

# Prise en main de l'application Shiny M@rsouin

L'application Shiny M@rsouin est une plateforme web pour la visualisation et l'étude des données d'enquêtes M@rsouin. L'intérêt de cette plateforme est de mettre à disposition un outil permettant de manipuler les données d'enquête et de réaliser des traitements statistiques de manière simple et efficace sans prérequis de programmation. Elle est accessible à tous via un navigateur web, et repose sur le langage de programmation R et le package Shiny.

Cette aide succincte vise à présenter la plateforme et à permettre une prise en main rapide de celle-ci, mais part du principe que les méthodes statistiques sous-jacentes sont connues.

Voici à quoi ressemble son interface.

## Visualisation des données d'enquête Marsouin



Sélectionnez les options ci-dessous

Quelle enquête ?  
PME2015

Variable 1 :  
Appartient à un groupe

Variable 2 :  
Evolution du chiffre d'affaires

Type de graphique :  
Barres

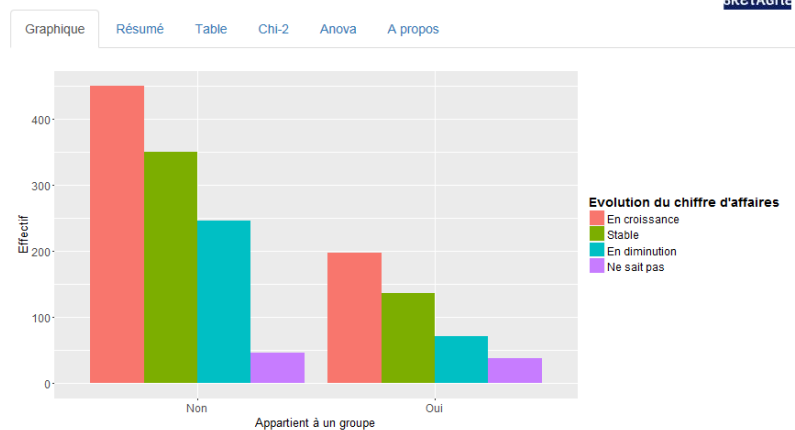
Afficher les points  
 Afficher les valeurs  
 Axe des ordonnées en Pourcentage

Valeur max de l'axe X :  
0 1,000 2,000

Valeur max de l'axe Y :  
0 1,000 2,000

Note: Pour accéder aux téléchargements, vous devez accepter les conditions d'utilisation de la plateforme définies dans l'onglet A propos

J'accepte les conditions d'utilisation



Cette interface se décompose en 2 parties : à gauche le volet de contrôle et à droite la zone de visualisation.

Le volet de contrôle permet de sélectionner par des menus déroulants l'enquête à étudier, la variable d'intérêt et de la croiser avec une 2<sup>ème</sup> variable.

Il permet également de sélectionner le type de graphique souhaité et de le personnaliser à travers plusieurs options sous forme de cases à cocher.

Sélectionnez les options ci-dessous

Quelle enquête ?  
PME2015

Variable 1 :  
Appartient à un groupe

Variable 2 :  
Evolution du chiffre d'affaires

Type de graphique :  
Barres

Plusieurs types de graphiques sont disponibles : boxplot, diagramme en violons, histogramme, densité, diagramme en barres et cartographie. Les options de manipulation des graphiques auront un sens (ou non) en fonction du graphique sélectionné.

Les graphiques et traitements statistiques (sauf l'anova) prennent en compte les éventuelles pondérations assurant la représentativité des échantillons.

La première option « Afficher les points » permet dans le cas d'un graphique sous forme de boxplot ou de violon de faire apparaître les individus sous forme de points.

