

## Presentación de hallazgos hechos en el dataset

### Alumnos matriculados en la Universidad Nacional de Ingeniería - [UNI]

Grupo conformado por: Enzo Huanca y Anthony Rivera

#### Información (utilizada) del dataset:

Tenemos columnas que permanecen en su mayoría igual a las que había en un principio:

- Tipo de Matrícula: *MATRICULA\_T*
- Departamento de Domicilio: *DEPA\_HOME*
- Modalidad de Ingreso: *MODALIDAD*
- Facultad: *FACULTAD*
- Especialidad: *ESPECIALIDAD*
- Ciclo Relativo: *C\_R*

Y después tenemos columnas nuevas, hechas a partir de columnas previas.

- ID del estudiante: *ID*

Código resumido por cada *IDHASH* del estudiante, para facilitar la visualización de los datos.

Ejemplo: *IDHASH*: 0005769B3CA7B05EB2305248FE46234C4EB3ECC0934686FD ... → *ID*: 1

- Semestre: *SEMEST*

Combinación del Año con el Período de matrícula, para evitar usar muchas columnas. Ejemplo:

*ANIO*: 2016, *PERIODO*: 1 → *SEMEST*: 2016-1

- Edad y Rango de Edad del estudiante: *AGE* y *AGE\_R*

Obtenemos la edad (*AGE*) al restar el año actual por la fecha de nacimiento del estudiante. A partir de ahí, agrupamos las edades por rangos, de la siguiente manera:

$$AGE\_R = \begin{cases} 50 + & , 50 \leq AGE \\ 40 - 50 & , 40 \leq AGE < 50 \\ 30 - 40 & , 30 \leq AGE < 40 \\ -30 & , AGE < 30 \end{cases}$$

#### Información adicional:

Además de lo que se podía entender del dataset original, utilizamos información adicional:

- Tasas de pobreza:

Para facilitar la visualización de datos por departamentos de domicilio (en la columna *DEPA\_HOME*), los agrupamos por las tasas de pobreza de cada uno de los departamentos. Para el Callao, que es considerada una provincia constitucional, usamos el [Reporte Regional de Indicadores Sociales del Callao](#), y para los demás departamentos, usamos el [Informe Técnico sobre la Evolución de la Pobreza Monetaria](#). Una vez asignamos las tasas de pobreza (en porcentajes: *TASA\_P*) a cada dato en *DEPA\_HOME*, los agrupamos por el nivel de pobreza (*LVL\_P*), el cual se calcula así:

$$LVL_P = \begin{cases} \text{Muy alta} & , 0.375 < TASA\_P \\ \text{Alta} & , 0.300 < TASA\_P \leq 0.375 \\ \text{Moderada} & , 0.150 < TASA\_P \leq 0.300 \\ \text{Baja} & , 0.075 < TASA\_P \leq 0.150 \\ \text{Muy baja} & , TASA\_P \leq 0.075 \end{cases}$$

Ya en el tablero de control, se decidió que los indicadores que usarían *LVL\_P* llevarían el nombre de “Tasa de Pobreza”, y que los valores que usaran *TASA\_P* llevarían el nombre de “% Pobreza”.

- Significado de Ciclo Relativo para la UNI:

El término era nuevo para nosotros, y la explicación que había en el diccionario de datos no nos parecía suficiente explicación. De acuerdo con una [Resolución Rectoral de la Facultad de Ciencias del 2017](#), el ciclo relativo nos indica el ciclo del estudiante según cuántos créditos (generalmente, cuántos cursos) haya aprobado.

Es decir, si el estudiante aprobó todos los cursos del primer ciclo, su siguiente matrícula indicará que está en ciclo relativo 2. Sin embargo, hay estudiantes que adelantan cursos y los aprueban, por lo que también puede suceder lo siguiente:

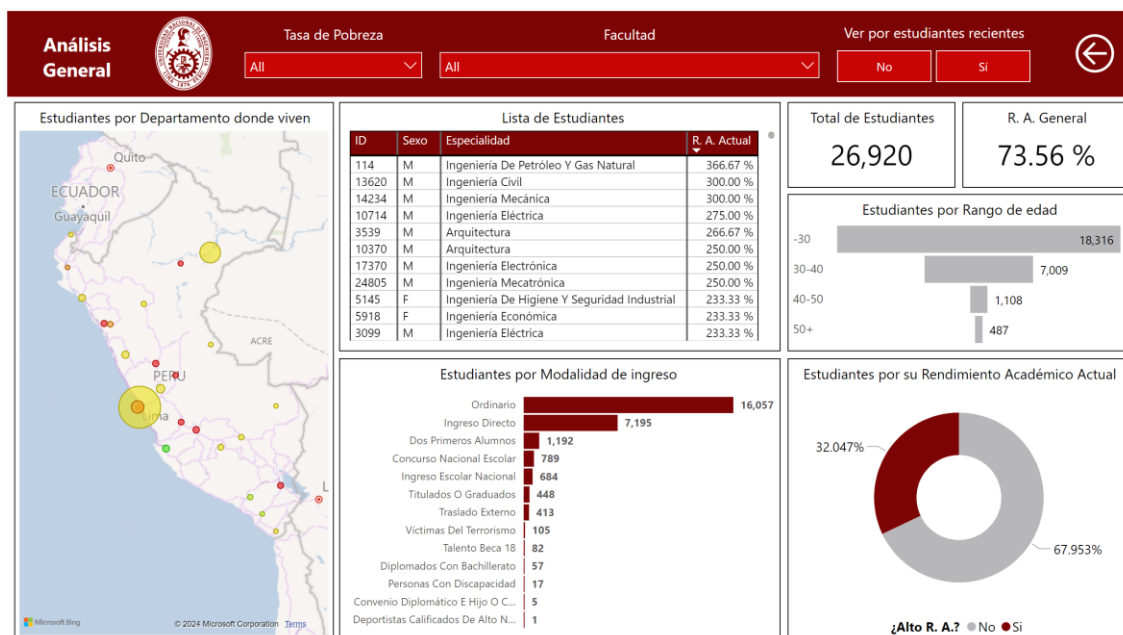
Un estudiante lleva los cursos de segundo y algunos de tercer ciclo en su segundo semestre en la universidad, y los aprueba. En el tercer semestre, lleva los que le faltaban de tercer ciclo y todos los de cuarto ciclo, y también los aprueba. En el cuarto semestre, su matrícula indicará que no está en ciclo relativo 4, sino 5.

Pero lo que no puede suceder, es que el ciclo relativo del estudiante disminuya (pase de ciclo relativo 5 a 3, por ejemplo). Las excepciones serían cuando el estudiante se cambia de especialidad, o ya se graduó y decidió estudiar una especialidad más.

## Explicación del tablero de control (Dashboard):

Este consiste en 3 páginas. La primera que aparece es un menú principal con 2 botones: General y Detallado. Cada uno lleva a las otras páginas, según el análisis que quiera ver el usuario, y cuentan con un botón en su esquina superior derecha, para regresar al menú principal.

### a) Análisis General:



La página de análisis general muestra información tomando en cuenta a todos los estudiantes, mostrando detalles generales como porcentajes, proporciones y hasta la ubicación de sus domicilios.

Cada estudiante cuenta con un porcentaje llamado “R. A. Actual”, o “Rendimiento Académico Actual”. Éste es un porcentaje individual para cada estudiante, y sirve para ver qué tan bien le está yendo al estudiante en la universidad. El rendimiento académico es individual por cada estudiante, y se calcula tomando en cuenta lo siguiente:

1. Por cada estudiante, tomamos sólo las matrículas de su especialidad más actual.
2. Tomamos el valor del ciclo relativo actual,  $LC$ , y del primer ciclo relativo,  $FC$ , (usualmente será 1, pues es donde la mayoría de los estudiantes empieza).
3. Obtenemos el avance del estudiante hecho mediante esta fórmula:

$$TA = LC - FC + 1$$

- a. Ejemplo 1: El estudiante empezó en el ciclo relativo 1 y su última matrícula dice que está en el ciclo relativo 4. Entonces su avance sería:  $4 - 1 + 1 = 4$ .
  - b. Ejemplo 2: El estudiante empezó en el ciclo relativo 1 y su última matrícula aparece en el ciclo relativo 5. Entonces su avance sería:  $5 - 1 + 1 = 5$ .
4. Dividimos el avance del estudiante por el total de matrículas con esa especialidad, y multiplicamos el resultado por el 100%.

$$RA = \frac{TA}{\text{Total Matrículas (TM)}} \cdot 100\%$$

5. Obtenemos el rendimiento académico actual del estudiante.
- a. Ejemplo 1: El estudiante sólo tiene 3 matrículas para esa especialidad. Entonces su rendimiento académico sería:  $\frac{4}{3} \cdot 100\% = 133.33\%$ .
  - b. Ejemplo 2: El estudiante tiene 7 matrículas para esa especialidad. Entonces su rendimiento académico sería:  $\frac{5}{7} \cdot 100\% = 71.43\%$
6. Repetir este proceso con cada estudiante

El escenario deseable sería que los estudiantes tengan un rendimiento académico del 100%, pues eso indicaría que siempre aprueban el ciclo completo en cada semestre que se matriculan, lo cual indicaría que se graduarían en los 5 años esperados.

Sin embargo, un escenario incluso mejor sería que los estudiantes tengan un rendimiento académico por encima del 100%, pues eso indicaría que son capaces de aprender mucho más de lo esperado, y tienen potencial que merece ser aprovechado.

