CompTIA DataSys+ Certification Exam Objectives

EXAM NUMBER: DS0-002 V1

- Pre-draft Exam Objectives summarize the tasks and skills identified in the Job Task Analysis (JTA) workshop that provide directional information about the upcoming exam version.
- The Draft Exam Objectives will replace the Pre-draft Exam Objectives after approximately two months when the skills have been peer-evaluated and validated through a JTA survey of job role practitioners.
- Pre-draft Exam Objectives may contain typos and errata that will be corrected during the development process.
- CompTIA will not accept feedback on the Pre-draft Exam Objectives document. If errors are found, please wait until the Draft Exam Objectives are posted, and then provide feedback using the Draft Exam Objectives Feedback form.

About the Exam

The CompTIA DataSys+ DS0-002 V1 certification exam will certify that the successful candidate has the knowledge and skills required to:

- Design, deploy, manage, and maintain databases
- Demonstrate skills in data acquisition and integration
- Apply and explain scripting and programming concepts in a database environment
- Adhere to security and business continuity best practices

This is equivalent to 2-3 years of hands-on experience as a database administrator.

These content examples are meant to clarify the exam objectives and should not be construed as a comprehensive listing of all the content of this examination.

EXAM ACCREDITATION

The CompTIA DataSys+ exam is accredited by the ANSI National Accreditation Board (ANAB) to show compliance with the International Organization for Standardization (ISO) 17024 standard and, as such, undergoes regular reviews and updates to the exam objectives.

EXAM DEVELOPMENT

CompTIA exams result from subject matter expert workshops and industry-wide survey results regarding the skills and knowledge required of an IT professional.

CompTIA AUTHORIZED MATERIALS USE POLICY

CompTIA Certifications, LLC is not affiliated with and does not authorize, endorse, or condone utilizing any content provided by unauthorized third-party training sites (aka “brain dumps”). Individuals who utilize such materials in preparation for any CompTIA examination will have their certifications revoked and be suspended from future testing in accordance with the CompTIA Candidate Agreement. In an effort to more clearly communicate CompTIA’s exam policies on use of unauthorized study materials, CompTIA directs all certification candidates to the CompTIA Certification Exam Policies. Please review all CompTIA policies before beginning the study process for any CompTIA exam. Candidates will be required to abide by the CompTIA Candidate Agreement. If a candidate has a question as to whether study materials are considered unauthorized (aka “brain dumps”), they should contact CompTIA at examsecurity@comptia.org to confirm.

PLEASE NOTE

The lists of examples provided in Bulleted format are not exhaustive lists. Other examples of technologies, processes, or tasks pertaining to each objective may also be included on the exam, although not listed or covered in this objectives document. CompTIA is constantly reviewing the content of our exams and updating test questions to be sure our exams are current, and the security of the questions is protected. When necessary, we will publish updated exams based on existing exam objectives. Please know that all related exam preparation materials will still be valid.

TEST DETAILS

Required exam	DataSys+ DS0-002 V1
Number of questions	TBD
Types of questions	Multiple-choice and performance-based
Length of test	TBD
Recommended experience	2-3 years of hands-on experience as a database administrator

EXAM OBJECTIVES (DOMAINS)

The table below lists the domains measured by this examination and the extent to which they are represented.

DOMAIN		PERCENTAGE OF EXAMINATION
1.0	Database Fundamentals	
2.0	Database Deployment	
3.0	Database Management and Maintenance	
4.0	Data and Database Security	
5.0	Business Continuity	
6.0	Data Integration	
Total		**100%**

1.0 Database Fundamentals

1.1 Compare and contrast database and data types.

- Relational vs. non-relational databases
- NoSQL types
  - Document
  - Key-value stores
  - Column-oriented
  - Graph
  - Vector
  - Time series
  - Object-oriented
- Data types
  - Structured
  - Unstructured
  - Semistructured

1.2 Given a scenario, develop, modify, and run SQL code.

- SQL subtypes
  - Data Definition Language (DDL)
  - Data Manipulation Language (DML)
  - Transaction Control Language (TCL)
  - Data Query Language (DQL)
    - Windowing
    - Joins
    - Nested
- Create Read Update Delete (CRUD) principle
- Set-based logic
- SQL programming
  - Triggers
  - Stored procedures
  - Functions

1.3 Compare and contrast scripting methods and environments.

- Script purpose and runtime location
  - Server side
  - Client side
- Languages
  - PowerShell
  - Python
  - Unix shell
  - Perl script
- Command-line scripting
- Integrated development environment (IDE) scripting

1.4 Explain the impact of programming on database performance.

- Object-relational mapping (ORM)
  - Hibernate
  - Entity Framework
  - SQLAlchemy
  - Data build tool (dbt)
  - Ebean

- Measuring impact
  - Review SQL code generated by ORMs
  - Confirm validity of code
  - Determine impact to database server
  - Remediate

