## A FRESH SQUEEZE ON DATA

PROBLEM SOLVING WITH DATA

0

•

-

0



## A FRESH SQUEEZE **ON DATA**

### PROBLEM SOLVING WITH DATA



One sunny day, Clara runs home and tells her family:

That's a nice idea! Let me know if you need any help!

**BIG News!** I want to raise money for the hospital. I'm going to open a lemonade stand!

Can I help?

Sure, you're great at making signs!



The next day, Clara and Alex wake up early. They set up a table with a sign in front of their house. They lay out a jug of fresh lemonade and a stack of cups.

No one bought lemonade from them for two hours. Alex looks sadly at his sister and says,



-emongde

## Why didn't we sell anything, Clara?

Clara begins to think about what she might have missed in her planning. What could she have done differently to bring more people to their lemonade stand?

Suddenly, Clara has an idea.

Clara thinks about how her mom solves problems at work. Her mom is a data scientist and works at a company that uses data to solve all kinds of problems.

· ,



Maybe **data** can help us solve our problem!

What is data and how can it help?

> Data is information. It's all around us. We create data every day and people like mom use it to solve problems.

> > Okay, but how can data help us sell more lemonade?

> > > 0

Data can tell us what people want to buy at our lemonade stand. We don't have any data yet. We have to collect data by asking questions.



### What questions?

Clara and Alex ask their friends what they would buy at a lemonade stand.

Clara and Alex talk to their friends Carlos, Isabella, and Ethan. Carlos wants strawberry lemonade. Isabella wants stickers for her notebooks. Ethan wants regular lemonade and some painted rocks for his garden.



### Clara and Alex go back home to collect stickers and paint rocks.



Hoping to get more visitors to their lemonade stand, they reopen with lemonade, stickers, and painted rocks.

They wait and wait but only sell two cups of regular lemonade.



Mr. Williams, their mom's coworker, sees them looking sad at their stand.

Why are you all so sad?

Tell me more. Who did you ask?

Hi, Mr. Williams. Our things are not selling well. We asked our friends what we should sell, but it didn't help.

Did you ask any adults?

We didn't ask the grown-ups. I thought data could come from anyone.

### We asked three of our friends.

I asked Carlos, my friend from school!

It's true that data can come from anyone. But to make good predictions, or guesses about what will happen, we have to talk to a lot of people!

What an adult might want is different from what a kid might want. Everyone will have different ideas depending on what they like.

Not just our friends?

Another piece of advice is that you want to find a place that has more people walking by.

 $\mathbf{O}$ 

•

You're right. We should ask other people, too.

We know what to do now! Thanks, Mr. Williams! Clara and Alex begin asking people they meet in the neighborhood, adults and kids.

What do you want? What would you like? 1 0 0 -U:D PI





3



1.1

The neighbors give Clara and Alex lots of different answers about what they'd like to buy, including cookies, cupcakes, lemonade, stickers, painted rocks, and yo-yos! No other people ask for strawberry lemonade.

Wishlist
Lemonade
Stickers
Lemonade, painted
Lemonade, painted
Lemonade, pointed
Stickies & cupcake
cookies & cupcake
; Yo-YDs, Lemonade · vo-yos Lemonade & cupcakes! cookies! Lemonade & cookies.

Item Lemonade stickers painted rocks cookies cupcakes 40-40s

### What three things do people want most?



Not many people walk past our house. Let's gather our friends and look for places that have more people walking by.

12

Our neighbors want freshly squeezed lemonade, cookies, and yo-yo's the most. Let's sell these things at our lemonade stand!



Clara and Alex meet up with Noah, Lucas, and Mia, who are all excited to join the Lemonade Crew!

Noah lives by the grocery store, Lucas lives across from the park, and Mia lives near the school.

6

6

Let's count how many people go by our houses in one hour. Cases

Together, they count the number of people walking by their houses. The kids know that more people going by means more customers.



Which location had the most people passing by?

### Can you link the correct place to the chart?





Clara and Alex realize that Lucas' house



The Lemonade Crew has so many customers, and everyone is busy!



## Lemonade



Everyone is so proud of the good things they have done. The Lemonade Crew now understands the power of data science. They are on their journey to make the world better using data!

Please use this money to help people get better.

The next day, the kids give the hospital a big envelope with a letter and all the money they raised.

✓ - the Lemonade Crew

## **CAN YOU TELL US MORE ABOUT DATA?**

6 6

Data scientists use a lot of data to solve bigger problems like selfdriving cars and ocean pollution.

Clara, I like the way that you and Alex used a small amount of data to set up the lemonade stand.





