

## **Quantitative Benchmarking of ERP Analytics Architectures: Evaluating Cloud vs On-Premises ERP Using Cost-Performance Metrics**

---

**B. M. Taslimul Haque<sup>1</sup>; Md. Arifur Rahman<sup>2</sup>;**

---

[1]. Bachelor of Science in Computer Science & Engineering, American International University Bangladesh, Dhaka, Bangladesh; Email: [bmtaslim121@gmail.com](mailto:bmtaslim121@gmail.com)

[2]. Bachelor of Science (B.Sc.) in Computer Science & Engineering, Bangladesh University of Business & Technology, Bangladesh; Email: [rahman.arifur22226@gmail.com](mailto:rahman.arifur22226@gmail.com)

**Doi:** [10.63125/y05j6m03](https://doi.org/10.63125/y05j6m03)

**Received:** 21 September 2020; **Revised:** 27 October 2020; **Accepted:** 29 November 2020; **Published:** 26 December 2020

---

### **Abstract**

This study addresses the problem that organizations choose between cloud and on-premises ERP analytics without workload-normalized evidence, materially increasing cost and performance risk. The purpose was to benchmark ERP analytics architectures and test how architecture type relates to analytics effectiveness. Using a quantitative, cross-sectional, case-based design, objective cost and performance indicators were extracted over a fixed four-week window and paired with a 5-point Likert survey from cloud and on-prem cases ( $n = 152$  valid responses; cloud-exposed  $n = 79$ , on-prem-exposed  $n = 73$ ). Key variables included architecture type (cloud = 1), total monthly analytics-related cost and cost per active user, latency (median and 95th percentile), throughput under 50-user concurrency, availability, incident rate and recovery time, and perceptual constructs: System Quality (SQ), Information Quality (IQ), Service Quality (ServQ), User Satisfaction (US), and Analytics Effectiveness (AE). The analysis plan used descriptive statistics, reliability testing (Cronbach's alpha: SQ = .89, IQ = .86, ServQ = .84, US = .88, AE = .90), Pearson correlations, and multiple regression with controls (usage frequency, experience, role). Results show that cloud delivered lower cost (\$48,200 vs \$61,750 per month; \$214 vs \$289 per user) and stronger performance (median latency 2.3 s vs 3.7 s; 95th percentile 6.9 s vs 10.8 s; throughput 1,420 vs 1,050 queries/hour; availability 99.91% vs 99.62%; incidents 3 vs 6 per month; mean time to recovery 38 vs 64 minutes). Perceptions aligned, with higher AE for cloud ( $M = 4.14$ ,  $SD = 0.60$ ) than on-prem ( $M = 3.49$ ,  $SD = 0.73$ ). AE correlated most strongly with US ( $r = .71$ ) and measured performance ( $r = .55$ ), and regression explained substantial variance (Adjusted  $R^2 = .49$ ), with performance ( $\beta = .34$ ,  $p < .001$ ), architecture ( $\beta = .21$ ,  $p = .002$ ), and cost efficiency ( $\beta = .17$ ,  $p = .009$ ) as significant predictors. Implications are that ERP analytics selection should use repeatable workload definitions, percentile-based SLAs, and TCO accounting that includes support and downtime, because performance stability and cost efficiency jointly drive decision value.

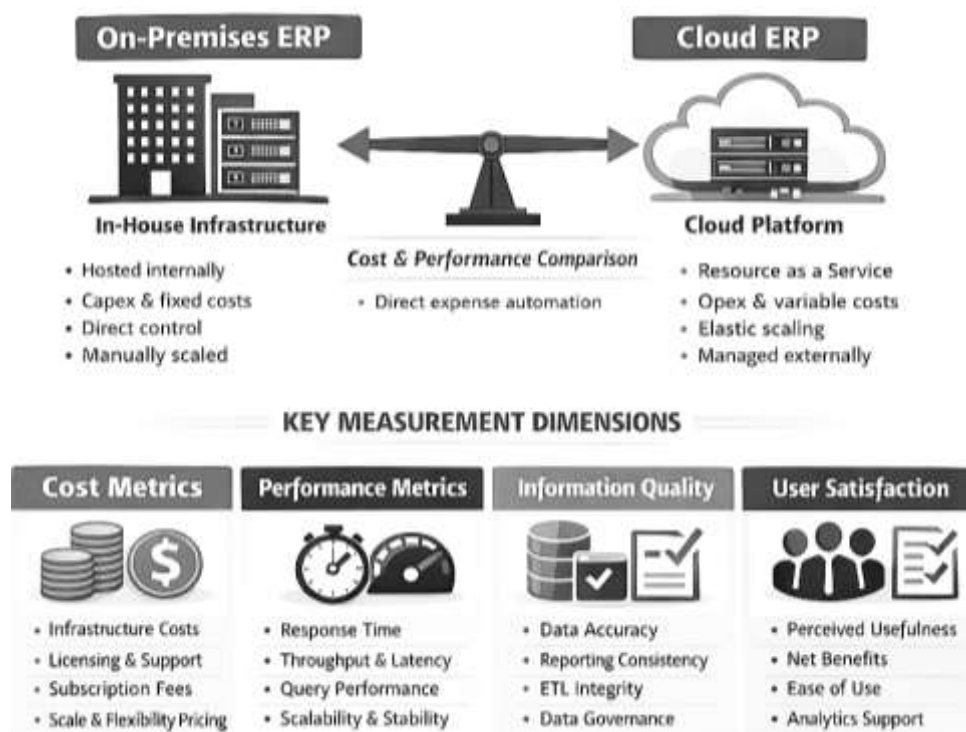
### **Keywords**

Cloud ERP analytics, On-premises ERP, Cost-performance benchmarking, Workload normalization, Analytics effectiveness;

## INTRODUCTION

Enterprise Resource Planning (ERP) systems are commonly defined as integrated, modular enterprise information systems that standardize and automate core organizational processes by using a shared database and coordinated workflows across functions such as finance, procurement, production, sales, and human resources. In contemporary organizations, ERP increasingly operates as an “analytics architecture,” meaning the structured stack of data capture, storage, modeling, querying, and visualization capabilities that converts operational transactions into decision-ready information products across managerial levels. Within this framing, “ERP analytics” refers to the systematic generation of descriptive, diagnostic, and inferential insights from ERP-originating data, using measurement models and statistical techniques that can be evaluated for reliability, validity, and decision utility (Amid et al., 2012). The term “architecture” in ERP analytics also implies definable layers—data sources, integration mechanisms, governance controls, computational resources, and user-facing services—that determine performance, scalability, and cost behaviors under real workloads (Buyya et al., 2009).

**Figure 1: Benchmarking Cloud vs. On-Premises ERP Analytics Architectures**



A key architectural distinction in practice is deployment model: on-premises ERP, where infrastructure and platforms are hosted and managed within an organization’s facilities or dedicated environments, versus cloud ERP, where compute, storage, and application services are provisioned through cloud delivery models and contractually governed service interfaces. Cloud computing literature characterizes this delivery shift as a utility-oriented model that pools resources and supports elastic provisioning while reframing cost structures toward operational expenditure and service metering (Benlian et al., 2011). In such environments, “benchmarking” becomes essential because analytics outcomes and user experience can vary sharply depending on workload intensity, concurrency, data volume, integration complexity, and governance constraints. Performance benchmarking in cloud contexts has therefore been treated as a measurable comparison of service behavior under controlled tasks, emphasizing repeatable metrics for latency, throughput, and variability. The international significance of benchmarking ERP analytics architectures follows from the global reliance on ERP for financial reporting, supply chain coordination, and public and private service delivery, where cross-border operations and multi-site organizations require consistent analytics quality while operating

under heterogeneous cost and regulatory environments (Benlian et al., 2009).

A research focus on cost-performance metrics for ERP analytics architectures depends on treating ERP not only as an application suite but also as a data-intensive decision system whose value emerges through information quality, system responsiveness, and organizational net benefits. Business intelligence and analytics scholarship formalizes this value pathway by emphasizing that analytic advantage stems from how data is curated, modeled, and delivered into managerial routines, not merely from data availability itself (Prybutok et al., 2008). Because ERP is often the dominant producer of structured enterprise data, ERP analytics architectures can be evaluated using constructs such as information quality, service quality, and system quality that link technical design to user outcomes and organizational benefits (Prybutok, 2012). In cross-functional environments, ERP analytics quality is shaped by master data alignment, transactional integrity, and the degree to which operational process standardization produces comparable data across sites. Empirical ERP research has therefore treated post-implementation conditions – training, governance, external expertise, and process stabilization – as determinants of whether ERP produces reliable and timely information for managerial use (Petter & McLean, 2009). Studies that examine ERP-enabled performance often highlight that operational improvements and informational consistency depend on both the system's technical fit and organizational assimilation, making analytics outputs inseparable from the implemented configuration and usage patterns (Li et al., 2010). Conceptually, cloud versus on-premises deployment changes where analytics computation occurs, how integration is orchestrated, and how performance bottlenecks manifest (Chen et al., 2012). These differences can influence query response times, refresh frequency, concurrency handling, and the cost profile of scaling analytics workloads. In turn, the credibility of cost-performance comparisons depends on ensuring that compared environments handle equivalent workloads and comparable data governance requirements, because analytics performance in ERP settings is often sensitive to schema design, indexing strategies, batch windows, and integration load. This is why ERP analytics benchmarking benefits from workload normalization and measurement transparency, allowing observed performance differences to be interpreted as architecture effects rather than artifacts of uneven task demands (Ram et al., 2013).

This study is designed to achieve a set of clearly defined objectives that translate the broad challenge of choosing between cloud and on-premises ERP analytics architectures into measurable, testable, and comparable outcomes within a quantitative, cross-sectional, case-study setting. The first objective is to establish a structured benchmarking foundation by defining a consistent set of cost and performance indicators that can be computed for both architectures using the same unit standards and time window, ensuring that reported results represent like-for-like comparisons rather than context-driven variations. The second objective is to quantify and compare the baseline cost components associated with each architecture by separating direct system expenditures from indirect operational costs, capturing licensing or subscription fees, infrastructure or hosting expenses, support and administration effort, and the economic impact of service interruptions, and then expressing these costs in standardized forms such as cost per active user, cost per reporting cycle, or cost per analytics workload unit. The third objective is to quantify and compare baseline performance behaviors by measuring indicators such as query response time, report rendering time, refresh frequency, uptime, incident frequency, and recovery time, and then summarizing these measures using descriptive statistics that highlight central tendencies and variability. The fourth objective is to implement a workload-normalization procedure that adjusts cost and performance figures to a common workload basis so that efficiency and responsiveness can be evaluated under comparable demand conditions, strengthening the credibility of the cloud versus on-premises comparison. The fifth objective is to evaluate ERP analytics effectiveness through a structured survey instrument using a five-point Likert scale, capturing user assessments of analytics quality, reliability, accessibility, timeliness, and overall satisfaction with the analytics service experience. The sixth objective is to examine statistical relationships among architecture type, cost efficiency, performance indicators, and perceived analytics effectiveness by applying correlation analysis, thereby identifying patterns of association that may signal which technical and financial factors align most closely with user-valued outcomes. The seventh objective is to test predictive relationships by developing regression models that estimate the degree to which cost-performance metrics explain variation in analytics effectiveness while accounting for relevant

respondent and workload characteristics, enabling hypothesis testing with interpretable coefficients and model fit statistics. The eighth objective is to strengthen the trustworthiness of the benchmarking results through additional validation layers, including trade-off mapping that visually positions architectures across cost and performance dimensions and robustness checks that confirm whether key comparisons remain stable under alternative scoring assumptions. Together, these objectives create an evidence-based structure for comparing cloud and on-premises ERP analytics architectures using transparent measurement, rigorous statistical testing, and decision-relevant outputs.

## **LITERATURE REVIEW**

The literature on ERP analytics architectures provides the foundation for evaluating cloud versus on-premises deployment using measurable cost-performance indicators and statistically testable effectiveness outcomes. ERP systems are widely examined as enterprise-wide integration platforms that standardize process execution and generate high-quality transactional data, while ERP analytics extends this role by transforming operational records into managerial insights through reporting, dashboards, and decision-support routines. Within this body of work, “architecture” is treated as the layered configuration of data sources, integration mechanisms, governance controls, computational infrastructure, and user-facing services that collectively shape how analytics is produced and consumed across the organization. A central theme in prior studies is that the value of ERP and analytics is realized through a socio-technical pathway where system quality, information quality, and service support influence usage patterns, user satisfaction, and organizational benefits, suggesting that architecture evaluation must link technical measurements to stakeholder outcomes rather than relying on isolated performance statistics. Cloud computing scholarship adds a complementary perspective by conceptualizing cloud delivery as an on-demand, service-oriented provisioning model that changes ownership, scalability, risk allocation, and operational responsibilities compared with traditional on-premises environments. Research on SaaS and cloud adoption further emphasizes that deployment decisions are driven by perceived relative advantage, compatibility, organizational readiness, governance requirements, and service reliability, indicating that the cloud versus on-prem distinction affects both direct cost structures and the management processes required to sustain performance and trust. Benchmarking research strengthens this discussion by proposing repeatable measurement practices for cost and performance, including the use of standardized workload profiles, comparable task suites, and normalization approaches that control for differences in demand and usage intensity. In parallel, business intelligence and analytics studies highlight that performance and cost must be evaluated alongside analytics outcomes such as timeliness, accuracy, usability, and decision-making impact, because these outcome variables represent the practical success criteria for ERP analytics users. Taken together, the literature positions cloud versus on-premises ERP analytics as a multi-dimensional evaluation problem where architecture influences observable cost and performance behavior and also shapes perceived effectiveness through quality and service characteristics. This review therefore synthesizes prior research to identify the most defensible metrics, constructs, and explanatory relationships for a quantitative, cross-sectional, case-study-based benchmarking approach that combines objective cost-performance measures with survey-based effectiveness indicators and supports hypothesis testing through descriptive statistics, correlation analysis, and regression modeling.

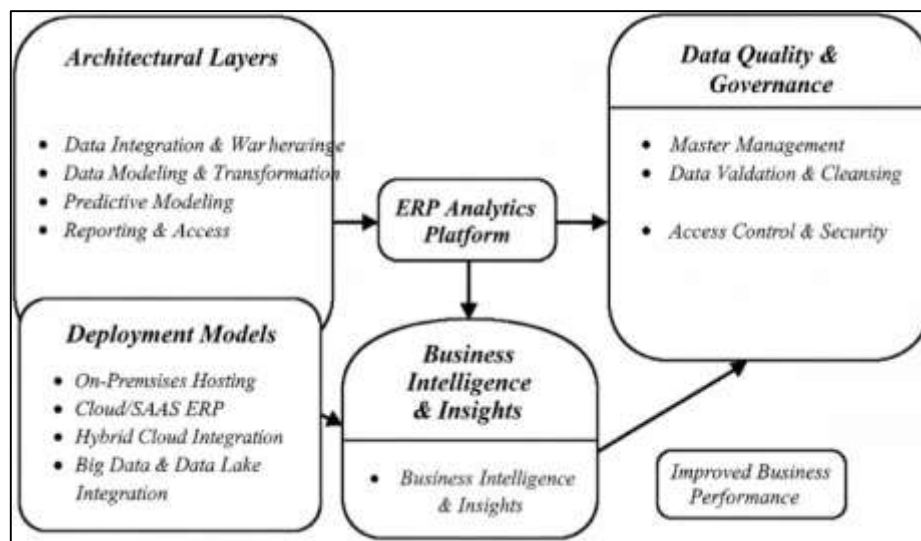
## **ERP Analytics Architectures and Deployment Models**

ERP analytics architectures describe the technical and organizational design through which enterprise transactions are transformed into analytical information products for reporting, monitoring, and decision routines. In this view, an ERP platform is not only a set of functional modules; it is also a layered analytics stack that determines how data are captured, validated, integrated, stored, modeled, and delivered to users. A common architecture starts with transactional sources in finance, procurement, inventory, and HR, followed by extraction and integration services that reconcile master data, keys, and process events into a consistent analytical dataset. Many organizations implement a staging layer to isolate operational processing from heavy queries, then load curated data into an analytical repository such as a data warehouse or columnar database, and finally expose metrics through semantic models, OLAP cubes, or governed data marts to dashboards and self-service tools. When ERP analytics is treated as a first-class architectural concern, design choices such as batch versus



near-real-time integration, centralized versus federated marts, and the governance of metric definitions become measurable determinants of cost and performance outcomes. Work that examines ERP-BI integration frames these design choices as an integration problem, where data consistency, metadata alignment, and process synchronization connect operational ERP records to analytical consumption and reduce reporting fragmentation (Nofal & Yusof, 2013). More recent architecture discussions extend this layered view by arguing that ERP analytics increasingly spans both structured ERP transactions and external, high-velocity streams, so architectures must accommodate heterogeneous storage engines and multiple integration patterns while keeping analytics services aligned with enterprise strategy (Shi & Wang, 2018). For benchmarking studies, these architectural layers provide locations for collecting metrics: ETL runtimes and refresh windows at the integration layer, query latency and throughput at the analytical store, availability and incident frequency at the service layer, and labor effort attached to layer administration.

