**L’emploi au féminin**

**Situation-problème 1 : Écart entre les salaires hommes-femmes**

Identification des variables :

$$x:salaire horaire moyen des hommes $$

$$y:salaire horaire moyen des femmes$$

**Utilisation d’un outil technologique**

Coefficient de corrélation : $r≈0,90$

La corrélation est forte et positive. L’estimation est fiable.

Droite de régression : $y=0,63x+12,5$

Salaire horaire moyen d’une avocate :

 $y=0,63\left(75\right)+12,5=59,75$

**59,75 $/h**

**Approche graphique**

L’élève trouve le coefficient de corrélation et la droite de régression avec la méthode du rectangle (voir le travail de l’élève).

Salaire horaire moyen, comparaison hommes/femmes

****

Salaire horaire moyen des femmes

Salaire horaire moyen des hommes

**Donnée éloignée**

La donnée (55,88 ; 47,71) sort du lot. Le coefficient de corrélation est plus faible si on refait l’analyse sans tenir compte de cette donnée.

**Analyse avec un outil technologique**

Coefficient de corrélation : $r≈0,71$

La corrélation est moyenne et positive. L’estimation est relativement fiable.

Droite de régression : $y=0,69x+10,48$

Salaire horaire moyen d’une avocate :

 $y=0,69\left(75\right)+10,48=61,95$

**61,95 $/h**

**Commentaires**

La comparaison des deux nuages (avec et sans la donnée éloignée) peut apporter des éléments de discussion :

* Extrapolation similaire dans les 2 cas mais une des estimations est moins fiable
* L’écart entre les estimations s’agrandit lorsque les données en « x » sont plus grandes
* L’écart entre les salaires hommes-femmes semble plus grand lorsque le salaire de l’homme est plus élevé (professeur d’université, avocat); on pourrait analyser cette corrélation
* Etc.

**Situation-problème 2 : Des études dans le domaine de la construction**

**Utilisation d’un outil technologique pour le coefficient de corrélation**

Coefficient de corrélation : $r≈0,85$

La corrélation est forte et positive. L’estimation est fiable.

Droite de régression : $y=16,89x+1472,54$

Pour 10 000 hommes, environ 505 femmes

**Démarche algébrique**

Par exemple, la méthode de Mayer pour trouver la droite de régression.

**1. Ordonner les couples en ordre croissant des abscisses**

|  |
| --- |
| **Nombre de diplômes émis pour les programmes d'étude pouvant mener à un****métier de la construction selon le sexe de l’élève, années de diplomation 1997 à 2011** |
| **Féminin** | **Masculin** |
| 133 | 3995 |
| 136 | 4040 |
| 164 | 4267 |
| 240 | 4911 |
| 257 | 5184 |
| 298 | 5858 |
| 311 | 5436 |
| 319 | 7902 |
| 322 | 8578 |
| 331 | 6786 |
| 340 | 6422 |
| 353 | 8976 |
| 354 | 7765 |
| 379 | 7229 |
| 384 | 7705 |

**2. Moyenne des abscisses et des ordonnées des deux groupes**

$$x\_{1}=\frac{133+136+164+240+257+298+311+319}{8}=232,25$$

$$y\_{1}=\frac{3995+4040+4267+4911+5184+5858+5436+7902}{8}=5199,13$$

$$x\_{2}=\frac{322+331+340+353+354+379+384}{7}=351,86$$

$$y\_{1}=\frac{8578+6786+6422+8976+7765+7229+7705}{7}=7637,86$$

Points : $P\_{1}\left(232,25;5199,13\right) et P\_{2}\left(351,86;7637,86\right)$

**3. Droite de régression**

$$a=\frac{7637,86-5199,13}{351,86-232,25}=\frac{2438,73}{119,61}=20,39$$

$y=20,39\left(75\right)+b$ $7637,86=20,39\left(351,86\right)+b$ $b=463,43$

$y=20,39x+463,43$

**4. Pour 10 000 diplômés masculins**

$10 000=20,39x+463,43$

$$20,39x=10 000-463,43$$

$$x≈468$$

**On peut estimer le nombre de diplômés féminins à 468.**

**Note : La droite obtenue avec un outil technologique est différente et plus précise. La droite de Mayer est une approximation.**

**Approche graphique**

Si l’élève utilise cette approche, voici un exemple de la représentation graphique de la distribution:

**Situation-problème 3 : La main d’œuvre dans le domaine de la construction**

**Utilisation d’un outil technologique pour le coefficient de corrélation et la droite de régression**



**Modèle linéaire**

Coefficient de corrélation : $r≈0,98$

La corrélation est forte et positive.

Droite de régression :$y=159,43x+107,56$

**Modèle quadratique**

Coefficient de corrélation : $r≈0,99$

La corrélation est forte et positive.

Droite de régression :$y=9,39x²+74,92x+220,25$

**Estimation de l’année pour 2 067 travailleuses**

Les deux modèles offrent une corrélation forte.

L’année 2011 : 14 ans après 1997

Modèle linéaire : $y=159,43\left(14\right)+107,56≈2 340$

Modèle quadratique : $y=9,39\left(14\right)^{2}+74,92\left(14\right)+220,25≈3 110$

**Le modèle linéaire semble offrir une estimation plus près de l’objectif réel. Le modèle quadratique offre une croissance plus importante au fil du temps.**