

Statistiques descriptives et esquisse de plan d'exposé

Coulibaly_Khadidiatou

2024-03-07

A. Calcul de certaines statistiques descriptives

- Statistiques concernant la taille de la population:

Valeurs	
Minimum	4.5
Quartile 1	15.0
Médiane	18.0
Quartile 3	17.7
Moyenne	18.0
Maximum	27.7

- Statistiques concernant le solde de la balance commercial:

Valeurs	
Minimum	-150.0
Quartile 1	-100.0
Médiane	-58.5
Quartile 3	-66.2
Moyenne	-20.0
Maximum	-10.0

B. Quelques tableaux...

Voyons une répartition des pays en fonction du niveau de vie.

Table 3: Répartition des pays en fonction du niveau de vie

	Niveau de vie	Pourcentage de pays
High	169	56.3
Low	88	29.3
Med	43	14.3

Voyons une répartition des pays en fonction du sexe dominant.

Table 4: Répartition des pays en fonction du sexe dominant

	Sexe	Pourcentage de pays
F	150	50
H	150	50

En croisant le secteur dominant à la situation géographique, on obtient le tableau suivant:

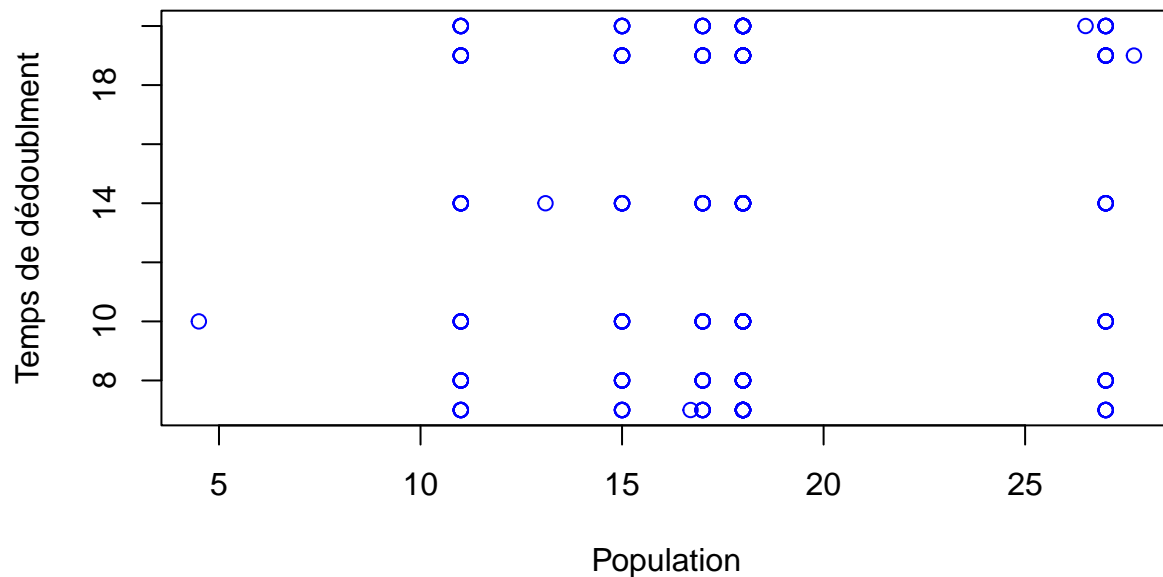
Table 5: Tableau croisé

	Equateur	Nord	Sud	Total
Primaire	0	30	60	90
Secondaire	30	60	30	120
Tertiaire	60	0	30	90

C. Quelques représentations graphiques...

Il serait intéressant de construire un nuage de points pour étudier la relation entre la population (en millions) et le temps de dédoublement de celle-ci.

