How to Check That Pay Is Fair: A Simple Guide to Pay Equity Audits

A plain-language walk through every step of a pay equity audit — what happens, why it happens, and what the math really tells us — told through the story of one small ice cream company.

What's Inside

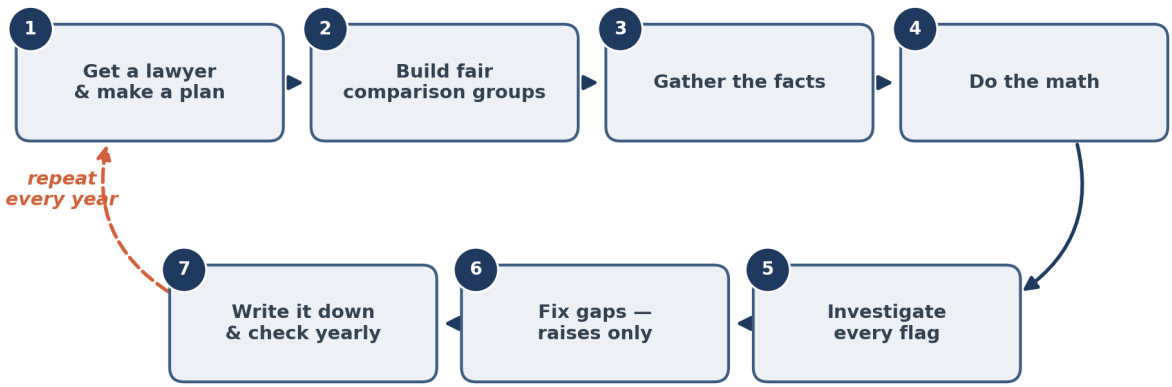
What Is a Pay Equity Audit?

Imagine a company is a big team. Everyone on the team works hard, and everyone gets paid. A pay equity audit is a careful check-up to make sure teammates who do the same kind of work get paid fairly. It asks one big question: if two people do the same job just as well, are they paid about the same? If not, is there a fair reason, like one person has worked there longer? Or is the difference tied to something unfair, like whether someone is a man or a woman, or the color of their skin?

To see how it works, let's follow a made-up company called Sunny Scoops, an ice cream company with scoopers, shift leaders, and store managers.

The Seven Steps at a Glance

Here is the whole journey on one map. An audit is not a mystery; it is seven steps in a row, and the rest of this guide walks through each stop one at a time.



The seven steps of a pay equity audit. The dashed arrow is the most important one: the journey repeats every year.

Step 1: Get a Lawyer's Help and Make a Plan

Before looking at any numbers, Sunny Scoops asks a lawyer to lead the check-up. This may sound odd, but there is a smart reason. When a lawyer leads the work, the early notes and rough findings can stay private while the company fixes any problems it finds. Lawyers call this "privilege." Think of it like a practice test. You get to find your mistakes and fix them before the test that really counts. Companies that skip this step can get in trouble later, because their rough notes could be used against them.

With the lawyer's help, the company also makes a plan. It decides which kinds of pay to check (like regular pay and bonuses), which workers to include, and what date to use for the numbers, like taking one clear photo of everyone's pay on the same day.

Step 2: Compare Apples to Apples

Next, the company puts workers into fair groups. You can only compare people who do the same kind of work. It would not be fair to compare a scooper's pay to a store manager's pay, because those are very different jobs. So Sunny Scoops makes groups: all the scoopers in one group, all the shift leaders in another, and all the managers in a third.

Making good groups takes care. If a group is too big and mixed up, real problems can hide inside it, like trying to find one wrong answer in a giant messy pile of papers. If a group is too small, just two or three people, the math in the next steps won't work well. Very small groups get checked one person at a time instead.

Step 3: Gather the Facts

Now the company collects facts about every worker: how much they are paid, how many years they have worked there, how much experience they had before, their schooling or training, and which city they work in. These are called fair reasons, because they are reasons that can fairly explain why one person earns more than another.

The facts must be correct, so the team double-checks everything, just like checking your homework before turning it in. One warning here: some facts can secretly carry old unfairness inside them. For example, if the company sets pay based on what a worker earned at their last job, and that old pay was unfair, the unfairness sneaks in the door. The team must watch out for tricky facts like that.

