{\frac{\left(-(x-0.165)^2\right)}{0.016}} dx = 8.27 * 10^{-19};$$

considered highly unlikely

Note: correction for the boundaries of the integral: (1-0.05), (1+0.05)

Probability of getting a packet with all red m&ms (Probability that proportion of red m&ms in a randomly selected packet = 1)

$\int_{0.95}^{1.05} 5.011 * e^{\frac{(-(x-0.153)^2)}{0.0126}} dx = 5.004 * 10^{-24}$; also considered highly unlikely

It was found that the probability of getting a packet with all blue m&ms is relatively more likely than getting a packet of all red m&ms, however, both outcomes are considered to be highly unlikely.

4. Discussion & Conclusion

Based on the histograms, the closest to normally distributed color of m&ms is yellow, and red is also approximately normal. Further, at a constant sample size, kurtosis and skewness vary. Since kurtosis and skewness are random variables themselves, they have their own distributions. Further investigation focusing on kurtosis and skewness could examine if the distributions of kurtosis and skewness are the same as the distribution of the random variable used to calculate them.

Sample size does not make kurtosis and skewness closer to normally distributed based on these results. In some colors, there was an observed pattern between sample size and an increase in kurtosis, but many of the groups of bags were still far from normally distributed. It is possible that since the group sizes were small (5-45 bags), the changes to kurtosis and skewness were not observable. This experiment could be repeated with larger sample sizes to detect a more pronounced change in kurtosis by sample size.

The results demonstrate that the Shapiro-Wilk test lacks concordance with kurtosis and skewness scores. The kurtosis and skewness scores for many groups were not normally distributed, but the test statistic indicated that they were. Other groups that seemed to have kurtosis and skewness scores near normally distributed were not normal according to this test.

Due to the assumptions made and integral boundary corrections, this investigation has some limitations. Experimentation with larger sample sizes would be valuable in refining the probability density functions, thus calculating more accurate probabilities.

There may have been some intervention by Mars, Inc. in controlling the proportions of m&m colors as many of the colors were not normally distributed. Thus, these findings support the alternate hypothesis.

References

(n.d.). Retrieved July 18, 2022, from http://www.cs.uni.edu/~campbell/stat/prob9.html

(n.d.). Retrieved July 18, 2022, from https://sites.nicholas.duke.edu/statsreview/.

Chen, J. (2022, July 8). *What is skewness in Statistics?* Retrieved July 18, 2022, from https://www.investopedia.com/terms/s/skewness.asp#:~:text=Skewness%20refers%20to%20a%20distortion,is%20said%20to%20be%20skewed.

Davis, B. (2021, April 30). *Home*. Retrieved July 18, 2022, from https://www.mvorganizing.org/what-is-the-minimum-sample-size-for-experimental-design/#:~:text=1%2C200%20is%20desirable).

Excel: How to calculate kurtosis with Kurt Function. (n.d.). Retrieved July 18, 2022, from https://indxar.com/contents/1404#:~:text=*The%20kurtosis%20of%20the%20normal,to%20be%200%20in%20Excel.

Glen, B. (2015) *Kurtosis Distributions*, Image, Stack Exchange,, https://stats.stackexchange.com/questions/84158/how-is-the-kurtosis-of-a-distribution-related-to-the-geometry-of-the-density-fun

Hogan, R. (2020, June 16). *Probability distributions for measurement uncertainty*. isobudgets. Retrieved July 18, 2022, from https://www.isobudgets.com/probability-distributions-for-measurement-uncertainty/

Information Technology Laboratory. (2022, June 27). Retrieved July 18, 2022, from https://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm.

Kurtosis. (2022, May 7). Retrieved July 18, 2022, from https://corporatefinanceinstitute.com/resources/knowledge/other/kurtosis/.

Measurement statistics: Color distribution in M&MS

- *la salle university*. (n.d.). Retrieved March 26, 2023, from http://www1.lasalle.edu/~prushan/Measurement%20 Statistics-lab2.pdf

Microsoft. (n.d.). Retrieved July 18, 2022, from https://support.microsoft.com/en-us/office/skew-function-bdf49d86-b1ef-4804-a046-28eaea69c9fa
Oxford languages and google - english. Oxford

Languages. (n.d.). Retrieved July 18, 2022, from https://languages.oup.com/google-dictionary-en/.
SPSS Shapiro-Wilk Test – quick tutorial with example. SPSS tutorials SPSS ShapiroWilk Test Quick Tutorial with Example Comments. (n.d.). Retrieved July 18, 2022, from https://www.spss-tutorials.com/spss-shapiro-wilk-test-for-normality/#:~:text=The%20null%20hypothesis%20for%20the,sample%20from%20a%20normal%20distribution

TodayShow. (2016, March 3). *Check out these 3 new peanut M&M's flavors, plus 5 M&M's fun facts*. TODAY.com. Retrieved July 18, 2022, from https://www.today.com/food/5-things-you-didn-t-know-about-m-m-s-t77591

Zach. (2021, March 2). *Can kurtosis be negative?* Statology. Retrieved July 18, 2022, from https://www.statology.org/can-kurtosis-be-negative/.

Zaiontz et al., Retrieved July 18, 2022, from https://www.real-statistics.com/free-download/real-statistics-resource-pack/.