

Statamarkdown [RStudio]

Marc Thevenin

5 October, 2022

Table of contents

1	Introduction: Markdown et Statamarkdown dans RStudio	1
2	Statamarkdown de Doug Hemken	4
2.1	Ressources	4
2.2	Installation	4
2.3	Exemple	5
2.3.1	Suite de commandes dans un bloc	5
2.3.2	Suite de commandes dans plusieurs blocs	6
2.4	Commandes externes	7
2.5	Inserer un graphique	8
2.6	Résultat de régression avec collect	9
2.6.1	Programme Stata	9
2.6.2	Insertion du tableau avec <code>\input</code>	10

Remarque: à l'exception de l'insertion d'un tableau au format *Latex*, le document est quasiment mot pour mot identique à celui en format *html*.

1 Introduction: Markdown et Statamarkdown dans RStudio

- **Markdown.**
 - Ce n'est vraiment pas très compliqué, le but, mais il faut un minimum de pratique en édition avec **markdown** (md). Il y a pléthore de tutoriels (un exemple: <https://www.markdownguide.org/cheat-sheet/>)

- Stata permet d’exécuter dans un `.do` un programme documenté en md depuis de nombreuses années: se reporter par exemple à [markstat](#) de *Germán Rodríguez*, de loin la solution la plus satisfaisante. C’est mon avis, mais les solutions dans l’environnement Stata sont et je pense resteront largement en deçà de celles proposées dans un environnement R (*Rmarkdown*, *Bookdown*).
- La dernière évolution en date, **Quarto**, continue de creuser le fossé, en particulier pour générer des documents nativement en Latex. Ce document, diffusé en pdf, est également mis à disposition en format html. Le code source est disponible sur github [lien].....

- **La durée d’exécution**

- Cette présentation montre rapidement comment dans *Rstudio* on peut éditer un document en exécutant simultanément *Stata*. Les durée d’exécution ne sont malheureusement pas optimales malgré les améliorations récentes du logiciel. L’auteur du package *Statamarkdown* *Doug Hemken* l’indique clairement. *Stata* est exécuté en mode batch en s’ouvrant et se fermant systématiquement à chaque bloc de commandes. Cette durée d’exécution est encore plus élevée lorsqu’on compile un document en format pdf, donc nativement en *Latex*. Même si *Statamarkdown* présente un intérêt, je pense que l’ensemble du matériel d’analyse, comme les tableaux et graphiques, devrait resté produit sous environnement *Stata*, et dans un deuxième temps intégré au document.

i Markdown versus Markup

Le markdown est langage d’édition à balisage léger créé en 2002. **Markdown** parce qu’il existe du **Markup**, le plus célèbre étant le HTML (c’est le M), et avec *Stata* on a le très pénible *SMCL* (Stata Markup Control Language).

Le *Latex* n’est pas officiellement il me semble un langage d’édition de type Markup, mais il reste très gourmand en balises. Tel qu’il est intégré à **Rmarkdown** mais plus encore avec **Quarto**, le balisage Latex devient minimal pour ne pas dire inexistant (`\newpage` et `\input` reste très utile).

Bloc de code

- Dans ce document les lignes de commandes seront bien évidemment reportées, ce qui n’est pas le cas si on souhaite par exemple éditer un article, un rapport, une thèse ou un ouvrage. Un bloc de codes ou un fragment de programme est ouvert avec `{nom_langage, options}**` et fermé par `**`. Les options `echo=FALSE`, `warning=FALSE` et `message=FALSE` permettent de ne pas afficher le programme et les messages qui peuvent être affichés dans l’output.

