Technology Companion

Doing Statistics in StatCrunch

A hands-on companion to

You're Smarter Than You Think: Statistics

Safaa Dabagh

Keyed chapter-by-chapter to the main book

"The computer does the arithmetic. Your job is to ask the right question and read the answer. You can absolutely do that."

2026

Before You Click a Single Thing

If the word "software" just made your shoulders climb toward your ears, breathe. StatCrunch is not a wall of code. It is a spreadsheet with two friendly menus on top. You put your numbers in the grid, you pick a menu item that names what you want (a mean, a graph, a test), you click, and it answers. If you pick the wrong menu, nothing breaks. You close the little window and try another menu. That is the entire loop.

This companion follows the main book chapter for chapter. When you finish Chapter 4 in the book, open Chapter 4 here and watch the same ideas happen in StatCrunch. You never have to guess which tool goes with which idea.

One promise: every click-path in this book is short, and every one is followed by what the answer looks like and what it means. You are never left staring at output wondering what you are seeing.

Getting Into StatCrunch

There is nothing to install. StatCrunch is a web app that runs in your browser.

1. **Through your course.** Most students reach StatCrunch through **Pearson MyLab** (also called MyStatLab). Log in to your course and look for the StatCrunch link in the menu or inside an assignment.

2. **Directly.** You can also go to `statcrunch.com` and sign in with a StatCrunch account.

3. Either way, you land on the same tool: an empty **data table** (a spreadsheet grid) with a row of menus above it.

The parts you'll use:

- **The data table** (the grid): each column is a variable, each row is one person or observation. This is where your numbers live.
- **The Data menu:** load files, take a sample, simulate, or compute a new column.
- **The Stat menu:** every number result. Summary Stats, Tables, Z Stats, T Stats, Proportion Stats, Variance Stats, ANOVA, Regression, Goodness-of-fit, and Calculators all live here.
- **The Graph menu:** every picture. Bar plots, histograms, boxplots, scatterplots.

Your very first click

Type three numbers into the first column of the grid, one per row: `2`, `4`, `6`. Then follow this path.

StatCrunch (click this path)

```
Stat > Summary Stats > Columns
```

Pick that first column, click Compute!, and read the mean.

What you'll see

```
Mean
   4
```

You just ran StatCrunch. Every result in this whole book is the same three steps: numbers in the grid, pick a menu, click Compute!. That is the whole scary part, done.

Getting Data Into StatCrunch

Two ways, from simplest to most real. You will use both across the book.

1. Type it into the data table yourself. Click a column header to name it (say, `score`), then type one value per row down the column. For a second variable, use the next column. That is it, you now have data.

2. Open a file (a spreadsheet saved as `.csv`). Use the Data menu.

StatCrunch (click this path)

```
Data > Load > from a computer file
```

Choose the `.csv` file, and its columns drop straight into the grid. You can also load a file that lives online.

StatCrunch (click this path)

```
Data > Load > from a URL
```

Paste the web address of the `.csv` and click Load. Either way, the column names come along, so you can pick them by name later.

That is the whole toolkit. Now let's walk the book.

Chapter 1

What's the Story in the Data?

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Get data into the grid, tell the difference between the *words* column (categories) and the *numbers* column, and look at a frequency table without panic.

In Chapter 1 you learned the difference between a population and a sample, and between categorical and quantitative variables. In StatCrunch, that difference shows up in the tools it offers: a words column gets a frequency table, a numbers column gets an average.

Load `class.csv` (or type in `study_spot` and `quiz_score`). To count the categories in `study_spot`, use a frequency table.

StatCrunch (click this path)

```
Stat > Tables > Frequency
```

What you'll see

study_spot	Frequency
Home	8
Library	5
Cafe	3

▪ READ THE OUTPUT

The frequency table counts each category: Home appears most often (8 of 16), then Library, then Cafe. That is the "words" variable, summarized by counting. The other column, `quiz_score`, is numbers you can average, so it belongs to Summary Stats instead (next chapter). Same table, two kinds of variable, two kinds of tool.

▪ YOUR TURN

Make a frequency table of a categorical column you type yourself: five majors of your choice, one per row. Which major is the most common in your little sample? Could you take a mean of that column? Why not?

Chapter 2

Where Did This Data Come From?

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Take a *simple random sample* in StatCrunch, and see why setting a seed makes randomness repeatable.