El tablero de control cuenta con 2 indicadores:

- Total de Estudiantes: Número total de estudiantes desde el 2016-1 hasta el 2024-1.
- R. A. General: Rendimiento Académico General. Al igual que el R. A. individual de los estudiantes, éste sólo toma en cuenta las especialidades actuales de cada estudiante. Para calcular el R. A. General, se suman todos los avances hechos por los estudiantes, lo dividimos por el total de matrículas y lo multiplicamos por el 100%.

También cuenta con 3 segmentaciones de datos:

- Taza de Pobreza: Muestra los estudiantes cuyos departamentos pertenezcan a una tasa de pobreza muy baja, baja, moderada, alta y/o muy alta.

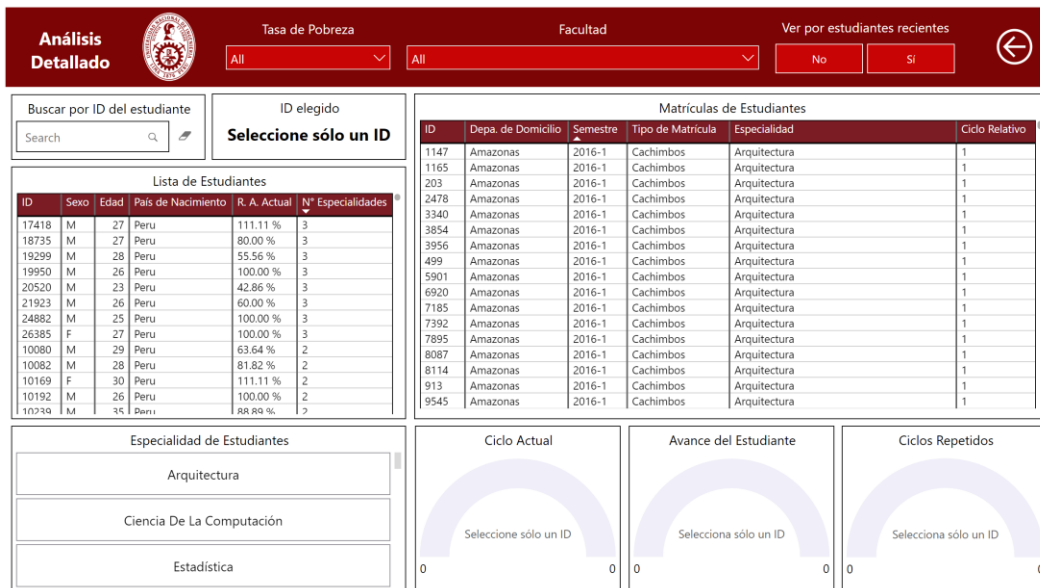
- Facultad: Muestra los estudiantes por sus facultades.
- Ver por estudiantes recientes: Muestra los estudiantes cuya última matrícula, por lo menos, fue en el semestre 2023-1, o hasta el semestre 2022-2.

Como información visual, cuenta con 4 gráficos:

- Gráfico de dona: Muestra los estudiantes que tienen un R. A. actual de al menos un 100%, o por debajo del 100%.
- Gráfico de embudo: Muestra los estudiantes por su rango de edad: menores de 30 años, entre 30 y 39 años, entre 40 y 49 años, y de 50 años a más.
- Gráfico de barras horizontales: Muestra los estudiantes por su modalidad de ingreso.
- Gráfico de mapa: Muestra los departamentos de domicilio de los estudiantes, por colores según la tasa de pobreza en la que se encuentren, y por tamaño según el número de estudiantes que viva o haya vivido allí.

Y por último, cuenta con 1 tabla para mostrar los estudiantes por su ID resumido, su sexo, especialidad y R. A. actual.

#### b) Análisis Detallado:



La página del análisis detallado muestra información tomando en cuenta a los estudiantes de manera individual. Los gráficos se ajustarán cuando el usuario seleccione a uno de los estudiantes en la tabla de la izquierda. Una vez hecho esto, se mostrará la información relacionada con él/ella.

El tablero cuenta con 4 segmentaciones de datos, donde 3 de ellos son los mismos que hay en el análisis general, junto con uno adicional:

- Especialidad de Estudiantes: Muestra los estudiantes por la especialidad en la que estudian o hayan estudiado. Es de utilidad cuando seleccionamos estudiantes con más de una especialidad.