Sélectionnez les options ci-dessous

Quelle enquête ?  
EPCI2014

Variable 1 :  
Nombre d'agents

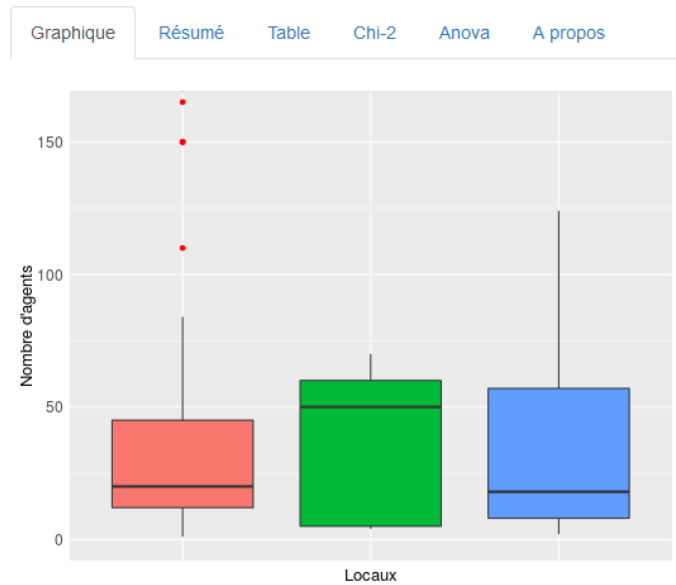
Variable 2 :  
Locaux

Type de graphique :  
Boxplot

Afficher les points

Afficher les valeurs

Axe des ordonnées en Pourcentage



Sélectionnez les options ci-dessous

Quelle enquête ?  
EPCI2014

Variable 1 :  
Nombre d'agents

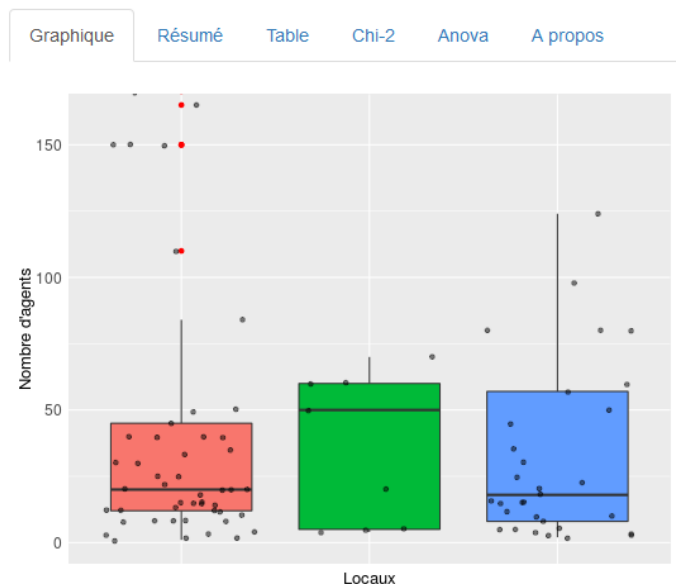
Variable 2 :  
Locaux

Type de graphique :  
Boxplot

Afficher les points

Afficher les valeurs

Axe des ordonnées en Pourcentage



La 2<sup>ème</sup> option « Afficher les valeurs » permet d'afficher l'effectif composant chaque barre.

