

MEMORIA

PRÁCTICA FINAL – PROGRAMACIÓN EN SCRIPTING

MANUEL ROJAS GARCÍA
UOC: mrojasgar

Semestre: Sep 2023/ Feb 2024

UOC

*PROCESAMIENTO DE DATOS
ACADEMICOS DE ESTUDIANTES
POR GÉNERO Y AÑO.
PROCESAMIENTO DE DATOS
PARA EL ESTUDIO DEL RATIO
POR GÉNERO Y AÑO PARA EL
ANÁLISIS DE LA PARIDAD.*

Universitat Oberta
de Catalunya



Índice

<i>A. Explica y justifica brevemente el motivo que te ha llevado a elegir la Operación de categorización/agrupación que has usado en el apartado D).</i>	3
<i>B. Adjunta una captura de pantalla con el informe que contiene los resultados del apartado D y haz una interpretación de los mismos.</i>	3
<i>C. A partir de los datos que has obtenido en los apartados B), C) y D) haz una valoración y establece las conclusiones a la que has llegado.</i>	5
<i>D. Indica las principales dificultades que hayáis encontrado y proponed técnicas para mejorar o extender el proyecto</i>	6
<i>Bibliografía</i>	7

A. Explica y justifica brevemente el motivo que te ha llevado a elegir la Operación de categorización/agrupación que has usado en el apartado D).

La elección de la operación de categorización/agrupación en el apartado se fundamenta en la necesidad de comprender y contextualizar mejor los datos, especialmente frente a la observación inicial de que alrededor del 40% de los alumnos han suspendido o no se han presentado. Este porcentaje elevado indica una situación crítica que puede afectar significativamente el análisis general de los datos académicos.

Para abordar esta problemática, se optó por realizar una categorización más detallada de las notas, asignando valores medios a las categorías de matrícula de honor, excelente, notable, aprobado y suspenso. Esta operación permitirá interpretar de manera más equitativa la información y mitigar el impacto de la gran cantidad de suspensiones en la evaluación general del rendimiento académico. Asignar valores específicos a cada categoría facilitará la comparación y el análisis más preciso de las notas.

Además, se considera crucial determinar el ratio entre hombres y mujeres, ya que esto proporcionará información valiosa para evaluar la paridad de género en el centro educativo. Entender si hay desequilibrios significativos entre el número de hombres y mujeres puede ayudar a la institución a tomar medidas específicas para promover la igualdad de oportunidades y condiciones educativas.

La elección de la categorización/agrupación se justifica por la necesidad de adaptar y contextualizar los datos frente a un alto porcentaje de suspensiones, así como por la importancia de analizar la paridad de género para implementar acciones correctivas si es necesario. Estas operaciones permitirán obtener una visión más precisa y significativa del rendimiento académico y la distribución de género en el centro educativo.

B. Adjunta una captura de pantalla con el informe que contiene los resultados del apartado D y haz una interpretación de los mismos.

Informe de Notas Medias por Año y Género

Año	Género	Nota Media	Alumnos por género
2015	Femenino	5.06	42.02%
2015	Masculino	5.13	57.98%
2004	Femenino	5.14	38.39%
2004	Masculino	5.14	61.61%
2021	Femenino	5.07	36.69%
2021	Masculino	5.16	63.31%
2009	Femenino	5.11	40.79%
2009	Masculino	5.17	59.21%
2011	Femenino	5.09	40.54%
2011	Masculino	5.12	59.46%
2000	Femenino	5.15	38.27%
2000	Masculino	5.10	61.73%
2016	Femenino	5.08	40.19%
2016	Masculino	5.14	59.81%
2005	Femenino	5.14	39.09%
2005	Masculino	5.12	60.91%

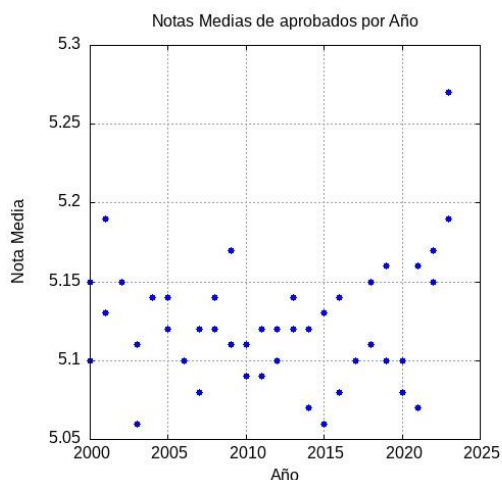
El informe tiene una cantidad completa de los años y se ha subdividido en dos capturas.