Como información visual, cuenta con 3 gráficos de medición radial:

- Ciclo actual: Muestra el ciclo actual del estudiante.
- Ciclos avanzados: Muestra el avance individual ( $TA$ ) del estudiante.
- Ciclos repetidos: Muestras las matrículas donde haya un ciclo repetido de las anteriores matrículas del estudiante. Se calcula así:

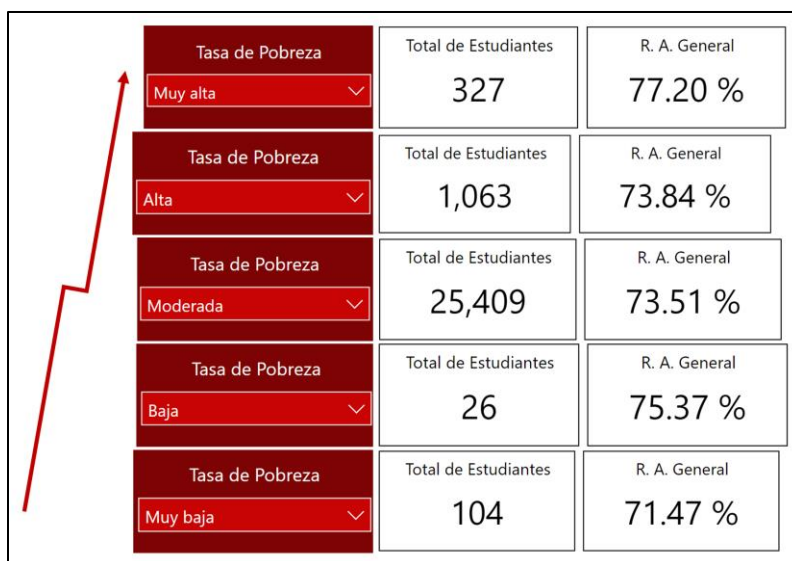
$$\text{Ciclos Repetidos} = \begin{cases} TM - TA & , TM - TA \geq 0 \\ 0 & , TM - TA < 0 \end{cases}$$

Y por último, cuenta con 2 tablas:

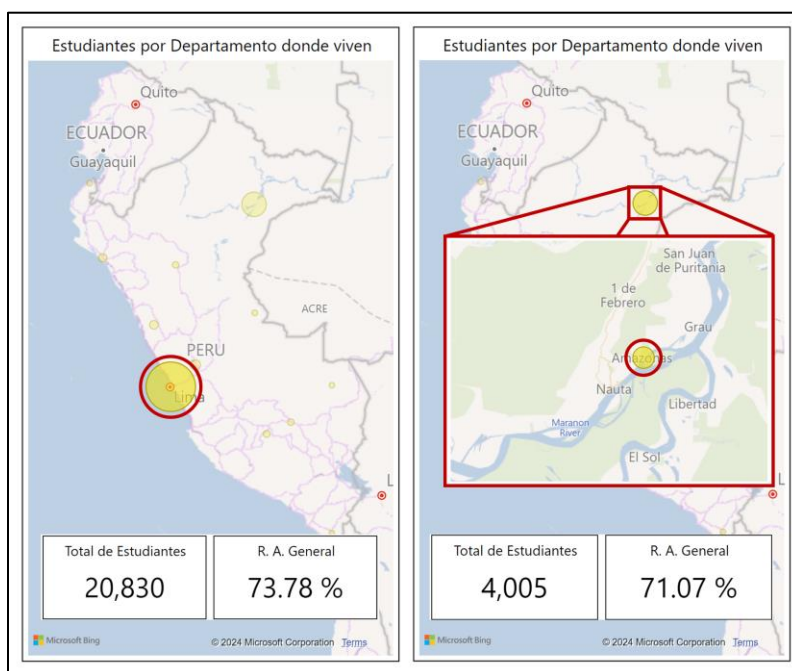
- La tabla de la izquierda, para mostrar los estudiantes por su ID resumido, su sexo, especialidad y R. A. actual. Aquí el usuario seleccionará una de las filas para revisar la información importante del estudiante
- La tabla de la derecha, para mostrar las matrículas de los estudiantes por el ID resumido, su departamento de domicilio, el semestre en el que fue hecha la matrícula, el tipo de matrícula, la especialidad del estudiante y su ciclo relativo.

### **Hallazgos del tablero de control**

1. A excepción de la tasa de pobreza baja, el R. A. general de los estudiantes aumenta mayormente cuanto más alta sea la tasa de pobreza de sus departamentos de domicilio.



2. La mayor parte de los estudiantes viven o han vivido en departamentos con una tasa de pobreza moderada, con 25409 estudiantes. Sin embargo, una enorme porción de estos estudiantes pertenece a dos departamentos: Lima y Amazonas.



3. La facultad con el mayor R. A. general es la de Ingeniería Ambiental, seguida de la Ingeniería Geológica, Minera y Metalúrgica; y de la de Ingeniería de Petróleo, Gas Natural y Petroquímica.