Nuage de points entre la population et le temps de dédoublement

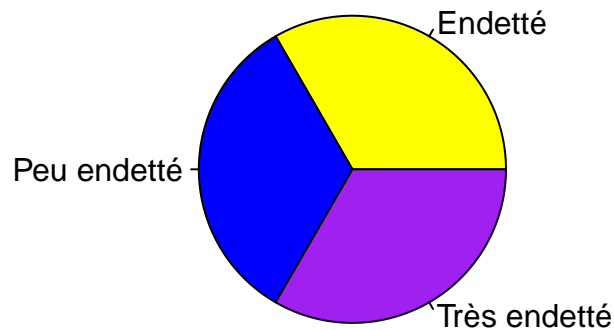


A la vue d'un tel nuage, il est évident qu'il n'y a pas de relation linéaire entre la taille de population et le temps de dédoublement.

Et si l'on regardait de plus près le niveau d'endettement des pays ? Quelle est l'importance de leurs dettes ?

```
#Pour cette opération, nous avons choisi d'afficher le code.
pie(table(My_data$Dette),labels =names(table(My_data$Dette)) ,
main="Niveau d'endettement des pays", col=c("yellow","blue", "purple" ))
```

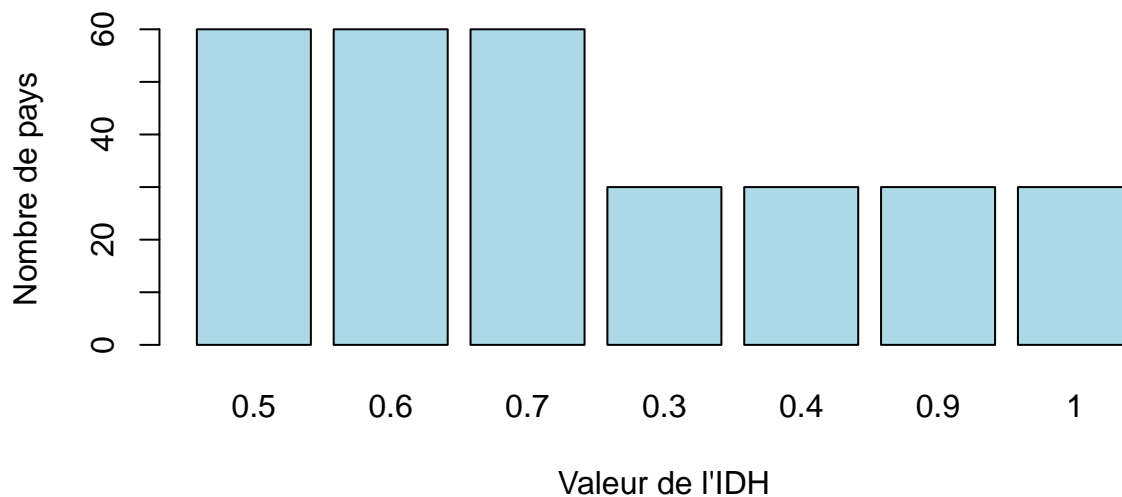
Niveau d'endettement des pays



Il apparait alors une égale distribution du niveau d'endettement.

Pourquoi ne pas nous intéresser à l'IDH des différents pays ? Quelle valeur de l'indice est la plus occurente ? Le diagramme en barres ci-dessous nous donne la réponse à la question.

diagramme en barres de l'IDH



Nous remarquons que l'IDH prend un nombre fini de valeurs.

C'est sur ce graphique que nous terminons nos analyses.

D. Esquisse de plan sur l'exposé sur le thème : *Statistiques descriptives et visualisation des variables catégorielles avec ggplot2*

Nous proposons l'esquisse de plan suivante :

1. Présentation générale d'une variable catégorielle

Il s'agira de donner la définition, caractéristiques...

2. Analyses et tests possibles

On parlera des statistiques descriptives et tests statistiques possibles avec une variable catégorielle.

3. Représentations graphiques:

On présentera les différentes représentations graphiques et leur l'implémentation su R, notamment avec ggplot.

4. Cas pratiques

On utilisera des bases de données pour un usage des outils présentés.

NB: Les exposants sont *Samba Dieng* et *Khadidiatou Coulibaly*.