

Laboratori integrat 1 (Informàtica)

1er de grau de Biotecnologia

<http://mat.uab.cat/~infobiotec>



Universitat Autònoma de Barcelona

Índex

| | | |
|----------|--|-----------|
| 1 | Introducció al sistema operatiu GNU-Linux | 3 |
| 1.1 | Iniciant una sessió en GNU-Linux | 3 |
| 1.2 | Primers passos | 3 |
| 1.2.1 | Navegació amb <i>dolphin</i> | 4 |
| 1.2.2 | Navegació amb terminal | 4 |
| 1.3 | Sistema de directoris. Crear, moure i esborrar | 6 |
| 1.4 | Edició de fitxers | 7 |
| 1.5 | Compressió de fitxers | 8 |
| 1.6 | Obtenció d'ajuda | 8 |
| 1.7 | Sistema de fitxers: permisos | 9 |
| 2 | Manipulació de fitxers | 11 |
| 2.1 | Entrada i sortida. Les "pipes" | 11 |
| 2.2 | <i>sort</i> , ordenació de dades | 12 |
| 2.3 | <i>grep</i> i expressions regulars | 13 |
| 2.4 | <i>awk</i> i manipulació de dades | 15 |
| 2.4.1 | Expressions de l' <i>awk</i> | 15 |
| 2.4.2 | Instruccions de l' <i>awk</i> | 16 |
| 3 | Calc (I) | 19 |
| 3.1 | Text estructurat en LibreOffice Calc: Taules | 19 |
| 3.2 | Funcions Estadístiques | 21 |
| 3.3 | Altres Funcions | 24 |

| | | |
|----------|--|-----------|
| 4 | Calc (II) | 27 |
| 4.1 | Més funcions estadístiques: Regressió lineal | 27 |
| 4.2 | Funcions lògiques | 29 |
| 4.2.1 | Exemple | 29 |
| 4.3 | Taules dinàmiques | 30 |
| 4.3.1 | Exemple 1 | 30 |
| 4.3.2 | Exemple 2 | 31 |
| 4.4 | Exercici d'ampliació | 32 |
| 5 | Aplicació | 33 |
| 5.1 | Introducció | 33 |
| 5.2 | Un experiment | 33 |
| 5.3 | Múltiples experiments | 34 |
| 5.4 | Càlcul de la recta de regressió | 35 |



GNU-Linux¹

1 Introducció al sistema operatiu GNU-Linux

Aquesta primera part del curs veurem instruccions utilitzades en ordinadors i dispositius amb sistemes operatius tipus UNIX. Això inclou els ordinadors d'escriptori i portàtils Apple, els GNU-Linux, la majoria de servidors i els dispositius Android.

Aquestes instruccions serviran per a poder analitzar i estructurar el sistema de fitxers en directoris ordenats, així com per a aplicar el mateix procediment a tots els fitxers i directoris amb una sola instrucció.

1.1 Iniciant una sessió en GNU-Linux

Inicieu l'ordinador i, a la pantalla de tria del sistema operatiu, cliqueu [Iniciar Linux](#). Per a entrar heu de posar el vostre N.I.U. i la vostra paraula clau.

Nomenclatura: D'ara endavant

- `<NIU>` denota el vostre NIU.
- `K` denota la icona que hi ha a la cantonada inferior esquerra de la pantalla (logo de la UAB).

Aclariment: El directori *Home* és: `/home/samba/homes/<NIU>` (normalment, en un ordinador personal, seria `/home/<NIU>`). Els fitxers d'aquest directori es guarden de manera local a l'ordinador on esteu i (possiblement) s'esborra cada cop que entreu, pel que no és aconsellable deixar arxius que vulgueu conservar a aquest espai. El vostre directori permanent (al que podeu accedir des de tots els ordinadors, tant amb GNU-Linux com amb Windows, i des de casa) és el directori [Carpeta Personal en Red](#) que apareix a l'escriptori. Si esteu en un terminal la instrucció és `cd /home/samba/homes/<NIU>/smbhome`.

1.2 Primers passos

Observeu que a la pantalla hi ha una barra inferior, amb moltes icones. Aquesta barra es diu el *plafó*, i conté molts elements. A l'esquerra de tot es troba el menú `K`, que conté tots els programes, repartits en diferents categories. Dediqueu uns moments per a localitzar els programes *firefox*, *kate* i *konsole*.

A la dreta del menú `K` es troba una zona que permet fer servir varis escriptoris diferents a la mateixa sessió per a tenir les finestres més ordenades. A la dreta hi ha un rellotge, i d'altres icones. Aquesta disposició no està fixada *de fàbrica*, sinó que podem variar el contingut i la posició dels objectes (o estris). Per exemple, per a modificar la disposició heu d'anar primer a la part de dalt a la dreta i clicant amb el ratolí triar l'opció [Desbloqueja Estris](#). Ara apareix a la part dreta del plafó una icona que si la cliqueu obre una barra que permet configurar el plafó.

Noteu, però, que els canvis que fem a la configuració del plafó i de l'escriptori es queden en un fitxer de configuració local de l'ordinador, així que si canvieu d'ordinador (o el reinstal·len) no trobareu la mateixa configuració.

¹Aquestes pràctiques estan pensades per a que es puguin fer en un ordinador amb un Linux instal·lat. Al principi parlem de l'escriptori KDE per a que us familiaritzeu amb el sistema. Algunes de les indicacions que es donen a aquesta part tant sols són vàlides per a algunes versions de KDE, però això no afecta al contingut principal de l'assignatura.

Els menús del KDE estan agafat de la versió en català, però no afecta als conceptes i tots són fàcilment interpretables si tenim una versió en castellà o anglès.

1.2.1 Navegació amb *dolphin*

1. Obriu un navegador *dolphin* al vostre directori permanent clicant la icona adequada. Polseu la tecla **F9** per a fer que aparegui/desaparegui a l'esquerra la informació de directoris i unitats connectades al sistema.
2. Creeu un directori que es digui *prac1* (observeu que al sistema d'arxius del GNU-Linux les lletres majúscules són diferents de les minúscules), clicant amb el botó dret del ratolí i seleccionant l'opció **Crea nou → Carpeta**. Entreu-hi.
3. Creeu un fitxer nou que es digui *<NIU>.txt*, clicant amb el botó dret del ratolí i seleccionant el menú **Crea nou → Fitxer de text**. Borreu l'arxiu que acabeu de crear, arrossegant-lo a la paperera.
4. Torneu al directori anterior (fent clic a la fletxa a dalt).

1.2.2 Navegació amb terminal

Per a canviar de directori s'utilitza la instrucció *cd* i per a veure el contingut d'un directori la instrucció *ls*.

1. Obriu un terminal (el programa *konsole*). Executeu *pwd* ("print working directory") per a veure a quin directori esteu.
2. Aneu al directori permanent, mitjançant la instrucció *cd /home/samba/homes/<NIU>/smbhome*. Creeu un altre directori, que es digui *altredir*, amb la instrucció *mkdir altredir* i un fitxer nou, amb *touch altrefitxer.txt*. Feu *ls -l* per a veure que efectivament heu creat el fitxer i el directori (observeu que la línia corresponent a *altredir* comença amb una *d*).
3. Comproveu a la finestra de *dolphin*, que hem fet servir anteriorment, que efectivament s'han creat el directori i el fitxer (si cal s'actualitza prement **F5**).
4. Esborreu tant el fitxer com la carpeta, mitjançant les instruccions *rm altrefitxer.txt* i *rmdir altredir*. Torneu a crear el fitxer *<NIU>.txt*.

Tanqueu el terminal que teniu oberta i obriu-ne una altra. Per a "navegar" pels directoris usarem principalment el terminal. Per tant, ens hem d'acostumar a les seves instruccions. Executeu, successivament:

```
cd .
pwd
cd Desktop
pwd
cd
pwd
cd Desktop/..
pwd
cd /home/samba/homes/<NIU>/smbhome/prac1 ; pwd
cd /home/samba/homes/<NIU>/smbhome/prac1/../../../../ ; pwd
cd /usr; pwd
cd ~ ; pwd
cd /etc/X11; pwd
cd $HOME ; pwd
cd -; pwd
```

Observació: El “fitxer” `.` (punt) denota el directori actual, sigui quin sigui, i el “fitxer” `..` denota el directori del nivell immediatament superior.

Notes:

1. La primera instrucció acaba en “.”.
2. El caràcter `~` equival sempre a *Home* (aquí `/home/samba/homes/<NIU>`) i s’obté prement simultàniament les tecles `Alt Gr` i `Ñ`).
3. `$HOME` és una variable d’entorn que té valor `/home/samba/homes/<NIU>`, però li podem canviar el valor.
4. Podeu obrir un terminal al directori que esteu veient amb el *dolphin* mitjançant la tecla `F4`.

Exercici 1.1

Què passa quan es posen dues (o més) instruccions a la mateixa línia separades per “;” com al final del bloc d’instruccions anterior?

Exercici 1.2

Perquè la primera instrucció (`cd .`) i la setena (`cd Desktop/..`) no fan res?

Exercici 1.3

Trobeu les tres instruccions `cd` del bloc anterior que són equivalents, *independentment* del directori on estiguen situats. Què fan?

Exercici 1.4

Aneu al vostre *Home* (executant `cd` sense cap altre argument). A la llista d’instruccions següent, n’hi ha diverses que fan el mateix. Agrupeu les instruccions equivalents.

```
ls
ls Desktop/..
ls .
ls ..
ls /home/samba/homes/<NIU>/smbhome
ls ~/Desktop
ls $HOME/Desktop
ls /home/samba/homes/<NIU>/smbhome/prac1
ls /
ls /home/samba/homes
```

Finalment, per tal de continuar amb les seccions següents, esborreu el directori `prac1`.

1.3 Sistema de directoris. Crear, moure i esborrar

Hem vist que hi ha un sistema de directoris o carpetes, on cada directori se separa del superior i inferior mitjançant el caràcter / (barra de dividir). També hem vist com moure'ns pels directoris amb la instrucció `cd`.

Per a crear directoris i fitxers en blanc mitjançant la consola s'utilitzen les instruccions `mkdir` i `touch` respectivament.

Exercici 1.5

Creeu al vostre directori de treball *home* un directori que es digui `pract1`, i entreu a aquest directori. Executeu la instrucció `pwd` per a comprovar que esteu a aquest directori.

Per a copiar fitxers i directoris s'utilitza la instrucció `cp`. Podeu consultar el seu funcionament amb el manual de la instrucció: `man cp`.

Exemple 1.1

Si volem copiar el directori `/etc/apt` al directori actual hem d'executar `cp -R /etc/apt .`

Observacions

1. El punt del final de la instrucció és necessari. La instrucció `cp` interpreta l'últim directori (o fitxer) que entrem com el directori (o fitxer) destí.
2. L'opció `-R` és necessària sempre que vulgueu copiar directoris, i vol dir *recursiu*.
3. Podeu veure que ha creat un directori `apt` al directori on som.
4. Possiblement han sortit missatges d'error quan intentàvem copiar fitxers dels que no tenim permís de lectura.

Exercici 1.6

Des del directori `pract1`, executeu la instrucció `cp -R /etc/apt/* .`
Què s'ha copiat ara? Què vol dir l'asterisc de la comanda?

