

Tipologia i cicle de vida de les dades

Pràctica 2. Neteja i anàlisi de dades

Solució

Aitor Ferrus Blasco [aferrus]
Alonso López i Vicente [alopezvic]

05/01/2021

Contents

1. Descripció del dataset.	2
2. Integració i selecció de les dades d'interès a analitzar.	2
3. Neteja de les dades	2
3.1. Les dades contenen zeros o elements buits? Com gestionaries aquests casos?	2
3.2. Identificació i tractament de valors extrems.	2
4. Anàlisi de les dades.	2
4.1. Selecció dels grups de dades que es volen analitzar/comparar (planificació dels anàlisis a aplicar).	2
4.2. Comprovació de la normalitat i homogeneïtat de la variància.	2
4.3. Aplicació de proves estadístiques per comparar els grups de dades. En funció de les dades i de l'objectiu de l'estudi, aplicar proves de contrast d'hipòtesis, correlacions, regressions, etc. Aplicar almenys tres mètodes d'anàlisi diferents.	2
5. Representació dels resultats a partir de taules i gràfiques.	2
6. Resolució del problema. A partir dels resultats obtinguts, quines són les conclusions? Els resultats permeten respondre al problema?	2
7. Codi: Cal adjuntar el codi, preferiblement en R, amb el que s'ha realitzat la neteja, anàlisi i representació de les dades. Si ho preferiu, també podeu treballar en Python.	2
8. Contribucions	3

1. Descripció del dataset.

Perquè és important i quina pregunta/problema pretén respondre?

Resposta

2. Integració i selecció de les dades d'interès a analitzar.

Resposta

3. Neteja de les dades

3.1. Les dades contenen zeros o elements buits? Com gestionaries aquests casos?

Resposta

3.2. Identificació i tractament de valors extrems.

Resposta

4. Anàlisi de les dades.

4.1. Selecció dels grups de dades que es volen analitzar/comparar (planificació dels anàlisis a aplicar).

Resposta

4.2. Comprovació de la normalitat i homogeneïtat de la variància.

Resposta

4.3. Aplicació de proves estadístiques per comparar els grups de dades. En funció de les dades i de l'objectiu de l'estudi, aplicar proves de contrast d'hipòtesis, correlacions, regressions, etc. Aplicar almenys tres mètodes d'anàlisi diferents.

Resposta

5. Representació dels resultats a partir de taules i gràfiques.

Resposta

6. Resolució del problema. A partir dels resultats obtinguts, quines són les conclusions? Els resultats permeten respondre al problema?

Resposta

7. Codi: Cal adjuntar el codi, preferiblement en R, amb el que s'ha realitzat la neteja, anàlisi i representació de les dades. Si ho preferiu, també podeu treballar en Python.

Resposta

8. Contribucions

Contribucions	Firma
Investigació prèvia	Aitor Ferrus Blasco, Alonso López i Vicente
Redacció de les respostes	Aitor Ferrus Blasco, Alonso López i Vicente
Desenvolupament codi	Aitor Ferrus Blasco, Alonso López i Vicente