Data scientists teach **self-driving cars** to recognize things on the road.

Like people, cars, and traffic lights?

You're right! Lanes, turn signals, everything a human driver needs to know about.

6.



We help train the computer in the car to make decisions. We give it **labels**, so it can put things into different groups, or what data scientists call **classes**.

Great job! A computer has to see a ton of images with labels like these to learn new things.







traffic light



But how do you teach a car?

SELF-PRIVING

I can label these images!









Can computers learn without labeled data?

Yes! We can also teach a computer to learn from patterns. You and I can look at lemons and strawberries and know they're different kinds of fruit without anyone telling us.







OUTPUT



We give computers enough information or **input** to provide the answers, or **output**, we want.

When we use patterns, we won't have the data labeled for the computers, but we have millions of images that can provide features for computers to look at.

If we give the computer a lot of images of lemons and strawberries and ask the computer to separate these images into two categories, the computer could look at the color, shape, or other features that the images provide.









FEATURE







We also learned that we need to ask both kids and adults for information for our lemonade stand.

If you only include one group of people's ideas, the data that you collect could be biased. **Bias** means that the data doesn't give a fair picture of what all the people think. We need to hear more people's voices because the decisions we will make might affect them as well.

We ask people questions and watch what they do. It's just like what you did with the lemonade stand, Clara and Alex. You asked what people wanted to buy and then saw what they actually bought.



How might kids' taste in food be biased?

I want to be a doctor when I grow up. Can data help me with medicine?

Yes. With the data collected from patients, researchers can study and find cures for new diseases. Besides music and medicine, data science helps people make better decisions.

For example, we **predict** the amount of milk people might buy next week at the grocery store, so the store knows how much to stock. We predict the traffic in the next hour and find the fastest way to get you to your school in the morning.



How does this data help us in our daily lives?

There are so many ways that data helps every aspect of our lives. Part of the data scientist's job is to collect the desired data to solve problems.

I love music. Can data help me with music?

Yes. With data collected from your playlist, we can help computers recommend the next song that you might like.





# **REAL-WORLD PROBLEMS**

Data scientists use data and similar methods to solve larger problems that affect more people.

Try the activities below. Use data to answer the questions.

### **Question**: Which type of video game should a company create?



Answer:

**Question:** How does data help my deliveries arrive faster?



**Answer**:

### **Question**: How can I use data to identify a dog's breed?



Hint: What are some of the features of each breed?

### Answer:

**Question**:



Answer:

### Can data science help cure diseases?

### **Question**: How do we decide what food to take to space?

you want to answer with data:

	MASS (how much space does the food take up)	calories (how much energy can it provide)	
FOOD A	100	200	My question is:
<b>FOOD B</b>	50	150	
FOOD C	200	250	
Answer			~~~~~

## Can you think of one question that



## A Letter to the Reader



### Hello Everyone,

In our world, it's very important that we understand where data comes from and how it can be used to make good predictions and decisions. Data can help us find the best solutions for people, our natural resources, and our communities. There are a lot of problems we can solve by analyzing data. The possibilities are limitless – just as they are for you!

Thanks for starting early on your path to appreciating the value of data and developing your data literacy. You're on your way - keep it going!

Scott Aronson

Thank you so much for reading this book! My sons, Flynn and Jedd, and I really enjoyed learning from Clara and Alex and hope you did too.

Proud father of two young boys & Chief Operating Officer at Cloudera

### GLOSSARY

**Bias (Data bias):** The available data does not fairly represent the people or objects in the study.

Classes: The different groups that we sort things into.

**Data:** A collection of facts or things you know, such as numbers, words, images, or just descriptions of things.

**Input:** Information given to the computer. In computer science, inputs are usually data you wish to analyze.

Labeled data: Information that has been tagged with one or more labels, such as a picture of a cat with the label "cat".

**Output:** Data generated by a computer. In machine learning, outputs are predictions generated by the model. Usually, output changes when the input is changed.

Predict: To guess about what you think will happen.

**Self-driving cars (autonomous car):** A vehicle that is capable of sensing its environment and moving by itself.

**Unlabeled data:** Data that comes with no tag. For example, a picture of a cat without the label "cat".

www.freshsqueezekids.com www.readyai.org/a-fresh-squeeze-on-data

Contributors: Annabel Hasty, Dave Tourezky, Ethan Chen, Joel Wilson, Juno Schaser, Kelli Lawless, Madge Miller, Melanie Beck, Roozbeh Aliabadi, Santiago Giraldo, Sushil Thomas Author and Illustrator: Shanshan Jin



## CLOUDERA

.:

· · · ·

In partnership with Reody

f.

i.