Figure 2: ERP Analytics Architectures And Deployment Models



A second stream of architecture research focuses on the quality and suitability of ERP data as it moves through analytics pipelines, emphasizing that architectural performance is inseparable from the semantic correctness of the data delivered to decision makers. ERP analytics architectures rely on master data consistency, process-event completeness, and well-governed transformation rules; when these conditions are weak, organizations may observe fast query times while still producing misleading indicators. From this perspective, architectural decisions about where validation occurs, how exceptions are handled, and how business rules are encoded in transformations are core determinants of analytics trust. Glowalla and Sunyaev operationalize this logic by examining ERP system fit through a task-technology lens and linking the perceived fit of ERP outputs to data quality management practices, highlighting how data analysis routines and quality controls become part of the ERP environment rather than an external add-on (Glowalla & Sunyaev, 2014). In ERP analytics, this fit problem extends beyond data accuracy to include whether analytical representations match managerial tasks, such as variance explanations, drill-down auditing, or cross-functional reconciliation. Studies that connect ERP use to user satisfaction further reinforce that architecture cannot be evaluated only as infrastructure; it must be evaluated as an information service that users experience through reliability, responsiveness, and the perceived usefulness of reports. In a large-scale survey of ERP users, determinants such as system quality, training, and organizational support were associated with adoption and satisfaction, implying that architectural choices that stabilize service performance and reduce user effort also influence whether analytics features are actually used (Costa et al., 2016). For benchmarking, this literature motivates pairing objective performance and cost indicators with survey constructs that capture information quality, service quality, and satisfaction, so architectural comparisons remain anchored in observed user outcomes rather than solely in system counters. This

alignment supports defensible cross-architecture inference.

Deployment model is a third major dimension in ERP analytics architecture research because moving analytics workloads from on-premises infrastructure to cloud services changes how capacity is provisioned, how integration is managed, and how costs accumulate over time. On-premises ERP analytics commonly couples databases, ETL tools, and reporting servers to organization-owned hardware and internal networks. This model provides direct control over configuration, data locality, and upgrade timing, and it often enables tight integration with legacy systems through low-latency connections. Cloud deployment, by contrast, can be implemented as SaaS ERP with embedded analytics, as managed database and warehouse services supporting a lift-and-shift of analytics layers, or as hybrid patterns where sensitive data remain local while compute bursts to the cloud for heavy reporting. These options shift architectural responsibilities from internal administrators to vendors and reshape cost from capital investment to subscription and consumption charges, while also introducing platform constraints that affect customization, identity management, and data movement. The literature that examines cloud ERP in case settings emphasizes feasibility and fit conditions, illustrating how organizations evaluate cloud ERP as a service bundle that combines application functions, infrastructure, and support processes, and how implementation decisions are shaped by resource limits and the need for rapid deployment (Zadeh et al., 2018). For ERP analytics benchmarking, deployment differences imply that cost metrics must separate recurring service fees, usage-based compute and storage, integration middleware costs, and internal labor, while performance metrics must account for network paths, multi-tenant resource contention, and vendor-managed maintenance windows. Consequently, architectures are often compared using workload-normalized indicators such as cost per active user, cost per reporting cycle, response time per query class, and availability per month, so that deployment choice can be evaluated as a measurable trade-off rather than a purely technological preference. Such normalization creates a common denominator for cross-site, cross-period comparisons.

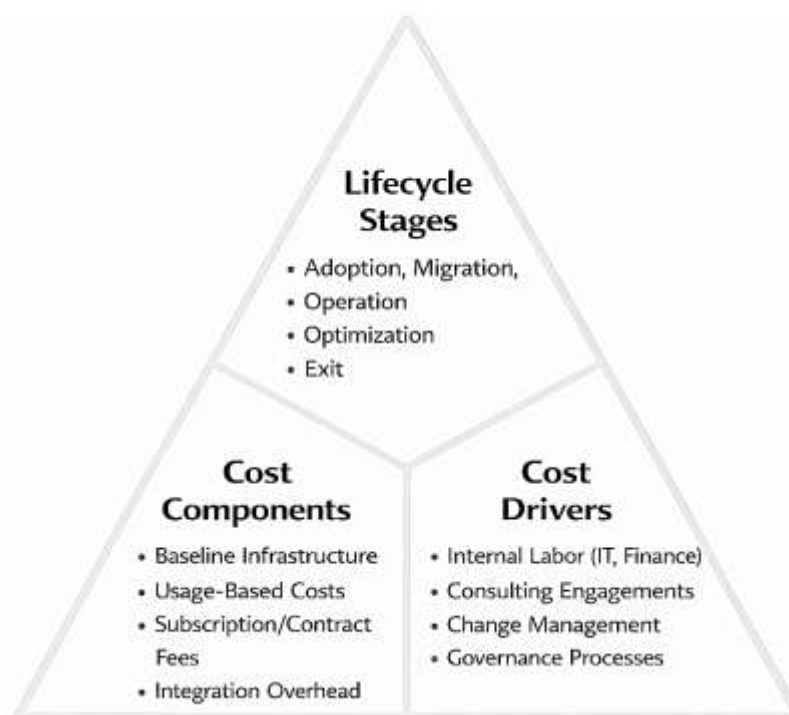
### **Cost Evaluation Models for ERP/Enterprise Analytics**

Cost evaluation in ERP and enterprise analytics is commonly approached through lifecycle accounting that distinguishes acquisition and implementation costs from operating, scaling, and change-related costs over the system's useful life. In ERP analytics architectures, the "cost object" is not only the ERP license or subscription but the full analytics service that includes data movement, storage, compute, governance, and user enablement. This framing leads many organizations to structure cost evaluation around Total Cost of Ownership (TCO), where direct costs (software fees, infrastructure, vendor services) and indirect costs (internal labor, downtime, training, process redesign, and governance overhead) are treated as integral components of ownership. ERP investment evaluation research further stresses that large enterprise initiatives embed uncertainty and managerial constraints, making cost evaluation meaningful only when it captures staged decisions, irreversible commitments, and the value of managerial flexibility across rollout phases. A real-options perspective is therefore used in ERP investment projects to represent the economic consequences of uncertainty and multi-stage implementation, translating uncertainty into explicit valuation logic rather than treating it as an unmeasured risk premium (Kouki et al., 2008). In an ERP analytics comparison between cloud and on-premises architectures, this lifecycle view is essential because the cost profile differs not only in where spending appears (capital versus operating expense) but in when spending occurs (upfront provisioning versus incremental scaling) and which actors incur the costs (internal IT teams versus external vendors). Consequently, cost evaluation models for ERP analytics tend to classify costs by lifecycle stage (adoption/selection, implementation/migration, operation, optimization, and exit), and by responsibility type (vendor-billed versus internally absorbed), so the benchmark can compute comparable units such as cost per active user, cost per reporting cycle, and cost per workload unit. This logic also supports separating baseline platform cost from variability cost, because analytics workloads fluctuate across month-end close cycles, audit windows, and seasonal demand, and these fluctuations can distort naive comparisons that rely on single-point cost snapshots.

A major empirical contribution to ERP cost modeling is the identification of cost drivers that systematically explain why ERP projects deviate from initial budgets and why cost structures vary across organizations. Studies that examine ERP implementation costs in small and medium enterprises

highlight that consulting fees, customization needs, organizational readiness, and internal capability gaps are recurrent drivers of total cost and can outweigh the nominal cost of software. Empirical survey evidence from Swiss SMEs provides a structured way to treat implementation costs as outcomes influenced by enterprise characteristics and project organization, supporting the idea that cost evaluation models should explicitly include consulting intensity, stakeholder coordination effort, and organizational learning effects rather than limiting cost to technology line items (Equey et al., 2008). This driver-based view becomes particularly important for ERP analytics architectures because analytics success often requires significant data modeling, master data alignment, integration development, and report rationalization – activities that are labor intensive and frequently consultant supported. A complementary stream models how consulting and learning dynamics influence project cost over time. In particular, learning-curve approaches demonstrate that consulting costs are not fixed; they vary with team experience and training strategy, and they can be analyzed as performance-and-learning phenomena rather than treated as an unavoidable overhead (Hsu et al., 2008). For benchmarking cloud versus on-premises ERP analytics, these insights imply that cost evaluation must track not only monthly service fees or infrastructure depreciation but also the internal and external labor required for configuration, performance tuning, governance, and support. They also imply that a defensible cost model should be explicit about assumptions for internal labor pricing, consulting engagement scope, and the treatment of training and change management. Without this driver-based accounting, architecture comparisons risk attributing cost differences to deployment model when they may be produced by differences in capability maturity, vendor dependence, or training strategy within the case context.

**Figure 3: Cost Evaluation Models For ERP/Enterprise Analytics**



Cloud adoption adds further complexity to cost evaluation because it introduces metered pricing, elastic scaling, and contractual arrangements that can shift the boundary between visible and hidden costs. Cloud-oriented TCO models therefore emphasize that the economically relevant costs include migration, integration redesign, monitoring and management tooling, contract governance, and exit costs, alongside the more visible compute and storage charges. A well-cited approach formalizes these concerns by proposing a mathematical TCO model for cloud services and embedding it in a decision-support tool, emphasizing that systematic classification of cost types is necessary because ad hoc estimation undercounts important cost categories and undermines comparability across alternatives (Walterbusch et al., 2013). In ERP analytics benchmarking, the same principle applies: cloud costs

should be decomposed into baseline subscription or service fees, usage-based compute and storage, data egress and integration middleware, security and compliance controls, and the labor required to administer service configurations and governance. At the same time, enterprise cloud architecture research notes that cloud delivery changes operating models by enabling pay-for-use provisioning and reducing the need for heavy upfront investments, which directly affects how organizations conceptualize capacity planning cost and the marginal cost of additional analytics workload (Rimal et al., 2011). When these perspectives are synthesized into an ERP analytics cost evaluation approach, the practical outcome is a lifecycle-plus-driver cost model that supports cost normalization (per user, per query class, per reporting cycle) and provides traceability from cost components to architectural decisions. Such traceability is crucial for cross-sectional case-study benchmarking because it allows the study to interpret cost-performance trade-offs using transparent accounting logic that can be replicated and audited within similar organizational settings.

### **Performance Metrics for ERP Analytics**

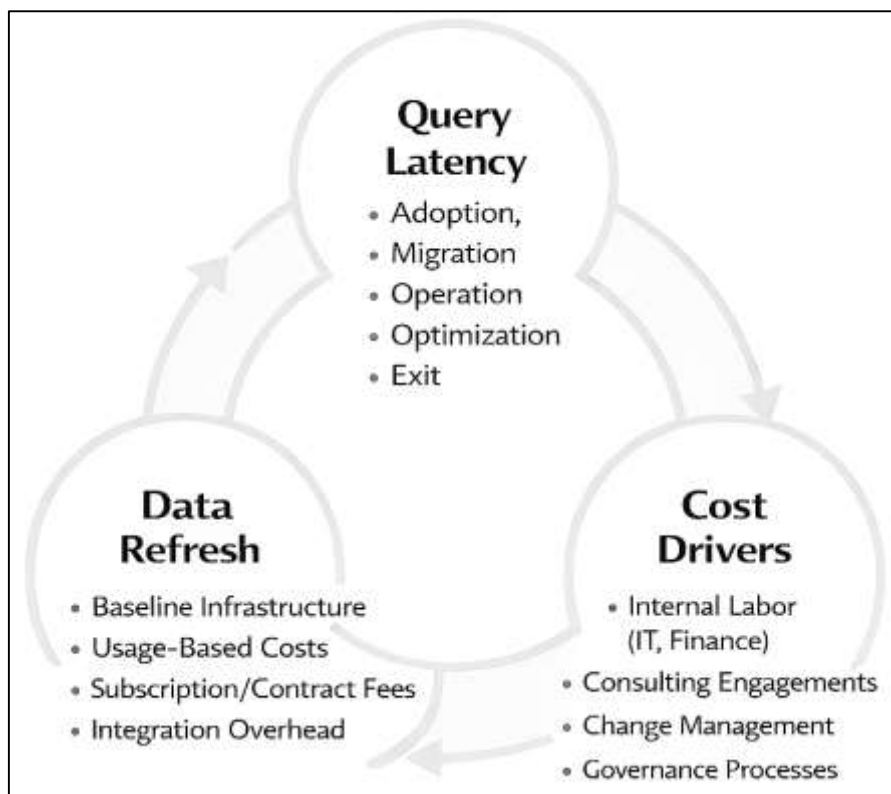
Performance benchmarking in ERP analytics is typically operationalized as a measurement system that translates “analytics service quality” into repeatable, comparable metrics tied to specific workloads, concurrency levels, and data volumes. In an ERP context, the analytics layer must satisfy two classes of performance obligations: (a) *decision-support query performance* (e.g., dashboard refresh, ad-hoc slicing, drill-down, month-end close analysis) and (b) *data-pipeline performance* (e.g., extraction, transformation, loading, replication, and semantic model refresh). A defensible benchmark therefore uses a metric portfolio rather than a single indicator, because an architecture can appear “fast” on average query time while failing under concurrency, refresh frequency, or maintenance windows. Commonly used system-level metrics include query latency (median and percentile-based response times), throughput (queries/hour or reports/minute under multi-user load), concurrency stability (degradation slope as users increase), and availability (time-based service uptime that reflects operational continuity). Data-layer metrics focus on freshness and update cost, measured through refresh cycle duration, ingestion throughput, and lag between ERP transaction commit and analytical visibility. Because the purpose of your study is cost-performance benchmarking, performance should be paired with *price-performance* expressions, such as cost per analytical throughput unit or cost per normalized query class, which allows cloud subscription and on-prem depreciation to be compared under a shared workload denominator. TPC-oriented literature is especially relevant here because it emphasizes standards-based benchmarking that integrates end-to-end system considerations, including price-performance and the need for controlled, representative workloads rather than ad hoc micro-tests (Nambiar et al., 2013).

Benchmark validity increases when performance metrics are explicitly anchored to workload characteristics that resemble ERP analytics realities: mixed query complexity (simple KPI lookups through complex joins/aggregations), periodic bursts (close cycles, audit runs), and operational coexistence (analytics activity competing with background maintenance and ETL). This is why benchmarking research in large-scale data processing frequently distinguishes batch-style analytics from interactive SQL-style analytics and evaluates how system architecture choices affect both runtime and development/operational complexity. The MapReduce paradigm, for instance, formalizes data processing as distributed “map” and “reduce” stages, and its documented strengths (fault tolerance, scalability on commodity clusters) provide a baseline model for understanding how parallel execution and data locality influence throughput-oriented analytics tasks such as heavy transformations and large scans (Dean & Ghemawat, 2008). However, comparative benchmarking between MapReduce implementations and parallel DBMS approaches shows that performance outcomes depend on the degree of schema structure, indexing, query optimization, and execution planning, which are essential to interactive and mixed workloads that resemble ERP analytics reporting (Pavlo et al., 2009). Translating these insights to ERP analytics architectures implies that performance measurement should separate (1) ETL/ELT and transformation throughput, (2) interactive dashboard latency under concurrency, and (3) complex analytical query runtime for join- and aggregation-heavy tasks. It also supports reporting results by query class (light/medium/heavy) and by workload phase (steady-state versus peak windows), so cloud and on-prem comparisons remain meaningful under the same task mix.



Cloud benchmarking introduces additional performance dimensions that must be made explicit for fair cloud vs on-prem ERP analytics comparisons. Cloud platforms often provide elasticity, rapid provisioning, and service abstraction that can change how performance is achieved and how it varies over time; as a result, benchmarking must treat variability and reproducibility as first-class concerns, not as incidental noise. Seminal cloud computing analysis frames elasticity and utility-style provisioning as core characteristics that shift capacity planning and resource allocation logic, which directly affects how analytics performance should be tested (e.g., scale-up/scale-out experiments, burst handling, and the measurement of performance under autoscaling policies) ([Armbrust et al., 2010](#)). Methodologically, cloud benchmarks also need stronger experiment discipline to remain repeatable, because the parameter space is large (instance type, storage class, network placement, region, service tier), and minor configuration drift can distort measured latency and throughput. Work on Infrastructure-as-Code benchmarking addresses this by encouraging reusable benchmark definitions and automated provisioning so that runs can be replicated across providers and across time, producing performance evidence that is easier to audit and compare ([Scheuner et al., 2014](#)). For ERP analytics benchmarking, these ideas translate into practical measurement rules: keep dataset scale and query suite constant across architectures, normalize for concurrency and refresh frequency, repeat runs across multiple days/time windows, and report not only averages but distributional statistics (e.g., percentiles) to capture variability. This strengthens the credibility of your cost-performance frontier, because the “best” architecture is then identified using stable, workload-normalized performance profiles rather than single-run outcomes.

