1. Which of the following is NOT a true statement about SAS programming?
a) Statements can begin and end anywhere
b) Statements must end with a colon
c) Several statements may be on the same line
d) Blanks separate the components in a SAS program statement

2. Which of the following file extensions is used for a SAS data file?
a) .sas
b) .log
c) .lst
d) .sas7bdat

3. Which statement signals the creation of a new data set in SAS?
a) SET
b) INPUT
c) DATA
d) PROC

4. What is the maximum length of a SAS variable name?
a) 8 characters
b) 16 characters
c) 32 characters
d) 64 characters

5. Which of the following is the default variable type in SAS?
a) Character
b) Numeric
c) Date
d) Boolean

6. What is the purpose of the INPUT statement in SAS?
a) To read data into SAS
b) To write data out of SAS
c) To sort data in SAS
d) To print data in SAS

7. Which input method allows data fields to be defined and read in any order?
a) Freeform list input
b) Compact method
c) Column input

d) Formatted input

8. What is the purpose of the INFORMAT statement in SAS?
a) To specify how to read data
b) To specify how to output data
c) To sort data
d) To format data for printing

9. What is the purpose of a SAS library?
a) To store temporary datasets
b) To store permanent datasets
c) To store SAS procedure output
d) To store SAS logs

10. What does the KEEP statement do in a DATA step?
a) Specifies variables to drop from the output dataset
b) Specifies variables to keep in the output dataset
c) Sorts the data by the specified variables
d) Renames the specified variables

11. What is the purpose of the LABEL statement in SAS?
a) To assign labels to datasets
b) To assign labels to variables
c) To assign formats to variables
d) To assign lengths to variables

12. What does the PROC PRINT procedure do?
a) Sorts a dataset
b) Summarizes data
c) Displays observations from a dataset
d) Merges datasets

13. What is the purpose of the BY statement in SAS procedures?
a) To subset data based on a condition
b) To calculate statistics separately for groups
c) To merge datasets
d) To transpose datasets

14. Which procedure is used to calculate simple descriptive statistics for quantitative data?
a) PROC FREQ

b) PROC MEANS
c) PROC UNIVARIATE
d) PROC SUMMARY

15. Which procedure is used to produce frequency tables for categorical data?
a) PROC FREQ
b) PROC MEANS
c) PROC UNIVARIATE
d) PROC SUMMARY

16. What does the CLASS statement do in PROC MEANS?
a) Specifies the variables to analyze
b) Specifies the statistics to calculate
c) Specifies the grouping variables
d) Specifies the output options

17. What is the purpose of the TABLES statement in PROC FREQ?
a) To specify the variables to analyze
b) To specify the statistics to calculate
c) To specify the output dataset
d) To specify the output options

18. What does the WEIGHT statement do in PROC FREQ?
a) Specifies the variable containing frequency weights
b) Specifies the variable to use for percentages
c) Specifies the variable to use for sorting
d) Specifies the variable containing sampling weights

19. Which SAS procedure is used to create scatter plots?
a) PROC PLOT
b) PROC GPLOT
c) PROC SCATTER
d) PROC CHART

20. What is the purpose of the INFILE statement in SAS?
a) To read data from an external file
b) To write data to an external file
c) To sort an external file
d) To merge external files

21. Which PROC IMPORT option specifies the Excel worksheet to import?
a) SHEET
b) GETNAMES
c) DATAROW
d) GUESSINGROWS

22. What does the DBMS option specify in PROC EXPORT?
a) The file type to export
b) The output file path
c) The variable names
d) The delimiter to use

23. What is the purpose of the FORMAT statement in SAS?
a) To specify the format of input data
b) To specify the format of output data
c) To convert character data to numeric
d) To convert numeric data to character

24. What does the NOTSORTED option do in PROC SORT?
a) Sorts the data in descending order
b) Sorts the data by formatted values
c) Suppresses the sort and just copies the dataset
d) Sorts by a specific locale

25. What is the purpose of the OUT option in PROC MEANS?
a) To create an output dataset with summary statistics
b) To specify the statistics to calculate
c) To specify the class variables
d) To specify the analysis variables

26. What does the TYPE option control in an OUTPUT statement?
a) The statistics to output
b) The name of the output dataset
c) The variables to output statistics for
d) The level of aggregation for the output statistics

27. Which function returns the number of days between two dates in SAS?
a) YRDIF
b) DATDIF
c) DATEDIF

d) DAYDIF

28. What does the INTCK function calculate in SAS?
a) The difference between two dates in years
b) The difference between two dates in days
c) The number of intervals between two dates
d) The number of weekdays between two dates