- Si on souhaite cacher intégralement les résultats, on peut ajouter l'option (normalement `quietly` doit suffire) **results=FALSE**.

```
```${stata}
di "{bf: *****AU SECOURS LE SMCL*****}"
```
```

Affichage dans le document:

```
di  "{bf: *****AU SECOURS LE SMCL*****}"
```

```
*****AU SECOURS LE SMCL*****
```

Sans affichage du programme: `echo=FALSE`

```
```${stata, echo=FALSE, warning=FALSE, message=FALSE}
di "{bf: *****AU SECOURS LE SMCL*****}"
```
```

Affichage dans le document :

```
*****AU SECOURS LE SMCL*****
```

Cacher les résultats: `result=FALSE`

```
```${stata, echo=FALSE, warning=FALSE, message=FALSE, results=FALSE}
di "{bf: *****AU SECOURS LE SMCL*****}"
```
```

Pas d'affichage dans le document

Dans la suite du document, les informations qui ouvrent le bloc de codes seront reportées sous forme de titre de cette zone, par exemple :

“{stata, warning=FALSE, message=FALSE, results=FALSE}

```
di  "{bf: *****AU SECOURS LE SMCL*****}"
```

! profile.do

Si vous avez des paramètres de session dans un fichier *profile.do* ([tutoriel](#)), je conseille très vivement de mettre ce fichier dans le répertoire où se trouve le fichier que vous éditez avec *Rmarkdown*, *Bookdown* ou comme ici *Quarto*.

Par exemple, si comme moi le répertoire **adoplus** des commandes externes a été déplacé et son emplacement chargé dans un fichier *profile*, Statamarkdown ne pourra pas chercher ces commandes et les exécuter.

2 Statmarkdown de Doug Hemken

2.1 Ressources

- [Doug Hemken](#)
- [Site tutoriel](#)
- [Dépôt github](#)

2.2 Installation

Installation du package

```
library(devtools)
install_github("hemken/Statamarkdown")
```

Chargement de l'exécutable Stata

Le code R qui suis est à utiliser systématiquement dans chaque document.

```
```{r, echo=FALSE, message=FALSE}
library(Statamarkdown)
stataexe <- "C:/Program Files/Stata17/StataSE-64.exe"
knitr::opts_chunk$set(engine.path=list(stata=stataexe))
```
```

2.3 Exemple

2.3.1 Suite de commandes dans un bloc

“{Stata}

```
sysuse auto, clear
describe
sum price, d
```

(1978 automobile data)

Contains data from C:\PROGRA~1\Stata17\ado\base/a/auto.dta

```
Observations:      74      1978 automobile data
Variables:         12      13 Apr 2020 17:45
                        (_dta has notes)
```

| Variable
name | Storage
type | Display
format | Value
label | Variable label |
|------------------|-----------------|-------------------|----------------|------------------------|
| make | str18 | %-18s | | Make and model |
| price | int | %8.0gc | | Price |
| mpg | int | %8.0g | | Mileage (mpg) |
| rep78 | int | %8.0g | | Repair record 1978 |
| headroom | float | %6.1f | | Headroom (in.) |
| trunk | int | %8.0g | | Trunk space (cu. ft.) |
| weight | int | %8.0gc | | Weight (lbs.) |
| length | int | %8.0g | | Length (in.) |
| turn | int | %8.0g | | Turn circle (ft.) |
| displacement | int | %8.0g | | Displacement (cu. in.) |
| gear_ratio | float | %6.2f | | Gear ratio |
| foreign | byte | %8.0g | origin | Car origin |

Sorted by: foreign

| Price | | |
|-------|-------------|----------|
| | Percentiles | Smallest |
| 1% | 3291 | 3291 |

| | | | | |
|-----|--------|---------|-------------|----------|
| 5% | 3748 | 3299 | | |
| 10% | 3895 | 3667 | Obs | 74 |
| 25% | 4195 | 3748 | Sum of wgt. | 74 |
| 50% | 5006.5 | | Mean | 6165.257 |
| | | Largest | Std. dev. | 2949.496 |
| 75% | 6342 | 13466 | | |
| 90% | 11385 | 13594 | Variance | 8699526 |
| 95% | 13466 | 14500 | Skewness | 1.653434 |
| 99% | 15906 | 15906 | Kurtosis | 4.819188 |

2.3.2 Suite de commandes dans plusieurs blocs

Si on exécute la suite du programme dans un autre bloc:

“{Stata}

```
regress price mpg turn if foreign==1
```

| Source | SS | df | MS | Number of obs | = | 22 |
|----------|------------|----|------------|---------------|---|--------|
| | | | | F(2, 19) | = | 8.51 |
| Model | 68197168.3 | 2 | 34098584.2 | Prob > F | = | 0.0023 |
| Residual | 76166044.4 | 19 | 4008739.18 | R-squared | = | 0.4724 |
| | | | | Adj R-squared | = | 0.4169 |
| Total | 144363213 | 21 | 6874438.7 | Root MSE | = | 2002.2 |

| price | Coefficient | Std. err. | t | P> t | [95% conf. interval] | |
|-------|-------------|-----------|-------|-------|----------------------|-----------|
| mpg | -164.397 | 84.536 | -1.94 | 0.067 | -341.332 | 12.538 |
| turn | 607.204 | 372.318 | 1.63 | 0.119 | -172.067 | 1386.475 |
| _cons | -1.10e+04 | 14587.768 | -0.76 | 0.458 | -4.16e+04 | 19489.258 |

Cela ne marche pas.

On doit ajouter **collectcode=TRUE** en entrée du premier bloc: {Stata, collectcode=TRUE }

```
```{stata, collectcode=TRUE }
sysuse auto, clear
```
```

```
““{Stata, collectcode=TRUE }
```

```
sysuse auto, clear
```

(1978 automobile data)

Puis

```
““{Stata}
```

```
regress price mpg turn if foreign==1
```

| Source | SS | df | MS | Number of obs | = | 22 |
|----------|------------|----|------------|---------------|---|--------|
| | | | | F(2, 19) | = | 8.51 |
| Model | 68197168.3 | 2 | 34098584.2 | Prob > F | = | 0.0023 |
| Residual | 76166044.4 | 19 | 4008739.18 | R-squared | = | 0.4724 |
| | | | | Adj R-squared | = | 0.4169 |
| Total | 144363213 | 21 | 6874438.7 | Root MSE | = | 2002.2 |

| price | Coefficient | Std. err. | t | P> t | [95% conf. interval] | |
|-------|-------------|-----------|-------|-------|----------------------|-----------|
| mpg | -164.397 | 84.536 | -1.94 | 0.067 | -341.332 | 12.538 |
| turn | 607.204 | 372.318 | 1.63 | 0.119 | -172.067 | 1386.475 |
| _cons | -1.10e+04 | 14587.768 | -0.76 | 0.458 | -4.16e+04 | 19489.258 |

2.4 Commandes externes

Si vous avez modifié l’emplacement du répertoire *adoplus* et que l’appel à ce répertoire est dans un fichier *profile.do* ([tutoriel](#)), Statamarkdown ne sera pas capable d’exécuter les commande externes.

La solution est simple: il suffit de coller votre fichier *profile.do*, normalement dans votre répertoire *user*, dans le répertoire du document que vous êtes en train de rédiger.

2.5 Insérer un graphique

Le graphique exporter avec **graph export**. Il est conseillé de l'enregistrer dans le même répertoire que le document généré ou dans un sous répertoire. Pour l'insérer, il suffit d'utiliser l'intégration standard avec markdown: **![nom graph](nom_graph.png){option}** s'il est dans le même répertoire ou **![nom graph](img/nom_graph.png){option}** s'il se trouve dans un sous répertoire, par exemple *img*

