Ripening Fruit

November 8, 2022

1 STATS 604: Ripening Fruit

1.1 Team Members: Alex Kagan, Daniele Bracale, Josh Wasserman, Xinhe Wang, and Yash Patel

We present herein the findings from our explorations on the ripening of bananas. The specific experiment of interest was to see if bananas' ripening is affected by exposure to different "accompanying" fruits, specifically apples and tomotoes. In this report, we present some our experimental design along with the raw data processing pipeline, statistical analyses, and conclusions. The report is, therefore, broken into the following sections:

- Experimental Design
- Data processing
- Permutation test analyses
- Conclusion

Each section is annotated with the corresponding code and figures in the sections below along with associated exposition.

1.2 Introduction

Fruits in general ripen through the conversion of starch to sugars. Ideally, we'd be able to measure this conversion process chemically, but we don't have access to refractometers. We can define the "true" ripeness as being this ratio of sugars to starches. So, we use a number of proxies that correlate strongly with this ripeness.

Starch, a complex carbohydrate, is broken into simple sugars through a reaction with ethylene. However, this release of ethylene also results in a number of other reactions that are clearly visible. Specifically, ethylene additionally breaks down pectin, which is responsible for the structural integrity of the banana. This is why bananas also grow softer as they ripen. Similarly, the release of ethylene results in the breakdown of chlorophyll, which then gives rise to the characteristic yellow and eventual browning color of bananas. This is surprisingly the same process that happens with leaves, although not all the same pigments are present in banana skins, which is why there is only really a progression from yellow to brown.

So, we want to conceptually measure:

- Color
- Firmness

1.3 Experimental measures

Color: The color of bananas characteristically changes in two ways in ripening: becoming more yellow and subsequently developing brown spots. We, therefore, have two measures towards this end. For both, we took a picture of the banana on a white piece of paper, placing a "registration" grey card (printed) to serve as a fixed normalization point for white balance in post-processing. In post-processing, after normalizing colors, we segmented the banana in the image and ran a connected components segmentation analysis to find the "brown spots," computing the area as the number of pixels (as a proportion to the total area). For the remainder of the banana, note that (idealized) green in RGB is (0, 255, 0), whereas (idealized) yellow is (255, 255, 0), so the average red channel should serve as a rough proxy for the ripeness (w.r.t. color). Part of the reason for this bifurcated analysis is to account for the fact brown does not have a greater red value than yellow, meaning there is an inflection in the red channel with respect to ripeness which would skew results were the average run pre-segmentation.

Firmness: For this, on the final day of the experiment, we unpeeled the banana and cut off a 2" segment from the middle. This segment was then glued on a white piece of paper. We then placed a fixed weight (a cast iron pan) atop it for a fixed period of time (5 seconds). We took a picture before and after exposure to the weight and measured the percentage change in the resulting number of pixels as a measure of how much the banana got "squished."

1.4 Raw Data

Data were collected in the following fashion:

1. Take a picture of each banana as described in the "color" portion of the experimental measures section
2. Randomly place N/3 bananas in each of three bags
3. Place nothing additional in one bag, 3 apples in another, and 3 tomatoes in the last
4. Leave these bags for 5 days, taking a picture each day as described in the "color" portion of the experimental measures section
5. Measure the above listed "experimental measures" on the final day

The raw data are available in the corresponding Day_* folder.

1.5 RGB Analysis

In this section we process the data in the following way:

1. Collect all the bananas' pictures for each day.
2. For each picure, make a selection, i.e. select the banana only and put it in a completely white background. This helps in having pixels not corresponding to a banana location be equal to $(255, 255, 255)$.
3. Read each picture from previous step in R and count the number of non-white pixels, i.e. the pixels corresponding to the banana.
4. Transform the picture into "greyscale" in order to have a unique channel and being able to select the brown spots with a uniform measure. Count the number of pixels with values less than 0.5 (recall that in R the maximum RGB value is 1 instead of 255) and divide this number by the number of non-white pixels in step 2. The result is the proportion of black spots in the

banana. Then 1 minus this proportion is the proportion of yellow (non-brown) spots. Take also the average of the red channel values corresponding to the non-brown spots.

5. Collect the data for the proportion of brown spots, yellow spots and the average of red spots for each banana and output a csv file. For each banana we also add a variable that specify the treatment: C for control, A for apple and T for tomato.

1.6 Squish Analysis

To perform the analysis, there are two stages of the processing:

- Segment bananas out from the remainder of the image
- Count the masked pixel area of this segmented section

We perform the first by running an edge detector algorithm followed by a watershed algorithm to extract out the complete mask of the object. Note that there are some nuances about how this edge detection is implemented: a vanilla Canny edge detector fails to work because of small issues with leaving "gaps" in the object boundary.

