

**Réalisé par :**

**Hajar El Jaouhari**

1. **Introduction**
2. **Dictionnaire des variables et des modalités.**
3. **Méthodes utilisées**
4. **Résultats et interprétations**
5. **Conclusion**
6. **Introduction**

L’accident vasculaire cérébral (AVC) est, selon la définition internationale, « un déficit neurologique soudain d’origine vasculaire ». Tout AVC comporte donc d’une part une lésion cérébrale responsable du déficit neurologique, d’autre part une lésion vasculaire sous-jacente, qui est la cause immédiate de l’accident et en explique la soudaineté. Contrairement à ce que pourrait laisser supposer le terme AVC, il ne s’agit pas d’« accidents » mais de complications brutales de maladies chroniques des vaisseaux évoluant souvent depuis des années, voire des décennies. En revanche, les complications vasculaires des traumatismes crâniens (hématomes extra et sous-duraux) sont exclues du champ des AVC ainsi que les lésions cérébrales des arrêts cardiaques, quelle qu’en soit la cause.

La fréquence, la gravité et le coût des accidents vasculaires cérébraux (AVC) en font u n problème de santé publique considérable. *Troisième cause de mort après l’infarctus du myocarde et les cancers, les AVC constituent la première cause de handicap acquis chez l’adulte, la deuxième cause de démence* (après la maladie d’Alzheimer avec laquelle ils sont fréquemment intriqués et dont ils aggravent les manifestations cliniques) *et une cause majeure de dépression* tant chez les patients que dans leur entourage , Compte tenu de leur fréquence et de la gravité de leurs séquelles, les AVC sont parmi les affections les plus coûteuses qui existent .

Alors on va suggérer quelques hypothèses et on va les confirmer par la suite ,

Hypothèse 1 : il existe une relation entre l' AVC et Type de résidence

Hypothèse 2 : les hommes sont plus susceptibles de contracter la maladie

Hypothèse 3 : l' hypertension et le work\_type sont des causes d ' AVC

Hypothèse 4 : il existe une relation entre le tabagisme et les maladies cardiaques

Hypothèse 5 : les personnes âgées sont plus sensibles aux maladies

**II. Dictionnaire des variables et des modalités.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Code de variable | Nom de variable | descriptif de variable | Type de variable | Position de variable | Largeur de variable | Code de modalité | Nom de modalité | Descriptif de modalité |
| id | identifiant | Identifiant unique | ordinal | 1 | 4 |  |  |  |
| gender | sexe | Genre | nominale | 2 | 1 | 1  \_\_\_\_\_\_\_2  \_\_\_\_\_\_\_  3  \_\_\_\_\_\_\_  9 | Female  \_\_\_\_\_\_\_  Male  \_\_\_\_\_\_\_  Other  \_\_\_\_\_\_\_  Valeur maunqante | \_\_\_\_\_\_\_  Masqué pour garantir  L'anonymat |
| age | âge | Âge | ordinale | 3 | 2 | Entre 1 et 82 |  |  |
| hypertension | hypertension | Fonction binaire d'hypertension | nominale | 4 | 1 | 0  \_\_\_\_\_\_\_  1  \_\_\_\_\_\_\_  9 | Non  \_\_\_\_\_\_\_  Oui  \_\_\_\_\_\_\_  Valeur maunqante |  |
| heart\_disease | maladie cardiaque | Caractéristique binaire des maladies cardiaques | nominale | 5 | 1 | 0  \_\_\_\_\_\_\_  1  \_\_\_\_\_\_\_  9 | Non  \_\_\_\_\_\_\_  Oui  \_\_\_\_\_\_\_  Valeur maunqante |  |
| ever\_married | mariage | Le patient a-t-il déjà été marié? | nominale | 6 | 1 | 1  \_\_\_\_\_\_\_  2  \_\_\_\_\_\_\_  9 | No  \_\_\_\_\_\_\_  Yes  \_\_\_\_\_\_\_  Valeur maunqante |  |
| work\_type | Le type de travail | Type de travail du patient |  | 7 | 1 | 1  \_\_\_\_\_\_\_  2  \_\_\_\_\_\_\_  3  \_\_\_\_\_\_\_  4  \_\_\_\_\_\_\_  5  \_\_\_\_\_\_\_  9 | Children  \_\_\_\_\_\_\_  Govt\_job  \_\_\_\_\_\_\_  Never\_worked  \_\_\_\_\_\_\_  Private  \_\_\_\_\_\_\_  Self-employed  \_\_\_\_\_\_\_  Valeur maunqante |  |
| Residence\_type | Residence | Type de résidence du patient | nominale | 8 | 1 | 1  \_\_\_\_\_\_\_  2  \_\_\_\_\_\_\_  9 | Rural  \_\_\_\_\_\_  Urbain  \_\_\_\_\_\_\_  Valeur maunqante |  |
| avg\_glucose\_level | glucose | Type de résidence du patient | ordinale | 9 | 4 |  |  |  |
| bmi | indice de masse corporelle | Indice de masse corporelle | nominale | 10 | 3 | Entre 1 et 419  \_\_\_\_\_\_\_  9999 | \_\_\_\_\_\_\_  Valeur maunqante |  |
| smoking\_status | tabagisme |  | nominale | 11 | 1 | 1  \_\_\_\_\_\_\_  2  \_\_\_\_\_\_\_  3  \_\_\_\_\_\_\_  4 | formerly smoked  \_\_\_\_\_\_\_  never smoked  \_\_\_\_\_\_\_  Smokes  \_\_\_\_\_\_\_  Unknown |  |
| stroke | AVC |  | nominale | 12 | 1 | 0  \_\_\_\_\_\_\_  1 | Non  \_\_\_\_\_\_\_  Oui |  |

