SAS Econometrics Procedures

Time-Series and Forecasting Models	CARIMA– Fits autoregressive integrated moving average (ARIMA) models with covariates. CESM– Leverages an Exponential Smoothing Method. CSSM– Implements state-space models for time-series forecasting. DYNAMICLINEAR – Fits dynamic linear models for time-series forecasting. ECM– Develops an economic capital model. HMM– Implements hidden Markov models for sequential data analysis. TSINFO– Extracts key time-series characteristics for analysis. TSMODEL– Builds and evaluates time-series models UCM – Fits unobserved components models for trend and seasonal decomposition.
Causal Inference Models	CAUSALDISCOVERY– Identifies causal relationships in observational data. DEEPCAUSAL– Performs causal inference, policy evaluation, and policy comparison by applying DNNs, when the treatment variable is binary. DEEPPRICE– Performs causal inference, policy evaluation, and policy comparison by applying DNNs, when the treatment variable is continuous.
Panel and Cross-Sectional Analysis	CNTSELECT– Implements count data regression models with model selection CPANEL– Implements linear panel data models CQLIM– Estimates quantitative and limited dependent variable models
Bayesian and Stochastic Models	CNTSELECT– Implements Bayesian inference for count data regression models CQLIM– Implements Bayesian inference for quantitative and limited dependent variable models HMM– Implements hidden Markov models for sequential data analysis. SMC– Applies sequential Monte Carlo methods for Bayesian inference.

Economic and Financial Modeling	CCDM- Estimates aggregate loss that occurs over a period of time DEEPPRICE– Uses deep learning for price elasticity estimation. FRONTIER– Computes stochastic frontier models for efficiency analysis. SEVSELECT – Selects severity models for risk assessment.
Spatial and Attribution Analysis	CSPATIALREG– Performs spatial regression analysis for geographically correlated data. MKTATTRIBUTION– Analyzes marketing attribution using econometric techniques.

SAS Visual Forecasting Procedures

Data Filtering	DFIL – Performs data filtering for time-series forecasting.
Forecasting Models	KTMONITOR– Monitors multivariate processes over time to assess stability. KTTRAIN– Trains models using stable process data for monitoring. MTS - Performs fault detection and diagnostics of multivariate data. MTSSCORE – Performs monitoring for fault detection, and it provides diagnostics (or root-cause analysis) of the faults for the ongoing process. SMCALIB– Calibrates forecasting models for improved accuracy. SMPROJECT– Manages forecasting projects and model repositories. SMSCORE– Scores new data using calibrated forecasting models. SMSELECT – Selects the best forecasting models for scoring. SMSPEC – Specifies forecasting model parameters and configurations.
Time Series	TSCUSTINT– Loads user-defined custom time intervals that are used for time series analysis and forecasting. TSGLOBALRECON– Performs global reconciliation of hierarchical forecasts simultaneously across all levels of a hierarchy. TSINFO– Evaluates a variable for its suitability as a time ID variable in SAS procedures and solutions that are used for time series analysis. TSMODEL–Constructs, transforms, analyzes, and forecasts time series data via distributed processing employing function-based packages such as Automatic Time Series Modeling (ATSM), Time Series Analysis (TSA), Time Frequency Analysis (TFA), etc., and executing user-defined programs. TSMODREPO– Enables users to migrate model repositories that are created by SAS Forecast Server or SAS Forecast Server procedures to a format that is compatible for use by the cloud-enabled SAS Visual Forecasting procedures. TSRECONCILE– Reconciles forecasts at two different levels of a hierarchy in a top-down fashion. TSSELECTLAG– Calculates the lag at which two different time series are maximally correlated. This can determine whether two time series have essentially the same shape.

SAS Optimization Procedures

SAS Optimization provides access to SAS/OR procedures.

CLP– Solves constraint satisfaction problems using finite-domain constraint programming.
OPTLP– Solves linear programming problems using simplex and interior-point methods
OPTMILP– Solves mixed integer linear programming problems with branch-and-bound techniques.
OPTMODEL– Provides a modeling environment for building and solving optimization problems.
OPTNETWORK– Solves network-based problems such as shortest paths and minimum-cost flows.
OPTQP– Solves quadratic programming problems with linear constraints.

