



# SAS A00-225

SAS ADVANCED ANALYTICS PROFESSIONAL CERTIFICATION QUESTIONS & ANSWERS

---

Exam Summary – Syllabus – Questions

---

**A00-225**

**SAS Certified Advanced Analytics Professional Using SAS 9**  
**50-55 Questions Exam – 67% Cut Score – Duration of 110 minutes**

[www.AnalyticsExam.Com](http://www.AnalyticsExam.Com)

## Table of Contents

Know Your A00-225 Certification Well: .....	2
A00-225 SAS Advanced Analytics Professional Certification Details: .....	2
A00-225 Syllabus: .....	3
<b>Neural Networks - 20%</b> .....	3
<b>Logistic Regression - 30%</b> .....	5
<b>Predictive Analytics on Big Data - 40%</b> .....	7
<b>Open Source Models in SAS - 10%</b> .....	9
SAS A00-225 Sample Questions: .....	10
Study Guide to Crack SAS Advanced Analytics Professional A00-225 Exam: .....	12

## Know Your A00-225 Certification Well:

The A00-225 is best suitable for candidates who want to gain knowledge in the SAS Advanced Analytics Professional. Before you start your A00-225 preparation you may struggle to get all the crucial SAS Advanced Analytics Professional materials like A00-225 syllabus, sample questions, study guide.

But don't worry the A00-225 PDF is here to help you prepare in a stress free manner.

The PDF is a combination of all your queries like-

- What is in the A00-225 syllabus?
- How many questions are there in the A00-225 exam?
- Which Practice test would help me to pass the A00-225 exam at the first attempt?

Passing the A00-225 exam makes you SAS Certified Advanced Analytics Professional Using SAS 9. Having the SAS Advanced Analytics Professional certification opens multiple opportunities for you. You can grab a new job, get a higher salary or simply get recognition within your current organization.

## A00-225 SAS Advanced Analytics Professional Certification Details:

<b>Exam Name</b>	SAS Advanced Predictive Modeling
<b>Exam Code</b>	A00-225
<b>Exam Duration</b>	110 minutes
<b>Exam Questions</b>	50 to 55 Multiple choices or short answer questions
<b>Passing Score</b>	67%
<b>Exam Price</b>	\$180 (USD)
<b>Training</b>	<ol style="list-style-type: none"> <li>1. <a href="#">SAS Academy for Data Science: Advanced Analytics</a></li> <li>2. <a href="#">Neural Network Modeling</a></li> <li>3. <a href="#">Predictive Modeling Using Logistic Regression</a></li> <li>4. <a href="#">SAS Visual Statistics: Interactive Model Building</a></li> <li>5. <a href="#">Predictive Modeling Using SAS In-Memory Statistics</a></li> <li>6. <a href="#">SAS Enterprise Miner High-Performance Data Mining Nodes</a></li> <li>7. <a href="#">Using SAS to Put Open Source Models into Production</a></li> </ol>
<b>Exam Registration</b>	<a href="#">Pearson VUE</a>
<b>Sample Questions</b>	<a href="#">SAS Advanced Analytics Professional Certification Sample Question</a>
<b>Practice Exam</b>	<a href="#">SAS Advanced Analytics Professional Certification Practice Exam</a>