1. **Méthodes utilisées**

les Méthodes utilisées dans cette analyse statistique est les suivantes :

* [Statistiques descriptives univariées](http://elearning.esi.ac.ma/course/view.php?id=173#section-3)
* Analyse Statistique bivariées
* Tableaux croisés
* Création de Variable
* Recodage de variables
* Explorer
* le test d'indépendance Khi-2
* tableau de contingence
* le V de Cramer
* Diagramme en bâtons

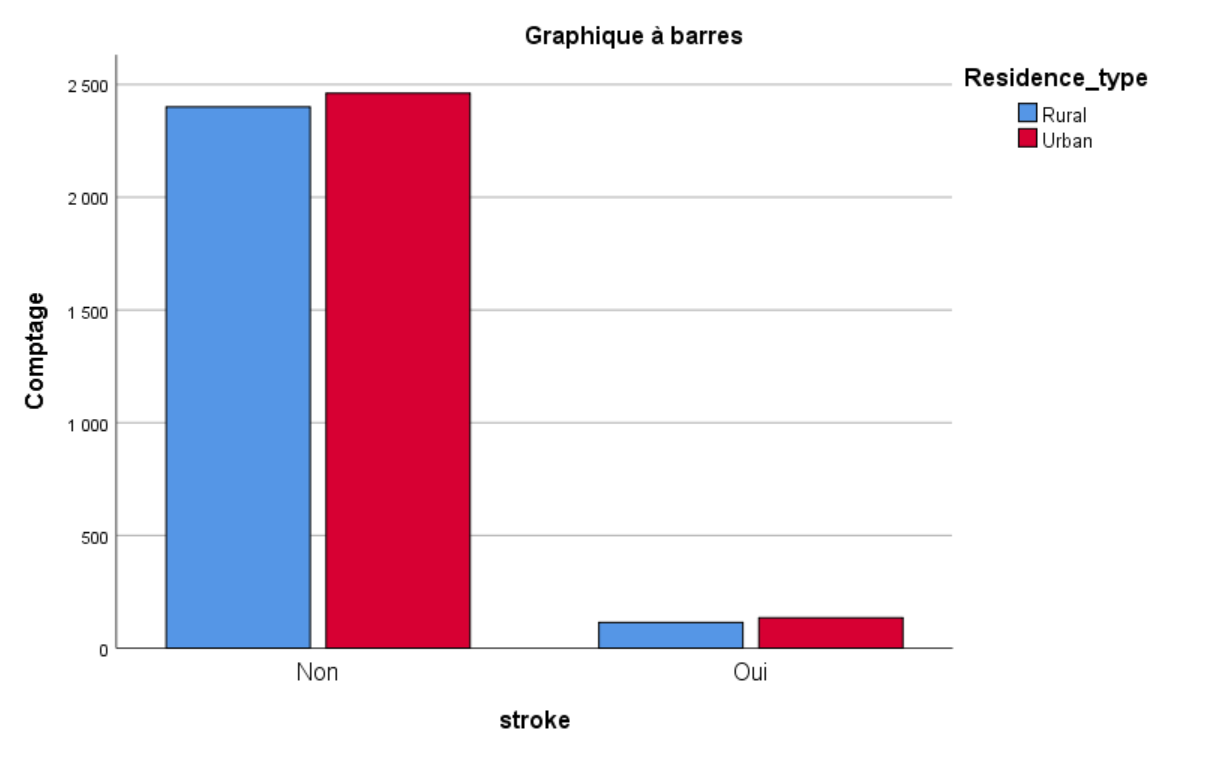
**IV. Résultats et interprétations**

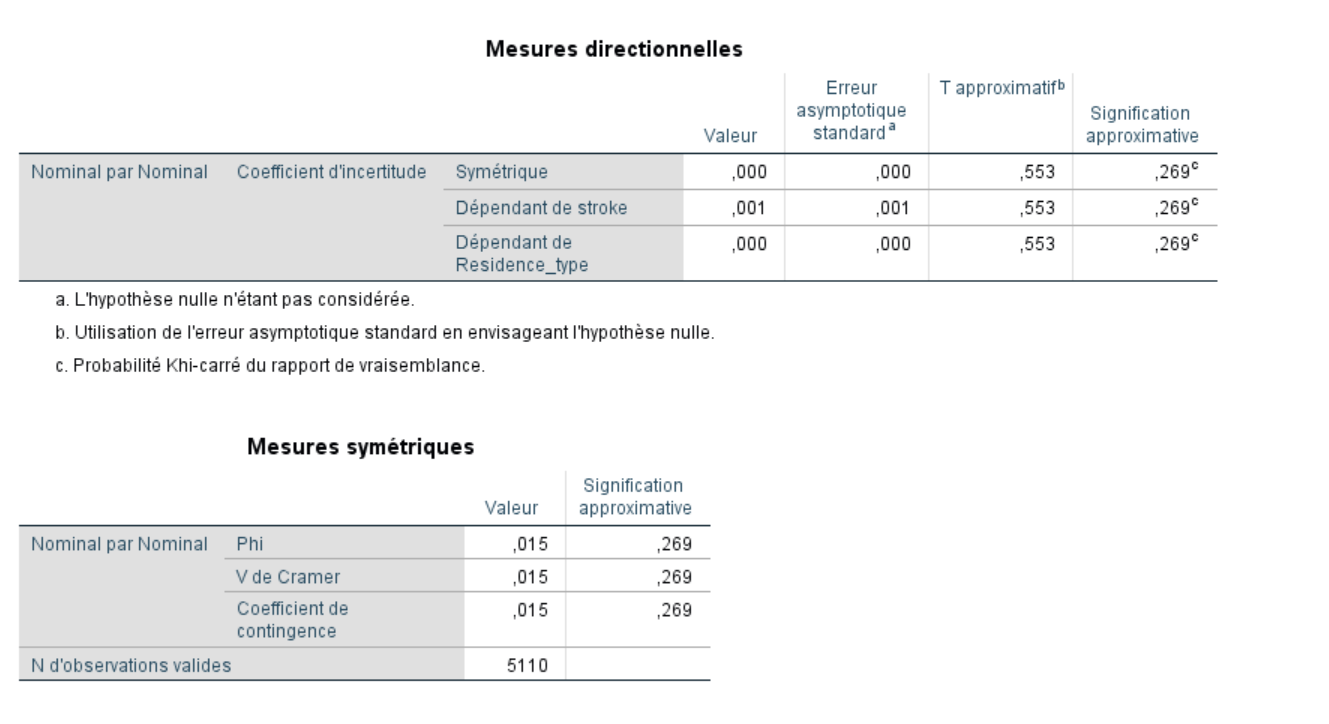
On va commencer par la vérification de la relation entre les 2 variables suivantes :

* type de résidence
* AVC

On a 2 hypothèses :

* H0 : il n'y a pas de relation entre les deux variables
* H1: il y'a une relation entre les deux variables