Se ha realizado un análisis detallado para evaluar la nota media de los estudiantes, considerando una calificación de suspenso fijada en 3.5. Este informe se ha dividido por género y año académico. Además, se ha incorporado un gráfico que visualiza el rango de la nota media. Es relevante destacar que la nota media es significativamente baja, con una variación mínima, indicando que

la mayoría de los estudiantes ha obtenido una calificación de aprobado (A). Al considerar 3.5 como la media de los alumnos con calificación suspensa (SU), se puede afirmar con alto grado de certeza que aquellos estudiantes con una nota media inferior a este valor tendrán una calificación media final de suspendido.

Aunque se puede determinar que el último año académico presenta un desempeño levemente superior y se identifica una pequeña disparidad entre géneros, la variación general en las notas medias es mínima. Este análisis proporciona información valiosa para comprender la distribución

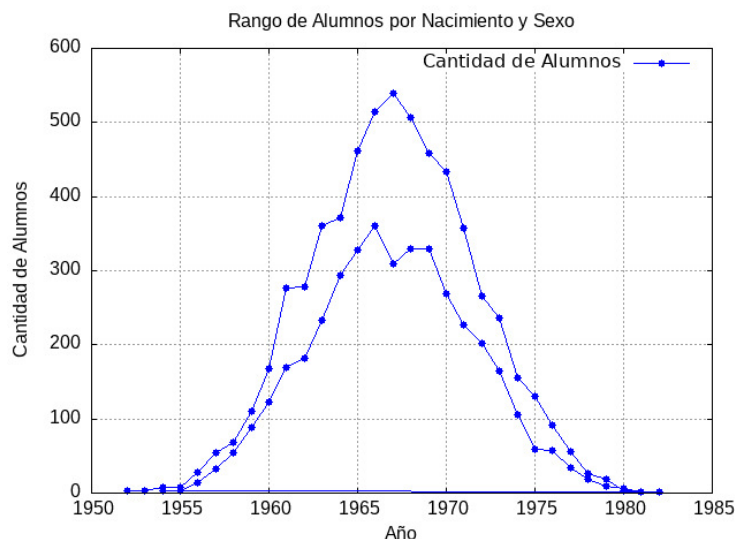
2006	Femenino	5.10	40.56%
2006	Masculino	5.10	59.44%
2023	Femenino	5.19	37.50%
2023	Masculino	5.27	62.50%
2013	Femenino	5.12	42.45%
2013	Masculino	5.14	57.55%
2002	Femenino	5.15	38.51%
2002	Masculino	5.15	61.49%
2018	Femenino	5.11	39.03%
2018	Masculino	5.15	60.97%
2007	Femenino	5.12	39.90%
2007	Masculino	5.08	60.10%
2014	Femenino	5.07	40.88%
2014	Masculino	5.12	59.12%
2003	Femenino	5.06	39.11%
2003	Masculino	5.11	60.89%
2020	Femenino	5.08	38.69%
2020	Masculino	5.10	61.31%
2019	Femenino	5.16	38.88%
2019	Masculino	5.10	61.12%
2008	Femenino	5.14	41.43%
2008	Masculino	5.12	58.57%
2010	Femenino	5.09	41.07%
2010	Masculino	5.11	58.93%



de calificaciones y la probabilidad de obtener una aprobación, considerando 3.5 como el punto de referencia crítico para una nota media superior al 5.

Como punto conclusivo de nuestro análisis, se destaca la notable disparidad de género entre los individuos nacidos en el periodo comprendido entre 1961 y 1975, alcanzando una brecha máxima entre

aquellos nacidos entre 1960 y 1970. En todos los años considerados, la proporción de hombres supera o se mantiene prácticamente equivalente a la de mujeres. Este aspecto merece una consideración especial en futuros análisis, ya que podría proporcionar insights significativos acerca de diversos aspectos del centro educativo en cuestión.



