



SAS Base Interview Questions

First EDITION: Contain 165+ Questions

About book

SAS® is one of the fastest growing and matured software solutions for the analytics worlds and recent development in the Machine Learning and Artificial intelligence made this SAS software even more useful and well-integrated with BigData computing world. It has its own programming languages which is popularly known as Base SAS and if you want to learn and become expert for the SAS then you must learn this SAS Base programming. In this book we are covering around 165 SAS Base interview questions and answers which are popularly asked in the interview and must aware all this concept covered. In this book we are not covering advanced concepts like Machine Learning, Data science, Artificial intelligence, Big Data etc., there would be separate book launched for the same. This book also helps for the learners who are preparing for the SAS certification like A00-215, A00-231 & A00-232 global SAS certification which include both multiple choice as well as project-based questions and answers. However, for complete questions and answer please visit our website and you can get the same questions and answer in video cum audio book. You must go through this Question and Answer before your real SAS interview questions and keep this book handy if you are working or plan to work in the SAS world. On regular basis we would be updating this book based on the learners feedback and more interview questions would be added, hence it is always recommended that you have access to the latest edition of the book.

Feedback

This is a full-length book from <http://hadoopexam.com> and we love the feedback so that we can improve the quality of the book. Please send your feedback on hadoopexam@gmail.com or admin@hadoopexam.com

Even you want to share your experience and story with the preparation of the real exam certification please share the same. It would help other candidates as well.

Restrictions

Entire content of this book is owned by HadoopExam.com and before using it or publishing anywhere else either digitally on web or printing and distribution require prior written permission from HadoopExam.com. You **cannot** use the code or exercises from this book in your software development or in your software product (commercial as well as open source) and you need to take prior written permission to use the same.

Copyright© Material

This book contents are copyright material and it is hard work and many years of experience working with disruptive technologies, which helps in producing this material. All rights are reserved on the material published in this book. You are not allowed to any part of this material to be reproduced, stored in a retrieval system, and must not be transmitted in any form or by any means, without the prior written permission of the author and publisher, except in the case of brief quotations embedded in critical articles or

online and off-line reviews. Wherever, you use contents make sure full detail of the book is mentioned.

Author had tried as much as his capacity in preparing of this book so that accuracy can be maintained in the presented material. The material sold using this book does not have any warranty or guaranty either express or implied. Neither of the author, publisher, dealer and distributors will be held liable and responsible (explicit/implicit these all parties mentioned are not liable and responsible) for any damages caused or alleged to be caused directly or indirectly by this book. You should note, this material is part of your learning process and as time passes material can be outdated and you should wait or look for that latest material.

Author and publisher have endeavored to provide trademark information about all of the companies and products mentioned in this book. However, we cannot guarantee the accuracy of this information.

Disclaimer:

1. Hortonworks® is a registered trademark of Hortonworks.
2. Cloudera® is a registered trademark of Cloudera Inc
3. Azure® is a registered trademark of Microsoft Inc.
4. Oracle®, Java® are registered trademark of Oracle Inc
5. SAS® is a registered trademark of SAS Inc
6. IBM® is a registered trademark of IBM Inc
7. DataStax® is a registered trademark of DataStax
8. MapR® is a registered trademark of MapR Inc.
9. Apache® is a registered trademark of Apache Foundation

10. Databricks® is a registered trademark of Databricks Inc

Publication Information

First Version Published: Dec 2019

Edition: 1.0

Piracy

We highly discourage the piracy of copyright material especially it happened online on the internet. Piracy causes the damages to all, first of all it damages yourself by not honestly using the correct material, generally pirated material is edited and wrong information is presented which can make big damage as part of your learning process. As well as when you become author and honestly write similar material, piracy will damage your material with the same extent or more. Hence, don't encourage piracy. If piracy is reduced cost of material will automatically decreases. It also makes damages to author, publisher, dealer and distributors. If you come across any illegal copies of this works in any form on the Internet, then please share the detail with URL, location or website name immediately on email id hadoopexam@gmail.com we really appreciate your help in protecting author's hard work and also help in reducing the cost of material.

Author/Trainer required

Corporate Trainer: We have many requirements, where our corporate partners need their team to be trained on particular skill sets. If you are already providing corporate trainings for any skills set, then please become our onsite training partner and fill in the form mentioned above and our respective team will contact you soon. You will get very good revenue for sure. However, what we want, you must be able to train our corporate partner resources. What matters to us? Your proficiency in a particular domain/skill and good oral communication skills. You must be able to accessible to learners as well.

Online Trainer: If you are a working professional and master or proficient in any particular skills and feel that, you are capable of giving online virtual trainings e.g. 2 hrs a day until course contents are completed. Please send an email at admin@hadoopexam.com . You will get a very good revenue share for sure. What matters to us? Your proficiency in a particular domain/skill and good oral communication skills. It will certainly not impact your daily work.

Self-Paced Trainings: Ok, you want to work as per your comfortable time and at the same time sharpen your skills. You can consider this option. You can create self-paced trainings on particular domain/skills. Please send an email at admin@hadoopexam.com with us as soon as possible. Before somebody else connect with us for the same skill set. Your commitment is very important for us. We respect your work and we will not sell your work in just \$10 or less to acquire more resources. As we know, it takes a good amount

of time and you will provide quality material, so we charge reasonable on that so, you will feel motivated with your work and effort. We respect you and your skill.

Certification Material: You may be already certified professional or preparing for particular certification in a specific domain/skill. So why not use this to make money as well as sharing your effort with other learners globally. Please connect with us by filling form or send email at admin@hadoopexam.com and our respective team will contact you soon.

Author: Yes, we are also looking for authors. Who can write books on a particular technology and what you can get certainly a very good revenue sharing and you can bring the same on your resume or linked in profile to show your excellence? Yes, we are not in need of very good oral communication skills, but good writing skill. However, team will also help you to get work done. Author can be more than one for a particular book. However, we wanted you to be in long relationship. So that you don't just write a single e book, but can create an entire series for a particular domain or skill. Good royalty for sure...

Trending Skills (Not limited these):

Hadoop Spark AWS Cloud Azure Cloud Google Cloud	EMC NetApp VMWare CISCO HP	Adobe Alfresco Apple AppSense AutoDesk	Data Analysis Django Docker Drupal Graphics	Infrastructre Automation Internet of Things (IOT) ISO Development Java Java Script
JQuery Kali Linux Laravel Linux Machine Learning	Mobile Application Development NodeJS Android Angular JS Arduino	IBM Watson IBM BPM WebMethod Gemfire Liferay	Scala Python Java SQL/PLSQL Ruby	SAP SAS Salesforce Oracle Cloud Redhat

Audio Cum Video Book

Most of our interview preparation material comes in these two formats

- [eBook \(Online PDF access\)](#)
- [Video cum Audio Book \(Listen & Watch anywhere\)](#)

Premium & Pro subscription

We have two popular subscription model which you can use to get access all the online available material from the HadoopExam.com and the duration for those certifications are [Quarterly](#), [Semi-Annually](#), [Premium Annually](#) and [Pro Subscriptions](#).

Currently we have around 50,000+ subscriptions across all products and we regularly update our products and during your active subscription you get access to all this material.

Question-1: What is the importance of the Base SAS?

Answer: Base is the core of SAS software, most of the SAS other component depend on the Base SAS.

Question-2: What all you can do using the Base SAS?

Answer: Using the Base SAS we can accomplish the following activities.

- You can manage the data in the SAS software
- You can do data analysis and report generations.

Question-3: What is the SAS Data set?

Answer: All the data in the SAS is organized in the rectangular form which is also known as table and this table is called SAS Data set. Table contains your actual data.

Question-4: Can you please explain what is the variable in the SAS Data set?

Answer: Variables are the column header in the SAS Dataset. Let's see the below example of the SAS Dataset of HadoopExam sample learner.

ID	Name	CourseName	Fee
1001	Amit	Hadoop	8000
1002	Rakesh	SAS	7000
1004	Nutan	Spark	7000
1003	John	Python	6000
1005	Venkat	Scala	8000

In this data all the column 4 header (ID, Name, CourseName & Fee) are known as variable. Variable is also known as features.