Per a moure fitxers tenim la instrucció `mv`. El seu funcionament és molt semblant a la instrucció `cp`.

Exemple 1.2

Si anem al directori de treball *home* i volem moure el directori `pract1` a `prac1` tant sols hem d'escriure: `mv pract1 prac1`

Exercici 1.7

Si encara no heu esborrat el directori `apt` que ara ha d'estar al directori `prac1`, expliqueu què fa la comanda `mv prac1/apt` .

La mateixa instrucció serveix per a canviar el nom d'un fitxer.

Exemple 1.3

Si volem renombrar el fitxer `sources.list` a que hi ha al directori `apt` al nom `llista` podem fer:

```
cd apt
mv sources.list llista
```

Si el que volem és esborrar fitxers tenim la instrucció `rm`, on a la utilització hem de dir el llistat de fitxers i directoris que volem esborrar. Per a esborrar directoris també hi ha la instrucció `rmdir`, però aquesta exigeix que el directori que volem esborrar estigui buit.

Observació important: La instrucció `rm` esborra de manera permanent el que li diguem. No ho envia a cap part de l'ordinador en que es pugui recuperar. A més, segons com estigui configurat no demana confirmació abans d'esborrar.

Exercici 1.8

Des del directori `home` esborreu el directori `apt` i tot el seu contingut. Proveu-ho primer amb la instrucció `rmdir` (que no hauria de funcionar ja que el directori no està buit) i finalment feu-ho amb la instrucció `rm` (podeu consultar l'ajuda per a veure com es pot esborrar de manera recursiva d'un directori cap endavant).

1.4 Edició de fitxers

Un programa que permet editar (no processar) fitxers és el `kate`. Podeu obrir-lo buscant-lo al menú `K` o bé des del terminal. A més, des del terminal, la instrucció `kate nomdelfitxer &` permet obrir el fitxer dessitjat. També hi ha editors que viuen a dintre del terminal, per exemple l'editor `nano`, que és molt senzill però és útil si esteu treballant amb algun servidor remot i només disposeu del terminal. Obriu el fitxer `<NIU>.txt` que havíeu creat anteriorment, i escriviu-hi la informació següent:

```
Nom: El vostre nom
NIU: El vostre NIU
Grup: El vostre grup de pràctiques
Data: La data actual
```

Guardeu el fitxer, per exemple clicant sobre l'icona `disquet` a la barra superior.

Si només volem llegir un fitxer, no cal obrir-lo amb un editor, podem més senzillament fer que l'ordinador ens mostri el seu contingut. Aneu a un terminal, navegueu fins el directori on es troba el fitxer que acabeu d'editar (`/home/samba/homes/<NIU>/smbhome/prac1`) i executeu `cat <NIU>.txt` i `less <NIU>.txt` . Feu el mateix amb un fitxer més llarg, com ara el fitxer `/etc/X11/rgb.txt` (un fitxer de configuració de l'ordinador).

Exercici 1.9

Quina diferència hi ha entre les instruccions `less` i `cat`?

Observació: Veurem aquestes instruccions amb més detall a la pràctica 2.

1.5 Compressió de fitxers

Per a comprimir un o varis fitxers en un arxiu (d'ara endavant farem servir sempre la diferència de notació: *fitxer* = document o text; *arxiu* = resultat d'una compressió) es fa servir el programa `tar`. Si bé el `tar` és un programa extremadament complex (el seu manual té més de 150 capítols) el seu ús bàsic és força senzill.

Per a comprimir uns fitxers (o directoris) a un arxiu, només cal executar la instrucció `tar cvzf nomdelarxiu.tar.gz llista dels fitxers i directoris per a comprimir` on les opcions `cvzf` indiquen:

- c** Creació de l'arxiu. En alternativa es poden afegir fitxers a un arxiu existent fent servir la opció `a` en lloc de `c`;
- v** Verboositat. Indica que el programa ens ha de llistar tots els fitxers que està compriment.
- z** Compressió (de tipus `gzip`, reflexat també pel fet que l'arxiu acaba per `.gz`. Si s'omiteix, l'arxiu es crearà sense compressió. També es poden fer servir altres tipus de compressions, com ara la de `bzip2`, amb opció `j` i arxiu acabat en `.bz2`, que comprimeix molt més, però és menys coneguda.
- f** Indica que el que ve just després és el nom que li volem donar a l'arxiu. No es pot omitir en l'ús bàsic.

En canvi, per a descomprimir un arxiu existent es fa servir la instrucció següent:

```
tar xvzf nomdelarxiu.tar.gz
```

i els fitxers i directoris es crearan dintre del directori actual. Si no esteu segurs del contingut d'un arxiu, sempre és aconsellable crear un directori nou i descomprimir l'arxiu a dintre, per a evitar confondre els nous fitxers amb els antics. En realitat no és necessari especificar el tipus de compressió a utilitzar en fase de extracció, ja que es pot deduir des de l'extensió.

Observació: També podeu comprimir fitxers i directoris amb el programa gràfic `ark` (si està instal·lat) i, depenent també de la instal·lació, amb els menús que ofereix el botó dret del ratolí des del `dolphin`.

1.6 Obtenció d'ajuda

Gairebé totes les instruccions del GNU-Linux tenen una o més formes d'obtenir ajuda sobre el seu ús. La més ràpida, generalment, és `instrucció --help`, que dona un breu resum de les opcions disponibles. La forma més àmplia és fent servir les pàgines de `info`, que inclouen manuals breus i d'altres més extensos.

Si executeu `ls --help | less` (el signe “|” s'obté amb `AltGr` `1`), veureu que hi ha una llarga llista d'opcions que serveixen per a fer que la instrucció `ls` ens obeeixi. Aneu amunt i avall amb les tecles de direcció, i sortiu prement `q`.

Si executeu `info`, en canvi, entrareu en un programa força complex per a veure manuals, que es diu justament `info`. També us podeu moure a dintre del manual amb les fletxes i sortir amb `q`. A moltes distribucions modernes `ls` té un manual més complet, al que es pot accedir amb la instrucció `info coreutils ls`.

La pàgina inicial de l'`info` conté la llista de tots els manuals instal·lats, i es pot accedir a qualsevol d'ells desplaçant-se (amb les fletxes) a sobre del nom del manual i prement l'`Intro`. Cada manual pot contenir

diferents capítols, que estan marcats amb el signe “*”, i s’hi pot accedir també movent-s’hi a sobre amb el cursor i prement l’[Intro](#). Per a tornar enrere, només cal premer la tecla `1` (de *last*, “anterior”).

Exercici 1.10

Executeu `man gzip`. Quina diferència hi ha entre el que esteu veient i el contingut del manual que s’obté amb `info gzip`?

Exercici 1.11

A partir de la informació del manual de `ls` i de l’efecte de les mateixes instruccions, deduiu què fan la instrucció `ls *txt`, la instrucció `ls -ohX` i la instrucció `ls -mQu`.

Exercici 1.12

Trobeu la forma, fent servir el manual de `info (man info)`, de bolcar a un fitxer el contingut d’un capítol d’un manual. Volqueu al fitxer `core.txt` un capítol del manual de `coreutils`.

1.7 Sistema de fitxers: permisos

El *GNU-Linux* és un sistema multiusuari i té un sistema de fitxers que permet garantir el funcionament del sistema i la privadesa de tots els seus usuaris. Molts dels *servidors* utilitzen aquest sistema operatiu i tenen molts usuaris que treballen a l’hora en aquell ordinador. A més hi ha un usuari (en general s’anomena `root`) que té poder per a administrar la màquina i per tant per a modificar, llegir o esborrar totes les dades del sistema.

Donant els permisos convenients als fitxers i directoris evitem que un altre usuari del sistema pugui veure o esborrar els nostres fitxers².

Per a veure els permisos d’un fitxer o directori utilitzem la opció `-l` de la instrucció `ls`, i mirem la primera columna de les dades.

Exemple 1.4

Mirem la sortida de:

```
ls -l /usr/bin
```

veiem els permisos de cada fitxer/directori a la primera columna.

El primer caràcter (un guió per a la majoria) ens diu si es tracta d’un fitxer (guió), un directori (una lletra `d`) o un enllaç (una lletra `l`).

Llavors venen nou caràcters que s’agrupen de 3 en 3. Cada cadena de 3 està format per `rwX` (`r`read, `w`write, `X`execute), o bé substituint algunes d’aquestes lletres per guions. El primer bloc de 3 lletres afecta al propietari

²Això depèn de com s’ha formatat el dispositiu. Per exemple una clau USB s’acostuma a formatar en sistema `FAT32` que no admet permisos: qualsevol usuari pot esborrar els fitxers.

del fitxer (en aquest cas `root`). El segon bloc al grup (en aquest cas torna a ser `root`, però podria ser diferent) i el tercer bloc a la resta d'usuaris.

Exemple 1.5

La cadena `rwrx-x---` correspondria a un fitxer on el propietari pot llegir, escriure i executar el fitxer; els usuaris del grup assignat poden llegir-lo i executar-lo i la resta d'usuaris no poden ni llegir-lo, ni escriure-hi, ni executar-lo.

Per a canviar els permisos d'un fitxer o directori tenim la instrucció `chmod`. Una de les sintaxis per a executar aquesta instrucció és assignant un número del 0 al 7 a cada bloc, i assigna els permisos depenent de com és en base 2. A la taula següent hi ha el número en base 10, en base 2 i els corresponents permisos:

| Base 10 | Base 2 | Permisos |
|---------|--------|----------|
| 0 | 000 | --- |
| 1 | 001 | --x |
| 2 | 010 | -w- |
| 3 | 011 | -wx |
| 4 | 100 | r-- |
| 5 | 101 | r-x |
| 6 | 110 | rw- |
| 7 | 111 | rxw |

Ara l'ús de la instrucció `chmod` es redueix a dir quins permisos volem per *usuari-grup-tots* en tres dígit seguits entre el 0 i el 7, seguit de l'arxiu o directori del que volem canviar els permisos.

Exemple 1.6

Si tenim el fitxer `hola.txt` i volem que només nosaltres el puguem veure i modificar executem:
`chmod 700 hola.txt`

Exercici 1.13

Quina opció permet canviar els permisos de tots els fitxers i directoris d'un en endavant?

Exercici 1.14

La lletra `x` del les inicials `rxw` en el cas dels directoris no vol dir executar (els directoris no s'executen). Esbrineu que vol dir.

Creeu un directori nou i copieu un fitxer qualsevol dins. Doneu els permisos `rw-rw-rw-` al directori. Podeu entrar al directori? I veure el fitxer que conté?

Ara doneu els permisos `-wx-wx-wx` al mateix directori. Podeu entrar? I llistar el contingut?

2 Manipulació de fitxers

2.1 Entrada i sortida. Les “pipes”

Els programes de terminal GNU-Linux (gairebé) *sempre* tenen una entrada estàndard (`standard input-stdin`), que està connectada al teclat, i una sortida estàndard (`standard output-stdout`), que està connectada a la pantalla.

Hi ha una forma comú de connectar la sortida d'un programa a l'entrada d'un altre, fent servir el connector “pipe”: `|`. D'aquesta manera es formen unes “pipes” de dades, que passen per l'acció de diferents “motors” (programes).