Step 4: Do the Math

Here comes the detective work. The team uses a math tool (grown-ups call it "regression analysis") that acts like a super-smart sorting machine. It looks at all the fair reasons at once and asks: after we account for years worked, experience, schooling, and location, is there still a pay gap left over that lines up with gender or race?

Here is a simple example. Maya and Jake are both scoopers at Sunny Scoops. Both have worked there three years and do a great job. Maya earns 15 dollars an hour. Jake earns 17 dollars an hour. The fair reasons are the same for both of them, so nothing explains the two-dollar gap. The math tool raises a flag on that group, like a referee blowing a whistle. The whistle doesn't mean someone cheated for sure. It means "stop and take a closer look."

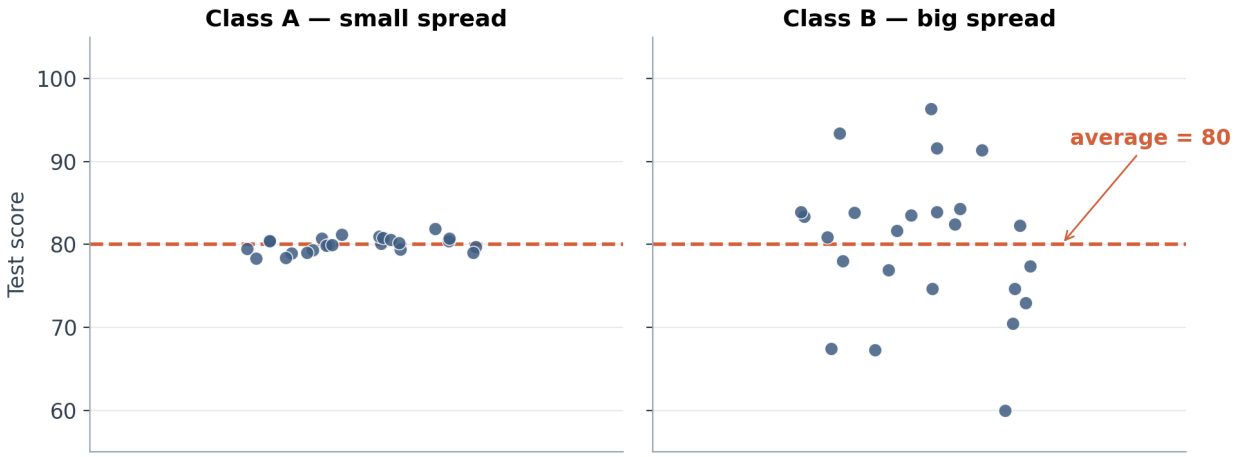
The Math Toolbox: A Closer Look

Step 4 said the team "does the math," but what math, exactly? This section opens the toolbox. Inside are five tools: averages, spread, the t-test, the p-value, and regression. Each one answers a different question, and together they tell the team which pay gaps are probably real and which

are probably just chance. Each tool comes with a picture, because these ideas are easier to see than to read.

Averages and Spread

An average is what you get when you add up everyone's pay and divide by the number of people. Spread tells you how far people's pay usually sits from that average. Two groups can have the same average but very different spreads. Picture two classrooms that both average 80 on a test. In the first class, everyone scored between 78 and 82, so the spread is tiny. In the second class, scores ran all the way from 60 to 100, so the spread is huge. Spread matters for pay because a one-dollar gap between two groups is a big deal when everyone's pay is bunched tightly together, but it might mean nothing in a group where pay is scattered all over the place. Grown-ups measure spread with a number called the "standard deviation."



