

# Commandes SPSS associées aux outils statistiques

## Une variable qualitative (VD)

### Analyse descriptive : Résumé numérique

- Mesures de position : Mode et Médiane (qualitative ordinale)
  - *Analyze* → *Descriptive Statistics* → *Frequencies (Statistics)*
  - *Analyze* → *Reports* → *Case Summaries* (décocher «*Display cases*»)
  - *Analyze* → *Descriptive Statistics* → *Explore*
- Distribution : Tableau de fréquences (+ cumulés si ordinale)
  - *Analyze* → *Descriptive Statistics* → *Frequencies*

### Analyse descriptive : Résumé graphique

- Diagramme en barres
  - *Graphs* → *Legacy Dialogs* → *Bar (Simple)*

### Inférence (non-paramétrique)

- Variable à deux niveaux ( $k=2$ ): Test sur une proportion
  - *Analyze* → *Non Parametric Tests* → *Binomial*
  - *Analyze* → *Non Parametric Tests* → *Chi Square*
- Variable à plus de deux niveaux ( $k \geq 2$ ) : Test d'ajustement  $\chi^2$ 
  - *Analyze* → *Non Parametric Tests* → *Chi Square*

## Une variable quantitative (VD)

### Analyse descriptive : Résumé numérique

- Mesures de position : Moyenne, mode, médiane
  - *Analyze* → *Reports* → *Case Summaries* (décocher «*Display cases*»)
  - *Analyze* → *Descriptive Statistics* → *Explore*
  - *Analyze* → *Descriptive Statistics* → *Descriptives*
- Mesures de dispersion : Étendue, EIQ, variance, écart-type, CV
  - *Analyze* → *Reports* → *Case Summaries* (décocher «*Display cases*»)
  - *Analyze* → *Descriptive Statistics* → *Explore*
  - *Analyze* → *Descriptive Statistics* → *Descriptives*
- Distribution : Tableau de fréquence + cumulés (éventuellement par classes)
  - *Analyze* → *Descriptive Statistics* → *Frequencies*

### Analyse descriptive : Résumé graphique

- Histogramme - visualiser distribution
  - *Graphs* → *Legacy Dialogs* → *Histogram*
  - *Graphs* → *Chart Builder* → *Histogram*
- Box plot
  - *Graphs* → *Legacy Dialogs* (→ *Interactive*) → *Boxplot*
  - *Graphs* → *Chart Builder* → *Boxplot*
- Diagramme en barres (si quantitative discrète)
  - *Graphs* → *Legacy Dialogs* → *Bar (Simple)*
- Analyse de la normalité : QQ plot
  - *Analyze* → *Descriptive Statistics* → *Q-Q Plots*

### Inférence (paramétrique)

- Test t et intervalle de confiance sur une moyenne
  - *Analyze* → *Compare Means* → *One-Sample T test*
- Test  $\chi^2$  et intervalle de confiance sur une variance
  - *Non disponible avec le module de base de SPSS*

## Deux variables qualitatives

### Analyse descriptive : Résumé numérique

- Table de contingence
  - *Analyze* → *Descriptive Statistics* → *Crosstabs*
- V et Phi de Cramer
  - *Analyze* → *Descriptive Statistics* → *Crosstabs*

### Analyse descriptive : Résumé graphique

- Diagramme en barres par catégories
  - *Graphs* → *Legacy Dialogs* → *Bar (Clustered/Stacked)*

### Inférence (non paramétrique)

- Test de comparaison de deux proportions ou d'homogénéité
  - *Analyze* → *Descriptive Statistics* → *Crosstabs (Statistics:Chi-Square)*
- Test d'indépendance Chi2
  - *Analyze* → *Descriptive Statistics* → *Crosstabs (Statistics:Chi-Square)*

## Une variable quantitative VD et une variable qualitative VI

### Analyse descriptive : Résumé numérique

- Indices de position et dispersion par catégories de la variable qualitative
  - *Analyze* → *Reports* → *Case Summaries (décocher «Display cases»)*
  - *Analyze* → *Descriptive Statistics* → *Explore*