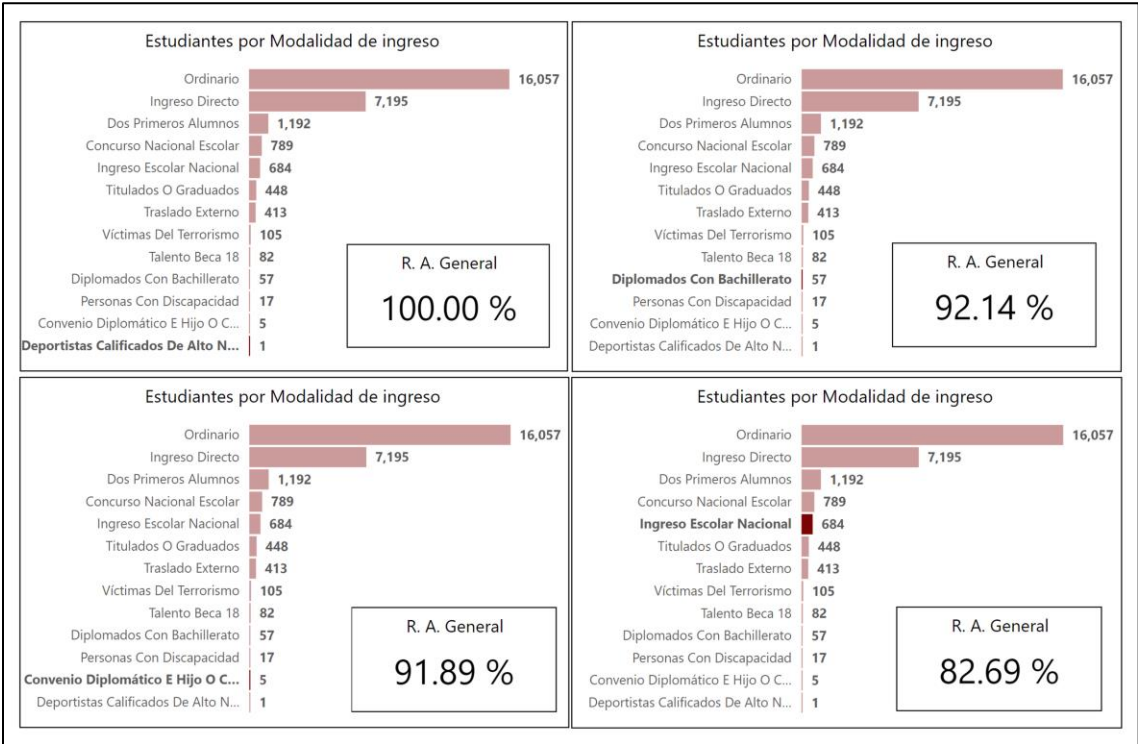
Facultad Arquitectura, Urbanismo Y Artes	Total de Estudiantes 1,814	R. A. General 62.86 %
Facultad Ciencias	Total de Estudiantes 2,480	R. A. General 63.08 %
Facultad Ingeniería Ambiental	Total de Estudiantes 1,883	R. A. General 81.48 %
Facultad Ingeniería Civil	Total de Estudiantes 3,370	R. A. General 75.58 %
Facultad Ingeniería De Petróleo, Gas Natural Y Petroquímica	Total de Estudiantes 682	R. A. General 77.85 %
Facultad Ingeniería Económica, Estadística Y Ciencias Sociales	Total de Estudiantes 2,036	R. A. General 76.32 %
Facultad Ingeniería Eléctrica Y Electrónica	Total de Estudiantes 3,360	R. A. General 73.66 %
Facultad Ingeniería Geológica, Minera Y Metalúrgica	Total de Estudiantes 2,208	R. A. General 80.72 %
Facultad Ingeniería Industrial Y De Sistemas	Total de Estudiantes 3,798	R. A. General 74.15 %
Facultad Ingeniería Mecánica	Total de Estudiantes 3,863	R. A. General 74.40 %
Facultad Ingeniería Química Y Textil	Total de Estudiantes 1,779	R. A. General 72.57 %

- La mayor parte de los estudiantes estudian o han estudiado en la facultad de Ingeniería Mecánica, seguida de Ingeniería Industrial y de Sistemas; y de Ingeniería Civil, con más de 3300 estudiantes en cada una.
- La suma de los totales de estudiantes por las tasas de pobreza da 26929, pero el total real es de 26920. Este exceso de 9 estudiantes ocurre porque hay estudiantes que con el paso del tiempo se han mudado a otros departamentos, e incluso a departamentos con distinta tasa de pobreza. Por ejemplo, el estudiante con el ID 16429 vivía en Ayacucho (con tasa de pobreza muy alta) en el semestre 2017-2; para cuando se matriculó en 2018-2, se encontraba ahora viviendo en Junín (con tasa de pobreza moderada).