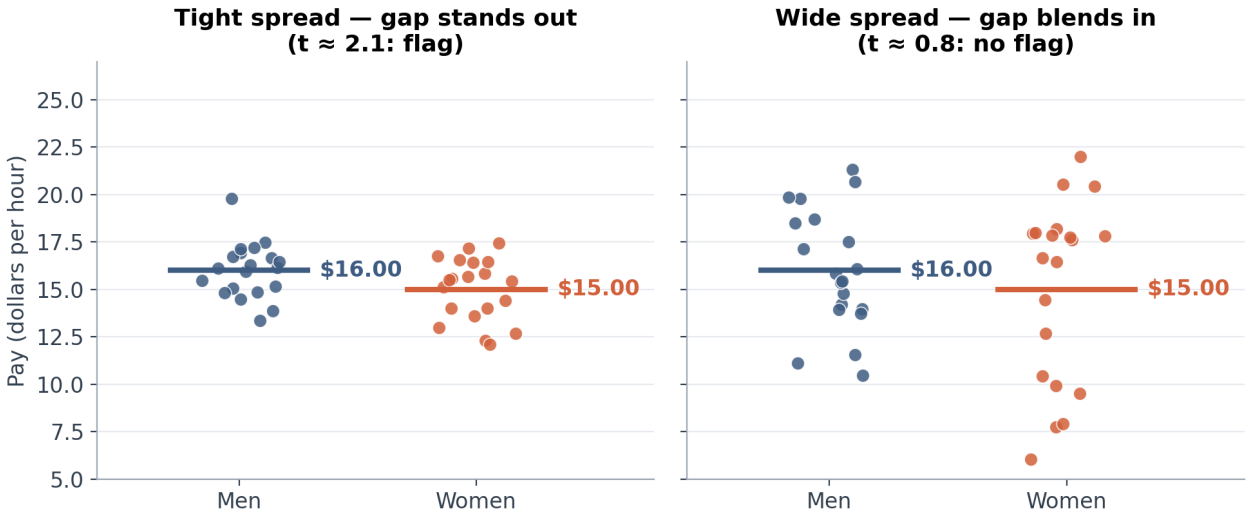
Two classes with the same average score of 80. Class B's scores are far more spread out, so a gap of a few points means much less there.

The t-Test: Is the Gap Between Two Groups Real?

The t-test is a tool for comparing the averages of two groups. At Sunny Scoops, the team can use it to compare the average pay of the 20 men who scoop ice cream with the average pay of the 20 women who scoop ice cream. The t-test asks: is the gap between these two averages too big to shrug off, once we consider the spread and the size of the groups? Its answer comes as a single number called a t-score. The bigger the t-score, the more surprising the gap. As a rough rule, a t-score of about 2 or more means "this gap is surprising, pay attention."

Here is how it plays out. The men average 16 dollars an hour and the women average 15 dollars an hour, and pay in both groups has a spread of about a dollar fifty. The t-test turns that into a t-score of about 2.1, which is over the line, so the group gets flagged. Now imagine the same one-dollar gap, but pay is scattered widely, with a spread of about four dollars. The t-score drops to about 0.8, well under the line, because in such a jumbled group, a one-dollar difference

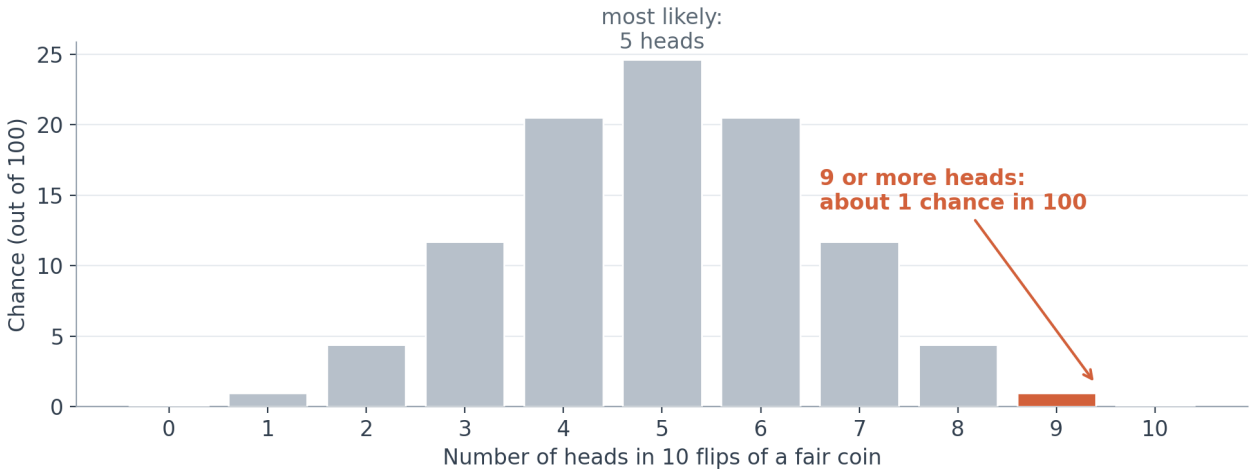
between averages could easily happen by ordinary shuffle. Same gap, different meaning, and that is exactly why the t-test looks at spread and group size, not just the gap itself.



