As part of our ongoing efforts to optimize Risk (SVAR) under Basel 2.5, there has been a focus on several high-impact areas such as enhancing data sources in SPG and increasing time series granularity in Credit. Alongside these, onboard CDX options for hedging, one recurring theme is **Data Quality (DQ)**.

To move forward,please review the angles below (to add clarity and enable **senior management** to make informed decisions) and get back to us on where you'd be willing to support or invest based on technical feasibility, on **operating model appetite, prioritization, and resource commitment**

As we look to clarify the next steps for our data quality checks, I want to outline two potential paths forward.

**Path 1: Leverage Existing Methods**

* **Action:** Use our current statistical functions without adding new benchmarking or advanced analytics.
* **Pros:** No additional hiring or resources needed.
* **Cons:** Requires manual remediation by SMEs, which can be time-consuming and may lead to recurring issues.

**Path 2: Enhance with Advanced Analytics**

* **Action:** Implement machine learning and advanced analytics to improve data quality checks.
* **Pros:** Provides more robust and automated solutions, reducing manual efforts.
* **Cons:** Requires hiring for research, development, and productionization, and involves building new workflows and governance processes

Please review these options and let us know your thoughts on prioritization, critical path considerations, and resource allocation. Your input will help us shape the roadmap and ensure we make the best decision for our team's goals.

Great — you’ve outlined a comprehensive, multi-angle framework for evaluating market data quality and vendor time series. Here's how I would **structure this into a professional checklist** or **validation framework** for internal use or for review with Risk/Quants/Front Office/MRO:

**✅ Comprehensive Market Data Quality & Vendor Evaluation Framework**

**1. Qualitative Assessment of Vendor Methodology**

✔️ Used to assess model robustness, construction logic, and alignment with use-case

| **Criteria** | **Key Questions** |
| --- | --- |
| **Bond Selection Criteria** | Are rules transparent and consistent across time and rating buckets? |
| **Quote Sources** | Are prices based on evaluated quotes, dealer quotes, composite prices, or trades? |
| **Curve Construction** | Is the methodology factor-based (e.g., PCA/index modeling like S&P) or curve fitting (e.g., BBG cohort)? |
| **Coverage** | Does the vendor cover sectors, regions, and ratings that match current and future exposure? |
| **Documentation** | Do we have white papers explaining methodology and rationale? |

**2. Market Event Testing**

✔️ To verify that the time series reacts to real-world stressors and behaves intuitively

| **Event Type** | **What to Check** |
| --- | --- |
| **Crisis Dates (COVID, Tariffs, 2008)** | Does the data reflect meaningful and realistic moves? |
| **Vendor Benchmarking** | Compare to known standards (e.g., IG OAS / HY OAS) across vendors |
| **Root Cause Analysis** | Are deviations explainable by methodology, stale pricing, or gaps? |
| **SME Judgment** | Get feedback from desks, MROs, or quants on plausibility |

**3. Data Quality Diagnostics**

✔️ Core statistical and rule-based validation

| **Test** | **What to Check** |
| --- | --- |
| **Missing Data** | Are there gaps? What’s the frequency of non-reporting days? |
| **Outlier Detection** | Use z-score to identify and flag spurious moves or spikes |
| **Monotonicity of Ratings** | Are yields ordered correctly by rating (e.g., AAA < BBB)? |

**4. Benchmarking & Cross-Checks**

✔️ Ensure consistency and reasonability of data across time and sources

| **Test** | **Description** |
| --- | --- |
| **Vendor-to-Vendor** | Cross-check for same sector/rating across vendors over time |
| **Trader Mark Comparison ("PLA-style")** | Compare time series-based PnL with PnL derived from trader marks |
| **Sector Consistency** | Are intra-sector moves in line with economic intuition? |
| **Time Consistency** | Are ranking and levels stable across time, barring shocks? |

**5. Statistical Behavior & Correlation Testing**

✔️ Deep-dive into how the data behaves across sectors and ratings

| **Test** | **What to Look For** |
| --- | --- |
| **Volatility (StDev, 99th pct)** | Identify high-volatility names or vendors |
| **1D/10D Correlation within Ratings** | Are different sectors in same rating too correlated or too independent? |
| **Cross-Rating Correlation** | Is the downgrade/uprating behavior logical? |