The final "squish ratio" is defined as:

$$\frac{\%pixels_{after}}{\%pixels_{before}}$$

1.7 Permutation Test

Justified by the random allocation of bananas, we performed permutation tests for each of our treatment conditions.

First, we tested if tomatoes accelerate the ripening process of bananas with the following competing hypotheses:

$$H_0: \text{ Tomatoes do not accelerate the ripening process of bananas}$$

verses

$$H_a: \text{ Tomatoes accelerate the ripening process of bananas.}$$

For this test, we only used the data of tomato and control groups, randomly permuted the treatment labels for 20,000 times, and used the difference-in-means of tomato verses control as our test statistic. More specifically, we computed the contrast between the mean outcomes under each treatment label for each permutation. Then we compare the observed difference-in-means with the permuted distribution and find out if our observations are unusual under the null hypothesis.

The story is similar with testing the effect of apples. The competing hypotheses we are interested in are

$$H_0: \text{ Apples do not accelerate the ripening process of bananas}$$

verses

$$H_a: \text{ Apples accelerate the ripening process of bananas.}$$

We only used the data of apple and control groups. We permuted the treatment labels for 20,000 times and for each permutation, computed the difference-in-means of apple verses control as the test-statistic.

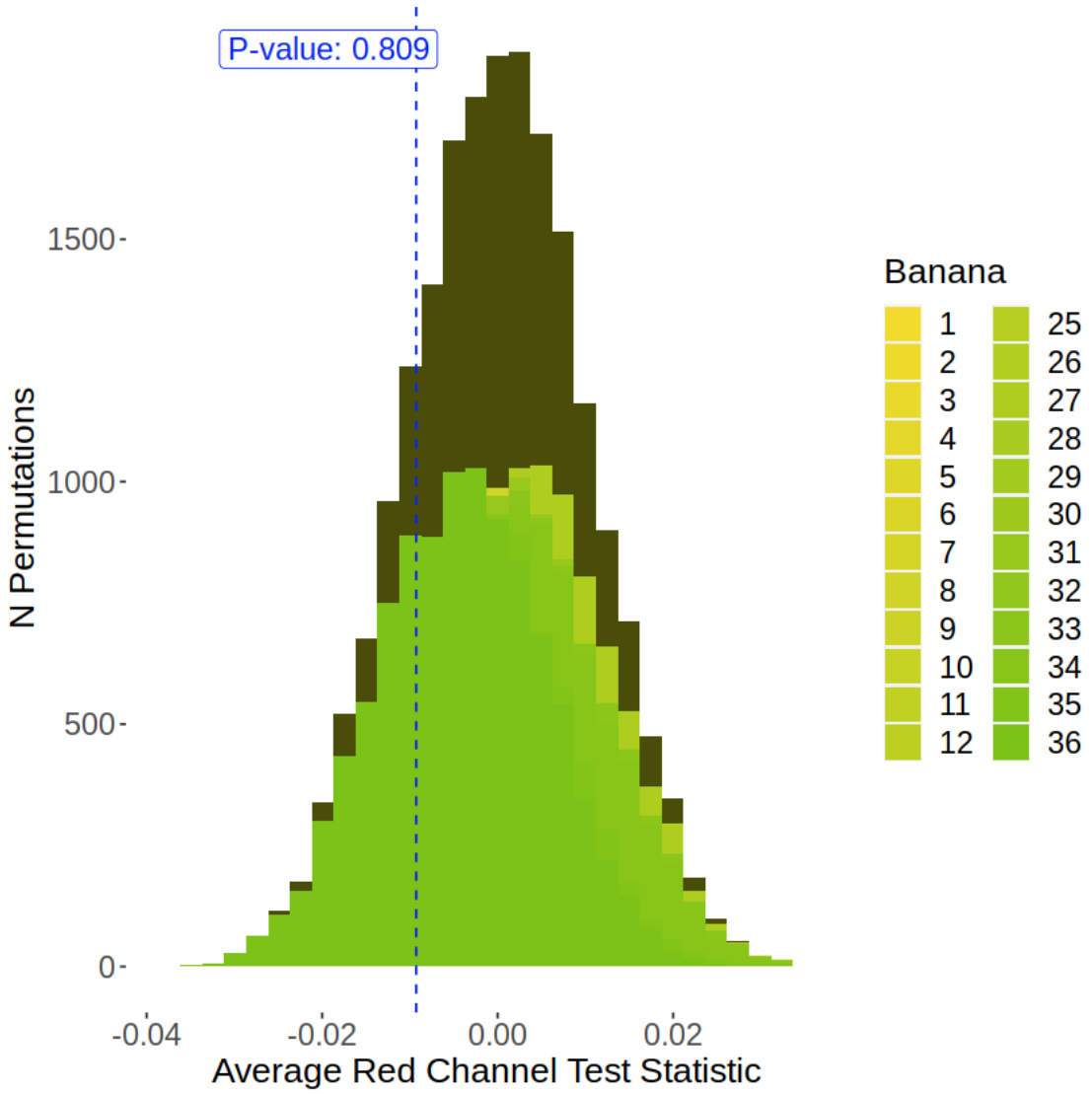
The computed p-values are shown in the following table, which we will explain in more detail in the next section.

