



Universitat Oberta
de Catalunya

Anàlisi de dades òmiques (M0-157)

Primera prova d'avaluació contínua.

Presentació i objectius

Aquesta PEC completa la introducció a les òmiques mitjançant un exercici de repàs i ampliació que ens permet treballar amb algunes de les eines que hem introduït durant el curs: Bioconductor i l'exploració multivariant de dades.

Per dur-la a terme, prèviament haureu d'estar familiaritzats amb:

- Les tecnologies òmiques.
- [Bioconductor](#) i les classes que s'utilitzen per emmagatzemar dades òmiques, com els `expressionSets`.
- Git com a eina de control de versions, així com GitHub.
- Les eines estadístiques d'exploració de dades introduïdes al final del primer repte.

L'objectiu d'aquesta PEC és que planifiqui i executeu una versió simplificada del procés d'anàlisi de dades òmiques, alhora que practiqueu amb algunes de les eines i mètodes que hem treballat durant el primer repte.

Descripció de la PEC

A continuació, es detallen les tasques que haureu de dur a terme:

1. Seleccioneu i descarregueu un dataset *de metabolòmica*, que podeu obtenir de [metabolomicsWorkbench](#) o d'[aquest repositori](#) de GitHub.
2. Crea un objecte de classe `SummarizedExperiment` que contingui les dades i les metadades (informació sobre el dataset, les seves files i columnes). La classe `SummarizedExperiment` és una extensió d' `ExpressionSet`, utilitzada per moltes aplicacions i bases de dades (com és el cas de

metabolomicsWorkbench). Quines són les seves principals diferències amb la classe ExpressionSet?

3. Porteu a terme una anàlisi exploratòria que us proporcioni una visió general del dataset en la línia del que hem vist en les activitats d'aquest repte.
4. Elabora un informe que descrigui el procés que heu realitzat, incloent-hi la justificació de la selecció del dataset, la seva incorporació al summarizedExperiment, l'anàlisi exploratòria de les dades i la interpretació dels resultats des del punt de vista biològic. L'extensió màxima d'aquest informe (sense tenir en compte els Annexos) ha de ser de **10 pàgines**, en format PDF.
5. Creau un repositori de GitHub¹ que contingui:
 - l'informe,
 - l'objecte de classe SummarizedExperiment que contingui les dades i les metadades en format binari (. Rda),
 - el codi R per a l'exploració de les dades degudament comentat (el control de versions del mateix s'ha de realitzar amb Git)
 - dades en format text i
 - les metadades acompanyades d'una breu descripció en un arxiu markdown.

El nom del repositori ha de ser 'Apellido1-Apellido2-Nombre-PEC1'. La direcció (*URL*) del repositori haurà d'estar inclosa en l'informe de manera clara. Teniu en compte que a través de CANVAS deveu lliurar únicament l'informe.

Recursos

Els recursos per dur a terme la PEC són els que s'han proporcionat en les diferents activitats del primer repte. A més, haureu de familiaritzar-vos amb la classe SummarizedExperiment. Per això podeu utilitzar [aquest tutorial de Bioconductor](#).

Criteris d'avaluació

Recordeu que la PEC és un exercici de síntesi i aprenentatge, que intenta valorar la vostra capacitat per resoldre un problema molt similar als que es troba un bioinformàtic en el seu dia a dia. Això vol dir que per a algunes de les tasques que haureu de realitzar no hi ha una única manera de procedir. Plantegeu la vostra pròpia solució i expliqueu per què creieu que és l'adequada. Entre altres coses valorarem:

- Capacitat de definir correctament els objectius a assolir.

¹ Si no teniu compte de GitHub n'haureu de crear una, és gratuït i molt senzill.

- Capacitat d'organitzar l'obtenció i preparació de les dades i la seva posterior anàlisi.
- Domini adequat de les eines pròpies del tema (tècniques d'exploració de dades, R, Bioconductor, RMarkdown)
- Capacitat d'explicar què es fa en cada pas i per què es fa.
- Capacitat d'interpretar els resultats obtinguts.
- Capacitat crítica per identificar i discutir les possibles limitacions de l'estudi.
- Capacitat de síntesi i de presentació d'un treball científic en un document amb l'estructura i el format adequats, llegible i ben organitzat.

Aquesta PEC suposa el 30% de la nota final de l'assignatura.

Codi d'honor

Quan presenteu exercicis individuals us adhiereu al codi d'honor de la UOC, amb el qual us comprometeu a no compartir la vostra feina amb altres companys o a demanar que ells ho facin per vosaltres. Així mateix, accepteu que, de procedir així, la qualificació de la PEC serà de zero.