THE FUTURE-READY CFO MASTERCLASS

MODULE 3 OF 20 | PILLAR I: AI, TECHNOLOGY AND DIGITAL INTELLIGENCE

Data, Analytics and Decision Intelligence

MODULE SUMMARY

This module covers the data foundations behind sound decisions: architecture, ERP integration, reporting tools, statistical reasoning, predictive models, data quality, and governance. It explains distributions, sampling, base rates, regression, overfitting, leakage, and drift, with Python verified examples on fictional pricing and churn data, using plain, simple definitions throughout. It supplies a data quality scorecard, a ten question challenge for any dashboard, and a method for linking data to decisions, decision rights, and feedback. A fictional case, boardroom questions, a data governance charter with a dashboard checklist, cited research from leading firms, and a ten question self assessment complete the module.

MODULE AT A GLANCE

Core question	How does a CFO build one trusted set of numbers that leaders use every day, and how does the CFO test the statistics, models, and dashboards built on those numbers before the organization acts on them?
Why it matters to the CFO	Every forecast, covenant report, pricing decision, and AI tool rests on data. A CFO who accepts a dashboard without challenge carries risk that no tool can remove.
Key sources	KPMG global survey (May 2026), PwC Ireland (August 2026), BDO Audit Innovation Survey, Workiva benchmark survey, Deloitte research, and NIST.
Executive artifact	A data governance charter and a dashboard challenge checklist, each shown as a completed example for the reader to adapt.

Important notice. This material is provided for educational and informational purposes only and does not constitute accounting, tax, legal, investment, or other professional advice. Readers should consult qualified professionals in accounting, tax, legal, and other relevant services before acting on any information presented. Fictional companies and figures are used for teaching and are labeled as such. The maturity ratings are an educational framework and have not been validated against a broad population of finance executives.

1. Executive overview and learning objectives

The core question. How does a chief financial officer build one trusted set of numbers that leaders use every day, and how does the CFO test the statistics, models, and dashboards built on those numbers before the organization acts on them? The question is older than any current technology, since a ledger has always been only as good as the entries in it. What has changed is the speed and volume with which numbers now reach decision makers, and the willingness of leaders to act on a chart without asking where it came from.

Why this matters over the next five years. KPMG reported in May 2026, from an online survey fielded in March 2026 of 1,013 senior finance leaders in 20 countries, that 75 percent of organizations now use AI actively in finance, up from 30 percent in 2024, and that 36 percent name data quality as both their biggest barrier and their biggest opportunity. The same survey found that fewer than half of organizations, 42 percent, were fully ready to provide assurance over AI enabled finance processes. PwC Ireland made the point in August 2026 in plain terms: data is no longer simply a reporting asset but a strategic one, and speed creates value only when leaders trust the information they act upon. A CFO who deploys the best tools on unreliable data will produce wrong answers faster, and will do so under a more impressive interface.

Terms used in this module. Data architecture is the design of how data is collected, stored, connected, and delivered to users. **Analytics** is the use of data and statistical methods to describe what has happened and to estimate what may happen. **Decision intelligence** is the discipline of connecting data and analysis to specific decisions, with named owners and a record of whether the decisions worked. Each term receives a fuller definition where it first matters. The module treats the three as links in one chain, because a perfect dataset that informs no decision has no value, and a well designed decision process fed by poor data merely automates error.

Learning objectives. After completing this module the reader should be able to do the following. Each objective is tested by the self assessment in Section 8.

1. Describe the layers of a finance data architecture, including sources, integration, warehouse or lakehouse, semantic layer, and master data, and explain what a single source of truth does and does not mean.
2. Explain the main ways an ERP ecosystem is integrated with other systems, and the reconciliation controls that keep reporting tied to the ledger.
3. Apply core statistical ideas to finance questions, including distributions, sampling and confidence intervals, base rates, correlation versus causation, and regression, and interpret a regression table with confidence.
4. Recognize the pitfalls of predictive models, namely overfitting, leakage, and drift, and require a backtest and a naive benchmark before relying on any model.
5. Measure data quality across defined dimensions with a scorecard, and set up data governance with owners, stewards, lineage, access, retention, and privacy rules.
6. Question the numbers behind any dashboard with a ten question challenge, and link data to decisions through decision rights, decision logs, and feedback loops.

How the module connects to the others. Module 1 explains why generative systems err with confidence, and the discipline in this module is the practical defense, since a model fed clean, defined, traceable data errs less often and can be checked when it does. Module 2 describes the continuous close and predictive planning, none of which can be trusted without the data foundations here. Module 10 treats enterprise risk, privacy, and model risk, and the governance charter in this module is the finance side of that work. Module 19 builds the performance system, including the design of key measures and unit economics, on the numbers whose definitions this module teaches the CFO to fix.

2. Conceptual foundations and frameworks

The single source of truth, correctly understood. The phrase is used loosely and often misleadingly. A company rarely has one database. It has one *agreed definition* of each number that matters, one named owner of that definition, and one place from which the approved figure is published. Revenue, a customer, a headcount, and churn may be stored in several systems, and that is acceptable, provided that everyone who reports them draws from the same definition and the same approved calculation. The goal is therefore a trusted set of numbers and not a single piece of technology. Deloitte, in a management reporting survey of more than 600 senior finance professionals published in March 2020, observed that weak data governance undermines finance outputs and leads to rework and manual manipulation, and that fixing it typically requires changes to culture, technology, and process together.

Data architecture for finance. Exhibit 1 shows the layers in the order that data travels. A **data warehouse** is a store of structured, cleaned data organized for reporting, in which the structure is decided before the data is loaded. A **data lake** holds raw files of any type, including documents and logs, with the structure applied when the data is read. A **lakehouse** is a design that places warehouse style controls and query performance over lake style storage, so that one platform can serve both reporting and data science. The terms are used differently by different vendors, and a CFO should ask which controls a proposal actually provides and not rely on the label. The **semantic layer** is the most important layer to a CFO. It is the place where business definitions, such as net revenue or active customer, are written once as calculations so that every report and every analyst uses the same one. **Master data** is the small set of reference records that describe the business itself, namely customers, vendors, products, employees, legal entities, and the chart of accounts.

Exhibit 1 Layers of a Finance Data Architecture

Layer	What it does	Finance example	Principal failure
Sources	Systems that create transactions and records	General ledger, subledgers, billing, CRM, payroll, bank feeds, spreadsheets	Shadow spreadsheets that hold the real number and are known to one person
Integration	Moves and standardizes data between systems on a schedule or event	Nightly extract of invoices from three ERP instances. Daily bank files	Silent load failures. Totals that no longer tie to the source
Storage: warehouse or lakehouse	Holds history in a governed form for analysis	Seven years of invoice lines, mapped to one chart of accounts	A lake of unlabeled files that no one can safely use
Master data	Holds one clean record per customer, vendor, product, and account	One customer record linked to every ERP account for that customer	Duplicates that split revenue, credit exposure, and payments
Semantic layer	Stores approved business definitions as shared calculations	Net revenue, gross margin, active customer, churn, days sales outstanding	Each team builds its own formula and the results differ
Consumption	Delivers reports, dashboards, models, and extracts to users	Board pack, covenant report, sales dashboard, forecast model	Users export to a spreadsheet and the controls end there
Governance and security (all layers)	Sets ownership, quality rules, access, lineage, and retention	Named owner for net revenue. Access review. Lineage from report to ledger	Rules written but not operated, with no evidence

Source: author's synthesis of common practice. The definitions of warehouse, lake, lakehouse, and semantic layer are the author's own and vendors use them differently.

