

TD1 Commandes Statistica

La version 10 de Statistica propose deux types d'affichage des menus : l'affichage « classique » (analogue à celui des versions antérieures) et l'affichage « ruban ». On peut passer de l'un à l'autre dans le menu **Affichage**.

Les documents de TP sont rédigés pour l'affichage classique¹.

Les commandes des analyses statistiques sont accessibles dans les menus **Statistiques** et **graphiques**.
Chaque fois que le logiciel demande si l'on veut poursuivre une analyse ou en démarrer une nouvelle, choisir d'en commencer une nouvelle.

1^{re} partie

3. Paramétrage du logiciel

Sélectionnez la commande **Options** dans le menu **Outils**. Dans la boîte de dialogue **Options**, sélectionner l'onglet **gestionnaire de Sorties**.

Sélectionner l'option **Classeur**, puis l'option

Même classeur que celui qui contient le fichier de données.

Cocher l'option **placer les résultats automatiquement dans le classeur**.

4. Manipulations de la feuille des données et du classeur

Créer le nouveau classeur à partir du menu **Fichier / Nouveau**. Ne pas cocher la case **Insérer une feuille de données vierge**.

Ouvrir la feuille de données à partir du menu **Fichier / Ouvrir**.

Pour insérer la feuille de données dans le classeur, sélectionner la commande **Insérer** dans le menu **Classeur** (ou bien par un clic-droit sur l'arborescence du classeur), puis cocher l'élément **Toutes les fenêtres** et cliquer sur **OK**.

Pour rendre la feuille de données active, faire un clic-droit sur son nom **salaires.sta** affiché dans l'arborescence du classeur, puis cliquer sur **Feuille de données active**. Lorsqu'elle est active, l'icône de la feuille est encadrée par un carré rouge.

5. Valeurs-texte

Pour la variable **minority** : dans la feuille de données, sélectionner la colonne **minority** en cliquant sur l'en-tête de la colonne ; dérouler le menu **Variables** et sélectionner **Valeurs-Texte** ; dans l'éditeur de **Valeurs-Texte**, taper **Non** dans le champ **Valeurs-Texte**, puis **0** dans le champ suivant **V. Numériques**. Saisir ensuite la valeur-texte **Oui**, puis traiter la variable **sex**.

Cliquer sur le bouton  pour afficher les valeurs-texte.

2^e partie : étude descriptive d'une variable

1. Pour la variable **minority**

(a) *Tableau des effectifs et des fréquences*

Ouvrir la boîte de dialogue **Tables de fréquences** par la commande **Statistiques/ Statistiques Élémentaires / Tables de fréquences** ; cliquer sur **OK** pour afficher la boîte. Sélectionner la variable **minority** (bouton **Variables**) ; cliquer sur **Synthèse : Tables de fréquences** pour afficher les résultats.

Catégorie	Table de fréquences : minority (salaires.sta)			
	Effectif	Effectifs Cumulés	%age	%age Cumulé
Non	370	370	78,05907	78,0591
Oui	104	474	21,94093	100,0000
VM	0	474	0,00000	100,0000

1. Certaines fenêtres ou boîtes de dialogues présentées dans le document peuvent être légèrement différentes à l'écran.

Pour n'afficher que les effectifs et les proportions : dans l'onglet *Options* désélectionner *Effectifs cumulés* et *Pourcentages cumulés*, ainsi que *Comptage et report des valeurs manquantes (VM)*.

(b) *Camembert et diagramme en barre*

– Pour obtenir le camembert : utiliser la commande *Graphiques / Graphiques en 2D / Camemberts (Diagr. circulaires)*.

Dans l'onglet *Avancé* : sélectionner la variable *minority* ; sélectionner ensuite *Légende du camembert : Texte et %age*, cocher les cases *Type : 2D* et *Forme : circulaire*, puis cliquer sur *OK*.