Sélectionnez les options ci-dessous

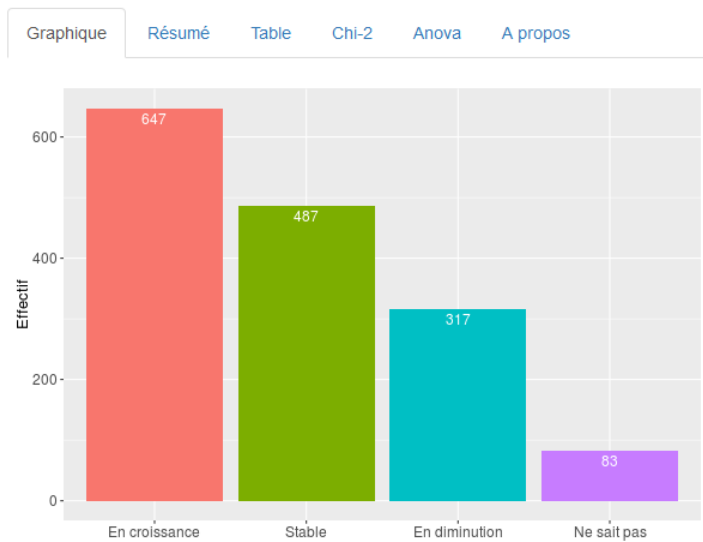
Quelle enquête ?  
PME2015

Variable 1 :  
Evolution du chiffre d'affaires

Variable 2 :  
Evolution du chiffre d'affaires

Type de graphique :  
Barres

Afficher les points  
 Afficher les valeurs  
 Axe des ordonnées en Pourcentage



La 3<sup>ème</sup> option « Axe des ordonnées en Pourcentage » permet quant à elle de passer, sur l'axe des ordonnées, de l'affichage des effectifs aux pourcentages.

Sélectionnez les options ci-dessous

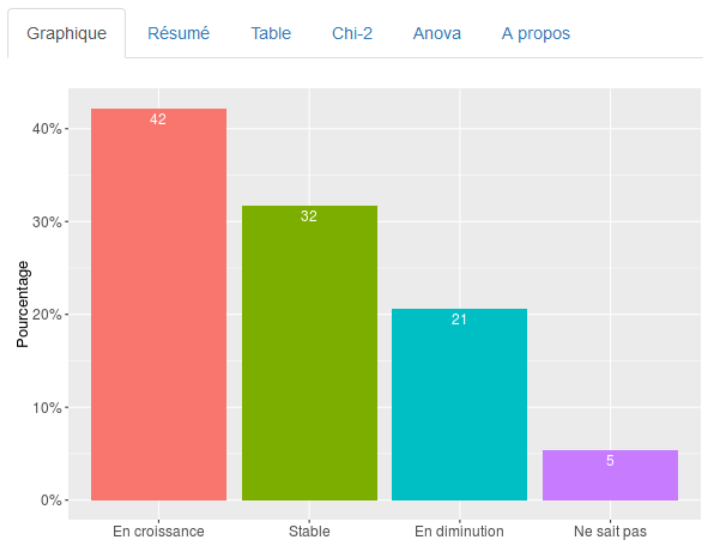
Quelle enquête ?  
PME2015

Variable 1 :  
Evolution du chiffre d'affaires

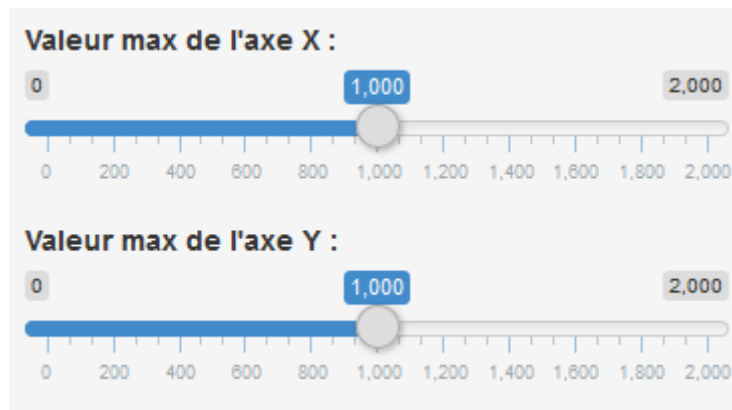
Variable 2 :  
Evolution du chiffre d'affaires

Type de graphique :  
Barres

Afficher les points  
 Afficher les valeurs  
 Axe des ordonnées en Pourcentage



Juste en dessous 2 curseurs permettent d'ajuster les axes horizontaux et verticaux des graphiques et ainsi de choisir le niveau de zoom désiré. Ces 2 curseurs se révèlent particulièrement pratiques pour les graphiques du type boxplot, violon, densité et histogramme.



