

PAC3 Desenvolupament del treball - Fase 1

Vasyl Druchkiv

Estudiant del Màster de Bioestadística i Bioinformàtica

20 d'Abril 2019

Índice

1	Descripció de l'avenç del projecte	1
1.1	Instal·lació de l'aplicació	1
1.2	Manual i les ajudes del programa	3
1.3	Memoria	6
2	Grau de compliment dels objectius.	7
	Biblilografia	7

1 Descripció de l'avenç del projecte

El comentari del professor després de la primera PAC incluïa els punts següents:

- Pensar en la memòria on he d'explicar els mètodes que utilitzo;
- Pensar en el manual i les ajudes del programa. És fonamental que sigui el més autoexplicatiu possible;
- Preparar un instal·lador de forma que no calgui entrar en el codi. De fet alguns paquets no s'instalen com els altres i això fa que l'execució de l'app no sigui immediata.

Els punts són importants i van definir el full de ruta de la PAC actual. Explico per punt el que he pogut fer per realitzar aquests punts.

1.1 Instal·lació de l'aplicació

La solució més plausible i ràpida era empaquetar tota l'aplicació dins d'un paquet R i fer-la disponible d'aquesta manera en el GitHub. Hi havia també dues opcions més serioses:

- Publicar l'aplicació a CRAN
- Publicar l'aplicació en un servidor Shiny

La primera opció, publicació en CRAN, no he conteplat encara, perquè la solució no és immediata, sino és un procés que no és fàcil i pot tardar fins que el paquet està publicat amb éxit. Com comenta [Wickham, 2015] “submitting to CRAN is a lot more work than just providing a version on github, but the vast majority of R users do not install packages from github, because CRAN provides discoverability, ease of installation and a stamp of authenticity. The CRAN submission process can be frustrating, but it’s worthwhile...”. Normalment els paquets han de ser en perfectes condicions abans d’entregar-los i seran revisats manualmet per a equip dels voluntaris. D’aquesta manera l’aplicació no seria avaluable dins del marc temporal previst per a treball de master. A més a més considero que podria millorar encara més l’aplicació abans d’entregar-lo.

La segona opció, publicació via Shiny Server, és molt interessant, però implicaria un treball considerable per configurar el servidor. Perque ho faria per primera vegada, no puc assegurar que tot estaria preparat a temps. A més a més encara tinc pendents la redacció del manual (que explicaré més endavant) i de la memoria.

Per tant, el paquet **PathwayApp** es pot instal·lar del repositori GitHub seguint els pasos següents:

1. Instal·lar, si encara no està fet, la versió actual de R;
2. Instal·lar, si encara no està fet, el Bioconductor;
3. Instal·lar, si encara no està fet, el paquet **devtools**

```
install.packages( ‘ ‘ devtools ’ ’ )  
library( devtools )
```

4. Instal·lar el paquet **PathwayApp**

```
devtools :: install_github( " vdruchkiv / TFM / 5 _ Packages / PathwayApp / PathwayApp " )
```

5. Iniciar l'aplicació

```
PathwayApp :: runPathwayApp( )
```

La funció **runPathwayApp()** iniciarà la comprovació dels paquets necessaris i començarà l'aplicació. Els paquets següents seran instal·lats, si no ho són encara:

Paquet	Font
clusterProfiler	Bioconductor
ReactomePA	Bioconductor
pathview	Bioconductor
pathviewPatched	GitHub vdruchkiv/TFM
dplyr	CRAN
ggplot2	CRAN
knitr	CRAN
kableExtra	CRAN
formattable	CRAN
shiny	CRAN
shinydashboard	CRAN
shinyhelper	CRAN
shinycssloaders	CRAN

1.2 Manual i les ajudes del programa

Per facilitar l'ús d'aplicació he pensat com es podria fer-ho de manera més intuïtiva possible. Primer cal destacar que com a lengua de manual he elegit anglès per poder fer l'ús d'aplicació més inclusiu possible. Segon, l'usuari pot accedir tan al manual com a ajudes, que es guarden en els arxius .Md separats. Per accedir al manual l'usuari ha de clicar al símbol interrogatori a prop del títol **Pathway analysis**:

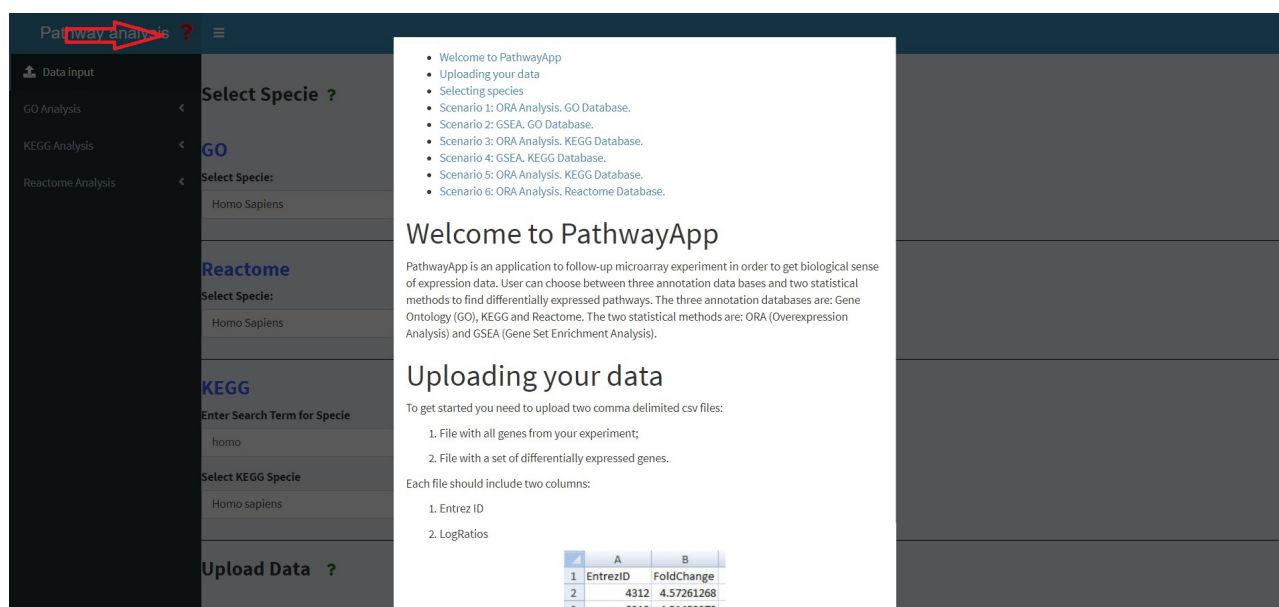


Figure 1: Manual per a aplicació

Com es veu hi ha apartats diferents. Depen dels objectius d'usuari l'usuari pot seleccionar l'apartat que més l'interessa. Així, si l'usuari vol fer l'anàlisi ORA amb l'anotació KEGG ell pot navegar a la secció —textbfScenario 3: ORA Analysis. KEGG Database.

