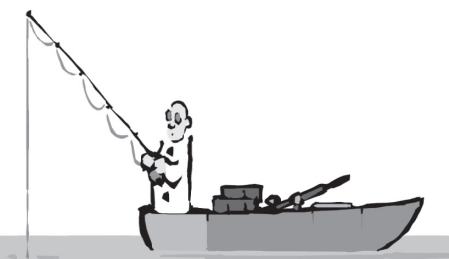


BY **GRADY KLEIN** AND **ALAN DABNEY, Ph.D.**

**THE
CARTOON
INTRODUCTION TO
STATISTICS**

THE CARTOON INTRODUCTION TO STATISTICS

BY **GRADY KLEIN** AND **ALAN DABNEY, Ph.D.**



A NOVEL GRAPHIC FROM HILL AND WANG
A DIVISION OF FARRAR, STRAUS AND GIROUX
NEW YORK

HILL AND WANG

A DIVISION OF FARRAR, STRAUS AND GIROUX
18 WEST 18TH STREET, NEW YORK 10011

TEXT COPYRIGHT © 2013 BY GRADY KLEIN AND ALAN DABNEY
ARTWORK COPYRIGHT © 2013 BY GRADY KLEIN
ALL RIGHTS RESERVED
PRINTED IN THE UNITED STATES OF AMERICA
FIRST EDITION, 2013

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

TK

WWW.TWITTER.COM/FSGBOOKS • WWW.FACEBOOK.COM/FSGBOOKS

1 3 5 7 9 10 8 6 4 2

FOR ANNE AND LIAM AND BENJAMIN
—GK

FOR ELLIOTT AND LOUISE AND NICK
—AD

CONTENTS

INTRODUCTION: THEY'RE EVERYWHERE! ...1

PART ONE: GATHERING STATISTICS ...15

- 1. NUMBERS ...17**
- 2. RANDOM RAW DATA ...25**
- 3. SORTING ...39**
- 4. DETECTIVE WORK ...51**
- 5. MONSTER MISTAKES ...67**
- 6. FROM SAMPLES TO POPULATIONS ...81**



PART TWO: HUNTING PARAMETERS ...89

- 7. THE CENTRAL LIMIT THEOREM! ...91**
- 8. PROBABILITIES ...105**
- 9. INFERENCE ...121**
- 10. CONFIDENCE ...131**
- 11. THEY HATE US ...143**
- 12. HYPOTHESIS TESTING ...161**
- 13. SMACKDOWN ...175**
- 14. FLYING PIGS, DROOLING ALIENS, AND FIRECRACKERS ...191**



CONCLUSION: THINKING LIKE A STATISTICIAN ...205

APPENDIX: THE MATH CAVE! ...213

MOST OF US ENCOUNTER STATISTICS **EVERY DAY...**



AWESOME!

ONE BOWL OF CHOCOLATY FROSTO BOMBS HAS **1200%** OF MY DAILY RECOMMENDED SUGAR.



... EVEN IF WE DON'T **CRUNCH NUMBERS** FOR A LIVING.

STATISTICS RADIATE FROM OUR **TELEVISIONS...**

THIS SHOW HAS AN **ESTIMATED AUDIENCE OF 4.8 MILLION!**

IT MUST BE GOOD.



... EMANATE FROM OUR **PHONES...**

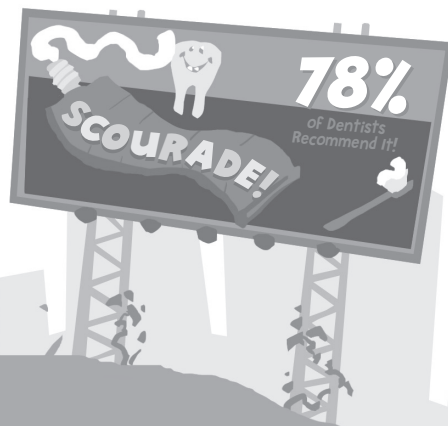
YOU SENT MORE **TEXT MESSAGES** THIS MONTH THAN THE **ENTIRE NATION OF CHAD.**



... AND LITTER OUR **ROADS.**

... SEEP FROM OUR **RADIOS...**

POLLS SHOW SENATOR NEERDORPH WITH A **40 POINT LEAD!**



IT'S IMPOSSIBLE TO ESCAPE THEM.