The same one-dollar gap between men's and women's average pay. With a tight spread (left) the gap stands out clearly; with a wide spread (right) the same gap disappears into the noise.

The p-Value: Could It Be Just Luck?

Every t-score comes with a partner number called the p-value, and the p-value answers the question a smart skeptic would ask: could this just be luck? Think about flipping a coin. If you flip a fair coin 10 times, you expect around 5 heads. If you get 9 heads, you get suspicious, because a fair coin does that only about 1 time in 100. That "1 time in 100" is a p-value. It is the chance of seeing a result at least that extreme if nothing unfair is going on at all.

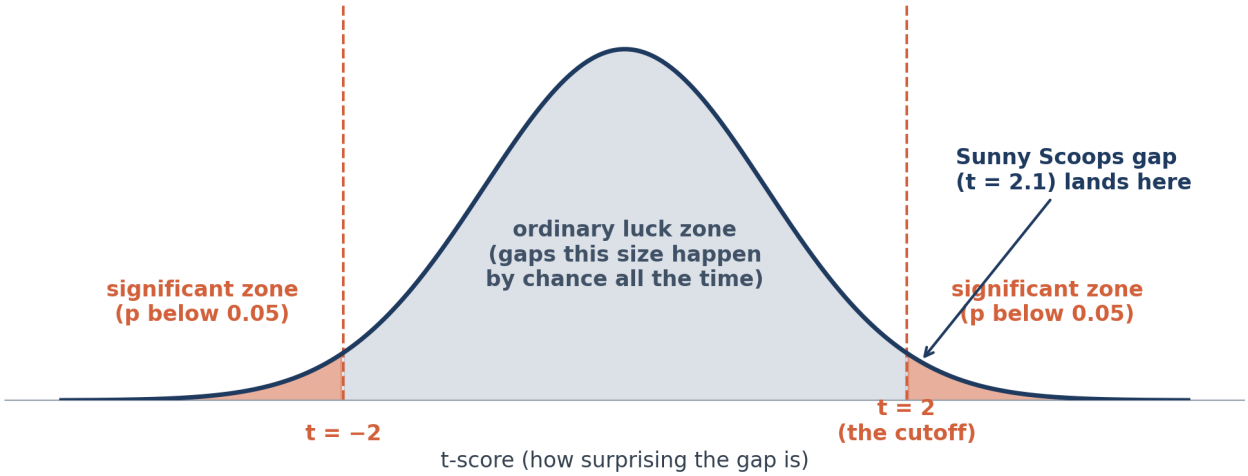


The chance of each result when flipping a fair coin 10 times. Nine or more heads happens only about 1 time in 100. That small chance is a p-value, and it is why 9 heads makes you suspicious of the coin.

Pay works the same way. The p-value tells us: if men and women in this group were truly paid evenly, what is the chance we would see a gap this big anyway, just by luck? A small p-value means luck is a poor explanation. In the Sunny Scoops example, the t-score of 2.1 comes with a p-value of about 0.04, which means that if pay were truly even, a gap this large would show up only about 4 times out of 100. Luck is possible, but it is not likely, and that is what earns the flag.