29. Which function is used to convert a character variable to proper case in SAS?
a) UPCASE
b) LOWCASE
c) PROPCASE
d) NAMECASE

30. What does the SCAN function do in SAS?
a) Searches for a substring within a string
b) Extracts a word from a string
c) Counts the number of words in a string
d) Removes trailing blanks from a string

31. What does the SUBSTR function do in SAS?
a) Searches for a substring within a string
b) Extracts a substring from a string
c) Replaces a substring within a string
d) Compares two substrings

32. Which function returns the position of a substring within a string in SAS?
a) FIND
b) INDEX
c) SEARCH
d) MATCH

33. What is the purpose of the LAG function in SAS?
a) To return the value of the variable from the next observation
b) To return the value of the variable from the previous observation
c) To return the first value of the variable within a BY group
d) To return the last value of the variable within a BY group

34. What does the RETAIN statement do in a DATA step?
a) Keeps variables from being set to missing at the beginning of each iteration

b) Specifies the variables to keep in the output dataset
c) Specifies the variables to drop from the output dataset
d) Renames variables in the output dataset

35. Which function generates random numbers from a normal distribution in SAS?
a) RAND
b) NORMAL
c) RANNOR
d) GAUSS

36. What does the CALL STREAMINIT statement do in SAS?
a) Specifies the random number generator to use
b) Specifies the seed value for random number generation
c) Specifies the stream to use for random number generation
d) Initializes the random number stream

37. What is the purpose of the SAMPSIZE option in PROC SURVEYSELECT?
a) To specify the number of observations to select
b) To specify the sampling rate
c) To specify the sampling method
d) To specify the input dataset

38. Which sampling method selects units with equal probability and without replacement?
a) SRS
b) PPS
c) SYS
d) SEQ

39. What does the STRATA statement do in PROC SURVEYSELECT?
a) Specifies the variables to use for stratified sampling
b) Specifies the sampling method to use within strata
c) Specifies the number of strata to create
d) Specifies the allocation method for the strata

40. What is the purpose of an index variable in a SAS DO loop?
a) To specify the number of iterations
b) To control the sequence of values for each iteration
c) To subset the data within the loop
d) To accumulate a sum within the loop

41. What does the WHILE statement do in a SAS DO loop?
a) Specifies the number of iterations
b) Specifies the sequence of index values
c) Specifies a condition for continuing the loop
d) Specifies a condition for exiting the loop

42. What is the purpose of the BY statement in PROC SQL?
a) To subset the data based on a condition
b) To order the output by the specified variables
c) To group the data by the specified variables
d) To join tables by the specified variables

43. Which PROC SQL statement is used to create a new table?
a) SELECT
b) FROM
c) CREATE TABLE
d) INSERT INTO

44. What does the HAVING clause do in PROC SQL?
a) Selects rows based on a condition
b) Orders the output
c) Groups the output
d) Filters grouped output based on a condition
45. Which join type includes all observations from the left table and matching observations from the right table?
a) Inner join
b) Left join
c) Right join
d) Full join

46. What is the purpose of the UNION operator in PROC SQL?
a) To select distinct rows from the output
b) To concatenate tables vertically
c) To join tables horizontally
d) To intersect tables

47. What does the CORR option do in PROC REG?
a) Requests correlation statistics
b) Requests corrected total sums of squares
c) Requests a correlation matrix of the input data

d) Requests confidence limits for the parameter estimates

48. Which plot is used to assess the normality of residuals in regression?
a) Residual plot
b) Normal probability plot
c) Fit plot
d) Leverage plot

49. What does the I option do in PROC GLM?
a) Requests an inverse matrix
b) Requests I-to-I plots of least squares means
c) Requests confidence intervals for parameter estimates
d) Requests an iteration history

50. Which statistic measures the proportion of variance explained by a regression model?
a) R-square
b) Adjusted R-square
c) Mean square error
d) F-value

1. Which of the following is true about SAS statements?
a) Case sensitivity matters for all SAS statements
b) Blanks are not allowed within a statement
c) Semicolons are optional at the end of statements
d) A RUN statement is required to end each DATA step

2. What is the default length for a character variable in freeform input?
a) 8 characters
b) 16 characters
c) 32 characters
d) 64 characters

