

1 Introducció

Mètodes empírics 2

11/04/2023

Avui

- Organització del curs
- Sinopsi del curs
- Replicabilitat i reproduïbilitat
- R
- Estadística descriptiva

Organització

Informació bàsica

- 5 ECTS (~125-150 hores)
- Material i comunicació: <https://aulaglobal.upf.edu/>
- Equip
 - Thomas Brochhagen (thomas.brochhagen@upf.edu)
<https://brochhagen.github.io>
 - Guillermo Montaña (guillermo.montana@upf.edu)
- Estructura
 - Dimarts: Conceptes i discussió
 - Divendres: Pràctica i més discussió

Horari grup gran

Per facilitar que els estudiants tinguin temps de canviar d'aula, caldrà deixar sempre 10 minuts de cortesia a l'inici de la sessió [...]

⇒ 12:40 - 14:25

Avaluació

- Exercicis setmanals (20%, no recuperable)
- Exercici pràctic (20%, no recuperable)
- Revisió per parells (10%, no recuperable)
- Informe: disseny i execució d'anàlisi (50%, recuperable)

Avaluació

Exercicis setmanals (20%, no recuperable)

- A través d'Aula Global
- Correcció automàtica
- 5 en total

Avaluació

Exercici pràctic (20%, recuperable)

- *take home exam/study*
- A través d'Aula Global
- A meitat de curs

Avaluació

Disseny i execució d'anàlisi (50%)

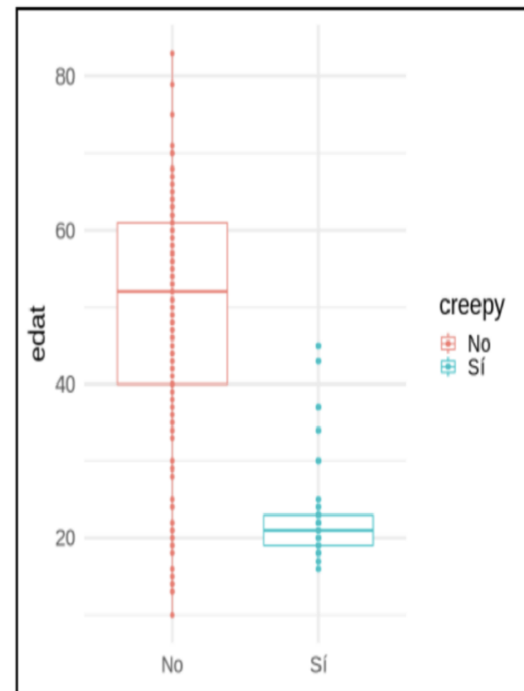
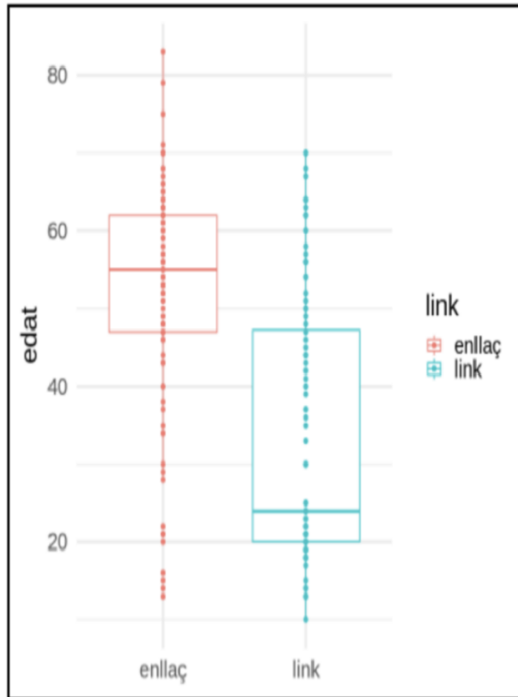
- Individual o en grup
- Al final del curs
- "Recuperable" com a examen
- Criteris d'avaluació i exemples: Aula Global

Relació entre edat i ús d'anglicismes

2022-06-28

Introducció

Estudis previs suggereixen que hi ha hagut un increment en l'ús d'anglicismes en altres llengües, però actualment encara no s'ha estudiat la relació entre l'ús d'anglicismes i l'edat dels parlants, i menys l'efecte que té sobre el costatè. Segons Serrano (2017), encara que sempre hi ha hagut interaccions entre



Altres exemples

- "Carrota vs. pastanaga en el territori alt empordanès"
- "S'ha substituït més de pressa el dialecte balear pel central que el nord-occidental en venir a Barcelona?"
- "Intercomprensión románica o plurilingüísmo: análisis de identificación de lenguas"
- "Evaluating linguistic laws on the Harry Potter saga"
- ...