The 5 Percent Rule: What Counts as "Significant"

So how small does a p-value need to be? The common cutoff is 0.05, or 5 percent, which is the same thing as a t-score of about 2, or what statisticians call "two standard deviations." If the p-value falls below that line, the result is called "statistically significant." This is not just a habit among math fans; it is the same line that courts and government agencies in the United States have long used when they weigh pay cases. Two things are worth remembering about this rule. First, significant does not mean discrimination is proven; it only means the gap is probably not luck, so someone must investigate. Second, the cutoff is a guide, not a magic wall. A large gap in a small group can just miss the cutoff and still deserve a careful look.

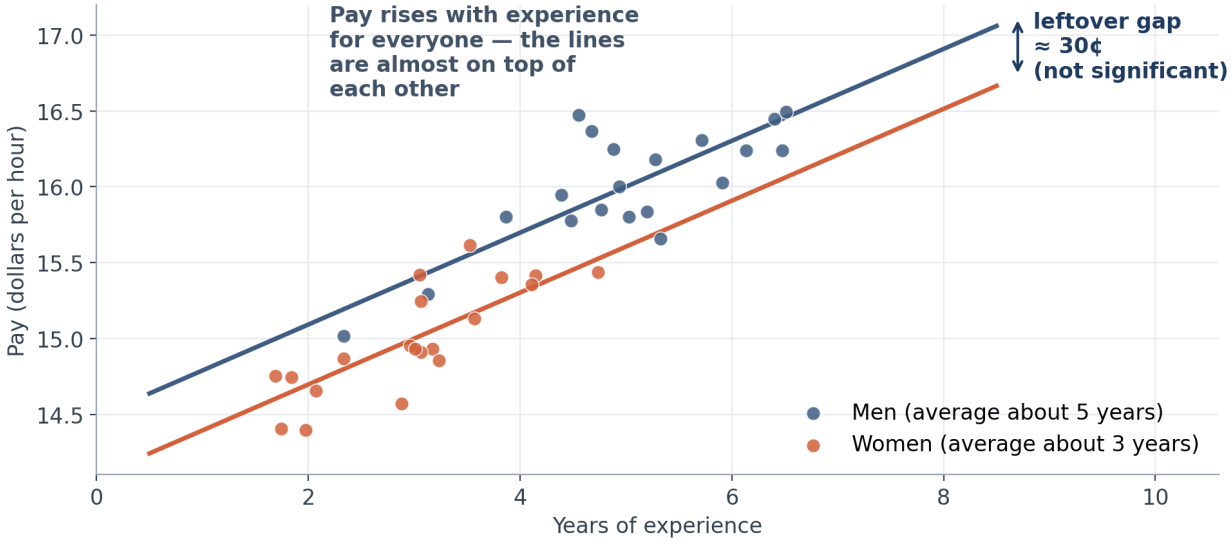


The bell curve of gaps you would expect from luck alone. Most gaps land in the middle. Gaps past a t-score of 2 fall in the significant zone, and the Sunny Scoops gap (t = 2.1) lands just inside it.

Regression: The Fair-Comparison Machine

The t-test has a weakness: it compares only two averages and ignores all the fair reasons. Maybe the men at Sunny Scoops earn more simply because, by chance, they have more years on the job. The t-test cannot see that, but regression can. Regression is the big machine of the toolbox. It takes every fair reason at once, like years worked, experience, schooling, and location, and figures out what each person's pay would be expected to look like based on those reasons alone. Then it asks the key question: after all of that is accounted for, does being a man or a woman, or being one race or another, still help predict who earns more? If pay is fair, the answer should be no.

Regression's answer also comes with its own t-score and p-value. Suppose the raw gap between men and women scoopers is one dollar, but regression discovers the men really do average two more years of experience, which the company fairly pays for. Once experience is accounted for, the leftover unexplained gap shrinks to 30 cents with a p-value of 0.35, meaning luck explains it easily, so there is no flag. But if the leftover gap had stayed at 80 cents with a p-value of 0.02, the whistle would blow, because now the gap cannot be blamed on luck or on any fair reason the company measured.



Each dot is one scooper. Pay rises with years of experience for everyone, and the men simply have more years on average. Once the trend lines account for experience, the leftover gap between the lines is only about 30 cents.