C. A partir de los datos que has obtenido en los apartados B), C) y D) haz una valoración y establece las conclusiones a la que has llegado.

El script B presenta un análisis detallado de los resultados de suspensos, categorizados por asignatura y género, revelando que las mujeres exhiben un porcentaje ligeramente superior de suspensos en las asignaturas A48 y A12.

Al consolidar los datos de todas las asignaturas analizadas, se revela que el porcentaje de suspensiones por parte de las mujeres es del 40.82%, mientras que para los hombres es del 40.52%. La diferencia entre ambos géneros en el conjunto total de asignaturas es mínima, situándose en un 0.3%. Estos resultados proporcionan una visión detallada de las tendencias de rendimiento en suspensos, permitiendo una comprensión más profunda de las disparidades de género en este ámbito educativo.

El script C, aunque no arroja resultados relativos a los rangos de edad, proporciona información exhaustiva sobre el porcentaje de alumnos en relación con el total, la media de graduados en un rango dado, y la media de asignaturas cursadas en dicho intervalo. Además, ofrece estadísticas globales, incluyendo el total de alumnos, la media de graduados y la media de asignaturas cursadas en la universidad. Es importante señalar que, a pesar de posibles desviaciones en los cálculos, el script C no se centra en las diferencias de género, sino en la totalidad de estudiantes en la institución académica. A pesar de estas consideraciones, los resultados son de interés, ya que proporcionan una visión integral de la matriculación, graduación y distribución por edad de los estudiantes. Como conclusión, podemos determinar que el “grueso” de los alumnos es de una edad avanzada, no es un centro educativo con alumnos jóvenes menores a 35.

El informe del script D presenta una baja nota media general entre los estudiantes, con la mayoría obteniendo aprobado y poca variación. Se establece 3.5 como punto crítico para aprobar. Aunque hay un rendimiento ligeramente superior en el último año académico, la diferencia no es significativa. Se observa una pequeña disparidad de género, con mujeres mostrando un pequeño porcentaje de suspensos. El informe en la tercera columna concluye una notable disparidad de género entre los nacidos entre 1961 y 1975, especialmente entre 1960 y 1970. Se sugiere considerar este aspecto en futuros análisis para obtener insights sobre diversos aspectos del centro educativo.

En un análisis global de los datos, se desprende que la edad promedio de los alumnos se sitúa en un nivel considerablemente avanzado. Se observa una preocupante incidencia de suspensos y no presentados, lo que sugiere una situación que podría considerarse alarmante. Además, se constata un desequilibrio significativo en la distribución de género, con una marcada predominancia masculina en comparación con la participación femenina. Asimismo, al evaluar las calificaciones medias, excluyendo los elementos asociados a los suspensos y no presentados, se evidencia una tendencia hacia calificaciones considerablemente bajas.

Como conclusión final y en un plano más personal, se destaca la necesidad imperativa de implementar una campaña de atracción para el centro educativo, con un enfoque particular en facilitar el acceso y la participación de mujeres. Asimismo, se aboga por llevar a cabo una revisión exhaustiva tanto del plan docente de las asignaturas como del cuerpo docente, con el fin de identificar y abordar el notable porcentaje de suspensiones. La evaluación del contenido de los recursos educativos también emerge como una prioridad, dada la elevada proporción de suspensos y las calificaciones notoriamente bajas. Estas acciones son fundamentales para mejorar la calidad educativa y garantizar un entorno académico más equitativo y eficiente.

D. Indicad las principales dificultades que hayáis encontrado y proponed técnicas para mejorar o extender el proyecto.

El proyecto se fundamenta en la utilización de dos archivos CSV, cuya integración en un único conjunto de datos resultaría altamente beneficiosa. Las convenciones de nomenclatura de ambos archivos presentan ciertas complejidades que podrían ser simplificadas, particularmente en lo que respecta a la columna de semestre. La limpieza de datos en Bash se revela como una tarea ardua, y las expresiones regulares (regex) pueden resultar visualmente complejas para aquellos estudiantes que están incursionando por primera vez en esta disciplina.

