

The logo for Exam Good, with the word 'EXAMGOOD' in a bold, sans-serif font. Each letter is a different color: 'E' is black, 'X' is pink, 'A' is blue, 'M' is green, 'G' is orange, and 'O' is orange. The 'O's are hollow.

# EXAMGOOD

## QUESTION & ANSWER

Exam Good provides update free of charge in one year!

The background of the page is a large, abstract geometric pattern composed of many overlapping triangles in various colors including purple, red, orange, yellow, and blue. The pattern is centered and fills most of the page.

Accurate study guides  
High passing rate!

<http://www.examgood.com>

**Exam** : **000-419**

**Title** : IBM InfoSphere Quality  
Stage v8 Examination

**Version** : DEMO

1. When running Word Investigation, producing a pattern report will help you do what?

- A. Refine a standardization rule set.
- B. Decide which fields to use for blocking fields.
- C. Discover inconsistencies in data type representation.
- D. Identify which patterns should survive.

**Answer: A**

2. In a sample data set of 10,000,000 records, the Social Security number has 90 percent distinct values. To ensure that the Social Security number does not over-discriminate based upon its rare occurrence, which statement is true?

- A. Change the u-probability on the match command screen from .01 to .0000001.
- B. Set special variable handling "CRITICAL".
- C. Set special variable handling "NOFREQ".
- D. Do not use the Social Security number in a match command.

**Answer: C**

3. Click the Exhibit button.

Examine the statistics shown in the exhibit.

How many candidate groupings were processed for the data in Match Pass - CUSTOMER\_ZIP3?

Match Pass - CUSTOMER_NAME_ADDR		
Chart	Type	Value
<input type="checkbox"/>	Data records read	23
<input type="checkbox"/>	Blocks processed	19
<input type="checkbox"/>	OVERFLOW blocks	0
<input type="checkbox"/>	Maximum Data block size (including overflow)	2
<input type="checkbox"/>	Average Data block size (not including overflow)	1
<input checked="" type="checkbox"/>	Pseudo matches	3
<input type="checkbox"/>	Data duplicates	1
<input type="checkbox"/>	EXACT Data duplicates	0
<input checked="" type="checkbox"/>	Clerical pairs	2
<input checked="" type="checkbox"/>	Data residuals (including SKIPS & MISSING)	20

Match Pass - CUSTOMER_ZIP3		
Chart	Type	Value
<input type="checkbox"/>	Data records read	23
<input type="checkbox"/>	Blocks processed	14
<input type="checkbox"/>	OVERFLOW blocks	0
<input type="checkbox"/>	Maximum Data block size (including overflow)	4
<input type="checkbox"/>	Average Data block size (not including overflow)	2
<input checked="" type="checkbox"/>	Pseudo matches	2
<input type="checkbox"/>	Data duplicates	2
<input type="checkbox"/>	EXACT Data duplicates	0
<input checked="" type="checkbox"/>	Clerical pairs	1
<input checked="" type="checkbox"/>	Data residuals (including SKIPS & MISSING)	20

- A. 4
- B. 19
- C. 14
- D. 23

**Answer: C**

4. Which two steps occur when executing a match? (Choose two.)

- A. summing records
- B. isolating a subset of records to process using blocking strategy
- C. incrementing a group of records using hashing strategy
- D. scoring records

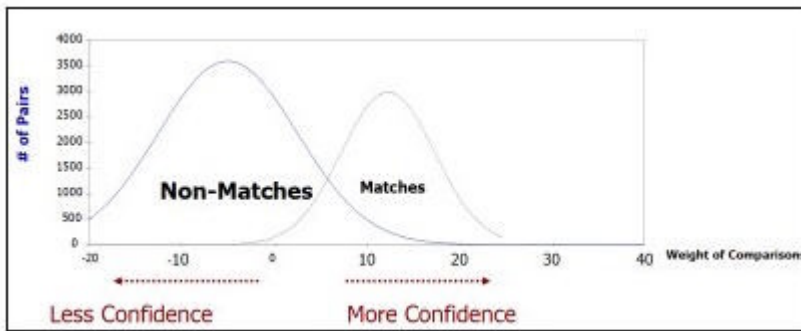
**Answer: BD**

5. Click the Exhibit button.

The exhibit displays an example of distribution of weights for matched and unmatched records.