What the Math Can and Can't Tell Us

The math is powerful, but it has honest limits. It finds patterns; it does not explain them. A significant result says "something here needs a closer look," and that is exactly why Step 5 exists. The math also cannot see fair reasons nobody put in the data, and it can be fooled by tricky facts that carry old unfairness inside them, like pay from a past job. Finally, wise teams look at both size and significance together: a tiny gap can be significant in a very large group, and a large gap can miss the cutoff in a very small one, and both deserve human judgment before anyone celebrates or panics.

Step 5: Look Closer at Every Flag

A flag is a question, not an answer. For each flagged group, the team looks at each person's story. Maybe Jake earns more because he has a special ice cream safety certificate that the company pays extra for, and it was written down when he was hired. That would be a fair reason, and the flag goes away. But if the team looks and looks and finds no fair, written-down reason for

Maya's lower pay, then the gap is a real problem that must be fixed. Made-up excuses invented after the fact don't count.

Step 6: Fix It the Right Way

When the company finds unfair gaps, there is only one right fix: raise the pay of the people who were paid too little. It is against the law to cut anyone's pay to make things even. So Sunny Scoops gives Maya a raise to close her gap.

Good companies also fix the habits that caused the problem in the first place. That might mean making clear rules for starting pay, so managers don't just guess, and checking every raise and promotion for fairness before it happens. Fixing the leak matters as much as mopping the floor.

Step 7: Write It Down and Check Again Every Year

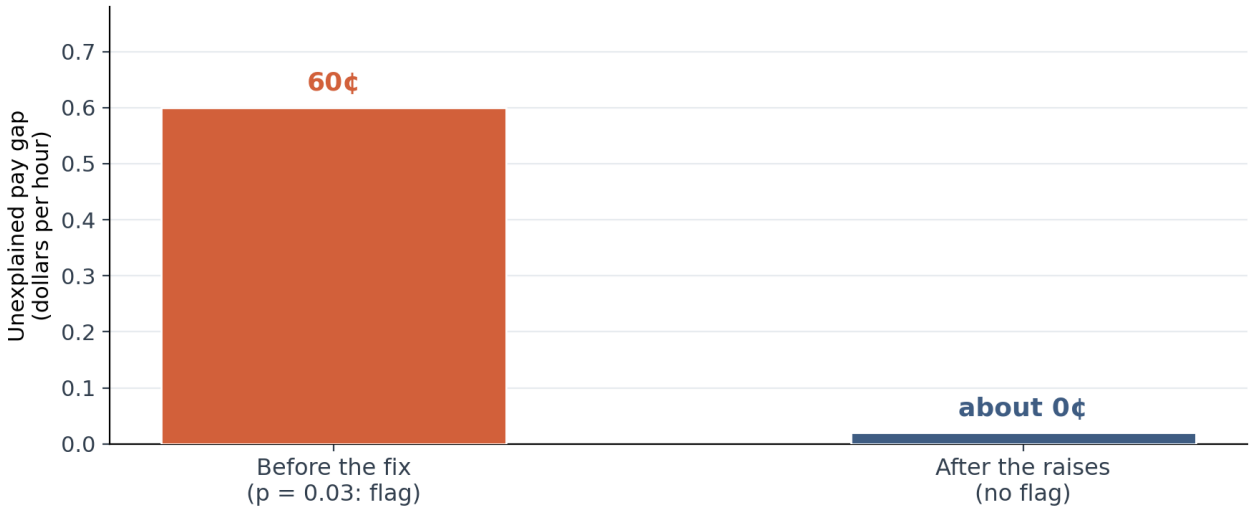
Finally, the company writes down what it did: how it made the groups, what facts it used, what it found, and how it fixed things. Then it marks its calendar to do the whole check-up again next year, and any time something big changes, like joining with another company. A pay equity audit works best like brushing your teeth: not once, but again and again.

How the Story Ends at Sunny Scoops

So what actually happened when Sunny Scoops ran its audit? The whole project took about six weeks. With the lawyer leading, the team built three groups: 40 scoopers, 12 shift leaders, and 5 store managers. The first-look t-test found no worrying gap among the shift leaders. The manager group was far too small for statistics, so those five were reviewed one by one. The scoopers, as you have seen, got flagged: a one-dollar gap between men and women, with a t-score of 2.1 and a p-value of 0.04.

