

DRAFT — UNOFFICIAL — NOT FOR OPERATIONAL USE

PUBLICATION

EXAM-TM20-PRE



PRE-TEST — SL 2: BUILDER

Maven Smart System (MSS) — USAREUR-AF

HEADQUARTERS
UNITED STATES ARMY EUROPE AND AFRICA
(USAREUR-AF)
Wiesbaden, Germany

DRAFT — NOT FOR OFFICIAL USE. FOR TRAINING PLANNING PURPOSES ONLY.

26 MARCH 2026

DRAFT — UNOFFICIAL — NOT FOR OPERATIONAL USE

PRE-TEST — SL 2: BUILDER

MAVEN SMART SYSTEM (MSS) — USAREUR-AF

Field	Detail
Course	SL 2: Builder
Form	Pre-Test
Level	SL 2 (Intermediate)
Audience	All staff — prerequisite: SL 1
Time Allowed	30 minutes
Passing Score	N/A — diagnostic only

INSTRUCTIONS

This diagnostic assessment establishes your baseline knowledge before training. Your score does not affect course eligibility. Answer honestly — results help the instructor tailor instruction to gaps.

SECTION 1 — MULTIPLE CHOICE

Circle the letter of the best answer. (2 points each)

1. In a relational database, a "primary key" is:

A. A field or combination of fields that uniquely identifies each record in a table
B. The first column in any table
C. The most frequently queried field in a dataset
D. A field that stores the date a record was created

2. When joining two database tables, the join field must:

A. Contain only numeric values
B. Be a date field in at least one of the tables
C. Be the same data type in both tables and represent the same entity
D. Exist in both tables as the first column

3. A pipeline that reads raw source data, transforms it, and writes clean output to a target system is generally called:

A. An ETL (Extract, Transform, Load) process B. A dashboard C. An API endpoint D. A version control workflow

4. In data terms, "null" means:

A. Zero (the numeric value) B. The absence of a value — the field contains nothing C. An empty string with no characters D. A value that failed a type validation check

5. "Type casting" in a data pipeline refers to:

A. Converting a field's data type (e.g., from text to number) B. Assigning roles to columns in a join operation C. Adding a new calculated column to a dataset D. Filtering records based on a field's data type

6. In a data schema, "cardinality" describes:

A. The maximum number of rows a table can hold B. The order in which columns appear in a table C. The relationship between the number of records on each side of a table join (one-to-one, one-to-many, etc.) D. The compression ratio of a stored dataset

7. Access control in a data system is best described as:

A. Defining who can read, write, or modify specific data or system components B. Encrypting data at rest C. Auditing user login history D. Preventing external systems from querying the database

8. A "calculated column" in a data transformation is:

A. A column imported from an external system B. The primary key column of a dataset C. A column derived from arithmetic or logical operations on existing columns D. A column that stores running totals

9. Which of the following is an example of a "filter" step in a data pipeline?

A. Adding the values of two columns together B. Changing a column's name C. Keeping only records where a field value equals "Active" D. Merging two datasets on a common field

10. In project data management, "naming conventions" are important because:

A. They prevent unauthorized users from finding data B. They compress file sizes for faster data transfer C. They are required by Army Regulation for all digital files D. They ensure consistent, discoverable, and interpretable naming across all artifacts

11. A "union" operation in data transformation:

A. Joins two tables side-by-side based on a shared key B. Creates a new column based on a formula C. Filters one table using values from another table D. Stacks two tables with the same schema on top of each other to combine rows

12. In a Workshop-style dashboard, "conditional formatting" refers to:

A. Filtering displayed records based on user input
B. Automatically changing a cell or widget's color/style based on a value condition
C. Requiring user confirmation before displaying sensitive data
D. Configuring which users can see which widgets

13. An "audit trail" in a data system is:

A. A log of all automated pipeline runs
B. A list of approved data exports from the system
C. A summary report of data quality checks
D. A record of who made what changes to data, and when

14. "Data governance" in an Army data platform context most directly refers to:

A. The physical security of servers storing the data
B. The policies, roles, and processes that ensure data is accurate, accessible, and used appropriately
C. The encryption standards applied to data in transit
D. The backup and recovery procedures for system data

15. When a data pipeline "fails," the BEST practice is to:

A. Immediately restart the pipeline without reviewing the error
B. Review the error log, identify the root cause, fix the issue, and then re-run
C. Delete the pipeline and rebuild it from scratch
D. Notify the data owner that the data is permanently unavailable

SECTION 2 — SHORT ANSWER

Answer in 2–5 sentences. (6 points each)

SA-1. Explain the difference between a "one-to-many" and a "many-to-many" relationship in a relational data model. Give a military example of each.