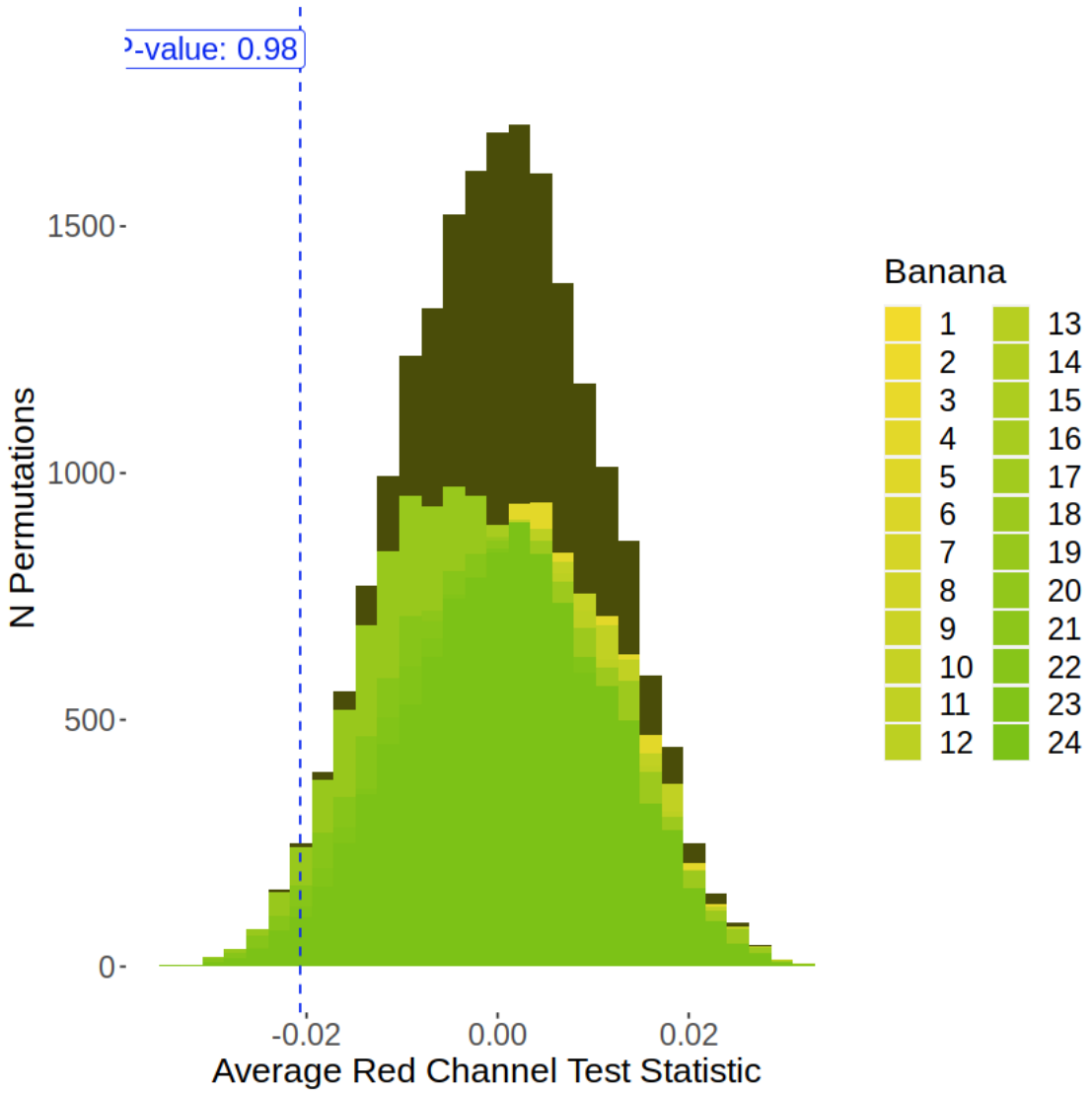
	Average.R	Brown.Percentage	Squish.Ratro
tomato	0.80945	0.38175	0.2371
apple	0.98025	0.90625	0.2655