Avaluació

Revisió per parells (10%, no recuperable)

- Dues parts:
 1. Esborrany del vostre anàlisi final
 2. Feedback sobre el esborrany de dos companyes unes setmanes més tard
- Només opció d'aprovar, o no
- No s'aprova si no es lliuren els dos documents; o si qualsevol no s'adequa a un estàndard mínim de qualitat

Avaluació

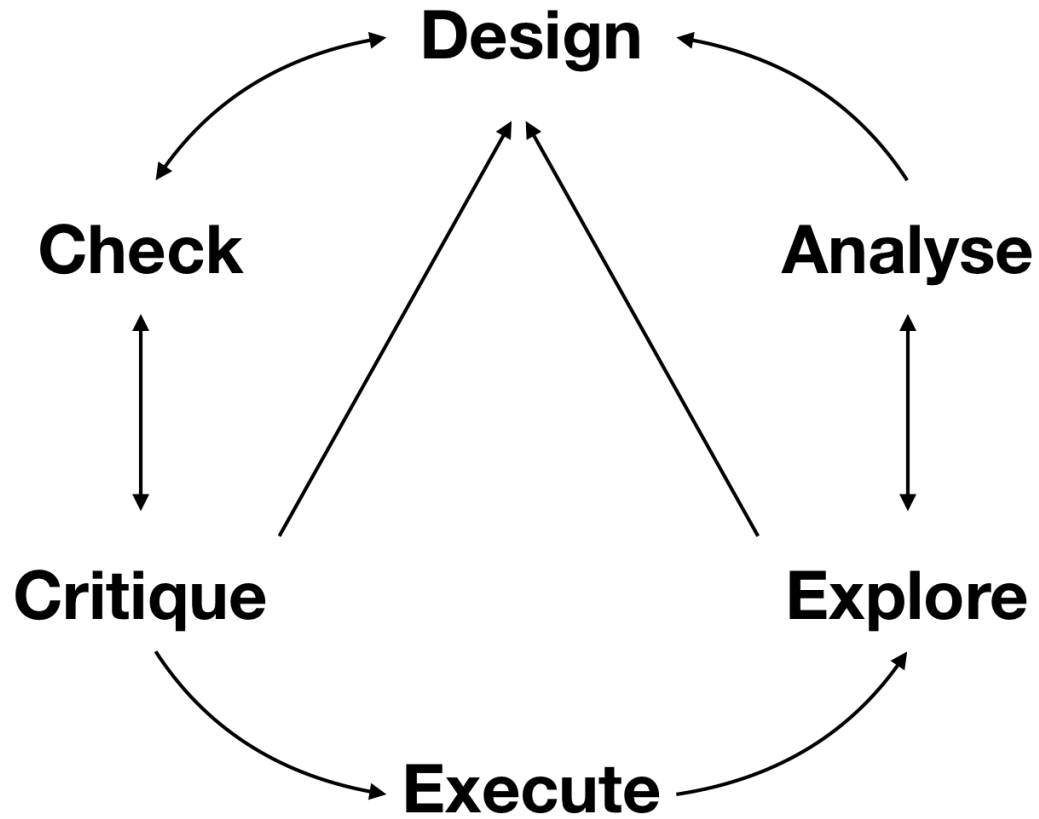
- Exercicis setmanals (20%, no recuperable)
- Exercici pràctic (20%, recuperable)
- Revisió per parells (10%, no recuperable)
- Informe: disseny i execució d'anàlisi (50%, recuperable)

Sinopsi del curs

Què és un mètode empíric?

Doneu un exemple

Cicle d'anàlisi



A efectes pràctics

- Manipulació i interpretació de dades
- Comparació de grups
- Predicció
- Visualització

... aplicat a ciències del llenguatge

Anàlisi inferencial

Què és un anàlisi descriptiu?

Doneu un exemple

Anàlisi inferencial (vs. descriptiu)

- Inferència de propietats més enllà de la mostra
- Predicció
- Comparació
- Causa-efecte