## A00-225 Syllabus:

Objective	Details
<b>Neural Networks - 20%</b>	
<b>Describe key concepts underlying neural networks</b>	<ul style="list-style-type: none"> <li>- Use SAS procedures to perform nonlinear modeling                             <ul style="list-style-type: none"> <li>• Use the NLIN procedure for non-linear regression</li> </ul> </li> <li>- Explain advantages and disadvantages of using neural networks compared to other approaches                             <ul style="list-style-type: none"> <li>• Explain two ways to respond to the black-box objection</li> <li>• Compare and contrast variable selection, degrees of freedom to traditional approaches</li> <li>• Explain advantages of the Widrow-Hoff Delta rule</li> </ul> </li> </ul>
<b>Use two architectures offered by the Neural Network node to model either linear or non-linear input-output relationships</b>	<ul style="list-style-type: none"> <li>- Define the linear perceptron neural network                             <ul style="list-style-type: none"> <li>• Define combination functions (linear, additive, equal slopes)</li> <li>• Define activation functions (logistic, tanh, arctan, softmax, exponential, identity)</li> <li>• Explain the difference between activation and link functions</li> </ul> </li> <li>- Be able to demonstrate how a linear perceptron is a generalized linear model that is able to model many target distributions                             <ul style="list-style-type: none"> <li>• Explain the difference between a general and generalized model</li> <li>• Demonstrate the power of the NEURAL procedure in SAS</li> </ul> </li> <li>- Construct multilayer perceptrons                             <ul style="list-style-type: none"> <li>• Define the three layers in a basic multilayer perceptron (input, hidden, output)</li> <li>• Explain how you can obtain a skip-layer network</li> </ul> </li> <li>- Construct radial basis function networks                             <ul style="list-style-type: none"> <li>• Compare ordinary and normalized</li> </ul> </li> <li>- Identify advantages of using a radial basis function network over using a multilayer perceptron (invert order)</li> </ul>
<b>Use optimization methods offered by the SAS Enterprise Miner Neural</b>	<ul style="list-style-type: none"> <li>- Describe the problem of local minima</li> <li>- Explain the rationale behind the initialization settings</li> <li>- Explain how early stopping and weight decay can be used to help avoid bad local minima</li> </ul>

Objective	Details
<p><b>Network node to efficiently search the parameter space in a neural network</b></p>	<ul style="list-style-type: none"> <li>- Describe parameter estimation methods and determine best method to use</li> <li>- List the assortment of error functions that are available in the Neural Networks node and determine the appropriate one to use based upon statistical considerations                             <ul style="list-style-type: none"> <li>• Find the parameter set that minimizes the specified error function</li> <li>• Ordinary least squares</li> <li>• Maximum likelihood /Minimizing Deviance</li> <li>• Robust estimation methods Huber's M-estimation (HUBER)</li> <li>• Determine the appropriate activation and error function combination to apply based on the target data</li> </ul> </li> <li>- List the optimization (training) techniques available in the Neural Networks node and determine the appropriate method to use based upon statistical considerations                             <ul style="list-style-type: none"> <li>• iterative updating</li> <li>• back propagation                                     <ul style="list-style-type: none"> <li>- Conjugate gradient</li> <li>- Quasi-Newton</li> <li>- Levenberg-Marquardt</li> </ul> </li> </ul> </li> </ul>
<p><b>Construct custom network architectures by using the NEURAL procedure (PROC Neural)</b></p>	<ul style="list-style-type: none"> <li>- Working with SAS Enterprise Miner, use selected NEURAL procedure statements and PROC DMDB to construct neural networks                             <ul style="list-style-type: none"> <li>• ARCHI</li> <li>• CONNECT</li> <li>• HIDDEN</li> <li>• INPUT</li> <li>• PRELIM</li> <li>• TARGET</li> <li>• TRAIN</li> </ul> </li> <li>- Define Sequential Network Construction (SNC) and use it to build an MLP(Multilayer Perceptron)</li> <li>- Use weight interpretation to select relevant input variables</li> <li>- Define a generalized additive neural network (GANN) and be able to explain the use of the GANN paradigm</li> </ul>
<p><b>Based upon statistical considerations, use either time delayed neural networks, surrogate models to</b></p>	<ul style="list-style-type: none"> <li>- Given a particular scenario/problem, use the time delayed neural network (TDNN) model to conduct time series analysis</li> <li>- Apply a surrogate model to help understand a neural network's predictions</li> </ul>