Enfin, l'acceptation des conditions d'utilisation de la plateforme entrainera l'apparition de 2 boutons en bas du volet de contrôle. Ces boutons permettent respectivement de télécharger le jeu de données (réponses à l'enquête) et de télécharger le graphique nouvellement réalisé.

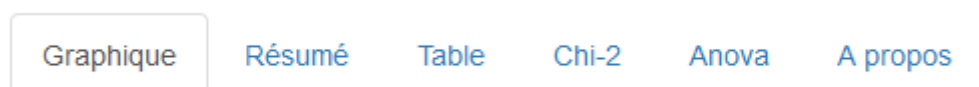
Note: Pour accéder aux téléchargements, vous devez accepter les conditions d'utilisation de la plateforme définies dans l'onglet *A propos*

J'accepte les conditions d'utilisation

[Données](#) [Graphique](#)

Ces conditions d'utilisation sont données dans l'onglet *A propos*.

Au-dessus de la zone de visualisation se trouve une barre de 6 onglets permettant d'accéder aux diverses fonctionnalités proposées par la plateforme. Jusqu'ici, nous nous sommes intéressés au 1<sup>er</sup> onglet proposant la création de graphiques.



Intéressons-nous maintenant aux onglets suivants.

Le 2<sup>ème</sup> onglet « Résumé » donne les statistiques descriptives univariées (tri à plat) de la variable étudiée (Variable 1). Dans le cas d'une variable qualitative, un tableau regroupant l'effectif, la fréquence et la fréquence cumulée de chaque modalité sera donné. Dans le cas d'une variable quantitative, ce sera un résumé de sa distribution : le minimum et le maximum, les 1<sup>er</sup> et 3<sup>ème</sup> quartiles, la moyenne et la médiane de la variable.

Graphique				Résumé	
Graphique				Résumé	
	Effectif	Fréquence	Fréquence cumulée	Distribution	
Ne sait pas	83.21	5.42	5.42	Min.	10.00
En diminution	316.81	20.65	26.08	1er quartile	13.00
Stable	486.73	31.73	57.81	Médiane	18.00
En croissance	647.18	42.19	100.00	Moyenne	30.32
				3e quartile	35.00
				Max.	249.00

Le 3<sup>ème</sup> onglet « Table » permet de visualiser sous la forme d'un tableur les réponses à l'enquête. Par défaut, 25 réponses sont présentées par page (Show 25 entries), mais il est possible d'afficher 50 ou 100 réponses par page en cliquant sur la liste déroulante. Il est également possible de trier les colonnes par ordre croissant ou décroissant grâce aux 2 petites flèches grises présentes à droite de chaque colonne, d'effectuer une recherche dans le tableur grâce au champ « Search » situé en haut à droite, ainsi que de filtrer les réponses grâce aux champs filtres présents sous chaque colonne.

Graphique						Résumé		Table		Chi-2		Anova		A propos	
Show 25 entries						Search: <input type="text"/>									
Numéro.département	Fonction.du.répondant	Fonction.autre	Nombre.d.agents	Locaux	Locau:										
56	Autre	Chargé de mission	80	partagés avec une autre structure	associati										
56	DGS		40	exclusifs à l'EPCI											
22	DGS		8	exclusifs à l'EPCI											

Showing 1 to 25 of 94 entries

Le 4<sup>ème</sup> onglet « Chi-2 » regroupe à la fois un tableau croisé (tri croisé) des 2 variables sélectionnées dans le volet de contrôle et un test du chi-2 d'indépendance de ces 2 variables.

Le tableau croisé est un tableau de contingence enrichi puisqu'il donne pour chaque croisement de modalités : les effectifs (présents classiquement dans un tableau de contingence), les pourcentages, les pourcentages ligne et les pourcentages colonne, ainsi que les totaux, en ligne, en colonne et globaux (parfois aussi appelés marges).