Regression told the next part of the story. Some of that dollar was fair: the men really did average more years of experience, which the company pays for. But even after experience was counted, a 60-cent gap remained, with a p-value of 0.03, too big and too unlikely to be luck. The file review found a written reason for one person (Jake's food-safety certificate) but nothing at all for Maya and two of her coworkers. At the next pay cycle, all three received raises that closed their gaps. The one-by-one manager review also turned up a long-serving manager paid under her peers for no recorded reason, and she was adjusted too. All the fixes together cost about a quarter of one percent of total payroll.

Then Sunny Scoops fixed the leak, not just the puddle: it wrote simple starting-pay ranges for every job so managers stop guessing, and it marked the calendar to run the numbers again every summer. When the team re-ran the math after the raises, the unexplained gap had shrunk to roughly zero.



The scooper group's unexplained gap, before and after Sunny Scoops gave the missing raises.

Questions People Often Ask

How long does an audit take?

For a small company like Sunny Scoops, a few weeks. For a company with thousands of workers, several months. The slow part is almost never the math; it is gathering and double-checking the facts in Step 3.

Who gets to see the results?

While the work is happening, only the small team led by the lawyer sees the drafts, and privilege keeps those rough notes private. Workers see the outcome: raises where they were owed and fairer rules going forward. Keeping the drafts private is exactly what lets a company look honestly without being punished for finding its own problems.

What if the audit finds nothing wrong?

That is a happy ending, not a wasted trip. The company gains written proof that its pay holds up under real scrutiny, and some states give legal credit to companies that check themselves in good faith. The audit still repeats next year, because every new hire, raise, and promotion can slowly push pay off course.

Why would a company check on itself?

Four reasons. The law requires fair pay. New pay transparency laws mean workers can see pay ranges and spot gaps themselves. Finding and fixing a problem quietly costs far less than a lawsuit. And people work harder, and stay longer, at places they trust.

How much do the fixes usually cost?

It varies, but for a first audit the raises commonly add up to a small slice of payroll, often between one-tenth and one-half of one percent. Companies that fix their habits, like using starting-pay ranges, usually pay less each year after.

Can the math ever be wrong?

Yes, in two opposite ways. About 5 times out of 100, pure luck will cross the significance line and raise a false alarm. And in very small groups, a real problem can hide below the line. That is exactly why flags go to humans in Step 5, and why the audit repeats instead of happening once.

Why It All Matters

Fair pay is not just a rule; it is a promise a company makes to its team. When workers know the company checks its own fairness, they trust it more, work happier, and stay longer. And when everyone doing the same work gets a fair deal, the whole team wins.

Words to Know

Pay equity audit — A careful check-up to make sure people doing similar work are paid fairly, and that any differences have fair reasons.

Privilege — The legal shield that keeps the audit's rough notes private when a lawyer leads the work, so a company can look honestly and fix what it finds.

Comparator group — A group of workers who do the same kind of job under the same conditions, so their pay can fairly be compared. Apples with apples.

Fair reasons — Things that can fairly explain a pay difference, like years worked, experience, schooling, performance, or location. Grown-ups call these "legitimate factors."

Average — Everyone's pay added up and divided by the number of people.

Spread — How far people's numbers usually sit from the average. Measured with a number called the standard deviation.

t-test — A tool that compares two groups' averages and asks whether the gap is too big to blame on chance, given the spread and the group size.

t-score — The t-test's answer. The bigger the score, the more surprising the gap. Around 2 or more means "pay attention."

p-value — The chance of seeing a gap at least this big if pay were truly even. A small p-value means luck is a poor excuse.

Statistically significant — What we call a result whose p-value falls below 0.05. It means "probably not luck, so investigate," not "guilt proven."

Regression — The big machine that accounts for all the fair reasons at once, then checks whether gender or race still predicts who earns more.

Flag — A group or person the math marks for a closer human look. A flag is a question, not an answer.

Remediation — The fix: raising the pay of people who were underpaid, never cutting anyone's pay, and repairing the habits that caused the gaps.