Objective	Details
<b>augment neural networks</b>	<ul style="list-style-type: none"> <li>• Interpret a neural network with a continuous target</li> <li>• Interpret a neural network with a discrete target</li> </ul>
<b>Use the HP Neural Node to perform high-speed training of a neural network</b>	
<b>Logistic Regression - 30%</b>	
<b>Score new data sets using the LOGISTIC and PLM procedures</b>	<ul style="list-style-type: none"> <li>- Use the SCORE statement in the PLM procedure to score new cases</li> <li>- Use the CODE statement in PROC LOGISITIC to score new data</li> <li>- Describe when you would use the SCORE statement vs the CODE statement in PROC LOGISTIC</li> <li>- Use the INMODEL/OUTMODEL options in PROC LOGISTIC</li> <li>- Explain how to score new data when you have developed a model from a biased sample</li> </ul>
<b>Identify the potential challenges when preparing input data for a model</b>	<ul style="list-style-type: none"> <li>- Identify problems that missing values can cause in creating predictive models and scoring new data sets</li> <li>- Identify limitations of Complete Case Analysis</li> <li>- Explain problems caused by categorical variables with numerous levels</li> <li>- Discuss the problem of redundant variables</li> <li>- Discuss the problem of irrelevant and redundant variables</li> <li>- Discuss the non-linearities and the problems they create in predictive models</li> <li>- Discuss outliers and the problems they create in predictive models</li> <li>- Describe quasi-complete separation</li> <li>- Discuss the effect of interactions</li> <li>- Determine when it is necessary to oversample data</li> </ul>
<b>Use the DATA step to manipulate data with loops, arrays, conditional statements and functions</b>	<ul style="list-style-type: none"> <li>- Use ARRAYS to create missing indicators</li> <li>- Use ARRAYS, LOOP, IF, and explicit OUTPUT statements</li> </ul>
<b>Improve the predictive power of categorical inputs</b>	<ul style="list-style-type: none"> <li>- Reduce the number of levels of a categorical variable</li> <li>- Explain thresholding</li> <li>- Explain Greenacre's method</li> <li>- Cluster the levels of a categorical variable via Greenacre's method using the CLUSTER procedure</li> </ul> <ul style="list-style-type: none"> <li>• METHOD=WARD option</li> <li>• FREQ, VAR, ID statement</li> </ul>

Objective	Details
	<ul style="list-style-type: none"> <li>• Use of ODS output to create an output data set</li> <li>- Convert categorical variables to continuous using smooth weight of evidence</li> </ul>
<b>Screen variables for irrelevance and non-linear association using the CORR procedure</b>	<ul style="list-style-type: none"> <li>- Explain how Hoeffding's D and Spearman statistics can be used to find irrelevant variables and non-linear associations</li> <li>- Produce Spearman and Hoeffding's D statistic using the CORR procedure (VAR, WITH statement)</li> <li>- Interpret a scatter plot of Hoeffding's D and Spearman statistic to identify irrelevant variables and non-linear associations</li> </ul>
<b>Screen variables for non-linearity using empirical logit plots</b>	<ul style="list-style-type: none"> <li>- Use the RANK procedure to bin continuous input variables (GROUPS=, OUT= option; VAR, RANK statements)</li> <li>- Interpret RANK procedure output</li> <li>- Use the MEANS procedure to calculate the sum and means for the target cases and total events (NWAY option; CLASS, VAR, OUTPUT statements)</li> <li>- Create empirical logit plots with the GPLOT procedure</li> <li>- Interpret empirical logit plots</li> </ul>
<b>Apply the principles of honest assessment to model performance measurement</b>	<ul style="list-style-type: none"> <li>- Explain techniques to honestly assess classifier performance</li> <li>- Explain overfitting</li> <li>- Explain differences between validation and test data</li> <li>- Identify the impact of performing data preparation before data is split</li> </ul>
<b>Assess classifier performance using the confusion matrix</b>	<ul style="list-style-type: none"> <li>- Explain the confusion matrix</li> <li>- Define: Accuracy, Error Rate, Sensitivity, Specificity, PV+, PV-</li> <li>- Explain the effect of oversampling on the confusion matrix</li> <li>- Adjust the confusion matrix for oversampling</li> </ul>
<b>Model selection and validation using training and validation data</b>	<ul style="list-style-type: none"> <li>- Divide data into training and validation data sets using the SURVEYSELECT procedure</li> <li>- Discuss the subset selection methods available in PROC LOGISTIC</li> <li>- Discuss methods to determine interactions (forward selection, with bar and @ notation)</li> <li>- Create interaction plot with the results from PROC LOGISTIC</li> <li>- Select the model with fit statistics (BIC, AIC, KS, Brier score)</li> </ul>
<b>Create and interpret graphs (ROC, lift, and gains charts) for model comparison and selection</b>	<ul style="list-style-type: none"> <li>- Explain and interpret charts (ROC, Lift, Gains)</li> <li>- Create a ROC curve (OUTROC option of the SCORE statement in the LOGISTIC procedure)</li> <li>- Use the ROC and ROCONTRAST statements to create an overlay plot of ROC curves for two or more models</li> </ul>