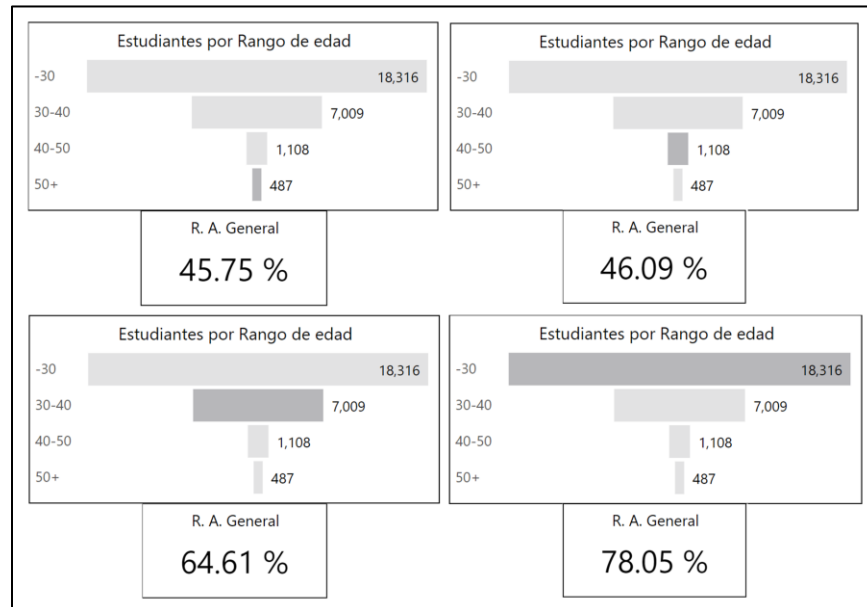


Matrículas de Estudiantes					
ID	Depa. de Domicilio	Semestre	Tipo de Matrícula	Especialidad	Ciclo Relativo
16429	Ayacucho	2017-2	Cachimbos	Ingeniería Metalúrgica	1
16429	Junín	2018-2	Cachimbos	Ingeniería De Minas	1
16429	Junín	2019-1	Regulares	Ingeniería De Minas	2
16429	Junín	2019-2	Regulares	Ingeniería De Minas	3
16429	Junín	2020-1	Regulares	Ingeniería De Minas	3
16429	Junín	2020-2	Reincorporados	Ingeniería De Minas	4
16429	Junín	2021-1	Regulares	Ingeniería De Minas	5
16429	Junín	2021-2	Regulares	Ingeniería De Minas	6
16429	Junín	2022-1	Regulares	Ingeniería De Minas	7
16429	Junín	2022-2	Regulares	Ingeniería De Minas	8
16429	Junín	2023-1	Regulares	Ingeniería De Minas	9
16429	Junín	2023-2	Regulares	Ingeniería De Minas	10
16429	Junín	2024-1	Regulares	Ingeniería De Minas	10

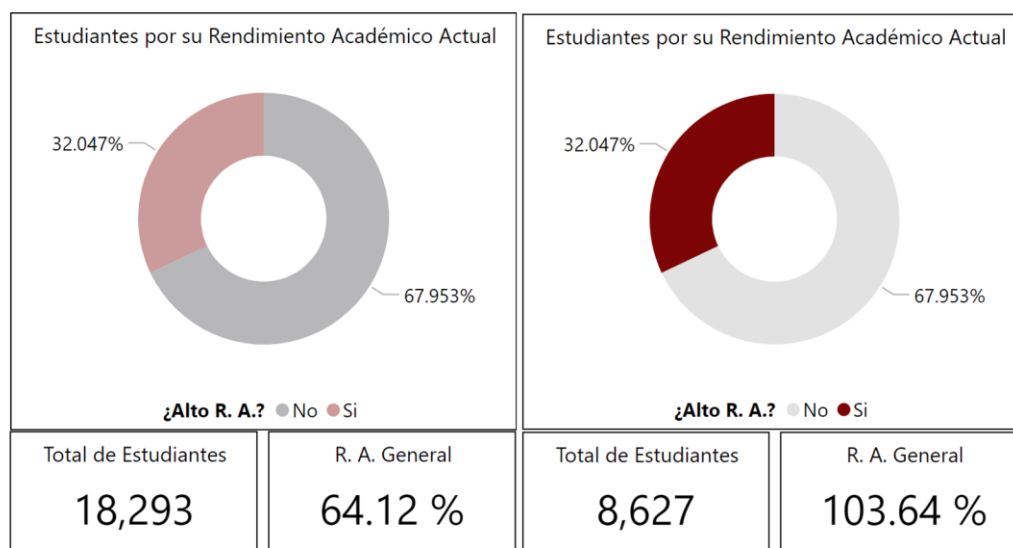
- La suma de totales de estudiantes por las facultades es de 27273 estudiantes. El exceso de 353 estudiantes también ocurre porque ha habido estudiantes que se cambiaron de especialidad, o terminaron la suya y decidieron estudiar otra más. Por supuesto, algunos estudiantes que pertenecen a este grupo especial se cambiaron a una especialidad que también pertenecía a otra facultad.
- La modalidad de ingreso con mayor R. A. general es la de Deportistas Calificados de Alto Nivel, pero ésta sólo incluye a un estudiante. Le sigue la modalidad de Diplomados con Bachillerato; la de Convenio Diplomático e Hijo o Cónyuge de Diplomático; y la de Ingreso Escolar Nacional.