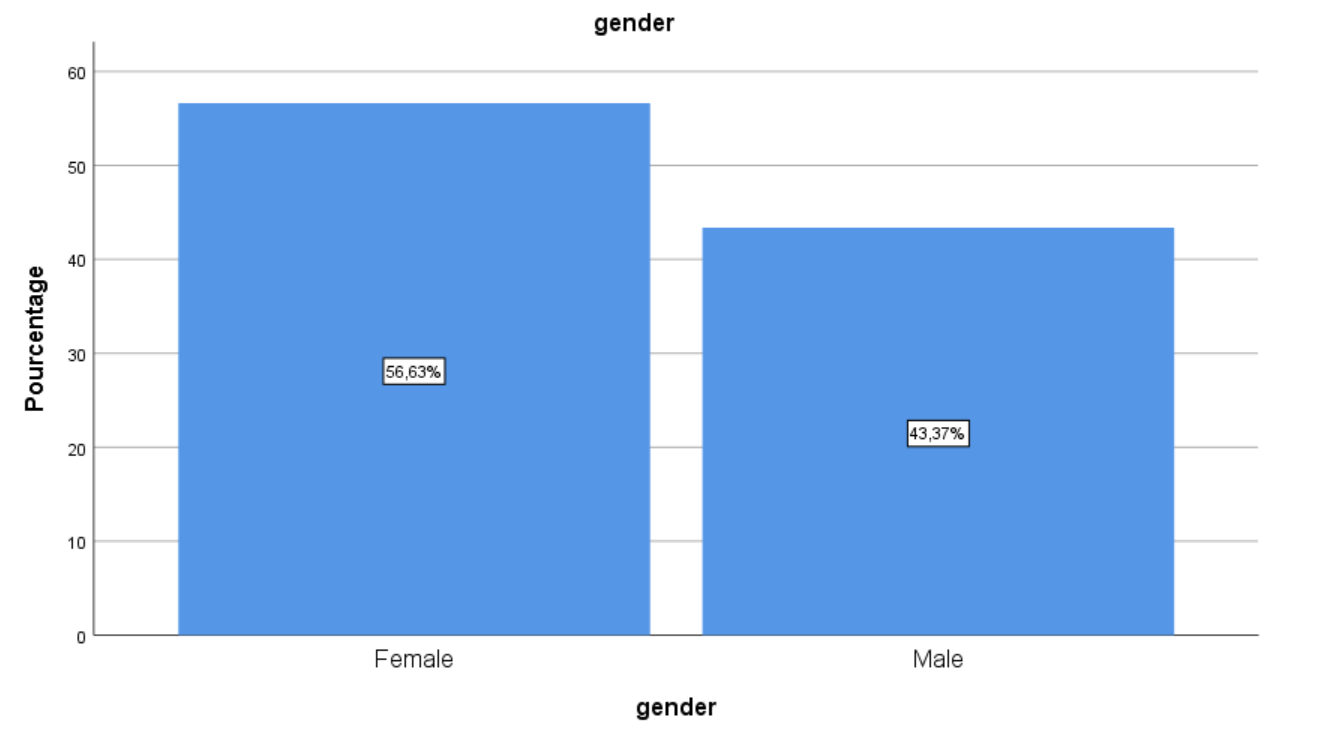


Selon le diagramme en bâton juxtaposé affichent l'AVC selon le milieu nous avons constaté que les % des individus en milieu urbain pour chaque patient qui souffre de l'AVC est supérieur au % des individus issus du milieu rural

Hypothèse : le milieu de résidence influence la maladie

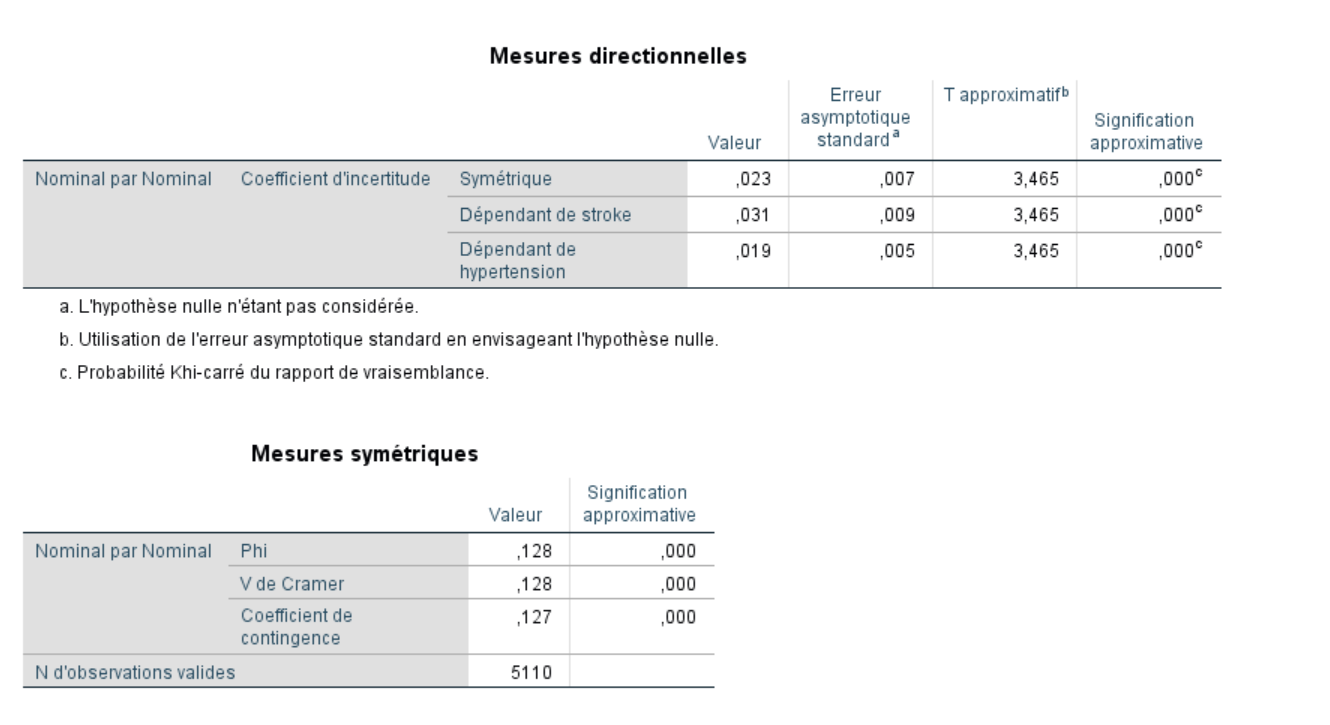
on a V de Cramer est plus proche de 0 alors il y a indépendance entre les 2 variables étudiées la relation est faible