“{Stata, message=FALSE}

```
local mopts mc(%80) mlc(black) mlw(*.2) jitter(4)

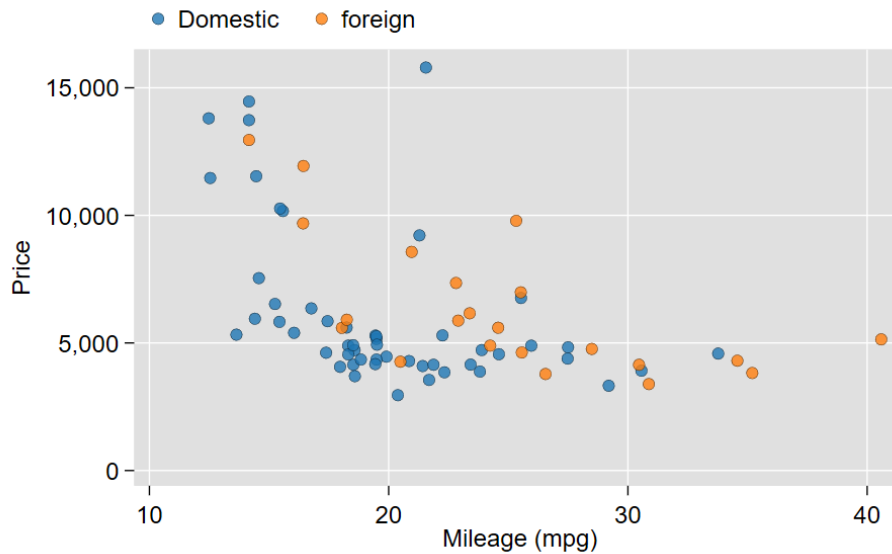
tw scatter price mpg if !foreign, `mopts' ///
|| scatter price mpg if foreign, `mopts' ///
|| , legend(order(1 "Domestic" 2 "foreign") region(color(%0)) pos(11)) ///
    ylabel(, angle(0)) ///
    title("Graphique avec Statamarkown" "dans un document Quarto", pos(11))

qui graph export ///
    "D:\D\Marc\SMS\STATA INED\Programmation\R\statamarkdown\g1.png", replace
```

Directement dans le document

```
![titre graph](g1.png){width=80%}{fig-align="left"}
```


Graphique avec Statamarkown dans un document Quarto



2.6 Résultat de régression avec `collect`

- Se reporter à la courte présentation de `collect` ([lien](#)) introduite avec la version 17 de Stata. On prendra le premier exemple.
- Pour ce document html on utilisera le package R `htmltools` pour insérer un tableau généré avec `collect` et exporter avec `collect export` en html. Pour la version pdf, le tableau sera de type *Latex*. On doit absolument ajouter l'option `tableonly` avec `collect export`.

:::callout-important

On doit absolument ajouter l'option `tableonly` avec `collect export`.

2.6.1 Programme Stata

```
“{Stata, message=FALSE, warning=FALSE}
```

```
qui use https://www.stata-press.com/data/r17/nhanes21, clear  
qui collect clear
```

```

qui collect use multimod, replace

qui: collect _r_b _r_se _r_p: regress bpsystol i.agegrp i.region
qui: collect _r_b _r_se _r_p: regress bpsystol i.agegrp i.region i.sex

qui collect label values cmdset 1 "Modèle 1" 2 "Modèle 2"

qui collect export "D:/D/Marc/SMS/STATA INED/Programmation/R/statamarkdown/ex1", as

```

2.6.2 Insertion du tableau avec \input

Directement dans le corps de texte:

```
\input{ex1}
```

| | | Modèle 1 | | | Modèle 2 | | |
|-----------|--------|-------------|------------|---------|-------------|------------|---------|
| | | Coefficient | Std. error | p-value | Coefficient | Std. error | p-value |
| Age group | 20–29 | 0.000 | 0.000 | | 0.000 | 0.000 | |
| | 30–39 | 2.891 | 0.664 | 0.000 | 2.917 | 0.661 | 0.000 |
| | 40–49 | 9.580 | 0.716 | 0.000 | 9.584 | 0.713 | 0.000 |
| | 50–59 | 18.324 | 0.713 | 0.000 | 18.384 | 0.710 | 0.000 |
| | 60–69 | 24.184 | 0.574 | 0.000 | 24.193 | 0.571 | 0.000 |
| | 70+ | 30.851 | 0.780 | 0.000 | 30.958 | 0.777 | 0.000 |
| Region | NE | 0.000 | 0.000 | | 0.000 | 0.000 | |
| | MW | -0.023 | 0.595 | 0.969 | 0.033 | 0.592 | 0.956 |
| | S | -0.303 | 0.591 | 0.608 | -0.227 | 0.588 | 0.699 |
| | W | -0.777 | 0.601 | 0.196 | -0.745 | 0.599 | 0.213 |
| Sex | Male | | | | 0.000 | 0.000 | |
| | Female | | | | -4.015 | 0.402 | 0.000 |
| Intercept | | 117.632 | 0.589 | 0.000 | 119.672 | 0.621 | 0.000 |