Descriptive Statistics

1) Dotplot

Store data in C1 and name C1 'data'
Graph → Dotplot… →
Variables: data
/ No grouping
→ OK

2) Stem and leaf display

Store data in C1 and name C1 'data'
Graph → Stem-and-Leaf… →
Variables: data
→ OK

3) Histogram
   a) Frequency histogram

   Store data in C1 and name C1 'data'
   Graph → Histogram… →
   Graph Variables: data
   → OK

   b) Percentage histogram (Note: *percentage* = *relative frequency* × 100)

   Store data in C1 and name C1 'data'
   Graph → Histogram… →
   Graph Variables: data
   → Options… →
   Type of Histogram:
   / Percent
   → OK → OK

4) Mean, median, standard deviation, minimum, maximum, lower quartile, upper quartile

Store data in C1 and name C1 'data'
Stat → Display Descriptive Statistics… →
Variables: data
→ OK

Confidence Intervals and Hypothesis Tests

1) One population
   a) Parameter: Mean
      i) Large sample size (i.e. $n \geq 30$)

Store data in C1 and name C1 'data'
Stat → Basic Statistics → 1-Sample Z… →
Variables: data
Sigma: s
Test mean: μ_0
→ Options… →
Confidence level: $100(1-\alpha)$ e.g. 95.0
Alternative: less than / not equal / greater than
→ OK → OK

      ii) Small sample size (i.e $n < 30$)

Store data in C1 and name C1 'data'
Stat → Basic Statistics → 1-Sample t… →
Variables: data
Test mean: μ_0
→ Options… →
Confidence level: $100(1-\alpha)$ e.g. 95.0
Alternative: less than / not equal / greater than
→ OK → OK

   b) Parameter: Proportion

Stat → Basic Statistics → 1 Proportion… →
/ Summarized data
Number of trials: n or denominator of $\hat{p}$ (before simplified)
Number of successes: numerator of $\hat{p}$ (before simplified)
→ Options… →
Confidence level: $100(1-\alpha)$ e.g. 95.0
Test proportion: p_0
Alternative: less than / not equal / greater than
☑ **Use test and interval based on normal distribution**
→ OK → OK

2) Two populations
   a) Parameter: Mean
      i) Independent Samples
         (1) Small sample size (i.e. $n < 30$)

Store data in C1, C2 and name C1 'data1', C2 'data2'
Stat → Basic Statistics → 2-Sample t… →
/Samples in different columns
First: data1
Second: data2
☑ **Assume equal variances**
→ Options… →
Confidence level: $100(1-\alpha)$ e.g. 95.0
Test mean: D_0 e.g. 0.0
Alternative: less than / not equal / greater than
→ OK → OK

      ii) Paired Samples
         (1) Small sample size (i.e $n < 30$)

Store data in C1, C2 and name C1 'data1', C2 'data2'
Stat → Basic Statistics → Paired t… →
First sample: data1
Second sample: data2
→ Options… →
Confidence level: $100(1-\alpha)$ e.g. 95.0
Test mean: D_0 e.g. 0.0
Alternative: less than / not equal / greater than
→ OK → OK

   b) Parameter: Proportion

Stat → Basic Statistics → 2 Proportions… →
/ Summarized data

	Trials:	**Successes:**
First sample:	n_1 or denominator of $\hat{p}_1$	numerator of $\hat{p}_1$
Second sample:	n_2 or denominator of $\hat{p}_2$	numerator of $\hat{p}_2$

(before simplified)

→ Options… →
Confidence level: $100(1-\alpha)$ e.g. 95.0
Test difference: D_0 e.g. 0.0
Alternative: less than / not equal / greater than
☑ **Use pooled estimate of p for test** (if test difference is 0.0)
→ OK → OK

χ^2 - Test of Independence

Store data in C1, C2, …
Stat → Tables → Chi square Test… →
Columns containing the table: C1 C2 …
→ OK

Example:
Given the following contingency table. Test if gender and smoking habit are related.

	Smokes?	
	No	Yes
Female	27	8
Male	37	19

Store data in C1, C2 and name C1 'no', C2 'yes'

	C1	C2
	no	yes
1	27	8
2	37	19

Stat → Tables → Chi square Test… →
Columns containing the table: no yes
→ OK

Chi-Square Test: no, yes

```
Expected counts are printed below observed counts

            no      yes    Total
    1       27        8       35
         24.62    10.38

    2       37       19       56
         39.38    16.62

Total       64       27       91

Chi-Sq =  0.231 +  0.548 +
          0.144 +  0.342 = 1.265
DF = 1, P-Value = 0.261
```