WINKS FOR DATA ANALYSIS

Read each table from left to right.

Descriptive Statistics		
	Data Type	Procedure
Describing one variable	Normal	Descriptives – Mean, SD, Min, Max, etc
	Not Normal	Descriptives, Median, Histogram, Stem and Leaf
	Categorical	Frequencies
	Over Time	Line Plot/Time Series
Describing two related variables	Normal	Pearson's Correlation
	Not Normal	Spearman's Correlation
	Categorical	Crosstabulations

Relational Analyses (Correlation and Regression)		
	Data Type	Procedure
You want to analyze the relationship between two variables.	Normal	Pearson Correlation, Simple Linear Regression
	Not Normal	Spearman Correlation
	Categorical	Contingency Coefficient
	Mixed	Spearman Correlation
Analyze the relationship between three or more variables.	Normal	Multiple Regression
	Not Normal	Kendall Partial Rank
	Categorical	Discriminant Analysis (not in SDA)

Comparison Tests (t-test/ANOVA)		
	Data Type	Procedure
Comparing a SINGLE SAMPLE to a norm (gold standard)	Normal	Single Sample t-test
	Not Normal	Sign Test
	Categorical	Goodness-of-Fit
Comparing two groups – Samples PAIRED	Normal	Paired t-test
	Not Normal	Wilcoxon
	Categorical	McNemar
Comparing two groups – Samples INDEPENDENT	Normal	Ind. Gp. t-test
	Not Normal	Mann-Whitney
	Categorical	Chi-Square
More than two groups - REPEATED MEASURES	Normal	Rep. Measures ANOVA
	Not Normal	Friedman ANOVA
	Categorical	Cochran's Q
More than two groups – INDEPENDENT	Normal	One-Way ANOVA
	Not Normal	Kruskal-Wallis
	Categorical	Chi-Square

www.texasoft.com

TEXASOFT SOFTWARE

WINKS SDA STATISTICAL DATA ANALYSIS

regression

t-tests

correlation

BASIC EDITION REFERENCE

Survival Analysis

Logistic Regression

CONTINGENCY TABLES

STATISTICAL ANALYSIS

ANOVA

graphis

Analysis

write-up

WINKS STATISTICAL DATA ANALYSIS

Opening Data Files

SDA recognizes the following file types:

⇒ **SDA** – native data set format.

⇒ **DBF- dBase/KWIKSTAT/WINKS** data set format Open using File/Open Dataset and select "Files of Type" as .dbf (Data files from previous WINKS versions.)

⇒ **XLS – Excel spreadsheet format.** Open using File/Open Dataset and select "Files of Type" as .xls (Excel 97 or newer.)

⇒ **CVS (TXT or DAT) –** ASCII comma delimited data sets. To be opened/imported into SDA, data sets in the XLS or CVS formats *should include a first row (record) containing a list of the variable names.*

Analysis of Variance

One-Way ANOVA
(Analyze/t-Tests and ANOVA – Independent Group t-test/ANOVA)

A comparison of means where subjects in each group are different (independent):

Ho: The difference in the means of the groups is zero. ($\mu_1 = \mu_2 = \ldots = \mu_n$)

Ha: The difference in the means of the groups is not zero. (At least one $\mu_i \neq \mu_j$)

Repeated measures ANOVA
(Analyze/t-Tests and ANOVA – Paired Rep. Measures)

Used for three or more repeated measures; a generalization of the paired t-test.

Chi-Square Analysis

(Analyze/Crosstabs, Frequencies, Chi-Square/ Crosstabulations/ Chi-Square)

Ho: The variables are independent of each other. (There is no association between them.)

Ha: The variables are not independent of each other.

Or, an hypothesis of homogeneity can be performed (math is the same.)

t-tests

Independent Group t-test
(Analyze/t-Tests and ANOVA – Independent Group t-test/ANOVA)

Independent group analysis is appropriate when observations are taken from groups in which subjects in one group do not appear in another group. You are testing the hypotheses:

Ho: The difference in the means of the groups is zero. ($\mu_1 = \mu_2$)
Ha: The difference in the means of the groups is not zero. ($\mu_1 \neq \mu_2$)

Paired t-test
(Analyze/t-Tests and ANOVA – Paired Rep. Measures (t-test/ANOVA))

A paired t-test is performed when there are two repeated (paired) measures from the same or matched subjects The hypotheses being tested are:

Ho: The mean of the differences is zero. ($\mu_{difference} = 0$)
Ha: The mean difference is not zero. ($\mu_{difference} \neq 0$)

Single Sample t-test Analysis
(Analyze/t-Tests and ANOVA – Single Sample t-test)

The single sample analysis tests a specified mean differs from an hypothesized mean. The hypotheses are:

Ho: The mean equals the hypothesized value.
Ha: The mean does not equal the hypothesized value.

Nonparametric Procedures

Nonparametric procedures are useful as an alternative to parametric tests when you suspect that the data are not normal or that variances between groups are unequal.

Parametric test	*Nonparametric*
Ind. Group t-test	Mann-Whitney
One-Way ANOVA	Kruskal-Wallis
Paired t-test	Wilcoxon Signed Rank
Repeated Measures ANOVA	Friedman's ANOVA
Pearson's r	Spearman's r

Correlation

Correlation is used to measure the strength of association between two variables. It is measured using r (Pearson's correlation.) Where r ranges from –1.0 to 1.0. A t-test for significance uses the hypotheses (where *rho* is the population correlation.)

Ho: rho = 0
Ha: rho ≠ 0

Linear Regression

(Analyze/Regression and Correlation/Simple Linear Regression or Multiple Regression Analysis)

Simple linear regression is used for predicting a value of a dependent variable using an independent variable. Hypotheses tested are:

Ho: $\beta = 0$ (The slope is zero; no linear relationship.)
Ha: $\beta \neq 0$ (The slope is not zero; is a linear relationship.)

Multiple linear regression is used to predict values of a dependent variable using two or more independent variables. The hypothesis are:

Ho: $\beta_1 = \beta_2 = \ldots \beta_k$ (No relationship between dependent variable and the set of independent variables.)

Performing & Interpreting a Statistical Test

1. **State a null hypothesis** (Ho) (and usually an alternative hypothesis (Ha)).
2. **Perform an analysis** to test the hypothesis (the statistical test).
3. **Interpret the test** and make a decision, using a decision criterion (usually the p-value) based on the assumption that the null hypothesis has been satisfied.

Typically, the reported p-value is the most convenient way to decide on statistical significance. For most cases, if the p-value is less than 0.05 ($p < 0.05$), you reject the null hypothesis in favor of the alternative.