Objective	Details
	<ul style="list-style-type: none"> <li>- Explain the concept of depth as it relates to the gains chart</li> </ul>
<p><b>Establish effective decision cut-off values for scoring</b></p>	<ul style="list-style-type: none"> <li>- Illustrate a decision rule that maximizes the expected profit</li> <li>- Explain the profit matrix and how to use it to estimate the profit per scored customer</li> <li>- Calculate decision cutoffs using Bayes rule, given a profit matrix</li> <li>- Determine optimum cutoff values from profit plots</li> <li>- Given a profit matrix, and model results, determine the model with the highest average profit</li> </ul>
<p><b>Predictive Analytics on Big Data - 40%</b></p>	
<p><b>Build and interpret a cluster analysis in SAS Visual Statistics</b></p>	<ul style="list-style-type: none"> <li>- Assign roles for cluster analysis</li> <li>- Set cluster matrix properties (number, seed, etc)</li> <li>- Select the proper inputs for the k-means algorithm for a given cluster analysis scenario</li> <li>- Choose the number of clusters for a given cluster analysis scenario</li> <li>- Set Parallel coordinate properties for cluster analysis</li> <li>- Interpret a cluster matrix</li> <li>- Interpret a parallel coordinates plot</li> <li>- Display summary statistics for clusters</li> <li>- Interpret summary statistics for clusters</li> <li>- Assign cluster IDs to the data within Visual Statistics</li> <li>- Score observations into clusters based on the results from Visual Statistics</li> </ul>
<p><b>Explain SAS high-performance computing</b></p>	<ul style="list-style-type: none"> <li>- Identify limitations of traditional computing environments</li> <li>- Describe the characteristics of SAS High-Performance Analytics procedures</li> <li>- Compare SMP and MPP computing modes</li> <li>- Distinguish between HPA and the LASR related operation</li> </ul>
<p><b>Perform principal component analysis</b></p>	<ul style="list-style-type: none"> <li>- Explain how principal component analysis is performed</li> <li>- List the benefits and problems of principal component analysis</li> <li>- Distinguish between clustering, variable clustering, and principal component analysis</li> <li>- Determine the number of principal components to retain</li> <li>- Compare IMSTAT, Visual Statistics, and High Performance Computing nodes in Enterprise Miner</li> </ul>
<p><b>Analyze categorical targets using logistic regression in SAS Visual Statistics</b></p>	<ul style="list-style-type: none"> <li>- Assign roles for logistic regression</li> <li>- Assign properties for logistic regression</li> <li>- Filter data used for logistic regression</li> <li>- Interpret logistic regression results (fit summary, residual plots, ROC/Lift charts, etc)</li> </ul>