Visual Machine Learning Procedures

Model Assessment	ASSESSBIAS– Evaluates bias in machine learning models to ensure fairness in predictions. RECASSESS– Assesses recommendations in recommender systems.
Deploy Models	ASTORE– Saves and deploys analytical models for scoring in different environments. REGISTERMODEL– Registers models for deployment and management.
Classification and Regression Models	BOOLRULE– Generates Boolean rules for classification tasks. FASTKNN– Implements fast k-nearest neighbors for classification and regression. FOREST– Builds random forests for classification and regression. GRADBOOST – Implements gradient boosting for predictive modeling. GPCLASS– Uses Gaussian processes for classification tasks. GPREG– Applies Gaussian processes for regression analysis. LIGHTGRADBOOST– Uses a lightweight gradient boosting framework for efficient modeling. NNET– Builds neural networks for predictive modeling. SVMACHINE– Implements support vector machines for classification.
Clustering and Dimensionality Reduction	FPCA- Performs functional principal component analysis on dense, regularly spaced functional data. FPCASCORE- Applies a previously trained functional principal component analysis model to new functional data to compute functional principal component scores. GMM– Implements Gaussian mixture models for clustering. GVARCLUS– Performs variable clustering using Gaussian methods. KPCA– Conducts kernel principal component analysis for dimensionality reduction. MWPCA – Conducts moving window principal component analysis for time-series data. NOMINALDR- Implements dimensionality reduction methods for nominal data. RPCA– Performs robust principal component analysis. SPARSEML – Conducts sparse machine learning for high-dimensional data. TSNE – Applies t-distributed stochastic neighbor embedding for visualization.

Bayesian and Sequential Modeling	BNET– Implements Bayesian networks for probabilistic modeling and inference. DYNBNET– Builds dynamic Bayesian networks for time-series analysis. SEQMC– Implements sequential Monte Carlo methods.
Feature Engineering and Model Interpretation	FISM– Performs feature importance scoring for machine learning models. LIME- Implements the local interpretable model-agnostic explanations (LIME) method that explains how a machine learning model behaves near a specific query observation. MTLEARN– Implements multi-task learning for shared feature modeling. OPTBINNING– Optimally bins continuous variables for better predictive performance. PARTIALDEPEND– Computes partial dependence plots for model interpretability. SHAPLEY- Estimates Shapley values of a query on the basis of the information that is provided in a reference table.
Recommender Systems and Market Analysis	MBANALYSIS– Performs market basket analysis for association rule mining. RECENGINE – Implements recommendation engines for personalized predictions.
Unsupervised Learning and Factor Analysis	FACTMAC– Conducts factor analysis to identify latent variables. SEMISUPLEARN– Applies semi-supervised learning techniques.
Graph and Network Analytics	NETWORK – Analyzes network structures and relationships. PATHING– Analyzes paths in sequential data.
Deep Learning and Generative Models	FITTEDQNET– Fits quantile regression neural networks. STYLEGAN– Generates synthetic images using StyleGAN. TABULARGAN– Generates synthetic tabular data using GANs.
Text Mining and Anomaly Detection	SVDD– Uses support vector data description for anomaly detection. TEXTMINE– Performs text mining and natural language processing. TMSCORE– Computes text similarity scores.
Pattern Mining	CSPADE– Performs sequential pattern mining on large datasets.
Deep Learning	DLMZEXPORT– Exports trained deep learning models from the Model Zoo for deployment. DLMZSCORE – Applies a trained deep learning model to new data for scoring. DLMZTRAIN– Trains deep learning models using the Model Zoo framework.