8. La modalidad de ingreso con mayor número de estudiantes es la de Ordinario, seguida de la de Ingreso Directo; y la de Dos Primeros Alumnos, con más de 1000 estudiantes en cada una.
9. El R. A. general de los estudiantes aumenta cuanto más jóvenes son los estudiantes.



10. El número de estudiantes disminuye cuanto más viejos son los estudiantes.
11. El 67.95% de los estudiantes no alcanzan un alto rendimiento académico, con un R. A. general del 64.12%. Por otro lado, el 32.05% de los estudiantes sí alcanzan un alto rendimiento académico, con un R. A. general del 103.64%.



12. Tenemos 12763 estudiantes cuya última matrícula fue en el semestre 2022-2, con un R. A. general del 79.82%. Por otro lado, tenemos 14157 estudiantes recientes, con un R. A. general del 68.69%

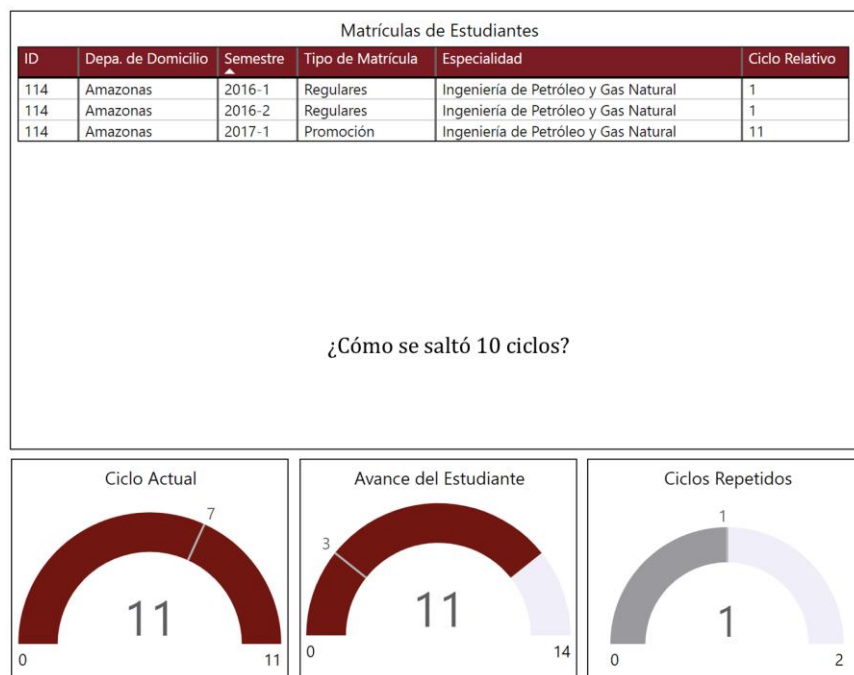
Ver por estudiantes recientes		Total de Estudiantes	R. A. General
No	Si	12,763	79.82 %
Ver por estudiantes recientes		Total de Estudiantes	R. A. General
No	Si	14,157	68.69 %

13. Tenemos varios estudiantes con un R. A. actual por encima del 100%. Sin embargo, aunque eso sea normal, los porcentajes demasiado altos no se ven normal.

Analicemos los primeros 3 estudiantes con el R. A. actual más alto:

- a. Estudiante con ID 114:

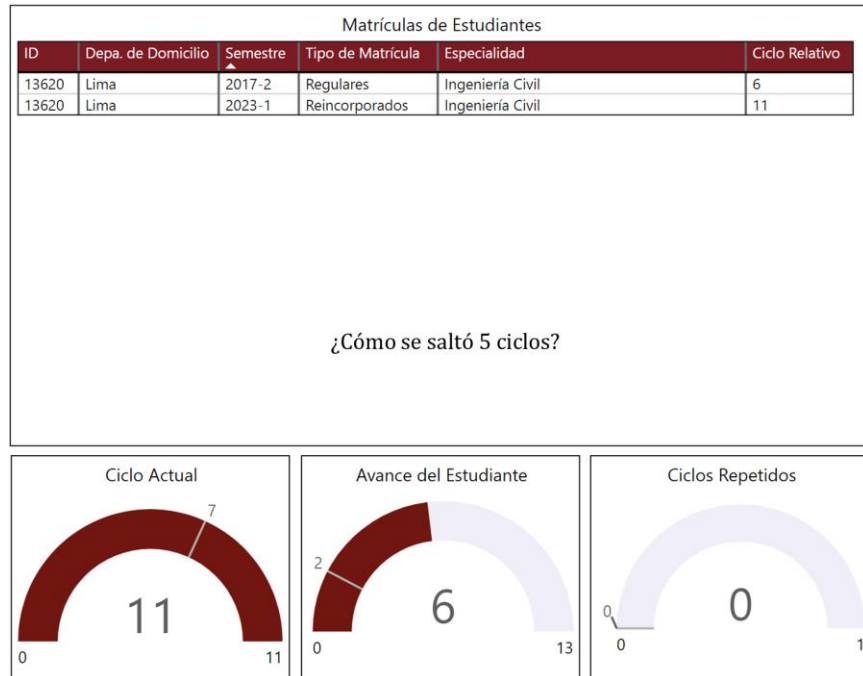
Este estudiante permaneció en un ciclo relativo 1 durante su primer y segundo semestre. Sin embargo, en el tercer semestre, pasó del primer al último ciclo ¿Cómo es posible esto?