Chapter 2 is about how good data is collected. The gold standard is the simple random sample: everyone has an equal chance. StatCrunch can draw one for you.

Load `population40.csv` (40 students). Then sample five of them.

StatCrunch (click this path)

```
Data > Sample
```

What you'll see

```
Sample of 5 rows from student_id:
S9, S37, S1, S23, S6
(sampled without replacement)
```

▪ READ THE OUTPUT

In the Sample window you pick the column to sample, set the sample size to 5, and click Compute!. StatCrunch draws *without replacement* (nobody is picked twice), which is exactly a simple random sample. If you type a number in the seed box first, the "random" draw becomes repeatable: the same seed gives the same five students, so another person can check your work. Change the seed, get a different sample.

▪ YOUR TURN

Draw a random sample of 8 students. Then open Data > Sample again with the same seed. Same eight? Now clear the seed box and sample twice. What changes?

Chapter 3

Show Me a Picture

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Make the three core graphs from the chapter: a bar plot for categories and a histogram for numbers (plus a boxplot).

StatCrunch keeps every picture in one place: the Graph menu. Bar plots for categories, histograms for numbers.

StatCrunch (click this path)

```
Graph > Bar Plot > With Data
Graph > Histogram
Graph > Boxplot
```

What you'll see

```
Bar Plot: Home tallest, then Library, then Cafe
Histogram of quiz_score: tallest bars over 14 to 15
Boxplot of quiz_score: box centered near 14
```

▪ READ THE OUTPUT

For the bar plot, pick `study_spot`: Home gives the tallest bar, just like the frequency table in Chapter 1. For the histogram, pick `quiz_score`: the scores pile up around 14 to 15, exactly where the book said they would. The boxplot shows the middle, the spread, and any outliers as one tidy summary. Each graph opens in its own window; click Options > Save to keep it.

▪ YOUR TURN

In the histogram window, change the bin width (Options) so there are more bars, then fewer. Does the "pile" still sit near 14 to 15 no matter how you slice it?

Chapter 4

The Typical and the Spread

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Compute the center (mean, median) and the spread (standard deviation, IQR) in one window.

Chapter 4's formulas are a single window in StatCrunch. You do not compute them by hand here; you check your by-hand work against StatCrunch.

StatCrunch (click this path)

```
Stat > Summary Stats > Columns
```

What you'll see

```
Column       Mean   Median   Std.Dev.   IQR
quiz_score    14      14       1.15     1.25
```

▪ READ THE OUTPUT

Pick `quiz_score`, and in Statistics choose the ones you want (Mean, Median, Std. Dev., IQR, and the quartiles). Here the mean and median are both 14, which says the scores are fairly symmetric. If the mean were pulled well above the median, that would warn you of a right skew or a high outlier, exactly the reasoning from the chapter. The standard deviation, about 1.15, is the typical distance from the mean.

▪ YOUR TURN

Add one unusually high score, say 40, in the next empty row, and run Summary Stats again. Which moved more, the mean or the median? That is why we report the median for skewed data.

Chapter 5

What Are the Chances?

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Build a two-way table, add its totals, and read probabilities and conditional probabilities straight off it.

Chapter 5's two-way tables are where probability gets concrete. StatCrunch turns counts into proportions for you.

Load `caffeine_sleep.csv` (100 people). Then cross the two columns.

StatCrunch (click this path)

```
Stat > Tables > Contingency > With Data
```

What you'll see

	Slept:Yes	Slept:No	Total
Caffeine:Yes	30	20	50
Caffeine:No	10	40	50
Total	40	60	100

▪ READ THE OUTPUT

Pick `caffeine` for rows and `slept_well` for columns. StatCrunch shows the counts plus the row, column, and grand totals. In the same window, check Row percent to get the *conditional* view: *given* caffeine (the top row), 30 of 50, or 60%, slept well. That row-percent option is conditional probability, the trickiest idea in the chapter, made into one checkbox.

▪ YOUR TURN

Turn on Column percent instead. Now you are conditioning on whether they slept well. In plain words, what question does that version of the table answer?

Chapter 6

Predictable Randomness

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Get exact binomial and normal probabilities without a table in the back of the book.

Chapter 6's distribution tables are replaced by two calculators. No data file needed; you just type the numbers into the calculator window.