THEY'RE **EVERYWHERE!**

AT THE MALL

WE'RE PLAYING
THAT BACKGROUND
MUSIC...

... BECAUSE STUDIES
SHOW IT'LL MAKE
YOU BUY 10% MORE
CLOTHES!



AT SCHOOL

YUP, I'M
GRADING ON
A CURVE!



IN THE KITCHEN

WHY DO I HAVE TO DO
THE DISHES, LIKE, 75%
OF THE TIME!

BECAUSE I COOK
DINNER 99% OF
THE TIME,



IN THE BEDROOM

THIS WEBSITE WILL
FIND MY PERFECT
MATCH...

... IF I ENTER MY
VITAL STATISTICS,



STATISTICS ARE WITH US
WHEN WE'RE BORN...

95% OF ALL BABIES ARE
DELIVERED BETWEEN
WEEKS 38 AND 42...

... SO THAT'S WHEN
WE'LL DELIVER
YOURS,



...AND LIKE IT OR NOT, WE WILL **BECOME STATISTICS** WHEN WE DIE.

SO SAD,

AT LEAST HE LIVED
LONGER THAN THE
AVERAGE BEAGLE.



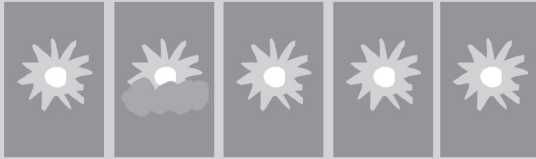
FORTUNATELY, THERE'S
A **GOOD REASON** FOR ALL THIS.

STATISTICS ARE EVERYWHERE
BECAUSE THEY'RE SO USEFUL.

STATISTICS HELP PEOPLE
PREDICT THE **WEATHER...**

THERE'S A **95%**
CHANCE IT'LL BE SUNNY
TOMORROW!

BUT ALSO A **3%**
CHANCE IT'LL
RAIN FROGS!



...AND ORGANIZE THE
INTERNET...

BASED ON YOUR BUYING
HISTORY, I'VE GOT
RECOMMENDATIONS
FOR YOU,

HOW DID IT KNOW
I WANTED A WILLIAM
SHATNER DOLL?



...AND DEVELOP **MEDICINES...**

OUR STUDIES SHOW THAT
THIS PILL IS ONLY **2.5%** MORE
LIKELY THAN A PLACEBO TO
PREVENT CANCER, WITH A
MARGIN OF ERROR OF **12%...**

...BUT IT WORKS
GREAT AS A
LAXATIVE!

GREAT, WHAT
SHOULD WE
CALL IT?



...AND **INFLUENCE FASHION.**

I USED STATISTICS TO
DETERMINE THAT **DENIM**
JACKETS ARE LIKELY TO
COME BACK IN STYLE
THIS YEAR...

LIKE, WOW YOU
ARE SO **1987**,
I LOVE IT!

BUT LOSE THE
BELL BOTTOMS,

AND THAT'S **NOT ALL.**



**STATISTICS ALSO HELP PEOPLE
WIN ELECTIONS...**

ONLY 23% OF MY
CONSTITUENTS THINK
I'M AN OBSEQUIOUS
JERK!



**...AND BUILD POWER
PLANTS...**

WILL OUR NUCLEAR
FACILITY CAUSE NEARBY
RESIDENTS TO MUTATE?



...AND MAKE MONEY...

IF CURRENT MARKET
TRENDS HOLD...

... I'LL BE BETWEEN 12%
AND 15% MORE FILTHY
RICH TOMORROW!



**...AND SHOW THEIR
SUPERIORITY.**

HA, I HIT MORE HOME
RUNS THAN YOU DID.

HA, YOU USED STEROIDS,
AND HERE ARE THE STATISTICS
TO BACK ME UP!



SO **WHAT MAKES STATISTICS**
SO INCREDIBLY USEFUL?

