

SAS CONTENT

S.NO	TOPICS	
1	SAS for Analyst Programmers	
	SAS/BASE	
	Introduction	
		Introduction to SAS
		Introduction to Analytics
		Editor File
		Log File
		Output File
		Result, Explorer windows
		Permanent SAS datasets
		Environments where SAS runs
		SAS Libraries
		SAS Data types
		Step boundaries and run-group processing
	Advanced Input Features	
		Read Data and Raw Data files
		INFILE Statement
		Input Statement
		DATALINES, CARDS
		List INPUT
		Column INPUT
		Formatted INPUT
		Mixed INPUT
		:(colon) and & (Ampersand) Modifiers
		double trailing @@
		Black box SAS programming
		/ and # pointers,Single trailing @, Character Pointer
		MISCOVER,TRUNCOVER,SCANOVER
		DLM,DSD,Firstobs,Obs
	Data Step functions	
		Character functions (left,right,tranwrd,translate ,index, indexw, indexc,find,trim, catx,catt, cat,cats,compress, scan,substr, upcase,lowcase,strip,etc.
		Numaricfunctions (int,ceil,floor,Round, abs, Min,Max, Sum,mean, lag, dif).

	Date function (Today, Datetime, Time, Timepart, Datepart, Day, Month, Year, Qtr, MDY etc).
	SAS System options
SAS Structures and Flow	
	Data step overview
	LENGTH
	FORMATS, IMFORMATS, LABELS
	Titles and Footnotes
	reading existing SAS datasets with SET
	Assignment statements
	RENAME
	DROP and KEEP
	Subsetting observations
	Subsetting IF statement
	If, If-then, Else-if, If then Do.
	WHERE statement
SAS Do-loops	
	Syntax
	How to create Do loops
	Conditional Do loops(DO until, DO while, by clause).
	Nesting Do loops
	Arrays
SAS Executable Statements	
	Accumulating totals
	RETAIN and SUM
	SUM statement
	SELECT statement
	Deleting observations
	Numeric-character conversion
	OUTPUT, PUT
Procedures	
	Proc Print
	Proc Contents
	Proc Sort
	Proc Means
	Proc Freq
	Proc Report
	Proc Tabulate
	Proc Printto
	Proc Dataset
	Proc Compare

	Proc Transpose
	Proc Format (input, put function, creating permanent formats, informats)
	Proc Import
	Proc Export
	Proc Append
	Proc Summary
Merging SAS Datasets	
	Syntax
	one to one merges
	Match merging
	Multiple OBS with the same BY variable
	Merging with identical variable names
	Merging without a common variable
	Update statement
SAS/ODS	
	ODS/Trace/Select/Exclude
	ODS/HTML FILE
	ODS/PDF FILE
	ODS/RTF FILE
	PROC PRINT WITH STYLE OPTION
	PROC REPORT WITH STYLE OPTION
	PROC TABULATE WITH STYLE OPTION
SAS/GRAPH	
	Proc PLOT
	Proc GPLOT
	Proc CHART
	Proc GCHART
SAS/SQL	
	SQL DDL Statements (CREATE,ALTER,DROP)
	SQL DML Statements (INSERT,UPDATE,SELECT,DELETE)
	SQL Filter Clauses (WHERE, GROUP BY, HAVING, ORDER BY)
	SQL Horizontal Joins (INNER, LEFT,RIGHT, FULL, FULL with condition, Cartesian product)
	SQL Vertical Joins (UNION,UNION ALL,INTERSECT,EXCEPT)
SAS/MACROS	
	An Introduction to SAS Macros
	Functions of the SAS macro processor
	Macro processor flow
	Macro and macro variable