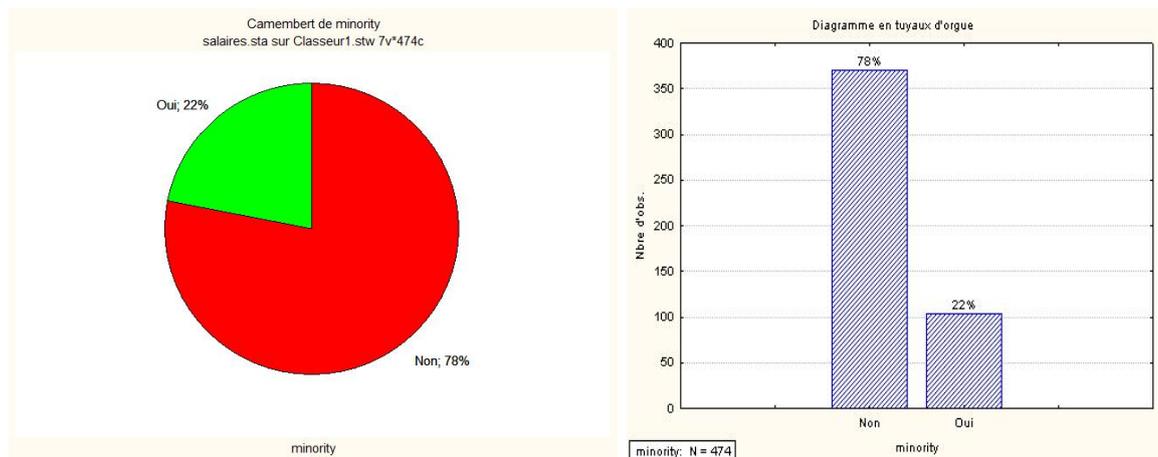
– Le diagramme en barres (ou en tuyaux d'orgue) d'une variable qualitative est obtenu à partir de la commande *Histogrammes*.

Utiliser la commande *Graphiques / Histogrammes...* puis cliquer sur l'onglet *Avancé* ; sélectionner la variable *minority* ;

sélectionner *Type de graphique : Simple*, *Type d'ajustement : Inactif*, cocher la case *Effectif total* pour l'afficher dans le graphique.

Cocher la case *Afficher les %ages* pour ajouter les pourcentages au dessus de chaque barre.

Cocher la case *Rupture entre les colonnes* pour séparer les barres du graphique, conformément à la représentation standard d'une variable qualitative. Cliquer sur *OK*.



2. *Résumés statistiques des variables quantitatives*

Ouvrir la boîte de dialogue *Statistiques descriptives* par la commande *Statistiques / Statistiques Élémentaires/Statistiques descriptives* ; cliquer sur *OK* pour afficher la boîte. Sélectionner les variables 1-5 (bouton *Variables*).

Dans l'onglet *Avancé*, cocher les résumés statistiques demandés : *Moyenne*, *Médiane*, *Ecart-type*, *Variance*, *Minimum & Maximum*, *1er et 3ème quartile* ; cliquer sur *Synthèse* pour afficher les résultats.

Remarque : le logiciel calcule la variance en divisant par $N - 1$ (formule-échantillon).

Variable	Statistiques Descriptives (salaires.sta)								
	N Actifs	Moyenne	Médiane	Minimum	Maximum	1er Quartile	3ème Quartile	Variance	Ecart-type
salary	474	34419,57	28875,00	15750,00	135000,0	24000,00	37050,00	291578214	17075,66
salbegin	474	17016,09	15000,00	9000,00	79980,0	12450,00	17490,00	61946945	7870,64
jobtime	474	81,11	81,00	63,00	98,0	72,00	90,00	101	10,06
prevexp	474	95,86	55,00	0,00	476,0	19,00	139,00	10938	104,59
educ	474	13,49	12,00	8,00	21,0	12,00	15,00	8	2,88

3. Graphiques des salaires actuels, variable salary

(a) Histogramme

L'histogramme fourni par Statistica représente par défaut la densité des effectifs de la variable. Les intervalles sont de même longueur. L'axe Y affiche l'échelle des effectifs. Si l'on choisit l'option % pour l'axe Y, les effectifs de l'échelle sont convertis en pourcentages.