THIS THING IS
AMAZING!

IT'S GOT A **FORK** AND A
KNIFE AND A **SPOON** AND
A **RAKE** AND A **STRAW**...

...AND A **DRILL** AND
FINGERNAIL CLIPPERS
AND A **PENCIL** AND...



THE **SIMPLE ANSWER** IS THAT STATISTICS HELP US COME TO GRIPS WITH
LARGE NUMBERS OF IMPORTANT THINGS...

94% OF ALL HUMANS
WHO EVER LIVED
ARE DEAD...

...AND 200 MILLION
OF THEM DIED IN
THE PLAGUE...

...AND TRAFFIC KILLS
MILLIONS MORE
EVERY YEAR...

...AND YOUR ODDS OF BEING
STRUCK BY LIGHTNING ARE
WAY **HIGHER** IF YOU PLAY GOLF!



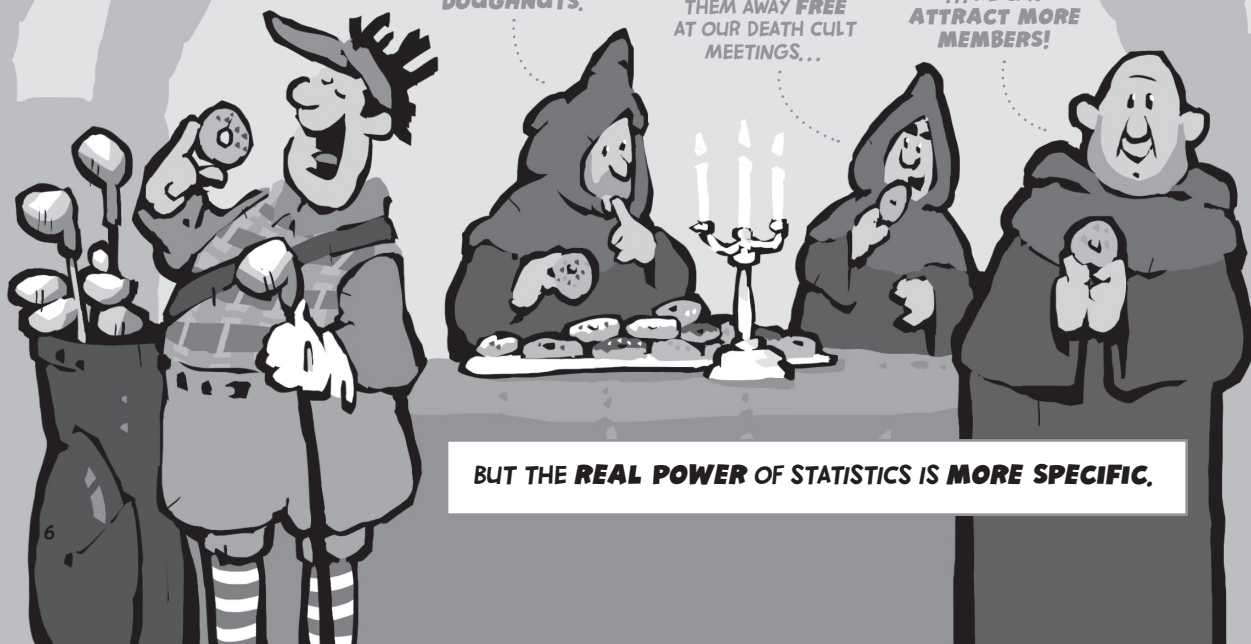
... WHICH IN TURN CAN HELP US **BETTER**
UNDERSTAND OUR COMPLEX WORLD...

... AND **MANIPULATE IT.**

STUDIES SUGGEST THAT
78% OF ALL PEOPLE LIKE
DOUGHNUTS.

SO IF WE GIVE
THEM AWAY **FREE**
AT OUR DEATH CULT
MEETINGS...

... WE CAN
ATTRACT MORE
MEMBERS!



BUT THE **REAL POWER** OF STATISTICS IS **MORE SPECIFIC.**

HERE'S THE **REAL REASON**
EVERYONE USES STATISTICS:

STATISTICS HELP US
MAKE CONFIDENT
DECISIONS...