ERP ecosystems and integration. An enterprise resource planning system, or ERP, is the software that records the core transactions of a business, such as orders, invoices, payments, inventory, and the ledger. Few companies have only one. Acquisitions leave several, and each is surrounded by billing tools, planning tools, expense tools, and specialist applications. That collection is the **ERP ecosystem**, and the quality of the connections among its parts matters as much as the quality of any single system. Exhibit 2 describes the common integration patterns and the control that each requires.

Exhibit 2 Common ERP Integration Patterns and Their Controls

Pattern	How it works	Strength	Risk and control
Manual export to spreadsheet	A person downloads reports and combines them by hand	Fast to start and flexible	Error, version confusion, no audit trail. Control: lock, review, and reconcile to the ledger
Point to point interface	One system sends a defined file or message directly to another	Simple for two systems	Becomes a tangle as systems grow. Control: inventory of interfaces and an owner for each
Integration platform	A central service moves and transforms data among many systems under common rules	One place to monitor, log, and change	Single point of failure. Control: monitoring, alerting, and change approval
Extract to warehouse or lakehouse	Scheduled loads copy data into a governed store for reporting	Separates reporting from transaction systems	Timing differences and missed loads. Control: record counts and amount totals reconciled after each load
Application programming interface and event streams	Systems exchange data in near real time through defined connections	Current data for continuous close and monitoring (Module 2)	Volume and security exposure. Control: access tokens, rate limits, logging

Source: author's synthesis of common practice.

The reconciliation control. Whatever pattern is used, one control is indispensable. After each load, the total of every key measure in the reporting store must be compared with the total in the source system, and any difference above a stated tolerance must stop publication until explained. This control costs little and catches a surprising range of failures, including a load that ended early, a currency table that did not update, and an entity that was left out of a join. Many dashboard errors that reach a board began as a load that nobody checked.

Business intelligence and self service reporting. Business intelligence, or BI, is the set of tools that turn data into reports, charts, and dashboards. **Self service** means that business users build their own analyses without waiting for a technical team. The benefit is speed and curiosity, and the hazard is that fifty people will compute fifty versions of revenue. A sound design offers tiers. In Exhibit 3 the top tier is a small set of certified reports built on the semantic layer, owned and tested by finance. Below it are departmental analyses that use certified data but add local logic, and below that are personal workspaces that are labeled as unreviewed. The rule that makes the tiers work is that anything shown to the board, a lender, an auditor, or a regulator must come from the certified tier.

Exhibit 3 **Tiers of Self Service Reporting**

Tier	Who builds it	Data used	Control and label
Certified	Finance or a central analytics team	Governed warehouse and semantic layer only	Tested, versioned, owner named, certified label, reconciled to the ledger each load
Departmental	Trained analysts in a business function	Certified data plus approved local sources	Peer review. Definitions must match the semantic layer or state the difference
Personal	Any user	Whatever the user can reach within access limits	Marked unreviewed. May not be shown outside the team or used for external reporting

Source: author's synthesis of common practice.

Statistical reasoning for CFOs. A CFO is not asked to compute a regression by hand. A CFO is asked to read one, and to notice when a chart implies more than the data can support. Exhibit 4 lists the ideas that matter most, the question to ask about each, and the error that most often follows. A **distribution** is the pattern of how values are spread, and in finance it is frequently skewed, with a few very large items pulling the average above the typical case. **Correlation** is the tendency of two measures to move together, and it says nothing by itself about whether one causes the other. A **sample** is a subset used to learn about a larger population, and a **confidence interval** is a range that, under stated assumptions, would contain the true value in a stated share of repeated samples. A **base rate** is how common something is in the population before any new evidence is considered. **Regression** is a method of estimating how one measure changes, on average, when another changes, while holding others constant.

Exhibit 4 **Statistical Ideas Every CFO Should Be Able to Test**

Idea	Question a CFO should ask	Common error
Distribution	Is the average typical, or is it pulled by a few large items? What do the median and the spread show?	Planning on the mean of a skewed set, such as customer size or deal value
Correlation and causation	What else changed at the same time? Could the relationship run the other way?	Crediting a price move for a volume change that a promotion caused
Sampling	How was the sample drawn, how large is it, and how wide is the interval around the result?	Reporting a sample average as if it were exact, or sampling only the easy items
Confidence interval	Does the range include zero, or the threshold that would change our decision?	Treating the middle of the range as certain
Base rate	How common is the event before the signal? What share of alerts are real?	Trusting an alert with a strong sensitivity when the event is rare
Regression	What is the effect size, its uncertainty, and what was held constant? Does it hold out of sample?	Reading a high R squared as proof of cause, or omitting a driver that matters
Statistical versus practical significance	Is the effect large enough to change a decision at our scale?	Acting on a tiny effect because it is statistically detectable in large data

Source: author's synthesis. The definitions are standard statistical concepts stated from the author's own knowledge.

Predictive modeling and its pitfalls. A predictive model learns a pattern from past data and applies it to new cases, for example to forecast revenue, to estimate which customers will leave, or to flag a risky payment. The

model is useful only to the extent that the pattern it learned will continue to hold, and four failures account for most disappointments. **Overfitting** occurs when a model is tuned so closely to past data that it memorizes noise and fails on new data. **Leakage** occurs when the model is given, during training, information that would not be available at the moment of the real decision, which makes it look brilliant in testing and useless in use. **Drift** occurs when the world changes and the relationships the model learned no longer hold. **Backtesting** is the defense: the model is built on an earlier period, tested on a later period it has not seen, and compared with a simple naive alternative.

Exhibit 5 Predictive Model Pitfalls, Warning Signs, and Defenses

Pitfall	What happens	Warning sign	Defense
Overfitting	Model fits history almost perfectly and misses new periods badly	In sample error far below out of sample error. Many parameters for little data	Hold out recent data. Prefer the simpler model that performs about as well
Leakage	A field that is only known after the event is used to predict the event	Accuracy that seems too good. A top driver that is a consequence of the outcome	Time stamp every input and use only what was known at decision time
Drift	Customers, prices, or processes change and the model decays	Stability measures move. Errors rise. Overrides increase	Monitor input and error stability. Set a retraining and retirement rule
Survivorship and selection	Only the cases that remained in the data are studied	History that omits lost customers, closed entities, or failed projects	Include the exits. Ask what is missing from the sample
Ignoring the base rate	Alerts are trusted although the event is rare	Many alerts, few findings	Report precision and recall, and size the review team accordingly

Source: author's synthesis of standard practice. Module 10 treats model risk governance and Module 1 treats the limits of generative systems.

Data quality. Data quality means that data is fit for the use to which it is put. It has several dimensions, and a company that measures only one will be surprised by the others. **Completeness** asks whether the required fields are filled. **Validity** asks whether values follow the required format and rules. **Uniqueness** asks whether each real thing appears once. **Timeliness** asks whether the data is current enough for the decision. **Consistency** asks whether the same fact agrees across systems. **Accuracy** asks whether the data matches reality, and it is the hardest to test because it requires external evidence such as a contract, a bank record, or a physical count. Section 3 turns these into a scorecard.