**Tableau croisé**

		<b>En croissance</b>	<b>Stable</b>	<b>En diminution</b>	<b>Ne sait pas</b>	<b>Total</b>
<b>Ne sait pas</b>	eff	9	8	2	9	28
	%	0.61	0.51	0.15	0.59	1.85
	%lig	33.03	27.46	7.95	31.56	
	%col	1.45	1.61	0.71	10.79	
<b>Non</b>	eff	329	291	194	45	860
	%	21.47	18.98	12.65	2.96	56.06
	%lig	38.30	33.86	22.56	5.28	
	%col	50.90	59.82	61.23	54.60	
<b>Oui</b>	eff	308	188	121	29	645
	%	20.11	12.24	7.86	1.88	42.08
	%lig	47.78	29.08	18.68	4.46	
	%col	47.65	38.57	38.05	34.61	
<b>Total</b>	eff	647	487	317	83	1534
	%	42.19	31.73	20.65	5.42	100.00

Le test du chi-2 permet de tester l'indépendance entre 2 variables qualitatives. Le résultat de ce test est présenté sous le tableau croisé, notamment la valeur observée de la statistique du chi-2, le nombre de degrés de liberté et la p-value, permettant de conclure à la dépendance ou à l'indépendance des 2 variables en rejetant ou non l'hypothèse H0: « les variables sont indépendantes ».

### Test du chi-2

```
Pearson's Chi-squared test

data: tabCroisPond
X-squared = 53.145, df = 6, p-value = 1.097e-09
```

### Résidus standardisés

	En croissance	Stable	En diminution	Ne sait pas
Ne sait pas	-1.00	-0.49	-1.69	6.21
Non	-3.48	2.02	2.08	-0.28
Oui	3.78	-1.90	-1.63	-1.42

L'analyse des résidus permet ensuite de déterminer quels croisements contribuent le plus au rejet de l'indépendance des variables. Les croisements significativement influents (5%) ont été colorés en vert lorsqu'il y a sur-représentation (l'effectif observé est plus grand qu'attendu), et en rouge lorsqu'il y a sous-représentation (l'effectif observé est plus petit qu'attendu).

### Attention

```
chr "Chi-squared approximation may be incorrect"
```

Enfin, un message d'alerte apparaît lorsque l'approximation du chi-2 peut-être incorrecte, notamment lors du non-respect des conditions d'application du test du chi-2.

Le 5<sup>ème</sup> onglet « Anova » permet tester l'égalité des moyennes prises par une variable quantitative en fonction des modalités d'une variable qualitative. La variable quantitative est à sélectionner dans « Variable 1 » et la variable qualitative dans « Variable 2 ».

Graphique    Résumé    Table    Chi-2    **Anova**    A propos

### Test de comparaison de moyennes

One-way analysis of means (not assuming equal variances)