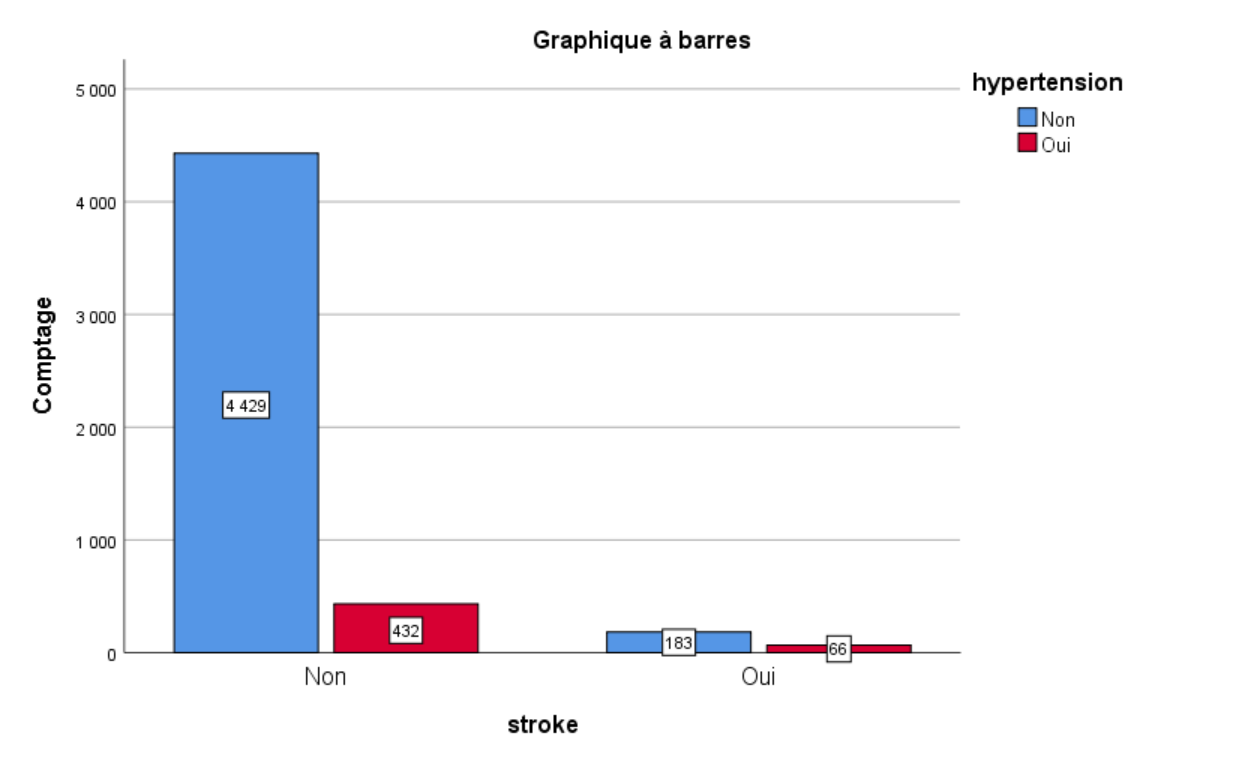
On va vérifier maintenant cette hypothèse : les hommes sont plus susceptibles de contracter la maladie