Question-5: What is the observation in the SAS Dataset?

Answer: As you can see in the previous data, each record in the table is known as observation. There are currently 5 observations.

Question-6: What is a Data value?

Answer: Each separate information for a variable and an entity is known as data value.

Question-7: Can you please explain, what is an entity and a data value?

Answer: As you can see in the below SAS Dataset, it has detail about the 5 different persons.

FirName	LastName	Gender	Height	Weight
Amit	Jain	Male	162	58
Rakesh	Gupta	Male	159	57
Nutan	Jain	Female	148	51
John	Fernandes	Male	178	78
Venkat	Upala	Male	167	71

Here, each person is known as an entity. Hence, this dataset represents the 5 entities. Where each entity has some characteristics, which varies from entity to entity (person to person) for example their name, height etc. Hence, the characteristics or the features that varies needs to be

represented by a variable. And each individual entity has some value for each variable like Amit has weight 58 kg. Hence, here 58 and Amit all are the Data values. So, an observation contains the Data value for one entity. In the given Dataset we can say

- Total 5 Entity or 5 observations
- Total 5 variables or features
- Total 25 Data values.

Question-8: How do you create a new Data set in the SAS?

Answer: In the SAS we use the DATA step to create a new SAS Data set.

Question-9: Please show me an example, how do you create a SAS Data set?

Answer: To create a SAS Dataset we need to write a program. Let's see from the below pseudo code.

```
data he_learner;  
    input ID 1-4 Name $ 6-12 CourseName $ 13-18 Fee $;  
    datalines;  
    1001  Amit   Hadoop8000  
    1002  Rakesh SAS    7000  
;  
run;
```

The entire program above represents a SAS Data step. Which is creating the new Data set named "he_learner" using the data provided below datalines.

Question-10: Please tell me what all common elements you know about the Base SAS programming?

Answer: Common programming elements in the SAS Base are statements, expressions, functions, call routines, options, formats, and informats.

Question-11: What all are the basic rules for SAS statements?

Answer: Lets see the few rules below

- SAS statement would always end with the semicolon.
- SAS statement can be written using lowercase, uppercase or mixture of two.
- You can write more than one statement in a single line.
- It is not necessary that your SAS statement can start from the first column of the line.
- Your SAS statement can be written in more than one line, as soon as it found semicolon then statement ends.

Question-12: What all are the rules for the SAS names?

Answer: SAS could have various names like

- Name of the variable
- Name of the Data sets

So, each name in the SAS has to follow some rules which are as below.

- Name must have length between 1 to 32 character.
- Name's first character can be either “_” underscore or any other letter.
- From 2nd character to 32, you can even have the number e.g. “_he_2_data”

- You can not have blank in the names.

Question-13: Is case sensitivity matters for the variable name and how internally they are represented?

Answer: In case of SAS variable name, you can use mix of uppercase or lower case letter. But internally it does not matter whether you use "CourseName", "COURSENAME" or "coursename" internally all are the same and would be referred to the same variable.

Question-14: What is the SAS Procedure?

Answer: SAS has built in programs known as SAS procedures. Using the SAS procedure, you can do data analysis for the data stored in the SAS Data sets. And generate reports, which save your lot of effort by not writing entire procedure your own.

Question-15: What is the purpose of PRINT procedure?

Answer: PRINT procedure as name suggests display the Data values for the Dataset variables in simple organized form.

Question-16: What is the difference between procedure and proc step?

Answer: Proc step which starts with the PROC (procedure) and ends with the run statement or (if next PROC or Data step found), is called the proc step. Hence, Proc step is a collection of statements which must include the procedure in it.

Question-17: Then what is the proc statement?

Answer: In the proc step we write which procedure to use and end it with the semicolon (remember each SAS statement must end with semicolon) is known as Proc statement. See the below example for entire Proc step comprising proc statement in the first line

```
proc print data=he_learner;  
    title 'HadoopExam Learner List'  
run;
```

All 3 lines are part of PROC step and the first line are known as proc statement. And the 'PRINT' is a procedure, which is in-built in SAS to print the content of the SAS Data set.

Question-18: What is the purpose of the "RUN" statement?

Answer: A RUN statement in the SAS base program indicates that all the preceding statements are ready to be executed.

Question-19: What all are the possible output from the SAS program?

Answer: A SAS program can generate following types of the output

- A new SAS Data set
- SAS Log
- Report or new Data listing
- Catalog files
- External data files
- Creating new entries in external databases like RDBMS (Oracle, MySQL, SQL Server, excel etc.)

Question-20: What all information and data can be stored by the SAS Data sets?

Answer: As we have seen previously SAS data set can store variables and observations. And few other things also it stores like descriptive information about the Data.

Question-21: What is in the descriptive information?

Answer: Descriptive information is an information about the SAS Data sets, which include the

- name for all the variables
- How these variables are arranged (Order of the variables)
- How many total numbers of the observations in the SAS Dataset?
- When this Dataset was created?

Question-22: What do you mean by the SAS log?

Answer: Whenever you run the SAS program it generates some messages about the SAS program execution, and this has quite useful information. And these logs can be created on the

- Disk
- Display in the monitor
- Hardcopy listing

Question-23: What are the catalogs, can you give some examples?

Answer: Catalog contains the information that can not be easily represented in the tabular format. You can see below are the examples

- Function key settings (e.g. F9, F11, F10 etc.)

- Letters produced by SAS/FSP software
- Displays that are produced by SAS/GRAPH software

Question-24: Which SAS software you can use to update or insert the entries in the external database?

Answer: SAS/ACCESS software

Question-25: What is the ODS?

Answer: In SAS ODS is a short form of Output delivery system, which helps in producing the output in variety of formats like

- As an HTML file
- Traditional SAS listing (monospace)
- A PostScript files
- Generating an RTF (Rich Text file, can be opened in MS office)
- Generating output Dataset

Question-26: What is the table definition in the ODS?

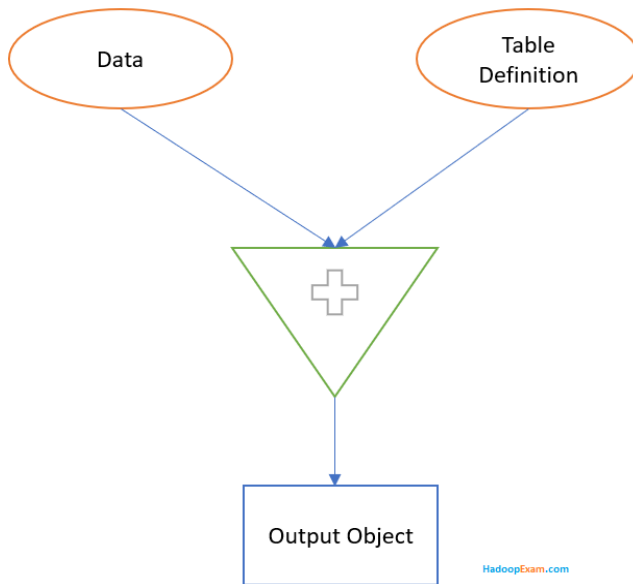
Answer: Table definition provides the instructions that describe how to format the data which may include following

- In what order the column should be generated
- How to format the data
- Font sizes and font faces etc.

Question-27: What is the “Output object” in ODS?

Answer: ODS uses the table definitions (data formatting) and data which is combined by the ODS and produce the output object. An output object contains the information like name, a label and a path. It is not necessary that you always

provide table definition (formats), its optional. You can see in picture as below.



Question-28: What is ODS destinations?

Answer: Using the ODS destination it is specified what kind of output needs to be generated e.g. RTF, HTML or Printer

Question-29: What is the non-interactive mode to run the SAS program?

Answer: You write your entire program in a file and save it which include your SAS statements and Data step, proc step etc. and then you submit that file to be executed by operating environment immediately. As soon as this program execution start you can not interact the session in which the program is running. Once the program finishes it generates

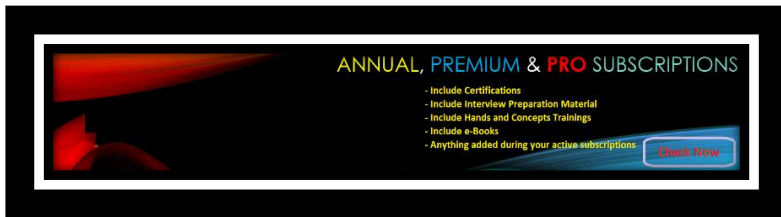
the logs, errors and output at predefined destinations in operative environment.

