

Hiver 2019

Professeur : Amor Amami

## Pratique-1 : introduction à ggplot2

### Objectif :

Créez des graphiques en utilisant **ggplot2**. Ne vous inquiétez pas si vos graphiques ne correspondent pas exactement à ce qui est présenté ci-dessous.

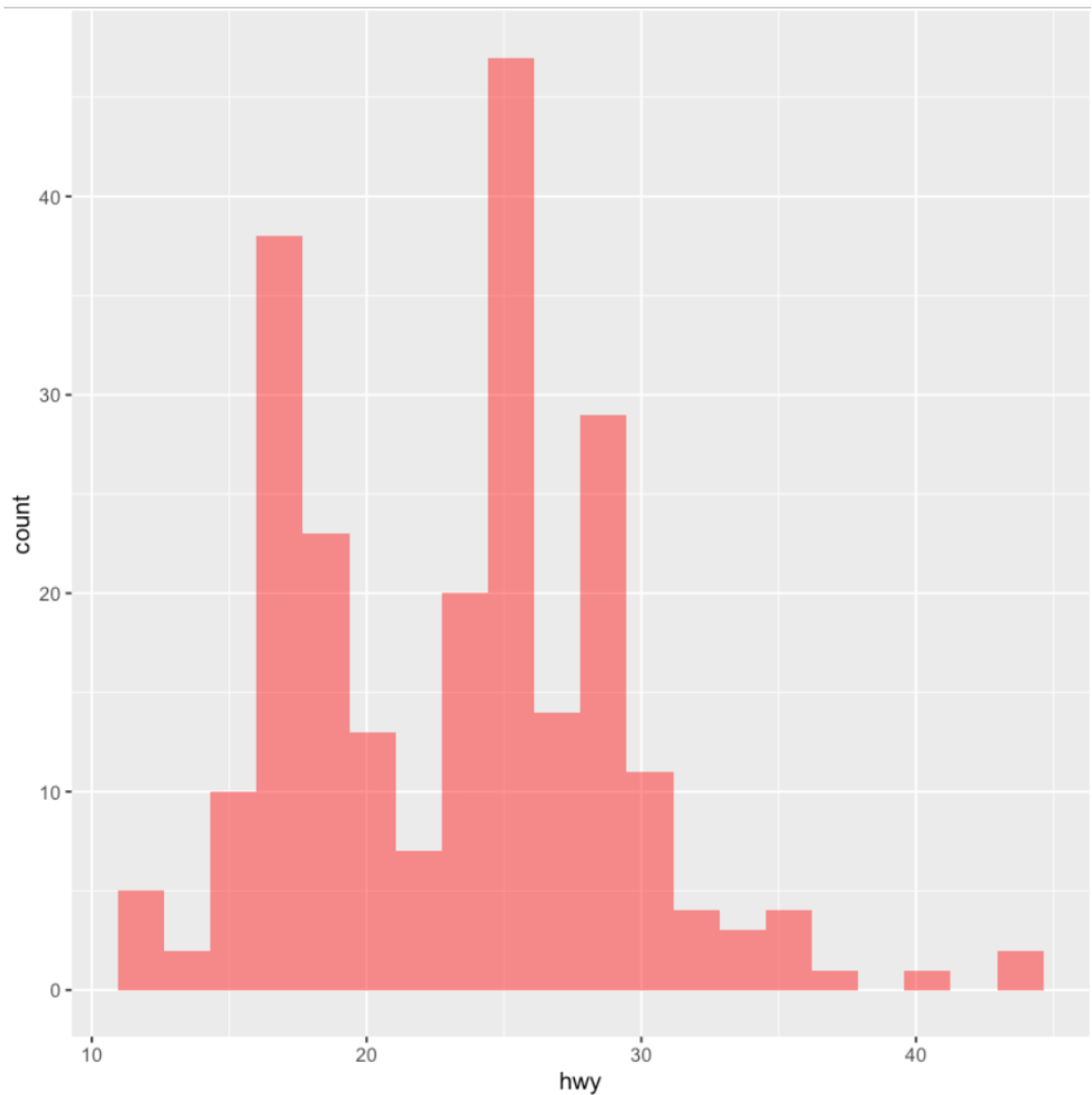
1. Pour les premiers graphes, utilisez le jeu de données mpg

```
library(ggplot2)
library(ggthemes)
head(mpg)
```

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	p	compact
2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	p	compact
3	audi	a4	2	2008	4	manual(m6)	f	20	31	p	compact
4	audi	a4	2	2008	4	auto(av)	f	21	30	p	compact
5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	p	compact
6	audi	a4	2.8	1999	6	manual(m5)	f	18	26	p	compact