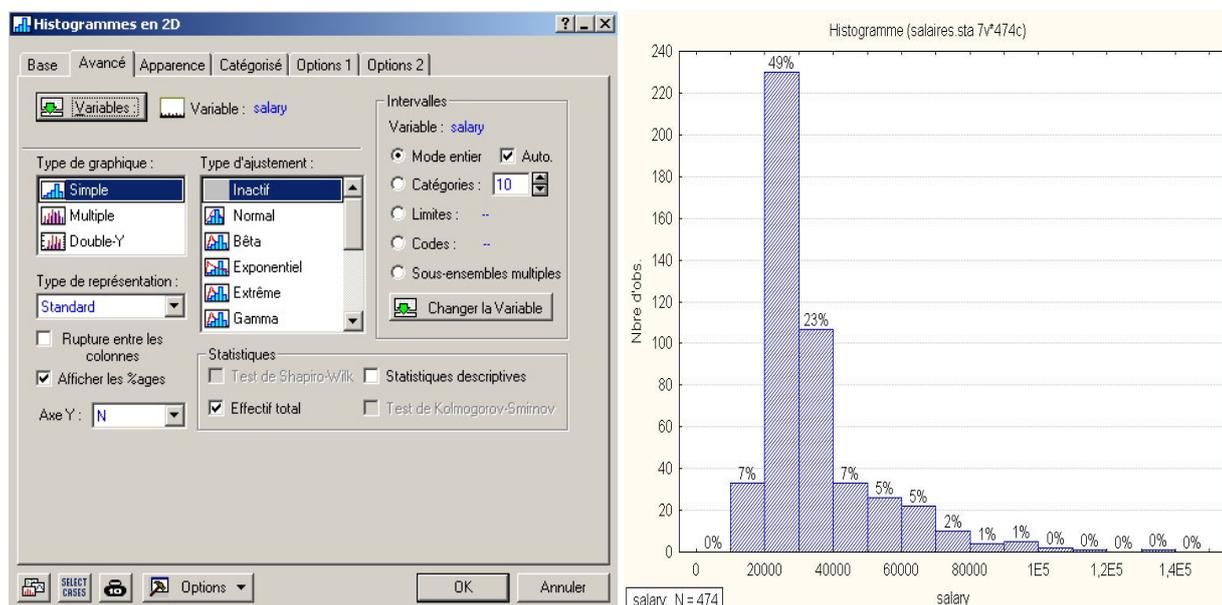
Pour faire l'histogramme :

utiliser la commande **Graphiques / Histogrammes...** puis sélectionner l'onglet **Avancé** ; sélectionner la variable **salary**, sélectionner **Type de graphique : Simple**, cocher la case **Effectif total** pour l'afficher dans le graphique.

Pour superposer la courbe de la loi normale, sélectionner **Type d'ajustement : Normal** . Pour ne pas afficher la courbe, sélectionner **Type d'ajustement : Inactif**.

Pour afficher les pourcentages au dessus de chaque rectangle, cocher la case **Afficher les %ages**.

Cliquer sur **OK**.



(b) Coefficients d'aplatissement et d'asymétrie

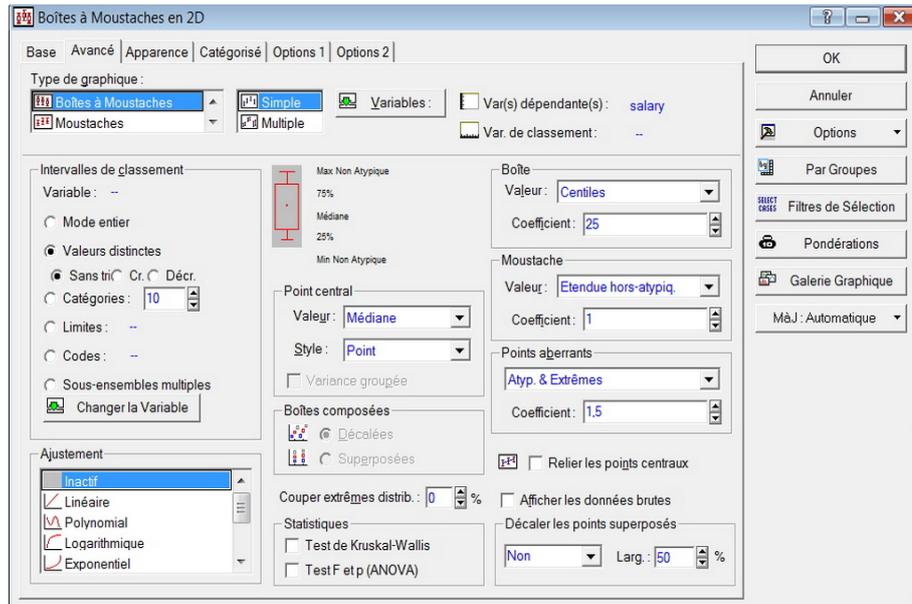
Dans l'onglet **Avancé** de la boîte de dialogue **Statistiques descriptives**, sélectionner la variable **salary** et cocher les cases **Asymétrie** et **Aplatissement**.