...WHEN WE
HAVE **LIMITED**
INFORMATION.

BUT LET'S EXPLAIN
WHAT THAT
MEANS...



IMAGINE THAT WE WANT TO KNOW
THE **AVERAGE WEIGHT**...

...OF **ALL THE FISH IN A LAKE**.

HERE FISHY, HERE
FISHY, FISHY...

IF WE FIND OUT HOW
MUCH EACH FISH WEIGHS
ON AVERAGE,...

...WE CAN FIGURE OUT
ABOUT HOW MANY FISH
WE NEED TO CATCH EACH
DAY TO **KEEP OUR CATS**
FROM STARVING!

IF WE **DRAINED THE LAKE** AND
WEIGHED EVERY SINGLE FISH...

...WE'D HAVE **ALL THE INFORMATION** WE NEED TO CALCULATE THE ANSWER.

BUT FOR OBVIOUS REASONS, WE **CAN'T DO THAT**.

MAYBE
THAT WASN'T SUCH
A **GOOD IDEA**.

ON THE OTHER HAND, IF WE CATCH A
SAMPLE OF 100 FISH AND WEIGH THEM...

THESE **100 FISH**
WEIGH **247 POUNDS**.

SO THE **AVERAGE**
FISH IN THIS SAMPLE
WEIGHS **2.47 POUNDS!**

... WE'LL HAVE ONLY **LIMITED INFORMATION** ABOUT ALL THE FISH.

SO NOW WE KNOW THE
AVERAGE WEIGHT IN THIS
SAMPLE PILE...

... BUT WE STILL DON'T
KNOW THE AVERAGE WEIGHT
OF THE **REST OF THE FISH**
IN THE LAKE.

BUT HERE'S THE
COOL THING:

WITH THE TOOLS OF STATISTICS WE CAN
USE THIS LIMITED INFORMATION...

... TO MAKE **CONFIDENT STATEMENTS**
ABOUT **ALL THE FISH** IN THE LAKE.

STATISTICS IS ABOUT
USING THE FISH
WE DID CATCH...