Data governance. Data governance is the set of decisions about who may do what with which data, and the means of making those decisions stick. Exhibit 6 sets out its elements. The central idea is **ownership**. An owner is a senior business person accountable for a data domain, such as customers or revenue, who approves its definitions and quality targets. A **steward** is the practitioner who maintains the data day to day, resolves issues, and enforces the standards for the owner. A **custodian**, usually in technology, operates the systems that hold the data. **Lineage** is the documented path from a reported number back through each transformation to its source, so that anyone can see how it was made. PwC Ireland, in August 2026, framed the CFO's agenda as strengthening the data foundation and giving data clear ownership, and the Workiva survey of February 2026 reported that 96 percent of respondents believed the CFO, the chief information officer, and the chief sustainability officer must unite around a shared data governance strategy.

Exhibit 6 Elements of Data Governance

Element	Purpose	What good looks like	Evidence
Ownership and stewardship	Fixes accountability for each data domain	Named owner and steward for customers, vendors, products, revenue, and the chart of accounts	Owner register approved by the executive sponsor
Definitions and standards	Gives each measure one meaning	Business glossary tied to the semantic layer, with change history	Glossary entry with owner and effective date
Data quality management	Detects and corrects defects	Scorecard by domain, targets, and a log of issues with owners and dates	Quarterly scorecard and issue log
Lineage and change control	Shows how a number is made and who changed the logic	Traceable path from dashboard measure to ledger account. Changes approved and tested	Lineage map. Change tickets
Access and security	Limits data to those with a need	Least privilege roles, periodic review, logging of sensitive access	Access review sign off (Module 10)
Retention and disposal	Keeps data as long as required and no longer	Schedule by record type set with legal counsel, with documented disposal	Retention schedule and disposal log
Privacy	Protects personal information	Inventory of personal data, lawful basis recorded, minimization, response process for individual requests	Privacy inventory reviewed with counsel

Source: author's synthesis of common practice. Legal requirements for retention and privacy vary by jurisdiction and industry, and the reader must consult qualified legal counsel.

Decision intelligence. Most companies measure what they can easily count and decide by custom. Decision intelligence reverses the order. It begins with the decisions that matter, for example whether to raise prices, extend credit, hire, or acquire, and works backward to the data and analysis each requires. Three practices define it. First, each recurring decision has a **decision right**, which names who recommends, who decides, who must be consulted, and who is told. Second, the decision is recorded with the data used, the expected result, and a date for review, in a **decision log**. Third, the actual result is compared with the expectation, and the difference feeds back into the data, the model, or the judgment, which forms a **feedback loop**. Exhibit 7 shows the design for four common finance decisions.

Exhibit 7 Linking Data to Decisions: Rights and Feedback

Decision	Data and analysis needed	Decision right	Feedback measure and review date
Raise list prices by a stated percent	Price and volume history with promotions controlled, regression effect and interval, margin impact	Chief commercial officer recommends. CFO and CEO decide. Sales is consulted	Realized volume and margin against estimate, at two quarters
Extend credit above standard terms	Payment history, credit score, concentration, aging, base rate of loss for similar accounts	Credit manager recommends. CFO decides above a stated limit	Loss and days late against expectation, at twelve months
Approve a hire in a function	Workload driver, productivity trend, forecast of volume, cost per head	Function head recommends. CFO and head of people decide	Output per head and volume handled, at two quarters
Retire or retrain a forecast model	Backtest results, bias, stability of inputs, override rate	Model owner recommends. FP&A head decides. Internal Audit is informed	Error against the naive benchmark each quarter

Source: author's synthesis. The decision rights shown are examples and each company must set its own to fit its governance (Module 9 and Module 13).

3. Practical applications and tools

The six methods that follow are each shown with a worked example on fictional data. Northgate Supply, Inc. is a fictional distributor with three ERP instances and about 4,800 invoices in the period examined. Every figure was computed in Python, and the arithmetic in the text and exhibits was generated from that computation.

Method one: look at the distribution before the average. Exhibit 8 summarizes the annual contract value, in thousands of dollars, of a list of 25 Northgate customers. The mean is 80.6 thousand dollars and the median, the middle value, is 29.0 thousand dollars. The gap is large because a few customers are very large. The three largest account for 60.0 percent of the total, and 20 of the 25 customers are below the mean. The standard deviation, a measure of spread, is 138.3 thousand dollars, larger than the mean itself, which signals a heavily skewed set.

Exhibit 8 Northgate Supply, Inc. (Fictional): Distribution of Annual Contract Value

Measure	Value (thousands of dollars unless stated)
Number of customers	25
Total annual contract value	2,015
Mean	80.6
Median	29.0
Standard deviation	138.3
90th percentile	182.0
Largest customer	620
Skewness (zero for a symmetric set)	2.94
Share of total held by the three largest, percent	60.0
Customers below the mean	20 of 25

Source: fictional data, computed in Python with population skewness. Contract values in thousands of dollars: 12, 14, 15, 16, 18, 18, 19, 21, 22, 24, 25, 27, 29, 31, 34, 38, 42, 47, 55, 68, 90, 140, 210, 380, 620.

What the distribution teaches. A pricing or service plan built for the average customer of 80.6 thousand dollars would fit almost no one, since the typical customer is nearer 29.0 thousand dollars and the revenue is concentrated in a handful of accounts. The CFO should draw two conclusions. Averages of skewed data should be reported with the median beside them, and concentration is itself a risk that belongs on the board agenda, since the loss of one of the top three would remove a large share of the total.

Method two: size a sample and state the interval. Suppose an analyst examines 40 of Northgate's 4,800 invoices and records the billing error on each, defined as the amount billed less the correct amount, in dollars. The mean error is -5.72 dollars per invoice with a standard deviation of 49.23 dollars. The standard error, which is the standard deviation divided by the square root of the sample size, is 7.78 dollars. With 39 degrees of freedom the multiplier for a 95 percent interval is 2.023, so the interval runs from -21.47 to +10.02 dollars per invoice. Because the interval includes zero, this sample cannot show whether the average error is an overcharge, an undercharge, or nothing at all. Scaling to all invoices gives an estimated net error of -27,480 dollars, with an interval from -103,051 to +48,091 dollars, which is far too wide to book anything.

Exhibit 9 Northgate Supply, Inc. (Fictional): How the Sampling Interval Narrows with Sample Size

Sample size	Half width of the 95 percent interval, dollars per invoice	Relative to the 40 invoice sample
10	35.22	2.24
40	15.74	1.00
160	7.69	0.49
640	3.82	0.24

Source: fictional data, computed in Python holding the sample standard deviation of 49.23 dollars constant. Student t multipliers are used.

The lesson of the interval. Quadrupling the sample roughly halves the width of the interval, so precision becomes expensive quickly. A sample of 160 would narrow the half width from 15.74 to about 7.69 dollars, and one of 640 to about 3.82 dollars. The CFO should therefore decide first how precise an answer is worth, and then size the sample. The same reasoning applies to a customer survey, a quality inspection, and an audit test. A statement such as the error rate is 2 percent has no meaning until the interval and the sampling method are stated, and a sample drawn from the easiest items is not a random sample at all.