Treballarem diversos exemples perquè pugueu veure la importància de les “pipes”. Necessitem un fitxer llarg que podem crear amb la comanda `ls -l /usr/bin > llarg.txt`. Podem tenir el fitxer obert amb la instrucció `less` per a veure el seu contingut.

1. Creeu el directori `prac2` i executeu `ls -l /usr/bin > llarg.txt`.
2. Executeu `cat llarg.txt`.
3. Executeu `cat llarg.txt | tac`. Aquesta instrucció imprimeix l'entrada en ordre invers, des de l'última línia a la primera.
4. Executeu `cat llarg.txt | nl`; aquesta instrucció renumera les línies.
5. Executeu `cat llarg.txt | nl | tac` i després `cat llarg.txt | tac | nl`. Com veieu, invertint l'ordre dels factors (els motors) canvia el resultat!
6. Executeu `cat llarg.txt | fmt -40`. A que és interessant? Hem agafat un text i l'hem reescrit en una amplada menor!
7. Executeu `cat llarg.txt | fmt -36 | pr -2`. D'aquesta forma podem formatjar un text en dues columnes!!!

Torneu a executar les instruccions anteriors afegint `| less` al final, per poder desplaçar-vos més fàcilment pel text de sortida de la instrucció.

No només existeixen aquestes instruccions, n'hi ha moltes més. A un GNU-Linux basat en Debian (per exemple, l'Ubuntu) podeu trobar una llista d'instruccions bàsiques executant `info coreutils`. Tots aquests programes es poden encadenar i les dades passaran d'un a l'altre com en una cadena de muntatge.

Un altre programa típic que es fa servir amb “pipes”, generalment al final, és el `wc` (de *word count*). Executar `man awk | wc` donarà un recompte del nombre de línies, paraules i caràcters que componen el manual de l'`awk`.

No us explicarem totes les instruccions possibles, si no que intentarem donar les eines necessàries per a poder connectar programes. Només explicarem en detall el funcionament de tres programes: el `grep`, el `sort` i l'`awk`.

Una altra forma de “pipe” és la que deriva la sortida d'un programa cap a un fitxer, fent servir el connector `>` (o `>>` si volem afegir les dades al final del fitxer). També es pot derivar el contingut d'un fitxer cap a l'entrada d'un programa (amb `programa < fitxer`), encara que és totalment equivalent a `cat fitxer | programa`.

Una altra eina important quan estem fent servir “pipes”, és la “T”, que és un “canal secundari” pel qual podem tenir una còpia de les dades que estan passant per la “pipe”. Com sempre al GNU-Linux, el nom d'aquesta eina és senzill, no és més que el nom anglès de la T, `tee`. Fent

`...instruccions | tee UnFitxer | més instruccions`

obtindrem una còpia de les dades, passant per la “pipe”, al fitxer que es diu *UnFitxer* sense afectar de cap forma al funcionament de les altres instruccions. Per exemple

`man awk | tee awk.txt | wc`

dóna, com abans, un recompte del nombre de línies, paraules i caràcters que componen el manual de l'*awk*, però a més tindrem una còpia del manual de l'*awk* al fitxer *awk.txt*.

2.2 `sort`, ordenació de dades

A vegades necessitem ordenar unes dades per algun ordre (alfabètic, numèric, etc.). Si aquestes dades es troben a un fitxer de text podem fer servir la instrucció de terminal `sort`, que vol dir “ordena”.

Normalment, `sort` apareix com a filtre (és a dir com a component d’una “pipe” — o bé al final o bé al mig).

L’ús habitual de `sort` és el següent:

```
...| sort opcions clau1 clau2 ...| ...
```

on cada *clau* indica la columna o columnes que volem fer servir per a ordenar, i les opcions indiquen quin tipus d’ordenació volem. Per exemple, per a ordenar un fitxer alfabèticament, però en el qual hi ha una primera columna (per exemple el NIU d’una persona) seguida del nom, haurem de fer

```
cat fitxer | sort -k2
```

Aquesta instrucció és equivalent a `sort -k2 fitxer` i vol dir que volem una ordenació alfabètica però fent servir de la columna 2 endavant. Per una ordenació numèrica, o bé afegim la opció “-n” o bé afegim la lletra n després de la clau. Per exemple, per ordenar per la columna 2 numèricament i, quan el valor de la columna coincideix, ordenar per la columna 4, alfabèticament, es pot fer servir

```
...| sort -k2,2n -k4,4
```

Les lletres per indicar la forma d’ordenar són les següents:

- n** ordena de forma numèrica. Per exemple, alfabèticament “10” va abans que “9” però numèricament va després.
- r** ordena a l’invers, per exemple de gran a petit, de desembre a gener o de “Z” a “A”, segons les altres lletres posades.
- b** indica que no s’han de considerar espais en blanc. D’aquesta forma les columnes només contenen les dades, sense els espais que les delimiten.
- d** indica que s’ha d’ordenar com si fos un diccionari, tenint en compte només lletres, nombres i espais. Altrament, tot caràcter es pren en consideració.
- f** indica que no s’han de considerar majúscules i minúscules com diferents. Aquest tipus d’ordenació té problemes en algunes instal·lacions de GNU-Linux, ja que de vegades no es reconeix que “ç” és la versió minúscula de “Ç”.

Tota lletra pot servir tant com a “opció” per una clau, o com a “opció general”, si li afegim un guió abans.

Exercici 2.1

Quin tipus d’ordenació representa la instrucció

```
sort -k4,4nr -k2,2f?
```

Doneu una estructura de fitxer de dades on es pugui aplicar.

Per ordenar un fitxer, en lloc de les dades d'una "pipe", només cal afegir al final de tot el nom del mateix fitxer. De tota forma, el mateix resultat s'obté afegint `cat fitxer |` abans de la instrucció d'ordenació. D'altra banda, si volem que les dades ordenades vagin cap a un fitxer, ho hem de fer explícitament, amb un "pipejat" cap a un fitxer (`>`).

Exercici 2.2

Doneu el llistat de fitxers del directori `/usr/bin` ordenats de més vell a més nou en una sola instrucció.

2.3 grep i expressions regulars

El programa `grep` serveix per a extreure les línies que compleixen una condició donada d'entre les línies d'un fitxer o, també, de la sortida estàndard d'una instrucció anterior.

En primer lloc estudiarem el programa `grep` per a processar la sortida estàndard d'una instrucció anterior (és a dir *com a filtre*).

Durant la primera pràctica vàrem executar la instrucció `ls *txt`. La cadena `*txt` és un cas particular del que es diu *expressió regular*, que indica diferents paraules o frases. Per exemple, `*txt` vol dir "qualsevol cosa que acabi amb `txt`" o, més precisament, "qualsevol cosa seguit de `txt`".

La "paraula" buscada pel `grep` també és una expressió regular. Per tant, tot seguit farem un resum de les opcions que pot contenir una expressió regular.

•: Cada caràcter (per exemple `a`) es representa ell mateix.

Rangs: `[acds]` representa qualsevol caràcter en el rang escrit, en aquest cas representa o bé una `a`, o una `c` o una `d` o una `s`. Els rangs seguits es poden indicar amb un guió (`f-m` és el mateix que `fghijklm`). Per exemple `[0-9]` representa qualsevol dígit i `[A-La-l]` representa qualsevol caràcter entre `a` i `l`, tant majúscula com minúscula.

Executeu `ls /usr/bin | grep [r-z]ip` per veure el resultat.

Si afegim un accent circumflex al principi del rang (`[^a-z]`) representa que NO volem els caràcters del rang.

A dintre d'un rang, es poden donar conjunts preestablerts, per exemple `[:alpha:]` que és igual a `a-zA-Z`, però altres que no es podrien escriure d'altra manera, com ara `[:space:]`: els espais i tabulacions, o `[:print:]`: tots els caràcters que es poden veure a pantalla. Per exemple, `a[[:space:]][[:punct:]]a` representa dues lletres `a` separades per un espai o un signe de puntuació.

Especials: Per indicar l'inici de la línia es fa servir `^` (fora d'un rang!) mentre que `$` representa el final de la línia. `\<` i `\>` representen l'inici i el final d'una paraula. Executeu

```
ls /usr/bin | grep zip$
```

per trobar els fitxers que acaben en "zip", i

```
ls /usr/bin | grep "\<zip\>"
```

per trobar els fitxers que es diguin *exactament* "zip".

Repeticions: `?` indica que l'objecte anterior (si cal, entre parèntesis) pot aparèixer o no, però no més d'una vegada. Per exemple `^(go)?ogle` correspon a les línies que comencen per `google` i per `ogle`.

`+` indica que l'objecte anterior ha d'aparèixer una o més vegades. Per exemple `(go)+ogle` correspon tant a `gogoogle` com a `google`, però no a `pirgle`.

`*` indica que l'objecte anterior pot aparèixer o no, i un nombre arbitrari de vegades. Per exemple `g(o)*gle` correspon a `ggle` com a `gooooooooogle`, però no a `gaagle`.

Podem expressar fins i tot la quantitat exacta de vegades que l'objecte ha d'aparèixer. `{5}` volem 5 repeticions. Si hi posem `{5,8}` vol dir que pot aparèixer de 5 a 8 vegades, mentre si posem `{7,}` vol dir que ha d'aparèixer al menys 7 vegades.

Exercici 2.3

Què vol dir la expressió regular següent?

```
^([[[:alnum:]]+[[[:space:]]*]){3}[[[:alnum:]]][.]?zip\>[[[:print:]]].*\<(des)?fer
```

Feu exemples de frases que corresponen a l'expressió i que no hi corresponen. En el segon cas, justifiqueu-ne la manca de correspondència.

L'explicació completa de les expressions regulars es troba al manual de `grep` ([man grep](#)).

Perquè `grep` entengui les expressions regulars en tota la seva potència, és necessari fer servir la opció `-E`, i també les cometes al voltant de l'expressió regular.

L'altra opció important del `grep` és `-v`, que serveix per seleccionar les línies que NO es corresponen amb l'expressió donada. Per exemple `ls | grep -v [aA]` dóna tots els fitxers sense la lletra "a" en el seu nom. Això també ho podeu fer sense fer servir expressions regulars amb la comanda `ls | grep -vi a`.

Practicant amb `grep`: Pels següents exercicis necessitareu el fitxer `observ.txt`. Creeu un directori anomenat `prac2` al vostre directori d'usuari permanent. Seguidament, baixeu-vos el fitxer auxiliar de la pràctica ([practica02-aux.tar.gz](#)) des de la plana web de l'assignatura (<http://mat.uab.cat/~infobiotec/>) i guardeu-lo i descomprimiu-lo al directori que acabeu de crear. Hi trobareu el fitxer `observ.txt` que conté una sèrie d'observacions d'uns investigadors.