**6. Quantitative Impact Analysis**

✔️ Understand how data affects risk and capital

| **Test** | **Purpose** |
| --- | --- |
| **VaR/SVaR Snapshots** | Use different vendor time series and run risk on the same day |
| **Hypothetical Portfolio Testing** | Simulate portfolios (e.g., AA bonds across sectors) to see vendor impact |
| **Backtest VaR PnL Explain** | Compare PnL explain using vendor series vs internal/trader marks |

**7. Data Required to Complete These Tests**

✔️ What’s needed from vendors or internal sources

| **Required Dataset** | **Notes** |
| --- | --- |
| **Vendor White Papers** | Explains methodology (construction, interpolation, sources) |
| **Full Time Series History** | Across: |

* All sectors for a given rating
* All ratings for a given sector
* All regions for a given rating/sector |  
  | **Clean PnL and Risk Factor Mapping** | For PLA-style benchmarking |  
  | **Trader marks** | For PnL explain cross-validation |

Let me know if you'd like this turned into:

* ✅ A **formal checklist spreadsheet**
* ✅ A **PowerPoint framework** for stakeholder alignment
* ✅ A **Python/Pandas template** for implementing some of the statistical checks

This can also serve as the **foundation for onboarding any new vendor** (e.g., Bloomberg, S&P, ICE, JPM) into your RFDM or MDSOR platforms.

Exactly — you're spot on.

What you've outlined is a **pragmatic, risk-aware alternative** to going down the full **machine learning (ML)** and **over-engineered analytics** path. It emphasizes **reliability, transparency, and actionability** over complexity.

Here’s how you can **frame this route** to senior management or tech leadership as a **valid and efficient strategy**:

**✅ “Smart Simplicity” Over Complexity – A Risk-Aligned DQ Strategy**

**🎯 Why This Approach Makes Sense**

| **Concern** | **How This Simpler Route Solves It** |
| --- | --- |
| ⚠️ ML may be overkill | Simple z-scores, drift, and logical tests already catch 80–90% of issues |
| ⚠️ Black-box risk | These rules are transparent, auditable, and explainable |
| ⚠️ Resource constraints | Avoids hiring data scientists or building MLOps pipelines |
| ⚠️ Time-to-value | Can be implemented incrementally and productionized quickly |
| ⚠️ Governance complexity | Avoids new model governance, calibration policies, or signoffs |

**🔁 This Route Still Delivers Value**

Even without ML, you can have:

* ✅ **Weekly automated suspecting with clear thresholds**
* ✅ **Monitoring of curve and rating logic**
* ✅ **Alerting on vendor drift or major data moves**
* ✅ **Integration into workflow: override, approve, escalate**

These can be designed to **scale and evolve**, without betting on a PoC that may not be accepted or resourced.

**⚖️ When This Approach is Enough**

This route is **more than sufficient when**:

* The number of **critical risk factors** is known and not expanding rapidly
* SME oversight is strong, and the **cost of false negatives** is manageable
* You have **defined vendor relationships** with periodic benchmarking
* The focus is on **maintaining data quality** rather than deriving insights from data

**🚧 Where It May Need to Evolve Later**

You may revisit advanced analytics **only if:**

* You're managing **tens of thousands** of series with shrinking teams
* There's a **pattern of recurring false positives or missed signals**
* Regulatory pressure requires **auditable justification of every override**
* You shift from rules-based to **fully automated proxy selection or risk factor detection**

**🧩 Framing It to Senior Management**

“Rather than prematurely investing in complex AI/ML-based solutions, we propose a high-impact, low-friction path forward: strengthen weekly calibration and ongoing validation using a rules-based, explainable control framework. It’s scalable, auditable, and keeps us focused on fixing what matters — not building a model for the sake of it.”

Let me know if you’d like this in:

* 📊 A visual **1-slide justification**
* 📄 A one-page **strategy note to share with tech/data leadership**
* ✅ A **decision tree** to show when this path is enough vs when ML is justified

Great follow-up. Let’s walk through each stage—**pre-onboarding, post-purchase, and ongoing monitoring**—and address your questions with practical guidance, along with ideas for what may need to be **productionized** long-term.