SA-2. A Soldier in your unit submits data in a form and accidentally enters a text value ("N/A") in a field that should contain a numeric mileage reading. What data quality problem does this create, and how should it be handled in a pipeline?

SA-3. Describe the purpose of role-based access control in a collaborative data project. What problems does it prevent?

SA-4. What is a data "schema," and why does schema consistency matter when multiple teams contribute data to the same pipeline?

SA-5. Your S6 asks you to "just dump all the raw data into the dashboard." Explain why raw, untransformed data is typically not suitable for direct use in a commander's dashboard.

SCORING SUMMARY

Section	Questions	Points Each	Total Points
Multiple Choice	15	2	30
Short Answer	5	6	30
Total	—	—	60

Passing: N/A — Pre-test is diagnostic only.

ANSWER KEY — INSTRUCTOR USE ONLY

Do not distribute to students.

Multiple Choice: 1. A — Primary key uniquely identifies each record; it is not necessarily the first column or a date field. 2. C — Join fields must match in data type and represent the same entity for a valid join. 3. A — ETL (Extract, Transform, Load) is the standard term for a source-to-target data pipeline. 4. B — Null means the absence of a value, not zero or empty string. 5. A — Type casting converts a field's data type from one type to another. 6. C — Cardinality describes the numeric relationship between joined tables (1:1, 1:M, M:M). 7. A — Access control defines who can read, write, or modify data/system components. 8. C — A calculated column is derived from operations on existing columns. 9. C — Filtering to keep only "Active" records is a filter step; the other options describe other transform types. 10. D — Naming conventions ensure consistent, discoverable, interpretable artifact naming. 11. D — Union stacks rows from two compatible schemas; join combines columns on a key. 12. B — Conditional formatting changes visual styling based on field value conditions. 13. D — Audit trail records who made what changes to data and when. 14. B — Data governance covers policies, roles, and processes for accurate and appropriate data use. 15. B — Review error log, identify root cause, fix, then re-run is the correct pipeline failure response.

Short Answer Guidance:

SA-1. Full credit: one-to-many = one record in Table A relates to multiple records in Table B (e.g., one unit has many vehicles); many-to-many = multiple records in Table A relate to multiple records in Table B (e.g., many Soldiers can be assigned to many training events). Both examples must be military-relevant for full credit.

SA-2. Full credit: type mismatch problem — text in a numeric field causes type cast errors or null coercion; pipeline should include input validation or a type-check step that either rejects the record, flags it for review, or substitutes a null with documentation. Partial credit (3 pts) for identifying the problem without describing a solution.

SA-3. Full credit: RBAC ensures users can only access and modify data appropriate to their role; it prevents unauthorized edits, accidental data corruption by viewers, and insider data exposure; in a collaborative project, it maintains data integrity while enabling sharing. Partial credit (3 pts) for one valid prevention purpose without broader explanation.

SA-4. Full credit: a schema defines the structure of a dataset (column names, data types, required fields); schema consistency matters because mismatched schemas break joins, cause pipeline failures, and produce incorrect aggregations when multiple teams' data is combined. Partial credit (3 pts) for correct definition without explaining consistency implications.

SA-5. Full credit: raw data typically contains nulls, type errors, duplicate records, inconsistent naming, and non-standardized values; commanders need clean, validated, aggregated, and clearly labeled data; a dashboard built on raw data may display incorrect metrics, crash on type errors, or mislead decision-makers. Partial credit (3 pts) for identifying one data quality issue without connecting to dashboard impact.

USAREUR-AF Operational Data Team TM-20 Pre-Test | Version 1.0 | March 2026

DRAFT