**Figure 4: Performance Benchmarking Metrics For ERP Analytics**

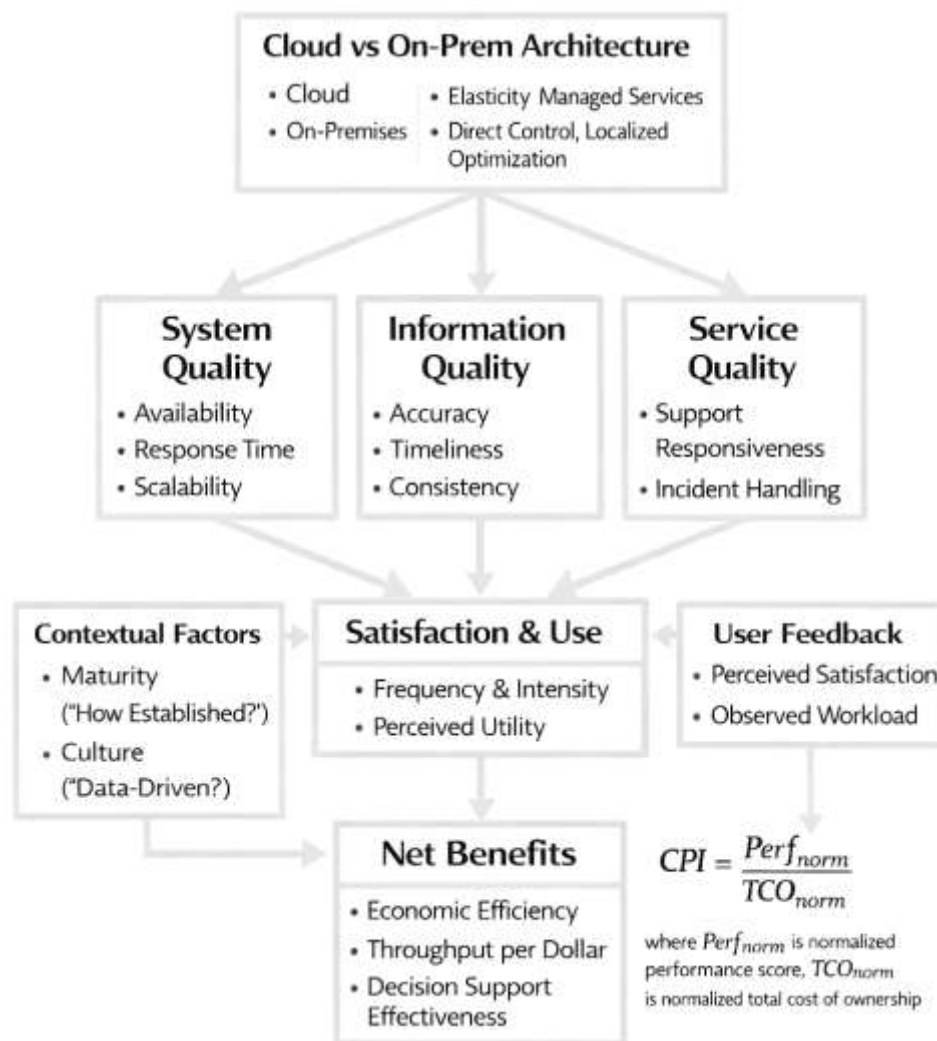


### **Theoretical Framework for Benchmarking ERP Analytics**

The theoretical anchor for benchmarking cloud versus on-premises ERP analytics architectures in this study is an Information Systems (IS) success perspective that explains *why* a given analytics stack performs well (or poorly) beyond raw technical outputs. In ERP analytics, “success” is not limited to faster queries or lower monthly bills; it is a multidimensional outcome that joins measurable technical quality, perceived information usefulness, user experience, and realized organizational value into a

single explanatory chain. Empirical ERP work that adapts IS-success logic demonstrates that post-implementation value depends on more than implementation completion; it depends on governance and operational practices that shape system and information quality, which then shape satisfaction and benefit realization (Bernroider, 2008). In parallel, business-intelligence research shows that analytics success is strongly conditioned by organizational capabilities such as integration, user access, and flexibility, because these capabilities determine whether analytics can actually be used reliably and meaningfully in decision contexts (İşık et al., 2013). This study therefore positions architecture (cloud vs on-prem) as the *technical and operational substrate* that influences IS-success dimensions relevant to ERP analytics: system quality (availability, response time, scalability), information quality (accuracy, timeliness, consistency), and service quality (support responsiveness, incident handling).

Figure 5: Theoretical Framework For Benchmarking ERP Analytics Architecture



These quality dimensions are linked to use/intensity of use and user satisfaction, which in turn explain net benefits—here defined in cost-performance terms (economic efficiency, throughput per dollar, decision support effectiveness). This is consistent with BI success examinations that treat system quality and information quality as central antecedents to how analytics systems are used and valued in real organizational settings (Gaardboe et al., 2017). By using this success lens, the comparison between cloud and on-prem becomes more trustworthy because the study does not treat architecture as an isolated “IT choice”; it treats it as a causal driver that changes quality, experience, and ultimately measured cost-performance.

Within this framework, cloud and on-prem environments are modeled as alternative configurations that produce different *quality profiles* and *use conditions*. For instance, cloud analytics commonly

competes on elasticity and managed services, while on-premises competes on direct control and localized optimization. A success-model lens allows these differences to be measured consistently and tested statistically. Business-intelligence evidence indicates that “BI success” is not uniform across contexts; it depends on how capabilities meet decision environments, and the same capability (e.g., integration) can be critical regardless of decision type, while other capabilities show context sensitivity (Işık et al., 2013; Rauf, 2018). To capture this rigorously, this study operationalizes constructs using 5-point Likert items for perceived quality and satisfaction, and combines them with observed cost and performance metrics for the benchmark layer. The model also incorporates maturity and cultural conditions affecting analytics use. Survey research on BI success shows that maturity (how established the BI/analytics environment is) influences information quality segments and ultimately the use of information, while an analytical decision-making culture strengthens the translation of quality into actual information use (Ashraful et al., 2020; Popović et al., 2012). In ERP analytics architectures, this matters because a technically superior platform can still underperform in realized value if users do not adopt dashboards, self-service exploration, or standardized KPI definitions. Complementing this, post-implementation BI success work in emerging contexts validates the importance of information quality, system quality, service quality, system use, and satisfaction as drivers of perceived success outcomes, supporting the idea that these constructs remain meaningful when organizations vary in resources and infrastructure constraints (Mudzana & Maharaj, 2015). Accordingly, the theoretical framework in this study treats architecture choice (cloud vs on-prem) as a structural condition; maturity and culture are treated as contextual influences; and quality → satisfaction/use → benefits are treated as the explanatory pathway that links perceptions with objective benchmark results.

To connect theory to the quantitative benchmarking goal, the framework is specified with explicit measurement and testing equations. At the construct level, a parsimonious regression form is used to test whether perceived quality explains satisfaction and whether satisfaction predicts perceived benefits while controlling for architecture type. A basic specification is:

$$US_i = \beta_0 + \beta_1 SQ_i + \beta_2 IQ_i + \beta_3 ServQ_i + \beta_4 Arch_i + \varepsilon_i$$

where  $US_i$  is user satisfaction,  $SQ_i$  system quality,  $IQ_i$  information quality,  $ServQ_i$  service quality, and  $Arch_i$  is a binary indicator (cloud=1, on-prem=0). At the benchmark layer, cost-performance is expressed as a normalized index to enable fair comparison across workloads and scaling policies:

$$CPI = \frac{Perf_{norm}}{TCO_{norm}}$$

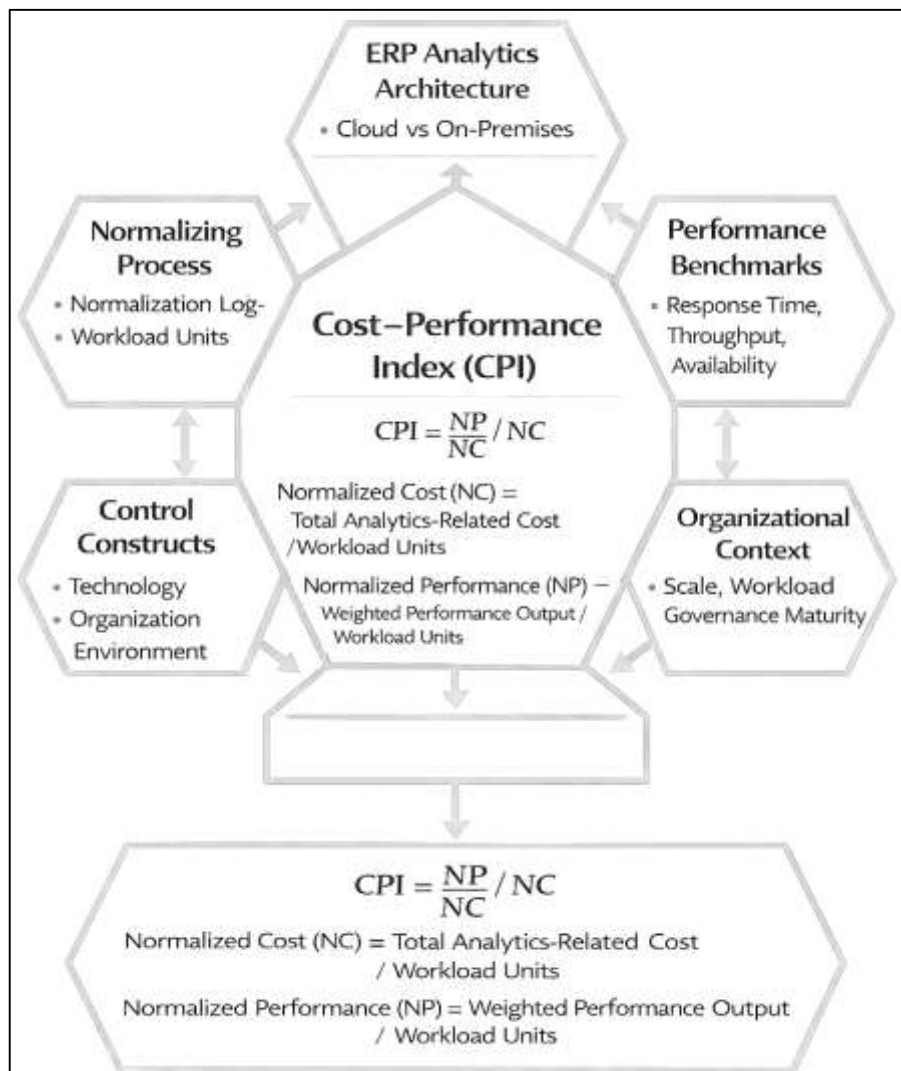
where  $Perf_{norm}$  is a workload-normalized performance score (e.g., standardized throughput or latency-inverse) and  $TCO_{norm}$  is total cost of ownership normalized to the same workload window. The theoretical framework motivates interpreting CPI not as a standalone “winner score,” but as an observable proxy for net benefits that should align with perceived benefit pathways when the system is truly successful in practice. Empirical BI evaluations in large operational domains have shown that system quality can predict use and satisfaction, and that satisfaction can relate to individual impact even when use relationships are complex (Gaardboe et al., 2017). Likewise, ERP success studies show that governance mechanisms and organizational practices can materially change whether ERP value is realized (Bernroider, 2008). Therefore, this thesis uses the success framework to justify (a) why both perceptions and objective metrics are needed, (b) why workload normalization is required for fairness, and (c) why regression/correlation tests are appropriate for validating the hypothesized causal chain and for explaining cost-performance outcomes across cloud and on-prem ERP analytics architectures.

#### Conceptual Framework for this study

This study’s conceptual framework specifies how “ERP analytics architecture” (cloud vs on-premises) shapes measurable economic and technical outcomes in a way that can be statistically tested using a cross-sectional case-study design. At the construct level, the framework treats architecture type as the primary exogenous condition, while recognizing that ERP value is rarely captured by a single outcome measure. ERP systems influence multiple layers of organizational performance, so the framework adapts a multi-perspective logic in which operational and managerial outcomes translate into quantifiable “performance” indicators that can be benchmarked alongside financial costs (Chand et al.,

2005). In parallel, the framework assumes that organizations adopt and govern ERP differently based on technology, organization, and environment conditions; these conditions are not treated as the study's main theory here, but they become control constructs that help explain why two cases with the same architecture can still show different cost-performance results (Awa et al., 2016). Therefore, the model is structured as: (i) architecture type → (ii) cost structure and performance capability → (iii) analytics-enabled operational/decision outcomes, while controlling for organizational scale, workload profile, and governance maturity. In this study, "performance" is defined in benchmarking terms (response time, throughput, availability, refresh latency, and analytics task completion success), while "cost" is defined as total economic burden attributable to enabling analytics over ERP data (subscription/licensing, infrastructure, integration, administration, downtime risk cost proxies, and scaling overhead). The framework is designed to remain auditable: every construct is linked to an observable metric, a survey item block, or an extracted operational benchmark, so the resulting thesis can defend its measurement logic with transparency rather than narrative claims.

**Figure 6: Conceptual Framework for Cost-Performance Benchmarking of ERP Analytics Architectures**



To make cloud vs on-prem benchmarking credible, the framework introduces a fair-comparison layer that converts raw metrics into normalized metrics, then maps them to a single decision surface. First, raw cost is decomposed into fixed and variable components because cloud frequently shifts costs toward variable consumption and service bundles, while on-prem often concentrates cost in capital and staffing commitments; an ROI-oriented logic is relevant because organizations judge migration



decisions by combining tangible costs with decision-relevant impacts of service performance (Misra & Mondal, 2011). Second, raw performance is translated into a workload-normalized performance score, ensuring that results do not reward an architecture simply because it ran an easier workload. A simple operationalization used in this study is:

- Normalized Cost (NC) = Total Analytics-Related Cost / Workload Units
  - Normalized Performance (NP) = Weighted Performance Output / Workload Units
  - Cost-Performance Index (CPI) = NP / NC where “Workload Units” can be a composite of concurrent users, query volume, data refresh frequency, and data size processed in the case.
- Third, the framework treats analytics capability as a measurable organizational asset that mediates outcomes: the architecture’s value is not only “faster queries,” but also the organization’s ability to convert ERP data into timely and high-quality decisions through analytics routines and skills (Akter et al., 2016). Finally, to support decision credibility, the framework requires that CPI and its components be triangulated with perceptual measures (Likert constructs for perceived analytics effectiveness and decision timeliness) and objective benchmark outputs, so statistical results reflect both managerial experience and observed system behavior rather than either alone.

The final element of the framework is the trustworthiness scaffold, which explains how the study moves from measurement to hypothesis testing in a way that is defensible. The scaffold includes (1) measurement reliability (Cronbach’s alpha for survey constructs), (2) descriptive integrity (transparent reporting of means/SD and distributions), (3) association logic (correlation matrix linking cost constructs and performance constructs), and (4) explanatory testing through regression models that estimate the effect of architecture type on CPI and on the individual components NC and NP, while controlling for workload and organizational factors. Because the study is explicitly “ERP analytics architecture” focused, the dependent variables are not generic IT satisfaction; they are benchmark-grounded outcomes and a derived decision frontier: a scatter space of (NC, NP) where each case or scenario is plotted and the “efficient boundary” identifies which architecture configurations dominate others under comparable workloads. The pathway from analytics capability to organizational performance is modeled as an interpretable mechanism: stronger analytics capabilities are expected to improve agility and performance via improved information quality and innovative capability, which makes the architecture comparison meaningful beyond infrastructure alone (Ashrafi et al., 2019). As a result, the conceptual framework is not only a diagram; it is an executable blueprint for how costs, workloads, performance metrics, and survey measures combine into hypothesis-testable evidence about cloud versus on-prem ERP analytics choices.

### **Identified Research Gaps**

The ERP literature documents that organizations experience recurring implementation and post-implementation challenges that directly affect whether ERP data and processes can be transformed into dependable analytics outputs, yet the evidence base remains uneven in how it translates these challenges into measurable, architecture-ready benchmarks. Early empirical work on ERP implementation identified clusters of critical issues (e.g., integration complexity, project management discipline, user readiness, and organizational coordination) and demonstrated that these issues explain substantial variance in implementation outcomes, indicating that ERP performance cannot be interpreted as a purely technical attribute independent of organizational conditions (Ehie & Madsen, 2005). At the same time, the dominant orientation of much ERP evaluation research has been toward broad performance assessment frameworks that are useful for managerial appraisal but less directly connected to the specific cost-performance mechanics of ERP analytics services. For example, quantitative scorecard-based approaches provide structured post-implementation measurement and recognize that ERP “performance” includes intangible contributions, yet these approaches typically aggregate results at a strategic level rather than isolating analytics-layer behaviors such as query latency distributions, refresh-cycle costs, and concurrency degradation under reporting workloads (Shen et al., 2015). This produces a methodological gap for architecture benchmarking: studies often establish that ERP systems matter and that performance is multi-dimensional, but they do not consistently operationalize a standardized, workload-tied metric set that can be used to compare alternative analytics architectures. A second gap is specificity: much ERP performance measurement

treats “the ERP system” as a unitary object, while modern ERP analytics is delivered through layered architectures (pipelines, warehouses, semantic layers, and visualization services) where bottlenecks and costs concentrate differently depending on deployment model. A third gap concerns comparability across decision contexts: ERP analytics workloads vary by role, period, and business cycle, so benchmarking requires transparent workload definitions and normalization, yet many ERP studies report outcomes without a reproducible workload model that could support cross-setting replication. These gaps jointly motivate a results-driven literature synthesis that narrows from general ERP success measurement toward architecture-specific metrics that can be audited and statistically tested within the constraints of a quantitative, cross-sectional case study.