1. Llisteu tots els arxius del directori `/usr/bin`, seleccionant i mostrant només els que contenen la paraula `zip` al seu nom. La instrucció és `ls /usr/bin | grep --color=auto zip`.
2. El programa `cal` ens mostra el calendari d'un mes o un any arbitrari. Executeu `cal 6 2015` per veure el calendari de juny. Seleccioneu la setmana que conté el vostre aniversari del 2030. Per fer-ho, heu d'executar `cal 7 2030 | grep "<25>"` si el vostre aniversari és el 25 de juliol.
3. Seleccioneu les línies del fitxer `observ.txt` que contenen la paraula "Lleó" al principi. Això es pot fer de tres formes:


```
cat observ.txt | grep "^Lleó"
grep "^Lleó" < observ.txt
grep "^Lleó" observ.txt
```

Nota: A la última forma `grep` no entra en una "pipe". Treballa directament sobre un fitxer. De fet és l'ús més habitual del `grep`.
4. De les línies anteriors, seleccioneu les que la segona columna (dia d'observació) és 2:


```
cat observ.txt | grep "^Lleó grep -E "[[:alpha:]]+[[[:space:]]]+2"
o bé
cat observ.txt | grep -E "^Lleó[[[:space:]]]+2"
```

Exercici 2.4

Construïu una expressió regular que, aplicada a la comanda

```
ls -l | grep expressió
```

doni tots els fitxers que tenen data del mes actual, però de qualsevol any. Executeu aquesta instrucció aplicada a qualsevol directori de l'ordinador i bolqueu el seu resultat al fitxer anomenat `llistafitxers.txt`.

Exercici 2.5

Esbrineu què fa la instrucció `find`. Què us dona la sortida de la instrucció `find / ?` Què hi podeu afegir a aquesta instrucció per a obtenir un llistat amb els fitxers que contenen la paraula “zip” al seu nom?

2.4 awk i manipulació de dades

L'objectiu de l'`awk` és el d'executar una acció per cada línia de text d'un fitxer (o de l'entrada estàndard) que compleix una condició prefixada. Si no s'especifica cap acció s'imprimeix la línia llegida. Com a conseqüència d'això l'`awk` emula el `grep`:

```
awk '/expressió/' fa el mateix que grep expressió.
```

Nota: Segons quina instal·lació de GNU-Linux s'utilitzi, hi ha diverses versions de l'`awk` instal·lades. Usualment s'anomenen “awk”, “mawk” o “gawk”.

Suposem que volem escriure, de les línies del fitxer d'observacions, només les que es refereixen a lleons, però només volem saber les dates en que s'han fet les observacions (2na i 3ra columna). Per fer això hem d'escriure

```
awk '/Lleó/ {print $2, $3}' observ.txt
```

Aquest exemple ens mostra el patró general d'una instrucció per l'`awk`, que és:

```
'/expressió regular/instruccions'
```

i aquest patró general es pot repetir. Per exemple, la següent instrucció:

```
awk '/^Lleó/ {print "Lleons:",$3} /^Tigre/ {print "Tigres:",$4}' observ.txt
```

imprimirà per pantalla només les línies que comencin amb “Lleó” o “Tigre” i d'aquestes, respectivament, només la columna 3 o la 4. Com a cas especial, la “columna 0” (`$0` indica tota la línia llegida).

En general, el funcionament de l'`awk` és el següent:

1. Agafa una línia de l'entrada (tan d'una “pipe” com d'un fitxer).
2. Revisa si aquesta entrada coincideix amb alguna de les expressions.
3. Per a cada expressió amb la qual l'entrada coincideix, executa les instruccions corresponents.

Hi ha dues expressions, especials que corresponen a l'inici i el final de l'entrada. Són, respectivament, `BEGIN` i `END`. Aquestes expressions no s'han d'incloure entre signes “/” ja que no són expressions regulars.

2.4.1 Expressions de l'awk

A part de la busca d'una expressió regular, l'`awk` pot comprovar condicions molt més complexes. Per exemple, pot veure si una certa expressió regular coincideix amb una certa columna de l'entrada. La instrucció

```
cat observ.txt | awk '$5 ~ /20/ {print}'
```

imprimeix les línies que contenen les lletres "20" a la cinquena columna. A més, es poden fer comparacions, com ara `$1=="Tigre"` o `$4<5`.

També es poden donar condicions sobre el nombre de línies que ja s'han mirat (una mena de comptador) amb la variable `NR` (Number of Rows) i sobre el nombre de columnes de la línia actual, amb la variable `NF` (Number of Fields). L'expressió `$i` fa referència a la columna `i` del fitxer mentre que `$0` fa referència a tota la línia (registre) llegida.

Executeu `awk '{print NR, $0}' observ.txt`. Com veieu, `NR` es pot utilitzar dins d'una instrucció. Executeu ara `cat observ.txt | awk 'NR==6 {print}'`, veureu que només s'imprimeix la línia 6 (compareu amb la sortida de la instrucció anterior).

2.4.2 Instruccions de l'awk

Per a cada expressió que coincideix amb la línia llegida, s'han de posar una o més instruccions que s'han d'executar. La més senzills és `print`, que simplement imprimeix a la pantalla la línia. També es pot "pipejar" l'escriptura, fent servir els operadors `|`, `>` i `>>` com ja heu fet abans. Per exemple, la instrucció `awk '{print $1 >> "primeracolumna.txt"; print $2 >> "segona.txt"}' observ.txt` crea el fitxer `primeracolumna.txt` que conté la primera columna del `observ.txt` i el `segona.txt` que conté la segona.

Es pot instruir `awk` per que faci moltes més coses. Per exemple, la següent instrucció `awk 'BEGIN {c=0;p=0} {c=c+length($0);p=p+NF;} END {print c, p, NR}' observ.txt` escriu el nombre de caràcters, paraules i línies de l'entrada. És a dir, és equivalent a `wc observ.txt`.

En aquest exemple hem fet servir la suma. Podem utilitzar les quatre operacions bàsiques de l'aritmètica (`+` `-` `*` `/`), i a més el "mòdul" `%`, que dóna la resta de la divisió entera, i la potència `^`. A més, hem introduït les variables "`c`" i "`p`", que anem canviant a cada lectura. A l'hora d'imprimir, una variable s'imprimeix com qualsevol altra cosa, posant-la després d'un `print`. També es poden imprimir cadenes de caràcters posant-les entre cometes.

Hi ha també una sèrie de funcions predefinides que es poden utilitzar. Per exemple:

length: Dóna la llargada del seu argument. Exemples: `var = length($0)` dóna a `var` el nombre de caràcters de la línia actual i `length($3)` dóna la llargada de la tercera paraula de la línia actual.

split: Separa l'argument en trossos amb el separador escollit. Exemple: `split($1,trossos,",")` guarda en `trossos[1]...trossos[5]` les diferents parts de `$1` separades pel caràcter `","`. Més precisament, si `$1` és `1,34,hola,t1,5666666`, llavors `trossos[1]=1`, `trossos[2]=34`, `trossos[3]='hola'`, `trossos[4]='t1'`, `trossos[5]=5666666`.

`split($0,columna)` ja sabem que fa: aïlla els trossos de `$0`, que són `$1`, `$2`,..., però amb l'avançatge que després podem fer servir la notació `columna[1]` en lloc de `$1`.

Com veieu, un "programa" `awk` es pot allargar i complicar bastant. Tenim la possibilitat d'*automatitzar* l'`awk` fent que llegeixi les instruccions d'un fitxer, de forma que no cal escriure-les cada vegada. Això es fa amb la opció `-f`. Per exemple, feu servir la instrucció de l'`awk` que trobeu al material auxiliar: `awk -f filtre.awk observ.txt`

Obriu el fitxer d'instruccions `filtre.awk` amb l'editor.

Exercici 2.6

Quin efecte creieu que té l'execució d'aquestes instruccions sobre el fitxer de dades? Comenteu el fitxer.

Com veieu: múltiples instruccions per al mateix bloc s'han de dividir mitjançant signes `;` i es poden posar comentaris, iniciant la línia amb un caràcter `#`.

Exercici 2.7

Modifiqueu el fitxer `filtre.awk` per tal que imprimeixi la mitjana dels pesos dels exemplars observats per a cada grup. Guardeu el resultat com a `mitjanes.awk`. Podeu suposar que no hi ha més de 10 columnes.

Suggerència: el nombre d'exemplars, per a cada línia, es troba a partir de la variable `NF`, restant-li el nombre de columnes que no contenen dades d'observació, el nom de la espècie i la data.)



LibreOffice Calc³

3 Calc (I)

Descàrrega del material de la pràctica

1. Creeu un directori nou de treball que es digui [prac3](#) al vostre directori permanent d'usuari al servidor ([/home/samba/homes/<NIU>/smbhome](#)) o en un subdirectoriu seu.
2. Connecteu-vos a la plana web de l'assignatura: <http://mat.uab.cat/~infobiotec/>; baixeu el fitxer auxiliar del material de la pràctica al directori de treball, [prac3](#), i descomprimiu-lo.
3. Comproveu que heu obtingut el fitxer [calc2.csv](#).

3.1 Text estructurat en LibreOffice Calc: Taules

Iniciu una sessió de *LibreOffice Calc*. Fixeu-vos que un document consta per defecte de tres fulls. De moment treballarem al primer full del document.

D'alguna manera, cada cel·la del *LibreOffice Calc* funciona com un petit text (o una porció de text), i com a tal se li pot donar format. Vegem-ho amb un exemple:

1. Creeu una taula com la següent a partir de la cel·la [B4](#). És en aquesta cel·la on cal escriure "Tipus de lesió":

| Tipus de lesió | Homes | Dones |
|---------------------|-------|-------|
| Contusions | 22 | 17 |
| Fractures | 7 | 8 |
| Inflamacions | 31 | 22 |
| Laceracions | 2 | 1 |
| Torçades | 50 | 62 |
| Estirades musculars | 72 | 59 |
| Altres | 15 | 23 |

2. Adapteu l'amplada de cada columna al text.
3. Començarem donant format a les capçaleres de cada columna: seleccioneu les cel·les [B4-D4](#), i, al menú visual, cliqueu la icona amb l'opció [Negreta](#), indicada per una lletra "B" en negreta.
4. Sempre mantenint aquestes cel·les marcades, cliqueu també la icona corresponent a l'opció [Alinea al centre horitzontalment](#).

³Aquestes pràctiques estan pensades per a que es puguin fer amb diferents versions del LibreOffice i OpenOffice Calc i fins hi tot amb altres fulls de càlcul. Tot i això els noms dels menús estan agafats del LibreOffice 4.2 en català i en altres versions pot ser que algunes de les funcions es trobin a menús lleugerament diferents.

El nom de les funcions matemàtiques estan en català, però la traducció al castellà és, en general, directe.

Si quan l'executeu surt en castellà i voleu canviar la configuració de l'idioma de castellà a català heu d'anar al menú [Herramientas-Opciones-Configuración de idioma-Idiomas-Interfaz de usuario](#) i triar [Catalán](#).

5. Des del menú visual també podeu modificar el tipus de lletra, la mida, tal com si es tractés de text normal. Fixeu-vos que si teniu un seguit de cel·les seleccionades, els canvis es realitzaran a totes aquestes cel·les. Seleccioneu la primera columna de la taula, sense incloure la capçalera, i escolliu el tipus de font **DejaVu Sans** de mida 10pt. Si cal, torneu a modificar l'amplada de la columna **B**.
6. Seleccioneu ara les dues columnes de dades (sense incloure les capçaleres) i cliqueu també la icona amb l'opció **Alinea al centre horitzontalment**.
7. Per últim, seleccioneu la primera columna (igualment sense la capçalera) i cliqueu la icona amb l'opció **Cursiva**, indicada per una lletra "i" en cursiva.