data: variable and group

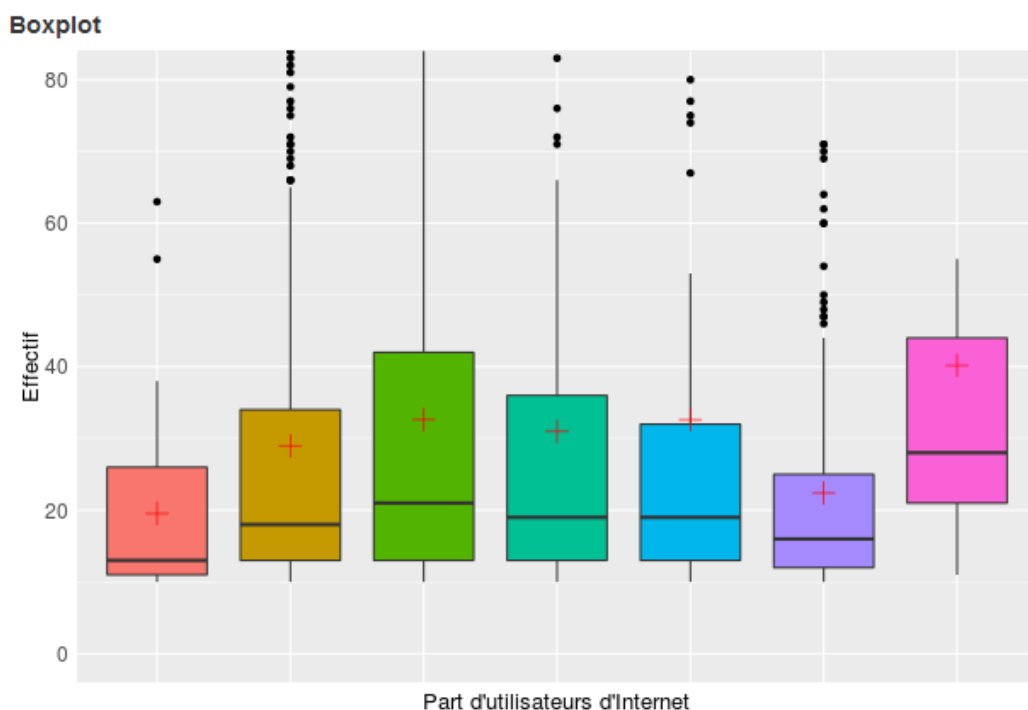
F = 5.6391, num df = 6.00, denom df = 148.53, p-value = 2.668e-05

### Moyennes observées par modalité

Personne	Moins de 25%	25 à 50%
19.54	28.98	32.62
50 à 75%	75 à 100%	La totalité du personnel
30.98	32.60	22.40
Ne sait pas		
40.18		

Le résultat du test est affiché avec la valeur de la statistique de test, le nombre de degrés de liberté et la p-value permettant ainsi de rejeter ou non l'hypothèse H0 : « les moyennes sont égales ».

Les moyennes par modalité sont également affichées, ainsi qu'un boxplot permettant de visualiser directement les distributions de la variable quantitative pour chaque modalité, et en particulier de comparer les moyennes, représentées par un + rouge.





Et pour terminer, le 6<sup>ème</sup> onglet « A propos » donne une description succincte de l'application Shiny M@rsouin et propose une aide à la prise en main de l'application, téléchargeable au format PDF.

Les conditions d'utilisation de la plateforme et des résultats sont décrites à cet endroit.

Il est rappelé également que la plateforme est en développement, et qu'elle continue d'évoluer (date de la dernière mise à jour).

Enfin si vous souhaitez exprimer vos impressions, me faire des suggestions, etc., vous pouvez le faire grâce au contact présent dans cet onglet.

[Graphique](#) [Résumé](#) [Table](#) [Chi-2](#) [Anova](#) [A propos](#)

Shiny M@rsouin est une application de visualisation et d'étude des données recueillies lors des enquêtes réalisées par le GIS M@rsouin auprès des individus, des PME et des collectivités bretonnes.

Une aide est disponible [ici](#).

L'utilisation de la plateforme Shiny M@rsouin est soumise aux conditions d'utilisation suivantes :

*L'utilisateur s'engage à citer le GIS M@rsouin lors de la publication de résultats issus du traitement des données d'enquêtes M@rsouin ou de la plateforme Shiny M@rsouin, et à informer le GIS M@rsouin desdites publications. Aucune utilisation commerciale ne peut être faite de la plateforme ou des résultats.*

Cette application continue à être développée. D'autres enquêtes seront incorporées ultérieurement.

N'hésitez pas à me faire part de vos impressions et suggestions. Bientôt de nouvelles améliorations...

Version v0.2

Actualisée le 09/03/2016



**Contactez l'auteur**

[Nicolas Deporte](#)

Statisticien

M@rsouin

Bonne utilisation de Shiny M@rsouin !