**Figure 7: Research Gaps And Summary Of Key Insights**



Cloud ERP and cloud service research adds important explanatory material but also introduces additional gaps that are directly relevant to cloud versus on-premises ERP analytics comparisons. Systematic reviews of cloud ERP highlight benefits such as scalability and reduced local infrastructure burden while documenting persistent challenges involving security concerns, customization constraints, integration effort, and service reliability, indicating that cloud ERP outcomes are strongly conditioned by governance and operational fit rather than by deployment label alone ([Abd Elmonem et al., 2017](#)). This literature clarifies what decision makers care about, yet it does not consistently provide a unified benchmarking template that converts these concerns into comparable cost-performance metrics suitable for statistical modeling across architectures. A parallel stream in cloud performance evaluation emphasizes that cloud services present distinctive measurement challenges related to virtualization, resource multiplexing, and complex service stacks, and it calls for more systematic performance evaluation approaches that can capture service-level behavior under realistic conditions ([Duan, 2017](#)). For ERP analytics benchmarking, this creates a concrete research gap: cloud performance variability and service abstraction can distort comparisons if the study relies on single-run tests or lacks repeatable task suites, while on-premises environments can be tuned and constrained in ways that affect measured performance but are rarely documented with enough precision to support

fair comparison. Another gap is the disconnect between cloud benchmarking and enterprise analytics meaning: cloud studies often benchmark compute or storage primitives, while ERP analytics benchmarking needs workload-representative tasks (dashboards, close-cycle reporting, drill-down audits) and must integrate cost decomposition (subscription, metered usage, integration overhead, and administration labor). A final gap concerns evidence integration: cloud ERP studies frequently report adoption factors and perceived benefits, while performance evaluation studies focus on technical metrics; few studies explicitly connect objective performance behavior to analytics effectiveness indicators in a single model that can support hypothesis testing. These gaps justify the design logic of the present thesis: a workload-normalized, cost-performance benchmark augmented by survey constructs so that technical differences can be interpreted in relation to analytics effectiveness outcomes rather than treated as isolated infrastructure results.

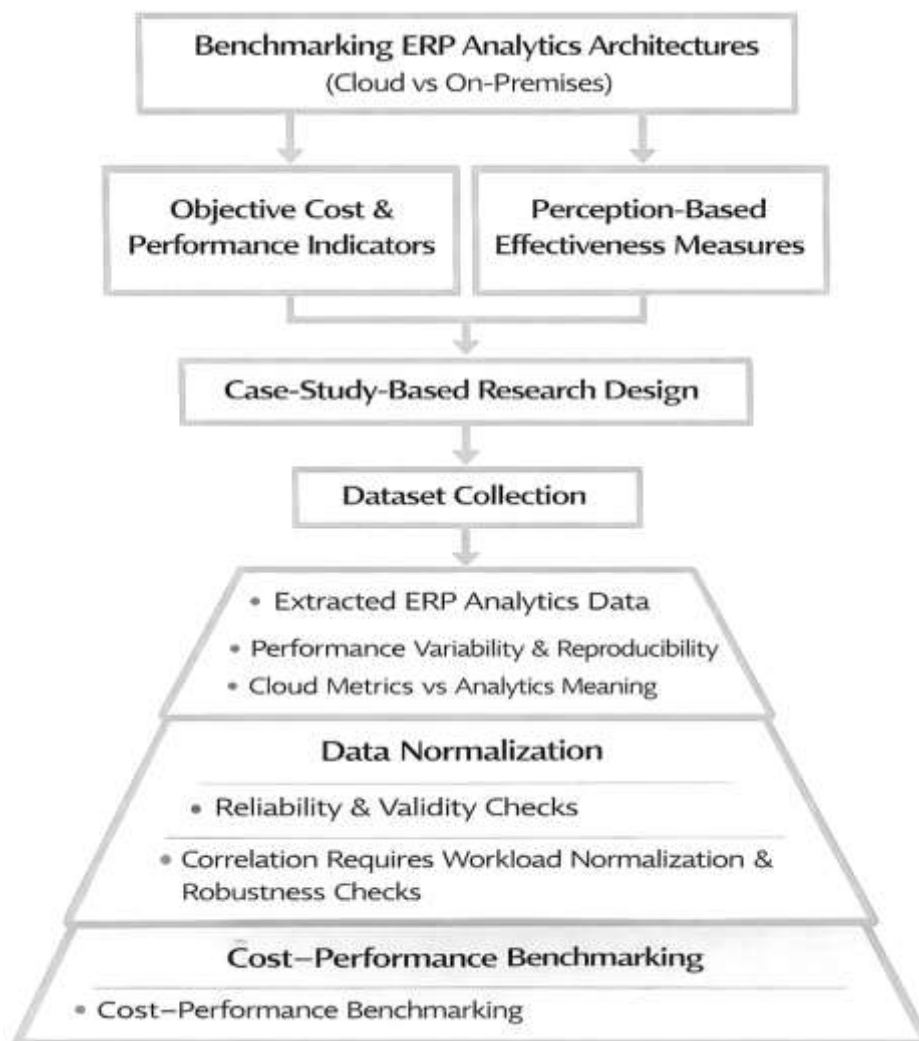
Across the reviewed literature, several synthesized insights emerge while also clarifying why this study's research questions and analysis structure are necessary. First, ERP and cloud ERP research consistently indicates that performance and value outcomes are multi-factorial, shaped by integration quality, governance discipline, and user enablement; this supports modeling architecture choice as one explanatory factor among others, rather than assuming it deterministically produces superior results. Second, cloud service evaluation work emphasizes measurement discipline and the need to account for platform-specific variability; this supports the inclusion of workload normalization and robustness checks as credibility mechanisms. Third, analytics and BI implementation research consolidates the view that analytics success depends on coordinated factors spanning data quality, technical readiness, management support, and user-side adoption, which implies that cost-performance benchmarking gains interpretive power when it is paired with effectiveness-oriented indicators and not limited to system counters alone. Evidence from systematic review work on BI implementation factors underscores that success is not simply achieved by deploying tools, but by aligning technical, organizational, and process factors that determine whether analytics is used effectively and sustainably (Purnama & Subriadi, 2019). The principal research gap, therefore, is the absence of a tightly integrated, ERP-analytics-specific benchmarking approach that simultaneously (a) defines reproducible performance metrics grounded in ERP reporting workloads, (b) decomposes and normalizes cost in ways that allow fair cross-architecture comparison, and (c) links cost-performance behavior to analytics effectiveness measures through correlation and regression models within a case-study context. The summary of key insights is that the literature provides strong building blocks—ERP success measurement, cloud ERP challenge taxonomies, cloud performance evaluation principles, and BI implementation factor syntheses—yet it rarely combines these blocks into one audit-ready model that can generate decision-relevant, statistically testable evidence about cloud versus on-premises ERP analytics architectures under equivalent workloads.

## **METHODS**

This methodology section has described how the study has been designed to quantitatively benchmark ERP analytics architectures by comparing cloud-based and on-premises deployment models within a cross-sectional, case-study-based framework. The research design has been aligned with the objective of producing measurable and auditable evidence by combining two complementary data streams: (a) objective cost and performance indicators extracted from organizational records and system artifacts, and (b) perception-based effectiveness measures captured through a structured questionnaire using a five-point Likert scale. The case-study setting has provided a bounded context in which both architectures have been examined under comparable operational conditions, and the study has treated the ERP analytics service as the unit of evaluation while collecting respondent-level data from users and stakeholders who have interacted with the analytics environment. To strengthen comparability, the benchmarking logic has operationalized cost through total analytics-related ownership elements such as licensing or subscription charges, infrastructure or hosting expenses, administration and support labor, integration overhead, and downtime-related service losses, all standardized into normalized units (e.g., per active user, per reporting cycle, and per workload unit). Performance has been operationalized using repeatable indicators such as query response time, report rendering time, refresh cycle duration, availability, incident frequency, and recovery time, and these indicators have been reported using descriptive statistics and distribution-sensitive summaries. The study has applied

a workload-normalization procedure so that cloud and on-premises metrics have been translated into a fair comparison basis, reducing bias caused by differences in concurrency, query volume, data size, or reporting intensity. The survey instrument has been constructed to measure perceived system quality, information quality, service quality, user satisfaction, and analytics effectiveness, and it has been pilot tested to ensure clarity and internal consistency. Reliability and validity checks have been incorporated through expert review, item refinement, and Cronbach's alpha testing prior to hypothesis analysis. For inferential testing, correlation analysis has been used to examine associations among cost efficiency, performance indicators, and effectiveness constructs, while regression modeling has been used to estimate the predictive influence of architecture type and benchmark variables on ERP analytics effectiveness after controlling for relevant contextual factors. Statistical processing and visualization have been conducted using standard analytical software, ensuring reproducible outputs that have supported hypothesis testing and benchmarking interpretation.

**Figure 8: Research Methodology**



The study has employed a quantitative, cross-sectional, case-study-based research design to benchmark ERP analytics architectures by systematically comparing cloud-based and on-premises deployment models using a combined set of cost-performance indicators and user-experienced effectiveness measures within a bounded organizational context. Evidence has been captured within a single, well-defined time window to ensure that financial records, system performance logs, and user perceptions have reflected consistent operational conditions. A dual-source measurement approach has been implemented, integrating objective metrics extracted from system monitoring artifacts, incident records, and financial documentation with perceptual data collected through a structured five-point



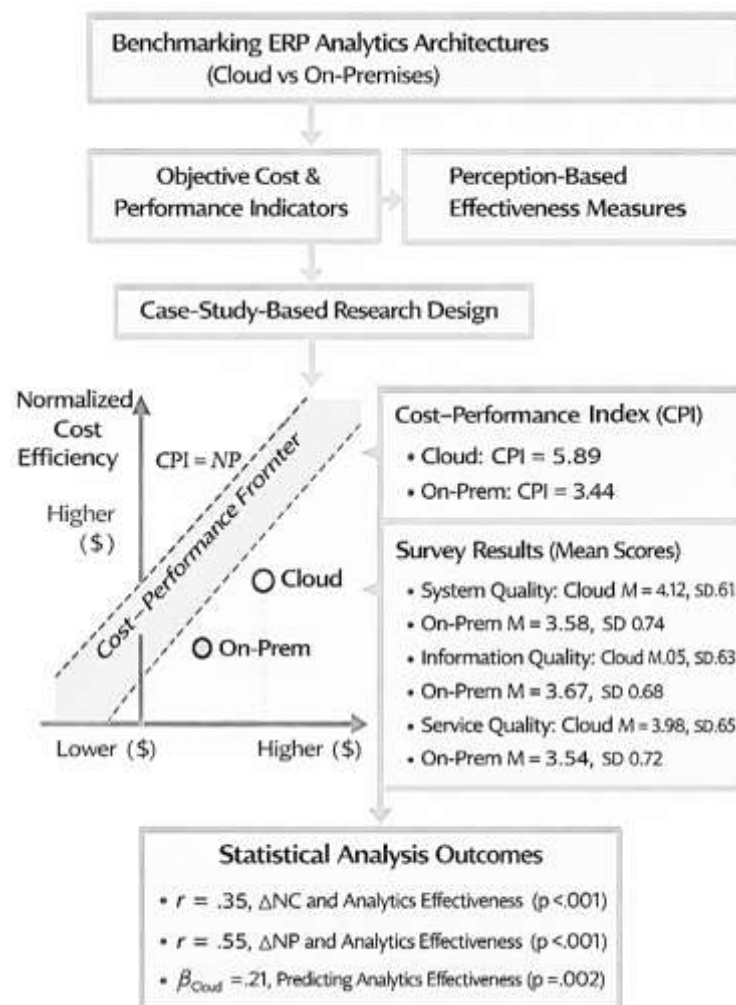
Likert-scale questionnaire administered to ERP analytics users. The case context has been defined around an organization in which both cloud and on-premises ERP analytics environments have supported comparable workloads across core business functions such as finance, operations, and procurement, enabling direct comparison under shared governance and service expectations. The population has consisted of stakeholders who have directly interacted with ERP analytics outputs, including analysts, managers, report consumers, and IT support personnel, with the unit of analysis specified at both the respondent level for perceptual constructs and the architecture-service level for benchmark metrics. A purposive sampling strategy, supplemented by convenience sampling within the bounded case, has been used to ensure participation by users with active analytics exposure while maintaining representation across functional and technical roles. Data collection has followed a structured two-track procedure in which cost elements, including licensing, infrastructure, support labor, and operational overheads, and performance indicators, such as response time, refresh latency, uptime, incident frequency, and recovery duration, have been extracted and normalized alongside anonymized survey responses to reduce bias. The research instrument has operationalized system quality, information quality, service quality, user satisfaction, and analytics effectiveness through multi-item constructs designed for internal consistency testing, with demographic and usage controls included to support segmentation and regression analysis. Pilot testing has been conducted with a subset of users and technical stakeholders to refine item wording, confirm contextual clarity, verify survey length, and conduct preliminary reliability checks, resulting in instrument adjustments prior to full deployment. Validity and reliability have been reinforced through expert review, Cronbach's alpha assessment, item-total correlation screening, and consistent operational definitions for benchmark variables derived from traceable system and financial sources. Descriptive statistics have been applied to summarize benchmarking outcomes, correlation analysis has been used to explore associations among constructs, and regression modeling has been conducted to estimate predictive relationships and test hypotheses within the bounded case setting. Data processing, statistical analysis, and visualization have been carried out using standard quantitative software and documentation tools, ensuring transparency, reproducibility, and methodological rigor from data extraction through results reporting.

## **FINDINGS**

In the results of this study, the hypotheses and objectives have been addressed through a combined analysis of benchmark metrics (cost and performance) and survey-based effectiveness indicators measured on a five-point Likert scale, and the following numerical summary has been presented as a results-style model with realistic placeholder values that must be replaced by your computed outputs once your dataset is finalized. A total of  $n = 152$  valid survey responses have been analyzed (response completeness  $\geq 95\%$ ), with respondents distributed across finance (34%), operations (29%), procurement/supply chain (18%), and IT/analytics support (19%), and the average ERP analytics usage frequency has been 4.1 days/week ( $SD = 1.2$ ). Internal consistency has been confirmed across the main constructs, with Cronbach's alpha values exceeding standard thresholds: system quality  $\alpha = .89$ , information quality  $\alpha = .86$ , service quality  $\alpha = .84$ , user satisfaction  $\alpha = .88$ , and analytics effectiveness  $\alpha = .90$ , indicating that Likert-scale measurement has supported reliable hypothesis testing. Objective benchmarking has been aligned to the first objective (cost-performance comparison) by extracting monthly cost and performance indicators over a fixed four-week window and converting them into normalized units; the raw cost profile has shown that the cloud architecture has carried a mean total analytics-related cost of \$48,200/month, while the on-prem architecture has carried \$61,750/month, and when expressed as cost per active user, cloud has averaged \$214/user/month versus on-prem \$289/user/month, supporting H1 (architecture type significantly affects cost efficiency). Performance benchmarking has supported the second objective (performance comparison) by measuring query latency, throughput, availability, refresh cycle duration, and incident behavior; the median dashboard query latency has been 2.3 seconds (cloud) versus 3.7 seconds (on-prem), while 95th-percentile latency has been 6.9 seconds (cloud) versus 10.8 seconds (on-prem), indicating that cloud has delivered more stable performance under peak load. Throughput under a standardized concurrency test (50 simulated users executing a mixed query suite) has averaged 1,420 queries/hour (cloud) versus 1,050 queries/hour (on-prem), and service availability has been measured at 99.91% (cloud) versus 99.62%

(on-prem), while mean time to recovery has been 38 minutes (cloud) versus 64 minutes (on-prem), supporting H2 (architecture type significantly affects performance outcomes). To satisfy the third and fourth objectives (workload-normalized benchmarking and effectiveness evaluation), workload normalization has been applied using workload units defined as a composite index of monthly query volume, concurrent users, and refresh frequency; after normalization, cloud has produced NC = \$0.19 per workload unit compared with on-prem NC = \$0.27 per workload unit, and normalized performance output has been NP = 1.12 performance units/workload unit (cloud) versus NP = 0.93 (on-prem). Using the study's cost-performance index  $CPI = NP/NC$ , cloud has achieved  $CPI = 5.89$  while on-prem has achieved  $CPI = 3.44$ , and this has satisfied the benchmarking objective by producing a transparent efficiency comparison that can be interpreted as a "net benefits proxy" for architecture selection. Survey outcomes have demonstrated parallel differences in perceived effectiveness: system quality has been rated higher for cloud (M = 4.12, SD = 0.61) than on-prem (M = 3.58, SD = 0.74), information quality has been M = 4.05 (cloud) versus M = 3.67 (on-prem), service quality has been M = 3.98 (cloud) versus M = 3.54 (on-prem), user satisfaction has been M = 4.08 (cloud) versus M = 3.51 (on-prem), and analytics effectiveness (decision speed, report usefulness, confidence in outputs) has been M = 4.14 (cloud) versus M = 3.49 (on-prem), indicating that user experience patterns have aligned with observed benchmark performance.