De moment, en aquest exercici, no farem coses gaire complicades. De fet, anem a fer un parell d'exercicis fàcils i a acabar de donar format a aquesta taula:

1. Al costat de l'última columna, inseriu-ne una de nova, amb la capçalera "Total x lesió" i feu la suma de cada fila: per exemple, la cel·la **E5** ha de contenir la fórmula $= C5 + D5$, etc.
 2. Igualment, sota l'última fila, inseriu-ne una de nova, amb la capçalera "Total x sexe", i feu la suma per a cada columna.
 3. Volem que la columna "Total x lesió" mantingui el format de les columnes "Homes" i "Dones". Per a fer-ho, seleccioneu per exemple la columna "Dones" de la taula, i escolliu **Copia**. Ara situeu-vos sobre la cel·la "Total x lesió" i escolliu l'opció **Enganxament especial**. Desactiveu l'opció **Enganxa-ho tot**, i deixeu només seleccionada la cel·la **Formats**. Cliqueu **D'acord**. Si cal, modifiqueu l'amplada de la columna "Total x lesió" per a que es vegi el contingut de les cel·les.
 4. Formateu també la fila "Total x sexe" com la resta de files.
 5. Finalment, després de l'última columna, inseriu-ne una última amb la capçalera "Percentatge".
 6. Per a introduir els percentatges de cada lesió situeu-vos a la cel·la **F5**, seleccioneu la icona "%" del menú visual, que canvia el format de la cel·la, i introduïu la fórmula següent: $=E5/E12$. Arrossegueu la fórmula fins omplir tota la columna... Us sortirà un error. A què és degut? reviseu el contingut de les cel·les marcades amb error: al arrossegar la fórmula, també ha variat el número **E12**. Anem a corregir aquest petit problema: col·loqueu-vos a la cel·la **F5** de nou, i substituïu la fórmula que hi apareix per la següent: $E5/E\$12$.
- IMPORTANT: Si en una fórmula voleu fixar un número de fila o una lletra de columna (o tots dos), només cal que introduïu davant del número o lletra en qüestió el símbol "\$".**

Ara que la taula ja està completa, podem acabar de formatar-la.

1. Seleccioneu tota la taula. Clicant amb el botó dret del ratolí us apareixerà un llistat d'opcions. Escolliu **Formata les cel·les**.
2. Aquí, aneu a l'opció de **Vores**, i, fent servir el diagrama que teniu a l'esquerra de la finestra de format, seleccioneu els següents marges:
 - Les línies horitzontals interiors, marqueu-les com a contínues de 1,00 pt d'amplada.
 - Les verticals interiors, marqueu-les com a contínues de 0,5 pt d'amplada.

- Les línies exteriors de la taula poseu-li doble marge de 3,00 pt d'amplada.

3. I ara l'últim retoc: seleccioneu les cel·les **B2-F2**. Amb el menú **Format → Fusiona les cel·les** combineu les cel·les seleccionades. Ara podeu posar un títol a la taula: "Lesions Esportives de la Temporada". Centreu el títol, poseu-lo en negreta i doneu-li mida 14.

Ja hem acabat amb el format de cel·les (de moment). Per acabar aquest exercici introductori, farem un primer gràfic a partir de les dades de la taula:

1. Seleccioneu el menú **Insereix → Diagrama...**⁴. Us apareixerà el quadre de diàleg **Auxiliar de diagrames** on podreu escollir el rang de les dades i quines capçaleres tindrà el gràfic.⁵
2. Ara podeu escollir el **Tipus de diagrama**. Escolliu un gràfic de **Columna → Apilat** (si manteniu el cursor a sobre de qualsevol dels models durant un parell de segons us apareixerà el nom de cada model). Cliqueu a **Següent**.
3. Per a triar l'**Interval de dades** teniu un petit botó amb una fletxeta cap amunt. Cliqueu-hi. Ara podeu seleccionar sobre la pàgina que vulgueu el rang de les dades. Seleccioneu les cel·les **B4-D11** (les dades inicials de la taula) i torneu a clicar al botonet petit (com que tenim seleccionades les opcions **La primera fila com a etiquetes** i **La primera columna com a etiquetes**, també hi ha marcades la fila i la columna a la selecció). Cliqueu a l'opció **Següent** dos cops.
4. Estem a l'últim pas abans de crear el gràfic. Aquí podeu posar el **Títol** que vulgueu pel gràfic, per exemple "Lesions Esportives".
5. Cliqueu **Finalitza**. Us apareixerà el gràfic a la pantalla. Podeu moure'l i col·locar-lo on vulgueu, així com ampliar-lo, reduir-lo, etc. Fins i tot, podeu modificar els títols, només clicant-hi dues vegades.
6. Observeu també que si ara aneu a la taula i modifiqueu alguna de les dades inicials (per exemple, a la cel·la **D8** substituiu el número que hi figura per "54"), no només es modifiquen els percentatges a la taula, si no que a més es modifica el gràfic automàticament!

Exercici 3.1

Guardeu el document amb format natiu del *LibreOffice Calc* amb el nom **calc1.ods** en el directori **prac3**.

3.2 Funcions Estadístiques

Ara que ja estem més situats podem fer una miqueta més: veurem com fer càlculs estadístics (no només percentatges!) amb el *LibreOffice Calc*.

Aneu ara al segon full del document.

Comenceu important el fitxer **calc2.csv** que es troba al directori **prac3**.

⁴Segons la versió del *LibreOffice* el menú és **Insereix → Objecte → Diagrama...**.

⁵Segons com feu aquest pas us trobareu que no podeu triar l'**Interval de dades** (pas 3.). En aquest cas, us aconsello que seleccioneu les cel·les **B4-D11** abans i després feu **Insereix → Diagrama...**.

Per a importar el fitxer passeu per les opcions del menú **Fitxer → Obre**. Obtindreu el **Diàleg** per a obrir arxius. A **Tipus de fitxer** haureu de seleccionar l'opció **Text CSV**. Obtindreu el **Diàleg** per la importació d'arxius on haureu de fixar les opcions d'importació. En **Opcions del separador** del **Diàleg** heu de marcar **només** l'opció **Punt i coma** (desmarqueu l'opció de coma) i accepteu.

Veureu que, en importar dades, el *LibreOffice Calc* ha obert un nou document.

Anem a traslladar el full del nou document que conté les dades importades al document amb el que estem treballant:

1. Cliqueu amb el botó dret del ratolí sobre la pestanya **Calc2**.
2. Escolliu l'opció **Mou/copia**.
3. Aquí podeu escollir a quin document de destí voleu enviar el full, i en quina posició voleu deixar-lo. Col·loqueu-lo després del Full 1 del document **calc1.ods**.

I ara...

1. A les cel·les **A53** i **A54** escriviu respectivament "Mitjana" i "Desv. Típica".
2. I ara, anem a calcular aquests estadístics: situeu-vos a la cel·la **B53** i seleccioneu el menú **Insereix → Funció...**.
3. A la secció **Categoria** escolliu **Estadística**. A la llista, escolliu la funció que desitgeu i cliqueu-hi dues vegades (per a calcular la mitjana cal escollir la funció **MITJANA**⁶).
4. Ara cal que especifiqueu a quins valors voleu aplicar la fórmula. Comenceu aplicant-la a les dades de la columna **Obs1** (després podreu repetir el procés o simplement arrossegar la fórmula).
5. Seleccioneu **D'acord** i ja està. Feu el mateix per a calcular les desviacions estàndards amb la funció **DESVEST**.
6. També volem calcular el coeficient de correlació entre les dades de la primera columna i les de la segona. Situeu-vos a la cel·la **A56**, i torneu a inserir una funció.
7. A la llista de funcions heu d'escollir **CORREL**⁷. Pel que fa a les dades a les quals s'aplica la fórmula, heu de seleccionar, per **Data_1** les cel·les de la columna **Obs1** (sense la capçalera) i el mateix amb la columna **Obs2** a **Data_2**. Feu **D'acord**.

Com sabreu, el coeficient de correlació mostrat entre dues mostres indica quin nivell de relació lineal hi ha entre les dades d'una mostra i les de l'altre. Si el coeficient de correlació surt proper a 0, vol dir que gairebé no hi ha relació.

Igualment es podrien calcular molts altres estadístics, però ja només es tractaria d'escollir l'opció adequada al problema que vulgueu resoldre...

Finalment, calcularem les freqüències dels valors d'**Obs1**. Les dades d'**Obs1** van de 0 a 1 i es tracta de veure quants individus tenen un valor **Obs1** entre 0 i 0,1, o entre 0,1 i 0,2, etc. Haurem d'utilitzar la funció **FREQUENCIA** que es troba a la categoria de funcions tipus **Matrius**.

⁶En castellà la funció mitjana s'obté amb la instrucció **PROMEDIO**.

⁷A la versió en castellà és **COEF.DE.CORREL**.

1. Escriviu a les cel·les **G1** i **H1** els rètols “Intervals” i “Freqüències”, respectivament.
2. A la columna “Intervals” hem d’escriure els extrems superiors dels intervals dels que volem calcular la seva freqüència. És a dir, els nombres 0,1, 0,2, 0,3, 0,4,...,1. Per a escriure aquesta seqüència de nombres podem recórrer a l'utilització de les referències a les cel·les.
3. Si a la cel·la **G2** escriviu 0,1 i a la cel·la **G3** escriviu la fórmula = **G2 + 0,1** obtindreu el 0,2. Aquesta fórmula pot semblar una manera complicada d’escriure 0,2 però si copieu la cel·la **G3** fins a la cel·la **G11** obtindreu la seqüència completa de nombres.
4. Per a calcular les freqüències, situeu-vos a la cel·la **H2** i inseriu la funció **FREQUENCIA**. A **dades** fixeu el rang de les dades **B2:B51** i a **classes** fixeu el rang dels intervals **G2:G11**
5. Observeu que per defecte el *LibreOffice Calc* afegeix un interval per valors superiors als intervals fixats. També proporciona la seva freqüència.

Exercici 3.2

Fer una taula de freqüències semblant per **Obs2** però que consti només dels intervals “0-0,25”, “0,25-0,5”, “0,5-0,75” i “0,75-1”. Escriviu a la cel·la **J1** el rètol “Intervals” i a les cel·les inferiors els extrems superiors dels intervals. A la cel·les **K1** fins a **K5** escriviu els rètols “Resposta”, “Baixa”, “Moderada”, “Alta” i “Molt alta”, respectivament. A la cel·la **L1** escriu el rètol “Freqüències” i a les cel·les inferiors calculeu les freqüències.

Per a continuar amb el tema dels gràfics, a partir de les dades que tenim, crearem altres tipus de gràfics.

1. El primer gràfic que volem crear és de columnes, similar al de l'exercici anterior, a partir de les freqüències que hem calculat per **Obs1**. Ara, però, volem que les barres siguin verticals.
2. Seleccioneu el rang de cel·les **G1:H11** i cliqueu la icona per a inserir un diagrama.
3. Obtindreu el diàleg **Auxiliar de diagrames**. Heu d'especificar que a la primera fila del rang consten etiquetes i també a la primera columna. Cliqueu el botó **Següent** Seleccioneu **Columnes** com a tipus de gràfic (igualment podeu modificar el títol, etc.).

Ara, farem un gràfic de sectors o circular a partir de la taula de freqüències d'**Obs2**.