Le coefficient d'asymétrie est positif et grand : asymétrie prononcée avec un étalement à droite de la queue de la distribution.

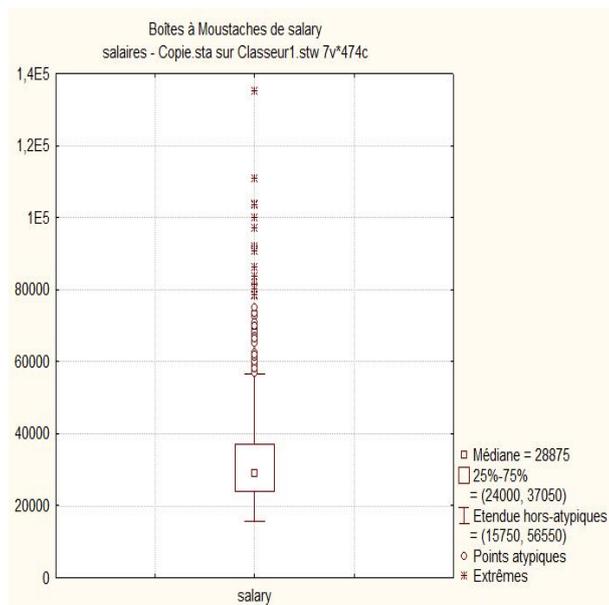
Le coefficient d'aplatissement est positif et grand : distribution plus pointue que la loi normale, avec des queues de distribution épaisses.

(c) *Boîte à moustaches des salaires*

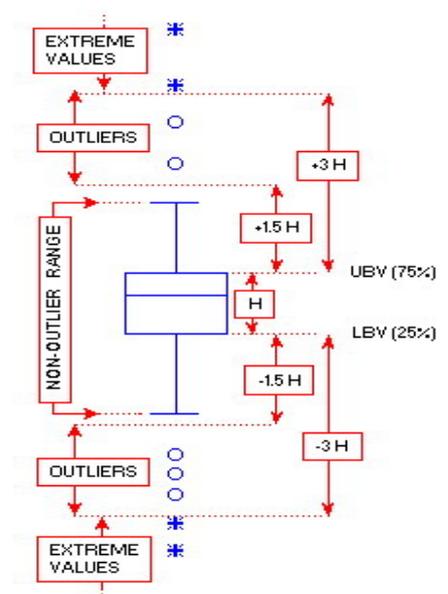
Sélectionner l'onglet **Avancé** de **Graphiques / Graphiques en 2D / Boîtes à Moustaches...** ; sélectionner la variable **salary** comme variable dépendante et les différentes options du graphique :



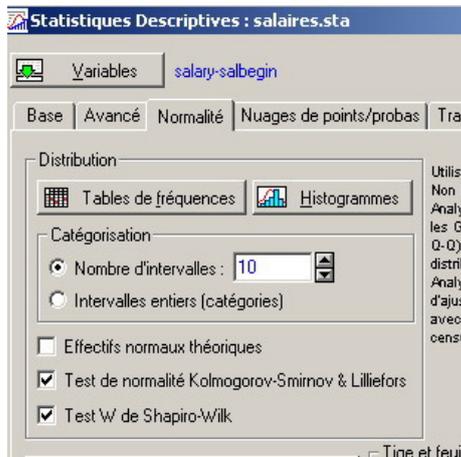
Boîte à moustaches :



Graphique d'aide de Statistica :



