

Università degli Studi di Milano
Laurea Specialistica in Genomica Funzionale e Bioinformatica
Corso di Linguaggi di Programmazione per la Bioinformatica

Package

Giorgio Valentini
e –mail: *valentini@dsi.unimi.it*

DSI – Dipartimento di Scienze dell' Informazione
Università degli Studi di Milano

1

Package in R

I package forniscono uno strumento semplice ed efficiente per gestire collezioni di funzioni e di dati (librerie) e la relativa documentazione.

I package R forniscono librerie liberamente disponibili scritte da sviluppatori esperti in diversi domini applicativi.

Caratteristiche dei package R:

- Caricamento dinamico in memoria: vengono caricati in memoria quando necessario e possono essere scaricati in qualsiasi momento.
- Facili da installare ed aggiornare: le funzioni, dati e documentazione sono installati con un singolo comando che può essere eseguito dall' interno o dall' esterno di R.
- Estendibili ed adattabili alle esigenze degli utenti: gli utenti possono creare propri package (si veda il manuale disponibile on line “*Writing R extensions*”)

2

Installare ed aggiornare package

- I package possono essere installati:
 1. direttamente dal sito CRAN: cran.r-project.org.
selezionando dal menu *Packages* della finestra *RGui* l'opzione *Install packages from CRAN*
 2. da file locali compressi tramite il programma *zip* (file “zippati”) selezionando dal menu *Packages* della finestra *RGui* l'opzione *Install packages from local zip files*
- Analogamente l'aggiornamento di package già installati può essere effettuato selezionando dal menu *Packages* l'opzione *Update packages from CRAN*.

3

Caricare in memoria i package

- Le funzioni ed i dati dei package devono essere caricati in memoria perchè possono essere utilizzabili.
- Per caricare un package si può scegliere uno dei modi seguenti:
 - Dal menu *Packages* selezionare l'opzione *Load package*
 - Dal prompt digitare: `library(nome_package)`

Esempio:

```
> library() # elenca i package correntemente installati
> library(cluster) # package per il clustering
> library(marray) # package per il preprocessing di
  dati di cDNA microarray
```

4

Documentazione sui package

- Per avere accesso alla documentazione sui package installati nel sistema si può selezionare dal menu *Help* l'opzione *HTML help*. Selezionando *Packages* dalla pagina del browser si può scegliere dalla lista dei package il package desiderato ed accedere alle informazioni dettagliate in formato HTML.
- Dal prompt si può anche digitare *help(package=xxx)* per avere informazioni sul package *xxx*
- Spesso i package dispongono della medesima documentazione in formato pdf.
- Ogni funzione, insieme di dati, classe ed in generale ogni oggetto disponibile nei package *dispone di una propria documentazione specifica*.

5

Struttura tipo di una pagina di documentazione

Indipendentemente dal formato della documentazione (HTML, pdf, etc), nella pagina di documentazione di una funzione si trovano solitamente le seguenti sezioni:

- **Nome** della funzione e **package** nella quale è contenuta
- **Description**: descrizione sintetica delle caratteristiche generali della funzione
- **Usage**: sintassi (formato della chiamata della funzione)
- **Arguments**: elenco e descrizione degli argomenti della funzione
- **Value**: valore ritornato dalla funzione
- **Details**: descrizione più dettagliata della funzione (se necessaria)
- **References**: riferimenti bibliografici utili
- **See also**: link ad altre funzioni, oggetti correlati alla funzione
- **Examples**: alcuni esempi di utilizzo della funzione

6

I package di Bioconductor

- **Bioconductor** è un progetto software internazionale il cui obiettivo consiste nel fornire strumenti software di alta qualità per attività professionali e di ricerca nell'ambito della bioinformatica
- Nell'ambito del progetto, sono stati sviluppati e sono tuttora in sviluppo diversi package R per la bioinformatica, liberamente scaricabili dal sito di Bioconductor:
www.bioconductor.org
- Nelle prossime lezioni vedremo in particolare alcuni package per l' *analisi dei dati di espressione genica* e per l' *analisi di biosequenze*.