2.0 Database Deployment

2.1 Compare and contrast aspects of database planning and operations.

- Requirements gathering
  - Gap analysis
  - Resource projection
  - System specifications
    - Storage considerations
      - Size
      - Speed
      - Type
      - Cloud-based vs. on premises vs. hybrid
    - Number of users
    - Type of users
  - Service-level agreement (SLA)
    - Key performance indicator (KPI)
    - Escalation procedures
    - Reporting
- Phases of deployment
  - Installation and configuration
    - Database prerequisites
    - Provisioning
    - Upgrading
    - Modifying
    - Importing
- Database objectives
  - Types of cloud-hosted environments
    - Platform as a service (PaaS)
    - Software as a service (SaaS)
    - Infrastructure as a service (IaaS)
    - Database as a service (DBaaS)
  - Computational persistency
- Testing
  - Database quality check (columns and tables)
  - Schema compatibility
  - Stress testing
    - Database
    - Application
  - Regression testing
- Database validation
  - Index analysis
  - Data mapping
  - Data values
  - Referential integrity

2.2 Given a scenario, implement techniques related to database design and documentation.

- Database schema
  - Logical
  - Physical
- Design documentation
  - Data dictionary
  - Entity relationship diagram
- Online transaction processing (OLTP) vs. Online analytical processing (OLAP)

- Standard operating procedure (SOP) documentation
  - Organizational compliance documentation
  - Maintenance documentation
  - Third-party compliance documentation
- Data consistency
  - Atomicity Consistency Isolation Durability (ACID)
  - Basically Available, Soft state, Eventual consistent (BASE)

2.3 Explain connectivity concepts related to databases.

- Networking concepts
  - Load balancing
  - Domain name service (DNS)
  - Client/server architecture
    - Firewall and perimeter network considerations
    - Static and dynamic IP addressing
    - Multizone region
  - Ports/protocols

3.0 Database Management and Maintenance

3.1 Explain the purpose of monitoring and reporting for database management and performance.

- System alerts and notifications
  - Growth in size/storage limits
  - Job completion/failure
  - Database backup alerts
- System health
  - Daily usage
  - Baseline configuration
  - Throughput
  - Log files
- Resource utilization
  - OS performance
  - CPU usage
  - Memory
  - Disk input/output operations per second (IOPS)
  - Disk space
  - Data metering
- Deadlock monitoring
- Connections and sessions
  - Concurrent connections
  - Failed/attempted connections

3.2 Explain common database maintenance processes.

- Patch management
- Integrity checks
  - Data corruption checks
- Periodic review of logs
- Performance tuning
  - Index optimization
  - Query optimization
  - Transaction volumes
- Partitioning
- Change management
  - Release schedules
  - Continuous integration/continuous deployment (CI/CD)
  - Change approval
  - Database refresh
  - Version control

3.3 Given a scenario, implement data management tasks.

- Data management
  - Modify data
  - Define data
  - Append columns
  - Create views
  - Create index
  - Create statistics
  - Create data tables
    - Table isolation levels
  - Create data relationships

- Database normalization and denormalization
- Computed columns

4.0 Data and Database Security

4.1 Explain data security concepts.

- Encryption
  - Data in motion
    - Client-side encryption
    - In-transit encryption
  - Data at rest
    - Keep your own key (KYOK)
    - Bring your own key (BYOK)
- Confidential computing
- Data protection
  - Masking
  - Anonymization
  - Suppression
  - Destruction techniques
  - Security audit
    - Expired accounts
    - Connection requests
    - Forensics
- Code auditing
  - SQL code
  - Credential storage checks

4.2 Explain the purpose of governance and regulatory compliance.

- Data loss prevention
- Data retention policies
- Data sovereignty
- Personally identifiable information (PII)/personal health information (PHI)
- Payment Card Industry Data Security Standard (PCI DSS)
- Global regulations
  - General Data Protection Regulation (GDPR)
- Regional regulations
  - Data residency compliance

4.3 Given a scenario, implement policies and best practices related to authentication and authorization.

- Role-based access control (RBAC)
- Attribute-based access control (ABAC)
- Authentication
  - Multifactor
  - Single sign-on (SSO)
  - Kerberos
    - Service Principal Names (SPN)
- Identity and access management (IAM)
  - Federated identity
  - Open Authorization (OAuth)
  - OpenID Connect
- SSL certificates

4.4 Explain the purpose of database security.

- Physical

- Logical
  - Firewall
  - Perimeter network
  - Port security
- Zero Trust architecture
- Attack surface management
  - Vulnerability scans
  - Patching
  - Remediation
- Preventing attacks
  - SQL injection
  - Data poisoning
  - Malware
  - Social engineering

5.0 Business Continuity

5.1 Given a scenario, implement backup and restoration processes.

- Scheduling and automating backups
- Backup types
  - Full
  - Incremental
  - Differential
- Test and restore backups
  - Validate backup hash
  - Restore points
- Database dumping
- Reports and alerts
- Storage location
  - On-site vs. off-site
- Archiving
- Snapshot