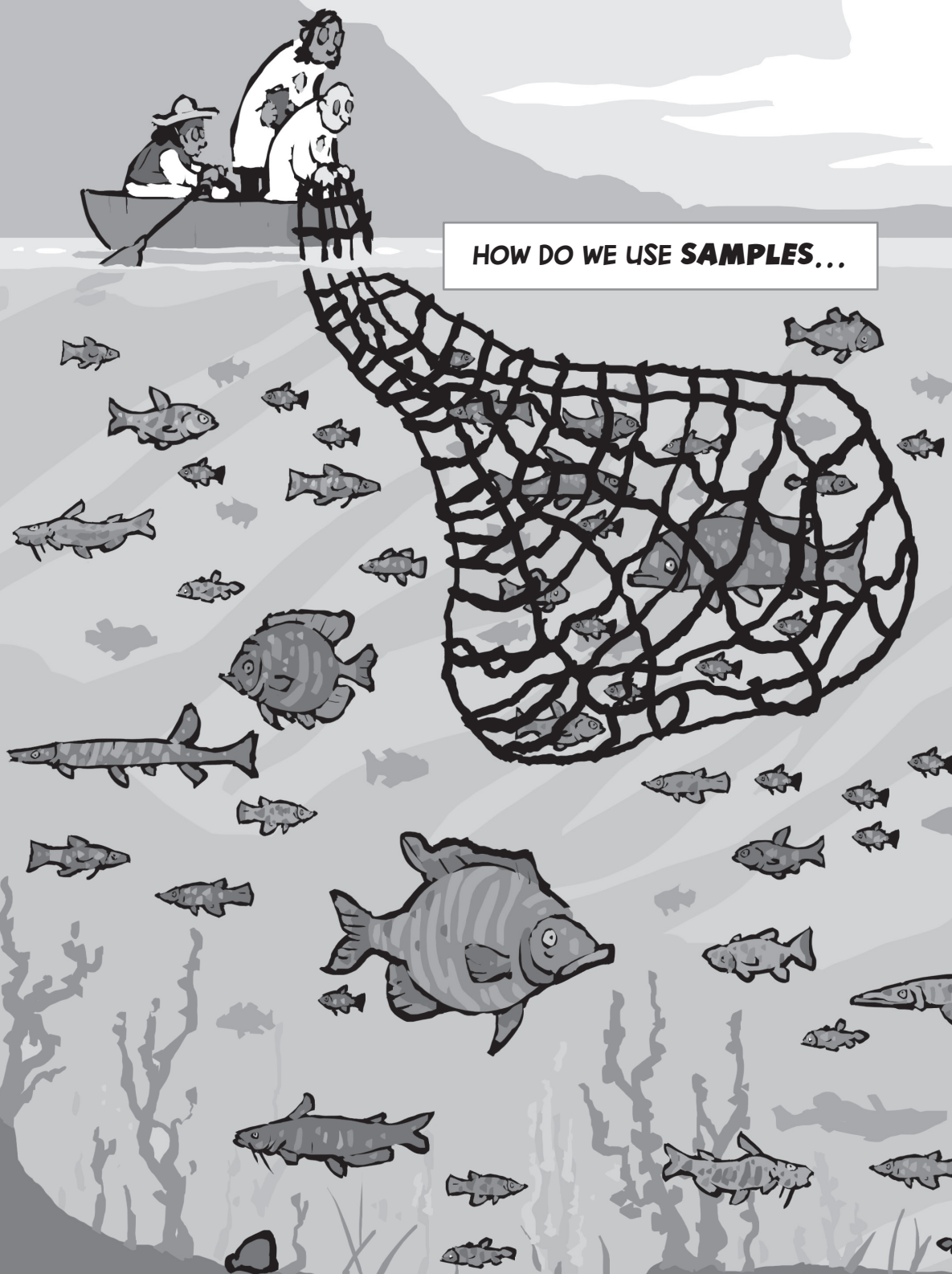
... TO SAY THINGS
ABOUT THE FISH
WE DIDN'T.

REALLY?
HOW DOES
THAT WORK?

THAT'S WHAT
THIS BOOK IS
ABOUT!

THIS BOOK IS ABOUT THE
FUNDAMENTAL QUESTION OF STATISTICS:

HOW DO WE USE **SAMPLES...**



...TO MAKE **CONFIDENT STATEMENTS**
ABOUT **ENTIRE POPULATIONS?**

ALL STATISTICS
PROBLEMS LOOK
LIKE THIS!