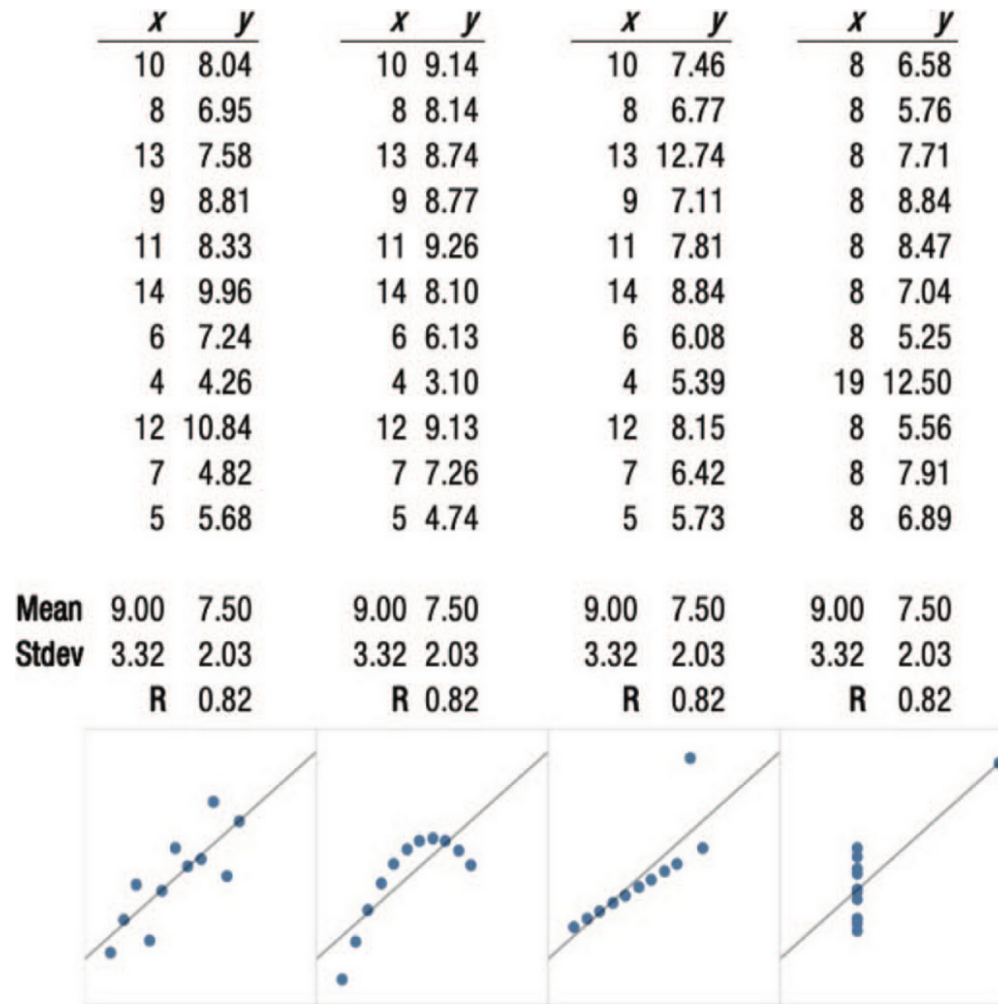


Fig. 1 de Franconeri et al. 2021 [The Science of Visual Data Communication: What Works](#)

Exemples

Estimació de l'efecte d'un o més *predictores* en un *resultat*

- Fluïdesa en L2 dependent de L1
- Probabilitat de l'ús d'un (pro)nom a la base del seu entorn lingüístic
- Comparació d'un mètode d'aprenentatge de la llengua contra un altre
- Comparació de polítiques lingüístiques i les seues conseqüències
- ...

Replicabilitat i reproduïbilitat

Crisi de replicabilitat i reproduïbilitat

Replication is one of the central issues in any empirical science.

To confirm results or hypotheses by a repetition procedure is at the basis of any scientific conception. A replication experiment to demonstrate that the same findings can be obtained in any other place by any other researcher is conceived as an operationalization of objectivity.

It is the proof that the experiment reflects knowledge that can be separated from the specific circumstances (such as time, place, or persons) under which it was gained.

-Stefan Schmidt

Replicabilitat

Que es puguin obtenir resultats consistents amb les mateixes dades d'entrada; passos computacionals; mètodes; codi; i condicions d'anàlisi

Reproduïbilitat

Que es puguin obtenir resultats consistents en diferents anàlisis que busquen respondre la mateixa pregunta, cadascú amb les seves pròpies dades

Crisi de replicabilitat i reproduïbilitat

- Una quantitat important de resultats no s'han pogut ni replicar ni reproduir (50%-80%, depenent del camp)
- Va comportar un gran canvi en metodologia i documentació (estem en això)
- Independentment, és central documentar totes les teves decisions i manipulacions; de principi a fi
- Llengües de programació són una eina ideal per
 - conduir anàlisi empírica, i
 - assegurar-se que sigui replicable

R

```
x = c(5, 3, 8, 4) # vector of values  
mean(x)           # mean of x
```

```
## [1] 5
```

En aquest curs

- Ús rudimentari i interactiu
- Principalment: intuïció sobre estructures de dades i la seva manipulació
- Manera d'assegurar replicabilitat
- Manera de millorar comunicació

Estadística descriptiva

Estadística descriptiva

- Tendències centrals (mitja; mitjana; moda)
- Dispersió (variança; desviació típica)
- Variació conjunta (correlació, covariança)

Si cal fer una revisió: Franke (2021) *An Introduction to Data Analysis*, capítol 5

Propera sessió de seminari:

- Petit resum d'estadística descriptiva
- R i collaborate

Si cal fer una revisió: Franke (2021) *An Introduction to Data Analysis*, capítol 5

Propera sessió

- Si n'hi ha, preguntes generals sobre el curs
 - Si n'hi ha, preguntes generals sobre l'informe
-

- **Disseny d'anàlisi**