Thanks for considering SAS A00-215 Certification Preparation Material from <http://hadoopexam.com> . You can get access to all the 150+ questions and answer for your real exam preparation from below link. Where you would find the same quality of questions and correct answers with the detailed explanation. Below are most popular SAS Certification preparation material and we are continuously updating and adding more questions.



You can check all the available products across the website and have Annual Pro subscription which includes

- Trainings
- Certification Preparation
- Interview Preparation material
- Interview Preparation Audio and Video Books
- eBooks on various subjects
- Project Implementations
- Anything newly added during your active subscription also you can get.



Please join this google group to get notified whenever new questions are added

JOIN GOOGLE GROUP TO RECEIVE UPDATE NOTIFICATIONS

Question-30: What is the major difference between batch mode and interactive mode?

Answer: In case of interactive mode program runs immediately and occupies your current workstation session and you can not work until the program runs. On the other hand, in the batch mode once you submit the program you can start working on the other task in the workstation and operating environment run the program. Similarly, once the

program finishes you can look or check the logs in predefined destinations.

The best feature of the batch mode is that batch execution is completely separate from the other activities at your workstation. You can not see the program while it is running, and you can't make any correction while it is running. If you want to edit the program, edit it and create a new batch and re-submit it. This is the good approach for the large programs as well as the one which takes more time.

Question-31: Can you please explain again the Proc step?

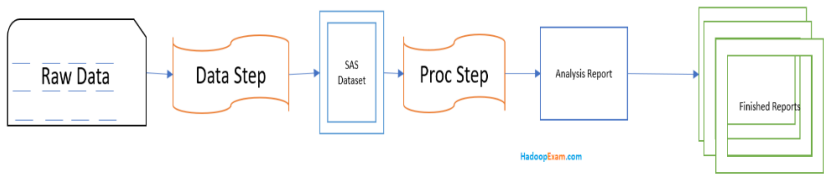
Answer: Lets have basic syntax for the PROC step

```
PROC <Proc_Name> Data = <Data_Set_name>
```

Here, both PROC and DATA are the SAS keyword. PROC statement needs a procedure name e.g. PRINT and the Data specifies the dataset on which procedure would be applied. However, "DATA=" is optional, if you don't provide this it would take most recently generated Dataset in the same session.

Question-32: Can you please explain the Data step in more detail?

Answer: The basic purpose of the SAS is to analyze or process the data in some way. However, before processing the data, it should be in a proper format which SAS can understand and it is the responsibility of the DATA step to generate data in the required format from the RAW data. You can see entire flow as below



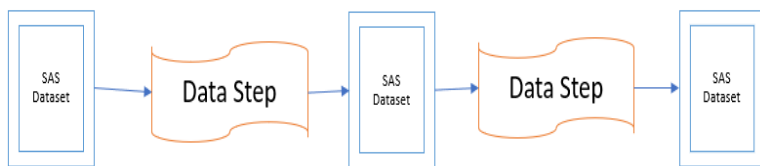
Here, raw data can be your csv file, text file or any other file, which you wanted to convert into SAS Dataset. Here first 3 block you can consider part of Data step and remaining three are the proc step.

Question-33: Data step should always be followed by proc step?

Answer: No, it is not necessary. A Data step can be followed by another Data step or Proc step. Similarly, a proc step can be followed by a Data step. In a program you can have as many Data step and proc step in any order as per the requirement.

Question-34: Can we create a new SAS Data set from existing SAS Data set?

Answer: Yes, you can create a new SAS Data set from the existing SAS data set. Which can be done using the various SAS statements like SET, MERGE, MODIFY or UPDATE. You can see below image to understand more.



Question-35: Why do you need Data step followed by another Data step sometime?

Answer: We need to sometime format the complex data before doing any analysis using the procedure. In this case we can divided Data step in multiple data step or before analyzing the data we wanted to further process the data.

Question-36: In the SAS dataset we know rows are called observation and column name are variables, but what are those in raw file?

Answer: In the raw file we can say each observation is equivalent to the records and each column name is known as field. As you can see in below Raw data first line represent all the fields and 5 records, each data value is separated by “~”

```
ID~Name~CourseName~Fee~~~FirName~LastName~Gender~Height~We  
1001~Amit~Hadoop~8000~~~Amit~Jain~Male~162~58  
1002~Rakesh~SAS~7000~~~Rakesh~Gupta~Male~159~57  
1004~Nutan~Spark~7000~~~Nutan~Jain~Female~148~51  
1003~John~Python~6000~~~John ~Fernandes~Male~178~78  
1005~Venkat~Scala~8000~~~Venkat ~Upala~Male~167~71
```

Question-37: What happen if some records do not have a value for every fields?

Answer: When Raw data is converted into SAS Data set and if any data value is missing then it can be either represent by blank space or by the period (“.”). If missing data value is numeric type then it would be shown as period and if missing data value is character type then it is shown as “ ” blank space.

Question-38: What is permanent and temporary dataset?

Answer: Any dataset which is created under the Work data library then it is temporary and as soon as you close your session these data would be deleted. And permanent dataset are those dataset which remain persisted even you close your session and re-open it. Any dataset you would be creating other than Work data library are permanent dataset.

Question-39: How do you specify in which data library the data should be created?

Answer: You can specify dataset using two names separated by "." Period. Where the name before the period is data library name and after the period is dataset name. Let see the example "he_sas.he_learner". Here, he_sas is a data library name and he_learner is a dataset name. Another example is "Work.he_data" . In this case he_data is created in the work data library, hence this "he_data" is a temporary dataset and would be deleted as soon as session is closed. If you create a dataset without the library name e.g. "data he_data" then it is always temporary and would be created in the work data library.

Question-40: What is the use of LIBNAME SAS statement?

Answer: You can use the LIBNAME to assign a libref (library reference) to a SAS data library for example below

libname he_sas

'C:\Users\HadoopExam\he_sasdata';

As you can see "libname" is a keyword, "he_sas" is a libref and 'C:\Users\HadoopExam\he_sasdata' is physical path for the data library where the actual data would be stored. Here, libref (he_sas) is a logical name to refer the data stored in the directory/folder specified.

Question-41: What all are the restriction for creating libref?

Answer: It has following rules

- Can only be 1 to 8 character long.
- It can begin with the letter or underscore
- It can contain only letter, number and underscore.

Question-42: Can you give an example creating permanent dataset?

Answer: You can have Data step as below to create a permanent dataset

```
libname he_sas 'C:\Users\HadoopExam\he_sasdata';  
data he_sas.he_learner
```

You can create permanent dataset named he_learner in the he_sas libref, which would be stored in the 'C:\Users\HadoopExam\he_sasdata' folder.

Question-43: Why it is said that Data step is a loop?

Answer: Data step is said as loop because it keeps iterating on all the records in the Raw data file until all the records are iterated to create new SAS dataset. Hence, if your raw file contains 1000 records then data step would be iterated 1000 times.

Question-44: What do you mean by SAS Dataset first phase is compile and what does it do in compile phase?

Answer: Whenever you submit a Data step for execution it would first check the syntax of the SAS statements and check whether these are correct or not and then compile it. Compiling the SAS statements means converting them into the machine instructions which computer processor can

understand also known as generating machine code. And during the compile phase of the Data step SAS generate following three items.

- Input buffer
- Program data vector
- Descriptor information

Question-45: Can you please elaborate all these items created in the compile phase and what is the use of them?

Answer: These all three things are containers for the various data and would be used during the execution phase. Let's discuss each of them

- **Input buffer:** Whenever SAS reads a record from the RAW data (during execution phase) it first places it into the input buffer. Input buffer is a logical area in memory (RAM). Input buffer at a time has only one record from raw file.
- **Program data vector:** This is also a logical area in the memory but SAS uses it to create a Data set. First raw data one record is read from a raw file and added in the input buffer as a next step it assigns variables to each data value and writes the entire single observation in the program data vector. Program data vector at a time keeps only one observation.
In addition to the observation, program data vector also has two automatic variables created which are `_N_` and `_ERROR_`
- **Descriptor information:** This holds the information about the entire dataset like all the variables information, how many observations in the data set etc.