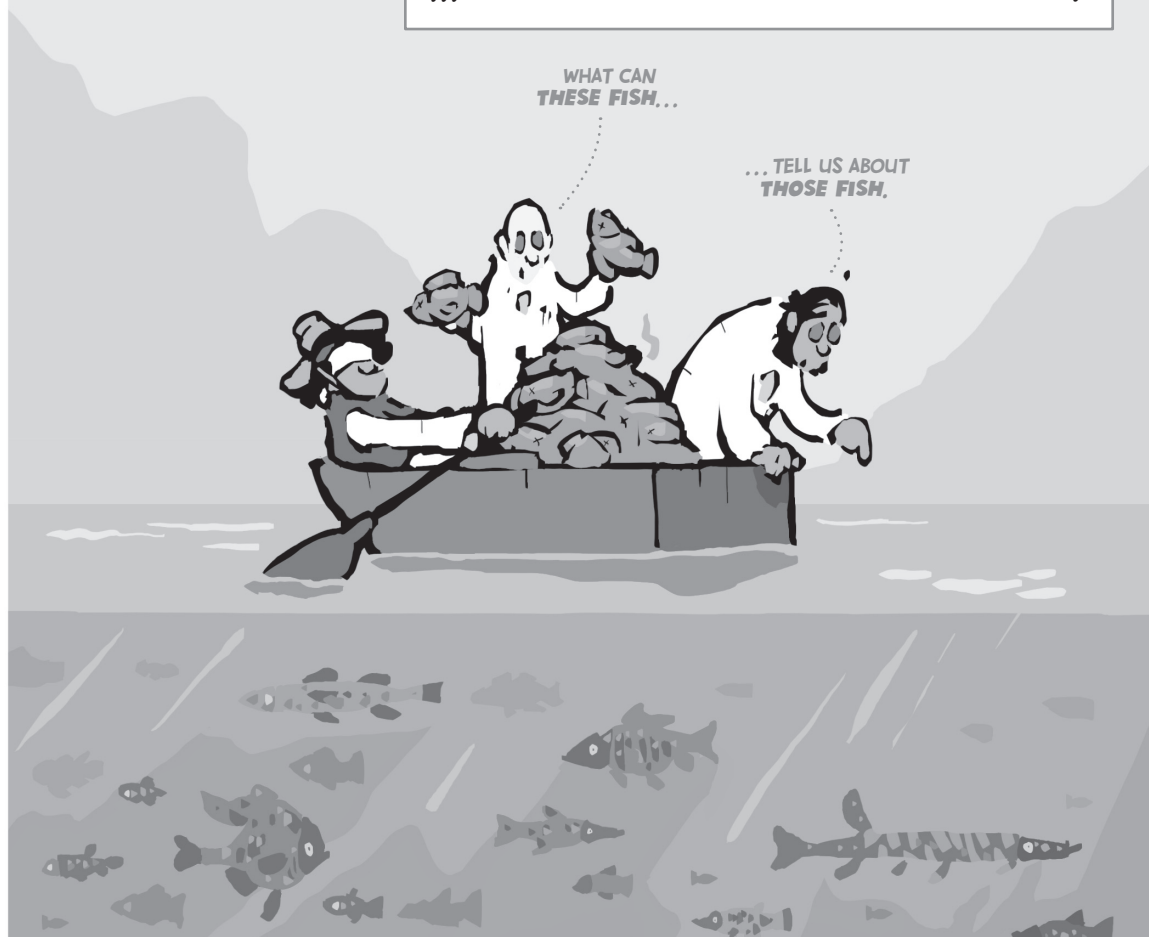
IN **PART ONE** WE'LL LEARN HOW TO
GATHER SAMPLES...

...AND **INVESTIGATE THEM.**



THEN, IN **PART TWO**, WE'LL LEARN HOW TO USE SAMPLES TO
HUNT FOR QUALITIES IN POPULATIONS...

... USING A PROCESS CALLED **STATISTICAL INFERENCE.**



ALONG THE WAY WE'LL
LEARN TO SIFT THROUGH
BIG PILES OF DATA...

ARGH!
WE'RE
SKEWED!

THAT'S NOT
NORMAL!

... CALCULATE **CONFIDENCE
INTERVALS**...

I'M 95% CONFIDENT
THAT WE HATE YOU
ABOUT THIS MUCH,

... AND **TEST
HYPOTHESES**.

I'M 3% CONFIDENT
THAT MY EVIL MACHINE
ISN'T BUSTED!

AND MORE GENERALLY, WE'LL GET A SENSE OF
THE KIND OF THINGS YOU **CAN**...

... AND **CAN'T**...

WE CAN USE
STATISTICS TO
MAKE **CONFIDENT
GUESSES**...

... BUT YOU CAN **NEVER**
USE THEM TO **ACHIEVE
CERTAINTY**.

... DO WITH STATISTICS.

IF WE DON'T CATCH
ALL THE FISH...

... WE'LL NEVER KNOW
FOR CERTAIN WHAT'S
DOWN THERE,

IN THIS BOOK WE'RE GOING TO
FOCUS ON THE BASIC CONCEPTS.

LIKE STANDARD
DEVIATIONS...

...AND SAMPLING
DISTRIBUTIONS...

...AND PROBABILITIES...

...AND **CONFIDENCE!**

BUT IF YOU'RE ALSO CURIOUS ABOUT
THE **TECHNICAL DETAILS...**

LIKE WHAT THE HECK DO
THESE **FORMULAS** AND
SYMBOLS MEAN?

... YOU CAN FIND THOSE IN A SECTION AT
THE END CALLED THE **MATH CAVE**.

PART ONE
GATHERING
STATISTICS

NO PEEKING.