Method three: apply the base rate. Northgate screens 20,000 payments for problems with an analytic test. Assume that 1 percent of payments, or 200, are genuinely wrong, that the test catches 90 percent of the wrong ones, and that it falsely flags 5 percent of the correct ones. These are illustrative rates and not measured ones. Exhibit 10 follows the counts. The test flags 1,170 payments, of which 180 are real, so only 15.4 percent of the flags are genuine, even though the test sounds accurate. The other 990 flags are false alarms that consume the review team's time, and 20 real problems are still missed.

Exhibit 10 Northgate Supply, Inc. (Fictional): A Sensitive Test on a Rare Event

Step	Count	Basis
Payments screened	20,000	Given
Genuinely wrong payments	200	1 percent of 20,000
Wrong payments flagged (true positives)	180	90 percent of 200
Correct payments flagged (false positives)	990	5 percent of 19,800
Total flags raised	1,170	180 plus 990
Share of flags that are genuine, percent	15.4	180 divided by 1,170
Wrong payments not flagged	20	200 less 180

Source: fictional rates, computed in Python. Sensitivity of 90 percent, false positive rate of 5 percent, and prevalence of 1 percent are assumptions for illustration.

What the base rate teaches. When an event is rare, even a good test produces mostly false alarms. A CFO who reads that a fraud model is 90 percent accurate has learned very little until the share of flags that are real is stated, and the review team must be sized for the false alarms as well as the findings. The remedy is to measure precision, the share of flags that are real, and recall, the share of real problems caught, on the company's own history, and to tune the threshold with the cost of a miss and the cost of a review both in view.

Method four: read a regression, and separate correlation from cause. Northgate wishes to know how volume responds to a change in list price. Exhibit 11 lists 16 region quarters, each with the price change in percent, whether a promotion ran, and the resulting change in unit volume in percent. Note that promotions

ran in 7 of the 16 cases and almost always together with a price cut.

Exhibit 11 Northgate Supply, Inc. (Fictional): Price Changes, Promotions, and Volume by Region Quarter

Region quarter	Price change, percent	Promotion ran	Volume change, percent
R01	+2.9	No	-2.7
R02	-3.9	Yes	+10.2
R03	+5.9	No	-4.0
R04	-3.9	Yes	+11.1
R05	-2.0	Yes	+7.8
R06	+2.1	No	-1.8
R07	+3.3	No	-2.9
R08	-3.8	Yes	+12.0
R09	+1.8	No	-1.5
R10	-1.0	Yes	+6.7
R11	+4.8	No	-3.9
R12	+5.3	No	-5.0
R13	-3.9	Yes	+10.1
R14	-3.1	Yes	+8.1
R15	+4.3	No	-3.7
R16	-1.0	No	+1.6

Source: fictional data used in the Python regression below. The correlation between price change and volume change is -0.97, and between price change and the promotion flag is -0.88.

The simple regression of volume change on price change gives a slope of -1.70, meaning that each one percentage point rise in price is associated with a fall of about 1.70 percentage points in volume, with an R squared of 0.94. This looks decisive, and it is misleading, because price cuts and promotions arrived together and the promotion itself lifted volume. When the promotion is added as a second explanatory variable, the price slope becomes -0.99, and the promotion carries an effect of +5.79 percentage points of volume. The first answer credited price for a volume change that the promotion had caused, so that the market looked more sensitive to price than it is.

Exhibit 12 Northgate Supply, Inc. (Fictional): Regression Results Without and With the Promotion Control

Term	Model 1 coefficient	Model 1 standard error	Model 2 coefficient	Model 2 standard error
Intercept	+3.46	0.41	+0.58	0.49
Price change, per percentage point	-1.70	0.12	-0.99	0.12
Promotion ran (1 if yes)	Not included		+5.79	0.90
95 percent interval for price	-1.95 to -1.46		-1.26 to -0.72	
t statistic for price	-14.8		-7.9	

Term	Model 1 coefficient	Model 1 standard error	Model 2 coefficient	Model 2 standard error
R squared	0.940		0.986	
Observations and degrees of freedom	16 and 14		16 and 13	

Source: fictional data, ordinary least squares computed in Python with numpy, and 95 percent intervals from the Student t distribution. The slope is in percentage points of volume per percentage point of price.

From the regression to the decision. Consider a proposal to raise list prices by 3 percent. Under the first model, volume would fall by about 5.1 percent and revenue would change by -2.3 percent, which argues against the increase. Under the second model, volume would fall by about 3.0 percent and revenue would change by -0.1 percent, so the increase is roughly neutral for revenue and positive for margin, since fewer units are sold at a higher price. A slope of minus one is the point at which a small price rise leaves revenue unchanged, and the second model sits close to it. The decision turns entirely on a variable that the first analysis omitted, and the CFO who asks what else changed at the same time will have asked the most valuable question in statistics. The interval for the second slope, -1.26 to -0.72, also deserves attention, since a decision that is sound only at one end of the interval should be tested in a small pilot first. Even then the result is an association in 16 observations and is not proof of cause. A controlled test, in which prices differ across comparable regions by design, would be stronger evidence.

Method five: backtest, and compare with the naive benchmark. Exhibit 13 uses a fictional revenue index for 24 months. Each model was built on the first 18 months only, and then tested on the last 6, which it had not seen. Four polynomial trend models of rising complexity were fitted, and the naive forecast, which repeats the last known value, is shown for comparison. Error is the mean absolute percentage error, the average of the absolute percentage miss.

Exhibit 13 Fictional Revenue Index: In Sample and Out of Sample Error by Model

Model	In sample error, percent	Out of sample error, percent	Reading
Straight line trend	3.10	1.35	Simple. Stable out of sample. Beats naive
Second degree curve	2.57	10.21	Better fit in sample, worse outside
Fourth degree curve	1.91	29.82	Memorizing noise
Sixth degree curve	1.37	88.81	Best fit in sample and useless outside
Naive, repeat last value	3.14	1.79	Minimum standard to beat

Source: fictional monthly series, computed in Python with numpy polynomial fitting on a scaled time axis. In sample error covers months 1 to 18 and out of sample error covers months 19 to 24. The naive in sample error covers months 2 to 18.

Reading the backtest. The sixth degree curve had the lowest in sample error, 1.37 percent, and an out of sample error of 88.81 percent, which is about 66 times that of the straight line at 1.35 percent. A buyer shown only the in sample figure would choose the worst model. The straight line also beat the naive forecast outside the sample, 1.35 percent against 1.79 percent. The CFO should require a vendor or an internal team to show out of sample results by period, to name the benchmark, and to state how much data was set aside before the model was built. A good result in a later test, on a period that arrived after the model was frozen, is the best evidence available.

Detecting drift. A model that passed its backtest can still decay. A common monitoring measure is the **population stability index**, or PSI, which compares the distribution of an input today with its distribution when

the model was built. For each band, the index adds the product of the difference in shares and the natural logarithm of the ratio of shares. Exhibit 14 applies it to the mix of invoice sizes at Northgate. Practitioners often treat a value below 0.10 as stable, between 0.10 and 0.25 as a shift to investigate, and above 0.25 as a major shift, but these are conventions and not laws, and the CFO should set thresholds with the model owner.

Exhibit 14 Northgate Supply, Inc. (Fictional): Population Stability Index of Invoice Size Mix

Invoice size band, dollars	Share when model was built	Share this quarter	Contribution to PSI
Under 1,000	20.0 percent	12.0 percent	0.041
1,000 to 4,999	30.0 percent	24.0 percent	0.013
5,000 to 19,999	25.0 percent	26.0 percent	0.000
20,000 to 49,999	15.0 percent	20.0 percent	0.014
50,000 and over	10.0 percent	18.0 percent	0.047
Population stability index	**100.0 percent**	**100.0 percent**	**0.116**

Source: fictional shares, computed in Python. The thresholds of 0.10 and 0.25 are common practitioner conventions stated from the author's own knowledge and not from a source read.