4. Tests de normalité

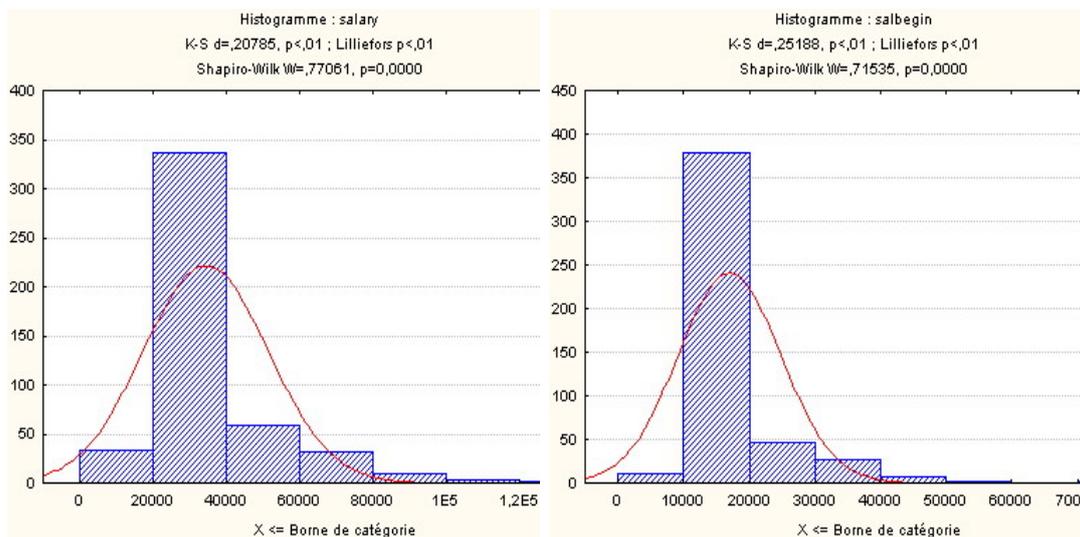


Sélectionner l'onglet Normalité du module Statistiques Descriptives.

Sélectionner les variables salary et salbegin.

Cocher la case Test de Kolmogorov-Smirnov & Lilliefors et la case Test W de Shapiro & Wilk.

Cliquer ensuite sur le bouton Histogrammes.



Dans *Statistica*, une autre procédure possible est de tracer l'histogramme des données. En option, on peut superposer la courbe de la loi normale et inclure les résultats des principaux tests d'ajustement à la loi normale (Kolmogorov-Smirnov ; Lilliefors ; Shapiro & Wilk).

3^e partie : étude des variables *sex* et *salary*

1. Résumés statistiques catégorisés Hommes et Femmes

Ouvrir la boîte de dialogue Statistiques par groupe (Décomposition) par la commande Statistiques / Statistiques Élémentaires / Décompositions & ANOVA à 1 facteur.

Sélectionner la variable *salary* comme variable dépendante et la variable *sex* comme variable de classement.

Cliquer sur OK pour ouvrir la boîte de dialogue Statistiques par groupe - Résultats.

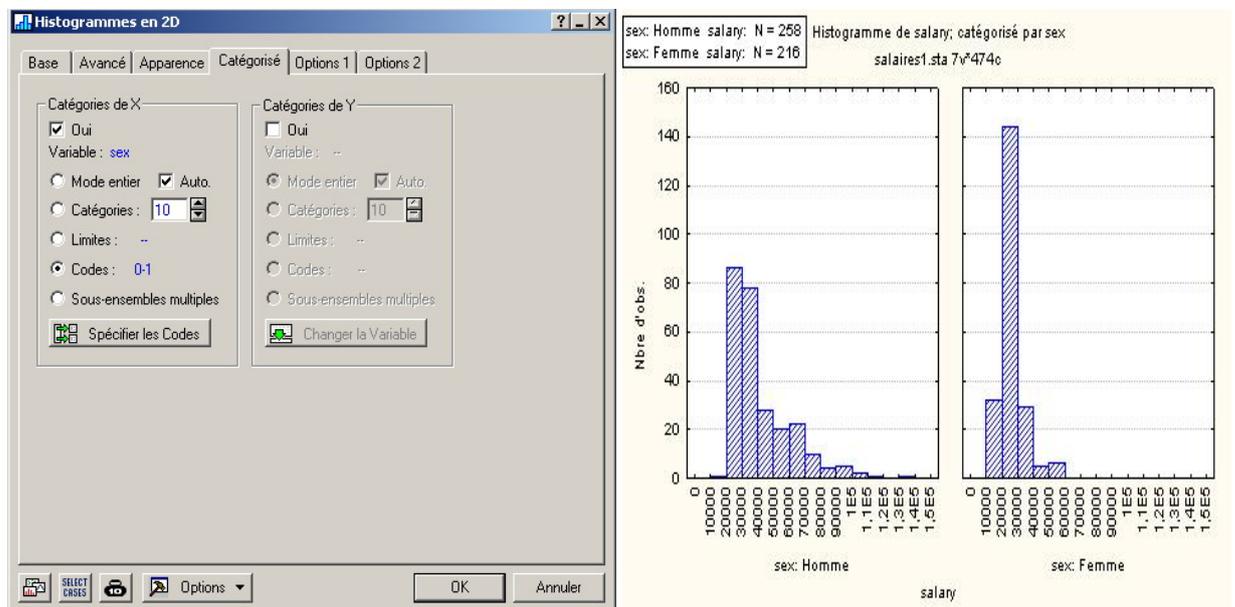
Dans l'onglet Stats descriptives, cocher les cases N actifs, Ecart-types, Médiane et quartiles. Cliquer sur Synthèse.