Objective	Details
	<ul style="list-style-type: none"> <li>- Use Group-By variables to perform binary logistic regression</li> </ul>
<p><b>Analyze categorical targets using decision trees in SAS Visual Statistics</b></p>	<ul style="list-style-type: none"> <li>- Assign roles for decision trees</li> <li>- Assign properties for decision trees</li> <li>- Interpret decision trees results (trees, leaf statistics, assessment, etc)</li> <li>- Identify variable importance with decision trees for use in other analysis techniques</li> <li>- Splitting criteria used by Visual Statistics</li> </ul>
<p><b>Analyze categorical targets using decision trees in PROC IMSTAT</b></p>	<ul style="list-style-type: none"> <li>- Use the DECISIONTREE statement to create decision trees</li> <li>- Define input variables with the INPUT and NOMINAL options</li> <li>- Create and retrieve saved trees for input data scoring with the SAVE, TREETAB, and ASSESS options</li> <li>- Evaluate the output of ODS tables (DTREE, DTreeVarImpInfo, DTREEScore, etc) from decision trees</li> <li>- Use the ASSESS statement to create data sets for evaluating the decision tree model</li> <li>- Perform honest assessment on PROC IMSTAT decision trees</li> <li>- Assess decision trees using ODS statistical graphics (SGPLOT)</li> </ul>
<p><b>Analyze categorical targets using logistic regression in PROC IMSTAT</b></p>	<ul style="list-style-type: none"> <li>- Assign variables to roles for logistic regression in PROC IMSTAT</li> <li>- Create logistic regression in PROC IMSTAT using the LOGISTIC statement</li> <li>- Use selected options of the LOGISTIC STATEMENT (ROLEVAR, INPUTS, SCORE, CODE, SHOWSELECTED, SLSTAY=)</li> <li>- Assess logistic regression models using ODS statistical graphics (SGPLOT)</li> <li>- Perform honest assessment on PROC IMSTAT logistic regression</li> </ul>
<p><b>Build random forest models with PROC IMSTAT</b></p>	<ul style="list-style-type: none"> <li>- Describe random forests</li> <li>- Use the RANDOMWOODS statement to build a forest of trees</li> <li>- Score data with the RANDOMWOODS score code</li> <li>- List benefits of forests</li> <li>- Interpret random forests</li> <li>- Identify variable importance with forest for use in other analysis techniques</li> </ul>
<p><b>Analyze interval targets with SAS Visual Statistics</b></p>	<ul style="list-style-type: none"> <li>- Build linear regression models in SAS Visual Statistics</li> <li>- Assign roles for linear regression models</li> <li>- Set properties for linear regression models</li> <li>- Assess a linear regression model (evaluate Fit summary statistics, residual plot, influence plot, summary table, etc)</li> </ul>

Objective	Details
	<ul style="list-style-type: none"> <li>- Assess linear model assumption violations and recognize when linear model is inadequate</li> <li>- Build generalized linear models in SAS Visual Statistics</li> <li>- Assign roles for generalized linear models</li> <li>- Set properties for generalized linear models</li> <li>- Assess a generalized linear model (evaluate Fit summary statistics, residual plot, assessment, etc)</li> </ul>
<p><b>Analyze interval targets with PROC IMSTAT</b></p>	<ul style="list-style-type: none"> <li>- Use GENMODEL and GLM statements</li> <li>- Distinguish between GENMODEL and GLM statements and the results of each procedure</li> <li>- Assign variables to roles for GENMODEL and GLM statements in PROC IMSTAT</li> <li>- Create models with GENMODEL and GLM statements in PROC IMSTAT</li> <li>- Use selected options of the GENMODEL and GLM statements in PROC IMSTAT</li> <li>- Assess models using ODS statistical graphics (SGPLOT)</li> <li>- Perform honest assessment on PROC IMSTAT linear models</li> </ul>
<p><b>Analyze zero inflated models with HPGLM in Enterprise Miner</b></p>	<ul style="list-style-type: none"> <li>- Identify when it would be appropriate to use mixture distribution</li> <li>- Describe the link functions and distributions available in the HP GLM node</li> <li>- Build a zero inflated generalized linear model in EM</li> <li>- Describe restrictions on roles and levels in input data sources for generalized linear models in EM</li> <li>- Assess a zero inflated generalized linear model (evaluate Fit summary statistics, residual plot, assessment, etc)</li> </ul>
<p><b>Open Source Models in SAS - 10%</b></p>	
<p><b>Incorporate an existing R program into SAS Enterprise Miner</b></p>	<ul style="list-style-type: none"> <li>- Enable R language statements to connect SAS to R</li> <li>- Use the Open Source Integration node in SAS Enterprise Miner                             <ul style="list-style-type: none"> <li>• Modes of operation (training, output)</li> <li>• Use Predictive Modeling Markup Language (PMML) in Open Source Integration Node</li> </ul> </li> <li>- Use Enterprise Miner variable handles to alter an R script                             <ul style="list-style-type: none"> <li>• Use Enterprise Miner to run a random forest in R</li> </ul> </li> </ul>
<p><b>Incorporate an existing Python program into SAS Enterprise Miner</b></p>	<ul style="list-style-type: none"> <li>- Determine steps to perform in SAS to incorporate a Python model</li> <li>- Determine nodes in Enterprise Miner to incorporate a Python model</li> <li>- Determine the necessary set up requirements for running Python models in SAS</li> </ul>

## SAS A00-225 Sample Questions:

### Question: 1

After a logistic regression has been created in SAS Visual Statistics, you discover that not all observations were used to create the model. How would you run the model on all of the data?