### Analyse descriptive : Résumé graphique

- Box plot par catégories
  - *Graphs* → *Chart Builder* → *Boxplot*
  - *Analyze* → *Descriptive Statistics* → *Explore*

### Inférence quand la variable qualitative (VI) a deux niveaux : k=2

- Groupes indépendants – tests paramétriques
  - Test t de comparaison de deux moyennes (variances égales ou différentes)
    - *Analyze* → *Compare Means* → *Independent-Samples T Test*
  - Test F de comparaison de 2 variances
    - *Non disponible avec le module de base de SPSS*
  - Test de Levene de comparaison de 2 variances
    - *Analyze* → *Compare Means* → *Independent-Samples T Test*
- Groupes indépendants – tests NON paramétriques
  - Test de la somme des rangs de Wilcoxon ou test de Mann-Whitney pour comparer deux valeurs centrales
    - *Analyze* → *Nonparametric Tests* → *2 Independent Samples (Mann-Whitney)*
- Données paires – tests paramétriques
  - Test t de comparaison de deux moyennes pour données paires
    - *Analyze* → *Compare Means* → *Paired Samples T Test*
- Données paires – tests NON paramétriques
  - Test des rangs de Wilcoxon pour données paires
    - *Analyze* → *Nonparametric Tests* → *2 Related Samples (Wilcoxon)*
  - Test du signe
    - *Analyze* → *Nonparametric Tests* → *2 Related Samples (Sign)*

## Inférence quand la variable qualitative (VI) a deux niveaux ou plus : $k \geq 2$

- Groupes indépendants - outils paramétriques
  - ANOVA I pour la comparaison des moyennes des traitements
    - *Analyze* → *Compare Means* → *One-Way ANOVA*
  - Tests de comparaisons multiples (Bonferroni, Tukey, Newman-Keuls)
    - *Analyze* → *Compare Means* → *One-Way ANOVA (Post Hoc)*
  - Tests sur des contrastes
    - *Analyze* → *Compare Means* → *One-Way ANOVA (Contrasts)*
  - Test de Levene de comparaison de variances
    - *Analyze* → *Compare Means* → *One-Way ANOVA (Options)*
  - ANOVA I avec correction de Welch en cas de variances non homogènes
    - *Analyze* → *Compare Means* → *One-Way ANOVA (Options)*
- Groupes indépendants - outils NON paramétriques
  - Analyse de la variance à un critère de classification de Kruskal-Wallis
    - *Analyze* → *Nonparametric Tests* → *k Independent Samples (Kruskal-Wallis)*
- Données répétées - outils paramétriques
  - ANOVA I pour mesures répétées
    - *Analyze* → *General Linear Models* → *Repeated Measures*
  - Tests de comparaisons multiples (Bonferroni)
    - *Analyze* → *General Linear Models* → *Repeated Measures (Options – Compare main effects)*
- Données répétées - outils NON paramétriques
  - Test des rangs de Friedman pour k échantillons dépendants
    - *Analyze* → *Nonparametric Tests* → *k Related Samples (Friedman)*

## Explication-modélisation

L'ANOVA I peut être considérée comme un modèle où on "explique" la variable Y à partir du niveau pris par X (voir chapitre sur les modèles GLM)

## Deux variables quantitatives

### Analyse descriptive : Résumé numérique (2 VD)

- Coefficient de corrélation de Pearson, Spearman
  - *Analyze* → *Correlate* → *Bivariate*
- Coefficient de corrélation partielle
  - *Analyze* → *Correlate* → *Partial*

### Analyse descriptive : Résumé graphique (2VD ou 1VD et 1VI)

- Graphe X-Y - Scatter Plot
  - *Graphs* → *Legacy Dialogs* → *Interactive* → *Scatterplot*

### Inférence (paramétrique) (2VD)

- Test sur un coefficient de corrélation de Pearson, Spearman
  - *Analyze* → *Correlate* → *Bivariate*
- Test sur un coefficient de corrélation partielle
  - *Analyze* → *Correlate* → *Partial*