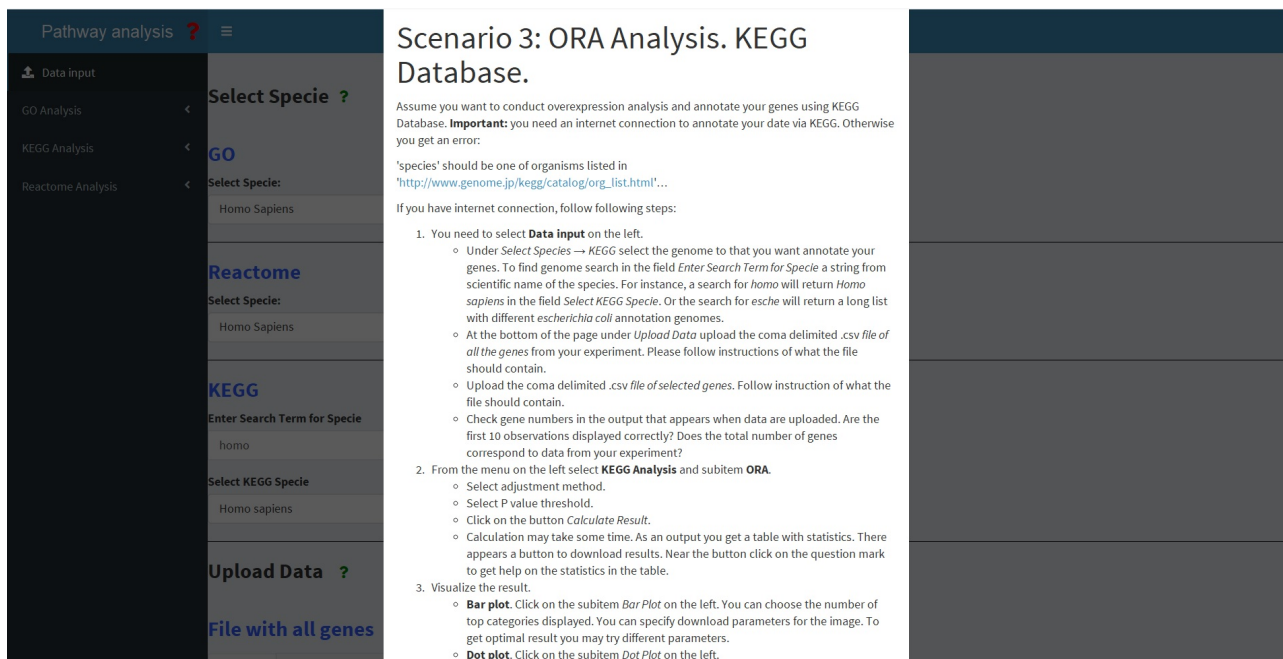


Figure 2: Manual per a l'anàlisi ORA amb l'anotació KEGG

També, l'usuari pot accedir a ajuda clicant als símbols interrogatoris distribuïts per la aplicació en als llocs, que penso poden generar dubtes.

Per fer-ho possible s'utilitza el paquet **shinyhelper** que s'instal·la en executar la funció **runPathwayApp()**.

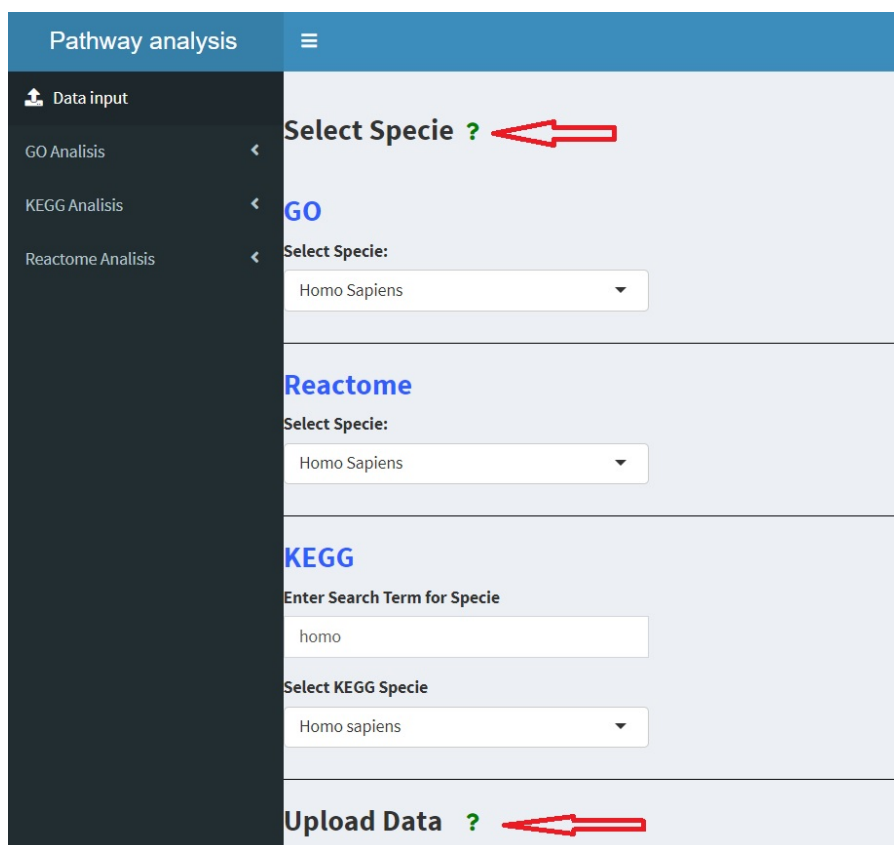


Figure 3: Senyals d'ajuda

El clic en aquests senyal es fa que aparegui una finestra amb la informació d'ajuda.