StatCrunch (click this path)

```
Stat > Calculators > Binomial
Stat > Calculators > Normal
```

What you'll see

```
Binomial(n=10, p=0.7):  P(X = 8)   = 0.2335
Binomial(n=10, p=0.7):  P(X <= 8)  = 0.8507
Normal(mean=70, sd=8):  P(X < 80)  = 0.8944
Normal(mean=70, sd=8):  90th pctile = 80.25
```

▪ READ THE OUTPUT

In the Binomial calculator, set `n = 10` and `p = 0.7`, then read `P(X = 8) = 0.233` and `P(X <= 8) = 0.851`. In the Normal calculator, set mean 70 and standard deviation 8: asking `P(X < 80)` gives 0.894 (about 89% of scores fall below 80). To run it backward, type the probability 0.90 and let it solve for the score: the 90th percentile is 80.25. One calculator takes a value and gives a probability; the same calculator can take a probability and give the value.

▪ YOUR TURN

In the Normal calculator, find the probability a score is *above* 85. (Hint: switch the inequality to "greater than," or read "below" and subtract from 1.)

Chapter 7

From Sample to Truth

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Watch the Central Limit Theorem happen by simulating many samples and graphing their means.

Chapter 7 claims something wild: even from a lopsided population, the averages of samples pile up into a bell. In StatCrunch you can see it, not just trust it.

Use the built-in applet, which draws thousands of samples and graphs their means for you.

StatCrunch (click this path)

```
Applets > Sampling distributions
```

What you'll see

```
Population: skewed
Sample means (n = 30): bell-shaped pile centered near the population mean
Standard error of the mean = s / sqrt(n)
```

▪ READ THE OUTPUT

Pick a skewed population, set the sample size to 30, and click to draw 1000 samples. The applet stacks each sample's *mean* into a histogram, and that histogram is bell-shaped even though the population was lopsided. That is the Central Limit Theorem, live. The spread of those means is the *standard error*, and it follows the chapter's formula: the sample standard deviation divided by the square root of the sample size. (If the applet is not in your version, Data > Simulate can generate a column and you can take repeated means the slower way.)

▪ YOUR TURN

Change the sample size from 30 to 5 and draw again. Does the bell get wider or narrower? Bigger samples give smaller standard error, exactly what the chapter promised.

Chapter 8

How Sure Are We?

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Build a confidence interval for a mean and for a proportion, and read it the way the chapter taught.

A confidence interval is a range of believable values for the truth. StatCrunch hands you the whole interval.

For the mean, load `class.csv` and use `quiz_score`. For the proportion, 18 of 40 said "yes."

StatCrunch (click this path)

```
Stat > T Stats > One Sample > With Data
Stat > Proportion Stats > One Sample > With Summary
```

What you'll see

```
Mean quiz_score: 95% CI = (13.39, 14.61)
Proportion 18/40: 95% CI = (0.296, 0.610)
```

▪ READ THE OUTPUT

For the mean, pick `quiz_score`, choose Confidence interval, and read about `13.4 to 14.6`: we are 95% confident the true average score sits in there. For the proportion, type 18 successes out of 40 and read about `0.30 to 0.61` for the true "yes" rate. Want 90% instead of 95%? Change the level in the same window and watch the interval shrink.

▪ YOUR TURN

Re-run the mean interval at 99% confidence. Wider or narrower than 95%? Why does being *more* confident cost you a *wider* net?

Chapter 9

Testing What We Believe

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Run a one-sample hypothesis test for a mean, a proportion, and a variance, and find the p-value.

Chapter 9 is the logic of the p-value. StatCrunch runs the test and prints the p-value plainly; your job is to decide what it means.

Test whether the true mean score is 13, using `quiz_score`.

StatCrunch (click this path)

```
Stat > T Stats > One Sample > With Data
Stat > Proportion Stats > One Sample > With Summary
Stat > Variance Stats > One Sample > With Data
```

What you'll see

```
One-sample t-test, H0: mean = 13
t = 3.46,  df = 15,  P-value = 0.003
```

▪ READ THE OUTPUT

For the mean, choose Hypothesis test, set the null value to 13, and read the output: `t = 3.46` and `P-value = 0.003`. That p-value is smaller than 0.05, so we reject the claim that the mean is 13; the evidence says the true mean is higher. Same move for the proportion test (18 of 40 versus 0.5) and the variance test. Read the p-value, compare it to your significance level, decide. StatCrunch never decides for you, and that is the point of the chapter.

