SAS - Questions and Answers

Dear readers, these SAS Interview Questions have been designed specially to get you acquainted with the nature of questions you may encounter during your interview for the subject of SAS programming. As per my experience good interviewers hardly plan to ask any particular question during your interview, normally questions start with some basic concept of the subject and later they continue based on further discussion and what you answer –

What is the difference between do while and do until?

The DO WHILE expression is evaluated at the top of the DO loop. If the expression is false the first time it is evaluated, then the DO loop never executes. Whereas DO UNTIL executes at least once.

How many data types are there in SAS?

Character, Numeric

How to limit decimal places for variable using PROC MEANS?

By using MAXDEC= option

What is the difference between PROC MEANS and PROC Summary?

The difference between the two procedures is that PROC MEANS produces a report by default. By contrast, to produce a report in PROC SUMMARY, you must include a PRINT option in the PROC SUMMARY statement.

How to specify variables to be processed by the FREQ procedure?

By using TABLES Statement.

Purpose of double trailing @@ in Input Statement ?

The double trailing sign (@@) tells SAS rather than advancing to a new record, hold the current input record for the execution of the next INPUT statement.

How to include or exclude specific variables in a data set?

Using DROP, KEEP Statements and Data set Options.

What are the default statistics that PROC MEANS produce?

It produce the default statistics of MIN, MAX, MEAN and STD DEV.

What is Program Data Vector (PDV)?

PDV is a logical area in the memory.PDV is created followed by the creation of input buffer. SAS builds dataset in the PDV area of memory

What is DATA _NULL_?

It can also be used to write output without creating a dataset.

How to remove unique and duplicate values?

By using PROC SORT with NODUPKEY and NODUP Options.

What is the difference between NODUP and NODUPKEY Options?

The NODUPKEY option removes duplicate observations where value of a variable listed in BY statement is repeated while NODUP option removes duplicate observations where values in all the variables are repeated (identical observations)

How to sort in descending order?

Use DESCENDING keyword in PROC SORT code.

How to convert a numeric variable to a character variable?

By creating a differently-named variable using the PUT function.

How to convert a character variable to a numeric variable?

By creating a differently-named variable using the INPUT function.

What is Difference between SET and MERGE?

SET concatenates the data sets where as MERGE matches the observations of the data sets

What are SYMGET and SYMPUT?

SYMPUT puts the value from a dataset into a macro variable where as SYMGET gets the value from the macro variable to the dataset.

Which date function advances a date, time or datetime value by a given interval?

INTNX function advances a date, time, or datetime value by a given interval, and returns a date, time, or datetime value

what is the purpose would you use the RETAIN statement?

A RETAIN statement tells SAS not to set variables to missing when going from the current iteration of the DATA step to the next. Instead, SAS retains the values.

Difference between %EVAL and %SYSEVALF

%EVAL cannot perform arithmetic calculations with operands that have the floating point values. It is when the %SYSEVALF function comes into picture.

what Proc glm does?

It performs simple and multiple regression, analysis of variance (ANOVAL), analysis of covariance, multivariate analysis of variance and repeated measure analysis of variance.

what is SAS informats?

SAS INFORMATS are used to read, or input data from external files known as Flat Files ASCII files, text files or sequential files).The informat will tell SAS on how to read data into SAS variables.

What are the parameters of Scan function?

Scan(argument,n,delimiters)

Name few SAS functions?

Scan, Substr, trim, Catx, Index, tranwrd, find, Sum.

What is the function of tranwrd function?

TRANWRD function replaces or removes all occurrences of a pattern of characters within a character string.

What is Difference between INPUT and INFILE?

The INFILE statement is used to identify an external file while the INPUT statment is used to describe your variables.

Difference between Missover and Truncover

When the MISSOVER option is used on the INFILE statement, the INPUT statement does not jump to the next line when reading a short line. Instead, MISSOVER sets variables to missing

How to print observations 4 through 8 from a data set?

Using the FIRSTOBS = and OBS = statements.

What is the difference between '+' operator and SUM function?

SUM function returns the sum of non-missing arguments whereas "+" operator returns a missing value if any of the arguments are missing.

What SUBSTR function does?

The SUBSTR function is used to extract substring from a character variable.

What is difference between CEIL and FLOOR functions?

The ceil function returns the smallest integer greater than/equal to the argument whereas the floor returns the greatest integer less than/equal to the argument.

What is the Difference between SCAN and SUBSTR?

SCAN extracts words within a value that is marked by delimiters. SUBSTR extracts a portion of the value by stating the specific location. It is best used when we know the exact position of the sub string to extract from a character value.

How to count unique values by a grouping variable?

Uisng PROC SQL with COUNT(DISTINCT variable_name) to determine the number of unique values for a column.

How to debug SAS Macros

There are some system options that can be used to debug SAS Macros:MPRINT, MLOGIC, SYMBOLGEN

How to save log in an external file?

Use PROC PRINTTO

How Data Step Merge and PROC SQL handle many-to-many relationship?

Data Step MERGE does not create a cartesian product incase of a many-to-many relationship. Whereas, Proc SQL produces a cartesian product.

What is the smallest length for a numeric and character variable respectively ?

2 bytes and 1 byte.

What is the difference between SAS PROCs and the SAS DATA STEP.

Procs are sub-routines with a specific purpose in mind and the data step is designed to read in and manipulate data.

How can you write a SAS data set to a comma delimited file ?

PUT (formatted) statement in data step.

Which SAS statement does not perform automatic conversions in comparisons?

where statement

Difference between Input and Put function?

Input function – Character to numeric conversion- Input(source,informat) and put function – Numeric to character conversion- put(source,format)

If a variable contains letters or special characters, can it be numeric data type?

No, it must be character data type.

What can be the size of largest dataset in SAS?

The number of observations is limited only by computer's capacity to handle and store them.

How to limit decimal places for variable using PROC MEANS?

By using MAXDEC= option

What is the difference between CLASS statement and BY statement in proc means?

BY processing requires that your data already be sorted or indexed in the order of the BY variables.

How to specify variables to be processed by the FREQ procedure?

By using TABLES Statement.

What would be the denominator value used by the mean function if two out of seven arguments are missing?

Five

What is the difference Between Proc Means and Proc Summary?

proc means will give descreptive statitstics. By default it will give output in output window.but proc summary will not give output as default.we need give an option print then only it will give the output.

HOW TO TEST THE DEBUGGING IN SAS?

Use the Debug clause after '/' in the data statement.

If reading a variable length file with fixed input, how would you prevent SAS from reading the next record if the last variable didn?t have a value?

Using the MISSOVER keyword.