Statistiques Descriptives par Groupe N=474 (Aucune VM dans la liste des VD)						
sex	salary Moyennes	salary N	salary Ec-Type	salary 1er Q	salary Médiane	salary 3ème Q
Homme	41441,78	258	19499,21	28050,00	32850,00	50550,00
Femme	26031,92	216	7558,02	21525,00	24300,00	28500,00
TsGrpes	34419,57	474	17075,66	24000,00	28875,00	37050,00

2. Histogrammes catégorisés Hommes et Femmes

Il n'existe pas d'échelle de pourcentages commune pour les histogrammes multiples (catégorisés). Pour des effectifs totaux différents, on peut comparer les formes des histogrammes mais pas les hauteurs des rectangles.

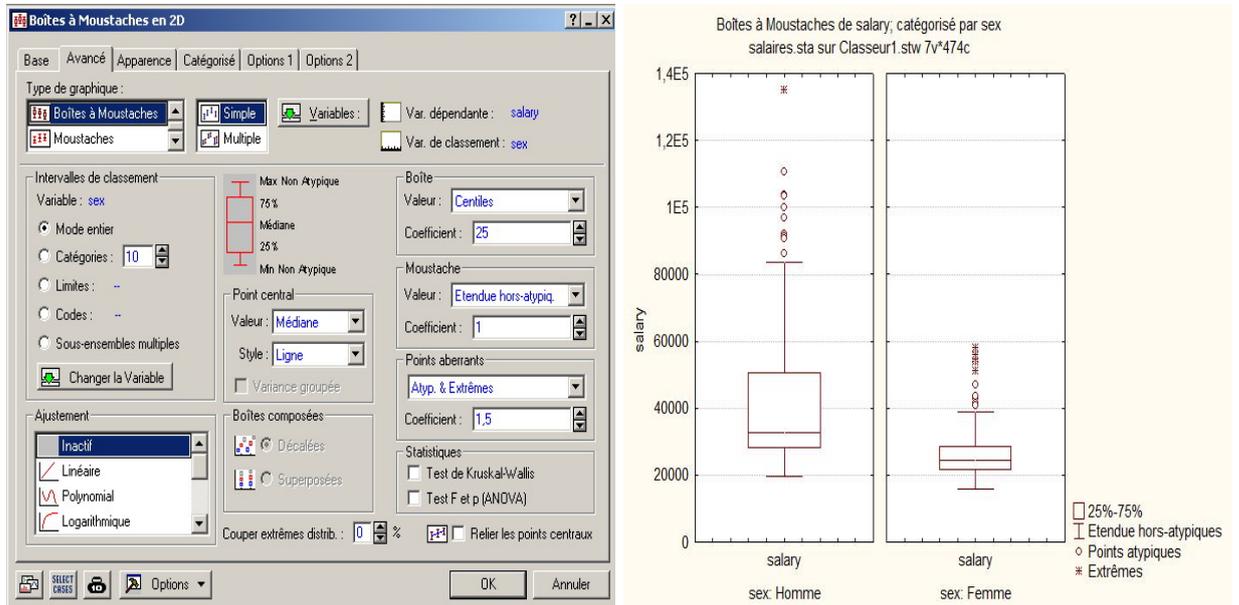
Définir la variable et le type de graphique comme dans la question 2^e partie 3. (a) ; puis, dans l'onglet Catégorisé sélectionner la variable sex :



Remarque : il est possible d'obtenir des histogrammes catégorisés (aucune option n'est disponible) à partir de la boîte de dialogue Statistiques par groupe utilisée dans la question précédente.

3. Boîtes à moustaches catégorisées Hommes et Femmes

Sélectionner l'onglet **Avancé** de **Graphiques / Graphiques en 2D / Boîtes à Moustaches...** ; sélectionner la variable **salary** comme variable dépendante, la variable **sex** comme variable de classement et les différentes options du graphique :



*Remarque : comme pour les histogrammes, il est possible d'obtenir des boîtes à moustaches catégorisées (sans options) à partir de la boîte de dialogue **Statistiques par groupe**.*