Aquí són informacions de l'apartat **Data Input**:

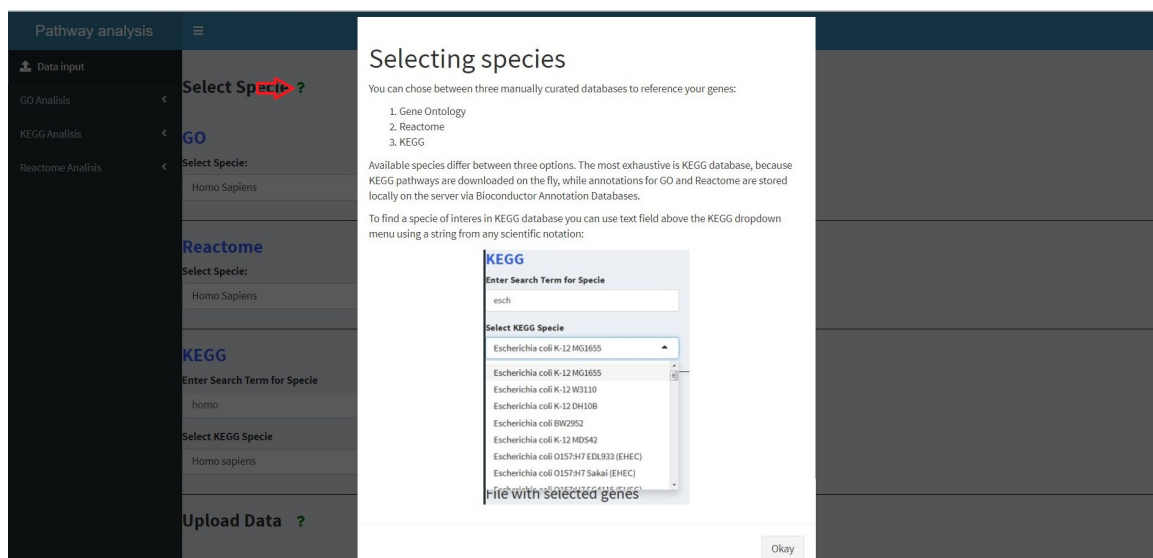


Figure 4: Ajuda per a elecció de l'espècie

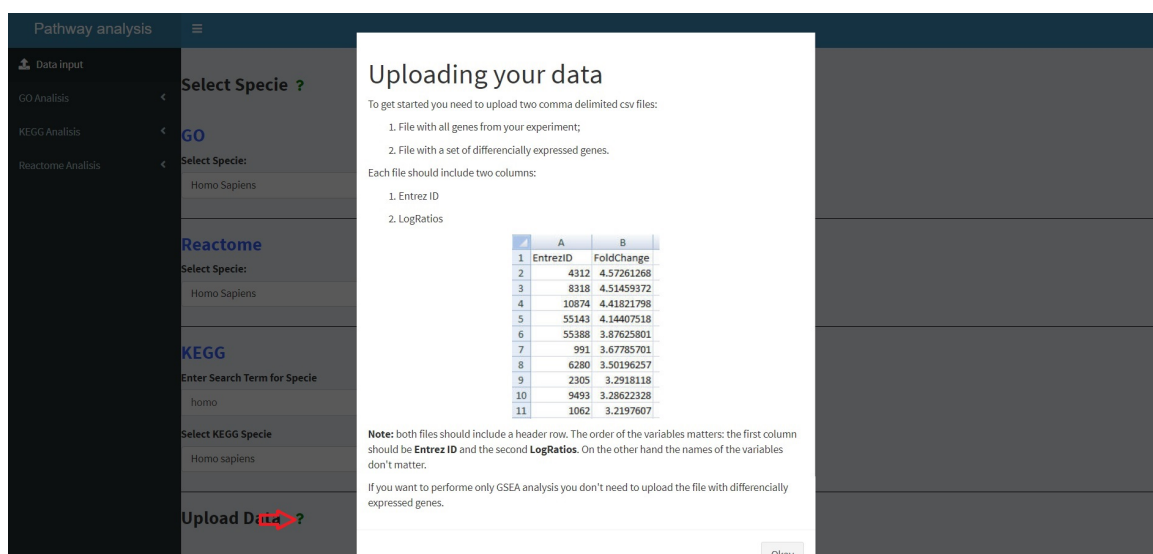


Figure 5: Ajuda per pujar les dades

Les informacions per a apartat ORA són següents:

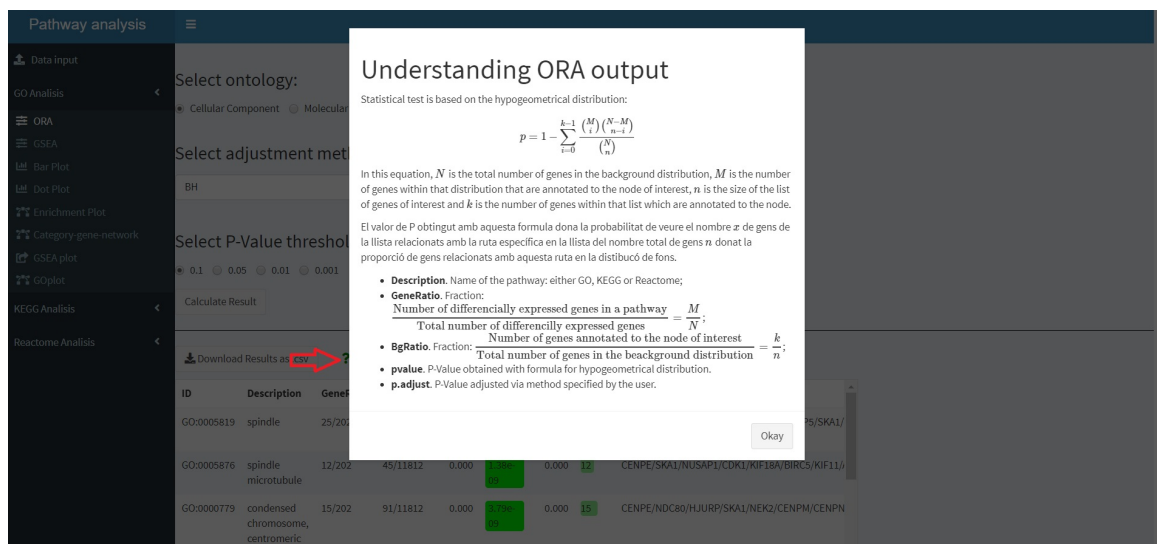


Figure 6: Informació per la interpretació d'anàlisi ORA

Aquí cal destacar que les formulas, depen de l'ordinador, no apareixen degudament en el RStudio Browser. Sí que apareixen bé quan l'aplicació s'obra via el internet browser. L'usuari ha de tenir la connexió amb internet perquè l'aplicació pugui decodificar la formula via MathJax. Encara no he trobat la causa per la qual el Rstudio Browser en alguns ordenadors no visualitza bé les formules. Pot ser un problema amb Java, que s'ha d'actualitzar? Ho estic investigant.