### Explication-modélisation (1VD et 1VI)

- Régression linéaire
  - Tests et intervalles de confiance sur les coefficients d'une droite de régression
    - *Analyze* → *Regression* → *Linear (Statistics)*
  - Prédiction et intervalle de confiance sur la réponse moyenne prédite et intervalle de prédiction
    - *Analyze* → *Regression* → *Linear (Save – Prediction intervals)*
  - Outils pour valider les hypothèses (QQ plot résidus, graphe des résidus)

- *Analyze → Regression → Linear (Plots)*
  - *(Plots: X=ZPRED Y=ZRESID)*
  - *(Plots: Normal probability plot)*

## **2 variables quantitatives et une variable qualitative**

### Analyse descriptive : Résumé numérique

- Coefficient de corrélation par niveau de la variable catégorielle
  - *Data → Split File (var catég) + Analyze → Correlate → Bivariate*

### Analyse descriptive : Résumé graphique

- Graphe X-Y avec des symboles différents pour les niveaux de la variable qualitative
  - *Graphs → Legacy Dialogs → Scatterplot (Simple Scatter)*

### Inférence quand la variable qualitative a deux niveaux

- Test d'égalité de deux coefficients de corrélation
  - *Non disponible avec le module de base de SPSS*
- Test de comparaison de deux pentes de droites de régression
  - *Non disponible avec le module de base de SPSS*

### Explication-modélisation

- Modélisation GLM ou analyse de covariance pour expliquer une variable quantitative à partir d'une variable qualitative et une variable quantitative
  - *Analyze → General linear model → univariate*

## **2 variables qualitatives A et B et une variable quantitative X**

### Analyse descriptive : Résumé numérique

- Mesures de position et dispersion par combinaisons des niveaux de A et B
  - *Analyze → Reports → Case Summaries (décocher «Display cases»)*

### Analyse descriptive : Résumé graphique

- Box plot par combinaisons des niveaux de A et B
  - *Graphs → Legacy Dialogs → Boxplot (Clustered)*

### Inférence - groupes indépendants

- ANOVA II pour tester les effets principaux et l'effet d'interaction des deux variables qualitatives A et B sur la variable quantitative X
  - *Analyze → General Linear Models → Univariate*
- Autres outils disponibles avec ANOVA I (comparaisons multiples, contrastes, test de Levene ...)
  - *Analyze → General Linear Models → Univariate (Post Hoc – Contrasts – Options)*

### Inférence - plusieurs traitements par individu

- Modèle pour mesures répétées avec un facteur within et une facteur between ou avec deux facteurs within
  - *Analyze → General Linear Models → Repeated Measures*

## k variables quantitatives

### Analyse descriptive : Résumé graphique

- Graphe X-Y matriciel (scatter plot matrix)
  - *Graphs → Legacy Dialogs → Scatterplot (Matrix Scatter)*

### Relation, Visualisation par projection

- Analyse en composantes principales
  - *Analyze → Dimension reduction → Factor*

### Recherche de groupes d'individus

- Clustering
  - *Analyze → Classify → Hierarchical cluster*

### Explication-modélisation

- Régression linéaire multiple
  - Tests et intervalles de confiance sur les coefficients d'une droite de régression
    - *Analyze → Regression → Linear (Statistics)*
  - Prédiction et intervalle de confiance sur la réponse moyenne prédite et intervalle de prédiction
    - *Analyze → Regression → Linear (Save – Prediction intervals)*
  - Outils pour valider les hypothèses (QQ plot résidus, graphe des résidus)
    - *Analyze → Regression → Linear (Plots)*
      - *(Plots: X=ZPRED Y=ZRESID)*
      - *(Plots: Normal probability plot)*
  - Outils de détection d'outliers et de points influents : leverage et statistique de Cook
    - *Analyze → Regression → Linear ( Save: Préd unstd & Leverage) + Scatter/Dot: X=Préd unstd Y=Lev)*
  - Indice VIF pour mesurer la colinéarité
    - *Analyze → Regression → Linear (Statistics)*