To further separate the distributions and improve the distinction of matched and unmatched records,

which three statements are true? (Choose three.)



- A. Add disagreement weight overrides which will decrease the match composite score.
- B. Decrease the number of variables in the match to allow remaining match fields to have more discriminating power.
- C. Add more variables to the match to provide additional information in calculating the match composite score.
- D. Lower the clerical cutoff and increase the match cutoff.
- E. Add agreement weight overrides which will increase the composite match score.

**Answer: ACE**

6. You are obtaining product data from a new parts vendor. In your processing design you are using the Reference Match stage to ensure these new parts are valid. This stage can deliver up to six outputs as a result of this match processing. What are three of these outputs? (Choose three.)

- A. Data Duplicate
- B. Data Filtered
- C. Clerical
- D. Sorted
- E. Match

**Answer: ACE**

7. What are the two Match Specification probabilities that can be configured? (Choose two.)

- A. m probability
- B. u probability
- C. r probability

D. pi probability

**Answer: AB**

8. What is a result when you increase the match cut-off setting?

- A. It minimizes the false positives.
- B. It minimizes the false negatives.
- C. It adds agreement weight.
- D. It subtracts agreement weight.

**Answer: A**

9. Some records in the Name column may contain a valid data value such as "Jane Doe" and some records contain a NULL. The Name column is not used as a blocking variable. Which action can be performed in the match specification to penalize records for containing a NULL in the name column?

- A. Increase the Match Cut-off setting.
- B. Change the Data Source Missing Weight.
- C. Add a Weight Override for disagreement weight.
- D. Change the Weight Override Conditional Data setting.

**Answer: B**

10. When testing a reference match specification, which two statements accurately describe what should be performed? (Choose two.)

- A. Match frequencies must be generated to have no overlap.
- B. Match frequencies must be generated for the source file.
- C. Match frequencies must be generated for the reference file.
- D. Match frequencies must be generated to have some overlap.

**Answer: BC**

11. Given the survivorship rule shown below:

COLUMN3: (SIZEOF(TRIM c.COLUMN3) >= 5) AND (SIZEOF(TRIM c.COLUMN1) > 0)

How will this survivorship rule be interpreted?

- A. COLUMN3 of the previous record should be retained if the column contains five or more characters and COLUMN1 has any contents.
- B. COLUMN3 of the previous record should not be retained if the column contains five or more characters and COLUMN1 has any contents.
- C. COLUMN3 of the current record should be retained if the column contains five or more characters and COLUMN1 has any contents.
- D. COLUMN3 of the current record should not be retained if the column contains five or more characters and COLUMN1 has any contents.

**Answer: C**

12. How does the Survive stage relate records to be survived?

- A. by the target columns of each survive rule
- B. by the group identification column
- C. by the column to be analyzed for each rule
- D. by the match type ID column

**Answer: B**

13. For an Unduplicate Match, how does the Match stage determine the master record?

- A. the record within a block with the most unhandled data
- B. the record within a block with the most values populated
- C. the record within a block that matched to itself with the highest weight
- D. the record within a block that occurred last

**Answer: C**

14. Which two statements are true about QualityStage survivorship? (Choose two.)

- A. A rule can have only one target and a TRUE condition expression.
- B. The input record requires a group identifier.
- C. All input records must be combined into one input link.
- D. A Survive job requires that you link a single input stage with multiple output stages.
- E. The output file definition is the same as the input file definition.

**Answer: BC**

15. A survive rule requires which two components? (Choose two.)

- A. master record
- B. column to analyze
- C. match type
- D. technique to apply

**Answer: BD**

16. When multiple rules are specified for the same target in a Survive stage, how is precedence determined?

- A. The first rule to evaluate as true has priority.
- B. The rules appearing later in the list have precedence.
- C. The value in the rule priority option determines the precedence.
- D. The master record of the match set has priority.