Question-46: What is the SAS automatic variable?

Answer: During the Data step SAS creates a program data vector which can hold actual observation as well as two automatic variables named `_N_` and `_ERROR_`

Where

- `_N_`: holds how many times data step is iterated.
- `_ERROR_`: During the execution phase if there are any error happened or not.

Both the automatic variable can only be populated during the execution phase.

Question-47: What is the execution phase?

Answer: During the compilation phase only machine instructions and logical memory (input buffer, program data vector) would be created but they would not be executed. Machine instructions would be executed during the execution phase only. Suppose your raw file has 1000 records then your data step would be compiled only once but executed 1000 times, once for each record.

Question-48: What would happen during the execution phase of the data step?

Answer: During the execution phase of data step following things would be done in each iteration

- SAS reads raw record in input buffer.
- Assign values from record in the program vector and in the specific variable.
- If you are creating a new variable in the program by applying some calculation then this would also be created in the program data vector.

- Once all the records iteration is done then following 3 things would be done at the end of each execution step
 - o Dumping the program data vector (one observation) in the SAS data set
 - o Program go back to the next record and start from the Data step again.
 - o Variable in the program data vector are specified or reset to missing values.
- Following two things would not be reset in each iteration
 - o Automatic variable `_N_` and `_ERROR_`
 - o If you have specified any variable with the `RETAIN` statement
- Once all the records are processed and SAS data set would be closed for further write, and SAS goes to next DATA or PROC step, if specified as next.

Question-49: What is the use of automatic variable generated in the Data step?

Answer: There are two automatic variables generated in the Data step which are as below, this both variables has initial value set as missing. As on when Data step process each step they are updated accordingly

- `_N_` : It represents how many time DATA step has been iterated.
- `_ERROR_` : It has binary value either 0 or 1 , if while processing the records if there is any error then this would be set as 1 else remain 0.

Both of this variable are not printed as output, once the DATA step completed.

Question-50: Can SAS proc process the RAW data directly?

Answer: Generally, not, your RAW data first needs to be converted into SAS Dataset then SAS proc can process this SAS Dataset.

Question-51: What all are the possible ways to tell the SAS for reading data?

Answer: There are mainly three input styles in the SAS

- **List input:** This style is used when each field in the raw data is separated by at least one space and field value does not have embedded space in it. You have to provide the list of the variable names.
- **Column Input:** In this case each column has fixed width.
- **Formatted input:** In this case you have to specify the format of the data. For instance, numeric data can contain special symbol.

Question-52: What is informats?

Answer: While using Formatted input read style you have to provide specific instructions about the format of the data. These instructions are known as informats. For example what is the format of the date.

Question-53: What is the pointer control, while reading the RAW data?

Answer: When SAS reads the data, you can specify using the pointer control from which column it should start reading for the next item.

Question-54: What is the use “YEARCUTOFF” option in the SAS?

Answer: When SAS reads the date values and if data has two-digit year like 19, then SAS should consider it as 1919 or 2019. It can be decided using the “YEARCUTOFF” option value. Lets say you define below system option

```
OPTION YEARCUTOFF=1950
```

In this case two-digit dates would be considered between 1950 to 2049. It means if two-digit year 19 would be considered 2019, and 71 would be consider as 1971. However, it is recommended in your RAW data if possible then have 4-digit year and avoid the confusion.

Question-55: What all are the attributes for the SAS variables?

Answer: SAS variables can have following attributes

- Name of the variable
- Type of the variable
- What is the length of the variable
- Informat & format of the variable
- Label for the variable
- What is the position in an observation
- Index type

Question-56: What is the default type is considered, if you don't define the type for a variable?

Answer: If you don't define the type of the variable then it would be considered as integer and can hold upto 8 bytes. In the below statement all the 4 variables are considered as Numeric, because you are not providing any explicit type information.

input Id Name fee location;

Question-57: How can you define character variable and upto what length you can have character variable?

Answer: To define a character variable we need to add \$ sign after the variable name and we can have character variable upto 32767 bytes long. To specify the length we have to supply an informat or use the column numbers.

Question-58: How do you specify that variables can be printed or displayed in specific format?

Answer: You can use two attributes format and label which can define how the variables values can be printed and displayed.

Question-59: What all are the possible ways from which SAS can read the data to create SAS Dataset?

Answer: You can use following 4 ways by which you can specify the location of the data for creating a new Dataset

- **Stream:** You can provide raw data inline using the DATALINES statement. In this case DATALINES statement and actual data must occur last in the DATA step statements. As you can see in below example

```
data he_data;
input Id 1-4 Name $ 6-24 Location $;
datalines;
    1000 Amit Jain          Mumbai
    1001 Rakesh Kumar      Chennai
;
```

- **FILE:** If data is in the file then you have to use INFILE statement. See example below

```
data he_data;  
  infile 'c:/Users/hadoopexam/he_data.csv';  
  input Id 1-4 Name $ 6-24 Location $;  
run;
```

- **SAS Dataset:** You can use existing SAS DATASET to create new DATASET. You can use following statement and with the existing DATASET

- o **SET , MERGE, MODIFY, UPDATE**

In the following example you are creating new dataset named he_data_new from existing dataset he_data and in the new data set you are adding one additional variable "FEE" in each observation.

```
data he_data_new;  
  set he_data;  
  FEE=8000;  
run;
```

- **DBMS file:** You can create Dataset from the Database management System as well. In this case you have to use SAS/ACCESS software to bring this data into a SAS data set. You can make the connection with the Oracle or any other DBMS using libname statement and helib is a library reference, you can create a new SAS Dataset named he_data from the learner table in the database.

```
libname helib oracle user=hadoopexam password=hadoopexam path='he';  
data he_data;  
  set helib.learner;  
run;
```

Question-60: What is the "Fileref" , when you should use and what all are the advantages?

Answer: If you wanted to refer to external file then you can use the fileref for that external file. This is you can consider as a shorthand for referring external file in your SAS program. Once fileref is created you can use it in your program like in the FILE or INFILE statement.

Advantages:

- If you are referring the same file at multiple places in a program and file name changes for the external file then you don't have to make changes to many places. Only once place you change it wherever fileref is created.

As you can see in below example

```
filename heref 'c:/Users/hadoopexam/he_data.csv';  
  
data he_data;  
infile heref;  
    input Id 1-4 Name $ 6-24 Location $;  
;
```

“heref” is a fileref which represent external file at 'c:/Users/hadoopexam/he_data.csv'. And same fileref (heref) you can use in your infile statement.

Question-61: Can I create fileref for the entire directory which has multiple files in it?

Answer: Yes, you can create the same way as you created in above example, but instead of specific file change the path to the directory and in the infile statement you have to specify the filename as in below example.


```
filename heref 'c:/Users/hadoopexam/';  
  
data he_data;  
infile heref (he_data.csv);  
    input Id 1-4 Name $ 6-24 Location $;  
;
```

Question-62: What is the SAS Library?

Answer: SAS Library is a collection of SAS files, which can include SAS data sets or SAS files. And this entire collection is referenced as a unit. In Windows env. One SAS Library is equal to one folder and in Unix based env. It is equal to a directory.

Question-63: Which SAS statement you can use to create the SAS libref?

Answer: We can use LIBNAME statement to create the libref. In the single LIBNAME statement we can create more than one libref as well. LIBNAME statement is a Global statement.

Question-64: How long libref is effective?

Answer: Once libref is created it would remain effective until you change it, cancel it or until your SAS session is alive. The libref is mandatory for all the permanent SAS library which you want to refer and Work is also a libref for the temporary SAS Library.

Question-65: Which of the proc you can use to check the contents of the SAS Library and what all the information you check with that?

Answer: To check the contents of a particular SAS library we can use the “CONTENTS” proc, which can give following information.

- What all are the contents in the library (all the SAS Data files)
- Descriptor information for each SAS Data set in that library.

Question-66: Using CONTENTS proc, which variable you can use to list all the files in a SAS library?