1. Seleccioneu el rang de cel·les **K1:L5** i cliqueu la icona per a inserir un diagrama. També ara heu d'especificar que a la primera fila i a la primera columna del rang consten etiquetes. Cliqueu el botó **Següent**
2. Seleccioneu **Diagrames de sectors** com a tipus de gràfic. A continuació, formateu convenientment el gràfic.

El *LibreOffice Calc* té encara més tipus de gràfics i diagrames. L'elecció del gràfic dependrà del que vulgueu fer en cada moment.

Guardau els canvis al fitxer **calc1.ods**.

3.3 Altres Funcions

El *LibreOffice Calc* no és només una eina estadística! Veurem ara com podem fer-lo servir també per a modelar problemes concrets.

Les equacions de Lodka-Volterra és un sistema d'equacions diferencials no lineal de primer ordre que modela un entorn presa-depredador. L'exemple més habitual és utilitzar conills com a preses (x) i guineus com a depredadors (y). Aquest sistema és:

$$\frac{dx}{dt} = \alpha x - \beta xy \quad (1)$$

$$\frac{dy}{dt} = \delta xy - \gamma y \quad (2)$$

on:

- x i y són el nombre de conills i guineus respectivament,
- $\frac{dx}{dt}$ i $\frac{dy}{dt}$ és la variació de conills i guineus respectivament (derivades),
- α els el coeficient de reproducció dels conills i β la seva facilitat per a ser caçats,
- δ és la capacitat de caçar de les guineus i γ el sue coeficient de mortalitat.

Una manera d'estudiar numèricament aquest sistema d'equacions diferencials és suposar que estem a (x_t, y_t) , decidir un increment de temps fixat i calcular la població després d'aquest increment ϵ obtenint $(x_{t+\epsilon}, y_{t+\epsilon})$. Substituint a les equacions queda:

$$\frac{x_{t+\epsilon} - x_t}{\epsilon} = \alpha x_t - \beta x_t y_t \quad (3)$$

$$\frac{y_{t+\epsilon} - y_t}{\epsilon} = \delta x_t y_t - \gamma y_t \quad (4)$$

I aïllant $x_{t+\epsilon}$ i $y_{t+\epsilon}$:

$$x_{t+\epsilon} = x_t + \epsilon(\alpha x_t - \beta x_t y_t) \quad (5)$$

$$y_{t+\epsilon} = y_t + \epsilon(\delta x_t y_t - \gamma y_t) \quad (6)$$

Anem a fer un exemple. Considerem que tenim (en milers) $x_1 = 3$, $y_1 = 1$, els paràmetres $\alpha = 0,04$, $\beta = 0,03$, $\delta = 0,005$ i $\gamma = 0,037$ i l'increment de temps $\epsilon = 1$.

Volem estudiar com evolucionaran aquestes espècies (teòricament, clar):

1. Obriu un document nou.
2. A la primera columna escrivim el nom dels paràmetres a les cel·les **A1-A4**: "Coeficient de reproducció dels conills", "Facilitat de ser caçats", "Capacitat de caçar" i "Coeficient mortalitat guineus".
3. A les cel·les **B1-B4** escrivim els 4 coeficients corresponents.
4. A les cel·les **A6**, **B6** i **C6** entrem, respectivament, "Temps", "Conills (mils)" i "Guineus (mils)".
5. A les cel·les **A7** fins a la **A506** escrivim, respectivament, els números 1, 2, ..., 500.
6. A les cel·les **B7** i **C7** escrivim 1 i 3, que són el nombre inicial de conills i guineus.

- A les cel·les **B8** i **C8** entrem les fórmules corresponents per a calcular quants conills i guineus hi haurà en temps 2: s'ha de fer a partir dels coeficients i el nombre de conills i guineus en temps 1 segons les fórmules (5) i (6).
- Ompliu la resta de cel·les entre **B9** i **C506**.

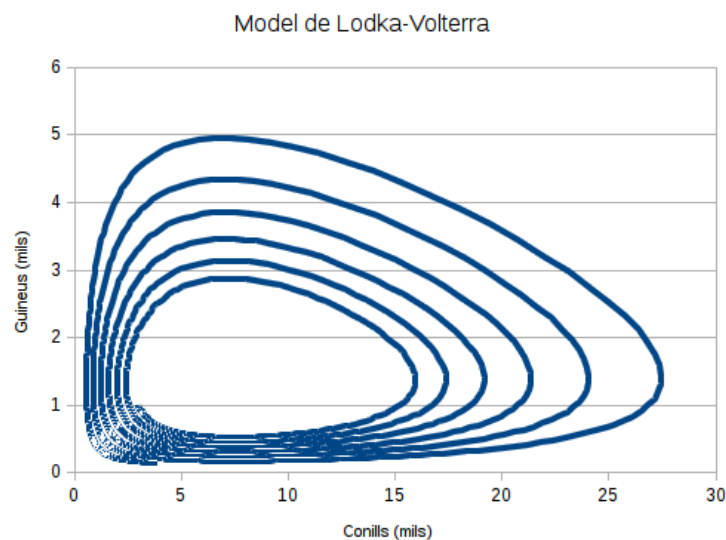
Aprofitarem aquest exemple per veure algun detall més: per defecte, ens han sortit els resultats amb un nombre fixat de decimals, però tant ens podria interessar obtenir més decimals, com obtenir-ne menys. Seleccionem les columnes de dades i cliqueu-hi amb el botó dret. A **Formata → Cel·les...** podeu escollir quin tipus de contingut tindrà la cel·la (text, números, dates, etc.), així com altres detalls relatius al tipus que hageu escollit. Si seleccionem **Números**, veureu que teniu una opció **Nombre de decimals** on podeu indicar els decimals que desitgeu.

En aquest tipus d'estudis sempre és interessant representar gràficament els valors que han resultat. Crearem un gràfic de dispersió amb el qual és molt més fàcil veure l'evolució de cada espècie en funció del temps:

- Seleccionem el rang **A7:C507** i cliqueu la icona per inserir un gràfic. A continuació delimitem un rang on inserir el gràfic.
- Obtenim la finestra **Auxiliar de diagrames**. Seleccionem **Diagrama XY** com a tipus de gràfic. Seleccionem la variant **Punts i línies**. Cliqueu el botó **Següent**.
- Heu d'especificar que a la primera fila del rang consten etiquetes. Cliqueu **Següent** dos cops.
- Desactiveu les dues opcions de **Mostra graella** i poseu al gràfic els títols convenients.
- Una vegada creat el gràfic podeu reformatar-lo, activant el gràfic amb un doble clic i seleccionant cada un dels objectes que el componen per modificar els atributs que considereu. En aquest aspecte és molt útil el botó dret del ratolí per desplegar el menú contextual.

Exercici 3.3

A partir de les dades que heu calculat, feu un gràfic com el següent:



4 Calc (II)

Descàrrega del material de la pràctica

1. Creeu un directori nou de treball que es digui [prac4](#) al vostre directori permanent d'usuari al servidor ([/home/samba/homes/<NIU>/smbhome](#)) o en un subdirectoriu seu.
2. Connecteu-vos a la plana web de l'assignatura: <http://mat.uab.cat/~infobiotec/>; baixeu el fitxer auxiliar del material de la pràctica al directori de treball, [prac4](#), i descomprimiu-lo.
3. Comproveu que heu obtingut els fitxers [regressio.csv](#) i [clients.ods](#).

4.1 Més funcions estadístiques: Regressió lineal

LibreOffice Calc incorpora en el llistat de funcions pròpies un ventall bastant extens de funcions per a l'anàlisi de dades. A continuació descriurem la utilització de les funcions estadístiques per a realitzar la regressió lineal simple d'una variable Y (resposta) sobre una variable X (explicativa).

1. Inicieu una sessió de *LibreOffice Calc*.
2. Importeu el fitxer [regressio.csv](#) que es troba al directori [prac4](#). Aquest fitxer conté el diàmetre del tronc de 31 cirerers a un metre d'alçada i el volum de fusta obtingut després de la tal·la. En el problema de regressió que plantegem ens interessa comprovar si la regressió lineal del *Volum* sobre el *Diàmetre*, permet pronosticar adequadament el volum de fusta que produirà un cirerer a partir del diàmetre del tronc abans de tal·lar-lo. En aquest cas, la variable resposta és el *Volum* i la variable explicativa és el *Diàmetre*.
3. Observeu que el fitxer [regressio.csv](#) presenta dues columnes. La primera columna amb el rètol *Diàmetre* (cel·la [A1](#)) conté els valors observats dels diàmetres (cel·les [A2](#) fins [A32](#)) i la segona anomenada *Volum* (cel·la [B1](#)) conté el valors observats dels volums de fusta (cel·les [B2](#) fins [B32](#)).
4. Formateu amb el menú [Format → Estils i formatació](#) les cel·les [A1](#) i [B1](#) amb l'estil de cel·la [Encapçalament](#). Ajusteu l'amplada de les columnes al text.
5. Una manera gràfica d'explorar si la regressió lineal entre dues variables X i Y és escaient consisteix en realitzar un gràfic de dispersions de Y sobre X per a visualitzar la configuració de punts que resulta. El *LibreOffice Calc* anomena Diagrama XY als gràfics de dispersions d'una variable Y sobre una variable X . Realitzeu un gràfic de dispersió de la variable *Volum* sobre *Diàmetre*.
 - (a) Seleccioneu el rang [A1:B32](#).
 - (b) Trieu el menú [Insereix → Diagrama](#) per a obtenir el diàleg de l'[Auxiliar de diagrames](#).
 - (c) Trieu el tipus de gràfic [Diagrama XY](#). Seleccioneu la variant [Només punts](#). Cliqueu a [Següent](#).
 - (d) Observeu que indica el rang de cel·les que alimenten el gràfic i observeu també que també remarca que prendrà el contingut de la primera fila com etiquetes. Cliqueu [Següent](#) dos cops per a continuar formatant el gràfic.