SAS Visual Statistics Procedures

Model Assessment and Validation	ASSESS– Evaluates model performance using various assessment metrics. MVOUTLIER– Detects multivariate outliers in datasets. SANDWICH– Computes robust standard errors using sandwich estimators. SPC– Conducts statistical process control for quality monitoring.
Regression and Predictive Modeling	BART– Implements Bayesian Additive Regression Trees for flexible modeling. GAMMOD– Fits Generalized Additive Models for flexible regression analysis. GAMSELECT – Selects optimal GAM models using variable selection techniques. LMIXED– Fits linear mixed models for hierarchical data analysis. LOGSELECT– Selects logistic regression models using variable selection. NLMOD– Fits nonlinear models for complex relationships. PHSELECT– Selects proportional hazards models for survival analysis. PLSMOD– Implements partial least squares regression for predictive modeling. QTRSELECT– Selects quantile regression models. REGSELECT– Performs variable selection for regression models.
Clustering and Dimensionality Reduction	EFA– Performs exploratory factor analysis to identify latent structures. ICA– Conducts Independent Component Analysis for feature extraction. KCLUS– Performs k-means clustering for data segmentation. MBC – Implements model-based clustering for data segmentation. NMF– Performs nonnegative matrix factorization for dimensionality reduction. PCA– Conducts principal component analysis for dimensionality reduction. VARREDUCE– Reduces variable dimensionality for efficient modeling.

Data Preparation and Feature Analysis	BINNING– Groups continuous variables into discrete bins for analysis. CARDINALITY– Analyzes categorical variable cardinality to optimize model performance. CATTRANSFORM– Transforms categorical variables using binning and encoding techniques. CORRELATION – Computes correlation coefficients between variables. FREQTAB– Generates frequency tables for categorical data. MODELMATRIX– Creates model matrices for regression and machine learning. PARTITION– Splits datasets into training and validation subsets. VARIMPUTE– Imputes missing values using various techniques.
Causal Inference and Econometrics	CAEFFECT– Estimates causal effects in observational studies. SUPERLEARNER– Combines multiple models for improved predictive accuracy.
Optimization and Model Selection	GENSELECT– Implements genetic algorithms for model selection.
Decision Trees and Rule Based Learning	TREESPLIT– Splits decision trees for classification and regression.
Simulations and Systems Analysis	SIMSYSTEM– Simulates complex systems for scenario analysis.

SAS Visual Text Analytics Procedures

This Cheat Sheet shows SAS Visual Text Analytics procedures in SAS Viya 4

BOOLRULE– Extracts Boolean rules from large-scale transactional data.

BUILDINDEX - Constructs an index that is based on a specified schema.

GETINDEXSCHEMA - Gets the schema from an existing index table.

IDGENERATION - Generates unique identifiers (IDs) for an input data table.

IDVALIDATION - Validates the uniqueness of each identifier (ID) in an input data table.

LANGUAGEID - Performs language identification tasks on textual data in a specified input table.

RELATEDTERMS – Identifies terms that are closely related to a specified term in a data set.

SEARCHINDEX - Searches for a query against an index table and retrieves records that are relevant to that query.

TERMMAP – Extracts and visualizes patterns and relationships within textual or transactional data.

TEXTBERT - Performs text classification using deep learning transformer models.

TEXTCATEGORY - Provides development packages for building category models.

TEXTCATSCORE - Scores text documents against a category model.

TEXTCONCEPT - Validate language interpretation for textual information (LITI) syntax and to compile an LI binary from a table of LITI rules.

TEXTCRF - Trains conditional random field (CRF) models by using labeled text data.

TEXTCRFSCORE - Enables you to score new text data by using a trained conditional random field (CRF) model.

TMCOOCCUR - Calculates a co-occurrence association score for selected pairs of terms in the document collection.

SAS Visual Text Analytics PROC

This Cheat Sheet shows SAS Visual Text Analytics procedures in SAS Viya 4

TEXTMINE - Integrates natural language processing and statistical analysis to analyze large-scale textual data in SAS Viya.

TEXTPROFILE - Describes statistical characteristics of textual data in order to provide a profile of a data set and enable comparisons between data sets.

TEXTRULE - Generates language interpretation for textual information (LITI) concept or fact rules by using annotated data.

TEXTSENTIMENT - Validates language interpretation for textual information (LITI) syntax and to compile an LI binary from a table of LITI rules.

TEXTSENTSCORE - Uses a language interpretation for textual information (LITI) model to score text in an input data table.

TEXTSUMMARY - Uses natural language processing (NLP) techniques to summarize a document by selecting representative sentences from within the document to become the summary.

TMSCORE - Scores textual data in SAS Viya.

TRANSCRIPT - transcribes audio input in Waveform Audio File (WAV) format by using the acoustic model and the language model that you specify.

TRANSCRIPTERROR - Matches the hypothetical word sequences from an input hypothesis table to the true word sequences from an input reference table in order to measure the performance of a speech recognition system.