# Descripción de los datos

Disponemos de varios archivos que contienen los datos de alumnos de la UPV matriculados este año de todas las escuelas y, si no todas las carreras, de un elevado número. Los datos están anonimizados, teniendo un hash de los DNIs. Algunos de las transformaciones están en Datos/Iniciales/Transformaciones.pdf

Disponemos de toda la información relacionada con datos demográficos del estudiante (año de nacimiento, sexo, nota de acceso a la universidad…), de actividad académica (créditos matriculados este año, créditos superados, rendimiento…) y datos referidos a su uso de los recursos de la universidad y de poliformat (eventos de poliformat, conexiones al wifi…, tiempo total en poliformat…)

Datos identificativos:

|  |  |
| --- | --- |
| dni\_hash | character |
| asi\_hash | character |
| tit\_hash | character |
| titnom | character |
| asinom | character |

Datos demográficos

|  |  |
| --- | --- |
| nacionalitat | character |
| data\_nac | integer |
| sexe | character |
| alta\_universitat | integer |
| prov\_origen | character |
| anyo\_ingreso | integer |
| tipo\_ingreso | character |
| nota10 | numeric |
| nota14 | numeric |
| campus | character |
| estudios\_p | integer |
| estudios\_m | integer |
| dedicacion | character |
| desplazado | integer |
| discapacidad | numeric |
| becado | integer |

Datos Académicos

|  |  |
| --- | --- |
| preferencia\_seleccion | integer |
| baja\_fecha | character |
| caca | integer |
| grupos\_por\_tipocredito | character |
| matricula\_activa | integer |
| nota\_asig | numeric |
| cod\_centro | character |
| curso\_mas\_bajo | integer |
| curso\_mas\_alto | integer |
| cred\_matx | numeric |
| cred\_sup\_normal | numeric |
| cred\_sup\_espec | numeric |
| cred\_sup | numeric |
| cred\_mat\_normal | numeric |
| cred\_mat\_movilidad | numeric |
| cred\_ptes\_acta | numeric |
| cred\_mat\_practicas | numeric |
| cred\_mat\_sem\_a | numeric |
| cred\_mat\_sem\_b | numeric |
| cred\_mat\_anu | numeric |
| cred\_mat\_total | numeric |
| cred\_sup\_sem\_a | numeric |
| cred\_sup\_sem\_b | numeric |
| cred\_sup\_anu | numeric |
| cred\_sup\_total | numeric |
| rendimiento\_cuat\_a | numeric |
| rendimiento\_total | numeric |
| exento\_npp | integer |
| anyo\_inicio\_estudios | integer |
| es\_retitulado | numeric |
| es\_adaptado | numeric |
| cred\_sup\_x | numeric |
| practicas | numeric |
| Actividades | numeric |
| Ajuste | numeric |
| cred\_sup\_tit | numeric |
| cred\_pend\_sup\_tit | numeric |
| impagado\_curso\_mat | numeric |
| asig1 | numeric |
| pract1 | numeric |
| activ1 | character |
| total1 | numeric |
| ajuste1 | integer |
| rend\_total\_ultimo | character |
| rend\_total\_penultimo | numeric |
| rend\_total\_antepenultimo | numeric |

Datos de poliformat y recursos

|  |  |
| --- | --- |
| pft\_events\_2024\_x | numeric |
| pft\_days\_logged\_2024\_x | numeric |
| pft\_visits\_2024\_x | numeric |
| pft\_total\_minutes\_2024\_x | numeric |
| n\_wifi\_days\_2024\_ | numeric |
| n\_resource\_days\_2024\_x | numeric |
| pft\_test\_submissions\_2024\_x | numeric |

**Notas de los datos:**

Los datos referidos a poliformat y recursos, a excepción de las conexiones al wifi, todos dependen de la asignatura y tienen un valor único.

Por lo pronto, hay que quitar asignaturas del cuatrimestre B. No tienen datos, no tienen sentido. Sin embargo, no tengo ninguna variable que me permita eliminarla, más que observar que no hay eventos.

Las cosas de poliformat son 7x7=49+10=59 columnas

**Preguntas**