▪ CHECK YOUR VOICE

When the results table appears, the anxious voice says "I don't know what any of this means." You do. You are looking for one number, the P-value, and asking one question: smaller than 0.05 or not? Everything else in the table is supporting detail.

Chapter 10

Comparing Two Groups

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Compare two group means (independent and paired) and two proportions.

Chapter 10 asks: is the difference between two groups real, or just noise? One menu path per situation.

Load `two_groups.csv` for the independent test, and `before_after.csv` for the paired test.

StatCrunch (click this path)

```
Stat > T Stats > Two Sample > With Data
Stat > T Stats > Paired
Stat > Proportion Stats > Two Sample > With Summary
```

What you'll see

```
Two-sample t-test (Group A vs B):
mean A = 78.2,  mean B = 66.5,  P-value = 0.036
```

▪ READ THE OUTPUT

For the independent test, pick `score` split by `group`: the two means are about 78 and 67, and the p-value `0.036` is below 0.05, so the gap is bigger than noise would easily explain. T Stats > Paired is for the *same* people measured twice (`before` and `after`); it looks at each person's change. Choosing independent versus paired is the chapter's key decision; in StatCrunch it is just which menu item you click.

▪ YOUR TURN

Which design, independent or paired, matches a weight-loss study that weighs the *same* people in January and June? Which menu item would you click for it?

Chapter 11

When Three or More Groups Disagree

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Run a one-way ANOVA and, if it is significant, find *which* groups differ.

When there are three or more groups, you do not run many t-tests. You run one ANOVA. Load `three_groups.csv`.

StatCrunch (click this path)

```
Stat > ANOVA > One Way
```

What you'll see

```
One-way ANOVA:
F = 15.44,  P-value = 0.004
(Tukey pairwise comparisons available in the same output)
```

▪ READ THE OUTPUT

Pick `score` as the response and `group` as the factor. Read the `P-value`, `0.004`, below 0.05, so at least one group's mean really differs. But ANOVA does not say which one. Turn on the Tukey comparisons (a checkbox in the same window) and StatCrunch compares each pair (A vs B, A vs C, B vs C) and flags the real gaps. That two-step logic, "is there a difference, then where," is the heart of the chapter.

▪ YOUR TURN

Look at the Tukey output. Which pair of groups has the largest difference? Does its confidence interval include 0? (An interval that misses 0 means a real difference.)

Chapter 12

Counts, Categories, and Surprises

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Run both chi-square tests: goodness-of-fit (one variable) and independence (two variables).

Chapter 12 is about counts that surprise us. Are the categories as evenly split as expected? Are two categorical variables related? Two flavors of one test.

For goodness-of-fit, load `dice60.csv` (a die rolled 60 times). For independence, reuse the caffeine table.

StatCrunch (click this path)

```
Stat > Goodness-of-fit > Chi-Square Test
Stat > Tables > Contingency > With Data
```

What you'll see

```
Goodness-of-fit (fair die):
X-squared = 2.8,  df = 5,  P-value = 0.73
```

▪ READ THE OUTPUT

For the die, feed the six observed counts and let StatCrunch compare them to "all faces equally likely." The p-value `0.73` is large (well above 0.05), so we do *not* have evidence the die is unfair; the ups and downs are ordinary luck. For the caffeine-and-sleep table, run Contingency again and check the Chi-Square option: a small p-value would say the two variables are related, not independent. Two questions, the same chi-square idea, exactly as the chapter frames them.

▪ YOUR TURN

Change the die counts to something lopsided, like 2, 3, 4, 20, 6, 25, and re-run the goodness-of-fit test. Does the p-value drop below 0.05 now? What would you conclude about that die?

Chapter 13

Drawing the Line

▪ WHAT THIS CHAPTER DOES IN STATCRUNCH

Measure correlation, fit a regression line, read its slope and R^2, and predict.

The last chapter draws the line through a scatterplot and asks what it means. Load `study_hours.csv`. StatCrunch gives you the correlation, the line, and a prediction.