3. Which of the following is NOT an advantage of column input?
a) Data fields can be skipped
b) Embedded blanks are allowed in character values
c) Data values must be in fixed column positions
d) Character values can be up to 200 characters long

4. What does a "$" indicate in an INPUT statement?
a) A numeric variable
b) A character variable
c) A date variable
d) A hexadecimal value

5. What is the purpose of a SAS format?
a) To read data into SAS
b) To specify how data is stored
c) To display data in a report
d) To convert data from one type to another

6. What is the difference between a temporary and permanent SAS dataset?
a) Temporary datasets are deleted at the end of the SAS session
b) Permanent datasets cannot be deleted
c) Temporary datasets cannot be used in procedures
d) Permanent datasets can only be accessed in the current SAS session

7. Which function returns the date value from a month, day and year in SAS?
a) MDY
b) DATE
c) TODAY

d) YEAR

8. What does the DIF option calculate in PROC MEANS?
a) The difference between the maximum and minimum values
b) The difference between each value and the mean
c) The difference between each value and the previous value
d) The difference between the upper and lower quartiles

9. What does the NOPRINT option do in PROC FREQ?
a) Suppresses all output
b) Suppresses percentage statistics
c) Suppresses frequency counts
d) Suppresses chi-square statistics

10. Which plot is used to compare the distribution of a variable across levels of a class variable?
a) Histogram
b) Box plot
c) Bar chart
d) Scatter plot

11. What does the DLM option specify in an INFILE statement?
a) The delimiter character
b) The name of the input data file
c) The number of lines to skip before reading data
d) The maximum length of input lines

12. Which option specifies the output dataset name in PROC IMPORT?
a) OUT
b) DATA
c) DBMS
d) REPLACE

13. What does the DELIMITER option specify in PROC EXPORT?
a) The delimiter character for output data
b) The file extension for the output file
c) The output file path
d) The type of file to export

14. What is the purpose of the PUT function in SAS?

a) To write data to an external file
b) To convert numeric values to character
c) To convert character values to numeric
d) To display values in a report

15. Which procedure is used to transpose a dataset in SAS?
a) PROC TRANSPOSE
b) PROC SORT
c) PROC PRINT
d) PROC DATASETS

16. What does the NWAY option do in PROC MEANS?
a) Suppresses the overall statistics
b) Suppresses the statistics for each class level
c) Calculates n-way combinations of class variables
d) Calculates separate statistics for each BY group

17. What is the purpose of the WAYS statement in PROC MEANS?
a) To specify the number of class variables to combine
b) To specify the statistics to calculate
c) To specify the output dataset options
d) To specify the class variables to analyze

18. Which function calculates the number of years between two dates in SAS?
a) YEAR
b) YRDIF
c) AGE
d) DIF

19. What does the SCAN function return in SAS?
a) The position of a substring within a string
b) The length of a string
c) A word from a string
d) The number of words in a string

20. What is the purpose of the FIRST.variable in a BY-group processing DATA step?
a) To identify the first observation in each BY group
b) To identify the first variable in the dataset
c) To subset the first observation from the dataset
d) To rename variables with a common prefix

21. Which function generates uniform random numbers in SAS?
a) UNIFORM
b) RANUNI
c) RAND
d) URAND

22. What does the SEED option specify in a random number function?
a) The multiplicative constant
b) The additive constant
c) The modulus
d) The starting value for the random number stream

23. Which of the following is NOT a valid sampling method in PROC SURVEYSELECT?
a) SRS
b) PPS
c) POISSON
d) MULTISTAGE

24. What does the ALLOC option specify in PROC SURVEYSELECT?
a) The allocation method for stratified sampling
b) The number of replicates to select
c) The sample size allocation for each stratum
d) The input dataset to be sampled

25. What is the purpose of the OUTPUT statement in PROC SURVEYSELECT?
a) To create an output dataset with the selected sample
b) To specify the sampling units
c) To specify the number of replicates to select
d) To specify the sampling method

26. Which of the following increments the index variable by 2 in a SAS DO loop?
a) BY 2
b) STEP 2
c) INCREMENT 2
d) ADD 2

27. What does the UNTIL statement specify in a SAS DO loop?
a) The number of iterations
b) The sequence of index values

c) A condition for continuing the loop
d) A condition for exiting the loop

28. Which PROC SQL clause is used to specify the rows to include in a query?
a) SELECT
b) FROM
c) WHERE
d) HAVING

29. What is the purpose of the DISTINCT keyword in PROC SQL?
a) To remove duplicate rows from the output
b) To select rows based on a condition
c) To group the output by a set of columns
d) To join tables based on a common column