		Defining and using a macro
		Creating macro variables- 3 ways
		Local and global Macro
		Automatic macro variables
		Avoid macro errors
		Positional macro parameters
		Keyword macro parameters
		Call Symput and symget
		System options for debugging macro
	SAS/STATISTICS	
		Proc UNIVARIATE
		Proc MEANS
		Proc FREQ(Chi-Square)
		Proc GLM
		Proc RANK
		Proc ANOVA
		Proc REG
		Proc LOGISTIC(logistic Regression)
		Proc TTEST(Paired)
		Proc CORR (correlation)
	SAS/ACCESS	
		How to connect with data server
		libname code
		Sql Code
	Projects	
		3 Projects have to be completed
2	Advance Analytics Using SAS	
	Introduction to Statistics/analytics	
		Need for analytics
		Analytics use in different industries
		Challenges in adoption of analytics
		Overview of Course Contents
	Data understanding	Data types (nominal, Ordinal, Interval and ratio)
	Parametric & Non-Parametric test	
	Estimation	
	Descriptive statistics	
		Tabular & Graphical Method, Summary statistics, Means, Freq, Correlation, Rank etc
	Linear Regression	
		fit a multiple linear regression model using the REG and GLM procedures

	Analyze the output of the REG procedure for multiple linear regression models
	Use the REG procedure to perform model selection
	Assess the validity of a given regression model through the use of diagnostic and residual analysis
Logistic Regression	
	Perform logistic regression with the LOGISTIC procedure
	Optimize model performance through input selection
	Interpret the output of the LOGISTIC procedure
	Score new data sets using the LOGISTIC and SCORE procedures
Introduction to some statistical terminologies and inferences	
	Population, Sample and random variables
	Point and interval Estimations
	Probability
	Discrete/Continuous probability Distributions
Hypothesis Testing	
T-Test	
	One-Tailed, Two-Tailed, Z-Test
Anova	
	Verify the assumptions of ANOVA
	Analyze differences between population means using the GLM and TTEST procedures
	perform ANOVA post hoc test to evaluate treatment effects
	Detect and analyze interactions between factors
CHI-SQUARE	
Prepare Inputs for Predictive Model Performance	
	Identify potential problems with input data
	use the DATA step to maipulate data with loops, arrays,conditional statements and functions
	Reduce the number of categorical levels in a predictive model
	Screen variables for irrelevance using the CORR procedure
	Screen variables for non-linearity using empirical logit plots
Measure Model performance	
	Apply the principles of honest assessment to model performance measurement

		Assess classifier performance using the confusion matrix
		Model selection and validation using training and validation data
		Create and interpret graphs for model comparison and selection
		Establish effective decision cut-off values for scoring
	Cluster Analysis	Case study on cluster Analysis
	factor Analysis	Case Study on factor Analysis
3	SQL FOR ANALYST	
	SQL SERVER 2008	
	Introduction to SQL Server Concepts	
	Introduction to DBMS & RDBMS Concepts	
		SQL Introduction
		Sql Commands
	Data Types	
	Data Definition Languages	
		Create table command
		Alter table command
		Truncate table command
		Drop table command
	Constraints	
		Not Null
		Unique Key
		Primary key
		Foreign Key
		Check
		Default
	Data Manipulation Language	
		Select Command
		Insert Command
		Update Command
		Delete command
	Filter Clause	
		WHERE Clause
		GROUP BY Clause
		Having Clause
		Order by clause
	Operators	
		Arithmetic Operators

	Comparison Operators
	Logical Operators
Range Operators	
	IN/NOT IN
	Between
Set Operators	
	Union
	Union All
	Intersect
	Except
SQL Functions	
	Aggregate Function
	Date Function
	String Function
Identity Properties	
Column and table Alias Joins	
	Simple join:
	Non equi join
	Equi join
	Self join
	inner join
	Outer join
	Left Outer join
	Right Outer join
	Full Outer join
	Cross join
Different Types of queries	
	Simple Queries
	Sub Queries
	Nested Sub Queries
	Correlated Sub Queries
	Temporary tables
	Common Table Expressions(CTEs)
	Derived Tables
	Scalar Expressions
INDEXES	
VIEWS	
Stored Procedures	
Triggers	
	DDL,DML,LOGON,Triggers
Cursors	

	Search expressions like	
	Dealing with Nulls views and Derived Tables	
	Exercises and Real Time examples	
4	R For Analyst	
	Introduction to R	
		Why R is used
		Interfaces in R
	Programming Constructs in R	
		Syntax
		Data Types
		Functions
	Manipulations using R	
		Extractions of Data
		Storage
		Executions
	Analytics using R	
		Calculating various Descriptive Statistics
		Analyzing Data using R
		Regression Analysis using R
		Cluster Analysis using R
		Factor Analysis using R