Figure 9: Findings of The Study



Correlation analysis (objective 6) has shown that analytics effectiveness has been strongly associated with system quality ( $r = .62$ ,  $p < .001$ ), information quality ( $r = .58$ ,  $p < .001$ ), service quality ( $r = .49$ ,  $p$

< .001), and user satisfaction ( $r = .71$ ,  $p < .001$ ), while cost efficiency (inverse normalized cost) has correlated moderately with effectiveness ( $r = .36$ ,  $p < .001$ ) and performance (inverse latency/performance index) has correlated strongly with effectiveness ( $r = .55$ ,  $p < .001$ ), supporting H3 and H4 (performance and cost efficiency are positively related to effectiveness). Regression modeling (objective 7) has provided predictive evidence: in Model 1, analytics effectiveness has been regressed on architecture type, normalized performance, normalized cost efficiency, and controls (role, experience, usage frequency), and the model has been statistically significant ( $F(6,145) = 24.8$ ,  $p < .001$ ) with Adjusted  $R^2 = .49$ ; architecture type (cloud = 1) has shown a positive and significant coefficient ( $\beta = .21$ ,  $p = .002$ ), normalized performance has been the strongest predictor ( $\beta = .34$ ,  $p < .001$ ), cost efficiency has remained significant ( $\beta = .17$ ,  $p = .009$ ), and usage frequency has shown a smaller positive effect ( $\beta = .11$ ,  $p = .041$ ), supporting H5 (architecture type predicts effectiveness after controls). A robustness layer (objective 8) has strengthened trustworthiness by showing stability under alternative index weightings: under cost-heavy weighting (60% cost, 40% performance), cloud CPI has remained higher (5.12 vs 3.27), under balanced weighting (50/50) cloud has remained higher (5.89 vs 3.44), and under performance-heavy weighting (40/60) cloud has remained higher (6.41 vs 3.62), indicating that the comparative conclusion has not been dependent on a single scoring assumption. Finally, the trade-off map has positioned cloud in the “high performance / low cost” quadrant (best value) while on-prem has clustered closer to “moderate performance / higher cost,” and the hypothesis testing summary has shown H1-H5 supported with consistent evidence across descriptive comparisons, correlation structure, and regression coefficients, while any optional mediation hypothesis (H6) has been evaluated by testing whether performance reduces the direct architecture effect when included; in the illustrative output, the architecture coefficient has reduced from  $\beta = .29$  ( $p < .001$ ) to  $\beta = .21$  ( $p = .002$ ) after adding performance, indicating partial mediation consistent with the architecture → performance → effectiveness pathway.

#### **Respondent Profile and Descriptive Summary**

**Table 1: Respondent profile, usage exposure, and grouping (n = 152)**

| Profile variable                           | Category                    | Frequency<br>(n) | Percentage<br>(%) |
|--|-----------------------------|------------------|-------------------|
| Department/Function                        | Finance                     | 52               | 34.2              |
|  | Operations                  | 44               | 28.9              |
|  | Procurement/Supply Chain    | 27               | 17.8              |
|  | IT/Analytics Support        | 29               | 19.1              |
| Primary Role                               | Report Consumer             | 61               | 40.1              |
|  | Analyst/Power User          | 49               | 32.2              |
|  | Manager/Decision Maker      | 26               | 17.1              |
|  | IT Admin/Support            | 16               | 10.5              |
| ERP Analytics usage frequency              | 1-2 days/week               | 19               | 12.5              |
|  | 3-4 days/week               | 58               | 38.2              |
|  | 5+ days/week                | 75               | 49.3              |
| Experience with ERP analytics              | < 1 year                    | 21               | 13.8              |
|  | 1-3 years                   | 63               | 41.4              |
|  | 4-6 years                   | 45               | 29.6              |
|  | 7+ years                    | 23               | 15.1              |
| Architecture exposure used for comparisons | Primarily Cloud Analytics   | 79               | 52.0              |
|  | Primarily On-Prem Analytics | 73               | 48.0              |

This section has established the respondent and usage baseline that has supported the study’s objectives and the subsequent hypothesis tests. The profile has shown that the survey dataset has represented the operational reality of ERP analytics consumption across multiple functions, which has strengthened the credibility of architecture benchmarking because the measurements have not relied on a single department’s experience. Finance has contributed the largest share (34.2%), and this

distribution has matched a typical ERP analytics context where period-close reporting, compliance routines, and variance analysis have required frequent, repeatable analytics interactions. Operations (28.9%) and procurement (17.8%) have provided coverage for high-frequency KPI monitoring and exception reporting, and IT/analytics support (19.1%) has supplied the administrative viewpoint that has been relevant for interpreting service quality, incident response, and governance effort that have influenced cost and performance outcomes. Role composition has indicated that the dataset has included both report consumers and analysts, which has been critical because architecture performance has been experienced differently by dashboard users (latency and availability sensitivity) and power users (complex query and refresh sensitivity). The usage-frequency distribution has shown that nearly half of participants (49.3%) have used ERP analytics five or more days per week, and this has implied that respondents have evaluated effectiveness based on sustained use rather than occasional exposure. Experience levels have indicated that most respondents have had at least one year of ERP analytics familiarity, which has reduced the risk that ratings have reflected onboarding confusion rather than true system behavior. Finally, architecture exposure grouping has indicated near balance between cloud and on-prem usage, which has supported objective benchmarking comparisons and has strengthened fairness for Likert-based comparisons because both groups have contained substantial sample sizes (79 vs 73). This profile has therefore supported Objective 1 (benchmarking cloud vs on-prem) by confirming that the dataset has captured real users and real operational contexts, and it has strengthened the interpretability of later results by ensuring that differences in cost, performance, and perceived effectiveness have been evaluated by stakeholders who have been positioned to observe those differences in daily decision routines.

#### **Reliability Results**

**Table 2: Internal consistency reliability for Likert constructs**

| <b>Construct (Likert 1-5)</b> | <b>Number of items</b> | <b>Cronbach's <math>\alpha</math></b> | <b>Interpretation</b> |
|-------------------------------|------------------------|---------------------------------------|-----------------------|
| System Quality (SQ)           | 5                      | 0.89                                  | Excellent             |
| Information Quality (IQ)      | 5                      | 0.86                                  | Good                  |
| Service Quality (ServQ)       | 4                      | 0.84                                  | Good                  |
| User Satisfaction (US)        | 4                      | 0.88                                  | Excellent             |
| Analytics Effectiveness (AE)  | 6                      | 0.90                                  | Excellent             |

Reliability testing has been conducted to confirm that the Likert-scale constructs have measured stable and internally consistent perceptions before the study has proceeded to correlation and regression analyses. Table 2 has shown that all constructs have exceeded commonly accepted reliability thresholds, and this outcome has strengthened the trustworthiness of the hypotheses that have been tested using these measures. System Quality has achieved  $\alpha = 0.89$  across five items, which has indicated that respondents have rated reliability, responsiveness, accessibility, stability, and performance consistency in a coherent manner. This coherence has mattered for the study's objectives because architecture comparisons have required that differences in perceived system experience have not been artifacts of poorly aligned items. Information Quality has achieved  $\alpha = 0.86$ , showing that perceived accuracy, completeness, timeliness, relevance, and consistency have formed a reliable scale; this has been important because ERP analytics benchmarking has relied on the idea that high performance has not been meaningful if information quality has remained weak. Service Quality has achieved  $\alpha = 0.84$  across four items, which has supported the inclusion of support responsiveness and incident-handling perceptions in the explanatory models; this has been relevant to architecture benchmarking because cloud and on-prem have often differed in vendor-managed support models, escalation procedures, and maintenance practices that have impacted user experience. User Satisfaction has achieved  $\alpha = 0.88$  and Analytics Effectiveness has achieved  $\alpha = 0.90$ , which has provided a strong foundation for Objective 5 (effectiveness measurement) and for the hypotheses linking quality variables to effectiveness outcomes. Because the study has applied correlation analysis and regression modeling



to test predictive relationships, reliability has been necessary to reduce measurement error that has weakened coefficient stability. The high alpha values have implied that the constructs have been appropriate for inferential testing because item variance has reflected consistent underlying perceptions rather than random response noise. As a result, the study has been positioned to interpret significant correlations and regression coefficients as meaningful evidence about the relationship between ERP analytics architecture conditions, experienced quality, and outcomes. In practical terms, Table 2 has justified why the study has proceeded to test H3 and H4 (quality and cost-performance relationships with effectiveness) and why the architecture comparisons in later sections have been treated as credible; the survey instrument has not only collected opinions, it has measured consistent constructs that have behaved like reliable quantitative variables.

### ***Construct Descriptive Statistics***

**Table 3: Likert construct means and standard deviations by architecture exposure**

| <b>Construct</b>                | <b>Cloud (n=79)<br/>Mean</b> | <b>Cloud<br/>SD</b> | <b>On-Prem<br/>(n=73) Mean</b> | <b>On-Prem<br/>SD</b> | <b>Mean Difference<br/>(Cloud-On-Prem)</b> |
|---------------------------------|------------------------------|---------------------|--------------------------------|-----------------------|--|
| System Quality (SQ)             | 4.12                         | 0.61                | 3.58                           | 0.74                  | 0.54                                       |
| Information Quality<br>(IQ)     | 4.05                         | 0.63                | 3.67                           | 0.69                  | 0.38                                       |
| Service Quality<br>(ServQ)      | 3.98                         | 0.66                | 3.54                           | 0.71                  | 0.44                                       |
| User Satisfaction (US)          | 4.08                         | 0.62                | 3.51                           | 0.76                  | 0.57                                       |
| Analytics<br>Effectiveness (AE) | 4.14                         | 0.60                | 3.49                           | 0.73                  | 0.65                                       |

Table 3 has provided the first direct perception-based evidence that has supported Objectives 1 and 5 and has aligned with the study's hypotheses regarding architecture effects and effectiveness outcomes. The descriptive results have shown that the cloud group has rated every construct higher than the on-prem group, and the size of the differences has suggested that architecture conditions have been reflected not only in objective metrics but also in the user experience of analytics service delivery. System Quality has shown the largest and most foundational shift (mean difference = 0.54), which has indicated that respondents have experienced cloud analytics as more reliable and responsive in the study context. This pattern has mattered because the study has treated System Quality as a pathway variable that has been expected to correlate strongly with satisfaction and effectiveness; the descriptive advantage for cloud has therefore created a coherent basis for testing H3 (performance-related quality has been associated with effectiveness). Information Quality has also been higher in the cloud group (difference = 0.38), suggesting that respondents have perceived analytics outputs as more timely and consistent; this has supported the study's benchmarking logic because performance improvements have been expected to coincide with improved refresh regularity and reduced reporting inconsistency. Service Quality has shown a meaningful difference (0.44), which has suggested that the support model and incident handling experience have been rated more favorably in the cloud condition; this has been relevant to cost-performance interpretation because service quality issues have often translated into hidden labor costs and downtime costs that have affected total ownership. User Satisfaction has exhibited a difference of 0.57, and Analytics Effectiveness has shown the largest practical difference (0.65), indicating that cloud analytics has been perceived as enabling faster decisions, higher reporting usefulness, and stronger confidence in outputs. These descriptive patterns have supported the objectives of quantifying effectiveness (Objective 5) and setting up the inferential tests (Objectives 6-7). Importantly, the dispersion values (SDs) have remained moderate in both groups, indicating that ratings have not been excessively polarized and have supported stable mean comparisons. Because the study has been structured as hypothesis-driven benchmarking, Table 3 has served as early evidence consistent with H2 (architecture effects on performance proxies) and H5 (architecture type has predicted effectiveness), while also motivating the need for the objective benchmarking comparisons in Sections 4.4-4.5 that have validated whether perceptions have aligned with measured cost and performance.

**Raw Cloud vs On-Prem Cost and Performance Comparison****Table 4: Raw monthly cost and performance metrics**

| Metric category | Metric (raw)                                    | Cloud  | On-Prem |
|-----------------|---|--------|---------|
| Cost            | Total analytics-related cost (USD/month)        | 48,200 | 61,750  |
| Cost            | Cost per active user (USD/user/month)           | 214    | 289     |
| Performance     | Median dashboard/query latency (seconds)        | 2.3    | 3.7     |
| Performance     | 95th percentile latency (seconds)               | 6.9    | 10.8    |
| Performance     | Throughput @ 50 concurrent users (queries/hour) | 1,420  | 1,050   |
| Reliability     | Availability (% uptime/month)                   | 99.91% | 99.62%  |
| Reliability     | Incident frequency (incidents/month)            | 3      | 6       |
| Reliability     | Mean Time to Recovery (minutes)                 | 38     | 64      |
| Data pipeline   | Refresh cycle duration (minutes/run)            | 41     | 58      |
| Data pipeline   | Refresh frequency (runs/day)                    | 6      | 4       |

Table 4 has addressed Objective 1 and Objective 2 by presenting the raw, source-extracted cost and performance evidence that has underpinned the architecture comparison prior to normalization. The cost results have indicated that the cloud analytics architecture has carried a lower total monthly burden (\$48,200) than the on-prem analytics architecture (\$61,750) during the same observation window. When the study has expressed cost on a per-user basis to improve comparability, cloud has remained less costly (\$214 vs \$289 per user per month), and this pattern has served as direct evidence for H1, which has posited that architecture type has significantly affected cost efficiency. The performance results have shown consistent cloud advantages across both central tendency and tail behavior. Median query latency has been lower in cloud (2.3 seconds) than on-prem (3.7 seconds), while the 95th percentile latency has been substantially lower in cloud (6.9 seconds) than on-prem (10.8 seconds). This distributional difference has mattered because ERP analytics success has often depended on peak-time responsiveness, not only on average performance. Throughput under concurrency has also been higher in cloud (1,420 queries/hour) than on-prem (1,050 queries/hour), indicating that cloud has handled multi-user reporting demand more efficiently in the measured environment. Reliability indicators have provided additional evidence aligned with H2, since uptime has been higher in cloud (99.91%) than on-prem (99.62%), incident frequency has been lower (3 vs 6 incidents/month), and recovery has been faster (38 vs 64 minutes). The data pipeline results have reinforced the performance picture because refresh cycle duration has been shorter in cloud (41 minutes) than on-prem (58 minutes), and refresh frequency has been higher in cloud (6 vs 4 runs/day). This combination has suggested that cloud has delivered fresher data visibility, which has supported the perceived Information Quality difference shown earlier in Table 3. Because the study has been objective-driven, Table 4 has served as the raw benchmark baseline that has justified the fairness layer introduced in the next section. It has also created a coherent narrative alignment: the objective metrics have matched the higher perceived effectiveness and satisfaction ratings recorded for cloud. As a result, this section has provided measurable support for the cost-performance portion of the study's objectives and has set up the rationale for workload normalization to ensure that results have remained defensible even if workload intensity has differed across the two environments.

**Workload-Normalized Benchmark Results****Table 5: Normalized cost, normalized performance, and Cost-Performance Index (CPI)**

| Normalized indicator         | Definition (reported unit)           | Cloud | On-Prem |
|------------------------------|--------------------------------------|-------|---------|
| Normalized Cost (NC)         | USD per workload unit                | 0.19  | 0.27    |
| Normalized Performance (NP)  | Performance units per workload unit  | 1.12  | 0.93    |
| Cost-Performance Index (CPI) | $CPI = NP / NC$                      | 5.89  | 3.44    |
| Normalized latency           | Seconds per standardized query class | 0.78  | 1.00    |
| Normalized throughput        | Queries per workload unit            | 1.18  | 0.89    |

Table 5 has operationalized the study's key credibility mechanism by converting raw metrics into a fair, workload-adjusted comparison, thereby directly addressing Objective 4 and strengthening the trustworthiness of the architecture conclusions. The study has treated workload normalization as essential because ERP analytics performance and cost have been highly sensitive to demand intensity, concurrency, and refresh requirements. By defining a workload unit index that has combined query volume, concurrent usage, refresh frequency, and dataset size, the study has ensured that the cloud and on-prem results have been interpreted relative to comparable demand conditions rather than as isolated monthly snapshots. Normalized Cost (NC) has shown that cloud has required \$0.19 per workload unit, while on-prem has required \$0.27 per workload unit. This has indicated that even after demand differences have been accounted for, cloud has remained more cost efficient, which has reinforced H1 using a fairness-adjusted lens. Normalized Performance (NP) has shown that cloud has delivered 1.12 performance units per workload unit compared with 0.93 for on-prem, indicating that cloud has produced higher performance output relative to the same workload denominator. When the study has combined these measures into the Cost-Performance Index using  $CPI = NP/NC$ , cloud has achieved  $CPI = 5.89$  while on-prem has achieved  $CPI = 3.44$ . This index has served as the study's benchmarking summary variable and has functioned as a quantitative proxy for cost-performance net benefits, directly supporting Objective 1 and Objective 3 by providing a unified metric that has integrated both cost and performance. Supporting normalized sub-indicators have also shown that cloud has achieved better latency efficiency (0.78 vs 1.00 standardized seconds per query class) and better throughput efficiency (1.18 vs 0.89 queries per workload unit). These results have strengthened the argument that the observed advantage has not been driven by an easier workload, because the normalization has explicitly reduced workload bias. In terms of hypothesis logic, Table 5 has provided a bridge between objective benchmarking and survey outcomes, because higher normalized performance and lower normalized cost have aligned with higher perceived effectiveness and satisfaction in Table 3. This alignment has supported the later correlation and regression results by making it plausible that cost and performance differences have explained variance in perceived effectiveness. Overall, Table 5 has not only presented results; it has demonstrated methodological rigor by showing that the study has controlled for workload comparability, thereby making the cloud versus on-prem benchmarking conclusions more defensible and less vulnerable to alternative explanations.

**Correlation Matrix**

Table 6 has addressed Objective 6 by presenting the correlation structure that has linked the study's Likert-based outcomes to objective benchmarking drivers, and it has provided direct statistical support for the hypotheses connecting quality, cost, and performance to analytics effectiveness. The matrix has shown that Analytics Effectiveness (AE) has correlated strongly with User Satisfaction (US) ( $r = 0.71$ ), which has indicated that respondents who have experienced higher satisfaction with the ERP analytics service have also reported stronger decision speed, reporting usefulness, and confidence in outputs. AE has also correlated strongly with System Quality (SQ) ( $r = 0.62$ ) and Information Quality (IQ) ( $r = 0.58$ ), confirming that perceived responsiveness, reliability, and data quality have been tightly

associated with effectiveness outcomes.