30. Which of the following is an inner join in PROC SQL?
a) JOIN
b) INNER JOIN
c) CROSS JOIN
d) NATURAL JOIN

31. What does the INTERSECT operator do in PROC SQL?
a) Returns rows that are common to both tables
b) Returns rows that are unique to each table
c) Returns all rows from both tables
d) Returns rows from the first table that have a match in the second table

32. Which PROC REG option requests a normal probability plot of residuals?
a) PLOT
b) NORMAL
c) RESIDUAL
d) QQ

33. What does the CLI option request in PROC REG?
a) Confidence limits for individual predicted values
b) Confidence limits for the mean predicted values
c) Confidence limits for the parameter estimates
d) Confidence limits for the R-square statistic

34. Which PROC REG option specifies the significance level for adding variables in a stepwise regression?
a) INCLUDE
b) SLENTRY
c) SLSTAY
d) SIGNIFICANCE

35. What does the ADJUST option specify in a LSMEANS statement in PROC GLM?
a) The method for multiple comparison adjustments
b) The confidence level for the means
c) The type of means to calculate
d) The denominator degrees of freedom method

36. Which of the following statistics is NOT produced by the MEANS statement in PROC GLM by default?
a) Mean
b) Standard deviation
c) Coefficient of variation
d) Minimum

37. What is the purpose of the CONTRAST statement in PROC GLM?
a) To specify a priori hypothesis tests
b) To specify post-hoc multiple comparisons
c) To specify the model effects
d) To specify the error term for tests

38. Which option specifies the significance level for comparison tests in PROC GLM?
a) ALPHA
b) SIGNIFICANT
c) LEVEL
d) THRESHOLD

39. What does the E option request in PROC GLM?
a) The error sums of squares and crossproducts matrix
b) The overall ANOVA table
c) The estimated regression coefficients
d) The Type I sums of squares

40. Which of the following is used to specify an output dataset for score statistics in PROC LOGISTIC?

a) OUTPUT
b) OUT
c) SCORE
d) PREDICTED

41. What does the CLODDS option request in a PROC LOGISTIC MODEL statement?
a) Confidence limits for the odds ratios
b) The odds ratio for a one unit change in the predictors
c) The estimated regression coefficients
d) The probability of the response level

42. Which option specifies the significance level for a likelihood ratio test in PROC LOGISTIC?
a) ALPHA
b) SIGNIFICANT
c) PLEVEL
d) THRESHOLD

43. What does the SCALE option specify in PROC LOGISTIC?
a) The method to correct for overdispersion
b) The type of confidence limits to compute
c) The algorithm to estimate regression coefficients
d) The link function to use

44. Which option specifies the convergence criterion in PROC LOGISTIC?
a) EPSILON
b) CONVERGE
c) TOLERANCE
d) CRITERION

45. What is the purpose of the INFLUENCE option in PROC REG?
a) To request influence diagnostic statistics
b) To specify influential observations to exclude
c) To specify the type of residuals to calculate
d) To request collinearity diagnostic statistics

46. Which PROC REG option requests variance inflation factors?
a) COLLIN
b) TOL
c) VIF

d) COLLINOINT

47. What does the P option request in PROC REG?
a) P-values for the parameter estimates
b) A matrix of predicted values and residuals
c) Partial regression leverage plots
d) Press residuals and diagnostics

48. Which of the following is used to specify the response variable in PROC LOGISTIC?
a) MODEL
b) RESPONSE
c) TARGET
d) DEPENDENT

49. What is the purpose of the UNITS statement in PROC LOGISTIC?
a) To specify the change in predictors for computing odds ratios
b) To specify the scale of the predictors
c) To specify the frequency variable for grouped data
d) To specify a data step for constructing the model

50. Which option plots the influence of individual observations in PROC LOGISTIC?
a) IPLOTS
b) INFLUENCE
c) DFBETAS
d) LEVERAGE

Answers:

1. d
2. a
3. c
4. b
5. c
6. a
7. a
8. a
9. a
10. b
11. a
12. a
13. a
14. b
15. a
16. a
17. a
18. b
19. c
20. a
21. b
22. d
23. c
24. a
25. a
26. a
27. d
28. c
29. a
30. b
31. a
32. b
33. a
34. b
35. a
36. c
37. a
38. a

39. a
40. a
41. b
42. a
43. a
44. a
45. a
46. c
47. b
48. a
49. a
50. a