**Answer: B**

17. What are the values of a survived record's columns if no input records pass any of the survive rules?

- A. blanks
- B. values of the first record processed
- C. values of the last record processed
- D. values of the master record of the matched set

**Answer: A**

18. You modified a rule set to correct a problem with address standardization. Subsequent job runs show no changes in standardization. Which two techniques will troubleshoot the problem? (Choose two.)

- A. Provision the rule set.
- B. Use a reject file for the input sequential file.
- C. Use a reject file for the output data set.
- D. Use the Rules Analyzer to test the rules logic.



**Answer: AD**

19. You built and tested a match job. Now you want to move it to the production server. Which three components should be moved to the production server? (Choose three.)

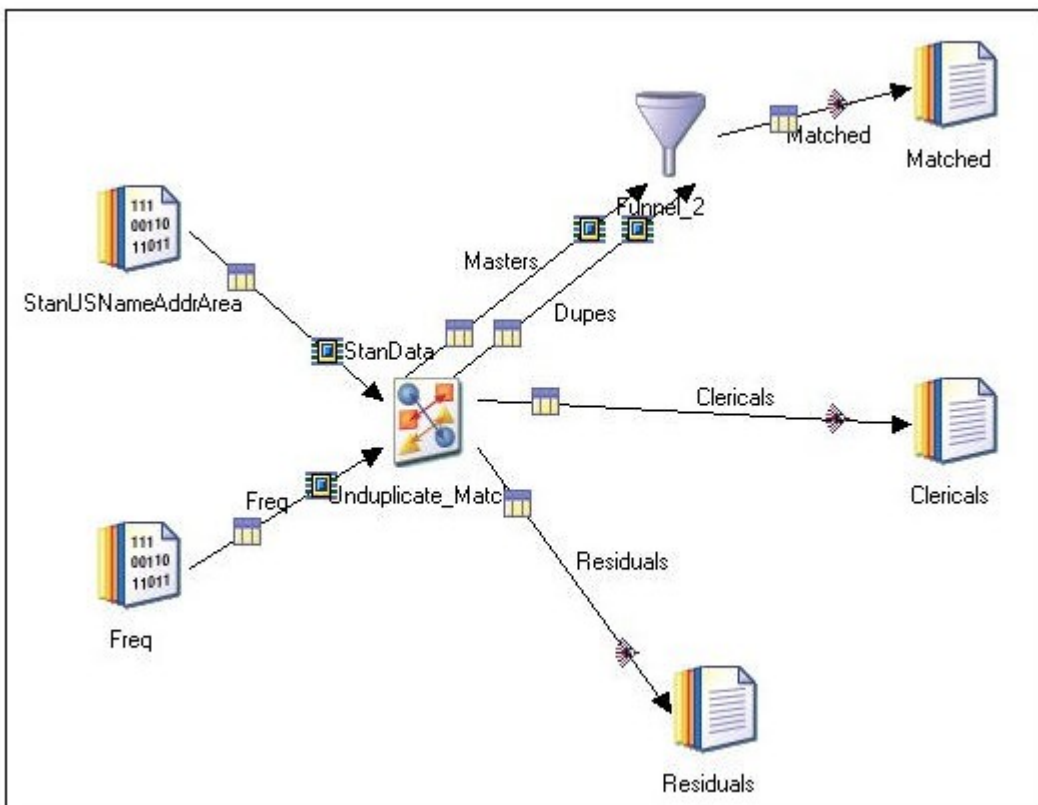
- A. job
- B. match passes
- C. rule sets
- D. match specifications
- E. Match Designer evaluation database

**Answer: ABD**

20. Click the Exhibit button.

In the exhibit you created and ran an unduplication job that produces masters, duplicates, clericals, and residuals. The resulting residuals appear to contain records that should have matched.

Which technique will help you find the problem?



A. Add a reject link to the Matched stage to capture records with invalid data.

- B. Add a reject link to the Unduplicate stage to capture records with invalid data.
- C. Insert a Peek stage between the Unduplicate and Residuals stage to audit records.
- D. Examine the match statistics message for the job in Director and look for block overflow.

**Answer: D**