The index of 0.116 falls in the range that calls for investigation. Invoices of 50,000 dollars and above grew from 10 to 18 percent of the mix, and a model trained when large invoices were rare may now misjudge them. The right response is not automatic retraining. It is to find out why the mix changed, whether a new customer segment or a pricing change explains it, and whether the model still performs on the new mix.

Method six: score data quality. Exhibit 15 is a scorecard for Northgate's customer master of 5,000 records. Each of the six dimensions is tested with a stated rule on a stated number of records, compared with a target, and weighted by its importance to the decisions the data supports. The weighted score is 90.9 percent against a weighted target of 97.8 percent, and none of the six dimensions meets its target. The duplicate rate of 4.2 percent deserves attention first, since duplicates split credit exposure and revenue across records and can lead to double payment.

Exhibit 15 Northgate Supply, Inc. (Fictional): Data Quality Scorecard for the Customer Master

Dimension	Test applied	Tested	Failed	Result, %	Target, %	Weight	Status
Completeness	Tax identifier, billing country, and payment terms all populated	5,000	385	92.3	98.0	20	Below target
Validity	Postal code and tax identifier pass format and checksum rules	5,000	240	95.2	99.0	15	Below target
Uniqueness	Records that are not duplicates of another active customer	5,000	212	95.8	99.0	20	Below target
Timeliness	Active customer records reviewed or updated within twelve months	5,000	650	87.0	95.0	10	Below target
Consistency	Customer name and credit limit agree across the three ERP instances	1,800	306	83.0	97.0	20	Below target

Dimension	Test applied	Tested	Failed	Result, %	Target, %	Weight	Status
Accuracy	Sample of 250 records agrees with contract and bank evidence	250	22	91.2	98.0	15	Below target
Weighted score				**90.9**	**97.8**	**100**	

Source: fictional data, computed in Python. Tested is the number of records tested and Failed the number failing. Result is records passing divided by records tested, in percent. The weights and targets are proposed for illustration and each company should set its own according to the use of the data.

Using the scorecard. The scorecard is only useful if it drives action. Each failure should map to a root cause, such as a free text field that allows a missing tax identifier, or a customer created in two ERP instances without a check. The owner then sets a remediation date, and the score is re-measured every quarter. Targets should differ by the importance of the data, so that the fields used for credit and tax have tighter targets than a marketing field. Finally, the CFO should insist that a score be tied to a use. A customer record that is 99 percent complete but wrong in credit limit is worse than one that is 95 percent complete and right.

The dashboard challenge: ten questions. A dashboard is a claim in graphical form. The CFO should treat it as such, and ask the ten questions in Exhibit 16 before accepting what it shows. The questions are the practical summary of the module, and Section 6 turns them into a checklist that the reader completes.

Exhibit 16 The Dashboard Challenge: Ten Questions

No.	Question	What a good answer sounds like
1	What decision is this dashboard meant to inform, and who owns that decision?	A named decision and owner. Not 'to give visibility'
2	What is the exact definition of each measure, and where is it written?	A glossary entry tied to the semantic layer, with an owner and a date
3	Where does each number come from, and does it reconcile to the ledger or source system?	A lineage map and a reconciliation that ran after the last load
4	How current is the data, and when was it last refreshed successfully?	A visible refresh time and a rule for stale data
5	What is excluded, filtered, or missing?	A list of entities, segments, and late items left out, with reasons
6	What is the comparison, and is it fair?	Same period, same basis, same definition, with the base rate stated
7	How large is the change against normal variation?	The usual range of the measure, so that noise is not read as a trend
8	Is the relationship causal, or only correlated, and what else changed?	A test or a control, or an honest statement that cause is not shown
9	Who last changed the logic, and what control governs changes?	Change history with approver and test evidence
10	What would we do differently if the number were 10 percent better or worse?	A stated threshold for action, and the record of past actions

Source: author's synthesis of common practice. The ten questions are a teaching tool and not a published standard.

Putting the method to work. A CFO can apply the six methods in a fixed order when a new analysis arrives. Begin with the definitions and the data quality of the inputs, then look at the distribution, then ask how large the sample was and how wide the interval is, then examine the base rate, then ask what else changed, and finally ask how the analysis performed out of sample. An analysis that survives all six is rarely wrong in a way that matters, and one that fails the first is not worth examining further. The order is deliberate, since the early questions are cheap to ask and the late ones are expensive to answer.

4. Executive case study and decisions

Fictional case: Cedar Vale Distribution, Inc. The names, companies, and figures in this case are fictional and exist only for teaching. Cedar Vale is a distributor with annual revenue of 480 million dollars, owned by a private equity sponsor, with three ERP instances inherited from acquisitions, two billing tools, and a customer base that appears in the three systems under inconsistent names. The sponsor requires a monthly report to its lenders that includes customer churn, and the lender covenant package refers to it.

The problem. The sales dashboard shows churn for the quarter at 6.1 percent, down from 8.0 percent in the prior quarter, and the chief commercial officer has attributed the improvement to a new retention program. The CFO is asked to confirm the figure for the lender package. A review of the logic finds that the dashboard counts a customer as lost only when a formal cancellation notice is received. Of the 1,250 customers at the start of the quarter, 76 gave notice. A further 29 customers placed no order for more than 90 days and have not reordered, and they are counted as active because they never cancelled. The prior quarter's figure was computed by the finance team, which treated 90 days without an order as a loss.

Exhibit 17 Cedar Vale Distribution, Inc. (Fictional): Churn Under Two Definitions

Measure	Dashboard definition	Finance definition
Customers at start of quarter	1,250	1,250
Customers who gave formal notice	76	76
Customers with no order for 90 days and no notice	Counted as active	29
Customers lost	76	105
Churn rate, percent	6.1	8.4
Prior quarter churn on finance definition, percent	8.0	8.0
Change on a like for like basis, percentage points	Not comparable	+0.4

Source: fictional data, computed in Python. Churn is customers lost divided by customers at the start of the quarter.

Reading the table. On a like for like basis the churn rate rose slightly, from 8.0 to 8.4 percent, and the reported improvement came entirely from a change of definition. No one had acted in bad faith. The dashboard had been built by a sales operations analyst for sales purposes, with a reasonable definition for a sales team that wishes to count only confirmed losses. The failure was that the same word, churn, had two meanings and the figure was about to be reported to lenders. Five of the ten dashboard questions would have caught the problem: the definition, the lineage, what is excluded, whether the comparison is fair, and who changed the logic.

The decision. The CFO must correct the figure, and must also decide how to prevent the next occurrence. The three ERP systems cause the root difficulty, since customer records are duplicated and sales and finance each maintain their own mapping. Exhibit 18 compares three options over five years, in millions of dollars, with a net present value at 10 percent and a second value in which benefits are 70 percent of plan. Benefits are

fictional estimates of avoided write offs, fewer duplicate payments, reduced reconciliation effort, and better pricing decisions.