Answer: We can use <libref>._ALL_ to list all the files in a library which is represented by libref.

Question-67: What is the use of NODS in the CONTENTS proc?

Answer: Using the NODS you can suppress the detailed information about each file when you use <libref>._ALL_

Question-68: How does DATASETS proc is different than CONTENTS proc?

Answer: Yes, this is true you can use DATASETS proc as well to view the contents of the SAS library or a SAS Data sets. But DATASETS proc has additional data management capabilities like copying, deleting and modifying SAS files. And to end the DATASETS proc you would be using QUIT statement.

Question-69: What is the use of VARNUM variable in the CONTENTS or DATASETS proc?

Answer: If you see by default CONTENTS proc print all variable information in the alphabetical order but if you want to print them in their creation order or logical position then

you have to use VARNUM option in the CONTENTS proc. For instance, you can as below

```
proc datasets;  
contents data=he_sas.le_learner varnum;
```

Question-70: What is the difference between libref and fileref?

Answer: We use the fileref to temporary point to the external file location while libref reference to the SAS data set. Similarly, libref is created using the LIBNAME while fileref is created using FILENAME statement.

Question-71: What is the use of IMPORT proc?

Answer: IMPORT proc as name suggests, import the data from external data source and writes it to a SAS Data set. Using this you can import both structured and unstructured data. Example files which you can import are

- Delimited files (Comma, tilde, blank, tab etc.)
- Excel files
- JMP 7 files (If using SAS 9.4)

However, by default IMPORT proc assume that first row has the header information. And to guess the correct number of variables, their informats and formats it scans first 20 (which you can change) rows in the file. Internally IMPORT proc create DATA step to read the contents of the file. The generated DATA step code you can see in the log information generated as part of the SAS IMPORT statement.

Question-72: What is the use of VALIDMEMNAME= system option?

Answer: Using system option VALIDNAME= we can enhance the rule for having SAS Name datasets like having special character or international character in the SAS Data set name if we use VALIDMEMNAME=EXTEND.

Question-73: Can you specify WHERE= data set option while importing csv file using "IMPORT" statement?

Answer: No, when importing delimited, comma separated, or tab-delimited external files we can not use any SAS Data set options like ALTER=, PW=, READ=, WRITE= and WHERE= etc.

Question-74: Can you import specific sheet of the excel file as SAS Data set, if yet then how?

Answer: To import the excel data we can use the IMPORT statement and we need to specify DBMS=XLSX. If we want to specify specific sheet from the xlsx then we need to use sheet=<sheet_name> option for example as below.

```
Option validvarname=v7;  
Proc import  
  datafile='c:\Users\HadoopExam\he_learner.xlsx'  
  Dbms=XLSX  
  Out = work.he_learner  
  Sheet=courses  
Run;
```

Here validvarname option helps in converting spaces into underscores, while converting column name to variable name.

Question-75: What is the use of SAS System option "OBS="?

Answer: Using the SAS OBS= option you can specify how many observations to read. Suppose OBS=5 then it read 5 observations. If you want to read entire or all observations then you have to specify OBS=max.

Question-76: Why does SAS XLSX engine does not support RANK proc?

Answer: The XLSX engine allows data to be read in sequence and not randomly. And RANK proc required data to be read randomly to order the data, hence while reading the data RANK proc or any other which need data to be read randomly are not supported.

Question-77: In the DATA Step what is the implicit Output and explicit output?

Answer: As we know SAS Data step runs in a loop, hence, at the end of each iteration there is an implicit output statement which writes data to the Dataset which is generated in the current data step. If we explicitly put output statement in the Data step it means then it overrides the automatic output and SAS adds observation to a data set only when an explicit OUTPUT statement is executed. Once you use an OUTPUT statement to write an observation to any one data set, however, there is no implicit OUTPUT statement at the end of DATA step. We can use the OUTPUT statement alone as part of an IF-THEN or SELECT or DO-LOOP processing.

Question-78: When and where the OUTPUT statement writes the observation?

Answer: SAS DATA step output statement tells the SAS to write current observation to a SAS data set immediately, not the end of the DATA step. If you are not providing any name of the DATA set then output statement would write data in the data sets which is listed in the DATA statement.

Question-79: What delimited file by default read by the IMPORT proc?

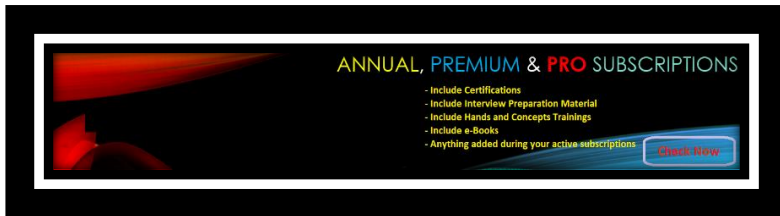
Answer: By default, IMPORT proc reads delimited files as varying record length files.

Thanks for considering SAS A00-215 Certification Preparation Material from <http://hadoopexam.com> . You can get access to all the 150+ questions and answer for your real exam preparation from below link. Where you would find the same quality of questions and correct answers with the detailed explanation. Below are most popular SAS Certification preparation material and we are continuously updating and adding more questions.



You can check all the available products across the website and have Annual Pro subscription which includes

- Trainings
- Certification Preparation
- Interview Preparation material
- Interview Preparation Audio and Video Books
- eBooks on various subjects
- Project Implementations
- Anything newly added during your active subscription also you can get.



Please join this google group to get notified whenever new questions are added

JOIN GOOGLE GROUP TO RECEIVE UPDATE NOTIFICATIONS

Question-80: While importing fixed length file how to use IMPORT proc?

Answer: If your external file has fixed length format, then use the OPTIONS statement before the PROC IMPORT statement that includes the RECFM=F and LRECL=options.

Question-81: What is the Syntax and semantic errors in SAS?

Answer: Syntax errors occur when SAS program does not confirm the programming language of the SAS for instance having a wrong spelling for SAS keyword, missing RUN statement, missing semicolon, unbalanced quotation marks etc.

And semantic errors mean you are using language element which is not valid for a particular usage.

Logic error: logic errors means that program successfully executed but not generate the expected results or incorrect

results. In this PUTLOG statement can help to find the cause of the issue, which can help you to print the messages in the SAS log.

Question-82: What is the use of the STOP statement in SAS?

Answer: STOP statement of the SAS program can help you in stopping the currently running DATA step and run the or resume the processing statement which comes after the end of current DATA step.

Question-83: What is the effect if you use “drop=” option in DATA statement and SET statement?

Answer: When you use “drop=” option in the set statement it means variable mentioned in the “drop=” option **would not be even processed and would not be part of the new data set.**

But if you want to process the variables but do not want them to be appear in the output data set then use the “drop=” option in the data statement.

Question-84: What all are the data types in SAS?

Answer: SAS has two data types NUMERIC and CHARACTER.

Question-85: What is the difference between function and procedure?

Answer: In SAS the fundamental difference between function and proc is that function expects the argument values across an observation in a SAS DATA set and procedure expects one variable value per observation.

```
data he_learner ;
```

```
set he_fee ;  
    avgfee = mean( of F1 - F6 ) ;  
run ;  
  
proc sort ;  
    by month ;  
run ;  
proc means ;  
    by month ;  
    var avgfee ;  
run ;
```

In the above code you can see calculate the average fee for each observation in a day for 6 courses. And the proc means we are passing the same average fee for each day and it calculates the mean value across 30/31 observation in the month.

Question-86: What is the difference between sum function and + operator?

Answer: In case of SUM function it would do the summation for all non-missing values while with the “+” operator, if there is any missing value then it would generate missing value as a final output.

Question-87: What is the difference between PROC MEANS and PROC SUMMARY?

Answer: Both PROC SUMMARY and PROC MEANS produces the exactly same output dataset the difference between these two procedures is that PROC MEANS produces a report by default while PROC SUMMARY produces an output

dataset by default. Hence, if you want to report printed to the listing use the PROC MEANS and if you want that info to be passed to a dataset for further then use proc summary.

Question-88: When SAS can not convert numerical value from character value automatically?

Answer: When it has a non-numeric value like \$ sign, for that you have to use informat.