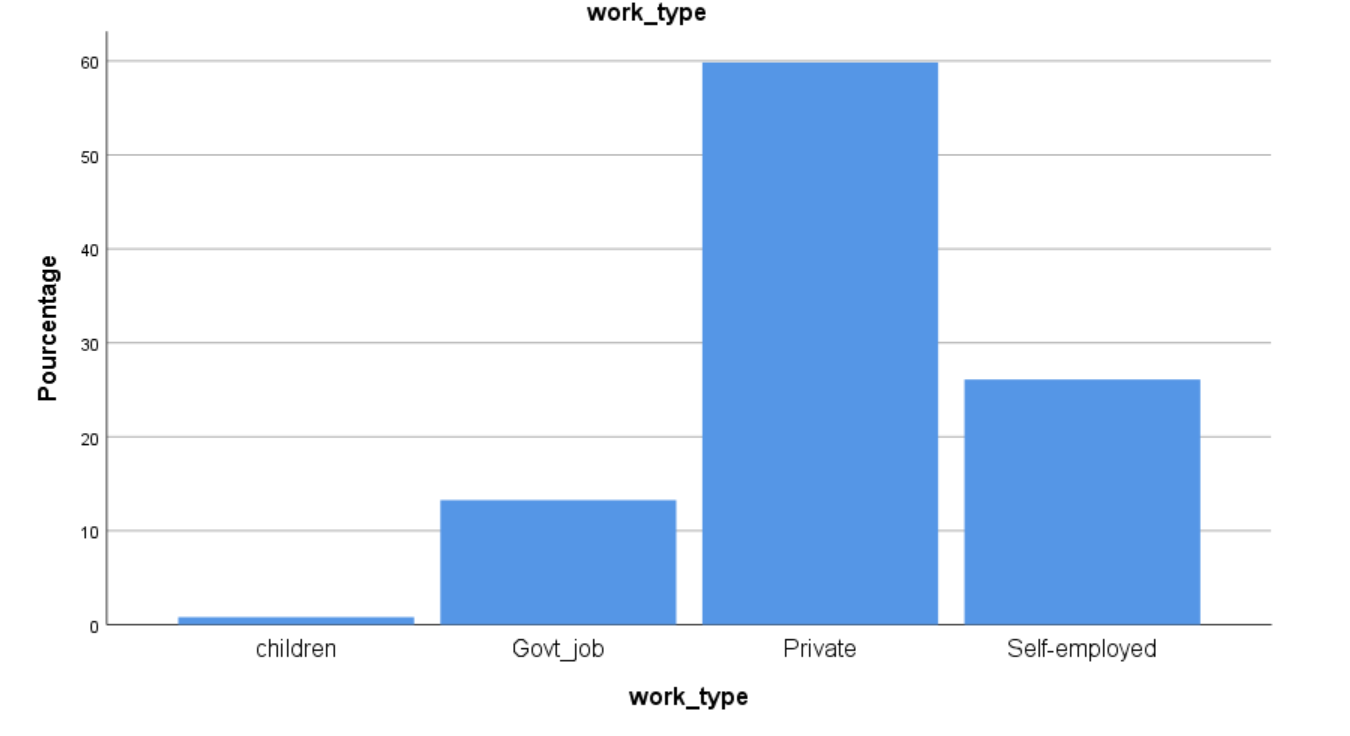
le purcentage des femmes (56,63%) est plus élevé au pourcentage des hommes (43,37) alors les femmes sont plus susceptibles de contracter la maladie