Exhibit 18 Cedar Vale Distribution, Inc. (Fictional): Options Compared, Millions of Dollars

Option	One time cost	Year 1 net	Year 5 net	Payback (years)	NPV at 10 percent	NPV, benefits at 70 percent
A. Replace all three ERPs and build one enterprise lakehouse	8.2	(4.50)	4.50	4.40	0.16	(3.38)
B. Master data, semantic layer, and governance in waves	3.6	(1.40)	2.90	3.00	3.54	0.95
C. New dashboard tool on top of current data	0.5	(0.45)	0.45	2.00	0.89	0.20

Source: fictional example, computed in Python. Option A: one time 4.0, 3.4, 0.8, run 0.5, 1.3, 1.5, 1.5, 1.5, benefits 0, 1.2, 4.0, 6.0, 6.0. Option B: one time 1.6, 1.4, 0.6, run 0.2, 0.5, 0.7, 0.7, 0.7, benefits 0.4, 1.6, 3.0, 3.6, 3.6. Option C: one time 0.5, run 0.25 each year, benefits 0.3 then 0.7. Net equals benefit less run cost less one time cost.

Reading the options. Option A is the most ambitious. It requires 8.2 million dollars of one time spending, a net present value of 0.16 million dollars in the base case, and a value of (3.38) million dollars when benefits fall 30 percent short, and it would take two years before any process in the business changed. Option B costs 3.6 million dollars, pays back in 3.00 years, has the highest net present value at 3.54 million dollars, and remains positive at 0.95 million dollars in the downside. Option C is cheapest and has a net present value of 0.89 million dollars, and 0.20 million dollars in the downside, but its small benefit is the most doubtful of the three, because a new dashboard tool on top of the same inconsistent data produces more attractive charts and the same disagreements. It would have made the churn problem easier to display and no easier to prevent.

The recommended path. The CFO should first correct the lender figure with a written reconciliation of the two definitions, and tell the sponsor and the chief commercial officer plainly before the lender sees it. The CFO should then recommend Option B. The reasoning has four parts. First, the root cause is the absence of an owner and a definition, and no platform supplies either. Second, master data for customers, a semantic layer holding the approved churn, revenue, and margin definitions, and a certified tier of reports can be delivered within a year and cover the figures that lenders and the board see. Third, the downside case shows that Option B is the only choice that stays positive if benefits fall short, and that Option A has the largest regret. Fourth, a staged program lets the company stop or redirect at each gate, and it leaves the larger question of consolidating the three ERP instances to be decided later on evidence, with Module 4 supplying the technology return framework.

Governance actions that accompany the choice. The CFO should name the chief commercial officer as the owner of the customer domain and the finance controller as the owner of the revenue and churn definitions, with the chief information officer as custodian. Every board and lender measure should carry a glossary entry and a certified label. A change to any certified definition should require approval from the owner, a recorded test, and a note in the change log, and a restated prior period should be disclosed to those who received the earlier figure. Customer matching rules should be applied across the three systems, with the duplicates found in the first pass reviewed by a steward and not merged automatically.

What a weaker answer misses. A weaker answer reports the improved figure because the dashboard said so, and congratulates the sales team. It then buys a better dashboard tool, since the visible symptom was a chart. It treats the lender package as a formatting task and not a control, and it does not ask who owns the definition. It commits the entire budget to a platform on the first day, ignores the downside case, and does not record the decision or its expected result. Each of these omissions is common and each is avoided by the discipline of

definitions, owners, certified reports, and gates.

Questions for reflection. Which of the ten dashboard questions would have caught the problem soonest, and what would that have cost? How would you tell the chief commercial officer that the improvement did not occur, and what would you say to the lender? If the sponsor insisted on Option A because it promised a single platform, what evidence would change your recommendation? Write your answers before reading the boardroom questions in Section 5.

5. Boardroom discussion and key takeaways

Questions directors should ask, and the CFO should be ready to answer. A director need not understand a data warehouse to ask these questions, and a CFO who cannot answer them is not yet ready to ask for investment in analytics.

1. Which numbers in the board pack and the lender package come from certified reports, and who owns the definition of each? When did a definition last change, and who approved the change?
2. How do the reporting figures reconcile to the general ledger after each data load, and what happens when they do not?
3. What is the customer duplicate rate, the share of vendors with complete records, and the trend in our data quality scorecard? Which dimension is weakest and who owns the fix?
4. How wide is the interval around the estimates that drive this decision, and is the decision sound at both ends of it?
5. When a model or an alert system is proposed, what is its out of sample error, how does it compare with a naive forecast, and what share of its flags are genuine?
6. What has changed in the data feeding our key models since they were built, and how do we monitor for drift?
7. Where does self service reporting stop, and what prevents an unreviewed analysis from reaching this board or our lenders?
8. Which recurring decisions have named decision rights and a decision log, and what did we learn from comparing the results with the expectations?
9. How does our privacy and retention practice work, and when did counsel last review it?

Key takeaways. The ten lessons that follow condense the module.

1. A single source of truth is an agreed definition with a named owner and one approved place of publication. It is not a single database.
2. Define the number before building the chart. A dashboard that uses a word in two senses will mislead without anyone intending it.
3. Reconcile the reporting store to the ledger after every load. This inexpensive control prevents many errors that reach a board.
4. Report the median beside the mean for skewed data, and treat concentration as a risk.
5. An estimate without an interval is incomplete. Precision costs money, and quadrupling a sample only halves the interval.
6. When an event is rare, most alerts will be false. Measure precision and recall and size the review team for both.

7. Correlation is a prompt to ask what else changed. A regression that omits a driver can reverse a decision.
8. Backtest every model on data it has not seen, compare it with a naive benchmark, and watch for leakage and drift.
9. Score data quality across defined dimensions, tie every defect to an owner and a date, and re-measure each quarter.
10. Connect data to decisions. Record who decides, what was expected, and what happened, and let the difference improve both the data and the judgment.

A closing thought for the section. A careful investor reads a set of accounts with three questions in mind: what is being measured, how was it measured, and what has been left out. The same questions, asked of a dashboard, protect a company from the quiet confidence that attractive charts can create. The CFO cannot read every number in the enterprise, and does not need to. The task is to build an environment in which the numbers that matter have owners, definitions, and evidence, and to be the person in the room who asks the plain question before the plain mistake is made.

6. The executive artifact: a data governance charter and a dashboard challenge checklist

The artifact has two parts. The first is a data governance charter, a short document that the CFO, the chief information officer, and the executive sponsor sign, and that the audit committee can read in five minutes. The second is a checklist that applies the ten questions to one dashboard that the board actually uses. Exhibit 19 is a filled in charter for Northgate Supply, Inc. (fictional), and Exhibit 20 applies the checklist to the Cedar Vale churn dashboard before and after correction. The reader should complete both for the company's own data and a dashboard that leaders use.