StatCrunch (click this path)

```
Stat > Summary Stats > Correlation
Stat > Regression > Simple Linear
Graph > Scatter Plot
```

What you'll see

```
Correlation r = 0.9985
Simple linear regression:
  intercept = 46.47,  slope = 5.69,  R-sq = 0.997
  predicted grade at hours = 3.5  ->  66.4
```

▪ READ THE OUTPUT

`r = 0.9985` is a very strong positive relationship. In Simple Linear, pick `hours` as X and `grade` as Y: the slope `5.69` means each extra study hour adds about 5.7 points, and `R-sq = 0.997` says hours explain about 99.7% of the variation in grade (unusually clean, because this is teaching data). Type 3.5 in the Predict box to read the line at 3.5 hours. Slope, strength, prediction: the whole chapter, on one screen. Check fitted line in the scatterplot to see it drawn.

▪ CHECK YOUR VOICE

You just fit a regression model, the thing that sounded impossible on day one. Notice that. The tools got simpler as the ideas got bigger, because the computer carries the arithmetic. You carried the understanding the whole way.

The Datasets, Chapter by Chapter

Every example in this book comes with a ready-made data file, so you can load the real numbers instead of typing them. All nine files live in a folder called `data`, sitting right next to this PDF. They are plain `.csv` files, which means they open in **StatCrunch**, **R**, **jamovi**, Excel, Google Sheets, and (typed into lists) the **TI-84**. One set of data, every tool.

How to load a dataset (do this once per file)

In StatCrunch, the friendliest way is the menu: Data > Load > from a computer file, pick the file, and its columns drop into the grid. Or point to a file that lives online:

StatCrunch (click this path)

```
Data > Load > from a computer file
Data > Load > from a URL
```

Once it is loaded, each column is named at the top of the grid, so you can pick it by name in any Stat or Graph window. To peek at a single column, just scroll down it in the table. If StatCrunch cannot find your file, make sure you chose the `.csv` inside the `data` folder and try the load again.

The nine datasets

File	Chapters	What's inside (columns)
`class.csv`	1, 3, 4, 8, 9	16 students: `student`, `study_spot` (Home/Library/Cafe), `quiz_score`
`population40.csv`	2	40 students to sample from: `student_id`, `gpa`
`caffeine_sleep.csv`	5, 12	100 people: `caffeine` (Yes/No), `slept_well` (Yes/No)
`survey40.csv`	8, 9	40 people: `person`, `response` (Yes/No; 18 said Yes)
`two_groups.csv`	10	Two independent groups: `id`, `group` (A/B), `score`
`before_after.csv`	10	Same people twice: `student`, `before`, `after`
`three_groups.csv`	11	Three groups for ANOVA: `id`, `group` (A/B/C), `score`
`dice60.csv`	12	A die rolled 60 times: `face` (1 to 6), `count`
`study_hours.csv`	13	`hours` studied and the `grade` earned (6 students)

Two chapters need no file. Chapter 6 (Predictable Randomness) uses StatCrunch's built-in calculators (Stat > Calculators > Binomial and Normal), and Chapter 7 (From Sample to Truth) *generates* its own data with the sampling-distributions applet. Nothing to download for those two; the tool makes the numbers.

Loading each one is the same move. Point Data > Load at the file, then open the Stat or Graph window you need and pick the columns by name. For example, load `caffeine_sleep.csv` and build the Chapter 5 table with Stat > Tables > Contingency > With Data, or load `two_groups.csv` and run the Chapter 10 test with Stat > T Stats > Two Sample > With Data. Same file, same grid, every chapter.

Quick Reference: The Whole Book in StatCrunch

One path per idea

Idea (chapter)	StatCrunch
Categories, frequency (1)	Stat > Tables > Frequency
Random sample (2)	Data > Sample (set a seed)
Bar plot / histogram (3)	Graph > Bar Plot / Graph > Histogram
Center & spread (4)	Stat > Summary Stats > Columns
Two-way table (5)	Stat > Tables > Contingency (row %)
Normal / binomial (6)	Stat > Calculators > Normal / Binomial
Sampling distribution (7)	Applets > Sampling distributions
Confidence interval (8)	Stat > T Stats > One Sample (CI)
One-sample test (9)	Stat > T Stats / Proportion Stats (test)
Two groups (10)	Stat > T Stats > Two Sample / Paired
ANOVA (11)	Stat > ANOVA > One Way (Tukey)
Chi-square (12)	Stat > Goodness-of-fit / Contingency
Regression (13)	Stat > Regression > Simple Linear

You did statistics, and you did it by pointing and clicking with intent. That is not a small thing. Keep this page next to you; it is the whole course in thirteen menu paths.