Hypothèse 3 : l' hypertension et le work\_type sont des causes d ' AVC



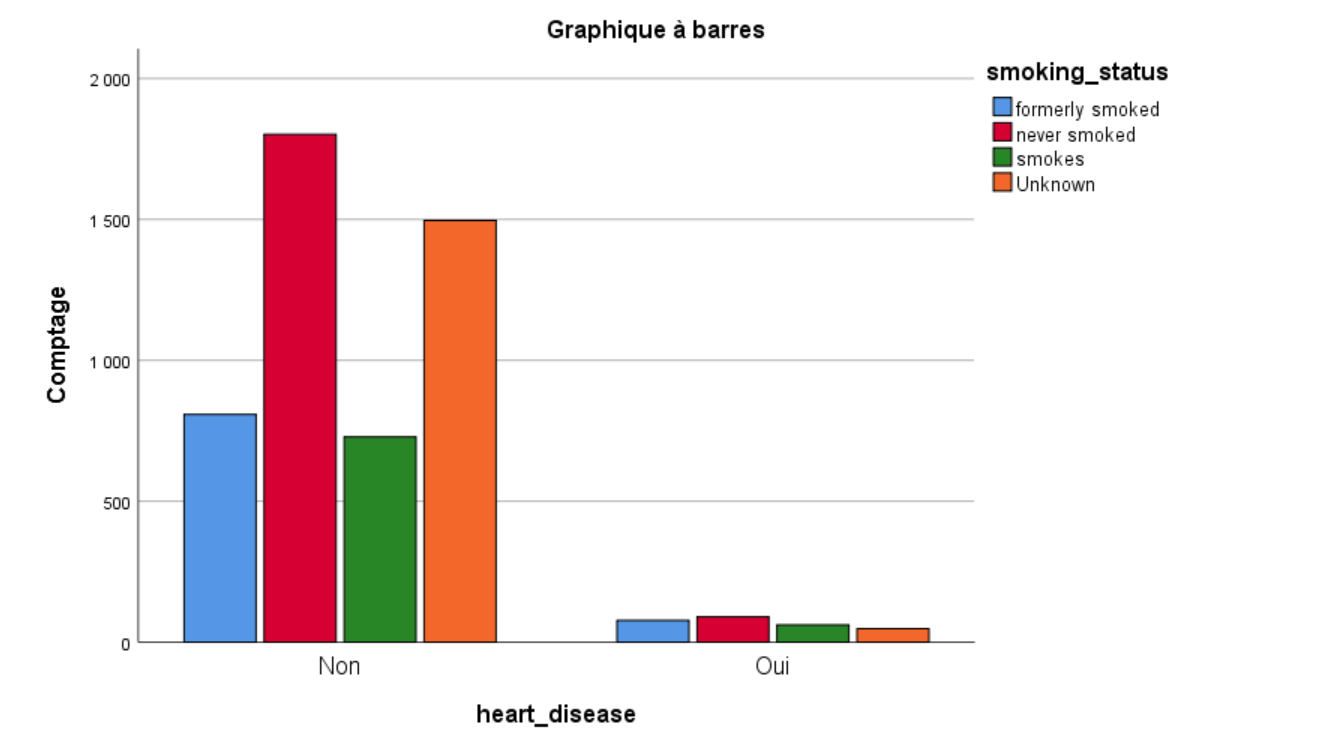
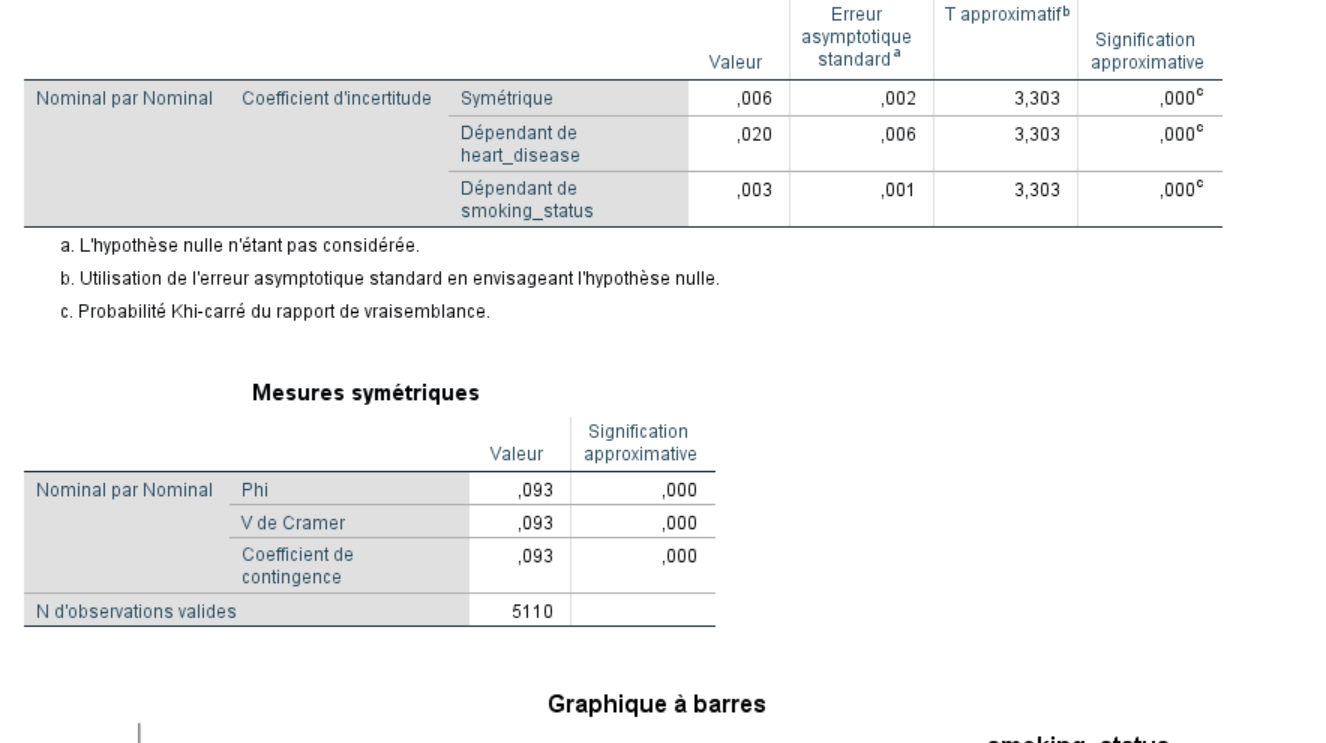


d'après les résultats on constate que les individus qui ne souffre pas de l'AVC ont l'hypertension plus que les individus qui souffre de la maladie



Selon le diagramme en bâton juxtaposé affichent les individus qui souffrent de l 'AVC et le pourcentage du type de travail , nous avons constaté que les % des individus qui ont un travail privé est supérieur au % des autres type de travail .