Question-89: What all are the main components of the SAS programming?

Answer: There are two major steps DATA and PROC steps which you use in SAS programming and third component is a variable.

Question-90: What is PDV?

Answer: It is a sort form of Program Data Vector; we have already covered detail about this.

Question-91: Can you please again explain me the difference between function and procedure?

Answer: In SAS, even many other programming this concept is almost same, function may or may not take input from the individual observation to generate the output. However, in case of procedure it requires only one single variable.

Question-92: What is the difference between format and informat?

Answer: Format is used while writing data to a file and informat is used while reading the file. INFOMATS also input the data from external files known Flat Files ASCII.

Question-93: How can you create a permanent dataset in SAS?

Answer: While creating dataset in SAS use the library reference "libref" which should be other than "WORK" to store the SAS Dataset.

Question-94: Can a numeric variable be of the 'character' data type?

Answer: Yes, it is possible and based on the context where it is being used. It happens when numeric variable does not represent several items, but rather a position on some sort of a scale.

Lets take an example you have been given a variable named "classType" which has 3 values 1,2 & 3 where

1 -> means OnDemand class from

<http://hadoopexam.com>

2 -> means Annual Subscription training from HadoopExam.com

3 -> means Classroom training conducted by HadoopExam.com

Question-95: Can you explain the PROC contents and PROC print procedure?

Answer: Using the PROC contents you can display all the information processed in the recent Dataset or dataset explicitly mentioned. While PROC PRINT all the contents can be printed as well as make sure that mentioned information is correct or not.

For example, with the CONTENTS you can not use the WHERE option because CONTENTS proc does not process any observations from the DATASET.

Question-96: What all are the basic categories in which you can divide the SAS procedures?

Answer: You can divide SAS procedures in below three categories.

- **Report writing:** These are mainly for displaying the data like CALENDAR, PLOT, SQL, PRINT, SUMMARY etc.

- **Statistics:** Using these procedure elementary statistics can be calculated which include descriptive statistics which is based on moment, quantiles, confidence intervals, frequency counts, crosstabulations, correlations, distribution tests and ranking the standardize the data.

- Utilities: These are for various utility like
 - o Create, edit, sort, transpose data sets.
 - o Creating user-defined functions.
 - o Various file maintenance operations.Examples are: CATALOG, APPEND, SORT, DELETE, MIGRATE, PDS etc.

Question-97: How do you specify the input dataset for a procedure?

Answer: Many Base SAS procedure requires input SAS dataset, you can specify the input SAS dataset by using the DATA= option in the procedure statement as below.

```
proc print data=he_data
```

If we omit the data= option the procedure uses the value of the SAS system option `_LAST_` which represent the most recently created SAS dataset in the current session.

Question-98: What do you mean by threaded processing of SAS Base procedure?

Answer: Run small tasks in parallel or concurrently to improve the performance. Using the threaded processing enables multiple pieces of the executable code to run simultaneously, many SAS procedure specially in SAS/STAT and HPA (High performance Analytics) procedure supports the threaded processing. But keep in mind not all procedure supports the threaded processing.

Question-99: Why ODS was introduced and what is the use of it?

Answer: ODS stands for Output delivery System, which gives you the flexibility in generating, storing and reproducing SAS procedure and DATA step output, ODS provides the formatting capability which is not available from individual procedures or from the DATA step alone. ODS overcomes these limitations and enables you to format your output more easily.

Question-100: What is the use of QUIT statement?

Answer: QUIT statement executes the any statement which is not yet executed and ends the procedure.

Question-101: What do you mean by in-database processing?

Answer: In-database is processing which is done in or on the database itself rather than done in the SAS, which has several benefits like

- Increased security
- Less data transfer over the network
- Potential for faster data processing

Question-102: Why do you say that in-database processing is secure?

Answer: Increased security is possible because you are not transferring sensitive data over the network and processing in the database itself. If there is no data transfer over the network then man in the middle attack can be avoided.

Question-103: Why in-database processing is considered faster?

Answer: In database processing can be considered faster because of following reasons.

- Data is manipulated locally, on the data source, using high-speed secondary storage devices instead of being transported across a relatively slow network connection.
- The data source might have more processing resources at its disposal.
- The data source might be capable of optimizing a query for execution in a highly parallel and scalable fashion.

Question-104: What is CAS processing?

Answer: Since SAS 9.4M5, some BASE SAS procedure are enhanced to process data inside SAS Cloud Analytics Service (CAS) which is known as CAS actions. Which can result in faster processing times.

Question-105: What is the reason that processing using the CAS is faster?

Answer: There are following reason because of which CAS is considered faster

- The CAS server processes the in-memory tables instead of transferring the data across a relatively slow network connection between the server and the SAS client machine.
- The hardware for a CAS server typically has greater processing resources at its disposal.
- The majority of CAS actions are scalable for multithreaded processing. For large data volumes, distributed servers use multiple hosts to perform massively parallel processing.

Question-106: What is the basic principal of the CAS actions?

Answer: The principle is to summarize and analyze large data volumes in the in-memory tables in the CAS server. The smaller, summarized results, are transferred from the server to the SAS client. The procedure then post-processes the summarized results to produce additional statistics, Output Delivery System (ODS) objects, and so on.

Question-107: Can you give example of the base procedure which uses the CAS actions

Answer: There are quite a few such procedure please some of them below

- APPEND
- CONTENTS
- COPY
- TABULATE etc.

Question-108: What is the difference between INPUT and INFILE statement?

Answer: INFILE statement you can use to identify an external file to be read with an INPUT statement.

With the INFILE statement you provide the location of the external file and the type of the file like csv, text etc. And in the INPUT statement you are providing list of fields to be read in.

Question-109: What is the difference between FILENAME and INFILE?

Answer: Using the FILENAME we can create a fileref which is used to locate the external file and then in the INFILE statement we can use this fileref to refer that external file.

Question-110: Can you explain all the below 4 INFILE options?

FLOWOVER, MISSOVER, STOPOVER, TRUNCOVER

Answer: INFILE has various options to read the INFILE data and how to take care when the field size is not as expected.

- **FLOWOVER:** This is a default behavior in this case SAS DATA step simply reads the next record into the input buffer, attempting to find values assign to the rest of the variable names in the INPUT statement. If

you don't want SAS to move on to a new row if there are not enough data in the current row, you need to use one of the other INFILE option.

- **MISCOVER:** this option prevents the DATA step from going to next line if it does not find values in the current record for all the variables in the input statement. And for all the incomplete data a missing value would be assigned.
- **STOPOVER:** causes the DATA step to stop processing if an INPUT statement reaches the end of the current record without finding values for all variables in the statement. If your data should be complete and the process should not continue if the data are incomplete this is a good option.
- **TRUNCCOVER:** causes the DATA step to assign the raw data value to the variable even if the value is shorter than expected by the INPUT statement. If, when the DATA step encounters the end of an input record, there are variables without values, the variables are assigned missing values for that observation.

Question-111: What is the use of “@@” at the end of input statement?

Answer: Using the DOUBLE training sign in the INPUT statement, until all the fields in the record is processed it holds the current record for the execution of the next INPUT statement, as you can see in the below example where “Hadoop 8000 Spark 8500 SAS 9000” holds 6 fields which can be converted into 3 records having two fields for each record.

```

DATA he_data;
  input CourseName $ Fee @@;
  cards;
Hadoop 8000 Spark 8500 SAS 9000
Python 7600 Scala 9500
  ;
RUN;
PROC PRINT DATA=he_data;

```

Obs	CourseName	Fee
1	Hadoop	8000
2	Spark	8500
3	SAS	9000
4	Python	7600
5	Scala	9500

Question-112: What all are the by default statistics can be produced by the MEANS proc in SAS?

Answer: SAS can produce N, MIN, MAX, MEAN and STD DEV can be produced.

Question-113: What is the use of _NULL_ in SAS?

Answer: In SAS _NULL_ is a reserved keyword which specifies a SAS data set that has no observations and no variables. For example, if you specify _NULL_ as an output dataset then it would not no output generated. The _NULL_ data set is often used when you want to execute DATA step code that displays a result, defines a macro variable, writes a text file, or makes calls to the EXECUTE subroutine. In those cases, you are interested in the "side effect" of the DATA step and rarely want to write a data set to disk.