- a) Include the correct interaction term.
- b) Select Informative Missingness on the properties tab.
- c) Include the correct offset term.
- d) Select Use Variable Selection on the properties tab.

**Answer: b**

### Question: 2

The softmax activation function is appropriate for which type of target?

- a) Continuous
- b) Binary
- c) Ordinal
- d) Multinomial

**Answer: d**

### Question: 3

McCulloch-Pits neurons used which activation functions?

- a) hyperbolic tangent
- b) logistic function
- c) Elliot function
- d) step function

**Answer: d**

### Question: 4

In order to take advantage of a neural network's ability to model a nonlinear relationship between inputs and outputs, which feature of the network is necessary?

- a) The inclusion of direct connections between the input and output units in the network.
- b) At least one hidden layer with a non-linear activation function.
- c) A non-linear combination function in the output units.
- d) A sigmoidal activation function in the output units.

**Answer: b**

**Question: 5**

Why is a decision tree an ideal surrogate model for a neural network?

- a) The if-then rules are easy to interpret.
- b) A decision tree is a black-box.
- c) A decision tree can be used to do variable selection.
- d) A decision tree is a parametric model.

**Answer: a**

**Question: 6**

Which statement is true with respect to the DECISIONTREE statement in PROC IMSTAT?

- a) Only binary target variables are supported.
- b) By default, pruning is based on assessment of a holdout sample.
- c) The C4.5 decision tree methodology is employed to derive the decision tree.
- d) Pruning can be controlled using the GREEDY option.

**Answer: c**

**Question: 7**

What is the primary purpose of weight decay?

- a) Prevent overfitting.
- b) Prevent underfitting.
- c) Avoid bad local minima.
- d) Increase convergence speed.

**Answer: a**

**Question: 8**

Which SAS Enterprise Miner node should you use to run a Python script?

- a) Open Source Integration node
- b) Model Import node
- c) SAS Code node
- d) Register Model node

**Answer: c**

**Question: 9**

In a forest, what is the out of bag (OOB) sample?

- a) A random partition of the validation data
- b) The partition of training data not used in growing an individual tree
- c) The partition of training data not used in growing any tree
- d) The partition of the validation data most closely resembling the data used to train a tree

**Answer: b**

**Question: 10**

In the Open Source Integration node in SAS Enterprise Miner, which Output Mode(s) creates SAS DATA step score code for the user?

- a) Predictive Modeling Markup Language (PMML)
- b) None
- c) Merge
- d) Both PMML and Merge

**Answer: a**

## Study Guide to Crack SAS Advanced Analytics Professional A00-225 Exam:

- Getting details of the A00-225 syllabus, is the first step of a study plan. This pdf is going to be of ultimate help. Completion of the syllabus is must to pass the A00-225 exam.
- Making a schedule is vital. A structured method of preparation leads to success. A candidate must plan his schedule and follow it rigorously to attain success.
- Joining the SAS provided training for A00-225 exam could be of much help. If there is specific training for the exam, you can discover it from the link above.
- Read from the A00-225 sample questions to gain your idea about the actual exam questions. In this PDF useful sample questions are provided to make your exam preparation easy.
- Practicing on A00-225 practice tests is must. Continuous practice will make you an expert in all syllabus areas.

## Reliable Online Practice Test for A00-225 Certification

Make AnalyticsExam.Com your best friend during your SAS Advanced Predictive Modeling exam preparation. We provide authentic practice tests for the A00-225 exam. Experts design these online practice tests, so we can offer you an exclusive experience of taking the actual A00-225 exam. We guarantee you 100% success in your first exam attempt if you continue practicing regularly. Don't bother if you don't get 100% marks in initial practice exam attempts. Just utilize the result section to know your strengths and weaknesses and prepare according to that until you get 100% with our practice tests. Our evaluation makes you confident, and you can score high in the A00-225 exam.

**Start Online Practice of A00-225 Exam by Visiting URL**

**<https://www.analyticsexam.com/sas-certification/a00-225-sas-advanced-predictive-modeling>**