Hypothèse 4 : il existe une relation entre le tabagisme et les maladies cardiaques

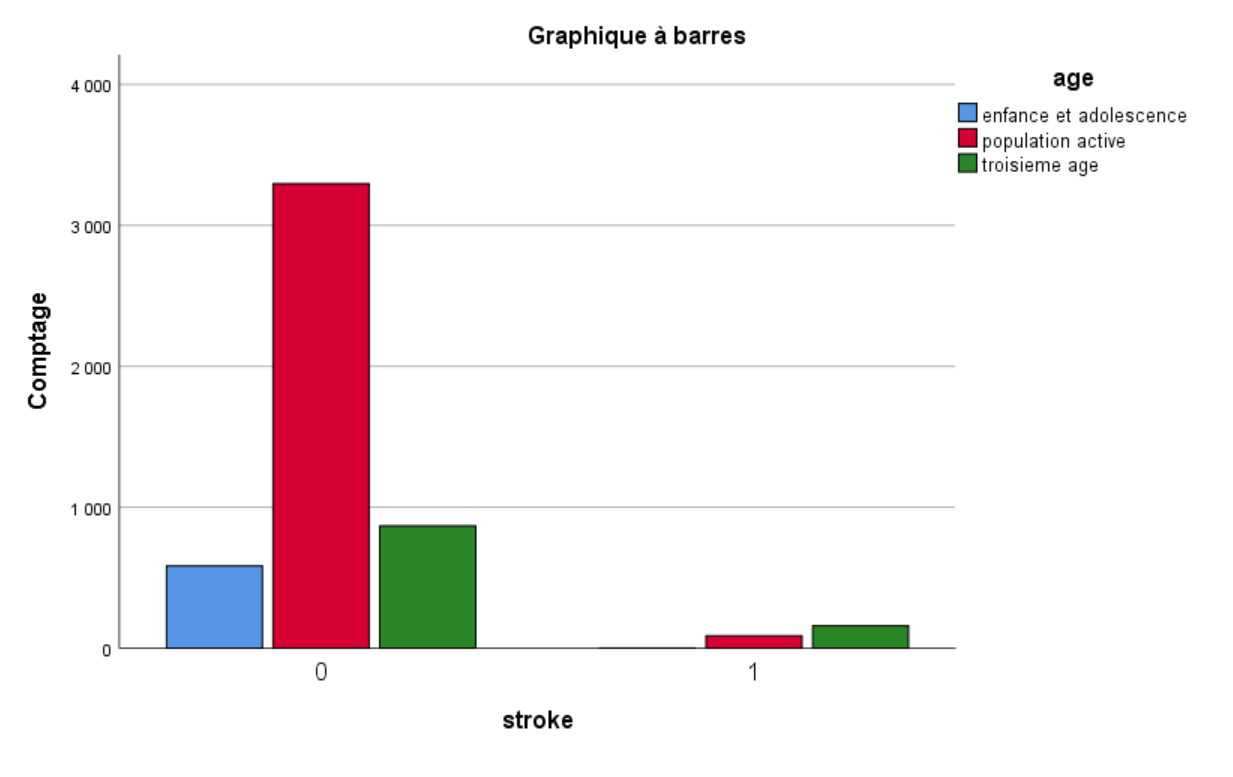


Selon le diagramme en bâton juxtaposé affichent heart\_disease selon le smoking\_status nous avons constaté que les % des individus qui n'ont jamais fumé est supérieur au % des individus anciennement fumé et qui fument .

Hypothèse : le tabagisme influence la maladie

on a V de Cramer est plus proche de 0 alors il y a indépendance entre les 2 variables étudiées la relation est faible .

Hypothèse 5 : les personnes âgées sont plus sensibles aux maladies



on constate que la tranche du troisième âge sont plus sont plus susceptibles de contracter la maladie sont plus susceptibles de contracter la maladie et aussi la population active .

**V. Conclusion**

Cet ensemble de données est utilisé pour prédire si un patient est susceptible d'avoir un accident vasculaire cérébral en fonction des paramètres d'entrée tels que le sexe, l'âge, diverses maladies et le tabagisme. Chaque ligne des données fournit des informations pertinentes sur le patient, On constate d'après cette analyse que l'âge et le milieu de résidence sont parmi les causes qui influence cette maladie .