Exhibit 19 Data Governance Charter: Completed Example

Charter element	Entry for Northgate Supply, Inc. (fictional)
Purpose	To ensure that the numbers used to run, finance, and report on the business are defined, owned, accurate, secure, and traceable
Scope	Customer, vendor, product, employee, legal entity, chart of accounts, and revenue data across all ERP instances, the warehouse, and certified reports
Executive sponsor	Chief Financial Officer, with a quarterly report to the audit committee
Data council	CFO (chair), CIO, Chief Commercial Officer, Controller, head of Internal Audit (observer), privacy lead. Meets monthly. Approves definitions, targets, and exceptions
Data owners	Customer: Chief Commercial Officer. Vendor and chart of accounts: Controller. Product: Chief Operating Officer. Employee: head of people. Each owner approves definitions and quality targets for the domain
Stewards and custodians	One named steward per domain maintains standards and works the issue log. The CIO's team operates systems and applies access and retention rules
Definitions	Every measure in a board or lender report has a glossary entry with formula, owner, effective date, and change history. The semantic layer holds the approved calculation
Quality standards	Scorecard by domain on six dimensions (Exhibit 15) with targets set by use. Any critical domain below target for two quarters is reported to the audit committee

Charter element	Entry for Northgate Supply, Inc. (fictional)
Lineage and change control	Each certified measure is traceable to the ledger or source. Changes to certified logic require owner approval, testing, and a log entry. A restatement is disclosed to prior recipients
Access and security	Least privilege roles. Quarterly access review by owners. Logging of access to sensitive data (Module 10)
Retention and privacy	Retention schedule by record type approved with legal counsel. Personal data inventory kept current. Disposal documented. Counsel reviews annually
Reporting tiers	Certified, departmental, and personal tiers (Exhibit 3). Board, lender, auditor, and regulator reports come only from the certified tier
Model and analytics standards	Models require a backtest on held out data, a naive benchmark, documented inputs, a drift monitor, and an owner (Module 1 and Module 10)
Measures and review	Quality score, reconciliation breaks, open issues by age, share of board measures certified. Charter reviewed each year

Source: fictional example prepared for teaching. The reader should replace every entry with the company's own owners, scope, and standards, and should have counsel review the retention and privacy provisions.

How to complete the charter. Start with the owners, since a charter without named individuals is a wish. Ask each owner to write the definition of the two or three measures in the domain that matter most, and to say what would make the data fit for its use. Set the first quality targets from a measured baseline, so that they are achievable, and tighten them as the data improves. Keep the document short enough that directors will read it, and attach the detail as schedules. Then run the council for a quarter before judging it, because the first meetings will mostly resolve old disagreements about definitions, and that work is the real value.

Exhibit 20 Dashboard Challenge Checklist: Completed Example for the Cedar Vale Churn Dashboard (Fictional)

No.	Question	Dashboard as received	Dashboard after correction
1	Decision and owner	Concern. Described as visibility for sales	Pass. Lender covenant report and retention spending. Owner: Controller
2	Definition of churn	Fail. Formal notice only, not written down	Pass. Glossary entry: notice or 90 days without an order
3	Source and reconciliation	Concern. Customer list from one ERP instance	Pass. Three instances matched and reconciled after each load
4	Currency of data	Pass. Refresh time displayed	Pass. Refresh time displayed and stale data flagged
5	What is excluded	Fail. Lapsed customers silently active	Pass. Exclusions listed on the page
6	Fair comparison	Fail. Prior quarter used a different definition	Pass. Prior quarters restated on one definition
7	Variation	Concern. No range shown	Pass. Four quarter range displayed
8	Cause and correlation	Fail. Retention program credited without test	Concern. Program effect to be tested against a comparison group
9	Change control	Fail. Logic changed by one analyst, no log	Pass. Owner approval and change log
10	Action threshold	Concern. None stated	Pass. Churn above 9 percent triggers account review

Source: fictional example. Ratings are Pass, Concern, or Fail, assigned by the reviewer from evidence. The reader should apply the same ten questions to a real dashboard and record the evidence for each rating.

How to use the checklist. Choose one dashboard that the board or the lenders see, and ask its builder to answer the ten questions in writing, with evidence. Rate each answer, and treat any Fail on questions 2, 3, or 5 as a stop, since a number whose definition, source, or exclusions are unknown should not be reported externally. Repeat the review when the logic changes and at least once a year. The point of the exercise is not to embarrass the builder. It is to make asking the questions a habit of the organization, so that the discovery of a definition problem happens in a working session and not in a lender call.

7. Research and Perspectives from the Leading Firms

The publications below were found through search, and each page was opened before it was described. Where a page returned only part of its content, or where a statistic was read from a news report of a survey and not from the publisher, the description says so. Several are surveys by firms that sell related services, and a reader should weigh them as evidence of what finance leaders report and not as independent measurement.

Exhibit 21 Publications of the Leading Firms and Standards Bodies

Source	Title and URL	What it says
KPMG	AI adoption in finance doubles, but assurance readiness determines who wins, KPMG finds, press release, May 11, 2026. https://kpmg.com/xx/en/media/press-releases/2026/05/ai-adoption-in-finance-doubles-but-assurance-readiness-determines-who-wins.html	Survey of 1,013 senior finance leaders in 20 countries, fielded March 2026. 36 percent cite data quality as both biggest barrier and opportunity. 42 percent are fully assurance ready for AI enabled finance. Assurance ready organizations report 33 percent error reduction against 6 percent for others
PwC Ireland	The CFO's new mandate, August 11, 2026. https://www.pwc.ie/services/consulting/insights/cfo-trust-data-ai.html	Argues that the CFO is now expected to build trust in the data, governance, and insights behind decisions, that data is a strategic asset, and that clear ownership and a strong data foundation let organizations move faster. Contains no statistics
BDO USA	BDO Research: Audit Technology Achieves Maturity, but More Guardrails Needed, news release, November 17, 2025 (date taken from the release identifier in the web address). https://www.businesswire.com/news/home/20251117316336/en/BDO-Research-Audit-Technology-Achieves-Maturity-but-More-Guardrails-Needed	2025 Audit Innovation Survey of 210 senior finance leaders at U.S. companies with revenue of 250 million to 10 billion dollars, fielded in June 2025. Finance leaders rating their data governance as mature fell from 55 percent in 2024 to 46 percent in 2025. 43 percent have a formal AI governance framework
Workiva (reporting software vendor)	Workiva Executive Benchmark Survey Finds Instability is Accelerating Data Automation and Governance in 2026, press release, February 3, 2026, read through CFO Dive. https://www.cfodive.com/press-release/20260203-workiva-executive-benchmark-survey-finds-instability-is-accelerating-data-a-1/	Survey of 1,497 professionals. 79 percent prioritize data automation and governance. 96 percent say the CFO, CIO, and chief sustainability officer must unite around a shared data governance strategy. Sponsor is a vendor of related software

Source	Title and URL	What it says
Deloitte	Delivering insightful management reporting, March 23, 2020 (the page metadata uses the title management information). https://www.deloitte.com/an/en/services/financial-advisory/perspectives/gx-delivering-insightful-management-information.html	Based on a survey of over 600 senior finance professionals. Says weak data governance undermines finance outputs and causes rework and manual manipulation, and that poor data quality, ineffective technology, and talent shortages are closely linked. Older than the others and cited for its enduring point
Deloitte Belgium	Global Finance Trends 2026: Navigating the expanded scope of Finance, October 8, 2025. https://www.deloitte.com/be/en/services/consulting/perspectives/finance-trends-2026.html	Reports that 63 percent of finance departments have fully deployed AI solutions but significantly fewer have achieved measurable return, and that 64 percent of finance leaders plan to add AI, data, and scenario modelling skills in 2025 to 2026
NIST	NIST AI Risk Management Framework (AI RMF 1.0, NIST.AI.100-1), released January 26, 2023, landing page. https://www.nist.gov/itl/ai-risk-management-framework	Voluntary framework for managing risks of AI systems. The landing page, which was all that was read, states that the framework is being revised as part of the White House AI Action Plan and lists the Generative AI Profile of July 26, 2024. It does not discuss data quality

Each page was opened and the descriptions follow what was returned. No EY, Grant Thornton, Baker Tilly, Crowe, Forvis Mazars, CBIZ, Moss Adams, CohnReznick, CLA, RSM US, SEC, PCAOB, COSO, FASB, or OECD publication was read for this module, and none is cited. The full NIST framework text was not read.

