FIVE QUESTIONS FOR SELECTING A STATISTICAL METHOD

How many variables do you have?

1. How many variables do you have --- just one, more than one, or way too may?
2. What is your statistical objective – describe, identify/classify, compare/test, predict, or explain?
3. What scales of measurement do the variables use?
4. Are there dependent and independent variables?
5. Are the samples autocorrelated by location (1D, 2D, or 3D) or time/sequence (1D)?

JUST ONCE

DESCRIBE

Descriptive Statistics
Discrete Scales
Counts, proportions
Continuous Scales
Means, standard deviations, medians, ranges, quartiles
Distribution Fitting Statistics:
Skewness, kurtosis coefficient of variation
Graphics:
Histograms, box plots, probability plots
Tests:
Kolmogorov-Smirnov, Anderson-Darling, Shapiro-Wilk Lillifors

CLASSIFY

Filtering
Sorting
Cluster analysis

COMPARE

One-Population Test
Discrete Scale
Nonparametric statistics
Continuous Scales
t-tests, control charts

PREDICT

[none]

EXPLAIN

[none]

WAY TOO MANY

Conduct a cluster analysis to select representative variables
or
Conduct a principal components analysis, factor analysis, correspondence analysis, or multidimensional scaling to reduce the number of variables needed to represent most of the variability

MORE THAN ONE

Are some variable dependent and others independent?

No → (DESCRIBE / CLASSIFY / COMPARE / PREDICT / EXPLAIN branch)

Yes → **Are the dependent variables autocorrelated?**

No

DESCRIBE

Same as for single variables
Cross-tabulations
Correlations
Discrete Scales
Spearman R. Kendall Tau, gamma
Continuous Scales
Pearson Product Moment Correlation

CLASSIFY

Cluster analysis, discriminant analysis

COMPARE

Same as for single variables
Multi-Population Tests
ANOVA, ANCOVA, nonparametric tests

PREDICT

Discrete scale dependent variable Logistic regression, classification trees, discriminant analysis
Continuous scale dependent variables
Regression

EXPLAIN

Discrete scale dependent variable Logistic regression classification trees, discriminant analysis
Continuous scale dependent variables
Regression, canonical correlation

(No — right branch)

DESCRIBE

Same as for single variables
Correlations
Discrete Scales
Spearman R. Kendall Tau, gamma
Continuous Scales
Pearson Product Moment Correlation

CLASSIFY

Same as for Single variables

COMPARE

Same as for Single variables

PREDICT

[none]

EXPLAIN

Cluster analysis, multi dimensional scaling, principal components analysis, correspondence analysis

Yes

Time or location dependent?

LOCATION DEPENDENT

DESCRIBE

Same as for non-autocorrelated variables with a location-dependent variable as a grouping factor

CLASSIFY

Same as for non-autocorrelated variables with a location-dependent variable as a grouping factor

COMPARE

Same as for non-autocorrelated variables with a location-dependent variable as a grouping factor

PREDICT

Smoothing interpolation, trend-surfaces, geostatistics (variogramming and kring)

EXPLAIN

Trend-surfaces geostatistics (variogramming and kriging)

TIME DEPENDENT

DESCRIBE

Same as for non-autocorrelated variables with a time-dependent variable as a grouping factor

CLASSIFY

Same as for non-autocorrelated variables with a time-dependent variable as a grouping factor

COMPARE

Same as for non-autocorrelated variables with a time-dependent variable as a grouping factor

PREDICT

Smoothing interpolation, time-series regression, ARIMA, spectral analysis, neural networks

EXPLAIN

Time-series regression, ARIMA, spectral analysis