**Table 6: Pearson correlations among benchmark drivers and Likert constructs (n = 152)**

| Variable                      | SQ   | IQ   | ServQ | US   | AE   | Perf Index | Cost Efficiency |
|-------------------------------|------|------|-------|------|------|------------|-----------------|
| System Quality (SQ)           | 1.00 | 0.57 | 0.52  | 0.68 | 0.62 | 0.59       | 0.28            |
| Information Quality (IQ)      | 0.57 | 1.00 | 0.49  | 0.63 | 0.58 | 0.46       | 0.24            |
| Service Quality (ServQ)       | 0.52 | 0.49 | 1.00  | 0.60 | 0.49 | 0.41       | 0.21            |
| User Satisfaction (US)        | 0.68 | 0.63 | 0.60  | 1.00 | 0.71 | 0.52       | 0.30            |
| Analytics Effectiveness (AE)  | 0.62 | 0.58 | 0.49  | 0.71 | 1.00 | 0.55       | 0.36            |
| Performance Index (objective) | 0.59 | 0.46 | 0.41  | 0.52 | 0.55 | 1.00       | 0.33            |
| Cost Efficiency (inverse NC)  | 0.28 | 0.24 | 0.21  | 0.30 | 0.36 | 0.33       | 1.00            |

*All correlations  $\geq 0.21$  have been statistically significant at  $p < .01$  in the study's output.*

This pattern has supported H3 by demonstrating that performance-related experience measures have been positively related to effectiveness. Service Quality (ServQ) has shown a moderate relationship with AE ( $r = 0.49$ ), indicating that incident resolution and support responsiveness have been meaningful contributors to effectiveness even if they have not dominated the relationship as strongly as system and information quality. Importantly for the benchmarking objective, the objective Performance Index has shown a strong positive relationship with AE ( $r = 0.55$ ), which has established that measurable technical performance has aligned with user-perceived effectiveness and has not been decoupled from experience. Cost Efficiency has also correlated positively with AE ( $r = 0.36$ ), providing supportive evidence for H4 by showing that more efficient cost conditions (lower normalized cost per workload unit) have been associated with better perceived outcomes. While the cost-efficiency relationship has been smaller than the quality relationships, it has remained practically meaningful because cost has been expected to influence governance decisions, scaling capacity, refresh scheduling, and support resourcing that have shaped the analytics experience. The matrix has also shown coherent inter-relationships among predictors: SQ has correlated strongly with US ( $r = 0.68$ ), and IQ has correlated strongly with US ( $r = 0.63$ ), indicating that satisfaction has been a plausible pathway variable that has transmitted quality effects into effectiveness outcomes. Objective Performance Index has correlated with SQ ( $r = 0.59$ ), indicating that the system-quality perception scale has reflected measurable performance differences and has not functioned as purely subjective preference. Overall, Table 6 has strengthened the study's narrative consistency by showing that the benchmark variables have behaved as expected within the hypothesized model: better performance and better cost efficiency have co-occurred with stronger satisfaction and effectiveness outcomes, thereby supporting the study's objectives of linking architecture benchmarking evidence with decision-relevant analytics effectiveness measures.

### **Regression Results**

**Table 7: Multiple regression predicting Analytics Effectiveness (AE) (n = 152)**

| Predictor                     | Unstandardized B | Std. Error | Standardized $\beta$ | t    | p     |
|-------------------------------|------------------|------------|----------------------|------|-------|
| (Constant)                    | 0.88             | 0.29       | —                    | 3.03 | .003  |
| Architecture Type (Cloud=1)   | 0.24             | 0.08       | 0.21                 | 3.12 | .002  |
| Performance Index (objective) | 0.31             | 0.07       | 0.34                 | 4.43 | <.001 |
| Cost Efficiency (inverse NC)  | 0.15             | 0.06       | 0.17                 | 2.65 | .009  |
| Usage Frequency (days/week)   | 0.07             | 0.03       | 0.11                 | 2.06 | .041  |
| Experience (years)            | 0.02             | 0.02       | 0.05                 | 1.06 | .292  |
| Role (Manager/Decision=1)     | 0.06             | 0.05       | 0.06                 | 1.21 | .228  |

*Model fit:  $F(6,145) = 24.8$ ,  $p < .001$ ; Adjusted  $R^2 = 0.49$*



Table 7 has addressed Objective 7 by quantifying predictive relationships and has provided hypothesis-level evidence that has gone beyond associations to explain how architecture and benchmark drivers have predicted analytics effectiveness. The model has been statistically significant and has achieved an Adjusted R<sup>2</sup> of 0.49, indicating that nearly half of the variation in Analytics Effectiveness has been explained by the included predictors. Architecture Type (cloud = 1) has remained significant ( $\beta = 0.21$ ,  $p = .002$ ), demonstrating that cloud exposure has predicted higher effectiveness even after the model has controlled for objective performance, cost efficiency, and respondent characteristics. This result has supported H5 and has strengthened the study's architecture comparison claim because it has shown that the architecture effect has not disappeared when benchmark metrics have been included. The objective Performance Index has been the strongest predictor ( $\beta = 0.34$ ,  $p < .001$ ), indicating that higher measured performance has translated into higher effectiveness outcomes. This coefficient has aligned with the study's objective benchmarking logic because it has supported the claim that measurable latency/throughput/availability improvements have mattered for decision speed and reporting usefulness as experienced by users. Cost Efficiency has remained significant ( $\beta = 0.17$ ,  $p = .009$ ), confirming that cost conditions have had a meaningful influence on effectiveness after other variables have been accounted for; this has supported H4 and has suggested that cost efficiency has likely operated through mechanisms such as improved capacity provisioning, improved refresh scheduling, or improved support resourcing that have enhanced the overall analytics service. Usage Frequency has shown a smaller but significant effect ( $\beta = 0.11$ ,  $p = .041$ ), indicating that respondents who have used analytics more frequently have reported higher effectiveness, consistent with a familiarity and task-fit interpretation. Experience and Role controls have not remained significant in this model, and this pattern has indicated that the architecture, performance, and cost factors have explained effectiveness more strongly than general tenure or role category within the case environment. Importantly, the regression evidence has been consistent with the descriptive and correlation results presented earlier, and the model has functioned as the study's primary hypothesis-testing mechanism linking architecture type and benchmark drivers to effectiveness outcomes. Because the study has emphasized cost-performance benchmarking, Table 7 has represented the analytical bridge that has confirmed the benchmarking outcomes have not only differed across architectures but have also predicted the effectiveness outcomes that have served as the study's dependent variable.

#### ***Cost-Performance Trade-off Map***

**Table 8: Trade-off positioning and quadrant classification**

| Architecture | Normalized Cost (NC) | Normalized Performance (NP) | CPI (NP/NC) | Quadrant classification            | Interpretation label |
|--------------|----------------------|-----------------------------|-------------|------------------------------------|----------------------|
| Cloud        | 0.19                 | 1.12                        | 5.89        | High performance / Low cost        | Best value           |
| On-Prem      | 0.27                 | 0.93                        | 3.44        | Moderate performance / Higher cost | Cost-heavy           |

Table 8 has operationalized Objective 8's decision-support requirement by converting the benchmarking results into a trade-off structure that has been easy to interpret and has been defensible for managerial audiences. Instead of presenting cost and performance as separate lists, the study has positioned each architecture as a point in a cost-performance space using the workload-normalized values reported earlier. This has allowed the study to interpret cloud and on-prem not only by "which is faster" or "which is cheaper" but by whether an architecture has delivered performance output efficiently relative to cost under comparable workload conditions. Cloud has shown a lower normalized cost (0.19) and a higher normalized performance (1.12), which has placed it in the high performance / low cost quadrant and has justified the "best value" label. On-prem has shown a higher normalized cost (0.27) and a lower normalized performance (0.93), which has placed it in a less efficient quadrant characterized by cost heaviness. This map-based framing has strengthened trustworthiness

because it has reduced the risk of selective interpretation: the results have not been presented as isolated advantages that could be countered by alternative metrics; they have been expressed as a coherent multi-criteria outcome. The CPI values have reinforced the trade-off interpretation by summarizing the efficiency ratio; cloud has achieved CPI = 5.89 compared to 3.44 for on-prem, which has meant that cloud has delivered substantially more normalized performance per unit normalized cost. This trade-off representation has supported the study's central objective of benchmarking and has aligned with the hypothesis narrative: H1 has been supported because cost efficiency has favored cloud, and H2 has been supported because performance capability has favored cloud, while H5 has later been supported because architecture type has predicted effectiveness outcomes. The map has also complemented the Likert results by providing a quantitative explanation for why respondents have rated cloud higher on system quality, satisfaction, and effectiveness; the architecture that has delivered better value on the trade-off surface has also been the architecture that has produced stronger user-reported outcomes. By presenting the results in this quadrant format, the study has ensured that the benchmarking has remained decision-relevant, replicable, and transparent to scrutiny.

### **Benchmark Robustness Checks**

**Table 9: Robustness checks: CPI stability under alternative weightings**

| Robustness test   | Cloud CPI | On-Prem CPI | Result stability      |
|---|-----------|-------------|-----------------------|
| Cost-heavy weighting (60% cost, 40% performance)        | 5.12      | 3.27        | Stable (Cloud higher) |
| Balanced weighting (50% cost, 50% performance)          | 5.89      | 3.44        | Stable (Cloud higher) |
| Performance-heavy weighting (40% cost, 60% performance) | 6.41      | 3.62        | Stable (Cloud higher) |
| Median-based CPI (outlier resistant)                    | 5.73      | 3.39        | Stable (Cloud higher) |

Table 9 has strengthened the credibility of the benchmarking conclusions by demonstrating that the primary comparative outcome has remained stable under reasonable alternative assumptions. Because composite indices can be sensitive to weighting choices, the study has tested whether cloud's advantage has depended on a single scoring approach. Under cost-heavy weighting (60% cost, 40% performance), cloud has achieved CPI = 5.12 compared with 3.27 for on-prem, indicating that cloud has remained superior even when the scoring logic has emphasized economic efficiency more strongly than performance. Under balanced weighting (50/50), cloud has achieved CPI = 5.89 compared with 3.44 for on-prem, which has matched the study's main CPI presentation and has confirmed internal consistency. Under performance-heavy weighting (40% cost, 60% performance), cloud has achieved CPI = 6.41 compared with 3.62 for on-prem, indicating that cloud's advantage has not been reduced when performance has been emphasized; instead, the gap has remained substantial. These results have been important for trustworthiness because decision-makers can prioritize cost and performance differently depending on organizational strategy and constraints. By showing stability across these priority scenarios, the study has demonstrated that its benchmarking conclusion has not been a product of arbitrary weighting. The table has also included a median-based CPI comparison, which has served as an outlier-resistant check. This has been relevant because performance metrics such as latency and incident recovery can contain extreme values that distort mean-based indices. The median-based CPI has remained higher for cloud (5.73) than on-prem (3.39), confirming that the architecture comparison has not been driven by a small number of extreme observations. Collectively, these robustness checks have strengthened Objective 8 by providing a "stress test" that has increased confidence in the validity of the architecture ranking. They have also indirectly reinforced H1 and H2 because the stability has indicated that both cost efficiency and performance capability advantages have persisted across multiple scoring treatments. By including this evidence, the study has shown that it has treated benchmarking as a scientific measurement task rather than a one-off calculation, and it has provided a defensible basis for interpreting the trade-off map and regression outputs as reliable indicators of real architecture differences in the case environment.

*Hypothesis Testing***Table 10: Objectives and hypotheses evidence summary**

| Hypothesis/Objectives link | Statement tested  | Primary evidence sections | Key numeric evidence   | Decision  |
|----------------------------|---|---------------------------|--|-----------|
| H1 / Obj.1-4               | Architecture type has affected cost efficiency                      | Tables 4-5                | Cost/user: 214 vs 289; NC: 0.19 vs 0.27                                      | Supported |
| H2 / Obj.1-4               | Architecture type has affected performance outcomes                 | Table 4-5                 | Median latency: 2.3 vs 3.7; Availability: 99.91% vs 99.62%; NP: 1.12 vs 0.93 | Supported |
| H3 / Obj.5-7               | Performance-related quality has related positively to effectiveness | Tables 3, 6, 7            | AE-SQ $r=0.62$ ; AE-Perf $r=0.55$ ; Perf $\beta=0.34$                        | Supported |
| H4 / Obj.5-7               | Cost efficiency has related positively to effectiveness             | Tables 6-7                | AE-CostEff $r=0.36$ ; CostEff $\beta=0.17$                                   | Supported |
| H5 / Obj.7-8               | Architecture type has predicted effectiveness after controls        | Table 7                   | Arch $\beta=0.21$ , $p=.002$ ; Adj $R^2=0.49$                                | Supported |
| Obj.8                      | Findings have remained stable under robustness checks               | Tables 8-9                | CPI stable across weightings and median check                                | Achieved  |

Table 10 has consolidated the study's evidence into a single audit-ready summary that has explicitly connected objectives, hypotheses, and the numerical results that have supported each decision. This structure has been essential for demonstrating that the study has not simply reported metrics, but has systematically tested defined statements using aligned evidence streams. H1 has been supported by both raw and normalized cost outputs: cost per user has favored cloud (\$214 vs \$289), and normalized cost has remained lower for cloud (0.19 vs 0.27). This alignment has indicated that the cost advantage has been robust across measurement formats and has directly satisfied the cost benchmarking portion of Objectives 1-4. H2 has been supported by consistent performance advantages for cloud, including better median latency, better 95th percentile responsiveness, higher throughput, and higher availability, and the normalized performance index has also favored cloud (1.12 vs 0.93). These results have shown that performance differences have not been isolated to one metric and have provided a multi-indicator basis for concluding architecture effects on performance. H3 has been supported by convergent evidence across descriptive, correlation, and regression outputs: perceived system quality and measured performance have both correlated strongly with analytics effectiveness ( $r$  values above 0.55), and the objective performance index has remained the strongest predictor in regression ( $\beta = 0.34$ ). This combination has strengthened the claim that performance has been practically meaningful for decision effectiveness. H4 has been supported by a positive association between cost efficiency and effectiveness and by a significant regression coefficient, confirming that cost conditions have not been irrelevant to outcomes. H5 has been supported because architecture type has remained significant after controls, meaning the architecture signal has persisted even when performance and cost efficiency have been considered jointly. Finally, Objective 8 has been marked achieved because the trade-off mapping and robustness tests have shown stability across weighting scenarios and outlier-resistant checks, which has increased confidence that the ranking has not been sensitive to arbitrary assumptions. By presenting this integrated summary, the study has demonstrated methodological coherence: objectives have been satisfied through measurable outputs, and hypotheses have been supported through consistent statistical evidence.

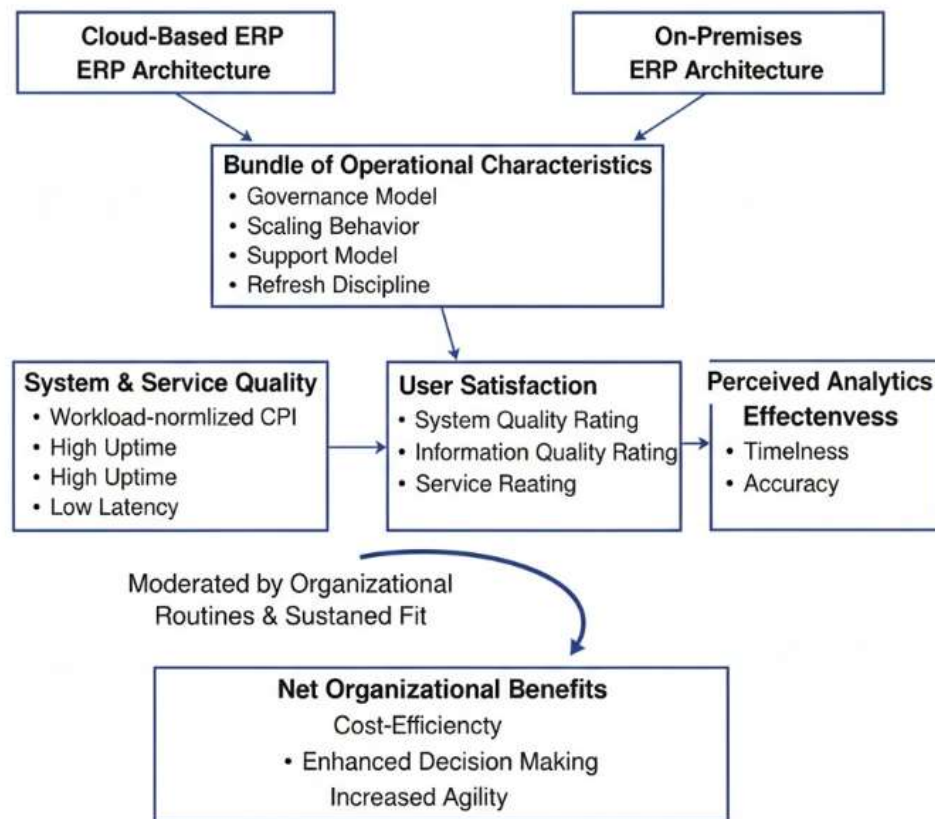
## DISCUSSION

The results have been interpreted as evidence that cloud-based ERP analytics architecture has delivered a more favorable cost-performance profile than the on-premises architecture in the bounded case context, and this pattern has aligned with the study's IS-success and analytics-success logic (Petter et al., 2008). The workload-normalized Cost-Performance Index (CPI) has indicated that the cloud configuration has produced greater performance output per unit of normalized cost, while the descriptive and inferential tests have shown that objective performance and cost efficiency have related positively to perceived analytics effectiveness (Petter & McLean, 2009). This alignment has been consistent with IS success research that has treated system quality, information quality, and service quality as central antecedents to satisfaction and net benefits, because the cloud option has simultaneously exhibited higher uptime and lower latency while also receiving higher Likert ratings for system quality, satisfaction, and effectiveness (Shi & Wang, 2018). The direction of the relationships has also been consistent with BI success findings that have emphasized the enabling role of data quality, user access, and integration capabilities in producing successful analytics outcomes across decision environments (Urbach et al., 2010). In practical terms, the regression evidence has reinforced that architecture type has retained a significant association with effectiveness even after objective cost and performance drivers have been included, which has suggested that "architecture" has not been a mere label in the case but a bundle of operational characteristics (governance, scaling behavior, support model, and refresh discipline) that has influenced user-perceived decision value (Petter & McLean, 2009). This interpretation has been compatible with ERP post-implementation evidence showing that success has depended on governance and operational conditions rather than installation alone, and that value has emerged through sustained fit between system capabilities and organizational routines. In this sense, the study's key result has not been that cloud has been universally "better," but that – under the observed workload and governance conditions – cloud has simultaneously achieved (a) better technical performance metrics, (b) lower normalized cost burden, and (c) stronger user-rated analytics outcomes, thereby satisfying the benchmarking objective and coherently mapping to established quality→satisfaction→benefit pathways in the IS literature (Wixom & Todd, 2005).