- (e) Podeu indicar que mostri els objectes de text (títols, llegendes, ...) en el gràfic de previsualització. Desactiveu les opcions de **Mostra la graella**. A **Títol de diagrama** escriviu-hi, per exemple, *Dispersió del Volum sobre el Diàmetre*, a **Mostra la llegenda** desactiveu l'opció, i a **Títol de l'eix** escriviu, a **Eix X**, *Diàmetre* i a **Eix Y**, *Volum*. Cliqueu **Finalitza**.
- (f) Observeu com resulta el gràfic en el full de càlcul. Si cliqueu en qualsevol posició del full fora el gràfic, el gràfic queda inserit en el full. Ara, si feu un simple clic en el gràfic veureu unes marques verdes a les vores del gràfic que us indiquen que podeu moure el gràfic en el full i redimensionar l'espai del gràfic. Moveu i redimensioneu el gràfic per tal que ocupi el rang de cel·les **G5** fins **K22**.
- (g) Feu doble clic en el gràfic. Ara heu activat el gràfic i podeu modificar-lo. Per modificar qualsevol element del gràfic heu de clicar sobre l'element i amb el botó dret del ratolí desplegar el menú contextual per a obtenir l'opció **Propietats de l'objecte**. Veiem un exemple.
- (h) Cliqueu sobre l'eix *Y*. Observeu com queda remarcat amb dues senyals (possiblement verdes). Amb el botó dret del ratolí seleccioneu les **Formata l'eix** i obteniu el diàleg per a modificar l'eix. A la pestanya **Escala** indiqueu que el mínim sigui zero. Cal desactivar l'opció **Automàtic**. A continuació modifiqueu l'eix *X* per tal que l'interval principal sigui d'una unitat.
6. A resultes de l'observació del gràfic de dispersió sembla que és adequat ajustar una recta de regressió per a relacionar *Volum* amb el *Diàmetre*. De manera que proposem el model següent: $Volum = a + b \text{ Diàmetre}$. És a partir de les observacions que estimarem els valors dels paràmetres *a* i *b*. El paràmetre *a* és el terme independent de la recta de regressió i el paràmetre *b* és el pendent de la recta de regressió. En general, parlem de coeficients de regressió.
7. Per a calcular el *terme independent* i el *pendent* de la recta de regressió utilitzareu funcions pròpies de *LibreOffice Calc*.
8. Obtenció del *terme independent*.
- (a) Escriviu a la cel·la **A35** el rètol *Terme independent* i formateu-la cel·la amb l'estil de cel·la **Resultat**. Ajusteu l'amplada de les columnes al text.
- (b) Cliqueu a la cel·la **A36** i inseriu una funció amb les opcions del menú **Insereix → Funció**. Obteniu l'assistent de funcions. Seleccioneu la categoria **Estadística**, de funcions d'anàlisi estadístic. De la llista de funcions busqueu la funció **INTERCEPCIO**⁸ i cliqueu dues vegades per a enganxar la funció a l'editor de fórmules.
- (c) A **dades_Y** seleccioneu el rang de la variable *Volum*, **B2:B32**. A **dades_X** seleccioneu el rang de la variable *Diàmetre*, **A2:A32**. Accepteu.
9. Obtenció del *pendent*.
- (a) Escriviu a la cel·la **A37** el rètol *Pendent* i formateu la cel·la amb estil de cel·la **Resultat**.
- (b) Cliqueu a la cel·la **A38** i inseriu una funció amb les opcions del menú **Insereix → Funció**. Obteniu l'assistent de funcions. Seleccioneu la categoria **Estadística**, de funcions d'anàlisi estadístic. De la llista de funcions busqueu la funció **PENDENT** i cliqueu dues vegades per a enganxar la funció a l'editor de fórmules.

⁸A la versió en castellà aquesta funció es diu **INTERSECCION.EJE**

- (c) Al camp `dades_Y` seleccioneu el rang de la variable *Volum*, `B2:B32`. Al camp `dades_X` seleccioneu el rang de la variable *Diàmetre*, `A2:A32`. Accepteu.
10. Heu obtingut la recta de regressió: $Volum = -37.57 + 5.09 \text{ Diàmetre}$. Podeu utilitzar una fórmula que implementi la recta anterior per a pronosticar un *Volum* donat un valor de *Diàmetre*.
- Situeu-vos a la cel·la `C1` i escriviu el rètol *Predicció*.
 - Situeu-vos a la cel·la `C2`.
 - Escriviu la fórmula `=A$36+A$38*A2`.
 - Copieu la fórmula en el rang `C2:C32`.
 - Heu obtingut el valors de *Volum* pronosticats per la recta corresponents als valors de *Diàmetre*.
11. Clar que també podeu utilitzar una funció pròpia del *LibreOffice Calc* per a realitzar pronòstics. A continuació efectuarem el pronòstic del *Volum* donat un valor de *Diàmetre* de 20.
- Situeu-vos a la cel·la `A39`. Escriviu el rètol *Valor de diàmetre*. Situeu-vos a la cel·la `A40` i escriviu 20.
 - Situeu-vos a la cel·la `B39`. Escriviu el rètol *Pronòstic*.
 - Situeu-vos a la cel·la `B40` i amb les opcions del menú `Insereix → Funció` crideu la funció `PRONOSTIC` de la categoria `Estadística`. Al camp `x` escriviu `A40`, al camp `dades_Y` seleccioneu el rang de la variable *Volum*, `B2:B32`, al camp `dades_X` seleccioneu el rang de la variable *Diàmetre*, `A2:A32`. Accepteu.

Exercici 4.1

Situeu-vos a la cel·la `D1`. Escriviu el rètol *Error*. Calculeu en el rang `D2:D32` la diferència entre els valors observats de *Volum* i els valors pronosticats en el pas 10 elevada al quadrat.

4.2 Funcions lògiques

El *LibreOffice Calc* incorpora una categoria de funcions anomenades de tipus lògic. A continuació implementarem algunes d'aquestes funcions seguint un exemple.

4.2.1 Exemple

A l'Exercici 1 heu calculat els valors de la variable anomenada *Error*. Ara, volem que el *LibreOffice Calc* escrigui en les cel·les del rang `E2:E32` un 1 si l'error que consta a la corresponent cel·la del rang `D2:D32` és superior a 2 i en cas contrari que escrigui 0.

- Situeu-vos a la cel·la `E1`. Escriviu el rètol *Indicador*.
- Situeu-vos a la cel·la `E2` i amb les opcions del menú `Insereix → Funció` crideu la funció `SI` de la categoria `Lògica`.

- Al camp **Verificació**⁹ heu de posar la condició de la funció **SI**. En aquest exemple la condició és **D2>2**.
- Al camp **Then_value**¹⁰ heu d'escriure l'acció si es satisfà la condició. En aquest exemple heu d'escriure un 1.
- Al camp **Otherwise_value**¹¹ heu d'escriure l'acció si no es satisfà la condició. En aquest exemple heu d'escriure un 0.
- Copieu la fórmula en el rang **E3:E32**.

Exercici 4.2

Situeu-vos a la casella **F1** i escriviu el rètol *Indicador 2*. Els valors que ha de prendre aquesta columna (**F2-F32**) són **Molt gran** si l'error que consta a la corresponent cel·la del rang **D2:D32** és superior a 10, **Gran** si està entre 2 (no inclòs) i 10 (inclòs), i **Petit** si és menor o igual a 2.

4.3 Taules dinàmiques

El *LibreOffice Calc* incorpora un eina bastant útil per processar dades i obtenir taules resum que permeten elaborar informes amb agilitat. Aquesta eina és la **Taula dinàmica**. A continuació veurem, seguint un exemple, com utilitzar-lo.

4.3.1 Exemple 1

Iniciu una sessió nova de *LibreOffice Calc* i obriu el fitxer **clients.ods**. Observeu que aquest fitxer conté a la columna **A** un número de comanda, a la columna **B** el codi del client que ha efectuat la comanda, a la columna **C** el número de setmana que ha efectuat la comanda i a la columna **D** les unitats venudes. Volem crear una taula que resumeixi per a cada client el total de comandes realitzades per setmana.

- Seleccioneu el rang **A6:D118**.
- Crideu la **Taula dinàmica** amb el menú **Insereix → Taula dinàmica**. Apareixerà la finestra **Selecciona la font**. Trieu l'opció **Selecció actual** i cliqueu a **D'acord**. Apareixerà el diàleg de la **Taula dinàmica**.
- Arrossegueu el botó **Client** fins la caixa anomenada **Camps de fila**.
- Arrossegueu el botó **Setmana** fins la caixa anomenada **Camps de columna**.
- Arrossegueu el botó **Unitats** fins la caixa anomenada **Camps de dades**. Observeu que per defecte escriurà **Suma-Unitats**.

⁹A la versió en castellà **Prueba_lógica**.

¹⁰A la versió en castellà **Valor_si_verdadero**.

¹¹A la versió en castellà **Valor_si_falso**.

6. Cliqueu sobre el botó **Font i destinació** que es localitza a la part inferior dreta de la finestra. Al camp **Resultats a** seleccioneu l'opció **Full nou**. Accepteu.
7. Observeu que la taula de resultats apareix en un full nou anomenat **Taula_dinamica_Full1_1**.
8. Podeu modificar la taula de resultats activant la **Taula dinàmica**. Per a activar-lo situeu-vos en el full **Taula_dinamica_Full1_1** i cliqueu a la cel·la **Suma-Unitats** i desplegueu el botó dret del ratolí. Trieu l'opció **Edita la disposició**. Obtindreu de nou la finestra de la **Taula dinàmica**.
9. Feu doble clic al botó **Suma-Unitats**. Seleccioneu la funció **Mitjana** i Accepteu. Observeu que ara a la caixa de **Dades** hi trobeu **Mitjana-Unitats**. Accepteu.
10. Observeu que la taula queda modificada amb les mitjanes.

4.3.2 Exemple 2

Seguim amb el fitxer `clients.ods`. Utilitzarem la **Taula dinàmica** per a obtenir una taula que proporcioni la mitjana d'unitats venudes per a cada client. La taula de resultats apareixerà a un full nou.

1. Seleccioneu el rang **A6:D118**.
2. Crideu la **Taula dinàmica** amb les opcions del menú **Insereix → Taula dinàmica**. A continuació trieu l'opció: **Selecció actual**. Apareixerà la finestra de la **Taula dinàmica**.
3. Arrossegueu el botó **Client** fins la caixa anomenada **Camps de Fila**.
4. Arrossegueu el botó **Unitats** fins la caixa anomenada **Camps de dades**. Observeu que per defecte escriurà **Suma-Unitats**.
5. Feu doble clic al botó **Suma-Unitats**. Seleccioneu la funció **Mitjana** i Accepteu. Observeu que ara a la caixa **Camps de dades** consta **Mitjana-Unitats**.
6. Cliqueu sobre el botó **Font i destinació** que es localitza a la part inferior dreta de la finestra. Al camp **Resultats a** seleccioneu l'opció **Full nou**. Accepteu.
7. Obtindreu una taula de dues columnes. La primera columna amb el codi dels clients i la segona columna amb la mitjana aritmètica d'unitats venudes.
8. Realitzem un diagrama de columnes amb la mitjana d'unitats venudes. Seleccioneu la taula continguda al rang **A2:B6**¹² del full actual.
9. Inserir en el full un diagrama de columnes normal, sense línies de quadrícula, sense llegenda, amb el títol principal *Mitjana de vendes*, amb el títol de l'eix X *Clients* i amb el títol de l'eix Y *Mitjana*.
10. Modifiqueu el gràfic per tal que cada barra mostri a la part superior la mitjana corresponent.
11. Copiarem el gràfic a un document de *LibreOffice Writer*. Sense sortir de *LibreOffice Calc*, inicieu una sessió de *LibreOffice Writer*.

¹²Si hi ha el **filtre** activat seran **A4:B8**

12. Al *LibreOffice Calc*, cliqueu sobre el gràfic i desplegueu el botó dret del ratolí. Seleccioneu Copia.
13. Situeu-vos al document de *LibreOffice Writer* i feu Edita → Enganxa.
14. Guardeu el document `clients.ods` i també el document de *LibreOffice Writer* amb el nom `grafic.odt`.