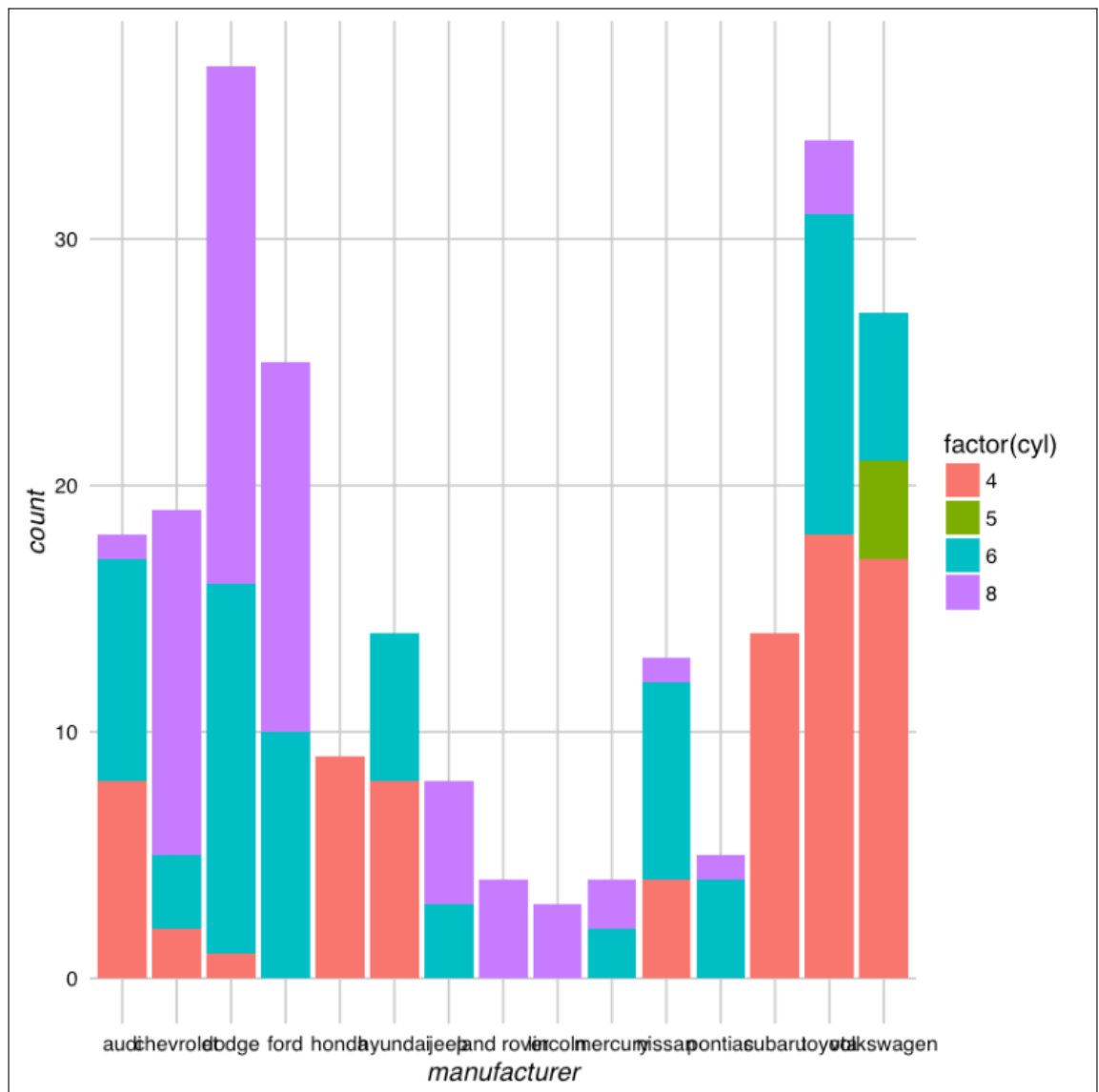
2. Histogramme des valeurs hwy mpg:

```
ggplot(mpg,aes(x=hwy)) + geom_histogram(bins=20,fill='red',alpha=0.5)
```



3. Diagramme à barres du nombre de voitures par fabricant avec un remplissage de couleur défini par le nombre de cylindres

```
ggplot(mpg, aes(x=manufacturer)) + geom_bar(aes(fill=factor(cyl))) + theme_gdocs()
```