**✅ 1. Pre-Onboarding: Data Selection Phase**

These tests are applied **before purchasing or onboarding** new data. The goal is to validate **fit-for-purpose** and **coverage**.

**Must-Have Pre-Onboarding Tests:**

* ✅ **Qualitative checks** (methodology, quote source, rating hierarchy)
* ✅ **Historical backfill checks** (coverage, depth, and format)
* ✅ **Market event reactivity** (COVID, Tariffs, etc.)
* ✅ **Benchmarking** (vs vendor peers, industry indices)
* ✅ **PnL Explain sanity checks** (if sample marks are available)

👉 *Yes, this is where the full checklist you already drafted is critical.*

**✅ 2. Post-Purchase: One-Time Statistical & Benchmarking Tests**

Once you've purchased the data, the goal is to **validate delivery quality** and **build trust** before feeding into VaR, pricing, or PLA.

**Recommended Post-Purchase Tests:**

* ✅ **Data quality scan** (missing points, spikes, stale fields)
* ✅ **Z-score or MAD outlier detection**
* ✅ **Curve shape logic** (monotonicity, spread level consistency)
* ✅ **Sector/rating correlation tests**
* ✅ **Backtest vs FO marks** for key names (e.g., top 10 by notional)

🔁 These can be **run once or quarterly**, unless new mappings are introduced.

**🔄 3. Ongoing Monitoring / Weekly Calibration**

This is where things evolve into BAU, and your question is key: are **statistical and benchmarking checks enough**?

**What Works Well:**

* ✅ **Weekly re-checks on Z-score outliers**
* ✅ **Weekly volatility band check (stdev and 99th pct)**
* ✅ **Drift monitoring** (e.g., 5-day change outside historical bands)
* ✅ **Auto-flags on inversion (e.g., BB yield > B)**

**Additional Angles to Consider:**

* ⚠️ **Sector/rating hierarchy consistency drift** (e.g., AA no longer consistently tighter than BBB)
* ⚠️ **Proxy divergence**: if you still use proxies, do they remain correlated to true series?
* ⚠️ **Survivorship bias**: new data from vendor may drop inactive tickers, causing silent loss of coverage

**🏗 4. What Should Be Productionized**

Not everything needs to be live in prod — but some definitely should. Here’s what to consider:

| **Function / Check** | **Should Be Productionized?** | **Notes** |
| --- | --- | --- |
| Weekly DQ scan (missing, stale, outliers) | ✅ Yes | Automatable in RFDM/MDSOR |
| Benchmark drift (vs last week or vendor) | ✅ Yes | Simple logic, valuable early flag |
| Sector correlation breakdown | ❌ Not initially | Run manually when needed |
| Inversion in yield curve or ratings | ✅ Yes | Triggers suspecting logic |
| FO/Risk alignment monitoring (PnL explain) | 🔶 Optional | Useful, but may require trader data |
| Rating/slope inconsistency checks | ✅ Yes | Important for curve-driven desks |
| Proxy divergence check | ✅ If proxies are used | Monitor correlation breakdown |

**🧠 Bonus: Governance / Workflow Considerations**

* 🗂 **Dashboards** (e.g., Tableau or internal UI) for viewing suspect data
* 👥 **Approval Workflow** (Desk / Risk / MROs sign-off on override or exclusions)
* 📬 **Alerting System** (email or platform notification when thresholds are breached)

**🚦 Summary**

| **Phase** | **Are Statistical & Benchmarking Tests Enough?** | **What Else to Add** |
| --- | --- | --- |
| **Pre-onboarding** | ✅ Yes | SME review on desk-fit |
| **Post-purchase** | ✅ Mostly | PnL explain, sector correlation |
| **Ongoing calibration** | ✅ For basic monitoring | Add weekly logic, productionize alerts/workflow for key risks |

Would you like this structured into:

* A production implementation checklist?
* Or a slide summarizing onboarding → calibration → BAU lifecycle?