4.4 Exercici d'ampliació

Es disposa de tres tancs industrials per a cultivar bacteris, que produeixen un antibiòtic. Cada tanc s'inocula a temps zero amb una solució del bacteri, i es fa el seguiment del cultiu prenent una mostra cada tres hores durant un dia (un total de 8 mostres per tanc). La lectura es realitza analitzant la concentració d'un segon metabòlit que produeixen els bacteris que és molt més fàcil de quantificar que l'antibiòtic. D'estudis anteriors es coneix l'expressió següent que permet relacionar la concentració d'antibiòtic a l'instant t , que s'indica amb c_t amb la concentració del metabòlit a l'instant t , que s'indica amb m_t :

$$c_t = 1 - 0.5m_t + 0.25m_t^2$$

A més a més, a partir de la segona mostra es vol conèixer l'increment relatiu d'antibiòtic entre dues mostres consecutives, que es defineix per:

$$\frac{c_t - c_{t-1}}{c_{24}}$$

Les dades experimentals per a m_t són:

| Hora | Tanc 1 | Tanc 2 | Tanc 3 |
|------|--------|--------|--------|
| 3 | 1.00 | 1.15 | 1.35 |
| 6 | 2.00 | 1.85 | 2.55 |
| 9 | 3.00 | 2.85 | 3.65 |
| 12 | 4.00 | 4.15 | 4.45 |
| 15 | 5.00 | 5.15 | 5.35 |
| 18 | 6.00 | 5.85 | 6.05 |
| 21 | 6.50 | 6.35 | 6.00 |
| 24 | 6.00 | 5.86 | 5.50 |

En un full de càlcul es demanen:

Per a cada tanc individualment una taula amb les dades següents:

- La primera columna amb l'hora que s'ha pres la mostra.
- La segona columna amb la concentració m_t de cada mostra.
- La tercera columna amb la concentració c_t calculada segons l'expressió anteriorment indicada.
- La quarta columna amb l'increment relatiu d'antibiòtic.
- La cinquena columna ha d'indicar "sup" si el tanc està produint antibiòtic per sobre la mitjana dels tres tancs i "inf" en cas contrari. La comparació s'efectua respecte la mitjana de cada hora.

Pels tres tancs conjuntament una taula amb les dades següents:

- La primera columna amb l'hora que s'ha pres la mostra.
- La segona columna amb la concentració mitjana de m_t que correspon a la mitjana dels tres tancs.
- La tercera columna amb la concentració mitjana de c_t que correspon a la mitjana dels tres tancs.

Aplicació

5 Aplicació

Aquesta pràctica utilitza les instruccions i procediments explicats a les pràctiques anteriors. Els càlculs que surten els haureu de tornar a fer en una altra pràctica de Laboratori Integrat 1, amb dades que heu de trobar vosaltres. Les dades que s'utilitza a aquesta aplicació corresponen a cursos anteriors.

5.1 Introducció

L'activitat de l'enzim fosfatasa àcida es va seguir mesurant l'absorbància en varis experiments independents.

L'equació de Michaelis-Menten

$$v = \frac{V_{max} \cdot [S]}{KM + [S]},$$

on V_{max} i KM són uns paràmetres cinètics, proporciona una expressió per a relacionar la velocitat, v , i la concentració de substrat, $[S]$.

Podem veure que la relació entre v i $[S]$ no és lineal, pel que l'aproximació per la recta de regressió no és una bona solució.

En canvi, podem manipular algebraicament l'equació de Michaelis-Menten (mètode de Lineweaver-Burk) i obtenir:

$$\frac{1}{v} = \frac{1}{V_{max}} + \frac{KM}{V_{max}} \frac{1}{[S]}$$

On veiem que la relació entre $1/v$ i $1/[S]$ sí que és lineal.

Això permet utilitzar els resultats experiments per a calcular el pendent ($\frac{KM}{V_{max}}$) i la intersecció ($\frac{1}{V_{max}}$) de la recta de regressió corresponent.

5.2 Un experiment

En un experiment realitzat amb un enzim d'aquest tipus es varen obtenir els valors següents:

| Concentració (μM) | Velocitat (nmol/min) |
|--------------------------------|----------------------|
| 0,005 | 4,190 |
| 0,02 | 7,449 |
| 0,05 | 8,931 |
| 0,08 | 9,312 |
| 0,12 | 12,204 |
| 0,16 | 13,473 |
| 0,2 | 14,8 |

Calculem ara V_{max} i KM per a aquest experiment:

1. Obriu un document nou de *LibreOffice Calc*.
2. Copieu en un full de càlcul la taula anterior amb les dades de l'experiment.
3. En un **Diagrama XY** representeu la velocitat, v , sobre la concentració de substrat, $[S]$.
4. A les columnes a continuació de la taula de dades, calculeu els recíprocs $1/[S]$ i $1/v$ dels valors experimentals.

5. Trobeu, aplicat les funcions de regressió, els valors per a $1/V_{max}$ i KM/V_{max} . Feu-ho a les files de sota de les dades, deixant lliures les columnes a la dreta de les dades.
6. Obtingueu KM i V_{max} .
7. Calculeu la velocitat de la reacció per a cada concentració $[S]$ utilitzant l'equació de Michaelis-Menten ajustada.
8. En un **Diagrama XY** representeu simultàneament la velocitat observada experimentalment i la velocitat segons la fórmula de Michaelis-Menten sobre la concentració de substrat.

Finalment, guardeu el document amb format *LibreOffice Calc* amb el nom `practica5A.ods`.

5.3 Múltiples experiments

L'arxiu `fosfatasa.tar.gz` conté fitxers amb mostres de concentració i velocitat fetes sota les mateixes condicions per companys vostres d'anys anteriors. A aquesta part volem ajuntar tots els fitxers en un mitjançant instruccions de `bash` i després procedir com a la secció anterior per a calcular KM i V_{max} i el nou coeficient de correlació per a saber si podem considerar que és una aproximació millor.

1. Descarregueu l'arxiu `fosfatasa.tar.gz` de la web de l'assignatura.
2. Descomprimiu-lo a un directori nou (el crea la mateixa instrucció de descomprimir).
3. Veiem que l'arxiu conté tot de fitxers en format `csv`. Amb la instrucció `cat` o `less` podem comprovar que cada fitxer té 2 columnes separades per punt-i-coma. La primera columna correspon a la `Concentracio(μM)` i la segona a la `Velocitat(nmol/min)`.
4. Ara volem crear un fitxer nou `fosfatasa.csv` que tingui totes les mesures. Per a no perdre a quin fitxer correspon cada mostra volem que l'estructura sigui: `fitxer;Concentracio(μM);Velocitat(nmol/min)`. Podem fer això de diverses formes, i aquí n'expliquem una. Aquesta va utilitzant fitxers temporals `temp*.csv` per a anar fent els passos.
 - (a) Creem el fitxer nou amb la línia `fitxer;Concentracio(μM);Velocitat(nmol/min)` fent:


```
echo "fitxer;Concentracio(μM);Velocitat(nmol/min)" > temp1.csv
```
 - (b) Creem un altre fitxer `temp2.csv` amb les línies que comencen per un número de tots els fitxers que en el seu nom segueixen el patró `fitxer*.csv`:


```
grep -E "^[0-9]" fitxer*.csv > temp2.csv
```
 - (c) Veiem que la separació entre el nom del fitxer i la concentració és un dos punts, pel que substituïm tots els dos punts per punt-i-coma al fitxer amb l'ajuda del `awk`:


```
awk 'sub(/:/, ";")1' temp2.csv > temp3.csv
```
 - (d) Ordenem el fitxer `temp3.csv` segons el valor de la concentració. Per a utilitzar la instrucció `sort` necessitem dir-li que els camps estan separats per punt-i-coma i són numèrics:


```
sort -t";" -k2,2n temp3.csv > temp4.csv
```
 - (e) Ajuntem ara els fitxers `temp1.csv` i `temp4.csv` en un que es digui `fosfatasa.csv`:


```
cat temp1.csv temp4.csv > fosfatasa.csv
```
 - (f) Esborrem els fitxers temporals que hem fet servir:


```
rm temp*
```

5. Ara podem importar el fitxer [fosfatasa.csv](#) des del *LibreOffice Calc* i fer els mateixos passos que a la secció anterior.

Després de fer tots els càlculs guardeu el document amb el nom [fosfatasa.ods](#).

5.4 Càlcul de la recta de regressió

Si volem calcular la recta de regressió sense utilitzar les funcions **PENDENT** i **INTERCEPCIO** del *LibreOffice* podem utilitzar les fórmules següents:

Si tenim el núvol de punts (x_i, y_i) amb $1 \leq i \leq n$ considerem:

La mitjana de les x_i : $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$.

La mitjana de les y_i : $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$.

La suma de quadrats de les x : $s_{x^2} = \sum_{i=1}^n x_i^2$.

La suma dels productes $x_i y_i$: $s_{xy} = \sum_{i=1}^n x_i y_i$.

Llavors la recta de regressió $y = \alpha x + \beta$ es pot calcular com:

$$\alpha = \frac{s_{xy} - n\bar{x}\bar{y}}{s_{x^2} - n\bar{x}^2} \quad \text{i} \quad \beta = \bar{y} - \alpha\bar{x}.$$

Com a exemple, obrim el document [practica5A.ods](#) i calculem els coeficients de la recta de regressió amb aquestes fórmules. Recordem que en aquest cas les coordenades x són a la columna 1/[S] (probablement la **C**) i les y a la 1/v (probablement la **D**).

1. Calculeu a les dues columnes següents, mitjançant una fórmula, els quadrats de les coordenades x_i (x_i^2) i els productes $x_i y_i$.
2. A la zona de sota de les columnes amb les dades, escriviu en columna [mitjana x](#), [mitjana y](#), [suma x^2](#), [suma xy](#) i calculeu al costat els resultats corresponents.
3. Utilitzeu els càlculs anteriors per a calcular α i β segons les fórmules i compareu-les amb el resultat de les funcions **PENDENT** i **INTERCEPCIO**.

Exercici 5.1

Creeu un full nou al fitxer [fosfatasa.ods](#) amb les pendents i les interseccions de les rectes de regressió corresponents a cada fitxer de dades originals.

Indicació: Després d'afegir les columnes com a l'exemple, feu una taula dinàmica amb les dades necessàries per a calcular els coeficients de la recta de regressió de cada fitxer.

Índex alfabètic

GNU-Linux, 11

- >, 11, 13
- ~, 5
- ., 5
- /, 6
- ;, 5
- awk, 15
- cal, 14
- cat, 7, 11
- cd, 3
- cp, 6
- echo, 34
- expressió regular, 13
- find, 15
- fmt, 11
- grep, 13, 14
- help, 8
- info, 8
- less, 7, 11
- ls, 4
- man, 6, 8
- mkdir, 4
- mv, 6
- nl, 11
- pr, 11
- pwd, 4
- rm, 4, 7
- rmdir, 4, 7
- sort, 12
- tac, 11
- tar, 8
- touch, 4
- wc, 11

LibreOffice

- Alinea, 19
- Copia, 22
- CORREL, 22
- CSV, 22
- Cursiva, 20
- DESVEST, 22
- Diagrama, 21
- Enganxament Especial, 20
- Fórmula, 20
- Fonts, 20

Formata les cel·les, 20

- FREQUENCIA, 22
- Fusiona cel·les, 21
- INTERCEPCIO, 28
- MITJANA, 22
- Mou, 22
- Negreta, 19
- PENDENT, 28
- PRONOSTIC, 29
- Regressió lineal, 27
- SI, 29
- Taula dinàmica, 30
- Taules, 19
- Vores, 20