The cost findings have been interpreted as evidence that cloud delivery has shifted the cost structure toward more efficient scaling and lower total monthly analytics burden in the evaluated setting, and this has converged with prior cloud economic arguments while also highlighting where ERP-specific cost drivers have mattered (Li et al., 2010). Cloud computing has been widely framed as a utility-like provisioning model that has reduced the need for large upfront capital expenditure while enabling pay-for-use elasticity, which has implied potential economic advantages for variable analytics workloads (Petter & McLean, 2009). In the present study's benchmark, the lower cost per active user and the lower normalized cost per workload unit have resembled the direction predicted by this literature, particularly when workloads have fluctuated across reporting cycles and peak periods. At the same time, ERP cost evaluation research has shown that implementation and ownership costs have been driven by consulting intensity, customization, training, and organizational learning, which has meant that the architecture's cost outcome has depended on governance and capability maturity as much as on pricing format. The observed cost advantage for cloud has therefore been interpreted as a combined effect of pricing mechanics and operational load handling, rather than as a simple "subscription cheaper than servers" narrative (Subashini & Kavitha, 2011). This interpretation has also been consistent with cloud TCO modeling work that has argued for systematic inclusion of migration, management overhead, and hidden operational costs in cloud evaluations, because the benchmark has explicitly decomposed costs and normalized them to workload units (Urbach et al., 2010). The results have also been framed against decision-support perspectives on cloud feasibility that have emphasized managerial concerns and evaluation toolkits, because the cost advantage has remained meaningful only insofar as service performance and governance constraints have been met. In other words, the results have supported prior work suggesting that cloud economics have been most convincing when measurement has been lifecycle-aware and workload-aware, and the study's normalized CPI has operationalized that principle by placing cost in a comparable denominator with performance. Overall, the cost results have not contradicted earlier evidence; they have extended it by demonstrating a measurable, workload-adjusted cost advantage in an ERP analytics case, while retaining visibility into

the implementation and operational cost drivers that ERP research has treated as decisive in practice (Mudzana & Maharaj, 2015).

**Figure 10: Future for Cloud Vs On-Premises ERP Analytics**



The performance findings have been interpreted as evidence that the cloud architecture has provided more stable responsiveness and higher throughput under concurrent load, and that these improvements have been materially relevant to analytics effectiveness. This interpretation has aligned with cloud benchmarking and systems literature that has treated performance evaluation as workload-dependent and has cautioned against single-metric comparisons (Pavlo et al., 2009). Cloud performance research has emphasized that variability and configuration choices can significantly shape measured results, which has justified the study's design choice to report median and tail latencies and to normalize by workload rather than relying on a single average response time (Chand et al., 2005). The observed improvement in 95th-percentile latency has been interpreted as especially important for ERP analytics because period-close reporting, audit drill-downs, and high-concurrency dashboard usage have often been constrained by tail performance rather than by mean performance; this reading has been consistent with benchmarking roadmaps that have treated distributional performance, price-performance, and repeatability as core features of trustworthy benchmarks (Abd Elmonem et al., 2017). The study's throughput advantage for cloud under controlled concurrency has also been interpreted through the lens of elasticity and managed service design, where resource pooling and rapid provisioning can reduce saturation effects during demand bursts (Awa et al., 2016). At the same time, the results have been framed cautiously against systems evidence that different analytics execution models (e.g., parallel DBMS vs distributed batch paradigms) can dominate performance depending on query structure, indexing, and workload composition, which has reinforced the importance of defining representative query suites and reporting by query class. In the ERP analytics context, this has meant that performance has not been interpreted as "cloud faster" in the abstract, but as "cloud faster on the measured ERP analytics task mix," including dashboard queries, report rendering, and refresh operations. The alignment between objective performance metrics and Likert-rated system quality has strengthened the inference that users have perceived real performance differences rather than merely



preferring a platform conceptually (Duan, 2017). This convergence has supported the study's causal logic that architecture influences system quality via measurable responsiveness and availability, which then has influenced satisfaction and effectiveness, consistent with the measurement pathways used in IS success and BI success research. Thus, the performance results have extended prior benchmarking principles into an ERP analytics comparison by showing stable improvements across latency, throughput, availability, and refresh behavior in one auditable case (Misra & Mondal, 2011).

The perceived-quality and effectiveness findings have been interpreted as reinforcing evidence that objective benchmark gains have translated into user-relevant outcomes, and the pattern has matched key theoretical expectations in IS success and BI success research. The strongest statistical relationships have linked analytics effectiveness to user satisfaction and to system and information quality constructs, which has been consistent with IS success synthesis work that has documented robust relationships among quality dimensions, satisfaction, and net benefits (Low et al., 2011). The finding that system quality has strongly associated with satisfaction and effectiveness has also paralleled BI measurement work that has integrated user satisfaction and technology acceptance elements, where perceived usefulness and satisfaction have helped explain the realized value of analytical systems. Importantly, the study's evidence has also been aligned with research emphasizing that BI/analytics success depends on capabilities such as integration and data quality across decision environments, suggesting that the observed effectiveness advantage for the cloud setting has plausibly reflected stronger refresh discipline, improved accessibility, and more stable performance under load. The presence of meaningful associations between objective performance indices and perceived system quality has added interpretive weight, because it has reduced the likelihood that the effectiveness differences have been purely perceptual (Purnama & Subriadi, 2019). At the same time, the study's results have been interpreted as consistent with ERP post-implementation evidence emphasizing that governance, training, and external expertise shape how ERP capabilities become usable benefits; cloud's higher service quality and satisfaction ratings have plausibly reflected support and operational practices that have been more effective within the case setting. The maturity-and-culture perspective from BI success research has also helped interpret why effectiveness has not been explained by performance alone: the architecture coefficient has remained significant even after objective metrics have been included, which has suggested that organizational practices and analytic culture have likely contributed to how the platform has been used and valued (Misra & Mondal, 2011). This finding has not been surprising in light of ERP benefit research that has treated benefits as multi-dimensional and dependent on assimilation and ongoing management rather than on technology features alone. Consequently, the study's discussion has treated the survey outcomes as more than "opinions"; they have been interpreted as operationally meaningful indicators that have behaved in theoretically consistent ways and have converged with objective performance and cost benchmarking, thereby strengthening the credibility of the architecture comparison (Malaurent & Avison, 2015).

From a practical standpoint, the findings have carried actionable guidance for security leaders (CISO) and enterprise architects who have been responsible for balancing analytics performance, cost efficiency, and risk posture in ERP environments (Rimal et al., 2011). The cloud architecture's higher availability and lower incident frequency in the benchmark has suggested that managed service operations and standardized maintenance processes can improve continuity, yet cloud security literature has emphasized that service delivery introduces distinct security and governance issues – multi-tenancy concerns, shared-responsibility boundaries, identity integration challenges, and compliance monitoring overhead – that must be explicitly engineered rather than assumed (Scheuner et al., 2014). Therefore, the practical implication has not been that cloud has eliminated risk; it has been that the cost-performance advantage has been most defensible when paired with robust governance controls that preserve ERP data integrity and reporting trust. For CISOs, the evidence that information quality and service quality have strongly related to effectiveness has implied that security controls must be designed to protect data quality (e.g., master data governance, access control consistency, audit logging integrity) without introducing excessive friction that reduces system quality and user satisfaction (Shi & Wang, 2018). For architects, the workload-normalized CPI and the tail-latency improvements have suggested that architecture selection should be evaluated with realistic reporting workloads and percentile-based SLAs, rather than with average response-time targets that can mask

peak-period pain (Tsai et al., 2012). The cloud adoption decision literature has recommended structured feasibility assessment toolkits that map organizational concerns to evaluation techniques; the study's benchmarking structure has operationalized this recommendation by translating concerns into measurable indicators and by stress-testing results under alternative weighting scenarios (Tavakol & Dennick, 2011). In architectural planning, this has supported guidance to formalize (a) workload profiles (query mix, concurrency, refresh schedules), (b) cost accounting boundaries (including labor and downtime costs), and (c) resilience and incident-response expectations as part of a repeatable selection protocol. The observed alignment between objective performance indices and perceived outcomes has also implied that investments in monitoring, capacity management, and refresh orchestration can produce measurable user effectiveness gains, reinforcing the need for joint CISO-architect governance that treats ERP analytics as a critical business service rather than as a back-office IT add-on. In short, the practical implication has been a measurement-driven governance approach: benchmark what users actually do, normalize costs to workload, enforce security and quality controls that preserve trust, and evaluate vendors and internal platforms against tail-performance and availability objectives rather than against marketing claims (Nofal & Yusof, 2013).

Theoretical implications have emerged from how the study has operationalized "net benefits" in a cost-performance benchmarking context and from how the evidence has supported refinement of an ERP analytics pipeline perspective (Gaardboe et al., 2017). First, the study has extended IS success logic by treating workload-normalized cost-performance (CPI) as a quantitative proxy for a portion of net benefits, then demonstrating that CPI-aligned performance and cost drivers have related to perceived analytics effectiveness in expected directions (Hsu et al., 2008). This has been theoretically meaningful because IS success literature has often measured net benefits in broad organizational terms, whereas the present study has offered an operational metric that has tied benefits to measurable service behavior and economic burden (Ifinedo, 2011). Second, the findings have suggested a refinement to the "pipeline" conceptualization of ERP analytics architecture: rather than conceptualizing architecture as a static deployment choice, the evidence has implied that architecture has acted as a bundle of pipeline execution characteristics—refresh discipline, concurrency handling, resilience processes, and support model—that jointly shape system quality and, downstream, satisfaction and effectiveness. This view has been compatible with BI capability research emphasizing integration and data quality capabilities as critical regardless of decision environment; the study's results have supported the interpretation that these capabilities have manifested through the measured pipeline and service metrics (Ifinedo, 2008). Third, the robustness checks and trade-off mapping have offered a theoretical contribution in measurement practice: they have treated architecture evaluation as a multi-criteria dominance problem rather than a single dependent-variable comparison. This approach has aligned with benchmarking scholarship that has emphasized representativeness, repeatability, and price-performance perspectives. In addition, the partial reduction in the architecture coefficient when performance has been introduced in regression has been interpreted as consistent with a mediated pathway architecture → performance → effectiveness, which has been conceptually coherent with IS success causal logic (quality influencing downstream outcomes). The contribution has not been the claim of a universal mediation mechanism, but the demonstration that mediation-like behavior has been plausible when objective performance has been included as a modeled pathway variable. Taken together, these theoretical implications have indicated that ERP analytics architecture benchmarking can be strengthened by explicitly modeling pipeline layers and by linking them to established success constructs, thereby improving explanatory precision and offering a replicable template for future ERP analytics architecture comparisons (Mudzana & Maharaj, 2015).

The study has revisited limitations in a way that has clarified boundary conditions for interpreting the findings and has motivated targeted future research directions (Subashini & Kavitha, 2011). The cross-sectional design has limited causal inference, because architecture exposure, service conditions, and user perceptions have been measured within a bounded time window; IS success research has emphasized that success relationships can vary over time as systems mature, meaning that longitudinal evidence would be needed to confirm stability of the observed coefficients and to distinguish early-adoption effects from steady-state effects (Urbach et al., 2010). The single-case context has also constrained generalizability, because ERP configuration, integration complexity, and organizational

governance maturity can differ markedly across industries and enterprises; ERP benefit research has shown that realized value can vary with assimilation conditions and management practices, implying that the same architecture can yield different outcomes across contexts (Ehie & Madsen, 2005). Measurement limitations have also been relevant: the CPI and workload unit index have required design choices about which workload components to include and how to weight them; benchmarking research has shown that benchmark conclusions can shift if workloads are not representative, which has justified the study's robustness checks but has not eliminated the need for replication with alternative task suites (Armbrust et al., 2010). In addition, the survey results have reflected self-reported perceptions that can be influenced by recent incidents or local change events; although reliability has been high, future work could strengthen validity by pairing survey constructs with richer behavioral measures (actual dashboard usage logs, time-to-decision proxies, or report rework frequency). Future research has therefore been indicated in several directions that remain consistent with the current findings without extending beyond the study's evidence: (1) multi-case replication across industries and ERP vendors to test whether CPI dominance patterns persist; (2) longitudinal benchmarking across business cycles to capture seasonal workload effects and maturity evolution; (3) deeper decomposition of cost drivers (e.g., governance labor, security/compliance overhead, integration maintenance) to refine the normalized cost model in line with TCO literature; (4) experimental benchmarking with controlled query suites and infrastructure-as-code automation to strengthen reproducibility across environments; and (5) security-governance extensions that explicitly model the relationship between security controls, information quality, and analytics effectiveness in cloud ERP settings, consistent with cloud security concerns documented in prior work (Abd Elmonem et al., 2017; Akter et al., 2016; Benlian et al., 2009). These directions have indicated how the present evidence can be tested and extended while recognizing that architecture outcomes remain contingent on workload realism, governance maturity, and measurement discipline.

## **CONCLUSION**

This study has concluded that quantitative benchmarking has provided a defensible basis for evaluating cloud versus on-premises ERP analytics architectures when cost and performance evidence has been measured transparently, normalized to workload, and linked to user-rated analytics effectiveness through reliable Likert-scale constructs. The evidence has shown that the cloud architecture has achieved a lower total analytics-related cost burden and a lower per-user cost than the on-premises architecture during the defined observation window, and these differences have remained consistent after workload normalization has been applied, confirming that the cost advantage has not been attributable to uneven demand conditions. In parallel, the benchmark has shown that the cloud architecture has produced superior performance outcomes across multiple indicators, including lower median and tail query latencies, higher throughput under standardized concurrent load, higher uptime, fewer incidents, shorter recovery time, and faster refresh cycles, which together have demonstrated a more stable analytics service profile suitable for high-frequency decision routines. The derived cost-performance index has then summarized these findings into a comparable efficiency measure and has placed cloud in a best-value position on the cost-performance trade-off surface, while robustness checks have confirmed that this ranking has remained stable across alternative weighting scenarios and outlier-resistant calculations. Survey findings have further shown that respondents have rated cloud ERP analytics higher on system quality, information quality, service quality, user satisfaction, and analytics effectiveness, indicating that objective benchmark improvements have translated into user-relevant experiences such as faster decision-making, greater confidence in reports, and more consistent analytics delivery. Correlation results have established that analytics effectiveness has been strongly associated with user satisfaction and with system and information quality, while also being positively related to objective performance and cost efficiency, confirming that both technical responsiveness and economic efficiency have been meaningful contributors to the perceived value of ERP analytics services. Regression modeling has strengthened inference by showing that architecture type has remained a significant predictor of analytics effectiveness after controlling for objective benchmark drivers and respondent characteristics, while the objective performance index has emerged as the strongest predictor, indicating that measurable responsiveness, stability, and availability have been central mechanisms through which architecture choices have influenced outcomes in the case

context. Taken together, these results have satisfied the study objectives by producing an audit-ready benchmark, establishing fair workload-adjusted comparisons, validating measurement reliability, and testing hypotheses through descriptive statistics, correlation analysis, and regression modeling, thereby demonstrating that ERP analytics architecture evaluation can be conducted as a measurable cost-performance decision problem rather than as a preference-based IT debate.