La respuesta puede ser sencilla: un error de tipeo, en el que escribieron doble 1.

b. Estudiante con ID 13620:

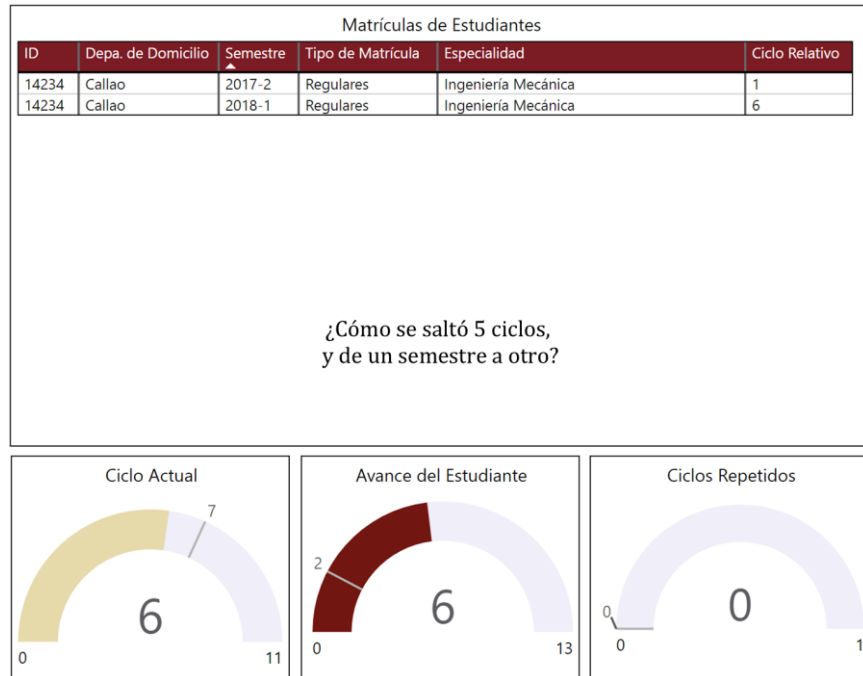
Este estudiante sólo tiene 2 matrículas. La primera es del semestre 2017-2, con ciclo relativo 6, y la segunda es del semestre 2023-1, ¡con ciclo relativo 11!



El hecho de que haya regresado a sus estudios en la UNI tras 5 años, entre el semestre 2018-1 y el 2022-2, podría indicar algo más: Puede ser que el estudiante dejara la universidad para continuar su carrera en otra casa de estudios, y cuando ya estaba muy adelantado, regresó a la UNI para concluir allí, convalidando la mayoría de los cursos hasta posicionarse como estudiante de onceavo ciclo.

c. Estudiante con ID 14234:

Este estudiante tiene sólo dos matrículas. Se puede ver que pasó del ciclo relativo 6 al 11 en tan solo 1 semestre.



Es posible que el estudiante venga de otra universidad, donde también estudió Ingeniería Mecánica, y no pudo convalidar sus cursos para el semestre 2017-2, pero sí para el semestre 2018-1.

Estos casos no son los únicos: Varios estudiantes con R. A. actual mayor al 100% presentan este mismo caso: es casi imposible que un estudiante salte 2 ciclos seguidos “a la manera legal”, pero decir que saltaron 4, 5 o hasta 10 ciclos es ridículo.

## Conclusiones del análisis

- En caso de que no se trate de corrupción, hay serios problemas con el registro de los ciclos relativos en las matrículas de los estudiantes. Se necesita mejorar el sistema que registra los datos de matrícula.
- Si bien el rendimiento académico puede indicar el potencial de los estudiantes para el aprendizaje, los valores exagerados pueden entorpecer la visión sobre este potencial, y hasta considerarlo irreal.
- Las matrículas con R. A. anormales vienen mayormente de los estudiantes matriculados hasta el 2022-2. Para llevar a cabo posibles soluciones, sería más efectivo tomar en cuenta los datos de matrículas que continúen más allá del semestre 2022-2.