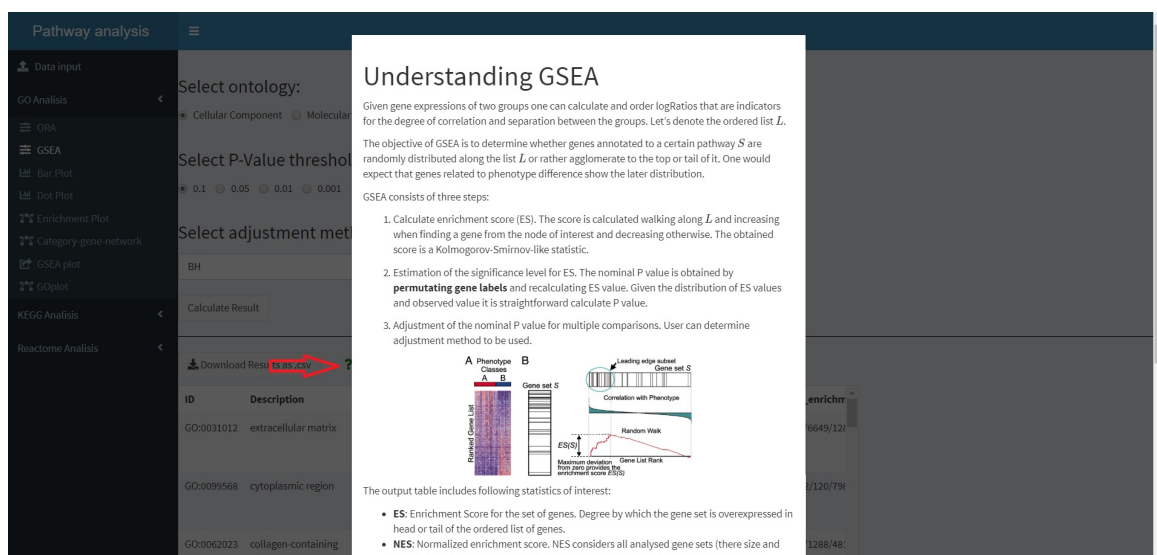


Figure 7: Ajuda per la interpretació de GSEA

1.3 Memòria

Ja he començat descriure els mètodes per a memòria. El draft de la memòria l'adjunto junts amb aquesta PAC. Clarament hi escriuré més informació. Per ara, amb el temps disponible, m'he centrat en la descripció dels mètodes ORA i GSEA. A més a més encara no tinc disponible l'estructura predeterminada d'acord amb el pla docent. En les pròximes dues setmanes tindrà temps suficient per acabar-la.

2 Grau de compliment dels objectius.

Mirem el calendari previst per a aquesta PAC

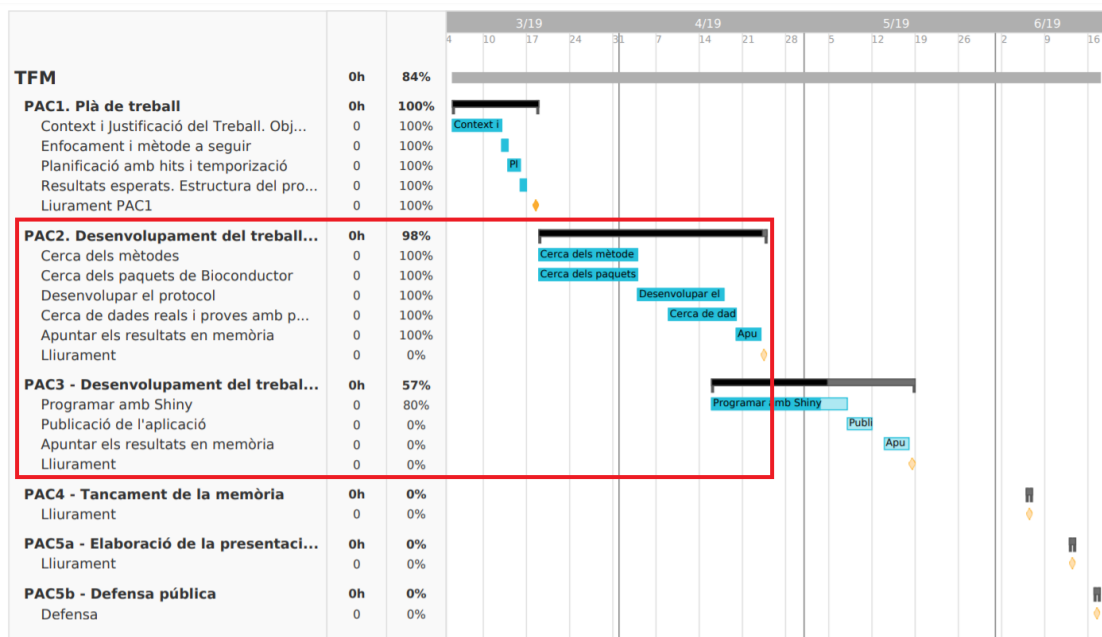


Figure 8: Calendar

Les activitats previstes eren:

1. Programar amb Shiny;

Crec que he pogut millorar l'aplicació afegint el manual i les ajudes. Temporalment m'ha costat trobar i implementar les ajudes però finalment ho he aconseguit i estic content amb el resultat. Les millores en aquesta direcció (afegir més informació o més apartats al manual) ja no trigaran gaire, perquè ja sé com ho fer.

2. Publicació de l'aplicació.

Com he explicat en el primer apartat, la opció de publicació elegida és la via GitHub. Les altres opcions són també possibles però al mateix temps impliquen una feina fora del límit temporal previst per el TFM.

3. Apuntar els resultats en memòria.

Ja ho he començat fent i estic segur que tenint l'aplicació acabada i rebent les instruccions determinades per part del professor acabaré amb la memòria més ràpidament.

Bibliografia

[Wickham, 2015] Wickham, H. (2015). R packages.