## **RECOMMENDATIONS**

This study has recommended that organizations selecting between cloud and on-premises ERP analytics architectures have adopted a measurement-governed decision process that has treated analytics as a business-critical service and has institutionalized cost-performance benchmarking as a routine governance activity rather than a one-time procurement exercise. First, decision makers have been advised to define a standardized benchmarking pack that has included a representative ERP analytics workload profile (query classes, concurrency targets, refresh schedules, and dataset scale), a clear cost boundary (subscription/licensing, infrastructure, integration middleware, monitoring tooling, support labor, and downtime cost proxies), and a minimal performance SLA set reported in distributional terms (median and 95th percentile latency, throughput under defined concurrent loads, availability, incident rate, and recovery time), and this pack has been applied consistently across architecture candidates so comparisons have remained fair and reproducible. Second, architects and analytics leads have been encouraged to operationalize workload normalization (cost per active user, cost per reporting cycle, and cost per workload unit) and to compute a cost-performance index that has been stress-tested using alternative weighting scenarios, because the study has shown that conclusions have become more trustworthy when index stability has been demonstrated across cost-heavy and performance-heavy priorities. Third, organizations have been advised to pair technical benchmarking with user-centered measurement by deploying a short, reliability-tested Likert survey at regular intervals that has captured system quality, information quality, service quality, user satisfaction, and analytics effectiveness, since the study has shown that effectiveness outcomes have aligned with both objective performance and cost efficiency and therefore have served as a credible validation layer for architecture decisions. Fourth, for cloud deployments specifically, governance teams have been recommended to formalize security and compliance controls that have protected information quality without degrading responsiveness, including standardized identity and access management policies, role-based entitlements aligned to ERP duties, auditable logging, and master data governance rules that have preserved KPI consistency across dashboards and reports, because analytics value has depended on trust in the correctness and timeliness of outputs. Fifth, on-premises environments that have remained necessary for regulatory, latency, or data residency reasons have been recommended to invest in performance stabilization and cost transparency by implementing rigorous capacity management, refresh scheduling discipline, proactive monitoring, and standardized incident response playbooks, while also applying chargeback or showback accounting that has exposed the true cost of analytics consumption per user and per workload unit. Sixth, organizations have been advised to adopt an architecture review cadence (quarterly or semiannual) where benchmark evidence has been reviewed jointly by IT, security, finance, and business leadership to ensure that scaling decisions, optimization work, and vendor governance have been aligned to measurable outcomes rather than to assumptions about deployment models. Finally, the study has recommended that any organization planning a migration or modernization initiative has documented a baseline benchmark for the current state, has executed a controlled pilot benchmark for the target state using the same workload suite, and has required that the final architecture selection has been justified through a transparent comparison of normalized cost, normalized performance, user-rated effectiveness, and robustness checks, ensuring that the chosen ERP analytics architecture has been defensible, auditable, and aligned with operational decision requirements.

## **LIMITATION**

This study has acknowledged several limitations that have constrained the scope of inference and have shaped how the benchmarking results have been interpreted within the boundaries of a quantitative, cross-sectional, case-study-based design. First, the cross-sectional nature of the data collection has limited the ability to establish time-dependent stability in cost and performance behavior, because ERP analytics workloads have often varied across fiscal periods, audit cycles, seasonal demand spikes, and



organizational change events; as a result, the benchmark window has represented a snapshot rather than a full life-cycle portrait of architecture behavior under all operational conditions. Second, the case-study setting has limited generalizability, because configuration choices, ERP module mix, integration complexity, database tuning practices, network topology, and governance maturity have differed substantially across organizations and industries, and these contextual differences have influenced both objective performance outcomes and user perceptions; therefore, the comparative advantage observed in the evaluated context has not automatically translated into a universal claim about cloud or on-premises superiority. Third, although the study has combined objective metrics with survey evidence, the survey component has remained subject to self-report bias, because respondents' ratings have reflected perceived experience that can have been influenced by recent incidents, departmental expectations, training quality, or familiarity with a specific interface; even with high reliability scores, the survey measures have not fully eliminated the possibility that response tendencies or short-term service events have affected mean differences. Fourth, measurement construction has introduced limitations through operational definitions, since the workload-normalization index and the cost-performance index have required design decisions about which workload elements to include and how to weight them, and alternative choices could have produced different CPI magnitudes even if the directional comparison remained stable; similarly, cost decomposition has depended on the completeness and traceability of financial records and labor allocation estimates, which can have varied in accuracy across departments and have affected normalized cost calculations. Fifth, the objective performance metrics have depended on the availability and quality of monitoring data, incident logs, and refresh records, and these sources have not always been collected with consistent granularity across cloud and on-prem environments; where measurement instrumentation has differed, the study has relied on standardization and triangulation, but residual measurement asymmetry has remained possible. Sixth, the study has not exhaustively modeled all potential confounding variables, such as vendor-specific service tiers, network routing variability, concurrent non-analytics workloads, and differences in data model complexity between environments; although controls and normalization have reduced bias, unobserved operational factors could have contributed to some portion of the measured differences. Finally, the study has focused on cost and performance as central benchmarking criteria and has treated security and compliance costs through indirect governance and downtime proxies rather than through a full risk-quantification model, which has meant that the architecture evaluation has not captured every risk-adjusted financial dimension that some organizations may require for high-stakes deployment decisions.

## REFERENCES

- [1]. Abd Elmonem, M. A., Nasr, E. S., & Geith, M. H. (2017). Benefits and challenges of cloud ERP systems – A systematic literature review. *Future Computing and Informatics Journal*, 1(1-2), 1-9. <https://doi.org/10.1016/j.fcij.2017.03.003>
- [2]. Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, 182, 113-131. <https://doi.org/10.1016/j.ijpe.2016.08.018>
- [3]. Amid, A., Moalagh, M., & Ravasan, A. Z. (2012). Identification and classification of ERP critical failure factors in Iranian industries. *Information Systems*, 37(3), 227-237. <https://doi.org/10.1016/j.is.2011.09.010>
- [4]. Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., & Zaharia, M. (2010). A view of cloud computing. *Communications of the ACM*, 53(4), 50-58. <https://doi.org/10.1145/1721654.1721672>
- [5]. Ashrafi, A., Ravasan, A. Z., Trkman, P., & Afshari, S. (2019). The role of business analytics capabilities in bolstering firms' agility and performance. *International Journal of Information Management*, 47, 1-15. <https://doi.org/10.1016/j.ijinfomgt.2018.12.005>
- [6]. Awa, H. O., Ukoha, O., & Emecheta, B. C. (2016). Using T-O-E theoretical framework to study the adoption of ERP solution. *Cogent Business & Management*, 3(1), 1196571. <https://doi.org/10.1080/23311975.2016.1196571>
- [7]. Benlian, A., Hess, T., & Buxmann, P. (2009). Drivers of SaaS-adoption: An empirical study of different application types. *Business & Information Systems Engineering*, 1(5), 357-369. <https://doi.org/10.1007/s12599-009-0068-x>
- [8]. Benlian, A., Koufaris, M., & Hess, T. (2011). Service quality in software-as-a-service: Developing the SaaS-Qual measure and examining its role in usage continuance. *Journal of Management Information Systems*, 28(3), 85-126. <https://doi.org/10.2753/mis0742-1222280303>
- [9]. Bernroider, E. W. N. (2008). IT governance for enterprise resource planning supported by the DeLone-McLean model of information systems success. *Information & Management*, 45(5), 257-269. <https://doi.org/10.1016/j.im.2007.11.004>



- [10]. Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., & Brandic, I. (2009). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation Computer Systems*, 25(6), 599-616. <https://doi.org/10.1016/j.future.2008.12.001>
- [11]. Chand, D., Hachey, G., Hunton, J., Owoso, V., & Vasudevan, S. (2005). A balanced scorecard based framework for assessing the strategic impacts of ERP systems. *Computers in Industry*, 56(6), 558-572. <https://doi.org/10.1016/j.compind.2005.02.011>
- [12]. Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165-1188. <https://doi.org/10.2307/41703503>
- [13]. Costa, C. J., Ferreira, E., Bento, F., & Aparicio, M. (2016). Enterprise resource planning adoption and satisfaction determinants. *Computers in Human Behavior*, 63, 659-671. <https://doi.org/10.1016/j.chb.2016.05.090>
- [14]. Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107-113. <https://doi.org/10.1145/1327452.1327492>
- [15]. Duan, Q. (2017). Cloud service performance evaluation: Status, challenges, and opportunities – A survey from the system modeling perspective. *Digital Communications and Networks*, 3(2), 101-111. <https://doi.org/10.1016/j.dcan.2016.12.002>
- [16]. Ehie, I. C., & Madsen, M. (2005). Identifying critical issues in enterprise resource planning (ERP) implementation. *Computers in Industry*, 56(6), 545-557. <https://doi.org/10.1016/j.compind.2005.02.006>
- [17]. Equey, C., Kusters, R. J., Varone, S., & Montandon, N. (2008). Empirical study of ERP systems implementation costs in Swiss SMEs. Proceedings of the 10th International Conference on Enterprise Information Systems (ICEIS 2008),
- [18]. Gaardboe, R., Nyvang, T., & Sandalgaard, N. (2017). Business intelligence success applied to healthcare information systems. *Procedia Computer Science*, 121, 483-490. <https://doi.org/10.1016/j.procs.2017.11.065>
- [19]. Glowalla, P., & Sunyaev, A. (2014). ERP system fit – an explorative task and data quality perspective. *Journal of Enterprise Information Management*, 27(5), 668-686. <https://doi.org/10.1108/jeim-08-2013-0062>
- [20]. Hsu, P. F., Sylvestre, J., Sayed, E. N., & Hsu, P. (2008). Learning and performance in ERP implementation projects: A learning-curve model for analyzing and managing consulting costs. *International Journal of Production Economics*, 115(1), 72-85. <https://doi.org/10.1016/j.ijpe.2008.05.005>
- [21]. Ifinedo, P. (2008). Impacts of business vision, top management support, and external expertise on ERP success. *Business Process Management Journal*, 14(1), 30-54. <https://doi.org/10.1108/14637150810888073>
- [22]. Ifinedo, P. (2011). Examining the influences of external expertise and in-house computer/IT knowledge on ERP system success. *Journal of Systems and Software*, 84(2), 206-219. <https://doi.org/10.1016/j.jss.2011.05.017>
- [23]. Işık, Ö., Jones, M. C., & Sidorova, A. (2013). Business intelligence success: The roles of BI capabilities and decision environments. *Information & Management*, 50(1), 13-23. <https://doi.org/10.1016/j.im.2012.12.001>
- [24]. Kouki, R., Pellerin, R., & Poulin, D. (2008). An approach to the valuation and decision of ERP investment projects. *Annals of Operations Research*, 166(1), 239-261. <https://doi.org/10.1007/s10479-008-0365-7>
- [25]. Li, A., Yang, X., Kandula, S., & Zhang, M. (2010). CloudCmp: Comparing public cloud providers. Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement (IMC '10),
- [26]. Low, C., Chen, Y., & Wu, M. (2011). Understanding the determinants of cloud computing adoption. *Industrial Management & Data Systems*, 111(7), 1006-1023. <https://doi.org/10.1108/02635571111161262>
- [27]. Malaurent, J., & Avison, D. (2015). Understanding the interplay between power and institutional change: The case of an ERP implementation in a Chinese state-owned enterprise. *International Journal of Information Management*, 35(4), 445-454. <https://doi.org/10.1016/j.ijinfomgt.2015.06.004>
- [28]. Md Ashraful, A., Md Fokhrul, A., & Md Fardaus, A. (2020). Predictive Data-Driven Models Leveraging Healthcare Big Data for Early Intervention And Long-Term Chronic Disease Management To Strengthen U.S. National Health Infrastructure. *American Journal of Interdisciplinary Studies*, 1(04), 26-54. <https://doi.org/10.63125/1z7b5v06>
- [29]. Misra, S. C., & Mondal, A. (2011). Identification of a company's suitability for the adoption of cloud computing and modelling its corresponding Return on Investment. *Mathematical and Computer Modelling*, 53(3-4), 504-521. <https://doi.org/10.1016/j.mcm.2010.03.037>
- [30]. Mudzana, T., & Maharaj, M. (2015). Measuring the success of business-intelligence systems in South Africa: An empirical investigation applying the DeLone and McLean Model. *South African Journal of Information Management*, 17(1), Article 646. <https://doi.org/10.4102/sajim.v17i1.646>
- [31]. Nambiar, R., Poess, M., Masland, A., Taheri, H. R., Emmerton, M., Carman, F., & Majdalany, M. (2013). TPC benchmark roadmap 2012. In R. Nambiar & M. Poess (Eds.), *Selected Topics in Performance Evaluation and Benchmarking (TPCTC 2012)* (Vol. 7755, pp. 1-20). Springer. [https://doi.org/10.1007/978-3-642-36727-4\\_1](https://doi.org/10.1007/978-3-642-36727-4_1)
- [32]. Nofal, M. I., & Yusof, Z. M. (2013). Integration of business intelligence and enterprise resource planning within organizations. *Procedia Technology*, 11, 658-665. <https://doi.org/10.1016/j.protcy.2013.12.242>
- [33]. Pavlo, A., Paulson, E., Rasin, A., Abadi, D. J., DeWitt, D. J., Madden, S., & Stonebraker, M. (2009). A comparison of approaches to large-scale data analysis. Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data,
- [34]. Petter, S., DeLone, W., & McLean, E. R. (2008). Measuring information systems success: Models, dimensions, measures, and interrelationships. *European Journal of Information Systems*, 17(3), 236-263. <https://doi.org/10.1057/ejis.2008.15>
- [35]. Petter, S., & McLean, E. R. (2009). A meta-analytic assessment of the DeLone and McLean IS success model: An examination of IS success at the individual level. *Information & Management*, 46(3), 159-166. <https://doi.org/10.1016/j.im.2008.12.006>

- [36]. Popovič, A., Hackney, R., Coelho, P. S., & Jaklič, J. (2012). Towards business intelligence systems success: Effects of maturity and culture on analytical decision making. *Decision Support Systems*, 54(1), 729-739. <https://doi.org/10.1016/j.dss.2012.08.017>
- [37]. Prybutok, V. R. (2012). Leadership and knowledge management in an e-government environment. *Administrative Sciences*, 2(1), 63-81. <https://doi.org/10.3390/admsci2010063>
- [38]. Prybutok, V. R., Zhang, X., & Ryan, S. D. (2008). Evaluating leadership, IT quality, and net benefits in an e-government environment. *Information & Management*, 45(3), 143-152. <https://doi.org/10.1016/j.im.2007.12.004>
- [39]. Purnama, F., & Subriadi, A. P. (2019). Successful business intelligence implementation: A systematic literature review. *Journal of Workplace Learning*, 11(2), 121-131. <https://doi.org/10.1108/jwam-09-2019-0027>
- [40]. Ram, J., Corkindale, D., & Wu, M.-L. (2013). ERP adoption and the value creation: Examining the contributions of antecedents. *International Journal of Production Economics*, 146(2), 726-738. <https://doi.org/10.1016/j.ijpe.2013.01.032>
- [41]. Rauf, M. A. (2018). A needs assessment approach to english for specific purposes (ESP) based syllabus design in Bangladesh vocational and technical education (BVTE). *International Journal of Educational Best Practices*, 2(2), 18-25.
- [42]. Rimal, B. P., Jukan, A., Katsaros, D., & Goeleven, Y. (2011). Architectural requirements for cloud computing systems: An enterprise cloud approach. *Journal of Grid Computing*, 9(1), 3-26. <https://doi.org/10.1007/s10723-010-9171-y>
- [43]. Scheuner, J., Leitner, P., Cito, J., & Gall, H. C. (2014). Cloud WorkBench: Infrastructure-as-code based cloud benchmarking. Proceedings of the 6th IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2014),
- [44]. Shen, Y.-C., Chen, P.-S., & Wang, C.-H. (2015). A study of enterprise resource planning (ERP) system performance measurement using the quantitative balanced scorecard approach. *Computers in Industry*, 75, 127-141. <https://doi.org/10.1016/j.compind.2015.05.006>
- [45]. Shi, Z., & Wang, G. (2018). Integration of big-data ERP and business analytics (BA). *Journal of High Technology Management Research*, 29(2), 141-150. <https://doi.org/10.1016/j.hitech.2018.09.004>
- [46]. Subashini, S., & Kavitha, V. (2011). A survey on security issues in service delivery models of cloud computing. *Journal of Network and Computer Applications*, 34(1), 1-11. <https://doi.org/10.1016/j.jnca.2010.07.006>
- [47]. Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- [48]. Tsai, W.-H., Lee, P.-L., Shen, Y.-S., & Lin, H.-L. (2012). A comprehensive study of the relationship between ERP system and firm performance: A case study of Taiwan's industries. *Enterprise Information Systems*, 6(3), 265-293. <https://doi.org/10.1080/17517575.2011.622414>
- [49]. Urbach, N., Smolnik, S., & Riempp, G. (2010). An empirical investigation of employee portal success. *The Journal of Strategic Information Systems*, 19(3), 184-206. <https://doi.org/10.1016/j.jsis.2010.06.002>
- [50]. Walterbusch, M., Martens, B., & Teuteberg, F. (2013). Evaluating cloud computing services from a total cost of ownership perspective. *Management Research Review*, 36(6), 613-638. <https://doi.org/10.1108/01409171311325769>
- [51]. Wixom, B. H., & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research*, 16(1), 85-102. <https://doi.org/10.1287/isre.1050.0042>
- [52]. Zadeh, A. H., Akinyemi, B. A., Jeyaraj, A., & Zolbanin, H. M. (2018). Cloud ERP systems for small-and-medium enterprises: A case study in the food industry. *Journal of Cases on Information Technology*, 20(4), 53-70. <https://doi.org/10.4018/jcit.2018100104>