CERTKILLERS

IBM

C2090-422

IBM Certified Solution Dev - InfoSphere QualityStage v8.5

QUESTION: 112
What stage ensures that each data element has a consistent representation and format?

A. Match
B. Survive
C. Investigate
D. Standardize

Answer: D

QUESTION: 113
Which of the following data problems can be addressed by QualityStage Standardization?

A. Wrong data.
B. Missingdata.
C. Misspelled data.
D. Miscalculated data.

Answer: C

QUESTION: 114
Which two statements describe steps in the Data Quality methodology? (Choose two)

A. End user data entry.
B. Evaluation of results.
C. Analyze the Source data.
D. Address Certification application.
E. Data Management Implementation.

Answer: B, C

QUESTION: 115
Which two steps in Data Quality Methodology often facilitates effective matching? (Choose two)

A. Parsing
B. Surviving
C. Data Mapping
D. Standardization
E. TokenSelection

Answer: A, D

QUESTION: 116
You are using Word Investigation with one of the product provided address rule sets to investigate a business address field called BUSINESS_ADDRESS. Which two of the following will the pattern report show you in regards to the BUSINESS_ADDRESS field? (Choose two)

A. Address type
B. Token Counts
C. Pattern Counts
D. Valid Addresses
E. Sample of fields values

Answer: C, E

QUESTION: 117
Which two report types are generated by a Word Investigation? (Choose two)

A. Token
B. Pattern
C. Analysis
D. Statistical
E. Word List

Answer: A, B

QUESTION: 118
If there are a large number of "@" and "?" symbols in a standardization pattern report, which report will assist with improving classifications?

A. Run Wordinvestigation with Advanced options.
B. Run the Character Discrete report on all columns.
C. Run the Character Discrete report on all unhandled patterns.
D. Run Character Concatenate investigation on selected columns.

Answer: A

QUESTION: 119

In orderto run Word investigations on Latin and Chinese data columns, what two tasks are required? (Choose two)

A. Use Chinese ruleset for Chinese data.
B. Use NLS map Shift_CN in job properties.
C. Install Information Server with NLS option.
D. Set QualityStageclient codepage to hinese".Set QualityStage client codepage to ?hinese".
E. Specify UTF16 as the default character set for the Information Server.

Answer: A, C

QUESTION: 120

You are looking to modify the USADDR address rule set for your companies QualityStage data cleansing application. How can the word investigation reports be used to help in the task of modifying the rule set?

A. Add tokens to allow more duplicates to be identified
B. Correct spelling errors in the reported address fields data
C. Identify new tokens to allow additional address words to be classified
D. Identify new matches to get additional address words to be used in match designer

Answer: C

QUESTION: 121

Click on the exhibit button.

qsInvColumnName	qsInvPattern	qsInvSample	qsInvCount	q
PHONE	(nnn)bnnn-nnnn	(408) 269-0922	18	7
PCONTACT	aaaaa	phone	17	7
EMAIL			11	4
PCONTACT			6	2
PHONE			5	2
EMAIL	aaaaaa@aaaaa.aaa	myself@gmail.com	2	8
EMAIL	aaaaaaa@aaaaaaa.aaa	anymail@hotmail.com	2	8
EMAIL	aaaaaa@aaaaaa.aaa	anders@olsson.com	2	8
EMAIL	aaaaaa@aaaa.aaa	curtis@home.com	1	4
EMAIL	aaa@aaa.aaa.aaa	dfx@usa.ibm.com	1	4
EMAIL	aaa@aaaaaaa.aaa	arc@hotmail.com	1	4
EMAIL	aaaa@aaaaaaa.aaa	afan@hotmail.com	1	4
EMAIL	aaann@aaaaa.aaa	jaz23@yahoo.com	1	4
EMAIL	a@aaaaa.aaa	a@gmail.com	1	4

After examining the results of this investigation as shown in the exhibit, which mask type was used to produce the results?

A. C
B. N
C. T
D. X

QUESTION: 122
If the incoming records in a group have 3 records and 2 of those records contained the same value, which of the following rules would allow that field's value to survive?

A. Most Often (Non-null)
B. Most Frequent (Non-null)
C. Most Frequent (Non-blank)
D. Most Common (Non-blank)

QUESTION: 123
How many operands are present in the following Pattern Action language pattern? ^ | D | ?| T

A. 0
B. 3
C. 4
D. 7

QUESTION: 124
A client has address data with country codes for United States, Japan and 45 other countries for which there are no domain-specific rule sets. Which approach would be the best way to handle processing for this data if the client had only allocated a few weeks to design and code the jobs?

A. Filter all data by country, use the USADDR rule for the US data, the JPADDR for the Japanese data and the AVI stage for all other data.
B. Filter all data by country, use the JPADDR rule for the Japanese data,default all other data to US and use the USADDR rule for US and all other data.
C. Filter all data by country, use the domain-specific rules for the US and Japanese data and create custom rules for only those countries have a large number of records.

D. Filter all data by country, use the domain-specific rules for the US and Japanese data and create custom rules for each country that does not have a domain-specific rule set for address.

Answer: A