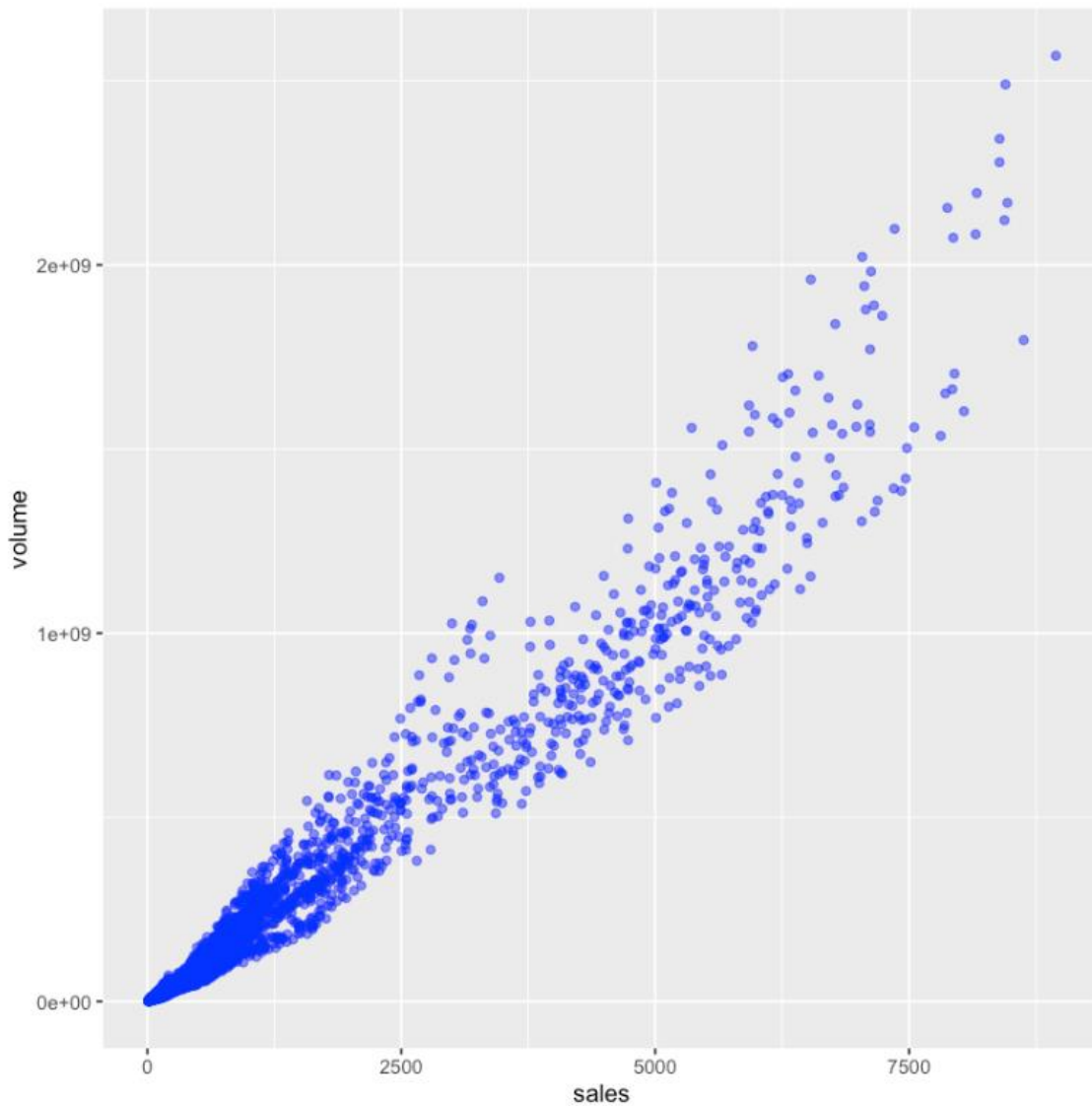
4. Basculez maintenant pour utiliser le jeu de données txhousing fourni avec ggplot2

```
head(txhousing)
```

	city	year	month	sales	volume	median	listings	inventory	date
1	Abilene	2000	1	72	5380000	71400	701	6.3	2000
2	Abilene	2000	2	98	6505000	58700	746	6.6	2000.083
3	Abilene	2000	3	130	9285000	58100	784	6.8	2000.167
4	Abilene	2000	4	98	9730000	68600	785	6.9	2000.25
5	Abilene	2000	5	141	10590000	67300	794	6.8	2000.333
6	Abilene	2000	6	156	13910000	66900	780	6.6	2000.417

5. Créez un diagramme de dispersion du volume par rapport aux ventes. Ensuite, jouez avec les arguments `alpha` et `color` pour clarifier les informations.

```
pl <- ggplot(txhousing,aes(x=sales,y=volume)) + geom_point(color='blue',alpha=0.5)
print(pl)
```



6. Ajoutez une ligne de lissage au nuage de points par le haut. Astuce:  
Vous devrez peut-être chercher `geom_smooth()`

```
pl + geom_smooth(color='red')
```