What the research agrees on. The sources differ in kind, from a global survey by a Big 4 firm to a vendor sponsored benchmark and a short perspective with no figures, and they should not be read as a single body of evidence. They agree on three points. Data quality is the constraint that finance leaders name most readily when asked about AI. Ownership and governance, and not technology alone, determine whether data can be trusted, and confidence in governance is not rising as fast as adoption, as the fall in the BDO measure from 55 to 46 percent suggests, though the firm itself cautions that this may reflect greater awareness of the standard required. And leaders expect the CFO to be a central figure in the answer. The KPMG finding that assurance ready organizations reported far greater error reduction is an association in a survey of self reported results, and it does not prove that assurance causes the improvement, which is itself an example of the correlation point made in this module.

8. The ten question self assessment

Rate each statement from 1 to 5 on **demonstrated capability**, using the five point scale of the program: 1 Unprepared (limited understanding, no meaningful application), 2 Developing (conceptual familiarity, relies heavily on others), 3 Competent (can independently apply the capability in conventional situations), 4 Advanced (integrates the capability into strategy and delivers measurable results), and 5 Future-ready (demonstrates enterprise wide impact, anticipates change, and helps others develop). Mark one box in each row. Ratings must rest on actual experience or documented evidence, such as a glossary you published, a data quality scorecard you ran, or a model you backtested, and not on confidence.

Exhibit 22 Module 3 Self Assessment: Rate Each Statement from 1 to 5

No.	Statement	1	2	3	4	5
1	I can describe my company's data architecture, from sources to semantic layer, and explain what our single source of truth is and is not.					
2	I know how our ERP systems are integrated and I require a reconciliation of reporting data to the ledger after each load.					
3	Every measure in our board and lender reports has a written definition, a named owner, and a certified source.					
4	I can read a distribution, a sample estimate with its confidence interval, and a base rate, and I explain them correctly to others.					
5	I can interpret a regression, distinguish correlation from cause, and name the omitted variable that could reverse a conclusion.					
6	I require a backtest, a naive benchmark, and a drift monitor before relying on any predictive model.					
7	I measure data quality on defined dimensions with a scorecard, and each defect has an owner and a date.					
8	I have established data governance with owners, stewards, lineage, access review, retention, and privacy practice, reviewed with counsel.					
9	I apply a structured challenge, such as the ten questions, to dashboards before they reach the board or lenders.					
10	Our recurring decisions have named decision rights, a decision log, and a review of results against expectations that improves the next decision.					
	Total score (maximum 50)					

Rating scale: 1 Unprepared, 2 Developing, 3 Competent, 4 Advanced, 5 Future-ready. Mark one box in each row and add the ten ratings.

Interpreting the total. The bands below are proposed program standards. They are an educational framework and have not been empirically validated against a population of finance executives.

Exhibit 23 Interpretation Bands (Proposed Program Standards)

Score out of 50	Band	What it suggests
10 to 19	Unprepared	Limited command of data foundations. Begin with definitions and ownership for three board measures
20 to 29	Developing	Conceptual familiarity. Practice the dashboard challenge and a first data quality scorecard
30 to 39	Competent	Can run conventional analysis and governance. Strengthen backtesting, drift monitoring, and decision logs
40 to 45	Advanced	Delivers measurable gains in data trust and decisions. Extend to all domains and to assurance of AI use

Score out of 50	Band	What it suggests
46 to 50	Future-ready	Enterprise wide impact and the ability to develop others. Share the charter and the method, and mentor peers

Source: proposed program standards for the Future-Ready CFO Maturity Index, not empirically validated. Module 3 is not designated a critical module.

Development actions by low score pattern. If statements 1, 2, and 3 score low, the weakness is in architecture and definitions. Action one is to draw the path of one board measure from the dashboard back to the ledger and list each system and transformation on the way. Action two is to run a reconciliation of the warehouse to the ledger for the last three loads. Action three is to write a glossary entry, with an owner, for the five measures that the board sees most often. If statements 4, 5, and 6 score low, the weakness is statistical reasoning and model discipline. Action one is to recompute one recent estimate with its confidence interval. Action two is to rerun a recent analysis with the most likely omitted driver added. Action three is to backtest the main forecast model on the latest six months and compare it with a naive forecast.

If statements 7 and 8 score low, the weakness is data quality and governance. Action one is to measure completeness, uniqueness, and consistency on the customer or vendor master. Action two is to name an owner and a steward for each of the five core domains. Action three is to review the retention and privacy practice with counsel and record what is missing. If statements 9 and 10 score low, the weakness is decision intelligence. Action one is to apply the ten questions to one dashboard and write the answers down. Action two is to start a decision log for the next three significant decisions. Action three is to review a decision from a year ago against what was expected and record what the comparison teaches.

9. Currency check and sources

Currency check: the publications in Exhibit 21 were opened and read through the web tools described in Section 7, and they are dated between 2020 and 2026. Survey statistics change from year to year, and standards and guidance for AI and data governance are being revised, including the NIST framework, so the reader should confirm current editions before relying on any of them. The following statements come from sources that were read: the KPMG survey findings of May 11, 2026 on adoption, data quality, and assurance readiness; the PwC Ireland perspective of August 11, 2026; the BDO survey findings of November 2025 on data governance maturity and AI governance; the Workiva survey of February 3, 2026, read as republished by CFO Dive; the Deloitte management reporting article of March 2020 and the Deloitte Belgium finance trends of October 2025; and the NIST landing page description. The following come from the author's own knowledge and were not confirmed in a source read: the definitions of warehouse, lake, lakehouse, semantic layer, master data, owner, steward, custodian, and lineage; the integration patterns and the reconciliation control; the self service tiers; the statistical definitions and the guidance on sampling, base rates, regression, and backtesting; the common 0.10 and 0.25 conventions for the population stability index; the data quality dimensions and scorecard design; and the ten question challenge and charter structure, which are teaching tools and not published standards. The NIST framework itself, and publications of COSO, the SEC, and the PCAOB were not opened. All companies, people, and figures in the exhibits, including Northgate Supply, Inc. and Cedar Vale Distribution, Inc., are fictional, and every calculation was computed in Python. This module is educational. It is not accounting, tax, legal, privacy, securities, or professional advice, and the reader should consult qualified professionals in accounting, tax, legal, and other relevant services, and the external auditor, before acting.

Sources: KPMG, AI adoption in finance doubles, but assurance readiness determines who wins (May 11, 2026); PwC Ireland, The CFO's new mandate (August 11, 2026); BDO USA, Audit Technology Achieves Maturity, but More Guardrails Needed (November 17, 2025); Workiva Executive Benchmark Survey press release (February 3, 2026); Deloitte, Delivering insightful management reporting (March 23, 2020); Deloitte Belgium, Global Finance Trends 2026 (October 8, 2025); NIST, AI Risk Management Framework landing page. URLs are in Exhibit 21.