Las dificultades inherentes al proyecto son notables, especialmente para estudiantes que se están iniciando en el uso de Bash. No obstante, es importante destacar que el proceso de aprendizaje ha sido significativamente enriquecedor y de rápido crecimiento. Existe la posibilidad de mejorar y ampliar el proyecto, incorporando gráficos, columnas adicionales e incluso la capacidad de realizar análisis a lo largo de rangos de años. Aunque el informe proporcionado es exhaustivo, se reconoce que la organización por años podría brindar una mayor coherencia al conjunto de resultados. Sin embargo, las restricciones temporales impuestas por el plan docente han limitado la expansión del proyecto en esta dirección.

Si me permitieran ofrecer una sugerencia con respecto a la asignatura, consideraría la posibilidad de reducir la duración de las demás Prácticas de Evaluación Continua (PECS), especialmente la primera y la segunda, y contemplar la inclusión de una PEC adicional que prepare de manera más integral a los estudiantes para el proyecto final. La complejidad inherente al desarrollo de scripts en Bash, particularmente en el contexto de proyectos de mayor envergadura, puede resultar abrumadora. Aunque reconozco la satisfacción de afrontar desafíos significativos y de cultivar la paciencia necesaria para superar las frustraciones asociadas, también observo que el esfuerzo requerido ha sido considerable en fechas concretas del calendario académico.

En conclusión, propongo la posibilidad de proporcionar códigos provisionales que permitan rastrear y evaluar la evolución del esfuerzo invertido a lo largo del tiempo. En mi caso específico, he abordado el ejercicio C en al menos cinco ocasiones y he dedicado a esta práctica más horas que el doble de la suma total de las horas invertidas en todas las PECS anteriores combinadas. Este aspecto podría ser ilustrativo de la complejidad y profundidad que ha adquirido el aprendizaje en el marco de este proyecto.

Bibliografia.

1. *Check number of arguments passed to a Bash script* [en línea] [fecha de consulta: 4 de enero de 2024]. Disponible en:
<https://stackoverflow.com/questions/18568706/check-number-of-arguments-passed-to-a-bash-script>
2. *How can I find encoding of a file via a script on Linux?* [en línea] [fecha de consulta: 4 de enero de 2024]. Disponible en:
<https://stackoverflow.com/questions/805418/how-can-i-find-encoding-of-a-file-via-a-script-on-linux>
3. *Curl vs Wget [Curl/Bash Code]* [en línea] [fecha de consulta: 4 de enero de 2024]. Disponible en: <https://reqbin.com/req/curl/c-ojpk15/wget-vs-curl>
4. *Linux Generate A MD5 String or Hash with md5sum Command* [en línea] [fecha de consulta: 4 de enero de 2024]. Disponible en:
<https://www.cyberciti.biz/faq/linux-md5-hash-string-based-on-any-input-string/>
5. *How can I find encoding of a file via a script on Linux?* [fecha de consulta: 4 de enero de 2024]. Disponible en:
<https://stackoverflow.com/questions/805418/how-can-i-find-encoding-of-a-file-via-a-script-on-linux>
6. *Join two csv files with key value* [fecha consulta 4 de enero de 2024]. Disponible:
<https://stackoverflow.com/questions/25875368/join-two-csv-files-with-key-value>
7. *Lawrence Berkeley National Laboratory* [en línea] [fecha de consulta 4 de enero de 2024]. Disponible en:
https://bl831.als.lbl.gov/~gmeigs/scripting_help/printf_awk_notes.txt
8. *Grep cheatsheet* [en línea] [fecha de consulta 4 de enero de 2024]. Disponible : <https://quickref.me/grep.html>
9. *AWK cheatsheet* [en línea] [fecha de consulta 4 de enero de 2024].
<https://www.shortcutfoo.com/app/dojos/awk/cheatsheet>
10. *Create HTML page from within bash script* [en línea] [fecha de consulta 4 de enero de 2024].
<https://stackoverflow.com/questions/57181863/create-html-page-from-within-bash-script>
11. <https://gist.github.com/l82Much/569452>
12. *How to put Gnuplot command in my bash script?* [fecha de consulta 4 de enero de 2024].
<https://stackoverflow.com/questions/55742939/cgi-how-to-put-gnuplot-command-in-my-bash-script>
13. <https://www.linuxquestions.org/questions/programming-9/cgi-bash-html-gnuplot-how-to-generate-a-graph-4175652415/>