1.7.1 Results

No permutation test yielded a significant p-value. The two smallest p-values came from the tests for the squish ratio test statistic: the permutation test of the bananas in the control environment and the bananas stored with apples returned a p-value of 0.24, and the permutation test of the control bananas and the bananas stored with tomatoes returned a p-value of 0.27. Using the percentage of brown spots test statistic, the test of the tomato treatment yielded a p-value of .382, and the test of the apple treatment yielded an even larger p-value of .906. The tests using average red channel test statistic produced the highest p-value on average across the tests of the two treatments: the test of the tomato treatment yielded a p-value of 0.81, and the test of the apple treatment yielded a p-value of 0.98.

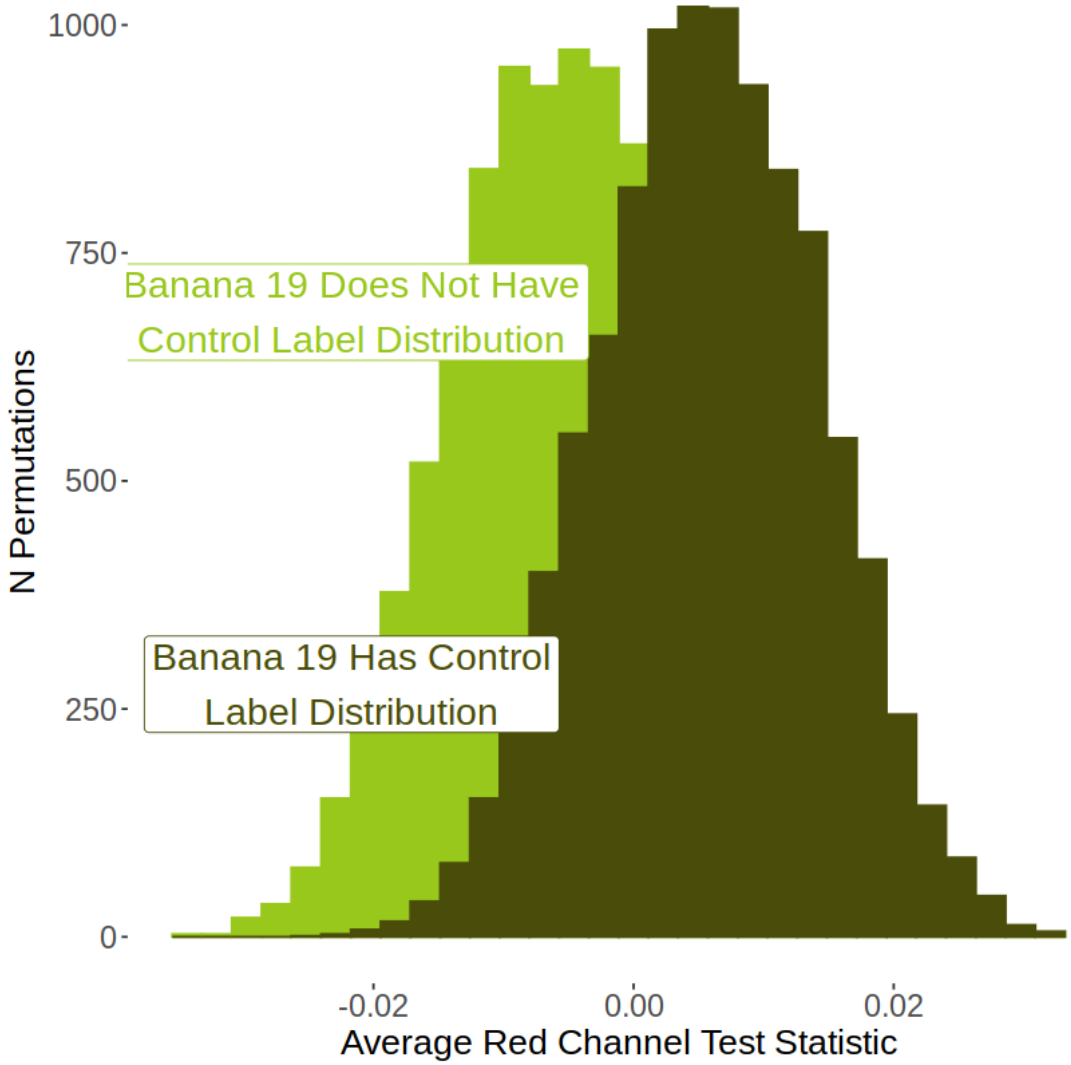
Despite the overwhelmingly large p-values, visualizations of the permutation distributions for the average red channel tests may reveal useful information about the experimental design. The histograms below showing the results from these permutation tests display the distribution of test statistics from all permutations in brown, then in different shades of yellow and green, the distribution of test statistics from permutations where each banana was labeled as a treatment banana. For the test of the tomato treatment, each of these "banana-specific" permutation distributions looks to have the same general shape as the overall permutation distribution. For the test of the apple treatment, however, there appears to be one "banana-specific" distribution whose shape differs somewhat notably from the distribution of all the permutations, which is made clearer in the histogram below.