Question-114: What do you mean by SAS expression?

Answer: SAS expression represent a single value that is used in the program, and it can be a constant, variable, or an array element.

Question-115: What is the difference between NODUPKEY and NODUP?

Answer: Please find the difference

- **NONDUP:** In this case you specify on which variable you want to sort the data, once data is sorted it removes the contiguous duplicate records.
- **NONDUPKEY:** As name suggest it first creates the keys in the data and then remove the observations which are having duplicate keys. You specify which all variable should be in the key using the BY statement

Question-116: What is an alternate way to remove duplicate data?

Answer: You can use the FIRST and LAST expressions to remove the duplicate data. Variable created when sorting containing Boolean for each BY variable.

- 1: It means current observation is either 'FIRST/LAST' in the By group.
- 0: It means current observation is **not** either 'FIRST/LAST' in the By group.

Question-117: If you want to remove the duplicate observations based on the few fields like out of 10 fields if you want to remove duplicated based on 3 fields. How can you achieve that?

Answer: Then you would be using the NONDUPKEY option and in the BY statement you would be using 3 variables based on which you want to remove the duplicate values.

Question-118: How would you find all the numeric and character variables in the current dataset?

Answer: You can use `_NUMERIC_`, `_CHARACTER_` and `_ALL_` automatic variables where `_ALL_` represent all the defined variable, `_NUMERIC_` gives list of all the numeric variable and similarly `_CHARACTER_` gives all the character variable list.

Question-119: What is the difference between SELECT and IF-THEN-ELSE statement?

Answer: As such there is no difference you can use both to generate the same output but you should prefer SELECT statement when you have series of conditions and all are independent because SELECT-WHEN is more CPU efficient compare to IF-THEN-ELSE statement.

Question-120: How can I convert a Numeric variable in a Character variable?

Answer: You can use the PUT() function to convert a numeric value to a character variable. As example you can see PUT(1000, 5) would generate ' 1000'

Question-121: Can you please explain PUT and INPUT in detail?

Answer: Whenever you want to convert NUMERIC to CHARACTER and vice versa you would be using PUT and INPUT function. However, it depends what is your target and source variable. Hence, keep the following things in mind.

- PUT() function always creates a new character variables
- INPUT() function can create character or numeric variables based on what informat you are using.
- The source format must match the source variable type in PUT()
- The source variable type for INPUT() must always be character variables.

Question-122: What is the difference when you define order using “-” single hyphen and double hyphen (“--”)?

Answer: When you use – hyphen then it would ignore any in-between variable which does not matches for example below

‘FEE1, FEE2, COURSNAME, FEE3, FEE4’

If you provide range as “FEE1-FEE4’ it would produce the following

‘FEE1, FEE2, FEE3, FEE4’

If you use double hyphen it would not remove the variable in between for example FEE1--FEE4 it would produce the output as below.

‘FEE1, FEE2, COURSNAME, FEE3, FEE4’

Question-123: When do you use %include statement?

Answer: You should use %include statement brings a SAS programming statement, data line or both into a current SAS

program. The %INCLUDE statement references lines from an external file and process them immediately like you have copied code from somewhere else and pasted it to be executed immediately.

Question-124: What is the difference between Match merge vs One-to-one merge?

Answer: Using the match merge you connate the two-dataset based on one column like ID. For example, first dataset has 3 fields and second datasets has 4 fields and based on the ID match both datasets would be combined and having each observation 7 fields, obviously it has some restriction like both data should have matching ID values.

And in case of ONE to ONE merge you have to first sort both the datasets on the same column and then merge them together.

Question-125: How does SAS find that current step is ended?

Answer: As you know in SAS each program is compiled and executed independently from every other step. As a step is compiled, SAS recognizes the end of the current step when it encounters one of the following

- A DATA or PROC statement, which indicates the beginning of a new step
- A RUN or QUIT statement, which indicates the end of the current step
- RUN statement is mandatory at the end of the last step, otherwise last step would not be executed.

Question-126: When you submit the SAS program and you get the error message like “A Quoted String has become too long or that the statement is ambiguous” what does that mean?

Answer: It means you have unbalanced quotation marks in your SAS program which you need to correct and resubmit the program.

Question-127: Is that true when SAS program has syntax error, it would be compiled but execution fails?

Answer: No, whenever syntax error happens program would not compile because before compiling the program checks the syntax and if these are not correct program will not compile and give error. Syntax errors occur because program does not follow the SAS Base programming rules correctly.

Question-128: If you have used invalid option in the SAS statement SAS program will ignore that?

Answer: No, whenever you have wrong option used with the SAS program then SAS program would identify that wrong option and notifies via error message that you have provided invalid option. In this case we need to recall the SAS program, remove or correct the invalid option and resubmit the program.

Question-129: What is the difference between PUTLOG and PUT statement?

Answer: We can use PUTLOG and PUT statement to help and identify the errors and print the messages in the logs. We can use the PUTLOG statement in a DATA step to write messages to the SAS log to help identify logic errors. And also use

temporary variables in the PUTLOG to assist in debugging. And using the PUT statement to examine variables and print your own message to the SAS log.

Question-130: In the PROC PRINT report, how can you use data value as an observation column?

Answer: If you don't provide any variable then by default OBS column would be generated and produce as first column. But if you don't want OBS as a first column and some other variable from the DATASET as an ID or OBS column then use the ID option and then variable name as below.

```
proc print data=he_sas.he_data;  
    Id courseName;  
run;
```

In this case courseName would be the first column in the report instead of OBS column.

Question-131: How many "WHERE" statement you can have in a step?

Answer: You should have only one WHERE statement in a step to filter the observation, however if you put more than one WHERE statement SAS would not give any error but consider the only last WHERE statement. And also keep in mind you can use the variable in the WHERE statement which is not defined in the var statement as well.

Question-132: What is the use of CONTAINS operator?

Answer: Using the "CONTAINS" operator you can check the substring value match. For example as below.

where courseName CONTAINS 'SQL';

where courseName ? 'SQL';

Instead of CONTAINS keyword you can use '?' symbol as well.

Question-133: Setting System options applies to?

Answer: Entire active SAS session and not to the specific program.

Question-134: If you want to process only observations between 3 and 10, how can you control that?

Answer: We can control this using OBS= and FIRSTOBS= options. In this case we would put FIRSTOBS=3 and OBS=10. Hence, it would start processing the 3rd observation go until 10th observations.

Question-135: Can I override the system option in the current session for specific data step?

Answer: Yes, we can do. For a particular data step by putting them in parenthesis at the end of statement for example as below.

```
proc print data=he_sas.he_data(firstobs=3 obs=10);
```

Question-136: While using the SORT procedure, how the missing values are treated by default?

Answer: By default, missing values are treated as smallest possible values.

Question-137: If you are sorting the data by multiple variables and put the DESCENDING keyword just before that first variable, what it does?

Answer: In this case data would be sorted in descending order by the first variable only, if you want to sort in the descending order for more than one variable then put the DESCENDING keyword in front of that each variable.

Question-138: What is the problem if you don't use the OUT= option in PROC sort?

Answer: In this case your existing data would be overwritten by newly sorted data.

Question-139: What is the use of PAGEBY statement?

Answer: If you use the PAGEBY option for any variable then it would create a new page in the output. Also, please note that whatever variable you have specified in the PAGEBY statement must also specified in the BY statement.

Question-140: How many footnote and title we can assign?

Answer: You can specify at the max 10 of each, and also please note that both FOOTNOTE and TITLE are global statements. You can place them anywhere withing or before the PRINT procedure. And they apply to all subsequent output. FOOTNOTE is equal to FOOTNOTE1 and similarly TITLE is equivalent to TITLE1.

Question-141: How do you cancel the previously assigned FOOTNOTE and TITLE statement?

Answer: As we discussed previously all the FOOTNOTE and TITLE statements are global, if you want to cancel those

before your next step, you have declared null FOOTNOTE and TITLE statement. By specify no values to title and footnote.

Question-142: In the output report you want to change the header of the column which is different than variable name, how can you achieve that?

Answer: You can use LABEL statement for that.

