

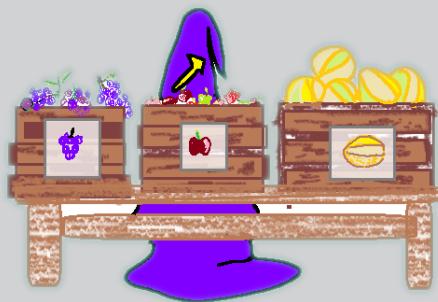
# 3

## Linear data

Inspired by some particularly tasty local produce, our wizard Lineum decided to sell fruit salad from a family recipe at the next farmers' market.

This would mean taking the bus, boarding at the 2nd stop and exiting on the last stop of the route.

Lineum wanted to arrive to the market by 8 a.m. to set up, but also not arrive too early and end up with warm fruit salad.



Bus A	6:00	6:16	6:18	6:24	6:30	6:34	6:44	7:06
Bus B	7:10	7:24	7:26	7:33	7:45	7:49	7:59	8:21

From the schedule, Lineum was disappointed by the only option: leave at 6:16 a.m. and arrive an entire hour early.

Warm fruit salad, what a disappointment!

Discouraged but determined, Lineum turned to the family's fruit salad recipe that makes 4 servings.

### Magic-Spice Fruit Salad

4 apples chopped into 1/2 inch squares

1/2 cantaloupe sliced and sectioned into 1 inch pyramids

3/2 cups of grapes (no seeds)

1/2 honeydew melon sliced and sectioned

Magic Spices, Honey drizzle

All food sales needed a nutrition label to sell at the market, but the Food and Drug Administration (FDA) nutrition chart [5] did not have any mention of the fruit salad!

	Calories	Fat	Sodium	Fiber	Carbs	Protein
1 Apple	130	0g	0mg	5g	34g	1g
1 Banana	110	0g	0mg	3g	30g	1g
1 Cantaloupe	200	0g	80mg	4g	48g	4g
1 cup Grapes	90	0g	15mg	1g	23g	0g
1 Honeydew	500	0g	300mg	10g	120g	10g

Lineum was ready to give up.

## The data science dilemma

Lineum's story is quite typical. We often do not have the data we want. This can be a disappointment, but it does not need to be. In some magical conditions we can actually create new information. This book, and much of linear algebra, is dedicated to the process of creating new information from the data we have. These days the software we use can do the arithmetic for us leaving some people to think that the job of data science is to "push the buttons" until new data comes out. And there certainly are many shiny buttons you will find people are pushing.



Do not buy into this misunderstanding of data science! As we explore in this chapter, it takes thought and practice to know what buttons produce **information**, not merely data.

Just then, Matica, Lineum's sibling, showed up. Matica's magical power was math. They looked over the challenge and suggested a strategy. Maybe they could use the table to create the missing entry for the fruit salad. Matica told Lineum that to work out the calories they could scale the amounts in the calorie column by the amount of each fruit in the recipe and then add up the results.

4 apples	$\times 130 \text{ calories/apple}$	= 520 calories
0 Bananas	$\times 110 \text{ calories/Banana}$	= 0 calories
1/2 cantaloupe	$\times 200 \text{ calories/cantaloupe}$	= 100 calories
3/2 Grape cups	$\times 90 \text{ calories/grape cup}$	= 135 calories
+ 1/2 honeydew	$\times 500 \text{ calories/honeydew}$	= 250 calories
		<hr/>
		1005 calories

Success! Each recipe had 1005 calories, meaning each serving has  $1005/4 \approx 250$  calories. As Matica left the room, Lineum applied the same algorithm to determine the amount of dietary fiber, sugar, protein, etc. in each serving. Now Lineum could properly label each salad.

Suddenly empowered by this new magic, Lineum looked back at the trip plan, paused, and smiled. Why leave at 6:16? Instead, take 1/2 of the 6:16 bus and 1/2 of the 7:24 a.m. bus to arrive at the market at time

$$\frac{1}{2}7:06\text{am} + \frac{1}{2}8:21\text{am} = \frac{1}{2}(7 \cdot 60 + 6) + \frac{1}{2}(8 \cdot 60 + 21) = 7:44\text{am}$$



Just enough time to set up and sell fresh magical fruit salad!

**Moral:** linear combinations are powerful magic but some uses are not approved for non-wizards.

# Linear Data

*How linear algebra informs us*

*Emily J. King  
James B. Wilson*