5.2 Explain the importance of disaster recovery (DR) and best practices.

- DR planning
  - DR documentation
    - System security plan (SSP)
    - Continuity of operation plan (COOP)
    - Specifications design document (SDD)
- Log shipping
- DR plan testing
- DR metrics
  - Recovery point objective (RPO)
  - Recovery time objective (RTO)
- Failover
- Failback

5.3 Compare and contrast fault tolerance operations.

- High availability
- Risk analysis
  - Single point of failure
- Multizone region (MZR)
- Redundancy
  - Replication
    - Georeplication
  - Mirroring
- Transaction logging

6.0 Data Integration

6.1 Given a scenario, use data acquisition techniques and methods.

- Data classification
- Extract load transform (ELT) and Extract transform load (ETL)
- Connectivity
  - Open database connectivity (ODBC)
  - Java database connectivity (JDBC)
  - Open data protocol (ODP)
  - File transfer protocol (FTP)
  - Network file share (NFS)
  - Secure shell (SSH)
  - Common internet file system (CIFS)
  - Remote procedure call (RPC)
  - Simple object access protocol (SOAP)
  - Application programming interface (API)
- Data sources
  - Streaming vs. non-streaming
  - Scraping
- Data formats
  - JSON
  - XML
  - Flat file

6.2 Given a scenario, troubleshoot common data acquisition issues.

- Data corruption
- Data unavailability
- Data format issues
  - Encoding
- Schema mismatch
- Connection-related issues
  - Permissions
  - Timeouts
- Encryption issues
  - Data encryption keys
  - Key encryption keys
- Infrastructure limitations
  - Platform
  - Software versioning
  - Drivers
- Programming errors
  - Syntax
  - Runtime

6.3 Explain emerging technologies and AI concepts related to data integration.

- Machine learning
  - Data manipulation libraries
    - NumPy
    - pandas
    - scikit-learn
    - Tidyverse
    - MATLAB
- Generative AI

    - Hallucinations
    - Prompt engineering
    - Retrieval augmented generation (RAG)
- Human-in-the-loop
- Virtual data warehouses
- Robotics processing automation (RPA)

CompTIA DataSys+ Acronym List

The following is a list of acronyms that appear on the CompTIA DataSys+ DS0-002 V1 certification exam. Candidates are encouraged to review the complete list and attain a working knowledge of all listed acronyms as part of a comprehensive exam preparation program.

ACRONYM	DEFINITION
ACID	Atomicity, Consistency, Isolation, and Durability
ANSI	American National Standards Institute
CPU	Central Processing Unit
CRUD	Create Read Update Delete
DAS	Direct-attached Storage
DB	Database
DBaaS	Database as a Service
DBMS	Database Management Service
DCL	Data Control Language
DDL	Data Definition Language
DHCP	Dynamic Host Configuration Protocol
DLP	Data Loss Prevention
DML	Data Manipulation Language
DNS	Domain Name Service
DoS	Denial of Service
DR	Disaster Recovery
ERD	Entity Relationship Diagram
FTP	File Transfer Protocol
GDPR	General Data Protection Regulation
IaaS	Infrastructure as a Service
IOPS	Input/Output Per Second
IP	Internet Protocol
LAN	Local Area Network
LDAP	Lightweight Directory Access Protocol
MySQL	My Structured Query Language
NAS	Network-attached Storage
NoSQL	Not Only Structured Query Language
ORM	Object-relational Mapping
OS	Operating System
PaaS	Platform as a Service
PCI DSS	Payment Card Industry Data Security Standard
PHI	Personal Health Information
PHP	Hypertext Preprocessor
PII	Personally Identifiable Information
RAM	Random Access Memory

ACRONYM	DEFINITION
REST	Representational State Transfer
RPO	Recovery Point Objective
RTO	Recovery Time Objective
SaaS	Software as a Service
SAN	Storage Area Network
SOP	Standard Operating Procedure
SQL	Structured Query Language
SSD	Solid-state Drive
SSH	Secure Shell
TCL	Transaction Control Language
TCP/IP	Transmission Control Protocol/Internet Protocol
UML	Unified Modeling Language
VLAN	Virtual LAN
VPC	Virtual Private Cloud
VXLAN	Virtual Extensible LAN

CompTIA DataSys+ Hardware and Software List

CompTIA has included this sample list of hardware and software to assist candidates as they prepare for the DataSys+ DS0-002 V1 certification exam. This list may also be helpful for training companies that wish to create a lab component for their training offering. The Bulleted lists below each topic are sample lists and are not exhaustive.

EQUIPMENT

- Desktop/laptop

SOFTWARE

- Free software/SQL environment to run scripts (e.g. MariaDB, DBeaver, SQL Server Management Studio [SSMS])
- Programming languages to practice connecting to a database (e.g., SQL, Python, PowerShell)
- Text editing software (e.g., Notepad++, Visual Studio code)
- Unified Modeling Language (UML) tools

OTHER

- Sample database (e.g., .csv files, Northwind) to practice imports
- Samples of technical (procedural or descriptive) documentation (e.g., data dictionary, entity relationship diagram [ERD])