P-value: 0.809
1500
1000
500
0
N Permutations
-0.04
-0.02
0.00
0.02
Average Red Channel Test Statistic
Banana
1
2
3
4
5
6
7
8
9
10
11
12
25
26
27
28
29
30
31
32
33
34
35
36



This banana had a change in average red channel between days 1 and 5 of -0.02, by far the lowest of all the bananas and theoretically impossible under the monotonicity assumption underlying our decision to use average red channel as a test statistic. This forces us to look more closely at few things, the first being whether our method of calculating average red channel needs to be improved. Do we need to adjust our method of calibrating the amount of light in each photo? Should we normalize the red channel of each pixel in a photo before taking the average of the pixels displaying the banana? Would these be possible, they should reduce the possibility of returning a negative value for this measurement. Even if the possibility of returning a negative value were completely removed though, this result also makes us consider whether a difference-in-means test statistic is appropriate for this measurement. Based on the measurements obtained in this experiment, a few bananas lagged behind noticeably in change in average red channel, which could be due to measurement error, or they might have been outliers. Were their lag not to be due to measurement error, it may be more appropriate to use an outlier-robust test statistic such as the difference-in-

medians in future tests of the average red channel measurement.



1.8 Discussion

1.8.1 Limitations of the experiment design

Our statistical analysis was conducted under several important assumptions about the experiment design; however, most of them were likely violated to some extend. Here, we present only the most important of these assumptions together with the reasons why they could be violated. To begin, the Stable Unit Treatment Value Assumption (SUTVA) assumption was violated: each of the bananas assigned to a single "treatment" were placaed in the same bag. As a result, these bananas, in addition to be potentially being impacted by the treatment of interest (exposure to the apple or tomato), likely was exposed to a great deal of ethylene simple from the presence of the remaining bananas, potentially overwhelming the treatment effect of interest and thus resulting in the same

ultimate measured "ripeness." In addition, the bananas assigned to the same treatments were not uniform: bananas were scattered throughout the bag, thus placing some closer to the tomato or apple than others.

In addition to these experimental shortcomings, there were likely some measurement issues, as indicated above in the report. Specifically, the RGB measurements seem to have issues of consistency across days: this is likely due to the ambient lighting conditions in the room, which affects the camera auto-exposure and white balance. While we tried to mitigate this using the fixed camera apparatus largely illuminated by a lamp rather than relying on inconsistent ambient light, it seems likely that there were still external undesired lighting variations. Future work may use a better camera calibration system. Further, in the squish test, the alignment of the weight (cast iron) and the subjected banana slice was not necessarily uniform across trials: while it is reasonably straightforward to place the cast iron roughly consistently between trials, because of the size discrepancy, the cast iron tends to become imbalanced in the squishing process. As a result, it needs to be manually "guided" to remain in the same horizontal position (centered on the banana). Since this guiding was done by hand, it is highly likely that this measurement was inconsistent across runs, although no systematic issue likely arose across the treatments, since squishing was conducted across bananas in a random order.

1.8.2 Further work

For future investigations, we would want to address the aforementioned shortcomings in the ways suggested. That is, we would separate treatments for the bananas into separate bags, placing the apple or tomato per bag. Images would be ideally calibrated better than in the current apparatus. Finally, using a smaller weight for the squish test where manual guiding would not be necessary would greatly improve the fidelity and signal in the corresponding data. In addition to these, there are some extensions we would investigate were we given an expanded budget. Specifically, we would measure the ripeness of bananas by their reaction to iodine, as stated in our pre-analysis. Such a reaction results in a more apparent visual marker of ripeness, which would contain more signal than the proxies of the RGB or black spots we are currently using.