Question-143: If you want to have label assigned permanently for a variable then how can you achieve that?

Answer: If you assign label in the DATA step itself then it would save the label with the data and in future whenever you use the PROC PRINT statement, it would use that label in the output. However, in case of PRINT PROC you still have to use label option to get displayed the label but with the other procedure this is not the requirement.

Question-144: When you use the SUM statement then it is mandatory to provide the var statement?

Answer: No, its not mandatory. It is optional. You do not need to name the variables in a VAR statement if you specify them in the SUM statement, but you can. If you choose not to name the variables in the VAR statement as well, then the SUM statement determines their order in the output.

Question-145: With the PROC SORT step which is statement is mandatory?

Answer: The 'BY' statement is mandatory.

Question-146: What is the default output look like for the PRINT proc?

Answer: By PROC PRINT displays all observations and variables in the data set, a column for observation numbers on the far left, and variables in the order in which they occur in the data set.

Question-147: During the compilation phase of DATA step what SAS creates?

Answer: During the compilation phase SAS creates the following items.

- Program Data Vector
- Descriptor information

Question-148: Is it true that PDV (Program Data vector) holds all the observations and then at the end of Data step it written to new Data set?

Answer: No, program data vector at a time contain only one observation which is being processed and as soon as processing is done. It is saved and next observation would be added in the program data vector.

Question-149: Which procedure we can use to check the descriptor information?

Answer: We can use 'CONTENTS' procedure to print the descriptor information about the Data sets.

Question-150: New Dataset is created after compilation or execution?

Answer: Compilation step checks the errors, type information of the variable and creates PDV etc. It does not write any observation; new observation would only be

written when execution phase started which process each record as a time.

Question-151: As soon as execution phase completed two automatic variables `_N_` and `_ERROR_` also stored with the output Datasets?

Answer: No, these two variables are not stored with the newly created output Datasets.

Question-152: The format to the variable is applied once the variables are added in the PDV during the execution phase?

Answer: No, the formats for each variable are applied before the SAS adds the values to the PDV. Even before each iteration PDV is reset to missing values.

Question-153: Which all variables are not reset in the PDV in each iteration during the execution phase?

Answer: following variables are not reset to the missing

- variables that are named in a RETAIN statement
- variables that are created in a sum statement
- automatic variables
- Variables that are created with options in a SET or MERGE statement also retain their values from one cycle of execution to the next.

Question-154: How FREQ procedure can be used to determine the errors in the Data?

Answer: You can use the FREQ procedure which detects the invalid character and numeric values by looking values. You can use the FREQ procedure to identify any variables that were not given an expected value.

Question-155: How does procedure means helps in finding the error in the data?

Answer: Using the MEANS procedure you can generate the summary statistics for a particular value and based on that statistics you can check whether the values are in range or not. Like if you get the max COURSEFEE around 700000 for Hadoop, it means something wrong. In the data, decimal is missed, actual value should be 7000.00 but that is an example of data entry error.

Question-156: What all is created during the compilation phased?

Answer: During the compilation phase following things are created.

- Program Data vector
- Two automatic variables in the PDV `_N_` and `_ERROR_`
- The descriptor portions of the new SAS data set is created at the end of the compilation phase. Which include
 - o Name of the Dataset
 - o Number of observations and variables.
 - o Name and attributes of the variables.
- Observations are only written after the execution steps.

Thanks for considering SAS A00-215 Certification Preparation Material from <http://hadoopexam.com> . You can get access to all the 150+ questions and answer for your real exam preparation from below link. Where you would find the same

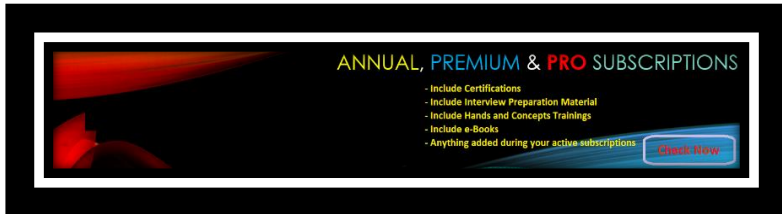
quality of questions and correct answers with the detailed explanation. Below are most popular SAS Certification preparation material and we are continuously updating and adding more questions.



You can check all the available products across the website and have Annual Pro subscription which includes

- Trainings
- Certification Preparation
- Interview Preparation material
- Interview Preparation Audio and Video Books
- eBooks on various subjects
- Project Implementations

- Anything newly added during your active subscription also you can get.



Please join this google group to get notified whenever new questions are added

JOIN GOOGLE GROUP TO RECEIVE UPDATE NOTIFICATIONS

Question-157: Wrong data values are considered Syntax errors?

Answer: No, wrong data values are data error and any program statement which does not follow the SAS programming rule are considered Syntax error like invalid options, variable names, missing or invalid punctuations, missing or misspelled keywords.

Question-158: If there are 3 errors in your data steps, then what would be value of automatic variable `_ERROR_`?

Answer: It would be 1. It does not matter how many times error occurs; it is set to 1. And if there is no error then automatic variable `_ERROR_` would have 0 value.

Question-159: What happens if SAS can not determine the syntax error?

Answer: When SAS can not determine the syntax error, it would compile the code but not execute it.

Question-160: What happens when SAS variable type is not defined correctly?

Answer: When SAS variable is not defined correctly then SAS would try to automatically convert that variable type. If it can not do that SAS continues processing and produces the output with missing values.

Question-161: What is the FIRST.variable and LAST.variable?

Answer: These are variables that SAS creates for each BY variable. SAS sets FIRST.variable when it is processing the first observation in a BY group, and sets LAST.variable when it is processing the last observation in a BY group. These assignments enable you to take different actions, based on whether processing is starting for a new BY group or ending for a BY group.

Question-162: In which scenario you don't have to preprocess the data if you are going to use the By-group processing?

Answer: Data does not require pre-processing if the observations in all of the data sets occur in one of the following patterns.

- ascending or descending numeric order
- ascending or descending character order
- not alphabetical or numerical order, but grouped in some way, such as by calendar month

If the observations are not in the order that you want, sort the data set before using BY-group processing.

Question-163: Can you explain the BY statement, in the SET statement?

Answer: When you use the BY statement in the SET statement of the DATA it would do the following things

- The data sets listed in the SET statement must be indexed or sorted by the values of the BY variable or variables.
- The DATA step automatically creates two variables, FIRST. and LAST., for each variable in the BY statement.
- FIRST. and LAST. identify the first and last observation in each BY group, respectively.

Question-164: What is the use of BY group statements?

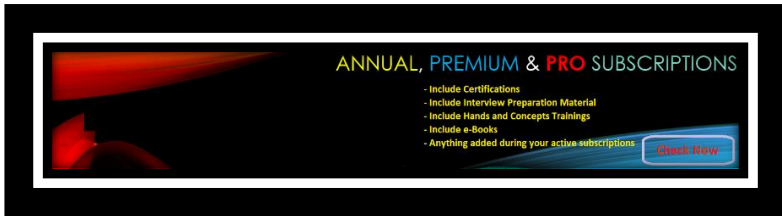
Answer: If you use the BY Group statement then it would create a group for the variable specified in the BY group statement and keep all the data with same value in the same group.

Thanks for considering SAS A00-215 Certification Preparation Material from <http://hadoopexam.com> . You can get access to all the 150+ questions and answer for your real exam preparation from below link. Where you would find the same quality of questions and correct answers with the detailed explanation. Below are most popular SAS Certification preparation material and we are continuously updating and adding more questions.



You can check all the available products across the website and have Annual Pro subscription which includes

- Trainings
- Certification Preparation
- Interview Preparation material
- Interview Preparation Audio and Video Books
- eBooks on various subjects
- Project Implementations
- Anything newly added during your active subscription also you can get.



ANNUAL, PREMIUM & PRO SUBSCRIPTIONS

- Include Certifications
- Include Interview Preparation Material
- Include Hands and Concepts Trainings
- Include e-Books
- Anything added during your active subscriptions

[Click Now](#)

Please join this google group to get notified whenever new questions are added

JOIN GOOGLE GROUP TO RECEIVE UPDATE NOTIFICATIONS